

國立中興大學資訊管理學系

碩士學位論文

應用特徵圖可視化和卷積神經網路

於花粉粒分類之研究

The study on applying the visualization of feature
map and convolutional neural network to pollen

grains classification

National Chung Hsing University

指導教授：蔡垂雄 Chwei-Shyong Tsai

陳信華 Hsin-Hua Chen

研究生：呂俊穎 Chun-Ying Lu

中華民國一〇七年七月

國立中興大學資訊管理學系

碩士學位論文

題目：應用特徵圖可視化和卷積神經網路於花粉粒分類之研究

姓名：呂俊穎

學號：7105029045

經 口 試 通 過 特 此 證 明

論文指導教授

蔡金龍 陳信亨

論文考試委員

林冠成

黃德權

吳忠味

蔡金龍

陳信亨

中華民國 107 年 7 月 12 日

誌謝

兩年的研究生生活轉瞬即逝，從一開始尋找研究方向，決定為機器學習後，讀了許多的教材，掌握了基本概念，接著搭配著實作程式，漸漸越來越上手，雖然說還有更多的知識值得我去學習，不過至少已打下了基礎，在這部分我首先要感謝我的指導老師蔡垂雄老師以及陳信華醫師，蔡老師提供了我許多的學習資源以及一些意見，讓我在學習時朝向好的方向邁進，陳醫師跟我分享了醫療資料的屬性值該如何解釋，解決了我在處理醫療資料上的一些問題。學習的路上我也遇到了不少幫助我的老師，讓我非常感謝，像是吳憲珠老師總是能在我每次報告完提出問題讓我思考，陳彥匡老師在我處理統計資料時也教了我處理方法，林冠成老師每次撥空與我討論論文都能點出我的問題並給我建議，使我獲益良多。除了老師們以外，也感謝學長姊們分享經驗以及傳承，還有我不可或缺的夥伴們，包含彥均、柏勳、永松、嘉嘉，有你們的互相扶持以及切磋才能讓我一路走來不寂寞並且不斷進步。最後，感謝我的家人，爸爸、媽媽以及妹妹對於我的支持與鼓勵，讓我有持續不斷努力的動力。

謹誌於中興

民國一百零七年七月

摘要

花粉粒在許多領域的應用非常廣泛，如蜂花粉含有高營養價值，有助於人類的美麗和健康，對其辨識並分類能使產品品質得到保證，花粉可能引發花粉過敏，藉由分類花粉可以協助醫生的診斷，花粉化石可用來研究古環境與古氣候的重建。花粉粒的分類傳統採人工方式，然而以人工分類不僅要由專家來執行，顯示花粉粒的分類需要高熟練度，而且當資料量龐大時，處理會非常耗時，因此產生了自動分類的概念。一般對於花粉粒的分類會先用影像處理方法提取出特徵，再以機器學習方法分類，準確率大概落在 64%至 95%，但人工提取特徵耗時，更重要的是，假如沒有提取到合適的特徵，分類的準確率會不符預期。近年來卷積神經網路(Convolution neural network, CNN)也被應用在花粉粒的研究上，CNN 的卷積層具有特徵提取的功能，能夠省去繁雜的前處理步驟，因此 Daood 在 2016 年以 CNN 對花粉粒分類，達到約 90%的準確率。本研究使用的資料集包含 805 張、23 類的花粉粒圖像，實驗結果分為兩階段，第一階段將 805 張影像分為灰階與彩色以三種不同架構的 CNN 分類，其中彩色圖的準確率皆比灰階圖高，而彩色圖以 Simple CNN 所得準確率 81.55%為最高，相較於使用同資料集的準確率 64%有顯著的提升。第二階段以資料擴充將圖像經旋轉、縮放、裁切等方式擴充到 8714 張，再以 Simple CNN 建立預測模型，實驗結果達到了 95.95%的準確率，相較於同樣以深度學習，包含 30 類資料集的所得 89.95%的差距不大，本研究也額外收集了 9 類花粉與原本 23 類資料合併成一 32 類的資料集，以同樣的實驗步驟得到 95.67%的準確率，說明了 Simple CNN 對於花粉粒分類的泛化性。由於傳統以特徵提取方式能夠對花粉粒特徵作觀察，本研究也透過卷積層的特徵圖可視化觀察花粉粒的特徵，發現同類的花粉粒其所得的特徵數以及過濾器的位置會相似，不同類的花粉粒可由顏色特徵、紋理特徵或形狀特徵去分辨其差異。

關鍵字:深度學習、卷積神經網路、資料擴充、特徵圖可視化

Abstract

Pollen grains is widely used in many fields. For example, bee pollen has high nutritional value and contributes to the beauty and health of human beings, pollen may cause pollen allergies, pollen fossils can be used to study the reconstruction of paleoenvironment and present a variety of paleoclimate features. The above application shows that the pollen grains have a very high research value. The classification of pollen grains is traditionally artificial. However, manual classification not only requires experts to classify pollen grains, showing that the classification of pollen grains requires high proficiency and cost, but also when the dataset is more than 1000 images, the classification task will become time-consuming. Therefore, the concept of automatic classification of pollen grains is generated. Generally, for the classification of pollen grains, the image processing method is used to extract features at first, after obtaining features, the next step is to classify with machine learning methods, the accuracy rate is approximately between 64% and 95%. As we mentioned, feature extraction is time consuming. Above all, if the appropriate features are not extracted, the accuracy of the classification will be affected. In recent years, due to the vigorous development of deep learning, Convolutional neural network(CNN) has also been applied to the study of pollen grains. Because the convolutional layer of the CNN has the feature extraction function, it is able to eliminate the need for complex pre-processing steps and achieving high classification accuracy. Daood et al. classified the pollen grains with CNN in 2016, reaching an accuracy of 89.95% after data augmentation and transfer learning. This study used a dataset containing a total of 805 images, 23 classes. The results of the experiment include two phases. In the first phase, 805 images were divided into grayscale and RGB with CNN classification of three different architectures. The best accuracy is RGB images with Simple CNN achieving 81.55%, which is a significant improvement over [1] using the same data set. The second part using general data

augmentation augmented the images to 8714 by rotating, resizing, and shearing, and then using Simple CNN to establish a predictive model, the experimental results have achieved an accuracy of 95.95% and a precision of 96.09%. Compared with the research using deep learning, there is not much difference in the accuracy of 89.95% obtained from [2] containing 30 types of dataset, also, our research collected 9 classes of pollen grains and combined with original dataset, following the research flowchart, the 32 classes dataset reached 95.67% of accuracy, proving the generalization of Simple CNN. Because the traditional feature extraction method can explain the importance of the extracted pollen features, this study also observed the features of the pollen grains through visualization of the convolution layer, and found that the same type of pollen grains, the locations of filters will be similar and different types of pollen grains can be distinguished by their different color features, texture features or shape features.

Keywords: Deep learning, convolutional neural network, data augmentation, visualization of feature map

目錄

第一章 緒論.....	1
1.1 研究背景與動機.....	1
1.2 研究目的.....	5
1.3 論文架構.....	5
第二章 文獻探討.....	6
2.1 花粉分類相關研究.....	6
2.1.1 影像處理做特徵提取.....	6
2.1.2 機器學習對於花粉粒分類的研究.....	8
2.2 深度學習.....	10
2.2.1 卷積神經網路.....	10
2.2.2 深度學習對於花粉粒分類的研究.....	13
2.3 資料擴充.....	13
2.4 特徵可視化.....	14
第三章 研究方法.....	16
3.1 研究流程描述.....	16
3.2 花粉粒圖像資料集.....	17
3.3 影像前處理(Pre-processing).....	18
3.4 分割資料.....	18
3.4.1 K-折交叉驗證.....	18
3.4.2 Holdout 驗證.....	19
3.5 卷積神經網路.....	19
3.5.1 卷積神經網路基本架構.....	19
3.5.2 LeNet-5 架構.....	19
3.5.3 AlexNet.....	20
3.5.4 Simple CNN.....	21
3.5.5 特徵可視化.....	22
3.6 資料擴充.....	23
第四章 實驗與結果.....	24
4.1 實驗環境.....	24
4.2 參數設定.....	24
4.3 實驗結果.....	26
4.3.1 第一階段實驗結果.....	26
4.3.2 特徵可視化.....	31
4.3.3 資料擴充.....	42
4.3.4 第二階段實驗結果.....	42
第五章 結論與建議.....	49

5.1 結論.....	49
5.2 未來展望.....	50
參考文獻.....	51



表目次

表 2.1 花粉粒分類的研究整理.....	9
表 4.1 參數設定.....	25
表 4.2 灰階圖以三種卷積神經網路搭配 10 折交叉驗證訓練結果.....	27
表 4.3 彩色圖以三種卷積神經網路搭配 10 折交叉驗證訓練結果.....	28
表 4.4 資料以交叉驗證分割並以資料擴充前後作比較.....	43
表 4.5 使用 Holdout 驗證並以資料擴充前後作比較	44
表 4.6 資料以交叉驗證分割並以資料擴充前後作比較(32 類).....	47
表 4.7 與使用一般機器學習方法比較.....	48
表 4.8 與使用深度學習的論文比較.....	48



圖目次

圖 3.1 研究流程圖.....	16
圖 3.2 23 種類的花粉粒資料集.....	17
圖 3.3 LeNet-5 示意圖(參考[20]的架構改寫繪成).....	20
圖 3.4 AlexNet 示意圖(參考[20]的架構改寫繪成).....	21
圖 3.5 Simple CNN 示意圖(參考[20]的架構改寫繪成)	22
圖 4.1 彩色圖以 Simple CNN 訓練所得混淆矩陣.....	29
圖 4.2 Arecaceae 圖像資料集	30
圖 4.3 Syagrus 圖像資料集.....	30
圖 4.4 Anadenanthera 以 Simple CNN，學習率=0.001 做特徵可視化 ..	31
圖 4.5 Anadenanthera 以 Simple CNN，學習率=0.01 做特徵可視化	32
圖 4.6 Anadenanthera 以 Simple CNN 做特徵可視化(灰階圖).....	33
圖 4.7 Anadenanthera 以 Simple CNN 做特徵可視化(相同類別比較)...	34
圖 4.8 Anadenanthera 以 Simple CNN 做特徵可視化(相同類別比較)	34
圖 4.9 Anadenanthera 以 Simple CNN 做特徵可視化(相同類別比較)...	35
圖 4.10 Arecaceae 以 Simple CNN 做特徵可視化(不同類別比較).....	36
圖 4.11 Syagrus 以 Simple CNN 做特徵可視化(不同類別比較)	36
圖 4.12 Eucalipto 以 Simple CNN 做特徵可視化(不同類別比較).....	37
圖 4.13 Matayba 以 Simple CNN 做特徵可視化(不同類別比較).....	38
圖 4.14 Anadenanthera 以 Simple CNN 做特徵可視化(三層卷積層).....	39
圖 4.15 Anadenanthera 以 AlexNet 前兩層卷積層做特徵可視化.....	40
圖 4.16 Anadenanthera 以 AlexNet 第五層卷積層做特徵可視化.....	41
圖 4.17 資料擴充範例.....	42
圖 4.18 資料擴充後的混淆矩陣.....	45
圖 4.19 從 The Global Pollen Project 所收集而來的 9 個類別的花粉粒 圖像資料集.....	46

第一章 緒論

1.1 研究背景與動機

花粉粒研究能帶來許多好處，其中[3]提到在考古學、醫學的領域中，考古學與花粉化石的觀察有關，藉由花粉化石的觀察來了解古生態環境的分布，醫學則是在尋找芳香烴化合物此類致癌物質有相當大的貢獻，甚至在化學領域等花粉也具有很大的商業意義。[4-6]也提到了花粉化石廣泛的用於古生態環境及古氣候的重建。花粉是造成過敏的一個因素，所以研究花粉對於有過敏體質的人也有不小的幫助[6-8]。而花粉粒可用於對蜜蜂採集的產物做地理相關的原理推導[6]和在蜂蜜孢粉學的研究上也有一定貢獻[1, 9]。[10, 11]則是由於某些植物瀕臨滅絕，對於植物多樣性會造成影響，研究植物的分布來達到對植物的保護，因而收集花粉來研究。

花粉粒的用途之多，而最常見的研究為對花粉粒做分類，如:[5]提到了一些在伊比利半島上的蕁麻科類的植物非常豐富，對於研究過敏的學者來說，區分出植物的種類可以提供給臨床醫生在診斷上的幫助，[6]提到了花粉的種類可用於定位特定物種，像是觀察不同種類的花粉粒可以發現到影響到特定物種的疾病以及掠食者，讓生物學家可以改善方法來保護這些特定物種，[12]也提到了花粉的分類與蜂蜜花粉學有相當程度的關連性，而蜂蜜花粉學在處理蜜蜂產品花粉粒研究有著很大的重要性，相異的花粉種類代表了蜜蜂產品的植物起源不同。

過去關於花粉粒分類的研究強調通過計算機視覺系統自動識別花粉種類的重要性，[13]認為自動化的花粉粒分類可帶來許多好處，如:更快速獲得結果、訓練大量的資料及較高的分辨率…等，[4]也提到以花粉粒用來幫助古環境的重建需要有更多的場地、數量大的資料集、更快的速度以及客觀性，說明了自動化的必要。

花粉粒的分類研究需要經過一連串的步驟，一開始必須先獲得花粉，取得花粉後先使用光學顯微鏡(Light Microscope, LM)或者電子顯微鏡(Scanning Electron Microscope, SEM)去觀察後，再利用軟體將含有花粉粒的部分切割出來，像是比較早期的 Langford 與 Treloar 都使用 SEM 圖片去做研究，[13]使用了由 SEM 所觀察到的現代花粉圖片，而[12]則是使用了 LM 圖片去做分析。[3]提到了 LM 圖片的品質一般相較起 SEM 圖片較差，可能會導致準確率的下降。[14]提到 SEM 圖片的取得較昂貴並且難以取得，所以最近的研究也說明了可以用 LM 圖片來做分類。[1]則談到蜂蜜孢粉學、外國孢粉學、古孢粉學通常都是利用 LM 圖片去對花粉粒做分類。

