

# SCIENTIFIC REPORTS



Correction: Author Correction

OPEN

## Deep-learning-based ghost imaging

Meng Lyu<sup>1,2</sup>, Wei Wang<sup>1,2</sup>, Hao Wang<sup>1,2</sup>, Haichao Wang<sup>1,2</sup>, Guowei Li<sup>1,2</sup>, Ni Chen<sup>1,2</sup> & Guohai Situ<sup>1,2</sup>

Received: 25 September 2017

Accepted: 5 December 2017

Published online: 19 December 2017

In this manuscript, we propose a novel framework of computational ghost imaging, i.e., ghost imaging using deep learning (GIDL). With a set of images reconstructed using traditional GI and the corresponding ground-truth counterparts, a deep neural network was trained so that it can learn the sensing model and increase the quality image reconstruction. Moreover, detailed comparisons between the image reconstructed using deep learning and compressive sensing shows that the proposed GIDL has a much better performance in extremely low sampling rate. Numerical simulations and optical experiments were carried out for the demonstration of the proposed GIDL.

Ghost imaging was first demonstrated as a manifest of quantum entanglement<sup>1</sup> as biphoton source was used. But soon after that it has been demonstrated that the quantum source is not necessary<sup>2</sup>. Despite of the debate on the physics, GI has been demonstrated further by using pseudothermal light generated by dynamically modulating the illumination laser beam with a spatial light modulator (SLM)<sup>3</sup>. Although the source changes, the final image are mostly reconstructed using the correlation of signals from the image arm and the reference arm. The 'reference' arm now can be physically unexisted as its function can be calculated with the knowledge of the random phase patterns displayed on the SLM. And thus this technique comes with the term of computational ghost imaging (CGI)<sup>3</sup>. CGI has been used in the study of lensless imaging<sup>4</sup>, X-ray imaging<sup>5,6</sup> imaging in low light<sup>7</sup> and harsh environments<sup>8</sup>. However, the requirement of large number of measurements is one of the main issues that prevent it from practical applications<sup>9–12</sup>. Many efforts have been made to reduce the sampling rate. For example, non-computational<sup>13–15</sup> and computational methods have been proposed to increase image quality under low sampling rate<sup>9–11,16–20</sup>. In particular, compressive sensing GI (CSGI)<sup>10,16–19</sup> and iterative GI<sup>11,20</sup> model the problem of image reconstruction in GI as an optimization problem.

In this letter, we propose a new framework of CGI for high quality image reconstruction under low sampling condition. The proposed method uses deep learning (DL) and thus we term it Ghost imaging using deep learning (GIDL). DL is a machine learning technique for data modelling, and decision making with a neural network trained by a large amount of data<sup>21,22</sup>. The application of machine learning techniques in optical imaging was first proposed by Horisaki *et al.*<sup>23</sup> who used Support Vector Degression (SVG) architecture to learn the scatterer. In the last two years, we have witnessed the rapid development of the application of deep learning in solving various inverse problems in optical imaging. For example, people have used it in fluorescence lifetime imaging<sup>24</sup> phase imaging<sup>25,26</sup> and imaging through scattering media<sup>27,28</sup>. By combining GI and DL, we show in this manuscript that GIDL can also decrease the number of measurements significantly as CSGI, but with much better reconstruction. Also, detailed comparisons between the performances, including the image quality and the noise robustness, of CSGI and GIDL are discussed. Our analysis suggests that the GIDL promises great potentials in applications such as imaging and sensing through harsh environments.

