

# 中華大學生物資訊學系系統開發專題報告

## 皮膚病變醫學影像辨識

### Skin disease medical image recognition

專題組員:呂妮珈、陳沛柔、簡心潔、王莉卉

專題編號:105007

指導老師:董其樺老師

#### 1. 摘要

人體最大器官為皮膚，但卻易被忽略。因這幾年皮膚癌罹患率提高，國人應更加重視及預防。本專題以影像辨識，運用機器學習方式，辨別皮膚上疾病或罹患皮膚癌風險程度。本專題研究成果最終提出醫學影像結合病人資訊的複合機器學習模型架構，可使準確率比僅用卷積神經網路辨識的方法更好。未來此專題能協助醫療人員快速、精確做初步診斷，減少醫療資源浪費，也能讓大眾了解該疾病對健康的重要，提高對皮膚癌的警覺性。

#### 2. 研究動機與研究問題

##### 2-1 研究動機

近年來皮膚癌罹患人數有上升趨勢，這是由於皮膚變化平常並不易注意，時常被大眾忽略。痣的成因是黑色素沉澱，但有可能轉變成為黑色素瘤，即是皮膚癌的其中一種。若長期忽略，嚴重時黑色素瘤可能會轉移至淋巴結、其他器官或其他皮膚，導致引發其他癌症，後果不堪設想。此專題是以人工智

慧影像辨識方式，辨別出皮膚上的痣與罹患皮膚癌風險的關聯。本專題研究未來希望可以讓人們在短時間內做簡易的自我檢測，不僅能解決民眾的疑慮，更可提高對皮膚癌的警覺達到及早發現及早治療的效果。

##### 2-2 研究問題

本專題研究皮膚相關疾病，主要採用7種皮膚症狀，其中包含惡性的皮膚癌以及良性皮膚疾病。

皮膚癌種類主要分為三種，如：黑色素細胞癌或稱黑色素瘤、基底細胞癌、鱗狀細胞癌，三者中最常見是基底細胞癌，其不易轉移。鱗狀細胞癌生長緩慢，忽視時可能造成嚴重的皮膚癌。在臺灣黑色素瘤雖然只占皮膚癌的7%，但其致死率極高。以上大部分發生於陽光易照到的位置，主要原因皆和紫外線曝曬有關。

另外還有四種較良性的疾病如：血管病變、良性角化病變、皮膚纖維瘤、黑色素細胞痣。血管病變為皮膚下層組織

的異常；皮膚纖維瘤大多為良性腫瘤；良性角化病變是良性的表皮增生；黑色素細胞痣即為常見的黑色痣。當皮膚上有任何異狀時，不可輕視黑痣的顏色或形狀的變化，若能及早治療可控制病情。

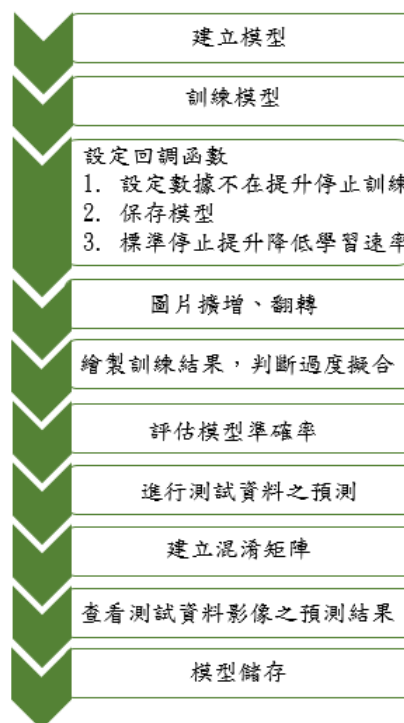
### 3. 文獻探討與回顧

過去曾有研究，使用卷積神經網路(CNN)，對肺部結核之生醫斷層掃描(CT)影像進行辨識[1]。論文最終結果是找出最佳參數組合，使其影像辨識準確率提高。結果也顯示隱藏層數目設定越多、花費的時間也越多，會使執行效率下降。從此論文中可得知，如何對CNN方法調整參數及層數是關鍵，將會影響準確率的高低。

