OE Optical Engineering

REVIEW

Review: Advances in lensless random phase encoded imaging for automated cell identification

Gregory Aschenbrenner, Saurabh Goswami, Kashif Usmani, and Bahram Javidi*
University of Connecticut, Electrical and Computer Engineering Department, Storrs, Connecticut, United States

ABSTRACT. We present recent advances in lensless random phase encoded imaging for cell identification as related to biomedical applications, accentuating the robustness of methodologies through an examination of previously published works. Following an explanation of the foundational principles of lensless random phase encoded imaging for cell identification, discussions proceed on the evaluation of lateral resolution capabilities, current computational approaches, and an exploration of the practical impacts of these advancements in cell identification for biomedical applications. In particular, we seek to foster a deeper understanding of the robustness of key biosensor parameters in non-reconstructive lensless imaging for cell identification and its role in computational imaging and biomedical diagnostics. To the best of our knowledge, this is the first review of lensless random phase encoding for automated cell classification and its robustness to noise, and key bio sensor parameters. These systems can be made into low-cost, compact, and field-portable biosensors that are attractive for constrained healthcare systems.

 [DOI: 10.1117/1.OE.63.11.111814]

Keywords: lensless imaging; random phase encoding; speckle patterns; cell identification; compact imaging systems

Paper 20240505SSV received May 16, 2024; revised Aug. 4, 2024; accepted Sep. 30, 2024; published Nov. 1, 2024.

1 Introduction

In this review, we discuss previously published works in automated cell identification using lensless random phase encoding systems. Lensless random phase encoding is a class of imaging techniques based on performing imaging tasks directly on a captured intensity pattern produced by the interaction between the light field and some thin optical encoding element. In such methodologies, the conventional lens in an imaging system is replaced with a phase or amplitude mask, and appropriate algorithms are applied to classify the input object and improve the overall system performance in terms of cost, compactness, and other metrics. This approach enables the development of compact, cost-effective, and versatile imaging solutions as the removal of bulky lenses cuts down on the size and weight of an imaging system. References 1–29 presented a variety of lensless imaging approaches, including holographic systems. However, our focus in this paper is on non-holographic lensless imaging systems that replace the conventional lens with a phase-modulating element.

Non-reconstructive lensless imaging corresponds to one of the two primary approaches in lensless imaging, the other being reconstruction-based lensless imaging. Each approach's respective use depends on the practitioner's desired goals. In reconstructive lensless imaging, data are captured utilizing a chosen encoding element and are then digitally reconstructed to produce two-dimensional (2D) images of the scene or three-dimensional (3D) volume data.[1,2] Utilizing a reconstructive setup allows the practitioner to inspect a natural image or volume resultant from a reconstruction, allowing for further processing in an imaging pipeline.

*Address all correspondence to Bahram Javidi, bahram.javidi@uconn.edu

A typical reconstructive mask-modulated lensless imaging system[1,2] is given in Fig. 1, where an object wavefront is manipulated by a fixed optical mask and captured by an imaging sensor. The captured measurements are computationally reconstructed to yield an estimated object reconstruction.

Illumination-modulated lensless systems are also used for shadow imaging, holographic lensless imaging,[15–20] and time-resolved lensless imaging, but these approaches generally fall outside the scope of discussions of this paper.

In reconstruction-based approaches, an important consideration is the reconstruction algorithm used, with a focus on the computation time and the quality of the reconstruction. Several prior review papers focusing on reconstructive lensless imaging have been published,[1,2] which provided significant detail on the topic and applications of reconstructive lensless imaging. Such reviews enumerated the use case and the primary advantages of reconstructive lensless imaging. Specifically, lensless imagers afford size, cost and weight savings, increased field of view, inherent visual privacy, and the use of compressive sensing devices due to the spreading of light across the sensor.[1,2]

The primary motivation behind non-reconstructive lensless imaging is to benefit from the inherent properties of lensless imagers while avoiding the reconstruction problem entirely by directly performing an imaging or sensing task on the captured data, without a reconstructive step.[3–10] Such systems may include single random phase encoding (SRPE) and double random phase encoding (DRPE), which have been used for classification, and they may utilize pseudo-random phase masks as their encoding element. Such systems have been shown to have a lateral resolution, which is robust to pixel pitch, sensor size, and object-to-diffuser-to-sensor distance,[8] and whose classification performance is robust to noise, data obstruction,[7] and data compression.[5]

If one does not need spatial localization of a specific region in the object plane afforded by reconstructive approaches but requires a computationally cheap approach to classify a sample, SRPE and DRPE approaches provide the previously enumerated benefits, as well as the benefits inherent to lensless imaging detailed earlier. Notably, in some cases, reconstruction-based approaches can yield poor downstream performance in an imaging pipeline due to propagation of reconstruction errors,[9] leading non-reconstructive approaches to outperform them. A general system diagram for an SRPE system is shown in Fig. 2.

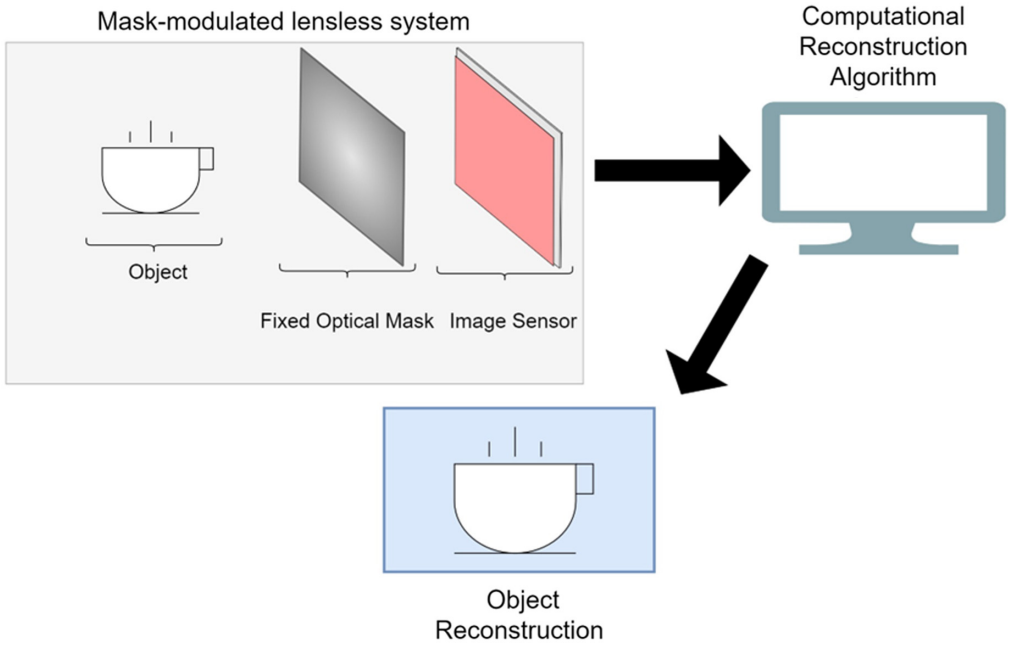

Fig. 1 Typical system diagram of a reconstructive mask-modulated lensless system.

