Generative Adversarial Network for Superresolution Imaging through a Fiber

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

[1] *Advanced Research Center for Nanolithography (ARCNL), Science Park 106, Amsterdam 1098 XG, Netherlands*

[2] *LaserLaB, Department of Physics and Astronomy, Vrije Universiteit Amsterdam, De Boelelaan 1081, Amsterdam 1081 HV, Netherlands*

(Received 15 March 2022; revised 17 August 2022; accepted 26 August 2022; published 27 September 2022)

A multimode fiber represents the ultimate limit in miniaturization of imaging endoscopes. However, such a miniaturization usually comes as a cost of a low spatial resolution and a long acquisition time. Here we propose a fast superresolution-fiber-imaging technique employing compressive sensing through a multimode fiber with a data-driven machine-learning framework. We implement a generative adversarial network (GAN) to explore the sparsity inherent to the model and provide compressive reconstruction images that are not sparse in a representation basis. The proposed method outperforms other widespread compressive imaging algorithms in terms of both image quality and noise robustness. We experimentally demonstrate machine-learning ghost imaging below the diffraction limit at a sub-Nyquist speed through a thin multimode fiber probe. We believe that this work has great potential in applications in various fields ranging from biomedical imaging to remote sensing.

DOI: 10.1103/PhysRevApplied.18.034075

I. INTRODUCTION

Optical fibers are broadly used in many imaging applications in various fields ranging from biomedical imaging to remote sensing. Multimode fibers (MMFs) together with advanced wave-front shaping [1] enable alternative ways to transmit information through a hair-thin probe [2,3]. Nowadays, the most popular approaches for imaging through a MMF exploit transmission matrix measurements, ghost imaging, and holographic light shaping [4–6]. On the other hand, computational methods can provide an elegant solution for imaging without the need for active control over the light propagation. Compressive imaging through a MMF improves both spatial resolution and imaging speed [7–9]. However, practical application of compressive imaging has been restricted by the strong assumption of sample sparsity and the demand to adjust the approach to different experimental conditions [10].

Recent years have witnessed the rise of deep learning as a powerful tool for computational imaging and metrology [11–14]. Deep neural networks have been successfully implemented to classify and reconstruct images distorted by propagating through a MMF [15–19]. However, these modern machine-learning approaches are only applied for coherent image transmission through a MMF and, for instance, cannot be used for common fluorescence imaging. Moreover, spatial resolution is still limited by the diffraction of light. These machine-learning approaches require a network to be trained with pairs of input images and their corresponding speckle patterns at the fiber output. As a result, it complicates the precalibration system and requires the training procedure to be repeated if the fiber configuration changes [20].

Recently, Bora *et al.* proposed to solve the compressive sensing problem with a generative adversarial network (GAN) and demonstrated sub-Nyquist imaging without any constraints on the sample sparsity [21]. The GAN is an unsupervised machine-learning framework and composed of two neural networks that compete with each other during the training process. A properly trained GAN can generate alternative data belonging to a certain class of objects. The sample image is produced by iterative minimization between the GAN-generated (fake) images and the measured data. Over the past years, the GAN approach for compressive imaging has seen a rapid development [22–26].

Here we propose the GAN-based framework for imaging through a hair-thin MMF. We use deep neural network to explore the sparsity inherent to the model and represent an image from a limited number of variables even for a sample that is not sparse in a measurement basis. In contrast to other methods of learning-based MMF

*w.li@arcnl.nl

imaging, our approach does not require the labeled data and therefore can be easily generalized on any speckle-based measurement system. Moreover, the proposed approach in theory delivers images with subdiffraction spatial resolution enhancement of 3.5×. We show that proper choice of architecture and loss functions allows for recovering images below the diffraction limit with sub-Nyquist imaging speed via a MMF. We experimentally demonstrate the superiority of our approach over standard l_1-norm optimization algorithms in terms of spatial resolution, noise tolerance, and sensitivity to the number of measurements.

II. MACHINE-LEARNING FIBER-BASED GHOST IMAGING

A. Theory

We focus on superresolution microscopy with a speckle illumination and single-pixel detection scheme, also known as superresolution ghost imaging [8]. The schematic imaging process is shown in Fig. 1(a). The sample is illuminated by a series of speckle patterns generated in a MM fiber or a scattering medium. The transmitted signal is recorded by a bucket (single-pixel) detector. The relationship between the intensity distribution of the speckle pattern and the bucket detector signal can be represented as a linear equation, and multiple measurements can be formulated as an underdetermined system of linear equations:

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

where each flattened speckle ($n \times 1$) corresponds to one row of measurement matrix A ($m \times n$). The responses from m different speckle patterns are measured by a bucked detector to form the measurement vector $\mathbf{y}$ ($m \times 1$). Finally, $\mathbf{x}$ is the one-dimensional (1D) unknown vector ($n \times 1$), where $n = N^2$ is the total number of pixels and $N \times N$ is the size of the two-dimensional (2D) digitized speckle pattern. We consider the case with $m \ll n$.

We reconstruct vector $\mathbf{x}$ by two computational approaches: a compressive sensing algorithm and a pretrained GAN. Reconstructed vector $\mathbf{x}$ ($n \times 1$) is resized to represent a 2D sample image ($N \times N$). We compare the results with the diffraction-limited images to evaluate the superresolution and sub-Nyquist capabilities of the two computational approaches. The reconstruction quality is quantified by the Pearson correlation coefficient r between the reconstructed image and the truth image. The correlation r ranges from -1 to 1, while 1 represents the perfect reconstruction, -1 represents the perfect negative reconstruction and 0 indicates that the reconstruction and ground truth are independent. The Structural similarity Index could be also used for quantitative analysis of the imaging quality with the more focus on structural information. However, our calculation show that the structural similarity index actually gives very similar results as the Pearson correlation coefficient but is more affected by the absolute pixel value and behaves less stable.

