tinyML® Summit

Miniature dreams can come true...

March 28-30, 2022 | San Francisco Bay Area

www.tinyML.org

EDGE IMPULSE

TINY ML

Making Things Smarter

EON Tuner: AutoML for constrained devices

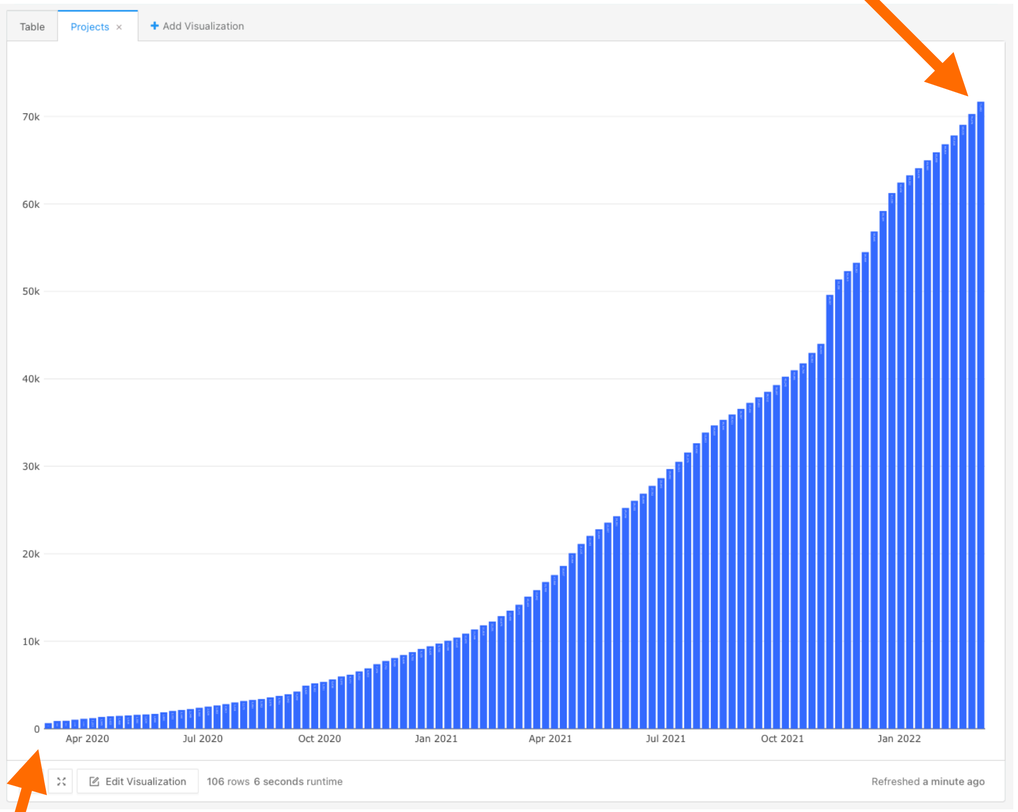
Hi, I'm Jan!

Co-founder and CTO at **Edge Impulse,** leading development platform for machine learning on edge devices.

Launched just before TinyML summit 2020…
Now 194 new projects daily (!)

TinyML Summit 2022

(71,682 projects)

