Meta-optical Front-end for Imaging and Computing

Luocheng Huang

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Arka Majumdar, Chair

Karl F. Böhringer

Mo Li

Sajjad Moazeni

Igor Novosselov

Program Authorized to Offer Degree:
Electrical and Computer Engineering

University of Washington

Abstract

Meta-optical Front-end for Imaging and Computing

Luocheng Huang

Chair of the Supervisory Committee:
Associate Professor Arka Majumdar
Electrical and Computer Engineering

This thesis delves into the world of meta-optics, with its immense potential for revolutionizing imaging and computing. The motivation arises from the limitations of conventional optics, inspiring the exploration of innovative meta-optic solutions. The work begins with a foundation in metasurfaces, examining issues such as chromatic aberrations in metalenses and the development of multiwavelength polychromatic metalenses. It further delves into dispersion engineering and computational imaging, setting the stage for the discussions that follow. Subsequent chapters explore the extension of depth of focus and the application of long-wavelength infrared (LWIR) meta-optics, examining design, fabrication, and characterization through both forward and inverse design approaches. The study also explores the exciting potential of meta-optics for optical computation, an area that could greatly benefit from the speed and efficiency of light-based calculations. The thesis concludes with an outlook on the future of meta-optics, discussing potential developments such as standardizing image quality benchmarking, designing larger aperture systems, enhancing efficiency, and future directions in computational meta-optics. Overall, this thesis illuminates the potential of meta-optics to transform imaging and computing, aiming to inspire further research and innovation in this burgeoning field.

Table of Contents

List of Figures

Acknowledgments

I would like to express my deepest gratitude to all those who have provided me their guidance, support, and encouragement throughout the duration of my PhD journey. I am humbly indebted to all the beautiful minds who have significantly contributed to my academic and personal growth.

Foremost, my deepest appreciation is extended to my advisor, Arka, for his boundless wisdom, patience, and kindness. The depth of your understanding has provided a solid foundation for my research, and your unending patience has been a source of comfort during challenging times.

To Shane Colburn, thank you for your invaluable teachings. Your knowledge has significantly influenced my scholarly development and has enriched my perspective of our field. I am grateful to Alan Zhan for meticulously going over his code line by line. Your detailed approach has elevated my comprehension and my abilities in our research work. I extend my heartfelt thanks to James Whitehead, a good friend and a persistent motivator. Your unwavering faith in me has been a beacon of hope in the darkest of times. Quentin Tanguy, your expertise and assistance in nanofabrication have been indispensable. Thank you for sharing your knowledge and for guiding me through complex processes. To Johannes Froch, thank you for generously sharing your profound knowledge and for inspiring me to strive for more.

I appreciate Zhuoran (Roger) Fang for instructing me on alignment in Ebeam lithography. Your guidance has been a vital component of my research work. A special thanks to Anna Wirth-Singh for her exceptional work in fabricating and characterizing LWIR optics. Your dedication and precision have contributed significantly to our research outcomes. Thank you, Forrest Miller, for the enlightening discussions on light. Your insights have broadened my understanding and have ignited my curiosity in ways I could not imagine. To David Rosser, your Economics magazine and intriguing stories have provided me a refreshing and insightful perspective outside of my research field. Your wisdom has enriched my experience. My thanks extend to Saswata Mukherjee for taking measurements on my behalf. Your invaluable assistance has made my research journey manageable. Albert Ryou, I appreciate your wisdom on life. Your philosophical insights have been a source of strength, enabling me to persevere during arduous times.

On a personal note, my wife Amber, thank you for your unwavering support and encouragement throughout this journey. Your love and patience have been my rock. You're the best mother our daughter could have, and I am grateful for your enduring love. Lastly, to Abby, our wonderful daughter. Your existence alone has been my greatest source of inspiration. Thank you for being the best daughter a father could ask for.

This journey has been a profound learning experience, and every individual mentioned here has played a crucial role in shaping me. For that, I am eternally grateful.

Dedication

to my wife, Amber

and my daughter, Abigail

Chapter 1

Introduction

1.1 Motivation

In the evolving world of optics and photonics, the advent of meta-optics has paved the way for a transformation in the field of imaging and computation. Meta-optics refers to the use of artificially structured, sub-wavelength scale materials—metamaterials—engineered to manipulate light in ways that natural materials cannot. Their potential lies in their ability to control multiple attributes of light, including phase, amplitude, polarization, and even frequency, thereby providing unprecedented abilities in managing light propagation and interaction. Here, we delve into the motivations behind adopting meta-optics for imaging and computation applications.

1.1.1 Meta-optics for Imaging

In modern life, cameras are indispensable, providing an invaluable functionality in preserving a scene as perceived by the human eye. With digital cameras now readily available to consumers, both professionals and hobbyists are able to experience how easily a photo can be captured, viewed, and shared. While already ubiquitous, many emerging applications, such as the internet of things, machine vision, virtual presence, angioscopy, and bioimaging, require new optical technologies that must dramatically miniaturize the form factor of exist-

ing cameras. A closer look at these requirements shows that the main bottleneck for further miniaturization comes from the lenses in these cameras [1, 2]. To maintain high spatial resolution and collect sufficient light, we need lenses with both large aperture and numerical aperture (NA), which push up the required overall volume of the cameras. While decreasing focal length can reduce the free-space volume between the optic and sensor, shorter focal-length refractive lenses require higher curvatures, which increase the thickness, weight, and manufacturing complexity of the optic itself. Finally, to ensure aberration-free imaging, often a compound lens system is used, which further increases manufacturing and packaging complexity as well as cost. While miniaturized imaging systems need new optics, extremely large optical systems, for example, space telescopes, can also benefit from novel, lightweight optical components.

To that end, there is a long history of trying to miniaturize optics by circumventing traditional lens designs. Refractive Fresnel optics were one of the early examples of such miniaturization, used to replace large spherical lenses in lighthouses. By treating the design in terms of discrete surface zones, rather than a continuous volume, these lenses could provide the same functionality of a spherical lens but with significantly lower weight and size. While different in their operating principle, diffractive Fresnel lenses also provide an even larger reduction in volume and weight for optical systems. Many of these Fresnel lenses exploit amplitude modulation, which limits the overall efficiency. Via phase modulation, the transmission efficiency can be significantly increased. Many of these diffractive optics are primarily used in nonimaging applications, such as solid-state lighting or beam shaping, mainly because of large chromatic aberration. In recent years, two types of flat diffractive optics have been touted for their capabilities in imaging: multilevel diffractive optics and subwavelength diffractive optics, also known as metaoptics. Multilevel diffractive optics provide different phase shifts to the incident optical wavefront using different heights, which generally need multistage or grayscale lithography. Most of these optics have superwavelength pitch and cannot guide all the light to the 0^{th}-diffraction order. Metaoptics, on the other hand, have subwavelength pitch, which enables all the light to diffract to the 0^{th} order.

In addition, phase shifts are imparted based on a different mechanism: exploiting variations in lateral size or orientation of the individual scatterers, while maintaining the same thickness throughout the design. This enables use of single-stage lithography for fabricating metaoptics. In recent years, several works explored the differences between metaoptics and multilevel diffractive optics. [3, 4] Regardless of the differences, both types of diffractive optics suffer from strong chromatic aberration, significantly limiting their usage in full-color imaging systems. With advances in scatterer engineering via computational electromagnetics, inverse design, and computational imaging, several research groups have reported broadband imaging using these diffractive optics.

In chapter **§ 2**, we review the current field of broadband imaging using metaoptics. We discuss different means for mitigating chromatic aberration and the limitations of these approaches. We first focus on full-color imaging in the visible regime, where images are captured using a red-green-blue (RGB) camera under broadband, incoherent illumination. These criteria are used as this is how most color cameras function. Then we shift focus to meta-optical imaging in the long-wave infrered regime. We also outline the outstanding challenges and future research directions on broadband metaoptical imaging.

1.1.2 Meta-optics for Computing

Meta-optics, beyond its remarkable advancements in traditional optics, has proven to be an extraordinary avenue for computational processes as well. As we move into a new epoch of technological advancement, the advantages of applying meta-optics in computational platforms are becoming increasingly evident. The power of meta-optics in computation stems from its unique ability to manipulate light at a nanoscale, a functionality that outshines conventional electronic computation [5, 6]. Processing information at the speed of light, meta-optics offers unmatched computational speed. This ushers in a new age of ultra-low latency computing, setting the stage for real-time data processing and analytics, a capability of paramount importance in areas such as autonomous driving, real-time financial trading, and advanced telecommunications.

Furthermore, the capability of meta-optics to perform complex calculations at the physical level through light manipulation introduces novel opportunities for power-efficient computation [7, 8, 9]. In an era where power consumption and thermal management are becoming critical bottlenecks in data centers and high-performance computing systems, the use of meta-optics can significantly reduce the power budget. This results not only in improved performance but also in greener computational technologies, a necessity in our environmentally conscious world. The transformative potential of meta-optics is not limited to speed and power efficiency. The technology also allows for the miniaturization of computational devices [10, 11]. This means that powerful, high-speed computing doesn't have to take up a significant amount of physical space. The miniaturization potential of meta-optics paves the way for compact, efficient computational devices, empowering everything from wearable technologies to space-constrained IoT devices. Additionally, meta-optics holds promise in driving the growth of emerging computational fields, such as quantum computing and artificial intelligence [12, 13, 14]. By addressing the challenges of scalability and speed in quantum computing, meta-optics can potentially accelerate the realization of fully-functional quantum computers. Similarly, in artificial intelligence, meta-optics' unique advantages could allow for faster machine learning and deep learning computations, speeding up AI training and inference processes. In this thesis, we cover a new kind of optical computational framework, enabling high throughput, light-speed meta-optics enabled computation.

1.2 Thesis outline

This thesis explores the capabilities of meta-optics, a technology that leverages artificially structured sub-wavelength meta-surfaces, opening up new paradigms in the world of optics and photonics.

The work begins by delving into the motivations behind using meta-optics for imaging and computing, discussing their potential to revolutionize traditional optical systems in chapter **§ 1**. The reader is provided with the necessary background on meta-surfaces in chapter **§ 2**, focusing on chromatic aberrations, multiwavelength polychromatic metalenses, dispersion

engineering, and computational imaging.

We then examine the process of extending the depth of focus using meta-optics in chapter **§ 3**, demonstrating improved performance in image acquisition. In addition, the thesis covers the design and fabrication of meta-optics, with a specific emphasis on long-wavelength infrared (LWIR) meta-optics. Forward and inverse design methodologies are employed to address the challenges and opportunities in meta-optic design in chapter **§ 4** and **§ 5**.

In the final section in chapter **§ 6**, the study takes a significant leap by exploring the potential of meta-optics for optical computation. An exciting hybrid system combining meta-optics with digital neural networks is proposed, marking a step towards the future of computational optics.

In conclusion (chapter **§ 7**), the thesis provides insights on the future of meta-optics, discussing prospects and challenges. It emphasizes the need for standardizing benchmarking of image quality, larger aperture, efficiency, real-world imaging applications, video capture, and the exciting field of computational metaoptics. This work bridges the gap between meta-optics and applications, leading us closer to a new era in optical technology.

Chapter 2

Background on Metasurfaces

2.1 Chromatic Aberrations in Metalens[1]

A conventional metaoptical lens, commonly known as a metalens, with a hyperboloidal phase profile, suffers from strong axial chromatic aberration: lights of different wavelengths focus at different distances from the lens. While the phase response of an individual scatterer in a metaoptic does depend on the optical wavelength, the primary reason for chromatic aberration is phase wrapping [15] (**Fig. 2.1**(a)). In a metaoptic, like any diffractive optics, the phase is wrapped to ε when the phase reaches $2\pi + \varepsilon$. The spatial locations where the phase value reaches $2\pi + \varepsilon$, however, depend on the design wavelength. In a metalens, phase wrapping occurs at specific spatial locations, making it very difficult to achieve the ideal phase distribution for every wavelength [15]. This can be easily explained for a quadratic lens phase profile $\phi(r)$, which is an approximation of the hyperboloid phase for $r << f, f$ being the focal length and r the radial coordinate)

$$\phi(r) = \frac{2\pi}{\lambda}\frac{r^2}{2f} \tag{2.1}$$

Assuming a nominal wavelength λ_o and a focal length f_0, the phase wrapping radii will

[1]The content of this chapter is based on the following publication:
Luocheng Huang, Shane Colburn, Alan Zhan, Arka Majumdar, "Full-Color Metaoptical Imaging in Visible Light." Adv. Photonics Res., 3: 2100265

be given by

$$r_{\mathrm{N}} = \sqrt{2N\lambda_0 f_0} \tag{2.2}$$

where N is an integer that governs the amount of phase shift. In a metalens, the phase-wrapping points are fixed, which makes the focal length of the metalens f inversely proportional to the optical wavelength λ. Thus, the chromatic aberration in a metalens comes from the overall phase distribution and not just the wavelength dependence of each scatterer. To achieve broadband operation of a metalens, we need to realize different phase distributions for each wavelength based on a fixed distribution of scatterers.

As this thesis will discuss, it is indeed possible to create such phase profiles for multiple different wavelengths by appropriately engineering the scatterers. Many of these scatterers consist of multiple subwavelength components, often called a metamolecule in analogy of a molecule being created from multiple atoms. By judiciously designing these metamolecules, one can engineer both the group delay and group delay dispersion in a metalens. Thus researchers have demonstrated broadband focusing using metaoptics [16, 17, 18, 19]. The aperture and NA of these demonstrated broadband metalenses remain very small, and thus the cameras are limited in terms of light collection and spatial resolution. Such limitations on aperture and NA of broadband metalenses have been extensively analyzed using the fundamental time-bandwidth product [20]. Specifically, researchers showed that the maximum achievable fractional optical bandwidth for a broadband metalens with scatterers behaving as optical waveguides is governed by the relation

$$\frac{\Delta\omega}{\omega_{\mathrm{c}}} \leq \frac{t\Delta n}{F} \frac{\sqrt{1-(\mathrm{NA}/n_{\mathrm{b}})^2}}{1-\sqrt{1-(\mathrm{NA}/n_{\mathrm{b}})^2}} \tag{2.3}$$

Here, $\frac{\Delta\omega}{\omega_{\mathrm{c}}}$ is the fractional optical bandwidth; t is the thickness of the metaoptic; F is the focal length; Δn is the index contrast between the metalens material and surrounding material; and n_{b} is the background material index. Such limits have also been explored for metaoptics using Panchratnam-Berry-phase elements [20, 21]. These analyses clearly show

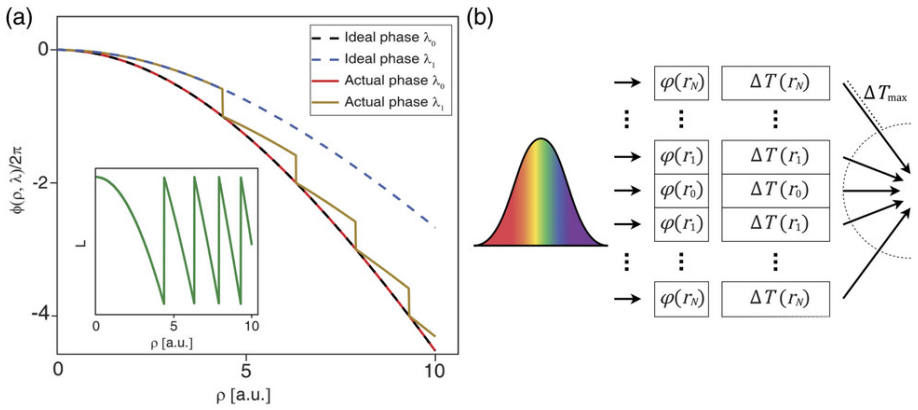

Figure 2.1: Chromatic aberration in a metalens. a) The phase wrapping in a metalens occurs at specific spatial locations, causing chromatic aberration. b) By considering the time taken by the light to reach the focal spot from the center and the edge of the lens, a fundamental limit on the fractional bandwidth of achromatic metaoptics can be derived. a)[15]. Copyright 2016 , Optical Society of America. b)[20]. Copyright 2020 , Optical Society of America.

that, for a metalens with a given thickness, there is a trade-off between aperture, NA, and the fractional bandwidth. We will review different ways to achieve broadband operation, which is bounded by this fundamental limit. We will also explore ways to possibly circumvent this limit using computational imaging.

2.2 Multiwavelength Polychromatic Metalens

Early attempts at creating broadband metaoptics resulted in polychromatic, multiwavelength elements [22, 23]. These polychromatic metalenses exhibit the same focal lengths for distinct target wavelengths, but the light does not focus on the same plane at intermediate wavelengths. Such optics could potentially provide achromatic operation, if the illumination could always be restricted to certain discrete wavelengths. The basic principle underpinning such polychromatic designs is to engineer metamolecules that simultaneously satisfy the required phase shifts at each spatial point in the metaoptic for the set of design wavelengths. Initial

reports for such designs were limited to infrared wavelengths (1300, 1550, and 1800 nm) and cylindrical lensing based on a diatomic grating structure [22] (**Fig. 2.2**(a)). A similar concept was explored for creating a 2D unit cell for polarization-insensitive, multiwavelength operation (200μm aperture and NA of 0.46 (**Fig. 2.2**(b)) [15] Here, the multiwavelength operation was demonstrated at 915 and 1550 nm. The reason for such different wavelengths was that the unit cell comprised two very different diameters of cylindrical scatterers to minimize mode coupling. Following similar techniques, various spatial multiplexing methods capable of creating polychromatic metalenses have been reported (**Fig. 2.2**(c)) [24]. While such techniques can be applied to larger apertures, they are not suitable for broadband operations and have not yet been demonstrated in color-imaging applications. Given that the essence of these designs is metamolecule engineering, inverse methods are well

Other than unit cell engineering, researchers have demonstrated multiwavelength operation by vertically stacking metaoptics (**Fig. 2.3**(a)) [26]. In one such approach, plasmonic scatterers are used to create three metalenses independently optimized for focusing light at 650,550 , and 450 nm. The aperture of the metalenses used was $\approx 200\mu$m and NA is 0.1. Similar stacked metaoptics have also been demonstrated for polychromatic imaging at infrared wavelengths [27]. In another approach, by engineering metamolecules near the zone boundaries, that is, where the phase wraps modulo 2π, polychromatic focusing was achieved for 658 , 532, and 488 nm using a 2 mm-diameter metalens with an NA of 0.7 (**Fig. 2.3**(b)) [28]. This work also demonstrated color projection of an image for applications related to virtual reality. We note that multichrome diffractive optics have been explored using multi-order diffractive optics before [29, 30, 31]. However, the design and operating principles of multiwavelength metaoptics are different due to lack of any higher orders of diffracted light.

2.3 Dispersion Engineering

While there are many impressive demonstrations of polychromatic lenses that do have applications in projection and while there is a large degree of control over the illumination wavelengths, these designs are not quite suitable for broadband imaging. The spectral filters

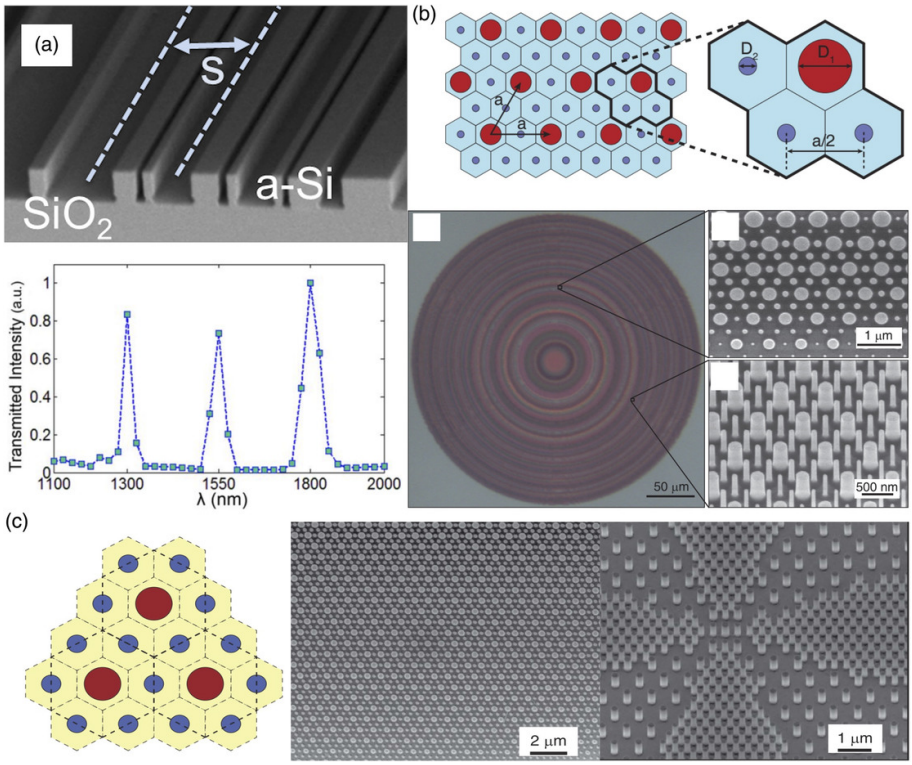

Figure 2.2: Polychromatic metaoptics via unit-cell engineering. a) Cylindrical metalens made of amorphous silicon focuses light at three specific wavelengths at the same plane. The pitch of the metaoptics is $S = 1\mu$m. b) By engineering the unit cell, a metalens operates at two different wavelengths. c) Polychromatic metalens can be realized via spatial multiplexing of the scatterers. a) [22]. Copyright 2015 , AAAS. b) [25]. Copyright 2016, Optical Society of America. c) [24]. Copyright 2016, Springer Nature Ltd. suited, as recently reported [25] in a 1 cm-aperture metaoptic used for projection in a virtual reality system, where the wavelengths of the light source can be easily controlled.

used in color cameras are quite broad, spanning from tens to hundreds of nanometers in bandwidth relative to the assumed discrete, laser-like illumination of polychromatic designs. As such, if polychromatic designs were used with color cameras, the broad bandwidth of the color filters would cause defocused light from other incoming wavelengths to arrive at the

sensor, blurring the resultant image. To ensure high-quality broadband imaging, one needs to ensure that the optics focus all colors in the desired wavelength band (e.g., the full visible spectrum) on the same plane. The phase profile $\phi(r,\omega)$ of a metalens can be expanded using a Taylor series near the design frequency ω_0

$$\phi(r,\omega) = \phi\left(r,\omega_{\mathrm{o}}\right) + \left.\frac{\partial\phi}{\partial\omega}\right|_{\omega=\omega_{\mathrm{o}}} \left(\omega-\omega_{\mathrm{o}}\right) + \frac{\partial^2\phi}{2\partial\omega^2}\left(\omega-\omega_{\mathrm{o}}\right)^2 + \ldots \tag{2.4}$$

Broadband imaging requires designing metamolecules to achieve both phase $\phi\left(r,\omega_{\mathrm{o}}\right)$ and group delay $\frac{\partial\phi}{\partial\omega}$ over a broad range of wavelengths. By exploiting the full electromagnetic interactions between scatterers in a metamolecule, it is indeed possible to achieve desired phase and group delay for different wavelengths. Such a metamolecule design, also known as dispersion engineering for broadband imaging, was first reported under reflection: a bandwidth of 140 nm nm at the central frequency of 1520 nm [16] and a bandwidth of 60 nm at the central frequency of 520 nm [32] were achieved. The reflective geometry was chosen to facilitate the engineering of meta-atoms. A reflective design, however, is far more limited in its applications for consumer cameras.

Progress was made soon after, with dispersion engineering exploited for broadband, transmissive metaoptics [17, 18]. By engineering the metamolecules, light over the entire visible spectrum can be focused at the same focal length. In one work, the metalens demonstrated broadband focusing over the $400-660$ nm range with an aperture of 100μm and NA of 0.106 [18] (**Fig. 2.4**(c)). In another work, a metalens of aperture 200μm and NA of 0.02 was used to focus the $470-670$ nm range (**Fig. 2.4**(a)). Both of these demonstrations depend on the polarization of the incident light. By engineering the meta-atoms, researchers demonstrated an polarization-insensitive achromatic metalens: while each anisotropic metamolecule has some polarization dependence, by restricting the rotation angle of each anisotropic element to either 0 or 90°, polarization-insensitive operation was achieved [33]. Here, the achromatic operation is demonstrated over $460-700$ nm using a 26.4μm-aperture metalens of NA of 0.2 . While such broadband focusing provides a route to color imaging, the demonstrated aper-

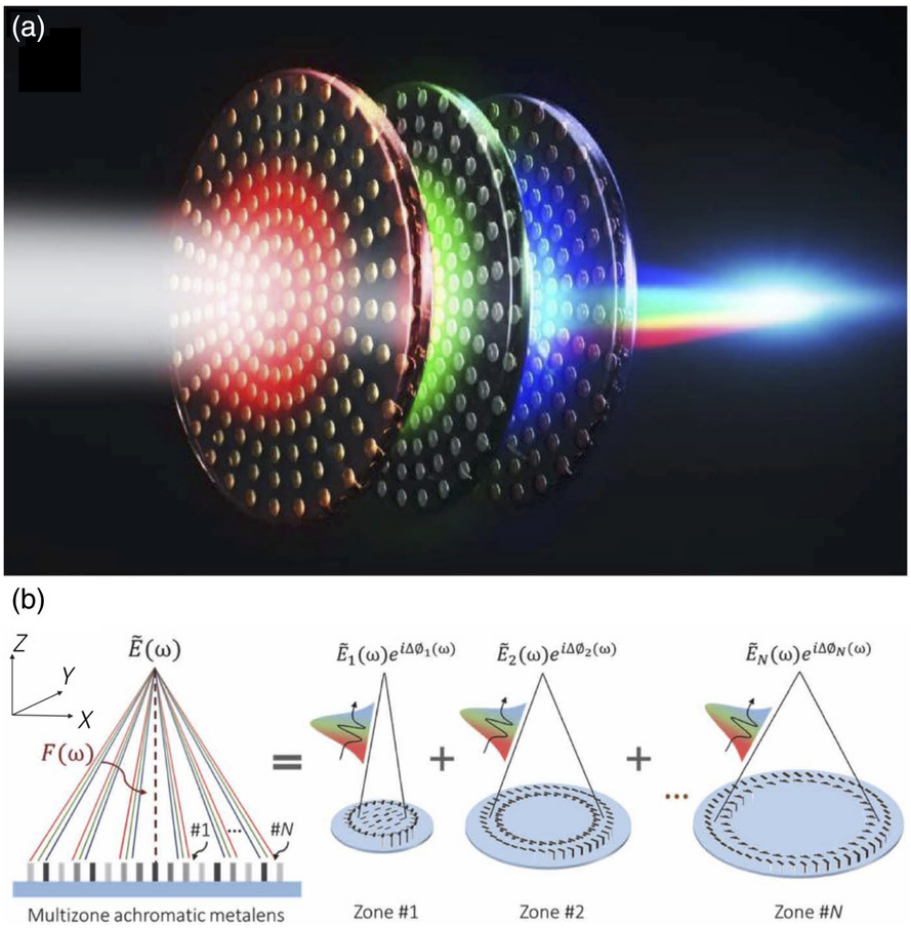

Figure 2.3: RGB polychromatic metalens. a) By stacking three metasurfaces, red, green, and blue (RGB) colors can be focused on the same plane. b) By engineering the metamolecules at the zone boundary, red, green, and blue light can be focused on the same plane. a) [26]. Copyright 2017, Springer Nature Ltd. b) [28]. Copyright 2021, AAAS.

tures and NAs are very small for most practical applications. Similar dispersion engineering methods have also been used for infrared wavelength range ($1200-1650$ nm), though these demonstrations equally suffer from very small apertures [19] (**Fig. 2.4**(b)). By exploiting the height as another parameter, 3D-printed achromatic metalenses have been demonstrated in the $1000-1800$ nm range, although the aperture of these lenses is less than 100μm.[25] Using

fishnet metaoptics, researchers also demonstrated achromatic focusing in the $640 - 1200$ nm range; [34]. however, the apertures of these lenses are very small, $\approx 50\mu$m, and no color imaging was reported. As such, dispersion-engineered metalenses suffer from a stringent trade-off between the NA and aperture size versus achievable optical bandwidth. A more detailed review of dispersion-engineered metaoptics can be found in other studies [35, 36].

2.4 Computational Imaging

Another approach to achieve full-color imaging is by exploiting computational imaging. Using metaoptics, we can capture an intermediate image that preserves as much information as possible over all the color channels in the sensor. This intermediate image can be subsequently decoded in software to extract a full-color image. In this way, the modulation transfer function (MTF) at one specific wavelength will always be inferior to that of a metalens specifically designed for that same wavelength; however, when considering the full visible spectrum, we can increase the total integrated MTF relative to a traditional metalens. In this process, we need to ensure the MTF is broad, that is, the area under the 1D MTF curve should be maximized, while also ensuring that the point spread function (PSF) is the same for all the wavelengths in a given color channel. This essentially means that we need to have a large Strehl ratio, defined as the ratio between the volume under the MTF surface of the metaoptics and the volume under the MTF surface for a diffraction-limited lens.

2.4.1 Forward Design

One way to achieve the aforementioned functionality is by exploiting extended depth of focus (EDOF) metalenses. These lenses produce an elongated focal zone or line instead of diffracting light to a focal spot. Due to chromatic dispersion, the centroid of the focal line changes with wavelength; however, if we can extend the depth of focus significantly, we can ensure that light at every wavelength reaches the sensor in a similar manner, mapping into a color-invariant PSF. One way to achieve this is using cubic metaoptics [37]. Extending the depth of focus using a cubic-phase plate was proposed more than 20 years ago and ar-

guably, pioneered the field of computational imaging [37, 38, 39, 40]. While the application of chromatic correction of refractive lenses using such wavefront coding was explored earlier, this was only recently applied for metaoptics (**Fig. 2.5**(a)) [37]. While an ordinary metalens can provide a broad MTF at one wavelength (here, green), for red and blue light, the MTF degrades significantly, exhibiting zeros in spatial frequency that represents an irrecoverable loss of scene content. Cubic metaoptics, however, not only produce a wavelength-invariant MTF but also capture a large range of spatial frequencies without any zeros, enabling deconvolution for image extraction. Unfortunately cubic EDOF metaoptics also produce an accelerating beam that induces lateral chromatic aberration. We demonstrate, in chapter **§ 3**, that this can potentially be circumvented using a rotationally symmetric EDOF lens, such as a log-asphere or shifted-axicon lens [41]. These demonstrated EDOF lenses based on canonical-phase masks unfortunately enable little control over the MTF.

2.4.2 Inverse Design

One way to increase the operating optical bandwidth and the total integrated MTF is to use inverse design to arrive at high-performance solutions with nonintuitive forms. Here, one can define the desired performance of the metaoptics using figure of merit (FOM) and optimize the scatterer distribution to reach the desired FOM. A recent work on an inverse-designed EDOF lens defines the FOM to maximize the intensity along a line in the optical axis [42]. The optical bandwidth can be defined as the wavelength range where the correlation function between the MTFs remains larger than 0.5. The inverse-designed metaoptics thus indeed showed larger optical bandwidth compared to that of a metalens or other canonical EDOF lenses (**Fig. 2.5**). In fact, a bandwidth of 290 nm was reported in inverse-designed EDOF metaoptics, at least twice larger than other EDOF metaoptics. The image quality captured via the EDOF metaoptics also outperforms the traditional EDOF metaoptics, measured via the structural similarity (SSIM) factor: for red, green, and blue channel, the calculated SSIM factors are $0.64(0.31), 0.8(0.65)$, and $0.48(0.47)$ for inverse designed (cubic) EDOF metaoptics. (a)

2.4.3 End-to-End Design

In most demonstrations of metaoptics for full-color computational imaging, the metaoptics and the reconstruction method are separately designed. Moreover, the deconvolution methods typically used are based on linear operators, the performance of which is solely determined by the area under the 1D MTF curve and requires higher signal to noise ratio (SNR) to achieve the same spatial resolution as in a traditional system. Researchers have already argued that joint optimization of optics and software together can reduce form factor and enhance performance [43]. Such codesign techniques, also known as “end-to-end" design, have already found applications in diffractive optics [44]. Recently, such efforts have been applied to metaoptics for imaging and polarimetry [45], as well as for broadband full-color imaging [46]. Here, simple cylindrical scatterers are used as the meta-atoms, and a differentiable proxy function is developed to connect the scatterer geometry to the phase distribution. The image formation via metaoptics and computational reconstruction is modeled in a differentiable pipeline (**Fig. 2.6**a) that can be optimized using automatic differentiation. To account for geometric aberration, spatially varying PSFs are considered. Here, the computational reconstruction techniques leverage a combination of classical, linear deconvolution that helps to preserve generality, while also utilizing state-of-the-art feature extraction networks and nonlinearities to maximize the image quality and denoising, while maintaining high spatial resolution. **Fig. 2.6**b shows the images captured using the end-to-end designed metaoptics and computational reconstruction. The end-to-end optimized metaoptic has an aperture of 500μm and a focal length of 1 mm and the deconvolution routine only required 58 ms. Figure 6c shows the same images captured using a commercially available six-element compound lens. The image quality obtained via metaoptics is comparable with that of the refractive lens, as quantitatively proven using SSIM and peak signal-to-noise ratio metrics [46].