B. Experimental setup

Our MMF-based single-pixel imaging setup is presented in Fig. 1(c). The light from a laser source (532 nm, Cobolt Samba, continuous wave) passes through a half-wave plate and a polarizing beam-splitter cube to control the incident power. We generate the speckles I_i by scanning the 50-μm input facet of the multimode optical fiber (NA = 0.22) with two galvomirrors as shown in Fig. 1(c). The first galvomirror ensures the vertical movement of the laser beam in the focal plane of the exciting objective (NA = 0.65, Olympus). It is optically projected on the second galvomirror with a 1:1 telescope. The second galvomirror ensures the horizontal movement of the laser beam in the focal plane of the coupling objective. The second 1:1 telescope optically projects the second galvomirror on the entrance aperture of the objective. We spirally scan the round input facet of the MMF in a point-by-point manner, fully covering the input facet of the fiber. The scanning grid is presented in Fig. 1(d) in comparison with the incoherent image of the output MMF facet shown in Fig. 1(e). The total number of points is 2013 with a mean distance between neighboring points 1 μm. The number of modes in the MMF for single polarization is $M_{\text{modes}} = 1050$ and the diffraction limit of the fiber is 1.2 μm. The 54×-magnified image of the speckle pattern is projected on the sample and the camera (pixel size = 2.4 μm, Basler acA 3088-57 um) by the objective (NA = 0.75, Olympus), a tube lens (f = 250 mm) and a beam splitter. The camera and sample need to be axially aligned within the Rayleigh length of the speckle pattern, which in this case corresponds to 1.2 mm [9]. The example of speckle intensity is presented in Fig. 1(f). The total intensity transmitted through the sample is measured by the avalanche photodiode (Thorlabs APD410A2) and forms measurement vector $\mathbf{y}$. The camera and the galvomirrors are triggered by a data acquisition board (NI-PCI 6353) and controlled by custom software.

C. Simulation of machine-learning fiber-based ghost imaging

As a first step, we perform simulations to evaluate the performance of the reconstruction algorithms for a different number of measurements, spatial frequencies, and levels of noise. We generate a random measurement matrix A by creating m random well-developed speckle fields (of size $n = 28 \times 28$ pixels), which are sampled from a circular Gaussian distribution [27]. The intensity distributions

