bioRxiv preprint doi: https://doi.org/10.1101/631101; this version posted May 9, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

# Revealing architectural order with polarized light imaging and deep neural networks

**Syuan-Ming Guo[1,*], Anitha Priya Krishnan[1,*], Jenny Folkesson[1,*], Ivan Ivanov[1], Bryant Chhun[1], Nathan Cho[1], Manuel Leonetti[1], and Shalin B. Mehta[1,✉]**

[1]Chan Zuckerberg Biohub
[*]Equal contribution

## Abstract

Architectural order across spatial and temporal scales is a defining characteristic of living systems. Polarization of light enables label-free imaging of sub-resolution order in diverse biological systems without perturbing their assembly dynamics or causing phototoxicity. However, identification of specific structures seen in these images has remained challenging. We report synergistic use of polarized light microscopy, reconstruction of complementary optical properties, and deep neural networks to identify ordered structures. We recover birefringence, orientation, brightfield, and degree of polarization contrasts simultaneously by using Stokes formalism to model image formation. We report computationally efficient U-Net architectures that exploit information in complementary contrasts and predict specific structures with high accuracy. We illustrate the performance of our models by predicting ordered F-actin and condensed DNA in morphological diverse components of a kidney tissue. Our open-source python software for reconstruction of optical properties and training the neural networks is available on GitHub.

**Label-free Microscopy | Polarization | Computational Imaging | Deep Learning**

**Correspondence:** ***shalin.mehta @czbiohub.org***

## Introduction

The function of living systems emerges from dynamic interaction of components that give rise to ordered structures over spatial scales of nanometers to meters and temporal scales of milliseconds to years. Methods for imaging ordered arrangement of molecules within the context of organelles, of organelles within the context of cells, and of cells within the context of tissues promise new insights in the function of biological systems that have been elusive from study of the individual components at a single scale.

Polarization of light provides sensitivity to architectural order below the spatial resolution of a given microscope. Transmitted-light polarization microscopy has enabled analysis of intrinsic order in live biological systems. It has led to the discovery of the dynamic microtubule spindle (1). It has been used in in vitro fertilization (IVF) clinics to assess structural integrity of meiotic spindles of oocytes (2). It has been used for label-free imaging of white matter in brain tissue slices (3, 4), and recently for imaging activity dependent structural changes in acute brain slices (5).

Key challenges in widespread adoption of label-free polarized light imaging are sensitive detection of biological structures in the presence of background and identification of specific structures that are detected. Fluorescence polarization microscopy lends itself to automated analysis of dynamic order (6) with molecular specificity. But a fluorescent reporter often compromises an ordered assembly, such as actin network (7), and limits the number of structures that can be analyzed at the same time, especially in live cells. Synergistic combination of label-free polarization-diverse imaging, accurate reconstruction algorithms, and deep neural networks can resolve these bottlenecks and reveal emergence of order among interacting structures in diverse biological systems.

**Related work.** We have previously employed liquid-crystal based transmitted polarized light microscopy (LC-PolScope) for sensitive detection of the specimen's birefringence and slow axis (8, 9) as well as diattenuation and the axis of maximum transmission (10). The reconstruction and background correction algorithms in earlier papers are based on Jones calculus that assumes coherent and fully polarized illumination, i.e. illumination with a plane wave (11, Ch.10). But LEDs or lamps used for these experiments actually lead to partially polarized illumination, which is not accounted for by Jones calculus. Fluorescence polarization microscopy also gives rise to partially polarized emission, since independent emission events imaged through the detection numerical aperture are mutually incoherent. Stokes vector representation of light and Mueller matrix representation of the optical components elegantly capture the full state of polarization of light, including partial polarization. We previously developed Stokes representation of fluorescence polarization for simultaneous recovery of concentration, alignment, and orientation of fluorophores imaged with instantaneous fluorescence polarization microscope (6), but partial polarization in transmitted light microscopy is yet to be exploited to retrieve information about the specimen.

Polarization-sensitive imaging has also been performed in reflection mode, most commonly with polarization sensitive optical coherence tomography (PS-OCT). PS-OCT has been used to measure round-trip birefringence and diattenuaton of diverse tissues, e.g., of brain tissue (12). But determination of the material axes in the reflection mode is confounded by the fact that light passes through the specimen in two directions. The reconstruction and background correction algorithms in PS-OCT primarily rely on Jones cal-

bioRxiv preprint doi: https://doi.org/10.1101/631101; this version posted May 9, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

culus, since OCT is a coherent interferometer and intensity recorded in individual speckle is fully polarized (13). However, PS-OCT practitioners employ degree of polarization uniformity (13) over several speckles to analyze depolarization due to multiple scattering.

Deep convolutional neural networks have recently enabled identification of structures that can be perceived in diverse label-free images, opening new opportunities to advance the state-of-the-art of biological imaging. These opportunities include increasing the imaging throughput via computationally multiplexed analysis of biological structures, comparative analysis of architecture of primary tissue that are difficult to label with consistency, and analysis of biological processes in non-model organisms that are difficulty to label genetically.

Recent papers report combination of brightfield or differential interference contrast (DIC) imaging with U-Net (14) for label-free identification of multiple organelles in cells; combination of phase-contrast, DIC and adaptation of Google's inception model for in silico labeling of nuclei, cell types, and cell state (15); combination of quantitative phase (16) and auto-fluorescence(17) with generative adverserial networks for prediction of histopathology images; combination of diffraction tomography with U-Net for segmentation of immunological synapse (18). In addition, identification of carcinoma region in colon tissue from Raman scattering images using a random forest classifier (19) has been reported.

However, rapid analysis of ordered biological structures remains challenging, because the aforementioned label-free imaging approaches (absorption, phase, and auto-fluorescence) are not as sensitive as polarized light microscopy in detecting ordered structures, and any machine learning algorithm can only learn the structural information that is present in the input data.

**Contributions.** We develop more accurate model of image formation in polarized light microscopy and corresponding algorithm for reconstruction of complementary optical properties using Stokes formalism. Our approach recovers brightfield, birefringence, orientation of dense axis, and degree of polarization images simultaneously. Casting the image formation and reconstruction in Stokes formalism also provides an elegant representation of the microscope in terms of an instrument matrix, which enables robust calibration and facilitates design of new polarization-resolved imaging algorithms.

We report a computationally efficient network architecture for 3D translation that combines information from the depth of field of the microscope and complementary label-free contrasts to predict fluorescent volumes with state-of-the-art accuracy, in contrast to some of the previous work that demonstrated 2D translation (15–17). Our 2.5D architecture achieves the same or better accuracy as the 3D translation architecture reported in (20) while taking significantly less time to train and without having to sacrifice in plane resolution by downsampling. In comparison to Google's 2D translation model (15), our 3D translation model uses significantly fewer parameters and predicts much larger dynamic range of gray levels, albeit for one translation task. We systematically evaluated how the contrasts and dimensions of the input affect the prediction accuracy and computational cost. We find that higher prediction accuracy is archived by combining the multiple label-free contrasts. We demonstrate prediction of fluorescence images of tissue, while previous work has reported prediction of fluorescence images of cultured cells or brightfield images of histochemically-stained tissue (15–17, 20). Image translation results reported in (15–17) were limited to 2D structures, likely because they use even more complex architectures than 3D U-Net model reported in (20).

In the results section, we first describe retrieval of complementary optical contrasts from polarization-resolved images and then discuss development of computationally-efficient image translation models.

## Results and Discussion

**Reconstructing complementary contrasts from polarization-diverse images.** Architectural order leads to variations in concentration and alignment of bio-molecules, which induce variations in optical path length, retardance[1], and scattering. We implemented automated polarized-light imaging protocol using LC-PolScope (Fig. 1A, methods) and developed two-step algorithm (Fig. 1B, methods) for simultaneous recovery of brightfield, retardance, slow axis, and degree of polarization contrasts. We also implemented background correction methods for sensitive imaging of small variations in the contrast due to the specimen (methods). Slight strain or misalignment in the optical components or the sample chamber can lead to background that masks contrast due to the specimen. The background typically varies slowly across the field of view and can introduce spurious correlations in the measurement. With our improved reconstruction approach, we computed background-free brightfield, retardance, slow axis, and degree of polarization images from polarization-resolved intensities. The python code for reconstruction is available at `https://github.com/czbiohub/reconstructorder`.