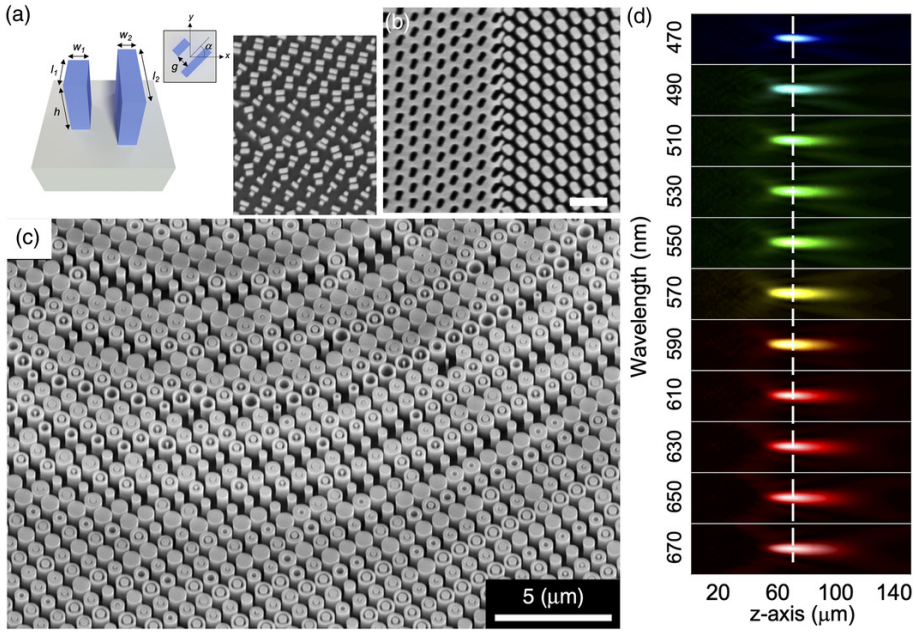

Figure 2.4: Dispersion-engineered metaoptics. a) A dispersion-engineered TiO_2 metalens can focus light over a continuous wavelength range (scale bar in the scanning electron micrograph (SEM) is 500 nm). b) A dispersion-engineered GaN metalens exploits both geometric and propagation phases to focus light over a broad wavelength range (scale bar in the SEM is 10μm). c) Metamolecule engineering in silicon allowed broadband focusing of infrared light in a polarization-insensitive fashion. d) Measured data for a dispersion-engineered metaoptics showing focusing of different colors at the same focal length. a,d) [17], Copyright 2018, Springer Nature Ltd. b) [18]. Copyright 2018, Springer Nature Ltd. c) [19]. Copyright 2018 , Springer Nature Ltd. full-color image. In this way, the modulation transfer function (MTF) at one specific wavelength will always be inferior to that of a metalens specifically designed for that same wavelength; however, when considering the full visible spectrum, we can increase the total integrated MTF relative to a traditional metalens. In this process, we need to ensure the MTF is broad, that is, the area under the 1D MTF curve should be maximized, while also ensuring that the point spread function (PSF) is the same for all the wavelengths in a given color channel. This essentially means that we need to have a large Strehl ratio, defined as the ratio between the volume under the MTF surface of the metaoptics and the volume under the MTF surface for a diffraction-limited lens.

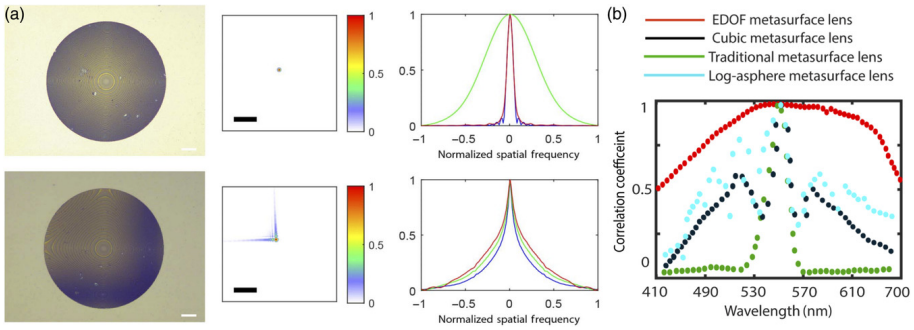

Figure 2.5: EDOF metalens for full-color imaging: a) Using a cubic-phase mask, the depth of focus can be extended enough to ensure that the MTFs at the sensor plane are identical for all the wavelengths in the visible range. Top row is ordinary hyperboloid metalens, with sharp focused PSF under green light in one plane (second column). However, other wavelengths are defocused in that plane, showing narrow MTF for red and blue (third column). For cubic metasurface, we see an extended PSF at the sensor plane, which makes the MTF narrower. The MTFs are similar for all wavelengths and have a higher cutoff spatial frequency than that of the metalens for red and blue. The scale bar for the optical images is 25μm. b) The achromatic imaging for this EDOF lens depends on how identical the MTFs are over the wavelength range. This can be quantified by the correlation function as a function of the wavelength. With inverse design, the bandwidth can be further increased, as shown by the correlation coefficient. a) [37]. Copyright 2018 , AAAS. b) [42]. Copyright 2021, De Gruyter.

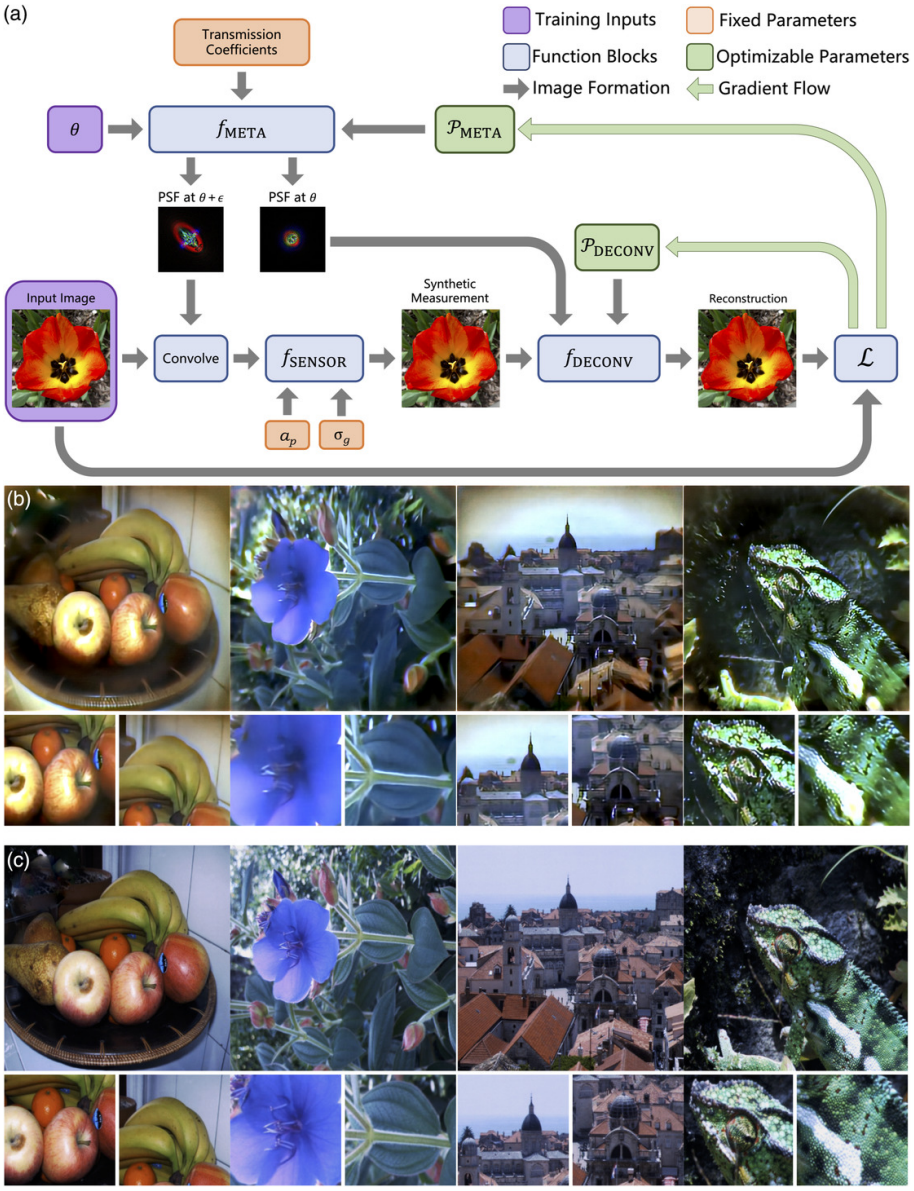

Figure 2.6: Co-optimized hardware-software for full-color imaging: a) Design framework to co-optimize the metaoptic and reconstruction algorithm via a differentiable pipeline. b) Images captured using a single 500μs, $f/2$ metaoptics. c) Same images captured using a six-element compound optic. [46]. Copyright 2021, Springer Nature Ltd.

Chapter 3

Extended Depth of Focus Meta-optics[1]

3.1 Background

As we briefly discussed in chapter **§ 2.4**, one effective technique for mitigating chromatic aberrations in meta-optics is to employ freeform metasurfaces and computational imaging [47] a paradigm that does not have the scaling challenges we discussed, although this process entails additional energy consumption and latency due to the need of computational reconstruction. With advancements in and the availability of fast and efficient computing, however, this can be done swiftly and with minimal energy consumption. Thus, computational imaging in conjunction with freeform metasurfaces is a promising avenue for mitigating metalens aberrations. As aforementioned in **§ 2.4.1**, full-color imaging in the visible wavelength regime was recently demonstrated using an extended depth of focus (EDOF) metasurface and post-processing deconvolution [37]. Here, a rectangularly separable cubic phase mask (CPM) [48] was added to the standard hyperboloidal metalens phase, generating a non-rotationally symmetric extended focal spot. The extended nature of the focal spots at different wavelengths is sufficient to compensate for the chromatic shift in the focal length. The EDOF property

[1]The content of this chapter is based on the following publication:
Luocheng Huang, James Whitehead, Shane Colburn, and Arka Majumdar, "Design and analysis of extended depth of focus metalenses for achromatic computational imaging," Photon. Res. 8, 1613-1623 (2020)

of the CPM enables the imaging system to capture useful spatial frequency information of the colored image so that computational reconstruction is possible [37]. The CPM is limited, however, in that it produces an asymmetric PSF that makes imaging sensitive to the orientation of the element, often manifesting as asymmetric artifacts even after deconvolution. Additionally, the CPM produces a lateral shift of the PSF with a change in wavelength, which can contribute to distortions in imaging. One potential solution to this limitation is to utilize a rotationally symmetric point spread function. Although there are previous works doing so based on extending the depth of focus of refractive optics, using log-asphere [49] and axicon-based lenses [50], they have not been used for correcting the strong chromatic aberrations encountered in a metasurface platform.

In this chapter, we extend the family of EDOF metasurfaces beyond a simple CPM. We design and fabricate four different types of EDOF metasurface lenses operating in the visible regime, including both rotationally symmetric and asymmetric phase profiles. We characterize the MTF for all of these lenses and demonstrate full-color imaging. A comparative analysis of all these lenses, is also provided, in terms of optical bandwidth and imaging performance. All our EDOF metasurfaces demonstrate at least an order of magnitude larger optical bandwidth. Full-color imaging in the visible range is achieved using all the EDOF lenses, outperforming the traditional metalens in terms of chromatic aberrations.

3.2 Design and Fabrication

An imaging system behaves as a linear system that maps the incoming light from a scene to the sensor. This mapping function of the optical element is modelled by a convolution with the element's point spread function (PSF). For imaging under incoherent light, the system can be considered to provide a linear mapping from the input intensity to the output intensity, which is captured on the sensor array. By scanning a point source throughout the object volume and measuring the resultant intensity across the image volume, a 3D intensity impulse response, or PSF, can be measured. This method can fully characterize the imaging function of the system. Assuming the PSF is shift-invariant, we can treat the PSF as a kernel

that convolves with the input to produce an image on the sensor plane. For an ideal lens, the PSF resembles a point, which enables capturing an exact replica of the scene. When the PSF deviates from a point, the captured image becomes blurry; however, with a PSF known a priori, an in-focus image can be retrieved via post-capture deconvolution if sufficient spatial frequency information is captured by the sensor. This information retrievability can be expressed in terms of the modulation transfer function (MTF) of the optical element, which is given by a slice of the magnitude of the Fourier transform of the PSF.

A broad MTF, one that does not drop to zero rapidly, signifies that a wide range of spatial frequencies is captured at the sensor plane, corresponding to PSF with a small spot size. On the other hand, a narrow MTF, one that decays rapidly, captures only a limited range of spatial frequency content, which precludes the possibility of computational reconstruction due to the zeros in the spatial frequency spectrum. A conventional, in-focus metalens exhibits a broad MTF when imaging with narrowband light, resulting in high-quality images. When imaging with a different wavelength at the same sensor plane, however, the spatial bandwidth of the MTF drastically decreases and the collected spatial frequencies from the scene are attenuated or eliminated. As some of the spatial frequencies are not collected, this results in an uncorrectable blur. With EDOF lenses, we can realize a similar PSF at the sensor plane for a broad and continuous range of wavelengths [51]. Moreover, the resulting MTF can capture a wider range of spatial frequencies, which is key to generating full-color images at high resolution.

Here, we designed four different EDOF lenses, namely EDOF cubic, shifted axicon, log-asphere, and SQUBIC lenses. All, except for the cubic, are axially symmetric. We fix the aperture for each of these metasurfaces at 200 μm and select a nominal focal length of 200 μm, making the numerical aperture (NA) close to 0.45 for all designs. We emphasize that the small aperture of the metalens is due to the prohibitive cost and time for fabricating large aperture lenses based on electron-beam lithography, and not due to scaling limitations as encountered with dispersion-engineered metalenses. The cubic metalens utilizes a focusing phase mask combined with a cubic term to produce a MTF insensitive to wavelength. [37]

Thus, the phase mask for an EDOF cubic lens is:

$$\phi(x,y) = \frac{2\pi}{\lambda}\left(\sqrt{x^2+y^2+f^2} - f\right) + \frac{\alpha}{R^3}\left(x^3+y^3\right) \tag{3.1}$$

where λ denotes the operational wavelength, x and y are the coordinates in plane, f is the nominal focal length, α represents the strength of the cubic term, and R represents the radius of the phase mask. We chose $\lambda = 550$ nm, $f = 200$ μm, $R = 100$ μm, and $\alpha = 55\pi$ for the design of the cubic metasurface. To compare the performance of the EDOF metasurfaces against the standard singlet metalens, we include a design with $\alpha = 0$ that imparts no cubic term to the wavefront. The log-asphere phase mask was inspired by a prior work [49] that divides the phase mask into annular zones with continuously varying focal lengths. The central annular zone has a focal length of s_1 and the outermost annular zone has a focal length of s_2. This design effectively extends the focal length from s_1 to s_2. The log-asphere phase mask is governed by the relation:

$$\phi\left(r\right) = \frac{2\pi}{\lambda}\int_0^r \frac{rdr}{\left\{r^2 + \left[s_1 + \left(s_2 - s_1\right)\left(\frac{r}{R}\right)^n\right]^2\right\}^{1/2}} \tag{3.2}$$

where $r = \sqrt{x^2+y^2}$ and R is the aperture radius of the phase mask. The parameter n changes the intensity distribution over the optical axis. For the log-asphere phase mask, we set $n = 2$, making the intensity distribution uniform across the line of foci. The parameters we chose for the log-asphere metasurface design are $\lambda = 550$ nm, $s_1 = 80$ μm, $s_2 = 300$ μm, and $R = 100$ μm. Similar to the log-asphere phase mask, the shifted axicon phase mask takes the same form of the Eq. (2). In this case, we set $n = 1$, $\lambda = 550$ nm, $s_1 = 80$ μm, $s_2 = 300$ μm. The axial intensity distribution of the shifted axicon resembles that of a diffractive axicon lens [52], hence the nomenclature. The SQUBIC metasurface was first demonstrated by Patwary to have EDOF properties [53]. The phase mask directs a collimated light beam into a line of discrete foci, achieving an EDOF, and is given by the equation:

$$\phi(x,y) = 2\pi A \left[\frac{\sqrt{1 - \left(\left(\frac{x}{R}\right)^2 + \left(\frac{y}{R}\right)^2 \right) \sin^2 \alpha} - 1}{1 - \cos\alpha} + \frac{1}{2} \right]^3 \tag{3.3}$$

where $\alpha = \sin^{-1}(NA)$, given that NA is the nominal numerical aperture of the optics. Here, A is a design parameter that determines the strength of the phase mask [53]. In our design, we set $A = 50$ and NA $= 0.45$. **Fig. 3.1**A shows the wrapped phase distributions of all the metasurfaces.

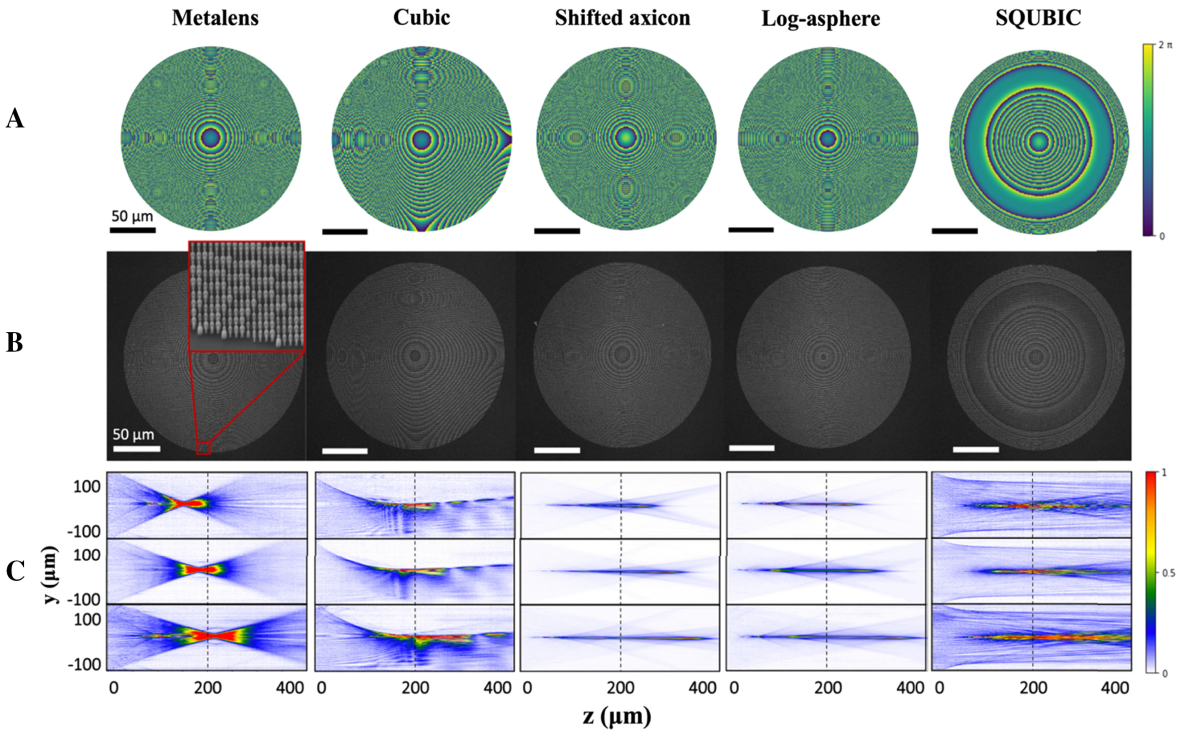

Figure 3.1: EDOF Metasurface Design and Measurements. (A) The phase masks of an ordinary metalens and four different EDOF metasurfaces. (B) Scanning electron micrographs of the fabricated metasurfaces. Inset shows the pillar distribution. (C) We experimentally measured the intensity along the optical axis where from top to the bottom panel represents illumination by 625, 530, and 455 nm wavelengths.

These EDOF lenses are then implemented using cylindrical Si_3N_4 nanopillars [54], [55] to ensure polarization insensitivity. These nanopillars are arranged on a square lattice. By varying the diameters of these nanoposts, the transmission coefficient imparted on incident light is modified as the coupling to and amongst different supported modes by the nanoposts changes, resulting in different phase shifts. Rigorous coupled-wave analysis (RCWA) is used to construct a library consisting of the diameters of the nanoposts and the corresponding phase shift and amplitude.

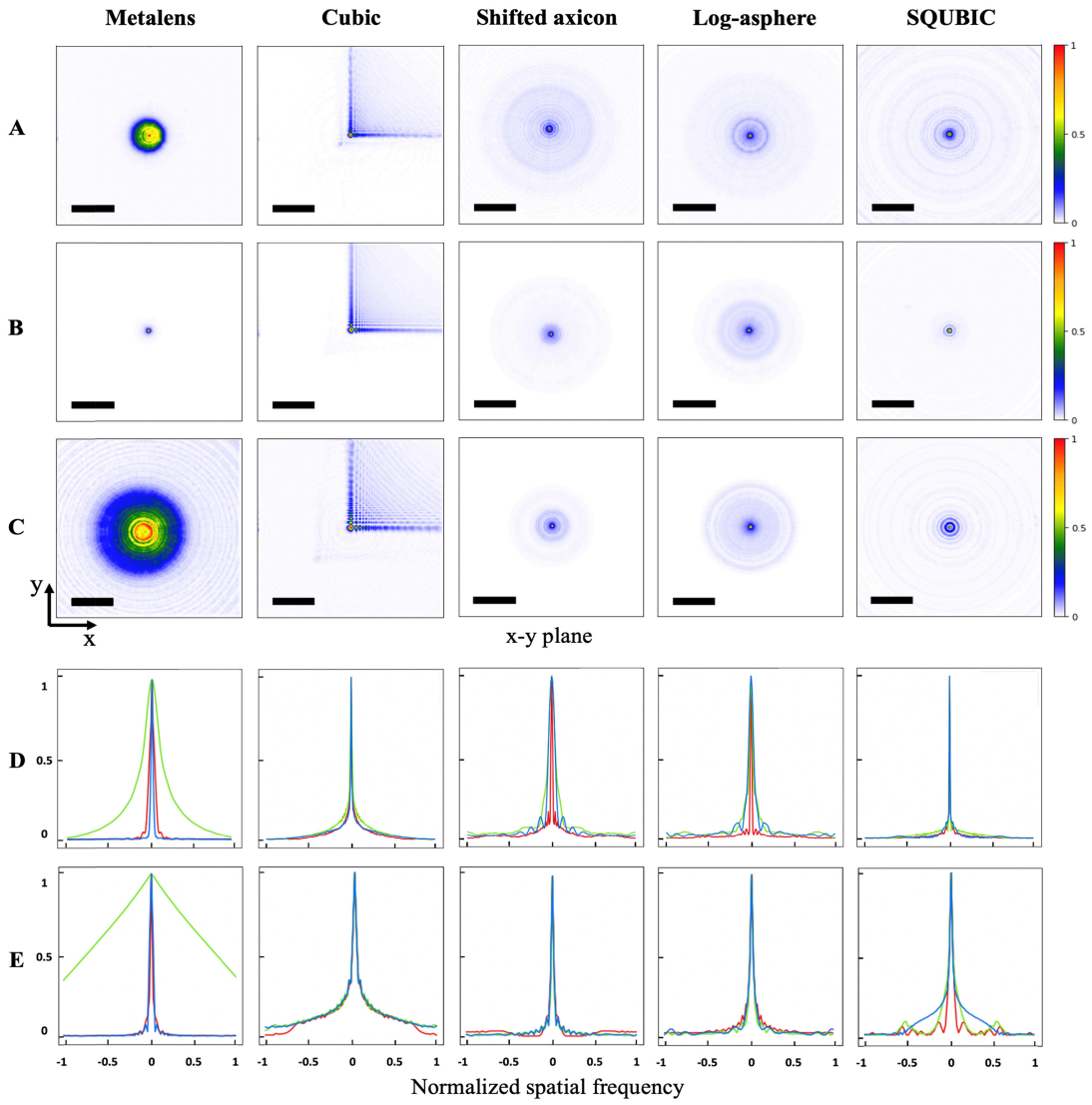

Figure 3.2: Characterization of the metasurfaces. The PSFs of the singlet metasurfaces were measured under 455 nm blue (A), 530 nm green (B), and 625 nm red (C). The corresponding experimental MTFs are displayed with red line from its PSF measured under red light, green lines under green light, and blue lines under blue light. (D) The theorical MTFs are displayed on row (E). The scale bar signifies 25 μm.

3.3 Fabrication, Characterization and Imaging

The designed metasurfaces are fabricated on the same sample. First, a double side polished fused silica wafer is cleaned with acetone and IPA. Plasma-enhanced chemical vapor deposition is used to deposit 623 nm of Silicon nitride (via SPTS). A layer of 200 nm of ZEP-520A is then spun on the wafer and 8 nm of Au/Pd is sputtered for charge dissipation. The design pattern is written using electron-beam lithography (JEOL JBX6300FS at 100 kV). Then the Au/Pd layer is removed by immersing in Gold Etchant Type TFA (Transene) and the chip is developed in amyl acetate. Next, 50 nm of aluminum is evaporated onto the developed pattern and then lifted off, leaving a patterned aluminum etch mask. The silicon nitride layer is etched through its full thickness in an inductively coupled plasma etcher using a fluorine chemistry (Oxford Plasmalab 100). Finally, the aluminum is removed, producing the metasurfaces. **Fig. 3.1**B shows the scanning electron micrograph (SEM) of all the metasurfaces.

The fabricated metasurfaces (one ordinary metalens and four EDOF metasurfaces) are then characterized experimentally. A fiber-coupled LED light source (Thorlabs M625F2, M530F2, M455F1) is used for illumination of the metasurfaces. A custom microscope mounted on a computer-controlled translation stage is used to take snapshots of the imaging plane. The focal length of the metalens is found to be 231 μm with the green LED. This deviation from the designed focal length (200 μm) could be due to the difference in the designed wavelength (550 nm) and the experimental light source wavelength (530 nm). The imperfect fabrication could have also contributed to this deviation. The nominal focal length is thereafter set to 231 μm for all five metasurfaces to accommodate this shift. To verify the extended depth of focus, PSF slices are measured along the optical axis (z-axis) spanning 0 to 400 μm from metasurface to construct a 3D PSF, shown in the **Fig. 3.1**C. The PSF measurement on the x-y plane is shown in **Fig. 3.1**, under different wavelength illumination. We note that for a standard metalens, the PSFs are very different for different colors, whereas, for the EDOF lenses the variation is significantly reduced. To quantitatively

understand this behavior, we calculate the MTF of all the metasurfaces at the nominal focal plane. As can be seen in **Fig. 3.2**, the MTF of an ordinary metalens well preserves spatial frequency information for green light (when focused at the sensor plane) but fails to preserve high frequency components for blue or red. On the other hand, the cubic metasurface presents a broad frequency preservation, exhibiting a higher cutoff frequency than the standard metalens. Similarly, the other EDOF metasurfaces demonstrate a broad frequency preservation with an oscillatory behavior: while there are some zeros, for several high spatial frequency components we have non-zero MTF over a wider range compared to the standard metalens. This also plays an important role, as when an image is captured under white light illumination, all the wavelengths are captured at the sensor array, and having a similar MTF enables computational reconstruction, as discussed below.

After characterization of the PSFs and MTFs, we captured colored images using all the metasurfaces. The same setup is used for imaging with an addition of an OLED monitor (SmallHD 5.5" Focus OLED HDMI Monitor). We have recorded the spectra of the OLED monitor as well as the fiber-coupled LEDs. The individual spectra are recorded when displaying a single color at a time using (IsoPlane SCT320). The OLED monitor is placed $\sim$15 cm away from the metasurface, which displays images shown on **Fig. 3.3**. The images are then captured using a camera (Allied Vision GT1930C). These raw images are then computationally reconstructed. Our EDOF imaging system presents a problem in the form of $O = Kx + n$, in which O is the observed image, K is the PSF or blur kernel, x is the latent image, and n is noise that corrupts the captured data. Although several methods exist to restore the image x, given K and O, we chose to use Wiener deconvolution due to its effectiveness and simplicity. To restore the captured raw images given the point spread function (PSF), a 3×3 median filter is first applied to the raw image to remove hot pixels. Then, Wiener deconvolution is applied to the image. We use the Wiener-Hunt deconvolution algorithm [56], which takes in the raw image and the PSF as input. Before it is passed into the algorithm, the PSF is normalized by dividing by the sum of all PSF pixels. The actual implementation uses the *skimage.restoration.wiener* module from the *Scikit-image* library

[57]. The Wiener filter estimates the desired image $\hat{x}$ as follows:

$$\hat{x} = \mathcal{F}^{-1}\left(|\Lambda_H|^2 + \lambda\,|\Lambda_D|^2\right)\Lambda_H \mathcal{F} y,$$

where Λ_H is the optical transfer function, Λ_D is a Laplacian filter that penalizes high frequencies, and λ is the parameter that tunes the regularization. The λ used for the deconvolutions is 1×10^{-4}. We note that, the normal metalens exhibits such a high level of chromatic aberration that we were unable to restore the raw images with any reasonably good results.

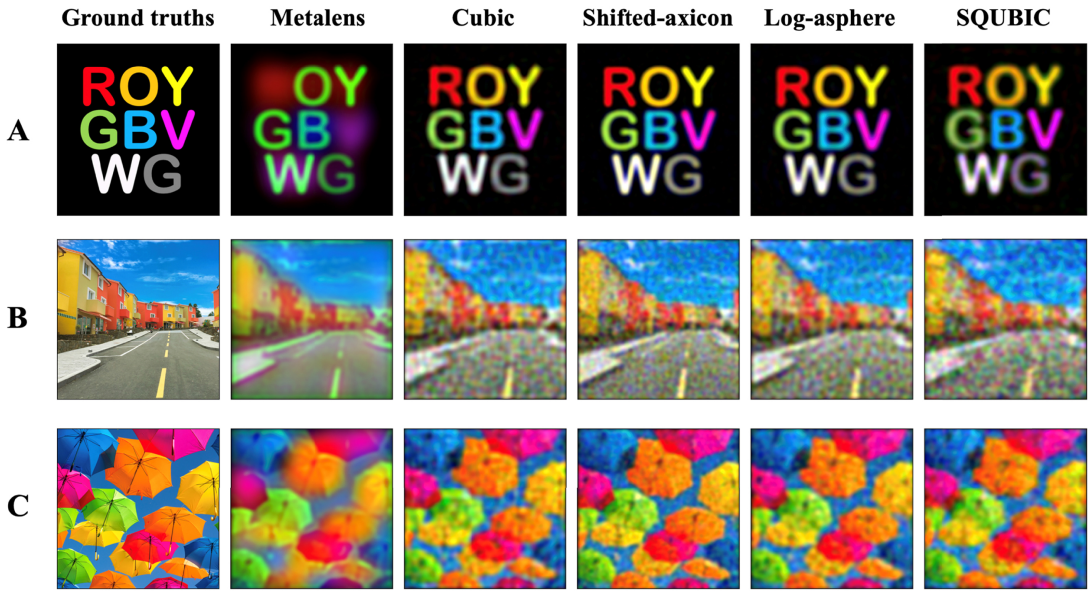

Figure 3.3: Simulated imaging performance after deconvolution. Deconvolved images captured by the EDOF imaging system, using the simulated images and PSFs. The experimental counterpart can be found in **Fig. 3.4**.

