High-fidelity imaging through multimode fibers via deep learning

Jun Zhao\(^1\), Xuanxuan Ji\(^2,\*\), Minghai Zhang\(\dagger\), Xiaoyan Wang\(\dagger\), Ziyang Chen\(\dagger\), Yanzhu Zhang\(\dagger\) and Jixiong Pu\(\dagger\)

\(^1\) School of Automation and Electrical Engineering, Shenyang Ligong University, Shenyang, Liaoning 110159, People's Republic of China
\(^2\) College of Information Science and Engineering, Fujian Provincial Key Laboratory of Light Propagation and Transformation, Huaqiao University, Xiamen, Fujian 361021, People's Republic of China
\(^3\) College of Engineering, Niigata University, Niigata 950-2181, Japan
\(^\dagger\) These authors contributed equally to this paper.

E-mail: syzd710471@163.com and jixiong@hqu.edu.cn

Keywords: deep learning, conventional neutral network (CNN), imaging, multimode fiber, speckle

Abstract

Imaging through multimode fibers (MMFs) is a challenging task. Some approaches, e.g. transmission matrix or digital phase conjugation, have been developed to realize imaging through MMF. However, all these approaches seem sensitive to the external environment and the condition of MMF, such as the bent condition and the movement of the MMF. In this paper, we experimentally demonstrate the high-fidelity imaging through a bent MMF by the conventional neural network (CNN). Two methods (accuracy and Pearson correlation coefficient) are employed to evaluate the reconstructed image fidelity. We focus on studying the influence of MMF conditions on the reconstructed image fidelity, in which MMF for imaging is curled to different diameters. It is found that as an object passes through a small bent diameter of the MMF, the information of the object may loss, resulting in little decrease of the reconstructed image fidelity. We show that even if MMF is curled to a very small diameter (e.g. 5 cm), the reconstructed image fidelity is still good. This novel imaging systems may find applications in endoscopy, etc.

1. Introduction

Optical fibers have extensively been used in telecommunication systems and endoscopy [1–3]. For the endoscopic imaging, the multimode fibers (MMFs) are frequently implemented, due to that more independent spatial modes can be used for information transmission, and wide-field endoscopic imaging can be achieved [4–6]. As well known, when light field passes through MMF, at the exit port of the MMF, a random speckle pattern takes place due to the effects such as defects, bending, coupling and superposition between the modes etc [7–11]. Several approaches have been developed for re-focusing the speckle into a focal spot, or reconstruction of the original information. These techniques include wavefront shaping, digital phase conjugation, transmission matrix, etc [12–17]. For example, Choi and his co-researchers proposed the so-called lensless microendoscopy by a single multimode fiber for realizing scanner-free wide-field endoscopic imaging. However all these approaches seem to be sensitive to the external environment and the condition of MMF, such as the bending condition and the movement of the MMF. Especially, when a femtosecond (fs) laser instead of a continuous wave (CW) laser through a long GRIN MMF, a speckle pattern will become more complex, due to the modal dispersion among the various modes [6].

With the development of the new technique—deep learning (DL), DL has been proved to be an effective approach in a revival of applications related to imaging [16–20]. For instance, by using DL, one can recover the imaging of an object transmitted through MMF. Also, the convolutional neural network (CNN) as one kind of DL has found applications in imaging system, for reconstructing image as the object passing through scattering medium [15–22]. In contrast to conventional imaging technique, CNN can learn the forward operator and regularizer implicitly through training process without having prior knowledge of them. CNN requires a training process to construct a computational architecture, that accurately images the object
transmitted through MMF or scattering media. To train the CNN, a large number of matched input (images of original objects) and output (the transported raw images) pairs are needed to optimize the parameters of the neural network and build a suitable computation architecture.

In this paper, we investigate the image reconstruction from a speckle generated by passing the light field with phase-only-type objects information through MMF. Two methods (accuracy, and Pearson correlation coefficient (PCC)) are employed to evaluate the reconstructed image fidelity. It is found that high-fidelity imaging reconstruction is obtained. We also study the influence of bending of MMF on the imaging reconstruction fidelity, in which the MMF are bent to different diameter. It seems that the bending of MMF does affect the imaging reconstruction fidelity slightly.

