tinyML® Talks

Enabling Ultra-low Power Machine Learning at the Edge

“Constrained Object Detection on Microcontrollers with FOMO”

Shawn Hymel - Edge Impulse

April 5, 2022

www.tinyML.org

tinyML Talks Strategic Partners

Additional Sponsorships available – contact Olga@tinyML.org for info

TINY ML TALKS webcast

Executive Strategic Partners

arm AI

Powering tinyML Innovation

Arm AI Virtual Tech Talks

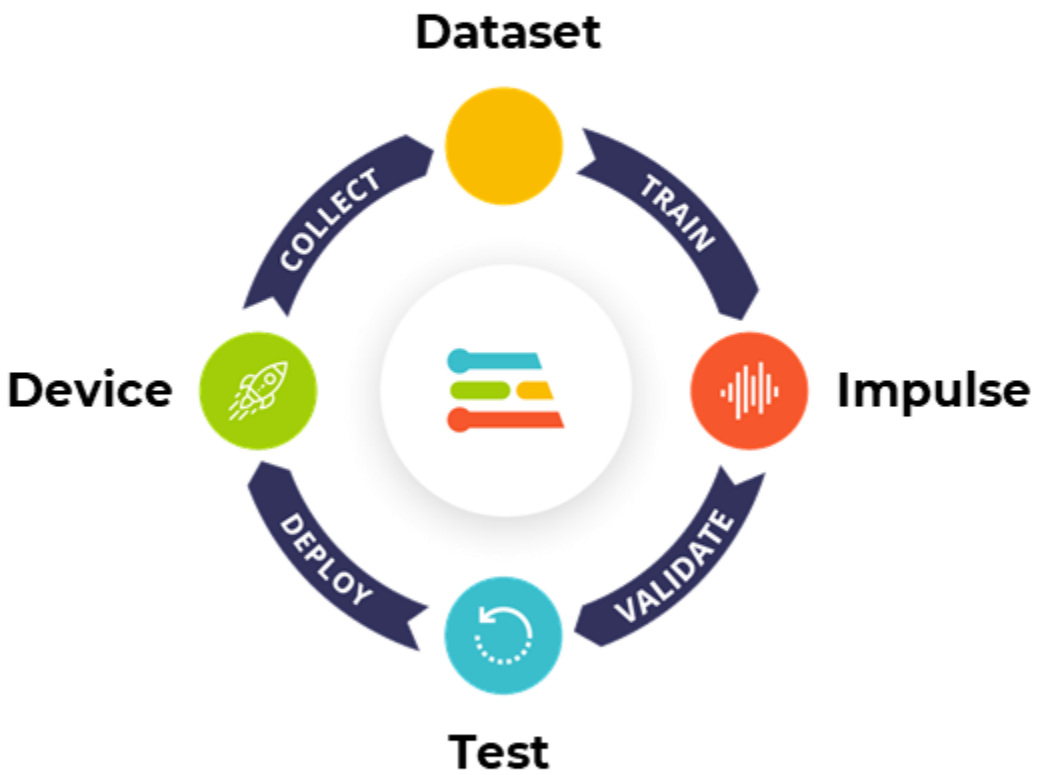
The latest in AI trends, technologies & best practices from Arm and our Ecosystem Partners.

Demos, code examples, workshops, panel sessions and much more!

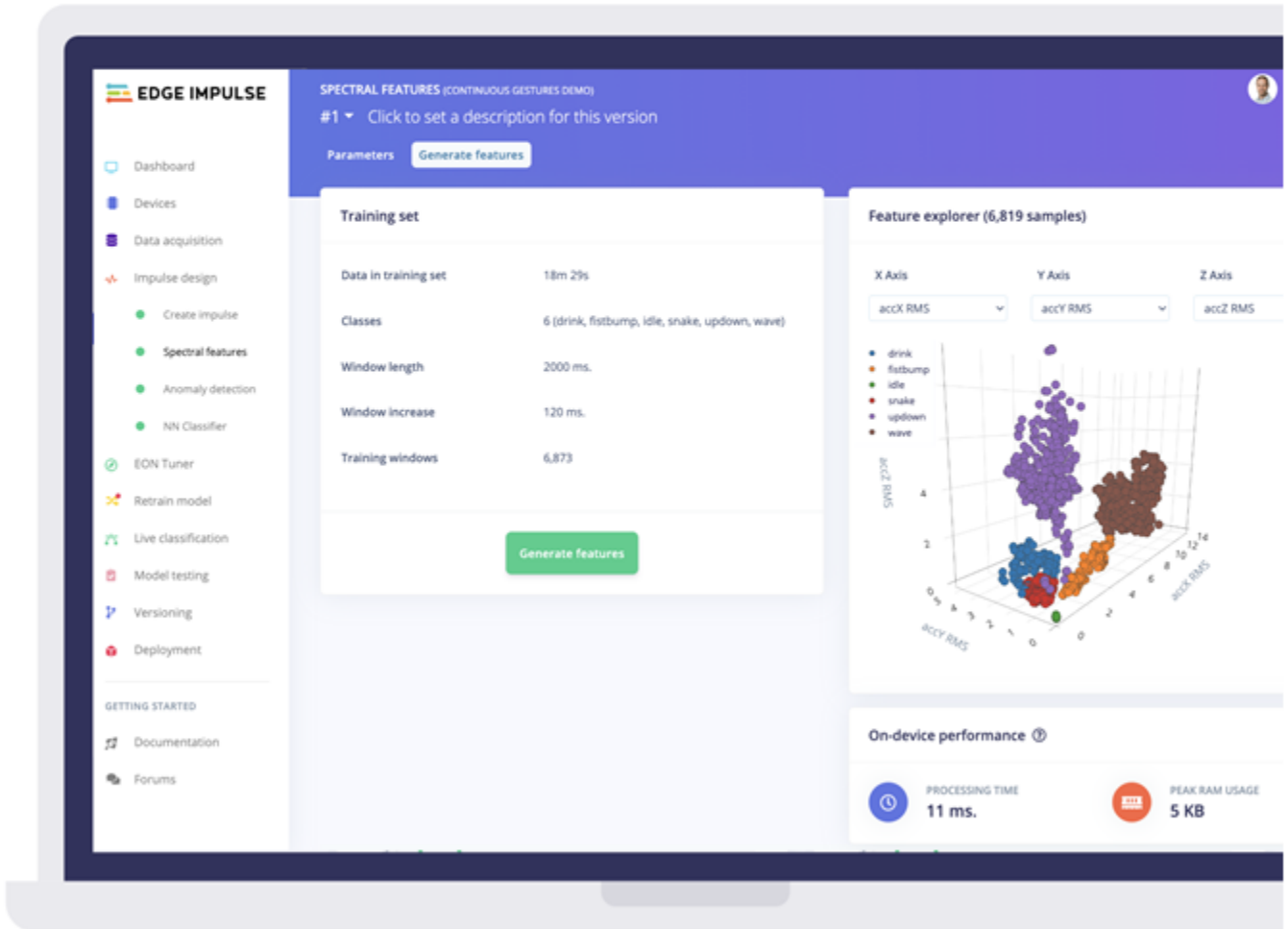
Fortnightly Tuesday @ 4pm GMT/8am PT

Find out more:
www.arm.com/techtalks

EDGE IMPULSE The leading edge ML platform



www.edgeimpulse.com



Qualcomm
AI research

Advancing AI research to make efficient AI ubiquitous

Power efficiency

Model design, compression, quantization, algorithms, efficient hardware, software tool

Personalization

Continuous learning, contextual, always-on, privacy-preserved, distributed learning

Efficient learning

Robust learning through minimal data, unsupervised learning, on-device learning

A platform to scale AI across the industry

Perception

Object detection, speech recognition, contextual fusion

Reasoning

Scene understanding, language understanding, behavior prediction

Action

Reinforcement learning for decision making

Edge cloud

Cloud

IoT/IIoT

Automotive

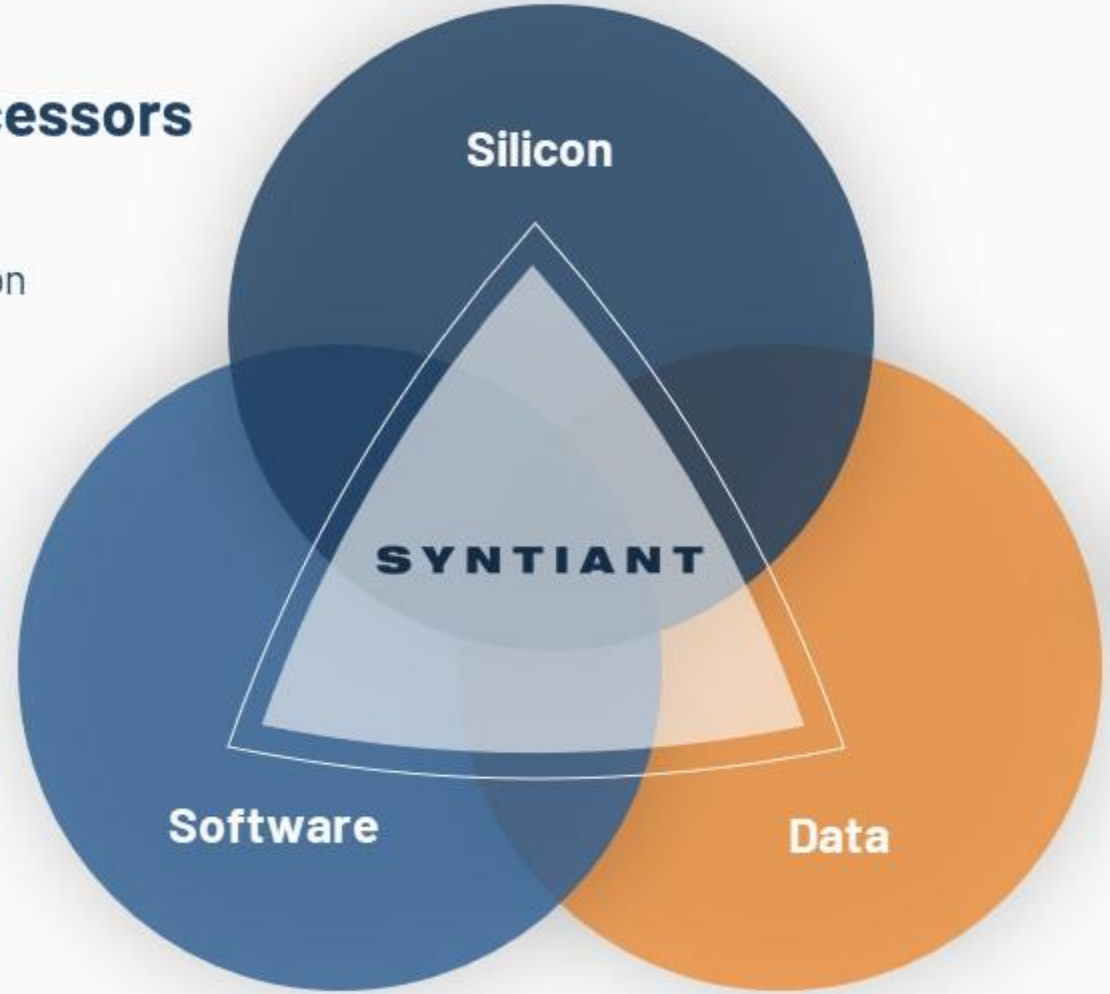
Mobile

SYNTIANT

End-to-End Deep Learning Solutions for TinyML & Edge AI

Neural Decision Processors

- At-Memory Compute
- Sustained High MAC Utilization
- Native Neural Network Processing



