



# Holo-UNet: hologram-to-hologram neural network restoration for high fidelity low light quantitative phase imaging of live cells

ZHIDUO ZHANG,<sup>1</sup> YUJIE ZHENG,<sup>1</sup> TIENAN XU,<sup>1</sup> AVINASH UPADHYA,<sup>1</sup>  YEAN JIN LIM,<sup>1</sup> ALEXANDER MATHEWS,<sup>2</sup> LEXING XIE,<sup>2</sup> AND WOEI MING LEE<sup>1,3,\*</sup> 

<sup>1</sup>Research School of Electrical, Energy and Materials Engineering, College of Engineering and Computer Science, The Australian National University, 31 North Road, Canberra, ACT 2601, Australia

<sup>2</sup>Research School of Computer Science, College of Engineering and Computer Science, The Australian National University, 108 North Road, Canberra, ACT 2601, Australia

<sup>3</sup>The ARC Centre of Excellence in Advanced Molecular Imaging, The Australian National University, Canberra, ACT 2601, Australia

\*[steve.lee@anu.edu.au](mailto:steve.lee@anu.edu.au)

**Abstract:** Intensity shot noise in digital holograms distorts the quality of the phase images after phase retrieval, limiting the usefulness of quantitative phase microscopy (QPM) systems in long term live cell imaging. In this paper, we devise a hologram-to-hologram neural network, Holo-UNet, that restores high quality digital holograms under high shot noise conditions (sub-mW/cm<sup>2</sup> intensities) at high acquisition rates (sub-milliseconds). In comparison to current phase recovery methods, Holo-UNet denoises the recorded hologram, and so prevents shot noise from propagating through the phase retrieval step that in turn adversely affects phase and intensity images. Holo-UNet was tested on 2 independent QPM systems without any adjustment to the hardware setting. In both cases, Holo-UNet outperformed existing phase recovery and block-matching techniques by ~ 1.8 folds in phase fidelity as measured by SSIM. Holo-UNet is immediately applicable to a wide range of other high-speed interferometric phase imaging techniques. The network paves the way towards the expansion of high-speed low light QPM biological imaging with minimal dependence on hardware constraints.

© 2020 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

## 1. Introduction

Quantitative Phase Microscopy (QPM) is a special class of light microscopy [1,2] that uses low levels of illuminating light to image cells in 3D; capturing biological activities in their natural environment with sub-cellular resolutions. An ideal low light QPM system for long term imaging (~weeks) needs to achieve high fidelity phase retrieval with minimal dependence on hardware constraints. Coherence and laser power from different illuminating sources and detectors changes the overall shot noise, pixel sampling and dynamic range of QPM. On top of that, living biological sample are constantly changing, creating time-varying intensity and phase changes. Together, the illumination sources and living biological sample can both increase the level of noise in the hologram resulting in poor interference fringe visibility and in turn phase fidelity [3].

The imaging performance (speed, illuminating power) of single shot off-axis digital hologram is ultimately shot-noise limited. The contrast and speed of QPM imaging systems are adversely affected by low illumination intensities due to photon shot noise and path length mismatch (low coherence sources). Photon shot noise follows a Poisson distribution, and is directly related to the number of photons impinging on detectors. Path length mismatch in QPM, using low coherence sources, are directly related to change of thickness in the sample. For QPM imaging of living biological sample, there is need to balance the hardware settings to reduce photon shot

noise and path length matching to obtain high quality holograms. These hardware adjustments include optimizing photodetector gain/exposure, path length matching and illumination power [3]. However, individual hardware optimisation for different digital imaging detectors and biological samples is not ideal for long term QPM imaging. Overall, specialised user knowledge and calibration of laser power, path length matching and detector create barriers to widespread adoption of QPM.

Although machine learning (ML) methods are making rapid strides in many topical areas of phase imaging such as classification, intensity-only and phase-only recovery, removal of coherent noise for optical diffraction tomography [4] and even speckle noise reduction [5], they do not address the critical issue of direct recovery of holograms. Previous uses of neural networks (NN) for phase imaging in QPMs are either end-to-end networks for direct phase retrieval [6–9] or for denoising phase images after phase retrieval [10–12]. The application of end to end networks for low light phase retrieval leads to low fidelity of the phase image [10]. On the other hand, the effect of shot noise in digital holograms makes post-phase retrieval and phase denoising more difficult. Each pixel of a digital hologram in the spatial domain can contain intensity shot noise that influences multiple pixels in the spatial frequency domain after Fourier transformation and can cause errors in the phase unwrapping. The propagation of noise in recorded hologram would affect the fidelity of phase after phase retrieval because it is much challenging to distinguish shot noise from object in phase recovered images. Hence, we argue that it is necessary to perform denoising on the raw hologram before phase retrieval. To our knowledge, no existing ML phase microscopy tools have been shown to directly recover single shot off-axis holograms that is close to the shot noise limit.

