Sub-Nyquist computational ghost imaging with deep learning

Item Type	Article
Authors	Wu, Heng; Wang, Ruizhou; Zhao, Genping; Xiao, Huapan; Wang, Daodang; Liang, Jian; Tian, Xiaobo; Cheng, Lianglun; Zhang, Xianmin
Citation	Heng Wu, Ruizhou Wang, Genping Zhao, Huapan Xiao, Daodang Wang, Jian Liang, Xiaobo Tian, Lianglun Cheng, and Xianmin Zhang, "Sub-Nyquist computational ghost imaging with deep learning," Opt. Express 28, 3846-3853 (2020)
DOI	10.1364/OE.386976
Publisher	OPTICAL SOC AMER
Journal	OPTICS EXPRESS
Rights	Copyright © 2020 Optical Society of America under the terms of the OSA Open Access Publishing Agreement.
Download date	27/08/2022 18:56:36
Item License	http://rightsstatements.org/vocab/InC/1.0/
Version	Final published version
Link to Item	http://hdl.handle.net/10150/643298

Sub-Nyquist computational ghost imaging with deep learning

HENG WU,[1,2,3] ⬤ RUIZHOU WANG,[1,3,4,*] GENPING ZHAO,[1,5] HUAPAN XIAO,[2] DAODANG WANG,[2] JIAN LIANG,[2] XIAOBO TIAN,[2] LIANGLUN CHENG,[1,5] AND XIANMIN ZHANG[3]

[1]*Guangdong Provincial Key Laboratory of Cyber-Physical System, School of Automation, Guangdong University of Technology, Guangzhou 510006, China*
[2]*College of Optical Sciences, University of Arizona, Tucson, AZ 85721, USA*
[3]*Guangdong Key Laboratory of Precision Equipment and Manufacturing Technology, School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510640, China*
[4]*School of Electro-mechanical Engineering, Guangdong University of Technology, Guangzhou 510006, China*
[5]*School of Computer, Guangdong University of Technology, Guangzhou 510006, China*
wangrz@gdut.edu.cn

Abstract: We propose a deep learning computational ghost imaging (CGI) scheme to achieve sub-Nyquist and high-quality image reconstruction. Unlike the second-order-correlation CGI and compressive-sensing CGI, which use lots of illumination patterns and a one-dimensional (1-D) light intensity sequence (LIS) for image reconstruction, a deep neural network (DAttNet) is proposed to restore the target image only using the 1-D LIS. The DAttNet is trained with simulation data and retrieves the target image from experimental data. The experimental results indicate that the proposed scheme can provide high-quality images with a sub-Nyquist sampling ratio and performs better than the conventional and compressive-sensing CGI methods in sub-Nyquist sampling ratio conditions (e.g., 5.45%). The proposed scheme has potential practical applications in underwater, real-time and dynamic CGI.

1. Introduction

Ghost imaging (GI), as an alternative to traditional imaging techniques, such as CCD and CMOS cameras, restores an object image nonlocally by a second-order correlation algorithm [1–6]. In 1995, the first GI experiment was achieved by Pittman *et al.* with entangled photon pairs [7]. Since then many improved GI schemes have been developed, for instance, iterative denoising GI [8], scalar-matrix-structured GI [9], compressive sensing (CS) GI [10,11], differential GI [12], and Hadamard GI [13,14]. Conventional GI systems include two light paths, the object and reference paths. However, in 2008, Shapiro proposed a single-path GI which only has an object path, and he named it computational GI (CGI) [15]. Over the past few years, although GI has been applied in various fields, such as THz imaging [16], X-ray imaging [17], turbulence imaging [18], remote sensing [19] and multispectral imaging [20], the enormous signal acquisitions (measurements) prevent GI developing into a more actual technique.

Recently, various methods have been proposed to reduce the signal acquisitions and improve the imaging performance [21–27]. One of the widely used methods is to use orthonormal patterns (basis) to reduce the sampling ratio (SR) and illumination patterns. Here, the SR is defined as $\eta = A/F$, where A and F represent the measurement number and total pixel of an image, respectively. For instance, Sun *et al.* presented an alternative CS method by optimizing the measurement order of the Hadamard basis with a "Russian Dolls" ordering method [21]. The method performed as well as CS techniques in low SR condition. Luo *et al.* introduced a post-processing method which used Gram-Schmidt process to orthonormalize the patterns [22].

