Supplementary Materials for

Ultracompact Wide-FOV Near-infrared Camera with Wafer-level Meta-Manufactured Aspheric Lens

Chuirong Chi, Qichao Hou, Guangyuan Zhao, Qiang Song*, Shengyuan Xu, Yanling Piao, Mengjie Qin, Yanan Hu, Chaoping Chen, Weiwei Cai, Yuan Chen, Xin Yuan**, Huigao Duan***

Corresponding authors: Qiang Song, songqiangshanghai@foxmail.com; Xin Yuan, xyuan@westlake.edu.cn; Huigao Duan, duanhg@hnu.edu.cn

This PDF file includes:

Note S1 The simulated and measurement analysis for meta-aspheric lens (MAL) with Zemax Optic Studio.

We utilize Zemax Optic Studio for the optical system design, employing the physical optics propagation method to optimize the phase distribution. Wide field-of-view (FOV) metalenses are popular in various applications including imaging, display, and sensing because they are capable of providing aberration-free, polarization-free, and high-quality imaging over large angular ranges. However, despite these advantages, designing an effective wide FOV metalens requires careful consideration of several critical factors, including image quality, form factor and process maturity. Recent research has shown that a single-layer quadratic metasurface offers more advantages in eliminating aberrations. Inspired by these findings, we aimed to design a wafer-scale wide FOV hybrid metalens. The design parameters are detailed in **Table S1**. The imaging diagram is shown in **Fig. S1.**

Tables S1 Optical design parameters

Item	Specifications
Working wavelength	940 ± 10 nm
FOV	101.5°
Working F-number	1.64
Total track length	3.39 mm
Focal length	1.4 mm
MTF	0< at 57 lp mm^{-1}
Size	2.564×2.39×3.39 mm
Distortion	<34%

The MAL comprises aspheric lens, refractive lens and metalens. We measured the actual surface profile of the aspheric lens and compared it with the theoretically fitted data. The surface fitting curve is shown in **Fig. S2a**, demonstrating that the fabricated product closely matches the theoretical values. Additionally, it is noteworthy that the MAL was fabricated using Wafer-Level Optics (WLO) technology. Subsequently, we measured the relative illumination, as shown in **Fig. S2b**. Across the entire FOV, the relative illumination of the fabricated MAL exceeds 15.86%.

We present the tangential ray trace of the MAL across 0–50.75° for 12 field angles in **Fig. S4a**. To achieve high-quality imaging, it is essential to obtain high-contrast and high-resolution images across the entire FOV. The MTF diagrams of sagittal and tangential beams are presented in **Fig. S4b** for 12 field angles. The MAL demonstrated a high MTF, exceeding 0.5 at 38 lp mm^{-1} and 0.31 at 50 lp mm^{-1}, meeting the diffraction limit for a 640×480 sensor array with 3 μm pixels. The MTF results of the metalens demonstrate its ability to provide the high resolution and contrast necessary for MAL. We measured the MTF of the MAL using the setup illustrated in **Fig. S3** and compared the experimentally measured MTF with the simulated MTF at 50 cycles mm^{-1}, and the results are shown in **Fig. S2c**. The measured data closely matches the simulation results, demonstrating the maturity of the fabrication technology. In contrast, traditional fabrication techniques often lead to a degradation in the actual MTF due to machining errors, further highlighting the advancement of the employed fabrication process.

The focusing performance of the MAL was intuitively demonstrated through its PSF intensity distribution at 12 angles within 0−50.75° FOV. As shown in **Fig. S4c**, the PSF exhibited a relatively uniform peak intensity across these angles. Meanwhile, we present the MTF values within the 30 nm design bandwidth for the 925 nm – 955 nm range, as shown in Fig S4d. **Figure**

S5a presents the spot size analysis of the entire FOV, where the geometric spot and RMS (root mean square) radius were found to be less than 15.27 μm and 5.53 μm, respectively. These results indicate excellent optical performance and alignment precision, further validating the high-quality design of the MAL. Additionally, we present the ray aberration analysis of the MAL, as shown in **Fig. S5b**. Across the entire FOV, the ray aberration is measured to be less than 20 μm. This result further demonstrates the high optical performance and precision of the MAL, highlighting its capability to maintain minimal aberrations over a wide angular range. The optical path difference was calculated and is shown in **Fig. S5c**, with values remaining within three wavelengths across the entire FOV. This indicates excellent wavefront consistency and optical performance, further validating the precision of the MAL design. Such low optical path difference values are critical for achieving high imaging quality and minimizing phase errors in optical systems.

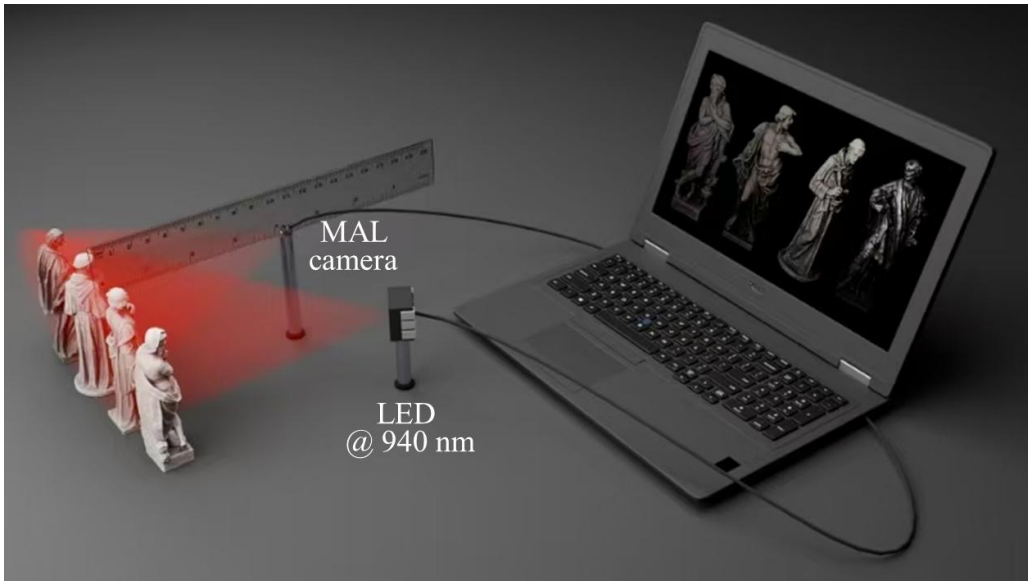


Fig. S1 Illustration of wide FOV of MAL.

