



Recovery of continuous 3D refractive index maps from discrete intensity-only measurements using neural fields

Renhao Liu^{1,5,7}, Yu Sun^{1,6,7}, Jiabei Zhu², Lei Tian^{2,3} and Ulugbek S. Kamilov^{1,4}✉

Intensity diffraction tomography (IDT) refers to a class of optical microscopy techniques for imaging the three-dimensional refractive index (RI) distribution of a sample from a set of two-dimensional intensity-only measurements. The reconstruction of artefact-free RI maps is a fundamental challenge in IDT due to the loss of phase information and the missing-cone problem. Neural fields has recently emerged as a new deep learning approach for learning continuous representations of physical fields. The technique uses a coordinate-based neural network to represent the field by mapping the spatial coordinates to the corresponding physical quantities, in our case the complex-valued refractive index values. We present Deep Continuous Artefact-free RI Field (DeCAF) as a neural-fields-based IDT method that can learn a high-quality continuous representation of a RI volume from its intensity-only and limited-angle measurements. The representation in DeCAF is learned directly from the measurements of the test sample by using the IDT forward model without any ground-truth RI maps. We qualitatively and quantitatively evaluate DeCAF on the simulated and experimental biological samples. Our results show that DeCAF can generate high-contrast and artefact-free RI maps and lead to an up to 2.1-fold reduction in the mean squared error over existing methods.

The refractive index (RI) measures the optical density, which determines the interaction between light and matter within a sample. The real part of the RI characterizes the phase, whereas its imaginary part characterizes absorption. The RI can thus serve as an endogenous source of optical contrast for imaging samples without staining or labelling. By quantitatively characterizing the three-dimensional (3D) distribution of the RI, one can visualize cellular or subcellular structures useful for morphogenesis¹, oncology², cellular pathophysiology³, biochemistry⁴ and beyond (see the review papers in refs. ^{5,6}).

Intensity diffraction tomography (IDT) is a recent technique for recovering the 3D RI maps of a sample by measuring the light it scatters. In the standard IDT set-up, a sample is illuminated multiple times from different angles and a set of two-dimensional (2D) intensity projections is captured by the camera (see Fig. 1c). A tomographic image reconstruction algorithm is then used to computationally reconstruct the desired 3D RI distribution from the set of 2D measurements. Unlike traditional optical diffraction tomography, which uses interferometry to record the complex-valued light fields^{7–9}, IDT only measures the squared amplitude of the scattered light, leading to an easy set-up on standard transmission optical microscopes with inexpensive hardware modifications. Such flexibility has spurred different IDT variants integrating object scanning^{10,11}, angled illumination^{12–15}, pupil engineering^{16,17} and multiple scattering^{18,19}. Set-ups achieving high resolution¹⁸ and fast acquisition²⁰ have also been reported.

Despite the rich literature on IDT, image reconstruction remains a fundamental challenge. The first issue is that the phase of the scattered light field is missing from the measurements, resulting in a

nonlinear measurement system that is not characterizable by classical linear Fourier diffraction theory²¹. This rules out the use of standard filtered-backprojection methods and calls for advanced computational algorithms. The second issue is the well-known missing cone problem, which causes elongation of the reconstructed object along the optical axis (z -dimension) and hence reduction of the axial resolution. The missing cone problem is a result of a limited-angle tomographic set-up, in which illuminations can come only from one side of the sample plane with a limited range for angle variation (less than $\sim 40^\circ$ in our set-ups). This leads to incomplete coverage of the 3D Fourier spectra with a missing, cone-shaped region in the axial direction. These missing phase and cone problems make image reconstruction in IDT a severely ill-posed inverse problem.

Regularization methods are commonly used for mitigating the ill-posed nature of many inverse problems. These methods are based on minimizing a cost function that consists of a data-fidelity term and a regularization term, where the former uses a physical model to quantify the mismatch between the predicted and acquired measurements, whereas the latter promotes solutions that are consistent with a priori knowledge of the sample. For example, the least-squares loss and Tikhonov regularizer (ℓ_2 -penalty) are widely used for obtaining closed-form solutions to inverse problems¹⁴. The work on plug-and-play priors has generalized the notion of image priors to implicit regularizers characterized by image denoisers^{22–25}. Deep learning has recently emerged as a powerful framework for image reconstruction. A traditional deep learning reconstruction is based on training a convolutional neural network (CNN) on a large dataset to learn a mapping from low-quality images to their

¹Department of Computer Science and Engineering, Washington University in St. Louis, St. Louis, MO, USA. ²Department of Electrical and Computer Engineering, Boston University, Boston, MA, USA. ³Department of Biomedical Engineering, Boston University, Boston, MA, USA. ⁴Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, MO, USA. ⁵Present address: Washington University in St. Louis, Google Inc., St. Louis, MO, USA. ⁶Present address: Washington University in St. Louis, Caltech, St. Louis, MO, USA. ⁷These authors have contributed equally: Renhao Liu, Yu Sun. ✉e-mail: kamilov@wustl.edu

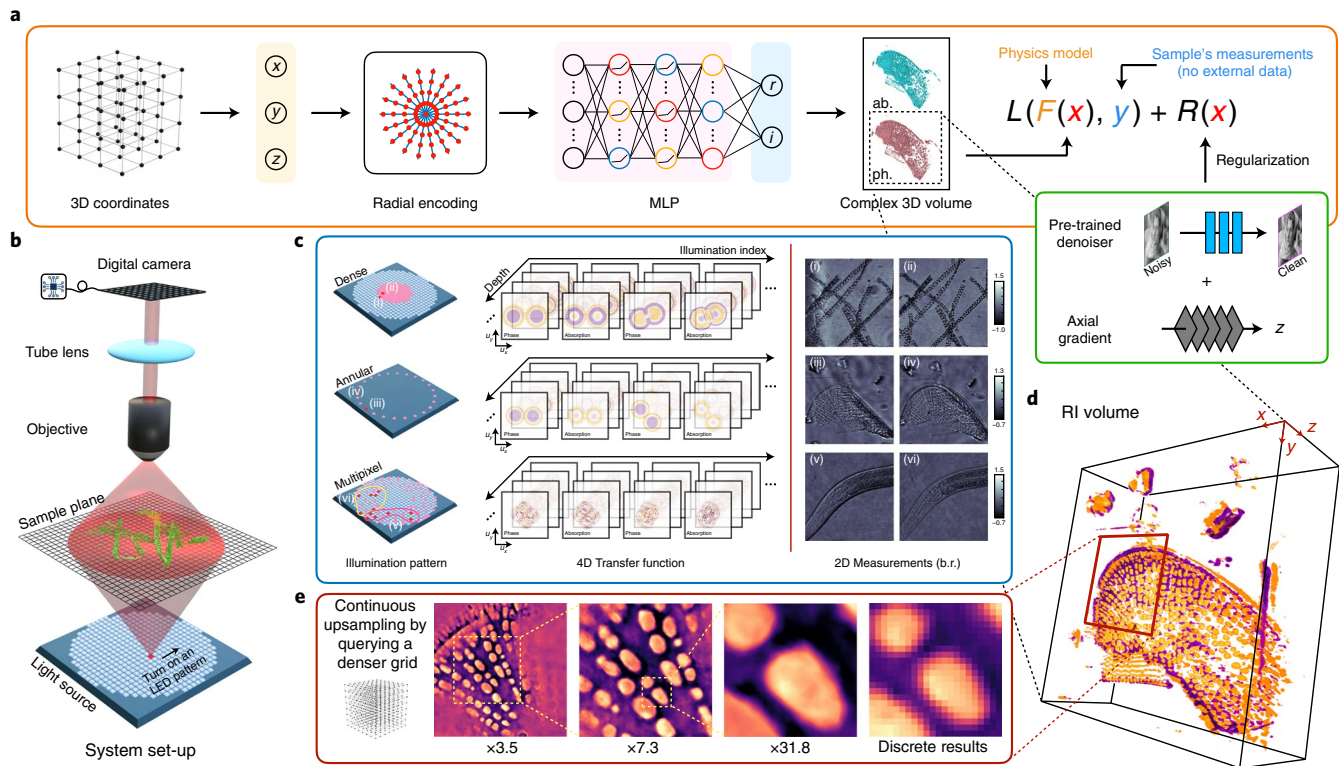


Fig. 1 | Conceptual illustration of DeCAF for IDT. **a**, DeCAF reconstructs the RI volume by learning a neural field parameterized by an MLP. The network is trained to map the 3D coordinates (x, y, z) to the corresponding RI value by minimizing a loss that penalizes measurement mismatch and imposes regularization. **b**, Our IDT system uses a programmable LED array to illuminate a sample from different angles and uses a digital camera to record the intensity measurements of the scattered light. By changing the illumination patterns, our system can implement different IDT modalities. **c**, Our experiments consider three illumination patterns (dense, annular and multiplexed illuminations), each of which corresponds to a different formulation of the 4D light transfer function (two separate 3D transfer functions for phase (ph.) and absorption (ab.), respectively). Example 2D intensity measurements are visualized for each illumination pattern. Note that the background light is removed (b.r.) from the image. **d**, DeCAF can reconstruct high-quality 3D RI maps from intensity-only and limited-angle measurements. **e**, DeCAF learns a continuous representation and can render samples on a pixel grid of the desired density (illustration with 3.5-, 7.5- and 31.8-fold upsampling).

high-quality counterparts. The state-of-the-art performance of such methods has been demonstrated in X-ray computed tomography^{26,27}, magnetic resonance imaging^{28,29}, optical tomography^{30,31} and seismic imaging³² (see refs. ^{33–35}). Although deep learning has considerably improved image reconstruction in many modalities, traditional deep learning methods are impractical for image reconstruction in IDT, where it is difficult to acquire high-quality ground-truth RI maps in experiments. Although a physics-based simulator has been proposed to generate datasets for training IDT artefact-suppressing CNNs, the results are still limited by the mismatch between the simulation and experiments³⁶.

