

tinyML[®] Talks

Enabling Ultra-low Power Machine Learning at the Edge

“Tiny machine learning with Arduino”

Massimo Banzi - Arduino

UK Area Group – April 20, 2021



www.tinyML.org



tinyML Talks Sponsors

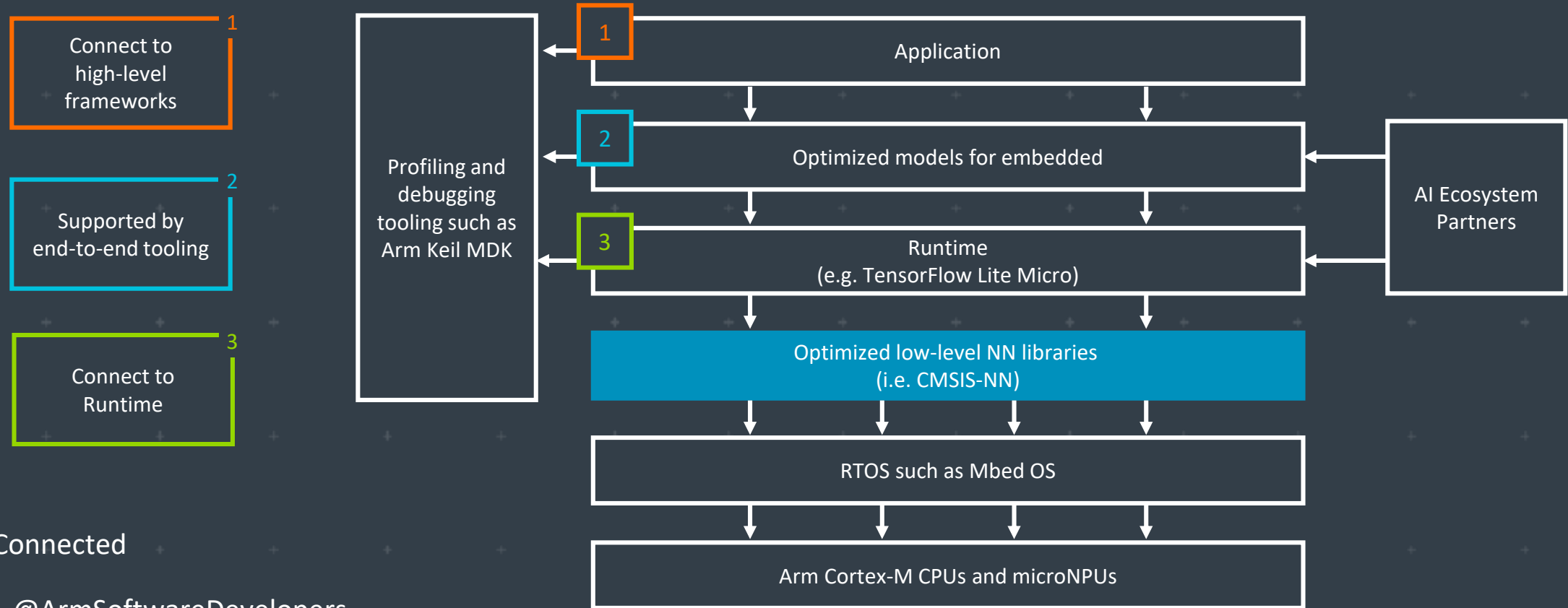


tinyML Strategic Partner



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

Arm: The Software and Hardware Foundation for tinyML



Stay Connected

 @ArmSoftwareDevelopers

 @ArmSoftwareDev

Resources: developer.arm.com/solutions/machine-learning-on-arm



WE USE AI TO MAKE OTHER AI FASTER, SMALLER AND MORE POWER EFFICIENT



Automatically compress SOTA models like MobileNet to <200KB with **little to no drop in accuracy** for inference on resource-limited MCUs



Reduce model optimization trial & error from weeks to days using Deeplite's **design space exploration**



Deploy more models to your device without sacrificing performance or battery life with our **easy-to-use software**

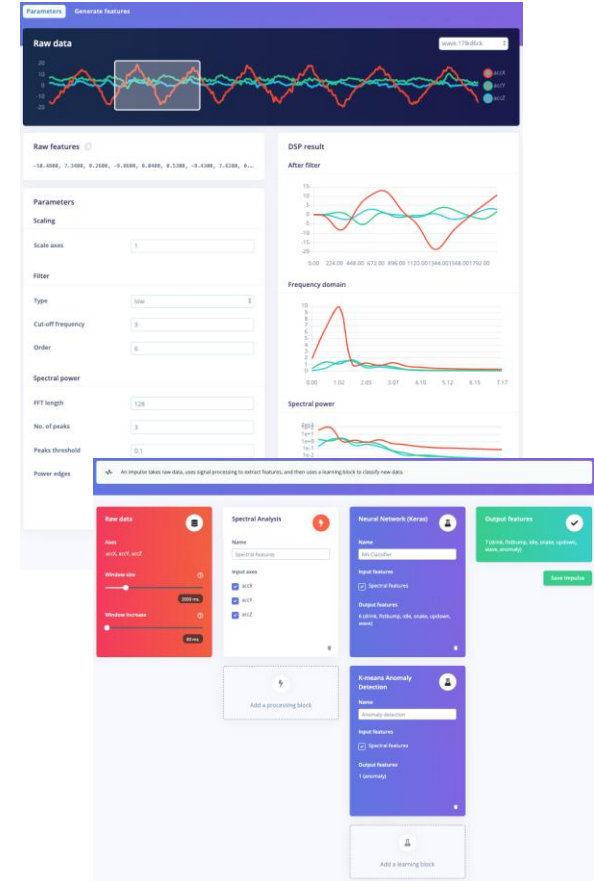
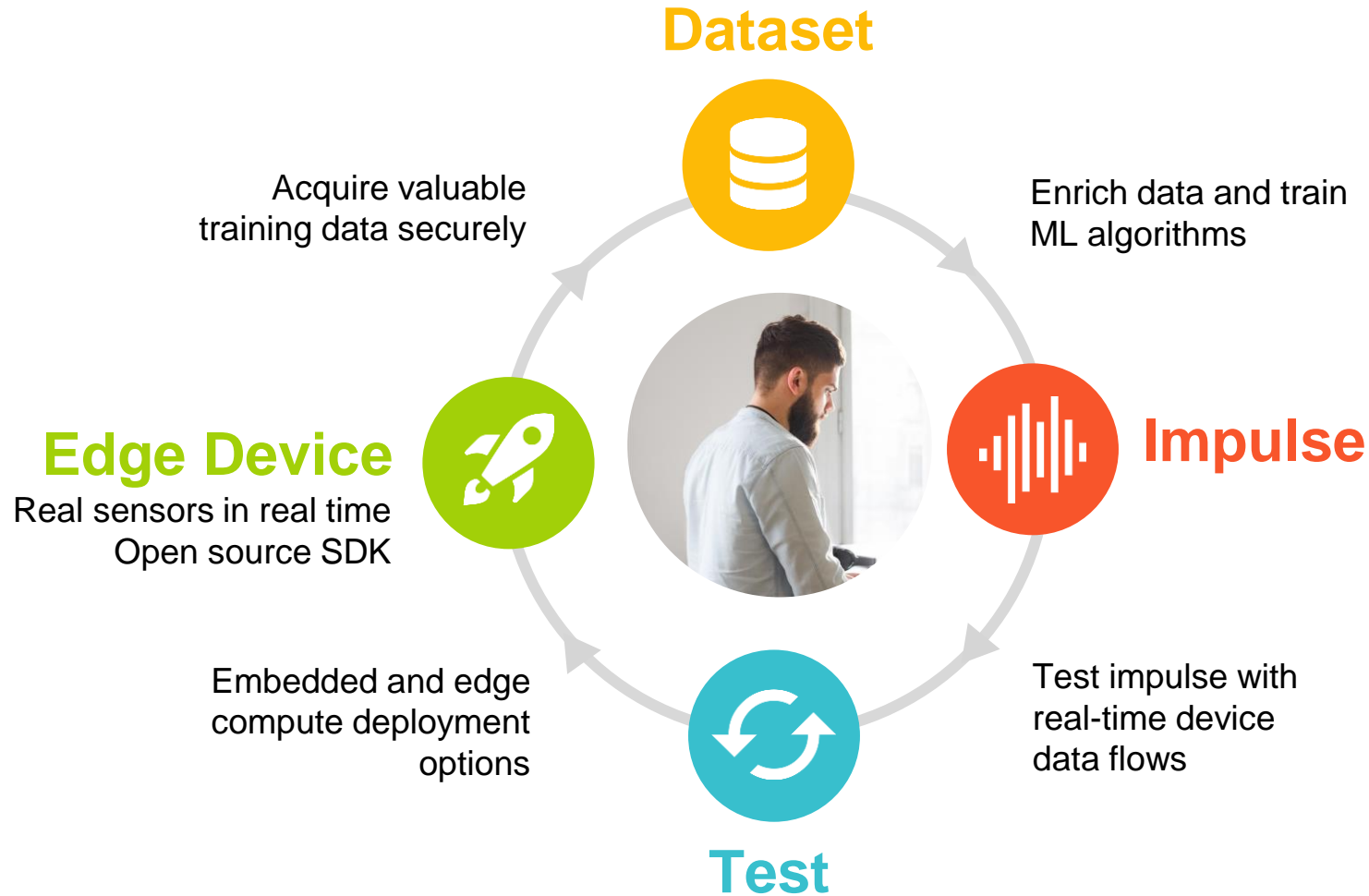
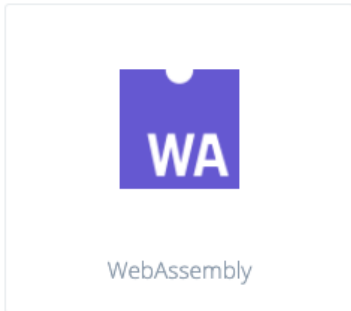
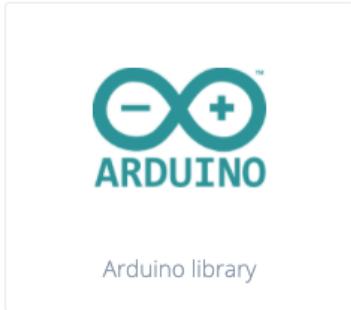
BECOME BETA USER bit.ly/testdeeplite

