Patch-Based U-Net Model for Isotropic Quantitative Differential Phase Contrast Imaging

An-Cin Li, Sunil Vyas, Yu-Hsiang Lin, Yi-You Huang, Hsuan-Ming Huang, and Yuan Luo

***Abstract*— Quantitative differential phase-contrast (qDPC) imaging is a label-free phase retrieval method for weak phase objects using asymmetric illumination. However, qDPC imaging with fewer intensity measurements leads to anisotropic phase distribution in reconstructed images. In order to obtain isotropic phase transfer function, multiple measurements are required; thus, it is a time-consuming process. Here, we propose the feasibility of using deep learning (DL) method for isotropic qDPC microscopy from the least number of measurements. We utilize a commonly used convolutional neural network namely U-net architecture, trained to generate 12-axis isotropic reconstructed cell images (i.e. output) from 1-axis anisotropic cell images (i.e. input). To further extend the number of images for training, the U-net model is trained with a patch-wise approach. In this work, seven different types of living cell images were used for training, validation, and testing datasets. The results obtained from testing datasets show that our proposed DL-based method generates 1-axis qDPC images of similar accuracy to 12-axis measurements. The quantitative phase value in the region of interest is recovered from 66% up to 97%, compared to ground-truth values, providing solid evidence for improved phase uniformity, as well as retrieved missing spatial frequencies in 1-axis reconstructed images. In addition, results from our model are compared with paired and unpaired CycleGANs. Higher PSNR and SSIM values show the advantage of using the U-net model for isotropic qDPC microscopy. The proposed DL-based method may help in performing high-resolution quantitative studies for cell biology.**

Manuscript received March 31, 2021; revised May 14, 2021; accepted June 12, 2021. Date of publication June 21, 2021; date of current version October 27, 2021. This work was supported in part by the Taiwan Ministry of Science and Technology (MOST) under Grant MOST 108-2221-E-002-168-MY4; in part by the National Taiwan University (NTU) under Grant 08HZT49001, Grant 108L7714, and Grant 109L7839; and in part by the NTU, SPARK Program under Grant MOST 109-2926-B-002-001. *(Corresponding authors: Hsuan-Ming Huang; Yuan Luo.)*

An-Cin Li, Sunil Vyas, Yu-Hsiang Lin, and Hsuan-Ming Huang are with the Institute of Medical Device and Imaging, College of Medicine, National Taiwan University, Taipei 10051, Taiwan (e-mail: r08458005@ntu.edu.tw; sunilk_vyas@yahoo.com; ggg77882@gmail.com; hmhuang1983@ntu.edu.tw).

Yi-You Huang is with the Department of Biomedical Engineering, National Taiwan University, Taipei 10051, Taiwan, and also with the National Taiwan University Hospital, Taipei 10051, Taiwan (e-mail: yyhuang@ntu.edu.tw).

Yuan Luo is with the Institute of Medical Device and Imaging, College of Medicine, National Taiwan University, Taipei 10051, Taiwan, also with the Department of Biomedical Engineering, National Taiwan University, Taipei 10051, Taiwan, also with the Molecular Imaging Center, National Taiwan University, Taipei 10672, Taiwan, and also with the Yong Lin Institute of Health, National Taiwan University, Taipei 10087, Taiwan (e-mail: yuanluo@ntu.edu.tw).

This article has supplementary downloadable material available at https://doi.org/10.1109/TMI.2021.3091207, provided by the authors.

Digital Object Identifier 10.1109/TMI.2021.3091207

***Index Terms*— Coded-illumination, deep neural network, differential phase contrast, patch-based, quantitative phase imaging, U-net model, CycleGAN.**

I. Introduction

OPTICAL microscopy provides a variety of medical applications, and has been widely utilized by researchers and clinicians. Many biological objects of interest are composed of weak phase features with minimal phase changes [1]. Hence, weak phase objects (i.e. living cells) show poor contrast, and they are barely observable without fluorescent labeling, which induces significant phototoxic side effects in live biological objects [2], [3]. Phase contrast microscopy [4] is a label-free imaging technique to observe small phase changes, which are invisible in a common bright-field microscope. Although a standard phase contrast microscope qualitatively makes phase changes visible, it does not provide true phase imaging in quantitative fashion, and this limits its use in imaging weak phase biological objects.

In contrast to phase contrast microscopy, quantitative phase imaging (QPI) quantifies phase changes of weak phase objects since phase of resultant label-free images can be directly retrieved from intensity measurements. Various QPI methods have been developed for biomedical applications [5], [6]. One of the most commonly used QPI methods is digital holographic imaging [1], [5], [7]-[9], which requires stringent interferometric experimental conditions, with inherent speckle noise effects that degrade image quality. Alternatively, among non-interferometric QPI methods, quantitative differential phase contrast (qDPC) microscopy retrieves accurate phase information with simple system configuration, and requires no interferometric measurements [4], [10]. In contrast to interferometric based QPI with limited spatial resolution, qDPC microscopy, under partially coherent light (i.e. LEDs), provides two times better resolution [1], [11], and alleviates the speckle problems raised by coherent laser light.

Typically, in qDPC microscopy, the quantitative phase is retrieved using an asymmetric half-circle illumination [4] with 2-axis pairwise (i.e. four-angle) intensity measurements. However, missing spatial frequencies in retrieved phase images produces artifacts due to anisotropic phase transfer function. In order to obtain missing spatial frequencies during quantitative phase recovery, efforts have been made to obtain isotropic (i.e. circularly symmetric) transfer function by developing different designs of pupil patterns [11]–[13]. Recently, qDPC imaging using 12-axis measurements, along twenty-four different angles [14], has been reported to obtain isotropic

