FOMO: Real-Time Object Detection on Microcontrollers

Jan Jongboom
CTO
Edge Impulse

Wait... What? Edge Impulse?

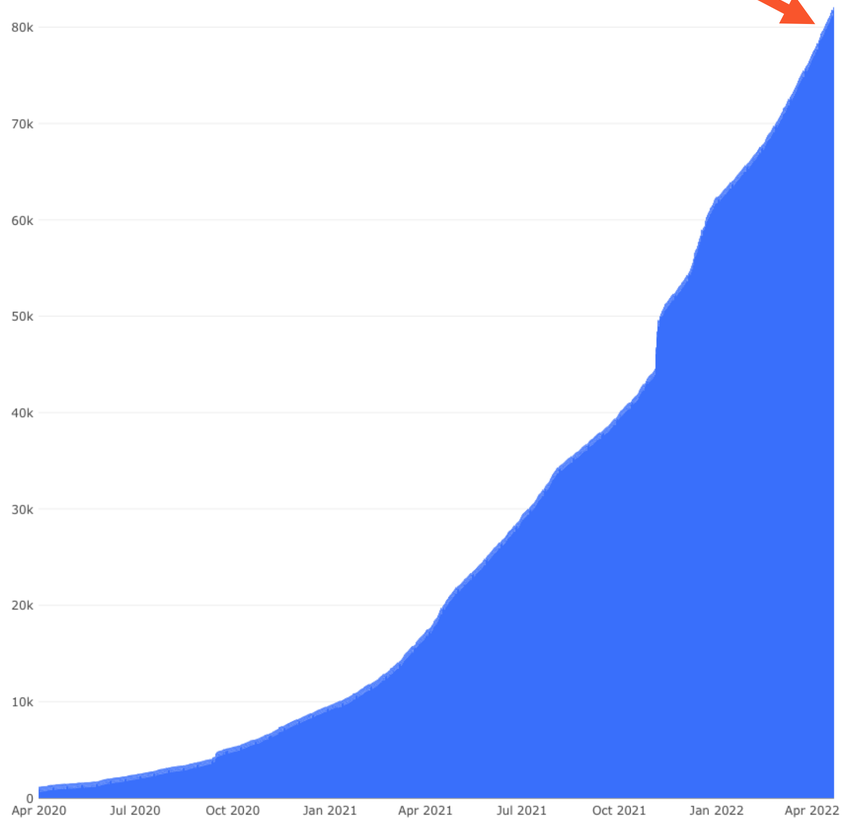
Leading development platform for machine learning on edge devices

Launched two years ago... Now 240 new projects daily (!)

40% of these are vision projects

Today

(82,069 projects)



Two years ago

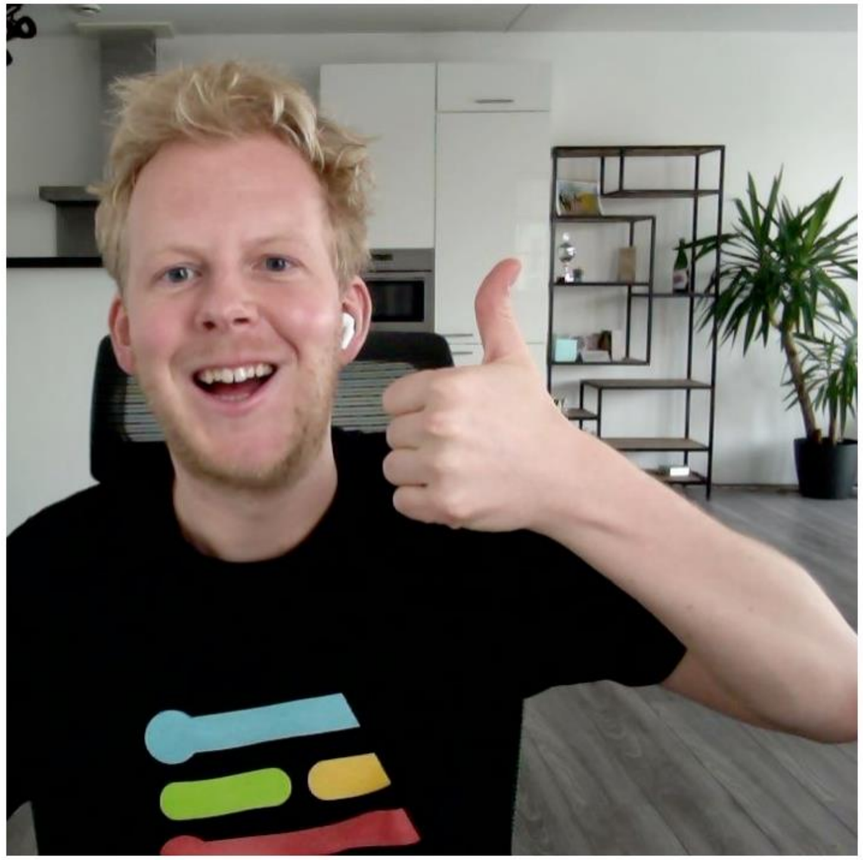
(1,163 projects)

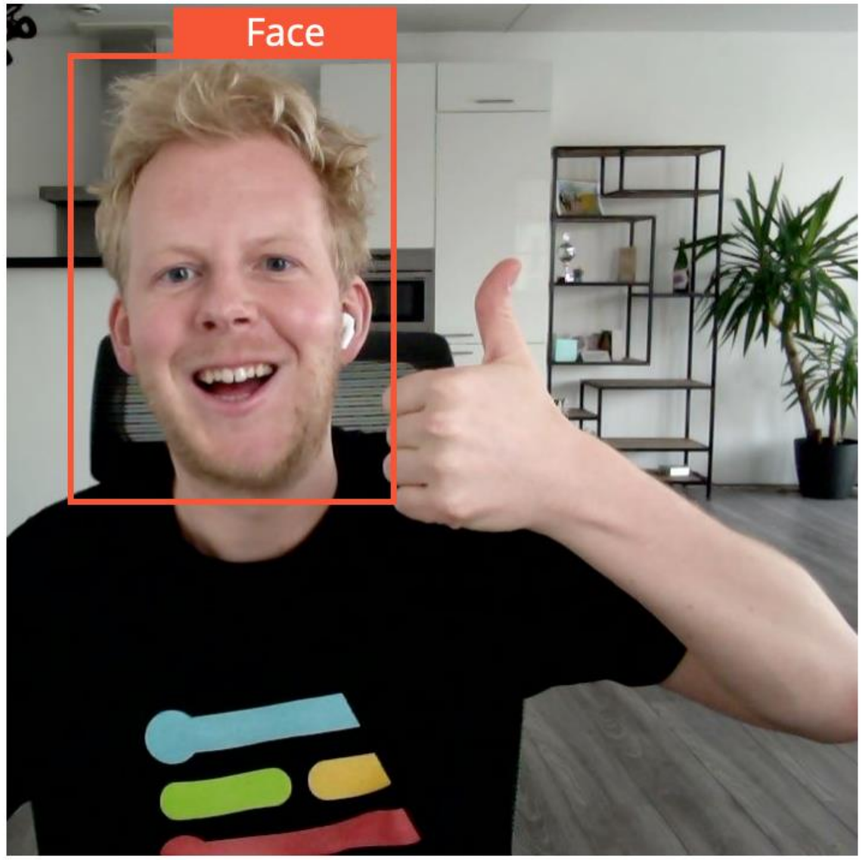
Edge Impulse project count

© 2022 Edge Impulse Inc.