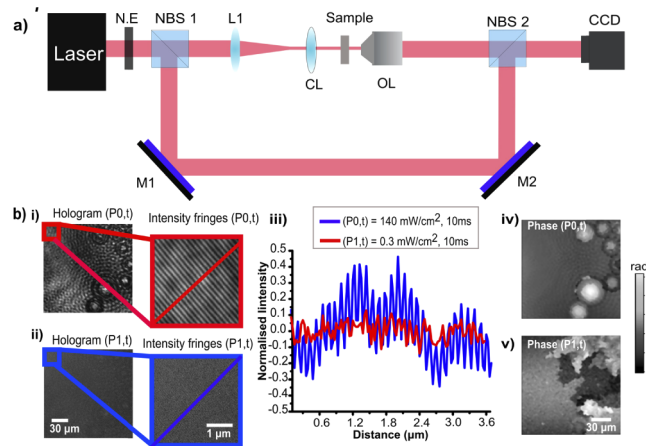
## 2. Holo-UNet

In this paper, we demonstrate for the first time a new ML hologram-to-hologram recovery tool, Holo-UNet, that operates at shot noise affected, sub-millisecond acquisition rates. In comparison to existing phase recovery neural network (PRNN) methods [11] that are applied onto the retrieved phase image, Holo-UNet is applied before the phase retrieval process. Holo-UNet can be broadly applicable to the problem of adverse effects in fringe visibility especially for path length mismatching in biological samples with varying thickness. Holo-UNet can dramatically increase acquisition rate of any high-speed quantitative phase microscopy by recovering holograms with short illumination times, and illustrates hologram denoising is a critical issue to address shot noise in the raw hologram and so cannot be ignored.

### 2.1. Physical setup

We validate Holo-UNet with 2 different off-axis QPMs each with a different sample type, illumination laser, and imaging objective lens. For the 6  $\mu\text{m}$  diameter microsphere samples, an off-axis QPM with laser source of wavelength 632.8 nm (JDS Uniphase 1144p-3581) was used. For imaging fibroblast cells (L929) a second QPM with laser source of wavelength 514 nm (OBIS coherent 1178771) was used. Both QPMs were based on a standard Mach-Zender interferometer as seen in Fig. 1 i) where the lasers were first split via a non-polarising beam-splitter into 2 independent optical paths before they were combined at the imaging plane (digital camera) using silver mirrors (M1 and M2). The cameras used for the QPM with 632.8 nm wavelength laser is a sCMOS (Thorlabs, CS2100M-USB) and the 514 nm wavelength QPM uses a CMOS (BFS-U3-32S4M-C). Each QPM uses a different imaging objective lens to achieve a suitable working distance, field of view and imaging resolution for imaging the sample: 6  $\mu\text{m}$  microspheres or  $\sim 10\text{-}20$   $\mu\text{m}$  fibroblast cells. Figure 1(a) the imaging objectives lenses (OL) used in the experiment were placed in the sample arm, collecting light that had passed through the sample of interest after the condenser lens (CL). For microsphere imaging, an objective lens with 40 $\times$ , magnification with numerical aperture (N.A) 0.70 was used (0.03  $\mu\text{m}/\text{pixel}$ ), whereas

a 20 $\times$ , NA 0.45, objective lens was used for fibroblast cell imaging (0.26  $\mu\text{m}/\text{pixel}$ ). To create different levels of low light illumination, we used combinations of neutral density (ND) filters (NE series Thorlabs Inc) to reduce light density from 140  $\text{mW}/\text{cm}^2$  (no filter) down to around 5, 0.6, 0.3  $\text{mW}/\text{cm}^2$ . A digital optical power meter (Thorlabs PM100d) was used to calibrate the laser power and another CCD camera (Thorlabs DCC1645) was used to capture the illumination area at the sample plane to calculate the illumination area on the sample. The exposure time of the imaging camera, in Fig. 1(a) is set at 10  $\text{ms}$  exposure time for video rate acquisition of 6  $\mu\text{m}$  microspheres suspended in water. For fibroblast cells imaging, the exposure time is reduced to  $\sim 200 \mu\text{s}$  with light density of  $\sim 5 \text{ mW}/\text{cm}^2$ . This is equivalent to around 300 photons/pixel. These imaging settings are held consistent for each imaging session. Figure 1(b) i) and ii) show the difference in the intensity fringes' contrast at two power levels, 140  $\text{mW}/\text{cm}^2$  and 5  $\text{mW}/\text{cm}^2$  at 10  $\text{ms}$  acquisition rate respectively. The comparison of the fringe visibility and phase retrieved can be seen in the normalised plot in Fig. 1(b) iii), b) iv) and b) v). At high shot noise (low light), the lower fringe visibility significantly reduced the fidelity of the retrieved phase image.

