tinyML. EMEA

Enabling Ultra-low Power Machine Learning at the Edge

tinyML EMEA Technical Forum 2021 Proceedings

June 7 – 10, 2021 Virtual Event







How to implement Distributed Deep Edge Al with STMicroelectronics?

STMicroelectronics

Agenda

- 1 ST strategy for TinyML
- The Context Awareness use case & function pack

- 3 Human Activity Recognition in Smart Sensor LSM6DSOX
- 4 Acoustic Scene Classification on STM32L4R9ZIJ6 MCU
- 5 Context Awareness Demo



ST strategy for TinyML



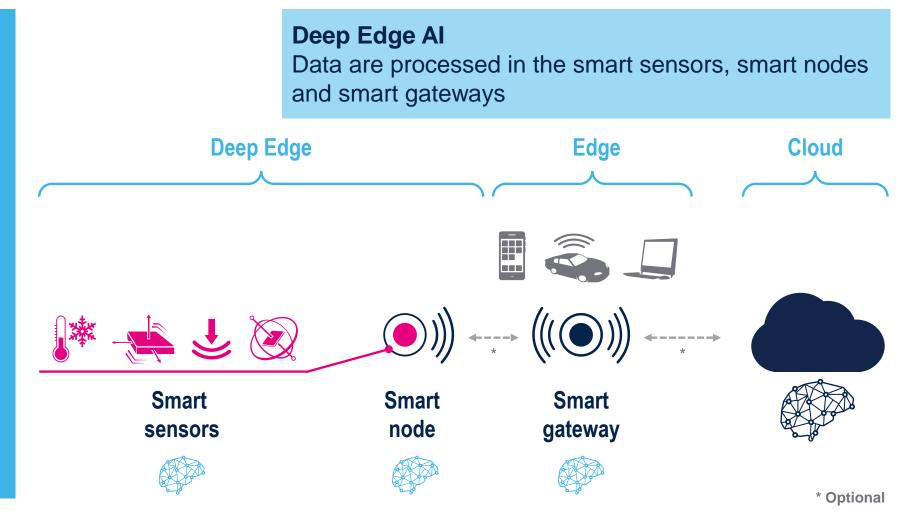
Moving closer to data source

Pros

- Responsiveness (fast)
- Bandwidth (low)
- Improved privacy (data stored locally)
- Energy saving

Cons

 Distribute the processing over multiple devices





Al in endpoints and gateways

Al complexity



Edge gateway



Sensor nodes Endpoints

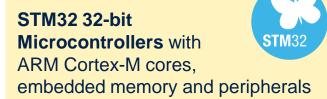


Smart sensors



Solutions

STM32 32-bit
Microprocessors with
ARM® Cortex® A+M cores,
embedded memory and peripherals



Advanced sensors
that can be used alone or
in combination with an MCU/MPU to
further improve power savings



Tools



TensorFlowLite & X-Linux-Al + STM32Cube.Al







STM32Cube.AI tool
to automatically optimize
Artificial Neural Networks
for STM32 MCUs and MPUs





Current generation of
Smart Sensors
with MLC and FSM
to enabling ML processing





Context awareness use case & function pack





Context awareness: where are we? what are we doing?

FP-AI-CXTAWARE1: the best power system saving solution

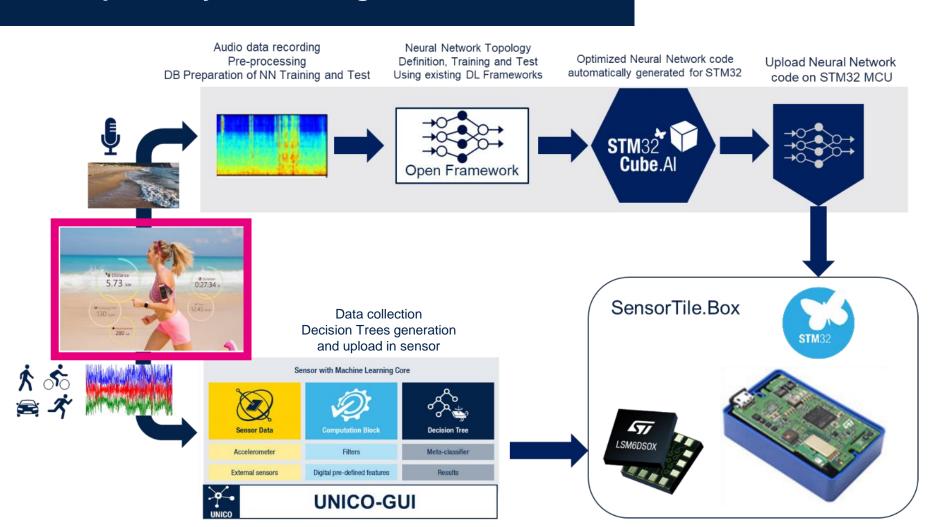


Sensor Human **Activity** Recognition









Context awareness activity recognition: What do you need?