Fig. 1C shows background-corrected images of a kidney tissue slice, U2OS cells (bone cancer cell line), and mouse brain slice in above four contrasts. The brightfield (BF) image reports dense structures that appear in positive contrast on one side of the focus and in negative contrast on the other. These intensity variations arise via transport of intensity relationship (22). The transport of intensity effect is noticeable in through focus brightfield images of kidney tissue (Supplementary Movie 1) and in through focus brightfield images of condensed chromosomes in cells (Supplementary Movie 2). The retardance image is proportional to liquid crystalline or orientational order among molecules. The orientation image reports the dense axis of the retardance. In kidney tissue, the retardance image highlights convolutions within glomelurus, capillary that is cut transversely, and tubules, among other

[1] optical path length and retardance are refractive index and birefringence integrated over a coherently illuminated volume within the specimen(21, Fig. 3)

bioRxiv preprint doi: https://doi.org/10.1101/631101; this version posted May 9, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

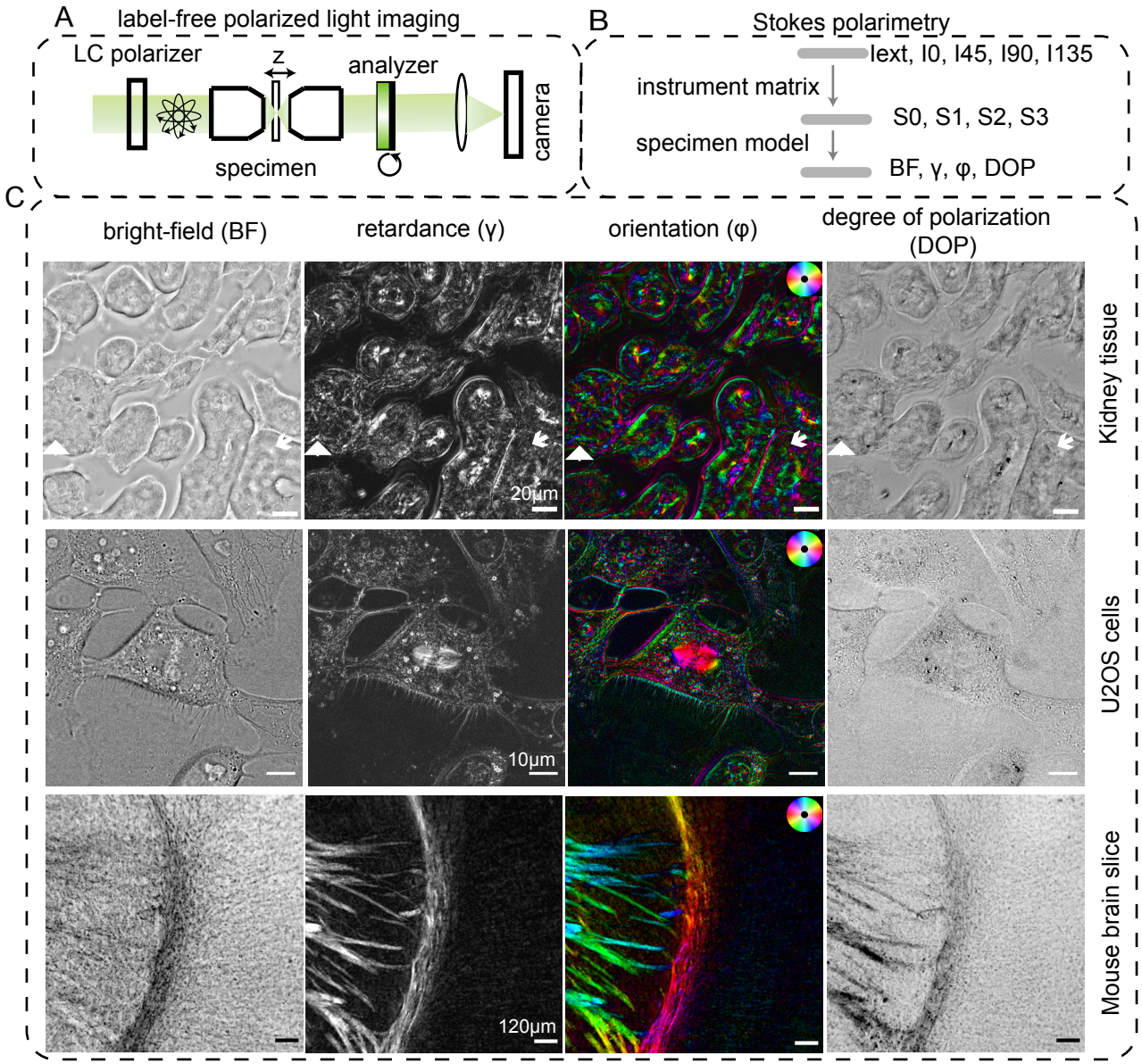


**Fig. 1. Retrieval of complementary contrasts from polarization-diverse images: (A)** Tunable liquid crystal (LC) is used to acquire polarization-diverse images in transmission. **(B)** We retrieve Stokes parameters of light at the image plane from polarization-resolved images and retrieve specimen properties from Stokes parameters assuming that the specimen is transparent. **(C)** Representative images of tubules in kidney tissue, organelles in U2OS cells, and axon tracts in mouse brain slice. The brightfield image visualizes structures with distinct optical path length, retardance and orientation images visualize structures with liquid crystalline order, and degree of polarization image visualizes structures that multiply scatter the light. The orientation images are color coded according to the color-wheel in top-right - the color represents the orientation and the brightness represents the retardance. The markers shown on kidney tissue images identify structures discussed in Fig. 2

components of the tissue. The nuclei appear in negative contrast in the retardance image because condensation of DNA leads to isotropic structure. In dividing U2OS cell (Supplementary Movie 2), the orientation image clearly shows dynamics of membrane boundaries, microtubule spindle, and lipid droplets. Lipid droplets in the U2OS cells and tubules in the kidney tissue significantly change the degree of polarization (DOP) as they multiply scatter the light. In brain tissue, the retardance and orientation images distinctly report axon tracts due to birefringence of myelinated axons (23). Supplementary Image 1C shows tiled image of the whole brain slice, in which not just the white matter tracts, but also changes in the orientation of axons within different cortical layers is visible. The DOP image of brain tissue primarily highlights large axon tracts that can multiply scatter light (Fig. 1C and Supplementary Image 1D).

It is worth clarifying the difference between retardance and degree of polarization measurements. The retardance variations arise from single scattering events within the sample that alter the polarization, but do not reduce the coherence. The degree of polarization on the other hand reports multiple scattering events that reduce the polarization of light and diattenuation that polarizes the light further. In the future, we are excited to develop models that account for diffraction and scattering effects in polarized light microscopy and develop methods that provide more quantitative estimates of the

bioRxiv preprint doi: https://doi.org/10.1101/631101; this version posted May 9, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

specimen properties.

Our approach enables visualization of large class of structures from their density, order, and scattering with diffraction-limited resolution and high sensitivity for the first time to our knowledge. In the next sections, we discuss how these complementary signatures of the specimen enable accurate prediction of the shape and expression of different types of structures.

**Optimization of image translation model.** Iterative optimization of optical contrast, architecture of the deep neural network, and the training process is key to successful analysis of structures of interest. We reasoned that complementary signatures acquired by our microscope should enable learning of structures of heterogeneous shapes in tissue. In order to develop our deep neural network models, we trained models to predict F-actin and DNA distribution within kidney tissue.

We adapted the widely successful residual U-Net model (24, 25) to translate structures represented in label-free images into fluorescence images. Prior work (20) on translation of brightfield images to fluorescence images has shown that 2D translation models result in discontinuous predictions along the z-axis as compared to 3D translation models. However, 3D U-Net model requires sufficiently large Z dimension as the input is isotropically downsampled in the encoding path of the U-Net. However, typical microscopy images acquired with Nyquist sampling have anisotropic dimensions due to anisotropic resolution in the microscope. Therefore, use of 3D translation models requires that either the data is downsampled in $XY$ or upsampled in $Z$. Downsampling in $XY$ leads to loss of resolution acquired by the microscope and maintaining the $XY$ resolution requires upsampling in $Z$, which increases data size without adding information. The increased data size makes training 3D translation model more computationally expensive. In our experiments, training 3D models using 100 training volumes required 3.5 days to converge with a cutting edge.

We sought to reduce the computational cost, while maintaining the optical resolution of the data and high accuracy of prediction. Structures with different physical properties, e.g., ordered vs condensed, give rise to different label-free signatures. We therefore evaluated the prediction accuracy as a function of the label-free contrast and dimensions of the input for an ordered structure (F-actin) and for a condensed structure (DNA) in kidney tissue.

We experimented with three types of U-Net models to predict fluorescence volumes: slice→slice (2D in short) models predicted 2D fluorescence slices from corresponding 2D label-free slices, stack→slice (2.5D in short) models predicted 2D fluorescence slices from a few (3,5, or 7) neighboring label-free slices, and stack→stack (3D in short) models predicted fluorescent volume from label-free volume. See methods for the description of the network architecture and training process. To evaluate the performance of the models and the effect of training parameters, we computed Pearson correlation coefficient and structural similarity index (26) between predicted fluorescent volumes and ground truth fluorescent volumes in the test set, which were never seen by the model during training (methods). table 1 reports these test metrics for models that predict F-actin volumes and table 2 reports test metrics for models that predict DNA volumes. The 2D models required 6-8 hrs to train on a GPU with 12GB RAM, 2.5D models required 24 hrs to train on GPU with 32GB RAM, and 3D models required 84 hrs to train on GPU with 32GB RAM. In the next two sections, we report main findings from our model optimization effort.

The python code for training our variants of image translation models is available at `https://github.com/czbiohub/microDL`.

