

# Deep Learning Enabled Scalable Calibration of a Dynamically Deformed Multimode Fiber

Pengfei Fan,\* Yufei Wang, Michael Ruddlesden, Xuechun Wang, Mohamed A Thaha, Jiasong Sun, Chao Zuo,\* and Lei Su

Multimode fibers (MMF) are miniaturized, flexible, and high-capacity information channels, promising to open up new applications in endoscopic imaging. However, precise light control through an MMF with continuous deformations is still a challenge. Here, a scalable calibration framework for a dynamically deformed MMF using deep learning is proposed. The proof-of-concept experiments demonstrate that the proposed continual generative adversarial model has the ability to characterize the MMF transmission states sequentially and detect the fiber deformation using proximal reflection in real-time synchronously, allowing self-adaptively cross-state focusing through a semi-flexible MMF without distal access after the scalable calibration. This framework is a continual learning scheme under extreme memory constraints where the model is able to synthesize training data and prevent forgetting the previously learned bending states. The proposed method paves the way for the experimental realization of scalable calibration of a dynamically deformed MMF.

and high-capacity information channels that may meet these extremely high demands due to their small diameter down to tens of microns and their ability to bend into acute angles.<sup>[1–9]</sup> However, the nature of MMF transmission leads to the scrambling of incident wavefronts resulting in random speckle patterns at the fiber output. A number of techniques in the adaptive optics domain have recently been developed to overcome this transmission degradation and to permit the desired light control through MMF.<sup>[3,10–12]</sup> Recently, advances in complex modulation of the phase or the intensity of a light beam were enabled by the development of light-shaping hardware such as spatial light modulators (SLM) and digital micromirror devices (DMD).<sup>[13,14]</sup> Subsequently, the transmission matrix (TM) measurement

of MMFs was explored,<sup>[2,3,8,9,15,16]</sup> allowing potential endoscopic applications.<sup>[8,9,17,18]</sup>

The MMF transmission state is highly sensitive to external perturbations and environmental changes.<sup>[19–21]</sup> When the fiber is disturbed, the transmission state changes, and the precalibrated TM fails to remain valid to the new MMF transmission states. Therefore, a completely rigid MMF endoscope was usually

## 1. Introduction

Miniaturized ultrathin endoscopes play a vital role in minimally invasive surgery and biological applications such as *in vivo* fluorescence microscopy. There is a need to develop nonrigid endoscopes to access different body cavities through a small incision. Multimode fibers (MMF) are miniaturized, flexible,

P. Fan, Y. Wang, M. Ruddlesden, X. Wang, L. Su  
School of Engineering and Materials Science  
Queen Mary University of London  
London E1 4NS, UK  
E-mail: p.fan@qmul.ac.uk

P. Fan, J. Sun, C. Zuo  
School of Electronic and Optical Engineering  
Nanjing University of Science and Technology  
Nanjing 210094, China  
E-mail: zuochao@njust.edu.cn

P. Fan, J. Sun, C. Zuo  
Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense  
Nanjing University of Science and Technology  
Nanjing 210094, China

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/adpr.202100304>.

© 2022 The Authors. Advanced Photonics Research published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/adpr.202100304

P. Fan, J. Sun, C. Zuo  
Smart Computational Imaging (SCI) Laboratory  
Nanjing University of Science and Technology  
Nanjing 210094, China

M. A. Thaha  
Blizard Institute  
National Bowel Research Centre  
Barts and the London School of Medicine & Dentistry  
Queen Mary University of London  
London E1 2AT, UK

M. A. Thaha  
Department of Colorectal Surgery  
Royal London Hospital  
Barts Health NHS Trust  
London E1 1FR, UK

used to overcome this limitation, where the fiber was calibrated once for one specific spatial conformation.<sup>[8,9,22,23]</sup> Nevertheless, a flexible or semi-flexible endoscope is necessary for some applications. In most biomedical applications, inserting the MMF into deep tissue would induce inevitable shape and temperature changes. Real-time compensation for distortions was proposed in ref.,<sup>[24]</sup> but additional feedback hardware was still needed. It was also reported that imaging with a fiber bent within a restricted radius of curvature range was achieved by using a particular S-shaped configuration.<sup>[25]</sup> A detailed analysis of the propagation-invariant modes within TMs was conducted but only using graded-index MMFs.<sup>[26]</sup> Recently, more compressive sensing MMF imaging schemes were proposed toward robustness of light transport or fast imaging speed.<sup>[27–32]</sup>

On the other hand, it has been widely recognized that reflected light transmitted back through the fiber is beneficial to detect fiber deformation and is promising for imaging without distal access,<sup>[33–37]</sup> which is critical for practical MMF imaging. One early attempt was by using a virtual coherent point light source placed at the distal fiber tip to dynamically compensate for bending, and the light was focused through a semi-flexible MMF for which the number of conformation is limited.<sup>[33]</sup> Subsequently, focusing was maintained as the fiber was maneuvered to the target site prior to imaging by the addition of a partial reflector to the distal fiber end.<sup>[34]</sup> Recently, Gordon et al. have proposed a method introducing a thin stack of structured metasurface reflectors at the distal facet of the fiber, to characterize MMF TMs for lensless imaging without distal access.<sup>[36]</sup> Also, the transpose relationship between the backward and forward transmission through an MMF was verified in ref.,<sup>[37]</sup> revealing that the direct retrieval of the forward TM from a round-trip measurement was impeded by the symmetry.

Deep learning techniques have been successfully applied to simple geometries (such as MNIST digits, letters) reconstruction<sup>[38–41]</sup> and spatially distributed data transmission<sup>[42,43]</sup> through MMFs. Particularly, it was also suggested that the neural network learned and generalized different transmission states when the MMF was bent or subject to continuous transmission characteristic variations.<sup>[38,39]</sup> Further, it was also reported that transmission of natural scenes through an MMF up to 10 m was achieved by statistically reconstructing the inverse TM for the fiber.<sup>[44]</sup> Turpin et al. demonstrated light scattering control in transmission and reflection with neural networks.<sup>[45]</sup>

To precisely control the light propagation using existing algorithms through a dynamically deformed MMF is very challenging. First, different from other scattering media, the transmission properties of MMF are extremely susceptible to the fiber's deformation. A precalibrated TM or a fixed neural network for the MMF transmission can only be applicable to the current fiber state. Second, previous studies about continuous MMF transmission variations have an assumption that these schemes have full access to data collected previously as the fiber is dynamically deformed,<sup>[39]</sup> which was impractical due to the limited memory on the machine. Third, most of the methods described above require access to the distal facet of the fiber, which is not feasible in most realistic usage scenarios without bulky distal optics.

Here, we propose and demonstrate a scalable calibration framework using deep learning for a dynamically deformed MMF under

