Demo: Low-Power Onboard Computer Vision for IoT Cameras

Chae Young Lee, Sara Achour, Zerina Kapetanovic

Stanford University
Stanford, USA
{chae,sachour,zerina}@stanford.edu

ABSTRACT

Wireless cameras can greatly benefit from onboard machine learning; however, conventional neural networks are typically power-hungry and require substantial computational resources. Brain-inspired hyperdimensional computing (HDC) is emerging as a promising alternative that is both energy-efficient and hardware-friendly. We demonstrate an ultra-low-power image classification pipeline based on HDC that runs entirely onboard wireless camera systems using about 50 KB of flash memory and 0.12-0.27 seconds for inference.

CCS CONCEPTS

• **Computing methodologies** → **Machine Learning**; • **Computer systems organization** → **Embedded systems**.

KEYWORDS

Wireless Cameras, Hyperdimensional Computing, TinyML

ACM Reference Format:

Chae Young Lee, Sara Achour, Zerina Kapetanovic. 2025. Demo: Low-Power Onboard Computer Vision for IoT Cameras. In *The 23rd Annual International Conference on Mobile Systems, Applications and Services (MobiSys '25), June 23–27, 2025, Anaheim, CA, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3711875.3734377

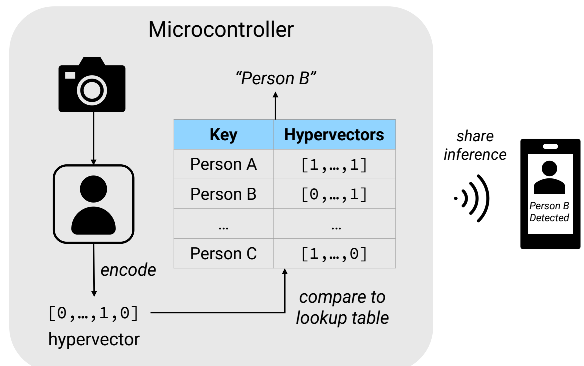


Figure 1: Onboard face detection using HDC.

1 INTRODUCTION

Coupled with advances in computer vision, wireless cameras are increasingly deployed for a wide range of practical applications such as agricultural monitoring, wildlife observation, and industrial safety [1, 2, 8]. These applications often rely on deep neural networks (DNNs), which require substantial computational resources and memory. However, wireless cameras are typically battery-powered and equipped with low-power microcontrollers that lack the capacity to run such models efficiently. As a result, many systems rely on cloud offloading, incurring latency, energy, and privacy concerns [4, 7, 9]. These drawbacks motivate onboard computer vision approaches, and developing DNNs that run onboard microcontrollers is an active area of research.

Hyperdimensional computing (HDC) emerges as a compelling alternative for low-resource hardware. By representing data with high-dimensional binary vectors and using bitwise operations, HDC achieves a hardware-friendly, energy-efficient form of computation [3, 5]. However, most existing HDC work has focused on time-series data rather than imagery, and naïve HDC methods can introduce substantial memory and latency overhead when encoding and classifying images. These challenges become especially pronounced on microcontrollers with highly constrained flat memory hierarchies and tight power budgets.

In this work, we demonstrate a real-time, low-power image classification pipeline that uses novel HDC encoding techniques to enable onboard classification on wireless cameras. It features two versions of a lightweight encoder introduced in [6] capable of producing and manipulating hypervectors without requiring large precomputed mappings, thus significantly reducing both model size and runtime latency. As shown in Figure 1, our system demonstrates face detection using its onboard camera and transmits inference results to a nearby smartphone using Bluetooth.

MobiSys '25, June 23–27, 2025, Anaheim, CA, USA

ACM ISBN 979-8-4007-1453-5/25/06
https://doi.org/10.1145/3711875.3734377

2 HDC-BASED FACE DETECTION SYSTEM & DEMO

We present a low-power wireless camera platform that performs onboard face detection with an optimized hyperdimensional (HD) classifier [6]. The core component of this classifier is a lightweight encoding pipeline featuring two encoder variants: a Bloom Filter encoder, designed for faster, more memory-efficient inference, and a Count Sketch encoder, which offers higher accuracy at the cost of slightly increased resource usage. The system supports both binary face detection (presence vs. absence) and multi-class face identification (recognizing specific individuals), although we limit the demonstration to the binary detection case for public audiences. The performance of HD and baseline classifiers on binary face detection task is compared in Table 1.

Model	Accuracy	Memory	Latency
`MobileNet V3`	88.18%	1640.00 KB	18.53 s
`MCUNet V3`	**99.88%**	1190.00 KB	6.70 s
HDC (Count Sketch)	92.98%	53.83 KB	0.27 s
HDC (Bloom filter)	92.73%	**42.91 KB**	**0.12 s**

Table 1: Performance on binary face detection.