**Predicting structures from multiple label-free contrasts improves accuracy.** We took advantage of the computational efficiency of 2D models to explore the effect of the label-free inputs, used independently and jointly, on the prediction accuracy of fluorescent structures (Fig. 2A). Ground-truth (Fig. 2B) and predicted (Fig. 2C) distributions of F-actin and DNA from a representative field of view containing glomerulus and convoluted tubules illustrate model performance. This field of view was chosen from the test set that was not seen by the model during the training or validation. Label-free inputs used for prediction are shown in Fig. 1C.

The 2D model accurately predicts small tubules bound by F-actin (arrowheads in Fig. 2B and C) from the retardance ($\gamma$) channel, but not from brightfield (BF) channel. On the other hand, closely-spaced nuclei within a glomerulus (triangleheads in Fig. 2B and C) are well-resolved in the prediction from BF channel, but not in the prediction from $\gamma$ channel. When $\gamma$, $\phi$, and BF images are jointly used as inputs, both structures are predicted with high fidelity. These variations in the prediction performance arise from the sensitivity of the retardance image to ordered F-actin and the sensitivity of the brightfield image to dense DNA. Consistent with models reported by Ounkomol *et. al.*(20), our brightfield model predicts large-scale F-actin structures, but not the small scale structure. Our models that use retardance and orientation as input, however, are able to predict fine F-actin structures - compare F-actin stress fibers in the last frame of z-stacks shown in supplementary movie 3A and supplementary movie 4A.

These observations from a representative field of view generalize to entire test set as illustrated by the distribution of Pearson correlations between slices of ground-truth and predictions ($r_{xy}$) in Fig. 2C. Correlation between ground truth and prediction over the test data has the highest median and the narrowest distribution when all three channels are used as input. Comparing median values of the correlation and structural similarity index for several models in table 1 and table 2 ($r_{xy}$ and $\text{SSIM}_{xy}$ columns) further shows that the prediction accuracy of 2D models robustly increased as the most informative channels were used jointly as input.

**Using label-free images over the depth of field improves prediction of 3D structure.** To evaluate our models' ability to learn complex three-dimensional structures, we take a closer look at glomeruli in the kidney tissue. Glomeruli

bioRxiv preprint doi: https://doi.org/10.1101/631101; this version posted May 9, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

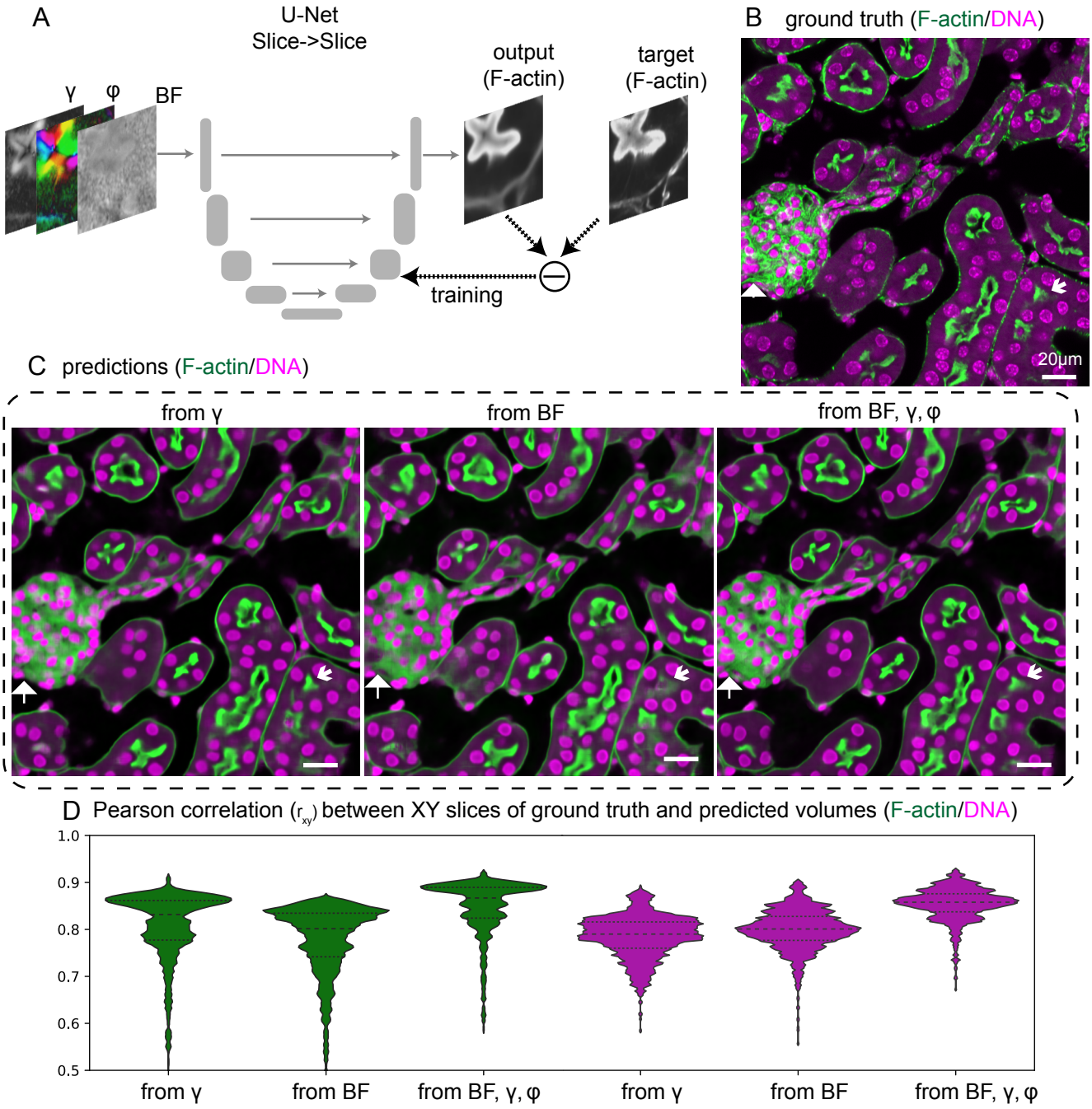


**Fig. 2. Prediction accuracy as a function of label-free contrast: (A)** We trained a slice→slice U-Net model to predict a fluorescent target from multiple label-free images. **(B)** Fluorescence ground truth for a field of view from the test set. Phalloidin-labeled F-actin in shown green and DAPI labeled DNA is shown in magenta. **(C)** F-actin and DNA distribution predicted with U-Net models trained on brightfield (BF), retardance ($\gamma$), and combination of BF, $\gamma$, and slow axis $\phi$ channels. **(D)** Violin plots show distribution of Perason correlation between the predicted XY slices and ground truth XY slices ($r_{xy}$) over the entire test dataset. The horizontal dashed lines in the violin plots indicate $25^{\text{th}}$ quartile, median, and $75^{\text{th}}$ quartile of the Pearson correlation. The triangle points to a collection of nuclei and the arrow points to F-actin structure seen in ground truth image, but that are predicted or not as a function of input channel used.

bioRxiv preprint doi: https://doi.org/10.1101/631101; this version posted May 9, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

| Translation model | Input(s) | Loss | $r_{xy}$ | $r_{xz}$ | $r_{xyz}$ | $SSIM_{xy}$ | $SSIM_{xz}$ | $SSIM_{xyz}$ |
|---|---|---|---|---|---|---|---|---|
| Slice->Slice (2D) | BF | L2 (Otsu) | 0.80 | 0.77 | 0.81 | 0.65 | 0.57 | 0.66 |
| | $\gamma$ | L2 (Otsu) | 0.83 | 0.80 | 0.83 | 0.69 | 0.61 | 0.69 |
| | $\phi_x$ | L2 (Otsu) | 0.82 | 0.79 | 0.83 | 0.68 | 0.60 | 0.69 |
| | $\phi_y$ | L2 (Otsu) | 0.82 | 0.80 | 0.83 | 0.69 | 0.61 | 0.70 |
| | DOP | L2 (Otsu) | 0.74 | 0.70 | 0.75 | 0.57 | 0.49 | 0.58 |
| | BF, $\gamma$ | L2 (Otsu) | 0.86 | 0.83 | 0.85 | 0.73 | 0.66 | 0.74 |
| | $\gamma$, $\phi_x$, $\phi_y$ | L2 (Otsu) | 0.86 | 0.83 | 0.85 | 0.74 | 0.67 | 0.74 |
| | BF, $\gamma$, $\phi_x$, $\phi_y$ | L2 (Otsu) | 0.87 | 0.84 | 0.86 | 0.75 | 0.69 | 0.76 |
| Stack-> Slice (2.5D, $z = 3$) | $\gamma$ | L1 (Rosin) | 0.86 | 0.83 | 0.86 | 0.74 | 0.68 | 0.75 |
| Stack-> Slice (2.5D, $z = 5$) | $\gamma$ | L1 (Rosin) | 0.89 | 0.85 | 0.87 | 0.75 | 0.69 | 0.76 |
| | BF, $\gamma$ | L1 (Rosin) | 0.90 | 0.88 | 0.90 | 0.79 | 0.73 | 0.80 |
| | $\gamma$, $\phi_x$, $\phi_y$ | L1 (Rosin) | 0.89 | 0.87 | 0.89 | 0.79 | 0.73 | 0.79 |
| | BF, $\gamma$, $\phi_x$, $\phi_y$ | L1 (Rosin) | 0.89 | 0.88 | 0.90 | 0.79 | 0.73 | 0.80 |
| Stack-> Slice (2.5D, $z = 7$) | $\gamma$ | L1 (Rosin) | 0.87 | 0.85 | 0.87 | 0.76 | 0.70 | 0.76 |
| Stack-> Stack (3D, $z = 96$) | $\gamma$ | L1 (Rosin) | 0.84 | 0.82 | 0.85 | 0.73 | 0.70 | 0.77 |