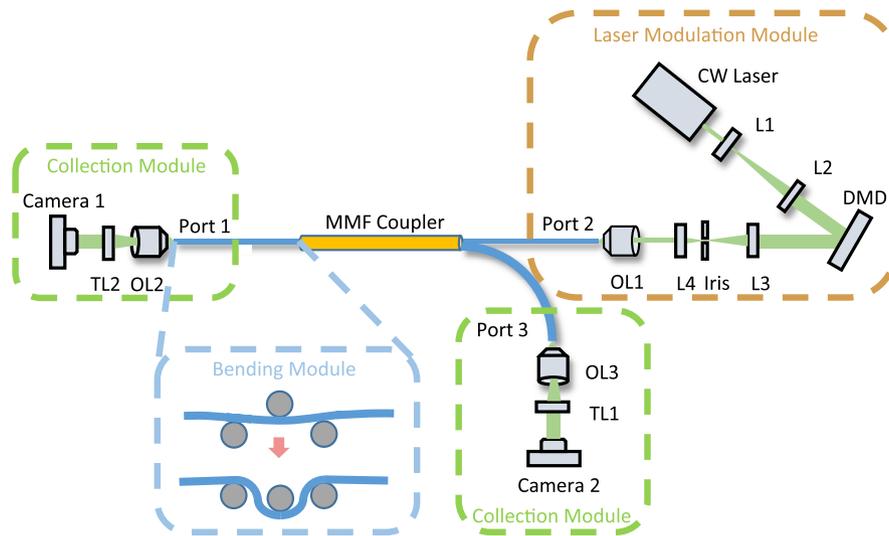
extreme memory constraints (EMC).<sup>[46,47]</sup> We term this deep-learning-based approach DI-GAN (Generative Adversarial Network for Deep Imaging) and use it to characterize the MMF transmission states sequentially and detect the fiber deformation in real-time synchronously, enabling self-adaptively cross-state focusing through a semi-flexible MMF without distal access after the scalable calibration. DI-GAN, modified based on CVAE-GAN,<sup>[48]</sup> is trained using a conditional GAN in a continual fashion with mutually matched triplets of 1) various randomly generated input patterns coupled into the proximal fiber end, 2) the corresponding transmitted speckle images, and 3) reflected speckle images captured at the distal and proximal ends, respectively, under different bending positions of a dynamically deformed MMF. The network architecture is plotted in Figure S3 (Section S4, Supporting Information) trained using data collected under the experimental setup shown in **Figure 1**. This framework accurately reconstructs the input patterns from the transmitted speckle images through a fiber with continuous deformations, naturally leading to the calibration of the fiber at the corresponding bending position and enabling the input excitation wavefront prediction for focusing light at the distal tip of the fiber. We also leverage the reflection information transmitted back to the proximal end of the fiber to detect the fiber conformation in real-time, and then spots can be self-adaptively generated on the distal fiber facet by projecting the corresponding input patterns inferred by the network without distal access. Therefore, access to the distal end (transmitted speckle images collection) is only needed during the training stage. But after that training, access to the distal end is not necessary as there is access to the reflection. For fully utilizing the reflection information and simplifying the setup, our proof-of-concept experiment deploys an MMF coupler instead of a standard MMF nor a reflector (Figure S2, more details in Section S3, Supporting Information).

The remainder of this paper is organized as follows. First, using DI-GAN, we implement the scalable calibration of a semi-flexible MMF under EMC by investigating the input–output relationships under different MMF transmission states. Second, we demonstrate DI-GAN-based transmission inference using reflection, which indicates the deterministic one-to-one mapping between the transmission and the reflection over the same input pattern. Third, we utilize the reflected speckle image to identify the current bending state and achieve self-adaptively focusing at an arbitrary state before and after MMF deformation. Furthermore, another important feature of DI-GAN is shown that it enables focusing performance monitoring at the proximal end without distal access.

## 2. Experimental Section

### 2.1. Experimental Setup

A schematic diagram of our experimental setup is depicted in Figure 1. The experiments are performed on an MMF coupler (Thorlabs, TM50R5F1A). Each port of the MMF coupler has a 0.8-m length fiber lead with a core diameter of 50  $\mu\text{m}$  and NA = 0.22. The 1  $\times$  2 MMF coupler is designed to split light between Port 2 and Port 3 with a 50:50 coupling ratio. In our setup, we use Port 2 as our signal input and Port 1 as a transmission output. Port 3 is used to capture the reflected light from



**Figure 1.** The experimental setup used to obtain the data under different bending states. The beam is expanded, collimated, and directed onto the DMD, the reflection of which is coupled into Port 2 of the MMF coupler. The transmission output from Port 1 of the MMF coupler is captured by Camera 1. Simultaneously, the reflection output from Port 3 of the MMF coupler is captured by Camera 2. The MMF pigtail in Port 1 is bent using three opposing optical posts, the middle one of which is mounted on a single-axis translation stage (shown in the inset below). L1-L4: bi-convex lenses; TL1, TL2: tube lens; DMD: digital micromirror device; OL1-OL3: objective lens; MMF: multimode fiber.

Port 1 using Camera 2. The system consists of three modules: 1) laser modulation module; 2) collection module; and 3) bending module. In the laser modulation module, a collimated laser beam from a continuous-wave diode-pumped laser (532 nm, Cobolt Samba 50) is expanded to match the area of our DMD (Vialux V-7001,  $\approx 22$  kHz). The DMD can modulate the laser beam into desired random patterns with binary spatial amplitudes. Each pixel of the input pattern occupies  $4 \times 4$  DMD micromirror pixels. For example, a pattern of  $24 \times 24$ -pixel can be obtained by maintaining the total micromirror pixel numbers at  $96 \times 96$ . Lenses L1, L2, L3, and L4 are placed in a  $4f$  setup to project the modulated beam to the back focal plane of a microscope objective lens OL1 (Nikon CFI Plan Achrom 20X, NA = 0.4). A pinhole filters the first diffraction order in the Fourier plane of L3, blocking the remaining orders. The beam is then coupled into Port 2 of the MMF coupler by OL1. In the collection module, the output from the fiber facet of Port 1 and Port 3 is both collected and collimated by a microscope objective (OL2, OL3, Olympus PLN 20X, NA = 0.4) and tube lens (TL1, TL2, Thorlabs AC254-200-A-ML). Camera 1 (QImaging optiMOS) and Camera 2 (JAI SP-500M-CXP4) are triggered simultaneously by the signal from the DMD. This allows two cameras to capture corresponding speckle images from the fiber facet when the DMD projects random patterns. In the bending module, as shown in Figure 1, the pigtail of the MMF in Port 1 is initially placed between three opposing optical posts. The post in the middle is mounted on a single-axis translation stage. The bending state of the MMF can be changed by tuning the position of the post at a displacement step of 1 mm. In our proof-of-concept experiments, the data sets are collected when the fiber pigtail is sequentially bent to 10 different positions (from 0 to 9 mm). To perform more continuous control of fiber deformation, one can use a three-axis motorized translation stage (more discussions in

Section 4). Examples of transmitted and reflected speckle images are illustrated in Figure S1a and S2a in Supporting Information, respectively.

## 2.2. Data Preprocessing and Preparation of Training Data

The system shown in Figure 1 operates at 200 frames per second (i.e., 200 input-transmission-reflection triplets per second), and 10 000 transmitted speckle images and reflected speckle images are collected for each fiber bending state sequentially. The 10 000 input patterns displayed on the DMD are randomly generated with an 'ON' to 'OFF' pixel ratio of 50:50 (implemented using random.randint function within the Python Numpy Library) in a  $24 \times 24$  square configuration. Both transmission and reflection images are then downsampled to  $96 \times 96$  pixels, using resize function with INTER\_AREA interpolation within the Python cv2 Library to speed up training and reduce computer memory usage, without losing image information.<sup>[49]</sup> The dynamic range of image intensity is subsequently normalized to  $[-1, 1]$  in order to stabilize the training of the GAN network. For each data set collected at each bending state, 90% of the data (i.e., 9000 data triplets) are randomly selected for training using the train\_test\_split function within the Python Scikit-learn Library, and the remaining 10% of data form the testing data set.

## 2.3. Continual Generative Adversarial Model under EMC: DI-GAN

The proposal focuses on the accessibility to training data and the scalability of the model. This issue is often termed continual learning or incremental learning.<sup>[50]</sup> One of the most straightforward solutions for this problem is sequential fine tuning (SFT), where the network is trained sequentially with a sequence of

independent tasks. The parameters of a network are first trained on previous data sets and then are fine tuned using the new data set to learn the current task. Nevertheless, this method will inevitably result in another problem, catastrophic forgetting,<sup>[51]</sup> where the network forgets its previous learning task. The reason is straightforward: the parameters of the network (optimized using the previous data set) have been tuned to adapt to the latest data distribution. Numerous efforts have been made to design network architectures and training algorithms that can prevent or at least alleviate catastrophic forgetting.<sup>[52]</sup> Based on how the task data are stored and utilized through the sequential learning process, there are three types of families to prevent forgetting: replay methods,<sup>[53–55]</sup> regularization-based methods,<sup>[51,56]</sup> and parameter isolation methods.<sup>[57,58]</sup> Inspired by these previous studies and given the impossibility of storing all data collected during the MMF bending, we propose our continual generative adversarial model under EMC. Instead of learning the different bending states of the fiber jointly at the same time in a single training process under the assumption that all data for training are available beforehand,<sup>[39]</sup> we develop a continual learning scheme under EMC where the model is able to synthesize training data for itself and the memory usage is extremely limited during the course of training. Compared with the conventional neural network, which results in forgetting the previous after transfer learning, our proposed scheme permits continual generalization of all bending states after the scalable calibration process.

