Research Article | Vol. 31, No. 9 / 24 Apr 2023 / Optics Express 14343

Optics EXPRESS

Adaptive inverse mapping: a model-free semi-supervised learning approach towards robust imaging through dynamic scattering media

XIAOWEN HU,[1] JIAN ZHAO,[2,3,*] JOSE ENRIQUE ANTONIO-LOPEZ,[1] STEFAN GAUSMANN,[1] RODRIGO AMEZCUA CORREA,[1] AND AXEL SCHÜLZGEN[1]

[1] *CREOL, The College of Optics and Photonics, University of Central Florida, Orlando, FL 32816, USA*
[2] *Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215, USA*
[3] *Current affiliation: The Picower Institute for Learning and Memory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[*] *jianzhao@knights.ucf.edu*

Abstract: Imaging through scattering media is a useful and yet demanding task since it involves solving for an inverse mapping from speckle images to object images. It becomes even more challenging when the scattering medium undergoes dynamic changes. Various approaches have been proposed in recent years. However, none of them are able to preserve high image quality without either assuming a finite number of sources for dynamic changes, assuming a thin scattering medium, or requiring access to both ends of the medium. In this paper, we propose an adaptive inverse mapping (AIP) method, which requires no prior knowledge of the dynamic change and only needs output speckle images after initialization. We show that the inverse mapping can be corrected through unsupervised learning if the output speckle images are followed closely. We test the AIP method on two numerical simulations: a dynamic scattering system formulated as an evolving transmission matrix and a telescope with a changing random phase mask at a defocused plane. Then we experimentally apply the AIP method to a multimode-fiber-based imaging system with a changing fiber configuration. Increased robustness in imaging is observed in all three cases. AIP method's high imaging performance demonstrates great potential in imaging through dynamic scattering media.

© 2023 Optica Publishing Group under the terms of the Optica Open Access Publishing Agreement

1. Introduction

Optical imaging through scattering media [1–3] is indispensable to many applications, ranging from underwater imaging [4] and biological tissue imaging [5] to imaging through the atmosphere [6] and non-line-of-sight imaging [7,8]. Unfortunately, light from the object undergoes multiple scattering and forms a noise-like speckle image at the detector [9]. An inverse problem must be solved in order to retrieve the object image. The problem becomes much harder when the dynamic nature of scattering media has to be incorporated, for example, in living tissues [10–12]. Time-varying scattering properties rapidly scramble the optical information and result in decorrelations of the speckle images. Many approaches have been proposed over the past decades to address this issue. Yet none of them work satisfactorily. In the methods utilizing phase conjugation [13,14], wavefront shaping [15,16], or transmission matrix (TM) [17–19], the dynamic change is compensated by fast spatial light modulators (SLMs) [20] or deformable mirror devices (DMDs) [21,22] in feedback control techniques. Although these methods are flexible to any dynamic change, they require access to both ends of the scattering medium, which is often unattainable in real-world applications. Deep learning methods [23–25] establish a robust inverse mapping by training a convolutional neural network (CNN) with numerous

#484252 https://doi.org/10.1364/OE.484252
Journal © 2023 Received 23 Dec 2022; revised 16 Mar 2023; accepted 22 Mar 2023; published 17 Apr 2023

Optics EXPRESS

pairs of speckle and object images collected under different conditions [26–34]. After training, the CNN can reconstruct object images from speckle images. Nevertheless, since the CNN is fixed at the time of the test, any dynamic change that has not been included during training can significantly degrade the reconstructed images. Memory effects for speckle correlation [35–38] enable single-shot imaging, overcoming the disadvantages of the above methods. However, the memory effect method is limited to thin films. As the scattering media get thicker, the speckle correlation drops rapidly and memory effects become neglectable.

Here, we present a general framework, the adaptive inverse mapping (AIP) method, to realize stabilized high-performance imaging through dynamic scattering media. Our AIP approach only requires speckle images and is not restricted to a few predetermined dynamic changes. We show that although the reconstructed images are scrambled, their relationship to the object images is preserved. By utilizing recently developed unpaired image-to-image translation [39], the AIP method enables the inverse mapping to keep close track of the scattering media variations by dynamic corrections. As a proof of concept, we test the AIP method on two numerical simulations: a dynamic scattering medium formulated as an evolving TM, and a telescope with a dynamic random phase mask at a defocused plane. We then experimentally apply the AIP method to a multimode optical fiber (MMF)-based imaging system with a changing fiber configuration. By closely monitoring the output speckle images, the AIP method preserves a high quality of reconstructed images in all three cases. With the universality shown, we see the great potential of the AIP method on increasing robustness in imaging through a wide range of dynamic scattering media.

2. Results

2.1. Principle

We illustrate the AIP method in Fig. 1. At any 'snapshot' of the dynamic scattering medium, e.g., state *i-1*, the scattering medium can be viewed as a forward mapping that takes object images and outputs speckle images. An inverse mapping is then applied to reverse this process to reconstruct object images. At the next state *i*, dynamic changes in the scattering medium perturb the forward mapping so that distorted images will be generated if the same inverse mapping is used. The AIP method corrects the inverse mapping by monitoring the output speckle images from the scattering medium (Fig. 1 (a)). More specifically, we initialize an inverse mapping of the scattering medium at state 0 by training a CNN_0 (see Methods for details) on m pairs of speckles and ground truth object images (Fig. 1 (b)). At any subsequent state i, m speckle outputs are passed through the inverse mapping of the previous state $\mathrm{CNN}_{\mathrm{i}-1}$, generating m distorted reconstructed images (Fig. 1 (c)). Among those m distorted images, n images are used, together with n reserved object images, to train an image restoration cycle-consistent adversarial network i (Restore-$\mathrm{CycleGAN}_{\mathrm{i}}$) (Fig. 1 (c) blue box and arrow). The n reserved object images are randomly selected from the m object images at state 0. Note that not only are these two sets of images unpaired, but the true objects of the n distorted images do not need to be included in the n reserved object images. After learning a translation between distorted images and object images, the Restore-$\mathrm{CycleGAN}_{\mathrm{i}}$ takes all the m distorted images and generates m object images. In this way, we have m pairs of speckle and object images. Finally, a $\mathrm{CNN}_{\mathrm{i}}$ learns a mapping from the m speckle images to the m object images (Fig. 1 (c) green boxes and arrow). As a result, the inverse mapping is adapted to the new state i to recover objects from the speckles. The AIP method is semi-supervised in the sense that paired images are only required at the initialization. In later states, all it needs is output speckle images.