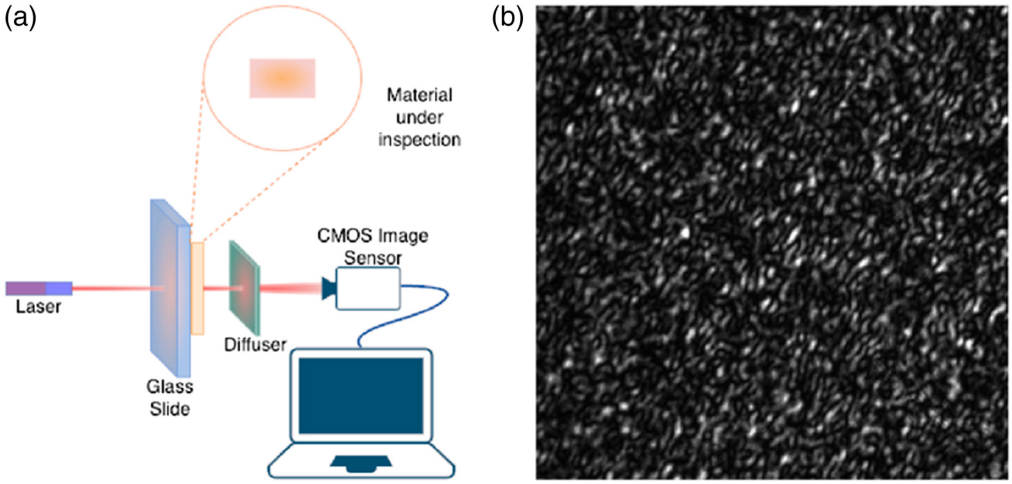

Fig. 2 (a) Schematic diagram of the SRPE imaging system and (b) a sample intensity pattern captured by the image sensor.[10]

Due to their portability, robustness, and accuracy, SRPE and DRPE systems are potential candidates for application to medical diagnostics, particularly for deployment in developing nations, where traditional biomedical imaging and diagnostics approaches may be cost-prohibitive, require specialized training, or are not designed for field use. As an example, a 3D-printed SRPE system used to identify sickle cell disease is presented in Fig. 3, measuring only 70 mm × 130 mm × 155 mm and weighing ∼155 g.[5] Even with these small dimensions, the system achieves 88.70% accuracy and 0.9622 area under the curve (AUC) while still being robust to data compression and obstruction.[5]

In this paper, we cover the theory and applications of SRPE and DRPE systems and briefly discuss where typical reconstructive approaches diverge from the non-reconstructive lensless imaging approach discussed. We progress to how machine learning can be applied to non-reconstructive lensless imaging and what considerations must be made in developing a machine learning pipeline. Finally, we conclude with the key benefits of the various approaches and with specific impacts on biomedical imaging. Applications to cell classification and disease identification are presented. To the best of our knowledge, this is the first review of lensless random phase encoding for automated cell classification and its robustness to noise, and key image sensor parameters. These systems can be made to be low-cost, compact, and field-portable biosensors, which are attractive for healthcare systems with constrained resources.

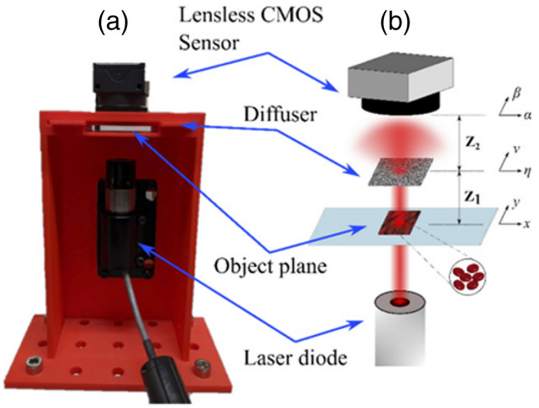

Fig. 3 (a) 3D-printed setup for SRPE system and (b) waveform propagation from the laser diode source to the object plane, to the diffuser, and then to the lensless CMOS detector.[7]

1.1 Theoretical Underpinnings and Basic Principles

The foundational theory behind lensless imaging draws upon wave optics and the modulation of an incoming wavefront by an encoding element. By manipulating the wavefront that interacts with an object, primarily done through an encoding element that encodes spatial information into resultant intensity patterns, the practitioner can subsequently process the captured intensity pattern pursuant to their end goal(s).

1.2 Non-reconstructive Lensless Imaging

Due to the lack of a reconstructive step, there is no need for an inverse model in non-reconstructive lensless imaging. A forward model is often utilized for the intensity captured at the sensor in an SRPE system based on the system's design.[3–8] SRPE systems use a pseudo-random phase mask, where the optical path length across a surface is randomly varied within a range to apply a pseudo-random phase factor to the incoming field. For the SRPE system in Fig. 2, the derivation for the intensity at the sensor is as follows:[5]

$$U_{\text{obj}}(x,y) = |A_{\text{obj}}(x,y)| \exp\{j\varphi_{\text{obj}}(x,y)\}, \tag{1}$$

where$U_{\text{obj}}(x,y)$ is the complex object waveform leaving the object plane with amplitude $A_{\text{obj}}(x,y)$ and phase $\varphi_{\text{obj}}(x,y)$. The complex waveform is propagated a distance z_1 under Fresnel conditions

$$U_{\text{diffuser}}(\eta,\nu) = [U_{\text{obj}}(x,y) * h_1(\eta,\nu)] \exp\{j\varphi_{\text{rand}}(\eta,\nu)\}, \tag{2}$$

where $U_{\text{diffuser}}(\eta,\nu)$ is the complex field leaving the diffuser, $\varphi_{\text{rand}}(\eta,\nu)$ is the random phase imposed by the diffuser, * denotes the convolution operation, and $h_1(\eta,\nu)$ is the Fresnel diffraction kernel as defined

$$h_1(\eta,\nu) = [\exp\{jkz_1\}/j\lambda z_1] \exp\{jk(\eta^2+\nu^2)/2z_1\}, \tag{3}$$