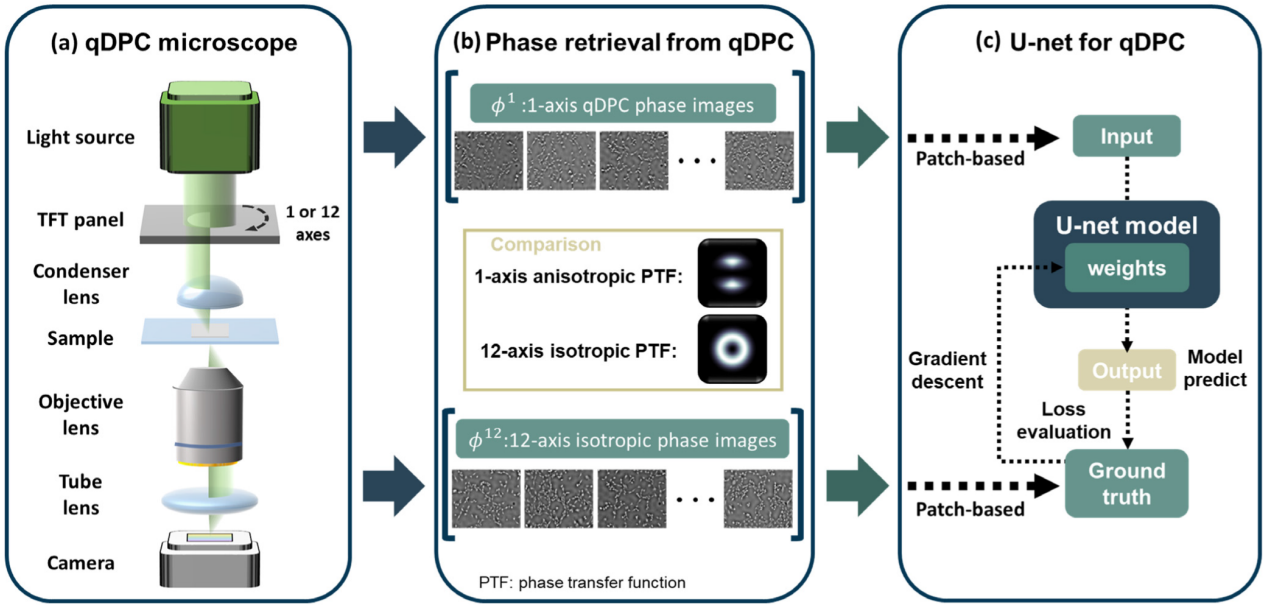

Fig. 1. Flow chart of DL based qDPC microscopy. (a) Schematic diagram of the qDPC microscope using an asymmetric half-circle pupil for intensity measurements. A TFT panel is positioned at the Fourier plane to generate an asymmetric half-circle pupil. (b) Experimentally obtained phase image dataset for training a U-net model, and comparison of phase transfer functions using 1-axis qDPC and 12-axis isotropic qDPC. (c) The U-net model for qDPC microscopy. For training U-net model phase reconstruction from 1-axis qDPC method is regarded as input dataset, and phase reconstruction from 12-axis isotropic qDPC method is regarded as ground-truth dataset, respectively. The U-net model takes the input data one-by-one and predicts each corresponding results. The output and ground-truth images are utilized for calculating the loss value for updating weights in each hidden layer of the model.

phase transfer function, with significantly improved stability and accuracy for quantitative phase recovery. On the other hand, additional intensity acquisition is time-consuming and causes slow phase recovery, which significantly limits temporal resolution for qDPC in high-speed operation, such as real-time observation and live cell imaging applications.

Data-driven methods of deep neural networks (DNN) are able to enhance efficiency and performance of optical imaging [14]–[16]. The deep learning (DL) approach has been applied for different optical microscopes [17]–[21], including interferometric based QPI [22]. Recently, a physics-based DL method has also used in qDPC to optimize illumination conditions for different pupil patterns [23]. Here, the primary objective is to demonstrate the implementation of deep learning to ease experimental requirement by performing image-to-image translation from anisotropic to isotropic phase images. We adopted a U-net model for isotropic quantitative phase retrieval in qDPC microscopy. Our U-net model results using 1-axis pairwise (i.e. two-angle) intensity measurements demonstrate similar accuracy as compared to 12-axis isotropic qDPC (i.e. twenty-four angle) measurements. In addition, utilization of the patch-based approach provides sufficient image data for the model training. The small size of patched images enables the U-net model to be trained with a higher efficiency.

To the best of our knowledge, this is the first report on applying a U-net model to predict the isotropic quantitative phase distribution, with the least number of intensity measurements. To demonstrate our approach, we conducted experiments with seven types of living cells. Using three different metrics, the performance of the U-net model for qDPC was evaluated from experimental data to train, validate and test using different living cells. Accuracy of recovered quantitative phase from our well-trained U-net model is measured for different loss functions. To further analyze quantitative features of cells using our approach, dry mass from measured values of the quantitative phase was calculated. The results show that our model provides accurate phase values, as well as dry mass density maps. In addition, results from our model are compared with paired and unpaired Cycle GANs. Higher PSNR and SSIM values show the advantage of using U-net for isotropic qDPC microscopy. The deviation of dry mass values predicted from our model is less than 5% of the estimated values.

II. Quantitative Differential Phase Contrast Microscopy

A. Quantitative Differential Phase Contrast Microscopy

The flow chart of DL based qDPC microscopy is shown in Fig. 1. A qDPC microscope is designed based on an inverted microscope (Leica, DMI 3000 B) as shown in Fig. 1 (a). A 10× objective lens (HC/X PL Fluotar 0.3NA) and a condenser lens (Leica, S28 module) are used. Half-circle pupil patterns are generated and controlled through a TFT panel (Seeed, 2.8" TFT Touch Shield V2.0), which is located at the front focal plane of the condenser lens [24]. Pairwise images are captured by a CCD camera (Tucsen, TCH-5.0ICE 5.0MP). In Fig. 1 (b), qDPC images using 1-axis and 12-axis intensity measurements are reconstructed for anisotropic and isotropic phase images, respectively. In Fig. 1 (c), the patch-based method is used for training our U-net model to predict the 12-axis isotropic qDPC images from 1-axis qDPC images. The U-net architecture will be discussed in detail in section III.

B. Principle of Quantitative Differential Phase Contrast Microscopy

Under partially coherent light, the intensity measurement at the image plane can be written as [12], [24]:

$$I(r) = \iint \left| \mathcal{F}\left\{ \mathcal{F}\left[\sqrt{LS(u)} e^{i2\pi ur} O(r) \right] PF(u) \right\} \right|^2 d^2u, \quad (1)$$

where $LS(u) = m(u)circ(u/\rho_c)$. $m(u)$ denotes half-circle amplitude mask function, $\rho_c = NA_{condenser}/\lambda$, and $circ(u/\rho_c)$ is the pupil function of an objective lens. λ denotes the central wavelength of illumination in a qDPC microscope. $r = (r_x, r_y)$ denotes spatial coordinate, and $u = (u_x, u_y)$ denotes the coordinates in the spatial frequency domain. An object can be expressed as a complex function $O(r) = exp(-\alpha(r) + i\phi(r))$, where $\alpha(r)$ and $\phi(r)$ denote the amplitude and phase distribution. $PF(u)$ is the pupil function of the objective lens. $\mathcal{F}$ denotes the Fourier transform induced by a condenser lens and an objective lens.

With weak object approximation, the object function can be simplified as $O(r) = 1 - \alpha(r) + i\phi(r)$. According to this assumption, the corresponding spectrum of (1) in the spatial-frequency domain can be re-written as [24]:

$$\tilde{I}(u) = \tilde{I}_{background} \cdot \tilde{\delta}(u) - \tilde{H}_{amplitude}(u) \cdot \tilde{\alpha}(u) + i\tilde{H}_{phase}(u) \cdot \tilde{\phi}(u), \quad (2)$$

where $\tilde{I}_{background}$ denotes background DC term, $\tilde{H}_{amplitude}(u)$ denotes the amplitude transfer function, and $\tilde{H}_{phase}(u)$ denotes phase transfer function. $\tilde{\delta}(u)$ is the Fourier transform corresponding to background term. $\tilde{\alpha}(u)$ and $\tilde{\phi}(u)$ are the amplitude and phase spectrums, respectively evaluated following earlier works [12], [24]. For the *j-th* axis measurement, pairwise images are captured on the camera as intensity distribution $(I_{1,j}, I_{2,j})$. Accordingly, the phase contrast image along the *j-th* axis is given by $I_{DPC,j} = (I_{1,j} - I_{2,j})/(I_{1,j} + I_{2,j})$. The retrieved phase is given as [24]:

