Supplementary information to
High-resolution and wide-field microscopic imaging with a monolithic meta-doublet under annular illumination

Jiacheng Sun[a,b †], Wenjin Shen[a †], Junyi Wang[a,b †], Rongtao Yu[a,b] , Jian Li[a], Chunyu Huang[a], Xin Ye[a], Zhaoyu Cheng[b], Jiefu Yu[b], Peng Wang[b], Chen Chen[a]*, Shining Zhu[a], and Tao Li[a]*

a National Laboratory of Solid State Microstructures, Key Laboratory of Intelligent Optical Sensing and Manipulation, Jiangsu Key Laboratory of Artificial Functional Materials, College of Engineering and Applied Sciences, Nanjing University, Nanjing, China, 210093
b MetaCV Technology Co., Ltd, Nanjing, China, 210012

†*These authors contributed equally.*
**Corresponding author: chenchen2021@nju.edu.cn, taoli@nju.edu.cn*

Supplementary note 1: Constraints of NA and FOV

It is well established that the numerical aperture (NA) of an imaging system is inversely related to its field of view (FOV): increasing the NA inevitably reduces the effective FOV. Additionally, the lens size significantly impacts the ratio of FOV to lens size, commonly referred to as the relative FOV. Larger metalenses generally exhibit a smaller relative FOV because, as the system scales up, the optical path difference between corresponding light rays grows proportionally. This leads to more pronounced wave aberrations, ultimately reducing the system's relative FOV.

To illustrate this relationship, we analyzed the dependence of relative FOV on lens diameter and object NA using a single metalens as a reference. The effective FOV was evaluated using the Rayleigh criterion, which defines satisfactory imaging performance as a Strehl ratio above 0.8 or an RMS wave aberration below one-quarter of the wavelength. The calculations were performed for a single metalens with a magnification factor of 3, mirroring the configuration of the metalens doublet discussed in the main text. The two-dimensional distribution of relative FOV as a function of NA and lens diameter is shown in Fig. S1a. The results indicate that the relative FOV is significantly constrained unless both the NA and lens size are kept small. This highlights the inherent challenge of achieving a large effective FOV with a single metalens.

To further clarify this limitation, Fig. S1b compares the relationship between relative FOV and NA for both a single metalens and an optimized metalens doublet, with lens diameters of 0.45 mm and 1.7 mm, respectively. Under identical conditions, the optimized doublet exhibits a marked improvement in effective relative FOV. However, even with this optimization, the effective FOV decreases rapidly as the NA increases. This trend underscores the difficulty in achieving both high NA and a large FOV, ultimately imposing a limit on the resolution enhancements achievable with the optimized meta-microscope.

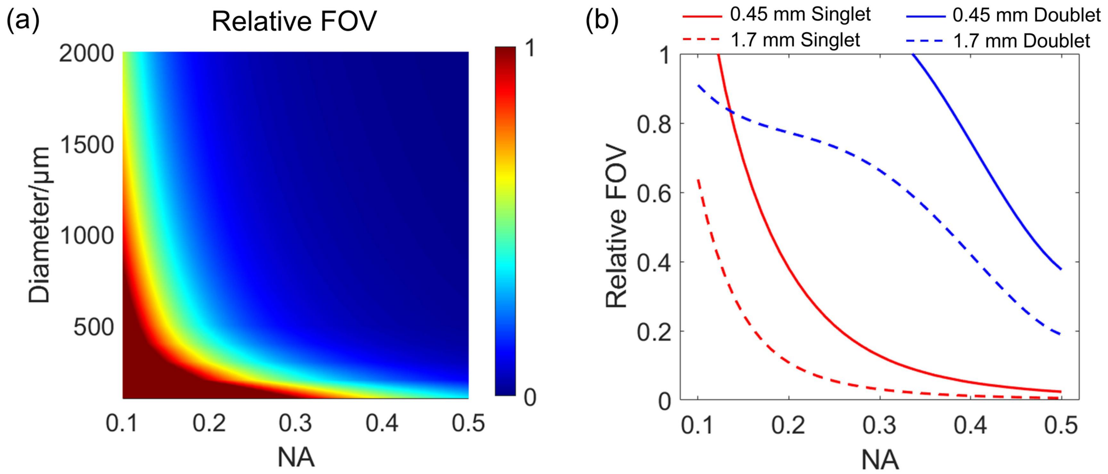

Fig. S1 (a) Distribution of relative FOV with respect to NA and size. The relative FOV is less than 0.1 in most cases. (b) Relationship between relative FOV and NA of both single metalens (red curves) and optimized metalens doublet (blue curves). The metalens doublet has impressive edge to the single lens, while the relative FOV decreases rapidly as NA increasing.