Neural fields (NF) is a recent deep learning framework that has gained popularity in computer vision and graphics for representing and rendering 3D scenes using coordinate-based deep neural networks^{37,38}. It is worth mentioning that although NF was deemed to be the most appropriate term^{39,40}, this idea currently goes by various names in the vision/graphics literature, including neural coordinate-based representations or neural implicit models. It has been shown that NF can learn a high-quality representation of a complex scene from a sparse set of data without any external training dataset. Motivated by this property, we propose Deep Continuous Artefact-free RI field (DeCAF) as a novel NF-based IDT method for learning a high-quality, continuous 3D RI map from intensity-only and limited-angle measurements without any external training dataset of ground-truth RI maps. Figure 1 provides a conceptual illustration of DeCAF, the key features of which are as follows:

- The central component of DeCAF is a multilayer perceptron (MLP)—a fully connected (non-convolutional) deep network—for learning a function that maps 3D coordinates (x, y, z) to the corresponding complex-valued RI values. The trained MLP thus provides a continuous neural representation of the RI map. The RI value at any spatial location can be retrieved by querying the trained MLP with the corresponding coordinate. By decoupling representation from an explicit voxel grid, DeCAF can efficiently store large 3D volumes.
- DeCAF is a self-supervised method, meaning that it does not need to be trained on an external dataset of ground-truth RI maps. This is possible as the same MLP is used at every 3D location, enabling it to learn natural redundancies and correlations within an RI volume. The MLP is trained directly at test time by using only the IDT measurements of the sample that we seek to reconstruct. The IDT forward model is used to map the output of MLP to the intensity measurements and use the gradient back-propagation to update the MLP weights.
- DeCAF enables easy integration of additional prior knowledge on the unknown sample using an explicit regularization term in the loss function. In this paper we explored the potential of such synergistic integration by including an anisotropic 3D regularizer that separately imposes penalties in the x - y plane and z direction. Specifically, the x - y penalty uses a deep denoising CNN pre-trained on natural images to

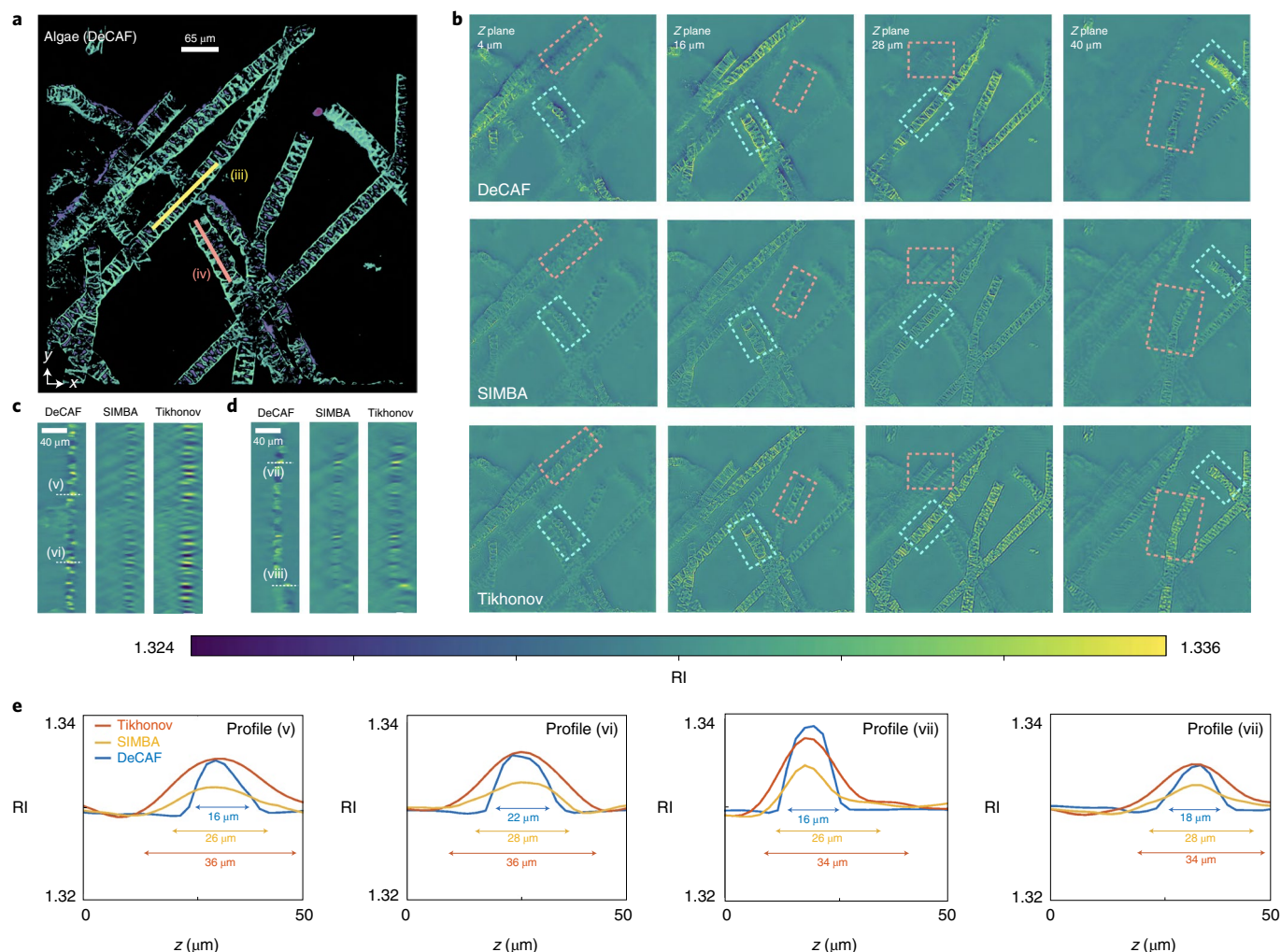


Fig. 2 | Reconstruction of *Spirogyra* algae acquired by dIDT. a, A 2D rendering obtained by accumulating all of the z slices from DeCAF. **b**, Visual comparison of the axial views at $z \in \{4, 16, 28, 40\}$ μm reconstructed by DeCAF, SIMBA and Tikhonov. **c,d**, Axial views corresponding to the coloured lines in **a**. **e**, Line profiles from the white dashed lines in **c** and **d**. This figure illustrates the ability of DeCAF to reconstruct high-contrast RI maps by also greatly reducing the missing cone artefacts. Note the quantitative demonstration of the reduction of elongation highlighted in **e**. Further examples can be found in the Supplementary Videos (spirogyra-decaf.mov, spirogyra-simba.mov and spirogyra-tikhonov.mov).

remove additive white Gaussian noise^{41,42}, whereas the z penalty is based on one-dimensional total variation. Although our denoising CNN was not trained explicitly on RI images, we show through ablation studies that it improves the performance by mitigating noise and imaging artefacts.

The pipeline of the proposed method is visually illustrated in Fig. 1a. In the training phase, the input of DeCAF is a set of spatial coordinates $\mathbf{c} = \{(x_i, y_i, z_i)\}_{i=1}^n$ taken from a pre-defined grid. DeCAF first maps the input coordinates to encodings using a non-trainable radial expansion, followed by a standard fully connected neural network to map the encodings to the RI values at the input coordinates. We introduced a novel type of encoding called radial encoding, which facilitates high-quality, artefact-free reconstruction of RI maps (see the Methods for details). DeCAF is trained to solve the following optimization with an objective consisting of a measurement loss $\mathcal{L}$ and regularizer $\mathcal{R}$

$$\begin{aligned} \phi^* &= \operatorname{argmin}_{\phi} \{ \mathcal{L}(\mathcal{F}(\mathbf{x}), \mathbf{y}) + \mathcal{R}(\mathbf{x}) \} \\ \text{such that } \mathbf{x} &= \mathcal{M}_{\phi}(\mathbf{c}), \end{aligned} \quad (1)$$

where $\mathbf{x}$ is the predicted RI map, $\mathbf{y}$ represents the intensity measurements of the test sample, $\mathcal{F}$ is the IDT forward model and $\mathcal{M}_{\phi}$ is the MLP (which includes the radial encoding) parameterized by weights ϕ . Note that the test measurements are the only input required in DeCAF. After the optimal ϕ^* is learned, one can render the test sample on a voxel grid with arbitrary density by simply querying $\mathcal{M}_{\phi^*}$ using the corresponding coordinates, as illustrated in Fig. 1e.

Past applications of NF include novel view synthesis^{43–45}, dynamic scene representation^{46–48}, object lightning^{49,50} and computed tomography⁵¹. Our work has several contributions to the existing NF literature: (1) DeCAF considers learning NF by accounting for the diffraction and scattering effects due to the wave nature of the light, whereas the existing work in the area has focused on ray-tracing models in graphics; (2) DeCAF extends the use of NF to the recovery of the phase information from intensity-only data; (3) DeCAF combines an implicit MLP regularization with an additional explicit image regularizer (for example, based on a deep denoiser) to achieve the best of both worlds, that is, to improve on the separate usage of an implicit and explicit regularization; (4) DeCAF introduces radial encoding as a novel type of encoding layer for improving the ability

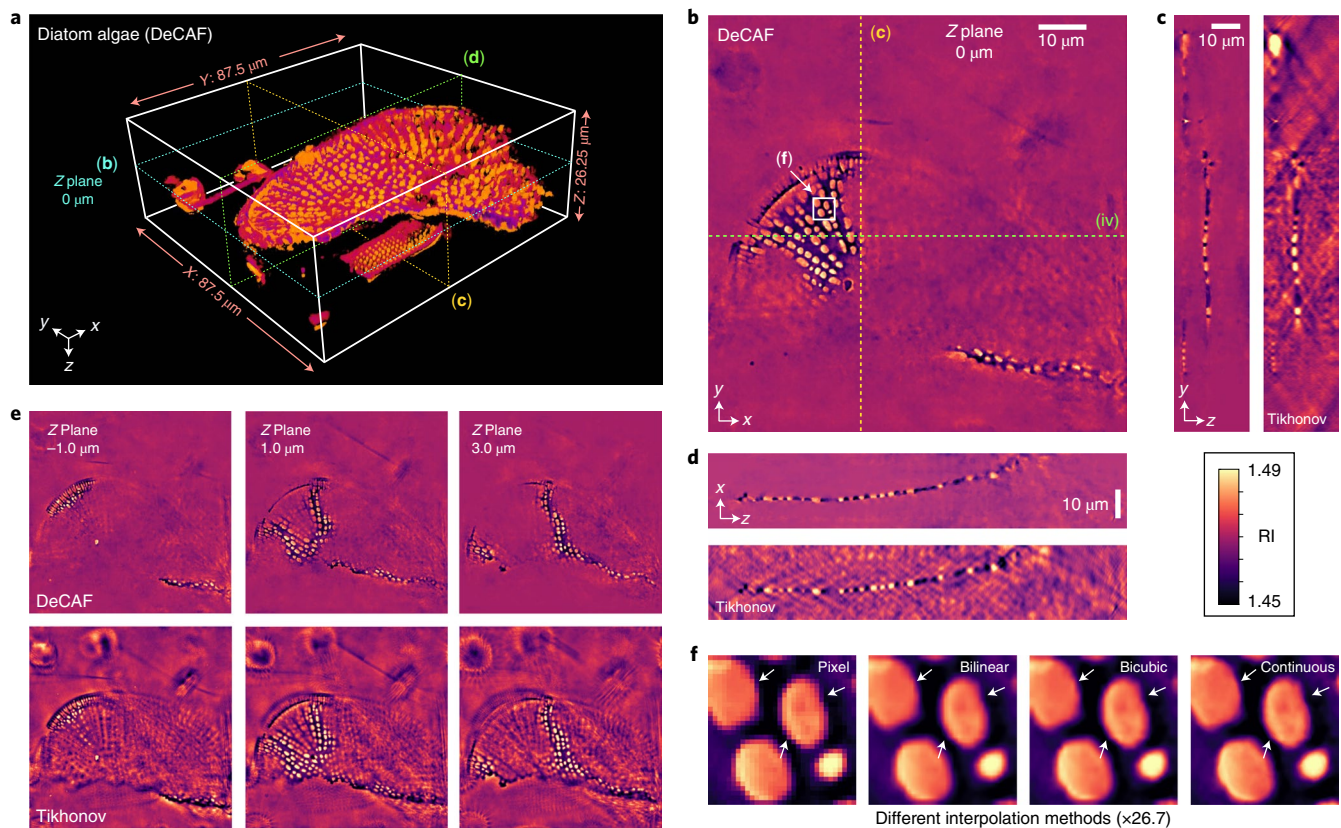


Fig. 3 | Reconstruction of diatom algae acquired by aIDT. a, A 3D illustration of the DeCAF reconstruction showing the overall structure of the sample. **b**, Axial view at $z = 0 \mu\text{m}$ (focal plane) reconstructed by DeCAF. **c, d**, Lateral views y - z (**c**) and x - z (**d**) corresponding to the colored paths in **b**. The Tikhonov results are also presented for reference. **e**, Visual illustrations of the axial views at $z \in \{-1, 1, 3\} \mu\text{m}$, highlighting better removal of artefacts than Tikhonov. **f**, Visual demonstration of DeCAF's ability to perform continuous RI upsampling along the x and y dimensions. DeCAF's results are consistent with that of the classic interpolation methods but provide finer details highlighted by the arrows. More examples are shown in the Supplementary Videos (diatom-decaf.mov and diatom-tikhonov.mov).

of NF to represent complex samples. Details on the network architecture and the learning procedure of DeCAF are provided in the Methods and the Supplementary Information. In the next section, we present both qualitative and quantitative results that demonstrate DeCAF's ability to reconstruct high-quality RI maps.

