Lab on a Chip

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/d2lc00084a

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]

Received 25th January 2022,
Accepted 11th May 2022

DOI: 10.1039/d2lc00084a

rsc.li/loc

The recent advent of whole slide imaging (WSI) systems has moved digital pathology closer to diagnostic applications and clinical practices. Integrating WSI with machine learning promises the growth of this field in upcoming years. Here we report the design and implementation of a handheld, colour-multiplexed, and AI-powered ptychographic whole slide scanner for digital pathology applications. This handheld scanner is built using low-cost and off-the-shelf components, including red, green, and blue laser diodes for sample illumination, a modified stage for programmable sample positioning, and a synchronized image sensor pair for data acquisition. We smear a monolayer of goat blood cells on the main sensor for high-resolution lensless coded ptychographic imaging. The synchronized secondary sensor acts as a non-contact encoder for precisely tracking the absolute object position for ptychographic reconstruction. For WSI, we introduce a new phase-contrast-based focus metric for post-acquisition autofocusing of both stained and unstained specimens. We show that the scanner can resolve the 388-nm linewidth on the resolution target and acquire gigapixel images with a 14 mm × 11 mm area in ~70 seconds. The imaging performance is validated with regular stained pathology slides, unstained thyroid smears, and malaria-infected blood smears. The deep neural network developed in this study further enables high-throughput cytometric analysis using the recovered complex amplitude. The reported do-it-yourself scanner offers a portable solution to transform the high-end WSI system into one that can be made widely available at a low cost. The capability of high-throughput quantitative phase imaging may also find applications in rapid on-site evaluations.

Introduction

The process of analysing histology slides using a regular light microscope is the gold standard in diagnosing a large number of diseases, including almost all types of cancers. However, this process is highly subjective and suffers from inter- and intra-observer variations despite its success in diagnostic applications. The current workflow in a clinical setting is also labour-intensive and mostly qualitative. It can be easily disrupted when a pathologist bumps a slide to a high magnification objective lens or switches to a different objective lens.

Quantitative characterization of histopathology imagery is vital for reducing observer errors in diagnoses. It can also streamline and standardize the clinical workflow. Thus, whole slide imaging (WSI) systems were developed to replace conventional light microscopes for quantitative and accelerated histopathological analyses.[1] The first WSI platform was developed based on a robotic microscope in 1990s.[2] With imaging hardware and software developments, the regulatory field for digital pathology using WSI systems has advanced significantly in the past decade. A key milestone was accomplished in 2017 when the Philips' whole slide scanner was approved for primary diagnostic use in the U.S.[3] Furthermore, the recent advancement of artificial intelligence (AI) in medical diagnostics promises further growth of this field in the coming decades. In particular, various deep learning approaches have been demonstrated for automatic analysis of whole slide images with performance comparable to human experts.[4,5]

A typical whole slide scanner consists of the following hardware components:[6] 1) a microscope with high numerical aperture (NA) objective lenses and a tube lens, 2) a motorized x–y stage for rapid lateral scanning of the slide, 3) a precise motorized z-stage for axial movement and rapid focusing of the objective lens, 4) one or more image sensors for image acquisition and autofocusing, and 5) a high-power light source

[a] Department of Biomedical Engineering, University of Connecticut, Storrs, CT, 06269, USA. E-mail: chengfei.guo@uconn.edu, guoan.zheng@uconn.edu
[b] Pathology and Laboratory Medicine, University of Connecticut Health Centre, Farmington, CT, 06030, USA

† Electronic supplementary information (ESI) available: Fig. S1–S10, Tables S1 and S2, Movie S1. See DOI: https://doi.org/10.1039/d2lc00084a

‡ These authors contributed equally to this work.

for motion-frozen sample illumination. The resultant scanner enables comprehensive digital rending of entire histology slides. In addition, the colour information is often obtained from the Bayer colour filter of the image sensor. The stitched whole-slide digital image can then be viewed, shared, navigated, and annotated with speed and ease on a computer monitor.

Several challenges are associated with using the conventional whole slide scanners for high-throughput imaging. First, the depth of field of high-NA objective lenses is on the micron scale. This limitation poses a challenge for tracking the axial topography variations of tissue sections, as specimens not placed within the depth of field of the objective lens will result in degraded quality of the acquired images. Ultimately, this leads to rescanning and workflow delays. To address this challenge, many whole slide scanners create a focus map before the scanning process. For each focus point, a system needs to acquire a z-stack by axially moving the objective lens to different z positions. The best focus position can then be inferred based on the captured intensity image with the highest Brenner gradient or other focus metrics.[6] However, most focus metrics employed in these existing systems are only valid for adequately stained specimens. For weakly stained or unstained samples, intensity-based focus metrics would not be a good indicator for the best focus position. This same limitation also applies to whole slide scanners with a secondary camera for autofocusing, including the widely adopted systems from Leica and Philips.[6] In a later section, we will introduce a new focus metric based on the quantitative phase contrast that works for both stained and unstained samples.

Second, a high-NA objective lens often implies a small field of view.[7] To accommodate this limitation, current scanners need to perform precise mechanical scanning of the sample at high speeds. The resultant motion blur presents an issue during the image acquisition process. Current solutions include 1) using an expensive time-delay-integration image sensor with precise synchronization with the sample motion, 2) using pulsed illumination or specially-designed high-power light sources to freeze the motion during the acquisition process, and 3) mechanically pausing the motion for image acquisition. The first two solutions require special hardware with high premiums and thus, are not available for cost-effective implementations. Alternatively, the motion acceleration and deceleration involved within the third solution would substantially decrease the scanning speed.

Third, challenges exist when upscaling the imaging throughput for existing whole slide scanners. For example, one can upscale the number of lenses used when building a microscope array. In this respect, a lenslet array has been utilized to demonstrate rapid WSI.[8] However, the large footprint, high cost, precise optical alignment, and other hardware challenges have prevented the adoption of these lenses in clinical settings. Another option is to scale up the size of the objective lens to achieve a higher space-bandwidth product.[9] Giant objective lenses have been used to demonstrate high-throughput, large-scale microscopic imaging.[10,11] However, these bulky and highly specialized lenses are not readily available to researchers due to their high costs and the required specialized acquisition hardware.

Here we report the design and implementation of a handheld, lensless ptychographic whole slide scanner. This handheld scanner is built using low-cost and off-the-shelf components that operate without using any lenses. It offers a portable solution to transform the high-end WSI system into one that is utilizable without compromising the performance.

The reported scanner integrates several innovations into a miniatured system. First, a synchronized image sensor pair is used for data acquisition. The first sensor is smeared with a monolayer of goat blood cells for high-resolution ptychographic imaging. The second sensor is utilized for providing precise lateral positional feedback during the scanning process. In contrast, conventional whole slide scanner employs a secondary camera for axial positional tracking.[6,12] Second, the reported device implements a novel focus map generation scheme after the data has been acquired. In particular, we adopt a new focus metric based on the quantitative phase contrast of the specimen. This new focus metric can find the correct focus positions for both stained and unstained specimens post measurement. Third, colour-multiplexed scheme is adopted in the reported device for ptychographic reconstruction. A compact coherent illumination module is developed using low-cost red, green, and blue laser diodes. The high illumination flux from this module allows image acquisition while the sample is in continuous motion.

The integration of these innovations addresses the challenges of the conventional scanner discussed above. In contrast with the small depth of field of the conventional scanner, the lensless nature of the reported device does not impose any limit on the depth of field. After reconstruction, we can digitally propagate the recovered complex wavefront to any axial position post measurement. The scanning mechanism of the reported device is also different from the conventional scanner. The reported device has a large field of view and the scanning step size in-between adjacent acquisitions is on the micron scale. Therefore, the sample can be in continuous motion without extra hardware. Lastly, multiple image sensors can be employed in the reported device to scale up the imaging throughput.

For application demonstration, we acquired whole slide images of different specimens in this study. We show that the scanner can resolve the 388 nm linewidth on the resolution target and acquire gigapixel images with a 14 mm × 11 mm area in ~70 seconds. We also show that the recovered phase information can be used to measure the precise height map of unstained fine needle aspiration (FNA) smears, thereby providing a powerful tool to visualize the cellular topographic structure in 3D. We note that these thick FNA cytology smears contain many slow-varying 2π wraps that are challenging to obtain using other common imaging techniques. To demonstrate the diagnostic potential in

resource-limited settings, we performed automatic tracking of malaria-infected blood cells over the entire microscope slide. The recovered phase and phase gradient were used to locate the malaria parasites in the blood cells. Finally, we developed a deep neural network that uses the recovered complex amplitude to perform high-throughput cytometric analysis. We envision that the reported DIY miniature can be used as a personal whole slide scanner for every pathologist in their offices.

Materials and methods

Design of the ptychographic whole slide scanner

Fig. 1a shows the schematic and the operation principle of the ptychographic whole slide scanner. We designed a three-layer structure for this miniature. The bottom layer contains the low-cost red, green, and blue laser diodes for multi-wavelength coherent illumination. The laser beam is coupled to a single-mode fiber using an optomechanical coupler. Other components housed in this layer include the current drivers for the stepper motor, the current drivers and temperature management module for the laser diodes, and the micro-controller for programmable positioning and synchronization of the image sensor pair. The middle layer contains a homemade motorized stage for lateral sample positioning and a pair of image sensors for lensless data acquisition. The motorized stage is modified from a low-cost manual stage off-the-shelf (SEMY60-AS, Amazon). We used two stepper motors (NEMA 11) with integrated lead screws to translate the stage's metal plates mounted on the linear guided sliding track. Lastly, the top layer of the device contains the illumination module for the image sensor pair.

