

Nonlocal flat optics for size-selective image processing and denoising

Received: 27 March 2025

Accepted: 2 May 2025

Published online: 14 May 2025



Sandeep Kumar Chamoli^{1,2,6}, Chunqi Jin^{1,2,6}, Yandong Fan^{1,2,6}, Mehdi Kiani^{3,4}, Heedong Goh³, Chen Huang^{1,2}, Shuyu Guo⁵, Yuntong Wang^{1,2}, Fei Zhu^{1,2}, Guohua Xing^{1,2}, Bo Li^{1,2}, Tian Bai⁵, Andrea Alù^{3,4} ✉ & Wei Li^{1,2} ✉

All-optical image processing based on metasurfaces is a swiftly advancing field of technology, due to its high speed, large integrability and inherently low energy requirements. So far, the proposed devices have been focusing on canonical operations, such as differentiations to perform edge detection across all objects in a complex scene. Yet, undesired background noise and clutter can hinder such operations, requiring target selection with digital post-processing which inherently limits the overall accuracy, efficiency and speed. Here, we introduce an optical solution for real-time size-selective image processing and experimentally demonstrate the concept with a metal-dielectric-metal film performing a spatial band-pass filter in momentum space. We show high-resolution ($\sim 0.9 \mu\text{m}$) edge detection and real-time dynamic denoising, ideally suited for bio-imaging applications and target recognitions. Our demonstrated k -space filtering metasurface expands the scope of nonlocal flat optics for analog image processing, ushering in opportunities for ultra-compact, cost-effective, and multifunctional image processors.

Amidst the growing demand for large-scale real-time data processing in the big data era, all-optical image processing^{1–6} has garnered significant attention. With its distinct advantages, such as rapid processing speed and low power consumption, this technology has rapidly become an attractive alternative to traditional digital approaches requiring analog-to-digital conversion^{7,8}. This research direction is steadfastly advancing, with applications spanning various scientific and engineering domains, including target recognition⁹, medical or biological imaging¹⁰, autonomous vehicles¹¹, and computer vision¹². Recently, leveraging the customized wavevector-dependent (nonlocal¹³) response of spatial optical filters, all-optical image processing has transitioned from using conventional bulky lenses and Fourier optics¹⁴ to employing spatial differentiation-based¹⁵ flat nanophotonic devices. These devices include multilayer films^{16,17}, metasurfaces^{15,18–29}, photonic crystals^{30,31}, plasmonic structures^{32,33}, the

spin Hall effect^{34–36} and topological photonics³⁷, contributing to the miniaturization of image processing systems and reduced demands in terms of precise positioning of optical filters and alignment.

The most explored operation in all-optical image processing—edge detection—is a key pre-processing step for machine vision, including target recognition. This operation extracts the relevant geometric information of the objects within an image, allowing machine learning models to recognize specific targets³⁸ with improved efficiency. Implemented with nanophotonic devices, optical analog spatial differentiators perform massively parallel edge detection on objects of varying sizes within an entire image^{16,30,39–46}. Inevitably, the edges associated with background noise and clutter⁴⁷ may be amplified unwantedly, leading to false positives or degraded recognition performance, and requiring further operations to achieve more accurate target recognition. The critical recognition improvement relies on

¹GPL Photonics Laboratory, State Key Laboratory of Luminescence Science and Technology, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, Jilin 130033, P. R. China. ²University of Chinese Academy of Sciences, Beijing 100049, China. ³Photonics Initiative, Advanced Science Research Center, City University of New York, New York, NY 10031, USA. ⁴Physics Program, Graduate Center of the City University of New York, New York, NY 10016, USA. ⁵College of Computer Science and Technology, Jilin University, Changchun 130012, China. ⁶These authors contributed equally: Sandeep Kumar Chamoli, Chunqi Jin, Yandong Fan. ✉ e-mail: aalu@gc.cuny.edu; weili1@ciomp.ac.cn

the capacity to selectively enhance specific targets while excluding unwanted noise and clutter. However, despite significant efforts^{5,13,23,48,49}, such capability remains challenging.

In this work, we present a real-time size-selective image processor, incorporating edge detection and denoising based on target sizes by introducing a band-pass optical transfer function (OTF). We implement the desired functionality through the customization of the nonlocal characteristics of a simple metal–dielectric–metal (MDM) film. Our work achieves the highest experimental resolution ($\sim 0.9\ \mu\text{m}$) for edge detection based on spatial differentiation and realizes real-time dynamic denoising. We demonstrate the potential of this approach to significantly enhance the efficiency of tasks such as bio-imaging and target recognition. Notably, the use of nonlocal multilayer films, with transversely homogeneous nature, eliminates position-dependent response and alignment issues. This makes them compatible with widely available optical thin films manufactured through standardized coating techniques. Compared to metasurfaces and photonic crystal platforms, this approach offers notable advantages in cost-effectiveness and scalability for large-scale production. Consequently, our approach expands the operational capabilities of nonlocal flat optics, providing an alignment-free retrofit for existing imaging platforms and opening opportunities for ultra-compact and multifunctional image processors.

Results

Design principles and simulation results

In the spatial frequency domain (k -space), the input electric field $E_{\text{in}}(x, y)$ is represented by its Fourier transform, $\tilde{E}_{\text{in}}(k_x, k_y) = \iint E_{\text{in}}(x, y) \exp(-i(k_x x + k_y y)) dx dy$, where x and y denote spatial coordinates, k_x and k_y represent the wavevector. Consequently, the light field is modulated by a wavevector-dependent OTF, with spatial frequency in k -space denoted as k_r/k_0 ($k_r = \sqrt{k_x^2 + k_y^2}$). Following Fourier optics¹⁴, the target in k -space encompasses spatial frequencies ranging from low to high, and their relationship with the input electric

field follows a *sinc* function. In particular, the radius Δr of the low-frequency region depends on the location of the first zero point of the *sinc* function: $\Delta r = A/\sqrt{\Delta x^2 + \Delta y^2}$, where Δx and Δy represent the size of the input in the x - and y -coordinate, and A is a system-specific constant. This analysis establishes a constant spatial frequency range for a given system. As the size of the input target (Δx or Δy) decreases in real space, the low-frequency radius in k -space expands, leading to a reduction in the high-frequency ratio, and vice versa. Further details, including the effects of noise and object positioning on the k -space distribution, are provided in SI Note 1. This property leads to distinct frequency distributions corresponding to different input sizes. Therefore, selecting an appropriate band-pass OTF that modulates and filters the information around the high- and low-frequency peaks facilitates size-selective image processing. Illustrated in Fig. 1, we analyze an ideal rectangular band-pass OTF denoted as $H(k_r)$, which can be expressed as:

$$H(k_r) = \text{rect}\left(\frac{k_r}{k_0}\right) = \begin{cases} 0.1, & 0 \leq \frac{k_r}{k_0} < \text{NA}_1 \\ 1, & \text{NA}_1 \leq \frac{k_r}{k_0} \leq \text{NA}_2 \\ 0.1, & \text{NA}_2 < \frac{k_r}{k_0} \leq 1 \end{cases} \quad (1)$$

where $\text{NA}_{1,2}$ denotes the numerical apertures for which the rectangular function induces a step change. Specifically, the range NA_1 to NA_2 corresponds to the transmitted frequency information, while the rest needs to be partially filtered for size-selective image processing.

