All-fiber ultrafast image detection enabled by deep learning: Supplementary Material.

Supplementary Note 1: Pulse Transmission

The max average power entering the MMF in the experiment is measured to be about 1.5 mW. Considering that the width of the pulses is 27.5 ps, the peak power of the pulses in the fiber is about 3.3 W. We simulate the evolution of a gaussian pulse with the same peak power in the 1-km MMF as shown in Fig. S1(a) without considering the intermodal dispersion. We see that the broadening caused by the chromatic dispersion and nonlinear effect is negligible. The calculation is based on the non-linear Schrodinger equation (NLSE)[1]. The simulation parameters are as follows. The dispersion coefficient $\beta_2 = 25 \times 10^{-3}\mathrm{ps}^2\mathrm{m}^{-1}$. the nonlinear coefficient is $\gamma = 5.8 \times 10^{-3}\mathrm{W}^{-1}\mathrm{m}^{-1}$ at 1064nm in the single mode fiber[2]. Considering $\gamma = \mathrm{n}_2\omega / \mathrm{cA}_{\mathrm{eff}}$, the γ is inversely proportional to the area the fiber core. Thus, in our MMF with 50-μm-diameter core, the γ is about 30 times that of the general single mode fibers with core diameter of around 9 μm. Therefore, we set the $\gamma = 2 \times 10^{-4}\mathrm{W}^{-1}\mathrm{m}^{-1}$ in the simulation. Besides, we also calculate the group delays for all the LP modes in the 1-km MMF that we use and the results are shown in Fig. S1(b). We see that different modes have different group delays. Compared with the pulse width, the delay differences between different modes are large enough that can cause the energy separating in these modes. Considering that the fastest and slowest modes have a delay difference of around 50 ns, we can predict that after transmitting through the 1-km MMF, a pulse will eventually split into a pile of isolated sub-pulses over a temporal range of 50 ns principally due to the intermodal dispersion.

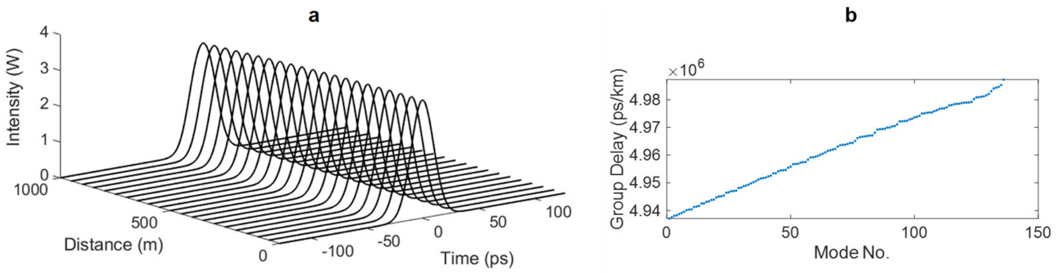

Fig. S1 (a) Evolution of the pulse in the MMF. (b) Group delays of different LP modes calculated by the Finite Difference Method.

Supplementary Note 2: Fiber Probe

The Section refractive index distribution of the fiber probe is shown in Fig. S2. It consists of three claddings. The first one is fluoride Doped layer that has a relative low index, so the signal light can be limited in the fiber core. The combination of a second silica cladding and a low-index coating layer allows the transmission of the illumination light.

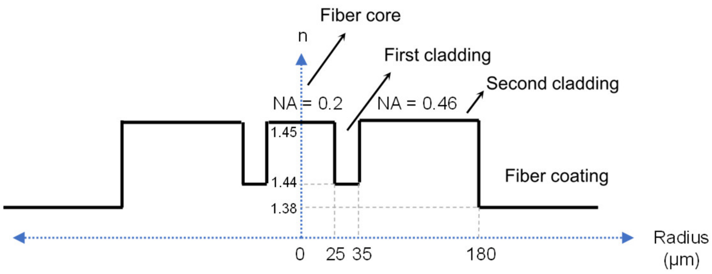

Fig. S2 Section refractive index distribution of the fiber probe.

Supplementary Note 3: Waveforms

Here we discuss about several features of the waveforms. When we put the ten waveforms in Fig. 3(b) of the main text together as shown in Fig. S3, we see that all the sub-pulses are overlapped, which proves that the mode dispersion actually dominates in the temporal evolution of the pulse so that the temporal positions of the sub-pulses will only be determined by the group delays of the corresponding modes. We observe from Fig. S3 that the burst of sub-pulses that covers a time range of around 45 ns, a little less than the 50 ns predicted in the Note 1. This may be attributed to that certain highest-order modes are harder than expected to be excited, or that a trivial deviation of the parameters of the real fiber from the idea values. Besides, we can see that every waveform contains about 40 sub-pulses (peaks in the waveform) which is much less than the number of all the LP modes in the MMF (136, see the calculation in Note 1). This is because some adjacent modes have very close group delays so that the light energy in these modes is not completely separated in the time domain. Instead, they will cause the broadening of the sub-pulses, which can be observed in the waveforms.

