Revealing architectural order with quantitative label-free imaging and deep neural networks

Syuan-Ming Guo[1,*], Li-Hao Yeh[1,*], Jenny Folkesson[1,*], Ivan E. Ivanov[1,#], Anitha Priya Krishnan[1,#,†], Matthew G. Keefe[2,#], David Shin[2], Bryant Chhun[1], Nathan Cho[1,‡], Manuel Leonetti[1], Tomasz J. Nowakowski[2], and Shalin B. Mehta[1,✉]

[1]Chan Zuckerberg Biohub
[2]Department of Anatomy, University of California, San Francisco
[*]S-M.G., L-H.Y., and J.F. contributed equally
[#] I.E.I., A.P.K., and M.G.K. contributed equally
[†]current address: Genentech
[‡]current address: University of California, San Francisco
[✉]correspondence: shalin.mehta@czbiohub.org

Abstract

Methods for imaging architecture of biological systems are currently not as scalable as genomic, transcriptomic, or proteomic technologies, because they rely on labels instead of intrinsic signatures. Label-free visualization of diverse biological structures is feasible with phase and polarization of light. However, distinguishing structures from information-dense label-free images is challenging. Recent advances in deep learning can distinguish structures from label-free images based on their shape. Here, we report joint use of polarization resolved imaging, reconstruction of optical properties, and deep neural networks for scalable analysis of complex structures. We reconstruct quantitative three-dimensional density, anisotropy, and orientation from polarization- and depth-resolved images. We report a computationally efficient variant of U-Net architecture to predict 3D fluorescent structure from its density, anisotropy, and orientation. We evaluate the performance of our models by predicting anisotropic F-actin and isotropic nuclei. We report label-free prediction of myelination in human brain tissue sections and demonstrate the model's ability to rescue inconsistent labeling. We anticipate that the proposed approach will enable quantitative analysis of architectural order across scales of organelles to tissues.

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

Introduction

The function of living systems emerges from dynamic interaction of components over spatial and temporal scales that range many orders of magnitudes. Although molecular components of biological systems can now be identified with scalable genomic, transcriptomic, and proteomic technologies, these technologies cannot capture dynamic interactions of components and are destructive end-point measurements. There remains an unmet need to develop technologies for mapping interactions among many components across spatial and temporal scales.

Light microscopy is uniquely suited to study dynamic arrangement of molecules within the context of organelles, of organelles within the context of cells, and of cells within the context of tissues. Advances in fluorescence and label-free microscopy, combined with complementary computational methods, continue to provide new insights in biological systems. The label is simultaneously the strength and the weakness of fluorescence microscopy. Fluorescent labels can report specific components in a crowded biological system beyond the resolution limit, but stochastic nature of labeling, perturbation due to labels, and broad spectra of fluorescent proteins make it difficult to analyze more than 5 to 7 fluorescent components in live systems (1). Further, labeling of primary patient specimens and diverse non-model organisms is difficult. On the other hand, label-free imaging can provide simultaneous and reproducible measurements of many biological components, even in live primary specimens or non-model organisms. Most widely used label-free imaging methods rely on intrinsic properties of light such as phase (2) and polarization (3, 4), which have been used for qualitative visualization of biological processes for almost a century. Recently developed quantitative phaes and polarization imaging approaches report physical properties of specimens not accessible with fluorescence and allow simultaneous visualization of many components with minimal photo-toxicity. Combining advantages of quantitative label-free and fluorescence microscopy via computation promises new avenues to study biology and pathology.

While label-free acquisition is facile, the information-dense data it generates is significantly more challenging to analyze as compared to fluorescence for two reasons: 1) In label-free images, optical properties of structures are encoded in complex modulation of the intensity, rather than intensity itself. 2) Many structures seen in label-free data have similar optical properties and are therefore difficult to distinguish with classical analysis tools. Human vision can interpret shape and dynamics, in addition to density and anisotropy to distinguish among structures. To overcome these challenges, we report a computational imaging method for quantitative analysis of density and structural anisotropy in 3D, which we collectively refer to as the architectural order in this paper. We also report a computationally efficient deep neural networks that predict 3D fluorescent structures from 3D architectural order. The predicted fluorescence can be analyzed with widely available tools.

Related work. Variations in concentration and anisotropic assembly of biomolecules gives rise to variations in phase and polarization of light. As opposed to qualitative methods mentioned above, one needs quantitative phase and polarization-resolved imaging to measure density and anisotropy. Detectors are sensitive only to intensity (i.e., brightness) and wavelength (i.e., color), not the relative times of arrival of light wave (i.e., phase) nor to the plane of oscillation of electric field (i.e., polarization). Quantitative methods consist of imaging systems that convert phase and polarization into intensity modulations, and inverse algorithms that reconstruct density and anisotropy using the model of image formation.

Phase can be converted into intensity modulation by propagation of light (5, 6), by interfering the light scattered by the specimen with another beam whose phase is controlled (7–10), and by illuminating the specimen from diverse angles (11–13). Recent advances in computational microscopy show that robust retrieval of phase from through-focus series is possible via weak object transfer function model of image formation (14–19). Polarization can be converted into intensity modulation using optical components that selectively transmit and change polarization (20–22). Quantitative imaging of structural anisotropy has been demonstrated by probing the specimen with complementary polarization states of light, which we have achieved with liquid-crystal based PolScope (LC-PolScope) (20, 23, 24). Accurate calibration is key to the successful use of quantitative phase and polarization microscopes, because phase and polarization of light are affected by not just the specimen, but also optical interactions in the rest of the light path.

Quantitative phase imaging has been used to analyze membrane mehanics, density of organelles, cell migration, and action potential (7, 25), among other processes. Polarized light imaging has enabled discovery of the dynamic microtubule spindle (4), assessment of structural integrity of meiotic spindles of oocytes in in-vitro fertilization (IVF) clinics (26), label-free imaging of white matter in brain tissue slices (22, 27), and imaging the activity dependent structural changes in brain tissue (28). Fluorescence polarization microscopy, which converts the anisotropy of the structure into anisotropy of fluorescent dipoles, lends itself to automated analysis of dynamic order (29) of molecules. However, a fluorescent reporter often compromises assembly of ordered structures, such as actin network (30).

Phase microscopes report the optical path length, which is usually modulated by the specimen density. On the other hand, polarization microscopes measure retardance, i.e., orientation-dependent optical path length, which is modulated by the structural anisotropy below the spatial diffraction limit of resolution [1] and has been used to make measurements well below the resolution limit (32). Shribak *et al.* (33) developed a method that employed multiple liquid crystal modulators for joint analysis of phase and retardance, which provided high-resolution 2D static label-free images of organelles in insect cells.

To our knowledge, approaches to measure density and anisotropy in live cells in 3D have not been reported.

Machine learning approaches have recently enabled identification of structures that can be perceived in diverse label-free images, opening new opportunities for scalable analysis of label-free images. 3D U-Net has been used to identify multiple organelles in cells in 3D from brightfield and differential interference contrast (DIC) images (34) for 3D of . Phase-contrast and DIC were combined with adaptation of the inception model for in silico 2D labeling of nuclei, cell types, and cell state (35). Quantitative phase (36) and auto-fluorescence (37) were combined with generative adverserial networks for prediction of 2D histopathological stains. Diffraction tomography was combined with U-Net for measuring 3D shape of immunological synapse (38). Raman scattering was combined with a random forest classifier (39) to identify of carcinoma regions in colon tissue. Among these approaches, the 3D U-Net models that translate brightfield stack to a fluorescence stack (40) are attractive for scalable analysis because they predict both localization and expression of specific molecules in 3D. However, learning anisotropic structures, such as cytoskeletal networks, could be challenging using these approaches (40) because they are not sensitive to structural anisotropy. Interpreting density information from brightfield images is also difficult because of the complex relation between the image contrast and the underlying density. Moreover, training and prediction using 3D U-Net with single contrast input is already computationally expensive and often requires downsampling the images with the cost of losing spatial resolution (40).

There is an unmet need of computationally efficient 3D image translation models that allows for exploiting multiple contrast inputs.

Contributions. We sought to measure density and anisotropy jointly in live 3D specimens. We reasoned that density variations can be encoded in the intensity using a depth-resolved acquisition, whereas aniostropy can be encoded in the intensity using a polarization-resolved acquisition. We also sought to develop more computation and memory efficient image translation models to analyze information-rich maps of density and anisotropy.

We report a computational imaging approach that recovers phase and retardance from single polarization- and depth-resolved acquisition. Depth-resolved acquisition in transmitted light microscopes is widely accessible, but the opportunity for quantitative measurements is often missed due to the lack of robust and accessible inverse algorithms to retrieve physical properties quantitatively. Our measurements reveal components across different biological scales, including lipid droplets, mitochondria, spindle, nucleoli, condensed chromosomes, nuclei, nuclear envelope, plasma membrane, glomeruli and tubules in mouse kidney tissue, axon tracts in mouse and human brain tissue, and cell morphology in human brain tissue. The reconstruction relies on the descrip-

[1] To the first approximation, optical path length and retardance measured by the microscope are refractive index and birefringence of the specimen, respectively, integrated over the coherently illuminated volume of the specimen (31)

tion of polarization of light by a Stokes vector (41) and recovery of phase from 3D brightfield stack via deconvolution. 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 background correction. The formalism also enables design of new polarization-resolved imaging algorithms. Using the Stokes formalism and deconvolution algorithms reported here, we measure phase, retardance, slow axis [2], and degree of polarization of transparent specimen.

Next, we report a computationally efficient deep neural network architecture for translating 3D distribution of optical properties to fluorescence intensities with state-of-the-art accuracy. Our 2.5D multi-channel models accurately translate brightfield, phase, retardance, and slow axis distributions over the depth of field of the microscope into fluorescence intensities in the corresponding 2D focal plane. 3D translation is achieved by using the model to predict fluorescence in each plane of the stack. Our 2.5D architecture follow intuition similar to architectures used for 3D segmentatation of medical imaging data from projections (42), but is informed by our understanding of image formation in light microscopes. Formulating the prediction problem over the depth of field allows us to work with substantially larger fields of view in a computationally efficient manner as compared to full 3D models. Using quantitative data as input removes spurious correlations that can arise between label-free intensities and fluorescence intensities due to background or image formation, which improves accuracy of prediction and allows cleaner interpretation of predicted structure. Further, measured density and anisotropy can be analyzed jointly with predicted fluorescent structure to gain new insights in biology.

We applied our approach to mapping the axon tracts and myelination in a developing human brain across different ages. In contrast to brains of model systems (e.g., mouse and fruit fly), scalable mapping of the architecture of the human and other large brains poses significant challenges. For example, genetic tools are limited and not scalable for tracing the projections at high-throughput. To overcome this issue, we reasoned that joint imaging of density and structural anisotropy will allow us to more thoroughly describe the structures of human brain tissue.

Collectively, above contributions establish a novel approach for imaging architectural order in living systems across biological scales and analyzing it with a judicious combination of physics-driven and data-driven modeling approaches.

Results

Reconstructing density and anisotropy. We implemented automated polarized-resolved and depth-resolved imaging, which allows retrieval of physical properties by illuminating specimens with 5 different polarization states, using LC-PolScope (fig. 1A, methods). We developed algorithms (fig. 1B, methods) for simultaneous recovery of brightfield, phase, retardance, slow axis, and degree of polarization contrasts. We developed background correction methods that enable detection of low contrast from the biological structures even in the presence of high, non-uniform background resulting from the optics or specimen (methods). The Python code for reconstruction is available at `https://github.com/mehta-lab/reconstruct-order`.

Simultaneous recovery of multiple properties requires a model of image formation accurate enough to account for partial polarization of light. Previous reconstruction algorithms (23) for transmission polarized light microscopy employed Jones calculus (41, Ch.10) to describe image formation, which assumes fully polarized illumination. However, LEDs or lamps used for transmission microscopy lead to partial polarized illumination that cannot be accounted for by Jones calculus. Stokes vector representation of light and Mueller matrix representation of the optical components cleanly 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 (29). Fluorescence polarization microscopy also gives rise to partially polarized emission, since independent emission events integrated through the imaging aperture are mutually incoherent. In this paper, we employ the Mueller calculus to describe transmission imaging with partial polarization. We exploit the same framework for calibration, background correction, and reconstruction of density and degree of polarization.

Figure 2 shows background-corrected images of a mouse kidney tissue slice, U2OS cells (bone cancer cell line), mouse brain slice, and developing human brain slice in brightfield, phase, retardance, and orientation contrasts. Both the phase and brightfield images report density variation in the specimen. However, optically dense structures appear in positive contrast on one side of the focus, almost no contrast at the focus, and negative contrast on the other side of the focus in the brightfield image. Note the contrast variations in images of nuclei in the XZ section of the brightfield image of the mouse kidney tissue and in Video 1. This contrast variation along the optical axis, described by the transport of intensity equation (5), makes proper interpretation of density from brightfield images challenging. We developed a phase reconstruction algorithm (Methods) that translates this axial contrast modulation into a more quantitative contrast proportional variations in density, such that denser structures consistently appear with positive contrast in the image. The retardance image is proportional to the degree of liquid crystalline or orientational alignment among molecules. The orientation image reports the slow axis orientation of the retardance. In mouse kidney tissue, the retardance image highlights capillaries within glomeruli, and brush borders in convoluted tubules, among other components of the tissue. The nuclei appear in negative contrast in the retardance image because of the relatively isotropic structure of condensed chromatin.

[2] Slow axis is the orientation along which the anisotropic structure is denser. Higher density leads to slower propagation of light when polarization of the light is oriented along the slow axis

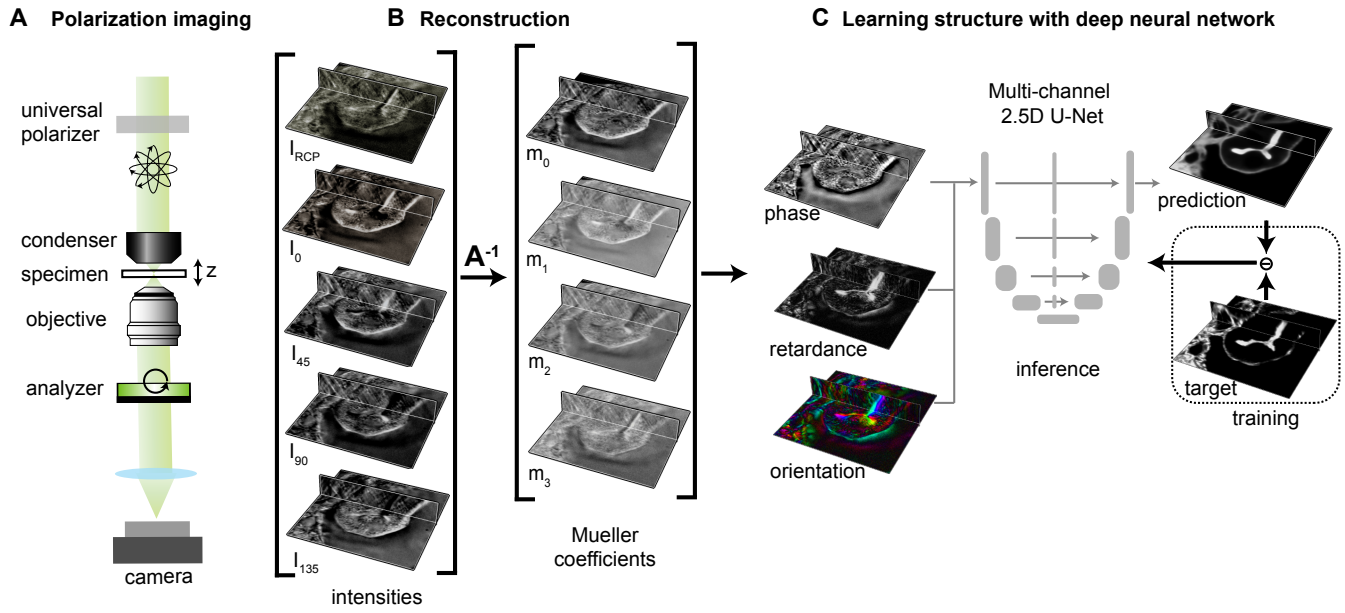


Fig. 1. Label-free prediction of a structure from density and anisotropy measurements: (A) Light-path of label-free polarization-resolved and depth-resolved imaging microscope. The measured intensities encode spatial distributions of refractive index and birefringence of the specimen. Polarization states are generated by a calibrated liquid-crystal universal polarizer that can be controlled electronically. **(B)** Mueller coefficients of the specimen are computed for every voxel from polarization-resolved intensities using the inverse of instrument matrix $\mathbf{A}^{-1}$ (see eq. (8)). Mueller coefficients retrieved from the background region are used to correct imperfections in the light path (not shown). Assuming that the specimen is transparent, complementary optical properties of the specimen are reconstructed from Mueller coefficients: retardance, orientation of slow axis, brightfield, phase, and degree of polarization (not shown). **(C)** Multi-channel, 2.5D U-Net model is trained to predict 3D fluorescent structures from label-free measurements. During training, pairs of label-free images and fluorescence images are supplied as inputs and targets to the U-Net model. The model is optimized by minimizing the difference between the model prediction and the target. During inference, only label-free images are supplied as input to the trained model to predict 3D fluorescence images.