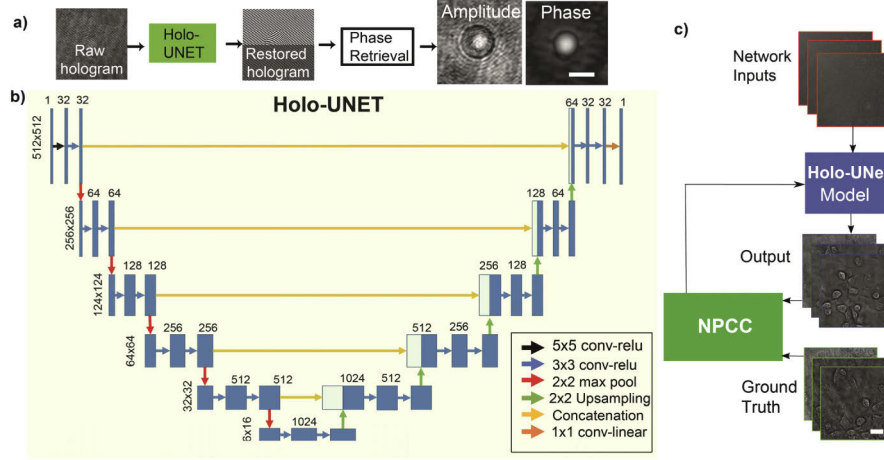


**Fig. 1. Off-axis QPM and low fidelity of phase retrieval under low light/high shot noise condition** (a) A schematic of the off-axis QPM setup used in the study. b) i) and ii) view of 6  $\mu\text{m}$  microspheres under linear intensity fringes captured with 140  $\text{mW}/\text{cm}^2$ , 0.3  $\text{mW}/\text{cm}^2$ , respectively. iii) cross sectional comparison of the visibility of the fringes. iv) and v) are the phase retrieved at the two power levels (P0 and P1). Red and blue box indicate the region in the hologram in b)i) and ii). Scale bar 30  $\mu\text{m}$ , 1  $\mu\text{m}$ .

## 2.2. Model architecture

To tackle the problem of hologram-to-hologram denoising at high shot noise, we propose a neural network inspired by the U-Net architecture [14], a form of fully convolutional network [15]. U-Net has successfully demonstrated accurate image segmentation of biological cells as well as image to image translation [16]. More recently, U-Net has been used in a number of optical imaging applications, for example low light fluorescence imaging [17], phase retrieval [6] and low light scenes [10]. U-Net utilises convolution blocks [18] to implement shift-invariant filtering for different parts of the image. By doing so, it enables parameter-sharing and reduces the number of parameters compared with fully connected layers. Through an iterative process of convolution and max-pooling, an input image is distilled into a sequence of information rich but successively lower resolution representations as illustrated in Fig. 2(b) from  $512 \times 512 \times 1$  down to  $16 \times 16 \times 1024$  blocks. At each step, the size of each convolution kernel's effective receptive field is increased. Further convolutional layers and upsampling are applied to return to

the input size. High resolution representations are combined with upsampled representations using concatenation to provide both fine- and coarse-grained information.



**Fig. 2.** Holo-UNet based on modified U-Net and learning process (a) Illustration of process converting hologram to phase. A hologram is passed through Holo-UNet, resulting in a ‘clean’ hologram. A common phase retrieval process [13] is used to then retrieve the phase and intensity (Scale bar 6  $\mu\text{m}$ ). (b) Modified U-Net architecture diagram. Data flows from left to right. Numbers on left vertical axis denote the kernel input size. Each blue box shows the resultant feature map, with kernel size above the box. (c) shows the training process, where low power (using ND filter) holograms are passed through the model. The result is compared with the ground truth (hologram without any ND filter) using NPCC and used to update the weights of the network arrows in (b). Scale bar 15  $\mu\text{m}$ .

The fundamental purpose of Holo-UNet is to predict intensity holograms that is distinctly different from simple low light intensity-only or phase-only images. It differs from existing ML methods that focus on intensity-only (CARE) [19] or phase-only recovery (PRNN) [11] because these technique (PRNN or CARE) fail to consider fringe improvement in shot noise limited holograms. Unlike traditional intensity images, off-axis holograms contain distinctive intensity fringes that visually resemble parallel stripes. The phase shifts are encoded onto small intensity variations along the primary carrier fringe as light passes through the sample experience a phase delay because of a change in refractive index. In this case, our network is modified to learn the changes along parallel intensity fringes upon encountering a phase-modifying object in the sample. The network is tasked with removing superfluous intensity changes (shot noise), and improving fringe visibility, thus improving the retrieved phase quality. To enforce fringe visibility in Holo-UNet, we employed a novel objective function using the Negative Pearson Correlation Coefficient (NPCC) in both the spatial and Fourier domains [16]. NPCC is the negative version of the standard Pearson’s Correlation Coefficient (PCC). Alongside the loss in the spatial domain, it is also imposed on the Fourier domain to optimise the three distinct peaks arising from the frequency spacing of the carrier intensity fringes in the frequency domain called the +1, 0 and -1 orders. Since phase is retrieved using the standard angular spectrum and adaptive Fourier filter [13], it is therefore logical to optimise for those peaks to increase the fringe visibility in the spatial domain. For given images  $X$ , and  $Y$ , NPCC is defined as:

$$NPCC(X, Y) = -\frac{cov(X, Y)}{\sigma_X \sigma_Y} \quad (1)$$

$$Loss(X, Y) = \frac{1}{2} * NPCC(X, Y) + \frac{1}{2} * NPCC[|FFT(X)|, |FFT(Y)|] \quad (2)$$

where  $\text{cov}(X, Y)$  is the covariance between images  $X$  and  $Y$ , and  $\sigma_X$ ,  $\sigma_Y$  are the standard deviations and  $\text{FFT}(X)$  and  $\text{FFT}(Y)$  are the Fast Fourier Transform (FFT) of the spatial domain. We also replace the one-hot output encoding step used by the original U-Net with a single linear final activation. This changes the output of the network from a stack of binary masks into a single real valued image. Furthermore, in holography, it is not only necessary to capture features of the sample but also finer details of parallel intensity fringes across the image which defines the final retrieved phase image. Hence additional zero padding was included to ensure input size matched output size and a deeper network consisting of 6 encoding layers is utilised to observe a larger receptive field. The network is applied before the phase retrieval process as seen in Fig. 2(a). The combined modification of Holo-UNet, shown in Fig. 2(b), is used to capture finer details of an off-axis hologram. In Fig. 5, we show the NPCC without considering the FFT has 10% lower performance average.

Shot noise has strongly distorted the fringes of the hologram. Hence, information lost cannot be directly retrieved without supervised learning. To train the network, randomly selected holograms of microspheres and fibroblast cells with pixel resolution of  $512 \times 512$  were cropped from a full frame of  $1920 \times 1080$ . In order to capture pairwise datasets of both low-power and high-power images of the same location. Image registration is necessary, and fringe vibrations need to be compensated for. Currently, this is achieved by recording 100 frames for each large training image, and manually matching of fringe positions is then used to achieve image registration. The microsphere experiment used a sequence of 30 different full frames at low light intensity, along with their corresponding holograms without any neutral density filter to train the network. From the 30 full frame images, randomly cropped images were used to generate approximately 800 images for training the network. To validate the network, we used four previously unseen full frame images from two different low light densities to create 120 cropped images that were split evenly between the validation and test set. For the fibroblast dataset the same method was used to generate 2100 training, 500 validation and 500 test images from 31 full frame images at 100 images per frame. Figure 2(c) shows an example of network training along with the use of NPCC to minimise the loss function between the output and ground truth (hologram without ND filter). The ground truth has a higher power at the sample due to the lack of ND filter and represents the best possible image from a standard DHM with minimum shot noise. We also compared the results to a common image denoising method used in holography called BM3D [20]. In all cases, phase is retrieved using the standard angular spectrum and adaptive Fourier filter [13]. The phase from BM3D and Holo-UNet was compared using Structural Similarity Index Metric (SSIM). SSIM and PCC are used to rank fidelity of the recovered phase and amplitude images respectively [11,21]. The model used in the experiments uses the same settings. Notably, the learning rate is controlled by the Adadelata [22] optimizer with a starting learning rate of  $1e-3$ . Due to GPU constraints, a batch size of 4 is used. Each model is run for 20 epochs, however, during testing the weights of the network from the epoch with the best validation results is used.

