



30科研新勢力

在明年迎來香港回歸祖國30周年的重要節點，特區政府正密鑼緊鼓準備以香港首份五年規劃「立」穩航道，穩步向前。年近30，往往也是青年科學家確立終身志業的重要時刻。香港文匯報「30科研新勢力」系列第二集專訪入圍本年度《福布斯》「30位30歲以下精英（30 Under 30）」亞洲區榜單、年僅28歲的香港理工大學學者陽行意，聆聽這位研究生成式人工智能（AI）模型「高效訓練」的新銳科研力量，如何在快速迭代的技術浪潮中穩住節奏，讓訓練和創造模型的能力不再集中於少數大型公司，而是交給更多研究者、企業和普通使用者，真正實現AI能力的高效普惠。 ●香港文匯報記者 陸雅楠



●陽行意 理大圖片



AI模型訓練往往需動用成千上萬枚高效能芯片長時間運作，而芯片與算力需求所耗用的電與水，令訓練過程被形容為「耗能巨獸」。為降低環境衝擊並控制成本，業界與學術界近年正加緊尋找更聰明、更省電的高效訓練方法。

專注研究生成式AI、世界模型與空間智能等方向的理大數據科學及人工智能學系助理教授陽行意指出，生成式AI的高效訓練，本質上屬於「系統問題」，而非單純的「演算法問題」。

陽行意解釋，過往談及高效訓練，討論焦點多落在縮小模型規模或設計計算量更低的演算法。然而，訓練大型模型更像在經營一條大型生產線，模型本身只是其中一部機器，整體成本與效率還要取決於設備如何選配、數據能否及時送達、不同設備如何協同，及投入的「原料」是否真正有價值，「若只優化其中一個環節，整體效率未必會明顯提高。」

模型設計需配合硬件能力

要做到高效訓練，他認為當前最有效策略不在於探尋單一技術解決所有問題，而是要把硬件、訓練系統、模型和數據各層面同時打通，並梳理出關鍵因素。首先是硬件匹配。不同計算設備各有特點，有些擅長同時處理大量任務，有些更適合特定類型運算，只有讓模型設計與硬件能力相配合，才能真正發揮其效率。

其次，是訓練平台與系統調度。很多時候模型並不是在「計算」，而是在「等待」，例如等待數據讀取、設備交換資訊，或資源重新分配等。他形容就像生產線中，即使某一台機器提速三成，但若原料供應跟不上，會令前後工序銜接不暢，總產能提升仍然有限。

其三是模型效率。模型不必每次都從頭訓練，可只調整與新任務最相關部分；亦可讓小模型吸收大模型既有能力，降低訓練成本。他形容部分模型雖規模龐大，但執行任

務時只需調用部分能力，猶如大型醫院按病情邀請相關專科介入，無須每次都召集所有科室，既能保留模型能力，也能降低計算成本。

最後是數據效率。大模型需要海量數據，但部分數據重複、低質或與目標任務關聯少，並非愈多愈好。若能從海量數據中篩選最具價值部分，便有望用更少數據訓練出同等甚至更佳模型表現。

他指近年包括DeepSeek、Kimi等模型的發展，讓外界更清楚看到效率提升往來來自多環節協同改進，「每一層的改進可能只有百分之十或百分之幾十，但當這些提升疊加起來，最終可能帶來重量級的效率提升。」

陽行意團隊的核心思路，是保留現有大模型的既有能力，以更少數據、算力與時間把模型快速遷移至新領域。他以改造大樓作比喻：一幢功能完備的大樓不必推倒重建，可在較低成本下調整內部結構，以適配醫院、學校等不同場景。循此思路，團隊從演算法、數據到系統層面同步降本增效，力求在有限資源下實現快速訓練與落地適配。

他強調，高效訓練最終要解決的，不只是少用幾張GPU（圖形處理器）的技術問題，也不是讓生成式AI始終只是少數大型企業之間的資本與算力競賽，而是希望把訓練和創造模型的能力交給更多研究者、企業和普通使用者，讓AI能被每個人使用、改造和創造。



●陽行意於學術會議上就深度模型重用及重組進行報告。 理大供圖

從學生到教授 發現科研世界更廣闊

去年博士畢業後加入理大的陽行意，從學生轉變指導老師，最大的感觸來自思維方式與責任邊界的徹底轉換。他形容，那就像突然看到了科研這座山的另一面，「山的這一邊，是自己做研究、寫論文和解決問題；翻過山之後，我看到一個更加廣闊的世界：建立自己的團隊，培養新的研究者，提出更長遠的問題。但與此同時，那裏也有更多未知、責任和艱難險阻。」

