

# 2021 年臺灣國際科學展覽會 優勝作品專輯

作品編號 190017  
參展科別 電腦科學與資訊工程  
作品名稱 以結膜影像判斷貧血之研究  
得獎獎項 大會獎 二等獎  
美國 ISEF 正選代表

就讀學校 臺北市立中山女子高級中學  
指導教師 黃升龍、黃建庭  
作者姓名 蔡伊甯

關鍵詞 人工智慧、貧血判斷、結膜影像

## 作者簡介



我是蔡伊甯，目前就讀中山女高二年級。從小學開始對資訊領域感到興趣，常常好奇著，可以如何透過開發軟體或電子裝置而使生活更便利。

第一次接觸人工智慧並建立模型，從零到有的過程非常激勵人心；尤其訓練很久（有時調個參數就要訓練一整天）而得到準確率更高之模型，賦予作品新生命的成就感很棒。

感謝黃升龍教授與助教們的細心教導與幫助，也很榮幸能夠參與這次國際科展，雖然結果不盡完美，但過程累積的學習更彌足珍貴。希望日後能夠延伸研究此專題，貢獻社會。

## 一、 中文摘要

全球貧血人口普遍，然許多人並不了解自身是否罹患貧血；長期患有貧血的病人，亦需定期抽血檢驗追蹤是否有貧血惡化達到需接受輸血的程度。研究顯示，結膜之顏色與貧血有絕對關係，結膜越白則貧血越嚴重，醫師也常使用結膜顏色推測是否有貧血情形。若能設計手機軟體自動分割結膜影像並分析其顏色，將有機會推測受試者是否罹患貧血。本研究收集 22 位無貧血者及 8 位貧血病人，並獲得其近期血紅素數值。以手機取得受試者之眼睛影像後，成功設計程式以深度學習完成結膜自動影像分割，對於分割影像以面積大小進行後期處理後，依其取得下眼瞼結膜之三原色平均，再利用 kNN 與 SVM 演算法判斷預測出該受試者是否具有貧血之症狀。本研究主要分為兩階段，其一為進行下眼瞼結膜分割模型訓練；其二為製作有無貧血之判斷模型。整合上述眼瞼分割模型(IoU=89.8%±0.02%)與貧血判斷模型(SVM 以 polynomial 核函數測出 準確值 93.3%±24.3%)後，可得貧血診斷準確率為 80%。此結果代表 AI 技術有機會透過結膜影像，判斷被拍攝者是否有貧血情形，未來若能增加研究人數，將可設計網頁版或手機 APP 加以推測血紅素值，供大眾居家篩檢。

## 二、 英文摘要 Abstract

Anemia is a common disease; however, many people do not realize whether they have anemia or not. Previous study has shown that the color of conjunctiva is associated with anemia. If we design a program in cell phone, which can identify the location of conjunctiva automatically, we are able to predict the existence and severity of anemia after taking eyes' photos. In this study, we enrolled 22 subjects without anemia and 8 patients with anemia; the recent hemoglobin levels are collected. After capturing their eye's photo using cell phone, we successfully design a program to segment the area of conjunctiva using a deep learning algorithm. The color of conjunctiva was then analyzed for the distribution of three primary colors, including red, green and blue. We confirmed that the distribution of the three primary colors of conjunctiva is different between those with and without anemia. After integration of segmentation model and color analysis model, the diagnostic accuracy for anemia achieved 80%. In the future, with increase of patient's number and optimization of models, creation of an APP or website for anemia detection is highly possible.

### 三、 前言

#### (一)、研究動機

貧血即指血液中的紅血球細胞濃度減少、血紅素降低或血液稀薄之狀態。台灣貧血人口相當普遍，常見的貧血類型有缺鐵性貧血及地中海型貧血。前者為鐵質不足所引起的貧血，台灣 19 至 44 歲之育齡期婦女中，患有缺鐵性貧血者的比例高達 60%，是我國最常見的貧血類型；後者屬遺傳性貧血，患者占我國總人數之 6~8%（大約 150 萬人）[1]。梁繼權等人發現萬里鄉與新莊市之 65 歲以上社區老人，其貧血比率分別為 34.8%及 21.8% [2]。世界各人口族群之貧血研究顯示（表 1），貧血之盛行率於不同族群介於 12.7%至 47.4%之間，而整體人口貧血盛行率為 24.8% [3]。另外，慢性腎臟病也常發生腎性貧血的併發症，發生率與嚴重度隨著腎臟病演進而惡化，嚴重的貧血會降低病患的生活品質及增加住院機率及死亡 [4]。台灣老人貧血之盛行率比美國還高，其部分原因可能是地中海型貧血盛行率在台灣較美國還高 [5]。

表 1: 世界各人口族群之貧血盛行率

| 人口族群        | 貧血盛行率       |                  | 罹病人數        |                  |
|-------------|-------------|------------------|-------------|------------------|
|             | 比例          | 95% CI           | 人數（百萬）      | 95% CI           |
| 學齡前兒童       | 47.4        | 45.7-49.1        | 293         | 283-303          |
| 學齡兒童        | 25.4        | 19.9-30.9        | 305         | 238-371          |
| 孕婦          | 41.8        | 39.9-43.8        | 56          | 54-59            |
| 非孕婦女        | 30.2        | 28.7-31.6        | 468         | 446-491          |
| 男性          | 12.7        | 8.6-16.9         | 260         | 175-345          |
| 年長者         | 23.9        | 18.3-29.4        | 163         | 126-202          |
| <b>整體人口</b> | <b>24.8</b> | <b>22.9-26.7</b> | <b>1620</b> | <b>1500-1740</b> |

貧血為紅血球減少或血紅素減少，造成血液稀薄的狀態。依世界衛生組織建議：成年男性血紅素含量低於 13.0 g/dL、女性低於 12.0 g/dL 即為貧血。一般貧血常見的症狀包括頭

暈、臉色泛白、心悸或心跳加速。貧血嚴重時，病患可能出現昏厥、呼吸急促，甚至導致大腦神經的損傷 [1]。缺乏紅血球也可能增加心臟輸送血液的負擔，進而引發包括心律不整、心臟肥大及心臟衰竭。雖然如此，許多人其實並不了解本身是否有貧血的情況。

現今如要進行有無貧血之檢測，往往要到大小醫院進行抽血，取得受試者之血液樣本，得知血紅素值 (hemoglobin)，即能判斷是否為貧血。貧血主要可分為三大主因：(1) 失血，由於外傷性失血及消化道出血等；(2) 紅血球製造數下降，如鐵質或維生素 B12 缺乏、地中海型貧血、慢性腎病、以及各種骨髓相關的腫瘤等；(3) 紅血球細胞分解增加(hemolysis)，如肝脾腫大、基因異常 (如鐮刀型貧血)、感染 (如瘧疾) 及自體免疫性疾病 [6]。梁瑞芳碩士曾發展一套自動系統 red blood cell segmentation method (RBCS method) 以協助醫生來對貧血疾病進行診斷 [7]。

貧血為慢性腎病患者最常見的併發症狀之一，其導因於慢性腎病造成紅血球生成素 (erythropoietin; EPO) 生成減少，進而使骨髓之紅血球製造能力下降。罹患嚴重腎病而需洗腎的病患，因 EPO 長期不足，因此多有或多或少的貧血情況 [8]。因此在貧血受試者之收錄上，本研究擬以洗腎病患為主要樣本來源，嘗試於洗腎中心取得貧血者之下眼瞼結膜影像進行後續分析。

三十年前已有研究顯示，結膜之顏色可用來輔助貧血的篩檢 [9]，而後續臨床診斷上，臨床醫師也常使用結膜的顏色來簡單推測是否病人有貧血的情形，並追蹤長期貧血之病患的貧血嚴重度的變化，結膜越白，貧血程度越嚴重 [10]。結膜包括眼瞼結膜和眼球結膜 [11]，由於結膜為一富含微血管的身體構造，因此若病人有貧血的狀況，則結膜的顏色可能便會因微血管內的紅血球個數較少，因此使結膜的顏色呈現較白的情況，因此以結膜的顏色來評估病人貧血之有無為合理的想法。若欲了解病人貧血的情況，通常以下眼瞼結膜的顏色為觀察重點部位 (圖 1)。