$$\phi(r) = \mathcal{F}^{-1}\left\{ \frac{\sum_{j=1}^{n} \tilde{H}_{DPC,j} \cdot \tilde{I}_{DPC,j}(u)}{\sum_{j=1}^{n} \left| \tilde{H}_{DPC,j} \right|^2 + \gamma} \right\}, \quad (3)$$

where n is the number of measurements, and γ is the regularization parameter for preventing singularity in the reconstructed phase. In addition, the reconstruction procedure involves multiple Fourier transform operations which reduces reconstruction speed. Phase transfer functions for each measurement can be written as:

$$\tilde{H}_{DPC,j}(u) = \left[\tilde{H}_{phase,1,j}(u) - \tilde{H}_{phase,2,j}(u) \right] / 2\tilde{I}_0, \quad (4)$$

where $\tilde{H}_{phase,1,j}(u)$ and $\tilde{H}_{phase,2,j}(u)$ are the pairwise phase transfer function from an asymmetric half-circle mask along the *j-th* axis direction. Fig. 2 (a) and (b) show comparison of different phase transfer functions, using 1-axis and 12-axis measurements, and the difference between 1-axis and 12-axis phase transfer functions is shown in Fig. 2(c)

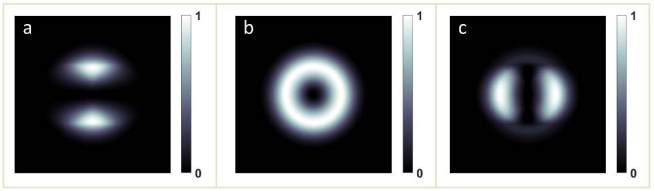

Fig. 2. Simulation results of phase transfer function ($\sum_i \left| \tilde{H}_{phase,i} \right|^2$) with an asymmetric half-circle pupil with different numbers of one and twelve measurements. (a) 1-axis qDPC phase transfer function, (b) 12-axis isotropic qDPC phase transfer function, and (c) difference of transfer function between (a) and (b).

III. U-Net Model for Quantitative Differential Phase Contrast Microscopy

A. U-Net Modeling

The U-net model has been often utilized in image transformation as an alternative method for inverse problems [25]–[27]. U-net has already been considered as one of the most useful architectures for biomedical imaging [28]. Compared to traditional CNN models, U-net is simple, effective, straightforward, less computationally expensive and faster with different scales of feature delivery. The skip connection of the U-net intends to provide the local information to the global information in the up-sampling process. Because of the symmetry of U-net, the network has a large number of feature maps in the up-sampling path, which allows it to transfer information. In addition, the U-net model has shown a better performance for phase contrast and DIC imaging in the previous work [29]. The purpose of the U-net in our work is to transform anisotropic quantitative phase images to isotropic quantitative phase images. The ground truth and input training images are obtained from isotropic and anisotropic phase reconstruction, respectively. The phase transfer function for ground truth images are isotropic whereas the phase transfer function for input images are anisotropic. In the training process, the U-net model is implicitly dependent on the phase transfer function shown by (4). A U-net model for obtaining isotropic qDPC images from only 1-axis qDPC image as shown in Fig. 1 (c) [29]. In addition, the use of multiple scales and skipping connections in the model help in retaining both low and high level information, for recovering fine features within cells.

As shown in Fig. 3, a U-net model consists of total 19 convolutional layers for feature extraction, with four maximum pooling (max-pool) layers, and four transpose convolutional layers to generate multi-scale features. The number of convolution layers in each scale was chosen from the empirical results following the previous works [28], [30]–[33]. The kernel size of each convolutional layer including transpose convolution is chosen as 3×3. The last layer for outputting results is a 1×1 convolutional layer. The number of filters used for feature extraction is shown at the upper part of each convolutional layer in Fig. 3. The rectified linear unit (ReLu) activation function is applied after each convolutional layer, except for the last layer which uses the tanh activation function. The max-pool layer takes the maximum value in the 2×2 area, and reduces the feature size by half.

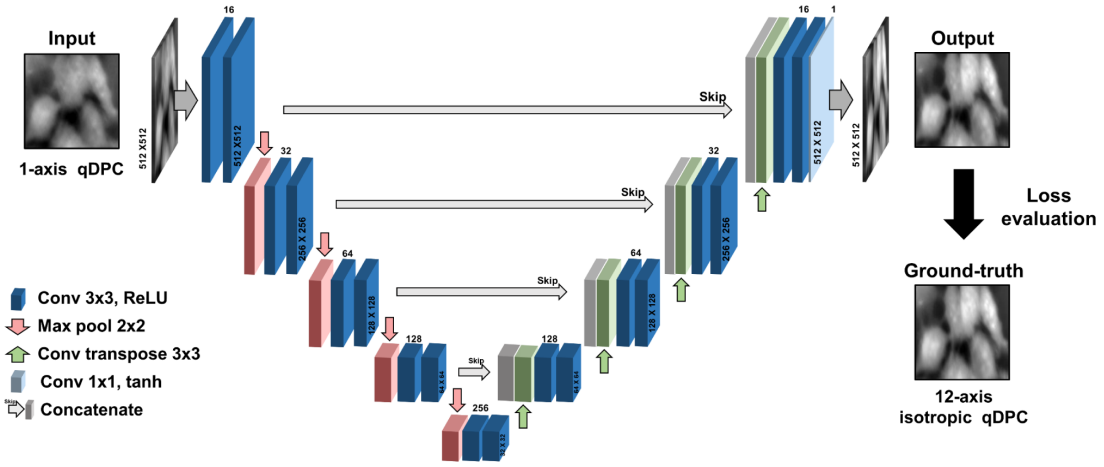

Fig. 3. The U-net architecture includes an encoding and decoding process with maximum pooling layer and a convolution transpose layer from left to right, respectively. Concatenate connections, between down-sampling and up-sampling procedure, are adapted for overall contextual information matching. The 1-axis qDPC images are used for input datasets, and 12-axis isotropic qDPC images are used for ground-truth datasets in the training process.

The stride in the transpose convolutional layer is equal to 2 with zero-padding to obtain consistent output feature size. The U-net model consists of multiple skipping connections to reuse features extracted during the encoding process. The encoded features are concatenated to the transpose convolutional layer, and provide low-level information while remaining high-level features [34].

In a traditional DNN model, the value $v_{x,y}^{l,k}$ at the *x, y-th* pixel position in *l-th* convolutional layer and *k-th* feature map is given by following [20]:

$$v_{x,y}^{l,k} = \sum_{pf} \sum_{p=-1}^{P-2} \sum_{q=-1}^{Q-2} w_{p+1,q+1}^{l,k} v_{x+p,y+p}^{l-1,pf}, \quad (5)$$

where $w_{p+1,q+1}^{l,k}$ is the weighting value in a convolutional kernel at *(p+1, q+1)* position for *l-th* convolutional layer, and *k-th* feature map. *P* and *Q* represent the size of convolutional kernels. $v_{x+p,y+q}^{l-1,pf}$ is the pixel value from *(l-1)-th* convolutional layer, and *pf* indicates the feature maps in the *(l-1)-th* layer. The weighting value $w_{p+1,q+1}^{l,k}$ determines the proportion of each pixel value that corresponds to the *l-th* layer. The *l-th* layer of convolution involves multiplication of weighting value and pixel value. In the training procedure, the weighting values are updated iteratively according to the loss function. The updated model can thus reduce the difference between the input image and the ground-truth image.