ML Training Pipeline

- Enables Production Quality Deep Learning Deployments

Data Platform

- Reduces Data Collection Time and Cost
- Increases Model Performance

partners@syntiant.com

www.syntiant.com

TINY ML TALKS webcast

Platinum Strategic Partners

TINY ML TALKS webcast

Deeplite

Fastest Video Analytics Solutions on Arm CPUs

AI Trailblazers Award Winner

TINY ML TALKS webcast

KLIKA·TECH

GLOBAL IOT SOLUTIONS

Reality AI®

Add Advanced Sensing to your Product with Edge AI / TinyML

https://reality.ai info@reality.ai @SensorAI Reality AI

Pre-built Edge AI sensing modules, plus tools to build your own

Reality AI solutions

Prebuilt sound recognition models for indoor and outdoor use cases

Solution for industrial anomaly detection

Pre-built automotive solution that lets cars "see with sound"

Reality AI Tools® software

Build prototypes, then turn them into real products

Explain ML models and relate the function to the physics

Optimize the hardware, including sensor selection and placement

BROAD AND SCALABLE EDGE COMPUTING PORTFOLIO

Microcontrollers & Microprocessors

Arm® Core

RA — Arm® Cortex®-M 32-bit MCUs
Arm ecosystem, Advanced security, Intelligent IoT

RZ — Arm®-based High-end 32 & 64-bit MPUs
High-resolution HMI, Industrial network & real-time control

RE — Arm® Cortex®-M0+ Ultra-low Power 32-bit MCUs
Innovative process tech (SOTB), Energy harvesting

Renesas Synergy™ — Arm®-based 32-bit MCUs for Qualified Platform
Qualified software and tools

Renesas Core

RL78 — Ultra-low Energy 8 & 16-bit MCUs
Bluetooth® Low Energy, SubGHz, LoRa®-based Solutions

RX — High Power Efficiently 32-bit MCUs
Motor control, Capacitive touch, Functional safety, GUI

RH850 — 40nm/28nm process Automotive 32-bit MCUs
Rich functional safety and embedded security features

Core technologies

AI

A broad set of high-power and energy-efficient embedded processors

Security & Safety

Comprehensive technology and support that meet the industry's stringent standards

Digital & Analog & Power Solution

Winning Combinations that combine our complementary product portfolios

Cloud Native

Cross-platforms working with partners in different verticals and organizations

BIG IDEAS FOR EVERY SPACE RENESAS

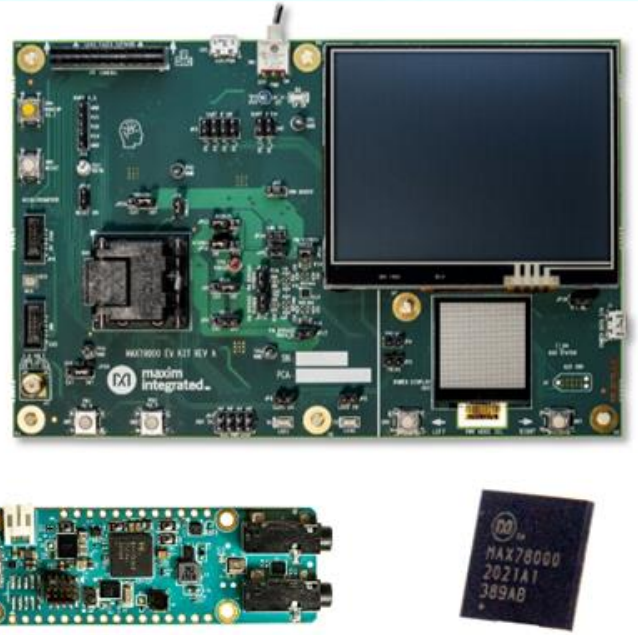
TINY ML TALKS webcast

Gold Strategic Partners

TINY ML TALKS webcast

Maxim Integrated: Enabling Edge Intelligence

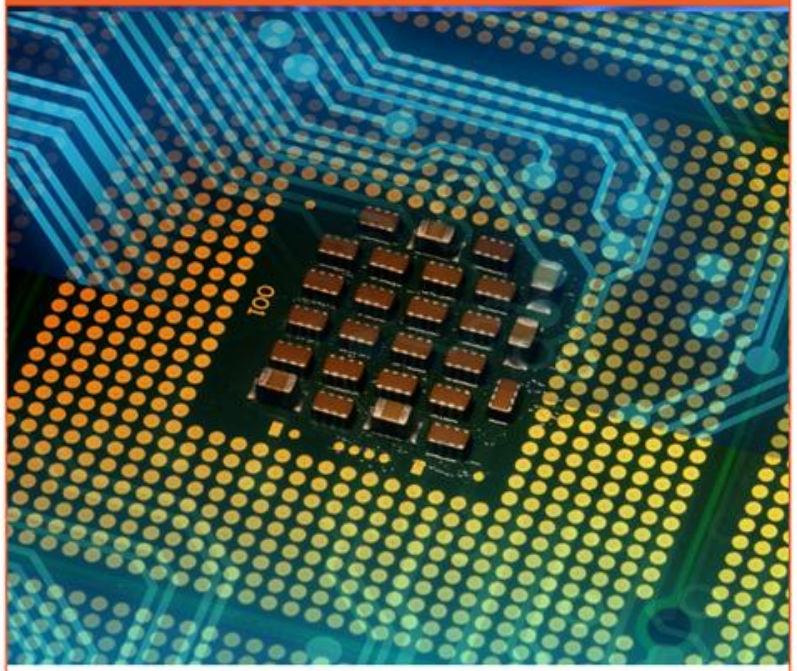
Advanced AI Acceleration IC



The new MAX78000 implements AI inferences at low energy levels, enabling complex audio and video inferencing to run on small batteries. Now the edge can see and hear like never before.

www.maximintegrated.com/MAX78000

Low Power Cortex M4 Micros



Large (3MB flash + 1MB SRAM) and small (256KB flash + 96KB SRAM, 1.6mm x 1.6mm) Cortex M4 microcontrollers enable algorithms and neural networks to run at wearable power levels.

www.maximintegrated.com/microcontrollers

Sensors and Signal Conditioning

Health sensors measure PPG and ECG signals critical to understanding vital signs. Signal chain products enable measuring even the most sensitive signals.

www.maximintegrated.com/sensors

LatentAI

Adaptive AI for the Intelligent Edge

Latentai.com

TINY ML TALKS webcast

TINY ML TALKS webcast

seeed studio

The IoT Hardware Enabler

SensiML™

Build Smart IoT Sensor Devices From Data

SensiML pioneered TinyML software tools that auto generate AI code for the intelligent edge.

- End-to-end AI workflow
- Multi-user auto-labeling of time-series data
- Code transparency and customization at each step in the pipeline

We enable the creation of production-grade smart sensor devices.

sensiml.com

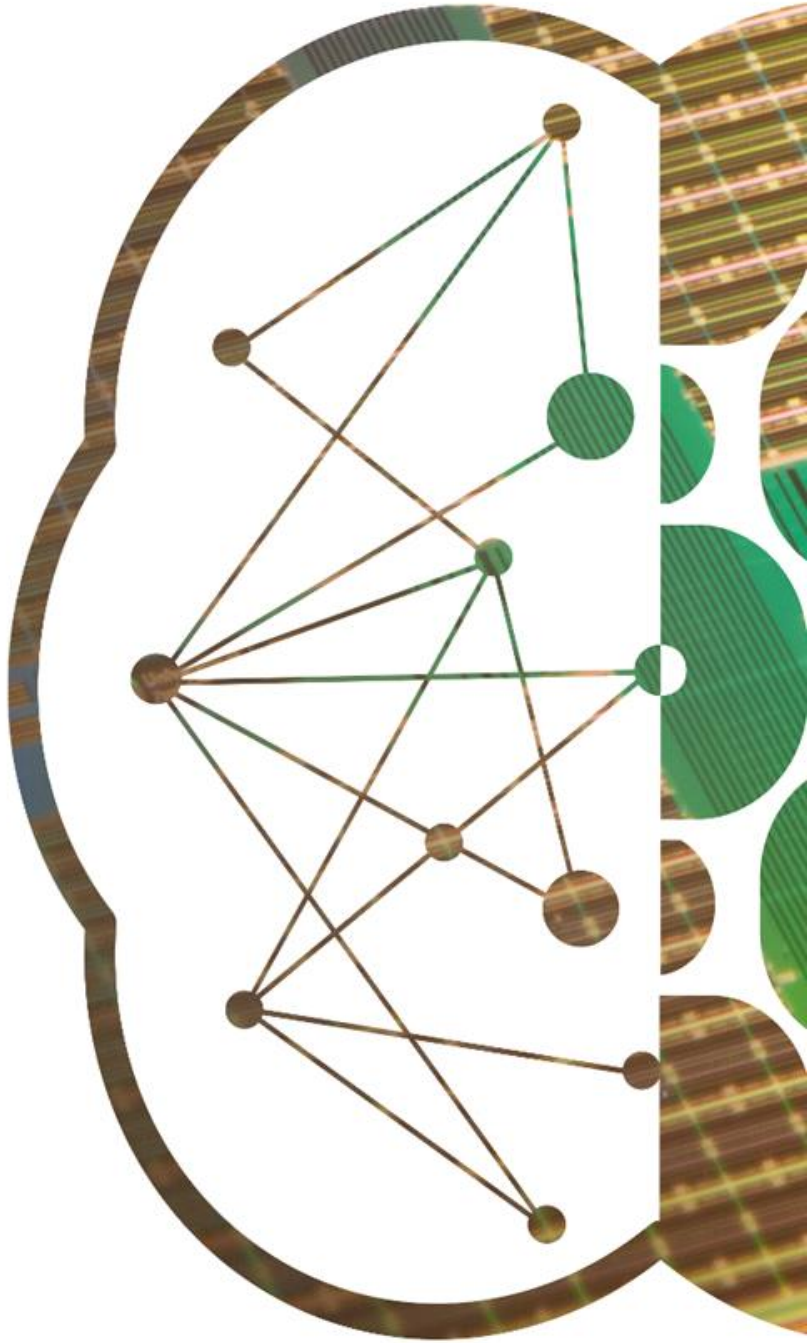
TINY ML TALKS webcast

life.augmented

SynSense

SynSense builds **sensing and inference** hardware for **ultra-low-power** (sub-mW) **embedded, mobile and edge** devices. We design systems for **real-time always-on smart sensing**, for audio, vision, IMUs, bio-signals and more.

https://SynSense.ai



TINY ML TALKS webcast

Silver Strategic Partners

AONdevIces

emza visual sense

GREENWAVES TECHNOLOGIES

Groovety Inc.

