PHYSICAL REVIEW APPLIED **20,** 064002 (2023)

Robust superresolution computational imaging with diffuser-refined illumination

Xinliang Zhai, Xiaoyan Wu,[*] Tailong Xiao, Jianhong Shi, and Guihua Zeng[†]
State Key Laboratory of Advanced Optical Communication Systems and Networks, and Center for Quantum Sensing and Information Processing, Shanghai Jiao Tong University, 200240 Shanghai, China

(Received 21 August 2023; revised 25 October 2023; accepted 16 November 2023; published 1 December 2023; corrected 22 December 2023)

Computational imaging allows image formation with single-pixel detectors that have no spatial resolution. It has wide-ranging potential applications in fields ranging from biomedical imaging to remote sensing. However, the practical applications of computational imaging are limited due to the high sampling ratio, weak robustness, and low resolution. In this paper, we propose a robust superresolution computational imaging scheme with diffuser-refined illumination and multibucket detection. To increase the reconstruction efficiency, a dual-stream neural network incorporated with a physical model is developed for this imaging scheme, which can reconstruct a high-resolution image in ultralow-sampling-ratio (1.56%) and ultrahigh-noise (4 dB) environments without any data for pretraining. The experimental results demonstrate that the proposed method can increase the spatial resolution by a factor of 2.59 with respect to the diffraction limit of the original imaging system. We also successfully apply this method to image a real biomedical sample and believe that this study provides a computational-imaging framework for complicated practical applications.

DOI: 10.1103/PhysRevApplied.20.064002

I. INTRODUCTION

Computational imaging is an indirect imaging technique that combines front-end optics and postdetection signal processing. In contrast to classical imaging, computational-imaging acquisition techniques have many advantages, including high detecting sensitivity, wide spectral range, and low cost [1–4]. As a well-known computational-imaging technique based on projected light patterns, single-pixel imaging (SPI) [5] or computational ghost imaging (CGI) [6] enables retrieval of the spatial information of an object without any spatially resolving array detectors. Benefiting from these advantages, computational imaging has great potential applications in many fields ranging from microscopy [7,8] to remote sensing [9,10].

However, such computational-imaging systems based on projected patterns naturally encounter a challenge of realizing reconstruction of high-resolution images with an ultralow sampling ratio in a low-signal-to-noise-ratio (SNR) channel environment. The most-direct way to address this limitation is to improve the imaging system, such as the method of detection [11,12] or the modulation rate [13]. For example, we can easily reduce the sampling by using multiple single-pixel detectors [14]. By use of heuristic optimization of structural illumination, it is possible to achieve adaptive ghost imaging [15]. In addition to improvement of the optical system, some studies have addressed this problem from the perspective of algorithms. Specifically, both compressive sensing–based [5,16,17] and deep learning–based solutions were proposed in succession [18,19]. By leveraging prior information, these methods can significantly reduce the sampling rate under the premise of increasing reconstruction or pretraining time. Recently, some methods have used known physical processes of imaging as *a priori* information to update the network and reconstruct superresolution results [20,21] at a low sampling ratio. The increase in resolution was attributed to the algorithm's ability to decode the high-frequency components of the target, which were highly compressed into one-dimensional bucket signals by random illumination. The increase in resolution is the result of a combination of the advanced reconstruction algorithm and fine-grained modulation speckles. Recently, Tian *et al.* [22] studied the influence of hybrid speckle patterns in ghost imaging and found that large speckles can increase the contrast-to-noise ratio and that small speckles help to increase the resolution during ghost-imaging reconstruction. Therefore, it is possible to improve the quality of CGI (SNR and resolution) by introducing an appropriate diffuser in the spatial modulation.

An analogous concept has been introduced in microscopy to attain superresolution capabilities [23]. Recently, both theoretical [24] and experimental [25] studies have demonstrated that periodic illumination schemes used in structured-illumination microscopy can be replaced by random speckles. Surprisingly, low-resolution images

---

[*] xiaoyanwu@sjtu.edu.cn
[†] ghzeng@sjtu.edu.cn

obtained under random speckled illumination can be processed to yield sample images with greater lateral resolution than sample images that can be obtained with classical widefield microscopy. Moreover, Haridas *et al.* [26] proposed a high-resolution-microscopy method using structured embedded speckle illumination. Their scheme provided sub-diffraction-limit-resolution imaging with a high SNR.

Inspired by the existing studies mentioned above, here we report a superresolution-computational-imaging scheme with diffuser-refined illumination and multibucket detection. Based on the traditional SPI or CGI system configuration, the proposed scheme introduces an additional diffuser before the programmable modulator. Besides, we use four single-pixel detectors in the proposed method. The use of multiple buckets not only intuitively brings benefits in reducing the sampling rate but also allows us to leverage redundant information by performing data fusion in the frequency domain, thereby further enhancing the quality of imaging. To accommodate the proposed imaging scheme, we further propose a physically enhanced dual-stream-neural-network model as a complementary reconstruction algorithm. One stream takes differential-ghost-imaging (DGI) [27] results obtained from different bucket detectors as the input and performs frequency-domain fusion, while the other stream takes the four bucket signals as a multichannel low-resolution image and performs upsampling. The two streams converge in the middle layer and undergo superresolution network processing to yield the final reconstruction result. The resulting neural network does not require pretraining on any dataset for image reconstruction. Experiments are performed to demonstrate that the proposed imaging scheme brings advantages in terms of accuracy, robustness, and resolution. Compared with widely used methods, the proposed approach yields superior outcomes under an extremely low detection SNR (4 dB) and sampling ratio (1.56%). Moreover, it enables a more than twofold increase in resolution beyond the resolution limit of the original imaging system.

II. IMAGING SCHEME

A schematic diagram of computational imaging with diffuser-refined illumination is shown in Fig. 1. Differently from the conventional CGI system, we insert a diffuser in front of the digital micromirror device (DMD) and use more single-pixel detectors [see Fig. 1(a)]. Through the diffuser, the laser beam is scattered and a pseudorandom light field is obtained and then modulated by the DMD. The reflected light field interacts with the target and is finally received by four single-pixel detectors.

