

# 人工智慧應用於皮膚疾病診斷的最新進展

<sup>1</sup>臺北榮民總醫院家庭醫學部 <sup>2</sup>國立臺灣大學電信工程學研究所 <sup>3</sup>臺北榮民總醫院皮膚部  
楊博欽<sup>1</sup> 陳育群<sup>1</sup> 施美如<sup>2</sup> 張雲亭<sup>3</sup>

## 前言

近年來，「人工智慧」、「機器學習」等名詞在報章雜誌中隨處可見，甚至不少新創事業應運而生，有些人採取觀望態度，有些人認為劃時代即將來臨，人工智慧將顛覆我們的生活。

「人工智慧」在過去的幾十年間於學術圈中醞釀發展，近年來的研究成果顯示人工智慧在各領域都有很好的應用。比較知名的例子像是人工智慧無人駕駛車、人工智慧圍棋程式「AlphaGo」打敗人類的世界冠軍，都曾大幅佔據了媒體的版面。在醫學領域也不例外，包括疾病的診斷、疾病預後的預測等，皆有許多新興的研究，甚至有人開玩笑地說，許多醫師的角色將被「人工智慧」所取代。醫學界對於「人工智慧」領域的發展普遍比較生疏，而本文將介紹「人工智慧」的基本概念，以及其於皮膚疾病的最新研究與應用。

## 人工智慧、機器學習、深度學習的基本概念

人工智慧(artificial intelligence, AI)，又稱為機器智慧(machine intelligence)，是指由機器所表現出來的智慧<sup>1</sup>。其核心問題包括建構和人類相似，甚至超越的推理、知識、規劃、學習、交流、感知、移動和操作物體等能力<sup>2</sup>。人工智慧的主要精神在於：如何使用電腦，來完成本來需由人類執行才能完成的工作。

機器學習(machine learning)，是人工智慧的一個分支，和統計學密切相關。藉由設計讓電腦可以「自主學習」的演算法，使機器可以

從一群資料或數據中自動分析，獲得規律，並利用規律對未知的資料進行預測，而非手動編寫含有特定指令的程式來完成某個特殊任務。傳統的機器學習需要設計一個特徵提取器(feature extractor)，將原始資料轉換為可以用來學習的形式或向量(vector)，再藉由分類器(classifier)來偵測或分類其他輸入的資料<sup>3</sup>。而分類(classification)方式有許多種，包括聚類(clustering)、關聯規則分析(association rule mining)、支持向量機(support vector machines)等。

機器學習的方式可分為監督學習(supervised learning)和非監督學習(unsupervised learning)。監督學習是最常見的方式，在給予機器一個輸入(input)後，同時也給予一個相對應的輸出(output)，訓練的過程中有告訴機器答案，是「有標籤」的資料。非監督學習則是給予機器一個輸入後，並沒有給予標準答案，讓機器自己摸索出輸入與輸出間的規律。

深度學習(deep learning)，是機器學習的其中一種，其使用了包含複雜的結構或者多重非線性函數構成的多個處理層，對資料進行高層抽象的演算法<sup>3</sup>。藉由這種演算法，可以強化輸入的某些重要部分，並抑制不重要的部分，以利辨識處理更複雜、更多維的資料。

在深度學習的眾多模型中，人工神經網路ANN (artificial neural network)，簡稱神經網路(neural network)，或是類神經網路，是一種受生物中樞神經系統的結構和功能所啟發而

產生的非線性數學計算模型<sup>4</sup>。典型的人工神經網路為一系列的處理單元所組成的多層神經元(neuron)結構，包括輸入層(input layer)、隱藏層(hidden layer)、輸出層(output layer)。其中，隱藏層可以有多層，數目越多則非線性越顯著。神經元之間的權重(weight)則用以表示其連結強度，概念類似於神經傳導時的突觸電位<sup>4</sup>。

和傳統的系統相比，ANN最大的不同點是它具有自動學習過程(training algorithm)。在訓練過程的開始，並沒有給予任何明確的資訊，而是給予大量已知結果的訓練樣本，其神經元間的權重會根據訓練演算法不斷改變，並通過大量訓練樣本的校正，對各個層的權重進行校正(learning)，以減少錯誤<sup>4</sup>。ANN已被廣泛使用於語音辨識、影像辨識、手寫辨識、分類、預測、記憶聯想等領域，在醫學上則被應用於藥物設計、藥物流病學、臨床診斷，以及疾病預後的預測<sup>5</sup>。