As shown in **Figure 2**, the DI-GAN updating is based on the mechanism of replaying memories from previous transmission states to consolidate them while learning new ones, preventing the models from forgetting the previous bending states.<sup>[59]</sup> Here, the memories refer to the images data synthesized by the previously trained model, differing from computer memory. If the synthesized images data match the data distribution of the ground truth very well, there is no need to allocate computer memory for data storage, and the training of the models satisfies the EMC. In our case, the training data set in each round contains 9000 newly collected data for the current bending state and  $9000 \times i$  network-synthesized data for the previous bending states, where  $i$  means the number of previous bending positions. For example, shown in **Figure 2**, after training using the data set collected at the first bending position, the models have had the ability to synthesize data of transmission and reflection,

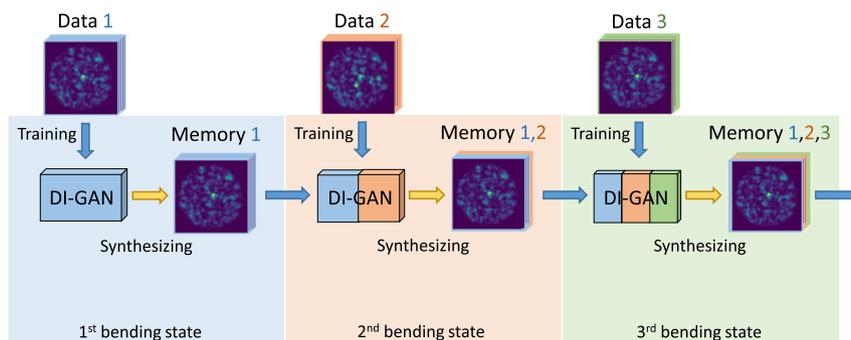
matching the data distribution of the first bending position very well. The fiber is then bent to the second position, so another data set is collected. In the next round of training, the newly collected data and the data synthesized by the previously trained models will be jointly used to train the new models. The training process continues until the MMF dynamic bending ends, and then the networks are able to generalize all bending states. See more details in Section S4 and S5, Supporting Information.

### 3. Results

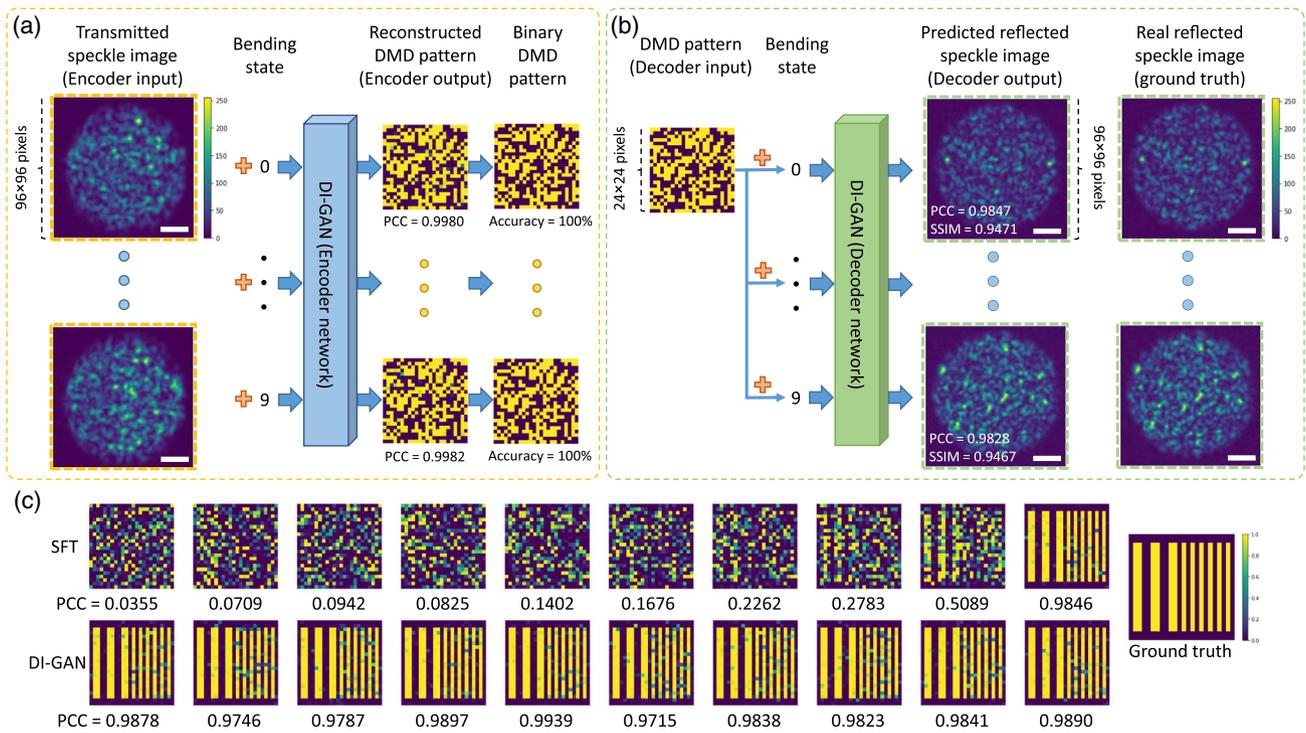
#### 3.1. Scalable Calibration of a Semi-flexible MMF Using DI-GAN

**Figure 3a** demonstrates DI-GAN-based DMD pattern reconstruction with high accuracy at different bending states (represented by digits from '0' to '9'), each one of which indicates the fiber conformation distance curved from its original position. By inputting the transmitted speckle images measured at the corresponding bending states into the DI-GAN (Encoder network), we recover the DMD patterns coupled into the fiber with  $\approx 100\%$  accuracy. Note that the high fidelity of the reconstructed DMD input patterns confirms the accurate characterization of the fiber with altered positions. To further quantify DI-GAN (Encoder network) output, we calculate various average performance evaluations (see details in Section S1, Supporting Information); **Table 1** illustrates that the reconstructed DMD patterns calculated from transmitted speckle images agree with the true binary DMD patterns at each MMF state very well.

Next, we apply DI-GAN to the data synthesis task for training future models continuously without forgetting the previous learning. Detailed in Section 2.3, the continual generative adversarial model, DI-GAN, is trained using the data triplets of images sequentially without direct access to the previously collected data under EMC. For these data triplets, we can easily generate a mass of random DMD patterns, and the trained DI-GAN will quickly synthesize realistic reflected speckle images (using Decoder network) and transmitted speckle image (using Generator network) over the corresponding bending states. The synthesized data triplets for the previous bending states are combined with the newly collected data triplets for the current state, and the combined outcome form the training data set for the next scalable calibration process. **Figure 3b** demonstrates that the predicted



**Figure 2.** Illustration of the continual learning scheme under EMC. Training is first done on the data set available at the first bending state. After that, when the fiber is bent to the second position, the new data set is collected and used together with the network-synthesized data to train the network. The training process continues until the MMF dynamic bending ends, and then the networks are able to generalize all bending states.