Figure 1 (d) shows the detailed flowchart of a Restore-CycleGAN. Similar to the original CycleGAN [39], two generator networks G_{obj} and G_{dis} try to learn a mapping from the distorted reconstructions to the objects and vice versa. Two discriminators D_{obj} and D_{dis} try to distinguish between the real images in the target domain and the fake images from the generators. The

Optics EXPRESS

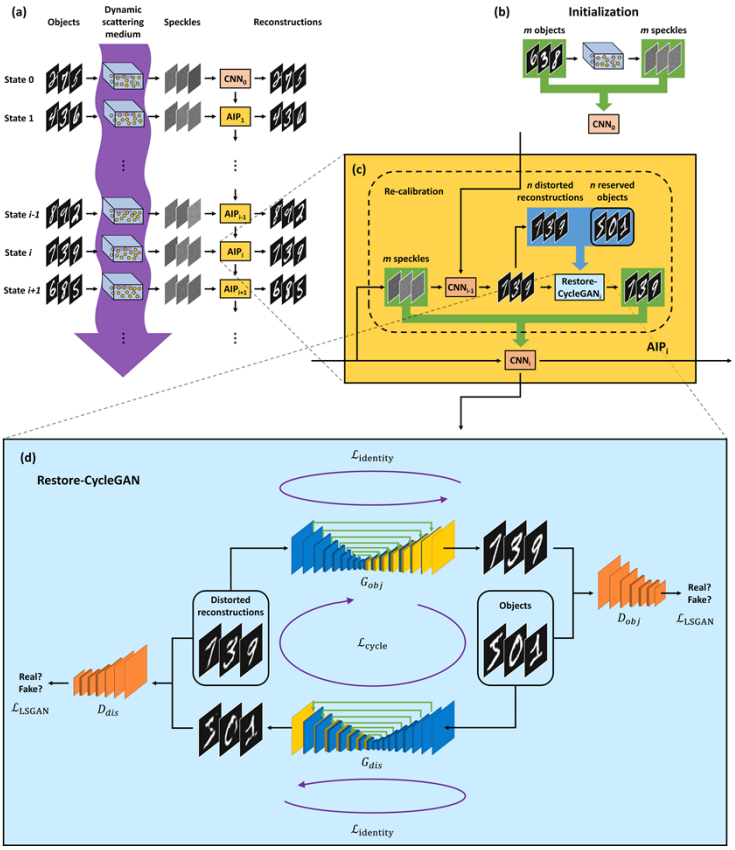

Fig. 1. (a) The diagram of applying the AIP method on imaging through dynamic scattering medium. The AIP method constantly corrects the inverse mapping by monitoring output speckle images from 'snapshot' states of the scattering medium. (b) Initialization of the AIP method. At state 0, m object images are passed through the scattering medium, generating m speckle images. A $\mathrm{CNN_0}$ is trained on these m pairs of images (green boxes) to establish an initial inverse mapping. (c) The detailed workflow of the $\mathrm{AIP_i}$. In the re-calibration stage (dashed box), m speckles among all the speckle images are passed through the inverse mapping $\mathrm{CNN_{i-1}}$ of the previous state *i-1*. Because of the dynamic changes in the scattering medium, $\mathrm{CNN_{i-1}}$ generates m distorted reconstructions. n images are randomly chosen among these m images, together with n reserved object images, to form the training set of $\mathrm{Restore\text{-}CycleGAN_i}$ (blue box). After learning a transition between those unpaired images, the $\mathrm{Restore\text{-}CycleGAN_i}$ takes all the m distorted images and generates m clear images. Therefore, we have m pairs of speckle images and object images (green boxes). Finally, a $\mathrm{CNN_i}$ is trained on these m pairs of images to re-establish the inverse mapping at the state i. Blue box: unsupervised learning. Green boxes: supervised learning. (d) The flowchart of the Restore-CycleGAN in (c). It consists of two generator-discriminator pairs: the object image generator G_{obj} and the discriminator D_{obj}, and the distorted image generator G_{dis} and the discriminator D_{dis}. The least square adversarial loss $\mathcal{L}_{LSGAN}$ is optimized in a min-max game, in which G_{obj} tries to fool D_{obj} by generating object images from distorted images, whereas D_{obj} distinguishes between the real objects and the fake objects generated by G_{obj}. Similarly, there is a $\mathcal{L}_{LSGAN}$ for the reversed direction. The cycle-consistent loss $\mathcal{L}_{cycle}$ enforces an identical output if an image passes through a full translation cycle. The identity mapping loss $\mathcal{L}_{identiy}$ regularizes the generators to have an identity mapping if the input is a real image from the target domain.

Optics EXPRESS

generators and discriminators are optimized in an adversarial game through the least square adversarial loss $\mathcal{L}_{LSGAN}$. Further, the generators are optimized through two more losses: the cycle-consistent loss $\mathcal{L}_{cycle}$ and the identity mapping loss $\mathcal{L}_{identiy}$. $\mathcal{L}_{cycle}$ requires that an image should be unaltered if it goes through a full cycle. $\mathcal{L}_{identiy}$ imposes an identical output if the input is an image from the target domain. We use UNets [40] as the generators in place of ResNets [41] in the CycleGAN [39], inspired by the observation that UNets with skip-connections have weaker priors than ResNets [42]. This enhances the performance of the Restore-CycleGAN on image restoration (Supplement 1). PatchGANs [43] are used as the discriminators. The details of the architecture and the training process of Restore-CycleGANs can be found in Methods.

2.2. Dynamic scattering imaging system as an evolving TM

We test the AIP method on a general case of dynamic scattering imaging systems, where the system is formulated as a complex-valued TM relating the input image to the output image (see Methods for details). We construct a TM by drawing its elements from a complex normal distribution [44–47] with a zero mean and a variance 1, i.e., $CN(0, 1)$. Dynamic changes are introduced by gradually replacing the elements in the TM with new elements from the same complex normal distribution (Fig. 2 (a)). The imaging objects are Modified National Institute of Standards and Technology (MNIST) handwritten digits [48] resized to 256×256. Starting from an initial inverse mapping CNN_0, the AIP method is applied with $m = 5000$ and $n = 1000$ every time when the percentage of the substituted elements p in the TM is increased by $\Delta p = 12.5\%$. The performance of image reconstruction at all states is evaluated on a separate set of 500 test images. The results are shown Fig. 2 (b-d). As the p increases, the reconstructed images by CNN_0 become more and more unrecognizable (Fig. 2 (b)). In comparison, the AIP method stabilizes image reconstruction by improving the inverse mapping from the preceding state. The improvement is quantified in Fig. 2 (c), where we plot the averages and standard deviations of the mean absolute errors (MAEs) between the reconstructions and the objects at different states. For every AIP_{i-1}, the MAE increases when the dynamic scattering system transforms to a later state i. The AIP_i then corrects the inverse mapping and lowers the MAE. Figure 2 (d) shows the output speckle decorrelations as a function of p. The amount of speckle decorrelation is evaluated through the Pearson correlation coefficient (PCC). When the output speckles at $p = 50\%$ are already decorrelated to the speckles at the first state, i.e., PCC < 1/e, good image reconstruction is still preserved. We attribute this to the fact that the speckles remain highly correlated to the neighboring state if the system is traced closely. This is further confirmed through a comparison with the results by the AIP method with an increased $\Delta p = 25\%$ (Supplement 1).