mobilityXlab

arm

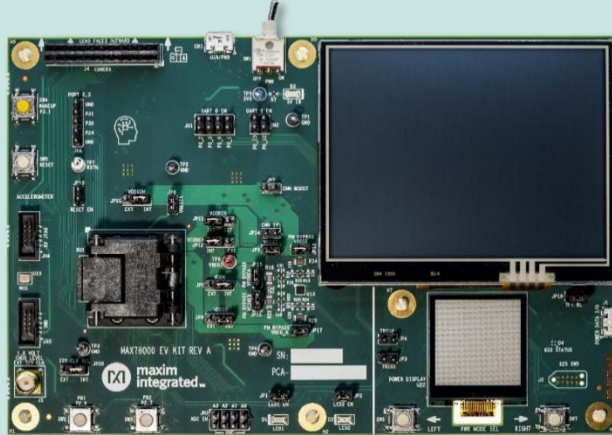


TinyML for all developers



Maxim Integrated: Enabling Edge Intelligence

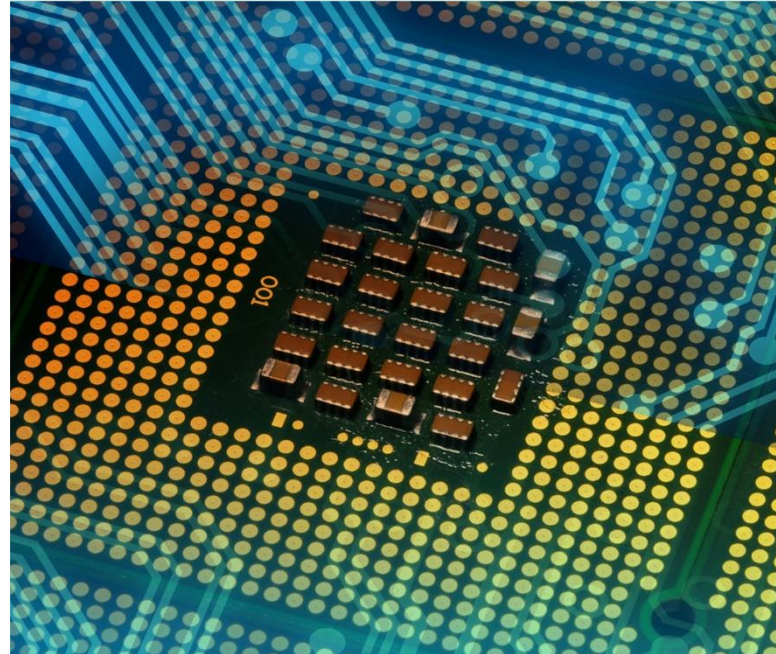
Advanced AI Acceleration IC



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/microcontroller
S

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/sensor
S

Qeexo AutoML

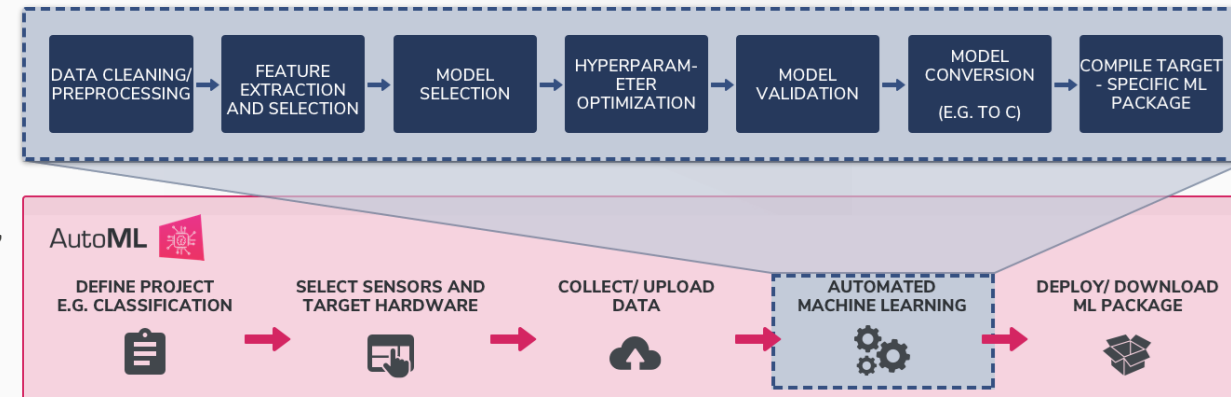


Automated Machine Learning Platform that builds tinyML solutions for the Edge using sensor data

Key Features

- Supports 17 ML methods:
 - Multi-class algorithms: GBM, XGBoost, Random Forest, Logistic Regression, Gaussian Naive Bayes, Decision Tree, Polynomial SVM, RBF SVM, SVM, CNN, RNN, CRNN, ANN
 - Single-class algorithms: Local Outlier Factor, One Class SVM, One Class Random Forest, Isolation Forest
- Labels, records, validates, and visualizes time-series sensor data
- On-device inference optimized for low latency, low power consumption, and small memory footprint applications
- Supports Arm® Cortex™- M0 to M4 class MCUs

End-to-End Machine Learning Platform



For more information, visit: www.qeexo.com

Target Markets/Applications

- Industrial Predictive Maintenance
- Smart Home
- Wearables
- Automotive
- Mobile
- IoT



Reality AI[®]

Add Advanced Sensing to your Product with Edge AI / TinyML

<https://reality.ai>



info@reality.ai



[@SensorAI](https://twitter.com/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



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>



Successful tinyML Summit 2021:

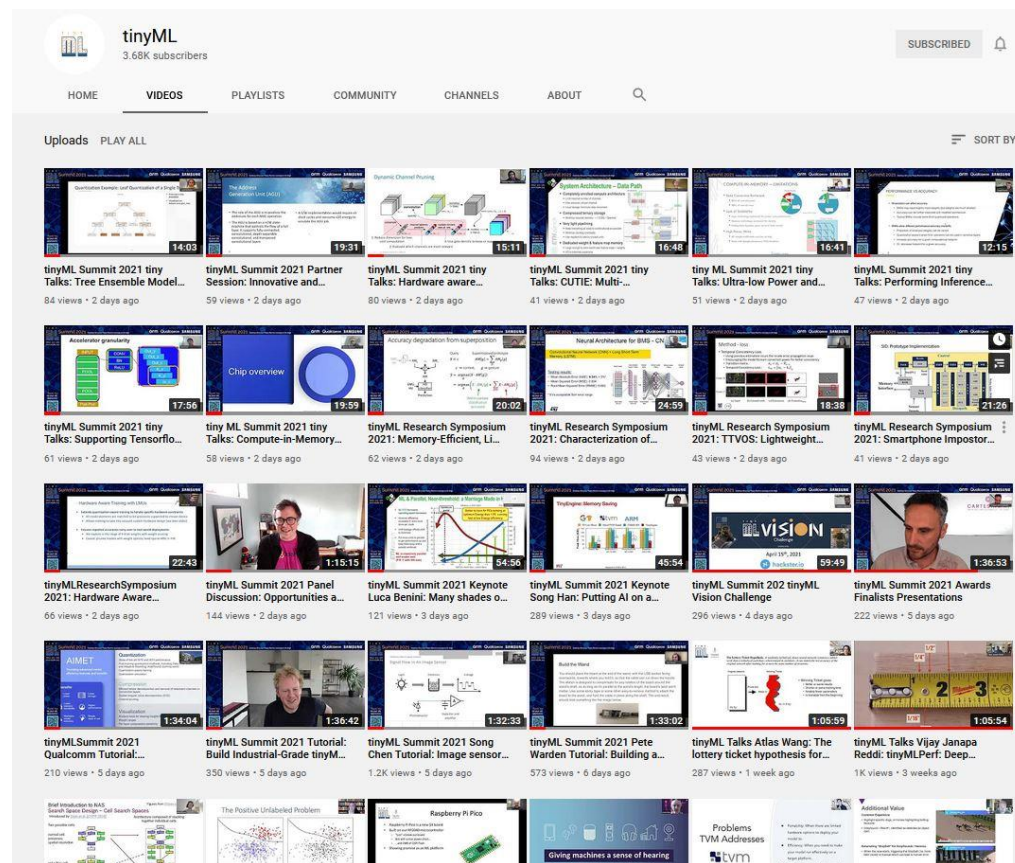
- **5** days of tutorials, talks, panels, breakouts, symposium

- **4** tutorials
- **6** keynotes & **6** plenary tinyTalks (more in breakouts)
- **2** panel discussions
- **5** disruptive news presentations
- **17** breakout/partner sessions
- **6** Best Product and Innovation Award Finalists & Presentations
- **89** Speakers



- **5006** registered attendees representing:
 - **104** countries, **1000+** companies and **400+** academic institutions
- **26** Sponsoring companies

www.youtube.com/tinyML with 150+ videos



tinyML Summit-2022, January 24-26, Silicon Valley, CA





EMEA

June 7-10, 2021 (virtual, but LIVE)
Deadline for abstracts: May 1

https://www.tinyml.org/event/emea-2021

TINY ML

Summit 2021 Research Symposium All Events



tinyML EMEA Technical Forum 2021

Enabling ultra-low Power Machine Learning at the Edge

June 7-10, 2021

Inaugural tinyML EMEA Technical Forum

Venue

Virtual - online

Sponsorships are being accepted: sponsorships@tinyML.org



Next tinyML Talks

Date	Presenter	Topic / Title
Tuesday, April 27	Michael Jo and Xingheng Lin Rose-Hulman Institute of Technology	Train-by-weight (TBW): Accelerated Deep Learning by Data Dimensionality Reduction

Webcast start time is 8 am Pacific time

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

Local UK Committee



Alessandro Grande

Developer advocate &
ecosystem manager, Arm



Dominic Binks

VP Technology, Audio Analytic



Gian Marco Iodice

ML Techlead, Arm



Neil Cooper

VP Marketing, Audio Analytic

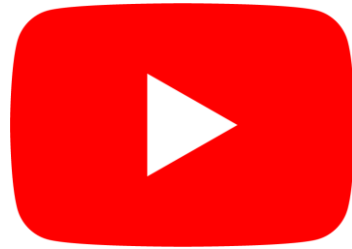


Reminders

Slides & Videos will be posted tomorrow

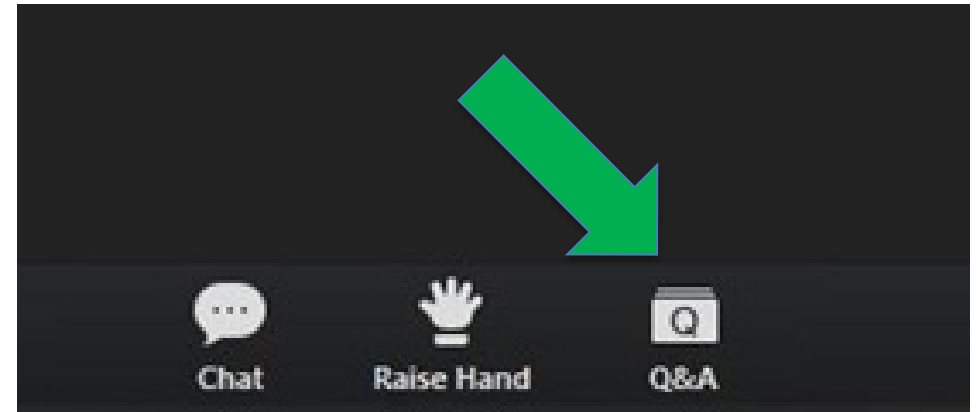


tinyml.org/forums



youtube.com/tinyml

Please use the Q&A window for your questions



Massimo Banzi



Massimo Banzi is the co-founder of the Arduino project. He is an Interaction Designer, Educator and Open Source Hardware advocate. He has worked as a consultant for clients such as: Prada, Artemide, Persol, Whirlpool, V&A Museum and Adidas. Massimo started the first FabLab in Italy which led to the creation of Officine Arduino, a FabLab/Makerspace based in Torino. He spent 4 years at the Interaction Design Institute Ivrea as Associate Professor. Massimo has taught workshops and has been a guest speaker at institutions all over the world. Before joining IDII he was CTO for the Seat Ventures incubator. He spent many years working as a software architect, both in Milan and London, on projects for clients like Italia Online, Sapient, Labour Party, BT, MCI WorldCom, SmithKlineBeecham, Storagetek, BSkyB and boo.com. Massimo is also the author of “Getting Started with Arduino” published by O’Reilly. He is a regular contributor to the Italian edition of Wired Magazine and Che Futuro, an online magazine about innovation. He currently teaches Interaction Design at SUPSI Lugano in the south of Switzerland and is a visiting professor at CIID in Copenhagen.



TinyML with Arduino

Massimo Banzi

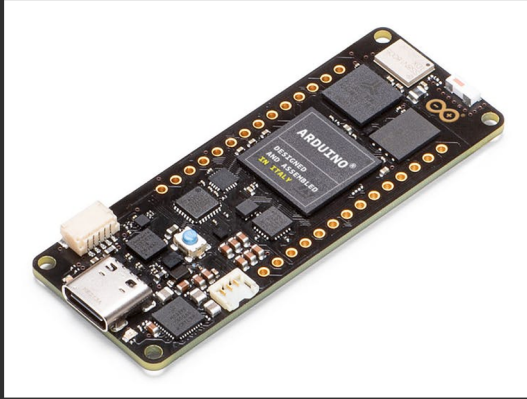
20.04.2021

A series of horizontal yellow lines of varying lengths on the left side of the slide.

Introduction

In this talk we walk through some examples of TinyML on Arduino devices to show how simple it is to build working examples of ML with minimal coding using three different platforms compatible with Arduino.

Arduino



Enabling anyone
to innovate by
making complex
technologies
simple to use.

