Does deep learning always outperform simple linear regression in optical imaging?

SHUMING JIAO,1 YANG GAO,1 JUN FENG,1 TING LEI,1,2 XIAOCONG YUAN1,3
1Nanophotonics Research Center, Shenzhen University, Shenzhen, Guangdong, China
2leiting@szu.edu.cn
3xcyuan@szu.edu.cn

Abstract: Deep learning has been extensively applied in many optical imaging applications in recent years. Despite the success, the limitations and drawbacks of deep learning in optical imaging have been seldom investigated. In this work, we show that conventional linear-regression-based methods can outperform the previously proposed deep learning approaches for two black-box optical imaging problems in some extent. Deep learning demonstrates its weakness especially when the number of training samples is small. The advantages and disadvantages of linear-regression-based methods and deep learning are analyzed and compared. Since many optical systems are essentially linear, a deep learning network containing many nonlinearity functions sometimes may not be the most suitable option.

© 2019 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

In recent years, deep learning receives much attention in many research fields including optical design [1,2] and optical imaging [3]. In previous works, deep learning has been extensively applied for many optical imaging problems including phase retrieval [4-7], microscopic image enhancement [8-9], scattering imaging [10-11], holography [12-18], single-pixel imaging [19,20], super-resolution [21-23], Fourier ptychography [24-26], optical interferometry [27,28], wavefront sensing [29,30], and optical fiber communications [31].

Despite the success, deep learning has its own limitations and drawbacks, like any other approach [32]. For example, a huge number of training samples is usually required to train a deep neural network, which may not be always available in practical applications. The optimization of connection weights in the network with many training samples takes a considerable amount of computational cost. The design of network structure and the tuning of network parameters are often implemented empirically and intuitively with weak explainability. A deep neural network trained and tested for one category of samples may fail to work when it is generalized to other different testing samples. In fact, it is likely that deep learning may perform worse than other machine-learning (or non-machine-learning) methods in certain application scenarios. In previous works, the deficiencies of deep learning in solving optical imaging problems, compared with other methods, were seldom investigated.

In recent works [33,34], deep learning has been employed to address the problems of attacking a random-phase-encoded optical cryptosystem [33] and blind reconstruction for single-pixel imaging [34]. A random-phase-encoded optical cryptosystem is a coherent imaging system with multiple diffractive optical elements such as lens and random phase masks. The input plaintext light field is sequentially modulated by each phase mask in the forward propagation and it is finally transformed to a ciphertext light field as the system output. The objective of attacking a optical cryptosystem is to recover the input image from the given output light field if the encoding of all the phase masks is unknown. In single-pixel imaging [35], the object image is sequentially illuminated by different structured light intensity patterns and the total light intensity of the entire object scene is recorded by a single-pixel detector for each pattern. Finally, the object image can be computationally reconstructed when both the illumination patterns and single-pixel intensity sequence are
known. However, in a blind reconstruction \cite{34}, the objective is to recover the object image from the intensity sequence when the encoding of all the illumination patterns is unknown.

The two systems \cite{33,34} stated above are both linear and can be regarded as a black box when the encoding of elements (phase masks or illumination patterns) is unknown. The random-phase-encoded optical cryptosystem is usually coherent while the single-pixel imaging is usually incoherent. In the previous work \cite{36}, it is shown that a multiple-phase-mask diffractive system and a single-pixel imaging system are similar from several aspects such as performing optical pattern recognition. In the works \cite{33,34}, each system is modeled by a different deep learning network optimized with many pairs of input and output training samples. Then the input image can be predicted by the network from an arbitrary given output. Since there is a linear relationship between the input and output of the two optical imaging systems mathematically, we point out that simple linear-regression-based methods can produce the same results as deep learning. A linear regression scheme can recover the object image more efficiently than deep learning for these two problems in some extent. The advantages and disadvantages of linear-regression-based methods and deep learning are analyzed and compared.

This paper is structured as follows. The linear regression model is described in Section 2. The two black-box optical imaging problems, i.e. attacking a random-phase-encoding-based optical cryptosystem and blind reconstruction in single-pixel imaging, are described in Section 3 and Section 4. The results and discussions about the comparison between linear-regression-based methods and deep learning are given in Section 5. A final conclusion is made in Section 6.

