Visual Anomaly Detection with FOMO-AD

Jan Jongboom

Co-founder & CTO

Edge Impulse

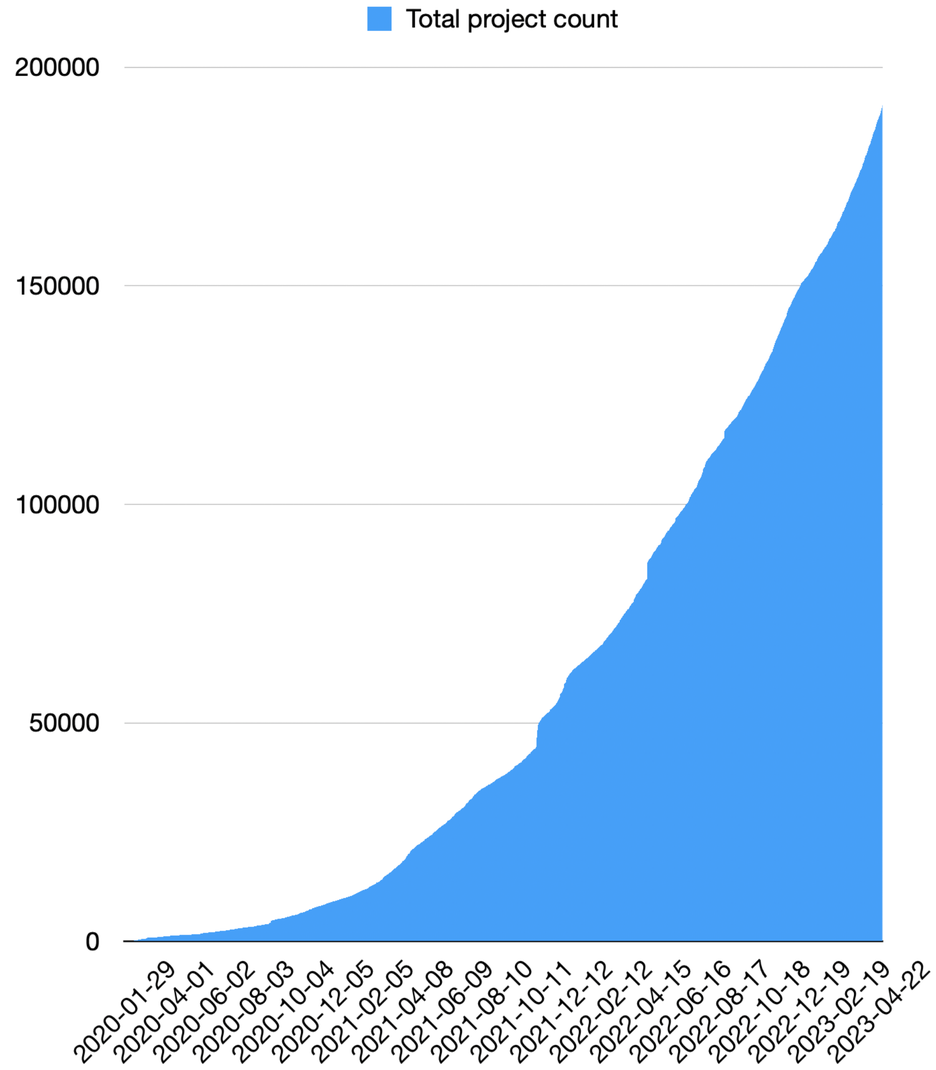
EDGE IMPULSE

Edge Impulse

Leading development platform for machine learning on edge devices

103,933 new projects (!) created since last Embedded Vision Summit

40% of these are vision projects



Edge Impulse project count

Can you trust ML models?

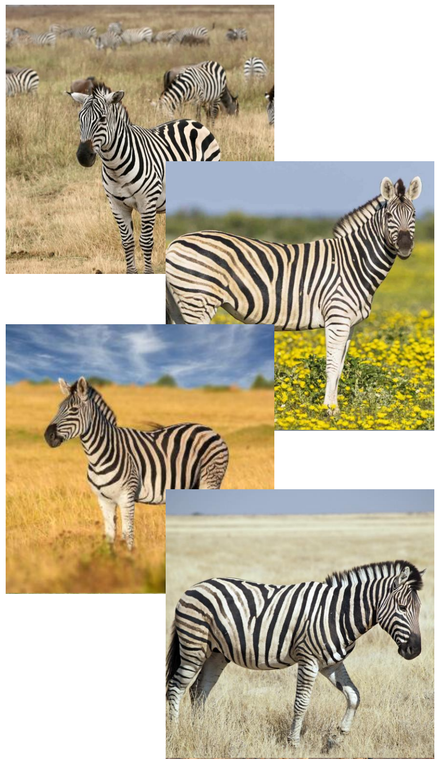
Day 3 — Llama detected!

Ernest Mwebaze gave an inspiring and witty tutorial on computer vision and deep learning. After training a convnet to recognize animals, he demonstrated the importance of using diverse and appropriate training data: A photo of a llama in the conservancy was misclassified as a giraffe (there aren't many llamas in East Africa).

