PHOTONICS Research

Imaging through unknown scattering media based on physics-informed learning

Shuo Zhu,^{1,†} Enlai Guo,^{1,2,†} Die Gu,¹ Lianfa Bai,¹ and Jing Han^{1,3}

¹Jiangsu Key Laboratory of Spectral Imaging and Intelligent Sense, Nanjing University of Science and Technology, Nanjing 210094, China ²e-mail: njustgel@njust.edu.cn

³e-mail: eohj@njust.edu.cn

Received 2 December 2020; revised 26 January 2021; accepted 21 February 2021; posted 22 February 2021 (Doc. ID 416551); published 15 April 2021

Imaging through scattering media is one of the hotspots in the optical field, and impressive results have been demonstrated via deep learning (DL). However, most of the DL approaches are solely data-driven methods and lack the related physics prior, which results in a limited generalization capability. In this paper, through the effective combination of the speckle-correlation theory and the DL method, we demonstrate a physics-informed learning method in scalable imaging through an unknown thin scattering media, which can achieve high reconstruction fidelity for the sparse objects by training with only one diffuser. The method can solve the inverse problem with more general applicability, which promotes that the objects with different complexity and sparsity can be reconstructed accurately through unknown scattering media, even if the diffusers have different statistical properties. This approach can also extend the field of view (FOV) of traditional speckle-correlation methods. This method gives impetus to the development of scattering imaging in practical scenes and provides an enlightening reference for using DL methods to solve optical problems. © 2021 Chinese Laser Press

https://doi.org/10.1364/PRJ.416551

1. INTRODUCTION

Object information is seriously degraded after being modulated by complex media [1,2]. The scattering of light in the diffusion is an established problem considered to be a common phenomenon in our daily life (i.e., seeing through dense fog to obtain the license plate and driver's facial information is crucial for a traffic monitor). Imaging with randomly scattered light is a challenging problem with an urgent requirement in different fields (e.g., astronomical observations through the turbulent atmosphere and biological analysis through active tissue) [3-7]. Conventional imaging methods based on geometric optics cannot work with the disordered light beam under scattering. Benefitting from the great progress of optoelectronic devices and computational techniques, many new imaging methods have been proposed for imaging through scattering media. The typical imaging techniques include the wavefront-shaping methods [8-11], reconstruction using the transmission matrix [12,13], single-pixel imaging methods [14–16], or techniques based on the point spread function (PSF) [17-19]. The methods listed here have made great progress in object reconstruction with invasive prior, and the scattering scenes are relatively stable. Speckle correlation based on the optical memory effect (OME) is an extraordinary method for noninvasive imaging through opaque layers [20] with only one frame of speckle pattern [21,22]. Object recovery based on speckle correlation methods uses phase retrieval algorithms such as hybrid input-output (HIO) [23], those based on the alternating direction method of multipliers (ADMM) [24], and phase retrieval based on generalized approximate message passing (prGAMP) [25]. The field of view (FOV) of speckle-correlation methods is limited by OME, and the recovery performance is also influenced by the recovery capability of phase retrieval algorithms.

Recently, with the advent of digital technology, big data, and advanced optoelectronic technology, deep learning (DL) has shown great potential in optics and photonics [26,27]. With powerful data mining and mapping capabilities, the datainformed DL methods can extract the key features and build a reliable model in many fields [28]. To date, the DL approach has been successfully applied in digital holography imaging [29-32], Fourier ptychographic imaging [33-36], computational ghost imaging [37,38], superresolution microscopic imaging [39-42], optical tomography imaging [43-45], photon-limited imaging [46,47], three-dimensional (3D) measurements with fringe pattern analysis [48-51], and imaging through scattering media [52-60]. Compared to traditional computational imaging (CI) technology, the learning-heuristic methods can not only solve complex imaging problems, but also can significantly improve the core performance indicators (i.e., spatial resolution, temporal resolution, and sensitivity). The great progress by DL is indicated by the rapidly increasing number of DL-related publications in photonics journal in the last several years [61]. However, the methods using DL are meeting several challenging problems, such as the choice of the DL framework tending to be empirical and the limited generalization capability.

Based on the nonlinear characteristics of deep neural networks (DNNs), DL methods have good performance in highly ill-posed problems, especially in imaging through random media [52–54,57]. IDiffNet is the first proposed method to reconstruct an object through scattering media via a densely connected DNN. The performance with a different type of training dataset and the loss function are systematically discussed [52]. A hybrid neural network is constructed to see through a thick scattering medium and achieves object restoration exceeding the FOV of OME [53]. The speckle patterns of single-mode fiber and multimode fibers are reconstructed and recognized successfully [54]. PDSNet is built to reconstruct complex objects through scattering medium and expands the FOV up to 40 times of the memory effect (ME) [55]. The methods above are mainly focused on a specified diffuser that has a limitation in complex and variable scattering conditions. Therefore, some DL methods have the potential to reconstruct objects through unstable media that mainly use different DNN structures, such as one-to-all with dense blocks, interpretable DL method, generative adversarial network (GAN), or a two-stage framework [57-60]. Li et al. [57] first proposed a DL technique to generalize from four different diffusers to more diffusers with raw speckles, which requires the unknown diffusers to have similar statistical properties and the structure of objects to be simple [57]. Almost all the DL methods for imaging through scattering media use speckle patterns directly, and more information might be further excavated with a traditional physical theory. The efficient physics prior can provide an optimized direction for DNN to find the optimal reconstruction solution in different scattering scenes. After being modulated by different diffusers, the scattering light with photons walk randomly brings about the great statistical difference of speckle patterns even with the same one object. Although it has been proven that the DL method, which focuses on DNN structure design, has the generalization capability to reconstruct hidden objects through unknown diffusers, it is still difficult to obtain an accurate object structure under the condition of fewer training diffusers and has a limitation in reconstructing the complex object [57]. At the same time, the generalized diffusers should have similar statistical characteristics. Therefore, in the absence of effective physical constraints and guidance, DL methods can hardly extract universal information from speckle patterns under highly degenerating conditions. Solely data-driven DL methods will lead to the limited generalization capability that the model is over-relying on training data. Thus, to solve the problems of imaging through multicomplex media, combining the scattering theory with DNN is a more efficient method than designing specific DNN structures.