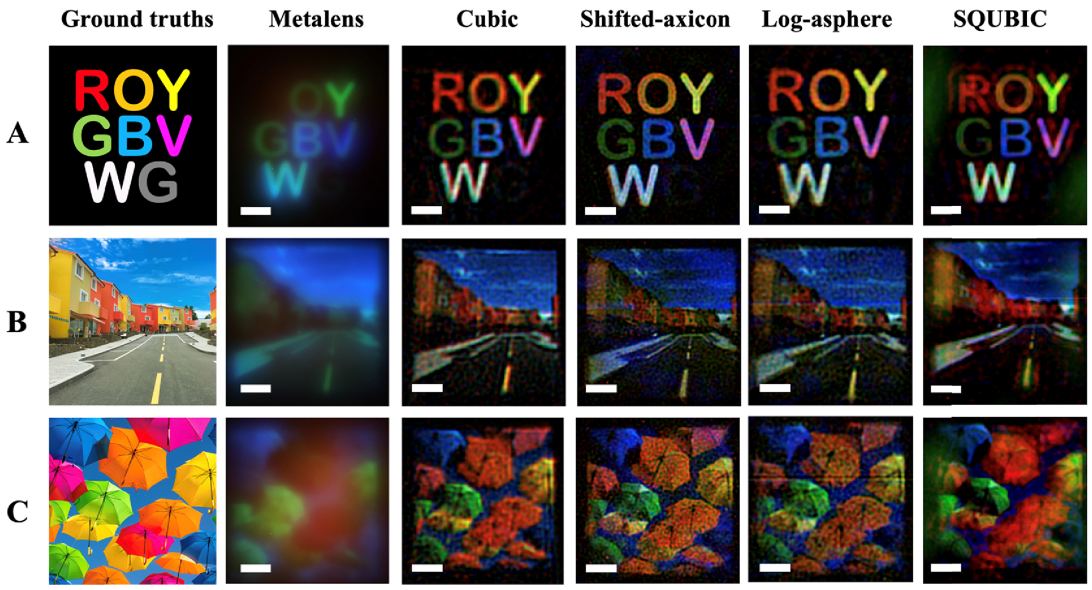

Figure 3.4: Imaging performance. Restored images taken from (A) an OLED display of colored letters in ROYGBVWG, (B) a colorful neighborhood, and (C) vibrant umbrellas against the sky. The scale bar signifies 20 µm. Note that the metalens images are raw and unrestored.

To obtain an objective imaging quality comparison across the five metasurfaces, structural similarity (SSIM) tests are carried out using the “ROYGBVWG" image shown on **Fig. 3.3**(A). SSIM calculations are sensitive to the subject's translation, scaling, and rotation, which are difficult to eliminate in an experimental setup. Hence, an intensity-based image registration method in MATLAB is utilized to transform the captured image to align with the ground truth using a similarity transformation. [58] This operation is performed for all three color-channels simultaneously for each capture. The SSIMs are then calculated between the ground truth (GT) and the output images for the red, green, and blue channel respectively. The values of the SSIM range from 0 to 1, 1 being a perfect match. As shown in **Fig. 3.4**, the imaging system with the singlet metalens exhibits the lowest SSIM score compared to the rest of the metasurfaces. Even though the SSIM score is at most 0.55, the scores on the devices with EDOF properties are significantly higher than that of the stan-

dard metalens. The metalens obtained poor imaging quality throughout all colors, although designed to work for green. The low SSIM of the metalens for green is due to the fact that the OLED display used has a nominal wavelength (511 nm) different from that of the light source (531 nm) used to measure the PSF for all the colors. The OLED monitor also has a much larger bandwidth (∼35 nm for green) compared to the simulated bandwidth for the metalens, leading to a blurry image even for the intended wavelength.

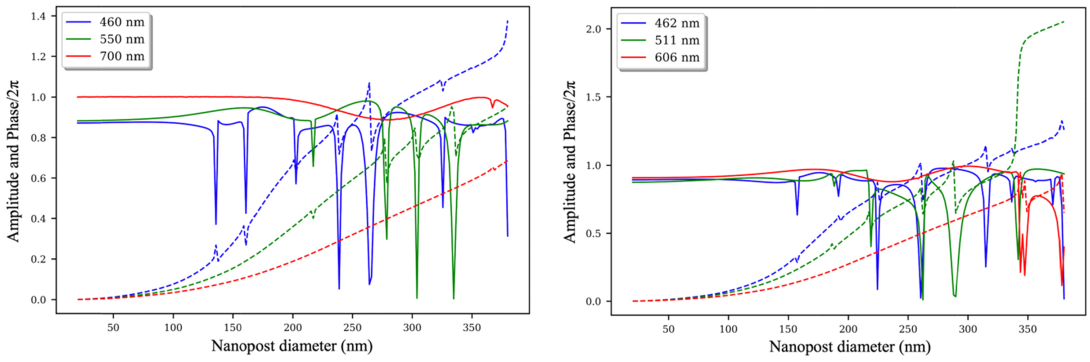

Figure 3.5: Phase (dashed lines) and amplitude (solid lines) response of the nanopillars, simulated using RCWA.

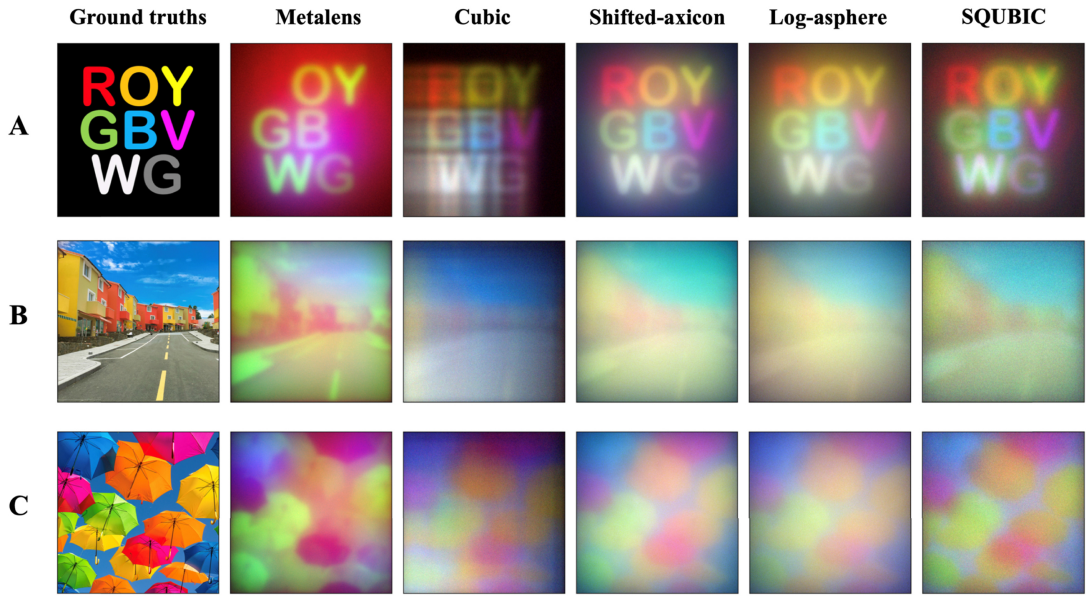

Figure 3.6: Simulated captured images before deconvolution. The experimental counterpart is shown in **Fig. 3.8**

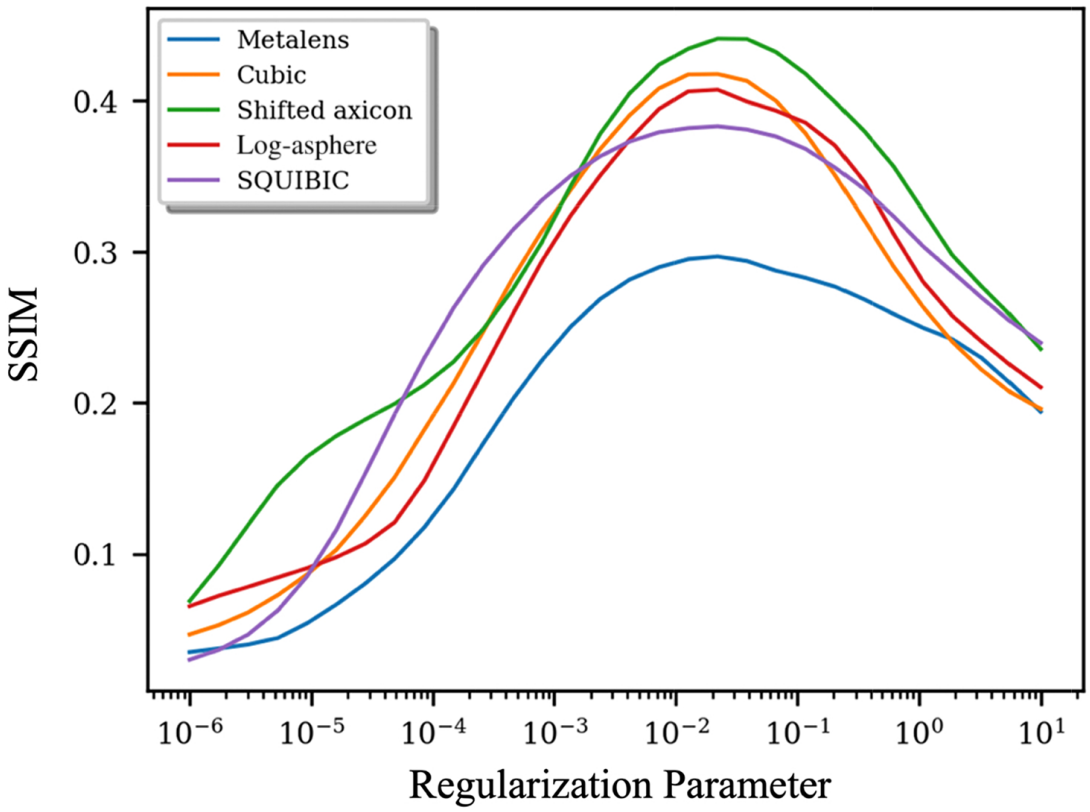

Figure 3.7: SSIM values when different regularization parameters are utilized.

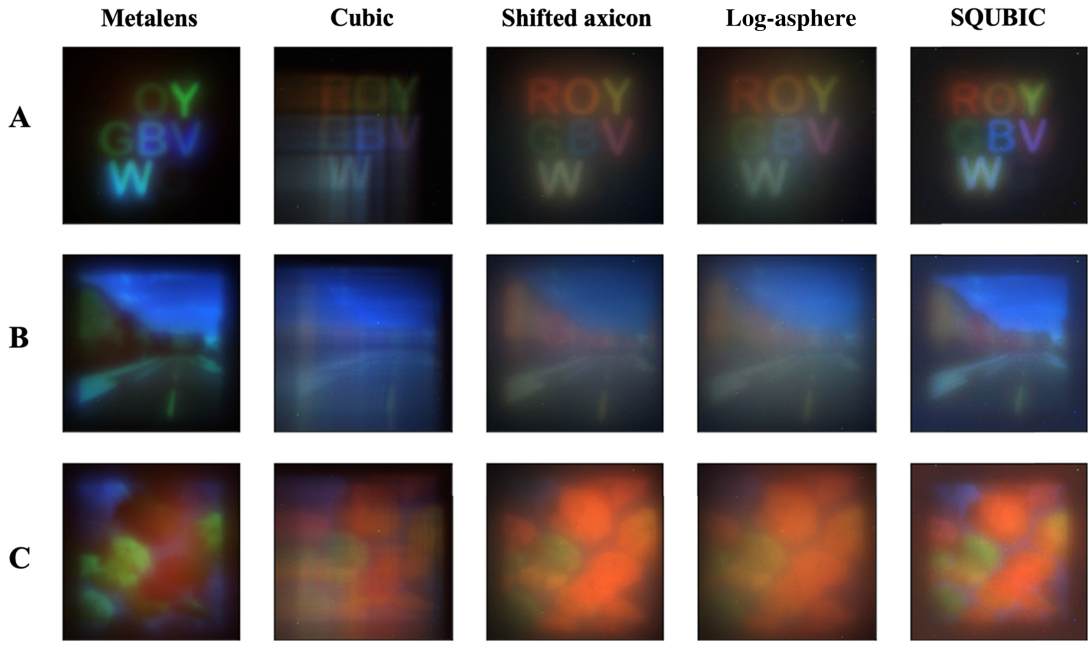

Figure 3.8: Raw images taken from an OLED display of colored letters in (A) ROYGBVWG, (B) a colorful neighborhood, and (C) vibrant umbrellas against the sky.

Table 3.1: Table Imaging bandwidth, defined as the bandwidth at one half of the maximum PSF similarity coefficient.

	Metalens	Log-asphere	Shifted Axicon	Cubic	SQUBIC
Bandwidth (nm)	15.2	233.3	233.3	112.1	157.6
Center Wavelength (nm)	543.9	553.0	547.0	540.9	530.3

As discussed before, it is desirable to have an invariant PSF across the full visible wavelength range to achieve full-color imaging. To capture the color invariance of the PSFs on the metasurfaces, we perform PSF correlation calculations with simulated data that estimate the spectral bandwidth. First, a wavelength-sensitive phase mask is generated consisting of transmission coefficients obtained through rigorous coupled-wave analysis, using the Stanford

S4 package [59]. Then, the wavefront is propagated to the image plane using the angular spectrum method. The correlation coefficient is calculated as the inner product between the baseline PSF and the test PSF at the test wavelength, normalized to the maximum coefficient. Table 3.1 shows the bandwidth (measured as the optical bandwidth where the correlation values falls to half of the maximum [60]) and the central wavelength of all the metasurfaces. Compared with the EDOF metasurfaces, the standard metalens has a much narrower spectral bandwidth, as is also evident from the MTF and SSIM measurements.

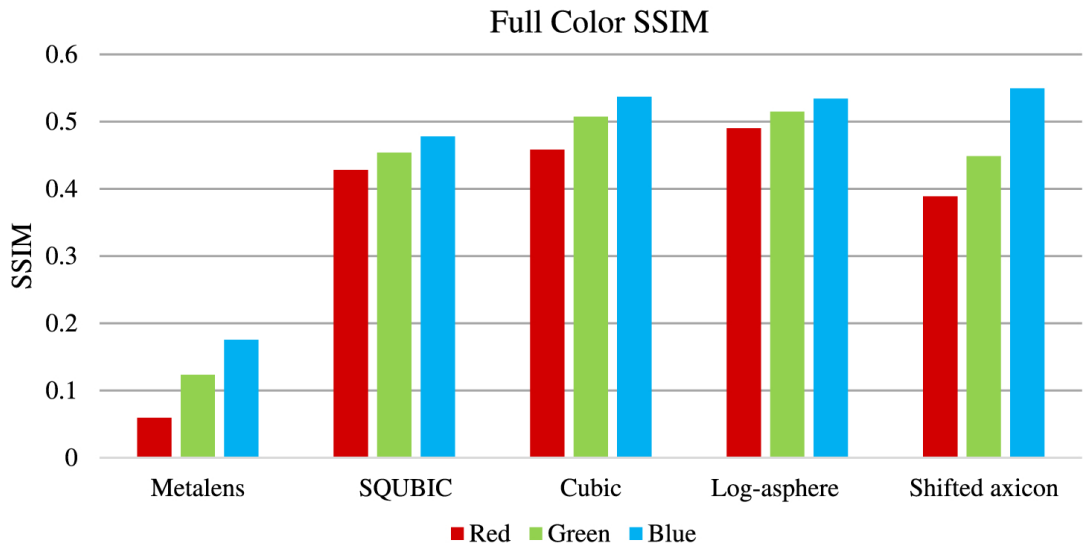

Figure 3.9: Full color SSIM. The restored captures are scaled, rotated, and translated to align with the ground truth; then SSIM is calculated for each color channel for the metasurface.

3.4 Discussion

The SSIM results agree with the poor imaging quality exhibited by the metalens in the visible regime, caused by its strong chromatic aberrations. In contrast, the EDOF metasurfaces demonstrate an impressive ability to maintain a highly invariant PSF across a large spectral bandwidth that makes them an excellent platform for computational imaging. The imaging results as well as the SSIM calculations indicate that EDOF metasurfaces significantly

outperform the standard metalens in full-color imaging. The log-asphere and shifted axicon designs both demonstrate the highest optical bandwidth for imaging among the tested EDOF designs. We note that recently a fundamental limit on the achievable optical bandwidth is reported given a thickness and numerical aperture of the lens [20]. For our parameters the fundamental limit becomes around ∼100nm (see appendix), which is smaller than all the bandwidth demonstrated using EDOF lenses. While it is indeed possible to increase the optical bandwidth at the expense of image quality, it is difficult to estimate the Strehl ratio for our EDOF lenses as the PSF is very different from a lens and the final image is obtained only after deconvolution.

Our image restoration process takes ∼2 seconds to restore a single image on a single CPU thread. This process can, however, be significantly accelerated by utilizing multiple CPUs, GPUs, or a dedicated field-programmable gate array. Preloading the PSF into memory will also speed up the computational reconstruction process. This work uses Wiener deconvolution due to its speed, however, more sophisticated deconvolution algorithms can be used to improve the image quality at the expense of time [61, 62], which could be applied in post-processing.

This work demonstrates the viability of four different metasurface designs for full-color imaging with minimal chromatic aberrations. We design and fabricate all of these EDOF metasurfaces, which when combined with deconvolution, correct for chromatic aberrations caused by the inherent diffractive nature of the metasurface. Imaging and analyses are carried out on our devices that compare the performance between the traditional metalens and the four different EDOF metasurfaces. Although previous works have demonstrated achromatic imaging with dispersion-engineered metasurfaces, our imaging system does not rely on any input polarization state and is generalizable to larger aperture applications. A larger aperture will enable higher signal-to-noise-ratio and faster shutter speed, which are crucial to any commercial application. Moreover, dispersion-engineered metasurfaces require high aspect ratio scatterers for large apertures and also have multiple meta-atoms per unit cell, making the fabrication more challenging. In our approach the aspect ratio of

the scatterers is relatively small, and thus our metasurfaces have more relaxed fabrication constraints.

Compared with the singlet metalens, our EDOF imaging system exhibits an exceptional ability to maintain imaging fidelity under broadband illumination. Our imaging platform combines the form factor of ultrathin diffractive metasurfaces and the flexibility of computational imaging, making our platform an attractive solution for novel imaging applications. The CMOS compatibility of our silicon nitride platform, combined with a high NA of ∼0.45, makes this approach ideal for miniaturized microscopy, smartphone cameras, and endoscopy. Moreover, it is possible to increase the aperture of our EDOF hybrid system while maintaining the same NA and imaging characteristics. This aperture scalability opens avenues to applications such as planar cameras and even satellite imaging.

3.5 Project Details

3.5.1 Theoretical Bandwidth Calculations

The theoretical bandwidth of each lens is calculated using Eq. (A1) and is listed in Table 3.2 [20]:

$$\Delta\omega \leq \omega_c \frac{L\Delta n}{f} \frac{\sqrt{1-\left(\frac{\mathrm{NA}}{n_b}\right)^2}}{1-\sqrt{1-\left(\frac{\mathrm{NA}}{n_b}\right)^2}},$$

where ω_c is the central frequency, L is the pillar height, f is the focal length, Δn is the refractive index difference between the background and the pillars, n_b is the background refractive index, and NA is the numerical aperture.

3.5.2 Transmission Coefficient of the Nanopillars

RCWA simulation is carried out on the nanopillars using the $S4$ library [61]. The periodicity of the unit cell is 0.4μm. The pillar height is 0.625μm. The metasurfaces are designed with the phase response of 550 nm wavelength.

Another RCWA simulation is carried out for the experimental phase masks. The metasurfaces are designed with the phase response of 606, 511, and 462 nm wavelengths corresponding to the red, green, and blue color channels on the OLED monitor.

3.5.3 Fabrication Process

The designed metasurfaces are fabricated on the same sample. First, a double side polished fused silica wafer is cleaned with acetone and isopropyl alcohol. Plasma-enhanced chemical vapor deposition is used to deposit 623 nm of silicon nitride (via SPTS). A layer of 200 nm of ZEP-520A is then spun on the wafer and 8 nm of Au/Pd is sputtered for charge dissipation. The design pattern is written using electron-beam lithography (JEOL JBX6300FS at 100kV). Then the Au/Pd layer is removed by immersing in gold etchant type TFA (Transene) and the chip is developed in amyl acetate. Next, 50 nm of aluminum is evaporated onto the developed pattern and then lifted off, leaving a patterned aluminum etch mask. The silicon nitride layer is etched through its full thickness in an inductively coupled plasma etcher using a fluorine chemistry (Oxford Plasmalab 100). Finally, the aluminum is removed, producing the metasurfaces. **Fig. 3.1**(B) shows the scanning electron micrograph of all the metasurfaces.

3.5.4 Image Capture Settings

The exposure durations are listed in Tables 3.3 and 3.5 for capturing the PSFs and imaging, respectively. No gain is used for capturing the PSFs, and 10 dB of gain is used for imaging. The captured image goes through demosaicing by Allied Vision capture software and is saved to a TIFF file with 12 bit per color channel. The camera gain is set to 0 dB.

3.5.5 Simulated Captured Images

The simulated images shown in **Fig. 3.6** are generated by convolving the experimental PSFs with the ground truth. Gaussian white noise with a standard deviation of 0.25 times the

Table 3.2: Theoretical Bandwidths of EDOF Metasurfaces

	Metalens	Cubic	Shifted Axicon	Log-Asphere	SQUBIC
Bandwidth (μm)	104.4	106.2	105.0	103.8	101.8

Table 3.3: Exposure Durations for PSF (in Seconds)

	Metalens	Cubic	Shifted Axicon	Log-Asphere	SQUBIC
Blue	0.006	0.016	0.009	0.009	0.011
Green	0.003	0.012	0.010	0.010	0.005
Red	0.050	0.025	0.014	0.014	0.009

Table 3.4: Exposure Durations for PSF (in Seconds)

	Metalens	Cubic	Shifted Axicon	Log-Asphere	SQUBIC
Exposure	7.0	7.0	15.0	15.0	7.0

average pixel value for each channel is then added to emulate the experimental measurement such that the SNR of the convolved image is 6.02 dB.

3.5.6 Simulated Restored Images

The SSIM and PSNR are calculated from the simulated restoration and ground truth in **Fig. 3.3**(A). The SSIM and PSNR results are listed in Tables 3.5 and 3.6. The deconvolution

Table 3.5: SSIM on Simulation-Restored Images

	Metalens	Cubic	Shifted Axicon	Log-Asphere	SQUBIC
Blue	0.169	0.440	0.445	0.392	0.502
Green	0.464	0.334	0.361	0.340	0.338
Red	0.115	0.259	0.292	0.261	0.131

Table 3.6: PSNR on Simulation-Restored Images (in dB)

	Metalens	Cubic	Shifted Axicon	Log-Asphere	SQUBIC
Blue	10.7	19.7	19.6	17.7	20.9
Green	21.3	18.9	18.5	17.4	18.1
Red	11.8	18.4	18.3	17.1	13.7

Table 3.7: SSIM on Simulation-Restored Images (Gray-Scale Ground Truth)

	Metalens	Cubic	Shifted Axicon	Log-Asphere	SQUBIC
Blue	0.101	0.317	0.322	0.312	0.385
Green	0.445	0.324	0.336	0.324	0.315
Red	0.137	0.317	0.359	0.314	0.197

Fig. 3.7 SSIM values when different regularization parameters are utilized. implementation for both the simulated and experimental images uses the skimage.restoration.wiener module from the Scikit-image library [62].

To simulate the performance of the metasurfaces on gray-scale images, we take the same ground truth used in **Fig. 3.3**(A) and convert it into a gray-scale image using the convention Gray $\leftarrow 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B$, where Gray is the output gray-scale pixel value; R, G, B are red, green, and blue channel pixel values, respectively. The simulation procedure is the same as conducted for Table 3.5. The SSIM values from the grayscale simulation-restored images are presented in Table 3.7.

3.5.7 Wiener Deconvolution Parameter Search

Fig. 3.7 shows the SSIM plotted against the regularization parameter (L) to find the best L.

3.5.8 Captured Raw Images

The captured raw images are presented in **Fig. 3.8**.

3.5.9 Diffraction Efficiency

We define the diffraction efficiency as the power ratio of the light at the focal plane to the light transmitted through the metalens. To calculate the efficiency, we first capture an image at the focal plane. Then, using the same setting, we capture an image at the plane of the metasurface. We then define a circular mask where the metasurface spans and sums all the pixels in this region to determine the total transmitted power. Finally, the ratio of the focal plane power and the transmitted power yields the focusing efficiency. The measured efficiencies are presented in Table 3.1. We have defined this efficiency in such a fashion as several of the EDOF lenses do not produce a nice focal spot by design. We note that, due to pixel noise, and overall nonlinear behavior of the pixel counts with integration time,

Table 3.8: Diffraction Efficiencies of the Metasurfaces

	Metalens	Cubic	Shifted Axicon	Log-Asphere	SQUBIC
Blue	80.1%	57.7%	38.2%	44.1%	24.9%
Green	84.0%	58.5%	34.9%	42.0%	18.4%
Red	75.6%	74.2%	38.2%	40.0%	38.4%

Table 3.9: Peak Signal-to-Noise Ratio (PSNR) in dB

	Metalens	Cubic	Shifted Axicon	Log-Asphere	SQUBIC
Mlue	9.7	15.7	15.4	15.9	16.5
Green	12.8	14.1	14.5	14.5	13.5
Red	13.1	13.1	13.9	13.8	13.1

the measured values are very noisy and have large error margins (±10%). As such, larger aperture of the metasurfaces will enable measuring more reliable efficiency numbers. We, however emphasize that the efficiency is overall not a good metric to assess the performance of the metasurfaces presented here, and one should primarily focus on PSNR and SSIM (see Table 3.8).

3.5.10 Peak Signal-to-Noise Ratio of Experimental Imaging

The PSNR values are calculated for the experimental imaging, between the restored images and the ground truths (see Table 3.9).

3.5.11 Methods for Calculating Bandwidths

To capture the color invariance of the PSFs on the metasurfaces, we perform PSF correlation calculations with simulated data that estimate the spectral bandwidth. First, a wavelength-sensitive phase mask is generated consisting of transmission coefficients obtained through rigorous coupled-wave analysis, using the Stanford S4 package [59]. Then, the wavefront is propagated to the image plane using the angular spectrum method. The correlation coefficient is calculated as the inner product between the baseline PSF and the test PSF at the test wavelength.

Chapter 4

Forward Designed LWIR Meta-optics[1]

4.1 Long-Wavelength Infrared Metalens

Long-wavelength infrared (LWIR) imaging is a critical technology to defense and national security, condition monitoring, medical imaging, and non-contact thermography [63, 64, 65]. Many of these applications, including wearables and autonomous drone surveillance, impose stringent criteria on size, weight and power (SWaP), and it is imperative to find solutions to create miniature LWIR imaging devices. However, it is difficult to miniaturize LWIR optics using the traditional refractive optics due to their volumetric nature. A refractive optical element relies on the precise surface curvature of the dielectric to bend the incident light and requires a large total thickness and added weight to accommodate high numerical aperture (NA) applications. One way to reduce the size of the LWIR optics is to use diffractive optical elements (DOE) such as a multilevel diffractive lens (MDL) [66]. Rather than introducing a continuous phase shift via varying the dielectric material thickness, MDL wraps the phase and only uses just enough Sagittal depth to cover the 0-2π phase shift. Though an effective way to reduce the overall thickness of the optic element, MDLs have several drawbacks

[1]The content of this chapter is based on the following publication:
Luocheng Huang, Zachary Coppens, Kent Hallman, Zheyi Han, Karl F. Böhringer, Neset Akozbek, Ashok Raman, and Arka Majumdar, "Long wavelength infrared imaging under ambient thermal radiation via an all-silicon metalens," Opt. Mater. Express 11, 2907-2914 (2021)

in the complexity of the multi-layer fabrication. Additionally, MDLs generally have large periodicity resulting in higher order diffraction.

As we have shown previously in chapter **§ 2** and **§ 3**, meta-optics have shown promising capabilities for full-color imaging in the visible regime. Similar endeavour in LWIR wavelength range using meta-optics is lacking. Recently, imaging under laser illumination has been demonstrated in the LWIR regime [67]. Broadband imaging under ambient light has also been recently demonstrated using MDL [68]. Here we test the broadband imaging capability of all-silicon LWIR metalens. We design, fabricate, and characterize our meta-optics with an NA of 0.45 and a diameter of 2 cm. The performance of the metalens is tested in mixed temperature environments (to emulate different wavelengths) such as ambient, hot, and cold conditions yet the proposed system has maintained good imaging quality throughout. As the detectors in the LWIR range do not have any color filters, and we do not often require aesthetically pleasing images, even a simple, chromatic metalens can provide acceptable imaging under ambient light.

4.2 Methods

4.2.1 Design of LWIR Metalens

To simplify the fabrication process, we choose silicon as the material for the meta-optics. Specifically, our meta-atoms are square silicon posts created by partially etching a silicon wafer. We note that, while silicon is slightly absorptive in the LWIR range, we still expect $\sim$ 80% of the light to be transmitted. The metalens is designed by placing appropriate scatterers in a square lattice. The schematic of a unit cell of the nanopost is shown on Fig. 4.1 (a). By adjusting the widths of these square nanoposts, we can modify the effective index of the nanoposts and thus change the transmission coefficient (both amplitude and phase) of the incident light. We perform rigorous coupled-wave analysis (RCWA) to map the nanopost width to the respective transmission coefficient [59], shown on Fig. 4.1 (b). To account for the variable etch-depth, we calculate the amplitude and phase responses of

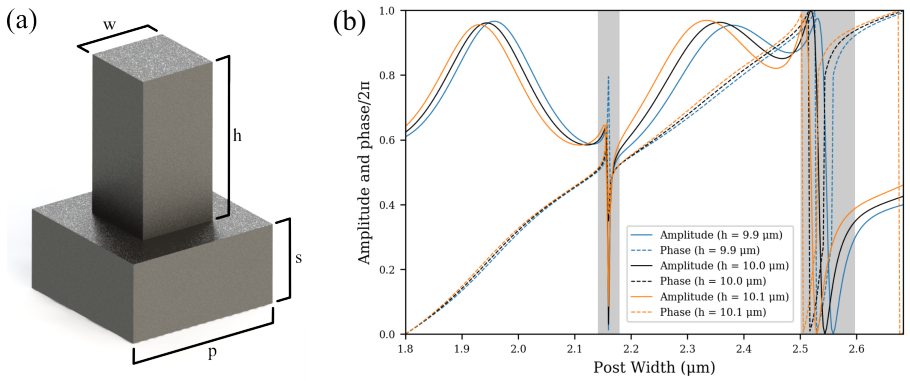

Figure 4.1: (a) The geometry of the individual silicon meta-atom is shown on a silicon substrate. The height h of the meta-atom is 10 µm. The substrate thickness s is 300 µm, the periodicity of the meta-atom p is 4 µm. (b) The transmission coefficient as a function of the width w of the meta-atom, for three different thicknesses. We exclude the post widths in gray regions on the plot to ensure acceptable level of transmission. This relation is used to map between the geometry of the meta-atoms and phase distribution of the metasurface.

posts with heights deviating by $\pm 100nm$ from the desired thickness. We found there is little change in the phase-values, making the scatterers robust to fabrication imperfections. Note that the resonances occur at some geometries where the amplitude takes a large dip, shown on the gray areas on 4.1 (b). These resonances are avoided during the scatterer assignment to create the meta-optics. We assign the metalens a standard hyperboloidal phase function:

$$\phi(x, y) = \frac{2\pi}{\lambda}(\sqrt{x^2 + y^2 + f^2} - f), \tag{4.1}$$

where λ is the nominal wavelength set to 10 µm, and the nominal focal length f is set to 2 cm; x and y indicate the coordinates on the metalens plane. A circular aperture stop made of aluminum is added to the metalens to block stray light.

4.2.2 Fabrication

We start with a double side polished Si wafer with a thickness of 300 µm and p-type doping with boron giving a sheet resistivity of 1 Ω·cm to 100 Ω·cm. We first fabricate the metal aperture stop. A negative photoresist is spin-coated onto the substrate before exposed to laser direct write lithography using Heidelberg DWL 66^+ and development. Next, 193 nm of aluminum is evaporated via E-beam evaporator CHA SEC-600, following lift-off, creating the aluminum aperture stop. Then, a positive photoresist is spin-coated on the substrate. We then again use laser direct writing with Heidelberg DWL 66^+ to create the metalens pattern inside the previously defined aperture. We develop the pattern and deep reactive-ion etch ∼10 µm the silicon using SPTS Rapier. We note that the pillar height is controlled by the etch time, which is calibrated beforehand to approximately achieve the desired thickness of the silicon scatterer. The residual photoresist is subsequently stripped by oxygen plasma in YES CV200 RFS Downstream Asher. The fabricated four metasurfaces can be seen on Fig. 4.2 (a). Scanning electron microscopy images are taken to reveal the top view (Fig. 4.2 (b)) and a high-magnification oblique view of the metalens (Fig. 4.2 (c)).

