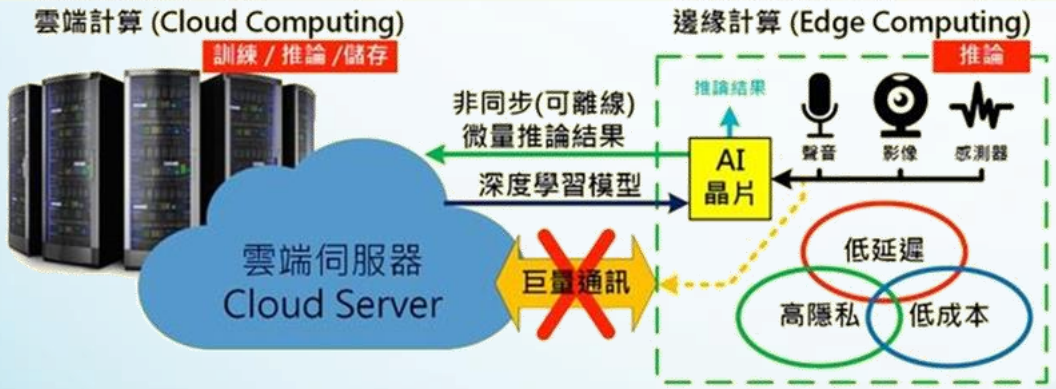
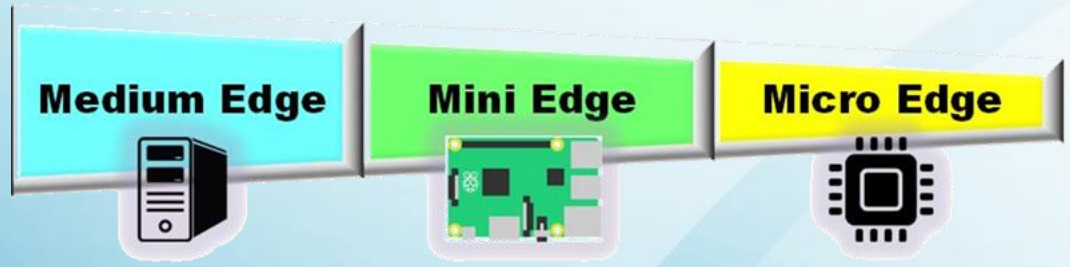
OmniXRI's Edge AI & TinyML 小學堂





【第7講】
微型機器學習簡介

歐尼克斯實境互動工作室 (OmniXRI Studio)
許哲豪 (Jack Hsu)

簡報大綱

- 7.1.嵌入式系統與微控制器
- 7.2. TinyML技術現況
- 7.3. TinyML開發平台
- 7.4. TinyML主要應用

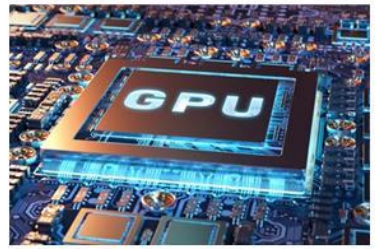
本課程完全免費，請勿移作商業用途！
歡迎留言、訂閱、點讚、轉發，讓更多需要的朋友也能一起學習。

完整課程大綱： https://omnixri.blogspot.com/2024/02/omnixris-edge-ai-tinyml-0.html
課程直播清單： https://www.youtube.com/@omnixri1784/streams

7.1. 嵌入式系統與微控制器

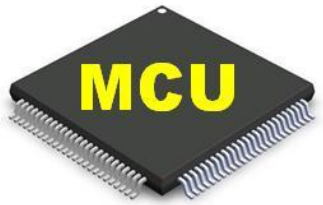
- 微型機器學習與生成智慧
- 嵌入式人工智慧定位
- 嵌入式系統架構
- 單晶片架構與指令集

