ARTICLE

Open Access

Unsupervised full-color cellular image reconstruction through disordered optical fiber

Xiaowen Hu[1], Jian Zhao [2][✉], Jose Enrique Antonio-Lopez[1], Rodrigo Amezcua Correa[1] and Axel Schülzgen[1]

Abstract

Recent years have witnessed the tremendous development of fusing fiber-optic imaging with supervised deep learning to enable high-quality imaging of hard-to-reach areas. Nevertheless, the supervised deep learning method imposes strict constraints on fiber-optic imaging systems, where the input objects and the fiber outputs have to be collected in pairs. To unleash the full potential of fiber-optic imaging, unsupervised image reconstruction is in demand. Unfortunately, neither optical fiber bundles nor multimode fibers can achieve a point-to-point transmission of the object with a high sampling density, as is a prerequisite for unsupervised image reconstruction. The recently proposed disordered fibers offer a new solution based on the transverse Anderson localization. Here, we demonstrate unsupervised full-color imaging with a cellular resolution through a meter-long disordered fiber in both transmission and reflection modes. The unsupervised image reconstruction consists of two stages. In the first stage, we perform a pixel-wise standardization on the fiber outputs using the statistics of the objects. In the second stage, we recover the fine details of the reconstructions through a generative adversarial network. Unsupervised image reconstruction does not need paired images, enabling a much more flexible calibration under various conditions. Our new solution achieves full-color high-fidelity cell imaging within a working distance of at least 4 mm by only collecting the fiber outputs after an initial calibration. High imaging robustness is also demonstrated when the disordered fiber is bent with a central angle of 60°. Moreover, the cross-domain generality on unseen objects is shown to be enhanced with a diversified object set.

Introduction

Optical fibers are well-known for transmitting remote information out of otherwise inaccessible areas. Because of their miniature sizes and flexibility, fiber-optic imaging systems (FOISs)[1] have become an indispensable tool in clinical practice and biological research, such as early detection of gastrointestinal cancers[2–5] and visualization of neuronal activities in freely moving animals[6–10]. Most common FOISs are based on optical fiber bundles or multimode fiber (MMFs). An optical fiber bundle consists of thousands of closely spaced cores in a shared cladding. Each core acts as a single-pixel detector to sample the object[6,11–14]. Due to the loss of information in the cladding,

the output images from an optical fiber bundle suffer from the honeycomb effect[15]. On the other hand, an MMF supports thousands of optical modes in a single core. Because of the mode coupling in MMF, object images are scrambled into speckle patterns. Supervised deep learning[16,17] has been successfully implemented in both cases to reconstruct high-quality images[18–21]. A Convolutional Neural Network (CNN) can "learn" an image reconstruction mapping from numerous pairs of fiber output images and the input object images. Despite its success, supervised deep learning imposes a heavy burden on FOISs. The collection of paired fiber outputs and input objects in the calibration step involves time-consuming and demanding alignments of the FOIS. Especially, a re-calibration is required for any system variations, which is infeasible for practical applications.

Unsupervised deep learning circumvents these hurdles by using unpaired training image data. Since the deep learning model has to uncover the hidden mapping between two

Correspondence: Jian Zhao (jianzhao@knights.ucf.edu)
[1]CREOL, The College of Optics and Photonics, University of Central Florida, Orlando, FL 32816, USA
[2]The Picower Institute for Learning and Memory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

image domains without paired images, image reconstruction using unsupervised deep learning is considered to be a challenging task. Recently, it has been demonstrated that if the two image domains are similar in the high-dimensional space, "generator" CNNs and "discriminator" CNNs can compete in adversarial games to find a "natural" translation between the two image domains[22,23]. To achieve this similarity in the high-dimension domain, the FOISs should have a point-to-point transmission between the input object and the fiber output with high sampling densities. Unfortunately, neither optical fiber bundles nor MMFs meet these requirements. Although optical fiber bundles can directly convey the images of the objects, they have limited sampling densities (~0.1 mode/μm^2). As more sampling points, i.e., more cores, are added, the core-to-core crosstalk becomes stronger and degrades the point-to-point transmission fidelity[24]. On the other hand, the input-output relationship in MMFs is far deviated from a point-to-point transmission due to the multimode interference. The recently proposed glass-air Anderson localizing optical fibers (GALOFs)[25–33] provide a promising alternative. With a disordered arrangement of air holes embedded in a silica matrix, GALOFs achieve local confinement of light and high sampling densities (~10 mode/μm^2) simultaneously[34] due to the transverse Anderson localization (TAL)[35,36]. Moreover, the TAL-supported modes are insensitive to external perturbations[37] or wavelength shifts[38], as opposed to the modes in optical fiber bundles[24,39,40] or MMFs[41–45]. Therefore, robust full-color image transport can be achieved.

Here, we demonstrate unsupervised full-color high-fidelity image reconstruction through a meter-long GALOF in both transmission and reflection modes. We show that a simple histogram equalization step is adequate to reveal the hidden objects in the GALOF outputs preliminarily due to the densely-distributed TAL modes. The objects' fine details can be further recovered by utilizing the unpaired image-to-image translation[22,23]. Unsupervised image reconstruction significantly simplifies the calibration step, where the object images only need to be collected once. Therefore, the GALOF-based FOIS is flexible towards different conditions. As a remarkable example, we show the system's consistent imaging performance within a working distance of at least 4 mm, with a simple one-step re-calibration that only requires GALOF outputs. Moreover, due to the robustness of the TAL-supported modes, high image quality is preserved even under substantial mechanical bending (~60° bent angle). Finally, we show that the cross-domain generalizability of unseen objects can be enhanced by increasing the objects' diversity.

Results
Principles
When propagating through a GALOF (Fig. 1a), the imaging information of the object is encoded by the TAL-supported modes. The light confinement provided by the TAL results in a nearly point-to-point transmission from the GALOF's input facet to the output facet. Due to the different mode losses, the GALOF output pattern is an unevenly weighted superposition of the TAL-supported modes. Reconstructing the object from the output pattern involves standardizing all the TAL-supported modes and solving an inverse imaging problem. This is a challenging task since the TAL-supported modes have a very high mode density. Instead, we tackle this problem by standardizing the pixels of the GALOF's output images. In the calibration step (Fig. 1a), we collect 1000 fiber output images and another unpaired 1000 objects' reference images ("Methods"). Especially, there does not exist a one-to-one correspondence between these unpaired data sets. Before standardization, we register the 1000 fiber outputs according to some arbitrarily chosen fiber outputs (Methods) to compensate for the image drift caused by the mechanical instability during experiments. As a result, each pixel in the fiber output has 1000 different values. Since a large area of the object is scanned, each pixel should have captured the comprehensive statistical features of the object. Statistically speaking, all these pixels should have the same Probability Mass Function (PMF) as those in the reference objects, despite being from different unknown objects. Therefore, we perform histogram equalization ("Methods") to each pixel in the fiber output image for each RGB (red, green, and blue) channel (Fig. 1b). We calculate the Cumulative Distribution Function (CDF) from the PMF of each pixel and look for the pixel value in the reference objects that has the same CDF. In this way, we generate a Look-Up Table (LUT) for each pixel to transform its value. Among the range of 0–255, zeros are assigned to the defective pixels whose maximum value is less than 10 or whose Standard Deviation (STD) value is less than 2. For example, the value of the blue channel is set to zeros, as shown in Fig. 1b. Next, we perform image inpainting ("Methods") on each processed image. For each RGB channel, we interpolate inward from the pixels whose values are less than 10. Fuzzy objects are recovered after inpainting.

