



Learning-based method to reconstruct complex targets through scattering medium beyond the memory effect

ENLAI GUO, SHUO ZHU, YAN SUN, LIANFA BAI, CHAO ZUO,  AND JING HAN*

Jiangsu Key Laboratory of Spectral Imaging and Intelligent Sense, Nanjing University of Science and Technology, Nanjing 210094, China

*eohj@njust.edu.cn

Abstract: Strong scattering medium brings great difficulties to image objects. Optical memory effect makes it possible to image through strong random scattering medium in a limited angle field-of-view (FOV). The limitation of FOV results in a limited optical memory effect range, which prevents the optical memory effect to be applied to real imaging applications. In this paper, a kind of practical convolutional neural network called PDSNet (Pragmatic De-scatter ConvNet) is constructed to image objects hidden behind different scattering media. The proposed method can expand at least 40 times of the optical memory effect range with a average PSNR above 24dB, and enable to image complex objects in real time, even for objects with untrained scales. The provided experiments can verify its accurateness and efficiency.

© 2020 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Scattering medium exists generally in biological tissues and is the main interference source in the field of astronomical imaging [1]. Many new imaging methods have been proposed to realize imaging through heavily disordered medium. The typical ones include optical coherence tomography [2,3], wavefront modulation [4,5], image reconstruction based on transmission matrix [6,7] or based on point spread function (PSF) [8–10]. Based on optical memory effect (OME), *Katz et al.* proposed speckle correlation imaging technology [1,11]. Different from the above methods, this method provides powerful capability to image through scattering media deeply, and can be applied to dynamic scattering media, meanwhile, no additional reference source is needed.

Due to the existence of OME, scattering medium can be regarded as a linear system within a settled FOV. Based on this principle, two kinds of algorithms have emerged: the first kind is based on the deconvolution of the PSF, which requires the measurement of the system PSF in advance and belongs to the invasive method. The second type is speckle correlation imaging technology [12,13], which uses the autocorrelation to concatenate speckle pattern with the original target directly, and finally utilizes a phase recovery algorithm to reconstruct the target hidden behind a diffuser. At present, typical phase recovery algorithms include HIO [14], ADMM-based [15], prGAMP [16], BM3D-prGAMP [17], etc. The FOV of an imaging system is strictly limited by OME and is inversely proportional to the effective thickness of the scattering medium, which limits the anti-scattering ability of various OME-based algorithms. The FOV of OME has been successfully expanded through several novel techniques, but most of them belong to invasive method, which relies on priori knowledge [18–20]. *Gardner et al.* introduced ptychography to image non-sparse objects by modulating the lighting source, which has the ability to expand FOV [21]. As far as we know, the best non-invasive method to image beyond the OME range is a double loop iterative algorithm proposed by *Wang et al.* [22], which can restore two targets whose total size is beyond OME, respectively. And the proposed algorithm can expand at least

three times of OME scope shown by experiments without prior information, but there are still some constraints for the target to be restored, e.g. the target must be in regular shape and in two independent OME regions. Since double iterative optimization is required, the algorithm is time-consuming, and it needs 14400s to reconstruct a image with a resolution of 300*300 pixels on MATLAB2018b.

To solve the original target distribution from the scattering image is to find the mapping relationship between them. Many scholars have also applied the deep learning algorithm to image through scattering medium in the context of OME, to solve the distribution of the tested targets [23–27]. "IDiffNet" proposed by *Shuai Li et al.* realized the reconstruction of speckle image [23], and discussed in detail the influence of the loss function, training set and other variables on the final reconstruction image quality. *Yunzhe Li et al.* constructed a CNN network with the "one-to-all" reconstruction capability of single optical statistical characteristics [24], which can reconstruct the speckle image generated by untrained ground glass, and the only condition is that untrained scattering medium has the same statistical characteristics as the ground glass used in making the training set. The scattered images of ground glass and single mode fiber are recovered in utilizing U-net realized by *Yang et al.* [25]. A GAN network was built by *Sun et al.* and fat emulsion solution was used to simulate dynamic scattering media [26]. The experiment proves that the network can realize the restoration of dynamic scattering images. *Lyu et al.* constructed a hybrid neural network (HNN) model [27], which is proved that targets larger than the FOV of OME can also be reconstructed. The experiments provided in Ref. [27] all employed a single ground glass as scattering medium, but the reconstruction ability of optical properties of different medium with one training process is not proved to possess. The above works prove that deep learning can be used to reconstruct objects within the OME range through random scattering media and break the OME range of a single scattering medium. At present, quantitative evaluation of the reconstructed image quality and the ability to expand ME range is lacked on learning-based work.

In order to introduce OME-based methods to living biological tissues and other practical applications, the limitation of this kind of methods in FOV is a problem demanding prompt solution. At the same time, the structural complexity of the actual target is very high, so reconstruction ability is demanded to be powerful enough to retrieve the image of hidden objects. In practical application, the scattering medium is complex, and both the scale of the target to be measured and the properties of the scattering medium are complex and changeable. In addition, the reconstructing algorithm should have prominent generalization ability to be adapted to tasks including different diffusers and targets of different scales. Therefore, removing barrier of the limitation of OME on the FOV, good reconstruction ability for complex targets and strong generalization ability, are the basic conditions for imaging through scattering media to be practical.

In this paper, a kind of convolutional neural network named PDSNet (Pragmatic De-scatter ConvNet) is constructed to remove barrier of the limitation of OME on FOV in a data-driven way, combining with the principle of traditional speckle correlation imaging algorithm to guide the design and optimization of the network. PDSNet is a neural network structure suitable for random scales and complex targets. The ability of PDSNet to recover hidden objects is tested experimentally. At least 40 times of optical memory effect range expansion is realized with a average PSNR above 24dB, and the average PSNR of the restored images is above 22dB in the condition of untrained scale. Complex target like human face can be reconstructed by the proposed neural network successfully.

The rest of the paper is organized as follows. Section 2 illustrates the design principles of PDSNet. Section 3 gives experiments. Section 4 concludes this paper.