卷積神經網路CNN (convolutional neural network)是目前ANN領域的發展主力，最早於1998年被提出，使用於辨認草寫的數字<sup>6</sup>，目前則廣泛使用在物體的辨認，尤其是圖片的辨認與分類。CNN與傳統的多層網路最大的差異在於多了卷積(convolutional)及池化(pooling)兩層結構，強化了模式辨識(pattern recognition)和相鄰資料間的關係，用以維持形狀資訊，並且避免參數大幅增加。CNN在影像、聲音等訊號類型的資料型態皆有非常好的應用，甚至已做到比人類還精準的程度<sup>7</sup>。

## AI過去於皮膚疾病的應用

在過去幾十年來，電腦輔助的皮膚病灶識別系統的研究主要是針對皮膚癌的早期診斷，更精確地說，是惡性黑色素細胞瘤辨認。也有少數被用在辨認日光性角化病<sup>8</sup>、乾癬<sup>9</sup>、痤瘡<sup>10</sup>的研究。

全球每年大約有2-3百萬的非黑色素細胞皮膚癌，以及約23萬的惡性黑色素細胞瘤的發生<sup>11</sup>。其中，表皮黑色素細胞瘤(cutaneous melanoma)是皮膚癌中常見的一種，以高度惡性的表現聞名，但若早期診斷早期治療，有機會治癒，而不留下後遺症<sup>12</sup>。然而，黑色素瘤可以長在任何部位的皮膚，外觀可能是扁平表淺的暗色斑塊，也可能以突起的結節腫瘤來表現。因此，早期黑色素細胞瘤和其他色素性皮膚病變有時相當難以鑑別診斷，即使是有經驗的皮膚科專科醫師<sup>13</sup>。而當外觀無法明確分辨，又有懷疑有皮膚癌的可能性的時候，就需要切片及顯微鏡檢查。

從前只能用肉眼來辨別皮膚疾病，而在1958年，最早的可攜帶式的皮膚透光顯微鏡(dermatoscopy或dermoscopy)，簡稱皮膚鏡問世<sup>14</sup>，供臨床檢查皮膚病灶使用。目前被廣泛使用的則是數位皮膚透光顯微鏡(digital epiluminescence microscopy, ELM)，是一種非侵入性、in vivo的技術，利用油浸的光學現象，使得皮膚表面的結構可以放大至10至40倍<sup>14</sup>，提供臨床醫師更精準的判斷。近年來有許多關於電腦輔助診斷皮膚惡性疾病(automated skin lesion diagnosis, ASLD)，即是利用直接拍

攝的皮膚影像或是藉由皮膚鏡所拍攝的影像，經過影像處理和機器學習的演算法，來幫助臨床醫師早期診斷皮膚腫瘤。

大部分的ASLD是依標準化的電腦輔助診斷 (computer-aided diagnosis, CAD) 進程序<sup>15</sup>。首先要獲得影像(image acquisition)，並進行資料處理(data preprocessing)，例如：偵測並排除會影響判讀的干擾訊號(例如毛髮、鏡檢照片的泡泡)，再將病灶處和正常健康的皮膚分開(segmentation)，接下來，病灶的特徵會被擷取(feature extraction)。擷取出來的病灶特徵，再根據一些設定的規則來進行分類(classification)。其中，監督學習通常只有在最後分類的步驟被使用，藉由呈現給電腦許多的樣本和參數而學習，再讓電腦幫助判讀不曾看過的影像。

過去相關的重要技術有依據大小、形狀、顏色、邊界來分類皮膚病灶<sup>16</sup>，也有根據病灶的形態特色來分類<sup>17</sup>。此外，也有根據面積、周長、甚至病灶顏色的波長與紅、藍、綠顏色梯度來分類的研究<sup>18</sup>。2007年時，有了結合ABCD規則來幫助分類皮膚疾病的電腦輔助診斷系統<sup>19</sup>，A = Asymmetry不對稱性，B = Border邊界，C = Color variation顏色差異，D = Diameter直徑。在此之後，有研究提出用皮膚影像的小波分析係數(wavelet coefficients)來辨別疾病種類<sup>20</sup>。也有人提出了以ABCD規則為基礎，再加上調整過加權計算過的全皮膚鏡分數(total dermoscopy value, TDV)來幫助判別<sup>21</sup>，或是加上病灶是否突起(E = Elevation)

