

# Reusability report: Unpaired deep-learning approaches for holographic image reconstruction

Received: 11 April 2023

Accepted: 13 January 2024

Published online: 15 February 2024

 Check for updates

Yuhe Zhang<sup>1</sup>, Tobias Ritschel<sup>2</sup> & Pablo Villanueva-Perez<sup>1</sup>

Deep-learning methods using unpaired datasets hold great potential for image reconstruction, especially in biomedical imaging where obtaining paired datasets is often difficult due to practical concerns. A recent study by Lee et al. (*Nature Machine Intelligence* 2023) has introduced a parameterized physical model (referred to as FMGAN) using the unpaired approach for adaptive holographic imaging, which replaces the forward generator network with a physical model parameterized on the propagation distance of the probing light. FMGAN has demonstrated its capability to reconstruct the complex phase and amplitude of objects, as well as the propagation distance, even in scenarios where the object-to-sensor distance exceeds the range of the training data. We performed additional experiments to comprehensively assess FMGAN's capabilities and limitations. As in the original paper, we compared FMGAN to two state-of-the-art unpaired methods, CycleGAN and PhaseGAN, and evaluated their robustness and adaptability under diverse conditions. Our findings highlight FMGAN's reproducibility and generalizability when dealing with both in-distribution and out-of-distribution data, corroborating the results reported by the original authors. We also extended FMGAN with explicit forward models describing the response of specific optical systems, which improved performance when dealing with non-perfect systems. However, we observed that FMGAN encounters difficulties when explicit forward models are unavailable. In such scenarios, PhaseGAN outperformed FMGAN.

Deep-learning approaches based on unpaired datasets have been demonstrated to be effective in various image reconstruction tasks, including biomedical imaging, where collecting paired datasets can be challenging or impossible due to practical considerations, such as dose. The key idea behind these approaches is to learn the underlying mapping between two different domains of images, such as between object and detector domains. This is usually implemented by learning a cyclic translation between the two domains using two generators. The forward generator learns the mapping from the first domain to

the second, and the backward generator learns the mapping in the reverse direction to enforce the cyclic consistency of the reconstructed images. Recently, Lee et al. have proposed a physically parameterized and unpaired approach for adaptive holographic imaging<sup>1</sup>, which replaces the forward generator with a forward physical model parameterized with the propagation distance of the probing light. For brevity, we refer to the parameterized physical forward model proposed by Lee et al. as FMGAN. The inclusion of a physical model stabilizes the learning process<sup>2</sup> for in-distribution (ID) data and increases tolerance

<sup>1</sup>Synchrotron Radiation Research and NanoLund, Lund University, Lund, Sweden. <sup>2</sup>University College London, London, UK. ✉e-mail: [yuhe.zhang@sljus.lu.se](mailto:yuhe.zhang@sljus.lu.se); [pablo.villanueva\\_perez@sljus.lu.se](mailto:pablo.villanueva_perez@sljus.lu.se)

**Table 1 | Experimental results for the reproducibility tests (tested on 44 images)**

Reconstruction results <sup>a</sup>		CycleGAN		PhaseGAN		PhaseGAN <sup>†</sup>		FMGAN		
Amplitude	FSIM	ID	0.84	(0.94×)	0.88	(0.99×)	0.87	(0.98×)	0.89	(1.00×)
		OOD	0.81	(0.95×)	0.79	(0.93×)	0.83	(0.97×)	0.85	(1.00×)
	PCC	ID	0.19	(0.28×)	0.67	(0.97×)	0.60	(0.87×)	0.69	(1.00×)
		OOD	0.24	(0.34×)	0.44	(0.61×)	0.45	(0.62×)	0.72	(1.00×)
Phase	FSIM	ID	0.62	(0.66×)	0.93	(1.00×)	0.91	(0.97×)	0.94	(1.00×)
		OOD	0.70	(0.74×)	0.85	(0.91×)	0.83	(0.88×)	0.94	(1.00×)
	PCC	ID	0.26	(0.29×)	0.88	(1.01×)	0.84	(0.95×)	0.88	(1.00×)
		OOD	0.36	(0.40×)	0.73	(0.81×)	0.67	(0.74×)	0.90	(1.00×)
Distance	MAE	ID	–	–	–	–	–	–	0.3	±0.2
		OOD	–	–	–	–	–	–	0.3	±0.2

ID, in-distribution; OOD, out-of-distribution. <sup>a</sup>The mean values of the FSIM and PCC between the ground truth and the predicted complex images are provided, along with the performance ratios between CycleGAN, PhaseGAN and PhaseGAN<sup>†</sup> relative to FMGAN. For FMGAN, the mean and standard deviation of the MAE for the predicted propagation distances are also reported.

for handling out-of-distribution (OOD) data because of the extension of the mapping space.