HOTG

imagimob

itemis

PROPHESEE METAVISION FOR MACHINES

Qeexo

REXEN technology

SAP

SA STREAM ANALYZE

TINY ML TALKS webcast

tinyML Summit 2022

Miniature dreams can come true…

March 28-30, 2022

Hyatt Regency San Francisco Airport

https://www.tinyml.org/event/summit-2022/

*The Best Product of the Year and the Best Innovation of the Year awards are open for nominations between **November 15 and March 14.***

tinyML Research Symposium 2022

March 28, 2022

https://www.tinyml.org/event/research-symposium-2022

More sponsorships are available: sponsorships@tinyML.org

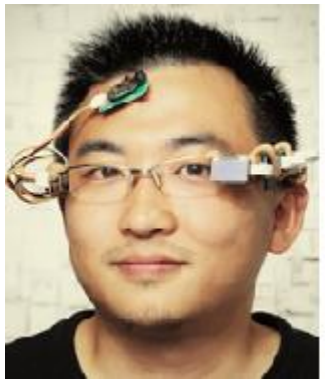
TINY ML TALKS webcast

Our next tinyML Trailblazers Series

Success Stories with Eric Pan

(Founder, Seeed Studio and Chaihuo Makerspace)

LIVE ONLINE **April 6th, 2022 at 8 am PST**



TINY ML TRAILBLAZERS

Register now!

Join Growing tinyML Communities:

8.8k members in
45 Groups in 35 Countries

tinyML - Enabling ultra-low Power ML at the Edge

https://www.meetup.com/tinyML-Enabling-ultra-low-Power-ML-at-the-Edge/

2.7k members
&
5.7k followers

The tinyML Community

https://www.linkedin.com/groups/13694488/

TINY ML TALKS webcast

Subscribe to tinyML YouTube Channel for updates and notifications

(including this video)

www.youtube.com/tinyML

YouTube

Next tinyML Talks

Date	Presenter	Topic / Title
Tuesday, April 5	Daniel Konegen, Embedded AI and data science engineer, Hahn-Schickard Marcus Rüb, Research assistant, Hahn-Schickard	AutoFlow - an open source Framework to automatically implement neural networks on embedded devices

Webcast start time is 9:30 am Pacific time

Please contact talks@tinyml.org if you are interested in presenting

TINY ML TALKS webcast

Reminders

Slides & Videos will be posted tomorrow

PDF

tinyml.org/forums

youtube.com/tinyml

Please use the Q&A window for your questions

TINY ML TALKS webcast

Shawn Hymel

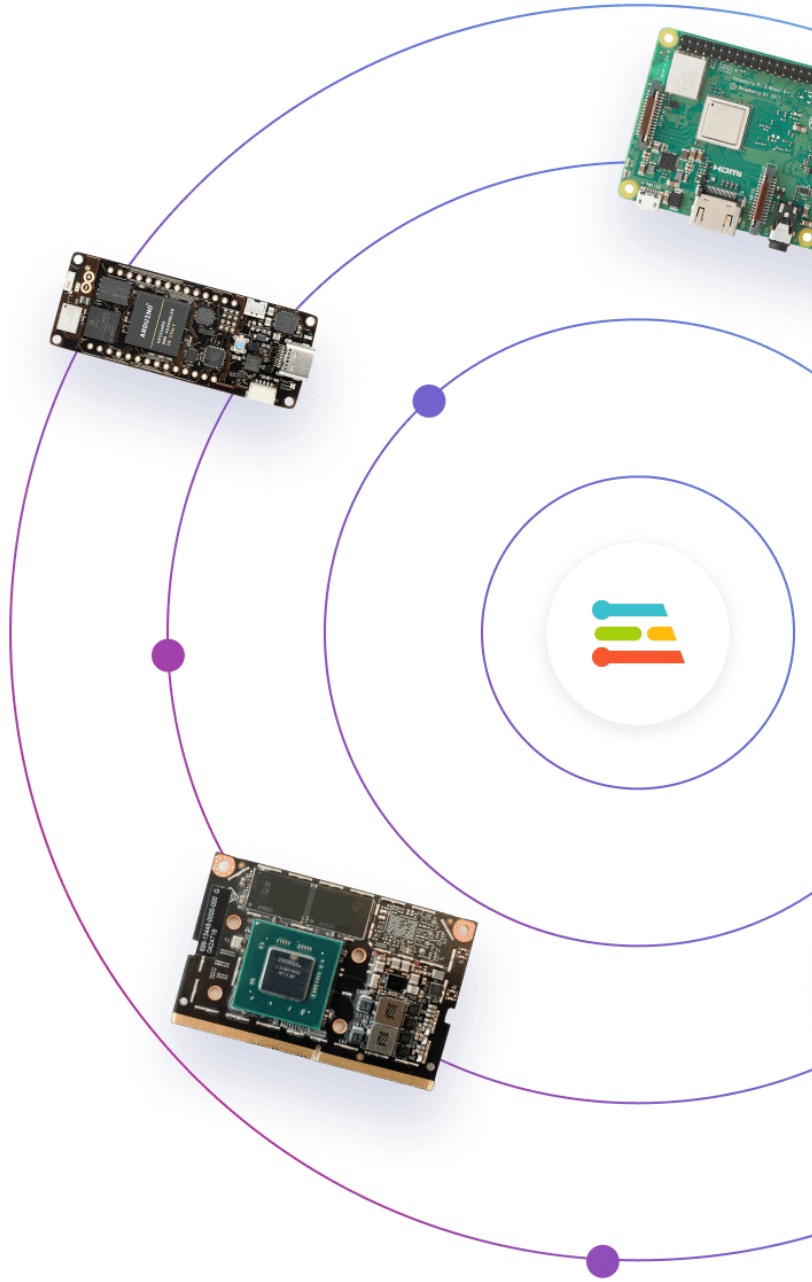
Shawn is a machine learning DevRel engineer, instructor, and university program manager at Edge Impulse. He creates compelling technical videos, courses, and blog posts around edge machine learning and embedded systems that inspire and teach engineers of all skill levels. Shawn is an advocate for enriching education through STEM and believes that the best marketing comes from teaching. He can be found giving talks, running workshops, and swing dancing in his free time.

EDGE IMPULSE

Constrained Object Detection on Microcontrollers with FOMO

Agenda

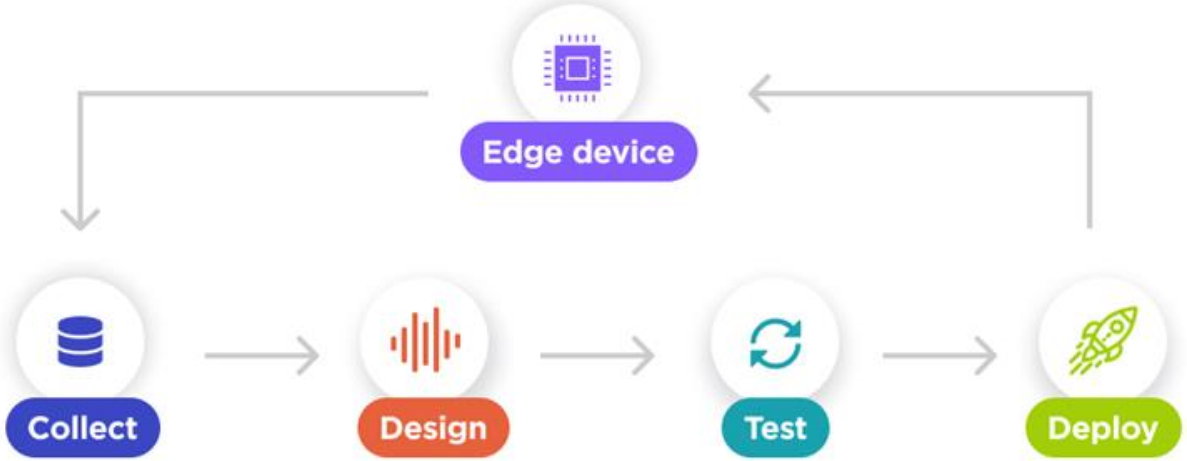
1. What is Edge Impulse?
2. Object Detection and Image Segmentation
3. Constrained Object Detection
4. Use Cases and Limitations
5. Live Demo