2. Linear Regression Model

For a linear optical imaging system, both the input X and output Y can be denoted by a column vector \(X = [x_1 x_2 \ldots x_M]^T\) and \(Y = [y_1 y_2 \ldots y_N]^T\). It is assumed that the input X has totally M pixels and the output Y has totally N pixels. The relationship between Y and X can be modeled as a matrix multiplication \(Y = WX\), given by Equation (1). The weighting matrix W consisting of \(M \times N\) elements can be employed to model a black-box optical system.

\[
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_N
\end{bmatrix} =
\begin{bmatrix}
w_{11} & \cdots & w_{1M} \\
\vdots & \ddots & \vdots \\
w_{N1} & \cdots & w_{NM}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_M
\end{bmatrix}
\tag{1}
\]

When a large set of training samples (many pairs of X and Y) is available, the elements in the matrix W can be estimated by optimization if they are not given. The elements in W can be iteratively optimized with a gradient descent algorithm. Initially, all the elements in W are set to be random values. Then each element in W can be updated in the following way based on the gradient descent for each training sample: \(w_{nm}' = w_{nm} + r(y_n - y^\prime_n)x_m\) \((1 \leq m \leq M, 1 \leq n \leq N)\), where \(y^\prime_n\) denotes the actual output generated from the input of one training sample by multiplying the current W, \(y_n\) denotes the target output of one training sample and r denotes the pre-defined learning rate. If the input values are complex-amplitude instead of real intensities, the algorithm needs to be slightly modified as: \(w_{nm}' = w_{nm} + r(y_n - y^\prime_n)\text{conj}(x_m)\) \((1 \leq m \leq M, 1 \leq n \leq N)\), where \(\text{conj()}\) denotes the conjugate of a complex value. After many iterations, the adaptively optimized W matrix multiplied with a given input will yield an output close to the target one. A linear regression model can be considered as one-layer fully connected neural network without nonlinear activation functions, shown in Fig. 1. In a true fully connected neural network shown in Fig. 2(b), both linear connections and nonlinear activation functions are densely interconnected in the form of multiple cascaded layers between the network input and output. Modern deep learning networks such as the ones proposed in the previous works \cite{33,34} shown in Fig. 3 and Fig. 5 usually have even more complicated structures than a fully connected neural network. Compared with a deep learning network, a linear regression model has very low complexity.
3. Problem 1: attacking a random-phase-encoding-based optical cryptosystem

As proposed in many previous works [33,37,38], an optical image encryption system can be constructed with an optical setup containing cascaded lens and random phase masks. Typical examples include Double Random Phase Encryption (DRPE) and Triple Random Phase Encryption (TRPE) [33]. In this work, the one with a more complicated structure, i.e. a TRPE system, is considered and its optical setup is shown in Fig. 2. In a TRPE system, the pixel intensities of the input light field represent the plaintext image O. Then the input light field is optically Fourier transformed and inverse Fourier transformed with a double-lens 4f setup. The light field in the output plane becomes the ciphertext C. The plaintext image can be decrypted from the ciphertext with the same setup by backforward light field propagation. Three random phase mask R₁, R₂ and R₃ are placed in the input plane, Fourier plane and output plane respectively. The pixel values of all the phase masks are encoded as random phases between [0 2π]. The three phase masks serve as the encryption and decryption key. The mathematical model of TRPE encryption and decryption is given by Equation (2) and (3).

\[
C = IFT \left\{ FT \left\{ O \cdot \exp \left( -i \cdot R_1 \right) \right\} \cdot \exp \left( -i \cdot R_2 \right) \right\} \cdot \exp \left( -i \cdot R_3 \right)
\]

\[
O = IFT \left\{ FT \left\{ C \cdot \exp \left( -i \cdot R_1 \right) \right\} \cdot \exp \left( -i \cdot R_2 \right) \right\} \cdot \exp \left( -i \cdot R_3 \right)
\]

where FT and IFT denotes Fourier transform and inverse Fourier transform, O denotes the input plaintext image and C denotes the encrypted light field (ciphertext).