2.3. Dynamic telescopic imaging system

Next, we numerically simulate the use of the AIP method on a dynamic telescopic imaging system, where a changing random phase mask is located at a defocused plane [25,26,30,49,50] (Fig. 3). The focal lengths f_1 and f_2 of the two lenses are chosen to be 250 mm and 150 mm, respectively. The object has a size of 10.24 mm × 10.24 mm. A random phase mask is placed z = 15 mm in front of the first lens L1. The transmittance of the phase mask $t(x, y)$ is formulated as Eqs. ((1)–(2)) given in Fig. 3. $\Delta n = 0.52$ is the refractive index difference between the phase mask and air. $\lambda = 632.8$ nm is the wavelength. $D(x, y)$ is a random height field. $W(x, y)$ is a set of random height values drawn from the normal distribution $N(\mu, \sigma_0)$ at discrete sample location (x, y), and $K(\sigma)$ is a zero-mean Gaussian smoothing kernel with a full width half maximum (FWHM) of σ. Moreover, the elements in the matrix $W(x, y)$ are gradually replaced by values from the same normal distribution (enlarged figure in Fig. 3). Thus, the phase mask is changing towards a different phase mask. μ, σ_0, and σ are chosen to be 16 µm, 5 µm and 4 µm, respectively. Imaging through the system is simulated using Fourier optics [51,52].

Optics EXPRESS

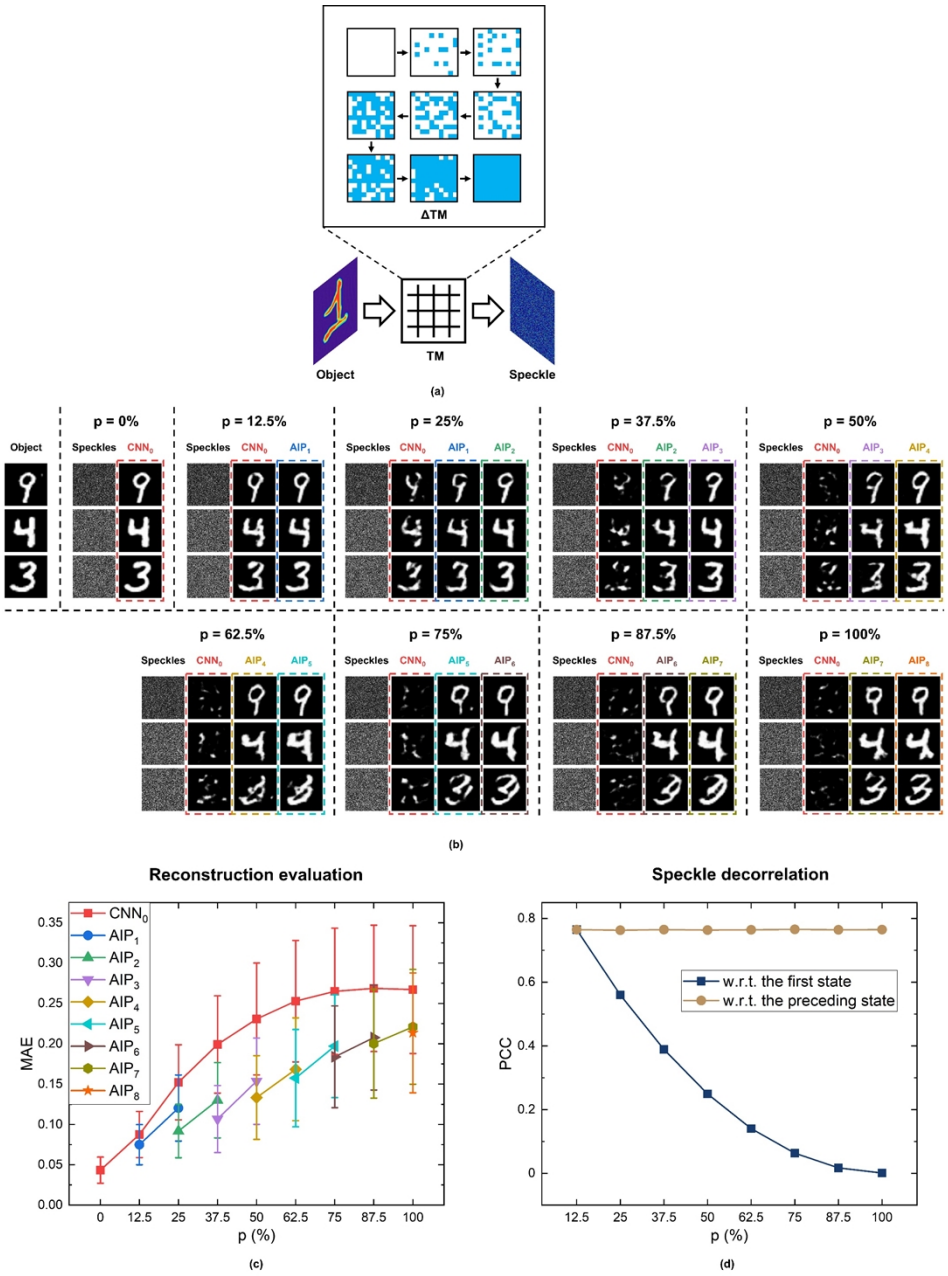

Fig. 2. (a) Schematic of a dynamic scattering system formulated as an evolving TM. The TM transforms objects into speckles. Enlarged figure: The dynamic changes in the TM. The elements in the original TM (white) are gradually replaced by new elements (blue). (b) The results of image reconstruction obtained by applying the AIP method to the dynamic scattering system shown in (a). Object: The input object images (top left column); Speckle: the output speckle images when the percentage of the substituted elements p in the TM is increased from 0% to 100% with a step of 12.5%; CNN_0 and AIP_i: the reconstructed images from CNN_0 and the ith AIP. The dashed bounding boxes with the same color represent the reconstructions from a particular AIP or CNN_0. (c) The averages and standard deviations of the MAEs of the test reconstructions from the AIP_i and CNN_0. The colors of the symbols correspond to the colors of the bounding boxes in (b). (d) The PCCs of the speckle images with respect to (w.r.t.) the speckle images from the first state $p = 0\%$ (dark blue line), and w.r.t. the speckle images from the preceding state (brown line).

Optics EXPRESS

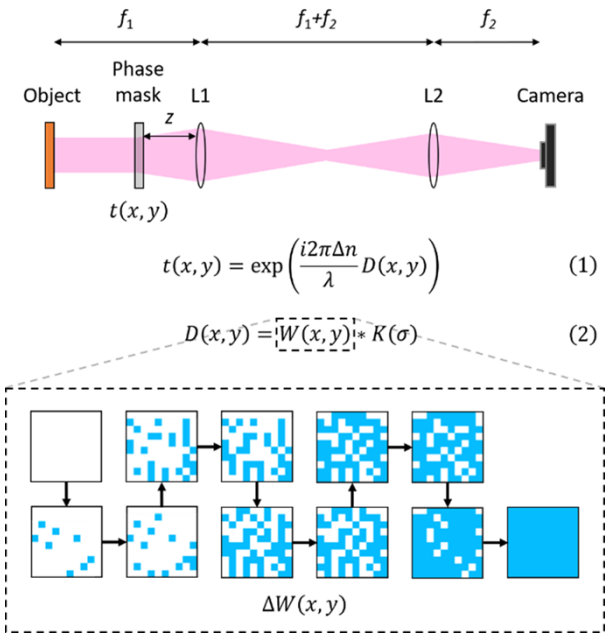

Fig. 3. Schematic of the simulated dynamic telescopic imaging system. A changing random phase mask is placed at a defocused plane. L1, L2: lenses. Eq. ((1)–(2)): formulas of the transmittance of the random phase mask $t(x, y)$. Δn: the refractive index difference between the phase mask and air. λ: wavelength. $D(x, y)$: a random height field. $W(x, y)$: a set of random height values drawn from the normal distribution $N(\mu, \sigma_0)$ at discrete sample locations (x, y). $K(\sigma)$: a zero-mean Gaussian smoothing kernel with FWHM value of σ. Enlarged: the elements in $W(x, y)$ (white) are gradually replaced by new elements (blue).

