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.

**MCUs are the ideal IoT platform:**
- Low-power: IoT devices are battery-operated.
- Low-cost
- Highly-flexible: SW programmable.

**But, MCUs present severe limitations:**
- **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.**
Deep Learning task chosen: Automatic License Plate recognition

ALPR challenge: predicting all the license plate characters.

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

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:
1. Detection CNN
2. Recognition CNN

Software:
- Hand-tuned NN topology
- Model Compression: 8-bit
- GAPflow: 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**

NEW SoA:
We achieve the most energy efficient MCU device for ALPR in literature with **117mW**

- **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

**Open-source**: GreenWaves-Technologies/licence_plate_recognition

Thank you for your attention!

Lorenzo Lamberti
Ph.D. student at University of Bologna, Italy
lorenzo.lamberti@unibo.it
Premier Sponsor
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.

Executive Sponsors
Arm: The Software and Hardware Foundation for tinyML

- Optimized models for embedded
- Optimized low-level NN libraries (i.e. CMSIS-NN)
- RTOS such as Mbed OS
- Arm Cortex-M CPUs and microNPUs
- Application
- Profiling and debugging tooling such as Arm Keil MDK
- Supported by end-to-end tooling
- Connect to high-level frameworks
- Connect to Runtime

Stay Connected
- @ArmSoftwareDevelopers
- @ArmSoftwareDev

Resources: developer.arm.com/solutions/machine-learning-on-arm
TinyML for all developers

Acquire valuable training data securely

Edge Device
Real sensors in real time
Open source SDK

Embedded and edge compute deployment options

Test impulse with real-time device data flows

Enrich data and train ML algorithms

Dataset

Test

www.edgeimpulse.com
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

Perception
- Object detection, speech recognition, contextual fusion

Reasoning
- Scene understanding, language understanding, behavior prediction

Action
- Reinforcement learning for decision making

A platform to scale AI across the industry
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
Platinum Sponsors
Part of your life. Part of tomorrow.

www.infineon.com
Add Advanced Sensing to your Product with Edge AI / TinyML

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

https://reality.ai  info@reality.ai  @SensorAI  Reality AI
Gold Sponsors
Adaptive AI for the Intelligent Edge

LatentAI.com
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
Silver Sponsors
Copyright Notice

The presentation(s) in this publication comprise the proceedings of tinyML® EMEA Technical Forum 2021. The content reflects the opinion of the authors and their respective companies. This version of the presentation may differ from the version that was presented at tinyML EMEA. 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