**Figure 3.** Scalable calibration of a semi-flexible MMF using DI-GAN. a) DI-GAN-based DMD pattern reconstruction with high accuracy ( $\approx 100\%$ ) at different bending states. By appending a bending state (digits from ‘0’ to ‘9’) to a transmitted speckle image (left) and passing it through a trained DI-GAN (Encoder network), a reconstructed DMD pattern (middle) can be accurately obtained. The output image is then processed to get a binary DMD pattern (right) using a threshold of 0.5. The Pearson correlation coefficients (PCC) of the predicted DMD pattern and the accuracy of the binary DMD pattern compared with the ground truth pattern are shown. b) From one same DMD pattern (left), diverse reflected speckle images (middle) synthesized at different bending states using DI-GAN (decoder network). The ground truth reflected speckle images (right) are shown for comparison, and the PCC and structural similarity index (SSIM) between the synthesis and ground truth are also given below. As in (a), color scale indicates intensity, which is normalized to a dynamic range of [0, 255] for better visualization. Evaluations were repeated with 1000 images, achieving similar results. Scale bars, 10  $\mu\text{m}$ . c) Bar examples of the scalability differences between SFT and the proposal, showing the reconstruction of standard bars pattern with different frequencies using SFT and DI-GAN, after sequentially training the ten tasks (left to right: 0 mm to 9 mm). The ground truth binary pattern of standard bars is shown (rightmost).

**Table 1.** Average accuracy, F1\_score, mean square error (MSE), PCC, SSIM, and output runtime for DMD patterns reconstruction at different bending states.

Bending state [mm]	Accuracy [%]	F1_score	MSE	PCC	SSIM	Network output runtime [s]
0	99.8052	0.9980	0.0032	0.9940	0.9924	0.0002
1	99.8323	0.9983	0.0026	0.9950	0.9940	0.0002
2	99.8038	0.9980	0.0030	0.9944	0.9930	0.0002
3	99.8217	0.9982	0.0027	0.9950	0.9939	0.0002
4	99.8569	0.9986	0.0024	0.9956	0.9945	0.0002
5	99.8018	0.9980	0.0031	0.9941	0.9927	0.0002
6	99.8007	0.9980	0.0031	0.9942	0.9928	0.0002
7	99.6800	0.9968	0.0041	0.9921	0.9904	0.0002
8	99.8247	0.9982	0.0028	0.9948	0.9936	0.0002
9	99.8220	0.9982	0.0029	0.9946	0.9932	0.0002

reflected speckle images synthesized from our model (Decoder network) have over 98% PCC and 94% SSIM on average with the ground truth reflected speckle images over ten bending states. It

should be noted that the ten different predicted examples shown in Figure 3b are synthesized over an identical DMD pattern, which indicates that the proposed models are capable of

**Table 2.** Average MSE, PCC, SSIM, and output runtime for reflected speckle image synthesis at different bending states.

Bending state [mm]	MSE	PCC	SSIM	Network output runtime [s]
0	8.8965e-04	0.9831	0.9451	0.0001
1	9.4072e-04	0.9698	0.9208	0.0001
2	8.8988e-04	0.9771	0.9319	0.0001
3	8.9076e-04	0.9795	0.9383	0.0001
4	8.7479e-04	0.9826	0.9442	0.0001
5	8.9429e-04	0.9837	0.9464	0.0001
6	8.8040e-04	0.9819	0.9449	0.0001
7	9.0806e-04	0.9814	0.9433	0.0001
8	9.4311e-04	0.9810	0.9414	0.0001
9	1.1118e-03	0.9764	0.9299	0.0001

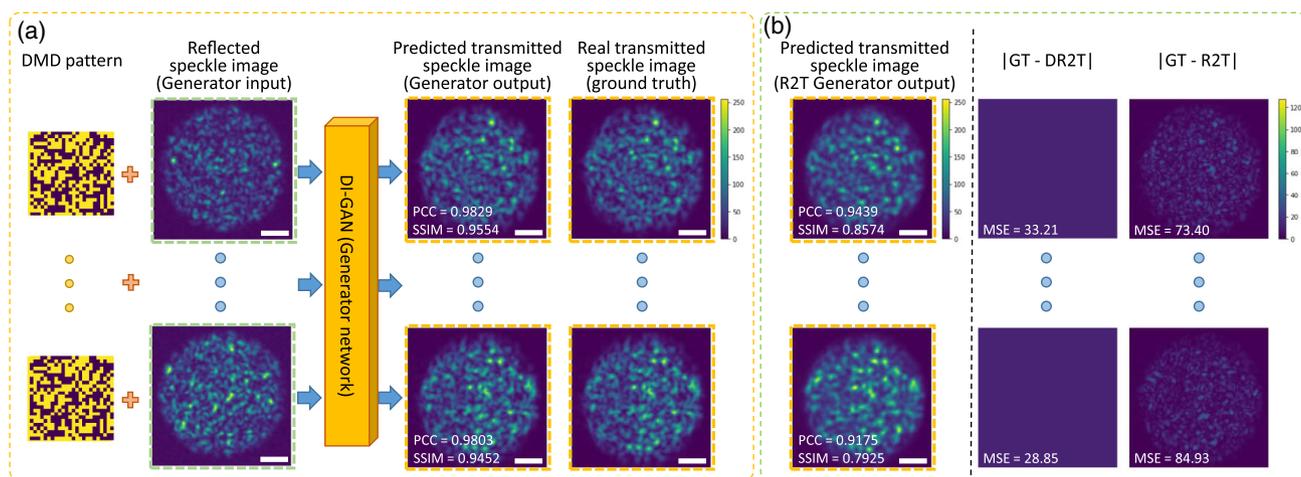
generating realistic and diverse samples very well. More evaluation results are presented in **Table 2**; it is worth mentioning that the MSE values are relatively small because the intensity values of the images have been scaled to the range of [-1, 1] for GAN training. Also, the generation of transmitted speckle images is vital for the training of the continual model and that is presented in Section 3.2.

As illustrated in Figure 3c, we compare the scalability between the proposal and a simple baseline, SFT, by showing the reconstruction of standard bars pattern after sequentially training the ten tasks (left to right: 0 to 9 mm). We observe that SFT completely forgets previous tasks in all data sets it learned before and can only reconstruct the DMD patterns at the last bending

state; while those input patterns recovered by DI-GAN are in general clear and much more recognizable. This demonstrates that our network has a certain extrapolation ability.

### 3.2. DI-GAN-based Transmission Inference Using Reflection

Toward complete data synthesis, DI-GAN can also be used to perform transmitted speckle image inference, where DI-GAN (Generator network) can be trained using data pairs of DMD patterns and reflected speckle images as the model inputs. **Figure 4** demonstrates our blind testing results for DI-GAN-based transmission inference using reflection. The trained Generator network digitally transfers the reflected speckle image into a transmitted speckle image, while at the same time retains high contrast, matching the ground truth to a high degree of similarity. For example, the predicted transmitted speckle image in the upper row of Figure 4a has a PCC of 0.9829 and an SSIM of 0.9554 with the ground truth image. To further qualify the performance of the proposed model architecture (using both DMD pattern and reflection as combined input, termed as DR2T-Generator), we also train a model only using reflection to infer transmission (termed as R2T-Generator). The R2T-Generator-inferred transmitted speckle images are also shown in the left column of Figure 4b for comparison, outputting more blurred images with smaller PCC and SSIM values. The pixels of DR2T-Generator-inferred transmitted speckle images are substantially sharper as compared to the R2T-Generator-inferred ones, providing a good match to the ground truth. The absolute difference images of the DR2T-Generator results and R2T-Generator results with respect to the corresponding ground-truth images are also provided on the right of Figure 4b, with MSE values reported, further demonstrating the success of



**Figure 4.** DI-GAN-based transmitted speckle image inference using reflected speckle image. a) Different reflected speckle images (left) appended with DMD patterns are digitally transferred into transmitted speckle image (middle), the resulting images provide a very good match to the corresponding ground truth images (right). The PCCs and SSIMs of the predicted images calculated with the ground truth are shown. b) For comparison, R2T-Generator outputs (left) are also shown, reporting a substantial blur in each case. As in a colour scale indicates intensity, which is normalized to a dynamic range of [0, 255] for better visualization. The absolute difference images of the DI-GAN (DR2T-Generator) output images and the R2T-Generator output images with reference to the corresponding ground truth (GT) images are also shown, with MSE values, further quantifying the performance of models. The former absolute difference images suggest that the mismatch formed by DI-GAN is negligible, while the latter absolute difference images report increased errors. Color scale for the absolute difference images indicates intensity, which is normalized to a dynamic range of [0, 127] for better visualization. Evaluations were repeated with 1000 images, achieving similar results. Scale bars, 10  $\mu\text{m}$ .