29M+

Users/Year on arduino.cc

900K+

Arduino Forum Users

17M+

IDE downloads/year

1.2M+

Users of Arduino Create Cloud

4000+

Libraries Available for Arduino



923K



302K



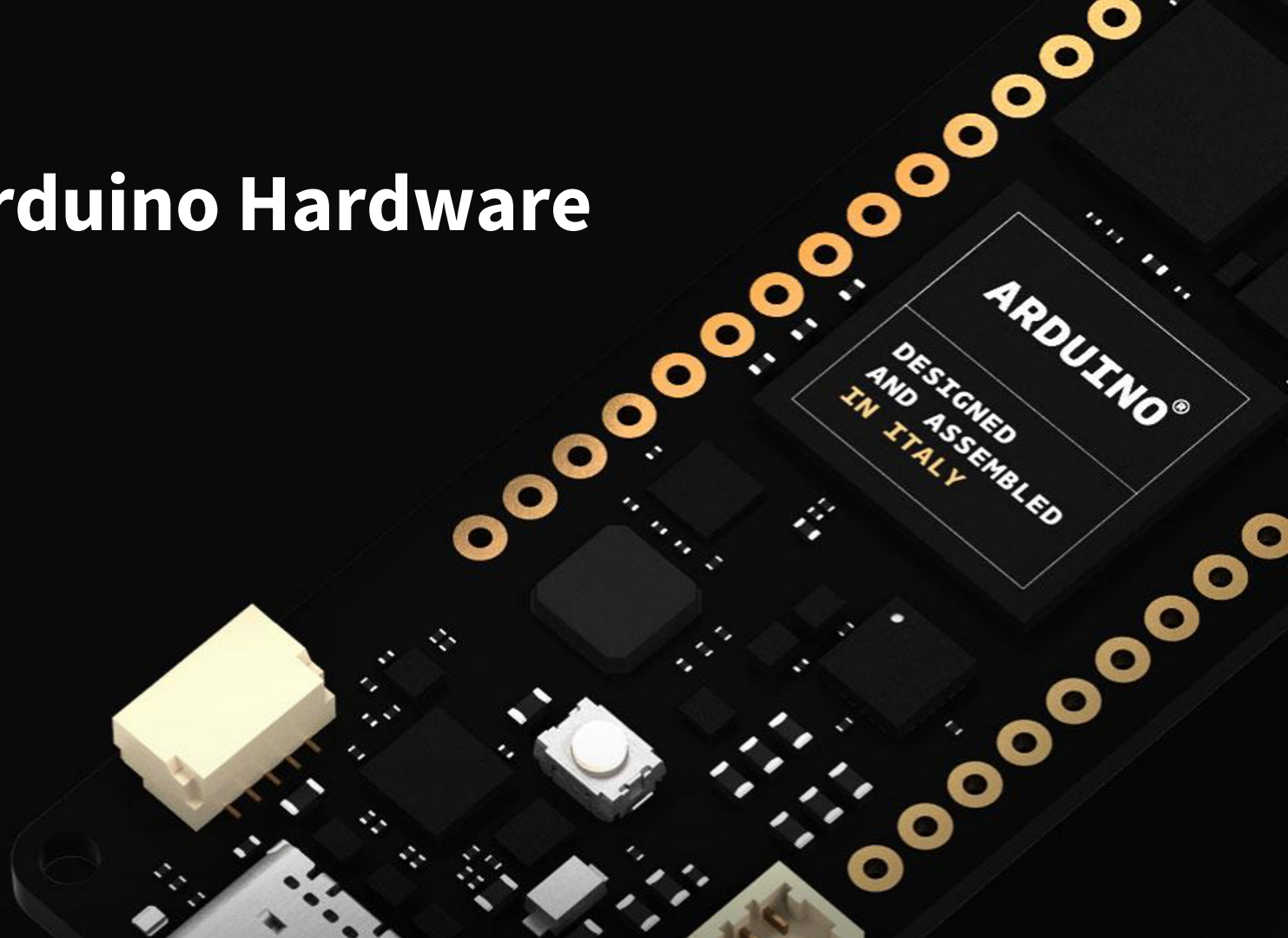
408K

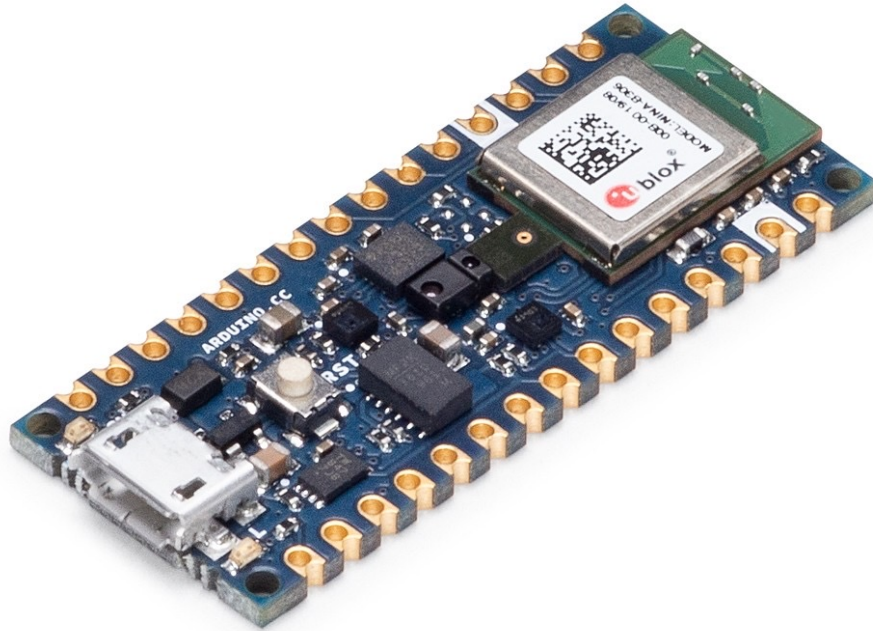
A series of horizontal yellow lines of varying lengths on the left side of the slide.

Arduino Today

**Easy for
Beginners,
Fast for
professionals**

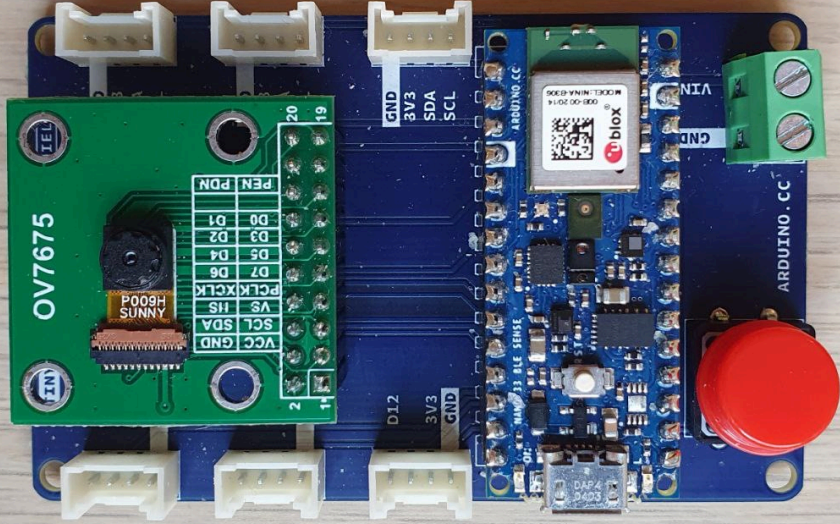
Arduino Hardware





Arduino Nano BLE Sense

- Nano form factor
- nRF52840 Arm Cortex M4f @ 64MHz
- 1MByte Flash
- 256Kb Ram
- Bluetooth 5
- Sensors: IMU, Microphone, Gesture, Light, Proximity, Pressure, Temperature , Humidity



Tiny Machine Learning Kit

- VGA Camera and Nano BLE Sense
- Used for the EdX course on TinyML



TINY MACHINE
LEARNING KIT

In partnership with:



TensorFlow

Catalog > Data Science Courses > HarvardX's Tiny Machine Learning (TinyML)

The Future of ML is Tiny and Bright



Professional Certificate in
Tiny Machine Learning (TinyML)

I'm interested ✓

<https://www.edx.org/professional-certificate/harvardx-tiny-machine-learning>

- Fundamentals of machine learning and embedded devices.

Self-paced



CS249r: Tiny Machine Learning

Applied Machine Learning for Embedded IoT Devices

Welcome

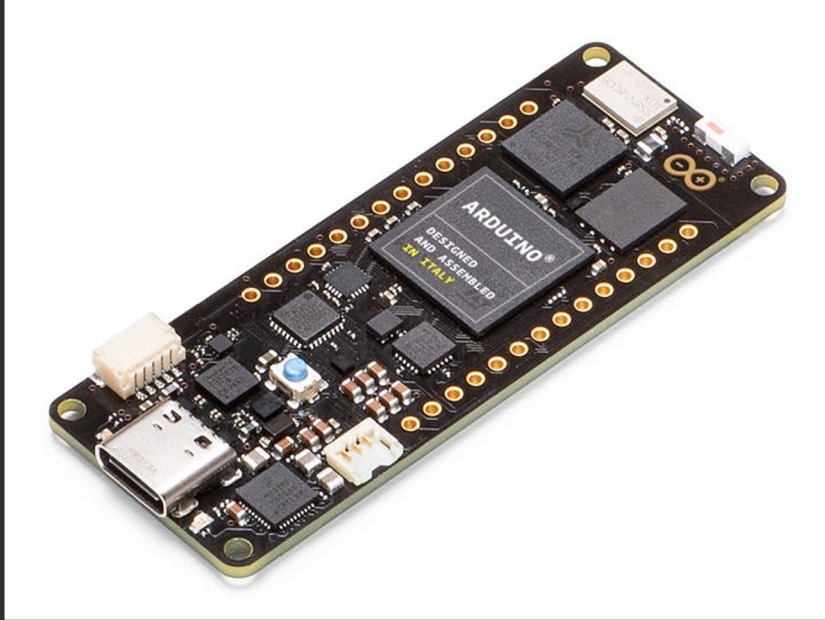


Course Overview

Tiny Machine Learning (TinyML) is an introductory course at the intersection of Machine Learning and Embedded IoT Devices. The pervasiveness of ultra-low-power embedded devices, coupled with the introduction of embedded machine learning frameworks like TensorFlow Lite for Microcontrollers, will enable the mass proliferation of AI-powered IoT devices. The explosive growth in machine learning and the ease of use of platforms like TensorFlow (TF) make it an indispensable topic of study for modern computer science *and* electrical engineering students.

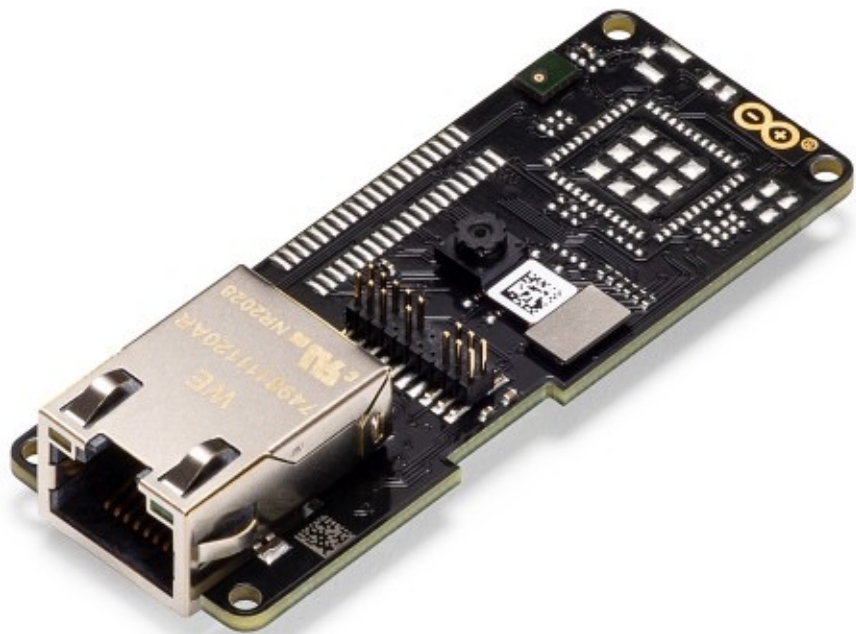
<https://sites.google.com/g.harvard.edu/tinyml/home>

The course will cover all things related to machine learning on embedded devices. The topics range from tinyML applications and algorithms to the design of the frameworks built for running ML on embedded systems to the microcontroller hardware designed, optimized, and built for ultra-low-power (milliwatts) computing.