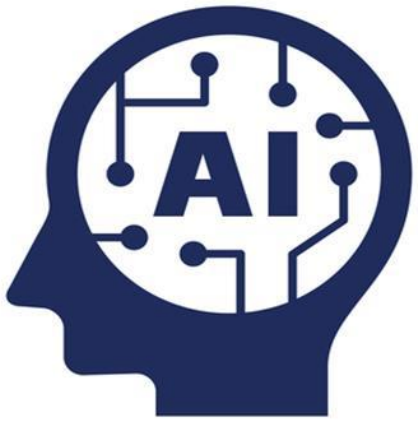
微型機器學習 vs. 生成智慧

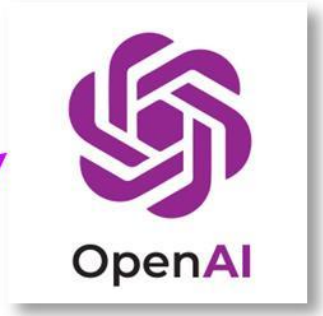


- 邊緣端極少資源
  儲存(Flash, RAM)、計算
- 智慧感測
  聲音、影像、運動...
- 超小模型、超低功耗

極小化





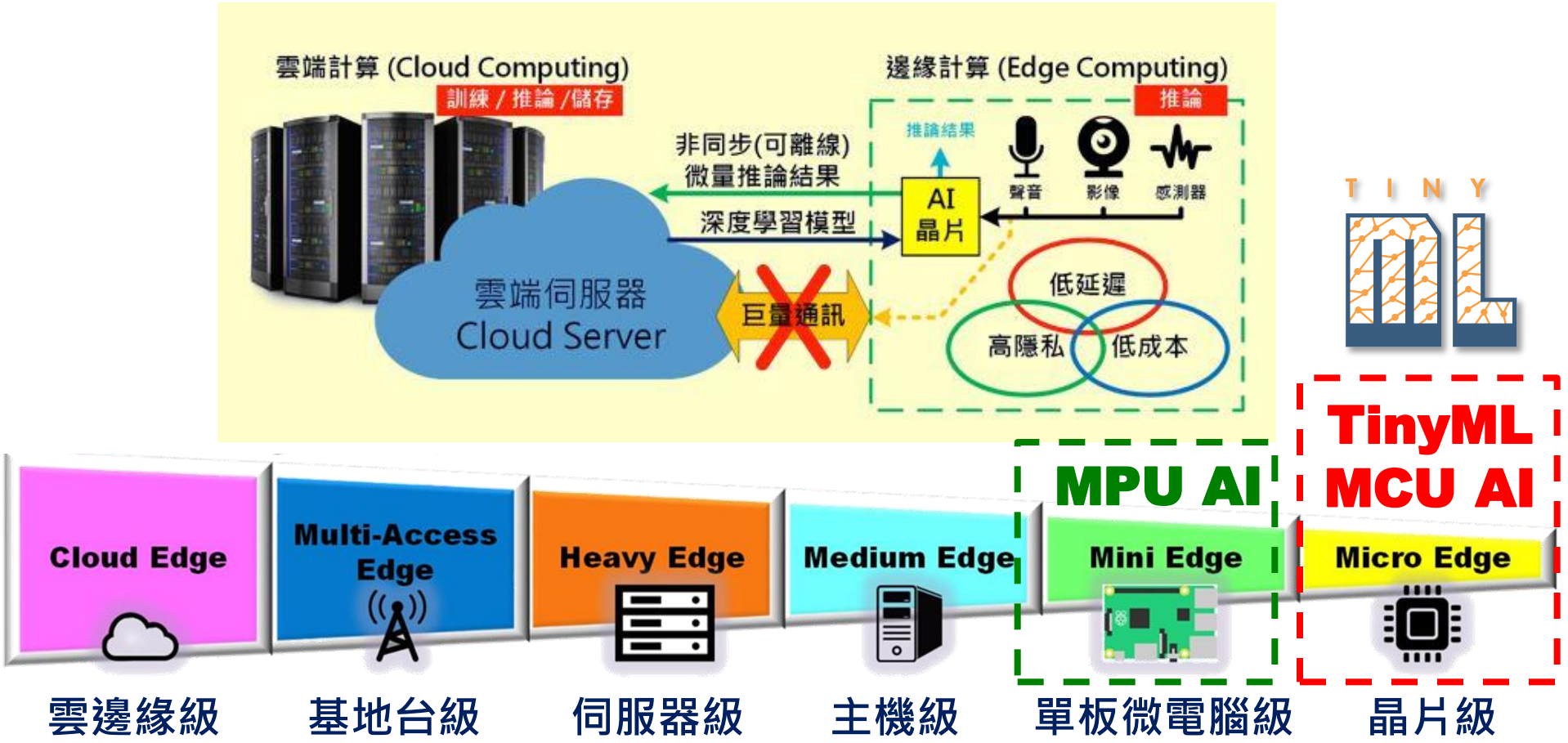


極大化

- 雲端無限資源
  儲存、計算、頻寬、功耗
- AI生成
  對話、影像、程式...
- 超巨大資料集、模型

資料來源：https://omnixri.blogspot.com/2023/04/20230420aiexpo-aiedge-ai.html

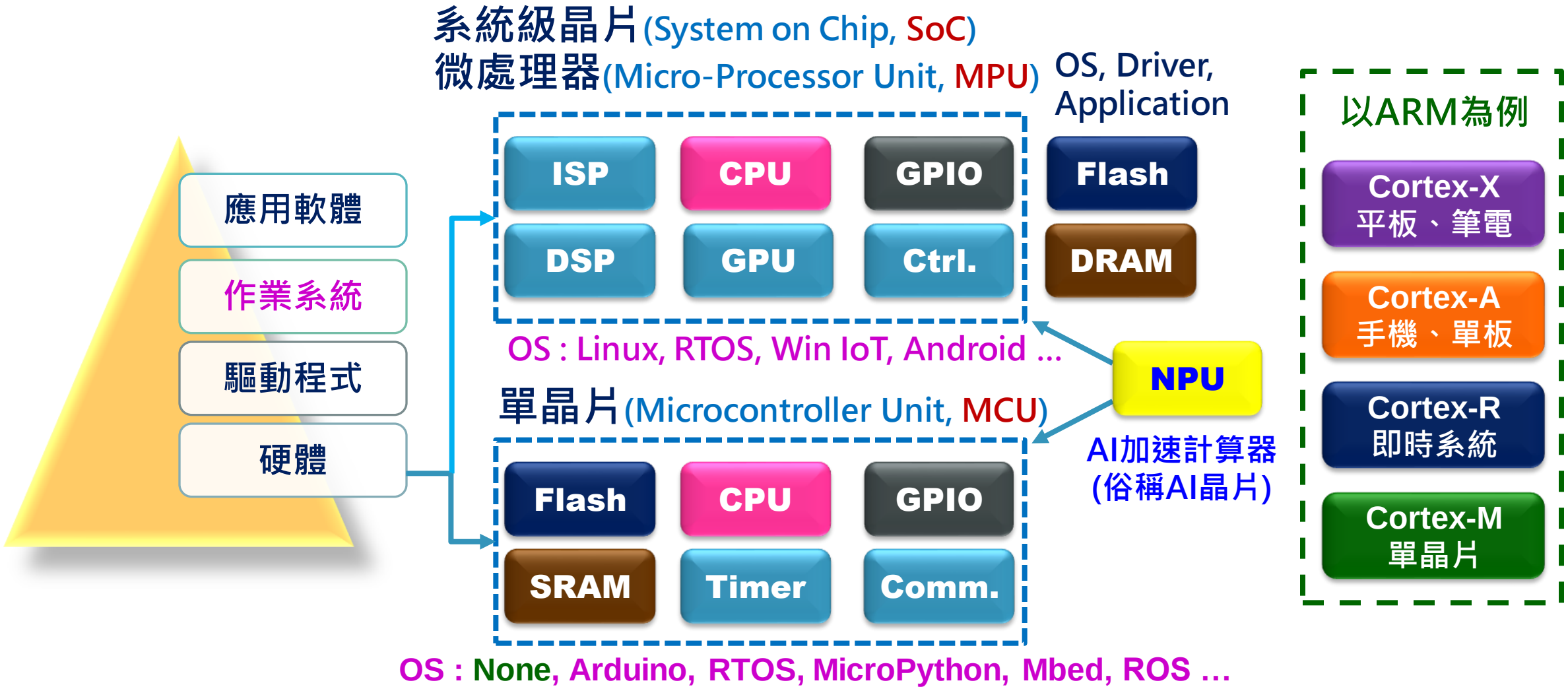
嵌入式人工智慧定位



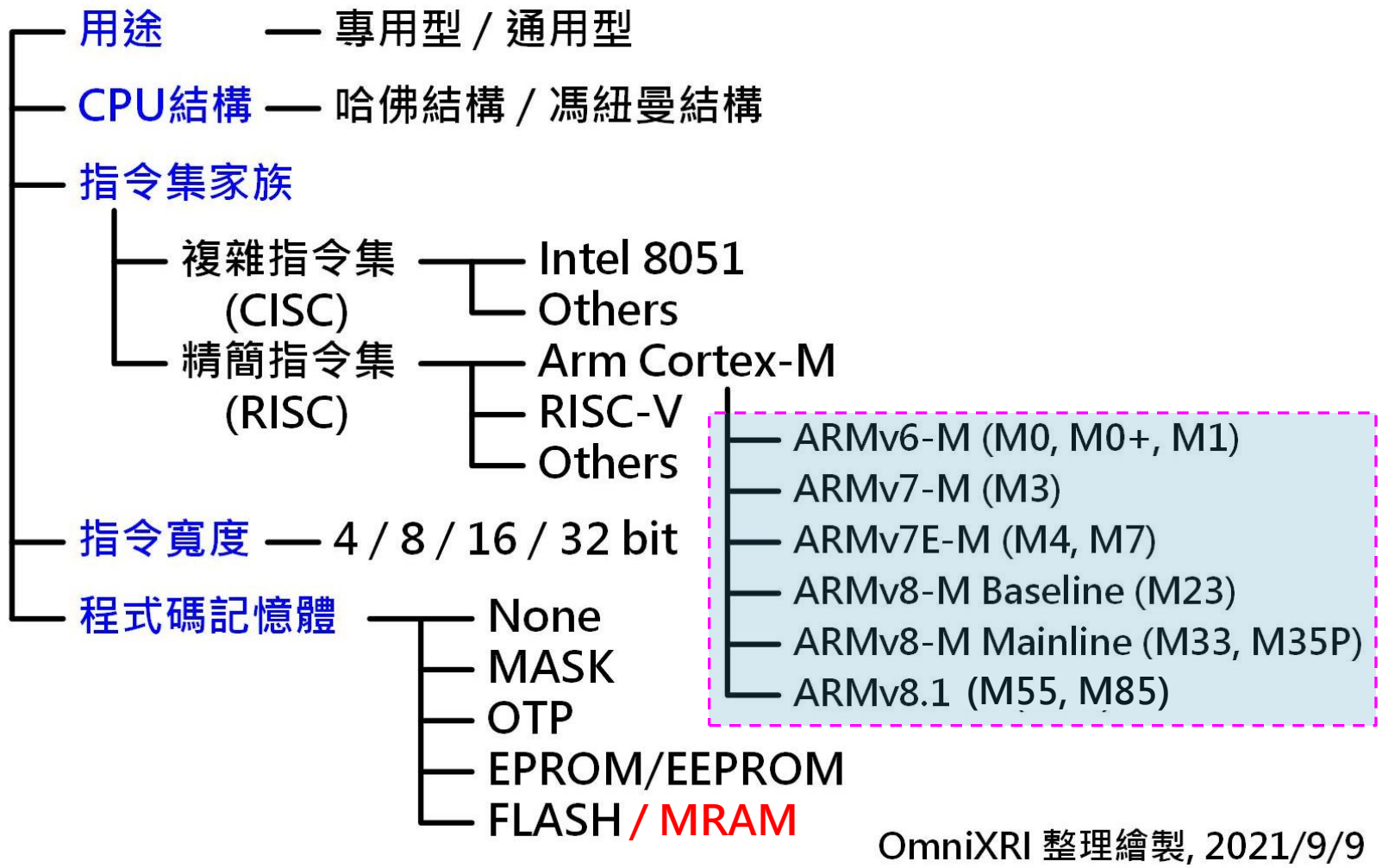
廣義邊緣智慧裝置：不上網就具備AI模型推論能力之裝置

狹義邊緣智慧裝置：低功耗、高性能、可移動、電池推動之具備AI模型推論能力之裝置

嵌入式系統架構

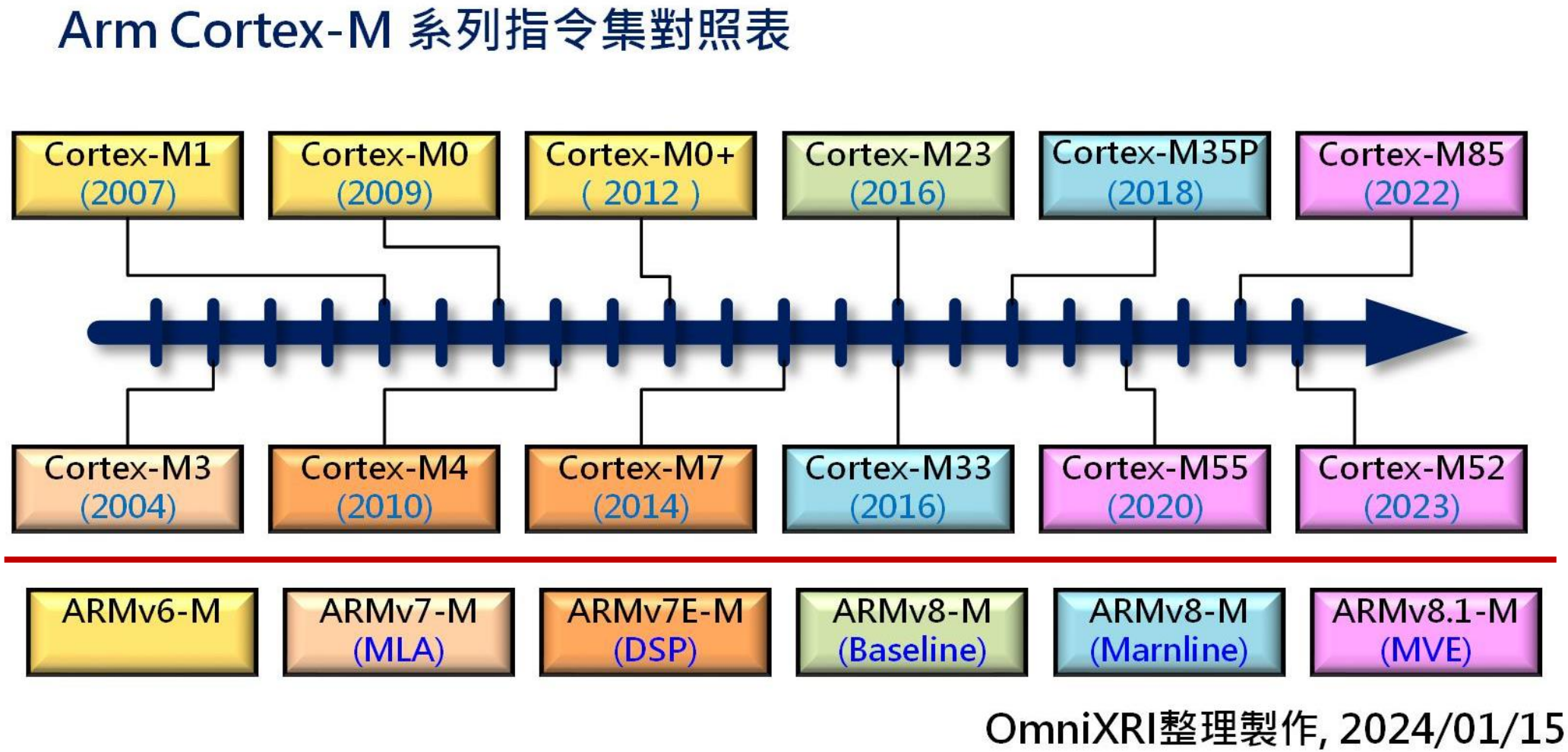

系統級晶片(System on Chip, SoC)
微處理器(Micro-Processor Unit, MPU)
OS, Driver, Application
應用軟體
作業系統
驅動程式
硬體
ISP
CPU
GPIO
DSP
GPU
Ctrl.
Flash
DRAM
OS : Linux, RTOS, Win IoT, Android ...
NPU
AI加速計算器
(俗稱AI晶片)
單晶片(Microcontroller Unit, MCU)
Flash
CPU
GPIO
SRAM
Timer
Comm.
OS : None, Arduino, RTOS, MicroPython, Mbed, ROS …
以ARM為例
Cortex-X
平板、筆電
Cortex-A
手機、單板
Cortex-R
即時系統
Cortex-M
單晶片

通用型單晶片(MCU)主要分類方式



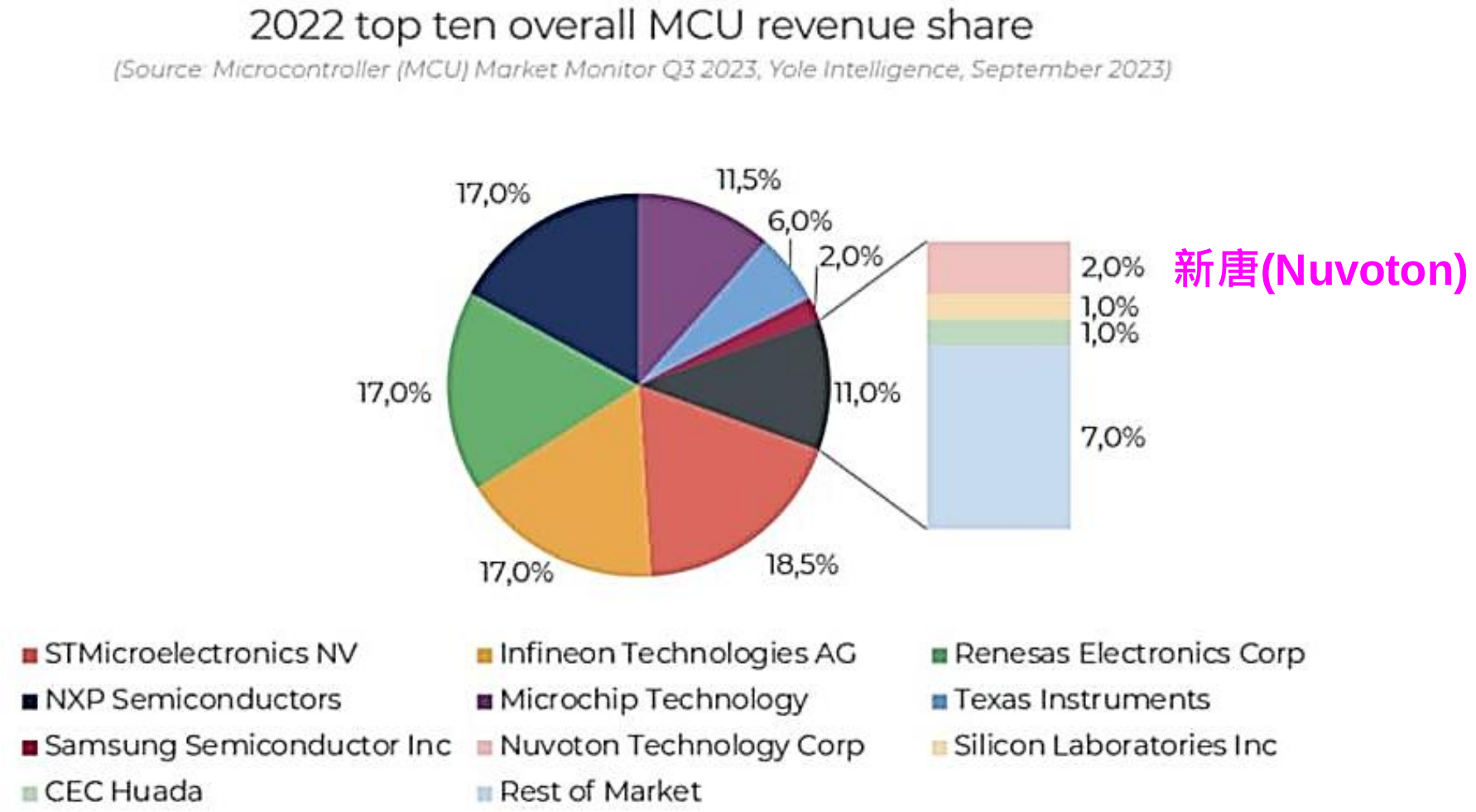
資料來源：https://omnixri.blogspot.com/2021/09/aiottinymlmcu.html