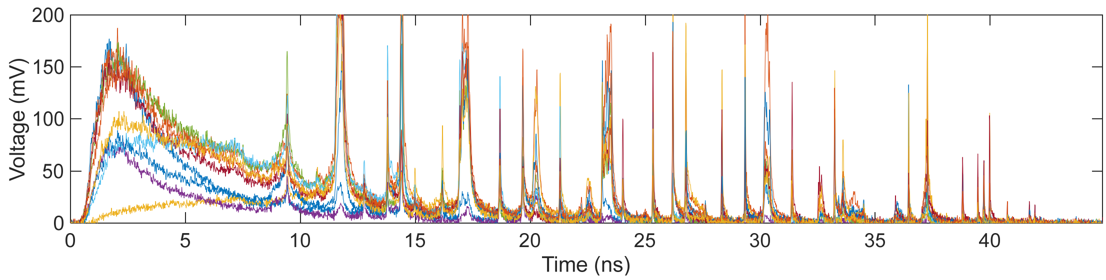

Fig. S3 Superposition of ten waveforms corresponding to different images.

Supplementary Note 4: Image Classification

We have tried two different networks for classification of the images of hand-written digits. The first is a CNN network as shown in Fig. S4(b). The waveforms reshaped into 64×64 matrices are as the input and the categories from 0 – 9 as the output of the CNN. The second is a combination of the U-net and the same CNN as shown in Fig. S4(a). The CNN model is pretrained with 60000 different images so it can act as a digit classifier, which is then used to directly classify the images recovered by the U-net. We use the 20000 images of digits and the recorded waveforms, including 17000 training set, 2000 valuation set and 1000 testing set to train and test the networks. The accuracy corresponding to the CNN and U-net + CNN is tested to be 91.5% and 82.0% respectively. The results show that the combination of two networks provide a higher accuracy, which is consistent with the previous research[3].

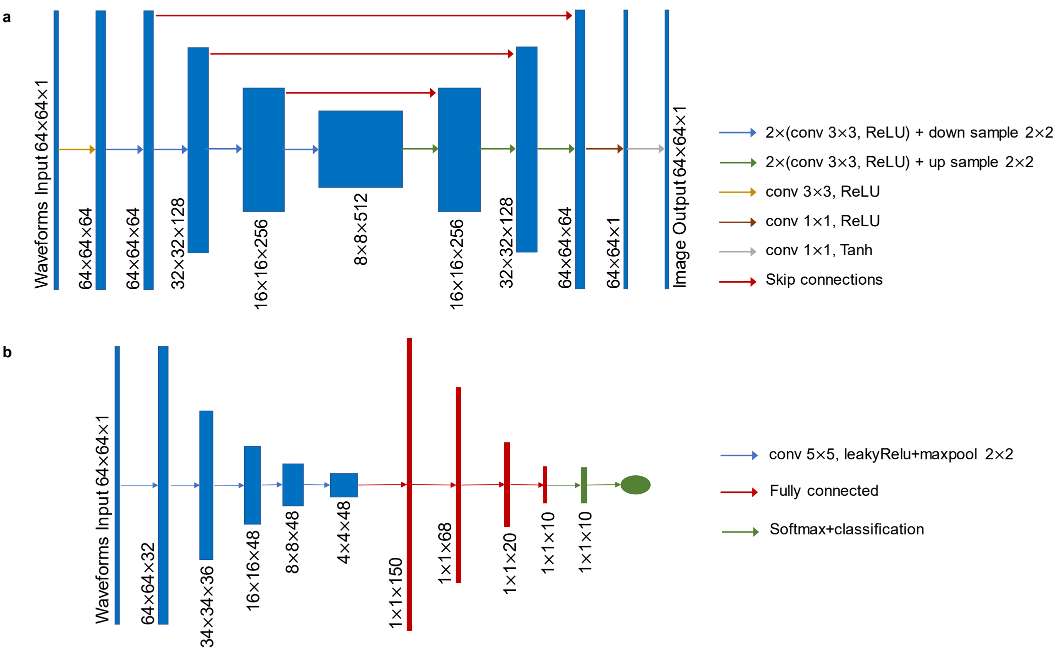

Fig. S4 (a) U-net structure. (b) CNN structure.