這個參數來幫助判別<sup>22</sup>。後來，有研究提出了使用系統發生樹(phylogenetic trees)的方法來偵測黑色素細胞瘤<sup>23</sup>。後續也有許多研究致力於除了ABCD規則外，其他能夠用來當做診斷皮膚腫瘤的準則，包括模式分析(pattern analysis)、Menzies方法(Menzies Method)、七點校驗(seven-point checklist)、質地分析(texture analysis)等<sup>12</sup>。同時，也有許多研究致力於影像處理的技術，以及量化各項形狀、邊界、顏色等為基礎的特徵，目標都是為了達到更準確的分類，以幫助臨床的診斷與治療。

然而，由於皮膚病灶的呈現形式太多樣，且資料庫通常太小，因此，電腦輔助診斷在臨床決策上的角色還是很有限。

近年來，由於大數據資料庫的發展，有了更足夠、大量的資料來訓練神經網路，以及硬體的進步，包括圖像處理單元與計算技術的發展，加上多種新的深度學習演算法，又有了許多嶄新的研究成果<sup>24</sup>。以下將分別介紹三個近年來相當有代表性的研究。

### 史丹福大學的皮膚腫瘤研究

史丹福大學研究團隊在2017年2月在Nature雜誌發表了非常創新的研究成果<sup>25</sup>，他們運用了CNN (卷積神經網路)，訓練出了可以分辨一般良性皮膚病灶與皮膚癌的人工智慧。這個研究中，使用了包含129,450張包含了由皮膚科醫師標記的皮膚疾病訓練影像的資料庫，其中包含了2,032種不同的皮膚疾病。並且將疾病以樹狀結構整理分類成757組，來訓



練名為GoogleNet Inception V3 的人工神經網路<sup>25</sup>。

研究團隊在開始時進行了兩次測試，來驗證CNN (卷積神經網路)的能力。第一次是三類型疾病分類測試(three-class disease partition test)，讓電腦區分良性腫瘤、惡性腫瘤、非腫瘤的皮膚疾病。在此階段，CNN達到了 $72.1 \pm 0.9\%$ 的正確率，而兩位皮膚專科醫師的正確率分別為65.56%與66.0%。第二次測試是九類型疾病分類測試(nine-class disease partition test)，讓電腦可以學習各類的疾病以及其中治療計劃的相似處。在此階段，CNN達到了 $55.4 \pm 1.7\%$ 的正確率，而同兩位皮膚專科醫師的正確率分別為53.3%與55.0%。然而，在這兩階段的測試所用的影像皆是由皮膚科醫師標記，不全然有切片證實，因此尚無法得到CNN能正確診斷疾病的結論，但已充分顯示卷積神經網路具有強大的學習能力<sup>25</sup>。

最後階段，則只使用了經切片證實的影像，來測試卷積神經網路及皮膚專科醫師能否分辨惡性或良性的表皮來源病灶(keratinocyte carcinoma或是benign seborrheic keratosis)，以及惡性或良性的黑色素細胞來源病灶(malignant melanoma或是 benign nevus)。並且計算其敏感性(sensitivity)與特異性(specificity)，再比較與21位皮膚專科醫師在區分病灶能力之差異。最終結果顯示，在判斷良性惡性、表皮來源或黑色素細胞來源的正確率方面，CNN在平均表現上贏過皮膚專科醫師<sup>25</sup>。

這個結果代表了人工智慧在區分良性或惡性疾病的能力不亞於皮膚專科醫師，而CNN的確可以幫助更快速精準地診斷臨床艱難的皮膚病灶。此項技術在臨床的實際運用情形仍需進一步研究，相同的概念亦可應用於眼科、耳鼻喉科、放射科及病理科的影像上<sup>25</sup>。

