

Electronic Supplementary Material (ESI) for Lab on a Chip

High-throughput digital pathology via a handheld, multiplexed, and AI-powered ptychographic whole slide scanner

Shaowei Jiang,^{‡a} Chengfei Guo,^{‡*a} Pengming Song,^{‡a} Tianbo Wang,^a Ruihai Wang,^a Terrance Zhang,^a Qian Wu,^b Rishikesh Pandey,^a and Guoan Zheng^{*a}

^aDepartment of Biomedical Engineering, University of Connecticut, Storrs, CT, 06269, USA

^bPathology and Laboratory Medicine, University of Connecticut Health Centre, Farmington, CT, 06030, USA

*Email: chengfei.quo@uconn.edu or guoan.zheng@uconn.edu

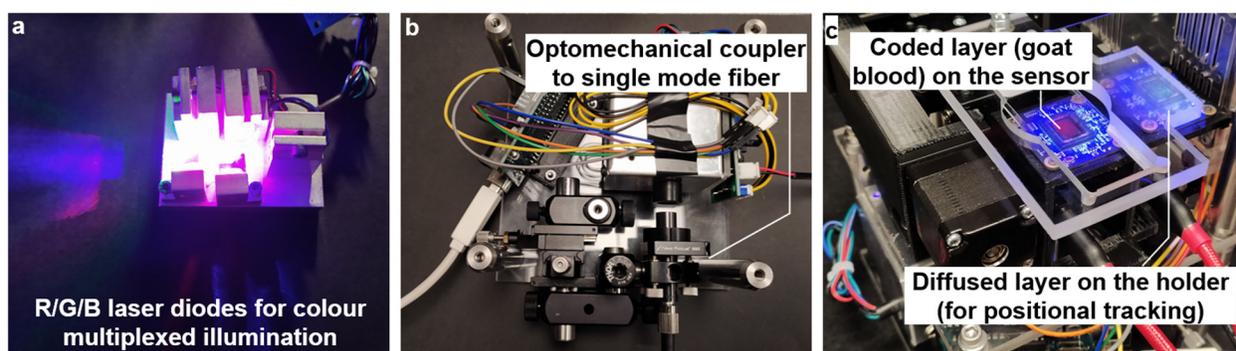


Fig. S1: System design for the handheld ptychographic whole slide scanner. (a) R/G/B laser diodes for colour-multiplexed illumination. (b) Optomechanical design for coupling the laser beam to the optical fiber. (c) The synchronized image sensor pair for image acquisition and positional tracking.

