


ARTICLE

Open Access

Far-field super-resolution ghost imaging with a deep neural network constraint

Fei Wang^{1,2}, Chenglong Wang^{1,2}, Mingliang Chen^{1,2}, Wenlin Gong^{1,2}, Yu Zhang¹, Shensheng Han^{1,2,3,4} and Guohai Situ^{1,2,3,4} 

Abstract

Ghost imaging (GI) facilitates image acquisition under low-light conditions by single-pixel measurements and thus has great potential in applications in various fields ranging from biomedical imaging to remote sensing. However, GI usually requires a large amount of single-pixel samplings in order to reconstruct a high-resolution image, imposing a practical limit for its applications. Here we propose a far-field super-resolution GI technique that incorporates the physical model for GI image formation into a deep neural network. The resulting hybrid neural network does not need to pre-train on any dataset, and allows the reconstruction of a far-field image with the resolution beyond the diffraction limit. Furthermore, the physical model imposes a constraint to the network output, making it effectively interpretable. We experimentally demonstrate the proposed GI technique by imaging a flying drone, and show that it outperforms some other widespread GI techniques in terms of both spatial resolution and sampling ratio. We believe that this study provides a new framework for GI, and paves a way for its practical applications.

Introduction

Conventional imaging methods exploit the light reflected or scattered by an object to form its image on a two-dimensional sensor that has millions of pixels. However, ghost imaging (GI), an advanced imaging modality based on the second-order correlation of quantum or classical light, uses a single-pixel detector instead to record the reflected or scattered light, yielding a one-dimensional (1D) bucket signal^{1–6}. In some cases, an additional position sensitive detector is required to measure the illumination patterns. Although neither detector directly records a resolvable image of the object, one can employ an intuitive linear algorithm to reconstruct its image by spatial correlating the acquired time-varying patterns and the synchronized bucket signal. As it uses a single-pixel

detector to collect the photons that interact with the object, GI has significant advantages over conventional imaging modalities in terms of detection sensitivity, dark counts, spectral range, and cost efficiency^{7,8}. In addition, with the aid of some prior information, e.g., sparsity, it is capable of sensing compressively during data acquisition^{9,10}. Such enhancements can provide significance in low-light imaging where the photon counts are very low due to scattering or absorption losses as in medical imaging or remote sensing; and in non-visible waveband imaging where the availability of silicon-based sensor becomes expensive or impractical as in infrared or deep ultraviolet regime.

However, in GI, a large amount of single-pixel measurements is necessary because one sampling only contains a little information about the object. Specifically, to obtain an N -pixel image one needs at least $M = N$ measurements to meet $\beta = M/N = 100\%$, where β represents the sampling ratio (the Nyquist sampling criterion). In many applications such as remote sensing¹⁰, a rotating ground glass (RGG) is frequently used

Correspondence: Guohai Situ (ghsitu@siom.ac.cn)

¹Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai 201800, China

²Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

Full list of author information is available at the end of the article

© The Author(s) 2022



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.



to generate speckle illumination patterns compared with other programmable modulation strategies, e.g., digital micromirror device¹¹ owing to its high power endurance and cost efficiency. In this case one needs $M \gg N$ measurements to improve the signal-to-noise ratio (SNR) of the reconstructed image due to the overlap of different patterns⁹. This inevitably leads to a paradox between the number of pixels occupied by the object and the data acquisition time. In addition, the spatial resolution of GI is physically limited by the grain size of the speckle pattern on the object plane¹². This is unfavorable for far-field imaging as the speckle grain becomes too large to distinguish the detailed structure of the object^{13,14}. Thus, an intuitive and longstanding goal in the study of GI is to decrease β while retaining good resolution, so as to reduce the burden of data acquisition and produce better imaging visual effects. However, the consequential incomplete sampling strategy usually lead to ill-posedness in GI reconstruction. Thus, suitable prior assumptions are needed to compensate the missing information.

One popular approach is based on compressive sensing (CS). CS uses sparsity as a general prior assumption and has become a popular signal reconstruction framework^{15–17}. It has been widely used in various imaging systems such as single-pixel cameras¹¹ and compressive holography¹⁸. Specifically, given the measurements y , the CS technique usually reconstructs the object x by solving the following iteration problem:

$$\min_x \|\Phi x - y\|_2^2 + \xi \|\Psi x\|_1 \quad (1)$$

where Φ is the random measurement matrix and Ψ is the transformation matrix that transforms x into a sparse domain such as discrete cosine transform (DCT) or wavelet. Ψx represents the corresponding transform coefficients regularized by the l_1 norm with the regularization parameter ξ . Owing to the sparsity of the image of the object and the randomness of the illumination patterns, CS is also suitable for GI reconstruction. Such GI using sparsity constraint, or GISC for short, enables the reconstruction of high-quality and high-resolution image when $\beta < 100\%$ ^{7,9,13,19,20} [Fig. 1f]. In the field of GI, CS has been used for resolution enhancement^{21–23}, remote sensing¹⁰, 3D imaging²⁴, and among many others^{7,8,19}. However, it is still a challenging problem for GISC to operate well in the case when β is less than the Cramer–Rao bound^{16,17,22}.

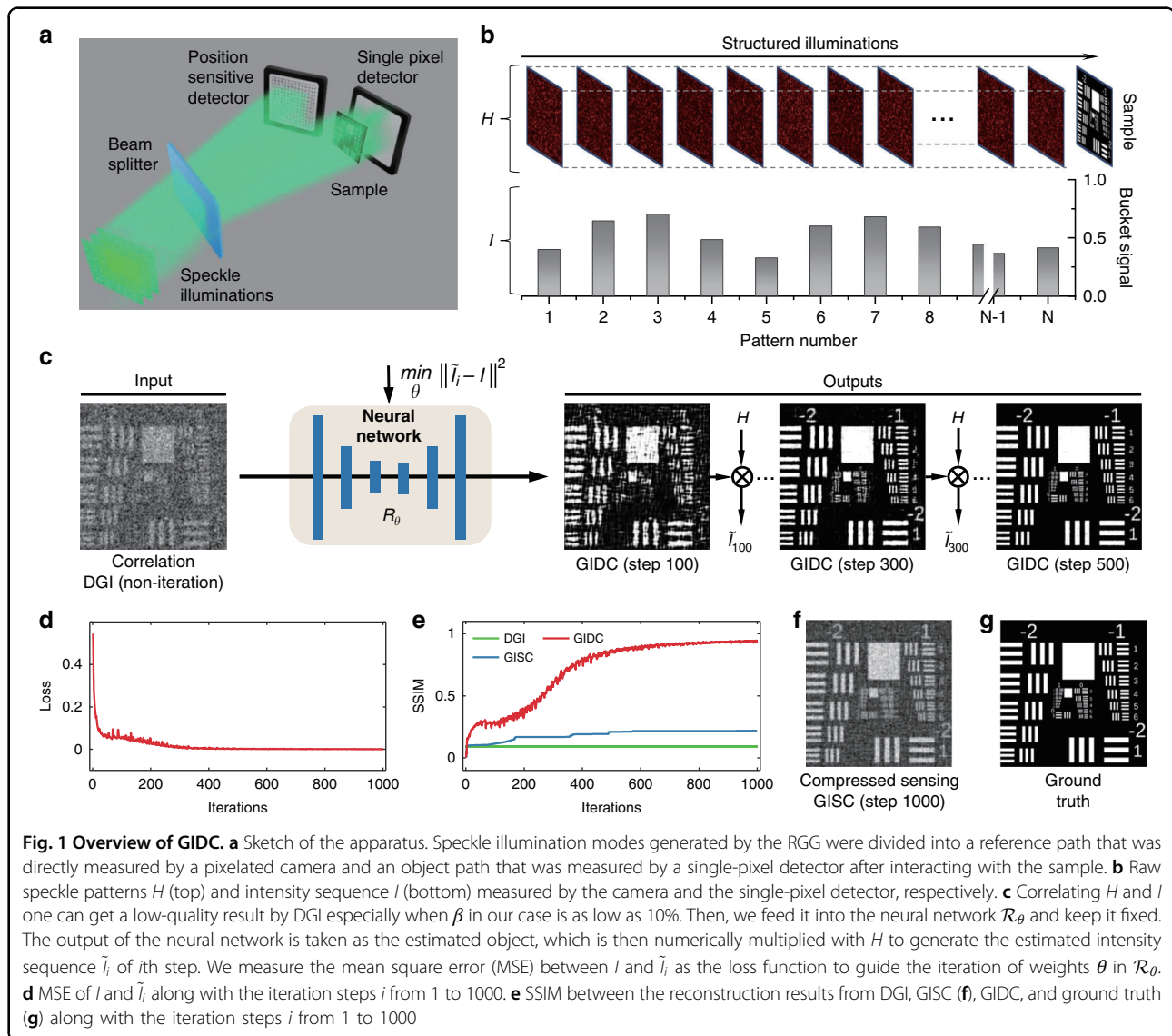
An alternative but increasingly important approach is deep learning that is based on data prior assumptions^{25–27}. Specifically, it has shown that it allows robust GI reconstruction of high-quality images even when $\beta < 10\%$ with high computational efficiency^{28–31}. Such GI based on deep learning (GIDL) technique uses a deep neural network

