

# Deep learning based brightfield image generation from a single hologram using unpaired dataset

DÁNIEL TERBE<sup>1,\*</sup>, LÁSZLÓ ORZÓ<sup>1</sup>, AND ÁKOS ZARÁNDY<sup>1</sup>

<sup>1</sup>*Institute for Computer Science and Control, Budapest H-1111, Hungary*

\**Corresponding author: terbed@sztki.hu*

Compiled October 28, 2021

We adopted an unpaired neural network training technique, namely CycleGAN, to generate brightfield microscope-like images from hologram reconstructions. The motivation for unpaired training in microscope applications is that the construction of paired/parallel datasets is cumbersome or sometimes not even feasible, for example, lensless or flow-through holographic measuring setups. Our results show that the proposed method is applicable in these cases and provides comparable results to the paired training. Furthermore, it has some favorable properties even though its metric scores are lower. The CycleGAN training results in sharper – from this point of view – more realistic object reconstructions compared to the baseline paired setting. Finally, we show that a lower metric score of the unpaired training does not necessarily imply a worse image generation, but a correct object synthesis yet with a different focal representation.

© 2021 Optical Society of America

<http://dx.doi.org/10.1364/ao.XX.XXXXXX>

## 1. INTRODUCTION

Recent development in the field of deep learning provided methods that were successfully applied in different imaging tasks in microscopy outperforming classical image processing algorithms. For example in coherent imaging, it was applied on phase recovery and hologram reconstruction [1–4], on phase-unwrapping [5, 6], on label-free sensing [7–9], on super-resolution [10], and even on transforming between coherent and incoherent imaging domains [11]. In the latter, the authors generate a brightfield microscope image from a single reconstructed hologram and thus combine the benefits of both domains. A brightfield image is sharper, visually more appealing and natural as our eyes are accustomed to incoherent images but its depth of field (DOF) is narrow (few  $\mu\text{m}$ ) while a hologram image encodes information of a larger volume (few hundred  $\mu\text{m}$ ) but the reconstructed hologram (at a given depth) is rich in artifacts

(caused by the characteristics of the coherent imaging and the noise of the in-line hologram reconstruction) which degrades the image quality. Wu et al. in [11] utilized the generative adversarial network (GAN) [12] training technique to learn the transformation between the two domains, after that they could generate a brightfield z-scan from a single hologram. In their training process, there is a generator network and a discriminator network that are trained together and compete with each other. The discriminator is trained to recognize generated (fake) and real images, while the generator is trained to synthesize images that can fool the discriminator. In addition to that, the generator network is conditioned with a reconstructed hologram and encouraged to generate the corresponding brightfield image minimizing the L1 loss between target and prediction. The adversarial training allows high-quality image synthesis while the L1 criterion constrains the generation so that the network won't create random brightfield images but that which corresponds to the given reconstructed hologram input. Thus, this method requires paired (and pixel-wise aligned) dataset. However, the collection of such a paired dataset is hard and sometimes it is not even possible. For example, in the case of a lensless holographic system [13], it is impossible to measure parallel holographic and brightfield images without significant modifications of the setup. Furthermore, in the case of holographic measuring setups that measure flow samples – where subtraction of the static background diffractions can considerably improve the quality of the acquired holograms – there is no way to implement the required parallel measurements. Unpaired training data, on the other hand, can be easily collected from different instruments or even from existing datasets bypassing the image alignment/registration issues. A recent technique called CycleGAN [14] enables us to train on unpaired dataset which has been already applied on many practical problems, for example virtual staining [15, 16], hologram reconstruction [17], and virtual brightfield and fluorescence staining of recovered fourier ptychographic microscopy (FPM) images [18]. Most recently, Zhang et al. proposed a physics-driven unpaired technique [19] for hologram phase retrieval and compared it with CycleGAN. In this study, our aim is to investigate the application of the CycleGAN technique for "brightfield holography", where a brightfield image is generated from a reconstructed hologram similarly to [11] but without the use of aligned training data. We also implement the method proposed in [11] which requires parallel dataset, and compare its results with the label-free training

method.