STEVAL-MKSBOX1V1 evaluation board

Multi-sensor kit with portable sensor box and smart sensor app

- Compact casing, IP54-compliant 57 x 38 x 20 mm (L x I x h) Additional cases (with flanges or hinges) available
- 500 mA-h Li-Po battery
- 8 GB µSD card as mass memory extendable to 64 GB
- Compatible with "ST BLE Sensor" app Available on Google Play and App Store





Where can you find the SensorTile.box (STEVAL-MKSBOX1V1)?

Evaluation kit available in ST eStore and Distributors:

https://estore.st.com/en/steval-mksbox1v1-cpn.html



STEVAL-MKSBOX1V1 ACTIVE

SensorTile.box wireless multi sensor development kit with user friendly app for IoT and wearable sensor applications

Learn More about STEVAL-MKSBOX1V1

| Quantity | \$ per Unit | Savings |
|----------|---------------|---------|
| 1 - 4 | \$42.88 | 0% |
| 5 - 500 | \$42.87 | 0% |
| 500 + | Contact Sales | |

Contents:

- STEVAL-MKSBOX1V1 with 2 different cases
- SD-Card and rechargeable battery
- JTAG20 to STDC14 adapter
- STDC14 cable





What's inside the SensorTile.box (STEVAL-MKSBOX1V1)?

Motion Sensors



Low-power 6-axis IMU, embedding Machine Learning LSM6DSOX

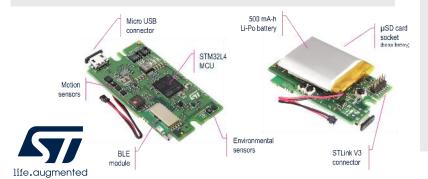


High-performance and low power accelerometers

LIS3DHH & LIS2DW12



Magnetometer LIS2MDL



Environmental Sensors



Altimeter / Pressure sensor LPS22HH



Accurate temperature sensor **STTS751**



Humidity sensor HTS221



Analog wide-band microphone MP23ABS1

Processing



STM32L4 low-power MCU STM32L4R9ZIJ6

Connectivity



Bluetooth Low Energy Module
SPBTLE-1S

Power management



Battery charger **STBC02**

Pre-integrated application example

The SensorTile.box is supported by several STM32Cube function packs



FP-SNS-STBOX1



For building custom applications using the the Pro Mode

FP-ATR-BLE1



For asset tracking using BLE connectivity

FP-SNS-ALLMEMS



Ultra-low power IoT node with BLE connectivity, digital microphone, environmental and motion sensors

FP-AI-SENSING1



Ultra-low power IoT node with artificial intelligence (AI) application based on audio and motion sensing

FP-AI-CXTAWARE1



Ultra-low power context awareness with distributed artificial intelligence (AI): acoustic analysis with NN on MCU and motion analysis with ML on IMU



What's inside the FP-AI-CTXAWARE1?

Complete firmware

Context awareness node with BLE connectivity, digital microphone, environmental and motion sensors, and perform real-time monitoring of sensors and audio data

Middleware library

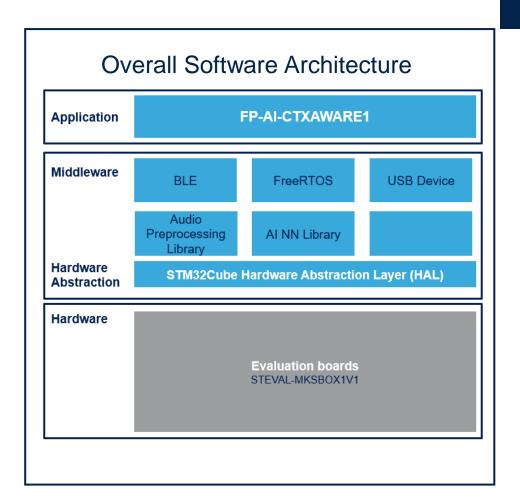
generated thanks to STM32CubeMX extension X-CUBE-AI, featuring example implementation of neural networks for real time acoustic scene classification (ASC) application

Concurrent execution ML & NN

with concurrent execution of the MLC for HAR and neural network for ASC

Features

Ultra-low power implementation based on the use of an RTOS Compatible with ST BLE Sensor application for Android/iOS Free, user-friendly license terms



Latest info available at www.st.com

FP-AI-CTXAWARE1

Setup & application examples Hardware prerequisites

- 1x STEVAL-MKSBOX1V1 evaluation board
- Laptop/PC with Windows 7, 8 or 10
- 1 x microUSB cable
- 1x USB type A to Mini-B USB cable
- ST-Link/V2 in-circuit debugger/programmer for STM8 and STM32