As shown in Figure 3a, the low-power wireless camera platform uses the evaluation board for the STM32U585 microcontroller, which has one ARM Cortex M-33 microprocessor, 2 MB of flash memory, and 736 KB of RAM. All non-critical components in the evaluation board were removed to save power. The Himax HM01B0 image sensor is used to capture 160x120 grayscale images. A custom printed circuit board is implemented to interface the microcontroller with the image sensor. In addition, we use the nRF52840 Bluetooth module to transmit data packets to a nearby base station.

Component	Active Current	Sleep Current	Voltage
STM32U585I	10.9 mA	74.9 μA	3.3V
Himax HM01B0	2.5 mA	1.3 mA	2.8V
nRF52840 Express	7.2 mA	1.4 mA	3.8V

Table 2: Power consumption per component.

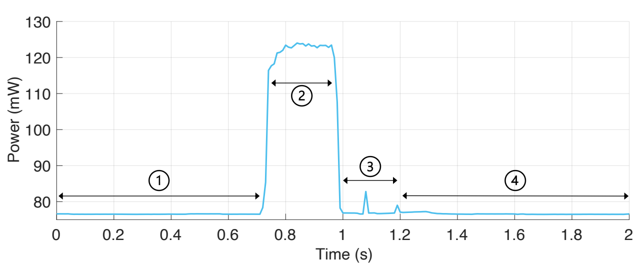
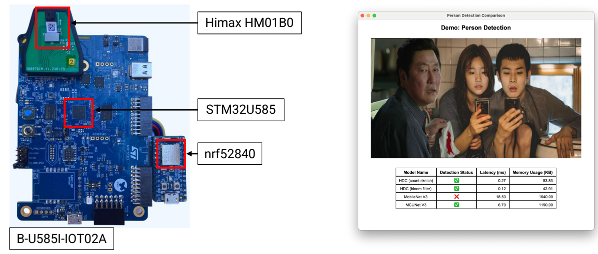


Figure 2: System power consumption. (1) camera platform in sleep mode (2) camera initialization, image capture, and inference, (3) data transmission, and (4) system returns to sleep mode.



(a) Hardware **(b) Screen display**

Figure 3: Demonstration setup.

The power consumption of the hardware is shown in Table 2 and Figure 2.

During the demonstration, our camera system is connected to a laptop that displays the captured image frame alongside the HD classifier's performance metrics in real time, as shown in Figure 3b. For comparison, we also show results from 2 baseline ML classifiers. Key metrics include the detection result, inference latency, and memory usage. The results should closely match those in Table 1.

REFERENCES

[1] Spartan Camera. 2024. Spartan GoLive Cellular Trail Camera. https://go.spartancamera.com/products/spartan-golive-4g-lte?srsltid=AfmBOop-EN_ehDVLtHK8nX_m46QsFOI_Z3nB8bkWI1fcWpw95tpO5wnQ.

[2] Instant Detect. 2024. Instant Detect - Multi-sensor monitoring system for protected areas. https://instantdetect.co.uk.

[3] Mohsen Imani, Abbas Rahimi, Deqian Kong, Tajana Rosing, and Jan M Rabaey. 2017. Exploring hyperdimensional associative memory. In *2017 IEEE International Symposium on High Performance Computer Architecture (HPCA)*. IEEE, 445–456. https://doi.org/10.1109/HPCA.2017.28

[4] Colleen Josephson, Lei Yang, Pengyu Zhang, and Sachin Katti. 2019. Wireless computer vision using commodity radios. In *Proceedings of the 18th International Conference on Information Processing in Sensor Networks*. 229–240.

[5] Jovin Langenegger, Geethan Karunaratne, Michael Hersche, Luca Benini, Abu Sebastian, and Abbas Rahimi. 2023. In-memory factorization of holographic perceptual representations. *Nature Nanotechnology* 18, 5 (2023), 479–485.

[6] Chae Young Lee, Pu, Yi, Maxwell Fite, Tejus Rao, Sara Achour, and Zerina Kapetanovic. 2025. HyperCam: Low-Power Onboard Computer Vision for IoT Cameras. arXiv:2501.10547 [cs.CV] https://arxiv.org/abs/2501.10547

[7] Saman Naderiparizi, Zerina Kapetanovic, and Joshua R. Smith. 2016. WISPCam: An RF-Powered Smart Camera for Machine Vision Applications. In *Proceedings of the 4th International Workshop on Energy Harvesting and Energy-Neutral Sensing Systems (ENSsys '16)*. Association for Computing Machinery, New York, NY, USA, 19–22.

[8] Arun Narenthiran Sivakumar, Sahil Modi, Mateus Valverde Gasparino, Che Ellis, Andres Eduardo Baquero Velasquez, Girish Chowdhary, and Saurabh Gupta. 2021. Learned Visual Navigation for Under-Canopy Agricultural Robots. In *Proceedings of Robotics: Science and Systems*. Virtual. https://doi.org/10.15607/RSS.2021.XVII.019

[9] B. Veluri, C. Pernu, A. Saffari, J. Smith, M. Taylor, and S. Gollakota. 2023. NeuriCam: Key-Frame Video Super-Resolution and Colorization for IoT Cameras. 1–17.