



面向新手的機器學習：現在無需攻讀資料科學學位即可構建邊緣 AI

探索嵌入式工程師如何在不斷壯大的生態系統支援下開發 AI 應用程式以輕鬆實現 ML 和 AI。

Microchip Technology Inc.
觸控界面業務部
行銷副總監
Yann LeFaou

協助嵌入式工程師進行 AI 開發

開發人工智能（AI）應用程式過去一直是資料科學家和專業工程師的領地，他們精通如何構建高效的機器學習（ML）AI 算法和深度學習神經網路。儘管備受矚目的生成式 AI 應用程式（如 ChatGPT）也同樣如此，但隨著在設備級部署 AI 的需求不斷增長，即使是幾乎甚至完全沒有 AI 相關經驗的嵌入式工程師也得以參與到構建此類機器學習應用程式的任務當中。

值得欣慰的是，硬體製造商和軟體開發商已經認識到這一趨勢，正在攜手打造一個不斷壯大的生態系統來提供各種設備、工具和支援。在這一生態系統的支援下，實現用於收集和組織資料、訓練神經網路和優化邊緣推理的 AI 和 ML 算法變得比以往任何時候都更加容易。

AI 開發方法的轉變

從歷史上看，大多數機器學習實現都依賴強大的計算機或雲端伺服器，不僅需要佔用大量資源，還需要用複雜的語言進行編寫。模型由具有相關資質的資料科學家開發，他們花費大量時間來開發能夠理解海量資料中的各種模式和相關性的算法，從而構建預測性 AI 模型。

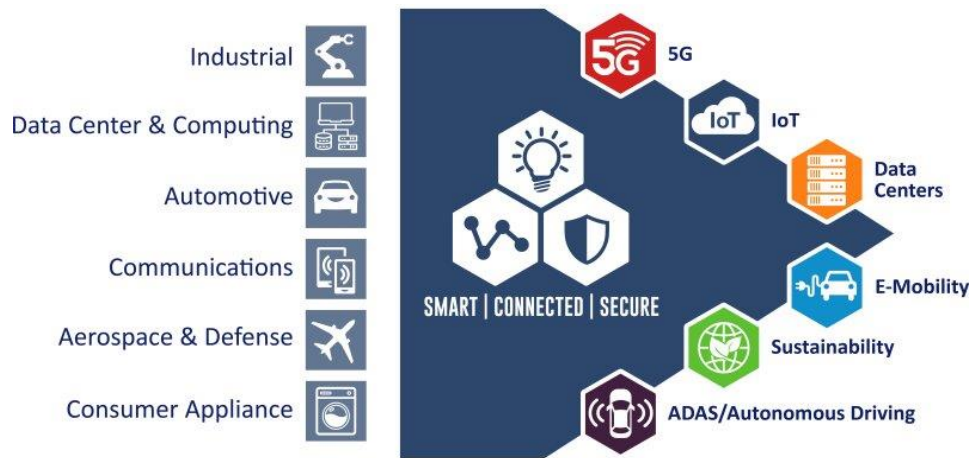


圖 1: AI/ML 的潛在應用

但是，隨著越來越多的應用（見圖 1）想要借力 AI，目前出現了兩大趨勢。

首先是將 AI 嵌入到智能互聯設備中，使 AI 脫離集中式伺服器。這種所謂的邊緣 AI 使得小型電池供電類設備能夠檢測複雜動作、識別聲音、分類圖像、查找感測器資料中的異常，以及快速處理現實世界中的大量其他輸入，而這一切都不需要連接到遠程伺服器。

其次是越來越多地需要嵌入式工程師來提供基於 AI 的解決方案，但他們中的大多數人幾乎甚至完全沒有 ML 或 AI 相關經驗。

支援邊緣 AI 開發

促進邊緣 AI 呈指數級加速增長的因素之一是「微型機器學習」（簡稱 TinyML）概念的誕生。根據 tinyML 基金會的定義，這是「快速發展的機器學習技術和應用領域，包括能夠以極低的功耗（通常在 mW 範圍內及以下）執行設備上感測器資料分析（視覺、音頻、IMU 和生物醫學等）的硬體、演算法和軟體，因此可以實現各種始終在線的應用，並以電池供電類設備為目標。」

TinyML 運動獲得了廣泛呼應並促進了相關工具和支援呈現爆炸式增長，這使得嵌入式設計工程師可以更加輕鬆地完成 ML 開發工作，不再必須依賴具備神經網路和複雜編程語言相關知識和經驗的資料科學家或專業工程師。

處理硬體

當然，邊緣 AI 應用程式依舊必須在處理器上運行，儘管顯然不是當今資料中心中的高功率、高效能計算（HPC）處理器。但事實上，成功實現邊緣 AI 的主要挑戰之一始終是如何在超低功耗預算和緊湊的尺寸下提供必要的處理能力。隨著模型的複雜度不斷增大，邊緣 AI 設備將面臨精度與資源需求之間的艱難權衡，這意味著連接的設備必須配備更為昂貴的處理晶片，但這又會消耗更多的電能。這剛好解釋了為何迄今為止，大多數部署在智能設備上的機器學習模型都要求小一點，並且 AI 任務僅限於解決簡單的問題。

幸運的是，最新的微控制器（MCU）、微處理器（MPU）和現場可程式化邏輯閘陣列（FPGA）技術即將解決這一難題，這些技術可以幫助開發人員以更低的功耗、更小的尺寸來構建更加複雜的邊緣 AI 應用。[整合專用攝相機介面的低功耗 32 位 MPU](#) 等技術進步使得開發人員能夠設計具有更精確深度感知能力的低功耗立體視覺應用，並且還允許在功耗與效能之間進行微調。