## 2. METHODS

The training method is illustrated in Fig. 1. Assume that we have unpaired samples from two image domain:  $A$  denotes the holographic domain and  $B$  denotes the brightfield domain. Our objective is to learn the  $A$  to  $B$  transformation:  $G_{AB} : A \rightarrow B$ . The process includes four neural networks: two generators ( $G_{AB}$  for  $A \rightarrow B$  and  $G_{BA}$  for  $B \rightarrow A$  transformation) and two discriminators ( $D_A, D_B$ ), but for inference only one generator ( $G_{AB}$ ) is used. This means that we have to train an additional generator network that transforms from the brightfield to the holographic domain even if our goal is only to transform from holographic to brightfield domain. The loss functions to be minimized for the generators ( $\mathcal{L}\{G_{AB}\}, \mathcal{L}\{G_{BA}\}$ ) and discriminators ( $\mathcal{L}\{D_A\}, \mathcal{L}\{D_B\}$ ) are the following respectively:

$$\mathcal{L}\{G_{AB}\} = \mathcal{L}_{adv}^{(AB)} + \lambda_A \mathcal{L}_{cyc}^{(ABA)} \quad (1)$$

where  $\mathcal{L}_{adv}$  denotes the adversarial loss (train the generator to fool the discriminator):

$$\mathcal{L}_{adv}^{(AB)} = \frac{1}{N} \sum_{x_A \sim A} \underbrace{[1 - D_B(\overbrace{G_{AB}(x_A)}^{\hat{x}_B})]^2}_{\text{fake sample} \rightarrow 1}$$

and  $\mathcal{L}_{cyc}$  denotes the cycle consistency loss:

$$\mathcal{L}_{cyc}^{(ABA)} = \frac{1}{N} \sum_{x_A \sim A} | \overbrace{G_{BA}(G_{AB}(x_A))}^{\tilde{x}_A} - x_A |$$

where  $N$  is the number of training samples,  $\hat{x}_B$  denotes the generated fake sample,  $\tilde{x}_A$  denotes the reconstructed sample, and  $\lambda_A$  is the weight parameter for the cycle consistency loss. And finally, the discriminator loss:

$$\mathcal{L}\{D_B\} = \underbrace{\frac{1}{N} \sum_{x_B \sim B} [1 - D_B(x_B)]^2}_{\text{real sample} \rightarrow 1} + \underbrace{\frac{1}{N} \sum_{x_A \sim A} D_B(\overbrace{G_{AB}(x_A)}^{\hat{x}_B})^2}_{\text{fake sample} \rightarrow 0} \quad (2)$$

The first part of this equation trains the discriminator to output 1 when the input is real, and the second part trains it to output 0 when the image is fake – note that the latter is the opposite of  $\mathcal{L}_{adv}$ . The losses in the reverse direction ( $\mathcal{L}\{G_{BA}\}$  and  $\mathcal{L}\{D_A\}$ ) can be similarly constructed. There is also an optional identity constrain for the generators which effect on the learning is also investigated in this study:

$$\mathcal{L}_{id}\{G_{AB}\} = \lambda_{id} \frac{1}{N} \sum_{x_B \sim B} |G_{AB}(x_B) - x_B| \quad (3)$$

and similarly for  $G_{BA}$ .  $\lambda_{id}$  is the weight parameter.

We used the Adam optimizer and the following hyperparameter settings:  $\lambda_A = \lambda_B = 10$ ,  $\lambda_{id} = 0.5$  or  $0$ ,  $lr = 0.0002$  (learning rate).

## 3. RESULTS AND DISCUSSION

To evaluate the different methods we utilized the frequently applied objective structural similarity index measure (SSIM) [20], root mean squared error (RMSE), and subjective human perceptual assessments too. The scores of the objective metrics are shown in Tab. 1 and few real and fake sample pairs are depicted in Fig. 2 along with the corresponding SSIM score<sup>1</sup>. We examined several UNet architectures with different layer depths for the generator networks: unet32 has 5 downsampling and upsampling layers (which halves and doubles the spatial dimensions respectively), unet64 has 6, and unet256 has 8. The baseline method [11] with labelled training used the unet32 architecture for the generator. For the discriminator networks we used the PatchGAN [21] architecture. We also inspected several training considerations: training with or without identity constraint (denoted with "widt" and "woidt" respectively); training with or without hologram phase information (if trained with phase it is denoted with "wang"); training with multi-scale structural similarity [22] as the distance measure in the cycle loss functions instead of L1 (denoted with "mssim").