https://medium.com/@damoncivin/arm-at-data-science-africa-2018-1071389e92d9

Countering with an 'unknown' state

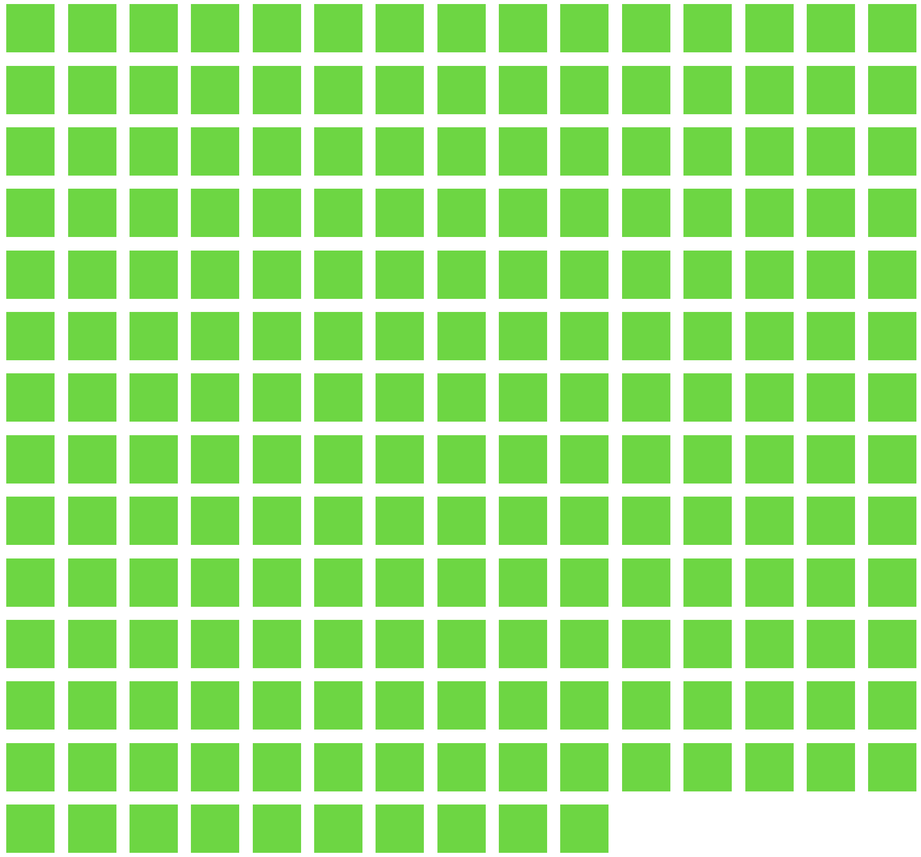
Giraffe



Zebra

Other

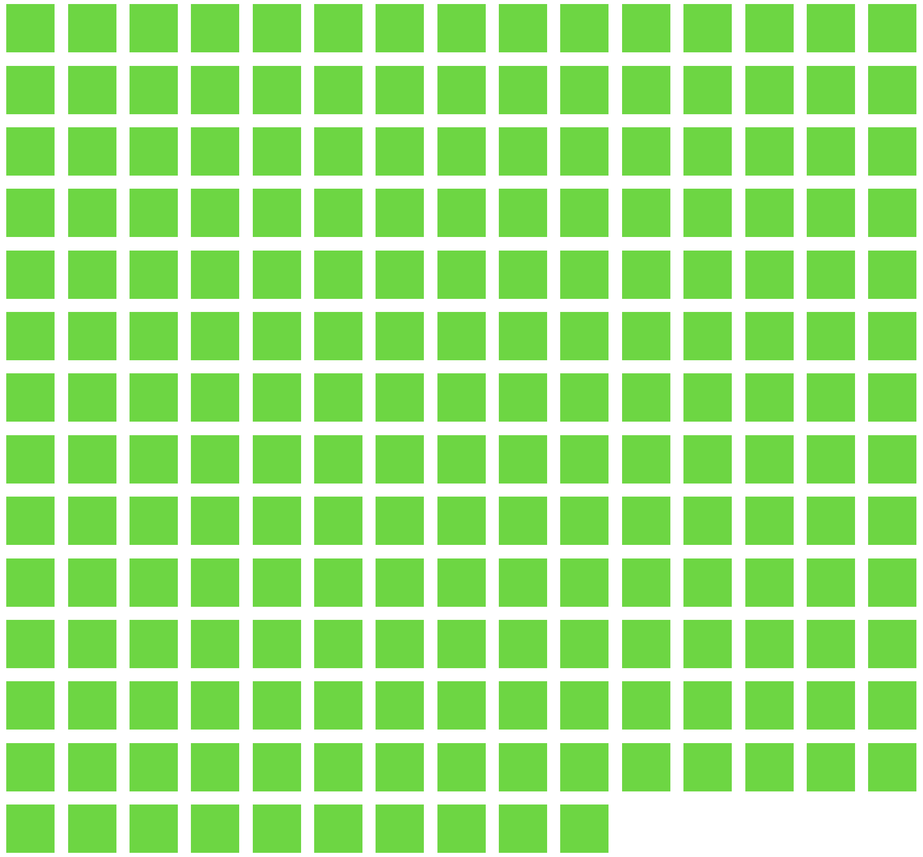
Dataset asymmetry



Normal operation

Faulty operation

Dataset asymmetry



Normal operation

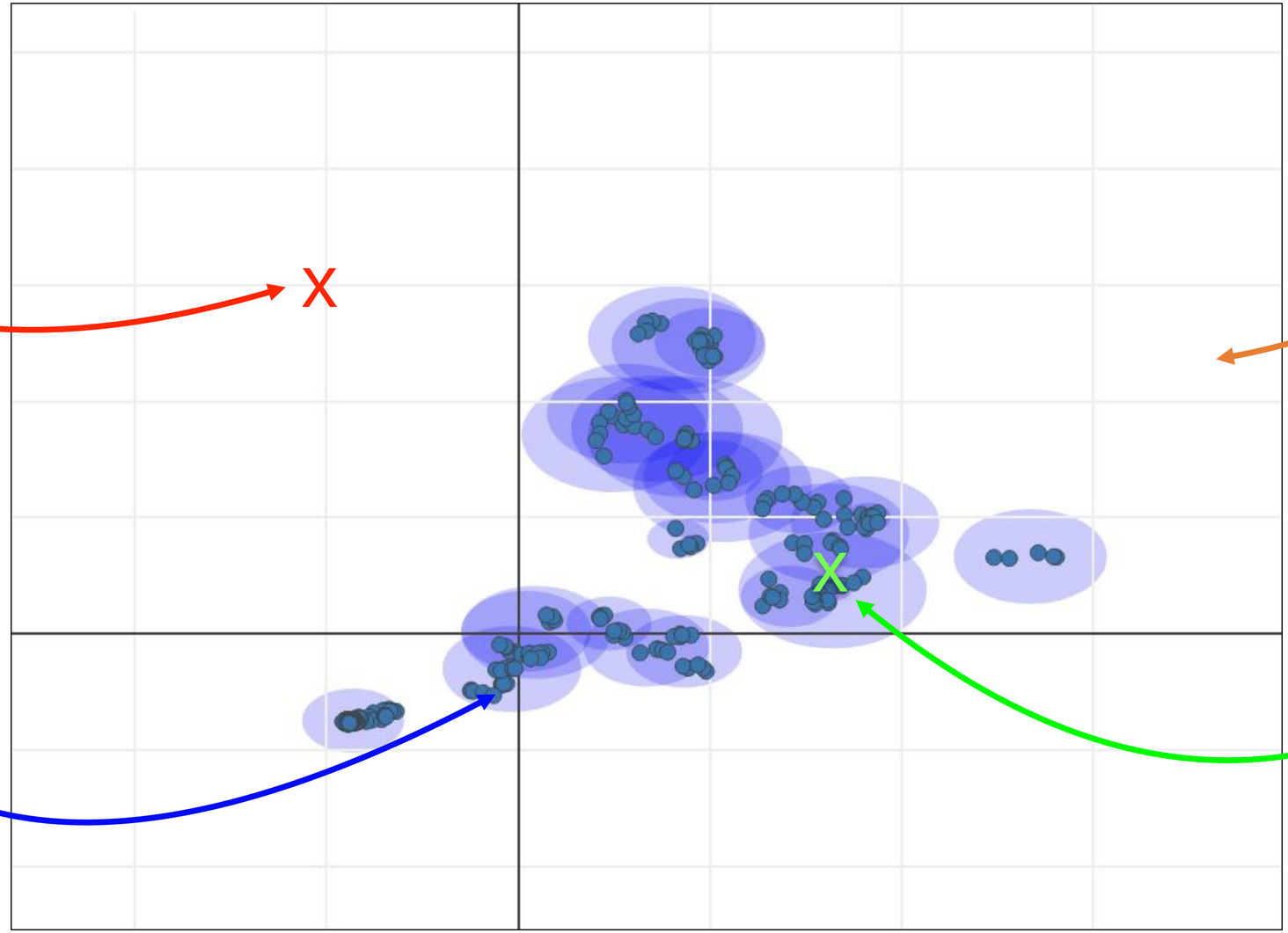
Fault state 1

Fault state 2

Fault state 3

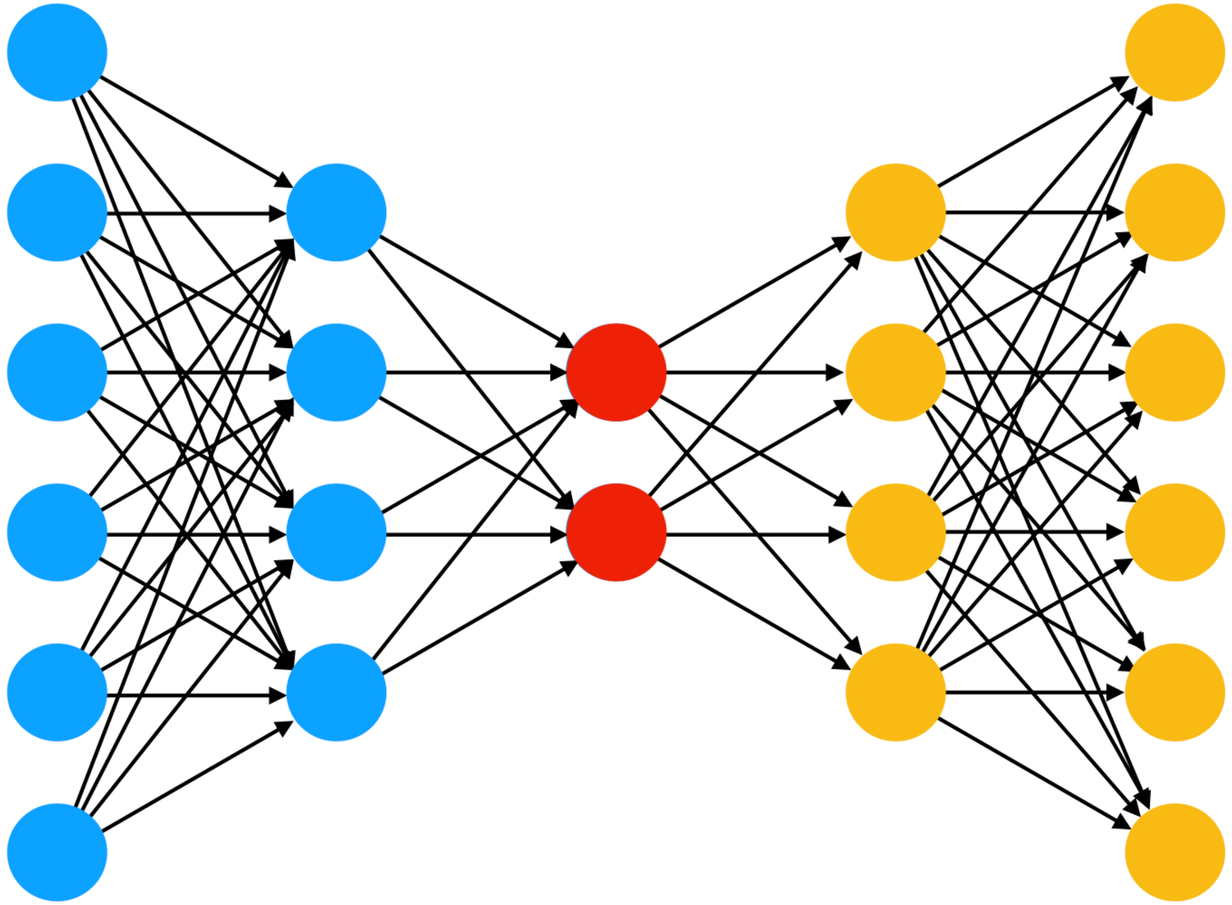
Fault state 4

Anomaly detection



Auto-encoders?

Input image



Reconstructed image

diff w/ input image:
similar? no anomaly.

Why auto-encoders don't work

- Computationally expensive, need both encoder/decoder.
- Working in pixel space is not great: poor evaluation metric, blurry images.
- Visual anomaly detection requires very high resolution images.

- Same accuracy: 10^6 parameters (auto-encoder) vs 10^3 parameters (our new approach).