In this report, we study the reusability of FMGAN, focusing on its robustness and adaptability under various conditions. First, we evaluate the reproducibility of the results reported in ref. 1, including a comparison of FMGAN with CycleGAN<sup>3,4</sup> and PhaseGAN<sup>2</sup>. Second, we study the performance of the three approaches for non-perfect optical systems, for example, with the presence of Poisson noise or under the effect of out-of-focus blurring for an untilted and a tilted detector plane. The blurring effect is simulated by a point spread function (PSF), a two-dimensional function that describes the response of an optical system to a point-light source. We finally consider the adaptability limit of FMGAN to a wide range of propagation distances.

## Reproducing the reported results

We study the reproducibility of the results on both in-distribution data and OOD data using the 3  $\mu\text{m}$  polystyrene microsphere dataset shared by Lee et al.<sup>5</sup> Note that additional datasets, including a red blood cell dataset and a histological slide dataset, are also available in ref. 5, but these datasets were not used in the current study. First, we studied the performance of the methods on in-distribution data, where the networks were trained on 600 diffraction intensities measured at propagation distances 7 to 17 mm in steps of 2 mm (100 images for each distance) and 300 complex field patches as ground truth. For testing, 44 diffraction intensity patches measured at propagation distances in the same range with 1 mm steps were used. Second, we reproduced the adaptability study of the methods on OOD data, where the networks were trained on 100 diffraction intensities measured at a fixed 13 mm propagation distance and tested on diffraction intensity patches measured at 7 to 17 mm distances with 1 mm spacing. For training each dataset, we used three networks: CycleGAN, PhaseGAN and FMGAN. We used U-Net as in ref. 2 for the networks of CycleGAN and PhaseGAN, and squeeze-and-excitation U-Net<sup>6</sup> as in ref. 1 for the network of FMGAN. The ADAM optimizer<sup>7</sup> with a mini-batch size of eight was used throughout the training. For CycleGAN and PhaseGAN, we set the learning rates to be 0.0002 for both generators and discriminators and decayed the learning rate by 0.5 for every 1,000 iterations. The cycle-consistency weights were set to 200. For PhaseGAN, we added Fourier ring correlation (FRC) losses and structural similarity index measure (SSIM) losses to better constrain the model<sup>1,2</sup>, with  $\lambda_{\text{FRC}} = 100$  and  $\lambda_{\text{SSIM}} = 0.5$ . Incorporating the FRC loss has been demonstrated to improve the reconstruction quality and enhance the training performance<sup>2,8</sup>. We also reported the results of PhaseGAN trained without an FRC loss, which is denoted by PhaseGAN<sup>†</sup>. For the training of FMGAN, we set the learning rates for the generator and the discriminator to be 0.001 and

0.0001, respectively. The learning rates reported in ref. 1 appear to be incorrect as the learning rate of a generator cannot be much lower than that of the discriminator. The learning rate was reduced by 0.95 for every 500 iterations. We set  $\lambda_{\text{SSIM}}$  to ten and gradient penalty loss weight  $\lambda_{\text{GP}}$  to 20. The rest of the parameters were identical to the ones reported in ref. 1.