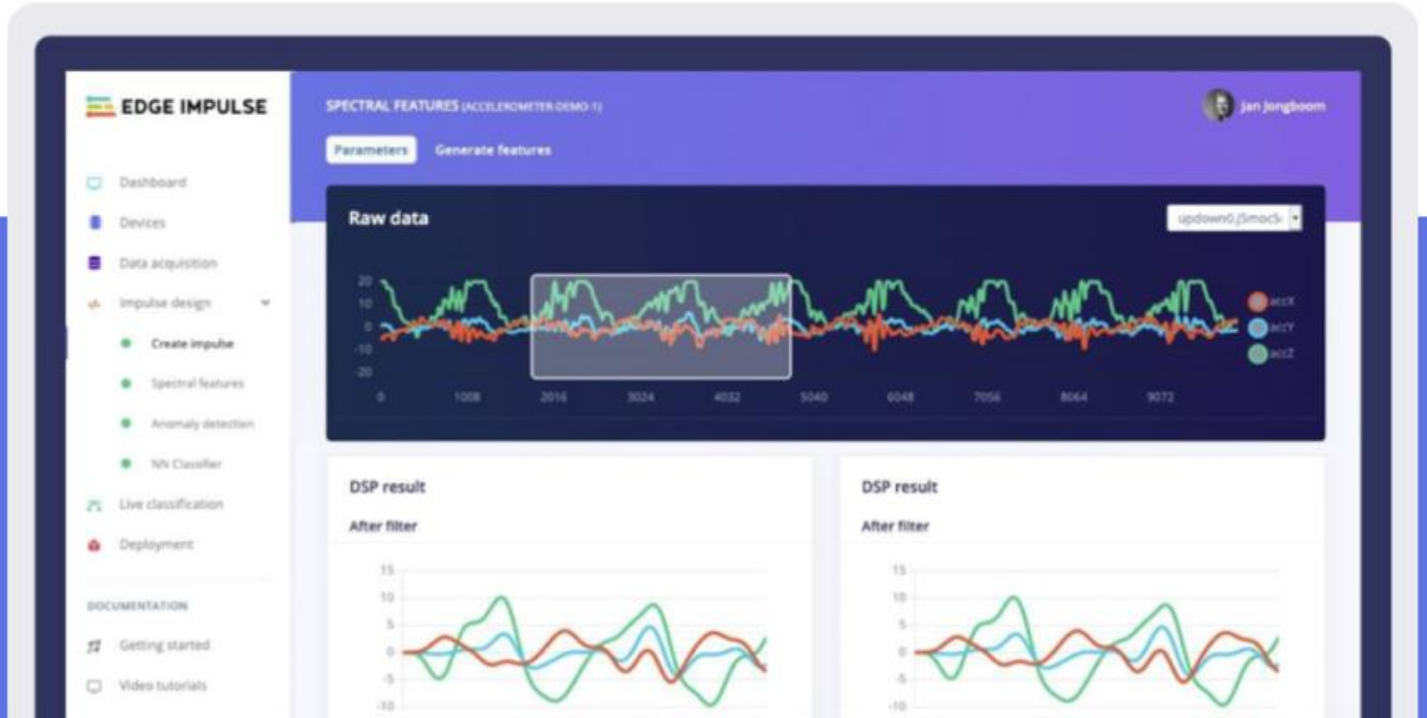


Edge Impulse

EDGE IMPULSE

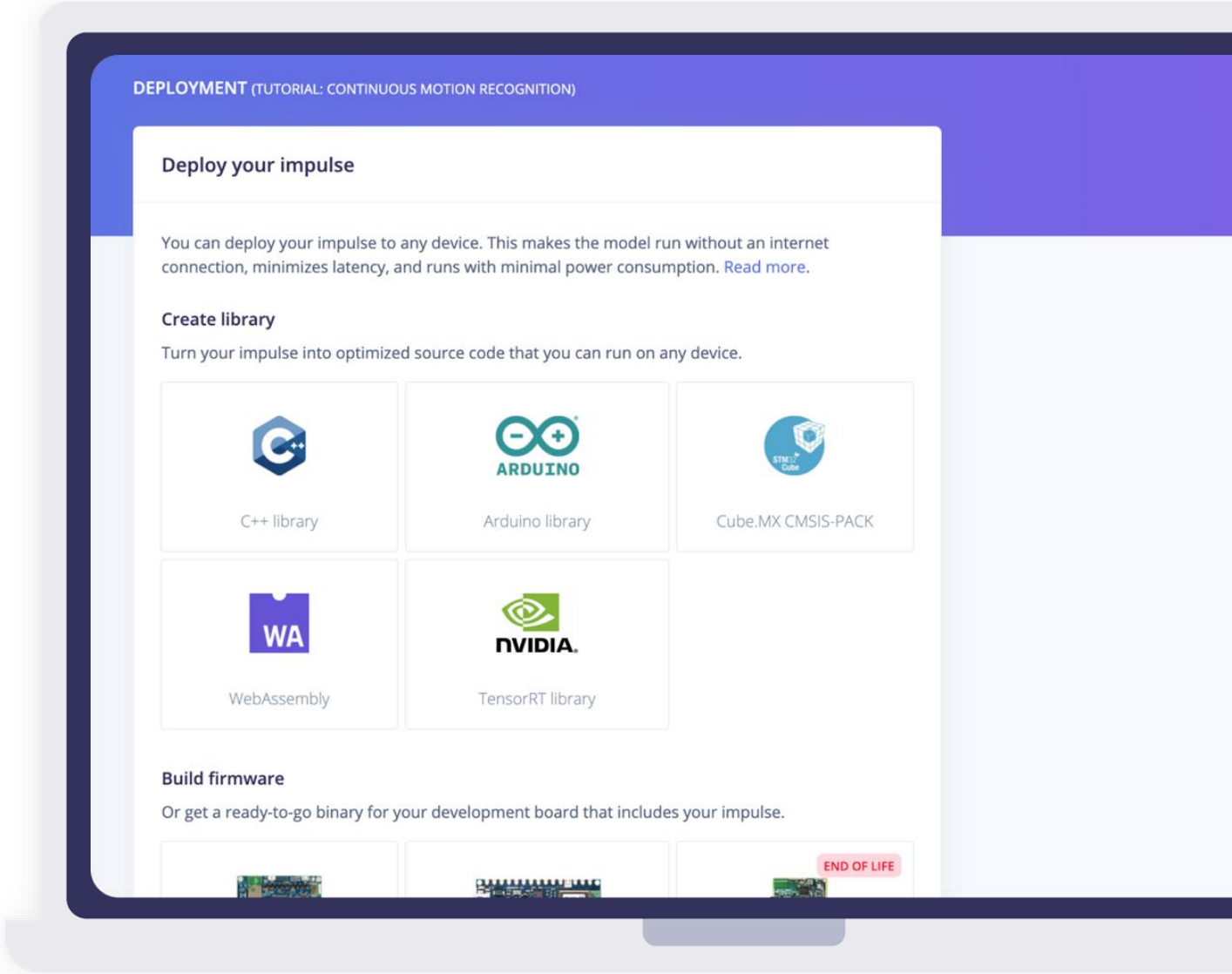
Go to market faster with confidence





Deploy to any edge device with ease

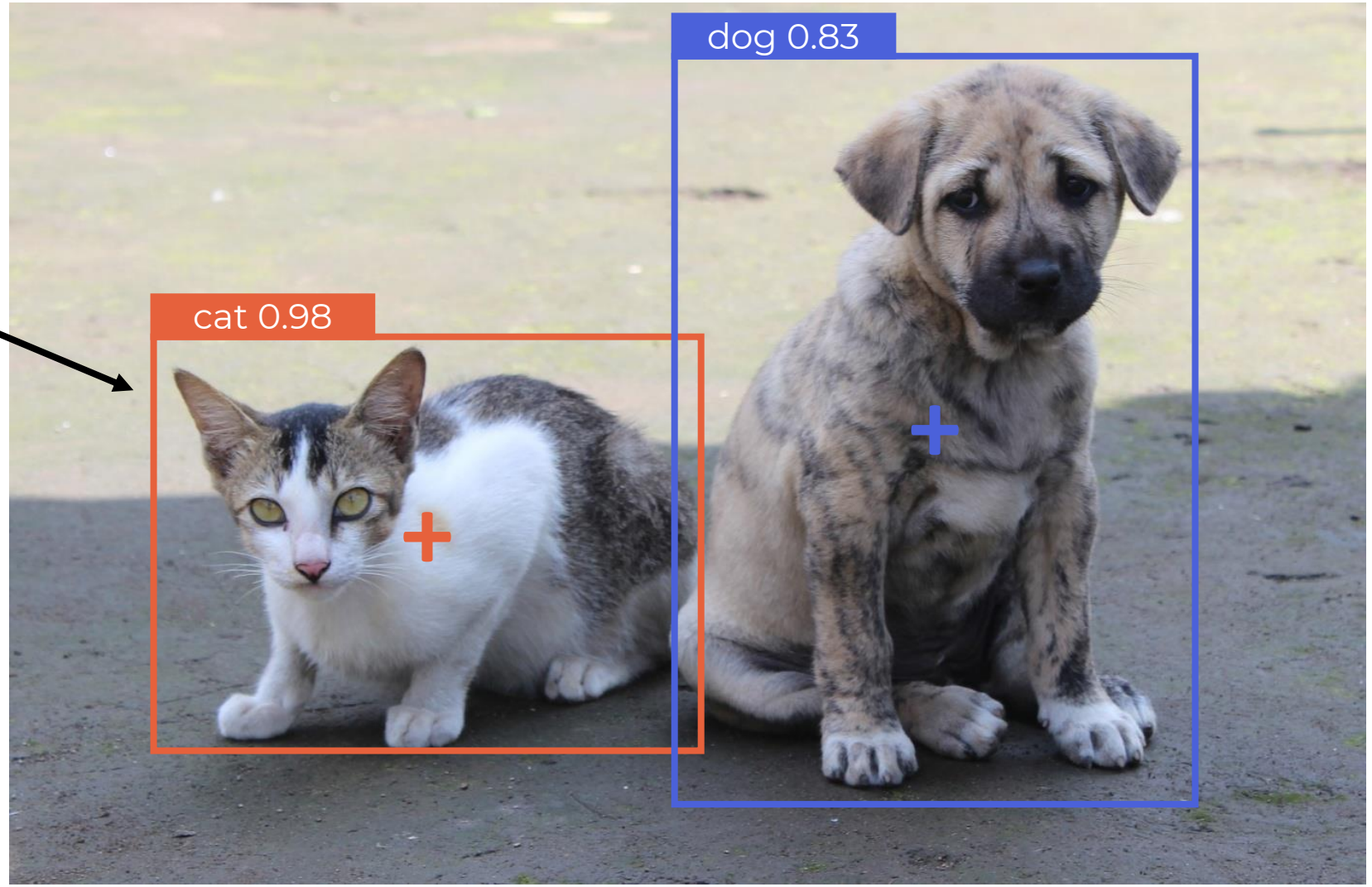
- The largest silicon ecosystem
- Award-winning compiler
- Access to device source code
- Full firmware integration for a number of devices