We used a paired dataset to be able to measure the accuracy of the unpaired training and the same time, to be able to compare it with the results of the paired training method. Of course, the paired property of the dataset was not exploited during the unpaired training. The training dataset contained approximately 3000 samples while the test set around 300 samples. The specimens are centrifuged and recorded in a plane, but due to the field curvature of the optics and due to the automatic focusing mechanism, our dataset contains not only in-focus but also slightly defocused objects<sup>2</sup> in both domain correspondingly. The methods are implemented using the PyTorch deep learning framework based on [14] and the training was run on an Nvidia RTX 2080-Ti graphics card (GPU). We let the training run for 200 epochs which took around 16 hours on the mentioned GPU, but the generated image quality was already satisfactory around epoch number 100.

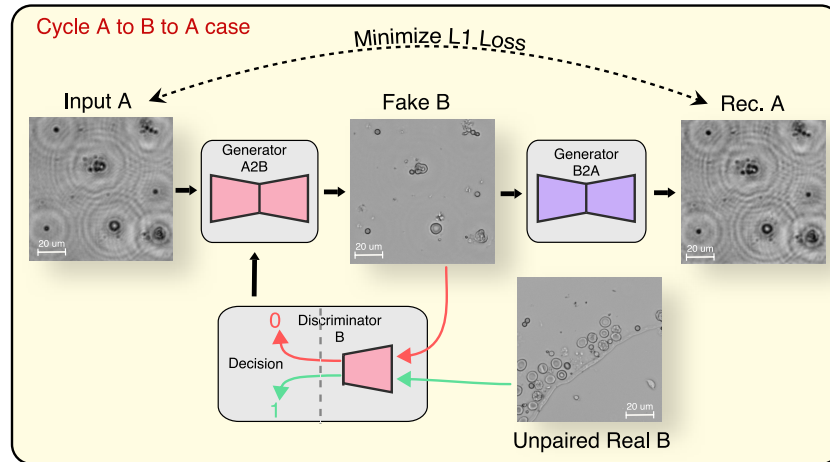
**Table 1.** Overall metric scores calculated on the test set for different models.

model	mode	RMSE	SSIM
UNET32-wang	paired	0.0623 ± 0.0242	0.741 ± 0.126
UNET32-woidt	unpaired	0.1136 ± 0.0348	<b>0.502 ± 0.133</b>
UNET32-wang	unpaired	<b>0.1072 ± 0.0332</b>	0.492 ± 0.134
UNET32-widt	unpaired	0.1121 ± 0.0376	0.475 ± 0.134
UNET32-mssim	unpaired	0.1275 ± 0.0384	0.461 ± 0.142
UNET64-widt	unpaired	0.1156 ± 0.0369	0.495 ± 0.135
UNET256-widt	unpaired	0.1248 ± 0.0377	0.473 ± 0.139

The unpaired training results were very similar with the different architecture sizes and training techniques (see Tab. 1 and samples in Fig. S2 and Fig. S3), therefore our choice of architecture is the smallest unet32. The best unpaired models based on the results in Tab. 1 are **UNET32-woidt** where we use hologram

<sup>1</sup>For more samples and a more detailed evaluation (utilizing more metrics and boxplots) we refer to the supplementary material.

<sup>2</sup>Note that we use the "slight defocus" term for that the objects are recorded at variable cross-sectional planes and not as a synonym for blur.



**Fig. 1.** Illustration of cycle consistent adversarial training in one direction. The generator is trained to fool the discriminator (and minimize the cycle constraint) while the discriminator is trained to distinguish real and fake samples.

TEST EXAMPLES

#	Paired		Real	Unpaired		Width (um)
	SSIM	Fake		Fake	SSIM	
(1)	0.75	(a)	(b)	(c)	0.34	10.6
(2)	0.83				0.44	11.8
(3)	0.86				0.28	10.6
(4)	0.91				0.48	11.4
(5)	0.22				0.66	14.2
(6)	0.67				0.25	13.4
(7)	0.61				0.61	9.8
(8)	0.90				0.47	6.6

**Fig. 2.** Test samples and corresponding SSIM values. The paired training achieves better SSIM scores but the generated images are more blurry. The generator with unpaired training synthesizes sharp brightfield images but with random focal representation.

amplitude data only and no identity constraint, and **unet32-wang** where we utilize hologram phase information too (as a two-channel input for the brightfield generator). Despite the slight variation in the metric values the generation quality perceptually was very similar (samples can be found in Fig. S2). We conclude that adding the hologram phase information does not considerably improve the model performance and that the identity constraint has only a minor degrading effect on learning. As we can see in Tab. 1 the paired training method outperforms the unpaired technique in objective metric scores. This is not surprising as the paired setup includes more information: we can directly show to the network what is the desired output

