Machine Learning for Embedded Software Engineers

Class 5: Near Real-time Machine Learning using Coral

April 26, 2019

Jacob Beningo

DesignNews

CEC CONTINUING EDUCATION CENTER

Presented by: Digi-Key ELECTRONICS

Course Overview

Topics:

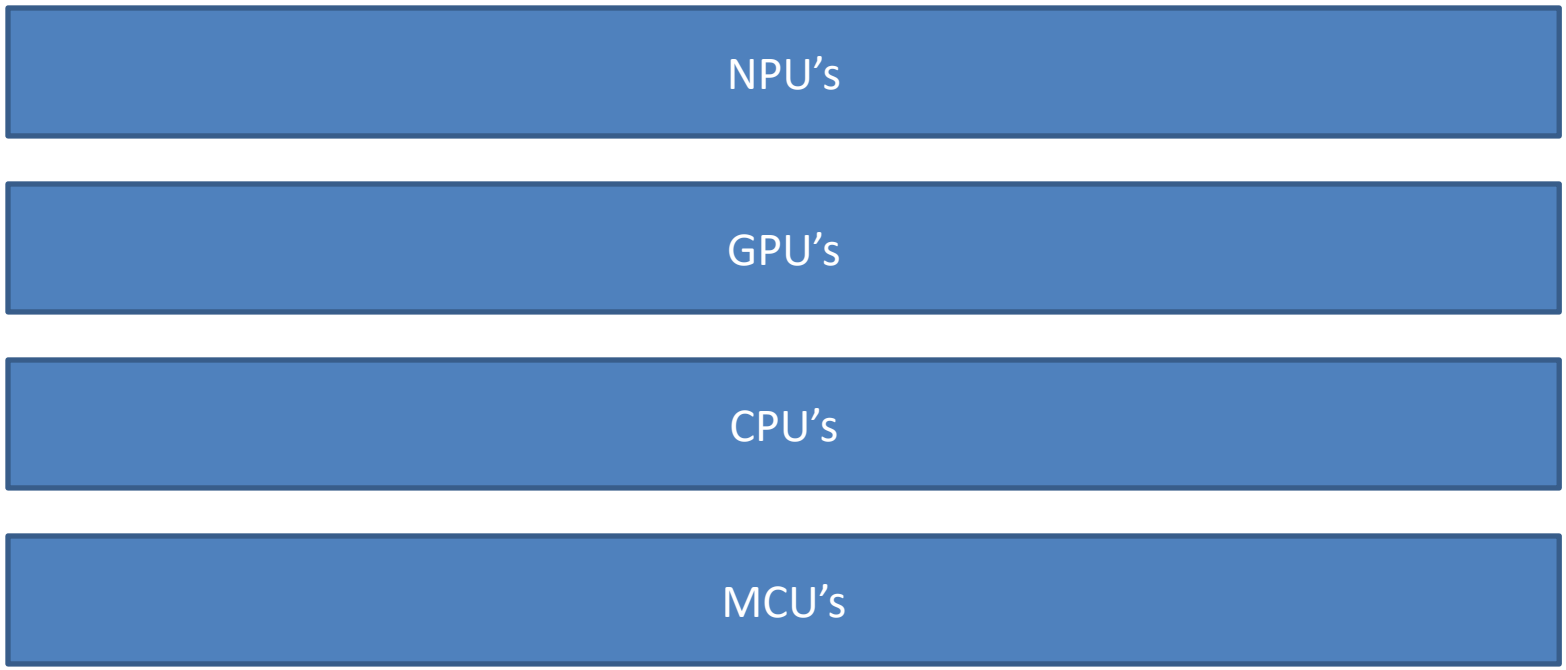
- Introduction to Machine Learning
- Machine Learning Architectures for Embedded Systems
- Machine Learning Applications: Vision and Speech
- Machine Vision with OpenMV
- **Near Real-time Machine Learning using Coral**

Session Overview

- ML Processor Options
- Google Coral
- Embedded ML Examples
- Where to go from here

Presented by:

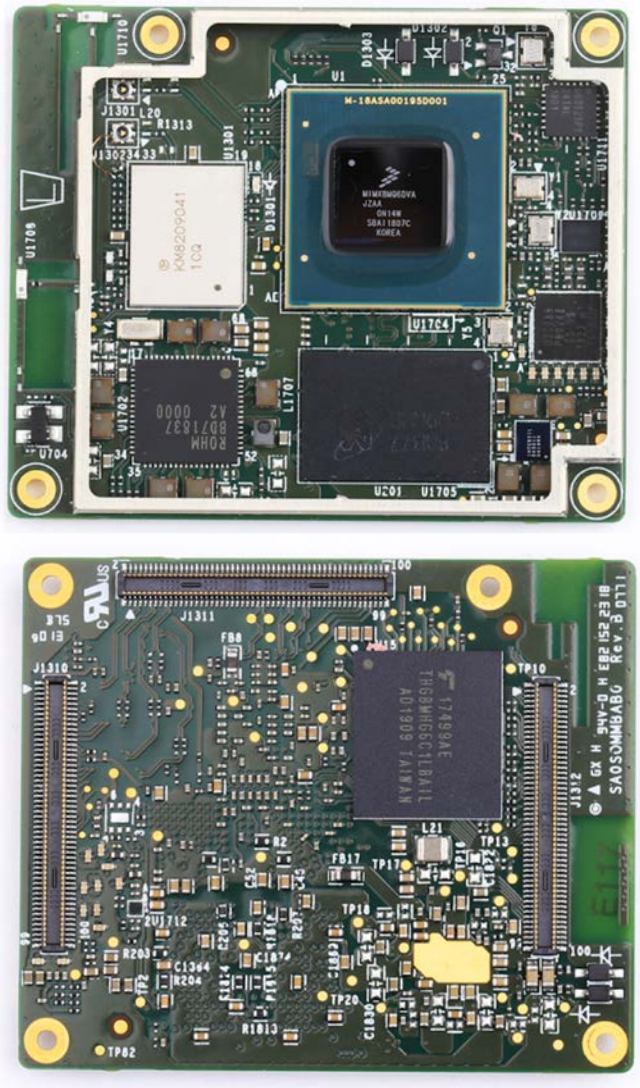
ML Processor Options



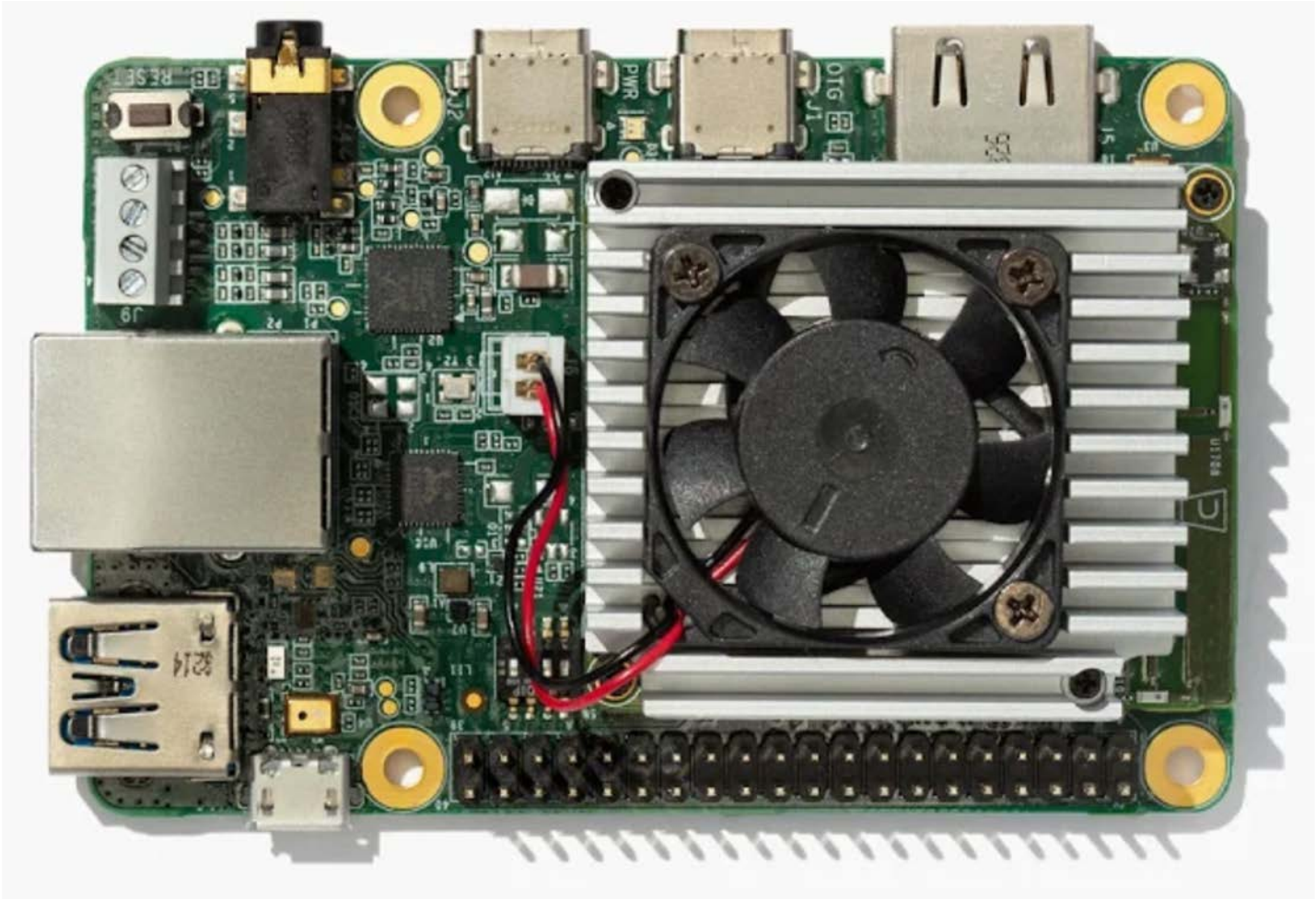
Google Coral TPU SOM

CPU	NXP i.MX 8M SOC (quad Cortex-A53, Cortex-M4F)
GPU	Integrated GC7000 Lite Graphics
ML accelerator	Google Edge TPU coprocessor
RAM	1 GB LPDDR4
Flash memory	8 GB eMMC
Wireless	Wi-Fi 2x2 MIMO (802.11b/g/n/ac 2.4/5GHz) Bluetooth 4.1
Dimensions	48mm x 40mm x 5mm