Ideally, the plaintext image O cannot be recovered from the ciphertext C if the key is not known and the information security is protected in this way. However, the encryption system can be cracked by a known-plaintext attack (KPA) if the attacker collects an adequate number of plaintext-ciphertext pairs. In KPA, the objective is to recover the plaintext O from the corresponding ciphertext C without knowing R₁, R₂ and R₃. The
entire system is linear and the ciphertext can be regarded as the input vector $X$ and the plaintext can be regarded as the output $Y$ in the linear regression model described in Section 2. The two-dimensional matrices $O$ and $C$ can be rearranged as one-dimensional vectors $X$ and $Y$. Consequently, a KPA to a TRPE system can be implemented with complex-amplitude linear regression (CLR), in addition to deep learning. In the previous work [33], the deep learning network structure shown in Fig. 3, referred to as DecNet, was employed for the KPA. In this work, CLR is compared with DecNet for the same KPA attack to a TRPE system.

![Deep learning network for attacking a TRPE system proposed in the previous work][33] (DecNet)

4. Problem 2: blind reconstruction in single-pixel imaging

In single-pixel imaging (SPI), the light intensity is recorded by a sensor containing only one single pixel, instead of a pixelated sensor array. A typical optical setup for a SPI system is shown in Fig. 4.

![Optical setup of a single-pixel imaging system](image)

The two-dimensional object image $O(x, y)$ is sequentially illuminated by $N$ varying two-dimensional structured light patterns $P_n(x, y)$ $(1 \leq n \leq N)$ and a single-pixel intensity sequence $I_n$ $(1 \leq n \leq N)$ will be recorded. Mathematically, each element in $I_n$ is the inner product between $O(x, y)$ and each pattern in $P_n(x, y)$. The object image $O(x, y)$ can be computationally reconstructed when both the illumination pattern sequence $P_n(x, y)$ and the recorded intensity sequence $I_n$ are known. It is assumed that the total number of pixels in $O(x, y)$ and $P_n(x, y)$ is $M$. The sampling ratio $S$ can be defined as $N/M$. In single-pixel imaging, various kinds of algorithms can be employed to reconstruct $O(x, y)$ from $P_n(x, y)$ and $I_n$ [39]. However, all the illumination patterns $P_n(x, y)$ are required to be known in these reconstruction algorithms. It is usually easier to reconstruct a high-quality object image when the sampling ratio $S$ is higher. A blind reconstruction in SPI by
deep learning was attempted in the previous work [34], where the object image \( O(x, y) \) is recovered from only \( I_n \) when the patterns \( P_n(x, y) \) are not given. The blind reconstruction in SPI is favorable for some applications such as scattering imaging [34, 40]. It is assumed that multiple pairs of different object images and single-pixel intensities are given for the fixed illumination patterns, which can be used as training samples in deep learning.

The blind reconstruction in SPI essentially contains two steps: (a) Recovery of the unknown illumination patterns \( P_n(x, y) \) from the training samples. This is similar to the KPA in random-phase-encoding-based optical encryption described in Section 3; (b) Object image reconstruction in SPI from a given \( I_n \) and the estimated \( P_n(x, y) \) obtained in Step (a). The deep learning approach in the previous work [34] is end-to-end and both two steps are realized within the network shown in Fig. 5.

![Deep learning network for blind image reconstruction in SPI proposed in the previous work [34] (Wang’s Net)](image)

In SPI, \( O(x, y) \) and \( I_n \) have a linear mathematical relationship. The \( M \) pixels in \( O(x, y) \) can be rearranged as the one-dimensional input vector \( X \) in Equation (1) and \( I_n \) is equivalent to the output vector \( Y \) in Equation (1). All the \( N \) illumination patterns \( P_n(x, y) \) will jointly constitute the weighting matrix \( W \) in Equation (1) and each pattern corresponds to one row in \( W \). Consequently, the unknown illumination patterns can be recovered from the training samples by linear regression for Step (a). Then a compressive sensing scheme with total variation minimization [39,41,42] can be employed for image reconstruction in Step (b) in this work. No training samples are required for compressive sensing reconstruction since it is not a machine learning process. Our proposed scheme is referred to as “Linear Regression + Compressive Sensing (LRCS)”. The LRCS scheme is compared with the deep learning network proposed in the previous work [34], refere to as Wang’s Net. It shall be noted that no linear regression is performed to recover the illumination patterns in the previous work [34], even though compressive sensing is adopted for image reconstruction by assuming the illumination patterns are known.