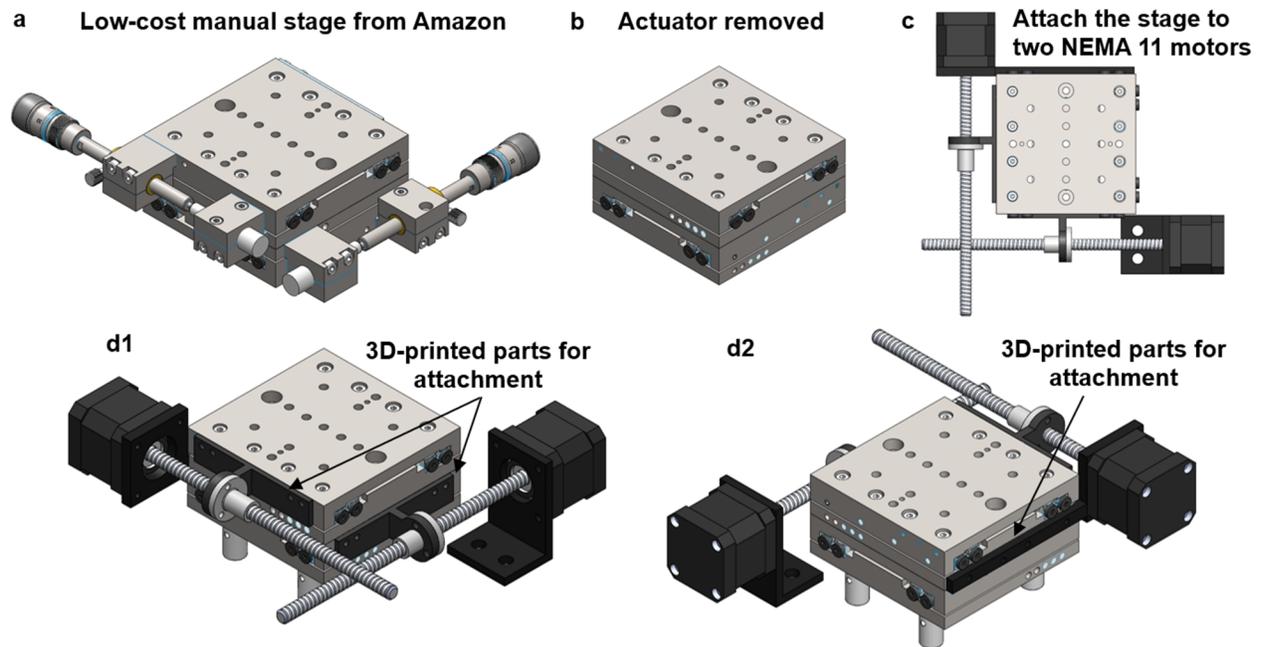


Fig. S2: Development of a DIY motorized stage for the ptychographic whole slide scanner. (a) The original low-cost manual stage from Amazon (<\$100). (b) Removing the actuator from the original stage. (c) Assembling two NEMA 11 stepper motors with the stage. (d) Isometric views of the assembled motorized stage.

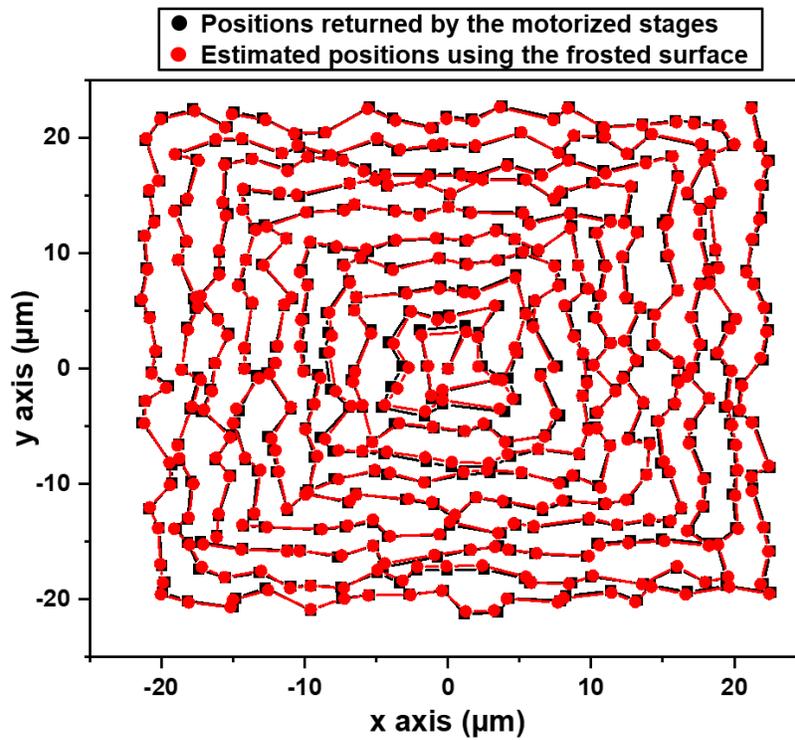


Fig. S3: Validating the positional tracking performance using the frosted surface. The average difference between the estimated positions and the positions returned by the ASI LS-50 motorized stage is $0.102 \mu\text{m}$, with a standard deviation of $0.076 \mu\text{m}$.

