Current Optics and Photonics
Vol. 4, No. 4, August 2020, pp. 286-292

ISSN: 2508-7266(Print) / ISSN: 2508-7274(Online)
DOI: https://doi.org/10.3807/COPP.2020.4.4.286

Deep Learning: High-quality Imaging through Multicore Fiber

Liqing Wu[1], Jun Zhao[2], Minghai Zhang[2], Yanzhu Zhang[2], Xiaoyan Wang[1], Ziyang Chen[1], and Jixiong Pu[1]*

[1]*College of Information Science and Engineering, Fujian Provincial Key Laboratory of Light Propagation and Transformation, Huaqiao University, Xiamen 361021, China*
[2]*School of Automation and Electrical Engineering, Shenyang Ligong University, Shenyang 110159, China*

(Received February 3, 2020 : revised April 13, 2020 : accepted May 8, 2020)

Imaging through multicore fiber (MCF) is of great significance in the biomedical domain. Although several techniques have been developed to image an object from a signal passing through MCF, these methods are strongly dependent on the surroundings, such as vibration and the temperature fluctuation of the fiber's environment. In this paper, we apply a new, strong technique called *deep learning* to reconstruct the phase image through a MCF in which each core is multimode. To evaluate the network, we employ the binary cross-entropy as the loss function of a convolutional neural network (CNN) with improved U-net structure. The high-quality reconstruction of input objects upon spatial light modulation (SLM) can be realized from the speckle patterns of intensity that contain the information about the objects. Moreover, we study the effect of MCF length on image recovery. It is shown that the shorter the fiber, the better the imaging quality. Based on our findings, MCF may have applications in fields such as endoscopic imaging and optical communication.

Keywords : Deep learning, Multi-core fiber, Imaging
OCIS codes : (060.2350) Fiber optics imaging; (060.4250) Networks; (100.3010) Image reconstruction techniques; (100.4996) Pattern recognition, neural networks

I. INTRODUCTION

It is commonly known that multicore fiber (MCF) has attracted much interest in biological and medical imaging, optical communication, *etc*. [1-4]. Due to the existence of mode coupling and mode superposition, only a random phase speckle pattern can be obtained at the exit of the fiber after a coherent beam passes through the MCF, each core of which is multimode. Thus, this phenomenon greatly hinders imaging through the MCF. To reduce the loss of information and break the transmission limitation, each core of a MCF can be regarded as an information channel, which is similar to wavelength-division multiplexing (WDM) in optical communication. Controlling the imaging through MCF has been shown to have important applications in the field of biological imaging [5]. Many techniques have been implemented to enable imaging using MCF. For example, scanning endoscopy and wide-field imaging through MCF are successfully achieved by using speckle correlation and memory effect [6-8]. Based on the digital phase conjugation and memory effect of the MCF, researchers realized high-resolution imaging using focusing and scanning of a spot, without a long calibration procedure [9]. Wide-field endoscopic imaging is achieved using a transmission matrix without a scanner [3]. However, these techniques crucially depend on the stabilization of the external environment and MCF. Recently a new and strongly robust technique, *deep learning*, has solved this deficiency successfully in multimode fiber (MMF) and scattering media [10-17].

Deep learning has dramatically facilitated visual object recognition, object detection, and many other domains [18]. In 1991, artificial neural networks (ANNs) were proposed to recover the images transmitting through MMF [19]. It

*Corresponding author: jixiong@hqu.edu.cn, ORCID 0000-0001-8781-6683
Color versions of one or more of the figures in this paper are available online.

has been demonstrated that a convolutional neural network (CNN) is capable of successfully solving the problem of optical imaging [20-24]. Based on a CNN, a common and stable network structure in deep learning, we can improve the network for our optical system to recover phase-type objects with high quality.

In this paper we use the state-of-the-art technique of deep learning to realize the reconstruction of input images with high quality through MCF, in which each core carries some information about the object, so as to increase the number of transmission channels. Moreover, we employ a database of 10,000 handwritten digits for training the network and 2000 digits for testing. Due to the sparse ground truth we use, the binary cross-entropy is adopted as the loss function in the optical system using the CNN. Accordingly, an efficient reconstruction effect is obtained in our experiment. Meanwhile, different lengths of fiber are used in our experiment, to show the influence of fiber length on image recovery and the evaluation function. The results in this paper may have applications in optical communication and endoscopic imaging, for example.