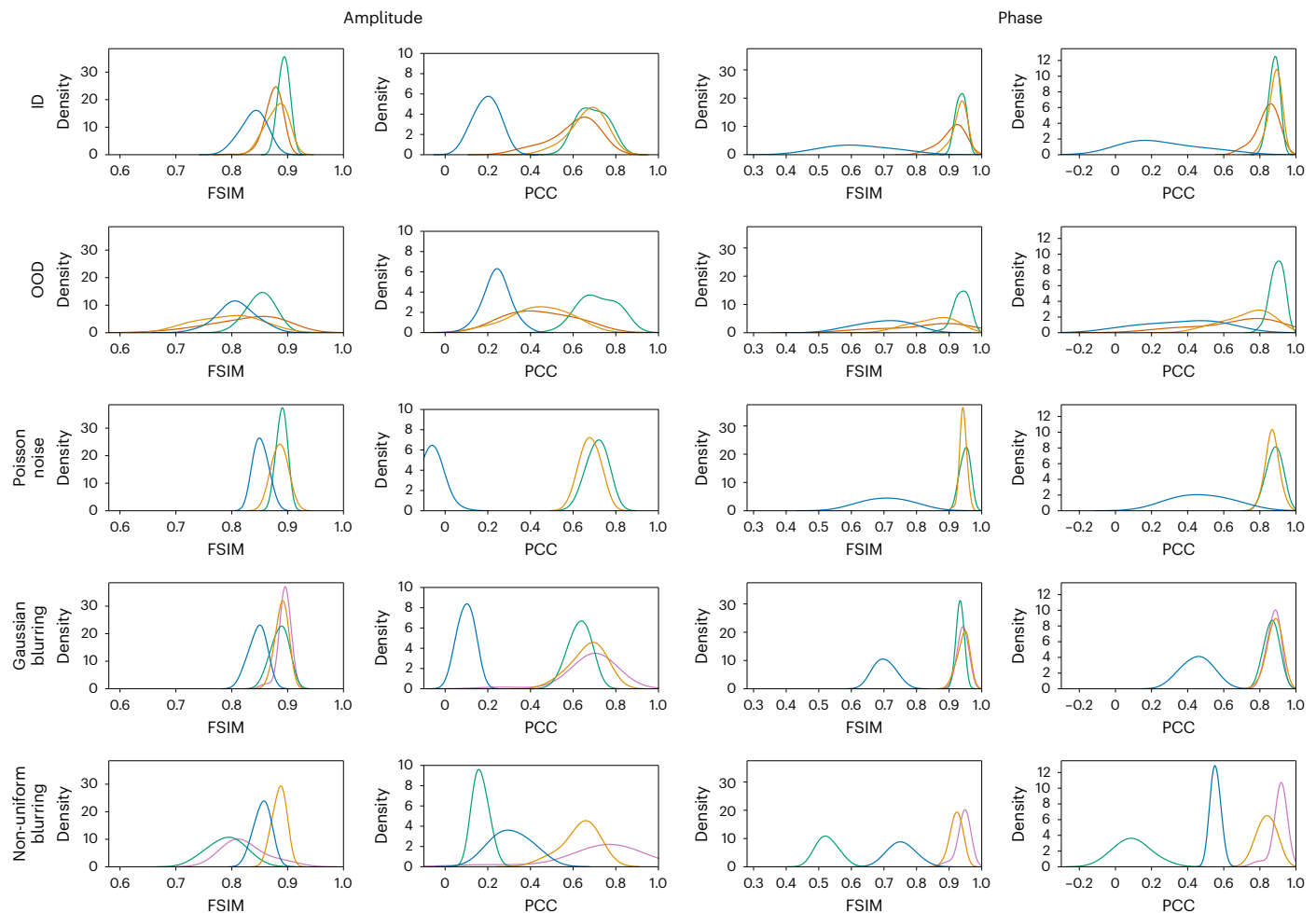
We ran five independent experiments with different random seeds for each method and evaluated the reconstructions of each method using the feature similarity index (FSIM)<sup>9</sup> and Pearson correlation coefficient (PCC)<sup>10</sup> as figures of merit. The results are presented in Table 1 and Fig. 1. Table 1 reported the mean of the FSIM and PCC between the predicted complex images and the ground truth, and also the ratio between CycleGAN, PhaseGAN and PhaseGAN<sup>†</sup> relative to FMGAN. For FMGAN, we also evaluated the predicted propagation distances by the mean absolute error (MAE). In Table 1, we report the mean of the MAE of the predicted propagation distances. The distribution of FSIM and PCC for the in-distribution and OOD data is shown in Fig. 1.

## Adaptability to non-perfect optical systems

In this section, we study the robustness of the approaches to three different non-perfect optical systems.

- (1) Poisson noise: we studied the noise tolerance of the approaches. We simulated noisy measurements by adding normalized Poisson noise to the detector images.
- (2) Gaussian blurring: we studied the generalizability of the approaches to a defocused optical system. We characterized the response of this optical system by using the PSF. The PSF was mathematically included as a convolution to the diffraction intensity images on the detector domain but not explicitly in the forward propagation process. We used a Gaussian PSF with a kernel size of 25 pixels and a standard deviation of three pixels.
- (3) Non-uniform blurring: we studied the generalizability of the approaches to non-uniform blurring, which could be because of a tilted image plane or defects in the optics. We simulated the blurring effect using Gaussian PSFs, which change from left to right on the detector plane. On the left side, we used a Gaussian PSF with a standard deviation of three pixels, and on the right side, we used that with a standard deviation of six pixels. The detector image was formed by a linear combination of the two blurring images.