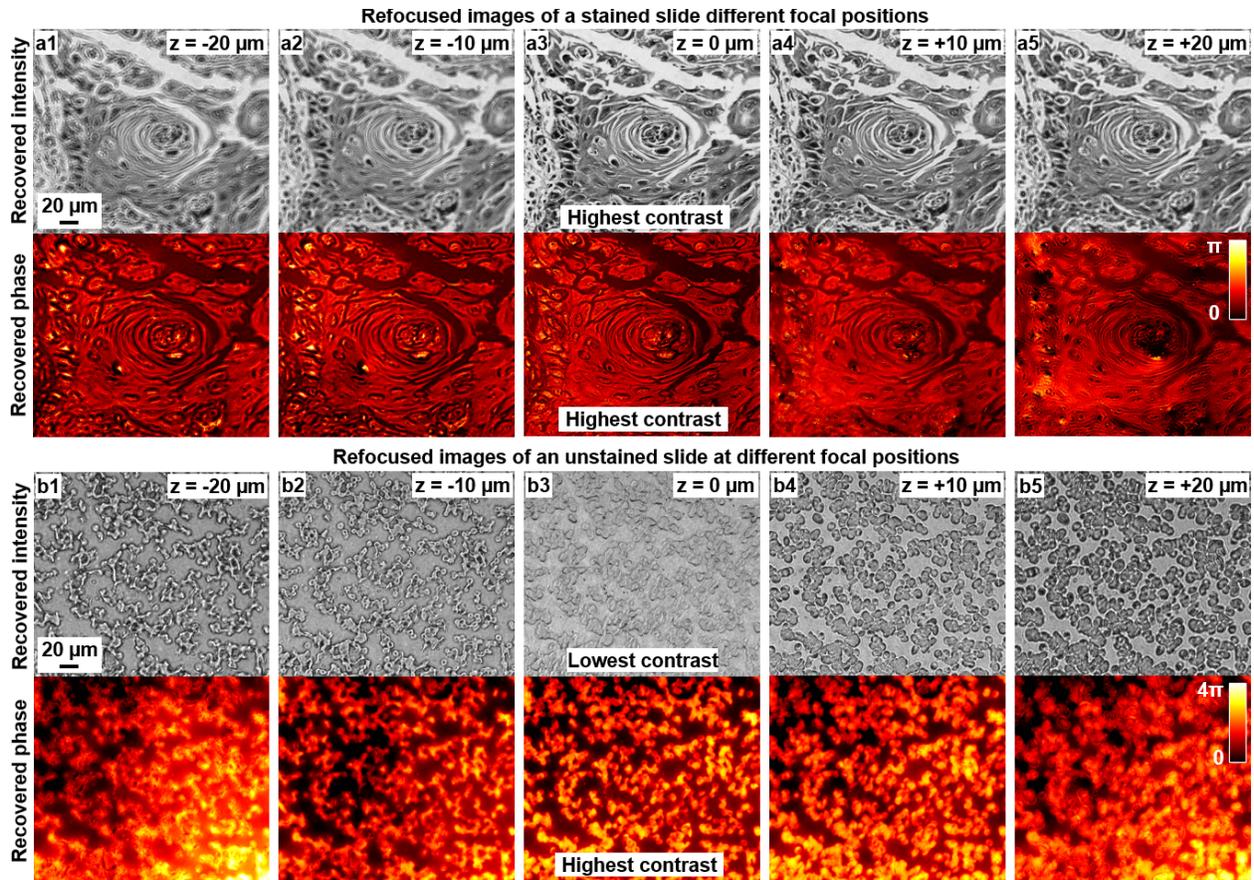


Fig. S5: Intensity and phase contrast comparison for stained and unstained slides. (a) The refocused images of an H&E-stained slide at different axial positions. (b) The refocused images of an unstained thyroid smear at different axial positions. The phase contrast correlates with the best focus positions for both cases.

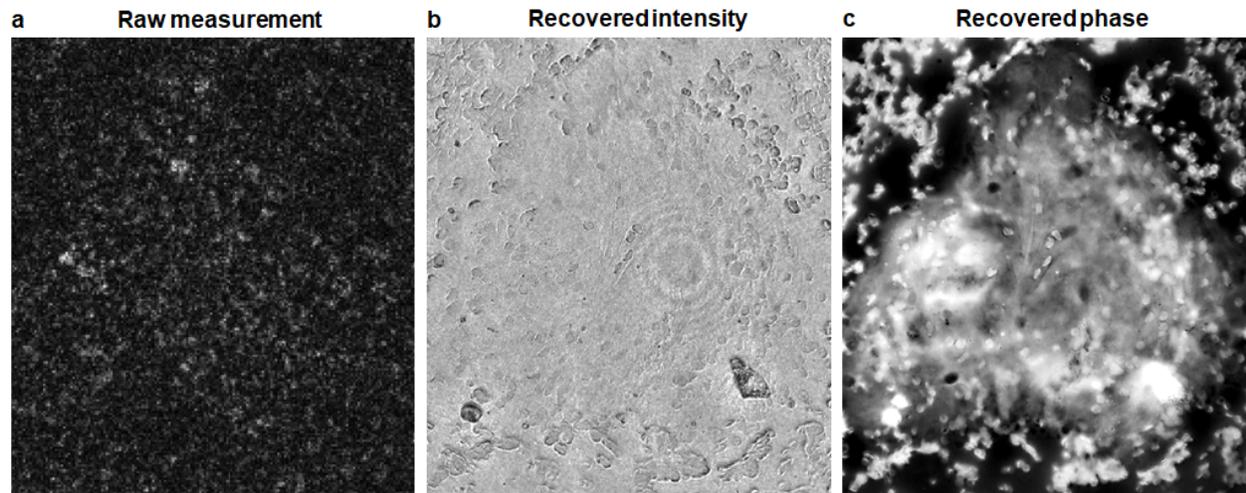


Fig. S6: Validating the imaging performance for label-free samples. (a) The captured raw image of an unstained thyroid sample. (b) The recovered intensity image. (c) The recovered phase image, where we can clearly see the cell cluster from the unwrapped phase.

