

Compute-First Optical Detection for Noise-Resilient Visual Perception

Jungmin Kim,* Nanfang Yu, and Zongfu Yu



Cite This: *ACS Photonics* 2025, 12, 1137–1145



Read Online

ACCESS |

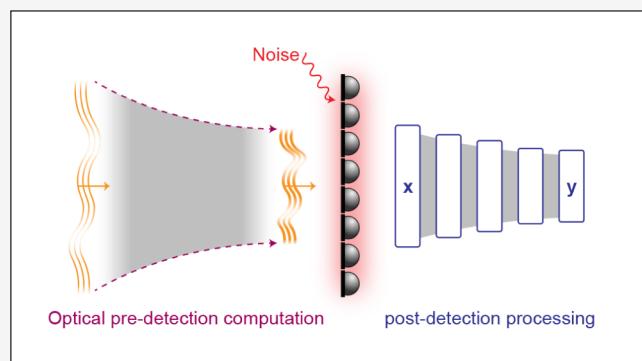
Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: During machine visual perception, the optical signal from a scene is transferred into the electronic domain by detectors in the form of image data, which are then processed for the extraction of visual information. In noisy environments, such as a thermal imaging system, however, the neural performance faces a significant bottleneck due to the inherent degradation of data quality upon noisy detection. Here, we propose a concept of optical signal processing before detection to address this issue. We demonstrate that spatially redistributing optical signals through a properly designed linear transformer can enhance the detection noise resilience of visual perception, as benchmarked with MNIST classification. A quantitative analysis of the relationship between signal concentration and noise robustness supports our idea with its practical implementation in an incoherent imaging system. This compute-first detection scheme can advance infrared machine vision technologies for industrial and defense applications.

KEYWORDS: *Optical processing, linear transformation, signal-to-noise ratio, metasurface, optical neural network, visual perception*



INTRODUCTION

Recent advances in infrared (IR) technologies around atmospheric windows have expedited various scientific and industrial fields, including night vision technologies based on thermal imaging^{1–4} and radiative cooling systems addressing the global climate crisis,^{2,5–8} which use infrared light as an information and heat carrier, respectively. These technologies commonly leverage transmission within the mid-IR regime, relying on blackbody radiation^{1,5} emitted from an object at around room temperature without external sources. However, the relatively weak IR power, compared with that of the daytime ambient light, has posed a challenge: the low signal-to-noise ratio (SNR) in thermal imaging in the presence of detection noise. Several studies have focused on the postprocessing of noisy images to overcome the low SNR issue by incorporating additional degrees of freedom, such as hyperspectral^{3,9–13} or polarimetric^{14–16} information, which involved developing apparatus for the fast acquisition and processing of large data sets.

Regarding visual perceptions¹⁷ such as object recognition and feature detection from a noisy environment, plenty of additional computing mechanisms in the optical domain based on diffractive,^{18–27} interferometric^{27–32} and/or optoelectronic analog^{33–36} devices can be applied to resolve this issue. The core idea is to obtain cleaner data rather than relying on digital postprocessing of noisy data. This is inspired by the Fourier-transform infrared (FTIR) spectrometer in comparison with a

grating-based monochromator.³⁷ Unlike the monochromator, which uses a diffraction grating to spatially separate each spectral component, the FTIR interferometer encodes the spectral information into a temporal pattern with a higher SNR, which is then decoded via a Fourier transformation.

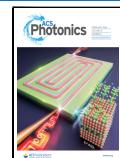
Traditionally, machine visual perception begins by capturing a scene with imaging devices. As displayed in Figure 1a, the wave signal of the image is then transferred to an electronic domain by “detection” with a photodetector (PD) array, and useful visual information (i.e., the feature) of the scene is extracted from the acquired image data through a series of data processing procedures. However, this detection-computing sequence places its computational load fully behind the detection, resulting in inherent vulnerability to noise, such as the thermally generated dark current in PDs.¹ In this work, we leverage additional optical computing mechanisms or resources to address this issue, as depicted in Figure 1b: enhancing SNR with a predetection optical processing unit (OPU) that is capable of concentrating the optical signal without loss of information. To validate the idea, we demonstrate a theoretical

Received: November 18, 2024

Revised: January 22, 2025

Accepted: January 23, 2025

Published: February 5, 2025



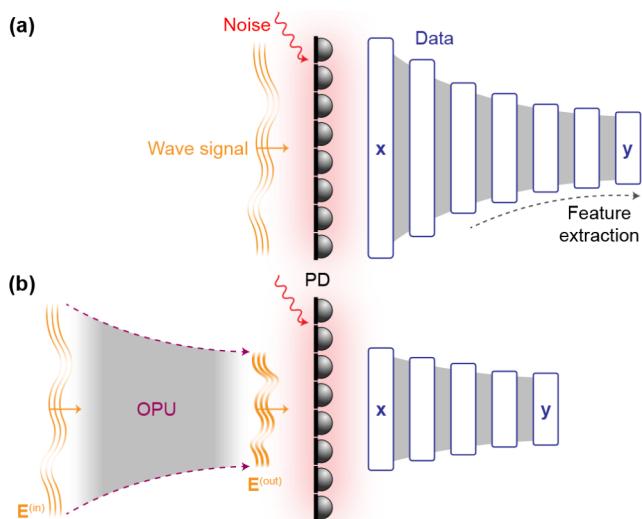


Figure 1. Concept of optical compute-first detection system for visual perception. (a) Conventional procedure: the wave signal from a scene is converted to image data by a photodetector (PD) array with additional detection noise. Subsequently, a digital processor processes the image data, extracting a latent feature of the scene. (b) Proposed scheme: the wave signal undergoes primary modulation ahead of detection through an optical processing unit (OPU). It is detected and then postprocessed in the digital domain to produce the final visual information. $E^{(in,out)}$, input and output state of waves; x , detected value in electronic domain; y , target feature.

framework where a properly designed optical neural network is integrated with a digital MNIST³⁸ classifier, revealing enhanced performance in terms of the resilience of classification accuracy against extreme dark noise. Quantitative evidence is then provided to establish the relationship between the robustness and the degree of concentrative modulation, along with the concept of detection pruning. Finally, we demonstrate an incoherent imaging system as a practical example, verifying the superior robustness against noise with the data-driven design of a metalens system.

RESULTS AND DISCUSSION

Model Definition. In the traditional approach to visual perception tasks, one needs to obtain the image of a scene, which is regarded as a tailored copy of a scene at a detection plane by imaging devices. As displayed in Figure 1a, the replicated wave signal is then transferred to a digital domain by “detection” with a photodetector array, and then the acquired image data results in the perception of the scene through neuromorphic image processing. However, this imaging-processing sequence places its computational load fully after detection, leading to inherent noise issues. Our objective is to leverage additional optical computing resources before detection, as depicted in Figure 1b, to enhance the SNR.

Despite the significance of nonlinearity in advanced neuromorphic devices,^{24,39} we exclusively focus on optical linear transformations to demonstrate how redistributing conserved total energy can be beneficial to signal preprocessing techniques. Based on the superposition principle for the electric field, therefore, we can streamline such linear optical devices through discretization: $E^{(out)} = f(E^{(in)}; P) = PE^{(in)}$, where the input and output vectors $E^{(in,out)} \in \mathbb{C}^{1 \times N}$ have finite spatial dimensions N . Assuming that the total energy is conserved during optical signal processing: $\langle E^{(in)} | E^{(in)} \rangle =$

$\langle E^{(out)} | E^{(out)} \rangle$, the transfer matrix P should be unitary, i.e., $P^\dagger P = I$. We note that this discrete and unitary constraint aligns with the solution of the coupled-mode equation for a waveguide system.⁴⁰ Importantly, any arbitrary unitary operation can be programmed using Mach–Zehnder interferometers and phase shifters with the same degrees of freedom (N^2). Well-known Clements⁴¹ and Reck⁴² designs serve as effective tools for achieving this programmability. Hence, a discrete unitary system emerges as an effective testbed for the analysis and demonstration of predetection optical processing.