Anomaly detection on sensor data

Input

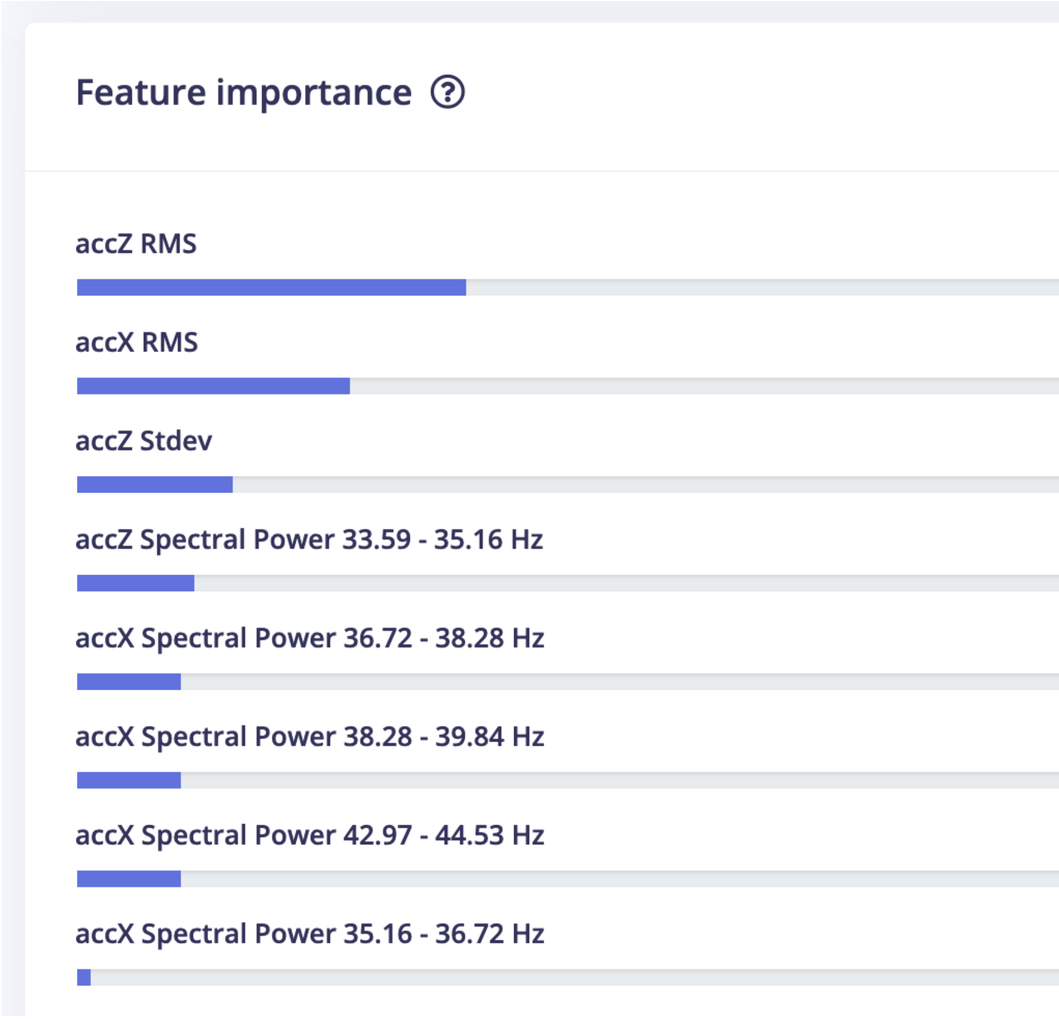
↓

DSP Features

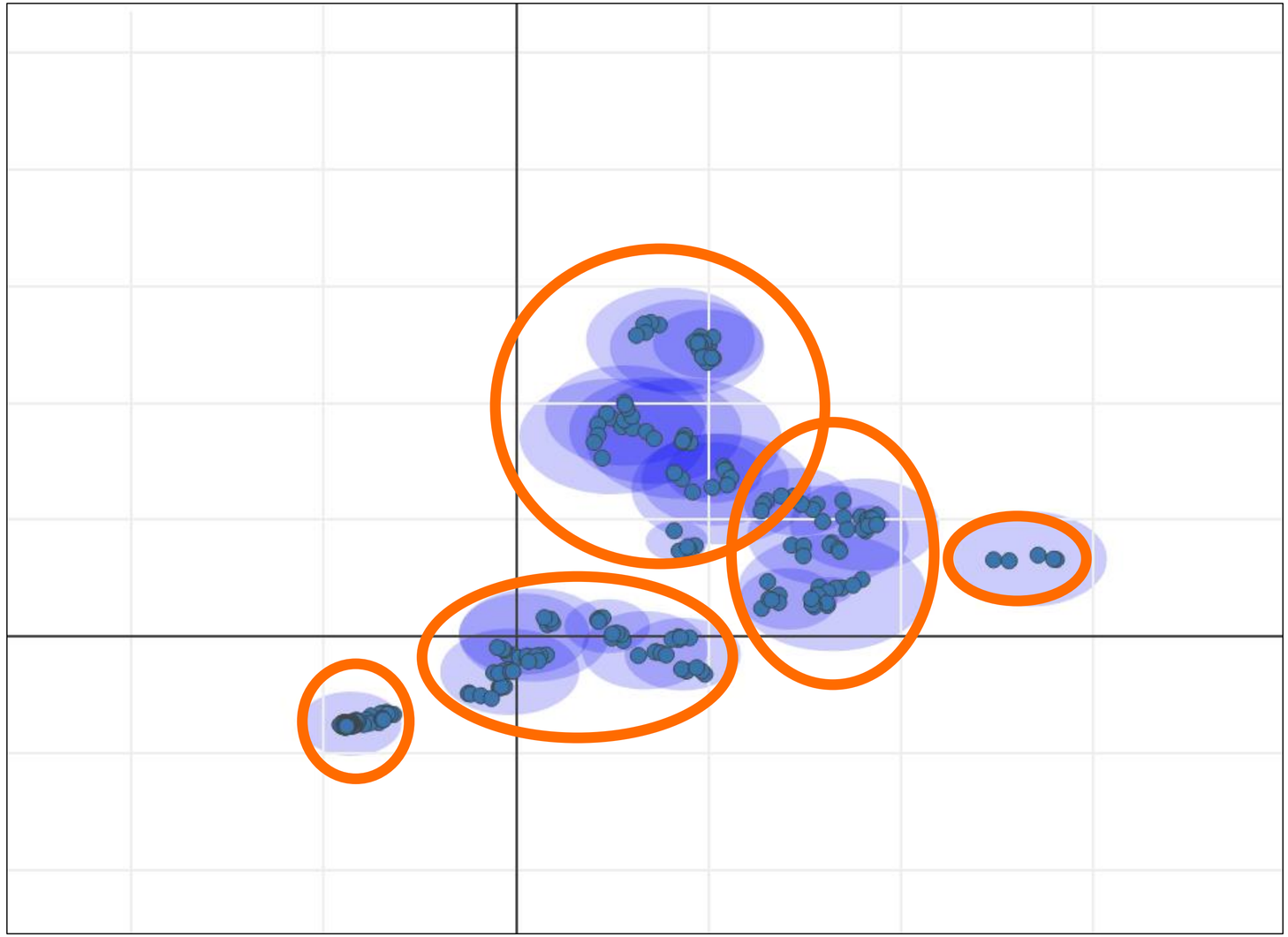
↓

Clustering algorithm

Great for basic sensor data
for which you can reason about features



Clustering with Gaussian Mixture Models



Applying this to visual AD

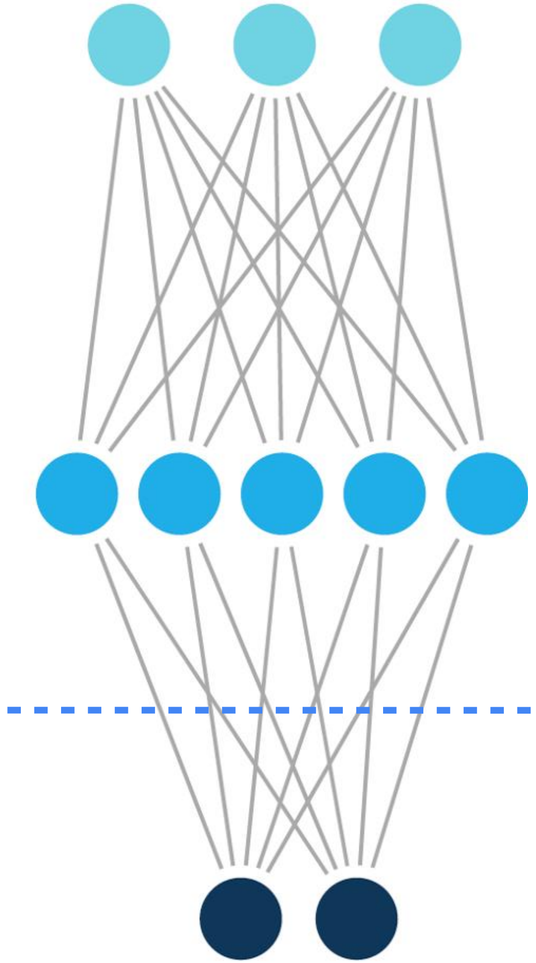
