

# Scan-less speckle encoded single pixel imaging over million frames per second assisted by dual optical frequency combs

Yangyang Wan<sup>1,†</sup>, Ziwen Long<sup>1,†</sup>, Xinyu Fan<sup>1,\*</sup>, Zuyuan He<sup>1</sup>

<sup>1</sup>State Key Laboratory of Advanced Optical Communication Systems and Networks, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China.

<sup>†</sup>These authors contributed equally to this work.

**Abstract.** Single-pixel imaging (SPI) has emerged as a powerful imaging technique that reconstructs two-dimensional images from one-dimensional (1D) signals detected by a single detector. Most SPI systems are constrained by a scanning paradigm to acquire the 1D signal, which limits the imaging frame rate to the modulation speed of spatial light modulators or tunable light sources. To address the challenge of low imaging rates, we propose a scan-less speckle encoded SPI (SSE-SPI) approach that leverages the advantages of speckle encoding and dual optical frequency combs (DOFCs). The spatial speckle encoding mask, generated from a disordered structure, facilitates high compression rates, enabling rapid imaging. Additionally, DOFCs can deploy spatial speckle encoding masks in parallel, thereby overcoming the limitations associated with serial scanning. By eliminating the scanning mechanism, we demonstrate an offline imaging rate of up to 20 MHz and a substantial spatial-temporal information flux of 15.68 GPixels/s through the application of a specially designed artificial neural network for image reconstruction. This SSE-SPI scheme holds promise for ultrafast imaging with a single detector and opens new possibilities for capturing transient processes in the fields of material science and life sciences.

**Keywords:** optical frequency comb, single pixel imaging, speckle, neural network.

\*Xinyu Fan, [fan.xinyu@sjtu.edu.cn](mailto:fan.xinyu@sjtu.edu.cn)

## 1 Introduction

Single-pixel imaging (SPI) is a promising imaging methodology that requires only one detector to replace two-dimensional (2D) sensors such as charge-coupled devices (CCDs) and complementary metal-oxide-semiconductor (CMOS) sensors.<sup>1</sup> By circumventing the constraints imposed by 2D sensors, SPI, when coupled with a high-performance detector, can offer notable advantages in terms of cost, detection efficiency, and response time.<sup>2</sup> This approach has been successfully demonstrated in various scenarios including infrared imaging,<sup>3,4</sup> terahertz imaging,<sup>5-7</sup> ultrafast imaging,<sup>8</sup> and long-range imaging.<sup>9-11</sup> In current SPI implementations, configurable spatial light modulators (SLMs) play a pivotal role in generating a sequence of pattern encoding masks to encode the illumination.<sup>12,13</sup> The signals corresponding to each pattern are captured by a single detector and demodulated by digital reconstruction algorithms to reconstruct the image.<sup>14</sup> While the use of a single detector theoretically facilitates ultrafast imaging in SPI, practical limitations emerge due to the refresh rate of the spatial light modulator (SLM). The digital micromirror device (DMD), which is the most commonly employed SLM in SPI, typically operates at a refresh rate of up to 3.2 MHz.<sup>15</sup> However, this refresh rate remains insufficient for applications that require high-speed imaging.

The imaging rate of SPI can be improved in two primary aspects. The first aspect involves the utilization of effective image reconstruction algorithms, such as differential ghost imaging (DGI),<sup>16</sup> compressed sensing (CS),<sup>17-20</sup> and deep neural networks (DNN).<sup>21-25</sup> These algorithms aim to

reduce the number of samples required for image reconstruction. While these reconstruction techniques can accelerate the image reconstruction process, the overall imaging rate remains constrained by the scanning mechanism employed.

The second aspect focuses on the advancement of rapid modulation hardware or mechanisms. In the context of terahertz and millimeter-wave SPI, fast rotating disks with etched coded patterns are utilized to achieve spatial modulation of illumination.<sup>26–28</sup> For instance, cyclic patterns have been proposed, achieving a spatial modulation rate of 2.4 MHz and an imaging rate of 72 frames per second (fps).<sup>29</sup> Additionally, an annular binary mask has been introduced to facilitate fast spatial modulation, realizing 100 fps with a frame size of  $32 \times 32$  pixels.<sup>30</sup> Despite the advantages of rapidly rotating disks for spatial modulation, their complex and inflexible structures present challenges compared to reconfigurable digital micromirror devices (DMDs). Furthermore, manufacturing larger coding patterns on a fixed-size disk is often problematic. An alternative method involves using arrays of light-emitting diodes (LEDs) to provide structured illumination patterns.<sup>31</sup> Due to their fast switching times, this method has demonstrated 2D imaging capabilities with a pixel size of  $32 \times 32$  at a frame rate of 2.5 Mfps.<sup>32</sup> However, the use of LED arrays is limited by the range of light sources available and poses difficulties in achieving high image resolution due to inherent limitations in focusing spatially low-coherent light. To overcome these challenges, a method has been proposed that involves rapidly deploying encoding masks in sub-regions of the DMD by scanning a laser.<sup>33</sup> With the aid of a developed compressed sensing reconstruction algorithm, this approach has achieved offline imaging at 1.2 Mfps and real-time video at 100 fps. Despite the introduction of various rapid spatial modulation methods, the necessity for scanning or switching coding patterns persists, presenting a significant obstacle to achieving imaging rates exceeding tens of Mfps in SPI.

In recent years, light scattering in disordered structures has transitioned from being perceived as a detrimental phenomenon to a valuable resource for various applications, including random projection,<sup>34</sup> reservoir computing,<sup>35,36</sup> spectrometry,<sup>37–39</sup> polarimetry,<sup>40</sup> radio frequency reception,<sup>41–43</sup> and hyperspectral imaging.<sup>44,45</sup> This shift in perspective highlights the potential of disordered media in advancing optical technologies. Given that a speckle pattern contains both spatial and spectral information, scan-less hyperspectral imaging can be achieved by capturing the speckle with a CCD camera and employing appropriate reconstruction algorithms.<sup>46</sup> Furthermore, the spatial pattern characteristics of a speckle, which vary with spectral content, can be utilized to create spatial encoding masks in SPI. The integration of a multimode fiber (MMF) with a tunable laser facilitates random speckle projection, which is instrumental in SPI applications.<sup>47,48</sup> However, the limited wavelength switching time of tunable lasers constrains the realization of ultrafast imaging speeds. To overcome this limitation, employing a phase modulator to modulate the incident laser enables random speckle projection at modulation rates exceeding GHz at the output of an MMF, which can be utilized for image-free classification.<sup>49</sup> Building on GHz-rate random speckle projection, image reconstruction is accomplished by combining a photonic reservoir chip with multiple photodetectors (PDs) and a digital neural network.<sup>50</sup> Despite these advancements, achieving effective image reconstruction at GHz-rate speckle encoding remains challenging when using a single PD. The aforementioned methods still rely on the serial scanning paradigm for deploying spatial pattern encoding masks, necessitating PDs with bandwidths exceeding 10 GHz and high-speed digital oscilloscopes. This requirement introduces complexity and cost, posing challenges for practical implementation in many scenarios.