As shown in Fig. 1(b), the original patterns loaded on the DMD are illuminated by the same pseudorandom light field, and we name the reflected combined illuminations as "shattered illuminations." Here we define μ as the lateral scale factor of the modulation pixel size, and then $\mu = d_s/d$, where d is the pixel size of original patterns on the DMD d_s is the average size of speckles generated by the diffuser. If tensor P denotes the original patterns, $P \in \mathbb{R}^{\sqrt{N}\times\sqrt{N}\times M}$, where N and M are the number of pixels and samplings, respectively. If matrix H denotes the speckle pattern, $H \in \mathbb{R}^{(1/\mu)\sqrt{N}\times(1/\mu)\sqrt{N}}$. Thus, the shattered illuminations can be expressed as $HP = [H \odot P_1, H \odot P_2 \cdots H \odot P_M]$, where $\odot$ denotes elementwise multiply. $P_M \in \mathbb{R}^{\sqrt{N}\times\sqrt{N}}$. Note that the numbers of pixels in P_M and H are not equal due to $\mu < 1$, so each pixel in P is refined into $1/\mu \times 1/\mu$ smaller pixels. In our experiment, $\mu \approx 0.4$; the effectiveness of DMD pixel refinement is visualized in Fig. 1(d).

Replacement of one bucket detector with multiple bucket detectors is the other change to the imaging system, which is an intuitive approach to decrease the sampling rate of ghost imaging [14]. As shown in Fig. 1(a), we use four single-pixel detectors located at different positions to collect the optical fluctuation signals in ghost imaging. These four bucket signals [see Fig. 1(c)] are, respectively, the measurements of four overlapping regions of the target. However, the imaging systems for such schemes are difficult to calibrate and suffer from boundary redundancy caused by diffraction. In this work, we leverage redundant information to obtain the high-resolution-imaging result. Thus, this solution does not require calibration and elimination of boundary redundancy. Specifically, the kth subregion of the target is detected by the kth single-pixel detector. Assuming the intensity transmission function of the object is $T(x)$, the corresponding bucket signal I_k can be expressed as

$$I_k = \int_{A_k} h * [H(x)P(x)]\, T(x) dx, \quad k = 1, 2, \ldots, K, \tag{1}$$

where h, $H(x)$, and $P(x)$ represent the point-spread function (PSF) from the interference factors (e.g., diffraction limit, turbulence, and aberration of system), the known speckle matrix of the diffuser, and the pattern loaded on DMD, respectively. The symbol "$*$" denotes convolution. Consequently, the intensity set collected by multiple single-pixel detectors can be represented as $\{I_0, I_1, \ldots, I_K\}$, where K is the number of single-pixel detectors. The product result, $H(x)P(x)$, is the proposed shattered illumination, mentioned above. A_k is the domain of integration (e.g., area of the illumination).

In conventional DGI theory [27], the image information of the object can be retrieved from the correlation calculation

$$\langle O_k \rangle = \langle \left(I_k - \frac{\langle I_k \rangle}{\langle R_k \rangle} R_k \right) ((HP)_k - \langle (HP)_k \rangle) \rangle, \tag{2}$$

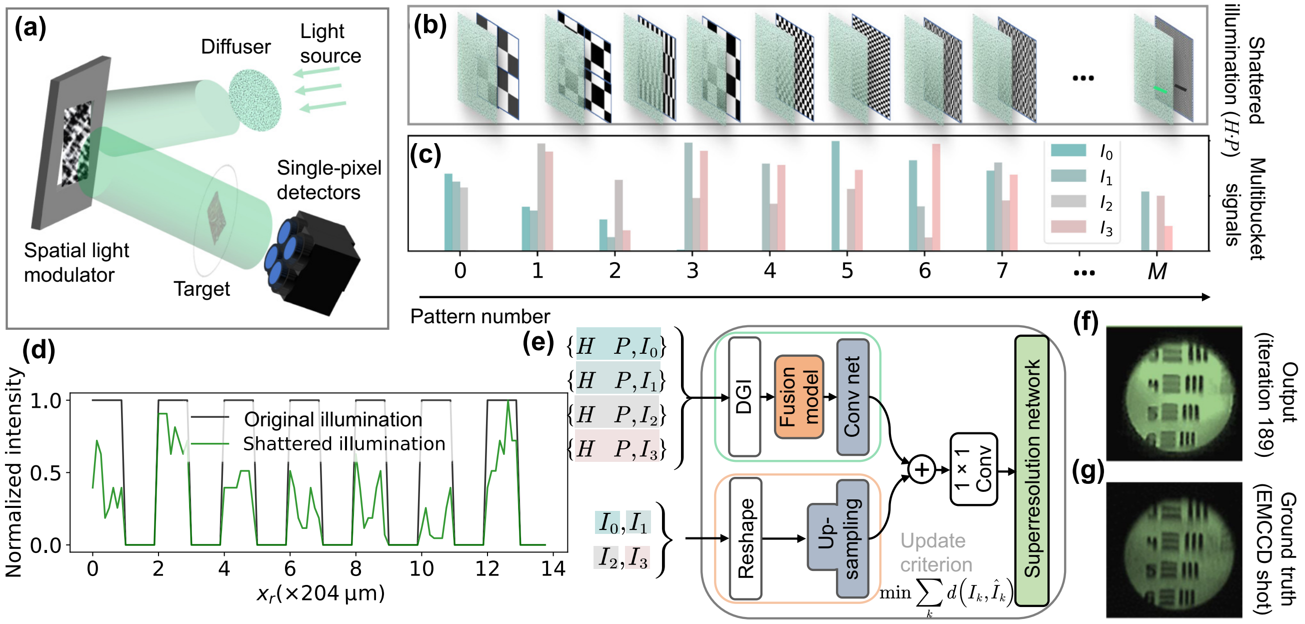

FIG. 1. Overview of the proposed computational-imaging scheme. (a) Sketch of the imaging system. Shattered-illumination modes are generated by a diffuser and a spatial light modulator. The modulated light is measured by multiple single-pixel detectors after interacting with the target. (b) Shattered-illumination modes generated by the combination of speckle pattern H(left) and modulated pattern P (right). (c) Intensity sequence measured by multiple single-pixel detectors. (d) Cross-section profile of the two lines in last illumination shown in (b). The black line represents the modulated pattern P and the green line represents the shattered illumination HP. (e) Dual-stream reconstruction network designed for the imaging scheme. The top stream takes DGI results as the input and the bottom stream takes raw intensities as the input. Then the two streams are aggregated in the middle layer and then output a superresolution image result. "Conv" denotes the convolution operation. The network is updated on the basis of the metric learning between measurements I_k and predictions $\hat{I}_k$. (f) Final output (128 × 128 pixels) of the dual-stream network at iteration 189. (g) Ground truth captured by the electron-multiplying CCD (EMCCD). Conv net, conventional network.