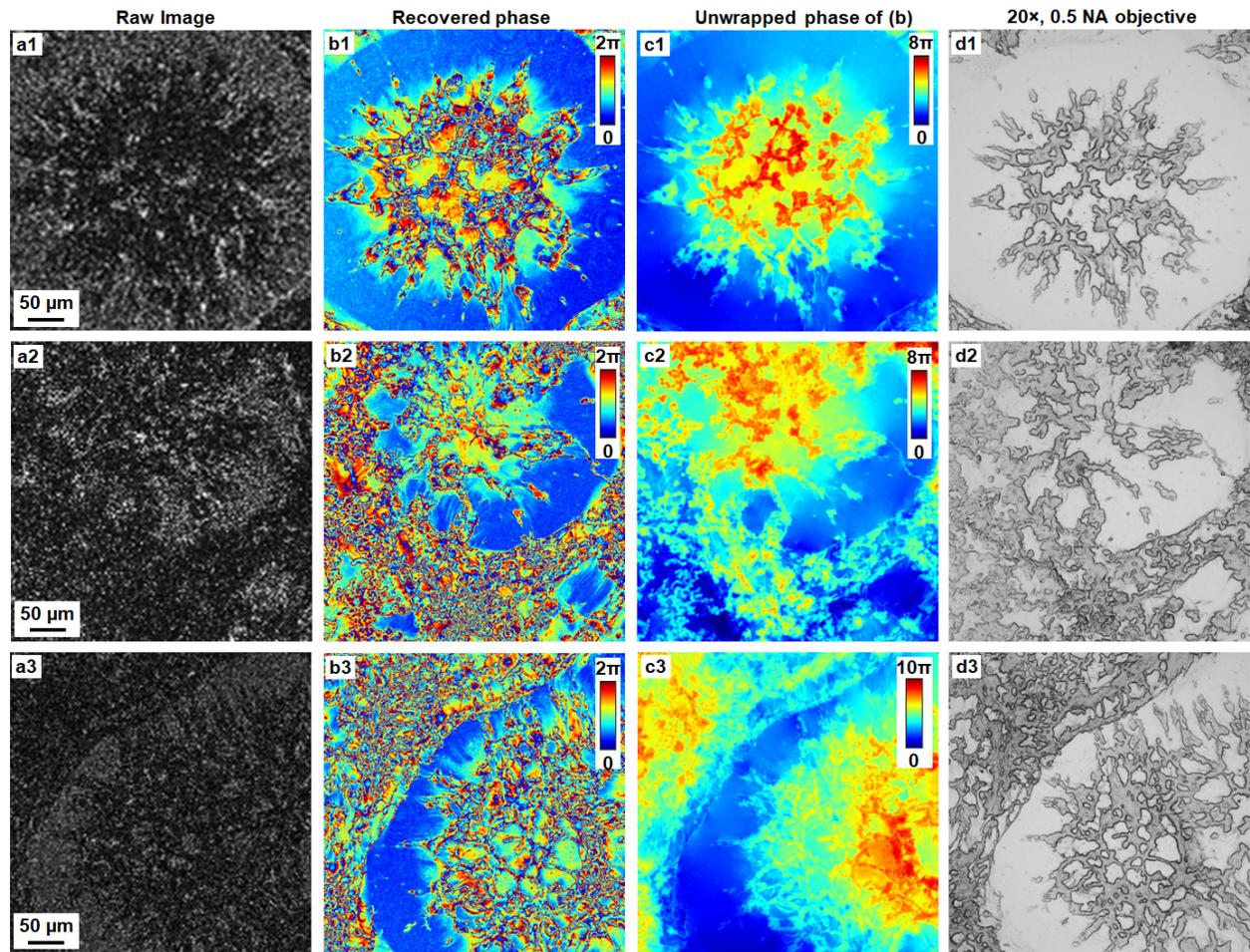


Fig. S7: Validating the imaging performance of unstained thyroid FNA smears. (a) The captured raw images. (b) The recovered phase images. (c) The unwrapped phase of (b). (d) The intensity images of the unstained thyroid smear captured using a 20x/ 0.5 NA objective lens.