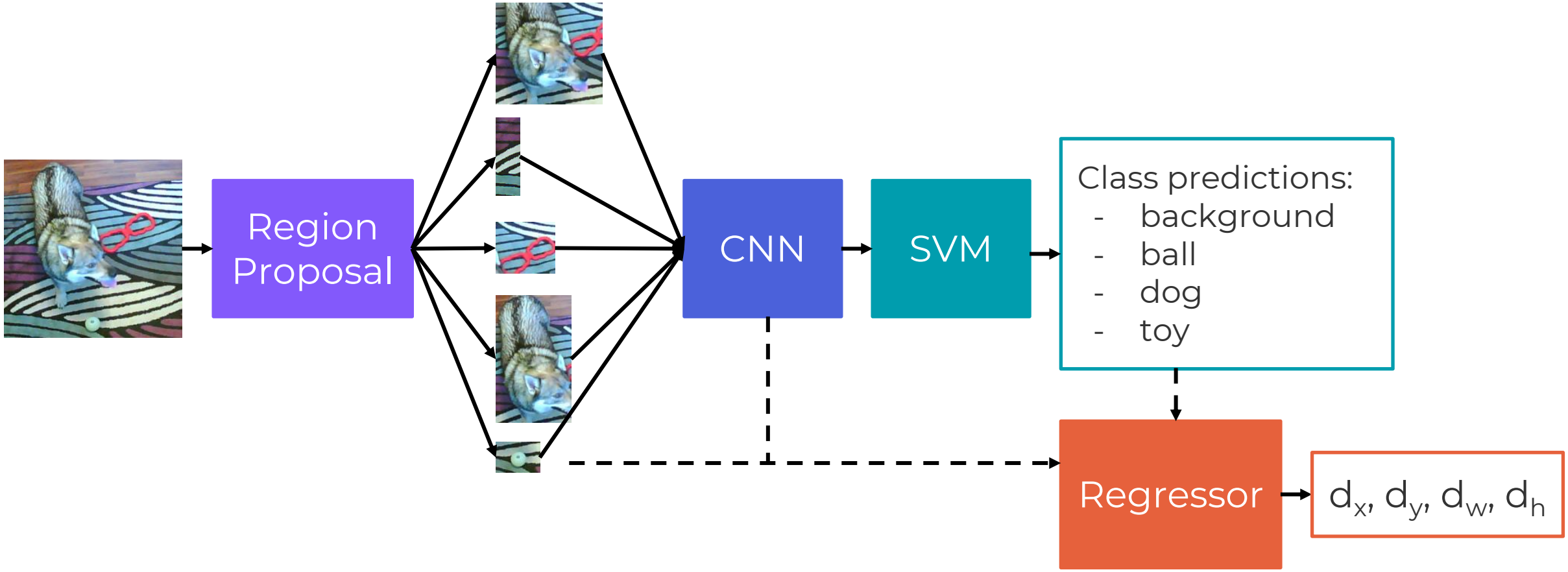
Object Detection + Image Segmentation

Object Detection

Bounding box



R-CNN



Single Shot MultiBox Detector (SSD)

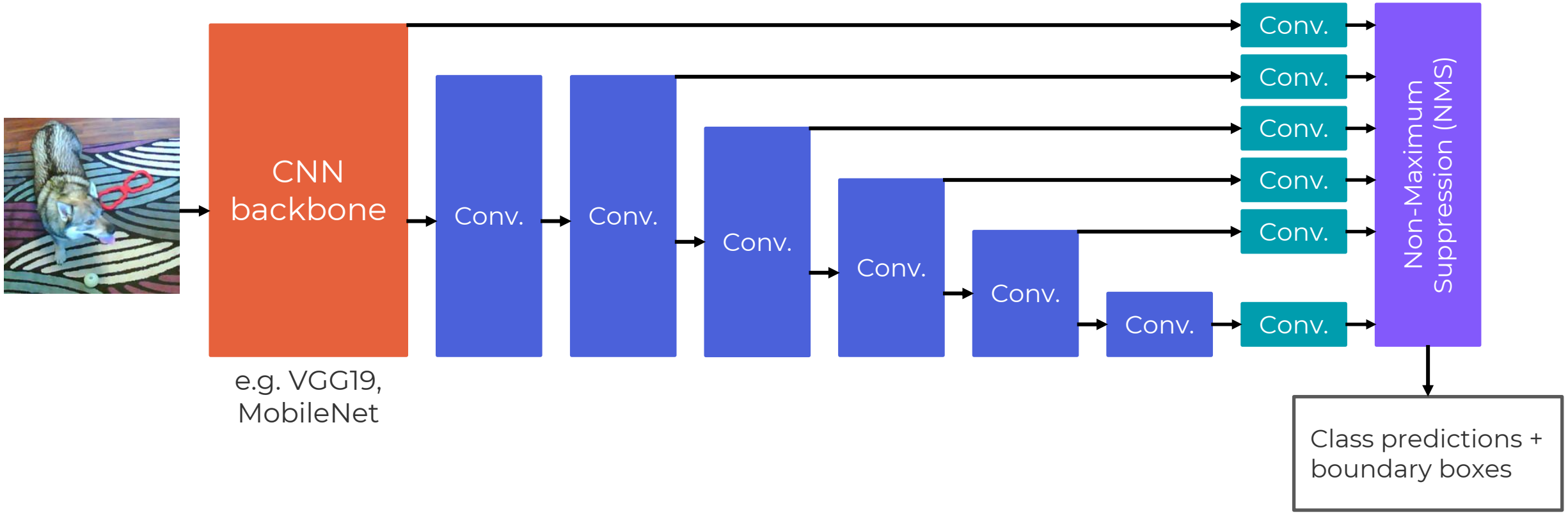
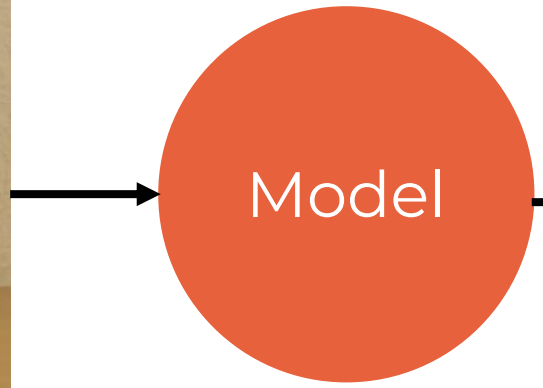
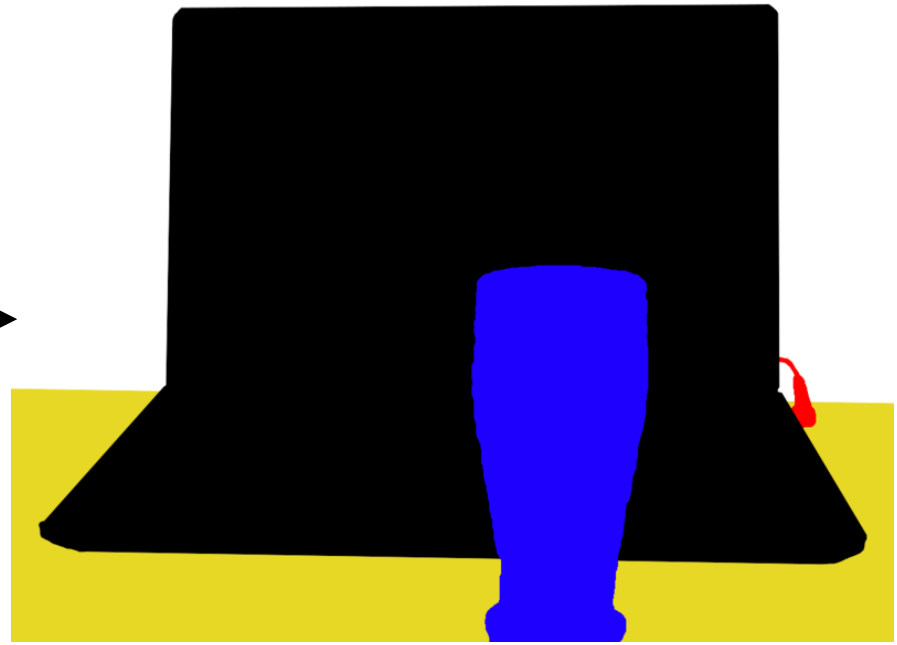


Image Segmentation





Constrained Object Detection

MobileNet V2

240x240

Convolutional layer(s)

Bottleneck residual block 1

Bottleneck residual block 2

Bottleneck residual block 3

Bottleneck residual block 4

...

Bottleneck residual block 17

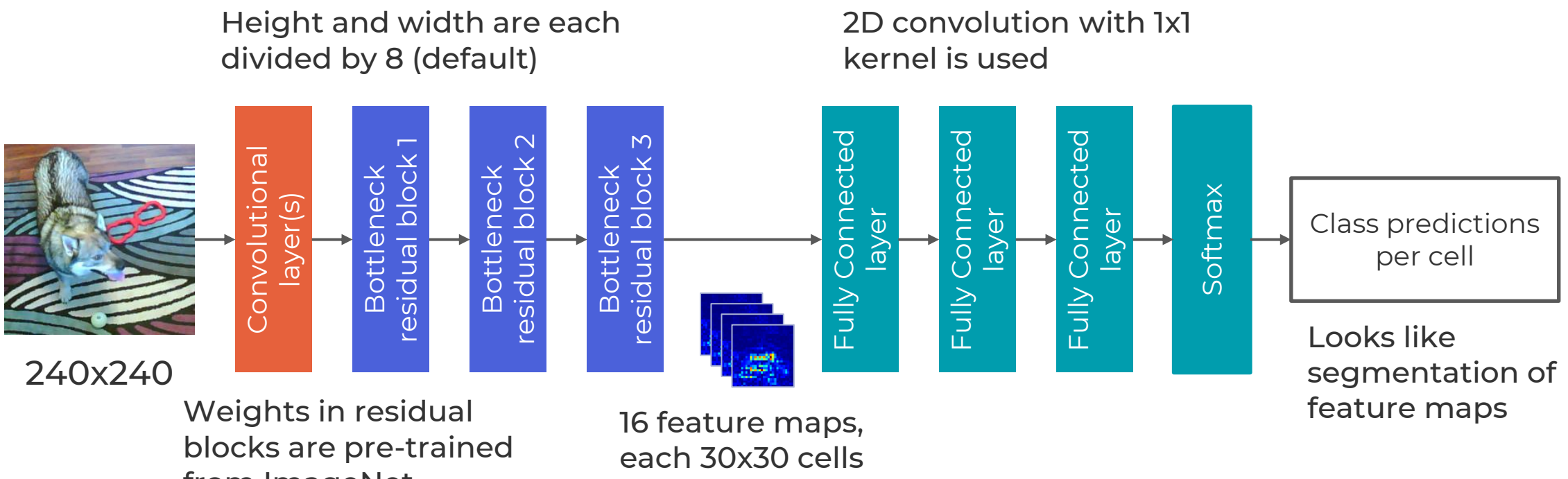
Convolutional layer(s)