Results. Experimental validation. We validated DeCAF's ability to recover high-quality RI maps with accurate biological features and minimal artefacts on experimentally collected IDT data. We used DeCAF on four biological samples, including *Spirogyra* and diatom algae, human buccal epithelial cells and *Caenorhabditis elegans* (*C. elegans*). We adopted the existing light-propagation models to formulate the inverse problems associated with the dense¹⁴, annular²⁰ and multiplexed⁵² illumination patterns. As absorption provides a lower contrast for the considered samples, we focus on comparing the phase images. In the subsequent sections we use x , y and z to denote length, width and depth, respectively.

We first show the effectiveness of DeCAF for dense IDT (dIDT) on stained *Spirogyra* (Fisher Scientific S68786, embedded in water, $\text{RI} \approx 1.33$)—unicellular algae containing a helical arrangement of chloroplasts oriented in the 3D space. In total we collected 89 bright-field intensity measurements using a 0.25 NA objective lens; two example measurements are presented in Fig. 1c(i,ii). In the experiment we compared DeCAF with two existing IDT reconstruction baselines: Tikhonov regularization¹⁴ and SIMBA⁵³, as both methods have been extensively validated under similar imaging settings. SIMBA is a recently proposed model-based algorithm that leverages a deep learning denoiser as an image prior. Each method's final

reconstructed RI volume consists of 40 axial slices of $1,024 \times 1,024$ pixels, equally spaced between $-30 \mu\text{m}$ and $50 \mu\text{m}$, forming a volume of $665.6 \times 665.6 \times 80 \mu\text{m}^3$. We define $z = 0 \mu\text{m}$ as the focal plane throughout the paper.

Figure 2 visualizes the experimental results. To demonstrate the overall structure of the sample, a rendered 2D image that accumulates all of the z layers of the DeCAF reconstruction is presented in Fig. 2a. As shown, DeCAF successfully reconstructed the spiral structure of the *Spirogyra*. Figure 2b compares the 2D axial slices obtained by DeCAF, SIMBA and Tikhonov at the depths $z \in \{4, 16, 28, 40\} \mu\text{m}$. The results show that DeCAF provides superior axial sectioning ability (that is, a pattern emerges only in the slices it belongs to and fades rapidly as we move axially to different depths) than the other two methods. This is demonstrated by the clarity and sharpness of the spirals (in the dashed cyan boxes) that appear at a specific depth, which show that DeCAF removes the artefacts (in the red dashed boxes) generated by the diffraction from the adjacent slices. These artefacts remain in the reconstructions by SIMBA and Tikhonov. We further evaluate the axial resolution of each reconstruction by comparing the lateral views that correspond to the cutlines shown in Fig. 2c,d. DeCAF substantially reduces the elongation artefacts caused by the missing cone problem. Line profiles presented in Fig. 2e quantitatively characterize the reduction of z -elongation by DeCAF.

We next applied DeCAF to annular IDT (aIDT) to explore its capability for efficient data processing. We imaged two distinct classes of biological samples, including diatom microalgae (S68786, Fisher Scientific) and unstained human epithelial buccal cells. The

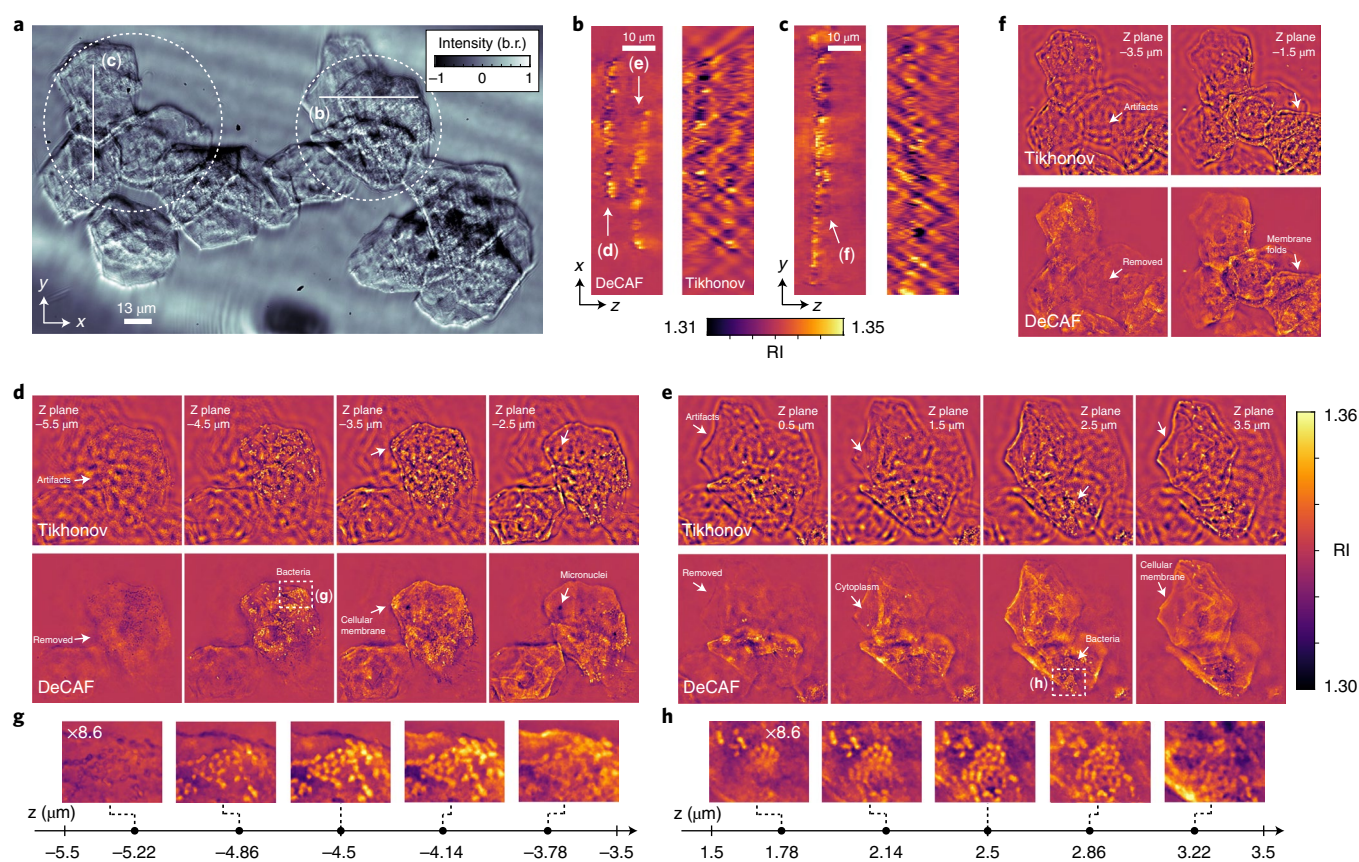


Fig. 4 | Reconstruction of a human buccal epithelial cell cluster acquired by aIDT. **a**, Example intensity measurement collected by our aIDT set-up for the cell cluster. Note that the background light is removed (b.r.) from the image. **b,c**, x - z (**b**) and y - z (**c**) lateral views of the DeCAF reconstruction associated with the paths in **a**. Superior performance in artefact removal and axial separation is demonstrated compared with Tikhonov regularization. **d-f**, The axial views at various depths of the two sub-cell clusters shown in **b** and **c**. These results further highlight the strong axial sectioning effects as well as the fine details preserved by DeCAF. **g,h**, Visual demonstration of DeCAF's ability to perform continuous RI upsampling along the z dimension. Note that $\{-5.5, -4.5, -3.5\} \mu\text{m}$ and $\{1.5, 2.5, 3.5\} \mu\text{m}$ are the only axial points used during training. Smooth and consistent transition is observed in the appearance of the bacteria. More examples are shown in Supplementary Videos (cell-decaf-b.mov, cell-decaf-c.mov, cell-tikhonov-b.mov and cell-tikhonov-c.mov).

former is a unicellular algae with regular arrangement of punctae, whereas the latter is a complex cell environment consisting of intracellular bacteria. We acquired 24 intensity images using a 0.65 NA objective lens under oblique illuminations for each sample. The diatom and cell cluster samples are fixed in glycerin gelatin ($\text{RI} \approx 1.47$) and water, respectively. We used Tikhonov as the baseline method for comparison.

Figure 3 presents the results for diatom algae. Two example measurements are provided in Fig. 1c(iii,iv). Both DeCAF and Tikhonov were configured to reconstruct 52 slices of 700×700 pixels equally spaced between $-10 \mu\text{m}$ and $16 \mu\text{m}$, forming a volume of $113.75 \times 113.75 \times 26 \mu\text{m}^3$. The 3D illustration of the volume reconstructed by DeCAF is presented in Fig. 3a, demonstrating the overall reconstruction quality. Figure 3e presents the slices reconstructed by each method at depths $z \in \{-1, 1, 3\} \mu\text{m}$. DeCAF demonstrates better sectioning capability than Tikhonov regularization. Superior removal of the missing cone artefacts is also shown in the lateral views in Fig. 3c,d. As DeCAF learns a continuous representation of the RI distribution, it can generate images on arbitrarily dense voxel grids without additional retraining. Figure 3f demonstrates this unique ability of DeCAF by interpolating 26.7-times more pixels than the original reconstruction in the x - y planar region shown in Fig. 3b. For comparison, we used nearest neighbour (Pixel), bilinear (Bilinear) and bicubic (Bicubic) interpolation methods to upsample the same region. Our results show that DeCAF is able to resolve

small features with strong cross-scale consistency while avoiding interpolation artefacts highlighted by the arrows.

We next present the results of epithelial buccal cell clusters in Fig. 4. A background-removed intensity measurement showing the distribution of the whole cell cluster is presented in Fig. 4a. In Fig. 4b,c we focus on two complex regions where cells overlap with each other to highlight the superior axial sectioning capability of DeCAF. The size of the two volumes are $81.25 \times 81.25 \times 16 \mu\text{m}^3$ and $97.5 \times 97.5 \times 16 \mu\text{m}^3$, discretized to 32 slices of 500×500 and 600×600 pixels, respectively. DeCAF successfully resolves different cells with clear separation, whereas Tikhonov generates strong artefacts that blur the boundaries. A visual demonstration of the axial slices of these cells is provided in Fig. 4d-f. In each reconstructed slice, DeCAF recovers clear cell membrane, cytoplasm, micronuclei and bacteria while removing the diffraction and scattering artefacts.