In the dividing U2OS cell (Video 2), the phase image shows contrast consistent with density of various cellular structures including cytoplasm, lipid vesicles, nucleolus, and chromosomes compared to the brightfield image where the contrast of the structures varies depending on their location relative to the focus. Similarly, in mouse and developing human brain tissue sections, the phase image identifies axon tracts better than brightfield image, because of variations in their density (fig. 2-suppelement S1). The retardace and orientation in U2OS cells (Video 2) show dynamics of membrane boundaries, spindle, and lipid droplets. The retardance and orientation images of mouse and human brain slices in fig. 2 distinctly report axon tracts. The birefringence of the axons arises from neurofilaments that have higher density along axon axis and myelin sheath that has higher density perpendicular to the axon axis (43). Due to the high birefringence of myelinated axons in mouse brain slice, we see a slow axis perpendicular to the direction of the axon tracts. Figure 2-supplement S2 shows stitched retardance and orientation images of a whole mouse brain slice, in which not just the white matter tracts, but also changes in the orientation of axons across different cortical layers are visible.

We show degree of polarization measurements in fig. 2-supplement S3. It is worth clarifying the difference between retardance images shown in fig. 2 and degree of polarization images. The retardance variations arise from single scattering events within the specimen that alter the polarization, but do not reduce the degree of polarization. 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 plan to pursue models that account for diffraction and scattering effects in polarized light microscopy and enable even more precise retrieval of specimen properties.

Data shown in fig. 2 and in Video 3 report simultaneous, quantitative measurements of density, structural anisotropy, and orientation in 3D biological specimens, for the first time to our knowledge. In the next sections, we discuss how these complementary label-free measurements enable prediction of fluorescence images of different types of structures.

Computationally efficient 3D prediction using 2.5D residual U-Net. Joint optimization of optical contrast, architecture of the deep neural network, and the training process is key to successful analysis of structures of interest. Label-free measurement of density and anisotropy simultaneously visualize several structures that are normally imaged with fluorescence labeling. To enable automated analysis of localization and expression of specific molecules, we sought to develop deep convolutional neural network models that translate 3D label-free stacks into 3D fluorescence stacks. We optimized our model architecture using mouse kidney tissue sections where F-actin and nuclei are labeled.

We adapted the U-Net architecture that has been widely successful in image segmentation and translation tasks (44, 45). U-Net's success arises from its ability to exploit image features at multiple spatial scales. It uses skip connections between the encoding and decoding blocks that give decoding blocks access to low-complexity, high-resolution features

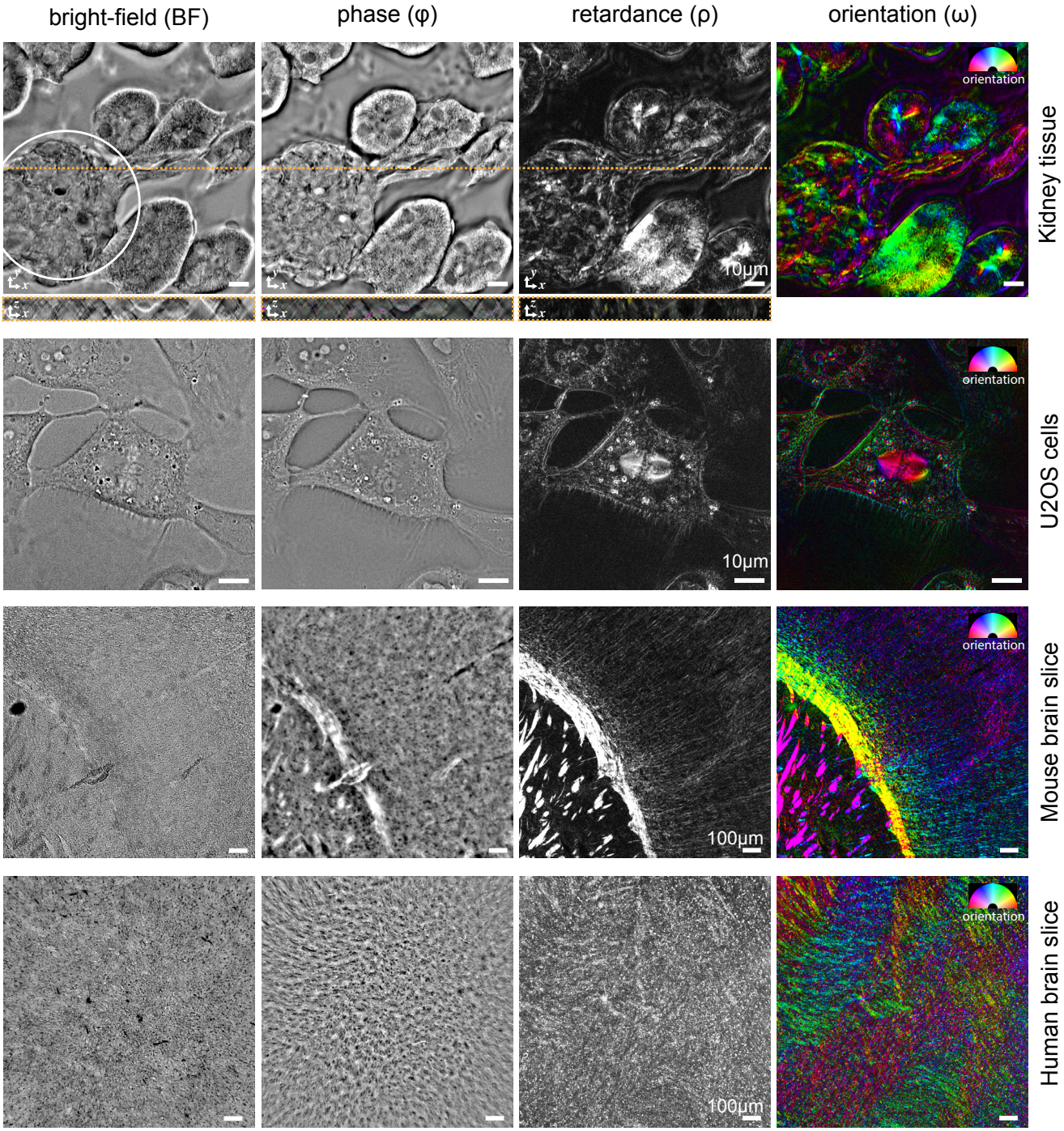


Fig. 2. Complementary contrasts of phase and retardance report diverse biological structures: Brightfield (BF), Phase (ϕ), retardance (ρ), and orientation (ω) images of mouse kidney tissue, mammalian cell line (U2OS), mouse brain tissue section, and human brain tissue section are shown. The orientation image represents slow axis of the specimen with color (hue) and retardance with brightness. In kidney tissue, nuclei appear as smooth patches of low density and low retardance, while tubules stand-out in retardance. Glomelurus, a network of small blood vessels, is identified by circular outline in the brightfield image of kidney tissue. In U2OS cells, chromatin, lipid droplets, membranous organelles, and cell boundaries are visible due to variations in density, while microtubule spindle, lipid droplets, and cell boundaries are visible due to their ordered structure. In mouse brain slices, axon tracts are more visible in phase, retardance, and orientation images compared to brightfield images, with slow axis perpendicular to the directions of the bundles. Similar contrast improvement was observed in developing human brain tissue slice with less ordered tracts due to the early age of the specimen. 3D stacks of mouse kidney tissue and U2OS cell were acquired with 63x 1.47 NA oil objective and 0.9 NA illumination, whereas images of mouse and human brain tissue were acquired with 10x 0.3 NA air objective and 0.2 NA illumination.

in the encoding blocks. We added a residual connection between the input and output of each block to speed up training of the model (45, 46). Prior work (40) on predicting fluorescence stacks from brightfield stacks has shown that 2D models result in discontinuous predictions along the z-axis as compared to 3D translation models. However, training transforming entire 3D stacks at a time is computation and memory intensive and, as we illustrate, not necessary. Further, typical microscopy stacks are bigger in their extent in the focal plane and smaller in extent along the optical

Translation model	Input(s)	Loss	r_{xy}	r_{xz}	r_{xyz}	$SSIM_{xy}$	$SSIM_{xz}$	$SSIM_{xyz}$
Slice-> Slice (2D)	ρ	L1	0.82	0.79	0.83	0.78	0.71	0.78
Stack-> Slice (2.5D, $z = 3$)	ρ	L1	0.85	0.83	0.86	0.80	0.75	0.81
Stack-> Slice (2.5D, $z = 5$)	ρ	L1	0.86	0.84	**0.87**	0.81	0.76	0.82
Stack-> Slice (2.5D, $z = 7$)	ρ	L1	**0.87**	**0.85**	**0.87**	**0.82**	**0.77**	0.83
Stack-> Stack (3D, $z = 96$)	ρ	L1	0.86	0.84	0.86	**0.82**	0.76	**0.85**

Table 1. Accuracy of 3D prediction of F-actin from retardance stack using different neural networks: Above table lists median values of the Pearson correlation (r) and structural similarity index (SSIM) between prediction and ground truth F-actin volumes. We report accuracy metrics for Slice→Slice (2D) ,Stack→Slice (2.5D), and Stack→Stack (3D) models trained to predict F-actin from retardance using Mean Absolute Error (MAE or L1) loss. We segmented target images with Rosin 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 (Methods) 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 entire test volumes (r_{xyz}, $SSIM_{xyz}$).

Translation model	Input(s)	Loss (mask)	r_{xy}	r_{xz}	r_{xyz}	$SSIM_{xy}$	$SSIM_{xz}$	$SSIM_{xyz}$
Stack-> Slice (2.5D, $z = 5$)	ρ	L1	0.86	0.84	0.87	0.81	0.76	0.82
	BF	L1	0.86	0.84	0.86	0.82	0.77	0.83
	ϕ	L1	0.87	0.85	0.88	**0.83**	0.78	0.84
	ϕ, ρ, ω_x, ω_y	L1	**0.88**	**0.87**	**0.89**	**0.83**	**0.80**	**0.85**
	BF, ρ, ω_x, ω_y	L1	**0.88**	**0.87**	**0.89**	**0.83**	0.79	**0.85**

Table 2. Accuracy of prediction of F-actin in mouse kidney tissue as a function of input channels: Above table lists median values of the Pearson correlation (r) and structural similarity index (SSIM) between prediction and target volumes of F-actin. We evaluated combinations of brightfield (BF), phase (ϕ), retardance (ρ), orientation x (ω_x), and orientation y (ω_y), as input. Model training conditions and computation of test metrics is described in table 1.

Translation model	Input(s)	Loss (mask)	r_{xy}	r_{xz}	r_{xyz}	$SSIM_{xy}$	$SSIM_{xz}$	$SSIM_{xyz}$
Stack-> Slice (2.5D, $z = 5$)	ρ	L1	0.84	0.85	0.85	0.81	0.76	0.82
	BF	L1	0.87	0.88	0.87	0.82	0.77	0.84
	ϕ	L1	0.88	0.88	0.88	0.83	0.78	0.85
	ϕ, ρ, ω_x, ω_y	L1	**0.89**	0.89	**0.89**	**0.84**	**0.80**	**0.86**
	BF, ρ, ω_x, ω_y	L1	**0.89**	**0.90**	**0.89**	**0.84**	**0.80**	**0.86**

Table 3. Accuracy of prediction of nuclei in mouse kidney tissue: Above table lists median values of the Pearson correlation (r) and structural similarity index (SSIM) between prediction and target volumes of nuclei. See table 2 for description.

axis. 3D U-Net model requires sufficiently large number of Z slices ($>= 64$) to extract meaningful high level features in Z-dimension as the input is isotropically downsampled in the encoding path of the U-Net. Therefore, use of 3D translation models often requires upsampling the data in Z, which increases data size and makes training 3D translation model more computationally expensive.

To reduce the computational cost while maintaining the optical resolution of the data and high accuracy of prediction, we evaluated the prediction accuracy as a function of model dimensions for an ordered structure (F-actin) and for a condensed structure (nuclei) in mouse kidney tissue. We evaluated three model architectures to predict fluorescence volumes: slice→slice (2D in short) models that predict 2D fluorescence slices from corresponding 2D label-free slices, stack→slice (2.5D in short) models that predict the central 2D fluorescence slices from a stack of neighbor label-free slices, and stack→stack (3D in short) models that predict fluorescent stacks from label-free stacks. See methods and fig. 3-supplement S1 for the description of the network architecture and training process. We used Pearson correlation coefficient and structural similarity index (SSIM) (47) between predicted fluorescent stacks and target fluorescent stacks from the test fields of view to evaluate the performance of the models (methods). We report these metrics on the test set (table 1,table 2,table 3), which was not used during the training.

Note that we could train 2.5D model with $\sim 3x$ more learnable parameters than 3D model (methods) using lesser computation and memory resources than 3D model. The python code for training our variants of image translation models is available at https://github.com/czbiohub/microDL.