To realize the required transfer function, we use a Fabry–Pérot cavity-based MDM film comprised of Ag and MgF_2 (Fig. 2a), effectively approximating a homogeneous medium (SI Note 2). According to Fresnel's equation, the transmission (T_{tot}) can be expressed as $T_{\text{tot}} = T^2/(1 + R^2 - 2\sqrt{R^2} \cos \delta)$ for the light incident on the film from air (refractive index n_0) at an angle θ_i . Here, R and T denote the reflection and transmission at the upper and lower interfaces of the film (details in

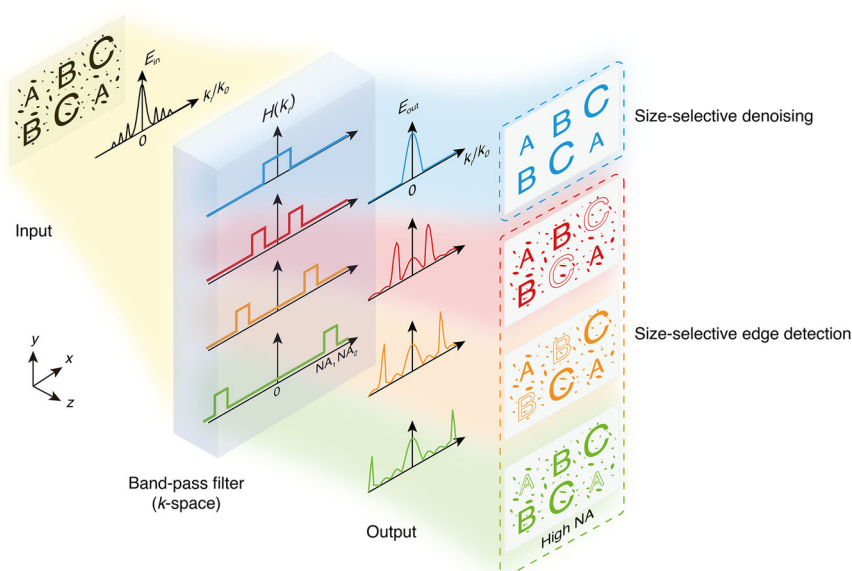


Fig. 1 | Schematic of a size-selective imaging processor acting as a rectangular band-pass filter operator. A conceptual illustration of a nonlocal flat optics-based filter, designed to execute the rectangular band-pass OTF ($H(k_r)$), capable of modulating and filtering specific spatial frequencies associated with targets of different sizes. This modulation is achieved by configuring the band-pass numerical aperture (NA) range (NA_1 – NA_2). Such customization facilitates differentiated image processing, including size-selective edge detection (red, yellow, and green parts),

specifically, high NA edge detection (green part), and denoising (blue part) tailored to different targets. For example, the red part can detect the edge of the larger image 'C' but allows images with other sizes to pass through unchanged; similarly, the yellow part applies to image 'B', and the green part applies to the smaller image 'A'. Meanwhile, the blue part effectively reduces noise without altering the images 'A', 'B', and 'C'.

