

智慧化企業整合

# Pneumonia

Using CNN Model to Classify Chest X-Ray

人工智慧醫療診斷模型

蘇靜琦

108034401

# 摘要

肺炎一直是國內十大死因之一，而且常位居第四或第五名，且初期症狀像感冒很容易就忽略。尤其是老人，有時症狀不太明顯，可能錯失治療先機。有時肺炎病情發展之快，令人措手不及。短短一天、甚至幾小時內，病菌就可能大量、快速繁殖，跑到血液裡變成菌血症、接著變成敗血症、敗血性休克、多重器官衰竭，十分危險。因此快速辨識是否為肺炎成為主要問題，透過建立 CNN 模型來辨識肺炎以輔助診斷。

# 目錄

一、研究背景

二、文獻回顧

三、研究方法

四、研究結果

五、結論

六、參考資料

## 一、研究背景

肺炎是現代非常常見的疾病且一直是國內十大死因之一，很多人都知道罹患肺炎後可能會發燒、咳嗽，卻不清楚它有致命風險。全球每年約有 4.5 億人（全球人口的 7%）罹患肺炎，每年約 400 萬人因此死亡。80 年後，儘管醫療和公共衛生進步，但肺炎的殺傷力依舊不容小覷。2013 年台灣有 9042 人死於肺炎，平均不到一小時就有一人因肺炎喪生，比死於肺癌、慢性阻塞性肺病的人還多。根據衛生福利部統計，台灣肺炎死亡率逐年攀升，10 年之間成長了兩倍，已在 2016 年成為台灣十大死因的第三名。

肺炎的初期症狀像感冒而很常被忽略，感冒是上呼吸道受到病毒感染而肺炎則是下呼吸道被細菌或病毒入侵，引起肺泡發炎。但肺炎初期症狀和感冒類似，如咳嗽、發燒、沒精神、呼吸喘等等，一開始不容易注意到，特別是老人，有時症狀不明顯，可能錯失治療先機。

感染肺炎的年齡層不拘，通常 5 歲以下幼兒、65 歲以上老人因免疫力較低，特別容易感染肺炎。老年人常常本來就有其他慢性病，再加上急性感染肺炎，病情往往一發不可收拾。死於肺炎的人當中有九成是 65 歲以上的長輩。肺炎病情可能快速惡化，短至一天、甚至幾小時內，病菌就可能快速大量繁殖，跑到血液裡變成菌血症、接著變成敗血症、敗血性休克、多重器官衰竭等十分危險。

台灣 X 光專科醫師人數有限，需要判讀的 X 光影像又非常多，傳統人力判讀的方式，不僅沒有效率也很容易發生疏漏。醫療導入人工智慧，這個在世界上議論已久的議題，利用人工智慧先挑出可能罹患肺炎的 X 光影像，再交由 X 光科醫師去判讀，才是兼顧效率與醫療價值的做法。

醫師若想確認病患有沒有得到肺炎，最快的方式就是胸部 X 光檢查，從影像去判讀肺部發炎的狀況。若肺炎能早期發現，並且確定病徵，及早投藥治療都能夠得到改善，因此希望建構人工智慧醫療診斷模型，提高辨認肺炎病徵的準確率。通過胸部 X 光影像數據的訓練學習，縮短細菌的鑑定時間，以爭取更多的時間搶救病患。

## 二、文獻回顧

### 1. 5W1H 分析法

5W1H 分析法是一種思考方法，對選定的項目、工序或操作，都要從對象（何事 WHAT）、時間（何時 WHEN）、人員（何人 WHO）、地點（何地 WHERE）、原因（何因 WHY）、方法（何法 HOW）等六個方面提出問題進行思考，且對問題進行綜合分析研究，從而得到更具建設性設的決策。利用此方法，主要想針對辨認肺炎的病徵來進行改善。

- (1) What: 決單獨由醫生判斷錯誤，可能準確率較低的問題。
- (2) When: 醫生確認並辨別病人肺炎的時候。
- (3) Who: 醫院看診醫生採用此模型，增加判斷準確率。
- (4) Where: 醫院的胸腔內科。
- (5) Why: 避免錯誤診斷導致病人看診時間延誤、提高醫院胸腔內科的名聲、給予病人更好的治療。
- (6) How: 建構神經網路模型，協助醫生進行病人肺炎病徵的判斷。