The ptychographic whole slide scanner prototype is shown in Fig. 1b. The compact size and the lensless operation allow us to hold the entire platform in one hand. Fig. 1c depicts different parts on the bottom and middle layers. Fig. 1d shows the laser module with the red (638 nm, 180 mW), green (520 nm, 120 mW), blue (445 nm, 500 mW) laser

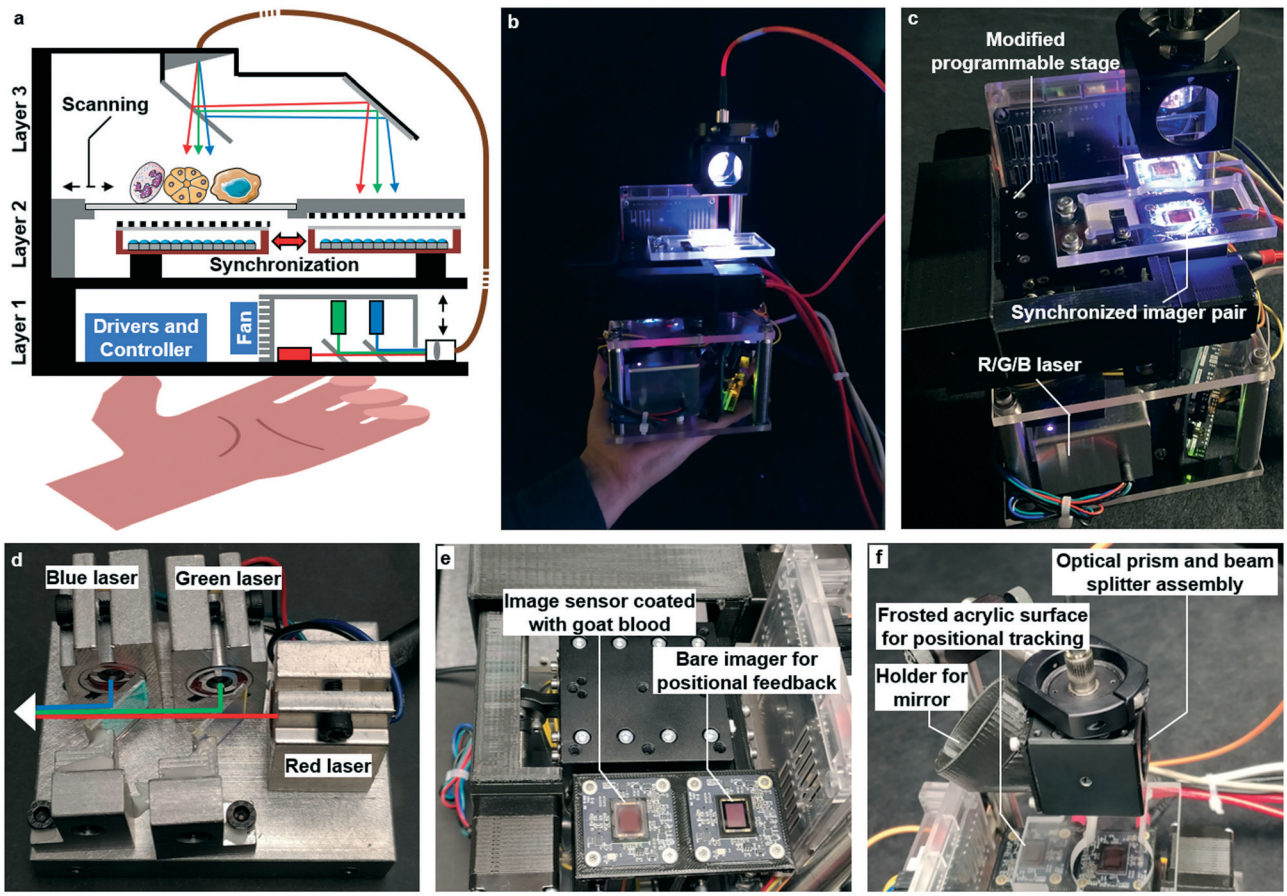

Fig. 1 The design and implementation of a handheld, multiplexed, and AI-powered ptychographic whole slide scanner. (a) The three-layer design of the system. Layer 1 contains the red, green, and blue laser diodes and the related optomechanical coupler to a single-mode fiber. It also contains the micro-controller, and the current drivers for the laser and the stepper motors. Layer 2 contains a modified *x*–*y* stage driven by two stepper motors for programmable positioning. A pair of synchronized image sensors are mounted on the stage: one for lensless ptychographic imaging and the other one as an absolute position encoder. Layer 3 contains the illumination module, including a prism for colour multiplexed illumination and a beam splitter for directing the laser beam to the secondary image sensor. (b) The handheld ptychographic whole slide scanner. (c) A zoomed-in view of different parts of the scanner. (d) The laser module at the bottom layer. (e) The synchronized image sensor pair at the middle layer. (f) The illumination module at the top layer. Refer to Fig. S1 and S2† for the system design and Movie S1† for the operation of the scanner.

diodes, and the corresponding dichroic mirrors. The use of high-power laser diodes enables continuous sample motion during the image acquisition process. The motion blur issue in conventional whole slide scanners is absent from the reported platform.

In Fig. S1a,† we used the digital pin from the microcontroller to control the transistor–transistor logic of the laser driver and adjusted the output power of the laser diodes accordingly. Fig. S1b† shows the design of the optomechanical coupler for directing the laser light to the single-mode fiber. Fig. S1c† shows a magnified view of the synchronized image sensor pair, where a monolayer of goat blood cells is smeared on top of the main image sensor (discussed in the next section). In Fig. S2,† we show the mechanical design of the scanning stage. Fig. S2a† shows the original low-cost stage with two manual actuators. We removed the actuators (Fig. S2b†) in our design and replaced them with two stepper motors for programmable control (Fig. S2c and d†). With the stepper motors, the maximum translation speed is ~3 mm s^{-1}. This simple modification enables the development of a handheld whole slide scanner for both lensless ptychographic imaging discussed here and regular lens-based WSI platforms.[13] Fig. 1e shows two 3D printed plastic cases that cover the lead screws of the two stepper motors.

Synchronized imager pair and image-based position encoder

A camera pair is often used for data acquisition in conventional whole slide scanners. This pair is composed of the main camera for high-resolution image acquisition and a secondary camera to track the defocus position of the specimen.[6,12] In particular, the most successful whole slide scanners from Philips and Leica place a tilted sensor at the conjugated image plane to enable autofocusing during the scanning process. This sensor generates a contrast curve based on the intensity gradient of the captured image and determines the best focus position for the main camera.[6]

The reported platform also adopts a synchronized image sensor pair (IMX226, 1.85 μm pixel size) for data acquisition. Our system utilizes the main sensor for lensless diffraction data acquisition and a secondary sensor to track the absolute position of the scanning stage. We synchronized the main and secondary image sensors by sending the triggering signal from the micro-controller.

For the main image sensor shown in Fig. 1e, we smeared a monolayer of blood cells as a coded scattering lens.[14–21] We also fixed the blood cells using alcohol to prevent their degradation over time. As we will discuss in the next section, we tested blood cells from both goat and fish in the experiment. The goat blood was obtained from Lampire Biological Laboratories and the fish blood was obtained from a supermarket. The coated monolayer layer of blood cells can down-modulate the diffracted light waves for sensor detection. The concept is similar to the operation of structured illumination microscopy,[22,23] where non-uniform structured light patterns are used to down-modulate the object information for detection. In the reported device, the monolayer of blood cells redirects the light waves with large diffraction angles into smaller angles, thereby improving the achievable resolution of the recovered images.

Robust and high-performance translational scanning often requires the use of encoders for measuring the absolute position of the scanning stage. Without the encoder, the open-loop operation would lose track of the position after several runs. This is a drawback in some of the previous demonstrations on low-cost microscope motorized stages.[13,24–27] In our implementation, a secondary image sensor is used as a non-contact position encoder for tracking the absolute position of the stage. As shown in Fig. 1f, we place this sensor under an acrylic slide holder with a frosted surface. When illuminated with a laser beam, the frosted surface generates a speckle pattern on the sensor. By scanning the slide holder to different lateral positions, we can recover the absolute position of the stage *via* cross-correlation analysis[19,20,28] with sub-pixel accuracy.[29] Fig. S3† shows the tracking performance using the secondary sensor. In this experiment, we used two motorized stages (ASI LS-50) to move the slide holder to different x–y positions. The recovered positions using the secondary sensor are plotted with the positions returned by the motorized stages. The average difference is ~100 nm. We note that part of the difference comes from the mechanical repeatability and backlash of the motorized stages.

We note that it would be challenging to retrieve this positional shift if a secondary sensor was not present. While one can produce a transparent region on the main sensor for positional tracking, the sample slide may not contain any detailed features on this region. The tracking would thus be subjected to the histology slide itself and would not be reliable when imaging the edge area. The empty region on the main sensor also reduces the effective field of view for lensless imaging. In contrast, the reported device performs positional tracking based on the frosted surface of the slide holder without relying on the object profile. One can precisely track the absolute position even if the object is an empty glass slide.