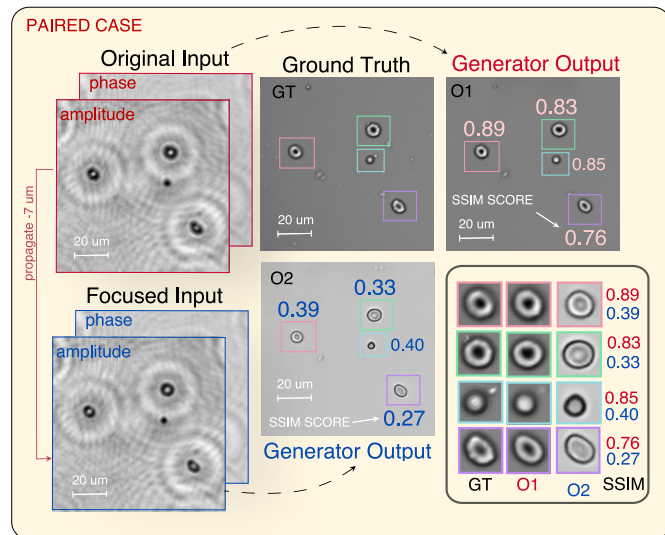
while in the unpaired setting it is indirect. In spite of all these, we claim that the generator neural network with unpaired training has some advantages over the other and that sometimes a lower metric score does not indicate per se inferior quality of the object synthesis. Primarily, it generates sharper images while the other one is more blurred (see Fig. 2 samples 1, 5, 7 and cut-line curves in Fig. S4). This blurriness may originate from the fact that classic paired training is known to average (blur) all possible solutions [23], while the unpaired method draws a unique sample from the learned distribution. Furthermore, as the SSIM scores and generated samples in Fig. 2 suggest, the synthesized images in the unpaired case may deviate from the real one, yet it is not always an erroneous discrepancy but may arise from that the generated image might have a different focal representation of the same object. In a brightfield image the properly focused objects usually appears to be gray (for example samples 1b, 2b, 4b in Fig. 2), while the slightly defocused ones shows up a bit brighter or darker (eg. samples 3b and 7b in Fig. 2, accordingly). In the paired setting the generator could mimic the imprecise, slightly defocused label (see sample 3a-b in Fig. 2) because of the direct L1 loss between the output and target<sup>3</sup> – which results in high SSIM score. On the other hand, in the unpaired setting the output brightfield image focus does not corresponds to the input hologram focus, but a slightly refocused object is synthesized, see samples 3, 4, 8 in Fig. 2. This discrepancy can be eliminated if we have only (or mainly) in-focus objects in our dataset.

The degradation of SSIM score in different focus representations is illustrated in Fig. 3. In this example we picked a defocused sample-pair from our dataset. The generator with paired training synthesizes the corresponding defocused brightfield image which results in a high SSIM score. It can produce high SSIM score, as the reconstruction and the reference correspond to each other, but both of them are out of focus. By adjusting the reconstruction distance of the input hologram, we can improve the focus of the hologram reconstruction and using it we can generate an in-focus brightfield image. Although, in this case the quality of the generated image considerably improves, its difference from the reference increases and the corresponding SSIM score drops. This phenomenon can be observed in Fig. 2 sample 3 where 3b and 3a is both out-of-focus (resulting in high

<sup>3</sup>and because the focus of the reconstructed holograms and the brightfield images are related in our dataset

SSIM score) while 3c is in-focus – but resulting in significantly lower SSIM score.

We also mention the slightly increased number of distorted predictions (when the synthesized objects differs from the real ones, e.g. Fig. 2. sample 6) in the unpaired setup that is caused by the looser constraints. There was also a decrease in the performance of the unpaired setup in case of dense samples when many objects were near to each other. In this case, the hologram reconstruction is more intermingled which makes the cycle transformations indeterminate.



**Fig. 3.** This figure illustrates that metric scores are not always a reliable measure of quality and that the generator in the paired training setup learned to generate a slightly defocused brightfield image if the input reconstructed hologram was also defocused. If we generate in-focus brightfield image from the same object the corresponding SSIM score will be smaller – even though the result is better.

Note that the byproduct of the unpaired training is an additional generator which transforms from the brightfield domain to the hologram domain. This generator synthesizes hologram like images, but not a true hologram in the sense that it can not be propagated – which is not an issue for us as we have interest only in the transformation regarding the opposite direction.