陽行意指，學生時代目標清晰，「把一個問題解決好，把代碼和實驗做好，再把論文寫出來投出去。」成為導師後則要思考更大的問題：「團隊未來幾年應該研究什麼？不同學生適合什麼方向？如何發揮每個人的長處，同時讓他們相互合作，形成真正有積累、有延續性的研究方向？」

他直言，幫助學生建立獨立研究能力，比獨立完成論文要困難得多，「每人性格、基礎和思維方式各異，教授需要引導他們明辨慎思，更需要耐心和不斷調整指導方式。」

今年上半年，陽行意招收了第一位博士生。此前兩人亦曾長時間合作，但主要是討論想法、做實驗和寫論文。成為對方的博士導師後，他便要為其研究尋找及計算資源或硬件支持，發掘合作機會和長期發展空間，「以前作為研究生，我通常只會關注眼前的研究問題；現在我才看到，一項研究能順利推進，背後還需要許多資源、協調和責任。」

肩上新擔更大，但陽行意卻覺得充滿動力，因為現在是AI研究的黃金時代，「但黃

金和沙子往往混在一起。年輕研究者真正需要的，是看清方向、作出判斷，然後勇敢行動。」關鍵在於有把握時代主線的能力，每個階段都會出現一些真正改變領域走向的技術，正如從2022年開始，大語言模型就是一條非常清晰的主線，「愈早看到這種變化，並真正參與其中，就愈有機會積累經驗。」與此同時，研究人員也需要從大量資訊中判斷什麼只是短期熱度，什麼是真正重要的問題，什麼適合自己。

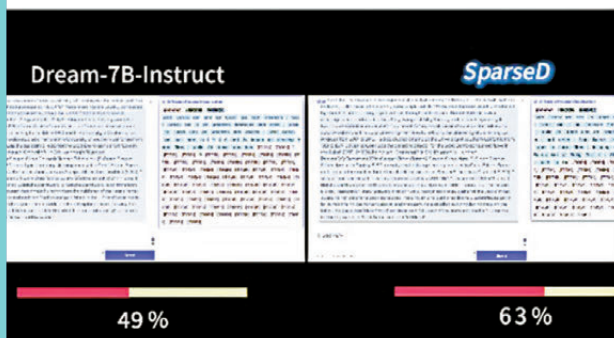
寄語後浪把握時代脈搏

技術發展得愈快，人愈容易焦慮，但陽行意認為很多時候真正讓人掉隊的，不是暫時選錯方向，而是一直站在岸邊分析水溫，卻始終沒有下水，「你可以完全進入一個新領域，也可以把新的技術帶回自己原來的研究中，但無論選擇哪一種，都需要先行動起來。」他寄語有興趣的青年研究人員：「要把握時代脈搏，過濾周圍的雜訊，然後勇敢地參與進去。找到真正值得做的問題之後，也要給自己一點耐心。AI迭代得很快，但真正好的研究，需要時間沉澱。」

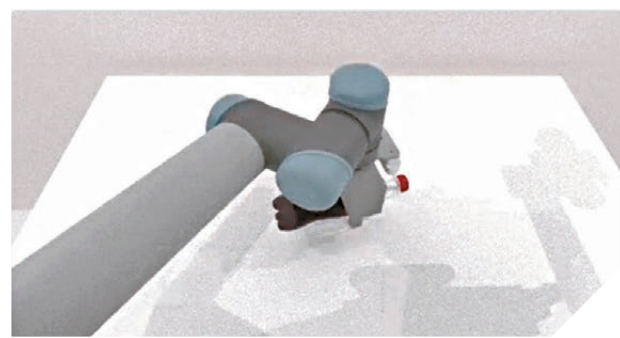


打破算力壟斷 實現AI高效普惠

善用遷移學習與知識蒸餾 理大陽行意致力攻克訓練難題



●陽行意研究項目：「高效AI—稀疏擴散」（左）及「具身AI—交互智能體」（右）。



網上截圖

獲學術圈外肯定 證研究具現實意義

對入選《福布斯》30 Under 30榜單，陽行意喜見來自學術圈以外的肯定。他形容，科研工作很多時比較「安靜」，或耗時數年深耕一個問題，價值未必更廣泛地被社會看見，認為登上榜單反映其研究成果獲更多人關注，也從側面印證高效AI訓練、生成式AI以至醫療健康等方向確具有現實意義。對陽行意來說，今次獲選既是階段性的鼓勵，亦是新的起點，若因此能有更多教育、科技界以至醫療領域人士了解其研究，並促成有價值的合作，意義遠勝於入選本身，並讓他更堅信要持續聚焦重要問題、踏實把工作做好，便有機會產生更大影響力。