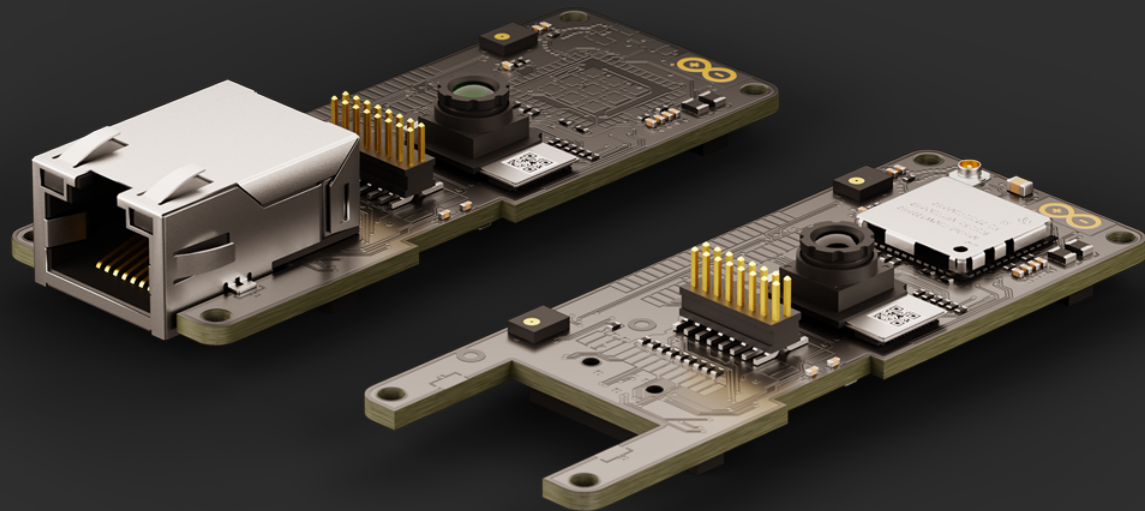
Portenta H7

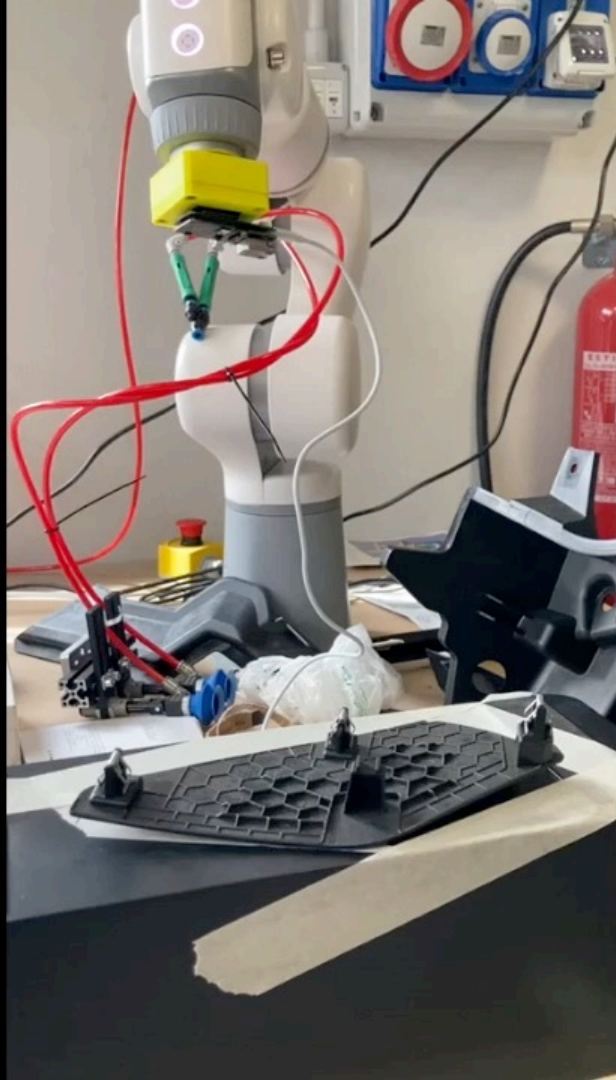
- Portenta form factor
- Dual core
 - Arm Cortex M7@480MHz
 - Arm Cortex M4@240MHz
- up to 64 MByte SDRAM
- up to 128 MByte QSPI Flash
- Common Criteria Crypto
- WiFi b/g/n 65Mbps + BT 5.1
- 100Mbit Ethernet
- CAN
- USB-C with DisplayPort output



Portenta Vision Shield

- Portenta form factor
- Low power cameras
 - Himax HM-01B0
 - Himax HM-0360
- Dual microphone (beamforming)
- Ethernet
- Lora
- SD Card