Although the illumination patterns were reduced, the method is sensitive to noise. Not long before, Yu proposed a new CGI approach based on CS [23]. By using a "Cake-Cutting" Hadamard basis optimization technique, a super sub-Nyquist sampling is realized and the acquisition time is significantly decreased. Another lately reported method is to use deep learning (DL) to decrease the SR [24–27]. In [24–26], three deep-learning-based methods are developed to lower the SR of CGI, where the input image is obtained by conventional GI. Most recently, Wang *et al.* demonstrated a neural network for CGI which reconstructs target images directly from the one-dimensional (1-D) bucket signals without the need of the illumination patterns [27]. The neural network includes 4 independent paths and lots of residual blocks, which can perfectly reconstruct an image with a SR of 6.25%. DL GI possesses many advantages over the conventional and CS GI, such as, without using illumination patterns, realizing high-quality imaging with a lower SR, imaging through scattering medium [28], etc. However, retrieving high-quality ghost images from the 1-D bucket signals with a sub-Nyquist SR (or fewer measurements) remains a challenge for the existing DL GI schemes.

In this paper, we propose a deep learning CGI scheme which can retrieve high-quality target images with a sub-Nyquist SR. We develop a deep neural network (DAttNet) to recover the target image directly from the 1-D light intensity sequence (LIS) in sub-Nyquist SR conditions. Compared with [27], DAttNet has a simpler architecture of the neural network because only one path is used. Besides, [27] used residual blocks in the network, but DAttNet utilizes dense blocks and skip connections which has a better generalization capability [28]. Moreover, the proposed scheme can recover a clear image with a lower SR (5.45%) than [27] (6.25%). The effectiveness of the proposed scheme is experimentally verified by using two types of targets (binary and grayscale). The image quality of the proposed scheme is evaluated by comparing the results with four state-of-the-art CGI methods.

2. Method

The experimental setup of the proposed scheme is shown in Fig. 1. The imaging system is composed of a digital light projector (DLP), test target, zoom lens (ZL), charge coupled device (CCD) and personal computer (PC). The ZL (focal length $f \in [0.2\ cm,\ \infty)$) is mounted before the CCD (Blackfly S USB3 CCD). The pattern projection of the DLP (Texas Instruments, DLP3010EVM-LC, 1280×720 pixels) and the light intensity collection of the CCD are controlled by a PC (Intel Core i7-8700K, 32 GB Memory). The distance between the test target and the DLP is 34.50 cm, and the distance between the CCD plane and the test target is 38.20 cm. The DLP projects the Hadamard pattern $P_k = P_k(x, y)$ onto the target. The reflective light is converged by a ZL and the light intensity I_k is recorded by a CCD, which is given by

$$I_k = \iint T(x, y) P_k(x, y) dx dy, \tag{1}$$

where $T(x, y)$ is the object function, $k = 1, 2, \cdots, K$, and K is the total number of measurement.
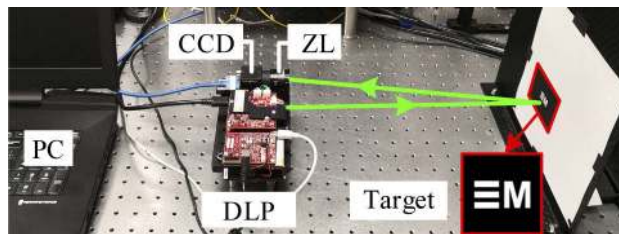

Fig. 1. Experimental setup. The target is printed on a white paper.

Based on the GI theory, the object image $O(x, y)$ can be recovered by [2–5,7],

$$O(x, y) = \frac{1}{K} \sum_{k=1}^{K} (I_k - \langle I \rangle) P_k(x, y), \tag{2}$$

where $I = [I_1, I_2, \cdots, I_K]$ is the LIS, and $\langle \cdot \rangle$ denotes the assemble average. Here, $P = [P_1, P_2, \cdots, P_K]$ is the pattern sequence.

With the pattern sequence P and LIS I, the target image $T(x, y)$ can also be restored by the CS methods [10,11,23].