In this paper, with the physics prior of scattering and the support of DL, a physics-informed learning method is proposed for imaging through unknown diffusers. By pre-processing, the data model based on the physics prior can solve the generalization problems in different scattering scenes, which can reduce

the data dependence of the DL model and robustly improve the feature extraction efficiency. The efficient physics prior can provide an optimized direction for DNN to find the optimal reconstruction solution in different scattering scenes. The DL method based on physics prior can help to learn and extract the statistical invariants from different scattering scenes. Instead of training with captured patterns directly, using the DL framework with speckle correlation prior to imaging through different diffusers is technologically reasonable. Employing the physics-informed learning method, scalable imaging through unknown diffusers can be achieved with high reconstruction quality. The scattering degradation of the sparse objects can even be modeled with one ground glass, and imaging through unknown ground glasses even with different statistical characteristics can be achieved. More complex objects (e.g., human faces) can be reconstructed accurately by slightly increasing the number of training diffusers. Meanwhile, it is hard to restore the objects efficiently exceeding the FOV of OME by the traditional speckle-correlation method. Based on the powerful capability in data mining and processing of DNN, the physics-informed learning method can also break through the FOV limitation for scalable imaging. Finally, we demonstrate the physics-informed learning scheme with an experimental dataset and present the quantitative evaluation results with multiple indicators. The results with the statistical average indicator show the accuracy and robustness of our scheme, and reflect the great potential of combining physical knowledge and DL.

2. METHODS

A. Physical Basic

The proposed model must be established to have general applicability for scalable imaging through unknown diffusers, and it is also one of the indispensable conditions to apply this method to practical complex scenes. The wave propagating through an inhomogeneous medium with multiple scattering will generate a fluctuating intensity pattern, and the universal physical law exists in different transmitted modes. The speckle correlation and memory effect in optical wave transmission through disordered media are proposed to observe and analyze the shift-invariant characteristic of speckle patterns [62,63]. The speckle patterns of scattered light through diffusive media are invariant to small tilts or shifts in the incident wavefront of light, and the outgoing light field still retains the information carried by the incoming beam within the range of ME [64]. Therefore, within the scope of ME, the scattering system can be considered as an imaging system with shift-invariant point spread function. The speckle pattern captured by the camera is given by the convolution of the object intensity pattern O(x) with the PSF (S), which can be calculated by

$$I = O * S, \tag{1}$$

where the symbol * denotes the convolution operator. Using the convolution theorem, the autocorrelation of camera pattern intensity can be defined as

$$I \star I = (O \star S) \star (O \star S) = (O \star O) \star (S \star S),$$
 (2)

where the \star is the correlation operator and $S \star S$ is a sharply peaked function representing the autocorrelation of broadband

noise. The autocorrelation of the speckle pattern is approximately equal to the autocorrelation of the object hidden behind the scattering media, and the speckle autocorrelation has an additional constant background term C [21]. Thus, Eq. (2) can be further simplified as

$$I \star I = (O \star O) + C. \tag{3}$$

When the object size exceeds the range of OME, the object can be divided into multiple objects O_i within the OME scope and *n* represents the object distributed in *n* different OME ranges (see Appendix A for details). Thus, the autocorrelation distribution of the speckle pattern exceeding OME can be defined as

$$I \star I = \sum_{i=1}^{n} (O_i \star O_i) + C'.$$
 (4)

To clarify the universal connection among different scattering scenes, speckle patterns are captured through different diffusers with the same object. In total, we take speckle patterns using nine different ground glasses in the experiment, including six 220 grit diffusers (D1, D2, D3, D4, D5, and D6), one 120grit diffuser (D7), one 600 grit diffuser (D8) produced by Thorlabs, and one 220 grit diffuser (D9) produced by Edmund. Among these diffusers, D4 to D9 are selected as testing diffusers. As shown in the first row in Fig. 1(a), even when the diffusers have different statistical characteristics and are made by different manufacturers, the autocorrelation of the object has a high degree of similarity with the autocorrelation of the speckle pattern within or exceeding the range of OME, and the difference is reflected in the different background terms. On the other hand, the correlation of different speckles is irregular by calculating the cross-correlation with D1. The autocorrelation of the speckle pattern exceeding OME is also similar among different diffusers to some extent. However, the cross-correlation with different diffusers almost has no



Fig. 1. Speckle statistical characteristics analysis of the same object corresponding to different testing diffusers. (a) First row and second row are the speckle autocorrelation of the object within or exceeding the OME range, the third row is the cross-correlation with D1, respectively. (b)–(d) Intensity values of the white dash lines in the first, second, and third rows of (a), respectively. The color bar represents the normalized intensity. Scale bars: 875.52 μ m.

similarities, even with similar statistical properties (e.g., D4, D5, and D6).