獲得 SEM 或 LM 的圖片後，接下來會使用特徵提取技術來取得花粉粒的特徵，而特徵提取的方法非常多，一般是使用影像處理的方法去對花粉圖像進行特徵提取，如[12]使用了小波轉換去先去對影像做處理後，再以分水嶺演算法去提取出紋理屬性，另外也用了 k 曲線演算法、形狀描述演算法提取出形狀屬性和 RGB、HSB 模型提取出顏色屬性。[10]使用了一連串的影像處理方法後，最後再以 Opencv 的 findContour 方法去提取特徵。[15]則是使用了 Gabor 轉換去計算紋理特徵和利用 invariant moment 獲得幾何特徵。[11]也使用了一連串的影像處理方法來取得紋理及幾何特徵。[1]使用了 Color, Shape, Texture(CST)、Bag of words(BOW)、灰度共生矩陣(GLCM)等技術提取出特徵。以上論文所使用的這些方法的好處在於可以對於自己想要觀察或認為重要性較高的特徵去做提取，而所提取的特徵大致上可以分為顏色特徵、形狀特徵、幾何特徵及紋理特徵，這些都是普遍在具有保證性的，不過有時還是會搭配一些其他方法去定義特徵的重要性，或者以降維的方法如主成分分析或線性判別分析將重要性低的特徵篩選掉。當上述步驟都完成，再替圖片標上類別標籤後，就等於建立好了花粉粒的資料集。

花粉粒資料集的類別數通常視所持有的資料多寡都不太一樣，如[13]使用了 6 類的現代花粉圖片，[3]使用了 4 類共 72 張的紋理圖片，[5]使用了共 3 類的圖

片，[6]使用了 3 類的圖片，[12]使用了從巴西中央區域收集而來的蜂蜜所提取出的花粉粒包含共 7 類的圖片，[16]使用了 7 個種類的圖片，以上都算是類別相較起來較少的資料集，多一點的類別資料如[11, 14]都使用到了位於哥斯達黎加的 17 種亞熱帶性的蜂蜜植物的種類共 345 張的圖片，再多一點如[1]使用了巴西薩凡尼亞花粉的圖片包含 23 類總共 805 張，不過此篇在分類準確率的結果上較差，推斷可能是在特徵提取的部分提取出的特徵重要性不足，並且沒有使用到合適的機器學習方法，而由於此資料集容易取得，並且類別上為過往研究中第二多的，因此本文的研究也採用同樣的資料集，嘗試是否能夠改善其結果。而目前最多的資料集則是[2]使用了 30 類共 1000 張的圖片。

當建立好資料集後，最後一個步驟就是利用到機器學習的方法來分類花粉粒的圖片並評估其分類的效果如何。機器學習的發展日新月異，如支援向量機、隨機森林、神經網路…等方法，近幾年隨著不斷的改良還是常常被用在各種不同的情況下做分類的工作，並且也都有被應用在花粉的分類上，像是[5]使用了神經網路，定義了三種多層感知機去對花粉粒做分類，[12]使用了三種方法分別是 K-最近鄰居法(K-nearest neighbor, KNN)、支援向量機(Support vector machine, SVM)、C4.5 決策樹，搭配 7 種不同的特徵排列組合方式去做分類，[1]使用了四種方法，包括支援向量分類器(C-Support vector classification, C-SVC)、序列最小優化算法(Sequential minimal optimization, SMO)、KNN、C48 決策樹，以 WEKA 軟體去做分類，[11, 17]則都使用了 SVM 跟人工神經網路(Artificial neural network, ANN)，其中[11]比較特別，先使用了獨立成分分析、主成分分析、線性判別分析等降維的方法才去做分類，[16]使用了 KNN，[18, 19]使用了 SVM，其它也有利用了隱藏馬可夫模型分類器[10]。這些方法各有其優缺點，而要如何去選擇是一件難事，通常都需要去嘗試過並使用一些評估方法後才能夠得知哪個方法所得的效果較好。除此之外，機器學習通常伴隨著其它比較耗時的資料前處理的方法，如前面所提到的許多特徵提取方法和特徵選擇，經常在使用機器學習方法訓練之前被應

用，而如何去定義特徵的重要性也是一件困難的事，甚至當資料的特徵太多時，也會使用到一些降維的方法。

由於上述的情況，深度學習開始變得受歡迎，其中最具代表性的就是卷積神經網路，最早提出的卷積神經網路模型架構為 1998 年由 Lecun 等人所提出的 LeNet-5，它包含了有著特徵提取功能的卷積層，有著特徵選取概念的池化層，藉由前兩部分將原始圖像轉換成一張張的特徵圖，最後再接到全連接層做分類。到了 2012 年，Krizhevsky 等人提出的 AlexNet 於大型的圖片資料庫 ImageNet 的分類競賽中以高出第二名 11% 的準確率奪冠後，引起了廣泛的關注。卷積神經網路的每一層都有著不同的功能，對於圖像資料的分類不用自己手動以影像處理的方式去提取特徵，也常常應用在人臉辨識、語音識別和物體偵測等方面，而且最重要的是，利用卷積神經網路可以減去一些繁瑣的前處理步驟，能夠避免神經網路的梯度消失問題，也能應付多類別的分類，在分類的結果上也能夠更加突出。基於以上優點，本研究決定以卷積神經網路對花粉粒圖片資料做分類。

過去的研究中也有人以卷積神經網路做過花粉粒的分類研究，[2]使用 MATLAB 以 MatConvNet 中的架構去對花粉粒做分類，不過使用深度學習時會存在的問題是當影像資料不足時，訓練結果可能會不符預期，而為了解決這個問題利用了資料擴充的方式將影像資料擴充到足夠的量，本研究也採用了資料擴充的技術。然而因為使用卷積神經網路不需經過手動的特徵提取前處理，無法對於特徵去做定義及描述，可說是一個黑盒子，所以本研究對於卷積層所卷積出來的特徵部分又進一步利用特徵可視化的概念去解釋卷積層的過濾器得到的特徵之間的差異。

1.2 研究目的

本研究的目的是通過使用卷積神經網絡的計算機視覺系統去自動分類各種巴西薩凡納花粉粒圖像，不僅可以提高分類的準確率，在資料前處理的過程中也提高了效率，雖然使用機器學習方法來分類花粉圖像是常見的，有時也可以達到高準確率，不過那是因為資料前處理提取了足夠的重要特徵來作為分類的依據。而每幅圖像都有很多不同的特徵，要將它們全部提取出來，並且定義每個特徵的重要性並非易事。相對來說，卷積神經網路可以提取圖片的像素作為特徵來分類圖像。本文在第一部分透過三種不同架構的卷積神經網路去對 805 張的花粉粒圖片資料集去做分類，由於認為所得結果還有提升的空間，所以在第二部分以資料擴充的方式將圖片擴充約 10 倍的張數去進行結果的比較，評估結果則以 10 折的交叉驗證去看分類準確率，也用了 Holdout 方式去做驗證。此外，本研究也透過特徵可視化去觀察到同類別花粉粒相似的特徵與不同類別的花粉粒之間的特徵差異，了解卷積神經網路是如何可以在分類上得到好的效果。

1.3 論文架構

本論文的第一章是緒論，主要介紹研究背景和動機，以及實驗的目的。第二章是相關的文獻探討，包括花粉粒相關研究、深度學習、資料擴充以及特徵可視化的介紹。第三章是本文使用到的研究方法，闡述了卷積神經網路三種架構的比較，還有資料擴充的概念及技術，第四章是實驗和結果，將本論文的想法付諸實踐並實作，第五章是結論和未來展望。

第二章 文獻探討

2.1 花粉分類相關研究

在花粉分類的相關研究中，大部分研究的基本流程都是先透過管道獲得透過光學顯微鏡或電子顯微鏡所觀察到的花粉或花粉粒的圖片，假如為花粉的圖片則要再透過影像切割的方法將花粉粒的部分切割出來，而得到的圖片經整理後變為一個影像的資料集，接著會對於影像資料及標上類別標籤後，做特徵提取的處理，提取出來的特徵再藉由機器學習的演算法將影像分類並獲得分類的準確率。總括來說，花粉粒的研究目的就是希望能夠比傳統人工方法達到更快速、並且更精確的自動分類。

2.1.1 影像處理做特徵提取

Langford 等人的研究[13]以測試自動化分類為目標，使用了六種現代花粉粒的 SEM 圖片，實驗首先將圖片數位化並提取出花粉粒的外部紋理得到 192 個樣本後去做紋理分析，接著計算每個樣本的灰度共生矩陣和計算紋理的測量值，將其作為分類程式的輸入，最後使用留一法和經過變量選擇的過程，得到 94.3% 的正確分辨率。

M. Rodr'iguez-Dami'an 等人的研究[6]提出了透過亮度以及各種形狀描述器去對花粉粒做分類。在研究過程中，資料使用了 Urticaceae 科的花粉粒，接著利用到了花粉粒的描述器去計算花粉粒二元圖的外圍邊界以及幾何特徵，幾何特徵包括了面積、周長、直徑...等，除此之外也利用了區域矩、邊界矩、傅立葉描述器和亮度來對花粉粒的特徵做描述，經過這些處理後使用到最小距離分類器(MDC)去做分類，其中以亮度特徵做分類達到 79% 的準確率，而以傅立葉描述器所得特徵經分類後達到了最高的準確率為 91%。

Diogo Soares da Silva 等人的研究[12]為了協助分析蜂蜜樣品的植物來源而對於花粉粒的自動分類提出了一些建議。資料集是由 7 種從巴西中央區域的蜂蜜中

所取出來的花粉粒，首先以高斯平滑去做花粉粒的偵測並使用小波轉換的方法切割出花粉粒的圖片，再來分別使用到 k 曲線演算法和形狀描述器提取出形狀特徵、RGB 模型及 HSB 模型提取出顏色特徵，以及分數樣條轉換提取出紋理特徵，經過特徵提取的步驟後，接著使用 Weka3.6 軟體中的機器學習方法包含 C4.5 決策樹、SVM 和 KNN 分類，而其分類結果以 F-score 來評估，達到了 79%。

Norma Monzón García 等人的研究[10]認為地球的生物多樣性由於人類所造成的污染而遭受考驗，所以透過自動分辨花粉種類の研究有助於特定物種的識別及相關的疾病和掠食者。資料集採用由哥斯大黎加的蜂蜜植物中收集而來 11 種不同家族，包含 17 類的花粉粒圖像。研究的設計上一開始先對於花粉粒以影像處理的方式做處理，分為二值化、去除噪音及輪廓提取三個步驟，處理完後接著做特徵提取，提取出角度的特徵，最後再以隱藏馬可夫模型分類器(HMM)做分類，達到了 98.77%的成功率。

Marcos del Pozo-Baños 等人的研究[14]提到世界上有 20%的植物瀕臨滅絕，保護地球的生物多樣性是全球性的重要工作，因此透過收集花粉的資料並在實驗室內研究藉此來觀察植物的分布來達到保護植物的目標。資料集的部分與[10]一樣也是使用哥斯大黎加的蜂蜜植中收集而來花粉粒圖像，包含 17 類共 345 張的圖像。研究的過程首先對影像做前處理的工作，分為解相關延伸將對比拉開、RGB 轉 HSV 後取出對比較明顯的 S(飽和度)通道、做直方圖均衡化、二值化以及遮罩，經過前處理工作後，接著套用到特徵提取的技術提取出兩大類的特徵，分別為幾何特徵和紋理特徵，其中幾何特徵包含面積、邊界框、凸面區域、周長…等共 24 種特徵，而紋理特徵則是利用 GLCM 去計算來得到，如:對比度、相關性、熵、同質性…等共 26 種特徵。當特徵提取的步驟完成後，ANN 被使用來做分類並搭配 3、5、7、10 折的交叉驗證做評估，其中得到最高的準確率的結果為 10 折交叉驗證所得的 92.81%。Marcos del Pozo-Baños 等人在之後又對此研究做了一些改進[11]，前處理以及特徵提取的部分保持不變，不過在將特徵利用到機

器學習方法的分類以前先使用了一些維度轉換的技術，有主成分分析、獨立成分分析、線性判別分析及離散餘弦轉換等方法，藉此達到降維的效果。然後在分類的部分除了原本的 ANN 外也用了 SVM 去觀察分類結果，評估方法則跟原本一樣用了 3、5、7、10 折的交叉驗證，其中最好的結果為使用了線性判別分析並搭配 SVM 結合 ANN 分類器，再以 10 折交叉驗證得到 95.27% 的準確率。

Ariadne Barbosa Gonçalves 等人的研究[1]談到花粉粒的分類在法醫孢粉學、考古孢粉學及蜂蜜花粉學等領域都非常重要，同時也有許多用途如控管蜂蜜產品的品質及犯罪證據的鑑定。此篇資料集是由巴西的一群養蜂人各提供了 1kg 的蜂蜜經分析並以 LM 觀察，再以 CorelDRAW® 切割出花粉粒的圖片，總共包含了 23 類，805 張的圖片。其研究的步驟首先利用到了兩種特徵提取的技術包括 CST 與 BOW，CST 取出的顏色特徵包括色相、飽和度跟亮度及紅色通道的部分，形狀特徵包括花粉粒的面積與環境的面積比、圓形率與 k 曲線演算法所取得的特徵，紋理特徵則是由灰度共生矩陣計算出的對比度和熵。而 BOW 是利用加速穩健特徵的技術去偵測並描述影像中的重要特徵並搭配 K-means 提取出來，再拿中心直代表此類別，而此類別也就是單字，將這些單字組合起來就是我們訓練出來的全部特徵。提取出特徵後，再來利用到 Weka3.7.9 軟體中的機器學習方法含 SMO、C-SVC、J48 決策樹、KNN 去做分類。最後再以正確分類率(CCR)、F-Measure 和 AUC 曲線做評估，並繪出混淆矩陣查看分類的情況，而其中所得的最高 CCR 為以 C-SVC 訓練所得的 64%。

2.1.2 機器學習對於花粉粒分類的研究

Ping Li 等人的研究[3]認為 SEM 圖片不易取得，所以以 LM 圖片去做研究，期望也能達到高準確率，他們使用了 4 類的紐西蘭花粉粒圖片包括 *Hoheria populnea*、*Phormium tenax*、*Phymatosorus novaezealandiae* 及 *Podocarpus totara*，每一類含有 18 張，4 類總共 72 張影像作為資料集，實驗先以紋理分析取出 8 個特