We further show the continuous representation learned by DeCAF by upsampling it along z , meaning that DeCAF is used to interpolate an entire axial slice that was not part of the grid used during training. Figure 4g,h presents the interpolated slices of the bacteria clusters highlighted in Fig. 4d,e. A z -axis is provided in each figure to show the axial location of each slice. Note that $\{-5.5, -4.5, -3.5\} \mu\text{m}$ and $\{1.5, 2.5, 3.5\} \mu\text{m}$ are the axial coordinates pre-defined in the training grid. The interpolated slices in Fig. 4g,h clearly show the appearance and disappearance of the

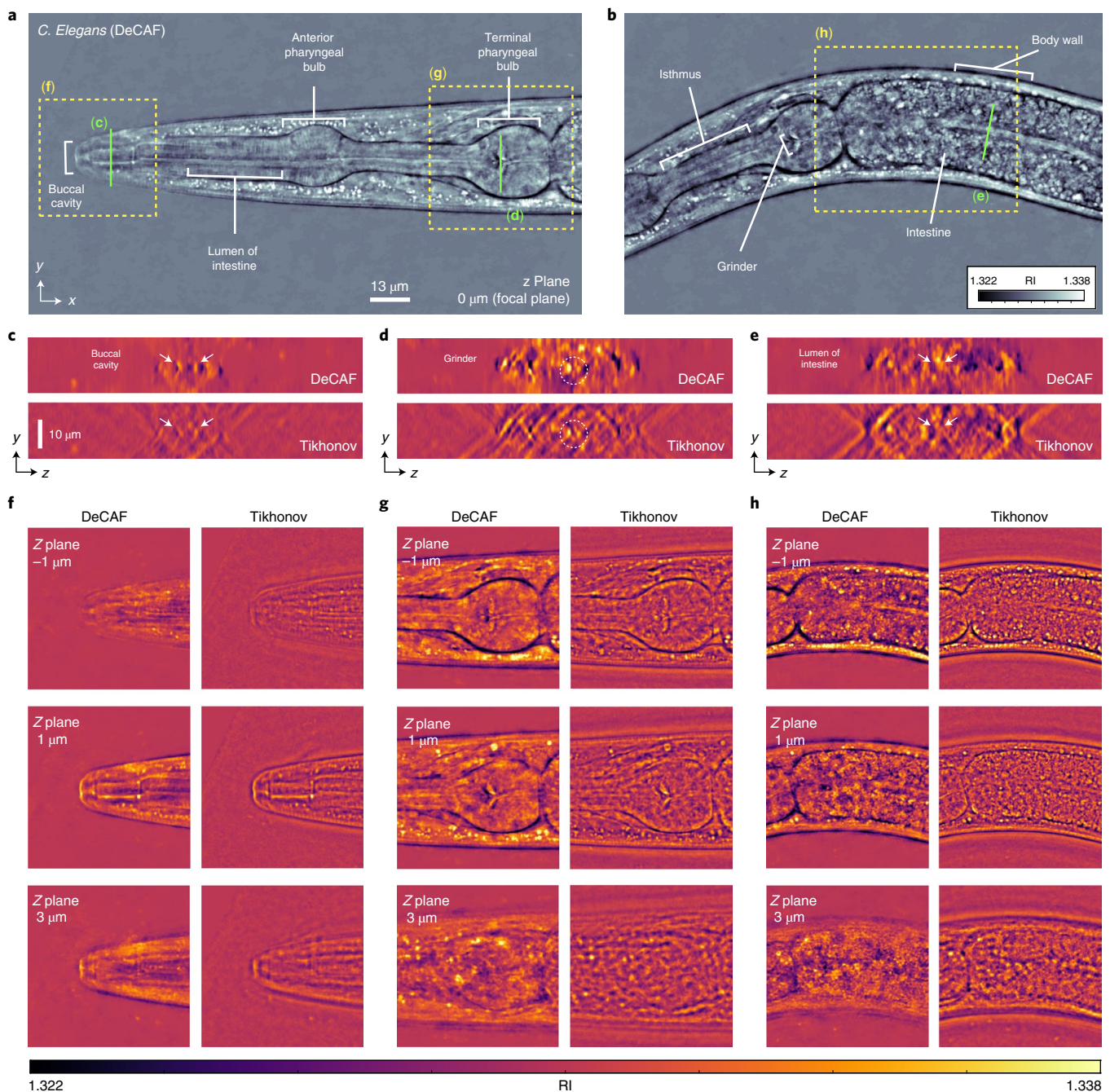


Fig. 5 | Reconstruction of *C. elegans* acquired by mIDT. a, b, The reconstructed RI distribution at $z = 0 \mu\text{m}$ (focal plane) by DeCAF. **c–e,** Lateral views corresponding to the paths shown in **a** and **b**. Biological structures are highlighted by the arrows and circles. **f–h,** Axial views of the regions highlighted in **a** and **b** at $z \in \{-1, 1, 3\} \mu\text{m}$. Note how DeCAF provides higher contrast and finer details than Tikhonov. More examples can be found in the Supplementary Videos (celegans-decaf-head.mov, celegans-decaf-body.mov, celegans-tikhonov-head.mov and celegans-tikhonov-body.mov).

bacteria clusters at different values of z . As shown, the interpolated biological features are consistent with those lying in the pre-defined grid, making the whole transition smooth across axial layers. This strong axial consistency preserved in DeCAF enables it to produce high-fidelity interpolations without any additional retraining.

We finally validate DeCAF on the multiplexed IDT microscopy (mIDT). This modality allows more rapid acquisition under the same time by simultaneously illuminating the sample from multiple angles for each intensity measurement. We imaged a *C. elegans* worm specimen by using a 0.65 NA objective lens to acquire 16 measurements, with each from the simultaneous illuminations of

six different light-emitting diode (LED) sources. Figure 1c(v,vi) shows two example measurements. The sample is challenging due to its thickness and complicated arrangement of organs. As the worm is alive and moving during the acquisition, we reconstructed two volumes of $162.5 \times 162.5 \times 20 \mu\text{m}^3$, discretized to 40 slices of $1,000 \times 1,000$ pixels, at different times to cover the worm bodies with interested biological features. Extended Data Fig. 1 also demonstrates DeCAF's ability to reconstruct a relatively thin diatom algae sample from mIDT measurements.

Figure 5a,b presents the reconstructed RI maps of *C. elegans* at the focal plane. DeCAF successfully recovers the sample's structure

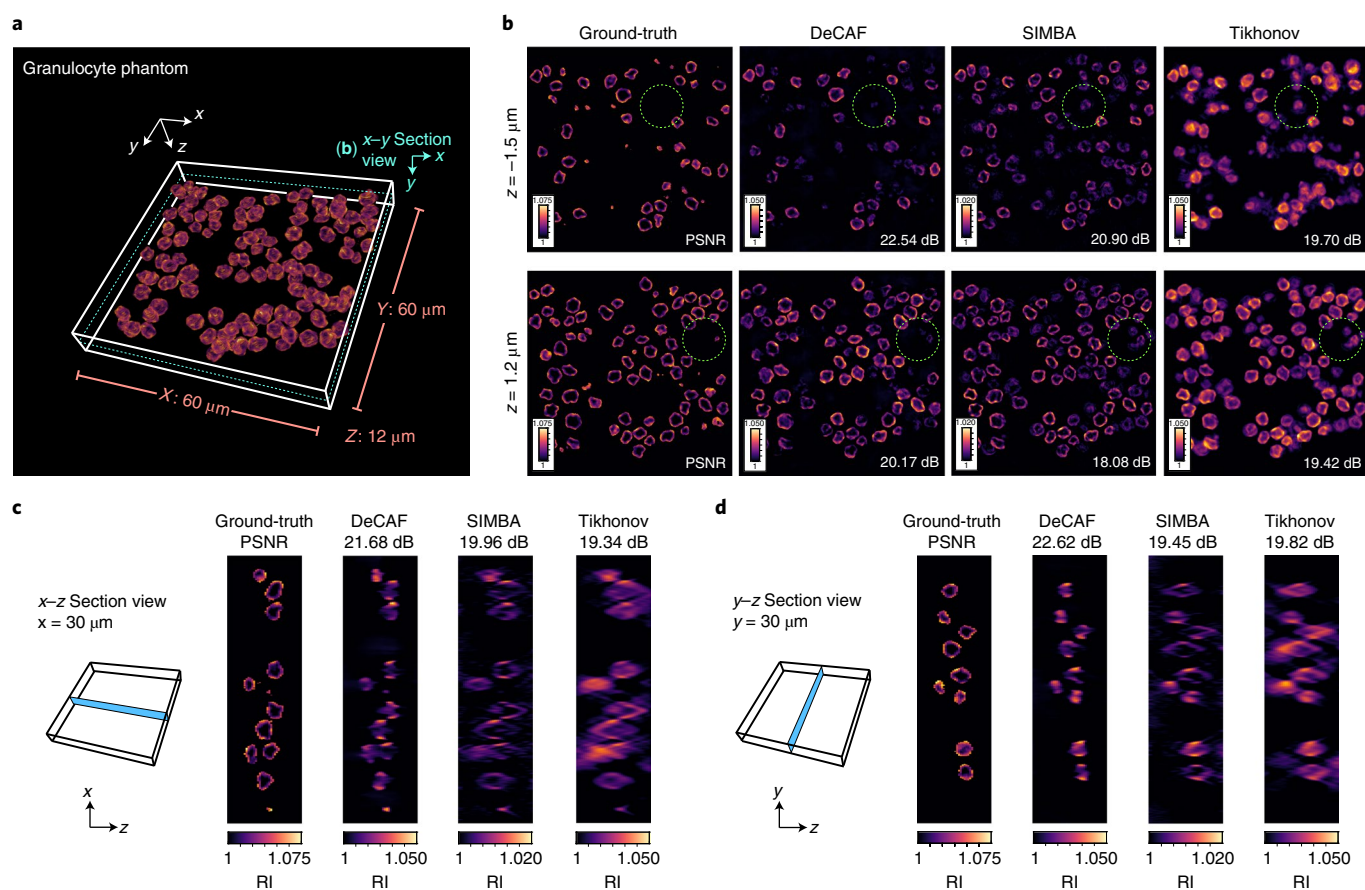


Fig. 6 | Quantitative and visual comparison of DeCAF, SIMBA and Tikhonov for the reconstruction of the granulocyte phantom from simulated data.

a, A 3D illustration of the phantom. **b**, Visual comparison of the axial views at $z \in \{-1.5, 1.2\}$ μm reconstructed using all three methods. The ground-truth is provided in the left-most column. Visual differences are highlighted in the green dashed circles. **c, d**, Visual comparison of the x - z (**c**) and y - z (**d**) lateral views reconstructed using each method. The corresponding position of each view is shown in the left of each figure. Note how DeCAF provides much higher PSNR values than both SIMBA and Tikhonov.

with clear quantification of the internal biological tissues. For example, the buccal cavity, anterior and terminal pharyngeal bulbs, isthmus and intestine are clearly restored in our reconstruction. Smaller features are also distinguishable with high contrast, as shown in the regions expanded in Fig. 5f–h. For example, lysosomes, a grinder and the lumen of intestine are accurately visualized with clear separation from the other tissues. Figure 5c–e shows the y - z lateral views, where the oval shape of *C. elegans* is reconstructed without the missing cone artefacts, and fine features such as buccal cavity and grinder are preserved and recovered at different axial layers.

Extended Data Fig. 2 highlights the space used for storing the MLP weights in DeCAF. As the representation is decoupled from a predefined voxel grid, DeCAF can be trained on a sparse grid to reduce the storage cost, but can still produce the final reconstruction on a grid of desired density. The storage reduction is demonstrated by comparing the memory requirements of DeCAF and Tikhonov for the reconstruction of the *C. elegans* worm. DeCAF retains the small memory size of 3 MB across different grid densities while that of Tikhonov increases as the grid becomes denser.