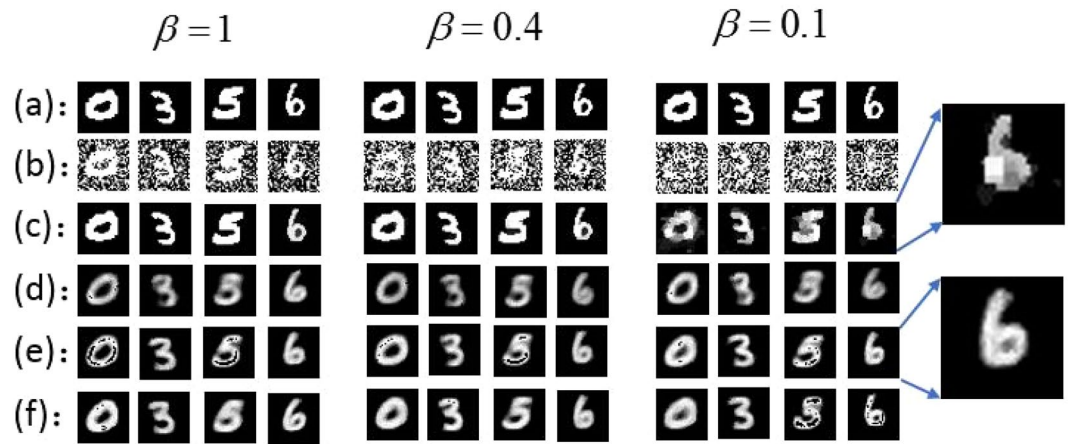
**Numerical Simulation.** In ghost imaging, the unknown object,  $T(x)$ , is illuminated by a sequence of speckle patterns,  $I_m(x)$ , where the subscript integer ( $m = 1 \dots M$ ) denotes the  $m^{\text{th}}$  illumination. Then, for the  $m^{\text{th}}$  speckle, the signal collected by a bucket detector can be written as  $S_m = \int I_m(x)T(x)dx$ . Traditionally, the image reconstructed using GI is obtained by the correlation of the signal fluctuation  $\delta S_m$  with the speckle patterns  $\delta I_m(x)$

$$O(x) = \langle \delta S_m \delta I_m \rangle. \quad (1)$$

In CGI, the speckle intensities  $I_m(x)$  are calculated numerically from the phase patterns displayed on the SLM.

<sup>1</sup>Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai, 201800, China.

<sup>2</sup>University of Chinese Academy of Sciences, Beijing, 100049, China. Meng Lyu and Wei Wang contributed equally to this work. Correspondence and requests for materials should be addressed to W.W. (email: wangweihust@126.com) or G.S. (email: ghsitu@siom.ac.cn)



**Figure 1.** Simulation results of GI, GICS and GIDL. (a) Top row: Ground truth objects, (b,c) Image reconstructed using GI and CSGI for different measurement ratios  $\beta$ . (d–f) Images reconstructed using GIDL (ghost imaging using neural networks) with the number of epochs 10, 100 and 500 respectively. Insets: zoomed in images reconstructed using CSGI and GIDL of the digit ‘6’.

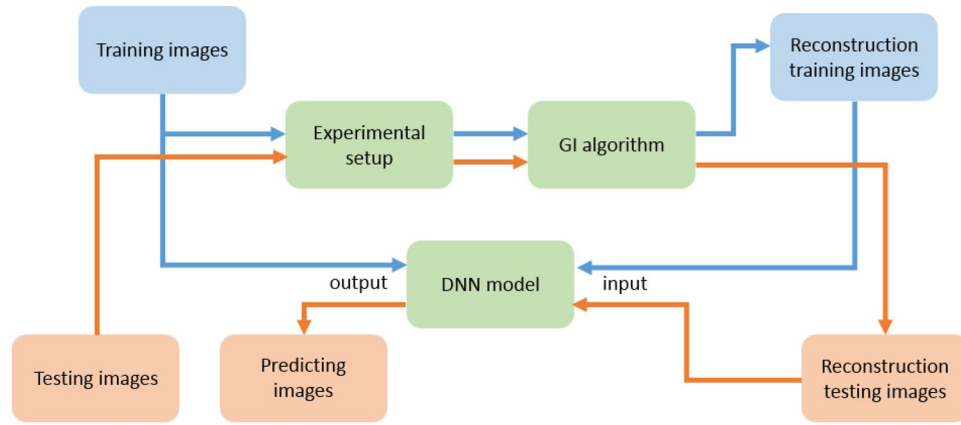
It has been demonstrated that the signal-to-noise ratio (SNR) of the image reconstructed in this way is proportional to the measurement ratio, i.e., the ratio between the number of illumination patterns  $M$  and the (average) number of speckle in each of these patterns  $N_{spec}$ <sup>9,11</sup>, namely,  $\beta = M/N_{spec}$ . To show how it works, we take the images (digits ‘0’, ‘3’, ‘5’ and ‘6’) shown in Fig. 1(a) as examples in our simulation study. These ground truth images have  $32 \times 32$  pixels. By using the algorithm defined by Eq. (1) one can reconstruct the images as shown in Fig. 1(b). The reconstructed image set in the columns are corresponding to the sampling ratio  $\beta = 1, 0.4$  and  $0.1$ , respectively. The results clearly suggest that, as the ratio  $\beta$  decreases from 1 to 0.1, the reconstructed images degrade significantly. The digits can be seen clearly when  $\beta = 1$ , although noise appears. But they are completely corrupted by noise when  $\beta = 0.1$ .