(DNN) to learn from a large number of input–output data pairs so as to establish a mapping relationship among them. The experimental acquisition of such a huge training set is time consuming and laborious because one needs at least thousands of measurements for one data pair even for a 64×64 image in a proof-of-principle experiment. Though the neural network can be trained on simulation data³⁰, the trained model only works well for the reconstruction of objects that resemble those in the training set. This challenge of generalization is one of the big issues that need to be addressed.

Recently, Ulyanov et al.³² proposed the deep image prior (DIP) framework that uses an untrained neural network as a constraint for image processing tasks such as denoising, inpainting, and super resolution. They demonstrated that a properly designed generator network architecture itself has an implicit bias towards natural images and thus can be used for solving ill-posed inverse problems³³. The most significant advantage of DIP is that a generator network can be used without training beforehand, and thus eliminating the need for tens of thousands of labeled data. A similar concept has also been used for computational imaging, such as phase retrieval^{34,35}, CS^{36,37}, and diffraction tomography³⁸.

Inspired by the idea of DIP, here we propose a new GI technique that incorporates the physical model of GI image formation into a DNN. We hypothesize that the image prior information introduced by an untrained DNN can be applied to achieve better GI reconstruction under much lower β . We term the proposed technique as GI using Deep neural network Constraint (GIDC). It utilizes an untrained DNN to generate high-quality and high-resolution results. The only input it requires are a 1D bucket signal sequence I from which one needs to reconstruct an image, together with the associated stack of illumination patterns H , which is easily accessible in a typical GI system [Fig. 1a]. The proposed GIDC technique is described as follows. First, we correlate the H [Fig. 1b, top] and I [Fig. 1b, bottom] by differential ghost imaging (DGI)^{39,40} and obtain a rough reconstruction of the image. Second, we feed the resulting DGI reconstruction into a randomly initialized neural network (untrained). Third, we take the output of the neural network as an estimation of a high-quality GI image and use it to calculate a bucket signal I by using a GI image formation model. Finally, we update the weights of the neural network to minimize the error between the measured and estimated bucket signal [Fig. 1c]. Along with the error reduction [Fig. 1d], the output of the neural network also converges to a good-quality image [Fig. 1e]. Compared with conventional DGI and GISC [Fig. 1f], the proposed strategy dramatically increases the quality and resolution of GI under much lower sampling ratio β . Compared with those state-of-the-art





deep-learning-based methods, GIDC does not need to train on any labeled data and thus is more flexible and does not bias towards a specific distribution. Specifically, our contributions include:

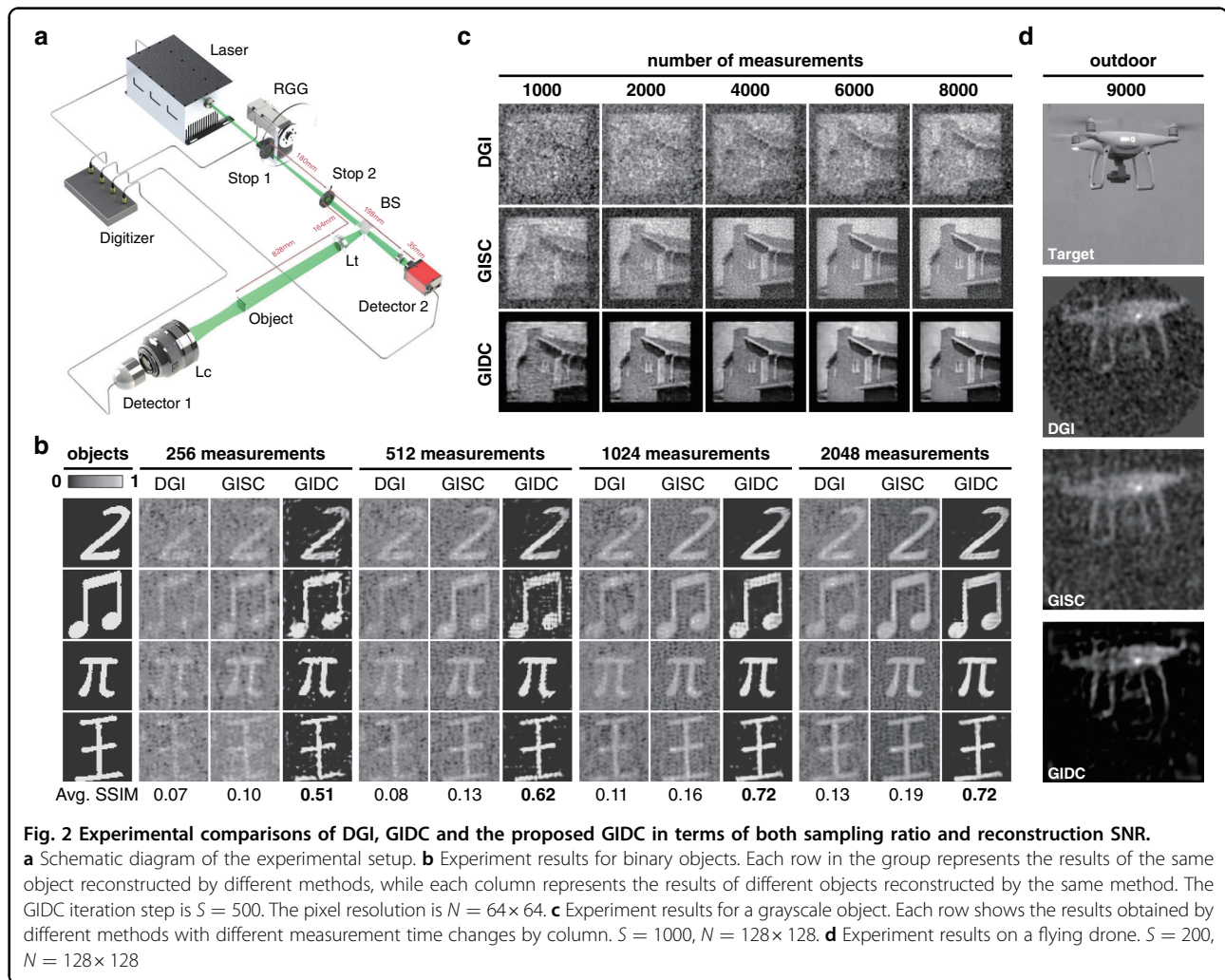
- We demonstrate that GIDC can reconstruct a dramatically high SNR GI image at a very low sampling ratio β .
- We demonstrate that GIDC can enhance the resolution of the reconstructed image even when the speckle grain size is larger, suggesting its potential to break the diffraction limit.
- We perform a comparative study on the base of a number of challenging real-world scenarios including a flying drone and synthesized dataset, and demonstrate that GIDC outperforms other widespread GI methods, including DGI, GISC, and GIDL.