Figure 3A and fig. 3B show XY and XZ slices of a glomerulus and surrounding tissue 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. Glomeruli are complex multi-cellular, three-dimensional structures in kidney that perform filtration (48). The predictions with 2D models show discontinuity artifacts in the structure along the depth (fig. 3C, Video 4), as also observed in prior work (40). The 2.5D model predicts smoother structures along Z dimension and improves the fidelity of F-actin prediction in XY plane, (fig. 3C, table 1), with higher prediction accuracy as the number of z-slices in the 2.5D model input increases. The 3D model further improves the fidelity and continuity prediction along the depth (fig. 3C and Video 4), however the additional gain in accuracy is not as significant as the transition from 2D to 2.5D model. Distribution of SSIM along XY and XZ slice (fig. 3D, table 1) shows the 2.5D model approach the prediction accuracy of the 3D model. In our experiments, training 3D models

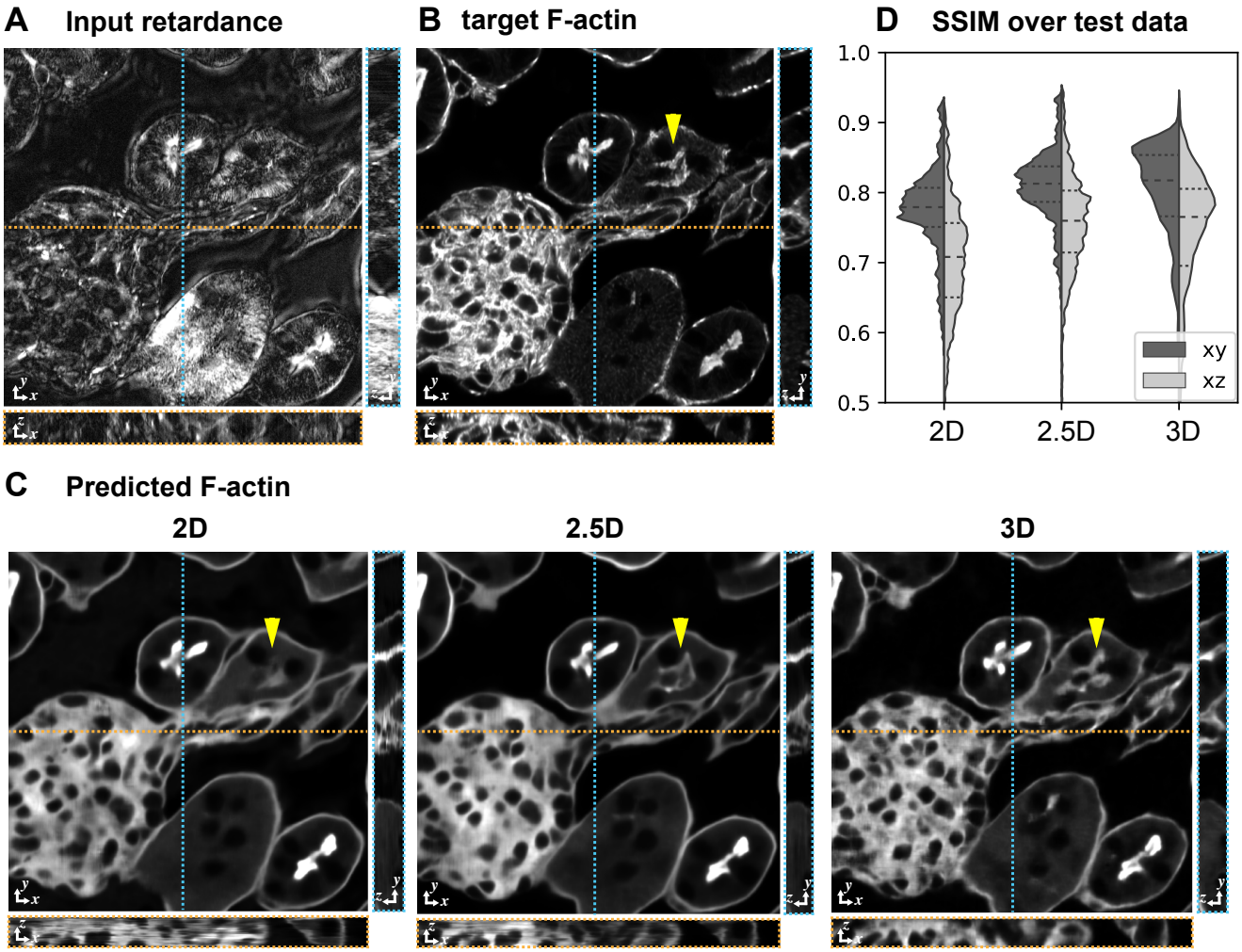


Fig. 3. Accurate 3D prediction of fluorescent structure with 2.5D U-Net: Orthogonal sections (XY - top, XZ - bottom, YZ - right) of 3D volumeric test images of (A) retardance, (B) Experimental fluorescence of F-actin stain (target image), and (C) F-actin predicted from the retardance image using different U-Net architectures. (D) Violin plots of structral-similarty metric (SSIM) between images of predicted and experimental stain in XY and XZ planes. The horizontal dashed lines in the violin plots indicate 25th quartile, median, and 75th quartile of SSIM. The yellow triangle in C highlights a tubule structure, whose prediction can be seen to improve as the model has access to more information along Z. The same field-of-view is shown in fig. 2, Video 1, Video 4, and Video 5.

using ~ 100 training volumes required 3.2 days to converge with a cutting edge GPU, while training a 2D model takes 6 hrs and the 2.5D model with 5 input z-slices takes 2 days.

Predicting structures from multiple label-free contrasts improves accuracy. Considering trade-off between computation speed and model performance, we adopted 2.5D models with 5 input Z-slices to explore how combinations of label-free inputs affect the accuracy of prediction of fluorescent structures.

We found that when multiple label-free measurements are jointly used as inputs, both F-actin and nuclei are predicted with higher fidelity compared to when only a single label-free measurement is used as the input (table 2 and table 3). Figure 4 A-C shows structural differences in the predictions of the same glomerulus as fig. 3. These observations from a representative field of view generalize to entire test set as illustrated by the distribution of SSIM between slices of ground-truth and predictions in fig. 4C. The continuity of prediction along z-axis improves as more label-free contrasts are used for prediction (Video 5). These results indicate that our model leverages information in complementary physical properties to predict target structures. Note that using complementary label-free contrasts boosted the performance of 2.5D models to exceed the performance of 3D single-channel models without significantly increasing the computation cost (compare table 1 and table 2). The prediction accuracy for fine F-actin structures also improves when complementary contrasts are used as input (Video 4 and Video 5).

Interestingly, when only single contrast is provided as the input, model trained on phase images has higher prediction accuracy than the model trained on brightfield images. This is possibly because the phase image has consistent, quantitative contrast along z-axis, while the depth-dependent contrast in brightfield images make the learning task more challenging. This improvement of using phase over brightfield images, however, is not observed when the retardance and orientation images are included in the input as well, possibly because the quantitative contrast in retardance and orientation images provides sufficient information for prediction

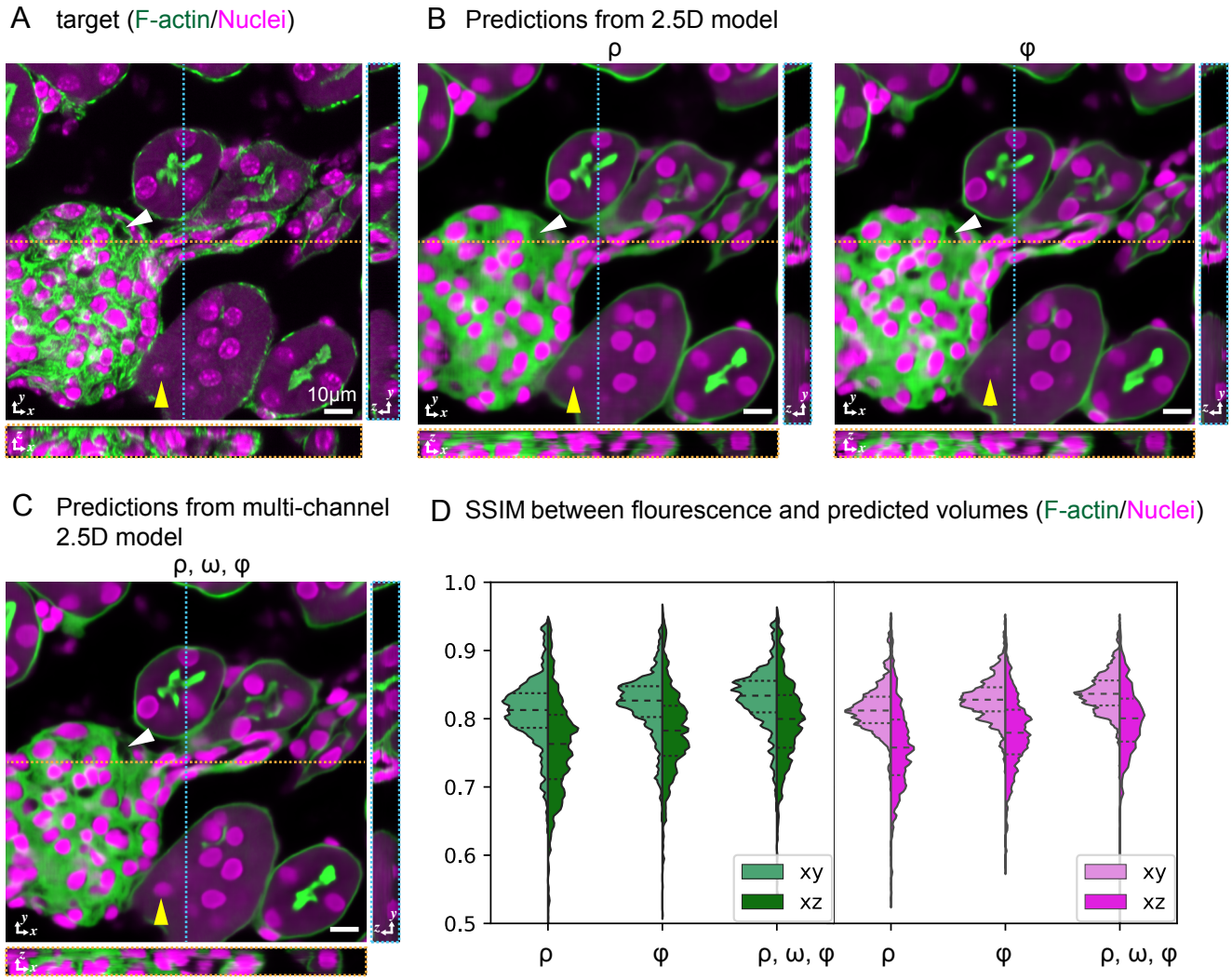


Fig. 4. Prediction accuracy improves with multiple label-free contrasts as inputs: 3D predictions of ordered F-actin and nuclei from different combinations of label-free contrasts using 2.5D U-Net model. **(A)** Fluorescent stain for a field of view from the test set. Phalloidin-labeled F-actin in shown green and DAPI labeled nuclei is shown in magenta. **(B)** F-actin and nuclei distribution predicted with models trained on retardance (ρ) and phase (ϕ) alone are shown. **(C)** F-actin and nuclei distribution predicted with models trained with combined input of retardance, orientation, and phase. **(D)** Violin plots of structral-similarty metric (SSIM) between images of predicted and experimental stain in XY and XZ planes. The horizontal dashed lines in the violin plots indicate 25th quartile, median, and 75th quartile of SSIM. The yellow triangle and white triangle point out structures missing in predicted F-actin and nuclei distributions when only one channel is used as an input, but predicted when all channels are used. Label-free inputs used for prediction are shown in fig. 2 and Video 1.

of nuclei. Note that the finest details of striated F-actin and nucleoli are missing even in the multi-contrast 2.5D model prediction (Video 5). This may be due to limited information content in the input images or due to the model not being trained on sufficient examples of fine features. Even the best model cannot predict the structural details if structures do not appear in the input or are not properly represented in training set.

In conclusion, above results show that 2.5D multi-contrast models predict 3D structures with as high an accuracy as 3D U-Net models, but have multiple practical advantages that allow scaling of the approach. In addition, we show that both ordered and condensed structures can be learned with high accuracy when complementary physical properties are used as inputs.

Mapping axon tracts and their myelination in sections of developing human brain. We next evaluate the possibility of mapping mesoscale connectivity and myelination in a postmortem human fetal brain without label. Studies in prenatal (49) and postnatal (50) human brains suggest that human brains develop and are wired very differently as compared to model systems such as mice. Development of myelination is essential for long-range neural transmission and, consequently, emergence of coordination among different parts of brain. Scalable imaging of myelination and connectivity in developing human brain has the potential to enable studies that address outstanding questions of fundamental importance and clinical relevance. Label-free imaging of white matter is already being pursued in adult human brains (22) with polarized light. Deciphering myelination and connectivity during human development requires higher sensitivity to changes in density and anisotropy, since the con-

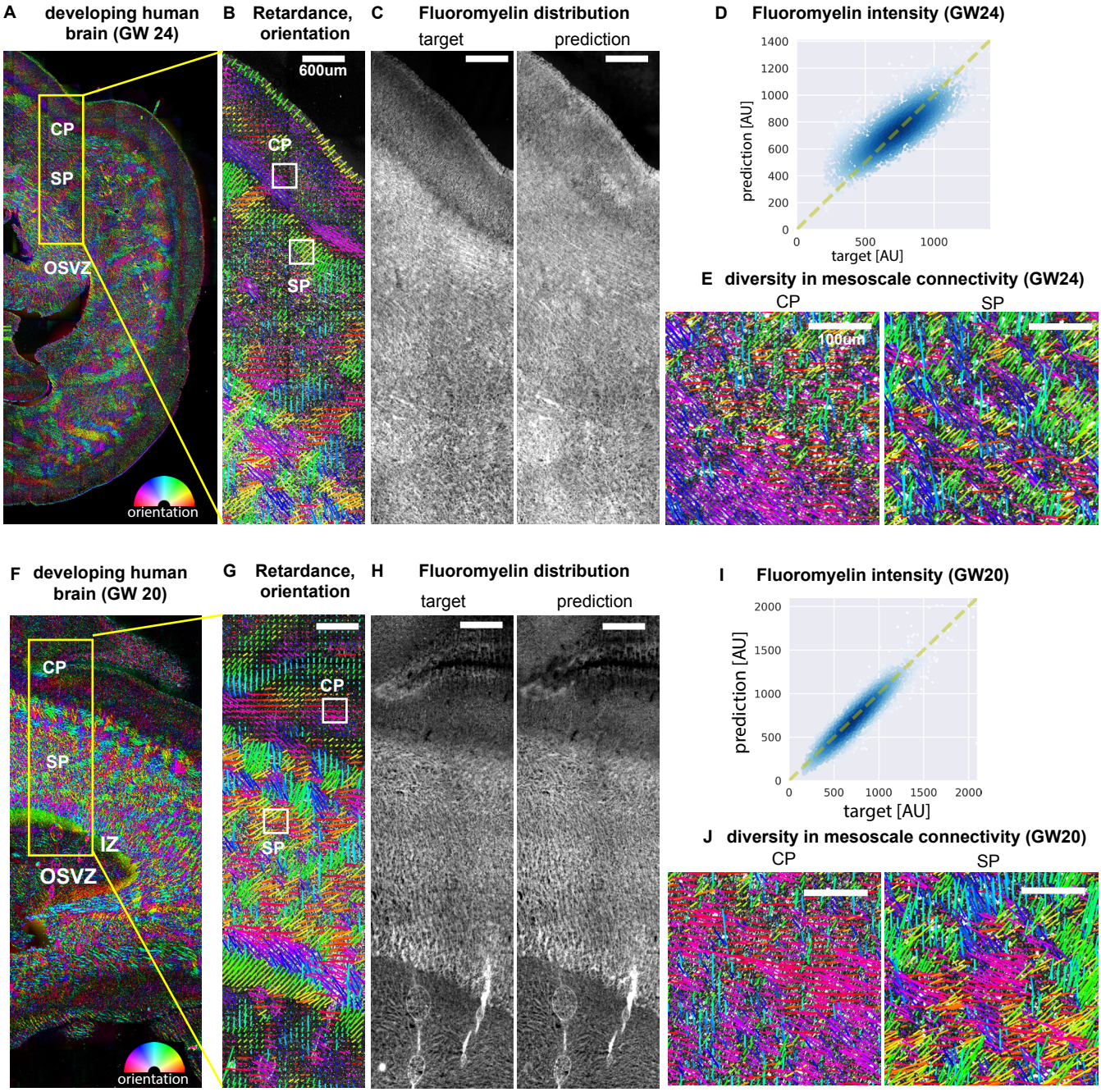


Fig. 5. Label-free mapping of myelination and axon tracts in developing human brain tissue section: (A) Stitched image of retardance and orientation of a gestational week 24 (GW24) brain section from the test set. Different regions of brain can be identified in the label-free images, CP: cortical plate; SP: subplate; OSVZ: outer subventricular zone. The orientation is encoded by color according to legend shown in bottom-right. **(B)** Region of interest over cortical plate and subplate showing retardance and orientation as lines. The orientation is represented by both colors and orientations of the lines, and retardance is represented by lengths of the lines. **(C)** Experimental FluoroMyelin stain (target) and the predicted fluoromyelin stain from the label-free contrasts with the most accurate model identified in table 4. **(D)** Scatter plot of target and predicted intensity. Yellow dashed line indicates the function y=x. **(E)** Insets from **(B)** of retardance and orientation highlight difference in pattern of orientation of tracts in cortical plate and mixed orientation of tracts in subplate. **(F – J)** Same as (A–E), but for GW20 sample. IZ: intermediate zone.