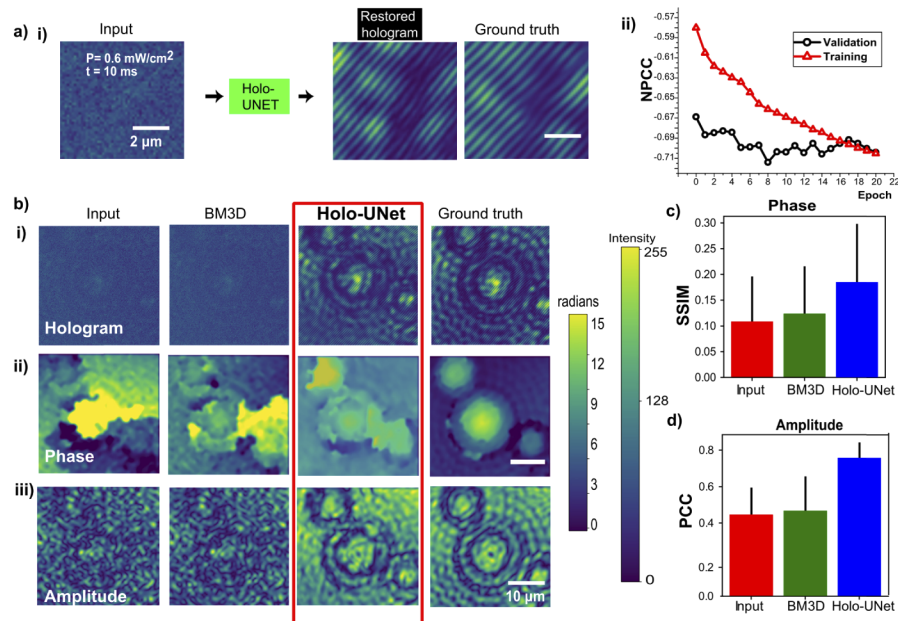
### 3. Results and discussion

#### 3.1. $6 \mu\text{m}$ diameter microspheres

All state-of-the-art ML phase imaging (e.g. PRNN) are focused upon denoising either phase-only or intensity-only images. We maintain fairness by using the PRNN model trained using the dataset's retrieved phase/intensity. Figure 3(a) i) shows the input of hologram taken at low light ( $0.6 \text{ mW/cm}^2$ ) at 10 millisecond acquisition time. Using Holo-UNet, we demonstrate the restoration of the off-axis interference fringes that is close to the ground truth hologram which is recorded at normal light levels, well above shot noise. Figure 3(a) ii) shows mean NPCC loss after supervised training data (red triangle) and the validation dataset (blue circle). Figure 3(b) i) shows a representative input hologram of  $6 \mu\text{m}$  diameter microspheres suspended in water. Figure 3(a) ii) shows mean NPCC loss after supervised training data (black square) and the

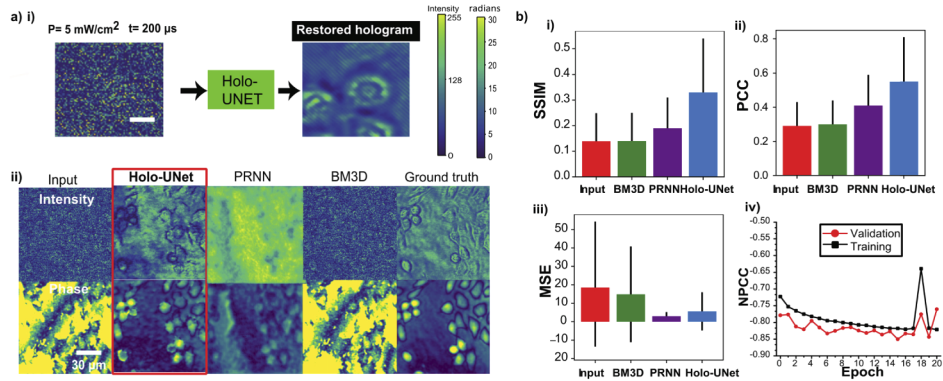


validation dataset (red circle) tracked over each epoch. For testing, the epoch that performed best on the validation dataset was chosen. The corresponding retrieved phase and amplitude of the holograms are shown in Fig. 3(b) ii) and iii) using Holo-UNet and BM3D, respectively. Figure 3(c) and (d) shows the averaged SSIM/PCC for the test set, which is used to compare the fidelity of the retrieved phase and amplitude. From the results in Fig. 3(a) i), the parallel intensity fringes in the recorded hologram exhibits low visibility that resulted in poor retrieval of phase and amplitude as compared to the ground truth. Upon visual inspection and comparing with BM3D, the recovered phase and amplitude using Holo-UNet exhibited improved image quality in terms of fringe visibility, recovered phase and amplitude of the microspheres. Based on the test dataset, we calculated the average SSIM of the retrieved phase increased almost 2 fold from 0.1 to  $\sim 0.18$  as shown in Fig. 3(b). We observed a near doubling of the PCC value from  $\sim 0.4$  to  $\sim 0.8$  in Fig. 3(c). While the network improved both phase and amplitude of the final restored hologram, the restoration performance displayed large variation with a standard deviation ( $\pm 0.09$ ). This variation can be attributed to influence of variation in fringe quality across the field of view under low-light illumination as a result of non-planar wave or amplitude variation. The increase in SD in the retrieved phase of our system is a result of its inability to improve upon some heavily noisy results. In addition, we also applied Holo-UNet without incorporating the Fourier transform loss, which resulted in the restored hologram that has phase image that is 10% lower in SSIM scoring (data not shown). To clarify, training was done separately for different samples and microscope to ensure validation of the method.



