Holographic virtual staining of individual biological cells

Yoav N. Nygate[a], Mattan Levi[a], Simcha K. Mirsky[a], Nir A. Turko[a], Moran Rubin[a], Itay Barnea[a], Gili Dardikman-Yoffe[a], Miki Haifler[a], Alon Shalev[b], and Natan T. Shaked[a,1]

[a]Faculty of Engineering, Department of Biomedical Engineering, Tel Aviv University, 6997801 Tel Aviv, Israel; and [b]QART Medical, 4366236 Ra'anana, Israel

Edited by John A. Rogers, Northwestern University, Evanston, IL, and approved March 3, 2020 (received for review November 14, 2019)

Many medical and biological protocols for analyzing individual biological cells involve morphological evaluation based on cell staining, designed to enhance imaging contrast and enable clinicians and biologists to differentiate between various cell organelles. However, cell staining is not always allowed in certain medical procedures. In other cases, staining may be time-consuming or expensive to implement. Staining protocols may be operator-sensitive, and hence may lead to varying analytical results, as well as cause artificial imaging artifacts or false heterogeneity. We present a deep-learning approach, called HoloStain, which converts images of isolated biological cells acquired without staining by holographic microscopy to their virtually stained images. We demonstrate this approach for human sperm cells, as there is a well-established protocol and global standardization for characterizing the morphology of stained human sperm cells for fertility evaluation, but, on the other hand, staining might be cytotoxic and thus is not allowed during human in vitro fertilization (IVF). After a training process, the deep neural network can take images of unseen sperm cells retrieved from holograms acquired without staining and convert them to their stainlike images. We obtained a fivefold recall improvement in the analysis results, demonstrating the advantage of using virtual staining for sperm cell analysis. With the introduction of simple holographic imaging methods in clinical settings, the proposed method has a great potential to become a common practice in human IVF procedures, as well as to significantly simplify and radically change other cell analyses and techniques such as imaging flow cytometry.

digital holography | deep learning | biological cell imaging

Digital pathology and cytology are emerging fields that are eventually expected to become fully automated and nonsubjective, with applications ranging from routine clinical tests of body fluids to more complex biological research. Part of these analyses is based on morphological evaluation of individual cells. Cells in vitro are mostly transparent under regular light microscopy, and therefore cannot be imaged well without external stains or contrast agents. However, cell staining is time-consuming and the staining materials might be harmful to the cells, resulting in the prohibition of chemical staining in certain medical procedures. Specifically, cell staining is not allowed during the selection of sperm cells for human in vitro fertilization (IVF), preventing high-quality intracellular morphology evaluation. Off-axis holography records the quantitative phase profile of the cell, which takes into account the cell refractive index and physical thickness, in a single camera exposure. This method creates great imaging contrast without the need for external contrast agents. The fact that the phase profile is quantitative and accounts for the cell internal refractive indices gives rise to new parameters with medical relevance that have not been available in imaging flow cytometry before, such as the dry mass of the cells (1, 2), even in addition to using contrast agents. Until recently, holographic cell imaging could not be implemented in clinical settings due to the bulkiness and nonportability of the optical system, as well as the requirement for specific optical skills to align and use it. In the last years, successful efforts have been made to make these wavefront sensors affordable to clinical use (3). Our approach, called interferometric phase microscopy (IPM), is based on the usage of microscopes already existing in medical clinics and attaching a portable interferometric module to their exit port (4). This wavefront sensor is compact, inexpensive, and easy to operate, making this technology accessible and affordable to clinicians' direct use. However, despite the potential of this technique to aid cell analysis, existing and well-established protocols for morphological cell evaluation are still based on chemical staining of the cell organelles, rather than on the quantitative topographic maps obtained by holography. Thus, despite its potential, digital holography is far from full integration into medical procedures and biological protocols.

In this paper, we propose a deep-learning approach for transforming quantitative phase maps of individual biological cells extracted from digital holograms to their virtual staining images, which are very similar to their actual chemical staining images. We have chosen to demonstrate this technique for stain-free sperm imaging, since there is an established World Health Organization (WHO) protocol for morphological evaluation of sperm cells during fertility evaluation. However, this protocol cannot be fully implemented in human IVF procedures due to the prohibition of using cell staining.

In the past several years, deep learning has emerged as a beneficial tool in the medical imaging field, simplifying many complex image analysis tasks (5). Deep learning enables the computer to learn specific tasks based on observed data. This is done by feeding the data through many processing layers, which, after a training procedure, are able to estimate complex data

Significance

We present a method for virtual staining for morphological analysis of individual biological cells based on stain-free digital holography, allowing clinicians and biologists to visualize and analyze the cells as if they have been chemically stained. Our approach provides numerous advantages, as it 1) circumvents the possible toxicity of staining materials, 2) saves time and resources, 3) optimizes inter- and intralab variability, 4) allows concurrent staining of different types of cells with multiple virtual stains, and 5) provides ideal conditions for real-time analysis, such as rapid stain-free imaging flow cytometry. The proposed method is shown to be accurate, repeatable, and nonsubjective. Hence, it bears great potential to become a common tool in clinical settings and biological research.

Author contributions: Y.N.N., A.S., and N.T.S. designed research; Y.N.N., M.L., S.K.M., N.A.T., M.R., I.B., G.D.-Y., M.H., and N.T.S. performed research; Y.N.N., M.L., and N.T.S. analyzed data; Y.N.N. and N.T.S. wrote the manuscript; all authors discussed the results and commented on the manuscript; and N.T.S. supervised the research.

Competing interest statement: We have submitted a patent application on this subject.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: nshaked@tau.ac.il.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1919569117/-/DCSupplemental.

First published April 13, 2020.

representations (6). Deep learning was already demonstrated as a beneficial method for performing segmentation of medical images (7–9) and solving various inverse problems in the medical imaging field (10). Furthermore, recently Pinkard et al. and Wu et al. have shown that using deep learning, one can perform single-shot autofocusing (11, 12). Generative Adversarial Networks (GANs) is a deep-learning framework, which allows the training of generative models by performing an adversarial process between two deep-learning networks, a generator network and a discriminator network (13). In particular, deep convolutional GANs (DCGANs) have been shown successful for training generative models for image generation tasks (14, 15).

