Optics Letters

Super-resolution multimode fiber imaging with an untrained neural network

Wei Li,[1,*] Ksenia Abrashitova,[1] and Lyubov V. Amitonova[1,2]

[1]Advanced Research Center for Nanolithography (ARCNL), Science Park 106, 1098 XG Amsterdam, The Netherlands
[2]LaserLaB, Department of Physics and Astronomy, Vrije Universiteit Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands
*w.li@arcnl.nl

Received 21 March 2023; revised 17 May 2023; accepted 22 May 2023; posted 23 May 2023; published 16 June 2023

Multimode fiber endoscopes provide extreme miniaturization of imaging components for minimally invasive deep tissue imaging. Typically, such fiber systems suffer from low spatial resolution and long measurement time. Fast super-resolution imaging through a multimode fiber has been achieved by using computational optimization algorithms with hand-picked priors. However, machine learning reconstruction approaches offer the promise of better priors, but require large training datasets and therefore long and unpractical pre-calibration time. Here we report a method of multimode fiber imaging based on unsupervised learning with untrained neural networks. The proposed approach solves the ill-posed inverse problem by not relying on any pre-training process. We have demonstrated both theoretically and experimentally that untrained neural networks enhance the imaging quality and provide sub-diffraction spatial resolution of the multimode fiber imaging system.

https://doi.org/10.1364/OL.491375

Introduction. Optical fibers are widely used in many disciplines from biomedical imaging to remote sensing. The combination of multimode fibers (MMFs) and advanced wavefront shaping techniques [1] offers minimally invasive investigation via hair-thin lensless endoscopes [2,3]. Endoscopic fiber imaging applications require the use of a single-pixel detector instead of conventional recording of the two-dimensional (2D) signal. Therefore, imaging through an ultra-thin MMF must be based on advanced computational algorithms such as the transmission matrix measurements, compressive ghost imaging, and/or holographic light shaping [4–6]. Compressive fiber imaging has significant advantages over other modalities in terms of detection sensitivity, spatial resolution, and imaging speed [7–9].

Compressive fiber imaging uses time-varying illumination patterns and the synchronized signal collected by the same fiber and recorded by a bucket detector to reconstruct the sample image. Since the computational algorithm must solve an underdetermined inverse problem, the choice of algorithm and imaging priors is very important. Traditionally, images are recovered by solving a convex optimization problem, minimizing both a least squares loss based on the physical model and a prior term, which enforces sparsity [10]. There is always the need to adjust the method to different experimental conditions when it concerns practical applications [11]. Over the last years, many different approaches have been developed for solving imaging inverse problems [10–14]. However, they are currently being outperformed by deep learning tools, where a network structure is better suited to image representation [15].

Machine learning has become a popular and powerful tool for computational imaging [16–19]. Recently, several deep learning frameworks have been demonstrated for MMF imaging [20–24]. However, these approaches typically require large datasets of labeled image and ground-truth pairs. It makes real-life applications nearly impossible as obtaining ground-truth datasets for a flexible MMF is impractical. As a solution to this problem, we have recently proposed the generative adversarial network (GAN) based MMF imaging framework [25]. This approach does not require labeled pairs but still relies on a large training dataset of expected unlabeled sample images.

Recent works on unsupervised learning with untrained networks, such as deep image prior (DIP) [26], are very promising. It has been proven that ill-posed inverse problems can be solved with a well-designed generator network [27]. Due to the nature of DIP, which directly uses the untrained neural network to solve phase retrieval [28,29] or compressive sensing [30,31] problems, design the diffractive optical element [32], or enhance the reconstruction of ptychography [33], there is no need for the pre-training of the neural network, and hence also no need for the preparation of either labeled or unlabeled training datasets. The structure of neural networks themselves serve as a prior on image statistics without any pre-training. The network is used as an image generator, outputting the recovered image, while its weights and biases are updated through a loss function comparing the results with the experimental data. Untrained neural networks for compressive optical imaging have been recently demonstrated [34,35].

Here we present an MMF compressive imaging framework based on an untrained neural network. We have shown the improved performance and better reconstruction results both theoretically and experimentally.