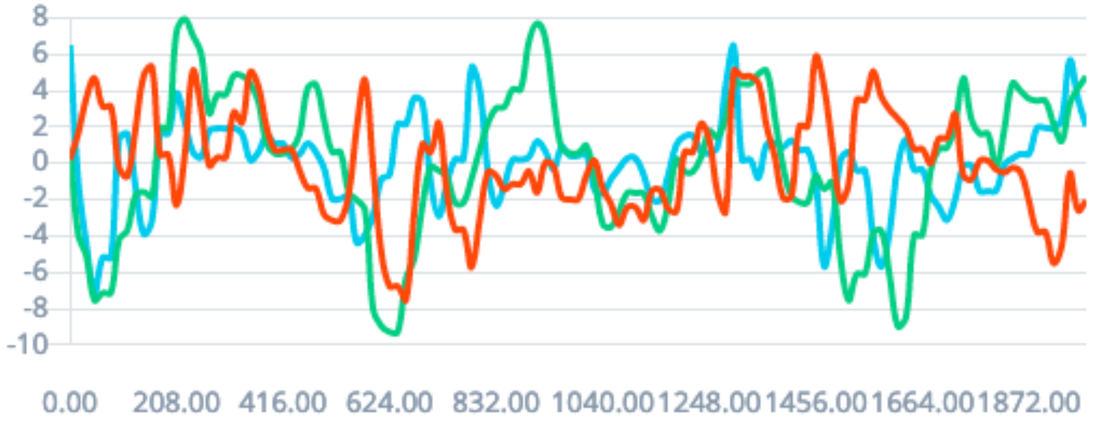


Edge Impulse project count

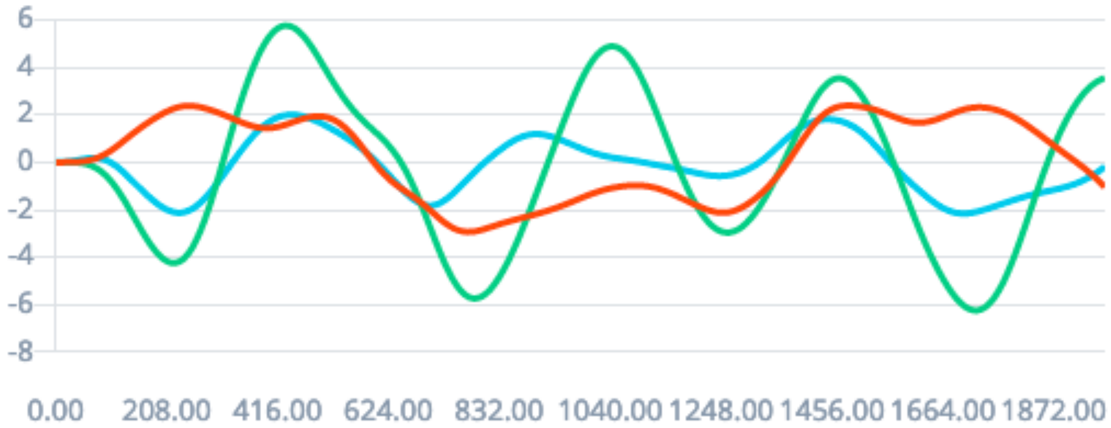
TinyML Summit 2020

(643 projects)

Signal processing + ML = 💚



Apply low-pass filter…



= much easier job
for the ML algo

Leveraging signal processing

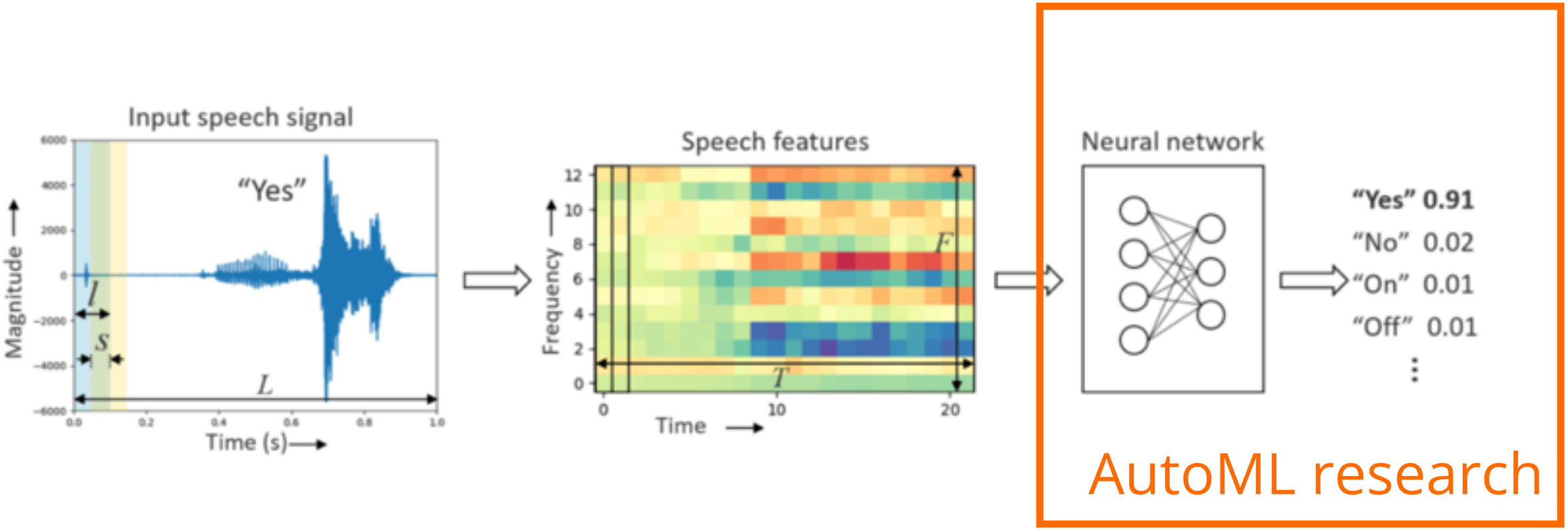
On-device intelligence is not new

Neural networks are inefficient, if you can preprocess? Do so!