Source: Google



Google Coral Dev Board



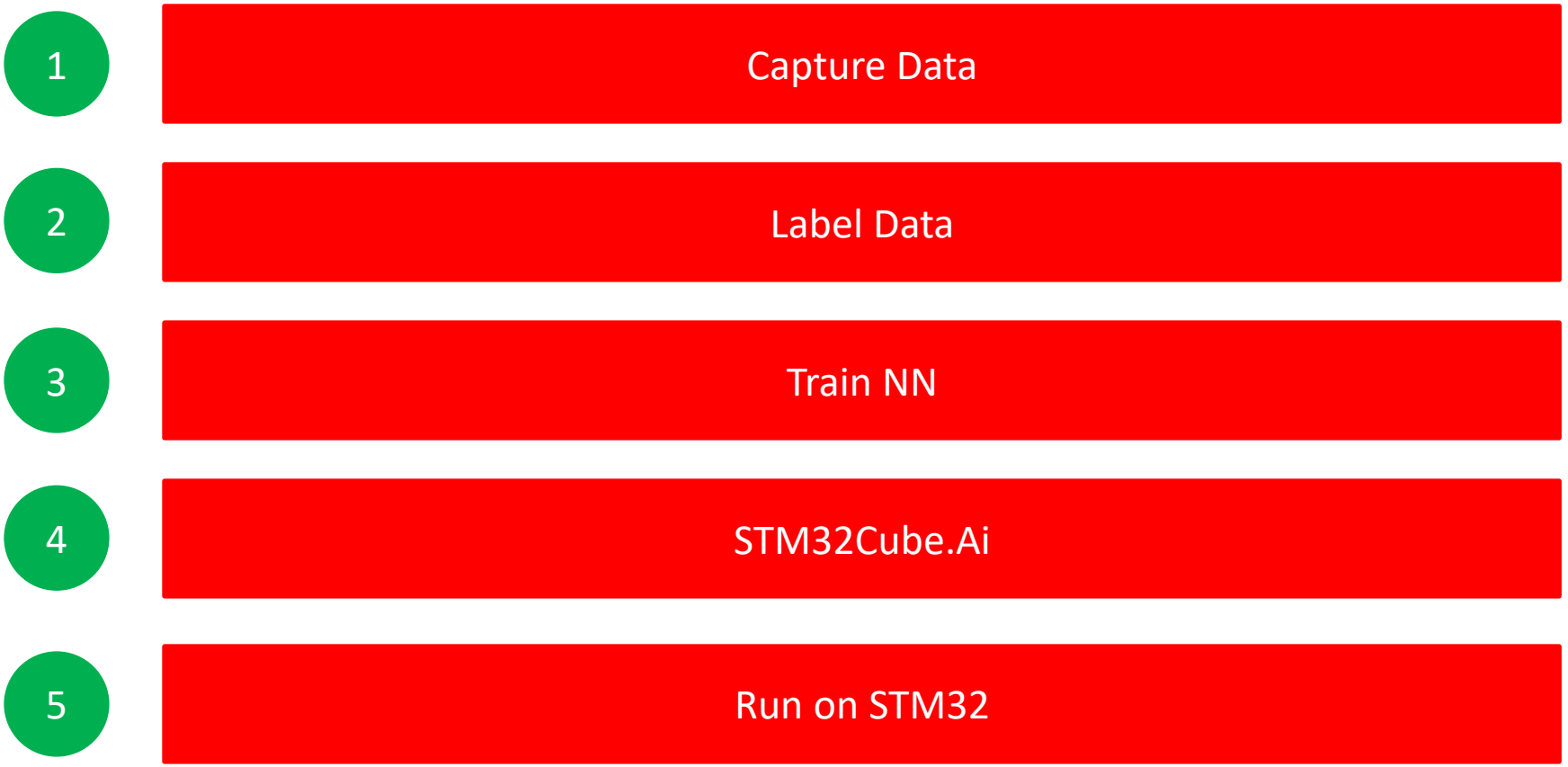
Source: Google

Google Coral Dev Board

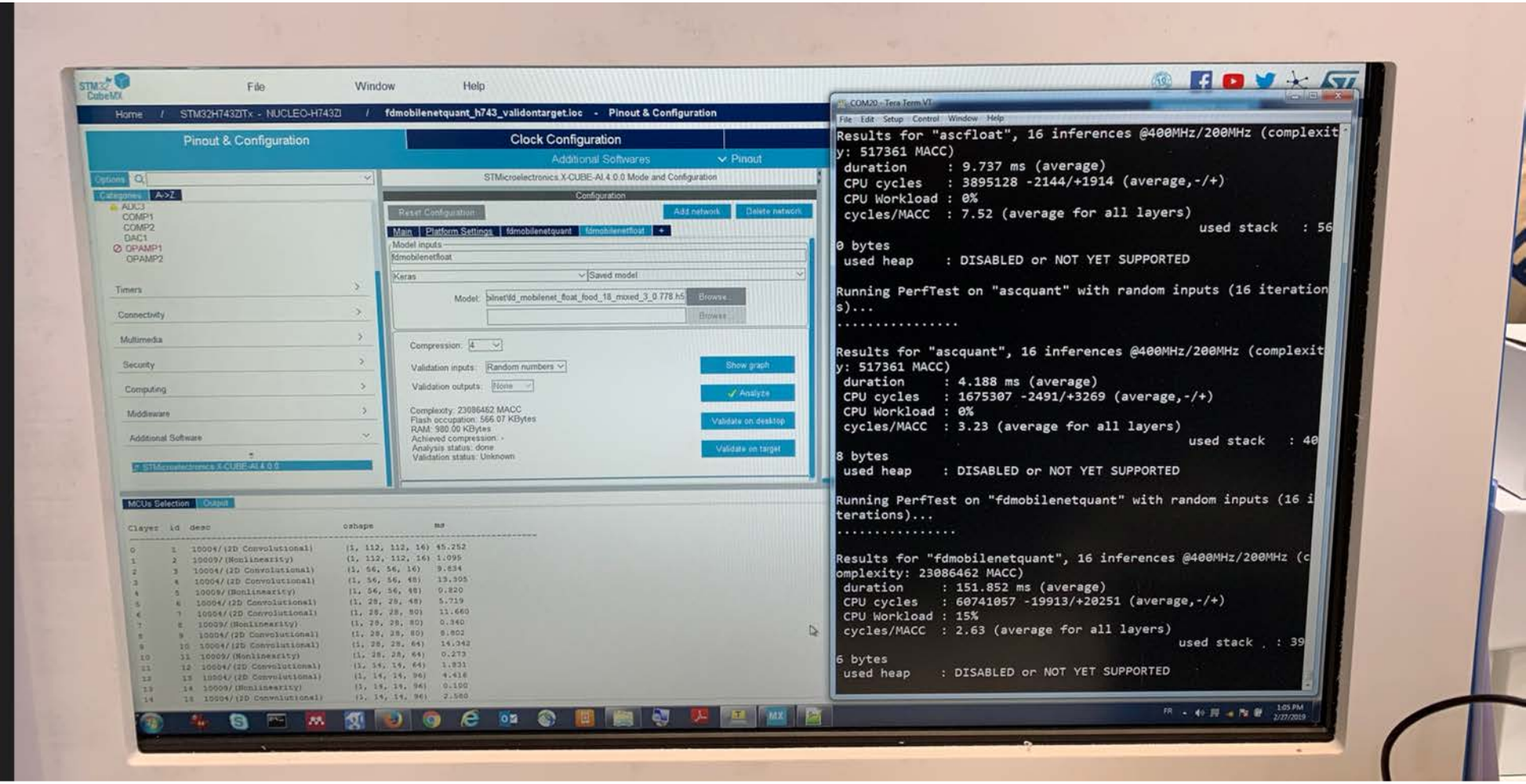
Flash memory	MicroSD slot
USB	Type-C OTG Type-C power Type-A 3.0 host Micro-B serial console
LAN	Gigabit Ethernet port
Audio	3.5mm audio jack (CTIA compliant) Digital PDM microphone (x2) 2.54mm 4-pin terminal for stereo speakers
Video	HDMI 2.0a (full size) 39-pin FFC connector for MIPI-DSI display (4-lane) 24-pin FFC connector for MIPI-CSI2 camera (4-lane)
GPIO	3.3V power rail 40 - 255 ohms programmable impedance ~82 mA max current
Power	5V DC (USB Type-C)
Dimensions	88 mm x 60 mm x 24mm

