Ghost Handwritten Digit Recognition based on Deep Learning
Xing He,1 Shengmei Zhao,1,2,1 and Le Wang1
1) Institute of Signal Processing and Transmission, Nanjing University of Posts and Telecommunications (NUPT), Nanjing, 210003, China
2) Key Lab of Broadband Wireless Communication and Sensor Network Technology (NUPT), Ministry of Education, Nanjing, 210003, China

(Dated: 7 April 2020)

We present a ghost handwritten digit recognition method for the unknown handwritten digits based on ghost imaging (GI) with deep neural network, where a few detection signals from the bucket detector, generated by the Cosine Transform speckle, are used as the characteristic information and the input of the designed deep neural network (DNN), and the classification is designed as the output of the DNN. The results show that the proposed scheme has a higher recognition accuracy (as high as 98.14% for the simulations, and 92.9% for the experiments) with a smaller sampling ratio (say 12.76%). With the increase of the sampling ratio, the recognition accuracy is enhanced greatly. Compared with the traditional recognition scheme using the same DNN structure, the proposed scheme has a little better performance with a lower complexity and non-locality property. The proposed scheme provides a promising way for remote sensing.

PACS numbers: 42.30.Sy, 42.50.Ar, 42.30.-d

I. INTRODUCTION

In recent years, Handwritten digit recognition is becoming an active research topic because it has many practical applications. But, Handwritten digit recognition is still a challenging problem due to the different handwriting qualities and the different styles. Several methods have been proposed for solving this problem, such as deep learning-based classification algorithms, artificial neural networks, and support vector machine classifier.

Among them, the image of the target handwritten digit should be firstly obtained, then the classification can be achieved with the characteristic information. Although the satisfactory recognition results could be achieved with the existing recognition methods, however, several digits misrecognitions are still inevitable due to the large variation of the individual writing style. Furthermore, at some cases, the characteristic information of the handwritten digit are impossible to be achieved before recognition.

In another context, Ghost imaging (GI), also called correlated imaging, is an intriguing optical imaging technique, where the images of objects can be achieved by correlating the fluctuations between the separated optical fields, without the need to record the image itself. GI offers great promise for its robustness against harsh environment, higher spatial resolution and higher detection sensitivity, and now has great applications in laser radars, microscopes, object edges extraction, image hiding methods, and optical encryption schemes.

Initially, GI was experimental demonstrated by using entangled-photon pairs, generated by spontaneous parametric down-conversion (SPDC), in 1995. Then, GI was also found to be successfully realized by using a classical pseudo-thermal light source, thermal source, and daylight. With the development of GI, Computational Ghost imaging (CGI) was introduced to compute the intensity offline, which greatly simplified GI’s configuration and generalized GI’s applications. It was also demonstrated that the single-pixel imaging (SPI) was the same technique as GI.

Recently, some works have studied the image classification by using single-pixel detector, such as, the authors proposed a reconstruction-free multi-class image classification framework to reveal the classification only based on the sequence of photodetector measurements was possible. Moreover, an optical diffractive neural network to perform machine learning tasks, such as number digit recognition, in an all-optical manner was presented. The single-pixel non-imaging object was experimentally demonstrated to achieve recognition by acquiring the Fourier spectrum. Additionally, a single neural network with multi-rate property for compressed domain classification, and a neural network for simultaneously learning the linear binary sensing matrix and the non-linear classification parameters were discussed for Single-pixel camera (SPC).

The most natural images, including the handwritten digit images, are sparse in frequency spectrum domain and mainly focus on the low-frequency components, that means their most coefficients in discrete cosine transform (DCT) basis are close to or equal to zero. At the same time, artificial neural networks are good at object recognition. Hence, we propose a novel handwritten digit recognition scheme using GI technique and DNN, named ghost handwritten digit recognition (GHDR), where a fewer detection signals, obtained from the DCT speckle pattern illuminations, are used as the feature information of the unknown handwritten digit image, and the input of the designed deep neural network (DNN). The output of the DNN is the recognition result.
And the fully connected DNN has one input layer, three hidden layers and one output layer.

The advantages of the proposed scheme are the following. With the non-locality of GI technique, the proposed scheme can obtain the recognition without obtaining the clear handwritten digit image at first. Secondly, the handwritten digit image (two-dimension) is projected and characterized by a smaller one-dimension set by using GI technique, that will simple the the designed DNN configuration, furthermore, that will greatly reduce the complexity of the recognition.