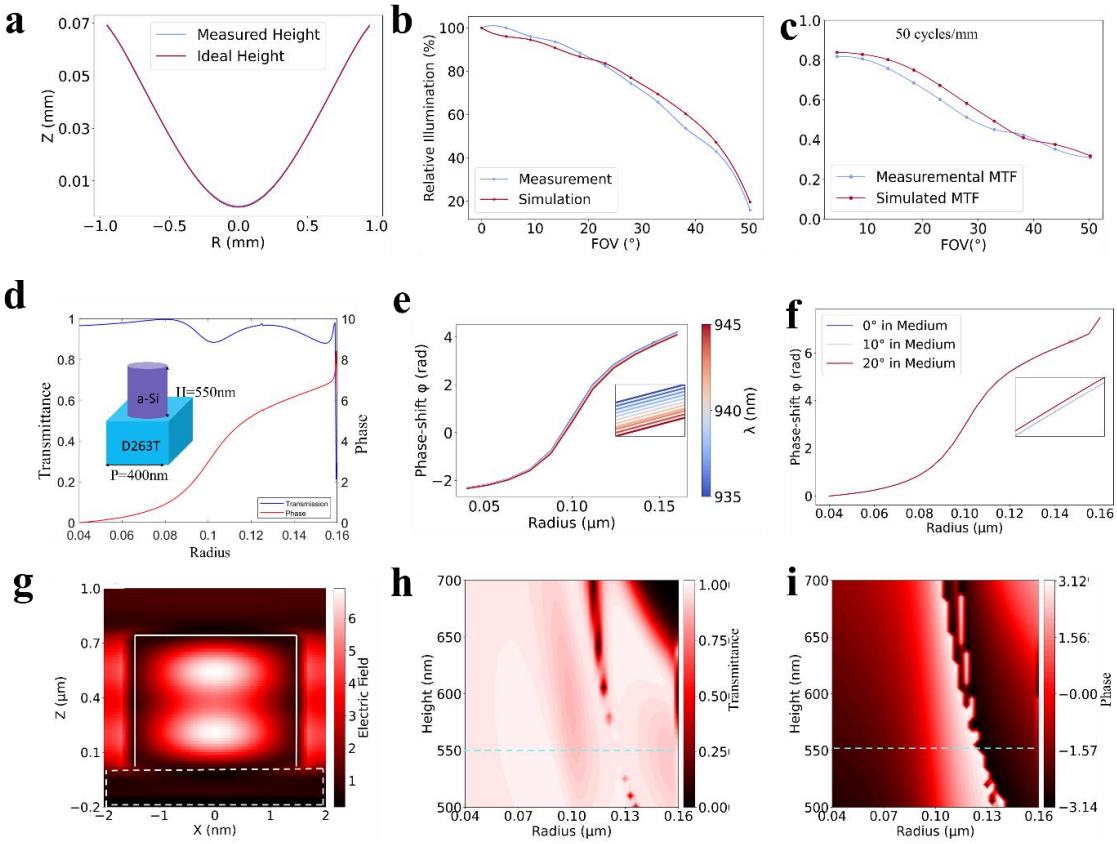


Fig. S2 Comparison of measured and simulated results for the characteristics of MAL, along with simulated results for the metalens. a Comparison of measured and simulated fitting shape of aspherical lens. **b** Comparison of measured and simulated relative illumination of MAL. **c** Comparison of measured and simulated MTF of MAL. **d** Transmission and h phase diagram versus radius at 550 nm nano-pillar height, and the unit cell structure of the metalens. **e** The relationship between the phase of a metalens and the radius of nano-posts at different wavelengths. **f** The relationship between the phase of a metalens and the radius of nano-posts at different angles. **g** Analysis of the electric field intensity distribution of individual nanorods and the coupling between adjacent nanorods. **h** Transmission and **i** phase diagrams as a function of different heights and radius at a wavelength of λ = 940 nm, with the dashed lines indicating the selected heights and radius of the nano-pillars.

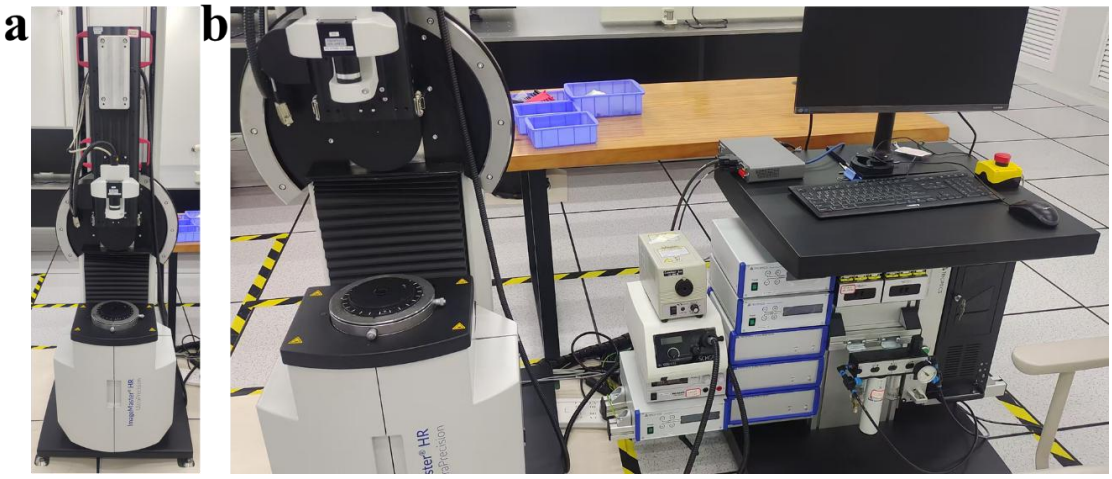


Fig. S3 Experimental setup used for measuring the MTF of MAL.