Input

↓

Neural network embeddings

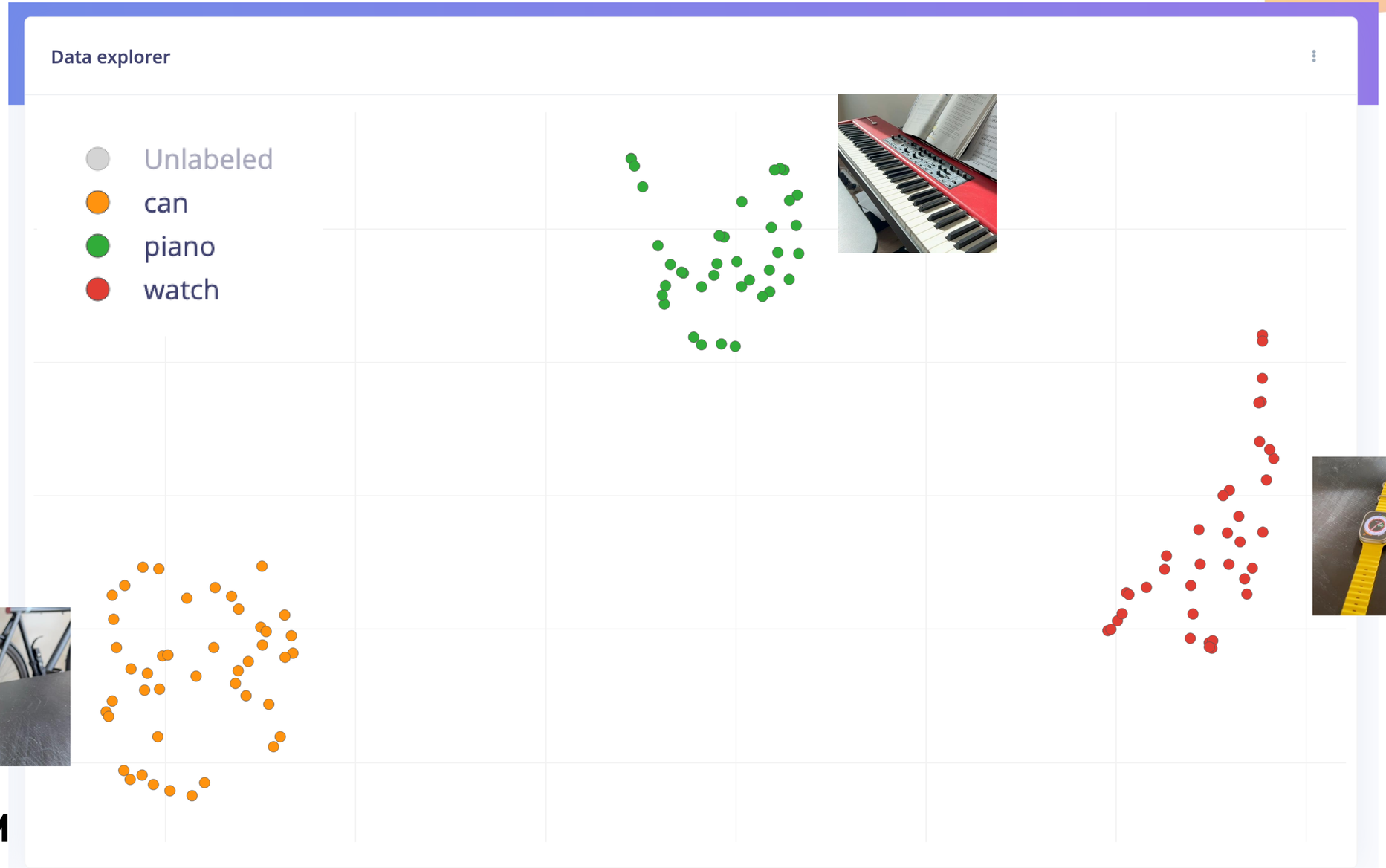
↓

Gaussian Mixture Models

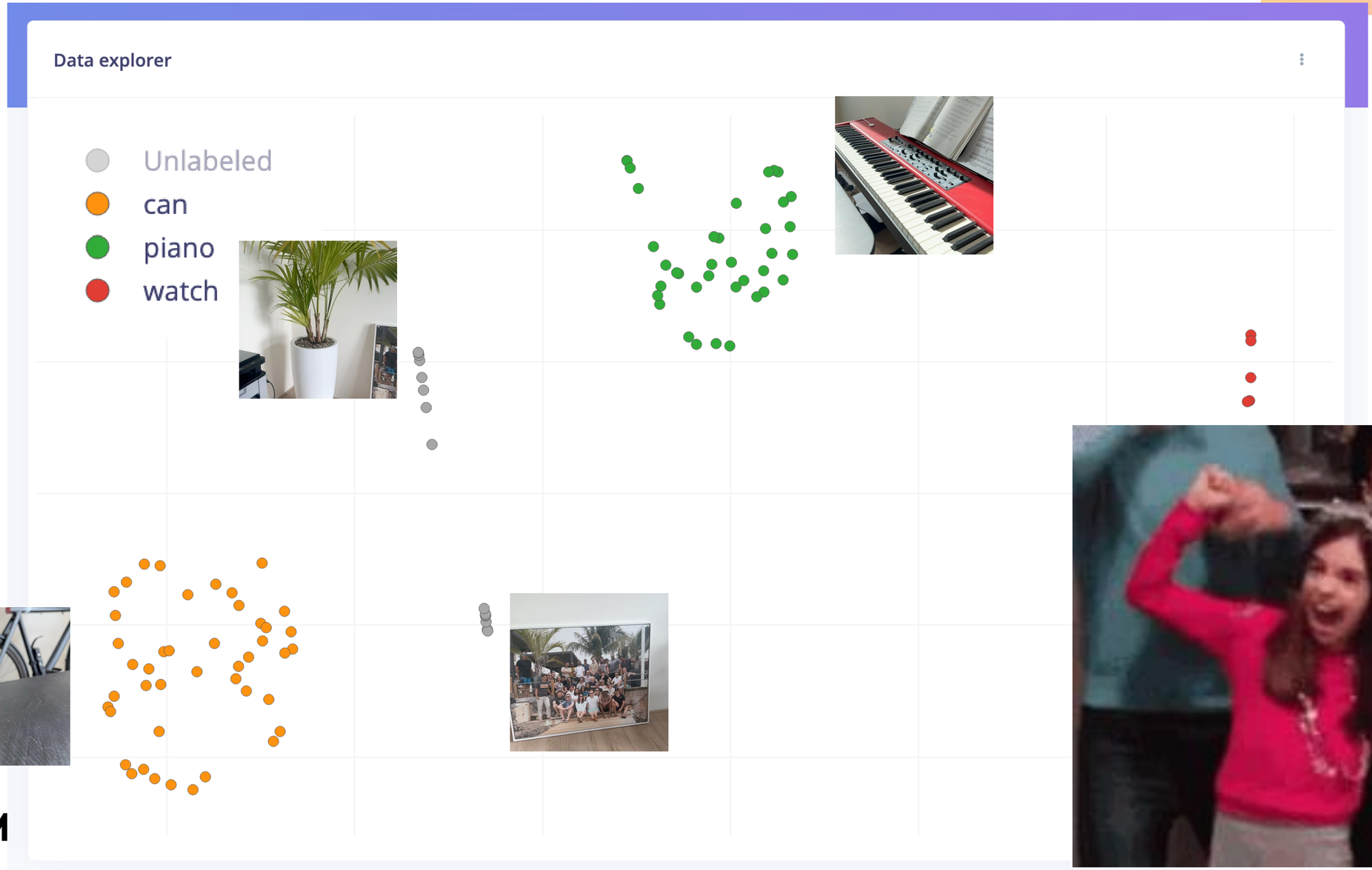


👍 Operates on feature space, not pixel space

Testing out this premise: MobileNet



Testing out this premise: MobileNet

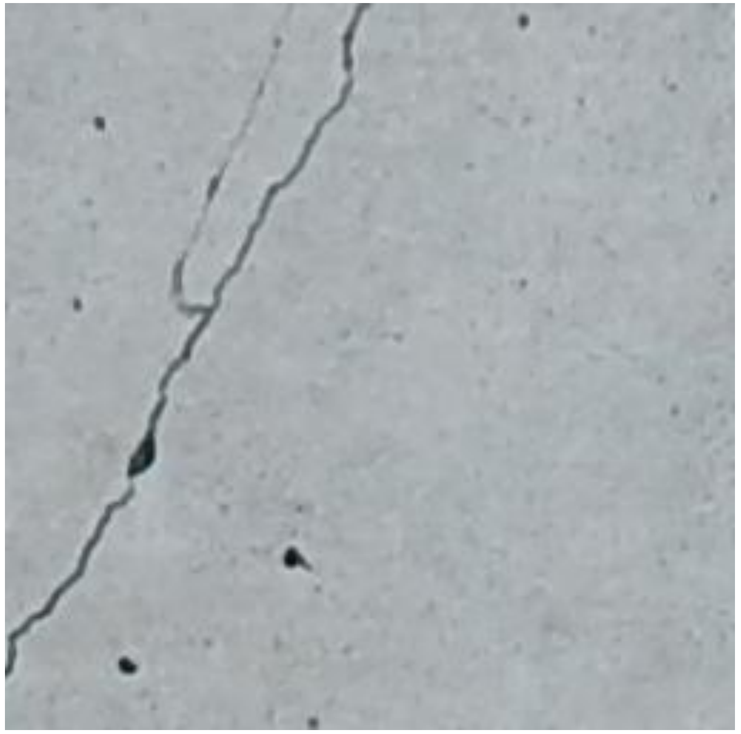


Where is the anomaly?