In order to increase the image quality, one usually sets  $\beta \gg 1$  in the conventional GI and CGI framework, so that the image acquisition procedure is very time-consuming. One popular solution to decrease the acquisition time is to combine GI and compressive sensing (CS) theory<sup>10,16,18</sup>. The CS theory allows the object to be recovered precisely from a smaller number measurements if it is sparse in a presentation domain<sup>29</sup>. So far several frameworks of CSGI have been demonstrated. But high-quality image reconstruction when  $\beta$  is small, i.e.,  $M \ll N_{spec}$ , is still a challenging problem<sup>10,16,18</sup>. In CSGI, one actually aims at solving the following inverse problem instead of calculating Eq. (1):

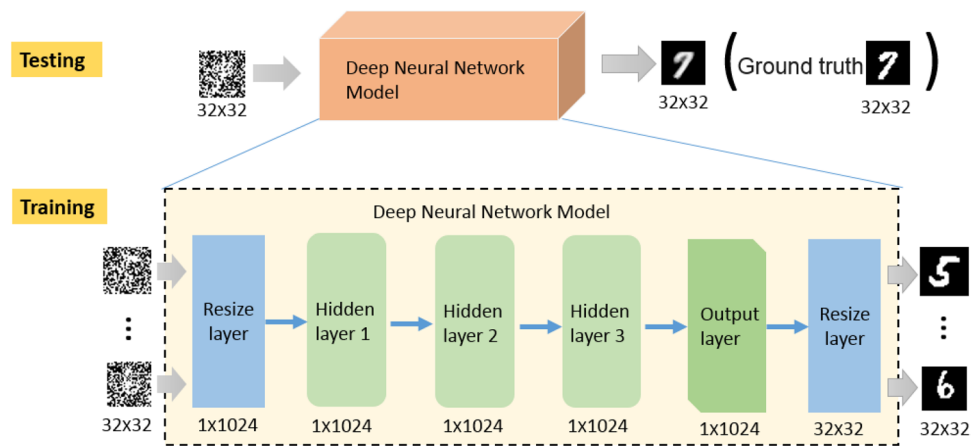
$$\min_T \left\| \nabla T \right\|_1 + \frac{u}{2} \left\| \mathbf{A}T - \mathbf{S} \right\|_2^2, \quad (2)$$

where  $\nabla T$  is the discrete gradient of  $T$ ,  $u$  is a weighting factor between the first term and the second term in Eq. (2), which represents the linear model between the image measurement matrix  $\mathbf{A}$  and the detected signal vector  $\mathbf{S} = [S_1, S_2, \dots, S_M]^T$ , where the symbol  $\perp$  denotes transposition. In this study, we solve Eq. (2) using the open source CS solver TVAL3<sup>30</sup> and reconstruct the images. The images reconstructed in this way are shown in Fig. 1(c). Because of the sparse constraint, the measurement ratio  $\beta$  for a good reconstruction of the object image can be decreased to 0.1 in our simulation. With a measurement ratio  $\beta = 0.4$ , the object can be recovered nearly precisely. However, the image reconstructed using a measurement ratio 0.1 is not so smooth due to the sparse regularization. This problem always exists in CS when the number of measurements is small<sup>29</sup>.