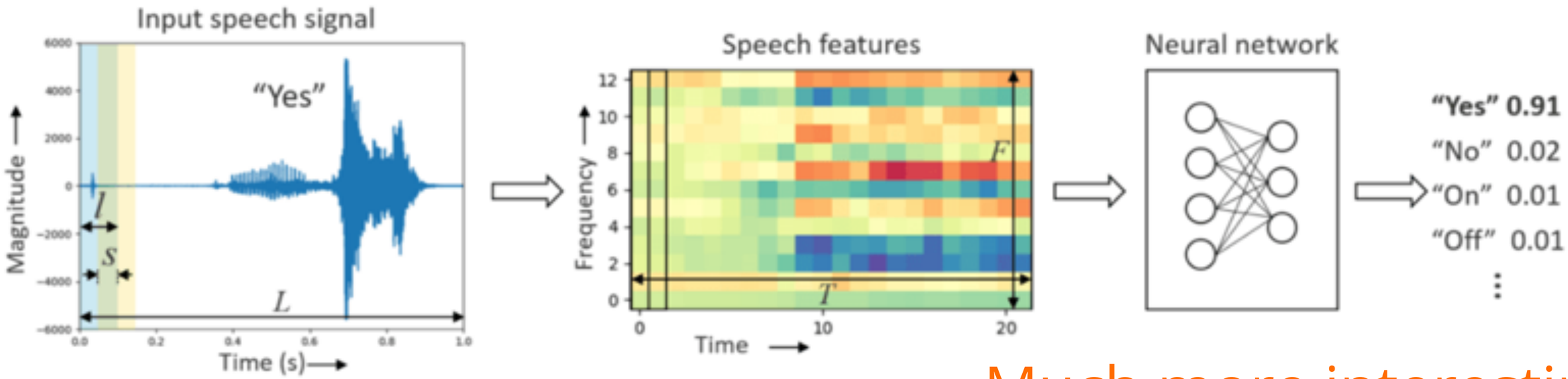
Significantly reduce input features, leading to smaller networks.

Cleans up input

ML Sensor pipeline is often combination



ML Sensor pipeline is often combination

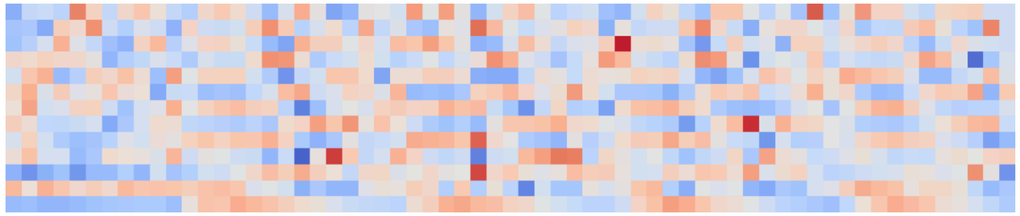


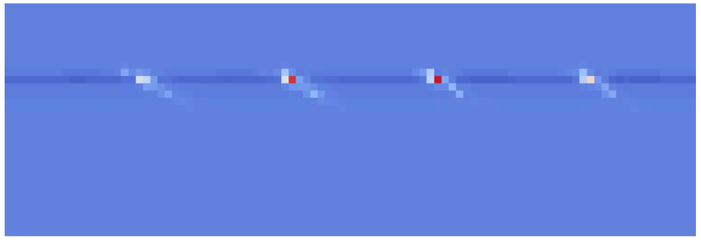
Much more interesting

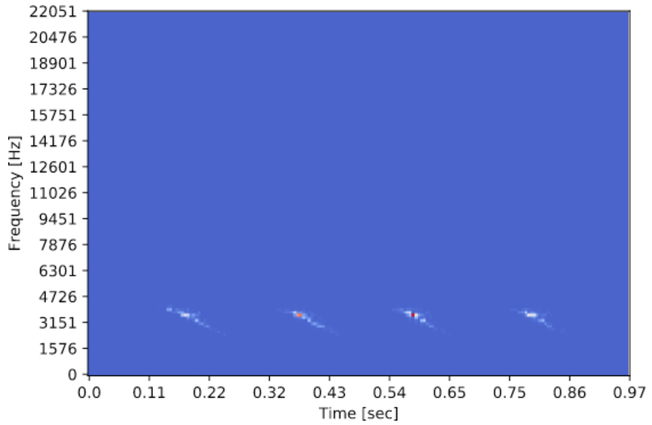
Wide range of parameters

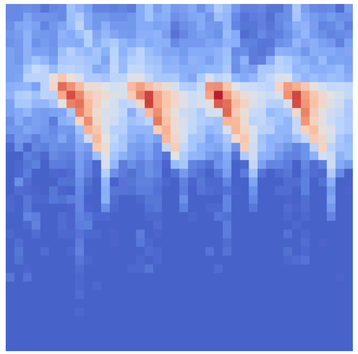
Window length, window step, downsample?