Finally, the reference objects are used again to further enhance the imaging quality of the 1000 fuzzy objects. We utilize our recently proposed image restoration cycle-consistent adversarial network (Restore-CycleGAN)[23] (Fig. 1c). As shown in our previous work, Restore-CycleGAN exhibits enhanced performance than the original CycleGAN[22] in extracting global information. In the Restore-CycleGAN, a U-Net[46] works as a generator G_1 to transform the fuzzy images into high-quality images, while a PatchGAN[47] works as a discriminator D_1 to distinguish the "real" reference objects from the "fake" ones produced by the generator G_1. The generator G_1 and the discriminator D_1 compete in an adversarial game through
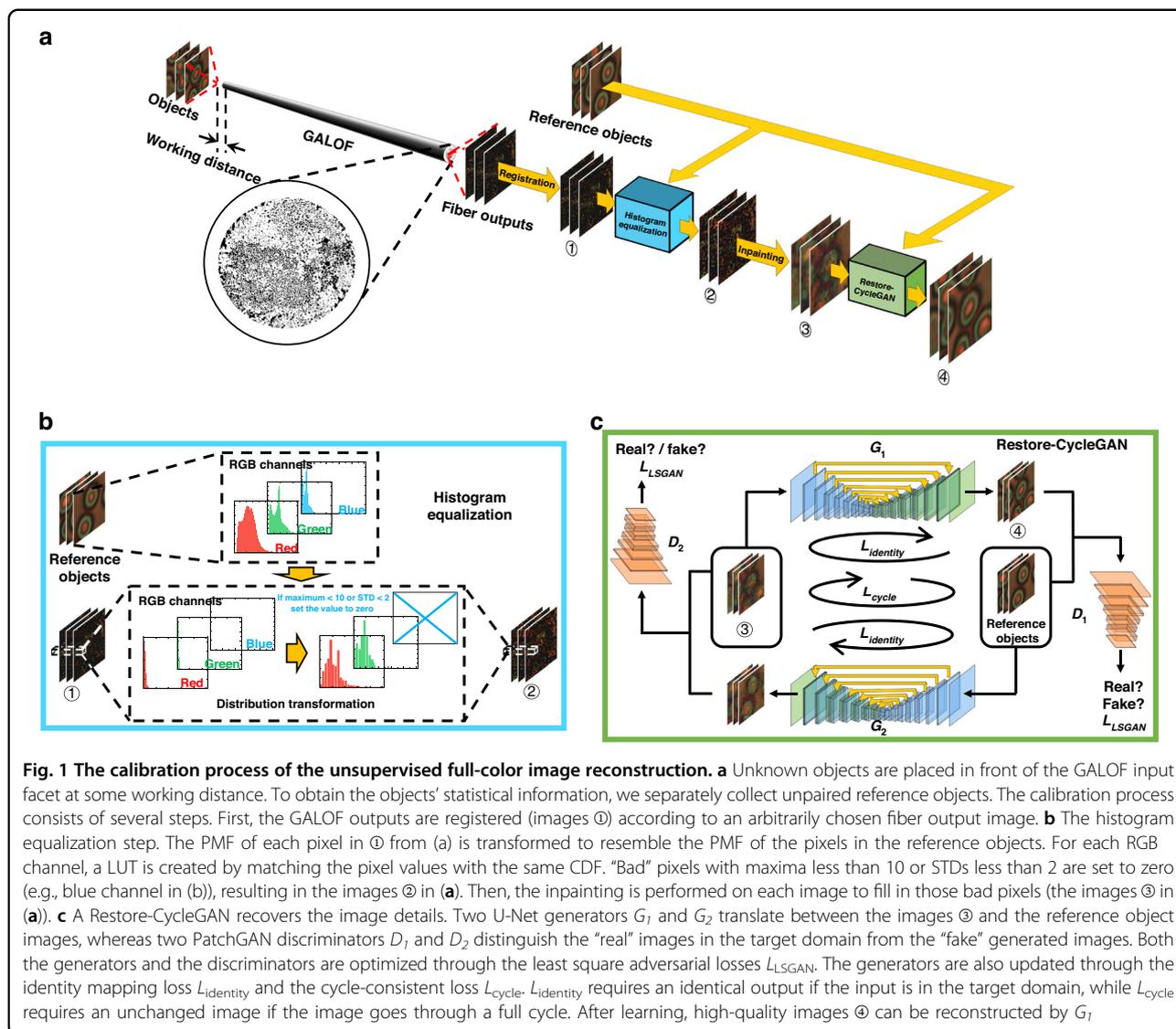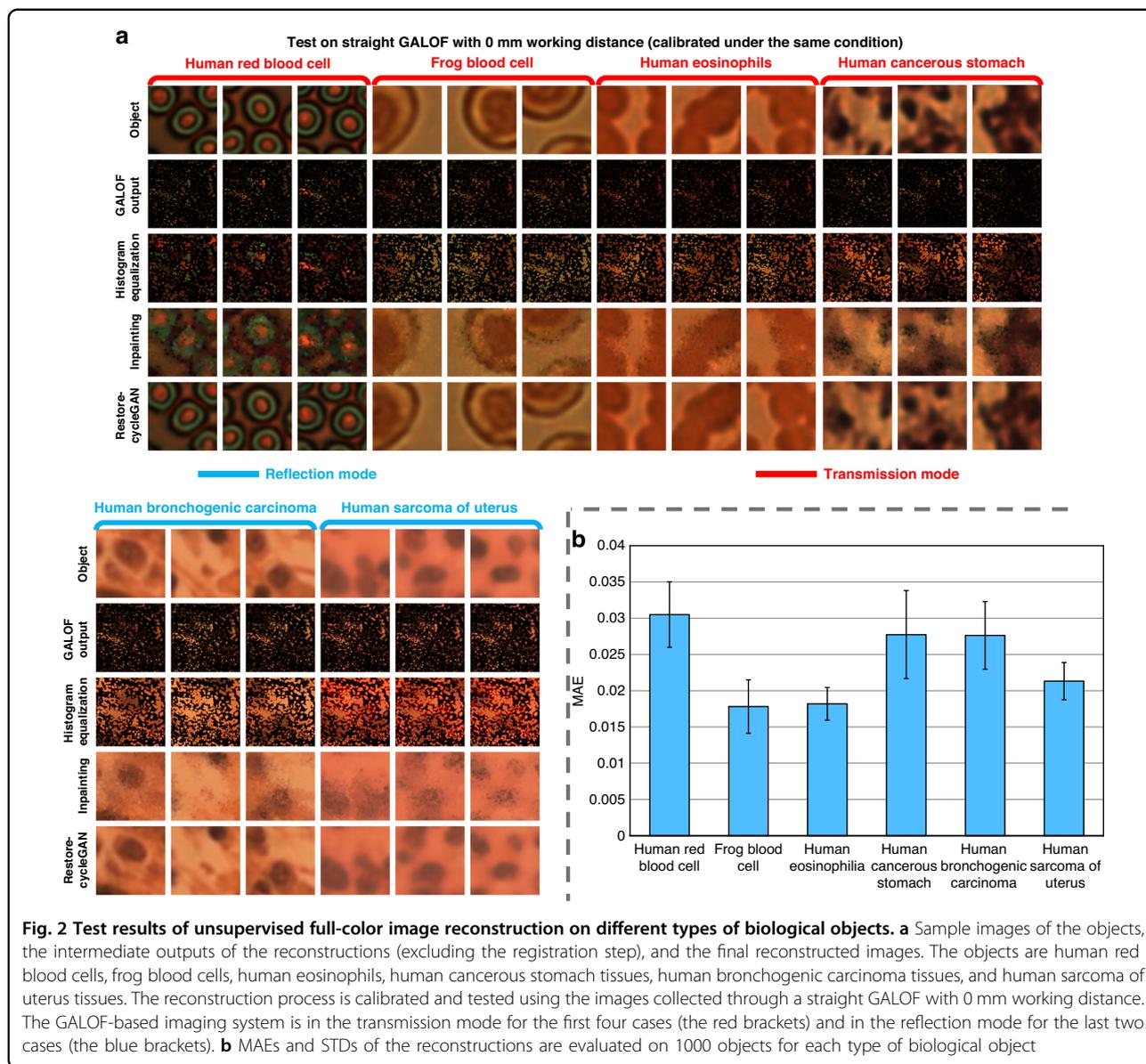
Fig. 1 The calibration process of the unsupervised full-color image reconstruction. a Unknown objects are placed in front of the GALOF input facet at some working distance. To obtain the objects' statistical information, we separately collect unpaired reference objects. The calibration process consists of several steps. First, the GALOF outputs are registered (images ①) according to an arbitrarily chosen fiber output image. **b** The histogram equalization step. The PMF of each pixel in ① from (a) is transformed to resemble the PMF of the pixels in the reference objects. For each RGB channel, a LUT is created by matching the pixel values with the same CDF. "Bad" pixels with maxima less than 10 or STDs less than 2 are set to zero (e.g., blue channel in (b)), resulting in the images ② in (a). Then, the inpainting is performed on each image to fill in those bad pixels (the images ③ in (a)). **c** A Restore-CycleGAN recovers the image details. Two U-Net generators G_1 and G_2 translate between the images ③ and the reference object images, whereas two PatchGAN discriminators D_1 and D_2 distinguish the "real" images in the target domain from the "fake" generated images. Both the generators and the discriminators are optimized through the least square adversarial losses L_{LSGAN}. The generators are also updated through the identity mapping loss $L_{identity}$ and the cycle-consistent loss L_{cycle}. $L_{identity}$ requires an identical output if the input is in the target domain, while L_{cycle} requires an unchanged image if the image goes through a full cycle. After learning, high-quality images ④ can be reconstructed by G_1