In the proposed scheme, the reconstruction is a two-step process. First, the image is reconstructed from the acquired data directly by solving Eq. (1). As shown in Fig. 1(b), the reconstructed image,  $O$ , in this way is usually very noisy when  $\beta$  is small. But the deep learning is then involved in the second step. The neural network attempts to reconstruct the object image  $T$  from the noisy, or even, corrupted,  $O$ . As schematically shown in Figs 2 and 3, the image reconstruction procedures of GIDL is also consist of two steps: training and testing. In the training step, we used a set of 2000 handwritten digits of  $32 \times 32$  pixels in size from the MNIST handwritten digit database<sup>31</sup> to train the network in our experiments. Some of the digits are shown in Fig. 3. To train the network, we first reconstructed the images of the digits in the training set according to Eq. (1). Then we fed these images together with the corresponding ground-truth digits into the neural network, and optimize the weighting factors that connect every two neurons in two neighboring hidden layers. In this work, we used a deep neural network (DNN) model with two reshaping layers, three hidden layers and one output layer. For demonstration, we used a very simple model. The reshaping layer at the input end shapes the  $32 \times 32$  input speckle pattern into a  $1 \times 1024$  vector. All the hidden layers and the output layer have 1024 neurons. The activate function of these neurons is rectified linear units (ReLU) which allow for faster and effective training of deep neural architectures on large and complex datasets compared with the sigmoid function<sup>32</sup>. The reshaping layer at the output end reshapes the  $1 \times 1024$  vector back to the  $32 \times 32$  image. The loss function and optimization in the DNN model is mean square error (MSE) and



**Figure 2.** The flowchart of GI using deep neural networks. The blue part represents the training stage and the orange part represents the testing stage.



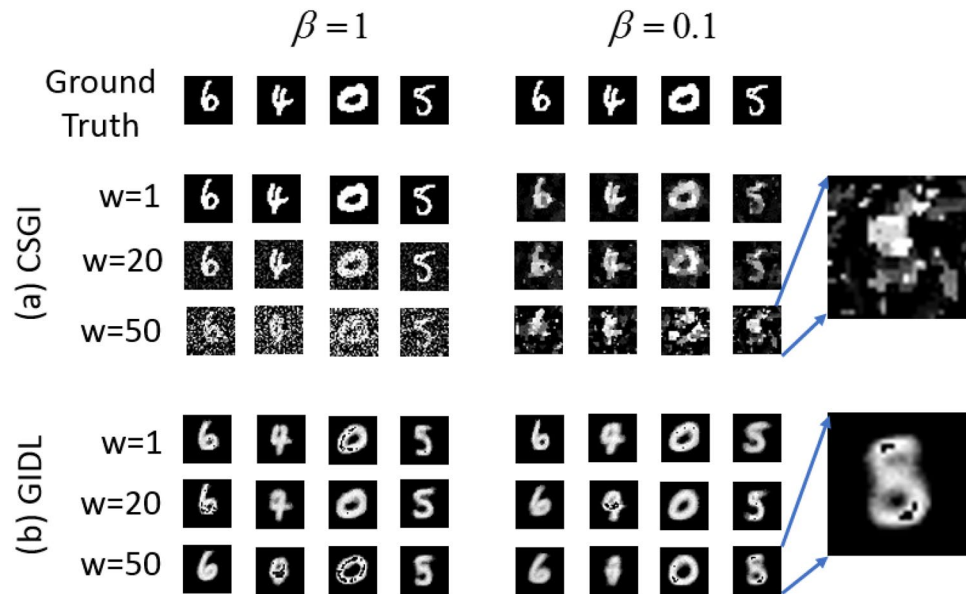
**Figure 3.** Framework of GI using deep neural networks.

stochastic gradient descent (SGD). Once the training is finished (after 500 epochs in our experiments), the DNN can be used to reconstruct the object image  $T$  from  $O$ . The program was implemented using Python version 3.5 and the DNN was implemented using Keras framework based on TensorFlow. The GPU-chip NVIDIA Tesla K20c was used to accelerate the computation.