where k is the wavenumber and λ is the wavelength. Then, the complex waveform is recorded at the sensor at a distance z_2 away, in the Fraunhofer region, is given by

$$U_{\text{sensor}}(\alpha,\beta) = \frac{\exp\{jkz_2\}}{j\lambda z_2} \exp\{jk(\alpha^2+\beta^2)/2z_2\} F\{U_{\text{diffuser}}(\eta,\nu)\}|_{f_\eta=\frac{\alpha}{\lambda z_2}, f_\nu=\frac{\beta}{\lambda z_2}}, \tag{4}$$

where $U_{\text{sensor}}(\alpha,\beta)$ is the complex waveform incident on the image sensor, $F\{\cdot\}$ is the Fourier transform (F.T), and f_η and f_ν are the spatial frequencies associated with the F.T. The image sensor records the intensity or magnitude squared of the complex waveform:

$$I(\alpha,\beta) = |U_{\text{sensor}}(\alpha,\beta)|^2. \tag{5}$$

The captured intensity is referred to as a pseudo-random intensity pattern or in biological contexts an opto-biological signature (OBS).[4–6] The use of a pseudo-random phase mask spreads the intensity across the entire sensor as a white pseudorandom pattern.[3–8]

This is a simplified approximation and does not incorporate the dimensions of the diffuser or sensor or the physical process of sampling with a sensor. This approach also assumes Fresnel and Fraunhofer regimes to simplify propagation, but angular spectrum propagation becomes necessary for simulation purposes. Thus, although the prior approach correctly models the essential optical processes in the SRPE-specific setups tested, a more detailed model would include these sampling and propagation concepts to accurately represent the system's behavior in a variety of conditions.

In simulation-based experiments, such as those presented in Ref. 8 and covered in Sec. 2.1, these considerations become crucial. This involves accounting for the pixel dimensions, sensor array layout, and the finite resolution of both the diffuser and sensor. These factors are essential in capturing the nature of light propagation, to ensure the simulation results, align closely with experimental observations.

1.3 Reconstruction-Based Lensless Imaging, Forward, and Inverse Models

Here, we provide a brief review of reconstruction-based lensless imaging formulations to establish the basics of the reconstruction process. Detailed discussions of this approach are presented in the cited reconstruction-based lensless imaging reviews[1,2]:

$$y = Hx, \tag{6}$$

where H is a system-dependent matrix, x is the unknown scene intensity to be estimated, and y is the measurement on the sensor.[1,2] In such a formulation, the computational intensity of the problem becomes clear, assuming H is of dimensions $N \times N$; then, the computational complexity of the forward model is $O(N^2)$.[12] A convolutional model has been suggested, which reduces the computational complexity of from $O(N^2)$ to $O(N \log N)$[12] with the cost being if the object-to-mask distance is too large, or if the required field-of-view for a scene becomes too large, the shift-invariance assumptions made in the model break down.[12] Similarly, if the utilized mask can be thought of as separable, that is, the 2D mask can be represented as a cross product of two 1D functions, the computational complexity can be reduced to $O(N)$.[13]

Xiong[14] formulated the forward and inverse problem in the context of the Fourier Transform.[14] Although the computational complexity is not investigated, the formulation is entirely based on the fast Fourier transform, making the process fast with high regularity and convergence.[14] Due to the complexity and differing nature of the derivation, we refer the reader to the literature.

With these basic forward problem formulations, we now formulate the standard inverse problem:

$$\hat{x} = \arg\min_x \|y - f_h(x)\|_2 + \lambda R(x), \tag{7}$$

where $\hat{x}$ is the image estimate, y is the sensor measurements, $f_h(x)$ is the forward model, λ is a Lagrange multiplier, and $R(x)$ is a regularization function that imposes priors on the estimated image. If $f_h(x)$ is linear, then the optimization problem can be made convex with the appropriate choice of $\lambda R(x)$.[1,2,12] Initial solutions were centered on using iterative algorithms to solve convex problems,[11, 12] such as the alternating direction method of multipliers (ADMM),[24] or the fast shrinking iterative shrinking-thresholding algorithm (FISTA).[25]

A variation of ADMM that arises from the aforementioned Fourier formulation of the forward and inverse problem,[14] referred to as Fourier ADMM, is the key to the approach's speed and convergence properties, with the additional property of reducing the effects of diffraction caused by small holes in the mask, due to the conjugated structure of the algorithm.

In general, these solutions' reliance on the physical models of light transport used and inability to be directly statistically tuned to an application has given rise to data-driven techniques from the machine learning field to be utilized to improve reconstructions.[1,2,21,22]

2 Assessment of Lateral Resolution

Lateral resolution is the ability of an imaging system to distinguish between two points in an object and is a critical parameter that determines the level of detail and clarity in the produced images. In 2D reconstructive approaches, the output yields natural images; thus, traditional resolution metrics, such as the Rayleigh or Sparrow criteria,[25,26] can be directly utilized on the recovered reconstructions.

In the case of non-reconstructive lensless SRPE imaging where there is no natural image corresponding to a scene, other methods must be developed to analyze resolution capabilities. Section 2.1 explores the principles of lateral resolution specific to non-reconstructive lensless SRPE imaging systems.

2.1 Principles of Lateral Resolution in Coherent Non-Reconstructive Lensless Single Random Phase Encoding Imaging

Traditional metrics for evaluating resolution are less applicable in non-reconstructive lensless SRPE imaging systems where intensity patterns are pseudorandom and spread across the imaging sensor. A correlation-based criterion is more apt for determining lateral resolution as the captured pseudo-random intensity patterns at the sensor are not directly interpretable by humans.[8,25,26] This difference in system characteristics is illustrated in the comparison to the captured intensity patterns of an SRPE and lensed system seen in Fig. 4; in the lensed configuration, the two-point sources at the object plane produce the expected overlapping Airy pattern(s), whereas in the SRPE configuration, the output appears as a pseudorandom intensity spread across the entire sensor. A characteristic arises due to the random path length variations