2. Image reconstruction method based on deep learning

2.1. Basics of image reconstruction

To remove barrier of the constraint of OME on the FOV, solving the original target distribution from the scattering image is the inverse problem corresponding to the imaging process, which can be expressed as

$$O = F^{-1}(M), \quad (1)$$

where O represents the target to be measured, M represents the output result obtained through the optical system, i.e., the scattering image, and F represents the forward model of the imaging process. The inverse problem is essentially an optimization problem, which can be expressed as

$$O = \operatorname{argmin} \|F(O) - M\|^2 + \lambda R(O), \quad (2)$$

where $R(O)$ is the regularization term. Starting with the speckle image, the original target to be restored is solved, which is the optimization problem represented by the above equation.

The light signal loaded with target information finally reaches to the camera after being modulated by the scattering medium. The process can be expressed as

$$I = \sum_{i=1}^n (O_i * S_i), \quad (3)$$

where n represents the target distributed in n different OME ranges, O_i is the i th imaging target, and S_i is the corresponding PSF. According to Ref. [1], the autocorrelation of PSF is an impulse function. When the target is completely in the range of OME, autocorrelation is performed on the speckle image as

$$I \star I \approx O \star O. \quad (4)$$

It can be inferred that, if the target is distributed in the corresponding range of n optical memories, the autocorrelation distribution of the speckle image can be expressed as

$$I \star I \approx \sum_{i=1}^n (O_i \star O_i), \quad (5)$$

which shows that even if the target to be restored crosses several OME ranges at the same time, there is a strong physical constraint between the target and the scattered image. Therefore, the optimization problem corresponding to Eq. 2 should have an optimal solution.

The traditional algorithms used in speckle correlation imaging have various limitations in practical use, e.g. HIO [14] needs multiple iteration cycles. This type of algorithms is highly dependent on the random initial values, so its optimization results often converge to the local optimal solution and the computational process takes a long time. Although improved algorithms such as ADMM-based [15], prGAMP [16] and BM3D-prGAMP [17] can achieve better reconstruction quality than traditional iterative phase recovery algorithms, but they are still time-consuming. Moreover, the reconstruction quality of those algorithms depend to some extent on prior information like target sparsity, which still cannot operate outside the range of memory effect.

A large target, which is bigger beyond the range of OME, can be split into several small targets within the range of OME. According to the existing traditional algorithm to solve optimization problems, the solution complexity of the optimization problem corresponding to Eq. 5 increases by an order of magnitude with the increase of target scale. Therefore, it is bound to introduce various constraints or prior information to assist the solution of the optimization problem, which will ultimately be more detrimental to the implementation and application of the technology. In

terms of optimization problem solving, deep learning method is better at using a large amount of data and mining the potential mapping relationship between independent variables and dependent variables compared with the traditional algorithms, which can often make the solution result converge to the global optimal solution with bigger probability.

2.2. Model design

In order to meet the practical needs of imaging through scattering media, a new encoder-decoder structure is designed. Through the skip connection, the local feature enhancement is further achieved, and the network structure is as shown in Fig. 1.

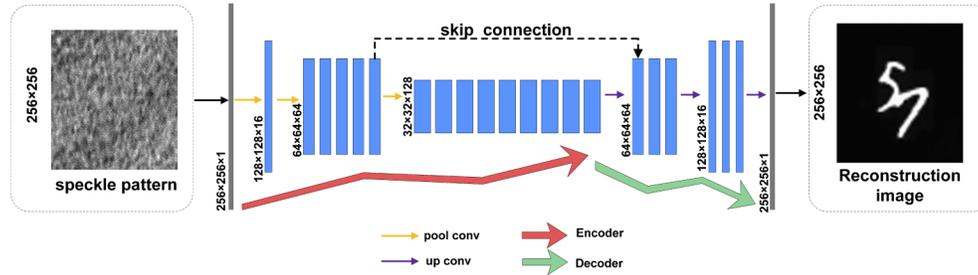


Fig. 1. The structure of the proposed PDSNet.

In solving the scattering problem in the range of OME, the commonly used network structure is U-shaped structure [23–25,28], e.g. U-net [29]. The network structure has a small number of parameters, which, on the one hand, is highly efficient in operation, and on the other hand, it also means that the network has limited ability in reconstructing the hidden objects. Especially, when the complexity of the solving problem increases, it is obvious that the reconstruction quality decreases. At the same time, the increase in the number of parameters will also lead to a decrease in the computational efficiency. Here, we adopt the method of factorized convolution in the chosen basic structure, which is a simple U-shaped structure with only encoder and decoder (not U-net) by using $n*1$ and $1*n$ to replace the usual $n*n$ convolution method. On the premise of ensuring the number of parameters, the computational complexity is reduced and the efficiency of the network computing is improved.

The structure of encoder-decoder in U-shaped network is a classic point-to-point network architecture. However, high-level semantic information is mainly involved in the final prediction, so low-level information that fully represents local details cannot well participate in the final prediction process, resulting in insufficient detail of reconstruction results.

According to the principle of traditional iterative phase recovery algorithm used in speckle correlation imaging technique, the initial image of the first iteration is generated by constraining the output image of the last iteration pixel by pixel. Take HIO algorithm as an example

$$\begin{cases} g_{k+1}(x, y) = g'_k(x, y), & \text{for } (x, y) \in \Gamma \\ g_{k+1}(x, y) = g_k(x, y) - \beta g'_k(x, y), & \text{for } (x, y) \notin \Gamma, \end{cases} \quad (6)$$

where $k < 0$, Γ is the set of points that satisfies the physical constraint, $g_k(x, y)$ is the k -th input, $g'_k(x, y)$ is the k -th output, and β is the feedback control parameter, which is similar to the learning rate in machine learning. Through several iterations, every pixel in the image is made to satisfy the physical constraint as much as possible. The calculation process of iterative projection introduces global constraints to the algorithm. The reconstruction process is the sequence from the local to the global.

During the encoding process, pixel level information evolves into higher level information gradually, and some feature information of speckle pattern is discarded inevitably after downsampling and dropout operation. Then global information is fed into the final prediction process by the decode process. The backward propagation process calculates the gradient information from the semantic-level to the pixel-level, and finally obtains the result in the input layer. The reconstructing process of encoder-decoder structure is also from the local to the global as the traditional speckle correlation imaging technique, which is believed as the reason that encoder-decoder structure can restore object from speckle pattern.