Quantitative evaluation. In this section we present quantitative evaluations of DeCAF using a high-fidelity cell phantom. We used CytoPacq⁵⁴ to generate a granulocyte phantom containing tens of granulocyte cells randomly distributed in a volume of $60 \times 60 \times 12 \mu\text{m}^3$, discretized to 40 slices of 454×454 pixels. The

maximum RI value in the cell is set to $n_{\text{re}} = 1.075$ and $n_{\text{im}} = 0$ (the real and imaginary parts of the sample's RI, respectively), meaning that there is no absorption. The immersion media is assumed to be air ($n_0 = 1$). The simulation is based on the aIDT set-up and used the split-step non-paraxial simulator to simulate the full wave propagation^{55,56}. The set-up includes an annular LED array at 515 nm wavelength for illumination and an objective lens with 0.65 NA. In total, 24 measurements are taken during the acquisition.

The result of quantitative evaluation are summarized in Fig. 6. Figure 6a shows the overall 3D structure of the phantom. Section views of axial and lateral planes are compared for DeCAF, SIMBA and Tikhonov, with quantitative evaluation of the peak signal-to-noise ratio (PSNR) values

$$\text{PSNR}(\mathbf{x}, \hat{\mathbf{x}}) = 10 \log \left(\frac{\max(\mathbf{x})}{\text{mse}(\mathbf{x}, \hat{\mathbf{x}})} \right), \quad (2)$$

where the mean squared error is computed by $\text{MSE}(\cdot, \cdot)$ and the maximum pixel value in the image is returned by $\max(\cdot)$. Figure 6b visualizes the axial slices reconstructed by each method at $z \in \{-1.5, 1.2\}$ μm . Figure 6c, d plots the x - z and y - z section views. The corresponding position of each view is shown in the left of each figure. DeCAF achieves better PSNR than both SIMBA and Tikhonov by reconstructing more accurate RI values, reducing the

missing cone artefacts, and removing the cell shadows due to axial elongations (highlighted using dashed circles).

Extended Data Fig. 3 visualizes the 3D volumes reconstructed by each method using Fiji⁵⁷ under the default configuration. From left to right, the figure displays the 3D volumes corresponding to the ground-truth, DeCAF, SIMBA and Tikhonov. Peak signal-to-noise ratio values are labelled on each volume in green. DeCAF clearly outperforms SIMBA and Tikhonov by reconstructing cells that look most similar to the ground-truth. For example, consider the cells highlighted in the zoom-in volume. DeCAF reconstructs these cells with clear shapes and sharp edges, while the reconstructions of SIMBA and Tikhonov are either axially elongated or blurry. Quantitative results further highlight the accuracy of DeCAF, showing PSNR improvements of 1.6 dB and 3.3 dB with respect to SIMBA and Tikhonov, respectively (equivalent to a 1.5- and 2.1-fold reduction in MSE).

Discussion. *Difference to SIMBA.* DeCAF offers several benefits from the existing SIMBA method. First, test-time learning: SIMBA does not adapt to the specifics of a test sample—it uses a fixed forward model and a fixed pre-trained prior. On the other hand, DeCAF is a test-time learning method in which the MLP weights are adjusted for each test sample, leading to a better reconstruction performance reported throughout this paper. Second, grid-free representation: SIMBA reconstructs a discrete volume on a pre-defined voxel grid. DeCAF decouples the representation of the reconstructed 3D RI from the grid by using MLP. This enables one to synthesize any part of the 3D RI volume ‘on demand’ on any grid by simply querying the relevant coordinates of MLP. Thus, the complexity of storing the sample reconstructed by DeCAF is decoupled from the voxel-grid. Third, internal and external regularization: unlike SIMBA, DeCAF synergistically uses internal and external regularization offered by MLP and a CNN denoiser, respectively. Our quantitative results show that MLP offers a substantial amount of regularization, even when no external regularizer is used; however, the best results are achieved when both regularizers are used.

Limitations of DeCAF. An obvious limitation of DeCAF is that it is based on the linear IDT forward models that are based on the first Born approximation. This limits the applicability of the current implementation to relatively thin and weakly scattering samples. This limitation can be observed in the reconstruction of a relatively thick *C. elegans* sample. Future work will explore the extension of DeCAF to thicker and stronger scattering samples by using forward models accounting for multiple scattering, such as the ones based on the variations of the beam propagation method^{12,58}. Another limitation of DeCAF is that it is currently slower than existing IDT reconstruction methods, Tikhonov and SIMBA, which is due to our implementation of the NF training. Our model takes less than a day (~20h) to infer each real sample, while the runtimes of Tikhonov and SIMBA are at the levels of several minutes and hours, respectively. Furthermore, DeCAF’s hyperparameters need to be tuned manually on real samples due to the lack of ground-truth, which potentially leads to further increases in runtime. The future work will explore faster DeCAF implementations that leverage recent progress in accelerating NF methods (for example, Instant Neural Graphics Primitives⁵⁹ suggests an order of magnitude acceleration).

Conclusion. We proposed a novel self-supervised deep learning method, DeCAF, for enabling high-quality 3D reconstruction of the RI distribution from intensity-only measurements. We extensively validated DeCAF on the experimentally collected datasets of multiple biological samples under three different IDT set-ups. The results show that DeCAF can mitigate the missing cone artefacts while maintaining the fine details of small biological features. We also provide quantitative evidence to further corroborate our

argument. Results show that DeCAF can reduce MSE by up to 2.1-fold. The continuous representation in DeCAF also allows to generate images at voxel grids of arbitrary density without retraining of the deep network, which is useful for addressing computational and memory bottlenecks in image reconstruction and analysis.

Methods

IDT experiments. *IDT resolution.* The lateral and axial resolutions of the IDT system are limited by the support of the optical transfer function, which is determined by the objective NA and illumination NA¹⁴. For both the aIDT and mIDT set-up, our maximum illumination angle is close to the objective NA; thus, the recovered lateral spatial frequency can reach the incoherent diffraction limit $4NA/\lambda$, and the axial Fourier coverage is up to $(2n_0 - 2\sqrt{n_0^2 - NA^2})/\lambda$, where n_0 is the RI of background media.

Dense IDT. Our dense IDT system consists of: a Nikon TE 2000-U microscope equipped with a custom programmable LED array (approximately illuminating the plane wave with a central wavelength of $\lambda = 632$ nm); a $\times 10/0.25$ NA objective (Nikon, CFI Plan Achromat); and an sCMOS camera (PCO.Edge 5.5). The LED array is placed about 79 mm away from the sample. It is controlled via a microcontroller and is synchronized with the camera. A small subset of the LEDs on the array—containing the 89 LEDs within the brightfield region—is used to illuminate the sample sequentially.

Annular IDT. Our annular IDT system consists of a Nikon ECLIPSE E200 microscope equipped with a programmable ring LED unit (Adafruit, 1586 NeoPixel Ring). The microscope objective is $\times 40/0.65$ NA (Nikon, CFI Plan Achromat), and each LED approximately provides a plane wave with a central wavelength of $\lambda = 515$ nm. The ring LED unit has 24 LED lights and is 60 mm in diameter. It is centered at the optical axis and placed approximately 35 mm away from the sample, which sets the angle between the wave vector and the optical axis to about 40° and complies with the microscope objective NA.

Multiplexed IDT. Our multiplexed IDT system has the same hardware specification as the dense IDT system except that the microscope objective is $\times 40/0.65$ NA (Nikon, CFI Plan Achromat). Besides, the subset of the LEDs used in the experiment changes to 96 LEDs corresponding to the NA range from 0.3 to 0.575. This design contains 16 disjoint illumination patterns and the multiplexed illumination quantity of each pattern is 6. The camera is synchronized with the LED array and captures 16 measurements corresponding to each illumination pattern.

Sample and data preparation. *Spirogyra algae.* This sample is a part of Fisher Science Education algae basic slide set S68786. We captured 89 intensity-only bright field measurements. We pre-processed each measurement by removing the background intensity followed by normalization. The same pre-processing procedure is also applied to other samples. We consider a reconstruction volume of $665.6 \times 665.6 \times 80 \mu\text{m}^3$, positioned between $-30 \mu\text{m}$ and $50 \mu\text{m}$ around the focal plane. The volume is discretized into 40 slices along the z -axis, with each slice having $1,024 \times 1,024$ pixels. Here, a single voxel corresponds to $6.5 \times 6.5 \times 2 \mu\text{m}^3$.

Diatom algae (aIDT). This sample is a part of Fisher Science Education algae basic slide set S68786. We captured 24 measurements and consider a reconstruction volume of $113.75 \times 113.75 \times 26 \mu\text{m}^3$, positioned between $-10 \mu\text{m}$ and $16 \mu\text{m}$ around the focal plane. The volume is discretized into 52 slices along the z -axis, with each slice having 700×700 pixels. Here, a single voxel corresponds to $0.1625 \times 0.1625 \times 0.5 \mu\text{m}^3$.

Diatom algae (mIDT). This sample is a part of Fisher Science Education algae basic slide set S68786. We captured 16 measurements, and each measurement used six LEDs. We consider a reconstruction volume of $130 \times 130 \times 30 \mu\text{m}^3$, positioned between $-15 \mu\text{m}$ and $15 \mu\text{m}$ around the focal plane. The volume is discretized into 60 slices along the z -axis, with each slice having 800×800 pixels. Here, a single voxel corresponds to $0.1625 \times 0.1625 \times 0.5 \mu\text{m}^3$.

Human buccal epithelial cells. This sample was swabbed from a researcher’s buccal. The individual rinsed the mouth with clean water and then twirled a wooden swab against the inner cheek. The end of the swab was immersed in a drop of purified water on a glass slide and covered by a coverslip. We captured 24 measurements of the cell cluster and consider two volumes in the region as shown in Fig. 4b,c. The former has $81.25 \times 81.25 \times 16 \mu\text{m}^3$ and the latter has $97.5 \times 97.5 \times 16 \mu\text{m}^3$. Both volumes are positioned between $-8 \mu\text{m}$ and $8 \mu\text{m}$ around the focal plane. They are discretized to 32 slices of 500×500 and 600×600 pixels. Here, a single voxel corresponds to $0.1625 \times 0.1625 \times 0.5 \mu\text{m}^3$.

C. elegans. Young adult *C. elegans* were mounted on 3% agarose pads in a drop of nematode growth medium buffer. Glass coverslips were then gently placed on top of the pads and sealed with a 1:1 mixture of paraffin and petroleum jelly.

As the *C. elegans* were alive and moving during data acquisition, we captured a video at 4 fps in which each frame contained 16 measurements and each measurement used six LEDs. We picked two frames at 1.5 s and 44 s for reconstruction, where the sample was relatively steady. We consider a unified reconstruction volume of $162.5 \times 162.5 \times 20 \mu\text{m}^3$, positioned between $-10 \mu\text{m}$ and $10 \mu\text{m}$ around the focal plane. The volume is divided into 40 slices along the z -axis, with each slice having $1,000 \times 1,000$ pixels. Here, a single voxel corresponds to $0.1625 \times 0.1625 \times 0.5 \mu\text{m}^3$.