徵，然後以三種多層感知機模型，架構從輸入層到輸出層的神經元數分別為 8-3-4、7-3-4、5-2-4 並搭配留一法去做分類，輸入層神經元數的減少也代表了使用的特徵數的減少，藉由多層感知機(Multi-layer perceptron, MLP)架構的改變，分類效果也隨之提升，因此文中提出了改善 LM 圖片分類準確率的途徑，透過簡化特徵、減少樣本來達到 100%的分類準確率並且勝過傳統 Fisher 的線性判別方法。

I. France 等人的研究[5]提到化石花粉的分類有助於重建偏遠古環境，並強調與先前的研究專注在如何去分類相比，他們是以如何用合適的影像當作輸入為重點。而研究步驟首先以感光耦合元件並搭配階層式的前處理過程後獲得完整的顯微鏡影像，接著將影像以物件辨識和把花粉以外的灰塵、木屑移除後，利用到天堂神經網路將花粉粒及其他物件分開並將相似的花粉粒分為同一類，其中天堂神經網路的架構包含輸入層、特徵提取層、圖樣偵測層，最後再連接到分類層，透過天堂神經網路訓練 204 張總共 3 類的圖片，可以達到平均 80%以上的準確率。

本文將上述有關花粉粒的相關研究整理為一表格(見表 2.1)內容包括資料及大小、特徵提取方法、提取出的主要特徵、所使用的機器學習方法以及評估方法，方便比較過去研究之間的差異。

參考文獻	資料集大小(張數)/類別	特徵提取方法	提取出的主要特徵	機器學習	評估方法
[13]	192/6	灰度共生矩陣	紋理特徵	統計方法	留一法
[3]	72/4	紋理分析	紋理特徵	MLP	準確率
[5]	204/3	神經網路	幾何特徵	NN	準確率
[6]	77/3	傅立葉轉換形狀描述器	幾何特徵、形狀特徵	MDC	準確率
[12]	210/7	K曲線演算法形狀描述器	顏色特徵、形狀特徵、紋理特徵	決策樹(C4.5)、SVM、KNN	F-Score
[10]	345/17	影像處理	幾何特徵、角度特徵	HMM	準確率
[14]	345/17	影像處理	顏色特徵、幾何特徵、紋理特徵、形狀特徵	ANN	準確率
[1]	805/23	CST、BOW	顏色特徵、幾何特徵、紋理特徵、形狀特徵	決策樹(J48)、SMO、SVC、KNN	正確分類率
[11]	345/17	影像處理	顏色特徵、幾何特徵、紋理特徵、形狀特徵	ANN、SVM	準確率

表 2.1 花粉粒分類的研究整理

2.2 深度學習

機器學習已經盛行的一段時間，從最早的人工智慧到機器學習再到現在的深度學習，不僅結構上越來越複雜，訓練的結果大部分也都能夠有效提升。在機器學習的歷史裡，具有一定代表性的就是神經網路，一開始的神經網路只有單層，發展一段時間後有了隱藏層的概念再加上倒傳遞演算法及梯度下降法後出現的多層感知機，現在則是有架構更為完善的卷積神經網路，在特徵提取的部分相較於傳統手工方法上更高效，可以利用在影像、視訊及語言等領域上，而在影像辨識的方面尤其常見，像是 MNIST、CIFAR10 皆是將影像的像素值作為輸入，中間經過卷積、池化、DROPOUT、歸一化等處理再輸出，而其中我們所知的像素值包含 0~255 的值，彩色圖又有 RGB 三個通道，在訓練上會較為耗時，因此深度學習在硬體上的設備也比較講究。

2.2.1 卷積神經網路

Yann Lecun 等人的研究[20]提出了 LeNet-5 卷積神經網路的概念，由於當時對於圖像資料及文件的辨識皆要透過多種的混合模型包括分割、特徵提取後才能做到辨識，而利用卷積神經網路達到上述工作的整合。傳統的神經網路輸入資料只能為一維，而相較於神經網路常用的 MLP 來說，CNN 也可以用倒傳遞演算法去訓練，但輸入資料可以不限於一維，因此保留了一些形狀及空間的資料，同時圖像資料是由許多像素值所組成，彩色圖像的像素值數量更是灰階圖的三倍，以 CNN 去做訓練的好處是可以達到局部感知、權值共享以及下採樣等功能，其中局部感知能夠做到特徵提取，藉由將範圍限制在局部去提取圖像的特徵；權值共享可以避免神經網路參數太多而造成的內存不足問題，不過相對來說就需要硬體設備上的支持；下採樣在實現了形狀不變性的情況下並且達到了降維的效果。LeNet-5 架構包含輸入層及輸出層共有 8 層，順序分別為輸入層→卷積層→池化層→卷積層→池化層→全連接層→全連接層→輸出層，卷積層的作用是特徵提取，

藉由設定一個框架，並以滑動的方式將特徵一一做提取。池化層的作用是對提取出的特徵做特徵選取的工作，可以降低網路的訓練參數以及過度擬合的程度。全連接層的作用則是對於最後提取出的特徵去做分類。而當時由於數據的量比較少並且缺乏運算能力，所以 LeNet-5 在對於較複雜或是規模較大的資料集上的問題表現較差。

Alex Krizhevsky 等人在 2012 年的研究[21]提出了 AlexNet 卷積神經網路的架構，相比之前的卷積神經網路架構有更多層的深度，也更為複雜，此多層的架構共包含了 5 個卷積層、3 個池化層、3 個全連接層，這些層的功能與原本並無不同，不過比較特別的地方在於有調整步長及填充的參數來控制卷積及池化，例如：在第一層卷積層設定步長為 4，這樣做是為了讓卷積的區域更少的重疊，保持較小的空間維度，而在第二個卷積層則是將填充設定為 2，代表在圖像邊界的上下左右方向各補上兩排像素值為 0 的填充，目的是為了最大化的儲存前一層輸入的資訊，並提取低層次的特徵。此外，在卷積層之後會連接到一個非線性層的處理，LeNet-5 架構時使用的是 sigmoid 函式，但在隨著層數的增加可能會遇到梯度消失的問題，梯度消失指的是神經網路在反向傳播的過程中是每一層對函數做偏微分的相乘，當激活函數使用的是 sigmoid 函數時，會是 0~1 之間的數相乘，相乘後會越來越小並接近於 0，導致淺層的權重無法更新。而選用線性整流函數 (ReLU) 作為激活函數是解決梯度消失問題的一個辦法，因其分段線性性質在輸入 $x \geq 0$ 時，會輸出 x ， $x < 0$ 時，則輸出 0，這樣子的做法除了可以解決梯度消失的問題外，也讓神經網路適度的有了稀疏性，可以緩解過度擬合的情況，而用了 ReLU 的另一個優點則在於不須對深度神經網路做預先的訓練，在資料量大時會更加有效率。所以 AlexNet 使用了 ReLU 作為替代，讓前一層的輸入當中的負值變為 0，藉此提高了模型的非線性特性。文中也提到了由於模型的參數太多，可能會有過度擬合的情況產生，所以也用到了兩種方法去降低過度擬合，第一種方法是資料擴充，第二種方法則是 Dropout，概念是透過讓網路中的一部分神經元

的權重暫時不做更新，減少神經元之間的關聯性，藉此達到泛化的效果。AlexNet 的架構出現後證明了深度 CNN 的有效性，同時也伴隨著演算能力的進步，讓卷積神經網路的發展有了更多的可能性。

Nitish Srivastava 等人的研究[22]提到具有大量參數的深度神經網路非常強大，不過存在著過度擬合的問題，表示模型對於訓練資料集有好的表現，但在應用到測試資料集時，準確率卻明顯的下降，而要降低過度擬合的其中一個方法就是使用到 Dropout 技術，Dropout 的作法為在神經網路訓練的過程中，適時的移除一部分的神經元，使其權重不會隨著正反向的傳播時而更新，這樣可以使神經元不會對於某些特徵的權重調整成太多，因而使訓練的結果發生過度擬合，並且由於移除神經元的過程是隨機選取的，可以使神經元之間的互相影響的作用降低。實驗也證明了使用 Dropout 相較於使用正規化方法在降低過度擬合的程度有更大的改進。Dropout 最常見的參數設定上是 0.5，代表了隨機移除選定的那一層的 50% 的神經元去做訓練。

深度學習中有許多種類的優化器，如 SGD、Adagrad、RMSprop、Adadelata 及 Adam...等，SGD 是常見的隨機梯度下降法，在梯度的更新速度上快，不過損失函數的值會比較不穩定，也容易受困在局部最佳解。Adagrad 可對低頻的參數做較大的更新，並在稀疏的數據上能有好的表現，不過缺點是學習率會隨訓練時間的增長而快速下降。RMSprop 與 Adadelata 都是 Adagrad 的改良版，可以解決學習率急遽下降的問題。Diederik P. Kingma 等人的研究[23]提出了 Adam 優化器的概念，Adam 結合了 Adagrad 與 RMSprop 兩者的優點，與兩者同樣是自適應學習的方法，對內存的需求小，收斂快，並能夠對於非平穩目標有好的處理，而當梯度變的稀疏時，效果也會更好，適用於大的數據集以及高維的空間，實驗證明與其他優化器相比更有優勢。

2.2.2 深度學習對於花粉粒分類的研究

Amar Daood 等人的研究[2]講到花粉的鑑定有助於法醫學家解決未破案的犯罪，為氣候變化建模者提供數據，甚至暗示潛在的石油探勘場所。而作者也提到過去對於花粉粒的特徵都是要先使用設計過的辦法所取得，但是取得的特徵可能不夠明顯，訓練的效果就會較差，因此使用了卷積神經網路去對花粉做分類。其資料集分為兩種，一個為包含 1000 張圖、30 類的 LM 資料集，另一個為含 1161 張圖、30 類的 SEM 資料集，而由於擔心資料不足可能會產生過度擬合的問題，所以將資料集透過資料擴充的方式將圖片旋轉產生對電腦來說不同的圖片，擴充到 14000 張及 15000 張，擴張後再以 MatConvNet 的架構去訓練，過程中為了改善分類準確率使用了轉換學習的方法，將卷積神經網路架構的前兩層以 MatConvNet 做預先訓練，其他層以 ImageNet 做預先訓練，不過由於原本的 ImageNet 模型參數太多，所以將過濾器的數量減少以降低計算的困難。最後由卷積神經網路所得的精確率 LM 資料集與 SEM 資料集分別可達到 92.04%及 95.00%。

National Chung Hsing University

2.3 資料擴充

資料擴充為當資料量不夠多時，訓練的準確率沒有得到預期成效，或者訓練的結果發生過度擬合的情況時可以考慮使用的一個方法，其概念為對原有的圖像資料去進行幾何轉換，如:旋轉、翻轉、平移、縮放...等方法，使圖像看起來與原圖相似，不過對於電腦來說是不同的圖像，而在這樣的處理過程中，圖片的像素值不會改變，確保了原本的特徵是不變的，但像素值的位置會不同。因此，藉由資料擴充的技術，可以將不足的資料擴張到自己認為足夠的數量。

Denise Klep 的研究[24]試圖利用最有效率的資料擴充的方式去對英文單字的手寫辨識資料集做擴充，觀察能否在訓練的效果上更加優化。步驟上首先取了 2400 張的手寫辨識圖像，對於其用不同的方法擴張到 9600 張的圖像，而方法包

括增加噪音、旋轉、翻轉、尺度變換、切割...等，其中也有以結合的方式去做擴充，當中所得到的最好的結果為使用逆時針旋轉 10 度，並向左平移六個像素。另外作者也有提到在此研究中，切割、增加高斯白噪及尺度變換這幾項方法的效果較差。

Jason Wang 等人的研究[25]提到了傳統的資料擴充方式簡單並且有效，不過使用生成對抗網路(GAN)及神經網路來做資料擴充也是值得研究的方法。實驗藉由將 ImageNet 的三種資料集包含狗與金魚、狗與貓以及 MNIST 資料集限制在一個小的子集，每一個子集被分割為兩部分，其中 80%為訓練資料集，20%為驗證資料集，來比較多種資料擴充的方式成效如何，結果狗與金魚的部分是神經網路加上沒有損失的準確率最高，狗與貓是用傳統的資料擴充方法所得準確率最高，MNIST 資料集則也是用神經網路加上沒有損失的準確率最高，不過其中狗與金魚和狗與貓的資料集不管用了哪種資料擴充方法準確率幾乎都是顯著增加，而對 MNIST 資料集效果不大，此篇作者推測應是原本的卷積神經網路架構對其的訓練已經結果夠好。從結果來看，不管使用傳統方法或是 GAN 亦或神經網路，資料擴充為一個提高分類準確率上有保障的方法。

2.4 特徵可視化

深度學習往往能夠獲得在分類上得到好的結果，不過解釋模型如何得到好的結果是一件重要的事，以花粉粒的分類為例，一般花粉粒的分類先做特徵提取再經機器學習演算法分類，因此可以透過對提取出的特徵做分析去解釋分類的結果，而深度學習是透過卷積層所卷積出特徵再連接到全連接層做分類，所以對於所卷積的特徵的觀察必須利用到視覺化的技術，其好處是可以了解模型是如何運作並判斷分類不精確的原因為何。

Matthew D. Zeiler 和 Rob Fergus 的研究 [26]認為 Krizhevsky 等人的論文展示了令人印象深刻的分類結果，然而他們沒有詳細解釋為何分類表現如此出色，

或者他們還有哪些改良的空間，所以 Zeiler 和 Fergus 使用了一種新的特徵可視化技術，透過此技術可深入了解到卷積神經網路每一層對於特徵的處理功能，選擇到合適的卷積核大小，以及分類器又是如何藉由特徵來分類，同時可視化技術也讓他們找到了勝過 Krizhevsky 等人研究的模型架構，並且此架構也適用於其他的數據集。

Luisa M Zintgraf 等人的研究[27]認為在醫療影像方面必須要精確，而被視為黑盒子分類器的卷積神經網路雖然能獲得好的分類結果不過需要對於模型有明確的解釋才能在醫學上提供決策上的輔助，因此他們提出了經由可視化深層的卷積神經網路對特定輸入所卷積出的圖像的差異來做預測及分析，該方法使用了更強大的多變量模型改進了以前方法的缺點，對於輸入影像的重要特徵區域有突出的效果，並且為分類器的決策過程中提供更多的訊息，讓卷積神經網路分類的效果提升，也增加了醫療領域上的接受性及可用性。