The simulation results plotted in Fig. 1(d–f) show the reconstructed images using GIDL after 10, 100 and 500 training epochs, respectively, for different measurement ratios  $\beta$ . From these images we can conclude that: First, as the number of iteration (epoch) increases, the DNN model is better optimized. As a consequence, the reconstructed images becomes clearer and brighter. However, when the number of epoch becomes too large, we observed over-fitting of the data, which yield bit error in the reconstructed images as evidenced by the black spots. Second, the GIDL is not very sensitive to  $\beta$ . The MSE values between the images in Fig. 1(f) and the corresponding ground truth images in Fig. 1(a) are all around 0.03 even when  $\beta = 0.1$ . This means that by using GIDL for image reconstruction, one can significantly reduce the number of measurements in the GI acquisition procedure. As a consequence, the time efficiency can be improved without sacrifice of image quality. We note that one can achieve the reduction of measurement by using the CSGI framework as well<sup>10</sup>. However, when one takes a closed look at the zoomed-in images of any of the reconstructed digits, say, digit ‘6’, in the inset of Fig. 1, it is clearly seen that the image reconstructed using CSGI is not so smooth because of the regulation, while GIDL gives much better reconstruction. This is one major difference between the images reconstructed using GIDL and CSGI.

An additional advantage of GIDL over other GI frameworks is its robustness against noise. Now we provide a theoretical analysis. For a sufficiently large number of photons, the observed signal  $S_m$  by the single pixel camera can be represented by an additive random Gaussian noise<sup>18</sup>

$$S_m = \int I_m(x)T(x)dx + w_{\sigma_m}\varepsilon_m, \tag{3}$$



**Figure 4.** Noise robustness of GIDL. (a,b) Images reconstructed using CSGI and GIDL under different levels of detection noises. Inset: zoomed in images of the digit object 5 reconstructed using CSGI and GIDL for a high noise level and a low measurement ratio.

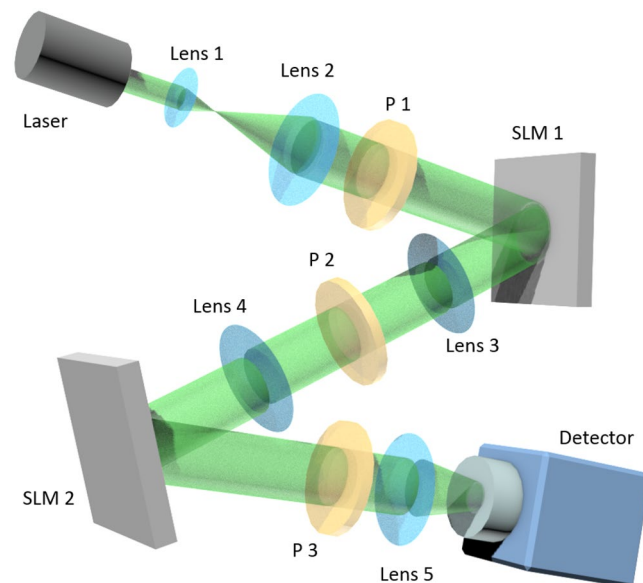
where the variance  $w^2\sigma_m^2 = w^2 \int I_m(x)T(x)dx$ , and  $\varepsilon_m$  is the standard Gaussian white noise. In the variance,  $w$  represents the noise level. A larger value of  $w$  will result in a worse detection image. For speckle field illumination of the same statistics,  $\sigma_m^2$  can be regarded as invariant so that one can replace it by a constant value,  $\sigma^2 \simeq \Sigma\sigma_m^2/M$ .

The simulation results are shown in Fig. 4. Figure 4(a) shows the images reconstructed using CSGI under different levels of detection noises. For CSGI, when the noise level  $w$  is small ( $w=1$ ), the reconstructed images are close to the ground truths, meaning that CSGI can tolerate low level noises. But as the noise increases to a certain level, CSGI fails. The reconstructed image quality is also influenced by the measurement ratio  $\beta$  in CSGI. For the case of  $w=50$  and  $\beta=0.1$ , the reconstructed images by CSGI are totally corrupted by noise according to our simulation.