As a representative task for machine visual perception, we benchmark the MNIST classification performance³⁸ using two cascaded networks: a deep neural network as a digital processor attached to the linear OPU, as illustrated in Figure 1b. The digital network, $y = g(x; Q)$, with a trainable parameter set Q , performs postprocessing of the optical intensity signal $x \in \mathbb{R}^{1 \times N}$ to generate an inferred probability distribution $y \in \mathbb{R}^{1 \times M}$ over M classes. Specifically, our target task is the classification of MNIST objects with 28×28 resolution; therefore, we have $N = 28^2$ and $M = 10$.

Meanwhile, there is assumed to be a physical detection process (i.e., a transition from an optical to an electronic signal) over PDs¹ between the two domains. The PD array typically measures the photon counts incident to each pixel, which is a function of the output intensity vector $x = h(E^{(out)})$ with element-wise operations:

$$x_\alpha = \left| E_\alpha^{(out)} \right|^2 + \Delta I_{\text{photon}} + \Delta I_{\text{dark}} \quad (1)$$

where $\alpha = 0, \dots, N - 1$ is the pixel index, and

$$\Delta I_{\text{photon}} \sim \frac{\text{Pois}\left(\Delta t \left| E_\alpha^{(out)} \right|^2\right)}{\Delta t} - \left| E_\alpha^{(out)} \right|^2 \quad (2)$$

$$\Delta I_{\text{dark}} \sim N(0, \sigma_{\text{dark}}^2) \quad (3)$$

represent two independent noise mechanisms typically involved in the optoelectronic detection: ΔI_{photon} is photon shot noise, a Poisson random process arising from the discrete nature of photons arriving at each detector within a time frame Δt ; ΔI_{dark} accounts for all other input-independent noises such as thermal and dark current noise, approximated by a Gaussian process with effective noise power σ_{dark} .

Noise Resilience of Compute-First Detection Scheme.

To demonstrate the robustness against detection noise achieved by the optical preprocessing, we investigate two different types of linear OPUs (i.e., designed P). First, P can be trained as part of the tandem optical-digital network $y = (g \circ h \circ f)(E^{(in)}; P, Q)$ using deep learning^{18–20,43} or adjoint-based optimization.^{39,44,45} We note, however, that gradient descent may converge to different local optima depending on the initial setup. Otherwise, P can be assigned a manually defined unitary matrix that is likely to concentrate the optical signal, as shown in ref. 46. On the contrary, we can set P as the identity matrix for the reference model, representing ideal imaging devices without proper optical treatment. For all cases, we optimize the digital network g through supervised learning, employing cross-entropy loss

$$L(y, \hat{y}) = - \sum_{m=0}^{M-1} y_m \log \hat{y}_m \quad (4)$$

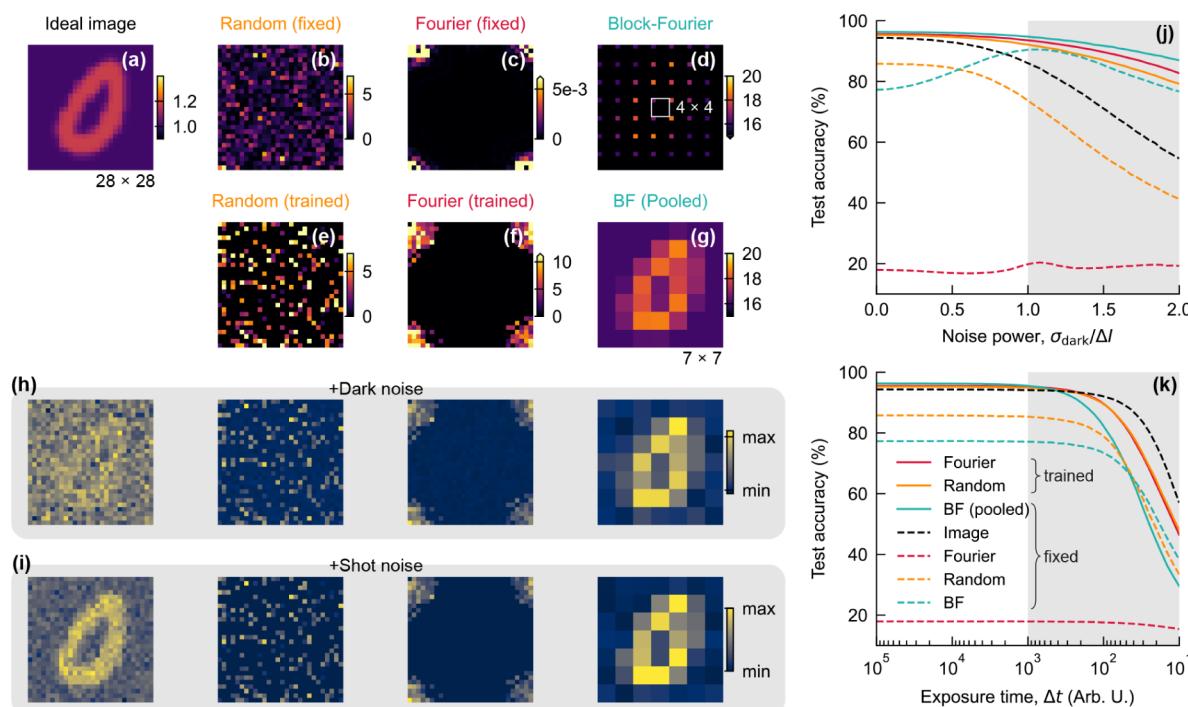


Figure 2. Noise robustness achieved by optical signal processing. (a–g) 2D representations of optical intensities before detection, $I_\alpha^{(\text{out})}$: ideal image of digit 0 (a), random matrix multiplied image (b), 2D Fourier image (c), block-wise 2D Fourier image (d), images with machine-trained unitary matrices [(e) and (f)] from the initialization with (b) and (c), respectively, and sampled (7^2 pixels) image from (d) by max-pooling (g). (h,i) Detected images with two different types of noise $x_\alpha = I_\alpha^{(\text{out})} + \Delta I_{\text{dark}} + \Delta I_{\text{photon}}$: dark noise [(h), $\Delta I_{\text{photon}} \sim 0$] and photon shot noise [(i), $\Delta I_{\text{dark}} = 0$], applied to (a), (e), (f), and (g) from left to right. (j,k) MNIST classification accuracies according to increasing test noise levels: dark noise power (j) and shot exposure time (k) for various optical preprocessing types (ideal image, black; fixed random/Fourier/block-Fourier operations, dashed orange/red/green; trained random/Fourier and pooled block-Fourier operations, solid orange/red/green). Gray regions indicate where test noises are stronger than training noises. The test accuracy is calculated over 10^4 balanced test samples with 20 repetitions. $\Delta I \sim 0.17$ is the intensity contrast in ideal images (a).

for a one-hot encoded class label \hat{y} . Further details on the model architectures and training process are available in Table S1 and Figure S1.