Additionally, there is a huge gap between the designed NA of metalens and the actual achieved resolution in a lot of works reported, especially when the NA is large. The corresponding parameters are shown as follows:

Table S1 NA and resolution in some meta-microscopy works reported

	Wavelength/nm	NA	Ideal Resolution/μm	Actual Resolution/μm
Li et al.[24]	800	0.37	1.32	2.19
Wang et al.[25]	620	0.7	0.54	0.78
		0.1	3.78	3.92
Kwon et al.[27]	850	0.4	1.30	1.74
Long et al.[28]	532	0.4	0.81	0.87
Kim et al.[30]	480-680	0.8	0.37	0.78
Arbabi et al.[33]	605	0.5	0.74	0.81

A substantial share of metalens-based microscopy research did not present a direct imaging result of a resolution chart to calibrate the system resolution. Simply measuring a point spread function can not characterize the actual imaging resolution of a metalens, because the imaging resolution is a combined effect of contrast and aberration. The imaging contrast is directly related to focusing efficiency, which is not a neglectable issue for metalens. For better clarification, the imaging resolution of metalens is characterized directly by a resolution chart in the main text, and the experimental setup is carefully adjusted to ensure the result is as good as possible. Several works using similar methods to calibrate imaging resolution are presented in Table S1. The discrepancy between the actual and ideal resolution presented above is obvious, and high NA metalens struggles more to reach the ideal resolution compared to low NA ones. This may result from factors such as a more severe drop in focusing efficiency, smaller FOV, etc. As a result, it is not practical to improve the imaging resolution by simply increasing the metalens NA while maintaining a suitable FOV.

Obviously, the actual resolution struggles to reach the ideal resolution in most works, especially when the NA is high. As the NA increases, the phase gradient at the edge of the metalens rises sharply, resulting in an insufficient sampling frequency at the edge. Moreover, the angular dispersion of sub-wavelength structures at the metalens edge becomes more pronounced with higher NA.

As a result, it is challenging for the metalens itself to achieve a large NA while maintaining excellent imaging quality, and additional methods are essential to achieve large FOV and high-resolution imaging.

Supplementary note 2: Methods

Zemax simulation and optimize: Optimization coefficients included polynomial coefficients, object distance, and image distance. Before optimization, all polynomial coefficients were set as variables with the initial value of 0, and the object and image distance were set 350 nm and 700 nm, respectively. At the same time, it was necessary to define default merit function surface (e.g., root mean square spot radius centroid Gaussian quadrature) in the figure of merit listing. The algorithms of orthogonal descent (OD) algorithm and damped least square (DLS) were used for optimization. The coefficients of the meta-microscopes are listed in Tables S2-S5.

Table S2 Optimal values of coefficients for the metalens-I of the doublet with NA 0.5

R_1/μm	a_1	a_2	a_3	a_4	a_5
225	-1306.4767	-1.6795	262.5454	-811.9746	-811.9746
	a_6	a_7	a_8	a_9	a_{10}
	-1667.7216	1204.9181	-537.1671	134.7347	-14.5440

Table S3 Optimal values of coefficients for the metalens-II of the doublet with NA 0.5

R_2/μm	a_1	a_2	a_3	a_4	a_5
120	-25.0646	23.4109	2.4607	0.1329	36.7115
	a_6	a_7	a_8	a_9	a_{10}
	-184.5449	385.1874	-411.1134	222.3454	-48.4965

Table S4 Optimal values of coefficients for the metalens-I of the doublet with NA 0.3

R_1/μm	a_1	a_2	a_3	a_4	a_5
850	-5395.4365	295.0900	149.6008	-1672.1012	4614.3747
	a_6	a_7	a_8	a_9	a_{10}
	-4774.7704	-1231.9241	7125.1037	-5866.5820	1613.2243

Table S5 Optimal values of coefficients for the metalens-II of the doublet with NA 0.3

R_2/μm	a_1	a_2	a_3	a_4	a_5
260	-124.5606	-47.3238	446.1214	-2080.1313	6002.3759
	a_6	a_7	a_8	a_9	a_{10}
	-11084.0034	13125.2512	-9636.2831	3989.8415	-711.2961

Köhler illuminating path: The experimental Köhler lighting consists of an LED light source (Jcoptix, LEM-W3), a filter ($\lambda_{\mathrm{center}}$ = 470 nm, bandwidth = 10 nm), three focusing lenses (Jcoptix, OLD2434-T2M, f = 4 cm; OLD3266-T2M, f = 20cm; OLD2438-T2M, f = 6cm) and two stops. Focusing and adjusting the filaments at the rear focal plane of the objective lens and the front focal plane of the condenser lens produce a maximum defocusing filament image on the specimen image plane, uniformly illuminating the entire field of view.

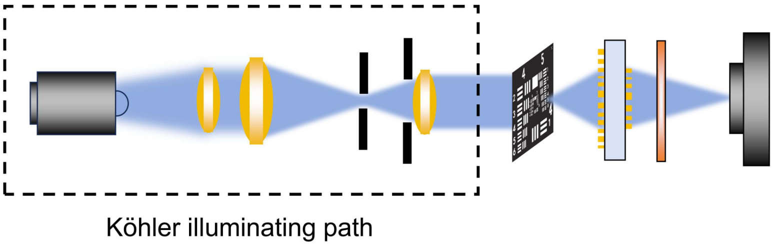

Fig. S2 Schematic drawing of the Köhler illuminating path.