### 2. 卷積神經網路(Convolutional Neural Networks, CNN)

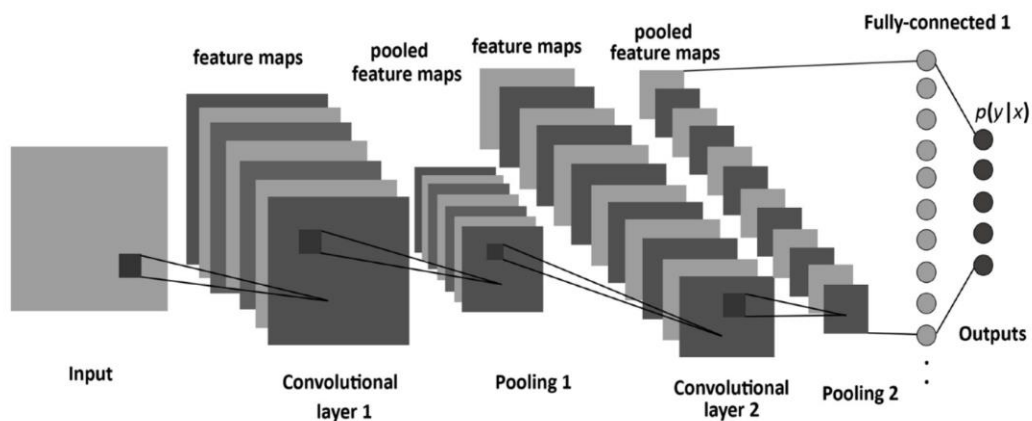
近年來，隨著人工智慧和大資料應用的發展，電腦輔助診斷技術越來越多的被應用，隨著到深度學習(Deep Learning)這一概念的興起，大批以卷積神經網路(Convolutional Neural Networks, CNN)代表的深度學習方法開始出現。

由於 CNN “自主”學習的特性，避免了機器學習中人工提取圖像特徵的局限性，大大提高了識別準確率。為進一步提高 CNN 的識別準確率，通常的做法是增大資料集、增加神經網路層數。雖然通過增加網路層數獲得了更加優異的分類效果，但是相應網路參數數量較淺層網路增加了幾十甚至幾百倍。

深度學習相較於傳統的簡單學習，其區別就在於前者可以通過一種多層的非線性結構“自主”學習，來表徵資料的特徵;而後者大多數需要人工提取特徵資訊，且因人工提取的特徵往往不能很好的表徵事物的本質，故學習效果難以提高。相較於簡單學習常用的決策樹、SVM 等方法，深度學習通過多層網路結構分層學習資料的特徵，再單獨對每一層的分類資訊進行微調，可以實現以簡單結構完成對複雜樣本的描述。CNN 是深度學習在影像處理領域的成功應用，作為一種具有深度結構的前向型神經網路，一般由卷積層、池化層、全連接層與輸出層組成。相較於傳統的人工神經網路，兩者主要的區別就在於前者通過卷積核及權值

共用大幅降低了參數數量與訓練難度。再結合大量資料的訓練，使 CNN 突破了傳統網路的限制，在影像處理領域取得了很大的進展。

CNN 是一個多層的神經網路，由多個卷積層、子採樣層交替組成，而每層由多個獨立神經元組成。CNN 框架由 2 個卷積層、2 個子採樣層交替組成。卷積層也稱為特徵提取層，每個神經元的輸入與前一層的局部感受野相連，並提取該局部的特徵，一旦該局部特徵被提取後，它與其他特徵間的位置關係也隨之確定下來。卷積層中有多個不同的二維特徵圖，一個特徵圖提取的是一種特徵，這表示提取多種不同的特徵。在提取特徵時，同一個特徵圖的權值是共用的，即使用相同的卷積核，不同的特徵圖使用不同的卷積核。卷積層將不同的局部特徵保存下來，使得提取出的特徵具有了旋轉、平移不變性。子採樣層也稱為特徵映射層，它負責將卷積層獲得的特徵進行子採樣，使提取的特徵具有縮放不變性。子採樣層只是做簡單的縮放映射，需要訓練的神經元權值相對較少，計算也比較簡單。在 CNN 的末層一般接上幾個全連接層，最終輸出節點個數就是分類目標個數，訓練的目的就是使 CNN 的輸出盡可能與原始的標籤相近。



