

產業焦點評論

Tesla AI Day 之觀察與評析

前言

美國時間 8 月 19 日，全球電動車龍頭 Tesla 舉辦的 Tesla AI Day 正式舉行。Tesla 在眾人期待下展示了尖端領先的軟硬體技術，從取代光達的純視覺自駕技術、透過台積電代工的 7 奈米自製 D1 晶片、躍居世界頂尖超級電腦行列的 Dojo，加上最後 Elon Musk 帶來的驚喜—人形機器人，宣示 Tesla 跨出造車行業，以頂尖人工智慧公司為新定位的戰略方向。活動中的技術發展與亮點，不僅符合 Tesla 招募人才的初衷，也使不久前身陷多起自駕爭議調查的陰霾得以稍微緩解。

陳建安

爭議中展示尖端技術、力求招賢納士

今年 7 月，Tesla 執行長 Elon Musk 宣佈 Tesla AI Day (人工智慧日) 將於 8 月 19 日舉行，活動中將說明 Tesla 最新人工智慧軟硬體發展，包含訓練、推理與應用的現況，但主要目的在於招募人才，說服最優秀的人工智慧人才加入，是活動的唯一目標。

Tesla AI Day 是繼 2019 年「Autonomy Day」與 2020 年「Battery Day」以來，Tesla 第三次的新技術展示會。但於活動即將舉辦前數日，Tesla 卻捲入了 FSD (Full Self-Driving) 和 Autopilot 兩項輔助駕駛產品的爭議，並由美國國家公路交通安全管理局 (NHTSA) 宣佈針對多起事故進行調查。

儘管遭受眾人質疑，但 Tesla AI Day 展示的技术結晶依然令人驚豔，以下將針對活動各亮點，概述其內容。

以純視覺神經網路取代雷達實現自駕

Tesla 仰賴神經網路 (Neural Network) 與深度學習 (Deep Learning) 分析來自 8 顆環景鏡頭的串流影像，以辨識車輛、道路、行人、號誌等重要資訊。Tesla 相信僅依據影像進行深度判斷、物件辨識的神經網路，雖然技術極為困難，但卻更貼近僅依賴視覺開車的人類，無須倚賴高精度地圖與光達等感測器，是可通用於世界各地的最佳方案。

Tesla 用於影像辨識的核心演算法，是稱為 HydraNets 的多工神經網路，可同時進行物件識別、號誌辨識與車道偵測三大任務。HydraNets 可同時辨識上千種物體，在共享骨幹網路下節省運算資源與測試時間，在訓練、測試或實際運行時皆能提昇大幅效率，以因應自駕的即時性要求。

雖然 HydraNets 辨識能力強大，但在實現 FSD 時 Tesla 發現仍有許多不足，物件識別在實務上作用有限，自駕真正需要的，是從影像中投射出現實世界的三維重建結果。

將 HydraNets 結合了 Transformer 與 Spatial RNN (Spatial Recurrent Neural Networks，空間遞迴神經網路) 兩種神經網路後，Tesla 可將多張平面影像即時轉換為鳥瞰視角，精確還原出周邊道路與場景資訊，並拼合、描繪出完整的地圖與運行路徑。據 Tesla 說法，目前純視覺神經網路的精確性已可與光達匹敵，並認為在一般自駕場景下已可捨棄光達，完全倚賴視覺進行場景還原。

圖一、Tesla 以多視角攝影機結合 HydraNets 還原鳥瞰向量空間



資料來源：Tesla · 2021 年 8 月

自動標註、模擬場景支持海量訓練集

Tesla 的神經網路架構多採用監督式學習 (Supervised Learning) 方法，儘管辨識效果優異，卻需要極大量的標註資料作為訓練素材。考量標註品質並因應快速攀升的大量需求，目前 Tesla 已成立千人內部團隊專職負責標註任務。

有別於傳統，Tesla 標註任務不再只是針對圖像空間，而是先運用 HydraNets 還原出場景後，直接在向量空間的四維度 (三維空間+一維時間) 上直接進行，維度的提昇使標註工作複雜度和資料量大幅提昇。對此 Tesla 開發了自動化標註系統，雖然並未提及細節，但該系統涉及神經網路、深度資訊與人工審核的組合，可透過拼接不同的剪輯畫面，產生充滿障礙物、牆壁、道路、行人等豐富元素的海量訓練集。目前 Tesla 累積了 1.5PB 的數據，包含 100 萬個 10 秒影片與 60 億個標註對象。