DeCAF framework. A linearized approximation of IDT forward measurement system can be described by equation (3)

$$\mathbf{y}_\rho \approx \mathbf{A}_\rho \Delta \epsilon, \quad (3)$$

where $\Delta \epsilon = \Delta \epsilon_{\text{re}} + j \Delta \epsilon_{\text{im}}$ is the unknown volume of complex-valued permittivity contrast, $\mathbf{y}_\rho$ is the collection of the background-removed intensity measurements corresponding to the LED illuminations emitted at a set of locations ρ , and $\mathbf{A}_\rho$ is the measurement matrices that model the sample-intensity mapping associated with these illuminations. The reconstruction of $\Delta \epsilon$ is equivalent to the reconstruction of the RI distribution via equation (4)

$$\mathbf{n}_{\text{re}} = \sqrt{\frac{1}{2} \left((\mathbf{n}_0^2 + \Delta \epsilon_{\text{re}}) + \sqrt{(\mathbf{n}_0^2 + \Delta \epsilon_{\text{re}})^2 + \Delta \epsilon_{\text{im}}^2} \right)} \quad \text{and} \quad \mathbf{n}_{\text{im}} = \frac{\Delta \epsilon_{\text{im}}}{2 \cdot \mathbf{n}_{\text{re}}}, \quad (4)$$

where $\mathbf{n}_{\text{re}}$ and $\mathbf{n}_{\text{im}}$ are the real and imaginary parts of the sample's RI, and $\mathbf{n}_0$ is the RI of the background medium (where the attenuation is often assumed to be zero). In equation (4), all operations are evaluated in an element-wise manner. We derived the formulations of $\mathbf{A}_\rho$ by following past works on dIDT¹⁴, aIDT²⁰ and mIDT⁵² (see the 'IDT forward model' section in the Supplementary Information).

The central piece of DeCAF is a coordinate-based MLP, $\mathcal{M}_\phi$, which maps the 3D coordinate (x, y, z) to the corresponding values of $\Delta \epsilon_{\text{re}}$ and $\Delta \epsilon_{\text{im}}$. We normalize the coordinate grid to a cube $[-1, 1]^3$ before feeding them into $\mathcal{M}_\phi$. The deep network $\mathcal{M}_\phi$ consists of two subnetworks, where the first one is an encoding layer $\gamma(x, y, z)$, pre-defined before training, and the second one is a standard MLP $\mathcal{N}_\phi: \gamma(x, y, z) \rightarrow (\Delta \epsilon_{\text{re}}, \Delta \epsilon_{\text{im}})$ parameterized by the trainable parameters ϕ . A visual illustration of the detailed network architecture is provided in the Extended Data Fig. 4a.

Radial encoding. It has been shown that a Fourier-type encoding of the spatial coordinates is essential for a MLP to represent high-frequency variations in the signal⁴³ and impose implicit regularization⁶⁰. In DeCAF, we consider a decomposition of the input coordinate (x, y, z) into (x, y) and z , and use different strategies to expand (x, y) and z . This is due to the non-isotropic resolution of the imaging system along the x - y plane and the z dimension. Our experiments showed that existing encoding strategies, such as positional⁴³ and Gaussian⁶⁰ encoding, lead to suboptimal reconstruction of RI images along the x - y dimensions. We propose radial encoding as an alternative for expanding $\mathbf{v} := (x, y)$

$$\gamma_{\text{rad}}(\mathbf{v}) = \begin{pmatrix} \sin(2^0 \pi \mathbf{R}_0 \mathbf{v}), \cos(2^0 \pi \mathbf{R}_0 \mathbf{v}), \\ \vdots \\ \sin(2^{L_{xy}-1} \pi \mathbf{R}_0 \mathbf{v}), \cos(2^{L_{xy}-1} \pi \mathbf{R}_0 \mathbf{v}) \end{pmatrix} \quad (5)$$

with $\mathbf{R}_0 = \left\{ \begin{bmatrix} \cos(\theta_k) & -\sin(\theta_k) \\ \sin(\theta_k) & \cos(\theta_k) \end{bmatrix} \right\}_{k=1}^K$.

Here, $\sin$ and $\cos$ compute the (element-wise) sinusoidal and cosinusoidal values, respectively, $\mathbf{R}_0$ denotes a collection of rotation matrices that translate the coordinates by the angles θ , and $L_{xy} > 0$ controls the number of the expanded frequency. By incorporating rotation, our strategy enables a frequency expansion that can efficiently account for the dependencies within the x - y plane (see the 'Radial Encoding' section in the Supplementary Information). The difference between radial encoding and positional encoding is conceptually illustrated in Extended Data Fig. 4b,c. In the experiments we observed that the radial encoding improves the representation of small textures that are otherwise lost by other encodings. We adopted the standard positional encoding for the expansion of z

$$\gamma_{\text{pos}}(z) = \begin{pmatrix} \sin(2^0 \pi z), \cos(2^0 \pi z), \\ \vdots \\ \sin(2^{L_z-1} \pi z), \cos(2^{L_z-1} \pi z) \end{pmatrix}, \quad (6)$$

where $L_z > 0$ denotes the total number of frequencies. We fine-tuned θ , L_{xy} , and L_z for every sample by running multiple sets of parameters and manually selecting

the set leading to the best visual quality. We summarize their values in Extended Data Fig. 5. The ablation experiment on the challenging *C. elegans* specimen (see Supplementary Fig. 4 for details) demonstrates the superior performance of the proposed encoding. More quantitative evidence can be found in the 'Ablation experiments' section in the Supplementary Information.

MLP architecture. The network architecture of $\mathcal{N}_\phi$ is illustrated in Extended Data Fig. 4a. Network $\mathcal{N}_\phi$ is composed of N fully connected layers. The first $N-1$ layers have M hidden neurons activated by the leaky rectified linear unit (leaky ReLU), while the last layer has M unactivated hidden neurons. A skip connection is implemented at the $\lfloor N/2 \rfloor^{\text{th}}$ fully connected layer to concatenate the original input of $\mathcal{N}_\phi$ with the intermediate outputs, which has been shown beneficial for improving the representation performance⁶². We used one network configuration for all biological samples, which is summarized in Extended Data Fig. 5.

Regularized loss function. At test time, we train $\mathcal{M}_\phi$ to minimize equation (7) by using a customized Adam⁶³ optimizer (see the 'Block-wise training of DeCAF' section in the Supplementary Information)

$$\begin{aligned} \mathcal{L}(\phi; \mathbf{y}_\rho) = & \underbrace{\| \mathbf{A}_\rho(\mathcal{M}_\phi(\mathbf{c})) - \mathbf{y}_\rho \|_1}_{\text{measurement consistency}} + \alpha \underbrace{\| \mathcal{M}_\phi(\mathbf{c}) - \mathbf{D}_\sigma(\mathcal{M}_\phi(\mathbf{c})) \|_2^2}_{x-y \text{ plane noise reduction}} \\ & + \beta \underbrace{\sum_j \| \mathcal{M}_\phi^j(\mathbf{c}) - \mathcal{M}_\phi^{j-1}(\mathbf{c}) \|_1}_{z\text{-dimension continuity}}, \end{aligned} \quad (7)$$

where $\mathbf{c} = \{(x_i, y_i, z_i)\}_{i=1}^n$ is a collection of all coordinates on the grid and $\mathcal{M}_\phi^j$ denotes the j th axial slice of the predicted RI map. The loss defined in equation (7) can be divided into three terms serving different purposes, with $\alpha \geq 0$ and $\beta \geq 0$ balancing their contributions. The first term is a widely used ℓ_1 -norm loss that ensures the consistency with the test measurements. The second and third terms are the regularizers imposing x - y plane noise reduction and continuity along z , respectively. The $\mathbf{D}_\sigma$ term denotes a 2D image denoiser with $\sigma > 0$ controlling the denoising strength. We selected a denoising convolutional neural network due to its state-of-the-art denoising performance⁶⁴. A detailed description of the architecture and training of the denoising convolutional neural network is presented in the 'Additional technical details' section in the Supplementary Information. We fine-tuned α and β for each sample by running multiple sets of parameters and manually selecting the set leading to the best visual quality. We summarized their values in Extended Data Fig. 5 and provide the empirical evidence to demonstrate the necessity of the explicit regularization for imaging the complex organism (see the visual results in Supplementary Fig. 6). Additional quantitative evaluations are provided in the 'Ablation experiments' section in the Supplementary Information.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data used for reproducing the results in the manuscript are available at <https://github.com/wustl-cig/DeCAF65>. We visualized the pre-processed raw intensity images of the relevant samples in Figs. 1 and 4.

Code availability

The code used for reproducing the results in the manuscript is available at <https://github.com/wustl-cig/DeCAF65>.

Received: 30 November 2021; Accepted: 5 August 2022;
Published online: 16 September 2022

References

- Kim, K. et al. Three-dimensional label-free imaging and quantification of lipid droplets in live hepatocytes. *Sci. Rep.* **6**, 36815 (2016).
- Yamada, K. M. & Cukierman, E. Modeling tissue morphogenesis and cancer in 3D. *Cell* **130**, 601–610 (2007).
- Kim, G. et al. Measurements of three-dimensional refractive index tomography and membrane deformability of live erythrocytes from pelophylax nigromaculatus. *Sci. Rep.* **8**, 9192 (2018).
- Cooper, K. L. et al. Multiple phases of chondrocyte enlargement underlie differences in skeletal proportions. *Nature* **495**, 375–378 (2013).
- Park, Y. K., Depeursinge, C. & Popescu, G. Quantitative phase imaging in biomedicine. *Nat. Photonics* **12**, 578–589 (2018).
- Jin, D., Zhou, R., Yaqoob, Z. & So, P. Tomographic phase microscopy: principles and applications in bioimaging. *J. Opt. Soc. Am. B* **34**, B64–B77 (2017).
- Park, Y. K. et al. Refractive index maps and membrane dynamics of human red blood cells parasitized by plasmodium falciparum. *Proc. Natl Acad. Sci. USA* **105**, 13730–13735 (2008).