B. Model Learning Objective

The learning objective is minimized using the following equation:

$$U_{learn} = \underset{U_w, w \in W}{argmin}\, F(U_w), \quad (6)$$

where U_{learn} is the optimal U-net model. U_w denotes initial U-net model weights, which are updated by training procedure, and monitored by the loss function F. W is the set of all possible parameters [25]. The loss function F used is the mean square error (MSE) and given as,

$$F(U_w) = \frac{1}{BNM} \sum_{b=1}^{B} \sum_{n=1}^{N} \sum_{m=1}^{M} \left\| \begin{matrix} U_w\left(\phi_b^1(n,m)\right) \\ -\phi_b^{12}(n,m) \end{matrix} \right\|_2^2, \quad (7)$$

where $\phi_b^1 \in \Re^{N \times M}$ and $\phi_b^{12} \in \Re^{N \times M}$ represent patched phase images from 1-axis qDPC and 12-axis isotropic qDPC, respectively. *N* and *M* are the width and height of the single patched image, and *(n, m)* are the pixel coordinates in each phase image. *B* is the number of mini-batches, and in our case B is equal to 50.

C. Evaluation Metrics

The performance of our U-net model can be evaluated by the peak signal to noise ratio (PSNR) which assesses the quality of prediction and can be calculated as follows:

$$\text{PSNR} = 10\ log_{10}\left(\frac{Maxv^2}{MSE}\right), \quad (8)$$

where *Maxv* is the maximum value in the image data. The calculation of *MSE* is the same as that shown in (7) without mini-batch index. It shows that a higher PSNR value provides a higher image quality as compared to the ground-truth images. On the other hand, a small value of the PSNR implies high numerical differences between two images.

Structural similarity (SSIM) is alternative evaluation metric, regularly used for DL, and can be calculated as [35],

$$\text{SSIM}(im_1, im_2) = \left[lm(im_1, im_2)^{\varepsilon_1} \cdot ct(im_1, im_2)^{\varepsilon_2} \cdot st(im_1, im_2)^{\varepsilon_3}\right], \quad (9)$$

where im_1 and im_2 represent the two images to be compared. In our case, im_1 is the input or predicted images and im_2 is the ground-truth image. *lm, ct,* and *st* represent the luminance, contrast, and structure, respectively. $\varepsilon_1, \varepsilon_2,$ *and* ε_3 are the weighting factors for emphasizing specific parameters, which are set to units. Here, PSNR is utilized to access the quality of images, while SSIM is used to quantify the similarity of the predicted and ground-truth images.

Since the phase value is related to non-aqueous content in the biological cells, the cell dry mass can be used to analyze the progress of cell growth, the changes of the ingredients in the cells, and the changes of cell populations [36]. For the

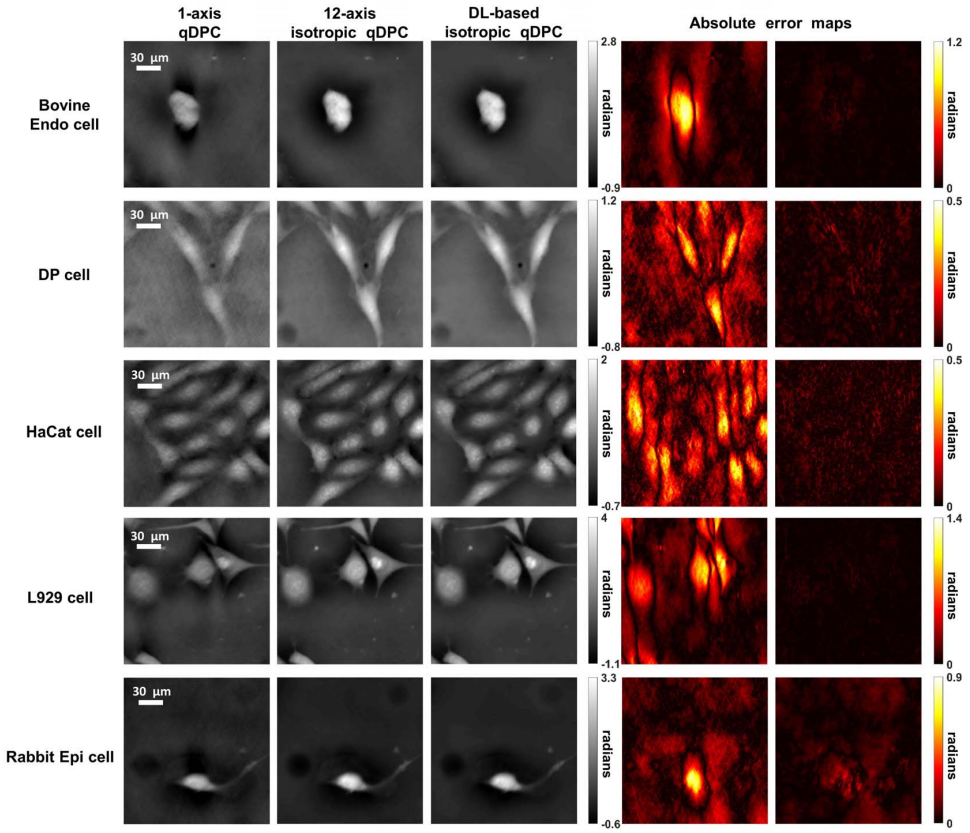

Fig. 4. Resultant images of different living cells, using 1-axis qDPC (input), 12-axis isotropic qDPC (ground truth), and DL-based qDPC (prediction). Images of HaCaT, Bovine Endo, L929, and DP cells were taken for training datasets, and Rabbit Epi cell images were taken for validation datasets. Absolute error maps on the left show the difference between 1-axis qDPC images and 12-axis isotropic qDPC images, and the error maps on the right shows the difference between DL-based qDPC images and 12-axis isotropic qDPC images.

1-axis qDPC and the DL-based isotropic qDPC, the cell dry mass is given by following [36]:

$$\text{Dry mass} = \frac{\lambda}{2\pi\alpha_{protein}} \iint_A \phi(n,m)\, dndm, \tag{10}$$

where $\alpha_{protein}$ is the refraction increment for proteins [37], and ϕ is the phase value at the position of (n, m). We chose region of interest in the images. Total quantitative phase of the ROI was evaluated by summation of phase values of all the pixels.