the least square adversarial loss L_{LSGAN}. In this adversarial game, G_1 gets rewarded if it successfully "fools" D_1, whereas D_1 gets rewarded if it differentiates the "real" from the "fake". G_1 is also optimized through the identity mapping loss $L_{identity}$, which requires a reference object to remain identical if it passes through G_1. Similarly, there is another pair of generator G_2 and discriminator D_2 in the opposite direction. To enhance cycle consistency, a cycle-consistent loss L_{cycle} is adopted to enforce an unaltered output if an image goes through the two generators successively. The details of the network architectures and the training processes can be found in the Methods. After training, only the generator G_1 is used. Therefore, unsupervised image reconstruction is achieved without paired training data. During the test, a fiber output image goes through the process of: (1) aligning with the arbitrarily chosen fiber output image; (2) pixel value transformation

using the LUTs; (3) inpainting; and (4) quality enhancement through the generator G_1. The set of reference objects is only needed to recalibrate the system for special cases, such as changing the working distances (Fig. 1a).

High fidelity

We perform the calibration processes for six different biological objects: human red blood cells, frog blood cells, human eosinophils, human cancerous stomach tissues, human bronchogenic carcinoma tissues, and human sarcoma of uterus tissues. All calibrations are performed using the same straight GALOF with a working distance of 0 mm. As shown in Fig. 2a, the data of the first four objects (human red blood cells, frog blood cells, human eosinophils, and human cancerous stomach tissues) are collected in the transmission mode, whereas that of the last two (human bronchogenic carcinoma tissues and human sarcoma of

Fig. 2 Test results of unsupervised full-color image reconstruction on different types of biological objects. a Sample images of the objects, the intermediate outputs of the reconstructions (excluding the registration step), and the final reconstructed images. The objects are human red blood cells, frog blood cells, human eosinophils, human cancerous stomach tissues, human bronchogenic carcinoma tissues, and human sarcoma of uterus tissues. The reconstruction process is calibrated and tested using the images collected through a straight GALOF with 0 mm working distance. The GALOF-based imaging system is in the transmission mode for the first four cases (the red brackets) and in the reflection mode for the last two cases (the blue brackets). **b** MAEs and STDs of the reconstructions are evaluated on 1000 objects for each type of biological object
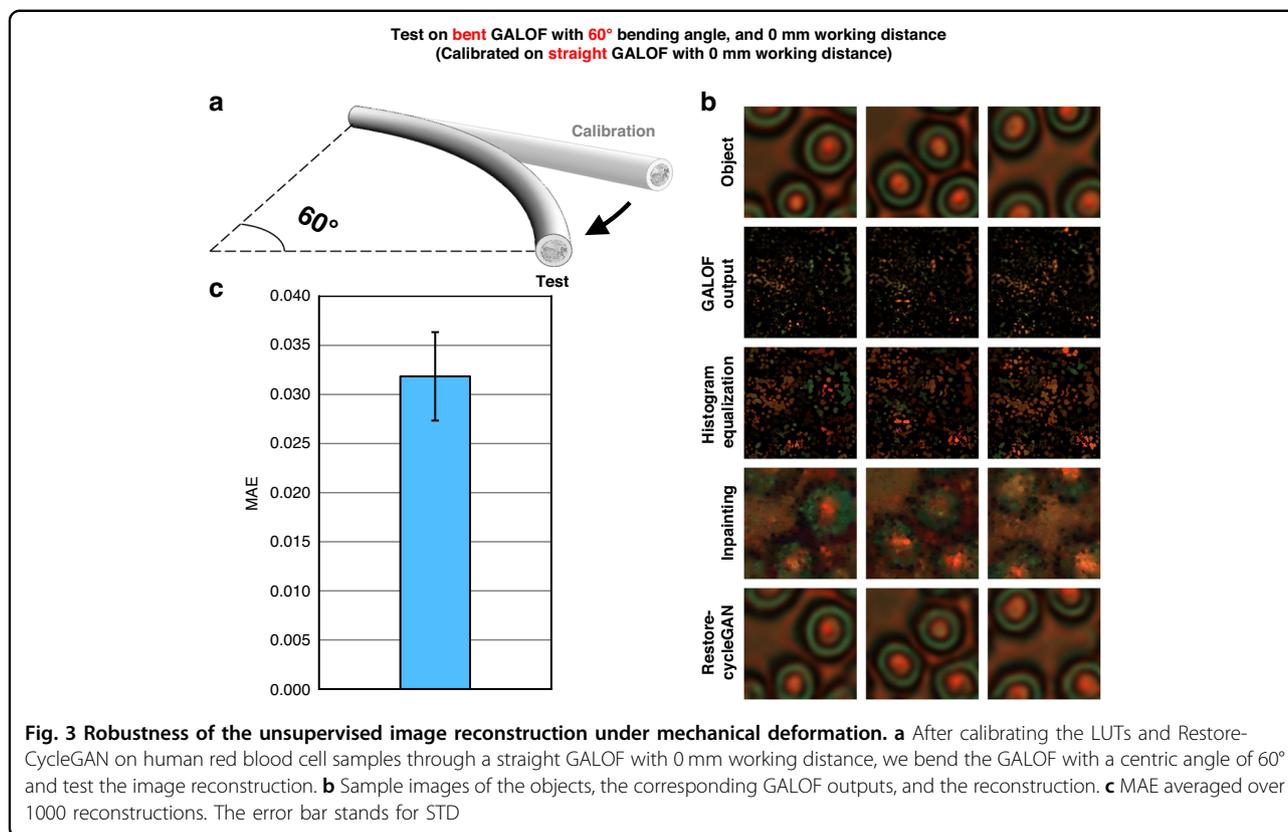
uterus tissues) are collected in the reflection mode. For each case, we separately collect 1000 object images and their GALOF outputs to evaluate the performance. The reconstruction time per image is about 1.6 s. Figure 2a shows some examples of the objects' reference images, the GALOF output images, and the results after each reconstruction step (without the registration step). Although the raw GALOF outputs are unrecognizable, they preserve the local information of the objects well at all RGB channels. This is made clear after the histogram equalization is applied to the pixels in the registered images. After the inpainting step, fuzzy images of the objects start to show up. Finally, Restore-CycleGANs further recover the fine details. The high fidelity of the reconstructions is quantitatively demonstrated in Fig. 2b, where we plot the mean absolute

errors (MAEs) and STDs of the 1000 reconstructions. In all six cases, the MAEs are below 0.035 (maximum ~1). In addition, we conduct a detailed analysis of the entire imaging pre-processing and reconstruction process, which reveals the cooperative impact of the pre-processing and the Restore-CycleGAN (Supplementary Information: Step-by-step analysis of unsupervised image reconstruction).

