Artificial Intelligence overview
1 What is AI at the edge

Smart objects are changing nearly every aspect of our daily lives, helping in our homes, workplaces, cities and factories. STMicroelectronics has been investing in AI for many years, recognizing that microcontrollers provide enough processing power to run AI software on smart objects and enable AI at the edge.

Local processing fixes many problems met when running AI algorithms purely in the Cloud:
- Latency and high communication costs on connected devices
- Reliability in case of limited network bandwidth or connectivity loss
Artificial Intelligence overview

- Power efficiency on battery-operated devices, as high-bandwidth connectivity requires large device batteries, leading to a rise in cost
- Data privacy since data is transformed locally and does not require to send all monitored data over the internet where security breach may occur during transmission or at Cloud storage

To simplify the development of AI algorithms on STM32 microcontrollers, STMicroelectronics has developed a solution called **STM32Cube.AI**.

### 2 STM32Cube.AI overview

STM32Cube.AI solution brings the following:

- An STM32CubeMX extension called **X-CUBE-AI** to convert a Neural Network into optimized code for an STM32 microcontroller.
- Software examples on STMicroelectronics development hardware boards are available for quick prototyping:
  - Audio and motion sensing with the **FP-AI-SENSING1** function pack working in combination of Bluetooth® Low Energy smart sensor application and supported on the following development boards:
    - STM32 Nucleo-64 development board with STM32L476RG MCU NUCLEO-L476RG + Bluetooth
    - Low Energy expansion board X-NUCLEOIB05A1 + Motion MEMS and environmental sensor expansion board X-NUCLEO-IKS01A2
    - SensorTile development kit STEVALSTLKT01V1
    - SensorTile.box multisensor kit with portable sensor box STEVAL-MKSBOX1V1
    - STM32L4 Discovery kit IoT node B-L475E-IOT01A
  - Vision on MCUs with the **FP-AI-VISION1** function pack supported on the following development board:
    - STM32H747I-DISCO Discovery kit connected to the STM32F4DIS-CAM camera daughterboard
  - Vision on Linux for MPU with the **X-LINUX-AI-CV** expansion pack supported on the following development board:
    - Discovery kit with STM32MP157C MPU STM32MP157C-DK2 with a USB camera
    - Evaluation board with STM32MP157A MPU STM32MP157A-EV1
    - Evaluation board with STM32MP157C MPU STM32MP157C-EV1 with their built-in camera module
Artificial Intelligence overview

- Avenger96 board from Arrow using Linaro™ 96Boards format based on the STM32MP157A microprocessor, either with a USB camera or the DesignCore® OV5640 camera mezzanine board from D3 Engineering
- STM32 community with dedicated Neural Networks topic
- AI Partner Program that brings expertise in Machine Learning and STM32 solutions

Check our list of resources for detailed information.

3 Videos related to STM32Cube.AI

4 STM32 supported by STM32Cube.AI

STM32Cube.AI supports all Arm® Cortex®-M4, Cortex®-M33 and Cortex®-M7-based MCUs, as well as STM32 MP1. Other MCUs are supported in our partner ecosystem or can be added on-demand.
STM32Cube.AI supports the following deep learning frameworks:

- **Keras**[^1]
- **TensorFlow**[^2]
- **Caffe**[^3]
- **Lasagne**[^4]
- **ONNX**[^5]

Supported versions and layers for each framework are detailed in X-CUBE-AI release note. This documentation is available in X-CUBE-AI local installation folder at this location:

```
STM32Cube/Repository/Packs/STMicroelectronics/X-CUBE-AI/5.0.0/Documentation/faqs.html
```

## 5 X-CUBE-AI overview

X-CUBE-AI is a tool provided to help users to:

- Generate STM32 code from a Python™ Neural Network described in supported AI frameworks. It supports quantization scheme and optimizations for STM32, reducing memory requirement up to a factor of 4 and decreasing latency and power consumption up to a factor of 3.
- Analyze the network and see if it fits on the target
- Validate the generated C code on the PC or on the target against the original model

It is available through STM32CubeMX Graphical User Interface, or via Command Line Interface.

### Software requirements

Supported operating systems and architectures

- Windows® 10 64-bit (x64)
- Linux® (tested on Red Hat®, Fedora® and Ubuntu® 16.04) 64 bits
- macOS® (minimum version High Sierra)

Prerequisite

- STM32CubeMX: 5.1.0 or higher is required

### License

STM32CubeMX and X-CUBE-AI extension are delivered under the Mix Ultimate Liberty+OSS+3rd-party V1 software license agreement SLA0048.

## 6 STMicroelectronics Resources

All resources are gathered on STM32Cube.AI web page.
7 How to and Examples

- Getting started with X-CUBE-AI Expansion Package for Artificial Intelligence (AI)
- How to integrate STM32Cube.AI generated code in OpenMV ecosystem

8 References

- Keras
- TensorFlow
- Caffe
- Lasagne
- ONNX