Recent attempts in virtual staining of individual cells based only on bright-field microscopy presented preliminary results (16, 17), but still lack the full information that is typical to chemical staining of individual cells under high-resolution light microscopy. Lately, the combination of holographic imaging and deep learning for classifying between different types of biological cells has been shown successful (18, 19). Furthermore, recently several deep-learning frameworks were used for performing virtual histology of biological tissue sections from autofluorescence signals and from quantitative phase images that were reconstructed from lens-free in-line holograms (20, 21). Specifically, ref. 21 has presented the PhaseStain method to virtually stain tissue slices, where the cells are arranged in the tissue structures, and based on this, tissue pathological analysis is performed. In this case, the cells and the inner organelles do not have the typical shapes as in individual cells. Hence, PhaseStain is suitable to replace histopathological tissue analysis by virtual staining, rather than analysis of single cells on a slide, sperm selection for IVF, or imaging flow cytometry. It thus remains unclear whether individual biological cells, rather than full tissue sections, can be virtually stained using only stain-free holographic imaging. In this paper, we show that this can be performed successfully, allowing its use for many medical and biological procedures including label-free in vitro diagnosis of individual cells. Our method, named HoloStain, uses DCGANs to transform quantitative phase images and phase-gradient images, extracted from stain-free digital holograms, to their stain-based versions that are similar to the conventional chemical staining images, making holographic imaging much more relevant for direct clinical use.

Results

Virtual Staining of Sperm Cells. We acquired 166 human sperm cells without staining using off-axis digital holographic microscopy, and then acquired the same cells after staining them by QuickStain using a conventional bright-field microscope. All images were acquired with a 60× oil-immersion microscope objective. The optical system details are presented in *Materials and Methods*. Next, we used image augmentation to create an eightfold increase in the dataset size. Overall, our dataset contained 1,328 image pairs of stain-free off-axis holograms of sperm cells, and their stain-based bright-field image counterparts. Each of the stain-free holograms was used to extract three images: a quantitative phase image and two synthetic gradient phase images in two orthogonal directions (see *Materials and Methods* for the digital processing used). These additional phase-gradient images were necessary for the success of the virtual staining process (thus, the quantitative phase images were not enough for the network convergence). It is important to note that only the stain-free digital hologram is acquired for the virtual staining. The quantitative phase image and the two phase gradients, which are fed through the model during the virtual staining stage, are derived from the digital hologram. These phase-gradient images can be regarded as hand-engineered features, which are used to improve the training process and overall create sharp, virtually stained images. Overall, for each cell, we had a batch of four images: stain-free quantitative phase image, two stain-free phase-gradient images, and chemical staining as the ground truth. We divided the data into 1,100 batches for training and 228 batches for testing. We then constructed a DCGAN model for obtaining virtual staining. The DCGAN framework, which is constructed from a generator network and a discriminator network competing with each other, is first trained on the 1,100 batches of sperm cells. The generator network receives as an input a batch of quantitative phase images and the two gradient-phase images, and outputs the generated virtually stained image. The discriminator network is trained to distinguish between the generated and the chemical staining images. It first receives both the generator input batch with the chemical staining image, and then receives the generator input batch with the generated output. By balancing between the loss functions of the generator and the discriminator, the generator is trained to create the correct virtually stained image. The full networks' architectures are given in *Materials and Methods*.

After training, the DCGAN model was tested on the 228 batches that were never seen by the model before. In this case, the generator was used in order to create the virtually stained images of the sperm cells, where the coinciding stain-based bright-field images were used for calculating a similarity metric between the real and generated images.

Fig. 1 presents examples of the results obtained by HoloStain on several sperm cells from the test dataset, never processed by the networks in the training step, having normal and pathological morphologies. Fig. 1*A* shows the stain-free off-axis holograms of the cells. Fig. 1 *B–D* show the coinciding quantitative phase images and gradient phase images, directly extracted from the holograms shown in Fig. 1*A*. This triplet of stain-free images is the input to the previously trained generator network. Fig. 1*E* shows the generated virtually stained images, the outputs of the trained generator network. Fig. 1*F* shows the chemical staining bright-field image of the coinciding cells, for comparison. The resulting virtual staining images in Fig. 1*E* have a similar color scheme to that of the chemical staining images in Fig. 1*F*. In addition, it can be seen that noise and debris are eliminated by the HoloStain method, resulting in a clean and even background surrounding the sample, which further eases the morphological examination of the cell.

For each of the 228 test images, the mean average error (MAE; *Materials and Methods*) was calculated between the virtually stained image and the chemical staining image, resulting in an overall MAE of 0.1566 ± 0.0446. In addition, in order to provide further insights regarding the comparison between the virtual and chemical staining images, a structural similarity (SSIM) index, as defined in ref. 22, was calculated. This resulted in an overall SSIM index of 0.8530 ± 0.0376.

In holographic imaging, the whole complex wavefront can be reconstructed from the captured holograms, allowing it to be propagated such that unfocused objects will come into focus. Thus, using HoloStain, we can now present the virtual staining images even if the cells have been out of focus during acquisition, which can help in increasing the acquisition throughput in comparison to bright-field imaging, even if the cells are chemically stained. Often, when imaging a certain population of cells, the clinician would need to constantly change the focus of the microscope in order to view all of the cells present. Using HoloStain, a single hologram can be captured with out-of-focus cells. Then, by reconstructing the whole complex wavefront, each cell can be propagated into focus and then virtually stained. Fig. 2 demonstrates an out-of-focus cell that is brought into focus by propagating the reconstructed complex wavefront (see *Materials and Methods* for details) and then virtually stained by HoloStain.

Classification of Sperm Cells. In order to assess the effectiveness of the virtually stained sperm cells for performing sperm-quality classification, five datasets of different sperm cells were created. The datasets for the analysis were created by sampling sperm