Multiplexed data acquisition and ptychographic reconstruction

A user can turn on one or more laser diodes for sample illumination to initiate the operation of the reported scanner (Movie S1†). The slide holder will be scanned at different lateral positions during the acquisition process. The imaging model can be explained as the following sequence: First, the histology slide is translated to the ith lateral position (x_i, y_i), and two images are acquired using both the main and the secondary sensors. Second, the image captured using the secondary sensor is used to recover (x_i, y_i) by performing a cross-correlation analysis with the reference image captured at (x_1, y_1). Third, the translated object profile $O(x–x_i, y–y_i)$ propagates to the blood-cell monolayer on the main sensor,

multiplies with this layer, and propagates to the surface of the pixel array, obtaining the complex lightwave $\psi_i(x, y)$. Fourth, the pixel array samples the complex light wave with both spatial and angular filtering, resulting in the captured intensity image of $I_i(x, y)$. The goal of the reconstruction process is to recover the complex object profile O based on all intensity measurements I_i (i = 1, 2, 3…).

In our implementation, we continuously translate the stage at a speed of ~60 μm s^{-1} and acquire the images at the full camera framerate at 30 frames per second. For one field of view of the image sensor, we typically acquire 450 images in 15 seconds. The stage then positions to the next field of view and repeats the process. For an area of 14 mm × 11 mm on a regular slide, we typically acquire 4 fields of view of the sensor and the acquisition time is ~70 seconds, including 10 seconds for motion overhead.

At the heart of our reconstruction process is an imaging technique termed ptychography.[30] The original concept was developed to address the missing phase problem in electron microscopy.[31] Its modern form was brought to fruition by adopting the iterative phase retrieval framework.[32] In a typical lensless implementation, the specimen is laterally translated through a spatially confined probe beam and the lensless diffraction patterns are recorded at the reciprocal space. It is also possible to swap these two spaces using a lens-based setup.[33,34]

The reconstruction process of ptychography iteratively imposes two different sets of constraints. First, the diffraction measurements serve as the Fourier magnitude constraints in the reciprocal space. Second, the confined probe beam limits the physical extent of the object for each measurement and serves as the support constraint in the real space. For the reported platform, we replaced the confined probe beam in conventional ptychography with the unconfined transmission profile of the blood-cell layer. We further employed a full-field illumination scheme to cover the entire sample for high-throughput optical imaging.[20,28]

A single wavelength is often adequate to obtain the whole slide image of unstained samples or samples stained with only two complementary colours. Using the recovered intensity and phase information, one can virtually stain the image according to different staining protocols post-measurement.[21,24,35–39] When working with samples with multiple stains, we can turn on the red, green, and blue laser sequentially. The recovered images can then be combined to form a colour whole slide image.[33]

Another strategy for colour imaging is to perform multiplexed illumination by turning on the red, green, and blue laser diodes simultaneously (Fig. 1b). In turn, the captured data becomes an incoherent summation of the diffraction patterns at three wavelengths. We can then recover the complex object profiles at the three wavelengths using a multiplexed ptychographic reconstruction strategy.[40–43] Fig. 1f shows the illumination module of the device, where the laser light from the single-mode fiber is split for both the main and the secondary image sensors. At the illumination path of the main sensor, we added a prism in-between the beam splitter and the specimen in Fig. 1f. Consequently, the incident angles of the laser beams are slightly different at different wavelengths. The resulting complex profiles of the blood-cell monolayer become uncorrelated at different wavelengths, thus breaking the ambiguities in mixed-state reconstruction[44] and facilitating colour-multiplexed ptychographic reconstruction.[40–43] Fig. S4† shows the procedure of the reconstruction scheme. The number of the colour channels reduces to one in Fig. S4† when we illuminate the sample with one wavelength.

A new focus metric for both stained and unstained specimens

Many commercially-available whole slide scanners generate a focus map prior to the scanning process. For each focus point, the system acquires a z-stack image of the slide by translating the objective lens to different z-positions. A specific figure of merit for each image is extracted to measure the quality of the estimated focus position. Conventional scanners often employ the image contrast or gradient of the captured intensity data as the focus metric. The logic behind this choice is that the captured intensity information should demonstrate the largest image contrast when the specimen is placed at the in-focus position. This logic, however, is only correct for the stained tissue sections. For unstained or weakly-stained sections, we show that the intensity contrast is not a proper indicator for the in-focus position. As a result, challenges arise when utilizing the conventional whole slides scanners to image unstained or weakly-stained specimens. To address this issue, we introduce a new focus metric based on the quantitative phase-gradient contrast of the recovered images. As we will discuss in a later section, this new focus metric works for both stained and unstained specimens.

Tracking malaria-infected red blood cells

We also demonstrate the use of the recovered phase to locate malaria-infected blood cells over the entire blood smear slide (*Plasmodium falciparum* smear from Carolina Biological Supply). In the segmentation process, we set a threshold of 1.2 radians for the recovered phase and a threshold of 0.42 radians per pixel for the phase gradient. The regions with the phase and phase gradient larger than these two thresholds were treated as candidates for malaria parasites. These two threshold values were chosen based on an analysis of more than 500 parasites from the recovered images. We then further examine the size of the candidate to rule out the large white blood cells.

AI-powered cytometric analysis

As we will discuss in a later section, we adopt a deep neural network to segment different cells of a slide stained with immuno-histochemical biomarkers. The segmented results allow us to perform cytometric analysis with an effective image acquisition throughput of >13 000 cells per second.

Results

Imaging performance characterization

We first validated the imaging performance of the reported scanner using a resolution target. In Fig. 2a–c, we show the results of using a monolayer of goat blood cells as the coded scattering lens. The cells were smeared directly on top of the coverglass of the main sensor. Once smeared, they are firmly attached to the surface and the transmission profile remains unchanged for a long period. In addition, this monolayer of blood cells allows the formation of a dense and thin layer with both intensity and phase modulation. Goat blood was chosen for our implementation because of its smallest cell size among all animals (2–3 μm). As a comparison, we also tested the fish blood cells as the coded scattering lens in Fig. 2d–f. Fig. 2a and d show the captured raw images of the resolution target using these two different types of blood samples. Fig. 2b and e show the recovered image of the resolution target. The recovered blood-cell monolayer profiles are shown in Fig. 2c and f. We can resolve the 388 nm linewidth using the goat blood cells and the 488 nm linewidth using the fish blood cells, corresponding to 776 nm and 976 nm full pitch resolution, respectively. While it is possible to use other microparticles or disorder-engineered surfaces to form the thin scattering lens,[14,15,19–21,45] smearing blood cells is easy to operate yet highly effective, enabling a DIY option for implementing lensless coded ptychographic microscopy. Disorder-engineered surfaces also requires sophisticated patterning and etching tools for fabrication.

Post-acquisition autofocusing *via* a phase-contrast metric

Once we recover the object exit wavefront from the diffraction measurements, we can digitally propagate it back to different axial positions for post-acquisition refocusing.[20] Unlike the conventional whole slide scanner, this process is performed computationally without involving any mechanical translation. In Fig. S5,† we refocused the recovered object exit wavefront to 5 different axial positions, with the in-focus position set as $z = 0$. For the hematoxylin–eosin (H&E) stained slide in Fig. S5a,† the refocused intensity and phase images show the highest contrast at the in-focus position. However, the unstained thyroid smear in Fig. S5b† exhibits the lowest contrast at the in-focus position. Therefore, the intensity contrast is not a good focus indicator for WSI, especially for weakly-stained or unstained slides.