**Table 1. Accuracy of prediction of F-actin in kidney tissue:** Above table lists median values of the Pearson correlation $(r)$ and structural similarity index (SSIM) between prediction and ground truth volumes. We evaluated combinations of brightfield (BF), retardance ($\gamma$), orientation x ($\phi_x$), orientation y ($\phi_y$), and degree of polarization (DOP) as input. Slice→slice (2D) translation models were trained using L2 (MSE) loss and stack→slice (2.5D) and stack→stack (3D) models were trained using L1 (MAE) loss. We segmented target images with Rosin and Otsu threshold to discard tiles that mostly contained background pixels. To dissect the differences in prediction accuracy along and perpendicular to the focal plane, we computed test metrics separately over XY slices ($r_{xy}$,$SSIM_{xy}$) and XZ slices ($r_{xz}$,$SSIM_{xz}$) of the test volumes, as well as over individual test volumes ($r_{xyz}$, $SSIM_{xyz}$).

| Translation model | Input(s) | Loss (mask) | $r_{xy}$ | $r_{xz}$ | $r_{xyz}$ | $SSIM_{xy}$ | $SSIM_{xz}$ | $SSIM_{xyz}$ |
|---|---|---|---|---|---|---|---|---|
| Slice->Slice (2D) | BF | L2 (Otsu) | 0.80 | 0.79 | 0.78 | 0.59 | 0.50 | 0.60 |
| | $\gamma$ | L2 (Otsu) | 0.79 | 0.76 | 0.76 | 0.59 | 0.46 | 0.57 |
| | $\phi_x$ | L2 (Otsu) | 0.78 | 0.76 | 0.76 | 0.60 | 0.46 | 0.56 |
| | $\phi_y$ | L2 (Otsu) | 0.80 | 0.78 | 0.78 | 0.61 | 0.48 | 0.58 |
| | DOP | L2 (Otsu) | 0.26 | 0.24 | 0.26 | 0.34 | 0.29 | 0.35 |
| | BF, $\gamma$ | L2 (Otsu) | 0.84 | 0.82 | 0.82 | 0.62 | 0.54 | 0.64 |
| | $\gamma$, $\phi_x$, $\phi_y$ | L2 (Otsu) | 0.83 | 0.81 | 0.81 | 0.62 | 0.52 | 0.62 |
| | BF, $\gamma$, $\phi_x$, $\phi_y$ | L2 (Otsu) | 0.86 | 0.84 | 0.84 | 0.63 | 0.55 | 0.65 |
| Stack-> Slice (2.5D, $z = 5$) | $\gamma$ | L1 (Rosin) | 0.85 | 0.85 | 0.84 | 0.61 | 0.54 | 0.63 |
| | BF, $\gamma$ | L1 (Rosin) | 0.91 | 0.90 | 0.90 | 0.71 | 0.63 | 0.74 |
| | $\gamma$, $\phi_x$, $\phi_y$ | L1 (Rosin) | 0.89 | 0.88 | 0.88 | 0.67 | 0.60 | 0.70 |
| | BF, $\gamma$, $\phi_x$, $\phi_y$ | L1 (Rosin) | 0.89 | 0.89 | 0.89 | 0.66 | 0.59 | 0.69 |

**Table 2. Accuracy of prediction of DNA in kidney tissue:** Above table lists median values of the Pearson correlation $(r)$ and structural similarity index (SSIM) between prediction and ground truth volumes. See table 1 for description of other columns.

are key components of kidney tissue that perform filtration. They are complex multi-cellular structures the size of a single cultured cell (27). Fig. 3A and Fig. 3B show XY and XZ slices through retardance volume and F-actin volume of the same glomerulus shown in Fig. 2B from the test set, while Fig. 3C shows XY and XZ slices through the F-actin volumes predicted using 2D, 2.5D, and 3D models trained on retardance as the input. The predictions with 2D models show discontinuity artifacts in the structure along the depth. These artifacts can also be observed from comparison of z-stacks of F-actin (supplementary movie 3A) and z-stack of 2D prediction (supplementary movie 4A). The 2.5D model predicts smoother structures along depth dimension and improves the fidelity of F-actin prediction in XY plane (Fig. 3B).

The 3D model further improves the continuity of prediction along the depth (Fig. 3C and supplementary movie 4C), however the gain in accuracy is not as significant as the transition from 2D to 2.5D model. As evaluated with distribution of Pearson correlations along XY and XZ slice in Fig. 3D, the 2.5D model performs almost as well as 3D model, although it is faster and more memory efficient. When evaluated with the median values of Pearson correlation over whole test set shown in table 1 2.5D models perform consistently better than 3D model along XY, XZ, and XYZ dimensions. However, when evaluated with SSIM, the 3D model performs slightly better along XZ and XYZ dimensions, whereas 2.5D model performs better along XY dimensions. Notably, 3D model took $3\times$ longer to train than 2.5D model due to the larger input Z dimension required by 3D model that significantly increases its memory footprint.

We reasoned that using complementary label-free contrasts can boost the performance of 2.5D models to match the performance of 3D single-channel models without significantly increasing the computation cost. The Pearson correlation and SSIM reported in table 1 for prediction of F-actin from 2.5D multi-channel models are consistently higher than

bioRxiv preprint doi: https://doi.org/10.1101/631101; this version posted May 9, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

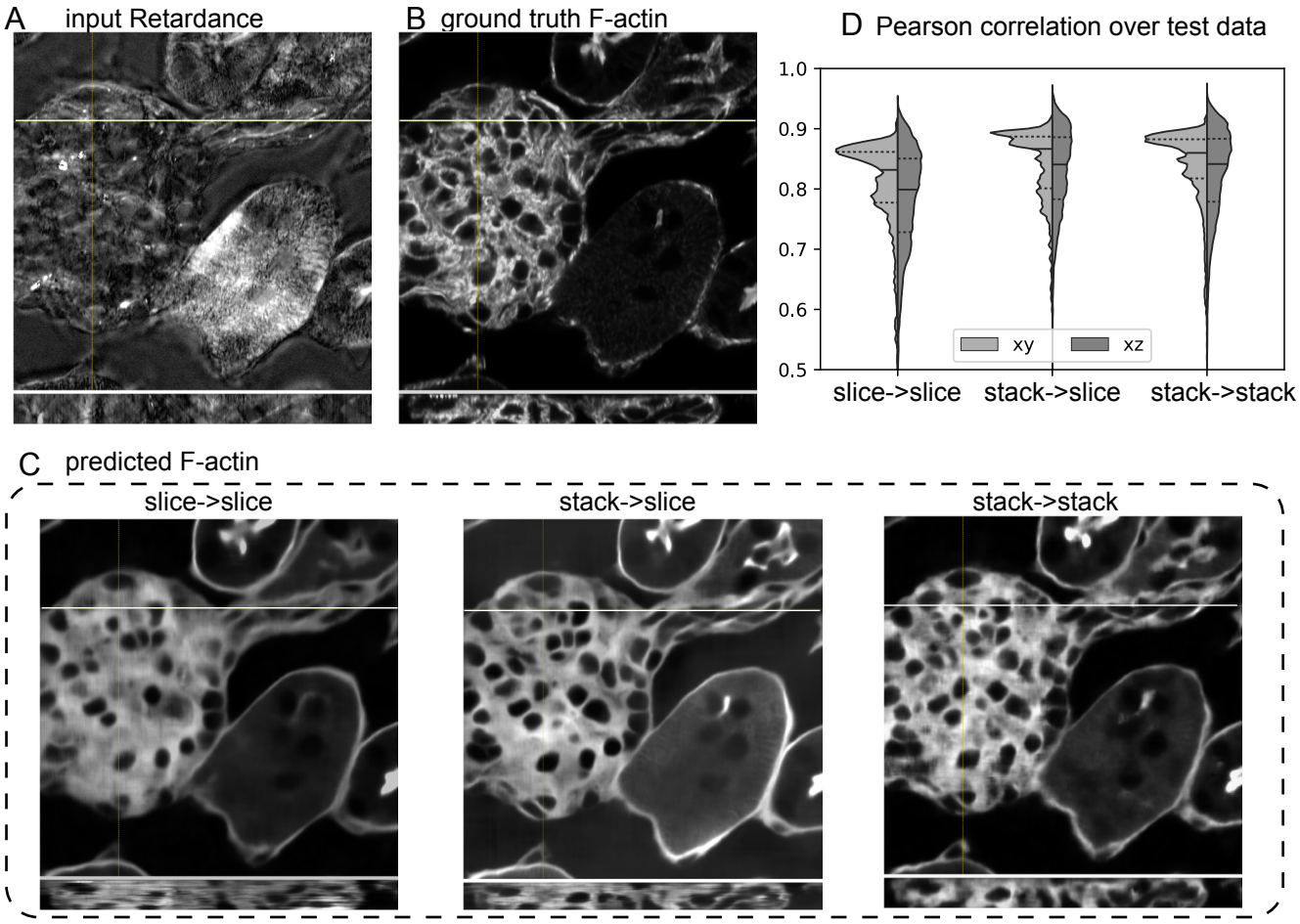


**Fig. 3. prediction accuracy as a function of U-Net architecture:** XY (top) and XZ (bottom) slices through a volume in test data shows (A) retardance, (B) ground truth fluorescence of F-actin and DNA, and (C) predictions made with different U-Net architectures. (D) shows violin plots of Pearson correlation distributions in the XY plane and XZ plane across the test dataset.