We study the performance of CycleGAN, PhaseGAN and FMGAN under the three simulated non-perfect optical systems. For the non-perfect optical systems 2 and 3, we introduced FMGAN<sup>†</sup>, where the blurring effects were added explicitly to the forward physical propagator of FMGAN accordingly. For each condition, we trained all of the



**Fig. 1 | Error distribution validation.** FSIM and PCC error distribution of amplitude (left two columns) and phase (right two columns) for CycleGAN (blue), PhaseGAN (yellow) and FMGAN (green) for in-distribution (row 1), OOD (row 2) and the non-perfect optical systems (rows 3 to 5). The results of PhaseGAN\* (red) and FMGAN\* (purple) are also shown where applicable.

networks using 600 simulated diffraction intensities with propagation distances between 7 and 17 mm in steps of 2 mm (100 images for each distance) and the ground truth. We performed data augmentation by applying random rotations and flips to the images. Then, we evaluated the performance of the approaches with these perturbations. For the tests of the last two systems with the Gaussian blurring, we used  $\lambda_{\text{Cyc}} = 400$ ,  $\lambda_{\text{FRC}} = 5$  and  $\lambda_{\text{SSIM}} = 2$  for the training of PhaseGAN, and  $\lambda_{\text{Cyc}} = 200$  for the training of CycleGAN. The rest of the training parameters were the same as those used in the previous section.

Figure 2 shows the simulated diffraction intensity images of a test image at a propagation distance of 8 mm, and the amplitude, as well as the phase, reconstructed by the different methods together with the ground truth. We also reported the mean of the FSIM and PCC between the predicted complex images and the ground truth, and the relative ratio between the CycleGAN, PhaseGAN, FMGAN\* and FMGAN results in Table 2. The MAE of the predicted propagation distances is also reported for FMGAN\* and FMGAN. The distribution of the FSIM and PCC between the reconstruction results and the ground truth images of the three or four different methods is shown in the last three rows of Fig. 1.

### Adaptability to the propagation distance

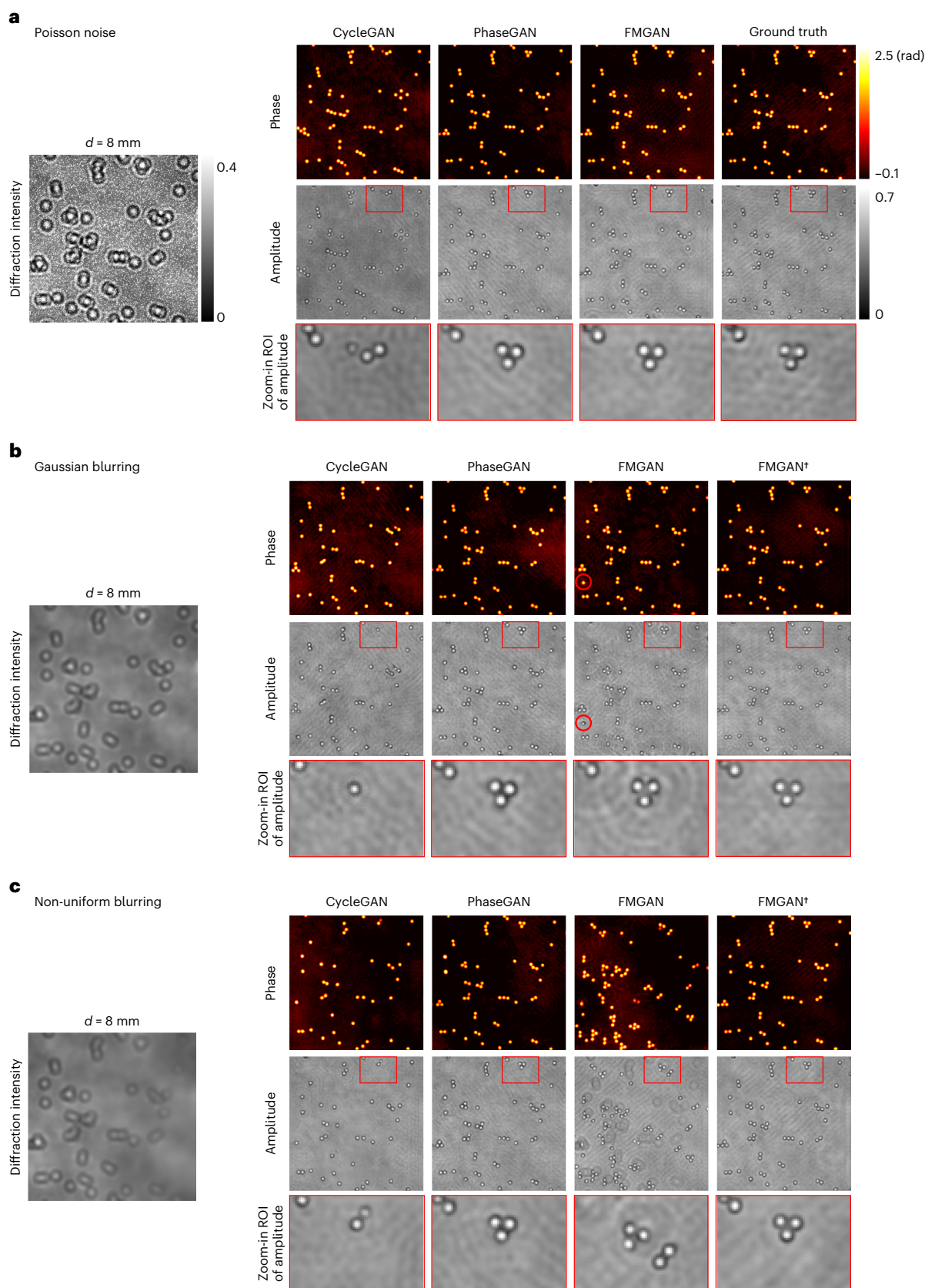
In this section, we investigate the adaptability of the FMGAN to varying propagation distances, specifically, to distances outside the range it was trained on. We trained the networks on diffraction intensity images measured at 13 mm, and tested the limit of the trained network on OOD

