

PHOTONICS Research

Physical origin and boundary of scalable imaging through scattering media: a deep learning-based exploration

XUYU ZHANG,^{1,2,†} SHENGFU CHENG,^{3,4,†} JINGJING GAO,^{2,5} YU GAN,^{2,5} CHUNYUAN SONG,^{2,5}
DAWEI ZHANG,^{1,8} SONGLIN ZHUANG,¹ SHENSHENG HAN,^{2,5,6} PUXIANG LAI,^{3,4,7,9} AND HONGLIN LIU^{2,4,5,*}

¹School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

²Key Laboratory for Quantum Optics, Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai 201800, China

³Department of Biomedical Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China

⁴Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen 518000, China

⁵Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

⁶Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou 310024, China

⁷Photonics Research Institute, The Hong Kong Polytechnic University, Hong Kong SAR, China

⁸e-mail: dwzhang@usst.edu.cn

⁹e-mail: puxiang.lai@polyu.edu.hk

[†]These authors contributed equally to this work.

*Corresponding author: hlliu4@hotmail.com

Received 21 March 2023; revised 12 April 2023; accepted 12 April 2023; posted 14 April 2023 (Doc. ID 490125); published 26 May 2023

Imaging through scattering media is valuable for many areas, such as biomedicine and communication. Recent progress enabled by deep learning (DL) has shown superiority especially in the model generalization. However, there is a lack of research to physically reveal the origin or define the boundary for such model scalability, which is important for utilizing DL approaches for scalable imaging despite scattering with high confidence. In this paper, we find the amount of the ballistic light component in the output field is the prerequisite for endowing a DL model with generalization capability by using a “one-to-all” training strategy, which offers a physical meaning invariance among the multisource data. The findings are supported by both experimental and simulated tests in which the roles of scattered and ballistic components are revealed in contributing to the origin and physical boundary of the model scalability. Experimentally, the generalization performance of the network is enhanced by increasing the portion of ballistic photons in detection. The mechanism understanding and practical guidance by our research are beneficial for developing DL methods for descattering with high adaptivity. © 2023 Chinese Laser Press

<https://doi.org/10.1364/PRJ.490125>

1. INTRODUCTION

Light scattering within and through complex media poses great challenges for many hotspot applications, including deep tissue imaging, antiscattered data transmission, etc. [1–3]. For example, biological tissues are usually optically turbid, which causes light to diffuse rapidly and prevents high-resolution focusing inside deep tissue. To ensure the imaging resolution, most optical microscopes [4–6] select ballistic photons for imaging with an imaging depth limited within an optical diffusion limit (~1 mm) [7]. Strong optical scattering also prevents explicit data communication through complex media where the input light is scrambled into a seeming random speckle pattern. Luckily, the process is still deterministic, which allows for the recovery of objects hidden behind scattering media.

Various methods for imaging through scattering media have been developed over the past two decades, such as object reconstruction via the transmission matrix (TM) [8,9], speckle correlation imaging [10–12], and single-pixel imaging [13,14]. These methods can retrieve the object information noninvasively, yet encounter limitations either in field of view (FOV) or reconstruction speed. Especially, they are all sensitive to the TM of scattering medium and any change may lead to model errors. Recently, deep learning (DL) approaches have been introduced to invert scattering [15,16] and reconstruct an object through complex media [17–21], showing superior recovery quality and extended FOV than the range of optical memory effect [22]. Initially, the DL-related studies are only applicable to a specific diffuser and cannot adapt to varying scattering

conditions. Later, efforts have been taken in overcoming the speckle decorrelation and achieving highly scalable imaging through scattering media [23–26], through optimizing the network model or training strategy. Those include adopting a “one-to-all” training strategy to enable a network learning the statistical information of multiple diffusers [23], integrating the prior knowledge of speckle correlation theory for physics-informed learning [25], or proposing a dynamic synthesis network (DSN) with robust 3D descattering ability [26], etc. However, the origin of such model generalization is unclear, and it sees limitations for thick scattering media or dynamic scattering conditions. Currently, there is a lack of research from the perspective of physics to reveal the origin and boundary of the scalability of a DL model, which is important in applying DL to scalable imaging through scattering media with high confidence.

In this paper, we investigate how the scattering property of a medium can influence the adaptivity of a reconstruction model. Specifically, we find the amount of (quasi-)ballistic photons in the output light field, which directly reflects the medium’s scattering property, is closely related to the general applicability of the model. Our findings are verified by both experimental and simulated results. In experiment, utilizing a homemade diffuser of relatively weak scattering, much improved adaptivity of a reconstruction network is obtained when trained with data sampled from different regions of the diffuser. To separately study the influences of model training strategy and the scattering property of a diffuser, simulations are performed in which different weights of ballistic component are tested thanks to the adjustable phase distribution of a simulated diffuser. It is revealed that the ballistic light plays a key role in the applicability of a model to unseen diffuser (region): it is lost when there is no ballistic component whatever training strategy is used; it is enhanced proportionally with increasingly larger weight of ballistic light even if the network only saw one specific diffuser

(region) before. The physical origin and boundary for the general applicability of a DL model are further clarified. In addition, our mechanism findings provide guidance for enhancing the generalization performance of DL to scalable descattered imaging.

