

Article

Θ-Net: A Deep Neural Network Architecture for the Resolution Enhancement of Phase-Modulated Optical Micrographs *In Silico*

Shiraz S. Kaderuppan^{1,*}, Anurag Sharma¹, Muhammad Ramadan Saifuddin¹, Wai Leong Eugene Wong²
and Wai Lok Woo³

¹ Faculty of Science, Agriculture & Engineering (SAGe), Newcastle University, Newcastle upon Tyne NE1 7RU, UK; anurag.sharma@newcastle.ac.uk (A.S.); ramadan.saifuddin@newcastle.ac.uk (M.R.S.)

² Engineering Cluster, Singapore Institute of Technology, 10 Dover Drive, Singapore 138683, Singapore; eugene.wong@singaporetech.edu.sg

³ Computer and Information Sciences, Sutherland Building, Northumbria University, Northumberland Road, Newcastle upon Tyne NE1 8ST, UK; wailok.woo@northumbria.ac.uk

* Correspondence: s.s.o.kaderuppan@newcastle.ac.uk; Tel.: +65-96-92-3371

Abstract: Optical microscopy is widely regarded to be an indispensable tool in healthcare and manufacturing quality control processes, although its inability to resolve structures separated by a lateral distance under ~200 nm has culminated in the emergence of a new field named *fluorescence nanoscopy*, while this too is prone to several caveats (namely phototoxicity, interference caused by exogenous probes and cost). In this regard, we present a triplet string of concatenated O-Net (‘bead’) architectures (termed ‘Θ-Net’ in the present study) as a cost-efficient and non-invasive approach to enhancing the resolution of *non-fluorescent* phase-modulated optical microscopical images *in silico*. The quality of the afore-mentioned enhanced resolution (ER) images was compared with that obtained via other popular frameworks (such as ANNA-PALM, BSRGAN and 3D RCAN), with the Θ-Net-generated ER images depicting an increased level of detail (unlike previous DNNs). In addition, the use of cross-domain (transfer) learning to enhance the capabilities of models trained on differential interference contrast (DIC) datasets [where phasic variations are not as prominently manifested as amplitude/intensity differences in the individual pixels unlike phase-contrast microscopy (PCM)] has resulted in the Θ-Net-generated images closely approximating that of the expected (ground truth) images for both the DIC and PCM datasets. This thus demonstrates the viability of our current Θ-Net architecture in attaining highly resolved images under poor signal-to-noise ratios while eliminating the need for *a priori* PSF and OTF information, thereby potentially impacting several engineering fronts (particularly biomedical imaging and sensing, precision engineering and optical metrology).

Keywords: deep neural networks; image denoising; computational phase-modulated nanoscopy; label-free optical imaging; biomedical imaging



Citation: Kaderuppan, S.S.; Sharma, A.; Saifuddin, M.R.; Wong, W.L.E.; Woo, W.L. Θ-Net: A Deep Neural Network Architecture for the Resolution Enhancement of Phase-Modulated Optical Micrographs *In Silico*. *Sensors* **2024**, *24*, 6248. <https://doi.org/10.3390/s24196248>

Academic Editors: Yajun Wang, Xiangchao Zhang and Jian Wang

Received: 30 August 2024

Revised: 23 September 2024

Accepted: 23 September 2024

Published: 26 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Over the past few decades, the field of artificial intelligence (AI) has witnessed considerable progress and deployments in numerous applications, such as computer vision [1,2], speech and text recognition [3,4] and cyber security [5,6], amongst others. Fueled primarily by developments in computing hardware resources (such as GPUs and memory [7]), a particular subset of AI algorithms (termed *deep neural networks* or DNNs for short) has played a fundamental role in driving this revolution. In this respect, it would be prudent to evaluate the impact of these DNNs in propelling developments in healthcare and ecological studies, two aspects which play prominent roles in circumventing present-day global dilemmas, such as pandemics [8] and climate change [9], amongst others. Of particular interest in this regard would be the role of these DNNs in image analysis [10], object detection and segmentation [11,12].

As a trusted primary approach formulated to address these healthcare and environmental issues, the role of optical microscopes cannot be undermined as they aid in the identification of host responses to disease-causing pathogens (e.g., the presence of tumors [13], impact of pathogens on native cellular metabolomics and molecular processes occurring *in vivo* [14,15], etc.), as well as the detection of ecologically essential microbiota (e.g., diatoms and phytoplankton, such as *Euglena* [16], coupled with soil microbes such as *P. alcaligenes* [17]) which play an essential role in the upkeep of marine and terrestrial ecosystems. Nonetheless, optical microscopes are plagued by a fundamental constraint—their lateral resolution (as they are often utilized) is confined to a minimum distance of 143nm, often described as the Abbe limit [18]. Numerous approaches (both optical and computational) have thus been proposed to circumvent this limitation. The former comprises mainly optical nanoscopic/super-resolution (SR) techniques [such as SIM [19], STED [20], PALM [21] and STORM [22] (amongst others)], realized through the addition of specialized but costly hardware attachments to the optical microscope, while the latter utilizes DNNs (such as ANNA-PALM [23], Deep-STORM [24], DeepZ [25] or 3D RCAN [26]) in seeking to enhance the resolution of micrographs acquired via conventional imaging techniques (often epi-fluorescence microscopy) *in silico*.

Despite the ubiquitous availability of DNN frameworks for SR imaging, we have come to realize that a number of the DNNs proposed (such as [23,24]) were developed to enhance the resolution of micrographs acquired via traditional widefield *epi-fluorescence microscopy*, where there is a clear distinction between the signal and the background (which is often dark). Such algorithms may thus also perform expectably well for other microscopical imaging modalities such as darkfield microscopy [27] or cross-polarized light microscopy [28], where the background is clearly demarcated by the extinction of optical signals. In contrast, for imaging modalities where there is little discrimination between the background and the signal being resolved, or where artifacts (such as halos or pseudo-relief features) are present in the image due to the imaging technique utilized (as is evident in DIC or PCM), some of these algorithms may *not* be suitable for increasing the resolution of the input images in this context (as depicted by the implementation of ANNA-PALM [23] in the Results Section of the present study). Consequentially, we have sought to develop a novel framework (named O-Net) which we described in our previously published study [29], although the potentiality of further improvements to the O-Net framework were noted to be plausible in its subsequent refinements. In this regard, we now propose an *extension* of the O-Net architecture (which we term Θ -Net) as a viable means of increasing the resolution of micrographs acquired through phase-modulated optical microscopical techniques (namely PCM and DIC microscopy). Θ -Net employs a concatenated architecture of multiple O-Nets (i.e., a computational ‘string of beads’), coupled with the transfer learning of features across these different phase-modulated microimaging modalities, so as to exploit the learning paradigms of phase-sensitive features resolvable through complementary phase microscopy techniques. Schematically, the Θ -Net architecture may be represented in Figure 1 as follows:

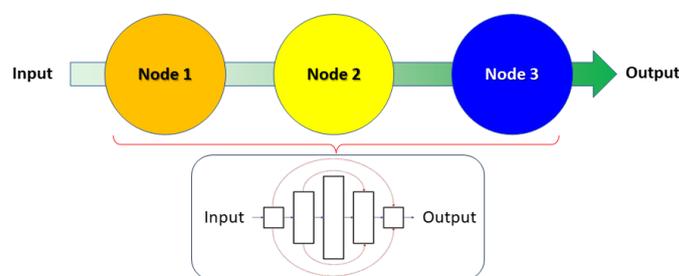


Figure 1. A generalized schematic (overview) of the Θ -Net architecture as proposed in the current context. Θ -Net adopts a ‘string of beads’ methodology of concatenating multiple O-Nets, thereby enhancing the DNN’s resilience to feature-based variations present in different samples that it might be trained with. Here (and as is presented in the current study), we employ a 3-node Θ -Net framework for model training and validation.

The testing of the proposed Θ -Net framework was performed using a separate dataset for each of the DIC and PCM imaging modalities. The results obtained in this respect (described in the following Figures 3–7 of the present study) seem promising, with ultra-structural details evident in these ER images.

2. Materials and Methods

2.1. Data Acquisition and Preparation (Light Microscopy)

The image datasets used for training were similar to those used for training O-Net, as expounded in our previous article in [29]. Generally, commercially available, prepared microscope slides (utilizing samples from a range of plant, animal and human tissues, as well as microbiota) were used for image acquisition to create the training datasets for the assessed DNNs. The images were acquired through 2 primary imaging modalities—PCM and DIC microscopy—using a Leica N PLAN L 20X/0.4 Corr Ph1 objective (Leica P/N: 506058) [for low-resolution (LR) images] and a Leica HCX PL Fluotar L 40X/0.60 Corr Ph2 objective (Leica P/N: 506203) [for high-resolution (HR) images] installed on a Leica DM4000M microscope, with a CMOS camera (RisingCam[®] E3ISPM12000KPA, RisingCam, China) having a pixel size of $1.85 \mu\text{m} \times 1.85 \mu\text{m}$ and an EK 14 Mot motorized stage (Märzhäuser Wetzlar GmbH & Co. KG, Wetzlar, Germany) mounted on it. The control of the motorized components of the microscope and camera settings was facilitated through a self-developed stage controller coupled with a desktop UI (also developed in C#). For each sample, similar regions of interest (ROIs) were imaged (under both PCM and DIC microscopes), with the acquired images registered and cropped using MATLAB R2020a (© 1984–2020, The MathWorks, Inc., Natick, MA, USA). Where present, shifts detected in the ROIs were mitigated through multi-layer image cropping using Corel PHOTO-PAINT X7 (© 2015 Corel Corporation), prior to splitting into $600 \text{ px} \times 600 \text{ px}$ RGB image tiles in MATLAB R2020a (© 1984–2020, The MathWorks, Inc.). These tiles were then conjoined into LR-HR image pairs [the LR images being the **Source** (to be transformed) and the HR images being the **Expected** (*target/ground truth*) images], resulting in a total of 3944 image pairs for each imaging modality employed (i.e., DIC or PCM), before being downsampled to a $256 \text{ px} \times 512 \text{ px}$ format and parsed into NumPy arrays for training the individual nodes of the proposed Θ -Net network in Python 3.8.

