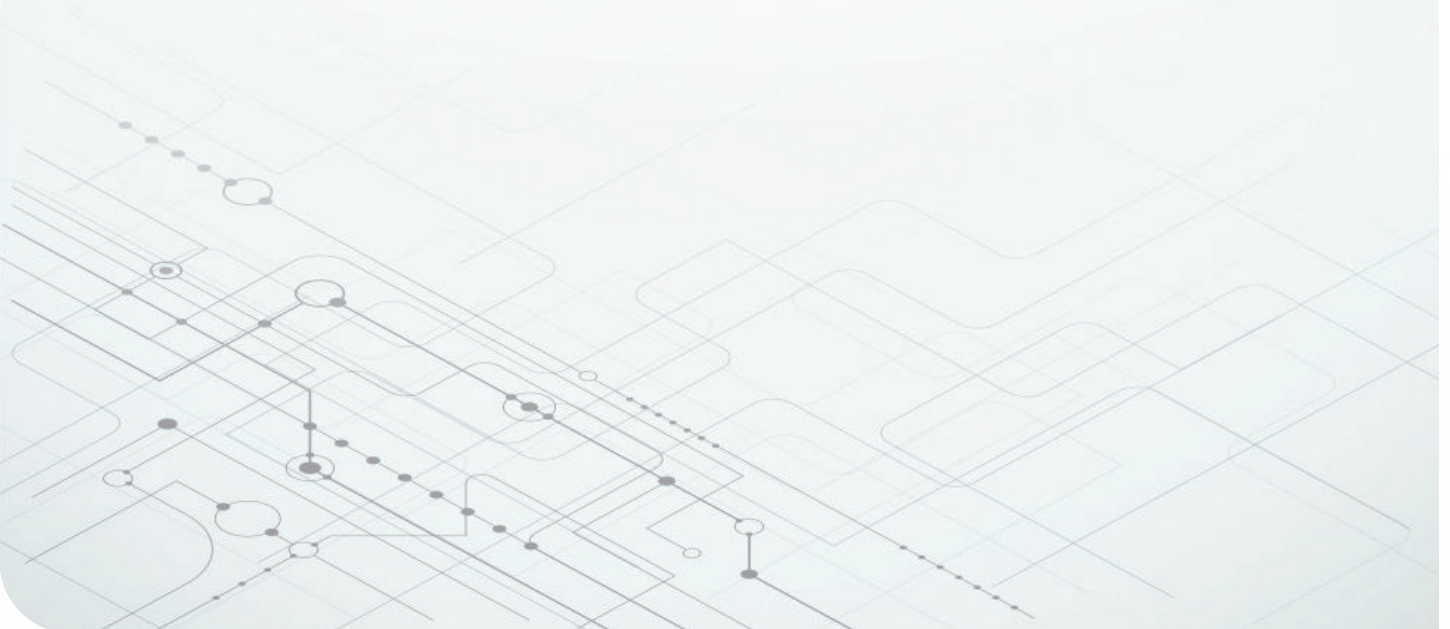


以嵌入式處理器提升邊緣智慧



Alec May
Systems Manager
Jacinto™ High-Performance Computing Processors



摘要

1 簡介

本白皮書探討了邊緣人工智慧 (AI) 的目的與優勢，以及嵌入式處理器與軟體的進步如何讓 AI 比以往更輕鬆地在各種應用中實作。

2 邊緣 AI 有哪些優點？

了解邊緣 AI 的優點，以及其在電子產品中實現本地推論的能力。

3 AI 如何走向邊緣

了解嵌入式硬體和軟體工程師所面臨的挑戰，以及 TI 如何因應這些挑戰。

4 邊緣 AI 的可擴展性

探索 TI 為因應可擴展性和可重複使用性挑戰而提供的硬體和軟體。

簡介

廣泛可用的雲端式 AI 解決方案具備可存取性和使用便利性，讓幾乎所有人都能更輕鬆地使用專為 AI 設計的模型和工具。

然而，並非所有 AI 創新都會在雲端進行。隨著嵌入式處理器設計的技術進步，AI 功能現在導入筆記型電腦、手機等消費產品以及其他電池供電應用：視訊門鈴、汽車系統的視覺處理，以及能源基礎架構與工業系統中的馬達。

邊緣 AI 是在資料來源附近本機執行 AI 模型的能力，可提升這些產品的回應性、效率、可靠性和安全性。嵌入式處理器讓這種雲端到邊緣的轉型成為可能，整合了用於數位訊號處理 (DSP) 的專用核心等元件，並由易於使用的 GUI 架構工具支援，可將 AI 導入邊緣所需的時間和專業知識降到最低。