The objects are extended MNIST (EMNIST) handwritten letters [53]. We initialize an inverse mapping CNN_0 of the system with the original phase mask. The AIP method is then adopted to stabilize the imaging reconstruction every time when the percentage of the substituted elements p in the $W(x, y)$ is increased by $\Delta p = 10\%$. We choose $m = 5000$, and $n = 1000$ in the AIP method. A separate set containing 500 images is used to test the performance of image reconstruction at all states. The results are shown in Fig. 4. As p increases, the quality of images reconstructed by CNN_0 degrades (Fig. 4 (a)). In comparison, the AIP method stabilizes the image reconstruction (the last column at each state in Fig. 4 (a)). Good visual quality is still maintained when the original phase mask has been completely replaced by a new phase mask ($p = 100\%$). In Fig. 4 (b), we plot the averages and standard deviations of the MAEs between the reconstructions and the objects at different states. Similar trends as in Fig. 2 (c) can be observed, where the AIP method corrects the inverse mapping of the preceding state. Figure 4 (c) shows the averages and standard deviations of the PCC scores between the output speckles. While the speckles become more and more decorrelated from the speckles at the first state (dark blue line), they remain highly correlated with the speckles from the preceding states (brown line). This indicates the necessity of tracing the state of the system closely. We compare to the results when the AIP method is applied directly at the final state. Degradation on the reconstructions can be seen from both the additional column in Fig. 4 (a) and the dark gray star in Fig. 4 (b) at $p = 100\%$.

Optics EXPRESS

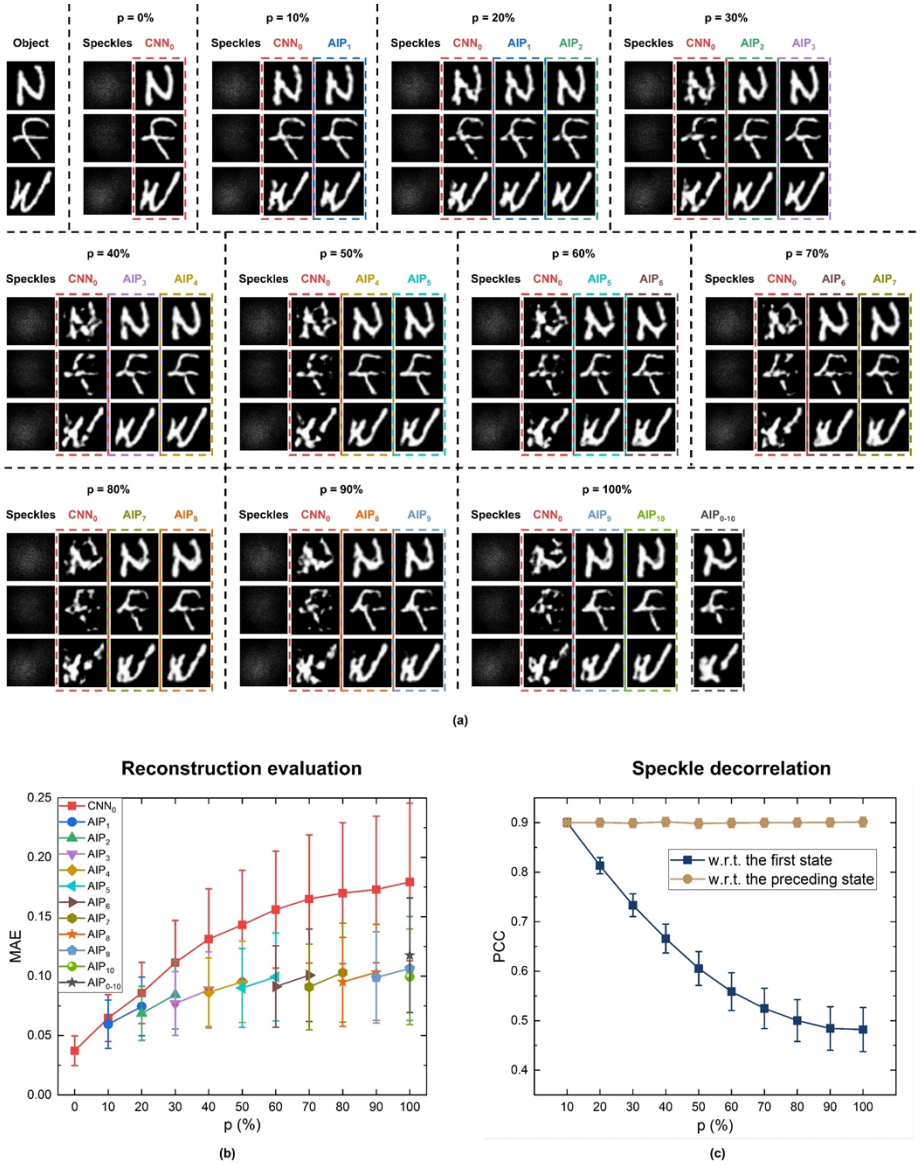

Fig. 4. (a) The results of image reconstruction obtained by applying the AIP method to the dynamic telescopic imaging system shown in Fig. 3. Object: the input object images (top left column); Speckles: the output speckle images when the percentage p of the substituted elements in $W(x, y)$ is increased from 0% to 100% with a step of 10%; CNN_0 and AIP_i: the reconstructed images from CNN_0 and the ith AIP. The dashed bounding boxes with the same color represent the reconstructions from a particular AIP or CNN_0. (b) The averages and standard deviations of the MAEs of the test reconstructions from the AIP_i and CNN_0. The colors of the symbols correspond to the colors of the bounding boxes in (a). (c) The PCCs of the speckle images w.r.t. the speckle images from the first state $p = 0\%$ (dark blue line), and w.r.t. the speckle images from the preceding state (brown line).

Optics EXPRESS

2.4. Dynamic MMF-based imaging system

Based on the numerical validations, we further experimentally apply the AIP method on a dynamic MMF-based imaging system. MMFs have shown great potential for endoscopic applications [54–57], thanks to their miniature sizes and high mode densities. Recently burgeoning deep learning technology further enhances the performance of MMF-based imaging systems [27, 29,58,59]. However, because of MMFs' intrinsic scattering properties, it is extremely challenging to realize image transport in a dynamic MMF [27,29], where the mode coupling is highly sensitive to perturbations. Here, we validate the AIP method by tackling the dynamic MMF-based image transport problem. The schematic of the setup is shown in Fig. 5. We illuminate a digital micromirror device (DMD) (ViALUX, V-7000) using a laser at 632.8 nm. The laser beam is expanded by lenses L1 and L2. MNIST handwritten digits [48] are binarized, resized to 128×128, and displayed on the DMD as the objects. We demagnify and couple the objects into a 50-cm-long MMF (Thorlabs, FG105LCA) by a tube lens L3 (f = 200 mm) and a 20x microscope objective MO1 (NA = 0.75). The MMF is placed on a translation stage with two optical posts (Supplement 1). The movement of the translation stage changes the fiber bending configuration, resulting in varied output speckles. The output speckles are magnified and projected onto a CCD camera (Manta G-145C) by a combination of 20x MO2 (NA = 0.75) and L3 (f = 200 mm). For image processing, both the objects and the speckles are resized to 256×256. After initializing an inverse mapping CNN_0, we apply the AIP method to stabilize the image reconstruction dynamically when the translation stage is moving with a step of 5 µm. m and n are chosen to be 5000 and 1000 for the AIP method implementation, respectively. A separate dataset containing 500 images is used to test the image reconstruction performance at all states.