2. Experimental setup

The optical experiment setup is illustrated in figure 1, the wavelength of the laser (Origami-10XP, One-five) is 1028 nm, the pulse width is 400 fs, and the repetition rate is 50 kHz. The laser beam expanded by the telescope system combined of two lenses $L_1$ and $L_2$ passes through a polarization plate to become a linearly polarized one. Then the laser beam illuminates the phase-type reflected spatial light modulator (SLM) (PLUTO-NIR-015), and reflected light from the SLM is then coupled into the input port of the GRIN MMF (M31L10, core diameter: 62.5 $\mu$m, cladding diameter: 125 $\mu$m, and NA = 0.275, THORLABS) by using a lens $L_3$ and an objective microscope ($O_1$). The output port of MMF is a speckle pattern, and is collected by another objective microscope ($O_2$) and subsequently recorded by a CCD (AVT PIKE, F421B). Here we use two different bending diameters of MMF, in which their diameters are (a) 20 cm, and (b) 5 cm, respectively. By using MMF of these two different bending diameters, we want to investigate the influence of the bending of MMF on the imaging fidelity.
Figure 2 illustrates how to generate a phase-only-type object from an SLM and its propagation through an MMF. Just as shown in figure 2(a), a laser beam is incident upon a phase-only-type SLM in which an object ‘7’ is imposed, and the reflected beam with object information seems different from the incident light beam. To train CNN, we need a lot of objects. Fortunately, a lot of objects can be available from datasets (MNITS handwritten digit [23] and Quickdraw Object [24] databases). As the reflected beam from SLM passes through MMF, the speckle pattern occurs at the output port of MMF. It is shown from figure 2(b) that the reflected beam with different phase objects are different, therefore speckle patterns generated from the light beams with different objects passing through MMF is certainly different. Based on trained CNN and the speckle pattern, one can reconstruct the image of the original object.

3. Method

The CNN is a common network structure in deep learning, including neural network with convolutional computation and deep structure. The CNN usually consists of a convolutional layer and a pooling layer. It is divided into supervised learning and unsupervised learning. Supervised learning trains samples with conceptual markers so as to learn outside the training sample set, while the unsupervised learning trains samples without markers are to discover the structure knowledge in training set. The network structure we used is a U-net [25], which is shown in figure 3. The U-net is a full CNN without a fully connection layer, which is often used for medical image segmentation. The network is mainly divided into two parts, being an encoder and a decoder. The U-net is consisted of 23 convolutional layers that each layer of convolution processing uses a $(3 \times 3)$ convolution kernel. Each convolution in the encoder is followed by a nonlinear activation function (ReLU) and a $(2 \times 2)$ maximum pooling operation with a stride of 2 for down-sampling, which doubles the number of feature channels during down-sampling. In the decoder process, the deconvolution technique is adopted, and thus the feature is deconvolution at each step. The convolution kernel of $(2 \times 2)$ is used to deconvolution so that the number of feature channels is halved, the size of the feature map is doubled, and then the relative feature layer of the encoder is performed. For preserving high frequency information, information at different spatial scales is transmitted through connections. Next, the convolution processing and activation functions are the same as those of the decoder.

CNN requires a training process to generate a computational architecture, so as to accurately reconstruct the images from the speckle patterns at the distal end of the MMF. Training of the CNN needs a large number of matched input (images of original objects) and output (the transported raw images, i.e. speckle patterns) pairs to optimize the parameters of the neural network and build a suitable computational architecture. In this paper, four sets of experiments are done, which are listed as follows:

Set 1: we take 12 000 speckle images of the digital objects (1–9), in which the length of MMF is 10 m, and bent diameter is 20 cm;
Set 2: we take 12 000 speckle images of the digital objects (1–9), in which the length of MMF is 10 m, and bent diameter is 5 cm;
Set 3: we take 12 000 speckle images of the graffiti objects 12 000, in which the length of MMF is 10 m, and bent diameter is 20 cm;
Set 4: we take 12 000 speckle images of the graffiti objects 12 000, in which the length of MMF is 10 m, and bent diameter is 5 cm.
Figure 4. (a) The accuracy for training process and validation process is a function of epoch, (b) The loss rate for training process and validation process is a function of epoch. In (a) and (b), the length of MMF is 10 m, and the curled diameter of the MMF is 20 cm. (c) The accuracy for training process and validation process is a function of epoch, (d) The loss rate for training process and validation process is a function of epoch. In (c) and (d), the length of MMF is 10 m, and the curled diameter of the MMF is 5 cm.