Translation model	Input(s)	Loss	r_{xy}	$SSIM_{xy}$
Slice->Slice (2D)	ϕ	L1	0.73	0.73
	ϕ, ρ, ω_x, ω_y	L1	0.83	0.78
	BF, ρ, ω_x, ω_y	L1	0.81	0.74
Stack-> Slice (2.5D, $z = 5$)	ρ, ω_x, ω_y	L1	0.83	0.76
	BF, ρ, ω_x, ω_y	L1	**0.85**	**0.79**

Table 4. Accuracy of prediction of FluoroMyelin in human brain tissue slices across two developmental points (GW20 and GW24): Above table lists median values of the Pearson correlation (r) and structural similarity index (SSIM) between predictions of image translation models and target fluorescence. We evaluated combinations of retardance (ρ), orientation x (ω_x), orientation y (ω_y), phase (ϕ), and brightfield (BF) as inputs. These metrics are computed over 15% of the fields of view from two GW20 datasets and two GW24 datasets that were not used during training. The 2D models take ~ 4 hours to converge, whereas 2.5D models take ~ 64 hours to converge.

trast is much lower than white matter. Calibration of elliptical polarization states and background correction approaches described here provide sensitive detection of changes in density and anisotropy of brain tissue. Deep learning models that translate this data into quantifiable maps of myelination can help facile analysis of the complex information contained in such data.

We imaged brain sections from two different ages, gestational week (GW) 24 (fig. 5A–E, fig. 5-supplement S1) and GW20 (fig. 5F–J, fig. 5-supplement S2), which correspond to the earliest stages of oligodendrocyte maturation and early myelination in the cerebral cortex. The stitched retardance and orientation images clearly show the orientation of the axon tracts in different regions of the brain that are accessible only with polarization-resolved imaging, with slow axis orientation perpendicular to the direction of the axon tracts, similar to the observations on the mouse brain section (compare with fig. 2-supplement S2). The retardance of the white matter in subplate is higher than cortical plates in both time points, which is consistent with the reduced myelin density in the cortical plate relative to the white matter. Importantly, thanks to the calibration and background correction methods described here, we are able to detect the orientation of axons in cortical plates even though myelin density is low early during development.

In contrast to predictions reported in fig. 3 and fig. 4, here we performed 2D instead of 3D phase reconstruction for human brain slice dataset as the archival tissue was thinner than the depth of field of the low magnification objective (10X) we used for imaging large areas. The stitched 2D reconstructed phase shows tissue morphology and axon tracts with significantly higher contrast when compared to the brightfield image (fig. 5-supplement S1A and fig. 5-supplement S2A). This is because 2D phase reconstruction is more sensitive to density variation integrated along the z-axis compared to the brightfield imaging.

Interestingly, the pattern of orientation in cortical plate at GW24 (fig. 5E) suggests layered arrangement of neurons, while the pattern of orientation in subplate at GW24 appears criss-crossing arrangement of neurons. These differences between orientation pattern in cortical plate and subplate are less pronounced in data acquired from GW20 tissue (fig. 5I) . To our knowledge, above data is the first report of label-free imaging of myelination and connectivity in pre-natal brain tissue.

Predicting myelination across different ages. We explored how information in these label-free signatures can be used to predict myelination, which serves as a proxy for development of white matter. We applied our image translation approach to the task of predicting myelination in the developing human brain across different ages. Myelination is visualized using lipophilic organic dye FluoroMyelin that preferentially binds to myelin (51).

To test how well our models generalize to different brain sections not seen by the model during the training stage, we ran model inference on images of a GW24 and GW20 sections that were not included in training, validation, or test set (results shown in fig. 5 and its supplements). The models were able to predict the axon tracts and myelination in these sections well, with increasing accuracy as we included more label-free channels as the input.

We trained multi-contrast 2D and 2.5D models with different combinations of label-free input contrasts and FluoroMyelin as the target to predict. To improve the model accuracy and generalize the model prediction to different development ages and different types of sections of the brain, we pooled imaging datasets from GW20 and GW24, with 2 different brain sections for each age. The pooled dataset was then split into training, validation, and test set. Similar to the observations on the mouse kidney tissue dataset, 2.5D model with brightfield, retardance, and orientation as the input has the highest SSIM and Pearson correlation over the test set (table 4). We compare the linearity between predicted and target FluoroMyelin with scatter plots (fig. 5D and fig. 5I) and find that our model reliably translates the complex information in density and anisotropy to expression of FluoroMyelin.

Notably, the 2D model with phase, retardance, and orientation as the input has scores close to the best 2.5D model but the training only takes 3.7 hrs to converge, while the best 2.5D model takes 64.7 hrs to converge (table 4). This is most likely because the 2D phase reconstruction captures the density variation encoded in the brightfield z-stack that is informative for model to predict axon tracts accurately.

Rescue of inconsistent label. Another useful feature of our image translation approach is that the predicted stains are robust to experimental variation in fluorescent stains. For examples, we found that antifade chemicals in the mounting media quenched the fluorescence of FluoroMyelin and created dark patches in FluoroMyelin (fig. 6) images. However, the quenching of dye had no impact on the physical

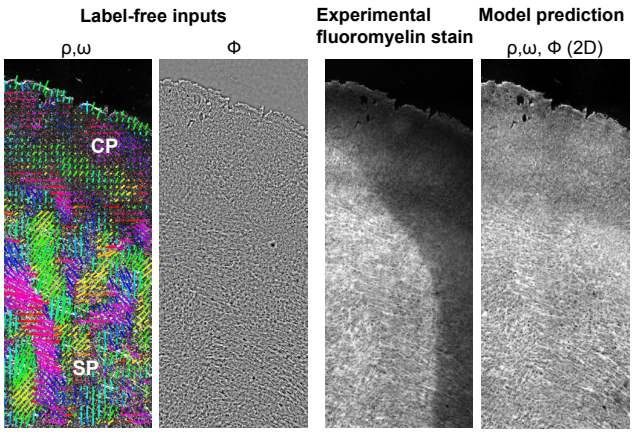


Fig. 6. Model predicted stain is robust to experimental variation in staining: ROI of GW24 brain section shows artifacts in FluoroMyelin stain due to quenching of fluorescence by mounting medium. FluoroMyelin stain predicted by the model does not show the artifacts observed in experimental FluoroMyelin stain. (Left to right) retardance, orientation, and phase input of the 2D model; experimental FluoroMyelin stain; predicted stain from multi-contrast 2D model. CP: cortical plate; SP: subplate.

properties we measure in the label-free channels. Therefore, the model trained on images without artifacts from fluorescence quenching correctly predicted the staining pattern even where experimental stain failed. This robustness is particularly valuable for precious tissue specimens such as archival prenatal human brain tissue in which obtaining multiple brain sections is not feasible.

Importantly, proper data normalization is essential for predicting the intensity correctly across different fields-of-view in large stitched images. Due to the batch variation of the staining and imaging process, we found that normalizing the images over whole dataset gives the most accurate intensity prediction (fig. 6-supplement S1).

Discussion

We have reported an innovative approach for label-free measurement of density and anisotropy from 3D polarization-resolved acquisition. We discuss how we elected to balance the trade-offs and the future directions of research.

The phase and degree of polarization information have always been present in the polarization-resolved acquisition, but can now be reconstructed thanks to our accurate forward models and corresponding reconstruction algorithms. As compared to joint imaging of density and anisotropy with orientation-independent differential interference contrast and orientation-independent PolScope (33), our method uses only one polarization modulator that simplifies calibration. However, our simpler light path does not achieve as high a depth sectioning as differential interference contrast. Our method provides diffraction-limited imaging of density and ansiotropy in live cells, as evident from the 3D movie of organelles (fig. 2, Video 2, Video 3). Our open-source python software is free to use for non-profit research. We anticipate the modularity of the optical path and the availability of software to facilitate adoption.

We note that recovery of phase from propagation of light reports relative phase in the field of view, and not the absolute phase. Measurement of absolute phase requires use of interference with a reference beam (8, 9). However, how biological components evolve and interact requires knowledge of local quantitative distribution of phase and not absolute phase. We employ partially coherent illumination, i.e., simultaneous illumination from multiple angles. Partially coherent illumination improves spatial resolution, depth sectioning, and robustness to imperfections in the light path away from the focal plane.

We also note that, similar to all present polarization-resolved imaging systems (22, 29), our approach reports projection of the anisotropy onto the focal plane. The anisotropic structure, such as an axon bundle, appears isotropic to the imaging system when it is aligned along the optical axis of the imaging path. Recovering true 3D anisotropy along with 3D density using forward models that account for diffraction effects in the propagation of polarized light is the topic of our future research.

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 (52). 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 calculus, since OCT is a coherent interferometer and intensity recorded in individual speckle is fully polarized (53). However, PS-OCT practitioners employ degree of polarization uniformity (53) over several speckles to analyze depolarization due to multiple scattering.

We have also reported novel deep learning models for efficient analysis of multi-dimensional 3D data we acquire. In contrast to other work on image translation that demonstrated only 2D prediction (35–37), our 2.5D architecture achieves 3D prediction with apparently similar or superior accuracy as the 3D prediction reported in (40) (Pearson correlation coefficient in 3D for nuclei prediction from brightfield images: 0.87 v.s. 0.7 reported in (40)), while being computationally more efficient. We note that Pearson correlation coefficient is affected by both the accuracy of the prediction as well as the noise in the target images. A more direct comparison of model performances on the same dataset would be useful in the future. Also, 2.5D network can be applied to image data that only has a few z-slices without up- or down-sampling the data, making it useful for microscopy data that often have asymmetric XY and Z dimension. Moreover, even though we focus on image translation in this work, the same 2.5D network can be used for 3D segmentation. 3D segmentation using the 2.5D network bears additional advantages over 3D network because annotation can be done only on a subset of slices rather than the whole 3D volume, which would save significant amount of annotation cost.

In comparison to Christiansen *et al.*'s 2D translation model (35) where the image translation was formulated as a

pixel-wise classification task of 8-bit classes, our 2.5D translation model formulates the image translation as a regression task that allows prediction of much larger dynamic range of gray levels. Also, our modeling strategy to train one model to predict only one target allowed us to use significantly smaller models that can fit into GPU memory for faster training.

We systematically evaluated how the dimensions and input channels affect the prediction accuracy. We find that higher prediction accuracy is achieved by combining multiple label-free contrasts, in contrast to previous work that predicts fluorescence images from single label-free contrast. Additionally, we demonstrated prediction of fluorescence images of tissue, while previous work has only reported prediction of fluorescence images of cultured cells or brightfield images of histochemically-stained tissue (35–37, 40). Image translation results reported in (35–37) were limited to 2D structures, likely because they use even more complex architectures than 3D U-Net model reported in (40).

A common shortfall of machine learning approaches is that they tend not to generalize well. We have shown that our data normalization and training process leads to models of myelination that generalize across developmental time 5. In contrast to the prediction by physical models, we note that the errors or artifacts in the prediction by machine learning models are usually heterogeneous and highly dependant on the quality of training data and the input. Therefore, prediction errors made by machine learning models can be difficult to recognize in absence of ground truth. For image translation in biological specimen with unknown structures, such as prenatal human brain, it would be crucial to predict not just the output, but also the confidence interval.

Several methods for tracing connectivity in the mouse brain at mesoscale have been developed (54), but they have not yet been scaled to human brain. The volume of fetal human brain during third trimester $(10^5\mathrm{mm}^3 - 4\times10^5\mathrm{mm}^3)$ is 3 orders of magnitude larger than the volume of an adult mouse brain ($\sim 5\times10^2\mathrm{mm}^3$). Our data shows that quantitative polarized light microscopy can provide label-free measurement of myelination and axonal tract orientation at high density with $\sim 1.5\mu\mathrm{m}$ resolution over centimeter spatial extents, in both primary human brain sections and mouse brain sections. Our results comparing human and mouse brain sections highlight the high degree of fasciculation of axonal bundles in the mouse brain compared with human. We anticipate that analysis of connectivity across large brain slices will be possible by adapting tractography algorithms developed for diffusion weighted-MRI measurements to our vectorial data. Scalable mapping of myelination and connectivity from archival tissue of postmortem brain may enable high-throughput mapping of mesoscale connectivity in large brains, leading to new insights into how information processing streams are organized across the developing and adult human brain.

Conclusion

In summary, we report an integrative computational imaging approach that combines label-free imaging, reconstruction of optical properties using physical models, and prediction of biological structures with data-driven deep neural network models. Our Stokes-formalism based reconstruction algorithms (https://github.com/mehta-lab/reconstruct-order) and computationally efficient U-Net variants (https://github.com/czbiohub/microdl) allow facile imaging and interpretation of label-free signatures across biological scales. Our approach enables simultaneous measurement of phase, retardance, orientation, and degree of polarization contrasts with diffraction-limited spatial resolution. These contrasts report variations in density, anisotropy, and scattering of the specimen. We demonstrated visualization of diverse biological structures: glomeruli and tubules in mouse kidney tissue, multiple organelles in cells, and axon tracts and myelination in mouse and human brain slices. We demonstrated multi-contrast 2.5D U-Net model for accurate and computationally efficient prediction of biological structures from label-free contrasts. Fluorescence predicted from label-free images is robust to inherent variability in labeling. We demonstrated accurate prediction of ordered F-actin and nuclei in heterogeneous structures within tissue, as well as myelination in prenatal human brain tissue. We anticipate that our approach will enable scalable analysis of architectural order that underpins healthy and disease states of cells and tissues.

Methods

Model of image formation. We describe dependence of the polarization resolved images on the specimen properties using Stokes formalism (55, Ch.15). This representation allows us to accurately measure the polarization sensitive contrast at the image plane for every focal plane. First, we retrieve the coefficients of the specimen's Mueller matrix that report linear birefringence, transmission, and depolarization. For brevity, we call them 'Mueller coefficients' of the specimen in this paper. Mueller coefficients are recovered from the polarization-resolved intensities and an instrument matrix that captures how Mueller coefficeints are related to intensities recorded by the microscope. Assuming that the specimen is mostly transparent, more specifically satisfies the first Born approximation (41), we reconstruct specimen phase, retardance, slow axis, and degree of polarization stacks from stacks of Mueller coefficients (fig. 1). The assumption of 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 is represented directly by a calibrated instrument matrix.

For sensitive detection of birefringence, it is advantageous to suppress isotropic background by illuminating the specimen with elliptically polarized light and image with circular state of opposite handedness (23). For experiments re-

ported in this paper, we acquired data by illuminating the specimen sequentially with right-handed circular and elliptical states and analyzed the transmitted light with left-handed circular state.

Forward model: specimen properties → Mueller coefficients. We assume that the specimen is transparent, and therefore, modeled by net retardance ρ, orientation of the slow axis ω, transmission t, and depolarization p. The Mueller matrix of the specimen is then given by,

$$\mathbf{M}_{\mathrm{sm}} = \begin{bmatrix} t & 0 & 0 & 0 \\ 0 & tp & 0 & 0 \\ 0 & 0 & tp & 0 \\ 0 & 0 & 0 & tp \end{bmatrix} \mathbf{M}_r = \begin{bmatrix} m_0 & 0 & 0 & 0 \\ 0 & * & * & m_1 \\ 0 & * & * & m_2 \\ 0 & -m_1 & -m_2 & m_3 \end{bmatrix} \tag{1}$$

where $\mathbf{M}_r$ is the Mueller matrix of a linear retarder that can be found in (55, Ch.14), and $*$ signs denote irrelevant entries that cannot be retrieved under our experiment scheme. The relevant entries that are retrievable can be expressed as a vector of Mueller coefficients, which is

$$\mathbf{m} = \begin{bmatrix} m_0 \\ m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} t \\ tp\sin 2\omega \sin\rho \\ -tp\cos 2\omega \sin\rho \\ tp\cos\rho \end{bmatrix} \tag{2}$$

This vector is coincidentally the Stokes vector when right-handed circularly polarized light passing through the specimen. The aim of the measurement we describe in the following paragraphs is to accurately measure these Mueller coefficients at each point in the image plane of the microscope by illuminating the specimen and detecting the scattered light with mutually independent polarization states. Once these Mueller coefficients map has been acquired with high accuracy, the specimen properties can be retrieved from above set of equations.