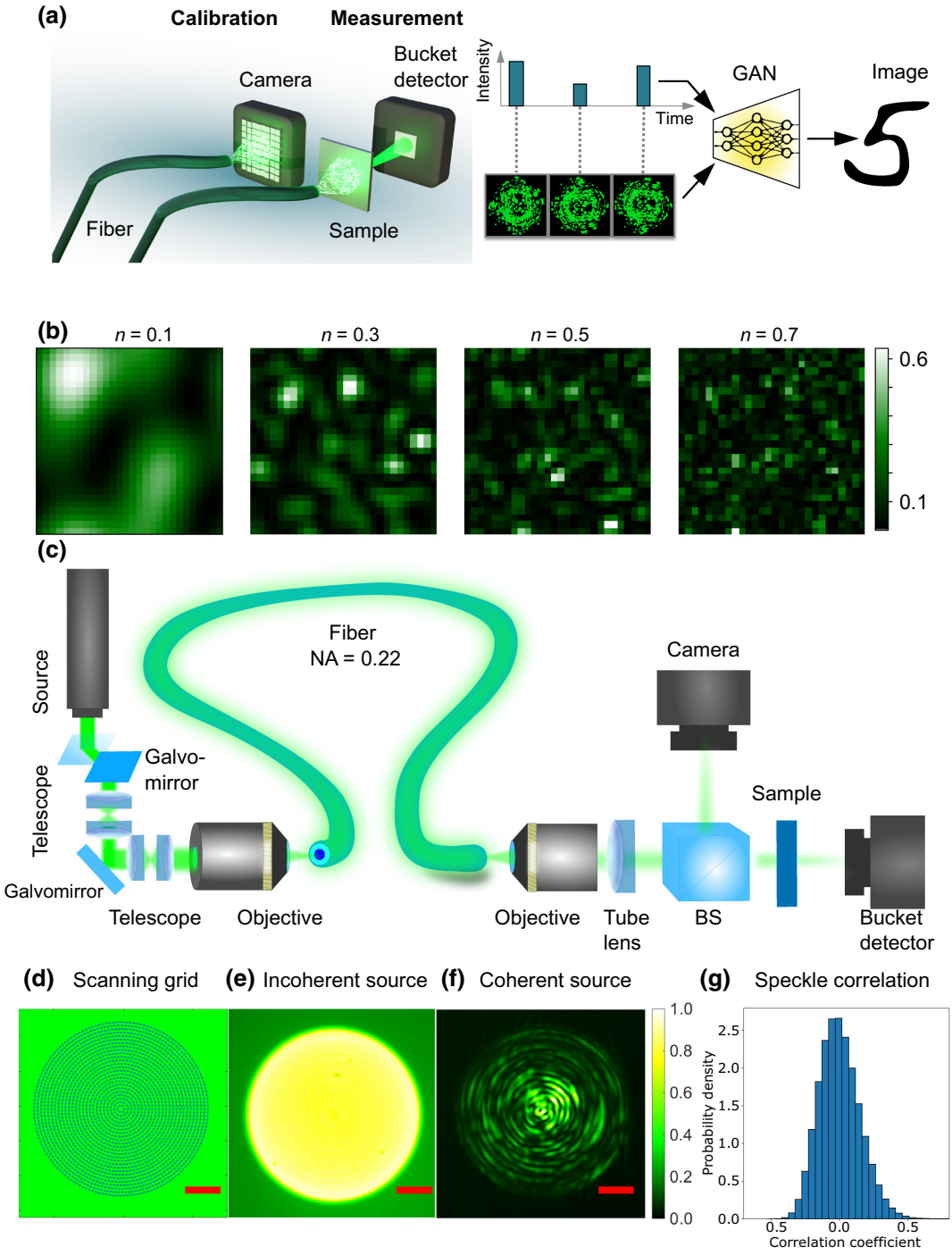

FIG. 1. (a) Simplified representation of machine-learning fiber-based ghost imaging. Different speckle patterns are sequentially projected on the sample. The corresponding total transmitted intensity for each speckle pattern is measured by a bucket detector. The GAN-based algorithm reconstructs the image from the intensity sequence measured by the bucket detector and speckle patterns measured by the camera. (b) Examples of simulated speckle patterns with different cutoff frequencies, ν. (c) Schematic representation of the experimental setup. BS, beam splitter. (d) Scanning grid. (e) Incoherent image of the output facet. (f) Coherent image of the output facet. The scale bar is 10 μm. (g) Probability density function of correlation coefficients for the experimental set of speckle patterns.

of these fields form the rows of our simulated measurement matrix. To study the diffraction limit effects, we apply a low-pass filter with a cutoff frequency ν to the random speckle fields, where $\nu = (0, 1]$ is normalized to the half of the spatial sampling frequency. As ν decreases the speckles get larger, hence less spatial information is obtained

from the measurement. Figure 1(b) shows examples of generated speckle patterns with different ν values.

As a sample, we use handwritten digits from the standard *MNIST* database [28], which contains 70 000 images (60 000 in the training dataset and 10 000 in the testing dataset) with an image size of 28×28, hence $n = 784$. The average sparsity of the complete dataset is 0.19.

In the noiseless case, the measurement signal from the simulated bucket detector is simply given by $\mathbf{y} = \mathbf{Ax}$, where $\mathbf{x}$ is the flattened ground truth image of a handwritten digit. Different levels of noise with the SNR of 40, 35, and 30 dB are added to both the measurement matrix and the measured signal. The noise follows a standard normal distribution with its power corresponding to the specific SNRs.

III. RECONSTRUCTION ALGORITHMS

A. Compressive imaging with GAN

The process of reconstruction by GAN consists of two main steps, as presented in Fig. 2.

1. Training step: training the network to generate the sample images corresponding to a certain class of objects.
2. Imaging step: searching for the image that fits the experimentally measured data.

The training step is completely independent of the reconstruction step and the properly trained network can be used for an arbitrary experimental system without any change.

B. GAN training

A GAN consists of a generator and discriminator, which are simultaneously trained. The generator network performs the transformation from a low-dimensional input latent representation $\mathbf{z}$ to the high-resolution image, whereas the discriminator network discriminates between the images that are "real" (from the training set) and "fake" (newly generated). These two networks learn from and compete with each other during the training process and improve their respective performance. When the GAN network is well trained, the generator can generate new images belonging to the same class as the training set from an arbitrary input $\mathbf{z}$.

For our deep convolutional generative adversarial network [29], we modify the design and the code from *TensorFlow* [30]. The *Keras Sequential API* is used to define both the generator and the discriminator architectures. The input latent representation $\mathbf{z}$ is a (100×1) vector whose entities follow the standard normal distribution. To transform $\mathbf{z}$ to the image (28×28), the generator uses the transposed convolution (deconvolution) layer several times, with the *LeakyReLU* activation for each layer [output layer uses hyperbolic tangent activation function (*tanh*)]. The discriminator is a convolutional neural-network-based image classifier, which gives positive values for real images and negative values for fake images.

Figure 2(a) shows the detailed training process of the deep convolutional GAN. The training dataset, which contains 60 000 handwritten digits, is shuffled and separated into smaller batches with the size of 256 to reduce computational complexity. The generator generates 256 fake images based on $\mathbf{z}$. Discriminator maps these 256 fake images and 256 real images to either positive (real output) or negative values (fake output). Fake output is used to calculate the loss function for the generator, while both the fake output and the real output are used to calculate the loss function of the discriminator. The generator and discriminator are updated by gradient descent with the calculated loss functions, implemented by the *Adam* optimizer [31] with the learning rate of 0.0001. The process of training for the whole training set is repeated 1000 times.

C. Reconstruction with GAN algorithm

Figure 2(b) shows the detailed reconstruction algorithm. The random input vector $\mathbf{z}$ (100×1) goes through the generator, which generates the corresponding image $G(\mathbf{z})$ (28×28). The loss function $L(\mathbf{z})$ can be calculated from the flattened image $G(\mathbf{z})$ combined with measured signal $\mathbf{y}$ and measurement matrix A:

$$L(\mathbf{z}) = ||\mathbf{AG}(\mathbf{z}) - \mathbf{y}||^2. \tag{2}$$

Following Bora *et al.* [21], we minimize the loss function, which quantifies the discrepancy between the guess generated by GAN and the experimental data. We randomly generate the starting value of the input latent representation $\mathbf{z} = \hat{\mathbf{z}}$ and search for the minimum point by gradient descent:

$$\hat{\mathbf{z}} \leftarrow \hat{\mathbf{z}} - \alpha \frac{\partial L(\mathbf{z})}{\partial \mathbf{z}}\bigg|_{\mathbf{z}=\hat{\mathbf{z}}}, \tag{3}$$

where α is the learning rate. Despite the fact that the loss function is highly nonconvex, hundreds of iterations combined with several random initiations of the input [21,32] gives a sufficiently good $\hat{\mathbf{z}}$. $G(\hat{\mathbf{z}})$ is the final reconstructed signal. In order to implement the gradient-descent method, we use *GradientTape* [33] to do the automatic differentiation and use *Adam* optimizer [31] with the learning rate of 0.1 to update the input latent representation $\mathbf{z}$. We use 2000 iteration steps in the gradient-descent optimization and repeat the procedure of optimization 10 times with new randomly generated starting point $\mathbf{z}$. We select the final reconstruction result $G(\hat{\mathbf{z}})$ out of 10 by choosing the $\hat{\mathbf{z}}$ with the lowest values of loss function $L(\hat{\mathbf{z}})$.