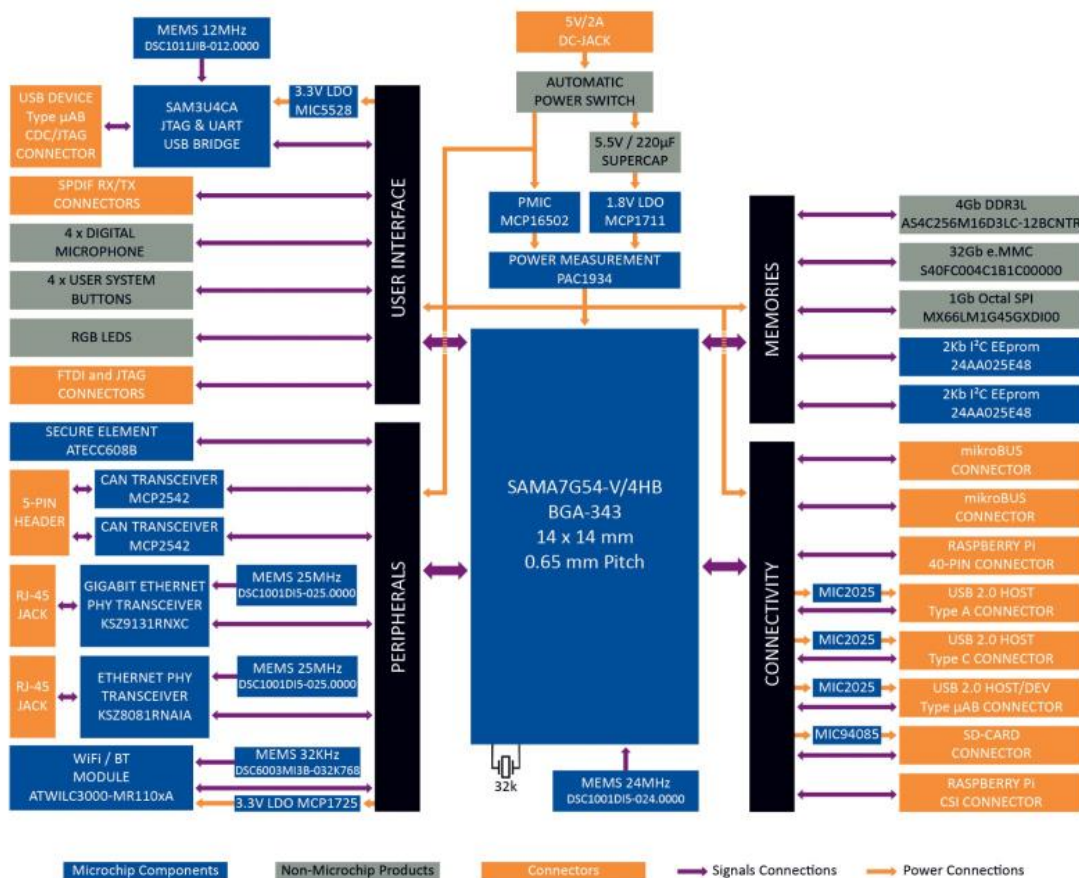


圖 2：具有嵌入式攝相機介面的低功耗、高效能邊緣 AI MPU

此外，如果將最新的半導體技術與能夠使 AI/ML 在邊緣設備上更高效地運行的專用演算法相結合，則意味著現在可以提供基於最低功耗 8 位元和 16 位元微控制器的 ML 解決方案。

無論是基於先進的 FPGA 還是成本最低的 8 位元 MCU 進行開發，都有大量額外的支援工具可幫助工程師簡化設計和測試、加速原型設計並縮短邊緣 AI 解決方案的整體上市時間。其中包括參考設計、用於編程節能神經網路的加速器軟體開發工具套件 (SDK) (無需任何 FPGA 設計經驗)，以及「開箱即用」的 ML 開發和評估板 (將 MCU、陀螺儀、加速計、運動檢測器和各種其他感測技術整合在一個單元中)。

展望未來

根據 GrandViewResearch, Inc. 分析師最新發佈的報告 1，全球邊緣 AI 市場預計從 2023 年到 2030 年將實現 21% 的複合年增長率 (CAGR)，達到 664.7 億美元的規模。報告指出，促進這一增長的主要因素是對基於 IoT 的邊緣計算服務的需求不斷增長，智能監控和安全、智能農業和工業製造等應用都需要依賴這些服務來分析大量感測器資料，進而實現營運決策自動化。

為了支援這一增長，我們需要進一步進行相關軟硬體開發，以簡化和加速邊緣 AI 應用程式的開發過程，讓每一位嵌入式工程師都能夠進行 AI 和 ML 設計。隨著開源資料集的不斷擴充，將不再需要花費大量的時間收集與特定應用相關的資料，進而進一步簡化開發過程中面臨的挑戰。

最後，正如生成式 AI 應用程式的興起所表明的那樣，AI 本身正在成為一種實用的工具，可輔助新手編寫複雜的應用程式，並且 ML 訓練與直接編碼相比可以縮短開發時間，減少重新設計並提高每位工程師的產能。

參考資料

1. *Edge AI Market Size, Share & Trends Analysis Report By Component (Hardware, Software, Edge Cloud Infrastructure, Services), By End-use Industry, By Region, And Segment Forecasts, 2023 - 2030: Grand View Research - April 2023.*