Theory. Super-resolution compressive endo-microscopy [8] is based on a speckle illumination and single-pixel detection scheme. The imaging system and reconstruction process is shown in Fig. 1. An MMF is used to generate a series of speckle patterns to illuminate the sample while a bucket (single-pixel)

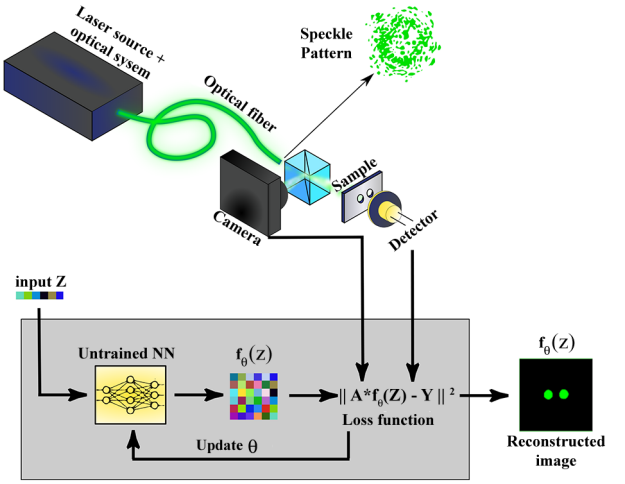

Fig. 1. Schematic illustration of the experimental setup for multimode-fiber-based compressive imaging (top) and of computational reconstruction process (bottom). The sample is illuminated with various speckle patterns generated in the multimode fiber. The corresponding transmitted intensity is recorded by the single-pixel detector. The camera records the illumination speckle patterns. These data are used for the computational reconstruction of the image.

detector records the total signal. The imaging process can be described with the forward model of a linear equation, while multiple measurements formulate the under-determined system:

$$\boldsymbol{y} = A\boldsymbol{x}, \tag{1}$$

where $\boldsymbol{x}$ is the unknown flattened sample of size $n \times 1$ and the flattened illumination speckle A_i (also $n \times 1$) turns out to be one row of the measurement matrix A of size $m \times n$. A total of m intensity measurements give the measurement vector $\boldsymbol{y}$ of size $m \times 1$ ($m \ll n$). The traditional ghost imaging approach solves this ill-posed inverse problem by correlating the intensity with the corresponding illumination pattern:

$$\boldsymbol{x} = \frac{1}{m}\sum_{i=1}^{m}(\boldsymbol{y}_i - \langle \boldsymbol{y} \rangle)A_i. \tag{2}$$

The more advanced basis pursuit (BP) method of compressive sensing explores the sparsity constraint and reconstructs the image by

$$\min_{\boldsymbol{x}} ||\boldsymbol{x}||_1 \quad \text{s.t.} \quad \boldsymbol{y} = A\boldsymbol{x}, \tag{3}$$

where the $||\boldsymbol{x}||_1$ is the l_1 norm of vector $\boldsymbol{x}$. In experiments, we use BPDN, which is the denoising version of BP. The ghost imaging and BP approaches are used to compare the image reconstruction quality with the untrained neural network based fiber imaging.

For the architecture of the untrained neural network, we use the U-net [36] structure modified from the study of Wang *et al.* [34], which consists of two convolutional layers in the encoder and one deconvolutional layer in the decoder with skip connections. The untrained neural network $f_\theta(z)$ solves Eq. (1) by minimizing the loss function:

$$L(\theta) = ||Af_\theta(z) - \boldsymbol{y}||^2, \tag{4}$$

where z is the input for the untrained neural network and θ contains the weights and biases. By training on the single data point,

