

NPU

(神經處理單元)



Reported
FAE Dept.

Date
2026.05.06



1. NPU技術簡介
2. NPU、CPU、GPU技術解析
3. NPU、CPU、GPU差異
4. NPU廠商
5. 案例

- **NPU**：(Neural Processing Unit, 神經處理單元)，負責神經網路運算 (Neural Network)，專為人工智慧 / 機器學習 (ML) 設計的處理器
- **技術特點：**
 - 模擬人類神經網路的運作方式、經由「近記憶體運算」
 - **三大核心特性：**
 - 1. 專業化架構設計
 - 2. 高能效比
 - 3. AI 推論專用引擎



NPU、CPU、GPU技術解析

	CPU	GPU	NPU
			
運算任務類型	一般應用及多樣化邏輯運算 最適合	高度平行與大量數據運算壇 長	專注AI推理與神經網路推算
功耗與效能需求	綜合通用性與靈活性較佳	提供強大運算力，適合大型 伺服器或AI訓練	低功耗高效能，適合移動或 邊緣裝置
成本與硬體整合	為系統核心處理中心，支持 多樣化架構	通常為獨立顯卡或內建於 CPU的顯示晶片，擴展性高	常與SoC整合，桌機及筆電 上通常與CPU同為一個品 片，降低系統複雜度與成本
適合應用場景	多任務處理、伺服器架構、 傳統軟體進行	影像處理、視覺運算、機器 學習訓練、高性能運算	AI推論、智慧手機、邊緣計 算、智慧家居與物聯網裝置

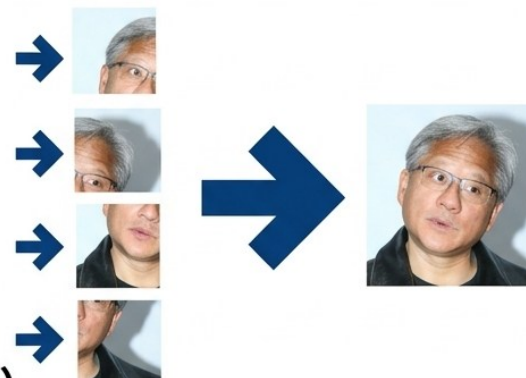
▶ CPU (中央處理器)

擅長處理複雜事務，
一次執行單一任務，
按順序進行。



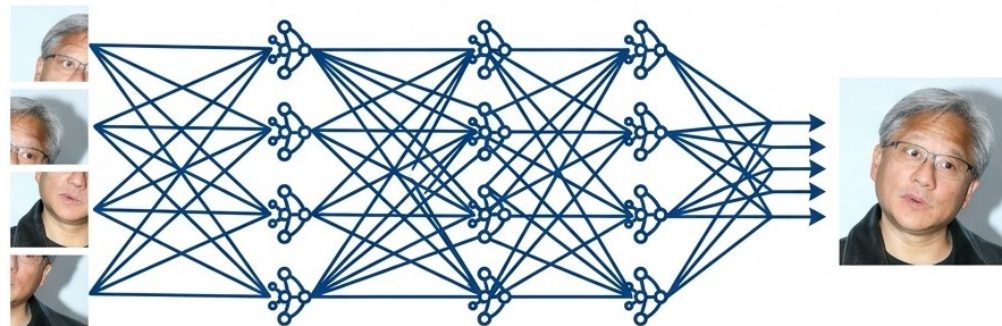
▶ GPU (圖形處理器)

擅長圖形與大規模並行運算，
可同時處理多項任務。



▶ NPU (神經網路處理器)

專為 AI 並行運算設計，大幅提升 AI 的執行效率。



廠商	NPU	效能(TOPS)
Intel	Lunar Lake	48
Supermicro	Ryzen AI 300	50
Qualcomm	Snapdragon X Elite	45
MediaTek	天璣9300	48
Apple	M4	38

TOPS

Tera Operations Per Second
每秒兆次運算

- 用來衡量AI硬體運算能力的單位
- 廣泛應用在自動駕駛、邊緣計算、VR/AR、雲端計算、AI加速器等，特別是NPU（高效神經處理）、深度學習方面

▲ 案例一：視訊會議 – 智慧視訊增強

AENEAS

NPU 功能應用：

- 背景模糊與更換功能由 NPU 即時處理，不吃 CPU 資源
- 自動構圖 (Auto Framing)：人物走動時畫面自動對焦
- 眼神校正 (Eye Contact)：即使看螢幕也像在看鏡頭
- 噪音抑制：即時消除背景聲音，如打字聲、狗叫聲



NPU 功能應用：

- 本地化執行 AI 文件摘要，不需每次上雲端查詢
- 自動完成電子郵件、安排行程
- 支援自然語言查詢功能：輸入「請幫我找出與上週銷售報告有關的檔案」即快速回應



NPU 功能應用：

- Auto Reframe 自動裁切影像主體，適配直式/橫式平台
- 一鍵背景移除，無需綠幕
- 即時濾鏡、影像強化（例如自動亮度修正）
- 預覽處理不延遲，即便使用筆電也順暢





Thank You!