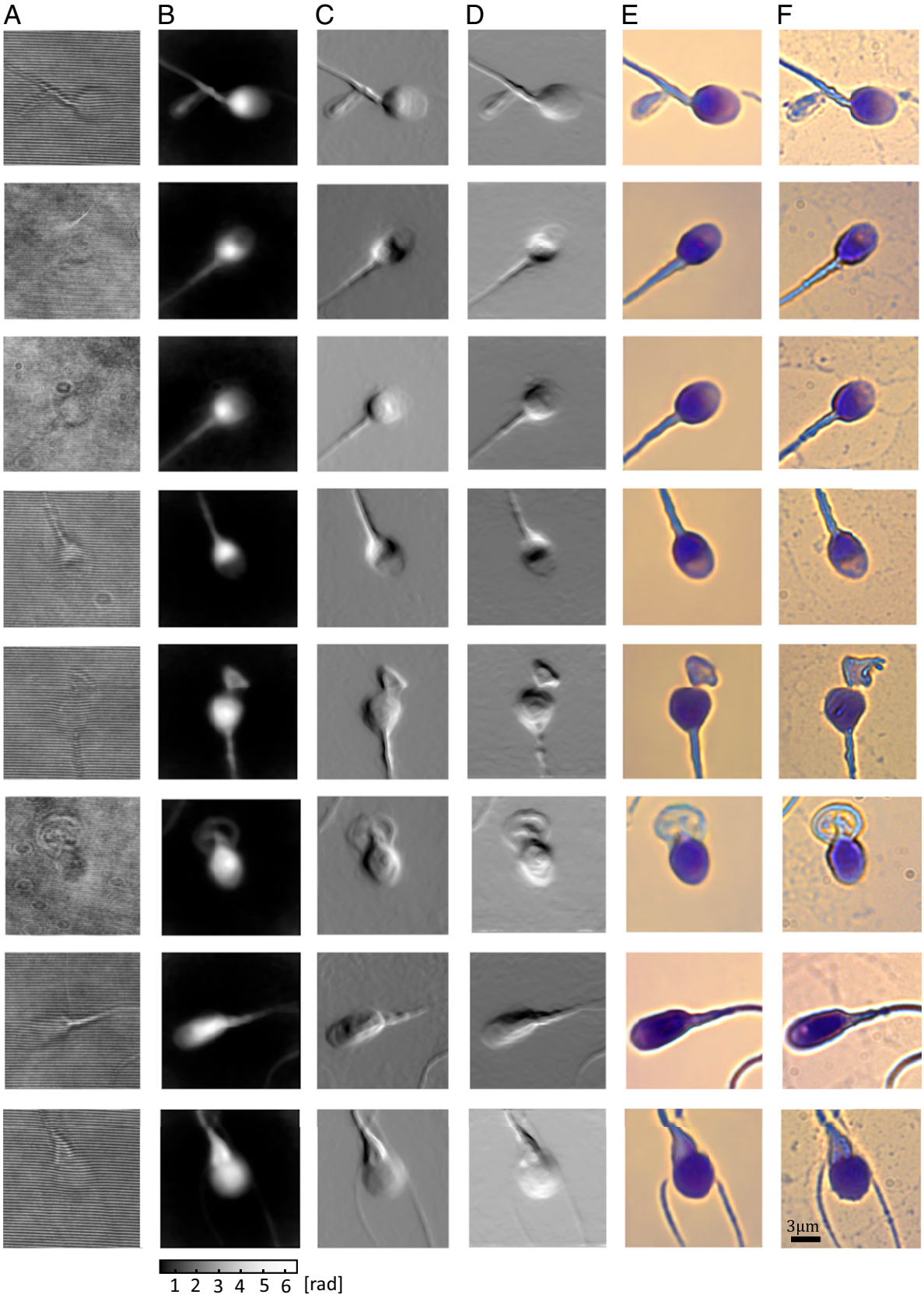

Fig. 1. Examples of HoloStain results for individual sperm cell imaging. The first four rows show morphologically healthy sperm cells. The last four rows show pathological cells. (*A*) Off-axis holograms of the cells acquired without staining. (*B*) The coinciding quantitative phase images extracted from the holograms. (*C*) The coinciding horizontal phase gradients extracted from the holograms. (*D*) The coinciding vertical phase gradients extracted from the holograms. (*E*) The coinciding virtual staining images, generated by the generator network, where *B*–*D* are the input to the generator. (*F*) The coinciding bright-field chemical staining images of the same cells, for comparison.

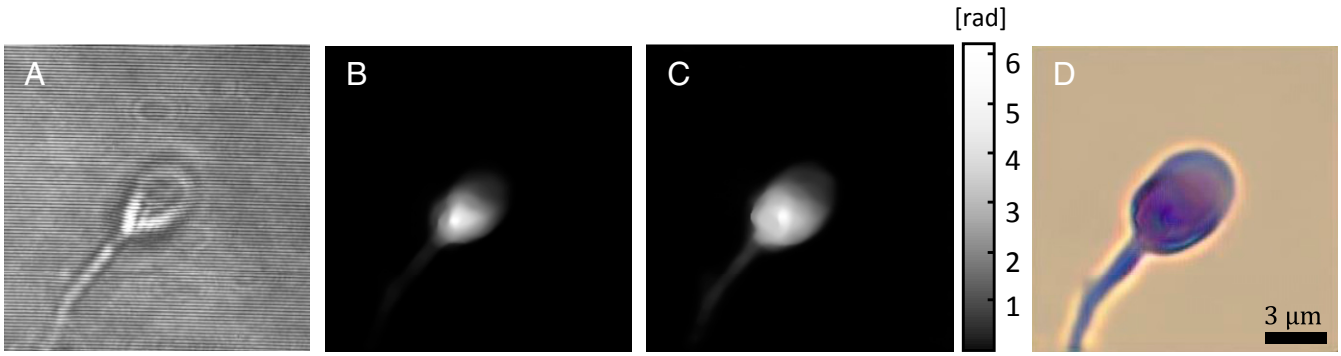

Fig. 2. Virtual staining of a sperm cell propagated into focus. (*A*) Hologram of an out-of-focus sperm cell. (*B*) Phase reconstruction of the out-of-focus sperm cell. (*C*) Phase reconstruction after propagating the complexed wavefront by $z = -1.6$ µm. (*D*) Virtual staining of the refocused sperm cell.

cells from nine different donors where the sample size for evaluation was chosen so as to be able to identify a difference of 15% with $P < 0.05$ and power of 80%. Furthermore, we chose to analyze datasets where the prevalence of morphologically healthy sperm cells was only 5%. By doing so, we were able to mimic the situation where healthy cells were rare, as this is the case in many IVF procedures. In addition, it is important to note that the goal in this analysis is to detect as many healthy cells as possible and not to give an overall diagnosis to a specific patient. The first dataset contained only bright-field images of sperm cells without staining, which is frequently the current practice in analyzing sperm cells today. The second dataset contained the bright-field images of the respective sperm cells, but now with staining using QuickStain. The third dataset contained images of the respective sperm cells that have been acquired without staining using off-axis holography and have been virtually stained using HoloStain. The fourth dataset contained the stain-free quantitative phase images of the respective sperm cells. The fifth dataset contained one of the stain-free phase-gradient images of the respective cells, which resembles differential interference contract (DIC) images. Since there is no well-established automatic standard for sperm cell morphological evaluation, we asked an experienced embryologist to analyze each sperm image in each of the five datasets and classify it, separately and independently, as normal or abnormal, using the WHO criteria for sperm cell analysis. The datasets were presented to the embryologist in a randomized and blinded manner up to four times in order to minimize the effect of subjective analysis. Four confusion matrices were calculated. Abnormal sperm cells were classified as negative labels - 0, normal sperm cells were classified as positive labels - 1, and the chemically stained sperm cells were regarded as the ground-truth labels.

