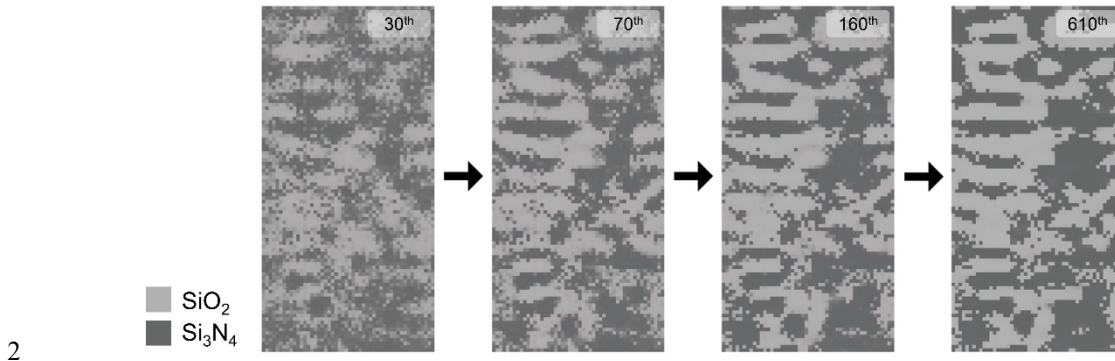


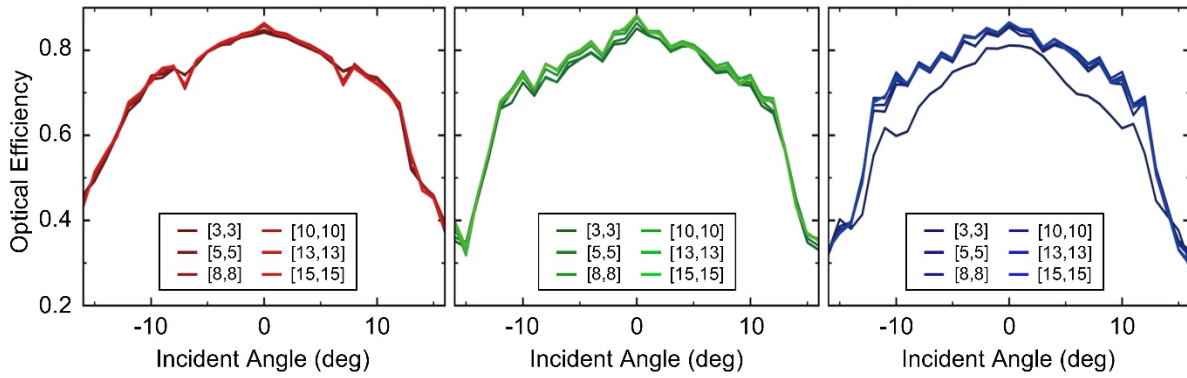


1 **Supplementary Section 1. Design Evolution Through Optimization**



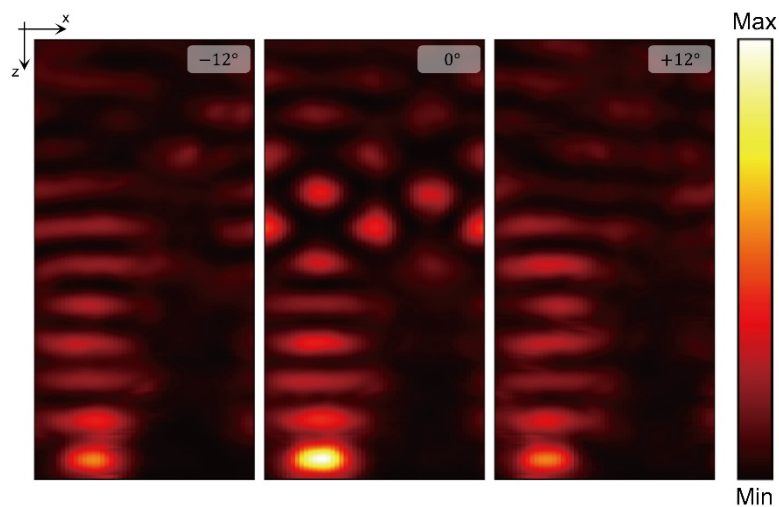
3 **Figure S1.** Evolution of the optimized structure at different iterations, showing the progressive changes in  
4 refractive index distribution. The structure gradually transitions from a random initial configuration toward a more  
5 binarized form composed of SiO<sub>2</sub> (light gray) and Si<sub>3</sub>N<sub>4</sub> (dark gray). Each vertical cross-sectional view (*x-z* plane)  
6 corresponds to the upper part of the section passing through the center of the R-G subpixel.

## 1 Supplementary Section 2. RCWA Convergence Verification



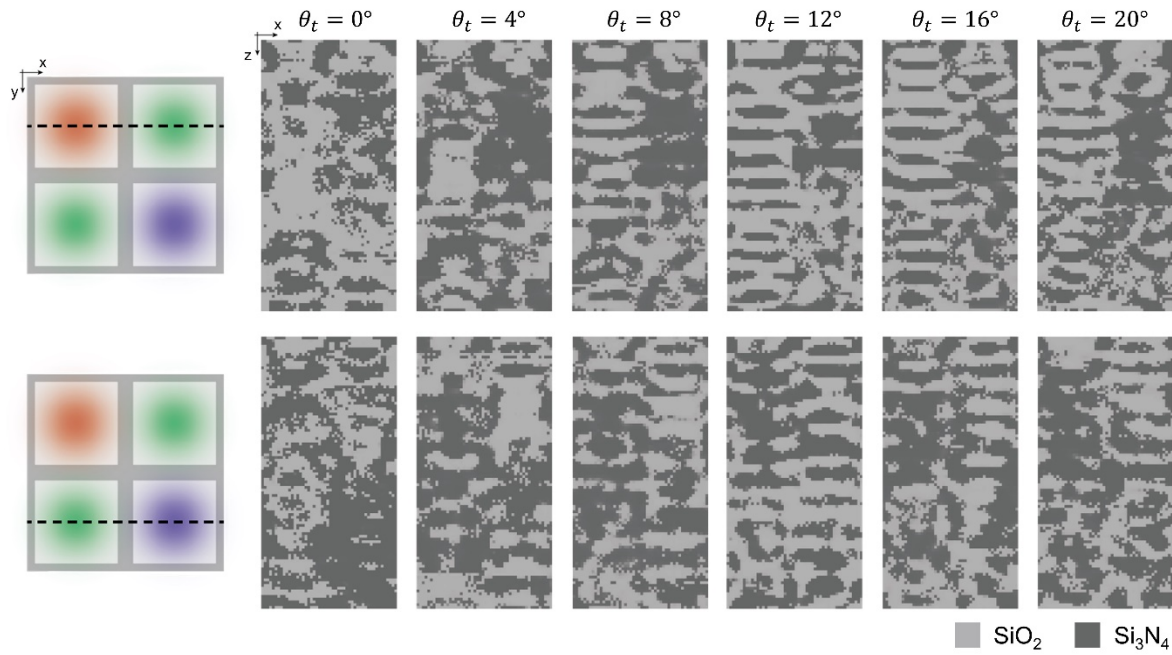
2  
3 **Figure S2.** Optical efficiency of the optimized color router calculated using different Fourier truncation orders  
4 ranging from [3,3] to [15,15] for red (left), green (center), and blue (right) wavelengths. The results show that  
5 convergence begins from [5,5], and [8,8] provides quantitatively stable results for all channels. These findings  
6 confirm the numerical convergence and validate the choice of truncation order used during optimization. The  
7 absolute relative error in optical efficiency between the two Fourier truncation orders [8,8] and [15,15] is 1.2 %  
8 on average across RGB channels.