### 韓國的灰指甲研究

韓國的研究團隊則是創造了一個指甲影像的標準化資料庫，並發表在2018年1月發表在PLOS ONE期刊<sup>26</sup>。由於指甲的照片影像並沒有標準的格式，且同一影像中可能包含了健康的指甲與受真菌感染的指甲，於是研究團隊使用了R-CNN (region-based convolutional neural network)這個技術。R-CNN是CNN的一種，其特色為可以辨別影像中病灶的位置，並且將其從背景中分離出來<sup>27</sup>。利用R-CNN定位影像中特定物件的能力，來產生更多可以用來訓練CNN模型的影像。由於CNN在訓練資料庫的影像數目越大時，準確率就會越高，這個研究便是使用了另一個人工智慧(R-CNN)來解決原本資料庫樣本缺乏的問題，並使原本的CNN得到更好的辨識能力。

影像庫有三個主要的來源，每個影像皆有皮膚科專科醫師進行標記。完成的資料庫中包含了49,567張影像。被訓練的CNN網路是微軟公司(Microsoft)的ResNet-152 (extremely deep 152-layer CNN model) 與牛津大學的VGG-19 (visual geometry group)，主要任務是辨別指甲是否遭受真菌感染。

研究結果顯示，藉由R-CNN，可以從現有的影像，建立更大的資料庫，並增加CNN的準確率。經由此種深度學習的人工智慧，在辨別指甲是否遭受真菌感染的能力優於42位皮膚科專家組成的小組(其中包括16名教授、18名臨床醫生和8名住院醫師)<sup>26</sup>。

該研究團隊另外提供了一個網站(<http://nail.medicalphoto.org>)，提供任何人從電腦或是行動通訊裝置上傳照片進行診斷測試，並藉此發現臨床使用者可能遇到的問題，希望藉此發展出更有臨床實用性的演算法<sup>26</sup>。

### 德美法團隊的黑色素細胞瘤研究

另一個由德國、美國、法國組成的研究團隊，也於2018年5月在Annals of Oncology發表了他們的研究成果<sup>28</sup>。這個研究也是用了深度學習的方式，以10萬多張的影像來訓練Google's Inception v4 這個CNN來辨別皮膚的黑色素細胞瘤，並和58名皮膚科醫師來比較辨別的精確率。

首先，研究人員先創造了一個含有300張影像的測試模組，這300張影像皆是從德國海德堡大學皮膚科資料庫取得的高畫質、驗證過的影像，其中有20%的黑色素細胞瘤，80%的良性黑色素痣以及其他型態類似黑色素細胞瘤的疾病。之後，由這300張照片中選出100張，分別由CNN以及來自17個國家的58名皮膚科醫師(其中30名為臨床經驗超過5年的專家)來進行判讀。皮膚科醫師的判讀分為兩階段，在第一階段的判讀測試中，只有提供皮膚鏡的照

片，而受測試者必須下診斷並決定治療計畫；而相同的受測試者在四週後進行第二階段的判讀測試，除了100張相同的影像外，還提供了相關的臨床資訊<sup>28</sup>。

研究結果發現，在第一階段的判讀測試中，皮膚科醫師的平均敏感性為86.6% (+/-9.3%)，特異性為71.3% (+/-11.2%)，而在第二階段的判讀測試中，有了更多的臨床資訊幫助診斷，皮膚科醫師的平均敏感性提升為88.9% (+/-9.6%)，特異性提升為75.7% (+/-11.7%)。而在同樣的敏感性設定下，CNN的特異性為82.5%<sup>28</sup>。由此可知，CNN經由訓練後的診斷能力比大部份皮膚科專科醫師更為傑出，但並非全部。對一般醫師而言，的確有可能藉由CNN影像分類辨識的幫助而更正確地診斷臨床皮膚疾病。

### 結語

目前已經有許多和「人工智慧」有關的臨床研究正在蓬勃發展中。其中，也已經有許多商品化的產品。目前看來，「人工智慧」尚無法完全取代醫師的工作，但的確可以大力輔助醫師的臨床判斷，尤其對於某領域的非專科的醫師，可以藉由「人工智慧」來彌補較為不足的臨床經驗，進而提升診斷疾病的精確率。期盼未來會有更多貼近臨床需求的研究與產品，使有需要的病人能夠得到最即時、適當的治療。