Fully Connected layer

Softmax

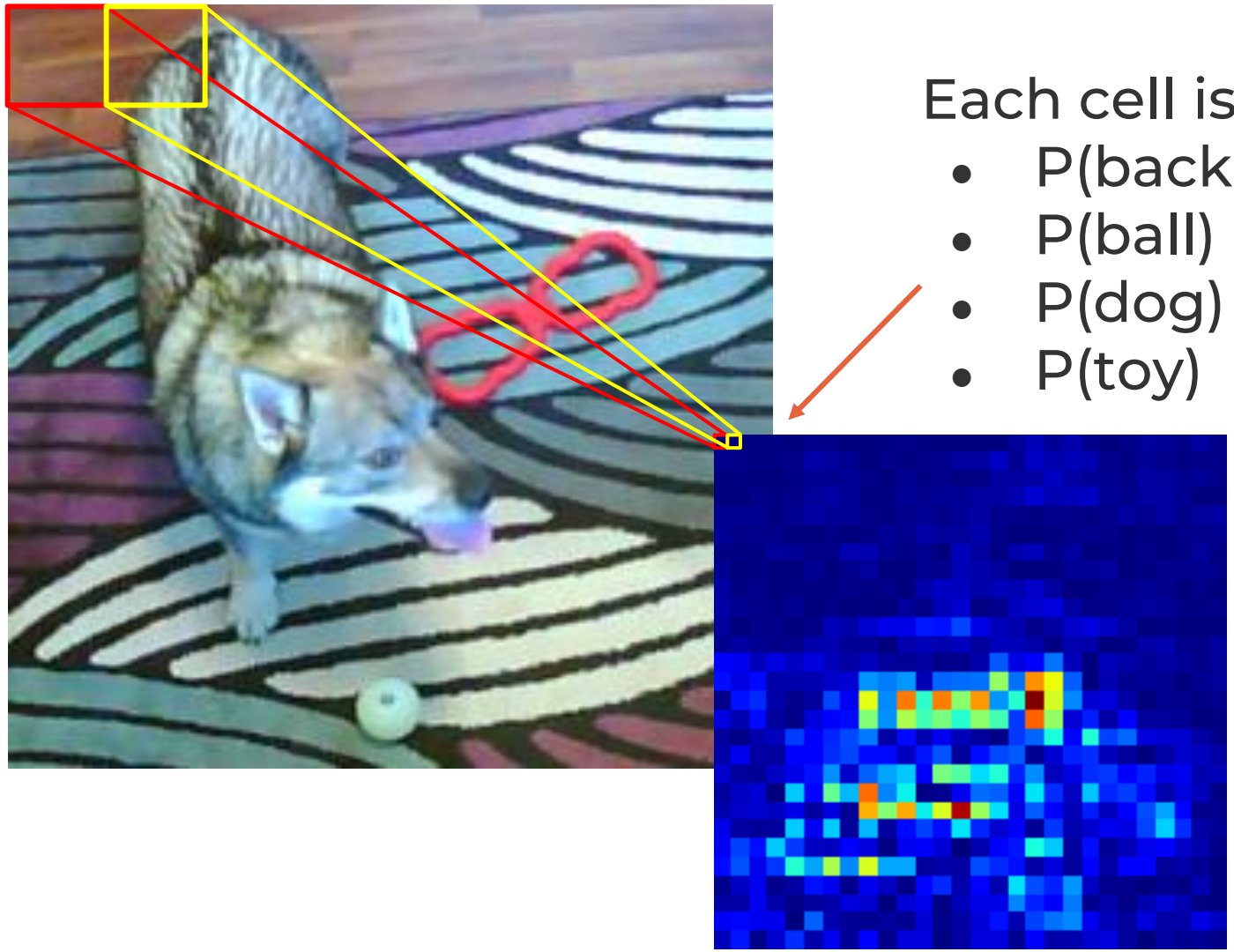
Class predictions for image

Faster Objects, More Objects (FOMO)



Faster Objects, More Objects (FOMO)

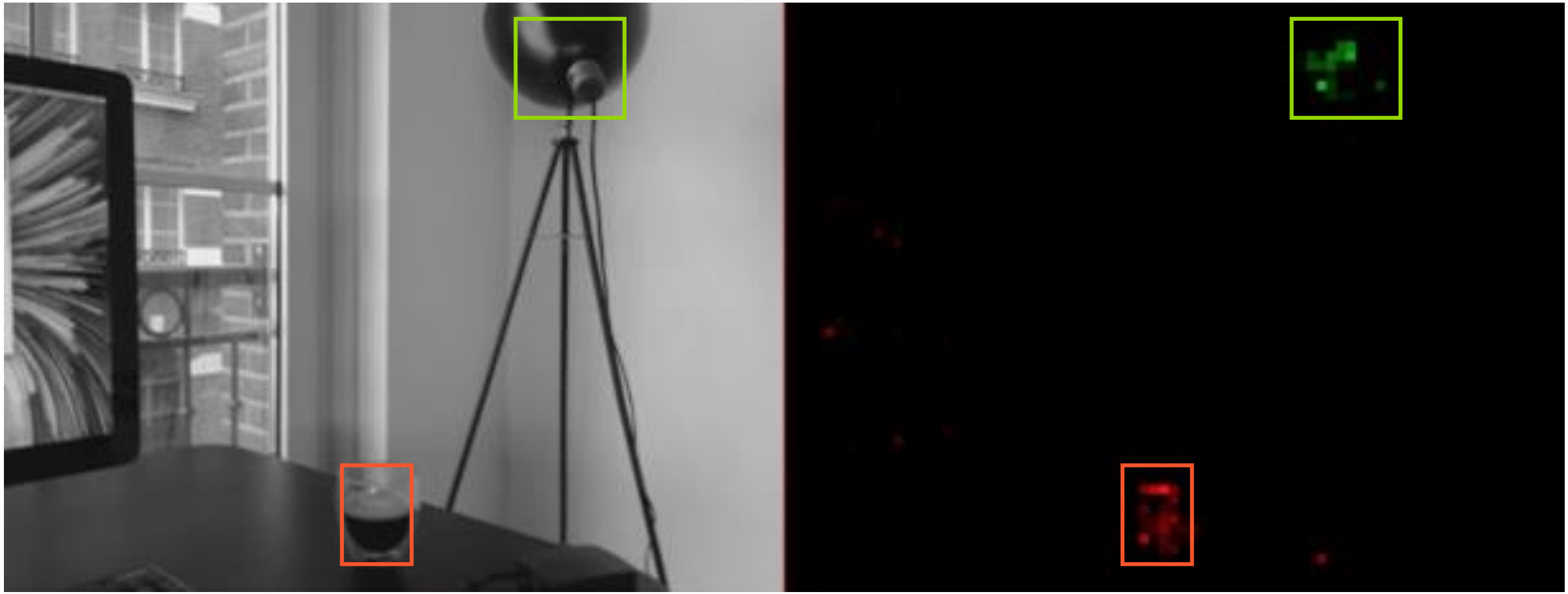
Receptive field



Each cell is given scores:

- P(background)
- P(ball)
- P(dog)
- P(toy)

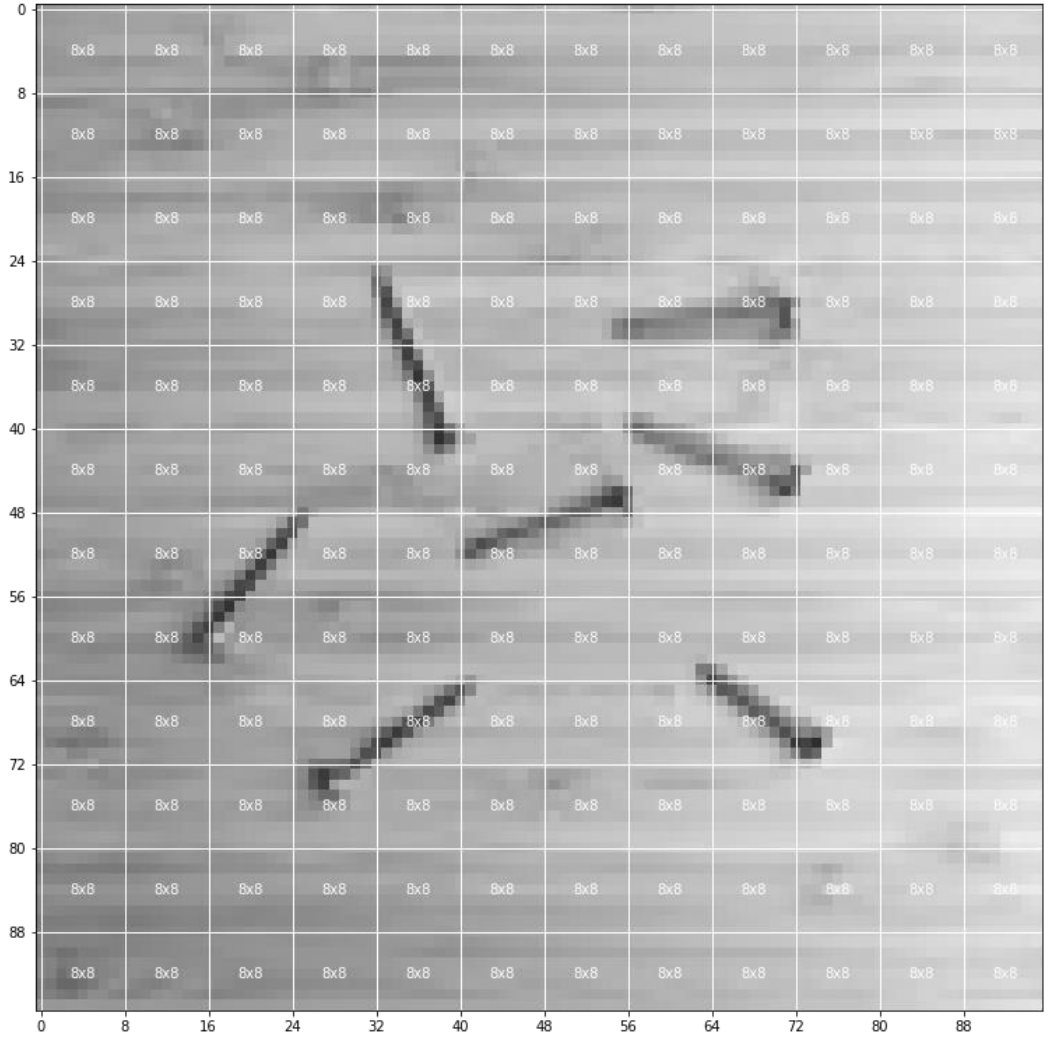
Faster Objects, More Objects (FOMO)