EXT. POWER TERMINAL BLOCK

MIPI 20T JTAG

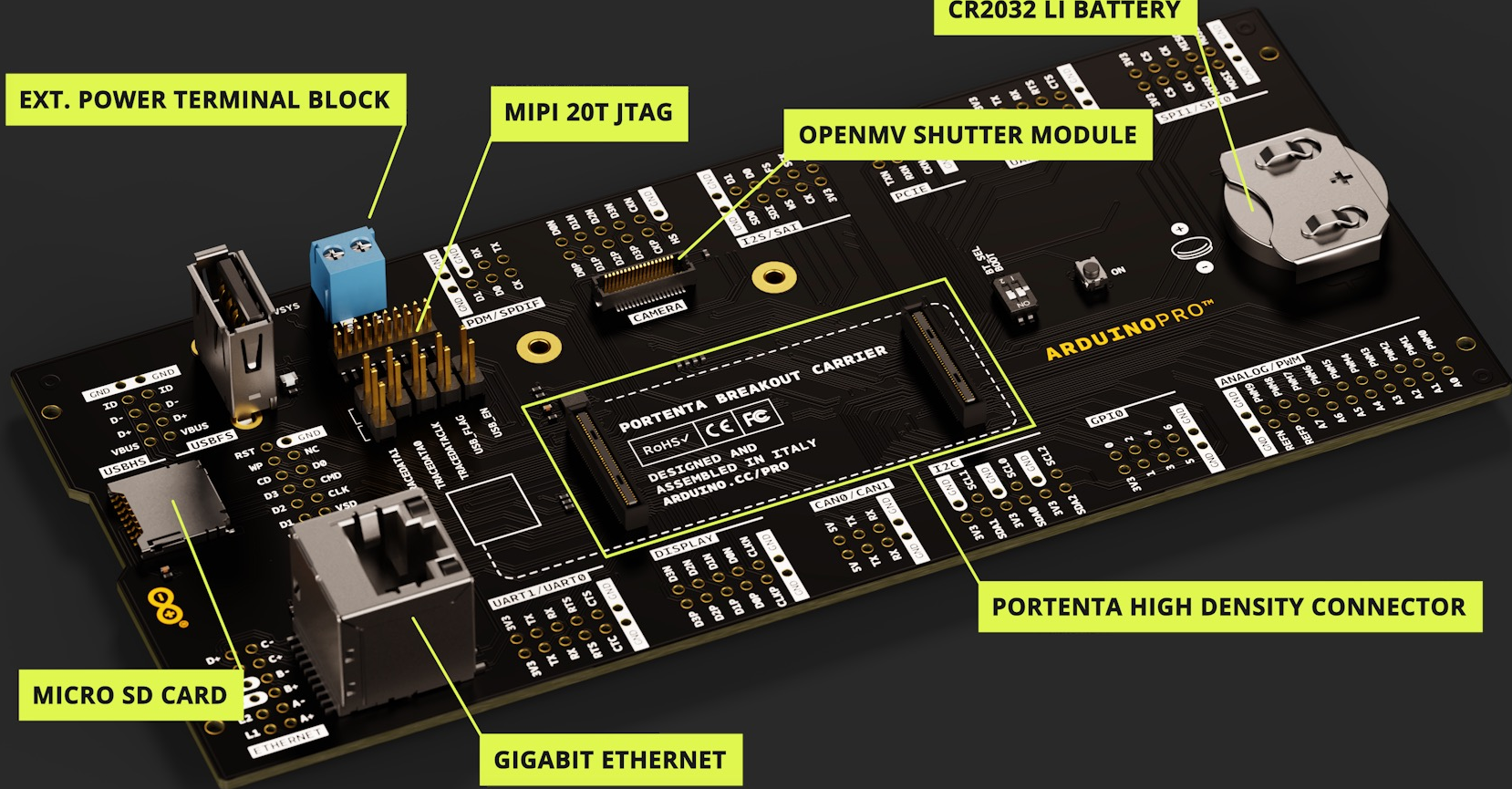
CR2032 LI BATTERY

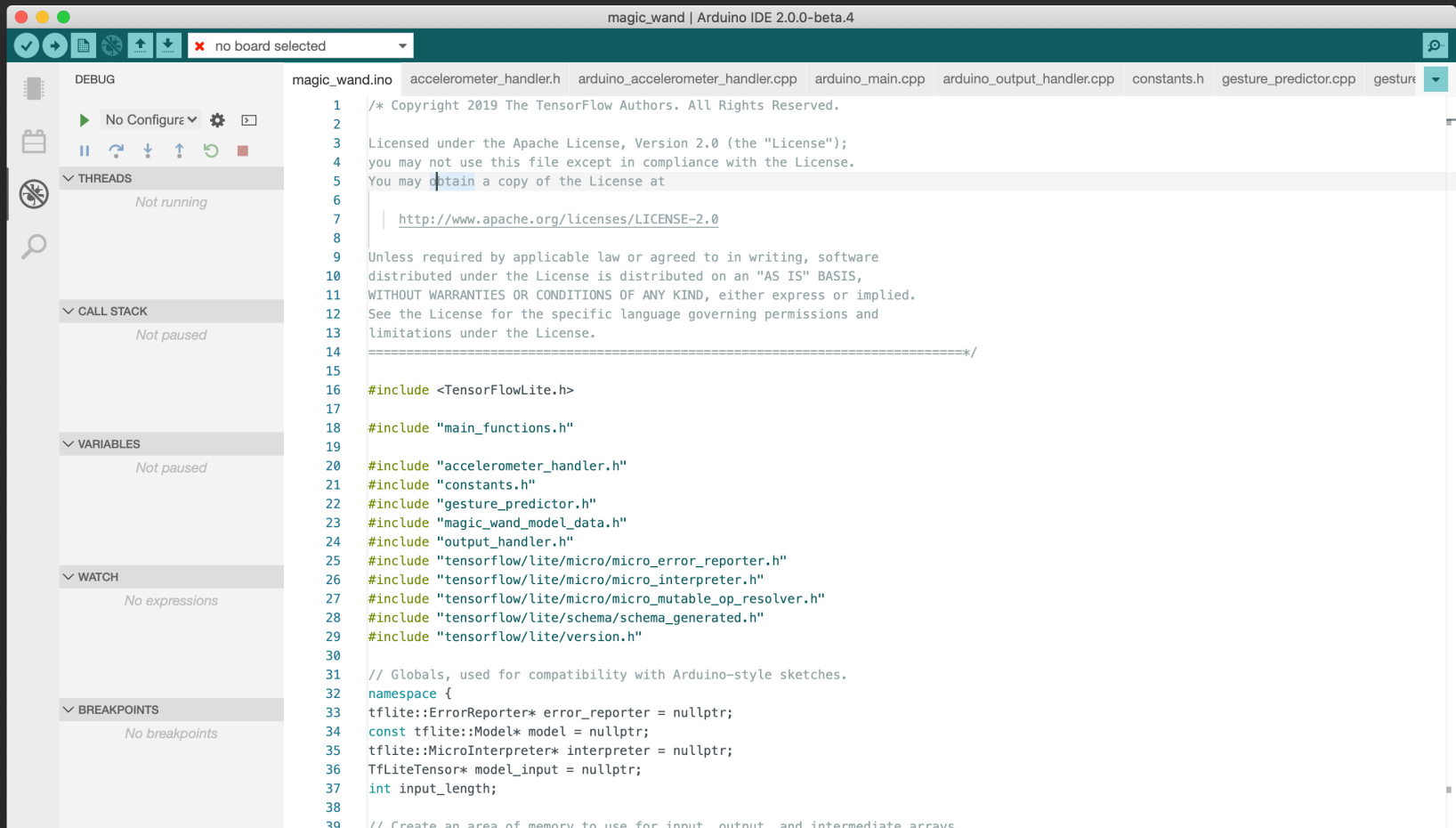
OPENMV SHUTTER MODULE

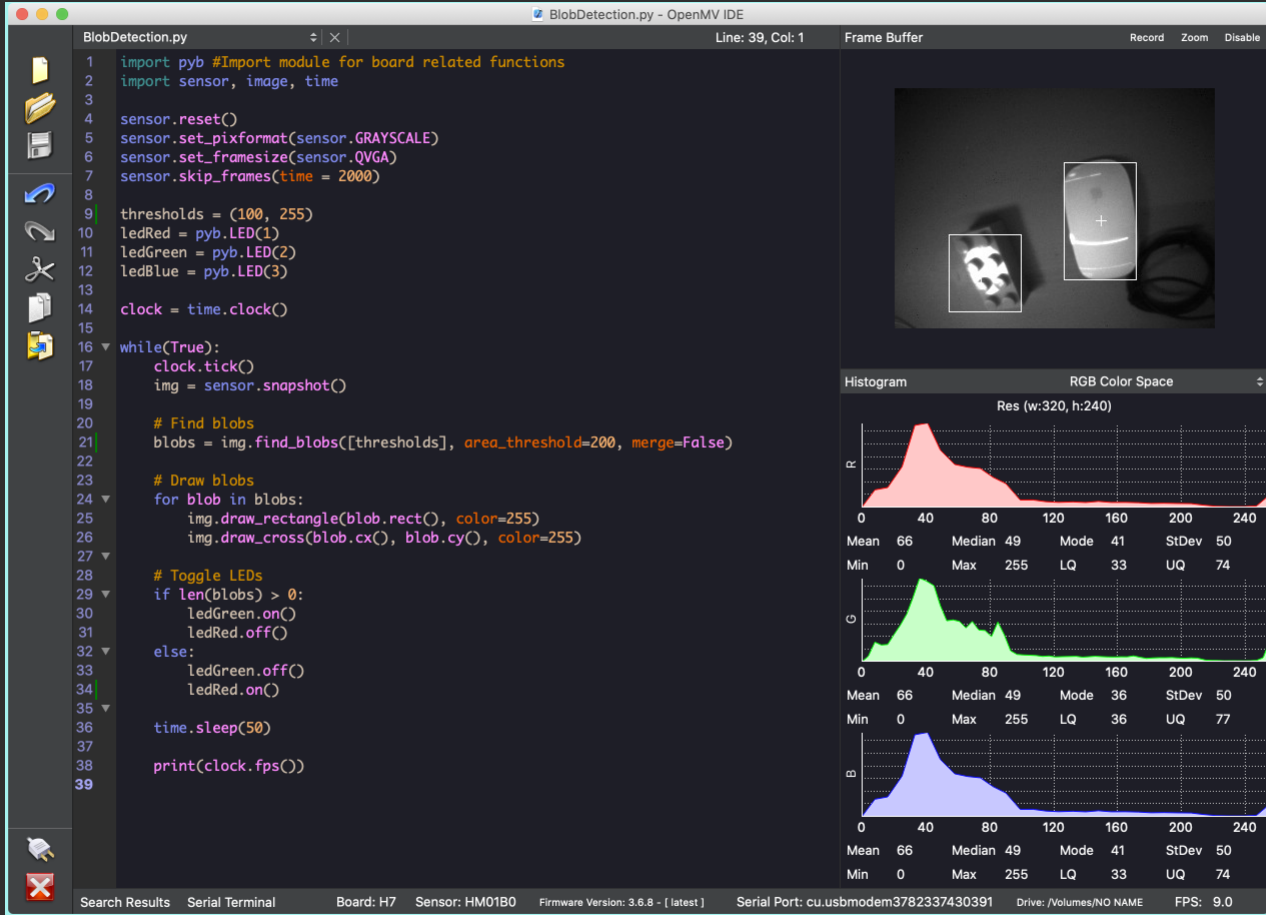
MICRO SD CARD

GIGABIT ETHERNET

PORTENTA HIGH DENSITY CONNECTOR









Examples

Arduino Tiny ML

For Mobile & IoT

Overview

Tutorials

Guide

Examples

Models

API

► Integrate models with Task Library
Customize input and output data processing

Performance

Best practices

Measurement

Delegates 

GPU delegate

Advanced GPU

NNAPI delegate

Hexagon delegate Core ML delegate 

Optimize a model

Overview

Post-training quantization

Post-training dynamic range quantization

Post-training integer quantization

Post-training float16 quantization

Quantization specification

Build TensorFlow Lite

Build for Android

Build for iOS

Build for ARM64

Build for Raspberry Pi

TensorFlow > Learn > For Mobile & IoT > Guide



TensorFlow Lite for Microcontrollers

Contents

[Get started](#)[Supported platforms](#)[Why microcontrollers are important](#)[Developer workflow](#)[Limitations](#)

TensorFlow Lite for Microcontrollers is an experimental port of TensorFlow Lite designed to run machine learning models on microcontrollers and other devices with only kilobytes of memory.

It doesn't require operating system support, any standard C or C++ libraries, or dynamic memory allocation. The core runtime fits in 16 KB on an Arm Cortex M3, and with enough operators to run a speech keyword detection model, takes up a total of 22 KB.

There are example applications demonstrating the use of microcontrollers for tasks including wake word detection, gesture classification from accelerometer data, and image classification using camera data.

Get started

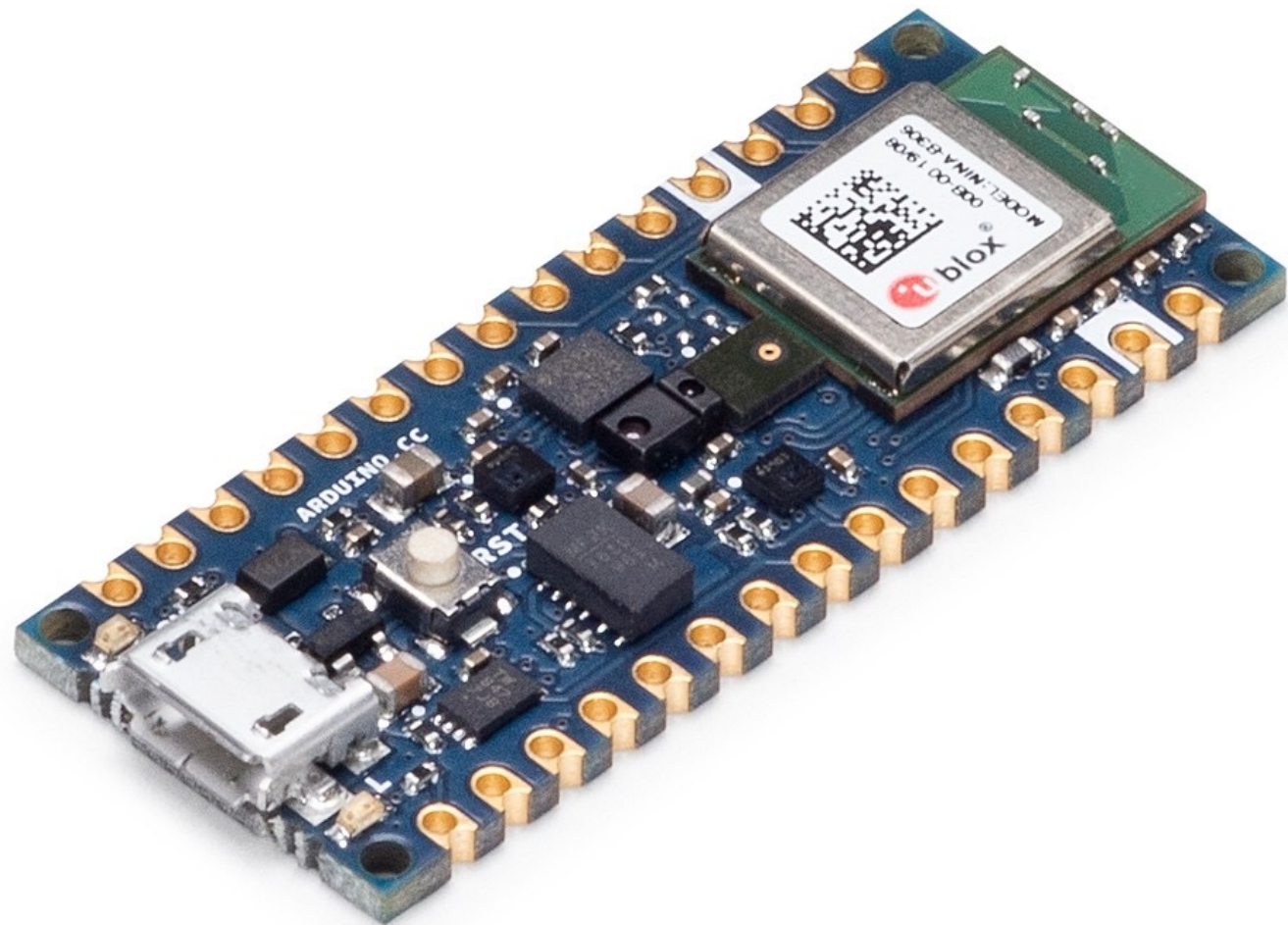
To try the example applications and learn how to use the API, read [Get started with microcontrollers](#).

Supported platforms

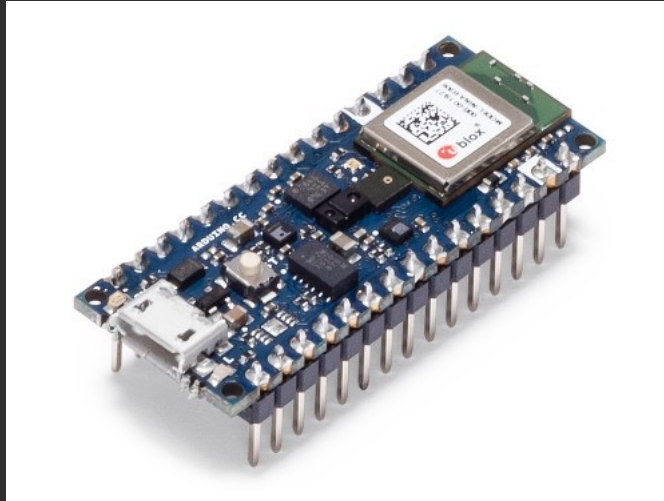
microcontrollers

[Overview](#)

TensorFlow Lite for Microcontrollers / Gesture Recognition



Gesture Recognition

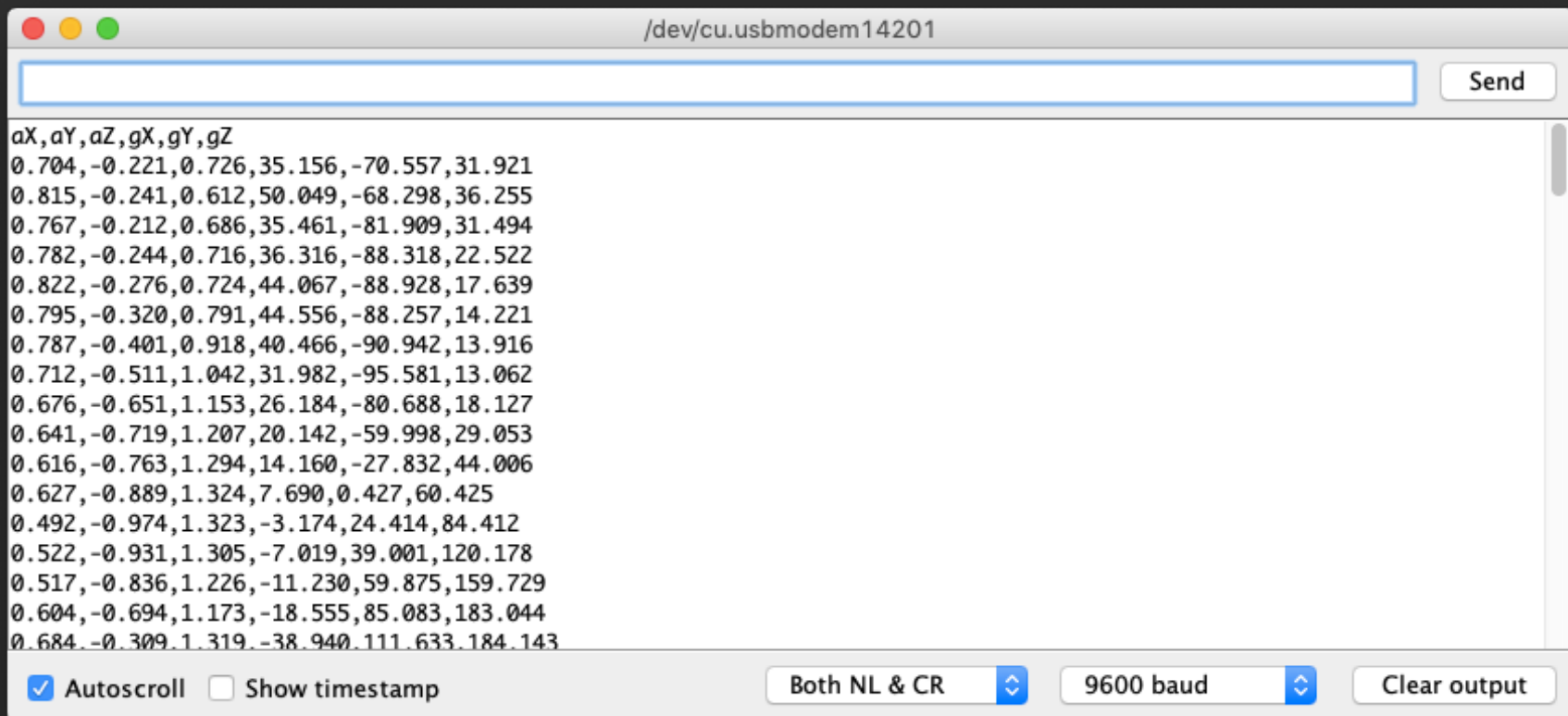


We're going to use the code provided here:

<https://github.com/arduino/ArduinoTensorFlowLiteTutorials/tree/master/GestureToEmoji>

Visualize the IMU Data

- Open **ArduinoSketches/IMU_Capture/IMU_Capture.ino** in the Arduino IDE.
- Compile the sketch and upload it to the board: Sketch -> Upload
- Open the Serial Monitor: Tools -> Serial Monitor
- Press the button, IMU data will be captured and outputted for 1 second
- Close the Serial Monitor window
- Open the Serial Plotter: Tools -> Serial Plotter
- Press the button, and perform a gesture
- You'll see a graph of the data capture
- Repeat capturing various gestures to get a sense of what the training data will look like
- Close the Serial Plotter