Object Detection

Image classification vs. object detection





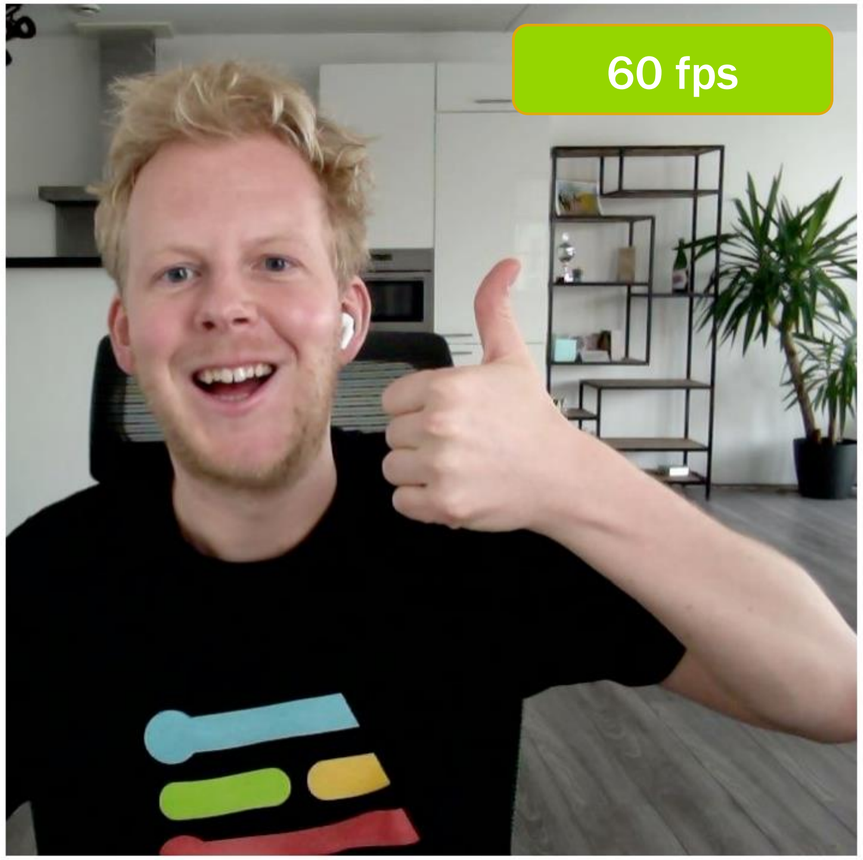
EDGE IMPULSE

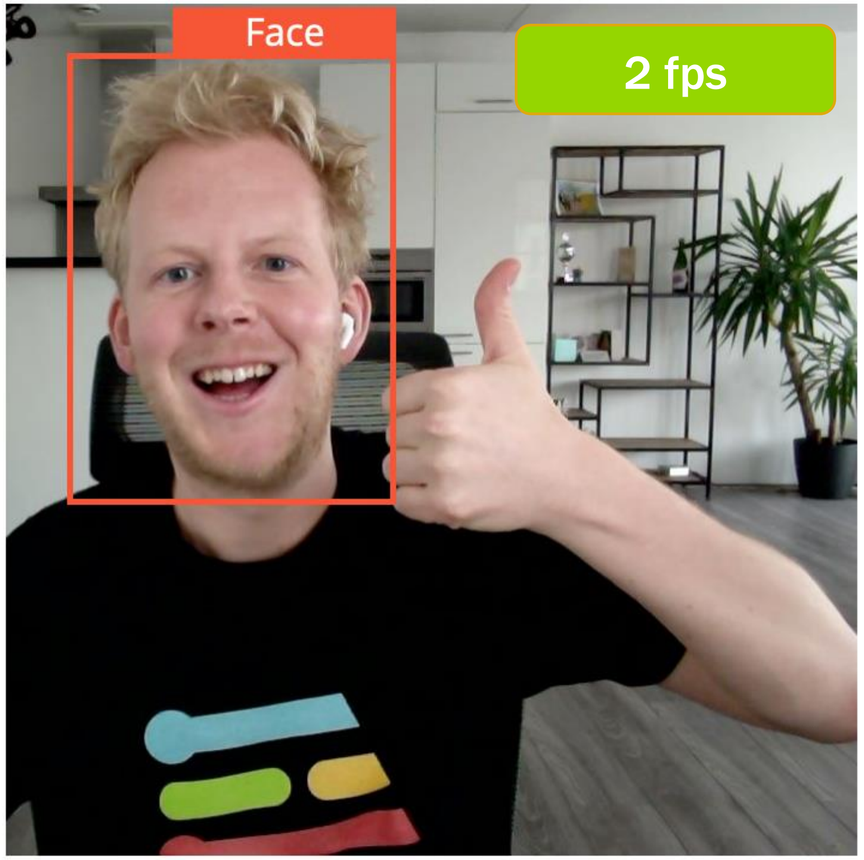
Face ✅

© 2022 Edge Impulse Inc.

Significantly Different Performance (Rpi4)

Image classification vs. object detection



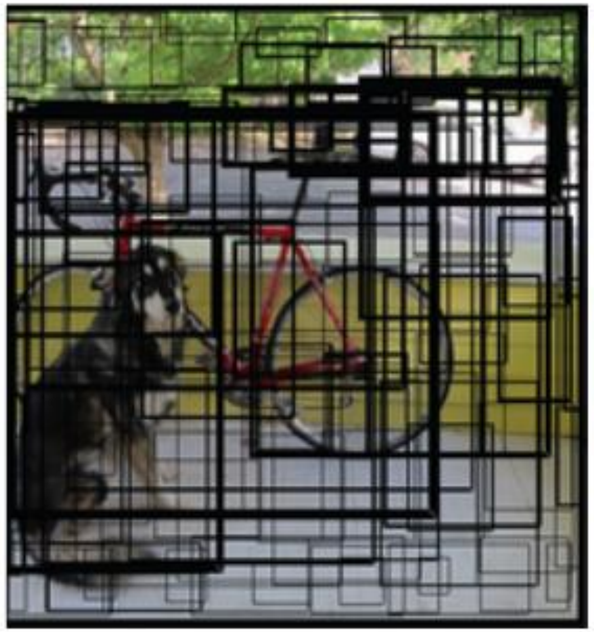


EDGE IMPULSE

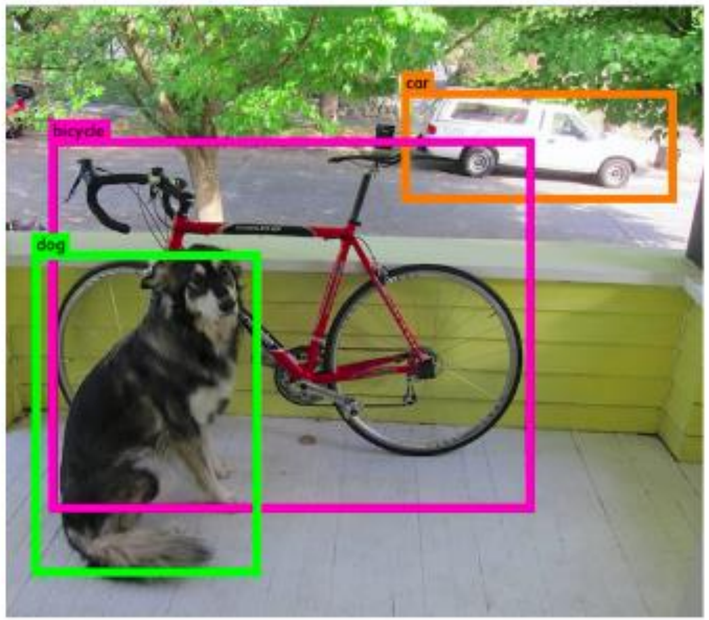
Face ✅

© 2022 Edge Impulse Inc.

But Why?



Multiple Bounding Boxes



Final Bounding Boxes

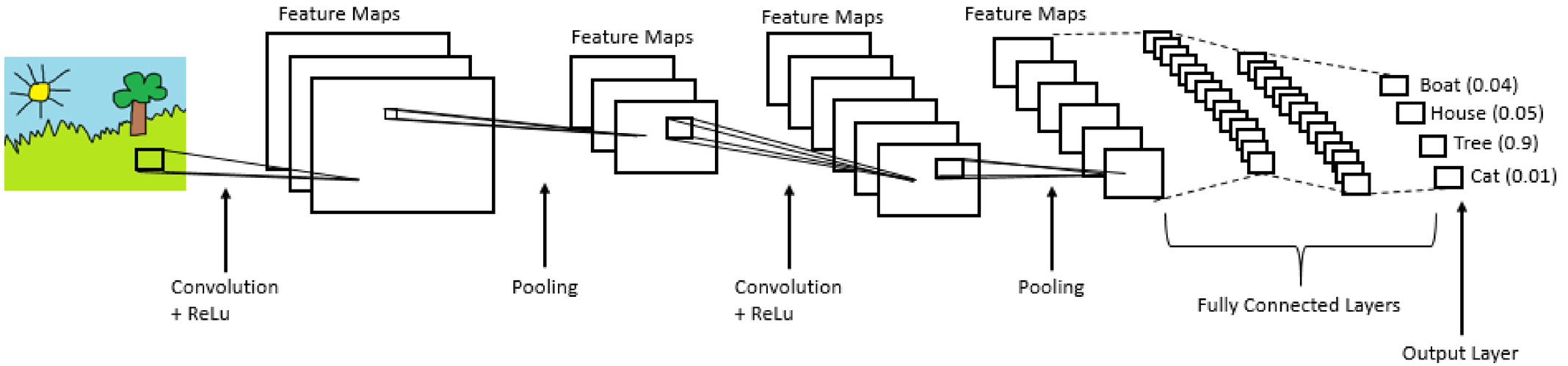
Source: https://pjreddie.com/darknet/yolov1/