Optical frequency comb (OFC) serves as a potential tool to address these issues. It is composed

of a series of longitude modes with equal frequency intervals, acting as a frequency ruler, which is widely utilized for spectroscopic measurement owing to its high-spectral resolution and broad spectral coverage.<sup>51,52</sup> By adopting the interference between two OFCs with slight different repetition rates, dual optical frequency comb (DOFC) interferometry enables simultaneous measurement of all comb lines without scanning mechanism for various applications of OFCs, such as spectra measurement,<sup>53</sup> distance measurement<sup>54</sup> and imaging.<sup>55</sup> In a notable development of DOFC SPI systems,<sup>56,57</sup> an one-to-one correspondence between the comb lines and the spatial positions was established by grating dispersion, and the images were reconstructed from the obtained interferograms. However, the one-to-one relationship between OFC modes and a 2D array of focal spots requires precise control of the optical system. Moreover, current DOFC SPI methods can only achieve a maximum imaging rate of kHz level due to the low repetition rate of the used OFCs.<sup>57</sup>

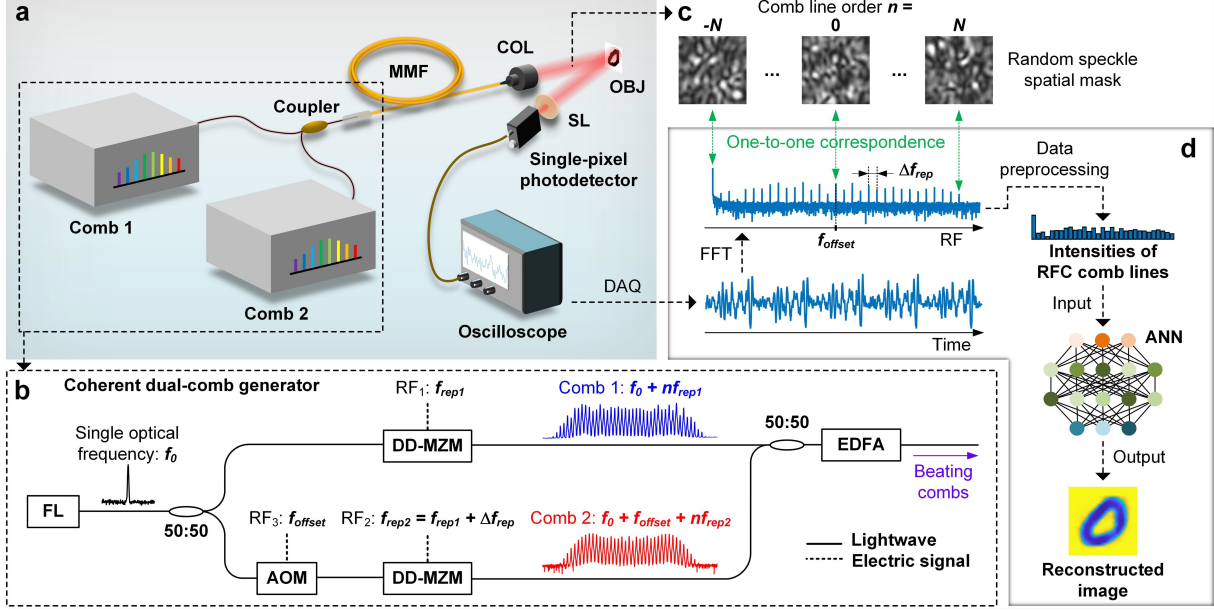
In this paper, we propose an ultrafast scan-less speckle encoded SPI (SSE-SPI) system that leverages the advantages of speckle encoding and DOFC detection. The two OFCs are employed by traversing an MMF, which serves as the disordered structure, to deploy spatial speckle encoding masks in parallel, thus eliminating the need for a SLM. This SPI framework presents a novel scan-less approach that overcomes the limitations of imaging speed imposed by scanning devices in traditional scanning paradigms, thereby rendering it suitable for ultrafast imaging applications. Experimental results demonstrate offline 2D imaging at speeds of up to 20 Mfps, with the potential for further enhancement to 100 Mfps. Notably, the system achieves ultrafast imaging using a PD that requires a bandwidth of only several hundred megahertz. In the absence of iterative optimization processes, we designed an artificial neural network (ANN) for image reconstruction, enabling real-time video capture at 10 kfps.

## 2 Results

### 2.1 System setup

The schematic of the proposed SSE-SPI system is illustrated in Fig. 1a. DOFCs with slightly different repetition rates were employed as the light source. In consideration of the flexibility and adjustability of the repetition rate, electro-optic frequency comb (EOFC) generation configuration was utilized in the system, as depicted in Fig. 1b. A continuous-wave (CW) fiber laser serves as the seed laser of the two OFCs. By controlling the radio frequency (RF) signals applied to the dual-drive Mach-Zehnder modulators (DD-MZMs), the repetition frequencies of the two OFCs can be selected arbitrarily. In our system setup, the repetition frequencies of the generated DOFCs were configured to  $f_{\text{rep1}} = 18$  GHz and  $f_{\text{rep2}} = f_{\text{rep1}} - \Delta f_{\text{rep}} = 17.98$  GHz, respectively. The total number of comb lines was 31. Additionally, an acousto-optic modulator (AOM) is employed to introduce a frequency offset between the two OFCs, distinguishing the positive and negative sidebands. The detailed spectra of the generated two OFCs are shown in Supplementary Note 1.

The two paths of OFCs are combined by a coupler and subsequently amplified by an erbium-doped fiber amplifier before traversing through a disordered structure, specifically an MMF. The spatial speckles are generated by the interference between the spatial transmission modes within the fiber. The MMF utilized in the proposed system functions as a wavelength-to-space converter. By varying the wavelength of the incident lightwave, distinct output spatial patterns can be achieved (See details in Methods). Since the OFC comb lines of different orders possess different wavelengths, speckle projections with different patterns are generated in parallel after passing



**Fig 1 Schematic of the proposed scan-less ultrafast SPI system.** **a** Experimental setup of the proposed system. MMF, multimode fiber; COL, collimator; OBJ, object; SL, spherical lens; DAQ, data acquisition. **b** Experimental setup for the generation of dual electro-optic frequency combs. FL, fiber laser; AOM, acousto-optic modulator; DD-MZM, dual-drive Mach-Zehnder modulator; RF, radio frequency; EDFA, erbium-doped fiber amplifier. **c** Spatial masks generated from MMF at different orders of comb lines. **d** Flow of the image reconstruction process. FFT, fast Fourier transform; RFC, RF comb; ANN, artificial neural network. The initial signal is a 1D time-domain interferogram. After FFT of the interferogram, the intensity of each RFC comb line can be extracted and used as input for the ANN to reconstruct the image.

through the MMF, as illustrated in Fig. 1c. These speckle projections serve as spatial encoding masks for SPI, enabling the spatial sampling of the object.

The central  $170 \times 170$  pixels of a DMD was used as the imaging object by displaying programmable reflectivity patterns. The pitch size of the DMD is  $10.8 \mu\text{m}$ , and the imaging area has a size of  $1.84 \text{ mm} \times 1.84 \text{ mm}$ . The light reflected back from the object is collected by a spherical lens, and a single PD is used to detect the interference light between the DOFCs, generating a electric signal which is then digitized by an oscilloscope. Subsequently, a short temporal span of data was acquired for processing. As shown in Fig. 1d, after applying a Fourier-transform (FT) to the obtained electric signal, a RF comb (RFC) can be obtained. There is an one-to-one correspondence between each RFC comb line and each OFC comb line. The RF interval between two adjacent RFC comb lines equals to  $\Delta f_{\text{rep}}$ . According to the principle of FT, the temporal span should be longer than  $1/\Delta f_{\text{rep}}$  to distinguish each RFC line,<sup>53</sup> thus limiting the imaging rate to  $\Delta f_{\text{rep}}$ .

Through the process of spatial sampling, the spatial reflectivity information of the imaging object is encoded into the intensities of the RFC comb lines, which can then be employed for image reconstruction. The physical model of the spatial sampling process can be expressed as:

$$\begin{aligned}
\mathbf{I} &= [I_{-N} \quad \cdots \quad I_0 \quad \cdots \quad I_N]_{(2N+1) \times 1}^T \\
&= 2G \cdot \begin{bmatrix} A_{1,-N}A_{2,-N} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_{1,N}A_{2,N} \end{bmatrix}_{(2N+1) \times (2N+1)} \cdot \begin{bmatrix} \mathbf{A}_{\text{MMF},N} \\ \vdots \\ \mathbf{A}_{\text{MMF},N} \end{bmatrix}_{(2N+1) \times XY} \cdot \mathbf{R} \quad (1) \\
&= \mathbf{H} \cdot \mathbf{R}
\end{aligned}$$

Here,  $\mathbf{I}$  is the matrix composed of the intensity of each RFC comb line,  $\mathbf{R}$  is a matrix with a dimension of  $XY \times 1$ , representing the flattened reflectivity function of an object with a pixel number of  $X \times Y$ . With  $n = -N, \dots, 0, \dots, N$ ,  $A_{1,n}A_{2,n}$  represents the amplitude-product of the  $n$ th-order lines of the two generated OFCs,  $\mathbf{A}_{\text{MMF},n}$  is a matrix with a dimension of  $1 \times XY$ , representing the  $n$ th-order spatial speckle encoding mask.  $G$  is the conversion gain. Detailed derivation process is presented in Supplementary Note 2.