II. METHODS

The CNN structure we used is based on the "U-net" architecture [25] and a dense net [26], to improve computation speed. Meanwhile, the dense block can converge rapidly for better training efficiency. As shown in Fig. 1, the collected speckle patterns are processed with downsampling, to retain core information, by the convolutional encoder block, which includes four dense blocks connected by a max pooling layer. Moreover, the dense block is composed of various layers, with each layer including batch normalization (BN), a rectified linear element (ReLU) with nonlinear activation, and convolution with 16 filters (conv), as shown in detail in Fig. 1. After downsampling, the feature maps of deep semantic information are transferred into the path of the decoder, connected by the upsampling convolution layer consisting of four additional dense blocks (see the green block in Fig. 1). Furthermore, to improve the performance of the image reconstruction, a skip connection is introduced in the network. Importantly, we design the convolution layer of a single channel for the last layer to output the reconstructed image, and thus it is easy to observe the recovery effect. The details of the CNN are presented in Table 1.

TABLE 1. The details of the network

	Name	Filter	Output size
Input			256 × 256 × 1
Encoding	Convolution	3 × 3 Conv	256 × 256 × 64
	Dense block	[5 × 5 Conv] × 4	256 × 256 × 128
	Convolution	3 × 3 Conv	128 × 128 × 128
	Dense block	[5 × 5 Conv] × 4	256 × 256 × 192
	Convolution	3 × 3 Conv	64 × 64 × 256
	Dense block	[5 × 5 Conv] × 4	64 × 64 × 320
	Convolution	3 × 3 Conv	32 × 32 × 512
	Dense block	[5 × 5 Conv] × 4	32 × 32 × 576
Bridge	Convolution	3 × 3 Conv	16 × 16 × 1024
	Dense block	[5 × 5 Conv] × 4	16 × 16 × 1088
Decoding	Convolution	3 × 3 Conv	32 × 32 × 512
	Dense block	[5 × 5 Conv] × 3	32 × 32 × 560
	Convolution	3 × 3 Conv	64 × 64 × 256
	Dense block	[5 × 5 Conv] × 3	64 × 64 × 304
	Convolution	3 × 3 Conv	128 × 128 × 128
	Dense block	[5 × 5 Conv] × 3	128 × 128 × 176
	Convolution	3 × 3 Conv	256 × 256 × 64
	Dense block	[5 × 5 Conv] × 3	256 × 256 × 112
	Convolution	3 × 3 Conv	256 × 256 × 32
	Convolution	1 × 1 Conv	256 × 256 × 1

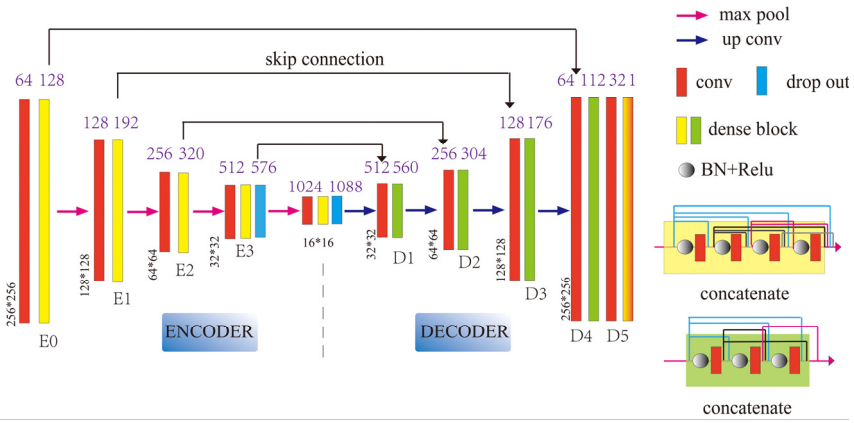
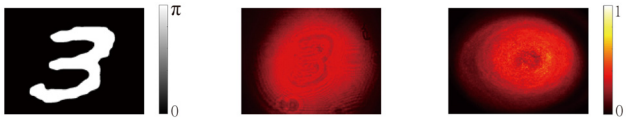


FIG. 1. The structure of the U-net is consisted of the general encoder and decoder parts. The input speckle pattern is downsampled by encoder and then sent to decoder for reconstruction.