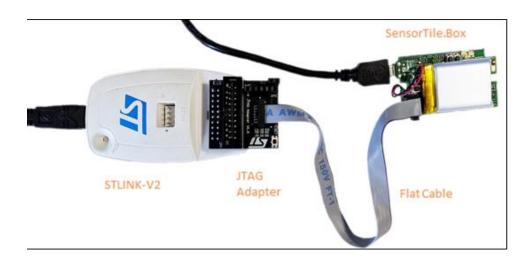




Mini USB



ST-Link/V2







Setup & application examples Software and other prerequisites

- STM32 ST-Link Utility
 - Download and install STSW-LINK004 from www.st.com
- FP-AI-CTXAWARE1
 - Copy the .zip file content into a folder on your PC. The package will contain source code example (Keil, IAR, STM32Cube IDE) based on **STEVAL-MKSBOX1V1**.
- ST BLE Sensor Application for Android or iOS, to download from Google Store / iTunes
- Serial line monitor, e.g. TeraTerm (https://ttssh2.osdn.jp/)



FP-AI-CTXAWARE1 setup overview



Select Function Pack: FP-AI-CTXAWARE1

Android™/iOS™ smartphone and ST BLE Sensor application or Serial Line Monitor









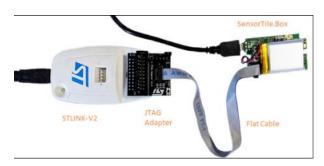




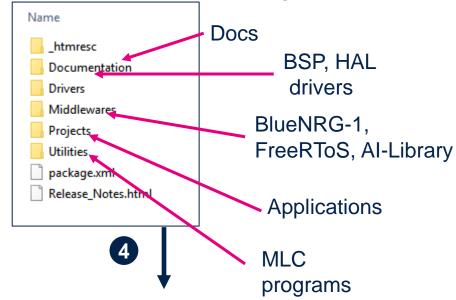


Projects\STM32L4R9ZI-SensorTile.box\Applications\BLELowPower\EWARM







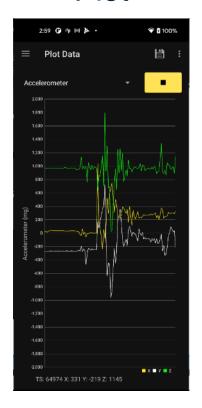


BLELowPower (FP-AI-CTXAWARE1) BLE version

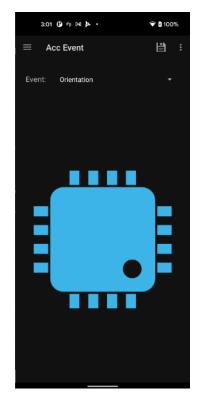
Environmental



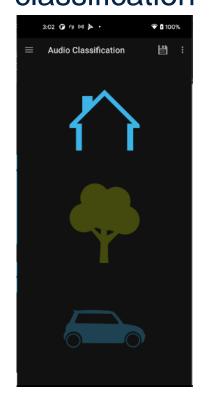
Plot



HW features



Audio scene classification



Activity recognition





BLELowPower (FP-AI-CTXAWARE1) BLE version

Human Activity Recognition (HAR) running on MLC provides indication about activity status ("Stationary", "Walking", "Jogging", "Biking") while Acoustic Scene Classification (ASC) is used for context awareness ("Indoor", "Outdoor", "In Vehicle")



The Human Activity Recognition (HAR) is always running on MLC (LSM6DSOX) without consuming power on STM32

The Acoustic Scene Classification (ASC) runs periodically for controlling the context, and changing the program on MLC if it is necessary

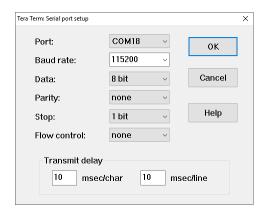
| Indoor/ Outdoor | In Vehicle | | |
|-----------------|------------|--|--|
| Stationary | Stationary | | |
| Walking | Walking | | |
| Jogging | Driving | | |
| Biking | | | |



Command Line AI Application (CLAi) (FP-AI-CTXAWARE1) Command Line version

Same functionalities as the Bluetooth version

Serial Line Monitor Configuration



At the beginning the Board will print out the usage:



- Run the Audio Scene Classification
- Run the Activity Recognition
- Run the Activity Recognition and the Audio Scene Classification in a combined mode
- Print out the dB measured by microphone



```
File Edit Setup Control Window Help

Runtime revision : 6.0.0
Tool revision : (6.0.0)
Network info...
nodes : 9
complexity : 509312 MACC
activation : 9296 bytes
params : 7900 bytes
inputs/outputs : 1/1
If01 s8, scale=0.313726, zero=127, 960 bytes, shape=(30,32,1)
OI01 s8, scale=0.313726, zero=128, 3 bytes, shape=(1,1,3)
Initializing the network asc
Activation buffer : 0x200028a0 (9296 bytes) internal
MLC Start
OK Init Accelero Sensor
-->Activity Recognition for LSM6DSOX MLC INDOOR and OUTDOOR
Program loaded inside the LSM6DSOX MLC Enabled LSM6DSOX INT1 Detection
Enabled LSM6DSOX INT2 Detection
ASC= 4x 9x 0x
ASC= 12x 16x 0x
MLC Activity->STATIONARY
ASC= 21x 21x 0x
ASC= 29x 27x 0x
ASC= 35x 35x 0x
```

Documents & related resources

All documents are available in the Design tab of the related products webpage

FP-AI-CTXAWARE1

- DB4483: STM32Cube function pack for ultralow power context awareness with distributed artificial intelligence (AI): acoustic analysis with NN on MCU and motion analysis with ML on IMU – databrief
- Software setup file

STEVAL-MKSBOX1V1

Gerber files, BOM, Schematic

DB3903: SensorTile.box wireless multi sensor development kit with user friendly app for IoT and wearable sensor applications – databrief

UM2580: How to use the wireless multi sensor development kit with customizable app for IoT and wearable sensor applications – user manual

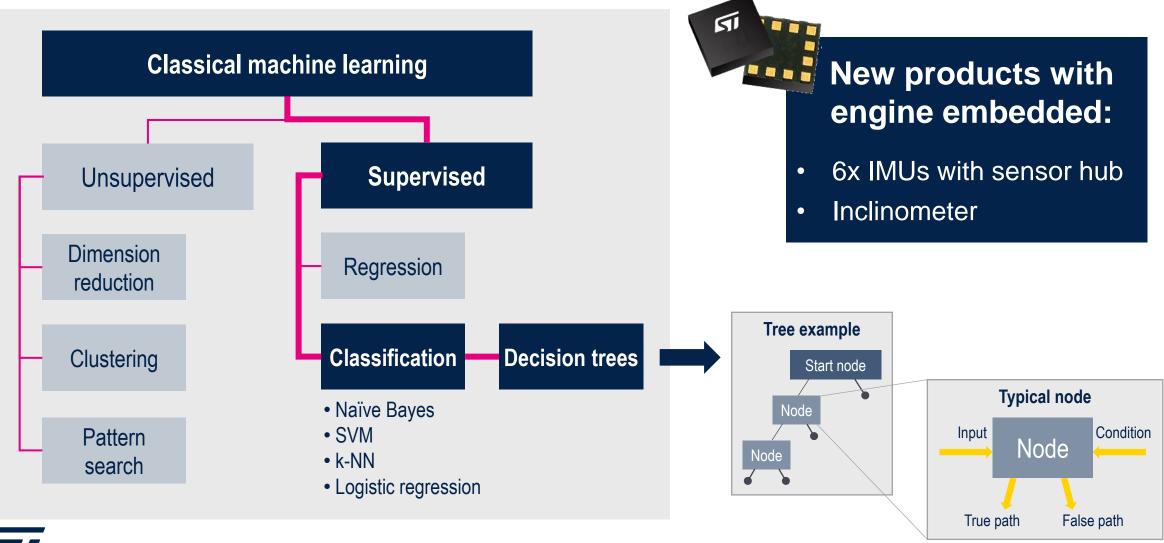


Human activity recognition in smart sensor LSM6DSOX





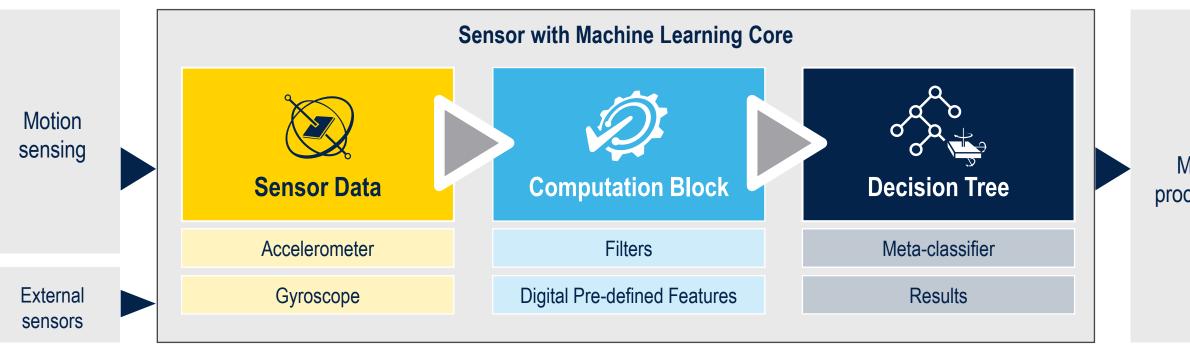
Machine learning embedded in ST sensors





What is a Machine Learning Core (MLC)?