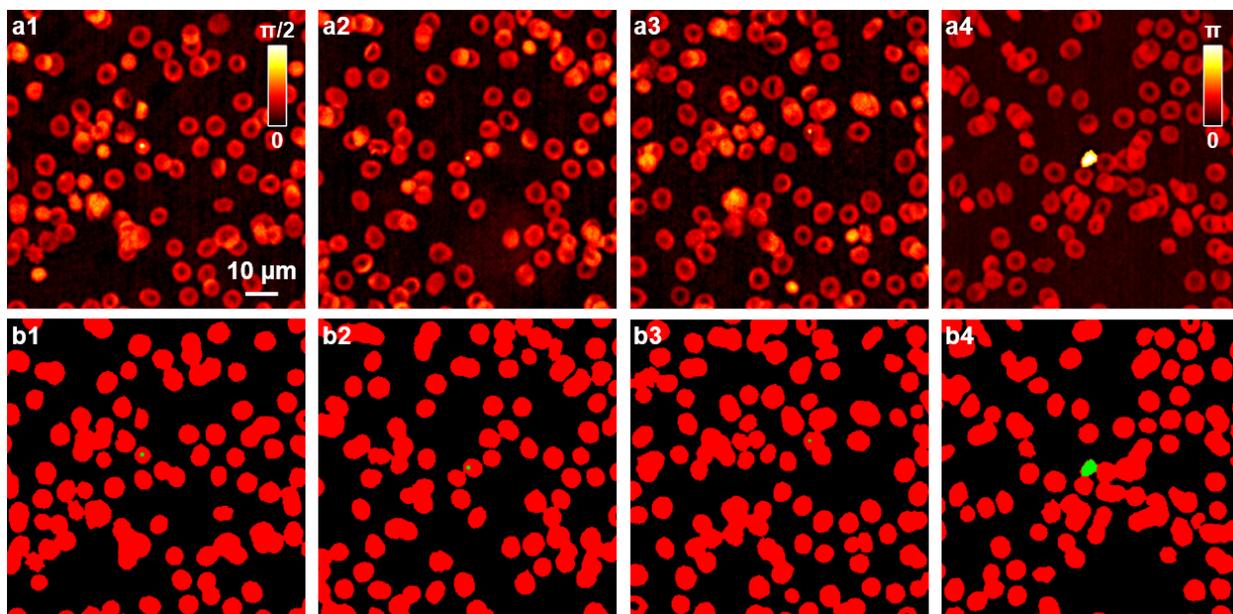


Fig. S8: Automatic tracking of malaria-infected blood cells. (a) The recovered phase images. (b) The cell segmentation masks, where the red regions represent the cell background and the green regions represent the parasite candidates. Based on the size of the green mask, we rule out the white blood cell in (b4) as a malaria parasite.