2. METHODS

A. Experimental Implementation

The experimental setup is illustrated in Fig. 1. A beam from a 532 nm solid-state laser (MGL-III-532–200 mW, Changchun New Industries Optoelectronics Tech.) is first expanded before being collimated onto a digital micromirror device (DMD, V-7001 VIS, ViALUX). The modulated light then illuminates a homemade 220 grit ground glass diffuser. An iris with a diameter of 5 mm is placed right after the diffuser, which creates a tunable window to control the region of imaging through the diffuser. The transmitted light travels a distance before being collected by an on-axis digital camera (DCU224M, Thorlabs). The distances from the beam expander to the DMD, from the DMD to the diffuser, and from the diffuser to the camera are $z_1 = 15$ cm, $z_2 = 16$ cm, and $z_3 = 10$ cm, respectively.

According to the Van Cittert–Zernike theorem [27], the spatial coherence length (SCL) in our diffraction imaging system, is described by $l_c = \frac{\lambda z_3}{D} = 10.64 \mu\text{m}$, where λ is the optical wavelength and D is the aperture size of the iris. SCL approximates the average size of speckle on the camera plane, which was about two to three times of the pitch of the camera pixel. To accelerate data processing, the central area of the originally acquired speckle pattern was cropped into 512×512 and then down sampled into a 256×256 array. To quantify the isoplanatic range, i.e., the average grain size of the diffuser, a point source generated by 3×3 binning pixels on the DMD was used to illuminate the diffuser. The diffuser was shifted

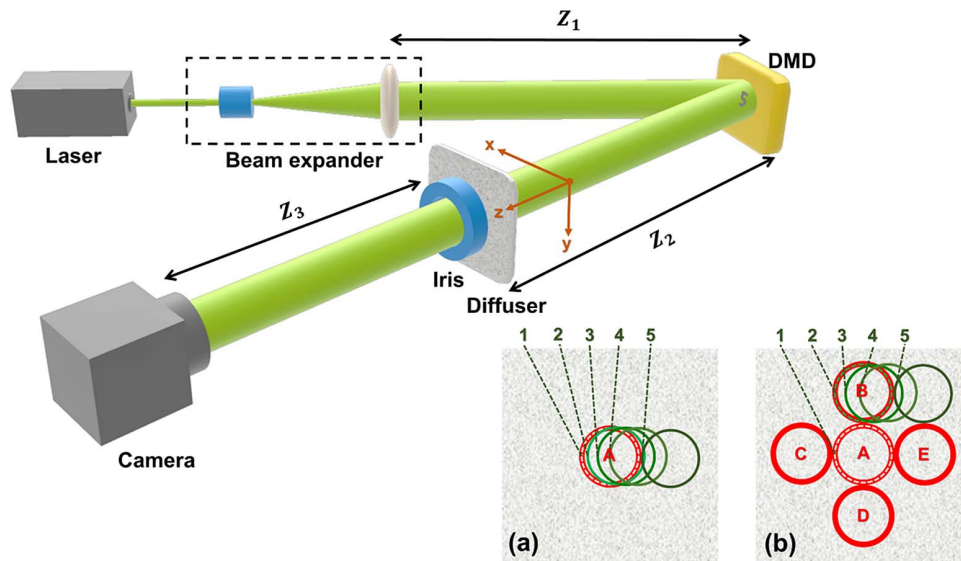


Fig. 1. Schematic of the experimental setup of imaging through a diffuser with the coordinate system labeled. The insets (a) and (b) show the settings of imaging regions to acquire the training and test data in Tests I and II, respectively. In Test I, the training data were obtained from region A (red circle) only with the test data from regions 1–5 (green circles). In Test II, the training data were obtained from regions A–E (red circles) with the test data from regions 1–5 (green circles).

horizontally with the speckle patterns recorded at each displacement (Δx) accordingly. The cross-correlation coefficient (CCC) between the speckle pattern recorded at each position and the one at the origin (i.e., $\Delta x = 0$) was calculated. The full width at half maximum (FWHM) of the fitted CCC curve was used to characterize the isoplanatic range, which reflects the spatially variant scattering property of the imaging system. The speckle pattern recorded with a displacement of diffuser larger than the isoplanatic range is regarded as unrelated to the one recorded at the origin.

In our experiment, handwritten digits from the Modified National Institute of Standards and Technology (MNIST) dataset [28] were used as the objects displayed on the DMD. A U-Net model [29] was used for object reconstruction from speckle data. Two tests were performed to validate the effect of adapting a one-to-all network training strategy [23]. The imaging regions for the acquisition of the training and test data in Tests I and II are indicated in the insets Figs. 1(a) and 1(b), respectively. In Test I, 20,000 pairs of training data were acquired only at region A. Data from five different regions were used for the network test in which region 1 overlapped with region A and regions 2–5 were 10, 40, 100, and 5000 μm , respectively, away from region 1 along the x axis. In Test II, 20,000 pairs of data acquired from regions A–E with 4000 pairs at each were used for network training. To test the network, five different sampling regions were also used with region 1 overlapping with region A, region 2 overlapping with region B, and regions 3–5 being 40, 100, and 5000 μm , respectively, away from region 2. For both Tests I and II, two groups of untrained MNIST digit images (each has 100 images) were selected with their corresponding speckle patterns for the network test.

B. Simulations

1. Diffuser Model

A random phase plate was used to model the diffuser according to the theory of light scattering from rough surfaces in which a Gaussian height distributed surface with a Gaussian autocorrelation function is assumed [30]. The original simulated phase mask has an array of 3000×3000 with a pitch size of 5 μm , whereas only a segment of the array was selected as the effective zone each time. The key parameters of the simulated phase mask include the SCL and the standard deviation of height. To model a 220 grit diffuser, the typical values for the two above parameters were 36 μm and 1.6 μm , respectively.