Wide range of parameters



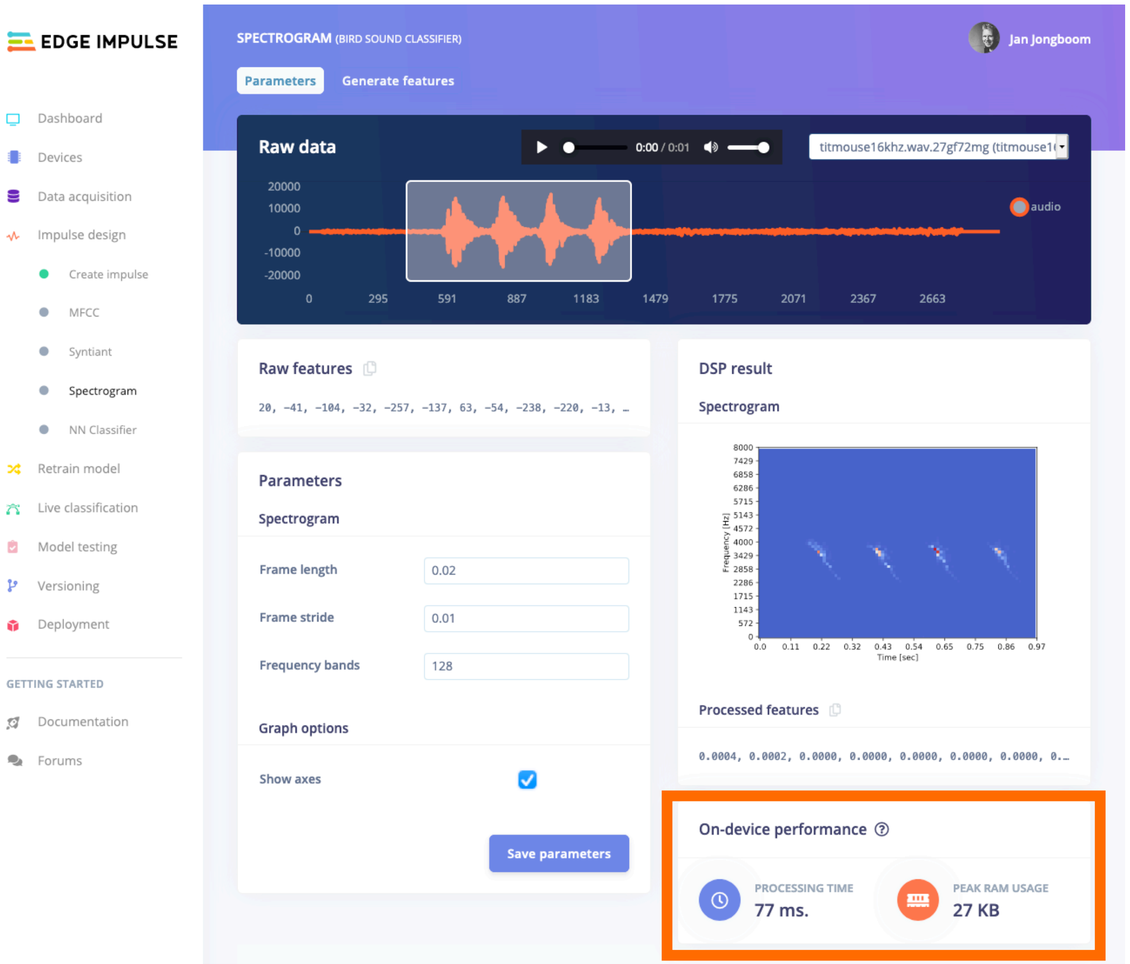


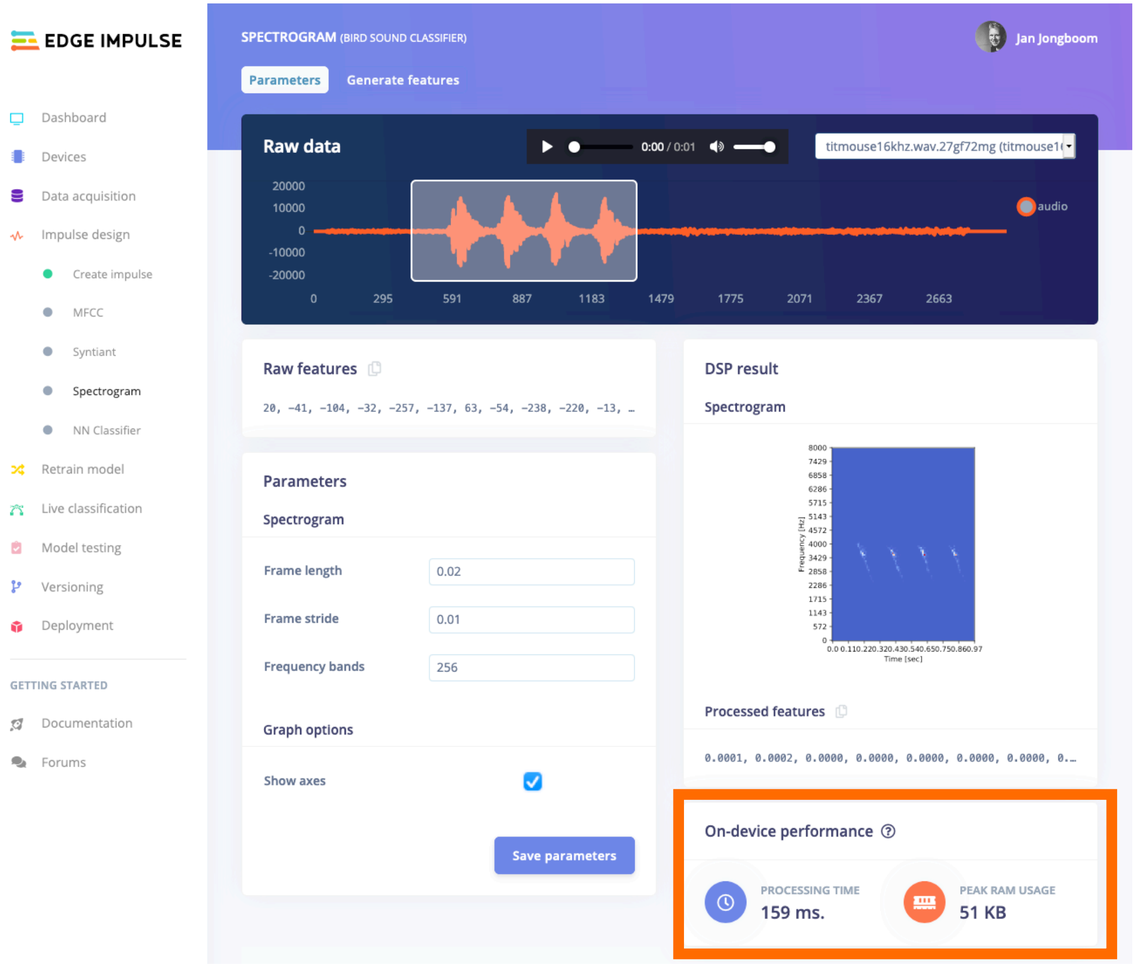




+ endless configuration options

Constrained targets - what's worth it?





Introducing the EON Tuner!

EDGE IMPULSE

- Dashboard
- Devices
- Data acquisition
- Impulse design
  - Create impulse
  - MFCC
  - NN Classifier
- Retrain model
- Live classification
- Model testing
- Versioning
- Deployment
- EON Tuner

EON TUNER (DCASE2018)

mathijs

EON Tuner

Config | Logs

Finding the most optimal architecture for your model (**17** model variants evaluated / **50** variants total)

cortex-m7-216mhz **1000 ms** RAM: **128kB** • ROM: **1024kB**

97% 10x10 | SYNTIA | Clone | a65

PERFORMANCE

Latency RAM ROM

ACCURACY

	coo	dis	eat	soc	vac	
coo	97	3	0	0	0	0
dis	2	94	4	0	0	0
eat	0	0	100	0	0	0
soc	4	0	2	93	0	0
vac	0	0	0	0	100	0

97% 10x10 | SYNTIA | Clone | 133

PERFORMANCE

Latency RAM ROM

ACCURACY

	coo	dis	eat	soc	vac	
coo	98	2	0	0	0	0
dis	4	88	8	0	0	0
eat	0	0	100	0	0	0
soc	2	2	0	96	0	0
vac	0	0	0	0	100	0

96% 10x10 | SYNTIA | Clone | 2f1

PERFORMANCE

Latency RAM ROM

ACCURACY

	coo	dis	eat	soc	vac	
coo	95	5	0	0	0	0
dis	2	92	4	2	0	0
eat	0	3	98	0	0	0
soc	2	0	2	96	0	0
vac	0	0	0	0	100	0