D. Basis pursuit

We compare the performance of our GAN reconstruction to a compressive sensing algorithm known as basis

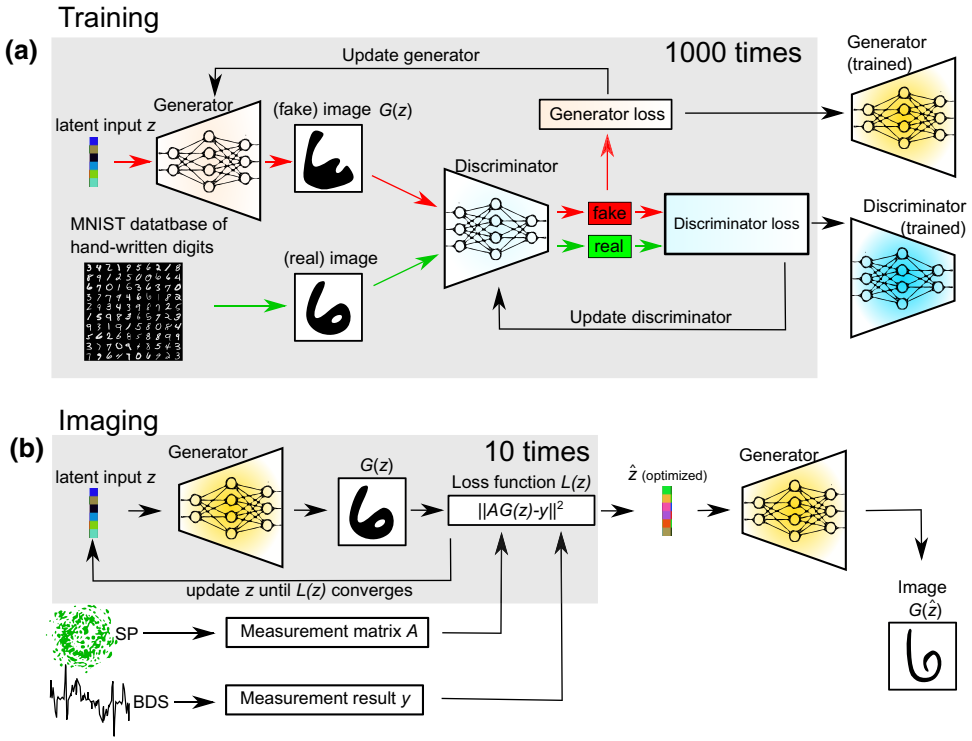

FIG. 2. Illustration of machine-learning fiber-based ghost imaging via GAN: (a) the training process and (b) the imaging step. SP, speckle patterns; BDS, bucket detector signal.

pursuit (BP). Basis-pursuit algorithm can solve the compressive sensing problem by

$$\min_{\mathbf{x}} ||\mathbf{x}||_1 \quad \text{such that} \quad \mathbf{y} = \mathbf{A}\mathbf{x}, \tag{4}$$

where $||\mathbf{x}||_1 = \sum_{\mathbf{i}} |\mathbf{x_i}|$ is the l_1 norm of vector $\mathbf{x}$ or the sum of the absolute values of vector entities. In the case where noise is present, the BP algorithm is substituted by the BP denoising (BPDN) algorithm:

$$\min_{\mathbf{x}} ||\mathbf{x}||_1 \quad \text{such that} \quad ||\mathbf{y} - \mathbf{A}\mathbf{x}||_2^2 \leq \delta, \tag{5}$$

where δ is the error-tolerance factor, which is dependent on the amount and nature of noise in the measurement system, the number of measurements m, and the number of nonzero elements (sparsity) of vector x. The $||.||_2$ is the euclidean norm. The BP denoising becomes BP when $\delta = 0$. We implement the BP algorithm in this study with the *spgl1* package [34,35].

IV. SIMULATION RESULTS

A. Noiseless simulations

We first study the performance of the GAN imaging framework in the ideal cases without noise. The results of the simulations performed on a handwritten digit "3" (randomly selected from the testing dataset) with $m = 40, 100, 200$ simulated speckle patterns and cutoff frequency $\nu = 0.2$ in the noiseless case are presented in Fig. 3. The ground truth, the diffraction-limited results and images reconstructed with the BP algorithm and the GAN are shown in Fig. 3, respectively. The GAN framework shows excellent quality improvement comparing to diffraction-limited imaging and reconstruction based on BP algorithm.

We study the performance of the reconstruction algorithms for each of the combinations of different cutoff frequencies $\nu = [0.1, 0.2, 0.3, 0.5, 0.7]$, number of measurements $m = [10, 40, 70, 100, 200, 300, 400, 500, 750]$, and samples. Each correlation coefficient between the reconstruction result and the ground truth is averaged over the digits of each kind $(0 - 9)$ randomly selected from the testing dataset. The results are shown in Figs. 4(a), 4(b), 4(f). In order to clearly show the performance difference at high r region, we use two color scales. The gray scale corresponds to r range between 0 and 0.9, while the blue-green color scale shows the detailed range between 0.9 and 1. The Nyquist limit, i.e., the minimum number of measurements required for conventional imaging, is shown by the red dashed line. In the case of diffraction-limited imaging, Fig. 4(a), the area below the Nyquist limit is not determined.

The GAN achieves a high correlation reconstruction ($r \geq 0.9$) with a much lower number of measurements and a lower cutoff frequency than BP, as shown in Figs.