Besides, we hope that the information missed during the encoding process can be more involved in the final prediction process. We use the skip connection method to integrate low-level information with high-level information to retrieve missing feature information, which realizes fully mining of pixel-level information and semantic-level features, and form a new convolutional neural network structure called PDSNet. Meanwhile, the combining high-level information with the low-level information is beneficial to prevent getting trapped in local optimality. The new encoder-decoder structure can integrate high-level semantic information with low-level local details, and improve the deficiency of the simple U-shaped structure in local details reconstruction as proved in Figs. 3 and 7.

Finally, starting from the simple U-shaped structure, PDSNet is formed by adding the design of factorized convolution and skip connection. The parameters number of PDSNet is 2,062,808 parameters (7.87MB), and it is much less than 34,525,889 parameters (131.71MB) of U-net.

3. Experiment

3.1. Experiment on simulation data

Figure 2 shows the experiment setup. In the process of designing convolutional neural network, simulation data is used to test and verify the performance of PDSNet.

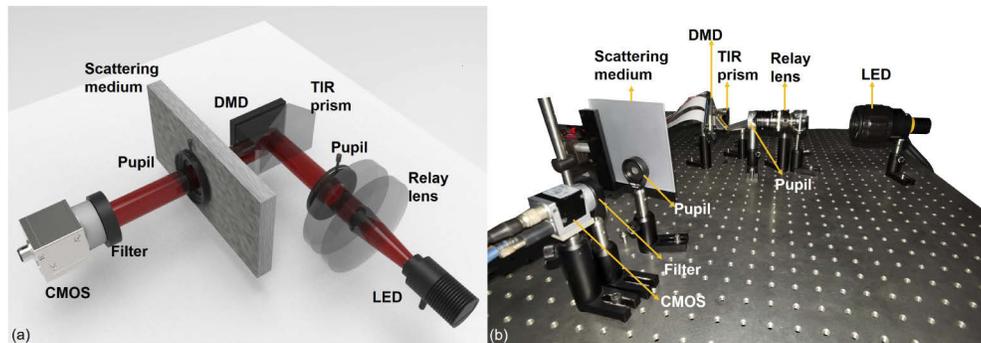


Fig. 2. Experiment setup uses an DMD as the object. (a) experiment setup; (b) the actual optical system.

The light signal of the target is modulated by the scattered medium after transmission in the first free space. After the second free space transmission process, the light signal emitted from the surface of the scattering medium reaches to the camera, and the camera obtains the seemingly disordered light field intensity information. The modulation effect of scattering medium on optical signal can be characterized by circular symmetric complex Gaussian random variables as characteristic matrix in the simulation process, of which variance is employed to represent the statistical characteristic of diffuser [30]. Fresnel diffraction theory is applied to simulate two free space transmissions.

We hope to construct a scattering image reconstruction network that can remove barrier of the OME on the FOV constraint. Therefore, it should be noted that all the simulation data and

experimental data mentioned in this paper are beyond the range of FOV restricted by OME with target in gray scale rather than binary. And the test environment is PyTorch 1.2.0 with RTX 2080Ti with I7-9700K CPU under ubuntu 16.04. Mean Absolute Error (MAE), Structural Similarity Index Measure (SSIM) and Peak Signal to Noise Ratio (PSNR) are used as objective indicators to assess the reconstruction quality. It's important to note that MAE is computed between the normalized images in which the gray value ranges from zero to one.

To demonstrate PDSNet's ability to reconstruct targets under different conditions, six datasets are generated through the simulation. The datasets 1-6 denote *the complex object dataset with 2 characters*, *the complex object dataset with 3 characters*, *the complex object dataset with 4 characters*, *the multi-medium A dataset with same property*, *the multi-medium B dataset with different property*, and *the human face dataset*, respectively. The cardinal targets commonly used in speckle correlation imaging based on deep learning are single handwritten characters [23–28]. In order to improve the complexity of the targets to be restored in the first five datasets, the handwritten characters in MNIST are randomly combined [31]. Numbers of characters to compose different dataset are shown in the third row of Table 1. To be noticed, overlap strategy is utilized in all the datasets to increase the diversity of targets. The variances of characteristic matrixes constructed in *the multi-medium A dataset* are maintained the same value to simulate the process of imaging with different diffusers of the same property. A similar role is played by the different variances of characteristic matrixes constructed in *the multi-medium B dataset*. The targets in *the human face dataset* are derived from FEI Face Databases [32]. The original targets acquired through the conditions shown in Table 1 is utilized as the input of the optic system to generate the corresponding emulational speckle image, which constitutes the simulation datasets together.

Table 1. The simulation datasets are generated according to the following conditions.

	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6
Basic element	MNIST	MNIST	MNIST	MNIST	MNIST	FEI Face
Number of characters	2	3	4	2	2	none
Number of diffusers	1	1	1	10	10	1
Variances of diffusers	One	One	One	Same	Different	One
Training set data (pairs)	3000	3000	30000	30000	30000	350
Test set data (pairs)	5000	5000	5000	5000	5000	40

The original targets of the datasets are employed as the ground truth (GT) of the neural network, while the speckle images are utilized as input images during the experiments. The size of both the GT image and the speckle image is 256*256.

Table 2. Results of Test Set of the Complex Object Datasets (Average MAE, SSIM and PSNR).