Figure 2a, for instance, shows the coherent input intensity distribution (and the identical output image for the reference model) $I_\alpha^{(\text{in})} \equiv |E_\alpha^{(\text{in})}|^2$ of Class 0, 2D-reshaped into 28^2 pixels. This input can be processed by a randomly generated unitary matrix P_R or the 2D discrete Fourier-transform (DFT) matrix P_F , resulting in the output intensity distributions $I_\alpha^{(\text{out})} = |E_\alpha^{(\text{out})}|^2$ (Figure 2b,c). Alternatively, a manually designed block-wise Fourier matrix $P_{BF,7}$ inspired by the microlens array structure⁴⁷ and dividing the domain into 7-by-7 blocks for local DFT (see Methods section for definition), can focus the output intensity into several representative pixels (Figure 2d; the white border indicates one of the square blocks).

Since the results in Figure 2b,c are not yet optimized, we further train $P_{R,F}$ through deep learning to $\tilde{P}_{R,F}$ resulting in the output intensity distributions optimal for the following neural inference as depicted in Figure 2e,f, respectively. Simultaneously, we apply max-pooling, i.e., dimensionality reduction by taking the maximum value for each subdomain, to the block-wise Fourier result (Figure 2g) to transfer only the DC component, i.e., maximum-intensity pixel per block, to the subsequent inference g. Considering that most optical energy is concentrated into the DC pixel through Fourier transforms in coherent systems with a constant phase, the max-pooling layer can be implemented by simply choosing detectors for the designated DC pixels. Both the machine-optimized (Figure

2e,f) and manually defined (Figure 2g) OPUs effectively concentrate the input signal distribution (Figure 2a) to boost the intensity contrast by an order of magnitude up to 10^1 . This is a marked increase compared to the input intensity contrast $\Delta I \equiv \max_\alpha(I_\alpha^{(\text{in})}) - \min_\alpha(I_\alpha^{(\text{in})}) \sim 0.17$, which reflects black-body radiation contrast within a temperature range from 300 to 310 K for a LWIR wavelength ($10 \mu\text{m}$). We also note that block-Fourier operations, while no longer unitary, still remain as a type of optical linear operation, as long as the position of the most intense pixel in each block is well-defined: here fixed to the DC pixel for the coherent input of constant phase.

Applying the dark noise ΔI_{dark} and the photon shot noise ΔI_{photon} independently upon optical-to-electronic transition, we can observe the capability of such optical treatments in compensating for unavoidable dark noise. For instance, the left to right inset of Figure 2h displays the detected signals x with dark noise of noise power $\sigma_{\text{dark}} = \Delta I$ for the ideal image (Figure 2a) and OPU output (Figure 2e–g), respectively. The reference result (left, Figure 2h) is almost masked by the strong dark noise, making it challenging to identify the digit “0.” In sharp contrast, the block-wise Fourier result (right, Figure 2h) can be interpreted as digit 0, despite its coarse mosaic effect due to the magnified output intensity contrast.

The distinct difference between dark-noise-screened signals can be analyzed through a quantitative MNIST benchmark. First, we train the combined optical-digital networks with a fixed degree of noise levels: $\sigma_{\text{dark}}^{(\text{tr})} = \Delta I$ and $\Delta t^{(\text{tr})} = 10^3$. Then,

we tested the trained networks with an increasing dark noise level σ_{dark} from zero, as shown in Figure 2j. The green solid lines for the manually designed block-Fourier matrices with 7 segments ($P_{\text{BF},7}$) exhibit extreme robustness against dark noise up to $\sigma_{\text{dark}} \sim 2\Delta I$. In contrast, the test accuracy for the ideal image without optical processing (black line) rapidly decreases with dark noise. The machine-optimized models with different initialization (\tilde{P}_F and \tilde{P}_R ; red and orange solid lines) also outperform the reference model. Notably, it is essential to further optimize the linear operations by such pooling or training to maximize the noise robustness, considering that raw-matrix results (colored dashed lines) do not outperform the untreated image (black dashed line), which parallels the effectiveness of “data-driven” binning in comparison to the conventional one (e.g., untreated block-Fourier) of image subpixels as demonstrated by Mennel et al.⁴⁸

Figure S2 also presents various results from different training and network parameter settings in comparison to Figure 2. It implies two key points: (1) The poor performance observed with the fixed Fourier transform (red dashed lines) is due to the extremely weak AC signals from nonsparse input (i.e., nonzero background signal, see Figure S2c); (2) with proper optical preprocessing—no matter whether it is machine-trained or manually defined—most pixels become redundant and even worsen inference performances due to overfitting to detection noise (see green dashed lines in Figure S2e compared to Figures 2j and S2g), which supports the necessity of posing signal processing ahead of noise.

Interestingly, the linear OPUs are not effective for photon shot noise in enhancing SNR, as the absolute noise power of the shot noise is simultaneously amplified when the signals are concentrated, as $\Delta I_{\text{photon}} \propto [I_a^{(\text{out})}]^{1/2}$. That is, the shot noise is not mitigated by redistributing signals as effectively as dark noise but is more related to the total computing energy per operation, as in ref. 29. This difference is evidenced by the noisy images in Figure 2i (ideal image on the left, machine-trained in the middle, and block-Fourier on the right), as well as by Figure 2k illustrating no superiority in noise robustness of various models with decreasing exposure time Δt in log scale.

Mutual Relationship Between Robustness and Concentration of Signals. For a deeper insight into the quantitative relationship between optical preprocessing and the immunity of visual inference to detection noise, we explore two scenarios of training networks and the corresponding evaluation methods. First, we investigate the influence of a predefined degree of concentration on the system’s resilience against noise during test inference. Second, we reciprocally assess the impact of training noise during the optimization process on the resulting optical network’s signal condensation.

For the first approach, we examine the block-wise Fourier transform of images with various numbers of segmentation, N_{seg} . Given that the original image consists of 28^2 pixels, we can consider a uniform segmentation along the width and height of the image with $N_{\text{seg}} = 1, 2, 4, 7$, and 14, which are all divisors of 28. Otherwise, a nonuniform segmentation is explored as well, for instance, $N_{\text{seg}} = 10$ for dividing 28 into 8 segments of width 3 and 2 segments of width 2 ($28 = 8 \times 3 + 2 \times 2$). Figures 3a,b, respectively, depict several intensity distributions $I_a^{(\text{out})}$ for block-wise Fourier operation $P_{\text{BF},N_{\text{seg}}}$ and the corresponding max-pooled images for $N_{\text{seg}} = 2, 4, 7, 10$, and 13. As N_{seg} decreases, notably, the output image becomes more compressive with the dimension reduced to

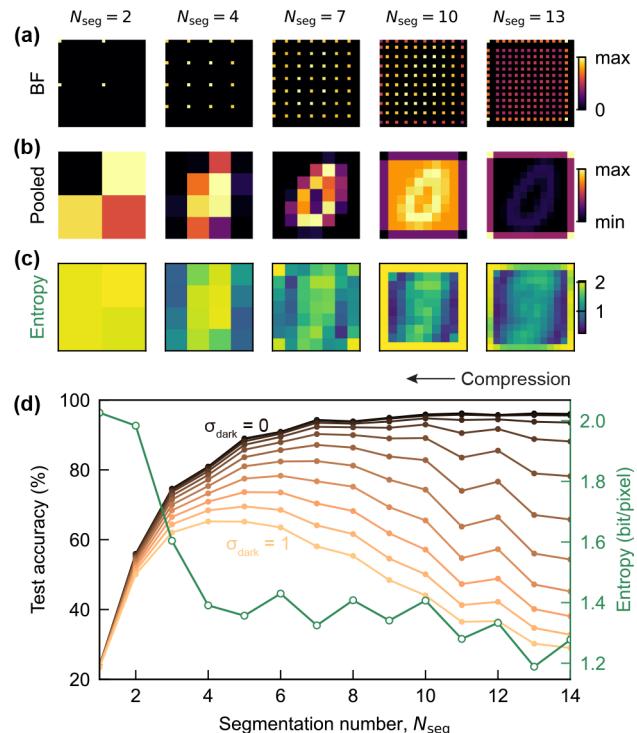


Figure 3. Concentration-induced noise robustness. (a,b) Output intensity distributions from an input example in class 0 after applying block-wise Fourier operations (a) and then max-pooling (b), with different segmentation numbers $N_{\text{seg}} = 2$ (left) to 13 (right). (c) Shannon entropy distributions given a data set and the operation with different N_{seg} . (d) MNIST classification accuracies as a function of N_{seg} with different test noise levels, from $I_{\text{dark}} = 0$ (black line) to $I_{\text{dark}} = 1$ (orange line). The average entropy per pixel is overlaid.

