Rivenson et al. *Light: Science & Applications* (2019)8:23
https://doi.org/10.1038/s41377-019-0129-y

Official journal of the CIOMP 2047-7538
www.nature.com/lsa

ARTICLE Open Access

PhaseStain: the digital staining of label-free quantitative phase microscopy images using deep learning

Yair Rivenson[1,2,3], Tairan Liu[1,2,3], Zhensong Wei[1,2,3], Yibo Zhang[1,2,3], Kevin de Haan[1,2,3] and Aydogan Ozcan[1,2,3,4]

Abstract

Using a deep neural network, we demonstrate a digital staining technique, which we term PhaseStain, to transform the quantitative phase images (QPI) of label-free tissue sections into images that are equivalent to the brightfield microscopy images of the same samples that are histologically stained. Through pairs of image data (QPI and the corresponding brightfield images, acquired after staining), we train a generative adversarial network and demonstrate the effectiveness of this virtual-staining approach using sections of human skin, kidney, and liver tissue, matching the brightfield microscopy images of the same samples stained with Hematoxylin and Eosin, Jones' stain, and Masson's trichrome stain, respectively. This digital-staining framework may further strengthen various uses of label-free QPI techniques in pathology applications and biomedical research in general, by eliminating the need for histological staining, reducing sample preparation related costs and saving time. Our results provide a powerful example of some of the unique opportunities created by data-driven image transformations enabled by deep learning.

Introduction

Quantitative phase imaging (QPI) is a rapidly emerging field, with a history of several decades in development[1,2]. QPI is a label-free imaging technique, which generates a quantitative image of the optical-path-delay through the specimen. Other than being label-free, QPI utilizes low-intensity illumination, while still allowing for a rapid imaging time, which reduces phototoxicity in comparison to, e.g., commonly used fluorescence imaging modalities. QPI can be performed on multiple platforms and devices[3–7], from ultra-portable instruments all the way to custom-engineered systems integrated with standard microscopes, with different methods of extracting the quantitative phase information. QPI has also been recently used for the investigation of label-free thin tissue sections[2,8], which can be considered a weakly scattering phase object, having limited amplitude contrast modulation under brightfield illumination.

Although QPI techniques result in quantitative contrast maps of label-free objects, the current clinical and research gold standard is still mostly based on the brightfield imaging of histologically labeled samples. The staining process dyes the specimen with colorimetric markers, revealing the cellular and subcellular morphological information of the sample under brightfield microscopy. As an alternative, QPI has been demonstrated for the inference of local scattering coefficients of tissue samples[8,9]; for this information to be adopted as a diagnostic tool, some of the obstacles include the requirement of retraining experts and competing with a growing number of machine learning-based image analysis software[10,11], which utilizes vast amounts of stained tissue images to perform, e.g., automated diagnosis, image segmentation, or classification, among other tasks. One possible way to bridge the gap between QPI and standard image-based diagnostic modalities is to perform digital

Correspondence: Aydogan Ozcan (ozcan@ucla.edu)
[1]Electrical and Computer Engineering Department, University of California, Los Angeles, CA 90095, USA
[2]Bioengineering Department, University of California, Los Angeles, CA 90095, USA
Full list of author information is available at the end of the article.
These authors contributed equally: Yair Rivenson, Tairan Liu, Zhensong Wei

© The Author(s) 2019
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

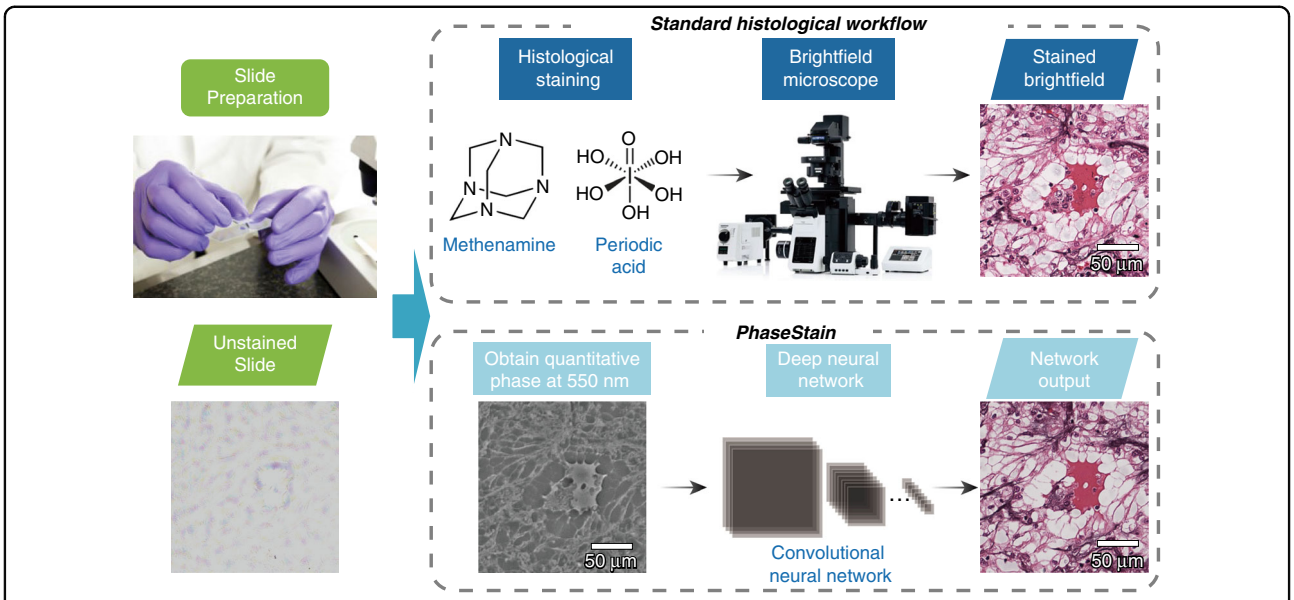

Fig. 1 PhaseStain workflow. A quantitative phase image of a label-free specimen is virtually stained by a deep neural network, bypassing the standard histological staining procedure that is used as part of clinical pathology

(i.e., virtual) staining of the phase images of label-free samples to match the images of histologically stained samples. One previously used method for the digital staining of tissue sections involves the acquisition of multimodal, nonlinear microscopy images of the samples, while applying staining reagents as part of the sample preparation, followed by a linear approximation of the absorption process to produce a pseudo-Hematoxylin and Eosin (H&E) image of the tissue section under investigation[12–14].

As an alternative to model-based approximations, deep learning has recently been successful in various computational tasks based on a data-driven approach, solving inverse problems in optics, such as super-resolution[15–17], holographic image reconstruction and phase recovery[18–21], tomography[22], Fourier ptychographic microscopy[23], localization microscopy[24–26], and ultrashort pulse reconstruction[27]. Recently, the application of deep learning for the virtual staining of autofluorescence images of nonstained tissue samples has also been demonstrated[28]. Following the success of these previous results, here, we demonstrate that deep learning can be used for the digital staining of label-free thin tissue sections using their quantitative phase images. For this image transformation between the phase image of a label-free sample and its stained brightfield image, which we term PhaseStain, we used a deep neural network trained using the generative adversarial network (GAN) framework[29]. Conceptually, PhaseStain (see Fig. 1) provides an image that is the digital equivalent of a brightfield image of the same sample after the histological staining process; stated differently, it transforms the phase image of a weakly scattering object (e.g., a label-free thin tissue section, which exhibits low amplitude modulation under visible light) into amplitude object information, presenting the same color features that are observed under a brightfield microscope, after the histological staining process.