For performing IVF procedures, where the selection of healthy sperm cells is considered critical, high precision for positive labels is required, where precision is defined by Eq. **1** below. Moreover, when the selection of several healthy sperm cells is needed, high recall is required as well, where recall (also called sensitivity) is defined by Eq. **2** below. Overall, an F1 score, defined by Eq. **3** below, can be calculated in order to quantify the balance between the precision and recall of the classified cells in each dataset. These three metrics are mathematically defined as follows:

$$Precision = \frac{TP}{TP + FP}, \quad [1]$$

$$Recall = \frac{TP}{FN + TP}, \quad [2]$$

$$\mathrm{F1} = 2 \times \frac{Precision \times Recall}{Precision + Recall}, \quad [3]$$

where *TP* signifies true positives: cells which are classified as positive and their corresponding chemically stained cells are classified as positive as well; *FP* signifies false positives: cells which are classified as positive and their corresponding chemically stained cells are classified as negative; *TN* signifies true negatives: cells which

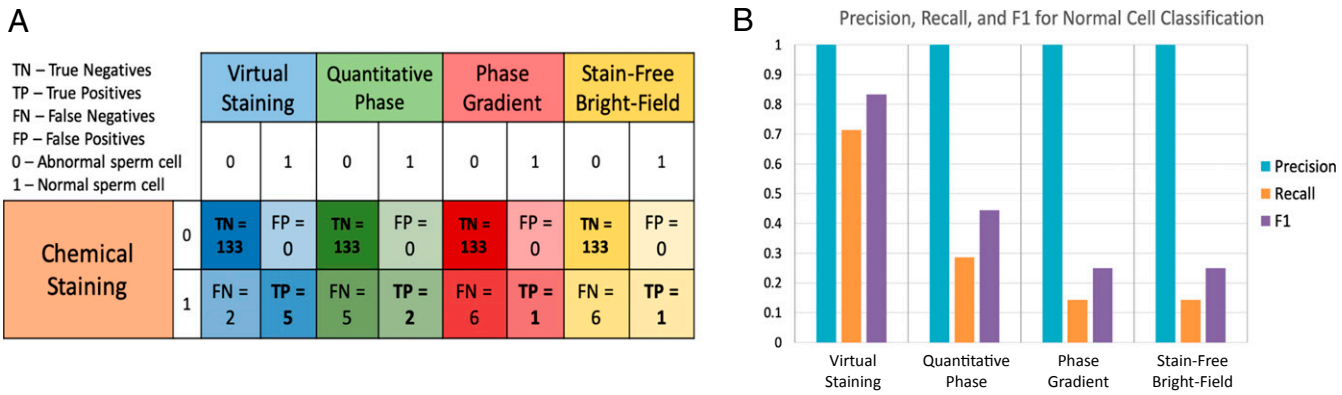

Fig. 3. Confusion matrices and metrics graph, when analyzing the sperm cells according to the WHO2010 protocol. (A) Confusion matrices for the classification of the virtually stained cells, the phase image of the cells, the synthetic phase-gradient image of the cells, and the stain-free bright-field image of the cells. (*B*) Comparison graph of the precision, recall, and F1 metrics for the four datasets.

are classified as negative and their corresponding chemically stained cells are classified as negatives as well; and *FN* signifies false negatives: cells which are classified as negative but their corresponding chemically stained cells are classified as positive.

From the confusion matrices in Fig. 3*A*, a precision of 1.0 was calculated across all datasets. This indicates that the embryologist was very conservative in classifying the cells, since he did not classify unhealthy cells as healthy cells in all methods in comparison to the chemical staining, even based on the stain-free bright-field images only. This indicated that his classification efficiency was low, since it would take him longer to choose morphologically healthy cells, especially in cases of pathologic sperm, where healthy cells are rare. Thus, virtual staining, providing contrast similar to chemical staining, is expected to make the embryologist's classification work in choosing healthy sperm cells less tedious by making him less conservative in classifying cells as healthy.

Furthermore, the recall gradually increased when advancing from stain-free bright-field imaging to virtual staining. The following recall values were calculated for the stain-free bright-field, phase-gradient, quantitative phase, and virtual staining images, respectively: 0.143, 0.143, 0.286, and 0.714. This shows that the virtual staining dataset enabled the detection of more normal-morphology sperm cells compared to the other datasets. Finally, the following F1 scores were calculated for the stain-free bright-field, phase-gradient, quantitative phase, and virtual staining images, respectively: 0.25, 0.25, 0.444, and 0.833. The F1 score signifies the overall accuracy of classifying healthy sperm cells by taking into account both the precision and recall. From the presented F1 scores, a gradual increase in classification performance can be seen, where out of the four stain-free methods analyzed, HoloStain enables classification results closest to the gold standard, the chemical staining method. A visualization of these metrics can be seen in Fig. 3*B*.

It should be emphasized that the analyses presented in Fig. 3 were performed using the strict protocol of WHO2010. In order to further highlight the advantage of virtual staining, we conducted an additional analysis, comparing our top two performing datasets—virtual staining and quantitative phase—to the chemical staining dataset. To do so, the experienced embryologist used the WHO1999 protocol, which employs less strict criteria, for selecting morphologically healthy cells. In this case, 51 healthy sperm cells were found in the chemical staining dataset. This resulted in a precision of 0.649 for the virtual staining dataset and a precision of 0.231 for the quantitative phase dataset. Furthermore, the virtual staining dataset received a recall of 0.585 and the quantitative phase dataset received a recall of 0.146. Thus, even when using less strict criteria, in which identified healthy cells are more abundant, the virtual staining dataset results in more than a twofold increase in precision and more than a fourfold increase in recall.

Moreover, other than achieving a level of analysis that is comparable with chemical staining, the experienced clinical embryologist (M.L.), who is also trained in analyzing sperm cells based on quantitative phase images, has highlighted several additional advantages of the HoloStain method. First, label-free bright-field images do not contain all of the necessary intracellular contrast needed to confidently provide accurate morphological analysis of sperm cells. Thus, taking advantage of the superior contrast and spatial information in the virtual staining images allows for a more efficient analysis procedure. Second, as previously shown, the HoloStain method creates a clean and even background free of noise and debris. This further simplifies the analysis process, even when compared to chemically stained cells, which often have contaminated backgrounds due to variability in the samples and the staining procedure. Third, it is important to note that the experienced clinical embryologist, who has analyzed the cells in this paper, has previous training and experience with analyzing quantitative phase images of sperm cells. Yet, the virtual staining images allowed for a more natural and straightforward analysis process, mostly due to the difficulty of seeing the outer dimensions and intracellular components of the cells in the quantitative phase and phase-gradient images. In addition, since the virtual staining images resemble the chemical staining images seen daily by clinicians, this capability eliminates the requirement for specifically training clinicians to accurately analyze quantitative phase images, and allows for an easier adoption process of holographic systems in existing biological and clinical laboratories.