In simulation, the weight of ballistic light (η) in the transmitted light field was calculated based on the spectrum of the simulated phase mask. Using a coefficient to control the phase distribution range of the mask, the ratios of ballistic and scattered light components were controllable. Specifically, we first calculate the power of the ballistic light, which will correspond to the center zero-frequency spectrum (W_0). However, there are also a few zero-frequency components of the scattered light, which are estimated by calculating the average of the adjacent power spectra around the center spectrum (W_{0_adj}). The accurate power of the ballistic light (W_b) is obtained by subtracting W_{0_adj} from W_0 , and η is calculated as the ratio of W_b to the total power ($\sum W_{spe}$),

$$\eta = \frac{W_b}{\sum W_{spe}} = \frac{W_0 - W_{0_adj}}{\sum W_{spe}}. \quad (1)$$

The capability of controlling the weight of the ballistic component of an output light field provides much convenience in allowing us to study the ballistic contribution to the adaptivity of a reconstruction model.

2. Output Light Field

The light field \mathbf{E} detected behind a scattering medium can be regarded as a weighted superposition of the ballistic and scattered light after free-space diffraction propagation,

$$\mathbf{E} = \alpha \cdot \mathcal{D}(\mathbf{E}_0) + \beta \cdot \mathcal{D}(\mathbf{T}\mathbf{E}_0). \quad (2)$$

Here, \mathbf{E}_0 is the input light field, \mathbf{T} represents the TM of the medium, $\mathcal{D}(\cdot)$ is the diffraction operator, and α , β are the weighting coefficients of ballistic and scattered light fields, respectively. The intensity pattern captured on the imaging plane is as follows:

$$I = |\mathbf{E}|^2 = \underbrace{\alpha^2 |\mathcal{D}(\mathbf{E}_0)|^2}_{I_b} + \underbrace{\beta^2 |\mathcal{D}(\mathbf{T}\mathbf{E}_0)|^2 + \alpha\beta \cdot 2 \text{Re}\{\mathcal{D}(\mathbf{E}_0)\mathcal{D}^*(\mathbf{T}\mathbf{E}_0)\}}_{I_s}, \quad (3)$$

where $|\mathcal{D}(\mathbf{E}_0)|^2$ is the diffracted input pattern, $|\mathcal{D}(\mathbf{T}\mathbf{E}_0)|^2$ is the scrambled speckle pattern, and $2\text{Re}\{\mathcal{D}(\mathbf{E}_0)\mathcal{D}^*(\mathbf{T}\mathbf{E}_0)\}$ is the cross term showing speckle appearance. Usually, the portion of ballistic light in the output field is very weak compared to that of the scattered light. Suppose the scattering mean free path of the medium is l_s , the transport mean free path is l_t , the medium thickness is d ($d \gg l_t > l_s$), the ballistic intensity is $I_b \sim \exp(-d/l_s)$, and the scattered intensity is $I_s \sim l_t/d$. That said, the ballistic light carrying the object information is independent of varying scattering conditions, which may relate to the adaptivity of a reconstruction model.

3. Simulation Settings

For a reconstruction model, it would be easy to extract object information purely from the ballistic component, and such an ability is also scalable to medium perturbation or generalizes to other unknown diffuser. However, with a totally diffused output light field, it would be hard for a model to be scalable if trained with data from a specific diffuser (region). The reason is the model can only learn the specific statistical information of diffuser, which is not generalizable. Therefore, it is natural to hypothesize that the portion of ballistic light in the output field impacts the model scalability. Thanks to the tunable ratio of ballistic component in simulation, we studied its influence on the model adaptivity separately with two tests conducted.

Simulation I. Under the condition of no ballistic component ($\eta = 0$), two comparative tests similar to the experimental settings were involved. For Test I, the U-Net was trained with 20,000 input–output pairs captured at one diffuser region only and tested on the data obtained from regions 1–9, which were horizontally away from the training region by 0, 10, 35, 40, 100, 1000, 2500, 3500, and 5000 μm . For Test II, data from five regions [also as indicated in Fig. 1(b)] with 4000 pairs at each were used for network training. Again, the trained model was tested on the data acquired from regions 1–9.

Simulation II. Although only allowed to see one diffuser region, the model was trained under varying weights of ballistic light ($\eta = 0.1, 0.3, 0.5, 0.7, 0.9, 1$). For each case, 20,000 data from 1 region were collected to train the U-Net. Additionally, data from six regions (with horizontal shifts of 0, 10, 35, 40, 100, and 5000 μm , respectively) were used for a network test.

For both experiment and simulation, we used the Python programming language and Keras/TensorFlow 2.0 framework for the construction of a U-Net model, which was running on the environment of GPU (NVIDIA RTX 3060 laptop edition). The size of speckle data was 256×256 . The total number of training epochs was 50, and the learning rate at the beginning was set as 2×10^{-4} . After five epochs, if the loss value did not decrease, the learning rate would be adjusted to one-tenth of the previous one until the learning rate was reduced to 2×10^{-6} . When the loss value did not decrease after 10 epochs, the training would be terminated. The averaged training duration of each epoch was 130 s.