Filters

Status

- [x] Pending
- [x] Running
- [x] Completed
- [x] Failed

View

Data set

- ( ) Validation
- ( ) Train

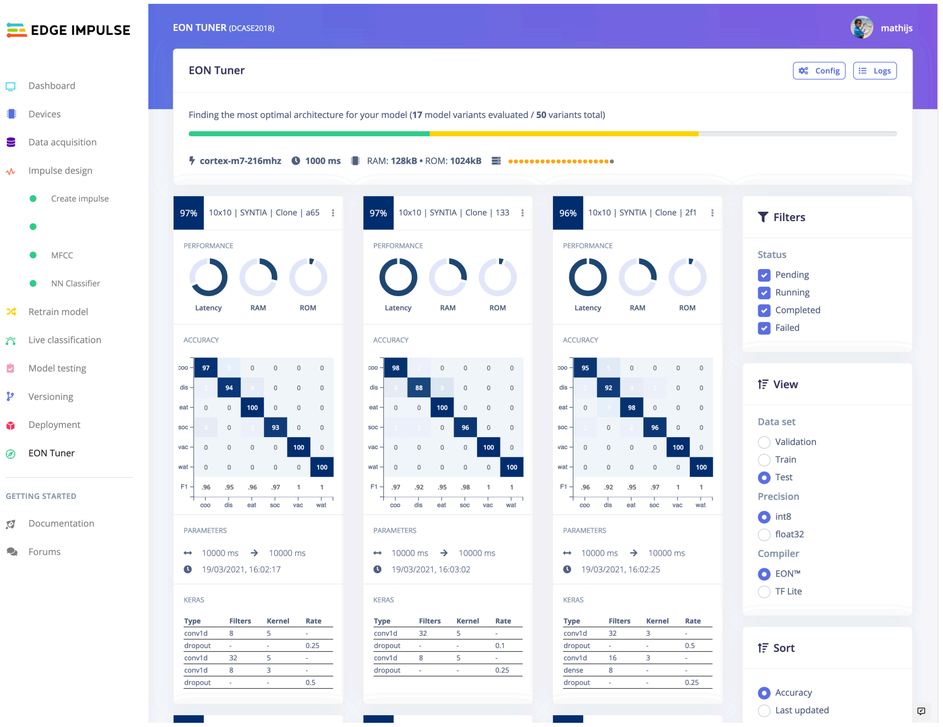
EON Tuner

Find best model for sensor data over mix of input blocks, DSP blocks and ML blocks

Specify device constraints

Extensible with your own DSP and ML blocks

Biggest win: "we found DSP configuration that works so well, we no longer need ML"



EDGE IMPULSE

That sounds too easy...

1. Understand the problem

Understand the business constraints
(response time = sampling time + inference)

See interesting frequencies?

Need custom DSP code?

Classification, regression or anomaly detection?

Still an engineering tool!

2. Dataset quality

3. Custom search space

Tuner comes with basic scenarios (keyword spotting, continuous audio, image classification, motion, etc.)

Customize search space (parameters, which preprocessing blocks, ML configuration)

Tied to performance APIs, to drop parts of the space that don't fit constraints

```
dspBlock: {
     type: 'spectrogram',
     noise_floor: [ -62, -52, -42 ],
     fft_length: [ 64, 128, 256 ]
 }
```

EDGE IMPULSE

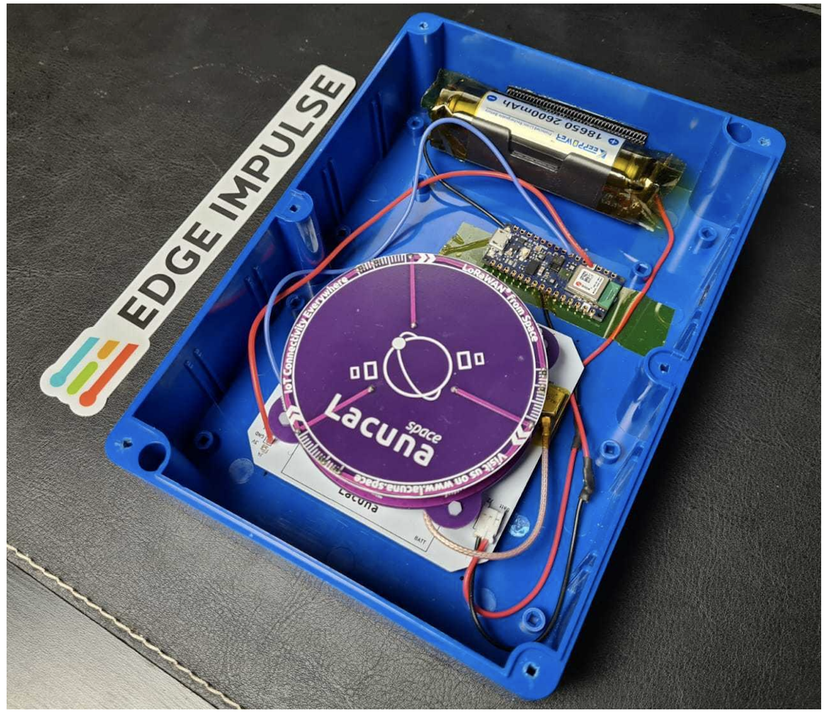
Examples

Bird sound classifier

Built with Lacuna Space, counts birds locally and relays results back over satellite

Model built by Dan (hi Dan!) from our team:
82.62% test set accuracy

EON Tuner found one w/ 92.66% test set accuracy (or even 95% if forego 1 sec. latency requirement)