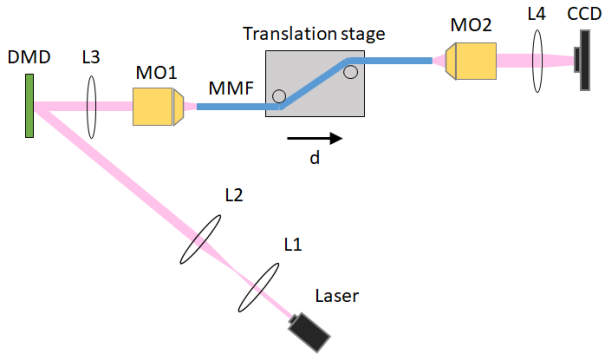

Fig. 5. Schematic of the MMF-based imaging system. DMD: Digital micromirror device. MO1, MO2: microscope objectives. L1, L2, L3, L4: lenses. CCD: CCD camera. The MMF is placed on a translation stage with two posts (the gray rectangle). Mechanical perturbations to the MMF are applied by translating the stage with a distance of d.

Figure 6 (a) shows the objects, the MMF outputs, and the reconstructed images of AIPs and CNN_0 at each translation distance d. As d increases, the output speckles from the MMF gradually decorrelate with the initial state (see the Visualization 1). During this dynamic process, the reconstructions from CNN_0 are increasingly distorted, while that from the AIP method maintains high visual qualities (the last column at each state in Fig. 6 (a)). The image qualities at each state are quantified through the MAEs between the ground truths and the reconstructions (Fig. 6 (b)). Similar to the observations in Fig. 2 (c) and Fig. 4 (b), the AIP method slows down the increase of the MAEs. In Fig. 6 (c), we calculate the PCC scores between the output speckles in the circular region of the MMF. Decorrelations to the initial state with PCC < 1/e occur when d is larger than 10 µm (dark blue line). Nevertheless, highly correlated neighboring states (brown line) ensure the success of the AIP method. This is further verified when we skip all the intermediate states

Optics EXPRESS

and directly apply the AIP method to the last state. Degraded image quality can be observed from both the additional column in Fig. 6 (a) and the dark brown triangle in Fig. 6 (b) at a translation distance of 25 μm.

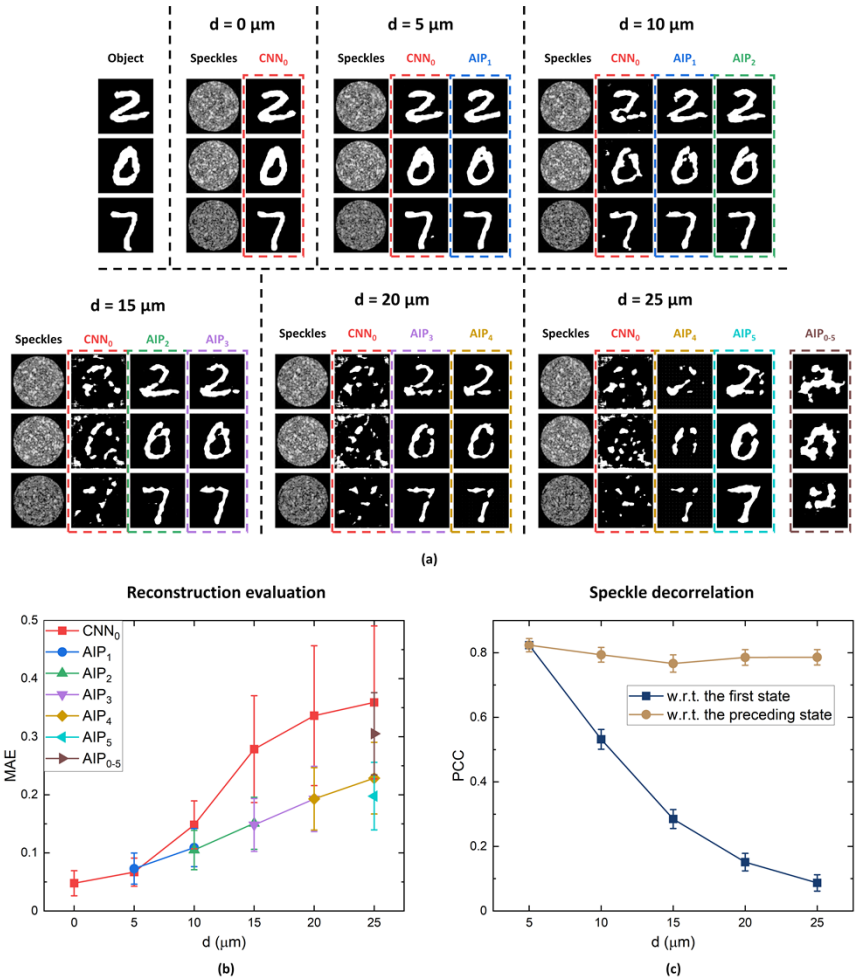

Fig. 6. (a) The results of image reconstruction obtained by applying the AIP method to the dynamic MMF-based imaging system shown in Fig. 5. Object: the input object images (top left column); Speckles: the output speckle images when the distance d is increased from 0 to 25 μm with a step of 5 μm; CNN_0 and AIP_i: the reconstructed images from CNN_0 and the ith AIP. The dashed bounding boxes with the same color represent the reconstructions from a particular AIP or CNN_0. (b) The averages and standard deviations of the MAEs of the test reconstructions from the AIP_i and CNN_0. The colors of the symbols correspond to the colors of the bounding boxes in (a). (c) The PCCs of the speckle images w.r.t. the speckle images from the first state $d = 0$ μm (dark blue line), and w.r.t. the speckle images from the preceding state (brown line). PCCs are calculated for the circular regions of the fiber outputs.

Optics EXPRESS

3. Discussion and conclusion

3.1. Semi-supervised learning

In the AIP method, object images are only collected at the initialization. They are used in two aspects. First, paired object images and speckle images form a training set to establish an initial inverse mapping through supervised learning. Second, a small part of object images is broadcast to the later states to stabilize the imaging performance. After the initialization, the AIP method only requires output speckle images from the dynamic scattering medium. Unsupervised learning is utilized to find a mapping from distorted reconstructed images to object images. Therefore, the AIP method eliminates the need for access to both ends of the dynamic scattering medium, as required in methods using feedback control [20–22]. This makes the AIP method easy to implement in most real-world applications, where only the output end is accessible during image acquisition.