3. RESULTS

A. Experiment Results

Figure 2 presents the results in experimental Tests I and II. A seemingly high visual similarity due to a lack of speckle details is found for the testing speckle patterns in Fig. 2(a). This could be attributed to the fact that the speckles were almost not amplified before being captured in our diffraction imaging system. However, the structural similarity index measure between

the testing speckle patterns resulting from different categories of MNIST digits is calculated to be less than 0.05, showing a very low level of correlation. For Test I [Fig. 2(a)], the qualitative reconstruction results for those sampled from regions 1–5 are gradually deteriorating. Region 1 has the best quality as it coincides with the training region, which also shows the success of the U-Net in extracting information about unseen objects. The reconstruction quality sees a decline at region 2 but is still acceptable, whereas, getting much poorer at regions 3 and 4, which are only partially “seen” by the network during training. At region 5, the recovered image can no longer be recognized as the network never sees that region before. By contrast, much improved generalization ability to different regions is found for the U-Net in Test II when adopting a one-to-all training strategy. Objects can be perfectly reconstructed at regions 1 and 2, and can still be visually recognized throughout regions 3–5 without much difference in reconstruction quality among them. Note region 5 has no overlapping with the training regions A–E, which means the network can generalize to the unknown region. The quantitative metrics of Pearson correlation coefficient (PCC) are plotted in Fig. 2(b) with those of Test I declining more rapidly than those of Test II. Especially, the average PCC for Test I drops below 0.3 when at $\Delta x = 5$ mm, whereas, Test II sees a relatively stable level of PCC at around 0.6 when the test region shifts horizontally by 40–5000 μm .

The above comparisons validate that the generalization capability of a reconstruction network can be considerably improved when trained with multisource data [18]. Figure 2(c)

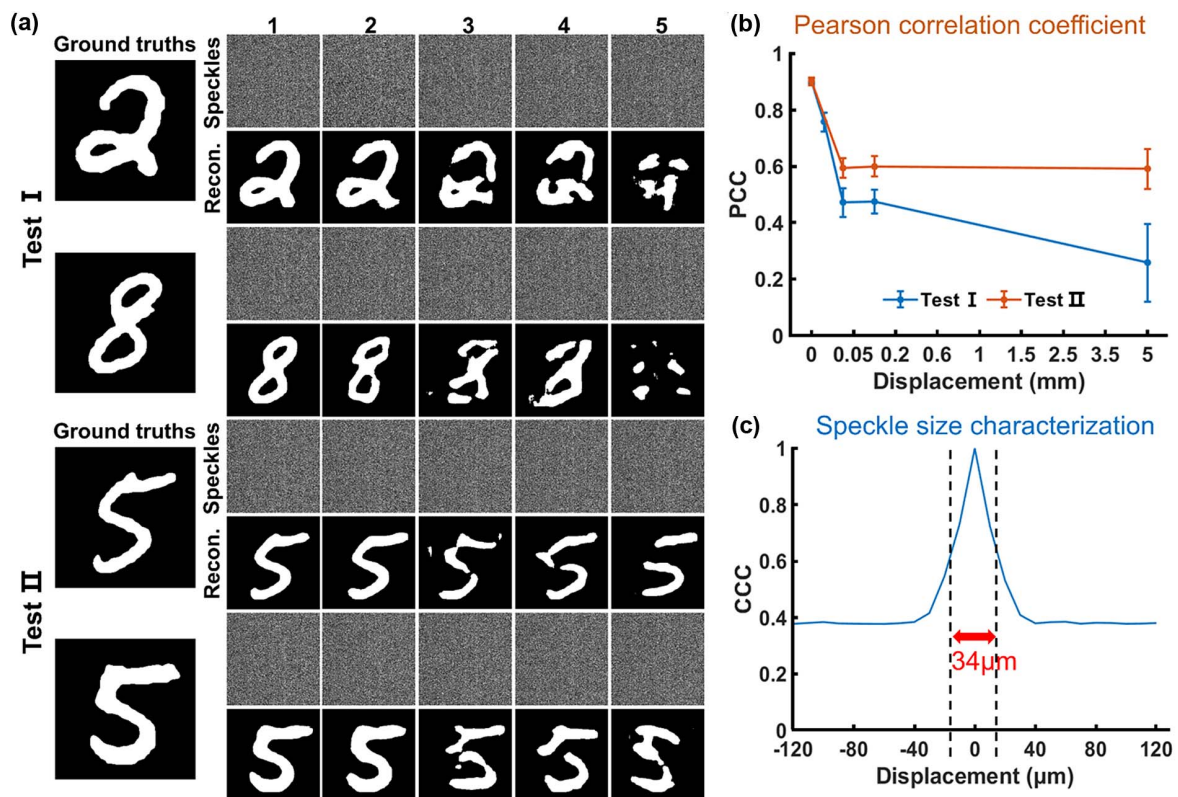


Fig. 2. Experimental results. (a) Image reconstruction through a homemade diffuser in Tests I and II. (b) Curves of the averaged PCC with error bar for 10 reconstructed images at each of the test regions 1–5. Note a nonuniform abscissa is adopted to better reflect the whole trend, given the nonuniform distributed displacements. (c) The CCC curve measured in experiment with a FWHM of $\sim 34 \mu\text{m}$.

gives the fitted CCC curve of the output light field, whose FWHM is used to denote the isoplanatic range, measured to be $\sim 34 \mu\text{m}$. This is consistent with the fast decay of PCC at around $40 \mu\text{m}$ displacement in Fig. 2(b). The fact that the CCC plateaus around 0.4 suggests there is still an important portion of ballistic light. It looks like that, such as in the presence of ballistic light, the adaptivity of reconstruction model can be enhanced using the one-to-all strategy. However, how ballistic component can contribute to the model adaptivity is still unclear. Therefore, we resort to simulated studies.