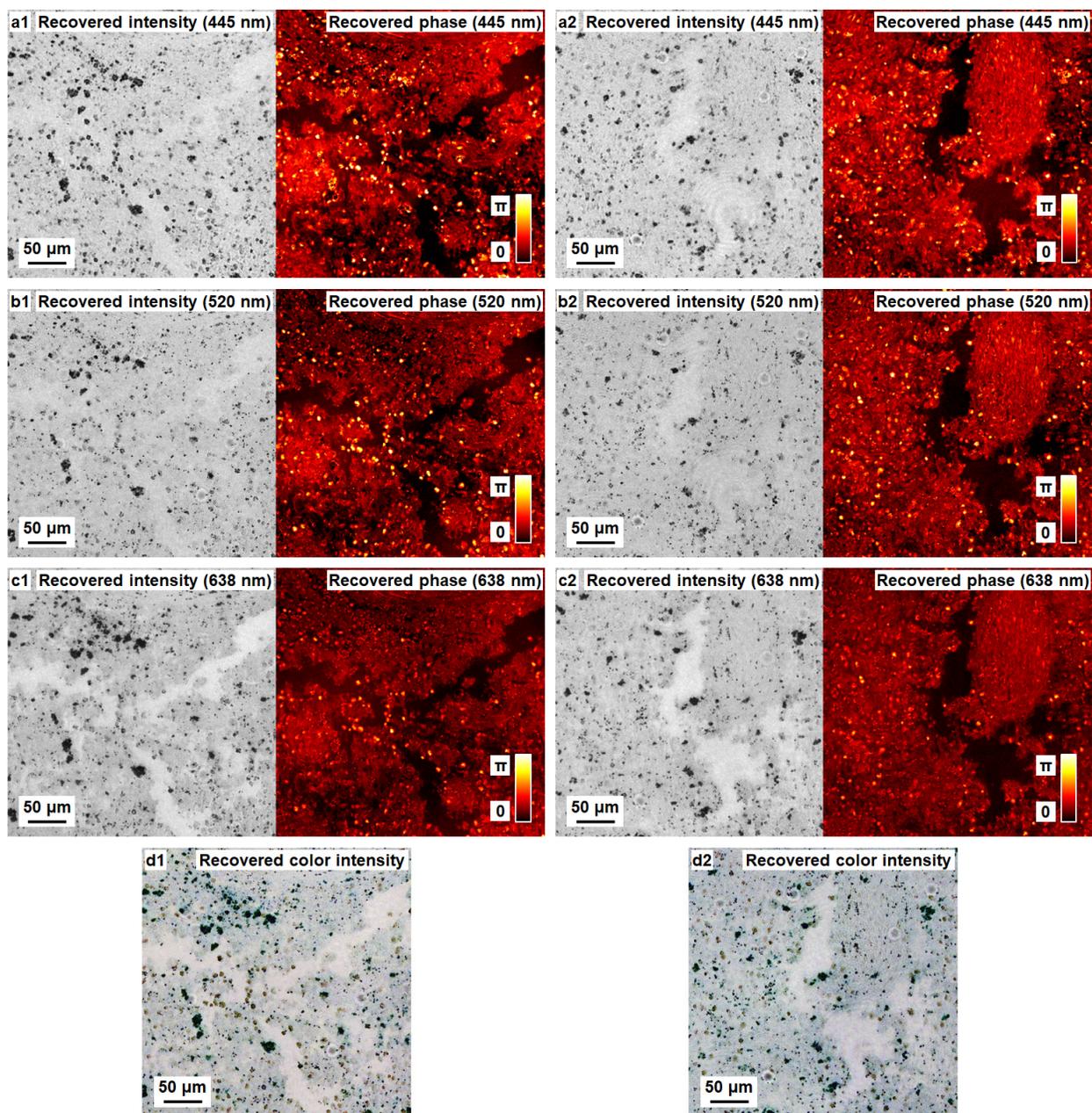


Fig. S9: The ptychographic reconstruction at the three wavelengths. (a) The recovered intensity and phase at the wavelength of 445 nm (a), 520 nm (b), and 638 nm (c). (d) The recovered colour images.

