

用於農場安全監控的人臉偵測及辨識系統

吳乙澤

國立臺灣大學生物機電工程學系

摘要

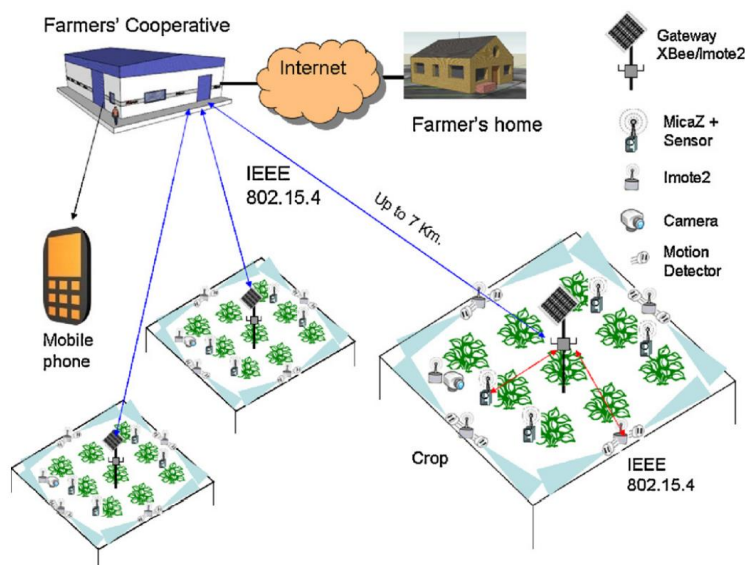
為保證農場安全，防範宵小入侵，偷取農民辛勤耕耘的蔬果。本研究打算以一個嵌入式系統完成農場安全監控。本系統會無間斷拍攝農場空間，偵測並記錄行人人臉，將之儲存在雲端計算平台 AWS (Amazon Web Service) 資料庫中。爾後，利用機器學習做人臉辨識。完成後，還會透過 AWS 架設網站，將系統所監控到的圖片顯示在該網站上，以利於農民查看。系統完成後，有放至實驗室環境做測試。測試結果顯示，人臉偵測準確度可達 0.99，人臉辨識驗證集準確度為 0.87，表該系統確實具備安全監控的能力。

關鍵字：人臉偵測、影像處理、雲端計算、安全監控系統

1. 簡介

1.1 研究背景

近幾年來，得益於圖形處理器及平行計算的發展，機器學習技術得以普及。連帶地，使得機器學習結合電腦視覺的方法，能順利地應用在精準農業的研究上，並有了眾多的成果[1]。Sabanci 等人在 2017 提出了一個電腦視覺系統。該系統可藉由人工神經網路，從影像中辨識出 *triticum aestivum* 和 *triticum durum* 兩種不同品種的小麥[2]。Pires 則於 2016 年，使用支持向量機(Support Vector Machine)，自動地偵測出大豆是否健康[3]。除了以影像的方式監控作物，也有學者以影片的方式進行精準農業的研究。見下圖一，Antonio Javier Garcia Sanchez 等人以多個嵌入式系統上的相機拍攝農地，並配合 XBee、IEEE 802.15.4 等通訊協定將畫面串流到手機及網路上，以偵測及識別入侵者，如動物或人類[4]。Tallam Charan Nikhil 等人則以四軸飛行機結合第一視角的相機，對廣大的農地進行攝影，並透過無線網路(Wi-Fi)將影片傳送到手機上。爾後，他們從該影片的圖像中，偵測是否有害蟲存在，提早得知蟲害以減少作物損失[5]。