With the speckle-correlation prior, the statistical invariants of the object through different scattering media can be effectively extracted, which informs the DNN to obtain useful information and reconstruct the object in different scattering scenes. Imaging through different scattering media with speckle-correlation prior can be used as a reference and a heuristic approach to design the DL methods in different optical problems.

B. Framework of Physics-Informed Learning

To solve the optical problems by DL methods, it is essential to make full use of optical physics prior. As shown in Fig. 2, the physics-informed learning framework consists of a speckle-correlation pre-processing step and a neural network post-processing step. The pre-processing step is mainly to obtain speckle autocorrelation R(x, y), is calculated by an inverse 2D Fourier transform of its energy spectrum, and the mathematical operations can be expressed as

$$R(x,y) = I(x,y) \star I(x,y) = FFT^{-1}\{|FFT\{I(x,y)\}|^2\}.$$
 (5)

After the speckle-correlation pre-processing step, the captured speckle pattern is adjusted and refactored, and the next step is post-processing by DNN to reconstruct the hidden object. By adding the speckle-correlation theory, the imaging model can make full use of the advantages of the neural network. The DL model is a simple convolutional neural network (CNN) of the U-Net type [65]. Comparing a specially designed DNN structure, the physics-informed learning method can achieve better imaging results with a simple U-Net without any other tricks.

In our experiments, multiple objects datasets with different levels of complexity are used to reconstruct through different diffusers, such as the modified National Institute of Standards and Technology (MINIST) dataset [66] and FEI face dataset [67]. An equilibrium constraint loss function is important for the training process, and we design a combination loss that includes negative Pearson correlation coefficient (NPCC) loss and mean square error (MSE). The Pearson correlation coefficient is an index used to evaluate the similarity between two variables, and the calculated value is distributed from -1 to 1. A negative value represents a negative correlation,



Fig. 2. Schematic of the physics-informed learning method for scalable scattering imaging.

a positive value represents a positive correlation, and 0 represents an irrelevant correlation. Since the optimization direction of the deep learning is optimized in the direction of loss value reduction, to obtain a positive reconstruction result, the NPCC is used for training [52]. The loss functions can be formulated as

$$Loss = Loss_{NPCC} + Loss_{MSE},$$
 (6)

Loss_{NPCC}

$$=\frac{-1\times\sum_{x=1}^{w}\sum_{y=1}^{h}[i(x,y)]-\hat{i}][I(x,y)]-\hat{I}]}{\sqrt{\sum_{x=1}^{w}\sum_{y=1}^{h}[i(x,y)]-\hat{i}]^{2}}\sqrt{\sum_{x=1}^{w}\sum_{y=1}^{h}[I(x,y)]-\hat{I}]^{2}}},$$
 (7)

$$Loss_{MSE} = Loss_I = \sum_{x=1}^{w} \sum_{y=1}^{h} |\tilde{i}(x,y) - I(x,y)|^2,$$
 (8)

where \hat{I} and \hat{i} are the mean value of the object ground truth Iand the DNN output i, respectively, and \tilde{i} is a normalized image of i. The combination loss function has a good capability to reconstruct objects with different complexity and sparsity through different scattering media. To train the DNN, an Adam optimizer is selected as the strategy to update the weights in the training process. The DNN is performed on PyTorch 1.4.0 with a Titan RTX graphics unit and i9-9940X CPU under Ubuntu 16.04.

3. EXPERIMENTS AND RESULTS

A. Experimental Arrangement and Data Acquisition

The optical configuration is schematically illustrated in Fig. 3. A mounted LED (M625L4, Thorlabs, Newton, NJ, USA) and a filter (FL632.8-1, central wavelength: 632.8 ± 0.2 mm, Thorlabs) are assembled as the light source, which can be employed as an approximate incoherent light source. A digital micromirror device (DMD) (pixel count: 1024×768 , pixel size:



Fig. 3. Experimental setup for the scalable imaging. Different diffusers are employed to obtain speckle patterns with different scattering scenes. The OME range of this system is also measured by calculating the cross-correlation coefficient [21]. See Appendix B for details.

13.68 μ m) is used to code and display the 8-bit objects. An industrial camera (acA1920-155um, Basler AG, Ahrensburg, Germany) is employed to obtain the patterns, which have lower data depth and relatively poor photo quality, instead of a scientific camera [8–13,17–19,52–54]. Thus, this method is more suitable for practical application. The ground glass is placed between the CMOS and DMD. The distance between the object and the diffuser is 30 cm, and the distance between the diffuser and the CMOS is 8 cm. The diameter of the iris behind the diffuser in the configuration is 8 mm, and the diameter of the iris combined with CL is 11 mm.

To obtain the speckle patterns in different scattering scenes, nine different ground glasses are used as the diffusers in the experiments, including six 220 grit diffusers (D1-D6), one 120 grit diffuser (D7), one 600 grit diffuser (D8) produced by Thorlabs, and one 220 grit diffuser (D9) produced by Edmund, like the configuration in Section 2.B. We choose one ground glass (D1) or the first three pieces of ground glasses (D1, D2, and D3) as the training diffusers, and the remaining ground glasses as the test diffusers. The objects are mainly selected from the MINIST database and FEI face databases. The character objects are selected randomly from the MINIST dataset to form the different complexity of single-character and double-character objects. For collecting the experimental data, 600 single characters, 600 double characters, and 400 human faces are used as objects hidden behind each diffuser. The first 500 characters are used as the seen objects and the remaining characters are used as the unseen objects. Similarly, the first 360 human faces are used as seen objects and the remaining faces are used as unseen objects. The autocorrelation pre-processing for speckle patterns is the first step for our method. As for the processing of the speckle patterns, we take the 512×512 camera pixels from the center pattern to calculate the autocorrelation and crop the center to 256×256 pixels autocorrelation pattern as the input autocorrelation image. All the objects, speckle patterns, and autocorrelation images are in grayscale in this experiment.