Arm Cortex-M 指令集對照表



資料來源：https://omnixri.blogspot.com/2024/01/vmaker-edge-ai-13-npuai.html

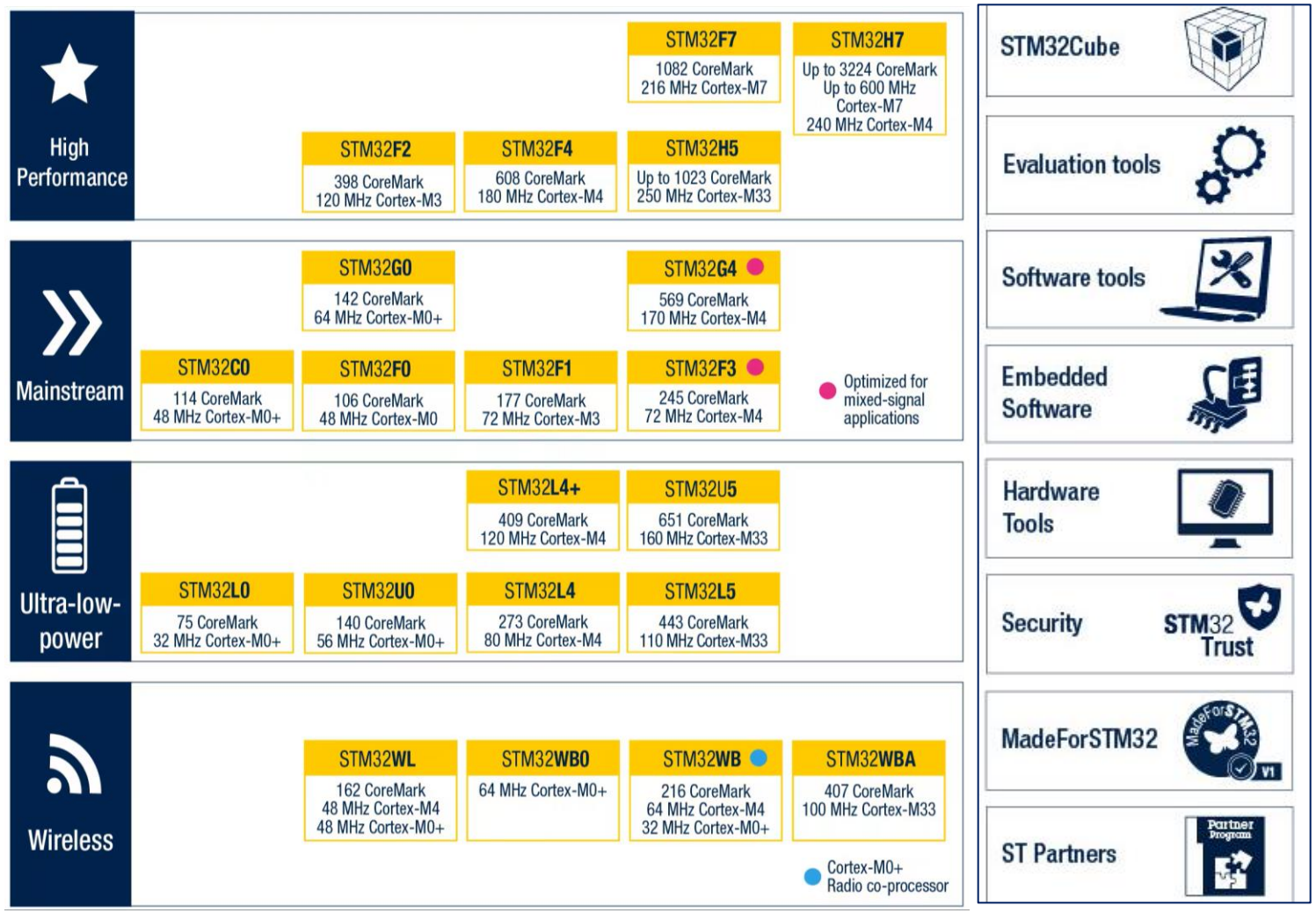
全球十大MCU供應商 (Q3 2023)



新唐(Nuvoton)

資料來源：https://36kr.com/p/2528780809938432

STM32 MCUs 32bit Arm Cortex-M系列



資料來源：https://www.st.com/en/microcontrollers-microprocessors/stm32-32-bit-arm-cortex-mcus.html

Cortex-M0
48MHz
Cortex-M0+
32M~56MHz
Cortex-M3
72M~120MHz
Cortex-M33
110M~250MHz
Cortex-M4
48M~240MHz
Cortex-M7
216M~600MHz

MIPS / DMIPS

2.3~6.1
CoreMark/MHz

台灣通用型MCU主要供應商 (2021/9)

公司名稱	產品系列	工作時脈	程式碼區	隨機記憶區
Intel 8051 家族 (8bit) [1T, 6T 指令週期]				
義隆 Elan	EM85	16K ~ 20MHz	16KB ~ 18KB	768B ~ 1024B
迅杰 ENE	TK18	28MHz	32KB	1280B
凌通 Generalplus	GPM831/8F37 GPM82(EEPROM)	48M ~ 64MHz 4M ~ 16MHz	32KB 2KB	256B ~ 1KB 128B
笙泉 Megawin	MG82/89/87/84	24M ~ 48MHz	8K ~ 64KB	256B ~5KB
世紀 Myson	CS89	16M ~ 25MHz	16K ~ 128KB	1K ~6KB
新唐 NuvoTon	N78, W78, N76E, N79E, W79E, MS51, ML51, ML54, N76E, ML56	10K ~ 24MHz	0K ~ 128KB (0K 表外掛)	256B ~ 4KB
陞達 Sentelic	STL5000	51.2MHz	8KB	8KB
松翰 Sonix	SN8F5	16K ~ 48MHz	16KB	512B
太欣 Syntek	STK60	2M ~ 32MHz	16K ~ 64KB	256B ~ 2304B
偉詮 Weltrend	WT51F/56F WT61	12M ~ 24MHZ 12MHz	4K ~ 48KB 64K ~ 128KB	256B ~ 1KB 1K ~ 4KB
Arm Cortex M0 家族 (32bit) [ArmV6-M 指令集]				
凌通 Generalplus	GPM32F0	80M ~ 92MHz	32K ~ 64KB	2K ~8KB
笙泉 Megawin	MG32F02	48MHz	32K ~ 132KB	4K ~ 16KB
新唐 NuvoTon	M03x/M0Ax/ Nano1xx/M05x NUC02x/1xx/2xx/	24M ~ 72MHz	16K ~ 512KB	2K ~ 96KB
松翰 Sonix	SN32F	16K ~ 72MHz	32K ~ 256KB	2K ~ 32KB
偉詮 Weltrend	WT32L	4M ~ 32MHz	32K ~ 64KB	8KB
Arm Cortex M0+家族 (32bit) [ArmV6-M 指令集]				
盛群 Holtek	HT32F0x/F5x/F6x	20M ~ 60MHz	16K ~ 256KB	4K ~ 32KB
九齊 Nyquest	NX1 (OTP) NX1 (Flash)	32MHz	32K~96KB 2M~32MB	4K ~ 12KB
Arm Cortex M3 家族 (32bit) [ArmV7-M 指令集]				
盛群 Holtek	HT32F1x	72M ~ 96MHz	32K ~ 256KB	8K ~128KB
Arm Cortex M4 家族 (32bit) [ArmV7E-M 指令集]				
新唐 NuvoTon	M48x/M45x/M4TK/ NUC4xx/NUC5xx	72M ~192MHz	40K ~ 2MB	16K ~ 128KB
Arm Cortex M7 家族 (32bit) [ArmV7E-M 指令集]				
新唐 NuvoTon	KM1M7	160MHz	512KB	128KB
Arm Cortex M23 家族 (32bit) [ArmV8-M Baseline 指令集]				
新唐 NuvoTon	M25x/ M261x/ M235x	32K ~ 32MHz	64K ~256KB	8K ~32KB

公司名稱	產品系列	工作時脈	程式碼區	隨機記憶區
晶心(Andes) RISC-V 家庭 (32bit)				
紘康 Hycon	HY16F18/19/39	2MHz	64KB	8KB
偉詮 Weltrend	WT58/59	1M ~ 25MHz	64K ~ 192KB	8K ~ 12KB
其它家族 (4bit)				
通泰 Tontek	TTR/TTU/TTP	910K ~ 4MHz	1K ~4KB	54B ~ 384B
其它家族 (8bit)				
祐華 Alpha	AM8EC (OPT)	14K ~ 16MHz	1K ~ 4KB	64B ~ 256B
義隆 Elan	EM78 (OTP) EM88	16K ~ 4MHz 16K ~ 32MHz	1.6K ~ 15KB 2K ~ 24KB	32B ~ 656B 112B ~ 1072B
凌通 Generalplus	GPM61 (MASK/OTP) GPM81	4M ~ 8MHz 16MHz	16K ~ 128KB 18K ~ 128KB	64B ~ 512B 256B
盛群 Holtek	HT46R (OTP) HT48F HT66F/68F	8M ~ 12MHz 400K ~ 12MHz 4M ~ 16MHz	1KB ~ 8KB 1K ~ 4KB 0.5K ~ 64KB	32B ~ 384B 64B ~ 160B 16B ~ 4KB
紘康 Hycon	HY1x	12K ~ 16MHz	2K ~ 8KB	128B ~ 640B
新唐 NuvoTon	KM101E KM101L	20MHz 10MHz	32K ~ 928KB 64KB	1K ~ 10KB 2K ~ 16KB
九齊 Nyquest	NY8A/B (OTP) NY8L (65C02)	20MHz	1K ~ 2KB 4K ~ 128KB	96B ~ 128B 64B ~ 128B
應廣 Padauk	PMS/PMC (OTP) PFS/PFC	32K ~ 8MHz 32K ~ 4MHz	0.5K ~ 4KB 2K ~ 3KB	64B ~ 256B 128B ~ 256B
松翰 Sonix	SN8F2 SN8Px (OTP)	16K ~ 16MHz 32K ~ 16MHz	4K ~ 64KB 1K ~ 24KB	128B ~ 6KB 48B ~ 512B
其它家族 (16bit)				
金麗科 RDC	R88/R11 (80C186)	25M ~ 80MHz	512KB	?
其它家族 (32bit)				
祐華 Alpha	AM32A0 (OTP) AM32A1/B2	14MHz 32K ~ 54MHz	32K ~ 2MB 32K ~ 64KB	4K ~ 8KB 4K ~ 16KB
新唐 NuvoTon	KM103	10MHz	64K ~ 512KB	4K ~ 64KB
金麗科 RDC	HB301/R86	100M~300MHz	4M ~16MB	128M~256MB (DDR2 外掛)
混合型 MCU				
凌陽 Sunplus	SP7021 (Arm Cortex-A7*4 + Arm 926 + 8051)	1 GHz (CA7) 202MHz(926) 202MHz(8051)	0KB (NOR, NAND Flash 外掛)	512KB(SRAM)+ 128MB/512MB (DDR3)

