

AI眼鏡自動答題 考試制度受衝擊

科大以大三程度考試測試 僅30分鐘獲92.5分名列前五



世界邁進人工智能（AI）急速發展的時代，不同的大模型於學習與評核層面相繼展現了

極強能力，在多種考試專業測試取得優秀成績，隨着人機協同技術發展，AI正為現實教育場景帶來顛覆性的衝擊。香港科技大學團隊早前透過演算法及網絡優化，讓商業智能眼鏡搭載先進的AI模型，成功令其變成具自動答題功能的「考試神器」。團隊安排戴上該AI眼鏡的學生應考大學三年級程度的「計算機網絡原理」專業課考試，並按眼鏡顯示內容轉錄作答，最終僅用30分鐘便在滿分100分中取得92.5分的驚人成績，在全班100多人中名列前茅。團隊認為是次測試引證了AI眼鏡於教育領域的潛力，但亦對學術誠信與教育評估體系公正帶來潛在威脅，建議應立即行動，在技術、監管與道德上取得平衡（見另稿）。 ●香港文匯報記者 史柳藝

在科技急速發展下，AI眼鏡已突破拍照、導航等基礎功能，具備深度知識推理、環境理解及人機互動的多元能力。科大電子及計算機工程學系教授張軍及助理教授孟子立團隊近日開展是次備受關注的實驗，團隊先評估現有的商業智能眼鏡及AI大模型，就拍攝與顯示功能、反應速度等作篩選，藉以實現自動答題，最終以Rokid眼鏡搭載GPT-5.2進行測試。

在實驗期間，戴上AI眼鏡的考生只需要低頭注視試卷，攝像頭快速拍照並進行圖像輸入，學生無須拿出手機操作，試卷照片透過手機App傳輸至遠端大模型，經分析推理得出答案後，再以相同路徑返回，最終清晰顯示在AI眼鏡的螢幕上，供學生轉錄作答。

實驗結果顯示，戴上AI眼鏡的考生，於選擇題和單頁解答題中均取得滿分。面對跨頁解答題時，儘管無法一次性把所有題目看全，但仍拿下了大部分的分數。值得關注的是，即使部分題目最終答案有誤，大模型依舊能給一部分推理過程，甚至還掌握了課上沒有學過的知識點，並提供了正確的答案。

獲97.5分學生耗3小時 與AI速度差距大

在該份考卷中，以AI眼鏡所獲的92.5分，遠超整體考生的72分平均分。雖然仍有學生成功「堅守」人類防線，最高分考獲97.5分，但其耗時為約6倍的3小時，與AI眼鏡的答題速度相比差距明顯。

香港文匯報記者就實驗獨家訪問了兩位科學家。張軍坦言對結果頗感驚喜：「這次實驗主要測試輕便但通信、計算能力受限的AI眼鏡是否能解決考試任務，過程中雖遇到一些技術問題，但結果證明，經過演算法及網絡優化的AI眼鏡已能在真實場景中應用，讓我們看到其短期內的推廣潛力。」

他補充，自身科研方向聚焦移動通信、移動計算及AI，「這個研究很好地揭示了智能眼鏡的應用落地需要很好地融合這三個方面的技術」，實驗結果既「展示了現有硬件和演算法能達到的能力」，也「揭示了還需解決的具體研究問題」。

科學家：已妥善解決實際輸入挑戰

孟子立指出，現有AI大模型解題能力已很強，只要不是數學奧林匹克競賽級別的試卷都能高效應對，獲高分是預期之內，「驚喜則來自於我們妥善解決了眼鏡拍攝不全、圖片模糊等實際輸入挑戰。」

他認為是次應用研究的價值驗證了技術可行性，「在用戶完全沒有和眼鏡交互的情況下，不需要語音對話、按鍵控制等操作，AI眼鏡可直接解題。這種主動式推送，相比現有AI眼鏡的被動式交互，用戶體驗更佳，希望未來能探索更多相關應用場景。」

他同時強調，AI在複雜知識任務中的應用潛力：

「像我們在程式開發中，AI輔助工具已成常態。港科大在亞太地區較早擁抱AI，ChatGPT問世後便部署了自己的GPT並鼓勵師生使用，包括我們還擁有香港自己的生成式AI大模型，未來相關應用只會愈來愈多。」



AI模型：

GPT-5.2 Thinking

場景：2026年日本大學入學統一考試15個主要科目

成績表現：平均得分96.9分（每科滿分100分），其中9個科目獲滿分，同期人類考生平均得分約58分

AI模型：

DeepSeek-V3.2-Speciale

場景：2025年國際數學奧林匹克（IMO）及中國數學奧林匹克

成績表現：達到金牌水平得分（相當於滿分或接近滿分），達到IMO金牌標準

近年各AI模型於不同考試所獲佳績

AI模型：

GPT-4

（2023至2024年進行）

場景：MBE+MEE+MPT 統一律師資格模擬考試

成績表現：400分中獲298分，約排第90百分位（即最佳的10%）

AI模型：

豆包1.6、DeepSeek

場景：2025年山東高考

成績表現：豆包理科獲648分，能考上985高校理工專業，文科獲683分，有望考上清華或者北大，DeepSeek於數學科獲145分

AI模型：

GPT-5、Grok 4、DeepSeek R1

場景：2025年韓國大學修學能力試驗（CSAT）數學科

成績表現：GPT-5獲100分滿分，Grok 4得97.8分，DeepSeek R1得95.7分，人類考生平均獲57.7分

資料來源：綜合各AI模型測試報告
整理：香港文匯報記者 史柳藝

學者用AI眼鏡量化未來教育新「視」界

就今次衝擊考試制度的AI眼鏡實驗，兩位科大學者分別與記者分享項目構思過程及未來預期。事情起因於一次孟子立在考場觀察到有學生佩戴墨鏡，令他萌生「AI是否能實現考試功能」的想法，遂於去年中啟動項目，團隊其後針對延遲控制、交互體驗、模型效果等方面反覆優化調試，最終於去年12月順利完成測試。