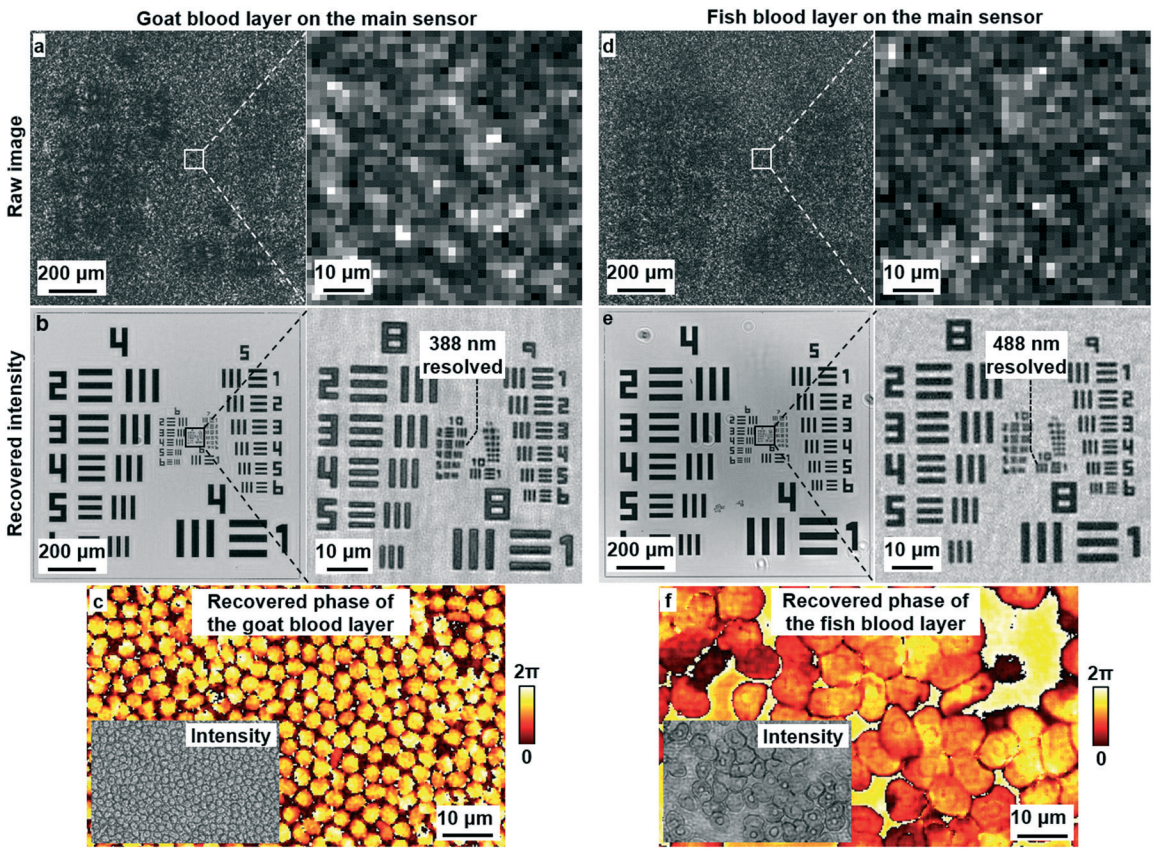

Fig. 2 Imaging performance characterization. (a) The captured raw image using a monolayer of goat blood cells as the coded scattering lens. The goat blood cells were smeared on top of the main image sensor and fixed with alcohol. (b) The reconstruction of (a). (c) The recovered transmission profile of the monolayer goat blood cells on the main sensor. (d) The captured raw image using a monolayer of fish blood cells as the coded scattering lens. (e) The reconstruction of (d). (f) The recovered transmission profile of the monolayer fish blood cells on the main sensor.

In Fig. 3a, we calculated the Brenner gradient[46] of the refocused intensity and phase images of the H&E stained slide. As a comparison, we show the same calculation of the unstained thyroid smear in Fig. 3b. We can see that the maximum points of the phase gradient correlate well with the in-focus positions of both slides. In contrast, the intensity gradient reaches a minimum point when the unstained slide is placed at the in-focus position. In the following, we adopt the Brenner gradient of the recovered phase image as the new focus metric for post-acquisition autofocusing. In our test, the axial range of this focus metric is ~80 μm. Beyond this range, the focus metric becomes saturated and provides little information on the focus position.

Fig. 3c shows the adaptive focus map generation scheme post-measurement. In the first iteration, we identified 4 focus points at the edge of the field of view, labelled by the blue dots. For each focus point, we refocused a small region (256 × 256 pixels) of the corresponding object exit waves to different axial positions. The searching process started with a coarse *z*-step size of 10 μm with *z*-range of 500 μm. The best focus position was determined by maximizing the phase-contrast metric. We then reduced the *z*-step size to 1 μm and performed a detailed search at a range of (−9 μm, +9 μm) with respect to the identified coarse focus position.

At the second iteration, we identified 5 focus points labelled by the red dots in Fig. 3c. We then estimated the in-focus positions of these red dots *via* a bilinear interpolation of the 4 focus points from the first iteration. The axial range was reduced to ~250 μm for searching the focus points of these red dots. For the entire field of view of the image sensor, we typically assign 13 focus points and perform 3 iterations of the searching process. The final focus map is generated based on a 2D interpolation of the identified focus points.

Fig. 4a shows the refocused image of the H&E-stained slide based on the interpolated focus map generated in Fig. 4b.

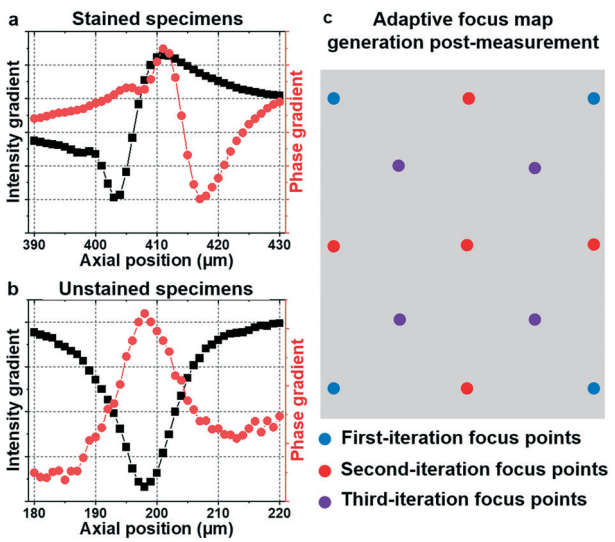

Fig. 3 The intensity and phase gradient of the stained slide (a) and unstained slide (b). (c) Adaptive focus map generation post-measurement.

Fig. 4c shows the recovered intensity images of the object exit waves at the plane of the blood monolayer. Fig. 4d and e show the corresponding refocused intensity and phase images of the two regions based on the focus map.

Quantitative phase imaging of thick unstained slides

For conventional in-line holography or multiple-height phase retrieval, it is challenging to restore the slow-varying phase profiles with many 2π wraps.[15] This drawback can be explained by the phase transfer function that characterizes the transfer property of the phase contents at different spatial frequencies.[47] For low spatial frequency contents, the phase transfer function is close to 0 for regular in-line lensless imaging setups.[47] Therefore, the corresponding phase information cannot be effectively converted into intensity variations for detection, and the recovered phase is not quantitative. Likewise, this same challenge also applies to blind ptychography where both the complex probe beam and the object profile need to be jointly recovered in the phase retrieval process.[48–51]

The reported platform encodes the slow-varying phase information into the spatial distortions of the intensity diffraction patterns. The operation of the blood-cell monolayer can be viewed as a special case of the microlens array on top of the Shack–Hartmann sensor,[52] which converts the slightly tilted wavefront into spatial displacements of the focused spots. To demonstrate its application in quantitative phase imaging, we acquired a whole slide phase image of the unstained thyroid smear in Fig. 5a. We can see that there are many regions containing slow-varying phase profiles in the right panel of Fig. 5a. Fig. 5b1 shows a zoomed-in view of the unstained smear, where we can recover the phase profile with many 2π wraps. Fig. 5b2 shows the unwrapped phase of Fig. 5b1. In Fig. S6,† we also compare the recovered intensity and phase of a small region of the unstained thyroid smear, where we can clearly see the cell cluster from the unwrapped phase but not from the intensity image. The unique capability of high-throughput quantitative phase imaging using the reported device can find applications in label-free rapid on-site evaluations of biopsy samples. It can also be used in other label-free biomedical applications.[53]

Assuming a refraction index of 1.43 for thyroid smear,[54] we can recover the height map of the slide in Fig. 5c. In Fig. S7,† we further show the difference between the captured raw image (Fig. S7a†), the recovered phase (Fig. S7b†), the unwrapped phase (Fig. S7c†), and the reference images captured using a regular light microscope (Fig. S7d†).

Tracking of malaria-infected blood cells

The reported device can benefit diagnosis in challenging environments, where access to physicians and good equipment can be limited. To demonstrate this point, we applied the prototype device to image a blood smear sample infected with *Plasmodium falciparum*, a particularly harmful type of malaria parasite and one of the major causes of death in the developing

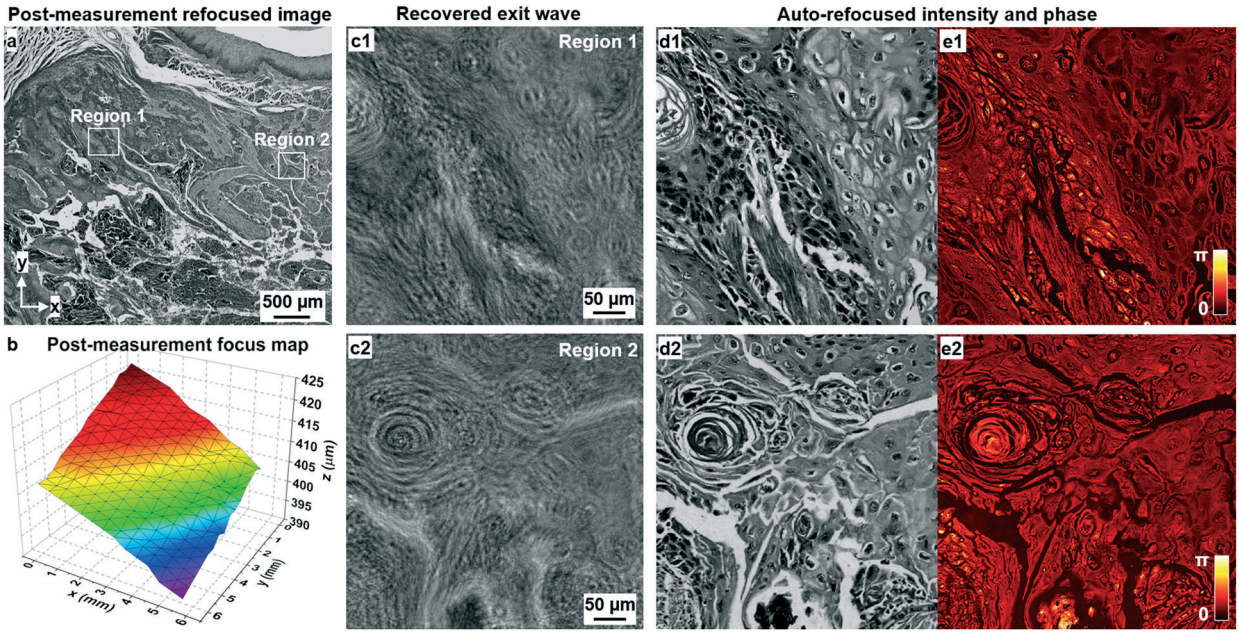

Fig. 4 (a) The refocused whole slide image using the Brenner gradient of the recovered phase as the focus metric. (b) The generated focus map. (c1 and c2) The recovered intensity of the object exit waves of regions 1 and 2 labelled in (a). The corresponding refocused intensity (d1 and d2) and phase images (e1 and e2) of the exit waves.