Results

Sampling ratio

We built a typical Pseudothermal GI system [Fig. 2a] for data acquisition. Here we show the reconstruction results of different objects using different methods at different sampling ratios. The first group of results is plotted in Fig. 2b. One can clearly see that all the binary objects have been successfully reconstructed by GIDC, with the number of measurements as low as 256 ($\beta = 6.25\%$). We also take DGI^{39,40} and GISC¹³ for comparison. For all the cases (different objects and β settings), GIDC outperform DGI and GISC both in terms of visual appearance and quantitative evaluation index (SSIM). We observe the same results in the cases that the object is in grayscale [Fig. 2c]. One can clearly see that the clean and high-contrast images reconstructed by GIDC, whereas the ones





recovered by DGI and GISC are dirty or even corrupted by strong noise in particular when β is low (see first two columns in Fig. 2c).

We also conducted an outdoor experiment to demonstrate the effectiveness of GIDC. The data were acquired by using a homemade GI LiDAR system⁴¹. The imaging target [Fig. 2d, top] is a flying drone (DJI, Phantom 4) hovering in the air, 50 m away from the GI LiDAR system. The main results are plotted in Fig. 2d. One can clearly see that the GIDC can successfully reconstruct the shape of the drone with very high contrast. The size of the reconstructed image is 128×128 , meaning that the sampling ratio $\beta = 9000/16,384 \approx 55\%$. The reconstructed image by DGI and GISC plotted as well for comparison. One can see that the image reconstructed by DGI or GISC is corrupted by noise, and the contrast is low.

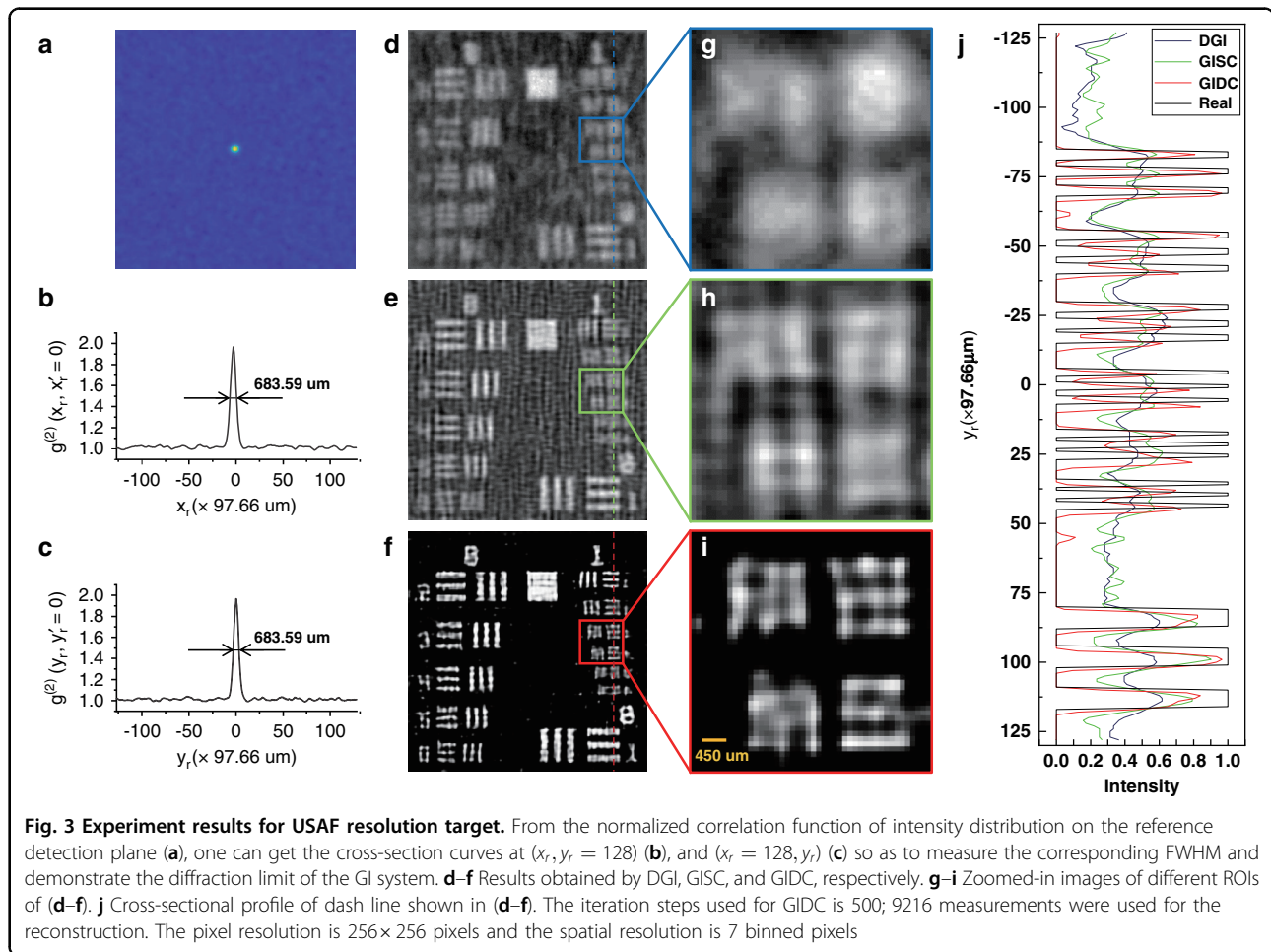
Resolution

We also experimentally demonstrated the spatial resolution that GIDC can offer. It is known that as an imaging

method based on the second-order (intensity) correlation of light, the spatial resolution of GI is theoretically limited by the width of the mutual correlation function of the illumination speckle patterns, measured at the object plane⁴². According to this, we first calculated the normalized correlation function [Fig. 3a] of the recorded speckle patterns⁴³, namely

$$g^{(2)}(x_r, x'_r = 0) = \frac{\langle H_1(x_r) H_2(x'_r = 0) \rangle}{\langle H_1(x_r) \rangle \langle H_2(x'_r = 0) \rangle} \quad (2)$$

Then, we measured the full-width at half-max (FWHM) to estimate the value of the speckle grain size on the object plane. We found that it occupies 7 binned pixels in both the horizontal and vertical directions [Fig. 3b, c], suggesting that the diffraction limit of our experimental GI system is $683.59\text{ }\mu\text{m}$. More details about the system configuration toward the GI system can be found in the section “Methods and Materials.”