3.2. Flexibility

Unlike other data-driven approaches [26–30], the AIP method does not make any assumptions on the system's dynamics. It can be applied to irregular or unpredictable system variations as long as the change between neighboring states is not significant. Under such a condition, there is an implicit connection between the distorted images and the objects. At any state i, the scattering imaging system takes object images x and generates speckles through a forward mapping $F_i(\cdot)$. If we take $\mathrm{AIP_i}$ as an approximation of the system inverse mapping ${F_i}^{-1}(\cdot)$, the distorted reconstructions at state $i+1$ can be represented by ${F_i}^{-1}(F_{i+1}(x))$. The weights of the generator in the Restore-CycleGAN are initialized in a way (see Methods 4.1) that the generator starts at the distorted reconstructions ${F_i}^{-1}(F_{i+1}(x))$ and looks for an image translation to a nearby local optimum. Under the condition of $F_{i+1}(\cdot) \approx F_i(\cdot)$, x would be the closest optimum to ${F_i}^{-1}(F_{i+1}(x))$ in the high-dimensional space. Therefore, the generator will translate the distorted reconstructions ${F_i}^{-1}(F_{i+1}(x))$ to their ground truths x, and correct the inverse mapping if the dynamic scattering medium is traced closely.

3.3. Universality

For linear propagation scattering media, the forward mapping operator $F_i(\cdot)$ reduces to a TM T_i. Thus, the dynamic changes to the medium result in different transformations of the TMs. In the first case of an evolving TM, dynamic changes simply replace elements in the TM. In the dynamic telescopic imaging system, the transformation of the TM is more implicit. In the dynamic MMF-based imaging system, under the assumption that the fiber deformation does not change the eigenmodes [60], the TM T_{i+1} at state $i+1$ can be written as:

$$T_{i+1} = P^{\dagger}\Lambda_{i+1}P, \tag{1}$$

$$\Lambda_{i+1} = D_{i+1}\Lambda_i, \tag{2}$$

where Λ_{i+1} is the diagonal eigenvalue matrix at the state $i+1$. P is the projection matrix to project the input onto the eigenmode basis. Fiber deformations introduce a diagonal deformation operator D_{i+1} that changes the eigenvalue matrix Λ_i of the previous state i. The AIP method corrects the inverse mappings for the TM transformations of all three cases. This shows the universality of the AIP method. The universality is further demonstrated in the Supplement 1, where we apply the AIP method to a disordered optical fiber imaging system with dynamic imaging depths.

3.4. Perspectives

Improvements of the AIP method can be made in several ways. First, while the AIP method could be applied in many imaging tasks satisfying the quasi-static assumption for re-calibrations,

Optics EXPRESS

such as fiber-based deep-brain imaging on neurons [55,56], it still faces challenges in imaging through fast-varying scattering systems. The speed of the current AIP method is limited by the acquisition time of collecting m speckle images (10^3 to 10^4 depending on the complexity of the objects) to correct the inverse mapping. During the image acquisition, the AIP method assumes a quasi-static scattering system. This means that the speckle image acquisition time should be much smaller than the speckle decorrelation time. The image acquisition time is determined by two factors: the frame rate of the camera and the number of images required. While the former is limited by the hardware, much effort can be made to reduce the latter. In the current method, the image reconstruction CNN and the Restore-CycleGAN operate separately. The CNN generates distorted reconstructions and the Restore-CycleGAN finds the connection between distorted reconstructions and object images. To make more efficient use of a reduced number of speckle images, in future studies, the training of the reconstruction CNN and the Restore-CycleGAN can be done interactively so that improvements on one lead to improvements on the other. Second, while the trained neural networks from the previous states can still be retrieved, the AIP method only uses the information from the last state to correct the inverse mapping of the current state. In future work, the information of all the preceding states could be utilized to establish a more robust inverse mapping to improve the image quality of new states. Moreover, the current work lacks start/stop criteria for the initiation and termination of the AIP method. Since the dynamics of the system are unknown to us, frequent executions of the AIP would be necessary to avoid losing track of system changes. In future work, additional metrics for evaluating the imaging performance should be developed to establish start/stop criteria so that the image quality could be preserved while minimizing the computational time. Finally, a fine-tuned initial CNN may not be necessary due to the Restore-CycleGAN's capabilities of correcting the inverse mapping. In the future, training a coarse CNN in the initialization step using simulations or simpler experiments could be investigated to further facilitate the AIP method.

In conclusion, we show that the inverse mapping of a dynamic scattering medium can be corrected through unsupervised learning if the medium is traced closely. We demonstrate the preservation of good image quality by the AIP method in three showcase dynamic scattering systems. The advantages of semi-supervised learning and its flexibility make the AIP method a promising candidate to improve imaging through dynamic scattering media without prior knowledge of dynamic changes.

4. Methods

4.1. Architectures and training processes of the Restore-CycleGAN

We use PatchGAN [43] as the discriminator network. The PatchGAN looks into patches of an input image and predicts whether they come from a real or a fake image. It consists of two input and output layers as well as five blocks in between (Supplement 1). The last four blocks consist of Convolutional/Instance-Normalization [61] /Leaky-ReLU layers, whereas the first block omits the Instance-Normalization layer. All convolutional filters in these five blocks, except the last one, have a size of 4×4 and a stride of 2. A convolutional layer is added after these five blocks to generate the final output.

We use UNet [40] as the generator network. The UNet has an encoder-decoder architecture with skip connections between layers in the encoder and decoder (Supplement 1). The input image is first down-sampled to a bottleneck layer by the encoder, which consists of convolutional layers with a kernel size of 4×4 and a stride of 2. The decoder then up-samples to the output image using transpose convolutional layers. Dropout layers with a rate of 0.5 are added to the decoder. The size of the input images is 256×256. All the weights in the PatchGAN and UNet are initialized through a random Gaussian distribution with a zero mean and a standard deviation of 0.02.

Optics EXPRESS

The real images from the target domain are labeled as '1', whereas the fake images from the generator are labeled as '0'. The generator network G_{obj} has the following loss function:

$$\begin{aligned}\mathcal{L}_{G_{obj}} &= \mathbb{E}_y[(D_{obj}(G_{obj}(CNN_i(y))) - 1)^2] \\ &+ \alpha_1 \mathbb{E}_x[||G_{obj}(G_{dis}(x)) - x||_1] \\ &+ \alpha_1 \mathbb{E}_y[||G_{dis}(G_{obj}(CNN_i(y))) - CNN_i(y)||_1] \\ &+ \alpha_2 \mathbb{E}_x[||G_{obj}(x) - x||_1]\end{aligned} \tag{3}$$

where x and y are the object and speckle images, respectively. The first term in the loss function is the least square adversarial loss $\mathcal{L}_{LSGAN}$. The second and third terms are the cycle-consistent losses $\mathcal{L}_{cycle}$ in both directions. The fourth term is the identity mapping loss $\mathcal{L}_{identiy}$. α_1 and α_2 control the weighting among the losses. α_1 and α_2 are chosen to be 10 and 5, respectively. The weights in D_{obj} and G_{dis} are fixed when G_{obj} is being trained. The loss function of the discriminator network D_{obj} is the least square adversarial loss $\mathcal{L}_{LSGAN}$:

$$\mathcal{L}_{D_{obj}} = \mathbb{E}_x[(D_{obj}(x) - 1)^2] + \mathbb{E}_y[D_{obj}(G_{obj}(CNN_i(y)))^2] \tag{4}$$

To train D_{obj}, the real object images are randomly selected from all the object images, whereas the fake images are randomly selected from a pool of 50 fake images generated by G_{obj}. The pool is randomly updated through newly generated fake images. The loss of the discriminators is weighted as one half to the loss of the generators. Similarly, we have the loss functions $\mathcal{L}_{G_{dis}}$ and $\mathcal{L}_{D_{dis}}$ for G_{dis} and D_{dis}, respectively. The discriminators and generators are trained for 100 epochs using a batch size of 1 and an Adam optimizer with a learning rate of 0.0002 and an exponential decay rate for the first momentum $\beta_1 = 0.5$. The training takes ~40 hours on a dual-GPU (GeForce GTX 1080 Ti) desktop.