N_{seg}^2 pixels. Figure 3d validates the noise robustness achieved through compression by presenting the classification accuracy as a function of N_{seg} and the test noise power σ_{dark} . As anticipated, more compressive processing with a larger N_{seg} leads to more robust classification accuracy. This is evidenced in the narrowing gap between the results for zero ($\sigma_{\text{dark}} = 0$, black) and high test noise levels ($\sigma_{\text{dark}} = 1$, yellow). However, it is noted that the ideal accuracy for zero test noise (black line) itself decreases due to the information loss caused by overcompression.

Notably, the way the image is segmented also impacts the overall task performance. That is, how much information each focused pixel carries depends on its block size and location, and it does not always follow a monotonic trend with N_{seg} —instead, it can fluctuate. To quantify this, we calculate the Shannon entropy⁴⁹ for each pixel as a measure of information:

$$H_\alpha \equiv - \int_{-\infty}^{\infty} dJ p_\alpha(J) \log_2 p_\alpha(J) \quad (5)$$

where $J_\alpha = [I_\alpha - \text{mean}(I_\alpha)]/\text{Var}(I_\alpha)^{1/2}$ represents the batch-normalized intensity of pixel α over the given validation set, and p_α is the probability distribution function for J_α . For example, if a pixel consistently produces a single output intensity regardless of the input class, $H_\alpha = 0$. On the contrary, an ideal pixel perfectly classifying into ten different output values depending on the input class has $H_\alpha \sim 3.3$ bits of information. The more compressive the segmentation is, the broader the region the max-pooled pixel covers, capturing

more information. However, pixel-wise entropy fluctuates in practice as indicated by the color variation in Figure 3c and the green line in Figure 3d on average. This fluctuation is likely to impact the total information transferred to the digital network, thereby affecting the overall noise performance, especially for $N_{\text{seg}} \geq 6$.

In the opposite direction, dark noise can induce a general optical linear transformation to be trained in a more compressive manner. In other words, when strong dark noise is applied during training, the output intensity distribution is more likely to be focused on fewer pixels with a high SNR based on the select-and-concentrate strategy. Figure 4a–c illustrates the entropy distribution over pixels trained from the same random initialization $P_R \in U(28^2)$ but with different training noise powers $\sigma_{\text{dark}}^{(\text{tr})}$ applied during the optimization process. While almost zero noise (Figure 4a) results in the equitable optimization of all pixels in terms of the degree of information contained, as represented by mostly flat yellow colors, strong noise (Figure 4c) leads to differential optimization over pixels, separating more (yellow) and less (navy) informative pixels.

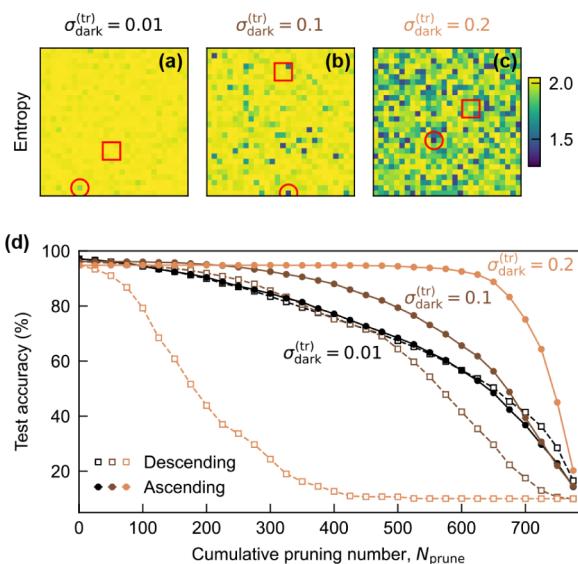


Figure 4. Training noise-induced emergence of hub detectors. (a–c), Shannon entropy distributions for trained $U(28^2)$ operations with the same random initialization but different training noise levels $\sigma_{\text{dark}}^{(\text{tr})} = 0.01$ (a), 0.1 (b), and 0.2 (c). Red circles and squares indicate the pixels with minimum and maximum entropy of each network, respectively. (d) MNIST classification accuracies according to the cumulative pruning of pixels (i.e., enforced zero output to the digital network regardless of input) with ascending (filled circles) or descending (empty squares) orders of entropy.

The concept of noise-induced compression can be proven by “pruning^{32,50} detections,” indirectly revealing the contribution of each pixel to the final inference. The pruning of a pixel means nullifying the corresponding detection by transferring only the precalculated batch-mean intensity $\bar{x}_\alpha = \text{mean}[I_\alpha^{(\text{out})}]$ instead of the exact detection value x_α . Starting from the minimum-entropy pixel, cumulative pruning of pixels in ascending order for the high-noise model (Figure 4c) does not significantly affect the classification performance until about 600 pixels are eliminated, as depicted by the orange solid line in Figure 4d. This result implies that the OPU is trained in

a way that only around 200 detectors are meaningful. Pruning in the opposite (descending) order, beginning with the most important detection, however, results in a rapid accuracy drop (orange dashed line) for the initial 400 cumulative prunings and eventually leads to $\sim 10\%$ accuracy which is equivalent to random guessing. In sharp contrast, the model with almost zero training noise (Figure 4a) undergoes a more linear-like performance degradation upon the cumulative pruning of detectors in both descending (black dashed line) and ascending (black solid line) orders. These results show the training noise-induced emergence of hub (high entropy) and periphery (low entropy) detectors of differential importance.

Practical Example: Incoherent Meta-Imaging System.

We have analyzed several conceptual results in the discrete model concerning the significance of optical preprocessing in mitigating vulnerability against dark noise. To validate our theoretical approach, we present a practical demonstration through the design of diffractive optical systems, termed the meta-imaging system. This system exhibits superior tolerance to dark noise compared to conventional imaging devices such as a simple $4f$ system.

Importantly, we extend our discussion to incoherent systems, given that the weak-signal characteristics in IR applications are due to the absence of an external coherent source. Therefore, it is worth remarking on the difference between the coherent and incoherent systems. Let us suppose a spatially incoherent input $E_\alpha^{(\text{in})}(t) = [I_\alpha^{(\text{in})}]^{1/2} \exp[i\phi_\alpha(t)]$ with a constant intensity $I_\alpha^{(\text{in})}$ and a time-varying phase $\phi_\alpha(t)$, extending our discussion to the more realistic passive environment where light usually originates from incoherent sources such as surface emission by blackbody radiation.⁵³ Given the assumption, the linear field relation $E^{(\text{out})} = PE^{(\text{in})}$ leads to a corresponding linear intensity relation^{43,54} when averaged over a long enough period:

$$\langle |E_\alpha^{(\text{out})}|^2 \rangle_t = \sum_\beta |P_{\alpha\beta}|^2 |E_\beta^{(\text{in})}|^2 \quad (6)$$

or, simply $\langle I^{(\text{out})} \rangle_t = SI^{(\text{in})}$, where $\langle \cdot \rangle_t$ denotes the time average and $S_{\alpha\beta} = |P_{\alpha\beta}|^2$ (see Note S1, Figures S3 and S4 for derivation and a detailed focusing example).