Forward model: Mueller coefficients → intensities. To acquire the above Mueller coefficients, we illuminate the specimen with a series of right-handed circularly and elliptically polarized light (23). The Stokes vectors of our sequential illumination states are given by,

$$\mathbf{S}_i = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}_{i=\mathrm{RCP}}, \begin{bmatrix} 1 \\ \sin\chi \\ 0 \\ \cos\chi \end{bmatrix}_{i=0}, \begin{bmatrix} 1 \\ -\sin\chi \\ 0 \\ \cos\chi \end{bmatrix}_{i=45}, \begin{bmatrix} 1 \\ 0 \\ \sin\chi \\ \cos\chi \end{bmatrix}_{i=90}, \begin{bmatrix} 1 \\ 0 \\ -\sin\chi \\ \cos\chi \end{bmatrix}_{i=135} \tag{3}$$

where χ is the compensatory retardance controlled by the LC that determines the ellipticity of the four elliptical polarization states.

After our controlled polarized illumination has passed through our specimen, we detected the light with the left-handed circular state by having a left-handed circular analyzer in front of our sensor. We express the Stokes vector before the sensor as

$$\mathbf{S}_{\mathrm{sensor},i} = \mathbf{M}_{\mathrm{LCA}}\mathbf{M}_{\mathrm{sm}}\mathbf{S}_i, \tag{4}$$

where $i = \{\mathrm{RCP}, 0, 45, 90, 135\}$ depending on the illumination states, and $\mathbf{M}_{\mathrm{LCA}}$ is the Muller matrix of a left-handed circular analyzer (55, Ch.14). The detected intensity images are the first component of Stokes vector at the sensor under different illuminations ($I_i = [\mathbf{S}_{\mathrm{sensor},i}]_0$). Stacking the measured intensity images to form a vector

$$\mathbf{I} = \begin{bmatrix} I_{\mathrm{RCP}} \\ I_0 \\ I_{45} \\ I_{90} \\ I_{135} \end{bmatrix}, \tag{5}$$

we can link the relationship between the measured intensity and the specimen vector through an 'instrument matrix' $\mathbf{A}$ as

$$\mathbf{I} = \mathbf{A}\mathbf{m}, \tag{6}$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 1 & \sin\chi & 0 & -\cos\chi \\ 1 & 0 & \sin\chi & -\cos\chi \\ 1 & -\sin\chi & 0 & -\cos\chi \\ 1 & 0 & -\sin\chi & -\cos\chi \end{bmatrix}. \tag{7}$$

Each row of the instrument matrix is determined by the interaction between various illumination polarization states and the specimen's properties. Any polarization-resolved measurement scheme can be characterized by an instrument matrix that transforms specimen's polarization property to the measured intensities. Calibration of the polarization imaging system is then done through calibrating this instrument matrix.

Computation of Mueller coefficients 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. 6),

$$\mathbf{m} = \mathbf{A}^{-1}\mathbf{I}, \tag{8}$$

Computation of background corrected specimen properties. We retrieved the vector of Mueller coefficients, $\mathbf{m}$, by solving Eq. 8. Slight strain or misalignment in the optical components or the specimen chamber can lead to background that masks out contrast from the specimen. The background typically varies slowly across the field of view and can introduce spurious correlations in the measurement. It is crucial to correct the vector of Mueller coefficients 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 ($i = \mathrm{sm}$) and background ($i = \mathrm{bg}$) vectors of Mueller coefficients as follows,

$$
\begin{aligned}
\overline{m_1}^i &= m_1^i/m_3^i, \\
\overline{m_2}^i &= m_2^i/m_3^i, \\
\mathrm{DOP}^i &= \frac{\sqrt{(m_1^i)^2+(m_2^i)^2+(m_3^i)^2}}{m_0^i},
\end{aligned} \tag{9}
$$

We then reconstructed the background corrected properties of the specimen: brightfield (BF), retardance (ρ), slow axis (ω), and degree of polarization (DOP) from the transformed specimen and background vectors of Mueller coefficients $\overline{\mathbf{m}}^{\mathrm{sm}}$ and $\overline{\mathbf{m}}^{\mathrm{bg}}$ using the following equations:

$$\overline{m_1} = \overline{m_1}^{\mathrm{sm}} - \overline{m_1}^{\mathrm{bg}} \tag{10}$$

$$\overline{m_2} = \overline{m_2}^{\mathrm{sm}} - \overline{m_2}^{\mathrm{bg}} \tag{11}$$

$$\mathrm{BF} = m_0^{\mathrm{sm}}/m_0^{\mathrm{bg}} \tag{12}$$

$$\rho = \arctan 2\left(\sqrt{\overline{m_1}^2+\overline{m_2}^2}\right) \tag{13}$$

$$\omega = \frac{1}{2}\arctan 2\left(\frac{\overline{m_1}}{-\overline{m_2}}\right) \tag{14}$$

$$\mathrm{DOP} = \mathrm{DOP}^{\mathrm{sm}}/\mathrm{DOP}^{\mathrm{bg}} \tag{15}$$

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

Phase reconstruction. As seen from Eq. 2, the first component in the vector of Mueller coefficients, m_0, is equal to the total transmitted intensity of electric field in the focal plane. Assuming a non-absorbing specimen, the intensity variations encode the average phase in the focal plane via the transport of intensity (TIE) equation (5). In the following, we leverage weak object transfer function (WOTF) formalism to retrieve 2D and 3D phase from this average TIE phase contrast and describe the corresponding inverse algorithm.

Forward model for phase reconstruction. The linear relationship between the 3D phase and the through focus brightfield intensity was established in (14) with Born approximation and weak object approximation. In our context, we reformulated as

$$m_0(\mathbf{r}) = m_{0,\mathrm{dc}} + \phi(\mathbf{r}) \otimes_{\mathbf{r}} h_\phi(\mathbf{r}) + \mu(\mathbf{r}) \otimes_{\mathbf{r}} h_\mu(\mathbf{r}), \tag{16}$$

where $\mathbf{r} = (\mathbf{r}_\perp, z) = (x, y, z)$ is the 3D spatial coordinate vector, $m_{0,\mathrm{dc}}$ is the constant background of m_0 component, $\otimes_{\mathbf{r}}$ denotes convolution operation over $\mathbf{r}$ coordinate, ϕ refers to phase, μ refers to absorption, $h_\phi(\mathbf{r})$ is the phase point spread function (PSF), and $h_\mu(\mathbf{r})$ is the absorption PSF. Rigorously speaking, ϕ and μ are the real and imaginary part of the scattering potential scaled by $\Delta z/2k$, where Δz is the axial pixel size of the experiment and k is the wavenumber of the incident light. When the refractive index of the specimen and that of the environment are close, the real and imaginary scaled scattering potential reduces to two real quantity, phase and absorption.

When specimen's thickness is larger than the depth of field of the microscope (usually in experiments with high NA objective), the brightfield intensity stack contains 3D information of specimen's phase and absorption. Without making more assumption or taking more data, this problem is ill-posed because we are solving two unknowns from one measurement. Assuming the absorption of the specimen is negligible (15, 18, 19), which generally applies to transparent biological specimens, we turn this problem into a linear deconvolution problem, where 3D phase is retrieved.

When specimen's thickness is smaller than the depth of field of the microscope (usually in experiments with low NA objective), the whole 3D intensity stack is coming from merely one effective 2D absorption and phase layer of specimen. We rewrite Eq. 16 as

$$
\begin{aligned}
m_0(\mathbf{r}) = m_{0,\mathrm{dc}} \quad &+ \quad \phi(\mathbf{r}_\perp) \otimes_{\mathbf{r}_\perp} h_\phi(\mathbf{r}_\perp, z) \\
&+ \quad \mu(\mathbf{r}_\perp) \otimes_{\mathbf{r}_\perp} h_\mu(\mathbf{r}_\perp, z).
\end{aligned} \tag{17}
$$

In this situation, we have multiple 2D defocused measurements to solve for one layer of 2D absorption and phase of the specimen.

Inverse problem for phase reconstruction. With the linear relationship between the first component of the Mueller coefficients vector and the phase, we then formulated the inverse problem to retrieve 2D and 3D phase of the specimen.

When we recognize the specimen as a 3D specimen, we then use Eq. 16 and drop the absorption term to estimate specimen's 3D phase through the following optimization algorithm:

$$\min_{\phi(\mathbf{r})} \sum_{\mathbf{r}} \left| m_0'(\mathbf{r}) - \phi(\mathbf{r}) \otimes_{\mathbf{r}} h_\phi(\mathbf{r}) \right|^2 + \tau_\phi \mathrm{Reg}(\phi(\mathbf{r})), \tag{18}$$

where $m_0'(\mathbf{r}) = m_0(\mathbf{r}) - m_{0,\mathrm{dc}}$, τ_ϕ is the regularization parameter for applying different degree of denoising effect, and the regularization term depending on the choice of either Tikhonov or anisotropic total variation (TV) denoiser is expressed

$$\mathrm{Reg}(\phi(\mathbf{r})) = \begin{cases} \sum_{\mathbf{r}} |\phi(\mathbf{r})|^2, & \text{Tikhonov} \\ \sum_{\mathbf{r}} \sum_{i=x,y,z} |\partial_i \phi(\mathbf{r})|, & \text{TV} \end{cases}$$

When using Tikhonov regularization, this optimization problem has an analytic solution as described by (15, 18, 19). As for TV regularization, we adopted alternating minimization algorithm proposed in (56) to solve the problem.

If we consider the specimen as a 2D specimen, we then turn Eq. 17 into the following optimization problem:

$$
\begin{aligned}
\min_{\phi,\mu(\mathbf{r}_\perp)} \sum_{\mathbf{r}} \Big| m_0'(\mathbf{r}) - \phi(\mathbf{r}_\perp) \otimes_{\mathbf{r}_\perp} h_\phi(\mathbf{r}_\perp, z) - \mu(\mathbf{r}_\perp) \otimes_{\mathbf{r}_\perp} \\
h_\mu(\mathbf{r}_\perp, z) \Big|^2 + \tau_\phi \mathrm{Reg}(\phi(\mathbf{r}_\perp)) + \tau_\mu \mathrm{Reg}(\mu(\mathbf{r}_\perp)),
\end{aligned} \tag{19}
$$

where we have an extra regularization parameter τ_μ here for the absorption. When Tikhonov regularization is selected, the analytic solution similar to the one described in (57) is adopted.

When the signal to noise ratio of the brightfield stack is high, Tikhonov regularization gives satisfactory reconstruction in a single step with computation time proportional to the size of the image stack. However, when the noise is high, Tikhonov regularization can lead to high- to medium-frequency artifacts. Using iterative TV denoising algorithm, we can trade-off reconstruction speed with robustness to noise.

Specimen preparation. Mouse kidney tissue slices (Thermo-Fisher Scientific) and mouse brain slices 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 nuclei was labeled with DAPI.

Human prenatal brain samples were fixed with 4% paraformaldehye in phosphate-buffered solution (PBS) overnight, then rinsed with PBS, dehydrated in 30% sucrose/OCT compound (Agar Scientific) at 4C overnight, then frozen in OCT at -80 °C. Frozen samples were sectioned at 12 μm and mounted on microscope slides. Sections were stained directly with red FluoroMyelin (Thermo-Fisher Scientific, 1:300 in PBS) for 20 minutes at room temperature, rinsed 3 times with PBS for 10 minutes each, then mounted with ProLong Gold antifade (Invitrogen) with a coverslip.

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.

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 (nuclei) at 1.6 mW, 100 ms exposure for 561 nm channel (F-actin) at 2.8 mW. The mouse brain slice were also imaged using 30 ms exposure for 5 polarization channels. U2OS cells were imaged using 50 ms exposure for 5 polarization channels. For training the neural network, we acquired 160 non-overlapping $2048 \times 2048 \times 45$ z-stacks of the mouse kidney tissue slice with Nyquist sampled voxel size $103\text{nm} \times 103\text{nm} \times 250\text{nm}$. Human brain sections were imaged with a 10X objective and 0.2 NA condenser with a 200 ms exposure for polarization channels, 250 ms exposure for 568 channel (FluoroMyelin). The full brain sections were imaged, approximately 200 images depending on the size of the section, with 5 z positions at each location.

Data preprocessing for model training. The images were flat-field corrected. For training 3D models, the image volumes were upsampled along Z to match the pixel size in XY using linear interpolation. The images were tiled into 256×256 patches with a 50% overlap between patches for 2D and 2.5D models. The volumes were tiled into $128 \times 128 \times 96$ patches for 3D models with a 25% overlap along XYZ. Tiles that had sufficient fluorescence foreground (2D and 2.5D: 20%, 3D: 50%) were used for training. Foreground was quantified the fraction of the combined masks obtained from Otsu thresholding of the nuclei and F-actin images in the case of mouse kidney tissue sections, and FluoroMyelin in the human brain sections.

Proper data normalization is essential for predicting the intensity correctly across different fields-of-views. We found the common normalization scheme where each image is normalized by its mean and standard deviation does not produce correct intensity prediction (fig. 5 - figure supplement S1). We normalized the images on the per dataset basis to correct the batch variation in the staining and imaging process across different datasets. To balance contributions from different channels during training of multi-contrast models, each channel needs to be scaled to similar range. We used the median and inter-quartile range (range defined by 25% and 75% quantiles) of each channel to normalized the channel because standard deviation underestimates the spread out of the distribution of highly correlated data such as pixels in images.

Neural network architecture. We experimented with 2D, 2.5D and 3D versions of U-Net models fig. 3-supplement S1. 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 (45, 46). 2×2 downsampling is applied with 2×2 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. Feature maps output by every level of encoding path were concatenated to feature maps in the decoding path at corresponding levels. 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, while the 3D U-Net consists of four layers with 16, 32, 64 and 128 filters each due to its higher memory requirement. The 2D and 3D versions use convolution filters of size of 3×3 and $3\times3\times3$ 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 similar architecture as the 2D U-Net with following differences:

1. The 3D features maps are converted into 2D using skip connections that consist of a N×1×1 valid convolution.

2. Convolution filters in the encoding path are N×3×3, where N = 3,5,7 is the number of slices in the input.

3. In the encoding path, the feature maps are downsampled across blocks using N×2×2 average pooling.

4. In 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.

The 2D , 2.5D, 3D network with single channel input consisted of 2.0 M, 4.8M, 1.5M learnable parameters, respectively.

Model training and inference. We randomly split the images in groups of 70%, 15%, and 15% for training, validation and test. The split are kept consistent across all model training to make the results comparable. All models are trained with Adam optimizer, mean absolute error (MAE) loss function, and a cyclic learning rate scheduler with a min and max learning rate of 5×10^{-5} and 6×10^{-3} respectively. The 2D, 2.5D, 3D network were trained on mini-batches of size 64, 16, and 4 to accommodate the memory requirements of each 3D model. Models were trained until the validation loss does not decrease for 20 epochs. The model with minimal validation loss was saved. Single channel 2D models converged in 6 hours, 2.5D model converged in 47 hours and the 3D model converged in 76 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 linear blending 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 coefficient between a target image T and a prediction image P is defined as

$$r(T,P) = \frac{\sigma_{TP}}{\sigma_T \sigma_P} \tag{20}$$

where σ_{TP} is the covariance of T and P, and σ_T and σ_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,

$$\mathrm{SSIM}(t,p) = \frac{(2\mu_t\mu_p + c_1)(2\sigma_{tp} + c_2)}{(\mu_t^2 + \mu_p^2 + c_1)(\sigma_t^2 + \sigma_p^2 + c_2)} \tag{21}$$

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 μ and σ^2 respectively, and the covariance between t and p is denoted σ_{tp}. We use $N = 7$. The total SSIM score is the mean score calculated across all windows, $\mathrm{SSIM}(T,P) = \frac{1}{M}\sum \mathrm{SSIM}(t,p)$ 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.

Importantly, it is essential to scale the the model prediction back to the original range before normalization for correct calculation of target-prediction SSIM. This is because unlike Pearson correlation coefficient, SSIM is not a scale-independent metrics.