We experimentally demonstrated the success of our PhaseStain approach using label-free sections of human skin, kidney, and liver tissue that were imaged by a holographic microscope, matching the brightfield microscopy images of the same tissue sections stained with H&E, Jones' stain, and Masson's trichrome stain, respectively.

The deep learning-based virtual-staining of label-free tissue samples using quantitative phase images provide another important example of the unique opportunities enabled by data-driven image transformations. We believe that the PhaseStain framework will be instrumental for the QPI community to further strengthen various uses of label-free QPI techniques[30–34] for clinical applications and biomedical research, helping to eliminate the need for histological staining, and reduce sample preparation associated time, labor, and costs.

Results

We trained three deep neural network models, which correspond to the three different combinations of tissue and stain types, i.e., H&E for skin tissue, Jones' stain for kidney tissue, and Masson's trichrome for liver tissue. Following the training phase, these three trained deep networks were blindly tested on holographically reconstructed quantitative phase images (see the Methods

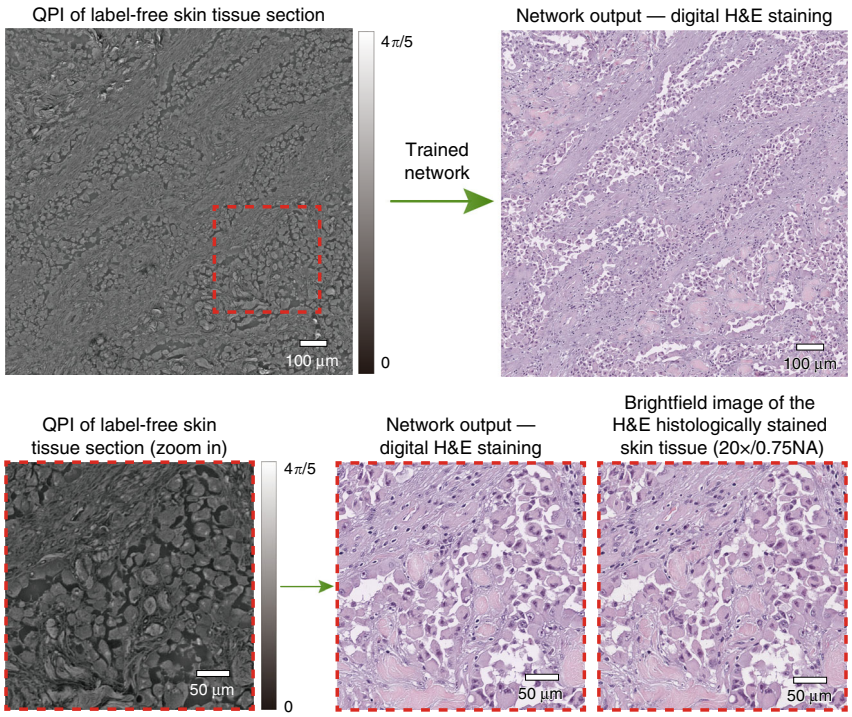

Fig. 2 Virtual H&E staining of label-free skin tissue using the PhaseStain framework. Top: QPI of a label-free skin tissue section and the resulting network output. Bottom: zoom-in image of a region of interest and its comparison to the histochemically stained gold standard brightfield image

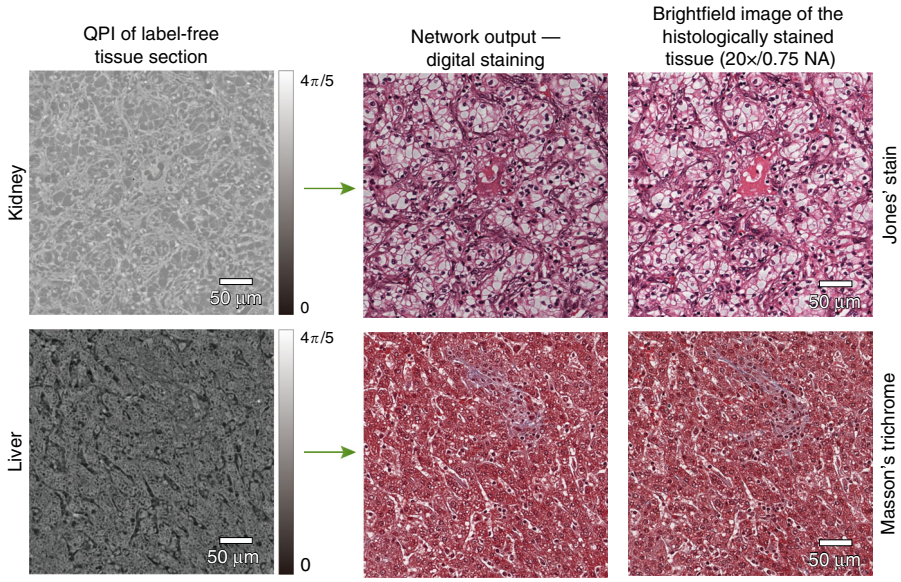

Fig. 3 PhaseStain-based virtual staining of label-free kidney tissue (Jones' stain) and liver tissue (Masson's Trichrome)

section) that were not part of the network's training set. Figure 2 shows our results for the virtual H&E staining of a phase image of a label-free skin tissue section, which confirms discohesive tumor cells lining papillary structures with dense fibrous cores. Additional results for the virtual staining of quantitative phase images of label-free tissue sections are illustrated in Fig. 3, for kidney (digital Jones' staining) and liver (digital Masson's Trichrome staining). These virtually stained quantitative phase images show sheets of clear tumor cells arranged in small nests with a delicate capillary bed for the kidney tissue section, and a virtual trichrome stain highlighting the

normal liver architecture without significant fibrosis or inflammation, for the liver tissue section.

These deep learning-based virtual-staining results presented in Figs. 2 and 3 visually demonstrate the high-fidelity performance of the GAN-based staining framework. To further shed light on this comparison between the PhaseStain results and the corresponding brightfield images of the histologically stained tissue samples, we quantified the structural similarity (SSIM) index of these two sets of images using:

$$SSIM(U_1, U_2) = \frac{1}{3}\sum_{i=1,2,3}\frac{(2\mu_{1,i}\mu_{2,i} + 2\sigma_{1,2,i} + c_2)}{\left(\mu_{1,i}^2 + \mu_{2,i}^2 + c_1\right)\left(\sigma_{1,i}^2 + \sigma_{2,i}^2 + c_2\right)} \quad (1)$$

where U_1 and U_2 are the PhaseStain output and the corresponding brightfield reference image, respectively, $\mu_{k,i}$ and $\sigma_{k,i}$ are the mean and the standard deviation of each image U_k ($k = 1,2$), respectively, and index i refers to the RGB channels of the images. The cross-variance between the i-th image channels is denoted with $\sigma_{1,2,i}$ and c_1, c_2 are stabilization constants used to prevent division by a small denominator. The result of this analysis revealed that the SSIM was 0.8113, 0.8141, and 0.8905, for the virtual-staining results corresponding to the skin, kidney, and liver tissue samples, respectively, where the analysis was performed on ~10 megapixel images, corresponding to a field-of-view (FOV) of ~1.47 mm^2 for each sample.