本白皮書探討邊緣 AI 的演進與優勢，以及支援邊緣 AI 的軟硬體之進步。

何謂 AI？

今天大多數人想到 AI 時，通常會聯想到文字和圖像生成器。但即使是最簡單的演算法，在技術上也是文字意義上 AI 的範例。

AI 的廣泛性及其多重使用案例催生了多個子領域，包括機器學習和深度學習，如 **圖 1** 所示。

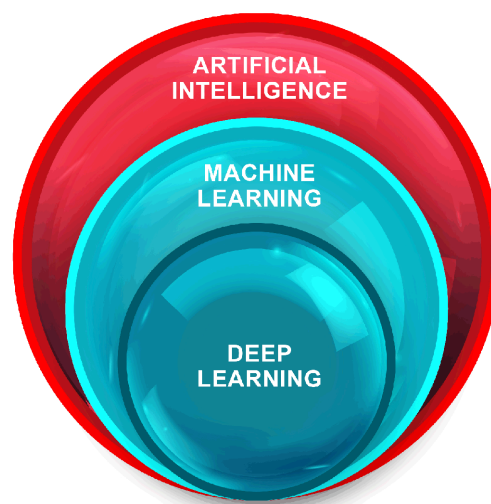


圖 1. 不同 AI 子領域之間的關係。

用於嵌入式應用的大多數 AI 為機器學習，這是機器和演算法「學習」如何在資料中解決問題的子領域；例如，車輛透過分析通用模式的影像資料來辨識行人與障礙物。機器學習模型可從訓練資料中學習，這些資料可能會標記地面真實資訊（即經驗證的準確數據），以更好地識別需要學習的模式。此訓練程序可讓機器學習模型能夠辨識資料中的模式，並可用於進行未來的推論。

在機器學習領域中，雖然深度學習需要大量的運算資源，但因其能準確地解決高度複雜的問題，已成為最受歡迎的實作方式之一。深度學習使用多層神經網路，這是受人腦神經元啟發的資料模型。神經網路可讓開發人員解決模式過於複雜、無法辨別或無法編寫自訂規則的問題。

邊緣 AI 有哪些優點？

AI 及其子領域通常可在雲端或本機伺服器上執行處理。雲端式 AI 在過去較為普遍，因為要在大型伺服器之外執行有影響力 AI 所需的運算能力並不是容易達成的條件。然而，

隨著嵌入式處理器的運算能力和電源效率不斷提升，邊緣 AI 也日益普及。

圖 2 顯示邊緣 AI 和雲端 AI 在接收和處理資料以及與雲端式資源互動的方式上有何不同。

邊緣 AI 通常會在開發期間使用雲端或桌上型電腦資源進行模型訓練。將模型部署至嵌入式裝置後，即可在本機獨立地對新資料進行模型推論和決策。

直到最近，最有意義的 AI 範例所需的處理能力均超出一般消費性電子產品所能提供的能力。這意味著機器學習模型通常是在雲端式資源上進行訓練和實作。雖然雲端式架構實作透過最大限度地減少硬體投資而提供了便利性，但也限制了 AI 的採用。雲端式 AI 實作將無法在沒有雲端存取（即網路連線）的任何應用程式中進行。此外，與雲端式 AI 相比，邊緣 AI 還可以提高保全性、安全性和回應能力。