指令位元數：
4/8/16/32 bit

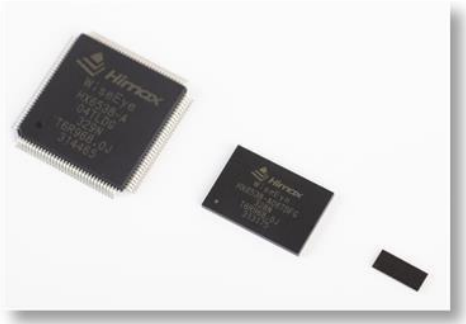
MCU家族：
8051
Cortex-M0
Cortex-M0+
Cortex-M3
Cortex-M4
Cortex-M7
Cortex-M23
RISC-V
Others.

資料來源：https://omnixri.blogspot.com/2021/09/aiottinymlmcu.html

Arm Based MCU + NPU解決方案

CES 2024 MCU AI 相關產品 (Arm Cortex-M55 + NPU Ethos U55)

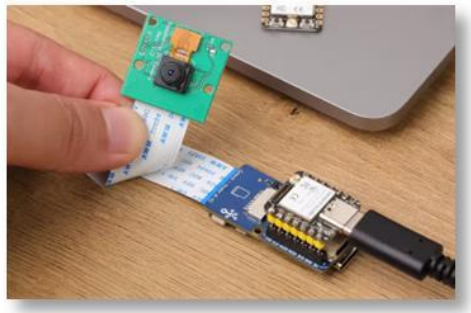
ALIF Ensemble
(E1/E3/E5/E7)



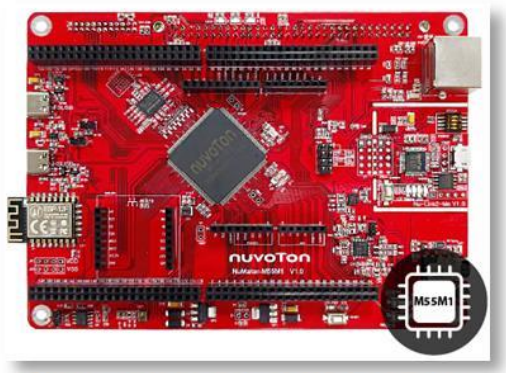
Himax WiseEye2
(HX6538)

Infineon PSoC Edge

Vision AppKit
(Telit + ALIF E3)



Seeed Grove Vision AI
v2 Kit (Himax HX6538)



Nuvoton NuMaker-M55M1

OmniXRI整理製作, 2024/02/15

資料來源：https://omnixri.blogspot.com/2024/02/vmaker-edge-ai-14-ces-2024-edge-aitinyml.html

智慧單晶片（MCU+NPU）

品牌	晶片名稱	CPU@Mhz	NPU(TOPS)	Flash	RAM
ADI	MAX78000	Cortex-M4F @100MHz RISC-V @60MHz	?	512KB	128KB
AONdevice	AON1120	RISC-V	NPU x2 (?TOPS)	?	?
ESPRESSIF	ESP32-S3	Xtensa LX7 @240MHz x2	DSP	8MB	8MB PSRAM
NXP	MCX-N54/N94	Cortex-M33 @ 150MHz *2	DSP + N1-16 (? TOPS)	2MB	512KB+16KB Catch
Realtek	RTL8735B	Arm v8M @500MHz	(0.4TOPS)	768K	512K +128MB DDR2
SiFive	X390	RISC-V	1024bit VLEN 512bit DLEN VCIX	?	?
ST	STM32N6	?	Neural-Art (?TOPS)	?	?

7.2. TinyML技術現況

- 何謂TinyML
- TinyML應用限制
- TinyML加速運算
- TinyML效能評比

何謂微型機器學習 (TinyML)

2019年3月舉辦首次tinyML峰會，有超過90家公司共同參與。

https://www.tinyml.org/

- **微型機器學習(Tiny Machine Learning)**被廣泛定義為一個快速發展的機器學習技術和應用領域，包括**硬體**（專用積體電路）、**算法**和能夠執行設備上感測器（視覺、聲音、運動IMU、生物醫學等）**數據分析**的軟體極低的功耗，通常在 **mW** 範圍內及以下，因此可實現各種**始終在線（Always On)**的用例並針對**電池供電**設備。
- **TinyML**同義字，**MCU AI**, Tiny AI, Micro AI, On Device AI, Embedded ML, Embedded AI, Smart Sensor, Intelligence Sensor ...

tinyML基金會2023主要贊助商

包含晶片(MCU, AI Chip)、開發工具、雲端平台、智慧感測及裝置等公司

TinyML應用限制

MCU等級跨度大

以Arm Cortex-M為例
M0+, M3, M4, M7, M55, M85, M33

指令速度從數十MHz到數百MHz
程式碼儲存空間從數KB到數MB
SRAM從數KB到數MB

圖片來源：https://www.arm.com/blogs/blueprint/ai-for-iot-devices

MCU等級TinyML

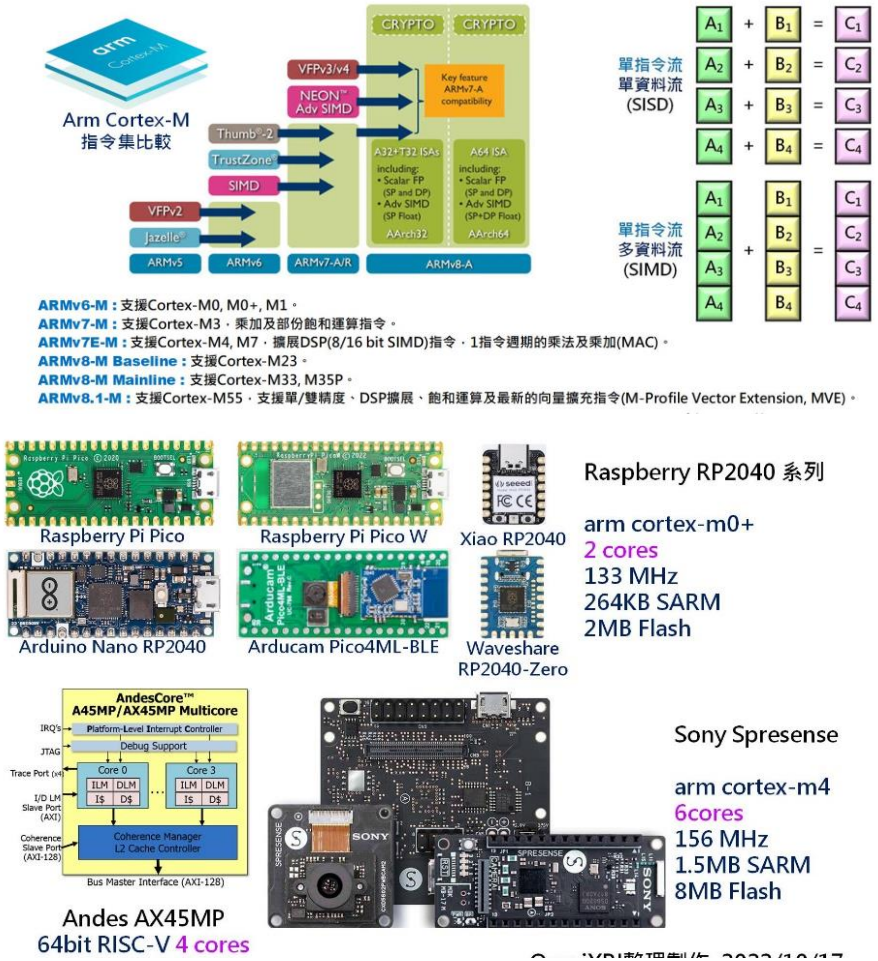
優點

- 低單價、功耗
- 低延時(反應快)
- 高隱私(免上網)
- 易連接各式感測器及通訊模組

缺點

- 速度、算力不足
- 記憶體不足
- 儲存能力小
- 難以在線訓練

TinyML加速運算 – 硬體增速



提高工作時脈速度

- 6MHz ~550MHz

平行/向量指令運算

- arm SIMD, NEON…
- RISC-V P, V Extension

多核加速

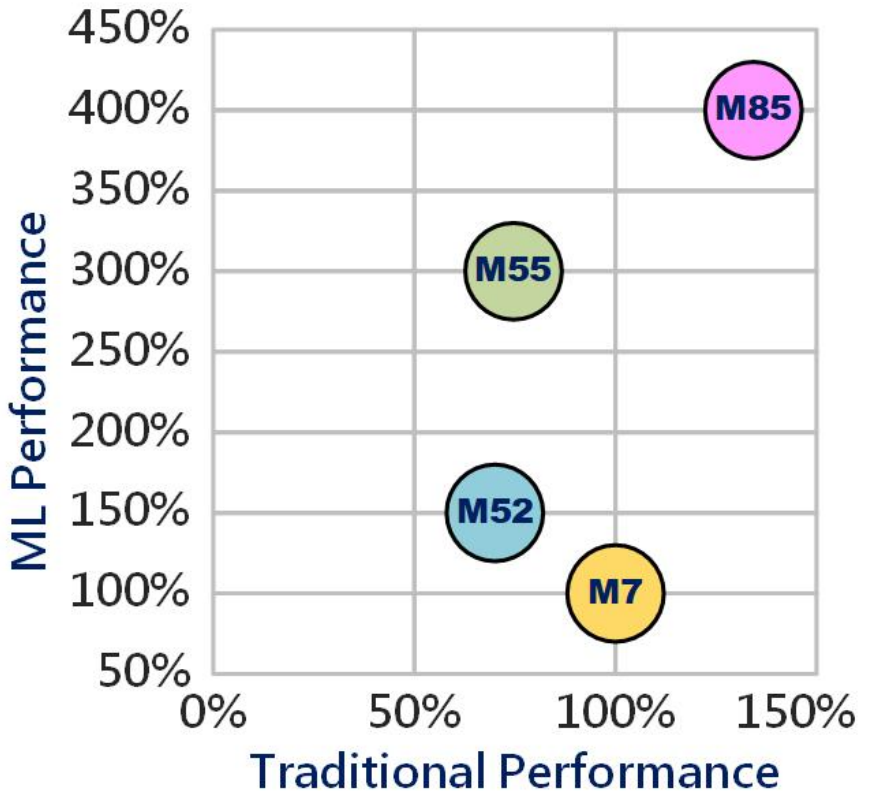
- 同質多核
- 大核加小核
- 異質(arm+RISC-V)
- CPU+GPU+DSP+NPU