MLC is an in-sensor classification engine based on Decision Tree logic



Main processor

MLC is able to increase accuracy with a better context detectability, offloading the main processor while the built-in sensors identify motion data



ML solutions in accelerometers and gyroscopes

Industrial IoT Personal Electronics LSM6DSRX LSM6DSOX ISM330DHCX **IIS2ICLX** LSM6DSO32X **O** Duration 5.73 km 0:27:34 12.45 km/h 130 bom **L**7/ LYI SM330DHCX LSM6DS0x LSM6DSRX 280 cal Structural health Virtual **Activity Smart antennas** monitoring Reality recognition Leveling Sensor **Gym activity** Industrial IoT instruments **Fusion** recognition **Equipment** Airplane mode **Vehicle stationary Dynamic** installation and detection detection inclinometers



monitoring

Machine Learning solutions in sensors: new developer model approach

Shorter development time and better accuracy with Machine Learning techniques (Decision Trees)

How it works in 5 simple steps and with an intuitive use case:



User defines **Classes** to be recognized



Label data and select **filters** and **features**



Build the decision tree based on a wide range of SW tools.



Program the decision tree into the MLC enabled Sensor



Run the MLC model and process incoming data in real time











MOH

UNICO-GUI



HAR in UNICO-GUI demo



Machine Learning solutions in sensors: ecosystem

A complete suite to create ML applications in sensors



Function packs for quick prototyping

FP-AI-CXTAWARE1



Getting start with ST development kit and GUI









st.com/mlc



Examples for motion recognition and context recognition

GitHub



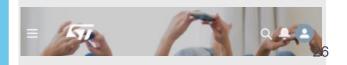
Videos, training material, in products campaign available







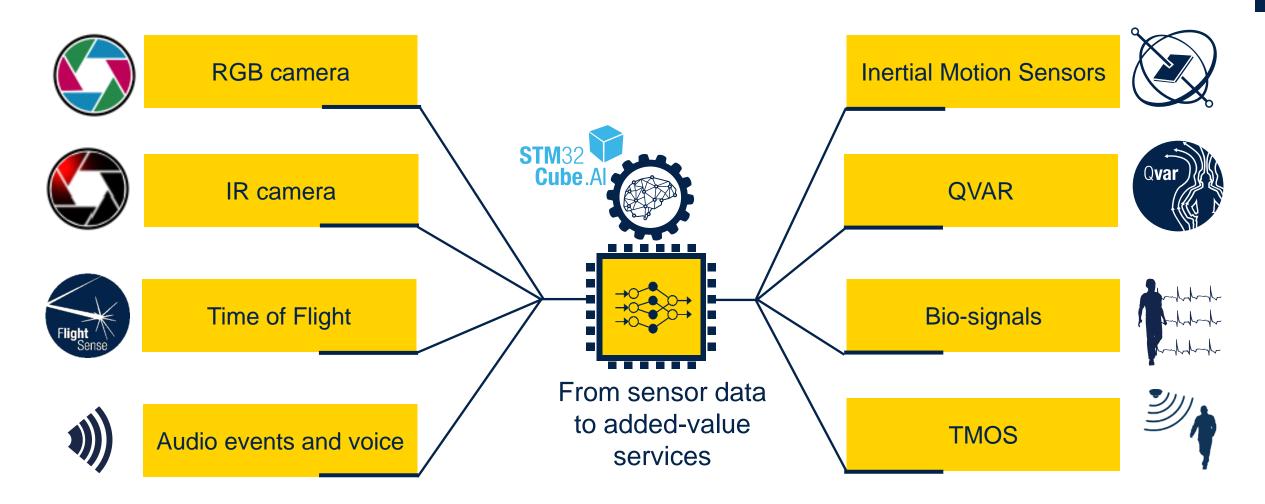
MEMS & Sensor community: MEMS Machine Learning & Al



Acoustic scene classification on STM32L4R9ZIJ6 MCU



Enhance sensor technologies with AI on MCU





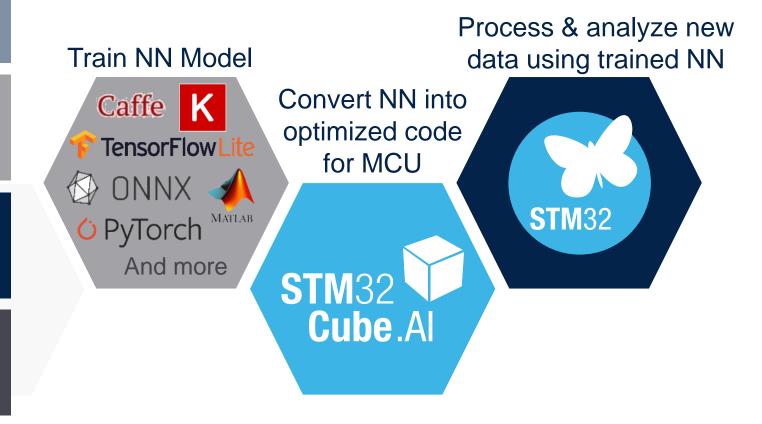
Convert models from any AI framework

Input your framework-dependent, pre-trained Neural Network into the **STM**32**Cube**.Al conversion tool

