

# Investigation of image plane for image reconstruction of objects through diffusers via deep learning

Takumi Tsukada and Wataru Watanabe\*

Ritsumeikan University, College of Science and Engineering,  
Department of Electrical and Electronic Engineering, Kusatsu, Shiga, Japan

## Abstract

**Significance:** The imaging of objects hidden in light-scattering media is a vital practical task in a wide range of applications, including biological imaging. Deep-learning-based methods have been used to reconstruct images behind scattering media under complex scattering conditions, but improvements in the quality of the reconstructed images are required.

**Aim:** To investigate the effect of image plane on the accuracy of reconstructed images.

**Approach:** Light reflected from an object passing through glass diffusers is captured by changing the image plane of an optical imaging system. Images are reconstructed by deep learning, and evaluated in terms of structural similarity index measure, classification accuracy of digital images, and training and testing error curves.

**Results:** The reconstruction accuracy was improved for the case in which the diffuser was imaged, compared to the case where the object was imaged. The training and testing error curves show that the loss converged to lower values in fewer epochs when the diffuser was imaged.

**Conclusions:** The proposed approach demonstrates an improvement in the accuracy of the reconstruction of objects hidden through glass diffusers by imaging glass diffuser surfaces, and can be applied to objects at unknown locations in a scattering medium.

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JBO.27.5.056001](https://doi.org/10.1117/1.JBO.27.5.056001)]

**Keywords:** deep learning; scattering imaging; diffuser; speckle.

Paper 210292SSR received Sep. 15, 2021; accepted for publication Jan. 31, 2022; published online May 4, 2022.

## 1 Introduction

Random media or scattering media such as frosted glasses and biological tissue significantly change and diffuse the properties of transmitted light, which reduces the quality of transmitted images. Visualization of objects inside or through scattering media is a challenging problem for various applications. In recent years, a variety of approaches such as point spread functions,<sup>1,2</sup> speckle correlation,<sup>3</sup> transmission matrices,<sup>4</sup> and wavefront shaping<sup>5,6</sup> have been demonstrated to visualize objects behind scatterers using randomly scattered light. However, these imaging methods do not perform well for objects obscured by complex scattering media.<sup>7</sup> Recently, data-driven approaches to imaging through scattering media have been proposed using support vector regression architecture<sup>8</sup> and deep learning.<sup>9</sup> Deep-learning-based methods solve inverse problems by introducing a large number of pairs of object images and speckle images to the deep-learning model.<sup>10,11</sup> The model trains an inverse scattering function, which enables the reconstruction of object images.<sup>12</sup> Various methods have been proposed for training networks. Among these networks, convolutional neural networks (CNNs) have been used as a more flexible and generalized technique to reconstruct objects behind scatterers or diffusers. CNNs have been used to retrieve amplitude objects behind a random medium using either a single<sup>13</sup> or multiple diffusers.<sup>14</sup> Deep-learning-based methods also enable imaging under low-photon conditions<sup>15</sup> and

---

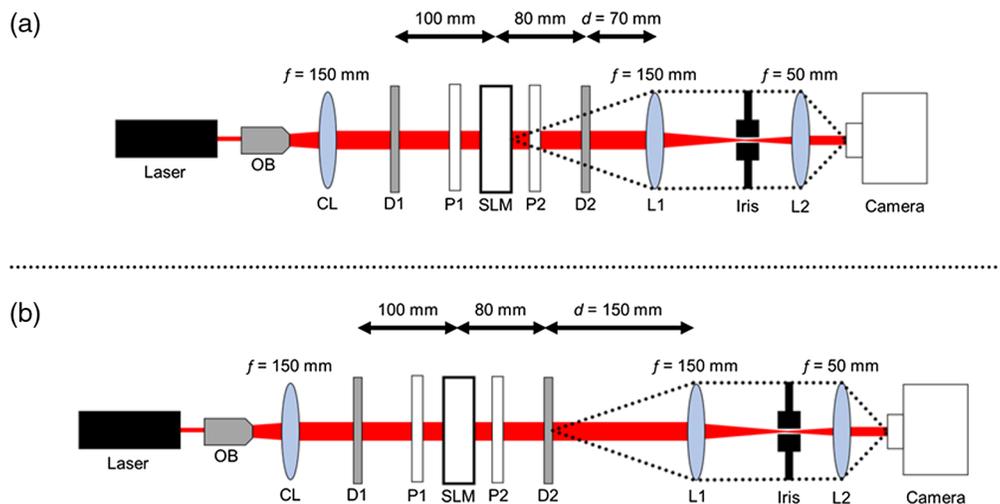
\*Address all correspondence to Wataru Watanabe, [wata-w@fc.ritsumei.ac.jp](mailto:wata-w@fc.ritsumei.ac.jp)