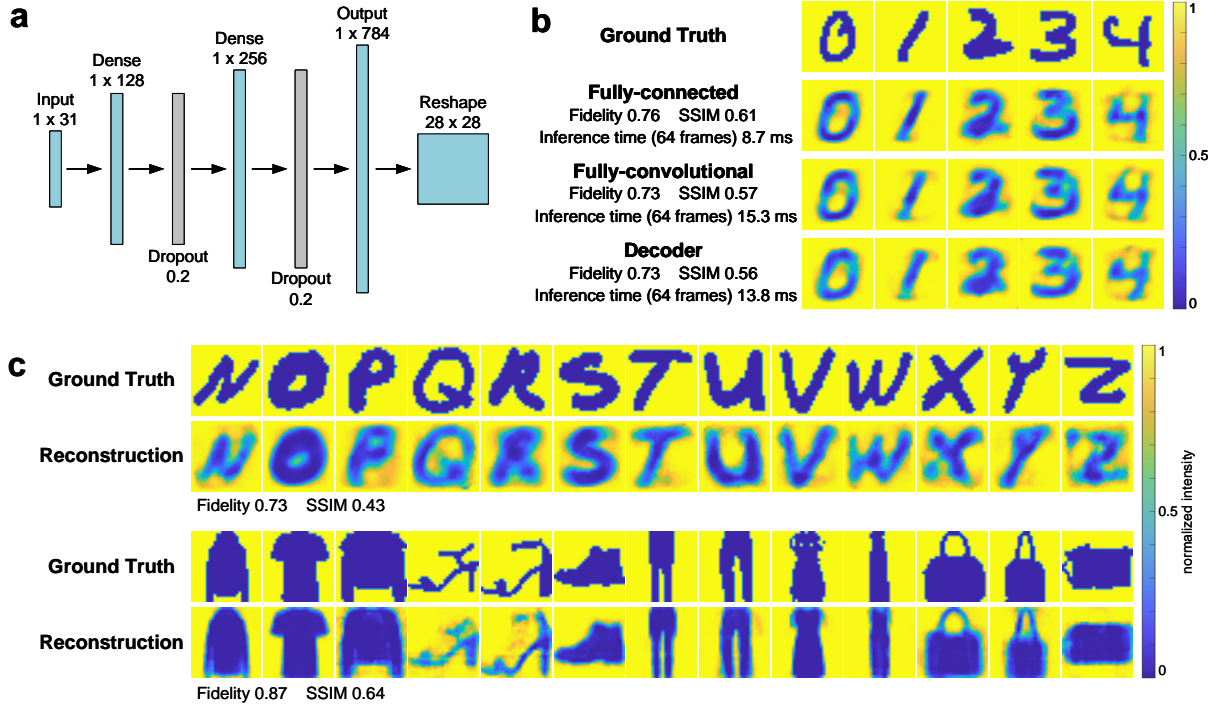
As shown in equation 1, the spatial sampling process of the system can be modeled as a matrix transformation of  $\mathbf{I} = \mathbf{H} \cdot \mathbf{R}$ , where  $\mathbf{H}$  represents the transmission matrix. It is known that, for an efficient information transfer from  $\mathbf{R}$  to  $\mathbf{I}$ , the un-correlation between the row-vectors of  $\mathbf{H}$  is required. This is realized by making the wavelength spacing of OFC comb lines larger than  $\lambda_{\text{decorr}}$ , where  $\lambda_{\text{decorr}} = 20$  pm, representing the minimum wavelength shift needed to de-correlate the output spatial mode of the MMF (See more details in Supplementary Note 3).

Spatial sampling is the process of encoding the information from  $\mathbf{R}$  to  $\mathbf{I}$ , while image reconstruction is the process of decoding the information encoded in  $\mathbf{I}$  to obtain an estimation of  $\mathbf{R}$ . In achieving this, a commonly used approach is the transmission matrix method. Specifically, after the establishment of  $\mathbf{H}$ , the estimation of  $\mathbf{R}$  is obtained through the matrix transformation equation of  $\hat{\mathbf{R}} = \mathbf{H}^{-1} \cdot \mathbf{I}$ , where  $\mathbf{H}^{-1}$  represents the pseudo-inverse matrix of  $\mathbf{H}$ . Experimental results indicated that this approach is applicable when the sampling ratio is high, however, it failed to achieve high-quality image reconstruction when the sampling ratio is low (See more details in Supplementary Note 2). Here, the sampling ratio is defined as  $(2N + 1)/(XY)$ . In recent years, neural networks have shown great potential in solving ill-posed problems. Therefore, the system employs an ANN for high-quality image reconstruction at low sampling ratio.

## 2.2 Image recovery

An ANN is constructed to recover images from intensities of the RFC comb lines, and the structure of the ANN is shown in Fig. 2a. The input and output of the ANN are the intensities of 31 RFC comb lines and the reconstructed images with a size of  $28 \times 28$  pixels, respectively. Therefore, the sampling ratio of the proposed system is  $31/(28 \times 28) = 3.95\%$ . The neural network comprises three fully connected layers. It was trained using 7200 images of handwritten digits from the MNIST dataset to learn the mapping relationship between the comb lines and the images. Other network structures such as convolutional neural network and decoder network were also trained with the same training dataset for comparison (See more details of networks in Supplementary Note 4). Fidelity and structural similarity<sup>58</sup> (SSIM) are used to evaluate image reconstruction performance (See details in Methods). The fully-connected network, convolutional network, and decoder network showed average fidelity of 0.76, 0.73 and 0.73, along with average SSIM of 0.61,



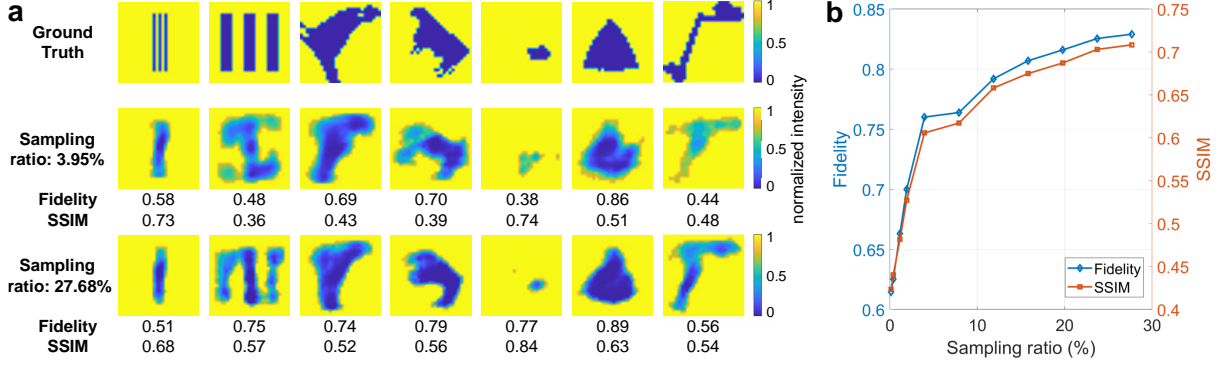


**Fig 2 Imaging results of neural networks.** **a** The structure of the proposed fully-connected network. **b** Comparison of imaging results from different neural networks. **c** Reconstructed images by fully-connected network on EMNIST and Fashion-MNIST dataset, respectively.

0.57 and 0.56 for 800 test images, respectively. Examples of reconstructed images are shown in Fig. 2b, fully-connected neural network has faster process speed and better image reconstruction performance. After similar training processes, reconstructed images of different types, such as handwritten letters from EMNIST dataset and fashion products from Fashion-MNIST dataset, are shown in Fig. 2c, along with the average fidelities and SSIMs. The above results demonstrate that neural network algorithms can achieve image reconstruction even with a sampling ratio of 3.95%. Moreover, the stability of the system in open environment was verified through a long-term measurement experiment (See more details in Supplementary Note 5).