2.2. Θ -Net Architecture

The currently proposed framework (Θ -Net) utilizes a ‘string of beads’ architecture, comprising multiple (in this context, three) O-Nets as described in [29]. The O-Net models utilized for each of the nodes in the Θ -Net scaffold were selected after multiple empirical runs evaluating model architectures of varying depths and trained over a range of epochs (up to a maximum of 320 epochs). Further details pertaining to the individual model architectures are described in the following sub-sections.

2.2.1. Generalized Θ -Net Structure

In the present study, we attempted to enhance the resolution of the input DIC images by employing a 7-layer O-Net model trained with the DIC dataset for both the 1st and 2nd nodes and a 7-layer O-Net model trained on *both* the PCM and DIC datasets (the latter of which is *transfer-learned*) for the 3rd node of the Θ -Net architecture. For the PCM images, we used a 5-layer O-Net model for the first node but a 7-layer O-Net model for the 2nd node. For the third node, we used the same transfer-learned O-Net model as was used for the DIC images. Figure 2 below depicts a general schematic of the Θ -Net framework utilized in this regard.

At this juncture, it would be essential to highlight that the individual O-Net nodes within a Θ -Net chain are founded on the Pix2Pix GAN architecture, as discussed in [29,30]. Similarly, Swish [31] and GELU [32] activation functions were also incorporated into the encoder and decoder blocks, respectively (the mathematical definitions of these functions are included in the accompanying Supplementary Materials for the interested reader). An

adaptive moments estimation (Adam) optimizer [33] was also employed for use in the GAN, the hyperparameters of which are as described in the associated Supplementary Materials.

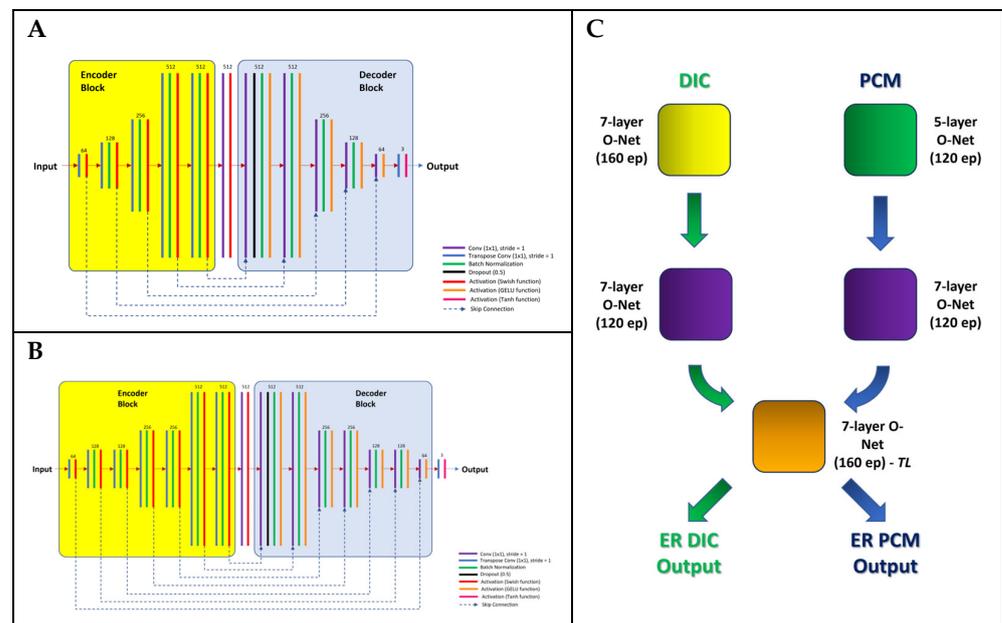


Figure 2. The structure of the 5-layer (Panel A) and 7-layer (Panel B) O-Net architectures utilized as nodes for Θ -Net (Panel C), as described in the present study. Each of the 3 nodes of Θ -Net (shown in Figure 1 previously) is an O-Net model specifically trained with the input image dataset for the imaging modality which it is intended to be deployed for use in. The exception here refers to the 3rd node utilized in the current Θ -Net framework—the O-Net model here was trained using both the DIC and PCM image datasets, via a *transfer learning* approach. As with the traditional O-Net architecture (described in [29]), skip-connections (*concatenations*) are used to join layers in the encoder block (consisting of transposed convolution operations) with their corresponding conjugates in the decoder block (comprising convolution operations).

The images generated by the Θ -Net models were also analyzed in MATLAB R2022b (© 1984–2022, The MathWorks, Inc.) using standard image quality metrics, such as the peak signal-to-noise ratio (PSNR), signal-to-noise ratio (SNR), image mean-squared error (IMSE) and structural similarity index (SSIM) (where plausible). The findings (and insights) gleaned from this assessment are henceforth depicted in the Results Section of this study.

2.2.2. Cross-Domain Learning

An interesting aspect of the present study refers to the employment of *cross-domain learning* [34] across 2 commonly utilized phase-modulated optical microscopy techniques (i.e., PCM and DIC microscopy). Here, we demonstrate that it is possible to extricate information gleaned from images acquired under each of these techniques to *enhance* the overall performance of the Θ -Net models in computational nanoscopy (as exemplified in the Results Section of this study). This may potentially be attributable to **both** DIC and PCM translating *optical path length* (OPL) variations in the sample into amplitude (image brightness/intensity) differences, albeit via different routes (PCM translates differences in OPL *magnitude* into brightness variations, while DIC converts OPL *gradient* differences into 3D relief effects [35]). In this respect, we seek to utilize *transfer learning* as a vehicle to facilitate cross-domain learning between these 2 microscopical imaging modalities, thereby spelling benefits for the Θ -Net-generated ER images to incorporate the ‘best of both worlds’ while seeking to circumvent the limitations faced when employing each technique individually (e.g., birefringence artifacts encountered in DIC, as well as the lack of lateral

and axial resolution coupled with unsuitability for specimens having high phase shifts for PCM imaging [35]).

2.2.3. Comparative Networks for ER Image Analysis

Separately, we also attempted to include models derived from other industry-leading DNN frameworks (such as ANNA-PALM [23] and 3D RCAN [26], the latter of which is employed for *in silico* nanoscopy in the Aivia deep learning suite [36]), as a comparative performance gauge of the proposed Θ -Net models. Nonetheless, it would be apt to emphasize at this juncture (as was also highlighted in the Introduction Section earlier) that these frameworks were *specifically* adapted for increasing the resolution of *widefield epi-fluorescence monochrome/grayscale* microscopical images (3D RCAN [26] uses grayscale image Z-stacks for generating SR images); hence, a somewhat *close* (albeit not very similar) comparative assay in this respect would require splitting the current 2D RGB image into a 3-channel grayscale image stack, which would then be simulated as a Z-stack for training and validating the 3D RCAN [26]-based models. For ANNA-PALM [23] validation, we utilized the ImageJ A-Net plugin downloaded from [37] and the supplied models trained for increasing the resolution of microtubules within ImageJ 1.52n (NIH, USA), as the structures we wanted to enhance the resolution of generally resembled (in part) the filamentous strands characteristic of microtubules. In addition, models developed on a third comparative framework (termed BSRGAN [38]) were also included in the present study, for verification purposes.

2.3. Image Denoising

In addition to ER images, we also attempted to evaluate the capability of the proposed Θ -Net models for image denoising, despite the models not being trained specifically for this purpose. To conduct this, salt-and-pepper noise was synthetically introduced into the validation images using MATLAB R2022b (©1984–2022, The MathWorks, Inc.), and the assayed Θ -Net models were assessed on their propensity to denoise the noise-infused images.

The codes and models used to generate and evaluate the images presented in this study are available for download in the accompanying Supplementary Materials.

3. Results

The models developed using the presently assayed Θ -Net architecture were contrasted against our previous O-Net models [29], as well as other state-of-the-art *in silico* SR architectures, such as 3D RCAN [26], BSRGAN [38] and ANNA-PALM [23], for each of the DIC and PCM imaging modalities. The generated images obtained from each of these models are depicted in the subsequent figures.

3.1. DIC Imaging

From Figure 3 and its associated Figure S11 in the accompanying Supplementary Materials, it may be clearly discerned that images generated via the Θ -Net-trained models exhibit an enhanced contrast and, consequently, an increased resolution of details and features as compared to those generated from the O-Net-trained models. Nonetheless, pseudo-relief depictions of ER features are also apparent in the Θ -Net-generated images, although network hallucination (a prominent flaw in DNN-generated images as highlighted in [39]) is not clearly evident in both the O-Net- and the Θ -Net-generated images. In some respects, it may be argued that the Θ -Net ER images provide greater clarity than even the **Expected** images, suggesting an *enhanced* feature detection realized through the application of the Θ -Net models, putatively due to the effect of the transfer learning of features from the PCM dataset.

Figure 4 below showcases the validation of both the O-Net and Θ -Net models with an *untrained* image dataset (comprising images which the models were never trained with). In addition, models adopting the popular U-Net architecture, 3D RCAN [26], BSRGAN [38] and ANNA-PALM [23] were also included (the said models representing the industry standard for most *in silico* ER image applications). Notably, it may be observed that

although both the O-Net and Θ -Net models seemingly outperform models adopting the U-Net [11], 3D RCAN [26], BSRGAN [38] and ANNA-PALM [23] architectures (in terms of the details resolvable visually), *none* of the models *may* be completely absolved from network hallucination artifacts, although this is subject to further evaluation and discussion (as detailed in a later section of this study).