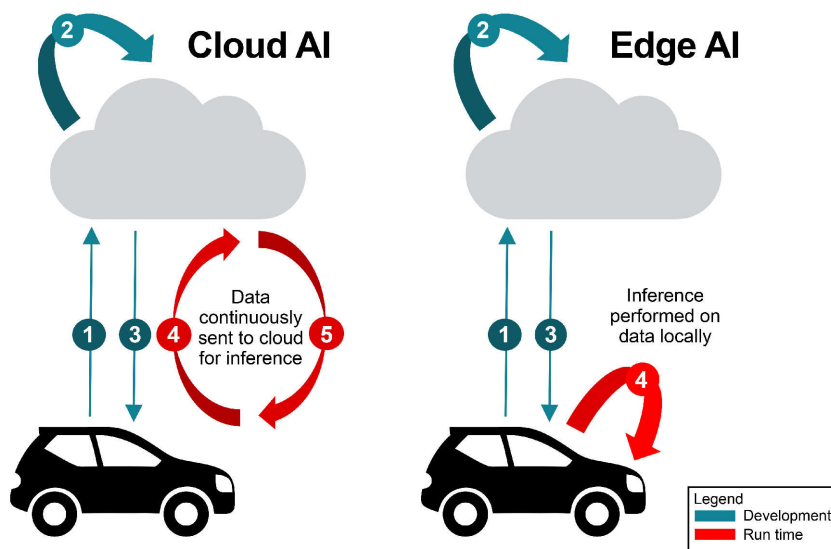


圖 2. 雲端式 AI 與邊緣 AI 的比較。

隨著半導體的進步和 AI 工具鏈的改善，現在已可將 AI 解決方案直接導入嵌入式處理器和微控制器 (MCU)，進一步將 AI 推展至邊緣。將 AI 帶入邊緣意味著運算和 AI 推論要在提供資料的感測器附近執行，這點很重要，因為電子裝置收集的感測器資料量正不斷增加。不斷成長的資料量使得純雲端資源變得不太實用，因為從雲端傳輸大量資料可能使成本高昂且過程複雜，並且會出現單點故障。

在網路邊緣執行 AI 模型通常可減少根據感測器資料進行推論和決策的延遲，例如車輛中用於碰撞偵測的攝影機感測器。運用邊緣 AI 功能，車輛可以更快速地進行推論，即時回應刺激，無需等待來自雲端的推論。此本機推論透過實體 AI 轉化為行動（如圖 3 所示），這個術語指感測、解讀並以實際世界動作回應的系統，例如機器人在工廠車間移動箱子，或是車輛自動應用煞車。



圖 3. 人形機器人中邊緣 AI 和實體 AI 的簡化比較

這些系統可使用邊緣 AI 將進階感知與機械致動結合，讓機器能與人類安全協作地運作。

與雲端式 AI 相比，邊緣 AI 具有其他多項優勢，包括減少對網路連線的依賴。邊緣 AI 可在無法存取雲端的應用中運作，並將網路中斷造成的潛在停機時間降至最低。此外，由於雲端式 AI 需要網路連線，因此可能會產生經常性的存取服務費，這在設計消費產品時可能是一個具有挑戰性的商業模式。

AI 如何走向邊緣

嵌入式處理器的處理與功耗限制，以及內部高階編程專業知識和資源，都限制了邊緣 AI 的普及性。可滿足 AI 運算性能需求的嵌入式裝置通常體積過大、功耗過高，且會產生過多熱能。

出現了專業硬體解決方案，可以更有效地加速邊緣 AI 所需的運算操作，但有幾項取舍限制了其在邊緣應用中的廣泛採用。圖形處理單元 (GPU)、現場可程式邏輯閘陣列 (FPGA) 和特定應用積體電路 (ASIC) 等專用硬體解決方案已獲得令人印象深刻的性能。但這些解決方案通常會受到高功耗（特別是在 GPU 和 FPGA 的情況下）或有限靈活性（就 ASIC 而言）的限制。

整合式神經處理單元 (NPU) 已成為可將 AI 功能直接導入嵌入式系統的解決方案。與通用處理器不同，NPU 專為執行構成現代神經網路骨幹的矩陣乘法、卷積運算和激活函數而打造。透過從主 CPU 卸載計算密集型任務，NPU 顯著提高推論速度，同時降低功耗，這是邊緣部署的兩個基本要求。

為了解這些元件的影響，我們來看看德州儀器專為邊緣 AI 市場不同領域設計的兩個 NPU：用於高性能應用的 C7™ NPU 和用於低功耗、可擴充裝置的 TinyEngine™ NPU。

C7 NPU

C7 NPU 是一款高性能、高電源效率的 AI 加速器，整合至 TDA54-Q1 和 TDA4VE-Q1 系統單晶片 (SoC)。TI 微處理器和 SoC 歷經數代產品迭代，均集成了 C7 NPU 的不同版本，以滿足視覺應用的計算需求。此 NPU 源自 TI 悠久的 DSP 歷史，使其能夠有效地驅動 AI 解決方案，而不犧牲成本或功耗。