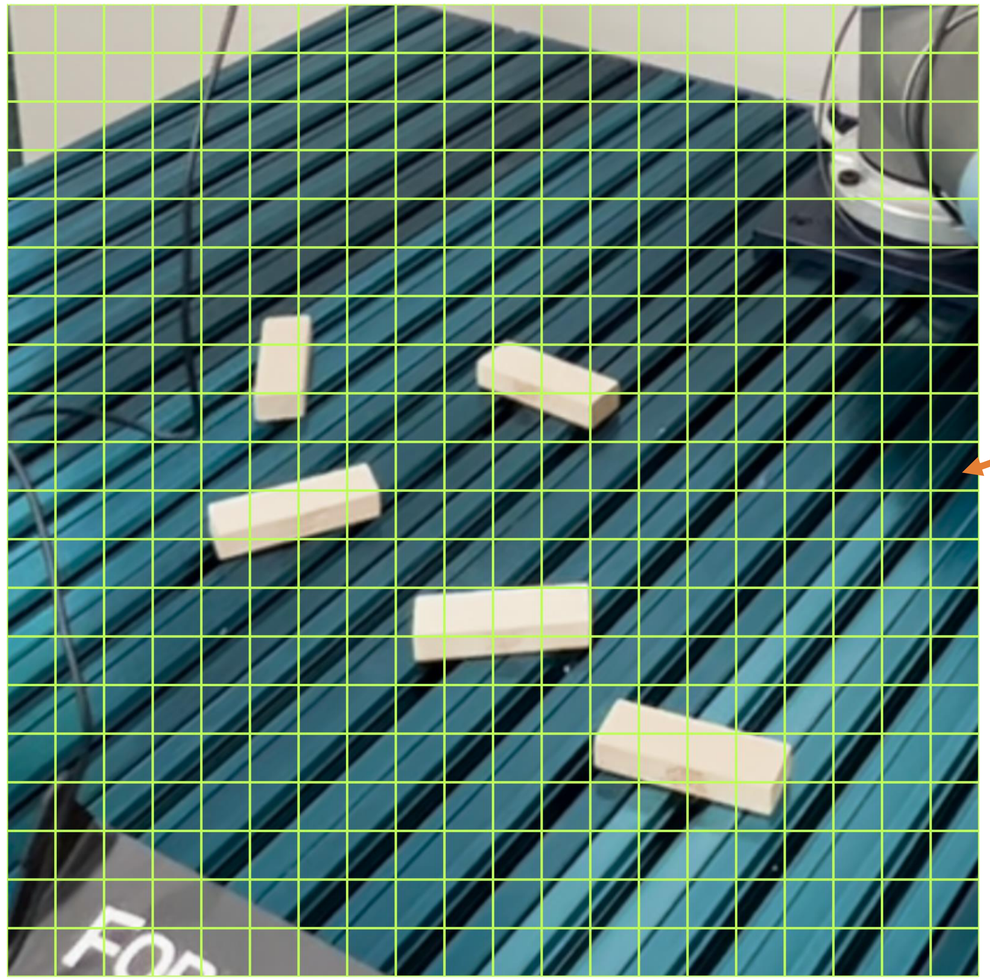
In lots of visual inspection cases only part of the image is anomalous.

Input image might be very large, if only 0.5% of your image covers an anomaly => hard to get your loss function right.

Knowing where a fault is, is super useful for humans!

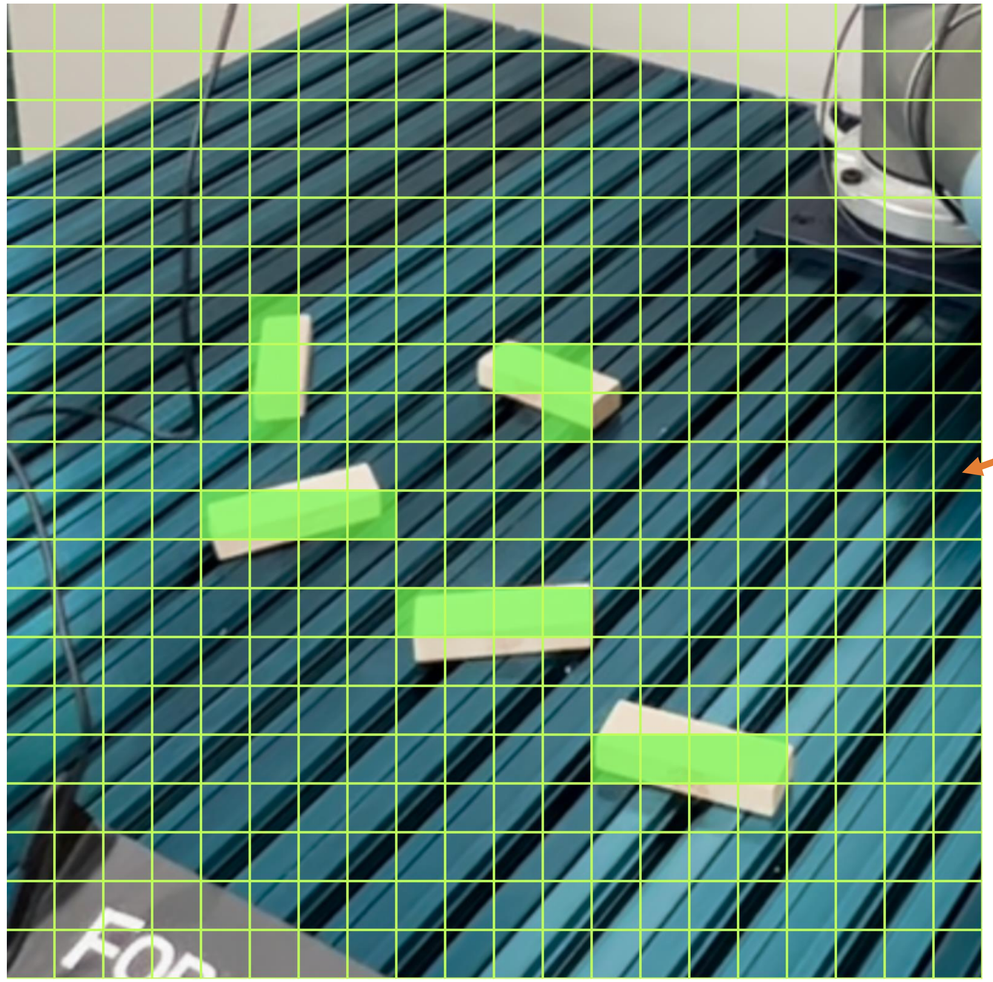


Edge Impulse FOMO



Each cell is a classifier

Edge Impulse FOMO



Each cell is a classifier

FOMO-AD

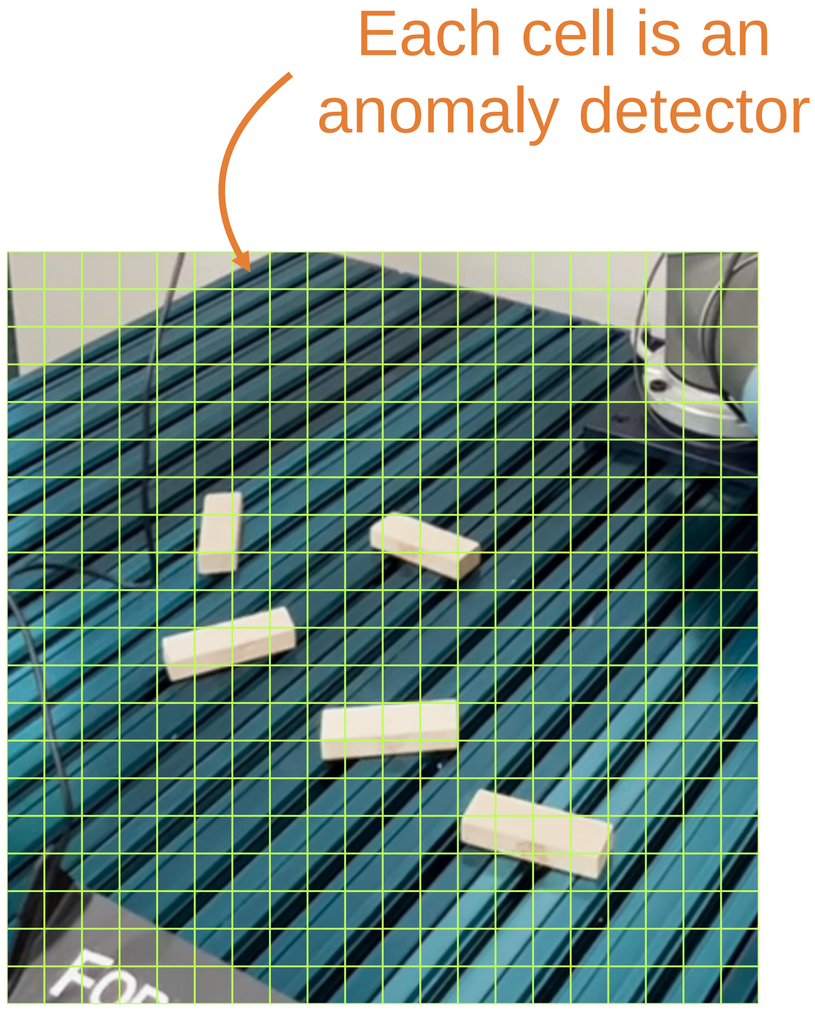
Replace classification with a GMM per cell.