5. Results and discussion

5.1 Attacking a random-phase-encoding-based optical cryptosystem

A complex-amplitude linear regression (CLR) is performed to crack a TRPE optical cryptosystem. For comparison, a DecNet [33] is constructed and the corresponding cracking results are obtained as well. The size of plaintext image, ciphertext and random phase masks is \( 32 \times 32 \) pixels. Plaintext images are randomly selected from the number-digit images in the MNIST dataset [43] and the fashion images in the Fashion-MNIST dataset [44]. The output ciphertext light fields corresponding to the plaintext images are generated from a simulated TRPE system. In the training, plaintext images are used as the target output and complex-amplitude ciphertexts are used as the input for both CLR and DecNet. Various
number of training samples are attempted: 50, 100, 200, 500, 2000 and 5000. In addition, 200 samples randomly selected from each dataset different from the training samples are employed to test the attacking capability of the CLR and DecNet after training. The peak-signal-to-noise-ratio (PSNR) between the original plaintext image and the recovered result from the ciphertext by these two methods is employed to evaluate their performance.

In CLR, the learning rate is 0.01 for the MNIST dataset and 0.001 for the Fashion-MNIST dataset. The number of iterations is set to be 300. The results of our complex-amplitude linear regression are compared with the ones using deep learning [33] in Table 1, Table 2 and Fig. 3. It can be observed that the performance of both methods will be improved as the number of training samples increases. However, CLR performs much better than DecNet when the number of training samples is small (e.g. from 50 to 500).

The DecNet can only yield satisfactory output results when the number of training samples is at least 2000 or 5000. The results from CLR with 200 training samples is close to the results from DecNet with 5000 samples. Evidently, CLR has significant advantages compared with DecNet in attacking a TRPE system when the number of training samples is inadequate. It can be observed from Fig. 3 that the recovered MNIST images by DecNet are contaminated with stripe noise and the recovered Fashion-MNIST images by DecNet are heavily blurred when the number of training samples is small. Theoretically, it is possible for a deep learning network like DecNet to accurately model a linear system. However, the network may be overfitted at a local optimal point in the training when the number of training samples is small. Since the global optimal solution is not reached, the network may yield unfavorable prediction results for the testing images.

In the previous work [33], the DecNet can still work when only ciphertext intensities are available as the network input, instead of both ciphertext intensities and phases. However, CLR will not work in this situation since the input-output relationship is no longer linear. This is one major limitation of CLR compared with DecNet.

---

Table 1 Comparison of recovered plaintext image quality with complex-amplitude linear regression (CLR) and DecNet for the MNIST dataset

| Number of training samples K | CLR   | DecNet |
|------------------------------|-------|--------|
| 50                           | 17.4439 dB | 11.5822 dB |
| 100                          | 19.4437 dB | 11.6669 dB |
| 200                          | 22.4502 dB | 12.0497 dB |
| 500                          | 27.2989 dB | 13.0242 dB |
| 2000                         | ----     | 13.6974 dB |
| 5000                         | ----     | 21.3453 dB |

Table 2 Comparison of recovered plaintext image quality with complex-amplitude linear regression (CLR) and DecNet for the Fashion-MNIST dataset

| Number of training samples K | CLR   | DecNet |
|------------------------------|-------|--------|
| 50                           | 15.8603 dB | 11.1401 dB |
| 100                          | 17.7190 dB | 12.1886 dB |
| 200                          | 19.8130 dB | 15.3303 dB |
| 500                          | 22.9465 dB | 17.4687 dB |
| 2000                         | ----     | 19.3498 dB |
| 5000                         | ----     | 20.6077 dB |
The similarities and differences between the proposed CLR scheme and the previously proposed DecNet for attacking a TRPE system are summarized in Table 3.