Supplementary Note 5: Feature Points

In the training process, the waveforms recorded by the oscilloscope have a resolution of 10 ps^{-1} which are used as the input of the network. There may exist some redundant information in the waveforms, such as noise and over-dense points. Thus, we select 256 data points near the peaks of the waveforms to down sample all the waveforms, which are marked in Fig. S5(a). This waveform is recorded when we use a mirror in front of the fiber probe instead of the DMD. We do so for exciting the modes in the MMF as much and as enough as possible, so that all the locations in the time domain that may appear sub-pulses can be marked. Thus, the original waveforms can be reshaped into 16×16 matrices that are much smaller. We then feed the feature points to the U-net model and some recovery results are shown in Fig. S5(b), which shows that the recovery with these feature points will only slightly reduce the image quality. This indicate that by selecting some feature points, most of the image information can be reconstructed.

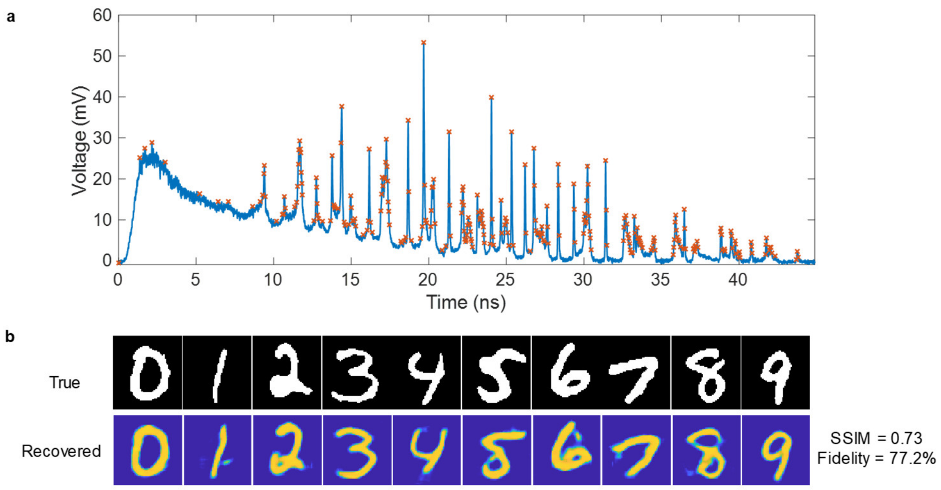

Fig. S5 (a) One recorded waveform and the red-cross-marked 256 feature points are located near the peaks in the waveform. (b) Some recovered images when using the feature points data.

Supplementary Note 6: Individual Illumination

Like most of the previous researches, we try to use an individual source to illuminate the images and an objective to couple light into the fiber probe as shown in Fig. S6. The source is collimated by a lens and illuminates the DMD where the images with size of around 5×5 mm^2 are displayed. The reflected light is coupled into the 1-km MMF via a 40× objective. The rest parts of the system are totally same with that presented in the main text. Compared with the all-fiber system, this individual-illumination system has the advantages of providing a brighter illumination and is suitable for detecting relatively large objects due to the use of an objective. After the same training and testing, the recovery results show a fidelity of 83.0% and SSIM of 0.76. The comparison of the recovery performance of the two different systems are shown in Fig. S7(a), which indicates that the two different illumination methods have similar performance and proves the high adaptability of our proposed method. The recovery of some other types of images of this system is also tested and shown in Fig. S7(b). Besides, we also test its classification performance and the confuse matrices are shown in Fig. S8.

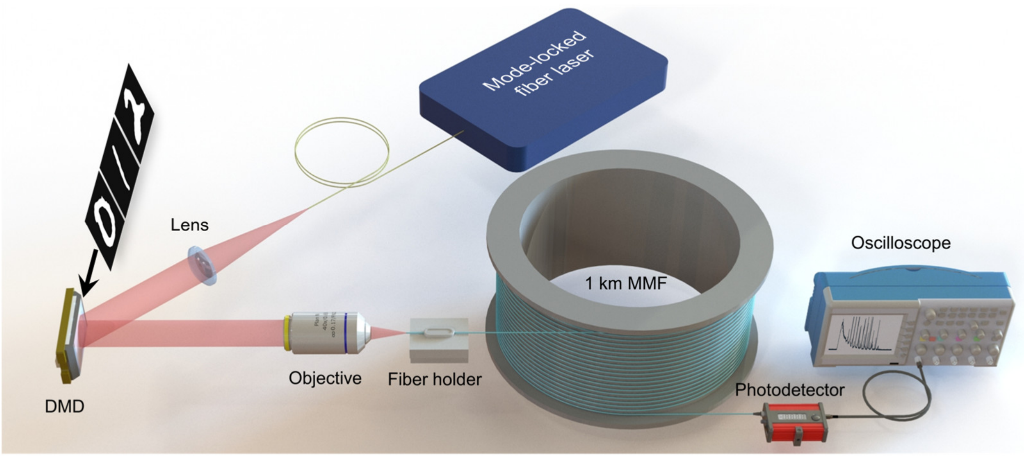

Fig. S6 Schematic of the experiment setup with an individual illumination.