Dataset name of untrained targets		Objective Indicators		
		MAE	SSIM	PSNR
Complex object dataset with 2 characters	PDSNet	0.0102	0.9283	29.9946dB
	U-net	0.0126	0.9265	27.3848dB
Complex object dataset with 3 characters	PDSNet	0.0175	0.8951	25.8168dB
	U-net	0.0216	0.8641	23.6850dB
Complex object dataset with 4 characters	PDSNet	0.0227	0.8767	24.1100dB
	U-net	0.0310	0.8304	21.3222dB

To demonstrate PDSNet's ability to reconstruct complex targets, PDSNet is tested on *the complex object dataset with 2 characters*, *3 characters* and *4 characters*. The relationship

between autocorrelation of the original GT and autocorrelation of the speckle image does not meet Eq. 4 as shown in Figs. 3(b) and 3(d), which proves that the target size in the dataset 1-3 at this time is beyond the range of OME. Both datasets are utilized to train PDSNet and U-net, and the reconstruction results of the test set are shown in Figs. 3(e) and 3(f), respectively. As complexity of the targets increases, a downward trend is shown by the objective indicators of both PDSNet and U-net, but even if the targets is generated with 4 characters, the hidden objects also can be reconstructed accurately by PDSNet, as shown in Table 2. And better details in Fig. 3(e) are produced by PDSNet than the images restored by U-net shown in Fig. 3(f). Better reconstruction results are obtained with fewer parameters by designing reasonable structure for PDSNet. The validity of the structure of the PDSNet is further proved.

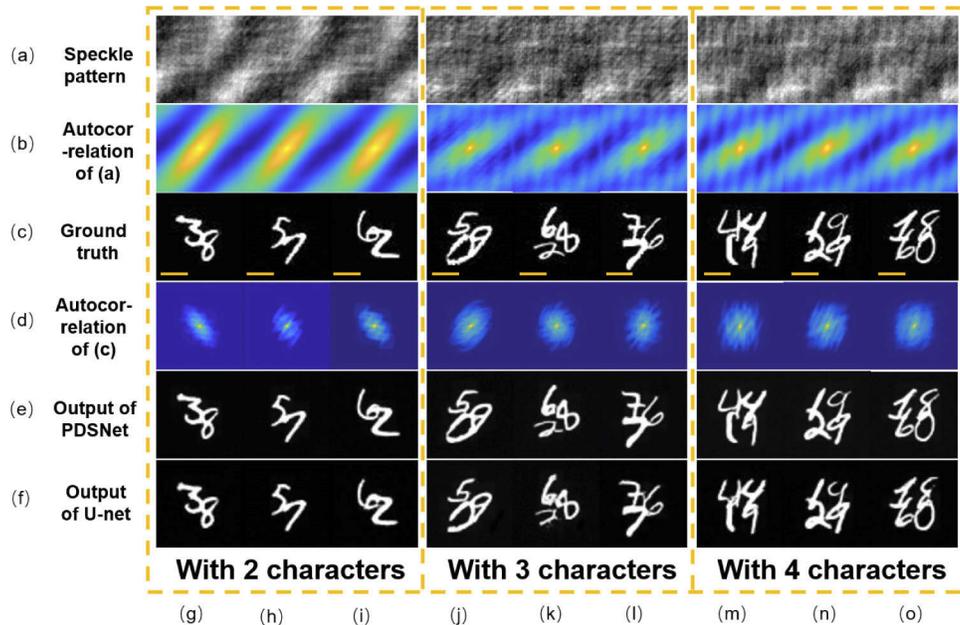


Fig. 3. Test result of *the complex object dataset with 2 characters, 3 characters and 4 characters*: (a) speckle image; (b) autocorrelation of (a); (c) original target; (d) autocorrelation of (c); (e) output image of PDSNet; (f) output image of U-net; (g)–(i) belong to the experiment of *the complex object dataset with 2 characters*; (j)–(l) belong to the experiment of *the complex object dataset with 3 characters*; (m)–(o) belong to the experiment of *the complex object dataset with 4 characters*. Scale bars: 75 pixels.

To further test the reconstruction capability of PDSNet for different scattering media with the same statistical characteristics, PDSNet is tested on *the multi-medium A dataset with same property*. The results of the test set are shown in Figs. 4(d)–4(g). Then, the reconstruction ability of PDSNet to scattering media with different statistical characteristics is tested, and *the multi-medium B dataset with different property* is constructed for this purpose. The test results are shown in Figs. 4(h)–4(k). The average MAE, SSIM and PSNR of *the multi-medium A dataset with same property* and *the multi-medium B dataset with different property* are shown in Table 3. As shown by the objective indicators, the reconstruction complexity of tasks with different scattering media is similar no matter the property of the different media is the same or different. Prominent reconstruction details, both global and local, are obtained for the multi-medium missions.

To further improve the complexity of the targets, *the human face dataset* is employed here. As shown in Figs. 5(c) and 5(d), for the targets as detailed as a human face, the original PDSNet

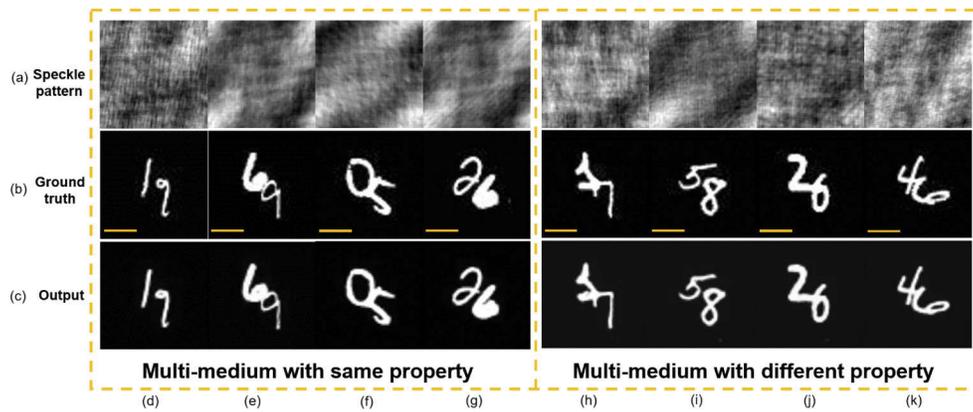


Fig. 4. Test result of *the multi-medium A dataset with same property* and *the multi-medium B dataset with different property*: (a) speckle image; (b) original target; (c) output image of PDSNet; (d)–(g) belong to the experiment of *the multi-medium A dataset with same property*; (h)–(k) belong to the experiment of *the multi-medium B dataset with different property*. Scale bars: 75 pixels.

Table 3. Results of the Test Set of Two Multi-medium Datasets (Average MAE, SSIM and PSNR).