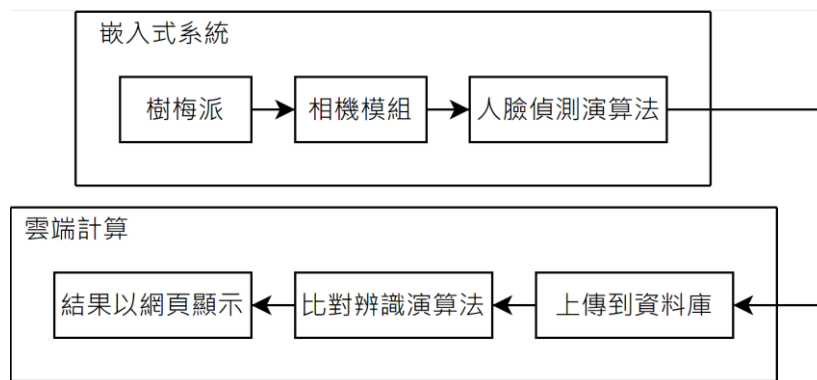
圖一、Antonio Javier Garcia Sanchez 等人的系統架構圖[4]

1.2 研究動機

由研究背景可知，以電腦視覺和機器學習技術應用在精準農業的研究中，大部分是對作物品質或害蟲做監控及辨識，少有農場安全性監控的研究論文。上網搜尋台灣有關作物的新聞，可以發現時不時有作物被偷的情形發生，尤其是高單價作物，例如蜜蘋果、玉荷包、芭樂等等，損失每每高達數千元，甚至萬元以上。雖有警方、農民定時巡邏田園，但是農園地廣人稀，難免有漏網之魚。因此，為減少農民的損失，本研究想以低成本的嵌入式系統，拍攝農場的照片或影片後，配合電腦視覺、影像處理與機器學習技術，偵測並儲存曾經出現在鏡頭內的人臉。最後，經過比對分析，找出非該農場的工作人員，達成農場安全的自動化監控。

2. 材料與方法

本章提出該研究使用到的材料及方法。見下圖二，該圖是本研究的系統架構圖。本系統由兩個主要部分組成，嵌入式系統及雲端計算。嵌入式系統的部分，打算以樹梅派(Raspberry Pi)為核心進行開發。除樹梅派外，會再加上相機模組以便於拍攝影像。拍攝到影像後，直接於樹梅派上做人脸偵測後存取人脸圖片。接著，把人脸資料送至 AWS 雲端資料庫中，並以 AWS 運算服務做人臉比對及人脸辨識。完成後，使用 AWS 架設簡陋的網頁應用程式(Web Application)，將拍攝畫面及偵測到的人臉顯示在網頁上。



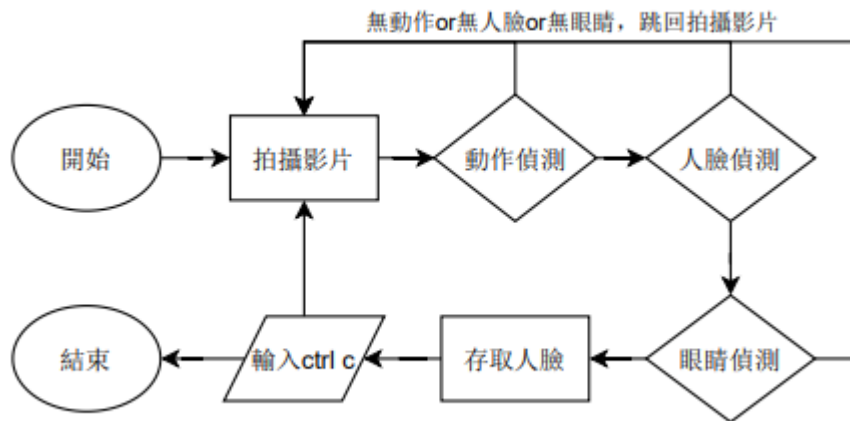
圖二、系統架構圖

2.1 硬體材料

本研究用到樹梅派、相機模組、防水盒子、壓克力基板等硬體材料。以下對樹梅派和相機模組進行介紹。樹梅派的型號為 Raspberry Pi 4 Model B，記憶體為 2GB，處理器頻率最高可到 1.5GHz，並且可透過 2.4GHz 與 5GHz 兩種頻段連接上無線網路。相機模組則利用 Raspberry Pi Camera Module v2.1，該相機可以使用 1080p、720p、480p 三種不同的解析度拍攝影片。