3D single-channel models along XY, XZ, and XYZ dimensions. The prediction accuracy for fine structures such as F-actin stress fiber also improves when complementary contrasts are used as input, as seen from last frames of z-stacks shown in supplementary movies 4A and 4B. We compare ground-truth distribution of F-actin and DNA (Fig. 4A) with predicted distributions (Fig. 4B) over the test field of view containing glomerulus. 2.5D model trained on $\gamma, \phi$, and BF is able to predict lumen locations (arrow head in Fig. 4) better than 2.5D model trained on $\gamma$ alone. This can be explained from the fact that lumen is an isotropic structure of lower density that appears similar to the background in retardance image.

We also reasoned that 2.5D multi-channel model can be robustly trained to predict diverse set of structures. All models trained just on retardance do not predict some of the gaps in the F-actin distribution, which may be nuclei or capillaries. As seen from metrics in table 1 and table 2, the prediction of both F-actin and DNA improves when 2.5D models are trained with complementary inputs. Note that the 2.5D model trained on $\gamma$ alone can miss nuclei (triangle head), which 2.5D model trained on $\gamma, \phi$, and BF is able to predict. Pearson correlation between ground truth and predictions also has higher medians and narrower distributions for 2.5D multi-channel models for both F-actin and DNA (Fig. 4C). Comparison z-stack of label-free inputs (supplementary movie 1), ground truth fluorescence (supplementary movie 3) and prediction from 2.5D multi-channel model (supplementary movies 4 and 5) further confirm our observations.

In conclusion, above results show that 2.5D multi-channel U-Net allows us to learn ordered and condensed structures with higher accuracy than reported before from complementary label-free contrasts. Our optimal network architecture can be applied to Nyquist sampled microscopy images with anisotropic or isotropic voxels.

## Methods

**Model of image formation.** We describe dependence of the polarization resolved images on the specimen properties using Stokes formalism (28, Ch.15). Based on this representation, we implement a two-step reconstruction of specimen properties from polarization-resolved intensities. First, we retrieve a background-corrected Stokes vector image of the specimen from the recorded images and an instrument matrix. Second, we convert the Stokes vector image into bright-field, retardance, slow axis, and degree of polarization images assuming that the specimen is transparent. The assumption of

bioRxiv preprint doi: https://doi.org/10.1101/631101; this version posted May 9, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

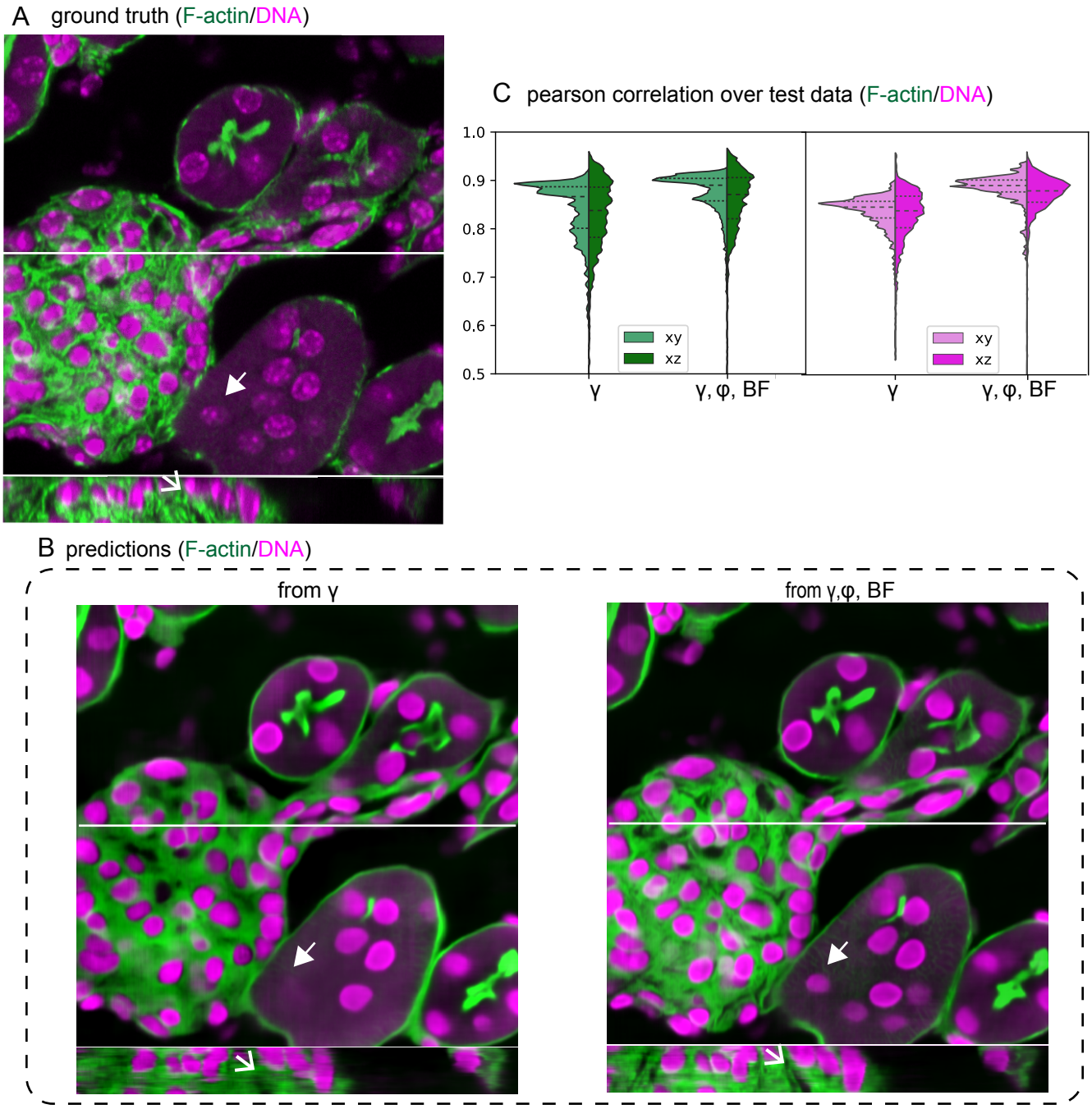


**Fig. 4. accurate prediction with multicontrast 2.5D models:** (A) Overlay of F-actin (green) and DNA (magenta) fluorescence for the volume shown in Fig. 3. (B) Predictions of F-actin and DNA from retardance along and from retardance, orientation, and brightfield images. (C) Violin plots of Pearson correlation distributions between XY slices and XZ slices of ground truth and predicted volumes across the test set. The triangle points to a nucleus seen in prediction from multiple channels, but not from single channel. The arrow points to a gap in actin distribution seen in prediction from multiple channels, but not from single channel.

transparency is generally valid for the structures we are interested in, but does not necessarily hold when the specimen exhibits significant absorption or diattenuation. To ensure that the inverse computation is robust, we need to make judicious decisions about the light path, calibration procedure, and background estimation. A key advantage of Stokes instrument matrix approach is that it easily generalizes to other polarization diverse imaging methods. A polarized light microscope (4-frame, 5-frame, instant, sequential) is represented directly by a calibrated instrument matrix.

For sensitive detection of birefringence, it is advantageous to illuminate the specimen with ellipitically polarized light and image with circular state of opposite handedness (9). For experiments reported in this paper, we acquired data by illuminating the specimen sequentially with left-handed circular and elliptical states and analyzed the light with right-handed circular state. However, for the sake of brevity, the following derivation assumes that the specimen is illuminated with right-handed circular state and analyzed with detectors sensitive to left-handed circular and elliptical

bioRxiv preprint doi: https://doi.org/10.1101/631101; this version posted May 9, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

states. Both of these systems are analogous in theory, but the later acquisition scheme with multiple detection states can be implemented both sequentially and parallely (6).