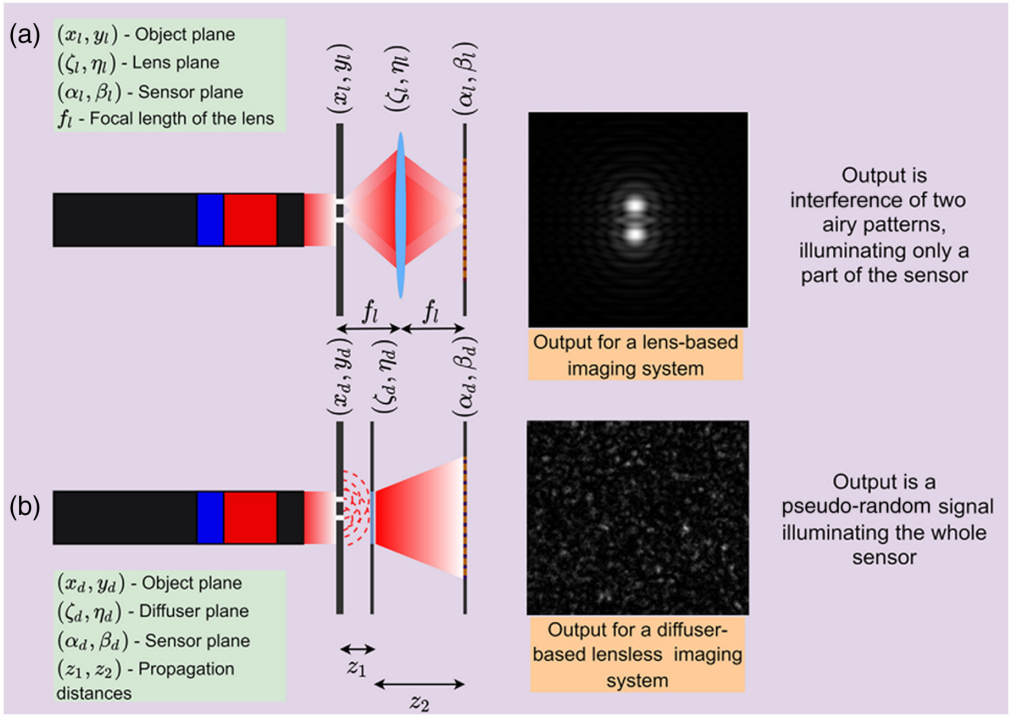

Fig. 4 Comparison of (a) lens-based imaging and (b) lensless SRPE imaging systems.[10]

utilized to implement a random phase diffuser. As mentioned earlier in Sec. 1.3, the intensity pattern captured at the sensor is a pseudo-random intensity pattern or in biological contexts an OBS.[5,7]

The approach in Ref. 8 leverages a simulation-based approach to delve into the capabilities of coherent SRPE systems in comparison to a similar lens-based system. The configuration(s) and captured intensity of the lens-based and SRPE systems are shown in Fig. 4, where a collection of two-point sources is used as input to each system. In Ref. 8, an actual laser and collection of point sources are not used as results are obtained through simulation. The SRPE simulation is based on the real parameters of an SRPE system of the form used in Refs. 5 and 7. To provide a comparison with the diffuser configuration, the lens has a focal length defined as

$$f_l = z_1 z_2/(z_1 + z_2). \tag{8}$$

Figure 5 illustrates the setup necessary to calculate the correlation criterion, where $\delta(x, y)$ represents an idealized point source in the object plane, $\delta(x - x_0, y)$ is a second shifted point

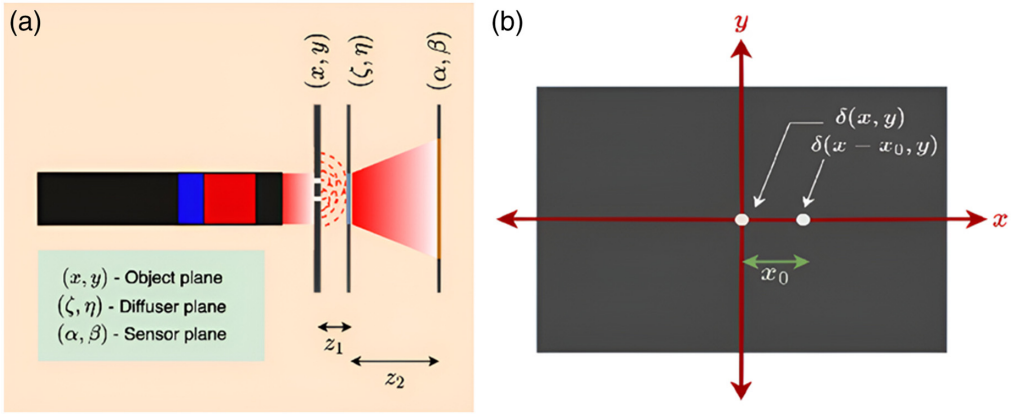

Fig. 5 Propagation of light due to two point-apertures at the object plane (a) to the output (image sensor) and placement of the point sources on the object plane (b).[10]

source, and x_0 is the lateral separation between the point sources in the object plane. The correlation criterion is calculated by first defining $x_0 = 0$ as the peak correlation, where the two point sources perfectly overlap; this corresponds to two overlapping point sources on the right side of Fig. 5. The distance between point sources, x_0, is then increased until a certain threshold of the peak correlation decrease is met, which is defined as the lateral resolution of the system under the proposed correlation criterion. The threshold of the peak correlation decrease is selected to be 35%, motivated by the equivalent lensed system used in comparisons having an Abbe diffraction limit corresponding to ~35% in the correlation criterion.[8]

The correlation plot for an SPRE system with parameters, wavelength$\lambda = 600$ nm, object to diffuser distance $z_1 = 3.6$ mm, diffuser to sensor distance $z_2 = 26.7$ mm, pixel size $p_\alpha = p_\beta =$ 0.6 m diffuser size $D_\zeta = D_\eta = 7.5$ mm and sensor size $S_\alpha = S_\beta = 7.5$ mm, is shown in Fig. 6. In this figure, the correlation drops below the ~35% threshold at 0.440 m for the base parameters tested. This is smaller than the pixel size of $p_\alpha = 0.6$ mwith subpixel resolution being possible due to the spreading of the input scene's information across the entire sensor by the diffuser as a white process [3-8], as opposed to focusing the light onto a particular pixel as in the lens-based imaging case.

2.2 Robustness of Lateral Resolution in Coherent Non-reconstructive Single Random Phase Encoding

Using the developed correlation criterion, the work in Ref. 8 discussed the lateral resolution of coherent lensless SRPE systems, specifically showing that the lateral resolution of these systems is robust to pixel pitch, sensor size, and object-to-diffuser-to-sensor distance. Each of the system parameters is tested for the system described in Fig. 5 and subsequently varies on the selected parameter and recalculating the correlation-based criterion as said parameters vary. These parameters over which resolution varies are visualized in Fig. 7. The parameters in which resolution remains relatively constant are visualized in Fig. 8.

Observing Fig. 7, the linear change in resolution one would associate with an increase in wavelength occurs. Similarly, Fig. 7 shows the distance between the object and the diffuser, and the physical size of the diffuser plays a significant role in the obtained resolution. This implies that in SRPE lensless systems, one should utilize larger diffusers whenever possible and keep the object-to-diffuser distance as small as possible to obtain the finest lateral resolution for a given SRPE configuration.