reconstruction of objects through dynamic scattering media.<sup>16</sup> However, these methods face several challenges, such as the selection of the training framework, scalability, and optical systems. The scope of this work is limited to optical systems for imaging through light-scattering obscuring objects using deep learning. Optical systems can be classified into those encoding objects as transmission mode<sup>8,14,15,17–19</sup> and reflection mode.<sup>9–11,13,16,20–24</sup> In both modes, optical systems, such as lensless imaging systems,<sup>8,10,17,20,21,23</sup> imaging systems with one or two lenses,<sup>9,11,14–16,18,19,22</sup> and imaging systems with a camera lens<sup>13,24</sup> have been investigated. Among these, the  $4f$  optical system consisting of two lenses has the advantage of removing scattered light by spatial filtering, a large field of view, ease of magnification by replacing the lens system, and light-collection efficiency.<sup>9,11,18,19,22</sup> In previous reports, the image plane was set as an object behind glass diffusers, assuming that the image plane was known.<sup>9,11,18</sup> In real space, the image plane of an object is unknown when the object is hidden by scattering materials or glass diffusers. In addition, when objects are located in thick scattering media, such as biomedical tissue, image reconstruction of objects at unknown depths must be considered.

In this study, we investigate the effects of the image plane of an optical system on the classification accuracy of reconstructed images through glass diffusers using deep learning. We investigate two cases, one in which the diffuser is imaged and the other where the object is imaged. The experimental results show that imaging on the diffuser plane improves the accuracy compared to imaging the object plane through the image reconstruction and image evaluation.

## 2 Method

A schematic of the optical system is shown in Fig. 1. We used a transmission-type spatial light modulator (SLM) (Holoeye LC2012,  $36\text{-}\mu\text{m}$  pixel pitch,  $1024 \times 768$ ). To obtain amplitude modulation, we placed a polarizer P2 whose polarization direction was perpendicular with respect to P1. The central  $160 \times 160$  pixels of the SLM were used in the experiments. Coherent light from a He–Ne laser source (LASOS LGK7654-8,  $632.8\text{ nm}$ ) was expanded with an objective lens (OB) with a  $20\times$  magnification, and was collimated by the collimating lens (CL) ( $f = 150\text{ mm}$ ). The  $4f$  optical system consisted of two lenses L1 ( $f = 150\text{ mm}$ ) and L2 ( $f = 50\text{ mm}$ ). Speckle images were recorded with a charged coupled device (CCD) camera (The Imaging Source DMK23U445, 8 bits,  $3.75\text{-}\mu\text{m}$  pixel pitch,  $1280 \times 960$ ), of which only the central  $256 \times 256$  portion was used in the experiments. In Fig. 1(a), the object plane on the SLM was imaged on the CCD camera. In Fig. 1(b), the diffuser plane (D2) was on the CCD camera. The distance between diffuser D1 and SLM was set to  $100\text{ mm}$ , and that between the



**Fig. 1** Schematic of imaging system for acquiring speckle images. OB: objective lens, CL: collimating lens, D: diffuser, P: polarizer, SLM: spatial light modulator, and L1, L2: lens. The distance between the diffuser D2 and the lens L1 was denoted by  $d$ . (a) SLM plane is imaged. (b) D2 plane is imaged.

SLM and diffuser D2 was set to 80 mm. An iris with a diameter of 7.5 mm was placed at the Fourier plane of the lens L1.

To evaluate our approach, we used the MNIST dataset of handwritten digits as input images for SLM.<sup>25</sup> For diffusers D1 and D2, we used a holographic diffuser and a white diffusive glass, as shown in Table 1.