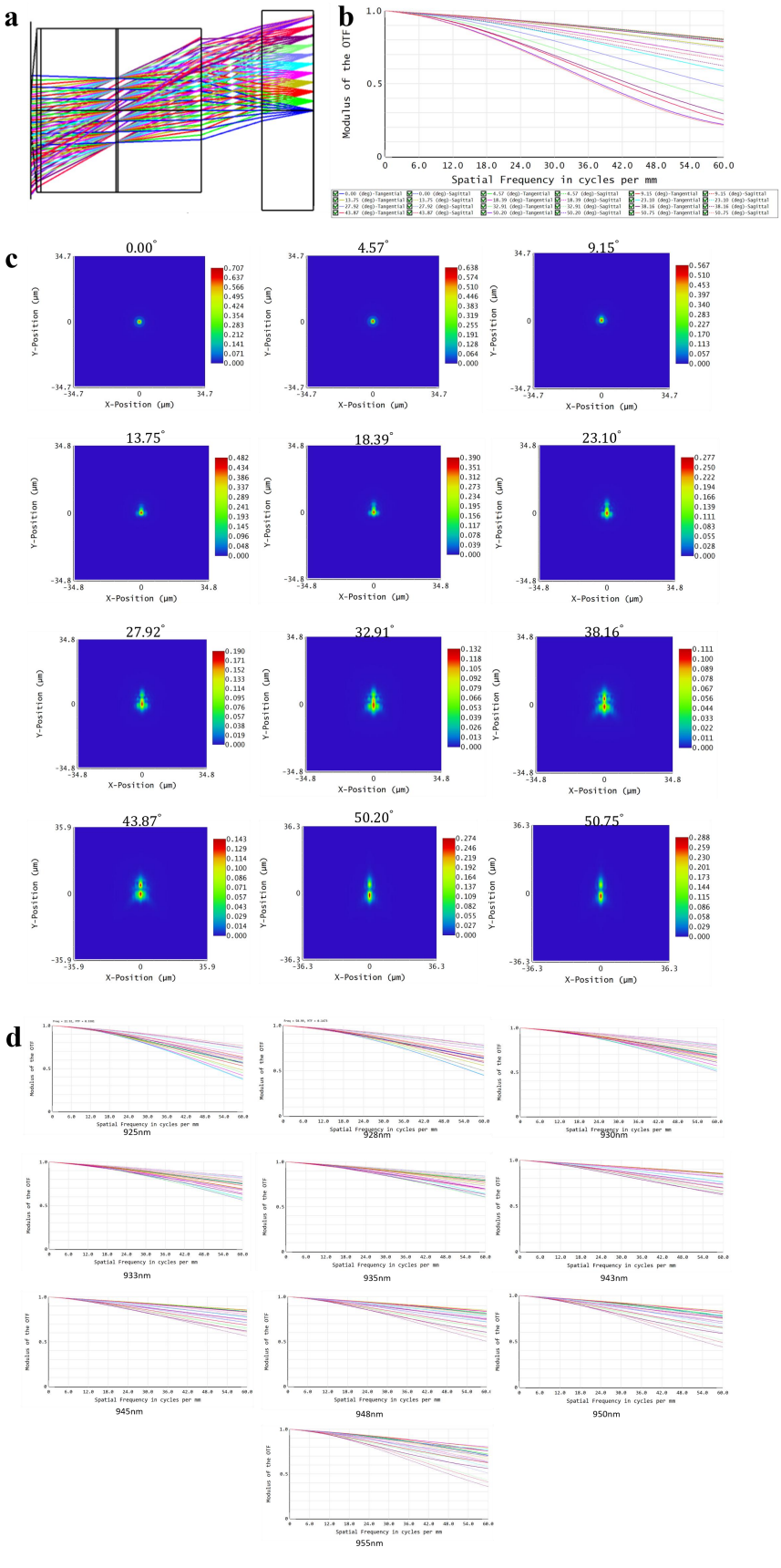

a
b
Modulus of the OTF
Spatial Frequency in cycles per mm
c
0.00°
4.57°
9.15°
13.75°
18.39°
23.10°
27.92°
32.91°
38.16°
43.87°
50.20°
50.75°
X-Position (μm)
Y-Position (μm)
d
925nm
928nm
930nm
933nm
935nm
943nm
945nm
948nm
950nm
955nm

Fig. S4 Analysis by ray tracing method of Zemax Optic Studio. a tangential ray trace of the MAL across 0–50.75° for 12 field angles. **b** MTF diagrams of sagittal and tangential beams of MAL across 0–50.75° for 12 field angles. **c** PSF characterization of MAL across 0–50.75° for 12 field angles. **d** Design MTF with a bandwidth ranging from 925 nm to 955 nm