**Table 3 Summary of similarities and differences between complex-amplitude linear regression (CLR) and DecNet for attacking a TRPE system**

| Similarity | Complex-amplitude Linear Regression (CLR) | DecNet |
|------------|-----------------------------------------|--------|
| Predict the plaintext image from the ciphertext light field for a TRPE system (or other similar optical cryptosystems) after the model is trained with a certain number of training samples | (1) Work with a small number of training samples | (1) Only work with a large number of training samples |
| | (2) Simple explicit model with few parameters | (2) Complicated black-box model with many parameters for tuning |
| | (3) Not work for intensity-only ciphertext | (3) Work for intensity-only ciphertext |

### 5.2 Blind Reconstruction in Single-pixel Imaging

In the simulation, the size of object image and each illumination pattern is $32 \times 32$ pixels. The pixel intensity values in each illumination pattern are randomly distributed between 0 and 1. Four different numbers of illuminations, $N=51$, $N=205$, $N=410$ and $N=1024$ corresponding to four different sampling ratios $S=0.05$, $S=0.2$, $S=0.4$ and $S=1$, are attempted. Various number of training images and 200 testing images are randomly selected from the MNIST dataset. The single-pixel intensity values can be obtained based on the SPI model described in Section 4. Both our proposed “linear regression + compressive sensing” (LRCS) scheme and Wang’s Net proposed in the previous work [34] are implemented to recover the original object image. In the linear regression step of LRCS, the learning rate is set to be 0.01 and the number of iterations is 300 for all the cases. The average PSNR of the blindly reconstructed images from simulated single-pixel intensity values for the 200 testing samples is presented in Table 4 and Table 5. Some examples of the reconstructed images are shown in Fig. 7.

It can be observed that the performance of LRCS will be enhanced as the sampling ratio increases and the number of training samples increases. The performance of deep learning will be significantly enhanced as the number of training samples increases but it will not be
necessarily improved as the sampling ratio increases. At a very low sampling ratio $S=0.05$, the reconstructed images by LRCS are very heavily degraded but most reconstructed images by deep learning still have acceptable visual quality when the number of training samples are adequate. Since deep learning can extract some high-level common features from the training images, the test object image can still be well recovered from these features when the sampling ratio is very low. On the other hand, the feature extraction and reconstruction may cause more unpredicted errors in the recovered images when the dimension of input data is higher. So the quality of recovered images will not always be worse at a lower sampling ratio and better at a higher sampling ratio for the deep learning approach.

Table 4 Blindly reconstructed image quality for SPI with LRCS for the MNIST dataset

| Number of training samples $K$ | $S=0.05$   | $S=0.2$   | $S=0.4$   | $S=1$   |
|-------------------------------|------------|------------|------------|--------|
| 50                            | 14.2419 dB | 16.8031 dB | 17.5811 dB | 18.3476 dB |
| 100                           | 15.1257 dB | 17.7134 dB | 20.0936 dB | 21.3368 dB |
| 200                           | 15.5376 dB | 20.1990 dB | 22.1549 dB | 24.2062 dB |
| 500                           | 15.3003 dB | 21.0111 dB | 23.6042 dB | 26.4277 dB |

Table 5 Blindly reconstructed image quality for SPI with Wang’s Net for the MNIST dataset

| Number of training samples $K$ | $S=0.05$   | $S=0.2$   | $S=0.4$   | $S=1$   |
|-------------------------------|------------|------------|------------|--------|
| 50                            | 11.9738 dB | 11.6227 dB | 12.2526 dB | 11.5919 dB |
| 100                           | 12.8692 dB | 12.6286 dB | 12.8774 dB | 12.6637 dB |
| 200                           | 13.0954 dB | 13.2059 dB | 13.2644 dB | 13.1486 dB |
| 500                           | 13.2826 dB | 13.3295 dB | 13.3089 dB | 13.2944 dB |
| 2000                          | 15.5567 dB | 14.7390 dB | 17.1544 dB | 15.2462 dB |
| 5000                          | 17.3145 dB | 19.3743 dB | 20.1935 dB | 19.7935 dB |