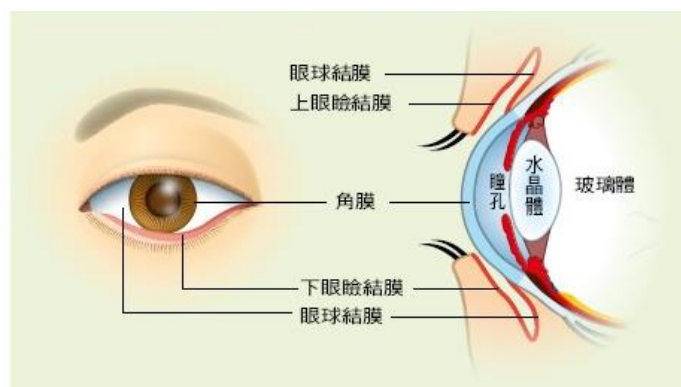


圖 1：下眼瞼結膜示意圖 [11]

近期紐西蘭團隊發現，使用手機拍攝眼睛結膜影像，再以手動方式定位結膜而後進行顏色分析，能達到不錯的貧血診斷的正確性，其 sensitivity 可達 0.57-0.74%，specificity 可達 0.71-0.83% [12]。最近美國普渡大學研究團隊開發一款應用程式，亦成功透過眼瞼拍照來判斷患者是否患有貧血 [13]。這些研究代表以 AI 應用於貧血之偵測是可行的，然而以上研究尚無法以軟體進行自動眼瞼範圍估計，而是手動以程式對眼瞼範圍圈選，或於拍攝時即對眼瞼進行人工定位，因此仍無法達到全自動分析。因此本研究將以智慧型手機拍攝無貧血者及貧血者的眼睛，獲得結膜的影像，並詢問其近期之血紅素值。而後嘗試設計程式，以深度學習演算法自動標定出下眼瞼的區塊，並進行後續之貧血及結膜顏色的相關性判讀。未來希望進而設計一個網站或手機 APP，讓使用者輸入拍攝結膜的影像，網站或 APP 之程式即能自動推測其是否有貧血的情況，如有貧血之可能性，則建議去醫院進行檢查。

## (二)、研究目的

目前貧血與否主要於醫療單位進行抽血檢驗，較為麻煩且具侵入性。本研究提出一個創新想法，透過 AI 人工智慧學習建立自動判別模型，運用普遍使用之智慧型手機進行影像拍攝，希望可達到居家簡易貧血檢測，若判斷為具有貧血可能性，即可進一步前往醫療院所檢查，提早預防以免錯失就醫良機。此方法將可能對疾病追蹤與預防醫療之使用者健康需求，提供具體的貢獻。本作品主要有三點目的：

1. 研究結膜影像自動分割之方法
2. 研究根據結膜顏色分布自動判別是否為貧血之方法
3. 整合與評估貧血居家檢測機制

## 四、 研究方法與過程

### (一) 資料收集

首先，需取得有貧血者及無貧血者的兩側下眼瞼結膜影像，與其近期血紅素數值（三個月內）資訊。由於此研究將使用到人體相關資訊，須通過倫理委員會之審查，於 2020 年 8 月底向倫理委員會申請此份研究。申請通過後，於醫院門診、病房、洗腎中心及院外收集無貧血者及貧血者共 30 人結膜樣本影像，使用一般智慧型手機拍攝取得，本研究以 iPhone 相機進行拍攝動作。

## (二) 影像處理

在對影像進行下眼瞼標記前，需先進行資料的整理，將一些解析度較低的影像予以排除，因其影像模糊將影響機器學習的判斷，但不需要排除因光線造成照片明暗的問題。而由於每個人的眼睛形狀、大小、位置與眼瞼結膜顏色皆為不同，需先以自己動手來標記每一張影像的下眼瞼位置。本研究以 ImageJ 進行標記動作，並取得以黑白區分兩類別的 ground truth data。ImageJ 為一基於 Java 的公共圖像處理軟體，能夠顯示、分析、處理、儲存圖片，本研究利用其手動標記並轉成以黑白兩色分割類別圖的功能，做出下眼瞼影像分割模型 ground truth。

## (三) 下眼瞼影像分割模型

由於手機拍攝的下眼瞼結膜在視域中會出現於不同位置，其區域大小亦會因人、手機品牌、手機位置而有所不同。因此，在軟體設計上，將以深度學習演算法，先分割出下眼瞼結膜的區塊，再做後續之貧血判讀。

深度學習是模擬人類腦神經系統建立的數學網絡模型，需要大數據來訓練，對電腦處理器必須進行大量的重複計算。因此，深度學習需仰賴好的 GPU 來加速計算工作。TWCC 使用的是刷新台灣世界排名紀錄的 aia 超級電腦《台灣杉二號》，其 GPU 系統能以最佳能源效率提供最高效能，與全球各地速度最快的超級電腦為同等級的系統。因此使用 TWCC 提供之開發環境能使運算加速，且有大量儲存空間及安全網路。

在下眼瞼結膜的影像分割上，以較常用的深度學習演算法 U-Net 做訓練，最後以 cross validation 加以驗證。

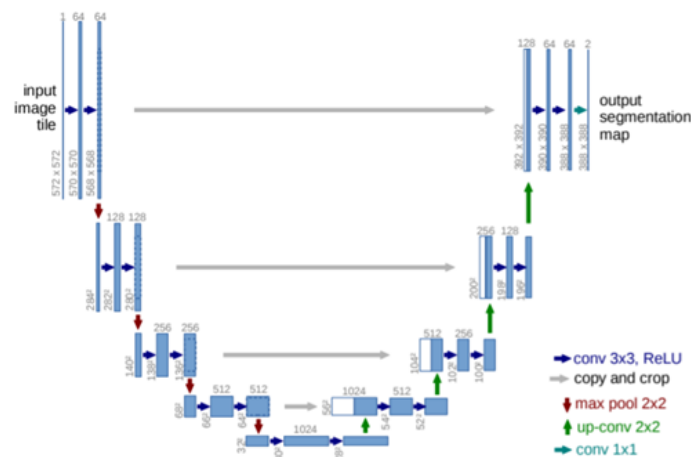


圖 2: U-Net 架構圖[14]

此模型採用的是 U-Net 演算法，為一種卷積神經網路 (convolutional neural network; CNN)。此種模型輸入為原始圖像，通過網路架構獲得訓練後的分割模型。它是由左側的收縮路徑與右側的擴展路徑組成一個類似英文字母 U 的形狀，因此稱作 U-Net。它與其他演算法較具差異性的部分為結構的後半部分，異於他者，它的網路結構只採用了卷基層，無全連接層，每個標準卷積層後都緊接著一個 Relu 激活函數層。

本研究在此模型將樣本資料分成三類，分別為 train data、validation data 以及 test data，筆數分配各為 85%、5%與 10%。先以 train data 進行訓練，再以 validation data 測試，且持續儲存 IoU 最高之模型，經過多輪測試後，將 test data 送進 IoU 值最高、結果最好的模型之中，並將執行成果之 0、1 矩陣存起 (0 代表非眼瞼部分，1 代表為眼瞼)，為 segmented result。

其中 IoU (Intersection over Union) 為一測量特定資料集準確度之標準，概念為兩區域之交集除以兩區域聯集的結果，其值越高，則代表資料預測愈準確。以圖 3 為例，A 為 CNN 模型跑出的下眼瞼分割影像，B 為以 ImageJ 畫出的 ground truth 影像。應此運算，可作為模型好壞的標準。

$$\text{IoU} = \frac{\text{Area of Overlap A, B 兩影像之交集}}{\text{Area of Union A, B 兩影像之聯集}}$$

圖 3：IoU 公式

而交叉驗證 (cross validation) 常被使用於評估一個機器學習模型的表現，避免依賴某一特定的訓練和測試資料而產生偏差，可供衡量模型表現好壞。其為統計學上，將樣本切割成多個小子集，以作為測試與訓練的一種方法。

本研究使用常用的 K-folder cross-validation 架構進行交叉驗證，其步驟說明如下：首先，將原始資料隨機平均分成 k 個集合，先將某一部分做為驗證集 (validation set)，其餘的 k-1 個集合做為訓練集及測試集 (training set, test data)；其次，訓練模型，且計算各模型的準確率，並重複以上步驟，運用不同的部分做為驗證集，直到每一個集合都曾被當作驗證集；最後的結果 (predication results) 再與真實答案 (ground truth) 進行成效比對 (performance comparison)