ACKNOWLEDGEMENTS

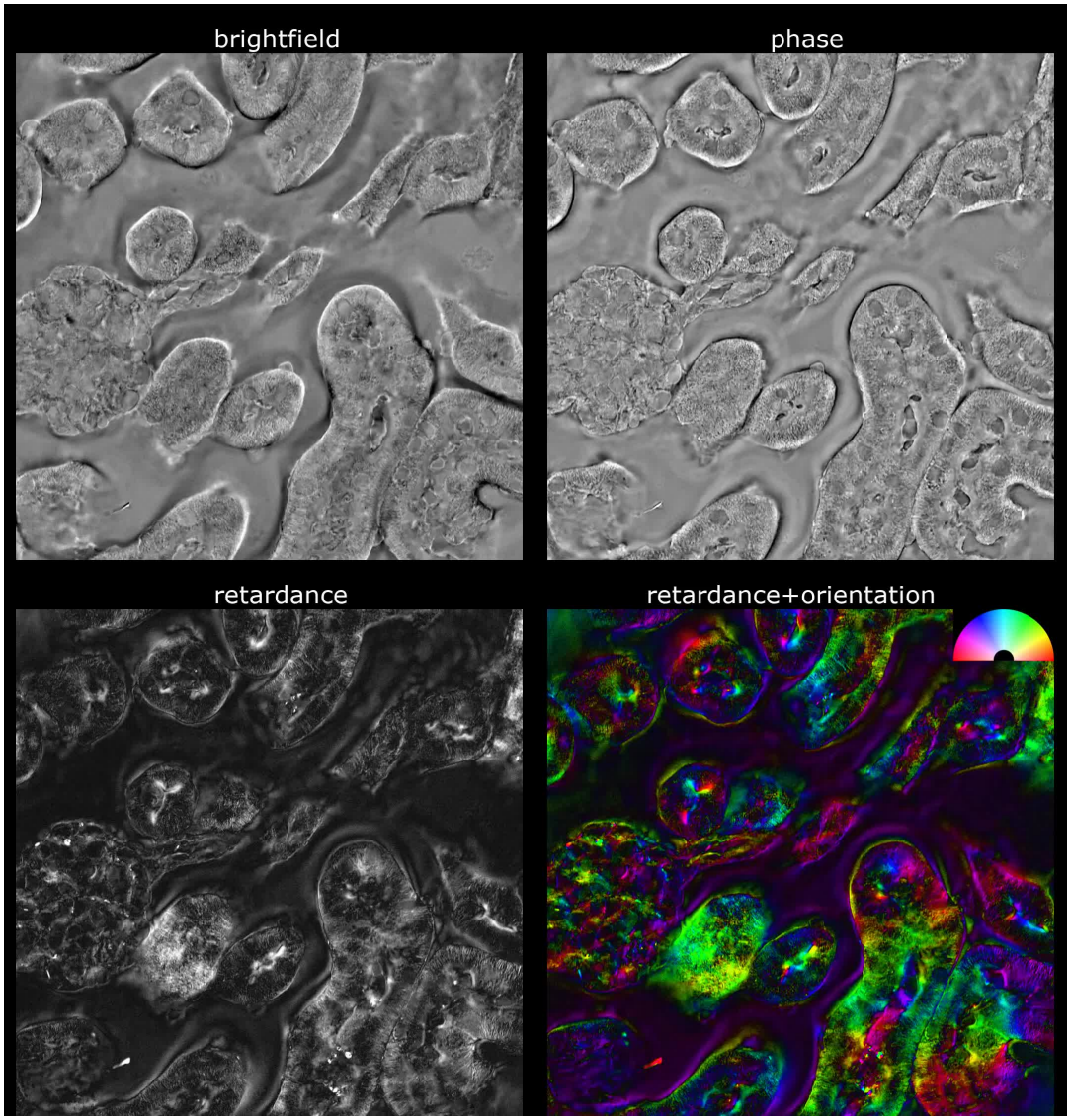
We thank Spyros Dermanis (CZ Biohub), Bing Wu (CZ Biohub), May Han (Stanford), and Ezzat Hashemi (Stanford) for providing the mouse brain slice used for acquiring data shown in Fig. 1. We thank Greg Huber, Loic Royer, Joshua Batson, Jim Karkanias, Joe DeRisi, and Steve Quake from the Chan Zuckerberg Biohub for numerous discussions. We also thank Eva Dyer from Georgia Tech for discussions about applications of the 2.5D models. This research was supported by the Chan Zuckerberg Biohub.

Bibliography

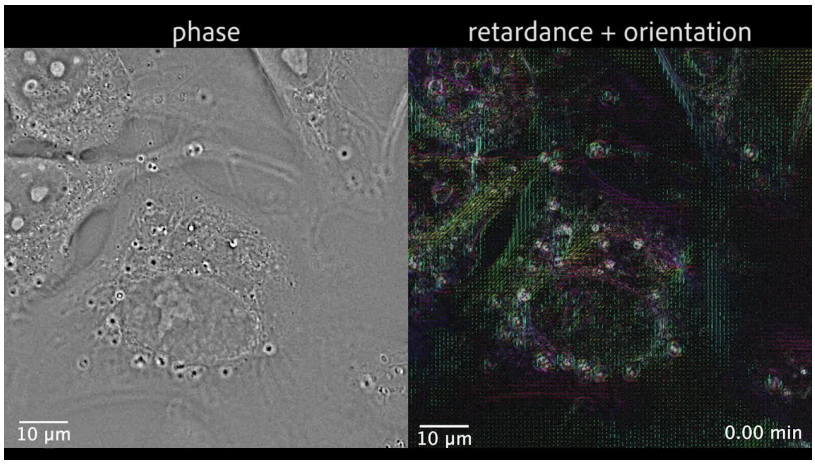
1. A. M. Valm, S. Cohen, W. R. Legant, J. Melunis, U. Hershberg, E. Wait, A. R. Cohen, M. W. Davidson, E. Betzig, and J. Lippincott-Schwartz, "Applying systems-level spectral imaging and analysis to reveal the organelle interactome," *Nature*, vol. 546, no. 7656, pp. 162–167, Jun. 2017.
2. F. Zernike, "How I Discovered Phase Contrast," *Science*, vol. 121, no. 3141, pp. 345–349, 1955.
3. W. J. Schmidt, *Die Bausteine des Tierkörpers in polarisiertem Lichte*. Bonn: F. Cohen, 1924, oCLC: 173148722.
4. S. Inoué, "Polarization optical studies of the mitotic spindle," *Chromosoma*, vol. 5, no. 1, pp. 487–500, Dec. 1953.
5. N. Streibl, "Phase imaging by the transport equation of intensity," *Optics communications*, vol. 49, no. 1, pp. 6–10, 1984.
6. L. Waller, L. Tian, and G. Barbastathis, "Transport of intensity phase-amplitude imaging with higher order intensity derivatives," *Optics express*, vol. 18, no. 12, pp. 12 552–12 561, 2010.
7. Y. Park, C. Depeursinge, and G. Popescu, "Quantitative phase imaging in biomedicine," *Nature Photonics*, vol. 12, no. 10, p. 578, Oct. 2018.
8. D. Kim, S. Lee, M. Lee, J. Oh, S.-A. Yang, and Y. Park, "Holotomography: Refractive index as an intrinsic imaging contrast for 3-D label-free live cell imaging," *bioRxiv*, p. 106328, Nov. 2018.
9. G. Popescu, T. Ikeda, R. R. Dasari, and M. S. Feld, "Diffraction phase microscopy for quantifying cell structure and dynamics," *Optics Letters*, vol. 31, no. 6, pp. 775–777, Mar. 2006.
10. S. B. Mehta and C. J. R. Sheppard, "Partially coherent image formation in differential interference contrast (DIC) microscope," *Optics Express*, vol. 16, no. 24, pp. 19 462–19 479, Nov. 2008.
11. ——, "Quantitative phase-gradient imaging at high resolution with asymmetric illumination-based differential phase contrast," *Optics Letters*, vol. 34, no. 13, pp. 1924–1926, 2009.
12. L. Tian and L. Waller, "3D intensity and phase imaging from light field measurements in an LED array microscope," *Optica*, vol. 2, no. 2, pp. 104–111, Feb. 2015.
13. L.-H. Yeh, S. Chowdhury, N. A. Repina, L. Waller, and L. Waller, "Speckle-structured illumination for 3D phase and fluorescence computational microscopy," *Biomedical Optics Express*, vol. 10, no. 7, pp. 3635–3653, Jul. 2019.
14. N. Streibl, "Three-dimensional imaging by a microscope," *Journal of the Optical Society of America A*, vol. 2, no. 2, p. 121, Feb. 1985.
15. T. Noda, S. Kawata, and S. Minami, "Three-dimensional phase contrast imaging by an annular illumination microscope," *Applied Optics*, vol. 29, no. 26, pp. 3810–3815, Sep. 1990.
16. R. A. Claus, P. P. Naulleau, A. R. Neureuther, and L. Waller, "Quantitative phase retrieval with arbitrary pupil and illumination," *Optics Express*, vol. 23, no. 20, pp. 26 672–26 682, Oct. 2015.
17. M. H. Jenkins and T. K. Gaylord, "Quantitative phase microscopy via optimized inversion of the phase optical transfer function," *Applied Optics*, vol. 54, no. 28, p. 8566, Oct. 2015.
18. ——, "Three-dimensional quantitative phase imaging via tomographic deconvolution phase microscopy," *Applied Optics*, vol. 54, no. 31, pp. 9213–9227, Nov. 2015.
19. J. M. Soto, J. A. Rodrigo, and T. Alieva, "Label-free quantitative 3D tomographic imaging for partially coherent light microscopy," *Optics Express*, vol. 25, no. 14, pp. 15 699–15 712, Jul. 2017.

20. R. Oldenbourg and G. Mei, "New polarized light microscope with precision universal compensator," *Journal of Microscopy*, vol. 180, no. 2, pp. 140–147, 1995.
21. B. Ge, R. Zhou, Y. Takiguchi, Z. Yaqoob, and P. T. C. So, "Single-Shot Optical Anisotropy Imaging with Quantitative Polarization Interference Microscopy," *Laser & Photonics Reviews*, vol. 12, no. 8, p. 1800070, 2018.
22. M. Axer, D. Grässel, M. Kleiner, J. Dammers, T. Dickscheid, J. Reckfort, T. Hütz, B. Eiben, U. Pietrzyk, K. Zilles, and K. Amunts, "High-Resolution Fiber Tract Reconstruction in the Human Brain by Means of Three-Dimensional Polarized Light Imaging," *Frontiers in Neuroinformatics*, vol. 5, Dec. 2011.
23. M. Shribak and R. Oldenbourg, "Techniques for fast and sensitive measurements of two-dimensional birefringence distributions," *Applied Optics*, vol. 42, no. 16, pp. 3009–3017, Jun. 2003.
24. S. B. Mehta, M. Shribak, and R. Oldenbourg, "Polarized light imaging of birefringence and diattenuation at high resolution and high sensitivity," *Journal of Optics*, vol. 15, no. 9, p. 094007, Sep. 2013.
25. T. Ling, K. C. Boyle, V. Zuckerman, T. Flores, C. Ramakrishnan, K. Deisseroth, and D. Palanker, "How neurons move during action potentials," *bioRxiv*, p. 765768, Sep. 2019.
26. D. Keefe, L. Liu, W. Wang, and C. Silva, "Imaging meiotic spindles by polarization light microscopy: Principles and applications to IVF," *Reproductive BioMedicine Online*, vol. 7, no. 1, pp. 24–29, Jan. 2003.
27. M. Menzel, J. Reckfort, D. Weigand, H. Köse, K. Amunts, and M. Axer, "Diattenuation of Brain Tissue and its Impact on 3D Polarized Light Imaging," *arXiv:1703.04343 [physics]*, Mar. 2017.
28. M. Koike-Tani, T. Tominaga, R. Oldenbourg, and T. Tani, "Instantaneous polarized light imaging reveals activity dependent structural changes of dendrites in mouse hippocampal slices," *bioRxiv*, p. 523571, 2019.
29. S. B. Mehta, M. McQuilken, P. J. L. Riviere, P. Occhipinti, A. Verma, R. Oldenbourg, A. S. Gladfelter, and T. Tani, "Dissection of molecular assembly dynamics by tracking orientation and position of single molecules in live cells," *Proceedings of the National Academy of Sciences*, vol. 113, no. 42, pp. E6352–E6361, Oct. 2016.
30. B. J. Belin, L. M. Goins, and R. D. Mullins, "Comparative analysis of tools for live cell imaging of actin network architecture," *BioArchitecture*, vol. 4, no. 6, pp. 189–202, Jun. 2014.
31. S. B. Mehta and R. Oldenbourg, "Image simulation for biological microscopy: Microlith," *Biomedical Optics Express*, vol. 5, no. 6, pp. 1822–1838, 2014.
32. R. Oldenbourg, E. D. Salmon, and P. T. Tran, "Birefringence of Single and Bundled Microtubules," *Biophysical Journal*, vol. 74, no. 1, pp. 645–654, Jan. 1998.
33. M. I. Shribak, J. LaFountain, D. S. C. Biggs, and S. Inoue, "Orientation-independent differential interference contrast microscopy and its combination with an orientation-independent polarization system," *Journal of Biomedical Optics*, vol. 13, no. 1, p. 014011, Jan. 2008.
34. C. Ounkomol, S. Seshamani, M. M. Maleckar, F. Collman, and G. Johnson, "Label-free prediction of three-dimensional fluorescence images from transmitted light microscopy," May 2018.
35. E. M. Christiansen, S. J. Yang, D. M. Ando, A. Javaherian, G. Skibinski, S. Lipnick, E. Mount, A. O'Neil, K. Shah, A. K. Lee, P. Goyal, W. Fedus, R. Poplin, A. Esteva, M. Berndl, L. L. Rubin, P. Nelson, and S. Finkbeiner, "In Silico Labeling: Predicting Fluorescent Labels in Unlabeled Images," *Cell*, vol. 173, no. 3, pp. 792–803.e19, Apr. 2018.
36. Y. Rivenson, T. Liu, Z. Wei, Y. Zhang, and A. Ozcan, "PhaseStain: Digital staining of label-free quantitative phase microscopy images using deep learning," *arXiv:1807.07701 [physics]*, Jul. 2018.
37. Y. Rivenson, H. Wang, Z. Wei, Y. Zhang, H. Gunaydin, and A. Ozcan, "Deep learning-based virtual histology staining using auto-fluorescence of label-free tissue," *arXiv:1803.11293 [physics]*, Mar. 2018.
38. M. Lee, Y.-H. Lee, J. Song, G. Kim, Y. Jo, H. Min, C. H. Kim, and Y. Park, "DeepIS: Deep learning framework for three-dimensional label-free tracking of immunological synapses," *bioRxiv*, p. 539858, Apr. 2019.
39. 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*, vol. 142, no. 8, pp. 1207–1215, Apr. 2017.
40. C. Ounkomol, S. Seshamani, M. M. Maleckar, F. Collman, and G. R. Johnson, "Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy," *Nature Methods*, vol. 15, no. 11, p. 917, Nov. 2018.
41. M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light.* Elsevier, Jun. 2013.
42. C. Angermann and M. Haltmeier, "Random 2.5D U-net for Fully 3D Segmentation," in *Machine Learning and Medical Engineering for Cardiovascular Health and Intravascular Imaging and Computer Assisted Stenting*, ser. Lecture Notes in Computer Science, H. Liao, S. Balocco, G. Wang, F. Zhang, Y. Liu, Z. Ding, L. Duong, R. Phellan, G. Zahnd, K. Breininger, S. Albarqouni, S. Moriconi, S.-L. Lee, and S. Demirci, Eds. Cham: Springer International Publishing, 2019, pp. 158–166.
43. B. de Campos Vidal, M. L. S. Mello, A. C. Caseiro-Filho, and C. Godo, "Anisotropic properties of the myelin sheath," *Acta Histochemica*, vol. 66, no. 1, pp. 32–39, Jan. 1980.
44. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, ser. Lecture Notes in Computer Science. Springer, Cham, Oct. 2015, pp. 234–241.
45. 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)*, Oct. 2016, pp. 565–571.
46. M. Drozdzal, E. Vorontsov, G. Chartrand, S. Kadoury, and C. Pal, "The Importance of Skip Connections in Biomedical Image Segmentation," *arXiv:1608.04117 [cs]*, Aug. 2016.
47. 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*, vol. 26, no. 1, pp. 98–117, Jan. 2009.
48. C. Schell, N. Wanner, and T. B. Huber, "Glomerular development–shaping the multi-cellular filtration unit," in *Seminars in Cell & Developmental Biology*, vol. 36. Elsevier, 2014, pp. 39–49.
49. I. Jakovcevski, R. Filipovic, Z. Mo, S. Rakic, and N. Zecevic, "Oligodendrocyte Development and the Onset of Myelination in the Human Fetal Brain," *Frontiers in Neuroanatomy*, vol. 3, Jun. 2009.
50. D. J. Miller, T. Duka, C. D. Stimpson, S. J. Schapiro, W. B. Baze, M. J. McArthur, A. J. Fobbs, A. M. M. Sousa, N. Šestan, D. E. Wildman, L. Lipovich, C. W. Kuzawa, P. R. Hof, and C. C. Sherwood, "Prolonged myelination in human neocortical evolution," *Proceedings of the National Academy of Sciences*, vol. 109, no. 41, pp. 16 480–16 485, Oct. 2012.
51. P. C. Monsma and A. Brown, "FluoroMyelin™ Red is a bright, photostable and non-toxic fluorescent stain for live imaging of myelin," *Journal of Neuroscience Methods*, vol. 209, no. 2, pp. 344–350, Aug. 2012.
52. H. Wang, C. Magnain, R. Wang, J. Dubb, A. Varjabedian, L. S. Tirrell, A. Stevens, J. C. Augustinack, E. Konukoglu, I. Aganj, M. P. Frosch, J. D. Schmahmann, B. Fischl, and D. A. Boas, "As-PSOCT: Volumetric microscopic imaging of human brain architecture and connectivity," *NeuroImage*, vol. 165, pp. 56–68, Jan. 2018.
53. J. F. de Boer, C. K. Hitzenberger, and Y. Yasuno, "Polarization sensitive optical coherence tomography – a review [Invited]," *Biomedical Optics Express*, vol. 8, no. 3, pp. 1838–1873, Feb. 2017.
54. H. Zeng, "Mesoscale Connectomics," *Current opinion in neurobiology*, vol. 50, pp. 154–162, Jun. 2018.
55. M. Bass, C. DeCusatis, J. M. Enoch, V. Lakshminarayanan, G. Li, C. MacDonald, V. N. Mahajan, and E. V. Stryland, *Handbook of Optics, Volume I: Geometrical and Physical Optics, Polarized Light, Components and Instruments(Set).* McGraw Hill Professional, Oct. 2009.
56. Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A New Alternating Minimization Algorithm for Total Variation Image Reconstruction," *SIAM Journal on Imaging Sciences*, vol. 1, no. 3, pp. 248–272, Jan. 2008.
57. M. Chen, L. Tian, and L. Waller, "3D differential phase contrast microscopy," *Biomedical Optics Express*, vol. 7, no. 10, pp. 3940–3950, Oct. 2016.

