

新冠肺炎疫情下的急中生智： 人工智慧輔助判讀肺炎

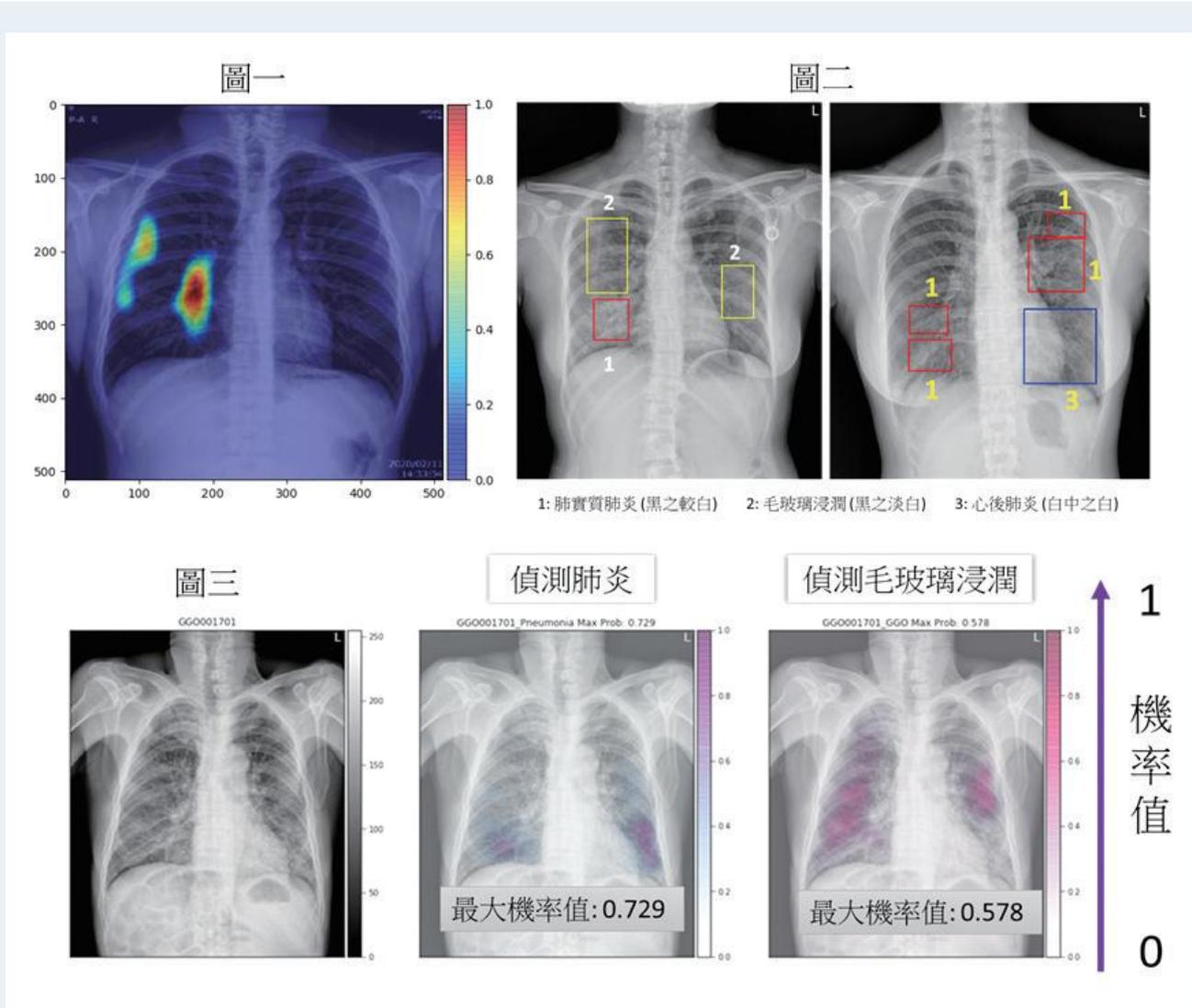
◎文 / 蔡依珊

新型冠狀病毒肺炎疫情肆虐全球，在政府超前佈署及阻絕境外的政策下，台灣為疫情下的清流淨土，人民能過著正常的生活，實屬萬幸。年初新型冠狀病毒肺炎疫情嚴峻，在 SARS 的慘痛經驗下，所有醫護人員都繃緊神經，一天晚上突然接到沈孟儒院長來電，問到目前嘗試架設的 X 光自動分類 AI 模型，能不能導入檢疫站來幫忙，心裡還想著院長腦筋動得真快，竟然把我們還在蘊釀的 AI 小苗，要拱上一線參戰，對抗的還是一個大家都摸不著頭緒的新興傳染病，當時也沒有想那麼多，只有回答院長，理論上十分可行，技術上我們來努力，沒想到一個不小心，踏上艱難的 AI 一條路，現在回想起來，也都不知道自己是哪股傻勁接下這困難的任務。