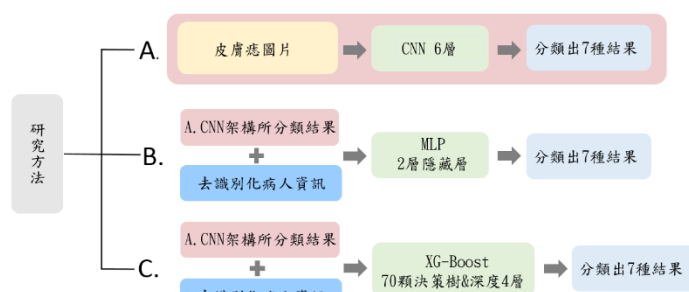
目前網路上有一個知名的網站：Kaggle，是數據建模和數據分析競賽的平台，在世界各地所有的企業或研究者可以在這個平台提供數據，同時也開放對於預測模型有興趣的參賽者們前來挑戰。在這平台上，有公布一份皮膚病變的資料集，內含10015張皮膚上痣的照片[2]。因此本專題研究的數據資料來源即是該資料集。此外，在Kaggle平台上也有公開來自全球各地研究團隊針對此資料集進行影像辨識研究時所使用的研究方法(Kernels)。他們所使用的模型架構都不同，如機器學習領域的隨機森林[3]、支持向量機[4]等，深度學習中的CNN與

其延伸出的Densenet[5]、MobileNet[6]輕量型架構。在這些Kernels中，其中一半使用卷積神經網路架構。經過我們統計歸納這些研究方法的準確率，平均落在75.45%。

### 4. 專題進行方式



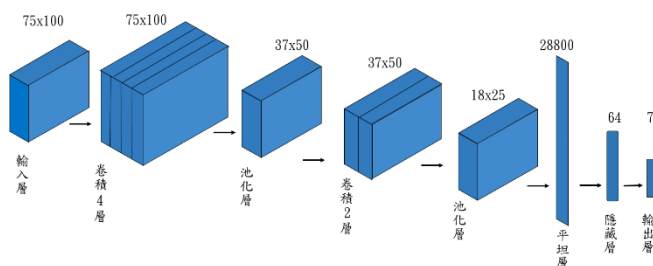
圖一：專題進行步驟



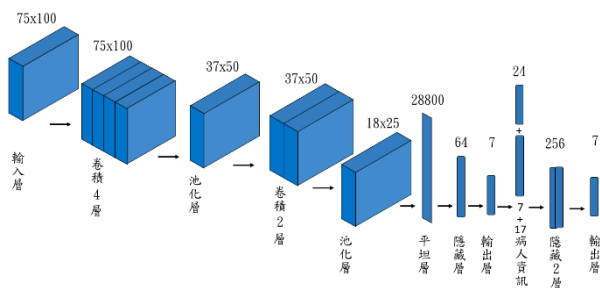
圖二：研究方法架構圖

如圖一為專題進行步驟，我們開發環境以python語言撰寫，資料來源是Kaggle網站Skin Cancer MNIST數據集，

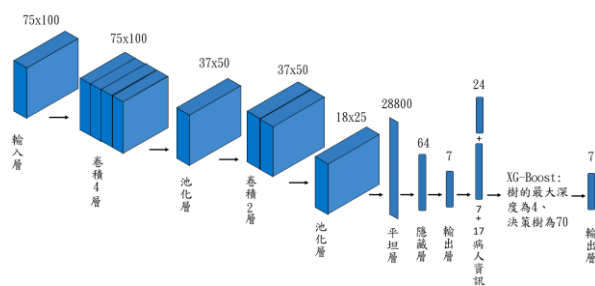
共10015筆皮膚鏡圖像。我們通過資料增強(Data augmentation)的方式，隨機翻轉照片，產生更多訓練資料，來降低過度擬合的問題，並將數據集分為學習/訓練(Training)資料80%、測試(Testing)資料20%。圖像中病例一共包含7種：Actinic keratoses (鱗狀細胞癌)、Basal cell carcinoma (基底細胞癌)、Melanoma (黑色素細胞癌/黑色素瘤)、Benign keratosis like lesions (良性角化病變)、Dermatofibroma (皮膚纖維瘤)、Melanocytic nevi (黑色素細胞痣)、Vascular lesions (血管病變)。