In the above imaging experiments, the test dataset and the training dataset for the neural network belong to the same type. Further validation is conducted on the generalization performance of the ANN. In order to avoid overfitting to a specific type of images, the training dataset comprises 7000 images selected from the MNIST, EMNIST, and Fashion-MNIST datasets. The images of these three datasets basically have the pattern characteristics of different types of images, such as dots, straight lines, curves, bulk areas, and more. After completing the training process, the trained network is used to reconstruct different types of test images that were not included in the training dataset, and the results are shown in Fig. 3a. The results indicate that the ANN has the generalization ability to reconstruct various types of images with high quality.

In the proposed system, the number of comb lines determines the sampling ratio (see details in Methods). Improving the sampling ratio to achieve better image reconstruction performance is an effective way. In the existing experimental setup, EOFC produces only 31 comb lines limited by the modulation efficiency of the electro-optical modulator. Higher sampling ratio can be realized by tuning the wavelength of the CW fiber laser and recording the corresponding intensities of RFC



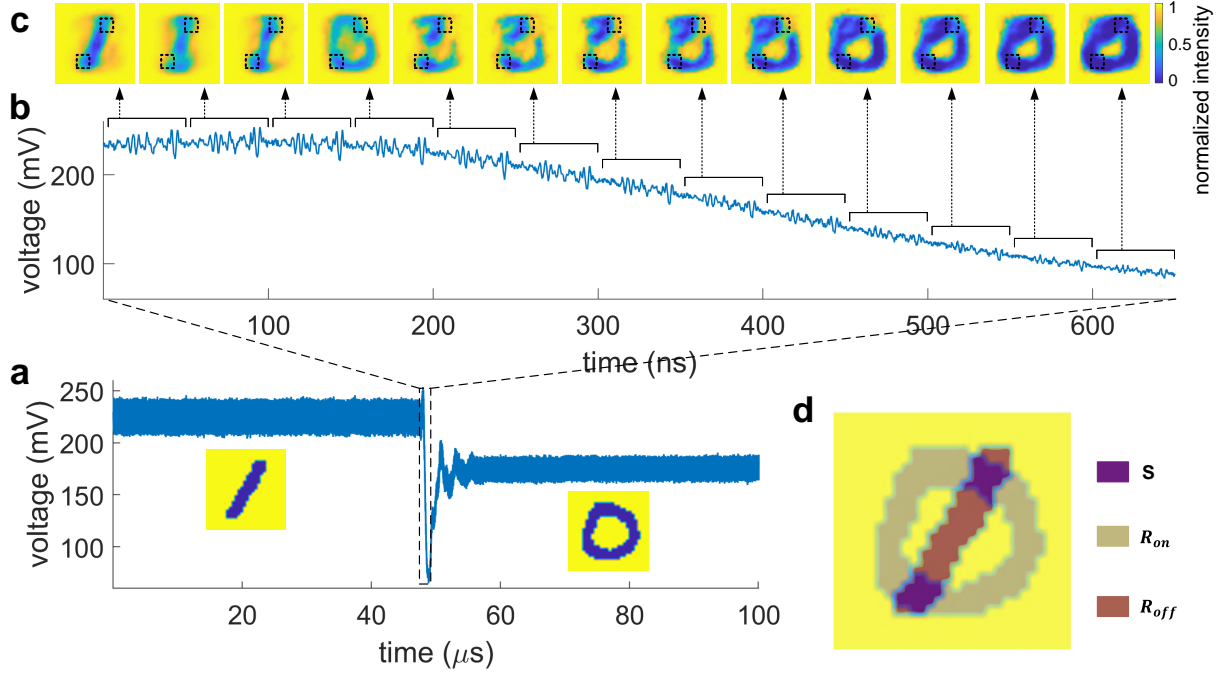
**Fig 3 Imaging results of the object with different types.** **a** Examples of reconstructed images of different types which are not included in the training dataset. **b** Relationship between the image reconstruction performance and the sampling ratio.

comb lines. The wavelength tuning step was  $\lambda_{\text{decorr}}$ . After 7 rounds of wavelength tuning by thermal tuning mode, intensities of 217 RFC comb lines were obtained for reconstruction for  $28 \times 28$  images, which corresponds to a sampling ratio of 27.68%. As shown in Fig. 3b, with the increase in the sampling ratio, the improvement of fidelity and SSIM goes through a process from rapid to slow (More reconstructed images at different sampling ratio are presented in Supplementary Note 6). In addition to sampling ratio, the signal-to-noise ratio (SNR) of the detected signals and the size of speckle grains also have an impact on the image reconstruction performance. Since the system noise are mainly additive Gaussian white noise introduced by the PD, SNR can be improved by averaging multiple detected signals to acquire reconstructed images with higher quality (See noise analysis in Supplementary Note 7).

### 2.3 Ultrafast imaging

The offline imaging frame rate of the proposed system is equal to the repetition frequency difference of the two OFCs ( $\Delta f_{\text{rep}}$ ). An experiment was performed to verify the ultrafast imaging of the system. In the experiment, the ultrafast imaging ability of the proposed system was demonstrated by capturing the transient image switching process of DMD, as shown in Fig. 4. The display of DMD was set to switch from the image of digit "1" to digit "0". The received electric signal is shown in Fig. 4a and Fig. 4b, which indicates that the DMD took approximately 600 ns to complete the image switching. During the display switching process from the image of digit "1" to digit "0", the micromirrors on the DMD can be divided into stationary and rotating states, as shown in Fig. 4d. The region  $S$  in Fig. 4d represents the overlapping part of digit "1" and digit "0", and the corresponding micromirrors remained the "on" state. The micromirrors in the region  $R_{\text{on}}$  and region  $R_{\text{off}}$ , which belong to the rotating state, were switched between the "on" state and the "off" state. The reconstructed images with a temporal interval of 50 ns are shown in Fig. 4c, exhibiting a well match with the state changing process of regions  $S$ ,  $R_{\text{on}}$  and  $R_{\text{off}}$ .

In offline imaging, higher speed can be realized by enlarging the repetition frequency difference of the two OFCs. By adjusting the driving RF signals applied to the EOFCs, the imaging rate of the proposed scan-less SPI can reach up to 100 MHz (See details in Supplementary Note 8). However, with a limited data acquisition bandwidth, the number of detectable RFC comb lines is constrained by the larger RF interval, which results in further reduction of sampling ratio. Moreover, a shorter temporal span of the interference signal is required for higher imaging rate, which



**Fig 4 Ultrafast imaging results.** **a** Received interference signal during the transient switching process from image 1 to 0 on the DMD. **b** Zoom-in figure of the received interference signal during the transient switching process. **c** Reconstructed images during the transient switching process. The dashed box represents the region where DMD remains stationary during the switching process. **d** Schematic of images of digit "1" and digit "0" displayed on the DMD. S: stationary region;  $R_{on}$ : the region in DMD that switched from the "off" state to the "on" state;  $R_{off}$ : the region in DMD that switched from the "on" state to the "off" state.

leads to a decrease of SNR. Image reconstruction performance needs to be improved at higher imaging rate.