**Table 3.** Average MSE, PCC, SSIM, and output runtime for transmitted speckle image generation at different bending states.

Bending state [mm]	MSE	PCC	SSIM	Network output runtime [s]
0	9.0188e-04	0.9651	0.9027	0.0004
1	6.5470e-04	0.9762	0.9256	0.0004
2	4.4676e-04	0.9885	0.9558	0.0004
3	4.8768e-04	0.9854	0.9482	0.0004
4	4.5858e-04	0.9870	0.9526	0.0004
5	4.5225e-04	0.9883	0.9558	0.0004
6	4.4004e-04	0.9888	0.9558	0.0004
7	4.4614e-04	0.9880	0.9533	0.0004
8	4.7748e-04	0.9862	0.9496	0.0004
9	5.6292e-04	0.9819	0.9389	0.0004

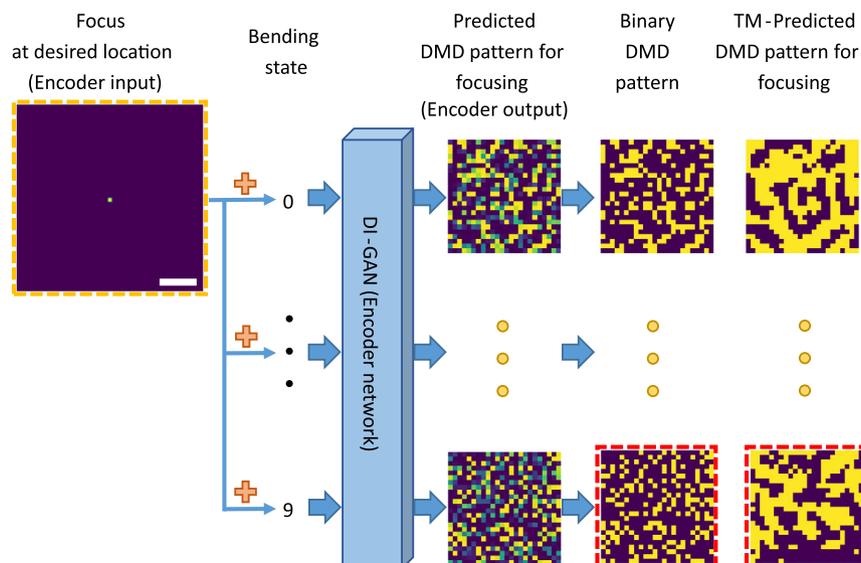
DI-GAN, matching the aim of its training. More evaluation results are presented in Table 3; it is worth mentioning that the MSE values are relatively small because the intensity values of the images have been scaled to the range of  $[-1, 1]$  for GAN training.

So far, in Section 3.1 and 3.2, we have blindly tested both transmission and reflection synthesis using DI-GAN (Decoder network and Generator network, respectively). Specifically, DI-GAN-based transmission inference using reflection indicates the deterministic one-to-one mapping between the transmission and the reflection over the same input pattern. Although not demonstrated here, the reflected speckle image also carries the complete information of the input pattern, which can be recovered through the reflection with high accuracy, too (similar to what has been shown in Figure 3a). We have evaluated the data

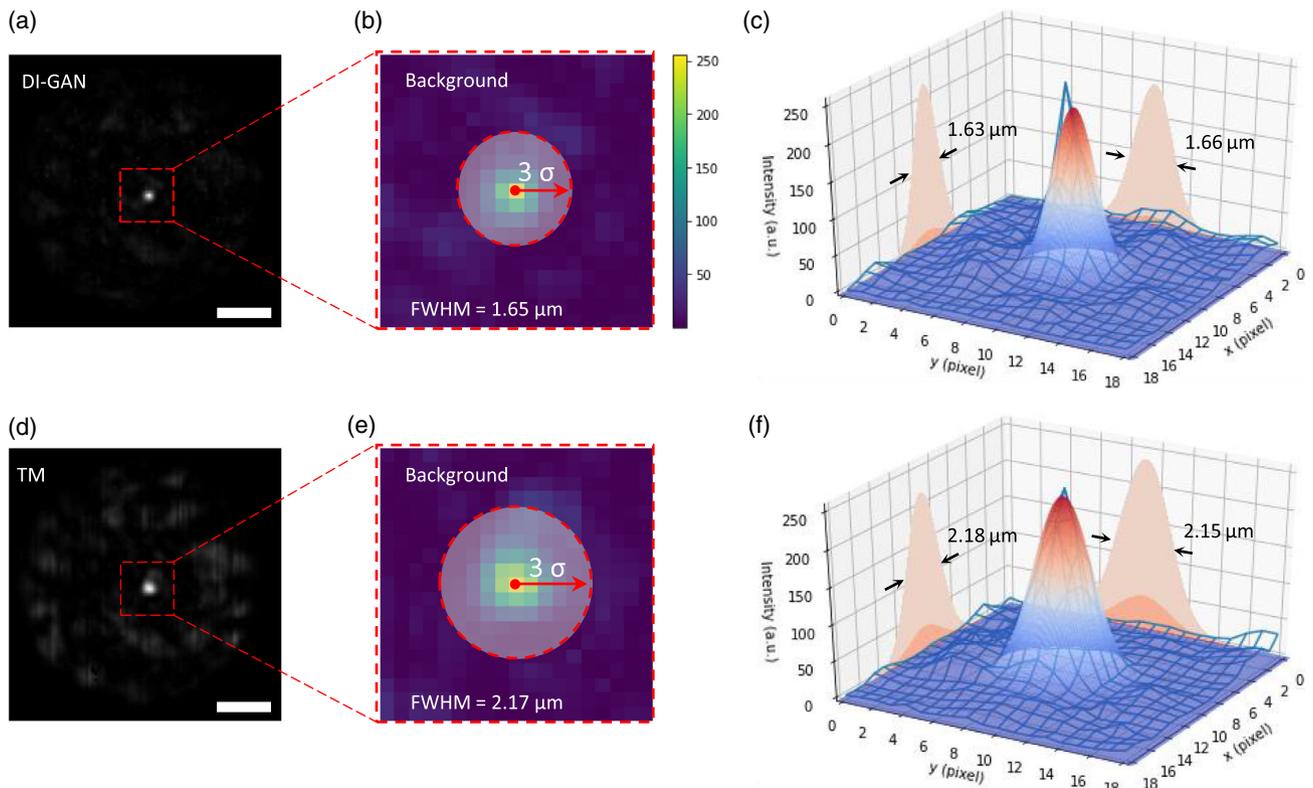
synthesis performance of DI-GAN under different metrics, which reveal the effectiveness and robustness of DI-GAN to the sequential learning task under EMC.

### 3.3. Self-adaptively Cross-state Focusing through a Semi-flexible MMF

In Section 3.1, we have demonstrated the ability of DI-GAN (Encoder network) to reconstruct the DMD input patterns with  $\approx 100\%$  accuracy. As demonstrated in ref., [45] the focus can be generated by using a neural network when the fiber conformation is kept stationary. To highlight the utility of DI-GAN for cross-state focusing through a dynamically deformed MMF, we take advantage of the DI-GAN to output the desired DMD input patterns for spot generation at the distal tip of the fiber. In Figure 3a, the inputs of the DI-GAN (Encoder network) are transmitted speckle images, while here, we directly replace the inputs with images with a focus at a desired location across the speckle. Illustrated in Figure 5, by appending the corresponding bending states, the different DMD patterns for focusing are digitally generated by the network. For one specific bending state, by inputting images with spots at different locations sequentially, the network can quickly output a series of input patterns for focusing. It has also been reported that focus could be generated at the distal end of MMF using TM information.<sup>[60,61]</sup> Specifically, the position of the focus is controlled by choosing the corresponding row in the TM; to implement the complex conjugate in one row of the TM as the DMD input field, the DMD pixels where the real part was positive are selectively turned on. For further comparison, we also show the input patterns generated by TM in the rightmost column in Figure 5. As can be seen in this comparison, the distribution of the pixels across DI-GAN-predicted input patterns is more separate and sparser than the TM-predicted ones, indicating that the DI-GAN enables



**Figure 5.** DMD input patterns generation for cross-state focusing. By appending a bending state (digits from ‘0’ to ‘9’) to an image with a focus at the desired location (left) and passing it through a trained DI-GAN (Encoder network), a predicted DMD input pattern for focusing (middle) can be digitally obtained. The output pattern is then binarily processed to get a binary DMD pattern (right) using a threshold of 0.5. The DMD input patterns predicted by TM are also provided on the right, with denser and larger patches. Scale bars, 10 μm.