1 **Supplementary Section 3. Electric Field Distribution in Design Region**



3 **Figure S3.** Electric field intensity distributions within the design region under three different incident angles:  
4  $-12^\circ$ ,  $0^\circ$ , and  $+12^\circ$ . The field distributions are obtained by illuminating the optimized structure ( $\theta_t = 12^\circ$ ,  
5  $N_L = 100$ ) with 450 nm plane wave at each incident angle. Each distribution is taken along the same vertical  
6 cross-section as shown in the structural views, corresponding to the upper part of the central R-G subpixel.

1 **Supplementary Section 4. Optimized Structures with Different Target Acceptance Angles**

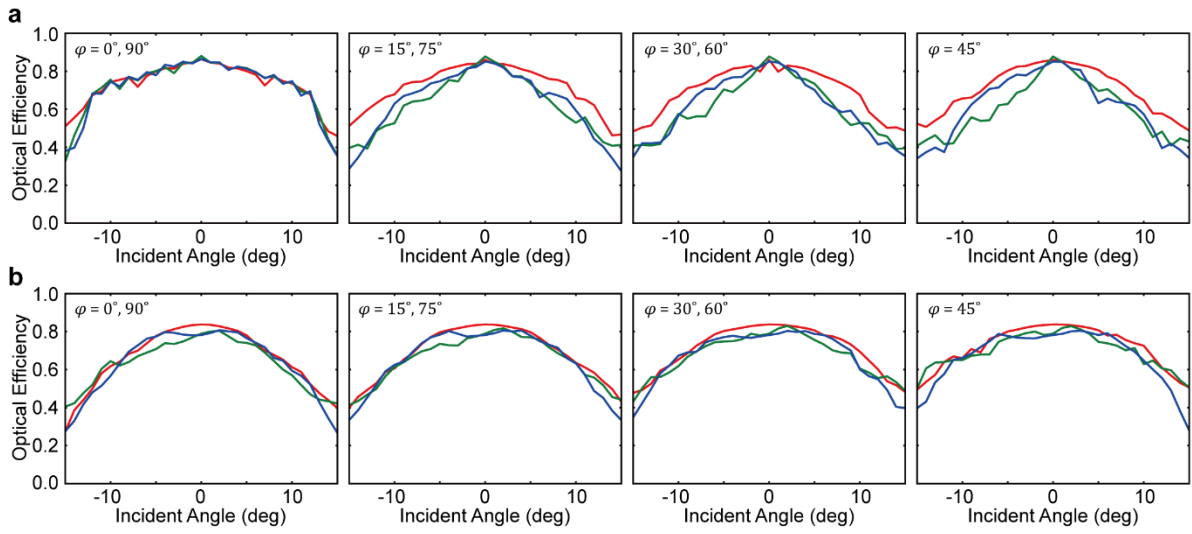


2

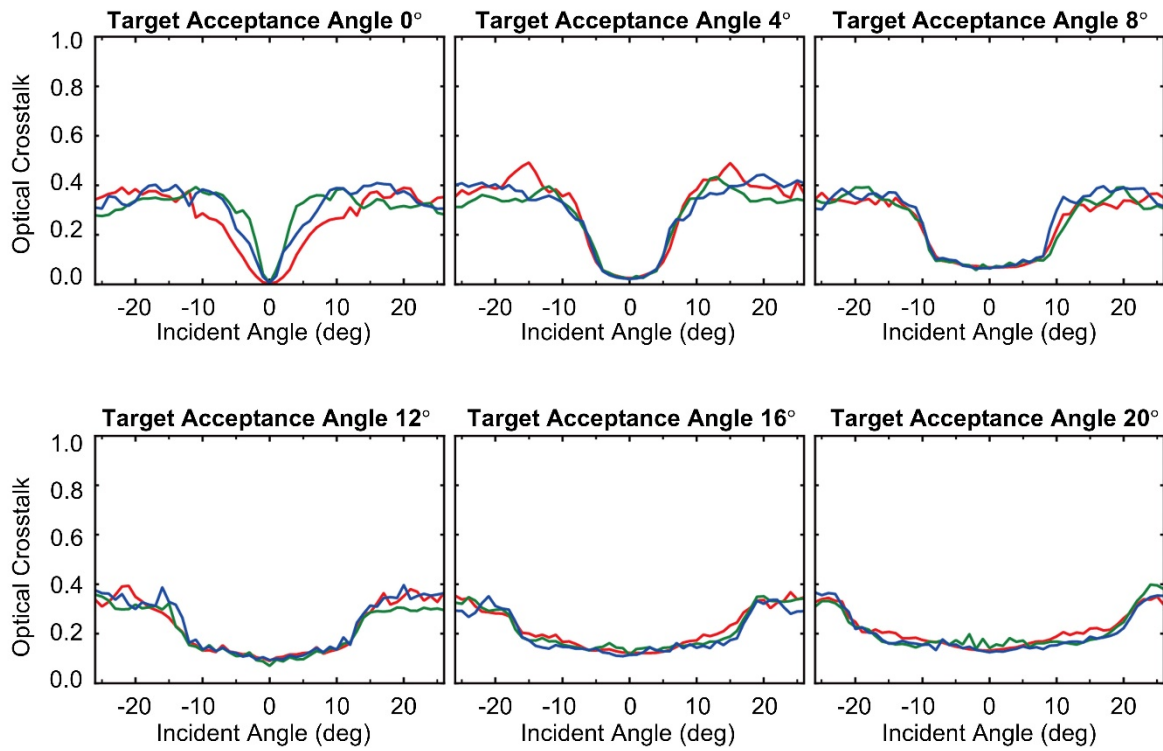
3 **Figure S4.** Optimized structures for each target acceptance angle  $\theta_t \in \{0^\circ, 4^\circ, 8^\circ, 12^\circ, 16^\circ, 20^\circ\}$ . Each vertical  
4 cross-sectional view ( $x$ - $z$  plane) corresponds to the dashed line locations shown in the top and bottom subpixel  
5 layers.