此種方法適合用於當數據集比較小的時候，因此非常適合進行本研究深度神經網路模型的效能測試。交叉驗證可充分利用有限的數據，找到適合的模型參數，防止過度擬合（over fitting）的情況產生。（圖 4）

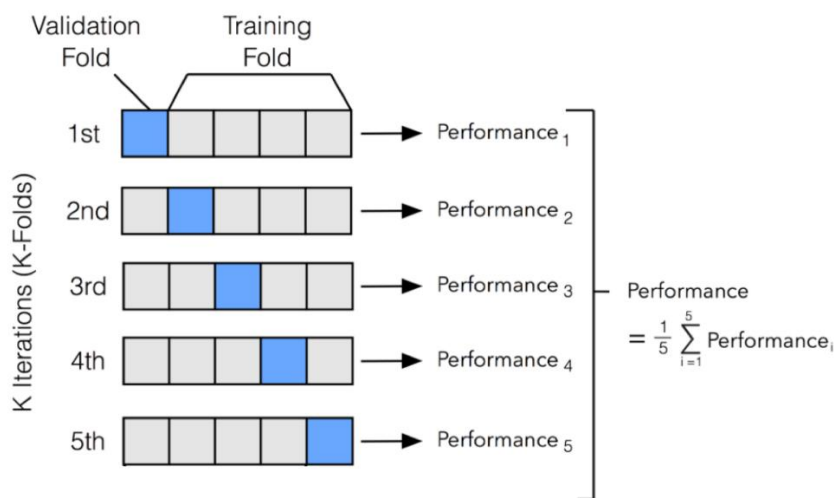


圖 4：K-fold cross-validation 架構圖[15]

#### (四) 進行結膜分割影像後期處理(post processing)

受限於計畫執行期間可收錄的受試者人數可能不足，深度學習訓練之結膜分割模型尚未達到完美之境界，因此，以分割影像進行輪廓面積大小比對，保留面積最大之範圍，判斷其為下眼瞼結膜位置。

#### (五) kNN 及 SVM 以 RGB 判斷是否為貧血模型

將資料進行分類是機器學習的重要環節，K-近鄰演算法（k nearest neighbor；kNN）與支援向量機（support vector machine；SVM）皆為分類與迴歸分析中分析資料的常見監督式學習模型演算法。而影像圖檔由許多「像素」、「畫素」所組成，大量的小光點排列整齊，構成一張「圖片」；因一個像素擁有三個數值，代表紅光、綠光、藍光的亮度，簡稱 RGB，取自紅綠藍的首字母。本研究針對結膜顏色之分析則先後以 kNN 與 SVM 運用 RGB 數值予以判斷。

kNN 為最簡單且耳熟能詳之分類演算法，根據附近 k 個鄰居評估新樣本類別，採用向量空間模型分類。此分類不需透過訓練得出模型，其參數僅 k 值一個，藉由選擇最佳 k 值產生 kNN 分類模型。以圖 5 為例，當 k=3，由於新樣本附近 3 鄰居由 class B 占多數，因此判斷其為 class b；當 k=7，因附近 7 鄰居中 class A 樣本佔多數而被分類為 class A。

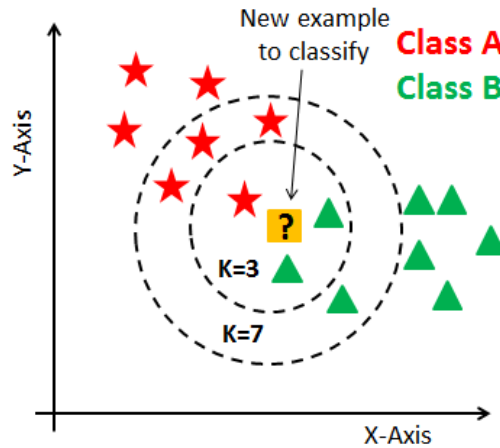


圖 5：kNN 分類示意圖 [16]

其中在  $k$  值越小，則代表由較小的臨域中進行預測，預測結果會對鄰近的數據非常敏感，容易受噪音干擾，易發生 over fitting。相對的， $k$  值越大，代表由較大臨域中進行預測，臨域中相異類別之樣本增加，容易造成 underfitting。

SVM 為基於統計學習理論的機器學習模型，概念於 1963 年由數學家 Vapnic 及 Chervonenkis 等人所提出，1992 年 Boser, Guyon, Vapnik 等人提出以供分類非線性數據的核技巧。SVM 為一種線性分類器，亦可通過將核技巧應用於最大間隔超平面建立非線性分類器（圖 6），效能高且空間成本低，於圖像、文字分類皆普遍作為分類器。

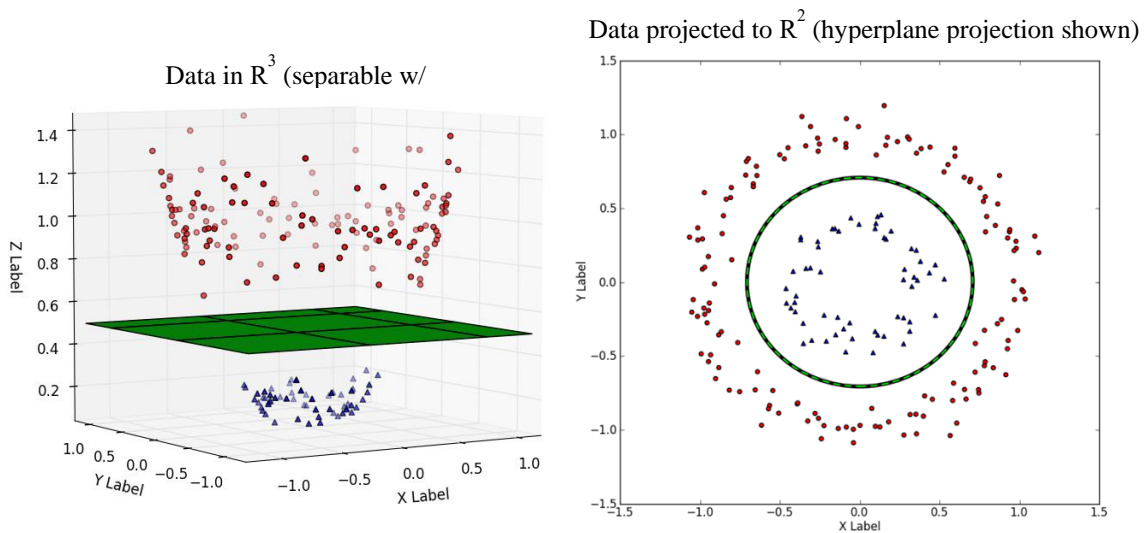


圖 6：SVM 非線性分類示意圖[17]

由於貧血者血紅素值較低，經研究顯示，血紅素值越低，下眼瞼之顏色也越白，因此以 RGB 平均數的三值當作指標，作為分辨是否為貧血的依據。本研究定義血紅素值  $< 11$  mg/dL 為貧血。

本研究先由 ground truth 配合原始影像取得各樣本結膜之 RGB 平均值，將數值結果分別以 kNN 及 SVM 演算法，調整參數進行分類訓練、驗證與測試。同於結膜分割模型，將樣本資料分 train data、validation data 以及 test data，筆數依人數分配各為 17 人、6 人、6 人。train data 供訓練，validation data 作驗證，test data 之原始影像利用由結膜分割模型跑出來之 segmented result 進行結膜 RGB 平均，透過模型判斷是否為貧血，最後經 cross validation 計算其準確率。

## 五、 研究結果

### (一) 資料收集

由於關乎人體之實驗審核較為嚴謹，經過反覆調整內容，終於十月底申請通過，得以前往台大醫院洗腎中心蒐集洗腎患者之資料。在申請通過之前，為更快取得數據以進行模型訓練的動作，先以研究者之親朋好友與自願健康受試者為樣本，而蒐集相關影像資訊與數據。

在影像樣本取得上，為了控制影像檔案的品質，避免因單純以肉眼對焦、調整拍攝大小可能造成較大誤差，影響後續模型的訓練效果，與日後進行有無貧血的判斷。本研究特別製作簡易模型（圖 5），設計水平與垂直校正參考線，以及取相參考（眼睛）定位圈；可將之套於拍攝用具（手機）上（圖 6），於拍攝受測樣本之兩側下眼瞼結膜時，利於固定所拍攝之眼睛大小與定位，也利於將來 CNN 深度學習的訓練。