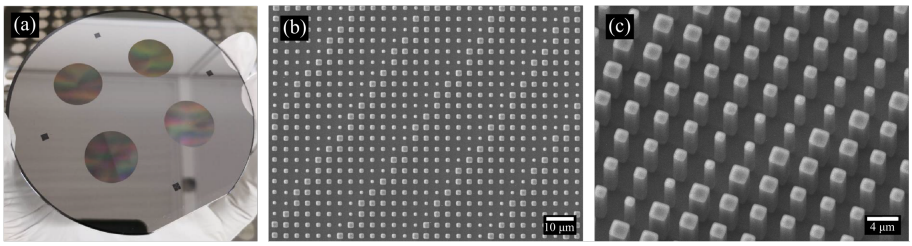

Figure 4.2: Fabricated devices: (a) Image taken with a cellphone camera, (b) scanning electron microscope (SEM) images of the fabricated structures, and (c) a zoomed-in view of the device.

4.3 Results

The resolution and contrast of an imaging system can be estimated by measuring the modulation transfer function (MTF). The MTF signifies how the imaging system preserves the spatial frequency from the scene, and is defined as the modulus of the optical transfer function (OTF), which is the Fourier transform of the point-spread function (PSF). We simulate the PSF by propagating the pupil function of the metalens using the fast Hankel transform method for a single wavelength [69]. We note that as our metalens has a very large aperture ($\sim$2 cm), we cannot directly use angular spectrum method to propagate the wavefront [41]. The broadband OTF is the average of OTFs sampled across the sensor spectral range and weighted by the spectral radiance of the heated body at each particular wavelength. The MTFs of our imaging system under different illumination conditions but at the same sensor plane are displayed on Fig. 4.3 (a). The MTF under narrowband illumination of 10 µm shows a near-diffraction-limited performance with a spatial cutoff frequency of $1/\lambda F$, which is strongly contrasted to that of the 12 µm source at the same sensor plane whose cutoff frequency is nearly zero. On the other hand, the broadband sources under hot water, soldering iron, and ice temperature all perform similarly at around 80% of the diffraction-limited cutoff frequency. The spectral radiance is calculated following Planck's Law [70]:

$$B_\nu(\nu, T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} \tag{4.2}$$

where B_ν is the spectral radiance, h is the Plank constant, c is the speed of light, k is the Boltzmann constant, ν is the wave frequency, and T is the temperature of the object. The normalized spectral radiance for different heated objects are displayed on Fig. 4.3 (b). Thus our metalens is strongly chromatic, but due to the weighted averaging, broadband MTFs preserve much more spatial frequency components than a narrowband source, that is different from the design wavelength. This observation is not necessarily new, but should be emphasized as we can still capture images from a scene under broadband illumination using chromatic optics. Especially without having any spectral discrimination, either in the source as a transmission filter or in the detector as a color-filter, an image under broadband

illumination can be captured.

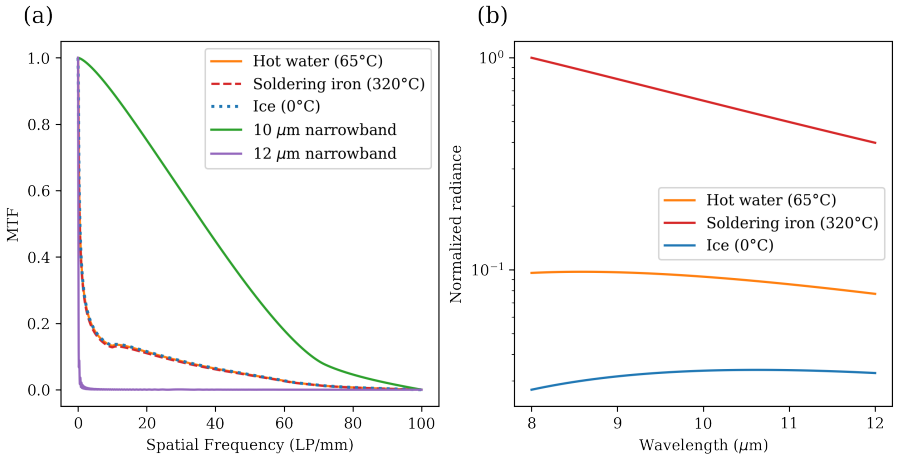

Figure 4.3: (a) The simulated MTFs of the metalens under different input sources. (b) The spectra of these input sources.

The simple metalens is tested using a minimal setup shown in Fig. 4.4 (a). On the left we have a LWIR camera mounted to an optical breadboard. The wafer is placed in front of the focal plane array at the appropriate focal distance ($f = 2cm$). We note that due to the low light from a pinhole and non-uniform noise in the LWIR detector, the measured PSF is noisy and as such does not provide much information. Hence, we characterize the metalens directly for imaging. The object distance x_o and the image distance x_i are determined by the thin lens equation $1/x_o + 1/x_i = 1/f$, where $f = 2\,\text{cm}$. First, we compare LWIR imaging capabilities in the ambient environment using the isothermal background broadband imaging setup shown on the right side of Fig. 4.4 (a). The black-body radiation source consists of a container full of water controlled at 65 °C using a thermal immersion circulator. Such container is covered with electrical tape to provide an emissivity of 0.95. A steel plate with precisely machined hole diameters is placed in front of the isothermal black-body source, creating contrasting circular cutouts (Fig. 4.4 (b)). The steel plate is placed 24 mm away from the metalens.

As shown on Fig. 4.4 (c), the broadband imaging result using the simple metalens already creates good contrasts for the circular cutouts. To compare the performance of narrowband imaging, a bandpass filter is applied at 10 µm and 12 µm in front of the metalens, shown on Fig. 4.4 (d) and (e), respectively. The imaging result at the designed nominal wavelength 10 µm shows better spatial resolution than broadband as it displays more recognizable cutouts of the metal gauge, however displaying with less contrast between the black-body background and the holey plate. This lower contrast is due to the limited number of photons in the filtered light. The imaging result at 12 µm conveys no recognizable shapes at the nominal imaging distance, however, when the metalens is moved 3 mm closer to the sensor, a clear image is formed with an imaging quality comparable to that of the 10 µm. This is expected in a metalens, where the focal length is inversely proportional to the wavelength.

The black-body radiation spectrum differs at different temperatures. To further characterize the performance of the metalens imaging system under more extreme temperatures, we include imaging of a hot soldering iron whose spectrum is shown in Fig. 4.3 (b). The soldering iron is heated to an estimated temperature of $\sim 320^{\circ}$C and is mounted on an optical post in front of a steel plate covered with electrical tape with 0.95 emissivity in a garage setting seen on Fig. 4.5 (a). It is imaged at 3.7 m away from the metalens. The imaging result on Fig. 4.5 (d) clearly shows the hot soldering iron contrasted by the backplate, along with the sawhorse and the panels of the garage door. Next, the same soldering iron is imaged but at a closer distance of 356 mm seen on Fig. 4.5 (b). At a closer distance, the border of the soldering iron is legible and the hot soldering iron is seen heating up a portion of the optical post assembly. To stretch the dynamic range, we include a person holding a cold ice pack and a burning blowtorch standing 3.7 m away from the metalens (Fig. 4.5 (c)) in a garage setting. The imaging result on Fig. 4.5 (f) clearly shows the contour of the person holding a cold item on the left with a clear border whilst the flame of the blowtorch is also visible.

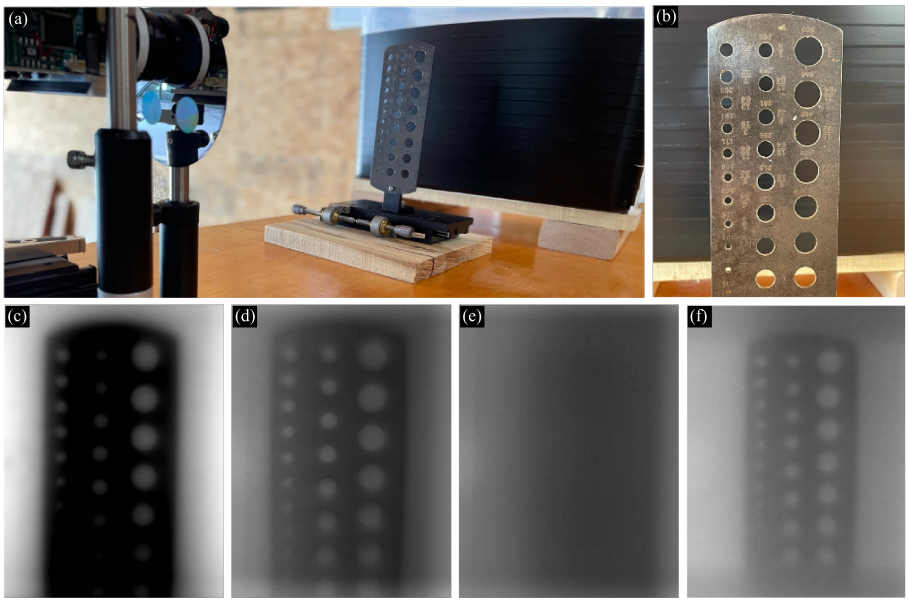

Figure 4.4: (a) The LWIR imaging setup. (b) The ground truth with a steel gauge in front of a isothermal radiation source controlled at 65 °C. (c) The broadband imaging result. (d) and (e) are the images when a bandpass filter centered at 10 µm and 12 µm are applied, respectively. The image is unintelligible in (e), however, when the image distance is adjusted, a clear image reappears, shown on (f).

4.4 Discussion

We are the first to demonstrate LWIR imaging using the simple metalens "in the wild", outside of the laboratory environment. Our simple metalens clearly shows a viable solution for imaging under real world ambient thermal emission applications. Here we discuss why a seemingly narrowband imaging system can be used for imaging under broadband, ambient thermal radiation. A simple metalens is known to have chromatic aberration under broadband illumination [37]: the focal length is proportional to $1/\lambda$, which is mainly caused by the discontinuities introduced by the phase-wrapping [16]. The shift in focal length across the color spectrum imparts blurs on the image. This phenomenon can be seen on Fig. 4.4

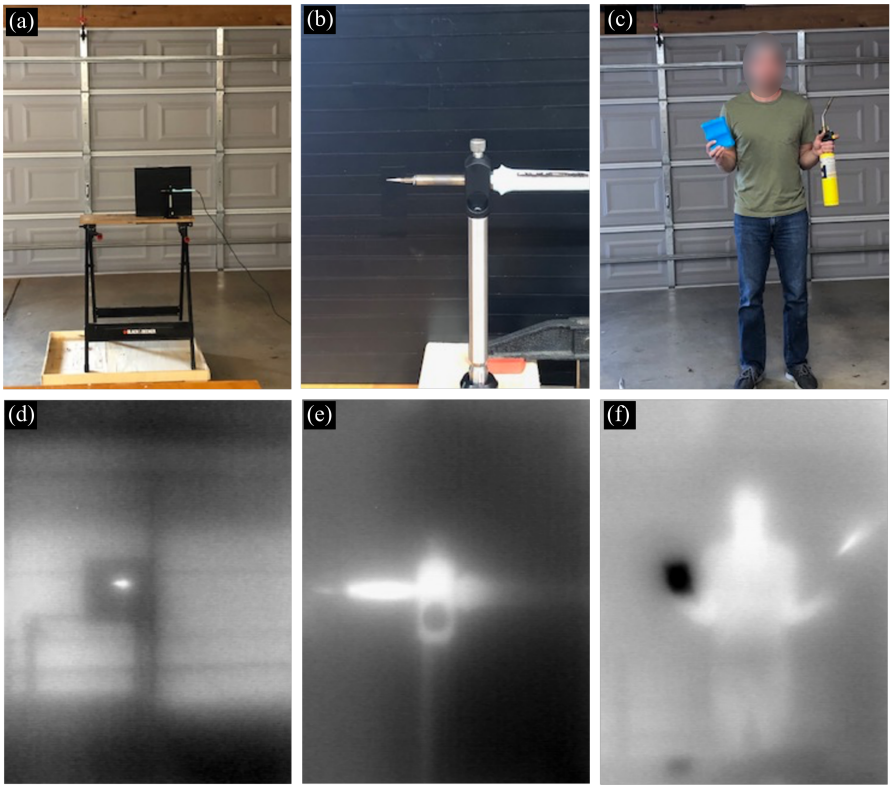

Figure 4.5: A hot soldering iron is seen on (a) in front of a steel plate covered with black tape with an emissivity of 0.95. A close up image of the soldering iron on a metal holder is seen on (b). A person holding an ice pack (left) and a blowtorch (right) is seen on (c). The LWIR imaging counterparts of these images are shown on (d), (e), and (f), respectively.

(e) where the input wavelength is longer than the nominal wavelength and is narrowband in which case the focal length gets effectively shortened. Under narrowband and non-nominal wavelength, the MTF of the imaging system characterizes and reflects the poor imaging quality capability of this chromatically aberrated system (Fig. 4.3 (a) purple curve), preserving very little spatial frequency.

In contrast, when imaging objects emitting black-body radiation at a variety of different temperatures, the MTFs of our imaging system drastically out-performs that of the 12 µm

narrowband lighting condition, shown on Fig. 4.3 (a). For temperatures ranging from 0 - 320°C the variation of the LWIR spectrum stays within an order of magnitude, seen on Fig. 4.3 (b). This relatively uniform spectral input ensures that there is always enough signal at 10 µm. The result gives rise to a viable imaging system with a theoretical cutoff frequency of $\sim$ 80% that of a diffraction-limited system. We note that the imaging performance is much worse in the visible wavelengths using colored camera, as the color filters distinguish different colors. Additionally the fractional bandwidth in the LWIR range ($\Delta\lambda/\lambda \sim 0.4$) is smaller compared to the visible range ($\Delta\lambda/\lambda \sim 0.6$).

In conclusion, we have demonstrated a simple-to-fabricate, large aperture, low f-number all silicon metasurface imaging platform capable of broadband imaging in the LWIR regime. We show that even with chromatic meta-optics, it is possible to image in the ambient light, including objects with different temperatures. This opens doors to applications such as wearable thermography devices, autonomous drones surveillance, and satellite imaging. Although we do not further process the captured images here, computational imaging combined with inverse designed meta-optics could significantly improve the overall imaging performance [71, 45, 72].

Chapter 5

Inverse Designed LWIR Meta-optics

5.1 Overview

We have since demonstrated the viability of meta-optical imaging in the LWIR regime in the previous chapter. In this chapter, we employ inverse-design to create a broadband meta-optics operating in the long-wave infrared (LWIR) regime ($8-12\mu m$). Via a deep-learning assisted multi-scale differentiable framework connecting meta-atoms to the phase, we maximize the wavelength-averaged volume under the modulation transfer function (MTF) of the meta-optics. Our design framework combines local phase-engineering via meta-atoms and global engineering of the scatterer in a single pipeline. We validate our design by fabricating and experimentally characterizing an all-silicon LWIR meta-optics. We also separately design a computational backend to ameliorate the captured images. We demonstrate 6-fold improvement of the wavelength-averaged Strehl ratio over the traditional hyperboloid metalens for broadband imaging.

5.2 Introduction

In recent years, LWIR meta-optics have been reported by several groups [73, 74, 75] to capture images in the ambient light. Unfortunately, the captured image quality remains poor, primarily due to strong chromatic aberrations. We have discussed in **§ 2.3** that the

traditional dispersion engineering approach to solve chromatic aberrations face fundamental limit [20], to overcome this fundamental limit and enhance the image quality beyond lens-only imaging, chromatic aberrations in a large-aperture meta-optics can be mitigated using our EDOF platform, as we discussed in chapter **§ 3**.

However, in most cases both approaches rely on a forward design methodology, for which the optical designer has to assign a spatial modulation distribution to the metasurface and therefore strongly rely on experience and intuition about the functionality of the meta-optics. As such, this method does not provide a clear path to further improve the imaging quality. Unlike forward design, inverse design approaches define the performance of the optical element via a figure of merit (FoM), and computationally optimize their structure or arrangement to maximize the respective FoM. The inverse design methodology has been very successful in creating non-intuitive yet functional meta-optics [42, 45, 76], including EDOF lenses for broadband imaging in the visible. A further refinement and extension of this approach is end-to-end design, where the meta-optics and computational backend are co-optimized with a FoM defined by the final image quality [71]. End-to-end optimization techniques couple the optimization of phase profiles and the computational backend. While this seems to be an advantage and takes the entire system into account, the downside is that we lack the insight into how and why the optic performs well. This can be detrimental when translating designs from the visible domain (with ample training data) to the thermal domain (with paucity of training data). As such, a new design paradigm is required for meta-optical imagers, which provides intuition on why such meta-optics can perform broadband imaging.

In this chapter, we report a 1 cm aperture, f/1 (numerical aperture: 0.45) broadband polarization-insensitive LWIR meta-optics, designed by a multi-scale optimization technique that maximizes the wavelength-averaged volume under the modulation transfer function (MTF) of the meta-optic. This method is "multi-scale" as distinct computational methods are used to calculate the electric fields at different length scales. The meta-optics are implemented in an all-silicon platform and we demonstrate that the captured images exhibit superior performance over traditional hyperboloid metalens. Unlike previous works, which

relied either on local engineering of meta-atoms (dispersion engineering) or global engineering of the phase-mask, our optimization technique employs both meta-atom and phase-mask engineering. We show that such an extended approach clearly outperforms a phase-mask engineering only approach.

5.3 Results

5.3.1 Inverse Design Framework

The key to achieve meta-optic broadband LWIR imaging is is a fully-differentiable inverse design framework that optimizes the binary structures (i.e. binary permittivity distribution) of the meta-atoms to achieve focusing of the desired wavelengths (**Fig. 5.1**(a)). However, designing a large-area ($\sim 10^3\lambda$) meta-optic is a computationally prohibitive problem, and full-wave simulation is not possible. Therefore, phase-mask optimization is generally used, but the scatterer-to-phase mapping on a sub-wavelength scale is not differentiable and a look-up-table based approach does not work for such an optimization methodology. In the past, polynomial proxy functions have been employed to connect scatterers to the corresponding phase [71], but were limited to only a monotonic relation. However, for broadband operations meta-atoms with a large phase diversity with multiple phase wrappings are required but suffer from multiple resonances at various wavelengths.

We resolve this obstacle by utilizing a more sophisticated approach, where a deep neural network (DNN)-based surrogate model maps the geometrical parameters ($\vec{\mathbf{p}}$) of the meta-atom to its complex transmission coefficient $\hat{\mathbf{E}}(\vec{\mathbf{p}}, \lambda)$. The utilization of DNNs enable modeling highly complicated functions, which is well suited for meta-optics. To emphasize this point, we compare meta-atoms with different degrees of parameterization: the "simple" scatterers are defined by one design parameter (the width of the pillar), whereas the "complex" scatterers have three parameters, **Fig. 5.1**(b). For each type of meta-atom, we sampled the parameters $\vec{\mathbf{p}}$ uniformly for each feature dimension and simulated it using rigorous coupled wave analysis (RCWA) [59], across different wavelengths λ. The simulated meta-atom library

is then taken as the training data set for the DNN that maps the parameterized features to the phase modulation. The RCWA simulations and the DNN fitting are pre-computed before the optimization loop, shown in **Fig. 5.1**(c). Although we are currently limited to parameterized meta-atoms, with the rapid progress of machine learning-enabled meta-optics design, we anticipate voxel-level meta-atom engineering will be possible in the near future. Also, our design currently relies on the local phase approximation, which is known to provide lower efficiency for meta-optics. However, recently a physics-inspired neural network has been used to design meta-optics without making such approximation and can be readily adapted for the scatterer-to-field mapping [77].

We define the FoM by computing the normalized MTFs first as $\hat{M}(k_x, k_y, \lambda, \theta)$, where k_x, k_y are defined as the spatial frequencies in the x and y direction, and θ signifies the incident angle (more details in section 5.5.1). The modified Strehl Ratio $\mathbf{S}$ can then be calculated using the following expression:

$$\begin{aligned}\mathbf{S}(\lambda, \theta) &= \int_{-k_{y1}}^{k_{y1}} \int_{-k_{x1}}^{k_{x1}} \hat{M}\left(k_x, k_y; \lambda, \theta\right) dk_x dk_y \\ &= \sum_i \sum_j \hat{M}\left(k_x(i), k_y(j); \lambda, \theta\right) \Delta k_x \Delta k_y\end{aligned} \tag{5.1}$$

Here, k_{x1} and k_{y1} denote the cutoff spatial frequencies in the x and y direction. Using the modifed Strehl Ratio, the FoM $\mathbf{F}$ is computed as:

$$\mathbf{F} = \log(\prod_i \prod_j S(\lambda_i, \theta_j)) \tag{5.2}$$

This FoM can be interpreted as the wavelength and incidence angle-averaged volume under the MTF surface, and we therefore term this optimization routine "MTF engineering". We emphasize that our FoM is maximized if all $S(\lambda_i, \theta_j)$ are maximized equivalently, thus constraining our meta-optics to have a uniform performance for the specified wavelengths without explicitly defining uniformity as an optimization criteria. Since the entirety of the forward computation of the FoM is implemented using differentiable operations, the gradient

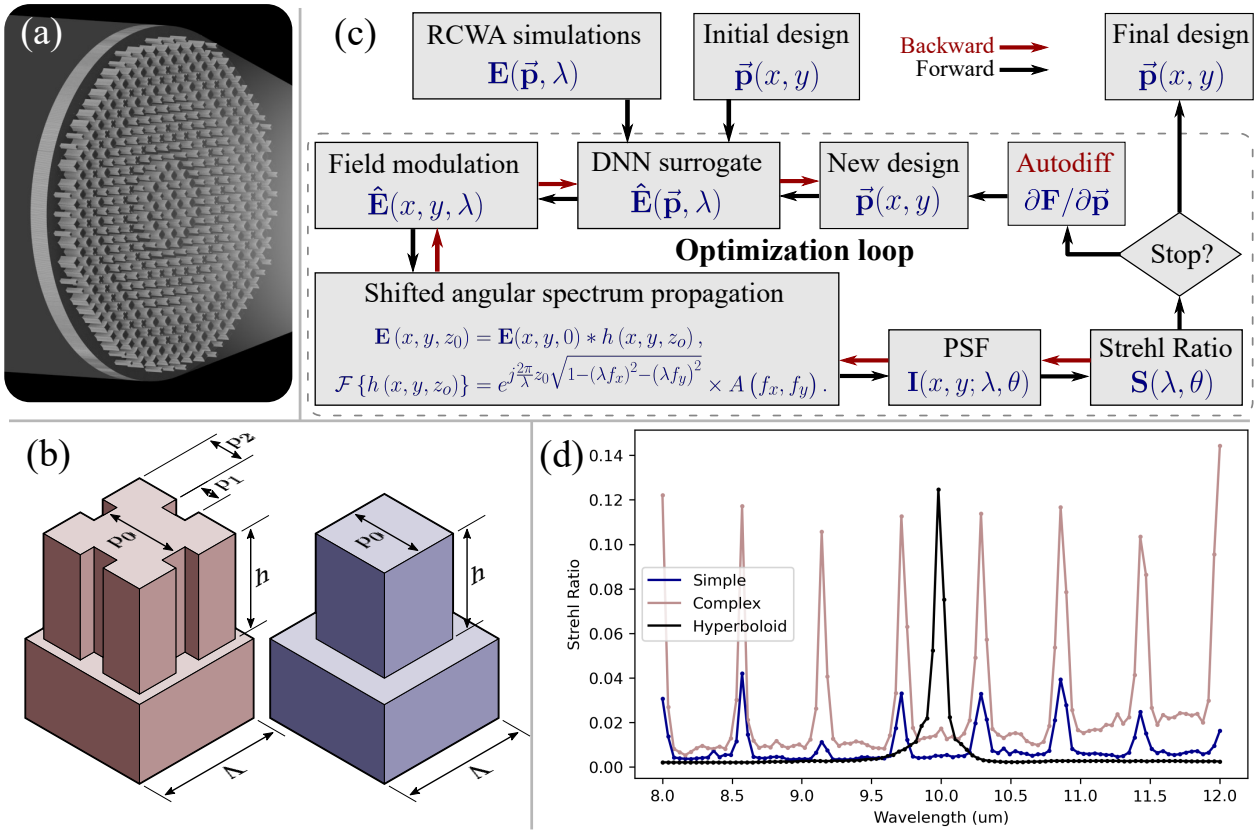

Figure 5.1: Design methodology to create the broadband LWIR meta-optics. (a) Our objective is a LWIR meta-optic that focuses broadband light at the same focal plane. (b) Parameterization of the "complex" (red) and the "simple" (blue) scatterers. For the complex scatterers, three variable parameters control the binary profile, namely, $\mathbf{p_0}, \mathbf{p_1}$, and $\mathbf{p_2}$. This method of parameterization ensures 4-fold symmetry of the geometry, thus ensuring polarization-insensitivity. The height $\mathbf{h}$ is kept constant at 10 µm. The meta-atoms sit on a Manhattan grid with a periodicity $\mathbf{\Lambda}$ set to 4 µm. (c) Block diagram of the optimization routine. (d) The simulated Strehl ratio of the "simple", "complex", and hyperboloid meta-optic.

of the FoM with respect to the geometry parameters, i.e. $\partial \mathbf{S} / \partial \vec{\mathbf{p}}$, can be obtained using automatic differentiation through chain rule in the direction indicated by the red arrows in **Fig. 5.1**(b). This gradient is then used by a stochastic gradient descent algorithm that

optimizes the structural parameters $\vec{\mathbf{p}}$. The optimization routine halts when the value of FoM converges, yielding a meta-optic design.

5.3.2 Meta-optics Design

Using the "MTF engineering" framework, we then designed the broadband LWIR meta-optics. For ease of fabrication, we chose an all-silicon platform, consisting of silicon pillars on a silicon substrate. We note that, while silicon is slightly absorptive in the LWIR range, approximately 80% of the light is still expected to be transmitted.

In more detail, we design two different broadband meta-optics with two scatterer archetypes, shown in **Fig. 5.1**(c), both of which are parameterized to exhibit fourfold symmetry as to ensure polarization insensitivity. We also forward-designed [41] a hyperboloid metalens with similar height and periodicity for a baseline comparison. All meta-optics have a nominal focal length of 10 mm, and a numerical aperture of 0.45. In simulation, we obtained orders-of-magnitude larger wavelength-averaged Strehl Ratios (using the 8 optimized wavelengths spanning 8 - 12 µm) in the optimized broadband meta-optics compared to a forward designed hyperboloid metalens: 0.15 (0.06) for the meta-optics with "complex" ("simple") scatterers against 0.0075 for a hyperboloid metalens. **Fig. 5.1**(d) illustrates the simulated Strehl ratios of the above-described optics with regard to the input wavelength. In this figure, individual meta-atoms were simulated using RCWA, while the DNN mapping was used for optimization only. Notably, the complex meta-optic design exhibits Strehl ratios at eight defined sampled wavelengths that are on-par with the Strehl ratio at a single operational wavelength of the hyperboloid metalens, and significantly exceeds the wavelength-averaged Strehl ratio of the hyperboloid meta-optic design by a factor of 6.

While spectral regions between the sampled wavelengths yield relatively lower Strehl ratios compared to the peak values, the wavelength-averaged Strehl ratio of the complex meta-optic design significantly exceeds that of the hyperboloid metalens . We also note that the Strehl ratios for the non-sampled wavelengths are significantly larger than the hyperboloid lens in those wavelengths. Our experimental results show that despite such

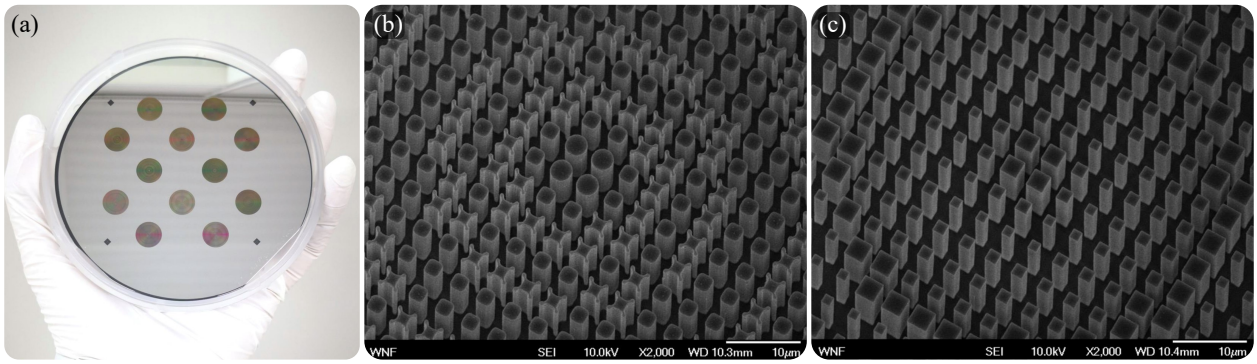

Figure 5.2: Images of the fabricated meta-optics: (a) Camera image of a fabricated wafer with several meta-optics. SEM images of the meta-optics with (b) complex scatterers and (c) simple scatterers.

polychromatic behavior, we can capture images under broadband ambient thermal radiation. This underlines the practical and versatile applicability of our broadband meta-optic designs in real-world scenarios.

We find that the remaining chromatic aberration can be mitigated effectively using computational deconvolution and denoising techniques. As a result, the broadband imaging capability of our meta-optics remains robust against spectral variations, which underpins the validity of our design framework to obtain a viable route to achieve broadband performing meta-optic designs.

5.3.3 Fabrication

We fabricated the LWIR meta-optics in all-silicon platform shown in **Fig. 5.2**(a) (details in Methods). The scanning electron microscope images of the fabricated complex and simple meta-optics are depicted on **Fig. 5.2**(b) and (c), respectively.

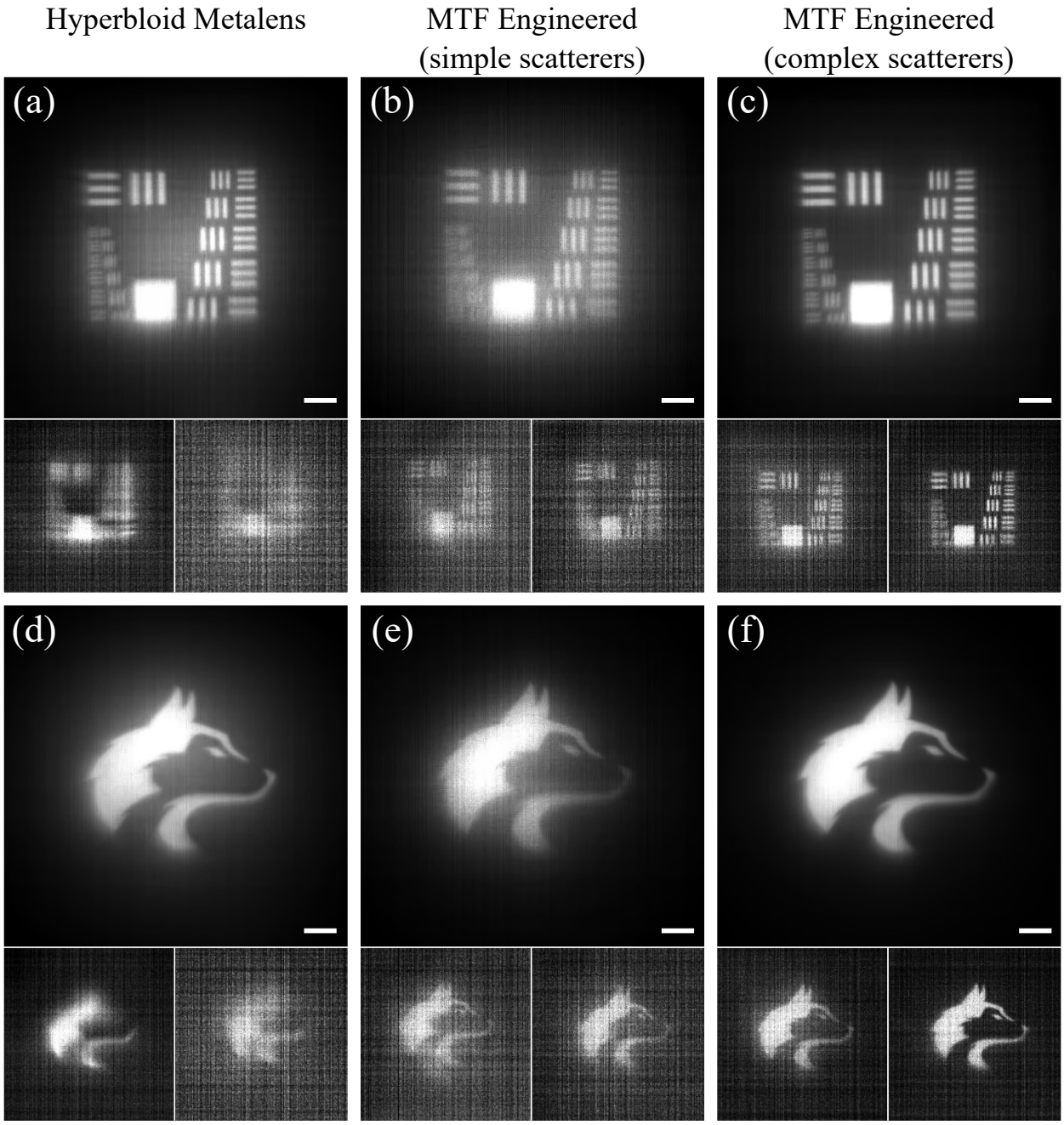

Figure 5.3: The broadband imaging results comparing the hyperboloid metalens (a, d), MTF engineered meta-optic with simple scatterers (b, d), and MTF engineered meta-optic with complex scatterers (c, f). The bottom images from each subfigure shows the imaging results without a filter (top), under a 10 µm bandpass filter (bottom left), and under a 12 µm (bottom right). The scale bar is 0.56 mm. Imaging results using two different targets are shown. They display a modified USAF 1951 pattern (top) and a UW husky logo (bottom).