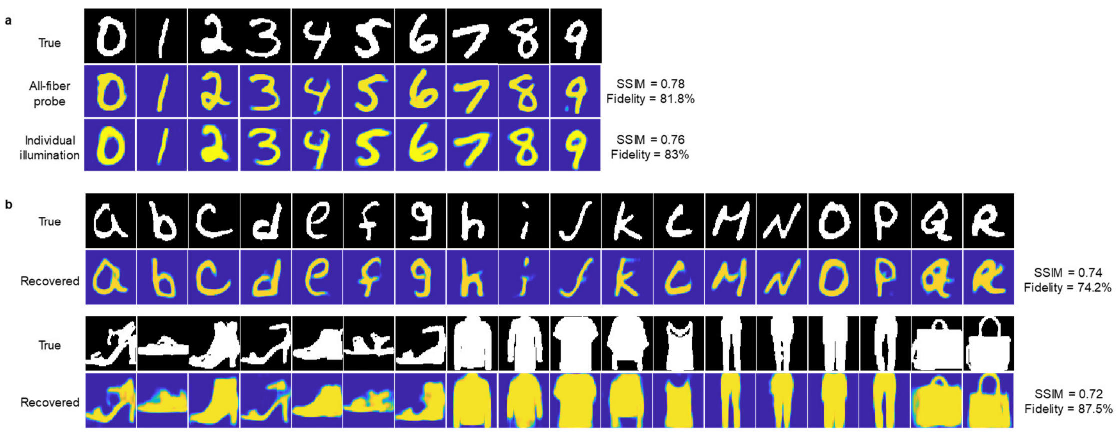

Fig. S7 (a) Comparison of recovery performance between the system shown in the main text and the system with an individual illumination. (b) Some example images of letters and clothes and the recovered results.

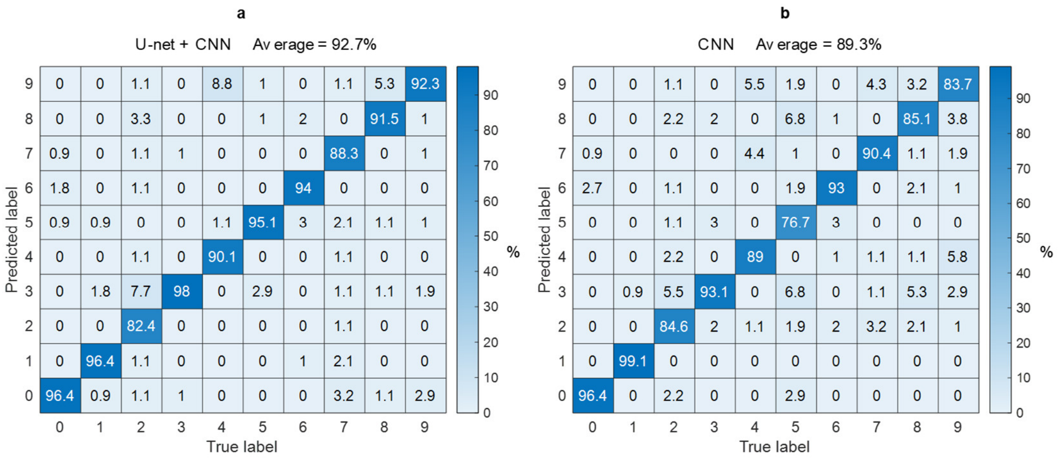

Fig. S8 Confusion matrixes for the system that uses the individual illumination setup. (a) The result of using the combination of U-net and CNN networks. (b) The result of using simply the CNN network. The average accuracy for 1000 test images of digits are shown in the top.

Supplementary References

1 Agrawal, G. P. *Nonlinear Science at the Dawn of the 21st Century Ch. 3*, (Springer, Berlin, 2000).

2 Kruglov, V., Peacock, A., Harvey, J. D. & Dudley, J. M. Self-similar propagation of parabolic pulses in normal-dispersion fiber amplifiers. *JOSA B* **19**, 461-469 (2002).

3 Borhani, N., Kakkava, E., Moser, C. & Psaltis, D. Learning to see through multimode fibers. *Optica* **5**, 960-966 (2018).