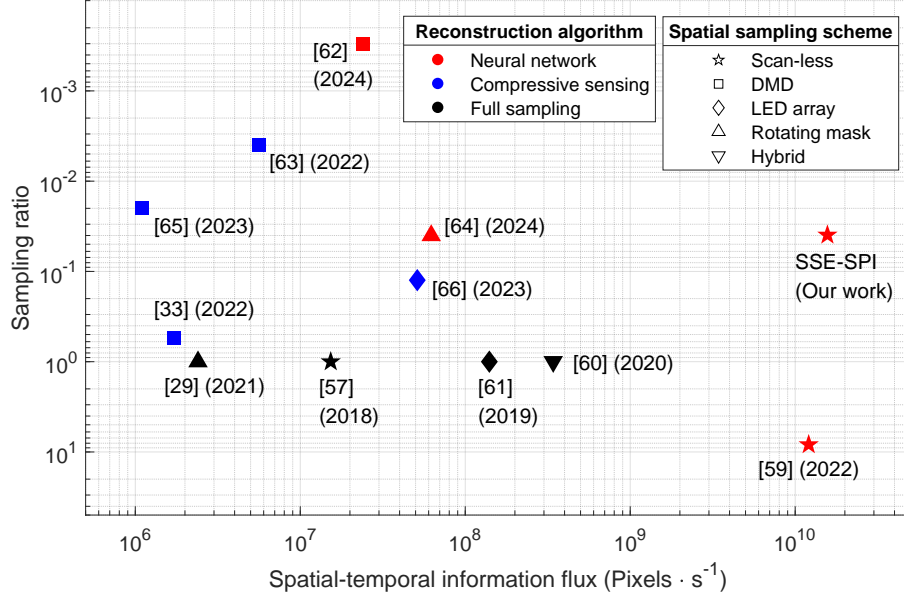
In real-time imaging, the frame rate is mainly constrained by the time consumption for data acquisition, data reading, and image reconstruction. Due to the rapid inference capability of the ANN, the system is capable of generating a real-time video at 3000 fps (See details in Supplementary Note 9 and 10).

### 3 Discussion

It is instructive to compare the proposed SSE-SPI with existing various high-speed SPI implementations,<sup>29,33,57,59–66</sup> as shown in Fig. 5. Two performance indexes of spatial-temporal information flux (STIF) and sampling ratio are compared. STIF, which is the product of the total number of imaging pixels and the imaging rate, demonstrates the total amount of imaging information provided by the system in per unit time. Sampling ratio demonstrates the information compression ability. System with lower sampling ratio can acquire the large amount of information with less demanding data acquisition equipment and less data. The STIF and sampling ratio of the proposed system are 15.68 GPixels/s and 3.95%, respectively.

The performance of the proposed system can be further improved from three primary aspects: OFC source, disordered structure, and image reconstruction algorithm.





**Fig 5 Performance metric chart of various SPI methods using sampling ratio and spatial-temporal information flux.** SSE-SPI, scan-less speckle encoded single-pixel imaging.

Given that the sampling ratio of the system depends on the number of OFC comb lines, the bandwidth of EOFC can be expanded through nonlinear effects to increase sampling ratio.<sup>67</sup> In addition, other types of OFC such as fiber mode-locked lasers and Kerr frequency combs can also be used as light source in this scheme.<sup>51</sup> Compared to EOFC, other types of OFCs typically possess more orders of comb lines, which is beneficial for a high sampling ratio. In the proposed scheme, higher sampling ratio achieves better image reconstruction performance without sacrificing imaging rate, but at the expense of a detector with large bandwidth. After two decades development, OFC can be generated over broad spectral regions including extreme ultraviolet, visible, infrared and terahertz domains.<sup>51</sup> Therefore, using OFCs as the light source for SPI can still maintain the advantage of general compatibility in various spectral regions. It is worth noting that the rapidly developing on-chip OFC in recent years may be an ideal light source for integrated ultrafast scan-less SPI.<sup>68</sup>

The disordered structure is one of the key devices in the proposed system for the deployment of spatial encoding masks via speckles. Here, an MMF is chosen as the disordered structure to generate speckles. The features of MMF speckles, such as the distribution and size of speckle grains, have an influence on the imaging quality and resolution of SPI. For better imaging reconstruction performance, the speckle grain size should closely match the object features (See details in Supplementary Note 11). In the future, wavefront shaping techniques can be used to customize the speckle,<sup>69</sup> aiming to attain the optimal spatial mask. Moreover, in order to generate more uncorrelated speckle masks simultaneously within limited bandwidth of OFC, an MMF with longer lengths should be adopted. Therefore, it is necessary to comprehensively consider the type, length and mode excitation method of MMF to ensure the image reconstruction performance of SPI. Other kinds of disordered structures can also be used to replace MMF, such as integrating spheres, metasurfaces, etc. Diverse disordered structures can bring different imaging effects to SPI in distinct spectral regions. Especially, some disordered structures, such as metasurfaces and silicon dioxide phase plate, can be designed and optimized to generate customized speckle masks for a better

imaging quality of SPI. System stability is a common problem for systems using disordered structures.<sup>38</sup> Changes of environmental temperature are the main factor for system instability. Further, active temperature control technology and passive methods using temperature insensitive materials can be adopted for the disordered structure to ensure long-term stability of the system.

Machine learning has demonstrated superior image reconstruction results in SPI due to its advantages in solving ill-posed inverse problems. In this scheme, a shallow ANN shows the ability to rapidly reconstruct images at a low sampling ratio. The training dataset used here is relatively small, and further experiments and generator networks can be used to obtain more training data with different noise and system settings in various environments for improving the robustness of image reconstruction. A deep learning network combined with physical-model informed method may be a solution to achieve higher quality of reconstructed images in noisy environments.

In conclusion, we have developed a novel mechanism for SPI by utilizing the advantages of DOFC and speckle encoding. In the proposed SSE-SPI system, an imaging rate of up to 20 MHz is realized, which is the fastest imaging rate of SPI to the best of our knowledge. Due to the fast image reconstruction process of the designed shallow ANN, real-time video with a frame rate reaching 3000 is demonstrated. The ANN has been verified to reconstruct images different from the types in the training dataset, demonstrating a certain degree of generalization. The long-term imaging test results show that this system has certain stability for practical applications. By further adopting a more advanced DOFC, a specific disordered structure and a deep learning network, SPI with faster and better image reconstruction quality in various spectral regions may be achieved. We believe that the proposed scan-less technology not only provides a novel implementation paradigm for SPI, but also brings a powerful solution for high-speed imaging.

## 4 Methods

### Experiments

Generation of dual electro-optic frequency combs (EOFC): A fiber laser (NKT Adjustik E15) is used to generate center frequency for EOFC with an output power of 16 dBm. The center frequency light is divided into two paths and passes through DD-MZMs (Fujitsu FTM7937) to generate dual-EOFC. The radio frequency (RF) signals driving DD-MZMs with frequencies of 18 GHz and 17.98 GHz are generated by two benchtop signal sources (R&S SMF 100A, Agilent E8257D). Before applying to DD-MZM, the RF signal needs to be amplified by electrical amplifiers (Mini-Circuits ZVE-3W-183+). In order to distinguish the positive and negative sidebands of dual-EOFC, an AOM (Gooch Housego T-M300-0.1C16J-3-F2P) is used to shift the center frequency of one EOFC. The 300 MHz RF signal with a power of 4 dBm driving AOM is generated by an arbitrary waveform generator (Siglent SDG6052X-E) and amplified by an electrical amplifier (R&K A3000-2H-R). The power of the beating combs was then amplified to 16 dBm by an EDFA (Discovery Optica EDFA-BA-23-FA-B).

Disordered structure: The light is transmitted from a single-mode fiber to the multimode fiber (Fibestar IR105/125-AC) which was spliced to the single mode fiber. In practical operation, we did not deliberately misalign the two fibers. Therefore, only the low-order mode of the MMF is excited at the splicing point. The light of the low-order mode gradually diffused to the higher-order mode with transmission.

Imaging part: The imaging targets were changed by loading different binary images on the DMD (Texas Instruments DLP550HE). The pitch size of the DMD is 10.8  $\mu\text{m}$ , the images were

loaded on the central  $170 \times 170$  pixels in the DMD. The photodetector (Thorlabs DET08C/M) has a response bandwidth of 5 GHz. The oscilloscope (Keysight DSOS204A) has a 2.1 GHz analog bandwidth and a sample rate of 10 G/s. During the process of acquiring training data for the neural network, the switching rate of the DMD was set at 30 Hz. Due to the non-uniform image switching interval of the DMD, a stepped trigger signal is necessary after loading an image onto the DMD. This process was achieved by inserting an all-black image between every two useful images. As a result, the acquisition rate for useful images is 15 Hz, and the entire data collection procedure took about 10 minutes to complete.