Next, to evaluate the sensitivity of the network output to phase noise in our measurements, we performed a numerical experiment on the quantitative phase image of a label-free skin tissue, where we added noise in the following manner:

$$\tilde{\phi}(m,n) = \phi(m,n) + \delta\phi(m,n) = \phi(m,n) + \beta r(m,n) * \frac{1}{2\pi L^2}\exp\left\{-(m^2+n^2)\Delta^2/\left[2(L\Delta)^2\right]\right\} \quad (2)$$

where $\tilde{\phi}$ is the resulting noisy phase distribution (i.e., the image under test), ϕ is the original phase image of the skin tissue sample, r is drawn from a normal distribution $N(0, 1)$, β is the perturbation coefficient, L is the Gaussian filter size/width, and Δ is the pixel size, which spatially smoothens the random noise into isotropic patches, as shown in Fig. 4. We chose these parameters such that the overall phase signal-to-noise-ratio (SNR) is statistically identical for all the cases and made sure that no phase wrapping occurs. We then used ten random realizations of this noisy phase image for four combinations of (β, L) values to generate $\tilde{\phi}$, which was used as the input to our trained deep neural network.

The deep network inference fidelity for these noisy phase inputs is reported in Fig. 4, which reveals that it is indeed sensitive to local phase variations and the related noise, and it improves its inference performance as we spatially extend the filter size, L (while the SNR remains fixed). In other words, the PhaseStain network output is more impacted by small scale variations, corresponding to, e.g., the information encoded in the morphology of the edges or the refractive index discontinuities (or sharp gradients) of the sample. We also found that for a kernel size of $L\Delta$~3 μm, the SSIM remains unchanged (~0.8), across a wide range of perturbation coefficients, β. This result implies that the network is less sensitive to sample preparation imperfections, such as height variations and wrinkles in the thin tissue section, which naturally occur during the preparation of the tissue section.

Discussion

The training process of a PhaseStain network needs to be performed only once, following which, the newly acquired quantitative phase images of various samples are blindly fed to the pretrained deep network to output a digitally stained image for each label-free sample, corresponding to the image of the same sample FOV, as it would have been imaged with a brightfield microscope, following the histological staining process. In terms of the computation speed, the virtual staining using PhaseStain takes 0.617 s on average, using a standard desktop computer equipped with a dual-GPU for an FOV of ~0.45 mm^2, corresponding to ~3.22 megapixels (see the implementation details in the Methods section). This fast inference time, even with relatively modest computers, means that the PhaseStain network can be easily integrated with a QPI-based whole slide scanner, since the network can output virtually stained images in small patches while the tissue is still being scanned by an automated microscope, to simultaneously create label-free QPI and digitally stained whole slide images of the samples.

The proposed technology has the potential to save time, labor, and costs, by presenting an alternative to the standard histological staining workflow used in clinical pathology. As an example, one of the most common staining procedures (i.e., H&E stain) takes on average ~45 min and costs approximately $2–5, while the Masson's Trichrome staining procedure takes ~2–3 h, with costs that range between $16 and $35, and often requires monitoring of the process by an expert, which is typically conducted by periodically examining the specimen under a microscope. In addition to saving time and costs, by circumventing the staining procedure, the tissue constituents would not be altered; this means that the unlabeled tissue sections can be preserved for later analysis, such as matrix-assisted laser desorption ionization by the microsectioning of specific areas[35] for molecular analysis

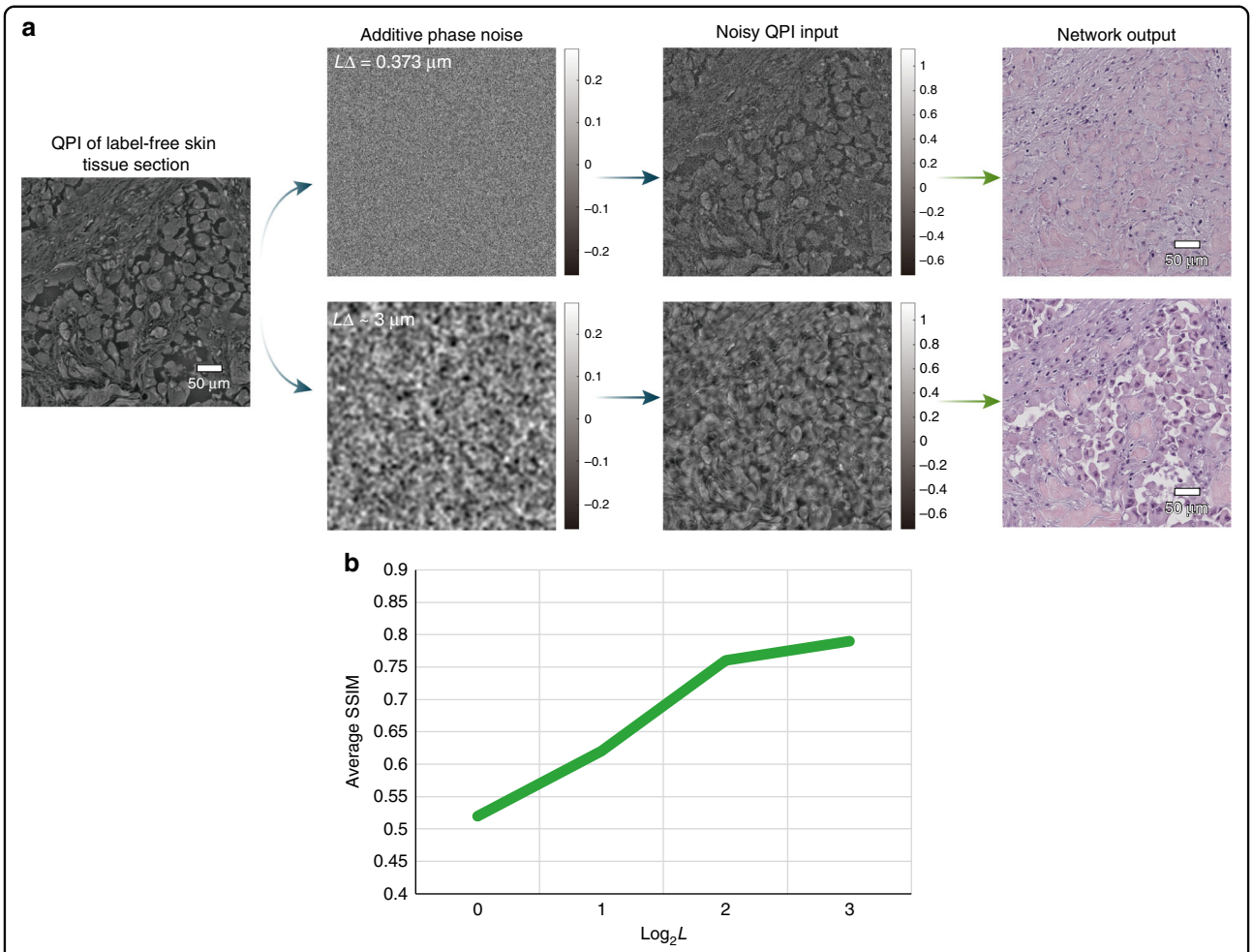

Fig. 4 PhaseStain results for noisy phase input images (ground truth shown in Fig. 2). a Top row: $L\Delta$~0.373 µm; second row: $L\Delta$~3 µm. **b** Analysis of the impact of phase noise on the inference quality of PhaseStain (quantified using SSIM), as a function of the Gaussian filter length, L (see Eq. (2))

or the micromarking of subregions that can be labeled with specific immunofluorescence tags or tested for personalized therapeutic strategies and drugs[36,37].