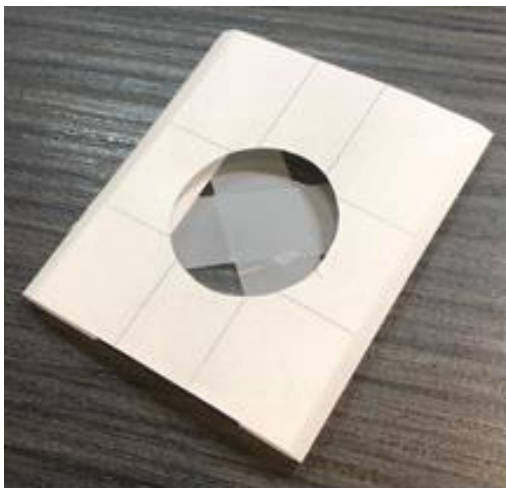


圖 7：定位簡易模型



圖 8：手機套上定位簡易模型之拍攝圖

受限於倫理委員會(IRB#202007060RINB)通過日期較晚，最終取得之樣本僅共 30 筆。樣本分布敘述（表 2）如下：依性別分，男 17 例、女 13 例；年齡介於 30~90 歲，可分為青壯年組（30~64 歲）23 例、老年組（65 歲以上）7 例；依貧血狀況分，貧血者 8 例、無貧血

者 22 例。研究者共拍攝 247 筆下眼瞼結膜影像資料，然為了提升機器深度學習效果，只選擇較為清晰之影像資料；扣除解析度較低、較模糊之資料，正式納入研究分析之影像資料共 198 筆，包含貧血者影像 78 張，與無貧血者 120 張。

表 2：樣本分布敘述

| 性別     | 年齡                  | 健康狀況      | 下眼瞼影像分析數   |
|--------|---------------------|-----------|------------|
| 男 17 例 | 青壯年組 (30~64 歲) 23 例 | 貧血者 8 例   | 貧血者 78 張   |
| 女 13 例 | 老年組 (65 歲以上) 7 例    | 無貧血者 22 例 | 無貧血者 120 張 |

## (二) 影像處理

本研究以 ImageJ 進行標記動作，並取得以黑白區分兩類別的 ground truth。將資料送入模型前，先將其資料初始化(normalize)，做以下幾項動作：進行將 4032x3024 影像 resize 成 1344x1344，並對 RGB 進行平均  $\text{mean}=[0.485, 0.456, 0.406]$ ，標準差  $\text{std}=[0.229, 0.224, 0.225]$  計算等動作。圖 9 之左側為手動標記下眼瞼結膜之示意圖，右側為取得之 ground truth；圖 10 將兩者疊合，驗證分割結果正確。

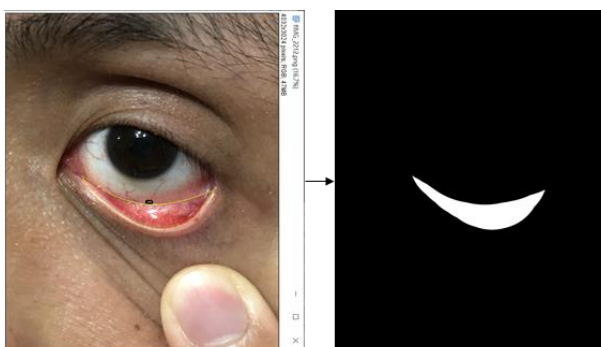


圖 9：以 ImageJ 進行標記動作

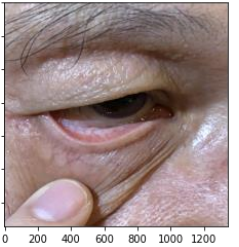
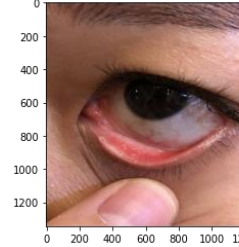
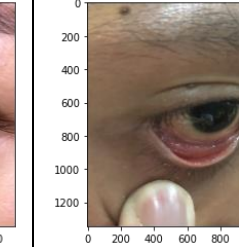
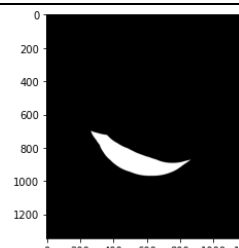
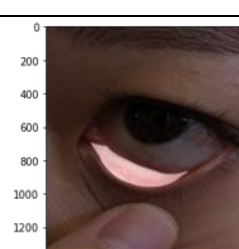
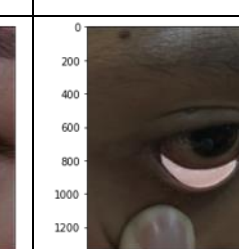
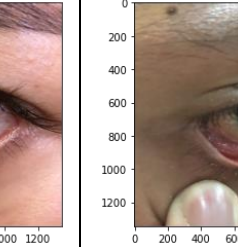
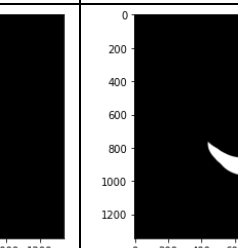
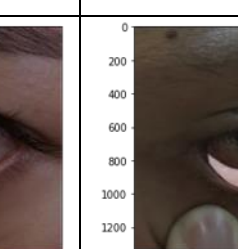
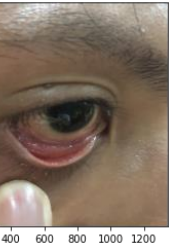
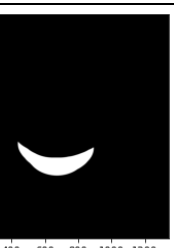



圖 10：圖 9 左右兩圖之疊合圖

### (三) 下眼瞼影像分割模型

本研究利用 CNN 深度學習，以 U-Net 演算法訓練下眼瞼影像分割模型。此模型將樣本資料依比例分成三類進行相關處理，先運用 85%比例之樣本資料列為 train data（168 張影像）進行訓練，再將 5%比例之樣本資料列為 validation data（10 張影像），反覆測試，驗證模型，持續儲存 IoU 最高之模型，最後以 10%比例之樣本資料列為 test data（20 張影像）以 IoU 值最高之模型作為結果參考。

為了訓練好下眼瞼影像分割模型，調整以下參數，分別為權重（weight）與學習率（learning rate, lr），epoch 與 lr 互相配合，隨著 lr 值減小而加大。經過 75 次反覆訓練嘗試不同結果，逐漸訓練出 IoU 較高之模型。

| 樣本                                                                                              | 樣本 1                                                                                 | 樣本 2                                                                                  | 樣本 3                                                                                  |
|-------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| Input (原始照片)                                                                                    |    |    |   |
| ground truth                                                                                    |  |   |  |
| Input + ground truth (疊合圖)                                                                      |  |  |  |
| 模型 A -<br>Input + segmented result<br>weight=[0.04,0.96]<br>lr=0.000001, epoch=100<br>IoU=0.486 |  |  |  |

|                                                                                                  |  |  |  |
|--------------------------------------------------------------------------------------------------|--|--|--|
| 模型 B -<br>Input + segmented result<br>weight=[0.06,0.94]<br>lr=0.00001, epoch=100<br>IoU=0.591   |  |  |  |
| 模型 C -<br>Input + segmented result<br>weight=[0.04,0.96]<br>lr=0.00001, epoch=100<br>IoU=0.682   |  |  |  |
| 模型 D -<br>Input + segmented result<br>weight=[0.043,0.957]<br>lr=0.00001, epoch=125<br>IoU=0.898 |  |  |  |

表 3：模型改良成果比較

從上表比較中，可明顯看出模型 D 之分割成果表現最佳，其分割區域幾乎完全對應到 ground truth；模型 A 對於特殊特徵影像，如皺紋較多、膚色較白、飽和度較高等，無法正確分辨下眼瞼；調整參數後，模型 B 明顯進步，但仍有許多誤判區域；模型 C 雖較模型 A、模型 B 為佳，但眼瞼邊緣亦較不清楚，尚無法判別飽合度較高之樣本影像。