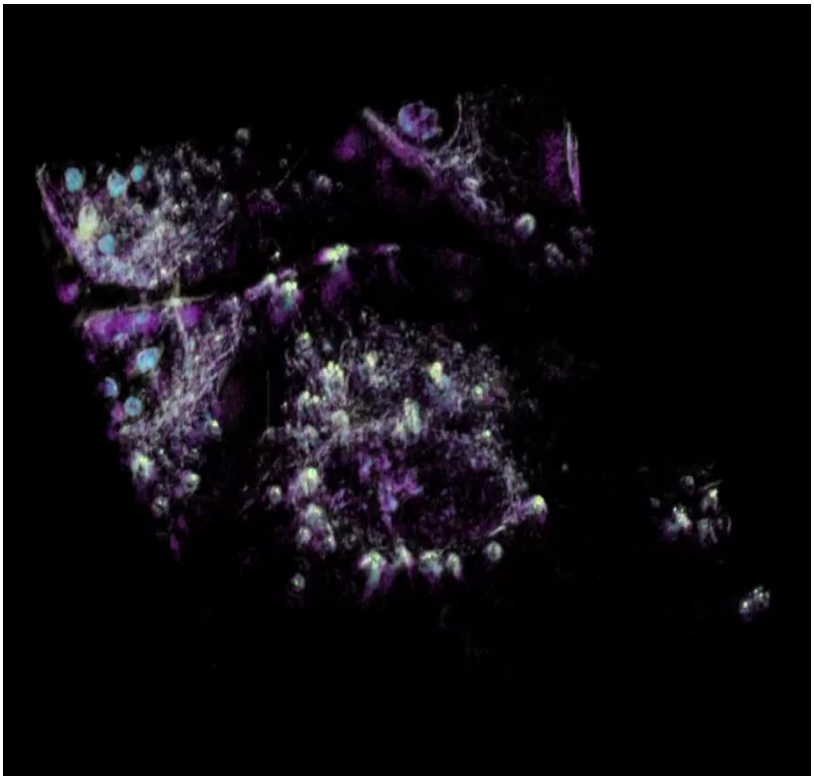
Videos



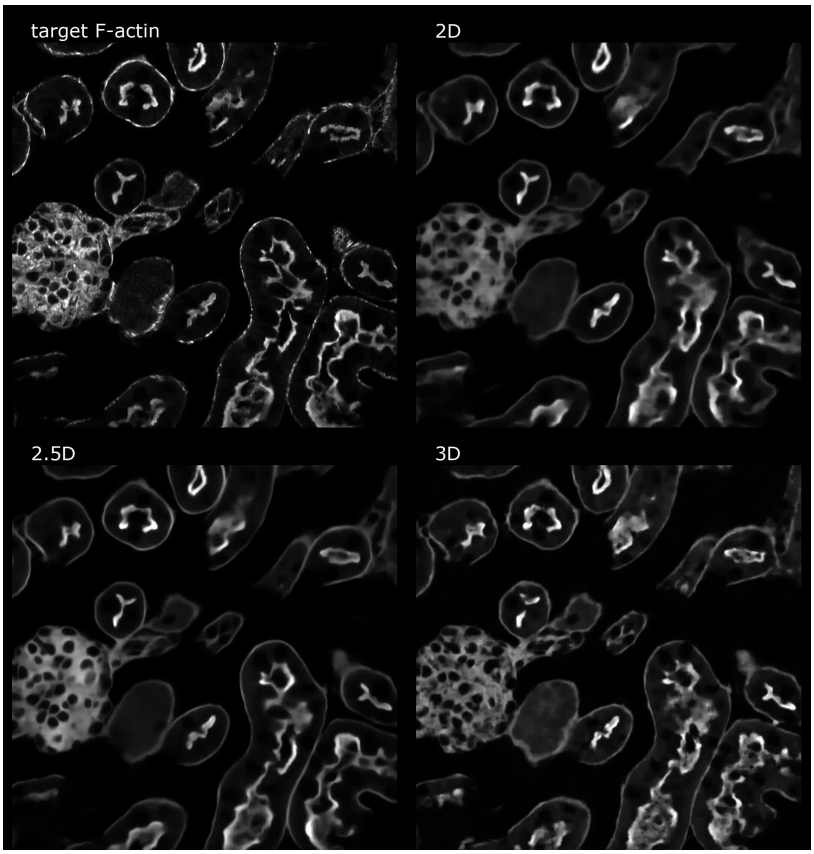
Video 1. Z-stacks showing brightfield , phase, retardance, and color-coded orientation images of mouse kidney tissue for field of view show in fig. 2, fig. 3, and fig. 4.



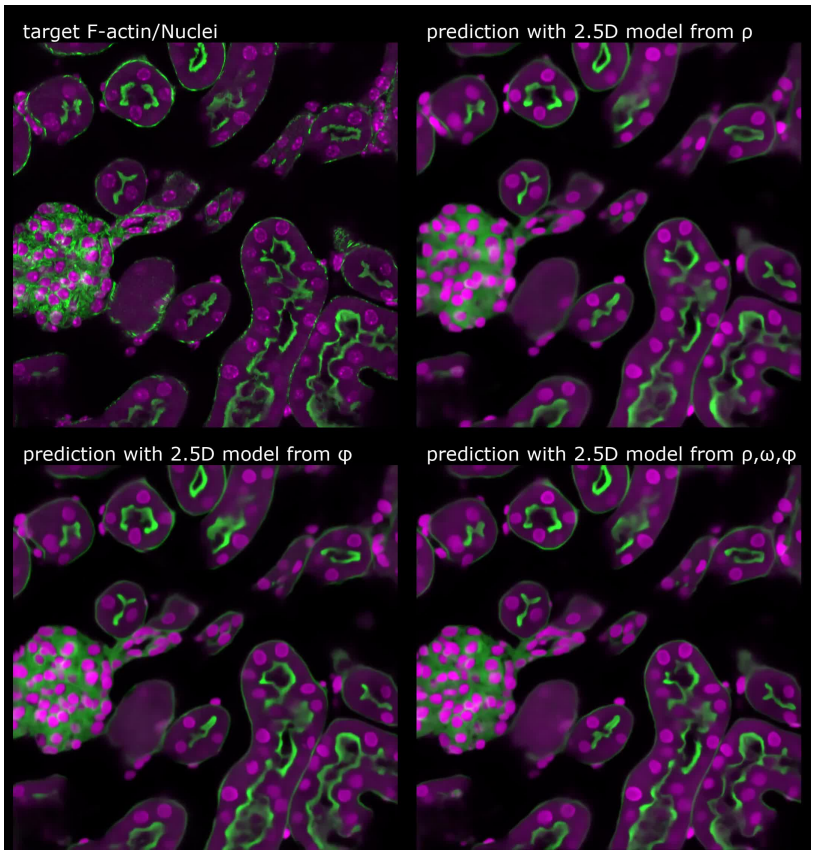
Video 2. Time-lapse of phase, retardance, and slow axis orientation in a dividing U2OS cell shows differences in density and anisotropy of organelles within cell.



Video 3. 3D rendering of the time-lapse showing diverse structures color coded by their retardance and phase in U2OS cells shown in fig. 2. Phase values from low to high are color-coded with magenta, cyan, and green, respectively. Retardance values are color-coded with yellow. This combination results in magenta cytoplasm, magenta to blue chromosomes as they condensed, blue nucleoli, green lipid vesicles surrounded by yellow rings, and yellow spindles.



Video 4. Through focus series showing 3D F-actin distribution in the test field of view shown in fig. 3: We show F-actin distribution (labeled with phalloidin-AF568) acquired on a confocal microscope (target), as well as predicted from 2D, 2.5D, and 3D models trained to translate retardance distribution into fluorescence distributions.



Video 5. Through focus series showing 3D F-actin and nuclei distribution in the test field of view shown in fig. 4: We show overlays of F-actin (phalloidin-AF568) distribution in green and nuclei (DAPI) distribution in magenta DAPI as acquired on confocal (target) and as predicted from models. Predictions were obtained from 2.5D models trained on retardance (ρ) alone, phase (ϕ) alone, and on combination of retardance, orientation, phase (ρ, ω, ϕ).

Supplementary figures

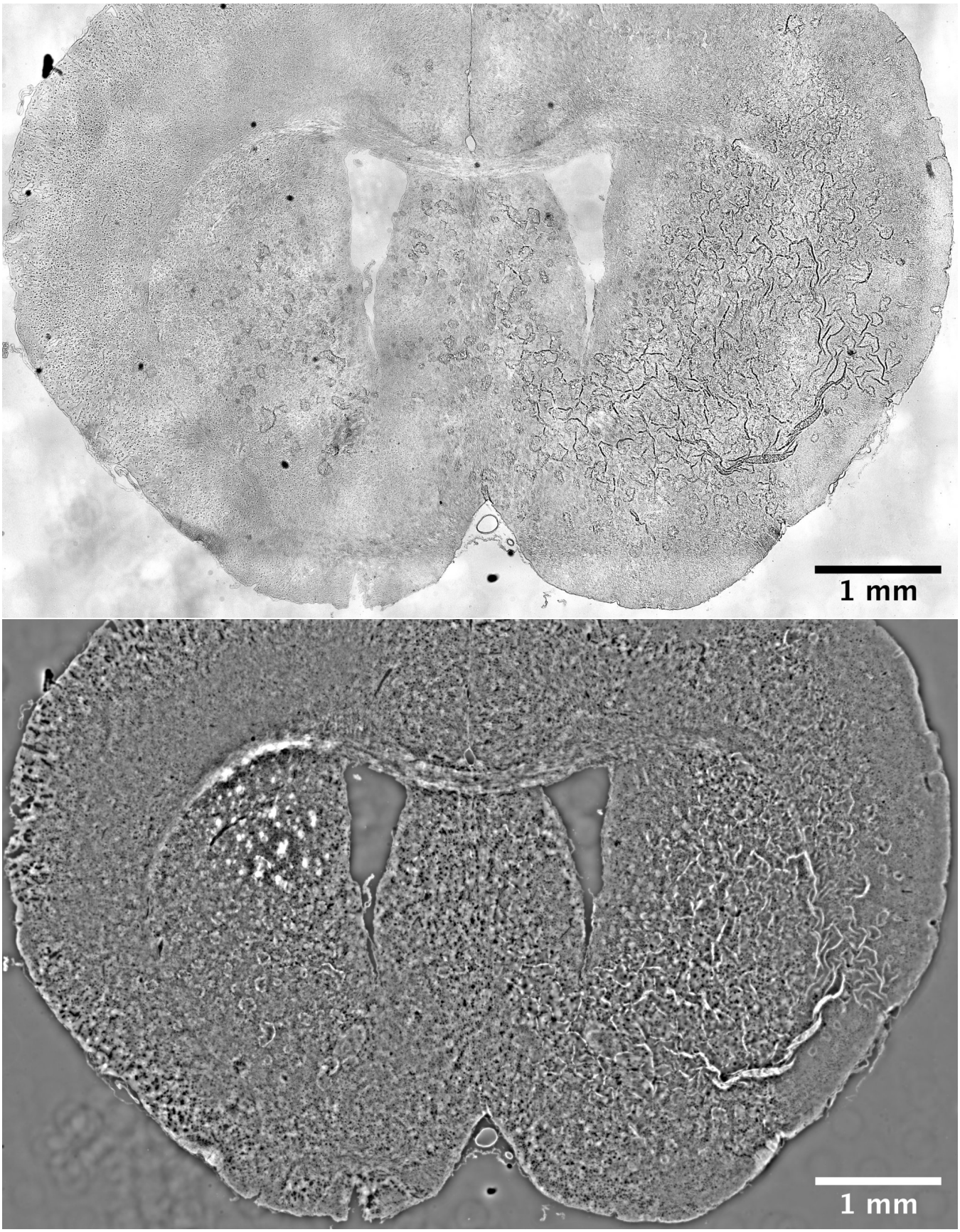


Figure 2 - supplement S1. Comparison of background-corrected brightfield image (top) with quantitative phase image (bottom) over a mouse brain slice. The phase image is more robust to intensity variations due to light path and reports density variations at higher contrast. These images are stitches of 48 fields of view and are substantially downsampled to reduce size.

Figure 2 - supplement S2. Retardance (top) and orientation (bottom) measurements over a mouse brain slice, which report structural anisotropy and slow axis orientation, respectively. We needed to compress the measured dynamic range of retardance and orientation by using gamma correction (0.5) to visualize less anisotropic gray matter in the presence of highly anisotropic white matter. These images are stitches of 48 fields of view and are substantially downsampled to reduce size. The peak retardance is $50\mathrm{nm}$.