***Forward model: Specimen properties → Stokes vector.*** The Stokes vector (28, Ch. 15) of right circularly polarized illumination is given by,

$$S_{RCP} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{1}$$

We assume that the specimen is transparent, and therefore, modeled by net retardance $\gamma$, orientation of the slow axis $\phi$, transmission $t$, and depolarization $p$. The Mueller matrix of the specimen is then given by,

$$M_{specimen} = \begin{bmatrix} t & 0 & 0 & 0 \\ 0 & tp & 0 & 0 \\ 0 & 0 & tp & 0 \\ 0 & 0 & 0 & tp \end{bmatrix} M_r \tag{2}$$

where, $M_r$ is the Mueller matrix of a linear retarder that can be found in (28, Ch.14).

The Stokes vector of light after it has interacted with the specimen is given by $M_{specimen} S_{RCP}$, which is

$$\mathbf{S} = \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} t \\ tp \sin 2\phi \sin\gamma \\ -tp\cos 2\phi \sin\gamma \\ tp\cos\gamma \end{bmatrix} \tag{3}$$

The aim of the measurement now is to accurately measure the Stokes vector of light at each point in the image plane of the microscope by analyzing it with mutually independent polarization states. Once the Stokes vector map has been acquired with high accuracy, the specimen properties can be retrieved from above set of equations.

***Forward model: Stokes vector → intensities.*** Any polarization-resolved measurement scheme (in imaging format or otherwise) can be characterized by an 'instrument matrix' $\mathbf{A}$ that transforms Stokes vector of light $\mathbf{S}$ to the measured intensities $\mathbf{I}$. Thus, we express 5 polarization images detected with left-handed circular and elliptical states ($I_{LCP}$, $I_0$, $I_{45}$, $I_{90}$, $I_{135}$) in terms of the specimen Stokes parameters $\mathbf{S}$ and instrument matrix $\mathbf{A}$.

$$\mathbf{I} = \mathbf{AS}, \tag{4}$$

where,

$$I = \begin{bmatrix} I_{LCP} \\ I_0 \\ I_{45} \\ I_{90} \\ I_{135} \end{bmatrix},$$

Each row of the instrument matrix is the Stokes vector of the polarization-state transmitted by detection optics. In other words, intensities recorded by the detector is the projection of the Stokes vector of light on the Stokes vector of the analyzed state. With some derivation, the instrument matrix for our 5-frame measurements turns out to be,

$$A = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 1 & \sin 2\pi\chi & 0 & -\cos 2\pi\chi \\ 1 & 0 & \sin 2\pi\chi & -\cos 2\pi\chi \\ 1 & -\sin 2\pi\chi & 0 & -\cos 2\pi\chi \\ 1 & 0 & -\sin 2\pi\chi & -\cos 2\pi\chi \end{bmatrix} \tag{5}$$

where, $\chi$ is the compensatory retardance that creates elliptical state instead of circular state of illumination (9).

***Computation of Stokes vector at image plane.*** Once the instrument matrix has been experimentally calibrated, the Stokes vector can be obtained from recorded intensities using its inverse (compare Eq. 4),

$$\mathbf{S} = \mathbf{A^{-1}I}, \tag{6}$$

***Computation of background corrected specimen properties.*** We retrieved the Stokes vector of specimen $\mathbf{S}$ by solving Eq. 6. We corrected the specimen Stokes vector for non-uniform background birefringence that were not accounted for by the calibration process. To correct the non-uniform background birefringence, we acquired background polarization images at the empty region of the specimen. We then transformed specimen and background Stokes vectors as follows,

$$\overline{s_{1,2}} = s_{1,2}/s_3 \tag{7}$$

$$DOP = \frac{\sqrt{s_1^2 + s_2^2 + s_3^2}}{s_0} \tag{8}$$

We then reconstructed the background corrected properties of the specimen: brightfield ($BF$), retardance ($\gamma$), slow axis ($\phi$), and degree of polarization ($DOP$) from the transformed specimen and background Stokes vectors $\overline{\mathbf{S}}^{sm}$ and $\overline{\mathbf{S}}^{bg}$ using the following equations:

$$\overline{s_1} = \overline{s_1}^{sm} - \overline{s_1}^{bg} \tag{9}$$

$$\overline{s_2} = \overline{s_2}^{sm} - \overline{s_2}^{bg} \tag{10}$$

$$BF = s_0^{sm}/s_0^{bg} \tag{11}$$

$$\gamma = \arctan 2\left(\sqrt{\overline{s_1}^2 + \overline{s_2}^2}\right) \tag{12}$$

$$\phi = \frac{1}{2}\arctan 2\left(\frac{\overline{s_1}}{-\overline{s_2}}\right) \tag{13}$$

$$DOP = DOP^{sm}/DOP^{bg} \tag{14}$$

In the case where the background cannot be completely removed using the above background correction strategy using a single background measurement, (i.e. the specimen

bioRxiv preprint doi: https://doi.org/10.1101/631101; this version posted May 9, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

has spatially varying background birefringence), we estimated the residual transformed background Stokes parameters by smoothing the transformed specimen Stokes parameters using a 401×401 Gaussian filter with standard deviation $\sigma = 60.5$ and performed another background correction with the estimated residual background.

**Image acquisition and registration.** We implemented LC-PolScope on a Leica DMi8 inverted microscope with Andor Dragonfly confocal for multiplexed acquisition of polarization-resolved images and fluorescence images. We automated the acquisition using Micro-Manager v1.4.22 and OpenPolScope plugin for Micro-Manager that controls liquid crystal universal polarizer (custom device from Meadowlark Optics, specifications available upon request).

We multiplexed the acquisition of label-free and fluorescence volumes. The volumes were registered using transformation matrices computed from similarly acquired multiplexed volumes of 3D matrix of rings from the Argolight test target.

Mouse kidney tissue slice (Thermo-Fisher Scientific) and mouse brain slice were mounted using coverglass and coverslip. U2OS cells were seeded and cultured in a chamber made of two strain-free coverslips that allowed for gas exchange. In the mouse kidney tissue slice, F-actin was labeled with Alexa Fluor 568 phalloidin and DNA was labeled with DAPI.

In transmitted light microscope, the resolution increases and image contrast decreases with increased numerical aperture of illumination. We used 63X 1.47 NA oil immersion objective (Leica) and 0.9 NA condenser to achieve a good balance between image contrast and resolution. The mouse kidney tissue slice was imaged using 100 ms exposure for 5 polarization channels, 200 ms exposure for 405 nm channel (DNA) at 1.6 mW, 100 ms exposure for 561 nm channel (F-actin) at 2.8 mW. The mouse brain slice were imaged using 100 ms exposure for 4 polarization channels. U2OS cells were imaged using 50 ms exposure for 5 polarization channels. For training the neural network, we acquired 160 non-overlapping 2048 x 2048 x 45 z-stacks of the mouse kidney tissue slice with Nyquist sampled voxel size 103 nm x 103 nm x 250 nm.

**Preprocessing.** The images were flat-field corrected and the volumes to 3D models were upsampled along Z to match the pixel size in XY using linear interpolation. The images were tiled into 256 x 256 patches with a 50% overlap between patches for 2D and 2.5D models. The volumes were tiled into 128 x 128 x 96 patches for 3D models with a 25% overlap along XY. Tiles that had sufficient fluorescence foreground (2D: 20%, 2.5D: 25%, 3D: 50%) were used for training. Foreground was quantified as the volume fraction of a mask obtained from Otsu thresholding in case of 2D and Rosin thresholding (29) for 2.5D and 3D models.

We evaluated the effect of Z-scoring the data at the tile scale, at the image scale, and at the stack scale on prediction accuracy. We found that Z-scoring the data at the image (for 2D models) or stack (for 2.5D and 3D models) scale recapitulated intensity variations in the fluorescent structures better than Z-scoring tiles. This effect can be attributed to preservation of the histogram of input and target distributions when the whole volume is Z-scored.

**Network architecture.** Here we experimented with 2D, 2.5D and 3D versions of U-Net models Fig. 5. Across the three U-Net variants, each convolution block in the encoding path consists of two repeats of three layers: a convolution layer, ReLU non-linearity activation, and a batch normalization layer. We added a residual connection from the input of the block to the output of the block to facilitate faster convergence of the model (25, 30). $2 \times 2$ downsampling is applied with 2x2 convolution with stride 2 at the end the each encoding block. On the decoding path, the feature maps were passed through similar convolution blocks followed by upsampling using bilinear interpolation and concatenation of feature maps from the same level of the encoding path. The final output block had a convolution layer only.

The encoding path of our 2D and 2.5D U-Net consists of five layers with 16, 32, 64, 128 and 256 filters respectively. The 3D U-Net consists of four layers with 16, 32, 64 and 128 filters each. The 2D and 3D versions use convolution filters of size of 3×3 and 3×3×3 with a stride of 1 for feature extraction and with a stride of 2 for downsampling between convolution blocks.

The 2.5D U-Net has the same architecture as the 2D U-Net except the convolution filters are N×3×3 in the encoding path with N=3,5,7 corresponding to an input stack with 3, 5, and 7 slices for feature extraction. The feature maps are downsampled across blocks using N×2×2 average pooling. The skip connections consists of a N×1×1 valid convolution, converting the 3D feature maps to 2D. On the decoding path, the feature maps were upsampled using bilinear interpolation by a factor of 1×2×2 and the convolution filters in the decoding path are of shape 1×3×3.