### 參考文獻

1. Nilsson NJ: Artificial Intelligence: A New

- Synthesis. 1998. Morgan Kaufmann.
2. Russell S, Norvig P: Artificial Intelligence: A Modern Approach, 3rd. 2010. Upper Saddle River, New Jersey. Prentice Hall.
  3. LeCun Y, Bengio Y, Hinton G: Deep learning. *Nature* 2015; 521: 436-44.
  4. Renders JM, Simonart T: Role of artificial neural networks in dermatology. *Dermatology* 2009; 219: 102-4.
  5. Patel JL, Goyal RK: Applications of artificial neural networks in medical science. *Curr Clin Pharmacol* 2007; 2: 217-26.
  6. LeCun Y, Bottou L, Bengio Y, et al: Gradient-based learning applied to document recognition. *Proc IEEE* 1998; 86: 2278-324.
  7. Krizhevsky A, Sutskever I, Hinton GE: Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst* 2012; 25: 1097-105.
  8. Spyridonos P, Gaitanis G, Likas A, et al: Automatic discrimination of actinic keratosis from clinical photographs. *Comput Biol Med* 2017; 88: 50-9.
  9. Shrivastava VK, Londhe ND, Sonawane RS, et al: Computer-aided diagnosis of psoriasis skin images with HOS, texture and color features: A first comparative study of its kind. *Comput Methods Programs Biomed* 2016; 126: 98-109.
  10. Khan J, Malik AS, Kamel N, et al: Segmentation of acne lesion using fuzzy C-means technique with intelligent selection of the desired cluster. *Conf Proc IEEE Eng Med Biol Soc* 2015; 2015: 3077-80.
  11. World Cancer Report 2014. World Health Organization. 2014; Chapter 5.14.
  12. Maglogiannis I, Doukas CN: Overview of advanced computer vision systems for skin lesions characterization. *IEEE Trans Inf Technol Biomed* 2009; 13: 721-33.
  13. Pariser RJ, Pariser DM: Primary care physicians' errors in handling cutaneous disorders. A prospective survey. *J Am Acad Dermatol* 1987; 17: 239-45.
  14. Argenziano G, Soyer HP: Interactive Atlas of Dermoscopy. 2000. Book and CD/Web Resource, Edra Medical Publishing and New Media, Milan, Italy.
  15. Wighton P, Lee TK, Lui H, et al: Generalizing common tasks in automated skin lesion diagnosis. *IEEE Trans Inf Technol Biomed* 2011; 15: 622-9.
  16. Green A, Martin N, Pfitzner J, et al: Computer image analysis in the diagnosis of melanoma. *J Am Acad Dermatol* 1994; 31: 958-64.
  17. Lee HC: Skin cancer diagnosis using hierarchical neural networks and fuzzy

- logic. Department of Computer Science, University of Missouri, Rolla, 1994.
18. Aitken JF, Pfitzner J, Battistutta SO, et al: Reliability of computer image analysis of pigmented skin lesions of Australian adolescents. *Cancer* 1996; 78: 252-7.
  19. She Z, Liu Y, Damato A: Combination of features from skin pattern and ABCD analysis for lesion classification. *Skin Res Technol* 2007; 13: 25-33.
  20. Fassihi N, Shanbehzadeh J, Sarafzadeh A, et al: Melanoma diagnosis by the use of wavelet analysis based on morphological operators. *Proceedings of the International Multiconference of Engineers and Computer Scientists* 2011; 16-8.
  21. Amaliah B, Fatichah C, Widyanto MR: ABCD feature extraction of image dermoscopic based on morphology analysis for melanoma skin cancer diagnosis. *J Comput Inf* 2012; 3: 82-90.
  22. Safi A, Baust M, Pauly O, et al: Computer-aided diagnosis of pigmented skin dermoscopic images. 2012. Springer-Verlag Berlin Heidelberg.
  23. Premaladha J, Ravichandran KS: Detection of melanoma skin lesions using phylogeny. *Natl Acad Sci Lett* 2015; 38: 333-8.
  24. Srivastava N, Hinton G, Krizhevsky A, et al: Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res* 2014; 15: 1929-58.
  25. Esteva A, Kuprel B, Novoa RA, et al: Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; 2; 542: 115-8.
  26. Han SS, Park GH, Lim W, et al: Deep neural networks show an equivalent and often superior performance to dermatologists in onychomycosis diagnosis: Automatic construction of onychomycosis datasets by region-based convolutional deep neural network. *PLoS One* 2018; 19: e0191493.
  27. Girshick R, Donahue J, Darrell T, et al: Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2014.
  28. Haenssle HA, Fink C, Schneiderbauer R, et al: Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Ann Oncol* 2018 May 28, doi: 10.1093/annonc/mdy166. 