#### 4. CONCLUSIONS

This study investigates the application of unpaired training technique (CycleGAN) for "brightfield holography" which aims to generate brightfield microscope-like images from single hologram reconstructions; furthermore, it compares the aforementioned method with a paired training technique [11]. The motivation for unpaired training is that it is much easier to collect unpaired data, and in some cases, it is infeasible to create a paired one. We tested several models and found the unet32 architecture to be ideal with L1 distance, and that including the hologram phase information or identity constraint has only a minor effect. The examinations show that: (1) the synthesized samples are sharper in the unpaired case; (2) the unpaired generator tends to output brightfield objects with random focal representation; (3) this results in the drop of the SSIM score. We conclude in this study that it is feasible to apply an unpaired training method for "brightfield holography".

**Acknowledgments.** Supported by the Ministry of Innovation and Technology NRDI Office, Hungary within the framework of the Artificial Intelligence National Laboratory Program.

**Disclosures.** The authors declare no conflicts of interest.

**Data availability.** Data underlying the results presented in this paper are not publicly available (may be obtained from the authors).

**Supplemental document.** See Supplement 1 for supporting content.

#### 5. REFERENCES

##### REFERENCES

1. A. Goy, K. Arthur, S. Li, and G. Barbastathis, *Phys. review letters* **121**, 243902 (2018).
2. Y. Rivenson, Y. Zhang, H. Günaydin, D. Teng, and A. Ozcan, *Light. Sci. & Appl.* **7**, 17141 (2018).
3. A. Sinha, J. Lee, S. Li, and G. Barbastathis, *Optica* **4**, 1117 (2017).
4. Y. Wu, Y. Rivenson, Y. Zhang, Z. Wei, H. Günaydin, X. Lin, and A. Ozcan, *Optica* **5**, 704 (2018).
5. G. Dardikman and N. T. Shaked, "Phase unwrapping using residual neural networks," in *Computational Optical Sensing and Imaging*, (Optical Society of America, 2018), pp. CW3B–5.
6. G. Spoorthi, S. Gorthi, and R. K. S. S. Gorthi, *IEEE Signal Process. Lett.* **26**, 54 (2018).
7. Y. Wu, A. Ray, Q. Wei, A. Feizi, X. Tong, E. Chen, Y. Luo, and A. Ozcan, *ACS Photonics* **6**, 294 (2018).
8. Y. Wu, A. Calis, Y. Luo, C. Chen, M. Lutton, Y. Rivenson, X. Lin, H. C. Koydemir, Y. Zhang, H. Wang *et al.*, *ACS Photonics* **5**, 4617 (2018).
9. Z. Göröcs, M. Tamamitsu, V. Bianco, P. Wolf, S. Roy, K. Shindo, K. Yanny, Y. Wu, H. C. Koydemir, Y. Rivenson *et al.*, *Light. Sci. & Appl.* **7**, 1 (2018).
10. T. Liu, K. De Haan, Y. Rivenson, Z. Wei, X. Zeng, Y. Zhang, and A. Ozcan, *Sci. reports* **9**, 1 (2019).
11. Y. Wu, Y. Luo, G. Chaudhari, Y. Rivenson, A. Calis, K. De Haan, and A. Ozcan, *Light. Sci. & Appl.* **8**, 1 (2019).
12. A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, *IEEE Signal Process. Mag.* **35**, 53 (2018).
13. Y. Wu and A. Ozcan, *Methods* **136**, 4 (2018).
14. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, (2017), pp. 2223–2232.
15. Z. Xu, C. F. Moro, B. Bozóky, and Q. Zhang, *arXiv preprint arXiv:1901.04059* (2019).
16. G. Lee, J.-W. Oh, N.-G. Her, and W.-K. Jeong, *Med. image analysis* **70**, 101995 (2021).
17. D. Yin, Z. Gu, Y. Zhang, F. Gu, S. Nie, J. Ma, and C. Yuan, *IEEE Photonics J.* **12**, 1 (2019).
18. R. Wang, P. Song, S. Jiang, C. Yan, J. Zhu, C. Guo, Z. Bian, T. Wang, and G. Zheng, *Opt. Lett.* **45**, 5405 (2020).
19. Y. Zhang, M. A. Noack, P. Vagovic, K. Fezzaa, F. Garcia-Moreno, T. Ritschel, and P. Villanueva-Perez, *Opt. Express* **29**, 19593 (2021).
20. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, *IEEE transactions on image processing* **13**, 600 (2004).
21. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2017), pp. 1125–1134.
22. Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, vol. 2 (Ieee, 2003), pp. 1398–1402.
23. C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2017), pp. 4681–4690.