where $\langle\cdot\rangle$ represents the ensemble average and R_k denotes the reference intensity, which is the mean value of all elements of HP. The whole result for the object is obtained by use of the total intensity of K bucket signals:

$$\langle O\rangle \approx \langle\left(\sum_k I_k - \frac{\langle\sum_k I_k\rangle}{\langle R\rangle}R\right)(HP - \langle HP\rangle)\rangle. \tag{3}$$

To fully utilize the physical prior information, we design a dual-stream neural network to reconstruct the object. The algorithm updates the parameters in a self-supervised manner without any training process, which is essentially a reparameterization of the neural network solving the optimization problem. In our proposed imaging scheme, K different bucket signals are first directly regarded as a low-resolution-image result with $\sqrt{K} \times \sqrt{K}$ pixels and M channels. Therefore, an upsampling network is built to process the upsampling and interpolation by use of the information from the multichannel input [see the bottom stream in Fig. 1(e)]. Meanwhile, K different bucket signals are then applied to reconstruct K different DGI results according to Eq. (2). As illustrated in the top stream in Fig. 1(e), these DGI results obtained by multiple detectors are first fused in the frequency domain, i.e., the output of the fusion model is

$$O_f = \mathcal{F}^{-1}\left(\frac{1}{K}\sum_k \mathcal{F}(\mathrm{Pad}(O_k))\right), \tag{4}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the Fourier transform and inverse Fourier transform, respectively, O_f is the output feature of the fusion model, and Pad$(\cdot)$ means the zero-padding operation for position alignment. The convolutional network then performs the initial feature extraction on O_f and maps it to the feature dimension as the output of the bottom branch for feature alignment. Then the outputs of the two streams merge through the element adding "$\oplus$" and finally are input to a well-designed superresolution network to produce a high-quality estimation. The detailed network architecture is demonstrated in Appendix A. In this work, the whole network plays the role of a solver for the optimization problem of computing inverse physical processes