(a) (b) (c)

FIG. 2. (a) The phase image imposed on the SLM. (b) The light-field pattern reflected from the SLM. (c) The detected speckle pattern at the exit of the fiber.

Owing to the limitations of the computer, the batch size is chosen to be 16 for network training. In particular, the Adam optimizer [27] is used to optimize deep convolutional networks to minimize loss functions. The momentum parameter is chosen to be 0.99 and the rate of learning 0.0001. During the process of training for the network, the rate of learning is appropriately adjusted. We experimentally test multiple groups to find the number of iterations that can achieve optimal network performance. Generally, the mean squared error (MSE) can be used to evaluate loss functions, though it is not ideal for sparse images [28]. Therefore, we choose as an alternative the binary cross-entropy as the loss function in our paper, which is shown as follows [29]:

$$\text{Loss}=\frac{1}{n}\sum_{i=1}^{n}\left(y_T \log\left(y_P\right)+\left(1-y_T\right)\log\left(1-y_P\right)\right), \quad (1)$$

where y_T represents the pixel of the true label image, y_P is the pixel of the predicted image, and n is the total number of pixels. In the optical system, we propose the selective binary cross-entropy, calculated for each pixel, to estimate the network performance for CNN. Hence, the network has a strong capacity for reconstructing phase object input from intensity-type speckle output.

Deep learning usually is divided two learning approaches: supervised learning and unsupervised learning. We adopt the former to train a network with speckle pattern as input. A CNN with a convolution layer and a pooling layer is one type of network for deep learning. A CNN model is built to learn the mapping relationship between the speckle pattern detected at the exit of the fiber (Fig. 2(c)) and the ground-truth image. The ground-truth image (Fig. 2(a)) of the network is imposed on SLM. Figure 2(b) is the light-field pattern reflected from the SLM. We build the training network by collecting 10,000 speckle patterns, in which 8000 images are used as training set for the CNN, and the other 2000 images are used as testing set. The network is so strong that it can recover arbitrary, nontrained handwritten digits of different types.

III. RESULTS

The experimental setup is shown in Fig. 3. An infrared beam of wavelength 1028 nm is emitted from the laser

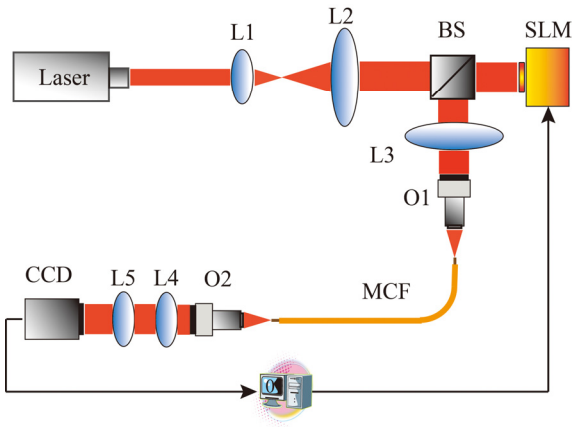


FIG. 3. Experimental setup. The focal lengths of lenses L1, L2, L3, L4, and L5 are 50, 150, 100, 100, and 50 mm respectively. The length of the MCF is 1 m or 2 m. The diameter of the MCF (seven cores) is about 1.75 mm, and the diameter of each core is about 0.55 mm. O1 and O2 are objective lenses with numerical aperture (NA) of 0.25.

(One five, 400 fs, 50 kHz). The lenses L1 and L2 are constructed to be a telescope system for expanding the light beam. The distance between L1 and L2 is 200 mm. The expanded light is split to the spatial light modulator (SLM, 1920 × 1080 pixels, Pluto-Vis, Holoeye) using a beam splitter (BS). The beam diameter of the light incident on the SLM is 7 mm, and the power is 30 mW. The SLM is phase-only, with a virtual phase object. The light beam carrying the information about the object is reflected from the SLM onto the proximal facet of the MCF (Thorlab, BF20LSMA02: each core's diameter, 550 ± 19 μm; effective core diameter, 1750 μm; NA = 0.22) by using a 4*f* imaging system (L3 and O1). The MCF is coiled in circles 10 cm in diameter. To obtain the object information, another 4*f* system (O2 and L4) is employed at the exit of the fiber. The light from the seven cores of the MCF coherently superpose to produce the speckle [30], which is detected by a CCD (PIKE F421B, 2048 × 2048, AVT). The pixel size of the CCD is 7.5 μm.