A USAF resolution target was used to test the resolution of different GI reconstruction methods. The main results are plotted in Fig. 3d–i. As expected, the image reconstructed by DGI suggests that the elements in Group 0 Element 5 are not resolvable because the linewidth ($629.96 \mu\text{m}$) is smaller than the diffraction limit ($683.59 \mu\text{m}$). It becomes a little bit better by using GISC, where some elements with their linewidth smaller than the diffraction limit (Group 1 Element 1, $500 \mu\text{m}$) can be distinguished. Evidentially, the proposed GIDC has the best performance in terms of both linewidth and sharpness exhibited in the reconstructed image. As shown in Fig. 3g–i, the line pairs in Group 1 Element 4 with the linewidth of $353.55 \mu\text{m}$ can be successfully reconstructed by GIDC, but neither DGI nor GISC achieves the same performance. This suggests that the proposed GIDC has the capability of enhancing the resolution by a factor of about 2 ($683.59/353.55 = 1.93$) with respect to the diffraction limit. More evidence can be found in Fig. 3j. In addition to the advantages of resolution, the image reconstructed by GIDC has much higher contrast as evidenced by the clean background.

Discussion

In this section, we make some more in-depth discussions on the performance of GIDC in comparison to DGI and GISC. GIDL trained on two different datasets were also considered. For the sake of quantitative evaluation, we examine on simulation data in this section.

Accuracy

Two different β settings were studied here, i.e., $\beta = 12.5\%$ and $\beta = 25\%$, corresponding to the number of measurements $M = 512$ and 1024 , respectively. The results are shown in Fig. 4. Apparently, the images reconstructed by GIDC have the best fidelity for all the sample objects we studied here. In the case of $\beta = 12.5\%$, we observed that the reconstructed grayscale images are not as good as the reconstructed binary images even using GIDC. This is probably because a grayscale image contains too much unknown information to be determined, and it seems to be unfeasible to achieve a good reconstruction with a small sampling ratio. However, the reconstructed images are much better when $\beta = 25\%$, which is in consistency with the optical experimental



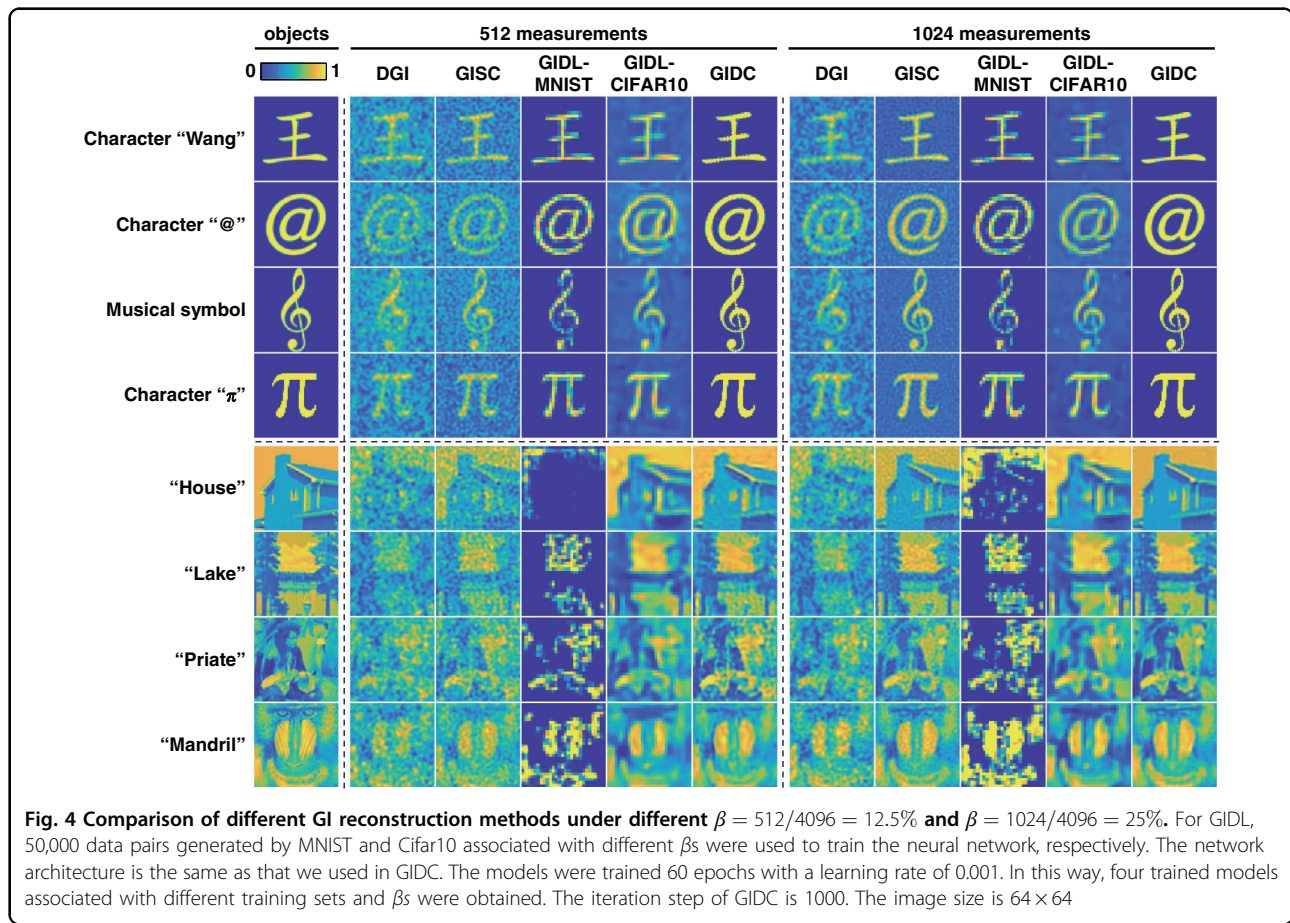


Table 1 The metrics of different GI reconstruction methods on SSIM when $\beta = 512/4096 = 12.5\%$ and $\beta = 1024/4096 = 25\%$

Number of measurements	512					1024				
	DGI	GISC	GIDL-MNIST	GIDL-CIFAR10	GIDC	DGI	GISC	GIDL-MNIST	GIDL-CIFAR10	GIDC
Character "Wang"	0.158	0.298	0.819	0.303	0.870	0.209	0.423	0.809	0.342	0.917
Character "@"	0.109	0.187	0.858	0.220	0.971	0.130	0.252	0.866	0.226	0.978
Musical symbol	0.109	0.179	0.827	0.177	0.898	0.131	0.232	0.843	0.215	0.889
Character " π "	0.089	0.138	0.836	0.193	0.993	0.100	0.208	0.866	0.237	0.995
"House"	0.179	0.397	0.016	0.452	0.653	0.218	0.502	0.102	0.470	0.790
"Lake"	0.177	0.431	0.075	0.423	0.606	0.254	0.634	0.119	0.468	0.697
"Priate"	0.267	0.513	0.130	0.412	0.504	0.337	0.657	0.189	0.493	0.726
"Mandril"	0.229	0.463	0.132	0.407	0.641	0.332	0.573	0.196	0.476	0.712