Device fabrication: A pair of alignment marks was patterned on the two sides of the substrates and aligned to each other using optical lithography (SUSS MA6BA6). Then the SiN$_x$ layer was deposited on the two sides of the fused silica substrate respectively using the plasma-enhanced chemical vapor deposition (PECVD) to a final thickness of 1000 nm. The aperture stop, which was composed of 100nm Cr and Au, was patterned using optical lithography. For the metalens-II, 220 nm PMMA A4 photoresist film is spin-coated onto the substrate and baked at 170°C for 5 min. A layer of water-soluble conductive polymer (AR-PC 5092) was spin coated on the resist for the dissipation of E-beam charges. The device pattern was written on an electron beam resist using an E-beam writer (Elionix, ELS-F125). An electron beam evaporated chromium layer was used to reverse the generated pattern with a lift-off process and was then used as a hard mask for dry etching the SiN$_x$ layer. Before etching, the Cr mask of the metalens-I was patterned on the backside of the substrate using a procedure similar to that described earlier. After that, the devices transferred into OXFORD PP100 Plasma Etcher and etched with SF$_6$ and CHF$_3$ (the flow rates of these two types of gases are 6sccm:60sccm). The chromium layer was removed by a stripping solution (ceric ammonium nitrate) later.

Supplementary note 3: Characterization of the Metalens Doublet

We characterized the imaging performance of the metalens doublet by utilizing a 1951 United States Air Force (USAF) resolution test chart as the object. The test chart was positioned in front of the metalens doublet and illuminated using a Köhler illumination setup, where the NA_{illum} was set 0, as shown in Fig. S3a. The image produced by the doublet was captured by a CMOS image sensor situated at the image plane behind the metalens doublet. Figure S3b presents the image captured by the CMOS sensor. Notably, the metalens doublet exhibits high clarity throughout all FOV. Figure S3c shows the normalized intensity profile of the vertical line pairs of group 9, where element 5, group 9 is distinctly identifiable, indicating a resolution exceeding 620 nm. Furthermore, our metalens doublet consistently achieves high resolution across the full FOV. We experimentally measured the focal spots in different FOVs. In the experiment, we used the objective lens to focus the normally incident collimated beam from a tunable continuous wave laser and employed the focal spot as the incidence point source for tests. Figure S3d shows the measured focal spots intensity profiles at different object heights, and the related x-direction cross sections are illustrated (bottom). The measurement results are close to the simulation results.

To quantitatively assess the imaging performance, the MTFs and Strehl ratios of the metalens doublet were calculated from the experimentally obtained PSFs. The MTFs were calculated through performing Fourier transform on the measured PSFs, as illustrated in Fig. S3e, which show diffraction limited performance across a wide FOV. Figure S3f presents both experimental and simulated Strehl ratios various object heights. These ratios are determined by comparing the average area under the MTF curves in the sagittal and meridional directions to that of the diffraction-limited MTF. The Strehl ratios are slightly lower than the theoretical predictions at all object heights, likely due to background noise and the angular dispersion caused by sub-wavelength structures. Nonetheless, the metalens doublet achieves a Strehl ratio above 0.8 for all object heights, indicating its strong imaging potential.