The CNN is trained at Huaqiao University processing with a GPU (NVIDIA, RTX 2080 SUPER), using Keras/Tensorflow. In our experiment, the ground-truth image of handwritten digit patterns is downloaded from the MNIST database [31]. The 12,000 handwritten digits imposed on the SLM are divided into the training and testing sets, which are 10,000 and 2000 in number respectively. It is worth noting that the number of digits for training the network can be chosen as 10,000 or 1000; the more digits we choose, the greater of the effect of generalization. In addition, to improve the speed of calculation, the speckle on the CCD and digits on the SLM are cropped to a 512 × 512 pixel window. Next, the collected speckle images are resized to 256 × 256 using the nearest-neighbor method, before being processed by the CNN. In the far field we

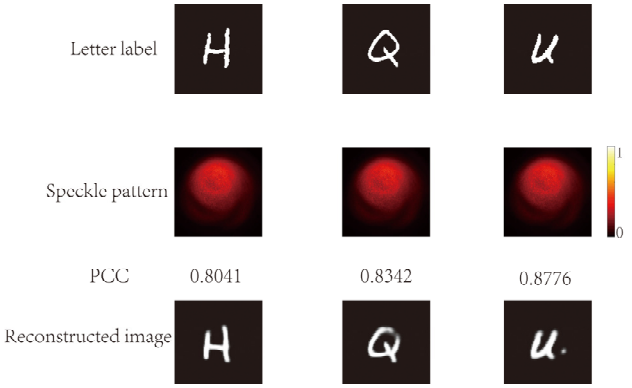


FIG. 4. Test results for the letters H, Q, and U, and their corresponding speckle patterns, based on the CNN.

can detect the coherent superposition of the seven light beams, so that each core's signal containing information about the object can be mixed. Therefore, the trained network of speckle patterns can be effectively reconstructed. For instance, as shown in Fig. 4, the letters H, Q, and U are generated on the SLM, and the corresponding speckle patterns after passing through 2 m of MCF are recorded by the CCD. Obviously, different object information is transmitted in the different modes of the MCF, so that the central intensity of speckle at the exit of the fiber is a little different for different objects on the SLM. For English letters, we need to train on a dataset based on the letter images before testing the speckle patterns of letters.

It has been shown that an amplitude object exhibits better reconstruction performance than a phase object, for the same training network [32]. Imaging nonsparse objects becomes more challenging, probably because the CNN is prone to reconstruct binary objects efficiently [20]. In our experiment, we impose a phase image on the SLM and obtain a high-quality reconstruction successfully, based on the improved CNN. A 4*f* system is placed to image the object onto the CCD. On account of the limitations of the computer, the speed of recording the output at the distal end of the fiber is 1 image per 1.5 s. Due to mode coupling, dispersion, and superposition of modes, the light field bearing the object information is scattered after passing through the MCF. However, the speckle acquired at the exit end of the MCF still contains the object information. To test the influence of the length of MCF on network reconstruction, we select 1-m and 2-m fibers in our experiment. Generally, speckle is sensitive to the external surroundings, such as temperature, mechanical vibrations *etc*. Hence, to better compare the effect of different fiber lengths on recovery, we set the same optical experimental conditions to train the identical network with two sets of speckle patterns as input, respectively. The recovery results for 1-m and 2-m MCFs with phase-modulated inputs are presented in Fig. 5. We can see that although the speckle intensity is not so uniform, the reconstructed image is still excellent.