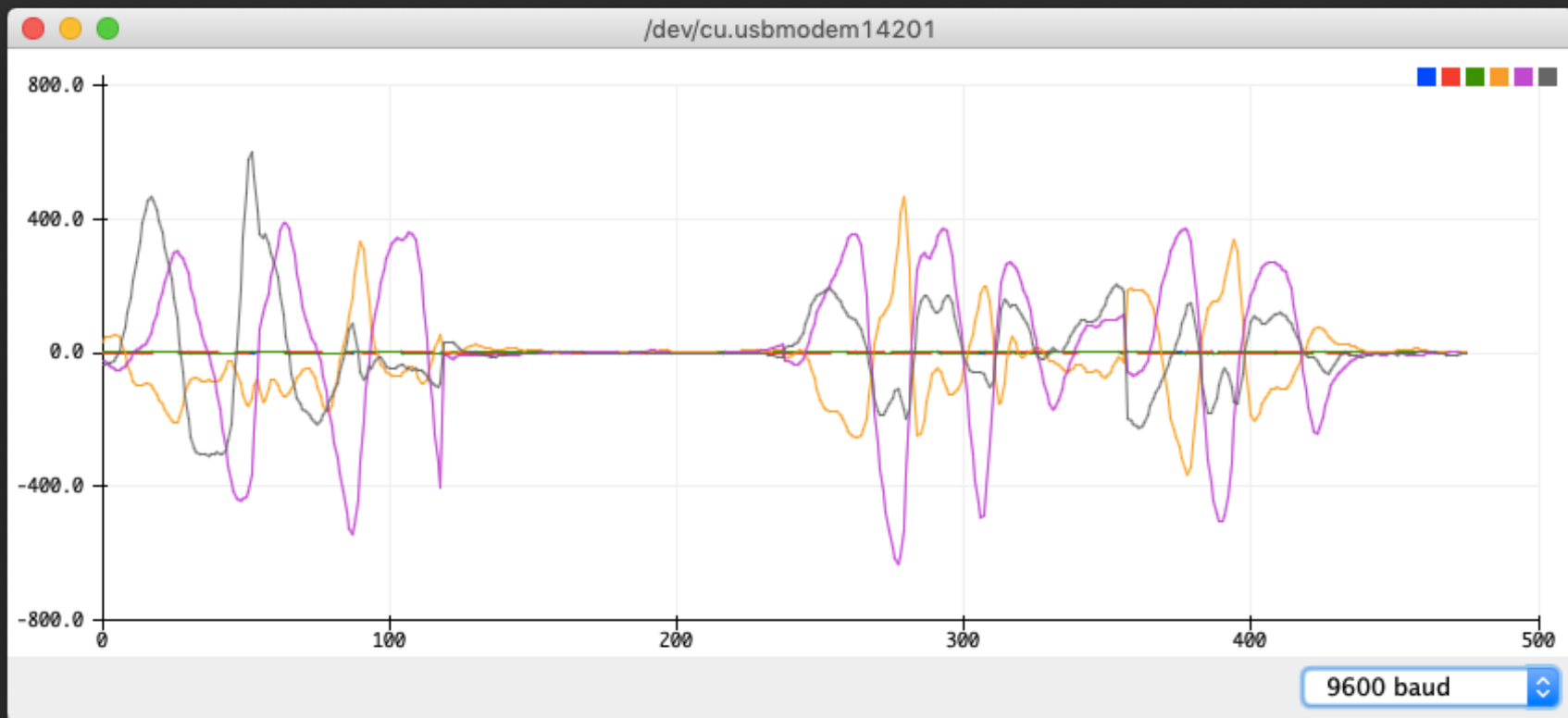


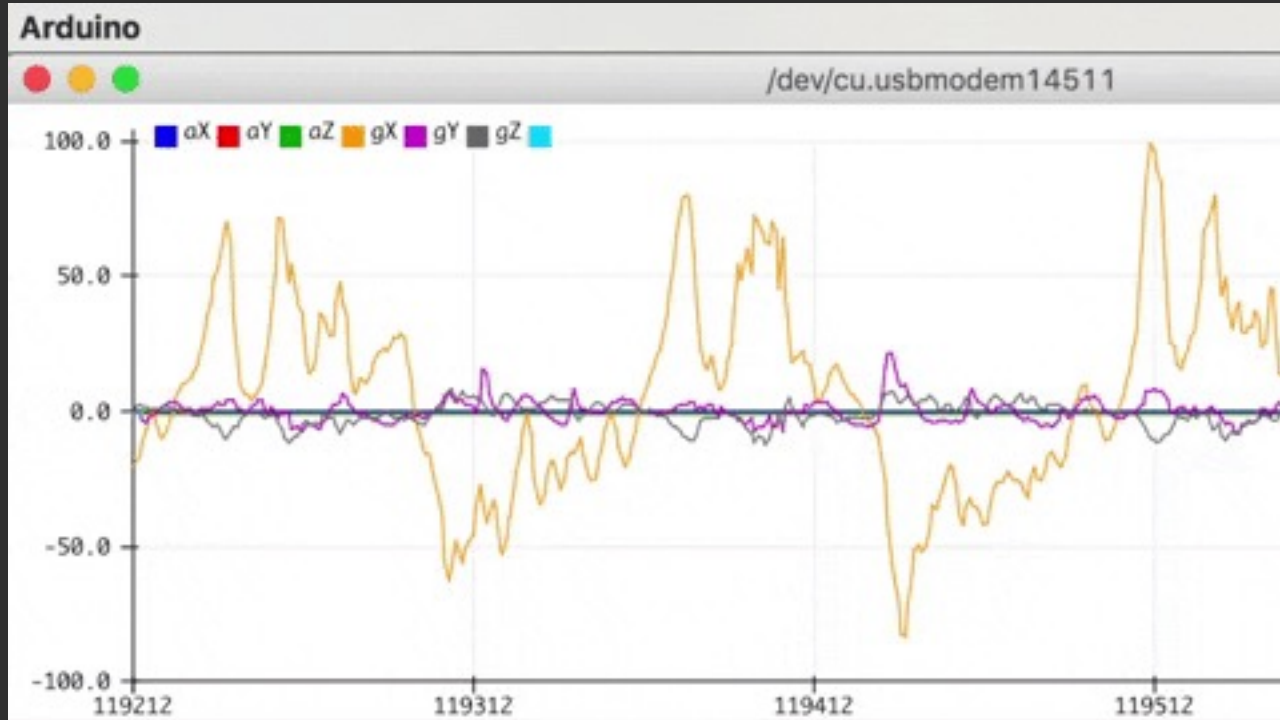
```

aX,aY,aZ,gX,gY,gZ
0.704,-0.221,0.726,35.156,-70.557,31.921
0.815,-0.241,0.612,50.049,-68.298,36.255
0.767,-0.212,0.686,35.461,-81.909,31.494
0.782,-0.244,0.716,36.316,-88.318,22.522
0.822,-0.276,0.724,44.067,-88.928,17.639
0.795,-0.320,0.791,44.556,-88.257,14.221
0.787,-0.401,0.918,40.466,-90.942,13.916
0.712,-0.511,1.042,31.982,-95.581,13.062
0.676,-0.651,1.153,26.184,-80.688,18.127
0.641,-0.719,1.207,20.142,-59.998,29.053
0.616,-0.763,1.294,14.160,-27.832,44.006
0.627,-0.889,1.324,7.690,0.427,60.425
0.492,-0.974,1.323,-3.174,24.414,84.412
0.522,-0.931,1.305,-7.019,39.001,120.178
0.517,-0.836,1.226,-11.230,59.875,159.729
0.604,-0.694,1.173,-18.555,85.083,183.044
0.684,-0.309,1.319,-38.940,111.633,184.143

```

☒ Autoscroll ☐ Show timestamp Both NL & CR 9600 baud Clear output





Gather Training Data

- Press the reset button on the board
- Open the Serial Monitor: Tools -> Serial Monitor
- Make a punch gesture with the board in your hand - you should see the sensor data log in the Serial Monitor
- Repeat 10 times to gather more data
- Copy and paste the data from the serial output to new text file called punch.csv using your favorite text editor
- Close the Serial Monitor
- Press the reset button on the board
- Open the Serial Monitor: Tools -> Serial Monitor
- Make a flex gesture with the board in your hand
- Repeat 10 times
- Copy and paste the serial output to new text file flex.csv using your favorite text editor

+ Code + Text

RAM Disk Editing

Table of contents Code snippets Files X

Tiny ML on Arduino

Setup Python Environment

Upload Data

Graph Data (optional)

Train Neural Network

Parse and prepare the data

Randomize and split the input and output pairs for training

Build & Train the Model

Verify

Graph the loss

Graph the loss again, skipping a bit of the start

Graph the mean absolute error

Run with Test Data

Convert the Trained Model to Tensor Flow Lite

Encode the Model in an Arduino Header File

Classifying IMU Data



Tiny ML on Arduino

Gesture recognition tutorial

- Sandeep Mistry - Arduino
- Don Coleman - Chariot Solutions

<https://github.com/arduino/ArduinoTensorFlowLiteTutorials/>

Setup Python Environment

The next cell sets up the dependencies in required for the notebook, run it.

```
[ ] # Setup environment
!apt-get -qq install xxd
!pip install pandas numpy matplotlib
!pip install tensorflow==2.0.0-rc1
```

Upload Data

1. Open the panel on the left side of Colab by clicking on the >
2. Select the files tab

Machine Learning

- Go to this colab:
- https://colab.research.google.com/github/arduino/ArduinoTensorFlowLiteTutorials/blob/master/GestureToEmoji/arduino_tinyml_workshop.ipynb
- Drag punch.csv and flex.csv files from your computer to the tab to upload them into colab.
- Follow the instructions on the Notebook and you'll end up with "model.h" file you can import into arduino

Classifying IMU Data

- Open ArduinoSketches/IMU_Classifier/IMU_Classifier.ino in the Arduino IDE.
- Switch to the model.h tab
- Replace the contents of model.h with the version you downloaded from Colab
- Upload the sketch: Sketch -> Upload
- Open the Serial Monitor: Tools -> Serial Monitor
- Press the button, and perform a gesture
- The confidence of each gesture will be printed to the Serial Monitor (0 -> low confidence, 1 -> high confidence)



A terminal window titled "/dev/cu.usbmodem14101" displays a series of sensor readings. The window has a title bar with red, yellow, and green window control buttons. At the top right is a "Send" button. The main area contains text output with a vertical scrollbar on the right. At the bottom, there are controls for "Autoscroll" (checked), "Show timestamp" (unchecked), a "Newline" dropdown menu, a "9600 baud" dropdown menu, and a "Clear output" button.

```
/dev/cu.usbmodem14101

punch: 0.529452
flex: 0.470548

punch: 0.917286
flex: 0.082714

punch: 0.877585
flex: 0.122415

punch: 0.000002
flex: 0.999997

punch: 0.176803
flex: 0.823197

punch: 0.999973
flex: 0.000027

punch: 0.077787
flex: 0.922213

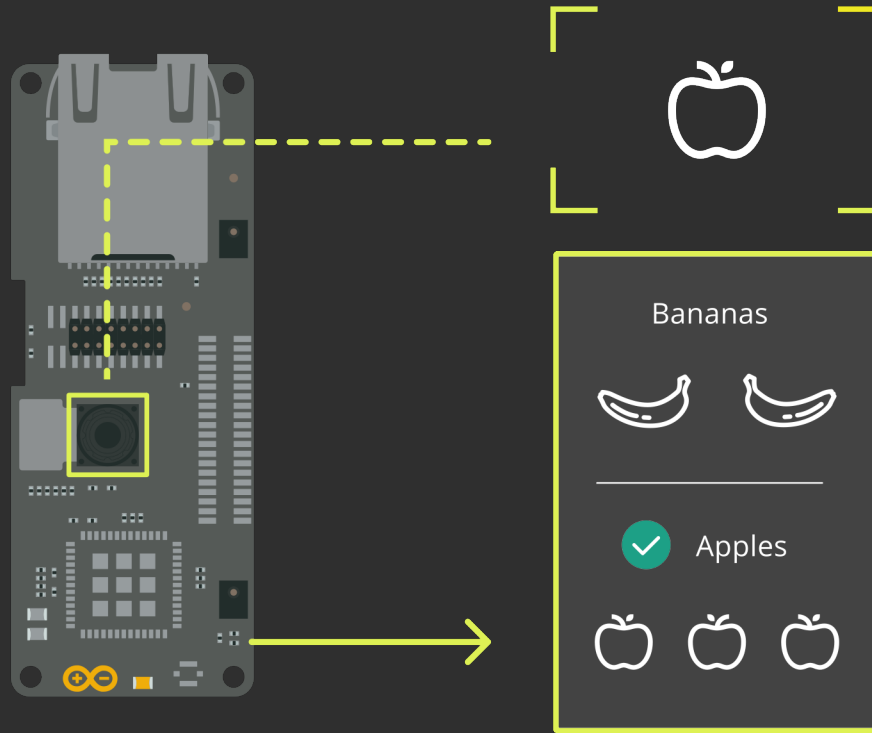
punch: 0.036865
flex: 0.963135

☒ Autoscroll ☐ Show timestamp
Newline 9600 baud Clear output
```