While in this study, we trained three different neural network models to obtain optimal results for specific tissue and stain combinations, this does not pose a practical limitation for PhaseStain, since we can also train a more general digital staining model for a specific stain type (H&E, Jones' stain, etc.) using multiple tissue types stained with it, at the cost of increasing the network size as well as the training and inference times[19]. Additionally, from the clinical diagnostics perspective, the tissue type under investigation and the stain needed for its clinical examination are both known a priori, and therefore the selection of the correct neural network for each sample to be examined is straightforward to implement.

It is important to note that, in addition to the lensfree holographic microscope (see the Methods section) that we used in this work, the PhaseStain framework can also be applied to virtually stain the resulting images of various other QPI techniques, regardless of their imaging configuration, specific hardware, or phase recovery method[2,6,7,38–41] that are employed.

One of the disadvantages of coherent imaging systems is "coherence-related image artifacts", such as speckle noise, or dust or other particles creating holographic interference fringes, which do not appear in the incoherent brightfield microscopy images of the same samples. In Fig. 5, we demonstrate the image distortions that, for example, out-of-focus particles create on the PhaseStain output image. To reduce such distortions in the network output images, the coherence-related image artifacts resulting from out-of-focus particles can be digitally removed by using a recently introduced deep learning-based hologram reconstruction method, which learns, through data, to attack or eliminate twin-image artifacts as well as the interference fringes resulting from out-of-focus or undesired objects[19,20].

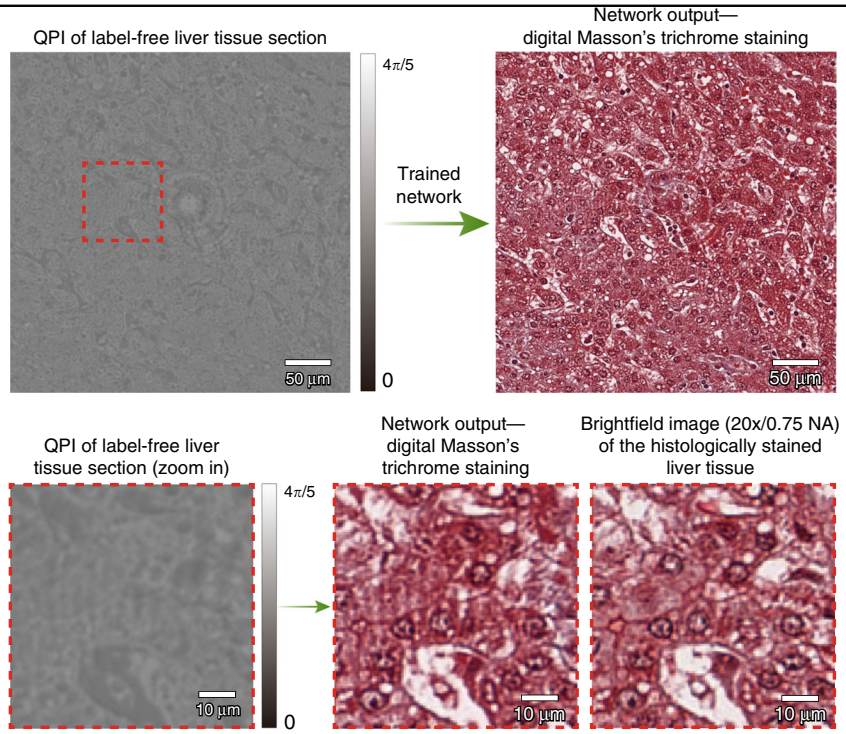

Fig. 5 The impact of holographic fringes resulting from out-of-focus particles on the deep neural network's digital staining performance. Top row: QPI of a label-free liver tissue section and the resulting network output. Bottom row: zoom-in image of a region of interest where the coherence-related artifact partially degrades the virtual staining performance

While in this manuscript, we demonstrated the applicability of the PhaseStain approach to fixed paraffin-embedded tissue specimens, our approach should also be applicable to frozen tissue sections, involving other tissue fixation methods as well (following a similar training process as detailed in the Methods section). Moreover, while our method was demonstrated for thin tissue sections, QPI has been shown to be valuable to image cells and smear samples (such as blood and Pap smears)[2,41], and the PhaseStain technique would also be applicable to digitally stain these types of specimens.

To summarize, our presented results demonstrate some of the emerging opportunities created by deep learning for label-free quantitative phase imaging. The phase information resulting from various coherent imaging techniques can be used to generate a virtually stained image, translating the phase images of weakly scattering objects such as thin tissue sections into images that are equivalent to the brightfield images of the same samples, after the histological labeling. The PhaseStain framework, in addition to saving time and costs associated with the labeling process, has the potential to further strengthen the use of label-free QPI techniques in the clinical diagnostics workflow, while also preserving tissues for, e.g., subsequent molecular and genetic analysis.

Materials and methods

Sample preparation and imaging

All the samples that were used in this study were obtained from the Translational Pathology Core Laboratory (TPCL) and prepared by the Histology Lab at UCLA. They were obtained after the de-identification of the patient related information and prepared from existing specimens. Therefore, this work did not interfere with standard practices of care or sample collection procedures.

Following formalin-fixing paraffin-embedding, the tissue block is sectioned using a microtome into ~2–4 µm thick sections. This step is only needed for the training phase, where the transformation from a phase image into a brightfield image needs to be statistically learned. These tissue sections are then deparaffinized using Xylene and mounted on a standard glass slide using CytosealTM (Thermo-Fisher Scientific, Waltham, MA, USA), followed by sealing of the specimen with a coverslip. In the learning/training process, this sealing step presents several advantages: protecting the sample during the imaging and sample handling processes and reducing artifacts such as sample thickness variations.

Following the sample preparation, the specimen was imaged using an on-chip holographic microscope to generate a quantitative phase image (detailed in the next

subsection). Following the QPI process, the label-free specimen slide was put into Xylene for ~48 h, until the coverslip can be removed without introducing distortions to the tissue. Once the coverslip was removed, the slide was dipped multiple times in absolute alcohol and 95% alcohol, and then washed in D.I. water for ~1 min. Following this step, the tissue slides were stained with H&E (skin tissue), Jones' stain (kidney tissue), and Masson's trichrome (liver tissue) and then coverslipped. These tissue samples were then imaged using a brightfield automated slide scanner microscope (Aperio AT, Leica Biosystems) with a 20×/0.75NA objective (Plan Apo), equipped with a 2×magnification adapter, which results in an effective pixel size of ~0.25 μm.