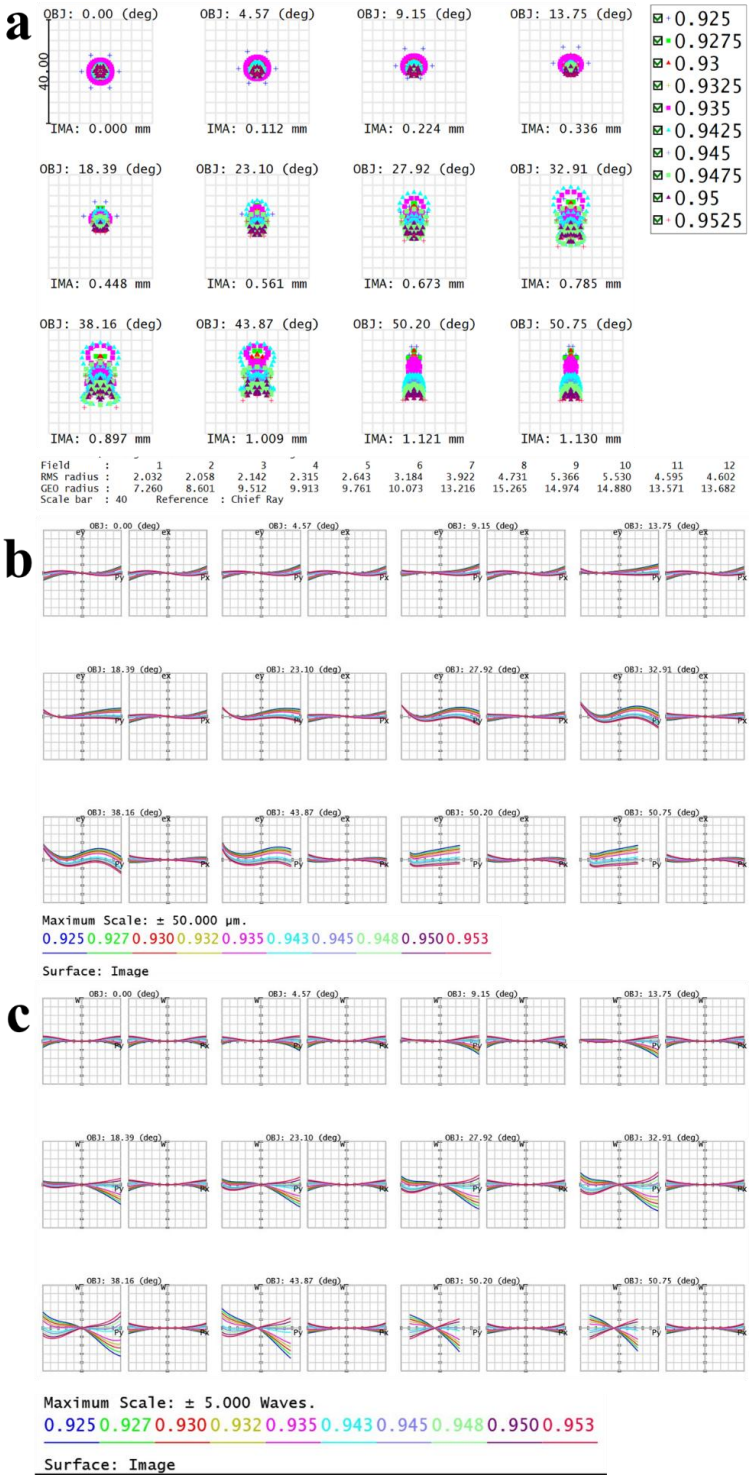


Fig. S5 Aberration analysis of the MAL. a Spot diagram analysis by ray tracing method of Zemax Optic Studio. **b** Ray aberration analysis by ray tracing method of Zemax Optic Studio. **c** Optical path difference analysis by ray tracing method of Zemax Optic Studio.

Note S2 The simulated analysis for metalens with Finite time-domain difference (FDTD).

We emplyed the Lumerical FDTD to conduct full-wave electromagnetic simulations for designing the nanostructures. To achieve a polarization-insensitive structure, while also considering the precision and error issues associated with the fabrication of sharp edges, this paper employed circular nanorods as the unit structure. The schematic diagram of the structure is shown in **Fig.S2d**. The D263T glass was selected as the substrate to balance performance and cost-effectiveness. For the nanorods, we used amorphous silicon (a-Si) for its high transmittance. To identify high-performance nano-pillars, we set the wavelength to 940 nm, the height to 550 nm, and the period to 400 nm. Metalens parameters are specified in **Table S2**. As illustrated in **Fig. S2h** and **Fig. S2i**, a fixed nano-pillar height of 550 nm and radii ranging from 40 to 160 nm can achieve a 2π phase shift while maintaining transmission above 90 %. The dashed line shows the selected height and radius of nano-pillars. The relationship between phase and radius under varying wavelengths and incident angles were analyzed in **Fig. S2e** and **Fig. S2f.** Furthermore, we analyzed the electric field intensity distribution of individual nanorods and the coupling between adjacent nanorods in **Fig. S2g.** When tracing to the metalens surface in Zemax Optic Studio, the maximum field of view of the light is approximately ±20°. We calculated the phase delays of the α-Si cylinder at different radii under incident conditions of 0°, 10° and 20°. As can be seen in the figure, the phase change is negligible within the range of 20°. Therefore, any reduction in imaging quality caused by the metalens' angular sensitivity and additional dispersion can be ignored. These analyses demonstrate the capability to achieve precise phase modulation and efficient light manipulation of metalens.

Table S2 Metalens specification

Item	Specifications
Nano-pillar height	550 nm
Nano-pillar aspect ratio	6.875
Period	400 nm
Nano-pillar radius	40-160 nm
Substrate material	D263T
Nano-pillar material	a-Silicon

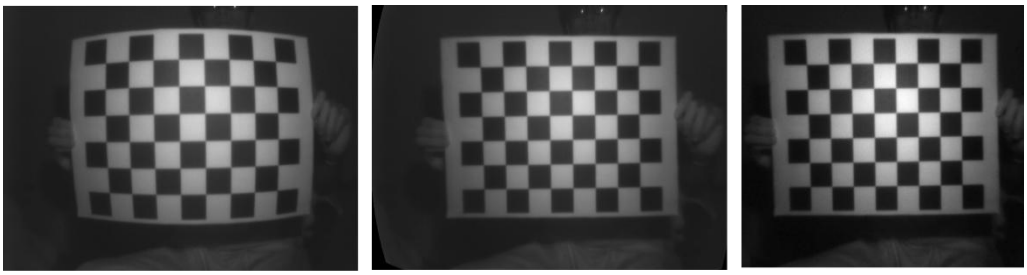
Note S3 The details of the MambaIR model for infrared image processing

Our proposed MAL provides excellent imaging quality and a sufficiently wide field of view (FOV) in the near-infrared band. However, infrared sensors typically have relatively low resolution, resulting in some details being lost in captured images. To address this issue, we introduce computational imaging to further enhance the image quality. Inspired by recent related work, we employed the MambaIR model to increase the resolution of the images captured by the MAL imaging system, thereby enabling more effective information extraction and analysis [1]. Notably, although the MambaIR model has been previously proposed, it was primarily designed for visible light images. To our knowledge, we are the first to adapt and apply it to improve the reconstruction quality of infrared images. Through this approach, we obtained high-quality images with a resolution that was four times greater. Compared with the traditional bilinear interpolation method, the MambaIR model can enhance image resolution while reducing the impact of noise, effectively improving the detailed texture and fidelity of the image. To assess the quality of images after super-resolution processing, the BRISQUE score [2], a no-reference image quality assessment method, was employed. This method quantifies image distortion by extracting natural scene statistical features, with scores ranging from 0 to 100, where lower

scores indicate better image quality. The details of the MambaIR network model and the evaluation indicators for image quality are described in detail below:

1.Image preprocessing Workflow

Before being fed into the network, images captured by metasurface cameras undergo a series of preprocessing operations to enhance their quality and provide higher-quality inputs for subsequent processing. The specific steps and corresponding image characteristics are illustrated in **Fig. S6**. **Figure S6a** shows the original, unprocessed image, which exhibits significant blurring and geometric distortion due to the optical properties of the metasurface camera. **Figure S6b** presents the undistorted image, where distortion correction has been applied to effectively rectify geometric distortions and restore the accurate geometric structure. **Figure S6c** displays the PSF deconvolution image, processed using point spread function (PSF) deconvolution to significantly reduce blurring and enhance sharpness and detail.



(a) Original (b) Undistortion (c) PSF Deconv.

Fig. S6: Preprocessing Results: (a) Original Image, (b) Undistorted Image, and (c) Image After PSF Deconvolution.

2. MambaIR model

The developed metasurface camera still faces issues with insufficient resolution and image blurriness during image capture, which may stem from limitations in optical system design, sensor performance, and external environmental interference. To address these challenges and enhance image quality, we have introduced the MambaIR1 algorithm, a highly efficient and powerful image restoration method. As far as we know, this is the first time this model has been applied to near-infrared imaging.