Robustness

To test the robustness of the unsupervised fiber-optic imaging, we bend the GALOF with a central angle of 60° (Fig. 3a). The output images are reconstructed using the LUTs and the Restore-CycleGAN calibrated on a straight GALOF. Both the calibration and test stages use the same type of human red blood cell samples at a working distance

Fig. 3 Robustness of the unsupervised image reconstruction under mechanical deformation. a After calibrating the LUTs and Restore-CycleGAN on human red blood cell samples through a straight GALOF with 0 mm working distance, we bend the GALOF with a centric angle of 60° and test the image reconstruction. **b** Sample images of the objects, the corresponding GALOF outputs, and the reconstruction. **c** MAE averaged over 1000 reconstructions. The error bar stands for STD

of 0 mm. As illustrated in Fig. 3b, high-fidelity reconstructions are achieved despite the large-angle fiber bending. We repeat the image reconstruction for 1000 GALOF outputs and calculate their MAE and STD with respect to the input objects (Fig. 3c). Similar values can be observed from the test results on a straight fiber. This indicates the consistency of the GALOF outputs, as attributed to the excellent light confinement of the TAL-supported modes.

Flexible working distance

Benefiting from the unsupervised image reconstruction, the reference objects only need to be collected once. When the working distance varies, we re-calibrate the LUTs and the Restore-CycleGAN using the same reference objects collected under a working distance of 0 mm. Figure 4 shows the test results on human red blood cells under the working distance of 1–6 mm with a step of 1 mm. Due to the loss of high-frequency information over the distance, the processed images after inpainting only demonstrate blurry profiles of the objects (Fig. 4a). Nevertheless, Restore-CycleGANs can still recover the images of objects with fine details. High-fidelity reconstructions are preserved up to a working distance of at least 4 mm. With increased working distances, the processed images after inpainting lost more information with significantly degraded imaging qualities, resulting in false blood cell reconstructions by the Restore-CycleGANs,

such as the reconstructions obtained at 6 mm working distance in Fig. 4a. To quantify the imaging performance, we calculate the MAEs and STDs of 1000 pairs of reconstructions and ground truths at each working distance (Fig. 4b). It shows that the increased working distance does not rapidly degrade the image quality. Therefore, our unsupervised image reconstruction approach enables flexible working distances through a simple re-calibration procedure. Imaging under various working distances serves as a showcase of the flexible re-calibration brought by unsupervised image reconstruction. The flexibility is further demonstrated in Supplementary Information, where we conduct numerical investigations to analyze the imaging performance under other extreme conditions, including low-light, high-noise, and uneven illuminations (Supplementary Information: "Imaging under low-light conditions", "Imaging under high-noise levels", and "Imaging under uneven illuminations"). The simulations demonstrate that a simple re-calibration can enable high-fidelity imaging even under low-light illuminations of 5% visibility, illuminations with additional Gaussian noise (0 mean and 50 variance), or Gaussian-distributed uneven illuminations.

Cross-domain generalizability

In the abovementioned results, the calibration and test are conducted on the same type of biological objects. Yet,
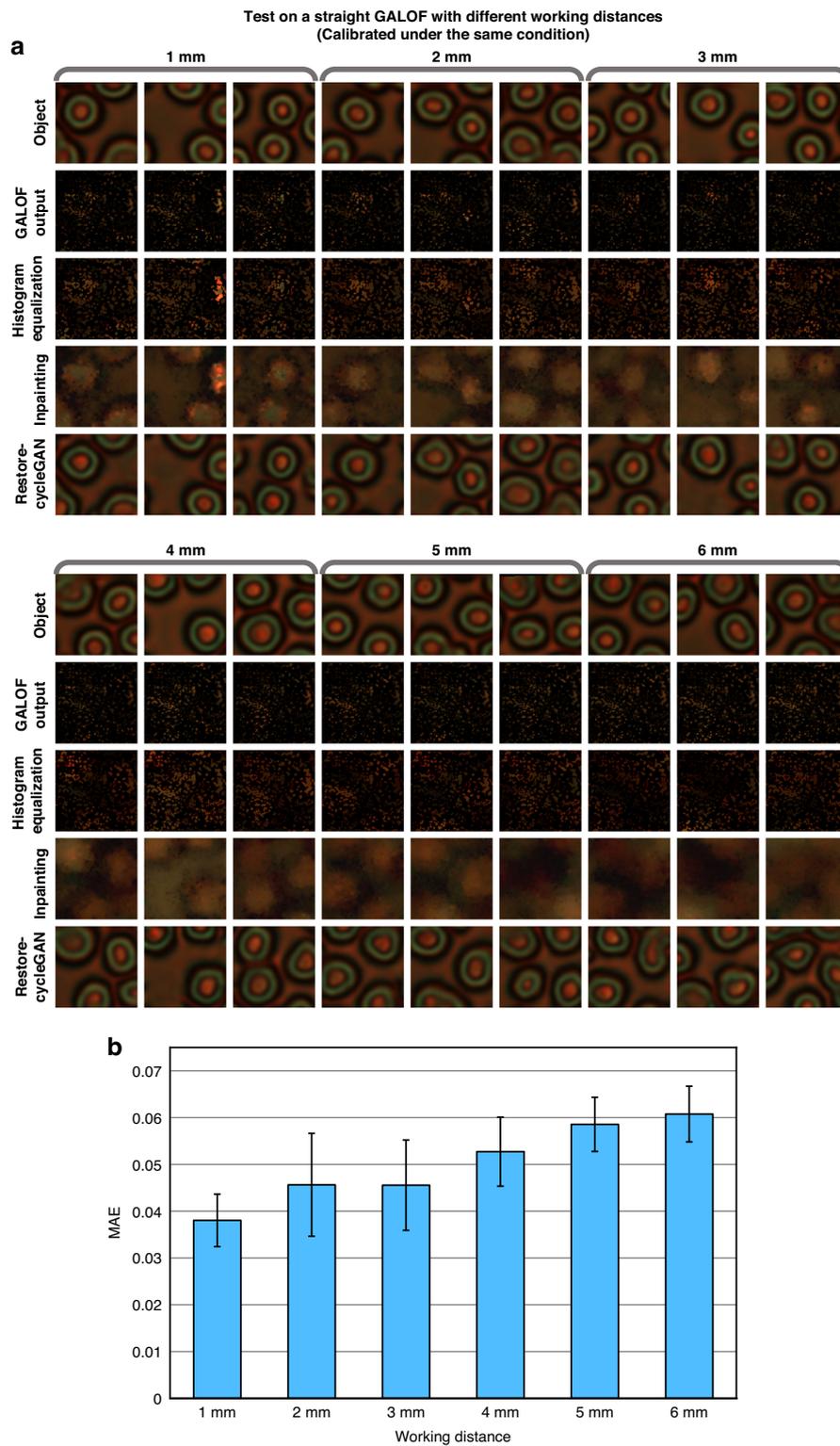
Fig. 4 Test results of unsupervised full-color image reconstruction with different working distances. The reconstruction process is calibrated and tested in a straight GALOF at each working distance. **a** Sample images of the objects (human red blood cells), the outputs of each reconstruction step (excluding the registration step), and the final reconstructed images. **b** MAEs and STDs of the reconstructions evaluated on 1000 objects for each type of biological objects

Fig. 5 Cross-domain generalizability on unseen objects. a In the calibration, both the GALOF outputs and the reference objects contain mixed images from four different types: human red blood cells, frog blood cells, human eosinophils, and human cancerous stomach tissues. We collect 300 unpaired and uncorrelated images for each type of sample from the GALOF outputs and the reference objects. Following the same procedure (Fig. 1), we obtain the LUTs for histogram equalization and the Restore-CycleGAN. Then we test the image reconstruction on GALOF outputs from unseen objects, i.e., bird blood cells. **b** Sample images of the bird blood cell images, the corresponding GALOF outputs, and the outputs of reconstruction. **c** MAE and STD averaged over 300 reconstructions

FOISs are expected to perform high-fidelity imaging on unseen objects in real-world applications. To enhance cross-domain generalizability, it is necessary to enrich the statistical information of the objects for image reconstruction. For this purpose, we include image data generated from various types of biological samples (Fig. 5a), such as human red blood cells, frog blood cells, human eosinophils, and cancerous stomach tissues. For each object type, we collect 300 reference object images and 300 GALOF output images from a straight GALOF with 0 mm working distance in the transmission mode. These two sets of images are unpaired and uncorrelated. We follow the same calibration procedure demonstrated in Fig. 1. After obtaining the LUTs and the Restore-Cycle-GAN, we test the unsupervised image reconstruction on GALOF outputs from unseen objects, i.e., bird blood cells. Figure 5b shows sample object images, the corresponding GALOF outputs, and the processed images after each reconstruction step (excluding the registration step). The profiles and orientations of the bird blood cells can be clearly identified, despite slightly degraded image quality. This corresponds to a higher MAE over 300 reconstructed images (Fig. 5c). The increase in MAE originates from the limited data size and object variations, which could be addressed by improving the training data further.