Different from the method in \cite{28}, which used the spatial light modulation to acquire the feature information and the convolutional neural network to learn the spatial features, our scheme adopts a fewer detection signals, generated by Cosine Transform speckle in GI system as the feature information, and only fully-connection DNN network is needed to obtain the classification. Both the dimension of the feature information and the corresponding recognition complexity are reduced.

II. GHOST HANDWRITTEN DIGIT RECOGNITION USING DEEP NEURAL NETWORK

The schematic diagram of the proposed scheme is illustrated in Fig. 1 which has the speckle patterns’ generation part, detector signal achievement part, and the handwritten digit recognition part. At speckle patterns’ generation part, some computed rows are selected from the normal DCT matrix and reshaped to two dimensional for constructing the patterns for the special speckles. At detector signal achievement part, these patterns are firstly modulated on the beam to generate the special speckles \( I_{(x, y)} \)\( \cdots \)\( I_M(x, y) \) by a Digital Light Projector (DLP). Then, these speckles are illustrated on the unknown handwritten digit image successively, after the expansion with a projector lens. The resultant beam is subsequently detected by a bucket detector to achieve the detection signals \( B_i, i = 1, \ldots , M \). At handwritten digit recognition part, the detection signals are used as the input to a designed DNN network, and the output is the classification result of the unknown handwritten digit image.

Since the classified object is a handwritten digit image, it should be inherently sparse in frequency spectrum domain and mainly focus on the low-frequency components. Hence, it is reasonable to use their low-frequency components for the feature extraction. To obtain the low-frequency components of the unknown handwritten digit image based on GI technique, the DCT speckles are preferable. One method to obtain the cosine speckle is from Discrete Cosine Transform (DCT), which is a discrete Fourier transform (DFT) with real transform coefficients. For one-dimensional DCT, the frequency spectrum can be expressed as

\[
F(\omega) = \sum_{x=0}^{N-1} g(x', \omega) f(x'),
\]

where \( f(x') \) represents discrete signals and \( g(x', \omega) \) denotes Cosine transform coefficient, expressed as

\[
g(x, \omega) = C(\omega) \sqrt{\frac{2}{N}} \cos \left( \frac{2x + 1}{2N} \right) \pi,
\]

where \( C(\omega) = \frac{1}{\sqrt{2}} \) when \( \omega = 0 \), otherwise \( C(\omega) = 1 \).

With GI technique, the detection signal \( B_i \) in the bucket detector can be expressed as

\[
B_i = \sum_{x} \sum_{y} I_{(x, y)} T(x, y) = \sum_{x'} I(x') T(x').
\]

where \( I_{(x', y)} \) denotes the illuminating intensity, \( T(x') \) denotes the unknown object. If \( I_{(x', y)} \) is given by Cosine transform coefficient, \( T(x') \) is the handwritten digit image, then \( B_i \) is one frequency spectrum value.

It is known that the low-frequency components are the most important frequency spectrum values for handwritten digit image, and these low-frequency components are mainly located at the left upper corner of the two-dimensional image frequency spectrum. Hence, it is suitable to consider the first \( b \) rows, \( a \) columns elements in the two-dimensional image frequency spectrum (total \( a \times b \) elements) be the most important frequency spectrum values. Consequently, the location of the corresponding speckle \( I_{(x, y)} \) in DCT matrix can be computed as

\[
i = N_y \times (f_a - 1) + f_b,
\]

where \( (f_a, f_b), 1 \leq f_a \leq a, 1 \leq f_b \leq b \), is the coordinate of the element in two-dimensional image frequency spectrum. In order to quantify the number of important frequency spectrum values used in the recognition, we define a sampling ratio \( SR \) as

\[
SR = \frac{M}{N_x \times N_y} = \frac{a \times b}{N_x \times N_y}.
\]
For simplification, $a$ is commonly chosen the same as $b$, $N_x$ is equal to $N_y$.

After the most important frequency spectrum values are obtained using GI technique, these detection signals are then formed a vector, and as the input to the designed and trained deep neural network (DNN) for handwritten digit image recognition. Fig. 2 shows the structure of the designed DNN for handwritten digit image recognition. The DNN is a feed forward artificial neural network with three hidden layers, and the full connections are used between the layers. The first layer is the input layer, which has $M$ neurons designed for $M$ detection signals $B_1, B_2, B_3, \cdots, B_M$. There are 2000 neurons in the first two hidden layers and 1000 neurons in the third hidden layer. There are 10 neurons in the output layer which correspond to the ten digits.