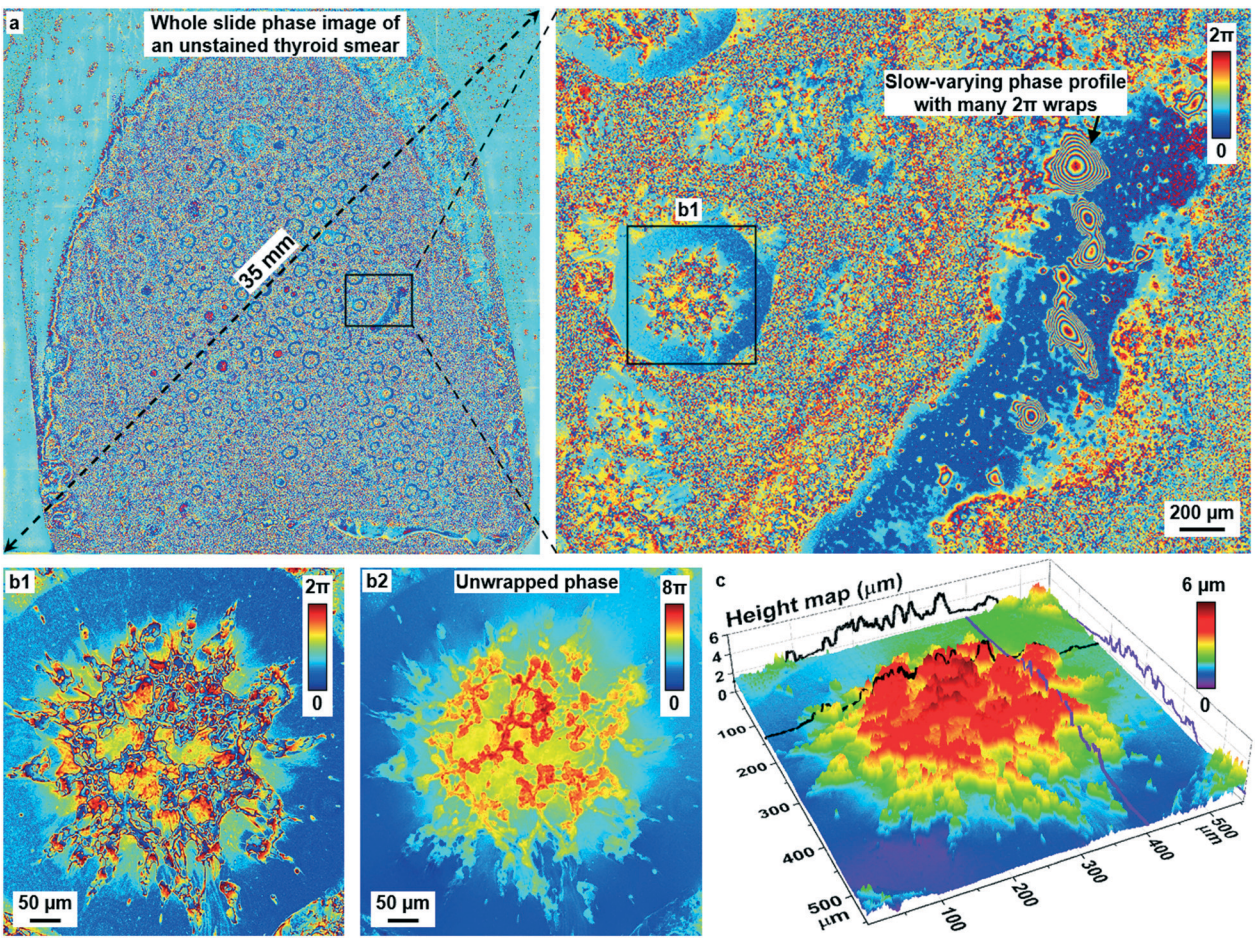

Fig. 5 The refocused whole slide phase image of an unstained thyroid smear. (b1) The zoomed-in view of (a) with many 2π phase wraps. (b2) The unwrapped phase of (b1). (c) The recovered height map of the smear.

world. Fig. 6a shows the recovered phase image of the entire blood smear. The inset of Fig. 6a shows the locations of the malaria parasites tracked using the reported device. Fig. 6b shows the zoomed-in view of the recovered phase, with the malaria parasites highlighted by the blue arrows. Sample images of malaria-infected blood cells are shown in Fig. 6c. In Fig. S8,† we also show the input phase image and the resultant mask for the parasite candidates.

Multiplexed illumination for lensless colour imaging

Fig. 7 shows the results of multiplexed illumination for colour imaging, where we turn on all three laser diodes and acquire the coherent state mixture for data reconstruction.[40–43] This process requires the blood-cell monolayer to be calibrated at the three wavelengths upfront. In our implementation, we used a human blood smear as the object in the calibration process. For each wavelength, we acquired 1500 images and jointly recovered both the human blood smear and the coded layer of goat blood cells on the main sensor. The recovered coded layers at different wavelengths were then enforced in the multiplexed reconstruction detailed in Fig. S3.† Fig. 7a shows the recovered whole slide colour image based on 600 raw measurements captured under the colour-multiplexed illumination. Fig. 7b shows the recovered phase at the green wavelength. The zoomed-in views of the two regions are shown in Fig. 7c. The colour images captured using a regular light microscope are shown in Fig. 7d as a reference. In Fig. S9,† we also show the recovered phase and intensity images at the three wavelengths.

AI-powered cytometric analysis

Ki-67 biomarker is a proliferation-associated nuclear protein that is only detected in dividing cells. The fraction of Ki-67 positive tumour cells is often correlated with the clinical course of cancer. In this study, we develop a deep neural network to perform cytometric analysis of Ki-67 cells using the recovered complex object amplitude. Fig. 8 shows the training and inference workflow of the network. In the training stage, we first recovered the intensity and phase images of the immunohistochemical slides labelled with the Ki-67 biomarkers (at the green wavelength). We then acquired the colour image of the same slides using a regular light microscope with a 40×/0.95 NA objective lens. A feature-based algorithm was implemented for image registration between these two sets of images.[55]

For colour images acquired using the regular light microscope, we first performed colour deconvolution[56] to separate the colour information into a brown channel and a blue channel, corresponding to the Ki-67 positive and

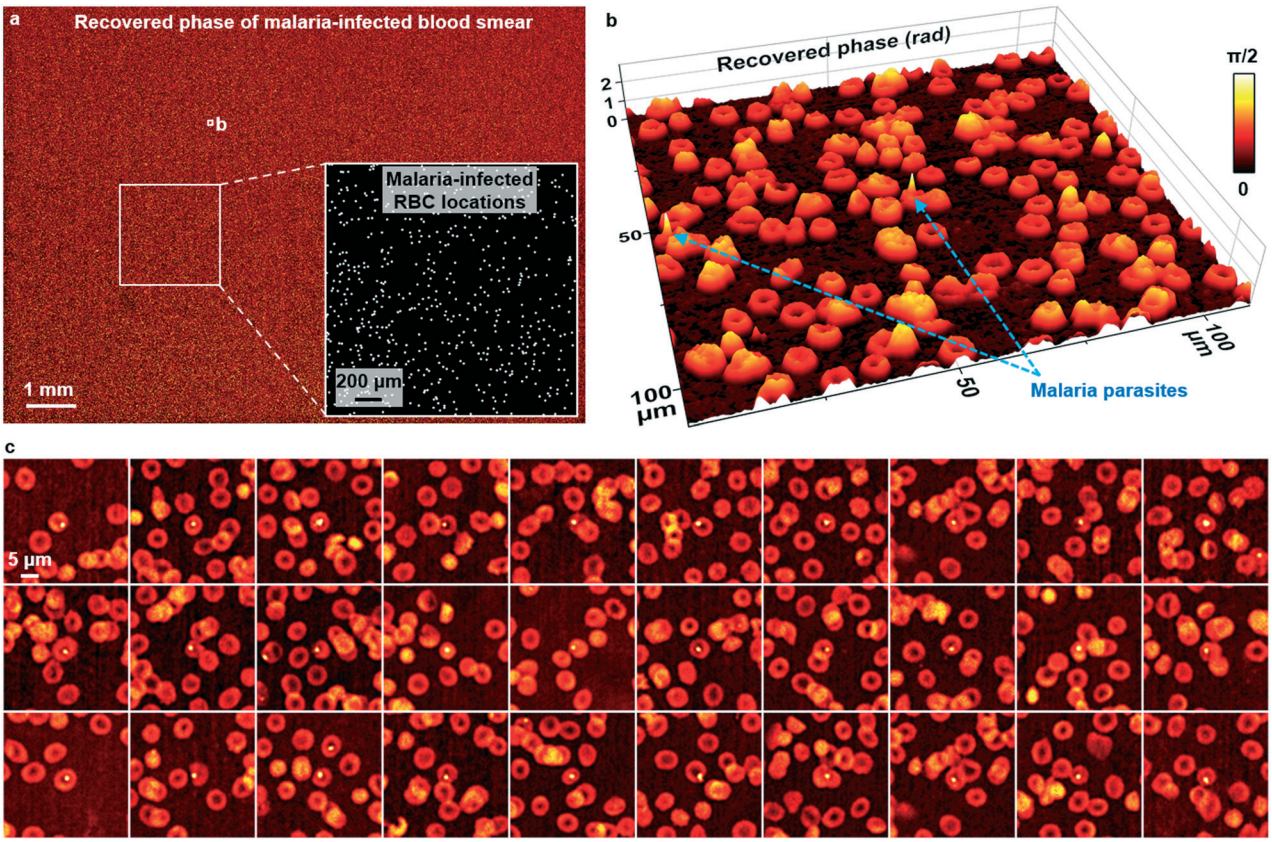

Fig. 6 Tracking malaria-infected blood cells over the entire microscope slide. (a) The recovered phase image of the sample. (b) The zoomed-in view of the recovered phase. (c) Sample images of malaria-infected blood cells located by the reported device.