### Speckle obtained from an MMF

An MMF is used as the disordered structure to generate speckle for spatial encoding in SPI. When a lightwave with fixed wavelength is input into a MMF, many guided modes are excited. The speckle is produced by interference between the guided modes at the output end of the MMF. The lightwave field at the output end can be expressed as:

$$E(r, \theta, \lambda, L) = \sum_m A_m \phi_m(r, \theta, \lambda) \exp[-j(\beta_m \lambda L - \frac{2\pi}{\lambda} t + \varphi_m)] \quad (2)$$

where  $A_m$ ,  $\phi_m$ ,  $\beta_m$  and  $\varphi_m$  are the initial amplitude, spatial profile, propagation constant and initial phase of the  $m^{th}$  guide mode, respectively. The number of guided modes supported by a MMF is determined by the refractive index distribution, core diameter, and numerical aperture of the fiber. The number of excited guided modes depends on the coupling method between the incident lightwave and the MMF. Besides excited guided modes, the speckle pattern is also related to the length of the fiber  $L$  and the wavelength of the incident lightwave  $\lambda$  according to the formula.

In the proposed system, multiple speckles at different wavelengths are generated simultaneously. The cross-correlation coefficient can be used to quantitatively describe the similarity between two speckles:

$$C(\Delta\lambda) = \left\langle \frac{\langle I(\lambda, x) I(\lambda + \Delta\lambda, x) \rangle}{\langle I(\lambda, x) \rangle \langle I(\lambda + \Delta\lambda, x) \rangle} - 1 \right\rangle_x \quad (3)$$

where  $\Delta\lambda$  is the difference between wavelengths,  $I(\lambda, x)$  is the intensity at a position  $x$  for input wavelength  $\lambda$ ,  $\langle \dots \rangle$  represents the average over wavelength  $\lambda$  and  $\langle \dots \rangle_x$  represents the average over position  $x$ .

### Image reconstruction algorithms

The fully connected neural network had one input layer, one output layer, and two hidden layers. The network was sequentially structured, with neuron numbers of the 4 layers from input to output being  $M$ , 128, 256 and 784, respectively. Here,  $M$  equals to the number of the combs used for imaging. To prevent over-fitting, dropout layers were added behind each hidden layer so that 20% of the neurons did not participate in training for each iteration. The activation function of the hidden layers was leaky-relu and the activation function of the output layer was sigmoid. Mean square error (MSE) is used as loss function. 30% of the training data was used as the validation set. When validation loss had not declined within 5 iterations, the learning rate was reduced by 5%, and when validation loss had not declined within 100 iterations, the training was stopped. The neural-network-related code was written in Python based on Tensorflow. We used a commercially-available laptop (Thinkpad X1 Carbon Gen 10; CPU, Intel Core i7-1280P) for neural network

training, which lasted approximately 5 min. When using the trained network for image restoration, the same laptop is used, while in the experiment of real-time video, a desktop computer equipped with a CPU of Intel Core i9-13900K is used as the computing unit.

### Fidelity and SSIM

Fidelity here refers to the two-dimensional correlation coefficient between the reconstructed image and the true image. The built-in ‘corr2()’ function with default parameters of matlab was used to calculate the fidelity, the expression is:

$$Fidelity(X, Y) = \frac{\sum_m \sum_n (X_{mn} - \mu_X)(Y_{mn} - \mu_Y)}{\sqrt{\left[ \sum_m \sum_n (X_{mn} - \mu_X)^2 \right] \left[ \sum_m \sum_n (Y_{mn} - \mu_Y)^2 \right]}} \quad (4)$$

where  $X$  and  $Y$  refer two-dimensional images,  $\mu_X$  and  $\mu_Y$  represent the mean value of  $X$  and  $Y$ .

The built-in ‘ssim()’ function with default parameters of matlab was used to calculate the structural similarity, the expression is:

$$SSIM(X, Y) = \frac{(2\mu_X\mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)}, \quad (5)$$

$$C_1 = (0.01L)^2, C_2 = (0.03L)^2$$

where  $X$  and  $Y$  refer two-dimensional images,  $\mu_X$  and  $\mu_Y$  represent the mean value of  $X$  and  $Y$ ,  $\sigma_X$  and  $\sigma_Y$  represent the standard deviation of  $X$  and  $Y$ ,  $\sigma_{XY}$  represent the cross-covariance between  $X$  and  $Y$ ,  $L$  represent the dynamic range of the images.

### Relationship between comb lines and sampling ratio

In the proposed system, each comb line produces a pattern encoding mask after passing through a disordered structure and samples the object once. When the image to be recovered has  $M$  pixels and the OFC has  $N$  comb lines, the theoretical sampling ratio of the system is  $N/M$ . The system generates another OFC with a repetition frequency difference of  $\Delta f_{\text{rep}}$  to achieve the detection of signals from each comb line by dual OFC interference technology. Constrained by the analog bandwidth  $B$  of the PD, the maximum sampling ratio is  $\frac{B}{\Delta f_{\text{rep}}}$ .

**Acknowledgements** This work is financially supported by National Natural Science Foundation of China (NSFC) under Grant No. 62275151 and No. 62405178. We thank Prof. Liangcai Cao from Tsinghua University for his valuable suggestions.

### Data, Materials, and Code Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

### Author contributions

Y. Wan conceived the idea, designed the experiments. Z. Long performed the simulations and experiments. Y. Wan and Z. Long analyzed the results and prepared the manuscript. X. Fan, and Z. He reviewed and revised original draft. Y. Wan, Z. Long, X. Fan, and Z. He discussed the manuscript. X. Fan supervised the study.

### Conflict of interest

The authors declare no competing interests.

## References

- 1 M. P. Edgar, G. M. Gibson, and M. J. Padgett, "Principles and prospects for single-pixel imaging," *Nature photonics*, vol. 13, no. 1, pp. 13–20, 2019.
- 2 Y. Wang, K. Huang, J. Fang, M. Yan, E. Wu, and H. Zeng, "Mid-infrared single-pixel imaging at the single-photon level," *Nature Communications*, vol. 14, no. 1, p. 1073, 2023.
- 3 M. P. Edgar, G. M. Gibson, R. W. Bowman, B. Sun, N. Radwell, K. J. Mitchell, S. S. Welsh, and M. J. Padgett, "Simultaneous real-time visible and infrared video with single-pixel detectors," *Scientific reports*, vol. 5, no. 1, p. 10669, 2015.
- 4 Q. Tong, Y. Jiang, H. Wang, and L. Guo, "Image reconstruction of dynamic infrared single-pixel imaging system," *Optics Communications*, vol. 410, pp. 35–39, 2018.
- 5 W. L. Chan, K. Charan, D. Takhar, K. F. Kelly, R. G. Baraniuk, and D. M. Mittleman, "A single-pixel terahertz imaging system based on compressed sensing," *Applied Physics Letters*, vol. 93, no. 12, 2008.
- 6 C. M. Watts, D. Shrekenhamer, J. Montoya, G. Lipworth, J. Hunt, T. Sleasman, S. Krishna, D. R. Smith, and W. J. Padilla, "Terahertz compressive imaging with metamaterial spatial light modulators," *Nature photonics*, vol. 8, no. 8, pp. 605–609, 2014.
- 7 R. I. Stantchev, B. Sun, S. M. Hornett, P. A. Hobson, G. M. Gibson, M. J. Padgett, and E. Hendry, "Noninvasive, near-field terahertz imaging of hidden objects using a single-pixel detector," *Science advances*, vol. 2, no. 6, p. e1600190, 2016.
- 8 Q. Guo, H. Chen, Z. Weng, M. Chen, S. Yang, and S. Xie, "Compressive sensing based high-speed time-stretch optical microscopy for two-dimensional image acquisition," *Optics Express*, vol. 23, no. 23, pp. 29639–29646, 2015.
- 9 W. Gong, C. Zhao, H. Yu, M. Chen, W. Xu, and S. Han, "Three-dimensional ghost imaging lidar via sparsity constraint," *Scientific reports*, vol. 6, no. 1, p. 26133, 2016.
- 10 M.-J. Sun, M. P. Edgar, G. M. Gibson, B. Sun, N. Radwell, R. Lamb, and M. J. Padgett, "Single-pixel three-dimensional imaging with time-based depth resolution," *Nature communications*, vol. 7, no. 1, p. 12010, 2016.
- 11 M.-J. Sun and J.-M. Zhang, "Single-pixel imaging and its application in three-dimensional reconstruction: a brief review," *Sensors*, vol. 19, no. 3, p. 732, 2019.
- 12 G. M. Gibson, S. D. Johnson, and M. J. Padgett, "Single-pixel imaging 12 years on: a review," *Optics express*, vol. 28, no. 19, pp. 28190–28208, 2020.
- 13 Y. Gao, F. Yang, and L. Cao, "Pixel super-resolution phase retrieval for lensless on-chip microscopy via accelerated wirtinger flow," *Cells*, vol. 11, no. 13, 2022.
- 14 Z. Qiu, Z. Zhang, J. Zhong, *et al.*, "Comprehensive comparison of single-pixel imaging methods," *Optics and Lasers in Engineering*, vol. 134, p. 106301, 2020.
- 15 A. B. Ayoub and D. Psaltis, "High speed, complex wavefront shaping using the digital micro-mirror device," *Scientific Reports*, vol. 11, no. 1, p. 18837, 2021.
- 16 F. Ferri, D. Magatti, L. Lugiato, and A. Gatti, "Differential ghost imaging," *Physical review letters*, vol. 104, no. 25, p. 253603, 2010.
- 17 W. Jiang, X. Li, X. Peng, and B. Sun, "Imaging high-speed moving targets with a single-pixel detector," *Optics Express*, vol. 28, no. 6, pp. 7889–7897, 2020.
- 18 W. Zhao, L. Gao, A. Zhai, and D. Wang, "Comparison of common algorithms for single-pixel imaging via compressed sensing," *Sensors*, vol. 23, no. 10, p. 4678, 2023.