Similarly, observing Fig. 8, we see that the resolution is relatively constant with respect to diffuser-to-image sensor distance and robust to the pixels of the image sensor. In these cases, resolution varies on the order of nanometers as the number of pixels varies. With respect to pixel

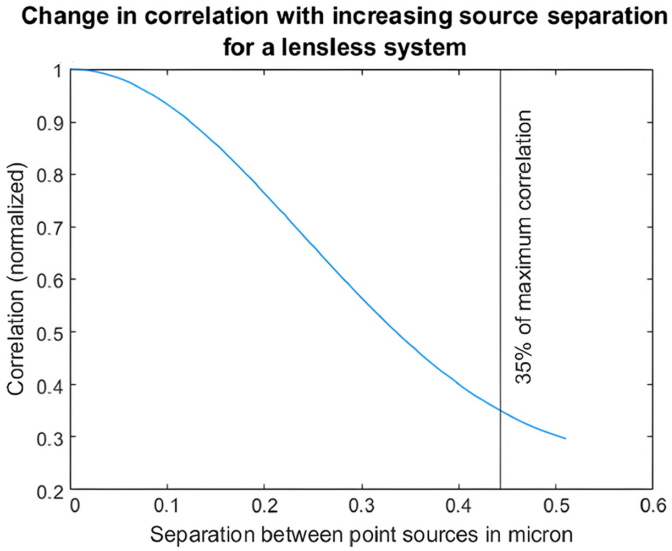

Fig. 6 Lateral resolution of lensless SRPE system estimated by change in correlation as a function of separation between two point sources (Fig. 5) for initial parameters of the simulation.[10]

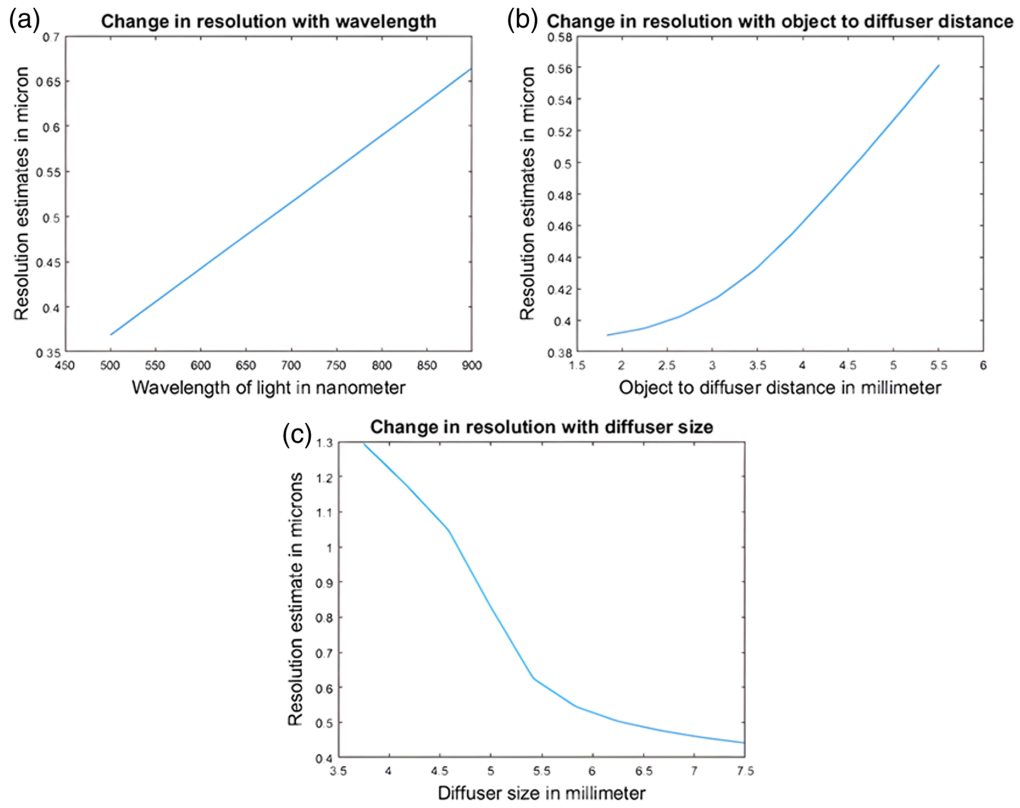

Fig. 7 Resolution estimate of the lensless SRPE system (see Fig. 5) as a function of (a) wavelength of the illuminating light, (b) distance between the object and the diffuser, and (c) size of the diffuser.[10]

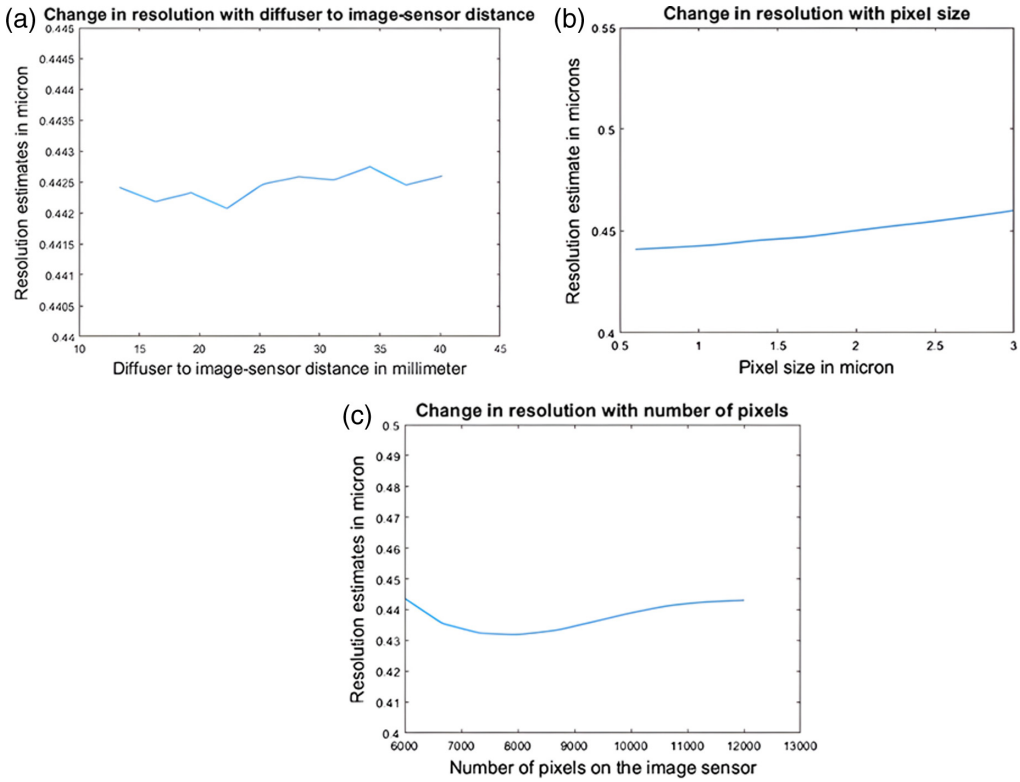

Fig. 8 Spatial resolution estimate of the lensless SRPE system (see Fig. 4) as a function of (a) distance from the diffuser to the image sensor, (b) size of the pixels on the image sensor, and (c) number of pixels on the image sensor.[10]

size, the resolution varies on the order of tens of nanometers. These variations can be regarded as statistical fluctuations in calculating the correlations.