"The tiny version of YOLO only uses 516 MB of GPU memory"

EDGE IMPULSE

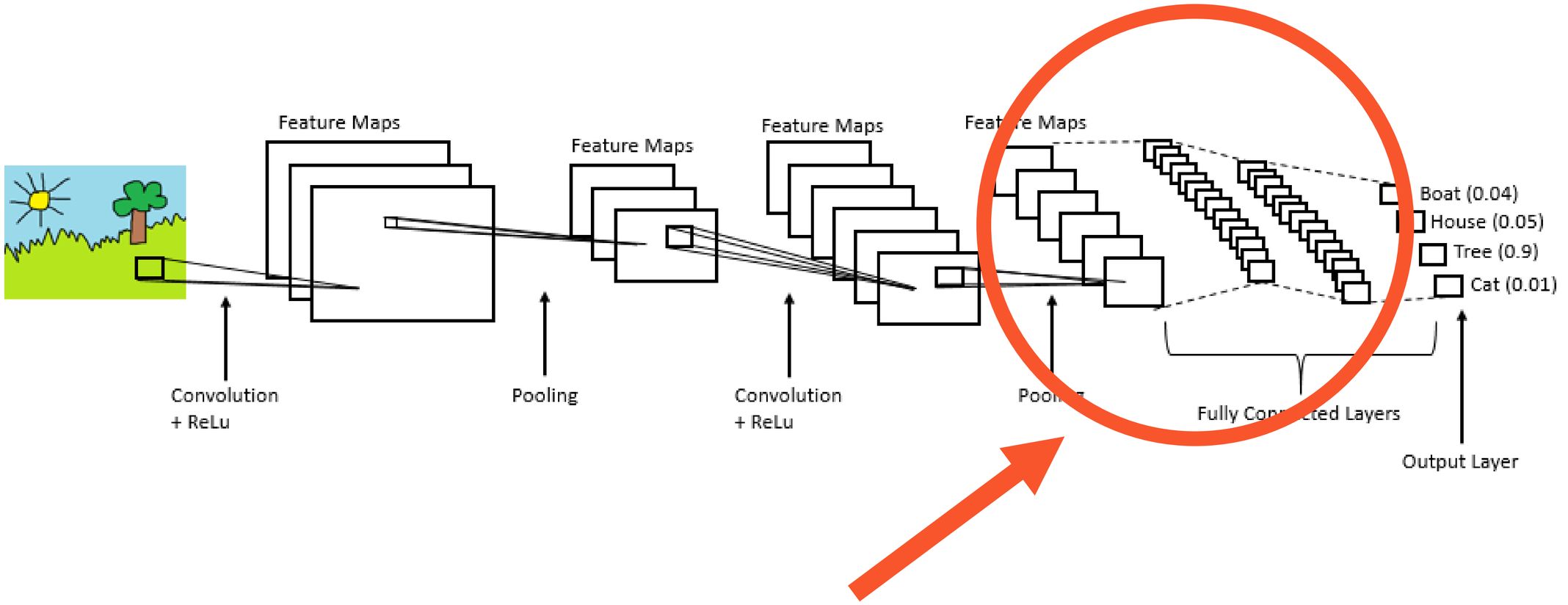
© 2022 Edge Impulse Inc.

Object Detection (MobileNet SSD, YOLO)



© 2022 Edge Impulse Inc.

FOMO



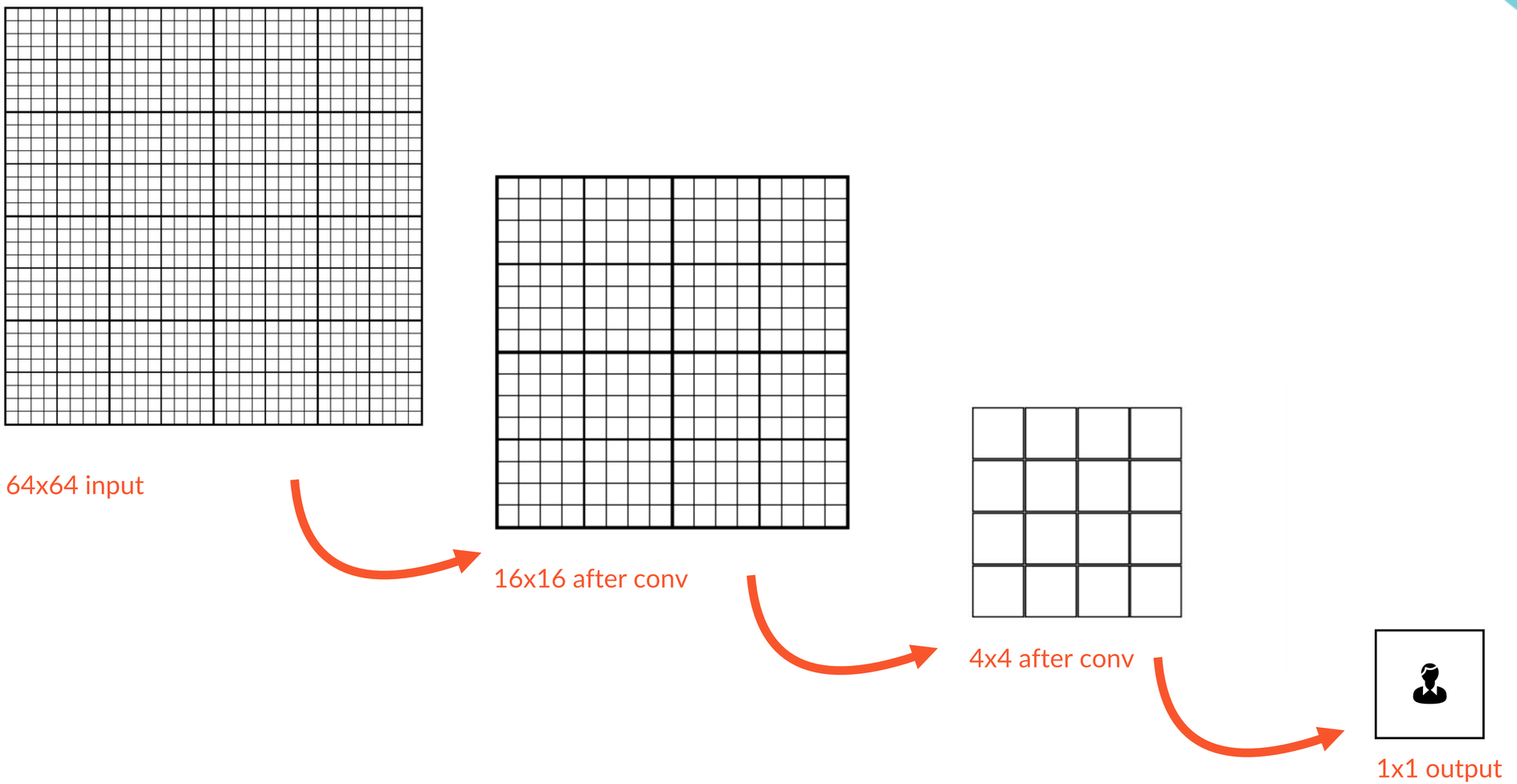
Replace with single per-region class probability map

© 2022 Edge Impulse Inc.

Single Per-Region Class Probability What?