Unlike the second-order correlation algorithm [2–5,7] and CS methods [10,11,23], the proposed scheme can recover target images directly from the 1-D LIS I. The mathematical model of the proposed scheme can be expressed as an implicit function,

$$\hat{T} = \hat{T}(x, y) = \Re(I; \Psi), \tag{3}$$

where Ψ represents the model parameters, and $\Re()$ denotes an implicit function that establishes a mapping relation between the input LIS I and the output image $\hat{T}$. Since it is difficult to figure out the explicit form of $\Re()$ and find the inverse solution of Eq. (3), we propose a DAttNet to solve the inverse mapping of Eq. (3) with a set of training data $\{(B_n; G_n)\}_{n=1}^{N}$. Here, N is the number of dataset, G_n is the nth ground truth (GT), $B_n = [I_1^n, I_2^n, \cdots, I_M^n]$ is the corresponding LIS, and M is the measurement number.

The architecture of DAttNet is illustrated in Fig. 2, which is inspired by the [27], IDiffNet [28], U-Net [29], Attention U-Net [30] and DenseNet [31]. The first layer is an input layer which inputs a normalized LIS I recorded by the CCD. Since we focus on the sub-Nyquist CGI whose SR is smaller than 10%, thus the maximum size of the input layer is $M = 16384 \times 10\% \approx 1638$. The second and third layers are fully connected layers whose sizes are 4096×1 and 16384×1, respectively. The fourth layer is a vector-to-matrix layer which reshapes the vector 16384×1 to a matrix (temporary image, size 128×128 pixels). Followed the fourth layer is a deep convolutional neural network (CNN) which is designed to recover the target image. In the CNN, we use six
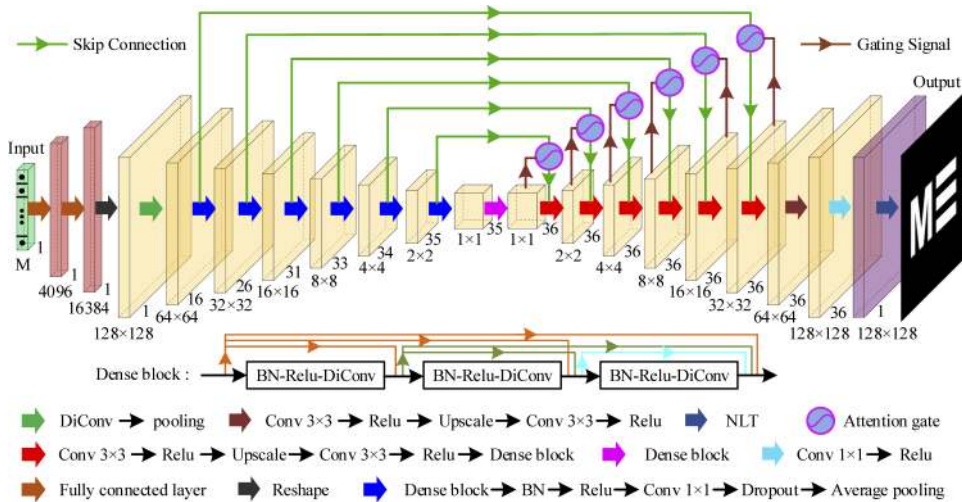

Fig. 2. The architecture of the proposed DAttNet. BN, batch normalization; DiConv, dilated convolution; Conv 3×3, convolution with filter size 3×3; Conv 1×1, convolution with filter size 1×1; Dropout, dropout rate is 0.05; ReLU, rectified linear unit; NLT, normalization and linear transformation; Average pooling and pooling, stride (2, 2); Upscale, factor 2.
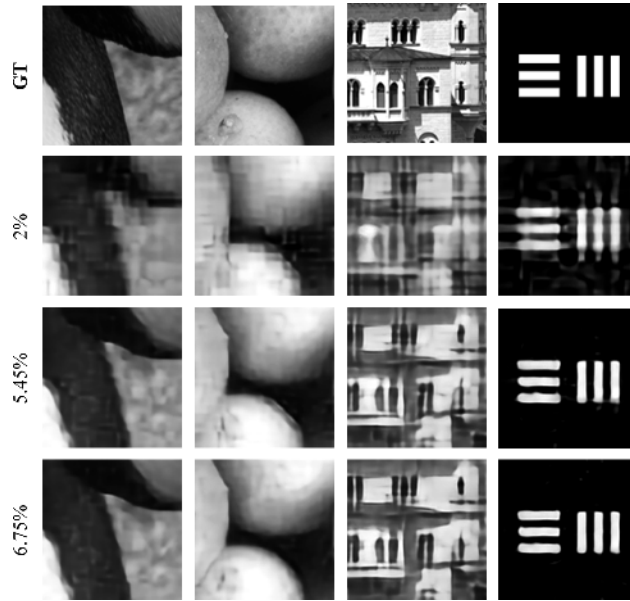
Fig. 3. Test results of the trained model with the SRs of 2%, 5.45% and 6.75%.