B. Simulation Results

In Simulation I, we controlled $\eta = 0$ such that the ballistic light through diffuser was depleted with only scattered light. The phase mask for modeling a diffuser with strong scattering is shown in Fig. 3(a). The characterized CCC curve of the simulated diffuser in Fig. 3(b) reveals an isoplanatic range of $\sim 36 \mu\text{m}$. Besides that, a base level of CCC at around 0 confirms the absence of ballistic light in the simulated output field. For both Tests I and II of Simulation I, image reconstruction results from only the odd test regions are presented in Fig. 3(c),

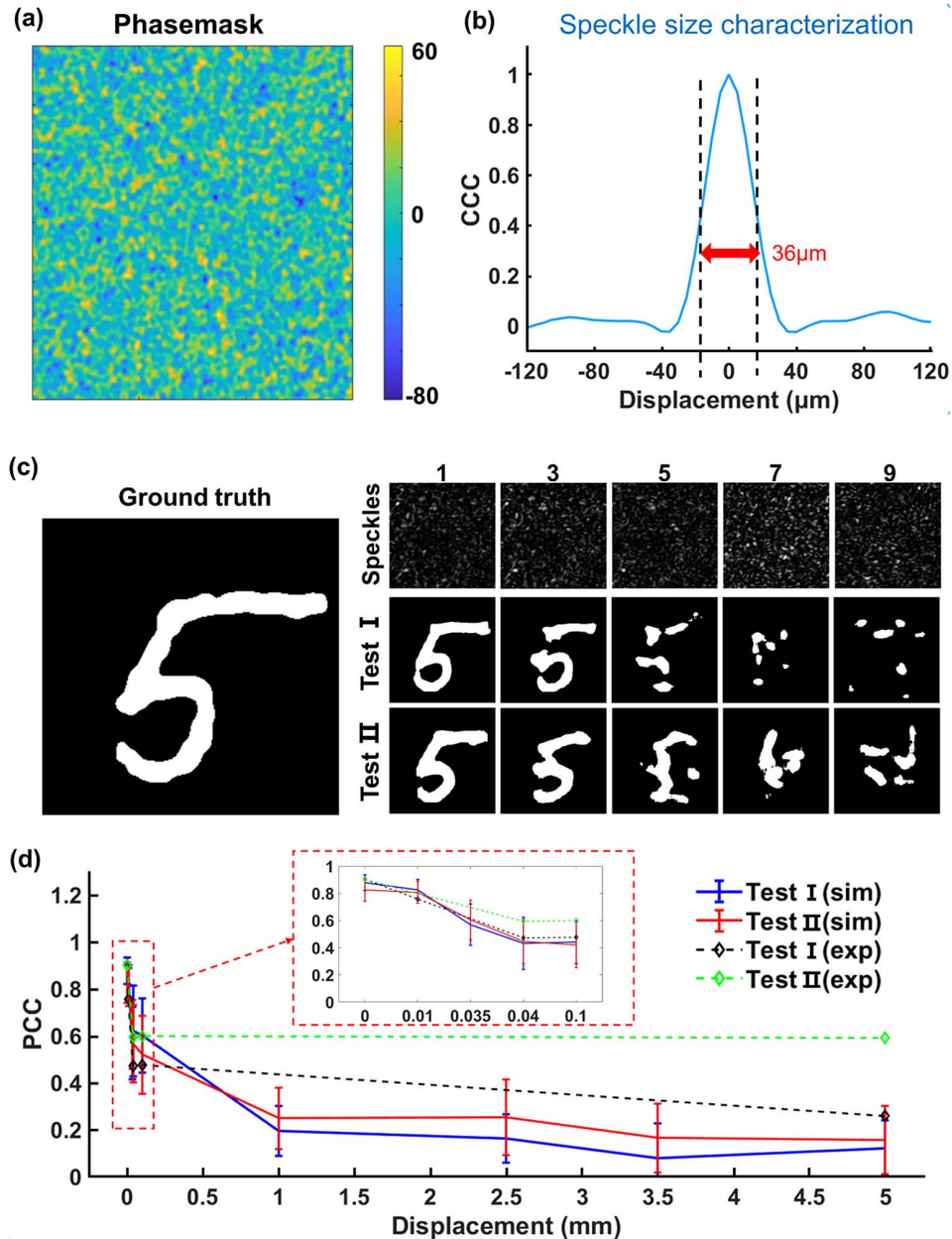


Fig. 3. Results of Simulation I where no ballistic light is involved. (a) The phase map of the simulated diffuser in which the color bar denotes the range of phase value in radian. (b) The characterized CCC curve of the simulated diffuser, which has a FWHM of $\sim 36 \mu\text{m}$ and a base level of zero. (c) The speckle patterns and predicted images in both Tests I and II at regions 1, 3, 5, 7, and 9 with the ground truth on the left. (d) Curves of averaged PCC with error bar for two tests in Simulation I in which experimental PCC results are also included for comparison. The right subplot shows the zoom-in area of the dash rectangle.

whereas all the reconstruction metrics are given in Fig. 3(d). We can see the objects can be well restored at regions 1 and 2 and recognizable at region 3 ($\Delta x = 35 \mu\text{m}$) although at a decaying quality. Interestingly, in both Tests I and II, the network seems not to reconstruct the objects from region 4 and above, once the displacement is beyond the isoplanatic range. This is confirmed by the high consistence of the PCC curves between the simulated Tests I and II and the experimental Test I within 0.1 mm displacement as shown in the enlarged subplot of Fig. 3(d). This suggests that under the condition of no ballistic component, a reconstruction model hardly generalizes to

an unknown diffuser (region) even if adopting a one-to-all training strategy in Test II.

Given the above results, we hypothesize the scalability of network observed in experimental Test II is preconditioned with ballistic light component. The scattering component can provide specific statistical information of a diffuser (region), i.e., a decryption key, but is unable to be used solely for training an adaptive network even from multiple sources. The immunity of the ballistic component to the change of scattering condition (e.g., the shift of diffuser region) may play an indispensable role in the model adaptivity.