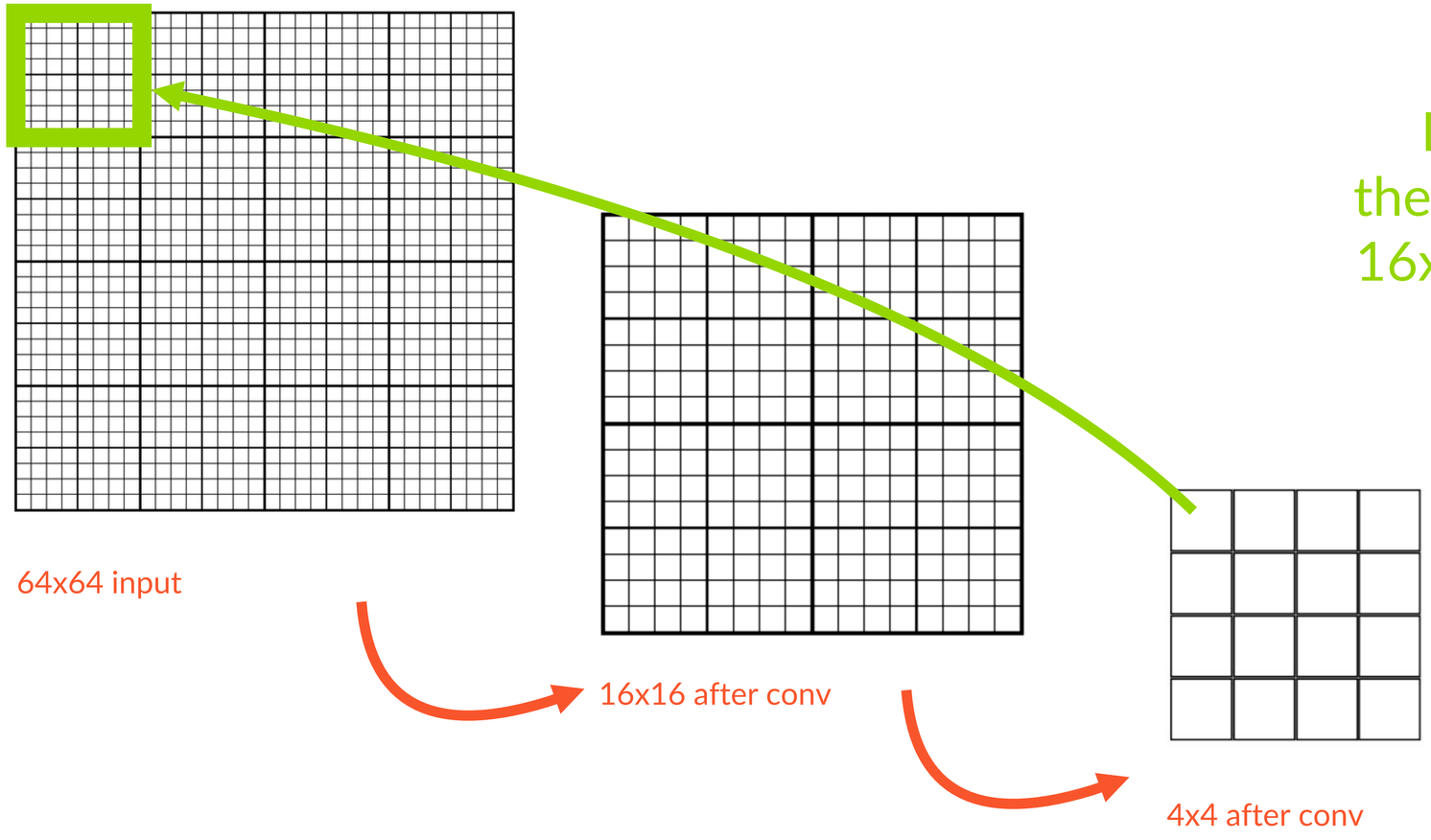
EDGE IMPULSE

Convolutional Image Classification Model



© 2022 Edge Impulse Inc.

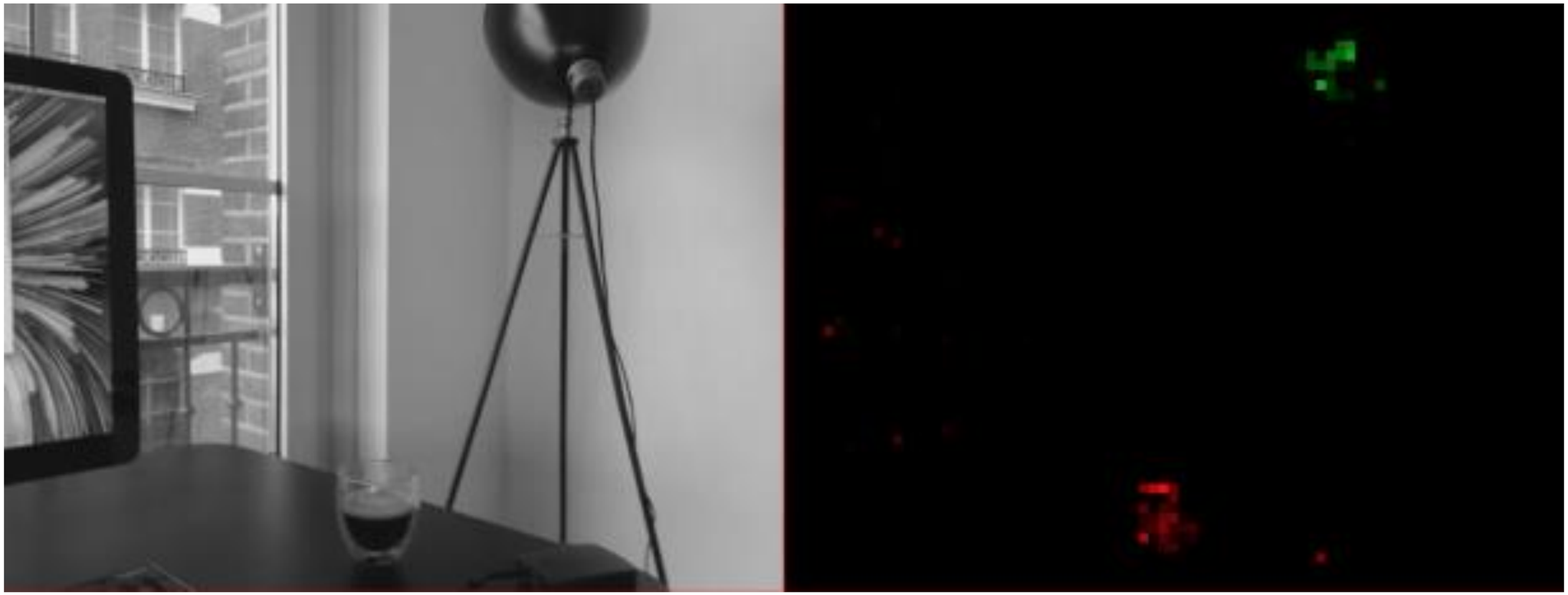
FOMO



Loss function that forces
the top-left tile here to map to
16x16 top-left tile in the input

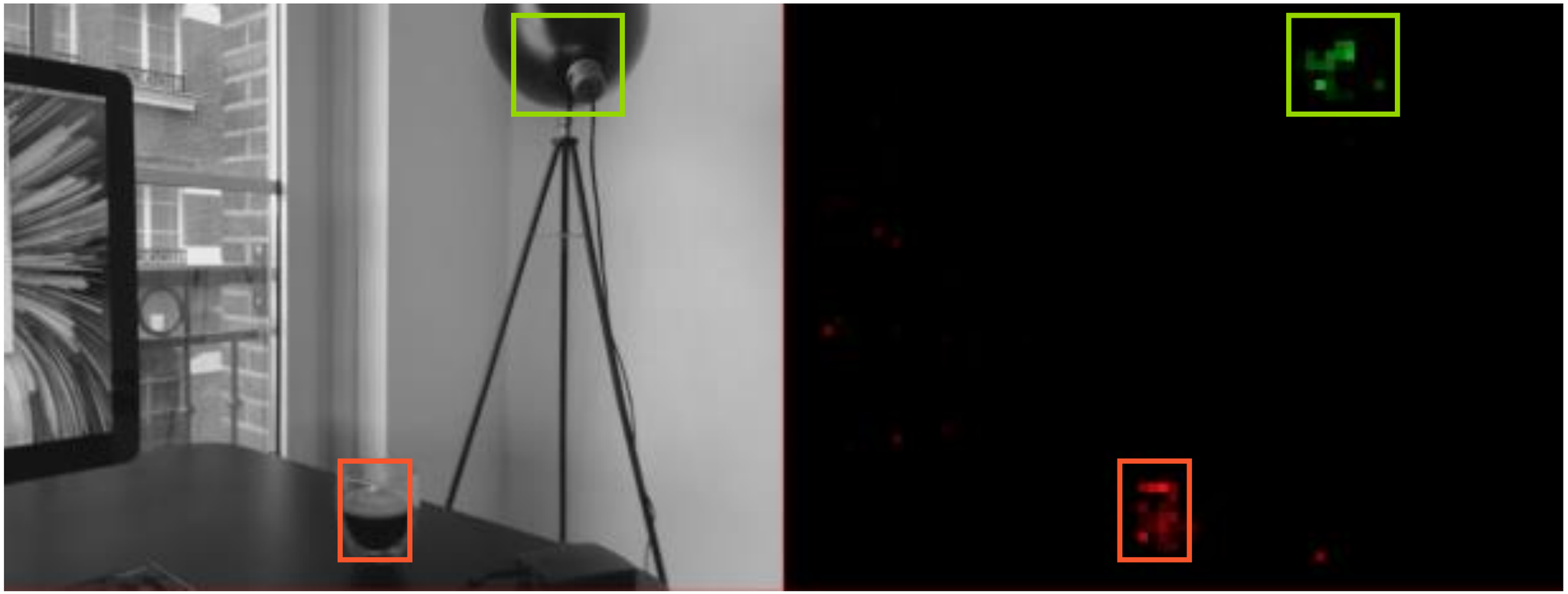
© 2022 Edge Impulse Inc.