- 19 J. Shin, B. T. Bosworth, and M. A. Foster, "Single-pixel imaging using compressed sensing and wavelength-dependent scattering," *Optics letters*, vol. 41, no. 5, pp. 886–889, 2016.
- 20 L. Bian, J. Suo, Q. Dai, and F. Chen, "Experimental comparison of single-pixel imaging algorithms," *JOSA A*, vol. 35, no. 1, pp. 78–87, 2018.
- 21 S. Rizvi, J. Cao, K. Zhang, and Q. Hao, "Improving imaging quality of real-time fourier single-pixel imaging via deep learning," *Sensors*, vol. 19, no. 19, p. 4190, 2019.
- 22 Y. Ni, D. Zhou, S. Yuan, X. Bai, Z. Xu, J. Chen, C. Li, and X. Zhou, "Color computational ghost imaging based on a generative adversarial network," *Optics Letters*, vol. 46, no. 8, pp. 1840–1843, 2021.
- 23 X. Zhang, C. Deng, C. Wang, F. Wang, and G. Situ, "Vgennet: variable generative prior enhanced single pixel imaging," *ACS Photonics*, vol. 10, no. 7, pp. 2363–2373, 2023.
- 24 C. F. Higham, R. Murray-Smith, M. J. Padgett, and M. P. Edgar, "Deep learning for real-time single-pixel video," *Scientific reports*, vol. 8, no. 1, p. 2369, 2018.
- 25 Y. Xu, L. Lu, V. Saragadam, and K. F. Kelly, "A compressive hyperspectral video imaging system using a single-pixel detector," *Nature Communications*, vol. 15, no. 1, p. 1456, 2024.
- 26 H. Shen, L. Gan, N. Newman, Y. Dong, C. Li, Y. Huang, and Y. Shen, "Spinning disk for compressive imaging," *Optics letters*, vol. 37, no. 1, pp. 46–48, 2012.
- 27 C. Li, J. Grant, J. Wang, and D. R. Cumming, "A nipkow disk integrated with fresnel lenses for terahertz single pixel imaging," *Optics express*, vol. 21, no. 21, pp. 24452–24459, 2013.
- 28 W. L. Chan, M. L. Moravec, R. G. Baraniuk, and D. M. Mittleman, "Terahertz imaging with compressed sensing and phase retrieval," *Optics letters*, vol. 33, no. 9, pp. 974–976, 2008.
- 29 E. Hahamovich, S. Monin, Y. Hazan, and A. Rosenthal, "Single pixel imaging at megahertz switching rates via cyclic hadamard masks," *Nature communications*, vol. 12, no. 1, p. 4516, 2021.
- 30 W. Jiang, J. Jiao, Y. Guo, B. Chen, Y. Wang, and B. Sun, "Single-pixel camera based on a spinning mask," *Optics Letters*, vol. 46, no. 19, pp. 4859–4862, 2021.
- 31 Z.-H. Xu, W. Chen, J. Penuelas, M. Padgett, and M.-J. Sun, "1000 fps computational ghost imaging using led-based structured illumination," *Optics express*, vol. 26, no. 3, pp. 2427–2434, 2018.
- 32 H. Huang, L. Li, Y. Ma, and M. Sun, "25,000 fps computational ghost imaging with ultrafast structured illumination," *Electronic Materials*, vol. 3, no. 1, pp. 93–100, 2022.
- 33 P. Kilcullen, T. Ozaki, and J. Liang, "Compressed ultrahigh-speed single-pixel imaging by swept aggregate patterns," *Nature Communications*, vol. 13, no. 1, p. 7879, 2022.
- 34 A. Saade, F. Caltagirone, I. Carron, L. Daudet, A. Drémeau, S. Gigan, and F. Krzakala, "Random projections through multiple optical scattering: Approximating kernels at the speed of light," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6215–6219, IEEE, 2016.
- 35 J. Dong, M. Rafayelyan, F. Krzakala, and S. Gigan, "Optical reservoir computing using multiple light scattering for chaotic systems prediction," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 1, pp. 1–12, 2019.
- 36 U. Paudel, M. Luengo-Kovac, J. Pilawa, T. J. Shaw, and G. C. Valley, "Classification of time-domain waveforms using a speckle-based optical reservoir computer," *Optics express*, vol. 28, no. 2, pp. 1225–1237, 2020.