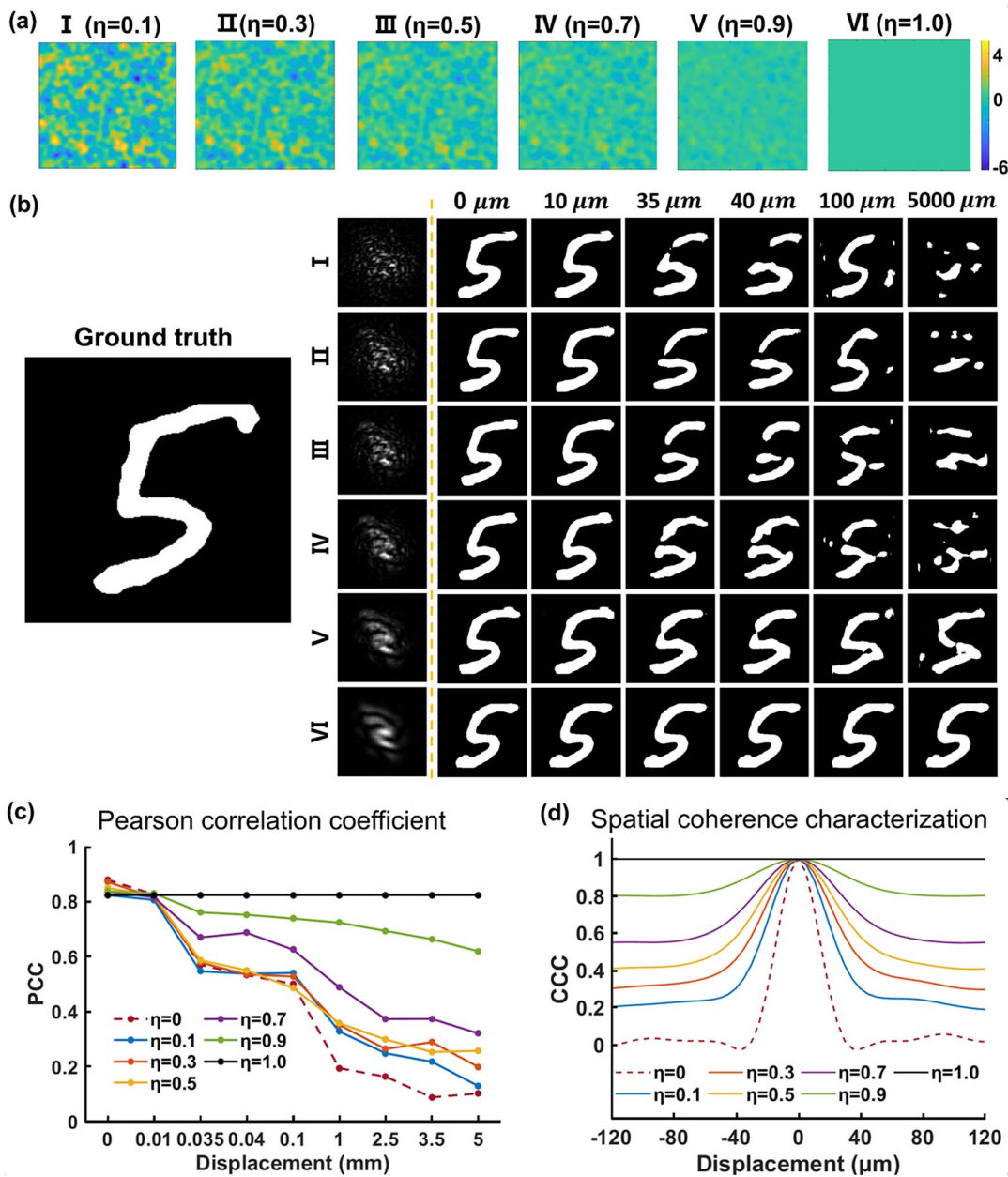


Fig. 4. Results of Simulation II that involves different weights of ballistic light (η). (a) Phase distributions of the simulated diffusers corresponding to different value of η . (b) The image reconstruction results on test regions 1–6 of varying Δx when the network is trained under different η . Note that rows I–VI correspond to $\eta = 0.1, 0.3, 0.5, 0.7, 0.9$, and 1 , respectively. (c) The curves of average PCC as a function of displacement for different η . (d) The CCC curves of output field under different η .

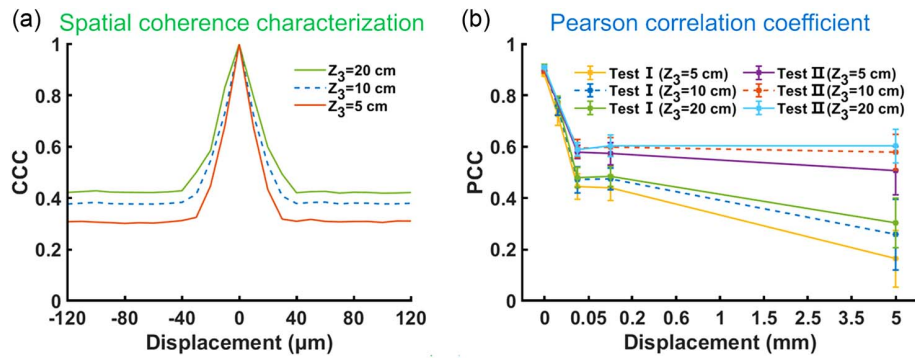


Fig. 5. Improved model generalization by increasing distance z_3 in experiment. (a) The CCC curves measured experimentally for $z_3 = 5, 10, 20$ cm, respectively. (b) The curves of average PCC for network testing on a series of regions in experimental Tests I and II under the case of different z_3 .