MambaIR leverages its core module—the Residual State Space Block (RSSB)—combined with local enhancement and channel attention mechanisms to effectively tackle the problems of image blurriness and insufficient resolution. Specifically:

a. Addressing insufficient resolution

MambaIR employs a global modeling capability to capture long-range dependencies within the image while extracting high-frequency detail information through its local enhancement module. This combination enables MambaIR to reconstruct finer details in low-resolution images, thereby significantly improving image clarity and resolution.

b. Addressing image blurriness

Image blurriness may result from focal length errors, motion blur, or optical system design flaws. MambaIR's channel attention mechanism dynamically adjusts the weights of different channels to highlight important features, while its residual structure minimizes information loss.

Through this approach, details in blurry areas are restored, and the overall sharpness of the image is enhanced.

c. Adapting to the unique characteristics of metasurface cameras

The imaging process of metasurface cameras has unique physical properties. MambaIR's efficient design allows it to flexibly adapt to these characteristics. By incorporating prior knowledge from image restoration tasks, MambaIR can maintain computational efficiency while activating a larger effective receptive field to capture more global and local information.

d. Efficiency and real-time processing

Compared to traditional deep learning methods, MambaIR has significant advantages in computational complexity and memory usage. This makes it suitable not only for offline image processing but also for real-time image restoration scenarios, providing technical support for the practical deployment of metasurface cameras.

By introducing MambaIR, we can effectively mitigate the limitations of metasurface cameras in terms of resolution and image blurriness, significantly improving imaging quality. Moreover, the application of this model to near-infrared imaging represents a pioneering step, further demonstrating its versatility and potential. This advancement lays a solid foundation for metasurface cameras to be utilized in scientific research, industrial inspection, and real-world applications.

As shown in **Fig. S7**, the network begins with an Extraction module, which processes the input image to extract initial features. These features are then passed through a series of Residual State Space Groups (RSSG), each containing multiple Residual State Space Blocks (RSSB). The processed features are further refined by additional RSSG layers, followed by a Convolutional (Conv) layer. The final output is obtained by adding the input and the processed features, ensuring residual learning.

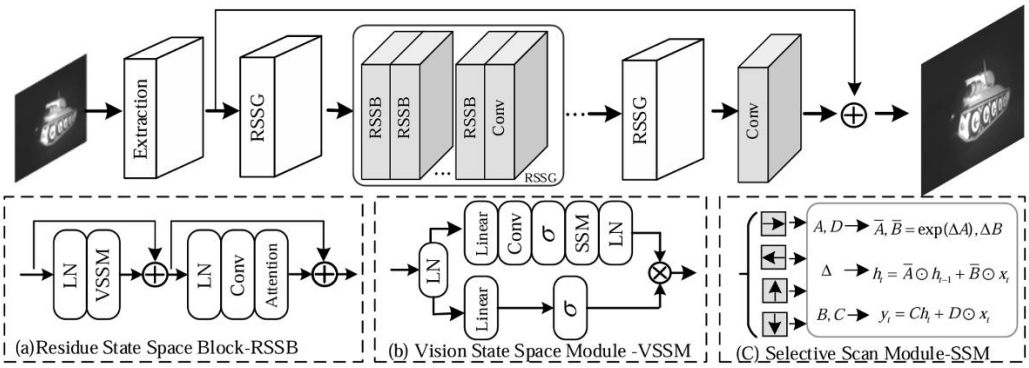


Fig. S7: The overall architecture of the MambaIR network.

Key Components:

(a) Residual State Space Block (RSSB):

Each RSSB includes the following sub-modules:

Layer Normalization (LN): Normalizes the features to stabilize training.

Vision State Space Module (VSSM): A core module designed to capture long-range dependencies and spatial relationships.

Convolutional (Conv) Layers: Extract local features to complement the global modeling of VSSM.

Attention Mechanism: Dynamically adjusts feature weights to enhance relevant information.

(b) Vision State Space Module (VSSM):

This module employs **Linear** and **Convolutional** operations to model both global and local dependencies. It incorporates a **Selective Scan Module (SSM)** to adaptively process spatial information while maintaining computational efficiency.

(c) Selective Scan Module (SSM):

The SSM dynamically updates features using the following equations:

$$\boldsymbol{A, D \rightarrow \overline{A}, \overline{B} = exp\,(\Delta A), \Delta B}$$

$$\boldsymbol{A \rightarrow h_t = A_0 h_{t-1} + B_0 x_t}$$

$$\boldsymbol{B, C \rightarrow y_t = C h_t + D_0 x_t}$$

These operations allow the module to selectively scan and update spatial features, ensuring efficient and effective feature extraction.

Purpose of Each Module:

The RSSB and VSSM work together to balance global and local feature extraction, addressing issues like image blurriness and low resolution. The SSM enhances the network's ability to capture spatial details adaptively. The Residual Learning mechanism ensures effective gradient flow and prevents information loss during feature processing.

3.Image Quality Evaluation by BRISQUE score

This paper conducts a comparative analysis of infrared (NIR) images processed using different methods and explores the relationship between image clarity and processing effectiveness by employing the no-reference image quality evaluation metric BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator).

BRISQUE is a no-reference image quality evaluation metric that predicts perceptual quality by analyzing an image's statistical properties without needing a reference image. Its key principles include:

- **Natural Scene Statistics (NSS):** High-quality images follow normal statistical distributions (e.g., Gaussian), while distortions cause deviations.
- **Feature Extraction:** BRISQUE uses Mean Subtracted Contrast Normalized (MSCN) coefficients to assess contrast, texture, and edges.
- **Regression Model:** A Support Vector Regression (SVR) model maps these features to a quality score—lower scores mean better quality, higher scores indicate more distortion.

BRISQUE's no-reference nature makes it ideal for evaluating infrared images, where reference images are often unavailable, ensuring fast and objective quality assessment. This method quantifies image distortion by extracting natural scene statistical features, with scores ranging from 0 to 100, where lower scores indicate better image quality.

As illustrated in **Fig. S8**, the comparison of image quality among the original images, Bicubic 4× interpolation, and MambaIR 4× super-resolution methods is evaluated using the BRISQUE score. To validate the reliability of the BRISQUE score for near-infrared (NIR) images, we also included a high-quality reference image captured by a commercial NIR camera for comparison.

The BRISQUE scores highlight the advantages of the MambaIR 4× method over traditional Bicubic interpolation. MambaIR achieves a significantly lower score, reflecting its superior ability to preserve image details, enhance clarity, and reduce artifacts. Moreover, the close proximity of the MambaIR score to the reference NIR camera score further validates its effectiveness in reconstructing high-quality images.

