Ghost Imaging Based on Recurrent Neural Network

Yuchen He, Sihong Duan, Jianxing Li, Hui Chen, Huabin Zheng, Jianbin Liu, Yu Zhou, and Zhuo Xu

Abstract—Benefit from the promising features of second-order correlation, ghost imaging (GI) has received extensive attentions in recent years. Simultaneously, GI is affected by the poor trade-off between sampling rate and imaging quality. The traditional image reconstruction method in GI is to accumulate the action result of each speckle and the corresponding bucket signal. We found that the image reconstruction process of GI is very similar to the Recurrent Neural Network (RNN), which is one of the deep learning algorithm. In this paper, we proposed a novel method that effectively implements GI on the RNN architecture, called GI-RNN. The state of each layer in RNN is determined by the output of the previous layer and the input of this layer, and the output of the network is the sum of all previous states. Therefore, we take the speckle of each illumination and the corresponding bucket signal as the input of each layer, and the output of the network is the sum of all previous speckle and bucket signal, which is the image of the target. The testing results show that the proposed method can achieve image reconstruction at a very low sampling rate (0.38%). Moreover, we compare GI-RNN with traditional GI algorithm and compressed sensing algorithm. The results of different targets show that GI-RNN is 6.61 dB higher than compressed sensing algorithm and 12.58 dB higher than traditional GI algorithm on average. In our view, the proposed method makes an important step to applications of GI.

Index Terms—Ghost imaging, recurrent neural network, basic correlation.

I. INTRODUCTION

GHOST imaging (GI), which was first demonstrated by Pittman and Shih in 1995, is regarded as a novel imaging technology different from conventional methods based on first-order interference [1]. By employing second-order correlation, GI provides some promising features such as lensless imaging, turbulence-free and high detection sensitivity. Consequently, GI has attracted much attention of researchers and produced many interesting results [2], [3], [4], [5], [6], [7], [8], [9], [10], [11]. However, GI is always trapped in the demand for a large number of speckles for high-quality imaging results, which has hindered the development of its applications. Compressed sensing (CS) method has been considered to solve this bottleneck problem [12], [13], [14], [15], [16], but large computational cost limits the application of CS in GI. Recently, artificial intelligence (AI) methods are more and more utilized to improve GI [17], [18], [19], [20], [21], [22], [23], [24].

In [17], [18], [19], the image obtained by traditional GI method is used as the input of the deep neural network, and the predicted by the network is the high resolution target image. [21] directly takes the bucket signal as the input of the network, and uses the multi-branched network with residual structure to reconstruct the target image. In training process, the ground truth needs to be resized and binarized into a resolution of 32*32 binary image. This method can reconstruct a clear image when the sampling rate is 6.25%. In [23], DAttNet structure is proposed to reconstruct the target image, and the structure is a network model similar to U-net. In training process, the resolution of ground truth is 128*128 and the size of train set contains 200011 samples. This method can reconstruct a clear image when the sampling rate is 5.45%. However, because the resolution of the ground truth of the network is 128*128, 893 detections are required to image a target even when the sampling rate is 5.45%.

In this paper, we proposed an improved GI method based on recurrent neural network (RNN) called GI-RNN, which can image the target at a very low sampling rate. Conventional neural networks can only individually process one input, and the previous input has nothing to do with the latter one. On the other hand, RNN can deal with the issue when the previous input is associated with the latter. In GI, the basic correlation method is the most widely used, which continuously accumulate the action results of speckle and bucket signal. In this sense, the result of the basic correlation method is the correlation with all previous states, that is, the input of each state is associated with the input of the previous one. In training, GI-RNN takes the speckle of each illumination and the corresponding bucket signal as the input of each state in RNN. Each state consists of the output of the previous state (speckle and bucket signal at time t-1) and the input of the current state (speckle and bucket signal at time t). In this way, the final output of RNN is the sum of all speckles and bucket signals, that is, the image of the target. In testing, the fixed sequence of speckle in training is employed to illuminate unknown target. To demonstrate the proposed method, we test it on MNIST at different sampling rates, and compare it with basic correlation algorithm and compressed sensing algorithm. The testing results show that the proposed method can achieve image reconstruction when the sampling rate is only 0.38%. Moreover, the results of different targets show that GI-RNN is 6.61 dB higher than compressed sensing algorithm and 12.58 dB higher than traditional GI algorithm on average.

The rest of this paper is organized as follows. Section II provides a brief survey of related work. In Section III, a comprehensive introduction to the proposed method is provided. In Section IV, the proposed method is demonstrated and compared by extensive experiments. The paper is concluded in Section V.