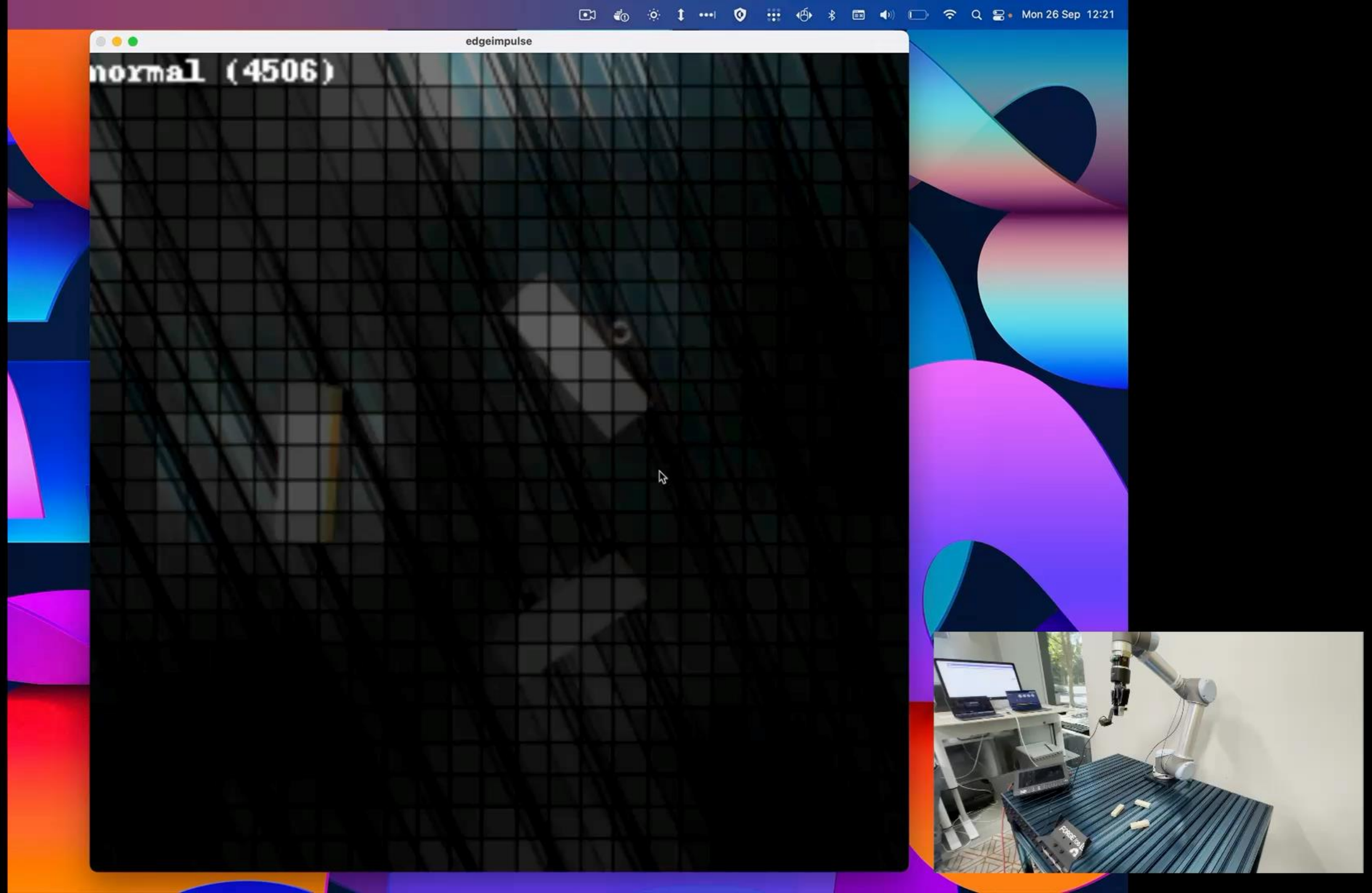
Each cell now has an anomaly score.

Fully convolutional, can train on patches of data.

Can train on only non-anomalous data.

Similar performance as FOMO:
Up to 30 fps on Cortex-M7, <200K RAM




edgeimpulse
normal (4506)
Mon 26 Sep 12:21

Performance

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

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

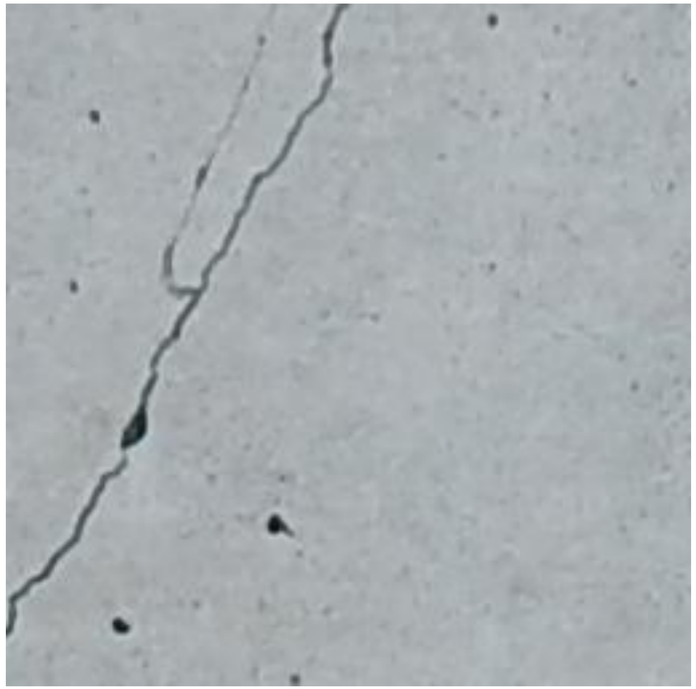
Himax DSP @ 400 MHz: **14 fps** (96x96 MobileNetV2 a=0.35)

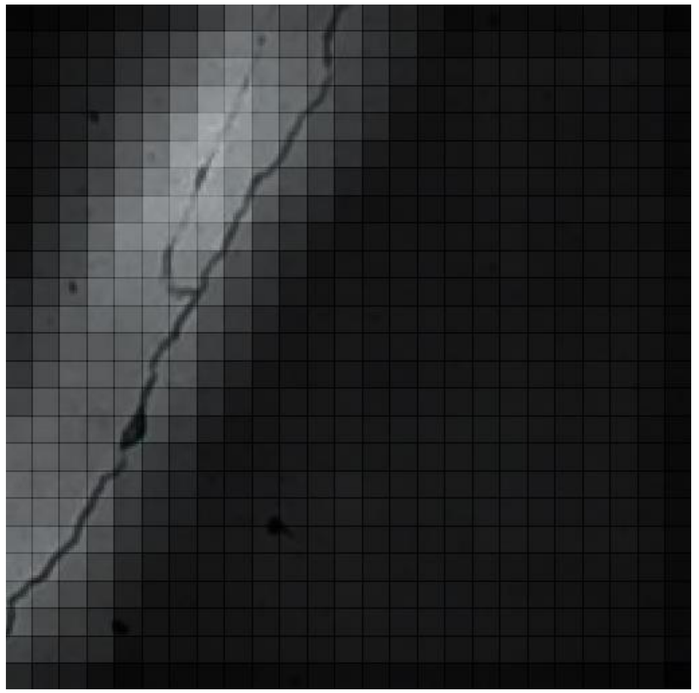
Cortex-M4F @ 156 MHz: **5 fps** (96x96 MobileNetV2 a=0.05)