simulated diffraction intensity images with the propagation distance ranging from 2 to 26 mm. We started from the propagation distance of 2 mm because interference fringes were not visible below that distance. For the simulation, we generated 27 diffraction intensity images from the test image as shown in Fig. 2. We used the same forward physics model as in ref. 1, based on the law of light propagation<sup>11</sup>, to simulate the diffraction intensity images.

The quantitative evaluation of the network on the simulated diffraction images as a function of the propagation distance is reported in Fig. 3a. We also plotted the relationship between the FMGAN predicted propagation distances and the real propagation distances in Fig. 3b. The mean and standard deviation of the FSIM and PCC between the predicted complex images and the ground truth and the MAE of the predicted distances below, within and above the distance range used in ref. 1 (7–17 mm) are reported in Table 3.

### Discussion and conclusion

We performed three tests in this report to analyse the reusability and adaptability of FMGAN. First, we investigated the reproducibility of FMGAN's results reported in the original paper and compared its performance with two other state-of-the-art image reconstruction networks, CycleGAN and PhaseGAN. We also studied the influence of including an FRC loss by comparing the performance between PhaseGAN and PhaseGAN\*. As shown in Table 1 and Fig. 1, our results are consistent with the results reported in ref. 1. For the in-distribution data, which



**Fig. 2 | Demonstration of the networks' adaptability to non-perfect optical systems.** Three different distortions are applied to the diffraction intensity images measured at a propagation distance of 8 mm to simulate the blurring effect of the three non-perfect optical systems: Poisson noise (a), Gaussian blurring (b) and non-uniform blurring (c). The reconstruction results of the

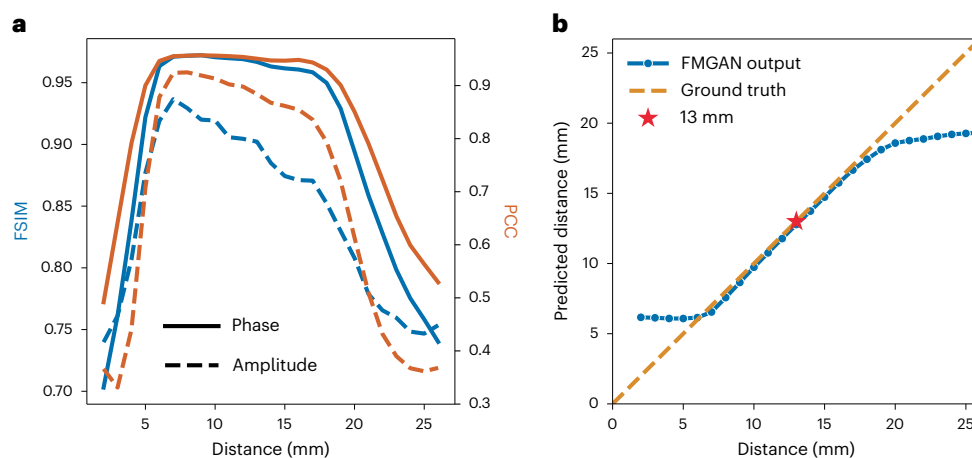
diffraction intensity image reconstructed by the three or four networks and the corresponding ground truth are shown on the right. Magnified views of the areas marked by the red boxes are shown at the bottom. Red circles mark an unexpected polystyrene bead reconstructed by FMGAN. ROI, region(s) of interest.