II. RELATED WORK

A. Ghost Imaging

At the beginning, GI collects data through coincidence detection of two arms. One arm uses a bucket detector without spatial resolution to receive the echo signal containing target information, and the other arm (without target) uses an array detector to collect the spatial information of the corresponding
position. Different from the traditional imaging technology which depends on intensity information, GI relies on intensity fluctuation information to achieve image reconstruction. Until 2018, Shapiro proposed computational ghost imaging (CGI) to obtain spatial information of corresponding position through calculation. The emergence of CGI simplifies the two arms of GI into one arm, which improves its practicability. CGI employs a sequence of random speckles to illuminate target, and detect the echo signal by a bucket detector without spatial resolution. Then, the image reconstruction process can be expressed as follow called basic correlation method

\[
T = \sum_{i=1}^{N} P_i \cdot B_i, \quad (1)
\]

where \( P_i \) represents the \( i \)-th speckle and \( B_i \) represents the \( i \)-th bucket signal. Through \( N \) times of illumination, target information \( T \) is obtained.

B. Recurrent Neural Network

RNN is a kind of neural network with sequence data as input, recursion in the evolution direction of the sequence, and all nodes are connected in a chain. RNN is very effective for the data with sequence characteristics. It can excavate the temporal information and semantic information in the data. By using this ability of RNN, the deep learning model has made a breakthrough in solving the problems in natural language processing (NLP) fields such as speech recognition, language model, machine translation and temporal analysis. The state of the system at time \( t \) can be expressed as

\[
h_t = f (h_{t-1} + x_t), \quad (2)
\]

where \( h_t \) represents the state of \( t \) and \( x_t \) represents the input of \( t \). \( h_{t-1} \) represents the state of the previous time. Eq. 2 shows that the state of \( t \) is determined by the output of the previous state \( t-1 \) and the input of \( t \).

III. THE PROPOSED METHOD

A. Principle

Inspired by the principle of basic correlation method and the process of RNN, we find that they are similar in form. For GI, the basic correlation method combines the speckle of each illumination by the corresponding bucket signal, and accumulates the results. For RNN, the output of the network is the sum of the previous states, and the state of each layer is determined by the output of the previous layer and the input of this layer. Fig. 1 shows the corresponding relationship.

In Fig. 1, \( h_t \) denotes the state of time \( t \) and \( T_t \) denotes the target image of time \( t \). \( x_t \) denotes the input of time \( t \) and \( P_t \) denotes the speckle pattern of time \( t \). \( O_t \) denotes the output of time \( t \) and \( (P_t, B_t) \) denotes the results of the interaction between speckle and bucket signal. As we can see, from expression to meaning, RNN is similar to GI. Consequently, the action results of each speckle and bucket signal in GI can be regarded as the input of each layer in RNN. Meanwhile, the imaging results of GI can be regarded as the output of the last layer of RNN. The schematic diagram of the proposed method is shown in Fig. 2.

B. Architecture

The proposed method consists of three parts: pre-processing module, RNN backbone and predictor. The pre-processing module is used to convert speckle and corresponding bucket signal into a form suitable for RNN input. In the proposed method, long short-term memory (LSTM) is used as the model of RNN. One of the key reasons for using LSTM is that the network can effectively learn long-term dependent information, which is very consistent with the imaging process of GI. GI-RNN is similar to the machine translation in NLP. A speckle-bucket pair is equivalent to a "word", and the feature vector of the target is output after all the "words" are input.

When imaging complex targets, the data sequence will be relatively long, even the sampling rate is low. Consequently, the network is required to have the ability to learn long-term dependent information, so LSTM is adopted. In addition, LSTM can also effectively solve the problems of gradient vanishing and exploding in the process of long sequence training. The predictor takes the feature vector output by LSTM as the input, and reconstructed image of target as the output.

IV. EXPERIMENTAL RESULTS

A. Training Settings

In this paper, we adopt the structure of multi-layer LSTM, and the number of circulating layers is 5. The network input size is 785 and the hidden state size is 1024. Moreover, the predictor input size is 1024 (equal to the LSTM hidden state) and output size is 784. During training, we resize the target image as a one-dimensional vector with a length of 784 as the ground truth. The mean-square error (MSE) is used as the loss function of network training between the reconstructed image and the ground truth, and we use Adam as the optimizer.

After a large number of experiments, we found that the optimal initialization learning rate is 0.0001, and the weight decay is 0. The train set in this work is MNIST with the image resolution of 28*28. We randomly selected 9000 images from MNIST for training, and the testing samples are randomly selected from the testing set of MNIST.
B. Results on MNIST

The experiments were compared at the sampling rates of 0.38%, 1.02%, 1.56%, 2%, 6.25%, 25% and 100%. Then, we selected 0.38%, 1.02%, 6.25% and 25% for comparison. The image reconstruction results are shown in Fig. 3.