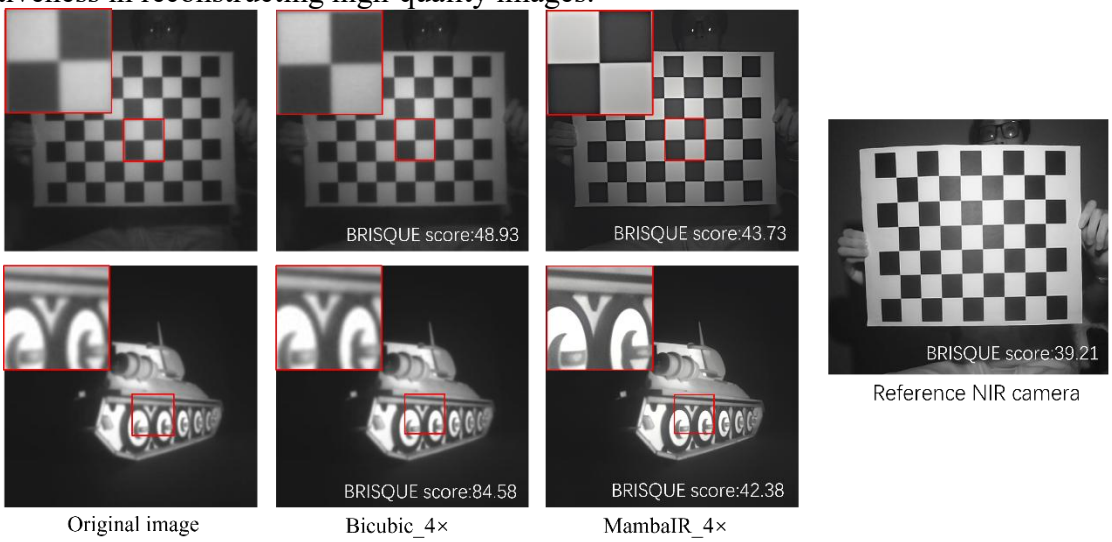


Fig. S8: BRISQUE-Based Image Quality Comparison for NIR Super-Resolution

Note S4 Applications of MAL Technology.

MAL technology opens transformative possibilities across industries by addressing the critical need for compact high-resolution imaging systems, Wide FOV (101.5°). In smartphone cameras, MAL enables ultra-slim modules with wide-field NIR imaging capabilities, enhancing low-light photography, depth sensing, and biometric authentication (e.g., iris scanning) without compromising device portability. For wearable AR glasses, the technology supports lightweight, wide FOV optics critical for immersive augmented reality experiences. The designed MAL can be deployed on AR glasses to enable highly precise eye movement tracking. The perception of large field of view (FOV) and NIR enables clear images of eye movement changes to be captured at close range and in low light conditions. Meanwhile, the compact structure of the MAL can be easily adapted and integrated into AR glasses without compromising their portability, appearance and user experience the device. In autonomous vehicles, MAL enhances LiDAR and environmental perception systems with miniaturized, high-resolution NIR sensors, enabling robust obstacle detection and navigation in challenging weather or low-visibility scenarios. Medical applications, such as endoscopic diagnostics, benefit from the lens's compact design and superior resolution, allowing deeper tissue imaging with smaller, less invasive probes. Furthermore, in smart home systems, MA-powered devices achieve secure, low-profile biometric authentication (e.g., facial recognition) with improved anti-spoofing capabilities through multispectral NIR imaging. By integrating computational imaging techniques like MambaIR -based super-resolution, MA systems further overcome physical resolution limits, enhancing image clarity for applications ranging from industrial inspection to scientific imaging. This synergy of wafer-scale manufacturability, advanced optics, and AI-driven processing positions MAL as a cornerstone for next-generation imaging systems, redefining performance benchmarks

in consumer electronics, healthcare, robotics, and beyond. **Figure S9** shows a conceptual demonstration of MAL deployed on AR glasses.

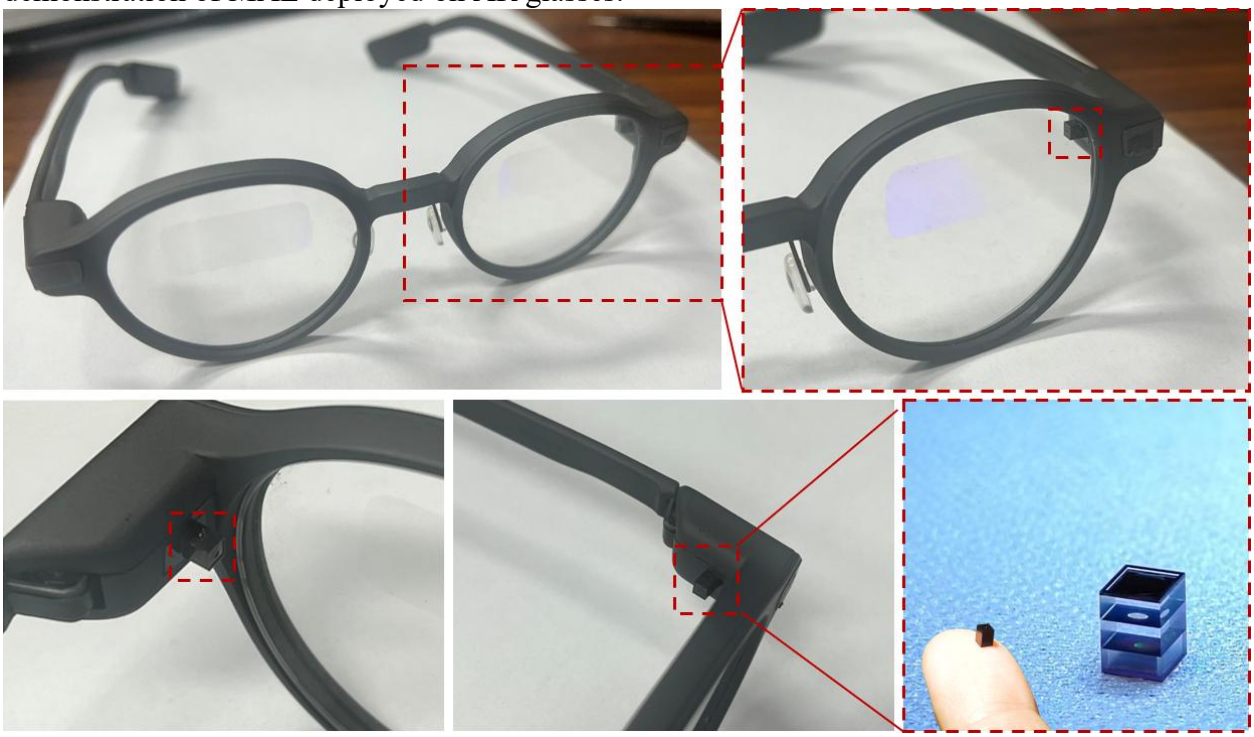


Fig. S9 Applications of MAL Technology. The MAL integrated into the AR glasses and its corresponding local magnified view.