然而現實獲取的資料，儘管數量龐大，但卻仍無法涵蓋所有的可能性。Tesla 展示幾可亂真的渲染技術，產生無數貼近現實的訓練資料，解決資料難以獲取、難以標註、訓練資料不足的窘境。結合神經網路渲染的模擬場景，可產出精確的感測器數據、逼真的動態影像，以及在任何天氣、地點中，加入任意的演員、障礙物等物品，快速產出可規模化、可擴展的場景訓練資料。目前 Tesla 模擬資料庫已包含了 3.71 億張訓練圖像，作為龐大訓練集的補充與輔助資源。

圖二、Tesla 神經網路渲染成果與模擬訓練場景



資料來源：Tesla · 2021 年 8 月

自製訓練晶片實現世排前五超級電腦

儘管 Tesla 軟體技術優異，但人工智慧對運算速度和訓練能力的需求永無止境，急需能改善延遲、減少昂貴訓練成本的途徑。因而 Tesla 推出了 Dojo 超級電腦作為硬體終極解方，有別於傳統分散式架構的延遲缺點，Dojo 以高效能訓練節點 (Training Node) 為最小單元，每個節點擁有 1,024 GFLOPS (BF16/CFP8) /64 GFLOPS (FP32) 的驚人算力，以及每秒 512GB 的超高頻寬，節點可匯集共同運算，達成高頻寬與低延遲性。

Dojo 核心是 Tesla 自行開發的 D1 晶片，採取台積電代工的 7 奈米製程，以具「 GPU 運算力、CPU 靈活性」為設計方向，並可達到頂尖網路交換晶片的兩倍 I/O 頻寬。D1 晶片具有高度擴充能力，四邊都有配置連結器，各邊頻寬高達 4TBps，可在不犧牲速度下自由擴展組合，並可虛擬化配置運算資源。

Tesla 以模組化方式進行 D1 晶片組裝，在 5×5 平面上整合後，下方輸入高達 18,000 安培的電流，上方為 15KW 的熱處理裝置，構成訓練磚 (Training Tile) 模組。12 個訓練磚以 4×3 構成一個訓練矩陣 (Training Matrix)，每個機櫃可容納 2 個訓練矩陣，最終由 10 個機櫃組成 Dojo 的運算系統—ExaPOD。

ExaPOD 運算系統包含 3,000 個 D1 晶片，綜合算力高達 1.086 Exa FLOPS (BF16/CFP8) /67.8 Peta FLOPS (FP32)，Exa 級算力即為 ExaPOD 之命名緣由。Tesla 稱 Dojo 為世界最強大的人工智慧訓練用超級電腦，可在相同成本下達成 4 倍訓練效能，並減少 5 倍的碳足跡，目前可在世界超級電腦中排行第 5，勝過美國國家能源研究科學計算中心 Perlmutter 的 64.59 Peta FLOPS (FP32)。

儘管 Dojo 預計明年度才啟用，Tesla 宣告已著手開發下一代的 Dojo，預計運算效能可再精進 10 倍。Dojo 主要的任務為運用於 Tesla 內部，執行神經網路訓練任務，處

理日益遽增的海量訓練集。除此之外，Tesla 也計畫在未來，將 Dojo 提供給外部人工智慧開發人員使用，亦不排除可創造新的商業模式。

推出人形機器人解決低階勞動力問題

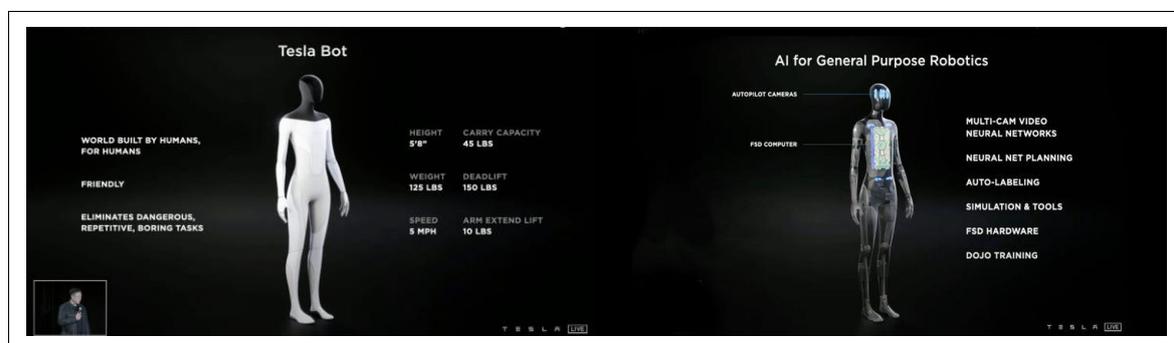
Elon Musk 宣佈 Tesla 將跳脫造車定位，以技術領先的人工智慧公司為導向，規劃推出人形機器人產品 Tesla Bot。Elon Musk 認為，Tesla 的電動車如同裝載輪子的半感知型機器人（Semi-Sentient Robot），並足以自稱為全世界最大的機器人公司。