Output = Heatmap



© 2022 Edge Impulse Inc.

Postprocessing in Software



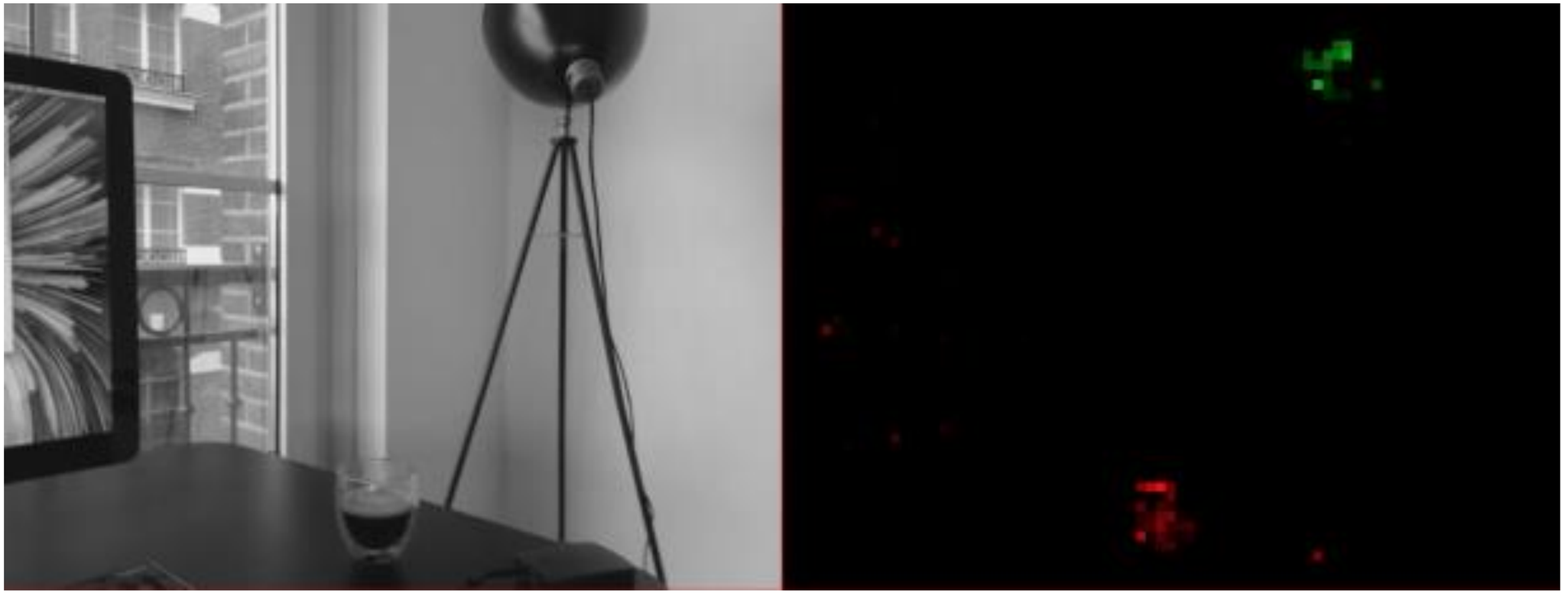
Not constrained to just bounding boxes!

© 2022 Edge Impulse Inc.

Heatmap Ratio is Configurable

© 2022 Edge Impulse Inc.

Flexibility in Input / Output Resolution

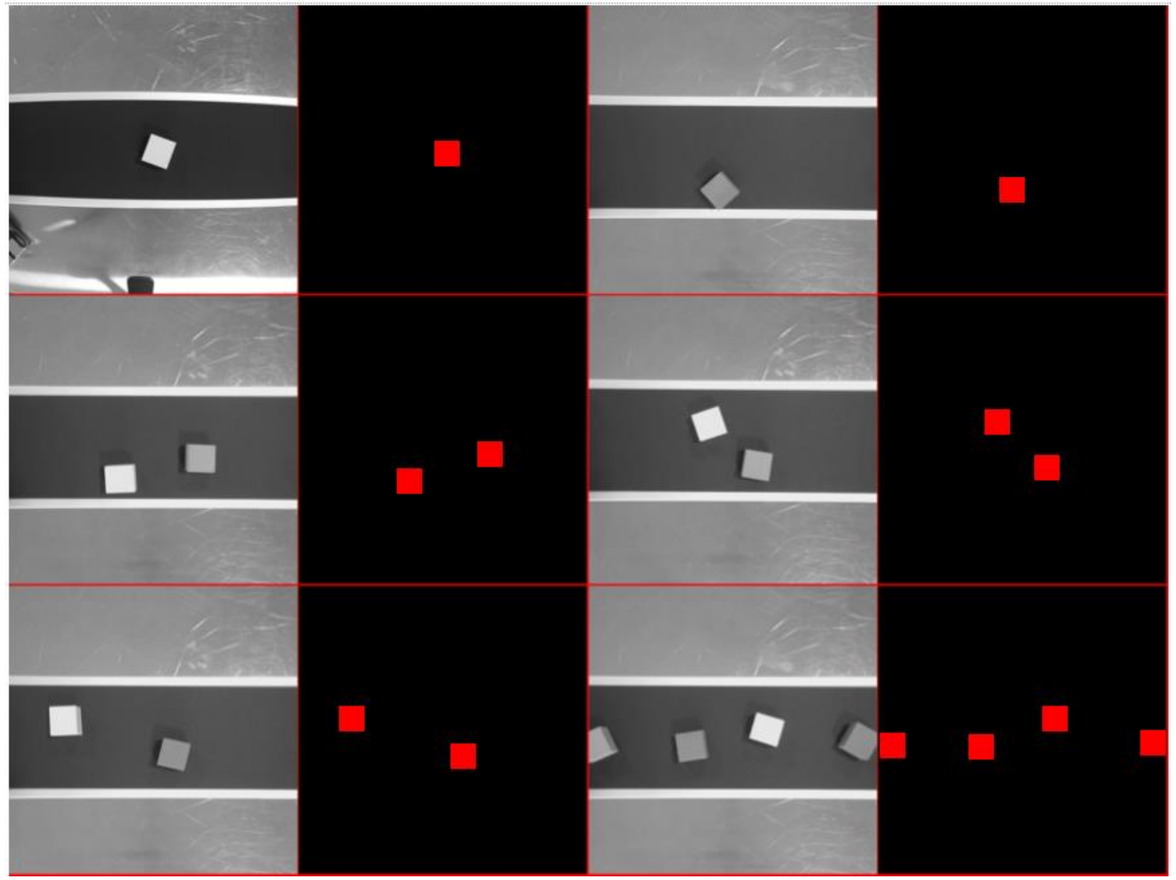


320 x 320 input => 80 x 80 output

EDGE IMPULSE

© 2022 Edge Impulse Inc.

Flexibility in Input / Output Resolution



96 x 96 input => 12 x 12 output

© 2022 Edge Impulse Inc.