As mentioned earlier, optical imaging systems such as the $4f$ system depicted in Figure 5a are typically linear, which allows us to describe the system through the linear operation between the electric field distributions at the input (object, $z = 0$) and output (image, $z = 4f_0$) planes for a coherent input, or through the linear intensity relation for an incoherent input as well. Especially, when two convex lenses (L1 and L3) of parabolic phase profile $\Phi(x,y) = -\pi(x^2 + y^2)/\lambda f_0$ are placed at $z = f_0$ and $3f_0$, where λ and f_0 are the wavelength and the focal length, respectively, the system operates as an ideal imager. The upper rows of Figure 5b–d illustrate the low-contrast incoherent images of MNIST objects without noise (Figure 5b) and with weak (Figure 5c) and strong dark noise (Figure 5d), calculated using eq 6 (see Methods section for derivation of P). Notably, the coherence length is naturally limited during the numerical calculations as $\langle E(\mathbf{r}_1)E^*(\mathbf{r}_2) \rangle = 0$ if $|\mathbf{r}_1 - \mathbf{r}_2| \geq \sqrt{2}\Delta x$, where $\Delta x = 1.5\lambda$ is the sampling distance chosen for this study.

While the $4f$ system produces a clear image before detection (Figure 5b), its low-intensity contrast is insufficient to withstand the pronounced detection noise (Figure 5d). To address this issue with the same strategy of concentrating optical energy into smaller meaningful regions, we introduce

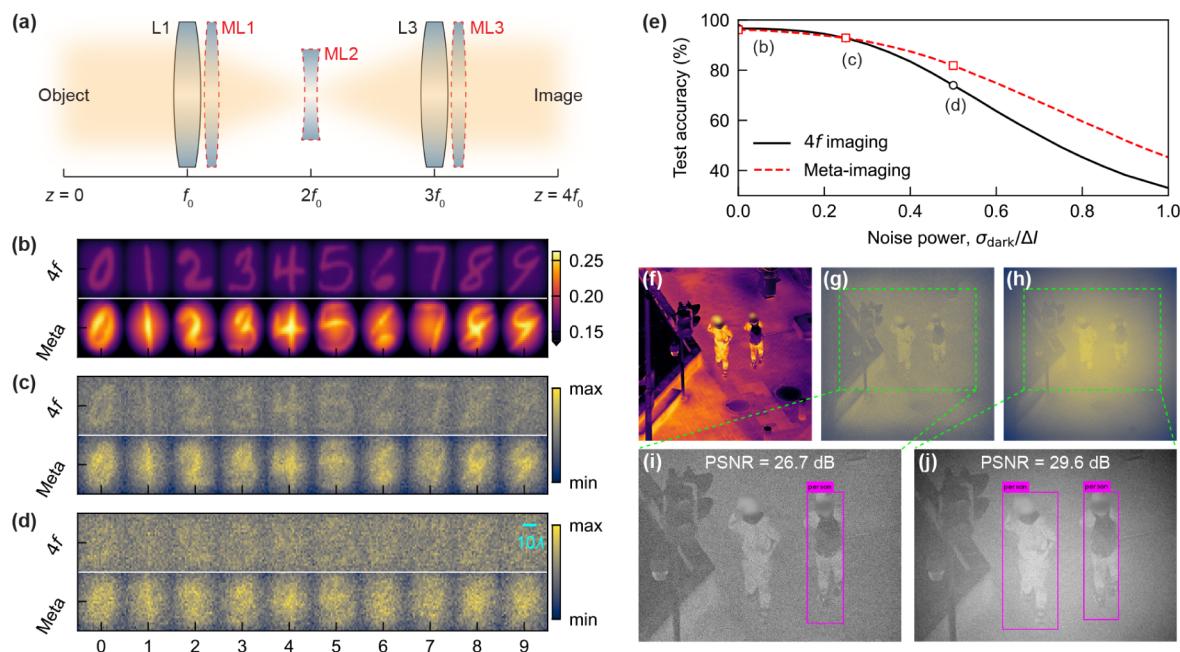


Figure 5. Simulation of incoherent meta-imaging systems. (a), Illustrations of a conventional $4f$ system (lenses; L1 and L3) and a meta-imaging system with additional trainable phase masks (metamaterials; ML1–3). (b–d), Pure images without noise (b) and noisy images with dark noise power $\sigma_{\text{dark}} = \Delta I/4$ (c) and $\Delta I/2$ (d), obtained by $4f$ (upper) and the optimized meta-imaging (lower) systems for digits 0 to 9. (e), MNIST classification accuracies of the conventional (black) and the meta-images (red) as a function of dark noise power. λ , wavelength; $f_0 = 300\lambda$ and NA ~ 0.22 , focal length and numerical aperture of L1 and L3; $\Delta I \sim 0.051$, constant for the intensity contrast in conventional images. (f–j), Example IR images in reality: a scene of pedestrians (f); its modified images based on the pixel-wise intensity ranges for the $4f$ (g) and the optimized meta-imaging (h) systems with the same level of additional Gaussian noises; and the object detection results [magenta boxes, (i) and (j)] using the YOLOv3 model⁵¹ for (g) and (h), respectively. Each image is normalized with its minimum and maximum values and further blurred around the facial features of the pedestrians. Adapted with permission from LLVIP data set introduced in ref. 52 Copyright 2021 Zhu et al.

additional phase-shift masks, namely metamaterials.⁵⁵ Positioned at $z = f_0$, $2f_0$, and $3f_0$, three metamaterials (ML1–3, Figure 5a) provide extra degrees of freedom for optimizing phase profiles, boosting classification accuracy under detection noise (see Table S2 and Figure S5 for optimization details). As a result, the trained meta-imaging system generates high-contrast incoherent images by vignetting the less informative area around the four corners while simultaneously highlighting the central part to mitigate the impact of dark noise (bottom row of Figure 5b). This enhanced intensity contrast certainly improves the machine perception of objects in the presence of strong dark noise, as compared by the bottom and top rows of Figure 5d. Figure 5e further quantifies the noticeable enhancement in the noise resilience achieved through the metamaterial-assisted classification (red dashed line), outperforming the conventional $4f$ imaging system (black line).

Our concept of metamaterials inspires the development of practical machine vision systems, as demonstrated in Figure 5f–j. Despite the lack of realistic considerations such as spectral and geometrical parameters in the previous $4f$ and meta-imaging systems and entirely different target purposes as well, their predetection modulation behaviors, specifically concentrating optical energy on the central part, can be adapted to a scene⁵² of two pedestrians captured by an IR camera (Figure 5f). This gives rise to vignetting and Gaussian noise in the images produced by the $4f$ (Figure 5g) and the meta-imaging (Figure 5h) systems. Similar to the comparison as shown in Figure 5b, the meta-image in Figure 5h exhibits a brighter central area. This optical treatment enhances the image quality against the same level of noise, as represented by the peak signal-to-noise ratio ($\text{PSNR} \equiv 20 \log_{10}[\max(I_a^{(\text{out})})]$)

of 29.6 dB (Figure 5h) compared to 26.7 dB (Figure 5g), which resultingly facilitates machine vision applications in industrial settings. Indeed, employing a pretrained model (YOLOv3⁵¹) allows for the detection of both pedestrians (Figure 5j) in the noisy meta-image (Figure 5h), while only one pedestrian (Figure 5i) is detectable in the conventional image (Figure 5g) with the same level of Gaussian noise. It is noted that the result shown in Figures 5f–j is not optimal but merely a single example showing a clear difference, demanding detailed optimization with a feasible design strategy in future work.