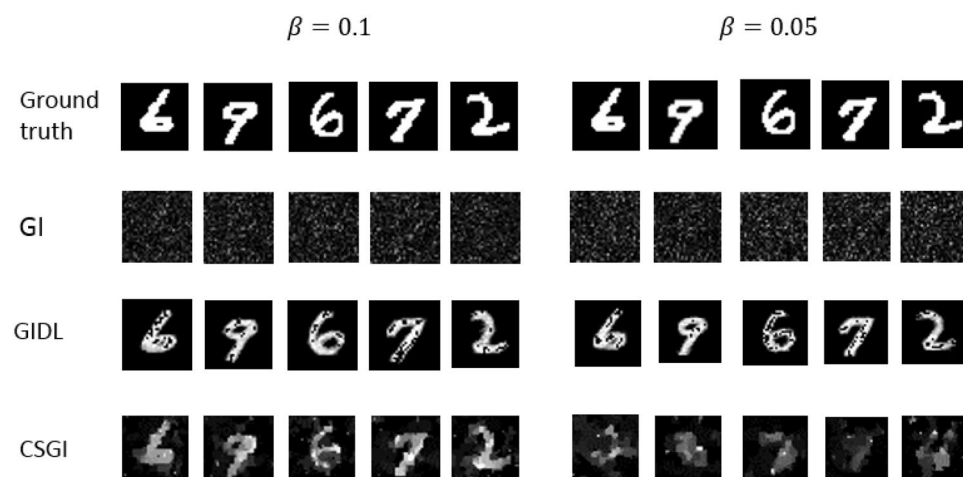
In contrast, GIDL has a much better performance. Figure 4(b) shows the images reconstructed using GIDL under different levels of detection noises. In consist with Fig. 2(d–f), all the images in Fig. 4(b) are smooth as compared to the ones reconstructed using CGSI. The inset shows the zoomed-in images of the digit object ‘5’ reconstructed using CSGI and GIDL for a high noise level and a low measurement ratio ( $\beta=0.1$ ). It is seen that the feature details of the digit ‘5’ is clearly recognizable in the image reconstructed by GIDL while it is not by CSGI. This demonstrates the advantage of GIDL over CSGI for imaging and sensing in harsh environments. Although the sparse constraint can be used to decrease the influence of the random detection noise to a certain level, CSGI can not work for high noise levels in which case the linear model Eq. (2) is affected severely. In contrast, in GIDL, the deep learning architecture takes all the noise into account in building up the network model and fits sharply all the partially reconstructed  $O$  to the corresponding object image  $T$ . However, when the noise level keep increasing together with the reduction of  $\beta$ , the effect of the additive noise cannot be ignored completely. As shown by the digit images ‘4’ and ‘5’ in Fig. 4(b), the reconstructed image becomes blurred, distorting the feature of the object.

**Experiment.** Now we demonstrate the proposed GIDL using some proof-of-principle experiments. We adopted a setup of ghost imaging as the one illustrated in Fig. 5. A laser beam with the wavelength  $\lambda = 532 \pm 2$  nm (Verdi G2 SLM, Coherent, Inc.) was expanded using a 4f system consisting of lens 1 and lens 2. An SLM 1 (Pluto-Vis, Holoeye Photonics AG) was used to subsequently display the phase distributions that generate speckle illumination  $I_m$ , whereas the objects were displayed onto an SLM 2 (Pluto-Vis, Holoeye Photonics AG). The collimated laser beam shone onto SLM 1 and was modulated by the speckle displayed on it. The beam reflected from it was projected onto SLM 2 using the other 4f system consisting of lens 3 and lens 4. In the setup, P1, P2 and P3 are linear polarizers. P1 and P3 are vertically polarized, and P2 is horizontally polarized, with respect to the laboratory coordinate, so that to achieve amplitude-only modulation for the SLMs. We displayed different digits from the MNIST database<sup>31</sup> on SLM 2, serving as the objects in our experiments. The beam reflected from SLM 2 was collected using a sCMOS camera (Zyla 4.2 PLUS sCMOS, Andor Technology Ltd.) because we do not have a bucket detector. We integrated each acquired intensit patterns to produce  $S_m$ . This does not affect the experimental results except the frame rate and signal amplification because the integration of a recorded intensity pattern acquired by the camera is proportional to the optical power.