5.3.4 Imaging

To demonstrate that the MTF-engineered optics have indeed superior imaging quality compared to the forward designed metalens, we compared their respected imaging quality under a controlled environment (details in section 5.5.5). As shown in **Fig. 5.3**(c), imaging results show the highest details on the MTF-engineered meta-optic with the complex scatterers, as all the triplet stripes are clearly visible and well separated.

To clearly show the broadband capability of MTF-engineered meta-optics, we use 10 µm and a 12 µm bandpass filters (linewidth of 500nm), while keeping the meta-optics to the sensor distance fixed. Due to chromatic aberration, the focal length of a hyperboloid metalens changes, and we place the sensor at the plane where the broadband imaging results were taken. These imaging results (without any denoising) are displayed on the bottom of each sub-figure on **Fig. 5.3**. We find that the MTF-engineered meta-optic with complex scatterers again outperforms the simple scatterers and the forward-designed optics for both the USAF and the husky targets. Interestingly, the MTF-engineered meta-optic with simple scatterers can be seen to outperform the forward-designed hyperboloid metalens as it shows better details in both the 10 µm and 12 µm cases and also for both patterns.

To further demonstrate the capabilities of the MTF-engineered LWIR meta-optics, we performed imaging testing “in the wild" under indoor, as well as ambient daylight conditions, shown in **Fig. 5.4**. In each case, we captured a single image, and performed a numerical deconvolution (details in the supplementary). The MTF-engineered with complex scatterers and forward-designed meta-optics were tested against a refractive lens for the ambient imaging. First, we imaged a truck parked outdoors shown in **Fig. 5.4**(a). The vehicle details are clearly visible such as the front grill. Further, background details such as the windows on the top right corner of the image, as well as the metal pillars on the left side of the image. In contrast, the hyperboloid image suffers from strong glowing artifacts in the center and fails to recover the sharp details. **Fig. 5.4**(b) shows image of a parked car. As with the previous example, the details are clearly visible, including the tyres on the car, and the branches of

the tree in the background with the MTF-engineered lens. **Fig. 5.4**(c) shows image of a person with hands stretched out. The difference between MTF-engineered lens and hyperboloid metalens is distinctinctly visible here. The sharp image features are clearly visible in the MTF-engineered lens, including variable heat patterns on the shirt. In contrast, the image formed by the hyperboloid lens is noisy with no clear features on the person.

5.4 Discussion

In this work, we have developed an inverse design methodology for broad band imaging meta-optics, guided by an easily translatable, intuitive and universal objective function, given by the volume of a multi-chromatic MTF surface. Using this MTF-engineering approach we achieve a sophisticated yet easy fabricable, large aperture, broadband LWIR meta-optic design, suitable for imaging "in the wild" under ambient temperature conditions. We experimentally verified this framework and demonstrated polarization insensitive broadband LWIR meta-optics with 1 cm diameter and NA of 0.45. Although previous works based on forward designed LWIR metalenses have demonstrated the imaging capabilities [41], they failed to resolve fine features due to strong chromatic aberrations. In contrast, the MTF-engineered meta-optics, show significantly improved performance over a broadband spectral range as well as narrowband imaging capabilities for wavelengths outside the center wavelength.

We further elucidate in simulation and experiment how a drastic performance enhancement can be achieved for the MTF-engineered meta-optics, if more structural degrees of freedom are considered, thus allowing a higher complexity of the meta-atoms. This more complex parameterization broadens the solution space in the optimization process and therefore yields a higher likelihood of reaching an improved FOM. This paves a clear pathway for future design works to utilize a performance boost by employing a higher degree of parameterization for the meta-optic scatterer, coupled with the large scale optimization of the meta-optic. We note, that previous works have primarily assumed one aspect or the other, while the potential of synergistic effects was overlooked or assumed to be negligible. However, we clearly demonstrat that by taking advantage of structural diversity can easily yield

Figure 5.4: LWIR imaging “in the wild". Three scenes were imaged using the refractive LWIR lens (left), the hyperboloid metalens (middle), and the MTF engineered meta-optic with complex scatterers (right). Scenes (a) and (b) were captuered outdoors on a sunny day while (c) was captured indoors.

a three-fold performance improvement. Yet, a clear downside of this is that, one must keep in mind that the sampling complexity increases exponentially as one increases the structure parameters. And beyond that the fabrication resolution becomes more stringent with

increased complexity of the meta-atoms. However, this can be potentially be overcome by using a more clever parameterization of the meta-atom, akin to what has been achieved in dispersion engineering approaches.

Overall, the developed MTF-engineering framework offers a general methodology to create large area broadband meta-optics. Our imaging and characterization results show a clear advantage of using the MTF-engineering methodology over the traditional forward design. We show that the MTF-engineered flat optics opens a new avenue for miniaturization of LWIR imaging systems with potential applications in unmanned aerial vehicles, target acquisition and designation sights, forward looking infrared cameras, and non-contact thermography.

5.5 Methods

5.5.1 MTF Calculation

Each meta-atom is assigned a position (x, y) on a Manhattan grid with a periodicity of $\mathbf{\Lambda}$ and a height of $\mathbf{h}$, whose geometric parameters are $\vec{\mathbf{p}}$. These geometric parameters are given random assigned initial values within the fabrication limited bounds. The spatial modulation of the meta-optic $\hat{\mathbf{E}}(x, y, \lambda)$ is calculated given the meta-atom position and the wavelengths of the incident fields. The spatial modulation is then multiplied by the incident field (planewave with angle θ) and propagated to the sensor plane centered to the chief ray, using the shifted angular spectrum method [78]. The intensity, $I(x, y; \lambda, \theta)$ for all permutations of the wavelengths and angles of incidence of the incident fields. This intensity is the point spread function (PSF) of the imaging system. The MTF (M) is computed from the PSF given by $\mathbf{M}(k_x, k_y; \lambda, \theta) =\mid \mathcal{F}(I(x, y; \lambda, \theta)) \mid$. The diffraction limited MTFs are calculated as $D(k_x, k_y, \lambda)$. This is to normalize the MTFs such that $\hat{M} = M(k_x, k_y, \lambda)/D(k_x, k_y, \lambda)$.

5.5.2 Optimization

For the design with the complex scatterer, we used 40 samples per feature dimension in $\vec{\mathbf{p}}$ while keeping h a constant of 10 µm. Combining with 101 samples of wavelengths λ, there are totally $40^3 \times 101$ RCWA simulations conducted for training of the DNN model. For the simple scatterer case, we sampled 1000 values for the only structure variable p_1 and 101 wavelengths, also keeping h constant of 10 µm, resulting in 1000×101 training samples for the surrogate. For training the DNN surrogate model, the Adam optimizer is used with a learning rate of 10^{-3} for 1000 epochs. To span a diameter of 10 mm, a total of 2500×2500 meta-atoms are used with a periodicity of 10 µm. During optimization, a total of 13 samples of wavelengths are used uniformly spanning from 8 µm to 12 µm. We selected 0° and 10° to be the angles of incidence.

The DNN surrogate model uses four layers of fully connected layers that have 256 units and hyperboloid tangent as the activation function for each layer. To ensure a differentiable mapping, the fitted surrogate model uses all differentiable operations.

The structure parameters $\vec{\mathbf{p}}$ are assigned initial values within the fabrication limited bounds. There are two conditions for the optimization loop to halt: (1) $\vec{\mathbf{p}}$ converges, i.e. $\vec{\mathbf{p}}_n - \vec{\mathbf{p}}_{n-1} < \mathbf{T}$, where $\mathbf{T}$ denotes a some threshold, and n represents the iteration number; (2) the number iteration reaches a certain integer $\mathbf{N}$. When the halting condition is met, the final meta-atom binary structure parameters are outputted, and converted to GDSII files for fabrication.

5.5.3 Solving the deconvolution inverse problem

Fig. 5.4 shows images in the wild captured with our setup. In each case, we deconvolved the measurements by solving an inverse problem using a data-free image prior. Each 640×512 image was modeled as output of an implicit neural representation (INR) [79, 80] $\mathcal{N}_\theta$ with parameters θ. INRs are continuous functions that map local coordinates $(x, y$ for images) to a given output (image intensity). A unique property enabled by certain INRs, particularly

ones equipped with a complex Gabor filter activation function, is the bias for images. This implies that the output tend to look more like images than noise. We leveraged this property to regularize the inverse problem. The specific inverse problem we solved is

$$\min_{\theta,K} \sum_{n=1}^{N} \sum_{m=1}^{M} \|I_{\text{obs}}(m,n) - (K * \mathcal{N}_\theta)(m,n)\|^2 + \eta_{\text{TV}} \text{TV}(\mathcal{N}_\theta(m,n)), \tag{5.3}$$

where I_{obs} is the observed 2D image, K is the PSF of the optical system, TV() is the total variation loss function, and η_{TV} is the weight for the total variation loss. We perform a semi-blind deconvolution where we initialize K to be the analytical PSF from our simulations, and then solve for the parameters of the network and the PSF together.

LWIR sensors based on microbolometers suffer from fixed pattern noise, which cases horizontal and vertical striations. Inspired by recent work on removing fixed pattern noise in thermal images [81], we modeled it as a low-rank image. We then simultaneously solved for the fixed pattern noise, the PSF, and the parameters of the network, giving us

$$\min_{\theta,K,F} \sum_{n=1}^{N} \sum_{m=1}^{M} \|I_{\text{obs}}(m,n) - F(m,n)(K * \mathcal{N}_\theta)(m,n)\|^2 + \eta_{\text{TV}} \text{TV}(\mathcal{N}_\theta(m,n)), \tag{5.4}$$

$$\text{s.t} \quad \text{rank}(F) = r, \tag{5.5}$$

where F is the fixed pattern noise. We used the recently developed wavelet implicit neural representations (WIRE) [80] for the INR architecture as it resulted in highest qualitative accuracy. An overview of the reconstruction pipeline is shown in **Fig. 5.5**.

5.5.4 Characterization and Imaging

To measure the PSFs of the fabricated optics, we used a collimated quantum cascade laser as the light source, which is focused to a 30 µm diameter pinhole by a Ge aspheric lens with an NA of 0.63. The pinhole is placed $2f$ away from the meta-optic, where f is the focal length of the meta-optic. A Ge aspheric lens and a ZnSe plano-convex lens are used to scale up the physical size of the PSF.

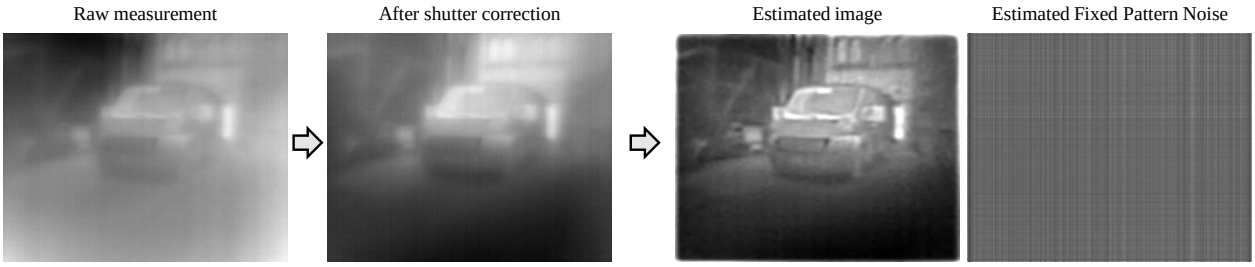

Figure 5.5: Deconvolution pipeline. We used an external shutter (flat black body) to remove any internall reflections. After shutter correction, we solved an inverse problem to simultaneously estimate a sharp image, and a fixed pattern noise.

5.5.5 In-lab Imaging

To assess the improved imaging performance of the MTF-engineered meta-optics over the forward-designed metalens, we captured imaged under broadband illumination using blackbody radiation from a hotplate with high-emissivity fiberglass tape heated to 150°C as the light source. Custom aluminum targets were made using laser cutting and finished with matte black paint to prevent reflections. These targets were placed in front of the hotplate, allowing patterned LWIR light to go through, creating contrast. A Boson 640 was placed on the imaging plane of the meta-optic in testing and sent the data to a PC for further post-processing, which included background subtracting, contrast stretching, and block-matching denoising. Through this predefined post-processing routine, we were able to improve the dynamic range and reduce microbolometer array artifacts.

5.6 Experimental Details

5.6.1 Alignment Sensitivity

To study the robustness of the optimized optics with respect to the shift in the sensor plane, we conducted simulations with varying sensor distances. **Fig. 5.6** presents the spectral Strehl ratio for our complex meta-optics at sensor distances of 10.0 mm, 10.5 mm, and 11.0 mm. The

wavelength-averaged Strehl ratios of these simulated systems are 0.0277, 0.0230, and 0.0157. It is observed that the peak of the Strehl ratio - denoting the system's highest performance - blueshifts, or shifts towards shorter wavelengths, with an increase in the focal plane. This phenomenon is governed by the principle that the wavelength of light and the focal length are inversely proportional [54]. We find that even though the wavelengh-averaged performance is robust against shift in the sensor plane, the individual peaks are still susceptible to the shift.

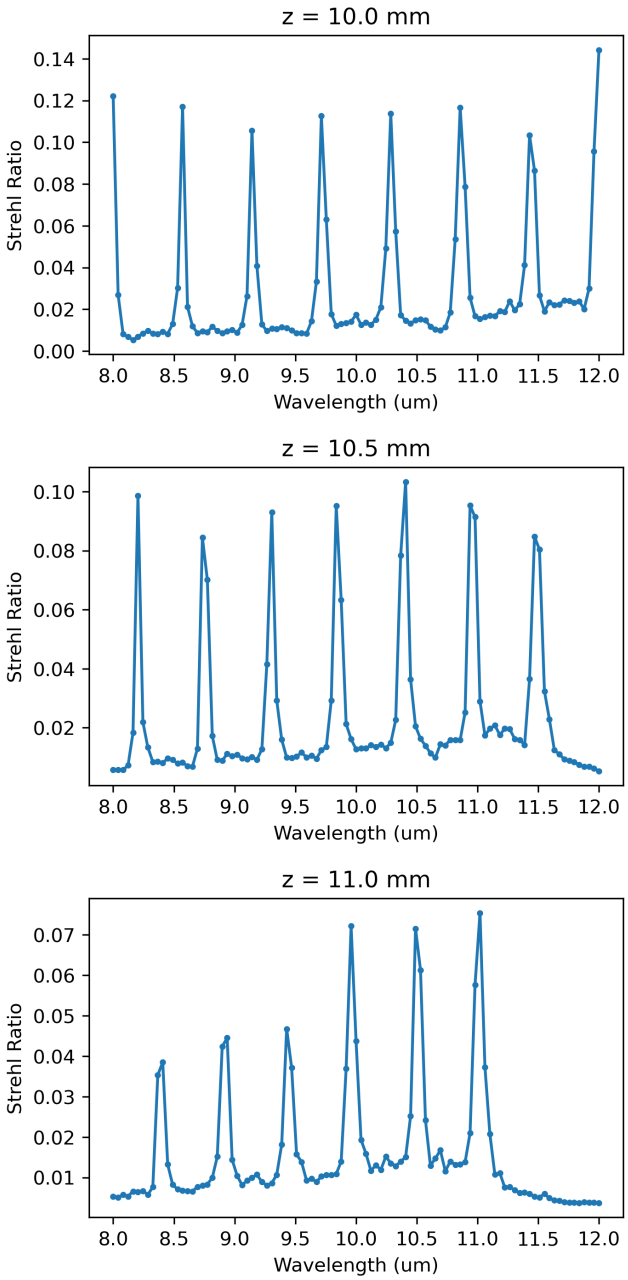

Figure 5.6: The Strehl ratio of the complex meta-optic as a function of the input wavelength at different focal planes. From to bottom, the focal lengths are 10.0, 10.5, and 11.0 mm.

5.6.2 Characterization

Our initial experimental step involved measuring the point spread functions (PSFs) of the fabricated devices across various wavelengths, which can be seen in **Fig. 5.7**. These measurements were taken in light of the alignment sensitivity expounded upon in Section 5.6.1, necessitating sparse sampling of PSFs across the spectrum at specific wavelengths—namely, 8.0, 8.5, 9.5, 10.0, and 10.5 micrometers. This selection of wavelengths was strategically chosen to facilitate a direct comparison with simulated PSFs, thereby enabling us to validate our models effectively. It's important to note that due to the aforementioned alignment sensitivity (as discussed in Section 5.6.1), we deliberately refrained from capturing PSFs at the optimized wavelengths. Subsequently, the experimentally obtained PSFs were evaluated in conjunction with their corresponding modulation transfer functions (MTFs) and compared against their simulated counterparts. Gratifyingly, the empirical results displayed a strong correlation with the simulations, endorsing the accuracy and reliability of our computational models.

The hyperboloid metalens performs well at the 10 µm as the PSF spot size is very compact at 10 µm, also its MTF curve shows a cutoff frequency of $\sim$ 20 lp/mm, compared to a diffraction limited cutoff frequency of 100 lp/mm. Although the experimental setup incurs additional aberrations from the relay optics effectively decreasing the cutoff frequency. The beam size of the hyperboloid metalens quickly broadens as the operating wavelength deviates from 10 µm showing a cutoff frequency less than 2 lp/mm for other wavelengths. This is due to the inherent chromatic aberration of the diffractive nature of the metalens [82]. In contrast, both MTF-engineered met-optics maintain a reasonably sized focal spot throughout the entire spectrum and maintain a cutoff frequency over 5 lp/mm for all wavelengths. We emphasize this is exactly what we intended in our design: rather than optimizing for one wavelength, we wanted to have a uniform MTF over the whole wavelength range.

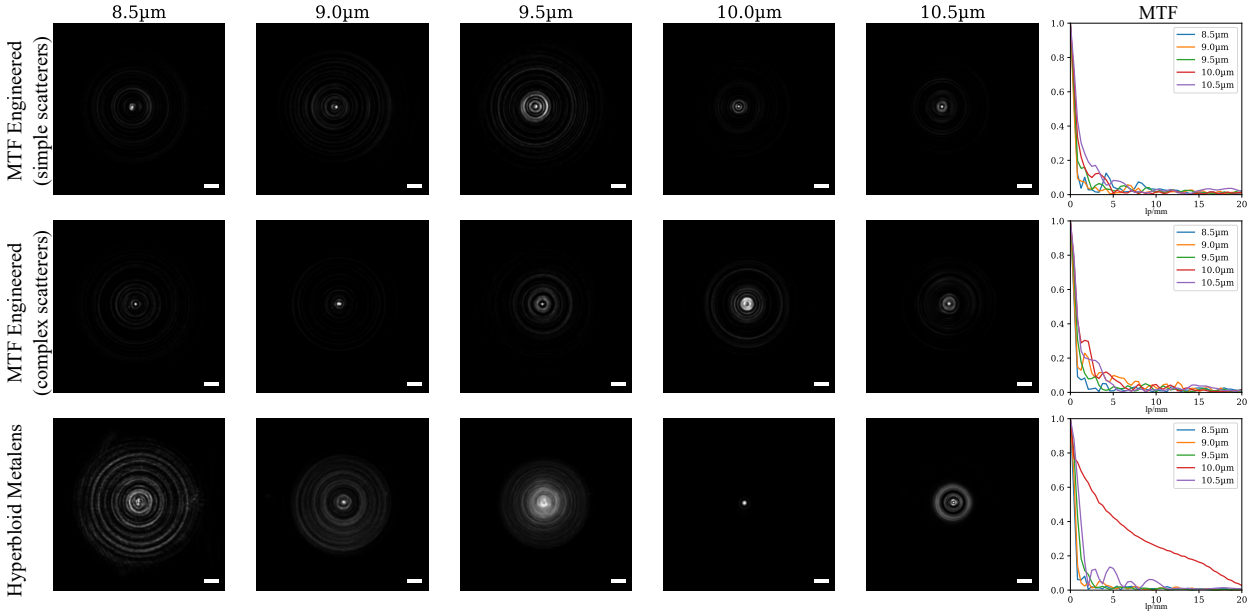

Figure 5.7: PSF and MTF characterizations for the hyperboloid metalens (top), MTF engineered meta-optics with simple scatterers (middle), and MTF engineered meta-optic with complex scatterers (bottom). The scale bar is 200 µm in width.

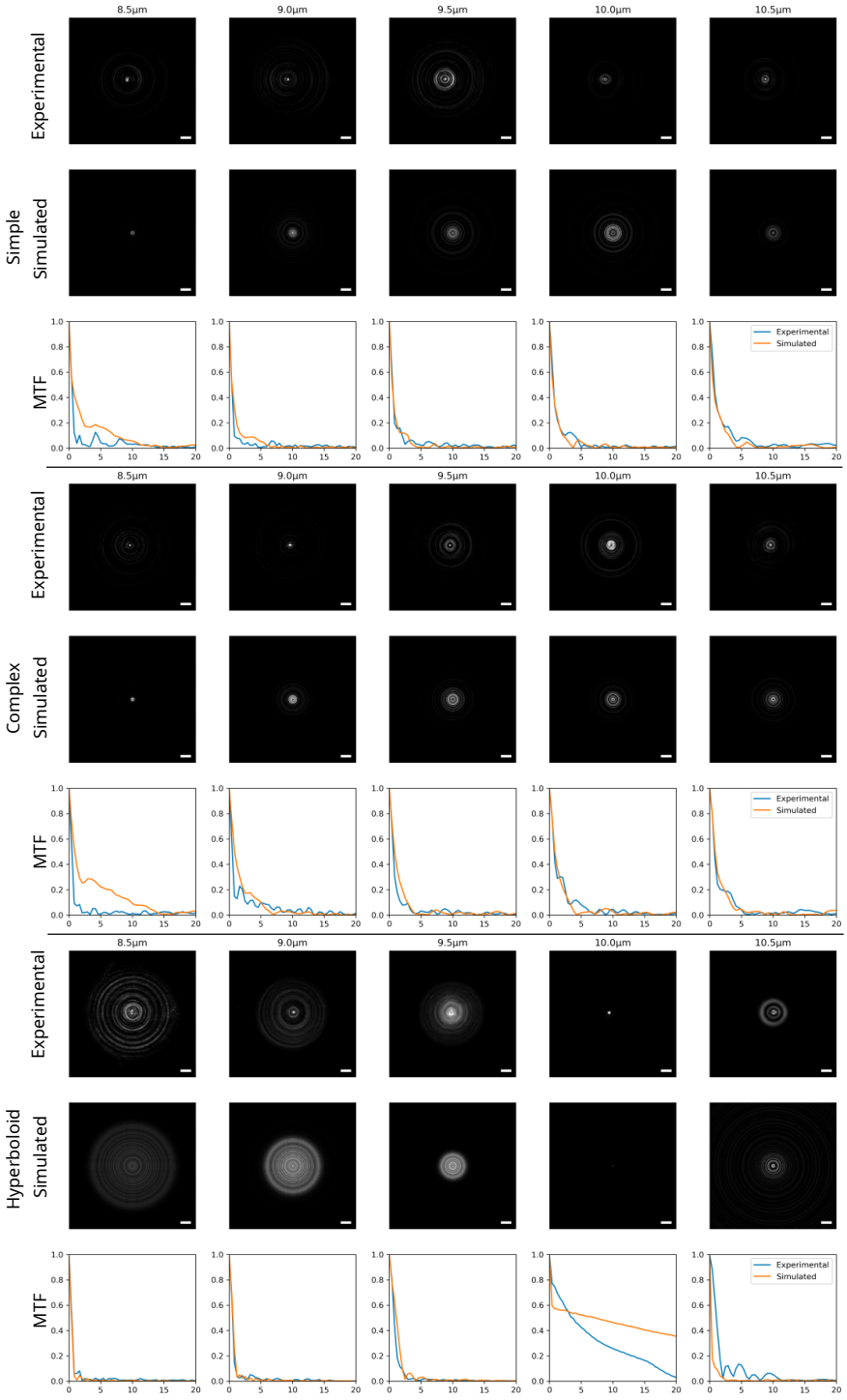

Figure 5.8: Simulated vs. experimental PSFs and MTFs. Comparing between "simple", "complex", and 'hyperboloid' meta-optics, from top to bottom, respectively. The scale bar is 200 µm in width.

5.6.3 Additional Imaging Results

To demonstrate the imaging capabilities of the MTF-engineered LWIR meta-optics, in addition to the imaging results in the main text, we conducted further imaging testing "in the wild" under ambient daylight conditions, shown on **Fig. 5.9**. The MTF-engineered and forward-designed meta-optics were tested against a refractive lens as the ground truth. First, we imaged a vehicle parked outdoors shown on **Fig. 5.9**(a). In this scene, the vehicle is clearly visible using the MTF engineered meta-optics with good details on the subject as well as the background. The hyperboloid metalens, however, does not capture the background with any fidelity, also leaving parts of the vehicle without contrast. The second series of captures on **Fig. 5.9**(b) show a scene of a water fountain in front of building. This series re-demonstrates the increase in details via the MTF engineered meta-optic capture, allowing observers to see the trees behind the water fountain, which the hyperboloid metalens fails to demonstrate. The third series on **Fig. 5.9**(c) shows a woman resting against a window. The hyperboloid metalens fails to capture the subject's glasses, which shows up on the MTF engineered case. Finally, we captured a metal "W" sculpture against vegitations in the background shown on **Fig. 5.9**(d). The MTF engineered capture displays better subject seperation against the background when compared to the hyperboloid metalens.

Figure 5.9: LWIR imaging "in the wild". Four scenes are imaged using the refractive LWIR lens (left), the hyperboloid metalens (middle), and the MTF engineered meta-optic with complex scatterers (right). The four scenes were all captured during daytime on a sunny day depicting (a) a parked car on a sunny day; (b) a fountain in the foreground aginst a building in the background; (c) a woman resting against a window; (d) a metallic structure.

5.6.4 Meta-atom Modeling

Fig. 5.10 presents the training loss evolution for our simple and complex meta-optic designs over multiple epochs. The models were trained using a 10-layer fully connected network (FCC) with 128 units per hidden layer and a ReLU activation function. Both models were trained using the Adam optimizer, with a learning rate set at 0.0006. The simple meta-optic design concludes training with a remarkably low loss of 0.007, while the more complex design, despite its additional intricacies, achieves a respectable final loss of 0.048. The consistent decrease in training loss over time, seen in both scenarios, validates the robustness and effectiveness of our deep-learning assisted differentiable framework in managing the design complexities inherent in meta-optics.

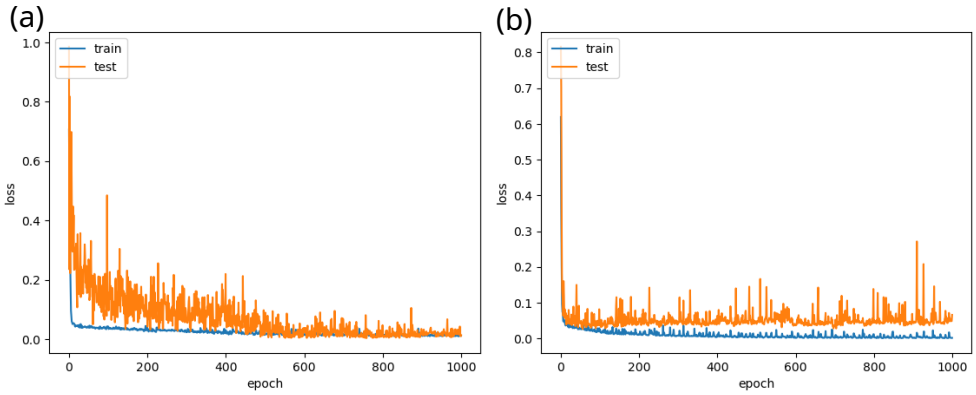

Figure 5.10: Meta-optic training loss progress for the simple (a) and the complex (b) meta atoms: each plot shows the downward trend of training loss as a function of epochs, indicating model optimization.

5.6.5 Fabrication

The meta-optics are fabricated on a 500 µm thick double-side polished silicon wafer, lightly doped with boron, giving a sheet resistivity of 1-10 Ω-cm. Direct-write lithography (Heidelberg DWL 66+) defines the aperture locations in the photoresist covering the wafer surface. A 220 nm thick aluminum layer is deposited onto the patterned photoresist via electron beam evaporation (CHA Solution) and lifted off to construct the metal mask around the circular apertures, helping to reduce noise during the experiments. The wafer is coated with another photoresist and patterned with the meta-optic scatterers via direct-write lithography, aligned inside the defined apertures. The photoresist pattern of meta-optics is transferred to the bulk silicon to a depth of 10 µm by deep reactive ion etching (SPTS DRIE), utilizing its capability to etch with high aspect ratios and vertical sidewalls. After etching, the photoresist residue is stripped, and the fabricated meta-optics are ready for characterization. This process is shown on **Fig. 5.11**.

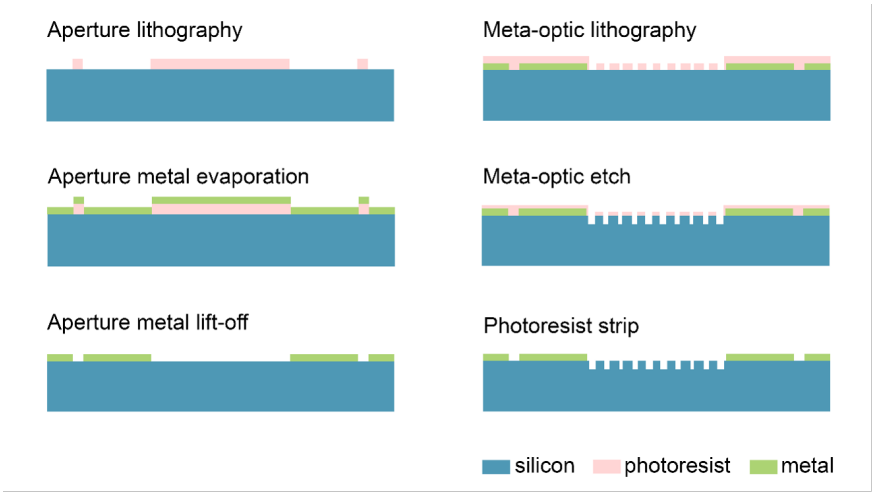

Figure 5.11: Pictorial representation of fabrication process.

Chapter 6

Meta-optics for Optical Computation[1]

6.1 Overview

Light's ability to perform massive linear operations parallelly has recently inspired numerous demonstrations of optics-assisted artificial neural networks (ANN). However, a clear advantage of optics over purely digital ANN in a system-level has not yet been established. While linear operations can indeed be optically performed very efficiently, the lack of nonlinearity and signal regeneration require high-power, low-latency signal transduction between optics and electronics. Additionally, a large power is needed for the lasers and photodetectors, which are often neglected in the calculation of energy consumption. Here, instead of mapping traditional digital operations to optics, we co-optimized a hybrid optical-digital ANN, that operates on incoherent light, and thus amenable to operations under ambient light. Keeping the latency and power constant between purely digital ANN and hybrid optical-digital ANN, we identified a low-power/ latency regime, where an optical encoder provides higher classification accuracy than a purely digital ANN. However, in that regime, the overall classification accuracy is lower than what is achievable with higher power and latency. Our

[1]The content of this chapter is based on the following publication:
Luocheng Huang, Quentin AA Tanguy, Johannes E Froch, Saswata Mukherjee, Karl F Bohringer, Arka Majumdar, "Photonic Advantage of Optical Encoders", arXiv:2305.01743 (2023)

results indicate that optics can be advantageous over digital ANN in applications, where the overall performance of the ANN can be relaxed to prioritize lower power and latency.