## 三、研究方法

### 1. 研究流程

建構人工智慧醫療診斷模型的過程大致可以分為七個步驟：

- (1)資料收集、分類
- (2)資料前處理
- (3)建立神經網路架構
- (4)訓練神經網路模型
- (5)透過測試集預測神經網路成效
- (6)不斷評估模型，優化此模型
- (7)提出結果

#### (1)資料收集、分類

為了解決辨認肺炎病徵的問題，使用了網路上開源的資料集，共有 2 種不同病徵的肺炎照片，共 5856 張。其中包括正常的肺部胸腔 X 光片 1583 張和罹患肺炎的肺部胸腔 X 光片 4273 張，將其作為模型的投入。

#### (2)資料前處理

由於資料的張數分布不均，且照片容易受到拍攝角度、光線、大小等等的影響，造成模型訓練的效果不好，因此透過各種不同的前處理方式，將資料進行處理，使模型的準確率能夠提高。原始模型並沒有進行資料的前處理，因此這個改善後模型的前處理可分為下面幾個步驟：

- i. 將放置在各個不同的資料夾的圖片進行讀取
- ii. 將所有圖片裁減為相同大小，使每張的像素點數一致，才能夠放入模型中
- iii. 將區分資料集的隨機性固定(為了後續模型比較，避免因為抽取不同樣影響訓練出來的結果)
- iv. 將資料集的 80% 做為訓練集，20% 做為測試集。訓練集當中再取出 10% 做為驗證集
- v. 將圖片的像素值除以 255(將像素值介於 0-255 之間)，使像素值介於 0-1 之間
- vi. 在訓練過程，將圖片隨機剪切 20 度，隨機放大縮小 20%，圖片水平隨機翻轉，加強模型的訓練。

### (3)建立神經網路架構

將資料做完前處理後，必須決定使用的神經網路類型以及其內部架構，由於投入神經網路模型的項目為圖片，使用 CNN 模型能夠進行圖片特徵的截取，幫助保留各種不同病徵的圖片特徵，從而進行分類。此章節說明之模型架構為 kaggle 網站中他人所提出的原始模型，而這次 project 著重在於如何提高這個模型的準確率，分析過程及最終模型，將在第 4 章分析與改善提出。



## 四、研究結果

### 1. 原始模型當中，CNN模型的架構為:

- (1)照片輸入(150, 150, 3)
- (2)一層卷積核大小為 3×3，個數為 32 的卷積層
- (3)一層池化大小為 2×2 的池化層
- (4)進行比率為 0.4 的 Dropout
- (5)二層卷積核大小為 3×3，個數為 32 的卷積層
- (6)一層池化大小為 2×2 的池化層
- (7)進行比率為 0.25 的 Dropout
- (8)三層卷積核大小為 3×3，個數為 64 的卷積層
- (9)一層隱藏單元為 64 個的全連接層
- (10) 進行比率為 0.5 的 Dropout
- (11) 各層激活函數:relu
- (12) 輸出層激活函數:softmax
- (13) 損失函數: categorical\_crossentropy

訓練集共:5216 張      測試集共:624 張

根據此架構，訓練 50 個 Epoch、batch size= 16，每訓練完一次 Epoch 驗證一次結果，可以根據訓練狀況來確認是否有過擬合的情況產生。

### 2. 結果輸出

根據此模型的訓練結果，最後一次訓練集的準確率為 0.7957，驗證集準確率為 0.6875，測試集準確率為 0.7708。從原始模型當中的結果可以看出病徵分類的準確率仍偏低，因此將透過進行不同的前處理、優化器、損失函數、激活函數、參數、層數等等，使模型的分類準確率能夠提高。

### 3. 分析與改善

#### (1) 決定損失函數和激活函數的組合

原先的模型激活函數是使用 softmax，而損失函數使用 categorical\_crossentropy，categorical\_crossentropy 通常是用於多分類，而 binary\_crossentropy 通常是用於二分類，但這兩個損失函數也有其輸出層適合的激活函數，因此可以看出 binary\_crossentropy 在訓練結果明顯比 categorical\_crossentropy 好很多，同時使用 sigmoid 做為激活函數結果最好。

#### (2) 決定優化器

原先使用的優化器為 rmsprop，再加入了 adam 這種不同的優化器來進行比較。根據訓練出來的結果可以看出，仍然是 adam 做為優化器在這個模型的結果為最佳。