Discussion

The capability of virtually staining label-free biological samples has great potential for replacing conventional staining techniques of individual cells, including fluorescence and histochemical staining. Virtual staining saves preparation time, it is less prone to variability caused by different staining protocols and environmental conditions, and it provides a solution for circumstances where staining is prohibited. Yet, it gives the clinician or the biologist cell visualization similar to actual chemical staining, so that established protocols for diagnosis or research can be directly applied. Our deep-learning–based technique, HoloStain, achieves virtual staining of quantitative phase images of individual biological cells acquired using a portable, clinic-ready off-axis holographic system that does not require cell staining. The reason we chose to perform virtual staining on label-free holography-based phase images rather than conventional phase contrast or label-free bright-field images is twofold. First, as opposed to standard phase-contrast methods, the phase images extracted from the holograms are quantitative images. Other than providing the necessary information to perform virtual staining, the quantitative nature of the images allows the creation of quantitative features that can then be used to assist with the overall analysis of the cells (23). Second, as shown in Fig. 3, performing morphological analysis of label-free bright-field images is nonsufficient. This is mostly due to the fact that label-free bright-field images do not contain all of the relevant contrast information that is necessary to perform accurate morphological analysis of sperm cells compared to the gold-standard staining method (24). As a result, rather than using label-free bright-field images, we chose to focus on using quantitative phase images for performing virtual staining since the necessary information required to create virtually stained images that are similar to the QuickStain method do not exist in the label-free bright-field images. Standard, nonquantitative phase imaging, such as Zernike's phase contrast and DIC, as the input to the network is not expected to yield good virtual staining results due to the lack of quantitative refractive-index information inside the organelles and distinctive typical imaging artifacts in these nonquantitative simple phase-imaging methods (such as halo and shadow effects). As demonstrated, HoloStain, which is based on quantitative phase imaging, is able to generate images of sperm cells that are similar to the conventional chemical staining method. Since reconstructing the full complex wavefront of the imaged samples is possible using holographic systems, the virtual staining of out-of-focus cells can be generated as well. As a result, constantly focusing the microscope during imaging is not needed, simplifying the analysis process for the clinician and increasing the analysis throughput. We have demonstrated that analyzing the virtually stained sperm cells by an experienced embryologist achieves similar performance compared to the classification of the coinciding chemically stained sperm cells, where the latter is currently regarded as the gold standard for performing morphological analysis in sperm cells. Furthermore, ref. 23 demonstrates the possibility of using machine learning for the automated analysis of individual sperm cells. Therefore, the combination of machine-learning algorithms and our HoloStain method can further ease the analysis process for clinicians. This can be done by utilizing the virtual staining images to review ambiguous borderline cases that could not be classified with high confidence using the machine-learning algorithm. Overall, we believe that HoloStain will provide a valuable tool for both

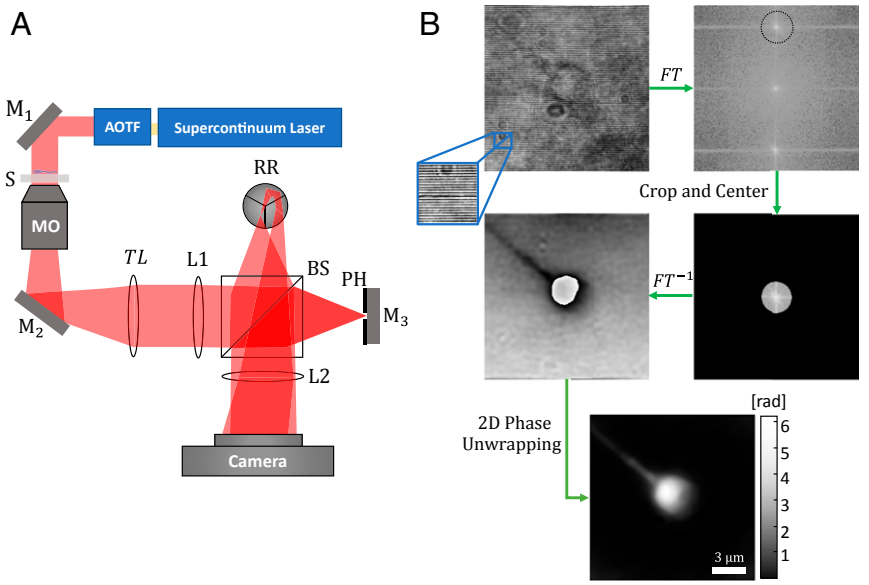

Fig. 4. Schematics of the optical setup and the reconstruction process. (*A*) The τ-Interferometer positioned at the output of a commercialized microscope, which consists of the following elements: a supercontinuum laser together with an acousto-optic tunable filter (AOTF) are used as the light source. M1, M2, and M3 are mirrors. S is the sample, MO is a microscope objective, L1 and L2 are lenses, and PH is the pinhole. (*B*) The reconstruction process: the captured hologram is first Fourier transformed (FT), then one of the cross-correlation terms is cropped and centered, then it is inverse Fourier transformed (FT^{-1}) and the phase argument is extracted. Finally, the phase argument undergoes a 2D unwrapping algorithm.

researchers and clinicians for performing stain-free morphological analysis of biological cells, saving them valuable preparation time and allowing them to perform a more accurate analysis when chemically staining cells is prohibited or is too expensive to perform. Although we have demonstrated using HoloStain for sperm imaging, the same platform can be adapted in imaging other types of cells, paving the way for stain-free digital pathology and stain-free imaging flow cytometry.

Materials and Methods

Sample Preparation and Imaging of Stain-Free Sperm Cells. The experiment was approved by the Tel Aviv University institutional review board (IRB) for human research. Human sperm cells were obtained from the ejaculate of 18–40-y-old anonymous donors, after they signed an IRB-approved informed-consent form. A drop of 5–10 μL of sperm was smeared onto several clean microscopic slides with a 80-μm × 80-μm square grid painted onto them for localization of the sperm cells when transferring the samples between the optical systems. These smeared drops were then left to dry for 5 min and then fixed to the slides with 98% ethanol for 10 min. The slides were then imaged using the IPM system, which is shown in Fig. 4*A*. This system consisted of the τ-Interferometer connected at the output of an inverted microscope. A supercontinuum fiber light source (SC-400–4 Fianium) connected to an acousto-optic tunable filter (SC-AOTF, Fianium) was used as the light source for the inverted microscope, emitting wavelengths of 532 ± 3.1 nm. The beam first passed through the sample, then magnified using the microscope objective MO (63×, 1.4 numerical aperture, oil immersion, infinity-corrected) and passed through a spherical tube lens TL (150-mm focal length). Then, it passed through lens L1 (100-mm focal length), which Fourier transformed the beam, and beam splitter BS split the beam into two separate beams. One beam passed straight through the BS and then reflected back and shifted by retroreflector RR. This beam was then reflected by the BS and inverse Fourier transformed by lens L2 (150-mm focal length) onto a digital camera (1,280 × 1,024 pixels, pixel size of 5.2 μm, DCC1545M, Thorlabs). This beam acted as the sample beam in this interferometric setup. The second beam was reflected by the BS onto a mirror-pinhole configuration, PH and M3, which spatially filtered the beam, thus erasing the sample information, creating the reference beam. This beam was then reflected back and passed through the BS, where it was then inverse Fourier transformed by lens L2 and interfered with the sample beam on the camera. The final result was an off-axis interference pattern, which was then transferred to the computer for further digital analysis.

