

IoT and Edge Computing: Future directions for Europe WORKSHOP

11 Sept 2020
10:00AM CEST



Alliance for
Internet of Things
Innovation



Csaba Kiraly
IoT Engineer
Digital
Catapult

Interactive Session on Enabling Technologies

AIoT and TinyML: The low-power compute revolution at the very edge.

Csaba Kiraly
Digital Catapult, UK

The Thing to Cloud Continuum

The four domains of compute

Cloud servers

- AI “as you know it”
- Big data
- ML Training
- APIs / services

Edge servers

Still servers, just closer to the edge

- MEC, etc.

Embedded computers at the edge

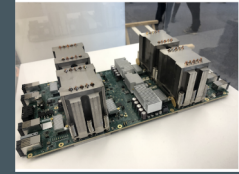
Not servers, just powerful nodes

- Smartphones (AI “as you use it”)
- IoT Gateways
- Embedded SBCs

Things at the very edge

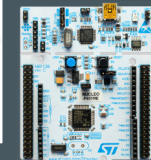
Low cost, low power, high volume, MCU driven
Often energy-constrained

CATAPULT
Digital



5G Small Cell Edge

- Expandable add-on FHFL PCIe modules for 5G
- Wider operating temperature from -5°C to 55°C



TinyML is changing the AIoT game

TinyML: ultra-low power (<mW) machine learning technology enabling battery-operated on-device analytics **at the very edge**

Technology Trends

Fast progress on enabling technologies

- **Model compression** (quantization, pruning, etc.) to <250 KBytes
- **HW** architectures
 - Near-memory, in-memory, on-sensor AI
- **Software** frameworks and toolsets

Powerful data-driven programming for embedded devices

- Enabling new **use cases**
- Enabling wider **developer ecosystem**

Emerging Use Cases



Task Category	Use Case	Model Type
Audio	Audio Wake Words Context Recognition Control Words Keyword Detection	DNN CNN RNN LSTM
Image	Visual Wake Words Object Detection Gesture Recognition Object Counting Text Recognition	DNN CNN SVM Decision Tree KNN Linear
Industry / Telemetry	Segmentation Anomaly Detection Forecasting Activity Detection	DNN Decision Tree SVM Linear

The future of compute is Tiny

MCU based Things are everywhere

- >250 billion microcontrollers in the world today

MCUs and SoCs are getting cheaper

- 0.50\$ Average MCU Selling Price in 2020

No good data should be left behind

- 5 Quintillion bytes of data produced every day by IoT

Source: Cisco

MCU compute is now cheaper than comms

- 32x32 Image classification @ 0.15 mJ per inference
- LPWAN image transfer @ 10-100 mJ per image

Early adopters are here (representative examples)

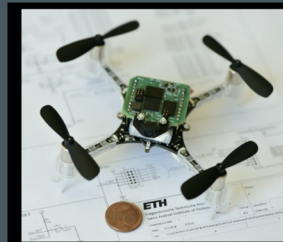
CATAPULT
Digital

PIR and People
Counting
© GreenWaves

IR people detection

- 80 x 80 IR Image - LynRED ThermEye
- Image preprocessing + human detection
- 62ms ~4.4mW / frame / second
- 99% accuracy on internally collected training set.
- A full solution for people counting / occupancy detection on a battery for > 5 years

Autonomous Drones
© ETH Zürich



PULP-based nano-UAV
Weight: 27+5 grams
Deep Learning-powered
On-board CNN:
- 6 fps @ 64 mW
- up to 18 fps @ 272 mW



We present a nano-sized, ultra-low power, autonomous UAV that can navigate through indoor and outdoor environments with on-board sensing and computing.

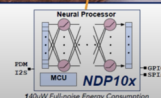
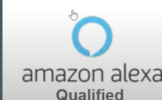
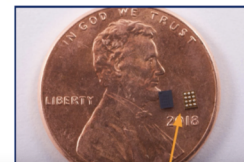
Copyright 2019 © ETH Zürich

Always-on Voice
Commands
© SYNTIANT

SYNTIANT® NDP100

Always-on Speech For Any Battery-powered Device

- Best Performance
- Negligible Battery Life Impact
- Quickest Time-To-Market
- Smallest Footprint
- Most Cost Effective



SYNTIANT

SYNTIANT® - COPYRIGHT 2020

Research and Innovation priorities - a look ahead

What we do to encourage innovation

- Performance **benchmarking**: **tinyMLPerf**
 - Encourage innovation to improve the state-of-the-art low-power ML
 - Accelerate progress in TinyML via fair and useful measurement
- **Accelerator programmes** for start-ups and scale-ups

The big picture

- Enable data-driven innovation in previously untapped domains
- Research directions
 - Sustainable energy-neutral deploy-and-forget devices
 - Thing-to-cloud orchestration of compute and communication resources
 - From inference to low-power learning at the very edge



Thank you!

Csaba Kiraly

csaba.kiraly@digicatapult.org.uk