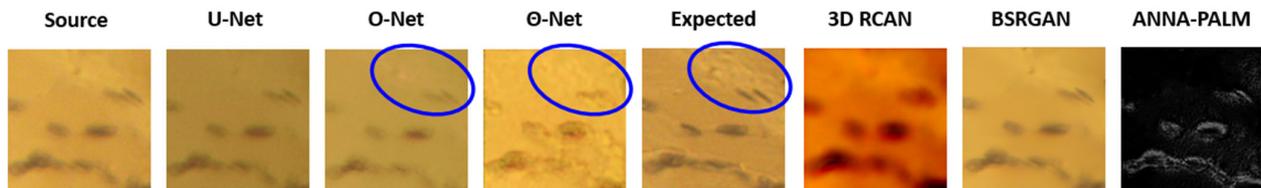


Figure 3. The validation of *in silico* ER images obtained through various models, including those developed using O-Net [29] and the presently surfaced Θ -Net. The sample shown here consists of highly magnified views of skeletal muscle tissue (L.S.) (adapted from Figure S11 of the Supplementary Materials, with further evaluation images being presented in this figure for the interested reader as well). The O-Net model was trained over 101 epochs, while the Θ -Net model assimilated O-Net models trained over 160 epochs (for the 1st node), a 120-epoch-trained O-Net model (for the 2nd node) and a 160-epoch-trained and transfer-learnt O-Net model (for the 3rd node). Notice the closer resemblance of the Θ -Net-generated image to the **Expected** (ground truth) image, as compared to the O-Net models (highlighted within the blue ellipses). **N.B.:* The **Source** image (input) was acquired via a 20X/0.40 Ph1 objective, while the **Expected** (ground truth) image was obtained using a 40X/0.60 Ph2 objective. Images generated through models founded on other frameworks (namely 3D RCAN [26], BSRGAN [38] and ANNA-PALM [23]) were also included for comparison purposes (the ANNA-PALM [23] model developed for increasing the resolution of grayscale photomicrographs of microtubules was utilized as an extension within ImageJ 1.52n (NIH, USA), while the 3D RCAN [26] model was trained over 250 epochs with 1972 steps per epoch and 2 residual groups). **N.B.:* The Θ -Net models utilized for generating the ER images in this figure and Figure S11 implement optional *node scaling* for each node, differing from the rest of this study. The supplied code (described in the accompanying Supplementary Materials) allows the user to select whether node scaling should be applied (or not), based on the user's discernment of their image dataset and acquisition parameters.

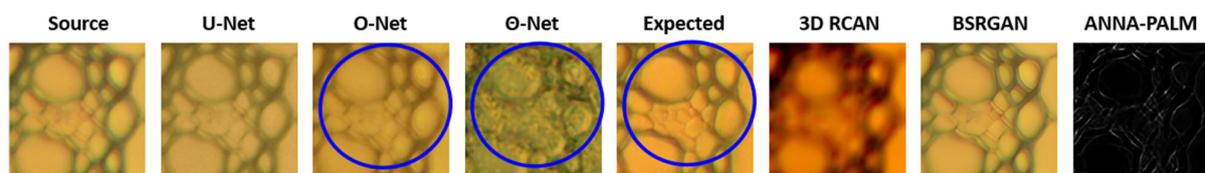


Figure 4. *In silico* ER images obtained through several models, including those developed using U-Net [11], O-Net [29] and Θ -Net (similar O-Net and Θ -Net models as described in Figure 3 previously were used for the ER images here as well). As with Figure 3, the **Source** image (input) was acquired via a 20X/0.40 Ph1 objective, while the **Expected** (ground truth) image was obtained using a 40X/0.60 Ph2 objective. **N.B.:* This figure was sourced from Figure S12 of the accompanying Supplementary Materials. Notice that some features (such as the cell walls within the blue ellipses as shown here or the shadow-like striations within the green ellipses in Figure S12) are resolved differently in the Θ -Net models as compared to that from O-Net—in the Θ -Net images, the cell walls appear more granular around the periphery, while the striations are less visible. This might be postulated to be due to the Θ -Net architecture adopting a transfer-learnt model from PCM; hence, pseudo-relief artifacts characteristic of DIC imaging would be less pronounced (leading to a reduced accentuation of the striations), while phasic variations identified in the image are characterized pixel-wise (accounting for the granular edges of the cell wall when adopting Θ -Net models for ER images). **N.B.:* The deployment of the Θ -Net model utilized for generating the ER images in this figure does *not* employ node scaling, differing from that used for Figure 3.

Approaching from a computational standpoint, we show the plots of the respective loss functions, namely the discriminator losses on real samples (dR) and generated samples (dG) and the generator (g) loss for the training runs in the various nodes of the Θ -Net models in Figure 5.

Conversely, an analysis of the images acquired via PCM revealed the following observations (detailed in the subsequent sections).

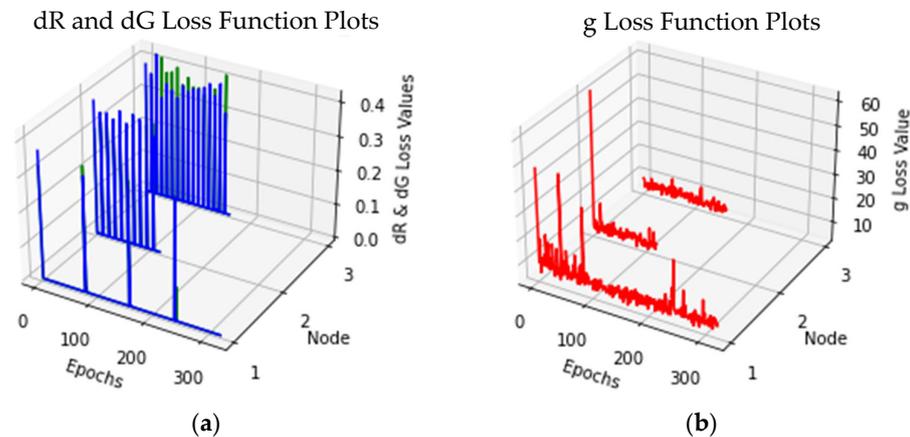


Figure 5. Loss functions for DIC micrographs imaged under the Θ -Net architecture. (a) Discriminator losses on *both* the real and generated samples, represented by dR (green line) and dG (blue line) losses, respectively, as well as (b) the generator (g) loss plotted as a red line. The spikes observed in the dR and dG loss function plots (a) demonstrate the commencement of a subsequent training run, implicating the selection of a new random seed for DNN training. Moreover, this approach of intermittently training the DNN models over multiple runs (rather than a single continuous run) is preferred [40] as it *restructures* the loss function error landscape, allowing the model to reach its global minimum (even if it may be trapped in one of the local minima during an earlier training run).

3.2. PCM Imaging

Analyses of the images portrayed in Figure 6 (coupled with its corresponding Figure S13 in the accompanying Supplementary Materials) depict a relatively similar trend to that observed for Figure 3 (i.e., that Θ -Net-generated images more closely resemble the ground truth images as compared to O-Net-produced images), favoring the ‘*string of beads*’ architecture hypothesis (characteristic of Θ -Net) proposed in the present study to attain ER images in computational nanoscopy. Nonetheless, as with Figure 3 previously, there still exists room for improvement when comparing the effective resolution of the Θ -Net-generated images with the ground truth images, implying that more nodes/*beads* could putatively be added to the ‘*string of beads*’ Θ -Net architecture in a bid to improve ER images.

It would also be prudent to highlight at this juncture that the dilemma of network hallucination cannot be totally absolved from any of the model-generated images as depicted by the ROIs within the red and green ellipses in Figure S13, where some features evident within the **Expected** image were not detected in any of the DNN-processed images (including O-Net and Θ -Net). However, a closer inspection of the Θ -Net ER images does reveal a slight variation in pixel intensities at the points correlating to these features. Moreover, the ROI indicated by the green ellipse in the Θ -Net-generated image of Figure S13 seems to suggest potential artifacts caused by network hallucination within the yellow-colored channel, although the observed variations might also potentially arise as a consequence of phase variations in the sample instead.

Figure 7 (and its associated parent Figure S14) portrays how the Θ -Net models perform (in relation to other models, including O-Net and U-Net) with regards to increasing the resolution of PCM images for which they were *not* explicitly trained on (i.e., none of the models depicted here were trained with any of these images; hence, a verifiable deduction may be drawn in this respect). From these figures, we may observe that Θ -Net models gen-

erally provide increased spatial resolution (when compared to the other models, including U-Net and O-Net), although a closer inspection between the Θ -Net-derived images and the ground truth may reveal some slight differences between these two images. In this regard, one may be prompted to conclude that this is a consequence of network hallucination artifacts in Θ -Net, but it would be prudent to mention that this might not necessarily be so, as the Θ -Net model architecture adopts a transfer-learned model (in its final O-Net mode) which incorporates similar features learnt differently from both DIC and PCM images. This allows Θ -Net to be more resilient to phase variations in samples when imaged across different phase-modulated microscopical modalities—an aspect which is not addressed through any one microscopical approach alone, thereby subjecting these imaging modalities to potential artifacts [such as the ‘halo’ effect (in PCM) or the pseudo-relief topography (observed in DIC)]. In the Θ -Net-generated images, however, these phase differences are converted into amplitude differences (when they fall within a certain range), allowing features to be resolved when these are often occluded by the ‘halo’ (or pseudo-relief) effects characteristic of PCM (and DIC microscopy).