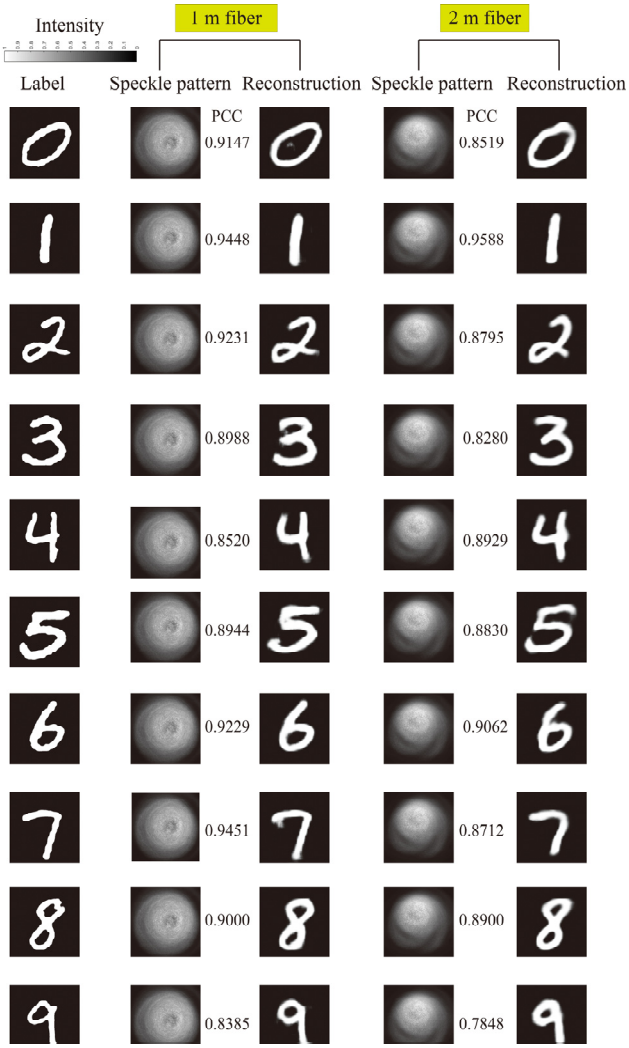


FIG. 5. Comparison of 1-m fiber to 2-m fiber, in the results of reconstruction using the U-net network.

To measure the results of reconstruction, an important parameter is introduced: the accuracy of the network, which is computed as percent mean square error. In addition, we also use the so-called Pearson correlation coefficient (PCC) between the ground-truth object and reconstructed image to evaluate the reconstruction effect. Notably, for each digit the fidelity of the 1-m fiber is mostly higher than that of the 2-m fiber, as shown in the Fig. 5. Furthermore, we plot a bar chart for both fibers with mean and standard deviation of PCC in Fig. 6, which reveals clearly the Pearson correlation coefficient for 1-m and 2-m fibers. This indicates that the longer the distance through which the light is transmitted in the fiber, the larger the loss of object information, and the more affected it is by the environment. The longer fiber is susceptible to environmental influences, so the speckle pattern contains more noise, which affects image reconstruction. We also perform an experiment on the recovery of images of English letters passed through MCF. We find from our experimental

results that the accuracy of digit imaging is higher than that of letter imaging. The average accuracy of the reconstructed digits is greater than 0.930, while the accuracy of the reconstructed letters H, Q, and U is estimated to be 0.8560, 0.8942, and 0.9176 respectively (see the reconstructed images in Fig. 4).

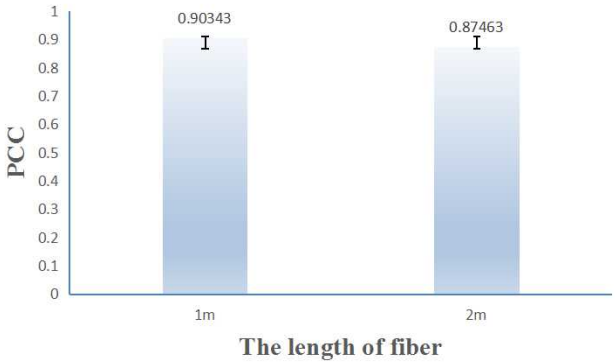


FIG. 6. Bar chart showing the mean and standard deviation of the PCC of imaging reconstruction through MCF of lengths 1 m and 2 m.

