

A Lens-free Microscope with Edge AI Acceleration

Lena Zhukova^{1,2}, Luqi Yan³, Agus Dharmawan³, J. Daniel Prades³, Ismael Benito-Altamirano^{1,3,4}

¹ e-Health Centre & Faculty of Computer Science, Multimedia

and Telecommunications, Universitat Oberta de Catalunya, Barcelona, Spain

² Centre for Sensors, Instruments and Systems Development (CD6), Universitat Politècnica de Catalunya, Barcelona, Spain

³ Institute of Semiconductor Technology (IHT) & Laboratory for Emerging Nanometrology (LENA), Technische Universität Braunschweig, Braunschweig, Germany

⁴ Department of Electronics and Biomedical Engineering, Universitat de Barcelona, Barcelona, Spain
ibenitoal@uoc.edu

Summary:

Lens-free microscopy offers a compact lab-on-a-chip solution but typically requires high external computational power for image reconstruction. This work introduces a lens-free system utilizing the Sony IMX500 sensor within the Raspberry Pi AI Camera. By removing the integrated lens, we leverage the sensor's internal AI accelerator to process in-line holographic images directly. Instead of traditional diffraction-based back-propagation, we employ Convolutional Neural Networks (CNNs) for real-time cell counting. Using a synthetic dataset for training, we applied post-training quantization to deploy our model onto the IMX500 hardware. Results demonstrate that the on-sensor AI performs comparably to a GPU, establishing a localized, low-power processing pipeline for portable lab-on-a-chip sensing.

Keywords: lens-free, microscopy, cell count, edge AI, deep learning

Introduction

Lens-free microscopy has become an easy solution for lab-in-a-chip solutions[1], despite this lens-less microscopy brings computational cost to the table which is often computed outside the lab-in-a-chip or in resource-intensive FPGA development boards. The post-processing pipeline for such images involve computer vision tasks (i.e. optical propagation, particle localization and counting, ...)[2].

In this work we introduce a new lens-free microscope based in the IMX500 sensor from Sony, packaged inside the Raspberry Pi AI Camera. The Sony IMX500 is a CMOS sensor that incorporates an AI accelerator inside the same chip, enabling running models like YOLO for human-activity recognition (see Fig.1).

By a small hardware modification of this camera, removing its lens system, we integrated it in our lens-free microscopy setup (see Fig.2a). We propose to use the AI accelerator inside the camera sensor for lens-free image processing. Rather than computing a back-propagation of the optical field using diffraction laws[2], we process the images with a CNN model to solve different task, like particle or cell counting directly from the in-line holographic image.

To train our CNN model, we created a synthetic dataset using the optical field propagation laws for several distributions of simulated samples resembling those from our laboratory (see Fig.b and Fig.c). Then, we trained the model on GPU, performed a quantization to the model, using the post-training quantization technique, and, finally, compiled the models for the IMX500 hardware. Models shown similar behavior in GPU and inside the IMX500 AI accelerator, enabling this technology for lab-on-a-chip sensing.

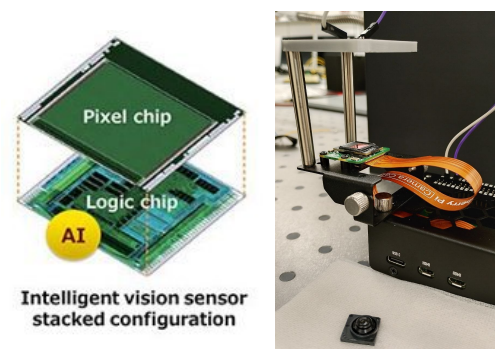


Fig. 1. (Left) The Sony IMX500 hardware stack (image from Sony). (Right) Provisional setup for a lens-less microscope based on a modified Raspberry Pi AI Camera. Image show the removed lens at the bottom.

