

CASE STUDY | NATIONAL CHIAO TUNG UNIVERSITY

交通大學電控所運用 NVIDIA 深度學習加速器 提升智慧影像分析研究效能



IVA Lab 有鑑於健康管理商機巨大，今後將持續透過 GPU 深度學習進行訓練，期許開闢更寬廣的應用加值空間，為民眾福祉做出更大贡献。

摘要

- > 交通大學電控工程研究所混沌系統與訊號處理實驗室一直致力於智慧控制以及影像處理上的研究，全國第一輛自動駕駛車即出自於此。
- > 主持人吳炳飛特聘教授也負責交大與研華公司的「物聯網智慧系統研究中心」，智慧影像分析實驗室。目前主要研究領域分成人臉影像辨識與分析、影像式行為分析、影像式生理訊號截取以及智慧購物等四大區塊，精進各種相關演算法，為這個智慧化時代盡一份心力。

軟體技術

深度學習框架：Caffe, TensorFlow

硬體設施

GS2020 伺服器

J-Castle 雲城堡

NVIDIA Tesla M40 24G * 2

人工智慧 (AI) 概念誕生迄今雖已有數十年，然近年隨著 GPU 平行運算、雲端運算、大數據分析和機器與深度學習等技術出現，呈爆炸性成長，舉凡 Google、微軟、蘋果、特斯拉 (Tesla) 等科技巨擘，皆已投入 AI 發展。同時，AI 亦成為學術研究領域的熱門議題。瞄準物聯網與智慧應用浪潮，交大與研華公司透過產學合作催生「物聯網智慧系統研究中心」，設立智慧影像分析實驗室 Intelligent Video Analytics Lab. (IVA Lab.)。由交通大學電控工程研究所 (簡稱交大電控所)，吳炳飛特聘教授 (IEEE Fellow) 主持，借助 AI 與深度學習方法，以突破傳統影像分析演算法的極限。

交大電控所憑藉控制技術與系統研發的基礎，近年也針對智慧型機器人、智慧車輛、系統晶片 (SoC) 與生醫科技及能源技術等新興產業應用主題，端出許多豐碩研究成果。

挑戰

綜觀 AI「深度學習」是成長最快的領域，其訴求透過多層級神經網路，讓電腦能模擬人腦思考行為，對複雜事物進行觀察、學習及反應，並理解文字、聲音及影像等資料的真實意涵。

初期神經網路猶如小孩牙牙學語，必須不斷被灌輸訓練素材、修正模型，才能慢慢成長為某一領域的專家，而在反覆匯入訓練資料及驗證演算法的過程，需消耗龐大運算資源，因此若研究單位僅憑藉一般處理器 (CPU) 進行培訓，勢必曠日費時。

交大電控所特聘教授吳炳飛表示，IVA 智慧影像分析實驗室的主要方向為基於深度學習的影像分析演算法，藉由 Caffe、TensorFlow 等演算法框架進行主題研究；不論所用的框架或開發工具為何，在執行訓練過程中，皆需借助高達數百萬張圖片素材，前後耗時數星期之久，時效急待提升。

有鑑於此，IVA Lab 決定改採 GPU，引進搭載兩個 NVIDIA® GeForce® GTX TITAN X 核心的電腦系統並採用 CUDA 平行運算架構，期望藉此增進訓練速度。運用 NVIDIA® GeForce® GTX TITAN X 後，在訓練速度上有近百倍的突破。



採用 GPU 理由

- > 大幅縮減深層模型的訓練時間
- > 可處理資料量大
- > 計算時間縮短，使得電腦得以即時執行更多更複雜的演算法

解決方案

交大「物聯網智慧系統研究中心」IVA Lab 進一步期盼以更複雜的模型以及更巨量的資料來大幅提高深度神經網路的學習能力，評估一些 GPU 運算卡後，最終決定以兩張 NVIDIA® Tesla® M40 運算卡的伺服器做為深度學習訓練平台。

NVIDIA® Tesla® M40 蘊含 3,072 個 NVIDIA® CUDA® 核心，24G 的記憶體，因此很適合批次處理較大量的資料，對於複雜的模型訓練來說，可以較容易學習到複雜資料的潛在特徵，而且也較快達到訓練的要求。搭配 NVIDIA Docker 使用，可以讓多位開發人員透過遠端登入系統，共享計算資源，而且互相不會干擾。