4.2. Architectures and training processes of the CNN

The CNNs used to establish the inverse mapping at all states have the same architecture as the generators in the Restore-CycleGAN. The weights of the CNNs are also initialized in the same way. The MAE is chosen as the loss function. The m pairs of speckle images and the corresponding Restore-CycleGAN reconstructions at each state are split into a training set and a validation set. The CNN is only trained on the training set, while the training process is monitored through the validation set. We train the CNNs for 200 epochs using an Adam optimizer with a learning rate of 0.005 and the exponential decay rate for the first momentum $\beta_1 = 0.9$. The batch size is 512.

4.3. Generating output images through a TM

The intensity of an input image is first converted to an electric field matrix E_{in}. The matrix is flattened into a vector and multiplied by the complex-valued TM (Eq. (7)). The resulting vector is then rearranged to an output electric field matrix E_{out}. Finally, the output electric field matrix is converted back to the output intensity.

$$E_{out} = \begin{pmatrix} E_{out,1} \\ E_{out,2} \\ \vdots \\ E_{out,n} \end{pmatrix} = \begin{pmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & \ddots & \vdots \\ t_{n1} & \cdots & t_{nn} \end{pmatrix} \begin{pmatrix} E_{in,1} \\ E_{in,2} \\ \vdots \\ E_{in,n} \end{pmatrix} = TE_{in} \tag{5}$$

Optics EXPRESS

4.4. Metrics

The definitions of the metrics used in this paper to evaluate the similarity between two images are the following:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|x_i - y_i| \tag{6}$$

$$PCC = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} \tag{7}$$

where the pixel values of the two images are flattened to two vectors: x and y, both containing n elements. $\bar{x}$ and $\bar{y}$ are the average values of elements in x and y, respectively.

Disclosures. The authors declare no conflicts of interest.

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

Supplemental document. See Supplement 1 for supporting content.

References

1. A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, 1978).
2. *Waves and Imaging through Complex Media*, 1ed. (Springer, Dordrecht, 2001), pp., 460.
3. V. I. Tatarski and R. A. Silverman, *Wave Propagation in a Turbulent Medium* (Dover Publications Inc., 2016).
4. M. Sheinin and Y. Y. Schechner, "The Next Best Underwater View," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2016), 3764–3773.
5. A. P. Gibson, J. C. Hebden, and S. R. Arridge, "Recent advances in diffuse optical imaging," Phys. Med. Biol. **50**(4), R1–R43 (2005).
6. J. S. Lee, L. Jurkevich, P. Dewaele, P. Wambacq, and A. Oosterlinck, "Speckle filtering of synthetic aperture radar images: A review," Remote Sensing Reviews **8**(4), 313–340 (1994).
7. I. Freund, "Looking through walls and around corners," Phys. A **168**(1), 49–65 (1990).
8. O. Katz, E. Small, and Y. Silberberg, "Looking around corners and through thin turbid layers in real time with scattered incoherent light," Nat. Photonics **6**(8), 549–553 (2012).
9. J. W. Goodman, *Speckle Phenomena in Optics: Theory and Applications, Second Edition* (SPIE Press, 2020).
10. M. Jang, H. Ruan, I. M. Vellekoop, B. Judkewitz, E. Chung, and C. Yang, "Relation between speckle decorrelation and optical phase conjugation (OPC)-based turbidity suppression through dynamic scattering media: a study on in vivo mouse skin," Biomed. Opt. Express **6**(1), 72–85 (2015).
11. Y. Liu, P. Lai, C. Ma, X. Xu, A. A. Grabar, and L. V. Wang, "Optical focusing deep inside dynamic scattering media with near-infrared time-reversed ultrasonically encoded (TRUE) light," Nat. Commun. **6**(1), 5904 (2015).
12. M. M. Qureshi, J. Brake, H.-J. Jeon, H. Ruan, Y. Liu, A. M. Safi, T. J. Eom, C. Yang, and E. Chung, "In vivo study of optical speckle decorrelation time across depths in the mouse brain," Biomed. Opt. Express **8**(11), 4855–4864 (2017).
13. Z. Yaqoob, D. Psaltis, M. S. Feld, and C. Yang, "Optical phase conjugation for turbidity suppression in biological samples," Nat. Photonics **2**(2), 110–115 (2008).
14. T. R. Hillman, T. Yamauchi, W. Choi, R. R. Dasari, M. S. Feld, Y. Park, and Z. Yaqoob, "Digital optical phase conjugation for delivering two-dimensional images through turbid media," Sci. Rep. **3**(1), 1909 (2013).
15. A. P. Mosk, A. Lagendijk, G. Lerosey, and M. Fink, "Controlling waves in space and time for imaging and focusing in complex media," Nat. Photonics **6**(5), 283–292 (2012).
16. D. Aizik, I. Gkioulekas, and A. Levin, "Fluorescent wavefront shaping using incoherent iterative phase conjugation," Optica **9**(7), 746 (2022).
17. S. M. Popoff, G. Lerosey, R. Carminati, M. Fink, A. C. Boccara, and S. Gigan, "Measuring the Transmission Matrix in Optics: An Approach to the Study and Control of Light Propagation in Disordered Media," Phys. Rev. Lett. **104**(10), 100601 (2010).
18. S. Popoff, G. Lerosey, M. Fink, A. C. Boccara, and S. Gigan, "Image transmission through an opaque material," Nat. Commun. **1**(1), 81 (2010).
19. M. Kim, W. Choi, Y. Choi, C. Yoon, and W. Choi, "Transmission matrix of a scattering medium and its applications in biophotonics," Opt. Express **23**(10), 12648–12668 (2015).
20. E. Tajahuerce, V. Durán, P. Clemente, E. Irles, F. Soldevila, P. Andrés, and J. Lancis, "Image transmission through dynamic scattering media by single-pixel photodetection," Opt. Express **22**(14), 16945–16955 (2014).
21. D. B. Conkey, A. M. Caravaca-Aguirre, and R. Piestun, "High-speed scattering medium characterization with application to focusing light through turbid media," Opt. Express **20**(2), 1733–1740 (2012).

