tinyML® EMEA

Enabling Ultra-low Power Machine Learning at the Edge

tinyML EMEA Technical Forum 2021 Proceedings

June 7 – 10, 2021

Virtual Event

www.tinyML.org

tinyML EMEA Technical Forum 2021
June 7-10, 2021

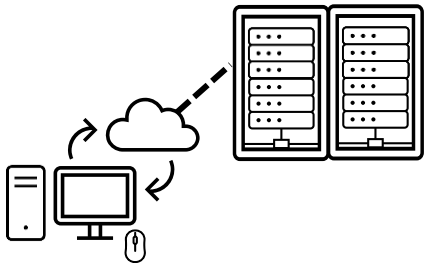
Low-Power License Plate Detection and Recognition on a RISC-V Multi-Core MCU-based Vision System

Lorenzo Lamberti , PhD student at University of Bologna, Italy

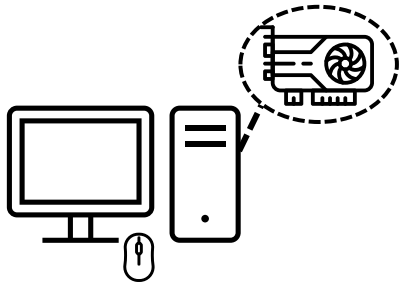
June 10, 2021

Deep Learning on Tiny Devices

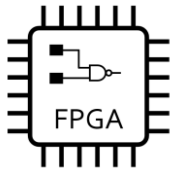
Deep Learning visual tasks are usually too computational intensive to be implemented on IoT devices



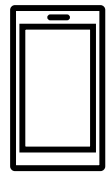
HPC servers ✓



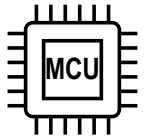
GPU ✓



FPGA ✓



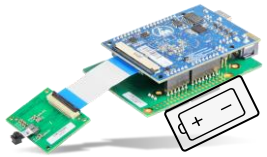
Mobile ✓



MCU ✗

<u>MCUs are the ideal IoT platform</u>:

- Low-power: IoT devices are battery-operated.
- Low-cost
- Highly-flexible: SW programmable.



<u>But, MCUs present severe limitations:</u>

- **Memory** is limited to few MB. (DL models ≳100MB)
- **Computational power** is limited:
  - Single-core Low clock frequency (≲500MHz)
  - Low inference throughput

Focus of our work: overcome these challenges.

Automatic License Plate Recognition (ALPR)

Deep Learning task chosen: Automatic License Plate recognition

ALPR challange: predicting all the license plate characters.

Why ALPR? No evidences of ALPR on low-power MCUs yet!

prediction

京QC8152

We present a HW/SW co-design flow that enable the deployment of intensive DL workloads on a 100mW budget:

1. Determination of HW and SW building blocks → **2. Optimization of the SW pipeline** → **3. In-field Testing and SoA comparison**

Hardware:

Software:

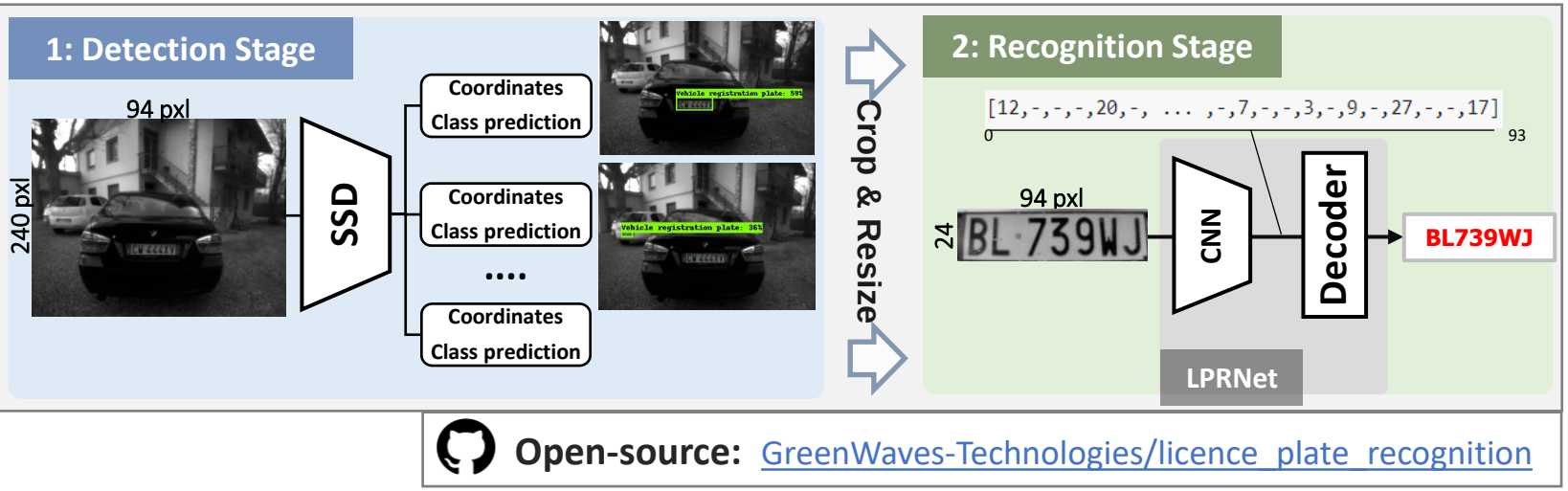
1. Detection CNN
2. Recognition CNN

- **Hand-tuned NN topology**
- **Model Compression:** 8-bit
- **GAP*flow:*** Efficient automated deployment on GAP8

We prove the accuracy of the system:
- on public datasets + in-field testing

We prove the energy-efficiency of our system
- Comparing the SoA of ALPR

Conclusions



Open-source: [GreenWaves-Technologies/licence_plate_recognition]

NEW SoA:

We achieve the most energy efficient MCU device for ALPR in literature with 117mW

73x less power consumption compared with previous SoA system [1] (**3.12W**)!

- **1.1FPS** inference @175MHz performing **687M MAC**
- **4.1MB** footprint (8-bit quantization applied)
- **Accuracy:** 39% mAP for LP detection, >99.13% for char recognition
- **Max recognition distance**: 4m for detection, 1m for recognition

[1] D. M. F. Izidio et al., "An embedded automatic license plate recognition system using deep learning," in 2018 VIII Brazilian Symposium on Computing Systems Engineering (SBESC), 2018, pp. 38–45.

Thank you for your attention !

Lorenzo Lamberti

Ph.D. student at University of Bologna, Italy
lorenzo.lamberti@unibo.it

TINY ML EMEA
Virtual
JUNE 7-10,2021

Premier Sponsor

Intelligent Agent
Neuton

Automated TinyML

Zero-code SaaS solution

Create tiny models, ready for embedding, in just a few clicks!

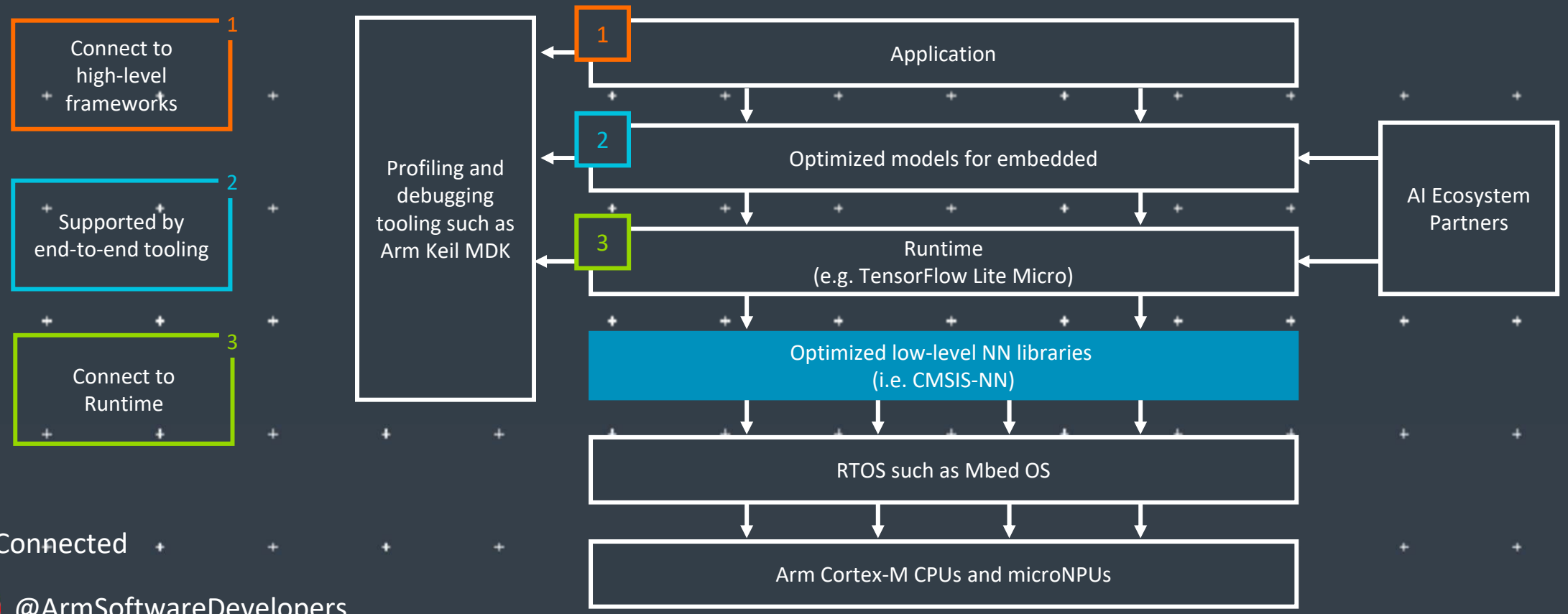
Compare the benchmarks of our compact models to those of TensorFlow and other leading neural network frameworks.