**Fig. 3. Restoration of hologram using Holo-UNet and quantitative comparison of phase and amplitude.** a) i) shows hologram imaged under  $0.6 \text{ mW/cm}^2$  and 10 millisecond that is then restored by Holo-UNet and respectively ground truth ii) Mean NPCC loss after each epoch of training during the bead experiment. b) shows hologram of  $6 \mu\text{m}$  diameter microspheres under  $0.6 \text{ mW/cm}^2$  using  $633 \text{ nm}$  laser source and performance of standard denoising (BM3D, Holo-UNet and ground truth) and Holo-UNet method against the ground based for the (i) hologram, (ii) retrieved phase and (iii) amplitude. (c) and (d) Quantitative comparison of SSIM and PCC of retrieved phase and intensity respectively averaged over test dataset (60 images). a) Scale bar  $2 \mu\text{m}$  b) Scale bar  $10 \mu\text{m}$  and c) error bars are standard deviations from the mean.

To demonstrate the applicability of U-Net for low light biological holographic imaging, we used fibroblast cells cultured on a coverslip. Fibroblasts play a key role in repair and remodelling of the extra-cellular matrix and is highly active in the production of collagen. In this experiment, we limit the light density down to  $5 \text{ mW/cm}^2$  on the sample and performed the NN learning process as shown in Fig. 2(c). Figure 4(a) i) shows a representative input hologram under low light conditions. before they are restored by the Holo-UNet. The examples in the inset demonstrate how Holo-UNet can restore cell holograms under ultralow light conditions as low as  $\sim 300$  photons/pixel at a video rate acquisition time of  $200 \mu\text{s}$ .



**Fig. 4.** Restoration of hologram of fibroblast cells using Holo-UNet a) i) Input of hologram of fibroblast sample imaged under  $5 \text{ mW/cm}^2$  at  $200 \mu\text{s}$  and after Holo-UNet is used to restore hologram. a) ii) shows the retrieved phase and intensity of the first ROI using BM3D, PRNN [11] and Holo-UNet as compared to the ground truth. b) i) ii) iii) are quantitative comparison of the retrieved phase as measured by SSIM, MSE and PCC averaged over the test dataset (500 images) respectively. iv) Mean NPCC loss after each epoch of training during the fibroblast experiment a) Scale bars  $50 \mu\text{m}$ ,  $30 \mu\text{m}$ , inset (scale bar  $5 \mu\text{m}$ ) and b) error bars are standard deviations from the mean.

### 3.2. Fibroblast cells

Figure 4(a) ii) shows visual comparison of the retrieved phase and amplitude cropped from a larger field of view using three methods, BM3D, Holo-UNet and an example PRNN [11] in comparison to the ground truth. BM3D is applied to the same hologram inputs as Holo-UNet, while PRNN is applied to the phase and intensity images after standard phase retrieval process on the low light input images. Moreover, PRNN is trained using the same dataset as our method using the architecture proposed in [11] before comparison. In Fig. 4(b), SSIM, Mean Squared Error (MSE) and PCC are used on the test dataset to provide objective comparison of the improvement of the retrieved phase. Figure 4(b) i) shows SSIM is increased from 0.14 to 0.33, which is more than twice that attained by BM3D, and  $\sim 34\%$  better than PRNN. Figure 4(b) iii) shows the similarity with ground truth using MSE and a similar trend is observed in comparison to SSIM, with a 3-fold reduction from 18 to 5. However, PRNN outperforms Holo-UNet from 2.9 to 5.62. Using PCC in Fig. 4(b) ii), the intensity was also improved from 0.3 to 0.55, in comparison to 0.3 (BM3D) and 0.41 (PRNN). In both the input and BM3D results, the majority of the cells across the images were hidden by shot noise. However, Holo-UNet revealed high fidelity phase information of the cells that would otherwise not be possible using existing denoising methods (BM3D and PRNN). In our test dataset, U-Net outperforms PRNN in all metrics except MSE. However, on observation of test images, U-Net creates sharper images but also more distinct errors in comparison to PRNN. This is supported by U-Net's improved SSIM score which also measures the standard deviation across the image. We believe a key of part of the performance

difference lies in the application of our network before phase retrieval, which combined with FFT-NPCC prevents the blurring effect seen in PRNN. Table 1 summarises the quantitative results from SSIM, MSE and PCC for all the samples using BM3D, PRNN and Holo-UNet. Figure 4(b) iv) shows mean NPCC loss after supervised training data (black square) and the validation dataset (red circle). While the SSIM show significant variations, there is qualitative improvement in retrieve phase as shown the line plots of the amplitude and phase of selected retrieved images and the ground truth in the Fig. 6.