圖三:研究方法 A 之卷積神經網路架構示意圖



圖四:研究方法 B 中多層感知器架構示意圖



圖五: 研究方法 C 中 XG-Boost 架構示意圖

研究方法主要分為三種 (如圖二)。

A. 第一種方式將皮膚鏡像圖片直接放入卷積神經網路 (CNN)進行學習 (如圖三)，經不斷修改模型，以小的模型為主要訓練架構，圖片大小75x100放入輸入層，先建立四層卷積，以32個卷積核提取特徵值，經過第一層池化圖片降維後變37x50，再建立二層卷積37x50，卷積核為64個，放入第二層池化圖片降維成18x25，最後得到特徵值總數為28800像素值，接著建立平坦層、一層隱藏層共64個節點並加入30%機率隨機丟棄節點 Dropout (0.3)，最後輸出層為7個分類結果。整體模型架構需要訓練的參數數量總計1,927,783個。

B. 第二種方式，以多層感知器 (MLP) 將 CNN 方式所預測的七種結果再加上17個病患資訊，包含性別、年齡、發生位置，最後再度預測分類七種疾病 (如圖四)。MLP 中的輸入層為24 (7+17)個神經元，接著兩層隱藏層皆設定256個節點並且 Dropout 設定0.25，最後輸出層為7個。

C. 第三種方式，則是改用整體學習演算法 XG-Boost 取代 MLP，預測新的分類結果 (如圖五)。XG-Boost 訓練所用的參數是我們利用 GridSearchCV 的技術，進行超參數調校。最後參數測試結果是以70顆決策樹、每顆樹的深度4層之參數組合，作為最終機器學習所用的參數。

	02月	03月	04月	05月	06月	07月	08月	09月	10月	11月	12月
題目選定											
kaggle 圖片資料集分析 資料匯整與處理											
模型測試與調教											
研究方法: 加入MLP, XG-Boost 期末報告撰寫											

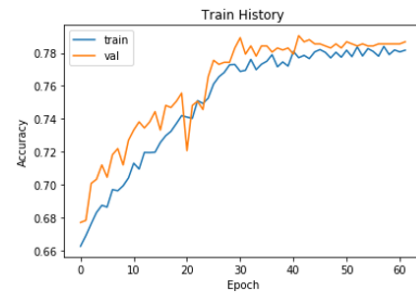
圖六:研究時程甘特圖

圖六表示本專題進行的時程。我們於3月份開始分析Kaggle的圖片疾病分類為哪幾種，找出其分佈狀況，並且尋找數據集是否有空值，將缺失值進行處理。男女的空值我們採用取多數者為填補選項，年齡則採用取平均值。在暑假期間不斷測試模型，將準確率高者留下，並觀察其折線圖是否有過度擬合的狀況。而9月份我們加入去識別化的病人資訊使用MLP方式的訓練。因其效果不如預期改用XG-Boost，可發現準確率有提升。

## 5. 主要成果

A. 發現使用CNN分析圖像，模型學習完之後最終準確率可達77.2%，並且沒有過度擬合(overfitting) (如圖七)，混淆矩陣結果如圖八。成功辨識第0類 akiec (鱗狀細胞癌)為26筆、第1類 bcc (基底細胞癌)62筆、第2類 bk1 (良性角化病變)94筆、第4類 nv (黑色素細胞痣)1243筆、第5類 mel (黑色素瘤)101筆、第6類 vasc (血管病變)21筆。從混淆矩陣中可發現 mel 最容易誤判為 nv 共有95筆。另外，第三類為 df (皮膚纖維瘤)沒

有顯示預測成功筆數，可能的原因是原始資料中各類疾病圖像數量分佈不均，其中又以 df 數據最少，且其不易學習出皮膚纖維瘤之明顯特徵。從混淆矩陣中發現其最容易誤判為 bcc 共12筆，可能和 bcc 特徵相近，如:粉紅略帶紫且不規則形狀、中間有潰瘍的情形，因而導致機器辨識錯誤。