There is a linear relationship between the input and output of each layer, including weight $w$ and bias $b$. In order to enhance the expression ability of DNN, the nonlinear transformation function $\sigma(\cdot)$ is added to the linear transformation result $O_j$, which is usually called the activation function. Assume there exists $m$ neurons in the $l-1$ layer. The connection between the input and output of the $j$th neuron in the $l$ layer, $a^l_j$ can be expressed as

$$a^l_j = \sigma(O^l_j) = \sigma(\sum_{k=1}^{m} w^l_{jk} a^{l-1}_k + b^l_j),$$  

where $w^l_{jk}$ is the weight between the $k$th neuron in the $l-1$ layer and $j$th neuron in the $l$ layer, $b^l_j$ is the basis for $j$th neuron in the $l$ layer, and $a^{l-1}_k$ represents the $k$th neuron in the $l-1$ layer. When $l-1$ equals 1, $a^{l-1}_k$ represents the $k$th input value in the input layer. The activation function used for each hidden layer is a tanh function $f(O_j) = \frac{e^{O_j} - e^{-O_j}}{e^{O_j} + e^{-O_j}}$, while it is a Softmax function in the output layer $\text{Softmax}(j) = \frac{e^{O_j}}{\Sigma_{j=0} e^{O_j}}$.

The adopted loss function is the cross entropy function, which can be defined as

$$L = \text{Loss}(w, b, x, y) = - \sum_j y_j \ln (\text{Softmax}(j)),$$

where $x$ is the input vector with dimension $M$, $y$ is the output classification result, from 0 to 9. the adaptive moment estimation (ADAM) optimization is adopted to update each layer’s weights and offsets in back propagation, the dropout function is also used to prevent the overfitting.

### III. EXPERIMENTAL AND SIMULATION RESULTS

In this section, we do experiments and simulations to testify the proposed handwritten digit recognition scheme. The numerical simulations was implemented by Keras framework based on TensorFlow with CPU of Intel Core i7-4790 (Dell Optiplex 920 with intel core 3.6GHz and memory 24GB). The language was Python version 3.5(64bits). All the handwritten digit images were from MNIST handwritten digit database, including 60,000 training images and 10,000 test images, all of which were 28 pixel × 28 pixel grayscale images.

The experiment is developed with the schematic diagram in Fig. 3. With the designed method proposed above, the DCT patterns with the locations for the most important frequency spectrum values, were firstly produced from DCT transform. Then, they were regarded as the speckle pattern, realized by a digital light projector (DLP) (TI Digital light Crafter 4500), to modulate the beam. The beam was expanded by a projecting lens (focal length is 200 mm) and illuminated on the unknown handwritten digit image, and was then focused by collecting lens (focal length is 200 mm). Finally, the detection result $B_i$ was obtained from the bucket detector (Thorlabs power meter S142C). Here, the bucket detector recorded the most important frequency spectrum values for the unknown handwritten digit image. The above process was repeated for all the handwritten digit images in the training set, and all the bucket detector output data were as the training data for the designed DNN. When the designed DNN was trained, the above process was executed for the testing handwritten digit images.

Fig. 4 shows the experimental and simulation results with the proposed handwritten digit image scheme for different sampling ratios (SR); (a) by simulations, where 60000 handwritten digit images are adopted for training the designed DNN network, and 10000 handwritten digit images are used for testing; (b) by experiments,
where 9000 handwritten digit images are used for training, and 1000 handwritten digit images are adopted for testing. The results showed that the proposed scheme had a higher recognition accuracy for the handwritten digit images. When $SR = 12.76\%$, the average recognition accuracy was 98.14\% for the simulation results, and was 92.9\% for the experimental results. Simultaneously, the recognition accuracy had increased with the sampling ratio both for simulation and experimental results. It was also shown that some digits were easy for recognition, such as the digit 0 and 1, and some digits were hard to recognize, for example, the digit 8 and 9.

Then, we compare the accuracy of the proposed recognition scheme with different classification algorithms in Fig. 5, where Deep Neural Network, \textit{k}-nearest neighbor, multiple Bernoulli model naive, Gaussian naive Bayesian, decision tree classifier, gradient boosting classifier, and random forest classifier are discussed. The experimental results showed that the proposed recognition scheme had different accuracies for the different classification algorithms even though the input data, the DNN structure and the sample ratio were the same. From the results, it was seen that back propagation and Gradient Boosting Classifier had a better performance among all the classification algorithms.