6.2 Introduction

Over the last decade the fields of artificial intelligence (AI) and deep learning have experienced accelerated progress, revealing the potential and capabilities of artificial neural networks (ANN) for a variety of applications, with recent demonstrations even advancing to the public spotlight in the form of chat software and artistic rendering programs. Their recent success can be traced back to major breakthroughs, both in terms of computational algorithms and digital hardware such as graphics processing units (GPU) [83]. While impressive, the scaling of power and latency of digital implementations of deep learning turned out to be unfavorable with the size of the ANN. This poses a serious limitation for further scaling of [84, 85] and applicability to low-power, realtime problems.

Light may be the answer to this scaling, thanks to its inherent parallelism, speed, and analog nature, thus providing an attractive alternative to electronic implementations to build energy efficient and fast artificial neural networks (ANNs). This has been recognized early on and several experiments reported optical ANNs already back in the 1990s [86, 87]. Unfortunately, progress stalled due to technological and fundamental reasons, which can be broadly classified into intrinsic and extrinsic problems. Intrinsic problems with optics had been the large size and poor tolerance to misalignment of optical components; limited space bandwidth product of spatial light modulators; and lack of nonlinear activation. The extrinsic problems originated from poor understanding of AI algorithms and adaptive learning, as well as the meteoric rise of electronic computing systems.

Given the current limitations of electronic hardware and our increased understanding of AI, the extrinsic problems are somewhat alleviated. In parallel, the advancement of nano-fabrication facilities, and the availability of sophisticated electromagnetic simulators have led to the highvolume manufacturing of multi-functional nano-optics, such as flat meta-optics [54, 88] and integrated photonic devices [89]. Emerging material systems coupled with

these nano-optical structures enable monolithic photonic integrated circuits (PIC) analogous to electronic ICs [90]. These innovations in nanophotonics and AI, combined with severe limitations of digital implementation of ANNs have generated strong interest in recent years in recreating optics assisted ANNs [91, 92, 93, 94, 95, 96, 97, 98]. However, thus far, none of the reported works has demonstrated a clear advantage of optics over digital ANNs for inference. Most implementations have only shown the substitution of a small linear part with an optical counterpart, while the rest was kept in the digital electronics. Although there is a clear advantage of optics for implementing a small sub-system, often the linear part, the power and latency in a complete ANN include the transduction of the signal between optical and electronic domains [99], i.e. the detector readout power and laser power, many of which are often neglected. In fact, an analysis considering these energy costs shows that implementing only one convolutional layer in optics does not provide any advantage, unless the input has a very large dimension [99]. Additionally, a large body of works demonstrated classification for extremely simple "toy" problems, for which no digital benchmark exists [94, 95]. Comparing the power and latency of an application specific optical ANN to a GPU (optimized for universal operations) is unfair. There are many ways to drastically reduce the power and latency of a digital ANN, including replacing matrix multiplication with XNOR operations [100]. Many pruning algorithms also exist to reduce the number of computations needed for inference. As such, there has been no clear demonstration where an optics-assisted ANN shows an advantage over a purely digital framework optimized for solving a specific problem. One challenge is that it is impossible to exactly define the computational complexity of an ANN, hence the exact calculation of power and latency in the digital part is dependent on both training and technology.

Here, we develop a framework to exactly compare the inference performance of a pure digital ANN against a hybrid optical-digital ANN. In both ANNs, we ensure the same power and latency, and thus by comparing the classification accuracy, we can clearly assess the relative advantage. **Fig. 6.1** shows the schematic of the two cases: the pure digital and the hybrid optical-digital. We encode the input in incoherent light, as the optical frontend of

the ANN can work with ambient light without incurring any additional energy. In a pure digital case, a lens-based sensor captures an image of an object under incoherent light, then the image is transferred to a digital ANN. For the hybrid case, we use an engineered optic - namely the optical encoder, instead of a lens, that captures the image in a different basis and sends the data to a digital backend. Instead of implementing a digital sub-system, such as convolutional operations in optics, we co-optimize the optical frontend (implemented via a sub-wavelength diffractive meta-optics), along with the digital backend using an end-to-end design framework (detail in **§ 6.5.1**, **§ 6.5.3**)[46]. The topology and resources (i.e., the same number of nodes, layers, and nonlinearities) used in the digital ANN are kept the same in both cases, though with different weights and biases. Thus, we ensure that the latency and power consumption in both cases remain identical.

Here, we tested the classification accuracy for MNIST data sets for different values of N, which represent the binned size of the image captured in the sensor either via a lens or the optical encoder. As the latency and power increase with the input dimensionality N of the data sent to the digital ANN, we found that classification accuracy increases in both cases, and there is no advantage from an optical frontend for large N. However, for smaller N, where the system power and latency are also lower, we found an increase in validation accuracy (10%) with a hybrid optical-digital ANN. We experimentally validated our theoretical model. Our work clearly demonstrates a photonic advantage for ANN inference, albeit such an advantage is observed when overall system performance is lower than the highest achievable performance.

6.3 Results

Our digital backend consists of three fully connected layers: $N \times 256$ (input), 256×256 (hidden) and 256×10 (output). The first two layers are each followed by a rectified linear unit (ReLU) nonlinearity and the output layer has a sigmoid nonlinearity. For the pure digital case, every image is converted to an N-pixel image by averaging the pixels. We chose 8 different N ranging from 1 to 100 , to assess the performance of the system with increasing

data input. We train the digital network by back-propagating the loss function defined by the cross-entropy between the output and the ground truth. In simulation, we obtained a validation classification accuracy of up to 98% (detail in). We note that, in prior works, to achieve a similar accuracy with the MNIST dataset, several layers were used (17), which we attribute to inefficient training. For the hybrid case, we model the optical frontend using a sub-wavelength diffractive meta-optics, although any freeform optical surface could suffice for implementation. The fabricated optical frontends with different output dimensionalities are shown in **Fig. 6.2**(b). We train the meta-optics along with a digital backend with the same neural network topology (details in method), following a similar framework used before for imaging (20). As expected, we observed an increase in classification accuracy with increasing N. We also found that for $N > 8 \times 8$, the digital and hybrid ANN demonstrate identical classification accuracies. However, at a lower value of N, the classification accuracy of the hybrid ANN surpasses that of the digital ANN. Example classification confusion matrices are shown in **Fig. 6.3**(a), comparing the experimental validation accuracies between a hybrid and a digital ANN with the same input size, $N = 3 \times 3$. Theoretically, we observe an increase in classification accuracy by up to $\sim$ 20% when an optical frontend is incorporated. A validation accuracy comparison chart can be seen on **Fig. 6.3**(b). We note that, even with a single data-point sent to the digital backend, we expect to see a higher classification accuracy with our optical frontend. This is because that single input can have 256 different values for an 8-bit precision sensor.

To validate the design, we fabricated the meta-optics (detail in) and measured their performance experimentally, where we projected images of the MNIST data set using an OLED display in green (detail in). The incoherent green light passes through the meta-optics, and we capture the data on the sensor with 8-bit precision. We then binned the captured image to create the N data-points that are passed to the digital backend. An experimental sample on **Fig. 6.2**(c) shows the signal processing of the 3×3 encoder. Due to fabrication imperfections, and misalignments, we retrained the digital backend (with the same topology) using the captured data. Our experiment matches the theory very well for

$N \geq 3 \times 3$. We note that the meta-optics optimized for $N = 8 \times 8$ was damaged and we could not collect data on that. At smaller N, the deviation from the theory is attributed to experimental noise. While a single point can provide more information to the digital backend, it is corrupted by the quantization noise, undermining the effect of the optical encoder and we obtained a similar classification accuracy, as we would have expected from a pure digital backend.

6.4 Discussion

By employing an incoherent light source and a meta-optical frontend, we created a framework, enabling us to compare the performance of a digital ANN to an optics-assisted ANN in the same footing. While keeping the power and latency constant in both cases, we showed that optical encoding does provide more information to the digital backend, resulting in $\sim 10\%$ more classification accuracy in the experiment. While our result is primarily applicable to the MNIST dataset, we believe that it indicates the conditions for which an optical frontend is beneficial to increase the performance of an ANN (more discussion in). Without any constraints on latency and power, one can arbitrarily increase N, and always find a digital solution that is better than the hybrid option. One way to rationalize this is that any optical implementation can be modelled digitally and therefore without any constraint a digital solution can be found with an accuracy in the same order of magnitude or higher than its optical counterpart. The higher classification accuracy of optics-assisted ANN in several reports is most likely a manifestation of poor training of the fully digital ANN. However, under the constraints of latency or power, we need to work with an intermediate value of N, where the optical frontend can provide a more efficient solution, albeit at overall lower accuracy.

6.5 Methods

6.5.1 Design of the Hybrid Optical-digital NN

We implemented an end-to-end differentiable pipeline to compute the gradient of the phase distribution on the meta-optical encoder with respect to the classification accuracy of the hybrid optical-digital neural network. The pipeline consists of three stages, namely, point spread function (PSF) simulation, imaging simulation, and classification. The loss function is given by the crossentropy between the output of the hybrid neural network and the ground truth. We used TensorFlow 2.8 as the automatic differentiation engine to implement the forward computations.

In the PSF simulation stage, light with normal incidence is propagated through the meta-optics, with the phase modulation distribution denoted by $\phi(x, y)$ ("Encoder" in **Fig. 6.4**). The phase distribution is parameterized by z_j, linear combinations of Zernike polynomials R_n^m, in which:

$$R_n^m(\rho) = \sum_{k=0}^{\frac{n-m}{2}} (-1)^k \begin{pmatrix} n-k \\ k \end{pmatrix} \left(\frac{n-m}{2} - k \right) \rho^{n-2k} \tag{6.1}$$

We used 200 terms of the Zernike polynomials to parameterize the phase distribution surface such that $\phi = \sum_{j=1}^{200} z_j R_j$, where $R_j \to R_n^m$, given $j = \frac{n(n+2)+m}{2}$.

These polynomials are precomputed and stored into the memory, kept until the end of the entirety of the optimization loop. The Zernike polynomials are orthogonal bases that provide spatial regularization to the phase modulation of the metasurface. We found that the employment of such basis functions, instead of optimizing phase value at each spatial location, prevents the optimization from getting stuck in local minima during optimization. After computing the phase distribution, we propagate the complex field using the bandwidth-limited angular spectrum method to obtain the intensity at the focal plane, i.e., the PSF.

The second stage of the end-to-end pipeline is to convolve the PSF with the batched input MNIST images. The input images are first up-sampled using the bilinear interpolation to

be the same size as the PSF intensity array. These input images are then stored in memory for the rest of the optimization routine. After that, the PSF is convolved with the input images using the Fourier convolution theorem such that $O = \mathcal{F}^{-1}\{\mathcal{F}\{PSF\} \cdot \mathcal{F}\{I\}\}$, where O denotes the output image, and I denotes the input image. The output O can be seen on the "Monitor" of **Fig. 6.4**. We note that this way of modelling imaging is valid only for imaging under incoherent emission, as we aim to demonstrate in this work.

The output images are then fed into a trainable digital backend ANN with sequential layers including an $N \times N$ average pooling layer, two fully connected layers with 256 units each with ReLU activation function, and finally a fully connected layer with 10 units with softmax activation function as the output layer. Note that this digital backend ANN architecture is kept the same between the hybrid optical-digital ANN as well as the purely digital ANN. Note that the $N \times N$ is equal to the input size of the digital backend. While the exact latency and power of the digital backend will depend on the technology (software, GPU or an ASIC optimized for a specific ANN), both of which are expected to monotonically increase with increasing value of N.

After the forward computation, we obtain the cross-entropy loss $\mathcal{L}$ between the output of the ANN and the ground truth label of the MNIST data set. The automatic differentiation algorithm then starts the backward computation in which we obtain the gradient of the loss with respect to the Zernike polynomial coefficients, namely $\partial\mathcal{L}/\partial z_j$. This process is visually represented in 6.4 by the red arrows. Then we apply this gradient to the parameters z_j multiplied by a factor provided by the Adam optimizer with a learning rate of 0.001 .

The training loop includes the forward computation of the loss function, as well as the backward computation of the loss gradient. This training loop is done iteratively until the loss converges. We find that 200 iterations are sufficient for a convergence. Both the forward and backward computations are done on an Intel Xeon @ 2.20 GHz, accelerated by a Nvidia Tesla P100 with 16GB of RAM. The optimization ran 4 times for each input size ($N \times N$).

The training and validation accuracies are displayed on **Fig. 6.5**. The training confusion matrices are shown in **Fig. 6.6**.

6.5.2 Design of Purely Digital neural network

The purely digital artificial neural network is designed with the same sequential architecture as the hybrid neural network's digital backend, comprising four layers: an $N \times N$ average pooling layer; two layers of fully connected neurons with 256 units and ReLU activation functions; and a 10-unit fully connected layer with a softmax activation. The optimization of the neural network follows an identical routine, including the optimizer scheme and the number of iterations, and is conducted using the same hardware. The optimization routine consists of 150 epochs using the Adam optimizer with a learn rate of 0.001 . Most of the optimizations converge within $\sim$ 50 epochs. The training and validation accuracies are displayed on **Fig. 6.7**. The training confusion matrices are shown in **Fig. 6.8**.

6.5.3 Design of the meta-optics

The metasurface comprises a 2D array of SiN meta-atoms on a Manhattan grid, with a periodicity of 350 nm in both x and y directions. Each meta-atom is shaped as a square pillar, with a fixed height of 775 nm and lateral width parameterized to span the periodicity of the grid. This

configuration is simulated using rigorous coupled-wave analysis [59], and the phase modulation of the metasurface is mapped to the corresponding meta-atom that yields the closest phase modulation to the target $\phi(x, y)$. The meta-atoms are situated on a quartz wafer with a thickness of about 500 μm.

6.5.4 Fabrication of the meta-optics

For fabrication, we first deposited a 775 nm thick SiN film on a 500μm thick quartz wafer using plasma enhanced chemical vapor deposition (PECVD) in a SPTS PECVD chamber. A positive resist (ZEP 520A) was then spun onto the wafer, followed by baking at 180°C for 3 minutes. To minimize charging effects during patterning, a conductive polymer layer (DisCharge H2O) was subsequently spun on top. The resist layer was then patterned using

a 100kV

layer (80 nm) of AlO_x was deposited using electron beam evaporation. After overnight liftoff in NMP heated at 90°C, the SiN layer was etched to the depth of 700 nm$(+/-2$ nm) with a remaining AlOx thickness of ~ 10 nm. using a mixture of C_4 F_8/SF_6 in an inductively coupled reactive ion etcher (Oxford PlasmaLab System 100). For SEM imaging a thin conductive Au/Pd layer was deposited to prevent charging.

6.5.5 Meta-optical encoder experimental details

For the hybrid neural network experimental measurements, we first measured the label-feature pairs. The MNIST features, i.e., handwritten black and white images, are scaled up using the nearest neighbor interpolation then displayed on an OLED monitor (SmallHD 5.5 in. Focus OLED HDMI Monitor), as shown on **Fig. 6.9**. The image is first displayed via an OLED monitor, set 10 cm away from the meta-optical encoder. The signal then goes through a custom microscope to transfer the output of the meta-optical encoder to the sensor. The power consumption of the OLED display is ~ 15 W. Albeit, since the sole purpose of the incoherent light source here is to provide an emulation of the real-world object, the actual power budget of the classification system should not include the OLED monitor.

Some examples of the experimental captures are shown in **Fig. 6.11**. First, the input is displayed on the OLED monitor shown on the top row. Then the encoder processes the input and projects the signal on the sensor, shown on the second row. The third row shows the encoder output after the average pooling. The fourth row shows what the purely digital ANN receives when there is no optical encoder.

For each meta-optical encoder, we first measured the PSF responses which match up well with the simulated PSFs, seen on **Fig. 6.10**. To calibrate the image height on the OLED monitor, we displayed a cross calibration pattern with varying scaling factors and chose the closest image that resembled the simulated output. The images are then judiciously cropped to reflect the correct physical extents of the simulation. The recorded image is then decimated into an $N \times N$ image using average pooling and fed to the digital neural network.

Note that one could also implement the average pooling step optically, by using a large pixel size. Such large pixel size can potentially provide benefit for low-light operations, which is important in application in mid/long wave infrared regime. We captured 11,000 images for the training set and 2,000 images for validation. We used the captured validation data set to first calculate the accuracy of the theoretical hybrid neural network, whose accuracies are shown on **Fig. 6.3**(b). The training data set is then used to train a hardware-inthe-loop hybrid ANN with the save topology as the original hybrid ANN. This procedure is done 4 times with different random seeds during training, the validation accuracies of which are recorded. The standard deviation of the validation accuracies is also calculated and shown as the distance between the error bars shown on **Fig. 6.3**(b) in the main text.

6.5.6 The demonstration of benefit of optics in the existing work

While a large body of works exists today on optics-assisted neural network, we argue that none of them showed an advantage over pure digital electronic ANN. Here, our classification of digital electronic ANN encompasses pure software solution (running on a GPU) or an application specific integrated circuit (ASIC) based accelerator. Essentially, as long as we are using CMOS-transistor based hardware, we are classifying them as digital backend. Any nano-electronic solutions, like memristors-based solutions are not included in this analysis. Based on the current works, there are largely three classes of optical neural networks, which are listed as follows.

Integrated photonics + digital: There is an extensive amount of works that rely on integrated photonic based optical operations and digital backends [92, 95, 91, 94, 101, 102]. Most of them depend on some form of arrays of switches, made of either Mach-Zehnder interferometers or ring resonators. Most of these networks are limited in terms of space-bandwidth product, which is the same as the number of waveguides. This essentially comes from the lack of dimensionality, as in integrated photonics, we effectively have one spatial dimension, making the space-bandwidth product to be $N \sim A/\lambda, A$ being the dimension of the chip and λ being the optical wavelength. While the number of waveguides currently is much smaller

due to technical limitations, integrated photonics is fundamentally limited by the achievable space-bandwidth product. Even with wavelength division multiplexing (WDM), the dimension of input vector remains small. As such, WDM can only provide a linear scaling of N. This limited space-bandwidth product necessitates time-domain multiplexing to send the data in batches, which also requires combining the data in the backend. As such most of these works did not report any excess electrical power and latency originating from the control circuit for the multiplexing. Moreover, some of the solved problems do not have an electronic benchmark and thus it is unclear if any advantage in the system level is achieved. Most of the works essentially demonstrated a similar classification accuracy against a digital ANN, but the power and latency of an application specific digital IC were not calculated.

Free space + Digital: Using free-space optics (either spatial light modulator or digital micromirror devices), researchers can achieve much larger-space bandwidth product [97, 98, 103, 104, 105, 106, 107]. Thanks to the two dimensions, the space-bandwidth product scales as $N \sim \left(\frac{A}{\lambda}\right)^2$, A being the aperture of the optics and λ being the optical wavelength. This is a much more favorable scaling than the linear scaling in integrated photonics (even with WDM). But most of the reported works neglect the power and latency coming from the conversion between the optics and electronics. Additionally, the use of a spatial light modulator can add substantial amount of power. Comparing these power numbers with a GPU is also unfair, as the GPU is designed to be a solution for many different problems, and thus have large redundancy. For a given problem, one can optimize and design an application specific integrated circuit, with pruning/ XNOR-operations and can require much lower power.

All-optical: Finally, there are a few demonstrations (both free-space and integrated photonics [95, 97, 103, 108] of all-optical neural network, which can be implemented without a digital backend. Some of them are fully linear (without any nonlinearity) and as such they cannot be used for complicated datasets. Moreover, even there, multiple layers of optics are generally used, each with additional scattering losses. Thus, the input power might be high, and they also require a laser to encode the electronic information. It is unclear what

kind of wall-plug efficiency has been achieved in these cases. Thus, the true power is not reported. Others employed nonlinearity either based on saturable absorption, electromagnetic induced transparency in cold atoms, or optoelectronic nonlinearity (image intensifier or optically induced electro-optic nonlinearity). While some of them did demonstrate operation with MNIST data set, the power consumption of all the possible sources is not accounted for, for example, no laser powers are accounted for to cool down the atom. As such, a true estimation of system level power/ latency is missing in all these works, and thus the claims of benefit over purely digital ANN were not established.

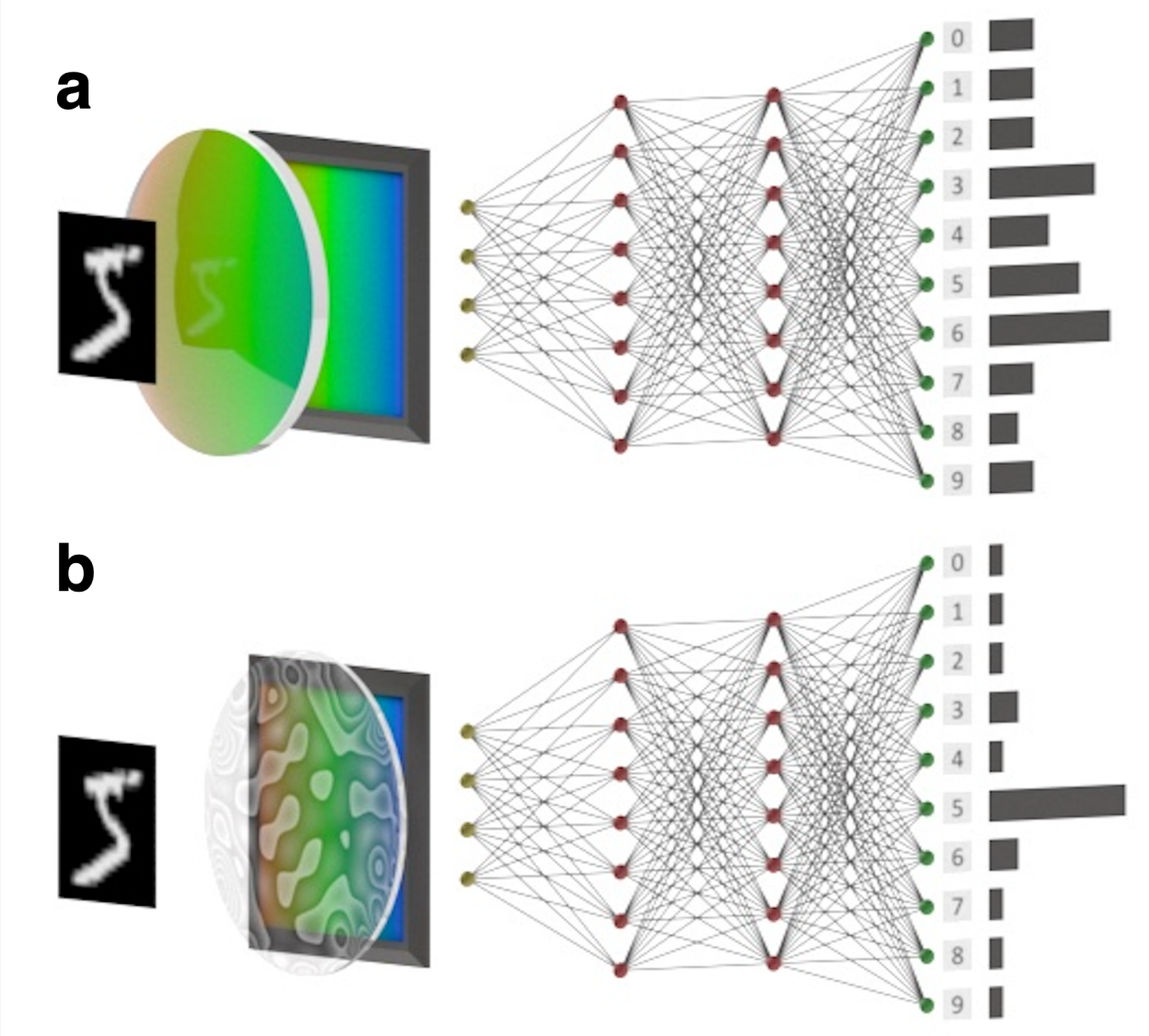

Figure 6.1: Schematic of the optical encoder and pure digital neural network. (a) Purely digital artificial neural networks operate on captured images in a lensed sensor. (b) Instead of using a lens, a designed optics can perform additional linear operations on the captured data. In both cases, the power and the latency of the sensor are the same. Using the digital computational backend with the same resources (number of layers and neurons), we ensure the same power and latency, both of which monotonically scale with the dimensionality of the input data to the digital backend.

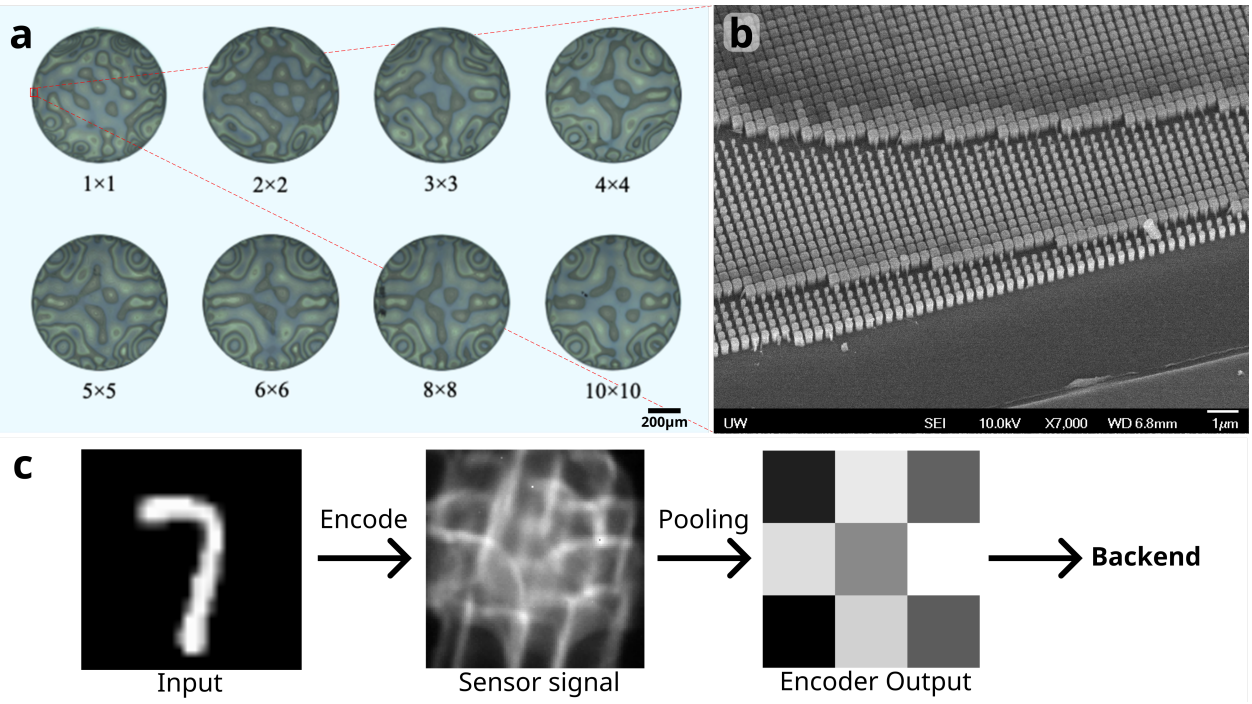

Figure 6.2: Fabrication and characterization of the meta-optical encoder: (a) Optical microscope images of the meta-optical encoders for different input sizes. (b) Scanning Electron Microscope (SEM) image of the optical encoder, region denoted by the red box on device 1×1. (c) The experimental input, sensor signal, and output of the meta-optical encoder.

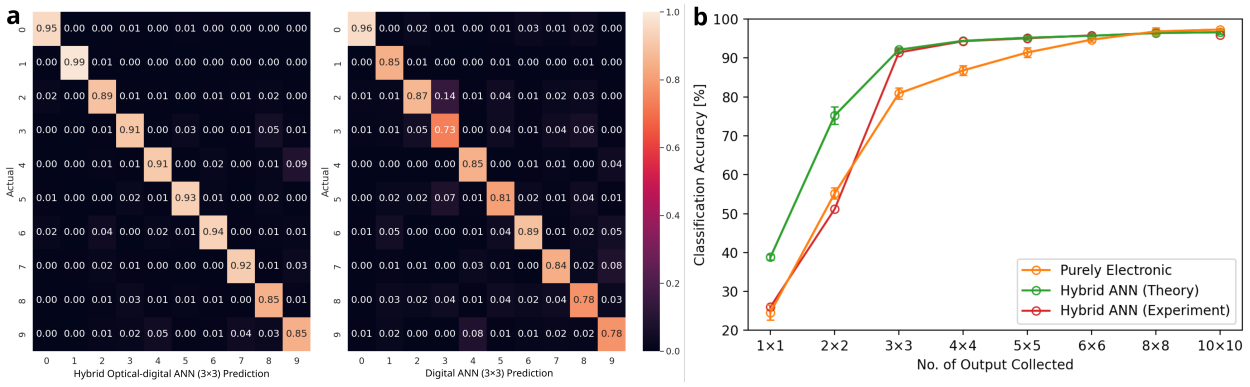

Figure 6.3: Performance comparison of the digital and hybrid ANN. (a) Confusion matrices comparing the experimental performances of the hybrid optical-digital against the pure digital ANNs for the case of $N = 3 \times 3$. (b) Validation classification accuracies of the purely electronic and hybrid optical-electronic ANNs. The error bar is shown to represent the range of one standard deviation.

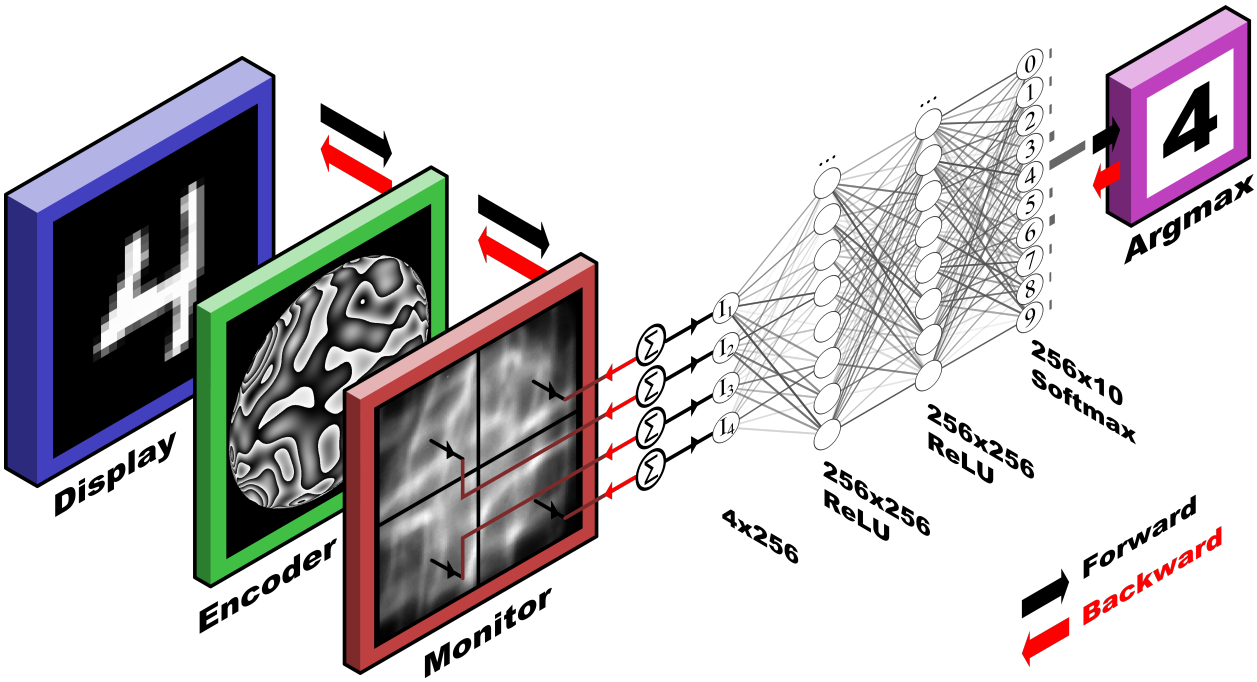

Figure 6.4: The hybrid optical-digital neural network is designed iteratively using an end-to-end differentiable pipeline. Each iteration consists of a forward computation of the loss, and a backward propagation of the loss. An example of the 2×2 encoding is shown here.