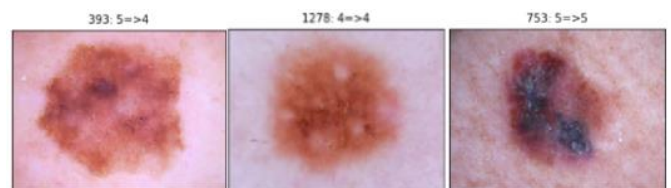


圖七:方法 A, CNN 訓練模型結果

	predict 0	1	2	4	5	6
label 0	26	11	11	7	6	0
label 1	9	62	11	21	2	6
label 2	3	3	8	94	66	26
label 3	3	3	12	5	7	2
label 4	7	9	39	1243	52	1
label 5	3	1	23	95	101	0
label 6	0	2	0	7	0	21

辨識正確筆數

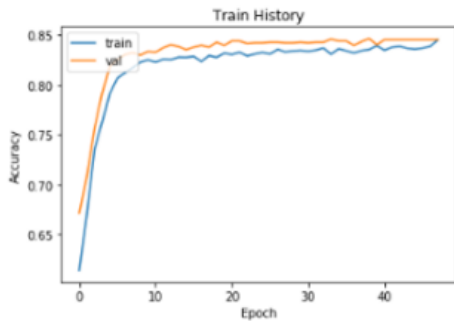
圖八: 方法 A 結果之混淆矩陣



圖九(左)CNN 預測錯誤的影像;(中)CNN 預測正確 nv 影像;(右)CNN 預測正確 mel 影像

另外 CNN 錯將黑色素瘤預測成黑色素細胞痣有95筆。以圖九為範例，三張圖片相較下錯誤者(左圖)與預測正確黑色素細胞痣(中圖)以肉眼觀察兩張圖相當雷同，顏色多為棕褐色且分布不均勻。辨識錯誤之影像(左圖)與黑色素細胞瘤(右圖)，明顯顏色差

異大，mel 顏色偏黑且邊緣有一點紫，但預測錯誤圖中無黑帶紫的顏色。



圖十：方法 B，MLP 訓練模型結果

	predict	0	1	2	3	4	5	6
	label							
	0	26	13	15	0	2	5	0
	1	14	70	12	1	6	5	3
	2	7	10	98	2	40	41	0
	3	6	5	3	7	5	3	0
	4	5	6	46	0	1239	55	0
	5	3	2	34	0	74	109	1
	6	0	1	0	3	3	2	21

mel 預測錯誤為 nv (row 4, column 5)  
辨識正確筆數 (diagonal line)

圖十一：方法 B 結果之混淆矩陣

B. 使用多層感知器 (MLP) 分析加入病患資訊後準確率達 78.6% (如圖十)。建立混淆矩陣可查看測試結果 (如圖十一)。成功辨識 akiec 為 26 筆、bcc: 70 筆、bkl: 98 筆、df: 7 筆、nv: 1239 筆、mel: 109 筆、vasc: 21 筆。從混淆矩陣中可發現最容易誤判為 nv 共 74 筆。

	predict	0	1	2	3	4	5	6
	label							
	0	25	12	15	0	2	7	0
	1	8	71	13	1	10	5	3
	2	2	12	106	0	45	32	1
	3	5	6	3	10	4	1	0
	4	3	11	41	1	1251	43	1
	5	2	3	24	2	81	110	1
	6	1	1	1	0	3	0	24

mel 預測錯誤為 nv (row 4, column 5)  
辨識正確筆數 (diagonal line)

圖十二：方法 C 結果之混淆矩陣

C. 因前述方法 B 預測準確率的提升並無如預期般明顯，所以我們改採用整體學習演算法 XG-Boost。XG-Boost 最後的準確率可達到 79.73%，較前兩者方法高。建立混淆矩陣可查看測試結果 (如

圖十二)。成功辨識 akiec 為 25 筆、bcc: 71 筆、bkl: 106 筆、df: 10 筆、nv: 1251 筆、mel: 110 筆、vasc: 24 筆。從混淆矩陣中可發現最容易誤判為 nv 共 81 筆。