Discussion

Robust full-color high-fidelity image transport using unsupervised learning is achieved through the combined effects of the GALOF's properties, the pre-processing steps (registration, standardization, and inpainting), and the Restore-CycleGAN. First, the high-density localized modes result in a point-to-point transmission of the object with a high sampling ratio, which makes the inverse imaging problem well-suited for unsupervised learning. In addition, the TAL-supported modes have flat responses to different wavelengths[38], enabling full-color imaging. In contrast, an optical fiber bundle designed to transmit images at one wavelength may not be suitable for another wavelength[13]. Moreover, the robustness of the TAL-supported modes makes the imaging performance stable under large fiber deformations, whereas a translation of a few millimeters in one end of a meter-long optical fiber bundle or MMF tends to significantly degrade the output image[24,41–43]. Second, the pre-processing steps help bring the two image domains of the GALOF outputs and the objects closer, enabling the Restore-CycleGAN to find a 'natural' translation[23], since unsupervised image-to-image translation often fails when an extreme transformation occurs between the two image domains[22] (Supplementary Information: Step-by-step analysis of unsupervised image reconstruction). Third, the Restore-CycleGAN determines the high-fidelity reconstructions using the pre-processed images. Its crucial role is even more significant under high-noise conditions (Supplementary Information: Imaging under high-noise levels).

Free from strictly paired training imaging data, unsupervised image reconstruction streamlines the system design and calibration process, facilitates simpler and

faster data acquisition, and enhances GALOF-based FOIS as a flexible and efficient imaging platform for practical applications in various circumstances. Without the constraint of paired data, the reference object images in our system can be used repeatedly, with only the GALOF outputs needed for re-calibrations when the system changes. For example, a wide range of working distances is desirable in endoscopy applications to reduce penetration damage. System re-calibration by supervised learning is impractical since it requires collecting paired images for a changed working distance. As shown in this work, unsupervised learning enables simple re-calibrations of the system to acquire high-quality images up to a working distance of at least 4 mm. Moreover, the amount of data needed in calibrations is dramatically reduced. In our experiments, we only acquire 1000 GALOF outputs and 1000 reference object images for one calibration. In contrast, supervised learning typically requires tens of thousands of paired images to train a CNN.

Further improvement can be made in the GALOF fabrications and the unsupervised image reconstruction process. Currently, there are many defective pixels in the GALOF outputs, which lead to the loss of information. Future work can be devoted to investigating methods of eliminating these defective pixels. Moreover, the geometrical parameters of the GALOF can also be improved. The GALOF used in this work has an air-filling fraction of ~28 %. In contrast, an air-filling fraction of ~50% has been shown to be favorable for reducing the localization lengths and improving spatial resolution[48,49]. On the other hand, there is still much room for improvements in cross-domain generalizability. We expect a larger and more diversified image dataset would enhance the image reconstruction of unseen objects in future work. Furthermore, while it only takes the Restore-CycleGAN about 70 ms to reconstruct an image, the steps preceding the Restore-CycleGAN add a significant amount of time to the whole reconstruction process. The reconstruction time per image is about 1.6 s. Future studies can focus on speeding up the processing speed for steps preceding the Restore-CycleGAN to realize real-time imaging.

Future studies could also investigate the behaviors of unsupervised learning-based fiber imaging under extreme conditions, both experimentally and numerically, to develop system enhancement solutions. As detailed in the Supplementary Information, our unsupervised-learning-based fiber imaging method maintains high-quality imaging capabilities under low-light, high-noise, or uneven illuminations, demonstrating significant resistance to these challenging conditions. However, we also show that our imaging method fails beyond certain critical thresholds, such as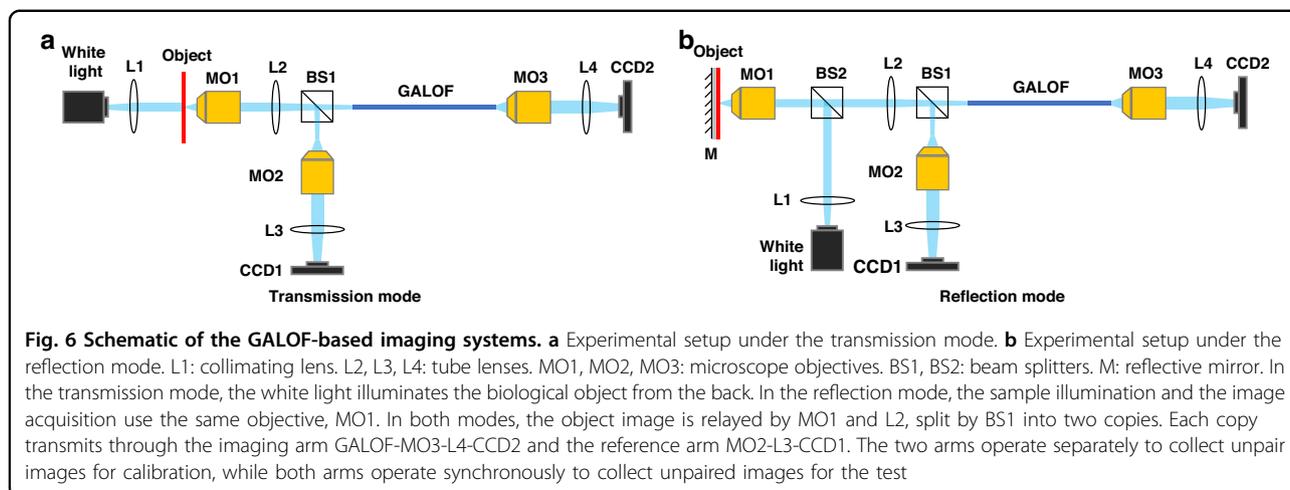 6 mm long working distance, extremely low-light illumination of 2% visibility, or high-level Gaussian noise of 100 variances, due to the significant alternations in the statistical features of GALOF's outputs. This raises the question of how to quantitatively monitor the entire imaging process and evaluate the reconstruction fidelity. In Supplementary Information: "Confidence metric for image reconstruction", we propose the correlation coefficient between the pre- and post-processed images by the Restore-CycleGAN as a confidence metric for alerting to model failures. The correlation coefficient demonstrates great conformity to image fidelity. Despite the progress, it remains an open question whether the limitations of unsupervised learning-based fiber imaging have been fully understood and whether a more suitable confidence metric could be employed. Consequently, further systematic experimental and numerical investigations are necessary to uncover better answers.

In conclusion, we achieve unsupervised image reconstruction in a meter-long GALOF based on its unique property of point-to-point transmission with high sampling densities. Full-color high-fidelity image transport is demonstrated on different types of biological samples in both transmission and reflection modes. The image quality is preserved when the GALOF is substantially bent with an angle of 60°. Enabled by unsupervised image reconstruction, the GALOF-based FOIS is flexible to different circumstances. High image quality is maintained within a working distance of at least 4 mm using a much-simplified re-calibration. Increased cross-domain generalizability on unseen objects is also shown by including diversified objects. Based on these results, we see the GALOF-based FOISs as promising candidates for the next-generation FOISs.