To confirm the hypothesis, Simulation II was further performed to investigate the influence of ballistic light on the adaptivity of a model when trained with single-source data. The range of phase distribution of the simulated diffuser was adjusted to control $\eta = 0.1, 0.3, 0.5, 0.7, 0.9$, and 1, respectively, as seen in Fig. 4(a). The image reconstruction results at different test regions (denoted by Δx) for each case of η are summarized in Figs. 4(b) and 4(c). Qualitatively, it can be observed from Fig. 4(b) that the generalization capability of network to displacements is enhanced with increasing weight of ballistic light (η). Such a trend is more straightforward from the PCC curves shown in Fig. 4(c) where one corresponding to a larger η is generally at a higher level among all the test regions.

At $\Delta x = 0 \mu\text{m}$, the average PCCs for different η are almost the same, although a slight bias for the case of $\eta = 1$ is observed. The reason may be that the information extraction efficiency of a network from speckle patterns is slightly higher than from diffraction patterns directly as more high-frequency components of an object could be encoded by the former due to the larger scattering angles. Regarding the performance of network generalization at the existence of scattering component (i.e., $\eta \neq 1$), which originates from a specific encryption key, the poorer recovery at a test region is mainly due to the mismatch between the encryption and the decryption keys. We show that such mismatch can be mitigated by increasing weight of ballistic light. According to the CCC curves under different η [Fig. 4(d)], the isoplanatic range of the output field grows proportionally with increasing ballistic light. In the extreme case of $\eta = 1$, the output field is solely the diffracted object [Fig. 4(b) VI] as no scattering is induced by the “flat” diffuser [Fig. 4(a) VI]. Consequently, data from different regions all show the same diffraction characteristic, which means $\text{CCC} = 1$ among the output field and the trained network can generalize to unknown diffuser regions freely. Through the above simulation tests, the roles of scattered and ballistic light on the model generalization are clarified. In particular, the latter contributes to the spatial coherence of output field to impact the network scalability.

Our findings provide practicable guidance in enhancing a DL model for scalable imaging through scattering media. In our experiment, a simple way for increasing the weight of

ballistic light is to increase the distance z_3 since the scattered photons of relatively large divergence angle can be partially filtered out during free-space propagation. This is verified by the experimental results shown in Fig. 5 where both the base level and the FWHM of the CCC curve are improved [Fig. 5(a)], meaning stronger spatial coherence in the output field with larger z_3 . Consequently, the performance of network in the generalization Tests I and II is also improved with an increase in z_3 [Fig. 5(b)].

4. DISCUSSIONS AND CONCLUSION

Although there have been many DL studies aiming at improving the model generalization for descattered imaging, the origin and boundary of such model scalability from the perspective of physics were still unclear. In this paper, through both experimental and simulated tests, we found the ballistic component was closely related to the model adaptivity. Furthermore, the different mechanisms of scattered and ballistic light for a reconstruction model were revealed. The scattering component was encrypted by a diffuser (region) from which a model learned a decryption key specific to the diffuser property. Although training the model with data from multiple encryption keys seemed to allow it to learn the statistical information of all diffusers with enhanced scalability, it was preconditioned with the ballistic component that offered a physical meaning invariance among the speckle data. Additionally, the model scalability was enhanced with larger weight of ballistic light as the spatial coherence of output field can be stronger. Based on the above findings, the network generalization ability was enhanced in experiment by increasing the detected ballistic component.

A few more discussions about this paper were clarified herein. First, it seemed the previous DL works about scalable imaging through scattering media [23–26] did not differentiate the roles of scattered and ballistic light; thus, were not aware of the ballistic contribution as a precondition. Usually, it was not the case to have a purely diffuse system in experiment when the diffuser was at the focal plane of subsequent imaging optics. There did exist a part of ballistic light, which could be estimated by comparing the experimentally measured CCC curve with the simulated ones [Fig. 4(d)]. Second, in our paper,

instead of using multiple diffusers, multiple regions of one diffuser were employed, which corresponded to different TMs, whereas, having the same mean scattering characteristics. Finally, in addition to revealing the prerequisite of ballistic light during one-to-all model training for better scalability, our paper also tried to define the physical boundary for a DL model trained under varying scattering conditions. It was hypothesized that the number of encryption keys involved during network training, resulting from internal or external perturbations, a change in incident beam, etc., cannot be unlimited. This may confuse a network due to the interference among the keys or even exceed its recognition capability. Besides, DL for scalable imaging through a thick scattering medium also saw limitation because of a lack of ballistic contribution. Enough data under each encryption key (corresponding to various scattering conditions) with an invariant correlation (i.e., ballistic component) among them, will define the boundary for the construction of a scalable DL model and further application. For example, a recent work for 3D adaptive descattering via a DSN [26] still relied on training data with sufficient ballistic component by detecting mostly the single-scattered photons in holographic particle imaging.

To summarize, our findings added new knowledge to the physical mechanisms of utilizing DL for scalable imaging through scattering media in which the roles of scattered and ballistic light components were revealed for contributing to the origin and physical boundary of the model adaptivity. The paper also offered practical guidance for improving the DL scalability by gating the ballistic photons. This can be performed by increasing the diffraction distance in our setup (also adopted for Refs. [19,25]) or introducing a spatial filter in a general setup [8,18,23] where the diffuser is imaged onto a camera via an objective lens or $4f$ system. Nowadays, the one-to-all training strategy had become the mainstream way to increase network generalization. Herein, our research results deepened the cognition on this mainstream method and provided guiding significance for the related areas. It reminded us that the invariant correlation among the multisource speckle data was a prerequisite for successful one-to-all training. Besides, the physical boundary of applying DL to descattering with general applicability was also defined to prescribe the scope of application. The mechanism understanding and guidance value of our research were beneficial for developing DL frameworks for scalable imaging under dynamic scattering scenarios.