## FULL REFERENCES

1. A. Goy, K. Arthur, S. Li, and G. Barbastathis, "Low photon count phase retrieval using deep learning," *Phys. review letters* **121**, 243902 (2018).
2. 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**, 17141–17141 (2018).
3. A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica* **4**, 1117–1125 (2017).
4. Y. Wu, Y. Rivenson, Y. Zhang, Z. Wei, H. Günaydin, X. Lin, and A. Ozcan, "Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery," *Optica* **5**, 704–710 (2018).
5. G. Dardikman and N. T. Shaked, "Phase unwrapping using residual neural networks," in *Computational Optical Sensing and Imaging*, (Optical Society of America, 2018), pp. CW3B–5.
6. G. Spoorthi, S. Gorthi, and R. K. S. S. Gorthi, "Phasenet: A deep convolutional neural network for two-dimensional phase unwrapping," *IEEE Signal Process. Lett.* **26**, 54–58 (2018).
7. Y. Wu, A. Ray, Q. Wei, A. Feizi, X. Tong, E. Chen, Y. Luo, and A. Ozcan, "Deep learning enables high-throughput analysis of particle-aggregation-based biosensors imaged using holography," *ACS Photonics* **6**, 294–301 (2018).
8. Y. Wu, A. Calis, Y. Luo, C. Chen, M. Lutton, Y. Rivenson, X. Lin, H. C. Koydemir, Y. Zhang, H. Wang *et al.*, "Label-free bioaerosol sensing using mobile microscopy and deep learning," *ACS Photonics* **5**, 4617–4627 (2018).
9. Z. Göröcs, M. Tamamitsu, V. Bianco, P. Wolf, S. Roy, K. Shindo, K. Yanny, Y. Wu, H. C. Koydemir, Y. Rivenson *et al.*, "A deep learning-enabled portable imaging flow cytometer for cost-effective, high-throughput, and label-free analysis of natural water samples," *Light. Sci. & Appl.* **7**, 1–12 (2018).
10. T. Liu, K. De Haan, Y. Rivenson, Z. Wei, X. Zeng, Y. Zhang, and A. Ozcan, "Deep learning-based super-resolution in coherent imaging systems," *Sci. reports* **9**, 1–13 (2019).
11. Y. Wu, Y. Luo, G. Chaudhari, Y. Rivenson, A. Calis, K. De Haan, and A. Ozcan, "Bright-field holography: cross-modality deep learning enables snapshot 3d imaging with bright-field contrast using a single hologram," *Light. Sci. & Appl.* **8**, 1–7 (2019).
12. A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: An overview," *IEEE Signal Process. Mag.* **35**, 53–65 (2018).
13. Y. Wu and A. Ozcan, "Lensless digital holographic microscopy and its applications in biomedicine and environmental monitoring," *Methods* **136**, 4–16 (2018).
14. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, (2017), pp. 2223–2232.
15. Z. Xu, C. F. Moro, B. Bozóky, and Q. Zhang, "Gan-based virtual restaining: a promising solution for whole slide image analysis," *arXiv preprint arXiv:1901.04059* (2019).
16. G. Lee, J.-W. Oh, N.-G. Her, and W.-K. Jeong, "Deepfcs++: Bright-field to fluorescence microscopy image conversion using multi-task learning with adversarial losses for label-free high-content screening," *Med. image analysis* **70**, 101995 (2021).
17. D. Yin, Z. Gu, Y. Zhang, F. Gu, S. Nie, J. Ma, and C. Yuan, "Digital holographic reconstruction based on deep learning framework with unpaired data," *IEEE Photonics J.* **12**, 1–12 (2019).
18. R. Wang, P. Song, S. Jiang, C. Yan, J. Zhu, C. Guo, Z. Bian, T. Wang, and G. Zheng, "Virtual brightfield and fluorescence staining for fourier ptychography via unsupervised deep learning," *Opt. Lett.* **45**, 5405–5408 (2020).
19. Y. Zhang, M. A. Noack, P. Vagovic, K. Fezzaa, F. Garcia-Moreno, T. Ritschel, and P. Villanueva-Perez, "Phasegan: a deep-learning phase-retrieval approach for unpaired datasets," *Opt. Express* **29**, 19593–19604 (2021).
20. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing* **13**, 600–612 (2004).
21. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2017), pp. 1125–1134.
22. Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, vol. 2 (Ieee, 2003), pp. 1398–1402.
23. C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2017), pp. 4681–4690.