Fig. 7 Comparison of reconstructed image results for a SPI system with LRCS and deep learning in the simulation.
Fig. 7 show that some reconstructed images by LRCS have quality degradation but the shapes of the digits match with the original groundtruth MNIST images. On the other hand, some reconstructed images by Wang’s Net can be noise-free but the digits have distorted shapes, which will cause a lower PSNR. From the results, it can be observed that the recovered image quality of LRCS with 200 training samples is comparable with the ones using deep learning with 5000 samples, except when the sampling ratio is very low. It takes only 20 minutes to train the linear regression part in the LRCS scheme with 200 training samples in a Matlab R2018a environment with Intel(R) Core(TM) i5-7200U CPU (2.50 GHz) and 8GB RAM. In contrast, it takes at least around 6 hours to train Wang’s Net with 5000 training samples in a CPU environment. Due to the model simplicity, LRCS is much more computationally efficient than deep learning in the training step. The similarities and differences between LRCS and deep learning are summarized in Table 6.

Table 6 Summary of similarities and differences between LRCS and Wang’s Net for blind reconstruction in a SPI system

| Similarity | Linear Regression + Compressive Sensing (LRCS) | Deep learning (Wang’s Net) |
|------------|-----------------------------------------------|-----------------------------|
| Blindly reconstruct the object image from the single-pixel intensity sequence at different sampling ratios for a SPI system after the model is trained with a certain number of training samples, when the illumination patterns are not known | (1) Work with a small number of training samples  
(2) Simple explicit model with few parameters  
(3) Two-step, not end-to-end  
(4) Poor performance at very low sampling ratio | (1) Only work with a large number of training samples  
(2) Complicated model with many parameters for tuning  
(3) End-to-end reconstruction  
(4) Possibly reconstruct high-quality results at very low sampling ratio |

In this work, LRCS and Wang’s Net are evaluated based on the experimentally recorded data as well. The SPI experiments are conducted using the optical setup shown in Fig. 8. Each object image is printed on a paper card and illuminated by the patterns projected by a JmGO G3 projector. The single-pixel intensity values are recorded by a Thorlabs FDS1010 photodiode detector and a NI USB-6216 data acquisition card. Totally ten different object images are tested in the experiment.

![Optical setup of our SPI experiment](image)

The reconstruction results with LRCS and Wang’s Net from the experimentally recorded data are shown in Fig. 9. It is reported in the previous work [34] that deep learning is significantly more robust to the noise in the experimental data. Since the optical setup in this work is different from the one in the previous work [34], the type of noise and its strength can be different in the experiment. For example, no laser illumination is employed in this work and the speckle noise contamination will not occur. In our observation, the performances of both LRCS and Wang’s Net are slightly degraded due to the extra...
experimental noise that do not appear in the simulated training data. But it is still evident that Wang’s Net performs better than LRCS at a low sampling ratio and LRCS perform better than Wang’s Net when the number of training samples is small.

Fig. 9 Comparison of reconstructed image results with LRCS and Wang’s Net for a SPI system based on the recorded data in real optical experiments

6. Conclusion

In this work, we point out that linear-regression-based methods can be used to solve two black-box optical imaging problems that were previously addressed by deep learning approaches. For attacking a TRPE optical cryptosystem, a complex-amplitude linear regression (CLR) scheme is proposed. For the blind image reconstruction in a SPI system, a “linear regression + compressive sensing (LRCS)” scheme is proposed. In these two problems, linear-regression-based methods show some advantages than deep learning such as small number of training samples and short training time. Simulation and experimental results indicate that deep learning does not always outperform linear regression in this type of black-box optical imaging problems and each approach has its own advantages and disadvantages. The similarities and differences between linear-regression-based methods and deep learning are analyzed and summarized.

Funding
National Natural Science Foundation of China (61805145, 11774240); Leading Talents Program of Guangdong Province (00201505); Natural Science Foundation of Guangdong Province (2016A030312010).