attention gates (AG) to enhance the salient features which pass through the skip connections, and eliminate the irrelevant and noisy responses in skip connections [30,31]. Besides, the dilated convolutions (DiConv, dilation rate 2, filter size 5×5) are used for convolution calculation. Additionally, dropout layers are used to prevent overfitting [32] and batch-normalization (BN) layers are added to speed up the convergence of loss function [33]. The last layer is a normalization and linear transformation layer which is used to normalize the output matrix and linearly transform it into an image. After the temporary image passes through the CNN, a ghost image is restored.

3.   Experimental results and analysis

We train the DAttNet in the High-Performance Computing of the University of Arizona with Pytorch (Python version 3.6). The training dataset includes 200 011 GT-LIS pairs. Here, GT (128×128 pixels) is collected from DIV2K [34] and COCO (2017 Train) [35] databases. LIS I is generated by calculating the light intensity I_k, where the Hadamard basis is sorted by the "Cake-Cutting" method [23]. The measurement number is $M = 1638$. Note that 1638 independent Hadamard patterns are used in the experiment and the length of LIS is equal to the measurement number. In the training phase, the ADAM algorithm [36] is utilized to optimize the following loss function,

$$L(\Psi) = \frac{1}{2N} \sum_{n=1}^{N} ||\Re(B_n; \Psi) - G_n||^2, \tag{4}$$

where the mini-batch size and training epochs (TE) are set as 90 and 200, respectively. The learning rate (LR) is set as follows. If TE $\leq$ 50, LR $= 10^{-3}$, else if $50 <$ TE ≤ 60, LR $= 10^{-4}$, else LR $= 10^{-6}$.

The trained model is first tested by the simulated LIS which is obtained by the GT shown in the first row of Fig. 3. It is important to note that the LIS generation method is the same as that in preparing the training dataset. Different sampling ratios (SRs) are adopted to generate the LIS. Here the SR is defined as $\eta = M/16384$. The reconstructed images are shown in the 2-4 rows

of Fig. 3. We find that the quality of recovered images gets better as the increasing of SR, and high-quality images are obtained when SR is equal to 5.45%.

The trained model is then used to restored images from the LIS which is obtained by the experimental setup in Fig. 1. Here, the LIS generation method is also the same as that in preparing the training dataset. We verify the proposed scheme (DL) using different SRs. The experimental results of the proposed scheme (DL) are compared with those of four state-of-the-art methods, fast Walsh-Hadamard transform CGI (FWHT) [14], "Russian dolls" CGI (RD) [22], compressive sensing RD (CSRD), and "Cake-Cutting" CGI (CC) [23].
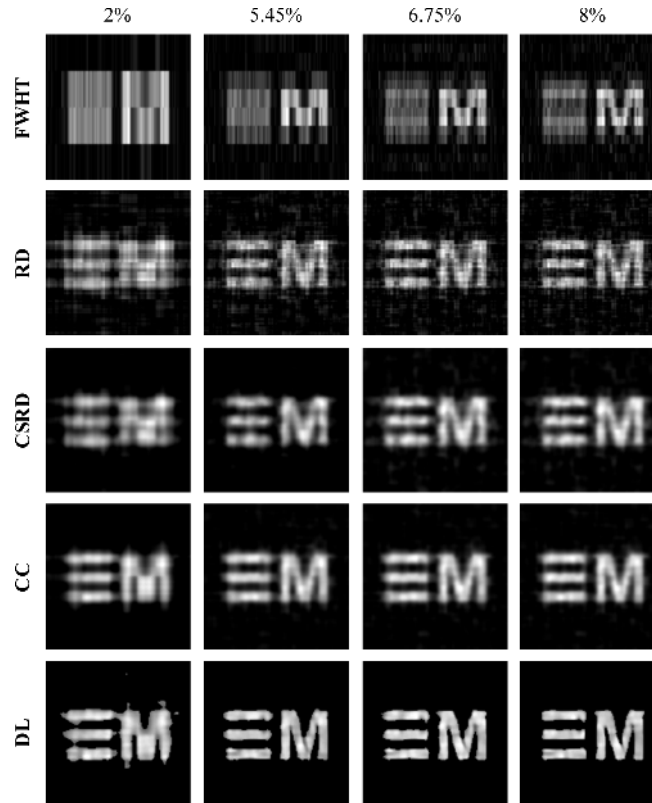