IV. Results

Phase images from seven different types of living cells were used for training, validation, and testing datasets for our model. The training datasets consisted of human keratinocyte cell line HaCaT, Bovine endothelial (Bovine Endo), L929, and DP cells. Images of Rabbit epithelial (Rabbit Epi) cells were used for validation datasets, while both 3T3 mouse fibroblasts cells and rat mesenchymal stem cells (RMSC) were used for testing. Twelve different microscope images of each cell type were obtained, and each microscope image contained 2048×1536 pixels. An out-of-memory error often occurs when training such huge microscope images. To solve this problem, the patch-based method was employed [38]–[40]. The full size image was then disassembled before the training process, and the patch size of 512×512 was adopted. We tested images with the size of 128×128, 256×256 and 512×512, and obtained the lowest loss value for 512×512 size images. The image was cropped from the top left position, and overlapped by half of the image area for each neighbor patch. The stride for image cropping was equal to 256 pixels. The same cropping method was repeated along both horizontal and vertical directions. The number of patches per microscope image was 35. For the training datasets, the number of patched images for each cell type was 420. During testing, the input of U-net was a patched-image, instead of the original image. Due to patch images, there will be small seam errors, which was reduced by overlapping the neighboring patched-images and reducing the number of patches. Four different cell types, including HaCaT, Bovine Endo, L929, and DP cells, were used for the training datasets; hence, total number of patched images was 1680. For the validation and testing datasets, the number of patched images were 420 and 840, respectively.

Our U-net model was trained by Adam optimization method [41]. The learning rate was set to be 1e-3. The beta1, beta2 were the optimization parameters for Adam, and their values were equal to 0.9 and 0.999, respectively. The number of epochs was set to 1000. In order to reduce the loss value, during the training process, the training datasets were utilized for updating the weights in each kernel, and we saved model's weights at the epoch with the lowest validation loss. After finishing the training step, the testing datasets were then used for evaluating the model performance. Our model was implemented with TensorFlow and processed on a NVIDIA Titan XP GPU.

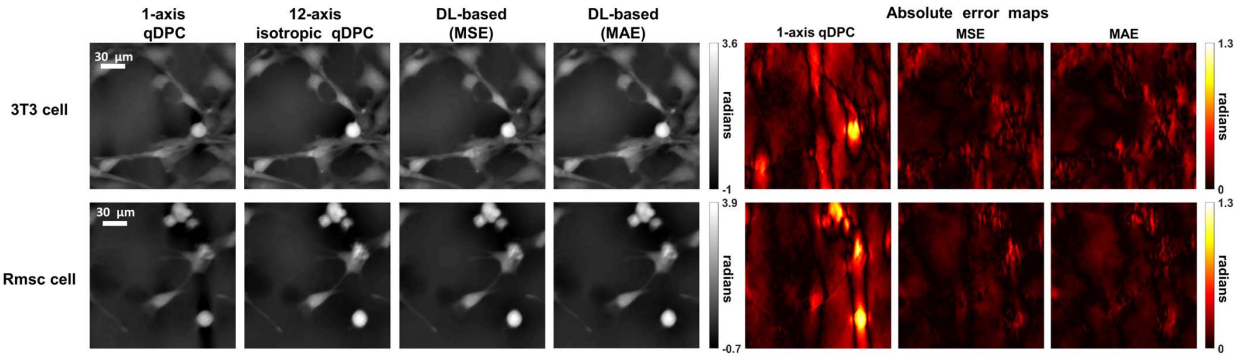

Fig. 5. Both images of 3T3, and Rmsc cells were taken for testing datasets. Resultant images of 3T3 and Rmsc living cells, using 1-axis qDPC (input), 12-axis isotropic qDPC (ground truth), and DL-based qDPC (prediction). The comparison of MSE and MAE loss functions is shown. The absolute error map on the left show the difference between 1-axis qDPC and 12-axis isotropic qDPC, and the error maps on the middle and right show the difference between DL-based qDPC with MSE and MAE losses and 12-axis isotropic qDPC images, respectively.

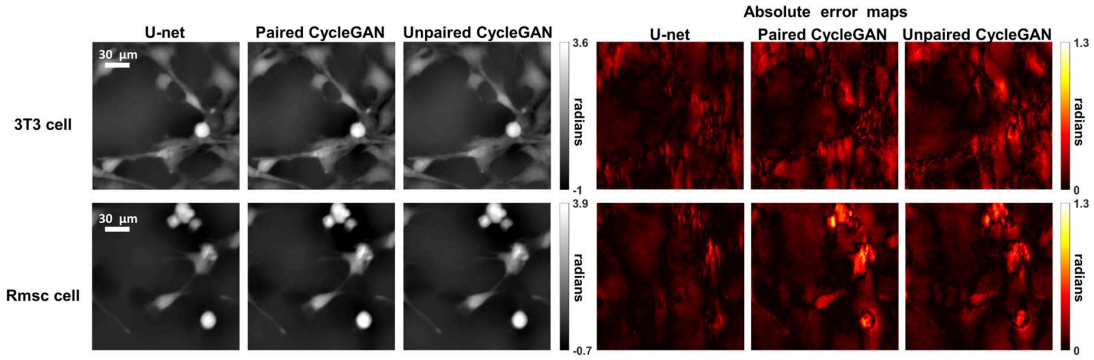

Fig. 6. Comparison of results from different deep learning models. The U-net, paired CycleGAN, and unpaired CycleGAN have been trained by the same datasets as displayed in Fig 5. The performance of three models are shown by absolute error maps.

A. Training Results

The loss for training and validation for our U-net model is plotted and the details can be found in supplemental information (1). Figure 4 shows a comparison of input (i.e. 1-axis qDPC images), ground truth (i.e. 12-axis isotropic qDPC images), and prediction (i.e. DL-based qDPC reconstructions). Four different types of living cells, including HaCaT, Bovine Endo, L929, and DP cells were used for training, and images of Rabbit Epi cells were used for validation. Each cell has different structures, we have trained the model for different combinations of seven distinct types of cells for training, validation and testing. Eventually, we showed the best result of utilizing four cell types for training, one cell type for validation, and two different cell types for final evaluation on our well-trained model. In Fig. 4, phase error maps on the left show the absolute difference between the input and the ground-truth images, while absolute phase error maps on the right show the difference between the predicted and the ground-truth images. Compared to the input images, the predicted images provide similar accuracy to the ground-truth. In addition, absolute error maps show that the trained U-net model recovers images at low spatial frequencies, as well as detailed fine features at high spatial frequencies in the input phase distribution. By a large number of training data images, the U-net model we applied in our study can extract the correct features during the encoding process in the model architecture. The features from the input images then can be converted to the 12-axis isotropic phase by the combination of multiple convolutions

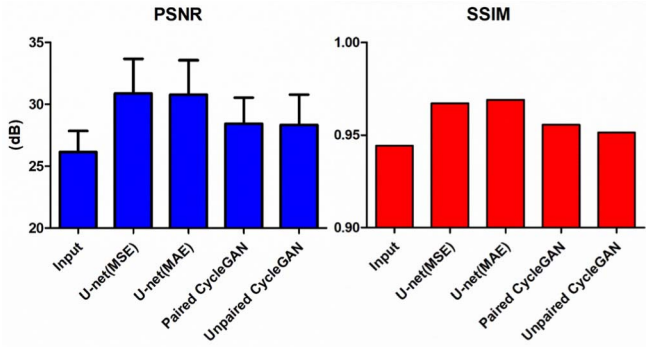

Fig. 7. Histogram of PSNR and SSIM from U-net and CycleGAN models.

and activation functions. The correspondence between the input data and the ground-truth data enables the realization of recovering the missing phase contents that are absent in the input datasets. The utilization of patch-wise approach provided more image features to be learned by model.