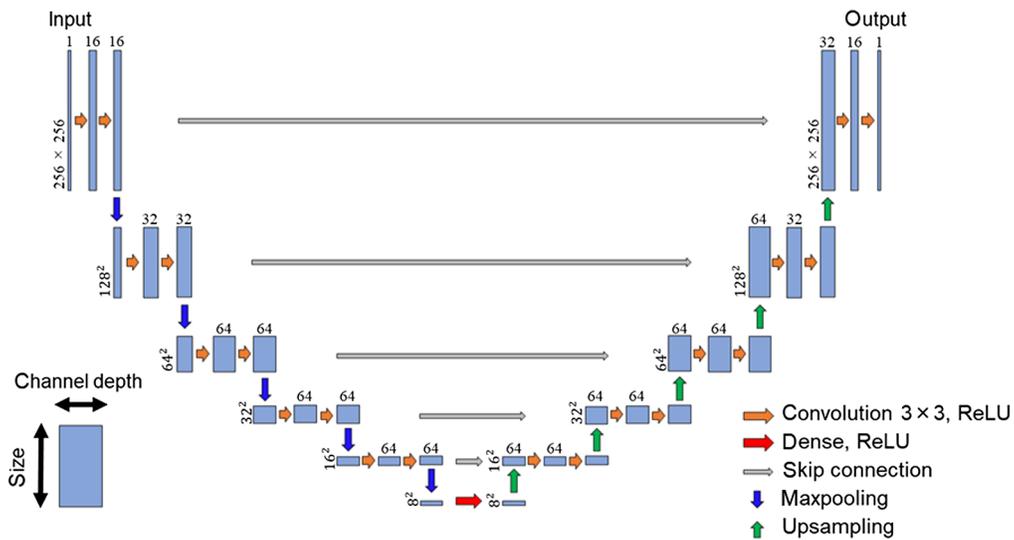
The deep-learning model used in this experiment is shown in Fig. 2. We used CNN, and our deep-learning model possessed a U-Net architecture.<sup>26</sup> The model was trained using speckle images of  $256 \times 256 \times 1$  ( $256 \times 256$   $xy$  size, 1 channel depth) acquired by the camera as inputs. In the encoding part, the input data were repeatedly passed through two convolutional layers with a  $3 \times 3$  kernel and a maxpooling layer with a  $2 \times 2$  kernel. The  $xy$  size was reduced and the channel depth increased. Finally, the input data resulted in  $8 \times 8 \times 64$  ( $8 \times 8$   $xy$  size, 64 channel depth). The encoding part and the decoding part were connected with a densely connected layer. In the decoding part, the channel depth was gradually reduced back to the same size as the input size, in contrast to the encoding part. Only the convolution layer before the output had a kernel size of  $1 \times 1$  to match the output size. The loss function was the negative Pearson correlation coefficient (NPCC), and we used Adam to update the weights.<sup>11</sup> NPCC is expressed as

$$\text{NPCC} = \frac{-1 \times \sum_{i=1}^w \sum_{j=1}^h (X(i, j) - \bar{X})(Y(i, j) - \bar{Y})}{\sqrt{\sum_{i=1}^w \sum_{j=1}^h (X(i, j) - \bar{X})^2} \sqrt{\sum_{i=1}^w \sum_{j=1}^h (Y(i, j) - \bar{Y})^2}}, \quad (1)$$

where  $w$  and  $h$  are the vertical and horizontal sizes of the images,  $X(i, j)$  and  $Y(i, j)$  are ground-truth images and reconstructed images, respectively, and  $\bar{X}$  and  $\bar{Y}$  are the mean values of pixels in these images. Keras was used to build the deep-learning model, which was run on a GPU (GeForce GTX 1660 Ti).

**Table 1** Description of diffusers used in experiments.

Name	Diffusion angle	Thickness
Holographic diffuser (Edmund Optics 47996)	5 deg	0.78 mm
White diffusing glass (Edmund Optics 34480)	120 deg	$1.25 \pm 0.1$ mm



**Fig. 2** The deep-learning model used for image reconstruction. The structure of the deep-learning model follows the U-Net architecture. The height of each block represents the size. The width of each block represents the channel depth.