Tesla 現場亦提供 Tesla Bot 規格，其身高 172.7 公分，體重為 56.7 公斤；面部沒有五官，但配有顯示螢幕；最高速度為時速 8 公里，相當於人類的慢跑速度；硬舉重量 56.7 公斤，承載重量 20.4 公斤。Elon Musk 表示，Tesla Bot 初衷為處理不安全、重複和枯燥的任務，協助 Tesla 解決低階勞動力問題，刻意設計為故障時可由人類壓制，不至於對人類造成威脅。

Tesla Bot 雖然以泛用人形機器人為定位，但並非設計為可進行人機複雜交流的 AI，初步構想是接受使用者簡單的指令，如更換燈泡、裝卸螺絲等。然而 Tesla Bot 卻匯集了目前 Tesla 所有尖端技術，其頭部裝了 8 個攝影鏡頭，胸腔內含搭載 FSD 的移動型電腦，而物件辨識、移動、路徑規劃等演算法，都是基於 Dojo 所訓練出來的神經網路，儼然是將 Tesla 電動車改裝為人形機器人的模樣。

儘管 Tesla Bot 是 Tesla AI Day 的大亮點，但令人扼腕的是，目前 Tesla Bot 還停留在構想階段。活動現場僅展示了一台 Tesla Bot 模型，預計明年會推出可運作的原型，但對於何時進軍機器人產業、何時進入產品商業化階段，Elon Musk 仍未提供明確的規劃。

圖三、Tesla Bot 集 Tesla 人工智慧技術大成



資料來源：Tesla，2021 年 8 月

MIC 觀點

AI軟硬體技術成為Tesla資源投入重中之重

與過去技術發表會不同，今年從宣佈召開 Tesla AI Day 開始，Elon Musk 就將活動設定為以招募人工智慧人才為目標，足見 Tesla 對於掌握人工智慧資源極為重視，而招募人才的主要目的，應是希望能共同解決在自駕技術所遭遇的難題。

從最核心的純視覺自駕演算法中，可以看出 Tesla 正逐漸以神經網路取代過去程式碼的歷史包袱，並屢屢強調運算效能和準確性的提昇，而這一切的努力，都是為了能夠擺脫對光達的依賴。然而 Tesla 純視覺自駕方案發展至今，雖然已可取代光達，但卻也付出須高度倚賴深層神經網路的代價。

從影像辨識、機器視覺和資料標註等流程，目標都是從海量的資料中訓練出最聰明的神經網路模型，為訓練階段的演算能力和資料吞吐能力帶來極大的壓力，因而 Tesla 必須選擇自製人工智慧晶片和超級電腦，以滿足市面方案所無法滿足的龐大內部需求。由 Tesla 展示的尖端軟硬體技術，及對後續規格升級的策略發展需求，可期待未來在超級電腦的晶片代工、系統架設、整合與維運等供應鏈方面將帶來利多。

純視覺深層神經網路自駕仍待市場實際考驗

一般而言，自動駕駛車輛會透過光達、雷達、攝影鏡頭、超音波聲納等感測器探索環境，但 Tesla 近年特立獨行，有別於坊間自駕系統，力求捨棄光達，堅持使用純視覺的神經網路方案。然而市場對 Tesla 純視覺技術並不買賬，今年美國車輛測試權威「消費者報告」(Consumer Report) 因 Tesla 改採純視覺方案，認知為缺乏關鍵先進安全功能，將 Model 3 從頂級汽車排行榜上移除。

深層神經網路在自動駕駛上是否能克服所有挑戰，目前仍未有明確結論，難以進行進行因果推理、基於黑箱模型的演算法架構，就算投入再多海量數據進行訓練，依然無法確保在現實中不可預測情境的絕對可靠性，使責任歸屬面臨相當大的阻礙。

回顧 Tesla 近期接受調查的事件，兩位美國參議員也同時向聯邦貿易委員會 (FTC) 呼籲調查 Tesla 的 FSD 與 Autopilot 兩項產品是否有廣告不實、誤導性行銷的問題，而 NHTSA 也就此向大眾發出提醒「目前沒有任何一輛市售車輛具備自動駕駛功能」。