Note S5 Details of wafer-level manufacturing

In the wafer-level fabrication of the metamaterial lenses, electron-beam lithography (EBL) was employed for wafer-scale patterning. Although this method is less efficient, it ensures reliable accuracy. Scanning electron microscopy (SEM) images of the top and side views of the fabricated metalens are shown in **Fig. S10**. The fabricated structures exhibit dimensions closely matching the designed values (design radius: 80–290 nm; height: 550 nm).

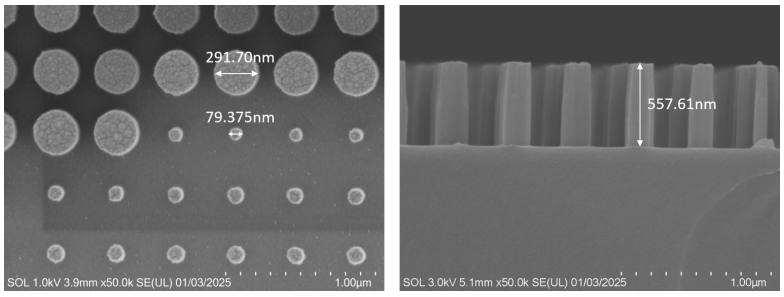


Fig S10 SEM images of the top and side views of the metalens

Side-view SEM images from different wafer regions are also provided in **Fig. S10**. The sidewall angles of the meta-atoms at all positions are <15°, and the edge roughness remains within an acceptable range, consistent with the current processing capability for metalens. During wafer-

level, large-area etching, however, depth non-uniformity arises; conservatively, the fabricated meta-atom heights vary by ±20 nm (530–570 nm).

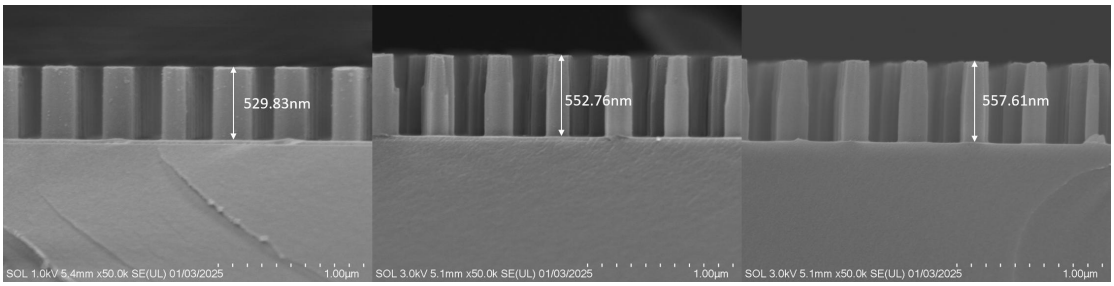


Fig S11 Side SEM images of metalens in different areas

To evaluate the influence of meta-atom height variation, we conducted simulation calculations for the height of the meta-atom ranging from 530 nm to 570 nm using FDTD. The results, given in **Fig. S13**, show that within this range the phase gradient is preserved and the transmittance remains high; consequently, the overall device performance is not appreciably affected.

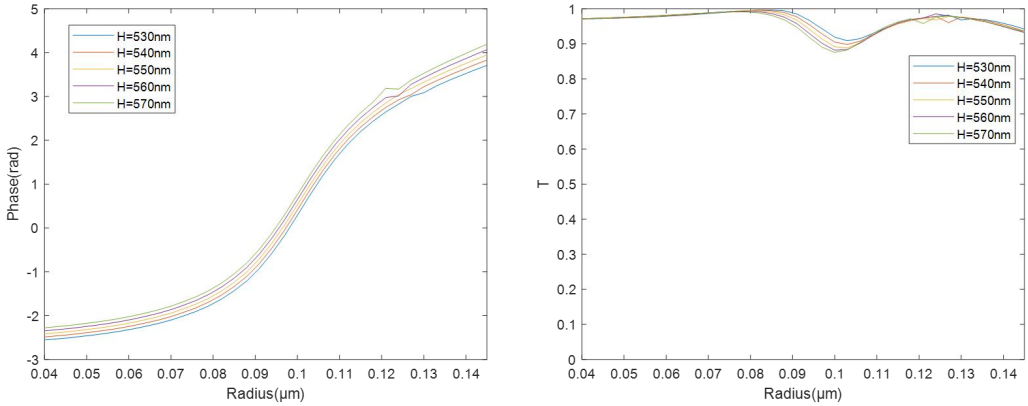


Fig S13 Distribution map of phase and transmittance of the nanocylinder with a height variation of ±20nm as the radius changes

To further illustrate the yield of our wafer-level manufacturing of MAL, we quantitatively evaluated the central diffraction efficiency of the additional 6*5 metalens and the aspheric lens profile errors in different areas during the wafer manufacturing process. The experimental setup for measuring the diffraction efficiency of the metalens array and the distribution map of the measured diffraction efficiency are shown in **Fig. S13** and **Fig. S14**. The average diffraction efficiency of the 6*5 array metalens central field measured is 84.96%, indicating that the differences between these metalens array devices fabricated at the wafer level are within an acceptable range.