We captured 3000 speckle images, of which 2220 were used for training, 560 for validation, and 200 for testing. We used structural similarity (SSIM) and accuracy (ACC) as indices of evaluation of the reconstructed images. ACC denotes the classification accuracy of the digital images, which is defined by the correct answer rates of identification in the number recognition program trained by CNN. SSIM is calculated as

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (2)$$

where  $\mu_x$  and  $\mu_y$  are the means of the pixel value,  $\sigma_x^2$  and  $\sigma_y^2$  are the variances of the pixel values,  $\sigma_{xy}$  is the covariance of the pixel values,  $C_1 = (255 \times 0.01)^2$  and  $C_2 = (255 \times 0.03)^2$  are the normalization constants. The SSIM and ACC in the following results are the average values of 200 reconstructed images.

### 3 Experimental Results and Discussion

#### 3.1 Image Reconstruction Using One Diffuser

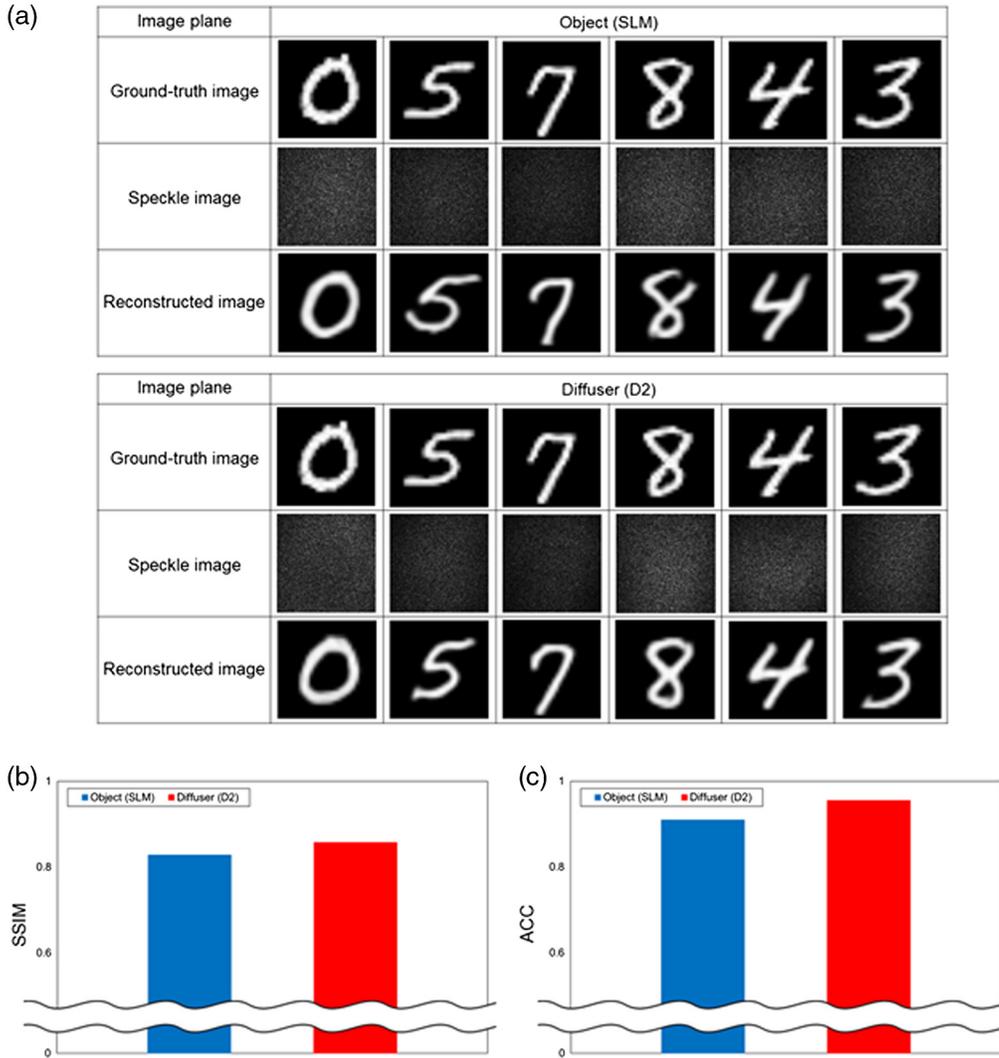
To investigate the image plane in the optical system, we first conducted experiments using one diffuser. We only used the white diffusive glass with a diffusion angle of 120 deg as D2 and no diffuser was placed at D1, as shown in Figs. 1(a) and 1(b). Figure 3 shows the reconstruction results and the evaluations of the reconstructed images. The images were reconstructed as high-quality images in both image planes, as shown in Fig. 3(a). Figures 3(b) and 3(c) show the SSIM and ACC evaluations, respectively. As shown in Fig. 3(b), the SSIM was  $\sim 0.828$  when the object was imaged and  $\sim 0.857$  when the diffuser was imaged. As shown in Fig. 3(c), the ACC was 91% when the object was imaged and 95.5% when the diffuser was imaged. Evaluations of the reconstructed images were better when the image plane was set as the diffuser and one diffuser was placed between the object and collecting lens L1 in a  $4f$  imaging system.