CONCLUSION

Although various types of noise and experimental errors, such as the misalignment of optical elements,^{27,56} can impact the final results, we have focused exclusively on two noise mechanisms—dark and photon noises—within the optoelectronic detection process. Particularly, the dark noise characteristics were analyzed to assess the effectiveness of SNR enhancement by predetection OPUs. However, as the size of the OPU increases, noises and errors generated within the OPU itself may accumulate, which requires further study in comparison to the detection noises.

Our results align with previous studies, which demonstrate that optical preprocessing using vortex-Fourier operations⁴⁶ or subpixel binning⁴⁸ can enhance noise robustness in image reconstruction tasks. We build on these important findings by incorporating various types of optical processing to articulate a seemingly obvious fact with quantitative analysis: redistributing optical signal power while preserving sufficient information is essential for noise resilience. It is also noteworthy that our

work extends the application of OPUs into general neural tasks, such as object recognition, beyond their role in achieving denoising effects in image reconstruction.^{26,46} Limiting our discussion to “linear” OPUs, however, this approach differs from the perspective of ref. 24 which emphasizes the importance of optical nonlinearity. While nonlinearity plays a crucial role in overall neural performance despite active energy consumption, it does not necessarily explain why it should be performed in the optical domain. Our work highlights the necessity of linear OPUs that cannot be replaced by digital counterparts after the impact of noise.

To summarize, we have verified the role of optical processing in advance of detection, which concentrates the optical signal power into a smaller region to address the low SNR challenge in noisy systems such as infrared devices. Through optical computing, the information redundancy in the original distribution of signal power is eliminated until the target performance is not maintained, while the detection power per detector is amplified due to the conservation of the total signal energy. Compared with the ideal imaging model, where the optical signal is mainly obscured by the severe dark noise, our proposed machine-learned and manually defined optical operations have demonstrated the ability to strategically redistribute optical signals to effectively compete with noise. This outcome underscores the imperative need for harnessing optical computation resources, not only for ultrafast and energy-efficient bosonic computing but also to navigate noisy environments that cannot be adequately addressed solely through postdetection digital processing.

METHODS

Definition of the Block-Wise Matrices. In general, a discrete Fourier transform (DFT) tensor in a rectangular domain ($N_x \times N_y$) is defined as

$$[P_F(N_x, N_y)]_{k,l,m,n} \equiv \frac{1}{\sqrt{N_x N_y}} \exp\left[2\pi i \left(\frac{km}{N_x} + \frac{ln}{N_y}\right)\right] \quad (7)$$

where (k, l) and (m, n) are the 2D coordinates satisfying $0 \leq k, m < N_x$ and $0 \leq l, n < N_y$. Using this definition as a building block, the block-wise DFT tensor can be written as

$$\begin{aligned} [P_{BF}]_{k,l,m,n} &\equiv \sum_{X,Y} [P_F(L_x, L_y)]_{k-x_0, l-y_0, m-x_0, n-y_0} \\ I_X(k)I_Y(l)I_X(m)I_Y(n) \end{aligned} \quad (8)$$

where $X = \{k \in \mathbb{Z}: x_0 \leq k < x_0 + L_x\}$ is iterated over disjoint subsets of the integer range $0 \leq k < N_x$ slicing the 2D domain into columns, Y is defined in the same manner, and $I_X(k)$ is the indicator function that returns 1 if $k \in X$ and otherwise 0. Reshaping the 2D indices into flattened 1D indices, $(k, l) \leftrightarrow \alpha$ and $(m, n) \leftrightarrow \beta$, the unitary matrices $[P_F]_{\alpha\beta}$ and $[P_{BF}]_{\alpha\beta}$ in the main text can be derived using (k_α, l_α) and (m_β, n_β) , which are the quotient-remainder pairs of integers α and β with N_y , respectively.

Diffractive Optics. Based on the Rayleigh–Sommerfeld diffraction integral, the spatial evolution of a scalar electromagnetic wave along z -direction can be described as

$$E(x, y, z_2) = \frac{1}{i\lambda} \int dx' dy' \frac{e^{ik_0 R}}{R} \frac{z_2 - z_1}{R} \left(1 + \frac{i}{k_0 R}\right) E(x', y', z_1) \quad (9)$$

where λ and $k_0 = 2\pi/\lambda$ are a free-space wavelength and the corresponding wavenumber, respectively, and $R = [(x - x')^2 + (y - y')^2 + (z_2 - z_1)^2]^{1/2}$ is the distance between the source (x', y') and observation (x, y) points at $z = z_2$ and z_1 planes, respectively. Sampling the continuous electric fields with rectangular basis functions as,

$$E(x', y', z_1) \approx \sum_{k,l} E_{k,l} \text{rect}\left(\frac{x' - k\Delta x}{\Delta x}\right) \text{rect}\left(\frac{y' - l\Delta x}{\Delta x}\right) \quad (10)$$

$$E(x, y, z_2) \approx \sum_{m,n} E_{m,n} \text{rect}\left(\frac{x - m\Delta x}{\Delta x}\right) \text{rect}\left(\frac{y - n\Delta x}{\Delta x}\right) \quad (11)$$

where $\text{rect}(a) \equiv 1$ if $|a| < 1/2$ and elsewhere 0, a discretized numerical linear relationship can be derived as

$$E_{m,n} = G_{m,n}^{k,l} E_{k,l} \quad (12)$$

where

$$G_{m,n}^{k,l} = \frac{1}{i\lambda} \int_{(k-1/2)\Delta x}^{(k+1/2)\Delta x} dx' \int_{(l-1/2)\Delta x}^{(l+1/2)\Delta x} dy' \frac{(z_2 - z_1) \exp(i k_0 R_{m,n}^{k,l})}{(R_{m,n}^{k,l})^2} \left(1 + \frac{i}{k_0 R_{m,n}^{k,l}}\right) \quad (13)$$

and $R_{m,n}^{k,l}(z_2 - z_1) = [\Delta x^2(k - m)^2 + \Delta x^2(l - n)^2 + (z_2 - z_1)^2]^{1/2}$. On top of that, lenses and metasurfaces in the main text are assumed to be infinitesimally thin and therefore lead to a point-by-point local phase jump, which can be described by $E_{m,n}(z = z_0^+) = \Phi_{m,n} E_{m,n}(z = z_0^-)$, where $z = z_0$ is the location of the lens. By multiplying these transfer relationships alternatively through the lens array, one can obtain the input-output relation of the entire optical system as

$$E_{\alpha_L}^{(\text{out})} = P_{\alpha_L}^{\alpha_0} E_{\alpha_0} = [G_{\alpha_L}^{\alpha_{L-1}} \Phi_{\alpha_{L-1}} G_{\alpha_{L-1}}^{\alpha_{L-2}} \Phi_{\alpha_{L-2}} \cdots G_{\alpha_1}^{\alpha_0}] E_{\alpha_0}^{(\text{in})} \quad (14)$$

where α_l for $0 \leq l \leq L$ is the flattened 1D index on planes $z = z_l$ including input ($z = z_0$) and output ($z = z_L$) planes.