According to different training data and testing data, different groups are used to characterize the generalization capability of the physics-informed DL method, respectively. All of the testing data are captured from unknown diffusers for emphasizing the generalization. The data can be roughly divided into four groups.

Group 1: The objects are the single characters within the OME. The training data can be divided into two types: training with one diffuser (D1) or three diffusers (D1–D3) with seen objects (the first 500 characters). The testing data can also be divided into two types: the seen objects and the unseen objects (the last 100 characters) with testing diffusers (D4–D9).

Group 2: The objects are the double characters within the OME. The data arrangement is similar to Group 1, except for the complexity of objects.

Group 3: The objects are the human faces within the OME. The training data can also be divided into two types: training with one diffuser (D1) or three diffusers (D1–D3) with seen objects (the first 360 faces). The testing data can also be divided into two types: the seen objects and the unseen objects (the last 40 faces) with testing diffusers (D4–D9).

B214 Vol. 9, No. 5 / May 2021 / Photonics Research

Group 4: The objects are the single characters extending the FOV to 1.2 times. The data arrangement is also similar to Group 1, except for the size and distribution of objects.

B. Scalable Imaging with Different Diffusers

After collecting and classifying the data into different types, the proposed method is used for training and testing. The objects within OME (i.e., the first three groups), are tested first, and the experimental variables are the numbers of training diffusers and the category of the objects with different complexity and sparsity. To prove the good generalization capability and robustness of the physics-informed learning method, the subjective evaluation with reconstruction and the statistical average of objective evaluation results are provided in this section. Before the quantitative evaluating, the output images of the model have first been normalized. The imaging results shown in this paper are randomly selected from testing data, and the mean absolute error (MAE), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM) are employed to quantitatively evaluate the generalization results and are presented in Table 1 with different groups. The results with Group 1 are presented with multiple examples in Fig. 4. Even with one training diffuser (D1), reliable generalization imaging results with unknown diffusers can be obtained, and the average PSNR is up to 23.41 dB. As shown in the reconstruction results of 0 in Fig. 4, it can be clarified that the seen objects testing results or unseen objects testing results with three training diffusers have higher fidelity than one training diffuser. Therefore, by increasing the quantity of training diffusers (D1-D3), the method can obtain higher accuracy and better generalization capability in unknown scattering scenes, and the average PSNR can reach over 40 dB with the training data of three diffusers.

To further verify the effectiveness of this method, more complex double-character objects and human faces are selected successively in this experiment. In addition to the singlecharacter objects commonly used in traditional scattering scenes, double-character objects and FEI face database are selected in this paper for scheme verification, and the structures are more complex and suitable for actual application scenes. As shown in Fig. 5, double-character objects formed by combining single characters randomly can also be restored accurately through unknown diffusers. In the same way, by increasing



Fig. 4. Testing results for generalization reconstruction of Group 1. Scale bars: $264.24 \ \mu m$.



Fig. 5. Testing results for generalization reconstruction of Group 2. Scale bars: 264.24 $\mu m.$

the quantity of training diffusers, the generalization capability can also be significantly improved. Compared to the character objects, the FEI faces database is more complex and difficult for the learning method. As shown in the second row in Fig. 6, the results of D4 are chosen as examples to show that the DL framework cannot reconstruct the human faces efficiently with

Table 1. Quantitative Evaluation Results of the Objects within OME

Туре	Training Set	Testing Set	MAE	SSIM	PSNR (dB)
Group 1	D1	Seen objects	0.0190	0.8477	23.41
	D1	Unseen objects	0.0288	0.8248	19.86
	D1, D2, D3	Seen objects	0.0033	0.9744	40.56
	D1, D2, D3	Unseen objects	0.0271	0.8088	20.58
Group 2	D1	Seen objects	0.0257	0.8608	20.75
	D1	Unseen objects	0.0548	0.6929	16.56
	D1, D2, D3	Seen objects	0.0072	0.9190	38.60
	D1, D2, D3	Unseen objects	0.0455	0.7678	16.76
Group 3	D1	Seen objects	0.1166	0.4672	15.59
	D1	Unseen objects	0.1234	0.5351	14.98
	D1, D2, D3	Seen objects	0.0323	0.8523	23.82
	D1, D2, D3	Unseen objects	0.0534	0.7410	20.67



Fig. 6. Testing results for generalization reconstruction of Group 3. Scale bars: 264.24 µm.

only one training diffuser. Once the quantity of training diffusers is improved to three diffusers, the generalization results are more accurate and reliable. In addition, one person has two photos with different micro-expressions in the FEI face database. From the seen objects testing results, with three training diffusers, the facial features and details of human faces can be clearly reconstructed. The scalable imaging of human faces has a high reduction degree, and the reconstructed faces with slight micro-expressions can also be identified and accurately distinguished.