Automatic and fast generation of an STM32-optimized library

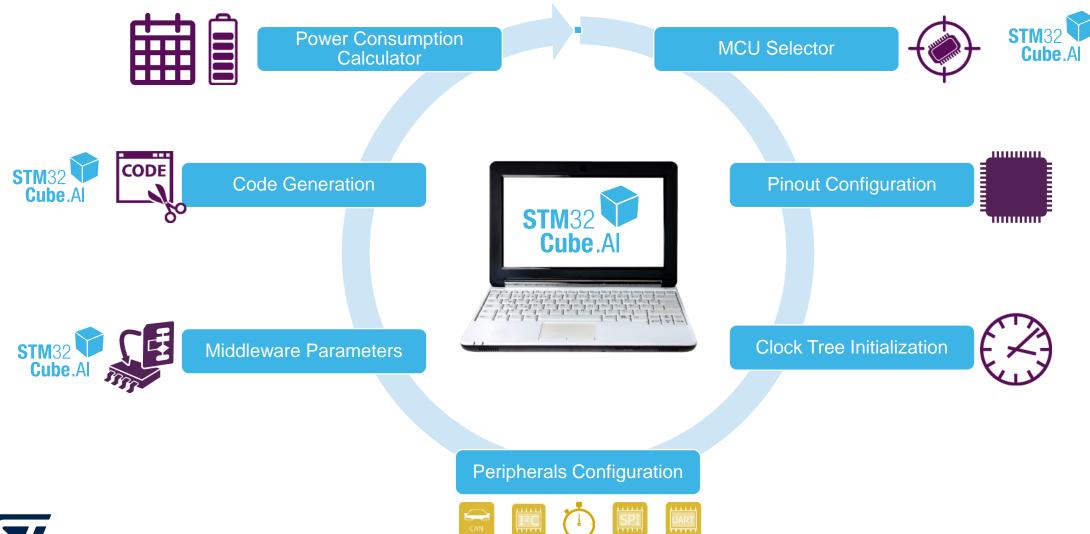
STM32**Cube**.Al offers interoperability with state-of-the-art Deep Learning design frameworks

Any framework that can export models in **ONNX** open format can be imported



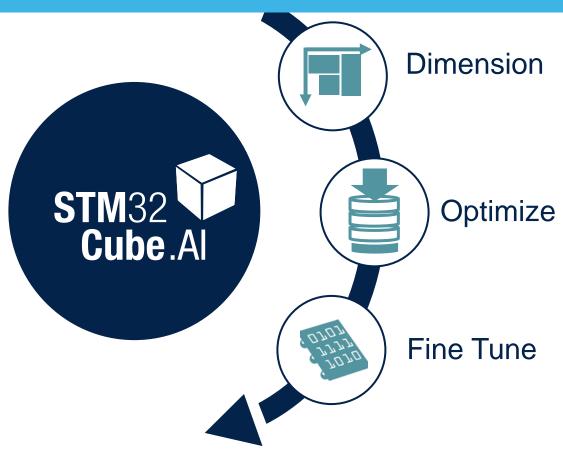


STM32Cube.AI, an STM32CubeMX expansion pack



STM32Cube.Al main features

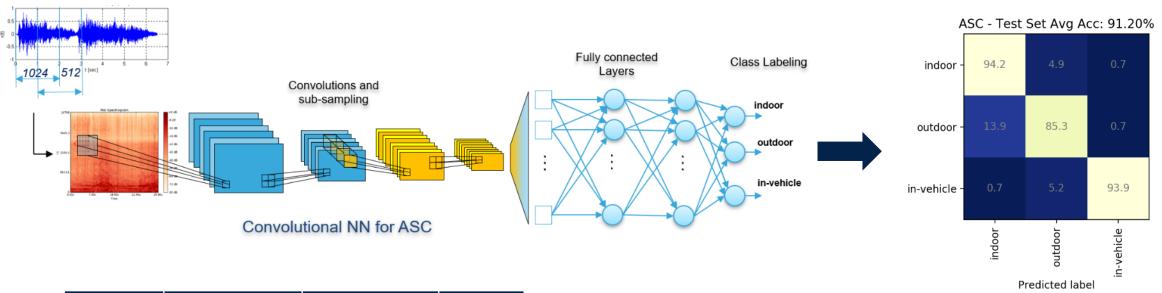
STM32Cube.Al is available both as a graphical and command line interface