8. Sung, Y. et al. Optical diffraction tomography for high resolution live cell imaging. *Opt. Express* **17**, 266–277 (2009).
9. Kamilov, U. S. et al. A learning approach to optical tomography. In *OSA Frontiers in Optics* (Optica Publishing Group, 2015); <https://doi.org/10.1364/LS.2015.LW31.1>
10. Gbur, G. & Wolf, E. Diffraction tomography without phase information. *Opt. Lett.* **27**, 1890–1892 (2002).
11. Jenkins, M. H. & Gaylord, T. K. Three-dimensional quantitative phase imaging via tomographic deconvolution phase microscopy. *Appl. Opt.* **54**, 9213–9227 (2015).
12. Tian, L. & Waller, L. 3D Intensity and phase imaging from light field measurements in an LED array microscope. *Optica* **2**, 104–111 (2015).
13. Chen, M., Tian, L. & Waller, L. 3D Differential phase contrast microscopy. *Biomed. Opt. Express* **7**, 3940–3950 (2016).
14. Ling, R., Tahir, W., Lin, H.-Y., Lee, H. & Tian, L. High-throughput intensity diffraction tomography with a computational microscope. *Biomed. Opt. Express* **9**, 2130–2141 (2018).
15. Li, J. et al. Three-dimensional tomographic microscopy technique with multi-frequency combination with partially coherent illuminations. *Biomed. Opt. Express* **9**, 2526–2542 (2018).
16. Wang, Z. et al. Spatial light interference microscopy (SLIM). *Opt. Express* **19**, 1016–1026 (2011).
17. Nguyen, T. H., Kandel, M. E., Rubessa, M., Wheeler, M. B. & Popescu, G. Gradient light interference microscopy for 3D imaging of unlabeled specimens. *Nat. Commun.* **8**, 210 (2017).
18. Chowdhury, S. et al. High-resolution 3D refractive index microscopy of multiple-scattering samples from intensity images. *Optica* **6**, 1211–1219 (2019).
19. Chen, M., Ren, D., Liu, H. Y., Chowdhury, S. & Waller, L. Multi-layer born multiple-scattering model for 3D phase microscopy. *Optica* **7**, 394–403 (2020).
20. Li, J. et al. High-speed in vitro intensity diffraction tomography. *Adv. Photon.* **1**, 1–13 (2019).
21. Kak, A. C. & Slaney, M. *Principles of Computerized Tomographic Imaging* (IEEE, 1988).
22. Venkatakrishnan, S. V., Bouman, C. A. & Wohlberg, B. Plug-and-play priors for model based reconstruction. In *IEEE Global Conference on Signal and Information Processing (GlobalSIP)* 945–948 (2013); <https://doi.org/10.1109/GlobalSIP.2013.6737048>
23. Sreehari, S. et al. Plug-and-play priors for bright field electron tomography and sparse interpolation. *IEEE Trans. Comput. Imag.* **2**, 408–423 (2016).
24. Chan, S. H., Wang, X. & Elgendy, O. A. Plug-and-play ADMM for image restoration: fixed-point convergence and applications. *IEEE Trans. Comput. Imag.* **3**, 84–98 (2017).
25. Ahmad, R. et al. Plug-and-play methods for magnetic resonance imaging: using denoisers for image recovery. *IEEE Signal Process. Mag.* **37**, 105–116 (2020).
26. Kang, E., Min, J. & Ye, J. C. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction. *Med. Phys.* **44**, e360–e375 (2017).
27. Jin, K. H., McCann, M. T., Froustey, E. & Unser, M. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. Image Process.* **26**, 4509–4522 (2017).
28. Zhu, B., Liu, J. Z., Cauley, S. F., Rosen, B. R. & Rosen, M. S. Image reconstruction by domain-transform manifold learning. *Nature* **555**, 487–492 (2018).
29. Aggarwal, H. K., Mani, M. P. & Jacob, M. MoDL: model-based deep learning architecture for inverse problems. *IEEE Trans. Med. Imag.* **38**, 394–405 (2019).
30. Sun, Y., Xia, Z. & Kamilov, U. S. Efficient and accurate inversion of multiple scattering with deep learning. *Opt. Express* **26**, 14678–14688 (2018).
31. Li, Y., Xue, Y. & Tian, L. Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media. *Optica* **5**, 1181–1190 (2018).
32. Zhang, Z. & Lin, Y. Data-driven seismic waveform inversion: A study on the robustness and generalization. *IEEE Trans. Geosci. Remote Sens.* **58**, 6900–6913 (2020).
33. Wang, G., Ye, J. C. & De Man, B. Deep learning for tomographic image reconstruction. *Nat. Mach. Intell.* **2**, 737–748 (2020).
34. Liang, D., Cheng, J., Ke, Z. & Ying, L. Deep magnetic resonance image reconstruction: Inverse problems meet neural networks. *IEEE Signal Process. Mag.* **37**, 141–151 (2020).
35. Adler, A., Araya-Polo, M. & Poggio, T. Deep learning for seismic inverse problems: toward the acceleration of geophysical analysis workflows. *IEEE Signal Process. Mag.* **38**, 89–119 (2021).
36. Matlock, A. & Tian, L. Physical model simulator-trained neural network for computational 3D phase imaging of multiple-scattering samples. Preprint at <https://arxiv.org/abs/2103.15795> (2021).
37. Sitzmann, V., Zollhöfer, M. & Wetzstein, G. Scene Representation Networks: Continuous 3D-structure-aware neural scene representations. In *Advances in Neural Information Processing Systems (NeurIPS)* 1121–1132 (2019); <https://doi.org/10.5555/3454287.3454388>
38. Sitzmann, V., Martel, J. N. P., Bergman, A. W., Lindell D. B. & Wetzstein, G. Implicit neural representations with periodic activation functions. In *Advances in Neural Information Processing Systems (NeurIPS)* 7462–7473 (2020); <https://doi.org/10.5555/3495724.3496350>
39. Hinton, G. How to represent part-whole hierarchies in a neural network. Preprint at <https://arxiv.org/abs/2102.12627> (2021).
40. Píala, M. & Clarck, R. TerminiNeRF: ray termination prediction for efficient neural rendering. Preprint at <https://arxiv.org/abs/2111.03643> (2021).
41. Sun, Y., Wu, Z., Xu, X., Wohlberg, B. & Kamilov, U. S. Scalable plug-and-play ADMM with convergence guarantees. *IEEE Trans. Comp. Imag.* **7**, 849–863 (2021).
42. Sun, Y., Liu, J., Wohlberg, B. & Kamilov, U. Async-RED: a provably convergent asynchronous block parallel stochastic method using deep denoising priors. In *International Conference on Learning Representations (ICLR)* (2021).
43. B., Mildenhall et al. NeRF: Representing scenes as neural radiance fields for view synthesis. In *The European Conference on Computer Vision (ECCV)* 405–421 (2020); https://doi.org/10.1007/978-3-030-58452-8_24
44. Martin-Brualla, R. et al. NeRF in the wild: Neural radiance fields for unconstrained photo collections. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 7206–7215 (2021); <https://doi.org/10.1109/CVPR46437.2021.00713>
45. Yu, A., Ye, V., Tancik, M., & Kanazawa, A. pixelNeRF: neural radiance fields from one or few images. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2021); <https://doi.org/10.1109/CVPR46437.2021.00455>
46. K., Park et al. Nerfies: deformable neural radiance fields. In *Proc. IEEE/CVF International Conference on Computer Vision (ICCV)* 5845–5854 (2021); <https://doi.org/10.1109/ICCV48922.2021.00581>
47. S., Peng et al. Neural body: Implicit neural representations with structured latent codes for novel view synthesis of dynamic humans. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2021); <https://doi.org/10.1109/CVPR46437.2021.00894>
48. Li, Z., Niklaus, S., Snavely, N. & Wang, O. Neural scene flow fields for space-time view synthesis of dynamic scenes. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 6494–6504 (2021); <https://doi.org/10.1109/CVPR46437.2021.00643>
49. P. P., Srinivasan et al. NeRV: neural reflectance and visibility fields for relighting and view synthesis. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2021); <https://doi.org/10.1109/CVPR46437.2021.00741>
50. Wizaradwongsa, S., Phongthawee, P., Yenphraphai, J. & Suwajanakorn, S. NeX: real-time view synthesis with neural basis expansion. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2021); <https://doi.org/10.1109/CVPR46437.2021.00843>
51. Reed, A. W. et al. Dynamic CT reconstruction from limited views with implicit neural representations and parametric motion fields. In *Proc. IEEE/CVF International Conference on Computer Vision (ICCV)* 2238–2248 (2021); <https://doi.org/10.1109/ICCV48922.2021.00226>
52. Matlock, A. & Tian, L. High-throughput, volumetric quantitative phase imaging with multiplexed intensity diffraction tomography. *Biomed. Opt. Express* **10**, 6432–6448 (2019).
53. Wu, Z. et al. SIMBA: Scalable inversion in optical tomography using deep denoising priors. *IEEE J. Sel. Topics Signal Process.* **14**, 1163–1175 (2020).
54. Wiesner, D., Svoboda, D., Maška, M. & Kozubek, M. CytoPacq: a web-interface for simulating multi-dimensional cell imaging. *Bioinformatics* **35**, 4531–4533 (2019).
55. Lim, J., Ayoub, A. B., Antoine, E. E. & Psaltis, D. High-fidelity optical diffraction tomography of multiple scattering samples. *Light Sci. Appl.* **8**, 82 (2019).
56. Zhu, J., Wang, H. & Tian, L. High-fidelity intensity diffraction tomography with a non-paraxial multiple-scattering model. *Opt. Express* **30**, 32808–32821 (2022).
57. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
58. Kamilov, U. S. et al. Learning approach to optical tomography. *Optica* **2**, 517–522 (2015).
59. Müller, T., Evans, A., Schied, C. & Keller, A. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph.* **41**, 102:1–102:15 (2022).
60. Sun, Y., Liu, J., Xie, M., Wohlberg, B. & Kamilov, U. S. CoIL: coordinate-based internal learning for tomographic imaging. *IEEE Trans. Comp. Imag.* **7**, 1400–1412 (2021).

61. Tancik, M. et al. Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in Neural Information Processing Systems (NeurIPS)* 7537–7547 (2020); <https://doi.org/10.5555/3495724.3496356>
62. Park, J. J., Florence, P., Straub, J., Newcombe, R. & Lovegrove, S. DeepSDF: learning continuous signed distance functions for shape representation. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 165–174 (2019); <https://doi.org/10.1109/CVPR.2019.00025>
63. Kingma, D. & Ba, J. Adam: a method for stochastic optimization. In *International Conference on Learning Representations (ICLR)* (ICLR, 2015).
64. Zhang, K., Zuo, W., Chen, Y., Meng, D. & Zhang, L. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. *IEEE Trans. Image Process.* **26**, 3142–3155 (2017).
65. Sun, Y. & Liu, R. *wustl-cig/DeCAF* (Zenodo, 2022); <https://doi.org/10.5281/zenodo.6941764>

Acknowledgements

This work was supported by the NSF award nos. CCF-1813910 and CCF-2043134 (to U.K.), and CCF-1813848 and EPMD-1846784 (to L.T.).

Author contributions

The project was conceived by Y.S., R.L. and U.S.K. The code of the model was implemented by R.L. and Y.S. The experiments were designed by R.L. and Y.S. The numerical results were collected by R.L. Data acquisition and preparation were

conducted by J.Z. and L.T. The manuscript was drafted by Y.S. with assistance from R.L. and J.Z. and revised by U.S.K. and L.T. All of the authors reviewed manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s42256-022-00530-3>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42256-022-00530-3>.

Correspondence and requests for materials should be addressed to Ulugbek S. Kamilov.