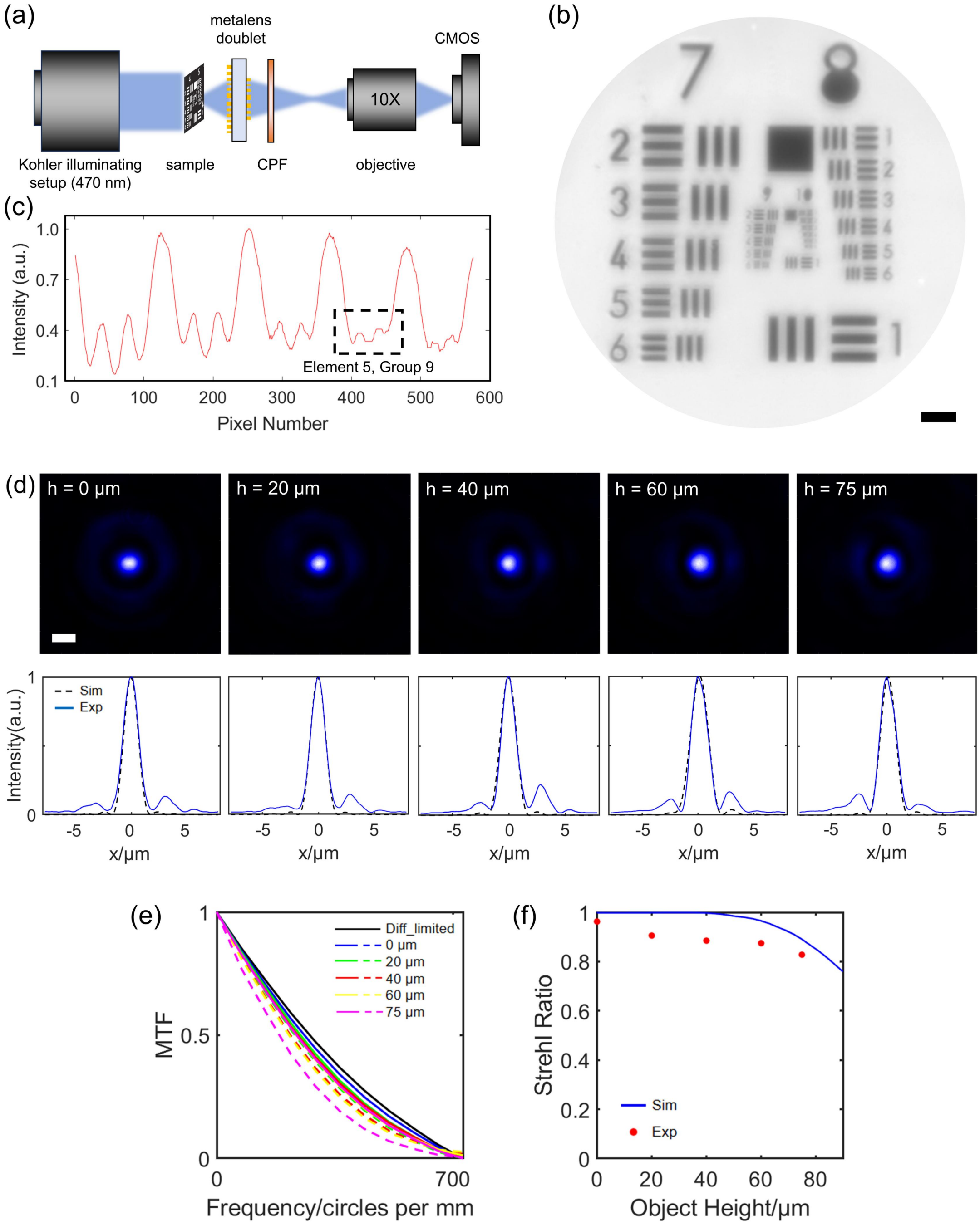

Fig. S3 (a) Schematic drawing of the measurement setup. (b) The image of the designed FOV captured by the CMOS. Scale bar = 15 μm. (c) The line cut of the vertical line pairs of group 9, where element 5 can be recognized. (d) Measured focal spot diagrams and focal spot intensity profiles at object heights of 0, 20, 40, 60, 75 μm. Scale bar = 2 μm. (e) MTF calculated from the PSFs, demonstrating diffraction-limited imaging performance. (f) Strehl ratios of different object heights.

Supplementary note 4: Frequency shift in optics

Any optical imaging system, including microscopy imaging systems, has a PSF with linear spatial translation invariance, and the image of an object $D(\boldsymbol{r})$ passing through a microscope can be described using a convolution process:

$$D(\boldsymbol{r}) = S(\boldsymbol{r}) * H(\boldsymbol{r}) \tag{S1}$$

where $S(\boldsymbol{r})$ is the function of the object to be imaged, and $H(\boldsymbol{r})$ is the PSF of the microscopic imaging system, which is the image of a point object formed by the system. For the convenience of mathematics, Fourier transform is used to transform the spatial domain into the Fourier domain spectrum. The formula can be transformed to

$$\widetilde{D}(\boldsymbol{k}) = \widetilde{S}(\boldsymbol{k}) \cdot \widetilde{H}(\boldsymbol{k}) \tag{S2}$$

where $\widetilde{D}(\boldsymbol{k})$ is the spectrum of the corresponding image, $\widetilde{S}(\boldsymbol{k})$ is the spectrum of the object, and $\widetilde{H}(\boldsymbol{k})$ is the transfer function of the system, which is the Fourier transform of PSF of the system and describes the ability of transferring different frequencies of the object. The spectrum of the object includes various spatial frequency components: low-frequency components usually correspond to the rough contour of the object and the high-frequency components correspond to the fine structure of the object.

As can be seen from formulas (S1) and (S2), the frequency spectrum of the obtained object image is the result of filtering the object spectrum through the optical system transfer function. Therefore, in theory, we can equate the microscopy imaging system as a circular low-pass filter, where only low-frequency information of objects within the filter's cut-off frequency can be transmitted to the image plane, while high-frequency information outside is cut off and cannot be transmitted. This leads to the loss of high-frequency information in the image plane, and the cutoff frequency represents the maximum resolution capability of the imaging system. In order to obtain more spectral information of objects in the image plane and improve system resolution, the most fundamental work is to expand the actual detectable spectral range of the microscopic imaging system, in order to obtain a larger range of high-frequency spectral components of objects as much as possible.