資料來源：https://omnixri.blogspot.com/2022/10/20221021ai.html

Arm Cortex-M 指令集與加速計算比較

IP	指令集	Helium	DMIPS/MHz	CoreMark/MHz
Cortex-M7	v7E-M	X	2.31	5.29
Cortex-M55	v8.1-M	Dual-beat	1.69	4.40
Cortex-M85	v8.1-M	Dual-beat	3.13	6.28
Cortex-M52	v8.1-M	Single-beat	1.60	4.30

註：純CPU比較，不含NPU



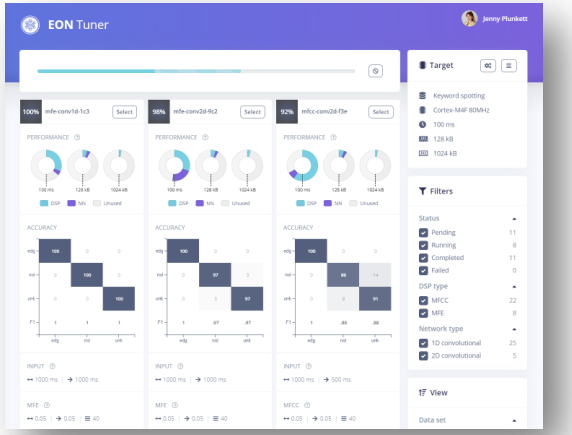
OmniXRI整理製作, 2024/01/15

資料來源：https://omnixri.blogspot.com/2024/01/vmaker-edge-ai-13-npuai.html

TinyML加速運算 – 模型優化 (AutoML)

以Edge Impulse EON Tuner為例

指定工作硬體，依據輸入資料找出潛在的信號處理模組和神經網路架構，並提供不同的計算時間、推論精度、裝置延遲及記憶體需求的排列組合給使用者自行選用。



KERAS

Type	Filters	Kernel	Rate
Data augmentation			
conv1d	16	3	-
conv1d	32	3	-
conv1d	64	3	-
conv1d	128	3	-
dropout	-	-	0.5
dense	64	-	-
dropout	-	-	0.5

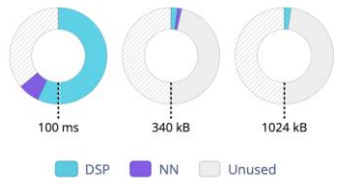
View

Data set
- Validation
- Train
- Test

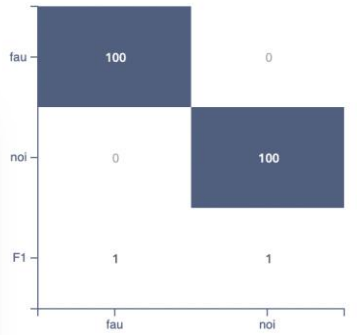
Precision
- int8
- float32

100% mfe-conv1d-352 Select

PERFORMANCE



ACCURACY



INPUT
1000 ms | 1000 ms

MFE
0.032 | 0.032 | 40

Target
- Continuous audio
- Cortex-M7 216MHz
- 100 ms
- 340 kB
- 1024 kB

Sort

General
- Accuracy
- Latency
- RAM
- ROM
- Last updated

F1-score
- faucet
- noise

Precision

Recall

資料來源：https://edge-impulse.gitbook.io/docs/edge-impulse-studio/eon-tuner

TinyML效能評比 — MLPerf Tiny

MLPerf v1.1 工作場景及效能評量項目

Task	Dataset	Model	Mode	Quality
Keyword Spotting	Google Speech Commands	DS-CNN	Single stream	90% (Top 1)
Visual Wake Words	Visual Wake Words Dataset	MobileNetV1 0.25x	Single stream	80% (Top 1)
Image classification	CIFAR10	ResNet-8	Single stream	85% (Top 1)
Anomaly Detection	ToyADMOS	Deep AutoEncoder	Single Stream	0.85 (AUC)

Scenario	Query Generation	Duration	Samples/query	Latency Constraint	Tail Latency	Performance Metric
Single stream	LoadGen sends next query as soon as SUT completes the previous query	1024 queries and 60 seconds	1	None	90%	90%-ile measured latency

OmniXRI整理製作, 2023/07/17

資料來源：https://omnixri.blogspot.com/2023/07/vmaker-edge-ai-07tinyml-mcu-ai.html

MLPerf Tiny v1.1 評比結果（2023/6）

評測項目	提交者 / 開發板名稱 / 軟體 / 主要處理器@工作頻率 / 推論速度(ms)
視覺喚醒詞 Data : Visual Wake Words Dataset Model : MobileNetv1(0.25x) Accuracy : 80% (Top 1)	Robert Bosch / NUCLEO-G0B1RE / HALE 1.0 / Arm Cortex-M0+ @64MHz / 1869.2
	Plumerai / NUCLEO-U575ZI-Q / Plumerai IE / Arm Cortex-M33 @160MHz / 59.5
	Nuvoton / NUMAKER-M467HJ / ONNC / Arm Cortex-M4F @200MHz / 98.7
	Plumerai / NUCLEO-H7A3ZI-Q / Plumerai IE / Arm Cortex-M7 @280MHz / 26.8
	Fpgaconvnet / ZC706 / fpgaConvNet / Arm Cortex-A9 @650MHz / 0.7
	Syntiant / NDP9120 / Sy
	Kai-Jiang / ZCU106 / --
影像辨識 Data : CIFAR-10 Model : ResNet-V1 Accuracy : 85% (Top 1)	Robert Bosch / NUCLEC
	Plumerai / NUCLEO-U5
	Plumerai / NUCLEO-L4F
	Plumerai / NUCLEO-H7
	Fpgaconvnet / ZC706 /
	Syntiant / NDP9120 / Sy
關鍵字偵測 Data : Google Speech Commands Model : DSCNN Accuracy : 90% (Top 1)	Robert Bosch / NUCLEO-G0B1RE / HALE 1.0 / Arm Cortex-M0+ @64MHz / 679.3
	Plumerai / NUCLEO-U575ZI-Q / Plumerai IE / Arm Cortex-M33 @160MHz / 30
	Nuvoton / NUMAKER-M467HJ / ONNC / Arm Cortex-M4F @200MHz / 46.2
	Plumerai / NUCLEO-H7A3ZI-Q / Plumerai IE / Arm Cortex-M7 @280MHz / 13.9
	Fpgaconvnet / ZedBoard / fpgaConvNet / Arm Cortex-A9 @650MHz / 0.3
	Syntiant / NDP9120 / Syntiant TDK+SDK / HiFi3 + M0 @98.7MHz / 1.5
異常偵測 Data : ToyADMOS Model : FC AutoEncoder Accuracy : 0.85 (AUC)	Robert Bosch / NUCLEO-G0B1RE / HALE 1.0 / Arm Cortex-M0+ @64MHz / 46.5
	Plumerai / NUCLEO-U575ZI-Q / Plumerai IE / Arm Cortex-M33 @160MHz / 3.7
	Plumerai / NUCLEO-L4R5ZI / Plumerai IE / Arm Cortex-M4 @120MHz / 4.1
	Plumerai / NUCLEO-H7A3ZI-Q / Plumerai IE / Arm Cortex-M7 @280MHz / 1.3

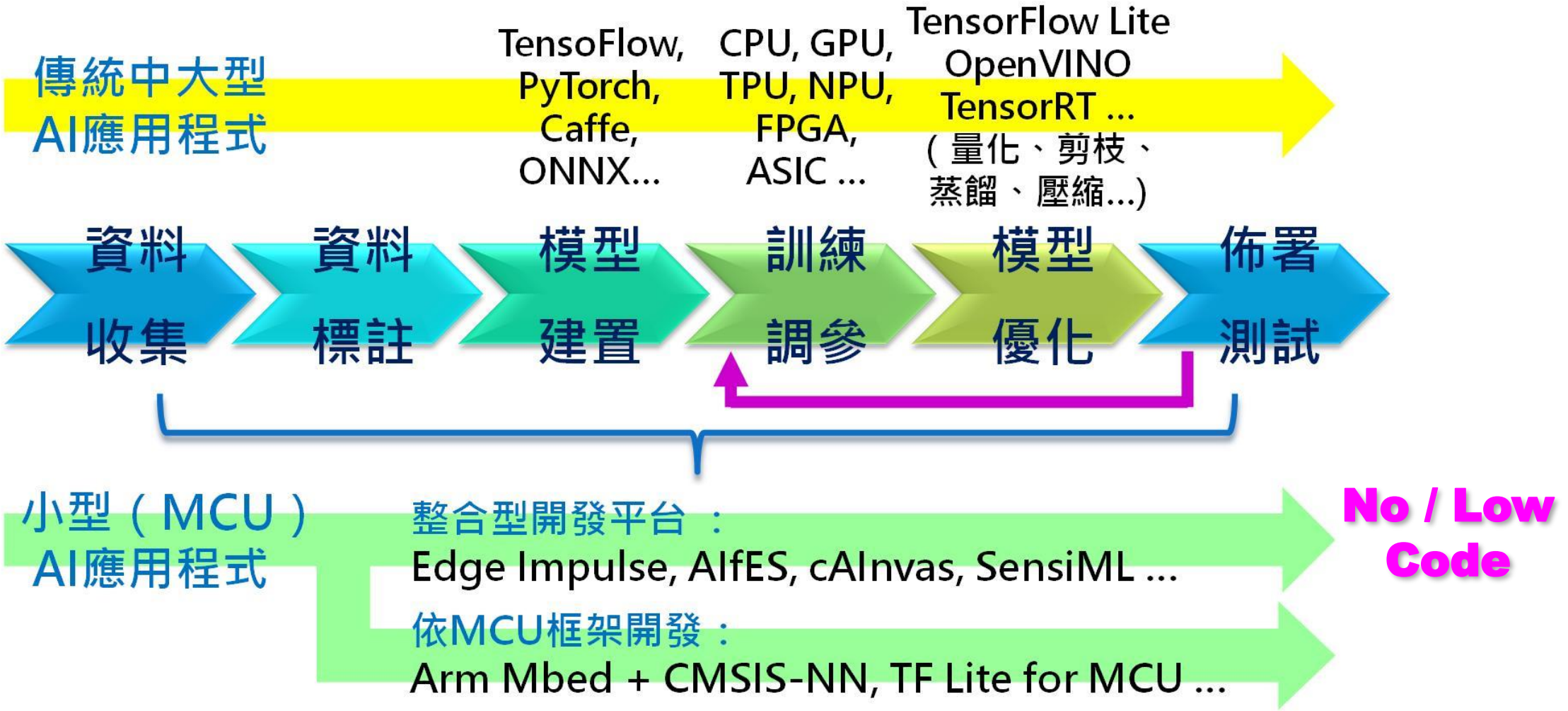
台灣新唐科技(Nuvoton), Nuvoton NUMAKER-M467HJ +臺灣發展軟體科技(Skymizer) ONNC, Arm Cortex-M4F @200MHz。VWW/KWS推論速度最快。

