# Project Proposal: Compact Computational Cameras by Joint Metasurface and Network Design

Kelsey Lee
EE367 | Winter 2026

## 1 Overview

The proposed project develops and evaluates hybrid optical-digital computational metalens cameras that function as optical neural networks (ONNs) for image classification. By allocating most of the necessary computation of the inference task to an analog optical frontend, the backend digital computational burden can be significantly reduced. We compare two lens architectures—a metalens array and a spatially multiplexed singlet metalens—under both standard shift invariant and spatially varying point spread function (PSF) models. Performance is analyzed with respect to key optical and digital metrics including field of view (FOV), effective focal length (EFL), and pixel sampling.

## 2 Motivation and Related Works

Optical computing has attracted growing renewed interest as modern machine learning workloads become increasingly compute- and data- intensive, creating significant pressure on conventional digital hardware. In computer vision applications in particular, low latency and compact form factors are necessary qualities for deploying sensors in real-time decision-making scenarios. Yet standard computer vision pipelines remain computationally expensive. Optical systems naturally perform key linear operations such as convolution and Fourier transforms in parallel and at the speed of light, making them a promising approach toward improved energy efficiency in compute tasks.

Optical computing platforms generally fall into two categories: free-space systems and integrated photonics. This project focuses on free-space optical computing, which is well suited for task-specific inference in applications such as robotics and embedded vision. In this context, metalenses—flat optical elements composed of subwavelength nanostructures that impart arbitrary phase control on incident light—are a natural candidate for compact optical computing implementations.

Recent works have shown significant progress in free-space optical computing for image classification on standard datasets such as MNIST, Fashion-MNIST, and CIFAR-10. These systems emulate convolution neural networks (CNNs) by engineering lens point spread functions (PSFs) to represent convolutional kernels. Two common configurations are the metalens array, where each lens implements a single convolutional kernel [1, 2], and a spatially multiplexed metalens, where one lens produces multiple spatially offset feature maps [3, 4]. While most prior work assumes shift-invariant PSFs, spatially varying models introduce additional degrees of freedom and have shown improved performance [2]. This project will systematically compare these approaches, with particular interest on the impact of spatially varying models on lens design and system performance.

## 3 Goals and Proposed Experiments

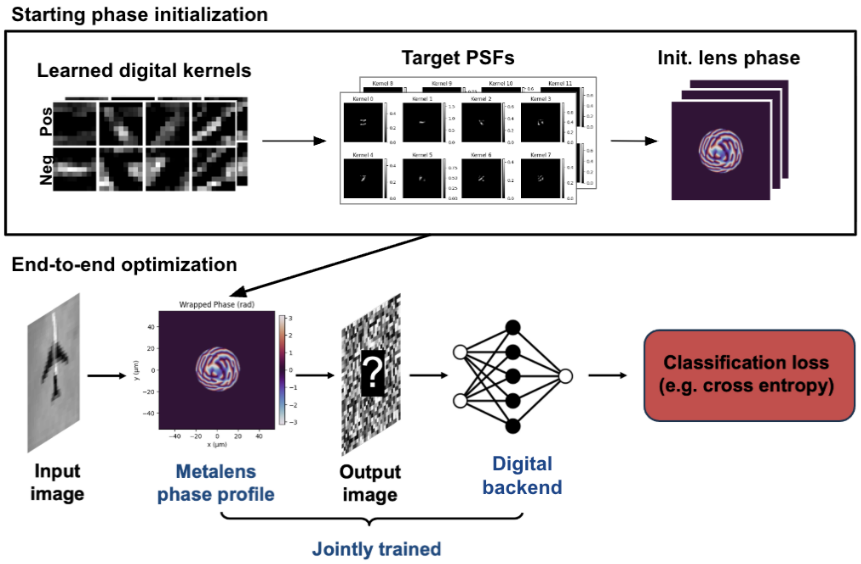


**Fig.** Sketch of the proposed simulation pipeline.

Classification will be evaluated on standard benchmarks (MNIST, Fashion-MNIST, and grayscale CIFAR-10) under a monochromatic imaging assumption for simplicity, since metalenses are dispersive and polychromatic optical encoders are nontrivial to design. The optical frontend is modeled using the angular spectrum method (ASM) to generate lens PSFs, followed by PSF-based image formation. For the spatially varying case, PSFs are computed at sparse field points and used with spatially varying convolution [5]. The digital backend consists of a nonlinear activation, pooling, and a fully connected layer, and the full system is trained end-to-end to jointly optimize the metalens phase profile and network weights. To provide a strong initialization, a digital CNN is first trained to obtain target kernels, which are split into positive and negative channels and converted into desired PSFs; the corresponding metalens phase is then initialized via Gerchberg-Saxton phase retrieval.

The goals of the project are as follows:

- **Demonstrate optical computing** using both a metalens array and a spatially multiplexed metalens singlet using a shift invariant optical model.
- **Implement spatially varying PSF model** and reoptimize phase profiles for both configurations.
- **Detailed ablation study** for the above approaches across digital parameters (e.g. kernel size and count) and optical parameters (e.g. EFL, FOV, and pixel sampling).

## 4 Project Timeline

**Week 7:** Build the shift invariant optical model and implement metalens array optimization.
**Week 8:** Implement optimization for the spatially multiplexed metalens singlet.
**Week 9:** Develop the spatially varying model and apply it to both lens configurations.
**Week 10:** Sweep digital and optical parameters and reoptimize. Prepare the poster and final document.

## 5 References

[1] Wirth-Singh, A., Xiang, J., Choi, M., Fröch, J. E., Huang, L., Colburn, S., ... & Majumdar, A. (2025). Compressed meta-optical encoder for image classification. *Advanced Photonics Nexus*, *4*(2), 026009-026009.

[2] Wei, K., Li, X., Froech, J., Chakravarthula, P., Whitehead, J., Tseng, E., ... & Heide, F. (2024). Spatially varying nanophotonic neural networks. *Science Advances*, *10*(45), eadp0391.

[3] Zheng, H., Liu, Q., Kravchenko, I.I. *et al.* Multichannel meta-imagers for accelerating machine vision. *Nat. Nanotechnol.* **19**, 471–478 (2024). https://doi.org/10.1038/s41565-023-01557-2

[4] Liang, R., Wang, S., Dong, Y., Li, L., Kuang, Y., Zhang, B., & Yang, Y. (2024). Metasurface-generated large and arbitrary analog convolution kernels for accelerated machine vision. *ACS Photonics*, *11*(12), 5430-5438.

[5] Denis, L., Thiébaut, E., Soulez, F., Becker, J. M., & Mourya, R. (2015). Fast approximations of shift-variant blur. *International Journal of Computer Vision*, *115*(3), 253-278.

[6] Choi, M., & Majumdar, A. (2025). Free-space optical encoder for computer vision. *npj Nanophotonics*, *2*(1), 36.