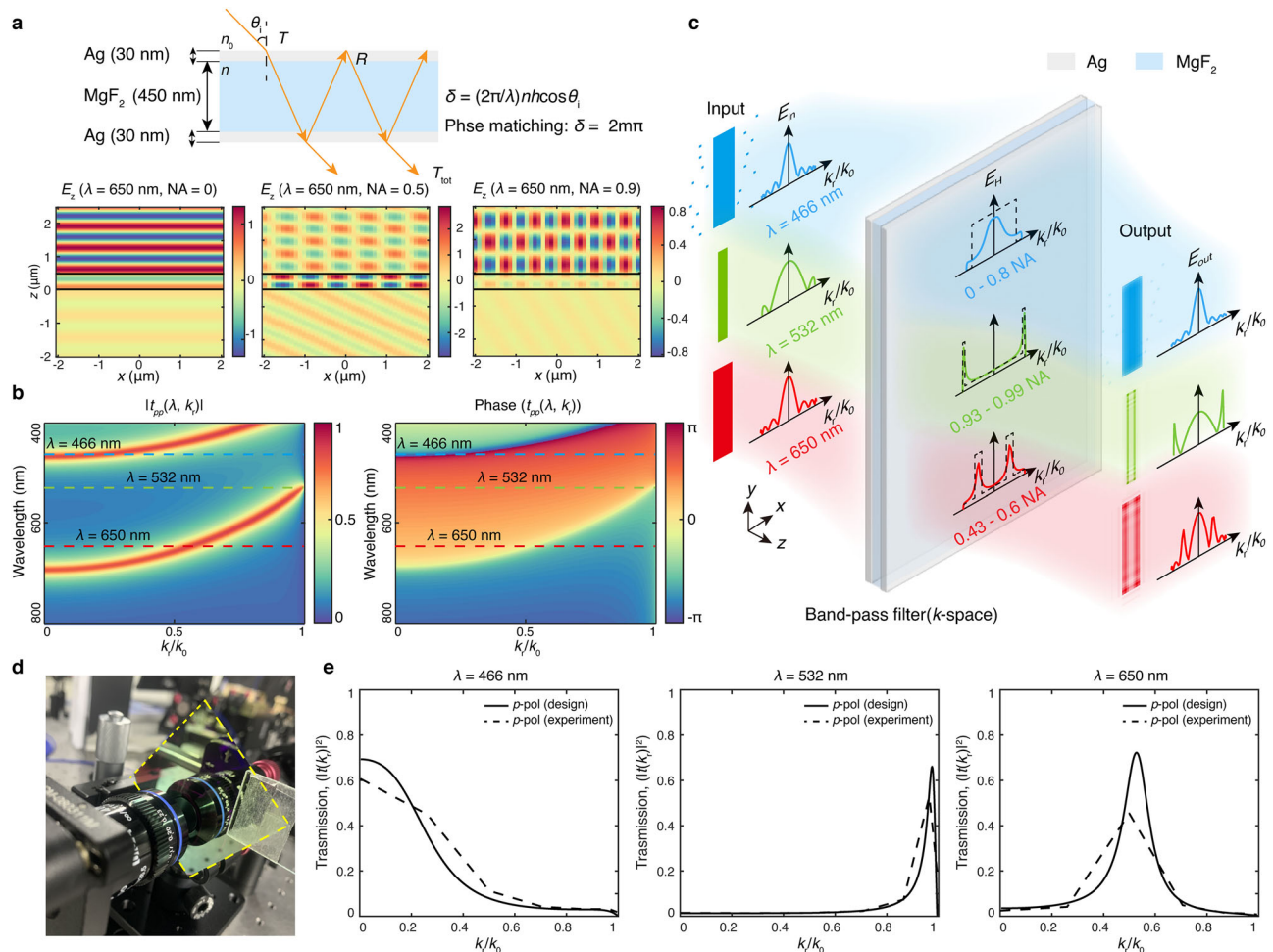


Fig. 2 | Size-selective image processor using a nonlocal MDM film. **a** The transmission principle involves light incidence at an angle θ_i into the MDM film, consisting of Ag and MgF₂ layers with phase accumulation δ . Simulation of electric fields at various incidence wavevectors at a 650 nm wavelength showcases that the nonlocal effect is observable only when the phase is matched (NA = 0.5), leading to maximum transmission, while conversely, transmission is nearly zero for NA of 0 and 0.9. **b** Simulated color-coded transmission coefficient $t_{pp}(\lambda, k_r)$, showcases a second-order resonance within the visible light for p polarization. The selected representative wavelengths include blue ($\lambda = 466$ nm), green ($\lambda = 532$ nm), and red ($\lambda = 650$ nm) light, indicated by the corresponding-colored dashed lines.

c The designed OTFs demonstrate Gaussian-type band-pass filtering at three operating wavelengths with NA ranges of 0–0.8 NA, 0.43–0.6 NA, and 0.93–0.99 NA. The solid-colored lines depict the simulated OTFs, and the dashed lines represent the corresponding ideal rectangular band-pass OTFs. The simulation results demonstrate its size-selective denoising and edge detection capabilities. **d** Photo of the fabricated two-inch size-selective image processor (yellow dashed box). **e** Wavevector-dependent transmission spectra ($|t(k_r)|^2$) by ellipsometry at 466 nm, 532 nm, and 650 nm wavelength under p -polarized light (dashed lines), which align well with the simulation results (solid lines).

SI Note 2), $\delta = (2\pi/\lambda)nh\cos\theta_i$ represents the phase accumulation within the film, where n and h denote the real part of the refractive index and thickness of the film, and λ is the incident wavelength. The transmission (T_{tot}) reaches its maximum when the phase is matched ($\delta = 2m\pi$, where m is an integer), revealing a periodic trend of decreasing wavelength with increasing incident angle. Accordingly, we identify three operating wavelengths—red ($\lambda = 650$ nm), green ($\lambda = 532$ nm), and blue ($\lambda = 466$ nm)—to realize size-selective imaging capability of the film. Through adjustments to the thicknesses of the metal and dielectric layers, we achieve the desired transfer functions at these wavelengths, resulting in corresponding thicknesses of 30 nm and 450 nm, respectively. As depicted in Fig. 2a, simulations are conducted to visualize color-coded electric fields corresponding to different incident wavevectors (NA = 0, 0.5, 0.9) at a wavelength of 650 nm. This result highlights that the film demonstrates a nonlocal effect and achieves maximum transmission only at the designated incident wavevector (NA = 0.5) for a specific operating wavelength ($\lambda = 650$ nm). Following this analysis, we numerically simulate the amplitude ($|t_{pp}(\lambda, k_r)|$) and

phase ($t_{pp}(\lambda, k_r)$) of the transmission coefficients of p -polarized light across a wavelength range of 400–800 nm. As shown in Fig. 2b, our findings reveal a second-order resonance of this MDM film within the visible light range. Results for s -polarized light are available in SI Note 3. At three operating wavelengths, the transfer function adheres to a band-pass Gaussian type with properties akin to a rectangular band-pass function (SI Note 4). The NA range of the band-pass filter (NA₁–NA₂) corresponds to 0.43–0.6 NA (red), 0.93–0.99 NA (green), and 0–0.8 NA (blue), respectively. Notably, the transmission coefficient phase at 650 nm wavelength covers a range of 2π but has few effects on edge detection due to the jump corresponding to the filtered-out information, showing greater design freedom compared with the differentiation operators. The simulation results in Fig. 2c demonstrate the ability of the film to perform selective image processing for targets of different sizes at these three operating wavelengths. The resolution (d) of the optical system follows the Rayleigh criterion ($d = 0.61\lambda/\text{NA}$), where a higher numerical aperture leads to improved resolution. Based on the NA range of the band-pass filter (NA₁–NA₂), selective denoising is achievable for

targets smaller than $0.35\ \mu\text{m}$ under blue light. Additionally, selective edge detection can be performed for targets ranging from 327–348 nm under green light and 660–922 nm under red light, with corresponding theoretical edge detection resolutions of 327 nm for green light and 660 nm for red light.

Fabrication and characterization of the processor

We fabricated a two-inch MDM film processor by using ion-assisted deposition (Fig. 2d) (Methods) to experimentally validate the design. To acquire the transmission spectrum of the sample as a function of the wavevector ($|t(k_r)|^2$), we conduct measurements using an ellipsometer at three operating wavelengths. The experimental transmission spectra under *p*-polarized light are shown in Fig. 2e, the trend of the measured spectra aligns well with the simulation. The measured transmission spectra for *s*-polarized light are shown in SI Note 5. It is worth noting that the measured transmission spectra under *s* polarization exhibit a similar trend to those under *p* polarization. Therefore, our processor demonstrates consistent image processing performance experimentally under *p*-polarized, *s*-polarized, and unpolarized light.

To validate the concept and quantify the resolution, we conducted imaging experiments on rectangular targets of various sizes, including their noise-perturbed counterparts, which allow us to characterize the processor's capabilities in selective edge detection and denoising. These experiments are carried out using a standard microscope imaging setup (Methods, SI Note 6). The target employed here is a chromium (Cr) mask containing various types of structures (Methods). The imaging results within the red and green dashed boxes in Fig. 3a demonstrate the processor's selective edge detection for targets of different sizes. The processor corresponds to edge detection sizes ranging from 1.6 to $2.3\ \mu\text{m}$ and 0.9 to $1.6\ \mu\text{m}$ under red ($\lambda = 650\ \text{nm}$) and green ($\lambda = 532\ \text{nm}$) light, respectively. The experimental edge detection resolution is measured as $0.9\ \mu\text{m}$ under green light and $1.6\ \mu\text{m}$ under red light. In addition, for the blue light ($\lambda = 466\ \text{nm}$), the blue dashed boxes in Fig. 3a illustrate the selective denoising capability of the processor when the rectangular structures are encircled by randomly distributed square-shaped noise of varying sizes. In this demonstration, noise is efficiently removed when the size is smaller than $0.3\ \mu\text{m}$, whereas noise of $0.4\ \mu\text{m}$ size is preserved.