#### (3) 調整卷積層數

由於不知道卷積神經網路模型應該要有多少層數才能使分類效果最好，因此透過一層一層增加卷積層的方式，找到模型分類效果最好的層數。根據訓練出來的結果，得到在進行第一次池化前的卷積層數為 1 層時，訓練效果最好。

#### (4) 其他調整

除了上述那些顯著影響預測結果的項目之外，同時也做了其他調整: 利用圖片生成器(Generator)，生成各種與原圖不同的照片。減少 Dropout 在卷積網路模型中讓在擷取特徵時，丟棄一些不重要的特徵，使訓練效果變好，因此來訓練的模型。

最後可以看出，使用的這兩種方式對於這個模型，都提高模型的準確率，因此後續就將此兩種方法納入考慮之中。

所改善過後的最終模型，架構如下：

- (1) 照片輸入(150, 150, 3)
- (2) 一層卷積核大小為 3×3，個數為 32 的卷積層
- (3) 一層池化大小為 2×2 的池化層
- (4) 二層卷積核大小為 3×3，個數為 32 的卷積層
- (5) 一層池化大小為 2×2 的池化層
- (6) 三層卷積核大小為 3×3，個數為 64 的卷積層
- (7) 一層池化大小為 2×2 的池化層
- (8) 三層卷積核大小為 3×3，個數為 64 的卷積層
- (9) 一層池化大小為 2×2 的池化層
- (10) 一層隱藏單元為 128 個的全連接層
- (11) 進行比率為 0.5 的 Dropout
- (12) 各層激活函數:relu
- (13) 輸出層激活函數:sigmoid
- (14) 損失函數: binary\_crossentropy

## 五、結論

### 1. 分析與改善小結

根據改善後所提出的模型，訓練集準確率達 0.9758、測試集準確率達 0.9311，比起原先所提出的模型訓練集準確率 0.7957、測試集 0.7708 有著非常顯著的提升，可見這次針對模型的改善，的確有使模型的預測能力提升。

根據這次所建立出的模型，有著 93.11% 的準確率，可以協助醫生在判斷病人肺炎病徵時，能夠協助醫生做出正確的判斷，讓病人能夠接受正確的治療。

雖然目前的準確率仍不到很高，但往後若能優化此模型，近一步提高準確率，能夠使未來運用在判斷病徵上達到非常好的效果。

### 2. 模型及研究限制

這次透過實驗設計所訓練出的模型，即使經過了幾十次的不同的架構，仍然無法使模型測試集的準確率提高到 94% 以上，主要受到下面幾點的原因所影響：

- (1) 兩種 X 光片的照片張數較少且不平均，而這樣會導致病徵模型學習的次數較少，導致分類預測不準確。
- (2) 這次訓練模型的過程，由於訓練一次(50 個 Epoch)需要 2-3 小時，且每次修改只能變動一個指標，因此沒辦法將所有可能優化模型的狀況皆跑過一次，因此這次所得到的結論只是根據所有試過的模型當中，所推得的結果，並不是所有可能組合當中最好的結果。

### 3. 後續優化方向

由於目前的研究結果準確率仍無法達到最佳，因此若要繼續優化此模型，可以藉由下面的幾種方式，來加強 CNN 模型準確率，甚至提高訓練效率。

- (1) 增加圖片張數較少的 X 光片數量，讓兩種分類結果都有被完整訓練。
- (2) 針對醫學影像分類特點，有一種使用 CNN 實現肺炎 X 光片圖像分類的方法。利用兩種結構、深度不同的神經網路模型 AlexNet 與 Inception V3，使用知識蒸餾方法得到優化後的模型 AlexNet\_S，其在分類性能上接近甚至超過 Inception V3 模型的同時大幅降低了模型對系統資源的需求。在未來可進一步提高模型的性能，同時減少模型對系統資源的需求。

## 六、參考資料

XU Shan-shan , LIU Ying-an\* , XU Sheng. Wood defects recognition based on the convolutional neural network. Journal of SHANDONG University (ENGINEERING SCIENCE) , 2013 , 43(2): 23-28.

Ciresan DC. , Meier U , Masci J , et al . Flexible , high performance convolutional neural networks for image classification. Proceedings of the 22nd International Joint Conference on Artificial Intelligence , Volume 2.AIPress , 2011:1237-1242 .