If the sample is illuminated by plane waves normally, as shown in Fig. S4a, the zero-order diffracted light scattered by the sample is in the center of the objective lens aperture, so only the diffracted light of low diffraction order within the scope of the objective lens aperture ($|\boldsymbol{k}|<k_c$) can pass through, limiting the spectrum range of the received object. If the illumination angle of the incident plane wave is changed, it becomes oblique illumination, as shown in Fig. S4b. In this case, the diffraction angle of the zero-order diffracted light corresponding to the scattering of the sample changes with the angle of the illumination light, in other words, there will be higher-order

diffracted light through the aperture of the objective lens, and the higher-order diffracted light often carries high-frequency information of the fine structures of the object. Therefore, the high-frequency spectrum of the object can be moved to the pass-band range of the objective lens with the help of oblique lighting, improving the resolution. Assume that the plane wave incident on the sample at angle θ, and its decomposed transverse wave vector is $\boldsymbol{p}$, the transverse illuminated plane wave light field can be expressed as $U_{in}(\boldsymbol{r})=U_0\exp(-i\boldsymbol{pr})$. After scattering of the sample, the outgoing light field can be expressed as:

$$U_{out}(\boldsymbol{r})=S(\boldsymbol{r})\cdot U_{in}(\boldsymbol{r})=U_0S(\boldsymbol{r})\exp(-i\boldsymbol{pr}) \tag{S3}$$

And the complex amplitude on the Fourier plane

$$\widetilde{U}_f(\boldsymbol{k})=U_0\widetilde{S}(\boldsymbol{k}-\boldsymbol{p})\widetilde{H}(\boldsymbol{k}) \tag{S4}$$

When the sample is illuminated with a plane wave whose transverse wave vector is $\boldsymbol{p}$, the spectrum of the object will produce a translation of $\boldsymbol{p}$, so that the high-frequency spectrum of the object will be moved into the pass-band range of the transfer function. The equivalent effect is that the transfer function of the system intercepts the high frequency component of the object in the opposite direction to obtain the high frequency pass-band $\widetilde{S}(\boldsymbol{k})\widetilde{H}(\boldsymbol{k}+\boldsymbol{p})$ of the object, as shown in the area intercepted by the blue circle in Fig. S4b. Therefore, in this direction, the pass-band of the system is extended from the case where the original cut-off frequency is k_c to $p+k_c$. Obviously, the extent of the spread of the pass-band depends on the size of the transverse wave vector of the illumination wave. Isotropic resolution can be improved in all directions by changing the illumination direction of the incident light, as shown in Fig. S4c.

The inner and outer diameters of the spectral space loop directly affect imaging performance, which can be obtained through parameters in real space. Specifically, the six-segment illumination metasurface deflects the incident beam with an angle of θ, and the six beams coincide at the object plane forming an annular illumination. We used an LED as the light source, which was slightly divergent with an angle of α after collimation. The illumination spectrum takes on a ring shape in the frequency domain, with its inner (k_1) and outer (k_2) frequency radius defined by the angle θ and α. The relationships are written as:

$$k_1=\frac{2\pi}{\lambda}(\sin\theta-\sin\alpha),\ k_2=\frac{2\pi}{\lambda}(\sin\theta+\sin\alpha) \tag{S5}$$

where λ is the wavelength, and the annular width is $4\pi\sin\alpha/\lambda$. The cut-off frequency (k_c) of the partially coherent illumination system is determined by the numerical aperture (NA) of the lens and k_2:

$$k_c=\frac{2\pi}{\lambda}NA+k_2 \tag{S6}$$

The maximum k_c is the diffraction limit 4πNA/λ, which directly influences the resolution performance.

As analyzed, when $\sin\theta$ < NA, the spatial frequency cutoff can be expressed as k_c = 2π(NA+sinθ)/λ + 2πsinα/λ. In this case, increasing the annular width contributes to an enhancement in resolution until k_c approaches the diffraction limit. However, when θ becomes sufficiently large (e.g., sinθ = NA), the cutoff frequency has already reached the diffraction-limited maximum. Under such conditions, further expanding the annular width (k_2-k_1) does not improve resolution but primarily affects image contrast.

For a quantitative evaluation, we introduce the weak object approximation to linearize the partially coherent imaging formation process, and derive the amplitude contrast transfer function (ACTF) to evaluate the resolution and contrast of a partially coherent imaging system. The ACTF is written as:

$$ACTF = TCC(\boldsymbol{u},0) + TCC^{*}(-\boldsymbol{u},0), \tag{S7}$$

where $TCC(\boldsymbol{u},0)=\int S(\boldsymbol{u}')H(\boldsymbol{u}'+\boldsymbol{u})H^{*}(\boldsymbol{u}')d\boldsymbol{u}'$, $H(\boldsymbol{u})$ is the coherent transfer function (CTF), $S(\boldsymbol{u})$ represents the intensity distribution in the ring aperture plane, and $\boldsymbol{u}$ is the two-dimensional frequency coordinate. We simulated the ACTFs under annular illumination with different ring size. In the simulation, the NA of the lens is 0.5, and the wavelength is 470 nm. The center frequency (k_0) of the annular illumination is set to match NA/λ, the annular width in frequency domain is defined as nk_0 (n = 0.004, 0.5, 1, 2). Figure S4(d)-(g) demonstrates the influence of annular width on the ACTF, indicating that the illumination wave vectors larger than k_0 do not contribute to the system's resolution, and the cut-off frequency of the system remain at the diffraction-limit of $2k_0$. The contrast increases with larger annular width, and ACTF changes to the modulation transfer function of an incoherent system when the annular width equals to $2k_0$.