Imaging of Stained Sperm Cells. After the sperm cells were imaged using the IPM system, they were stained using QuickStain (Biological Industries) and left to dry for 15 min. Then, using the 80-μm × 80-μm square grid, the field of views captured using the IPM system were located once again and imaged using a bright-field microscope (Axio Observer D1, Zeiss).

Digital Reconstruction of the Holograms. The off-axis interference pattern captured by the camera can be used to extract the complex wavefront. This reconstruction process is illustrated in Fig. 4*B*. Shortly, this off-axis hologram is digitally Fourier transformed, resulting in a zero order and two high-order cross-correlation terms. Each cross-correlation term contains the complex wavefront of the sample, which allows the extraction of the cell quantitative phase information. One of the cross-correlation terms is digitally cropped and inverse Fourier transformed. Then, in case the sperm image is out of focus, a digital propagation algorithm is applied. Our propagation method of choice was the Rayleigh–Sommerfeld propagation of the angular spectrum (25). Finally, the phase information was extracted from the argument of the resulting complex wavefront, which then underwent a two-dimensional (2D) phase-unwrapping algorithm (26).

Calculation of Synthetic Phase-Gradient Images from the Phase Images. In order to enhance high-frequency spatial information in the cell images, such as edges, and help with the training process, two synthetic phase-gradient images were created from each phase image. Those images were generated by shifting the quantitative phase images by one pixel in one of the spatial directions (x or y) and then subtracting the shifted image from the original phase image.

$$Grad\varphi_x = \varphi(x, y) - \varphi(x+1, y), \quad [4]$$

$$Grad\varphi_y = \varphi(x, y) - \varphi(x, y+1), \quad [5]$$

where φ is the quantitative phase of the sample extracted from the off-axis hologram. The result of these phase gradients resembles what can be obtained experimentally using a DIC microscope.

Digital Preprocessing. As with the holograms, the bright-field images of the stained sperm cells were cropped into 256 × 256 pixels. This resulted in two datasets, one containing the bright-field images of the stained sperm cells, and the other containing the quantitative phase images and the synthetic phase-gradient images of the same sperm cells. It should be noted that during the cropping stage, the fields of view of the two datasets were registered. This was attained by performing 2D correlation between the datasets and detecting the center of the target field of view in each dataset. As a result, an exact overlap between the fields of view of the two datasets was achieved.

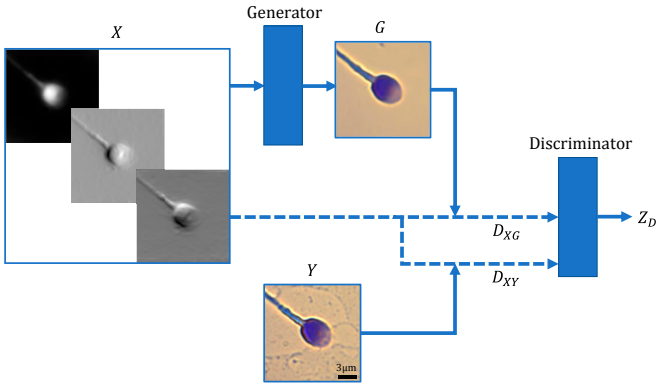

Fig. 5. Schematics of the training process. *X* is the input to the generator network consisting of the stain-free quantitative phase image and two synthetic phase-gradient images. *Y* is an image of the chemically stained sperm cell. The generator is trained to create *G*, the image of the virtually stained sperm cell. In one instance, *G* and *X* (marked as D_{XG}) are fed through the discriminator, which is trained to recognize this pair as fake images. In another instance, *Y* and *X* (marked as D_{XY}) are fed through the discriminator which is trained to recognize this pair as real images, while Z_D is the output of the discriminator.

After constructing the abovementioned dataset, it was further augmented by performing 90° rotations for each image, and then horizontally flipping all existing and new images in the dataset. Overall, this caused the original dataset to increase by eightfold.

Training and Testing Procedures. To train a deep-learning model to virtually stain sperm cells, a DCGAN framework was used. This framework consisted of a generator network that was trained to create the virtually stained images from the stain-free quantitative phase and synthetic phase-gradient images of the cells, and a discriminator network that was trained to discriminate between the generated and the chemical staining images.

As seen in Fig. 5, in order to train the generator and discriminator networks, the generator receives an input batch *X*, which is a concatenation between the quantitative phase image and the two synthetic phase-gradient images of the sperm cells, all extracted from the stain-free digital holograms of these cells. It is trained to generate *G*, which is the virtually stained image of the same sperm cell that has been fed through the generator network. Since the discriminator is trained to distinguish between the generated and chemical staining images, in one case it receives D_{XY}, which indicates that the generator input *X* is fed through the discriminator together with the chemically stained sperm cell image *Y*. In another case, the discriminator receives D_{XG}, which indicates that the generator input *X* is fed through the discriminator together with the generated virtual staining image *G*.

The losses for the networks use a combination of two error functions. The first is the MAE, also known as $\mathcal{L}_1$ loss, which is calculated as follows:

$$\mathcal{L}_1(Y,G)=\frac{\sum_{i=1}^{n}|Y_i-G_i|}{n}. \quad [6]$$

The second is the sigmoid cross-entropy (SCE), which is calculated as follows:

$$SCE(Z_D,Z)=\max(Z_D,0)-Z_D*Z+\log\{1+\exp[-abs(Z_D)]\}, \quad [7]$$

where Z_D is the output of the discriminator and *Z* is the designated Boolean (1 for real images and 0 for fake images).

Overall, the generator loss is calculated using the following equation:

$$\mathcal{L}_G=\beta\mathcal{L}_1(Y,G)+SCE(Z_{XG},1), \quad [8]$$

where Z_{XG} is the output of the discriminator when D_{XG} is fed through it, β is a multiplication factor used to give an emphasis on generating accurate virtual staining images; this value was set to 100 (15). The discriminator loss is calculated as follows:

$$\mathcal{L}_D=SCE(Z_{XG},0)+SCE(Z_{XY},1), \quad [9]$$

where Z_{XY} is the output of the discriminator when D_{XY} is fed through it.