Fig. 4. Comparison of experimental results obtained by FWHT, RD, CSRD, CC and DL with the SR at 2%, 5.45%, 6.75% and 8%, respectively.

Figure 4 shows the comparing experimental results, where the SRs are equal to 2%, 5.45%, 6.75% and 8%, respectively. The original target image (128 × 128 pixels), which is captured by a CCD camera and processed by MATLAB, is shown in Fig. 1. Note that the original target image is used as a reference image for calculating the peak signal-to-noise ratio (PSNR) [37] and Structural SIMilarity (SSIM) index [38] values in Fig. 5. As shown in Fig. 4, when η is 2%, all the five methods fail to reconstruct a ghost image. When η is 5.45%, only CC and DL can restore a clear ghost image, and the image quality of DL is better. As the SR increases, the image quality of DL is always superior to the other four methods. Here, the restored image quality is quantitatively evaluated by PSNR and SSIM. Figure 5 shows the PSNR and SSIM curves which is corresponding to the images in Fig. 4. Both the PSNR and SSIM of DL are higher than those of the other four methods, indicating that the image quality of DL is better than the other four. Note that the LIS generation method should be the same as that in preparing the training dataset. The reason is that DL is not suitable for arbitrary LIS generation methods in this paper.

Consequently, if the LIS generation method is not the same as that used to prepare the training set, the PNSR and SSIM of reconstructed results may decrease. From Figs. 4 and 5, it can be seen that DL possesses the best imaging performance among the five methods in sub-Nyquist SR condition (e.g., 5.45%). Compared with [27], DL can restore a clear image with a lower SR (5.45%) than [27] (6.25%), and has a smaller error. In DL, when SR is 5.45%, the root of mean square error (RMSE) [27] is 6.07. However, in [27], when SR is 5.45%, the RMSE is over 40. Here, the SR shape is independent.
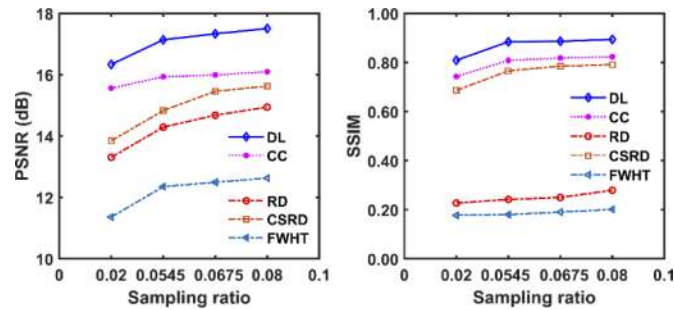

Fig. 5. Comparison of PSNR and SSIM acquired by FWHT, RD, CSRD, CC and DL using different SRs. The horizontal axis is not plotted in ratio.