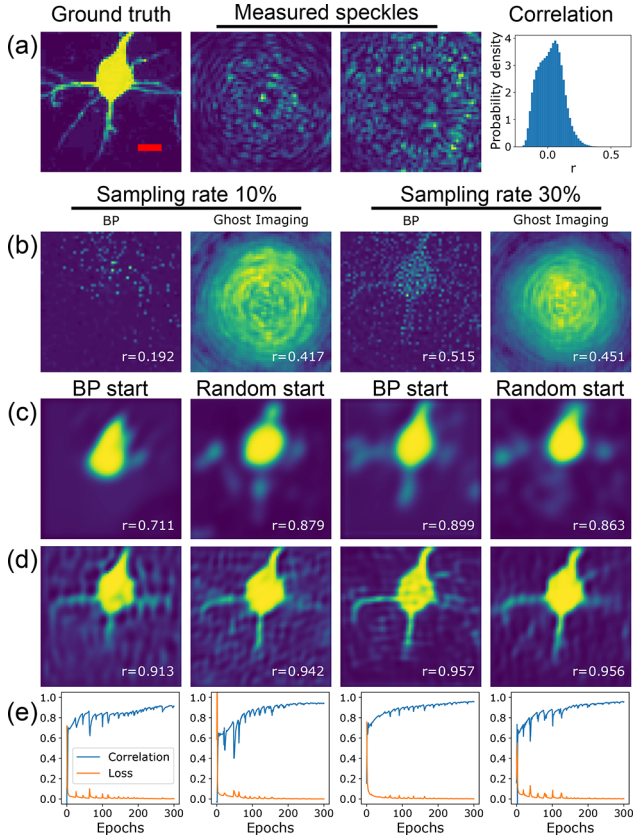

Fig. 2. Simulation results for MMF-based imaging. (a) Left, ground-truth image of the single neuron; middle, examples of the measured speckle patterns used in simulations; right, speckle pattern cross correlation histogram. (b) Image reconstructed with BP (first and third columns) and ghost imaging algorithms (second and fourth columns) for the sampling rate of 10% (left) and 30% (right). (c),(d) Imaging results of the proposed approach with the untrained neural network after (c) 100 and (d) 300 epochs for the sampling rate of 10% (left) and 30% (right) with the respective input z indicated on top. Correlation coefficient r with the ground truth is shown in each image. (e) Correlation with the original image and loss change during the reconstruction process as a function of the number of epochs for different sampling rates and inputs z. The scale bar is equal to the 5× diffraction limit ≈ 6 μm.

the solution of Eq. (1) is approximated as $\hat{\boldsymbol{x}} = f_\theta(z)$. The imaging reconstruction process using an untrained neural network is illustrated in Fig. 1. Image reconstructions have been done on a workstation computer with AMD 3970X CPU with 128-GB memory and the Nvidia RTX 2070 super GPU with 8-GB video memory. The Pearson correlation coefficient r between the reconstructed image and the ground-truth image is used to evaluate the reconstruction quality, due to its better performance than the structural similarity index (SSIM) [37].

Simulations. First, we test the performance of fiber-based compressive imaging with the untrained neural network, we perform the numerical experiment, and simulate the realistic imaging procedure. As a sample, we use a single neuron shown in Fig. 2(a, left). The ground-truth image has a size of 56 × 56 pixels and 35.6 × 35.6 μm. We simulate real imaging conditions by using a set of experimentally measured speckle patterns. The diffraction limit of an MMF is $d_{\text{diffr}} = \frac{\lambda}{2\text{NA}} \approx 1.2$ μm, where $\lambda = 0.532$ μm is the wavelength of light and NA= 0.22 is the

numerical aperture. The different speckles were generated by sequential scanning of the focused beam over the input fiber facet on a square grid with a mean distance between adjacent points of approximately 1.2 μm. The examples of experimentally measured speckle illumination fields are presented in Fig. 2(a, middle). We characterize the linear independence of the recorded speckle patterns by the Pearson correlation coefficients between each pair of the recorded intensity images. The results are presented in Fig. 2(a, right). The histogram shows that the real illumination patterns generated by a multimode fiber have a high level of decorrelation or that they are mainly linearly independent.

We simulate two different sampling rates of 10% and 30% of the total number of pixels by using 314 and 941 speckle patterns, respectively. The signal on the bucket detector ($\boldsymbol{y}$) has been simulated using Eq. (1) with measurement matrix A of size 314×3136 and 941×3136 with no additional noise. First, the reconstruction is done with the traditional BP method [Fig. 2(b), first and third columns] and ghost imaging algorithm [Fig. 2(b), second and fourth columns]. Ghost imaging fails to reconstruct the original sample for both sampling rates. The BP reconstructs the very noisy image only for a relatively high sampling rate of 30%.