Bold values indicates the highest quantitative metrics (SSIM)

results shown in Fig. 2. In order to quantitatively evaluate the results obtained by different methods, we calculated the SSIM value for each reconstructed image with respect to the corresponding ground truth. The SSIMs are listed

in Table 1. It is clearly seen that GIDC has the highest metrics values in most of the cases, suggesting that the reconstruction accuracy of GIDC outperforms the others. The performance of GIDL, however, depends strongly on

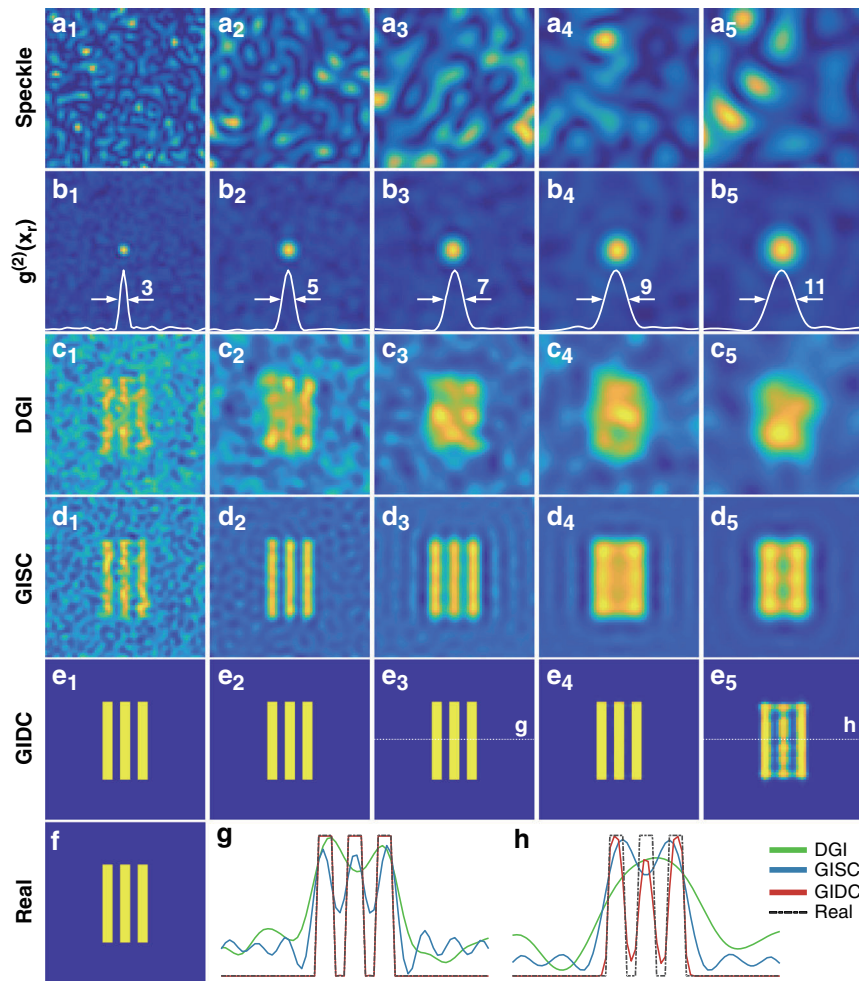


Fig. 5 Comparison of GI resolution using different reconstruction algorithms. **a₁–a₅** One of the speckle patterns used for generating the bucket signal. **b₁–b₅** The mutual correlation function was used for estimating the speckle grain size. From left to right the FWHM gradually increased. The unit is μm . **c₁–c₅**, **d₁–d₅**, and **e₁–e₅** The DGI, GISC, and GIDC results using the corresponding speckles. **f** The real object. **g**, **h** Line plots show image intensities along the dashed lines in **e₃** and **e₅**, respectively

the training set and the task in hand. For instance, relatively good performance can be achieved when using GIDL trained on MNIST to reconstruct binary characters. In contrast, the reconstructed images are severely corrupted in the cases of grayscale due to the limited generalization. Although this can be slightly relieved by training it on an alternative dataset such as Cifar10⁴⁴, it affects the accuracy of the reconstructed binary characters images as suggested by results shown in the fifth and tenth columns in Fig. 4. By contrast, GIDC is a general method that can be used to reconstruct different types of objects usually with a high accuracy and a low β .

Resolution

Here we will analyze the experiment result on resolution enhancement shown in the section “Resolution.” First we compare the resolution of the images reconstructed by

GIDC and other widespread GI algorithms from the same set of simulation data. We generated five groups of illumination speckles [Fig. 5a1–a5] with the grain size $x_s = \lambda z/D$ varying from 3 to 11 μm [Fig. 5b1–b5] to encode the object which was a triple-slit pattern shown in Fig. 5f. We set $\beta = 410/4096 \approx 10\%$. We found that the DGI cannot distinguish the slits well when $x_s > 5 \mu\text{m}$. As expected, GISC can enhance the resolution. As evidenced in Fig. 5d3, the slit pattern can still be recognized when x_s is as large as 7 μm . However, GISC fails when $x_s \geq 9 \mu\text{m}$. By contrast, the proposed GIDC reconstructs an almost perfect image under the same condition. Even when x_s is as large as 11 μm , the GIDC still provide a very good result [Fig. 5e5]. The cross-section of the reconstructed image when $x_s = 7 \mu\text{m}$ and $x_s = 11 \mu\text{m}$ was plotted in Fig. 5g, h, respectively. From the results, one can clearly conclude that the proposed GIDC can provide dramatically



resolution enhancement compared with DGI and GISC, in high consistence with the experimental results presented in Fig. 3.

Note that in the studies of phase imaging using untrained neural networks^{34,35}, we did not observe such a phenomenon of resolution enhancement. So it must have something to do with the imaging modality of GI. There are three unique features that GI possesses in comparison to phase imaging. First, the object is illuminated by a random beam. Second, the light scattered from the object is recorded with a bucket detector. Third, GI relies on the second-order correlation of the light field⁴⁵, whereas phase imaging relies on the first order. Each implementation of the random illumination can shift some of the high-spatial frequency components to the lower band⁴⁶. This means that the associated information beyond the diffraction limit can be efficiently encoded and transmitted to the detector. A similar concept has been introduced in microscopy to achieve super resolution as well⁴⁷. In the case of GI, however, the decoding of those high-frequency components is not so trivial due to the fact that they are highly compressed in the 1D bucket signal. Indeed, as shown in Fig. 5, none of those widespread GI algorithms can do this job. In contrast, GIDC endeavors to find a feasible solution that can reproduce the acquired bucket signal. Such a feasible solution has to contain those high-frequency components encoded in the bucket signal in order to decrease the loss function.