2.2 影像處理

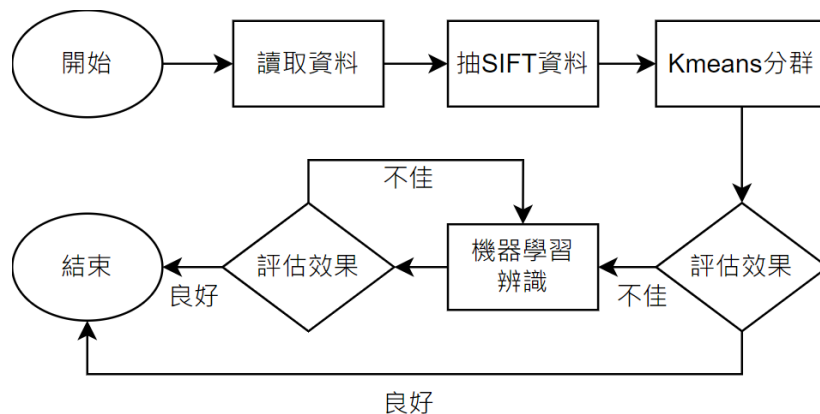
見下圖三，該圖是本系統用於存取人臉的演算法流程圖。若使用者不中斷程式的話，程式會無間斷地拍攝影片，並以三個過濾方法，動作偵測、人臉偵測、眼睛偵測，來確保偵測到的畫面為人臉。首先，動作偵測的部分，將前後兩個影格相減後，做二值化。接著，再經過形態學的細化，並得到輪廓後，便能從輪廓中得出移動物體的面積大小。最後，只要設定合理的面積閾值，便能有效地偵測出物體是否有在移動了。人臉偵測的部分，用到三種方法，分別為 Haar Cascades[6]、Dlib、DNN，這三種方法皆以函數庫實現，包括 OpenCV 和 dlib。DNN 使用 OpenCV 預先訓練好的 res10_300x300_ssd_iter_140000.caffemodel 模型做人臉偵測。使用者在執行程式時，可以自由從 Haar、Dlib、DNN 中，選擇一個演算法來偵測人臉。另外，眼睛偵測的部分，也是以 Haar Cascades 的方式來進行偵測。若偵測到的眼睛數量為 1 或 2，便將人臉存取至本機的資料夾中，等待雲端計算的程式將人臉上傳至資料庫。



圖三、人臉偵測演算法

2.3 雲端計算

使用目前市面上最為大宗的雲端計算平台 AWS，來上傳圖片、比對人臉、快速部屬網頁。上傳圖片的部分，以 AWS 雲端資料庫 S3(Amazon Simple Storage Service)來儲存圖片。S3 是種雲端資料庫，可以存放任何數量及形式的資料，適合 IOT 裝置與大數據分析[7]。系統每十分鐘上傳場域圖片，每三十分鐘上傳人臉圖片。上傳完畢後，將這些人臉作為機器學習的資料集使用。本研究以 AWS 的機器學習運算服務 SageMaker，來標記人臉資料集。接著，同樣以 SageMaker 對該資料集做人臉比對與辨識。見下圖四，該圖為人臉比對與人臉辨識的演算法流程圖。讀取存在 S3 的圖片後，以 SIFT (Scale invariant feature transform) 抓取特徵點，並把特徵點的座標丟入 Kmeans 模型做資料分群。接著，比對資料集的標籤與分群結果，以評估該法是否有效。若不佳，再以機器學習函數庫 Tensorflow，架構 CNN (Convolutional Neural Network) 網路模型以做人臉辨識。若 CNN 訓練結果不佳，重新修改模型的參數或架構，直到該模型具有一定的辨識準確度後，才結束該部分。



圖四、人臉比對及人臉辨識演算法

完成人臉辨識模型後，是快速部屬網頁的部分。本研究打算以 AWS 服務 Elastic Beanstalk 部屬網站，並以 Flask 為後端，HTML 為前端。在嵌入式系統上傳人臉圖片後，會在雲端自動以人臉辨識模型偵測可疑人物。若發現可疑人物，則將該人物之人臉顯示在網頁上並發送警告，以方便農民檢視並做進一步的處理。