In the second set of simulations, we use the same data, but the reconstruction is done by the unsupervised neural network, as shown in Fig. 1. As input z of the untrained neural network, we used either the random noise or the output of the BP algorithm. The image reconstruction results after 100 and 300 epochs (iterations) are shown in Figs. 2(c) and 2(d), respectively. The presented images are retrieved by the unsupervised neural network for the sampling rate of 10% (first and second columns) and 30% (third and fourth columns), and for z equal to the results of BP algorithm (first and third columns) and random noise (second and fourth columns). We see that the reconstruction result is getting better with the number of epochs. We always apply early stopping within the loss function plateau region to avoid overfitting for each simulated and experimentally measured dataset. The Pearson correlation coefficients between the reconstructed image and the original sample are shown in the figures providing a quantitative measure of imaging quality.

For 30% sampling rate, r reaches 0.95 after 300 epochs, while for 10% sampling rate, it is slightly lower. Compared with the BP and ghost imaging results shown in Fig. 2(b), the untrained neural network method shows superior performance. We also investigate the changes during the network-based reconstruction process. Figure 2(e) shows the correlation coefficient and loss function as a function of the number of epochs for different sampling rates and inputs z. It indicates that 300 epochs (iterations) are enough in this case to reach convergence. It takes approximately one minute to run 300 epochs and reconstruct the image on our PC. The image reconstruction speed is not significantly influenced by the original sampling rate and input z. The simulation also shows that while we could see slight improvement if we provide the network with more preliminary data by input z, the input does not significantly influence the final reconstruction quality, i.e., using BP results as input has the similar result of using random data.

Experimental validation. We perform fiber-based compressive imaging experiments with the untrained neural network. The setup is presented in Fig. 1. We scan a focal spot across the MMF input and project the resulting speckle pattern on the sample and a camera with 54× magnification. The total signal from the sample is recorded by a single-pixel detector. As samples, we used custom-made glass slides with a reflective aluminum film patterned with various shapes: dots and hand-written digits as presented in Fig. 3(a). The thickness of the aluminum film is adjusted to ensure a binary transmission function. The smallest features of samples 1–4 were 1.4, 1.8, 2.5, and 4 times smaller than the magnified diffraction limit, which is $\mathrm{M}\times\frac{\lambda}{2\mathrm{NA}} \approx 65$ μm, where M= 54 is the magnification of the optical system [38]. The computational reconstruction allows for the image pixel size to be chosen almost arbitrarily. We aim to resolve features ≈16 μm which is four times smaller than the magnified diffraction limit. Therefore, we went for a pixel size of 5 μm and 400×400 pixels in the image to ensure accurate representation and that the Nyquist criterion is satisfied. The sampling rate can be calculated as $961/400^2 \times 100\% = 0.6\%$. For samples 5 and 6, sampling rates were chosen to be $961/540^2 \times 100\% = 0.3\%$ and $2013/300^2 \times 100\% = 2.2\%$, respectively. Such extremely low sampling rates are the result of a large field of view combined with a limited number of illuminating speckle patterns.

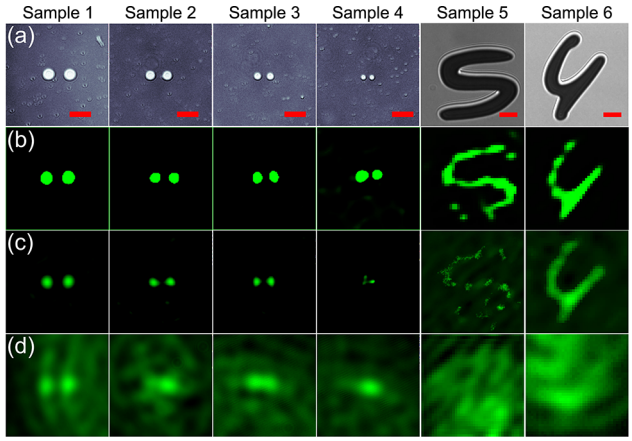

Fig. 3. Multimode fiber imaging experiments. (a) Inverse contrast reference images of six samples with different feature sizes and complexity. (b)–(d) Reconstruction results from the same dataset using the (b) untrained neural network; (c) BPDN algorithm; and (d) ghost imaging algorithm. The scale bar is equal to the magnified diffraction limit of the optical system ≈ 65 μm.