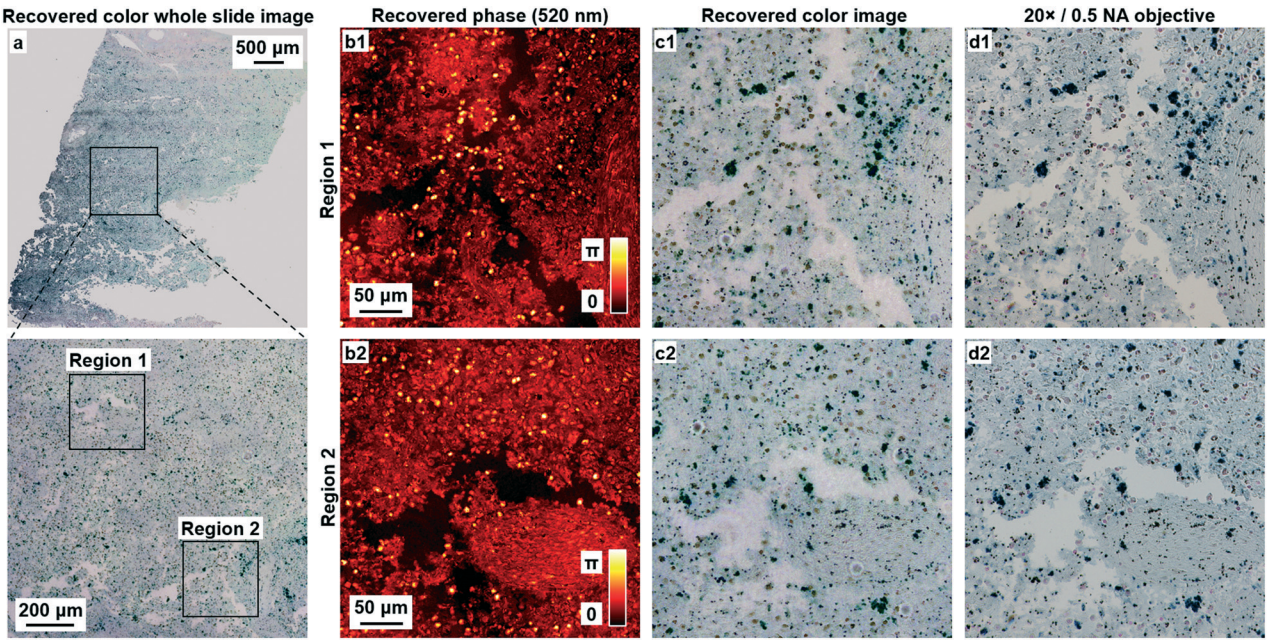

Fig. 7 Multiplexed illumination for colour imaging. (a) The recovered colour whole slide image. (b1 and b2) The recovered phase of regions 1 and 2 at the green wavelength. (c1 and c2) The zoomed-in views of the recovered colour images of regions 1 and 2. (d1 and d2) The corresponding images of regions 1 and 2 captured using a 20× objective.

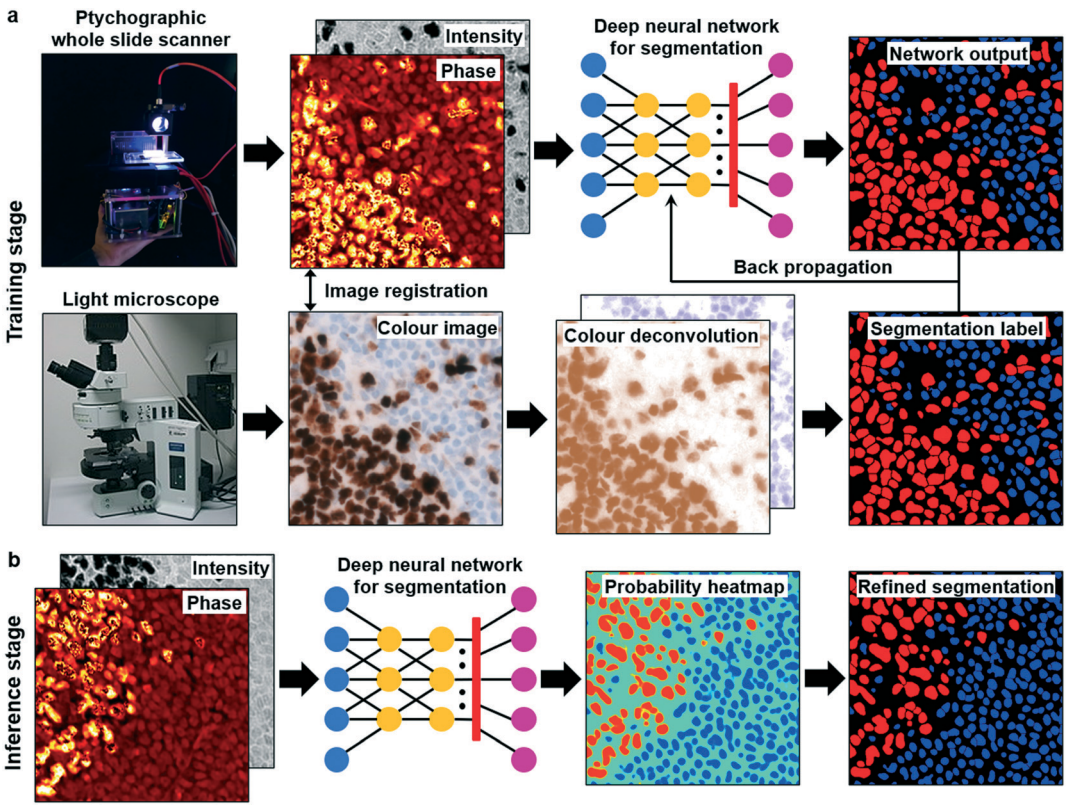

Fig. 8 Development of a deep neural network for automatic cell segmentation and counting based on the recovered intensity and phase images. (a) In the training stage, we minimized the difference between the network output and the ground-truth segmentation labels obtained from a regular light microscope. (b) The output of the network is a mask with the segmented positive cells, negative cells, and background of the slide.

negative cells, respectively. Based on the deconvoluted colour channels, we then performed marker-controlled watershed segmentation[57] to extract the cells and obtain the segmentation masks as the ground truth labels.

The employed deep neural network was built based on the DeepLabv3+ structure[58] with a ResNet-18 backbone.[59] This neural network can encode multi-scale contextual information by probing the incoming features at multiple rates and multiple effective fields of view. The input for this network contains two channels: the recovered intensity and phase images from the ptychographic scanner. The network output is a pixel-level probability map of the Ki-67 positive cells, negative cells, and the background of the slide. We trained the network by minimizing the difference between the network output and the microscope-based segmentation labels shown in Fig. 8a. Our training dataset includes 1500 aligned image pairs with a lateral dimension of 512 × 512 pixels each. We used the stochastic gradient descent approach with momentum acceleration for minimizing the loss function. The mini-batch size was set to 8 and the learning rate was set to 0.01. The optimization process took 200 epochs before the validation loss became saturated.

Fig. 8b shows the inference workflow, where the trained network generates a probability heatmap on the positive cells, negative cells, and the background. This heatmap is further processed by the watershed algorithm followed by image binarization. The final output is a mask with the segmented positive cells, negative cells, and the background of the slide.

Fig. 9a shows the recovered whole slide intensity and phase images of a slide labelled with the Ki-67 biomarkers. This slide has not been used in the training process. Fig. 9b shows the segmentation results using the deep neural network. The zoomed-in views of Fig. 9a2 and b are shown in Fig. 9c. With more than 200 000 cells identified in Fig. 9b, the effective imaging throughput is >13 000 cells per second in this demonstration.