- 37 H. Cao, "Perspective on speckle spectrometers," *J. Opt.*, vol. 19, no. 6, p. 060402, 2017.
- 38 Y. Wan, X. Fan, and Z. He, "Review on speckle-based spectrum analyzer," *Photonic Sensors*, vol. 11, pp. 187–202, 2021.
- 39 M. Chakrabarti, M. L. Jakobsen, and S. G. Hanson, "Speckle-based spectrometer," *Optics letters*, vol. 40, no. 14, pp. 3264–3267, 2015.
- 40 T. W. Kohlgraf-Owens and A. Dogariu, "Transmission matrices of random media: means for spectral polarimetric measurements," *Optics letters*, vol. 35, no. 13, pp. 2236–2238, 2010.
- 41 Y. Wan, X. Fan, B. Xu, and Z. He, "Microwave frequency measurement with high accuracy and wide bandwidth based on whispering-gallery mode barcode," *Optics Letters*, vol. 46, no. 19, pp. 5008–5011, 2021.
- 42 D. B. Borlaug, S. Estrella, C. T. Boone, G. A. Sefler, T. J. Shaw, A. Roy, L. Johansson, and G. C. Valley, "Photonic integrated circuit based compressive sensing radio frequency receiver using waveguide speckle," *Optics Express*, vol. 29, no. 13, pp. 19222–19239, 2021.
- 43 G. A. Sefler, T. J. Shaw, and G. C. Valley, "Demonstration of speckle-based compressive sensing system for recovering rf signals," *Optics express*, vol. 26, no. 17, pp. 21390–21402, 2018.
- 44 U. Kürüm, P. R. Wiecha, R. French, and O. L. Muskens, "Deep learning enabled real time speckle recognition and hyperspectral imaging using a multimode fiber array," *Optics express*, vol. 27, no. 15, pp. 20965–20979, 2019.
- 45 R. French, S. Gigan, and O. L. Muskens, "Speckle-based hyperspectral imaging combining multiple scattering and compressive sensing in nanowire mats," *Optics letters*, vol. 42, no. 9, pp. 1820–1823, 2017.
- 46 K. Monakhova, K. Yanny, N. Aggarwal, and L. Waller, "Spectral diffusercam: lensless snapshot hyperspectral imaging with a spectral filter array," *Optica*, vol. 7, no. 10, pp. 1298–1307, 2020.
- 47 R. Fan, L. Li, and Y. Zheng, "High-resolution single-pixel imaging based on a probe of single-mode fiber and hybrid multimode fiber," *Optics & Laser Technology*, vol. 175, p. 110732, 2024.
- 48 G. Hu, Y. Qin, and H. K. Tsang, "Multimode fiber speckle imaging using integrated optical phased array and wavelength scanning," *Journal of Lightwave Technology*, 2024.
- 49 J. Hanawa, T. Niiyama, Y. Endo, and S. Sunada, "Gigahertz-rate random speckle projection for high-speed single-pixel image classification," *Optics Express*, vol. 30, no. 13, pp. 22911–22921, 2022.
- 50 T. Yamaguchi, K. Arai, T. Niiyama, A. Uchida, and S. Sunada, "Time-domain photonic image processor based on speckle projection and reservoir computing," *Communications Physics*, vol. 6, no. 1, p. 250, 2023.
- 51 T. Fortier and E. Baumann, "20 years of developments in optical frequency comb technology and applications," *Communications Physics*, vol. 2, no. 1, p. 153, 2019.
- 52 K. Shibuya, T. Minamikawa, Y. Mizutani, H. Yamamoto, K. Minoshima, T. Yasui, and T. Iwata, "Scan-less hyperspectral dual-comb single-pixel-imaging in both amplitude and phase," *Optics express*, vol. 25, no. 18, pp. 21947–21957, 2017.
- 53 I. Coddington, N. Newbury, and W. Swann, "Dual-comb spectroscopy," *Optica*, vol. 3, no. 4, pp. 414–426, 2016.

- 54 Z. Zhu and G. Wu, “Dual-comb ranging,” *Engineering*, vol. 4, no. 6, pp. 772–778, 2018.
- 55 E. Vicentini, Z. Wang, K. Van Gasse, T. W. Hänsch, and N. Picqué, “Dual-comb hyperspectral digital holography,” *Nature Photonics*, vol. 15, no. 12, pp. 890–894, 2021.
- 56 W. Lu, Z. Zhu, B. Willenberg, J. Pupeikis, C. R. Phillips, U. Keller, and S.-c. Chen, “Scan-less 3d microscopy based on spatiotemporal encoding on a single-cavity dual-comb laser,” *Optics Letters*, vol. 49, no. 7, pp. 1766–1769, 2024.
- 57 E. Hase, T. Minamikawa, T. Mizuno, S. Miyamoto, R. Ichikawa, Y.-D. Hsieh, K. Shibuya, K. Sato, Y. Nakajima, A. Asahara, *et al.*, “Scan-less confocal phase imaging based on dual-comb microscopy,” *Optica*, vol. 5, no. 5, pp. 634–643, 2018.
- 58 Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- 59 Z. Liu, L. Wang, Y. Meng, T. He, S. He, Y. Yang, L. Wang, J. Tian, D. Li, P. Yan, *et al.*, “All-fiber high-speed image detection enabled by deep learning,” *Nature communications*, vol. 13, no. 1, p. 1433, 2022.
- 60 H. Kanno, H. Mikami, and K. Goda, “High-speed single-pixel imaging by frequency-time-division multiplexing,” *Optics Letters*, vol. 45, no. 8, pp. 2339–2342, 2020.
- 61 W. Zhao, H. Chen, Y. Yuan, H. Zheng, J. Liu, Z. Xu, and Y. Zhou, “Ultrahigh-speed color imaging with single-pixel detectors at low light level,” *Physical Review Applied*, vol. 12, no. 3, p. 034049, 2019.
- 62 S. Liu, B. Chen, W. Zou, H. Sha, X. Feng, S. Han, X. Li, X. Yao, J. Zhang, and Y. Zhang, “Compressive confocal microscopy imaging at the single-photon level with ultra-low sampling ratios,” *Communications Engineering*, vol. 3, no. 1, p. 88, 2024.
- 63 R. Stojek, A. Pastuszek, P. Wróbel, and R. Kotyński, “Single pixel imaging at high pixel resolutions,” *Optics Express*, vol. 30, no. 13, pp. 22730–22745, 2022.
- 64 L. Li, Y. Zhang, C. Hu, F. Wang, and G. Situ, “Learned spinning mask for high-speed single-pixel imaging,” in *Advanced Fiber Laser Conference (AFL2023)*, vol. 13104, pp. 1235–1242, SPIE, 2024.
- 65 M. Ni, H. Deng, X. Xiao, Y. Cai, and X. Gong, “Feature single-pixel imaging: What you see is what you want,” *Applied Physics Letters*, vol. 122, no. 15, 2023.
- 66 H.-X. Huang, L.-J. Li, Y.-X. Ma, and M.-J. Sun, “50k fps computational ghost imaging with an ultrafast led array,” in *Third International Computing Imaging Conference (CITA 2023)*, vol. 12921, pp. 1305–1311, SPIE, 2023.
- 67 B. Xu, X. Fan, S. Wang, and Z. He, “Broadband and high-resolution electro-optic dual-comb interferometer with frequency agility,” *Optics Express*, vol. 27, no. 6, pp. 9266–9275, 2019.
- 68 L. Chang, S. Liu, and J. E. Bowers, “Integrated optical frequency comb technologies,” *Nature Photonics*, vol. 16, no. 2, pp. 95–108, 2022.
- 69 N. Bender, M. Sun, H. Yilmaz, J. Bewersdorf, and H. Cao, “Circumventing the optical diffraction limit with customized speckles,” *Optica*, vol. 8, no. 2, pp. 122–129, 2021.

## List of Figures

- 1 **Schematic of the proposed scan-less ultrafast SPI system.** **a** Experimental setup of the proposed system. MMF, multimode fiber; COL, collimator; OBJ, object; SL, spherical lens; DAQ, data acquisition. **b** Experimental setup for the generation of dual electro-optic frequency combs. FL, fiber laser; AOM, acousto-optic modulator; DD-MZM, dual-drive Mach-Zehnder modulator; RF, radio frequency; EDFA, erbium-doped fiber amplifier. **c** Spatial masks generated from MMF at different orders of comb lines. **d** Flow of the image reconstruction process. FFT, fast Fourier transform; RFC, RF comb; ANN, artificial neural network. The initial signal is a 1D time-domain interferogram. After FFT of the interferogram, the intensity of each RFC comb line can be extracted and used as input for the ANN to reconstruct the image.
- 2 **Imaging results of neural networks.** **a** The structure of the proposed fully-connected network. **b** Comparison of imaging results from different neural networks. **c** Reconstructed images by fully-connected network on EMNIST and Fashion-MNIST dataset, respectively.
- 3 **Imaging results of the object with different types.** **a** Examples of reconstructed images of different types which are not included in the training dataset. **b** Relationship between the image reconstruction performance and the sampling ratio.
- 4 **Ultrafast imaging results.** **a** Received interference signal during the transient switching process from image 1 to 0 on the DMD. **b** Zoom-in figure of the received interference signal during the transient switching process. **c** Reconstructed images during the transient switching process. The dashed box represents the region where DMD remains stationary during the switching process. **d** Schematic of images of digit "1" and digit "0" displayed on the DMD.  $S$ : stationary region;  $R_{on}$ : the region in DMD that switched from the "off" state to the "on" state;  $R_{off}$ : the region in DMD that switched from the "on" state to the "off" state.
- 5 **Performance metric chart of various SPI methods using sampling ratio and spatial-temporal information flux.** SSE-SPI, scan-less speckle encoded single-pixel imaging.