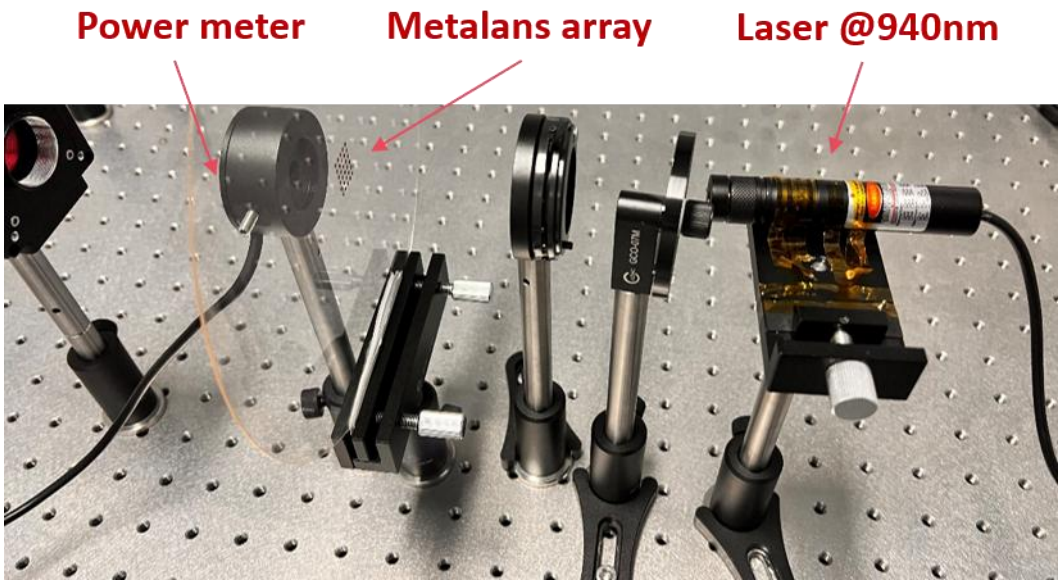


Fig S13 Diagram of the experimental setup for measuring the diffraction efficiency of a metalens

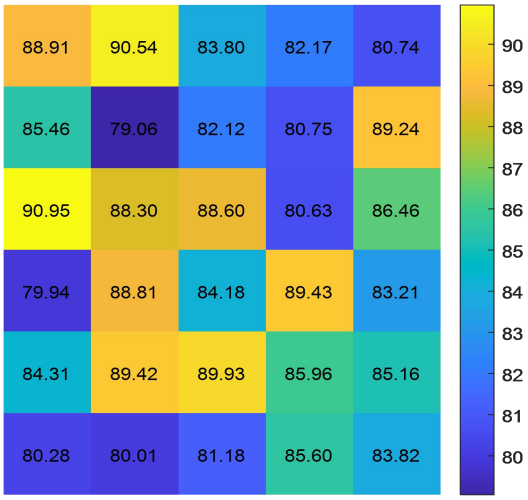


Fig S14. Efficiency distribution map of 6*5 array metalens

For the consistency of the aspheric lens array processing, **Figure S15** shows the comparison of the height data of aspheric lenses in four different areas on the wafer. For the concentrated areas of replication, we also measured the height error data of 2*2 array and 10*10 array of the aspheric lens. The 2 × 2 array exhibited a height-error PV (Peak-to-Valley) of 1.098 μm and RMS (Root Mean Square) of 0.169 μm, whereas the 10 × 10 array yielded a PV of 1.061 μm and RMS of 0.270 μm. These negligible deviations confirm the reliability of the employed NIL process for aspheric-lens-array fabrication.

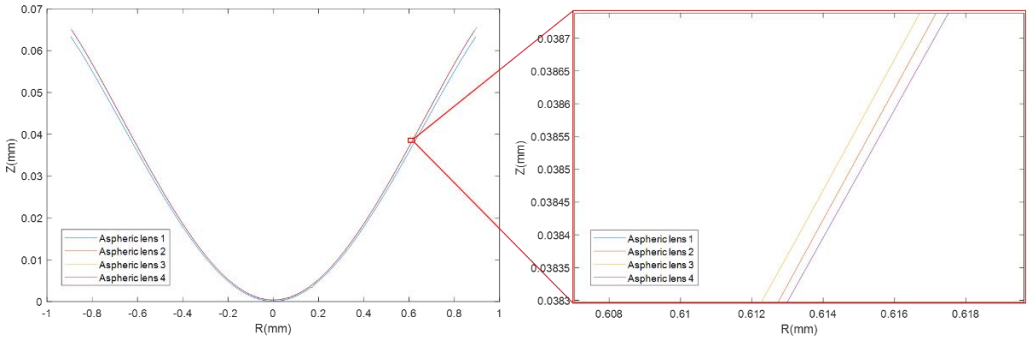


Fig S15 Height comparison of aspheric lenses in four different regions

Regarding the performance tests of MAL at different positions on the wafer, in addition to the MTF and relative illuminance of four randomly selected MALs that we presented in the manuscript, we also measured the effective focal lengths (EFL) of these four MALs. The distribution of the effective focal length data is as follows:

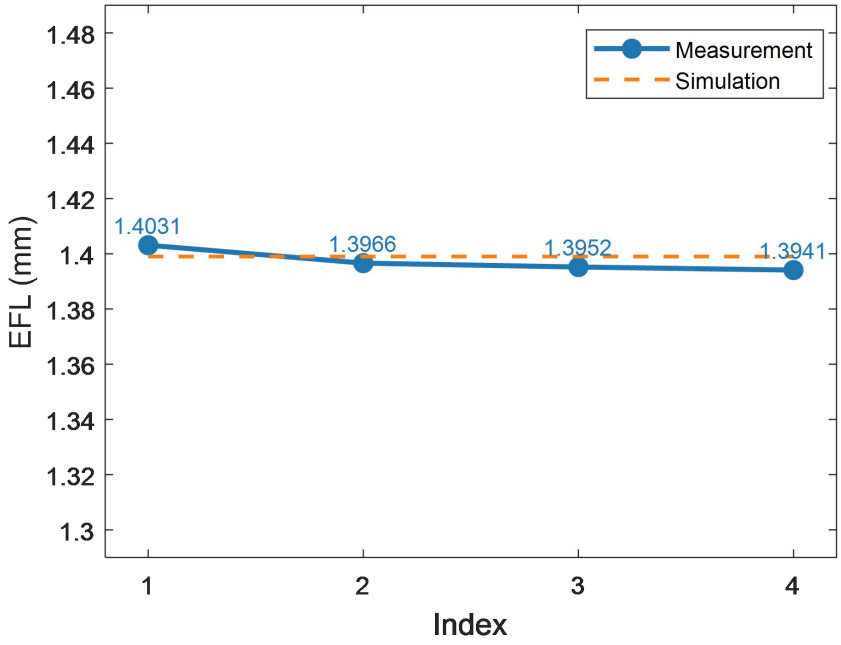


Fig S16 Effective focal length measurement values of MAL at four random positions on the wafer

References

[1] Hang G. et al. MambaIR: A Simple Baseline for Image Restoration with State-Space Model. *In Computer Vision – ECCV2024, Milan, Italy, 2024,* pp. 222–241.

[2] Anish. M. et al. No-Reference Image Quality Assessment in the Spatial Domain. *In IEEE Transactions on Image Processing,* **21**, 12, 4695-4708 (2012)