To further check the imaging performance of DL, a grayscale image (light tower) is adopted as the target. The SR η is set as 5.45%. Experimental results of the five methods are displayed in Fig. 6. Both FWHT and RD cannot provide a recognizable image. CSRD and CC can restore ghost images, but some details are vague, such as the window, and the top of the tower. DL can reconstruct the details, and presents a better performance than CSRD and CC.
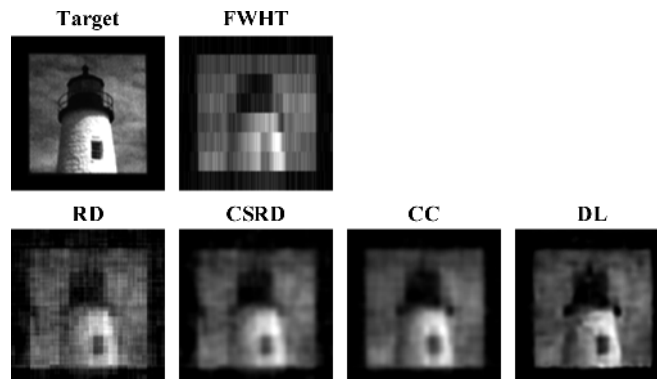

Fig. 6. Experimental results of a grayscale target reconstructed by FWHT, RD, CSRD, CC and DL with a SR of 5.45%.

4. Conclusion

We have experimentally presented a deep-learning-based CGI scheme, which obtains high-quality GI images in the sub-Nyquist SR condition. We develop a DAttNet that can recover the target image directly from the 1-D LIS without using the illumination pattern sequence. The DAttNet is trained with simulation data and tested using experimental data. We compare the imaging performance of the proposed scheme with four state-of-the-art CGI methods, FWHT, RD, CSRD and CC. The experimental results show that the image quality of the proposed scheme is better

than that of FWHT, RD, CSRD and CC in sub-Nyquist SR conditions. The proposed scheme can be applied in many areas, such as real-time CGI, scattering media GI and moving target CGI. The future work will focus on the optimizations and practical applications of the proposed scheme.

Funding

Disclosures

The authors declare no conflicts of interest.

References

1. S. Ma, Z. Liu, C. Wang, C. Hu, E. Li, W. Gong, Z. Tong, J. Wu, X. Shen, and S. Han, "Ghost imaging LiDAR via sparsity constraints using push-broom scanning," Opt. Express **27**(9), 13219–13228 (2019).
2. J. Cheng and S. Han, "Incoherent coincidence imaging and its applicability in X-ray diffraction," Phys. Rev. Lett. **92**(9), 093903 (2004).
3. W. Gong, "High-resolution pseudo-inverse ghost imaging," Photonics Res. **3**(5), 234–237 (2015).
4. X. Liu, X. Yao, R. Lan, C. Wang, and G. Zhai, "Edge detection based on gradient ghost imaging," Opt. Express **23**(26), 33802–33811 (2015).
5. M. Chen, H. Wu, R. Wang, Z. He, H. Li, J. Gan, and G. Zhao, "Computational ghost imaging with uncertain imaging distance," Opt. Commun. **445**, 106–110 (2019).
6. M. Sun and J. Zhang, "Single-pixel imaging and its application in three-dimensional reconstruction: a brief review," Sensors **19**(3), 732 (2019).
7. T. B. Pittman, Y. H. Shih, D. V. Strekalov, and A. V. Sergienko, "Optical imaging by means of two-photon quantum entanglement," Phys. Rev. A **52**(5), R3429–R3432 (1995).
8. X. Yao, W. Yu, X. Liu, L. Li, M. Li, L. Wu, and G. Zhai, "Iterative denoising of ghost imaging," Opt. Express **22**(20), 24268–24275 (2014).
9. C. Yang, C. Wang, J. Guan, C. Zhang, S. Guo, W. Gong, and F. Gao, "Scalar-matrix-structured ghost imaging," Photonics Res. **4**(6), 281–285 (2016).
10. H. Wu, X. Zhang, J. Gan, C. Luo, and P. Ge, "High-quality correspondence imaging based on sorting and compressive sensing technique," Laser Phys. Lett. **13**(11), 115205 (2016).
11. W. Yu, M. Li, X. Yao, X. Liu, L. Wu, and G. Zhai, "Adaptive compressive ghost imaging based on wavelet trees and sparse representation," Opt. Express **22**(6), 7133–7144 (2014).
12. Y. O-oka and S. Fukatsu, "Differential ghost imaging in time domain," Appl. Phys. Lett. **111**(6), 061106 (2017).
13. H. Wu, R. Wang, C. Li, M. Chen, G. Zhao, Z. He, and L. Cheng, "Influence of intensity fluctuations on Hadamard-based computational ghost imaging," Opt. Commun. **454**, 124490 (2020).
14. L. Wang and S. Zhao, "Fast reconstructed and high-quality ghost imaging with fast Walsh–Hadamard transform," Photonics Res. **4**(6), 240–244 (2016).
15. J. H. Shapiro, "Computational ghost imaging," Phys. Rev. A **78**(6), 061802 (2008).
16. R. I. Stantchev, B. Sun, S. M. Hornett, P. A. Hobson, G. M. Gibson, M. J. Padgett, and E. Hendry, "Noninvasive, near-field terahertz imaging of hidden objects using a single-pixel detector," Sci. Adv. **2**(6), e1600190 (2016).
17. H. Yu, R. Lu, S. Han, H. Xie, G. Du, T. Xiao, and D. Zhu, "Fourier-transform ghost imaging with hard X rays," Phys. Rev. Lett. **117**(11), 113901 (2016).
18. R. E. Meyers, K. S. Deacon, and Y. Shih, "Positive-negative turbulence-free ghost imaging," Appl. Phys. Lett. **100**(13), 131114 (2012).
19. W. Gong, H. Yu, C. Zhao, Z. Bo, M. Chen, and W. Xu, "Improving the imaging quality of ghost imaging lidar via sparsity constraint by time-resolved technique," Remote Sens. **8**(12), 991 (2016).
20. J. Huang and D. Shi, "Multispectral computational ghost imaging with multiplexed illumination," J. Opt. **19**(7), 075701 (2017).
21. M. Sun, L. Meng, M. P. Edgar, M. J. Padgett, and N. Radwell, "A Russian Dolls ordering of the Hadamard basis for compressive single-pixel imaging," Sci. Rep. **7**(1), 3464 (2017).
22. B. Luo, P. Yin, L. Yin, G. Wu, and H. Guo, "Orthonormalization method in ghost imaging," Opt. Express **26**(18), 23093–23106 (2018).