Source: Google

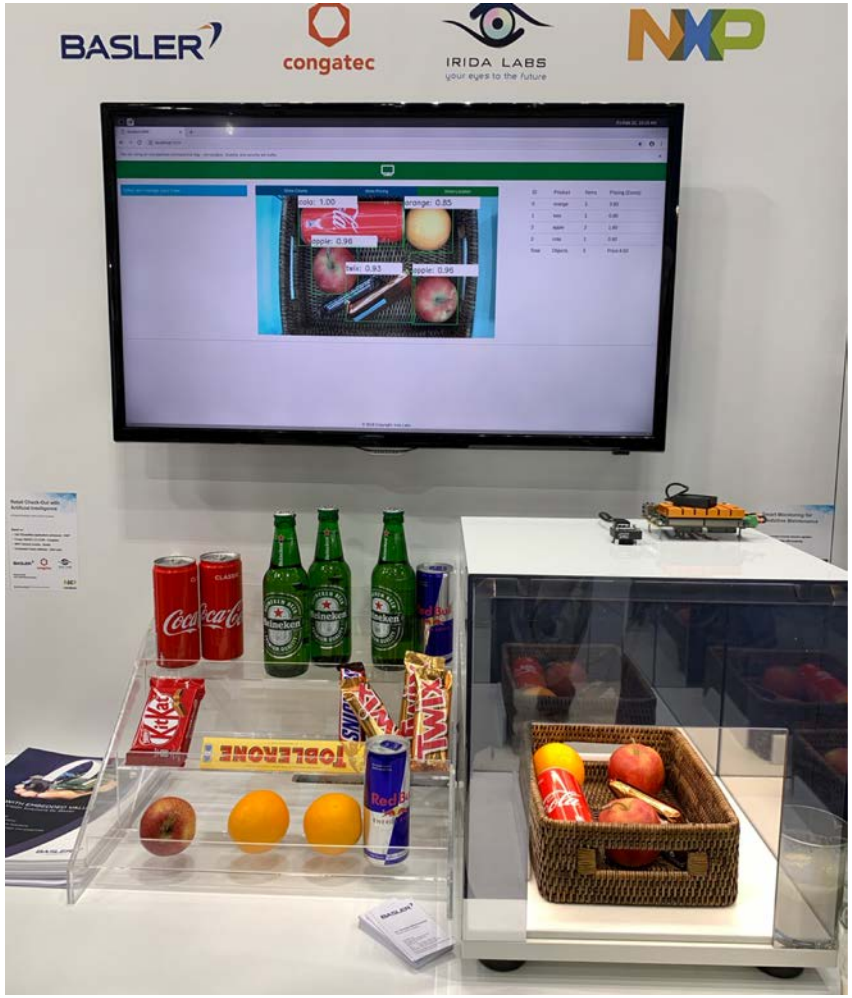
Embedded ML Examples – STM32



Embedded ML Examples – STM32



Embedded ML Examples - NXP



A few key take-a-ways

- Machine learning can be applied to nearly any embedded system
- Machine learning at the edge will be here faster than you think but slower than marketing is pushing
- Using a DSP capable processor can dramatically speed-up inference execution on MCU's

A few key take-a-ways

- There is a trade-off, balance between inference accuracy and memory/execution speed
- Testing and validation may be challenging in the foreseeable future
- There are many ML frameworks. Choose the one that best fits your needs
- Experiment with different frameworks and processors to get a good feel for ML capabilities

Where to go from here

Learning Neural Network Resources

- Videos
  - https://www.youtube.com/watch?v=aircAruvnKk
  - https://www.youtube.com/watch?v=IHZwWFHWa-w
  - https://www.youtube.com/watch?v=Ilg3gGewQ5U
  - https://www.youtube.com/watch?v=tIeHLnjs5U8
- Books
  - http://neuralnetworksanddeeplearning.com/
  - Deep Learning
  - Python Machine Learning by Example

Where to go from here

Experimentation

- Develop a Python ML example that analyzes hand writing digits
- Walk through caffe training and inference tutorial:
  - http://adilmoujahid.com/posts/2016/06/introduction-deep-learning-python-caffe/
- Apply a ML model to OpenMV module
- Setup and train an image classifier using Coral
  - Uses TensorFlow Lite for training

Where to go from here

Get Application Specific

- OpenMV
- STM32
- NXP Alexa Integration
- Keyword Spotting
- Object detection and classification
- Explore existing datasets

Additional Resources

- Download Course Material for
  - C/C++ Doxygen Templates
  - Example source code
  - Blog
  - YouTube Videos
- Embedded Bytes Newsletter
  - http://bit.ly/1BAHYXm

From www.beningo.com under

- Blog > CEC – Machine Learning for Embedded Software Engineers