Thus, when considering the lateral resolution in the system design of an SRPE system, the object-to-diffuser distance, z_1, and wavelength should be made as small as possible, whereas the diffuser should be made as large as possible. The effect of diffuser-to-sensor distance, z_2, pixel size, and number of pixels has less impact, thus allowing them to be chosen practically.

3 Computational Methods in Lensless Imaging

The convergence of machine learning with lensless imaging technologies has been pivotal for both reconstruction and non-reconstructive lensless imaging, allowing further enhancement of image reconstruction quality, resolution, and computation time,[1,2,21,22] as well as the facilitation of direct image classification without the necessity for traditional reconstruction.[3–10]

As described in Sec. 1.2, techniques such as the ADMM, Fourier-ADMM, or FISTA have been employed to solve the sparsity-constrained optimization problems inherent in recovering 2D/3D images from single-shot 2D captures.[1,2,12–14,24,25] These algorithms iteratively refine the reconstruction by imposing sparsity and non-negativity constraints, effectively enhancing the image quality and resolution.

Neural networks have been utilized to decrease reconstruction time and further improve reconstruction quality and resolution.[1,2,21,22] A primary drawback of supervised neural networks for used reconstruction is that they need ground truth for each training image, potentially complicating data collection setups. Exhaustive coverage of machine learning approaches used to improve the quality, resolution, and computation time is detailed within the previously mentioned review(s).[1,2]

3.1 Role of Machine Learning and Artificial Intelligence in Lensless Random Phase-Encoded Imaging

Neural networks similarly act as a natural step for use in direct classification. Convolutional neural networks (CNNs) replace earlier random forest (RF) classifier approaches as they have been shown to be more robust to noise.[4,6] Where RF classifiers use features extracted from data, CNNs are directly applied to the encoded patterns, directly extracting feature maps utilizing convolution and various other layer operations, and then classifying the underlying objects or features within the scene. By training CNNs on a dataset of known encoded patterns, this method bypasses the traditional reconstruction phase, facilitating real-time analysis and classification of images based on their unique optical signatures, and increases robustness to noise over RF classifiers. In such configurations, ground truth is not required; as there is no reconstructive step, only a label for the captured intensity pattern is necessary. Once data are captured as shown prior in Fig. 1, data can be processed in a SRPE system via CNN, as seen in Fig. 9.

To bolster the robustness of CNNs even further in these direct analysis tasks, especially under conditions where obstruction and noise can significantly degrade image quality, the application of preprocessing techniques becomes crucial. Among these, local binary pattern (LBP) has

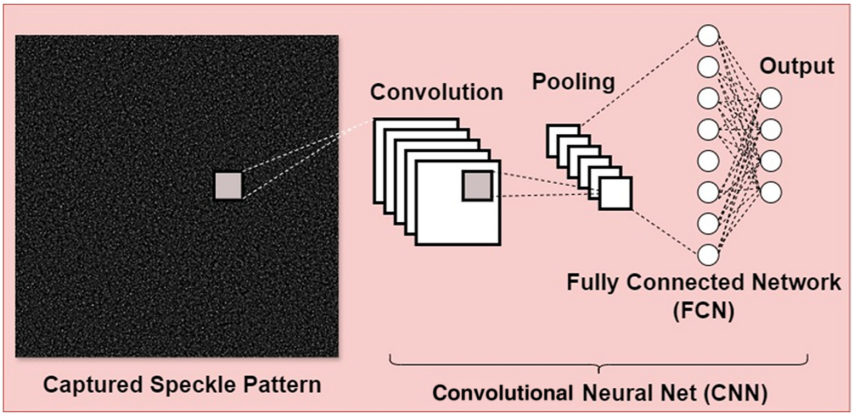

Fig. 9 Data processing for non-reconstructive lensless SRPE imaging.

emerged as a particularly effective method for enhancing the robustness of CNN-based classification systems against obstruction and data compression, as well as increasing general system performance.[5,11] The LBP operation consists of first defining a $N \times N$ neighborhood size about some pixel (x_0, y_0) and comparing every other pixel in the neighborhood according to

$$S(x_i, y_i) = \begin{cases} 1, & \text{if } I(x_i, y_i) \geq I(x_0, y_0) \\ 0, & \text{if } I(x_i, y_i) < I(x_0, y_0) \end{cases}. \tag{9}$$

The center pixel (x_0, y_0) is assigned a binary number constructed using $S(x_i, y_i)$, which is converted to decimal before the assignment

$$\mathrm{LBP}(x_0, y_0) = \sum_{i=0}^{N-1} S(x_i, y_i) * 2^i. \tag{10}$$

By applying this process to every pixel in the image, an LBP image is generated, where each pixel's value corresponds to its LBP. LBP is used to preprocess the OBS before they are fed into the CNN, emphasizing texture and structural information, a step taken to reduce the effects of noise. This preprocessing step proves heavily beneficial in cellular classification, where the fine details necessary for accurate classification can be obscured. A sample OBS and its calculated LBP are presented in Fig. 10, where the window size is selected to be 3×3.

The integration of LBP with CNNs for direct analysis in lensless imaging systems increases the robustness of the classification system to data obstruction and compression, ensuring that the predictive performance remains stable even in less-than-ideal imaging conditions[5,11] where significant portions of image data are lost. This combined approach streamlines the analysis process, allowing for real-time or near-real-time processing, which is essential for applications requiring rapid, onsite imaging processes.