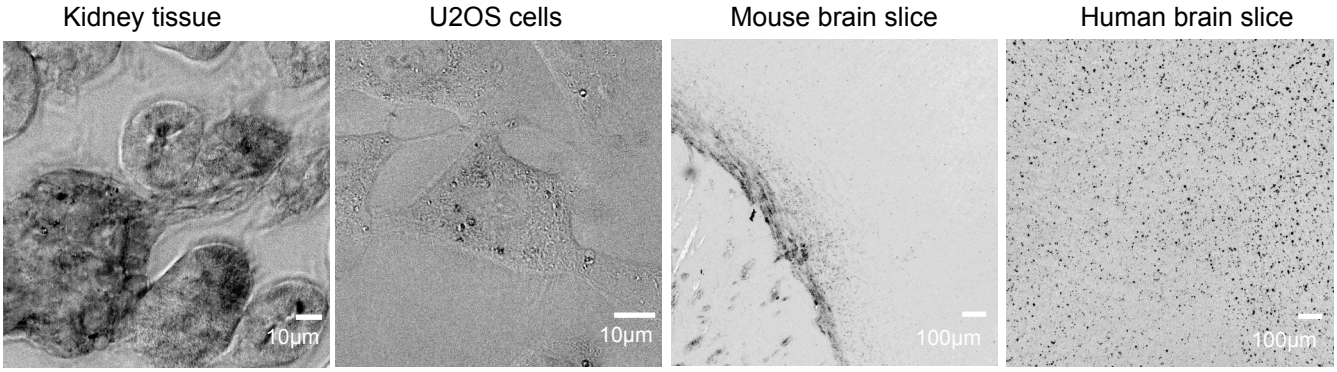


Figure 2 - supplement S3. Degree of polarization (DOP) images for four specimens shown in fig. 2. When the specimen does not exhibit multiple scattering or diattenuation, $\mathrm{DOP} = 1$. The DOP reduces when specimen exhibits multiple scattering and increases specimen exhibits diattenutation, i.e., polarization-dependent absorption.

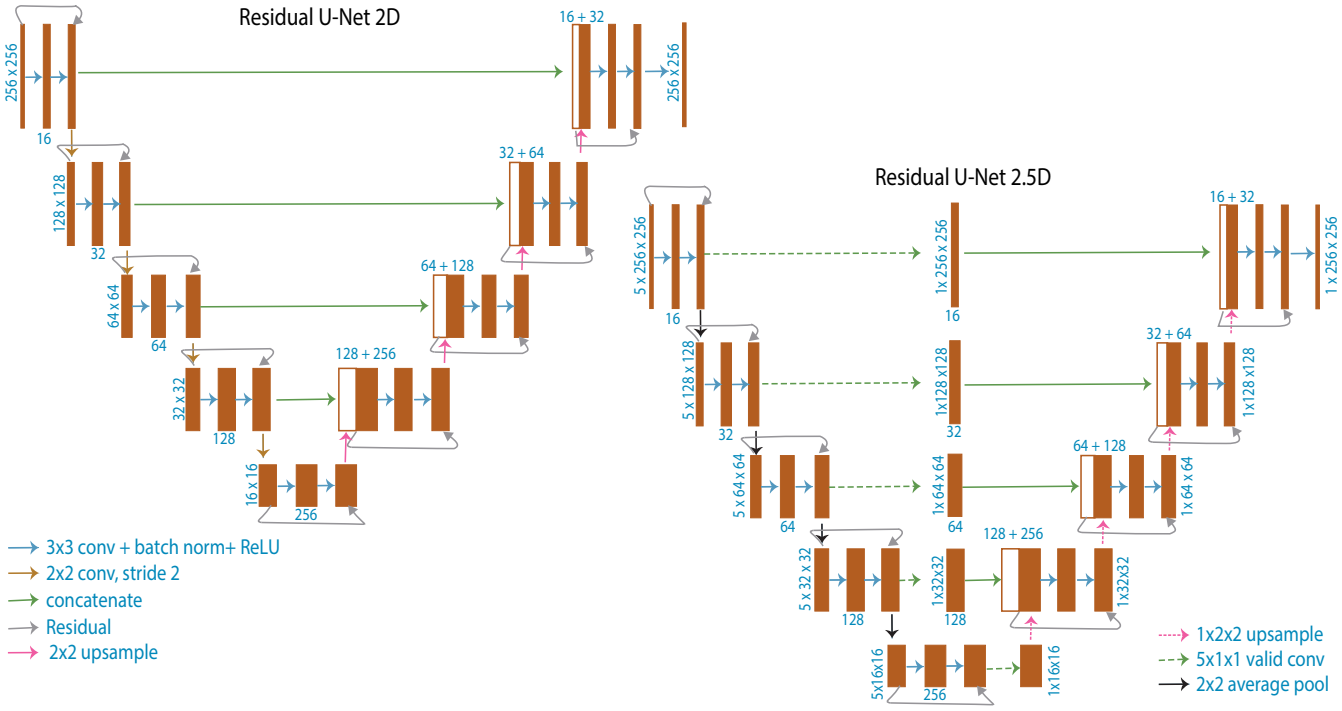


Figure 3 - supplement S1. Schematic illustrating U-Net architectures: Schematic of 2D U-Net model used for translating slice→slice and 2.5D U-Net model used for translating stack→slice. The 3D U-Net model used for translating stack->stack is similar to the 2D U-Net, but uses 3D convolutions instead of 2D and is 4 layers deep instead of 5 layers deep.

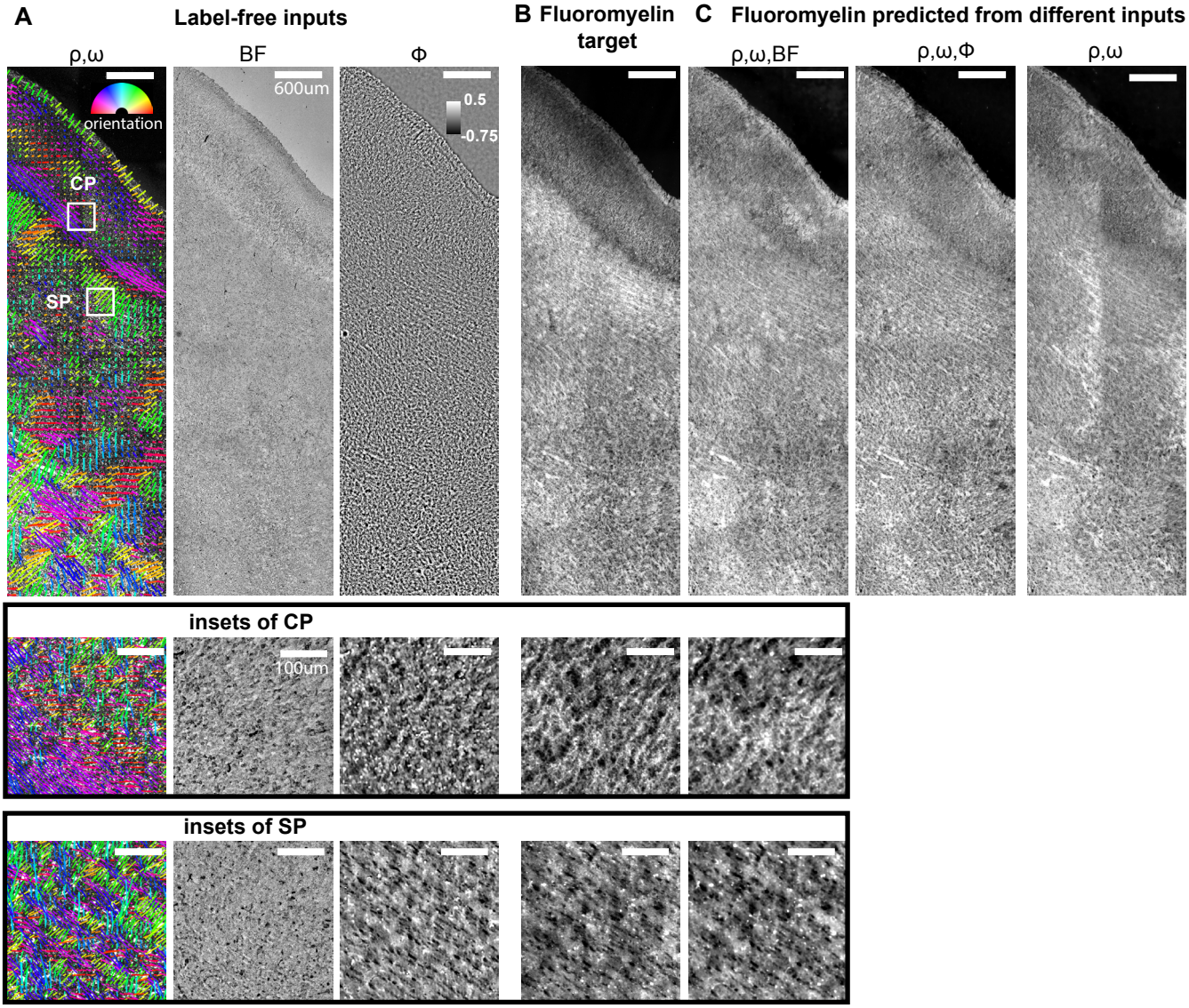


Figure 5 - supplement S1. Label-free images and predicted Fluoromyelin images of a slice of human brain at gestational week 24 (GW24): (A) retardance (ρ), orientation (ω), brightfield (BF), and phase (ϕ) images from a field of view drawn from a test dataset, (B) experimental Fluoromyelin (target), (C) Fluoromyelin predicted by image translation models using different label-free images as inputs: (left to right) retardance, orientation, and brightfield as input to 2.5D model; retardance, orientation, and phase as input of 2D model; retardance and orientation as input to 2.5D model. Insets at the bottom show label-free, target, and predicted images within cortical plate (CP) and subplate (SP) at higher magnification. Scale bars: (A-C) $600\mu\mathrm{m}$, (insets) $100\mu\mathrm{m}$.

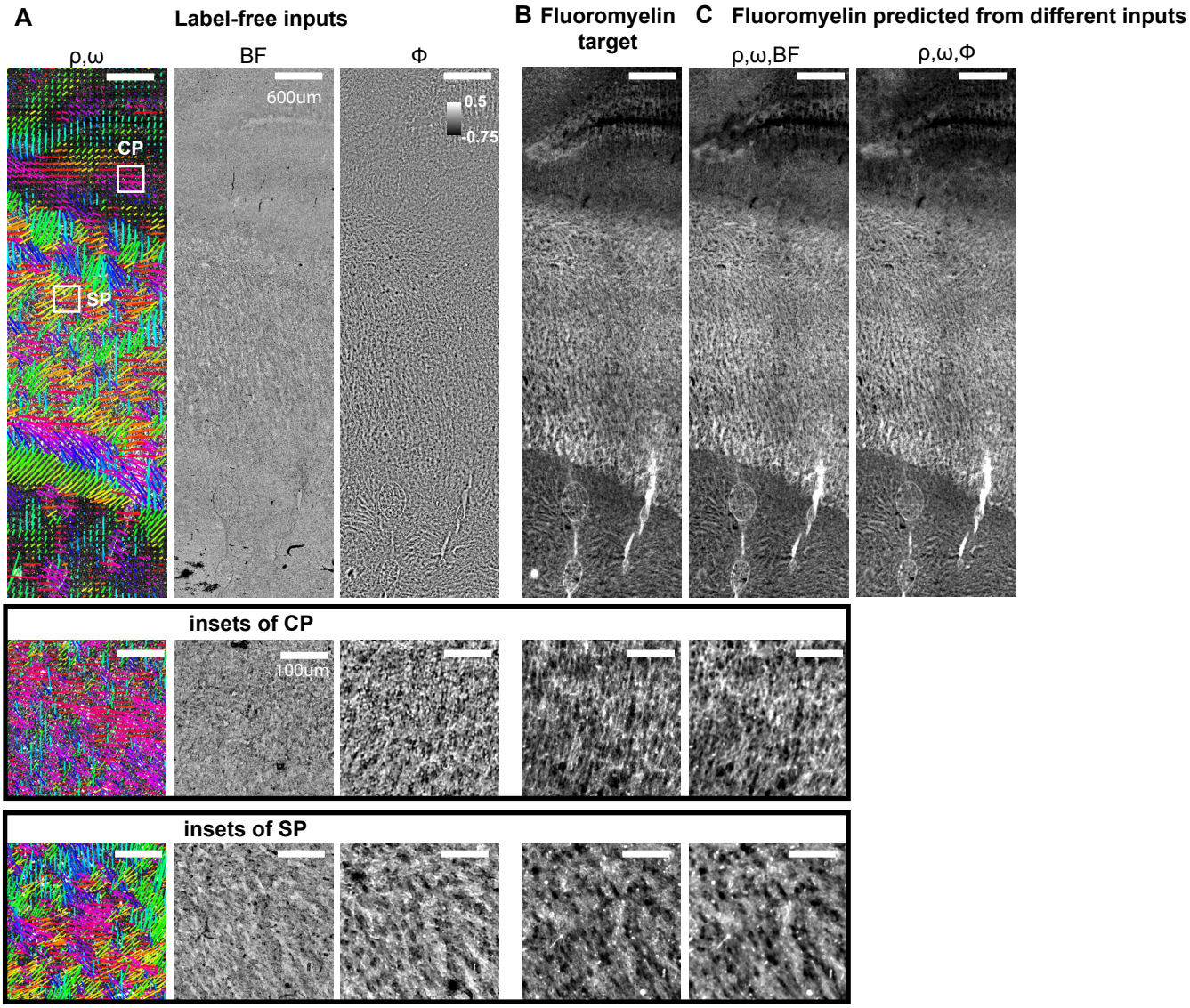


Figure 5 - supplement S2. Label-free images and predicted Fluoromyelin images of a slice of human brain at gestational week 20 (GW20): The data is presented in the same way as fig. 5-supplement S1.

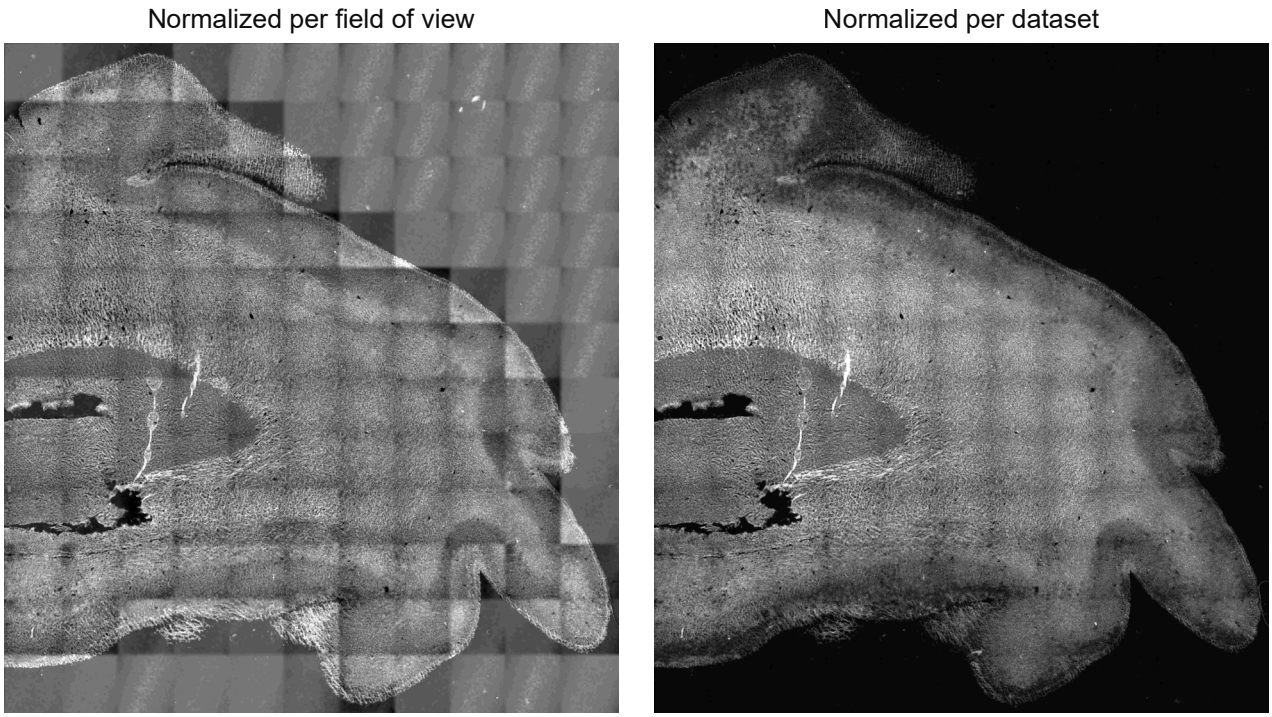


Figure 6 - supplement S1. Normalizing training data per dataset yields prediction with correct dynamic range of intensity: Predictions of FluoroMyelin in GW20 human brain tissue slice with training data normalized per field of view (left) and across dataset (right)