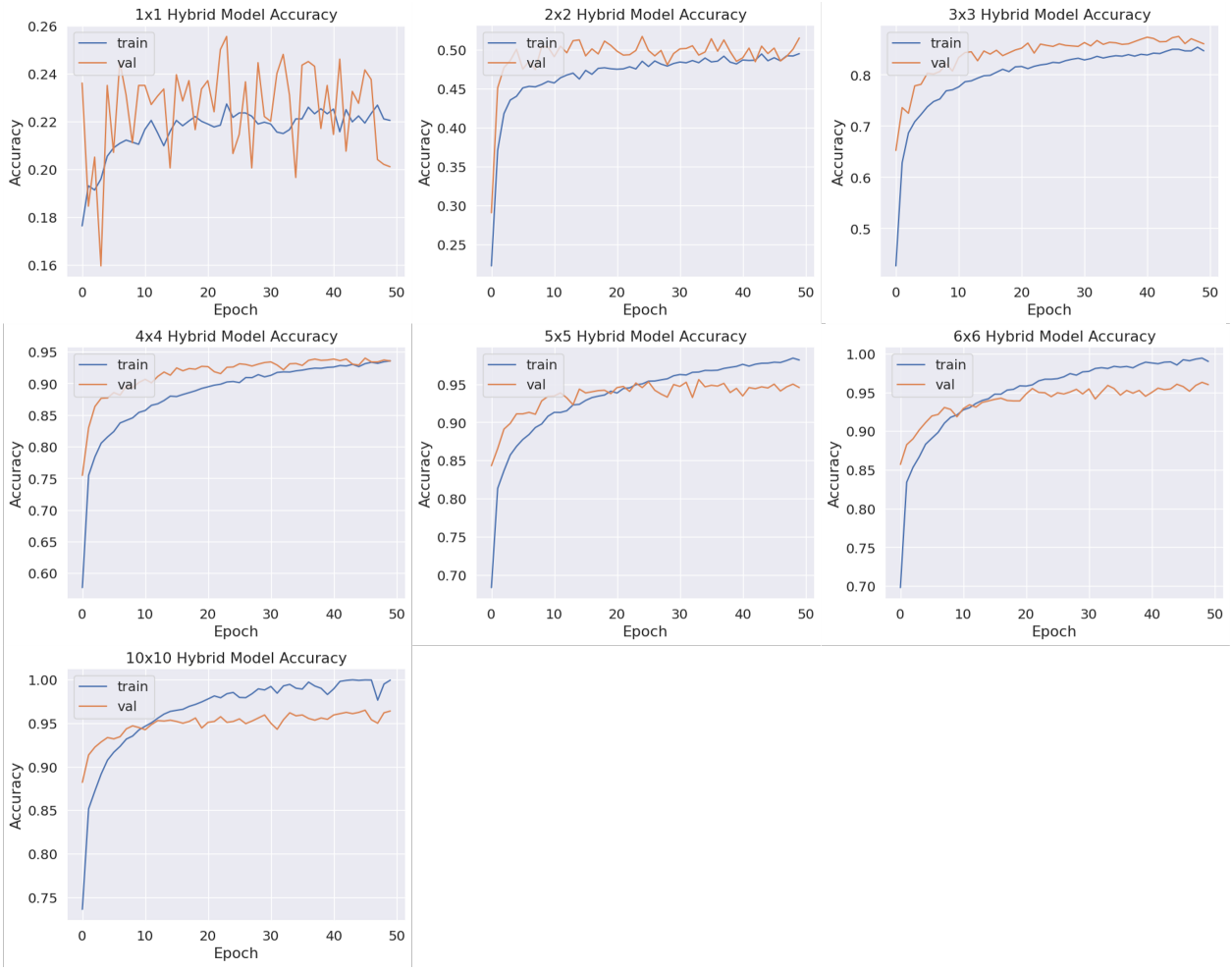

Figure 6.5: The training (train) and validation (val) accuracies of the hybrid artificial neural network classification.

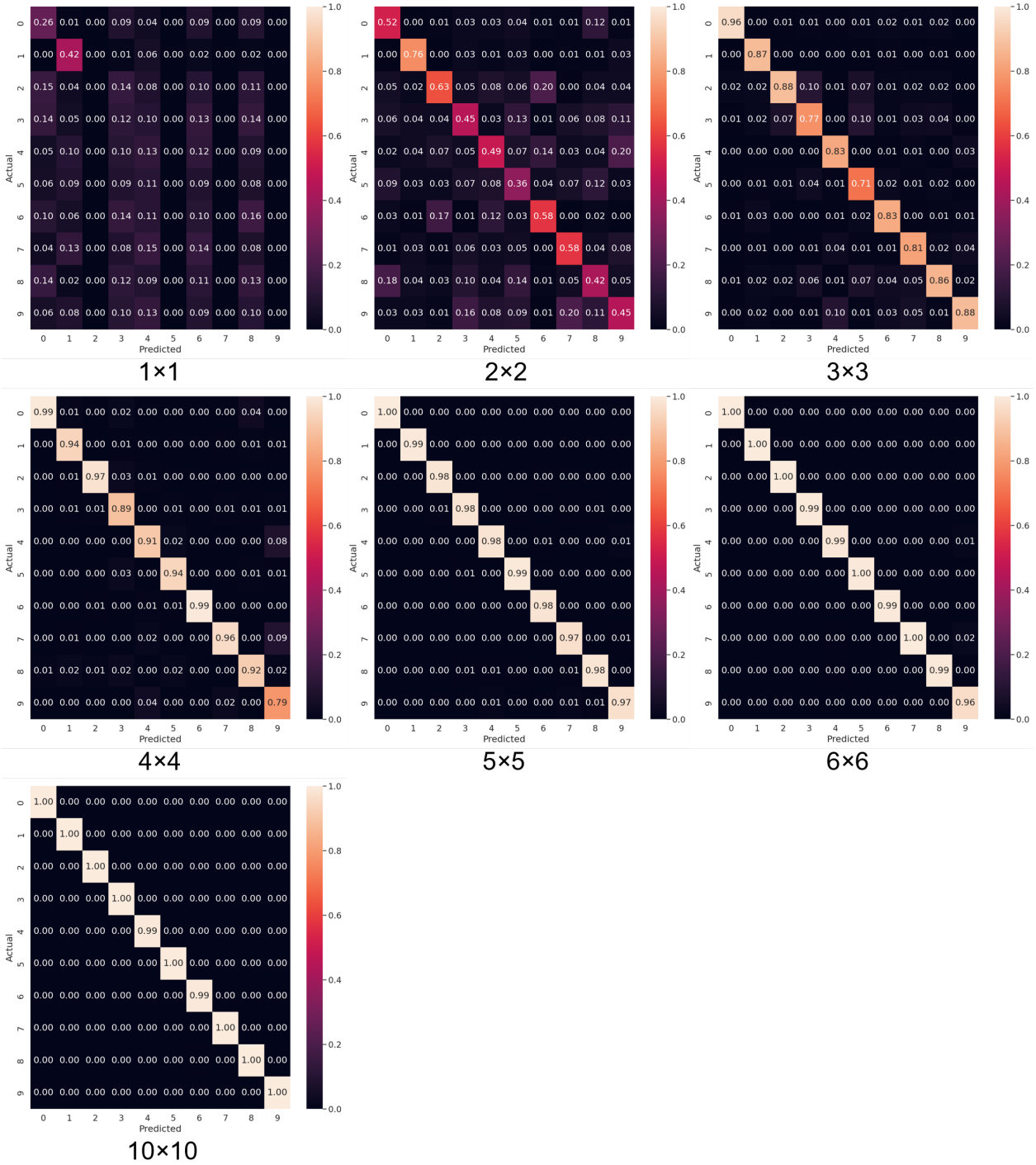

Figure 6.6: The training confusion matrices of the hybrid artificial neural networks.

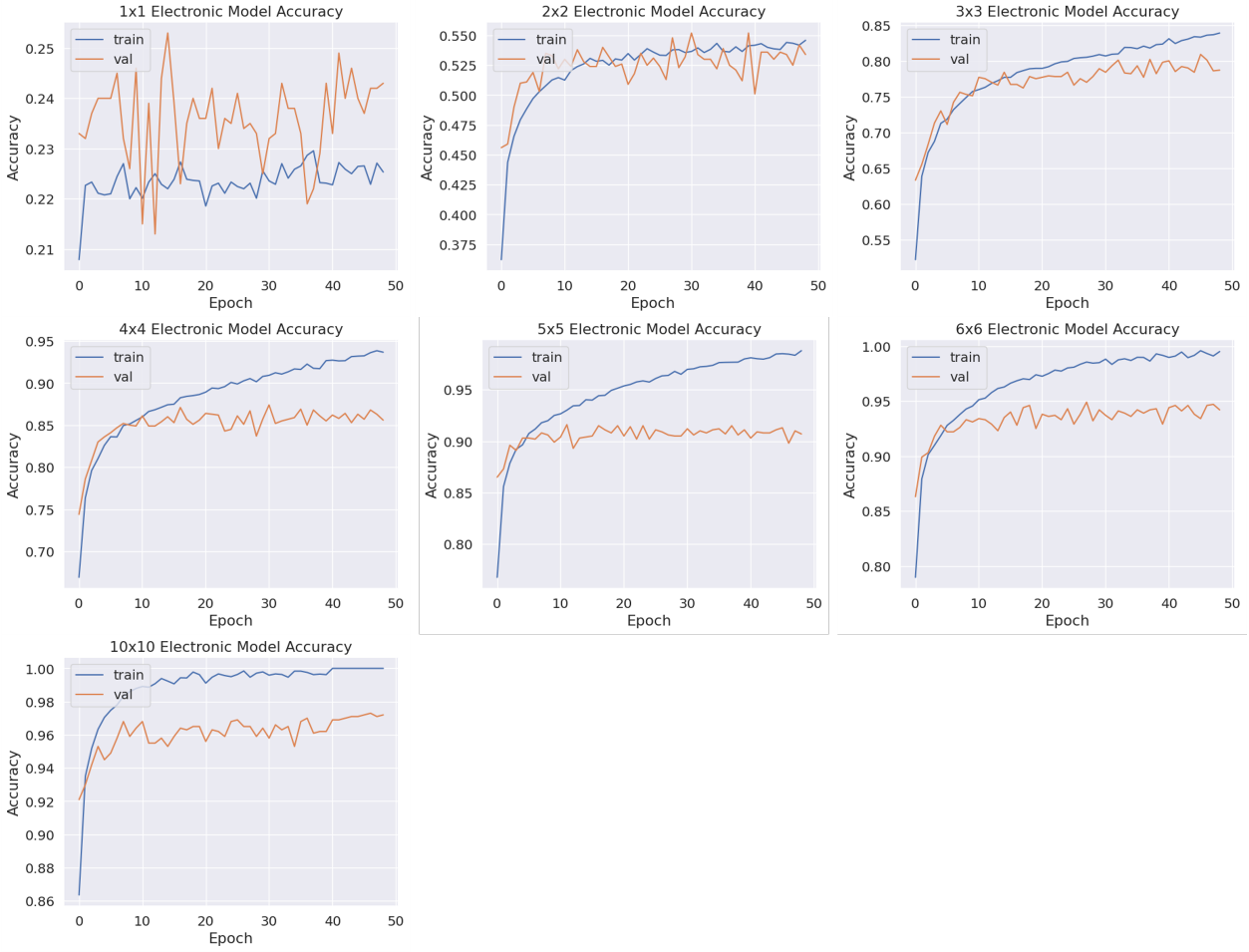

Figure 6.7: The training (train) and validation (val) accuracies of the purely electronic neural networks. We note that, to ensure the network is well trained, we started with ~ 5 layers, and reduced the layers and number of neurons and trained with many different inputs. Finally, we achieved a high classification accuracy (97%) only with two hidden layers. We emphasize that it is important to have a good training of the purely digital ANN, without which we can draw a wrong conclusion on the photonic advantage. We suspect that many of the reported optical neural network works have compared works with a poorly trained digital ANN, showing an improved classification accuracy.

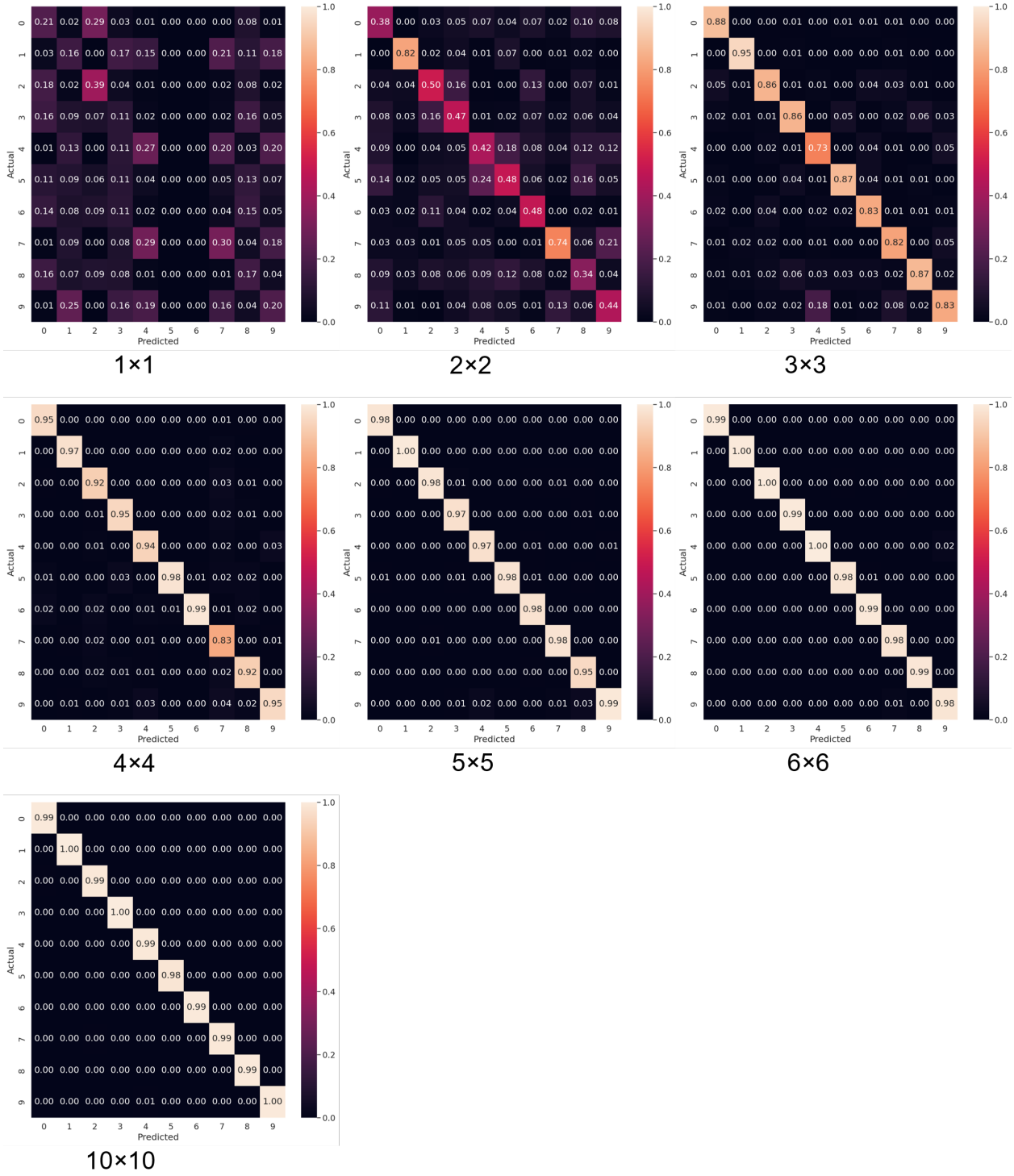

Figure 6.8: The training confusion matrices of the purely digital neural networks.

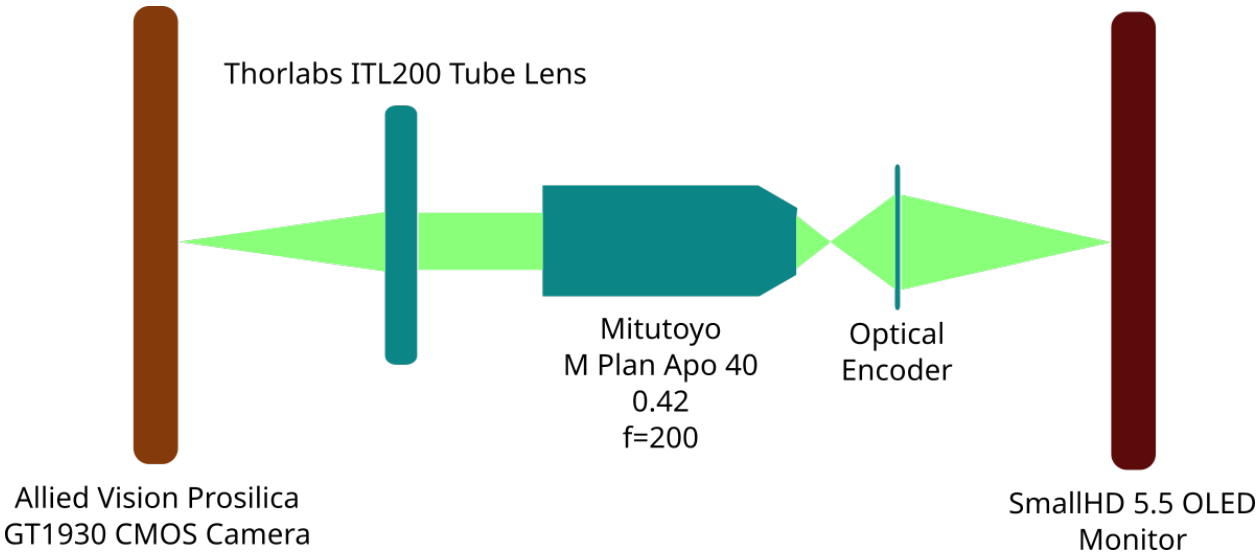

Figure 6.9: Schematic of the experimental setup.

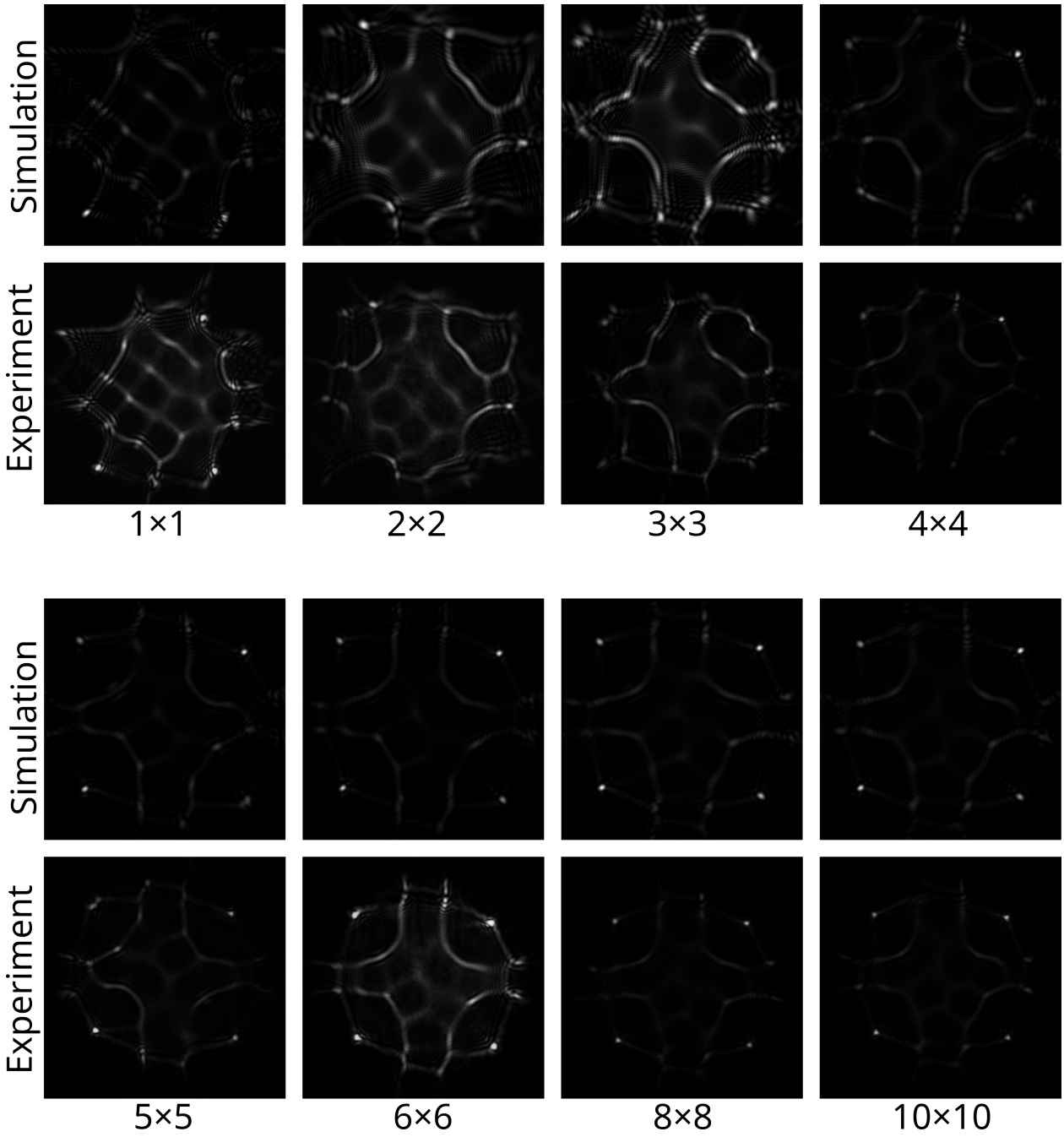

Figure 6.10: Simulation and experimental PSFs of the optical encoders.

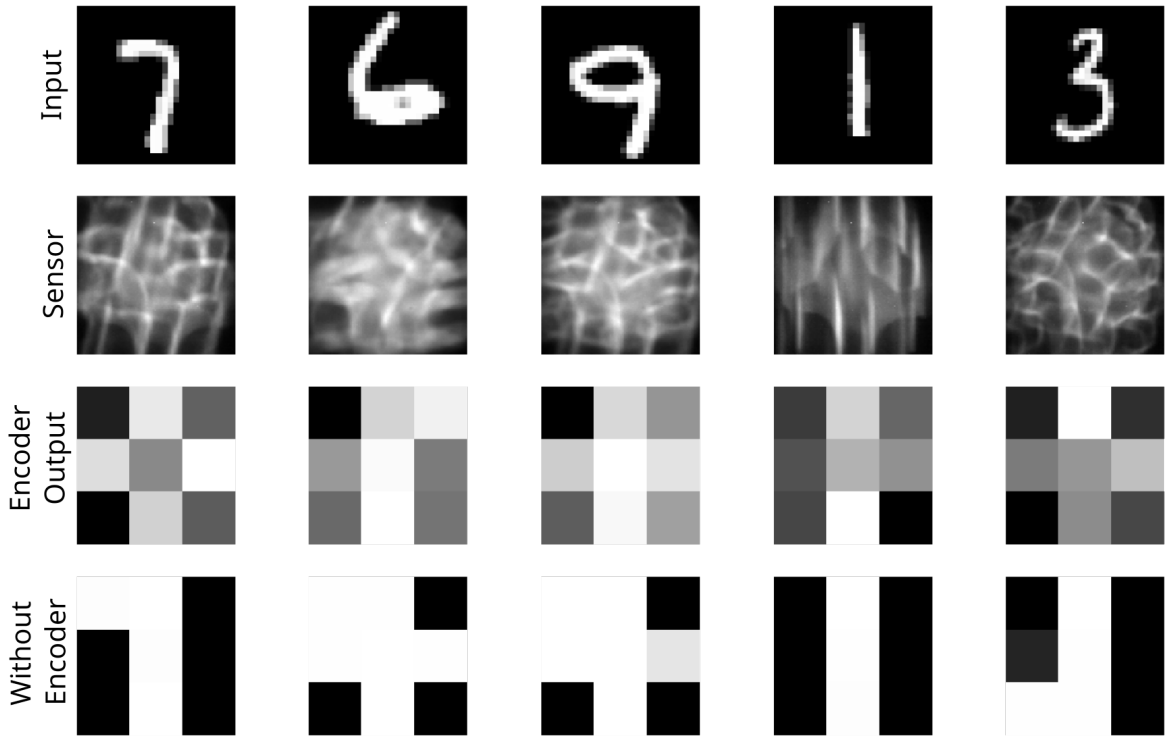

Figure 6.11: Experimental results of the displayed digits, captures, and output of the 3×3 encoder.

Chapter 7

Closing Thoughts

7.1 Conlusion

The aim of this thesis was to expand the understanding and potential of meta-optics for both imaging and computing applications, a goal which was explored through a range of investigations, methodologies, and experiments. The research focused on enhancing the functionalities of metasurfaces, applying them in innovative ways, and addressing several challenges related to these novel devices.

Starting with an introduction of the motivations and outlines of this research in Chapter **§ 1**, we delved into the specific mechanisms of meta-optics for imaging and computing. Then, in Chapter **§ 2**, the focus was on the background of metasurfaces, illuminating the unique challenges and opportunities provided by chromatic aberrations in metalens, dispersion engineering, multiwavelength polychromatic metalens, and computational imaging. Chapter **§ 3** introduced our novel approach towards achieving broadband imaging through the extended depth of focus imaging system with the design and fabrication of unique meta-optics devices. Following this, we explored forward and inverse design principles for Long-Wavelength Infrared (LWIR) Meta-optics in Chapters **§ 4** and **§ 5** respectively. In these chapters, we presented our methodologies, experimental results, discussions, and concluded with experimental details of the design, fabrication, characterization, and imaging process

of our meta-optics devices. In Chapter **§ 6**, we expanded our research scope to the emerging field of optical computation with meta-optics, where we presented an overview, introduction, results, discussion, and methods for designing Hybrid Optical-digital Neural Networks (NN), purely digital neural networks, and meta-optics. Chapter **§ 7** looked at the broader implications and potential future directions for meta-optical imaging and computing. This included the need for standardizing the benchmarking of image quality, opportunities for larger aperture, efficiency considerations, real-world imaging (imaging in the wild), video capture, and computational metaoptics.

The journey of this thesis has underscored the potential of meta-optics in transforming the fields of imaging and computing. However, it is important to note that this is an ongoing journey. Although substantial progress has been made, there remains much to be learned and discovered about meta-optics. As we continue to explore these devices' potential, it is crucial that we also consider the broader implications of their use, from practical challenges to ethical considerations. In conclusion, meta-optics represent an exciting frontier in the domain of imaging and computing. This thesis has contributed to that frontier by offering novel insights and innovations, which we hope will serve as stepping stones towards more advanced meta-optical applications in the future. Future research should aim to build on these findings, while also continuing to push the boundaries of what is possible with meta-optics. Ultimately, the full potential of meta-optics can only be realized through continued research, innovation, and experimentation. The path forward is challenging, but also filled with incredible potential.

7.2 Outlook on Meta-optical Imaging[1]

We have thus far covered various approaches to achieve broadband imaging. Unfortunately, the captured image quality often is poor, even though the focal length remains the same over

[1]The content of this chapter is based on the following publication:
Luocheng Huang, Shane Colburn, Alan Zhan, Arka Majumdar, "Full-Color Metaoptical Imaging in Visible Light." Adv. Photonics Res., 3: 2100265

a large optical bandwidth. **Fig. 7.1** shows published full-color images using metaoptics over the years, which, while impressive in their own right, are clearly inferior in quality to those produced by widely available commercial cameras based on refractive lenses. While some recent images are complex and comparable with those of refractive lens-captured images, they require computational reconstruction. The reason for this apparent discrepancy between demonstrated focusing behavior and actual imaging comes from noisy capture, poor MTF, and a limited field of view. Table 1 shows performance of various metaoptics reported so far for full-color imaging. In this section, we outline several challenges the metaoptics community must confront and possible directions for future developments.

7.2.1 Standardizing Benchmarking of Image Quality

We need a standardized metric to assess image quality using metaoptics between different methods going forward. While the information capacity of an imaging system is well captured by the MTF and SNR, these metrics alone are also known to not provide sufficient criteria for consumer photography, where almost always some computational amelioration of the images is performed, as extensively written by researchers from the industry [109]. Unfortunately, these metrics used for consumer photography, such as SSIM or peak signal-to-noise ratios, require imaging with a fixed dataset for fair comparison. This will require determining a fixed set of images to compare the performance for various different approaches in the metaoptics imaging community.

7.2.2 Larger Aperture

Any optical imaging system will need to collect a certain number of photons to be useful, and this requires a large aperture. For example, current cameras in smartphones have an aperture of at least 2 mm. Most achromatic metaoptics reported so far have sub-mm apertures. The increase in aperture from the mm to cm scale has four primary challenges. First, there are some fundamental limits in achieving broadband operation in an ultrathin form factor, as described by others. [11] With a computational back end, there may exist a route to

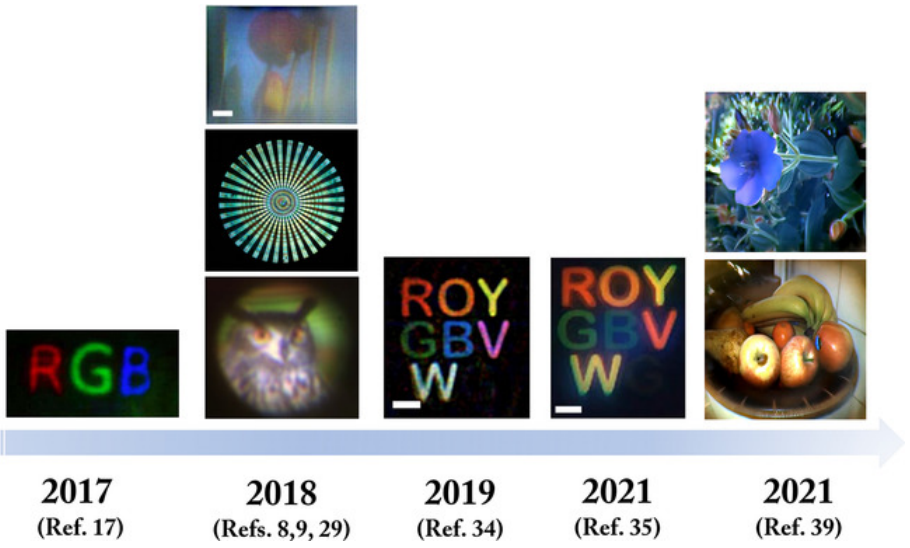

Figure 7.1: Full-color imaging using single metaoptics over the years shows progress, but the image quality is still worse than a simple refractive lens. While some recent results may be comparable with compound refractive lenses, this require a large amount of computation. While the academic community is focusing on many other applications of metaoptics, we believe not being able to capture high-quality color imaging will significantly limit the commercial application space of metaoptics. Hence, it is imperative to solve this problem, and we outline several possible directions toward that end. Images from left to right [26]. Copyright 2017, Springer Nature Ltd. The three images from bottom to top: [17]. Copyright 2018, Springer Nature Ltd. [18]. Copyright 2018, Springer Nature Ltd. [37]. Copyright 2018, AAAS. [41]. Copyright 2019, Optical Society of America. [42]. Copyright 2021, De Gruyter. [46]. Copyright 2021, Springer Nature Ltd.

circumvent this, although this has not been explicitly proven. Second, Seidel aberrations of a lens increase with aperture [110, 111]. This makes capturing high-quality images at larger aperture difficult, especially while maintaining a large field of view. A computational back end may also be of benefit here, while another option is to use multiple metaoptics as a means to increase the field of view without introducing large aberrations [15, 112, 113, 114, 115].

Table 7.1: Performance metrics for different full-color metaoptical imaging

Method	Aper. [μm]	$f/\#$	# of scatterers	Pitch [nm]	FOV	Bandwidth	Polarization [nm]	Design strategy
Tseng et al. [46]	500	2	1.6×10^6	350	40°	400 – 700	Insensitive	Neural design
Ndao et al. [34][a]	20	7.5	2.3×10^3	370	8°	640 – 1200	Insensitive	Phase slope and intercept
Chen et al. [33]	26.4	2.54	3.4×10^3	400	22°	460 – 700	Insensitive	Dispersion engineering
Colburn et al. [37][b]	200	1	1.6×10^5	443	20°	400 – 700	Insensitive	Computational imaging
Chen et al. [17][c]	220	25	2.4×10^5	400	30°	470 – 670	Sensitive	Dispersion engineering
Wang et al. [18][d]	50	4.7	5.2×10^4	120	12°	400 – 660	Sensitive	Dispersion engineering
Shrestha et al. [19][a,e]	200	4	N/A	N/A	14°	1200 – 1650	Insensitive	Dispersion engineering
Khorasaninejad et al. [32][a]	200	2.425	1.3×10^5	480	23°	490 – 550	Sensitive	Dispersion engineering
Wang et al. [116][a]	55.55	1.8	8.0×10^3	550	31°	1200 – 1680	Sensitive	Dispersion engineering
Arbabi et al. [16][a]	240	3.54	8.3×10^4	740	16°	1450 – 1590	Sensitive	Dispersion engineering

[a] FOV was not reported, so we estimated FOV assuming that aperture diameter equals sensor size;

[b] FOV is determined from off-axis simulations, which were taken up to $\pm 10°$;

[c] FOV is determined for a Strehl ratio of 0.8 ;

[d] In this design, the reported pitch corresponds to the side length in a hexagonal lattice;

[e] Reported for the largest aperture among the designs proposed in this work.