3. 結果與討論

本研究共有三項研究成果，包括硬體、影像處理、網頁等。硬體即為嵌入式系統的組裝成品，影像處理則為 Haar Cascades、Dlib、DNN 三種方法的比較，以及人臉比對和人臉辨識的結果。最後，網頁的部分，是將影像處理成果呈現的媒介。以下將分別展現這三項結果並討論之。

3.1 硬體結果

見下圖五，該圖是硬體的最終成果。黃色部分是防水盒外觀，該盒子前方有防水鏡片，可使盒內的相機拍攝場域。盒子上方有掛勾，可讓農民自由掛在農場出入口或蔬果場附近，以方便監控過往行人。



圖五、硬體成果

3.2 影像處理結果

我將本嵌入式系統放置於實驗室做測試，以比較三種方法，Haar Cascades、Dlib、DNN 的人臉存取準確度。見下表一，該表顯示出不同方法所抓到的圖片數量，以及經過人工視察，將無人臉及無法辨識為誰的人臉剔除後，確實為人臉的數量。人臉數量計算完畢後，將之相除後以無條件捨去法取小數至第二位，便能得到該方法的人臉偵測準確度。由下表可見，Dlib 和 DNN 的準確度極高，近乎百分之百都能抓到人脸，Haar Cascades 的準確度則只有 8 成左右。人臉偵測速度的部份，Haar 快於 DNN，Dlib 是這三種方法中偵測速度最慢的。另外，DNN 還能抓到帶有口罩的人臉，可謂這三法中最好的一個，兼具準確度與偵測速度。

表一、Haar Cascades、Dlib、DNN 比較表

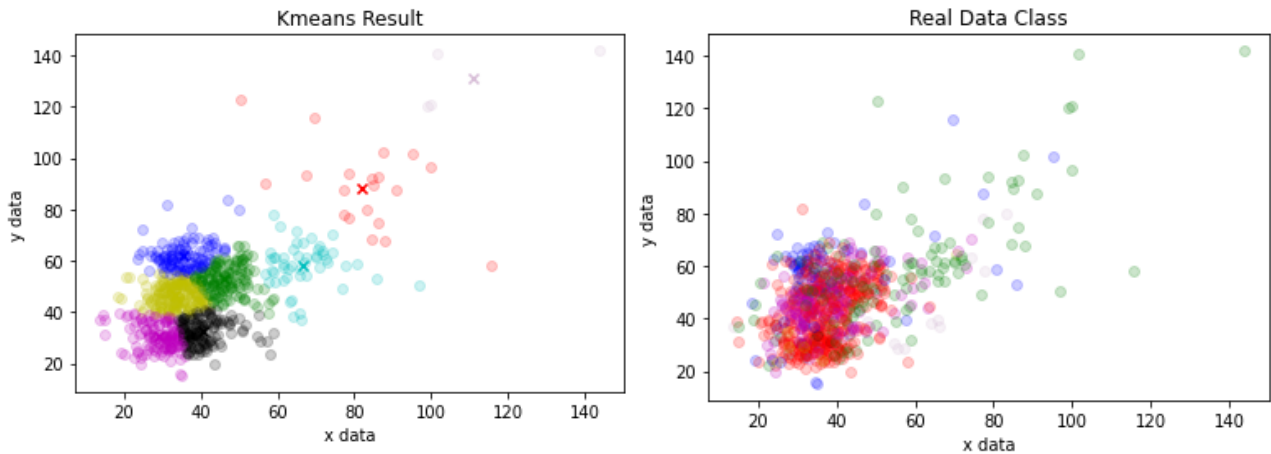
偵測方法	Haar Cascades	Dlib	DNN
總存取人臉數量	46	231	561
確實為人臉數量	37	230	555
準確度	0.80	0.99	0.98

除了三種人臉偵測演算法的比較外，還有人臉比對及人臉辨識演算法的結果討論。見下表二，這是本研究蒐集到的人臉資料集統計表，紀錄該資料集的類別標籤和該類別的圖片總數。本資料集共有八個類別，Outliers 與 Person 1 至 Person 7。Person 1 至 Person 7 為本實驗室的人員，由於個人隱私故以 Person 的代號顯示之，實際上是以姓名做標記的。Outliers 則代表該圖片裡沒有人臉、無法辨識，或者不為 Person 1 至 Person 7。Outliers 這個類別說明該人臉不為本實驗室的人員，因此是潛在的危險人物。由數量可以看出，有的類別圖片數過少，不到 10 張，有的類別卻過多，總數到了 400 多張。因此，可能導致模型在訓練完畢後，偏向某一類別。