Robustness

The robustness is evaluated by examining the effect of noise in the detection to the reconstructed image. There are different kinds of noise in the detection process⁴⁸, but the noise effect can be modeled as an additive Gaussian distribution with the standard deviation δ as a whole^{20,49}. Thus, one can define the detection SNR (dSNR) as follows⁵⁰:

$$\text{dSNR} = 10\log_{10} \frac{\langle I \rangle}{\delta} \quad (3)$$

to describe the degradation of the detected signal. Two cases were examined in our studies. In the first case, we fixed the dSNR to 26 dB, and see how the reconstructed image would be under different sampling ratio conditions. In the second one, we fixed the sampling ratio to be 60% for different dSNRs. In this analysis, eight standard grayscale images (Supplementary Fig. S1) were used as the target. We again used SSIM to measure the quality of the reconstructed image from the contaminated bucket signal. The results are plotted in Supplementary Fig. S2; one can clearly see that GIDC has the best performance among all the three, in particular when the noise level is high. For DGI, the SSIM value of the reconstructed image is linearly increased with the sampling ratio β as the SNR

of the reconstructed image is linearly proportional to the number of measurements^{39,40}. In addition, we observed that the averaged SSIM in this case is around 0.46 when $\beta = 60\%$. This noise-independence effect is highly consistence with the theoretical prediction^{39,40}. On the contrary, GISC is more sensitive to the detection noise¹⁵ as the SSIM drops from 0.862 to 0.544 when the dSNR is decreased from 30 to 22 dB. Some visualization results can be found in Supplementary Fig. S3.

Priors

The effect of priors is also examined here. Two types of priors were used in GIDC, the physical prior, i.e., DGI, and the total variation (TV) regularization. Here we analyze the effect of DGI and TV independently and in combination. When either DGI or TV is not used, the associated SSIM values are plotted as the bars in green and turquoise, respectively, in Supplementary Fig. S2. One can see that, in all the cases, the SSIM values are slightly less than the one associated with GIDC (orange). This suggests that the use of priors does have contribution to the quality of the reconstructed image. This can be more clearly seen by the yellow bars, which are associated with the cases that neither of them was used. But even in this case the reconstructed image is still far better than the one obtained from DGI alone, suggesting that the GIDC framework has good robustness performance. Some visualization results can be found in Supplementary Fig. S3.

Computational efficiency

It is necessary to compare the computational time for different approaches. Different image sizes were considered when β is set to 6.25%. Compared with DGI and GISC, GIDC provides the best results in terms of both visualizations [Supplementary Fig. S4a] and quantitative metrics [Supplementary Fig. S4c] under all pixel-resolution settings. However, as shown in Supplementary Fig. S4b, GIDC needs the longest time to optimize. For a 128×128 image, it needs about 5 min to restore a feasible result, while DGI and GISC only needs 0.221 and 12.29 s, respectively. Thus, the previous GISC and our GIDC are both not suitable for real-time applications, at least on the current computing platform. Despite this, for applications that allow post-processing offline but require fast data acquisition, GIDC yields the highest image fidelity at the lowest sampling ratio. The reconstructed image is associated with a SSIM value of 0.9 even when β is down to 6.25%. We also noticed that the computational time dramatically increases along with the increase of the image size. There are mainly two reasons for this. First, the width of the network will increase accordingly to accept the image as its input, process it and produce an output. Thus, it takes more time to forward infer during each iteration. Second, the size of the measurement



matrix H that is used to generate the estimated bucket signal will increase. Thus, it takes more time to calculate the gradient and update the network parameters.

There are several strategies that one can take into account to improve the GIDC computational efficiency. These include better design of the neural network architecture, the implementation of depth-wise convolution⁵¹, the employment of better initialization⁵² and learning⁵³ strategy. In addition, from a practical application point of view, the implementation of GIDC on a faster computing platform together with hardware speedup by using multiple GPUs will also help to significantly increase the computational efficiency.

Methods and materials

Formation of the reconstruction algorithm

For an object $O(x_t, y_t)$, the measurements of the pseudo-thermal GI system are the 1D bucket signal

$$I_m \propto \sum_{x_t, y_t} H_m(x_t, y_t) O(x_t, y_t) \quad (4)$$

measured by a single-pixel detector in the test arm, and the corresponding stack of random illumination patterns $H_m(x_t, y_t)$, where $m = 1, 2, \dots, M$, measured by a high-resolution camera in the reference arm. The conventional GI algorithm reconstructs the object image by computing the intensity correlation between H_m and I_m

$$O_{GI} = \langle H_m I_m \rangle - \langle H_m \rangle \langle I_m \rangle \quad (5)$$

where $\langle \cdot \rangle$ denotes the ensemble average approximately defined as $H_m = \frac{1}{M} \sum_{m=1}^M H_m$ and $I_m = \frac{1}{M} \sum_{m=1}^M I_m$.

For DGI^{39,40}, one uses $S_m = \sum H_m(x_t, y_t)$ to normalize the illumination patterns so as to improve the SNR

$$O_{DGI} = \langle H_m I_m \rangle - \frac{\langle H_m \rangle}{\langle S_m \rangle} \langle S_m I_m \rangle \quad (6)$$

For the proposed GIDC, the reconstruction of the object image is formulated as the following objective function

$$\mathcal{R}_{\theta^*} = \operatorname{argmin}_{\theta \in \Theta} \sum_{m=1}^M \left\| \sum H_m \mathcal{R}_{\theta}(O_{DGI}) - I_m \right\|^2 \quad (7)$$

where \mathcal{R}_{θ} is the DNN defined by a set of weights and biases parameters Θ . The goal of GIDC is to find a good configuration $\theta^* \in \Theta$ for the neural network that forces its output $O_{GIDC} = \mathcal{R}_{\theta^*}(O_{DGI})$ to produce a 1D sequence \tilde{I} according to the GI image formation physics (Eq.4) that resembles the experimentally acquired bucket signal I . As

it is an ill-posed problem, especially when $M \ll N$, there are in principle an infinite number of configurations that satisfies the objective function. Therefore, it is necessary to add prior information about the object so as to select a feasible solution from all the configurations. For example, in GISC, the prior information is about an assumption that the object is sparse in a certain domain. Different from GISC, the proposed GIDC is based on an untrained DNN prior. Although the theory for this has yet to be perfected, existing works has empirically suggested that a properly designed DNN with randomly initialized weights has an inherent bias toward natural images^{32,34–38}. We thus hypothesize that the DNN prior can be used to solve the ill-posed problem described by Eq. (7). We also argue that adding a conventional regularization terms such as the TV³⁸ in the GIDC framework would help improving the reconstruction results. So the final objective function (loss function) of GIDC is reformulated as follows:

$$\mathcal{R}_{\theta^*} = \operatorname{argmin}_{\theta} \left\| H \mathcal{R}_{\theta}(O_{DGI}) - I \right\|^2 + \xi T[\mathcal{R}_{\theta}(O_{DGI})] \quad (8)$$

where T stands for TV and ξ is its strength.