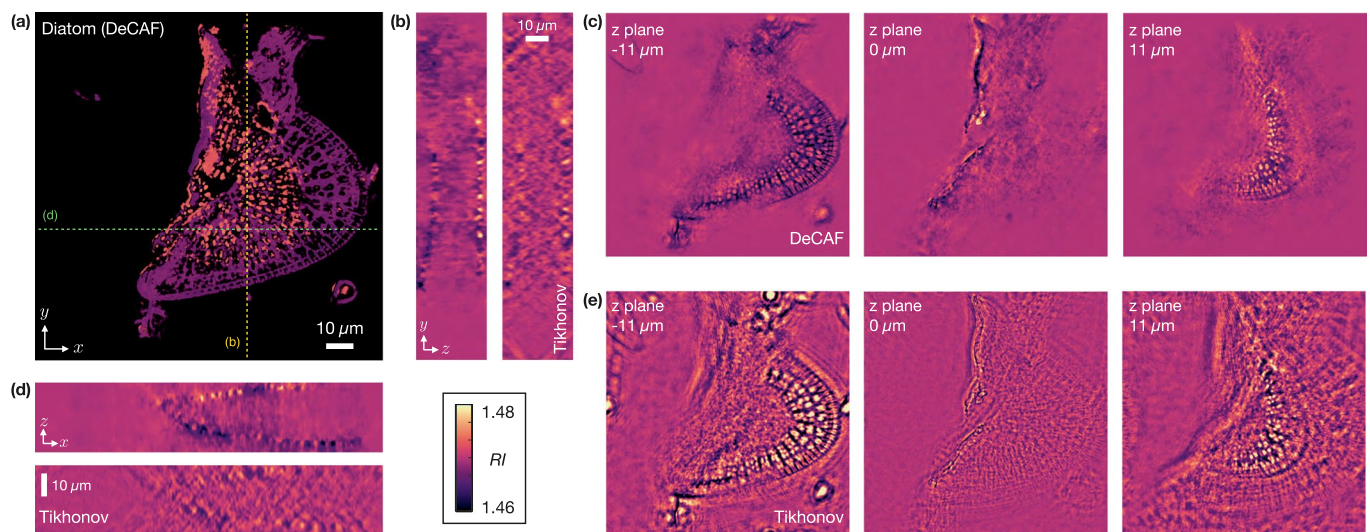
Peer review information *Nature Machine Intelligence* thanks Jan Funke, Jaejun Yoo and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

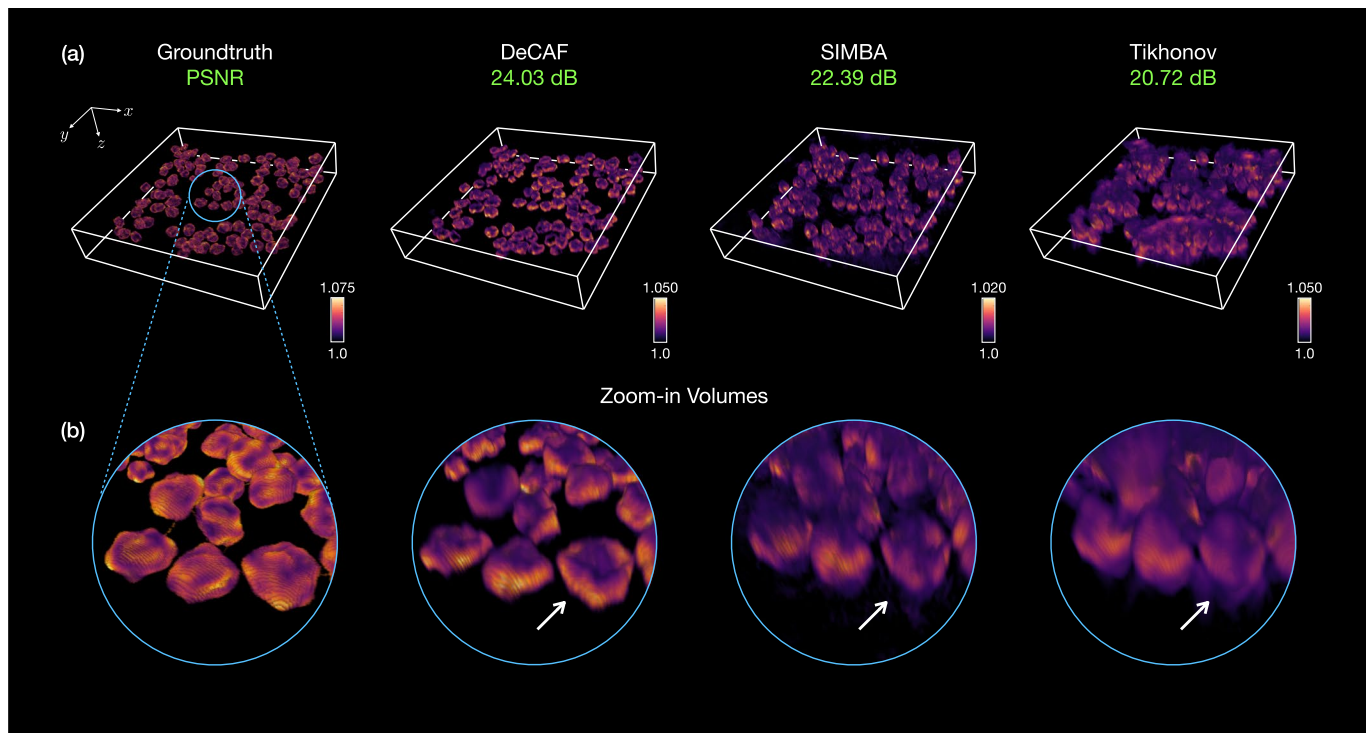
© The Author(s), under exclusive licence to Springer Nature Limited 2022



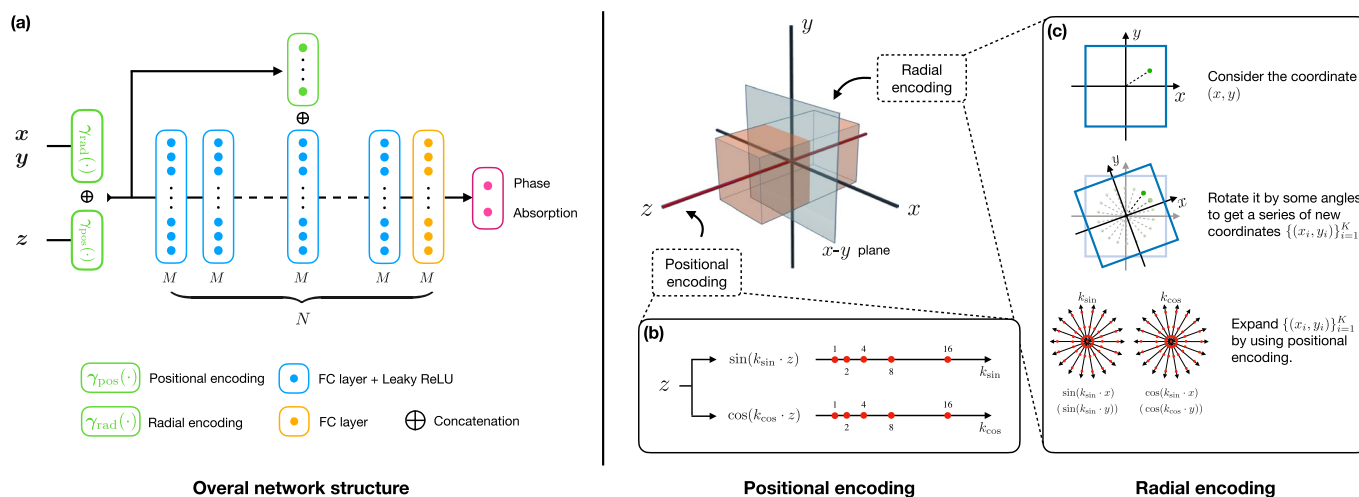
Extended Data Fig. 1 | Reconstruction of Diatom Algae acquired by mIDT. (a) 2D rendering obtained by accumulating all the z slices from DeCAF. Scale bar 10 μm . (b) & (d) Lateral views corresponding to the colored lines in (a). (c) & (e) Axial views at $z \in \{-11, 0, 11\}$ μm reconstructed by using DeCAF and Tikhonov, respectively. This figure illustrates the ability of DeCAF to reconstruct high-contrast RI maps for a relatively thin sample acquired by mIDT. Note how DeCAF successfully recovers the folding structure of the sample with two clear separate layers, which are barely recognizable in the Tikhonov reconstruction. Additional examples are shown in Supplementary Videos diatom-midt-decaf.mov and diatom-midt-tikhonov.mov.

Methods (Grid size)	Training	Inference		
	1000 × 1000 × 40	5000 × 5000 × 40	10000 × 10000 × 80	20000 × 20000 × 80
DeCAF	3MB	3MB	3MB	3MB
Tikhonov	-	3.73GB	29.84GB	119.36GB

Extended Data Fig. 2 | Quantitative Illustration of the scalability of DeCAF due to its off-the-grid feature using the *C. elegans* specimen as an example.
Note how the space required to store the reconstructed sample in DeCAF is independent of the reconstruction grid.



Extended Data Fig. 3 | Reconstruction of the 3D Granulocyte Phantom using DeCAF, SIMBA, and Tikhonov. (a) From left to right, 3D volumes correspond to Groundtruth, DeCAF, SIMBA, and Tikhonov, respectively. (b) Close-up views of the reconstructions at the location shown in (a). Note how DeCAF reconstructs sharper and better quality cell images compared to both SIMBA and Tikhonov.



Extended Data Fig. 4 | Visual illustration of the network structure and the encoding strategy used in DeCAF. (a) The overall structure of network $M\phi$. (b) Illustration of positional encoding for z coordinate. (c) Illustration of radial encoding for the coordinates in the (x, y) plane.

Hyperparameters		Algae	Diatom (aIDT)	Cells	Diatom (mIDT)	C.elegans
Modality type		IDT	aIDT	aIDT	mIDT	mIDT
Reconstruction Voxel Size		$1024 \times 1024 \times 40$	$1024 \times 1024 \times 40$	$700 \times 700 \times 52$	$800 \times 800 \times 60$	$1000 \times 1000 \times 40$
$ \theta $	Number of radial encoding lines distributed evenly around 360°	20	4	6	6	30
L_{xy}	Number of expanded frequencies for the x - y plane	7	9	8	9	9
L_z	Number of z expanded frequency	5	6	5	6	6
N	Number of layers in MLP	10	10	10	10	10
M	Number of Hidden neurons	208	208	208	208	208
	Skip layer	5	5	5	5	5
σ	Denoiser strength	1	1	1	1	1
α	Regularization strength of x - y noise reduction	1×10^{-5}	1×10^{-6}	3×10^{-7}	3×10^{-7}	3×10^{-6} (head) 1.5×10^{-6} (body)
β	Regularization strength of z -dimension continuity	1×10^{-6}	3×10^{-8}	1.5×10^{-7}	3×10^{-7}	3×10^{-6}

Extended Data Fig. 5 | List of algorithmic hyperparameters used by DeCAF for different biological samples. List of algorithmic hyperparameters used by DeCAF for different biological samples.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☒ ☐ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☒ ☐ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☒ ☐ A description of all covariates tested
- ☒ ☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☒ ☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☒ ☐ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection We implemented customized code to collect data.

Data analysis We implemented customized code using Matlab 2018a, Python 3.7.11, and Tensorflow 2.6 for pre-processing the raw intensity images and reconstruct the refractive index distribution of the sample. Also see README.md in the source code code.zip we provided.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The code used for reproducing the results in the manuscript is available at <https://github.com/wustl-cig/DeCAF>. We visualize the pre-processed raw intensity images of the relevant samples in Figure 1 and Figure 4.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We selected six biological samples.
Data exclusions	No data were excluded.
Replication	We ran the customized code for multiple times, and the samples can be successfully reconstructed in each trial.
Randomization	This is not relevant to the study because there is no biological analysis.
Blinding	This is irrelevant to the study because there is no biological analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	- Spirogyra Algae. This sample is a part of Fisher Science Education algae basic slide set S68786. - Diatom Algae (acquired by aIDT). This sample is a part of Fisher Science Education algae basic slide set S68786. - Diatom Algae (acquired by mIDT). This sample is a part of Fisher Science Education algae basic slide set S68786. - Human Buccal Epithelial Cells. This sample was swabbed from a researcher's buccal. The individual rinsed the mouth with clean water and then twirled a wooden swab against the inner cheek. - Caenorhabditis Elegans: Previously published dataset from the following paper: Alex Matlock and Lei Tian, High-throughput, volumetric quantitative phase imaging with multiplexed intensity diffraction tomography, Biomedical Optics Express Vol. 10, Issue 12, pp. 6432-6448 (2019).
Authentication	N/A
Mycoplasma contamination	N/A
Commonly misidentified lines (See ICLAC register)	N/A