**Table 1. Comparison of methods w.r.t the ground truth of both Microspheres and Fibroblast cells.**

Method	Microspheres			Fibroblast			
	Input	BM3D	Holo-UNet	Input	BM3D	PRNN	Holo-UNet
<b>Hologram PCC</b>	0.35	0.42	<b>0.86</b>	0.23	0.24	N/A	<b>0.64</b>
<b>PCC FFT Magnitude</b>	0.51	0.06	<b>0.87</b>	0.34	0.14	N/A	<b>0.85</b>
<b>PCC Retrieved Intensity</b>	0.45	0.47	<b>0.76</b>	0.29	0.3	0.41	<b>0.55</b>
<b>SSIM Retrieved Phase</b>	0.11	0.12	<b>0.18</b>	0.14	0.14	0.19	<b>0.33</b>
<b>MSE Retrieved Phase (lower is better)</b>	14.75	16.83	<b>13.48</b>	18.57	14.87	<b>3</b>	5.62

#### 4. Conclusion

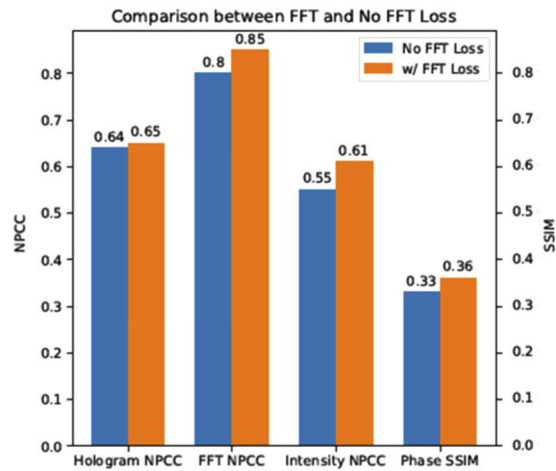
In conclusion, we demonstrate, for the first time, the use of a hologram-to-hologram neural network (Holo-UNet) to denoise digital holograms adversely affected by high shot noise under sub-mW/cm<sup>2</sup> intensities at sub-millisecond acquisition rates. In a shot noise hologram, the noise propagates into the spatial frequency domain after Fourier transformation and easily creates critical errors in both the Fourier domain filtering and phase unwrapping. It is therefore expected that the complete removal of the shot noise in the hologram is not possible because information has already been lost. However, we demonstrate that Holo-UNet can successfully generate sharper phase and intensity images (~ 1.8 folds higher SSIM) in comparison to other methods, which are less suited to the task of off-axis hologram denoising. Moreover, our findings show the importance of applying NPCC on both the hologram and the Fourier domain. We hypothesise that image-based transformations (e.g. FFT) as a part of the loss function will be useful in other holography imaging problems prone to noise such as highly sensitive electron interferometry [23]. Furthermore, path length matching in low coherence light sources can also affect phase quality, especially during long term imaging over several weeks. Holo-UNet deals with primarily intensity shot noise due to low light, instead of optical scattering in highly dense sample. Hence, Holo-UNet has the potential to perform long term cell imaging using a lower coherence light source (light emitting diode) in a portable QPM setup. Overall, the results showed that only a fraction of amount light is needed to retrieve a high quality phase image. Our technique opens up high speed QPM imaging without changing any hardware (illumination intensity or exposure rate). Aside from UNet, we also aim to investigate other neural net architectures and learning paradigms. This also paves the way for promising application in the quantitative phase imaging of photo-sensitive samples with different lasers and wavelengths for both biological and material sciences.

#### Appendix

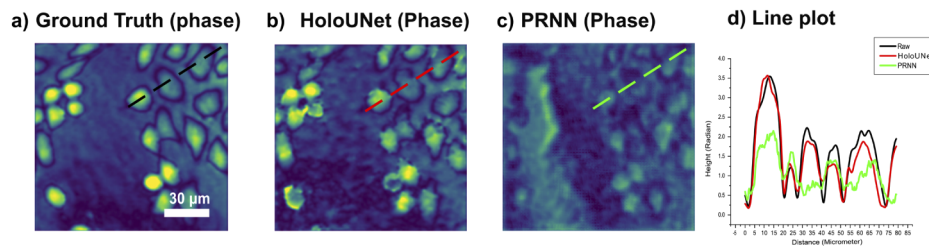
The addition of the NPCC loss improved the results of the network throughout the process. In particular, both reconstructed intensity and phase were improved by 10% on average as shown in Fig. 5. and in Fig. 6., we show the line plots of the amplitude and phase of selected retrieved



images and the ground truth to show the qualitative improvement of the Holo-UNet versus ground truth.



**Fig. 5.** Comparison between using NPCC with FFT vs NPCC without FFT loss on the fibroblast cells dataset.



**Fig. 6.** Comparison phase retrieved from holograms taken from ground truth and Holo-UNet on the fibroblast cells dataset. a-c) depict an example Fibroblast phase image with the dashed line denoting the position of cross-section. d) Shows the phase value along the dash-lines of a-c.