Build Fast. Build Once. Never Compromise.

TINY ML EMEA
Virtual
JUNE 7-10,2021

Executive Sponsors

Arm: The Software and Hardware Foundation for tinyML



Stay Connected

@ArmSoftwareDevelopers

@ArmSoftwareDev

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

arm

TinyML for all developers

C++ library

Arduino library

WebAssembly

Dataset

Acquire valuable training data securely

Enrich data and train ML algorithms

Edge Device

Real sensors in real time

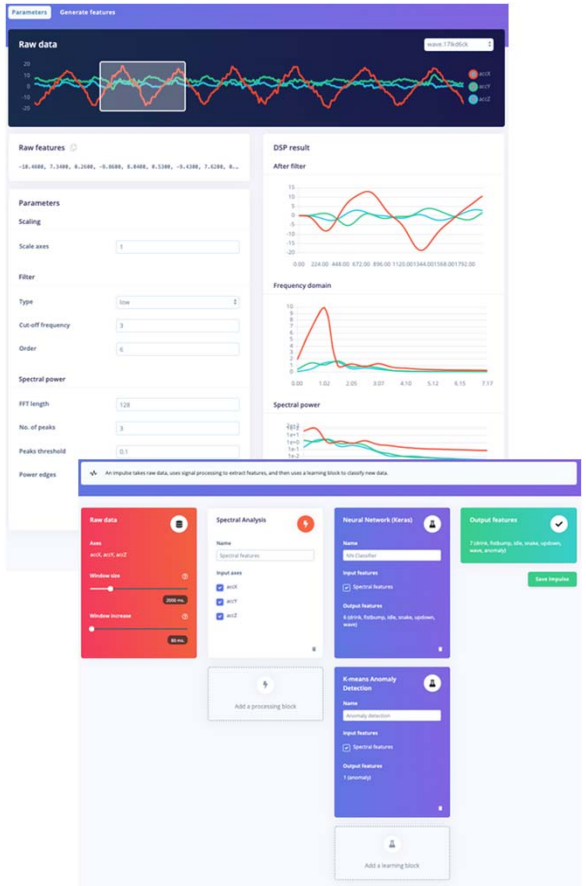
Open source SDK

Impulse

Embedded and edge compute deployment options

Test impulse with real-time device data flows

Test



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

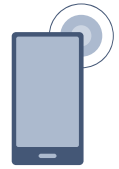
Reinforcement learning for decision making

Edge cloud

Cloud

IoT/IIoT

Automotive



Mobile

SYNTIANT

Syntiant Corp. is moving artificial intelligence and machine learning from the cloud to edge devices. Syntiant's chip solutions merge deep learning with semiconductor design to produce ultra-low-power, high performance, deep neural network processors. These network processors enable always-on applications in battery-powered devices, such as smartphones, smart speakers, earbuds, hearing aids, and laptops. Syntiant's Neural Decision Processors™ offer wake word, command word, and event detection in a chip for always-on voice and sensor applications.

Founded in 2017 and headquartered in Irvine, California, the company is backed by Amazon, Applied Materials, Atlantic Bridge Capital, Bosch, Intel Capital, Microsoft, Motorola, and others. Syntiant was recently named a CES® 2021 Best of Innovation Awards Honoree, shipped over 10M units worldwide, and unveiled the NDP120 part of the NDP10x family of inference engines for low-power applications.

www.syntiant.com

@Syntiantcorp

TINY ML EMEA
Virtual
JUNE 7-10,2021

Platinum Sponsors

Part of your life. Part of tomorrow.