Apart from the training datasets, in Fig. 4 Rabbit Epi cells show prediction results for validation datasets. To select an optimal model without overfitting, we need to use data that is absent in the training process for model validation. Here, one type of cells is utilized to validate the model with 15% of available data. Here, Rabbit Epi cell are used for validation since they have diverse shapes and structures as compared to other cell samples. Similar to the results of training, the error values in most of the regions were reduced by U-net model prediction. The PSNR values in 1-axis and the DL-based isotropic qDPC are respective 16.95 ± 1.76 dB and

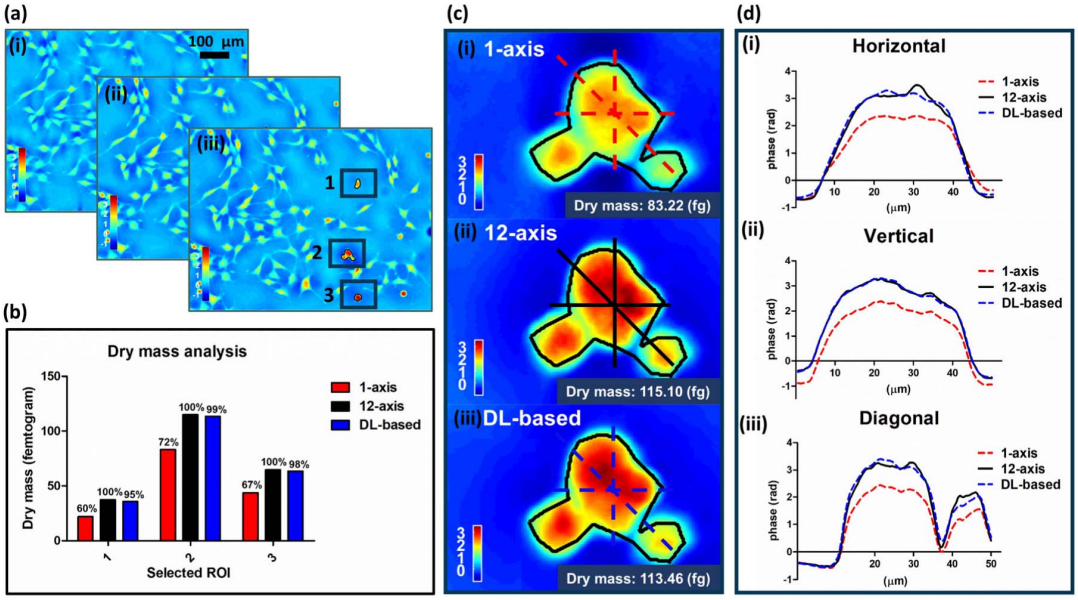

Fig. 8. (a) Phase distribution from (i) 1-axis qDPC image, (ii) 12-axis isotropic qDPC image, and (iii) DL-based isotropic qDPC image. Three regions of interest are high-lighted with black rectangles. (b) The dry mass corresponds to selected ROIs shown in (a, i-iii). (c) Zoomed-in images from (a) at ROI #2. (d) Cross-sections along horizontal, vertical, and diagonal directions in (c, i-iii). The corresponding profiles from input, prediction and ground-truth in dash red, blue lines and solid black line, respectively.

24.10 ± 1.65 dB for the validation datasets. The optimal model is generally biased towards the training datasets. Because of the exclusion of validation datasets in the training process, the hyper-parameters in the model were not adapted to the features appearing in the validation data.

B. Model Evaluation

To further evaluate the performance of the U-net model, additional unseen living cells of 3T3 fibroblast cells and RMSC cells were utilized for testing datasets. These two distinct cells have diversity in the cell density and shape. Fig. 5 shows phase error maps for MSE and MAE loss functions. The phase error maps on the left show absolute difference between the input and the ground-truth images. The phase error maps in the middle show absolute difference between the prediction of U-net, which is trained with MSE loss function, and ground-truth. The phase error maps on the right show absolute difference between the prediction of U-net, which is trained with MAE loss function, and ground-truth. Compared to the input images, the phase errors show similar accuracy for MSE and MAE predicted images.

PSNR is utilized to evaluate the performance of our U-net model, and is also evaluated for all predicted phase images from testing datasets. In Fig. 5, the PSNR values in 1-axis and the DL-based isotropic qDPC are respectively 26.15 ± 1.70 dB and 30.89 ± 2.78 dB, which are represented by the average and standard deviation. The calculated PSNR demonstrates that the image quality of our U-net model prediction on testing datasets is superior to the phase reconstruction from 1-axis qDPC. The results obtained from testing datasets show solid evidence that our trained U-net model has the capability to deal with different cell types with a variety of phase distributions.

In order to show the advantage of the U-net model, we compared it with an advantages model based on Cycle-Consistent Adversarial Networks (CycleGAN). CycleGAN [31], [42], [43] can be considered as the improved version of the original GAN network [44]. The results presented here can also be obtained using conditional GAN (cGAN) with some limitations such as hallucinations and unwanted information in the generated data [45]. The details of the CycleGAN architecture used in our work can be found in supplemental information (2). Our experimentally generated data have been used to compare the performance of both paired and unpaired CycleGAN approaches. Comparison of phase reconstructed results from different models is shown in Fig. 6. The U-net, paired CycleGAN, and unpaired CycleGAN have been trained by the same dataset as shown in Fig 5. PSNR and SSIM are used as performance metrics for the evaluation. The performance of the three models are measured by absolute error maps. The paired CycleGAN has lower errors than the unpaired. It is important to note that the phase error maps of the U-net model are smaller as compared with the CycleGAN models. Although the visual comparison of the phase images predicted from the three models look similar at the first glance.

However, the images generated by CycleGAN have more unwanted phase contrast information, which is absent in the ground-truth [31]. Two performance evaluation metrics for U-net and CycleGAN models are shown in Table I.

From the above Table, higher PSNR and SSIM values clearly show that the U-net model outperformed the CycleGAN models due to supervised learning approach. Above results show accuracy of U-net model is high as compared to other models which make it suitable for quantitative phase imaging for live cells. In addition, for better visualization, bar charts of PSNR and SSIM for U-net and CycleGAN models are shown in Fig. 7. The highest PSNR value is obtained from the prediction of U-net model with MSE loss. It denotes the small difference between the prediction and the ground-truth images which indicates