Materials and methods
Experimental setup

In both the transmission mode and the reflection mode (Fig. 6a, b), we use a quartz halogen lamp as the light source (wavelength: ~400 nm to ~2000 nm). A lens, L1, is placed in front of the lamp to collimate the light. In the transmission mode, the collimated light illuminates the object from behind. The object image is relayed by a 10× microscope objective (MO1) (infinity-corrected, NA = 0.3) and a tube lens L2 ($f = 200$ mm). The magnified image is then sent to two arms by a beam splitter BS1. In the reference arm, the image is further magnified by a 20x microscope objective (MO2) (NA = 0.75) and a tube lens L3 ($f = 200$ mm), and then collected by the CCD1 camera (Manta G-145C). In the imaging arm, the object image is delivered through the GALOF. The GALOF is fabricated by the stack-and-draw method. It has a disordered structure with a diameter of ~278 μm and an air-hole-filling fraction of ~28.5%. A segment of ~80 cm is used in the experiment. The GALOF output is

Fig. 6 Schematic of the GALOF-based imaging systems. a Experimental setup under the transmission mode. **b** Experimental setup under the reflection mode. L1: collimating lens. L2, L3, L4: tube lenses. MO1, MO2, MO3: microscope objectives. BS1, BS2: beam splitters. M: reflective mirror. In the transmission mode, the white light illuminates the biological object from the back. In the reflection mode, the sample illumination and the image acquisition use the same objective, MO1. In both modes, the object image is relayed by MO1 and L2, split by BS1 into two copies. Each copy transmits through the imaging arm GALOF-MO3-L4-CCD2 and the reference arm MO2-L3-CCD1. The two arms operate separately to collect unpair images for calibration, while both arms operate synchronously to collect unpaired images for the test

magnified by a combination of a 20x microscope objective (MO3) ($NA = 0.75$) and a tube lens L4 ($f = 200$ mm) before being collected by the CCD2 camera (Manta G-145C). In the reflection mode, the illumination light is coupled into the back aperture of the MO1 by a beam splitter (BS2), and focused onto the object. We place a mirror M as a highly reflective substrate behind the object without contact. Similar to the transmission mode, the reflected object image is magnified and sent to the two imaging arms. For both the transmission and reflection modes, the reference arm and the imaging arm collect images separately during calibration. They only operate synchronously during the test to evaluate the system's imaging performance.

GALOF outputs registration and inpainting

We first convert all the GALOF outputs to grayscale images. To find the transformation for registration, we use MATLAB "*imregtform*" function with monomodal registration and translation geometric transformation. Inpainting is based on MATLAB's "*regionfill*" function that can interpolates inward from the values of the pixels on the outer boundary of the regions.
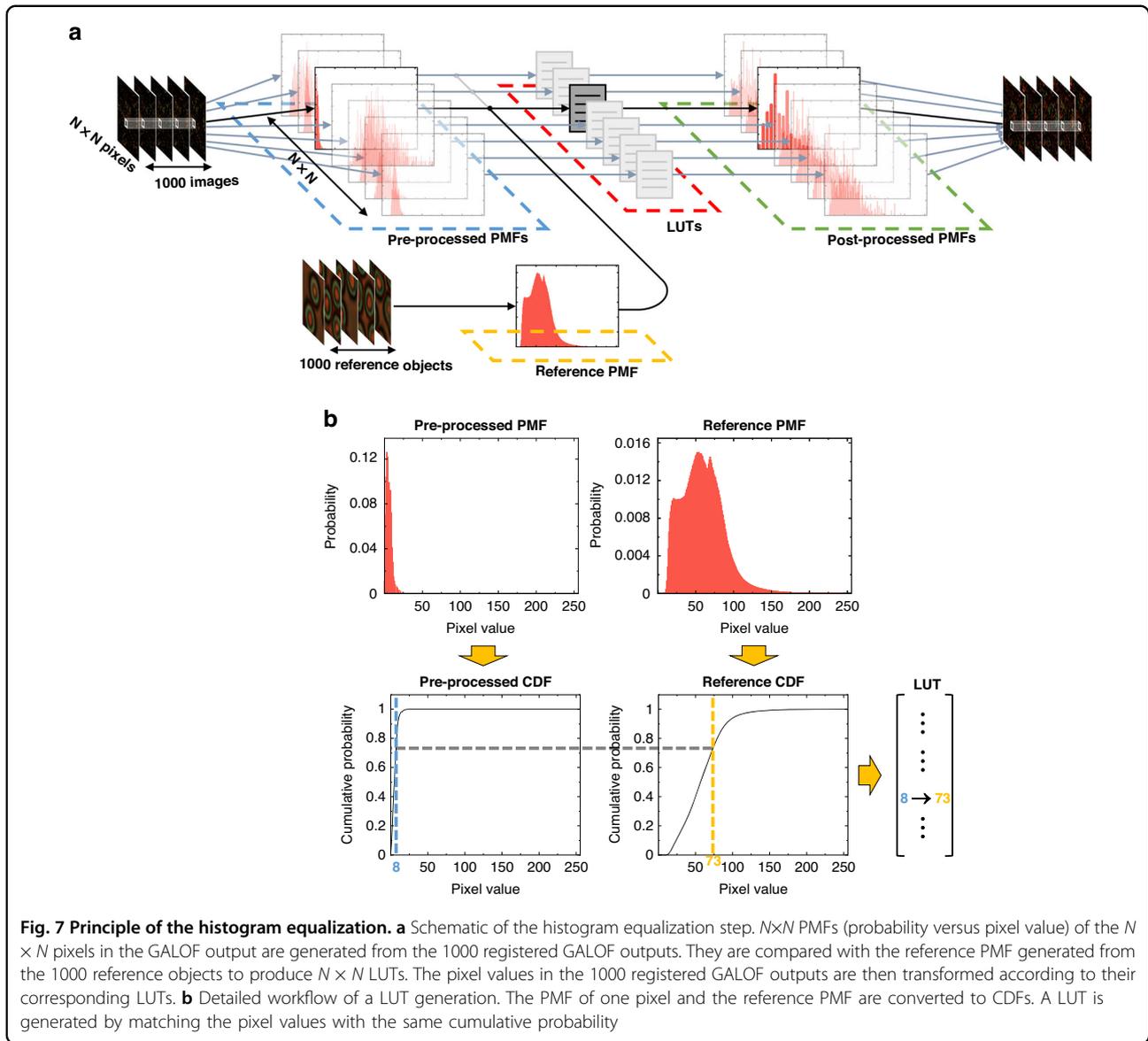
Histogram equalization

The detailed workflow of the histogram equalization step is illustrated in Fig. 7. For each RGB channel, every pixel in the 1000 registered GALOF outputs has 1000 values varying between 0 and 255 (Fig. 7a). We calculate the PMF of each pixel by counting the occurrence of a particular pixel value among the 1000 values. With the dimensions of one GALOF output being $N \times N$ ($N = 420$), there are $N \times N$ PMFs (probability versus pixel value). For the 1000 reference objects, we treat all the pixels equally and calculate one reference PMF from the $N \times N \times 1000$ pixel values. We convert each PMF to a CDF by summing up the probabilities of pixels whose

values are smaller or equal to a specific value (Fig. 7b). The $N \times N$ CDFs of the GALOF output pixels are then compared with the reference CDF individually. For each GALOF pixel, the pixel values are mapped to new values, such as 8 to 73, by matching the cumulative probabilities in the GALOF's CDF with those in the reference CDF. In this way, we obtain $N \times N$ LUTs. Consequently, the post-processed PMFs produced by the LUTs resemble the reference PMF distribution. Finally, the outputs from the histogram equalization step are generated by transforming the pixel values using the LUTs. It is noteworthy that, for any pixel with a maximum value of less than 10 or an STD value of less than 2 among the 1000 values (ranging from 0 to 255), we simply map all their values to 0. These pixels will be filled in during the inpainting step.