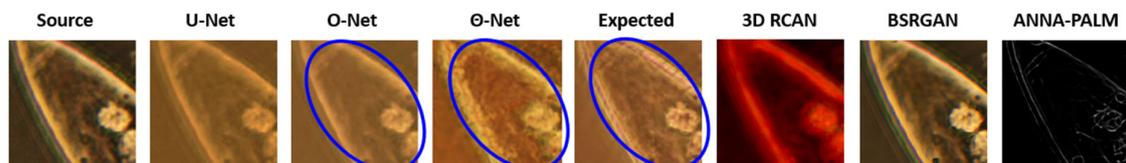


Figure 6. A comparison of the images obtained through utilizing models developed on the O-Net and Θ -Net architectures using PCM micrographs. Here, the O-Net model was trained over 120 epochs, while the Θ -Net framework incorporated the said O-Net model (in its 1st node), a 120-epoch-trained O-Net model (for the 2nd node) and a 160-epoch-trained and transfer-learned O-Net model (for the 3rd node), which is identical to that used for the DIC micrographs in Figure 3 previously. Also (as enunciated for the images in Figure 3), the **Source** (input) image was acquired via a 20X/0.40 Ph1 objective, while the **Expected** (ground truth) image was obtained with a 40X/0.60 Ph2 objective). **N.B.:* This figure was sourced from Figure S13 of the accompanying Supplementary Materials. As with the Θ -Net-generated DIC micrographs in Figure 3 previously, the region encircled within the blue ellipse of the Θ -Net-generated images depicts an enhanced level of detail (as compared to its O-Net counterpart), supporting the proposed Θ -Net framework as a viable improvement over O-Net (for producing computational models to facilitate *in silico* ER). Here too (as with Figure 3 previously), the ANNA-PALM [23] model was utilized as an ImageJ extension for microtubule SR, while the 3D RCAN [26] model was trained over 250 epochs with 1972 steps per epoch and 2 residual groups. Contrast enhancement (in MS PowerPoint) was also applied to the Θ -Net-derived image, making it easier to discern the features.

As with the Θ -Net models utilized for increasing the resolution of DIC images earlier, we sought to plot the respective loss functions (dR, dG and g loss) for the various Θ -Net PCM models, as shown in Figure 8 below. Here too, a similar legend is used for the plots—blue is indicative of dG loss, green for dR loss and red for g loss.

In this respect, it would be vital to mention that when evaluating the performance of a generative adversarial network (GAN) such as that of Θ -Net, there is *no* objective loss function which can be used to conduct this [41]; hence, we chose to utilize the discriminator losses on the real and generated samples (as well as the generator loss) as key determinants for assessing Θ -Net performance. We conducted this for the Θ -Net models trained on *both* the DIC and PCM datasets (as depicted in Figures 5 and 8, respectively). The formulae underlying these losses are indicated in the Discussion Section of this study.

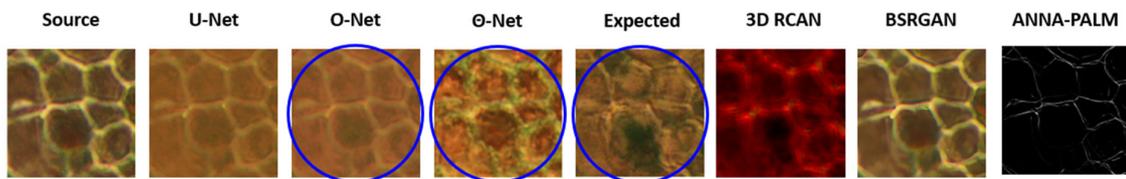


Figure 7. ER images obtained through several models, including those developed using O-Net [29] and Θ -Net. The O-Net model was trained over 101 epochs, while the Θ -Net model assimilated O-Net models trained over 120 epochs (for the 1st node), a 120-epoch-trained O-Net model (for the 2nd node) and a 160-epoch-trained and transfer-learned O-Net model (for the 3rd node). Notice the increased resolution evident in the Θ -Net-generated image, which seemingly surpasses that generated via O-Net and even the **Expected** (ground truth) images, as highlighted within the blue ellipses. **N.B.*: The **Source** image (input) was acquired via a 20X/0.40 Ph1 objective, while the **GT** (ground truth) image was obtained using a 40X/0.60 Ph2 objective. Images generated through models founded on other frameworks (namely 3D RCAN [26], BSRGAN [38] and ANNA-PALM [23]) were also included for comparison purposes (the ANNA-PALM [23] model developed for increasing the resolution of grayscale photomicrographs of microtubules was utilized as an extension within ImageJ 1.52n (NIH, USA), while the 3D RCAN [26] model was trained over 250 epochs with 1972 steps per epoch and 2 residual groups). Brightness enhancement (in MS PowerPoint) was applied to the Θ -Net-derived image, facilitating the discernment of the features. This figure was adapted from Figure S14 of the Supplementary Materials; further evaluation images are presented in Figure S14 for the interested reader).

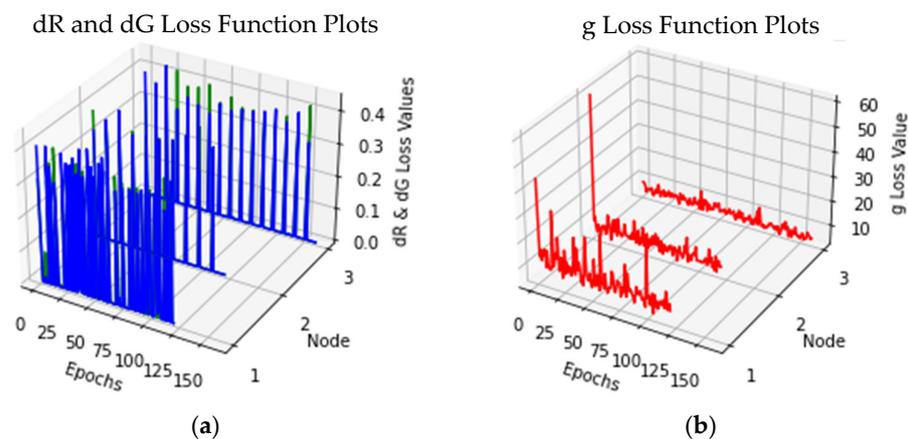


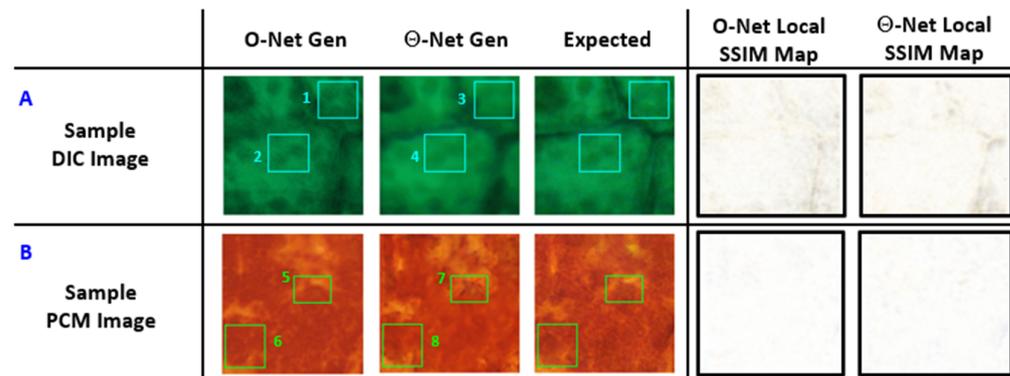
Figure 8. Loss function plots for Θ -Net models trained on PCM micrographs. (a) denotes the discriminator losses for real (dR) and generated (dG) samples, indicated by green and blue lines, respectively, while (b) represents the generator (g) loss (as red lines). As with Figure 5 previously, spikes present in the plots of the dR and dG loss functions indicate the start of a new training run (using the existing Python script), potentially implying the selection of a new random seed at the start of each run (model training for this dataset was also conducted intermittently).

3.3. Computation of Global and Local Image Metrics

A further evaluation of the global and local image metrics (for specific ROIs in a separate subset of the validation images) was subsequently conducted, and the results are depicted in Figure 9 as follows.

From the findings in Figure 9, we may deduce that Θ -Net generally performs relatively well in increasing the resolution of phase-modulated micrographs (the images chosen in this context were randomly selected for validating both the O-Net and Θ -Net model performance). Despite this fact, regions where Θ -Net was noted to have underperformed (when compared to O-Net) were prominent in the PCM image dataset, although further inspection revealed that this reduced performance may be attributed to Θ -Net seeking to enhance the resolution of fine structural details within the **Source** image, resulting in

lowered PSNR scores (since this metric is often used to penalize noise, while the latter may be reminiscent of *pseudo-noise*, as discussed in [29]). In this context, we would thus need to evaluate the performance of Θ -Net from the perspective of other metrics as well (such as SSIM), since these provide a holistic representation of Θ -Net as a DNN framework for *in silico* ER microscopy. Considering this perspective, we notice that the SSIM scores of Θ -Net-generated images closely approach those of O-Net (differing within 1%), implying that there is very little difference between the quality of Θ -Net- and O-Net-generated images for this (PCM) dataset



Panel	Image Type	ROI	Global Values	PSNR	SNR	IMSE	SSIM
A (Sample DIC Image)	O-Net-Generated DIC	1	PSNR: 26.8852 SNR: 19.6477 IMSE: 133.2161 SSIM: 0.9613	25.8452	15.8162	169.2621	0.9578
		2		27.0975	17.8477	126.8624	0.9724
	Theta-Net-Generated DIC	3	PSNR: 28.2752 SNR: 21.0377 IMSE: 96.7289 SSIM: 0.9705	28.3056	18.2766	96.0554	0.9743
		4		28.1441	18.8943	99.6953	0.9786
B (Sample PCM Image)	O-Net-Generated PCM	5	PSNR: 27.5383 SNR: 22.2778 IMSE: 114.6174 SSIM: 0.9895	27.3099	21.1366	120.8059	0.9898
		6		27.5839	20.6317	113.4213	0.9891
	Theta-Net-Generated PCM	7	PSNR: 26.5322 SNR: 21.2716 IMSE: 144.4989 SSIM: 0.9864	25.0905	18.9171	201.3876	0.9843
		8		26.4809	19.5287	146.2142	0.9857

Figure 9. Sample O-Net- and Θ -Net-generated images from (A) DIC and (B) PCM micrographs. For this figure, a random sample was selected and assayed, as a means of exercising stringency when evaluating potential mismatches between Θ -Net and ground truth (Expected) images. Here, we observe that the local SSIM maps for both O-Net and Θ -Net appear to be generally white, which is indicative of a high correlation between the respective DNN-generated images and the ground truth (Expected) images [the local SSIM maps are based on individual pixel mismatches within an 11-by-11 neighborhood [42] and range from black (0) to white (1)]. From the table, it may be observed that Θ -Net generally surpasses O-Net when increasing the resolution of DIC micrographs, while the reverse holds true for PCM images (at least for the assayed images in this instance). Nonetheless (for the PCM images), the image metrics (such as PSNR and SSIM) for Θ -Net seem to closely approach those of O-Net (differing by <1% for SSIM scores but a slightly higher margin of <9% for PSNR scores), implying that Θ -Net models might be more susceptible to *local* variations in individual pixel values and seeking to convert these into discernible features, thereby resulting in an imposed ‘penalty’ and reduced SSIM scores (as these metrics may interpret these features as noise). Further evidence of this is provided through the IMSE metric, which clearly shows a marked increase for the Θ -Net-generated images (when compared with those from O-Net) for the PCM images (Panel B).