第三章 研究方法

3.1 研究流程描述

本篇論文的實驗流程如圖 3.1 所示，大致可分為四個步驟。

- 1.前處理:將所有影像轉換成相同的尺寸 50*50 以及 227*227，接著保留一份彩色圖，複製一份圖像轉為灰階圖。
- 2.第一階段以三種 CNN 架構包含 LeNet-5、AlexNet、Simple CNN 去做訓練得到分類準確率以及混淆矩陣，將第一階段的結果儲存並做特徵可視化。
- 3.第二階段由於認為得到的準確率可再提升，所以將結果較好的 RGB 圖片做資料擴充，擴充到 8714 張圖片做 10 折交叉驗證，以及 Holdout 驗證，測試資料維持不變，訓練資料擴充到 6256 張圖片。
- 4.以前面所得結果最好的卷積神經網路架構去做訓練得到分類準確率。

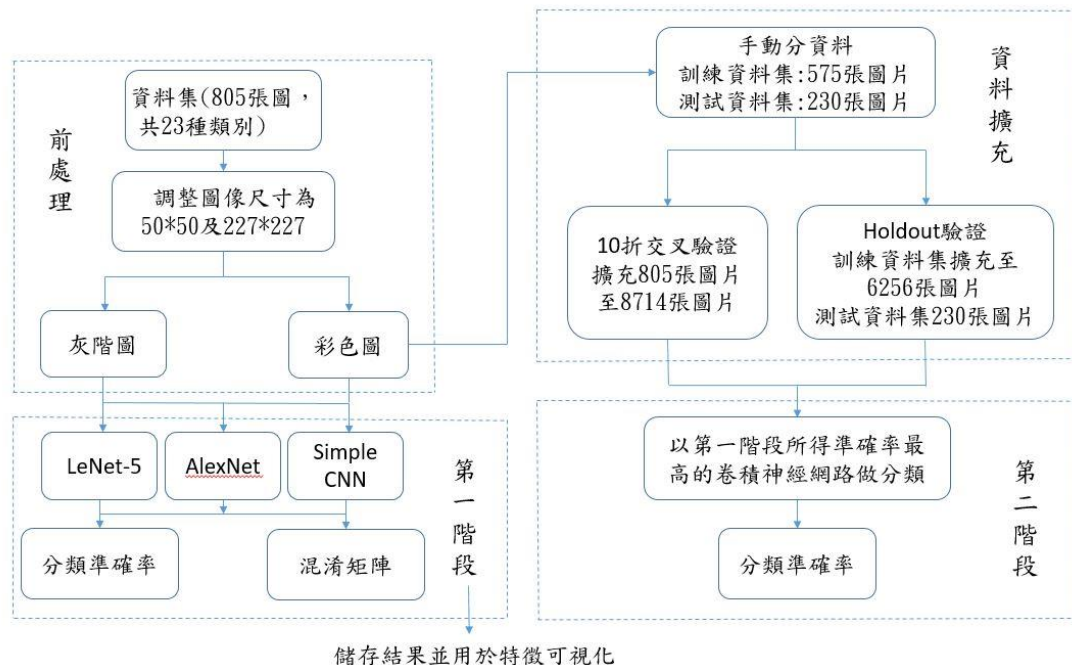


圖 3.1 研究流程圖

3.2 花粉粒圖像資料集





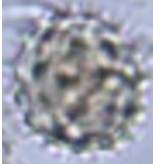

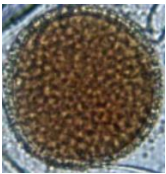
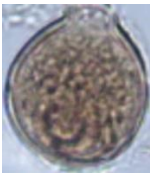


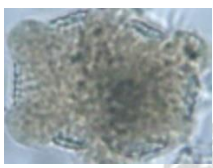
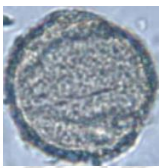
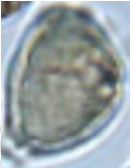

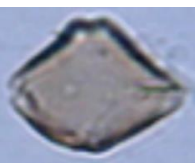






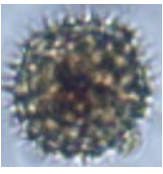
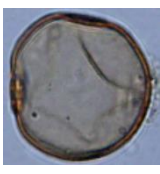
				
Anadenanthera	Arecaceae	Arrabidaea	Cecropia	Chromolaena
				
Combretum	Croton	Dipteryx	Eucalipto	Faramea
				
Hyptis	Mabea	Matayba	Mimosa	Myrcia
				
Protium	Qualea	Schinus	Senegalia	Serjania
				
Syagrus	Tridax	Urochloa		

圖 3.2 23 種類的花粉粒資料集

本文由於此資料集為公開的及容易取得，並且與過去研究相比，在圖像資料上類別為第二多的，分類上有一定的難度，所以使用了與[1]相同的資料集，而此資料集是由巴西的一群養蜂人各提供 1kg 的花粉經分析所產生，總共包含了 805 張共 23 類，如圖 3.2 所示。

3.3 影像前處理(Pre-processing)

在影像前處理的部分，由於圖像的大小皆不同，並且為因應卷積神經網路的輸入圖片必須要將圖片都調整成同樣的尺寸，所以在經過評估所有圖像的尺寸後，為了不降低比較小張圖片的圖像解析度，並保留適量的像素特徵，決定將所有圖像設定為 50*50 及 227*227。接下來將彩色圖像轉成灰階圖，雖然過去的論文也有不少以顏色作為特徵去做訓練，不過為了比較彩色圖像與灰階圖在分類準確率上的差距，並藉此來了解顏色特徵的重要性，因此實作了此轉換的步驟。此外，由於[15]在顏色的特徵提取部分只取了 R 通道，所以本篇也在前處理的部分特別將彩色圖片的 R、G、B 三個通道的像素值分別取出去做訓練。

3.4 分割資料

3.4.1 K-折交叉驗證

K-折交叉驗證的作法是先把資料集分為 K 份，其中一份作為測試資料集，另外 K-1 份則為訓練資料集，經過每一次交叉把其中一份作為測試資料集，其餘 K-1 份作為訓練資料集的步驟 K 次後，所得的 K 次結果會再以平均或其他結合的方式來獲得我們最終的結果。常用的 K 值包括 3、5、7、10 等，而本篇是以 K=10 做 10 折的交叉驗證。而一般在資料集較小時，使用交叉驗證可以從資料中獲得更多的訊息，同時防止樣本訓練過程中陷入局部最小值，以及降低過度擬合的情況產生。

3.4.2 Holdout 驗證

一般為了評估訓練好的模型對於新輸入的數據的預測能力，會將資料分為訓練資料集以及測試資料集，其中訓練資料集的資料會包含一個已知的輸出，模型會透過學習這些數據去建立一個準則，而測試資料集則是用來驗證模型對其的預測。常見的資料分割方式有訓練及測試分 60%/40%、70%/30%以及 80%/20%。

3.5 卷積神經網路

3.5.1 卷積神經網路基本架構

卷積神經網路的基本架構是由多個卷積層與一個全連接層所組成，而其特色為輸入的資料可以不限為一維，保留了一些形狀及空間的資料，因此在影像及語音辨識方面能獲得好的結果，卷積層又可稱為特徵提取層，會透過一個指定大小的視窗藉由滑動的方式提取出局部特徵再經過激活函數 Relu 輸出為特徵值做為下一層的輸入用，卷積層後通常會再連接池化層，目的是將圖片的尺寸縮小，並保留重要的特徵，同時達到降維的效果，而卷積層與池化層通常皆不一定只有一層，可以為多層。在卷積層之後通常都會連接上一個激活函數層，為的是讓輸入跟輸出脫離線性關係，傳統神經網路使用 sigmoid 或 tanh，卷積神經網路則以 ReLU 為主。經過上面幾層處理後，最後就會將特徵資訊平坦化並透過全連接層來進行分類。此外，層與層之間也常常會加上 Dropout 層，減少一些神經元的運作，來避免過度擬合情形的產生。

3.5.2 LeNet-5 架構

LeNet-5 架構由 Yann LeCun 所提出，如圖 3.3 所示。除了輸入層及輸出層以外總共含六層的架構。輸入層一開始輸入圖像，本篇使用 805 張 50*50 的圖像作

為輸入。接著第一層為卷積層，有六個卷積核，卷積核大小為 5×5 ，透過在圖像上滑動，步長為 1，從左到右、從上到下滑動後得出六張特徵圖，由於本篇所使用的圖像為 50×50 ，因此經第一層卷積後所得到的特徵圖尺寸會是 46×46 。第二層為池化層，可以使用平均池化或者最大池化，做法為先以 2×2 的格子來分區塊，如果是平均池化，就把每個區塊的 4 個格子的像素值作平均得出一個值，最大池化則是保留每個區塊最大的像素值，經處理後圖像大小會變為 23×23 。第三層還是卷積層，有 16 個卷積核，卷積核大小為 5×5 ，步長為 1，再做一次卷積的運算，得到 16 張 19×19 的特徵圖。第四層為池化層，同樣可使用平均池化或最大池化，以 2×2 的格子分區塊取平均值或保留最大像素值後，得到特徵圖的尺寸為 9×9 。第五層為全連接層，有 120 個神經元，透過此層將輸出的特徵圖轉換為 120 個 1×1 的特徵圖。第六層為全連接層，有 84 個神經元，而最後則視要分為幾類再接到輸出層，本篇有 23 類的花粉粒圖像，因此輸出層含 23 個神經元。

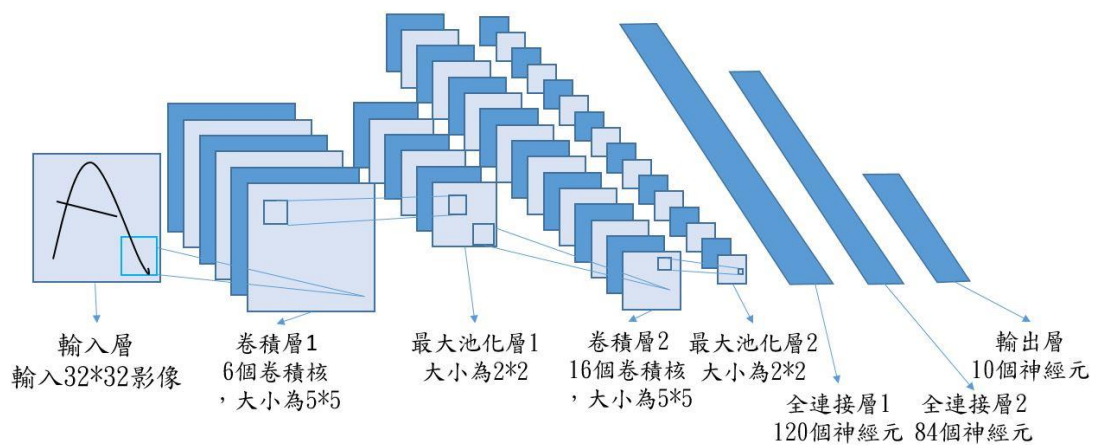


圖 3.3 LeNet-5 示意圖(參考[20]的架構改寫繪成)

3.5.3 AlexNet

AlexNet 架構由 Alex 在 2012 年所提出，證明了卷積神經網路在複雜模型下的有效性。而由於 AlexNet 模型較複雜，其輸入的圖像大小必須要有一定的尺寸，因此這邊在前處理的部分將圖像尺寸依照原 AlexNet 的輸入轉換為 227×227 。

AlexNet 為多層的架構，如圖 3.4 所示，第一層為卷積層，有 96 個卷積核，卷積核大小為 11×11 ，步長為 4，經過第一層的卷積後所得到的 96 張特徵圖大小為 55×55 。第二層為池化層，做最大池化，以 3×3 的格子分區塊，步長為 2，處理後得到 96 張尺寸 27×27 的特徵圖。第三層為卷積層，含 256 個卷積核，卷積核大小為 5×5 ，步長為 1，填充為 2，經卷積後得到 256 張大小為 27×27 的特徵圖。第四層為池化層，做最大池化，以 3×3 的格子分區塊，步長為 2，處理後得到 256 張大小為 13×13 的特徵圖。第五層及的第六層皆為卷積層，都有 384 個卷積核，卷積核大小為 3×3 ，步長為 1，填充為 2，經此兩層處理後得到 384 張大小為 13×13 的特徵圖。第七層為卷積層，有 256 個卷積核，大小為 3×3 ，步長為 1，填充為 2，經處理後得到 256 張 13×13 大小的特徵圖。第八層為池化層，做最大池化，以 3×3 的格子分區塊，步長為 2，處理後得到 256 張 6×6 的特徵圖。第九層與十層皆為全連接層，都含 4096 個神經元。第十一層為輸出層，含 1000 個神經元。近年來有許多深度比 AlexNet 更深，架構也更為複雜的模型，不過由於本篇在使用 AlexNet 架構所得的結果與期望相差甚遠，所以在評估過後決定不使用更為複雜的模型來實作。

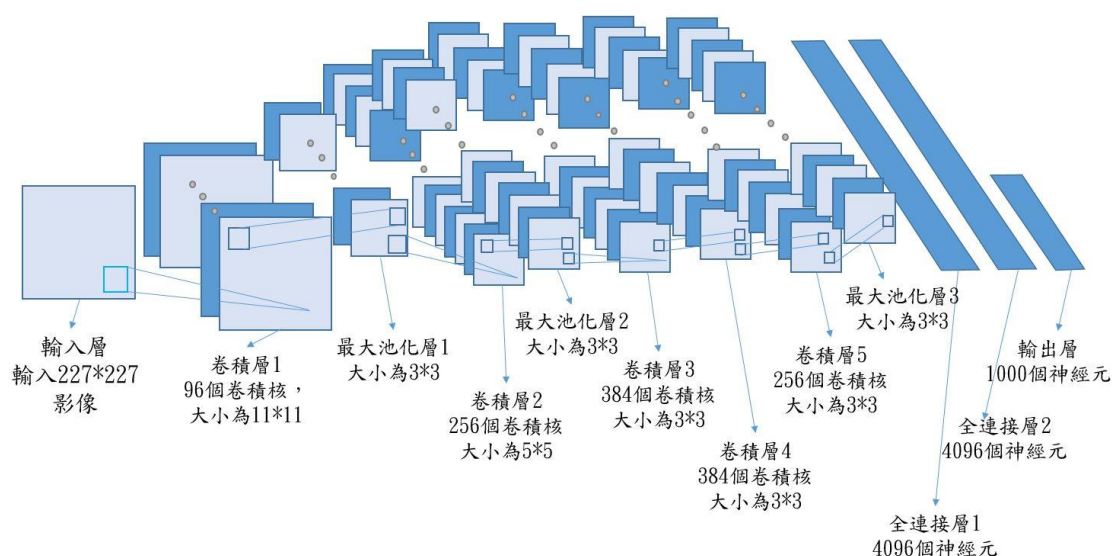


圖 3.4 AlexNet 示意圖(參考[20]的架構改寫繪成)