23. W. Yu, "Super Sub-Nyquist Single-Pixel Imaging by Means of Cake-Cutting Hadamard Basis Sort," Sensors **19**(19), 4122 (2019).

24. M. Lyu, W. Wang, H. Wang, H. Wang, G. Li, N. Chen, and G. Situ, "Deep-learning-based ghost imaging," Sci. Rep. **7**(1), 17865 (2017).

25. Y. He, G. Wang, G. Dong, S. Zhu, H. Chen, A. Zhang, and Z. Xu, "Ghost Imaging Based on Deep Learning," Sci. Rep. **8**(1), 6469 (2018).

26. S. Rizvi, J. Cao, K. Zhang, and Q. Hao, "Improving Imaging Quality of Real-time Fourier Single-pixel Imaging via Deep Learning," Sensors **19**(19), 4190 (2019).

27. F. Wang, H. Wang, H. Wang, G. Li, and G. Situ, "Learning from simulation: An end-to-end deep-learning approach for computational ghost imaging," Opt. Express **27**(18), 25560–25572 (2019).

28. 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).

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*, pp. 234–241, Springer, Cham, Oct 5, 2015.

30. O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, and B. Kainz, "Attention u-net: Learning where to look for the pancreas," arXiv preprint arXiv:1804.03999 (2018).

31. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708. 2017.

32. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," Advances in neural information processing systems, 1097–1105 (2012).

33. S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," arXiv preprint arXiv:1502.03167 (2015).

34. R. Timofte, S. Gu, J. Wu, and L. Van Gool, "Ntire 2018 challenge on single image super-resolution: Methods and results," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 852–863, June, (2018).

35. http://cocodataset.org/#download.

36. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980 (2014).

37. H. Liu, B. Yang, Q. Guo, J. Shi, C. Guan, G. Zheng, H. Mühlenbernd, G. Li, T. Zentgraf, and S. Zhang, "Single-pixel computational ghost imaging with helicity-dependent metasurface hologram," Sci. Adv. **3**(9), e1701477 (2017).

38. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. on Image Process. **13**(4), 600–612 (2004).