3.4. Image Denoising

To evaluate the efficacy of the proposed models in *denoising* input micrographs (even though the assayed models were **not** specifically trained for this purpose), we infused an artificial representation of noise into the **Source** images via a salt-and-pepper noise

algorithm in MATLAB [43], with the image denoising consequently performed using the same trained models in Python. The results of this trial [including the corresponding **Source** and ground truth (**Expected**) images] are depicted in Figure 10 as follows.

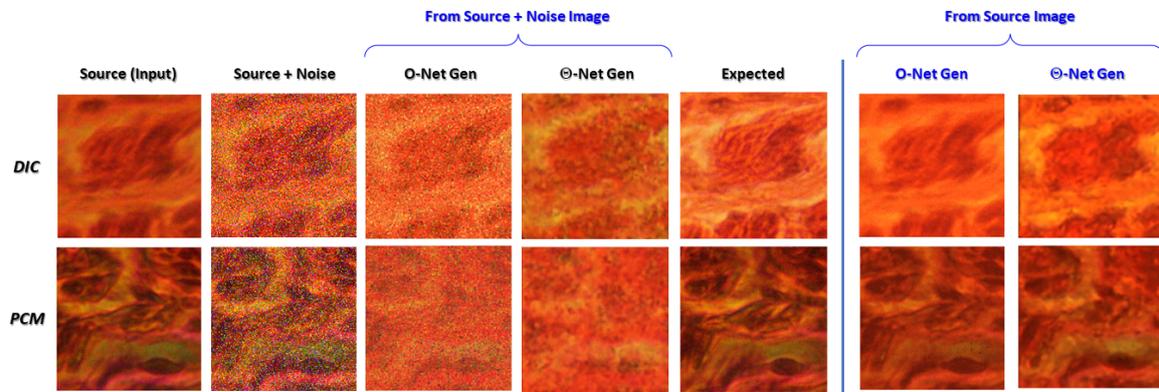


Figure 10. DIC and PCM photomicrographs infused with salt-and-pepper noise (noise density: 20%), which are labeled **Source + Noise**, and their corresponding ‘denoised’ images, as well as the ground truth (**Expected**) images utilized here. Here, we notice that Θ -Net models are relatively resilient to these noisy pixels, despite not being trained specifically to denoise images. This deduction is further corroborated by comparing these images against the Θ -Net-generated images in the absence of input noise (depicted within the violet ellipses). Nonetheless, it would be prudent to mention at this juncture that the Θ -Net models are still somewhat influenced by the noisy pixels in the image; hence, it would be preferred if the image denoising was performed *separately* prior to inputting the denoised images into Θ -Net for subsequent *in silico* ER.

3.5. Computational Complexity and Load

In addition to assessing the quality of ER images generated from each of the assayed architectures (in particular, Θ -Net), we also sought to compare the execution times for the afore-mentioned models, both as a means of quantifying their computational complexity and to determine the viable temporal resolution attainable by each of the models. The results following this analysis are presented in Table 1 below.

Table 1. Execution times (in seconds) indicated for each of the models trained under the assayed frameworks (i.e., O-Net and Θ -Net) in the present study. As Θ -Net comprises 3 nodes (each of which is an O-Net model), the implementation of Θ -Net for image post-processing is expected to take ~thrice as long as that of O-Net, although this is a non-linear relationship, being very much dependent on GPU capabilities.

Architecture	CPU	GPU	RAM	Storage	Average Execution Time
O-Net	Intel® Core™ i5-7200U CPU	NVIDIA GeForce® 920MX	20 GB	1 TB SSD	11.132 s (DIC)/11.463 s (PCM)
	2 x Intel® Xeon® Platinum 8170 (56C/112T)	NVIDIA Tesla K80	128 GB	500 GB SSD	2.437 s (DIC)/2.132 s (PCM)
	Intel® Core™ i9-10920X	NVIDIA GeForce RTX™ 3090	128 GB	512 GB SSD	0.422 s (DIC)/0.297 s (PCM)
Θ -Net	Intel® Core™ i5-7200U CPU	NVIDIA GeForce® 920MX	20 GB	1 TB SSD	51.756 s (DIC)/40.291 s (PCM)
	2 x Intel® Xeon® Platinum 8170 (56C/112T)	NVIDIA Tesla K80	128 GB	500 GB SSD	14.89 s (DIC)/19.613 s (PCM)
	Intel® Core™ i9-10920X	NVIDIA GeForce RTX™ 3090	128 GB	512 GB SSD	0.593 s (DIC)/0.703 s (PCM)

From the findings tabulated in Table 1, we may observe that the average execution times of Θ -Net (a DNN architecture comprising three O-Net nodes strung together in the

present context) expectably exceed those of O-Net, although this increase does not appear to be proportional to the node count of the Θ -Net framework used. Instead, the execution times seem to be highly dependent on the type of GPU used, ranging from a maximum multiplier of $\sim 9.2X$ (for the NVIDIA Tesla K80 GPU system) to a low of $\sim 1.4X$ (for the NVIDIA RTX 3090 GPU system). As GPUs are known to excel in parallel computationally intensive tasks, this might suggest that the individual O-Net nodes are being executed in a parallel (rather than a sequential) fashion.

4. Discussion

The results depicted in this study exhibit the significant potential of Θ -Net in attaining computational phase-modulated nanoscopy. Here, it may be observed that Θ -Net (as with O-Net in [29]) can enhance the resolution of both DIC and PCM micrographs while avoiding the formation of potential artifacts characteristic of these imaging modalities (namely the pseudo-relief effects in DIC microscopy [44] and the halo/shading-off effects of PCM [45]). In this regard, we surmise that the Θ -Net-based models likely conduct this (i.e., ER imaging) via a mapping function to reduce the PSF/OTF of the optical system (akin to that for O-Net as described in [29]), resulting in a hypothetical PSF which (when convolved with the ground truth representation) produces an ER image of the specimen in question (further details on this are expounded in [29] for the interested reader). However, Θ -Net conducts this through the *repeated functional mapping* of the acquired/generated PSF-convolved input image, which may be generally described by Equations (1) and (2) below:

Learning Phase

$$(f \otimes g) \xrightarrow[\underbrace{\dots}_{n \text{ times}}]{j_1 \dots j_n} (f_n \otimes g) \sim (h \otimes g) \quad (1)$$

where f_1 refers to the PSF of the optical system when using the 20X/0.4 Ph1 objective, g is the ground truth image of the specimen and h is the PSF of the optical system (when using the 40X/0.6 Ph2 objective). Here, f_n refers to the *generated* 'PSF' of the image when $(f \otimes g)$ is mapped under j_n [where $\lim_{n \rightarrow \infty} (f_n) = h$]. Like O-Net [29], the Θ -Net models thus attempt to learn the mapping function $j_{1..n}$ across n nodes (for which Θ -Net is defined) and (upon sufficiently learning this) deploy $j_{1..n}$ for ER images as follows:

Resolution Enhancement Phase

$$(f \otimes g) \xrightarrow[\underbrace{\dots}_{n \text{ times}}]{j_1 \dots j_n} \underbrace{(f_n \otimes g)}_{\substack{\text{enhanced resolution} \\ \text{image}}} \quad (2)$$

Nonetheless, a key aspect of our present study remains to highlight the feasibility of models adopting the Θ -Net framework over their predecessors employing the O-Net framework (as described in [29]) for *in silico* label-free ER microscopy—evidence for this being presented in Figures 3, 4, 6 and 7 (and their corresponding Supplementary Figures S11–14). As previously indicated, Θ -Net (in the present study) employs a *triple-node* architecture (each node being an O-Net) with a specialized *transfer-learned* O-Net model for the last node (the said node being trained on *both* the DIC and PCM datasets) (see Figure 2 for details). Here, the transfer learning process clearly aids in enhancing the ER capabilities of the Θ -Net models, putatively by allowing for a *transfer* of learnt phase–amplitude translations across two principally similar (yet methodologically different) microimaging modalities (i.e., DIC and PCM) to augment the trained models' abilities in detecting and translating these phase variations to the observable feature space. The empirical validation of this statement may be derived from Figures 4 and 7, where the granularity of some structural features (such as the surface of the cell wall) becomes distinctly visible upon processing with the Θ -Net models (as compared to the O-Net models), closely corroborating one's prior knowledge of such ultrastructural features (see [46] as an example for details).