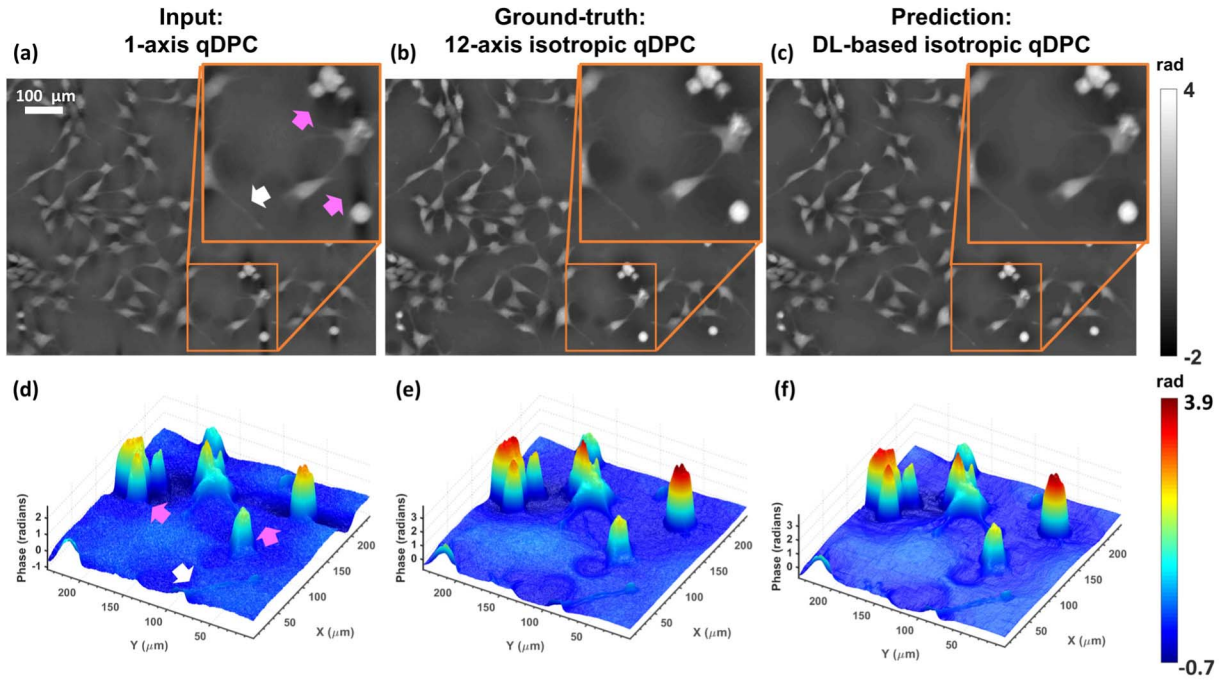

Fig. 9. Comparison of re-assembled images. (a) The input of the U-net model is the 1-axis qDPC. (b) Prediction from the U-net model. (c) The ground-truth of the U-net model from the 12-axis isotropic qDPC. (d-f) 3D profiles corresponding to each ROI (high-lighted in zoomed-in orange boxes in (a), (b), and (c)) for observation to emphasize differences between input and prediction. The apparent phase valleys in the images are highlighted by pink arrows. The trained model compensated the apparent phase valley along the vertical direction and enhanced the contrast of spindle-like structures marked by white arrow.

TABLE I
COMPARISON FOR PREDICTED PHASE IMAGES FROM DIFFERENT MODELS

	PSNR	SSIM
Input	26.15dB±1.70	0.9442
U-net with MSE	30.89dB±2.78	0.9671
U-net with MAE	30.80dB±2.76	0.9690
Paired CycleGAN	28.44dB±2.10	0.9556
Unpaired CycleGAN	28.34dB±2.46	0.9514

the better performance of U-net model compared to others. The highest SSIM value is obtained from U-net model with MAE loss which corresponds to high image quality.

For quantitative evaluation of phase images, we calculated the dry mass in three different ROIs as shown in Fig. 8 (a). In Fig. 8 (b), the dry mass was calculated by (10), and we used 12-axis isotropic phase values in each ROI as 100% dry mass recovery and compared to 1-axis phase values and DL-based phase values. The dry mass recovery of the input image does not exceed 73% compared to the ground-truth. However, in the prediction, the dry mass values calculated in each ROI are all well recovered above 94% compared to the ground-truth. The cross-sections along in different directions are shown in Fig. 8 (c). Each colored dash line and solid line corresponds to the input, prediction and ground-truth in Fig. 8 (d), and the cross-sections along three different directions show that the phase values obtained from the model prediction are well-matched with the ground-truth.

Figure 9 shows the comparison of re-assembled images. The histogram corresponding to Fig 9 can be found in supplemental information (3). The apparent anisotropic phase distribution in Fig. 9 (a) indicates that there is missing information in the phase transfer function of the 1-axis qDPC. The predicted phase distribution in Fig. 9 (b) provides a similar phase distribution as from the 12-axis isotropic qDPC results. Fig. 9 (d-f) shows 3D visualization of phase distribution corresponding to Fig. 9 (a-c). To enhance observation, the apparent phase valleys in the images are highlighted by pink arrows in Fig. 9 (a) and (d). The model compensated the errors and recovered phase is shown in Fig. 9 (e). In addition, the spindle-like structures extended from the cells are high-lighted by white arrows.

V. CONCLUSION

In conclusion, we implemented the U-net model with qDPC microscopy to generate high-resolution isotropic phase images of living cells from the least number of experimental measurements. To train, validate, and test our U-net model, all the datasets were obtained from the experiment by our qDPC microscopy. The patch-wise approach helps to acquire sufficient training data and prevent hardware limitation. After training, the U-net model significantly enhances acquisition speed, which may help in monitoring rapid changes in living cells. The correspondence between the input data and the ground-truth data enables the recovery of missing detailed phase features. The trained model accurately predicts quantitative phase images with fine features in cells. With DL-based methods we can achieve over ten times faster speed for phase reconstruction. We also compared our U-net model for qDPC with CycleGAN. Due to the supervised learning approach, and higher values of PSNR and SSIM, U-net model offers advantages in qDPC microscopy. The images generated by the CycleGAN models have more unwanted phase contrast information. The U-net model provides high accuracy for reconstructed quantitative phase images as compared to CycleGAN. Our model can perform dry mass analysis. The

accuracy of dry mass recovery is above 94% in all results. The recovery of accurate dry mass from predicted phase images can help in monitoring the cell growth progress. Additionally, the flexibility of our DL-based method can be applied to other pupil patterns to further enhance the performance of qDPC imaging. Our DL-based quantitative phase measurement approach has great potentials for important applications, including cell morphology, cell dynamics, cell tomography, cell pathophysiology, and neuroscience.

Acknowledgment

The authors would like to thank Pan-Chyr Yang and Huei-Wen Chen for valuable discussions, and Xuan-Ling Hsu for sample preparation.