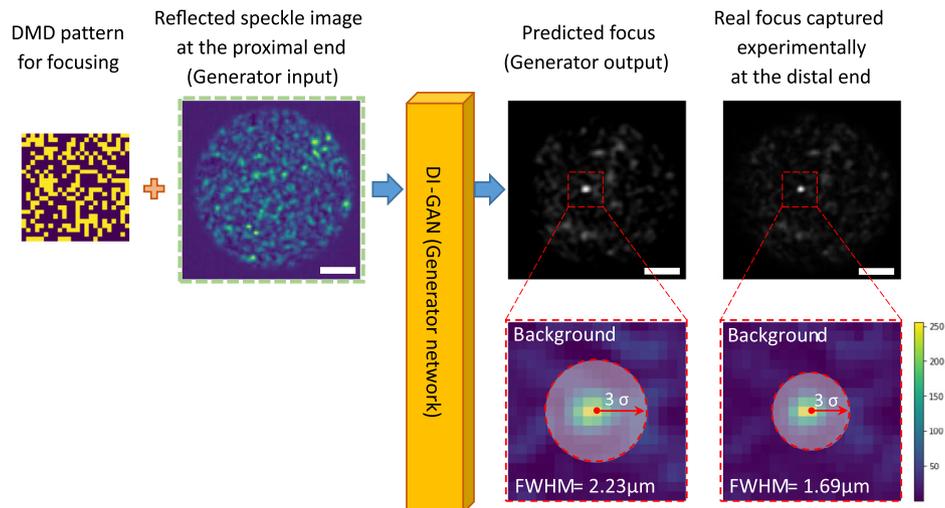
**Figure 6.** Focus generation around the fiber end center with DI-GAN and TM. a) and d) Foci generated by projecting the two DMD input patterns using DI-GAN and TM highlighted in the bottom row in Figure 5. b) and e) Zoomed areas from the ROI (red dotted squares, covering  $\approx 10 \times 10 \mu\text{m}^2$ , i.e.,  $18 \times 18$  pixels) in (a) and (d) with  $3\sigma$  circle area and FWHM values. c) and f) 2D Gaussian fit over (b) and (e). The blue mesh grid shows the intensity profile of the focus, and the multicolored smooth surface represents the intensity 2D Gaussian fit, which is then projected onto the  $xoz$  and  $yoz$  planes, respectively. FWHM values along the  $x$ -axis and  $y$ -axis are also provided. Experiments were repeated over 10 bending states, achieving similar results. Grid pitch:  $0.5537 \mu\text{m pixel}^{-1}$  (see calculation in Section S2, Supporting Information). Scale bars,  $10 \mu\text{m}$ .

to control light more precisely by providing better modulation patterns.

Next, we project the input patterns predicted by DI-GAN and TM onto the DMD, respectively, and observe the focus images captured at the distal fiber end. **Figure 6** illustrates the focusing performance of these two methods (two DMD patterns highlighted using red dotted squares from the bottom row in Figure 5), showing an example of a focus that is around the fiber end center. The first row in Figure 6a–c shows the focus evaluation of DI-GAN method, and the second row in Figure 6d–f shows the focus evaluation of TM method. For comparison, we calculated the enhancement factor (EF, detailed in Section S2, Supporting Information) of these two focuses. It has shown that the EF of DI-GAN-generated focus (about 74.45) is nearly double that of the TM-generated one (about 39.39). Besides, the full width at half maximum (FWHM, detailed in Section S2, Supporting Information) of the focus is calculated by performing a 2D Gaussian fit over the region of interest (ROI) around the focus center that covers  $\approx 10 \times 10 \mu\text{m}^2$  (Figure 6b, e, zoom-in areas in the red dotted squares from Figure 6a,d). The intensity profiles of the focuses along the  $x$  and  $y$  axes, plotted by fitting the data with a 2D Gaussian function, are also shown in Figure 6c,f, respectively. For example, in Figure 6c, the intensity profile of the focus in the ROI is shown using a blue

mesh grid, and the multi-colored smooth surface represents the intensity 2D Gaussian fit, which is then projected onto the  $xoz$  and  $yoz$  planes, respectively. It is clear that the focus generated using DI-GAN has smaller FWHM (about  $1.63 \mu\text{m}$  along the  $x$ -axis and  $1.66 \mu\text{m}$  along the  $y$ -axis) than the TM one (about  $2.18 \mu\text{m}$  along the  $x$ -axis and  $2.15 \mu\text{m}$  along the  $y$ -axis). The theoretical value of FWHM (Abbe diffraction limit, calculated as  $\lambda/(2NA)$ ) in our case is around  $1.21 \mu\text{m}$ , and the degradation result in a certain difference from the ideal case. Furthermore, we observe that the DI-GAN has significantly improved the signal-to-noise ratio (SNR, detailed in Section S2, Supporting Information) of the focus, from 26.06 to 52.58, more than twice the SNR of TM-generated one. In addition to focusing at the center of the speckle, a video showing the scanning with multiple spots enabled by the TM, and the proposed model can be found via the link given in Supporting Information as ref., [5] demonstrating that the CNN model has high enhancement and darker background.

DI-GAN not only substantially surpasses the TM in focusing ability but also permits focusing performance monitoring at the

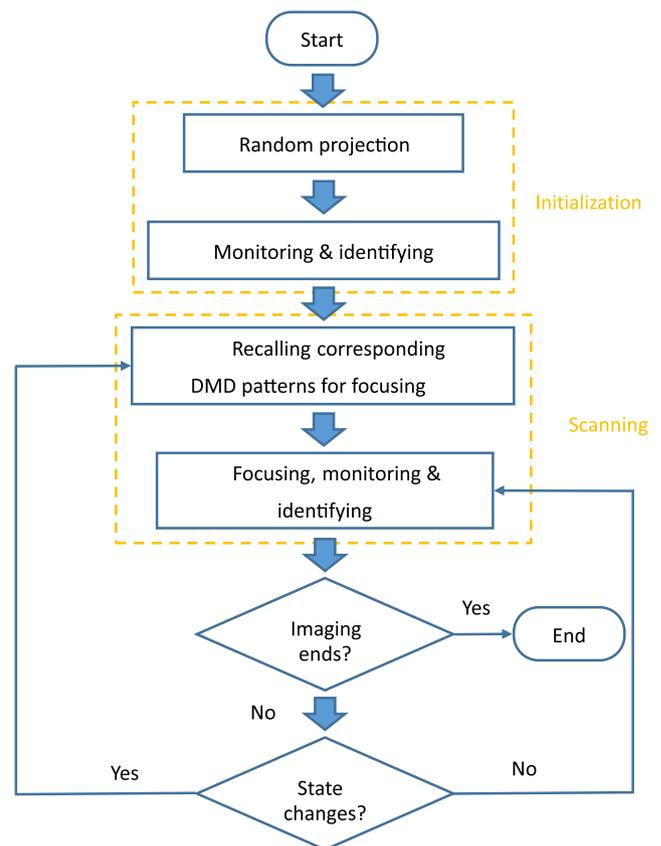


**Figure 7.** Focusing performance monitoring from the proximal end without distal access. A clear focus is predicted using DI-GAN (Generator network) by inputting the DMD pattern (left) for focusing used in the experiment and the reflected speckle image (middle) captured at the proximal end; the resulting focus provides a very good match to the corresponding real focus (rightmost) experimentally captured at the distal end. The FWHMs of the predicted and real focuses are shown. Scale bars, 10  $\mu\text{m}$ .

proximal end without distal access. For MMF-based endoscope application using scanning-based imaging technology, where the fiber tip would be inserted inside the deep tissue, the imaging quality primarily relies on the focusing performance that is impossible to be evaluated directly. We show for the first time, to the best of our knowledge, that the trained DI-GAN is able to monitor the focusing performance at the proximal end without distal access. Figure 4a has shown that the transmitted speckle images at the distal fiber end could be predicted using the corresponding DMD patterns and reflected speckle images; similarly, plotted in **Figure 7**, we use the DI-GAN (Generator network) to predict the focusing at the distal fiber end under the scenario where the distal tip is not accessible. A predicted focus is visible with high similarity to the experimentally captured focus, albeit with a 0.54  $\mu\text{m}$  larger FWHM.