儘管調查過程通常曠日費時，對 Tesla 而言仍有緩衝空間，但基於輿論影響下，Tesla 亦有可能被迫對自駕產品變更命名與行銷方式，改以「輔助駕駛」作為行銷口號，形同對 Tesla 車款施以限制，進而拉近 Tesla 與各傳統車廠競爭者 ADAS (Advanced Driver Assistance System, 先進駕駛輔助系統) 的差距。

自製D1晶片與Dojo超級電腦短期效益顯著

超級電腦通常被視作國家級的投入資源，以滿足特定國家資助單位的高規格需求為目的，其進入門檻非常高，晶片架構長久以來由 ARM、Intel 所壟斷。Tesla 憑一己之力，自行研發晶片與運算架構，擊敗現今 GPU、CPU 架構，甚至技壓 Google 的 TPU 人工智慧晶片，躍居世界排名前五的領先地位。

而下一代 Dojo 若順利研發成功，算力提昇 10 倍後，Tesla 將可擊敗現今冠軍 Fugaku 的 442.01 Peta FLOPS (FP32) 高標，成為擁有世界算力最強大超級電腦的機構，足見其技壓群雄的優異技術。而若 Tesla 願意對外釋出 D1 晶片授權，可想見將對高性能運算晶片的供應鏈生態帶來顛覆性的變化。

高度彈性、先進的虛擬化資源分配架構，更是 Dojo 令人超乎想像的特色。在短期策略上，Dojo 不只可用以處理 Tesla 內部海量的訓練任務，剩餘運算資源 Tesla 也將開放給外部使用。而 Dojo 在開發流程上與 Python、Pytorch 等開源語言與框架相容，底層架構更規劃了虛擬化動態分配運算資源，與 PaaS (Platform as a Service，平台即服務) 服務高度相容，快速進入商業化步驟，成為 Tesla 可觀的收益來源，即有可能對 Google、Amazon 等雲端機器學習平台業者造成相當衝擊。

人形機器人與公司新定位長期價值仍未明朗

儘管 Tesla Bot 讓眾人眼睛為之一亮，但進軍人形機器人領域對 Tesla 而言，是既太空與能源產業後又一個新挑戰。但目前 Tesla 最令眾人詬病的，無非是自駕產品疑慮仍未解決，以及面臨中美市場的品管召回、在德國設廠受阻、交車目標屢屢跳票等問題。此時投資人多希望可專注於造車工藝的精進，而非分散資源在非核心業務上。

Tesla 目前的規劃，僅是將大部分的電動車技術移植到人形機器人上，但機器人的導航與控制和電動車大不相同，運作場域多為室內而非室外，在資料與技術上難以取得太多綜效。Tesla 傾注資源押寶人工智慧作為公司新定位，是否能通過考驗、取信於市場，無論是實施新演算法、導入超級電腦或是投入專家人才，都需要歷經磨合和考驗，對 Tesla 的長期效益仍待時間證明。

以長期策略而言，機器人產業並不是容易跨足的領域，市面上已有許多人形機器人產品，但多半叫好不叫座，市場與獲利能力有限。而 Tesla Bot 預計明年才有原型推出，產品功能和實際上市時間也未定，可想見 Tesla 藉由人形機器人炒作話題，拉抬股價招募人才後，極可能暫先回歸造車本業，解決眼前棘手問題。對於機器人供應鏈相關之感測器、控制器、驅動元件等零組件而言，短期應無顯著影響。

附錄

英文名詞縮寫對照表

ADAS	Advanced Driver Assistance System
FSD	Full Self-Driving
PaaS	Platform as a Service
Spatial RNN	Spatial Recurrent Neural Networks

中英文名詞對照表

先進駕駛輔助系統	Advanced Driver Assistance System
深度學習	Deep Learning
神經網路	Neural Network
非凸性	Non-Convex
平台即服務	Platform as a Service
半感知型機器人	Semi-Sentient Robot
空間遞迴神經網路	Spatial Recurrent Neural Networks
監督式學習	Supervised Learning
訓練節點	Training Node



發行所	財團法人資訊工業策進會 產業情報研究所(MIC)
地址	台北市 106 敦化南路二段 216 號 19 樓
電話	(02)2735-6070
傳真	(02)2732-1353
全球資訊網	https://mic.iii.org.tw
會員服務專線	(02)2378-2306
會員傳真專線	(02)2732-8943
E-mail	members@micmail.iii.org.tw
AISP 會員網站	https://mic.iii.org.tw/aisp

以上研究報告經 MIC 整理分析所得，由於產業變動快速，並不保證上述報告於未來仍維持正確與完整，引用時請注意發佈日期，及立論之假設或當時情境。
著作權所有，非經 MIC 書面同意，不得翻印或轉載