3.5.4 Simple CNN

本篇所使用的 Simple CNN 為經過 LeNet-5 改良後所得，如圖 3.5 所示，扣掉輸入層及輸出層總共包含五層的架構。輸入層輸入 805 張大小為 50×50 的圖。第一層為之後透過卷積視覺化觀察花粉粒的特徵因此將卷積核數量設定多一些，有 32 個卷積核，卷積核大小為 3×3 ，步長為 1，經過第一層的卷積後所得到的 32 張特徵圖大小會是 48×48 。第二層與原本 LeNet-5 沒有差異，為最大池化層，同樣以 2×2 的格子來分區塊，保留每個區塊最大的像素值，處理後得到 32 張尺寸為 24×24 的特徵圖。而有了[9]此篇論文做特徵提取技術在幾何特徵及紋理特徵共提取出 50 個特徵的想法，第三層經嘗試後將卷積核設定為 64 個，去看能否取得更多的特徵去做觀察，而卷積核大小為 3×3 ，經卷積後得到 64 張尺寸為 22×22 的特徵圖。第四層再做一次最大池化，以 2×2 的格子分區塊，保留最大像素值，處理後得到 64 張 11×11 的特徵圖。第五層連接到全連接層，設定 128 個神經元，將前一層輸出的特徵圖轉換為 128 個 1×1 的特徵圖，最後再連接到含有 23 個神經元的輸出層。

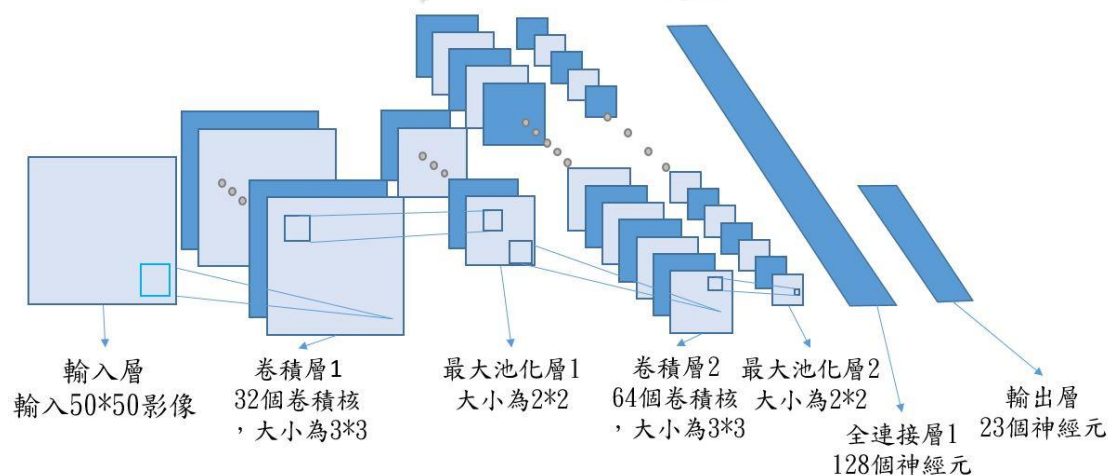


圖 3.5 Simple CNN 示意圖(參考[20]的架構改寫繪成)

3.5.5 特徵可視化

過去的論文對於花粉的分類都是先透過影像處理的方式去提取出特徵，再以機器學習的方法去訓練，此方式的好處在於可以對於特徵去做分析及觀察哪些特

徵的重要性可能較高，而同一類的花粉粒有哪些相似的特徵，不同類的花粉粒又是由於哪些特徵的差異而被分在不同類別。與傳統的方法不同，卷積神經網路不用先經過人工去作特徵提取的處理，就能夠在卷積層做到特徵提取的工作並連接到全連接層做分類，這樣的做法雖然方便，然而黑盒子的過程中花粉是如何透過特徵的不同而去分為不同類的我們卻無從得知，因此本文會透過對兩層卷積層所卷積出的特徵作視覺化，並以花粉粒常見的幾類特徵如顏色特徵、形狀特徵、幾何特徵及紋理特徵來描述不同類花粉之間特徵的差異。

3.6 資料擴充

不管是機器學習或是深度學習，有時在訓練資料難免會有過度擬合的情形產生，過度擬合指機器所學到的假設過度貼近於訓練資料，而使得訓練資料的準確率比測試資料的準確率高出許多，並且測試資料的誤差也比訓練資料來的高。產生過度擬合情況的原因可能是訓練資料的不足，或是訓練模型過於複雜，另外也可能是因為屬性值太多以及噪音數據所干擾導致。而解決過度擬合的方法不少，如利用特徵選擇演算法來降低特徵的數量，或利用正規化的方法，降低不重要的特徵的影響力，又或者使用 Dropout 技術。除了上述方法以外，資料擴充也是一個有效的辦法，對於深度學習來說當資料的數量上不足時，所得的準確率容易無法達到預期，透過資料擴充可以幫助我們彌補數量上的不足並且有效的避免過度擬合的情況。資料擴充常見的做法為將圖片經過旋轉、調整大小、調整比例或改變亮度等處理，雖然經這些處理後我們仍能夠辨識出與原本圖片是相似的，但對於機器來說輸入後會判定為不同的圖片。一般所使用的資料擴充為將圖片經平移、旋轉、縮放、調整比例、亮度...等方式處理後，使一張圖片產生出多張不同的圖片，以新的圖片與原本的資料集結合成為一個新的資料集作為輸入以卷積神經網路去做分類，通常能夠獲得更佳的準確率。

第四章 實驗與結果

4.1 實驗環境

本篇論文的實驗環境如下：

中央處理器:AMD Ryzen 5 1600 Six-Core Processor 3.20 GHz

記憶體:8GB

作業系統:Windows 10

硬碟:2TB

使用工具:Python, Tensorflow

4.2 參數設定

	LeNet5	Alexnet	Simple CNN
輸入層	50*50	227*227	50*50
卷積層 1	6(5,5)	96(11,11)	32(3,3)
最大池化層 1	(2,2)	(3,3)	(2,2)
卷積層 2	6(5,5)	256(5,5)	64(3,3)
最大池化層 2	(2,2)	(3,3)	(2,2)
卷積層 3		384(3,3)	
卷積層 4		384(3,3)	

卷積層 5		256(3,3)	
Dropout			0.5
全連接層 1	120	4096	128
Dropout		0.5	0.75
全連接層 2	84	4096	
Dropout		0.5	
全連接層 3		1000	
輸出層	23	23	23
其他參數設定	學習率:0.001 batch_size:128 Epoch:400 優化器:Adam	學習率:0.0001 batch_size:32 Epoch:50 優化器:SGD	學習率:0.001 batch_size:128 Epoch:400 優化器:Adam

表 4.1 參數設定

本研究的參數設定如表 4.1 所示，在參數設定上輸入層為配合 AlexNet 深層的架構將圖片尺寸設為 227*227，而 LeNet-5 架構與 Simple CNN 則以不使圖片因放大而失真為前提，將圖片尺寸設為 50*50。卷積神經網路架構的部分在 LeNet-5 以及 AlexNet 皆維持與原論文架構同樣的配置，Simple CNN 則是一個自定義的架構，搭配特徵可視化的實作，在卷積層的過濾器個數上設定為我們期望觀察到的特徵個數。而在其他參數設定上，學習率透過特徵可視化發現到當設置過大的學習率會導致特徵圖在經過激活函數層的處理後，出現太多值為零的情況，使得部分特徵無法順利顯示，因此將值調整的小一些，而由於本研究的實驗環境記憶體有限，因此將 AlexNet 的 batch_size 設定為 32，Epoch 則是在嘗試過後分別

設定，優化器部分 LeNET-5 與 Simple CNN 皆以使用 Adam 優化器的結果最好，AlexNet 避免收斂速度太快而使用 SGD。

4.3 實驗結果

4.3.1 第一階段實驗結果

本篇的實驗結果分為兩部分，第一階段為使用灰階圖以及彩色圖經過三種神經網路架構訓練後以交叉驗證的方式來評估所得的分類準確率以及混淆矩陣。從我們的表 4.2 與表 4.3 中可以見到以灰階圖去訓練的結果明顯比彩色圖來的差，因此說明了顏色特徵對花粉粒有一定的重要性，此外本篇也有以 RGB 三個通道分開去做訓練所得的結果與灰階圖的結果差不多，所以表示三個通道的顏色特徵皆重要，缺乏其中一個都會造成準確率的下降。而表格當中可見 AlexNet 的分類在缺少顏色特徵的情況下分類效果較 Simple CNN 好一些，但在有顏色特徵情況下的準確率比 Simple CNN 的準確率稍低一些，推測是由於 AlexNet 模型的複雜度高，並且在卷積層的部分圖片要以 227×227 的尺寸作為輸入，若以本篇使用的資料集來說，會導致一部分尺寸比 227×227 小的圖片在放大後失真而使得特徵變得模糊，進而影響到結果。另外可以看到使用彩色圖並以 Simple CNN 的效果最佳，達到了 81.55% 的準確率，說明了對於本研究使用的資料集來說，簡單的卷積神經網路模型所得分類結果比多層的 AlexNet 要來的好，準確率更穩定，訓練的時間也較短，而我們也特別將其混淆矩陣繪出查看分類結果，見圖 4.1。另外，在第一階段的實驗中，我們也嘗試對 Simple CNN 多加一層卷積層，其參數設定為 128 個卷積核，卷積核大小為 3×3 ，得到的準確率為 81.40%，與兩層卷積層的架構相近，因此在接下來的第二部分還是採用兩層卷積層架構的 Simple CNN。

	灰階圖以三種 CNN 架構訓練結果(%)		
10-Fold CV	LeNet-5	AlexNet	Simple CNN
1	52.17	70.65	68.47
2	67.39	77.17	71.73
3	67.39	78.26	83.69
4	66.30	74.99	71.73
5	63.04	61.95	77.17
6	53.62	76.81	63.76
7	52.17	76.81	66.66
8	63.76	73.91	81.15
9	62.31	76.81	71.01
10	55.07	75.36	68.11
平均	60.32±6.04	74.27±4.59	72.35±6.09

表 4.2 灰階圖以三種卷積神經網路搭配 10 折交叉驗證訓練結果

	彩色圖以三種 CNN 架構訓練結果(%)		
10-fold CV	LeNet-5	AlexNet	Simple CNN
1	63.04	80.43	78.26
2	67.39	83.69	80.43
3	67.39	64.13	80.43
4	64.13	82.60	85.86
5	67.39	66.30	84.78
6	65.21	78.26	84.05
7	57.97	84.05	78.26
8	69.56	82.60	84.05
9	55.07	81.15	84.05
10	68.11	72.46	75.36
平均	64.53±4.45	77.57±6.97	81.55±3.32

表 4.3 彩色圖以三種卷積神經網路搭配 10 折交叉驗證訓練結果

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
2	0	21	0	0	0	0	0	0	0	0	2	1	2	0	1	0	0	0	0	0	6	0	2
3	3	0	26	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	1	0	0	2	0
4	0	0	0	32	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0
5	0	0	0	0	35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	1	0	0	34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	32	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0
8	0	2	0	0	1	0	0	20	1	0	2	1	0	0	0	0	0	0	3	0	0	3	2
9	0	1	0	2	1	0	0	0	25	0	0	0	5	0	0	0	0	0	0	0	1	0	0
10	0	0	0	0	0	3	0	0	1	23	0	0	3	0	2	0	0	2	1	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	32	3	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	35	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	3	0	0	0	3	2	1	0	25	0	0	0	0	0	0	0	1	0	0
14	0	0	0	2	0	0	0	0	0	0	0	0	0	33	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32	0	0	1	0	1	0	1	0
16	2	0	0	0	0	0	0	0	0	1	0	0	0	0	0	23	0	7	0	1	0	0	1
17	0	0	2	0	1	0	0	1	1	0	0	0	0	0	0	0	30	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	2	0	0	0	0	1	4	0	28	0	0	0	0	0
19	0	0	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0	0	31	0	0	1	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	1	0	30	0	1	0
21	0	8	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	20	0	3
22	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	32	0
23	0	1	0	0	0	0	0	2	2	0	0	2	0	0	1	0	0	0	1	0	2	0	24

圖 4.1 彩色圖以 Simple CNN 訓練所得混淆矩陣

花粉類別 1:Anadenanthera 2:Arecaceae 3:arrabidaea 4:Cecropia 5:Chromolaena
6:Combretum 7:Croton 8:Dipteryx 9:Eucalyptus 10:Farama 11:Hyptis 12:Mabea
13:Matayba 14:Mimosa 15:Myrcia 16:Protium 17:Qualea 18:Schinus 19:Senegalia
20:Serjania 21:Syagrus 22:Tridax 23:Urochloa

圖 4.1 為繪出的混淆矩陣結果，可以觀察出 23 類的花粉粒在分類上錯誤分類到其他類的情形為何，其中我們可以看到在第 5 類以及第 12 的分類的結果都非常完美，沒有一張圖錯誤分類到其他類別，而第 1、4、6、7、11、14、15、17、19、20、22 類的分類情形皆不錯。在第 2、8、9、10、13、16、21 的分類情況則

較差，並且可以發現到第 2 類有 6 張圖錯誤分類到的 21 類，第 21 類有 8 張圖錯誤分類到第 2 類，說明這兩類圖的辨識容易混淆，如圖 4.2 與圖 4.3，兩類圖的顏色相近，Syagrus 此類的特點是花粉粒中間幾乎都會有一條裂縫，而在 Arecaceae 此類的資料集中也可看到有幾張花粉粒的中間也有裂縫並且形狀相近，使卷積神經網路在分類的結果相較於其他類別有一段差距。