The last feature of the proposed DI-GAN is that the Classifier network enables real-time bending state identification during monitoring or focusing. Using the reflected speckle image as the network input, the Classifier network can quickly output its corresponding fiber bending conformation (classification onto one of the 10 classes), with 100% accuracy in our case. Then, we are able to detect whether the current fiber configuration is changed, and recall the corresponding DMD input patterns for focusing, and recall the corresponding DMD input patterns for focusing. By doing so, we can achieve self-adaptively focusing at an arbitrary state before and after MMF bending. **Figure 8** shows the process of self-adaptively cross-state focusing through a semi-flexible MMF to solve the bending problem. Before the focusing starts, the framework initializes itself to identify the current fiber deformation. In the initialization, the DMD projects any patterns tentatively and the reflected speckle images are observed at the proximal end, so the fiber bending state can be identified in real-time with the help of DI-GAN (Classifier network). Next, the DMD input patterns over each corresponding bending state are recalled and projected for focusing. Simultaneously, the focusing performance is

monitored using DI-GAN (Generator network) and the fiber deformation is detected using DI-GAN (Classifier network) at the proximal end continuously. During the scanning, if the network detects a change of the fiber bending state, the different



**Figure 8.** The process of self-adaptively cross-state focusing.

DMD input patterns for focusing are recalled. Therefore, self-adaptively focusing at an arbitrary state before and after MMF bending can be achieved.

#### 4. Discussion and Conclusion

We have developed a unique framework, termed DI-GAN, powered by conditional GAN, that enables accurate input patterns reconstruction through a fiber with continuous deformations and permits input excitation wavefront prediction for self-adaptively focusing light at the distal tip of the fiber without distal access after the scalable calibration. This framework is a continual learning scheme under extreme memory usage limitation where the model is able to synthesize training data and prevent forgetting the previously learned bending states.

There are a few limiting factors in our proof-of-concept experiment. First, although it has been demonstrated that the deep learning method has surpassed the TM method in the ability to precisely control light propagation through MMF, it is also observed that the focus quality implemented using neural networks is not uniform across the fiber end. The DMD used in our experiment only supports binary modulation, which yields an inevitable difference between the ideal input pattern and the actual input pattern. This difference in input field at different positions results in a variation in the focus quality across the fiber end. For further improvement of the focus quality, more macro pixels on the DMD or a liquid crystal-based spatial light modulator (LC-SLM) could be used since both approaches can substantially improve the effective modulation. On the other hand, for the current setup and model structure in the text, the prediction of the DMD input pattern for focusing can be adjusted by changing the network input (focus intensity distribution), resulting in even, controllable, and measurable focus intensity delivery. Second, the number of MMF bending states in our proof-of-concept experiment is limited, and the cross-state calibration framework is only suitable for a semi-flexible MMF characterization. In the proposal, although we just demonstrate ten bending states within a small bending range (0 to 9 mm), the fiber can be bent in an arbitrary range because the different bending states are indicated using simple and discrete labels (digits from '0' to '9'). By appending the labels to the input images, the networks can generate the output images at the corresponding bending states. The separated labels are, however, not full featured and thus not suitable for referring to an arbitrary fiber deformation. To extend the framework to a totally flexible fiber in the future, a continuous and user-defined reference that can better indicate the fiber deformation needs to be explored for including them into the training of models. For example, the stage could be mounted on a three-axis motorized translation stage, allowing continuous, precise, and repeatable control of fiber deformation, and the data with a continuous reference could be collected accordingly. Third, we also note that the training of the continual learning model may be time consuming and challenging; however, this training process is a one-time shot, and the model can be rapidly deployed after the training is complete. Given the rapid development of high-performance computing technologies, neural networks may soon be able to be trained in a fraction of the time they currently require. Finally, the proposed network

architecture used in DI-GAN is fixed, which may fail to accommodate continually incoming tasks. A dynamic architecture expansion mechanism that ensures sufficient model capacity is considered feasible. In addition to the architecture, loss function redesign is also conducive to the data synthesis ability of the models.

In summary, we have presented a scalable calibration framework for a dynamically deformed MMF under EMC using deep learning. Our proof-of-concept experiments demonstrate that the proposed continual generative adversarial model enables us to characterize the MMF transmission states sequentially and detect the fiber deformation in real-time synchronously, allowing self-adaptively cross-state focusing through a semi-flexible MMF without distal access after the scalable calibration. The proposed method paves the way for the experimental realization of scalable calibration of a dynamically deformed MMF and will lead to future new flexible MMF-based endoscopes.

#### Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

#### Acknowledgements

This research was supported by Engineering and Physical Sciences Research Council (EP/L022559/1, EP/L022559/2, and EP/V050311/1), Royal Society (RG130230, IE161214), H2020 Marie Skłodowska-Curie Actions (790666), National Natural Science Foundation of China (61905115), and Leading Technology of Jiangsu Basic Research Plan (BK20192003). This research utilized Queen Mary's Apocrita HPC facility, supported by QMUL Research-IT (<http://doi.org/10.5281/zenodo.438045>).

#### Conflict of Interest

The authors declare no conflict of interest.

#### Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

#### Keywords

deep learning, light control, multimode optical fibers, neural networks, single multimode fiber imaging

Received: October 3, 2021

Revised: April 5, 2022

Published online:

- [1] D. J. Richardson, J. M. Fini, L. E. Nelson, *Nat. Photonics* **2013**, 7, 354.
- [2] Y. Choi, C. Yoon, M. Kim, T. D. Yang, C. Fang-Yen, R. R. Dasari, K. J. Lee, W. Choi, *Phys. Rev. Lett.* **2012**, 109, 203901.
- [3] T. Cizmar, K. Dholakia, *Nat. Commun.* **2012**, 3, 1027.
- [4] S. Bianchi, R. Di Leonardo, *Lab Chip* **2012**, 12, 635.