Acknowledgement
We would like to thank Dr. Fei Wang, Mr. Han Hai and Prof. Wenqi He's assistance in this work.

Disclosures
The authors declare no conflicts of interest.
Reference

1. K. Yao, R. Unni, and Y. Zheng, “Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale,” Nanophotonics 8(3), 339-366 (2019).
2. S. D. Campbell, D. Seli, R. P. Jenkins, E. B. Whiting, J. A. Fan, and D. H. Werner, “Review of numerical optimization techniques for meta-device design,” Opt. Mater. Express 9(4), 1842-1863 (2019).
3. G. Barbathitis, A. Ozcan, and G. Citti, “On the use of deep learning for computational imaging,” Optica 6(8), 921-943 (2019).
4. A. Sinha, J. Lee, S. Li, and G. Barbathitis, “Lensless computational imaging through deep learning,” Optica 4(9), 1117-1125 (2017).
5. A. Lucas, M. Biadis, R. Molina, and A. K. Katsaggelos, “Using deep neural networks for inverse problems in imaging: beyond analytical methods,” IEEE Sig. Process. Mag. 35(1), 20-36 (2018).
6. C. Ilg, F. S. Oktem, and A. Koč, “Deep Learning-Based Hybrid Approach for Phase Retrieval,” In Computational Optical Sensing and Imaging, Optical Society of America, pp. CTh2C-5 (2019, June).
7. G. Zhang, T. Guan, Z. Wang, T. Hu, D. Wang, Y. H. and N. Xie, “Fast phase retrieval in off-axis digital holographic microscopy through deep learning,” Opt. Express 26(15), 19388-19405 (2018).
8. Y. Rivenson, Z. Göricks, H. Günaydın, Y. Zhang, H. Wang, and A. Ozcan, “Deep learning microscopy,” Optica 4(11), 1437-1443 (2017).
9. B. Manifold, E. Thomas, A. T. Francis, A. H. Hill, and D. Fu, “Denoising of stimulated Raman scattering microscopy images via deep learning,” Biomed. Opt. Express 10(8), 3860-3874 (2019).
10. M. Lyu, H. Wang, G. Li, S. Zheng, and G. S. Liu, “Learning-based lensless imaging through optically thick scattering media,” Adv. Photon. 1(3), 036002 (2019).
11. Y. Li, Y. Xue, and L. Tian, “Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media,” Optica 5, 1181–1190 (2018).
12. Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng, and A. Ozcan, “Phase recovery and holographic image reconstruction using deep learning in neural networks,” Light Sci. Appl. 7(2), 1714 (2018).
13. H. Wang, M. Lyu, and G. S. Liu, “eHoloNet: a learning-based end-to-end approach for in-line digital holographic reconstruction,” Opt. Express 26(18), 22603-22614 (2018).
14. Z. Ren, Z. Xu, and E. Y. Lam, “End-to-end deep learning framework for digital holographic reconstruction,” Adv. Photon. 1(1), 016004 (2019).
15. Z. Ren, Z. Xu, and E. Y. Lam, “Learning-based nonparametric autofocusing for digital holography,” Optica 5(4), 337-344 (2018).
16. T. Ptkiuho, A. Manninen, and T. J. Naughton, “Focus prediction in digital holographic microscopy using deep convolutional neural networks,” Appl. Opt. 58(5), A202-A208 (2019).
17. S. Jiao, Z. Jin, C. Chang, C. Zhou, W. Zou, and X. Li, “Compression of phase-only holograms with JPEG standard and deep learning,” Appl. Sci. 8(8), 1258 (2018).
18. T. Shimobaba, D. Blinder, M. Makowski, P. Schelkens, Y. Yamamoto, I. Hoshi, T. Nishitsuji, Y. Endo, T. Kakue, and T. Ito, “Dynamic-range compression scheme for digital hologram using a deep neural network,” Opt. Lett. 44(12), 3038-3041 (2019).