Restore-CycleGAN

The architectures of the generator and discriminator networks in the Restore-CycleGAN are shown in Fig. 8a, b. The generator is a U-Net with skip connections. The input image has a size of 420 × 420. It passes through different convolutional layers to a bottleneck and then passes through different transposed convolutional layers to the final output. The discriminator is a PatchGAN, which looks into patches of an input image and decides whether they are from the real images in the target domain or from the fake images generated by the generator. The detailed operations in the convolutional layers and transpose convolutional layers are shown in Fig. 8c. The numbers of filters in the layers of the generator are 64-128-256-512-512-512-512-512-512-512-512-512-256-128-64-3. The numbers of filters in the layers of the discriminator are 64-128-256-512-512-1.

The weights in the generators and the discriminators are initialized by random Gaussian distributions with a zero mean and a standard deviation of 0.02. For

Fig. 7 Principle of the histogram equalization. a Schematic of the histogram equalization step. *N×N* PMFs (probability versus pixel value) of the *N* × *N* pixels in the GALOF output are generated from the 1000 registered GALOF outputs. They are compared with the reference PMF generated from the 1000 reference objects to produce *N* × *N* LUTs. The pixel values in the 1000 registered GALOF outputs are then transformed according to their corresponding LUTs. **b** Detailed workflow of a LUT generation. The PMF of one pixel and the reference PMF are converted to CDFs. A LUT is generated by matching the pixel values with the same cumulative probability

translating between the domain x and the domain y, there are two generator-discriminator pairs: $G_{x \to y}$ and D_y in the direction from x to y, and $G_{y \to x}$ and D_x in the other direction. The loss function of a generator $G_{x \to y}$ can be written as:

$$
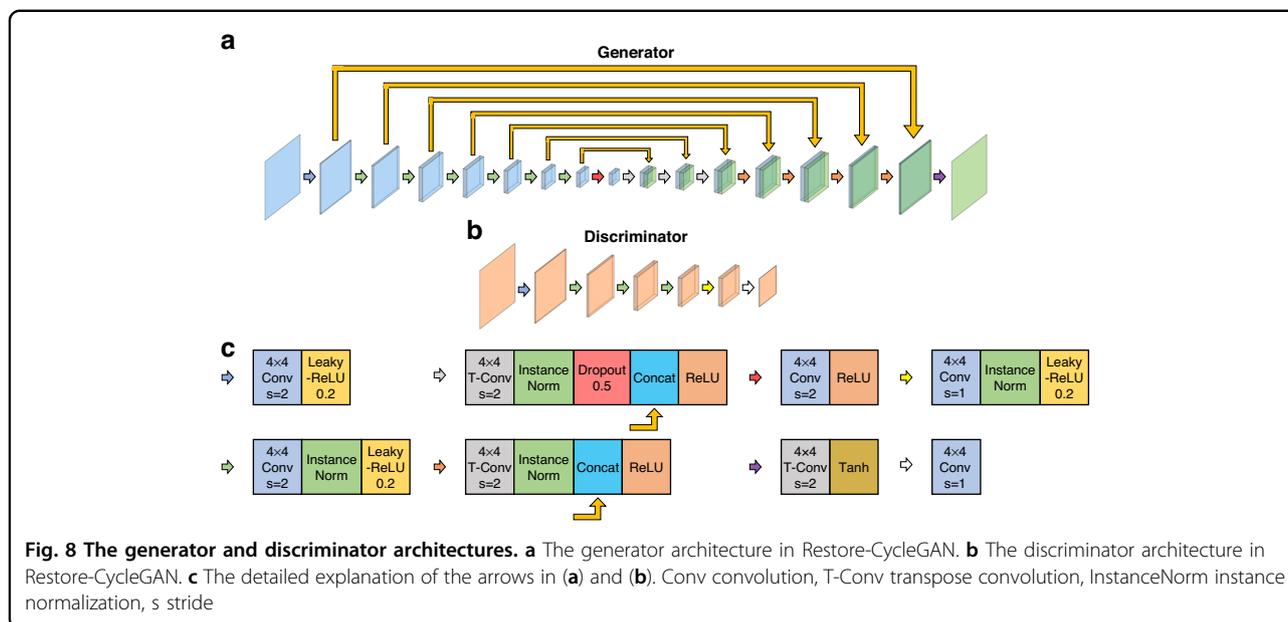\begin{aligned}
L_{G_{x \to y}} = & E_x[(D_y(G_{x \to y}(x)) - 1)^2] \\
& + \alpha_1 E_y[\|G_{x \to y}(G_{y \to x}(y)) - y\|_1] \\
& + \alpha_1 E_x[\|G_{y \to x}(G_{x \to y}(x)) - x\|_1] \\
& + \alpha_2 E_y[\|G_{x \to y}(y) - y\|_1]
\end{aligned}
\tag{1}
$$

The four terms in Eq. (1) are the least square adversarial loss $\mathcal{L}_{\text{LSGAN}}$, the cycle-consistent losses $\mathcal{L}_{\text{cycle}}$ in both directions, and the identity mapping loss $\mathcal{L}_{\text{identiy}}$,

respectively. α_1 and α_2 are the weights controlling the balance among the losses. We use $\alpha_1 = 10$, and $\alpha_2 = 5$. The weights in D_y and $G_{y \to x}$ are fixed when we train $G_{x \to y}$. The loss function of D_y is the least square adversarial loss $\mathcal{L}_{\text{LSGAN}}$:

$$
L_{D_y} = E_y[(D_y(y) - 1)^2] + E_x[D_y(G_{x \to y}(x))^2]
\tag{2}
$$

The weights in $G_{x \to y}$ are fixed when we train D_y. The real images y to train D_y are randomly selected from all the images in the target domain, whereas the fake images $G_{x \to y}(x)$ are randomly selected from a pool of 50 fake images. The pool is randomly updated through newly generated fake images. The loss of the discriminators is divided by half during training. The loss functions of $G_{y \to x}$

Fig. 8 The generator and discriminator architectures. a The generator architecture in Restore-CycleGAN. **b** The discriminator architecture in Restore-CycleGAN. **c** The detailed explanation of the arrows in (**a**) and (**b**). Conv convolution, T-Conv transpose convolution, InstanceNorm instance normalization, s stride

and D_x can be written in a similar way. We train the discriminators and generators for 100 epochs with a batch size of 1. We use an Adam optimizer with a learning rate of 0.0002 and the exponential decay rate for the first momentum $\beta_1 = 0.5$. The training takes ~40 h on a dual-GPU (GeForce GTX 1080 Ti) desktop.

Author contributions
X.H. developed the GALOF-based unsupervised full-color imaging framework, including the algorithms and the experimental systems, performed the data acquisitions, processed the experimental data, and wrote the first draft. X.H. and J.Z. proposed this project. J.Z. developed the prototype of the transmission imaging system, supervised the project, and revised the manuscript. J.Z. and J.E.A. developed and fabricated the GALOF used in the experiment. R.A.C. and A.S. supervised the GALOF development and fabrication. A.S. led the team, supervised the project, and revised the manuscript. All authors contributed to the final version of the manuscript.

Data availability
The datasets generated during the current study are available from the authors under reasonable request.

Code availability
The codes developed for this work are available from the authors under reasonable request.

Conflict of interest
The authors declare no competing interests.