Faster Objects, More Objects (FOMO)

Example: screws

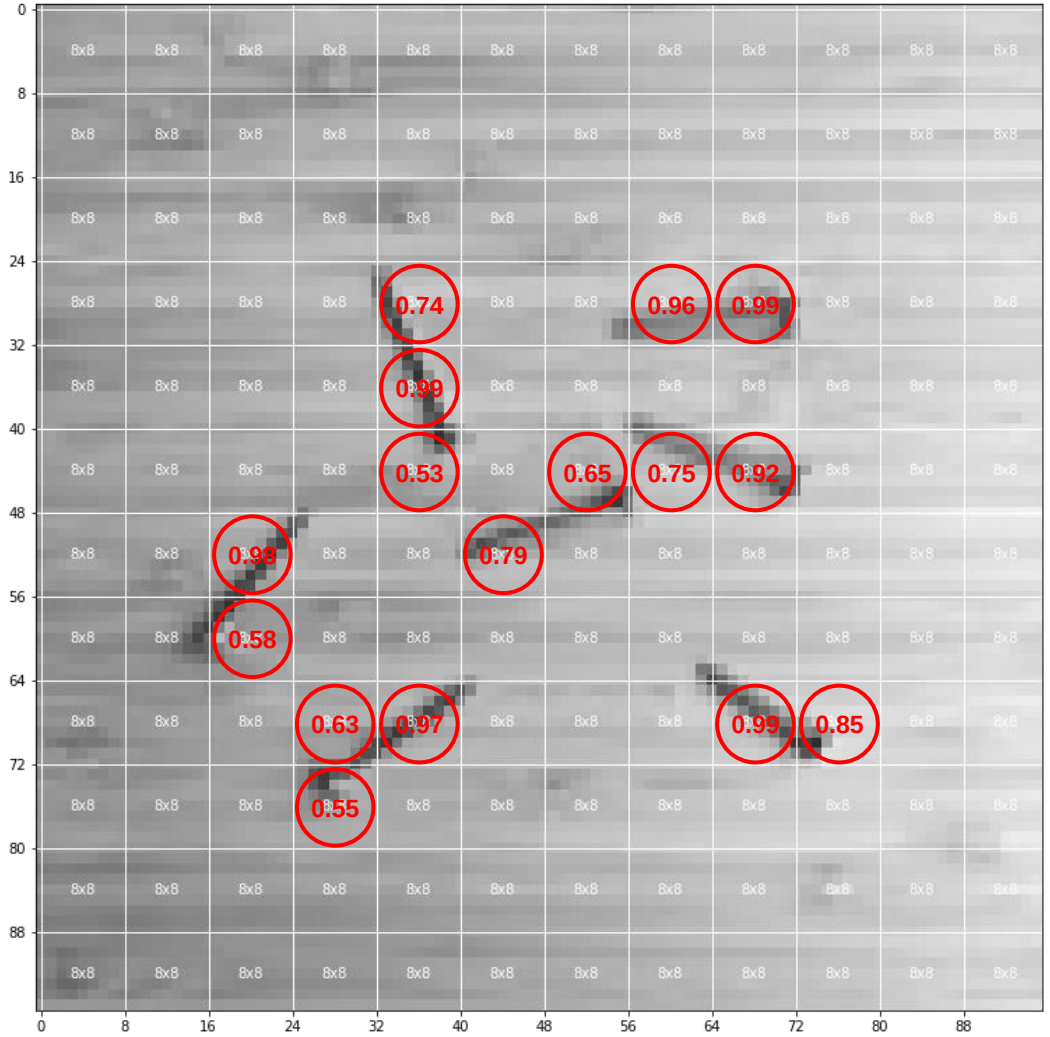
- Grayscale
- Image: 96x96
- Feature maps: 12x12



Faster Objects, More Objects (FOMO)

Example: screws

- Grayscale
- Image: 96x96
- Feature maps: 12x12

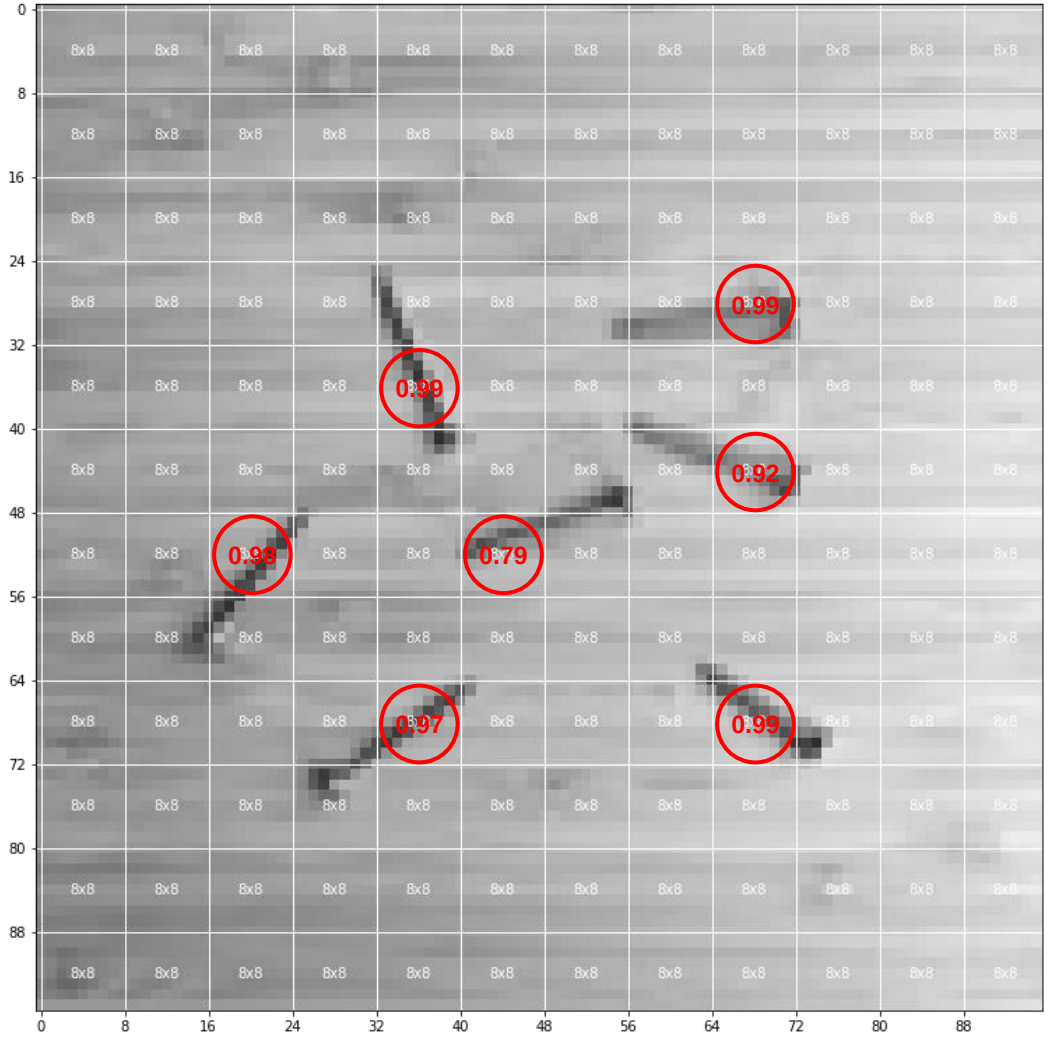


Faster Objects, More Objects (FOMO)

Example: screws

- Grayscale
- Image: 96x96
- Feature maps: 12x12

Neighboring cells with same class are removed (leaving highest scores)