Of particular importance to some would be the questions of why a different number of epochs was used for each of the O-Net nodes and what the significance of the chosen number of nodes in the Θ -Net architecture is. Here, it would be imperative to highlight that both of these factors were empirically determined for the acquired dataset—to minimize the model training duration while seeking to attain the optimal image accuracy (as described in Figures 4 and 7). In addition, one may also seek to question our choice of activation functions used in this study (namely Swish and GELU) and why more popular activation functions (such as ReLU or its leaky variant) were not utilized instead. To comprehend the rationale behind this, we would first need to take a step back and realize that elucidating the PSF of an optical system (such as a compound optical microscope) is a complex, non-linear problem, and the use of different activation functions (such as Swish and GELU) would potentially aid in mitigating some of the PSF variations introduced by sample inhomogeneities at different points in the specimen, thereby allowing the network to more closely approximate the convolution of the ground truth signal with the optical PSF at these different regions (as the light rays pass through the optical train of the microscope). Due to the highly complex nature of this problem [it is not possible to simply use a single formula to deduce the optical PSF and apply it via a deconvolution algorithm to the entire image (as is often the case for most computational deconvolution algorithms, whether blind or non-blind)], we postulate that the use of different activation functions for the discriminator and the generator networks (in particular, Swish and GELU which closely approximate the intensity distribution as introduced by the Airy function of the PSF) would allow for a convergence between these two networks into a more optimal solution. If only a single activation function (such as ReLU or Leaky ReLU) were used, these might not be able to sufficiently account for the spatiotemporal variations in the PSF, resulting in the production of output images which might not contain the required level of detail for SR images. It is, however, possible that only the use of ReLU or GELU might suffice (to a certain extent), although this approach was not validated in the present study.

From a computational perspective, the execution times of Θ -Net seem to be viable as well, generally not exceeding 1s in current advanced GPU systems, making it viable where processes having a temporal resolution exceeding 1s are generally encountered. Moreover, the loss function plots of the assayed Θ -Net models (for both DIC and PCM) suggest that the models were optimally trained, with the discriminator losses for both the real and generated samples (indicated by dR and dG, respectively) generally approaching 0 for all three nodes. The equations underpinning the computed losses [dR, dG and the generator loss (g)] may be mathematically defined as follows:

For the discriminator losses (i.e., both dR and dG), the **binary cross-entropy** loss (or log loss) ε was utilized, which may be expressed as below (from [29,47]):

$$\varepsilon = \frac{\sum_{i=1}^n [(b_i - 1) \log(1 - \hat{b}_i) - b_i \log \hat{b}_i]}{n} \quad (3)$$

where b_i is the label, and \hat{b}_i is the probability of $b_i = 1$ (derived from [47,48] and described in [29]).

In contrast, for computing the generator (g) loss, both ε and the **mean absolute error** (MAE)/ l_1 loss were used, the latter being computed as follows (from [29,49]):

$$\text{MAE } (l_1) \text{ loss} = \frac{\sum_{i=1}^n |a_i - b_i|}{n} \quad (4)$$

where n is the number of pixels in the image, and a_i and b_i refer to the target and the estimated values of the assayed parameter [e.g., pixel RGB (or HSL) intensities at pixel i , respectively].

Compounded together, the overall generator loss G may thus be expressed by the following equation (from [29]):

$$\text{Total generator loss, } G = \varepsilon + (\lambda \cdot l_1) \quad (5)$$

where $\lambda = 100$ [50] (to minimize potential artifacts, as described in [30]). ε and the MAE were utilized as viable loss functions for evaluating model training in this respect as they provide viable indications on how closely the generated images resembled the ground truth (**Expected**) image under differing representations (the MAE evaluates this from a linear perspective, while ε adopts a logarithmic standpoint).

On a separate note, the empirical verification of the image denoising capabilities of the Θ -Net models suggests that these models are relatively resilient to salt-and-pepper noise [43] present in the images, although there does exist a slight possibility of the models confounding this noise with sub-microscopic features which require a modulation transfer function (MTF) greater than that afforded by the microscope's optical train [51]. Evidence for this may be drawn from Figure 10, where some features encircled within the blue ellipse of the DIC micrograph depict an attempt by the said model to translate some of the noise into 'ER features' (a characteristic generally absent in traditional O-Net-based models). This finding thus portrays the need for Θ -Net users to denoise the input images separately, prior to increasing the resolution of them via the proposed models (to avoid introducing noise-based artifacts into the image). Despite this fact (and as previously already highlighted in the caption for Figure 10), the Θ -Net models perform significantly better in removing noisy pixels from the image (when compared to O-Net), although they were *not* specifically trained for this purpose (i.e., to conduct image denoising). This thus demonstrates a greater propensity of Θ -Net in image denoising applications, though potentially at the expense of incurring network hallucination artifacts.

Retrospectively, a further evaluation of the image quality metrics used (namely PSNR, SNR, IMSE and SSIM) seemingly suggests otherwise, with Θ -Net models generally surpassing O-Net model performance when increasing the resolution of DIC micrographs, although the opposite occurs for the PCM images. Nonetheless, this assessment proves contrary to the visual discrimination of the features present in the Θ -Net-generated PCM micrographs (when compared with their O-Net counterparts), casting doubt on the veracity of these metrics for quantifying the ER capabilities of a DNN algorithm developed for this purpose. This concern was also surfaced previously in [29], where a suggestion recommending the future development of a suitable metric for quantifying the "super-resolution" quality of an image was proposed, although this metric would probably be difficult to validate, as it would have to consider the illumination type and relative intensity of the light source (in addition to the specimen clarity and mounting procedure) being employed, amongst others. Moreover, Nyquist sampling criteria would also have to be satisfied [52] to alleviate potential network hallucination due to under-sampling operations which may result in sample noise being confounded as *pseudo*-noise [18].

It would also be noteworthy to mention that the 3D RCAN [26] models used for comparison against the evaluated O-Net and Θ -Net models in this study were developed based on an RGB image spectrally isolated into a 3-channel grayscale image stack within ImageJ 1.52n (NIH, USA), prior to training and executing the models on these images. This was because the 3D RCAN framework [26] was specifically developed to enhance the resolution of grayscale fluorescent micrographs in a 3D Z-stack, although our training images acquired were 2D 24-bit RGB images. The output image stack was subsequently re-merged in ImageJ into a single RGB image.

4.1. The Potential Advantages and Limitations of the Present Study

4.1.1. The Limitations of the Current Θ -Net Architecture

A prominent limitation in the current study (which is common to all DNN architectures) refers to the data source which was used to train the models. In the present context, our proposed Θ -Net models were trained on data acquired using a Leica DM4000M mi-

croscope with a RisingCam[®] CMOS camera. Should the models of the microscope and camera used to acquire the test images be different, the models would need to be retrained to compensate for the variations in the PSF experienced in the optical train in this respect. For this reason, readers who intend to use our supplied Θ -Net models for their own PCM and/or DIC photomicrographs would have to ensure that an identical microscope and camera setup is used for image acquisition (as highlighted in the present study), or a retraining of the models might be required.

On a separate note, we also *only* demonstrated the use of Θ -Net in increasing the resolution of PCM and DIC images—two highly popular phase-modulated optical microscopical approaches. Here, and as was also highlighted in [29], numerous other variants of phase microscopy techniques exist, notably Hoffman modulation contrast [53] and oblique illumination [54], as well as more recently developed (yet increasingly popular) quantitative phase imaging modalities [such as digital holographic microscopy (DHM) [55], amongst others]. As we did not verify the applicability of Θ -Net in increasing the resolution of images gleaned from each of these techniques, we are unable to hypothesize the putative performance of Θ -Net in this respect. It would also be noteworthy to highlight here that these techniques (i.e., PCM and DIC microscopy) are *semi-quantitative* approaches for representing phase variations in the sample and cannot be used to precisely *quantify* the phasic information due to the non-linear (and thus non-invertible) relationship between the phase and amplitude in different regions of the sample [56]. However, in our current study, we did not seek to specifically quantify these phase differences but simply use the said models to translate these phase variations in the sample (caused by the different PSFs in the optical train) into amplitude (brightness) differences for ER images.

4.1.2. The Advantages of the Current Θ -Net Architecture

As described in the captions for Figures 5 and 8 previously, the training of the Θ -Net models in the present study was performed *intermittently*, which resulted in the spikes observed in the loss function plots for the discriminator network (see Figures 5 and 8 for details). Here (and as was also demonstrated in [29,40]), intermittent model training generally results in more accurate models, visualized as the similarity between the generated (output) images and the ground truth (**Expected**) image datasets. This phenomenon may be postulated to be due to the minimization of the *trapping* of the loss function within local minima in the global error landscape, an aspect which we also surmised in [40].

Holistically, from the results gleaned and presented thus far, we may postulate that the proposed Θ -Net architecture poses significant usability in increasing the resolution of images acquired via PCM and DIC microscopy (two widely utilized label-free diascopic imaging techniques in the optical microscopical space today). In this respect, models developed using Θ -Net hold significant promise in future label-free optical nanoscopy applications, facilitating numerous potential extrapolations of engineering applications utilized in the industry today and moving towards the future (i.e., Industry 4.0 and beyond). Nonetheless, it would also be noteworthy at this juncture to address the perception that the Abbe diffraction limit *cannot* be surpassed through blind *in silico* approaches, as this information was effectively lost due to the physical limitations imposed by the imaging system. Proponents of this view may also suggest that a black-box model (such as a trained DNN) without any priors from the imaging physics or from the sample can only show the learnt features from the data (and nothing more). However, this perception may be challenged by the fact that there are numerous optical approaches (such as [57]) which demonstrated the circumvention of the Abbe limit, while our present study provides an alternative (non-optical) approach to achieving optical non-fluorescent nanoscopy (via phase-modulated *in silico* ER). All of these studies thus point to the fact that the information encoded by such high spatial frequencies is *not* entirely lost in the optical train (as might be perceived by these individuals), but it is simply *obscured* by other dominant (lower spatial frequency) impulse responses. Moreover, based on our current findings, we can somewhat

propose the ability of Θ -Net as a potential means of recovering some of this occluded high spatial frequency information, in a bid to achieve *in silico* ER so as to attain computational phase-modulated nanoscopy in the near future. This might thus be utilized to facilitate further developments in fields as diverse as *bioengineering* (for the synthesis of nano-scaffolds to direct proteins to specific sites for proper folding and functional deployment), *medical imaging* (for detecting and identifying nascent tumorigenic cells), *material sciences and electronics engineering* (for quality analysis and cost-friendly nanoscale defect inspection of semiconductor wafers and in nanolithography) and *optical engineering and photonics* applications (to facilitate the research and development of new optical modulators, lasers and diodes), where microscopes remain indispensable tools for both the diagnosis and quality inspection/analysis of parts (amongst others).