資料來源：https://omnixri.blogspot.com/2023/07/vmaker-edge-ai-07tinyml-mcu-ai.html

7.3. TinyML開發平台

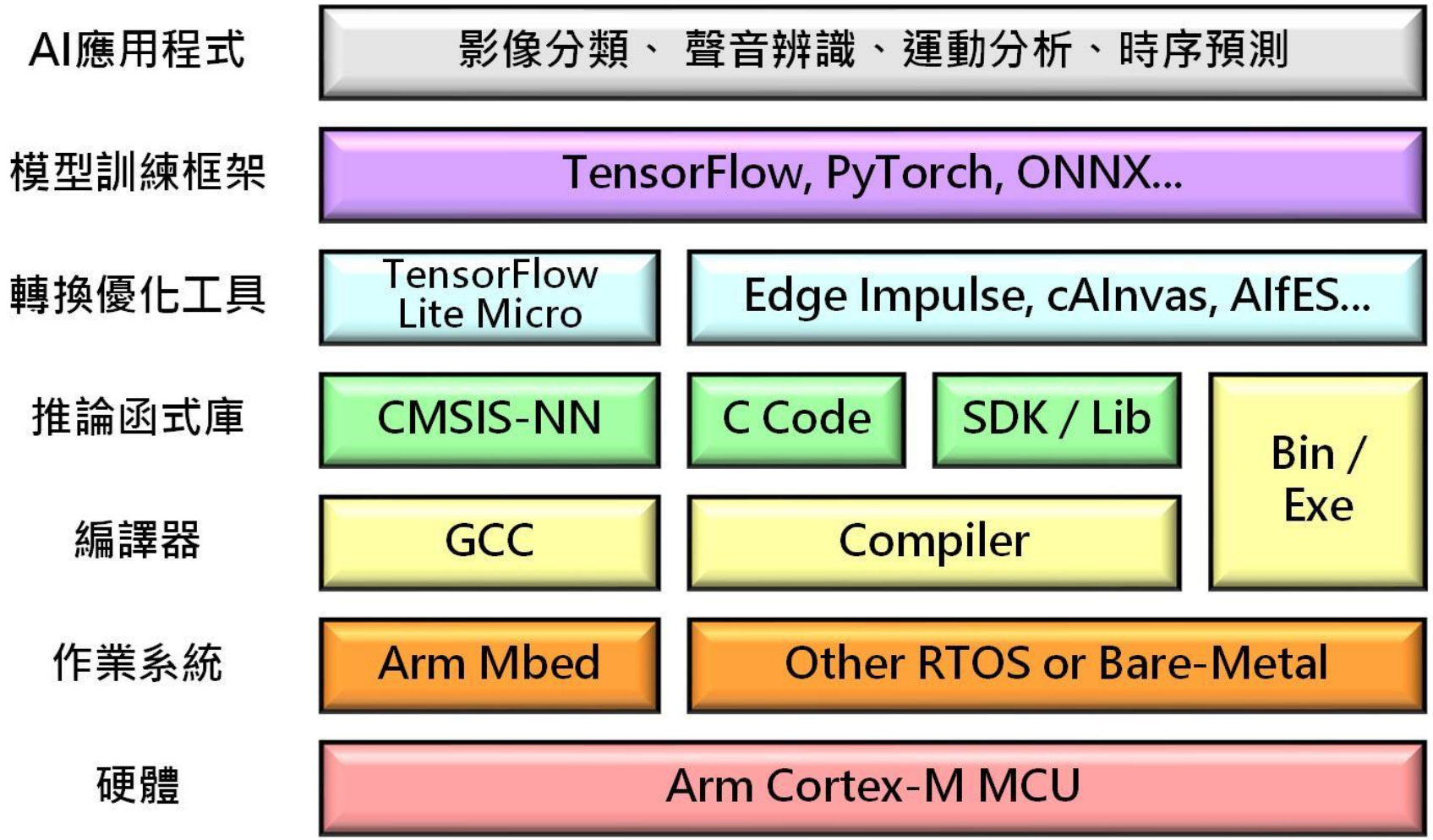
- 開發流程簡介
- 軟體堆疊架構
- 常見開發工具
- 常見開發板

Edge AI & TinyML開發流程



資料來源：https://omnixri.blogspot.com/2021/09/aiottinymlmcu.html

TinyML程式軟體堆疊架構 (Arm Based)