As shown in Fig. 3, the proposed method can obtain stable results (the nine targets in the testing set) at the sampling rate of 6.25%, and for some targets, the image can be obtained at 1.02% or even 0.38% (which means 3 illuminations). Fig. 4 shows the result comparison curve of PNSR, and the detailed numerical results are shown in Table I.

C. Comparison With Other Methods

To demonstrate the performance of the proposed method, we compare GI-RNN with traditional GI and compressed sensing algorithm at sampling rate of 25% (which means 196 illuminations). The image reconstruction results are shown in Fig. 5.

Fig. 5 shows that the traditional GI algorithm (basic correlation) can not reconstruct target at such low sampling rate. Compressed sensing algorithm (FISTA) can only partially
reconstruct targets. However, the proposed method can achieve stable results. The PSNR curve of reconstruction results are shown in Fig. 6.

Table I: PSNR comparison results under different sampling rates

| SR (SR%) | Target 0 | Target 1 | Target 2 | Target 3 | Target 4 | Target 5 | Target 6 | Target 7 | Target 8 | Target 9 |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 0.38%   | 10.964  | 14.932  | 11.293  | 14.077  | 12.156  | 12.320  | 11.882  | 13.416  | 11.855  | 12.254  |
| 1.02%   | 15.077  | 22.055  | 11.612  | 15.583  | 13.908  | 13.390  | 10.761  | 13.277  | 10.465  | 16.440  |
| 6.25%   | 19.428  | 27.067  | 14.048  | 19.930  | 18.681  | 19.216  | 15.114  | 20.579  | 15.034  | 19.320  |
| 25%     | 19.933  | 28.316  | 15.981  | 20.124  | 19.932  | 19.641  | 16.440  | 22.087  | 15.519  | 20.247  |

Fig. 6. PSNR comparison results of different methods.

Fig. 6 shows that the reconstruction results of GI-RNN are significantly better than the other two methods. Table II lists the detailed numerical results.

After statistics (the nine targets in the testing set). Table II shows that the proposed method is 6.61 dB higher than compressed sensing algorithm and 12.58 dB higher than traditional GI algorithm on average.

V. CONCLUSION

In this paper, we propose a novel GI method based on RNN called GI-RNN. The proposed method effectively integrates the traditional GI algorithm (basic correlation) into the architecture of RNN network, and can realize image reconstruction at a very low sampling rate. The process of basic correlation method is to accumulate the results of each speckle interacting with the corresponding bucket signal. For RNN, the state of each layer is determined by the input of this layer and the output of the previous layer, and the final output is determined by the accumulation of all previous layers. Therefore, we take each illuminated speckle and corresponding bucket signal in GI as the input of each layer in RNN, and the final output of RNN is target image. We demonstrate the proposed method at different sampling rates, and it can achieve image reconstruction when the sampling rate is only 0.38%. Moreover, we compare the proposed method with basic correlation method and compressed sensing method at sampling rate of 25%. Extensive experiments show that the proposed method is 6.61 dB higher than compressed sensing algorithm and 12.58 dB higher than traditional GI algorithm on average.