6

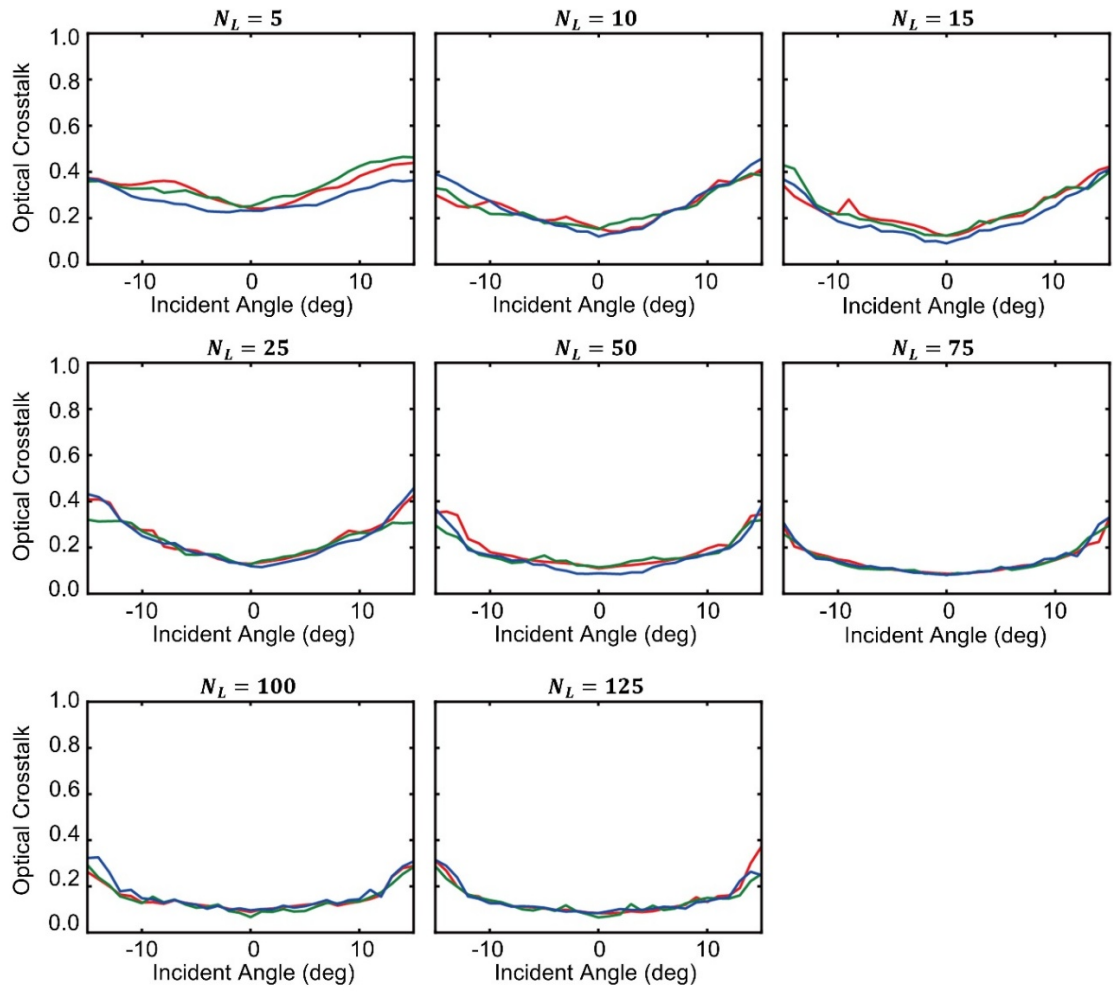
1 **Supplementary Section 5. Performance for Different Azimuthal Angles**



1 **Supplementary Section 6. Optical Crosstalk for Different Target Acceptance Angles and**  
2 **Number of Layers**



3  
4 **Figure S6.** Optical crosstalk for optimized structure with different target acceptance angles  $\theta_t$  from 0° to 20°.

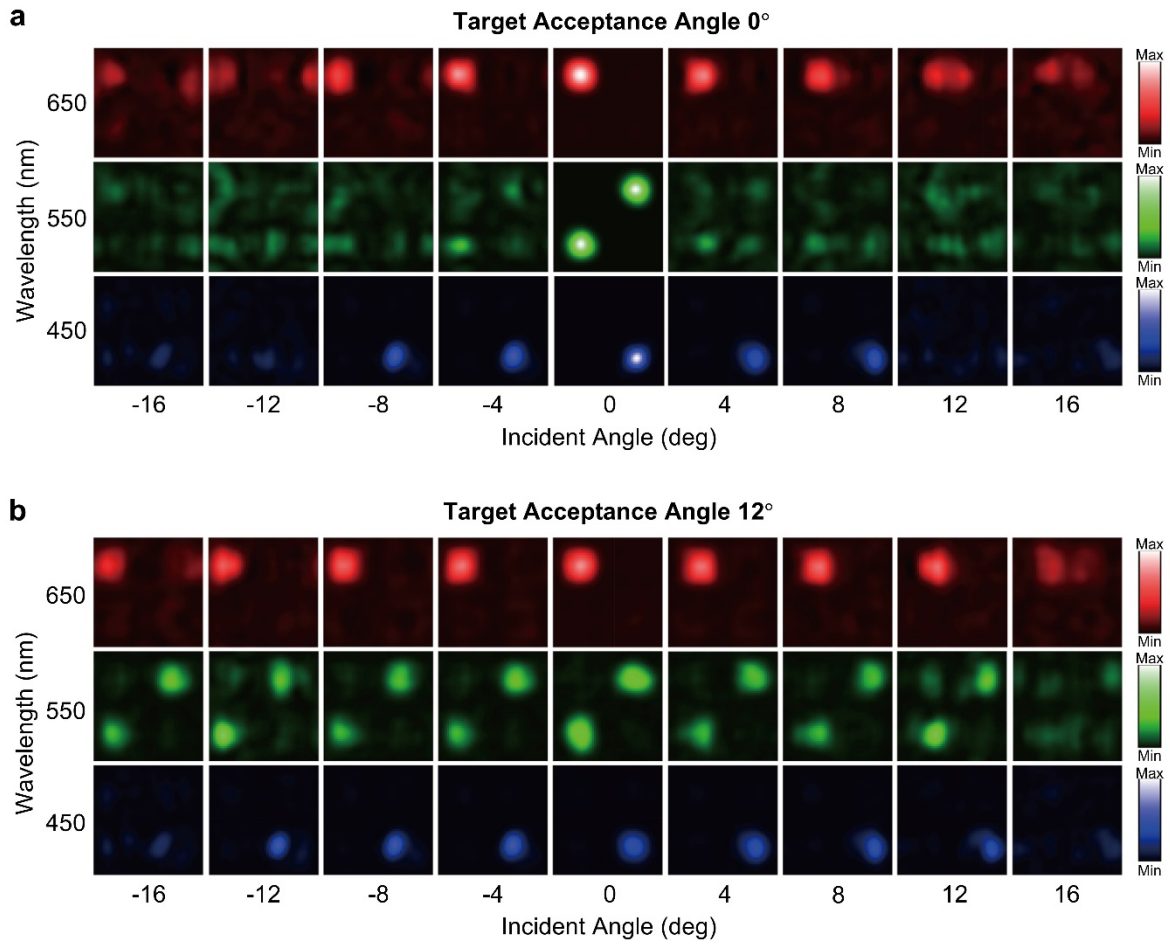


1

2 **Figure S7.** Optical crosstalk for optimized structure with number of layers  $N_L$  from 5 to 125.