Quantitative phase imaging

Lensfree imaging setup

The quantitative phase images of label-free tissue samples were acquired using an in-line lensfree holography setup[41]. A light source (WhiteLase Micro, NKT Photonics, Denmark) with a center wavelength at 550 nm and a spectral bandwidth of ~2.5 nm was used as the illumination source. The uncollimated light emitted from a single-mode fiber was used for creating a quasi-plane-wave that illuminated the sample. The sample was placed between the light source and the CMOS image sensor chip (IMX 081, Sony Corp., Minato, Tokyo, Japan, pixel size of 1.12 μm) with a source-to-sample distance (z_1) of 5–10 cm and a sample-to-sensor distance (z_2) of 1–2 mm. This on-chip lensfree holographic microscope has a submicron resolution with an effective pixel size of 0.37 μm, covering a sample FOV of ~20 mm^2 (which accounts for the entire active area of the sensor). The positioning stage (MAX606, Thorlabs Inc., Newton, NJ, USA), which held-the CMOS sensor, enabled the 3D translation of the imager chip for performing pixel super-resolution (PSR)[5,41,42] and multiheight-based iterative phase recovery[41,43]. All imaging hardware was controlled automatically by LabVIEW (National Instruments Corp., Austin, TX, USA).

Pixel super-resolution (PSR) technique

To synthesize a high-resolution hologram (with a pixel size of ~0.37 μm) using only the G1 channel of the Bayer pattern (R, G1, G2, and B), a shift-and-add based PSR algorithm was applied[42,44]. The translation stage that holds the image sensor was programmed to laterally shift on a 6 × 6 grid with a subpixel spacing at each sample-to-sensor distance. A low-resolution hologram was recorded at each position and the lateral shifts were precisely estimated using a shift estimation algorithm[41]. This step results in six nonoverlapping panels that were each padded to a size of 4096 × 4096 pixels, and individually phase-recovered, which is detailed next.

Multiheight phase recovery

Lensfree in-line holograms at eight sample-to-sensor distances were captured. The axial scanning step size was chosen to be 15 μm. Accurate z-steps were obtained by applying a holographic autofocusing algorithm based on the edge sparsity criterion ("Tamura of the gradient", i.e., ToG)[45]. A zero-phase was assigned to the object intensity measurement as an initial phase guess, to start the iterations. An iterative multiheight phase recovery algorithm[46] was then used by propagating the complex field back and forth between each height using the transfer function of free-space[47]. During this iterative process, the phase was kept unchanged at each axial plane, where the amplitude was updated by using the square-root of the object intensity measurement. One iteration was defined as propagating the hologram from the eighth height (farthest from the sensor chip) to the first height (nearest to the sensor) and then back propagating the complex field to the eighth height. Typically, after 10–30 iterations, the phase is retrieved. For the final step of the reconstruction, the complex wave defined by the converged amplitude and phase at a given hologram plane was propagated to the object plane[47], from which the phase component of the sample was extracted.

Data preprocessing and image registration

An important step in our training process is to perform an accurate image registration, between the two imaging modalities (QPI and brightfield), which involves both global matching and local alignment steps. Since the network aims to learn the transformation from a label-free phase retrieved image to a histologically stained brightfield image, it is crucial to accurately align the FOVs for each input and target image pair in the dataset. We perform this cross-modality alignment procedure in four steps; steps 1, 2, and 4 are done in MATLAB (The MathWorks Inc., Natick, MA, USA) and step 3 involves TensorFlow.

The first step is to find a roughly matched FOV between QPI and the corresponding brightfield image. This is done by first bicubic downsampling of the whole slide image (WSI) (~60 by 60 k pixels) to match the pixel size of the phase retrieved image. Then, each 4096 × 4096-pixel phase image was cropped by 256 on each side (resulting in an image with 3584 × 3584 pixels) to remove the padding that is used for the image reconstruction process. Following this step, both the brightfield and the corresponding phase images are edge extracted using the Canny method[48], which uses a double threshold to detect strong and weak edges on the gradient of the image. Then, a correlation score matrix is calculated by correlating each 3584 × 3584-pixel patch of the resulting edge image to the same size as the image extracted from the brightfield edge image. The image with the highest correlation score

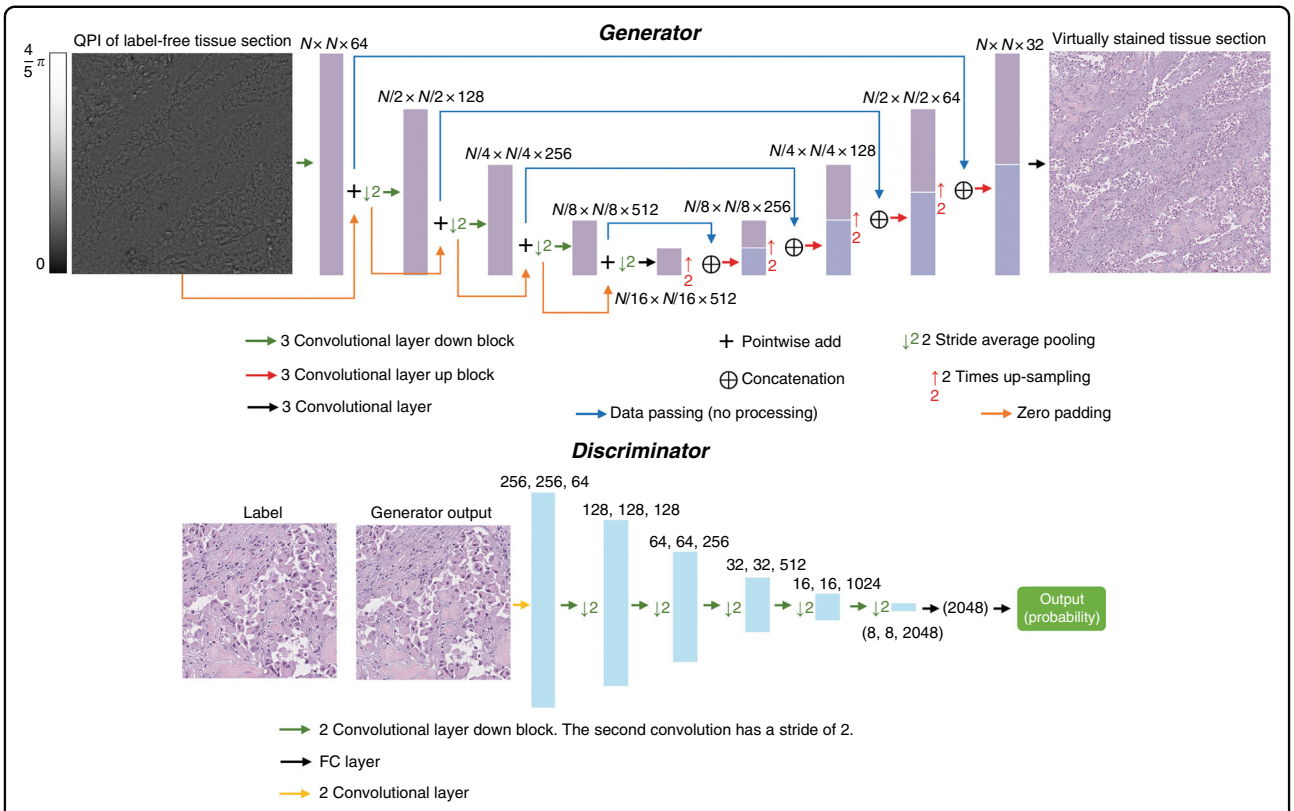

Fig. 6 Architecture of the generator and discriminator networks within the GAN framework

indicates a match between the two images, and the corresponding brightfield image is cropped out from the WSI. Following this initial matching procedure, the quantitative phase image and the brightfield microscope images are coarsely matched.