另一方面，亦可發現部分可能影響資料判讀之因素。如樣本 1、2 在模型 A、B、C 三種模型之分割成果皆較樣本 3 明顯不理想，推測其可能原因為背景顏色、特徵、飽和度等影像干擾因素較多而影響判別結果。然經過參數調整，包含權重 (weight) 與 learning rate (lr)，模型 D 已可呈現較良好之影像分割成果。

經過多次嘗試，發現當 weight=[0.043, 0.957]，lr=0.00001 時，得訓練出最佳之下眼瞼分割模型，因此停止調整參數，以一樣參數的模型進行 K-folder cross-validation。由於本研究將樣本以 8.5:0.5:1 之比例分成 train data、validation data 與 test data，將 k 設為 20，進行 20 個 fold 的訓練。

進行 cross-validation 前，以  $\text{weight}=[0.043, 0.957]$ ， $\text{lr}=0.00001$ ，epoch 調為 250 之模型（圖 11）作為參考，發現 loss 值大約於 epoch=40 開始以緩慢趨勢下降，epoch=250 之 loss 最小值 0.323 在 epoch=123 時已得出，epoch123 後 loss 值皆 0.323 及 0.346 間徘徊，可知其已接近飽和，因此選擇以 epoch=125 為標準訓練模型，也較省訓練時間與效率。

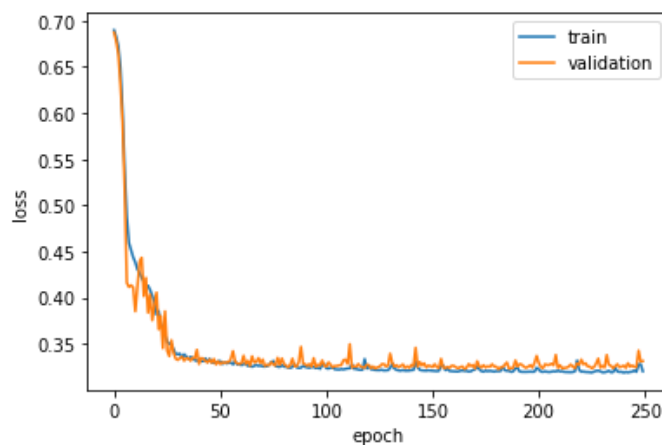


圖 11：隨 epoch 增加之 loss 趨勢圖

本研究最終採用的模型為  $\text{weight}=[0.043, 0.957]$ ， $\text{lr}=0.00001$ ，epoch=125。從圖 12 及圖 13 可知，loss 值隨著時間下降，大約於 epoch=40 處開始以極緩慢的趨勢下降，最後趨近飽和狀態。經過 125epoch 反覆訓練，其 loss 最終數值為 0.334（由 0.692 開始下降），其中更新了 29 次 model，其 IOU 由 0.169（epoch=0）上升至  $0.898 \pm 0.02$ （epoch=116）。

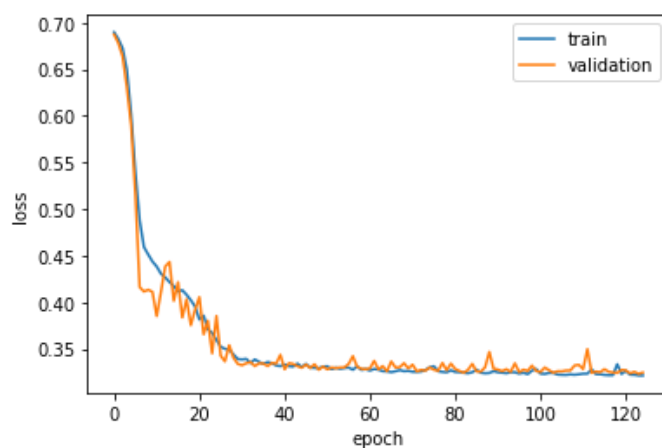


圖 12：隨 epoch 增加之 loss 趨勢圖

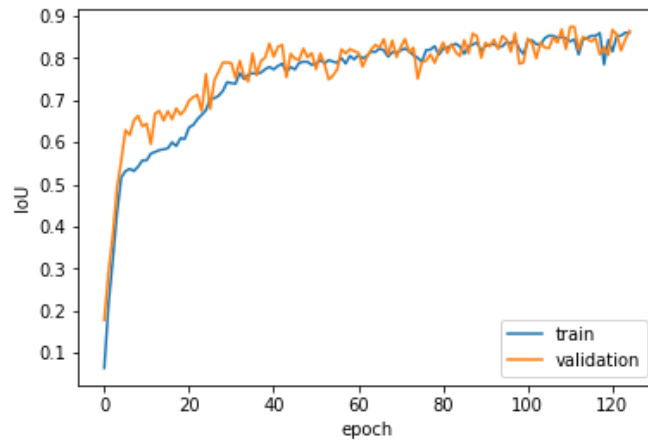


圖 13：隨 epoch 增加之 IoU 趨勢圖

■ Validation set  
■ Training set+Test set

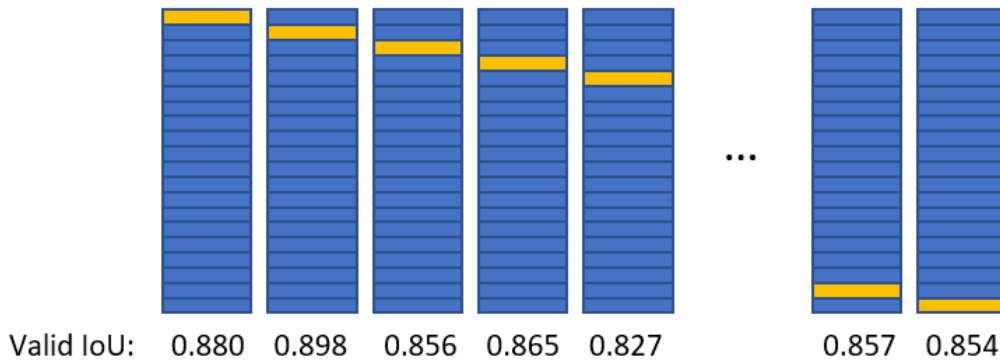


圖 14：下眼瞼分割模型 cross validation 示意圖

以 k-fold cross validation 驗證下眼瞼分割模型，最終平均 IoU 值為 0.864，標準差為 0.02，由此可知此模型具有不錯之分割水準。

#### (四) 進行結膜分割影像後期處理(post processing)

雖下眼瞼分割模型已進步許多，IoU 值也從提升至 0.898，仍有些影像無法完美分割。為不影響後續判斷有無貧血之作業，本研究將測試出的分割影像進行 post processing：偵測輪廓，並保留最大面積。由於分割模型訓練效果良好，得確保其最大面積皆為大致結膜範圍，因此以此動作確保下眼瞼分割模型之不足會強烈影響後續的分類作業，例如：眉毛、眼白等易大幅影響 RGB 值的區域 (圖 12 及圖 13)。



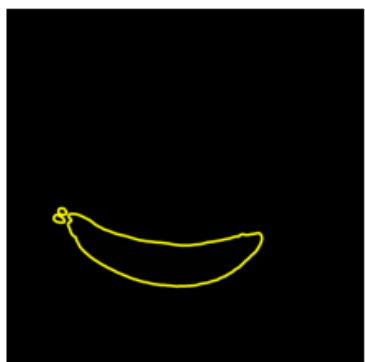


圖 15：跑結膜分割模型 segment data



圖 16：經過後期處理之成果

### (五) kNN 及 SVM 以 RGB 判斷是否為貧血模型

本研究於進行有無貧血判斷模型訓練時，亦將數據分為 train data、validation data 與 test data 三類。由於同一人結膜顏色在不同角度、明暗變化下，經過 normalize 平均後結果仍相似，因此需將同一人之樣本影像放於同一類別當中，以免造成模型測試時遇到學習過的特徵，造成模型準確率高於實際準確率的現象。由於收集到的資料量少，非貧血者 21 人（199 張影像），貧血者僅 8 人（78 張影像），因此以下列方式分配資料集：train data 為 17 人（13 位非貧血者；4 位貧血者）、validation data 及 test data 各為 6 人（4 位非貧血者；2 位貧血者），並進行 5 fold 之 cross validation。