References

[1] J. W. Goodman, *Introduction to Fourier Optics*. Greenwood Village, CO, USA: Roberts and Company Publishers, 2005.
[2] S. Skylaki, O. Hilsenbeck, and T. Schroeder, "Challenges in long-term imaging and quantification of single-cell dynamics," *Nature Biotechnol.*, vol. 34, no. 11, pp. 1137–1144, Nov. 2016.
[3] J. Icha, M. Weber, J. C. Waters, and C. Norden, "Phototoxicity in live fluorescence microscopy, and how to avoid it," *BioEssays*, vol. 39, no. 8, Aug. 2017, Art. no. 1700003.
[4] F. Zernike, "Phase contrast, a new method for the microscopic observation of transparent objects," *Physica*, vol. 9, no. 7, pp. 686–698, Jul. 1942.
[5] Y. Park, C. Depeursinge, and G. Popescu, "Quantitative phase imaging in biomedicine," *Nature Photon.*, vol. 12, no. 10, pp. 578–589, Oct. 2018.
[6] H. Majeed *et al.*, "Quantitative phase imaging for medical diagnosis," *J. Biophotonics*, vol. 10, no. 2, pp. 177–205, Feb. 2017.
[7] D. Gabor, "A new microscopic principle," Nature Publishing Group, London, U.K., Tech. Rep., 1948, doi: 10.1038/161777a0.
[8] G. Popescu, *Quantitative Phase Imaging of Cells and Tissues*. New York, NY, USA: McGraw-Hill, 2011.
[9] B. Kemper *et al.*, "Towards 3D modelling and imaging of infection scenarios at the single cell level using holographic optical tweezers and digital holographic microscopy," *J. Biophotonics*, vol. 6, no. 3, pp. 260–266, Mar. 2013.
[10] H. H. Barrett, and K. J. Myers, *Foundations of Image Science*. Hoboken, NJ, USA: Wiley, 2013.
[11] Y.-Z. Lin, K.-Y. Huang, and Y. Luo, "Quantitative differential phase contrast imaging at high resolution with radially asymmetric illumination," *Opt. Lett.*, vol. 43, no. 12, pp. 2973–2976, 2018.
[12] H.-H. Chen, Y.-Z. Lin, and Y. Luo, "Isotropic differential phase contrast microscopy for quantitative phase bio-imaging," *J. Biophotonics*, vol. 11, no. 8, Aug. 2018, Art. no. e201700364.
[13] Y.-H. Chuang *et al.*, "Multi-wavelength quantitative differential phase contrast imaging by radially asymmetric illumination," *Opt. Lett.*, vol. 44, no. 18, pp. 4542–4545, 2019.
[14] A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica*, vol. 4, no. 9, pp. 1117–1125, 2017.
[15] G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica*, vol. 6, no. 8, pp. 921–943, 2019.
[16] S. Li, M. Deng, J. Lee, A. Sinha, and G. Barbastathis, "Imaging through glass diffusers using densely connected convolutional networks," *Optica*, vol. 5, no. 7, pp. 803–813, 2018.
[17] H. Wang *et al.*, "Deep learning enables cross-modality super-resolution in fluorescence microscopy," *Nature Methods*, vol. 16, no. 1, pp. 103–110, Jan. 2019.
[18] N. Borhani, A. J. Bower, S. A. Boppart, and D. Psaltis, "Digital staining through the application of deep neural networks to multi-modal multi-photon microscopy," *Biomed. Opt. Exp.*, vol. 10, no. 3, pp. 1339–1350, Mar. 2019.
[19] X. Dong *et al.*, "Deep learning-based attenuation correction in the absence of structural information for whole-body positron emission tomography imaging," *Phys. Med. Biol.*, vol. 65, no. 5, Mar. 2020, Art. no. 055011.
[20] Y. Rivenson, Y. Zhang, H. Günaydin, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light, Sci. Appl.*, vol. 7, p. 17141, Feb. 2018.
[21] F. Wang *et al.*, "Phase imaging with an untrained neural network," *Light, Sci. Appl.*, vol. 9, no. 1, pp. 1–7, Dec. 2020.
[22] Y. Jo *et al.*, "Quantitative phase imaging and artificial intelligence: A review," *IEEE J. Sel. Topics Quantum Electron.*, vol. 25, no. 1, pp. 1–14, Feb. 2019.
[23] M. R. Kellman, E. Bostan, N. A. Repina, and L. Waller, "Physics-based learned design: Optimized coded-illumination for quantitative phase imaging," *IEEE Trans. Comput. Imag.*, vol. 5, no. 3, pp. 344–353, Sep. 2019.
[24] L. Tian and L. Waller, "Quantitative differential phase contrast imaging in an LED array microscope," *Opt. Exp.*, vol. 23, no. 9, pp. 11394–11403, 2015.
[25] M. T. McCann, K. H. Jin, and M. Unser, "Convolutional neural networks for inverse problems in imaging: A review," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 85–95, Nov. 2017.
[26] C.-C. Liu and H.-M. Huang, "Partial-ring PET image restoration using a deep learning based method," *Phys. Med. Biol.*, vol. 64, no. 22, Nov. 2019, Art. no. 225014.
[27] M. Yamaguchi *et al.*, "Dose image prediction for range and width verifications from carbon ion-induced secondary electron bremsstrahlung X-rays using deep learning workflow," *Med. Phys.*, vol. 47, no. 8, pp. 3520–3532, Aug. 2020.
[28] T. Chang *et al.*, "Calibration-free quantitative phase imaging using data-driven aberration modeling," *Opt. Exp.*, vol. 28, no. 23, pp. 34835–34847, 2020.
[29] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention*. Cham, Switzerland: Springer, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.
[30] L. Jiao, L. Huo, C. Hu, and P. Tang, "Refined UNet: UNet-based refinement network for cloud and shadow precise segmentation," *Remote Sens.*, vol. 12, no. 12, p. 2001, Jun. 2020.
[31] Y. Li, W. Li, J. Xiong, J. Xia, and Y. Xie, "Comparison of supervised and unsupervised deep learning methods for medical image synthesis between computed tomography and magnetic resonance images," *Bio-Med Res. Int.*, vol. 2020, pp. 1–9, Nov. 2020.
[32] Z. Zhang *et al.*, "Holo-UNet: Hologram-to-hologram neural network restoration for high fidelity low light quantitative phase imaging of live cells," *Biomed. Opt. Exp.*, vol. 11, no. 10, pp. 5478–5487, 2020.
[33] X. Zhao *et al.*, "Use of unmanned aerial vehicle imagery and deep learning UNet to extract rice lodging," *Sensors*, vol. 19, no. 18, p. 3859, Sep. 2019.
[34] Z. Zhang, Q. Liu, and Y. Wang, "Road extraction by deep residual U-Net," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 749–753, May 2018.
[35] A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *Proc. 20th Int. Conf. Pattern Recognit.*, Aug. 2010, pp. 2366–2369.
[36] C. Hu and G. Popescu, "Quantitative phase imaging (QPI) in neuroscience," *IEEE J. Sel. Topics Quantum Electron.*, vol. 25, no. 1, pp. 1–9, Jan. 2019.
[37] P. Y. Liu *et al.*, "Cell refractive index for cell biology and disease diagnosis: Past, present and future," *Lab Chip*, vol. 16, no. 4, pp. 634–644, 2016.
[38] M. Havaei *et al.*, "Brain tumor segmentation with deep neural networks," *Med. Image Anal.*, vol. 35, pp. 18–31, Jan. 2017.
[39] A. de Brebisson and G. Montana, "Deep neural networks for anatomical brain segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) Workshops*, 2015, pp. 20–28.
[40] Q. Sun *et al.*, "Patch-based deep convolutional neural network for corneal ulcer area segmentation," in *Fetal, Infant and Ophthalmic Medical Image Analysis*. Cham, Switzerland: Springer, 2017, pp. 101–108.
[41] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980
[42] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2223–2232.
[43] Y. Zhang *et al.*, "PhaseGAN: A deep-learning phase-retrieval approach for unpaired datasets," 2020, *arXiv:2011.08660*. [Online]. Available: http://arxiv.org/abs/2011.08660
[44] I. J. Goodfellow *et al.*, "Generative adversarial networks," 2014, *arXiv:1406.2661*. [Online]. Available: http://arxiv.org/abs/1406.2661
[45] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1125–1134.