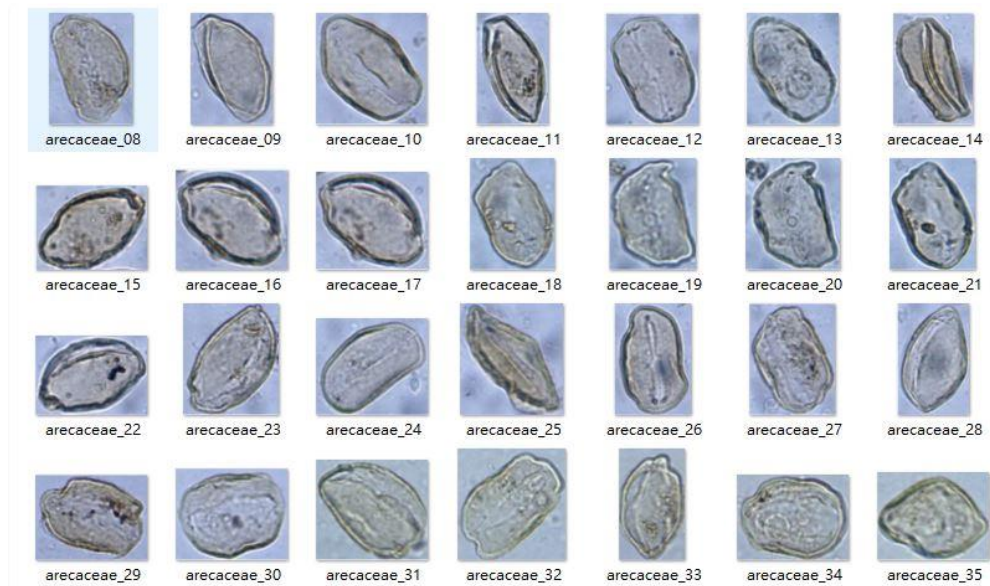


圖 4.2 Arecaceae 圖像資料集

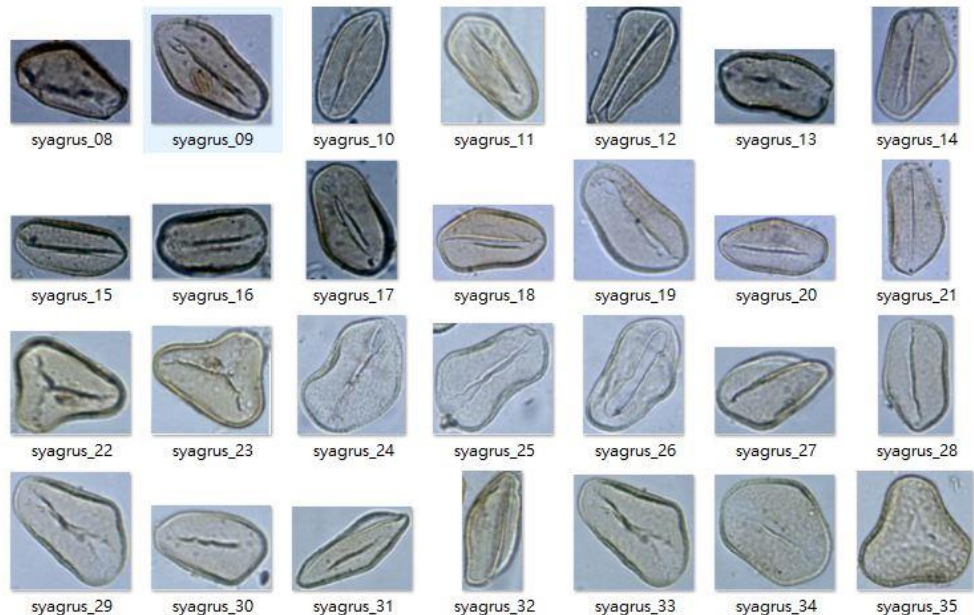


圖 4.3 Syagrus 圖像資料集

4.3.2 特徵可視化

本研究透過特徵可視化來驗證我們以卷積神經網路所得的分類結果，可視化的作法為將已訓練好的卷積神經網路，其每一卷積層卷積所得特徵圖經激活函數 ReLU 運算所得的輸出值作保存，再將圖像搭配輸出值後產生。首先我們可以從圖 4.4 及圖 4.5 觀察在不同學習率下以 Simple CNN 對於過濾器產生特徵的情況，在學習率偏大時可能導致圖 4.5 黑格較多，黑格代表在經過激活函數 ReLU 後大部分值為零，使得過濾器無法辨識到特徵，當認知的特徵較少時可能就會影響到分類的準確率，因此我們可以利用可視化來判斷是否在學習率的設定值上過大，進而來做調整。

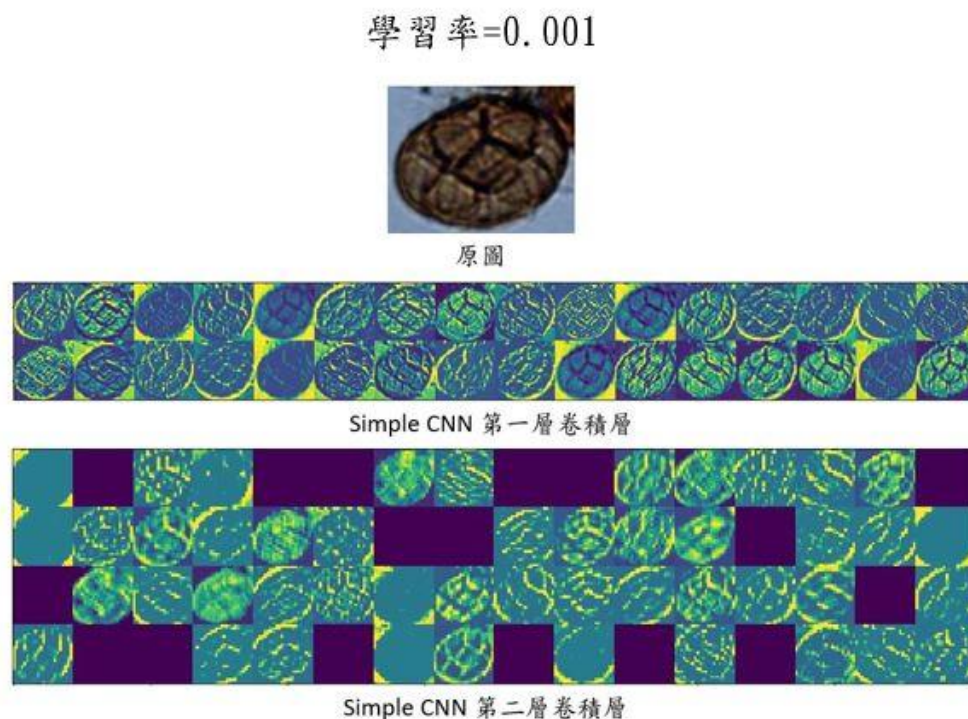
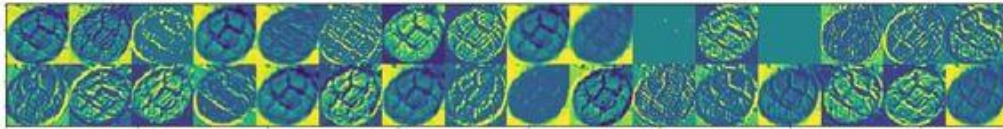


圖 4.4 *Anadenanthera* 以 Simple CNN，學習率=0.001 做特徵可視化

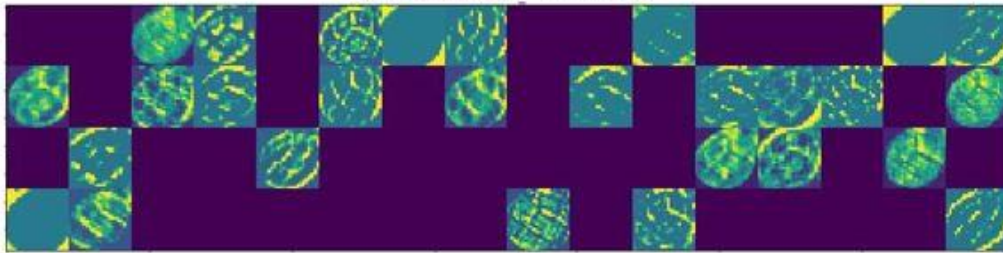
學習率=0.01



原圖



Simple CNN 第一層卷積層



Simple CNN 第二層卷積層

圖 4.5 Anadenanthera 以 Simple CNN，學習率=0.01 做特徵可視化

過去花粉粒的研究中從我們前面的表 2.1 可以看到大部分的研究所提取的特徵主要為顏色特徵、形狀特徵、紋理特徵，所以我們接著藉由可視化去觀察過濾器對於這幾類特徵的提取在同類以及不同類的花粉粒有哪些差異。顏色特徵我們可以透過觀察圖 4.6 以及圖 4.7 發現到在灰階圖的情況下，圖 4.6 過濾器所得到的特徵數明顯比圖 4.7 來的少，雖然在第一層卷積層所提取到的形狀特徵影響不大，不過對於第二層卷積層來說缺少了顏色特徵使得部分特徵無法與背景完全分離，特徵也較不明顯，皆影響了分類的結果。

相同類別花粉粒的比較我們可以透過圖 4.7、圖 4.8、圖 4.9 來觀察，從我們的三張圖當中可見 Anadenanthera 這一類的花粉粒在第一層卷積層所卷積出的特徵仍保留了我們原始圖中的所有信息，與我們的輸入相比起來沒有太大差距，還是完整的花粉粒樣貌，大多為比較直觀的特徵，像是傳統上以邊緣偵測器去提取出的輪廓和曲線的特徵的集合和形狀特徵，如紅框的部分為形狀特徵，其他有的則是分離背景、顯示色差，而在第二層卷積層我們可以看到更明顯的形狀特徵，

還有許多花粉粒外圍紋理的特徵，而黑框的部分是相對明顯的紋理特徵。從輪廓特徵部分我們可以看到與圖 4.7 相同類別的圖 4.8 的形狀相似，皆為橢圓狀，圖 4.9 則有些差異，偏像長條狀，不過可以看到在第二層左上角的兩個過濾器卷積出的特徵也稍微與橢圓狀相似。隨著層數的增加所提取出的特徵會越加抽象所提取的特徵也比較侷限在圖中的某個部分，能以人的視覺化角度去描述得特徵就變少一些，而第二層提取的特徵大多為紋理特徵，如圖 4.7 與圖 4.9 形狀上雖然比較不一樣，不過仍可在第二層看出兩者的紋理特徵類似，皆是網狀的圖案。除了以上對特徵的描述以外，本研究可以看到圖 4.7 及圖 4.9，對人來說可能不一定能以眼睛辨識出兩者為同類別的圖像，但在經過卷積後在第二層過濾器時有卷積出特徵的過濾器位置會一模一樣，並且卷積出的特徵數量也相近，代表其卷積層所卷積出的特徵圖的像素值差距不大，花粉粒的顏色相近。

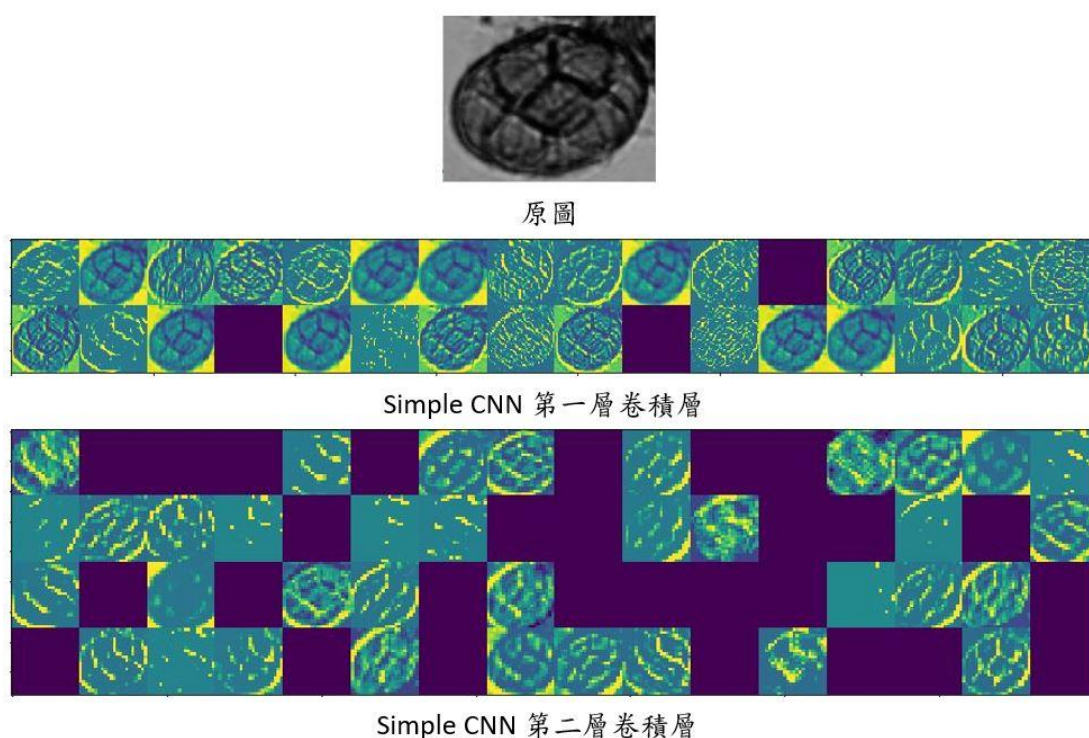
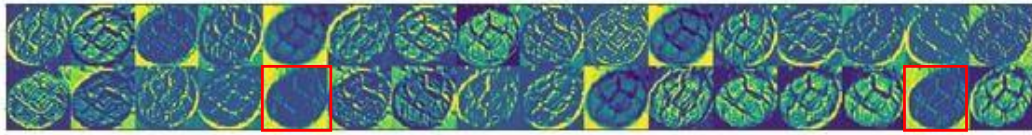