When the object's scale exceeds OME, the object information is also contained by the speckle pattern and can be described with the speckle-correlation theory as Eq. (4). If the scale of the object exceeds the FOV of OME, it is hard for the traditional speckle-correlation methods to recover the object via a single speckle pattern. Using the powerful datamining capabilities of CNN, the proposed method can extend the scope of OME, and have generalization capability for the large-scale objects shown in Fig. 7. As the quantity of training diffusers increases, the generalization effect can also be improved accordingly. From the quantitative evaluation results



Fig. 7. Testing results for generalization reconstruction of Group 4. Scale bars: $820.8 \ \mu m$.



Fig. 8. Generalization results for a single-character object with different scales and the scale of FOV is defined as the FOV/OME times. (a), (b) Results with different amounts of training diffusers, which are trained with one diffuser and three diffusers, respectively. (c) Reconstruction results with different scales and corresponding ground truth (GT).

in Table 2, the difficulty of scalable imaging exceeding OME is higher than Group 1, and the conclusion of the generalization capability is similar to the results within OME.

The FOV of scalable imaging beyond OME is affected by different configurations, such as the amount of the training diffusers, the complexity and sparsity of objects, and the camera hardware parameters. The generalization results with different scales are presented in Fig. 8, and the generalization conclusion with FOV is that the imaging effects and indicators are decreasing when the FOV is extending. As shown in Fig. 8(c), when the scale of the FOV extends to 1.8 times, the reconstructed characters are blurry and hard to distinguish. As shown in Fig. 8(b), with the quantity of training diffusers increasing, a similar conclusion can be drawn that the objective indicators of the generalization results are obviously improving.

4. ANALYSIS

A. Comparison to Traditional DL Strategy

To demonstrate the necessity of the physics-informed preprocessing step in imaging through unknown diffusers, the comparison images recovered by the end-to-end DL method without physics prior are presented in Fig. 9. As the aforementioned conclusions from Fig. 1, the speckle characteristics with different diffusers have a big difference. The unreliable imaging results are obtained from speckle patterns directly without

Table 2. Quantitative Evaluation Results of Objects Extending the FOV 1.2 Times

Туре	Training Set	Testing Set	MAE	SSIM	PSNR (dB)
Group 4	D1	Seen objects	0.0359	0.8439	17.53
	D1	Unseen objects	0.0432	0.6968	17.12
	D1, D2, D3	Seen objects	0.0085	0.9431	30.55
	D1, D2, D3	Unseen objects	0.0309	0.8706	18.90



Fig. 9. Comparison results without or with this pre-processing step for imaging through an unknown diffuser. Three ground glasses are selected as the training diffusers and another diffuser for testing.

physics prior, and the reconstruction indicators are 0.2111 in SSIM and 15.54 dB in PSNR. Although there are a few objects that can be distinguished, such as the digits "1" and "3," it is still hard for DNN to learn the speckle-correlation pre-processing step inside hidden layers automatically without an effective physics prior.

B. Performance with Different Number of Speckles

As shown in Fig. 9, the end-to-end DL method without physics prior can hardly work with raw speckle patterns for imaging through unknown diffusers. Thus, using the specklecorrelation prior is an effective step for the generalization of the imaging through unknown scattering media, and traditional speckle theory in speckle-correlation methods is also suitable to enhance the performance of physics-informed learning. Using relatively more speckles is an efficient way to reduce the statistical noise of the autocorrelation through scattering media [21]. The comparison results are shown in Fig. 10, using more speckles can also improve the performance of



Fig. 10. Results with different number of speckles via the physicsinformed learning method. Three ground glasses are selected as the training diffusers and another diffuser for testing.

Research Article

Table 5.	Objective indicators with Different Number of
Speckles	via the Physics-Informed Learning Method

Number of Speckles	MAE	SSIM	PSNR (dB)
256 × 256	0.0054	0.9342	38.38
512 × 512	0.0037	0.9539	44.53



Fig. 11. Generalization results of imaging exceeding OME range with different complexity objects.

the physics-informed learning method, and the objective indicators are also presented in Table 3.

C. Performance in Exceeding FOV

Using the existing optical system, the imaging capability beyond the FOV of OME is also tested up to five times. As shown in Fig. 11, when the size of the digits is 760 × 760 pixels on DMD for a 5× OME range, the hidden objects can also be accurately predicted, and the objective indicators are presented in Table 4. We can obtain reliable results through an unknown scattering medium with three training diffusers, and the recovery difficulty increases with more complex objects. As for more complex objects (e.g., the FEI face dataset), the physicsinformed learning method with a traditional U-Net cannot obtain reliable results. However, we can improve the generalization ability by using the PDSNet-L [55] for the

Table 4. Objective Indicators Corresponding to Fig. 11

Objects	Network	MAE	SSIM	PSNR (dB)
5× OME single characters	U-Net	0.0232	0.8741	21.70
5× OME double characters	U-Net	0.0361	0.8360	18.96
1.2× OME faces	U-Net	0.0673	0.6972	18.40
1.2× OME faces	PDSNet-L	0.0567	0.7050	21.16

neural network post-processing, which has a better reconstruction capability. Furthermore, we can further improve the generalization ability by using a better camera (a scientific CMOS or an electron-multiplying CCD) or gathering more training data.

5. DISCUSSION

According to the experimental results shown in Section 3, we have three key points.

(i) A physics-informed DL framework is proposed for scalable imaging through different scattering scenes in which the diffusers are previously untrained. The objects hidden behind the unknown diffusers are not limited to simple sparse characters, and more complex objects (e.g., human faces) can be reconstructed with high accuracy. The physics-informed learning method can also extend the FOV of conventional specklecorrelation methods.