For comparison, it is worthy of pointing out that GIDL uses a DNN as well. But it attempts to learn the mapping function \mathcal{R}_{θ} from a large number of labeled data pairs in the training set $S_T = \{(O_{DGI}^k, O^k) | k = 1, 2, \dots, K\}$, by solving

$$\mathcal{R}_{\theta^*} = \operatorname{argmin}_{\theta \in \Theta} \left\| \mathcal{R}_{\theta}(O_{DGI}^k) - O^k \right\|^2, \forall (O_{DGI}^k, O^k) \in S_T \quad (9)$$

GIDL learns to map the low-quality reconstructed images to a high-quality ones from the statistics of the training set S_T . Once trained, the neural net can be used directly to reconstruct objects that are similar with those in S_T .

By contrast, GIDC learns the mapping function through updating the weights and biases θ in the neural network to minimize the model-based fidelity term, which can be seen as an interplay between the GI physical model H and the DNN \mathcal{R}_{θ} . In this way, one can obtain a feasible solution $O_{GIDC} = \mathcal{R}_{\theta^*}(O_{DGI})$ without using any training data. That is to say, GIDC is an untrained method and does not bias toward any particular dataset. We note that the input of the neural network used in GIDC can be a coarse image recovered by any conventional GI algorithms^{20,34,39,40,46} or even random noise^{32,35,38}, here we use the result of DGI for convenience.

Network architecture and hyper parameters

The network architecture we employed in this work was derived from the U-net⁵⁴. More details of the network structure are provided in Supplementary Fig. S5. We adopted the Adam optimizer with a learning rate of



$\alpha = 0.05$, $\beta_1 = 0.5$, $\beta_2 = 0.9$, and $\epsilon = 10^{-9}$ to update the weights in the neural network. We also used an exponential decay with a decay rate of 0.9 and decay steps of 100. The momentum and epsilon parameters in the batch normalization were 0.99 and 0.001, respectively. The leak parameter of Leaky ReLU was 0.2. The regularization parameter of the TV was 10^{-10} . The code was run on a computer with an Intel Xeon CPU E5-2696 V3, 64 GB RAM, and an NVIDIA Quadro P6000 GPU. The main progress is illustrated in Algorithm 1. For the sake of comparison, we use the same network model for GIDC and GIDL. We also released our code at <https://github.com/FeiWang0824/GIDC>.

Algorithm 1 GI Reconstruction with GIDC. We use default values of $N = 1000$, $\lambda = 10^{-10}$, $\alpha = 0.05$, $\beta_1 = 0.5$, $\beta_2 = 0.9$, $\gamma = 0.9$, $\eta = 100$.

Input: H and I

Output: $O_{GIDC} = \mathcal{R}_\theta(O_{DGI})$

```

1: initialize: randomly initialize the parameters  $\theta$  in the
   neural network  $\mathcal{R}_\theta$ 
2:  $O_{DGI} = \text{DGI}(H; I)$ 
3: for  $step = 1, 2, \dots, N$  do
4:    $\tilde{I} \leftarrow H\mathcal{R}_\theta(O_{DGI})$ 
5:    $\mathcal{L}_\theta \leftarrow \text{MSE}(I, \tilde{I}) + \xi T[\mathcal{R}_\theta(O_{DGI})]$ 
6:    $\theta \leftarrow \text{Adam}(\nabla_\theta \mathcal{L}_\theta, \alpha, \beta_1, \beta_2)$ 
7:    $\alpha \leftarrow \alpha \gamma^{step/\eta}$ 
8: end for
```

Experimental details

Figure 2a presents the optical system we built for the experimental demonstration. L_t , L_r and L_c are lenses with focal length of 136.8, 30, and 75 mm, respectively. Detector 1 works as a single-pixel detector, whereas Detector 2 is a high-resolution camera. The light source is a solid-state pulsed laser with a $\lambda = 532$ nm center-wavelength, and a 10 ns pulse width at a repetition rate of 1 kHz. The pulsed beam emitted from the laser irradiated a RGG to produce pseudothermal light. The beam diameter on the RGG was D and can be adjusted by an optical stop (stop1). The distance between the RGG and the other optical stop (stop2) is about $z = 180$ mm. The shape of stop2 is a square with a side length equal to 5 mm. Owing to the RGG, a speckle field is fully developed at the plane of stop2. Then the speckle field is divided by a beam splitter into a test and a reference arms. In the test arm, we use an imaging lens L_t to project the speckle field at the stop2 plane to the surface of the object. The side length of the objects is about $L = 25$ mm (the magnification of L_t is 5). The transmitted light is collected by a lens L_c (Nikon AF-S NIKKOR 85 mm f/1.4 G) and finally recorded by a single-pixel detector (in our experiment, we actually used an AVT F504B camera to record the transmitted intensity, and generated the bucket signal by summing all the pixel values). In the reference arm, we use an image detector 2 (AVT F504B, with the pixel size

p_s of 3.45 μm) mounted with an imaging lens L_r to take a high-resolution image of the speckle pattern on the stop2 plane.

Three different types of objects were used to test the GIDC performance, i.e., transparent slices of various characters with binary value and of natural scenes in grayscale (a film after exposure a standard test image “house”), and a physical USAF resolution chart. The spatial resolution of our GI system can be adjusted by changing D as z is fixed through $x_s = \lambda z/D$. In addition, different pixel resolution N can be obtained by setting different resize factor $q = \sqrt{L/(\sqrt{N}p_s)}$ so that the binned pixel size is $q^2 p_s$. We set the resize factor of 10.64, 7.52, and 5.32 for binary characters, grayscale object, and USAF resolution chart to meet the pixel resolution of 64, 128, and 256, respectively. Then the binned pixel size is 390.63, 195.31, and 97.66 μm , respectively. For the experiment of USAF resolution chart, we set $D = 0.70$ mm to meet the spatial resolution of 683.59 μm (683.59/97.66 = 7 binned pixels).

Acknowledgements

We would like to thank Z. Tong (Shanghai Institute of Optics and Fine Mechanics) for discussion and helpful comments. This work was supported by the National Natural Science Foundation of China (61991452, 62061136005), the Key Research Program of Frontier Sciences of the Chinese Academy of Sciences (QYZDB-SSW-JSC002), and the Sino-German Center (GZ1391).

Author details

¹Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai 201800, China. ²Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing 100049, China. ³Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou 310024, China. ⁴CAS Center for Excellence in Ultra-intense Laser Science, Shanghai 201800, China

Author contributions

G.S. and F.W. conceived the idea. F.W. discovered the procedures, designed, performed, and analyzed most of the data. C.W. performed most of the experiments together with F.W. and C.W., M.C., and W.G. conducted the outdoor experiment. F.W. and G.S. wrote the manuscript. All authors discussed and commented on the manuscript. G.S. supervised the project.

Conflict of interest

The authors declare no competing interests.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41377-021-00680-w>.

Received: 17 May 2021 Revised: 28 October 2021 Accepted: 14 November 2021

Published online: 01 January 2022

References

- Pittman, T. B. et al. Optical imaging by means of two-photon quantum entanglement. *Phys. Rev. A* **52**, R3429–R3432 (1995).
- Strekalov, D. V. et al. Observation of two-photon “ghost” interference and diffraction. *Phys. Rev. Lett.* **74**, 3600–3603 (1995).
- Gatti, A. et al. Ghost imaging with thermal light: comparing entanglement and classical correlation. *Phys. Rev. Lett.* **93**, 093602 (2004).