In the experiments, we used the same training set and testing set as in the simulation. 2000 different digit images in the training set were used to train the network. To demonstrate the proposed scheme, we acquired a very small amount of data to reconstruct the testing digits. In order to speed up the convergence of the DNN model, we used the optimization Adam, and an algorithm for first-order gradient-based optimization of



**Figure 5.** Schematic setup of ghost imaging. P1, P2 and P3 are linear polarizers.



**Figure 6.** Experimental results under  $\beta=0.1$  and  $\beta=0.05$ . The images in the first row are the ground truth, the second row shows the images reconstructed using GI, the third row shows the predicted objects using GIDL, and the last row are the images using CSGI.

stochastic objective functions<sup>33</sup>, instead of SGD, in the training. The experimental results for  $\beta=0.1$  and  $\beta=0.05$  are plotted in Fig. 6. In this figure, the images in the first row are the ground truth images of four digits in the testing set. Due to the small  $\beta$  and noise in the system, the reconstructed images using the conventional GI are corrupted by noise as shown in the second row in Fig. 6. One cannot recognize any the visible feature about the target digits from these reconstructed images. However, when we sent them into the trained DNN model, we obtained the corresponding images shown in the third row in Fig. 6. Although they do not resemble exactly the ground truths, the image reconstructed by GIDL contain enough features to recognize. In contrast, the images reconstructed using CSGI are still recognizable at the measurement ratio  $\beta=0.1$ , but become totally corrupted when  $\beta=0.05$ . This suggests that GIDL has a better performance than CSGI at low measurement ratio.

## Conclusion

In conclusion, we have demonstrated the novel technique of GIDL using both numerical and optical experiments. We have analyzed the performance of conventional GI, CSGI and GIDL under different noise and measurement ratio conditions, and observed that GIDL has much better performance than the others especially when the measurement ratio  $\beta$  is small. This allows the significant reduction of data acquisition time in ghost imaging, giving a promising solution to this challenges that prohibits GI from practical applications. What's more, our study opens up new possibility for artificial intellectual techniques in the applications of ghost imaging, and in an even more sense, computational imaging.