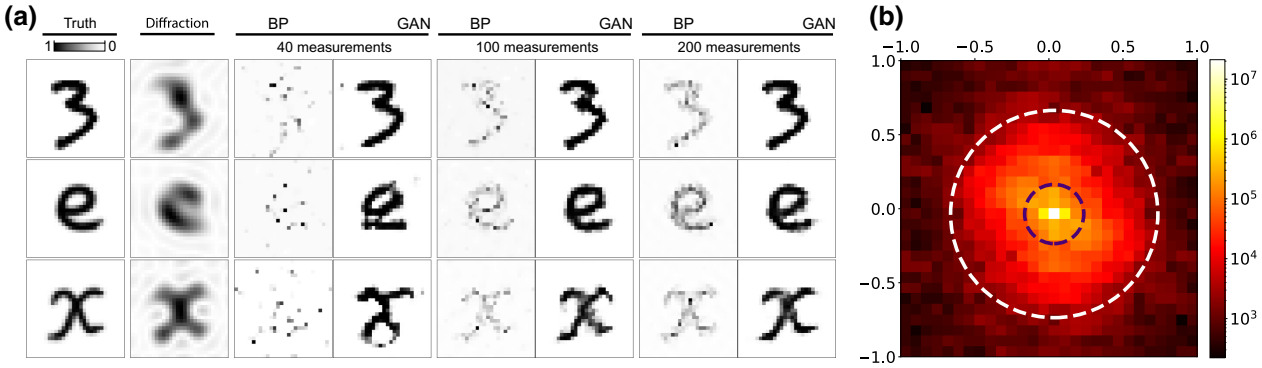

FIG. 3. (a) Results of the simulations performed on a handwritten digit "3" (randomly selected from the testing dataset) and handwritten letters "*e*" and "*x*" with $\nu = 0.2$ and $m = 40$, 100 and 200, respectively. The first column is the ground truth. The second column is the diffraction-limited image. Images reconstructed by the BP algorithm and by the GAN under a different number of measurements are shown from third to eighth column. (b) Average power spectrum for the reconstructed images for the cutoff frequency $\nu = 0.2$, $m \geq 40$. The violet dashed circle corresponds to $\nu = 0.2$, the white dashed circle marks the spatial frequency where the amplitude of the power spectrum decreases 4 orders of magnitude. All spatial frequencies are expressed in the units of half of the spatial sampling frequency.

4(b), 4(f). The correlation coefficient for the GAN quickly saturates to $r = 0.98$ at $\nu = 0.3$ and $m = 80$, while the BP algorithm requires $\nu = 0.6$ and $m = 400$ to reach similar performance. However, the nonconvex nature of the loss function $L(\mathbf{z})$, used to find the optimal solution in the GAN approach, makes it hard to reach the perfect reconstruction ($r = 1$). For $\nu \leq 0.4$, both the GAN and BP demonstrate superresolution imaging, as their image quality is better than the diffraction-limited imaging case. The lowest diffraction-limited cutoff frequency where we can get successful reconstruction is $\nu = 0.2$ for the number of measurements $m \geq 40$. For each successfully reconstructed image $\text{Image}_{ij}(x,y)$ (cross-correlation more than 0.9), we calculate the power spectrum. We find the mean power spectrum by ensemble averaging of power spectra corresponding to the reconstructed images of the different handwritten digits ("0", "1", .. "9") $i = 1..10$ and over a different number of measurements $j = 1..8$: $\text{PS}_{\text{mean}}(\nu_x, \nu_y) = < |\mathscr{F}(\text{Image}_{ij}(x,y)|^2 >$, where $\nu_x, \nu_y = [-1, 1]$ are the discrete spatial frequencies in the units of half of the spatial sampling frequency and $\mathscr{F}$ is the 2D Fourier transform. We define the maximum spatial frequency present in the reconstructed images $\nu_{\text{max image}}$ as a radius of a bright circle (the amplitude of the PS_{mean} within 4 orders of magnitude), as presented in Fig. 3(b) with the white dashed line. The corresponding average spatial frequency for reconstructed images is $\nu_{\text{max image}} = 0.7$, which provides 3.5× resolution enhancement for GAN. At the same time, only GAN overcomes the Nyquist criteria. For instance, at $\nu = 0.7$ we can enhance the imaging speed by roughly 8 times with GAN.

We use a different class of images to evaluate the transfer learning reconstruction performance and the flexibility of the GAN-based imaging approach. The results for handwritten letters "*e*" and "*x*," whose shapes are not close to the digits, are presented in Fig. 3(a). Therefore, the object is not of the same class as training data, although they are both gray scale and of a handwritten nature. The imaging conditions are the same as the handwritten digits. The GAN approach in these cases shows excellent performance demonstrating the flexibility of GAN-based imaging.

B. Simulations with noise

We investigate the performance of the proposed GAN approach under realistic conditions by adding noise on both the measurement matrix A and the measured signal $\mathbf{y}$ as described in Sec. II C. For a fair comparison, we replace BP with the BPDN algorithm, which is optimized for noisy measurements by tuning the error-tolerance factor δ. Meanwhile, the GAN does not require any additional parameter tuning. We repeat the reconstruction for different samples, m, ν and the noise levels. The results averaged over the digits of each kind $(0-9)$ are shown in Fig. 4.

The image quality of the BPDN algorithm deteriorates quickly with the increasing amount of noise, as can be seen in Figs. 4(c), 4(d), 4(e). The maximum achievable reconstruction quality of the BPDN decreases to $r_{\text{max}} = 0.8$ when SNR is 30 dB. At the same time, the GAN provides a good reconstruction (with r_{max} up to 0.97) even for the high noise level [Figs. 4(g), 4(h), 4(i)]. The GAN approach maintains superresolution and sub-Nyquist regimes.