#### 3.2 Image Reconstruction of an Object Placed between Two Diffusers

A difficulty arises in imaging an object placed between two diffusers, in that the object is illuminated with a diffused wave.<sup>14,27–29</sup> We investigated image reconstruction of an object placed between two diffusers, as shown in Fig. 1. We used the holographic diffuser as D1 and white diffusive glass as D2. Figure 4 shows the reconstruction results and the evaluations of the reconstructed images. As shown in Fig. 4(a), when the object was imaged, the reconstructed images were deformed. However, when the diffuser was imaged, the reconstructed images were of high quality even when the object was placed between the two diffusers. Figures 4(b) and 4(c) show the SSIM and ACC evaluations, respectively. As shown in Fig. 4(b), the SSIM was  $\sim 0.739$  when the object was imaged and  $\sim 0.846$  when the diffuser was imaged. As shown in Fig. 4(c), the ACC was 75.5% when the object was imaged and 92.5% when the diffuser was imaged. Based on the results, it may be noted that the accuracy was higher when the diffuser was imaged than when the object behind the diffuser was imaged.

The training and testing error curves of the CNN for each imaging condition are shown in Fig. 5, for the experimental conditions where the object was placed between two diffusers. The left panel in Fig. 5 shows the dependence of the loss value on the epoch number when the object was imaged, showing that the test loss value decreased with the epoch number and eventually reached a minimum  $-0.81$  in the training process of the CNN. The test loss converged at  $\sim 40$  epochs, and the difference between the training loss and test loss was large. The right panel in Fig. 5 shows the dependence of the loss value on the epoch number when the diffuser was imaged, showing that the loss value reached a minimum of  $-0.92$  in the training process of the CNN. The testing loss converged faster, i.e., epochs of  $\sim 20$  epochs, and the difference between the training loss and test loss was smaller than that when the object was imaged.

To gain insight into the dependence of the quality on the image plane, we calculated the Pearson correlation coefficient (PCC) between the ground-truth image and speckle images at