Gesture Controlled USB Emoji Keyboard

- The Emoji_Button example, ArduinoSketchs/Emoji_Button/Emoji_Button.ino, shows how to create a USB keyboard that prints and emoji character. (This only works on Linux and macos, so if you're running Windows, find a friend to work on this exercise.)
- Try combining the Emoji_Button example with the IMU_Classifier sketch to create an gesture controlled emoji keyboard. 🦊



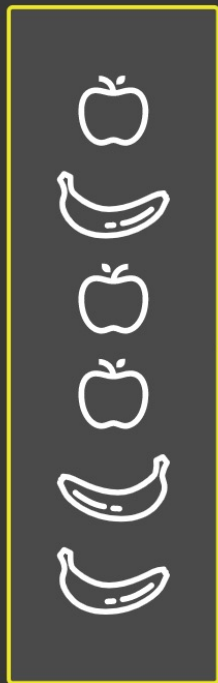


REQUIRED HARDWARE AND SOFTWARE

- Portenta H7 board
- Portenta Vision Shield - LoRa or Portenta Vision Shield - Ethernet
- USB-C cable (either USB-A to USB-C or USB-C to USB-C)
- An Edge Impulse account for training the ML model
- Fruits (or other objects) to create the classification model



INPUT DATA



Banana



Apple



LEARNING



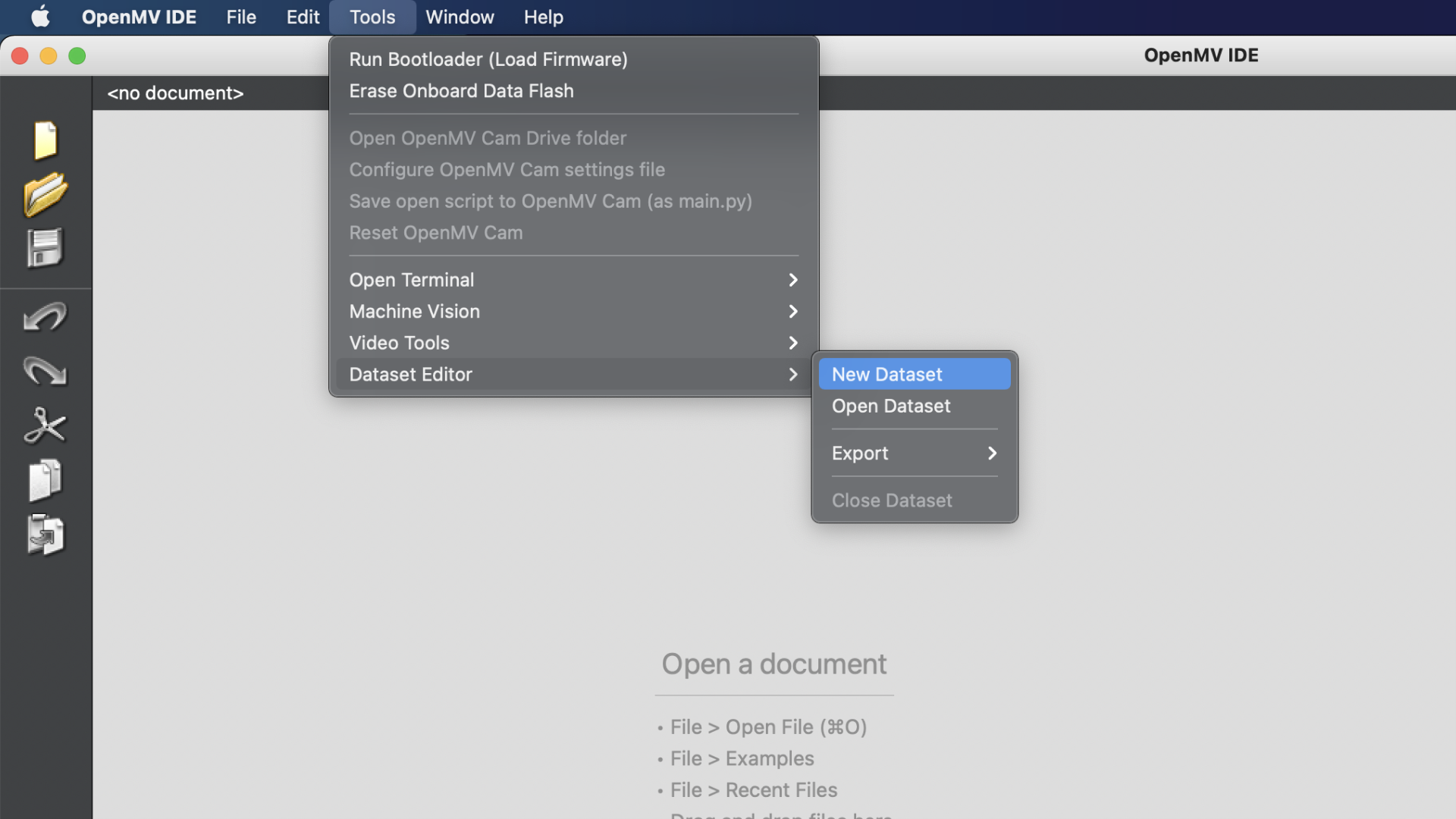
CLASSES

Bananas



Apples





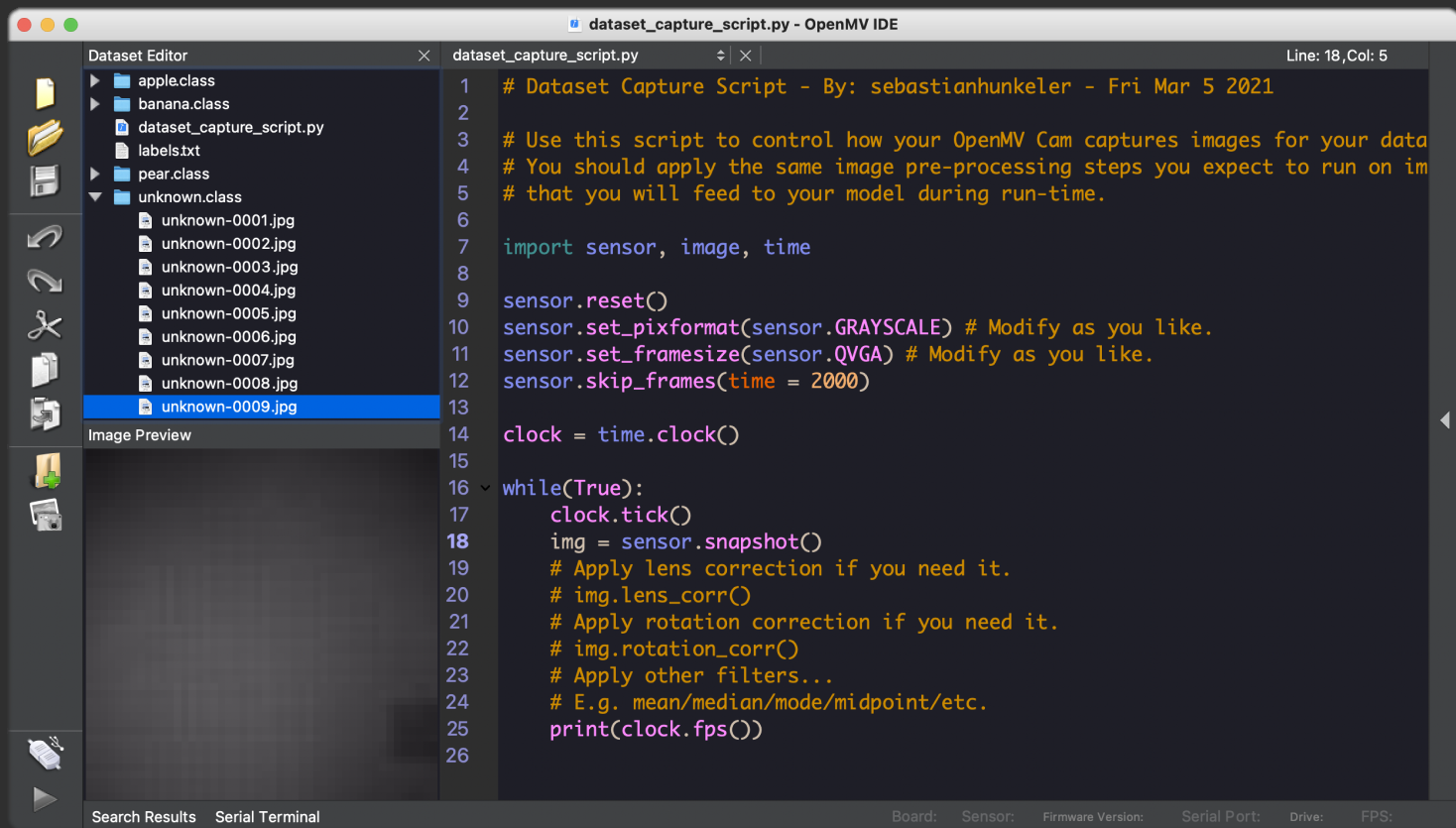
<no document>

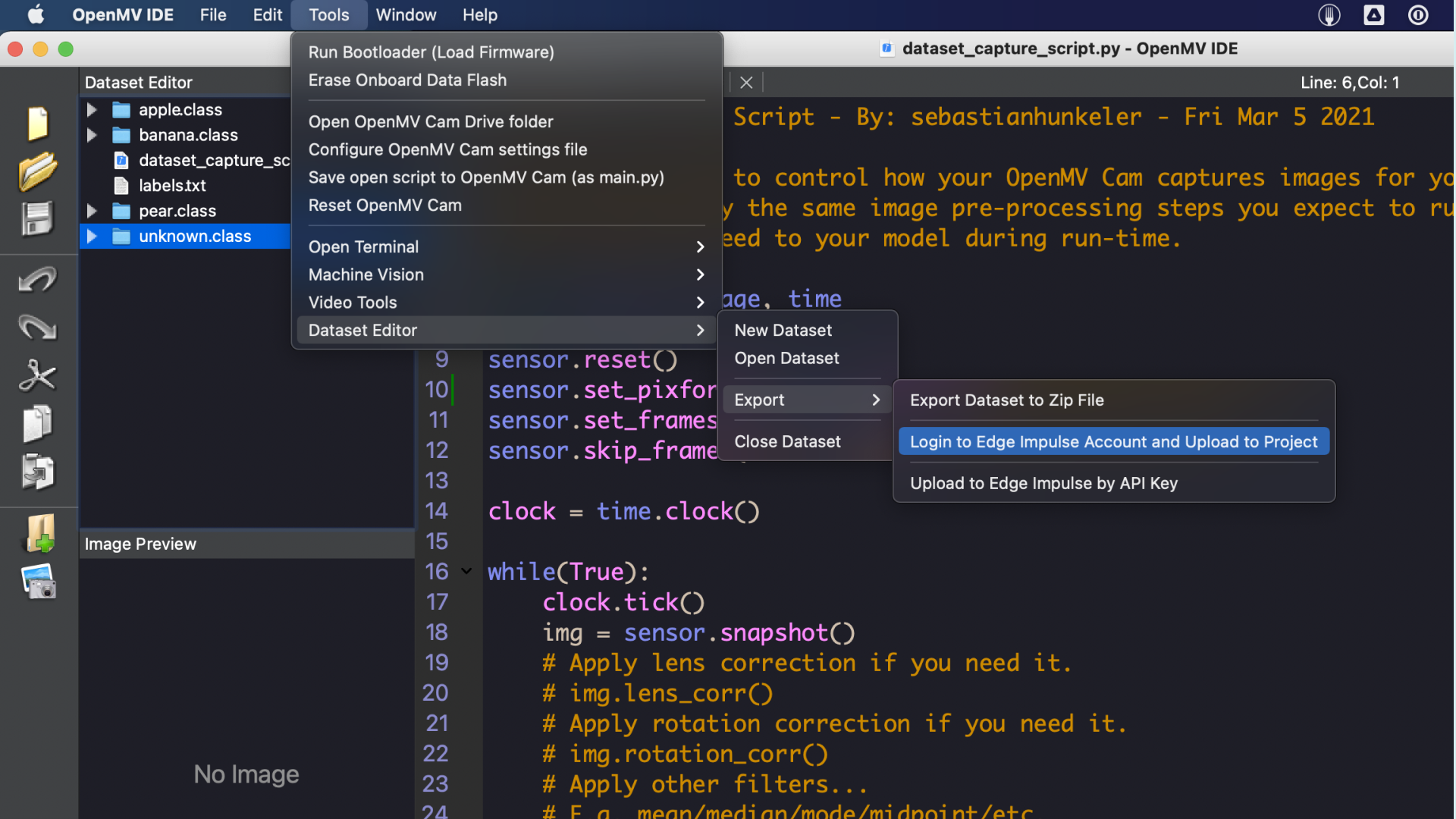
- Run Bootloader (Load Firmware)
- Erase Onboard Data Flash
- Open OpenMV Cam Drive folder
- Configure OpenMV Cam settings file
- Save open script to OpenMV Cam (as main.py)
- Reset OpenMV Cam
- Open Terminal >
- Machine Vision >
- Video Tools >
- Dataset Editor >

- New Dataset
- Open Dataset
- Export >
- Close Dataset

Open a document

- File > Open File (⌘O)
- File > Examples
- File > Recent Files
- Drag and drop files here





Test data



 No devices connected to the remote management API.