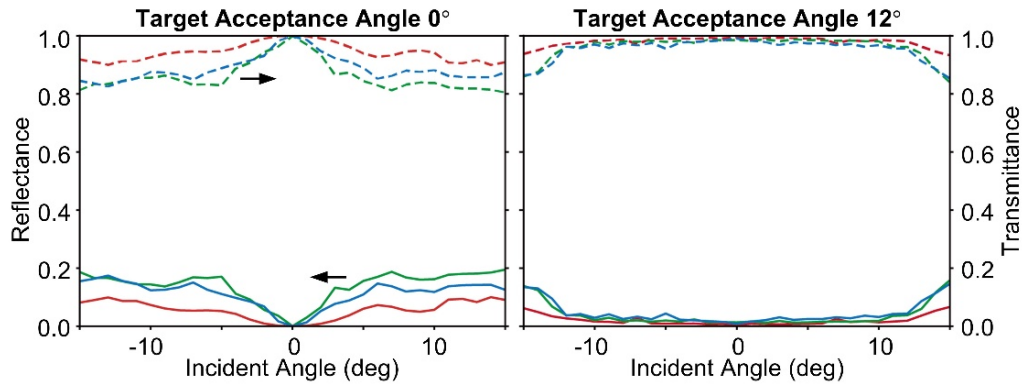
3

1 **Supplementary Section 7. Extended Comparison: With vs. Without Oblique Incidence**  
2 **Consideration**



3  
4 **Figure S8.** Extended comparison of the performance between structures optimized with and without consideration  
5 of oblique incidence. The field distributions of the z-component of the Poynting vector at the top of the  
6 photodetector are shown for a wider range of wavelengths and incident angles than those presented in the main  
7 text. In both cases, the structures were optimized with  $N_L = 100$  with (a)  $\theta_t = 0^\circ$  and (b)  $\theta_t = 12^\circ$ .

8

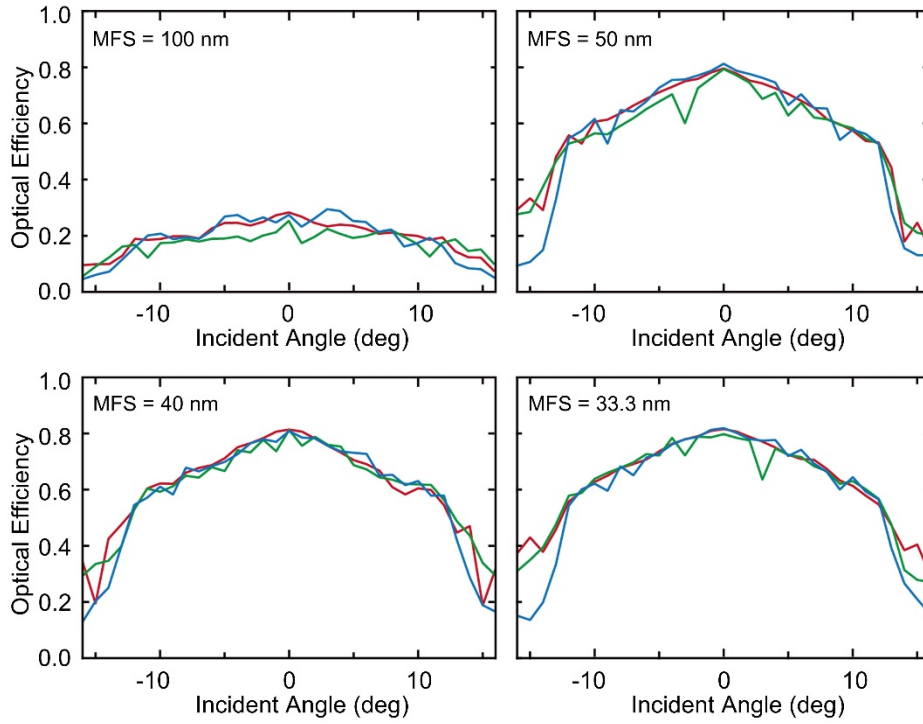


1

2 **Figure S9.** Spectra of each pixel at target acceptance angles of 0° (left) and 12° (right). The solid lines represent  
 3 the reflection (R) spectra, while the dashed lines show the transmission. The suppression of reflection within the  
 4 target acceptance angle demonstrates optimization's effectiveness in enhancing optical routing efficiency across  
 5 varying incident angles.