The second step is used to correct for potential rotations between these coarsely matched image pairs, which might be caused by a slight mismatch in the sample placement during the two image acquisition experiments (which are performed on different imaging systems, holographic vs. brightfield). This intensity-based registration step correlates the spatial patterns between the two images; the phase image that is converted to an unsigned integer format and the luminance component of the brightfield image were used for this multimodal registration framework implemented in MATLAB. The result of this digital procedure is an affine transformation matrix, which is applied to the brightfield microscope image patch, to match it with the quantitative phase image of the same sample. Following this registration step, the phase image and the corresponding brightfield image are globally aligned. A further crop of 64 pixels on each side to the aligned image pairs is used to accommodate for a possible rotation angle correction.

The third step involves the training of a separate neural network that roughly learns the transformation from quantitative phase images into stained brightfield images, which can help the distortion correction between the two image modalities in the fourth/final step. In other words, to make the local registration tractable, we first train a deep network with the globally registered images, to reduce the entropy between the images acquired with the two imaging modalities (i.e., QPI vs. brightfield image of the stained tissue). This neural network has the same structure as the network that was used for the final training process (see the next subsection on the GAN architecture and its training) with the input and target images obtained from the second registration step discussed earlier. Since the image pairs are not well aligned yet, the training is stopped early at only ~2000 iterations to avoid a structural change at the output to be learnt by the network. The output and target images of the network are then used as the registration pairs in the fourth step, which is an elastic image registration algorithm, used to correct for local feature registration[16].

GAN architecture and training

The GAN architecture that we used for PhaseStain is detailed in Fig. 6 and Supplementary Table 1. Following the registration of the label-free quantitative phase images to the brightfield images of the histologically stained tissue sections, these accurately aligned fields-of-view were

Table 1 Training details for the virtual staining of different tissue types using PhaseStain. Following the training, the blind inference takes ~0.617 s for an FOV of ~0.45 mm^2, corresponding to ~3.22 megapixels (see the Discussion section)

Tissue type	# of iterations	# of patches (256 × 256 pixels)	Training time (h)	# of epochs
Liver	7500	2500 training/625 validation	11.076	25
Skin	11000	2500 training/625 validation	11.188	18
Kidney	13600	2312 training/578 validation	13.173	39

partitioned to overlapping patches of 256 × 256 pixels, which were then used to train the GAN model. The GAN is composed of two deep neural networks, a generator and a discriminator. The discriminator network's loss function is given by:

$$\ell_{discrimnator} = D(G(x_{input}))^2 + (1 - D(z_{label}))^2 \quad (3)$$

where $D(.)$ and $G(.)$ refer to the discriminator and generator network operators, respectively, x_{input} denotes the input to the generator, which is the label-free quantitative phase image, and z_{label} denotes the brightfield image of the histologically stained tissue. The generator network, G, tries to generate an output image with the same statistical features as z_{label}, while the discriminator, D, attempts to distinguish between the target and the generator output images. The ideal outcome (or state of equilibrium) will be when the generator's output and target images share an identical statistical distribution, where in this case, $D(G(x_{\text{input}}))$ should converge to 0.5. For the generator deep network, we defined the loss function as:

$$\ell_{generator} = L_1\{z_{label}, G(x_{input})\} + \lambda \times TV\{G(x_{input})\} + \alpha \times (1 - D(G(x_{input})))^2 \quad (4)$$

where the $L_1\{.\}$ term refers to the absolute pixel-by-pixel difference between the generator output image and its target, $TV\{.\}$ stands for the total variation regularization that is being applied to the generator output, and the last term reflects a penalty related to the discriminator network prediction of the generator output. The regularization parameters (λ, α) were set to 0.02 and 2000 so that the total variation loss term, $\lambda \times TV\{G(x_{\text{input}})\}$, was ~2% of the L_1 loss term, and the discriminator loss term, $\alpha \times (1 - D(G(x_{\text{input}})))^2$ was ~98% of the total generator loss, $l_{\text{generator}}$.

For the generator deep neural network, we adapted the U-net architecture[49], which consists of a downsampling and an upsampling path, with each path containing four blocks forming four distinct levels (see Fig. 6 and Supplementary Table 1). In the downsampling path, each residual block consists of three convolutional layers and three leaky rectified linear (LReLU) units used as an activation function, which is defined as:

$$LReLU(x) = \begin{cases} x & for\ x>0 \\ 0.1x & otherwise \end{cases} \quad (5)$$

At the output of each block, the number of channels is increased by 2-fold (except for the first block that increases from 1 input channel to 64 channels). The blocks are connected by an average-pooling layer of stride two that downsamples the output of the previous block by a factor of two for both horizontal and vertical dimensions (as shown in Fig. 6 and supplementary Table 1).

In the upsampling path, each block also consists of three convolutional layers and three LReLU activation functions, which decrease the number of channels at its output by fourfold. The blocks are connected by a bilinear upsampling layer that upsamples the size of the output from the previous block by a factor of two for both lateral dimensions. A concatenation function with the corresponding feature map from the downsampling path of the same level is used to increase the number of channels from the output of the previous block by two. The two paths are connected in the first level of the network by a convolutional layer, which maintains the number of the feature maps from the output of the last residual block in the downsampling path (see Fig. 6 and supplementary Table 1). The last layer is a convolutional layer that maps the output of the upsampling path into 3 channels of the YCbCr color map.