www.infineon.com

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

TINY ML EMEA
Virtual
JUNE 7-10,2021

Gold Sponsors

LatentAI

Adaptive AI for the Intelligent Edge

Latentai.com

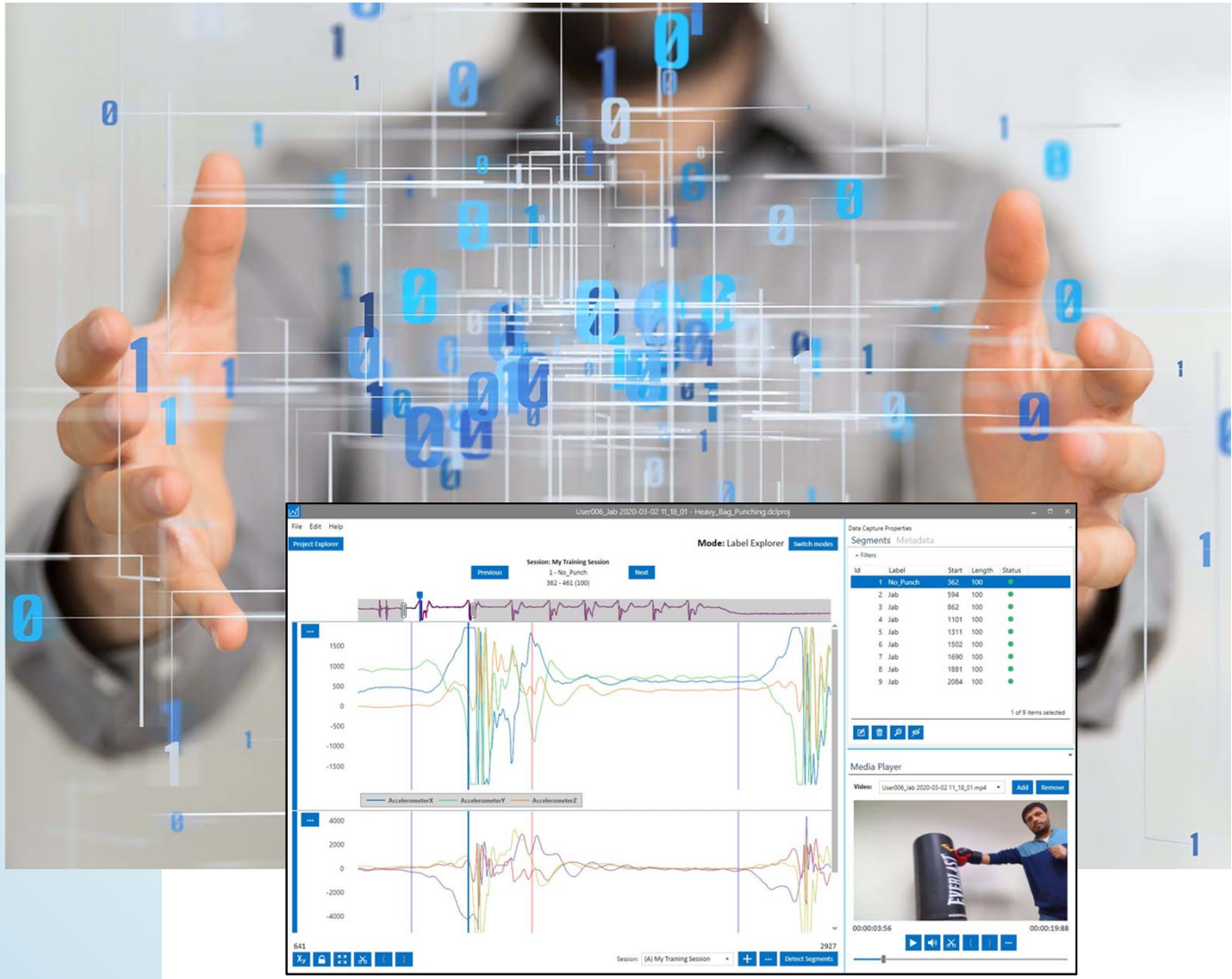
SensiML™

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

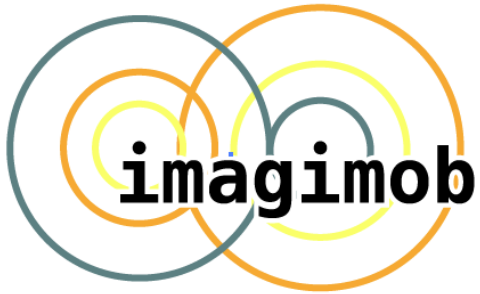
We enable the creation of production-grade smart sensor devices.



sensiml.com

TINY ML EMEA Virtual JUNE 7-10,2021

Silver Sponsors



Copyright Notice