Dataset name of untrained targets	Objective Indicators		
	MAE	SSIM	PSNR
The multi-medium A dataset with same property	0.0121	0.9235	27.8044dB
The multi-medium B dataset with different property	0.0121	0.9174	27.7602dB

is not capable of reconstructing the image because of the limitation of its parameter amount. Resnet-50 is utilized as the backbone network of PDSNet to enlarge the parameter amount formed PDSNet-L. The number of parameters increases from 2,062,808 parameters (7.87MB) in original PDSNet to 58,492,455 parameters (223.13MB) in PDSNet-L. The same dataset is employed to test PDSNet-L. As shown in Figs. 5(e)–5(h), better results are achieved by PDSNet-L. Although PDSNet is sufficient to reconstruct handwritten character, larger parameter amount is to the benefit of restoring targets with more details like human faces shown by the contrast experiment between PDSNet and PDSNet-L. As for the tasks with complex details, parameter amount of PDSNet can be enlarged by altering the backbone network or increasing numbers of layers to enhance reconstruct ability of the original PDSNet in recovering the hidden objects.

3.2. Experiment on real system

In this section, data from the actual optical system is collected, and several trainings and tests are conducted. Fig. 2(b) is the setup of the optical system. The light source is a LED with a central wavelength of 625nm (Thorlabs, M625L4). Ground glasses (Thorlabs, DG100X100-220) are used as the scattering media and placed between the target to be measured and the camera plane. The final camera (Balsar acA1920-155um) is used for collecting speckle images with information of the target to be restored. The distance between the target surface and the scattering medium is 25cm, and the distance between the scattering medium and the camera is 8cm. In the experiment, the target to be measured is generated by a DMD (resolution 1024*768pixel, mirror element size 13.68 $\mu\text{m}/\text{pixel}$).

First of all, the range of OME is tested. In order to test the range of shift-invariant, a series of speckle is collected while the horizontal displacement of the point object on the target surface is achieved. The cross-correlation coefficient between the speckle patterns and the PSF of the

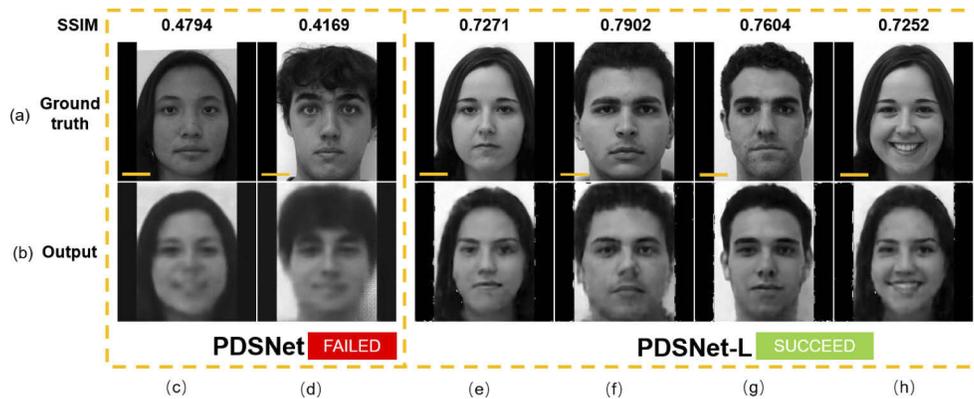


Fig. 5. Test result of *the human face dataset*: (a) original target; (b) output image; (c) and (d) results of PDSNet; (e)–(h) results of PDSNet-L. Scale bars: 50 pixels.

system, is calculated. 0.5 is chosen as the threshold value of the cross-correlation coefficient to determine the range of ME [19,20], and we define δp as the offset pixel number of image plane, which is 56 as shown in Fig. 6. The ME range of the system can be calculated by $2 * p * \delta p / \beta$ [19], β is the system magnification and p is the pixel size of the camera, which equals to $5.86 \mu\text{m}$. It can be concluded that the ME range of the system is $150 * 150$ pixels on DMD.

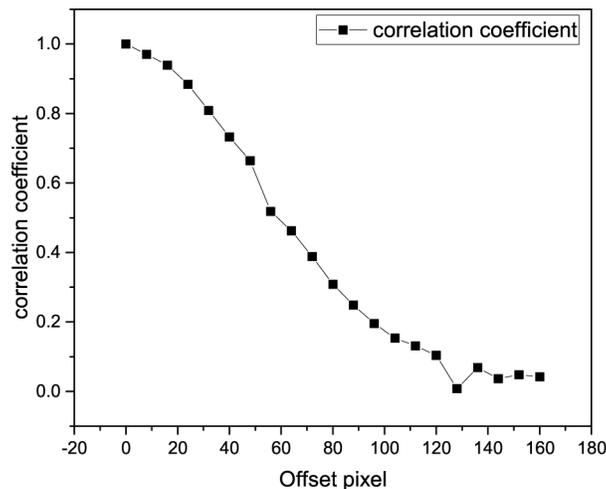


Fig. 6. The curve of cross-correlation coefficient to measure the ME range of the ground glass used in the following experiments.

3.3. Module function analysis

In this section, the need for a skip connection is first verified. In order to demonstrate the necessity of the skip connection on the condition of targets size out of the range of OME, the handwritten characters in MNIST are randomly combined in pairs with overlap to generate 8000 images, which are utilized as the targets displayed on DMD. The training set and the test set contain 7500 pairs data and 500 pairs data, respectively. It can be estimated from Fig. 7 that the autocorrelation distribution of the speckle images is quite different from that of the original GT image, which does not confirm to Eq. 4, that is, the targets have exceeded the range of OME.