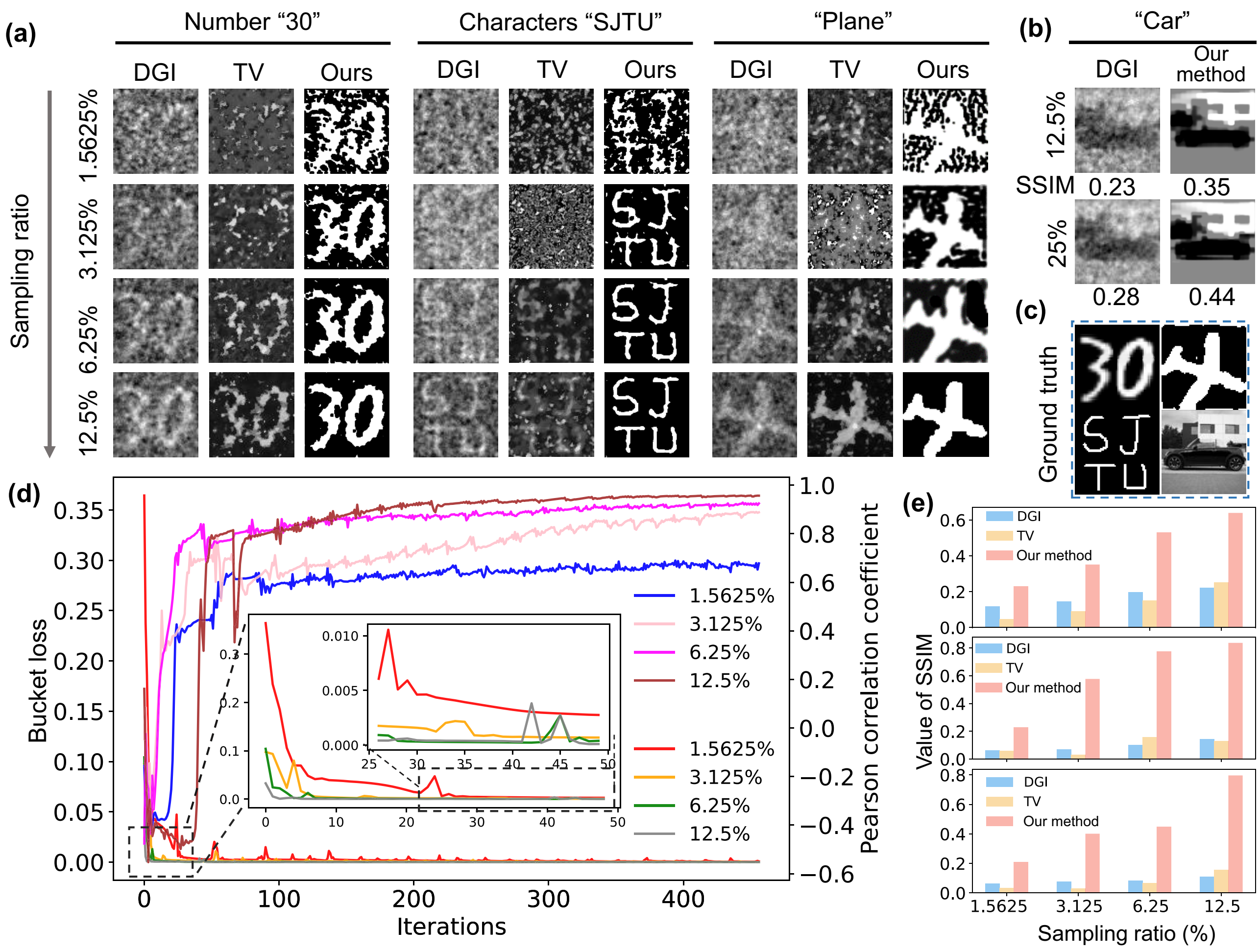

FIG. 2. Experimental comparisons of DGI, TV, and the proposed method in terms of both sampling ratio and reconstruction quality. (a) Experimental results for binary objects. Each row in the group represents the results reconstructed by different methods at the same sampling ratio, while each column represents the results for the objects reconstructed by the same method at different sampling ratios. The number of iteration steps is 200. The pixel resolution is 128 × 128. (b) Experimental results for a grayscale object (128 × 128) at two different sampling ratios. The SSIM values are shown at the bottom of the corresponding images. (c) Ground truth of objects. (d) Bucket loss between I_k and $\hat{I}_k$ and the PCC between the ground truth and the restoration along with the number of iteration steps from 0 to 450 under different sampling ratios for object characters "SJTU." There are two legends in (d), the top one is a legend for the PCC curves and the bottom one is a legend for the bucket loss curves. (e) Quantitative comparison of different methods at different sampling ratios for different objects (top, number "30"; middle, characters "SJTU"; bottom, "plane").

of imaging. The objective function can be formulated as

$$\min_{\theta} \frac{1}{K} \sum_{k} \|HP\mathcal{R}_{\theta}\left[I_k;(H,P)\right] - I_k\|_2^2 + \tau \sum_{x} |\nabla \mathcal{R}_{\theta}(I_k;(H,P))|, \qquad (5)$$

where $\mathcal{R}_\theta$ denotes the proposed model and τ is a weighting factor. The first term of the function is the mean square error between measurements and the bucket predictions. The second term is the total variation [28] of the output image, which is beneficial for converging and improving the reconstruction results.

III. RESULTS AND DISCUSSION

Here we perform a comparative study of the effectiveness of the proposed method, including the sampling ratio, noise level, and resolution.

A. Imaging results at different sampling ratios

To demonstrate the ability of the proposed computational-imaging scheme, an optical system is designed as in Fig. 5 in Appendix B. To calculate the experimental SNR, we need a dark room with as little background light as possible to acquire the ground truth. After the ground truth is obtained, the SNR (SNR $= 10\log_{10} V_s/V_n$, where V_s and

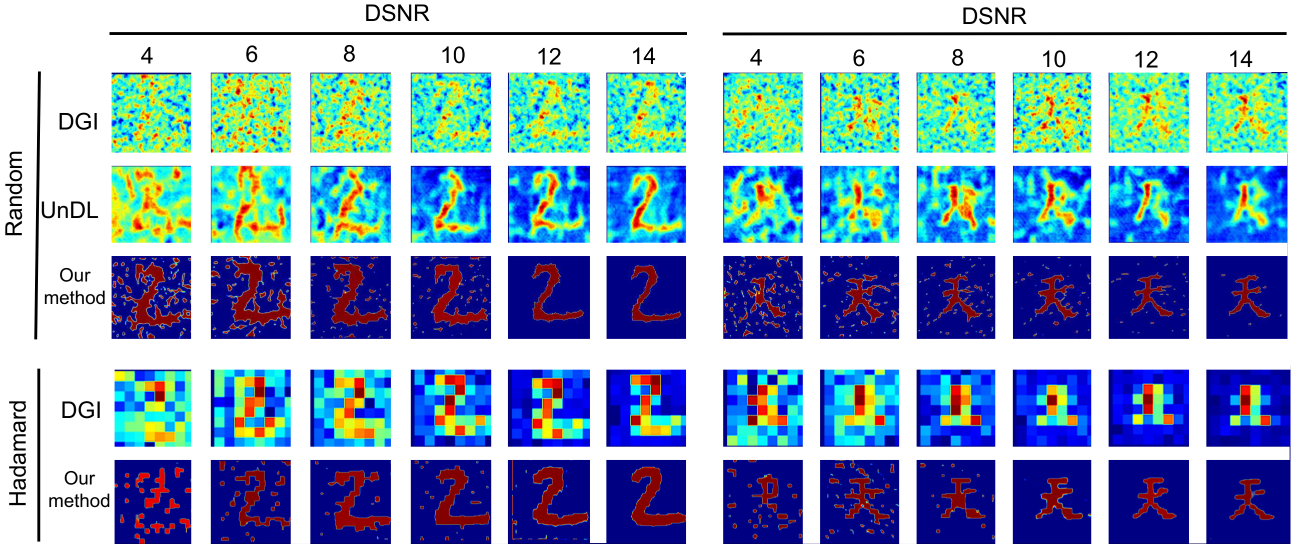

FIG. 3. Robustness comparisons of DGI, UnDL, and the proposed method in terms of noise level, modulated pattern, and reconstruction quality. The sampling ratio is set as 3.125% and the pixel resolution is 128 × 128. The six columns on the left show the imaging results for number "2" under different noise levels. The six columns on the right show the results for the Chinese character representing "Tian" under different noise levels. The top three rows show the results obtained with random patterns and the bottom two rows show the results obtained with Hadamard patterns.

V_n are the average intensity of pure signals and noise signals under multiple measurements) of our experimental environment is calculated to be approximately 10–13 dB. Because the sampling ratio is one of the most-important factors in computational imaging, here we experimentally demonstrate the reconstruction results at different sampling ratios.

The imaging results for three binary objects [number "30," characters "SJTU," and an airplane ("plane")] obtained with different methods (DGI [27] and total variation (TV) [28]) and sampling ratios (1.56%, 3.13%, 6.25%, and 12.50%) are shown in Fig. 2(a). For DGI and TV, the bucket signals from the four detectors are aggregated and processed as one bucket-signal vector. One can clearly observe that the reconstruction results of all methods degrade significantly as the sampling rate decreases. Visually, the proposed method gives better results (e.g., cleaner and higher contrast) than both DGI and TV under the same sampling ratio, even at the extremely low sampling ratio 1.56% [see the first row in Fig. 2(a)]. To quantitatively evaluate the accuracy of imaging, we calculate the structural-similarity-index measure (SSIM) [29] and depict the results in Fig. 2(e). From the top to the bottom are the results for number "30,", characters "SJTU," and "plane," respectively. The results show the proposed method significantly improves the imaging quality under all conditions. Furthermore, our imaging quality at a sampling ratio of 1.56% is also better than the results restored by DGI and TV at higher sampling ratios. Besides, we visualize the changes in bucket loss and imaging quality of characters "SJTU" during the optimization process in Fig. 2(d). The similarity between the reconstructed image and the ground truth is portrayed by use of the Pearson correlation coefficient (PCC) [30]. One can see that along with the loss reduction, all reconstructions under different sampling ratios converge to good-quality images. In addition, the higher the sampling rate, the faster it will converge to a higher PCC and a lower bucket loss. We also reconstruct a grayscale object under two different sampling ratios in Fig. 2(b). By comparing the values of the SSIM, we can obtain conclusions similar to those obtained for the imaging binary target.