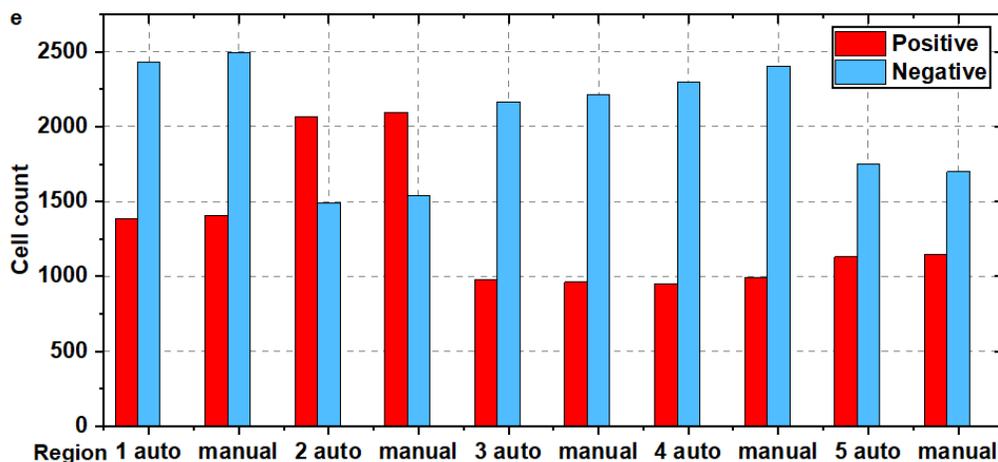
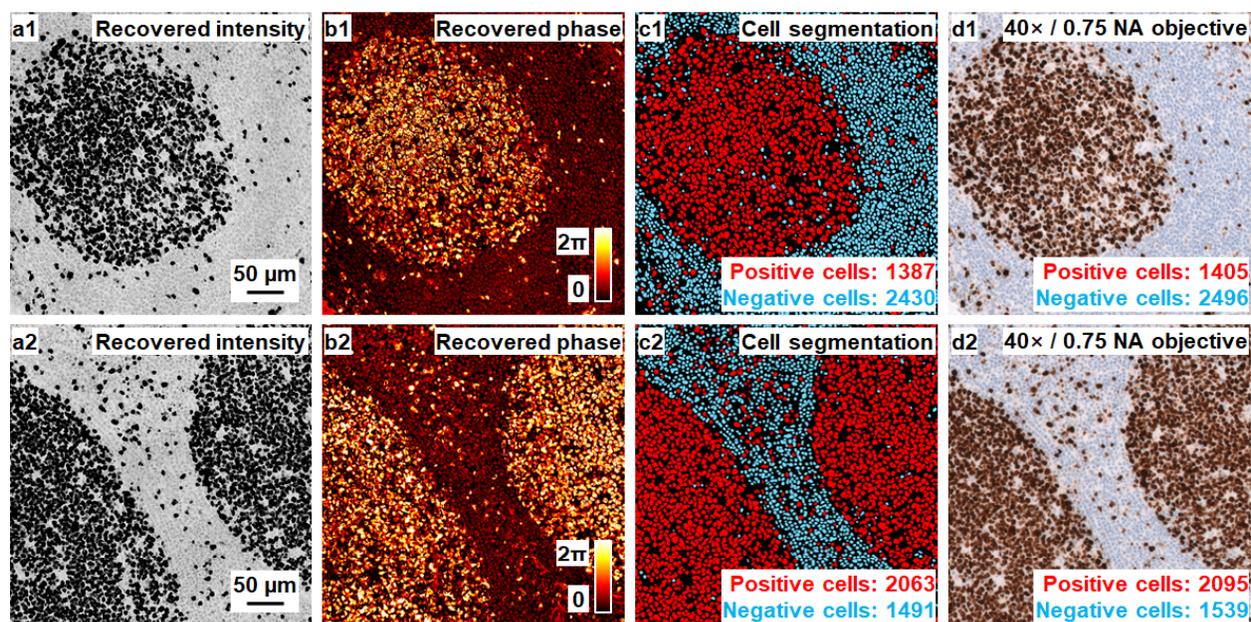


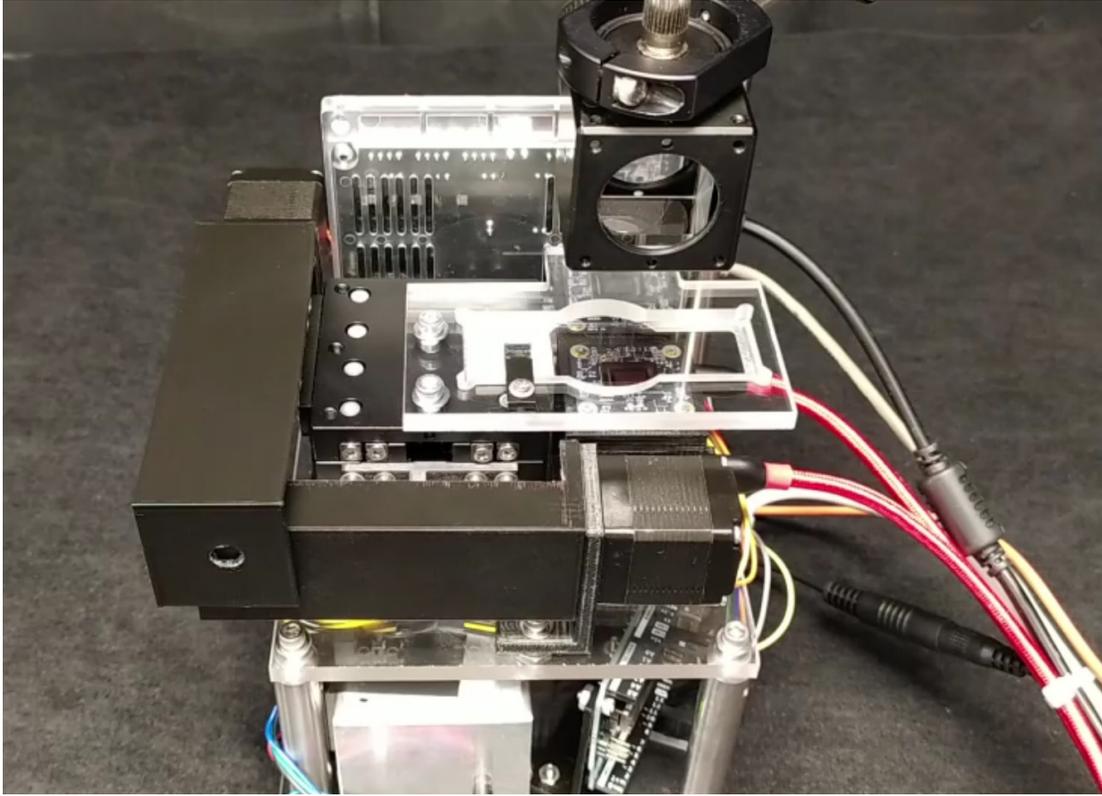
Fig. S10: Validating the accuracy of the segmentation results. (a) The recovered intensity. (b) The recovered phase. (c) The cell segmentation mask generated by the neural network. (d) The intensity image captured using a 40× / 0.95 NA objective. (e) The comparison of cell counting between the network-based segmentation and manual counting using the 40×, 0.95 NA objective lens. The difference between the two approaches is ~2.6% for 5 different regions.