During the training stage, the generator loss and the discriminator loss were minimized using the Adam optimizer (27). In addition, the generator

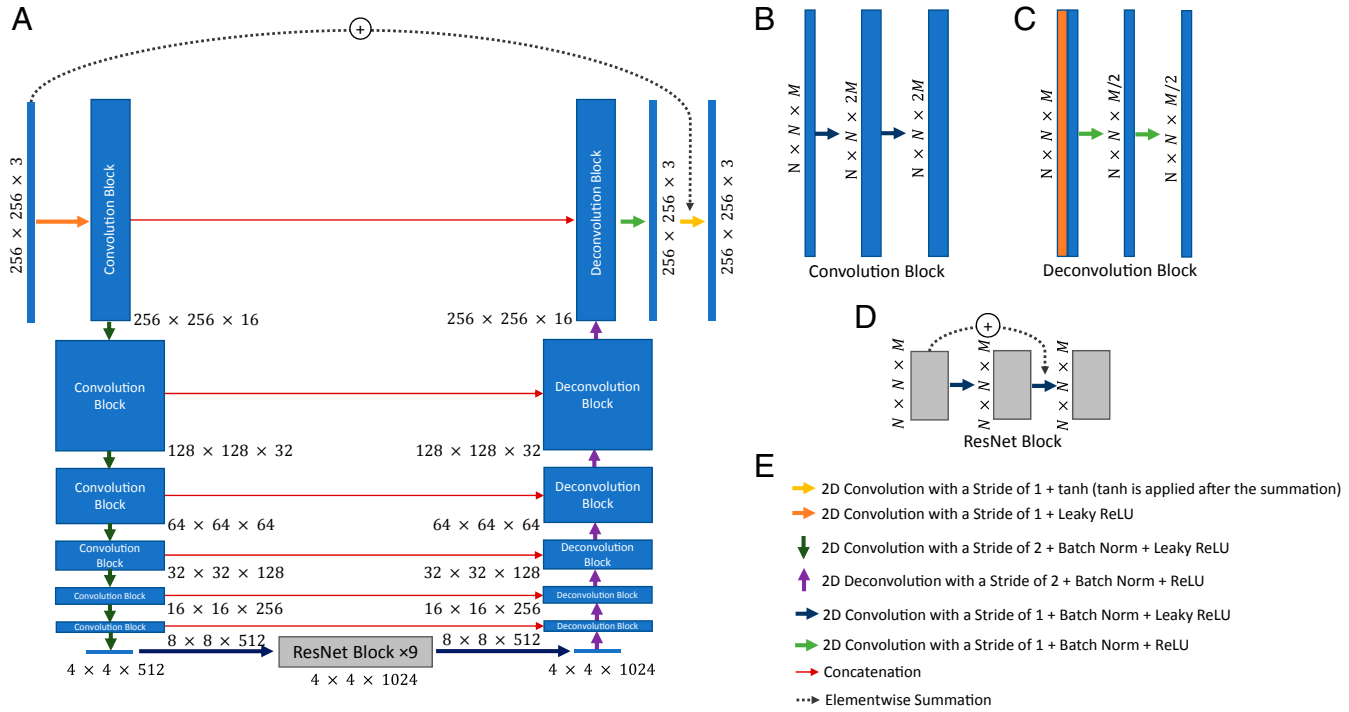

Fig. 6. Architecture of the generator network. (*A*) The overall architecture of the generator network. (*B*) The inner architecture of a convolution block. (*C*) The inner architecture of a deconvolutional block, where the orange rectangle signifies the concatenation process. (*D*) The inner architecture of a residual network block. (*E*) Legend explaining the signification of each arrow in the architectures.

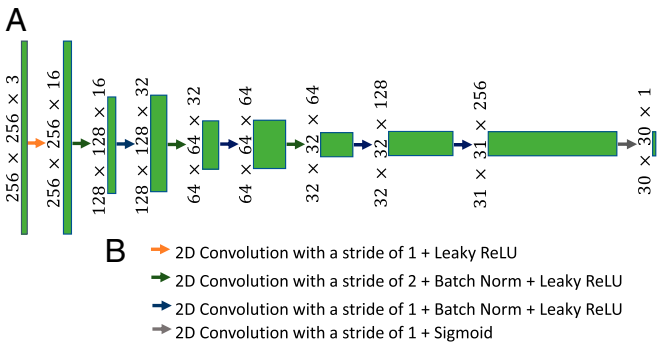

Fig. 7. Schematics of the discriminator network. (*A*) The overall architecture of the generator network. (*B*) Legend explaining the signification of each arrow in the architecture.

and discriminator use three types of activation functions. The first is a rectified linear unit (ReLU), which is calculated as follows:

$$ReLU(X) = \max(X, 0). \quad [10]$$

The second is a leaky ReLU, which is calculated as follows:

$$LeakyReLU(X) = \max(X, 0.2X). \quad [11]$$

The third is a sigmoid function, which is calculated as follows:

$$Sigmoid(X) = \frac{1}{1 + e^{-X}}. \quad [12]$$

Finally, also a hyperbolic tangent (tanh) is used, which can be calculated as follows:

$$\tanh(X) = \frac{e^{2X} - 1}{e^{2X} + 1}. \quad [13]$$

Internal Architecture of the Deep-Learning Networks. As seen in Fig. 6, the generator network is based on a U-Net architecture (28). This architecture consists of an encoder and a decoder with skip connections at every downsampling/upsampling stage. Each step of the encoder contains a convolutional block. Each convolutional block contains three sequences of a 2D convolution layer, a batch normalization layer (29), and a leaky ReLu activation function, as calculated in Eq. **11**. The first and second convolutions in every step of the encoder consist of a convolutional layer with a kernel of 4 and a stride of 1, and the third block contains a convolutional layer with a kernel of 4 and a stride of 2. Overall, in each step of the encoder there is an increase of the depth of the filters by a factor of 2 and a decrease by a factor of 2 in the height and width dimensions. After the encoding step, nine residual network (ResNet) blocks were added in order to assist with the image transformation training of the generator (30, 31). The decoding stage consists of deconvolutions, a concatenation step for the skip connections, and two additional convolutional layers. The deconvolution step is made up of a sequence of a transpose 2D convolution layer with a kernel of 4 and a stride of 2, a batch normalization layer, and a ReLu activation function, as calculated in Eq. **10**. This deconvolution step is followed by two sequences of a convolutional layer with a kernel of 4 and a stride of 1, a batch normalization layer, and a ReLu activation function. Overall, at each step of the decoder, the depth of the filters decreases by a factor of 2 and the height and width dimensions increase by a factor of 2. Furthermore, an additional skip connection is added at the final layer of the decoder, which performs an elementwise summation between the input image and the final layer of the generator, in order to decrease training time and achieve an image with a geometrical similarity to the input image (10).