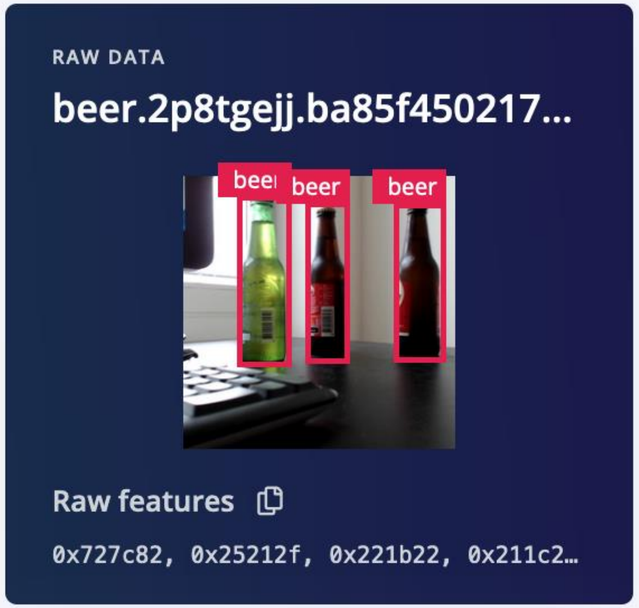
Centroids

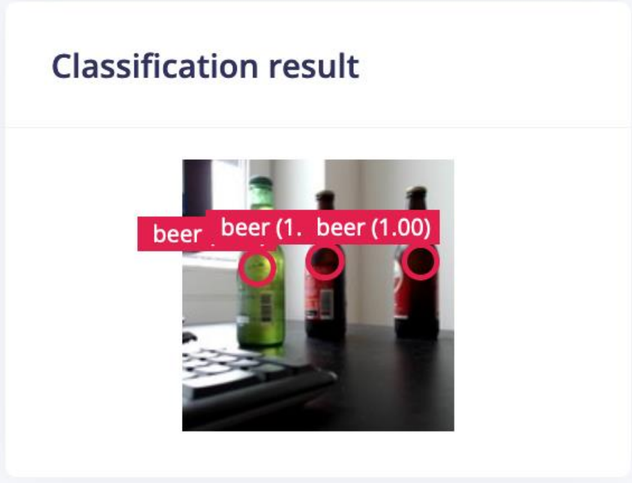
Bounding boxes are an implementation detail

Most of the time you just want to know **where** and **how many** objects there are

FOMO trains on the centroid of an object, w/ loss function allowing some error

Convolutional network, so will still look around the object, but lot less error prone against background

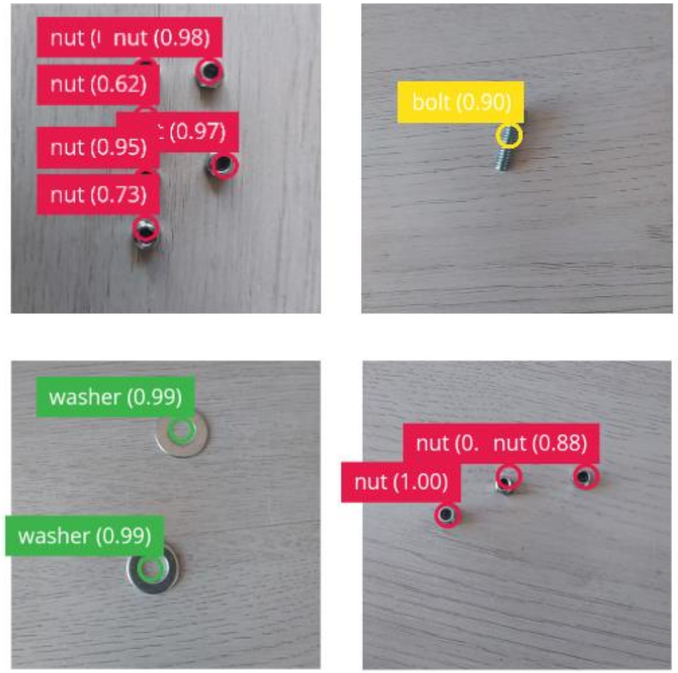




© 2022 Edge Impulse Inc.

Caveats

"Equal to an object detector, but bounding boxes are fixed and always square"



Constrained object detection for constrained problems

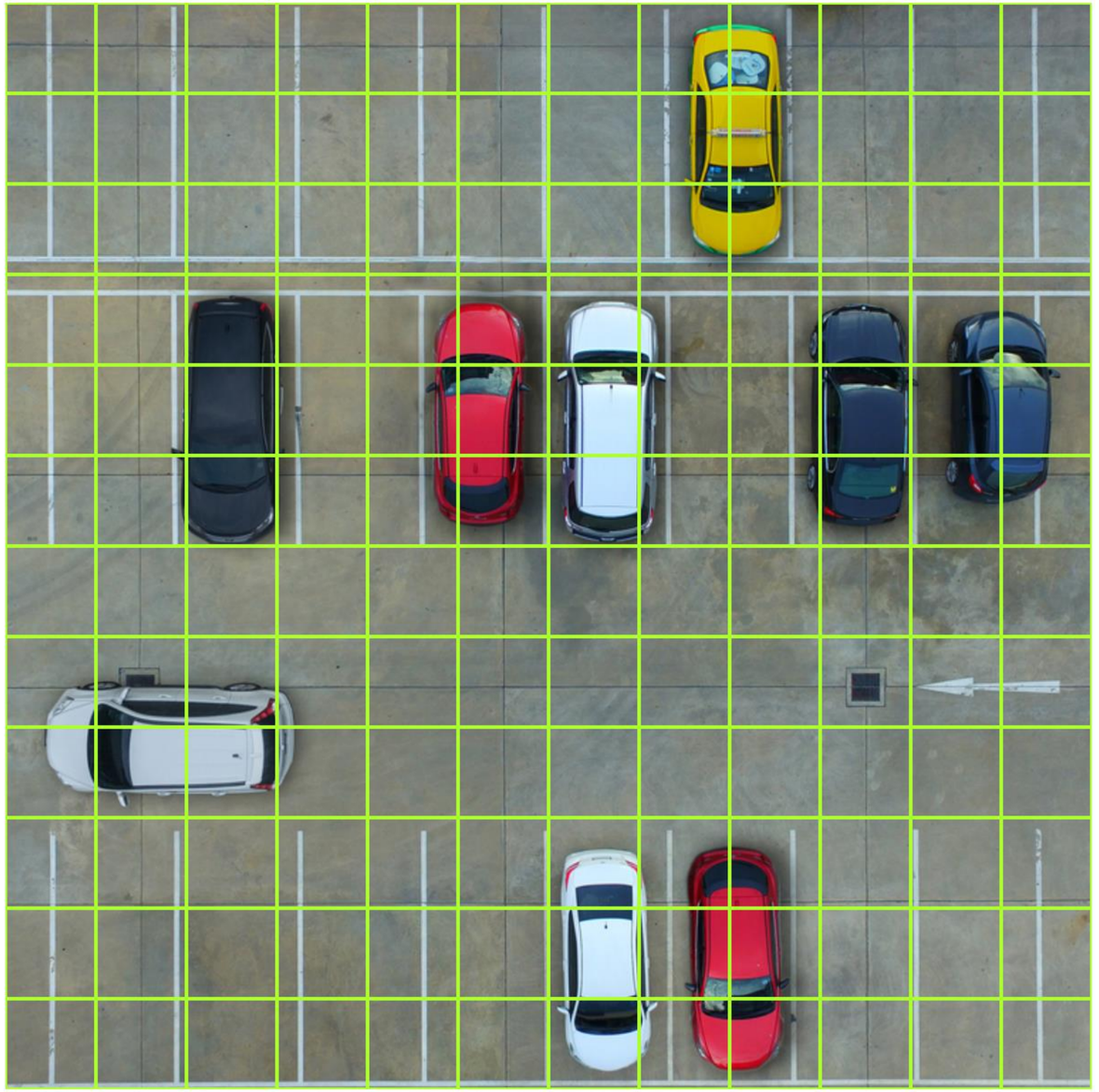
Works best with similar sized objects, at similar distances

One activation per heat map cell, so objects overlapping each other might be fused together

Few classes? Fine. 100s of objects across 60 classes? Use YOLOv5.