- ✓ Quickly assess model footprint requirements
- ✓ Select and configure MCU in STM32CubeMX
- ✓ Review model layers in STM32Cube.AI
- ✓ Generate C-code for pre-trained model
- ✓ Support quantized models to reduce RAM, flash and latency with minimal loss of accuracy
- ✓ Use light run-time libraries
- ✓ Optimize for performance
- ✓ Optimize memory allocation
- ✓ Fine control of weight mapping
- Split between internal and external memory
- Update model without full FW update



Audio scene classification using STM32Cube.Al



| Class | Training Set | Validation Set | Test Set |
|------------|--------------|----------------|----------|
| Indoor | 6:02:50 | 1:21:12 | 7:19:49 |
| Outdoor | 6:03:22 | 1:19:28 | 1:46:53 |
| In-vehicle | 6:03:29 | 1:20:51 | 3:58:13 |
| Total | 18:09:41 | 4:01:31 | 13:04:55 |

Dataset Composition

| Model | Avg. Acc. (%) | NVM (KB) | RAM (KB) | Inference Time (ms) |
|-----------------|------------------|-------------|-------------|---------------------------|
| ASC CNN int8 | 89.17 | 7.71 | 10.02 | 30.631 |

- Quantized ASC int8 model on STM32Cube.AI
- Performance on STM32L476 @ 80MHz



Model Accuracy (f32)

Context awareness demo



Demo video







Want to learn more?











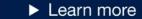
The 5 steps to deploy an ANN on STM32







Convert, Analyze & Deploy ANNs on Automotive MCUs



Our technology starts with You



ST logo is a trademark or a registered trademark of STMicroelectronics International NV or its affiliates in the EU and/or other countries. For additional information about ST trademarks, please refer to www.st.com/trademarks. All other product or service names are the property of their respective owners.





Premier Sponsor



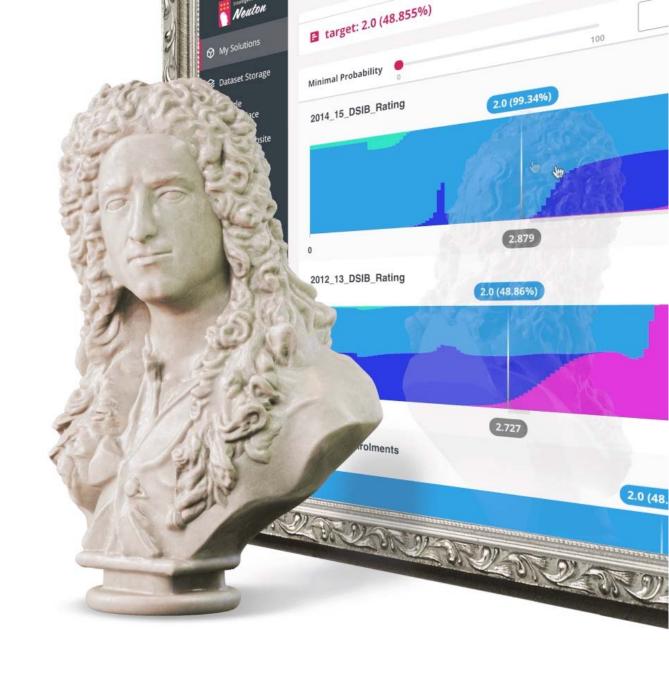
Automated TinyML

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.

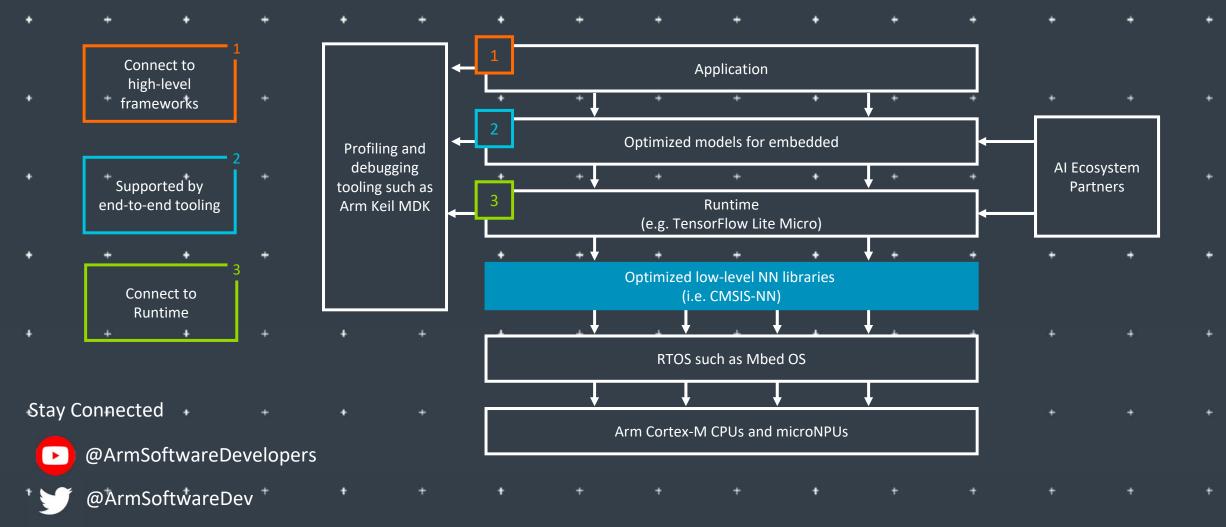
Build Fast. Build Once. Never Compromise.