According to compressive-sensing theory, the sparse *a priori* information of the natural image is one of the reasons why this scheme can perform high-accuracy computational imaging at very low sampling ratios. In addition, the use of multiple single-pixel detectors can be expected to trade spatial costs (number of single-pixel detectors) for less sampling time (number of measurements).

B. Imaging results at different noise levels

We also demonstrate the robustness of the proposed method under different noise conditions. Since noise sources in imaging systems are complex, it is hard to control the noise level accurately in experiments. Therefore, we modeled noise effects as an additive Gaussian distribution to explore the imaging ability under different noise levels. Specifically, because of the existence of detection noise (shot noise) and background noise in the actual imaging environment, we aggregate all the different types of noise and the object signals into an intensity sequence. The

TABLE I. The metrics of different GI reconstruction methods for the SSIM and the PSNR for the DSNR ranging from 4 to 14. The value on the left of the oblique is the PSNR and the value on the right of the oblique is the SSIM. Values in bold are the highest values of the quantitative metrics. (R) indicates the use of random patterns. (H) indicates the use of hadamard patterns.

		DSNR					
Method		4	6	8	10	12	14
Character “2”	DGI (R)	6.14/0.05	6.24/0.06	7.09/0.06	8.07/0.10	8.04/0.11	8.13/0.13
	UnDL (R)	6.58/0.06	7.57/0.11	9.10/0.13	11.88/0.17	11.87/0.19	12.10/0.21
	Ours (R)	**8.72/0.21**	**9.32/0.32**	**12.02/0.46**	**12.76/0.50**	**16.92/0.74**	**16.08/0.77**
	DGI (H)	8.37/0.06	9.17/0.09	9.41/0.09	11.09/0.10	12.47/0.11	12.84/0.12
	Ours (H)	**9.63/0.37**	**10.28/0.40**	**12.18/0.54**	**13.40/0.57**	**13.97/0.69**	**15.87/0.79**
Chinese character representing “Tian”	DGI (R)	6.19/0.02	6.54/0.04	6.63/0.04	6.71/0.04	6.65/0.06	6.87/0.08
	UnDL (R)	8.37/0.04	8.43/0.06	8.80/0.06	10.08/0.09	10.99/0.10	11.33/0.12
	Ours (R)	**10.76/0.25**	**11.96/0.43**	**13.64/0.46**	**14.32/0.58**	**17.42/0.67**	**18.11/0.84**
	DGI (H)	8.59/0.04	8.76/0.04	10.68/0.05	12.37/0.05	12.48/0.06	13.12/0.09
	Ours (H)	**10.83/0.51**	**10.75/0.52**	**13.36/0.65**	**13.91/0.77**	**15.99/0.85**	**17.28/0.88**

noisy bucket signal can be expressed as

$$I_k = \int_{A_k} H(x)P(x)T(x)dx + n(\sigma), \tag{6}$$

where $n(\sigma)$ is the additive noise sampled from a zero-mean Gaussian distribution with standard deviation of σ. To quantify the noise level, we also define the detection signal-to-noise ratio (DSNR) as

$$\gamma = 10\log_{10}\frac{I_k}{\sigma}. \tag{7}$$

We show the reconstruction results for different objects obtained with different methods at different noise levels in Fig. 3. The sampling ratio is set as 3.13%. The first group of results, obtained with random patterns, is plotted in the first three rows in Fig. 3 and the second group of results, obtained with Russian-doll-resorted Hadamard patterns [31], is plotted in the last two rows in Fig. 3. The six columns on the left show the imaging results for number “2” and the six columns on the right show the results for the Chinese character representing “Tian.” One can see that all objects have been successfully restored by the proposed approach, even with the DSNR as low as 4. We also take DGI and UnDL [32], which is an unsupervised-leaning-based antinoise GI method for comparison. The results show that the imaging quality of all methods improves with the increase of the DSNR, The proposed method outperforms DGI and UnDL for all cases (different DSNR settings and objects). To give a quantitative evaluation, we also use the peak signal-to-noise ratio (PSNR) and the SSIM as the indexes of image-quality assessment; the results are shown in Table I. We can see that the proposed method has the highest metric values in all cases, suggesting that the reconstruction accuracy of this method outperforms the other methods.

Consistent with the use of speckle-pattern illumination instead of illumination with a periodic grating pattern in fluorescence microscopy (e.g., structured-illumination microscopy), the object information is able to be better encoded by increasing the noise suppression and therefore an increased imaging SNR can be obtained with this shattered illumination [26,33]. The other reason for the robustness of the proposed method with regard to noise is because the statistical properties of the noise tend to zero mean. On this basis, the convergence result of the neural network with L_2 loss will converge to the mean value of the target distribution [34].