Furthermore, we validate the size-selective imaging functionality by capturing images of various artificial shapes at different operating wavelengths. The pattern depicted in Fig. 3b, formed by the arrangement of letters 'A' to 'I', illustrates the output of selective edge detection for targets of various sizes. Specifically, precise edge detection is observed for the letters 'D' to 'I' or 'A' to 'D' under red and green light, respectively. Similarly, for a pattern of an optical field vector diagram labeled 'E–H–K', only the edges of the letter 'H' and its axis are enhanced under red light, while the letter 'E' and its axis are enhanced under green light. Additionally, we showcase the artistic effect of a 'Schrödinger staircase' by designing the widths of the upper and side surfaces of the staircase. The results exhibit a visually perceived downward or upward effect under red and green light, respectively.

In the presence of blue light, the processor also exhibits significant denoising effects for targets surrounded by noise of $0.3\ \mu\text{m}$ in size, randomly distributed around them (Fig. 3c). In particular, for the letters 'RGB' and 'Satellite', the noise surrounding the letters 'B' and 'Satellite' is notably reduced with the incorporation of the processor compared to the imaging results in bright field. We further challenge the processor with a QR code embedded in noise. In bright-field imaging, the information within the code is unreadable. However, with the processor, the denoising operation significantly facilitates scanning, enabling smart devices to clearly identify the corresponding official 'Chinese Academy of Sciences' website. These results underscore the potential applications of the processor in diverse fields such as space remote sensing and consumer electronics. More size-selective imaging

results with various shapes are available in SI Note 7. Compared to previous optical denoising methods^{46,49}, our approach demonstrates a significantly broader operational size range. Furthermore, it offers advantages in alignment and fabrication, leveraging the inherent properties of optical thin films. Moreover, the spatial distribution of noise has minimal impact on the denoising results, as demonstrated in the analysis presented in SI Note 8.

Size-selective imaging in biological cells and target recognition

Bio-microscopic imaging stands as one of the applications for edge detection and denoising, primarily because biological cells pose challenges in direct imaging due to their properties, such as high transparency. Expanding upon the aforementioned capacity for size-selective imaging of artificial targets, we further illustrate its potential in bio-imaging. Yeast cells (size $\sim 3\ \mu\text{m}$) and sperm cells of *Simmental* cattle (size $\sim 1\ \mu\text{m}$) are selected as recognition targets based on their respective sizes. Experimental results manifest that the edges of yeast cells are enhanced under red light, with no such effect observed under green light and in bright field (Fig. 4a). Conversely, sperm cells show no noticeable edge enhancement under red light or in bright field, but enhancement is observed under green light (Fig. 4b). Leveraging these properties, we proceed to validate the processor's ability to recognize different cells in a mixed state (Fig. 4c). Specifically, under red or green light in the mixed state, the processor retains its proficiency in precisely enhancing the edges of yeast or sperm cells while preserving the full images of another cell type. Given the size spread of biological cells, the edge enhancement effect is expected to be more pronounced in practical applications. As illustrated in bright field imaging results presented in Fig. 4c, a mixture of cells generates a considerable amount of noise, like living nanobacteria (size $< 1\ \mu\text{m}$), resembling black, worm-like bacteria distributed around the target cells. This distribution adversely affects imaging quality and cell recognition. The processor's size-selective denoising capability under blue light precisely addresses this issue. Initially, we test the processor on yeast cells with tiny silver particles (size of $0.3\ \mu\text{m}$) added, simulating a noisy environment. As experimental results reveal in Fig. 4d, the processor successfully removes all the silver particle noise, making the yeast cells clearer. Furthermore, we experimentally showcase real-time denoising of dynamic nanobacteria distributed both inside and outside the cells by selecting three distinct instants. The observations reveal that, in the absence of the processor, the bacteria inside and outside the cells severely impact imaging quality, with the position of the bacteria outside the cells undergoing drastic changes over time. Upon the addition of the processor, the images of the nanobacteria both inside and outside the cells are substantially smoothed, thereby improving microscopy imaging quality and cell recognition ability (Fig. 4e). The movies provided in the Supplementary Information showcase the real-time ability and effectiveness of the processor in size-selective edge detection (Supplementary Movies 1–4) and denoising (Supplementary Movies 5 and 6), thus underscoring its potential for practical applications in microscopy imaging and recognition.

Edge detection and denoising are also crucial in target recognition, serving as preprocessing steps to enhance the efficiency of subsequent recognition processes. To validate the impact of size-selective imaging on enhancing target recognition capability and efficiency, we construct a recognition model utilizing the Visual Geometric Group-16 (VGG-16)⁵⁰ (Methods). The artificial letters targeted illustrated in Fig. 3b, c are used as a case study. Leveraging its depth and small-size filters, the VGG-16 model exhibits outstanding efficacy in the letter recognition task. Our analysis compares the model's performance in recognizing targets with uniform and size-selective edge-detection processing, and without and with denoising processing.

For edge detection processing, we choose the letter 'D' as an example recognition target for both the uniform and size-selective