Third, metaoptics presents a challenging multiscale electromagnetics problem: while each scatterer needs to be modeled rigorously using Maxwell's equations to maintain accuracy, a full metaoptic must be described in terms of ray or Fourier optics to be computationally feasible. Several recent Table 7.1. Performance metrics for different full-color metaoptical imaging.

works based on transfer matrix [117, 118, 119], scatterer proxy functions [46, 109, 110, 111, 15, 112, 113, 114, 115, 117, 118, 119, 120], or deep learning [121, 122] have shown large acceleration in the forward simulation and could be beneficial for designing achromatic metaoptics. Fourth, the fabrication of large-aperture metaoptics is difficult. Most current visible metaoptics are fabricated using electron beam lithography, which is not conducive to scalable manufacturing. While the dimensions of simple structures, such as cylindrical or square pillars, are within reach of deep-ultraviolet or immersion lithography systems [123, 124, 125] complicated metamolecules will be difficult to fabricate. Moreover, the lithogra-

phy systems are well suited to fabricate apertures at the die size, which is around 2.5 cm. Anything larger than that, which may find applications for space-based optics, will require step-and-stitch approaches combining multiple dies and can be very expensive. Approaches involving synthetic apertures could be beneficial for such large-scale metaoptics [126]. In fact, arrays of achromatic metaoptics have already been reported for lightfield imaging [127]. Another promising direction could be nanoimprint lithography, as recently demonstrated by NIL Technology. However, the index of the resist typically used in nanoimprint is low and may affect the efficiency of the metaoptics. Granted, using the resist as a mask for etching could potentially alleviate the problem, as recently demonstrated by NIL technologies for near-infrared lenses. Finally, metaoptics in visible wavelength requires a thin film of thickness $\approx 500 - 700$ nm. Creating high-quality films with such thickness could be a challenge. Current foundry services in semiconductor photonics, which are primarily centered around integrated photonics industries, generally work with $\approx 200 - 400$ nm-thick films. As such, highvolume manufacturing of metaoptics will remain an important milestone to achieve for any application and not just for full-color imaging.

7.2.3 Efficiency

Efficiency is an important metric for optical elements. Traditional optical elements are primarily characterized by transmission efficiency and Strehl ratio. In the metaoptical community, researchers also report focusing efficiency, which is defined as the ratio of the power inside a circle around the focal spot with a radius of three times the full-width-half maxima of the spot size and the total power in the focal plane. This metric is connected to the MTF and Strehl ratio. While high transmission efficiency (>90%) can be achieved using a transparent material in the visible range, such as SiN, GaN, or TiO_2, most full-color metaoptics suffer from either low focusing efficiency or low Strehl ratio, especially when implementing high NA or fast lenses. Multiwavelength metaoptics in particular, which exploit multiplexed structures in a unit cell, significantly suffer from low efficiency ($\approx 22 - 32\%$ in one wavelength) [15]. With dispersion engineering, the efficiency still remains relatively poor ($\approx 40\%$)

[17] Computational imaging-based approaches for full-color imaging seek to ensure all colors and angles reach the sensor in an identical manner, which can result in an optic with a lower Strehl ratio. However, the use of deconvolution and denoising methods in computational imaging can restore high-frequency image components with a well-designed MTF. Such computational approaches still require high-efficiency metaoptics to ensure adequate SNR for processing.

7.2.4 Imaging in the Wild

For a large number of practical applications, the imaging needs to be performed using ambient light. Unfortunately, in most works, lasers or light-emitting diodes, including organic light emitting diode (OLED) displays, were used in a controlled lab environment, often used in conjunction with an optical relay. In near- and long-wave-infrared wavelength range, some works did perform imaging using just a single metaoptic [15, 128, 73], but they used a filter to limit the wavelength range of the incident light. Similar filtered imaging methods using metaoptic integrated cameras have been reported recently for green light [129]. To the best of our knowledge, there has been no demonstration of full-color imaging using a metaoptics in the wild. This may need a departure from the "device engineering" mindset of researchers working in the metaoptics field. Without improving just metaoptics, a system-level performance needs to be assessed. Thankfully, there are several startups working on metaoptics, such as Metalenz, Leia, Lumotive, and Tunoptix. As such, we hope that full-color imaging systems involving metaoptics will be commercialized soon.

7.2.5 Video Capture

While full-color imaging is still an important achievement, many applications require video imaging with a frame rate of 33 frames s^{-1}. High-quality full-color video imaging has not yet been demonstrated using metaoptics. This problem is related to the use of relay optics, requiring longer exposure times. The need for relay optics is intimately related to the aperture as well, especially because most commercial sensors are covered with a glass slip

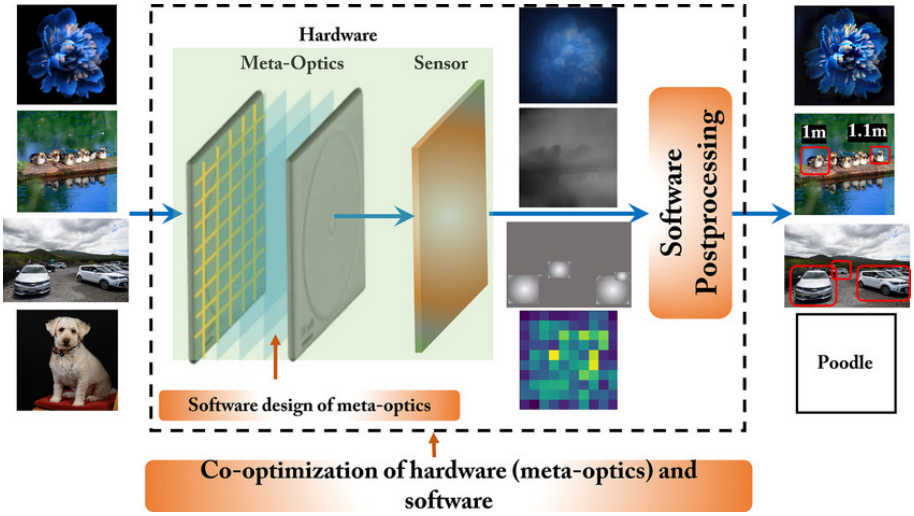

Figure 7.2: Outlook for co-optimized hardware-software platform: by optimizing a single metaoptic or a stack of metaoptics, along with a computational backend, a dramatic reduction in size, weight, power, and latency of image sensors can be achieved. Such sensors can either capture aesthetically pleasing images, as reviewed in this article, or capture more information from the scene, such as depth or spectral. We envision that some of them can even perform computation for object detection or scene understanding.

of thickness $\approx$ 1 mm. Thus to ensure small f-number, the aperture at least needs to be in the mm-scale. Achieving such apertures is difficult using dispersion engineering. While computational reconstruction may allow larger aperture, current computational times may not allow real-time video capture. Pruning of the computationally expensive algorithms or hardware acceleration techniques needs to be applied to solve this problem.

7.2.6 Computational Metaoptics

Finally, the computational framework developed in **§ 5.3.1** can be used for many other sensing and computational tasks (**Fig. 7.2**). We believe such co-optimized hardware and software can find applications in depth sensing, object detection, or for optical information process-

ing. One particularly promising direction will be optical neural networks [99, 130]. While all-optical diffractive neural networks have already been reported [131], a computational back end can potentially allow performing inference using incoherent light [132, 104].

7.3 Outlook on Meta-optical Computing

The recent development of meta-optics has opened up new possibilities for optics-assisted ANNs. Meta-optics is a subfield of optics that deals with the design and fabrication of optical components with subwavelength features, which can manipulate light in ways that are not possible with conventional optics. This has led to the development of all-silicon metalenses that can operate in the long-wavelength infrared regime under ambient thermal radiation [73]. Such advances in meta-optics can potentially overcome the intrinsic problems of optics, such as the large size and poor tolerance to misalignment of optical components, and limited space bandwidth product of spatial light modulators.

However, to fully realize the potential of meta-optics for optical computation, there is a need for co-optimization of the optical and digital components of the ANN. This can be achieved by developing end-to-end design frameworks that can optimize the optical and digital components simultaneously. Such frameworks can take into account the nonlinearity and signal regeneration requirements of the ANN, and optimize the optical and digital components to minimize the power and latency of the system. In addition, there is a need for further research to identify the applications where optics can provide an advantage over digital ANNs. Our results indicate that optics can be advantageous over digital ANN in applications where the overall performance of the ANN can be relaxed to prioritize lower power and latency. Such applications can include low-power, real-time image and video processing, where the accuracy requirements are not as stringent as in high-performance computing applications.

In conclusion, the development of meta-optics has opened up new possibilities for optics-assisted ANNs. However, to fully realize the potential of meta-optics for optical computation, there is a need for co-optimization of the optical and digital components of the ANN, and

further research to identify the applications where optics can provide an advantage over digital ANNs.

Bibliography

[1] Patrick Robert Gill et al. “A microscale camera using direct Fourier-domain scene capture”. In: *Optics letters* 36.15 (2011), pp. 2949–2951.

[2] Vivek Boominathan et al. “Lensless imaging: A computational renaissance”. In: *IEEE Signal Processing Magazine* 33.5 (2016), pp. 23–35.

[3] Jacob Engelberg and Uriel Levy. “The advantages of metalenses over diffractive lenses”. In: *Nature communications* 11.1 (2020), p. 1991.

[4] Sourangsu Banerji et al. “Imaging with flat optics: metalenses or diffractive lenses?” In: *Optica* 6.6 (2019), pp. 805–810.

[5] Timo Stolt et al. “Multiply-resonant second-harmonic generation using surface lattice resonances in aluminum metasurfaces”. In: *Optics Express* 30.3 (2022), pp. 3620–3631.

[6] Moysey Brio et al. “Scattering of a short electromagnetic pulse from a Lorentz–Duffing film: theoretical and numerical analysis”. In: *Wave Motion* 89 (2019), pp. 43–56.

[7] Tengfeng Zhu et al. “Plasmonic computing of spatial differentiation”. In: *Nature communications* 8.1 (2017), p. 15391.

[8] Yuan Meng et al. “Optical meta-waveguides for integrated photonics and beyond”. In: *Light: Science & Applications* 10.1 (2021), p. 235.

[9] Sajjad AbdollahRamezani et al. “Analog computing using graphene-based metalines”. In: *Optics letters* 40.22 (2015), pp. 5239–5242.

[10] Alexander S Berestennikov et al. “Active meta-optics and nanophotonics with halide perovskites”. In: *Applied Physics Reviews* 6.3 (2019).

[11] Sajjad Abdollahramezani, Omid Hemmatyar, and Ali Adibi. “Meta-optics for spatial optical analog computing”. In: *Nanophotonics* 9.13 (2020), pp. 4075–4095.

[12] Weiwei Fu et al. “Ultracompact meta-imagers for arbitrary all-optical convolution”. In: *Light: Science & Applications* 11.1 (2022), p. 62.

[13] Zhongwei Jin et al. “Complex inverse design of meta-optics by segmented hierarchical evolutionary algorithm”. In: *ACS nano* 13.1 (2019), pp. 821–829.

[14] Mu Ku Chen et al. “Artificial intelligence in meta-optics”. In: *Chemical Reviews* 122.19 (2022), pp. 15356–15413.

[15] Amir Arbabi et al. “Miniature optical planar camera based on a wide-angle metasurface doublet corrected for monochromatic aberrations”. In: *Nature communications* 7.1 (2016), p. 13682.

[16] Stewart D Jenkins et al. “Strong interactions and subradiance in disordered metamaterials”. In: *Physical Review B* 98.24 (2018), p. 245136.

[17] Wei Ting Chen et al. “Broadband achromatic metasurface-refractive optics”. In: *Nano letters* 18.12 (2018), pp. 7801–7808.

[18] Shuming Wang et al. “A broadband achromatic metalens in the visible”. In: *Nature nanotechnology* 13.3 (2018), pp. 227–232.

[19] Sajan Shrestha et al. “Broadband achromatic dielectric metalenses”. In: *Light: Science & Applications* 7.1 (2018), p. 85.

[20] Federico Presutti and Francesco Monticone. “Focusing on bandwidth: achromatic metalens limits”. In: *Optica* 7.6 (2020), pp. 624–631.

[21] Comrun Yousefzadeh et al. ""Achromatic limits" of Pancharatnam phase lenses". In: *Applied Optics* 57.5 (2018), pp. 1151–1158.

[22] Francesco Aieta et al. "Multiwavelength achromatic metasurfaces by dispersive phase compensation". In: *Science* 347.6228 (2015), pp. 1342–1345.

[23] Mohammadreza Khorasaninejad et al. "Achromatic metasurface lens at telecommunication wavelengths". In: *Nano letters* 15.8 (2015), pp. 5358–5362.

[24] Ehsan Arbabi et al. "Multiwavelength metasurfaces through spatial multiplexing". In: *Scientific reports* 6.1 (2016), p. 32803.

[25] Zhaoyi Li et al. "Inverse design enables large-scale high-performance meta-optics reshaping virtual reality". In: *Nature communications* 13.1 (2022), p. 2409.

[26] Ori Avayu et al. "Composite functional metasurfaces for multispectral achromatic optics". In: *Nature communications* 8.1 (2017), p. 14992.

[27] You Zhou et al. "Multilayer noninteracting dielectric metasurfaces for multiwavelength metaoptics". In: *Nano letters* 18.12 (2018), pp. 7529–7537.

[28] Zhaoyi Li et al. "Meta-optics achieves RGB-achromatic focusing for virtual reality". In: *Science Advances* 7.5 (2021), eabe4458.

[29] Tom D Milster et al. "Multiple-order diffractive engineered surface lenses". In: *Applied optics* 59.26 (2020), pp. 7900–7906.

[30] Donald W Sweeney and Gary E Sommargren. "Harmonic diffractive lenses". In: *Applied Optics* 34.14 (1995), pp. 2469–2475.

[31] Dean Faklis and G Michael Morris. "Spectral properties of multiorder diffractive lenses". In: *Applied Optics* 34.14 (1995), pp. 2462–2468.

[32] Mohammadreza Khorasaninejad et al. "Achromatic metalens over 60 nm bandwidth in the visible and metalens with reverse chromatic dispersion". In: *Nano letters* 17.3 (2017), pp. 1819–1824.

[33] Wei Ting Chen et al. "A broadband achromatic polarization-insensitive metalens consisting of anisotropic nanostructures". In: *Nature communications* 10.1 (2019), p. 355.

[34] Abdoulaye Ndao et al. "Octave bandwidth photonic fishnet-achromatic-metalens". In: *Nature communications* 11.1 (2020), p. 3205.

[35] Wenbo Zang et al. "Chromatic dispersion manipulation based on metalenses". In: *Advanced Materials* 32.27 (2020), p. 1904935.

[36] Xiaoying Lin. "Synthesis and Polymer-Mediated Regioselective Self-Assembly of Shaped Plasmonic Nanoparticles". PhD thesis. University of Maryland, College Park, 2021.

[37] Shane Colburn, Alan Zhan, and Arka Majumdar. "Metasurface optics for full-color computational imaging". In: *Science advances* 4.2 (2018), eaar2114.

[38] W Thomas Cathey and Edward R Dowski. "New paradigm for imaging systems". In: *Applied optics* 41.29 (2002), pp. 6080–6092.

[39] Joseph Van Der Gracht et al. "Aspheric optical elements for extended depth-of-field imaging". In: *Novel Optical Systems Design and Optimization*. Vol. 2537. SPIE. 1995, pp. 279–288.

[40] Joseph Van Der Gracht et al. "Broadband behavior of an optical–digital focus-invariant system". In: *Optics Letters* 21.13 (1996), pp. 919–921.

[41] Luocheng Huang et al. "Design and analysis of extended depth of focus metalenses for achromatic computational imaging". In: *Photonics Research* 8.10 (2020), pp. 1613–1623.

[42] Elyas Bayati et al. "Inverse designed metalenses with extended depth of focus". In: *ACS photonics* 7.4 (2020), pp. 873–878.

[43] David G Stork and M Dirk Robinson. "Theoretical foundations for joint digital-optical analysis of electro-optical imaging systems". In: *Applied Optics* 47.10 (2008), B64–B75.

[44] Vincent Sitzmann et al. "End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging". In: *ACM Transactions on Graphics (TOG)* 37.4 (2018), pp. 1–13.

[45] Zin Lin et al. "End-to-end nanophotonic inverse design for imaging and polarimetry". In: *Nanophotonics* 10.3 (2020), pp. 1177–1187.

[46] Ethan Tseng et al. "Neural nano-optics for high-quality thin lens imaging". In: *Nature communications* 12.1 (2021), p. 6493.

[47] Ojeda-Castaneda J, Landgrave J, and Escamilla H. "Annular phase-only mask for high focal depth". In: *Opt. Lett* 30 (13 July 2005), p. 1647. DOI: 10.1364/OL.30.001647.

[48] Dowski E and Cathey W. "Extended depth of field through wave-front coding". In: 34 (1995), pp. 1859–1866.

[49] Chi W and George N. "Electronic imaging using a logarithmic asphere". In: *Opt. Lett* 26 (12 2001), pp. 875–877.

[50] Zhai Z et al. "Extended depth of field through an axicon". In: *J. Mod. Opt* 56 (11 June 2009), pp. 1304–1308. DOI: 10.1080/09500340903082689.

[51] Wach H, Dowski E, and Cathey W. "Control of chromatic focal shift through wave-front coding". In: *Appl. Opt* 37 (23 1998), p. 5359. DOI: 10.1364/ao.37.005359.

[52] Wang Y et al. "Electromagnetic diffraction theory of refractive axicon lenses". In: *J. Opt. Soc. Am. A Opt. Image Sci. Vis* 34 (7 2017), p. 1201.

[53] Patwary N et al. "Experimental validation of a customized phase mask designed to enable efficient computational optical sectioning microscopy through wavefront encoding". In: *Appl. Opt* 56 (9 Mar. 2017), p. D14. DOI: 10.1364/AO.56.000D14.

[54] Alan Zhan et al. "Low-contrast dielectric metasurface optics". In: *ACS photonics* 3.2 (2016), pp. 209–214.

[55] Zhan A et al. "Metasurface Freeform Nanophotonics". In: *Sci. Rep* (Apr. 2017), pp. 1–9. DOI: 10.1038/s41598-017-01908-9.

[56] Orieux F, Giovannelli J.-F, and Rodet T. "Bayesian estimation of regularization and point spread function parameters for Wiener-Hunt deconvolution". In: *J. Opt. Soc. Am. A* 27 (7 July 2010), p. 1593. DOI: 10.1364/JOSAA.27.001593.

[57] Van Der S and Walt. "scikit-image: image processing in Python". In: *PeerJ* 2 (June 2014), e453. DOI: 10.7717/peerj.453.

[58] Nacereddine N, Tabbone S, and Ziou D. "Similarity transformation parameters recovery based on Radon transform". In: *Pattern Recognit* 48 (7 July 2015), pp. 2227–2240. DOI: 10.1016/j.patcog.2015.01.017.

[59] Liu V and Fan S. "S4 : A free electromagnetic solver for layered periodic structures". In: *Comput. Phys. Commun* 183 (10 Oct. 2012), pp. 2233–2244. DOI: 10.1016/j.cpc.2012.04.026.

[60] Osnabrugge G et al. "Generalized optical memory effect". In: *Optica* 4 (8 Aug. 2017), p. 886. DOI: 10.1364/OPTICA.4.000886.

[61] Getreuer P. "Total Variation Deconvolution using Split Bregman". In: *Image Process. Line* 2 (2012), pp. 158–174. DOI: 10.5201/ipol.2012.g-tvdc.

[62] Heide F. "FlexISP: a flexible camera image processing framework". In: *ACM Trans. Graph* 33 (6 Nov. 2014), pp. 1–13. DOI: 10.1145/2661229.2661260.

[63] Raiko Schulz et al. "Thermal imaging for monitoring rolling element bearings". In: *12th International Conference on Quantitative Infrared Thermography, Bordeaux, France*. 2014, pp. 7–11.

[64] E F J Ring and K Ammer. "Infrared thermal imaging in medicine". en. In: *Physiological Measurement* 33.3 (Mar. 2012), R33–R46. ISSN: 0967-3334, 1361-6579. DOI: 10.1088/0967-3334/33/3/R33. URL: https://iopscience.iop.org/article/10.1088/0967-3334/33/3/R33 (visited on 03/30/2021).

[65] Michael Vollmer and Klaus-Peter Möllmann. *Infrared thermal imaging: fundamentals, research and applications*. John Wiley & Sons, 2017.

[66] Monjurul Meem et al. “Broadband lightweight flat lenses for long-wave infrared imaging”. en. In: *Proceedings of the National Academy of Sciences* 116.43 (Oct. 2019), pp. 21375–21378. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1908447116. URL: http://www.pnas.org/lookup/doi/10.1073/pnas.1908447116 (visited on 03/30/2021).

[67] Qingbin Fan et al. “A high numerical aperture, polarization-insensitive metalens for long-wavelength infrared imaging”. In: *Applied Physics Letters* 113.20 (2018), p. 201104.

[68] Orrin Kigner et al. “Monolithic all-Silicon Flat lens for broadband LWIR Imaging”. In: *arXiv preprint arXiv:2103.14939* (2021).

[69] Vittorio Magni, Giulio Cerullo, and Sandro De Silvestri. “High-accuracy fast Hankel transform for optical beam propagation”. In: *JOSA A* 9.11 (1992), pp. 2031–2033.

[70] G.B. Rybicki and A.P. Lightman. *Radiative Processes in Astrophysics*. A Wiley-Interscience publication. Wiley, 1991, p. 22. ISBN: 9780471827597. URL: https://books.google.com/books?id=LtdEjNABMlsC.

[71] Ethan Tseng et al. “Neural Nano-Optics for High-quality Thin Lens Imaging”. In: *arXiv e-prints*, arXiv:2102.11579 (Feb. 2021), arXiv:2102.11579. arXiv: 2102.11579 [physics.optics].

[72] Elyas Bayati et al. “Inverse Designed Extended Depth of Focus Meta-Optics for Broadband Imaging in the Visible”. In: *arXiv preprint arXiv:2105.00160* (2021).

[73] Luocheng Huang et al. “Long wavelength infrared imaging under ambient thermal radiation via an all-silicon metalens”. In: *Optical Materials Express* 11.9 (2021), pp. 2907–2914.

[74] Halil Can Nalbant et al. “Transmission optimized LWIR metalens”. In: *Applied Optics* 61.33 (2022), pp. 9946–9950.

[75] Vishwanath Saragadam et al. *Foveated Thermal Computational Imaging in the Wild Using All-Silicon Meta-Optics*. 2022. DOI: 10.48550/ARXIV.2212.06345. URL: https://arxiv.org/abs/2212.06345.

[76] Sawyer D Campbell et al. "Review of numerical optimization techniques for meta-device design". In: *Optical Materials Express* 9.4 (2019), pp. 1842–1863.

[77] Maksym V. Zhelyeznyakov et al. *Large area optimization of meta-lens via data-free machine learning*. 2022. DOI: 10.48550/ARXIV.2212.10703. URL: https://arxiv.org/abs/2212.10703.

[78] Kyoji Matsushima. "Shifted angular spectrum method for off-axis numerical propagation". In: *Opt. Express* 18.17 (Aug. 2010), pp. 18453–18463. DOI: 10.1364/OE.18.018453. URL: http://opg.optica.org/oe/abstract.cfm?URI=oe-18-17-18453.

[79] Vincent Sitzmann et al. "Implicit neural representations with periodic activation functions". In: *Adv. Neural Info. Processing Systems* (2020).

[80] Vishwanath Saragadam et al. "WIRE: Wavelet Implicit Neural Representations". In: *arXiv preprint arXiv:2301.05187* (2023).

[81] Vishwanath Saragadam et al. "Thermal Image Processing via Physics-Inspired Deep Networks". In: *IEEE Intl. Conf. Computer Vision (ICCV)*. 2021.

[82] Shane Colburn et al. "Broadband transparent and CMOS-compatible flat optics with silicon nitride metasurfaces". In: *Optical Materials Express* 8.8 (2018), pp. 2330–2344.

[83] Yaser S Abu-Mostafa and Demetri Psaltis. "Optical neural computers". In: *Scientific American* 256.3 (1987), pp. 88–95.

[84] Nabil H Farhat et al. "Optical implementation of the Hopfield model". In: *Applied optics* 24.10 (1985), pp. 1469–1475.

[85] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning". In: *nature* 521.7553 (2015), pp. 436–444.

[86] Emma Strubell, Ananya Ganesh, and Andrew McCallum. "Energy and policy considerations for modern deep learning research". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. 09. 2020, pp. 13693–13696.

[87] Neil C Thompson et al. "The computational limits of deep learning". In: *arXiv preprint arXiv:2007.05558* (2020).

[88] Nanfang Yu and Federico Capasso. "Flat optics with designer metasurfaces". In: *Nature materials* 13.2 (2014), pp. 139–150.

[89] Lukas Chrostowski and Michael Hochberg. *Silicon photonics design: from devices to systems*. Cambridge University Press, 2015.

[90] Martijn JR Heck et al. "Hybrid silicon photonic integrated circuit technology". In: *IEEE Journal of Selected Topics in Quantum Electronics* 19.4 (2012), pp. 6100117–6100117.

[91] Yichen Shen et al. "Deep learning with coherent nanophotonic circuits". In: *Nature photonics* 11.7 (2017), pp. 441–446.

[92] Xingyuan Xu et al. "11 TOPS photonic convolutional accelerator for optical neural networks". In: *Nature* 589.7840 (2021), pp. 44–51.

[93] Johannes Feldmann et al. "Parallel convolutional processing using an integrated photonic tensor core". In: *Nature* 589.7840 (2021), pp. 52–58.

[94] Alexander Sludds et al. "Delocalized photonic deep learning on the internet's edge". In: *Science* 378.6617 (2022), pp. 270–276.

[95] Farshid Ashtiani, Alexander J Geers, and Firooz Aflatouni. "An on-chip photonic deep neural network for image classification". In: *Nature* 606.7914 (2022), pp. 501–506.

[96] Hanyu Zheng et al. "Meta-optic accelerators for object classifiers". In: *Science Advances* 8.30 (2022), eabo6410.

[97] Albert Ryou et al. "Free-space optical neural network based on thermal atomic nonlinearity". In: *Photonics Research* 9.4 (2021), B128–B134.

[98] Tianyu Wang et al. "An optical neural network using less than 1 photon per multiplication". In: *Nature Communications* 13.1 (2022), p. 123.

[99] Shane Colburn et al. "Optical frontend for a convolutional neural network". In: *Applied optics* 58.12 (2019), pp. 3179–3186.

[100] Mohammad Rastegari et al. "Xnor-net: Imagenet classification using binary convolutional neural networks". In: *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV*. Springer. 2016, pp. 525–542.

[101] Bhavin J Shastri et al. "Photonics for artificial intelligence and neuromorphic computing". In: *Nature Photonics* 15.2 (2021), pp. 102–114.

[102] Saumil Bandyopadhyay et al. "Single chip photonic deep neural network with accelerated training". In: *arXiv preprint arXiv:2208.01623* (2022).

[103] Tianyu Wang et al. "Image sensing with multilayer nonlinear optical neural networks". In: *Nature Photonics* (2023), pp. 1–8.

[104] Julie Chang et al. "Hybrid optical-electronic convolutional neural networks with optimized diffractive optics for image classification". In: *Scientific reports* 8.1 (2018), p. 12324.

[105] James Spall, Xianxin Guo, and Alexander I Lvovsky. "Hybrid training of optical neural networks". In: *Optica* 9.7 (2022), pp. 803–811.

[106] Deniz Mengu et al. "Analysis of diffractive optical neural networks and their integration with electronic neural networks". In: *IEEE Journal of Selected Topics in Quantum Electronics* 26.1 (2019), pp. 1–14.

[107] Hui Zhang et al. "An optical neural chip for implementing complex-valued neural network". In: *Nature communications* 12.1 (2021), p. 457.

[108] Shishi Lin et al. "Mapping the dark space of chemical reactions with extended nanomole synthesis and MALDI-TOF MS". In: *Science* 361.6402 (2018), eaar6236.

[109] Jonathan B Phillips and Henrik Eliasson. *Camera image quality benchmarking*. John Wiley & Sons, 2018.

[110] Adolf W Lohmann. "Scaling laws for lens systems". In: *Applied optics* 28.23 (1989), pp. 4996–4998.

[111] David J Brady and Nathan Hagen. "Multiscale lens design". In: *Optics express* 17.13 (2009), pp. 10659–10674.

[112] Zhenyu Huang et al. "Achromatic and wide-field metalens in the visible region". In: *Optics Express* 29.9 (2021), pp. 13542–13551.

[113] Dongliang Tang et al. "Achromatic metasurface doublet with a wide incident angle for light focusing". In: *Optics Express* 28.8 (2020), pp. 12209–12218.

[114] Andrew McClung, Mahdad Mansouree, and Amir Arbabi. "At-will chromatic dispersion by prescribing light trajectories with cascaded metasurfaces". In: *Light: Science & Applications* 9.1 (2020), p. 93.

[115] Changhyun Kim, Sun-Je Kim, and Byoungho Lee. "Doublet metalens design for high numerical aperture and simultaneous correction of chromatic and monochromatic aberrations". In: *Optics Express* 28.12 (2020), pp. 18059–18076.

[116] Shuming Wang et al. "Broadband achromatic optical metasurface devices". In: *Nature communications* 8.1 (2017), p. 187.

[117] Alan Zhan et al. "Inverse design of optical elements based on arrays of dielectric spheres". In: *Applied optics* 57.6 (2018), pp. 1437–1446.

[118] Alan Zhan et al. "Controlling three-dimensional optical fields via inverse Mie scattering". In: *Science advances* 5.10 (2019), eaax4769.

[119] Jinhie Skarda et al. "Simulation of large-area metasurfaces with a distributed transition matrix method". In: *CLEO: QELS_Fundamental Science*. Optica Publishing Group. 2022, FM5H–6.

[120] Raphaël Pestourie et al. “Inverse design of large-area metasurfaces”. In: *Optics express* 26.26 (2018), pp. 33732–33747.

[121] İpek Anıl Atalay et al. “A broad-band achromatic polarization-insensitive in-plane lens with high focusing efficiency”. In: *Acs Photonics* 8.8 (2021), pp. 2481–2488.

[122] Sensong An et al. “A Deep Learning Approach to Explore the Mutual Coupling Effects in Metasurfaces”. In: *CLEO: Applications and Technology*. Optica Publishing Group. 2021, JTu3A–75.

[123] Shane Colburn, Alan Zhan, and Arka Majumdar. “Varifocal zoom imaging with large area focal length adjustable metalenses”. In: *Optica* 5.7 (2018), pp. 825–831.

[124] Alan She et al. “Large area metalenses: design, characterization, and mass manufacturing”. In: *Optics express* 26.2 (2018), pp. 1573–1585.

[125] Joon-Suh Park et al. “All-glass, large metalens at visible wavelength using deep-ultraviolet projection lithography”. In: *Nano letters* 19.12 (2019), pp. 8673–8682.

[126] Jason Holloway et al. “SAVI: Synthetic apertures for long-range, subdiffraction-limited visible imaging using Fourier ptychography”. In: *Science advances* 3.4 (2017), e1602564.

[127] Ren Jie Lin et al. “Achromatic metalens array for full-colour light-field imaging”. In: *Nature nanotechnology* 14.3 (2019), pp. 227–231.

[128] Jacob Engelberg et al. “Near-IR wide-field-of-view Huygens metalens for outdoor imaging applications”. In: *Nanophotonics* 9.2 (2020), pp. 361–370.

[129] James EM Whitehead et al. “Fast extended depth of focus meta-optics for varifocal functionality”. In: *Photonics Research* 10.3 (2022), pp. 828–833.

[130] Gordon Wetzstein et al. “Inference in artificial intelligence with deep optics and photonics”. In: *Nature* 588.7836 (2020), pp. 39–47.

[131] Xing Lin et al. “All-optical machine learning using diffractive deep neural networks”. In: *Science* 361.6406 (2018), pp. 1004–1008.

[132] Carlos Mauricio Villegas Burgos et al. “Design framework for metasurface optics-based convolutional neural networks”. In: *Applied Optics* 60.15 (2021), pp. 4356–4365.

Vita

My email is luocheng@uw.edu.