

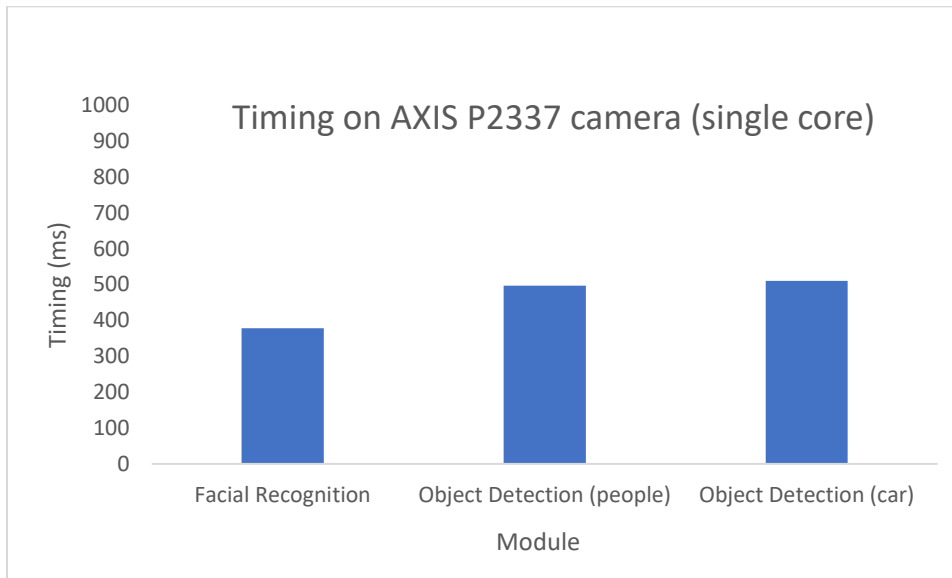
Karim Ali, Invision

Deep Learning on off-the-shelf cameras with examples in Top View, People Detection, Facial Recognition, and Car Detection

Problem: We are transforming leading commercial IP video cameras into intelligent devices that can detect, classify and objects in real-time by running Invision AI software using their internal processors only. The target cameras are Axis Communication IP security cameras as well as Ambarella S3L processors. Our greatest challenge is to fit our platform within the limited computational envelope available on commercial cameras. However, we have extensive experience in fitting our platform on production-grade low-power CPUs.

Technical approach and novelty: As part of partnerships with AXIS and Ambarella, we have access to the APIs required to port our various object detectors to camera CPUs. Our goal is to run our Top-Down people counter, people detection, facial recognition, and car detection modules on camera CPUs at 2-3fps on HD resolutions. We have developed fully convolutional deep neural network architectures that can fit within CPU constrained as well as memory constrained devices while providing state-of-the-art accuracy.

Results: Each module was ported to the [AXIS P2337 camera](#) which has a dual-core ARM CPUs (ARTPEC-6 - 512 MB flash, 512 MB RAM). Current timings for each detector is shown in Figure 1 on just one of the CPUs. Our Facial recognition achieves an accuracy of 97.8% on LFW while our object detector achieves between 92% and 94% TP at 0.01 FP per image. Our models typically utilize less than 5MB of memory and run in under 500ms on single core ARM A9 CPU.



Significance to the tinyML community: We are the first to demonstrate that deep learning models can perform reliably on off-the-shelf security cameras. To enable this, we are using our own C++ stack at run-time with as little as 10% memory overhead compared to the size of the network.