	Precision	Recall	Accuracy	F1 score
Region 1	0.980	0.940	0.970	0.959
Region 2	0.978	0.963	0.966	0.971
Region 3	0.972	0.926	0.968	0.948
Region 4	0.948	0.882	0.948	0.913
Region 5	0.974	0.935	0.962	0.954
Average	0.970	0.930	0.963	0.949

Table. S1: Validating the network-based cell classification performance with 4 different metrics -- precision, recall, accuracy, and F1 score. The metrics are defined as precision = $\frac{TP}{TP+FP}$, recall = $\frac{TP}{TP+FN}$, accuracy = $\frac{TP+TN}{TP+FN+TN+FP}$, and F1 score = $\frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$. TP, FP, TN, and FN represent true positive, false positive, true negative, and false negative, respectively.

	Cell area (μm^2 , Positive)	Cell area (μm^2 , Negative)	Eccentricity (Positive)	Eccentricity (Negative)
Region 1 output	23.1 \pm 12.4	15.7 \pm 7.7	0.63 \pm 0.17	0.56 \pm 0.21
Region 1 GT	23.9 \pm 12.6	18.7 \pm 9.9	0.64 \pm 0.20	0.59 \pm 0.21
Region 2 output	25.6 \pm 14.3	13.9 \pm 6.6	0.62 \pm 0.19	0.61 \pm 0.22
Region 2 GT	27.3 \pm 13.2	17.3 \pm 7.0	0.63 \pm 0.19	0.64 \pm 0.19
Region 3 output	19.4 \pm 12.1	15.9 \pm 8.2	0.62 \pm 0.23	0.59 \pm 0.19
Region 3 GT	19.8 \pm 11.3	18.2 \pm 10.3	0.63 \pm 0.21	0.60 \pm 0.22
Region 4 output	18.9 \pm 11.2	17.2 \pm 8.3	0.62 \pm 0.20	0.61 \pm 0.20
Region 4 GT	20.2 \pm 12.1	19.8 \pm 9.7	0.65 \pm 0.18	0.64 \pm 0.18
Region 5 output	19.2 \pm 10.6	16.1 \pm 6.8	0.61 \pm 0.20	0.61 \pm 0.18
Region 5 GT	21.5 \pm 13.3	18.3 \pm 8.3	0.63 \pm 0.22	0.62 \pm 0.20
Average GT	22.5	18.5	0.64	0.62
Average output	21.2 (94.2%)	15.8 (85.4%)	0.62 (96.9%)	0.60 (96.8%)

Table. S2: Comparison between the network-based segmentation (network outputs) and the ground-truth (GT) results. The GT results are obtained using images captured with a 40 \times , 0.95 NA objective lens.



Movie S1: The operation of the prototype platform. We converted the low-cost manual stage into a motorized stage for sample scanning. We used an Arduino UNO microcontroller to control the R/G/B laser, the motorized stage, and the image sensor pair. The motion speed in our implementation is $\sim 60 \mu\text{m/s}$ and it has been exaggerated for visualization in Movie S1.