The discriminator network consists of one convolutional layer, five discriminator blocks, an average-pooling layer, and two fully connected layers. The first convolutional layer receives 3 channels (YCbCr color map) from either the generator output or the target and increases the number of channels to 64. The discriminator blocks consist of two convolutional layers with the first layer maintaining the size of the feature map and the number of channels, while the second layer increases the number of channels by twofold and decreases the size of the feature map by fourfold. The average-pooling layer has a filter size of 8 × 8, which results in a matrix with a size of (B, 2048), where B refers to the batch size. The output of this average-pooling layer is then fed into two fully connected layers with the

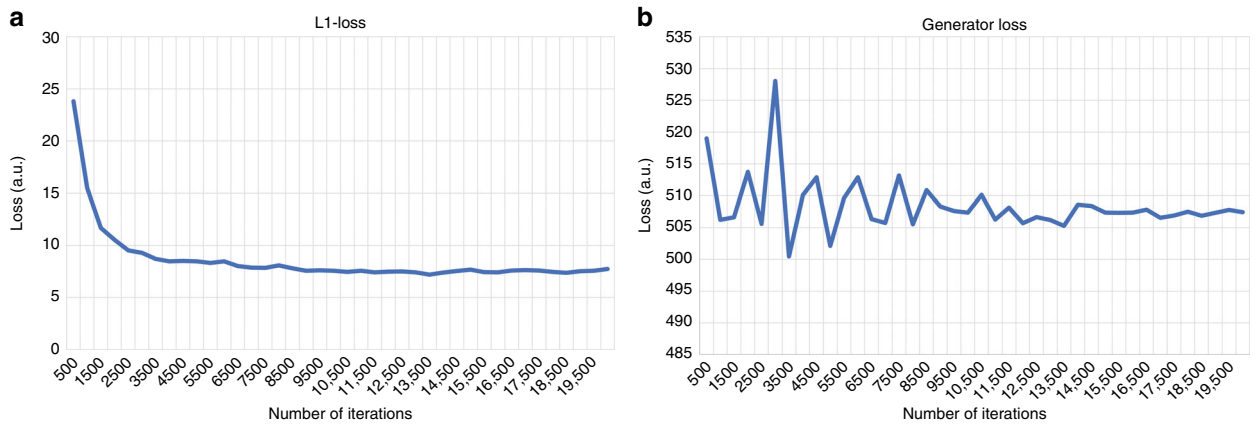

Fig. 7 PhaseStain convergence plots for the validation set of the digital H&E staining of the skin tissue. a *L*1-loss with respect to the number of iterations. **b** Generator loss, $l_{generator}$ with respect to the number of iterations

first layer maintaining the size of the feature map, while the second layer decreases the output channel to 1, resulting in an output size of (B, 1). The output of this fully connected layer is going through a sigmoid function, indicating the probability that the three-channel discriminator input is drawn from a histologically stained brightfield image. For the discriminator network, all the convolutional layers and the fully connected layers are connected by LReLU nonlinear activation functions.

Throughout the training, the convolution filter size was set to be 3×3. For the patch generation, we applied data augmentation by using 50% patch overlap for the liver and skin tissue images, and 25% patch overlap for the kidney tissue images (see Table 1). The learnable parameters including filters, weights, and biases in the convolutional layers and the fully connected layers are updated using an adaptive moment estimation (Adam) optimizer with a learning rate of 1×10^{-4} for the generator network and 1×10^{-5} for the discriminator network.

For each iteration of the discriminator, there were v iterations of the generator network; for the liver and skin tissue training, $v = \max(5, \text{floor}(7 - w/2))$, where we increased w by 1 for every 500 iterations (w was initialized as 0). For the kidney tissue training, we used $v = \max(4, \text{floor}(6 - w/2))$, where we increased w by 1 for every 400 iterations. This helped us to train the discriminator not to overfit to the target brightfield images. We used a batch size of ten for the training of the liver and skin tissue sections, and five for the kidney tissue sections. All the convolutional kernel entries are initialized using a truncated normal distribution. All the network bias terms are initialized to be zero. The network's training stopped when the validation set's L_1-loss did not decrease after 4000 iterations. A typical convergence plot of our training is shown in Fig. 7.

Implementation details

The number of image patches that were used for training, the number of epochs, and the training schedules are shown in Table 1. The network was implemented using Python version 3.5.0, with TensorFlow framework version 1.7.0. We implemented the software on a desktop computer with a Core i7-7700K CPU @ 4.2 GHz (Intel Corp., Santa Clara, CA, USA) and 64GB of RAM, running a Windows 10 operating system (Microsoft Corp., Redmond, WA, USA). Following the training for each tissue section, the corresponding network was tested with 4 image patches of 1792×1792 pixels with an overlap of ~7%. The outputs of the network were then stitched to form the final network output image of 3456×3456 pixels (FOV ~1.7 mm^2), as shown in, e.g., Fig. 2. The network training and testing were performed using dual GeForce GTX 1080Ti GPUs (NVidia Corp., Santa Clara, CA, USA).

Acknowledgments

The Ozcan Research Group at UCLA acknowledges the support of NSF Engineering Research Center (ERC, PATHS-UP), the Army Research Office (ARO; W911NF-13-1-0419 and W911NF-13-1-0197), the ARO Life Sciences Division, the National Science Foundation (NSF) CBET Division Biophotonics Program, the NSF Emerging Frontiers in Research and Innovation (EFRI) Award, the NSF INSPIRE Award, NSF Partnerships for Innovation: Building Innovation Capacity (PFI:BIC) Program, the National Institutes of Health (NIH, R21EB023115), the Howard Hughes Medical Institute (HHMI), Vodafone Americas Foundation, the Mary Kay Foundation, and Steven & Alexandra Cohen Foundation. The authors also acknowledge the Translational Pathology Core Laboratory (TPCL) and the Histology Lab at UCLA for their assistance with the sample preparation and staining, as well as Prof. W. Dean Wallace of Pathology and Laboratory Medicine at UCLA's David Geffen School of Medicine for image evaluations.

Author details

[1]Electrical and Computer Engineering Department, University of California, Los Angeles, CA 90095, USA. [2]Bioengineering Department, University of California, Los Angeles, CA 90095, USA. [3]California NanoSystems Institute (CNSI), University of California, Los Angeles, CA 90095, USA. [4]Department of Surgery, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA

Conflict of interest

A.O., Y.R., and Z.W. have a patent application on the invention reported in this manuscript.

Supplementary information is available for this paper at https://doi.org/10.1038/s41377-019-0129-y.

Received: 23 July 2018 Revised: 5 January 2019 Accepted: 11 January 2019
Published online: 06 February 2019