AI技術日新月異，身處不斷迭代的AI浪潮中，未足30歲理應是「後浪」的陽行意亦坦言，沒有人能保證自己的研究永遠不被取代，能做的只有在解題時盡量比眼前問題多想一步，同時保持快速學習與調整的能力，而有價值的研究不在於把既有方法提升幾個百分點，而在於提前洞察下一階段會出現的關鍵瓶頸，具體模型與技術或許迭替甚快，但效率、可靠性和落地等底層命題不會輕易消失，「只要研究的是這些長期問題，就不容易因為某個新模型的發布而立刻失去意義。」

保持資訊敏感度同樣關鍵，陽行意說AI領域變化太快，僅靠偶爾讀幾篇論文遠遠不夠，更要與同行交流，他也會使用AI工具協助搜尋與整理最新進

展，務求以更有效率方式跟上知識增長，並根據趨勢判斷是否要調整研究方向，更要「相信團隊，尤其是相信學生」，因為個人的時間與認知終究有限，導師不可能在所有議題上都最懂，更不應假裝全知全能。每個學生在各自方向上閱讀、實驗與思考，往往能帶回新的技術與視角，甚至推翻他原先的判斷，「一個好的團隊，不是所有人都聽一個人說話，而是每個人都能帶來新的資訊和視角。」

展望未來，在研究上，他有一個大膽科幻的構想：能否通過AI和演算法，創造新的智能形態，甚至可以把它理解為一種「硅基物種」？他指出，大語言模型在某種程度上為機器提供了類似大腦的能力，使其能理解語言、處理知識，並在部分問題上展現超出人類直覺的推演能力。下一步，他期望把感知與行動能力也納入其中，讓模型擁有在真實世界中觀察、學習並完成任務的能力。

望培養具解難能力新一代

在教育上，他希望培養具備獨立思考與解難能力的新一代，讓他們在科研過程中獲得思維訓練、資訊判斷、表達溝通及處理複雜問題的能力。他認為科研是創造新技術，也是培養能持續創造未來的人，「技術會迭代，論文會過時，但真正具備獨立判斷與創造力的人才，不會輕易被淘汰。」

論文「沉海」曾陷低谷 導師引導重建信心

談及科研之路的開端，陽行意憶述，自己是於東南大學本科時開始參與科研和產業項目，到升大四的暑期更投入海外研究計劃並長時間沉浸其中，反覆思考、嘗試和驗證，深深感受到科研的吸引力。然而，赴美讀碩士期間，陽行意亦曾經歷低谷，論文投稿幾乎全部石沉大海，常常自我懷疑，後於新加坡國立大學博士階段，他在導師引導下不斷嘗試和調整，才逐漸建立起對科研的信心與觸覺，「在博士階段，我終於發現自己願意長時間鑽研一個問題，同時想法比較活躍，常常從不同角度重新審視問題。直到今天，我依然很享受經過推理、實驗和不斷修正，像剝洋蔥一樣，一層一層地消除問題中的不確定性。」

在研究歷程中，陽行意有幾次有趣嘗試，都是把原本來自圖演算法或數學優化方法，放到神經網絡中重新使用。其中一個是做一種「可以自動變大小變小的神經網絡」，就像一把能自動伸縮的工具，其困難之處在於如何用數學語言描述「變大」和「變小」之間的關係，因為不同大小的神經



●陽行意研究項目：「非玩家角色（NPC）導向」，通過AI控制遊戲中的NPC角色，改變他們的動作策略。 理大圖片

網絡結構和參數排列都不一樣，不能簡單直接比較。

後來團隊想到用圖優化方法把這些網絡先「對齊」。他比喻說，就像比較幾幢大小和結構都不同的房子，需要先把屋頂和屋頂對齊、樓層和樓層對齊，再找出它們的相似部分，從而發現不同網絡之間共同的結構規律，再嘗試從中生成不同大小的模型，最終在小規模實驗中得到不錯的結果。雖然這方法未成為影響特別廣泛的技術，但陽行意認為，這是他做過最有趣的工作之一，「一個想法未必會變成最成功的成果，但在把它從『異想天開』變成『可以驗證』的過程中，你會真正體會到研究的樂趣。」