(ii) The DL framework has a reliable generalization capability in imaging through unknown thin scattering media using only one training diffuser for the sparse object. With the number of training diffusers increasing, the generalization capability of the methods is further improved. The proposed method can still reconstruct the overall structure and local details for human faces, even the slight micro-expressions can be clearly distinguished. However, the DL models are prone to preferentially fit the category of the training dataset, which limits the generalization capability of the physics-informed learning method with unknown category objects.

(iii) Benefitting from the great capability in data-mining and mapping of DNN, reliable generalization results can also be obtained through unknown diffusers with the extended FOV. Meanwhile, the FOV of the physics-informed learning method is also relevant to several factors, such as the number of training diffusers and the complexity of the hidden objects.

6. CONCLUSION

In this paper, a physics-informed learning method is introduced to imaging through diffusers. Specifically, an explicit framework is established to efficiently solve the generalization problems in different scattering scenes by combining the physics theories and DL methods. This is a new approach to solve scalable imaging with deep learning, which can reconstruct complex objects through different scattering media, and provide an expanded FOV for the real imaging scenes. In the future, more complex scenes and objects can be considered, which can be applied to volumetric multiple scattering, such as biological imaging and astronomical imaging.

APPENDIX A: THE FORMULA DERIVATION TO EXCEED THE OME RANGE

When the object size exceeds the range of OME, the object can be divided into multiple objects O_i within OME scope, and the PSFs produced from the different points become uncorrelated mutually. The autocorrelation of PSF can be approximately expressed as

$$\text{PSF}_i \star \text{PSF}_j \approx \begin{cases} \delta_{ij}, & i = j \\ 0, & i \neq j \end{cases}.$$
 (A1)

We assume that the distance between objects is beyond the single OME range. Taking the autocorrelation of the camera image and using the convolution theorem yields [68]

$$I \star I = \left(\sum_{i=1}^{n} O_i * \text{PSF}_i\right) \star \left(\sum_{i=1}^{n} O_i * \text{PSF}_i\right) = O_1$$

$$\star O_1 + C_1 + O_2 \star O_2 + C_2 + O_3$$

$$\star O_3 + C_3 + 2(O_1 \star O_2) * (\text{PSF}_1 \star \text{PSF}_2) + 2(O_2 \star O_3)$$

$$* (\text{PSF}_2 \star \text{PSF}_3) + 2(O_1 \star O_3) * (\text{PSF}_1 \star \text{PSF}_3) + \dots$$

$$= \sum_{i=1}^{n} (O_i \star O_i + C_i) = \sum_{i=1}^{n} (O_i \star O_i) + C'.$$
 (A2)

Thus, the autocorrelation distribution of a speckle pattern exceeding OME can be defined as

$$I \star I = \sum_{i=1}^{n} (O_i \star O_i) + C'.$$
 (A3)

APPENDIX B: OME RANGE CALIBRATION DETAILS

To calibrate the range of the shift-invariant, the distance from the object to the diffuser is changed to 15 cm and the image distance is maintained as 8 cm. A ground glass (DG100X100-220-N-BK7, Thorlabs) is used as the diffuser and placed between the object and the CMOS. A series of speckle patterns are collected while the horizontal displacement of the point object on the object surface is achieved. The cross-correlation coefficient between the speckle patterns and the PSF of the system is calculated. A threshold value of 0.5 is chosen as the crosscorrelation coefficient to determine the range of OME [68,69]. We define δp as the offset pixel number of the image plane, which is 30 pixels, as shown in Fig. 3. The OME range of the system can be calculated by $2 \times p \times \delta p/\beta$ [68], and β is the system magnification and p is the pixel size of the camera, which equals 5.86 μ m. It can be figured that the half-width at half maximum (HWHM) is 30 pixels. Because the distance from the object to the diffuser of the speckle collection system is 30 cm, the HWHM of the speckle collection system is 60 pixels. Then, the full width at half maximum of the speckle collection system is 120 pixels. Thus, the OME range of our speckle collection system is 152×152 pixels on DMD.

Funding. National Natural Science Foundation of China (62031018, 61971227); Jiangsu Provincial Key Research and Development Program (BE2018126); Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX20_0264).

Acknowledgment. The authors thank Qianying Cui, Yingjie Shi, Chenyin Zhou, Kaixuan Bai, and Mengzhang Liu for technical support and experimental discussions.

Disclosures. The authors declare no conflicts of interest.

[†]These authors contributed equally to this paper.

