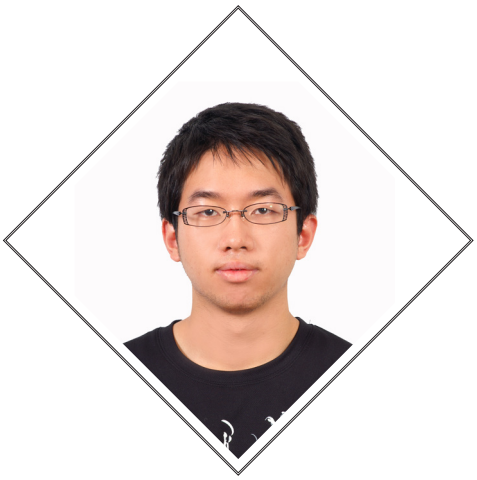

劉家榮 醫師
臺灣大學醫學院附設醫院新竹台大分院

人工智慧在分枝桿菌肺病影像判讀之應用

淺談人工智慧

2016 年，擁有多座國際圍棋賽事冠軍的韓國棋手李世石與人工智慧 AlphaGo 展開了五場圍棋對弈，最終結果相信讓許多人都大吃一驚，AlphaGo 以四比一的絕對優勢獲勝，在比賽結束之後，李世石甚至說出：「隨著人工智慧在圍棋比賽的登場，我意識到即使成為第一名，我也不會是最頂尖的。」經過這場對弈之後，也讓大家深刻意識到，人工智慧已經在一些領域上，有著比人類更好的表現。

人工智慧，簡單來說就是讓機器展現類似人類的智慧 [1]，而機器模型如何能展現類似人類智慧的方法又包含了機器學習和深度學習 [2]。以「機器學習」來說，今天我們如果想要訓練一個人工智慧模型來預測病人將來中風的機率是多少，我們需要先收集一組有中風跟一組沒中風的病人資料，接著透過人類的知識，將可能能夠拿來預測病患會不會中風的特徵（例如：年紀、性別、共病症）輸入訓練模型中，接著模型就會依據這些特徵去做運算及分類，最終得到病患中風的機率 [3]。舉個例子來說，大家很熟悉的 Logistic Regression 也可視為一種最簡單的機器學習模型。而「深度學習」算是機器學習的一個分支，它的設計理念是透過模擬人類的大腦神經網路，建立了多層次複雜的人工神經網絡模型，進而能產生遠比一般機器學習模型更強大的學習過程 [4]。「深度學習」的運作模式與「機器學習」最大的不同是，它不再需要依賴人類進行特徵擷取，我們人類只需要把資料輸入到訓練模型中，接著模型會自己進行特徵擷取、運算及分類，最終得到預測答案 [1]。

有了機器學習和深度學習這兩種不同的訓練模型方式之後，大家接著就想知道哪種學習模式比較好，

因而許多競賽便接踵而來，ILSVR (ImageNet Large Scale Visual Recognition Competition) 是由 ImageNet 所舉辦的大規模視覺識別挑戰賽，這個競賽會利用許多的圖片和影像，去測試人工智慧模型判讀圖像的能力，這個競賽每年都會舉辦，但從 2012 年之後，所有獲勝的人工智慧模型都是由深度學習的方式所建構，因此在近年來，若要判讀影像，深度學習所建構的人工智慧被視為較佳的模型 [5]。

人工智慧應用於肺結核之影像判讀

至今，肺結核仍是一個在世界上造成許多死亡的重要感染性疾病 [5]。一旦我們延遲幫病人診斷出肺結核，會增加傳染風險，也可能讓病人疾病嚴重度增加，造成治療更為困難，因此針對結核病高風險族群，結核篩檢是非常重要的 [6]。

根據世界衛生組織 (WHO) 和我們台灣的結核病診治指引，胸部 X 光檢查都是一項極為重要的結核篩檢工具，過去的研究指出，利用胸部 X 光檢查來偵測肺結核病灶的 sensitivity 和 specificity 可以高達 87% 和 89%[7, 8]，儘管胸部 X 光檢查的診斷表現良好，它仍然存在一些現實上的問題，像是要正確判讀胸部 X 光片常常需要胸腔科或影像科的專家，且判讀也需耗費許多時間和精力，而這樣的影像判讀專家在醫療資源匱乏的區域常常也會相對缺乏，導致無法進行適當的結核篩檢 [9]。為了因應這樣的臨床問題，在過去幾年間，學者們研發了許多人工智慧模型，希望透過這些模型自主判讀胸部 X 光片來抓出肺結核病患。舉其中的一個研究為例 [10]，此研究是在尼泊爾及喀麥隆進行，研究團隊收集了因懷疑肺結核而接受胸部 X 光片及 Xpert MTB/RIF 檢查的病人資料，並將這些胸部 X 光片輸入給三個利用深度學習所建構的人工智慧模型判讀（CAD4TB、qXR 及 Lunit INSIGHT），接著把人工智慧模型的判讀表現拿來和放射科醫師比較。此研究最終顯示，三個人工智慧模型對於肺結核都有極佳的鑑別力 (Area under the ROC Curve: 0.92-0.94)，更有趣的是，不論跟資深或資淺的放射科醫師比較，三個人工智慧模型都有著比放射科醫師更好的肺結核鑑別力。