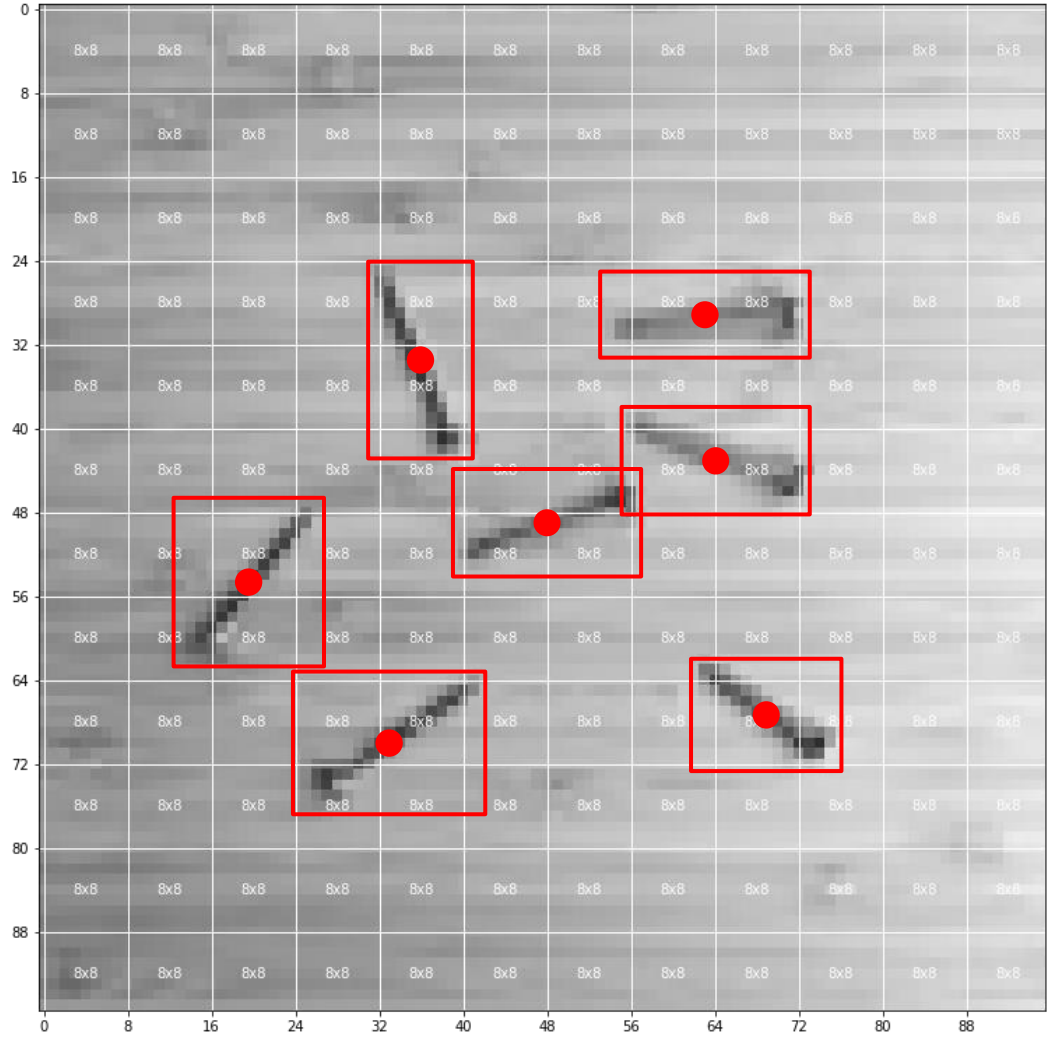


FOMO Ground Truth

Example: screws

- Grayscale
- Image: 96x96
- Feature maps: 12x12

User draws bounding boxes, tool picks cell with centroid of bounding box



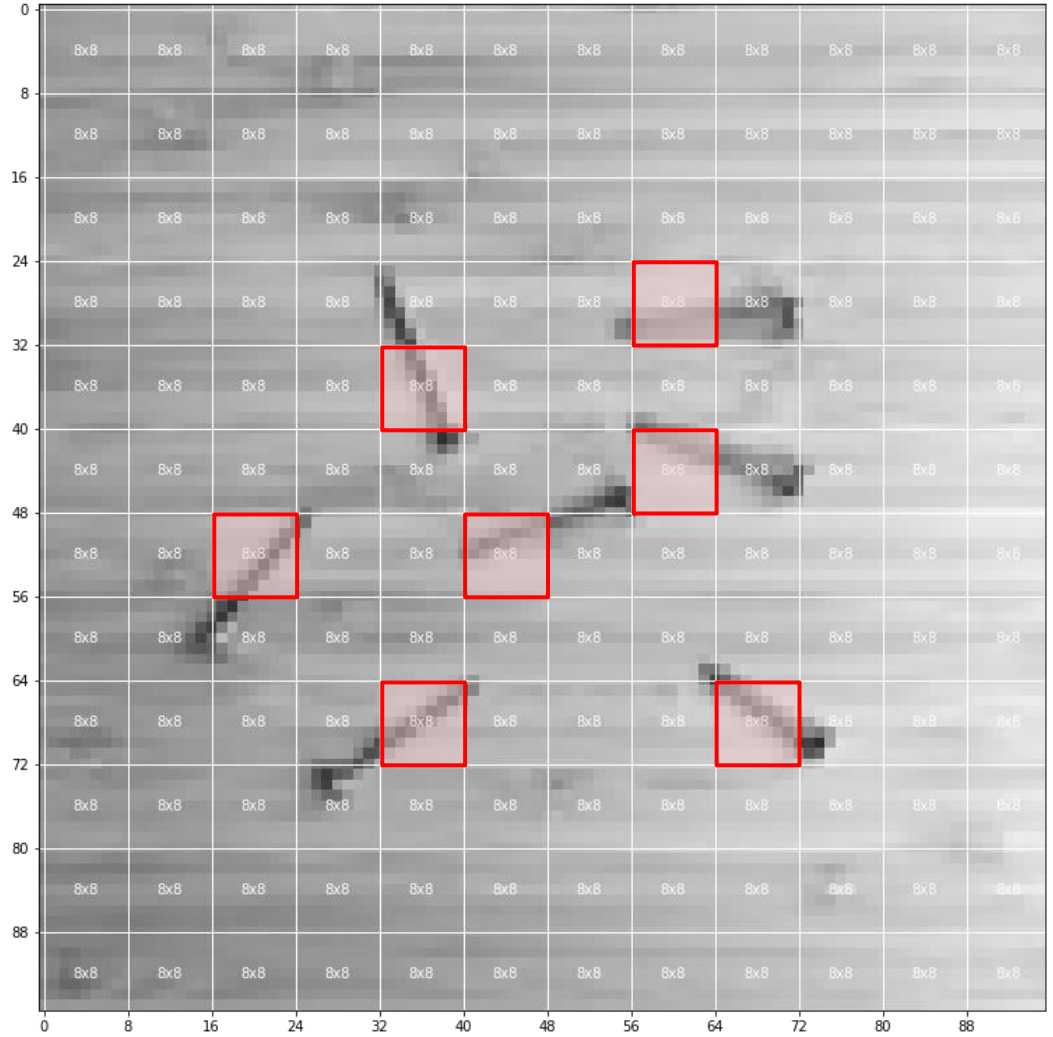
FOMO Ground Truth

Example: screws

- Grayscale
- Image: 96x96
- Feature maps: 12x12

User draws bounding boxes, tool picks cell with centroid of bounding box

Those cells are now representatives of that class



FOMO
Uses
+
Limitations

Use Cases

Want to know **where** and **how many** objects there are

Recommendations for success:

- Objects are same size
- Objects are square/round
- Objects take up 1 cell

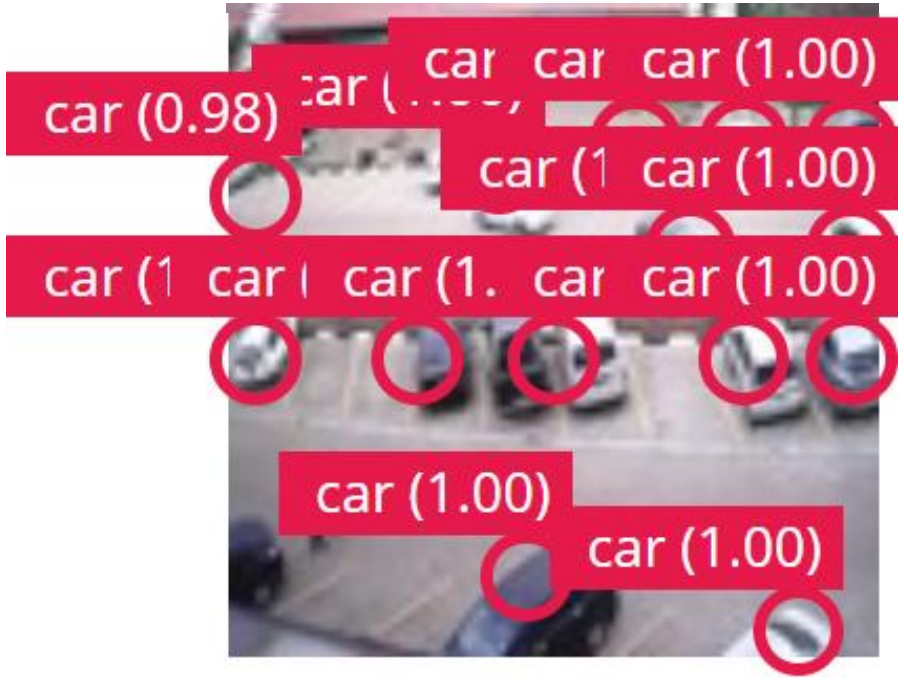
Very fast!

- Cortex-M7 at 480 MHz
- 240x240 image input
- 30 fps
- 245K RAM

https://matpalm.com/blog/counting_bees/

Limitations

- Each cell has its own classifier
- Small objects may be missed
- Neighboring objects may get lumped together
- Ends of oblong objects may be ignored
- Lots of objects/classes: use YOLOv5

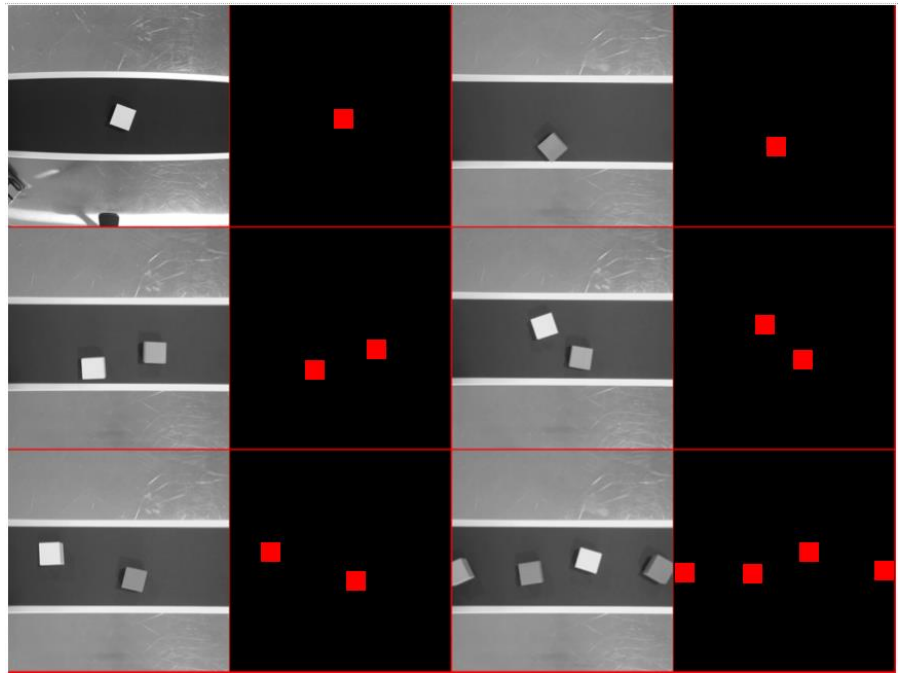


Demo

Getting Started

docs.edgeimpulse.com/docs/

- Tutorials > Counting objects using FOMO
- Various supported dev boards



EDGE IMPULSE

hello@edgeimpulse.com

3031 Tisch Way
110 Plaza West
San Jose, CA 95128
USA

TINY ML TALKS webcast

Please take the 5 question poll

and continue the conversation

@ tinyML.org/forums

tinyML Talks Strategic Partners

Additional Sponsorships available – contact Olga@tinyML.org for info

Next tinyML Talks

Date	Presenter	Topic / Title
Tuesday, April 5	Daniel Konegen, Embedded AI and data science engineer, Hahn-Schickard Marcus Rüb, Research assistant, Hahn-Schickard	AutoFlow - an open source Framework to automatically implement neural networks on embedded devices

Webcast start time is 9:30 am Pacific time

Please contact talks@tinyml.org if you are interested in presenting

Copyright Notice

www.tinyml.org

Copyright Notice

www.tinyml.org