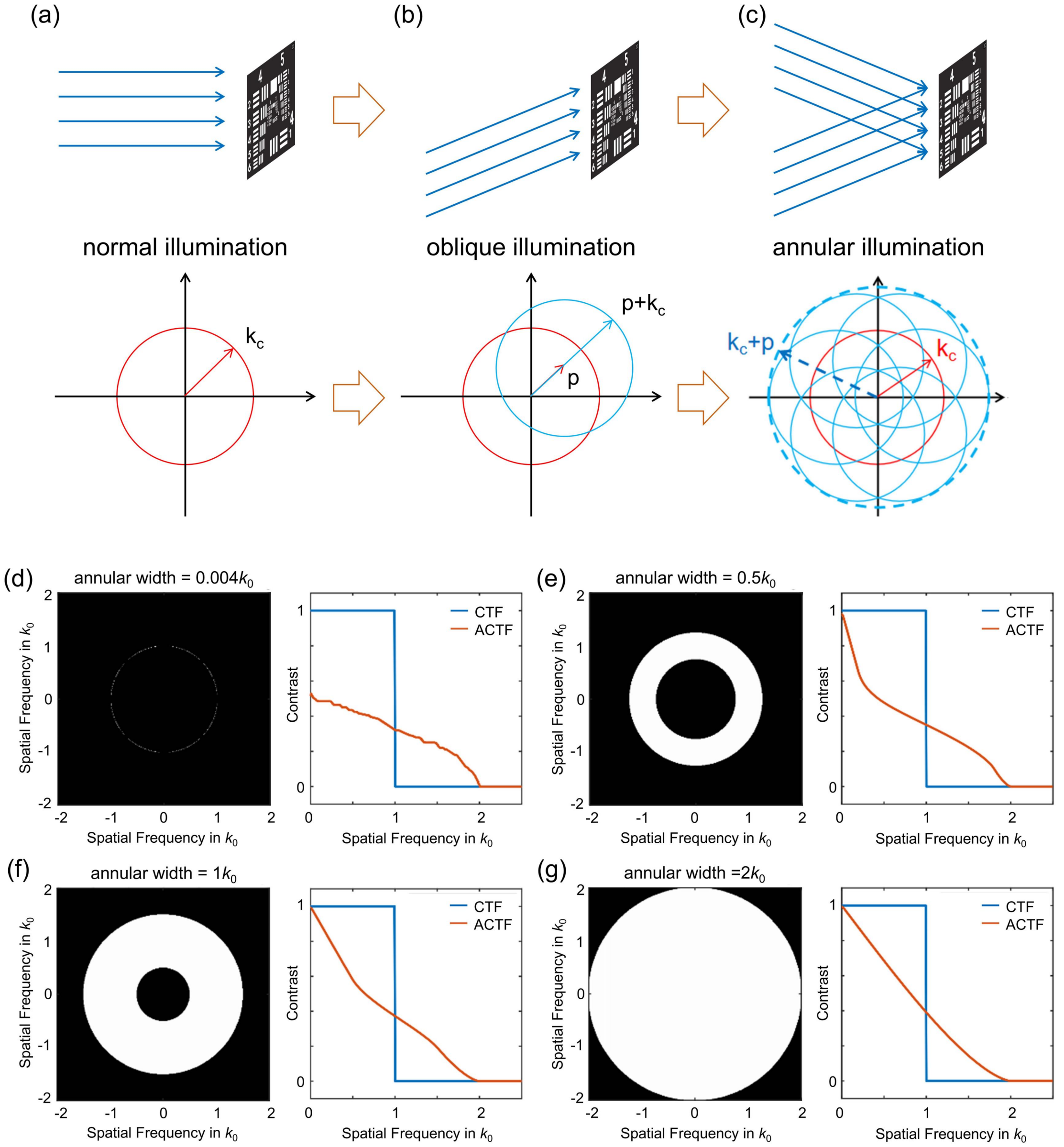

Fig. S4 (a)-(c) Annular lighting can expand the spectrum pass-band, so that high-frequency information can be received, thus improving the resolution. (d)-(g) The spectrums of the annular illumination with different annular width and comparison of corresponding ACTFs with CTFs. (d) The annular width equals 0.004k_0, meaning a near-perfect annular illumination without spectrum broadening. (e), (f) The annular widths equal 0.5k_0 and 1k_0. The contrast increases, but the cut-off frequencies of the two cases remain the same as 2k_0. (g) The annular width equals 2k_0. The ATF changes to the modulation transfer function of an incoherent system.