In Fig. S10a–d,† we further show the zoomed-in images of the two regions and their corresponding segmentation results. In Fig. S10e,† we compared the segmentation results with the manual counting results using the 40×, 0.95 NA objective lens. The average counting difference between the deep neural network and the objective lens is ~2.6% in this experiment. In Table S1,† we analysed the precision, recall, accuracy, and F1 score of the cell classification results using the network. In Table S2,† we quantified the morphological

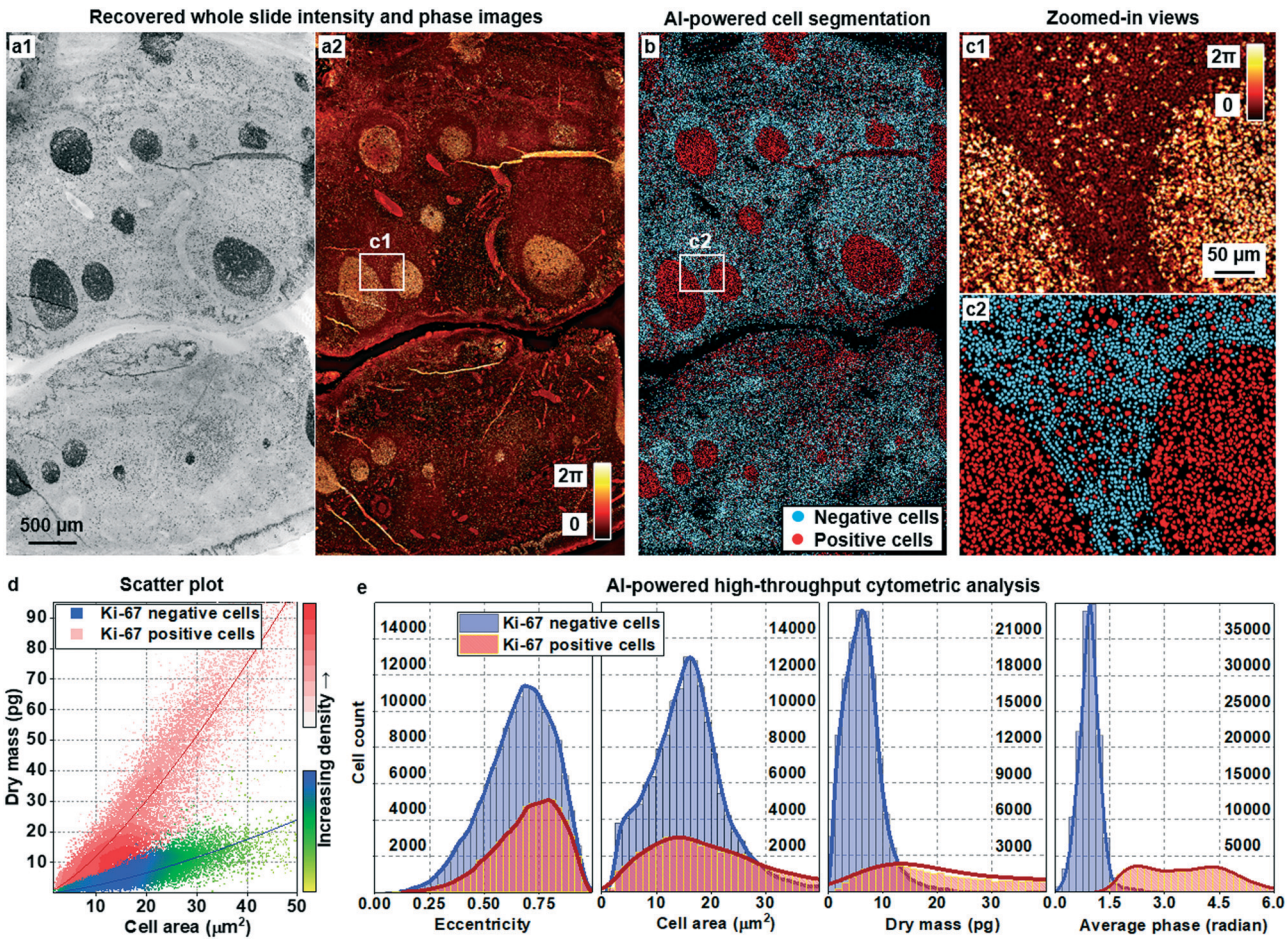

Fig. 9 AI-powered cytometric analysis using the ptychographic whole slide scanner. The recovered intensity (a1) and phase (a2) images of a slide labelled with the Ki-67 markers. (b) The segmentation results using the deep neural network. (c1 and c2) Zoomed-in views of the highlighted regions in (a2) and (b). (d) The measurement of dry mass and cell area for the Ki-67 positive and negative cells. (e) The histogram analysis of the cell eccentricity, cell area, dry mass, and average phase.

parameters with respect to the ground-truth results obtained using the 40×, 0.95 NA objective lens.

In Fig. 9d, we plot the measurement of dry mass and cell area for the segmented Ki-67 positive and negative cells. The dry mass (nonaqueous content) of the cell is linearly proportional to the optical phase shift accumulated through the cell. The dry mass density at each pixel can be calculated as $\rho(x,y) = (\lambda/2\pi\gamma)\phi(x,y)$, where λ is the center wavelength, γ is the average refractive increment of protein (0.2 mL g^{-1}), and $\phi(x,y)$ is the measured phase.[60] The total dry mass can then be obtained by integrating the density over the segmented cell. Due to the proliferation of the positive cells, the dry mass and cell area are higher than those of negative cells. However, we also note that Ki-67 biomarker may introduce additional mass to the cells. Therefore, the measured dry mass might not reflect the intrinsic cell mass without the markers. The cell mass change before and after biomarker labelling is an important research topic in the future.

In Fig. 9e, we plot the histogram analysis of the cell eccentricity, cell area, dry mass, and average phase for both the positive and negative cells. This demonstration shows the metrological versatility of AI-powered quantitative high-throughput cytometric analysis. The full clinical implications of the cell morphology distribution require further investigations.

Discussion and conclusion

In summary, we have demonstrated the design and implementation of a handheld lensless ptychographic whole slide scanner for digital pathology applications. In our design, we employed a synchronized image sensor pair for data acquisition. The main sensor was smeared with a layer of goat blood cells for coded ptychographic imaging. We found that the goat blood cells give us an excellent imaging performance as the cell size is the smallest among different animals. By using this goat-blood-cell encoder, we show that the reported scanner can resolve the 388-nm linewidth on the resolution target. The secondary sensor in our scanner acts as a non-contact image-based position encoder for tracking the absolute position of the stage. Our scanning stage is modified from a low-cost manual stage off-the-shelf. We replaced the original manual actuators with two stepper motors with integrated lead screws. When operated with the image-based position encoder, this simple modification offers a portable solution for WSI. Compared with the previous ptychographic implementations, we can precisely track the absolute position even if the object is an empty glass slide.

Conventional whole slide scanners often rely on maximizing the intensity-contrast for autofocusing. For unstained or weakly-stained slides, the intensity contrast, however, is not a proper indicator for the in-focus position. In this study, we introduce a new focus metric based on the quantitative phase contrast of the recovered images. We show that this new metric works for both stained and unstained specimens. To validate the performance of the reported scanner, we image various stained and unstained specimens. High-resolution gigapixel images with an area of 14 mm × 11 mm can be acquired in ~70 seconds. The achieved resolution and the image acquisition time are similar to those of conventional lens-based whole slide scanners.

One unique feature of the reported device is the recovery of quantitative phase information. We show that the recovered phase can be used to measure the height map of the unstained thyroid smear samples, providing a tool to visualize the cellular topographic structure in 3D. To demonstrate the application in resource-limited settings, we performed automatic tracking of malaria-infected blood cells using the recovered phase. Lastly, a deep neural network was developed for high-throughput cytometric analysis. This network takes the recovered complex amplitude as the input. The network output is a pixel-level probability map of different cells. With this network, we quantify millions of cells on immuno-histochemical-stained slides labelled by a proliferation-associated biomarker and demonstrate an effective image acquisition speed of >13 000 cells per second. We also demonstrate the metrological versatility of the reported platform for analysing the cell eccentricity, cell area, dry mass, and average phase.

One future research direction is to improve and optimize the phase sensitivity of the reported approach. The detection limit of a small phase variation is related to the complex profile of the encoder on the sensor. How to design an engineered surface to maximize phase sensitivity is an important research topic. A shorter distance between the object and the encoder may also help to improve the phase sensitivity limit.

Author contributions

Shaowei Jiang: methodology, validation, investigation, writing – original draft, visualization. Chengfei Guo: methodology, validation, investigation, writing – original draft, visualization. Pengming Song: methodology, validation, investigation. Tianbo Wang: methodology, validation, investigation. Ruihai Wang: methodology, validation, investigation. Terrance Zhang: investigation, writing – review & editing. Qian Wu: methodology, writing – review & editing. Rishikesh Pandey: methodology, writing – review & editing. Guoan Zheng: conceptualization, methodology, resources, supervision, validation, writing – original draft, writing – review & editing, funding acquisition, project administration.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

G. Z. acknowledges the support of the National Science Foundation 2012140. P. S. acknowledges the support of the Thermo Fisher Scientific fellowship.