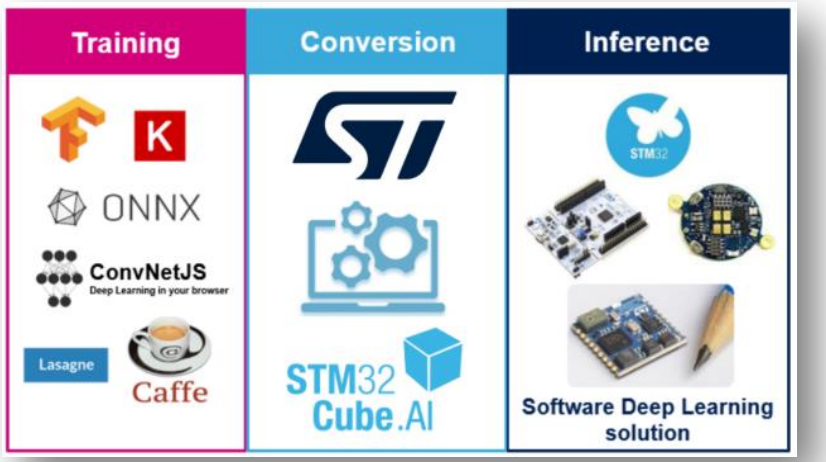
※特殊編譯器會考量模型運行時序不同記憶體需求量的優化

資料來源：https://omnixri.blogspot.com/2022/07/20220731coscupaimcutinyml.html

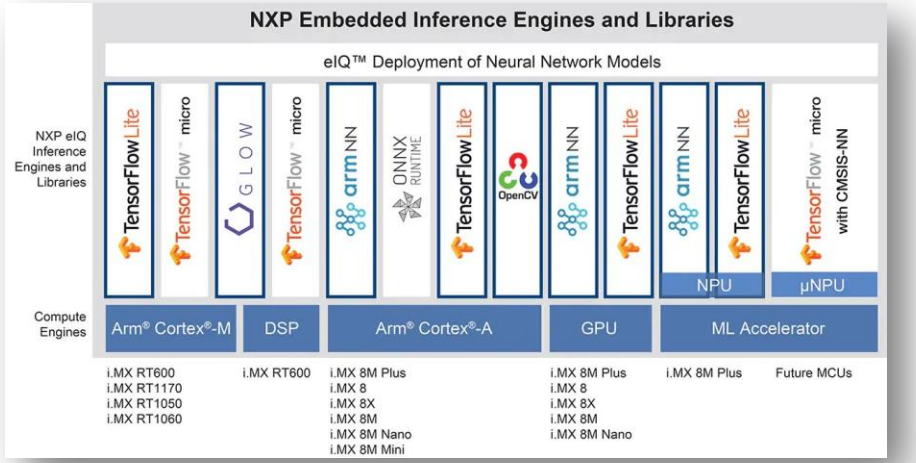
OmniXRI Aug. 2021整理繪製

常見TinyML 開發工具

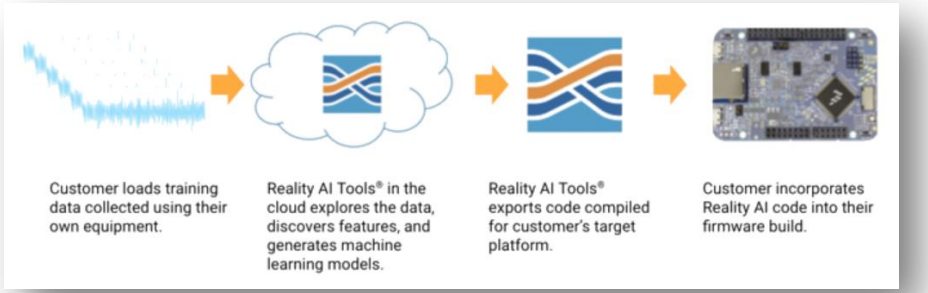
常見TinyML 開發工具 – MCU 專屬式工具



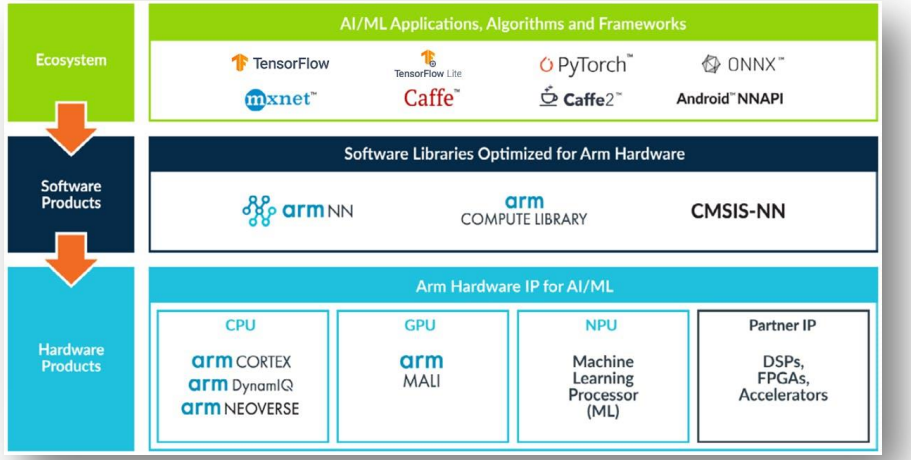
STM32 Cube.AI



NXP eIQ

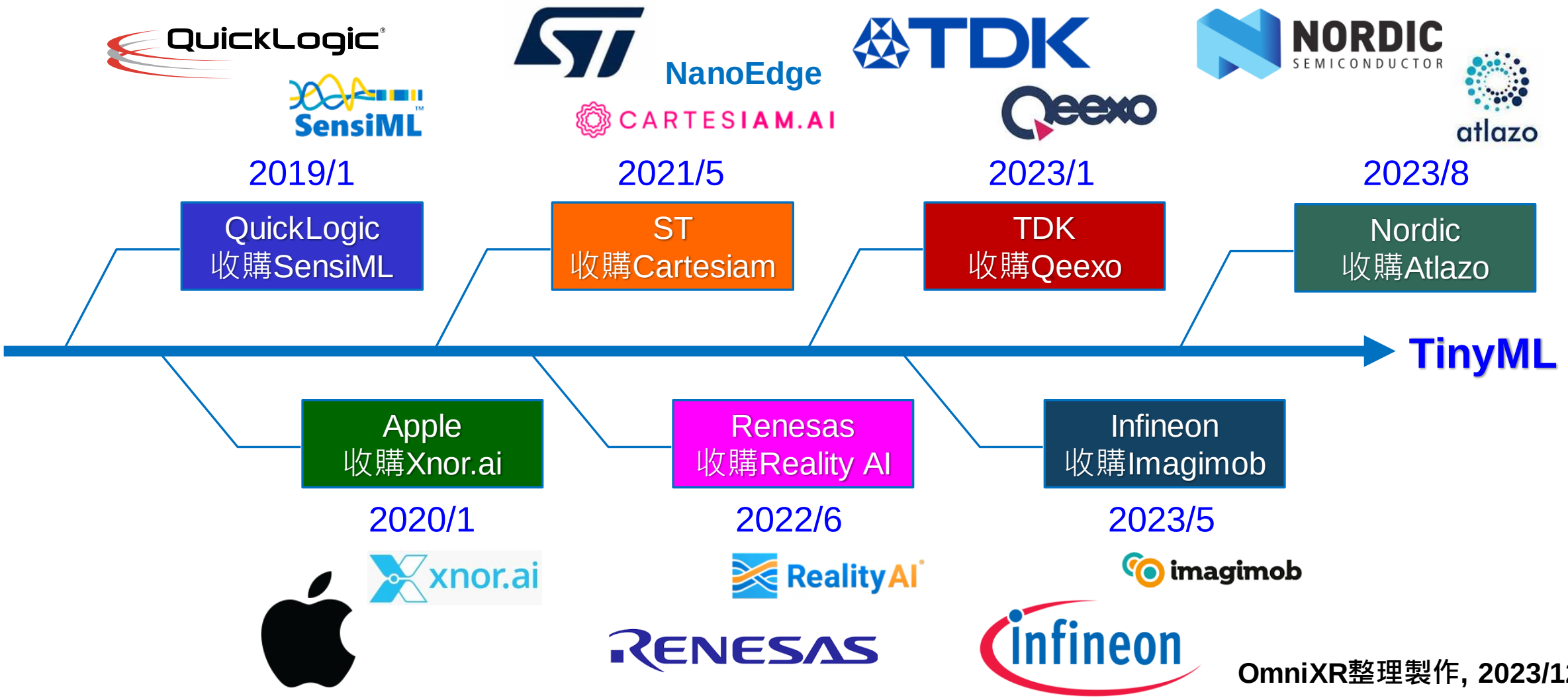


Renesas Reality AI



arm CMSIS-NN

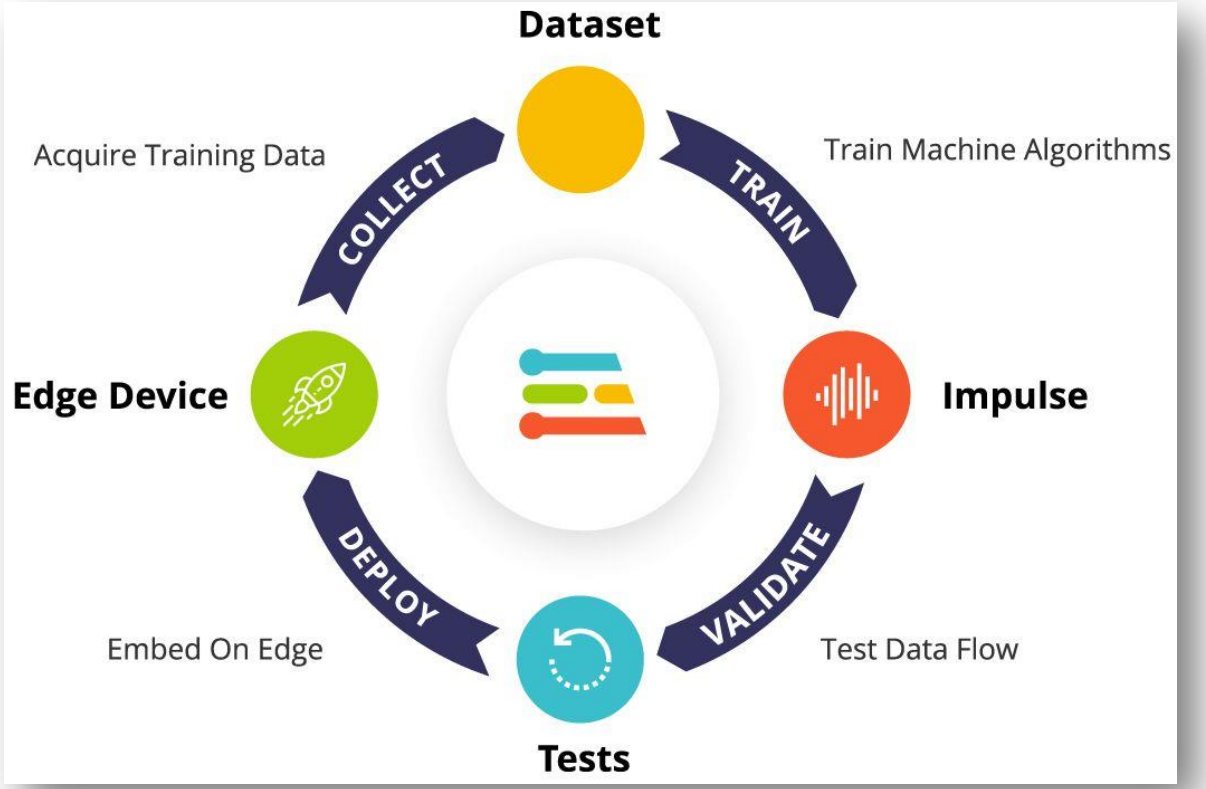
MCU硬體廠收購TinyML軟體開發商



OmniXR整理製作, 2023/12/13

Edge Impulse TinyML 雲端開發平台

EDGE IMPULSE

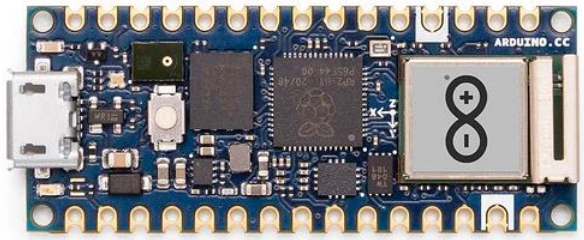


- 成立於2019年
- 超過70名員工
- 共獲US$54M投資（近17億多台幣）
- 免費TinyML雲端開發平台近30萬個專案於平台上運行
- 豐富教學文檔及影片

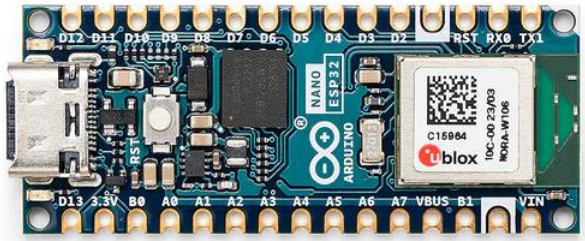
資料來源：https://omnixri.blogspot.com/2021/09/aiottinymlmcu.html

常見TinyML開發板

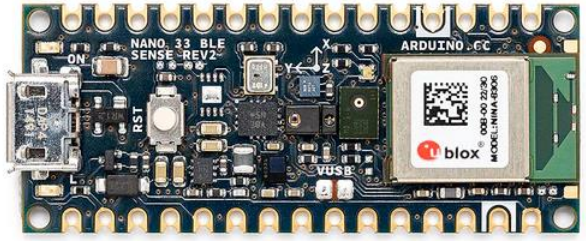
Nano RP2040 Connect



Nano ESP32 (S3)



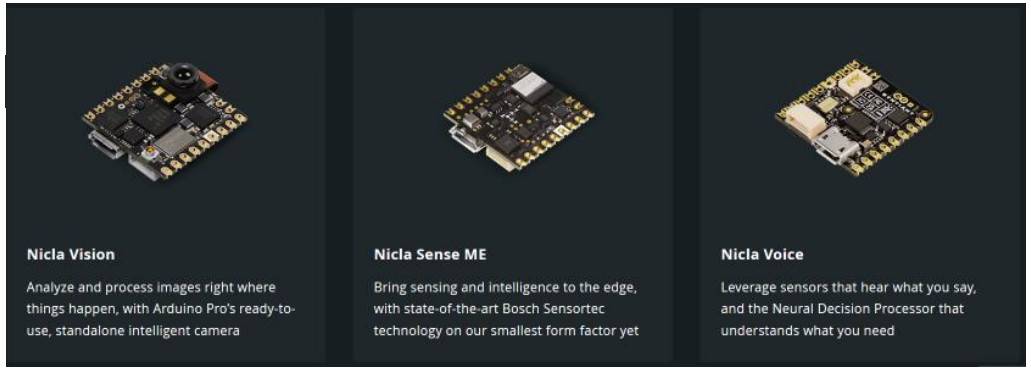
Nano 33 BLE Sense Rev2



Arduino PRO 系列

Protenta Family

Nicla Family



Seeed Xiao系列

RP2040

nRF52840 (Sense)

ESP32C3

ESP32S3 (Sense)

資料來源：https://www.arduino.cc/pro/platform-hardware/

資料來源：https://wiki.seeedstudio.com/cn/SeeedStudio_XIAO_Series_Introduction/

常見TinyML開發板（具攝影機）

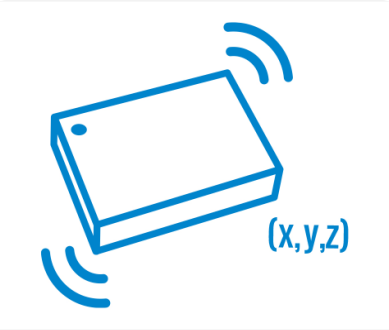
資料來源：https://omnixri.blogspot.com/2022/12/tinymlmcu.html

7.4. TinyML主要應用

- 聲音感測
- 運動感測
- 環境感測
- 影像感測
- 時序預測

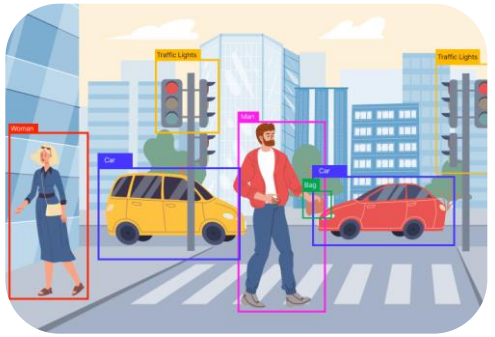
TinyML案例（以感測器分類）

聲音



運動

環境



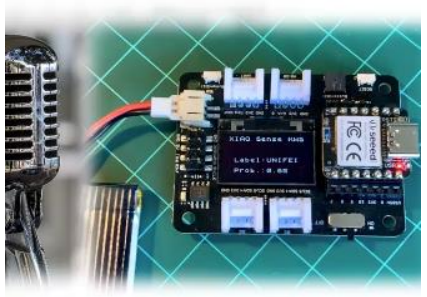
影像

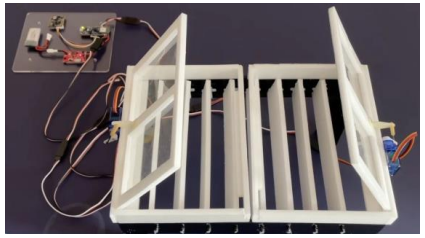
https://hackmd.io/@OmniXRI-Jack/tinyML_30_projects

聲音感測

* **Arduino BLE Sense 偵測語音命令控制LED亮滅**
* **CoreMaker-01 偵測語音命令控制LED亮滅**
* **Xiao nRF52840 Sense 關鍵詞偵測**
* **寶寶哭聲自動搖籃**
* **重金屬搖滾樂偵測器**
* **智慧蓮蓬頭**
* **智慧百葉窗**
* **狗叫停止器**

Edge Impulse 資料取樣及分割流程 OmniXRI整理繪製, 2021/10/5

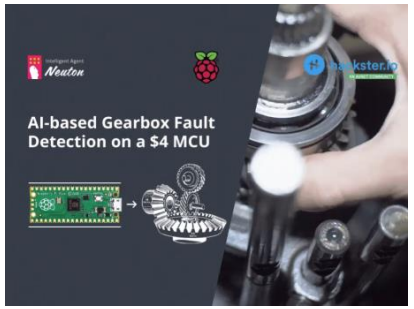


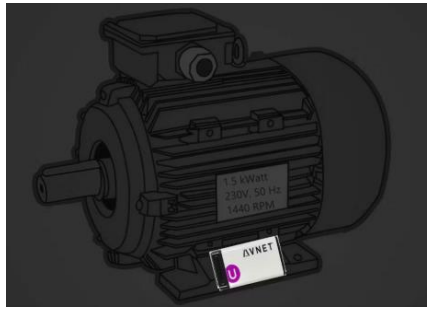


運動感測

* **Arduino Nano 33 BLE Sense - 手勢偵測**
* **手勢切換背包上點矩陣方向燈**
* **手勢遙控器**
* **包裹運送異常震動偵測**
* **齒輪箱異常振動偵測**
* **馬達異常振動偵測**
* **智能桌球拍（虛擬桌球教練）**