本研究運用所收集之 198 張樣本影像，依 ground truth 取得各影像之 RGB 平均值；將樣本分為 training data（157 張）、validation data 和 testing data（各 20 張）。將 157 張影像樣本之 RGB 平均值分別以 kNN 演算法、SVM 演算法進行分類訓練，20 張 validation data 驗證更新模型，再將最後 20 張利用結膜分割模型跑出來的 segmented result 進行下眼瞼 RGB 平均，並以模型測試是否為貧血；最後計算其判斷準確率。

本研究樣本影像之 RGB 平均值，為平均下眼瞼區域每個像素之 R、G、B 值，最終得到三平均值，代表該樣本 RGB 特徵。

由於不同環境可能造成影像明暗變化，造成 RGB 值判斷不準確，在平均結膜 RGB 值前，先將每像素之 RGB 三值進行 normalize：

$$\alpha_{ij}^2 \cdot R_{ij}^2 + \alpha_{ij}^2 \cdot G_{ij}^2 + \alpha_{ij}^2 \cdot B_{ij}^2 = 1$$

$$\alpha_{ij} = 1 / \sqrt{R_{ij}^2 + G_{ij}^2 + B_{ij}^2}$$

將所有下眼瞼影像之 RGB 平均值匯入，可產出影像 RGB 平均值三維分布圖（圖 17），綠點、紅點分別代表有、無貧血樣本。由此圖可見，綠點分布較為集中，紅點分布較為分散；換言之，貧血樣本分布較為集中，無貧血樣本之分布較為分散。

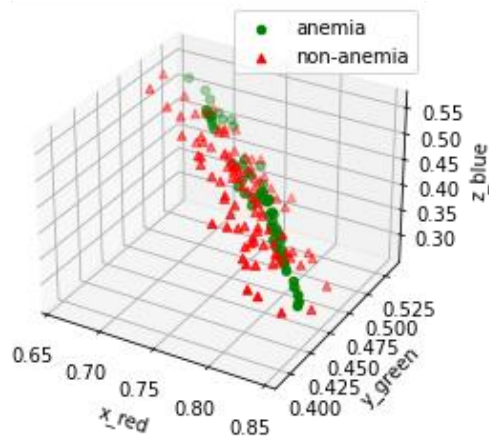


圖 17：影像平均 RGB 值三維分布圖

更進一步，將三原色各取兩色進行二維分布分析，橘點代表無貧血樣本，藍點代表貧血樣本。圖 18-20 中，以 RB 紅藍兩色之二維分布圖（圖 19）較為清楚，具有較高之相關性，較能看出貧血與無貧血之分類；另外兩種之兩色二維分布圖則較為凌亂，GB 綠藍兩色（圖 15）之相關性似乎最低。

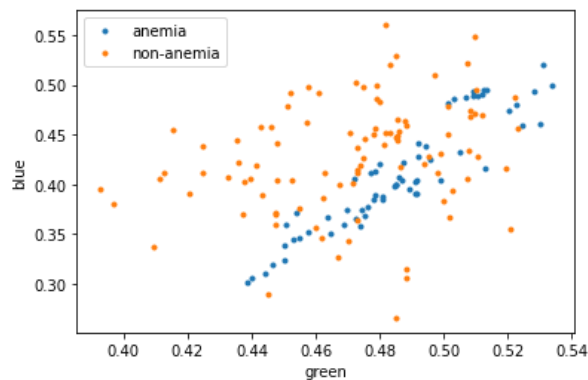


圖 18：影像平均 RGB 值之 GB 兩色二維分布圖

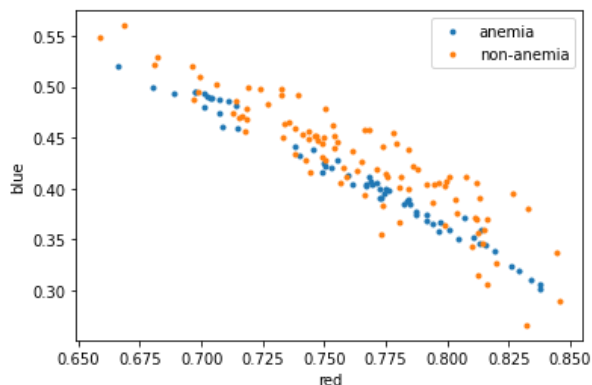


圖 19：影像平均 RGB 值之 RB 兩色二維分布圖

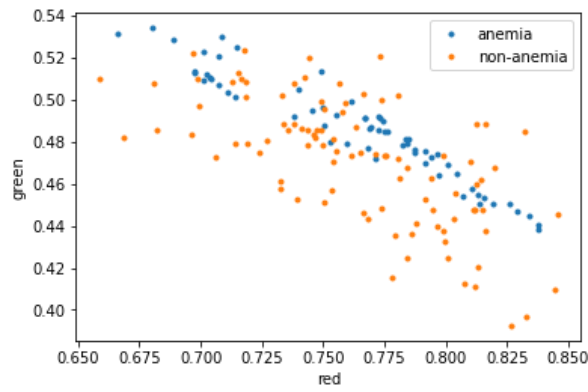


圖 20：影像平均 RGB 值之 RG 兩色二維分布圖

本研究首先以 kNN 演算法進行訓練，分別於 5 fold 中選擇錯誤率最低之 k 值，當作該有無貧血判斷模型之 k 值。其中排除 k 為偶數者，避免類別數被 k 整除，造成臨域中各類別具有相同樣本數即無法判斷。

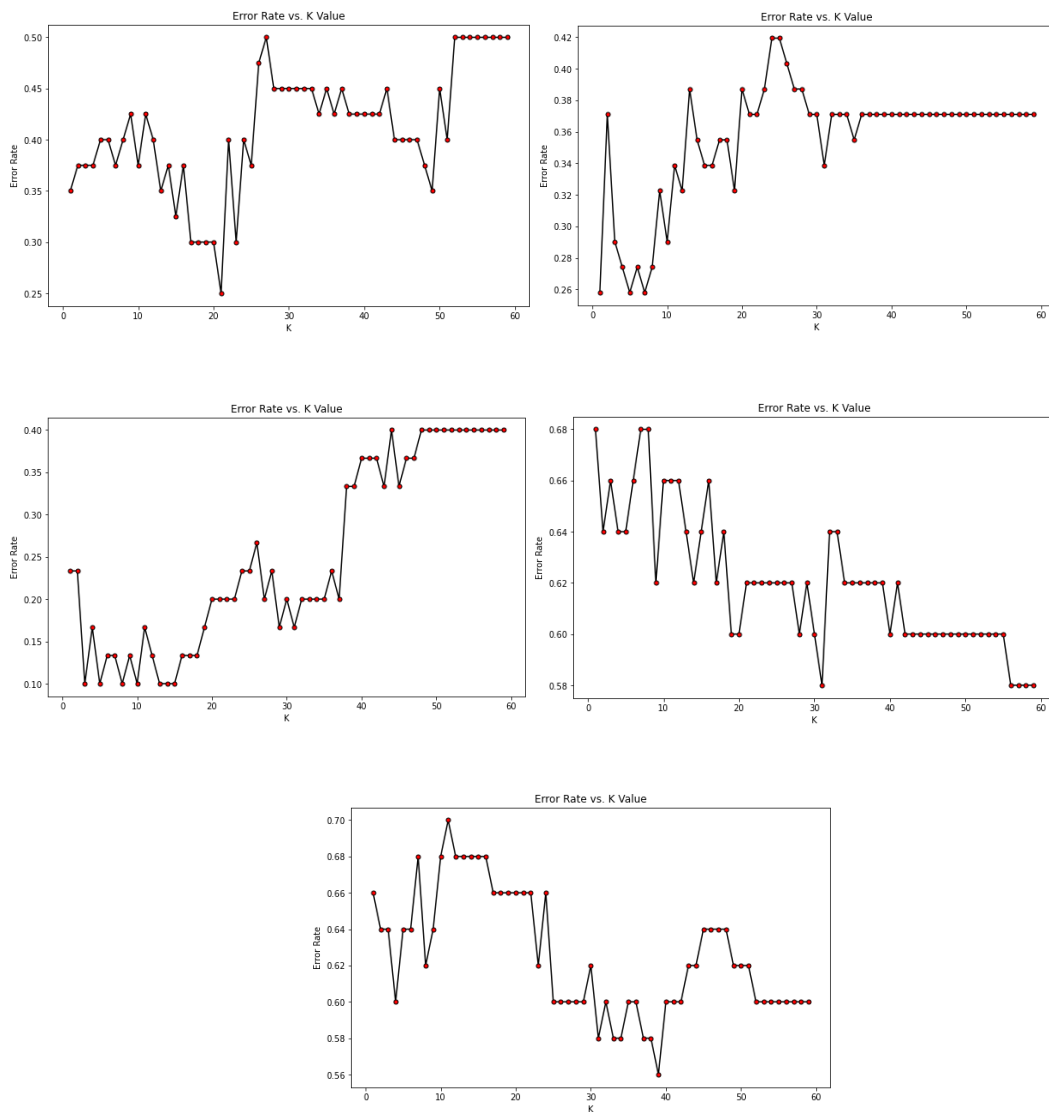


圖 21：5 fold error rate 隨 k 值增加而變動之趨勢圖

由此五張圖可觀察到，k 值在每個 fold 中皆不同，依序為 21、1、3、31、39，全距高達 38；其準確率分別為 36.0%、77.5%、69.3%、93.9%、34.3%， $\mu=0.622\%$ ， $\sigma=24.2$ ，結果差距大，推測其原因為樣本數不平衡、非稀有事件，且樣本數太少。