Notes and references

1 F. Ghaznavi, A. Evans, A. Madabhushi and M. Feldman, *Annu. Rev. Pathol.: Mech. Dis.*, 2013, **8**, 331–359.
2 R. Ferreira, B. Moon, J. Humphries, A. Sussman, J. Saltz, R. Miller and A. Demarzo, *Proceedings of the AMIA Annual Fall Symposium*, 1997, p. 449.
3 E. Abels and L. Pantanowitz, *J. Pathol. Inform.*, 2017, **8**, 23.
4 A. Janowczyk and A. Madabhushi, *J. Pathol. Inform.*, 2016, **7**, 29.
5 A. Madabhushi and G. Lee, *Med. Image Anal.*, 2016, **33**, 170–175.
6 Z. Bian, C. Guo, S. Jiang, J. Zhu, R. Wang, P. Song, Z. Zhang, K. Hoshino and G. Zheng, *J. Biophotonics*, 2020, **13**, e202000227.
7 J. Park, D. J. Brady, G. Zheng, L. Tian and L. Gao, *Adv. Photonics*, 2021, **3**, 044001.
8 R. S. Weinstein, M. R. Descour, C. Liang, G. Barker, K. M. Scott, L. Richter, E. A. Krupinski, A. K. Bhattacharyya, J. R. Davis and A. R. Graham, *Hum. Pathol.*, 2004, **35**, 1303–1314.
9 A. W. Lohmann, *Appl. Opt.*, 1989, **28**, 4996–4998.
10 G. McConnell, J. Trägårdh, R. Amor, J. Dempster, E. Reid and W. B. Amos, *eLife*, 2016, **5**, e18659.
11 J. Fan, J. Suo, J. Wu, H. Xie, Y. Shen, F. Chen, G. Wang, L. Cao, G. Jin and Q. He, *Nat. Photonics*, 2019, **13**, 809–816.
12 M. C. Montalto, R. R. McKay and R. J. Filkins, *J. Pathol. Inform.*, 2011, **2**, 44.
13 C. Guo, Z. Bian, S. Jiang, M. Murphy, J. Zhu, R. Wang, P. Song, X. Shao, Y. Zhang and G. Zheng, *Opt. Lett.*, 2020, **45**, 260–263.
14 Y. Choi, C. Yoon, M. Kim, W. Choi and W. Choi, *IEEE J. Sel. Top. Quantum Electron.*, 2013, **20**, 61–73.
15 C. Guo, S. Jiang, P. Song, T. Wang, X. Shao, Z. Zhang and G. Zheng, *Biomed. Opt. Express*, 2021, **12**, 7173–7184.
16 E. G. van Putten, D. Akbulut, J. Bertolotti, W. L. Vos, A. Lagendijk and A. Mosk, *Phys. Rev. Lett.*, 2011, **106**, 193905.
17 H. Yilmaz, E. G. van Putten, J. Bertolotti, A. Lagendijk, W. L. Vos and A. P. Mosk, *Optica*, 2015, **2**, 424–429.
18 Y. Choi, T. D. Yang, C. Fang-Yen, P. Kang, K. J. Lee, R. R. Dasari, M. S. Feld and W. Choi, *Phys. Rev. Lett.*, 2011, **107**, 023902.
19 Z. Bian, S. Jiang, P. Song, H. Zhang, P. Hoveida, K. Hoshino and G. Zheng, *J. Phys. D: Appl. Phys.*, 2019, **53**, 014005.
20 P. Song, S. Jiang, H. Zhang, Z. Bian, C. Guo, K. Hoshino and G. Zheng, *Opt. Lett.*, 2019, **44**, 3645–3648.
21 S. Jiang, C. Guo, P. Song, N. Zhou, Z. Bian, J. Zhu, R. Wang, P. Dong, Z. Zhang, J. Liao, J. Yao, B. Feng, M. Murphy and G. Zheng, *ACS Photonics*, 2021, **8**, 3261–3271.
22 M. G. Gustafsson, *J. Microsc.*, 2000, **198**, 82–87.
23 E. Mudry, K. Belkebir, J. Girard, J. Savatier, E. Le Moal, C. Nicoletti, M. Allain and A. Sentenac, *Nat. Photonics*, 2012, **6**, 312–315.
24 S. Jiang, C. Guo, P. Hu, D. Hu, P. Song, T. Wang, Z. Bian, Z. Zhang and G. Zheng, *Opt. Lett.*, 2021, **46**, 5212–5215.
25 H. Wang, H. C. Koydemir, Y. Qiu, B. Bai, Y. Zhang, Y. Jin, S. Tok, E. C. Yilmaz, E. Gumustekin and Y. Rivenson, *Light: Sci. Appl.*, 2020, **9**, 1–17.
26 R. A. Campbell, R. W. Eifert and G. C. Turner, *PLoS One*, 2014, **9**, e88977.
27 K. Guo, J. Liao, Z. Bian, X. Heng and G. Zheng, *Biomed. Opt. Express*, 2015, **6**, 3210–3216.
28 S. Jiang, J. Zhu, P. Song, C. Guo, Z. Bian, R. Wang, Y. Huang, S. Wang, H. Zhang and G. Zheng, *Lab Chip*, 2020, **20**, 1058–1065.
29 M. Guizar-Sicairos, S. T. Thurman and J. R. Fienup, *Opt. Lett.*, 2008, **33**, 156–158.
30 J. Rodenburg and A. Maiden, in *Springer Handbook of Microscopy*, Springer, 2019, ch. 17, pp. 819–904.
31 W. Hoppe, *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.*, 1969, 495–501.
32 H. M. L. Faulkner and J. Rodenburg, *Phys. Rev. Lett.*, 2004, **93**, 023903.
33 G. Zheng, R. Horstmeyer and C. Yang, *Nat. Photonics*, 2013, **7**, 739.
34 G. Zheng, C. Shen, S. Jiang, P. Song and C. Yang, *Nat. Rev. Phys.*, 2021, **3**, 207–223.
35 Y. Rivenson, T. Liu, Z. Wei, Y. Zhang, K. de Haan and A. Ozcan, *Light: Sci. Appl.*, 2019, **8**, 1–11.
36 Y. Zhang, K. de Haan, Y. Rivenson, J. Li, A. Delis and A. Ozcan, *Light: Sci. Appl.*, 2020, **9**, 1–13.
37 Z. Xu, C. F. Moro, B. Bozóky and Q. Zhang, 2019, arXiv preprint arXiv:1901.04059.
38 R. Wang, P. Song, S. Jiang, C. Yan, J. Zhu, C. Guo, Z. Bian, T. Wang and G. Zheng, *Opt. Lett.*, 2020, **45**, 5405–5408.
39 C. Guo, S. Jiang, L. Yang, P. Song, T. Wang, X. Shao, Z. Zhang, M. Murphy and G. Zheng, *Opt. Express*, 2021, **29**, 39669–39684.
40 P. Thibault and A. Menzel, *Nature*, 2013, **494**, 68–71.
41 D. J. Batey, D. Claus and J. M. Rodenburg, *Ultramicroscopy*, 2014, **138**, 13–21.
42 S. Dong, R. Shiradkar, P. Nanda and G. Zheng, *Biomed. Opt. Express*, 2014, **5**, 1757–1767.
43 P. Song, R. Wang, J. Zhu, T. Wang, Z. Bian, Z. Zhang, K. Hoshino, M. Murphy, S. Jiang and C. Guo, *Opt. Lett.*, 2020, **45**, 3486–3489.
44 P. Li, T. Edo, D. Batey, J. Rodenburg and A. Maiden, *Opt. Express*, 2016, **24**, 9038–9052.
45 M. Jang, Y. Horie, A. Shibukawa, J. Brake, Y. Liu, S. M. Kamali, A. Arbabi, H. Ruan, A. Faraon and C. Yang, *Nat. Photonics*, 2018, **12**, 84–90.
46 S. Yazdanfar, K. B. Kenny, K. Tasimi, A. D. Corwin, E. L. Dixon and R. J. Filkins, *Opt. Express*, 2008, **16**, 8670–8677.
47 C. Zuo, J. Li, J. Sun, Y. Fan, J. Zhang, L. Lu, R. Zhang, B. Wang, L. Huang and Q. Chen, *Opt. Lasers Eng.*, 2020, 106187.
48 A. M. Maiden and J. M. Rodenburg, *Ultramicroscopy*, 2009, **109**, 1256–1262.
49 P. Thibault, M. Dierolf, O. Bunk, A. Menzel and F. Pfeiffer, *Ultramicroscopy*, 2009, **109**, 338–343.
50 M. Guizar-Sicairos and J. R. Fienup, *Opt. Express*, 2008, **16**, 7264–7278.

51 H. Chang, P. Enfedaque and S. Marchesini, *SIAM J. Imaging Sci.*, 2019, **12**, 153–185.
52 B. C. Platt and R. Shack, *Journal of Refractive Surgery*, 2001, **17**, S573–S577.
53 Y. Park, C. Depeursinge and G. Popescu, *Nat. Photonics*, 2018, **12**, 578–589.
54 R. Khan, B. Gul, S. Khan, H. Nisar and I. Ahmad, *Photodiagn. Photodyn. Ther.*, 2021, 102192.
55 H. Bay, T. Tuytelaars and L. Van Gool, *European conference on computer vision (ECCV) 2006*, 2006, pp. 404–417.
56 A. C. Ruifrok and D. A. Johnston, *Anal. Quant. Cytol. Histol.*, 2001, **23**, 291–299.
57 M. Veta, A. Huisman, M. A. Viergever, P. J. van Diest and J. P. Pluim, *2011 IEEE international symposium on biomedical imaging: from nano to macro*, 2011, pp. 618–621.
58 L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff and H. Adam, *Proceedings of the European conference on computer vision*, 2018, pp. 801–818.
59 K. He, X. Zhang, S. Ren and J. Sun, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
60 M. Mir, Z. Wang, Z. Shen, M. Bednarz, R. Bashir, I. Golding, S. G. Prasanth and G. Popescu, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 13124–13129.