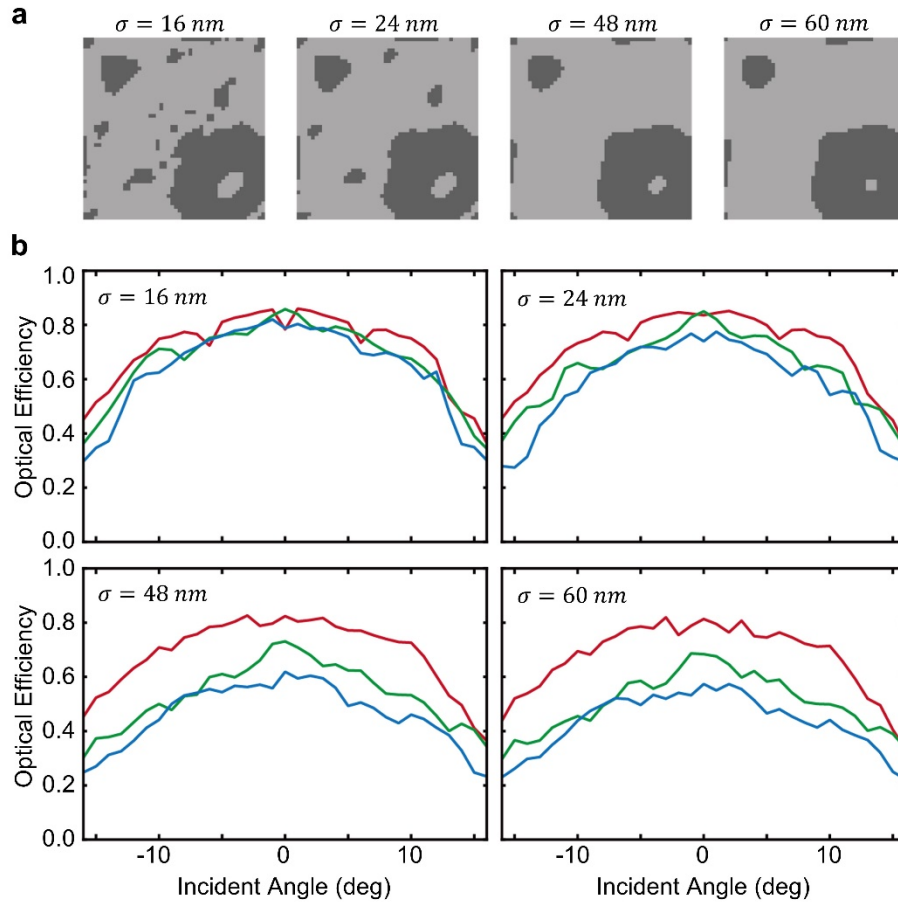
6

1 **Supplementary Section 8. Fabrication Feasibility Analysis**



2

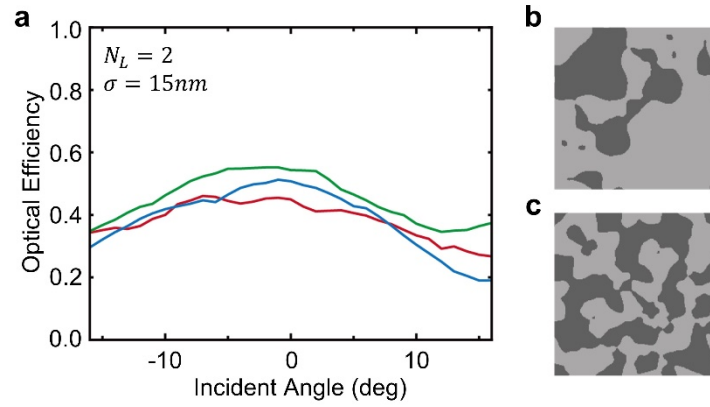
3 **Figure S10.** Optical efficiency as a function of incident angle for different minimum feature sizes (MFS) ranging  
4 from 100 nm to 33.3 nm. The average optical efficiency within  $\pm 12^\circ$  is 20.1% (MFS = 100 nm), 67.8% (MFS =  
5 50 nm), 70.1% (MFS = 40 nm), and 70.8% (MFS = 33.3 nm), respectively.



1

2 **Figure S11.** Optical efficiency as a function of incident angle for different minimum feature sizes (MFS),  
 3 implemented by Gaussian filtering. (a) Structures obtained by applying Gaussian filters with different standard  
 4 deviations ( $\sigma$ ) to the top layer. The light gray corresponds to  $\text{SiO}_2$ , and the dark gray to  $\text{Si}_3\text{N}_4$ . (b) Optical efficiency  
 5 of the final optimized design as a function of incident angle, evaluated after applying Gaussian filters with various  
 6 standard deviations ( $\sigma$ ).

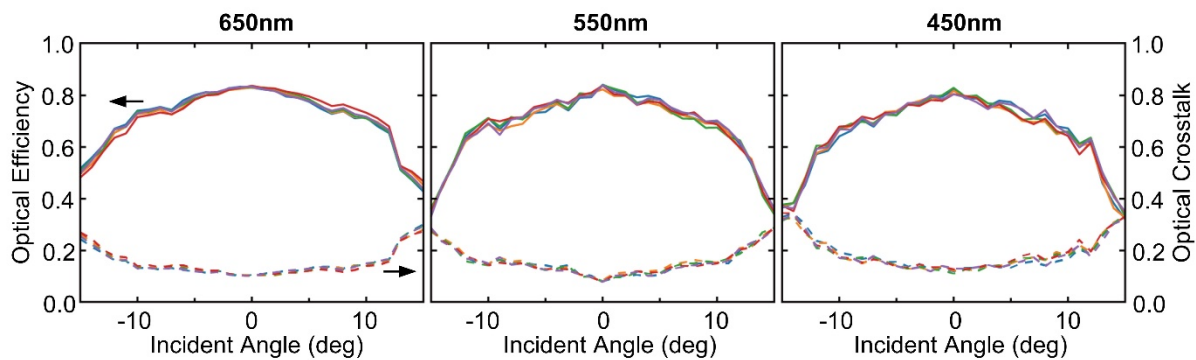
7



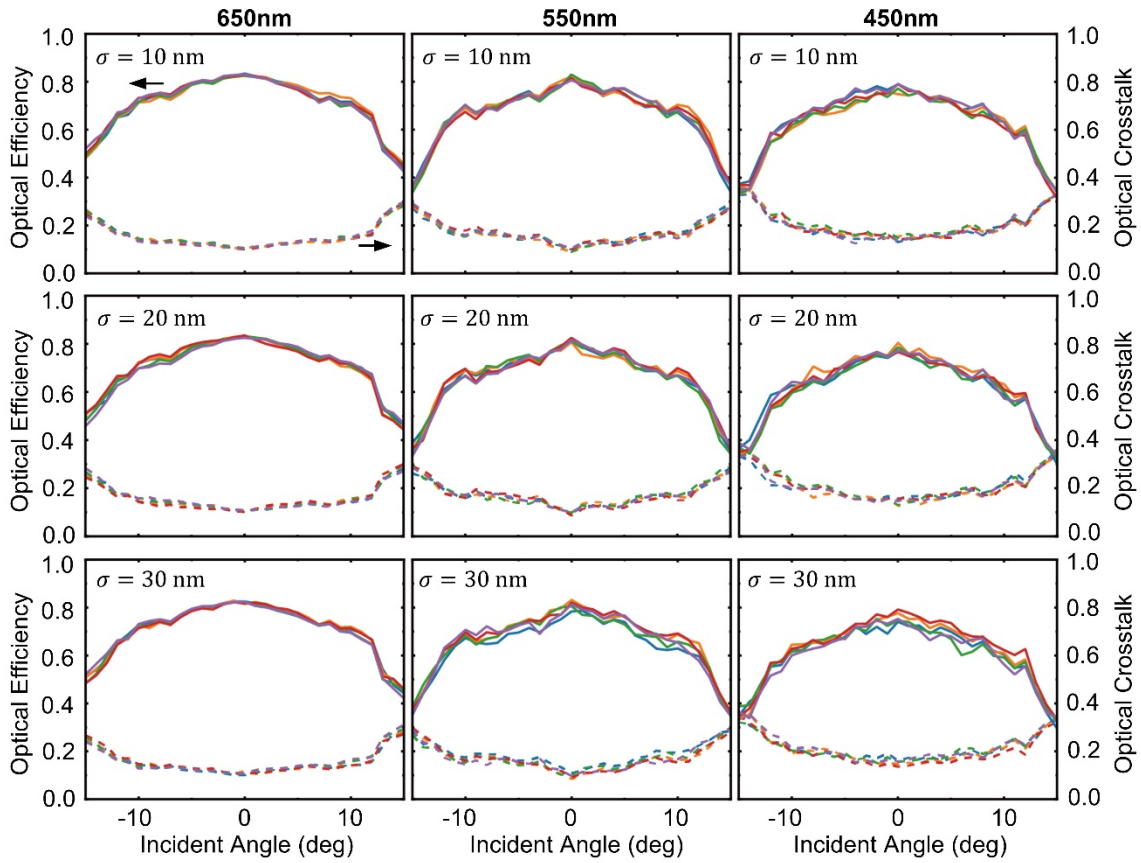
1

2 **Figure S12.** Performance of the optimized two-layer color router after applying a Gaussian filter with a standard  
 3 deviation of 15 nm. (a) Optical efficiency as a function of incident angle. (b) Top layer and (c) bottom layer of  
 4 the Gaussian-filtered structure, where light gray denotes SiO<sub>2</sub> and dark gray denotes Si<sub>3</sub>N<sub>4</sub>.