4. Cheng, J. & Han, S. S. Incoherent coincidence imaging and its applicability in X-ray diffraction. *Phys. Rev. Lett.* **92**, 093903 (2004).
5. Erkmen, B. I. & Shapiro, J. H. Ghost imaging: from quantum to classical to computational. *Adv. Opt. Photonics* **2**, 405–450 (2010).
6. Moreau, P. A. et al. Ghost imaging using optical correlations. *Laser Photonics Rev.* **12**, 1700143 (2018).
7. Edgar, M. P., Gibson, G. M. & Padgett, M. J. Principles and prospects for single-pixel imaging. *Nat. Photonics* **13**, 13–20 (2019).
8. Gibson, G. M., Johnson, S. D. & Padgett, M. J. Single-pixel imaging 12 years on: a review. *Opt. Express* **28**, 28190–28208 (2020).
9. Katz, O., Bromberg, Y. & Silberberg, Y. Compressive ghost imaging. *Appl. Phys. Lett.* **95**, 131110 (2009).
10. Zhao, C. Q. et al. Ghost imaging lidar via sparsity constraints. *Appl. Phys. Lett.* **101**, 141123 (2012).
11. Duarte, M. F. et al. Single-pixel imaging via compressive sampling. *IEEE Signal Process. Mag.* **25**, 83–91 (2008).
12. Ferri, F. et al. High-resolution ghost image and ghost diffraction experiments with thermal light. *Phys. Rev. Lett.* **94**, 183602 (2005).
13. Gong, W. L. & Han, S. S. High-resolution far-field ghost imaging via sparsity constraint. *Sci. Rep.* **5**, 9280 (2015).
14. Li, Z. P. et al. Super-resolution single-photon imaging at 8.2 kilometers. *Opt. Express* **28**, 4076–4087 (2020).
15. Candès, E. J., Romberg, J. K. & Tao, T. Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.* **59**, 1207–1223 (2006).
16. Donoho, D. L. Compressed sensing. *IEEE Trans. Inf. Theory* **52**, 1289–1306 (2006).
17. Eldar, Y. C. & Kutyniok, G. *Compressed Sensing: Theory and Applications* (New York: Cambridge University Press, 2012).
18. Brady, D. J. et al. Compressive holography. *Opt. Express* **17**, 13040–13049 (2009).
19. Han, S. S. et al. A review of ghost imaging via sparsity constraints. *Appl. Sci.* **8**, 1379 (2018).
20. Bian, L. H. et al. Experimental comparison of single-pixel imaging algorithms. *J. Optical Soc. Am. A* **35**, 78–87 (2018).
21. Gong, W. L. & Han, S. S. Experimental investigation of the quality of lensless super-resolution ghost imaging via sparsity constraints. *Phys. Lett. A* **376**, 1519–1522 (2012).
22. Li, W. W. et al. Single-frame wide-field nanoscopy based on ghost imaging via sparsity constraints. *Optica* **6**, 1515–1523 (2019).
23. Amitonova, L. V. & de Boer, J. F. Endo-microscopy beyond the Abbe and Nyquist limits. *Light: Sci. Appl.* **9**, 81 (2020).
24. Sun, M. J. et al. Single-pixel three-dimensional imaging with time-based depth resolution. *Nat. Commun.* **7**, 12010 (2016).
25. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* (Cambridge: MIT Press, 2016).
26. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
27. Barbastathis, G., Ozcan, A. & Situ, G. On the use of deep learning for computational imaging. *Optica* **6**, 921–943 (2019).
28. Lyu, M. et al. Deep-learning-based ghost imaging. *Sci. Rep.* **7**, 17865 (2017).
29. He, Y. C. et al. Ghost imaging based on deep learning. *Sci. Rep.* **8**, 6469 (2018).
30. Wang, F. et al. Learning from simulation: an end-to-end deep-learning approach for computational ghost imaging. *Opt. Express* **27**, 25560–25572 (2019).
31. Higham, C. F. et al. Deep learning for real-time single-pixel video. *Sci. Rep.* **8**, 2369 (2018).
32. Lempitsky, V., Vedaldi, A. & Ulyanov, D. Deep image prior. *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (Salt Lake City, UT, USA: IEEE, 2018).
33. Dittmer, S. et al. Regularization by architecture: a deep prior approach for inverse problems. *J. Math. Imaging Vis.* **62**, 456–470 (2020).
34. Wang, F. et al. Phase imaging with an untrained neural network. *Light: Sci. Appl.* **9**, 77 (2020).
35. Bostan, E. et al. Deep phase decoder: self-calibrating phase microscopy with an untrained deep neural network. *Optica* **7**, 559–562 (2020).
36. Van Veen, D. et al. Compressed sensing with deep image prior and learned regularization. Preprint at arXiv: 1806.06438 (2018).
37. Heckel, R. & Soltanolkotabi, M. Compressive sensing with un-trained neural networks: gradient descent finds the smoothest approximation. *Proceedings of the 37th International Conference on Machine Learning* (eds Ill, Hal, D. and Singh, A.), **119**, 4149–4158 <http://proceedings.mlr.press/v119/heckel20a/heckel20a.pdf> (PMLR, 2020).
38. Zhou, K. C. & Horstmeyer, R. Diffraction tomography with a deep image prior. *Opt. Express* **28**, 12872–12896 (2020).
39. Gong, W. L. & Han, S. S. A method to improve the visibility of ghost images obtained by thermal light. *Phys. Lett. A* **374**, 1005–1008 (2010).
40. Ferri, F. et al. Differential ghost imaging. *Phys. Rev. Lett.* **104**, 253603 (2010).
41. Wang, C. L. et al. Airborne near infrared three-dimensional ghost imaging LiDAR via sparsity constraint. *Remote Sens.* **10**, 732 (2018).
42. Bromberg, Y., Katz, O. & Silberberg, Y. Ghost imaging with a single detector. *Phys. Rev. A* **79**, 053840 (2009).
43. Scully, M. O. & Zubairy, M. S. *Quantum Optics* (Cambridge University Press, Cambridge, 1997).
44. Deng, M. et al. On the interplay between physical and content priors in deep learning for computational imaging. *Opt. Express* **28**, 24152–24170 (2020).
45. Zhang, P. L. et al. Improving resolution by the second-order correlation of light fields. *Opt. Lett.* **34**, 1222–1224 (2009).
46. Wang, W. et al. Gerchberg-Saxton-like ghost imaging. *Opt. Express* **23**, 28416–28422 (2015).
47. Mangeat, T. et al. Super-resolved live-cell imaging using random illumination microscopy. *Cell Rep. Methods* **1**, 100009 (2021).
48. Yariv, A. & Yeh, P. *Photonics: Optical Electronics in Modern Communications* (Oxford: Oxford University Press, 2006).
49. Healey, G. E. & Kondepudy, R. Radiometric CCD camera calibration and noise estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* **16**, 267–276 (1994).
50. Goodman, J. W. *Statistical Optics* (New York: Wiley-Blackwell, 2000).
51. Howard, A. G. et al. MobileNets: efficient convolutional neural networks for mobile vision applications. Preprint at arXiv: 1704.04861v1 (2017).
52. Glorot, X. & Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. *J. Mach. Learn. Res.* **9**, 249–256 (2010).
53. Ruder, S. An overview of gradient descent optimization algorithms. Preprint at arXiv: 1609.04747v2 (2017).
54. Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. *Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention* (Munich, Germany: Springer, 2015).