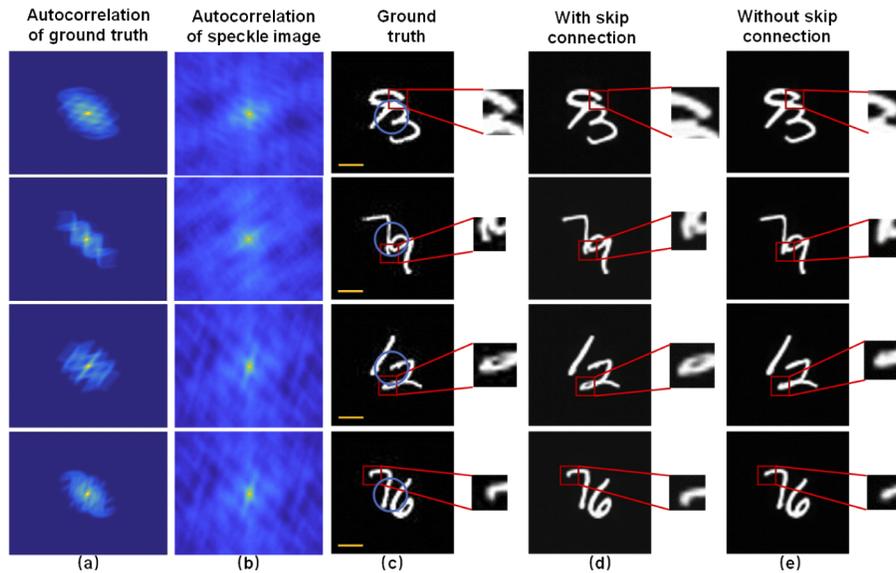


Fig. 7. Test result of PDSNet with and without skip connection: (a) autocorrelation of the original targets; (b) autocorrelation of the speckle images (c) original targets, and the blue circles are used to show the range of OME; (d) output images of PDSNet with skip connection; (e) output images of PDSNet without skip connection. Scale bars: 50 pixels.

As can be seen from the comparison in Fig. 7, the reconstructed results of PDSNet without skip connection basically restore the original target distribution. However, compared with GT, the detailed part of the target is not restored with higher accuracy. It can be found that the reconstruction quality of the detailed part has been significantly improved relatively by skip connection.

In the downsampling process of encoder, PDSNet without skip connection gradually focuses on high-level semantic information, corresponding to the overall shape distribution of the target has some defects in the reconstruction ability of local details. The skip connections, which can introduce low-level details into the final prediction process, are of great necessity.

3.4. Reconstruct ability test

To test PDSNet's ability to reconstruct targets in untrained size, a *multi-scale dataset* is acquired. The experimental result show that the range of OME is 150*150 pixels on DMD. Sizes of targets displayed on DMD are 300*300 pixels, 375*375 pixels, 420*420 pixels, 450*450 pixels, which are all beyond the range of OME. Targets in size of 300*300 pixels, 375*375 pixels, 450*450 pixels are used as the training set, when targets in size of 420*420 pixels are applied as the test set keep untrained. Although the target size in the dataset is different, target in GT images are maintained the same size. 22500 data pairs are employed as the training set, while 7500 data pairs are served as the test set. Test results of untrained targets are shown in Fig. 8. Even if the scale of targets in the test set is untrained, satisfactory details are produced by the PDSNet, and this conclusion is also sustained by Table 4. To further demonstrate the reconstruction ability of PDSNet, untrained targets with trained scales are also tested, and better objective indicators are obtained shown in Table 4, which prove the effectiveness of PDSNet for multi-scale tasks.

To test PDSNet's ability to reconstruct targets through several scattering media, 4 ground glasses are employed as scattering media to collect speckle images for *the multi-diffuser dataset*. Targets used here are also acquired by combining two handwritten characters in MNIST with

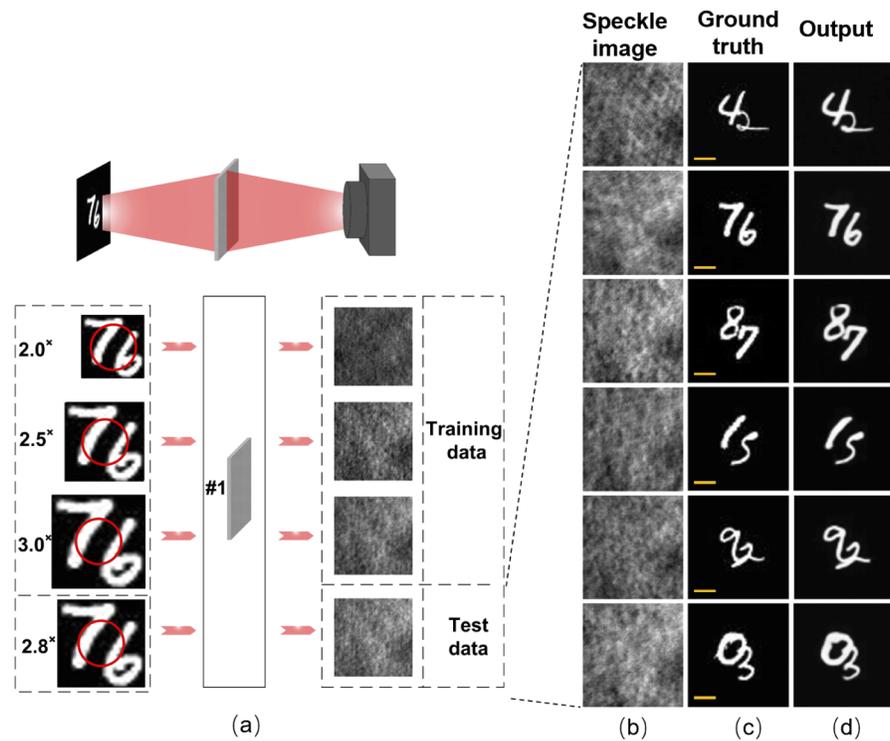


Fig. 8. Test PDSNet's ability to reconstruct targets in untrained size: (a) speckle patterns from different size of targets, and the red circles are used to show the range of OME; (b) speckle images; (c) original targets; (d) output images of PDSNet. Scale bars: 50 pixels.

Table 4. Results of the Test Set of Two Multi-medium Datasets (Average MAE, SSIM and PSNR).

Dataset name of untrained targets	Objective Indicators		
	MAE	SSIM	PSNR
The ability to reconstruct untrained scale	0.0207	0.8831	22.7068dB
The ability to reconstruct trained scale	0.0129	0.9091	27.4222dB
The ability to reconstruct targets through different ground glasses	0.0107	0.9249	29.1512dB

overlap. 30000 data pairs are employed as the training set, while 2000 data pairs are used as the test set. Test results of untrained targets are shown in Fig. 9. With a single training, PDSNet can restore speckle images generated through several different scattering media with rich details. Objective indicators show that multi-diffuser task is less complex than multi-scale task.