C. Imaging resolution with different methods

In this section, we first experimentally demonstrate the spatial resolution the proposed method can offer and then give the results of imaging a biomedical sample. The pixel size of the modulation system is one constraint on the resolution of the computational-imaging system. In the same nonzoomed-imaging-system configuration, use of the modulation device with smaller programmable microcells can encode finer information about the target. The point-spread function h is the other constraint on the resolution. Ideally, h is a δ function. However, in most practical computational-imaging scenarios, it cannot be approximated as a δ function. One reason is the diffraction limit and another reason is the disturbance caused by the imaging system’s defects and environmental noise. In GI schemes, the speckle grain size on the object plane can be obtained by measuring the normalized correlation function of the light-intensity distribution on the object plane [35]. Specifically, we can obtain the resolution limit of a GI system by measuring the full width at half maximum [21]. In our nonzoomed $4f$ imaging system, the pixel size of the modulator is 162 μm (15 binned pixels of the DMD in both the horizontal direction and the vertical direction). The full width at half maximum of the smallest programmable microcell on the object plane is 204 μm, which is suggested as the practical diffraction limit of the system. The

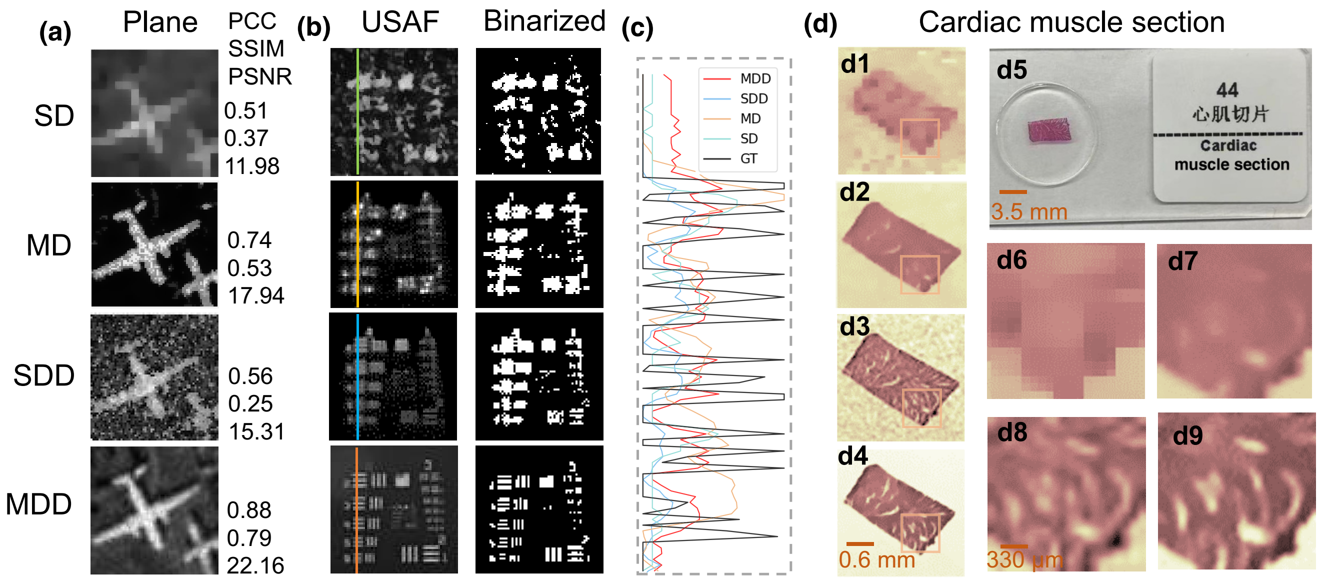

FIG. 4. Comprehensive comparison of different configurations including use of a single detector (SD), multiple detectors (MD), a single detector with a diffuser (SDD), and multiple detectors with a diffuser (MDD) in terms of both image quality and imaging resolution. (a) Experimental comparison of four different configurations (from top to bottom, SD, MD, SDD, and MDD). (b) Experimental results for the USAF resolution target (left column) and the corresponding binarized results (right column). (c) Cross-section profile of the colored lines shown in (b). (d) Results of imaging a cardiac muscle section. Panels d1, d2, d3, and d4 show the results for SD, MD, SDD, and MDD, respectively. Panel d5 shows a snapshot of the sample. Panels d6, d7, d8, and d9 show enlargements of panels d1, d2, d3, and d4, respectively. The number of iteration steps is 200, the sampling ratio is 6.25%, and the pixel resolution is 128 × 128. GT, ground truth.

speckle grain size d_s generated by the diffuser is about 65 μm. Thus, the lateral scale factor μ of the system is 0.4.

As shown in Fig. 4(a), we first demonstrate the accuracy metrics of different configurations including use of a single detector, multiple detectors, a single detector with a diffuser, and multiple detectors with a diffuser. The imaging conditions for the four configurations are the same, i.e., the number of iteration steps is 200, the sampling ratio is set as 6.25%, the modulated patterns are Hadamard patterns, and the experimental environment is noisy (SNR $\approx$ 11 dB). The imaging target is a two-plane image as shown in Fig. 4(a). From comparison of the results for multiple detectors without a diffuser and a single detector without a diffuser, one can clearly see that the result obtained by use of multiple bucket detectors is much better than the result obtained by use of a single bucket detector in terms of visual appearance and quantitative evaluation indexes (PCC, SSIM, and PSNR). From comparison of the results obtained with a single detector with a diffuser and a single detector without a diffuser, the introduction of the diffuser slightly improves the imaging result in terms of the PCC and the PSNR, but the result obtained with a single detector with a diffuser has more background noise. From the comparisons shown in Fig. 4(a), it is evident that the proposed scheme (multiple detectors with a diffuser) can provide the best reconstruction result among the four configurations.

To demonstrate the spatial resolution the four schemes can offer, a USAF resolution target is used as the object for imaging. The results show that the proposed method [multiple detectors with a diffuser, the last row in Fig. 4(b) and the red line in Fig. 4(c)] can achieve a resolution of about 78.75 um (group 2, element 5), which suggests that the method has the capability of increasing the resolution by a factor of about 2.59 (204/78.75) with respect to the system resolution limit. In addition, as shown in the third row in Fig. 4(b) and the blue line in Fig. 4(c), the image reconstructed by use of a single detector with a diffuser in group 2, element 2 is also resolvable, and the resolution is increased by a factor of about 1.83 (204/111.36).

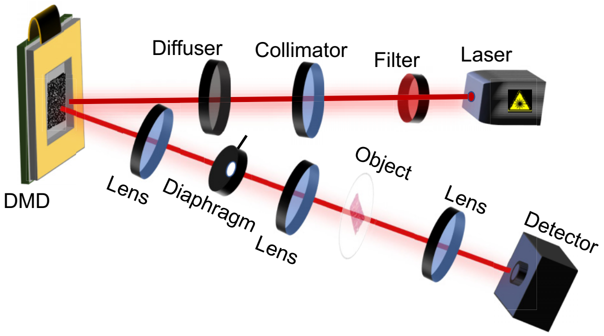

FIG. 5. Schematic diagram of the experimental setup.