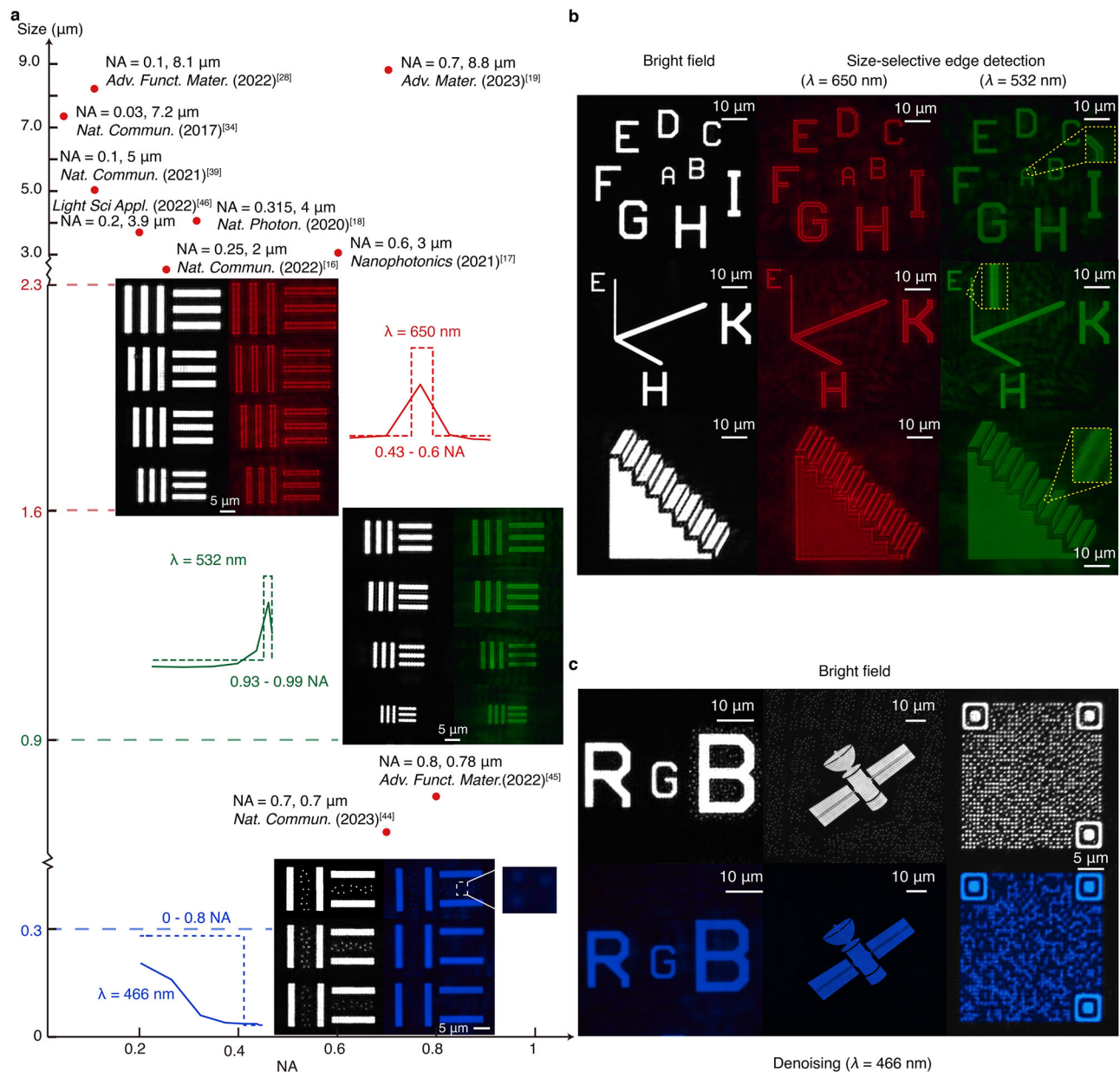


Fig. 3 | Characterizing size-selective imaging processor with artificial targets. **a** The experimental resolution of the size-selective imaging processor under red ($\lambda = 650 \text{ nm}$), green ($\lambda = 532 \text{ nm}$), and blue ($\lambda = 466 \text{ nm}$) lights. Under red and green lights, the processor demonstrates selective edge detection with sizes ranging from 1.6 to 2.3 μm and 0.9 to 1.6 μm , respectively. The insets depict the ideal optical transfer function of the image processor (dashed line) and the measured optical transfer function (solid line) at different operating wavelengths. Furthermore, under blue light, the processor efficiently engages in selective denoising of noise with a size less than 0.3 μm . Conversely, it preserves information when the noise size is 0.4 μm , as depicted in the inset that provides a magnified view of the retained noise. **b** The processor's size-selective edge detection capability with artificial targets of diverse letters and shapes under red and green lights, encompassing the letters 'A' to 'I', the optical field vector diagram denoted as 'E-H-K', and the

'Schrödinger staircase'. Specifically, under red light, the processor discernibly enhances the edges of letters 'D' to 'I' (size: 1.6–2.3 μm) and the letter 'H' with its axis (size of 2.2 μm). Conversely, under green light, the edge enhancement is observed on letters 'A' to 'D' (size: 1–1.6 μm) and the letter 'E' with its axis (size = 1.4 μm). Moreover, the 'Schrödinger staircase' achieves its artistic effect by designing the sizes of the upper (size of 2.3 μm) and side (size of 1.2 μm) surfaces. Consequently, the staircase imparts downward and upward visual effects under red and green light, respectively. The insets display locally magnified details of the structural features of the edges. **c** The size-selective denoising capability of the processor through the examination of the letters 'RGB', the 'satellite', and a QR code, with the size of the noise being 0.3 μm . In comparison to the results obtained through bright field imaging, the noise surrounding the letter 'B', the 'satellite', and the QR code is efficiently eliminated following the denoising.

edge-detected tasks, with representative datasets shown in Fig. 5a. We analyze the accuracy of the model in identifying letters from both tasks (Fig. 5a). After the size-selective edge detection, the recognition accuracy for the letter 'D' reaches 90% after 59 epochs, while requiring 112 epochs with uniform edge detection. These results demonstrate that size-selective edge detection accelerates model convergence and

reduces computational resources. For denoising processing, we recognize the letter 'B' with randomly distributed noise around it as well as the denoised one, respectively (Fig. 5b). The results show that the accuracy for the noisy image has low values around 40%, while after denoising the accuracy improves rapidly and converges to 96%. Furthermore, we conduct the ablation study and calculate the letter