In order to test the limitation of the proposed PDSNet, several additional experiments are conducted, the objects employed in those experiments are the same as *the complex object dataset with 2 characters*. Objects and speckle patterns on each scale are trained and tested, respectively. As shown in Fig. 10, the PSNR above 27dB is achieved by PDSNet while the scale of the objects is between 3.3 times-13 times of ME range. As well as the scale of the objects increases beyond 13 times of ME range, a downward trend is emerged. When the scale of the objects is beyond 20 times of ME range, there is a significant decline. Even if the average PSNR of the test dataset is 24.7628dB on 40 times of ME range, the reconstructed results are still satisfactory.

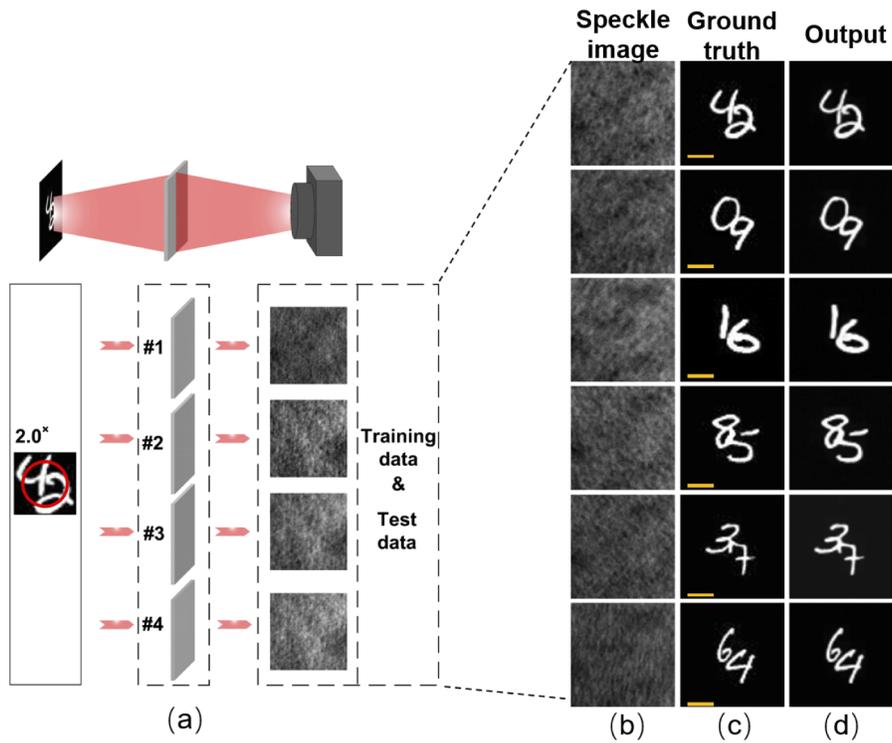


Fig. 9. Test PDSNet’s ability to reconstruct targets through several scattering media: (a) speckle patterns through different diffusers, and the red circle is used to show the range of OME; (b) speckle images; (c) original targets; (d) output images of PDSNet. Scale bars: 50 pixels.

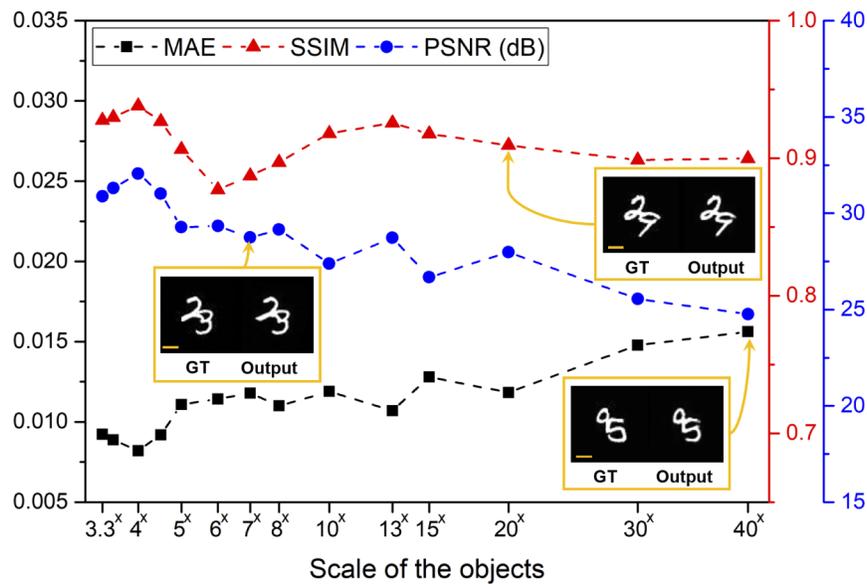


Fig. 10. Test results of *the complex object dataset with 2 characters* under different scales. Scale bars: 50 pixels (see Visualization 1).

16000 scattering patterns are employed only to test time-effectiveness of PDSNet, the conclusion is drawn that the average FPS (frames per second) of PDSNet is 105, which is beyond the standard of real time.

4. Conclusion

Inspired by the traditional algorithm based on OME, this paper proposes a novel convolutional neural network called PDSNet for the reconstruction of scattered images. The experiments show that scattered images can be restored accurately by PDSNet in real-time, and prominent reconstruction ability is realized for the target with complex structure. Numbers of layers of PDSNet can be increased to improve number of parameters to deal with targets with more details.

At the same time, PDSNet has a prominent generalization to reconstruct targets with untrained scales. For targets beyond the range of OME, the speckle images corresponding to targets at random scales can be reconstructed in one training, achieving "one-to-all" on the target scale. This ability enables us not to strictly require the target to be restored to have a uniform size, which is more conducive to the practical application of this technology in the future.