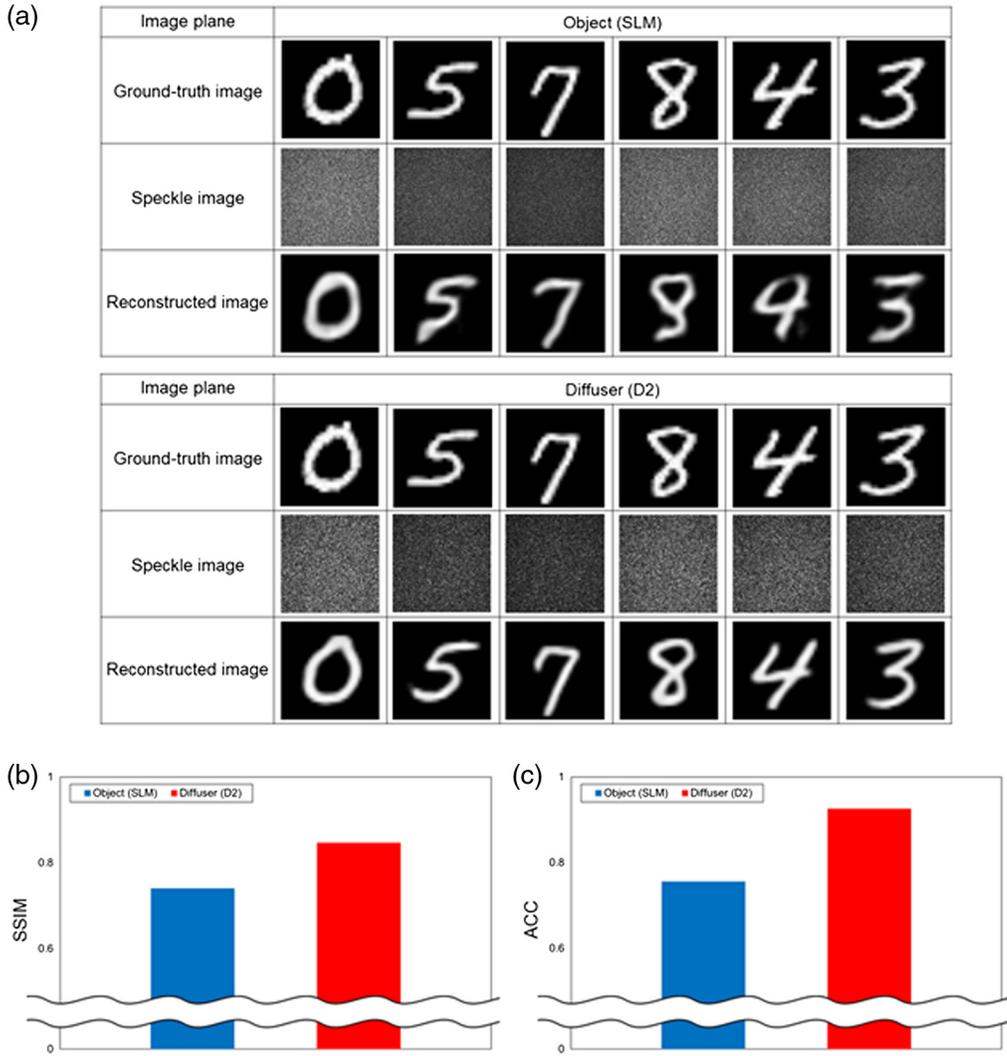


**Fig. 3** (a) Reconstruction results of objects behind the diffuser and (b) SSIM and (c) ACC evaluations of reconstructed images. The legends correspond to the image planes.

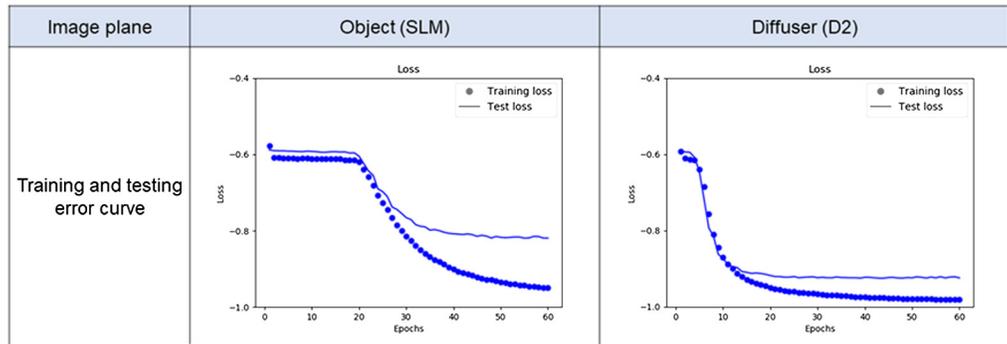
different image planes by changing the distance between the diffuser D2 and the lens L1,  $d$  (see Fig. 1). Figure 6 shows the PCC between ground-truth image and speckle images captured by changing the image plane. When  $d$  was 70 mm, the SLM was imaged [Fig. 1(a)]. When  $d$  was 150 mm, the diffuser D2 was imaged [Fig. 1(b)]. The PCC became higher as the image plane approached  $d = 150$  mm, and the highest PCC was obtained at the distance  $d$  of 150 mm. When the image plane was defocused from the diffuser D2 to the camera side, the PCC decreased. When the object was imaged, the propagated and diffracted light from the object was diffused by a diffuser. Therefore, imaging the object was associated with more complex training and reconstruction processes. When the diffuser was imaged, the propagation and diffracted light from the object is projected onto the diffuser. The diffuser plane is considered to be a secondary image plane by mapping the plane of the object image; therefore, the training becomes simple. The image plane affects the image reconstruction results and epochs in deep learning.