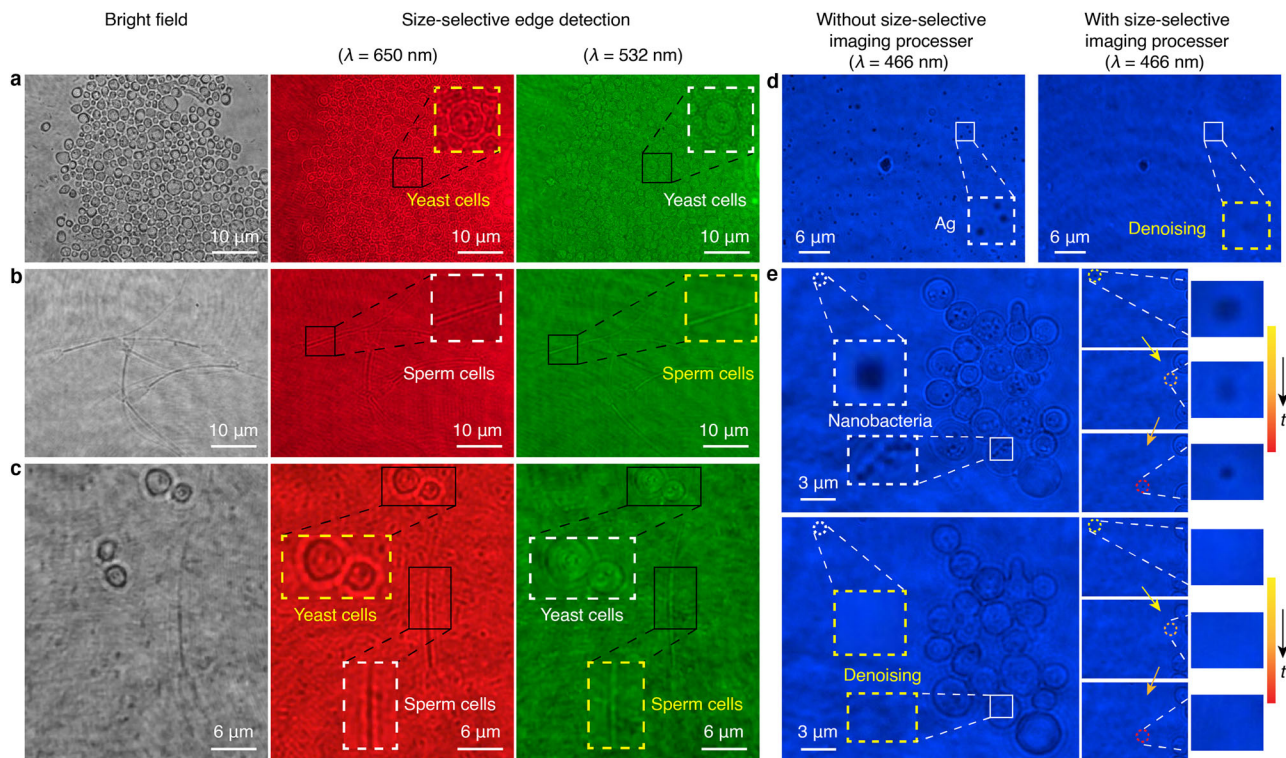


Fig. 4 | Characterization of real-time size-selective imaging in dynamic living biological cells. Size-selective edge detection experiments on yeast cells (size $\sim 3 \mu\text{m}$, (a)) and Simmental cattle sperm cells (size $\sim 1 \mu\text{m}$, (b)), both individually and in their mixed state (c). The edges of yeast cells are significantly enhanced under red light but remain nearly unchanged under green light. Conversely, the edges of sperm cells are significantly enhanced under green light but remain nearly unchanged under red light. The insets provide localized magnification results of selected cellular structures, highlighting effectively enhanced edges with yellow dashed boxes and the nearly unchanged parts with white dashed boxes. **d** Size-selective denoising experiments involving yeast cells surrounded by silver particles revealed a substantial impact on the imaging results in the absence of a processor. However, upon the insertion of the processor under identical imaging conditions,

the noise is markedly reduced. The white and yellow dashed boxes highlight zoomed-in views of the same areas in the image, one with the processor and one without, respectively. **e** Size-selective denoising experiments on dynamic living nanobacteria distributed both inside and outside yeast cells, with a significant and time-dependent change in the location of the living nanobacteria outside the cells. In the insets, localized enlargements of the same nanobacteria are presented, showcasing images without and with the addition of the processor, demarcated by a white dashed box and a yellow dashed box. These insets on the right offer zoomed-in views of the nanobacteria, displaying positional shifts and tracked images at three different moments, both without and with the processor in place, respectively. This highlights the processor's capability to remove noise and enhance the clarity of dynamic living nanobacteria imaging in real-time.

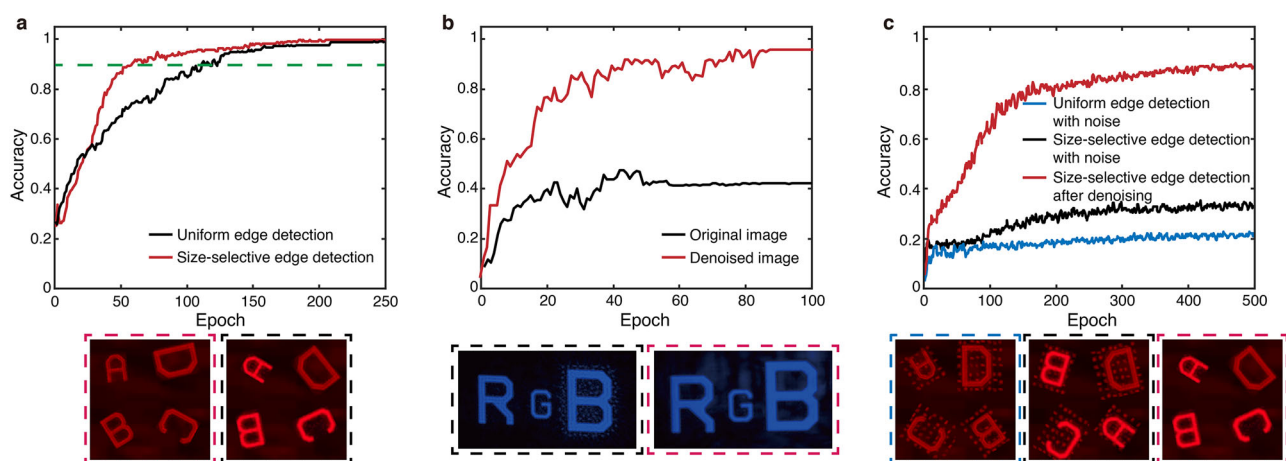


Fig. 5 | Improvement of target recognition performance by size-selective image processing and denoising. **a** The accuracy comparison between uniform edge-detected images (black line) and size-selective edge-detected images (red line) shows that size-selective edge detection achieves 90% recognition accuracy after 59 epochs, while uniform edge detection requires 112 epochs to reach the same level (green line). **b** Comparison of the recognition accuracy of the noisy images (black line) and the denoised images (red line). The accuracy of the noisy image is

about 40%, while the accuracy of the denoised image rapidly increases to 96%. This shows that denoising significantly improves recognition accuracy. **c** The recognition accuracy of the uniform edge-detected images with noise (blue line), size-selective edge-detected images with noise (black line), and size-selective edge-detected images after denoising (red line) is 90%, 33%, and 22.5%, respectively. Representative datasets for the tasks are shown at the bottom.

recognition accuracy in three cases: uniform edge detection with noise, size-selective edge detection with noise, and size-selective edge detection after denoising (Fig. 5c). After denoising and size-selective edge detection, the recognition accuracy reaches 90%. In contrast, for the other two cases, it is 33% and 22.5%, respectively, which further illustrates the improvement of size-selective imaging on target recognition performance. Upon analysis, size-selective edge detection and denoising exhibit several distinct advantages, including improved recognition accuracy and efficiency, as well as enhanced robustness. These findings highlight the potential role of size-selective image processing in enhancing the performance of target recognition.

Discussion

In conclusion, we have presented a size-selective image processor based on the target size by introducing a band-pass OTF. By employing a simple nonlocal thin film, we successfully attained real-time size-selective edge detection and denoising for both artificial targets and dynamic living biological cells. In addition, we demonstrated its potential application in the field of bio-imaging and target recognition. This approach achieves the highest experimental resolution ($\sim 0.9\ \mu\text{m}$) for edge detection based on spatial differentiation and realizes real-time dynamic denoising. Our processor can be seamlessly integrated with commercial imaging systems such as microscope objectives in an ultra-compact form, making it readily applicable to compact, alignment-free image processors.

Although the reported demonstrations are represented by three operating wavelengths, the number of operating wavelengths and the range of band-pass filtering can be extended arbitrarily, thanks to the continuous modulation properties of nonlocal thin films. In addition to the thin film schemes, our approach can be further implemented with photonic crystals⁵¹ and metasurfaces⁵², which can further enhance the performance of the OTF and provide greater operational flexibility for scenarios with specific target NAs. Furthermore, with the capability to tune wavevector domain responses, our approach may integrate high-dimensional photodetection⁵³, spatial compression^{54,55}, and distance ranging, thereby offering the potential for more integrated functionality.