- 1. J. W. Goodman, Speckle Phenomena in Optics: Theory and Applications (Roberts and Company, 2007).
- 2. M. C. Roggemann, B. M. Welsh, and B. R. Hunt, *Imaging Through Turbulence* (CRC Press, 1996).
- 3. R. K. Tyson, Principles of Adaptive Optics (CRC Press, 2015).
- E. J. McCartney, Optics of the Atmosphere: Scattering by Molecules and Particles (Wiley, 1976).
- 5. V. Ntziachristos, "Going deeper than microscopy: the optical imaging frontier in biology," Nat. Methods 7, 603–614 (2010).
- S. Yoon, M. Kim, M. Jang, Y. Choi, W. Choi, S. Kang, and W. Choi, "Deep optical imaging within complex scattering media," Nat. Rev. Phys. 2, 141–158 (2020).
- L. V. Wang and H.-I. Wu, *Biomedical Optics: Principles and Imaging* (Wiley, 2012).
- 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, 283–292 (2012).
- I. M. Vellekoop and A. Mosk, "Focusing coherent light through opaque strongly scattering media," Opt. Lett. 32, 2309–2311 (2007).
- S. Rotter and S. Gigan, "Light fields in complex media: mesoscopic scattering meets wave control," Rev. Mod. Phys. 89, 015005 (2017).
- K. Wang, W. Sun, C. T. Richie, B. K. Harvey, E. Betzig, and N. Ji, "Direct wavefront sensing for high-resolution *in vivo* imaging in scattering tissue," Nat Commun 6, 7276 (2015).
- S. Popoff, G. Lerosey, R. Carminati, M. Fink, A. 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**, 100601 (2010).
- A. Drémeau, A. Liutkus, D. Martina, O. Katz, C. Schülke, F. Krzakala, S. Gigan, and L. Daudet, "Reference-less measurement of the transmission matrix of a highly scattering material using a DMD and phase retrieval techniques," Opt. Express 23, 11898–11911 (2015).
- 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, 16945– 16955 (2014).
- Y.-K. Xu, W.-T. Liu, E.-F. Zhang, Q. Li, H.-Y. Dai, and P.-X. Chen, "Is ghost imaging intrinsically more powerful against scattering?" Opt. Express 23, 32993–33000 (2015).
- Q. Fu, Y. Bai, X. Huang, S. Nan, P. Xie, and X. Fu, "Positive influence of the scattering medium on reflective ghost imaging," Photon. Res. 7, 1468–1472 (2019).
- D. Lu, M. Liao, W. He, Z. Cai, and X. Peng, "Imaging dynamic objects hidden behind scattering medium by retrieving the point spread function," Proc. SPIE **10834**, 1083428 (2018).
- H. He, X. Xie, Y. Liu, H. Liang, and J. Zhou, "Exploiting the point spread function for optical imaging through a scattering medium based on deconvolution method," J. Innov. Opt. Health Sci. 12, 1930005 (2019).
- X. Xu, X. Xie, A. Thendiyammal, H. Zhuang, J. Xie, Y. Liu, J. Zhou, and A. P. Mosk, "Imaging of objects through a thin scattering layer using a spectrally and spatially separated reference," Opt. Express 26, 15073–15083 (2018).
- J. Bertolotti, E. G. Van Putten, C. Blum, A. Lagendijk, W. L. Vos, and A. P. Mosk, "Non-invasive imaging through opaque scattering layers," Nature 491, 232–234 (2012).
- O. Katz, P. Heidmann, M. Fink, and S. Gigan, "Non-invasive singleshot imaging through scattering layers and around corners via speckle correlations," Nat. Photonics 8, 784–790 (2014).
- A. Porat, E. R. Andresen, H. Rigneault, D. Oron, S. Gigan, and O. Katz, "Widefield lensless imaging through a fiber bundle via speckle correlations," Opt. Express 24, 16835–16855 (2016).
- J. R. Fienup, "Phase retrieval algorithms: a comparison," Appl. Opt. 21, 2758–2769 (1982).
- J. Chang and G. Wetzstein, "Single-shot speckle correlation fluorescence microscopy in thick scattering tissue with image reconstruction priors," J. Biophoton. **11**, e201700224 (2018).

- P. Schniter and S. Rangan, "Compressive phase retrieval via generalized approximate message passing," IEEE Trans. Signal Process. 63, 1043–1055 (2014).
- Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature 521, 436–444 (2015).
- I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning* (MIT, 2016), Vol. 1.
- G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," Optica 6, 921–943 (2019).
- Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," Light Sci. Appl. 7, 17141 (2018).
- Y. Wu, Y. Rivenson, Y. Zhang, Z. Wei, H. Günaydin, X. Lin, and A. Ozcan, "Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery," Optica 5, 704–710 (2018).
- Z. Ren, Z. Xu, and E. Y. Lam, "Learning-based nonparametric autofocusing for digital holography," Optica 5, 337–344 (2018).
- Z. Ren, Z. Xu, and E. Y. Lam, "End-to-end deep learning framework for digital holographic reconstruction," Adv. Photon. 1, 016004 (2019).
- T. Nguyen, Y. Xue, Y. Li, L. Tian, and G. Nehmetallah, "Deep learning approach for Fourier ptychography microscopy," Opt. Express 26, 26470–26484 (2018).
- A. Kappeler, S. Ghosh, J. Holloway, O. Cossairt, and A. Katsaggelos, "Ptychnet: CNN based Fourier ptychography," in *IEEE International Conference on Image Processing (ICIP)* (IEEE, 2017), pp. 1712–1716.
- S. Jiang, K. Guo, J. Liao, and G. Zheng, "Solving Fourier ptychographic imaging problems via neural network modeling and TensorFlow," Biomed. Opt. Express 9, 3306–3319 (2018).
- Y. F. Cheng, M. Strachan, Z. Weiss, M. Deb, D. Carone, and V. Ganapati, "Illumination pattern design with deep learning for single-shot Fourier ptychographic microscopy," Opt. Express 27, 644–656 (2019).
- M. Lyu, W. Wang, H. Wang, H. Wang, G. Li, N. Chen, and G. Situ, "Deep-learning-based ghost imaging," Sci. Rep. 7, 17865 (2017).
- Y. He, G. Wang, G. Dong, S. Zhu, H. Chen, A. Zhang, and Z. Xu, "Ghost imaging based on deep learning," Sci. Rep. 8, 6469 (2018).
- H. Wang, Y. Rivenson, Y. Jin, Z. Wei, R. Gao, H. Günaydın, L. A. Bentolila, C. Kural, and A. Ozcan, "Deep learning enables crossmodality super-resolution in fluorescence microscopy," Nat. Methods 16, 103–110 (2019).
- E. Nehme, L. E. Weiss, T. Michaeli, and Y. Shechtman, "Deep-storm: super-resolution single-molecule microscopy by deep learning," Optica 5, 458–464 (2018).
- W. Ouyang, A. Aristov, M. Lelek, X. Hao, and C. Zimmer, "Deep learning massively accelerates super-resolution localization microscopy," Nat. Biotechnol. 36, 460–468 (2018).
- C. Ling, C. Zhang, M. Wang, F. Meng, L. Du, and X. Yuan, "Fast structured illumination microscopy via deep learning," Photon. Res. 8, 1350–1359 (2020).
- L. Fang, D. Cunefare, C. Wang, R. H. Guymer, S. Li, and S. Farsiu, "Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search," Biomed. Opt. Express 8, 2732–2744 (2017).
- L. Waller and L. Tian, "Computational imaging: Machine learning for 3D microscopy," Nature 523, 416–417 (2015).
- T. C. Nguyen, V. Bui, and G. Nehmetallah, "Computational optical tomography using 3-D deep convolutional neural networks," Opt. Eng. 57, 041406 (2018).
- A. Goy, K. Arthur, S. Li, and G. Barbastathis, "Low photon count phase retrieval using deep learning," Phys. Rev. Lett. 121, 243902 (2018).
- C. Chen, Q. Chen, J. Xu, and V. Koltun, "Learning to see in the dark," in *IEEE Conference on Computer Vision and Pattern Recognition* (2018), pp. 3291–3300.
- S. Feng, Q. Chen, G. Gu, T. Tao, L. Zhang, Y. Hu, W. Yin, and C. Zuo, "Fringe pattern analysis using deep learning," Adv. Photon. 1, 025001 (2019).
- K. Wang, Y. Li, Q. Kemao, J. Di, and J. Zhao, "One-step robust deep learning phase unwrapping," Opt. Express 27, 15100–15115 (2019).