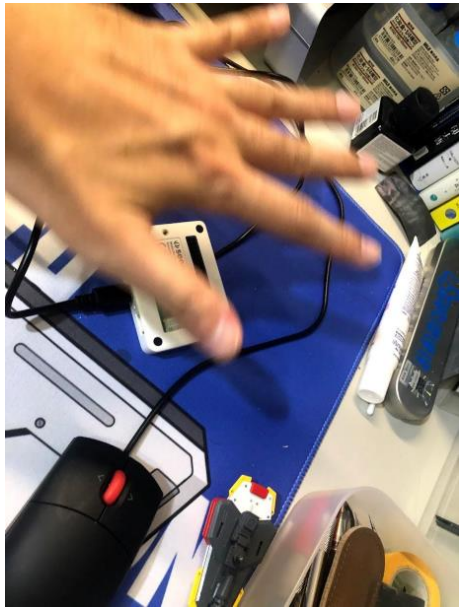
連續運動（手勢）辨識
OK(前後), NG(左右), Pass(畫圈), Idle(閒置)
OmniXRI整理繪製, 2021/10/9





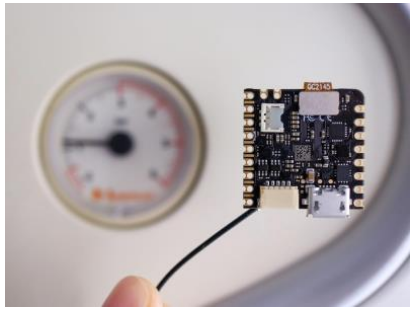
環境感測

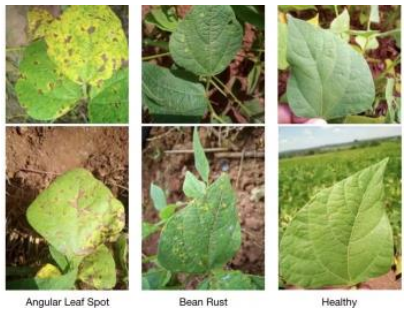
* Wio Terminal簡易氣象預測站
* Thingy:91簡易室內外環境分類器
* 使用Wio Terminal與光感測器偵測手勢
* 使用Wio Terminal及氣味感測器來偵測咖啡氣味
* 使用8bit單晶片進行空氣品質推論



影像感測

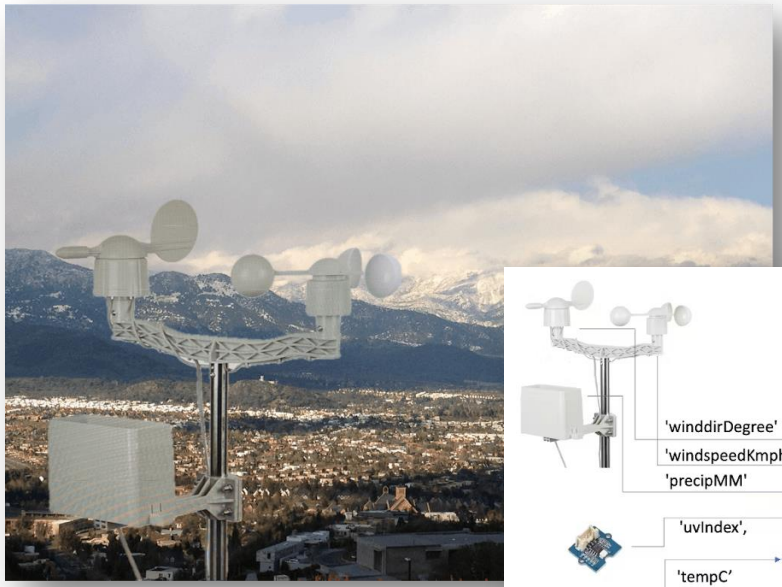
* 指針錶讀值分級
* 看到的是不是馬克杯
* 天空雲朵影像分類
* 使用OV7670進行影像數字分類
* 豆葉病害影像分類
* 電梯內乘客計數
* 影像口罩配戴偵測（物件偵測式）
* 自動追蹤人臉電風扇
* Xiao ESP32S3 Sense物件偵測
* 道路破損偵測

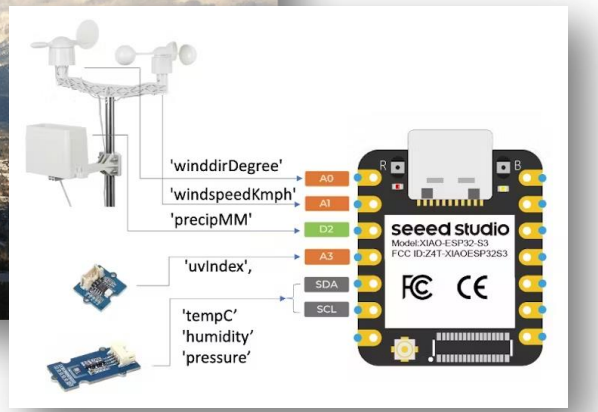


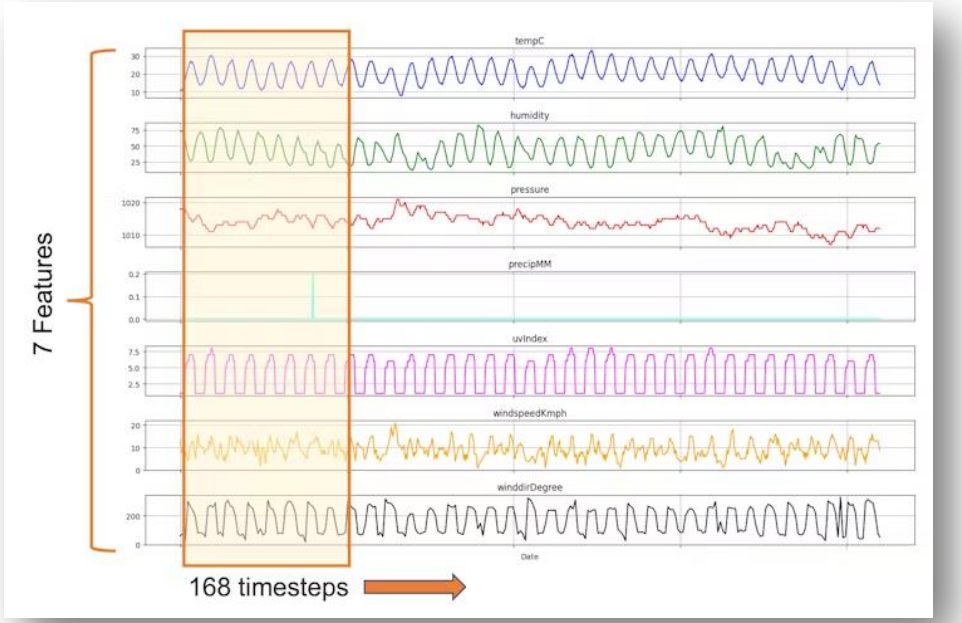


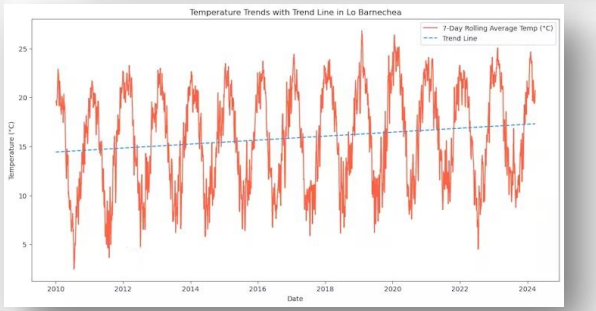
時序預測

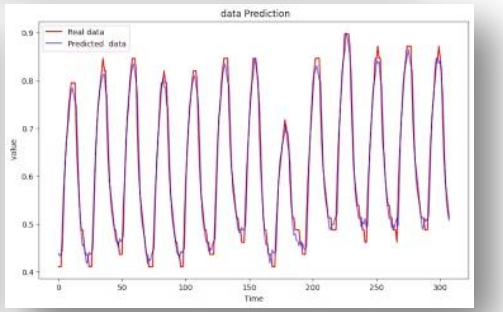
使用單個或多個感測器進行長時間資料收集，將其變化以RNN或LSTM類型時間序列模型進行訓練，以得到未來時間之變化預測。











資料來源：https://www.hackster.io/mjrobot/temperature-prediction-using-a-tinyml-lstm-model-264029

小結

- 透過本章節讓大家學習到嵌入式系統架構及常見單晶片規格及指令集之差異，並認識主要分類及相關供應商。
- 從TinyML基本定義了解其主要應用及使用限制，並了解如何透過硬體加速及模型優化來改善推論效能問題，最後說明如何評比及主要規範MLPerf Tiny。
- 了解TinyML開發流程及基礎軟體堆疊架構，並了解有那些開發工具及平台可以使用，並介紹各種適合用來開發的板子。
- 介紹TinyML的各種應用，包括聲音、運動、環境及影像感測器相關應用，最後還介紹時間序列預測之應用。

參考文獻

- 許哲豪，臺灣科技大學資訊工程系「人工智慧與邊緣運算實務」（2021~2023）

https://omnixri.blogspot.com/p/ntust-edge-ai.html

- 許哲豪，當智慧物聯網(AIoT)遇上微型機器學習(tinyML)是否會成為台灣單晶片(MCU)供應鏈下一個新商機！？

https://omnixri.blogspot.com/2021/09/aiottinymlmcu.html

- 許哲豪， MCU攜手NPU讓tinyML邁向新里程碑

https://omnixri.blogspot.com/2022/10/mcunputinyml.html

- 許哲豪，有了TinyML加持MCU也能開始玩電腦視覺了

https://omnixri.blogspot.com/2022/12/tinymlmcu.html

- 許哲豪，TinyML應用大全（30組案例分享）

https://hackmd.io/@OmniXRI-Jack/tinyML_30_projects

延伸閱讀

- Wevolver, 2024 State of Edge AI Report

https://www.wevolver.com/article/2024-state-of-edge-ai-report/introduction

- 許哲豪，【vMaker Edge AI專欄 #14】 從CES 2024 看Edge AI及TinyML最新發展趨勢

https://omnixri.blogspot.com/2024/02/vmaker-edge-ai-14-ces-2024-edge-aitinyml.html

- 許哲豪，【vMaker Edge AI專欄 #13】 誰說單晶片沒有神經網路加速器NPU就不能玩微型AI應用？

https://omnixri.blogspot.com/2024/01/vmaker-edge-ai-13-npuai.html

- 許哲豪，【vMaker Edge AI專欄 #07】TinyML (MCU AI) 運行效能誰說了算？

https://omnixri.blogspot.com/2023/07/vmaker-edge-ai-07tinyml-mcu-ai.html

- 許哲豪， 【心得筆記】Edge Impulse EON Tuner AutoML工具介紹

https://omnixri.blogspot.com/2022/05/edge-impulse-eon-tuner-automl.html

- 許哲豪，TinyML相關學術論文

https://hackmd.io/@OmniXRI-Jack/tinyML_papers

歐尼克斯實境互動工作室
(OmniXRI Studio)
許哲豪 (Jack Hsu)

Facebook : Jack Omnixri
FB社團： Edge AI Taiwan邊緣智能交流區
電子信箱：omnixri@gmail.com
部落格：https://omnixri.blogspot.tw
開　源：https://github.com/OmniXRI

YOUTUBE 直播： https://www.youtube.com/@omnixri1784/streams