Finally, we compare the performance of our scheme with the traditional recognition method using DNN. Both schemes have the same DNN structure (one input layer, three hidden layers, and one output layer) and the same neurons at the hidden layers and the output layer. At the same time, the same optimization algorithm are used for the two DNNs. Additionally, the back propagation algorithms is adopted. The numerical simulation results showed that our proposed scheme had better or almost the same well performance with the traditional DNN recognition scheme when the sampling ratio was 12.76\%. With the increase of the iteration number, both recognition schemes had a better recognition accuracy. However, the average recognition time for our proposed scheme was greatly reduced since only a few bucket detection signals were input to the DNN in our proposed scheme. The cost time was approximately 115 seconds for the traditional DNN recognition scheme, while it was 58 seconds for our proposed scheme.

\section{Conclusion}

In this paper, we have proposed a ghost handwritten digit image recognition scheme based on DNN network using GI. The experimental and simulation results have shown that the proposed scheme has had a higher recognition accuracy for the unknown handwritten digit image with a smaller sampling ratio. With the increase of the sampling ratio, the recognition accuracy has been increased greatly. Compared with the traditional recognition scheme using the same DNN structure, our proposed scheme has had almost better or almost the same well performance and less consumed time. Importantly, our proposed scheme has a lower complexity and is feasible for the unknown handwritten digit image. It is believable that the proposed scheme can be used for other image recognition, when the unknown objects pictures are from other datasets, such as Fashion-MNIST dataset and CIFAR-10 dataset, and the DNN are trained at first.
Y. Lecun, B. E. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel, “Handwritten digit recognition with a back-propagation network,” In Advances in Neural Information Processing Systems, 396–404 (1990).

Y. Wang, X. Wang, W. Liu, “Unsupervised local deep feature for image recognition,” Inf. Sci. 351, 67–75 (2016).

A. Goltsev, V. Gritsenko, “Investigation of efficient features for image recognition by neural networks,” Neural Netw. 28(28), 15–23 (2012).

R. Varatharajan, G. Manogaran, and M. K. Priyan, “A big data classification approach using LDA with an enhanced SVM method for ECG signals in cloud computing,” Multimed Tools Appl. 77(8), 10195–10215 (2018).

S. R. Kulkarni, B. Rajendran, “Spiking neural networks for handwritten digit recognition-Supervised learning and network optimization.” Neural Netw. 103, 118–127 (2018).

D. V. Strekalov, A. V. Sergienko, D. N. Klyshko, and Y. H. Shih, “Observation of two-photon “ghost” interference and diffraction,” Phys. Rev. Lett. 74, 3600–3603 (1995).

R. S. Bennink, S. J. Bentley, and R. W. Boyd, “Two-Photon Coincidence Imaging with a Classical Source,” Phys. Rev. Lett. 89, 113601 (2002).

A. Valencia, G. Scarcelli, M. D’Angelo, and Y. Shih, “Two-Photon Imaging with Thermal Light,” Phys. Rev. Lett. 94, 063601 (2005).

J. H. Shapiro, “Computational ghost imaging,” Phys. Rev. A. 76, 061802(R)—061819 (2008).

G. A. Howland, P. B. Dixon, and J. C. Howell, “Photon-counting compressive sensing laser radar for 3D imaging,” Appl. Opt. 50(31), 5917–5920 (2011).

N. Radwell, K. J. Mitchell, G. M. Gibson, M. P. Edgar, R. Bowman, and M. J. Padgett, “Single-pixel infrared and visible microscope,” Optica 1(5), 285–289 (2014).

N. Tian, Q. Guo, A. Wang, D. Xu, and L. Fu, “Fluorescence ghost imaging with pseudothermal light,” Opt. Lett. 36(16), 3302–3304 (2011).

L. Wang, L. Zou, and S.M. Zhao, “Edge detection based on subpixel-speckle-shifting ghost imaging,” Opt. Commun. 407, 181–185 (2018).

H. D. Ren, S. M. Zhao, J. Gruska, “Edge detection based on single-pixel imaging,” Opt. Express 26(5), 5501–5511 (2018).

H. D. Ren, L. Wang, S. M. Zhao, “Efficient edge detection based on ghost imaging,” OSA Continuum 2(1), 64–73, (2019).

L. Wang, S. M. Zhao, W. Cheng, L. Gong and H. Chen, “Optical image hiding based on computational ghost imaging,” Opt. Commun. 366, 314–320 (2016).