C7 NPU 也讓 TI 處理器可處理多個並行 AI 工作負載，這對於需要同時處理攝影機、雷達、光達和其他感測器資料的系統來說，是一項重要功能。

TDA54-Q1 使用 C7 NPU，在先進駕駛輔助系統、車載資訊娛樂系統和機器人中啟用邊緣 AI。

TinyEngine NPU

TinyEngine NPU（如圖 4 中的簡化方塊圖所示）是專為 MCU 設計的專用硬體加速器，可最佳化深度學習推論作業，以在處理資源受限裝置（包括電池供電產品）中的 AI 工作負載時，減少延遲和功耗。

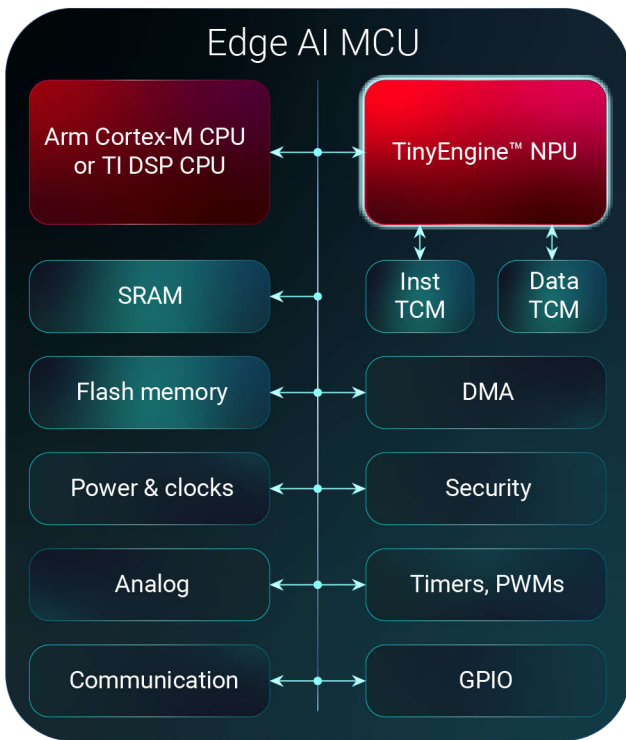


圖 4. 整合 TinyEngine NPU 的 TI 邊緣 AI MCU 簡化方塊圖

此 NPU 會與執行應用程式碼的主要 CPU 平行執行機器學習演算法。具有 TinyEngine NPU 的 MCU 可以運行模型，與沒有加速器的類似 MCU 相比，每次推論的延遲最多可降低 90 倍，能耗最多可降低 120 倍。

TMS320F28P550SJ C2000™ MCU 使用 TinyEngine NPU 進行馬達軸承和太陽能電弧故障偵測，進而由主要 CPU 處理即時馬達控制。**AM13E23019** 結合 TinyEngine NPU 與先進的即時控制架構（最多可包含四個馬達），可進行電器、機器人和工業系統中的適應性控制和預測性維護。**MSPM0G5187** Arm® Cortex®-M0+ 架構的 MCU 使用專用的 TinyEngine NPU，以獨立於主要 CPU 執行深度神經網路模型，進而在穿戴式健康監控器、電器和工業系統中實現邊緣 AI 功能，以進行預測性馬達維護。

邊緣 AI 軟體創新

隨著嵌入式裝置中高效率 AI 運算的硬體進步，開放原始碼社群和半導體製造商也讓測試和部署 AI 模型變得更加容易，只需具備最少的編程專業知識。讓 AI 更易於使用（在某些情況下是以 GUI 為基礎），有助於減少投資額外資源或訓練的需求。

對於較熟悉 AI 模型的設計人員而言，PyTorch 和 TensorFlow 等開放原始碼工具可以為其自訂資料集訓練模型架構，並可將模型匯出為方便嵌入式的格式，例如 ONNX 或 LiteRT（前稱 TensorFlow Lite）。接著，模型會在裝置上與等效的開放原始碼執行階段軟體一起執行。

這些開放原始碼工具透過抽象化嵌入式平台的細節，協助邊緣 AI 開發，並提供允許存取硬體加速後端（又稱為代理）的一致介面。這些後端可提供進一步的配置，讓設計人員能夠更充分地控制模型委派給硬體加速器的作業。

TI **CCStudio™ Edge AI Studio** 是一款 Web 架構工具系列，其使用遠端 TI 硬體和 GUI，簡化並加速 TI 嵌入式裝置上的邊緣 AI 應用開發。這些工具包括模型編輯器、模型分析器、模型選擇工具和模型製作器，可協助設計人員快速評估模型及其性能，而無需實際連接至評估板。