5



1  
 2 **Figure S13.** Optical efficiency and optical crosstalk when each of the 100 layers is subjected to a potential one-  
 3 pixel (20 nm) lateral shift with a 1/3 probability. The five overlapping curves in each plot, corresponding to five  
 4 independent random trials, are nearly indistinguishable from one another and from the ideal case without errors.

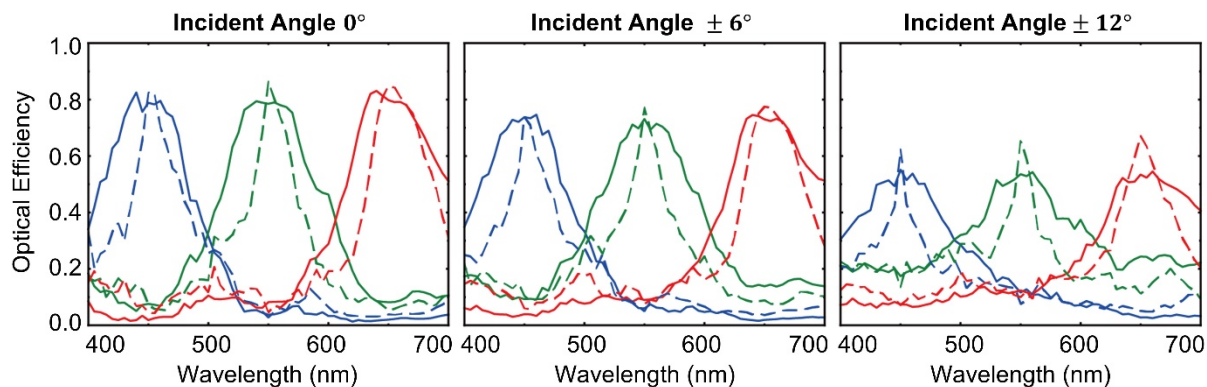


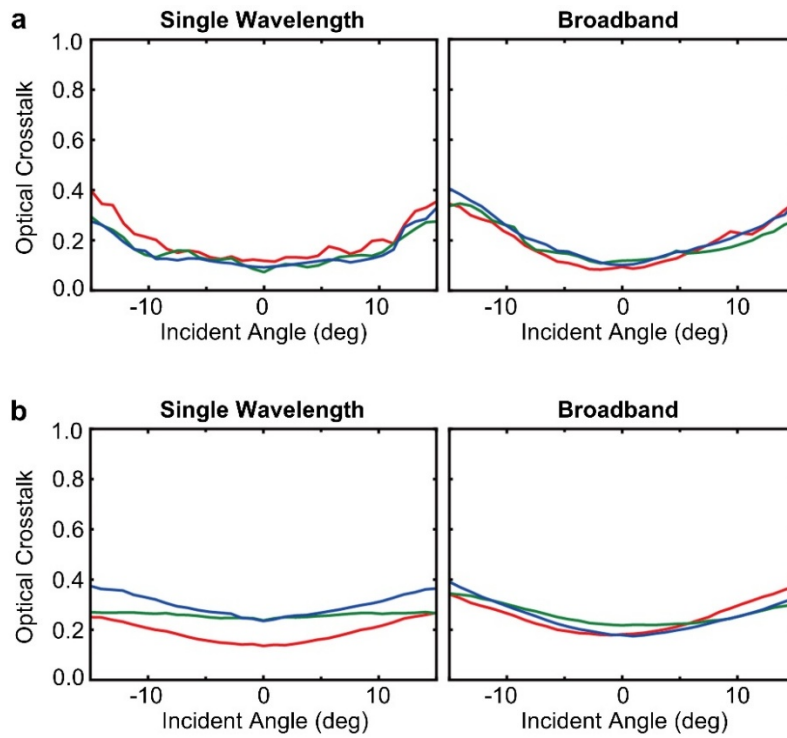
1

2 **Figure S14.** Optical efficiency (solid lines) and optical crosstalk (dashed lines) under a statistical misalignment  
 3 model where each layer is shifted according to a Gaussian distribution. The performance is evaluated for  
 4 increasing standard deviations ( $\sigma$ ) of 20 nm, 40 nm, and 60 nm, showing only minor degradation even at the  
 5 highest level of simulated error. For each  $\sigma$ , the Gaussian distribution corresponds to possible layer displacements  
 6 of up to two grid units (40 nm in maximum, with one unit representing 20 nm) along the x- and y-directions.

7

1 **Supplementary Section 9. Performance Comparison: Broadband vs. Single-Wavelength**  
2 **Optimization**