- H. Yu, X. Chen, Z. Zhang, C. Zuo, Y. Zhang, D. Zheng, and J. Han, "Dynamic 3-D measurement based on fringe-to-fringe transformation using deep learning," Opt. Express 28, 9405–9418 (2020).
- H. Yu, D. Zheng, J. Fu, Y. Zhang, C. Zuo, and J. Han, "Deep learningbased fringe modulation-enhancing method for accurate fringe projection profilometry," Opt. Express 28, 21692–21703 (2020).
- S. Li, M. Deng, J. Lee, A. Sinha, and G. Barbastathis, "Imaging through glass diffusers using densely connected convolutional networks," Optica 5, 803–813 (2018).
- M. Lyu, H. Wang, G. Li, S. Zheng, and G. Situ, "Learning-based lensless imaging through optically thick scattering media," Adv. Photon. 1, 036002 (2019).
- N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, "Learning to see through multimode fibers," Optica 5, 960–966 (2018).
- E. Guo, S. Zhu, Y. Sun, L. Bai, C. Zuo, and J. Han, "Learning-based method to reconstruct complex targets through scattering medium beyond the memory effect," Opt. Express 28, 2433–2446 (2020).
- E. Guo, Y. Sun, S. Zhu, D. Zheng, C. Zuo, L. Bai, and J. Han, "Singleshot color object reconstruction through scattering medium based on neural network," Opt. Lasers Eng. **136**, 106310 (2020).
- Y. Li, Y. Xue, and L. Tian, "Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media," Optica 5, 1181–1190 (2018).
- Y. Sun, J. Shi, L. Sun, J. Fan, and G. Zeng, "Image reconstruction through dynamic scattering media based on deep learning," Opt. Express 27, 16032–16046 (2019).
- M. Liao, S. Zheng, D. Lu, G. Situ, and X. Peng, "Real-time imaging through moving scattering layers via a two-step deep learning strategy," Proc. SPIE **11351**, 113510V (2020).

- Y. Li, S. Cheng, Y. Xue, and L. Tian, "Displacement-agnostic coherent imaging through scatter with an interpretable deep neural network," Opt. Express 29, 2244–2257 (2020).
- K. Goda, B. Jalali, C. Lei, G. Situ, and P. Westbrook, "Al boosts photonics and vice versa," APL Photon. 5, 070401 (2020).
- 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, 834–837 (1988).
- I. Freund, M. Rosenbluh, and S. Feng, "Memory effects in propagation of optical waves through disordered media," Phys. Rev. Lett. 61, 2328–2331 (1988).
- H. Liu, Z. Liu, M. Chen, S. Han, and L. V. Wang, "Physical picture of the optical memory effect," Photon. Res. 7, 1323–1330 (2019).
- O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, 2015), pp. 234–241.
- Y. LeCun, C. Cortes, and C. J. C. Burges, "The MNIST database of handwritten digits," http://yann.lecun.com/exdb/mnist/.
- C. E. Thomaz, "FEI face database," https://fei.edu.br/~cet/ facedatabase.html.
- C. Guo, J. Liu, W. Li, T. Wu, L. Zhu, J. Wang, G. Wang, and X. Shao, "Imaging through scattering layers exceeding memory effect range by exploiting prior information," Opt. Commun. 434, 203–208 (2019).
- D. Tang, S. K. Sahoo, V. Tran, and C. Dang, "Single-shot large field of view imaging with scattering media by spatial demultiplexing," Appl. Opt. 57, 7533–7538 (2018).