邊緣 AI 的可擴展性

使用嵌入式微控制器或微處理器開發產品時，請務必考慮產品如何隨時間演進與擴充。工程師不想花費數個月的時間在一個微處理器上開發解決方案，然後在將產品更新為更高性能的處理器時又必須從頭開始。

打造這些嵌入式裝置的半導體製造商必須開發在功能、性能和成本方面具有可擴展性的產品組合。此方法有助於確保其各種 AI 嵌入式處理器之間有無縫遷移策略，以讓開發人員盡可能簡單地在不同裝置上重複使用其工作。

邊緣 AI 也不例外。例如，製造居家機器人的設計師可能想要同時生產配備三部環景攝影機的高階版本，以及只配備單一前置攝影機的入門級版本。可擴充的邊緣 AI 加速裝置產品組合可將軟體從高階機型移植至入門級機型，將生產這兩種產品所需的資源量降至最低。可擴展性也允許開發人員隨著產品的發展，將研發投資從一個平台轉移至下一個平台。

結論

儘管邊緣 AI 目前仍是相對較新的技術，但其重塑我們日常生活的潛力正逐漸成為人們關注的焦點，特別是為幾乎所有應用帶來更高回應性和更高性能的能力。隨著低功率、經濟高效的嵌入式處理器以及直覺式軟體和模型訓練工具的進步，對於任何經驗程度的設計師而言，使用 AI 技術從未像今日這般容易。我們可以預期，這種情況將能持續到每一代邊緣 AI 裝置，以及管理我們與之互動和依賴的電子設備內的運作和資料收集的關鍵元件 (例如用於感測、電力傳輸和連接的半導體)。

其他資源

- 探索 [TI 的邊緣 AI 處理產品組合和設計資源](#)。
- 透過下列技術文章，了解 TI 的邊緣 AI 加速 MCU
 - [邊緣 AI 加速 Arm® Cortex®-M0+ MCU 如何為電子產品注入更強大的智慧](#)
 - [在工業自動化和家用電器設計中實現支援邊緣 AI 的馬達控制](#)
- 閱讀產品概覽 [《TI 的 TinyEngine™ NPU 為更多嵌入式系統解鎖邊緣 AI 加速功能》](#)，了解 TinyEngine NPU 及其對邊緣 AI 加速嵌入式設計的優勢。

重要聲明：本文所述德州儀器及其子公司相關產品與服務經根據 TI 標準銷售條款及條件。建議客戶在開出訂單前先取得 TI 產品及服務的最新完整資訊。TI 不負責應用協助、客戶的應用或產品設計、軟體效能或侵害專利等問題。其他任何公司產品或服務的相關發佈資訊不構成 TI 認可、保證或同意等表示。

所有商標均為其各自所有者的財產。

IMPORTANT NOTICE AND DISCLAIMER

TI PROVIDES TECHNICAL AND RELIABILITY DATA (INCLUDING DATASHEETS), DESIGN RESOURCES (INCLUDING REFERENCE DESIGNS), APPLICATION OR OTHER DESIGN ADVICE, WEB TOOLS, SAFETY INFORMATION, AND OTHER RESOURCES "AS IS" AND WITH ALL FAULTS, AND DISCLAIMS ALL WARRANTIES, EXPRESS AND IMPLIED, INCLUDING WITHOUT LIMITATION ANY IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT OF THIRD PARTY INTELLECTUAL PROPERTY RIGHTS.

These resources are intended for skilled developers designing with TI products. You are solely responsible for (1) selecting the appropriate TI products for your application, (2) designing, validating and testing your application, and (3) ensuring your application meets applicable standards, and any other safety, security, regulatory or other requirements.

These resources are subject to change without notice. TI grants you permission to use these resources only for development of an application that uses the TI products described in the resource. Other reproduction and display of these resources is prohibited. No license is granted to any other TI intellectual property right or to any third party intellectual property right. TI disclaims responsibility for, and you fully indemnify TI and its representatives against any claims, damages, costs, losses, and liabilities arising out of your use of these resources.

TI's products are provided subject to [TI's Terms of Sale](#), [TI's General Quality Guidelines](#), or other applicable terms available either on [ti.com](#) or provided in conjunction with such TI products. TI's provision of these resources does not expand or otherwise alter TI's applicable warranties or warranty disclaimers for TI products. Unless TI explicitly designates a product as custom or customer-specified, TI products are standard, catalog, general purpose devices.

TI objects to and rejects any additional or different terms you may propose.

Copyright © 2026, Texas Instruments Incorporated

Last updated 10/2025