**Model training and inference.** We randomly split the tiles in groups of 70%,15%, and 15% for training, validation and test. The 2D network with single channel input consisted of 2.0 M parameters. The network was trained on mini-batches of size 24 using Adam optimizer, Mean Squared Error (MSE) loss function, and a cyclic learning rate scheduler with a min and max learning rate of $5 \times 10^{-5}$ and $6 \times 10^{-3}$ respectively. Similarly the 2.5D network contained 4.8M parameters. It was trained using a mini-batch size of 20 using a masked Mean Absolute Error (MAE) as loss function, Nadam optimizer and a cyclic learning rate scheduler with a min and max learning rate of $10^{-4}$ and $6 \times 10^{-3}$ respectively. The masks used in the loss function were generated using Rosin thresholding. The 3D network consisted of 1.5M parameters and was trained using a similar set up as the 2.5D network but with a batch size of 4 to accommodate the increased memory requirements of the 3D model.

All the models were trained for up to 200 epochs or until the validation loss converged. The model with minimum loss

bioRxiv preprint doi: https://doi.org/10.1101/631101; this version posted May 9, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

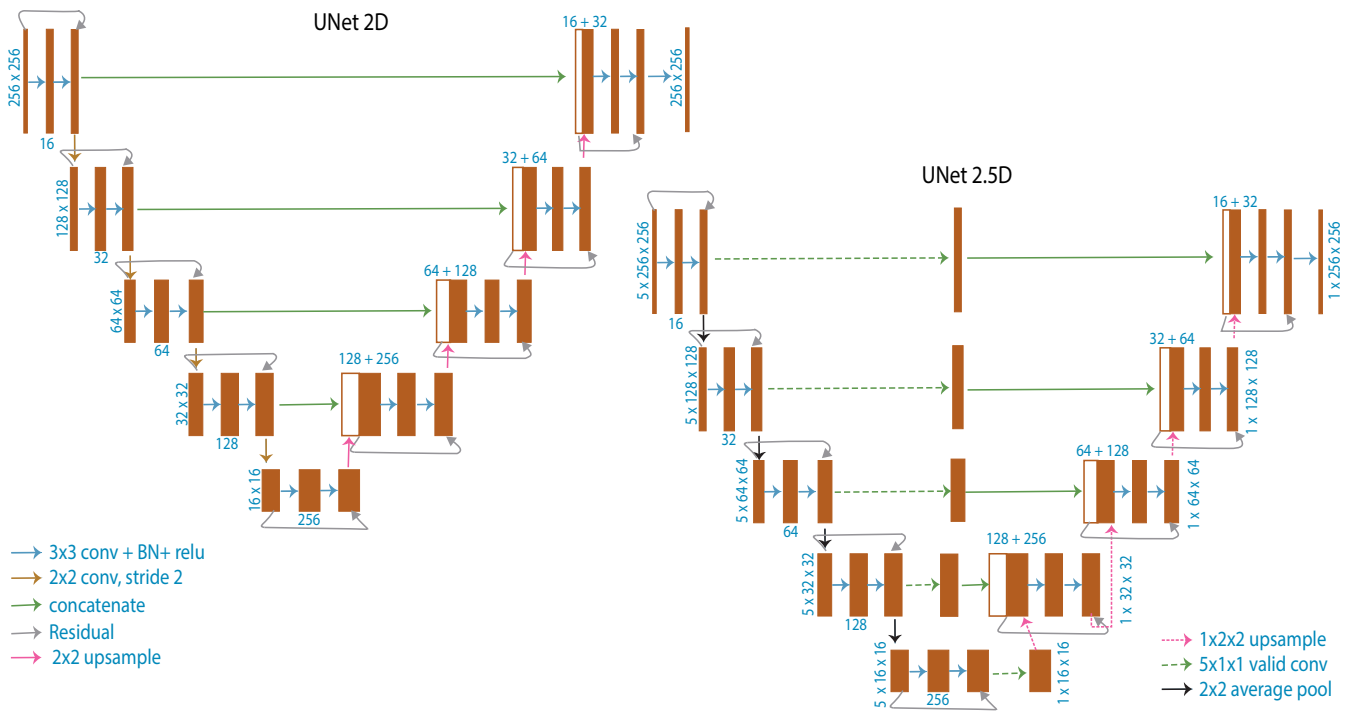


**Fig. 5. Schematic illustrating U-Net architectures:** Schematic of 2D U-Net model used for translating slice→slice and 2.5D U-Net model used for translating stack→slice. The 3D U-Net model used for translating stack->stack is similar to the 2D U-Net, but uses 3D convolutions instead of 2D and is 4 layers deep instead of 5 layers deep.

in the validation set was saved. The single channel 2D models converged in 16-23 hours on a workstation with NVIDIA Pascal Titan X GPU with 12GB RAM. The 2.5D models converged in 24 hours and the 3D model converged in 72 hours on NVIDIA Tesla V100 GPU with 32GB RAM.

As the models are fully convolutional, model predictions were obtained using full XY images as input for the 2D and 2.5D versions. Due to memory requirements of the 3D model, the test volumes were tiled along x and y while retaining the entire z extent (patch size: $96 \times 512 \times 512$) with an overlap of 32 pixels along x and y. The predictions were stitched together by weighted averaging the model predictions in the overlapping regions.

**Model evaluation.** Pearson correlation and structural similarity index (SSIM) along the XY, XZ and XYZ dimensions of the test volumes were used for evaluating model performance.

The Pearson correlation between a target image $T$ and a prediction image $P$ is defined as

$$r(T,P) = \frac{\sigma_{TP}}{\sigma_T \sigma_P} \tag{15}$$

where $\sigma_{TP}$ is the covariance of $T$ and $P$, and $\sigma_T$ and $\sigma_P$ are the standard deviations of $T$ and $P$ respectively.

SSIM compares two images using a sliding window approach, with window size $N \times N$ ($N \times N \times N$ for XYZ). Assuming a target window $t$ and a prediction window $p$,

$$\text{SSIM}(t,p) = \frac{(2\mu_t\mu_p + c_1)(2\sigma_{tp} + c_2)}{(\mu_t^2 + \mu_p^2 + c_1)(\sigma_t^2 + \sigma_p^2 + c_2)} \tag{16}$$

where $c_1 = (0.01L)^2$ and $c_2 = (0.03L)^2$, and $L$ is the dynamic range of pixel values. Mean and variance are represented by $\mu$ and $\sigma^2$ respectively, and the covariance between $t$ and $p$ is denoted $\sigma_{tp}$. We use $N = 7$. The total SSIM score is the mean score calculated across all windows, $\text{SSIM}(T,P) = \frac{1}{M}\sum \text{SSIM}(t,p)$ for a total of $M$ windows. For XY and XZ dimensions, we compute one test metric per plane and for XYZ dimension, we compute one test metric per volume.

## Conclusion

In summary, we report synergistic combination of image acquisition, model-driven reconstruction of optical properties, and data-driven prediction to reveal ordered structures from polarized light images. Our Stokes-model based reconstruction algorithms (https://github.com/czbiohub/reconstructorder) and computationally efficient variant of U-Net architecture (https://github.com/czbiohub/microdl) allows facile analysis of complementary label-free signatures of specimen. We report simultaneous recovery of background-corrected brightfield, birefringence, orientation, and degree of polarization contrasts with diffraction-limited spatial resolution. These contrasts report variations in density, order, and scattering by the specimen. We report rich imaging data that visualize diverse structures: glomeruli and tubules in kidney tissue, multiple organelles in cells, and axon tracts in white and gray matter of brain slice. We report memory and compute-efficient 2.5D U-Net models to reveal the structures using information in complementary contrasts. We demonstrate that our 2.5D U-Net performs as well as 3D U-Net when multiple label-free

bioRxiv preprint doi: https://doi.org/10.1101/631101; this version posted May 9, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

images are used as inputs. We demonstrate accurate prediction of ordered F-actin and condensed DNA in heterogeneous structures within a tissue. We anticipate that our approach will enable scalable analysis of architectural order that underpins healthy and disease states of cells and tissues.

**ACKNOWLEDGEMENTS**

We thank Spyros Dermanis and Bing Wu for providing the mouse brain slice used for acquiring data shown in Fig. 1. We thank Greg Huber, Loic Royer, Joshua Batson, Jim Karkanias, Joe DeRisi, and Steve Quake for helpful discussions. This research was supported by the Chan Zuckerberg Biohub.