## Funding

Australian Research Council (DE160100843, DP190100039, DP200100364).

## Acknowledgments

The authors acknowledge the JCSMR Imaging and Cytometry Facility for their technical support and advice. We thank David Tschärke for providing fibroblast cells and Elizabeth Gardiner for discussion in biological imaging using QPM.

## Disclosures

All authors declared no conflict of interest.

## References

1. V. Micó, J. Zheng, J. Garcia, Z. Zalevsky, and P. Gao, "Resolution enhancement in quantitative phase microscopy," *Adv. Opt. Photonics* **11**(1), 135–214 (2019).

2. N. M. Editorial, "Phototoxicity revisited," *Nat. Methods* **15**(10), 751 (2018).
3. F. Charrière, B. Rappaz, J. Kühn, T. Colomb, P. Marquet, and C. Depeursinge, "Influence of shot noise on phase measurement accuracy in digital holographic microscopy," *Opt. Express* **15**(14), 8818–8831 (2007).
4. G. Choi, D. Ryu, Y. Jo, Y. S. Kim, W. Park, H.-S. Min, and Y. Park, "Cycle-consistent deep learning approach to coherent noise reduction in optical diffraction tomography," *Opt. Express* **27**(4), 4927–4943 (2019).
5. C. A. Metzler, F. Heide, P. Rangarajan, M. M. Balaji, A. Viswanath, A. Veeraraghavan, and R. G. Baraniuk, "Deep-inverse correlative: towards real-time high-resolution non-line-of-sight imaging," *Optica* **7**(1), 63–71 (2020).
6. Z. Ren, Z. Xu, and E. Y. Lam, "End-to-end deep learning framework for digital holographic reconstruction," *Adv. Photonics* **1**(1), 016004 (2019).
7. Y. Rivenson, Y. Wu, and A. Ozcan, "Deep learning in holography and coherent imaging," *Light: Sci. Appl.* **8**(1), 85 (2019).
8. G. Zhang, T. Guan, Z. Shen, X. Wang, T. Hu, D. Wang, Y. He, and N. Xie, "Fast phase retrieval in off-axis digital holographic microscopy through deep learning," *Opt. Express* **26**(15), 19388–19405 (2018).
9. H. Wang, M. Lyu, and G. Situ, "eHoloNet: a learning-based end-to-end approach for in-line digital holographic reconstruction," *Opt. Express* **26**(18), 22603–22614 (2018).
10. A. Goy, K. Arthur, S. Li, and G. Barbastathis, "Low photon count phase retrieval using deep learning," *Phys. Rev. Lett.* **121**(24), 243902 (2018).
11. 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.* **7**(2), 17141 (2018).
12. S. Montrésor, M. Tahon, A. Laurent, and P. Picart, "Computational de-noising based on deep learning for phase data in digital holographic interferometry," *APL Photonics* **5**(3), 030802 (2020).
13. X. He, C. V. Nguyen, M. Pratap, Y. Zheng, Y. Wang, D. R. Nisbet, R. J. Williams, M. Rug, A. G. Maier, and W. M. Lee, "Automated Fourier space region-recognition filtering for off-axis digital holographic microscopy," *Biomed. Opt. Express* **7**(8), 3111–3123 (2016).
14. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, (Springer, 2015), 234–241.
15. J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2015), 3431–3440.
16. G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica* **6**(8), 921–943 (2019).
17. M. Weigert, U. Schmidt, T. Boothe, A. Müller, A. Dibrov, A. Jain, B. Wilhelm, D. Schmidt, C. Broaddus, S. Culley, M. Rocha-Martins, F. Segovia-Miranda, C. Norden, R. Henriques, M. Zerial, M. Solimena, J. Rink, P. Tomancak, L. Royer, F. Jug, and E. W. Myers, "Content-aware image restoration: pushing the limits of fluorescence microscopy," *Nat. Methods* **15**(12), 1090–1097 (2018).
18. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, (2012), 1097–1105.
19. M. Weigert, U. Schmidt, T. Boothe, A. Müller, A. Dibrov, A. Jain, B. Wilhelm, D. Schmidt, C. Broaddus, and S. Culley, "Content-aware image restoration: pushing the limits of fluorescence microscopy," *Nat. Methods* **15**(12), 1090–1097 (2018).
20. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. on Image Process.* **16**(8), 2080–2095 (2007).
21. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Process.* **13**(4), 600–612 (2004).
22. M. D. Zeiler, "Adadelta: an adaptive learning rate method," arXiv preprint arXiv:1212.5701 (2012).
23. M. Hýtch, F. Houdellier, F. Hüe, and E. Snoeck, "Nanoscale holographic interferometry for strain measurements in electronic devices," *Nature* **453**(7198), 1086–1089 (2008).