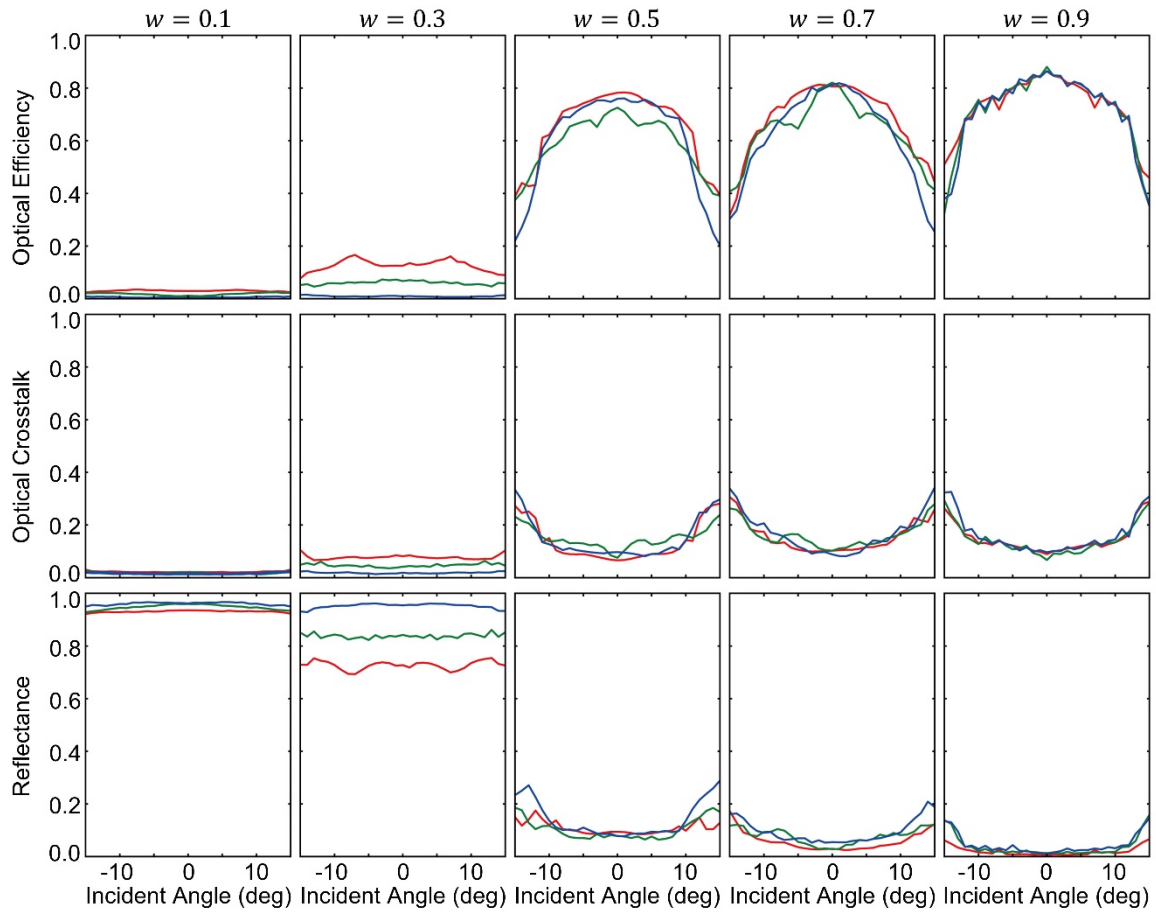


1

2 **Figure S16.** Optical crosstalk comparison of color routers optimized under broadband and single-wavelength  
 3 conditions. (a) Optical crosstalk at the R, G, and B center wavelengths. (b) Spectrally averaged optical crosstalk  
 4 of color routers optimized with single-wavelength and broadband conditions over the spectral ranges R (600–700  
 5 nm), G (500–600 nm), and B (400–500 nm).

6

1 **Supplementary Section 10. Impact of FoM Weighting on Optimization Outcome**



2

3 **Figure S17.** Optical performance comparison for devices optimized with different OE weights ( $w$ ). The plots  
4 show (top row) Optical Efficiency (OE), (middle row) Optical Crosstalk (OX), and (bottom row) Reflectance for  
5 each case. All optimizations are performed for a target acceptance angle of  $12^\circ$ .

6

## 1 Supplementary Section 11. Analysis of Interpixel Crosstalk

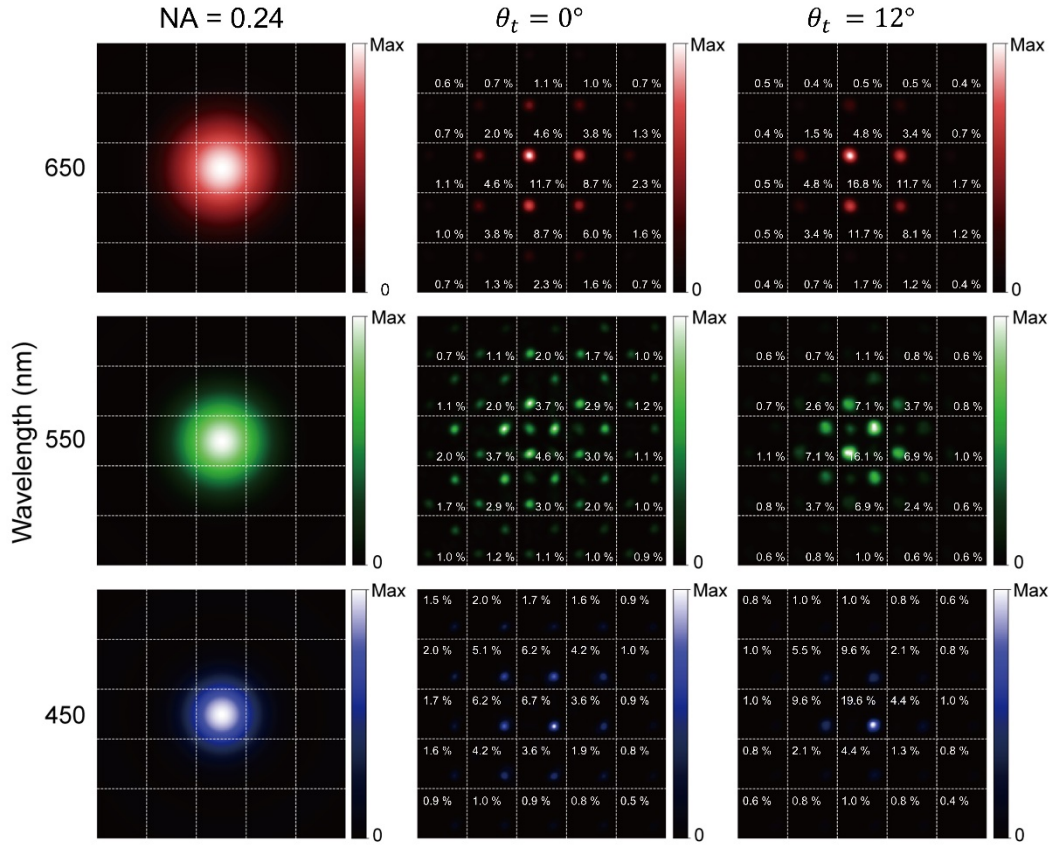
2 We propose a framework based on RCWA under periodic boundary conditions. As the RCWA framework  
3 inherently assumes periodicity, it is challenging to directly evaluate interpixel crosstalk. To achieve this within  
4 our RCWA framework, a quasi-Gaussian beam is implemented by coherently superimposing multiple plane waves  
5 with various polar ( $\theta$ ) and azimuthal ( $\varphi$ ) angles in-phase at the top surface center of the metasurface.

6 A technical challenge arises, as defining a 3D Gaussian beam with a consistent TE/TM polarization basis is  
7 problematic when  $\varphi$  varies. We address this by constructing x-polarized and y-polarized Gaussian beams  
8 independently. We account for the relative portion of  $k$ -space represented by each discrete incident condition  
9 when assigning weights. Since  $\varphi$  is uniformly sampled, the weighting along the  $\theta$ -direction is determined from  
10 the spacing between the midpoints of adjacent  $k \sin \theta$  values, ensuring that each sampled wave accurately  
11 represents its contribution to the overall angular spectrum.

12 The implemented localized field profiles of the quasi-Gaussian beams are shown in **Fig. S18** and **Fig. S19**. In this  
13 analysis, the azimuthal angle ( $\varphi$ ) is varied from  $-75^\circ$  to  $90^\circ$  in  $15^\circ$  steps, while the polar angle used for  
14 decomposition is swept over two ranges:  $[-12^\circ, 12^\circ]$  and  $[-20^\circ, 20^\circ]$ , both sampled at  $4^\circ$  intervals. The beam  
15 waist of the localized field varies depending on the polar-angle range and wavelength. These correspond to  
16 numerical apertures (NAs) of approximately 0.24 and 0.38, respectively, considering the  $k$ -space sampling. To  
17 evaluate the interpixel crosstalk for these input beams, we arrange the image sensor pixels in a  $5 \times 5$  array. The  
18 responses are averaged between Ex- and Ey-polarized Gaussian beams.

