ESP32forth and Artificial Intelligence

Marc PETREMANN

Version 1.0 - 03/02/26

Table of Contents

Preamble

This manual is intended for getting started with AI functions using ESP32forth and an ESP32-S3 board.

ESP32-S3: The Microcontroller Tailored for Edge AI

The ESP32-S3 is not just an update; it's Espressif's first SoC to integrate **vector instructions** (SIMD extensions). These instructions massively accelerate basic mathematical operations in neural networks (multiplication and accumulation), propelling inference performance far beyond its predecessors.

🛠️ Key Equipment Advantages

- **AI Acceleration:** Dedicated instructions for tensor calculation and signal processing.
- **Flexible Memory:** Supports up to **16MB of** external PSRAM (essential for loading vision or audio models).
- **Architecture:** Dual-core Xtensa® LX7 at 240 MHz, allowing one core to be dedicated to data acquisition and the other to inference.

📚 Essential Bookstore References

To transform this silicon into intelligence, you will need to rely on one of these three software pillars, depending on your desired level of control.

ESP-DL (Espressif Deep Learning)

the manufacturer's **native and optimized** library . It is designed to get the most out of the S3's vector instructions.

- **Usage:** High-performance model inference (face detection, object recognition).
- **Link :** espressif/esp-dl
- **The "Plus":** Contains a "Model Zoo" with already optimized models (Face Detection, Gesture Recognition).

ESP-NN

This is the low-level (back-end) layer used by ESP-DL and TensorFlow Lite Micro for the ESP32-S3.

- **Usage:** If you are developing your own inference engine and need ultra-fast convolution or activation functions.
- **Link :** espressif/esp-nn

TensorFlow Lite for Microcontrollers (TFLM)

Google's standard, adapted to the Espressif ecosystem via a specific component.

- **Usage:** Maximum portability. If you train your models on TensorFlow/Keras, this is the simplest way to deploy.
- **Link :** espressif/tflite-micro-esp-examples

ESP-Skainet (Speech Recognition)

Framework dedicated to offline voice interfaces.

- **Usage:** Keyword detection ("Wake Word") and voice commands (up to 200 customizable commands without retraining).
- **Link :** espressif/esp-skainet

🚀 Promising Application Areas

Thanks to these tools, the ESP32-S3 excels in three areas of embedded AI (TinyML):

Domain	Concrete example	Recommended bookstore
Vision	People counter, counter reading.	**ESP-WHO** (based on ESP-DL)
Audio	Local home automation assistant, noise detection.	**ESP-Skainet**
Sensation	Predictive maintenance (vibrations), ECG analysis.	**Edge Impulse** (No-Code tool compatible)