Four sets of data are sent to the CNN network for training, in which the parameters of the CNN network are chosen to be fixed. Based on the trained CNN, we can reconstruct the images from the speckles, due to that the relation between the reconstruction image and the speckle image has been built.

CNN is designed for achieving sparse images [17], however mean squared error (MSE) and mean absolute error (MAE) cannot be evaluated effectively on the sparse image [26], therefore we must introduce the cross-entropy as a network performance assessment-loss function, which is shown as follows [27]:

$$\text{Loss} = -\frac{1}{n} \sum_x (y \log(y^{(i)}))$$  \hspace{1cm} (1)$$

here $y$ represents the ground-truth image by the pixel value, $y^{(i)}$ is the pixel value of the predicted image, $x$ represents each pixel point, and $n$ is the total number of pixel points. In this paper, we calculate the loss function of every pixel point for evaluating the network performance. The CNN is trained in Huaqiao University processing with the GPU (NVIDIA, RTX 2080 SUPER) using Keras/Tensorflow.

In order to avoid over-fitting, the training sets are processed in 16 image batches for the reconstruction network. In general, an optimizer called as Adam optimizer, is used to optimize the deep convolution network for minimizing the loss function [28]. Additionally, the parameter of momentum is chosen to be 0.99 and the learning rate reaches $1 \times 10^{-4}$. It is found that the network can achieve optimal performance after 150 iterations.

4. Results and analysis

Figure 4 presents the accuracy and loss of the trained CNN as a function of epoch, showing the performance for achieving the reconstructed image of the input object (digital), in which the length of MMF is 10 m, and the curled diameters of MMF are 20 cm (figures 4(a) and (b)) and 5 cm (figures 4(c) and (d)), respectively. It is shown from figures 4(b) and (d) that with the increment of epoch, the loss of train and validation decreases to the value lower than 0.1. It is also shown from figures 4(a) and (c) that the accuracy of both train and validation increases to 0.981 (the curled diameter is 20 cm) and 0.971 (the curled diameter is 5 cm), respectively. This indicates that the smaller is the curled diameter of MMF, the lower is its accuracy. The reduced fidelity to the bending may attribute to that too small diameter of MMF results in the leak of
Figure 5. Image reconstruction of speckle images corresponding to digital and graffiti labels input to the MMF using CNNs. (a) Set 1: The MMF is curled to be diameter of 20 cm, and the input object is digital; Set 2: The MMF is curled to be diameter of 5 cm, and the input object is digital. (b) Set 3: The MMF is curled to be a diameter of 20 cm, and the input object is graffiti objects. Set 4: The MMF is curled to be a diameter of 5 cm, and the input object is graffiti objects (such as airplane, candle, cloud, glasses, and moon, etc). The length of MMF is 10 m. The yellow mark in the figures indicates the value of accuracy of reconstruction image. (c) The average accuracy of image for four sets of reconstructed images.

There are two methods for evaluating the reconstructed image fidelity. The first one is to compare the input object and the reconstructed image, by performing the pixel-by-pixel contrast. For simplicity, we call the parameter of this method for evaluating the reconstruction as accuracy of image \([29]\). The other is called as PCC. It is shown that these two methods can quantitatively represent the similarity between the input object and the reconstructed image \([30]\). It is commonly known that higher is parameter (accuracy or PCC), the better is the reconstructed image fidelity.