針對肺結核的胸部 X 光片篩檢，當大家發現人工智慧有比人類更好的表現之後，下一個問題就是人工智慧的好表現能不能拿來輔助人類，增進人類專家的肺結核診斷能力，因此在韓國就進行了這樣的一個研究 [11]，先讓人類專家判讀胸部 X 光片來判斷病患是否為肺結核，接著提供人工智慧模型判讀的結果給人類專家知道，後續再請人類專家作答一次，最終就比較人類專家在知道人工智慧預測結果之前和之後的肺結核診斷率是否有差異。此研究最終顯示出，針對非影像科的醫師，人工智慧可以顯著增加醫師診斷肺結核的診斷率和肺結核病灶的偵測率；而針對影像科醫師，人工智慧雖然不能增加醫師的診斷率，但仍可以顯著增加肺結核病灶的偵測率。

人工智慧應用於區分肺結核及非結核分枝桿菌肺病之影像判讀

近年來，已開發國家的肺結核盛行率是逐年下降；但相反的，非結核分枝桿菌肺病的患者卻有逐年上升的趨勢 [12]。然而醫療人員對於非結核分枝桿菌肺病的認知仍較少，導致了有些非結核分枝桿菌肺病的患者

在細菌培養結果出來前，就接受了肺結核的經驗性治療[13]，這樣的不適當治療可能會增加非結核分枝桿菌的抗藥性，導致後續治療更為困難[14]。此外，非結核分枝桿菌肺病患者不用進行空氣隔離和接觸者調查，若誤判為肺結核，也會導致不必要的醫療資源的浪費[15]。

因此，及時區分肺結核與非結核分枝桿菌肺病是非常重要的，然而這兩種疾病的臨床症狀和影像學變化都很相近，臨床上非常難以區分，大多還是要靠細菌培養的結果才能分辨，但細菌培養往往要花好幾週的時間才會有最終結果，導致目前要及時區分肺結核與非結核分枝桿菌肺病仍非常困難[16]。

為了解決這個臨床問題，學者們也開始嘗試研發人工智慧模型，利用判讀胸部影像來區分肺結核與非結核分枝桿菌肺病。在中國進行的一個研究指出[17]，使用深度學習所建構的人工智慧模型來判讀胸部電腦斷層影像，對於區分肺結核與非結核分枝桿菌肺病有相當不錯的鑑別力 (Area under the ROC Curve: 0.78-0.86)，且這個表現也比放射科醫師更好，更有趣的是，此研究也有進一步分析，判讀一位病人電腦斷層影像所需花費的時間，放射科醫師大約是 1 分鐘，而人工智慧模型只需 0.05 秒。

結論

隨著人工智慧的快速發展，人工智慧不再像過往一樣充滿距離感，反而已在各領域慢慢普及化。在分枝桿菌肺病的影像學判讀領域中，人工智慧模型在這幾年也開始展現了其強大的能力，甚至有著比人類專家更好的表現，不過值得一提的是，這些研究往往都是塑造了一個很合適人工智慧運作的環境，所以這些人工智慧模型能不能在現實世界中也維持這麼好的表現，仍待將來更多的研究來驗證。儘管如此，目前人工智慧模型發展的成果，仍是很令人期待，希望在不久的將來，人工智慧真的可以在分枝桿菌肺病的影像學判讀扮演一個重要角色，進一步幫助人類專家。