Methods

Fabrication

For the sample used in the experiments shown in Figs. 3, 4, the MDM film sample is fabricated using ion-assisted deposition on a two-inch-diameter fused silica substrate. For the imaging experiments shown in Fig. 3 and SI Note 7, the target mask is a chromium (Cr) mask fabricated by electron beam (E-beam), which includes a variety of letters and shapes with different sizes. For the imaging experiments shown in Fig. 4, the yeast cells and Simmental's sperm cells are commercially available.

Experimental setup

For imaging experiments (Figs. 3, 4), the setup we demonstrate is a common microscopy imaging setup. The light inputs are generated by a supercontinuum light source (YSL SC-PRO-7) coupled with an acoustic-optical tunable filter (YSL AOTF0019). The chrome mask is illuminated by an expanded laser beam by a beam expander (Thorlabs GBE05-A 5 \times), then a magnified processed image of the test object is then obtained with the combination of an objective (Olympus UPLXAPO Objective 40 \times -0.95NA), a tube lens, and a visible camera (Thorlabs CS165CU/M), as shown in SI Note 6.

Visual Geometry Group-16 (VGG-16) model and training details

In Fig. 5, the VGG-16 model employed for letter recognition is depicted, featuring a convolutional neural network architecture comprising three main components: Convolutional Blocks, Flatten Layer, and Fully Connected Layers. The Convolutional Blocks consist of five consecutive blocks, each incorporating multiple convolutional layers

utilizing a 3×3 kernel size, accompanied by a max pooling layer. These convolutional layers employ filters on the input image to extract features, including edges, shapes, and textures. The convolution layers have a stride of 1, and the padding is set to 'same', ensuring that these layers exclusively capture features from the image without downsizing. Additionally, five max-pooling layers are strategically interspersed among the convolutional layers. The function of these layers is to downsample the feature maps of the preceding layer, reducing their size to decrease computational complexity while retaining the most crucial information. Following the Convolutional Blocks, the output is processed through a Flattening Layer. This layer aids in reshaping the multidimensional output of the previous Convolutional Block into a one-dimensional vector, facilitating a seamless transition to the Fully Connected Layers. The architecture concludes with three Fully Connected Layers. The initial two layers employ Rectified Linear Unit (ReLU) activation functions and a dropout rate of 0.5 to introduce nonlinearity into the model and alleviate the vanishing gradient problem. The final layer employs a *softmax* activation function to establish a probability distribution across different classes. The class with the highest probability is selected as the final output of the model. The model adopts default parameters from ImageNet, including blocks 1 to 5. The model is implemented in TensorFlow version 2.10 using the Python programming language and trained on a machine with an H3C UniServer R4900 G5 CPU and an NVIDIA A30 GPU. The categorical cross-entropy loss is used to train all the connected artificial-neuron layers. The learning rate is initially 3×10^{-6} . For the other parameters of the Adam optimizer, the default TensorFlow settings are used.

Data availability

The data that support the plots in this paper are available from the corresponding authors. Source data are provided with this paper.

Code availability

The code for the letter recognition task in this study is freely available at https://github.com/YandongFan/Nonlocal_Flat_Optics_for_Size-Selective_Image_Processing_and_Denoising.git.

References

1. Solli, D. R. & Jalali, B. Analog optical computing. *Nat. Photon.* **9**, 704–706 (2015).
2. Woods, D. & Naughton, T. J. Photonic neural networks. *Nat. Phys.* **8**, 257–259 (2012).
3. Lin, X. et al. All-optical machine learning using diffractive deep neural networks. *Science* **361**, 1004–1008 (2018).
4. Kwon, H., Arbabi, E., Kamali, S. M., Faraji-Dana, M. & Faraon, A. Single-shot quantitative phase gradient microscopy using a system of multifunctional metasurfaces. *Nat. Photon.* **14**, 109–114 (2020).
5. Wang, X. et al. Advances in information processing and biological imaging using flat optics. *Nat. Rev. Electr. Eng.* **1**, 391–411 (2024).
6. Monticone, F. & Alù, A. Metamaterial, plasmonic and nanophotonic devices. *Rep. Prog. Phys.* **80**, 036401 (2017).
7. Caulfield, H. J. & Dolev, S. Why future supercomputing requires optics. *Nat. Photon.* **4**, 261–263 (2010).
8. Gonzalez, R. C. & Woods, R. E. *Digital Image Processing*. (Prentice Hall Press, 2008).
9. Wang, X., Deng, Y. & Duan, H. Edge-based target detection for unmanned aerial vehicles using competitive Bird Swarm Algorithm. *Aerosp. Sci. Technol.* **78**, 708–720 (2018).
10. Gebäck, T. & Koumoutsakos, P. Edge detection in microscopy images using curvelets. *BMC Bioinforma.* **10**, 75 (2009).
11. Assidiq A. A., Khalifa O. O., Islam M. R., Khan S. Real time lane detection for autonomous vehicles. *2008 International Conference on Computer and Communication Engineering*, Kuala Lumpur, Malaysia, 82–88 <https://doi.org/10.1109/ICCCE.2008.4580573> (2008).