In the CNN model, accuracy and loss are regarded as significant parameters for judging the network system. The number of iterations will also affect performance. In this paper, the network is trained for 80 iterations, and good recovery is obtained. In Fig. 7, we present the results for the accuracy and loss as a function of epoch, for 1-m and 2-m fibers respectively. Due to adjustment of the learning rate in the process of deep learning, sharp changes in the curves may occur, as seen in Fig. 7. Meanwhile, the accuracy of the training and validation are revealed in Figs. 7(a) and 7(c). It is seen that the results for training are better than those for validation. With increasing epochs, the graph at first changes rapidly, and then slowly approaches a steady state. Similarly, the loss first decreases quickly, then tends to a minimum value. On the whole, our trained CNN is capable of achieving high-resolution reconstruction of an object's image passing through MCF.

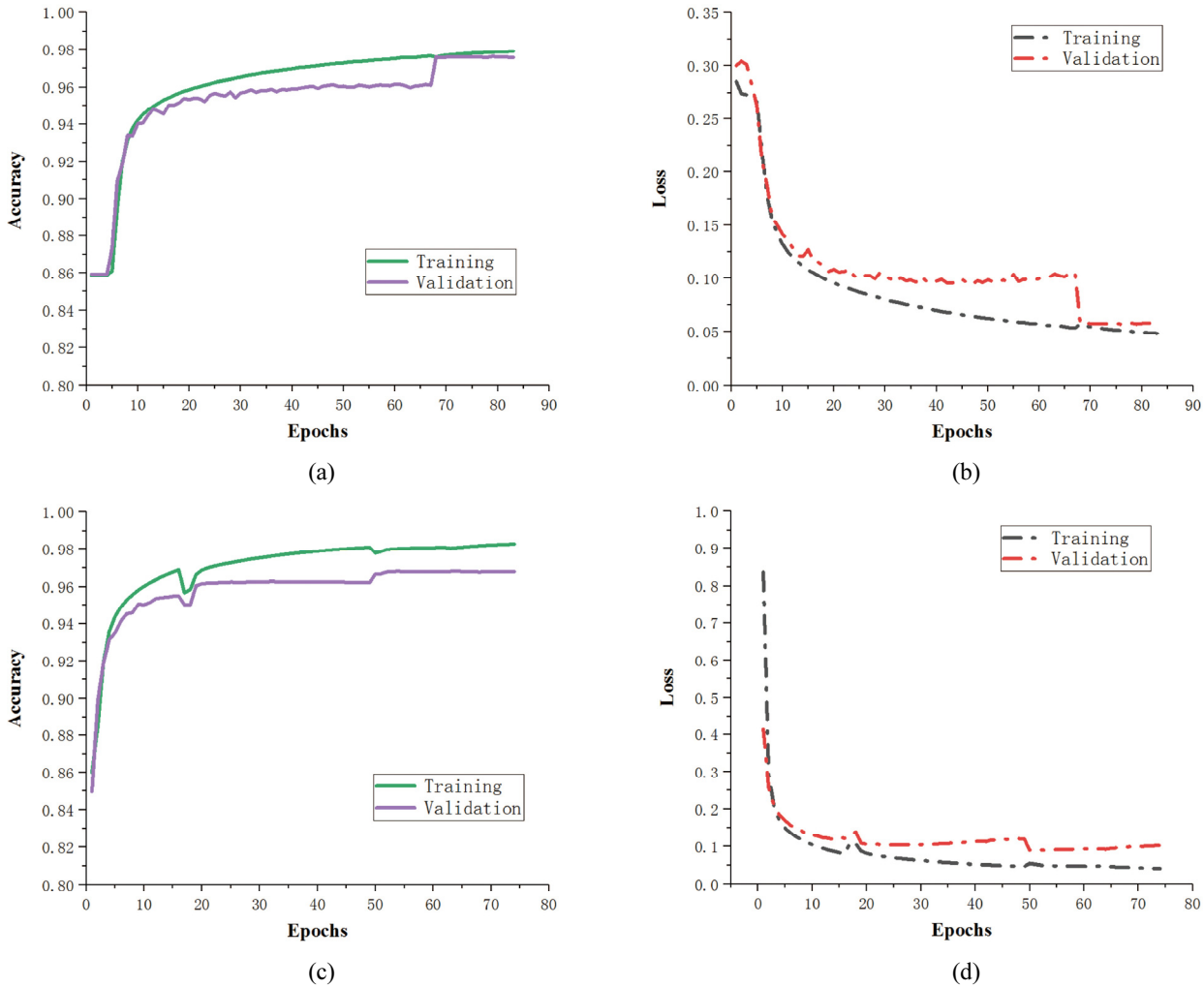


FIG. 7. Training and validation accuracy as a function of epoch, for the distal speckle intensity pattern of (a) 1-m fiber and (c) 2-m fiber; and loss as a function of epoch for the distal speckle intensity pattern of (b) 1-m fiber and (d) 2-m fiber.

IV. CONCLUSION

MCF plays an important role in endoscopic imaging and optical communication. The mode coupling and superposition of modes in MCF impede high-quality imaging. In recent years, several techniques have been implemented in computational imaging when light passes through MCF, but these techniques mostly are sensitive to the external environment and variation of the MCF. In this paper we employed the strong technique of deep learning to realize high-quality imaging through a MCF with seven cores. Based on a CNN, we obtained high-quality image reconstruction. It was shown that the deep learning technique was insensitive to variation of the MCF or the surroundings. The influence of the length of MCF on imaging quality was also studied, with the results demonstrating that the longer the length of MCF, the poorer the imaging quality. In these experiments we used accuracy and PCC to evaluate the performance of the network. The binary cross-entropy, which can efficiently estimate the reconstruction of a phase image, was used as the loss function for the optical system. It was estimated that the best accuracy for 1-m and 2-m fibers were 98.1% and 97.6% respectively. It was shown that the training set for our neural network was memorized very well. We envision that the technique and results of our experiments can be meaningful for biomedical imaging.

ACKNOWLEDGMENT

The authors acknowledge the National Natural Science Foundation of China (NSFC) (11674111, 61575070, 11750 110426), Fujian Province Science Funds for Distinguished Young Scholars (2018J06017), Promotion Program for Young and Middle-aged Teachers in Science and Technology Research of Huaqiao University (ZQN-PY209), and the Graduate Student Research and Innovation Ability Training Program of Huaqiao University (18013082015).