**Table 2 | Experimental results for the adaptability tests of non-perfect optical systems**

Reconstruction results <sup>a</sup>			CycleGAN		PhaseGAN		FMGAN <sup>†</sup>		FMGAN	
Poisson noise	Amplitude	FSIM	0.85	(0.96×)	0.89	(1.00×)	–	–	0.89	(1.00×)
		PCC	–0.10	(0.00×)	0.68	(0.95×)	–	–	0.71	(1.00×)
	Phase	FSIM	0.71	(0.75×)	0.94	(0.99×)	–	–	0.95	(1.00×)
		PCC	0.48	(0.54×)	0.87	(0.99×)	–	–	0.88	(1.00×)
Distance	MAE	–	–	–	–	–	–	0.2	±0.2	
Gaussian blurring	Amplitude	FSIM	0.85	(0.96×)	0.89	(1.01×)	0.89	(1.00×)	0.89	(1.00×)
		PCC	0.10	(0.15×)	0.67	(1.06×)	0.68	(1.08×)	0.63	(1.00×)
	Phase	FSIM	0.70	(0.75×)	0.94	(1.01×)	0.94	(1.01×)	0.93	(1.00×)
		PCC	0.45	(0.52×)	0.88	(1.02×)	0.88	(1.01×)	0.86	(1.00×)
Distance	MAE	–	–	–	–	0.2	±0.2	0.6	±0.4	
Non-uniform blurring	Amplitude	FSIM	0.85	(1.08×)	0.89	(1.12×)	0.82	(1.04×)	0.79	(1.00×)
		PCC	0.31	(1.87×)	0.63	(3.80×)	0.72	(4.31×)	0.17	(1.00×)
	Phase	FSIM	0.75	(1.43×)	0.92	(1.75×)	0.94	(1.80×)	0.53	(1.00×)
		PCC	0.55	(5.94×)	0.84	(8.98×)	0.91	(9.74×)	0.09	(1.00×)
Distance	MAE	–	–	–	–	0.3	±0.3	13	±3	

<sup>a</sup>For each test, the mean values of the FSIM and PCC between the ground truth and the predicted complex images are presented, along with the performance ratios between CycleGAN, PhaseGAN and FMGAN<sup>†</sup> relative to FMGAN. For FMGAN and FMGAN<sup>†</sup>, the mean and standard deviation of the MAE for the predicted propagation distances are also provided.



**Fig. 3 | Demonstration of FMGAN's adaptability to the propagation distance.** **a**, FSIM (blue) and PCC (orange) of the FMGAN reconstructed phase (solid line) and amplitude (dashed line) for each propagation distance. **b**, FMGAN predicted

distances (blue solid) as a function of the real propagation distances (yellow dashed). The red star marked the propagation distance where FMGAN was trained.

consisted of diffraction intensity images from the same distribution as the training data, both PhaseGAN (including PhaseGAN and PhaseGAN<sup>†</sup>) and FMGAN achieved high-precision reconstruction results, with FSIM and PCC values close to the values reported in the original paper. FMGAN achieved the highest mean and lowest standard deviation in FSIM and PCC values among the three networks, indicating a better performance compared to CycleGAN and PhaseGAN. For the OOD data, where the test images were from a different distribution from the training data, FMGAN also outperforms the other two methods in both FSIM and PCC, suggesting that FMGAN has a better generalizability than CycleGAN and PhaseGAN by the inclusion of the parameterized physical forward model. The performance of PhaseGAN dropped when tested on OOD data, as it is based on a precise propagation model with a known propagation distance, which better constrains the training and improves the performance if compared to CycleGAN but also limits its adaptability to OOD data due to the 'hard-coded' physical model. We also observed that PhaseGAN outperformed PhaseGAN<sup>†</sup>, providing evidence that the inclusion of an FRC loss leads to enhanced

reconstructions. CycleGAN failed to produce reliable results for both in-distribution and OOD data demonstrating the relevance of the physical model. As in ref. 1, we noticed that CycleGAN generated indistinguishable polystyrene bead images, but completely different from the ground truth. Furthermore, one can observe in Fig. 1 that the phase components were generally better reconstructed than the amplitude ones, as polystyrene beads are mainly phase objects.