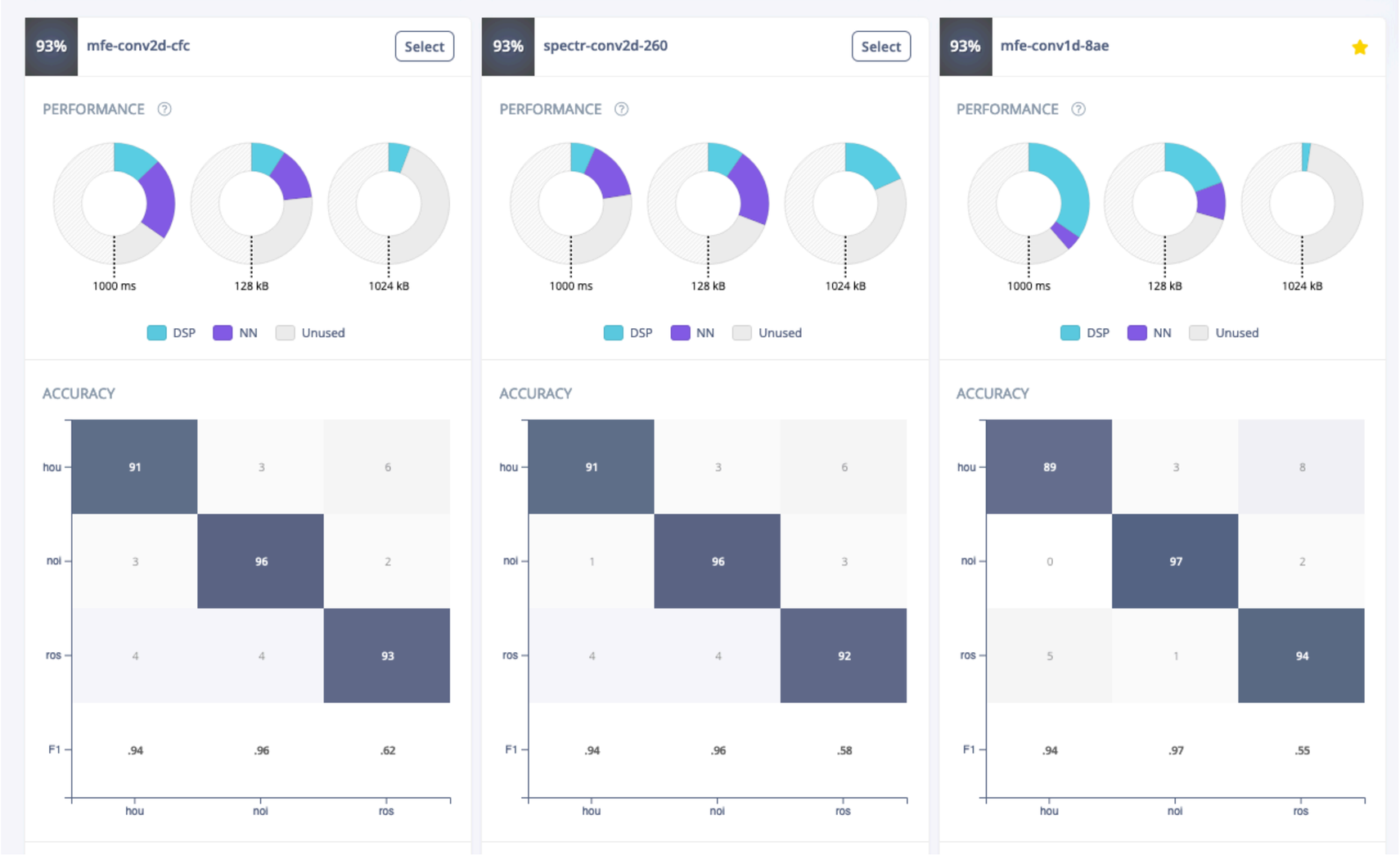


https://www.edgeimpulse.com/blog/introducing-the-eon-tuner-edge-impulses-new-automl-tool-for-embedded-machine-learning

Best models...


95% mfe-conv2d-37d Select
PERFORMANCE
1000 ms
128 kB
1024 kB
DSP NN Unused
ACCURACY
hou 93 3 5
noi 1 97 2
ros 1 1 97
F1 .95 .97 .68
hou noi ros
INPUT
4000 ms | 1000 ms
MFE
0.05 | 0.025 | 40
KERAS
Type Filters Kernel Rate
conv2d 8 3 -
conv2d 16 3 -
conv2d 32 3 -
dropout - - 0.5
dense 64 - -
dropout - - 0.5
27/07/2021, 08:33:33
93% mfe-conv2d-67d Select
PERFORMANCE
1000 ms
128 kB
1024 kB
DSP NN Unused
ACCURACY
hou 91 4 5
noi 3 96 1
ros 2 4 95
F1 .94 .96 .66
hou noi ros
INPUT
4000 ms | 4000 ms
MFE
0.05 | 0.05 | 40
KERAS
Type Filters Kernel Rate
conv2d 8 3 -
conv2d 16 3 -
conv2d 32 3 -
dropout - - 0.25
27/07/2021, 08:09:28
93% spectr-conv2d-528 Select
PERFORMANCE
1000 ms
128 kB
1024 kB
DSP NN Unused
ACCURACY
hou 91 2 7
noi 2 96 3
ros 2 2 95
F1 .94 .97 .56
hou noi ros
INPUT
4000 ms | 2000 ms
SPECTROGRAM
0.075 | 0.075 | -52
KERAS
Type Filters Kernel Rate
conv2d 8 3 -
conv2d 16 3 -
conv2d 32 3 -
dropout - - 0.25
dense 64 - -
dropout - - 0.25
27/07/2021, 08:10:41

Lots of DSP? Lots of NN?