孟子立表示：「科技不斷進

步，近兩年AI領域的爆發式增長已不可避免地滲透到日常生活。不論是教育工作者還是科研人員，未來都要學會與AI共存，就像當年Google問世時，我們鼓勵大家運用搜索引擎而非僅依賴圖書館查詢。AI時代下，我們更應將有限精力放在如何在教育、科研中用好AI，積極擁抱這一趨勢，才能進一步推動生產力提升。」

張軍表示，近年AI大模型的突破，深度影響眾多領域，是次研究的出發點是探討AI硬

件對教育的影響，「AI眼鏡考試場景，更是易於量化分析的測試場景。」

「教導學生運用這些新型工具應成為教育的一部分，未來能否有效利用這類工具將愈來愈重要。」他強調，教育一直以來受技術發展的驅動與影響，「我們需思考，如何利用新技術輔助學生學習，同時探討可能的負面影響。針對AI與智能硬件的發展如何改進課程設置是業界開放議題，希望這次研究能給教育界提供新角度。」

多地出現用AI出貓 反作弊「試」不宜遲

具備實時影像辨識及回答功能的AI眼鏡，引發對以新型科技作弊的關注，近期正持續衝擊全球各地的教育及考試體系。

在韓國今年5月舉行的TOEIC英語能力考試中，就首度查獲利用AI眼鏡作弊的案例，兩場考試各發現有考生因佩戴AI眼鏡被發現，涉事者將面臨長達4年的禁考處分。

台灣大學醫學院早前的入學申請甄選試也出現AI眼鏡作弊個案。一名重考生被發現在考試時換上一副黑框大眼鏡及長時間保持低頭姿勢，經檢查後發現眼鏡發燙，確認正使用AI眼鏡，隨即被判定0分並通報處理。內地華南農業大學近期也處理了學生利用智能眼鏡作弊的違規事件，校方通報相關案例並加強考場管理。

為加強監管，於上月舉行的內地高考，多地考點均在進入考場安檢時，要求考生於視頻監控下自行摘下眼鏡，由監考官實施專項檢查。部分考點引入升級的安檢門，可識別帶攝像頭的智能穿戴設備並觸發警報，甚至可偵測關機狀態下的設備，又會使用AI智能

巡查系統識別考生異常行為，由系統自動截取片段供審核員確認。

面對AI新興科技發展，對各地教育及監管單位來說，除被動地應對、加強監管遏止作弊以外，更要從根本開始，反思為何考試與如何考試在衡量學習成果上所扮演的關鍵角色，那將是AI時代新一輪教育變革的核心課題。 ●香港文匯報記者 史柳藝



●AI眼鏡答題測試情況。 課題組視頻截圖

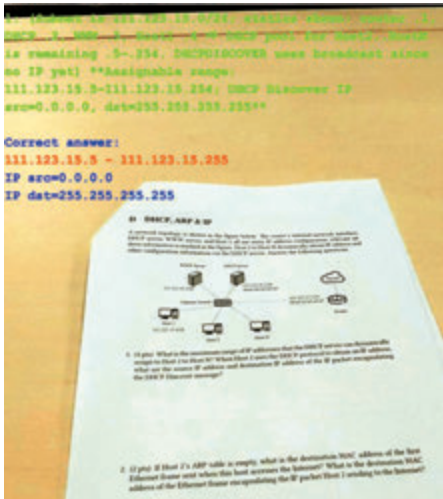
新科技如「雙刃劍」 靠機制監管

AI穿戴技術日趨成熟，今次實驗說明，AI眼鏡憑藉實時信息交互、場景化輔助等功能，正快速走進教育場景，但同時其技術所帶來的作弊可能性，也引起侵蝕學術誠信、衝擊教育評估體系的風險。

負責實驗的科大團隊形容，AI眼鏡在教育領域的崛起，是一把「雙刃劍」：積極影響是其潛在的教育技術創新，如個性化學習為學生提供定製化的學習輔助和資源，降低學習資源不平等現象，幫助需要額外支持的學生提高學術表現。但顯而易見其消極影響同樣嚴重：如考試時使用智能眼鏡將嚴重損害學術誠信，動搖整個教育評估體系的公正性。

籲做好技術監控檢測及教育宣傳

在此重大衝擊下，團隊呼籲教育者和政策制定者必須立即行動，建立合理監管機制，制定明確的政策法



●AI眼鏡答題時的顯示屏。 港科大Spark Lab供圖

技術防範，包括應試前嚴格檢查，禁止將AI眼鏡、手錶等電子產品帶入考場。而長遠來說亦需要深入研究評核形式的變革，例如摒棄傳統集體紙筆考試，轉而採用單對單考核或項目式評估，通過師生共同參與階段任務，強化過程表現評價，適配AI時代的教育評估需求。 ●香港文匯報記者 史柳藝