A similar, but distinct approach to using CNNs for the direct classification of captured data is the use of vision information transformers (ViTs).[10,27–29] ViTs are a relatively new computer vision approach, which does not use convolution to derive feature maps, but attention. Attention is a mechanism to dynamically weigh the importance of different elements in an input sequence. Although the specifics of the architectures of ViTs are outside the scope of this review, the general concept revolves around dividing an image into fixed-size patches; each of which is linearly embedded into a vector. These vectors, along with positional encodings to retain spatial information, are then fed into a standard transformer encoder. The encoder processes these patch embeddings through a self-attention mechanism of some form depending on the architecture, with the aim to increase the model's global context and relationships across the entire image.[10,27–29]

A significant conceptual note about ViTs is that the attention mechanism does not have a built-in assumption of locality as in convolution. This results in a lack of strong local bias and thus acts as a significantly weaker inductive process.[10,27–29] This results in the required size of data necessary to achieve state-of-the-art CNN-like performance being exceptionally high, with datasets often being millions or billions of images to achieve state-of-the-art performance.[27–29]

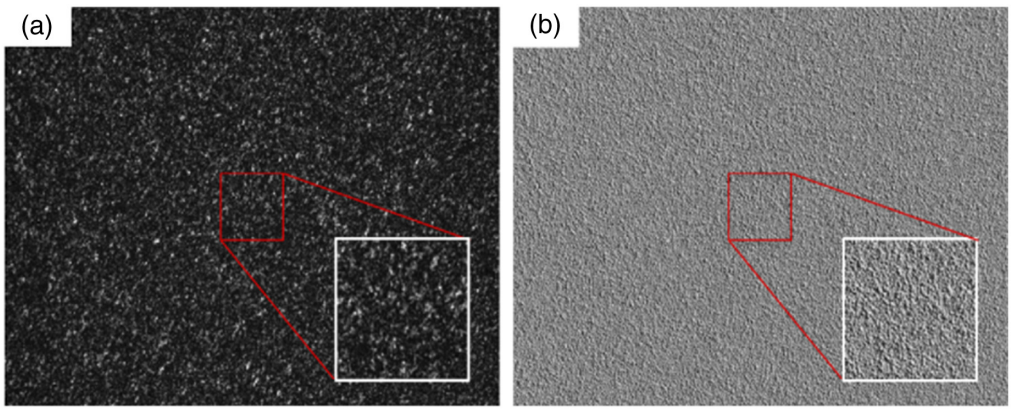

Fig. 10 (a) Raw OBS recorded at the image sensor of the SRPE system for a healthy red blood cell (RBC) and (b) generated LBP map of panel (a).[7]

The relatively large size of datasets is seen when a ViT is applied by Pan et al.[10] to the incoherent reconstruction-free lensless classification problem. The model was pretrained on ILSVRC-2012, with 1.3 million images and 1000 classes, and in two separate experiments trained on popular datasets MNIST, cats-vs-dogs (25,000 images), and separately on ∼8000 real images of fruit, respectfully. Pan reported that when trained on MNIST, with 60,000 train and 10,000 test images, an accuracy of 91.47% is obtained. When trained on cats-vs-dogs (12,500 of both cats and dogs, split 80% train, 20% test), Pan reported 94.26% accuracy and 0.9664 AUC.[10] We note that although these are standardized datasets and experiments are not biological in nature, the strong results obtained still exemplify the potential of non-reconstructive approaches.

3.2 Automated Cell Identification Using Lensless SRPE Imaging

The automated identification between cell types, i.e., between healthy and diseased cells or species types, using lensless SRPE imaging represents a significant breakthrough in biomedical diagnostics. These approaches initially leveraged statistical features extracted from the OBS for cell identification and then classified by an RF classifier.[4,6] Initial works achieve 97.9% accuracy when determining between cell types.[6] A particular limitation of RF classifiers identified in initial works is that they are not inherently robust to noise, motivating the use of CNNs and potential preprocessing of data with LBP.

When distinguishing between horse and cow red blood cells, an SRPE and a shearing digital holographic microscope[6,7] are compared, along with the proposed classification algorithm. In these experiments, AlexNet is used as the selected CNN architecture to be compared with RFs, all without LBP preprocessing. In the conducted experiments, the CNN is applied to an SRPE system and yields 88.99% accuracy and 0.9649 AUC, outperforming the other combinations of classification algorithms and optical methodologies tested.[7] In these experiments, the batch size is 24, the learning rate to 0.0006, and networks are trained for three epochs. Data collected in these experiments utilizing groups of red blood cells from each respective species are utilized, with a beam-splitting arm in the experimental setup to manually inspect slide content during calibration to ensure proper cell content.[7]

Further experiments were conducted in classifying sickle cell positive and negative red blood cells and compared the effect of LBP when utilizing CNN and RF classification algorithms.[5] Furthermore, the paper compares various CNN architectures (AlexNet, VGG19, SqueezeNet, ResNet50) and finds that AlexNet obtains the highest base performance metrics on the captured data. The CNN-based approach with LBP preprocessing yields 88.70% accuracy and 0.9622 AUC, outperforming the other tested methods. In these experiments, the networks were optimized using a grid search to select the best-performing mini-batch size (32, 64, 128, or 256), bias and weight learn rate factor (1, 5, or 10), and optimizer (Adam or SGDM); physical data consist of 3-μl wet mount slides from sickle cell positive and negative donors.

These advances address significant challenges in automated cell classification, including the need for compact and stable systems capable of distinguishing among various classes of microscopic objects without complex optics or bulky systems.

4 Conclusions

The progression from utilizing phase masks and diffractive elements to the integration of sophisticated computational approaches such as machine learning has not only enhanced the capabilities of lensless imaging systems but also broadened their applicability. By examining foundational principles, lateral resolution capabilities, and computational approaches, the review highlights the unique advantages and practical applications of non-reconstruction pseudo-random phase-encoded lensless imaging in the context of automated cell identification.

Particularly noteworthy is the use of direct classification and analysis of captured data via CNNs, which has opened new avenues in real-time imaging and diagnostics. The robustness of CNNs against noise and data obstruction underscores the advancements in making lensless imaging a viable, efficient solution for biomedical applications.

Automated disease identification, such as sickle cell disease, and the classification of red blood cells have demonstrated the clinical relevance of lensless SRPE imaging, showcasing its potential to significantly impact global healthcare, especially in underserved regions.

The compact, cost-effective nature of these systems, combined with their high accuracy and rapid processing capabilities, positions lensless imaging as a key technology in the democratization of healthcare diagnostics.

As we look forward, the continuous innovation in sensor technology, computational algorithms, and the increasing integration of artificial intelligence in lensless imaging promises to further enhance the resolution, speed, and versatility of these systems. The prospect of integrating lensless random phase encoding imaging technologies into portable diagnostic devices has the potential to benefit healthcare delivery, providing rapid, accurate diagnostics at the point of care.