## References

- Pittman, T. B., Shih, Y. H., Strekalov, D. V. & Sergienko, A. V. Optical imaging by means of two-photon quantum entanglement. *Phys. Rev. A* **52**, R3429 (1995).
- Valencia, A., Scarcelli, G., D'Angelo, M. & Shih, Y. Two-photon imaging with thermal light. *Phys. Rev. Lett.* **94**, 063601 (2005).
- Shapiro, J. H. Computational ghost imaging. *Phys. Rev. A* **78**, 061802 (2008).
- Basano, L. & Ottonello, P. Experiment in lensless ghost imaging with thermal light. *Appl. Phys. Lett.* **89**, 091109 (2006).
- Yu, H. *et al.* Fourier-transform ghost imaging with hard X rays. *Phys. Rev. Lett.* **117**, 113901 (2016).
- Pelliccia, D., Rack, A., Scheel, M., Cantelli, V. & Paganin, D. M. Experimental x-ray ghost imaging. *Phys. Rev. Lett.* **117**, 113902 (2016).
- Morris, P. A., Aspden, R. S., Bell, J. E., Boyd, R. W. & Padgett, M. J. Imaging with a small number of photons. *Nature Commun.* **6** (2015).
- Cheng, J. Ghost imaging through turbulent atmosphere. *Opt. Express* **17**, 7916 (2009).
- Ferri, F., Magatti, D., Lugiato, L. & Gatti, A. Differential ghost imaging. *Phys. Rev. Lett.* **104**, 253603 (2010).
- Katz, O., Bromberg, Y. & Silberberg, Y. Compressive ghost imaging. *Appl. Phys. Lett.* **95**, 131110 (2009).
- Wang, W. *et al.* Gerchberg-Saxton-like ghost imaging. *Opt. Express* **23**, 28416 (2015).
- Li, J. *et al.* Negative exponential behavior of image mutual information for pseudo-thermal light ghost imaging: observation, modeling, and verification. *Science Bulletin* **62**, 717–723 (2017).
- Scarcelli, G., Berardi, V. & Shih, Y. Phase-conjugate mirror via two-photon thermal light imaging. *Appl. Phys. Lett.* **88**, 061106 (2006).
- Chen, H., Peng, T. & Shih, Y. 100% correlation of chaotic thermal light. *Phys. Rev. A* **88**, 023808 (2013).
- Sun, M. J., Li, M. F. & Wu, L. A. Nonlocal imaging of a reflective object using positive and negative correlations. *Appl. Opt.* **54**, 7494–7499 (2015).
- Zhao, C. *et al.* Ghost imaging lidar via sparsity constraints. *Appl. Phys. Lett.* **101**, 141123 (2012).
- Hu, X., Suo, J., Yue, T., Bian, L. & Dai, Q. Patch-primitive driven compressive ghost imaging. *Opt. Express* **23**, 11092 (2015).
- Katkovnik, V. & Astola, J. Compressive sensing computational ghost imaging. *J. Opt. Soc. Am. A* **29**, 1556 (2012).
- Wu, H., Zhang, X., Gan, J., Luo, C. & Ge, P. High-quality correspondence imaging based on sorting and compressive sensing technique. *Laser Phys. Lett.* **13**, 115205 (2016).
- Wang, W., Wang, Y. P., Li, J., Yang, X. & Wu, Y. Iterative ghost imaging. *Opt. Lett.* **39**, 5150 (2014).
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. Imagenet classification with deep convolutional neural networks. In *Adv. Neural Info. Process. Sys.* 1097–1105 (2012).
- LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436 (2015).
- Horisaki, R., Takagi, R. & Tanida, J. Learning-based imaging through scattering media. *Opt. Express* **24**, 13738 (2016).
- Wu, G., Nowotny, T., Zhang, Y., Yu, H.-Q. & Li, D. D.-U. Artificial neural network approaches for fluorescence lifetime imaging techniques. *Opt. Lett.* **41**, 2561 (2016).
- Sinha, A., Lee, J., Li, S. & Barbastathis, G. Lensless computational imaging through deep learning. *Optica* **4**, 1117–1125 (2017).
- Rivenson, Y., Zhang, Y., Gunaydin, H., Teng, D. & Ozcan, A. Phase recovery and holographic image reconstruction using deep learning in neural networks. *arXiv preprint arXiv:1705.04286* (2017).
- Lyu, M., Wang, H., Li, G. & Situ, G. Exploit imaging through opaque wall via deep learning. *arXiv preprint, arXiv:1708.07881* (2017).
- Li, S., Deng, M., Lee, J., Sinha, A. & Barbastathis, G. Imaging through glass diffusers using densely connected convolutional networks. *arXiv preprint, arXiv:1711.06810* (2017).
- Candes, E. J., Romberg, J. K. & Tao, T. Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.* **59**, 1207 (2006).
- Li, C., Yin, W. & Zhang, Y. User's guide for TVL3: TV minimization by augmented lagrangian and alternating direction algorithms. *CAAM Rep.* **20**, 46 (2009).
- Deng, L. The MNIST database of handwritten digit images for machine learning research. *IEEE Signal Process. Mag.* **29**, 141 (2012).
- Vinod, N., & Geoffrey, H. Rectified linear units improve restricted boltzmann machines. *Proc. Int. Conf. Mach. Learn.* 807–814 (2010).
- Kingma, D. & Ba, J. Adam: A method for stochastic optimization. *arXiv preprint, arXiv:1412.6980* (2014).

## Acknowledgements

This project was funded by the Key Research Program of Frontier Sciences, Chinese Academy of Sciences (QYZDB-SSW-JSC002), the National Natural Science Foundation of China (61377005, 61327902), the China Postdoctoral Science Foundation (2015M580356), and the Natural Science Foundation of Shanghai (No. 17ZR1433800).

## Author Contributions

G.S. and W.W. conceived the idea. M.L. conducted the experiments. W.W. performed numerical simulation. M.L., W.W., H.W. analysed the results with the discussion with H.W., G.L., and N.C. W.W., M.L. and G.S. wrote the manuscript. G.S. supervised the project.

## Additional Information

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**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/>.

© The Author(s) 2017