V. EXPERIMENTAL VALIDATION

In order to test our GAN reconstruction framework in practice, we perform a MMF-based compressive imaging experiment on a real sample. Handwritten digits from

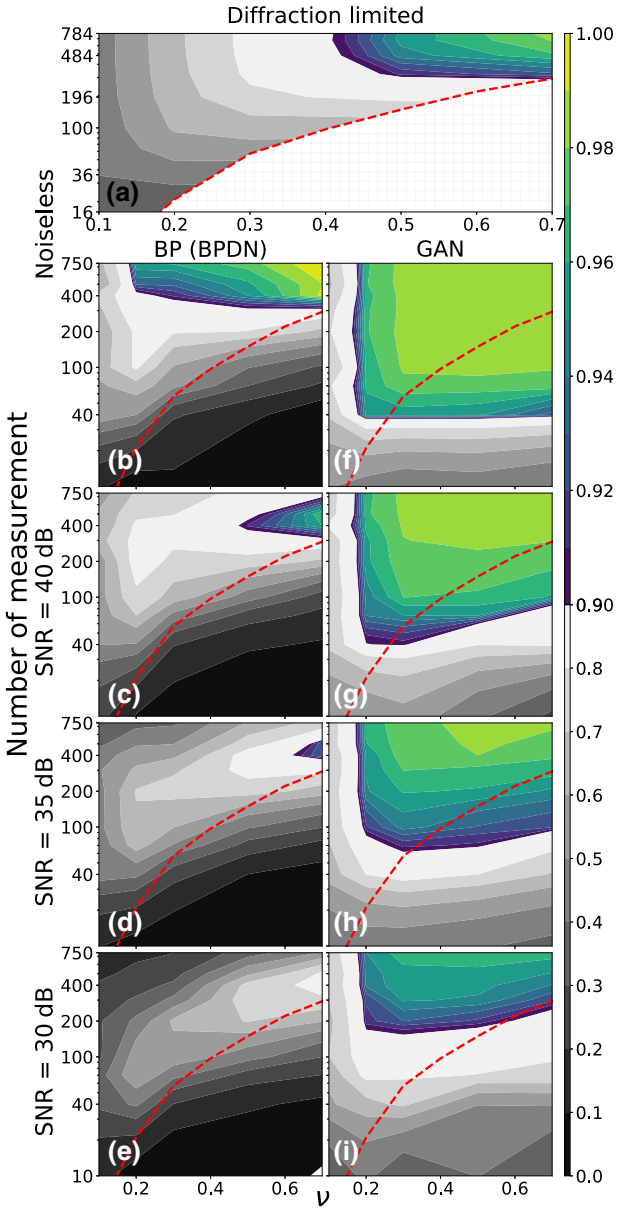

FIG. 4. Correlation between the reconstructed and ground truth images for diffraction-limited imaging, the BP (BPDN with noise) algorithm and GAN framework as functions of cutoff frequency ν and the number of measurement, averaged over the randomly selected digit of each kind $(0-9)$ for noiseless case (a),(b),(f) and for different noise levels: 40 dB (c),(g), 35 dB (d),(h), and 30 dB (e),(i) . The gray scale bar shows the 0–0.9 range and the colorbar shows the detailed 0.9–1 region. The Nyquist limit is indicated by the red dashed line.

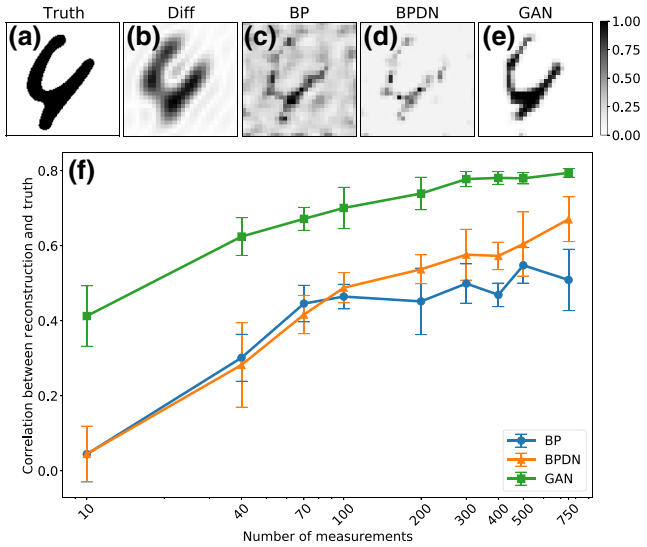

FIG. 5. Example of one validation image with $m = 300$. (a) is the bright-field image of the sample engraved in the aluminum film, which is randomly chosen from the testing dataset. (b) is the diffraction-limited imaging of (a) when $\nu = 0.2$. (c),(d),(e) show the reconstructed image from 300 measurements with BP, BPDN, and GAN, respectively. The average performance of reconstruction on the real samples of BP, BPDN, and GAN algorithms is shown in (f).

the testing subset of the MNIST database are prepared on a microscope slide using maskless UV photolithography (365 nm) and lift off of sputtered reflective 200-nm-thick aluminium film. To generate the ground truth, the separately measured bright-field image of the handwritten digit "4" (randomly selected from the testing dataset) engraved in the aluminum film is cropped and converted to the binary format, as presented in Fig. 5(a).

The detailed description of the experimental setup is provided in Sec. II B. Given the diffraction limit of our experimental setup, the effective cutoff frequency corresponds to $\nu \approx 0.2$. The diffraction-limited image, calculated from the ground truth by applying a low-pass filter with $\nu = 0.2$, is presented in Fig. 5(b). We perform $m = [10, 40, 70, 100, 200, 300, 400, 750]$ random measurements by scanning the input facet of the MMF and randomly choosing m speckle patterns and m corresponding detected intensities to process. The recorded speckle patterns are 4×4 binned to the size 300×300 pixels, cropped (40×40), and resized (28×28) to match the size of dataset images. The processed speckle images form measurement matrix A. We characterize the correlation coefficient between different speckle patterns I_i and I_j by Pearson correlation coefficient r_{ij} and calculate the probability density distribution as presented in Fig. 1(g). The mean value of correlation coefficient is approximately zero, which indicates that the speckle patterns are mostly uncorrelated.