Disclosures

The authors declare no conflicts of interest.

Code and Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed.

Acknowledgments

G.J. Aschenbrenner acknowledges support through the General Electric NextGen Scholar Fellowship. B. Javidi acknowledges support under the National Science Foundation (Grant No. 2141473) and the Office of Naval Research (ONR) (Grant Nos. N000142212375, N00014-22-1-2349, and AFOSR FA9550-21-1-0333).

References

1. V. Boominathan et al., "Recent advances in lensless imaging," *Optica* **9**(1) (2021).
2. S. Li et al., "Lensless camera: unraveling the breakthroughs and prospects," *Fundamental Research* 2024, .
3. A. Stern and B. Javidi, "Random projections imaging with extended space-bandwidth product," *IEEE J. Display Technol.* **3**(3), 315–320 (2007).
4. B. Javidi, A. Markman, and S. Rawat, "Automatic multicell identification using a compact lensless single and double random phase encoding system." *Appl. Opt.* **57**, B190–B196 (2018).
5. P. M. Douglass, T. O'Connor, and B. Javidi, "Automated sickle cell disease identification in human red blood cells using a lensless single random phase encoding biosensor and convolutional neural networks," *Opt. Express* **30**, 35965–35977 (2022).
6. B. Javidi et al., "Cell identification using single beam lensless imaging with pseudo-random phase encoding," *Opt. Lett.* **41**, 3663–3666 (2016).
7. T. O'Connor et al., "Red blood cell classification in lensless single random phase encoding using convolutional neural networks," *Opt. Express* **28**, 33504–33515 (2020).
8. S. Goswami et al., "Assessment of lateral resolution of single random phase encoded lensless imaging systems," *Opt. Express* **31**(7), 11213–11226 (2023).
9. Y. Zhang et al., "Hand gestures recognition in videos taken with a lensless camera," *Opt. Express* **30**, 39520–39533 (2022).
10. X. Pan et al., "Incoherent non-reconstructive object recognition with mask-based lensless optics and the transformer," *Opt. Express* **29**, 37962–37978 (2021).
11. X. Pan et al., "Lensless inference camera: incoherent object recognition through a thin mask with LBP map generation," *Opt. Express* **29**, 9758–9771 (2021).
12. N. Antipa et al., "DiffuserCam: lensless single-exposure 3D imaging," *Optica* **5** (2018).
13. M. J. DeWeert and B. P. Farm, "Lensless coded-aperture imaging with separable Doubly-Toeplitz masks," *Opt. Eng.* **54**(2), 023102 (2015).
14. Z. Xiong, "Super-resolution lensless imaging system based 509 on a fast anti-diffraction algorithm," *Opt. Express* **31**(23), 37395–37407 (2023).
15. B. Javidi et al., "Roadmap on digital holography [Invited]," *Opt. Express* **29**, 35078–35118 (2021).
16. T. O'Connor et al., "Digital holographic deep learning of red blood cells for field-portable, rapid COVID-19 screening," *Opt. Lett.* **46**, 2344–2347 (2021).
17. A. Anand, V. Chhaniwal, and B. Javidi, "Tutorial: common path self-referencing digital holographic microscopy," *Appl. Phys. Lett. Photon.* **3**, 071101 (2018).
18. M. K. Kim, "Principles and techniques of digital holographic microscopy," *SPIE Rev.* **1**, 8005 (2010).
19. U. Schnars et al., *Digital Holography*, Springer, pp. 39–68, 2015.
20. S. Utadiya et al., "Integrated self-referencing single shot digital holographic microscope and optical tweezer," *Nat. Light: Adv. Manuf.* **3**, 37 (2022).

21. J. Wu, L. Cao, and G. Barbastathis, "DNN-FZA camera: a deep learning approach toward broadband FZA lensless imaging," *Opt. Lett.* **46**, 130–133 (2021).
22. M. Haris, G. Shakhnarovich, and N. Ukita, "Deep back-projection networks for super-resolution," in *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 1664–1673, 2018.
23. B. Stephen et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.* **3**, 1–122 (2011).
24. A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.* **2**, 183–202 (2009).
25. M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th ed., Cambridge University Press, 1999.
26. C. M. Sparrow, "On spectroscopic resolving power," *Astrophys. J.* **44**, 76 (1916).
27. S. Srivastava and G. Sharma, "OmniVec: learning robust representations with cross modal sharing," in *IEEE/CVF Winter Conf. Appl. Comput. Vision (WACV)*, Waikoloa, Hawaii, USA, pp. 1225–1237, 2024.
28. J. Yu et al., "CoCa: contrastive captioners are image-text foundation models," arXiv e-prints (2022).
29. M. Wortsman, "Model soups: averaging weights of multiple fine-tuned models improves accuracy without increasing inference time," arXiv e-prints (2022).

Gregory Aschenbrenner is a PhD student at the University of Connecticut pursuing a degree in electronics, photonics, and biophotonics under Dr. Bahram Javidi; he received his BS degree in applied mathematics and his BSE (Honors Laureate) in electrical engineering from the University of Connecticut in 2023.

Saurabh Goswami is a PhD candidate at the University of Connecticut pursuing a degree in electronics, photonics, and biophotonics under Dr. Bahram Javidi; he received his BSE in electrical engineering in 2018 and his MS degree in image processing and computer vision in 2021 from IIT Madras.

Kashif Usmani is a PhD candidate at the University of Connecticut; he received his BS degree in physics from Shibli National College in 2016, his MS degree in optics/optical sciences from IIT Delhi in 2018, and his PhD in electronics, photonics, and biophotonics from the University of Connecticut in 2024.

Bahram Javidi is a Board of Trustees Distinguished Professor at the University of Connecticut. He received his PhD in electrical and electronics engineering from Penn State. He has more than 1100 publications, including 9 books, 58 book chapters, 540+ peer-reviewed journal articles, and 520+ conference proceedings. He has been named Fellow of the Institute of Electrical and Electronics Engineers (IEEE), Fellow of the American Institute for Medical and Biological Engineering (AIMBE), Fellow of the Optical Society of America (OSA), Fellow of the European Optical Society (EOS), Fellow of SPIE, Fellow of the Institute of Physics (IoP), Fellow of The Society for Imaging Science and Technology (IS&T), and Fellow of The Institution of Electrical Engineers (IEE). He has received the Fellow award by the John Simon Guggenheim Foundation, the Institute of Electrical and Electronics Engineers (IEEE) Donald G. Fink Prize Paper Award, The George Washington University's Distinguished Alumni Scholar Award, and the Humboldt and the Technology Achievement Award from SPIE.