Supplementary note 5: Comparison between doublet and singlet

To demonstrate the large field imaging capability of our meta-microscope, we propose the image of a single-layer spherical-aberration-free metalens with the same aperture diameter and focal length for comparison, as shown in Fig. S5. It is evident that the image from the singlet metalens shows considerable blurriness at a relatively large object height and is almost unable to observe tiny structures. To be more specific, Figure S5c presents the resolution assessed at different object heights (top) alongside the corresponding results from a spherical-aberration-free singlet metalens (bottom).

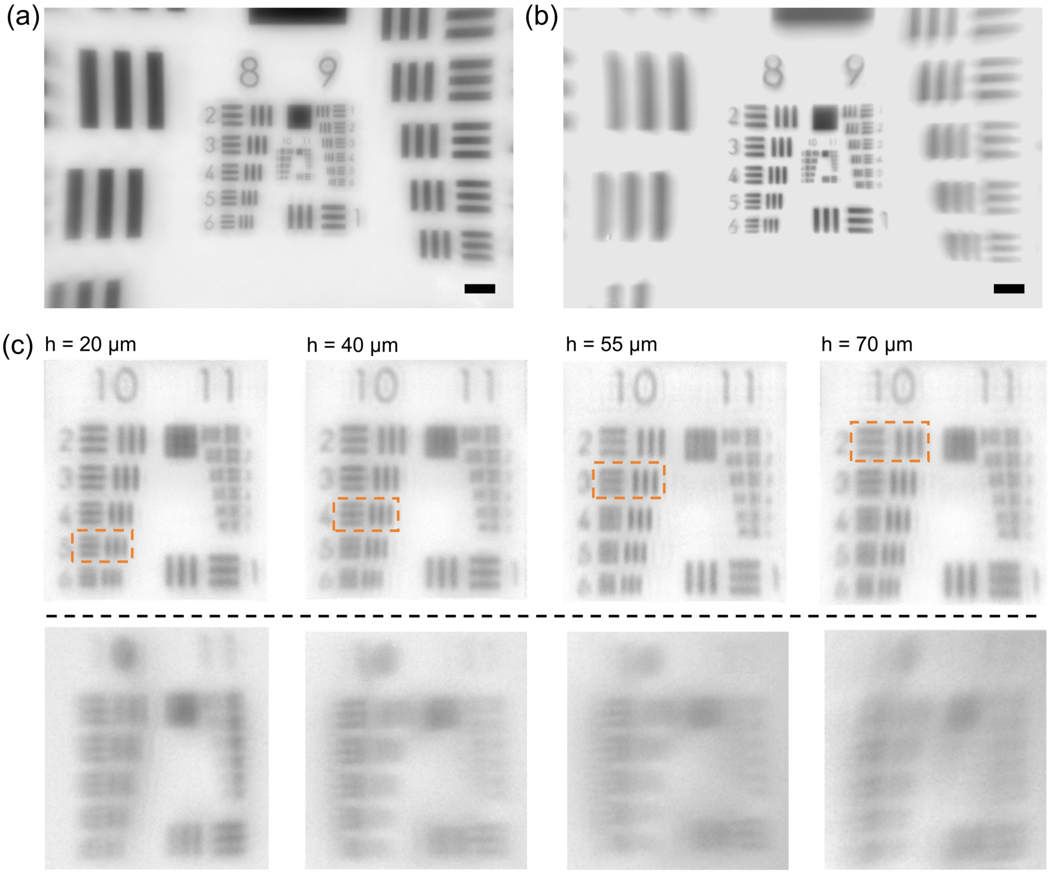

Fig. S5 Comparison between the metalens doublet and singlet. (a) Image of the metalens doublet. (b) Image of a single-layer spherical-aberration-free metalens. Scale bar = 15 μm for a&b. (c) The resolution assessed at different object heights (top) alongside the corresponding results from a spherical-aberration-free singlet metalens (bottom).