- [5] S. F. Liew, B. Redding, M. A. Choma, H. D. Tagare, H. Cao, *Opt. Lett.* **2016**, *41*, 2029.
- [6] D. Psaltis, C. Moser, *Opt. Photonics News* **2016**, *27*, 24.
- [7] I. T. Leite, S. Turtaev, X. Jiang, M. Šiler, A. Cuschieri, P. S. J. Russell, T. Čižmár, *Nat. Photonics* **2017**, *12*, 33.
- [8] S. Turtaev, I. T. Leite, T. Altwegg-Boussac, J. M. P. Pakan, N. L. Rochefort, T. Cizmar, *Light: Sci. Appl.* **2018**, *7*, 92.
- [9] S. A. Vasquez-Lopez, R. Turcotte, V. Koren, M. Ploschner, Z. Padamsey, M. J. Booth, T. Cizmar, N. J. Emptage, *Light Sci. Appl.* **2018**, *7*, 1.
- [10] R. Di Leonardo, S. Bianchi, *Opt. Express* **2011**, *19*, 247.
- [11] T. Čižmár, K. Dholakia, *Opt. Express* **2011**, *19*, 18871.
- [12] I. N. Papadopoulos, S. Farahi, C. Moser, D. Psaltis, *Opt. Express* **2012**, *20*, 10583.
- [13] D. Akbulut, T. J. Huisman, E. G. van Putten, W. L. Vos, A. P. Mosk, *Opt. Express* **2011**, *19*, 4017.
- [14] S. A. Goorden, J. Bertolotti, A. P. Mosk, *Opt. Express* **2014**, *22*, 17999.
- [15] M. Plöschner, T. Tyc, T. Čižmár, *Nat. Photonics* **2015**, *9*, 529.
- [16] M. Ploschner, B. Straka, K. Dholakia, T. Cizmar, *Opt. Express* **2014**, *22*, 2933.
- [17] D. Loterie, S. Farahi, I. Papadopoulos, A. Goy, D. Psaltis, C. Moser, *Opt. Express* **2015**, *23*, 23845.
- [18] M. Plöschner, V. Kollárová, Z. Dostál, J. Nylk, T. Barton-Owen, D. E. Ferrier, R. Chmellik, K. Dholakia, T. Čižmár, *Sci. Rep.* **2015**, *5*, 18050.
- [19] R. Olshansky, *Appl. Opt.* **1975**, *14*, 935.
- [20] L. Su, K. S. Chiang, C. Lu, *Appl. Opt.* **2005**, *44*, 7394.
- [21] L. G. Wright, Z. W. Liu, D. A. Nolan, M. J. Li, D. N. Christodoulides, F. W. Wise, *Nat. Photonics* **2016**, *10*, 771.
- [22] I. N. Papadopoulos, S. Farahi, C. Moser, D. Psaltis, *Biomed. Opt. Express* **2013**, *4*, 260.
- [23] A. M. Caravaca-Aguirre, R. Piestun, *Opt. Express* **2017**, *25*, 1656.
- [24] A. M. Caravaca-Aguirre, E. Niv, D. B. Conkey, R. Piestun, *Opt. Express* **2013**, *21*, 12881.
- [25] D. Loterie, D. Psaltis, C. Moser, *Opt. Express* **2017**, *25*, 6263.
- [26] D. E. Boonzajer Flaes, J. Stopka, S. Turtaev, J. F. de Boer, T. Tyc, T. Cizmar, *Phys. Rev. Lett.* **2018**, *120*, 233901.
- [27] L. V. Amitonova, J. F. de Boer, *Opt. Lett.* **2018**, *43*, 5427.
- [28] M. Lan, D. Guan, L. Gao, J. Li, S. Yu, G. Wu, *Opt. Express* **2019**, *27*, 12957.
- [29] A. M. Caravaca-Aguirre, S. Singh, S. Labouesse, M. V. Baratta, R. Piestun, E. Bossy, *APL Photonics*. **2019**, *4*, 096103.
- [30] C. K. Mididoddi, M. R. Hughes, in *Endoscopic Microscopy XIV*, **2019**, SPIE, San Francisco, California, United States.
- [31] M. Lan, Y. Xiang, J. Li, L. Gao, Y. Liu, Z. Wang, S. Yu, G. Wu, J. Ma, *Opt. Express* **2020**, *28*, 13662.
- [32] L. V. Amitonova, J. F. de Boer, *Light Sci. Appl.* **2020**, *9*, 1.
- [33] S. Farahi, D. Ziegler, I. N. Papadopoulos, D. Psaltis, C. Moser, *Opt. Express* **2013**, *21*, 22504.
- [34] R. Y. Gu, R. N. Mahalati, J. M. Kahn, *Opt. Express* **2015**, *23*, 26905.
- [35] S.-Y. Lee, B. E. Bouma, M. Villiger, in *2018 IEEE Photonics Conf. (IPC)* **2018**, IEEE.
- [36] G. S. D. Gordon, M. Gataric, A. G. C. P. Ramos, R. Mouthaan, C. Williams, J. Yoon, T. D. Wilkinson, S. E. Bohndiek, *Phys. Rev. X* **2019**, *9*, 041050.
- [37] S. Y. Lee, V. J. Parot, B. E. Bouma, M. Villiger, *APL Photonics* **2020**, *5*.
- [38] N. Borhani, E. Kakkava, C. Moser, D. Psaltis, *Optica* **2018**, *5*, 960.
- [39] P. Fan, T. Zhao, L. Su, *Opt. Express* **2019**, *27*, 20241.
- [40] B. Rahmani, D. Loterie, G. Konstantinou, D. Psaltis, C. Moser, *Light Sci. Appl.* **2018**, *7*, 69.
- [41] O. Moran, P. Caramazza, D. Faccio, R. Murray-Smith, in *Advances in Neural Information Processing Systems 2018* NeurIPS, Palais des Congrès de Montréal, Montréal CANADA.
- [42] P. Fan, L. Deng, L. Su, in *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conf. (IAEAC)* **2018**, IEEE, Chongqing.
- [43] P. Fan, M. Ruddlesden, Y. Wang, L. Zhao, C. Lu, L. Su, *Laser Photonics Rev.* **2021**, *15*, 2000348.
- [44] P. Caramazza, O. Moran, R. Murray-Smith, D. Faccio, *Nat. Commun.* **2019**, *10*, 2029.
- [45] A. Turpin, I. Vishniakou, J. D. Seelig, *Opt. Express* **2018**, *26*, 30911.
- [46] E. Fini, S. Lathuilière, E. Sangineto, M. Nabi, E. Ricci, in *European Conf. on Computer Vision*, **2020**, Springer, Cham.
- [47] H. Zhao, H. Wang, Y. Fu, F. Wu, X. Li, *arXiv preprint arXiv:2008.01411*. **2020**.
- [48] J. Bao, D. Chen, F. Wen, H. Li, G. Hua, in *Proc. of the IEEE Inter. Conf. on Computer Vision* **2017**, IEEE, Venice.
- [49] P. Fan, M. Ruddlesden, Y. Wang, L. Su, in *2020 Conf. on Lasers and Electro-Optics (CLEO) 2020*, OSA, San Jose, California, USA.
- [50] X. Liu, M. Masana, L. Herranz, J. Van de Weijer, A. M. Lopez, A. D. Bagdanov, in *2018 24th Inter. Conf. on Pattern Recognition (ICPR)* **2018**, IEEE, Beijing.
- [51] J. Kirkpatrick, R. Pascanu, N. Rabinowitz, J. Veness, G. Desjardins, A. A. Rusu, K. Milan, J. Quan, T. Ramalho, A. Grabska-Barwinska, D. Hassabis, C. Clopath, D. Kumaran, R. Hadsell, *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 3521.
- [52] M. Delange, R. Aljundi, M. Masana, S. Parisot, X. Jia, A. Leonardis, G. Slabaugh, T. Tuytelaars, *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, PP.
- [53] S.-A. Rebuffi, A. Kolesnikov, G. Sperl, C. H. Lampert, in *Proc. of the IEEE conference on Computer Vision and Pattern Recognition* **2017**, IEEE, Honolulu, HI, USA.
- [54] H. Shin, J. K. Lee, J. Kim, J. Kim, *arXiv preprint arXiv:1705.08690* **2017**.
- [55] D. Lopez-Paz, M. Ranzato, in *Advances in Neural Information Processing Systems 30 (Nips 2017)*, Vol. 30, **2017**, NeurIPS, Long Beach Convention Center, Long Beach.
- [56] Z. Li, D. Hoiem, *IEEE T Pattern Anal.* **2017**, *40*, 2935.
- [57] A. Mallya, S. Lazebnik, in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition* **2018**, IEEE, Salt Lake City, UT, USA.
- [58] A. Rosenfeld, J. K. Tsotsos, *IEEE Trans. Pattern Anal.* **2018**, *42*, 651.
- [59] C. S. Wu, L. Herranz, X. L. Liu, Y. X. Wang, J. van de Weijer, B. Raducanu, *Advances in Neural Information Processing Systems 31 (Nips 2018)*, Vol. 31, **2018**, NeurIPS, Palais des Congrès de Montréal, Montréal CANADA.
- [60] L. Deng, J. D. Yan, D. S. Elson, L. Su, *Opt. Express* **2018**, *26*, 18436.
- [61] T. Zhao, L. Deng, W. Wang, D. S. Elson, L. Su, *Opt. Express* **2018**, *26*, 20368.