Second, we investigated the effect of introducing non-perfect optical systems via noise and blurring effects on the performance of CycleGAN, PhaseGAN and FMGAN. Additionally, we examined the influence of introducing an explicit forward propagator by comparing the performance between FMGAN and FMGAN<sup>†</sup>. As shown in Figs. 1 and 2a,b and Table 2, our results showed that both FMGAN and PhaseGAN could reconstruct high-quality images under the influence of Poisson noise and uniform Gaussian blurring, whereas CycleGAN failed to capture many important details in the reconstructed images. However, we have observed phase artefacts in the reconstructions of FMGAN, as highlighted in the zoomed-in areas of Fig. 2. Additionally,

**Table 3 | Experimental results for the distance adaptability test**

Distance range (mm)	Amplitude		Phase		Distance MAE <sup>a</sup>
	FSIM	PCC	FSIM	PCC	
[2, 7)	0.85±0.07	0.6±0.3	0.86±0.11	0.8±0.2	2±2
[7, 17]	0.90±0.02	0.89±0.03	0.97±0.01	0.95±0.01	0.3±0.1
(17, 26]	0.85±0.07	0.7±0.2	0.91±0.08	0.9±0.2	3±2

<sup>a</sup>The mean values and standard deviations of the FSIM and PCC between the ground truth and the predicted complex images are reported for each distance range. In addition, the mean and standard deviation of the MAE for the predicted propagation distances are provided.

we have also noticed the occurrence of image hallucinations for the reconstructions of the non-perfect optical system 2, as marked by red circles in Fig. 2b. These phase artefacts and the hallucinations are not presented in the ground truth or in the reconstructions of PhaseGAN. For the reconstructions of non-perfect optical system 3, both CycleGAN and FMGAN failed to reconstruct the amplitude and phase of the object due to the non-uniform blurring of the detector images, as shown in Fig. 2c. Comparing the results between FMGAN and FMGAN<sup>†</sup>, it is noticeable that including the blurring effect explicitly to the forward propagator improved the reconstructions. FMGAN<sup>†</sup> correctly reconstructed the details of the sample images for both non-perfect optical systems 2 and 3, despite some minor artefacts on the background. These results indicate that FMGAN works well under noisy conditions. For optical systems characterized by a PSF, PhaseGAN outperformed FMGAN. The superior performance of PhaseGAN over FMGAN can be explained by the fact that FMGAN uses only one field generator, which makes it challenging to learn the performance of the optical system. By contrast, PhaseGAN uses two generators: a field generator and a detector generator. The detector generator of PhaseGAN can help it to learn the response of the optical system. Including the response of the optical system explicitly to the forward propagator is an effective way to improve the performance of FMGAN, but this information is not always available. There are conditions where previous knowledge of the response of the optical system is not available or difficult to measure, where PhaseGAN would be more suitable.

Third, we studied the limit of adaptability of FMGAN on propagation distances. As shown in Table 3 and Fig. 3, we observed that FMGAN could reliably reconstruct images within the distance range of 7 to 18 mm. This corresponded to the 7 to 17 mm test datasets used in ref. 1. Outside this range, the performance of FMGAN decreased, and it failed to reconstruct the phase and amplitude. Although FMGAN has been trained at the propagation distance of 13 mm, we have observed that it performs slightly worse at this propagation distance than shorter ones, for example, 7 to 10 mm. A possible explanation for this could be that FMGAN has difficulties in solving the holograms when the objects interfere with each other. We noticed that at shorter propagation distances (below 11 mm), the beads barely interfere with each other, leading to good reconstructions by FMGAN, even though not being trained at those propagation distances. From this propagation distance, the single-bead fringes started to interfere with the ones from other beads, increasing the hologram's complexity. This interference between the different beads seems to be difficult for FMGAN to learn, and this leads to deteriorated behaviour at longer distances.

In conclusion, our analysis of the FMGAN model indicates that it is highly reusable for holographic image reconstructions. Compared to CycleGAN and PhaseGAN, FMGAN demonstrated better performance in reconstructing in-distribution data in terms of the mean and standard deviation of FSIM and PCC values due to the inclusion of a parameterized physical forward model. Additionally, FMGAN showed good generalizability when reconstructing OOD data and noisy data, and our adaptability study indicated a reliable range of approximately 10 mm. However, we observed that when dealing with blurring non-perfect

systems, FMGAN struggled to learn the response of the optical system and generated hallucinations. By contrast, PhaseGAN appeared to be better suited for such scenarios due to its ability to learn the response of the optical system. Finally, CycleGAN performed poorly in all of our tests due to the lack of a physics model.