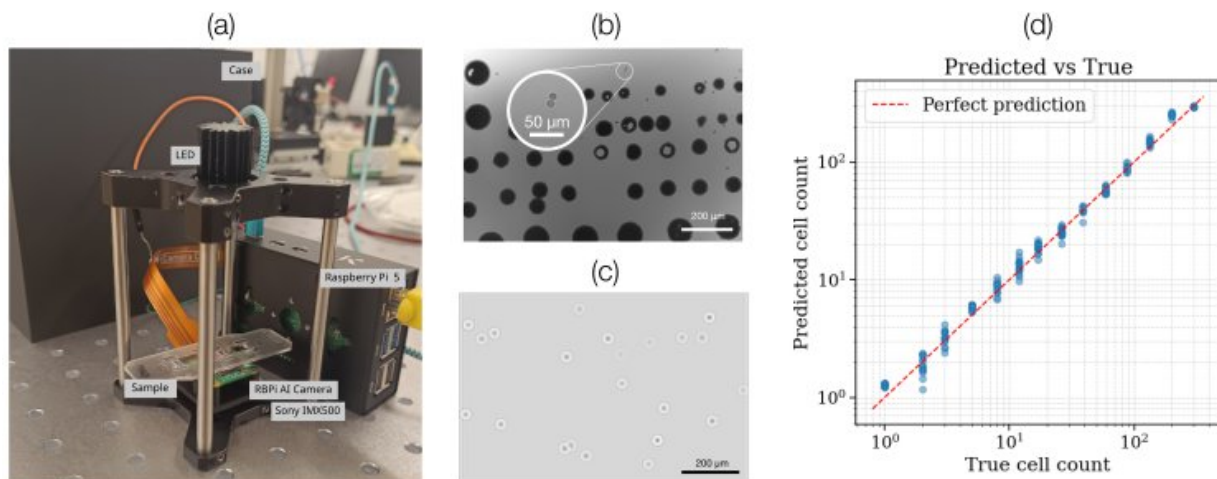


Fig. 2. (a) Our lens-free microscope setup, consists of: a case, an LED with a dissipation and a dampening filter, a sample, the modified AI camera a Raspberry Pi 5. (b) An actual image from a sample created with lithography from the IMX500 sensor. (c) A simulated image of a distribution of cells to train and to be predicted by the model. (d) Results from training a dataset of synthetic images.

Methodology

For the computational core of the system we defined a neural model, using PyTorch, this model consisted of a convolutional neural network (CNN) architecture designed for feature extraction through an 11-layer encoder, followed by a regression head for particle counting. To bridge the gap between high-level deep learning and resource-constrained sensing hardware, we quantized our model, this means we converted the model from a *float* model to a discrete logic model represented by *integer* logic. We used the Post-Training Quantization (PTQ) method[3], recommended by the IMX500 manufacturer. Using this technique, the model was optimized and compiled specifically for the IMX500 intelligent vision sensor's architecture. Enabling the sensor to perform on-chip inference directly from in-line holographic data.

The synthetic dataset was designed to cover a broad dynamic range of particle densities, utilizing a logarithmic distribution of cell counts ranging from 1 to 300. Each simulation modeled a 512×512 pixel window with a $2 \mu\text{m}$ pitch to match the IMX500 physical characteristics. To ensure model robustness and depth-of-field adaptability, we introduced stochastic physical parameters in the simulator: the sample-to-sensor distance was randomized between $400 \mu\text{m}$ and $1000 \mu\text{m}$, and spherical illumination was applied based on a fixed 70 mm source-to-sensor distance. The training data was generated by applying Fresnel propagations[2] to objects with randomized cell morphologies and positions, rendering a total of 1.400 images.

Results

The model training was performed on GPU (NVIDIA GeForce RTX 2080 Ti) on a training split (1.190 samples), and successfully converged within 31 epochs, achieving a final validation MSE of 0.00054 and a MAE of 0.044 in log10 space. Then we used 10 random samples from the training set to perform the post-training quantization on the model, obtaining the quantized model. Finally, we observed the correlation between predicted and true cell counts over the test split (210 samples) remains highly linear across the entire dynamic range (see Fig.2d), demonstrating that the quantized model maintains high fidelity for counting tasks even when deployed on the resource-constrained IMX500 hardware.

Acknowledgments

I.B. ack. the support for the Spanish MCIU under the CAS24/00369 José Castillejo grant. Also, and the RA 2026 program from UOC. J.D.P. ack. the support from the Alexander von Humboldt Professorship of the Humboldt Foundation and the Federal Ministry for Education and Research (Germany), and the zukunfft.niedersachsen program of the Lower Saxony MSC and the Volkswagen Foundation.

References

- [1]A. B. Dharmawan et al., "Nonmechanical parfocal and autofocus features based on wave propagation distribution in lensfree holographic microscopy," *Sci Rep*, vol. 11, no. 1, p. 3213, Feb. 2021, doi: 10.1038/s41598-021-81098-7.
- [2]I. Benito-Altamirano et al., "MicroVi: A Cost-Effective Microscopy Solution for Yeast Cell Detection and Count in Wine Value Chain," *Biosensors*, vol. 15, no. 1, p. 40, 2025.
- [3]J. Zhang, Y. Zhou, and R. Saab, "Post-training quantization for neural networks with provable guarantees," *SIAM journal on mathematics of data science*, vol. 5, no. 2, pp. 373–399, 2023.