Now we employ the accuracy of image to evaluate the reconstructed image fidelity. In figures 5(a) and (b), we present a part of our results, in which the number in the third column is the accuracy of image. We find that the reconstructed image fidelity is very good, e.g. for the case that the curled diameter of MMF is 20 cm, the value of accuracy for all reconstructed digits is larger than 0.9. Specially for reconstructed image of the digit ‘5’, the value of accuracy reaches 0.9813. In the case that the MMF is curled to be the diameter of 5 cm, the accuracy value is still high. However, for some digits (e.g. ‘2’, and ‘6’), the values of accuracy are a little low, e.g. 0.8640, and 0.8893. The reason for that may attribute to the loss of information of input objects as the laser beam with information transmits through the MMF of the smaller diameter, due to leaking of the laser beam from the MMF. In figure 5(c), we give the average accuracy of image for four sets of reconstructed images. It is shown that for the same objects (e.g. digital, or graffiti), the smaller curled diameter of MMF is, the smaller is the average accuracy of image. It is also shown from figure 5(c) that, the smaller diameter (Set 2 and Set 4, the curled diameter is 5 cm) is, the bigger error is.

Now we shift to use PCC for evaluating the reconstructed image fidelity. Figure 6 plots the PCC values of the reconstructed image fidelity for Set 1 experiment and Set 2 experiment. Comparing to the evaluation
Figure 6. PCC values of reconstructing different digital. (a) The PCC average value for Set 1 experiment is 0.833; the PCC average value for Set 2 experiment is 0.731. Set 1 experiment: The curled diameter of MMF is 20 cm. Set 2 Experiment: The curled diameter of the MMF is 5 cm. The length of MMF is 10 m. (b) The PCC values for reconstructing some input objects, such as airplane, candle, cloud, glasses, and moon. The average PCC value for Set 3 experiment is 0.91, and the PCC average value for Set 4 experiment is 0.83. The length of MMF is 10 m. And in Set 1 and Set 3 experiment, the curled diameter of MMF is 20 cm. In Set 2 and Set 4 experiment, the curled diameter of the MMF is 5 cm.

parameter of image (the accuracy), the values of the PCC become smaller for evaluating the same reconstructed image. This means that for evaluating the reconstructed image fidelity by PCC or accuracy, the value of the PCC is smaller than that of the accuracy. The average PCC value of Set 1 experiment is about 0.833, and the average value of Set 2 experiment is about 0.731. This indicates that the reconstructed image fidelity is better for the larger curled diameter of the MMF. Figure 6(b) presents the PCC values for reconstructing some input objects, which we classified as Set 3 experiment and Set 4 experiment. As shown in Figure 6(b), the values of PCC are quite high. The average PCC value for Set 3 experiment and Set 4 experiment are 0.91 and 0.83, respectively. It is shown that the smaller curled diameter of MMF damages the image reconstruction performance.

5. Conclusion and expectation

In this paper, we have experimentally demonstrated the high-fidelity imaging through a bent MMF by the CNN. It have been shown that the trained CNN can reconstruct high-fidelity images from the speckle patterns. The influence of MMF conditions on reconstructed image fidelity was also studied, in which MMF for imaging was curled to different diameters. It was found that as an object passes through the MMF with a small bent diameter, the information of the object may loss, resulting in somewhat decrease of the reconstructed image fidelity. We also found that even if MMF was curled to a very small diameter (5 cm), the reconstructed image fidelity was still good. We only tested the fidelity of reconstructed image at the condition that the curled diameter of the fiber is 5 cm, and 20 cm, respectively. It can be expected that when the curled diameter of the fiber is larger than 20 cm, the fidelity of the reconstructed image is nearly unchanged. Moreover, how to design the CNN for keeping the high fidelity of reconstructed image even when the bending radius of fiber is very small seems important to the applications of this technique. Further work on improving the performance of the CNN will be done.

The laser source we used is an fs one, and its wavelength is 1028 nm. The absorption of MMF at this wavelength is suitable, therefore we use this laser to perform the experiment. Such laser does have a reduced temporal coherence, which may blur the speckle. Fortunately, the pulse width is not too short, about 400 fs, and the temporal coherence is reasonable to achieve clear speckle pattern as a object passing through MMF. It is certain that if we use a higher coherence laser source (such as a He-Ne laser), the speckle patterns are sure to be clearer, and the reconstruction image fidelity is expected to be larger.