References
1. Flusberg, B. A. et al. Fiber-optic fluorescence imaging. *Nat. Methods* **2**, 941–950 (2005).
2. Wallace, M. B. & Fockens, P. Probe-based confocal laser endomicroscopy. *Gastroenterology* **136**, 1509–1513 (2009).
3. Wang, K. K. et al. Use of probe-based confocal laser endomicroscopy (pCLE) in gastrointestinal applications. A consensus report based on clinical evidence. *U. Eur. Gastroenterol. J.* **3**, 230–254 (2015).
4. Fugazza, A. et al. Confocal laser endomicroscopy in gastrointestinal and pancreatobiliary diseases: a systematic review and meta-analysis. *BioMed. Res. Int.* **2016**, 4638683 (2016).
5. McGoran, J. J. et al. Miniature gastrointestinal endoscopy: now and the future. *World J. Gastroenterol.* **25**, 4051–4060 (2019).
6. Flusberg, B. A. et al. High-speed, miniaturized fluorescence microscopy in freely moving mice. *Nat. Methods* **5**, 935–938 (2008).
7. Szabo, V. et al. Spatially selective holographic photoactivation and functional fluorescence imaging in freely behaving mice with a fiberscope. *Neuron* **84**, 1157–1169 (2014).
8. Zong, W. J. et al. Fast high-resolution miniature two-photon microscopy for brain imaging in freely behaving mice. *Nat. Methods* **14**, 713–719 (2017).
9. Turtaev, S. et al. High-fidelity multimode fibre-based endoscopy for deep brain in vivo imaging. *Light Sci. Appl.* **7**, 92 (2018).
10. Vasquez-Lopez, S. A. et al. Subcellular spatial resolution achieved for deep-brain imaging in vivo using a minimally invasive multimode fiber. *Light Sci. Appl.* **7**, 110 (2018).
11. Ghaemi, H. F. et al. Fiber image guide with subwavelength resolution. *Appl. Phys. Lett.* **72**, 1137–1139 (1998).
12. Muldoon, T. J. et al. Subcellular-resolution molecular imaging within living tissue by fiber microendoscopy. *Opt. Express* **15**, 16413–16423 (2007).
13. Hughes, M., Chang, T. P. & Yang, G. Z. Fiber bundle endocytoscopy. *Biomed. Opt. Express* **4**, 2781–2794 (2013).
14. Krstajić, N. et al. Two-color widefield fluorescence microendoscopy enables multiplexed molecular imaging in the alveolar space of human lung tissue. *J. Biomed. Opt.* **21**, 046009 (2016).
15. Perperidis, A. et al. Image computing for fibre-bundle endomicroscopy: a review. *Med. Image Anal.* **62**, 101620 (2020).
16. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
17. Barbastathis, G., Ozcan, A. & Situ, G. H. On the use of deep learning for computational imaging. *Optica* **6**, 921–943 (2019).
18. Shao, J. B. et al. Fiber bundle image restoration using deep learning. *Opt. Lett.* **44**, 1080–1083 (2019).
19. Shao, J. B. et al. Fiber bundle imaging resolution enhancement using deep learning. *Opt. Express* **27**, 15880–15890 (2019).

20. Borhani, N. et al. Learning to see through multimode fibers. *Optica* **5**, 960–966 (2018).

21. Rahmani, B. et al. Multimode optical fiber transmission with a deep learning network. *Light Sci. Appl.* **7**, 69 (2018).

22. Zhu, J. Y. et al. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proc 2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, pp. 2242–2251 (IEEE, 2017). https://doi.org/10.1109/ICCV.2017.244.

23. Hu, X. W. et al. Adaptive inverse mapping: a model-free semi-supervised learning approach towards robust imaging through dynamic scattering media. *Opt. Express* **31**, 14343–14357 (2023).

24. Weiss, U. & Katz, O. Two-photon lensless micro-endoscopy with in-situ wavefront correction. *Opt. Express* **26**, 28808–28817 (2018).

25. Karbasi, S. et al. Transverse Anderson localization in a disordered glass optical fiber. *Opt Mater. Express* **2**, 1496–1503 (2012).

26. Chen, M. H. & Li, M. J. Observing transverse Anderson localization in random air line based fiber. In: *Proc SPIE 8994, Photonic and Phononic Properties of Engineered Nanostructures IV*, (eds. Adibi, A., Lin, S.-Y. & Scherer, A.) 89941S (SPIE, San Francisco, CA, USA, 2014).

27. Zhao, J. et al. Image transport through meter-long randomly disordered silica-air optical fiber. *Sci. Rep.* **8**, 3065 (2018).

28. Zhao, J. et al. Deep learning imaging through fully-flexible glass-air disordered fiber. *ACS Photonics* **5**, 3930–3935 (2018).

29. Zhao, J. et al. Deep-learning cell imaging through Anderson localizing optical fiber. *Adv. Photonics* **1**, 066001 (2019).

30. Mafi, A. et al. Disordered anderson localization optical fibers for image transport—a review. *J. Lightwave Technol.* **37**, 5652–5659 (2019).

31. Hu, X. W. et al. Learning-supported full-color cell imaging through disordered optical fiber. In *Proc 2020 Conference on Lasers and Electro-Optics*, Washington DC, USA, p. SM2L.5 (Optical Society of America, 2020). https://doi.org/10.1364/CLEO_SI.2020.SM2L.5.

32. Hu, X. W. et al. Robust imaging-free object recognition through anderson localizing optical fiber. *J. Lightwave Technol.* **39**, 920–926 (2021).

33. Zhao, J. et al. Learning-based image transport through disordered optical fibers with transverse anderson localization. *Front. Phys.* **9**, 710351 (2021).

34. Abaie, B. et al. Disorder-induced high-quality wavefront in an Anderson localizing optical fiber. *Optica* **5**, 984–987 (2018).

35. Abdullaev, S. S. & Abdullaev, F. K. On propagation of light in fiber bundles with random parameters. *Radiofizika* **23**, 766–767 (1980).

36. De Raedt, H., Lagendijk, A. & de Vries, P. Transverse localization of light. *Phys. Rev. Lett.* **62**, 47–50 (1989).

37. Karbasi, S., Koch, K. W. & Mafi, A. Multiple-beam propagation in an Anderson localized optical fiber. *Opt. Express* **21**, 305–313 (2013).

38. Schirmacher, W. et al. What is the right theory for anderson localization of light? An experimental test. *Phys. Rev. Lett.* **120**, 067401 (2018).

39. Chen, X. P., Reichenbach, K. L. & Xu, C. Experimental and theoretical analysis of core-to-core coupling on fiber bundle imaging. *Opt. Express* **16**, 21598–21607 (2008).

40. Reichenbach, K. L. & Xu, C. Numerical analysis of light propagation in image fibers or coherent fiber bundles. *Opt. Express* **15**, 2151–2165 (2007).

41. Choi, Y. et al. Scanner-free and wide-field endoscopic imaging by using a single multimode optical fiber. *Phys. Rev. Lett.* **109**, 203901 (2012).

42. Ohayon, S. et al. Minimally invasive multimode optical fiber microendoscope for deep brain fluorescence imaging. *Biomed. Opt. Express* **9**, 1492–1509 (2018).

43. Caravaca-Aguirre, A. M. & Piestun, R. Single multimode fiber endoscope. *Opt. Express* **25**, 1656–1665 (2017).

44. Shabairou, N. et al. Color image identification and reconstruction using artificial neural networks on multimode fiber images: towards an all-optical design. *Opt. Lett.* **43**, 5603–5606 (2018).

45. Kürüm, U. et al. Deep learning enabled real time speckle recognition and hyperspectral imaging using a multimode fiber array. *Opt. Express* **27**, 20965–20979 (2019).

46. Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. In: *Proceedings of the 18th International Conference on Medical Image Computing and Computer-assisted Intervention*, 234–241 (Springer, Munich, Germany, 2015).

47. Isola, P. et al. Efros. Image-to-image translation with conditional adversarial networks. In P*roc 2017 IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 5967–5976. (IEEE, 2017).

48. Karbasi, S. et al. Detailed investigation of the impact of the fiber design parameters on the transverse Anderson localization of light in disordered optical fibers. *Opt. Express* **20**, 18692–18706 (2012).

49. Karbasi, S., Koch, K. W. & Mafi, A. Image transport quality can be improved in disordered waveguides. *Opt. Commun.* **311**, 72–76 (2013).