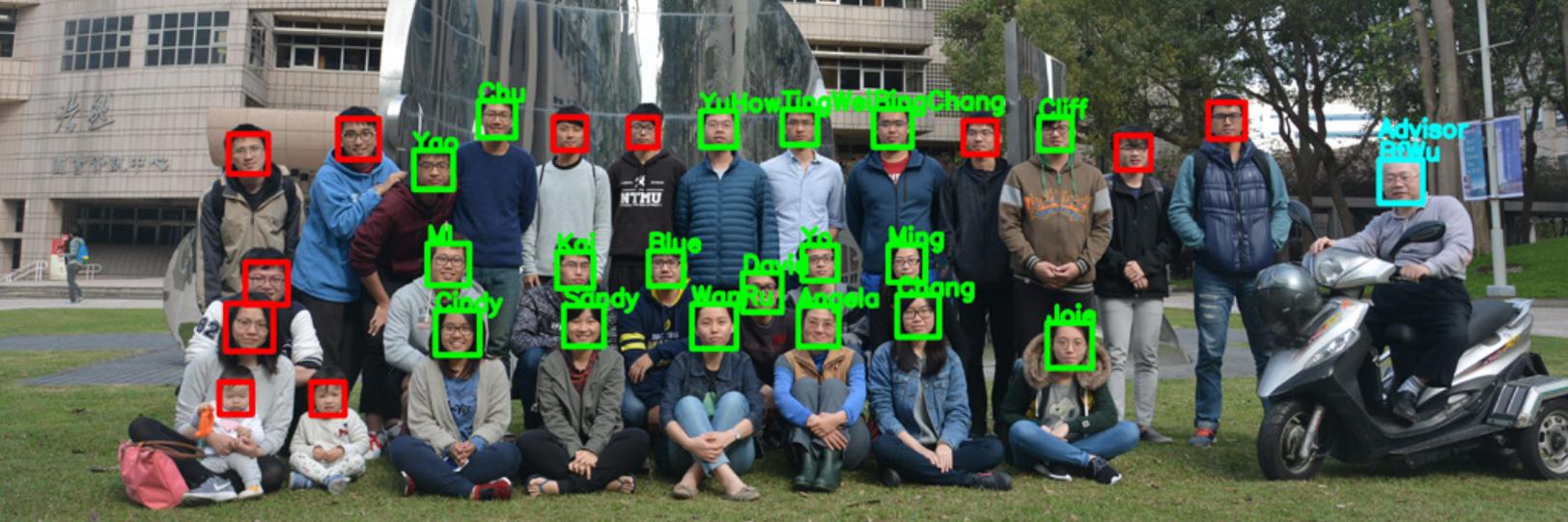
每當要利用深度學習來達成某一個目的地時，必須經過一段擺長的調參與測試的過程，原本 NVIDIA® GeForce® GTX TITAN X 每次只能供 1~2 個模型訓練，資料量也不能太大，如今在 NVIDIA® Tesla® M40 上可供 5 個以上的模型訓練，資料量可以到好幾 TB，而且在訓練的過程中，電腦還是可以跑得很順，因此開發效率可以增快 3~5 倍，以大量資料所訓練出來的模型也能有更好的能力。

影響

IVA Lab 成立至今，先後投入多項研究主題。初期主要著重零售賣場應用，研發智慧結帳系統，訴求使用者只要將商品放入購物車，無需借助傳統條碼掃描，便能自動藉由攝影機辨識商品並同步累計消費金額，使消費者無需因等候人工結帳而耗時排隊。

爾後 IVA Lab 延伸研究觸角，透過人臉與表情辨識，繼而結合不同應用領域的知識，判讀人臉表情變化的背後意涵，開創多元應用價值。如智慧電子數位看板，藉由看板上方的攝影機，持續捕捉與解析每位觀眾的關注區域、臉部表情變化，以利業主分析不同型態的客群，作為優化廣告投報率的依據。

車用電子則是另一應用方向，透過車內影像辨識裝置，可協助後台管理中心持續監測駕駛在行車過程中的行為與心跳變化，一旦察覺有危險的駕駛行為或心跳與基準值差異值過大，便立即做出對應的警訊，預先避免車禍憾事



發生。目前更進一步研發由車前攝影機所看到的影像，判斷感知車輛周邊發生的事件。例如，當辨識系統辨識出路上有許多車輛或人後，進一步判斷這些物件代表的可能發生車禍或是施工或是商店特賣等可能事件。

至於另一更具價值與商機的應用，則為健康管理，可謂「深度學習無所不在」的表現。以消費者為例，平日在家活動頻率最高的空間，不外是盥洗室與客廳，若在消費者經常直視的衛浴鏡或電視裝設智慧影像分析裝備，隨時量測生理訊息，持續上傳雲端，再結合專業醫護機構解析數據，便可望及早察覺病變徵兆，避免病情延誤而危及健康與生命。

IVA Lab 有鑑於健康管理商機巨大，今後將持續透過 CUDA / GPU 深度學習進行訓練，擴大智慧影像分析應用範圍，從現有的人臉、行為、商品、心跳等等，繼續往前推進，期許開闢更寬廣的應用加值空間，為民眾福祉做出更大贡献。