### 3.3 Discussion

We investigated the dependence of the distance between the SLM and diffuser D2 by shifting D2 along the optical axis when the object was imaged. When the distance between the SLM and D2 increased, the training speed decreased. As mentioned in Sec. 3.2, when the diffuser was placed in front of the object, the diffuser surface becomes the secondary image plane. Therefore, when

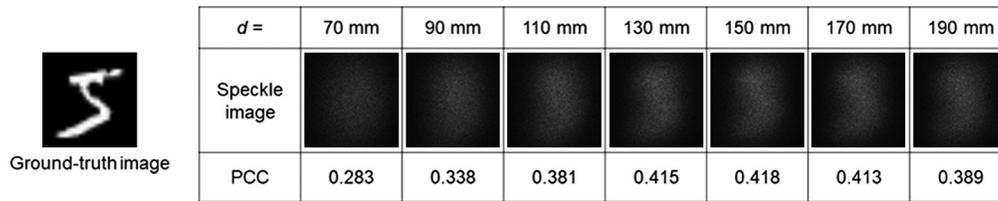


**Fig. 4** (a) Reconstruction results of objects placed between two diffusers and (b) SSIM and (c) ACC evaluations of reconstructed images.



**Fig. 5** Quantitative analysis with the training and testing error curves. Training and testing error of CNN using NPCC as the loss function.

the object was imaged, the closer the distance between the object (SLM) and the diffuser (D2), the smaller the effect of defocus and the faster the training speed. When the diffuser was imaged, the accuracy was better than when the object was imaged because the effect of defocus was the least.



**Fig. 6** The PCC between ground-truth image and speckle images captured by changing the image plane.

The position of the two diffusers in Sec. 3.2 was switched, and the experiment was also conducted with the white diffusive glass (diffusion angle of 120 deg) as D1 and the holographic diffuser (diffusion angle of 5 deg) as D2. When the diffusion angle in D2 was smaller, over 90% ACC was achieved in both image plane conditions. The results show that the diffuser placed behind the object, D2, had a strong effect on the image reconstruction. The diffusion angle in D2 was smaller, and the reconstruction accuracy improved.

We performed experiments with two diffusers placed in front of the object with respect to lens L1. In this case, no diffuser was placed at D1, as shown in Fig. 1, and two diffusers were placed between the object and lens L1. The reconstruction accuracy was also better when the diffusers were imaged than when the object was imaged, and more than 90% of the ACC was achieved by imaging the diffusers.

In these experiments, we demonstrated the feasibility of our approach by using diffusers with diffusion angles of 5 deg and 120 deg. Similar trends were obtained with the combination of two diffusers having diffusion angles of 5 deg, 10 deg, 20 deg, and 120 deg.

We verified our approach by using a  $4f$  optical system. This approach is versatile, being feasible with different optical systems. We performed experiments using a camera with a C-mount lens, which has a wide field of view. In this experiment, L1, L2, and the iris shown in Fig. 1 were removed, and a C-mount lens (The Imaging Source TCL 1216 5MP) was directly attached to the CCD camera. The image plane was set to be the diffuser surface, and the speckle images were captured by a camera with a C-mount lens with two diffusers. We used a holographic diffuser with a diffusion angle of 5 deg as D1 and white diffusive glass on D2. The results demonstrated that over 90% ACC was achieved.

## 4 Conclusion

We have demonstrated an imaging technique that allows the reconstruction of objects hidden behind glass diffusers by investigating their image planes in optical imaging systems. We reconstructed an object behind a diffuser or placed between two diffusers in a  $4f$  optical system with deep learning. The accuracy improved when the diffuser was imaged. By imaging the diffuser, scattering imaging can be performed regardless of the position of the object behind the diffuser. These results are expected to be applied to cases where the target object is located inside a scattering medium with an unknown position, such as in biological imaging.

## Disclosures

The authors have no relevant financial interests in this paper and no potential conflicts of interest to disclose.

## Acknowledgments

The authors thank T. Nishijima from Ritsumeikan University for assistance with the deep-learning model. This work was supported by JSPS KAKENHI Grant No. 21H01849.