由方法 A 之 CNN 預測結果的混淆矩陣來看 (圖八)，可知共有 1547 筆預測正確，有 456 筆預測錯誤。其中黑色素細胞痣預測正確筆數有 1243 筆，這可能是因為在原始資料中本身分佈約占 67%，數量相較其他種類較多。在觀察黑色素細胞痣後猜測其特徵較明顯不易混淆，如：邊緣較無輪廓線、顏色多為棕褐色或黑色但無透紫色色塊。再經過加入去識別化的病人資訊，使用 XG-Boost 可將 95 筆預測錯誤降為 81 筆。其中有 13 筆為原 CNN 預測錯誤經 XG-Boost 預測後改善。

## 6. 評估與展望

本專題研究最終我們是採用小型、淺層的 CNN 模型架構，再加入病人資訊進行第二階段的機器學習。這樣的研究架構，在 Kaggle 平台上來看，我們是第一個這樣做的。此外，學習的過程中，我們也有注意到避免過度擬合的狀況，這是因為我們用到了圖片擴增以及隨機丟棄 dropout 的技術。最後的準確率，若與在 Kaggle 平台其他團隊的 Kernel 相比，79.73% 的表現也在平均之上。

由於我們未做圖片前處理，若能先將圖片做處理，讓照片特徵更加明確，或照片中的痣大小更能一致，或許能在



機器學習中提高辨識準確率。再者，我們也發現七種疾病的照片數量差異很大，最多照片的疾病是黑色素細胞痣，共6705張；而張數最少的疾病是皮膚纖維瘤，僅有115張。若減少黑色素細胞痣的數量，讓資料分布較為平均，可降低學習資料差異過大的狀況。此外，我們未將相同位置拍攝的圖片做篩選，因為有來自相同病人同一個痣，但是多次拍攝的照片，這會在我們的模型中重複訓練。為了降低資料的重複和冗餘，應該要先篩選去除這些照片。最後，希望未來此專題可進一步延伸開發出檢測 App，可提供使用者對皮膚痣拍照即在手機上就能快速初步辨識皮膚相關疾病及罹患皮膚癌的風險，讓使用者不需要一點點小小且無害的皮膚症狀，就大驚小怪的前去大醫院要求檢驗，進而減少醫療資源的浪費。

## 7. 結語

現今已是資訊發達的時代，在我們周遭也已經有許多東西漸漸資訊化，這項研究是希望結合 AI 人工智慧的應用，可以讓皮膚科醫生在臨床診斷上有更便利、更快速、更準確的協助，也讓一般民眾可以更了解自己的健康狀況，並提高大眾對皮膚變化的警覺性，達到及早發現、治療的預防醫學目標。

## 8. 銘謝

感謝董其樺老師帶領我們完成專題

製作，帶著我們從懵懵懂懂的摸索到能明確建立模型架構，並且利用他少數空閒的時間解決我們專題上的煩惱及困惑，由於我們的電腦設備不足，老師還特別用心地幫我們準備一台電腦，能夠讓我們的專題順利進行，也謝謝組員們在討論時能夠互相提供想法及溝通，更要謝謝各位評審老師的指點，讓我們的專題有更多地方可以去學習及改進。

## 9. 參考文獻

- [1] 陳祥輝, "用卷積神經網路辨識 CT 影像內肺部結核之研究", 2018.
- [2] Kaggle. (2018). *Skin Cancer MNIST: HAM10000*. Available: <https://www.kaggle.com/kmader/skin-cancer-mnist-ham10000/kernels>
- [3] K. Mader, "Dermatology MNIST: CSV Overview(random forests)," 2018.
- [4] MrRobot1999. (2018/12). *Skin Cancer MNIST(SVM)*. Available: <https://www.kaggle.com/aryaman1999/skin-cancer-mnist>
- [5] X. Zhuang. (2019/5). *Skin lesion classification(acc>90%)(Pytorch)*. Available: <https://www.kaggle.com/xinruizhuang/skin-lesion-classification-acc-90-pytorch>
- [6] P. R. Gottimukkula. (2019/5). *Skin Lesion Analyzer + Tensorflow.js Web App(MobileNet)*. Available: <https://www.kaggle.com/pavanreddy1998/skin-lesion-analyzer-tensorflow-js-web-app>