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

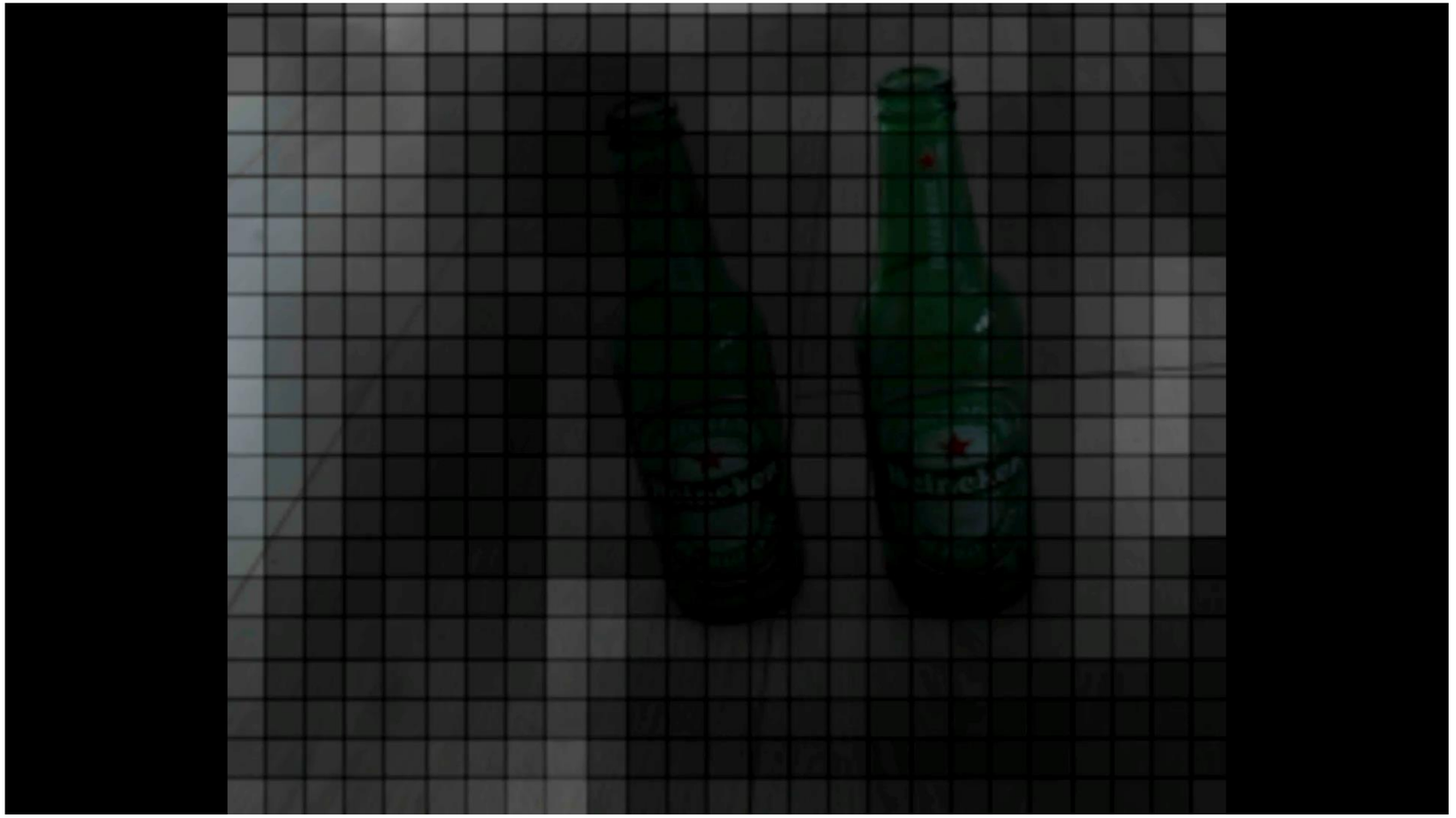
(Can be bolted on other CNNs, e.g. MobileNetV1)

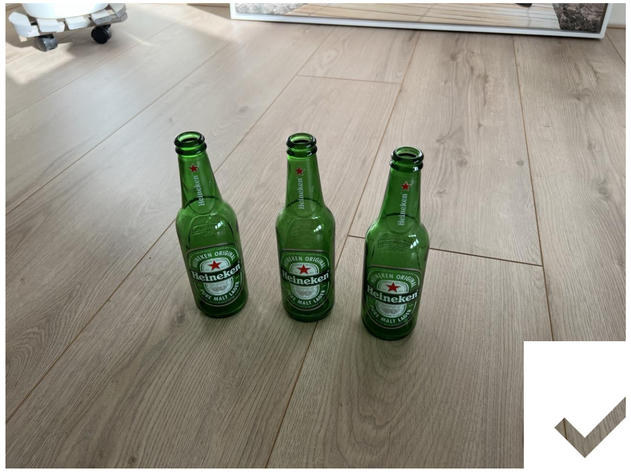
Example: crack QA

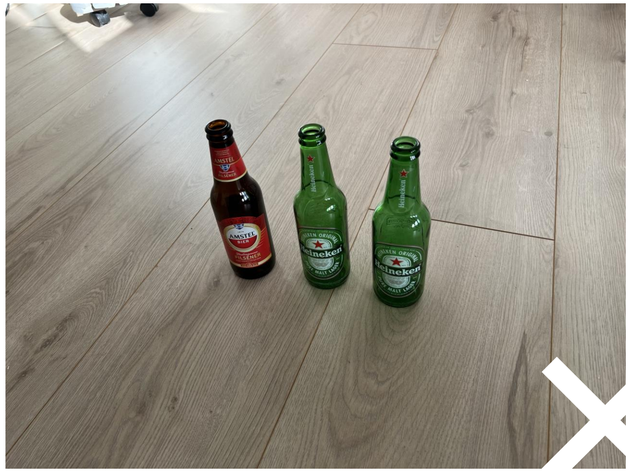




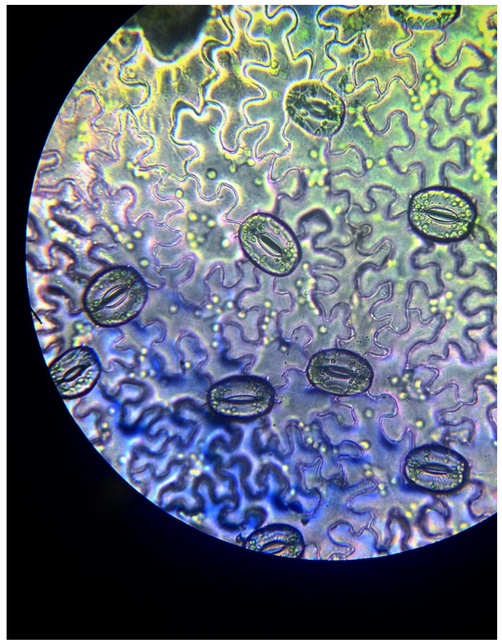
Example: production line monitoring







Where doesn't this work



Things that are not in ImageNet

Things larger than receptive field

Things that convolutions have trouble picking up

What we've learned

- Embeddings can have a large number of dimensions. Random projection helps cut this down, with little effect on accuracy.
- Calculating the embeddings can be shared between classifier and GMM. Just add two heads to your network.
- Training the anomaly detector can be done on the edge (if you have the compute power).
- Want to add custom code to a neural network graph (like an anomaly detector head)? Look at JAX[1].

[1] https://jax.readthedocs.io/en/latest/

Questions?

Wednesday 11:25AM

11:25am - 11:55am

Deploy Your Embedded Vision Solution on Any Processor Using Edge Impulse

By **Amir Sherman** Global Semiconductor Business Development Director, Edge Impulse

See live demos in our booth!

Docs: https://docs.edgeimpulse.com

Questions: jan@edgeimpulse.com