Figures 5(c)–5(e) show the imaging results for $m = 300$ random measurements and the reconstruction performed by the BP algorithm (c), the BPDN algorithm (d), and the proposed GAN (e). The GAN framework provides significantly better image quality compared to diffraction-limited microscopy and traditional compressive sensing algorithms. The GAN approach has a stronger

resemblance to the ground truth and demonstrates noiseless background.

To characterize the imaging performance, the correlation coefficients between the reconstructed images and the ground truth are calculated and presented in Fig. 5(f) for a different number of measurements m. For each m, we repeat the experiment and the reconstruction procedure 5 times and average the correlation coefficient over all realizations, calculating the mean and the standard deviation. For the BPDN, the error-tolerance factor δ is tuned to achieve the best quality of reconstruction for each number of measurements. As expected, for all the algorithms the quality of reconstruction increases with the number of measurements. The GAN approach shows better performance compared to the BP and BPDN algorithms, which is in full agreement with the results of the simulation. Meanwhile, the GAN approach shows consistently lower standard deviation comparing to the BP and BPDN algorithms, which indicates its stable performance. The BPDN algorithm performs better than BP, which can be explained by the fact that real measurements always contain noise that is not taken into account in the BP formulation, Eq. (4).

VI. DISCUSSION AND CONCLUSION

In this study, we experimentally demonstrate compressive fiber imaging with a deep convolutional GAN framework. State-of-the-art methods of machine learning allow reconstruction of images distorted by the MMF only in the case of coherent imaging [16–20]. Whereas the proposed GAN approach has no restrictions and can be used for popular incoherent methods of fiber imaging such as (auto)fluorescent microscopy. Moreover, the GAN framework does not require repetition of a training procedure with pairs of input and output images for different fiber configurations.

With the proposed GAN approach, our simulations demonstrate fiber-based compressive imaging with the 3.5× resolution enhancement. We show that the imaging quality outperforms diffraction-limited imaging approaches. The key component, compact optical fiber, ensures the potential of the proposed approach for endoscopic applications [36]. In contrast to conventional minimization algorithms typically used for reconstruction of compressed data with subdiffraction resolution, our approach does not require the samples to be sparse. The sparsity constraint is very general and can, in principle, be applied to many natural images. However, finding the sparsity domain for a certain sample is not always straightforward. The nature behind the resolution enhancement in compressive imaging is different from the structured illumination microscopy. In the case of fiber-based computational compressive imaging, the resolution is limited by the sparsity of the sample and the noise level [37].

We theoretically and experimentally demonstrated that the proposed GAN-based computational framework can achieve an enhanced quality of reconstruction compared to the traditional compressed sensing minimization algorithms. These results are more pronounced for the small number of measurements, while for the large number of measurements the GAN and the traditional BPDN perform similarly. The GAN is also proved to be tolerant to a large amount of noise. Moreover, the GAN in this study is capable of reconstructing images of handwritten letters as well. Hence the well-trained GAN can also be applied to reconstruct images from other testing datasets, provided the samples have similar features.

The GAN requires a properly tuned network architecture and the training process of the GAN demands a lot of computational power. The computational complexity can be reduced by either reducing the size of the training set or increasing the size of the latent space. However, both procedures may deteriorate the performance of the generator. In our study the size of the latent space is 100, which is less than an order of magnitude smaller than the overall number of pixels $n = 784$. We also perform the simulations with the latent space size of 50 and it also gives similar promising results.

No additional training is needed to make the transition from one experimental setup to another, as the GAN is trained to images without experimental imperfections. The approach can be easily generalized to any speckle-based single-pixel detection setup of imaging without retraining the GAN. As a result, the proposed GAN approach paves the way towards a broad implementation area with a subdiffraction imaging quality not limited by a sparsity of a sample. At the same time, the limitations and challenges cannot be ignored in the application of our GAN approach. The training dataset needs to capture the specific features of the samples of interest, which might limit the flexibility of the network. Also, for certain applications it might be challenging to obtain the required amount of training data. We suggest that efficiency and achievable enhancement of the GAN approach should still be studied in more detail for higher-resolution images.

ACKNOWLEDGMENTS

This work is 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 acknowledge the support from Nederlandse Organisatie voor Wetenschappelijk Onderzoek (WISE). We thank Sergey Amitonov (TU Delft) for custom-made samples, Mark Mol (ARCNL) and Marco Seynen (AMOLF) for their technical and programming support.

[1] I. M. Vellekoop, Feedback-based wavefront shaping, Opt. Express **23**, 12189 (2015).

[2] T. Čižmár and K. Dholakia, Shaping the light transmission through a multimode optical fibre: complex transformation analysis and applications in biophotonics, Opt. Express **19**, 18871 (2011).

[3] R. Di Leonardo and S. Bianchi, Hologram transmission through multi-mode optical fibers, Opt. Express **19**, 247 (2011).

[4] I. M. Vellekoop, A. Lagendijk, and A. Mosk, Exploiting disorder for perfect focusing, Nat. Photonics **4**, 320 (2010).

[5] T. Čižmár and K. Dholakia, Exploiting multimode waveguides for pure fibre-based imaging, Nat. Commun. **3**, 1 (2012).

[6] M. J. Padgett and R. W. Boyd, An introduction to ghost imaging: Quantum and classical, Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. **375**, 20160233 (2017).

