




ESP32forth et l'Intelligence Artificielle

Marc PETREMAN



Version 1.0 - 03/02/26

Table des matières

- Préambule..... 3
- ESP32-S3 : Le Microcontrôleur taillé pour l'Edge AI.....4
 -  Atouts Matériels Clés.....4
 -  Références des Librairies Essentielles.....4
 - ESP-DL (Espressif Deep Learning).....4
 - ESP-NN.....4
 - TensorFlow Lite for Microcontrollers (TFLM).....4
 - ESP-Skainet (Reconnaissance Vocale).....5
 -  Domaines d'Applications Prometteurs..... 5

Préambule

Ce manuel est destiné à la prise en main de fonctions IA avec ESP32forth et une carte ESP32-S3.

ESP32-S3 : Le Microcontrôleur taillé pour l'Edge AI

L'ESP32-S3 n'est pas une simple mise à jour ; c'est le premier SoC d'Espressif à intégrer des **instructions vectorielles** (extensions SIMD). Ces instructions permettent d'accélérer massivement les opérations mathématiques de base des réseaux de neurones (multiplications-accumulations), propulsant les performances d'inférence bien au-delà de ses prédécesseurs.

Atouts Matériels Clés

- **Accélération IA** : Instructions dédiées pour le calcul de tenseurs et le traitement de signal.
- **Mémoire Flexible** : Supporte jusqu'à **16 Mo de PSRAM** externe (indispensable pour charger des modèles de vision ou d'audio).
- **Architecture** : Dual-core Xtensa® LX7 à 240 MHz, permettant de dédier un cœur à l'acquisition de données et l'autre à l'inférence.

Références des Bibliothèques Essentielles

Pour transformer ce silicium en intelligence, vous devrez vous appuyer sur l'un de ces trois piliers logiciels, selon votre niveau de contrôle souhaité.

ESP-DL (Espressif Deep Learning)

C'est la bibliothèque **native et optimisée** par le fabricant. Elle est conçue pour tirer le maximum des instructions vectorielles du S3.

- **Usage** : Inférence de modèles haute performance (détection de visages, reconnaissance d'objets).
- **Lien** : [espressif/esp-dl](https://github.com/espressif/esp-dl)
- **Le "Plus"** : Contient un "Model Zoo" avec des modèles déjà optimisés (Face Detection, Gesture Recognition).

ESP-NN

C'est la couche de bas niveau (back-end) utilisée par ESP-DL et TensorFlow Lite Micro pour l'ESP32-S3.

- **Usage** : Si vous développez votre propre moteur d'inférence et avez besoin de fonctions de convolution ou d'activation ultra-rapides.
- **Lien** : [espressif/esp-nn](https://github.com/espressif/esp-nn)

TensorFlow Lite for Microcontrollers (TFLM)

Le standard de Google, adapté à l'écosystème Espressif via un composant spécifique.

- **Usage** : Portabilité maximale. Si vous entraînez vos modèles sur TensorFlow/Keras, c'est la voie la plus simple pour le déploiement.
- **Lien** : [espressif/tflite-micro-esp-examples](https://github.com/espressif/tflite-micro-esp-examples)

ESP-Skainet (Reconnaissance Vocale)

Framework dédié aux interfaces vocales hors-ligne.

- **Usage :** Détection de mots clés ("Wake Word") et commandes vocales (jusqu'à 200 commandes personnalisables sans réentraînement).
- **Lien :** espressif/esp-skainet



Domaines d'Applications Prometteurs

Grâce à ces outils, l'ESP32-S3 excelle dans trois domaines de l'IA embarquée (TinyML) :

Domaine	Exemple concret	Librairie conseillée
Vision	Compteur de personnes, lecture de compteurs.	ESP-WHO (basée sur ESP-DL)
Audio	Assistant domotique local, détection de bruits.	ESP-Skainet
Sensation	Maintenance prédictive (vibrations), analyse ECG.	Edge Impulse (outil No-Code compatible)