Acknowledgments

This work was supported by National Natural Science Foundation of China (NSFC) under Grant Nos. 11674111, 61575070, and Fujian Province Science Funds for Distinguished Young Scholar 2018J06017.
References

[1] Richardson D J, Fini J M and Nelson L E 2013 Space-division multiplexing in optical fibres Nat. Photonics 7 354
[2] Krupa K, Loutz C and Couder V 2016 Spatiotemporal characterization of supercontinuum extending from the visible to the mid-infrared in a multimode graded-index optical fiber Opt. Lett. 41 5785
[3] Picozzi A, Millot G and Wabnitz S 2015 Nonlinear optics: nonlinear virtues of multimode Nat. Photonics 9 289
[4] Gu R Y, Mahalati R N and Kahn J M 2015 Design of flexible multi-mode fiber endoscope Opt. Express 23 26905
[5] Loterie D, Farahi S and Papadopoulos I 2015 Digital confocal microscopy through a multimode fiber Opt. Express 23 23845
[6] Sivanukutty S, Andresen E R and Cossart R 2016 Ultra-thin rigid endoscope: two-photon imaging through a graded-index multi-mode fiber Opt. Lett. 24 825
[7] Gusachenko I, Chen M Z and Dholakia K 2017 Raman imaging through a single multimode fibre Opt. Express 25 10580
[8] Fan W R, Chen Z Y and Gezhi Z X B M 2019 Control the normalized polarization ratio of a focal spot in speckle field formed by on-polarization-maintaining multimode fiber J. Opt. 21 045704
[9] Gover A, Lee C P and Yariv A 1976 Direct transmission of pictorial information in multimode optical fibers J. Opt. Soc. Am. 66 306
[10] Cižmár T and Dholakia K 2012 Exploiting multimode waveguides for pure fibre-based imaging Nat. Commun. 3 1027
[11] Papadopoulos I N, Farahi S, Moser C and Psaltis D 2012 Focusing and scanning light through a multimode optical fiber using digital phase conjugation Opt. Express 20 10583
[12] Choi Y 2012 Scanner-free and wide-field endoscopic imaging by using a single multimode optical fiber Phys. Rev. Lett. 109 203901
[13] LeCun Y, Bengio Y and Hinton G 2015 Deep learning Nature 521 436
[14] Luu M, Wang H, Li G and Situ G 2017 Exploit imaging through opaque wall via deep learning arXiv Preprint. arXiv 1708 07881
[15] Yang M, Liu Z, Cheng Z, Xu J, Li C-F and Guo G-C 2019 Deep hybrid scattering image learning J. Phys.: D: Appl. Phys. 52 115105
[16] Horisaki R, Takagi R and Tamida J 2016 Learning-based imaging through scattering media Opt. Express 24 13758
[17] Li Y, Xue Y and Tian I. 2018 Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media Optica 5 1181
[18] Wang K et al 2019 Y-Net: a one-to-one deep learning framework for digital holographic reconstruction Opt. Lett. 44 4765
[19] Borhani N, Kakavava E, Moser C and Psaltis D 2018 Learning to see through multimode fibers Optica 5 960
[20] Fan P, Zhao T and Su L 2019 Deep learning the high variability and randomness inside multimode fibers Opt. Express 27 20241
[21] Rahmani B, Loterie D and Konstantinou G 2018 Multimode optical fiber transmission with a deep learning network Light-Sci. Appl. 7 69
[22] Redding B, Popoff S M and Cao H 2013 All-fiber spectrometer based on speckle pattern reconstruction Opt. Express 21 6584
[23] (MNIST: http://yann.lecun.com/exdb/mnist/)
[24] (QuickDraw: https://quickdraw.withgoogle.com/data)
[25] Lyu M, Wang H, Li G, Zheng S and Situ G 2019 Learning-based lensless imaging through optically thick scattering media Adv. Photonics 1 036002
[26] Kendall A and Gal Y 2017 What uncertainties do we need in Bayesian deep learning for computer vision? Advances in Neural Information Processing Systems
[27] Rubinstein R Y and Kroese D P 2004 The Cross-Entropy Method (Berlin: Springer)
[28] Kingma D and Ba J 2014. Adam: A method for stochastic optimization Proc. of the 3rd Int. Conf. on Learning Representations (ICLR 2015)
[29] González B C 2007 Digital Image Processing 3rd edn (Englewood Cliffs, NJ: Prentice Hall)
[30] Goshal A A 2012 Image Registration: Principles, Tools and Methods (Berlin: Springer)