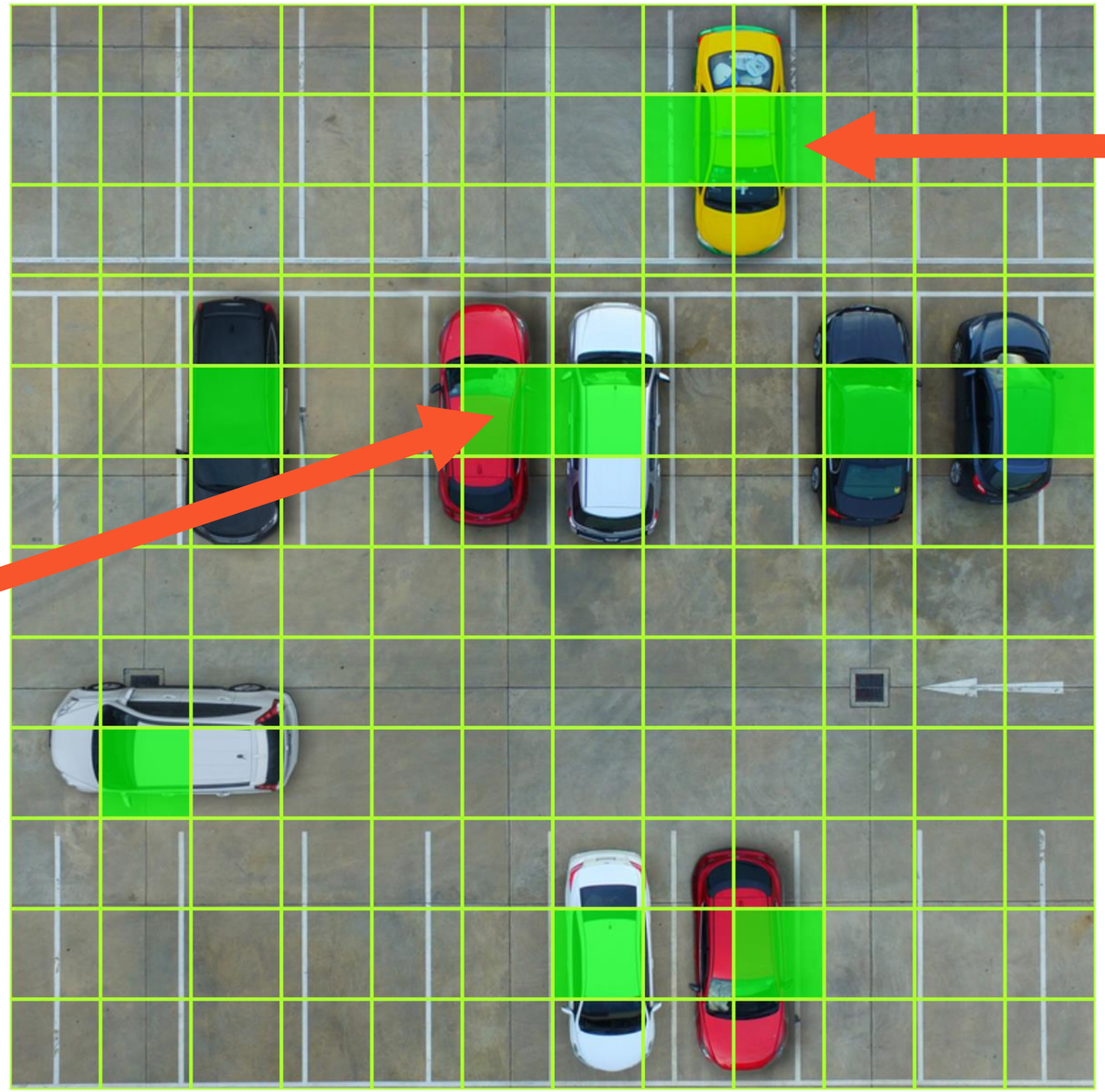
© 2022 Edge Impulse Inc.

Overlapping Objects?



EDGE IMPULSE

Fixed and Always Square



Potential issue, merge cells?