## Bibliography

1. Shinya Inoué. Polarization optical studies of the mitotic spindle. *Chromosoma*, 5(1):487–500, doi: 10.1007/BF01271498.
2. David Keefe, Lin Liu, Wei Wang, and Celso Silva. Imaging meiotic spindles by polarization light microscopy: Principles and applications to IVF. *Reproductive BioMedicine Online*, 7 (1):24–29, doi: 10.1016/S1472-6483(10)61724-5.
3. Markus Axer, David Grässel, Melanie Kleiner, Jürgen Dammers, Timo Dickscheid, Julia Reckfort, Tim Hütz, Björn Eiben, Uwe Pietrzyk, Karl Zilles, and Katrin Amunts. High-Resolution Fiber Tract Reconstruction in the Human Brain by Means of Three-Dimensional Polarized Light Imaging. *Frontiers in Neuroinformatics*, 5, doi: 10.3389/fninf.2011.00034.
4. Miriam Menzel, Julia Reckfort, Daniel Weigand, Hasan Köse, Katrin Amunts, and Markus Axer. Diattenuation of Brain Tissue and its Impact on 3D Polarized Light Imaging. *arXiv:1703.04343 [physics]*,
5. Maki Koike-Tani, Takashi Tominaga, Rudolf Oldenbourg, and Tomomi Tani. Instantaneous polarized light imaging reveals activity dependent structural changes of dendrites in mouse hippocampal slices. *bioRxiv*, page 523571,
6. Shalin B. Mehta, Molly McQuilken, Patrick J. La Riviere, Patricia Occhipinti, Amitabh Verma, Rudolf Oldenbourg, Amy S. Gladfelter, and Tomomi Tani. Dissection of molecular assembly dynamics by tracking orientation and position of single molecules in live cells. *Proceedings of the National Academy of Sciences*, 113(42):E6352–E6361, doi: 10.1073/pnas.1607674113.
7. Brittany J. Belin, Lauren M. Goins, and R. Dyche Mullins. Comparative analysis of tools for live cell imaging of actin network architecture. *BioArchitecture*, 4(6):189–202, doi: 10.1080/19490992.2014.1047714.
8. R. Oldenbourg and G. Mei. New polarized light microscope with precision universal compensator. *Journal of Microscopy*, 180(2):140–147, doi: 10.1111/j.1365-2818.1995.tb03669.x.
9. Michael Shribak and Rudolf Oldenbourg. Techniques for fast and sensitive measurements of two-dimensional birefringence distributions. *Applied Optics*, 42(16):3009–3017, doi: 10.1364/AO.42.003009.
10. Shalin B. Mehta, Michael Shribak, and Rudolf Oldenbourg. Polarized light imaging of birefringence and diattenuation at high resolution and high sensitivity. *Journal of Optics*, 15(9): 094007, Invited.
11. Max Born and Emil Wolf. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Elsevier,
12. Hui Wang, Caroline Magnain, Ruopeng Wang, Jay Dubb, Ani Varjabedian, Lee S. Tirrell, Allison Stevens, Jean C. Augustinack, Ender Konukoglu, Iman Aganj, Matthew P. Frosch, Jeremy D. Schmahmann, Bruce Fischl, and David A. Boas. As-PSOCT: Volumetric microscopic imaging of human brain architecture and connectivity. *NeuroImage*, 165:56–68, doi: 10.1016/j.neuroimage.2017.10.012.
13. Johannes F. de Boer, Christoph K. Hitzenberger, and Yoshiaki Yasuno. Polarization sensitive optical coherence tomography – a review [Invited]. *Biomedical Optics Express*, 8(3):1838–1873, doi: 10.1364/BOE.8.001838.
14. Chawin Ounkomol, Sharmishtaa Seshamani, Mary M Maleckar, Forrest Collman, and Gregory Johnson. Label-free prediction of three-dimensional fluorescence images from transmitted light microscopy. doi: 10.1101/289504.
15. Eric M. Christiansen, Samuel J. Yang, D. Michael Ando, Ashkan Javaherian, Gaia Skibinski, Scott Lipnick, Elliot Mount, Alison O'Neil, Kevan Shah, Alicia K. Lee, Piyush Goyal, William Fedus, Ryan Poplin, Andre Esteva, Marc Berndl, Lee L. Rubin, Philip Nelson, and Steven Finkbeiner. In Silico Labeling: Predicting Fluorescent Labels in Unlabeled Images. *Cell*, 173(3):792–803.e19, doi: 10.1016/j.cell.2018.03.040.
16. Yair Rivenson, Tairan Liu, Zhensong Wei, Yibo Zhang, and Aydogan Ozcan. PhaseStain: Digital staining of label-free quantitative phase microscopy images using deep learning. *arXiv:1807.07701 [physics]*,
17. Yair Rivenson, Hongda Wang, Zhensong Wei, Yibo Zhang, Harun Gunaydin, and Aydogan Ozcan. Deep learning-based virtual histology staining using auto-fluorescence of label-free tissue. *arXiv:1803.11293 [physics]*,
18. Moosung Lee, Young-Ho Lee, Jinyeop Song, Geon Kim, YoungJu Jo, HyunSeok Min, Chan Hyuk Kim, and YongKeun Park. DeepIS: Deep learning framework for three-dimensional label-free tracking of immunological synapses. *bioRxiv*, page 539858, doi: 10.1101/539858.
19. D. Petersen, L. Mavarani, D. Niedieker, E. Freier, A. Tannapfel, C. Kötting, K. Gerwert, and S. F. El-Mashtoly. Virtual staining of colon cancer tissue by label-free Raman microspectroscopy. *Analyst*, 142(8):1207–1215, doi: 10.1039/C6AN02072K.
20. Chawin Ounkomol, Sharmishtaa Seshamani, Mary M. Maleckar, Forrest Collman, and Gregory R. Johnson. Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy. *Nature Methods*, 15(11):917, doi: 10.1038/s41592-018-0111-2.
21. Shalin B. Mehta and Rudolf Oldenbourg. Image simulation for biological microscopy: Microlith. *Biomedical Optics Express*, 5(6):1822–1838,
22. Norbert Streibl. Phase imaging by the transport equation of intensity. *Optics communications*, 49(1):6–10,
23. Benedicto de Campos Vidal, Maria Luiza Silveira Mello, Alberto Costa Caseiro-Filho, and Carlos Godo. Anisotropic properties of the myelin sheath. *Acta Histochemica*, 66(1):32–39, doi: 10.1016/S0065-1281(80)80079-1.
24. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Lecture Notes in Computer Science, pages 234–241. Springer, Cham, doi: 10.1007/978-3-319-24574-4_28.
25. F. Milletari, N. Navab, and S. Ahmadi. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In *2016 Fourth International Conference on 3D Vision (3DV)*, pages 565–571, doi: 10.1109/3DV.2016.79.
26. Z. Wang and A. C. Bovik. Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures. *IEEE Signal Processing Magazine*, 26(1):98–117, doi: 10.1109/MSP.2008.930649.
27. C. Schell, N. Wanner, and T. B. Huber. Glomerular development–shaping the multi-cellular filtration unit. In *Seminars in Cell & Developmental Biology*, volume 36, pages 39–49. Elsevier,
28. Michael Bass, Casimer DeCusatis, Jay M. Enoch, Vasudevan Lakshminarayanan, Guifang Li, Carolyn MacDonald, Virendra N. Mahajan, and Eric Van Stryland. *Handbook of Optics, Volume I: Geometrical and Physical Optics, Polarized Light, Components and Instruments(Set)*. McGraw Hill Professional,
29. Paul L. Rosin. Unimodal thresholding. *Pattern Recognition*, 34(11):2083–2096, doi: 10.1016/S0031-3203(00)00136-9.
30. Michal Drozdzal, Eugene Vorontsov, Gabriel Chartrand, Samuel Kadoury, and Chris Pal. The Importance of Skip Connections in Biomedical Image Segmentation. *arXiv:1608.04117 [cs]*,

bioRxiv preprint doi: https://doi.org/10.1101/631101; this version posted May 9, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

## Supplementary movies and images

**Movie 1** Z-stacks showing (A) brightfield (BF), (B) retardance ($\gamma$),(C) color-coded orientation ($\phi$), and (D) degree of polarization (DOP) images of kidney tissue corresponding to Fig. 1C.

**Movie 2** 4D movie of brightfield and color-coded orientation images of U2OS cell undergoing mitosis. The movie shows following sequence: z-stack at prometaphase, timelapse from prometaphase to metaphase at speicific depth, z-stack at metaphase, timelapse from metaphase to cytokinesis at specific z, and z-stack at cytokinesis. Related to Fig. 1C.

**Movie 3** Z-stack of (A) F-actin and (B) DNA in a test field of view. Related to Fig. 2, Fig. 3, and Fig. 4.

**Movie 4** Z-stacks of predicted F-actin for the test field of view using (A) 2D model trained on BF,$\gamma$, $\phi$, (B) 2.5D model trained on BF,$\gamma$, $\phi$, and (C) 3D model trained on $\gamma$. Related to Fig. 2, Fig. 3, and Fig. 4.

**Movie 5** Through focus stacks of predicted DNA for the test field of view using (A) 2D model trained on BF,$\gamma$, $\phi$ and (B) 2.5D model trained on BF,$\gamma$, $\phi$. Related to Fig. 2 and Fig. 4.

**Image 1** Stitched fields of view of (A) brightfield, (B) retardance, (C) orientation, and (D) degree of polarization images of mouse brain slice. Each stitch is downsampled by 4.5x to reduce file size. The retardance and orientation images were gamma corrected (0.5) to visualize less ordered gray matter in the presence of highly ordered white matter.