表二、實驗室環境下人臉資料集類別及數量

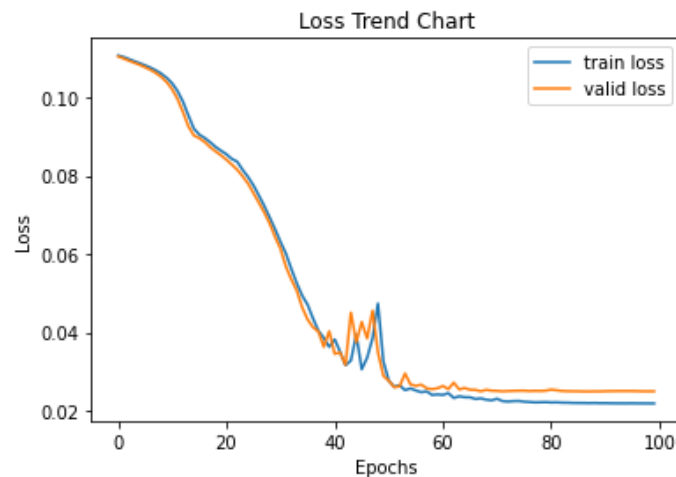
類別	Outliers	Person 1	Person 2	Person 3	Person 4	Person 5	Person 6	Person 7
總數量	98	123	408	5	124	7	7	15

見下圖六，左圖是提取完 SIFT 特徵後，以 Kmeans 做分群的結果。右圖則是真實數據的分布圖。本資料集有共八個類別，故將分群數設為 8 做 Kmeans 分群。圖中不同的顏色表示不同的類別。比較二圖後，可以看出該法效果不佳，真實數據與 Kmeans 的顏色分布完全對不起來，表示 SIFT 的關鍵點座標無法拿來做人臉辨識。因此，按照流程圖，進入機器學習的部分。



圖六、K-Means 結果圖

機器學習我以 Tensorflow 函數庫建構出 CNN 模型，並拿該模型訓練我在實驗室取得的人臉資料集。資料集總張數為 787，以 6 比 4 的比例，切分出訓練集與驗證集。見下圖七，這是模型的誤差趨勢圖。可以從中看出，訓練集與驗證集的誤差變化相似，沒有過擬合的問題。並且，到 60 個 epoch 後，誤差逐漸穩定下來，沒有 40 個 epoch 時的震盪現象。