We use experimentally measured matrices A and bucket detector intensities y also reported in [25,38] to reconstruct the transmission function of sample x and experimentally test the performance of the proposed untrained neural network based imaging [presented in Fig. 3(b)] and compare it with other reconstruction approaches: l_1 norm minimization by BPDN [Fig. 3(c)] and traditional ghost imaging [Fig. 3(d)].

Ghost imaging does not provide any resolution enhancement and can resolve two dots only for sample 1. For more complex samples 5 and 6, ghost imaging returns only the noise. The more advanced BPDN algorithm provides nearly perfect image contrast and resolution enhancement below the diffraction limit for samples 1–3. However, sample 4 is not resolved and the reconstruction quality of more complex samples 5 and 6 is mediocre. The proposed MMF imaging approach with an untrained neural network allows for resolving two dots for all four samples as presented in Fig. 3(b). It also provides better reconstruction of the more complex structure of samples 5 and 6, especially in the case of a very low sampling rate. We have always started from an

untrained network and run the reconstruction procedure for 300 epochs, applying early stopping within the loss function plateau region. As input z, we use a random noise. The reconstruction procedure takes around ten minutes.

We experimentally show that the proposed neural network based fiber imaging provides better image reconstruction than other methods resolving features with a size four times smaller than the diffraction limit. While two dots are resolved in the image, their sizes are not always reconstructed accurately and are enlarged compared to the ground truth. This could be the result of an extremely small sampling rate of 0.6% as well as the presence of experimental noise. To further improve the reconstruction quality and tolerate the error in the experimentally measured forward model, the regularization terms can be included into the loss function [34].

Conclusion. To summarize, we present the application of using an untrained neural network in the MMF compressive imaging framework. We have demonstrated that the untrained neural network improves the reconstruction quality of the images in the MMF compressive imaging framework both in simulation and in experiments with different samples. While all the samples are resolved by the untrained neural network, images could still suffer from artifacts. This is the result of an extremely low sampling rate as well as the presence of experimental noise. The untrained network outperforms conventional methods, especially in the case of the extremely low sampling rates. The proposed approach takes advantage of the training process of the neural network, and hence does not need a training dataset, which has a huge benefit over other traditional deep learning based methods.

Funding. Nederlandse Organisatie voor Wetenschappelijk Onderzoek (WISE); Dutch Ministry of Economic Affairs and Climate Policy (TKI).

Acknowledgments. This work has been carried out within ARCNL, a public-private partnership between UvA, VU, NWO, and ASML and was partly financed by "Toeslag voor Topconsortia voor Kennis en Innovatie (TKI)" from the Dutch Ministry of Economic Affairs and Climate Policy. We thank Sergey Amitonov (TU Delft) for custom-made samples, Mark Mol (ARCNL) for the technical support, and Marco Seynen (AMOLF) for his help in programming the software. We would like to thank Mark van Kraaij, Hugo Cramer and Lorenzo Tripodi for fruitful discussions.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