19 **Figure S18** compares the optical responses between the structures optimized for  $\theta_t = 0^\circ$  and  $\theta_t = 12^\circ$ , using a  
20 quasi-Gaussian beam decomposed over polar angles  $\theta \in [-12^\circ, 12^\circ]$ . For the structure optimized at  $\theta_t = 0^\circ$ , the  
21 average in-pixel optical efficiency and the interpixel crosstalk are 0.076 and 0.529, respectively. When the four  
22 first neighboring pixels are also considered as part of the in-pixel region, these values are 0.276 and 0.329,  
23 respectively. In contrast, for the structure optimized at  $\theta_t = 12^\circ$ , the average in-pixel optical efficiency and the  
24 interpixel crosstalk are 0.175 and 0.552, respectively. When the first four neighboring pixels are also considered  
25 as part of the in-pixel region, these values are 0.471 and 0.256., respectively. Considering the beam waist of the  
26 quasi-Gaussian illumination, the definition of the interpixel crosstalk region can vary depending on how the  
27 effective beam footprint is characterized. In this context, the first four neighboring pixels can also be considered  
28 as part of the in-pixel optical efficiency. Furthermore, the comparison shows that the angle-robust design exhibits  
29 better confinement of optical energy and thus superior suppression of interpixel crosstalk compared to the normal-  
30 only design.

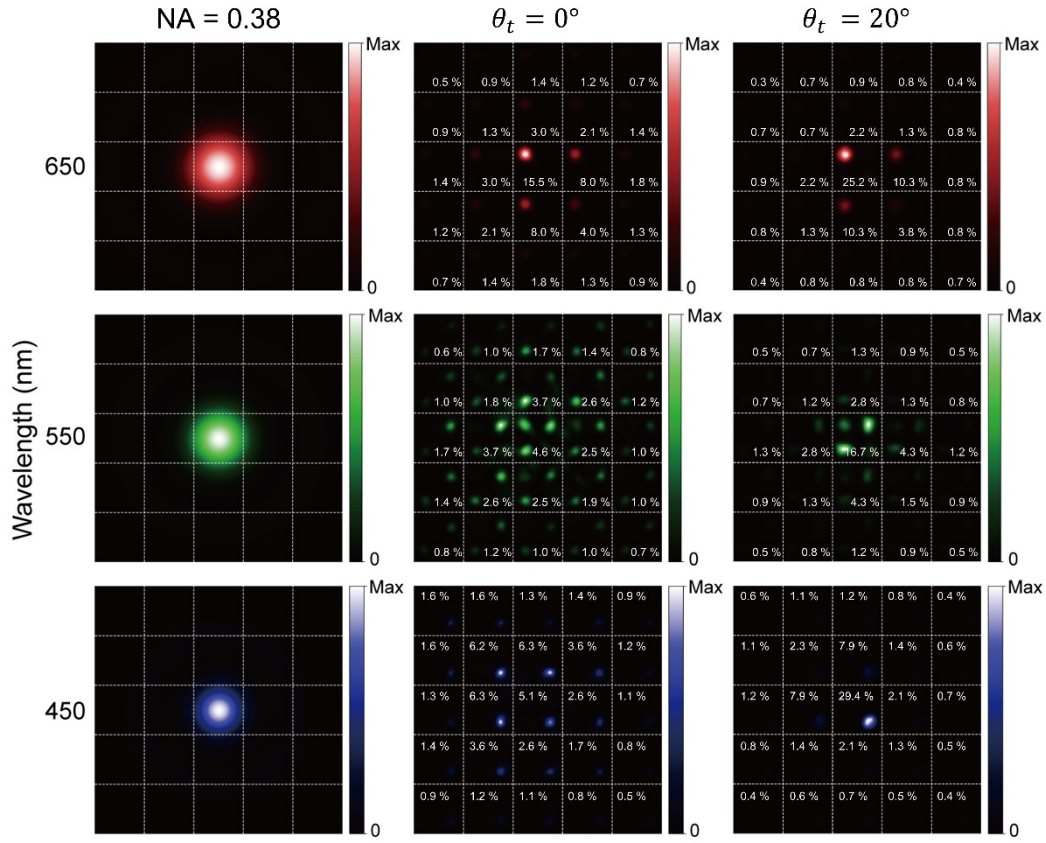
31 **Figure S19** presents the results for the beam decomposed over a wider polar range of  $\theta \in [-20^\circ, 20^\circ]$ , where the  
32 structure optimized for  $\theta_t = 0^\circ$  and  $\theta_t = 20^\circ$  are analyzed. For the structure optimized at  $\theta_t = 0^\circ$ , the average in-  
33 pixel optical efficiency and the interpixel crosstalk are 0.084 and 0.468, respectively. For the structure optimized  
34 at  $\theta_t = 20^\circ$ , the average in-pixel optical efficiency and the interpixel crosstalk are 0.238 and 0.383, respectively.  
35 These results indicate that as the design becomes more robust to high-angle illumination, the optical energy tends  
36 to be more effectively confined within the target pixel region, leading to a moderate reduction in interpixel  
37 crosstalk. This suggests that angle-robust optimization can contribute to improved spatial isolation between pixels  
38 under oblique illumination conditions.



1

2 **Figure S18.** Interpixel crosstalk for devices optimized with different target acceptance angles ( $\theta_t$ ), under varying  
3 illumination conditions. For  $\theta \in [-12^\circ, 12^\circ]$  with  $4^\circ$  increments, and azimuthal angles from  $-75^\circ$  to  $90^\circ$  in  $15^\circ$   
4 steps, a localized beam with  $\text{NA} = 0.24$  is generated through linear decomposition of plane waves. For each  
5 illumination condition, the ratio of the z-directed Poynting vector transmitted into the corresponding subpixel of  
6 each image sensor pixel was calculated relative to the input beam, and the resulting values are annotated within  
7 each pixel. The interpixel crosstalk is evaluated for the structure optimized at  $\theta_t = 0^\circ$  and  $\theta_t = 12^\circ$ .

8



1

2 **Figure S19.** Interpixel crosstalk for devices optimized with different target acceptance angles ( $\theta_t$ ), under varying  
3 illumination conditions. For  $\theta \in [-20^\circ, 20^\circ]$  with  $4^\circ$  increments, and azimuthal angles from  $-75^\circ$  to  $90^\circ$  in  $15^\circ$   
4 steps, localized beams with  $\text{NA} = 0.38$  are generated through linear decomposition of plane waves. For each  
5 illumination condition, the ratio of the z-directed Poynting vector transmitted into the corresponding subpixel of  
6 each image sensor pixel was calculated relative to the input beam, and the resulting values are annotated within  
7 each pixel. The interpixel crosstalk is evaluated for the structure optimized at  $\theta_t = 0^\circ$  and  $\theta_t = 20^\circ$ .