Executive Sponsors

Arm: The Software and Hardware Foundation for tinyML



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

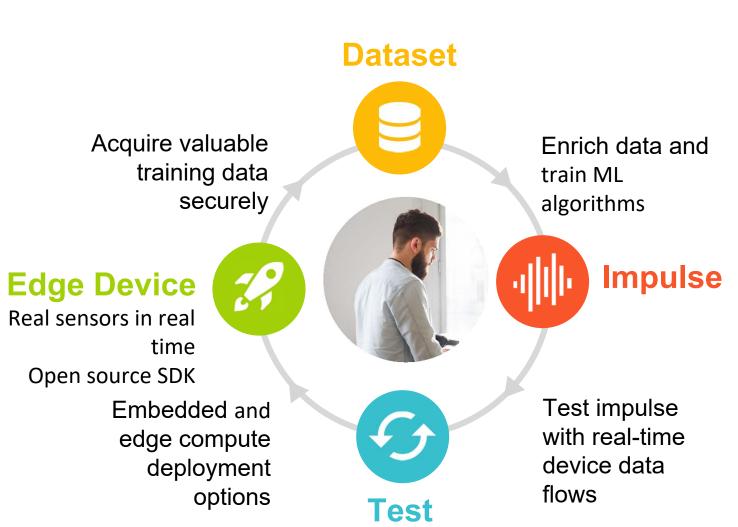


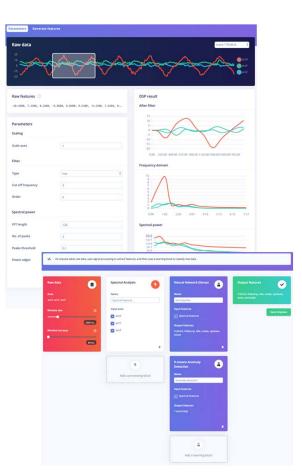
TinyML for all developers











Qualcom Al research

Advancing Al 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 Al across the industry



Perception

Object detection, speech recognition, contextual fusion

Reasoning

Action

Reinforcement learning for decision making



Edge cloud







Mobile

IoT/IIoT







SYNTIANT

<u>Syntiant Corp.</u> 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 ProcessorsTM 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

https://reality.ai







Pre-built Edge Al 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 Al 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



Gold Sponsors



Adaptive AI for the Intelligent Edge

Latentai.com

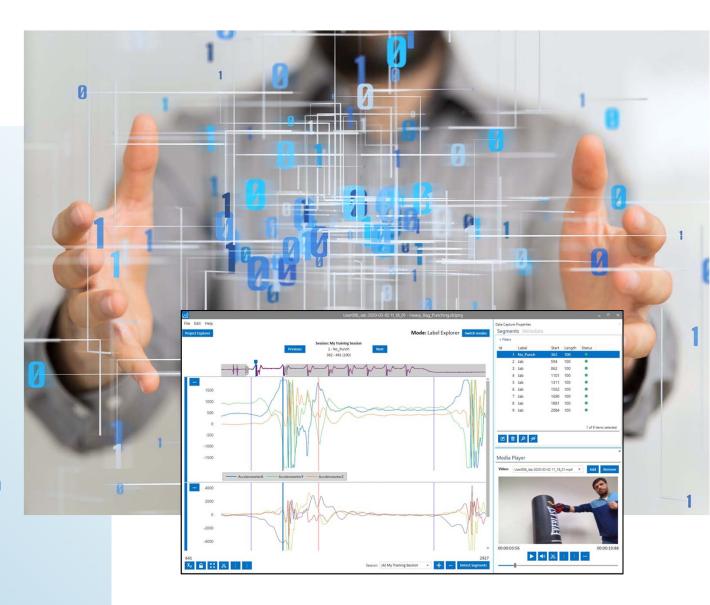


Build Smart IoT Sensor Devices From Data

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 productiongrade 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