Supplementary note 6: Details of the meta-microscope with NA 0.3

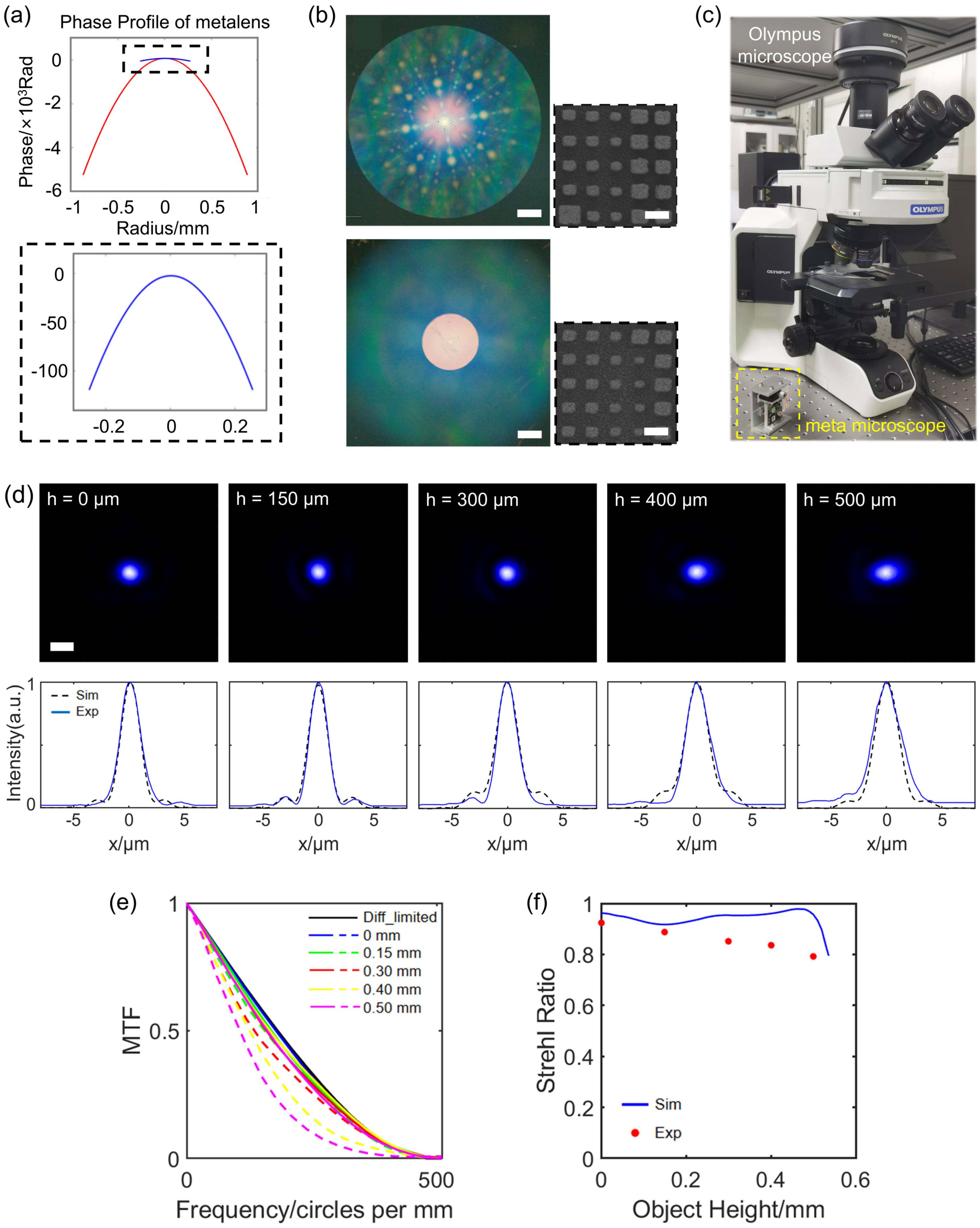

Fig. S6 Details of the meta-microscope with object NA=0.3. (a) Phase profiles of the metalenses. (b) Optical and top-view scanning electron microscope (SEM) images of the fabricated metalenses. Scale bar = 200 μm and 300 nm, respectively. (c) Photographic image of the traditional Olympus microscope together with the meta-microscope prototype. (d) Measured focal spot diagrams and focal spot intensity profiles at object heights of 0, 150, 300, 400, 500 μm. Scale bar = 2 μm. (e) MTF calculated from the PSFs. (f) Strehl ratios of different object heights.

Supplementary note 7: Detailed parameters from relevant studies for comparison

Table 1 Parameters and performance of some metalens microscope

		wavelength/ nm	resolution/ µm	FOV/µm	diameter of lens/µm	relative FOV	SBP	w/o FOV optimized
Li et al.[24]		800	2.19	80	2000	0.04	3.3×10^{3}	No
Wang et al.[25]		620	0.78	40*	200	0.2	6.5×10^{3}	No
Kwon et al.[27]		850	1.74	140	660	0.197	1.4×10^{4}	Yes
Long et al.[28]		532	0.87	50*	400	0.125	8.1×10^{3}	No
Liu et al.[29]		525	0.78	125	400	0.313	6.3×10^{4}	Yes
Kim et al.[30]		480-680	0.78	30*	1000	0.03	3.6×10^{3}	No
Wang et al.[31]		450	1.23	120*	2000	0.06	2.3×10^{4}	No
Sun et al.[32]		1550	3.78	600*	4000	0.15	6.2×10^{4}	No
Arbabi et al.[33]		605	0.81	90*	1600	0.056	3.0×10^{4}	No
Wang et al.[46]		920	0.92	45*	1000	0.045	5.9×10^{3}	No
Chu et al.[47]		532	2.76	320*	2000	0.16	3.3×10^{4}	No
This work	NA 0.5	470	0.31	150	450	0.333	5.8×10^{5}	Yes
	NA 0.3	470	0.62	1000	1700	0.588	6.4×10^{6}	Yes

*Since some studies did not explicitly define the FOV size, we estimated it by analyzing the captured images. The SBP is calculated as the product of the FOV area and the spatial frequency range[34]

$$SBP = \pi\left(\frac{FOV}{2}\right)^{2} \cdot \pi\left(\frac{1}{reslution}\right)^{2}$$