W. Chen, and X. Cheng, “Marked ghost imaging,” Appl. Phys. Lett. 104, 251109 (2014).

S. Jiao, C. Zhou, Y. Shi, W. Zou, and X. Li, “Review on optical image hiding and watermarking techniques,” Opt. Laser Technol. 109, 370–380 (2019).

P. Clemente, V. Durán, V. Torres-Company, E. Tjahjaurence, and J. Lancis, “Optical encryption based on computational ghost imaging,” Opt. Lett. 35, 2391–2393 (2010).

S. M. Zhao, L. Wang, W. Liang, W. Cheng and L. Gong, “High performance optical encryption based on computational ghost imaging with QR code and compressive sensing technique,” Opt. Commun. 353, 90–95 (2015).

M. Tanha, R. Kheradmand, and S. Ahmadi-Kandjani, “Gray-scale and color optical encryption based on computational ghost imaging,” Appl. Phys. Lett. 101(10), 101108 (2012).

P. Latorre-Carmona, V. J. Traver, J. S. Sanchez, E. Tjahjaurence, “Online reconstruction-free single-pixel image classification,” Image and Vision Computing, 86, 28-37 (2019).

S. Jiao, J. Feng, Y. Gao, T. Lei, Z. Xie, X. Yuan, “Optical machine learning with incomplete light and a single-pixel detector,” Opt. Lett., 44(21), 5186-5189 (2019).

H. Chen, J. Shi, X. Liu, Z. Niu, G. Zeng, “Single-pixel non-imaging object recognition by means of Fourier spectrum acquisition,” Opt. Commun., 413, 269-275 (2018).

Y. Xu, K. F. Kelly, “Compressed domain image classification using a multi-rate neural network,” arXiv preprint arXiv:1901.00929 (2019).

J. Bacca, C. V. Correa, E. Vargas, S. Castillo, H. Arguello, “Compressive Classification from Single Pixel Measurements Via Deep Learning,” 2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP), 1-6, (2019).

N. Ahmed, T. Natarajan, and K. R. Rao, “Discrete cosine transform,” IEEE Trans. Computers 100(1), 90–93 (1974).

Z. Zhang, X. Li, M. Yao, S. Zheng, G., Zheng, J. Zhong, “Image-free real-time classification of fast moving objects using ‘learned’ spatial light modulation and a single-pixel detector,” arXiv preprint arXiv:1912.01974 (2019).

D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, https://arxiv.org/abs/1412.6980 (2014).

Z. Wen, B. He, R. Kotagiri, S. Lu, and J. Shi, “Efficient Gradient Boosted Decision Tree Training on GPUs,” In 2018 IEEE International Parallel and Distributed Processing Symposium (IPDPS), 234–243 (2018).

N. Dogru and A. Subasi, “Traffic accident detection using random forest classifier,” In 2018 15th Learning and Technology Conference, 40–45 (2018).

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).

G. A. Howland, P. B. Dixon, and J. C. Howell, “Photon-counting compressive sensing laser radar for 3D imaging,” Appl. Opt. 50(31), 5917–5920 (2011).

N. Radwell, K. J. Mitchell, G. M. Gibson, M. P. Edgar, R. Bowman, and M. J. Padgett, “Single-pixel infrared and visible microscope,” Optica 1(5), 285–289 (2014).

N. Tian, Q. Guo, A. Wang, D. Xu, and L. Fu, “Fluorescence ghost imaging with pseudothermal light,” Opt. Lett. 36(16), 3302–3304 (2011).

L. Wang, L. Zou, and S.M. Zhao, “Edge detection based on subpixel-speckle-shifting ghost imaging,” Opt. Commun. 407, 181–185 (2018).

H. D. Ren, S. M. Zhao, J. Gruska, “Edge detection based on single-pixel imaging,” Opt. Express 26(5), 5501–5511 (2018).

H. D. Ren, L. Wang, S. M. Zhao, “Efficient edge detection based on ghost imaging,” OSA Continuum 2(1), 64–73, (2019).

L. Wang, S. M. Zhao, W. Cheng, L. Gong and H. Chen, “Optical image hiding based on computational ghost imaging,” Opt. Commun. 366, 314–320 (2016).

W. Chen, and X. Cheng, “Marked ghost imaging,” Appl. Phys. Lett. 104, 251109 (2014).