5. Conclusions

In the current study, the performance of our newly developed DNN architecture (Θ -Net) was assessed against other popular state-of-the-art DNN architectures (O-Net [29], 3D RCAN [26], ANNA-PALM [23] and BSRGAN [38]) with respect to its accuracy in increasing the resolution of PCM and DIC micrographs computationally. Our models depict a relatively high level of accuracy, generating images which come close to the ground truth (**Expected**) images. Notably, ER images generated by the Θ -Net models of untrained images (Figures 4 and 7, as well as Figures S12 and S14) closely agree with prior knowledge of cellular ultrastructure, exemplifying the usability of Θ -Net in this context. In this regard, one may be led to recommend Θ -Net-derived models for increasing the resolution of PCM and DIC images, although the other compared models (such as ANNA-PALM [23] and 3D RCAN [26]) might perform relatively well for other imaging modalities (such as epi-fluorescence microscopy stacks) (it would be noteworthy to mention here though that these models were not specifically trained for increasing the resolution of phase-modulated microscopical images, such as those acquired via DIC and PCM, as in the present study).

Despite this fact, there currently exists *no* viable image SR metric which may be used to objectively quantify the performance of a DNN model in resolving images. This dilemma is further exacerbated through the inappropriate use of popular image metrics (such as PSNR, SSIM and IMSE, amongst others) for this purpose in past studies exploring computational SR imaging. A probable resolution of this issue may thus lie in the development of such a metric specifically for ER/SR imaging, although this would have to consider multiple factors such as the effect of different illumination intensities, camera exposure and gain in the acquired image. A further treatment of this issue is hence presented in [29] for the interested reader.

Finally, we would also like to highlight that numerous recent cutting-edge research studies have demonstrated the potential of DNNs in various medical imaging approaches, ranging from ultrasound localization microscopy ([58]) to computer-aided detection ([59]) and MRI-CT ([60]), thereby providing an avenue for the potential applicability of our proposed Θ -Net architecture in a similar regard. Similarly, other far-ranging potential applications of Θ -Net may also be surmised in LiDAR-hyperspectral imaging and/or remote sensing applications (e.g., drone imaging), although the use of Θ -Net models for the latter might be subject to other potential constraints as well (as described at [61]) while also necessitating model retraining in this respect.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/s24196248/s1>, Figure S1: Diagram illustrating the model parameters for the O-Net GAN; Figure S2: Diagram illustrating the model parameters for the O-Net GAN discriminator (A) & generator (B) respectively; Figure S3: Diagram illustrating the model parameters for the 2nd node of the Θ -Net GAN; Figure S4: Diagram illustrating the model parameters for the 2nd node Θ -Net GAN discriminator (A) & generator (B) respectively; Figure S5: Diagram illustrating the model parameters for the 3rd node of the Θ -Net GAN; Figure S6: Diagram illustrating the model parameters for the 3rd node Θ -Net GAN discriminator (A) & generator (B) respectively; Figure S7:

Diagram illustrating the model parameters for the O-Net GAN; Figure S8: Diagram illustrating the model parameters for the O-Net GAN discriminator (A) & generator (B) respectively; Figure S9: Diagram illustrating the model parameters for the 2nd node of the Θ -Net GAN; Figure S10: Diagram illustrating the model parameters for the 2nd node Θ -Net GAN discriminator (A) & generator (B) respectively; Figure S11: Figure illustrating the model-generated images for the DIC training images dataset presented in this study; Figure S12: Figure illustrating the model-generated images for the DIC validation dataset presented in this study; Figure S13: Figure illustrating the model-generated images for the PCM training images dataset presented in this study; Figure S14: Figure illustrating the model-generated images for the PCM validation dataset presented in this study; Table S1: Equations considered for the activation functions; Table S2: Equations considered for the optimizer function. References [62–68] are cited in Supplementary Materials.

Author Contributions: All authors contributed to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study does not involve humans or animals.

Informed Consent Statement: This study does not involve humans.

Data Availability Statement: The data utilized for this study (figures, models and codes) may be downloaded at https://drive.google.com/file/d/1J0HYPE6-tOJOJ9F_fkifzomQ6oFxFJ-Mc/view?usp=sharing (accessed on 22 September 2024).

Acknowledgments: The authors would like to extend their sincere appreciation towards past studies conducted by other researchers in the field who laid some of the groundwork for attaining computational nanoscopy.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Chai, J.; Hao, Z.; Li, A.; Ngai, E.W.T. Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Mach. Learn. Appl.* **2021**, *6*, 100134. [[CrossRef](#)]
- Voulodimos, A.; Doulamis, N.; Doulamis, A.; Protopapadakis, E. Deep Learning for Computer Vision: A Brief Review. *Comput. Intell. Neurosci.* **2018**, *2018*, 7068349. [[CrossRef](#)] [[PubMed](#)]
- Chai, J.; Li, A. Deep Learning in Natural Language Processing: A State-of-the-Art Survey. In Proceedings of the International Conference on Machine Learning and Cybernetics (ICMLC), Kobe, Japan, 7–10 July 2019.
- Wang, Y.; Xiao, W.; Li, S. Offline Handwritten Text Recognition Using Deep Learning: A Review. *J. Phys. Conf. Ser.* **2021**, *1848*, 012015. [[CrossRef](#)]
- Aldhyani, T.H.H.; Alkahtani, H. Attacks to Automotous Vehicles: A Deep Learning Algorithm for Cybersecurity. *Sensors* **2022**, *22*, 360. [[CrossRef](#)] [[PubMed](#)]
- AbdulNabi, I.; Yaseen, Q. Spam Email Detection Using Deep Learning Techniques. *Procedia Comput. Sci.* **2021**, *184*, 853–858. [[CrossRef](#)]
- AMD Radeon™ RX Graphics Cards. Available online: <https://www.amd.com/en/graphics/radeon-rx-graphics> (accessed on 4 July 2021).
- Wu, Y.-C.; Chen, C.-S.; Chan, Y.-J. The outbreak of COVID-19: An overview. *J. Chin. Med. Assoc.* **2020**, *83*, 217–220. [[CrossRef](#)]
- World Meteorological Organization (WMO). United in Science 2022. Available online: <https://wmo.int/publication-series/united-science-2022> (accessed on 30 August 2024).
- He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016.
- Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Available online: <https://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a/> (accessed on 30 April 2019).
- Xie, Q.; Li, Y.; He, N.; Ning, M.; Ma, K.; Wang, G.; Lian, Y.; Zheng, Y. Unsupervised Domain Adaptation for Medical Image Segmentation by Disentanglement Learning and Self-Training. *IEEE Trans. Med. Imaging* **2024**, *43*, 4–14. [[CrossRef](#)]
- Jain, R.K.; Munn, L.L.; Fukumura, D. Dissecting tumour pathophysiology using intravital microscopy. *Nat. Rev. Cancer* **2002**, *2*, 266–276. [[CrossRef](#)]
- Vrieling, F.; Kostidis, S.; Spaink, H.P.; Haks, M.C.; Mayboroda, O.A.; Ottenhoff, T.H.M.; Joosten, S.A. Analyzing the impact of Mycobacterium tuberculosis infection on primary human macrophages by combined exploratory and targeted metabolomics. *Sci. Rep.* **2020**, *10*, 7085. [[CrossRef](#)]
- Castro-Moretti, F.; Gentzel, I.; Mackey, D.; Alonso, A. Metabolomics as an Emerging Tool for the Study of Plant–Pathogen Interactions. *Metabolites* **2020**, *10*, 52. [[CrossRef](#)]