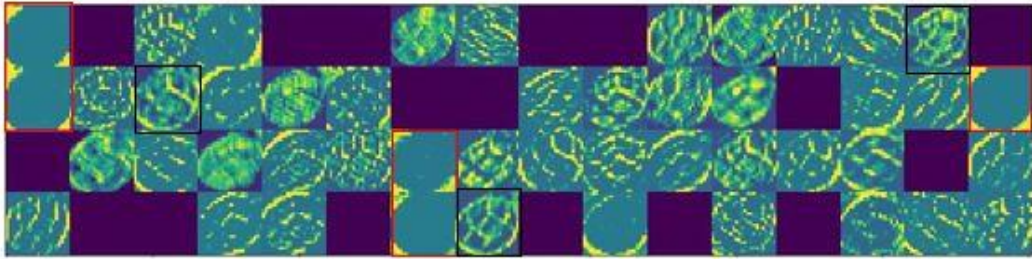
圖 4.6 Anadenanthera 以 Simple CNN 做特徵可視化(灰階圖)



原圖



Simple CNN 第一層卷積層

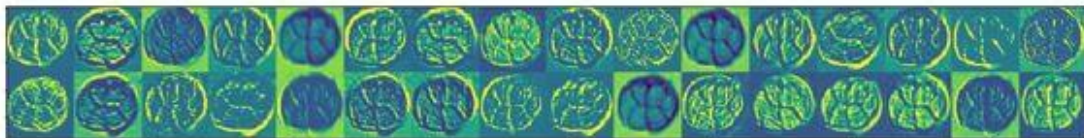


Simple CNN 第二層卷積層

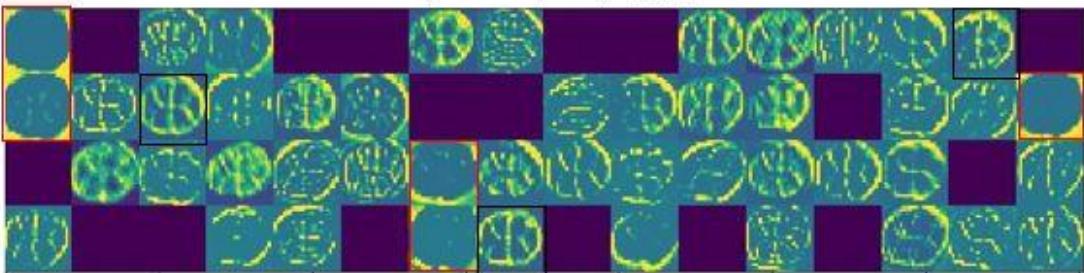
圖 4.7 *Anadenanthera* 以 Simple CNN 做特徵可視化(相同類別比較)



原圖



Simple CNN 第一層卷積層



Simple CNN 第二層卷積層

圖 4.8 *Anadenanthera* 以 Simple CNN 做特徵可視化(相同類別比較)

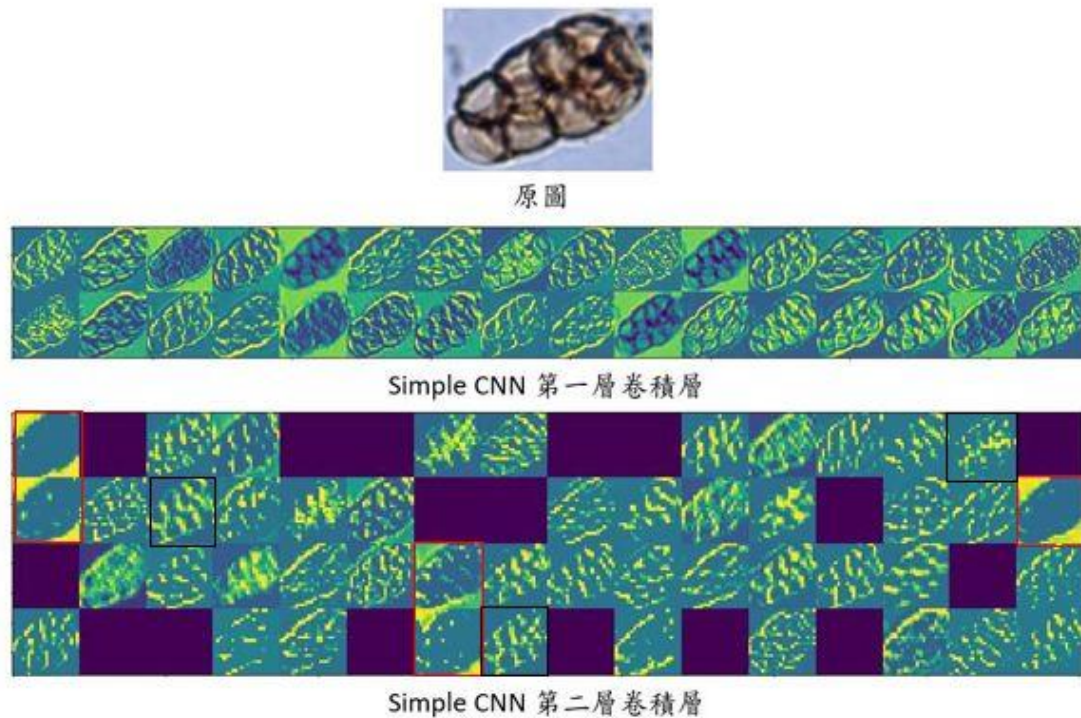


圖 4.9 Anadenanthera 以 Simple CNN 做特徵可視化(相同類別比較)

而在不同類別的花粉粒的比較上，我們可以透過圖 4.10 與圖 4.11 相似的圖來觀察，圖 4.10 與圖 4.11 分別是 Arecaceae 與 Syagrus 此兩類去做特徵可視化的結果，從我們前面的圖 4.1 的混淆矩陣中觀察到這兩類分類效果是較差的，其中 Arecaceae 有六張圖錯誤分類到 Syagrus，Syagrus 則有八張圖錯誤分類到 Arecaceae，以外觀上來看，Arecaceae 與 Syagrus 這兩類的顏色相近，皆為灰色且有些透明，形狀上也相似，而 Syagrus 此類在紋理特徵的部分中間裂縫較為明顯，另外在第二層卷積層的可視化結果卷積出的特徵也只有三個過濾器(以黑框表示)的差別，代表此兩類花粉粒顏色相近。我們可以從圖 4.10 及圖 4.11 中觀察到由於此兩類圖的花粉粒顏色與環境顏色的色差較小又有些透明，因此其形狀特徵(以紅框表示)與圖 4.7 相較起來提取的稍微模糊些，增加了卷積神經網路在分類上的難度。

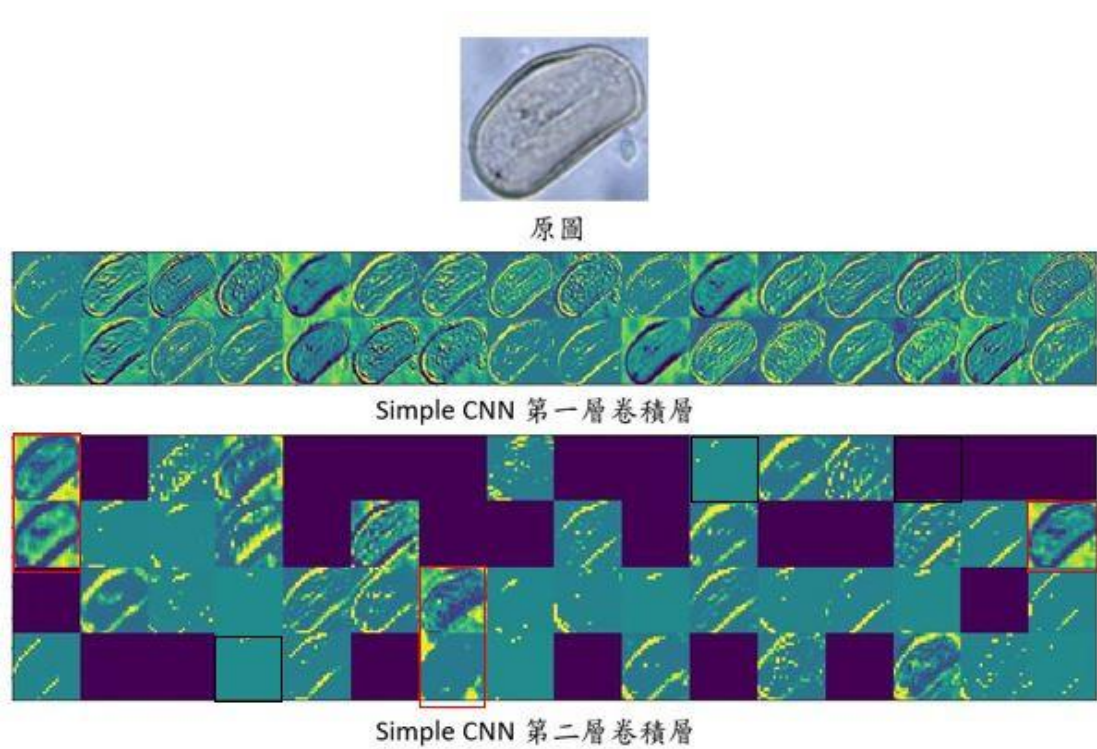


圖 4.10 Arcaceae 以 Simple CNN 做特徵可視化(不同類別比較)

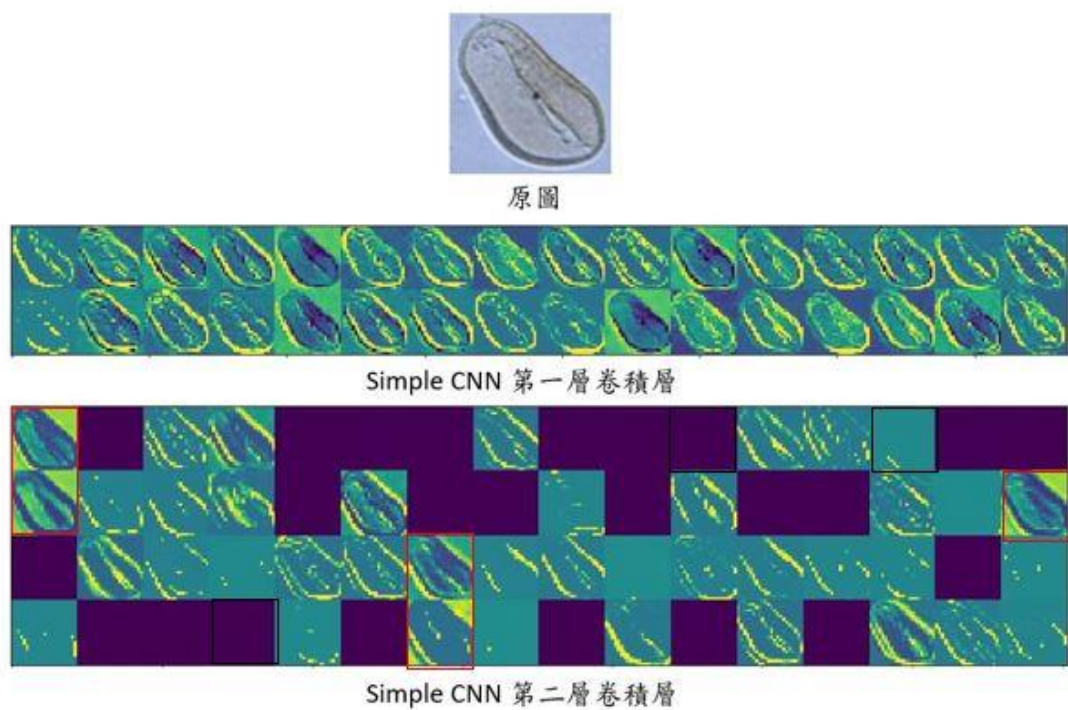


圖 4.11 Syagrus 以 Simple CNN 做特徵可視化(不同類別比較)

圖 4.12 與圖 13 也是分類結果較差的兩類，以此兩類相似的圖來觀察，其中

Eucalipto 有五張圖錯誤分類到 Matayba，Matayba 則有三張圖錯誤分類到 Eucalipto，此兩類在形狀上皆為三角形，顏色為灰色且有些透明，與圖 4.10 及圖 4.11 兩類花粉粒的顏色相近，所以我們可以看到卷積出特徵的過濾器位置此四類無太大差異。輪廓特徵的部分此兩類也有與前一部分相似的情況，色差較小且偏透明使得輪廓特徵的掌握上有些模糊，不過在紋理特徵部分可從圖中見到兩類是有差異的，從色差的部分可以看出 Eucalipto 中心有三角狀的特徵而 Matayba 則是蝴蝶形狀的特徵，因此在分類結果上相較於 Arecaceae 和 Syagrus 要來的好些。

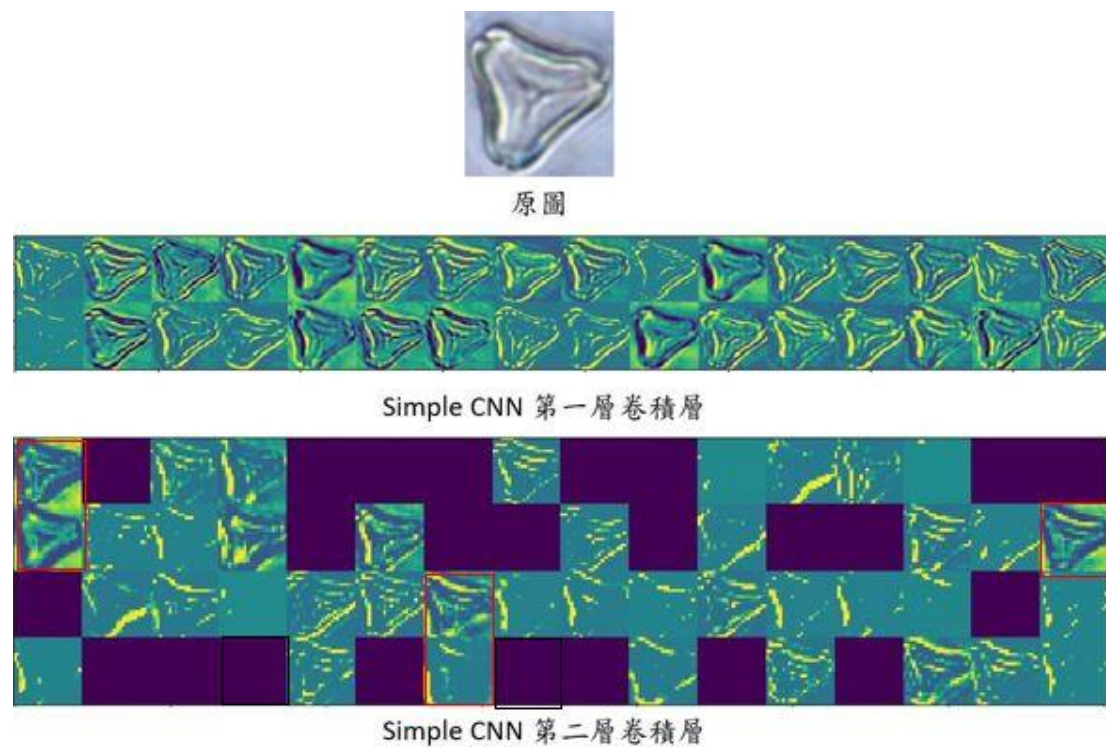


圖 4.12 Eucalipto 以 Simple CNN 做特徵可視化(不同類別比較)