12. Haralick, R. M. & Shapiro, L. G. *Computer and Robot Vision*. (Addison-Wesley Longman Publishing, 1991).
13. Shastri, K. & Monticone, F. Nonlocal flat optics. *Nat. Photon.* **17**, 36–47 (2022).
14. Goodman, J. W. *Introduction to Fourier Optics*. (Roberts and Company Publishers, 2005).
15. Silva, A. et al. Performing mathematical operations with metamaterials. *Science* **343**, 160–163 (2014).
16. Liu, Y., Huang, M., Chen, Q. & Zhang, D. Single planar photonic chip with tailored angular transmission for multiple-order analog spatial differentiator. *Nat. Commun.* **13**, 7944 (2022).
17. Jin, C. & Yang, Y. Transmissive nonlocal multilayer thin film optical filter for image differentiation. *Nanophotonics* **10**, 3519–3525 (2021).
18. Zhou, Y., Zheng, H., Kravchenko, I. I. & Valentine, J. Flat optics for image differentiation. *Nat. Photon.* **14**, 316–323 (2020).
19. Liang, X. et al. All-optical multiplexed meta-differentiator for tri-mode surface morphology observation. *Adv. Mater.* **35**, 2301505 (2023).
20. Tanriover, I., Dereshgi, S. A. & Aydin, K. Metasurface enabled broadband all optical edge detection in visible frequencies. *Nat. Commun.* **14**, 6484 (2023).
21. Cordaro, A. et al. Solving integral equations in free space with inverse-designed ultrathin optical metagratings. *Nat. Nanotechnol.* **18**, 365–372 (2023).
22. Cotrufo, M., Arora, A., Singh, S. & Alù, A. Dispersion engineered metasurfaces for broadband, high-NA, high-efficiency, dual-polarization analog image processing. *Nat. Commun.* **14**, 7078 (2023).
23. Zangeneh-Nejad, F., Sounas, D. L., Alù, A. & Fleury, R. Analogue computing with metamaterials. *Nat. Rev. Mater.* **6**, 207–225 (2021).
24. Cotrufo, M., Singh, S., Arora, A., Majewski, A. & Alù, A. Polarization imaging and edge detection with image-processing metasurfaces. *Optica* **10**, 1331 (2023).
25. Zhou, J. et al. Optical edge detection based on high-efficiency dielectric metasurface. *Proc. Natl Acad. Sci. USA* **116**, 11137–11140 (2019).
26. Zhou, J. et al. Metasurface enabled quantum edge detection. *Sci. Adv.* **6**, eabc4385 (2020).
27. Zhou, J. et al. Two-dimensional optical spatial differentiation and high-contrast imaging. *Natl Sci. Rev.* **8** (2020).
28. Zhou, J. et al. Nonlinear computational edge detection metalens. *Adv. Funct. Mater.* **32**, 2204734 (2022).
29. Kwon, H., Sounas, D., Cordaro, A., Polman, A. & Alù, A. Nonlocal metasurfaces for optical signal processing. *Phys. Rev. Lett.* **121**, 173004 (2018).
30. Guo, C., Xiao, M., Minkov, M., Shi, Y. & Fan, S. Photonic crystal slab Laplace operator for image differentiation. *Optica* **5**, 251 (2018).
31. Guo, C., Xiao, M., Minkov, M., Shi, Y. & Fan, S. Isotropic wavevector domain image filters by a photonic crystal slab device. *J. Opt. Soc. Am. A* **35**, 1685 (2018).
32. Zhu, T. et al. Plasmonic computing of spatial differentiation. *Nat. Commun.* **8**, 15391 (2017).
33. Mohammadi, H., Akbari, M. & Khavasi, A. All optical divergence and gradient operators using surface plasmon polaritons. *Opt. Express* **30**, 17806–17823 (2022).
34. Zhu, T., Huang, J. & Ruan, Z. Optical phase mining by adjustable spatial differentiator. *Adv. Photonics* **2**, 016001 (2020).
35. Zhu, T. et al. Generalized spatial differentiation from the spin Hall effect of light and its application in image processing of edge detection. *Phys. Rev. Appl.* **11**, 034043 (2019).
36. He, S. et al. Spatial differential operation and edge detection based on the geometric spin Hall effect of light. *Opt. Lett.* **45**, 877–880 (2020).
37. Zhu, T. et al. Topological optical differentiator. *Nat. Commun.* **12**, 680 (2021).
38. Liu, Q., Ayub, M. A., Ahmat Ruslan, F., Ab Patar, M. N. A. & Abdul-Rahman, S. An improved mask R-CNN algorithm for high object detection speed and accuracy. *Commun. Comput. Inf. Sci.* **1771** (2023).
39. Cordaro, A. et al. High-index dielectric metasurfaces performing mathematical operations. *Nano Lett.* **19**, 8418–8423 (2019).
40. Fang, Y. & Ruan, Z. Optical spatial differentiator for a synthetic three-dimensional optical field. *Opt. Lett.* **43**, 5893–5896 (2018).
41. Komar, A. et al. Edge detection with mie-resonant dielectric metasurfaces. *ACS Photonics* **8**, 864–871 (2021).
42. Kwon, H., Cordaro, A., Sounas, D., Polman, A. & Alù, A. Dual-polarization analog 2D image processing with nonlocal metasurfaces. *ACS Photonics* **7**, 1799–1805 (2020).
43. Wang, Z. et al. Single-layer spatial analog meta-processor for imaging processing. *Nat. Commun.* **13**, 2188 (2022).
44. Wang, X. et al. Single-shot isotropic differential interference contrast microscopy. *Nat. Commun.* **14**, 2063 (2023).
45. Kim, Y. et al. Spiral metalens for phase contrast imaging. *Adv. Funct. Mater.* **32**, 2106050 (2022).
46. Fu, W. et al. Ultracompact meta-imagers for arbitrary all-optical convolution. *Light Sci. Appl.* **11**, 62 (2022).
47. Liu, S., Liu, D., Muhammad, K. & Ding, W. Effective template update mechanism in visual tracking with background clutter. *Neuro-computing* **458**, 615–625 (2021).
48. Abdollahramezani, S., Hemmatyar, O. & Adibi, A. Meta-optics for spatial optical analog computing. *Nanophotonics* **9**, 4075–4095 (2020).
49. Işıl, Ç. et al. All-optical image denoising using a diffractive visual processor. *Light Sci. Appl.* **13**, 43 (2024).
50. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. Preprint at arXiv <https://doi.org/10.48550/arXiv.1409.1556>.
51. Fan, S. & Joannopoulos, J. D. Analysis of guided resonances in photonic crystal slabs. *Phys. Rev. B* **65** (2002).
52. Yu, N. & Capasso, F. Flat optics with designer metasurfaces. *Nat. Mater.* **13**, 139–150 (2014).
53. Fan, Y. et al. Dispersion-assisted high-dimensional photodetector. *Nature* **630**, 77–83 (2024).
54. Chen, A. & Monticone, F. Dielectric nonlocal metasurfaces for fully solid-state ultrathin optical systems. *ACS Photonics* **8**, 1439–1447 (2021).
55. Reshef, O. et al. An optic to replace space and its application towards ultra-thin imaging systems. *Nat. Commun.* **12**, 3512 (2021).

Acknowledgements

W.L., Y.F., F.Z., G.X. and B.L. were supported by the National Natural Science Foundation of China (grant Nos. 62134009, 62121005). M.K., H.G. and A.A. were supported by the Simons Foundation.

Author contributions

W.L., A.A. and S.K.C. conceived the project. W.L., S.K.C., Y.F., C.H., M.K., H.G., G.X. and C.J. performed the theoretical calculations and design. S.K.C., C.H. and Y.W. prepared the samples. S.K.C., C.H., Y.F. and C.J. performed the measurements. Y.F. and S.G. developed the program for applying the neural network. W.L., S.K.C., Y.F., C.J., M.K., H.G., C.H., F.Z., B.L., T.B. and A.A. discussed and analyzed the results. W.L., A.A., S.K.C., Y.F., C.J., M.K. and H.G. wrote and revised the paper. W.L. and A.A. supervised the project.

Competing interests

A patent (CN 116719111 B) has been granted related to this work by W.L., C.J., C.H. and S.K.C. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-59765-4>.

Correspondence and requests for materials should be addressed to Andrea Alù or Wei Li.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025