參考文獻

1. Y. Mintz, R. Brodie, Introduction to artificial intelligence in medicine, Minim Invasive Ther Allied Technol 28(2) (2019) 73-81.
2. R.Y. Choi, A.S. Coyner, J. Kalpathy-Cramer, M.F. Chiang, J.P. Campbell, Introduction to Machine Learning, Neural Networks, and Deep Learning, Transl Vis Sci Technol 9(2) (2020) 14.
3. Y. Liu, P.C. Chen, J. Krause, L. Peng, How to Read Articles That Use Machine Learning: Users' Guides to the Medical Literature, JAMA 322(18) (2019) 1806-1816.
4. Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521(7553) (2015) 436-44.
5. P. Lakhani, B. Sundaram, Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks, Radiology 284(2) (2017) 574-582.
6. S. Paramasivam, B. Thomas, P. Chandran, J. Thayyil, B. George, C.P. Sivakumar, Diagnostic delay and associated factors among patients with pulmonary tuberculosis in Kerala, J Family Med Prim Care 6(3) (2017) 643-648.
7. S. den Boon, N.W. White, S.W. van Lill, M.W. Borgdorff, S. Verver, C.J. Lombard, E.D. Bateman, E. Irusen, D.A. Enarson, N. Beyers, An evaluation of symptom and chest radiographic screening in tuberculosis prevalence surveys, Int J Tuberc Lung Dis 10(8) (2006) 876-82.
8. A.H. van't Hoog, H.K. Meme, K.F. Laserson, J.A. Agaya, B.G. Muchiri, W.A. Githui, L.O. Odeny, B.J. Marston, M.W. Borgdorff, Screening strategies for tuberculosis prevalence surveys: the value of chest radiography and symptoms, PLoS One 7(7) (2012) e38691.
9. A.H. Hoog, H.K. Meme, H. van Deutekom, A.M. Mithika, C. Olunga, F. Onyino, M.W. Borgdorff, High sensitivity of chest radiograph reading by clinical officers in a tuberculosis prevalence survey, Int J Tuberc Lung Dis 15(10) (2011) 1308-14.
10. Z.Z. Qin, M.S. Sander, B. Rai, C.N. Titahong, S. Sudrungrot, S.N. Laah, L.M. Adhikari, E.J. Carter, L. Puri, A.J. Codlin, J. Creswell, Using artificial intelligence to read chest radiographs for tuberculosis detection: A multi-site evaluation of the diagnostic accuracy of three deep learning systems, Sci Rep 9(1) (2019) 15000.
11. E.J. Hwang, S. Park, K.N. Jin, J.I. Kim, S.Y. Choi, J.H. Lee, J.M. Goo, J. Aum, J.J. Yim, C.M. Park, D. Deep Learning-Based Automatic Detection Algorithm, G. Evaluation, Development and Validation of a Deep Learning-based Automatic Detection Algorithm for Active Pulmonary Tuberculosis on Chest Radiographs, Clin Infect Dis 69(5) (2019) 739-747.

12. O.M. Rivero-Lezcano, C. Gonzalez-Cortes, M. Mirsaeidi, The unexplained increase of nontuberculous mycobacteriosis, Int J Mycobacteriol 8(1) (2019) 1-6.
13. M. Maiga, S. Siddiqui, S. Diallo, B. Diarra, B. Traore, Y.R. Shea, A.M. Zelazny, B.P. Dembele, D. Goita, H. Kassambara, A.S. Hammond, M.A. Polis, A. Tounkara, Failure to recognize nontuberculous mycobacteria leads to misdiagnosis of chronic pulmonary tuberculosis, PLoS One 7(5) (2012) e36902.
14. X. Wang, H. Li, G. Jiang, L. Zhao, Y. Ma, B. Javid, H. Huang, Prevalence and drug resistance of nontuberculous mycobacteria, northern China, 2008-2011, Emerg Infect Dis 20(7) (2014) 1252-3.
15. C. Swenson, C.S. Zerbe, K. Fennelly, Host Variability in NTM Disease: Implications for Research Needs, Front Microbiol 9 (2018) 2901.
16. W.J. Koh, C.M. Yu, G.Y. Suh, M.P. Chung, H. Kim, O.J. Kwon, N.Y. Lee, M.J. Chung, K.S. Lee, Pulmonary TB and NTM lung disease: comparison of characteristics in patients with AFB smear-positive sputum, Int J Tuberc Lung Dis 10(9) (2006) 1001-7.
17. L. Wang, W. Ding, Y. Mo, D. Shi, S. Zhang, L. Zhong, K. Wang, J. Wang, C. Huang, S. Zhang, Z. Ye, J. Shen, Z. Xing, Distinguishing nontuberculous mycobacteria from Mycobacterium tuberculosis lung disease from CT images using a deep learning framework, Eur J Nucl Med Mol Imaging 48(13) (2021) 4293-4306.

[試題]

1. 在人工智慧中，機器模型是如何學習到類似人類智慧的能力？
   (A) 機器學習。
   (B) 深度學習。
   (C) 以上皆是。
   (D) 以上皆非。

2. 根據目前的研究，針對人工智慧在分枝桿菌肺病影像判讀之應用，以下何者敘述有誤？
   (A) 目前人工智慧已可完全取代醫師。
   (B) 人工智慧可能有著比醫師更好的診斷力。
   (C) 在有人工智慧的幫忙下，可能可以增加醫師的診斷率。
   (D) 相較於醫師，人工智慧判讀電腦斷層影像所需花費的時間較少。

3. 關於肺結核與非結核分枝桿菌肺病，以下何者敘述有誤？
   (A) 這兩種疾病的臨床症狀和影像學變化都很相近，臨床上非常難以區分，大多還是要靠細菌培養的結果才能分辨。
   (B) 已開發國家的肺結核盛行率是逐年下降；但相反的，非結核分枝桿菌肺病的患者卻有逐年上升的趨勢。
   (C) 肺結核患者須進行空氣隔離和接觸者調查，但非結核分枝桿菌肺病患者則不用。
   (D) 這兩種疾病的治療相同，因此無法及時區分肺結核與非結核分枝桿菌肺病也沒關係。