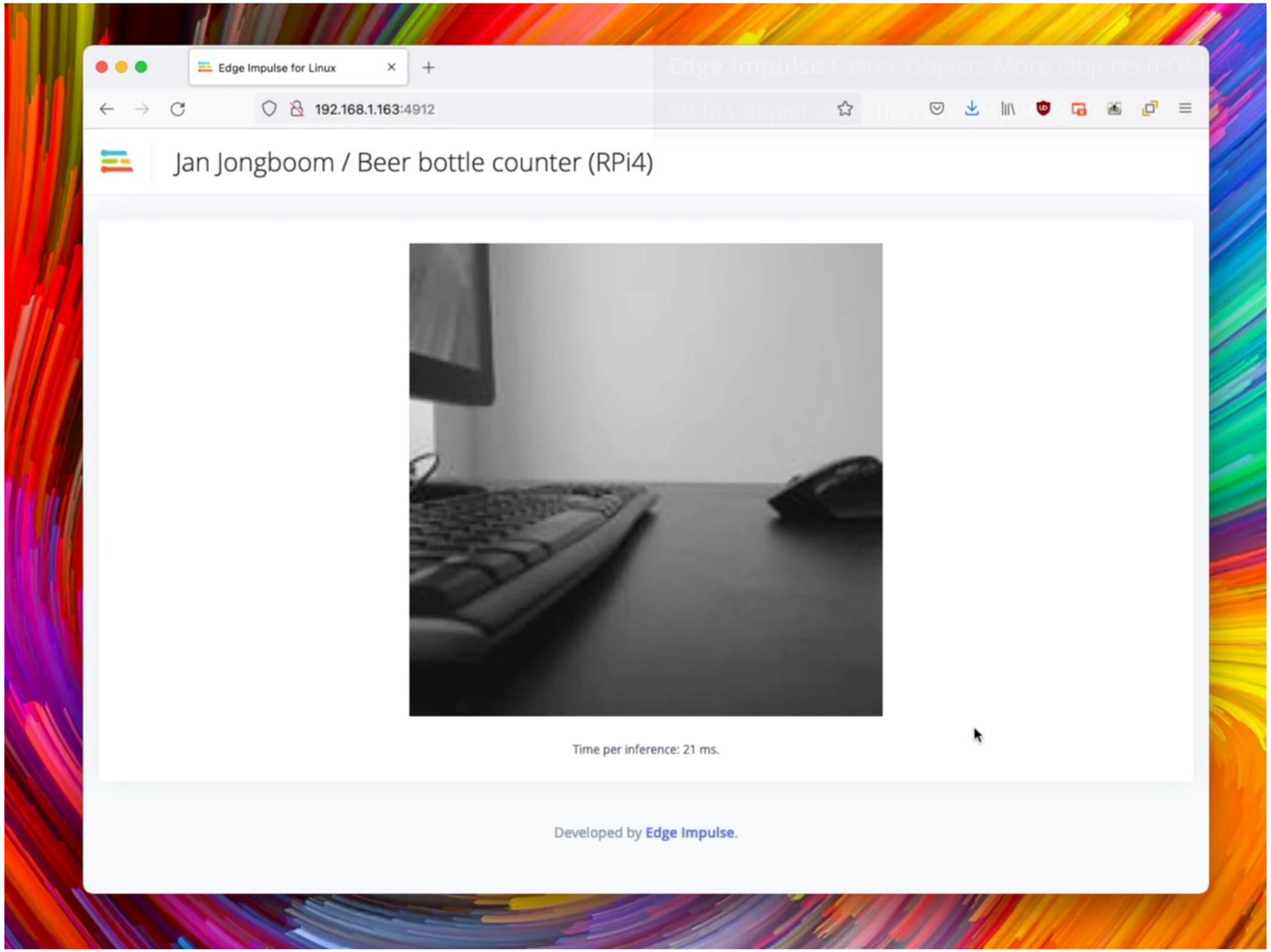
But... merging leads to issue here

EDGE IMPULSE

Centroids Should Have Meaning

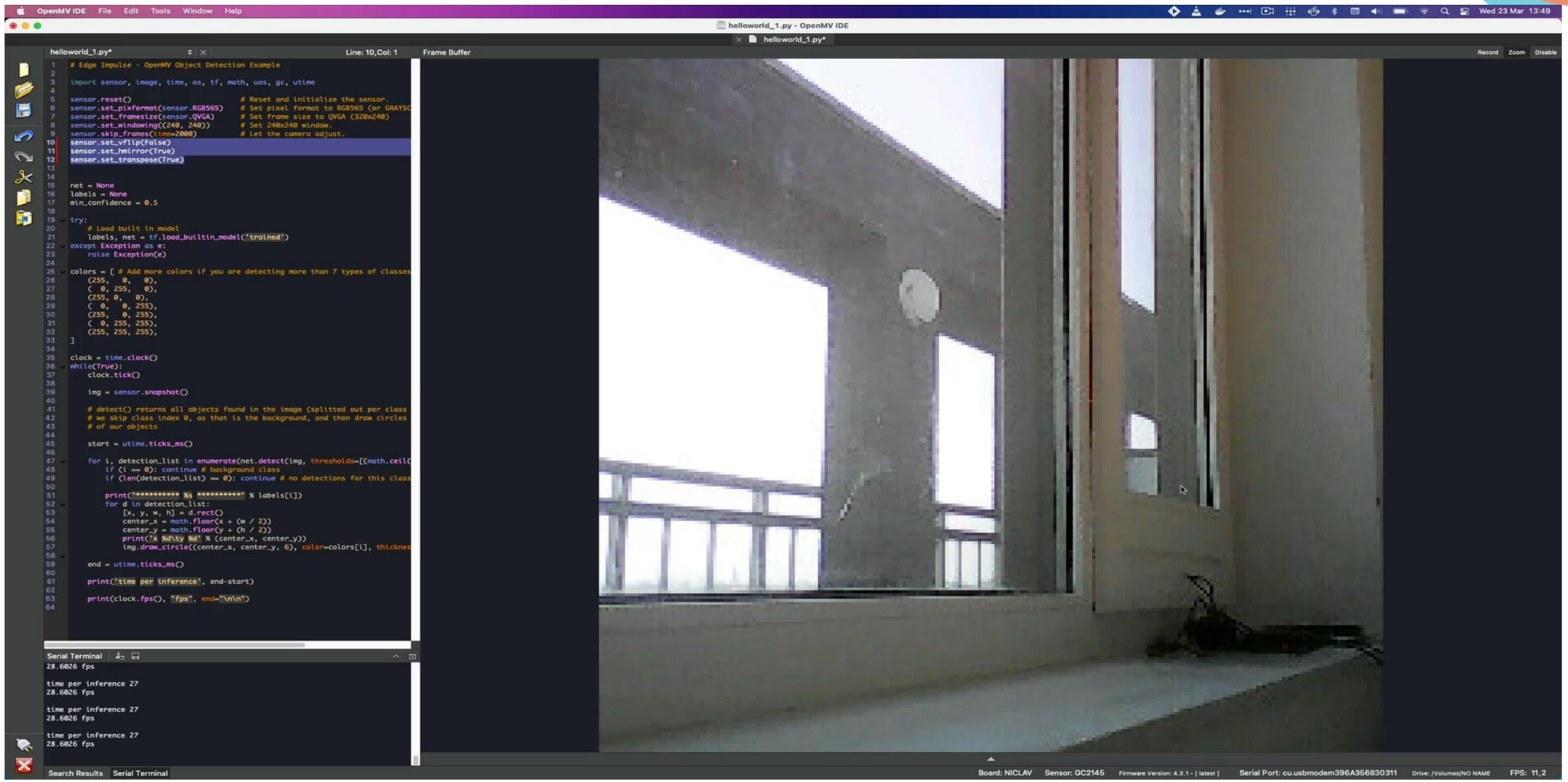
© 2022 Edge Impulse Inc.

Demo (Rpi4 on CPU, 160x160 input, 60 fps)



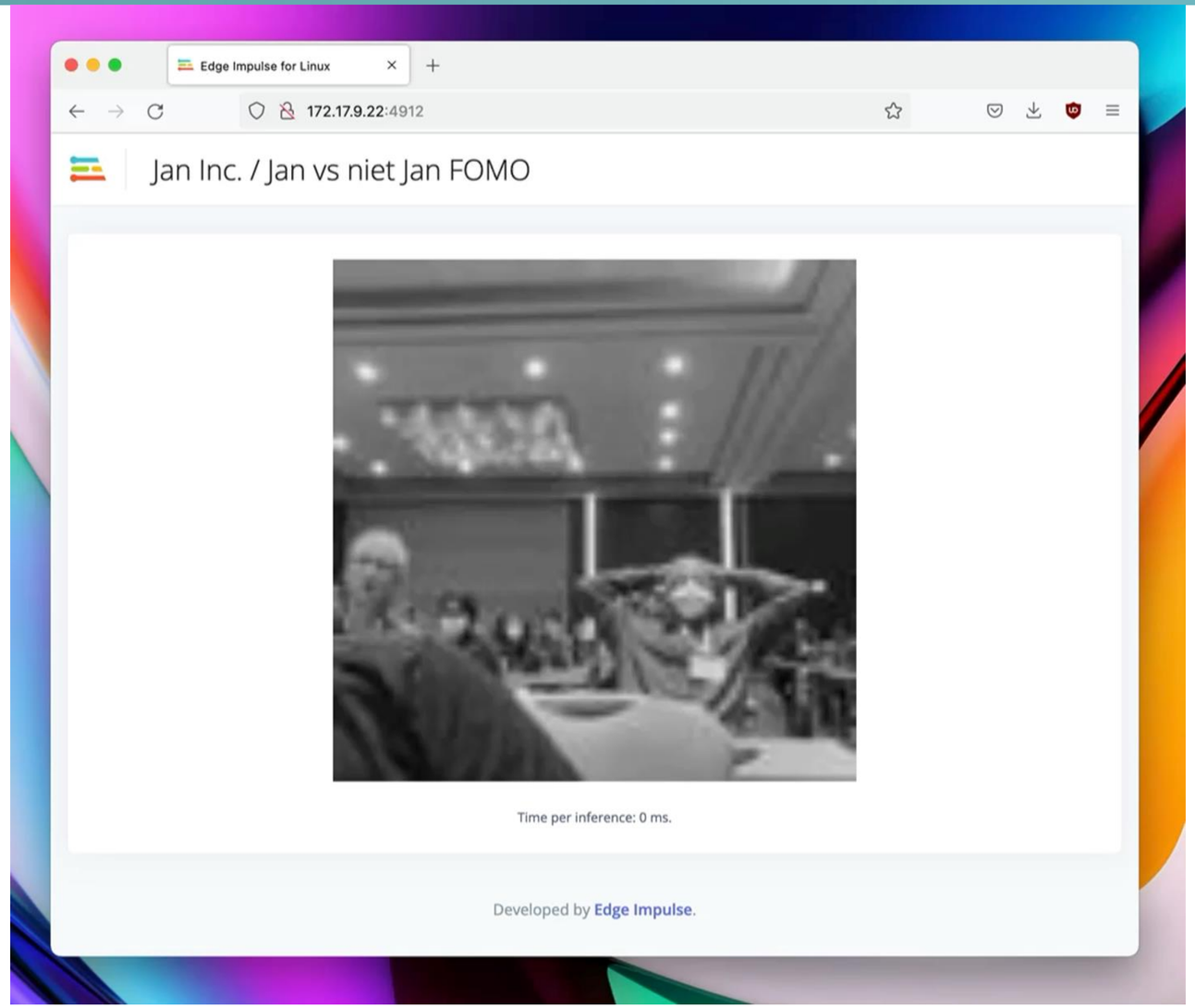
© 2022 Edge Impulse Inc.

Demo (Cortex-M7, 96x96 input, 28 fps)



© 2022 Edge Impulse Inc.

Demo (M1, 160x160, 1000 fps)



© 2022 Edge Impulse Inc.

Other Cool Features of FOMO

Add-on to any convolutional image network (incl. transfer learning models), so highly configurable

Fully convolutional, just the ratio is set

Can count objects

Can be a segmentation model

Performs much better on small objects than YOLOv5 or MobileNet SSD

© 2022 Edge Impulse Inc.

Performance

Cortex-M7 @ 480MHz: **30 fps** (96x96 MobileNetV2 a=0.1)

Raspberry Pi 4: **60 fps** (160x160 MobileNetV2 a=0.35)

Himax DSP @ 400MHz: **14 fps** (96x96 MobileNetV2 a=0.35)

Cortex-M4F @ 156MHz: **5 fps** (96x96 MobileNetV2 a=0.05)

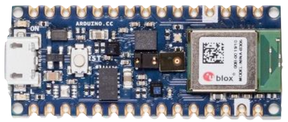
My Macbook: **1000 fps** :-)

(Can be bolted on other CNNs, e.g. MobileNetV1)

© 2022 Edge Impulse Inc.

Getting Started

- FOMO is available today for free: https://edgeimpulse.com/fomo
- Very wide range of dev boards, from Cortex-M4F to Jetson N
- Deploy to any device that has a C++ compiler
- Or use your phone!
- Be one of today's 194 new projects!




© 2022 Edge Impulse Inc.

Recap

Vision is such a cool sense

Classification is cool, locality and count matters

Want a 30x increase in performance? Use FOMO!

edgeimpulse.com

© 2022 Edge Impulse Inc.

Questions?

Full docs:
https://docs.edgeimpulse.com

FOMO
https://docs.edgeimpulse.com/docs/fomo-object-detection-for-constrained-devices

We're hiring!
https://edgeimpulse.com/careers

More questions:
forum.edgeimpulse.com / jan@edgeimpulse.com

© 2022 Edge Impulse Inc.