As a supervised learning method, PDSNet also needs to measure data in advance, but the prominent reconstruction ability of the network enables us to measure scattering media more flexible in practical use, even if objects of unmeasured scales can be recovered through data of several measured scales. At the same time, our method does not need to modulate the light source actively like Ref. [21]. It is a reconstruction method based on passive data acquisition, which can certainly reduce the difficulty of data acquisition. The test results of both simulation and experiment show that the proposed network also has prominent reconstruction potential for scattering media with different statistical characteristics. This is also a necessary ability for the practical application of deep learning to the reconstruction of scattering images of living biological tissues.

In the future, we will try to conduct experiments on living biological tissues that meet strong scattering conditions, hoping to be able to apply PDSNet to practical systems in fields like medical imaging and astronomical imaging.

Funding

National Natural Science Foundation of China (61727802, 61971227); Jiangsu Provincial Key Research and Development Program (BE2018126).

Acknowledgments

We thank Dr. Dongliang Zheng for his suggestions for the manuscript, and thank Yingjie Shi, Jie Gu and Qianying Cui for technical supports. We also thank the anonymous reviewers for their helpful comments to improve our paper.

Disclosures

The authors declare no conflicts of interest.

References

1. O. Katz, P. Heidmann, M. Fink, and S. Gigan, "Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations," *Nat. Photonics* **8**(10), 784–790 (2014).
2. Y. Mao, C. Fluoraru, S. Chang, D. P. Popescu, and M. G. Sowa, "High-quality tissue imaging using a catheter-based swept-source optical coherence tomography systems with an integrated semiconductor optical amplifier," *IEEE Trans. Instrum. Meas.* **60**(10), 3376–3383 (2011).
3. D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, and J. G. Fujimoto, "Optical coherence tomography," *Science* **254**(5035), 1178–1181 (1991).

4. 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**(1), 7276 (2015).
5. M. Nixon, O. Katz, E. Small, Y. Bromberg, A. A. Friesem, Y. Silberberg, and N. Davidson, "Real-time wavefront shaping through scattering media by all-optical feedback," *Nat. Photonics* **7**(11), 919–924 (2013).
6. M. Kim, W. Choi, Y. Choi, C. Yoon, and W. Choi, "Transmission matrix of a scattering medium and its applications in biophotonics," *Opt. Express* **23**(10), 12648–12668 (2015).
7. 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**(9), 11898–11911 (2015).
8. D. Lu, M. Liao, W. He, Z. Cai, and X. Peng, "Imaging dynamic objects hidden behind scattering medium by retrieving the point spread function," in *Speckle 2018: VII International Conference on Speckle Metrology*, vol. 10834 (International Society for Optics and Photonics, 2018), pp. 578–581.
9. 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. Innovative Opt. Health Sci.* **12**(04), 1930005 (2019).
10. 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**(12), 15073–15083 (2018).
11. 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**(7423), 232–234 (2012).
12. H. Li, T. Wu, J. Liu, C. Gong, and X. Shao, "Simulation and experimental verification for imaging of gray-scale objects through scattering layers," *Appl. Opt.* **55**(34), 9731–9737 (2016).
13. J. Xie, X. Xie, Y. Gao, X. Xu, Y. Liu, and X. Yu, "Depth detection capability and ultra-large depth of field in imaging through a thin scattering layer," *J. Opt.* **21**(8), 085606 (2019).
14. J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.* **21**(15), 2758–2769 (1982).
15. J. Chang and G. Wetzstein, "Single-shot speckle correlation fluorescence microscopy in thick scattering tissue with image reconstruction priors," *J. Biophotonics* **11**(3), e201700224 (2018).
16. P. Schniter and S. Rangan, "Compressive phase retrieval via generalized approximate message passing," *IEEE Trans. Signal Process.* **63**(4), 1043–1055 (2015).
17. C. A. Metzler, A. Maleki, and R. G. Baraniuk, "Bm3d-prgamp: Compressive phase retrieval based on bm3d denoising," in *2016 IEEE International Conference on Image Processing (ICIP)*, (IEEE, 2016), pp. 2504–2508.
18. L. Li, Q. Li, S. Sun, H.-Z. Lin, W.-T. Liu, and P.-X. Chen, "Imaging through scattering layers exceeding memory effect range with spatial-correlation-achieved point-spread-function," *Opt. Lett.* **43**(8), 1670–1673 (2018).
19. 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).
20. 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**(26), 7533–7538 (2018).
21. D. F. Gardner, S. Divitt, and A. T. Watnik, "Ptychographic imaging of incoherently illuminated extended objects using speckle correlations," *Appl. Opt.* **58**(13), 3564–3569 (2019).
22. X. Wang, X. Jin, J. Li, X. Lian, X. Ji, and Q. Dai, "Prior-information-free single-shot scattering imaging beyond the memory effect," *Opt. Lett.* **44**(6), 1423–1426 (2019).
23. S. Li, M. Deng, J. Lee, A. Sinha, and G. Barbastathis, "Imaging through glass diffusers using densely connected convolutional networks," *Optica* **5**(7), 803–813 (2018).
24. 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).
25. M. Yang, Z.-H. Liu, Z.-D. Cheng, J.-S. Xu, C.-F. Li, and G.-C. Guo, "Deep hybrid scattering image learning," *J. Phys. D: Appl. Phys.* **52**(11), 115105 (2019).
26. Y. Sun, J. Shi, L. Sun, J. Fan, and G. Zeng, "Image reconstruction through dynamic scattering media based on deep learning," *Opt. Express* **27**(11), 16032–16046 (2019).
27. M. Lyu, H. Wang, G. Li, S. Zheng, and G. Situ, "Learning-based lensless imaging through optically thick scattering media," *Adv. Photonics* **1**(03), 1 (2019).
28. N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, "Learning to see through multimode fibers," *Optica* **5**(8), 960–966 (2018).
29. 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.
30. J. W. Goodman, *Speckle phenomena in optics: theory and applications* (Roberts and Company Publishers, 2007).
31. Y. LeCun, C. Cortes, and C. J. C. Burges, "THE MNIST DATABASE of handwritten digits," <http://yann.lecun.com/exdb/mnist/>.
32. C. E. Thomaz, "FEI Face Database," <https://fei.edu.br/~cet/facedatabase.html>.