19. T. Shimobaba, Y. Endo, T. Nishitsuji, T. Takahashi, Y. Nagahama, S. Hasegawa, M. Sano, R. Hirayama, T. Kakue, A. Shiraki, and T. Ito, “Computational ghost imaging using deep learning,” Opt. Commun. 413, 147-151 (2018).
20. C. F. Higham, R. Murray-Smith, M. J. Padgett, and M. P. Edgar, “Deep learning for real-time single-pixel video,” Sci. Rep. 8(1), 2369 (2018).
21. Z. Ren, H. K. H. So, and E. Y. Lam, “Fringe Pattern Improvement and Super-Resolution Using Deep Learning in Digital Holography,” IEEE Trans. Ind. Inform., 1-1 (2019).
22. Z. Niu, J. Shi, L. Sun, Y. Zhu, J. Fan, and G. Zeng, “Photon-limited face image super-resolution based on deep learning,” Opt. Express 26(18), 22773-22782 (2018).
23. Z. Luo, A. Yurt, R. Stahl, A. Lambrecht, V. Reumers, D. Braeken, and L. Lagae, “Pixel super-resolution for lens-free holographic microscopy using deep learning neural networks,” Opt. Express 27(10), 13581-13595 (2019).
24. T. Nguyen, Y. Xue, Y. Li, L. Tian, and G. Neumannullah, “Deep learning approach for Fourier ptychography microscopy,” Opt. Express 26(20), 26470–26484 (2018).
25. S. Jiang, K. Gao, J. Liao, and G. Zheng, “Solving Fourier ptychography imaging problems via neural network modeling and TensorFlow Biomed,” Opt. Express 9(7), 3306–3319 (2018).
26. Y. F. Cheng, M. Strachan, Z. Weiss, M. Deb, D. Carone, and V. Ganapati, “Illumination pattern design with deep learning for single-shot Fourier ptychography microscopy,” Opt. Express 27(2), 644-656 (2019).
27. S. Feng, Q. Chen, G. Gu, T. Tao, L. Zhang, Y. Hu, W. Yin and C. Zuo, “Fringe pattern analysis using deep learning,” Adv. Photon. 1(2), 025001 (2019).
28. K. Wang, Y. Li, Q. Kernao, J. Di, and J. Zhao, “One-step robust deep learning phase unwrapping,” Opt. Express 27(10), 15100-15115 (2019).
29. Q. Xin, G. Ju, C. Zhang, and S. Xu, “Object-independent image-based wavefront sensing approach using deep learning images and deep learning,” Opt. Express 27(18), 26102-26119 (2019).
30. Y. Nishizaki, M. Valdivia, R. Horisaki, K. Kitaguchi, M. Saito, J. Tanida, and E. Vera, “Deep learning wavefront sensing,” Opt. Express 27(1), 240-251 (2019).
31. B. Karanov, M. Chagnon, F. Thouin, T. A. Eriksson, H. Bulow, D. Lavery, P. Bayvel, and L. Schmaler, “End-to-end deep learning of optical fiber communications,” J. Lightwave Technol. 36(20), 4843-4855 (2018).
32. M. Hutson, “AI researchers allege that machine learning is alchemy,” Science 360(6388), 861 (2018).
33. H. Hai, S. Pan, M. Liao, D. Lu, W. He, and X. Peng, “Cryptanalysis of random-phase-encoding-based optical cryptosystem via deep learning,” Opt. Express 27(15), 21204-21213 (2019).
34. 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).
35. M. P. Edgar, G. M. Gibson, and M. J. Padgett, “Principles and prospects for single-pixel imaging,” Nat. Photonics 13(1), 13–20(2019).
36. S. Jiao, J. Feng, Y. Gao, T. Lei, Z. Xie, and X. Yuan, “Optical machine learning with incoherent light and a single-pixel detector,” Opt. Lett. 44(21), 5186-5189 (2019).
37. S. Liu, C. Guo, and J. T. Sheridan, “A review of optical image encryption techniques,” Opt. Laser Technol. 57, 327-342 (2014).
38. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proc. IEEE 86(11), 2278-2324 (1998).
39. H. Xiao, K. Rasul, and R. Vollgraf, “Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms,” arXiv preprint arXiv:1708.07747(2017).