ASSOCIATED CONTENT

Data Availability Statement

The codes used to generate the results of this study are publicly available at <https://github.com/jmkim93/compute-first-optical-detection>.

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acspophotonics.4c02284>.

Network architectures (Tables S1–S2), notes on incoherent systems (Note S1), learning curves (Figure S1), testing results (Figure S2), incoherent input results with unitarity (Figure S3), incoherent focusing examples (Figure S4), and trained metasurface system (Figure S5) (PDF)

AUTHOR INFORMATION

Corresponding Author

Jungmin Kim – Department of Electrical and Computer Engineering, University of Wisconsin-Madison, Madison, WI 53706, United States; orcid.org/0000-0001-8567-5338; Email: jkim2325@wisc.edu

Authors

Nanfang Yu — Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027, United States;  orcid.org/0000-0002-9462-4724

Zongfu Yu — Department of Electrical and Computer Engineering, University of Wisconsin-Madison, Madison, WI 53706, United States;  orcid.org/0000-0002-4536-1526

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acsphtronics.4c02284>

Author Contributions

Z.Y. conceived the project. J.K. developed the theory, performed the numerical simulations, and wrote the first draft. All authors discussed the results, edited the manuscript, and approved the content.

Funding

The work was supported by the Army Research Office through a Multidisciplinary University Research Initiative program (Grant No. W911NF-22-2-0111).

Notes

The authors declare no competing financial interest.

■ REFERENCES

- (1) Rogalski, A. *Infrared Detectors*, 2nd ed.; CRC Press: Boca Raton, 2010.
- (2) Yang, J.; Zhang, X.; Zhang, X.; Wang, L.; Feng, W.; Li, Q. Beyond the Visible: Bioinspired Infrared Adaptive Materials. *Adv. Mater.* **2021**, *33* (14), 2004754.
- (3) Bao, F.; Wang, X.; Sureshbabu, S. H.; Sreekumar, G.; Yang, L.; Aggarwal, V.; Boddeti, V. N.; Jacob, Z. Heat-assisted detection and ranging. *Nature* **2023**, *619*, 743–748.
- (4) Marnissi, M. A.; Fathallah, A. GAN-based Vision Transformer for High-Quality Thermal Image Enhancement. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*; IEEE, 2023, pp. 817–825.
- (5) Raman, A. P.; Anoma, M. A.; Zhu, L.; Rephaeli, E.; Fan, S. Passive radiative cooling below ambient air temperature under direct sunlight. *Nature* **2014**, *515*, 540–544.
- (6) Fan, S.; Li, W. Photonics and thermodynamics concepts in radiative cooling. *Nat. Photonics* **2022**, *16*, 182–190.
- (7) Zhao, X.; Li, T.; Xie, H.; Liu, H.; Wang, L.; Qu, Y.; Li, S. C.; Liu, S.; Brozena, A. H.; Yu, Z.; Srebric, J.; Hu, L. A solution-processed radiative cooling glass. *Science* **2023**, *382*, 684–691.
- (8) Munday, J. N. Tackling Climate Change through Radiative Cooling. *Joule* **2019**, *3*, 2057–2060.
- (9) Amenabar, I.; Poly, S.; Goikoetxea, M.; Nuansing, W.; Lasch, P.; Hillenbrand, R. Hyperspectral infrared nanoimaging of organic samples based on Fourier transform infrared nanospectroscopy. *Nat. Commun.* **2017**, *8*, 14402.
- (10) Wang, Z.; Yi, S.; Chen, A.; Zhou, M.; Luk, T. S.; James, A.; Nogan, J.; Ross, W.; Joe, G.; Shahsafi, A.; et al. Single-shot on-chip spectral sensors based on photonic crystal slabs. *Nat. Commun.* **2019**, *10*, 1020.
- (11) Makarenko, M.; Burguete-Lopez, A.; Wang, Q.; Getman, F.; Giancola, S.; Ghanem, B.; Fratalocchi, A. Real-time Hyperspectral Imaging in Hardware via Trained Metasurface Encoders. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE, 2022, pp. 12682–12692.
- (12) Zhao, Y.; Kusama, S.; Furutani, Y.; Huang, W.-H.; Luo, C.-W.; Fuji, T. High-speed scanless entire bandwidth mid-infrared chemical imaging. *Nat. Commun.* **2023**, *14*, 3929.
- (13) Wang, X.; Yang, Z.; Bao, F.; Sentz, T.; Jacob, Z. Spinning metasurface stack for spectro-polarimetric thermal imaging. *Optica* **2024**, *11*, 73–80.
- (14) Gurton, K. P.; Yuffa, A. J.; Videen, G. W. Enhanced facial recognition for thermal imagery using polarimetric imaging. *Opt. Lett.* **2014**, *39*, 3857.
- (15) Yuffa, A. J.; Gurton, K. P.; Videen, G. Three-dimensional facial recognition using passive long-wavelength infrared polarimetric imaging. *Appl. Opt.* **2014**, *53*, 8514.
- (16) Deng, W.; Dai, M.; Wang, C.; You, C.; Chen, W.; Han, S.; Han, J.; Wang, F.; Ye, M.; Zhu, S.; et al. Switchable Unipolar-Barrier Van der Waals Heterostructures with Natural Anisotropy for Full Linear Polarimetry Detection. *Adv. Mater.* **2022**, *34*, 2203766.
- (17) Szeliski, R. *Computer Vision: Algorithms and Applications*, 2nd ed.; Springer International Publishing, 2022.
- (18) Lin, X.; Rivenson, Y.; Yardimci, N. T.; Veli, M.; Luo, Y.; Jarrahi, M.; Ozcan, A. All-optical machine learning using diffractive deep neural networks. *Science* **2018**, *361*, 1004–1008.
- (19) Wu, Z.; Zhou, M.; Khoram, E.; Liu, B.; Yu, Z. Neuromorphic metasurface. *Photon. Res.* **2020**, *8*, 46.
- (20) Wu, Z.; Yu, Z. Small object recognition with trainable lens. *APL Photonics* **2021**, *6*, 071301.
- (21) Bernstein, L.; Sludds, A.; Panuski, C.; Trajtenberg-Mills, S.; Hamerly, R.; Englund, D. Single-shot optical neural network. *Sci. Adv.* **2023**, *9*, No. eadg7904.
- (22) Chen, Y.; Zhou, T.; Wu, J.; Qiao, H.; Lin, X.; Fang, L.; Dai, Q. Photonic unsupervised learning variational autoencoder for high-throughput and low-latency image transmission. *Sci. Adv.* **2023**, *9*, No. eadg8437.
- (23) Fu, T.; Zang, Y.; Huang, Y.; Du, Z.; Huang, H.; Hu, C.; Chen, M.; Yang, S.; Chen, H. Photonic machine learning with on-chip diffractive optics. *Nat. Commun.* **2023**, *14* (1), 70.
- (24) Wang, T.; Sohoni, M. M.; Wright, L. G.; Stein, M. M.; Ma, S.-Y.; Onodera, T.; Anderson, M. G.; McMahon, P. L. Image sensing with multilayer nonlinear optical neural networks. *Nat. Photonics* **2023**, *17*, 408–415.
- (25) Zhou, T.; Wu, W.; Zhang, J.; Yu, S.; Fang, L. Ultrafast dynamic machine vision with spatiotemporal photonic computing. *Sci. Adv.* **2023**, *9* (23), No. eadg4391.
- (26) İşıl, C.; Gan, T.; Ardic, F. O.; Mentesoglu, K.; Digani, J.; Karaca, H.; Chen, H.; Li, J.; Mengü, D.; Jarrahi, M.; et al. All-optical image denoising using a diffractive visual processor. *Light: Sci. Appl.* **2024**, *13*, 43.
- (27) Zheng, Z.; Duan, Z.; Chen, H.; Yang, R.; Gao, S.; Zhang, H.; Xiong, H.; Lin, X. Dual adaptive training of photonic neural networks. *Nat. Mach. Intell.* **2023**, *5*, 1119–1129.
- (28) Shen, Y.; Harris, N. C.; Skirlo, S.; Prabhu, M.; Baehr-Jones, T.; Hochberg, M.; Sun, X.; Zhao, S.; Larochelle, H.; Englund, D.; Soljačić, M. Deep learning with coherent nanophotonic circuits. *Nat. Photonics* **2017**, *11*, 441–446.
- (29) Hamerly, R.; Bernstein, L.; Sludds, A.; Soljačić, M.; Englund, D. Large-Scale Optical Neural Networks Based on Photoelectric Multiplication. *Phys. Rev. X* **2019**, *9*, 021032.
- (30) Mourigas-Alexandris, G.; Moralis-Pegios, M.; Tsakyridis, A.; Simos, S.; Dabos, G.; Totovic, A.; Passalis, N.; Kirtas, M.; Rutirawut, T.; Gardes, F. Y.; et al. Noise-resilient and high-speed deep learning with coherent silicon photonics. *Nat. Commun.* **2022**, *13* (1), 5572.
- (31) Chen, Z.; Sludds, A.; Davis, R.; Christen, I.; Bernstein, L.; Ateshian, L.; Heuser, T.; Heermeier, N.; Lott, J. A.; Reitzenstein, S.; Hamerly, R.; Englund, D. Deep learning with coherent VCSEL neural networks. *Nat. Photonics* **2023**, *17*, 723–730.
- (32) Yu, S.; Park, N. Heavy tails and pruning in programmable photonic circuits for universal unitaries. *Nat. Commun.* **2023**, *14* (1), 1853.
- (33) Martel, J. N. P.; Müller, L. K.; Carey, S. J.; Dudek, P.; Wetstein, G. Neural Sensors: Learning Pixel Exposures for HDR Imaging and Video Compressive Sensing With Programmable Sensors. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 1642–1653.
- (34) Mennel, L.; Symonowicz, J.; Wachter, S.; Polyushkin, D. K.; Molina-Mendoza, A. J.; Mueller, T. Ultrafast machine vision with 2D material neural network image sensors. *Nature* **2020**, *579*, 62–66.