Funding. National Natural Science Foundation of China (81930048); Guangdong Science and Technology Department (2019BT02X105); Research Grants Council, University Grants Committee (15217721, C7074-21GF, R5029-19); Innovation and Technology Commission (GHP/043/19SZ, GHP/044/19GD); Hong Kong Polytechnic University (P0038180, P0039517, P0043485).

Acknowledgment. H. L. conceived the idea and designed the experiment and simulation. X. Z., J. G., Y. G., and C. S. implemented the experiment and simulation. X. Z., S. C., P. L., and H. L. analyzed the data and wrote the paper. All contributed to revising the paper.

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.

REFERENCES

1. S. Rotter and S. Gigan, "Light fields in complex media: mesoscopic scattering meets wave control," *Rev. Mod. Phys.* **89**, 015005 (2017).
2. J. Bertolotti and O. Katz, "Imaging in complex media," *Nat. Phys.* **18**, 1008–1017 (2022).
3. Z. Yu, H. Li, T. Zhong, J.-H. Park, S. Cheng, C. M. Woo, Q. Zhao, J. Yao, Y. Zhou, X. Huang, W. Pang, H. Yoon, Y. Shen, H. Liu, Y. Zheng, Y. Park, L. V. Wang, and P. Lai, "Wavefront shaping: a versatile tool to conquer multiple scattering in multidisciplinary fields," *Innovation* **3**, 100292 (2022).
4. S. W. Paddock, "Principles and practices of laser scanning confocal microscopy," *Mol. Biotechnol.* **16**, 127–149 (2000).
5. F. Helmchen and W. Denk, "Deep tissue two-photon microscopy," *Nat. Methods* **2**, 932–940 (2005).
6. D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, and C. A. Puliafito, "Optical coherence tomography," *Science* **254**, 1178–1181 (1991).
7. V. Ntziachristos, "Going deeper than microscopy: the optical imaging frontier in biology," *Nat. Methods* **7**, 603–614 (2010).
8. S. Popoff, G. Lerosey, M. Fink, A. C. Boccara, and S. Gigan, "Image transmission through an opaque material," *Nat. Commun.* **1**, 81 (2010).
9. Y. Choi, T. D. Yang, C. Fang-Yen, P. Kang, K. J. Lee, R. R. Dasari, M. S. Feld, and W. Choi, "Overcoming the diffraction limit using multiple light scattering in a highly disordered medium," *Phys. Rev. Lett.* **107**, 023902 (2011).
10. 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**, 784–790 (2014).
11. M. Chen, H. Liu, Z. Liu, P. Lai, and S. Han, "Expansion of the FOV in speckle autocorrelation imaging by spatial filtering," *Opt. Lett.* **44**, 5997–6000 (2019).
12. 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).
13. E. Tajahuerce, V. Durán, P. Clemente, E. Irlés, F. Soldevila, P. Andrés, and J. Lancis, "Image transmission through dynamic scattering media by single-pixel photodetection," *Opt. Express* **22**, 16945–16955 (2014).
14. 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).
15. Y. Luo, S. Yan, H. Li, P. Lai, and Y. Zheng, "Towards smart optical focusing: deep learning-empowered dynamic wavefront shaping through nonstationary scattering media," *Photon. Res.* **9**, B262–B278 (2021).
16. A. Turpin, I. Vishniakou, and J. D. Seelig, "Light scattering control in transmission and reflection with neural networks," *Opt. Express* **26**, 30911–30929 (2018).
17. N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, "Learning to see through multimode fibers," *Optica* **5**, 960–966 (2018).
18. 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).
19. 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).
20. S. Cheng, H. Li, Y. Luo, Y. Zheng, and P. Lai, "Artificial intelligence-assisted light control and computational imaging through scattering media," *J. Innov. Opt. Health Sci.* **12**, 1930006 (2019).
21. H. Li, Z. Yu, Q. Zhao, T. Zhong, and P. Lai, "Accelerating deep learning with high energy efficiency: from microchip to physical systems," *Innovation* **3**, 100252 (2022).

22. 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).
23. 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).
24. P. Fan, T. Zhao, and L. Su, "Deep learning the high variability and randomness inside multimode fibers," *Opt. Express* **27**, 20241–20258 (2019).
25. S. Zhu, E. Guo, J. Gu, L. Bai, and J. Han, "Imaging through unknown scattering media based on physics-informed learning," *Photon. Res.* **9**, B210–B219 (2021).
26. W. Tahir, H. Wang, and L. Tian, "Adaptive 3D descattering with a dynamic synthesis network," *Light Sci. Appl.* **11**, 42 (2022).
27. Y. Liu, F. Tang, X. Wang, C. Peng, and P. Li, "Applicability of the Van Cittert–Zernike theorem in a Ronchi shearing interferometer," *Appl. Opt.* **61**, 1464–1474 (2022).
28. L. Deng, "The MNIST database of handwritten digit images for machine learning research [best of the web]," *IEEE Signal Process Mag.* **29**, 141–142 (2012).
29. O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *18th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (Springer, 2015), pp. 234–241.
30. Y.-P. Zhao, I. Wu, C.-F. Cheng, U. Block, G.-C. Wang, and T.-M. Lu, "Characterization of random rough surfaces by in-plane light scattering," *J. Appl. Phys.* **84**, 2571–2582 (1998).