Optics EXPRESS

22. A. M. Caravaca-Aguirre, E. Niv, D. B. Conkey, and R. Piestun, "Real-time resilient focusing through a bending multimode fiber," Opt. Express **21**(10), 12881–12887 (2013).
23. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature **521**(7553), 436–444 (2015).
24. A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," Optica **4**(9), 1117–1125 (2017).
25. S. Li, M. Deng, J. Lee, A. Sinha, and G. Barbastathis, "Imaging through glass diffusers using densely connected convolutional networks," Optica **5**(7), 803–813 (2018).
26. Y. Li, Y. Xue, and L. Tian, "Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media," Optica **5**(10), 1181–1190 (2018).
27. N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, "Learning to see through multimode fibers," Optica **5**(8), 960–966 (2018).
28. Y. Sun, J. Shi, L. Sun, J. Fan, and G. Zeng, "Image reconstruction through dynamic scattering media based on deep learning," Opt. Express **27**(11), 16032–16046 (2019).
29. P. Fan, T. Zhao, and L. Su, "Deep learning the high variability and randomness inside multimode fibers," Opt. Express **27**(15), 20241–20258 (2019).
30. Y. Li, S. Cheng, Y. Xue, and L. Tian, "Displacement-agnostic coherent imaging through scatter with an interpretable deep neural network," Opt. Express **29**(2), 2244–2257 (2021).
31. Y. Sun, X. Wu, Y. Zheng, J. Fan, and G. Zeng, "Scalable non-invasive imaging through dynamic scattering media at low photon flux," Opt. Lasers Eng. **144**, 106641 (2021).
32. H. Wu, X. Meng, X. Yang, X. Li, and Y. Yin, "Single shot real-time high-resolution imaging through dynamic turbid media based on deep learning," Opt. Lasers Eng. **149**, 106819 (2022).
33. S. Zhu, E. Guo, J. Gu, L. Bai, and J. Han, "Imaging through unknown scattering media based on physics-informed learning," Photonics Res. **9**, B210 (2021).
34. S. Resisi, S. M. Popoff, and Y. Bromberg, "Image Transmission Through a Dynamically Perturbed Multimode Fiber by Deep Learning," Laser Photonics Rev. **15**, 1 (2021).
35. S. Feng, C. Kane, P. A. Lee, and A. D. Stone, "Correlations and Fluctuations of Coherent Wave Transmission through Disordered Media," Phys. Rev. Lett. **61**(7), 834–837 (1988).
36. I. Freund, M. Rosenbluh, and S. Feng, "Memory Effects in Propagation of Optical Waves through Disordered Media," Phys. Rev. Lett. **61**(20), 2328–2331 (1988).
37. O. Katz, P. Heidmann, M. Fink, and S. Gigan, "Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations," Nat. Photonics **8**(10), 784–790 (2014).
38. W.-Y. Chen, M. O'Toole, A. C. Sankaranarayanan, and A. Levin, "Enhancing speckle statistics for imaging inside scattering media," Optica **9**(12), 1408 (2022).
39. J.-Y. Zhu, T. Park, P. Isola, and A. A. E. Efros, "Unpaired Image-To-Image Translation Using Cycle-Consistent Adversarial Networks," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, (2017).
40. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, (Springer International Publishing, 2015), 234–241.
41. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2016), 770–778.
42. D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep Image Prior," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2018).
43. P. I. Isola, J.-Y. Z. Zhu, T. Z. Zhou, and A. A. Efros, "Image-To-Image Translation With Conditional Adversarial Networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2017).
44. J. W. Goodman, G. V. Skrockij, and A. A. Kokin, *Statistical Optics* (Wiley, 1985).
45. N. Garcia and A. Z. Genack, "Crossover to strong intensity correlation for microwave radiation in random media," Phys. Rev. Lett. **63**(16), 1678–1681 (1989).
46. J. B. Pendry, A. MacKinnon, and P. J. Roberts, "Universality classes and fluctuations in disordered systems," Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences **437**, 67–83 (1992).
47. J. Xu, H. Ruan, Y. Liu, H. Zhou, and C. Yang, "Focusing light through scattering media by transmission matrix inversion," Opt. Express **25**(22), 27234–27246 (2017).
48. Y. LeCun, C. Cortes, and C. Burges, *MNIST handwritten digit database*, (Florham Park, NJ, USA, 2010).
49. J. Mertz, H. Paudel, and T. G. Bifano, "Field of view advantage of conjugate adaptive optics in microscopy applications," Appl. Opt. **54**(11), 3498–3506 (2015).
50. J. Li, D. R. Beaulieu, H. Paudel, R. Barankov, T. G. Bifano, and J. Mertz, "Conjugate adaptive optics in widefield microscopy with an extended-source wavefront sensor," Optica **2**(8), 682–688 (2015).
51. J. W. Goodman and M. E. Cox, "Introduction to Fourier Optics," Phys. Today **22**(4), 97–101 (1969).
52. J. D. Schmidt, *Numerical Simulation of Optical Wave Propagation with Examples in MATLAB* (2010).
53. G. Cohen, S. Afshar, J. Tapson, and V. A. Schaik, "EMNIST: Extending MNIST to handwritten letters," in *2017 International Joint Conference on Neural Networks (IJCNN)*, (2017), 2921–2926.
54. Y. Choi, C. Yoon, M. Kim, T. D. Yang, C. Fang-Yen, R. R. Dasari, K. J. Lee, and W. Choi, "Scanner-Free and Wide-Field Endoscopic Imaging by Using a Single Multimode Optical Fiber," Phys. Rev. Lett. **109**(20), 203901 (2012).

Optics EXPRESS

55. S. Turtaev, I. T. Leite, T. Altwegg-Boussac, J. M. P. Pakan, N. L. Rochefort, and T. Čižmár, "High-fidelity multimode fibre-based endoscopy for deep brain in vivo imaging," Light: Sci. Appl. **7**(1), 2029 (2018).
56. S. A. Vasquez-Lopez, R. Turcotte, V. Koren, M. Plöschner, Z. Padamsey, M. J. Booth, T. Čižmár, and N. J. Emptage, "Subcellular spatial resolution achieved for deep-brain imaging in vivo using a minimally invasive multimode fiber," Light: Sci. Appl. **7**, 1 (2018).
57. P. Caramazza, O. Moran, R. Murray-Smith, and D. Faccio, "Transmission of natural scene images through a multimode fibre," Nat. Commun. **10**, 1 (2019).
58. N. Shabairou, E. Cohen, O. Wagner, D. Malka, and Z. Zalevsky, "Color image identification and reconstruction using artificial neural networks on multimode fiber images: towards an all-optical design," Opt. Lett. **43**(22), 5603 (2018).
59. B. Rahmani, D. Loterie, G. Konstantinou, D. Psaltis, and C. Moser, "Multimode optical fiber transmission with a deep learning network," Light: Sci. Appl. **7**, 69 (2018).
60. M. Plöschner, T. Tyc, and T. Čižmár, "Seeing through chaos in multimode fibres," Nat. Photonics **9**(8), 529–535 (2015).
61. D. U. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance Normalization: The Missing Ingredient for Fast Stylization," arXiv, arXiv:1607.08022 (2016).

132.170.212.17 {ts '2023-04-17 14:55:15'}