As seen in Fig. 7, the discriminator model consists of convolutional blocks that are similar to the ones in the encoder step of the generator. As the input image passes along the discriminator, its depth is increased by a factor of 2 and the height and width dimensions are decreased by a factor of 2 until a 32 × 32-pixel image is created. The final two convolutional layers create a 30 × 30-pixel image with a depth of 1, and by applying the sigmoid function given in Eq. **12**, each pixel in this image corresponds to the real or fake classification of overlapping patches within the input image. This framework was based on the PatchGAN discriminator (15), which decreases training time and improves the sharpness of the generated images.

Implementation. The hologram reconstruction, synthetic phase-gradient image calculation, and all of the digital preprocessing procedures performed on the images were implemented with MATLAB R2016b. All of the abovementioned processes were done on a desktop computer with an Intel Core i7-2600 CPU @ 3.40 GHz and 8.00 GB RAM, running on a Windows 10 operating system (Microsoft). The deep learning architecture and training/testing procedures were implemented in Python version 3.6.4 using the TensorFlow library version 1.10.0. The training and testing of the network were performed on a Tesla P100 GPU (NVIDIA) using the Google Cloud Platform. The framework was trained for 120 epochs, which lasted 31.5 h. Each image generation lasts ∼0.08 s on a NVIDIA Tesla P100 GPU.

Data Availability. All data and codes that support the results within this paper are available in *SI Appendix* and Dataset S1.

ACKNOWLEDGMENTS. This research was funded by the Horizon2020 European Research Council (ERC) 2018 Proof of Concept (PoC) Grant 838359 to N.T.S.

1. Y.-K. Park, C. Depeursinge, G. Popescu, Quantitative phase imaging in biomedicine. *Nat. Photonics* **12**, 578–589 (2018).
2. P. Girshovitz, N. T. Shaked, Generalized cell morphological parameters based on interferometric phase microscopy and their application to cell life cycle characterization. *Biomed. Opt. Express* **3**, 1757–1773 (2012).
3. M. Lee, O. Yaglidere, A. Ozcan, Field-portable reflection and transmission microscopy based on lensless holography. *Biomed. Opt. Express* **2**, 2721–2730 (2011).
4. P. Girshovitz, N. T. Shaked, Compact and portable low-coherence interferometer with off-axis geometry for quantitative phase microscopy and nanoscopy. *Opt. Express* **21**, 5701–5714 (2013).
5. H. Greenspan, B. V. Ginneken, R. M. Summers, Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Trans. Med. Imaging* **35**, 1153–1159 (2016).
6. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436–444 (2015).
7. S. Liao, Y. Gao, A. Oto, D. Shen, "Representation learning: A unified deep learning framework for automatic prostate MR segmentation" in *Medical Image Computing and Computer-Assisted Intervention: Lecture Notes in Computer Science*, K. Mori, I. Sakuma, Y. Sato, C. Barillot, N. Navab, Eds. (Springer, Berlin, Germany, 2013), pp. 254–261.
8. P. Liskowski, K. Krawiec, Segmenting retinal blood vessels with deep neural networks. *IEEE Trans. Med. Imaging* **35**, 2369–2380 (2016).
9. M. Havaei *et al.*, Brain tumor segmentation with deep neural networks. *Med. Image Anal.* **35**, 18–31 (2017).
10. M. T. Mccann, K. H. Jin, M. Unser, Convolutional neural networks for inverse problems in imaging: A review. *IEEE Signal Process. Mag.* **34**, 85–95 (2017).
11. H. Pinkard, Z. Phillips, A. Babakhani, D. A. Fletcher, L. Waller, Deep learning for single-shot autofocus microscopy. *Optica* **6**, 794–797 (2019).
12. Y. Wu *et al.*, Bright-field holography: Cross-modality deep learning enables snapshot 3D imaging with bright-field contrast using a single hologram. *Light Sci. Appl.* **8**, 25 (2019).
13. I. Goodfellow *et al.*, "Generative adversarial nets" in *Annual Conference on Neural Information Processing Systems: Advances in Neural Information Processing Systems 27*, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence and K. Q. Weinberger, Eds. (NIPS, Montreal, Quebec, Canada, 2014), pp. 2672–2680.
14. A. Radford, L. Metz, S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv:1511.06434 (7 January 2016).
15. P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, "Image-to-image translation with conditional adversarial networks" in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, Piscataway, NJ, 2017), pp. 1125–1134.
16. E. M. Christiansen *et al.*, In silico labeling: Predicting fluorescent labels in unlabeled images. *Cell* **173**, 792–803.e19 (2018).
17. C. Ounkomol, S. Seshamani, M. M. Maleckar, F. Collman, G. R. Johnson, Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy. *Nat. Methods* **15**, 917–920 (2018).
18. C. L. Chen *et al.*, Deep learning in label-free cell classification. *Sci. Rep.* **6**, 21471 (2016).
19. M. Rubin *et al.*, TOP-GAN: Stain-free cancer cell classification using deep learning with a small training set. *Med. Image Anal.* **57**, 176–185 (2019).
20. Y. Rivenson *et al.*, Deep learning-based virtual histology staining using auto-fluorescence of label-free tissue. arXiv:1803.11293 (30 March 2018).

21. Y. Rivenson *et al*., PhaseStain: The digital staining of label-free quantitative phase microscopy images using deep learning. *Light Sci. Appl.* **8**, 23 (2019).
22. Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004).
23. S. K. Mirsky, I. Barnea, M. Levi, H. Greenspan, N. T. Shaked, Automated analysis of individual sperm cells using stain-free interferometric phase microscopy and machine learning. *Cytometry A* **91**, 893–900 (2017).
24. M. Haifler *et al*., Interferometric phase microscopy for label-free morphological evaluation of sperm cells. *Fertil. Steril.* **104**, 43–7.e2 (2015).
25. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, 2005).
26. P. Girshovitz, N. T. Shaked, Fast phase processing in off-axis holography using multiplexing with complex encoding and live-cell fluctuation map calculation in real-time. *Opt. Express* **23**, 8773–8787 (2015).
27. D. P. Kingma, J. L. Ba, Adam: A method for stochastic optimization. arXiv:1412.6980 (30 January 2017).
28. O. Ronneberger, P. Fischer, T. Brox, *U-net: Convolutional Networks for Biomedical Image Segmentation* (Lecture Notes in Computer Science, Medical Image Computing and Computer-Assisted Intervention, 2015).
29. S. Ioffe, S. Christian, Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv:1502.03167 (2 March 2015).
30. K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual learning for image recognition" in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, Piscataway, NJ, 2016), pp. 770–778.
31. J.-Y. Zhu, T. Park, P. Isola, A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks" in *Proceedings of the IEEE International Conference on Computer Vision* (IEEE, Piscataway, NJ, 2017), pp. 2223–2232.