EDGE IMPULSE

Getting started

Getting started

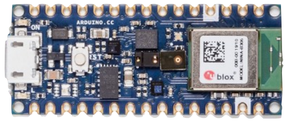
https://docs.edgeimpulse.com/docs

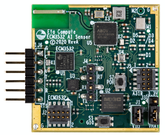
Very wide range of dev boards, from Cortex-M0+ to Jetson Nano

Deploy to any device that has a C++ compiler

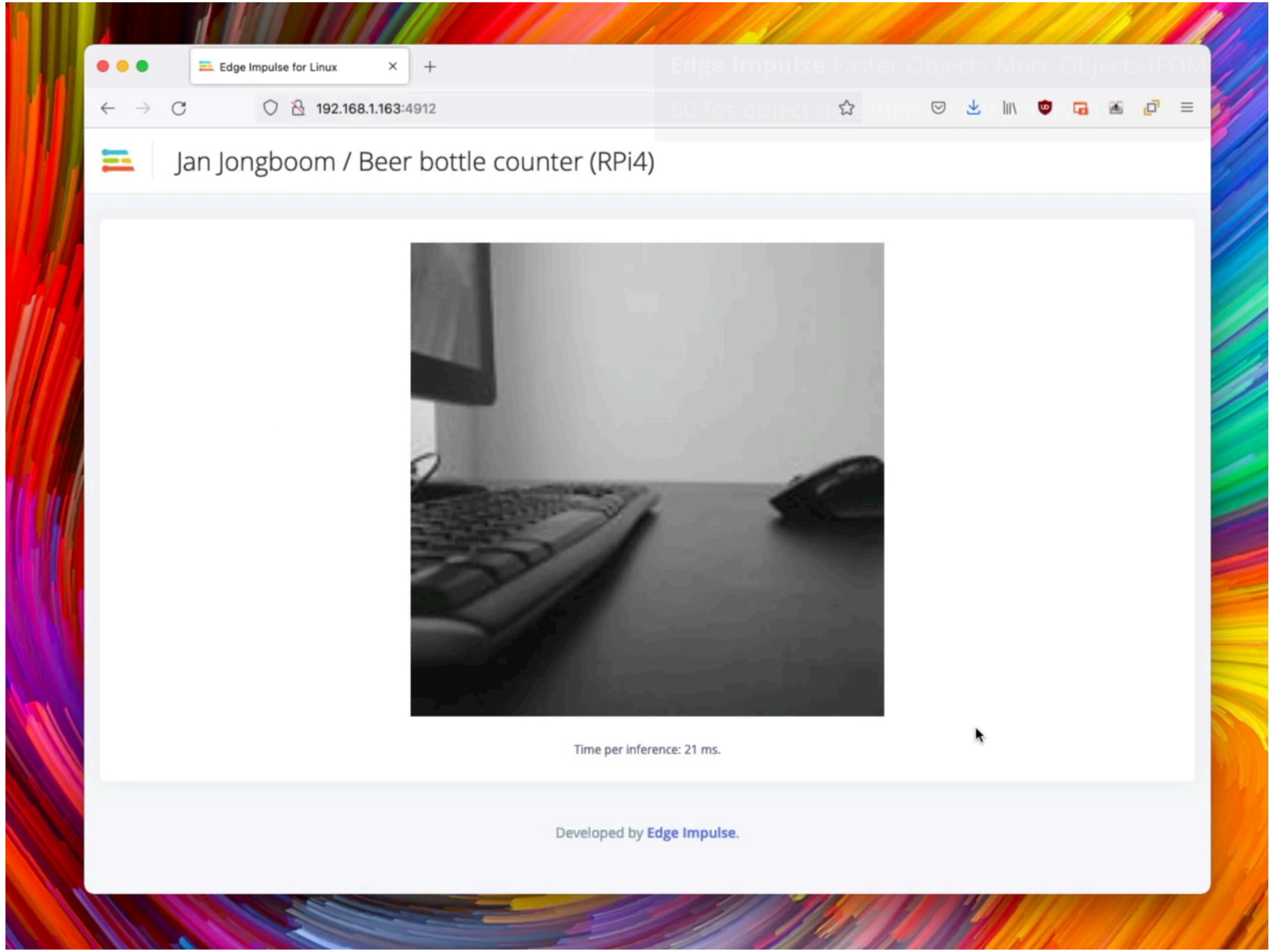
Or use your phone!

Be one of today's 194 new projects!





FOMO - Object detection for MCUs



Recap

We've come so far!

Domain knowledge is key.

It's the AutoML track, but it's still engineering.

edgeimpulse.com

Questions?

Full docs:

https://docs.edgeimpulse.com

Tuner intro:

https://docs.edgeimpulse.com/docs/eon-tuner

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

tinyML Summit 2022 Sponsors

Copyright Notice