## Future directions

In this section, we explore opportunities to improve the performance of holographic image reconstruction using deep learning. One promising approach could be to combine the strengths of both PhaseGAN and FMGAN. As we noticed in our analysis, PhaseGAN's ability to learn the response of the optical system could be beneficial in situations where the optical system is not ideal, while FMGAN's inclusion of a parameterized physical forward model improves the versatility and robustness when the propagator cannot be determined. A potential approach that combines a single generator to retrieve complex field images and the relevant propagator parameters and another generator to learn the response of the optical system can improve the capabilities of PhaseGAN and FMGAN. Such an approach could potentially overcome limitations observed in our analysis, such as the hallucinations generated by FMGAN in non-ideal optical systems, while keeping its generalization to OOD. Another aspect to consider for future improvements is related to the availability and quality of training data. As larger and more diverse datasets become accessible, new deep-learning architectures that capture complex spatial relationships in holographic images may be developed. Investigating the potential of transfer learning could also be valuable for future research, where pretraining a model on a large dataset of holographic images can reduce the required training data and improve performance in future applications. Finally, exploring different loss functions and their combination could lead to an improvement and provide more robust models. For instance, incorporating perceptual loss functions<sup>12</sup> could potentially enhance the visual quality of the reconstructed images, and including an FRC loss<sup>2,13</sup> could better constrain the images in the Fourier space.

## Data availability

The data used in this report are available in ref. 5.

## Code availability

The FMGAN code is available in ref. 14. PhaseGAN code is available in ref. 15. The FMGAN<sup>†</sup> code used in the non-perfect optical system tests is available in ref. 16.

## References

- Lee, C., Song, G., Kim, H., Ye, J. C. & Jang, M. Deep learning based on parameterized physical forward model for adaptive holographic imaging with unpaired data. *Nat. Mach. Intell.* **5**, 35–45 (2023).
- Zhang, Y. et al. Phasegan: a deep-learning phase-retrieval approach for unpaired datasets. *Opt. Express* **29**, 19593–19604 (2021).
- Zhu, J.-Y., Park, T., Isola, P. & Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proc. IEEE International Conference on Computer Vision* 2223–2232 (IEEE, 2017).
- Yin, D. et al. Digital holographic reconstruction based on deep learning framework with unpaired data. *IEEE Photon. J.* **12**, 3900312 (2020).
- Lee, C., Song, G., Kim, H., Ye, J. C. & Jang, M. 3 μm polystyrene bead, red blood cell and histological slide datasets. [figshare.com/articles/dataset/3um\\_polystyrene\\_bead\\_red\\_blood\\_cell\\_datasets/21378744](https://figshare.com/articles/dataset/3um_polystyrene_bead_red_blood_cell_datasets/21378744) (2022).
- Hu, J., Shen, L. & Sun, G. Squeeze-and-excitation networks. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* 7132–7141 (IEEE, 2018).
- Kingma, D. P. & Ba, J. ADAM: a method for stochastic optimization. Preprint at <https://arxiv.org/abs/1412.6980> (2014).

8. Mom, K., Langer, M. & Sixou, B. Deep Gauss–Newton for phase retrieval. *Opt. Lett.* **48**, 1136–1139 (2023).
9. Zhang, L., Zhang, L., Mou, X. & Zhang, D. FSIM: a feature similarity index for image quality assessment. *IEEE Trans. Image Proc.* **20**, 2378–2386 (2011).
10. Freedman, D., Pisani, R. & Purves, R. *Statistics (International Student Edition)* 4th edn (WW Norton & Company, 2007).
11. Goodman, J. W. *Introduction to Fourier Optics* (Roberts and Company Publishers, 2005).
12. Johnson, J., Alahi, A. & Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. In *Proc. Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, Part II, 14* 694–711 (Springer, 2016).
13. Van Heel, M. & Schatz, M. Fourier shell correlation threshold criteria. *J. Struct. Biol.* **151**, 250–262 (2005).
14. Deep learning based on parameterized physical forward model for adaptive holographic imaging: v1.0. *Zenodo* <https://doi.org/10.5281/zenodo.7220717> (2022).
15. Zhang, Y. PhaseGAN. *Zenodo* <https://doi.org/10.5281/zenodo.10440916> (2021).
16. Zhang, Y., Ritschel, T. & Villanueva-Perez, P. Reusability report: unpaired deep-learning approaches for holographic image reconstruction. *Code Ocean* <https://doi.org/10.24433/CO.6981228.v1> (2022).

## Acknowledgements

We thank Z. Matej for his support and access to the GPU-computing cluster at MAX IV. We acknowledge the use of QuillBot Paraphraser (QuillBot, <https://quillbot.com/>) for language refinement.

## Author contributions

Y.Z. contributed to the experimental work, data analysis and writing of the manuscript. T.R. contributed to the planning and writing of the manuscript. P.V.-P. contributed to the project planning, experimental work and writing of the manuscript.

## Funding

Open access funding provided by Lund University.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence and requests for materials** should be addressed to Yuhe Zhang or Pablo Villanueva-Perez.

**Peer review information** *Nature Machine Intelligence* thanks Suyeon Choi, Manu Gopakumar and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024