[7] L. V. Amitonova and J. F. De Boer, Compressive imaging through a multimode fiber, Opt. Lett. **43**, 5427 (2018).

[8] L. V. Amitonova and J. F. de Boer, Endo-microscopy beyond the Abbe and Nyquist limits, Light: Sci. Appl. **9**, 1 (2020).

[9] M. Pascucci, S. Ganesan, A. Tripathi, O. Katz, V. Emiliani, and M. Guillon, Compressive three-dimensional super-resolution microscopy with speckle-saturated fluorescence excitation, Nat. Commun. **10**, 1 (2019).

[10] G. Calisesi, A. Ghezzi, D. Ancora, C. D'Andrea, G. Valentini, A. Farina, and A. Bassi, Compressed sensing in fluorescence microscopy, Prog. Biophys. Mol. Biol. **168**, 66 (2021).

[11] G. Barbastathis, A. Ozcan, and G. Situ, On the use of deep learning for computational imaging, Optica **6**, 921 (2019).

[12] L. Li, H. Ruan, C. Liu, Y. Li, Y. Shuang, A. Alù, C.-W. Qiu, and T. J. Cui, Machine-learning reprogrammable metasurface imager, Nat. Commun. **10**, 1 (2019).

[13] M. Del Hougne, S. Gigan, and P. Del Hougne, Deeply subwavelength localization with reverberation-coded aperture, Phys. Rev. Lett. **127**, 043903 (2021).

[14] M. A. B. Abbasi, M. O. Akinsolu, B. Liu, O. Yurduseven, V. F. Fusco, and M. A. Imran, Machine learning-assisted lens-loaded cavity response optimization for improved direction-of-arrival estimation, Sci. Rep. **12**, 1 (2022).

[15] M. W. Matthès, Y. Bromberg, J. de Rosny, and S. M. Popoff, Learning and Avoiding Disorder in Multimode Fibers, Phys. Rev. X **11**, 021060 (2021).

[16] B. Rahmani, D. Loterie, G. Konstantinou, D. Psaltis, and C. Moser, Multimode optical fiber transmission with a deep learning network, Light: Sci. Appl. **7**, 69 (2018).

[17] H. Chen, Z. He, Z. Zhang, Y. Geng, and W. Yu, Binary amplitude-only image reconstruction through a MMF based on an AE-SNN combined deep learning model, Opt. Express **28**, 30048 (2020).

[18] N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, Learning to see through multimode fibers, Optica **5**, 960 (2018).

[19] P. Fan, T. Zhao, and L. Su, Deep learning the high variability and randomness inside multimode fibers, Opt. Express **27**, 20241 (2019).

[20] E. Kakkava, B. Rahmani, N. Borhani, U. Teğin, D. Loterie, G. Konstantinou, C. Moser, and D. Psaltis, Imaging through multimode fibers using deep learning: The effects of intensity versus holographic recording of the speckle pattern, Opt. Fiber Technol. **52**, 101985 (2019).

[21] A. Bora, A. Jalal, E. Price, and A. G. Dimakis, in *International Conference on Machine Learning* (PMLR, Sydney, Australia, 2017), p. 537.

[22] Y. Wu, M. Rosca, and T. Lillicrap, in *International Conference on Machine Learning* (PMLR, Long Beach, California, 2019), p. 6850.

[23] M. Kabkab, P. Samangouei, and R. Chellappa, in *Proceedings of the AAAI Conference on Artificial Intelligence* (2018).

[24] W. Gao, Q.-R. Yan, H.-L. Zhou, S.-T. Yang, Z.-Y. Fang, and Y.-H. Wang, Single photon counting compressive imaging using a generative model optimized via sampling and transfer learning, Opt. Express **29**, 5552 (2021).

[25] N. Karim and N. Rahnavard, SPI-GAN: Towards single-pixel imaging through generative adversarial network, arXiv preprint arXiv:2107.01330 (2021).

[26] Y. Ni, D. Zhou, S. Yuan, X. Bai, Z. Xu, J. Chen, C. Li, and X. Zhou, Color computational ghost imaging based on a generative adversarial network, Opt. Lett. **46**, 1840 (2021).

[27] J. W. Goodman, *Statistical Optics* (John Wiley & Sons, 2015).

[28] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, Gradient-based learning applied to document recognition, Proc. IEEE **86**, 2278 (1998).

[29] A. Radford, L. Metz, and S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks, arXiv preprint arXiv:1511.06434 (2015).

[30] TensorFlow, Deep convolutional generative adversarial network, https://www.tensorflow.org/tutorials/generative/dcgan (2021).

[31] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).

[32] P. Bojanowski, A. Joulin, D. Lopez-Pas, and A. Szlam, in *Proceedings of the 35th International Conference on Machine Learning*, Proceedings of Machine Learning Research, Vol. 80, edited by J. Dy and A. Krause (PMLR, Stockholm, Sweden, 2018), p. 600.

[33] TensorFlow, Record operations for automatic differentiation, https://www.tensorflow.org/api˙docs/python/tf/GradientTape (2021).

[34] E. Van den Berg and M. P. Friedlander, Sparse optimization with least-squares constraints, SIAM J. Optim. **21**, 1201 (2011).

[35] E. Van den Berg and M. P. Friedlander, Spgl1: A solver for large-scale sparse reconstruction., https://pypi.org/project/spgl1/ (2021).

[36] B. Lochocki, M. M. Verweg, J. J. Hoozemans, J. F. de Boer, and L. Amitonova, Epi-fluorescence imaging of the human brain though a multimode fiber, APL Photonics **7**, 071301 (2022).

[37] B. Lochocki, K. Abrashitova, J. F. de Boer, and L. V. Amitonova, Ultimate resolution limits of speckle-based compressive imaging, Opt. Express **29**, 3943 (2021).