本研究進一步運用 SVM 進行有無貧血判斷。以 SVM 分類樣本時，如數據線性不可分，則透過將輸入空間內線性不可分的數據，映射到一個高維的特徵空間內，使數據在特徵空間內成為線性可分。因內積的計算維度將非常大，故可引入核函數。核函數為一種快捷計算方式，可幫助更快地執行某些計算，否則將涉及高維空間中的計算，常用的核函數包含 Linear、Polynomial、RBF 與 Sigmoid。本研究分別將此四種核函數進行判斷與比較：

表 4：linear、polynomial、rbf、sigmoid 核函數模型比較

|                   | linear          | polynomial                | rbf                         | sigmoid                      |
|-------------------|-----------------|---------------------------|-----------------------------|------------------------------|
| Fold 0            | C=128 ; 26%     | C=8, gamma=2 ; 28%        | C=500, gamma=0.2 ; 26%      | C=0.004, gamma=0.004 ; 26%   |
| Fold 1            | C=0.004 ; 50%   | C=128, gamma=128 ; 75%    | C=7500, gamma=50 ; 77.5%    | C=0.004, gamma=0.004 ; 50%   |
| Fold 2            | C=0.004 ; 62.9% | C=10000, gamma=10 ; 72.6% | C=10000, gamma=100 ; 72.6%  | C=0.004, gamma=0.004 ; 62.9% |
| Fold 3            | C=243 ; 93.3%   | C=8, gamma=128 ; 93.3%    | C=243, gamma=9 ; 93.3%      | C=0.004, gamma=0.004 ; 93.3% |
| Fold 4            | C=512 ; 40.6%   | C=4096, gamma=128 ; 59.3% | C=8182, gamma=0.125 ; 46.9% | C=2, gamma=128 ; 40.6%       |
| 準確率 $\mu, \sigma$ | 54.6%, 25.6     | 65.6%, 24.3               | 63.3%, 26.7                 | 54.6%, 25.6                  |

經過多次參數之調整，最終發現以 polynomial 作為 SVM 核函數準確率相對較高（65.6%），亦高於 kNN 分類模型平均準確率，故本研究以 polynomial 作為 SVM 核函數，藉 RGB 數值判斷影像樣本是否為貧血（圖 22）。

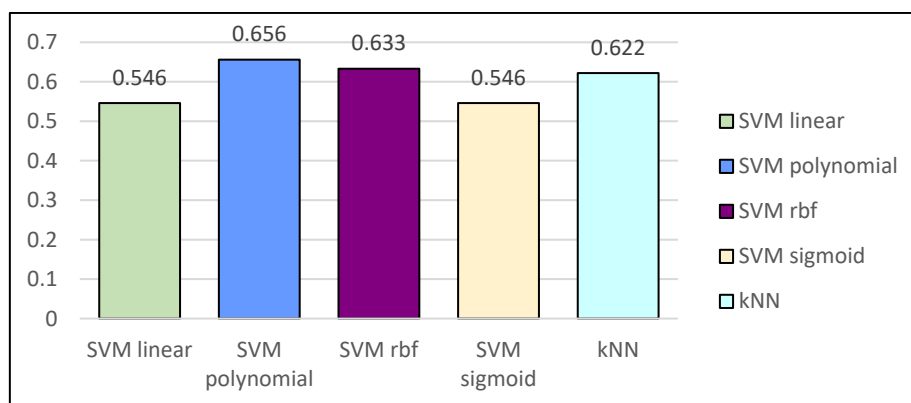


圖 22：四種核函數之準確率比較圖

## 5. 模型整合

最後一個步驟，將 IoU 為 0.898 之下眼瞼結膜分割模型與 SVM 以 polynomial 核函數測出準確值 93.3% 之有無貧血判斷模型，予以整合。最終以 test set (貧血者 8 張，非貧血者 12 張；共 20 張) 計算準確率，為 80%，其中 6 張貧血者與 10 張非貧血者影像判斷正確並得其混淆矩陣 (圖 23)。

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| non-anemia   | 0.83      | 0.83   | 0.83     | 12      |
| anemia       | 0.75      | 0.75   | 0.75     | 8       |
| accuracy     |           |        | 0.80     | 20      |
| macro avg    | 0.79      | 0.79   | 0.79     | 20      |
| weighted avg | 0.80      | 0.80   | 0.80     | 20      |

圖 23：混淆矩陣分析

使用者可上傳一含下眼瞼之影像，此模型將依序透過下眼瞼分割模型、後期處理、下眼瞼區域 RGB 值平均、有無貧血分類模型，最終得到一數值，0 為非貧血，1 為貧血。

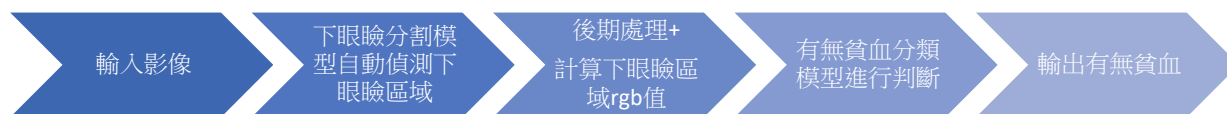


圖 24：整合模型流程圖

根據此模型，使用者可自行使用智慧型手機拍攝下眼瞼結膜影像，上傳照片至此模型，透過圖 24 之流程，回報其是否具有貧血之可能性。此模型運算結果之數值若為 1，則有貧血之可能，建議進一步至醫院檢查；數值若為 0，則無貧血之可能，為健康狀態。此模型可提供一般使用者進行居家檢測，方法簡單又迅速。

## 六、 研究討論

本研究主要分為兩階段，其一為進行下眼瞼結膜分割模型訓練；其二為製作有無貧血之判斷模型，未來在待取得更多樣本數，增加判斷準確率後，將有機會寫成網頁或應用程式，作為使用者介面，使用者得以傳入手機拍攝的結膜影像，即得到有無貧血之判斷。

此研究發現，AI 人工技術的確可由手機所拍攝的一般眼睛影像自動分割出結膜的位置，並進行顏色的深度分析。亦發現貧血者及無貧血者結膜顏色之三原色分布呈現差異，得以 kNN 與 SVM 進行有無貧血的分類。代表本研究的方法得以 AI 輔助預測是否有貧血，有助於潛在

患者居家簡易自我檢測，提早預防與就醫。此結果也在在顯示 AI 技術將可應用於未來人類的生活。

醫師過去常使用眼瞼顏色目測評估病人是否有貧血情形 [9,10]；近年亦有研究嘗試以手機拍攝眼瞼，成功以軟體分析貧血的可能性 [12,13]，然而過去這些研究仍需以手動方式定位眼瞼，降低了未來實際應用的潛力。本研究成功以 U-Net 深度學習準確完成眼瞼影像分割模型，對後續 APP 的應用將有很大的助益。

由於關乎人體之實驗審核較為嚴謹，倫理審查時程遠超過預期，導致樣本收集時間緊迫，相關實驗資料收集不易，最終只取得三十份樣本資料，尤其貧血者下眼瞼結膜影像明顯少於無貧血者下眼瞼結膜影像，因此樣本不足可能為造成貧血判斷準確率不高的原因之一。