As expected, the introduction of the diffuser in the illumination will increase the resolution of the imaging system. Recently, the same concept was used for high-resolution imaging and optical sectioning in fluorescence microscopy [26].

We also conduct a biomedical-sample imaging experiment to demonstrate the effectiveness of the proposed method. The imaging target, panel d5 in Fig. 4(d), is a stained-cardiac-muscle-section sample. As expected, the imaging quality for multiple detectors with a diffuser in panel d4 in Fig. 4(d) is better than with the other three configurations. From observation of the enlarged images of the regions in the orange rectangles in panels d1–d4 in Fig. 4(d), which are shown in panels d6–d9, respectively, both a single detector with a diffuser and multiple detectors with a diffuser can be used to successfully obtain resolvable results. Moreover, use of multiple detectors with a diffuser results in a higher-resolution and clearer result than use of a a single detector with a diffuser because because the dual-stream network is able to fully utilize redundant information from multiple detectors.

IV. CONCLUSION

We have proposed a computational-imaging scheme with diffuser-refined illumination and multibucket detection for robust superresolution imaging. Based on the conventional SPI or CGI system, a diffuser and four single-pixel detectors are introduced additionally into the optical system to increase the imaging resolution and robustness. Besides, a dual-stream model has been developed for image reconstruction. The shattered illumination provides the opportunity for sub-diffraction-limit resolution with a high SNR. Meanwhile, the multibucket detection reduces the sampling ratio effectively. The dual-stream neural network acts as the solver of the inverse imaging process with the constraint of the prior knowledge, so it can well reconstruct the object in sub-Nyquist sampling. The experimental results demonstrated that the proposed method is able to image an object at a very low sampling ratio (1.56%) and a very low DSNR (4 dB). Moreover, the experimental results also demonstrate an increase of resolution by a factor of 2.59. Through a comprehensive comparison with other methods, the proposed method has been demonstrated to be superior in both robustness and accuracy. In addition, we have imaged biological specimens using this approach, which serves as an illustrative example of its practical application. Both the proposed imaging scheme and the dual-stream reconstruction model can be applied for diverse structural illumination–based imaging systems.

However, it is worth pointing out that the proposed method relies on the scattering characteristic of the diffuser and the number of detectors. The diffuser speckle must have a smaller spot size than the modulator's pixel. Besides, the sampling rate can be further reduced by use of more detectors theoretically.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grant No. 61905140) and the Science and Technology Commission of Shanghai Municipality (Grant No. 22142201900).

APPENDIX A: NETWORK ARCHITECTURE AND HYPERPARAMETERS

Here we introduce the network architecture depicted in Fig. 1(e). In the top stream in Fig. 1(e), the convolutional network consists of four convolutional layers with the leaky-rectified-linear-unit (ReLU) activation function and batch normalization. The upsampling operation in the bottom stream is implemented by six transpose convolutional layers with the ReLU function and batch normalization. The superresolution network in Fig. 1(e) is derived from the enhanced deep residual networks for single image super-resolution (EDSR) net [36] with a scale of 1. The detailed network architecture is depicted in Table II.

TABLE II. The detailed network architecture. "Conv" denotes the convolutional operation.

Name	Layers	Kernel size	Stride	Output channel
Convolutional network	Conv_1	3 × 3	1	32
	Conv_2	3 × 3	1	32
	Conv_3	3 × 3	1	16
	Conv_4	3 × 3	1	16
Upsampling	Transposed conv_1	3 × 3	2	512
	Transposed conv_2	3 × 3	2	256
	Transposed conv_3	5 × 5	2	128
	Transposed conv_4	5 × 5	2	64
	Transposed conv_5	5 × 5	2	32
	Transposed conv_6	5 × 5	2	16
1 × 1 Conv	Conv_5	3 × 3	1	1
Superresolution network	EDSR	3 × 3	···	1

We use the Adam optimizer with learning rate $\alpha = 0.05$, $\beta_1 = 0.5$, $\beta_2 = 0.9$, and $\epsilon = 10^{-9}$ to update the proposed model. The weighting factor τ of TV is set as 10^{-5}. The leak parameter of leaky ReLU was 0.2. The algorithm runs on a computer with an Intel i7-6800K CPU, 64 GB RAM, and two NVIDIA Titan Xp graphics processing units with 24 GB virtual VRAM. The deep-learning platform is PyTorch version 1.9.1.

APPENDIX B: EXPERIMENTAL SETUP

A schematic diagram of computational imaging with diffuser-refined illumination is shown in Fig. 5. The light source is a pulsed laser with a filter set to $\lambda = 680$ as the wavelength. After collimation, the light passes through a diffuser (DG10-1200-MD, Thorlabs). Through the diffuser, the light is scattered and a pseudorandom light field is obtained and then embedded with the premodulated patterns on a DMD (Texas Instruments DLC9500P24, 1080×1920). The reflected light is projected on the object after a $4f$ system that consists of two lenses with a focal length of 100 mm. Finally, the object light is collected by another lens with a focal length of 100 mm and is detected by a four-pixel camera (Andor iXon Ultra 888, 1024×1024). We use 16×16 pixels of the camera with 8×8 binning to replace four single-pixel detectors.

[1] M. Levoy, Light fields and computational imaging, Computer **39**, 46 (2006).

[2] G. Barbastathis, A. Ozcan, and G. Situ, On the use of deep learning for computational imaging, Optica **6**, 921 (2019).

[3] O. Cossairt, M. Gupta, and S. K. Nayar, When does computational imaging improve performance?, IEEE Trans. Image Process. **22**, 447 (2012).

[4] B. Sun, M. P. Edgar, R. Bowman, L. E. Vittert, S. Welsh, A. Bowman, and M. J. Padgett, 3D computational imaging with single-pixel detectors, Science **340**, 844 (2013).

[5] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, Single-pixel imaging via compressive sampling, IEEE Signal. Process. Mag. **25**, 83 (2008).

[6] J. H. Shapiro, Computational ghost imaging, Phys. Rev. A **78**, 061802 (2008).