16. Borowitzka, M.A. Chapter 3—Biology of Microalgae. In *Microalgae in Health and Disease Prevention*; Levine, I.A., Fleurence, J., Eds.; Elsevier Inc.: Amsterdam, The Netherlands, 2018; pp. 23–72.
17. Feng, Y.; Khoo, H.E.; Poh, C.L. Purification and Characterization of Gentisate 1,2-Dioxygenases from *Pseudomonas alcaligenes* NCIB 9867 and *Pseudomonas putida* NCIB 9869. *Appl. Environ. Microbiol.* **1999**, *65*, 946–950. [[CrossRef](#)] [[PubMed](#)]
18. Kaderuppan, S.S.; Wong, E.W.L.; Sharma, A.; Woo, W.L. Smart Nanoscopy: A Review of Computational Approaches to Achieve Super-Resolved Optical Microscopy. *IEEE Access* **2020**, *8*, 214801–214831. [[CrossRef](#)]
19. Langhorst, M.F.; Schaffer, J.; Goetze, B. Structure brings clarity: Structured illumination microscopy in cell biology. *Biotechnol. J.* **2009**, *4*, 858–865. [[CrossRef](#)]
20. Hell, S.W.; Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: Stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **1994**, *19*, 780–782. [[CrossRef](#)]
21. Betzig, E.; Patterson, G.H.; Sougrat, R.; Lindwasser, O.W.; Olenych, S.; Bonifacino, S.; Davidson, M.W.; Lippincott-Schwartz, J.; Hess, H.F. Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **2006**, *313*, 1642–1645. [[CrossRef](#)] [[PubMed](#)]
22. Rust, M.J.; Bates, M.; Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **2006**, *3*, 793–796. [[CrossRef](#)]
23. Ouyang, W.; Aristov, A.; Lelek, M.; Hao, X.; Zimmer, C. Deep learning massively accelerates super-resolution localization microscopy. *Nat. Biotechnol.* **2018**, *36*, 460–468. [[CrossRef](#)]
24. Nehme, E.; Weiss, L.E.; Michaeli, T.; Shechtman, Y. Deep-STORM: Super-resolution single-molecule microscopy by deep learning. *Optica*. **2018**, *5*, 458–464. [[CrossRef](#)]
25. Wu, Y.; Rivenson, Y.; Wang, H.; Luo, Y.; Ben-David, E.; Bentolila, L.A.; Pritz, C.; Ozcan, A. Three-dimensional virtual refocusing of fluorescence microscopy images using deep learning. *Nat. Methods* **2019**, *16*, 1323–1331. [[CrossRef](#)]
26. Chen, J.; Sasaki, H.; Lai, H.; Su, Y.; Liu, J.; Wu, Y.; Zhovmer, A.; Combs, C.A.; Rey-Suarez, I.; Chang, H.-Y.; et al. Three-dimensional residual channel attention networks denoise and sharpen fluorescence microscopy image volumes. *Nat. Methods* **2021**, *15*, 678–687. [[CrossRef](#)]
27. Chambers, W.; Fellers, T.J.; Davidson, M.W. Darkfield Illumination. Nikon Instruments Inc. Available online: <https://www.microscopyu.com/techniques/stereomicroscopy/darkfield-illumination> (accessed on 4 July 2021).
28. Robinson, P.C.; Davidson, M.W. Polarized Light Microscopy. Nikon Instruments Inc. Available online: <https://www.microscopyu.com/techniques/polarized-light/polarized-light-microscopy> (accessed on 4 July 2021).
29. Kaderuppan, S.S.; Wong, E.W.L.; Sharma, A.; Woo, W.L. O-Net: A Fast and Precise Deep-Learning Architecture for Computational Super-Resolved Phase-Modulated Optical Microscopy. *Microsc. Microanal.* **2022**, *28*, 1584–1598. [[CrossRef](#)]
30. Isola, P.; Zhu, J.-Y.; Zhu, T.; Efros, A.A. Image-to-Image Translation with Conditional Adversarial Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
31. Ramachandran, P.; Zoph, B.; Le, Q.V. Swish: A Self-Gated Activation Function. *arXiv* **2017**, arXiv:1710.05941v1.
32. Hendrycks, D.; Gimpel, K. Gaussian Error Linear Units (GELUs). *arXiv* **2020**, arXiv:1606.08415.
33. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
34. Kim, D.; Saito, K.; Oh, T.-H.; Plummer, B.A.; Sclaroff, S.; Saenko, K. CDS: Cross-Domain Self-supervised Pre-training. In Proceedings of the International Conference on Computer Vision, Montreal, QC, Canada, 11–17 October 2021; Available online: <https://ieeexplore.ieee.org/document/9710187> (accessed on 30 August 2024).
35. Murphy, D.B.; Hinsch, J.; Spring, K.R.; Davidson, M.W. Comparison of Phase Contrast and DIC Microscopy. © by Michael W. Davidson and The Florida State University. Available online: <https://micro.magnet.fsu.edu/primer/techniques/dic/dicphasecomparison.html> (accessed on 30 August 2024).
36. Deep Learning—Artificial Intelligence for Image Based Applications (Aivia). Leica Microsystems, Inc. Available online: <https://www.aivia-software.com/deep-learning-aivia> (accessed on 15 July 2021).
37. Ouyang, W. Anet Plugin for ImageJ. GitHub. Available online: <https://github.com/imodpasteur/Anet-ImageJ> (accessed on 22 November 2022).
38. Zhang, K.; Liang, J.; van Gool, L.; Timofte, R. Designing a Practical Degradation Model for Deep Blind Image Super-Resolution. In Proceedings of the International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021.
39. Belthangady, C.; Royer, L.A. Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. *Nat. Methods* **2019**, *16*, 1215–1225. [[CrossRef](#)]
40. Kaderuppan, S.S.; Wong, E.W.L.; Sharma, A.; Woo, W.L. Impact analysis of deep neural network training methodology on computational nanoscopy. In Proceedings of the Focus on Microscopy (FOM), Online, 10–13 April 2022.
41. Brownlee, J. How to Evaluate Generative Adversarial Networks. Machine Learning Mastery Pty. Ltd. Available online: <https://machinelearningmastery.com/how-to-evaluate-generative-adversarial-networks/> (accessed on 23 October 2022).
42. The MathWorks, Inc. Obtain Local Structural Similarity Index. Available online: <https://www.mathworks.com/help/images/obtain-local-structural-similarity-index.html> (accessed on 23 October 2022).
43. Add Noise to Image—MATLAB Imnoise. The MathWorks, Inc. Available online: <https://www.mathworks.com/help/images/ref/imnoise.html> (accessed on 30 November 2023).
44. Bagnell, C.R.J. Chapter 11—Differential Interference Contrast Microscopy. Available online: <https://www.med.unc.edu/microscopy/files/2018/06/lm-ch-11-dic.pdf> (accessed on 2 November 2019).

45. Murphy, D.B.; Oldfield, R.; Schwartz, S.; Davidson, M.W. Introduction to Phase Contrast Microscopy. Nikon—MicroscopyU. Available online: <https://www.microscopyu.com/techniques/phase-contrast/introduction-to-phase-contrast-microscopy> (accessed on 24 April 2019).
46. Plant Epidermis Cellulose Cell Walls, SEM—Stock Image—C032/4389—Science Photo Library. Science Photo Library (SPL). Available online: <https://www.sciencephoto.com/media/801280/view/plant-epidermis-cellulose-cell-walls-sem> (accessed on 1 December 2023).
47. Murphy, K. *Machine Learning: A Probabilistic Perspective*; The MIT Press: Cambridge, MA, USA; London, UK, 2012; 246p.
48. Godoy, D. Understanding Binary Cross-Entropy/Log Loss: A Visual Explanation. Available online: <https://towardsdatascience.com/understanding-binary-cross-entropy-log-loss-a-visual-explanation-a3ac6025181a> (accessed on 22 November 2018).
49. Shukla, R. L1 vs. L2 Loss Function. Available online: <http://rishy.github.io/ml/2015/07/28/l1-vs-l2-loss/> (accessed on 30 August 2024).
50. pix2pix: Image-to-Image Translation with a Conditional GAN. TensorFlow. Available online: <https://www.tensorflow.org/tutorials/generative/pix2pix> (accessed on 4 July 2021).
51. Introduction to Modulation Transfer Function. Edmund Optics Inc. Available online: <https://www.edmundoptics.com/knowledge-center/application-notes/optics/introduction-to-modulation-transfer-function/> (accessed on 4 July 2021).
52. Murphy, D.B.; Davidson, M.W. *Fundamentals of Light Microscopy and Electronic Imaging*, 2nd ed.; John Wiley & Sons Inc.: Hoboken, NJ, USA, 2013.
53. Hoffman, R.; Gross, L. Modulation Contrast Microscope. *Appl. Opt.* **1975**, *14*, 1169–1176. [CrossRef] [PubMed]
54. Chambers, W.; Fellers, T.J.; Davidson, M.W. Oblique Illumination. Nikon Instruments Inc. Available online: <https://www.microscopyu.com/techniques/stereomicroscopy/oblique-illumination> (accessed on 4 July 2021).
55. Meng, H.; Hongpeng, Q.; Zhuqing, J. Real-time quantitative phase imaging by single-shot dual-wavelength off-axis digital holographic microscopy. *Appl. Opt.* **2021**, *60*, 4418–4425.
56. Zuo, C.; Li, J.; Sun, J.; Fan, Y.; Zhang, J.; Lu, L.; Zhang, R.; Wang, B.; Huang, L.; Chen, Q. Transport of intensity equation: A tutorial. *Opt. Lasers Eng.* **2020**, *135*, 106187. [CrossRef]
57. Chen, L.; Zhou, Y.; Wu, M.; Hong, M. Remote-mode microsphere nano-imaging: New boundaries for optical microscopes. *Opto Electron. Adv.* **2018**, *1*, 17000101–17000107.
58. Xing, P.; Poree, J.; Rauby, B.; Malescot, A.; Martineau, E.; Perrot, V.; Rungta, R.L.; Provost, J. Phase Aberration Correction for In Vivo Ultrasound Localization Microscopy Using a Spatiotemporal Complex-Valued Neural Network. *IEEE Trans. Med. Imaging* **2024**, *43*, 662–673. [CrossRef]
59. Shin, H.C.; Roth, H.R.; Gao, M.; Lu, I.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; Summers, R.M. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Trans. Med. Imaging* **2016**, *35*, 1285–1298. [CrossRef] [PubMed]
60. Ozbey, M.; Dalmaz, O.; Dar, S.U.H.; Bedel, H.A.; Ozturk, S.; Gungor, A.; Cukur, T. Unsupervised Medical Image Translation with Adversarial Diffusion Models. *IEEE Trans. Med. Imaging* **2023**, *42*, 3524–3539. [CrossRef] [PubMed]
61. Zheng, Q.; Tian, X.; Yu, Z.; Ding, Y.; Elhanashi, A.; Saponara, S.; Kpalma, K. MobileRaT: A Lightweight Radio Transformer Method for Automatic Modulation Classification in Drone Communication Systems. *Drones* **2023**, *7*, 596. [CrossRef]
62. Keras. LeakyReLU Layer. Available online: https://keras.io/api/layers/activation_layers/leaky_relu/ (accessed on 29 October 2022).
63. Keras. Layer Activation Functions. Available online: <https://keras.io/api/layers/activations/> (accessed on 29 October 2022).
64. Peak Signal-to-Noise Ratio (PSNR). Available online: <https://www.mathworks.com/help/images/ref/psnr.html> (accessed on 29 October 2022).
65. Solutions, C.P. What is Signal to Noise Ratio and How to Calculate It? Available online: <https://resources.pcb.cadence.com/blog/2020-what-is-signal-to-noise-ratio-and-how-to-calculate-it> (accessed on 29 October 2022).
66. Mean-Squared Error. Available online: <https://www.mathworks.com/help/images/ref/immse.html> (accessed on 29 October 2022).
67. Structural Similarity (SSIM) Index for Measuring Image Quality. Available online: <https://www.mathworks.com/help/images/ref/ssim.html> (accessed on 29 October 2022).
68. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef] [PubMed]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.