另外在取得影像後手動進行結膜位置圈選之步驟，當時只以眼球、眼角、下眼皮作為結膜位置判斷依據，並沒有排除眼睫毛、眼皮陰影等因素，也可能使 RGB 平均值造成落差，若後續有機會再修正，應在進行結膜位置圈選時排除以上狀況。

訓練結膜分割模型時發現，結膜拍攝影像之其他因素，可能影響影像分割之效果，如皺紋之線條干擾，膚色過白等，可能造成分割成果不精準；然透過參數調整，仍可找出準確率較高之模型參數。若能夠更進一步取得更多樣本資料，即可更加完美的分割結膜影像，提高準確率。

本研究雖然可驗證下眼瞼結膜影像之 RGB 與有無貧血可跑出相關性，但輸入影像自動判別貧血與否的準確率還可再加強。一方面可能貧血之樣本數太少，造成機器學習之偏頗誤差。亦或者是，輕度貧血對結膜顏色較無影響，介於臨界值附近之貧血或貧血者較無法以結膜顏色區分，若將來有機會收集更多居血紅素貧血指標附近之貧血樣本影像，即可驗證其與結膜顏色是否直接相關。若呈現直接相關，則可嘗試增加其影像量再進行訓練，效果較佳，可能可提高準確率。

雖然已完成下眼瞼結膜影像之自動分割、有無貧血判斷之模型，第二模型以結膜顏色判讀有無貧血準確率仍有待改善。本研究目前單以較為熟知之 kNN 及 SVM 進行分類判斷，然現今有許多演算法即參數供以參考，若有機會再進行此計畫，將會嘗試些更合適之演算法進行分類判斷。

梁瑞芳女士碩士曾建立一自動系統 red blood cell segmentation method (RBCS method) ，經

血液抹片影像使用影像技術計算一張血球抹片內紅血球的大小、數量及其雙凹盤比例，作為無貧血之判斷。其過程較本研究複雜且耗時。美國普渡大學研究團隊近期開發一款應用程式，透過眼瞼拍照來判斷患者時否患有貧血，其運用光譜分析(Spectral Analysis)，透過血紅素吸收可見光的特殊方式來檢測血紅素，由產生的光譜訊號得出血紅素含量，不過其團隊尚在努力加以改進。台灣尚未開發以結膜影像判斷是否患有貧血之應用程式或模型，期望得到更多樣本資料，待此研究更加完備、完整過後，能夠公開給大眾使用能夠更加精準以結膜 RGB 判斷使否患有貧血症狀。

## 七、 結論與應用

本研究發現，AI 人工技術的確可由手機所拍攝的一般眼睛影像自動分割出結膜的位置，並進行顏色的深度分析。亦發現貧血者及無貧血者結膜顏色之三原色分布呈現差異，得以 kNN 與 SVM 進行有無貧血的分類。代表本研究的方法得以 AI 輔助預測是否有貧血，有助於潛在患者居家簡易自我檢測，提早預防與就醫。此結果也在在顯示 AI 技術將可應用於未來人類的生活。

然而貧血者和無貧血者的結膜顏色雖稍有不同，但以本研究的架構尚不足以準確區分，其原因可能為 (1) 研究受試者人數不足，尤其是患有貧血之病人人數過少；(2) 貧血的嚴重度不足，或許以結膜顏色預測貧血，需要貧血相對嚴重時才能較為準確；(3) 尚有更加適合的演算法與參數以供分類。後續若能對上述問題釐清及修正，仍有機會獲得進一步的突破。

本研究成功建立兩個模型，第一個模型成功運用 CNN 深度學習訓練，完成了下眼瞼結膜影像分割模型，其根據使用者利用智慧型手機所拍攝之下眼瞼結膜影像，自動正確分割結膜區域；第二個模型成功根據 kNN 及 SVM 以下眼瞼 RGB 判斷是否具貧血之症狀。待訓練完畢，本研究更進一步將兩個模型整合成一個程式，寫入網頁中，方便未來使用者於一個步驟即可完成資料自動分割與自動判讀。只要傳入一張結膜影像，即可得知是否有貧血之疑慮。

因研究時間過短之限制，來不及自行架設網頁或應用程式，使大眾能夠直接使用。未來若有機會，應能運用此模型，開發成兼顧使用者介面優化之完整應用功能，造福一般民眾。利用簡單的三步驟，一拍照、二上傳、三判讀，簡易快速完成居家檢測，協助有需要進一步檢測之民眾提早就醫檢查，或是針對貧血風險較高之洗腎患者，提早檢測、儘早就醫，減少健康惡化之風險。

## 八、 參考文獻

1. 施珮筑. 認識貧血. 秀傳醫療體系: <https://www.scmh.org.tw/dept/lab/bing/Preview/page-105.html>
2. 梁繼權, 許詩典, 陳晶瑩, 陳慶餘. 都市與鄉村社區老人健康狀態。台灣醫學. 2002;6(4):523-533.
3. World Health Organization. Global anemia prevalence and number of individuals affected. [https://www.who.int/vmnis/anaemia/prevalence/summary/anaemia\\_data\\_status\\_t2/en/](https://www.who.int/vmnis/anaemia/prevalence/summary/anaemia_data_status_t2/en/)
4. 林之原、張弘育、林俊良, 2017年, 慢性腎臟病的貧血治療;腎臟與透析:29卷3期.
5. 謝美芬、顏兆熊. 老人貧血之臨床表現與病因. 台灣家庭醫學醫學會. [https://www.tafm.org.tw/ehc-tafm/s/w/ebook/people\\_other/journalContent/642](https://www.tafm.org.tw/ehc-tafm/s/w/ebook/people_other/journalContent/642)
6. 胡宛屏、李俊杰、柯逸青、蘇世斌, 2003年, 海洋性貧血與缺鐵性貧血之減速鑑別診斷;中華職業醫學雜誌:10卷1期 P25~30.
7. 梁瑞芳, 2012年, 貧血症血液抹片紅血球及其雙凹盤切割暨特徵計算;中興大學資訊管理系學系所學位論文。
8. National Kidney Foundation. Anemia and Chronic Kidney Disease. [https://www.kidney.org/atoz/content/what\\_anemia\\_ckd](https://www.kidney.org/atoz/content/what_anemia_ckd)
9. Strobach RS, Anderson SK, Doll DC, Ringengerg QS. The value of the physical examination in the diagnosis of anemia correlation of the physical findings and the hemoglobin concentration. Arch Intern Med 1988;148:831-832.
10. Sheth TN, Choudhry NK, Bowes M, Detsky AS. The Relation of Conjunctival Pallor to the Presence of Anemia. Journal of General Internal Medicine 2002;12:2.
11. 黃靜宜。你「眼紅」了嗎？結膜、角膜發炎怎麼辦？好健康第45期，全民健康基金會。<https://www.twhealth.org.tw/journalView.php?cat=45&sid=755&page=1>
12. Collings S, Thompson O, Hirst E, Goossens L, George A, Weinkove R. Non-invasive detection of anaemia using digital photographs of the conjunctiva. PLoS One. 2016;11:e0153286.
13. Wiles K. Smartphone app to help assess anemia by taking a picture of a person's eyelid. Purdue University Website. 2020 May: <https://www.purdue.edu/newsroom/releases/2020/Q2/smartphone-app-to-help-assess-anemia-by-taking-a-picture-of-a-persons-eyelid.html>
14. U-Net: Convolutional Networks for Biomedical Image Segmentation <https://arxiv.org/pdf/1505.04597v1.pdf><https://www.itread01.com/content/1543926966.html>
15. K-Fold Cross Validation: [http://ethen8181.github.io/machine-learning/model\\_selection/model\\_selection.html#K-Fold-Cross-Validation](http://ethen8181.github.io/machine-learning/model_selection/model_selection.html#K-Fold-Cross-Validation)
16. kNN: <https://ithelp.ithome.com.tw/m/articles/10224036>
17. SVM: <https://www.jeremyjordan.me/support-vector-machines/>



## 【評語】 190017

本作品巧妙應用機器學習技術於貧血的偵測，並且透過實際的系統開發與實驗設計，驗證所發展的系統具有良好的效能成果。本作品在研究方法與步驟上，皆展現良好的科學過程與精神，若能更進一步強化實驗的設計與討論，同時擴大實驗的規模，探討研究成果的可能延伸應用，將能更加凸顯本作品的貢獻，也能讓本作品的整體呈現更為完整。