REFERENCES

[1] T. B. Pittman, Y. H. Shih, D. V. Strekalov, and A. V. Sergienko, “Optical imaging by means of two-photon quantum entanglement,” Physical Review A, vol. 52, no. 5, 1995.
[2] A. Valencia, G. Scarcelli, M. D’Angelo, and Y. Shih, “Two-photon imaging with thermal light,” Physical Review Letters, vol. 94, no. 6, pp. 63.601–63.601, 2005.
[3] J. H. Shapiro, “Computational ghost imaging,” Physical Review A, vol. 78, no. 6, p. 61802, 2008.
[4] R. Meyers, K. S. Deacon, and Y. Shih, “Ghost-imaging experiment by measuring reflected photons,” Physical Review A, vol. 77, no. 4, p. 41801, 2008.
[5] Y. Bromberg, O. Katz, and Y. Silberberg, “Ghost imaging with a single detector,” Physical Review A, vol. 79, no. 5, p. 53840, 2009.
[6] F. Ferri, D. Magatti, L. A. Lugliato, and A. Gatti, “Differential ghost imaging,” Physical Review Letters, vol. 104, no. 25, pp. 253603–253603, 2010.
[7] R. E. Meyers, K. S. Deacon, and Y. Shih, “Turbulence-free ghost imaging,” Applied Physics Letters, vol. 98, no. 11, p. 111115, 2011.
[8] P. Ryczkowski, M. Barbier, A. T. Friberg, J. M. Dudley, and G. Genty, “Ghost imaging in the time domain,” Nature Photonics, vol. 10, no. 3, pp. 167–170, 2016.
[9] D. Pelliccia, A. Rack, M. Scheel, V. Cantelli, and D. M. Paganin, “Experimental x-ray ghost imaging,” Physical Review Letters, vol. 117, no. 11, pp. 113.902–113.902, 2016.
[10] R. Khakimov, B. Henson, D. K. Shin, S. Hodgman, R. Dall, K. Baldwin, and A. Truscott, “Ghost imaging with atoms,” Nature, vol. 540, no. 7631, pp. 100–103, 2016.
[11] S. Ota, R. Horisaki, Y. Kawamura, M. Ugawa, I. Sato, K. Hashimoto, R. Kamesawa, S. Setoyama, S. Yamaguchi, K. Fujih, K. Waki, and H. Noji, “Ghost cytometry,” Science, vol. 360, no. 6394, pp. 1246–1251, 2018. [Online]. Available: https://science.sciencemag.org/content/360/6394/1246
[12] O. Katz, Y. Bromberg, and Y. Silberberg, “Compressive ghost imaging,” Applied Physics Letters, vol. 95, no. 13, p. 131110, 2009.
[13] Y. Katkovnik and J. Astola, “Compressive sensing computational ghost imaging.,” Journal of The Optical Society of America A-optics Image Science and Vision, vol. 29, no. 8, pp. 1556–1567, 2012.
[14] M. Amann and M. Bayer, “Compressive adaptive computational ghost imaging,” Scientific Reports, vol. 3, no. 1, pp. 1545–1545, 2013.
[15] L. Long-Zhen, Y. Xu-Ri, L. Xue-Feng, Y. Wen-Kai, and Z. Guang-Jie, “Super-resolution ghost imaging via compressed sensing,” Acta Physica Sinica, vol. 63, no. 22, pp. 224.201–224.201, 2014.
[16] H. Zhang, Y. Xia, and D. Duan, “Computational ghost imaging with deep compressed sensing,” Chinese Physics B, 2021.
[17] M. Lu, W. Wang, H. Wang, H. Wang, G. Li, N. Chen, and G. Situ, “Deep-learning-based ghost imaging,” Scientific Reports, vol. 7, p. 17865, 2017.
| Target | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|-------|---|---|---|---|---|---|---|---|---|---|
| GI    | 6.6399 | 8.6899 | 6.4968 | 6.8234 | 6.9307 | 7.5898 | 6.9267 | 6.9045 | 6.9307 | 6.8669 |
| CS    | 10.4739 | 25.2921 | 10.7656 | 11.3144 | 13.0200 | 13.8379 | 11.3820 | 13.8439 | 9.6817 | 11.7075 |
| GI-RNN | 19.9336 | 28.3169 | 15.9812 | 20.1249 | 19.9325 | 19.6416 | 16.4140 | 22.0872 | 15.5198 | 20.2247 |

[18] 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,” Optics Communications, vol. 413, pp. 147–151, 2018.

[19] Y. He, G. Wang, G. Dong, S. Zhu, H. Chen, A. Zhang, and Z. Xu, “Ghost imaging based on deep learning.” Scientific Reports, vol. 8, no. 1, p. 6469, 2018.

[20] G. Wang, H. Zheng, W. Wang, Y. He, J. Liu, H. Chen, Y. Zhou, and Z. Xu, “De-noising ghost imaging via principal components analysis and compandor,” Optics and Lasers in Engineering, vol. 110, pp. 236–243, 2018.

[21] 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.” Optics Express, vol. 27, no. 18, pp. 25 560–25 572, 2019.

[22] T. Bian, Y. Yi, J. Hu, Y. Zhang, Y. Wang, and L. Gao, “A residual-based deep learning approach for ghost imaging.” Scientific Reports, vol. 10, no. 1, pp. 12 149–12 149, 2020.

[23] H. Wu, R. Wang, G. Zhao, H. Xiao, D. Wang, J. Liang, X. Tian, L. Cheng, and X. Zhang, “Sub-nyquist computational ghost imaging with deep learning.” Optics Express, vol. 28, no. 3, pp. 3846–3853, 2020.

[24] Z. Zhang, C. Wang, W. Gong, and D. Zhang, “Ghost imaging of blurred object based on deep-learning.” Applied Optics, vol. 60, no. 13, pp. 3732–3739, 2021.