圖七、模型誤差趨勢圖

見下表三，可以得知該模型的訓練結果。驗證集的準確度達 0.87，接近 9 成的準確度，有達本研究的標準，故不再修改模型架構及參數，結束人臉辨識的部分。

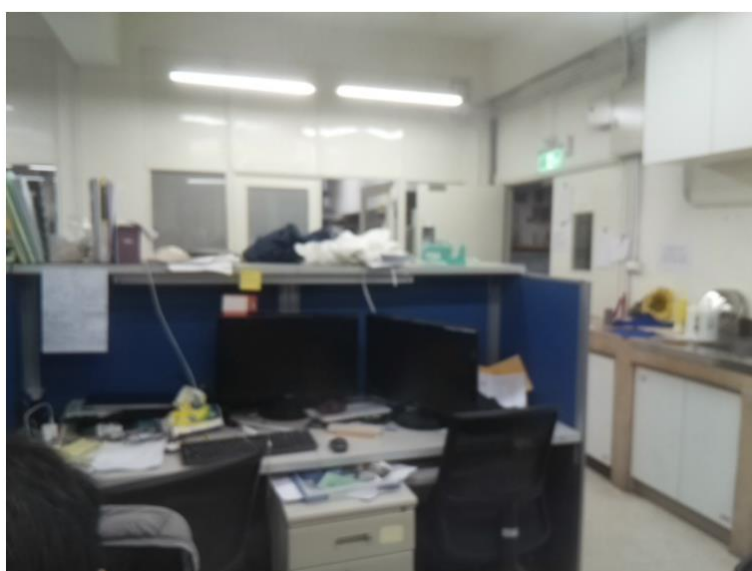
表三、模型訓練結果

種類	訓練集誤差	驗證集誤差	訓練集準確度	驗證集準確度
值	0.022	0.025	0.90	0.87

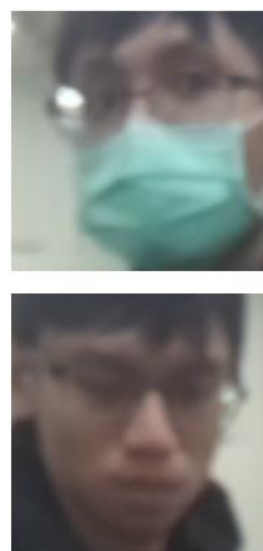
3.3 網頁結果

見下圖八，這是網頁成果的截圖畫面。左圖是本嵌入式系統所拍攝到的場域串流照片，每十分鐘更新一次。右圖則是人臉照片，這部分由於時間因素，沒有做到每三十分鐘更新一次的功能。因此，右方兩張人臉僅僅是把先前系統上傳到資料庫的人臉，靜態地置於網頁中。該網頁有實際上線，可點擊下圖註解中的網址進入查看。

Field Images



Face Images



圖八、網頁截圖成果，連結為 <http://flask-env.eba-m8y2qxrm.us-east-1.elasticbeanstalk.com/>

4. 結論

4.1 總結

本研究提出一個用於農場安全監控的人臉偵測及辨識系統。經過於實驗室的測試後，確實有實際成效。不論是 人臉偵測準確度，還是人臉辨識準確度，皆有達本研究當初設下的目標。尤其是本安全監控系統不必利用影片串流，就能成功偵測到人臉並儲存在雲端資料庫上，可謂同時節省了影片串流的網路流量與運算資源。

4.2 未來方向

關於資料集類別標籤的部分，未來考慮在標記類別時，直接把非人臉及無法辨識的圖片刪除。如此一來，不僅能減少資料量，更能使 Outliers 的定義更加明確，因為 Outliers 就只代表不為 Person 1 至 Person 7 的人臉。在以網頁顯示可疑人物時，資料的處理上也會簡單許多。另外，由於時間及個人能力限制，本研究在前端的部分，沒有做到人臉更新的功能，人臉辨識的模型也沒有部屬到雲端上。未來有機會，會再花時間完善本專題。

参考文献

- [1] Diego Inácio Patrícioa, Rafael Rieder, “Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review,” *Computers and Electronics in Agriculture*, Volume 153, pp. 69-81 (2018).
- [2] Kadir Sabanci, Ahmet Kayabasi, Abdurrahim Toktas, “Computer vision-based method for classification of wheat grains using artificial neural network,” *Journal of the Science of Food and Agriculture*, Volume 97, Issue 8 (2016).
- [3] Rillian Diello Lucas Pires, Diogo Nunes Gonçalves, Jonatan Patrick Margarido Oruê, Wesley Eiji Sanches Kanashiro, Jose F.Rodrigues Jr., Bruno Brandoli Machado, Wesley Nunes Gonçalves, “Local descriptors for soybean disease recognition,” *Computers and Electronics in Agriculture*, Volume 125, pp. 48-55 (2016).
- [4] Antonio JavierGarcia Sanchez, Felipe Garcia Sanchez, Joan Garcia Haro, “Wireless sensor network deployment for integrating video-surveillance and data-monitoring in precision agriculture over distributed crops,” *Computers and Electronics in Agriculture*, Volume 75, Issue 2, pp. 2888-303 (2011).
- [5] Tallam Charan Nikhil, Tallam Karthik, Tummuri Rajasekhar Reddy, B.K. Priya, “Agrifucus for Precision Farming,” *International Conference on Communication and Signal Processing, ICCSP* (2020).
- [6] Paul Viola, Michael Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features,” *Computer Vision and Pattern Recognition* (2001).
- [7] Amazon, “Amazon S3,” <https://aws.amazon.com/tw/s3/> (2020)