[7] G. Donnert, J. Keller, R. Medda, M. A. Andrei, S. O. Rizzoli, R. Lührmann, R. Jahn, C. Eggeling, and S. W. Hell, Macromolecular-scale resolution in biological fluorescence microscopy, Proc. Natl. Acad. Sci. **103**, 11440 (2006).

[8] R. O. Dror, R. M. Dirks, J. Grossman, H. Xu, and D. E. Shaw, Biomolecular simulation: A computational microscope for molecular biology, Annu. Rev. Biophys. **41**, 429 (2012).

[9] W. Gong, C. Zhao, H. Yu, M. Chen, W. Xu, and S. Han, Three-dimensional ghost imaging lidar via sparsity constraint, Sci. Rep. **6**, 26133 (2016).

[10] S. Royo and M. Ballesta-Garcia, An overview of lidar imaging systems for autonomous vehicles, Appl. Sci. **9**, 4093 (2019).

[11] R. I. Stantchev, X. Yu, T. Blu, and E. Pickwell-MacPherson, Real-time terahertz imaging with a single-pixel detector, Nat. Commun. **11**, 2535 (2020).

[12] Y. Wang, K. Huang, J. Fang, M. Yan, E. Wu, and H. Zeng, Mid-infrared single-pixel imaging at the single-photon level, Nat. Commun. **14**, 1073 (2023).

[13] W. Jiang, Y. Yin, J. Jiao, X. Zhao, and B. Sun, 2 000 000 fps 2D and 3D imaging of periodic or reproducible scenes with single-pixel detectors, Photonics Res. **10**, 2157 (2022).

[14] M.-J. Sun, W. Chen, T.-F. Liu, and L.-J. Li, Image retrieval in spatial and temporal domains with a quadrant detector, IEEE Photonics J. **9**, 1 (2017).

[15] B. Liu, F. Wang, C. Chen, F. Dong, and D. McGloin, Self-evolving ghost imaging, Optica **8**, 1340 (2021).

[16] C. Zhao, W. Gong, M. Chen, E. Li, H. Wang, W. Xu, and S. Han, Ghost imaging lidar via sparsity constraints, Appl. Phys. Lett. **101**, 141123 (2012).

[17] O. Katz, Y. Bromberg, and Y. Silberberg, Compressive ghost imaging, Appl. Phys. Lett. **95**, 131110 (2009).

[18] M. Lyu, W. Wang, H. Wang, H. Wang, G. Li, N. Chen, and G. Situ, Deep-learning-based ghost imaging, Sci. Rep. **7**, 1 (2017).

[19] F. Wang, H. Wang, H. Wang, G. Li, and G. Situ, Learning from simulation: An end-to-end deep-learning approach for computational ghost imaging, Opt. Express **27**, 25560 (2019).

[20] F. Wang, C. Wang, C. Deng, S. Han, and G. Situ, Single-pixel imaging using physics enhanced deep learning, Photonics Res. **10**, 104 (2022).

[21] F. Wang, C. Wang, M. Chen, W. Gong, Y. Zhang, S. Han, and G. Situ, Far-field super-resolution ghost imaging with a deep neural network constraint, Light: Sci. Appl. **11**, 1 (2022).

[22] T. Tian, Z. Sun, S. Oh, and C. Spielmann, Image-enhanced pseudo-thermal ghost imaging with hybrid speckle pattern, Photonics **10**, 7 (2023).

[23] T. Mangeat, S. Labouesse, M. Allain, A. Negash, E. Martin, A. Guénolé, R. Poincloux, C. Estibal, A. Bouissou, S. Cantaloube, *et al.*, Super-resolved live-cell imaging using random illumination microscopy, Cell Rep. Methods **1**, 100009 (2021).

[24] J. Idier, S. Labouesse, M. Allain, P. Liu, S. Bourguignon, and A. Sentenac, On the superresolution capacity of imagers using unknown speckle illuminations, IEEE Trans. Comput. Imaging **4**, 87, 2017.

[25] S. Labouesse, A. Negash, J. Idier, S. Bourguignon, T. Mangeat, P. Liu, A. Sentenac, and M. Allain, Joint reconstruction strategy for structured illumination microscopy with unknown illuminations, IEEE Trans. Image Process. **26**, 2480 (2017).

[26] A. Haridas, S. M. Perinchery, A. Shinde, O. Buchnev, and V. M. Murukeshan, Long working distance high resolution reflective sample imaging via structured embedded speckle illumination, Opt. Lasers Eng. **134**, 106296 (2020).

[27] F. Ferri, D. Magatti, L. Lugiato, and A. Gatti, Differential ghost imaging, Phys. Rev. Lett. **104**, 253603 (2010).

[28] A. Buades, B. Coll, and J.-M. Morel, A review of image denoising algorithms: With a new one, Multiscale Model. Simul. **4**, 490 (2005).

[29] A. Horé and D. Ziou, in *2010 20th International Conference on Pattern Recognition* (IEEE, Istanbul, Turkey, 2010), p. 2366.

[30] J. Benesty, J. Chen, Y. Huang, and I. Cohen, in *Noise Reduction in Speech Processing* (Springer, Berlin, Heidelberg, 2009), Vol. 2.

[31] M.-J. Sun, L.-T. Meng, M. P. Edgar, M. J. Padgett, and N. Radwell, A Russian dolls ordering of the Hadamard basis for compressive single-pixel imaging, Sci. Rep. **7**, 3464 (2017).

[32] X. Zhai, X. Wu, Y. Sun, J. Shi, and G. Zeng, Anti-noise computational imaging using unsupervised deep learning, Opt. Express **30**, 41884 (2022).

[33] J. Pawley, *Handbook of Biological Confocal Microscopy* (Springer Science & Business Media, New York, USA, 2006), Vol. 236.

[34] A. Krull, T. Vičar, M. Prakash, M. Lalit, and F. Jug, Probabilistic Noise2Void: Unsupervised content-aware denoising, Front. Comput. Sci. **2**, 5 (2020).

[35] M. O. Scully, M. S. Zubairy, and I. A. Walmsley, Quantum optics, Am. J. Phys. **67**, 7 (1999).

[36] B. Lim, S. Son, H. Kim, S. Nah, and K. Mu Lee, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (IEEE, Hawaii, 2017), p. 136.

Correction: The fourth author's name contained a typographical error and has been fixed.