REFERENCES

1. I. N. Papadopoulos, S. Farahi, C. Moser, and D. Psaltis, "High-resolution, lensless endoscope based on digital scanning through a multimode optical fiber," Biomed. Opt. Express **4**, 260-270 (2013).
2. T. Čižmár and K. Dholakia, "Exploiting multimode waveguides for pure fibre-based imaging," Nat. Commun. **3**, 1027 (2012).
3. Y. Choi, C. Yoon, M. Kim, T. D. Yang, C. Fang-Yen, R. R. Dasari, K. J. Lee, and W. Choi, "Scanner-free and wide-field endoscopic imaging by using a single multimode optical fiber," Phys. Rev. Lett. **109**, 203901 (2012).
4. K. Krupa, A. Tonello, B. M. Shalaby, M. Fabert, A. Barthélémy, G. Millot, S. Wabnitz, and V. Couderc, "Spatial beam self-cleaning in multimode fibres," Nat. Photonics **11**, 237-241 (2017).
5. M. Hughes, T. P. Chang, and G.-Z. Yang, "Fiber bundle endocytoscopy," Biomed. Opt. Express **4**, 2781-2794 (2013).
6. V. Tsvirkun, S. Sivankutty, G. Bouwmans, O. Katz, E. R. Andresen, and H. Rigneault, "Widefield lensless endoscopy with a multicore fiber," Opt. Lett. **41**, 4771-4774 (2016).
7. A. Porat, E. R. Andresen, H. Rigneault, D. Oron, S. Gigan, and O. Katz, "Widefield lensless imaging through a fiber bundle via speckle correlations," Opt. Express **24**, 16835-16855 (2016).
8. N. Stasio, C. Moser, and D. Psaltis, "Calibration-free imaging through a multicore fiber using speckle scanning microscopy," Opt. Lett. **41**, 3078-3081 (2016).
9. N. Stasio, D. B. Conkey, C. Moser, and D. Psaltis, "Light control in a multicore fiber using the memory effect," Opt. Express **23**, 30532-30544 (2015).
10. P. Fan, T. Zhao, and L. Su, "Deep learning the high variability and randomness inside multimode fibers," Opt. Express **27**, 20241-20258 (2019).
11. P. Caramazza, O. Moran, R. Murray-Smith, and D. Faccio, "Transmission of natural scene images through a multimode fibre," Nat. Commun. **10**, 2029 (2019).
12. N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, "Learning to see through multimode fibers," Optica **5**, 960-966 (2018).
13. R. Horisaki, R. Takagi, and J. Tanida, "Learning-based imaging through scattering media," Opt. Express **24**, 13738-13743 (2016).
14. K. Wang, J. Dou, Q. Kemao, J. Di, and J. Zhao, "Y-Net: a one-to-two deep learning framework for digital holographic reconstruction," Opt. Lett. **44**, 4765-4768 (2019).
15. H. Wang, M. Lyu, and G. Situ, "eHoloNet: a learning-based end-to-end approach for in-line digital holographic reconstruction," Opt. Express **26**, 22603-22614 (2018).
16. T. Nguyen, V. Bui, V. Lam, C. B. Raub, L.-C. Chang, and G. Nehmetallah, "Automatic phase aberration compensation for digital holographic microscopy based on deep learning background detection," Opt. Express **25**, 15043-15057 (2017).
17. Y. Wu, Y. Rivenson, Y. Zhang, Z. Wei, H. Günaydin, X. Lin, and A. Ozcan, "Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery," Optica **5**, 704-710 (2018).
18. Y. L. Cun, Y. Bengio, and G. Hinton, "Deep learning," Nature **521**, 436-444 (2015).
19. S. Aisawa, K. Noguchi, and T. Matsumoto, "Remote image classification through multimode optical fiber using a neural network," Opt. Lett. **16**, 645-647 (1991).
20. S. Li, M. Deng, J. Lee, A. Sinha, and G. Barbastathis, "Imaging through glass diffusers using densely connected convolutional networks," Optica **5**, 803-813 (2018).
21. X. Yuan and Y. Pu, "Parallel lensless compressive imaging via deep convolutional neural networks," Opt. Express **26**, 1962-1977 (2018).
22. C. Park, C. C. Took, and J.-K. Seong, "Machine learning in biomedical engineering," Biomed. Eng. Lett. **8**, 1-3 (2018).
23. C. M. Sandino, J. Y. Cheng, F. Chen, M. Mardani, J. M. Pauly, and S. S. Vasanawala, "Compressed sensing: from research to clinical practice with deep neural networks: Shortening Scan Times for Magnetic Resonance Imaging," IEEE Signal Process. Mag. **37**, 117-127 (2020).
24. K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock, and F. Knoll, "Learning a variational

network for reconstruction of accelerated MRI data," Magn. Reson. Med. **79**, 3055-3071 (2018).
25. O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Proc. Medical Image Computing and Computer-Assisted Intervention* (Munich, Germany, Oct. 2015), pp. 234-241.
26. G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Honolulu, USA, Jul. 2017), pp. 4700-4708.
27. P. D. Kingma and J. Ba, "Adam: a method for stochastic optimization," in *Proc. the 3rd International Conference for Learning Representations* (San Diego, USA, May 2015).
28. A. Kendall and Y. Gal, "What uncertainties do we need in Bayesian deep learning for computer vision?" in *Proc. Neural Information Processing Systems* (CA, USA, Dec. 2017), pp. 5574-5584.
29. B. Li, Y. Liu, and X. Wang, "Gradient harmonized single-stage detector," in *Proc. AAAI Conference on Artificial Intelligence* (Honolulu, USA, Jan. 2019), pp. 8577-8584.
30. L. Wu, W. Fan, Z. Chen, and J. Pu, "Focusing and polarized modulation of a laser passing through multi-core fiber," Opt. Rev. **26**, 531-536 (2019).
31. Y. L. Cun, C. Cortes, and C. J. C. Burges, *THE MNIST DATABASE of handwritten digits* (Y. L. Cun, C. Cortes, and C. J. C. Burges), http://yann.lecun.com/exdb/mnist/ (Accessed date: 2019.06.13.).
32. 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).