REFERENCES

1. I. M. Vellekoop, Opt. Express **23**, 12189 (2015).
2. T. Čižmár and K. Dholakia, Opt. Express **19**, 18871 (2011).
3. R. Di Leonardo and S. Bianchi, Opt. Express **19**, 247 (2011).
4. I. M. Vellekoop, A. Lagendijk, and A. Mosk, Nat. Photonics **4**, 320 (2010).
5. T. Čižmár and K. Dholakia, Nat. Commun. **3**, 1 (2012).
6. M. J. Padgett and R. W. Boyd, Philos. Trans. R. Soc., A **375**, 20160233 (2017).
7. L. V. Amitonova and J. F. De Boer, Opt. Lett. **43**, 5427 (2018).
8. L. V. Amitonova and J. F. de Boer, Light: Sci. Appl. **9**, 1 (2020).
9. M. Pascucci, S. Ganesan, A. Tripathi, O. Katz, V. Emiliani, and M. Guillon, Nat. Commun. **10**, 1 (2019).
10. B. Lochocki, K. Abrashitova, J. F. de Boer, and L. V. Amitonova, Opt. Express **29**, 3943 (2021).
11. G. Calisesi, A. Ghezzi, D. Ancora, C. D'Andrea, G. Valentini, A. Farina, and A. Bassi, Prog. Biophys. Mol. Biol. **168**, 66 (2021).
12. M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, IEEE J. Sel. Top. Signal Process. **1**, 586 (2007).
13. S. Gazit, A. Szameit, Y. C. Eldar, and M. Segev, Opt. Express **17**, 23920 (2009).
14. N. Kulkarni, P. Nagesh, R. Gowda, and B. Li, IEEE Trans. Circuits Syst. Video Technol. **22**, 778 (2012).
15. K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, IEEE Trans. on Image Process. **26**, 4509 (2017).
16. G. Barbastathis, A. Ozcan, and G. Situ, Optica **6**, 921 (2019).
17. L. Li, H. Ruan, C. Liu, Y. Li, Y. Shuang, A. Alù, C.-W. Qiu, and T. J. Cui, Nat. Commun. **10**, 1082 (2019).
18. M. Del Hougne, S. Gigan, and P. Del Hougne, Phys. Rev. Lett. **127**, 043903 (2021).
19. M. A. B. Abbasi, M. O. Akinsolu, B. Liu, O. Yurduseven, V. F. Fusco, and M. A. Imran, Sci. Rep. **12**, 1 (2022).
20. M. W. Matthès, Y. Bromberg, J. de Rosny, and S. M. Popoff, Phys. Rev. X **11**, 021060 (2021).
21. B. Rahmani, D. Loterie, G. Konstantinou, D. Psaltis, and C. Moser, Light: Science & Applications **7**, 69 (2018).
22. H. Chen, Z. He, Z. Zhang, Y. Geng, and W. Yu, Opt. Express **28**, 30048 (2020).
23. N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, Optica **5**, 960 (2018).
24. P. Fan, T. Zhao, and L. Su, Opt. Express **27**, 20241 (2019).
25. W. Li, K. Abrashitova, G. Osnabrugge, and L. V. Amitonova, Phys. Rev. Appl. **18**, 034075 (2022).
26. D. Ulyanov, A. Vedaldi, and V. Lempitsky, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2018).
27. S. Dittmer, T. Kluth, P. Maass, and D. Otero Baguer, J Math Imaging Vis **62**, 456 (2020).
28. F. Wang, Y. Bian, H. Wang, M. Lyu, G. Pedrini, W. Osten, G. Barbastathis, and G. Situ, Light: Sci. Appl. **9**, 1 (2020).
29. E. Bostan, R. Heckel, M. Chen, M. Kellman, and L. Waller, Optica **7**, 559 (2020).
30. D. Van Veen, A. Jalal, M. Soltanolkotabi, E. Price, S. Vishwanath, and A. G. Dimakis, "Compressed sensing with deep image prior and learned regularization," arXiv arXiv:1806.06438 (2018).
31. R. Heckel and M. Soltanolkotabi, in *International Conference on Machine Learning*, (PMLR, 2020), pp. 4149–4158.
32. Z.-Y. Chen, Z. Wei, R. Chen, and J.-W. Dong, Opt. Express **29**, 13011 (2021).
33. Q. Chen, D. Huang, and R. Chen, Opt. Express **30**, 39597 (2022).
34. F. Wang, C. Wang, M. Chen, W. Gong, Y. Zhang, S. Han, and G. Situ, Light: Sci. Appl. **11**, 1 (2022).
35. K. Monakhova, V. Tran, G. Kuo, and L. Waller, Opt. Express **29**, 20913 (2021).
36. O. Ronneberger, P. Fischer, and T. Brox, in *International Conference on Medical image computing and computer-assisted intervention*, (Springer, 2015), pp. 234–241.
37. V. Starovoitov, E. Eldarova, and K. Iskakov, Eurasian J. Math. Comput. Appl. **8**, 76 (2020).
38. K. Abrashitova and L. V. Amitonova, Opt. Express **30**, 10456 (2022).