RAW DATA





CREATE IMPULSE (FRUIT-DETECTOR)



An impulse takes raw data, uses signal processing to extract features, and then uses a learning block to classify new data.

Image data



Axes

image

Image width

48

Image height

48

Resize mode

Squash



i For optimal accuracy with transfer learning blocks, use a 96x96 image size.



Image



Name

Image

Input axes

☒ image



Transfer Learning (Images)



Name

Transfer learning

Input features

☐ Image

Output features

4 (apple, banana, pear, unknown)



Output features



4 (apple, banana, pear, unknown)

Save Impulse

Image - Fruit-Detector - Edge

+

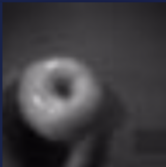
← → ↺ 🏠 🔒 https://studio.edgeimpulse.com/studio/20325/dsp/image/3 ☆ 🛡️ 📧 ABP ⚙️ 👤 ⋮

IMAGE (FRUIT-DETECTOR)

ParametersGenerate features

Raw data

apple.apple-0040.jpg.21rt7oob (apple) ▾



Raw features 📄

0x1f1d20, 0x201e21, 0x201e21, 0x211f22, 0x222023, 0x232124, 0x242...

Parameters


Image

Color depthGrayscale ▾

Save parameters

DSP result

Image



🗨️

Training set

Data in training set 1,142 items

Classes 4 (apple, banana, pear, unknown)

Generate features

Feature generation output

```
Mon Mar 29 16:07:20 2021 Construct embedding
  completed 0 / 500 epochs
  completed 50 / 500 epochs
  completed 100 / 500 epochs
  completed 150 / 500 epochs
  completed 200 / 500 epochs
  completed 250 / 500 epochs
  completed 300 / 500 epochs
  completed 350 / 500 epochs
Still running...
  completed 400 / 500 epochs
  completed 450 / 500 epochs
Mon Mar 29 16:07:25 2021 Finished embedding
Reducing dimensions for visualizations OK
Job completed
```

Feature explorer (1,142 samples)

X Axis

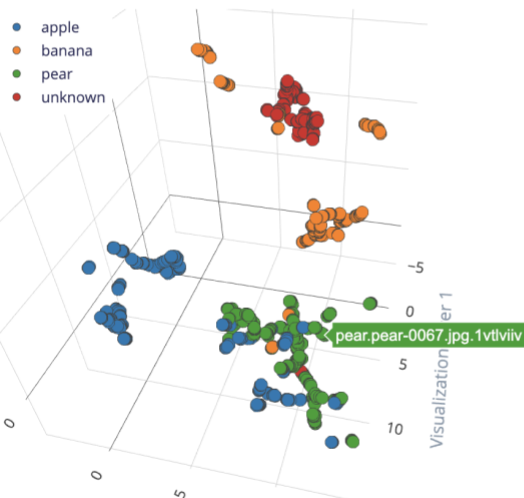
Y Axis

Z Axis

Visualization lay ▾

Visualization lay ▾

Visualization lay ▾

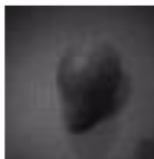


pear.pear-0067.jp...


Label: pear

[View sample](#)

[View features](#)



Input layer (2,304 features)


MobileNetV2 0.35 (final layer: 16 neurons, 0.1 dropout)

Choose a different model

Output layer (4 features)

Start training

Model

Model version: ⓘ Quantized (int8) ▾

Last training performance (validation set)

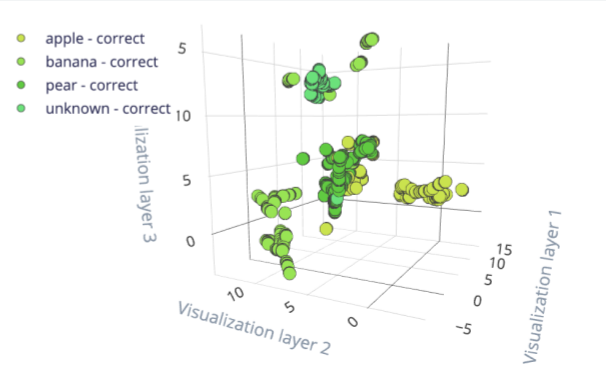
 ACCURACY
100.0%

 LOSS
0.00

Confusion matrix (validation set)

	APPLE	BANANA	PEAR	UNKNO...
APPLE	100%	0%	0%	0%
BANANA	0%	100%	0%	0%
PEAR	0%	0%	100%	0%
UNKNOWN	0%	0%	0%	100%
F1 SCORE	1.00	1.00	1.00	1.00

Feature explorer (full training set) ⓘ



Turn your impulse into optimized source code that you can run on any device.



C++ library



Arduino library



Cube.MX CMSIS-PACK



WebAssembly

Build firmware

Or get a ready-to-go binary for your development board that includes ~~also~~.



Eta Compute ECM3532
AI Vision



Himax WE-I Plus



OpenMV library

A library that runs your
impulse on OpenMV
cameras.

Build



```
confidence = predictions_list[i][1]
label = predictions_list[i][0]
print("%s = %f" % (label[2:], confidence))

if confidence > 0.9 and label != "unknown":
    print("It's a ", label, "!")
```

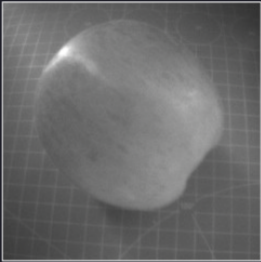
ei_image_classification.py - OpenMV IDE

Line: 33, Col: 30

Frame Buffer

Record Zoom Disable

```
13
14 clock = time.clock()
15 while(True):
16     clock.tick()
17
18     img = sensor.snapshot()
19
20     # default settings just do one detection... change them to search
21     for obj in tf.classify(net, img, min_scale=1.0, scale_mul=0.8, x=
22     print("*****\nPredictions at [x=%d,y=%d,w=%d,h=%d]" % obj
23     img.draw_rectangle(obj.rect())
24     # This combines the labels and confidence values into a list
25     predictions_list = list(zip(labels, obj.output()))
26
27     for i in range(len(predictions_list)):
28         confidence = predictions_list[i][1]
29         label = predictions_list[i][0]
30         print("%s = %f" % (label, confidence))
31
32     if confidence > 0.9 and label != "unknown":
33         print("It's a", label, "!")
34
35     print(clock.fps(), "fps")
36
```



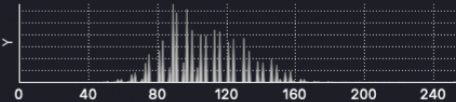
Serial Terminal

```
Predictions at [x=0,y=0,w=240,h=240]
apple = 1.000000
It's a apple !
banana = 0.000000
pear = 0.000000
unknown = 0.000000
5.63686 fps
```

Histogram

Grayscale Color Space

Res (w:240,h:240)



Mean	Median	Mode	StDev
110	105	89	30

Min	Max	LQ	UQ
40	255	89	124

Search Results

Serial Terminal

Board: H7

Sensor: HM01B0

Firmware Version: 3.9.4 - [latest]

Serial Port: cu.usbmodem364D346431391

Drive: /Volumes/PORTENTA

FPS: 5.6

<https://www.arduino.cc/pro/tutorials/portenta-h7/vs-openmv-ml>

https://bit.ly/portenta_person_detection

DOCUMENTATION

Getting Started
API and SDK references
What is embedded ML, anyway?

DEVELOPMENT BOARDS

Overview
ST B-L475E-IOT01A
Arduino Nano 33 BLE Sense
Eta Compute ECM3532 AI Sensor
Eta Compute ECM3532 AI Vision
OpenMV Cam H7 Plus
Himax WE-I Plus
Nordic Semi nRF52840 DK
Nordic Semi nRF5340 DK
SiLabs Thunderboard Sense 2
Arduino Portenta H7 + Vision shield (preview)
Raspberry Pi 4
NVIDIA Jetson Nano
Mobile phone
Porting guide

EDGE IMPULSE FOR LINUX

Edge Impulse for Linux
Node.js SDK
Go SDK

Responding to your voice

In this tutorial, you'll use machine learning to build a system that can recognize audible events, particularly your voice through *audio classification*. The system you create will work similarly to "Hey Siri" or "OK, Google" and is able to recognize keywords or other audible events, even in the presence of other background noise or background chatter.

You'll learn how to collect audio data from microphones, use signal processing to extract the most important information, and train a deep neural network that can tell you whether your keyword was heard in a given clip of audio. Finally, you'll deploy the system to an embedded device and evaluate how well it works.

At the end of this tutorial, you'll have a firm understanding of how to classify audio using Edge Impulse.

There is also a video version of this tutorial:

Build Your Own ML-powered keyword spotting model in 30 KB RAM

Watch later Share

Configure model

Last training performance

	helloworld	noise	UNKNOWN
ACCURACY	96	2	8
LOSS	0	103	8

Available optimization types: Unoptimized (float32)

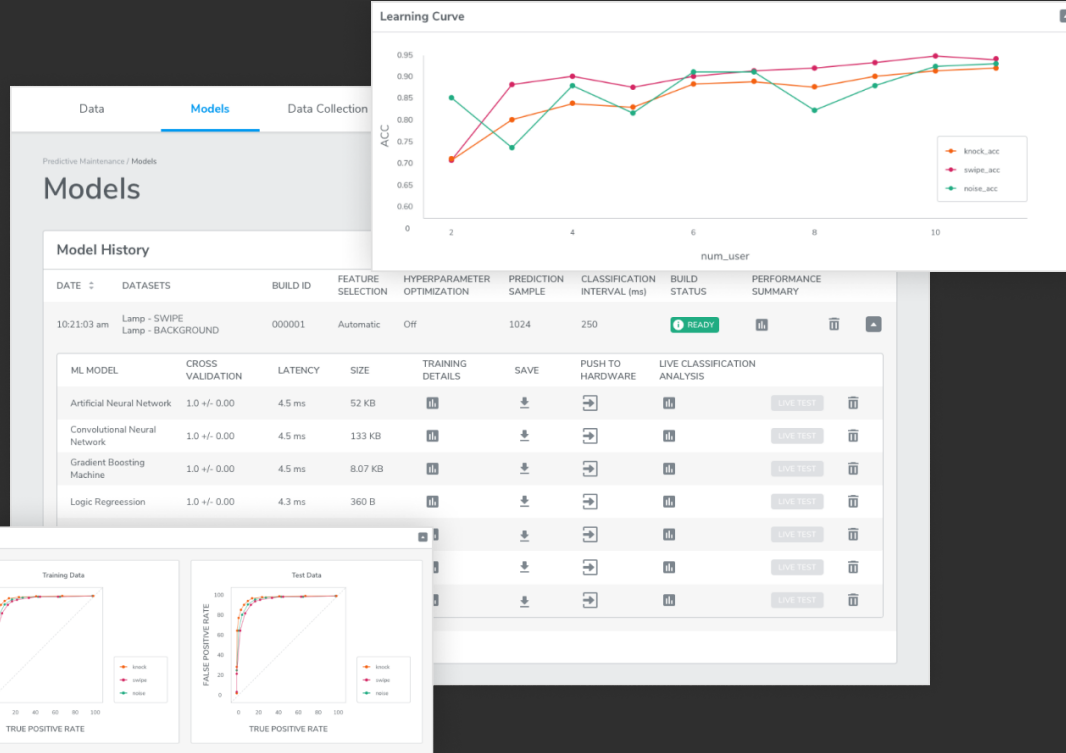
On-device performance

INFERRING TIME: 18 ms

PEAK RAM USAGE: 8.8K

Model performing poorly?

Keyword Spotting / <https://docs.edgeimpulse.com/docs/responding-to-your-voice>





That's a wrap,
Thank you!

Massimo Banzi

askmassimo@arduino.cc



Copyright Notice

This presentation in this publication was presented as a tinyML® Talks webcast. The content reflects the opinion of the author(s) and their respective companies. The inclusion of presentations in this publication does not constitute an endorsement by tinyML Foundation or the sponsors.

There is no copyright protection claimed by this publication. However, each presentation is the work of the authors and their respective companies and may contain copyrighted material. As such, it is strongly encouraged that any use reflect proper acknowledgement to the appropriate source. Any questions regarding the use of any materials presented should be directed to the author(s) or their companies.

tinyML is a registered trademark of the tinyML Foundation.

www.tinyML.org