References

1. Cuche, E., Marquet, P. & Depeursinge, C. Simultaneous amplitude-contrast and quantitative phase-contrast microscopy by numerical reconstruction of Fresnel off-axis holograms. *Appl. Opt.* **38**, 6994–7001 (1999).
2. Popescu, G. *Quantitative Phase Imaging of Cells and Tissues*. (McGraw-Hill, New York, 2011).
3. Shaked, N. T., Rinehart, M. T. & Wax, A. Dual-interference-channel quantitative-phase microscopy of live cell dynamics. *Opt. Lett.* **34**, 767–769 (2009).
4. Wang, Z. et al. Spatial light interference microscopy (SLIM). *Opt. Express* **19**, 1016–1026 (2011).
5. Greenbaum, A. et al. Imaging without lenses: achievements and remaining challenges of wide-field on-chip microscopy. *Nat. Methods* **9**, 889–895 (2012).
6. Zheng, G. A., Horstmeyer, R. & Yang, C. Wide-field, high-resolution Fourier ptychographic microscopy. *Nat. Photonics* **7**, 739–745 (2013).
7. Tian, L. & Waller, L. Quantitative differential phase contrast imaging in an LED array microscope. *Opt. Express* **23**, 11394–11403 (2015).
8. Wang, Z., Tangella, K., Balla, A. & Popescu, G. Tissue refractive index as marker of disease. *J. Biomed. Opt.* **16**, 116017 (2011).
9. Wang, Z., Ding, H. F. & Popescu, G. Scattering-phase theorem. *Opt. Lett.* **36**, 1215–1217 (2011).
10. Liu Y., et al. Detecting cancer metastases on gigapixel pathology images. ArXiv: 1703.02442 (2017).
11. Litjens G., et al. A survey on deep learning in medical image analysis. *Med. Image Anal.* **42**: 60–88 (2017).
12. Tao, Y. K. et al. Assessment of breast pathologies using nonlinear microscopy. *Proc. Natl Acad. Sci. USA* **111**, 15304–15309 (2014).
13. Giacomelli, M. G. et al. Virtual hematoxylin and eosin transillumination microscopy using epi-fluorescence imaging. *PLoS ONE* **11**, e0159337 (2016).
14. Orringer, D. A. et al. Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated Raman scattering microscopy. *Nat. Biomed. Eng.* **1**, 0027 (2017).
15. Rivenson, Y. et al. Deep learning microscopy. *Optica* **4**, 1437–1443 (2017).
16. Rivenson, Y. et al. Deep learning enhanced mobile-phone microscopy. *ACS Photon.* **5**, 2354–2364, https://doi.org/10.1021/acsphotonics.8b00146 (2018).
17. Wang H., et al. Deep learning enables cross-modality super-resolution in fluorescence microscopy *Nat. Methods* **16**, 103–110 (2019).
18. Sinha, A., Lee, J., Li, S. & Barbastathis, G. Lensless computational imaging through deep learning. *Optica* **4**, 1117–1125 (2017).
19. Rivenson, Y., Zhang, Y. B., Günaydin, H., Teng, D. & Ozcan, A. Phase recovery and holographic image reconstruction using deep learning in neural networks. *Light Sci. Appl.* **7**, e17141 (2018).
20. Wu, Y. C. et al. Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery. *Optica* **5**, 704–710 (2018).
21. Jo Y., et al. Quantitative phase imaging and artificial intelligence: a review. arXiv: 1806.03982 (2018).
22. Kamilov, U. et al. Learning approach to optical tomography. *Optica* **2**, 517–522 (2015).
23. Nguyen T., Xue Y. J., Li Y. Z., Tian L., Nehmetallah G. Deep learning approach to Fourier ptychographic microscopy. arXiv: 1805.00334 (2018).
24. Boyd N., Jonas E., Babcock H. P., Recht B. DeepLoco: fast 3D localization microscopy using neural networks. bioRxiv: 267096, 2018. https://doi.org/10.1101/267096.
25. Nehme, E., Weiss, L. E., Michaeli, T. & Shechtman, Y. Deep-STORM: super-resolution single-molecule microscopy by deep learning. *Optica* **5**, 458–464 (2018).
26. Ouyang, W., Aristov, A., Lelek, M., Hao, X. & Zimmer, C. Deep learning massively accelerates super-resolution localization microscopy. *Nat. Biotechnol.* **36**, 460–468, https://doi.org/10.1038/nbt.4106 (2018).
27. Zahavy, T. et al. Deep learning reconstruction of ultrashort pulses. *Optica* **5**, 666–673 (2018).
28. Rivenson Y., et al. Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. *Nat. Biomed. Eng.* (in the press).
29. Goodfellow I. J., et al. Generative adversarial nets. In Proceedings of the 27th International Conference on Neural Information Processing Systems. (MIT Press: Cambridge, MA, 2014) pp. 2672–2680.
30. Park, H. S., Rinehart, M. T., Walzer, K. A., Chi, J. T. A. & Wax, A. Automated detection of *P. falciparum* using machine learning algorithms with quantitative phase images of unstained cells. *PLoS ONE* **11**, e0163045 (2016).
31. Chen, C. L. et al. Deep learning in label-free cell classification. *Sci. Rep.* **6**, 21471 (2016).
32. Roitshtain, D. et al. Quantitative phase microscopy spatial signatures of cancer cells. *Cytometry A* **91**, 482–493 (2017).
33. Jo, Y. et al. Holographic deep learning for rapid optical screening of anthrax spores. *Sci. Adv.* **3**, e1700606, https://doi.org/10.1101/109108 (2017).
34. Javidi, B. et al. Sickle cell disease diagnosis based on spatio-temporal cell dynamics analysis using 3D printed shearing digital holographic microscopy. *Opt. Express* **26**, 13614–13627 (2018).
35. Tata, A. et al. Wide-field tissue polarimetry allows efficient localized mass spectrometry imaging of biological tissues. *Chem. Sci.* **7**, 2162–2169 (2016).
36. Cree, I. A. et al. Guidance for laboratories performing molecular pathology for cancer patients. *J. Clin. Pathol.* **67**, 923–931 (2014).
37. Patel, P. G. et al. Preparation of formalin-fixed paraffin-embedded tissue cores for both RNA and DNA extraction. *J. Vis. Exp.* **2016**, e54299, https://doi.org/10.3791/54299 (2016).
38. Ikeda, T., Popescu, G., Dasari, R. R. & Feld, M. S. Hilbert phase microscopy for investigating fast dynamics in transparent systems. *Opt. Lett.* **30**, 1165–1167 (2005).
39. Shaked, N. T., Zhu, Y. Z., Badie, N., Bursac, N. & Wax, A. Reflective interferometric chamber for quantitative phase imaging of biological sample dynamics. *J. Biomed. Opt.* **15**, 030503 (2010).
40. Watanabe, E., Hoshiba, T. & Javidi, B. High-precision microscopic phase imaging without phase unwrapping for cancer cell identification. *Opt. Lett.* **38**, 1319–1321 (2013).
41. Greenbaum, A. et al. Wide-field computational imaging of pathology slides using lens-free on-chip microscopy. *Sci. Transl. Med* **6**, 267ra175 (2014).
42. Bishara, W., Su, T. W., Coskun, A. F. & Ozcan, A. Lensfree on-chip microscopy over a wide field-of-view using pixel super-resolution. *Opt. Express* **18**, 11181–11191 (2010).
43. Luo, W., Zhang, Y. B., Göröcs, Z., Feizi, A. & Ozcan, A. Propagation phasor approach for holographic image reconstruction. *Sci. Rep.* **6**, 22738 (2016).
44. Farsiu, S., Robinson, M. D., Elad, M. & Milanfar, P. Fast and robust multiframe super resolution. *IEEE Trans. Image Process.* **13**, 1327–1344 (2014).
45. Zhang, Y. B., Wang, H. D., Wu, Y. C., Tamamitsu, M. & Ozcan, A. Edge sparsity criterion for robust holographic autofocusing. *Opt. Lett.* **42**, 3824–3827 (2017).
46. Greenbaum, A. & Ozcan, A. Maskless imaging of dense samples using pixel super-resolution based multi-height lensfree on-chip microscopy. *Opt. Express* **20**, 3129–3143 (2012).
47. Goodman, J. W. *Introduction to Fourier Optics*. (Roberts and Company Publishers, Englewood, 2005).
48. Canny, J. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-8**, 679–698 (1986).
49. Ronneberger O., Fischer P., Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N., Hornegger J., Wells W. M., Frangi A. F., (eds). Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015. (Springer, Cham, 2015) pp. 234–241 https://doi.org/10.1007/978-3-319-24574-4_28.