當然時間相當的緊迫，基於一個 AI 新手，只好請成醫 AI 資訊師先拿了 GitHub 開放源，利用 NIH chest X-ray dataset 訓練好的分類模型先來上手，有了模型也要有本院的資料來微調模型的參

數，用最短的時間召集影醫部的總醫師群及 AI 小組成員，收集了 3000 筆左右的成醫肺炎及正常案例，終於可以來校調模型。這時候問題來了，胸部 X 光判讀因照相品質差異性較大，本來就帶有灰色模糊地帶，當我們導入模型後發現模型怎樣校調都不好，這時想要解決的時候，的確遇到 AI 分類模型的黑盒子迷思，到底 AI 看到什麼了，於是我們把 AI 分類模型掛上熱區圖 (heatmap)，試著了解 AI 分類模型到底關注了那裡。不看還好，看了心涼了一大半，AI 判讀的圖都看了肩頰骨還有肺門血管（圖一），這模型真心不敢用在病人身上，特別是還要用在檢疫前線上。

幾經思索只好向蔣榮先資訊長的團隊求援了，既然單純的分類模型沒法符合臨床的使用習性，我想還是要回歸醫師判讀影像的邏輯，我們是在肺部裡找尋有浸潤的肺部才會高度懷疑有肺炎，接下來要用 segmentation 模型來挑戰，試著讓模型學到肺部然後要學到有生病的肺。起



了心動了念，AI 模型最麻煩的其中一部分就來了，我們需要有經驗的醫師來標註 (label) 病兆，以便能讓模型好好學習，這一動作可就又累慘一坨醫師，真是印證一句話，要 AI 前一定非常不 AI，所需投注的人力成本不斐。

但是標註其實是一門很大的學問，肺炎在胸部 X 光片上並不是只有一種型態，我們可以把它分成黑中之白（正常肺部

內的浸潤）、白中之白（心臟後方肺部的浸潤），最難標註還有所謂的毛玻璃浸潤（正常肺部內的淡淡輕微浸潤），而這種剛好是新冠肺炎早期的變化，經過幾次標註的試誤，最後終於將標註分成三類（圖二），希望能有好的訓練模型成效。最終，幾經波折終於做出了自己能信服也自己敢用的模型，因為在學理上站得住腳，也真真實實的用 heatmap 去呈現模型所

看的關注區（圖三），用在檢疫站裡至少能有效的偵測出有浸潤的肺炎，後來我們統計檢疫站的 AI 模型輔助診斷結果，並與影像醫學部醫師報告來比對，我們輔助偵測肺炎達到 83% 的準確率。

到了這個階段，雖然對自己團隊的模型比較放心，但面對別人的質疑，“這模型真的能診斷新冠肺炎嗎？”，其實不用別人我自己也很質疑，但礙於台灣的案例數，我們很難證實這件事。隨著疫情發展，各國越來越多案例累積，也有人釋放了一些新冠肺炎病例圖像到 GitHub，我想真是一個好機會來驗證挑戰這個成大團隊灌溉出來的模型，我們將開放源的圖像導入模型做為測試，以我為判定有無新冠肺炎來當作標準，模型偵測出肺炎的準確率達到 89%，這時我才真的放下心中的大石，我們的模型不是一個包裹著 AI 糖衣的花拳繡腿模型。當我們完成新冠肺炎測試的同時，由成大資訊工程系特聘教授蔣榮先，帶領碩博士生王麒詳、邱煌鑛、吳昭儀以及我組成的隊伍 MedCheX，參加「國際 COVID19 科技防疫黑客松大賽」，從 1560 個團隊中脫穎而出的好消息也傳來，是台灣唯一獲獎團隊為國爭光。

以人工智慧輔助判讀肺炎並非是來取

代醫師的工作，而是全年無休輔助判讀小幫手，這樣的協助判讀能量，對於人力較缺乏的區域醫院，更是彌足珍貴。特別在防疫期間，一線的醫事人員可以藉由 AI 輔助判讀系統獲得初步判讀建議，影像醫學部醫師則在接到簡訊警示有新冠肺炎報告待判讀後，可提升擅打報告的優先處理次序，加快報告流程，協助病人分流處置。

參與人工智慧輔助判讀肺炎，可說是有苦有樂，承受極大壓力下，也做出了能實際導入臨床服務的輔助工具，也從這整個過程中學習到，說 AI 很容易，但請不要隨意啟動一個 AI 計畫，務必確定好臨床使用的面向，使用者對象要廣泛才能帶來足夠的效益，因為一個 AI 模型都要需要團隊好好的灌溉與討論，才有機會做出一個落地實用度高的模型。