## References

1. E. Edrei and G. Scarcelli, "Memory-effect based deconvolution microscopy for super-resolution imaging through scattering media," *Sci. Rep.* **6**(1), 33558 (2016).
2. L. Li et al., "Imaging through scattering layers exceeding memory effect range with spatial-correlation-achieved point-spread-function," *Opt. Lett.* **43**(8), 1670–1673 (2018).
3. O. Katz et al., "Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations," *Nat. Photonics* **8**(10), 784–790 (2014).
4. M. Kim et al., "Transmission matrix of a scattering medium and its applications in biophotonics," *Opt. Express* **23**(10), 12648–12668 (2015).
5. I. M. Vellekoop and A. P. Mosk, "Focusing coherent light through opaque strongly scattering media," *Opt. Lett.* **32**(16), 2309–2311 (2007).
6. A. P. Mosk et al., "Controlling waves in space and time for imaging and focusing in complex media," *Nat. Photonics* **6**(5), 283–292 (2012).
7. 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).
8. R. Horisaki, R. Takagi, and J. Tanida, "Learning-based imaging through scattering media," *Opt. Express* **24**(13), 13738–13743 (2016).
9. 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).
10. A. Sinha et al., "Lensless computational imaging through deep learning," *Optica* **4**(9), 1117–1125 (2017).
11. S. Li et al., "Imaging through glass diffusers using densely connected convolutional networks," *Optica* **5**(7), 803–813 (2018).
12. G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica* **6**(8), 921–943 (2019).
13. C. A. Metzler et al., "Deep-inverse correlography: towards real-time high-resolution non-line-of-sight imaging," *Optica* **7**(1), 63–71 (2020).
14. Q. Li et al., "Imaging reconstruction through strongly scattering media by using convolutional neural networks," *Opt. Commun.* **477**(1), 126341 (2020).
15. L. Sun et al., "Photon-limited imaging through scattering medium based on deep learning," *Opt. Express* **27**(23), 33120–33134 (2019).
16. Y. Sun et al., "Image reconstruction through dynamic scattering media based on deep learning," *Opt. Express* **27**(11), 16032–16046 (2019).
17. K. Yamazaki, R. Horisaki, and J. Tanida, "Imaging through scattering media based on semi-supervised learning," *Appl. Opt.* **59**(31), 9850–9854 (2020).
18. T. Nishijima, H. Arimoto, and W. Watanabe, "Image reconstruction behind diffuser by deep learning and spatial filtering," *Proc. SPIE* **11521**, 1152116 (2020).
19. I. Kang et al., "Recurrent neural network reveals transparent objects through scattering media," *Opt. Express* **29**(4), 5316–5326 (2021).
20. M. Lyu et al., "Learning-based lensless imaging through optically thick scattering media," *Adv. Photonics* **1**(3), 036002 (2019).
21. E. Guo et al., "Learning-based method to reconstruct complex targets through scattering medium beyond the memory effect," *Opt. Express* **28**(2), 2433–2446 (2020).
22. M. Deng et al., "On the interplay between physical and content priors in deep learning for computational imaging," *Opt. Express* **28**(16), 24152–24170 (2020).
23. S. Zhu et al., "Imaging through unknown scattering media based on physics-informed learning," *Photonics Res.* **9**(5), B210–B219 (2021).
24. Q. Hu et al., "Object recognition for remarkably small field-of-view with speckles," *Appl. Phys. Lett.* **118**(9), 091103 (2021).
25. Y. LeCun, C. Cortes, and C. J. Burges, "The MNIST database," <http://yann.lecun.com/exdb/mnist/>
26. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).
27. X. Xu et al., "Imaging of objects through a thin scattering layer using a spectrally and spatially separated reference," *Opt. Express* **26**(12), 15073–15083 (2018).

28. W. Harm et al., “Lensless imaging through thin diffusive media,” *Opt. Express* **22**(18), 22146–22156 (2014).
29. T. Tsukada, T. Nishijima, and W. Watanabe, “Image reconstruction of objects through diffusers with different diffusion angles,” *Proc. SPIE* **11925**, 119250K (2021).

**Takumi Tsukada** received his BS degree from Ritsumeikan University in 2021. He is a first-year master’s student at the Graduate School of Science and Engineering, Ritsumeikan University. His research interests include computational imaging and deep learning.

**Wataru Watanabe** is a professor at Ritsumeikan University. He received his BS, MS, and DEng degrees from Osaka University in 1994, 1996, and 1999, respectively. During 1999–2006, he was an assistant professor at the Graduate School of Engineering, Osaka University. During 2006–2013, he was a researcher at the National Institute of Advanced Industrial Science and Technology (AIST). He joined Ritsumeikan University in 2013 as a professor. He is the author of more than 50 journal papers. His current research interests include ultrafast laser micromachining and biomedical optics. He is a member of SPIE.