- (35) Chen, Y.; Nazhamaiti, M.; Xu, H.; Meng, Y.; Zhou, T.; Li, G.; Fan, J.; Wei, Q.; Wu, J.; Qiao, F.; Fang, L.; Dai, Q. All-analog photoelectronic chip for high-speed vision tasks. *Nature* **2023**, *623*, 48–57.
- (36) Yuan, S.; Ma, C.; Fetaya, E.; Mueller, T.; Naveh, D.; Zhang, F.; Xia, F. Geometric deep optical sensing. *Science* **2023**, *379* (6637), No. eade1220.
- (37) Colthup, N. B.; Daly, L. H.; Wiberley, S. E. *Introduction to Infrared and Raman Spectroscopy*, 3rd ed.; Academic Press: San Diego, 1990, pp. 75–107.
- (38) Cohen, G.; Afshar, S.; Tapson, J.; van Schaik, A.; EMNIST: An extension of MNIST to handwritten letters. *arXiv*, **2017**.
- (39) Hughes, T. W.; Minkov, M.; Williamson, I. A. D.; Fan, S. Adjoint Method and Inverse Design for Nonlinear Nanophotonic Devices. *ACS Photonics* **2018**, *5*, 4781–4787.
- (40) Carolan, J.; et al. Universal linear optics. *Science* **2015**, *349*, 711–716.
- (41) Clements, W. R.; Humphreys, P. C.; Metcalf, B. J.; Kolthammer, W. S.; Walsmley, I. A. Optimal design for universal multiport interferometers. *Optica* **2016**, *3*, 1460.
- (42) Reck, M.; Zeilinger, A.; Bernstein, H. J.; Bertani, P. Experimental realization of any discrete unitary operator. *Phys. Rev. Lett.* **1994**, *73*, 58–61.
- (43) Rahman, M. S. S.; Yang, X.; Li, J.; Bai, B.; Ozcan, A. Universal linear intensity transformations using spatially incoherent diffractive processors. *Light: Sci. Appl.* **2023**, *12*, 195.
- (44) Piggott, A. Y.; Lu, J.; Lagoudakis, K. G.; Petykiewicz, J.; Babinec, T. M.; Vučković, J. Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer. *Nat. Photonics* **2015**, *9*, 374–377.
- (45) Khoram, E.; Chen, A.; Liu, D.; Ying, L.; Wang, Q.; Yuan, M.; Yu, Z. Nanophotonic media for artificial neural inference. *Photon. Res.* **2019**, *7*, 823.
- (46) Muminov, B.; Vuong, L. T. Fourier optical preprocessing in lieu of deep learning. *Optica* **2020**, *7*, 1079–1088.
- (47) Lin, R. J.; et al. Achromatic metalens array for full-colour light-field imaging. *Nat. Nanotechnol.* **2019**, *14*, 227–231.
- (48) Mennel, L.; Molina-Mendoza, A. J.; Paur, M.; Polyushkin, D. K.; Kwak, D.; Giparakis, M.; Beiser, M.; Andrews, A. M.; Mueller, T. A photosensor employing data-driven binning for ultrafast image recognition. *Sci. Rep.* **2022**, *12* (1), 14441.
- (49) Jaynes, E. T. Information Theory and Statistical Mechanics. *Phys. Rev.* **1957**, *106*, 620–630.
- (50) Han, S.; Pool, J.; Tran, J.; Dally, W. J. Learning both Weights and Connections for Efficient Neural Network. In *Proceedings of the 29th International Conference on Neural Information Processing Systems*, ACM, 2015, pp. 1135–1143.
- (51) Redmon, J.; Farhadi, A.; YOLOv3: An Incremental Improvement. *arXiv*, **2018**.
- (52) Jia, X.; Zhu, C.; Li, M.; Tang, W.; Zhou, W. LLVIP: A Visible-Infrared Paired Dataset for Low-Light Vision. In *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, IEEE, 2021, pp. 3489–3497.
- (53) Mehta, C. L.; Wolf, E. Coherence Properties of Blackbody Radiation. I. Correlation Tensors of the Classical Field. *Phys. Rev.* **1964**, *134*, A1143–A1149.
- (54) Filipovich, M. J.; Malyshev, A.; Lvovsky, A. I. Role of spatial coherence in diffractive optical neural networks. *Opt. Express* **2024**, *32*, 22986–22997.
- (55) Khorasaninejad, M.; Capasso, F. Metalenses: Versatile multifunctional photonic components. *Science* **2017**, *358* (6367), No. eaam8100.
- (56) Mengu, D.; Zhao, Y.; Yardimci, N. T.; Rivenson, Y.; Jarrahi, M.; Ozcan, A. Misalignment resilient diffractive optical networks. *Nanophotonics* **2020**, *9*, 4207–4219.

**CAS INSIGHTS™****EXPLORE THE INNOVATIONS SHAPING TOMORROW**

Discover the latest scientific research and trends with CAS Insights. Subscribe for email updates on new articles, reports, and webinars at the intersection of science and innovation.

[Subscribe today](#)