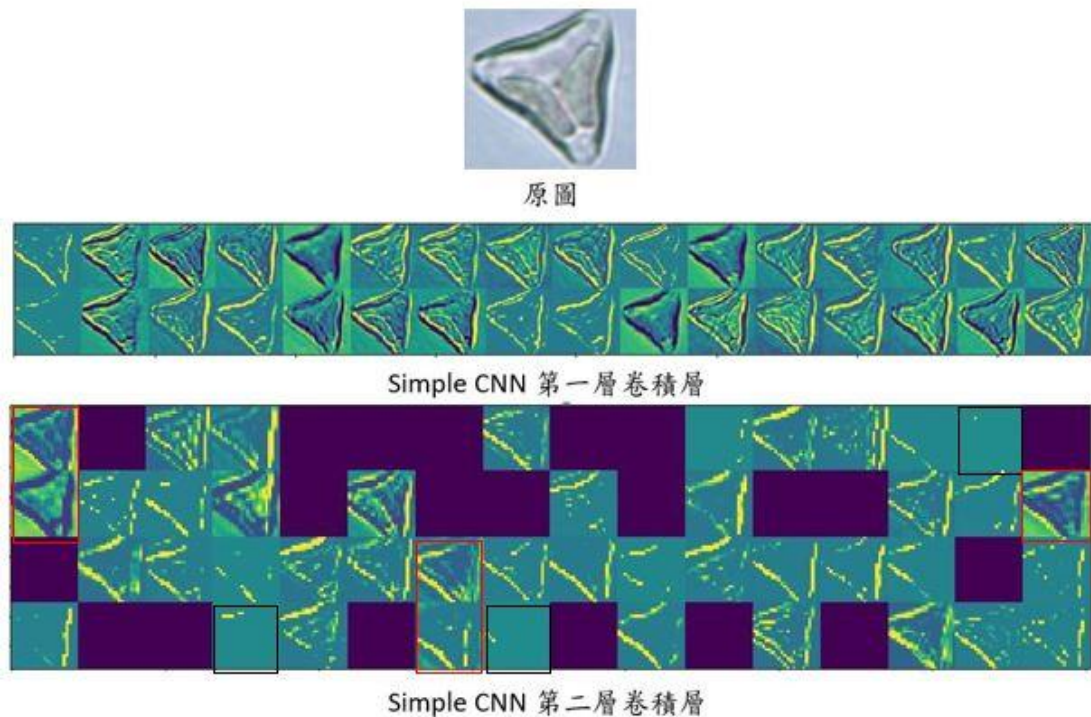
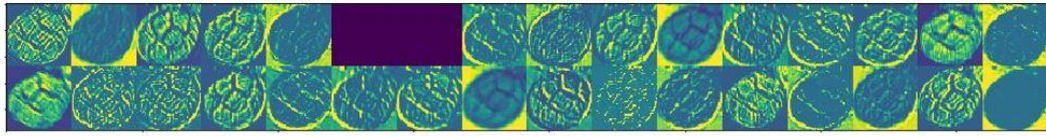


圖 4.13 Matayba 以 Simple CNN 做特徵可視化(不同類別比較)

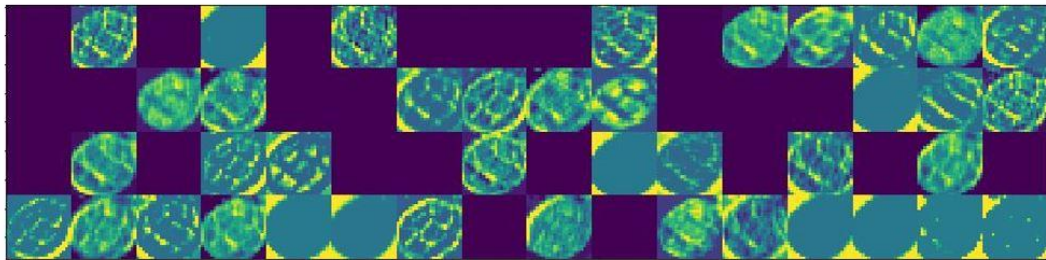
藉由圖 4.14 我們也可以觀察前面提到的再將 Simple CNN 卷積層的架構增加一層 128 個過濾器的卷積層後，經過可視化的結果。從圖中我們可以見到在第一層就有兩格過濾器沒有認知到特徵，第二層有認知到特徵的過濾器相較起兩層卷積層架構也較少，不過可以觀察到提取的輪廓特徵及紋理特徵皆非常明顯。而主要的第三層特徵提取的特徵有限，也較難以人類視覺所見的角度描述，皆為偏向圖像類別的信息。



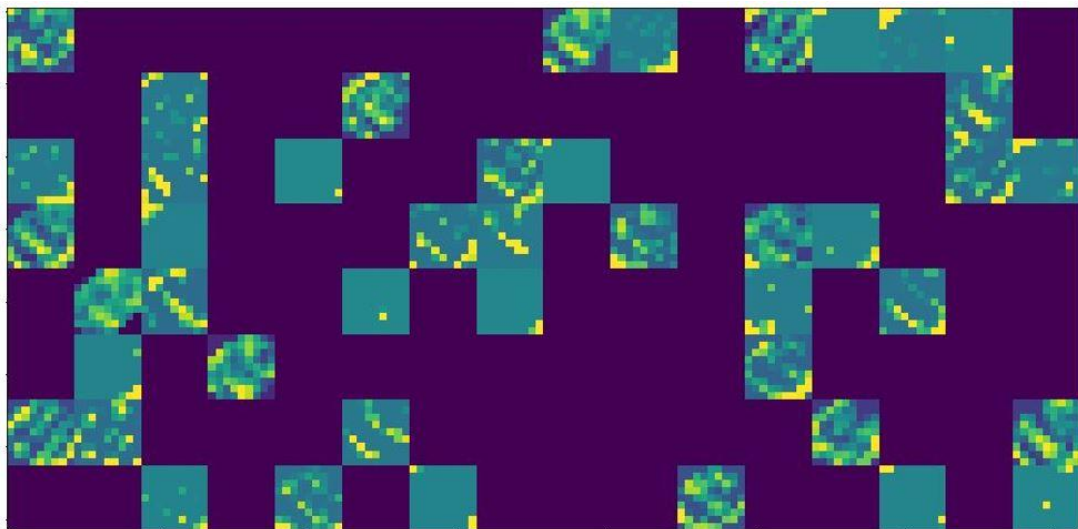
原圖



Simple CNN 第一層卷積層



Simple CNN 第二層卷積層



Simple CNN 第三層卷積層

圖 4.14 Anadenanthera 以 Simple CNN 做特徵可視化(三層卷積層)

除了上述的比較之外，我們也可以藉由圖 4.15 觀察 AlexNet 架構所卷積出的特徵，可以看到在第一層的部分所卷積出的邊緣集合的特徵與圖 4.7 的 Simple CNN 兩層架構沒有太大差別，而第二層的部分同樣皆偏向紋理的特徵，我們發現圖 4.15 經過過濾器得到特徵的比例更高，然而以表 4.3 來看，AlexNet 所得準確率卻沒有 Simple CNN 來的高，可以說明這些特徵對於花粉粒分類的重要性可

能相對於第一層的邊緣特徵的集合來的低。圖 4.16 觀察到 AlexNet 最後一層卷積層所得特徵，與圖 4.14 的第三層一樣皆是圖像類別的信息，以人眼的觀察角度較無法去解釋。

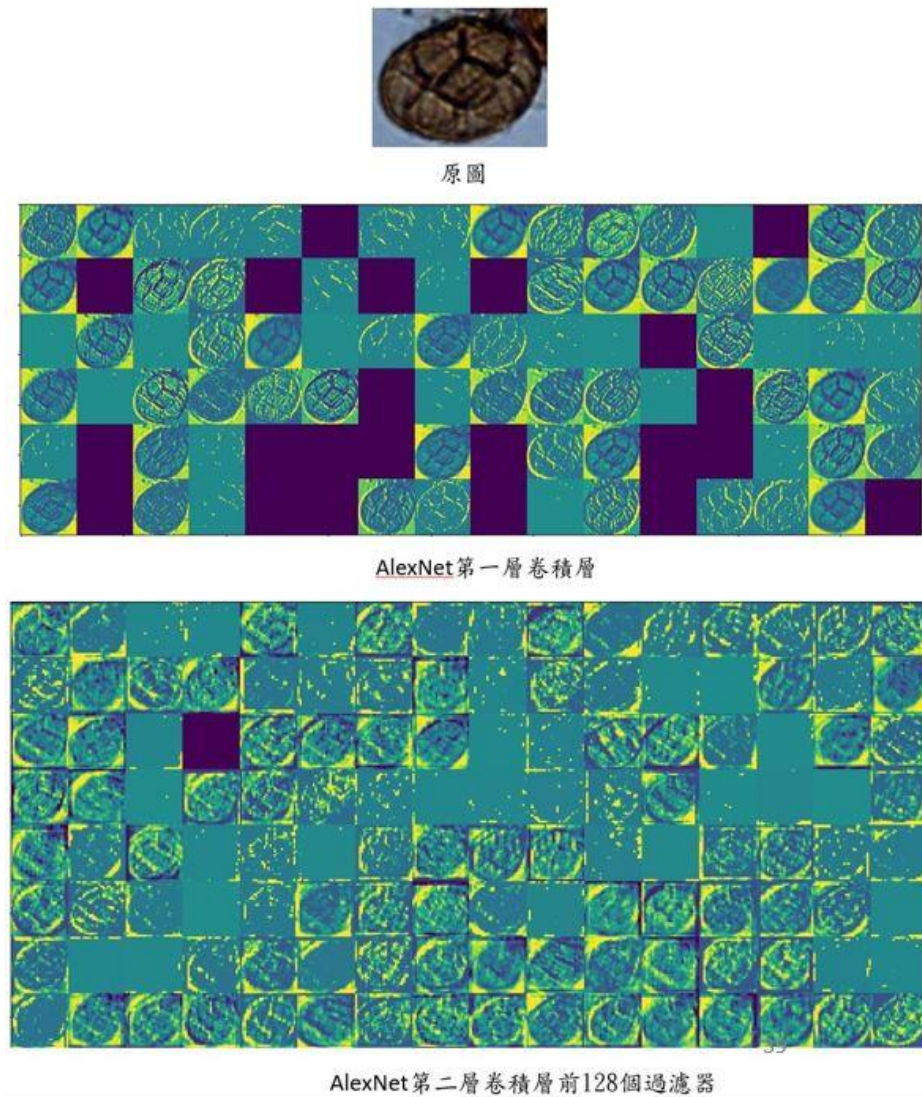
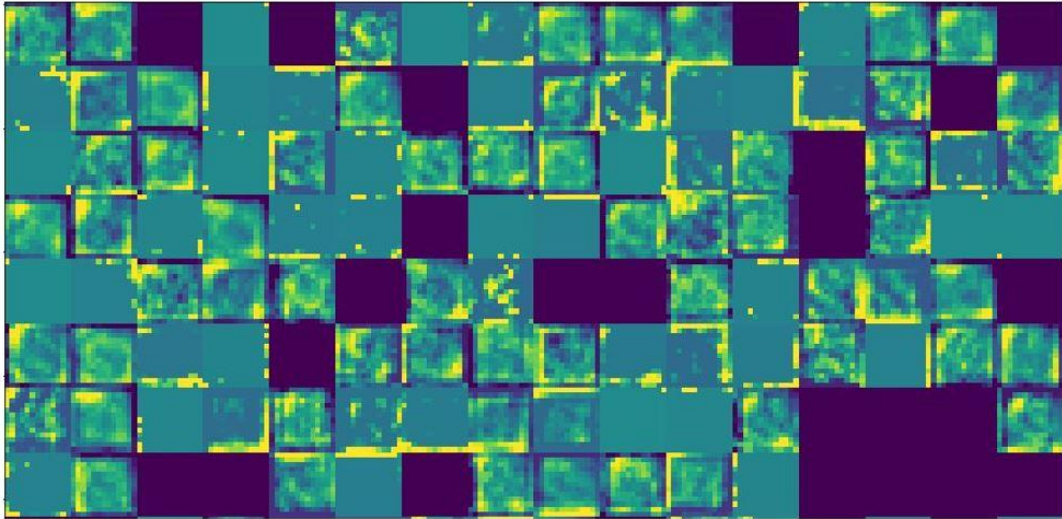


圖 4.15 Anadenanthera 以 AlexNet 前兩層卷積層做特徵可視化



原圖



AlexNet 第五層卷積層前128個過濾器

圖 4.16 Anadenanthera 以 AlexNet 第五層卷積層做特徵可視化

4.3.3 資料擴充

深度學習在資料上不足的情況下所得的準確率可能就不符預期，所以本研究使用了一般的資料擴充方式將我們圖片經過隨機旋轉 40 度、對長與寬做 10%幅度的平移、隨機縮小或放大 10%的幅度、逆時針裁切 10%的角度以及隨機水平翻轉等方式處理，如圖 4.17 所示，訓練資料集的圖片由原本 575 張擴充到了 6256 張，約比原本多出了約 10 倍的量，另外為了觀察利用交叉驗證所得到的準確率，所以將原本所有的 805 張圖片擴充到 8714 張的圖片，而在數量上原本所設定是將每一張圖皆擴充 10 倍的量，不過有一些圖片在以擴充方法處理時，花粉粒會超出原本圖片尺寸的框架，在辨識上會有困難，所以將這些圖片捨棄，因此每一張圖在擴充的量介於 6 到 10 張圖之間。將這些圖片以卷積神經網路進行訓練後，過度擬合的情況可以得到一些改善，同時準確率上也有提升的效果。

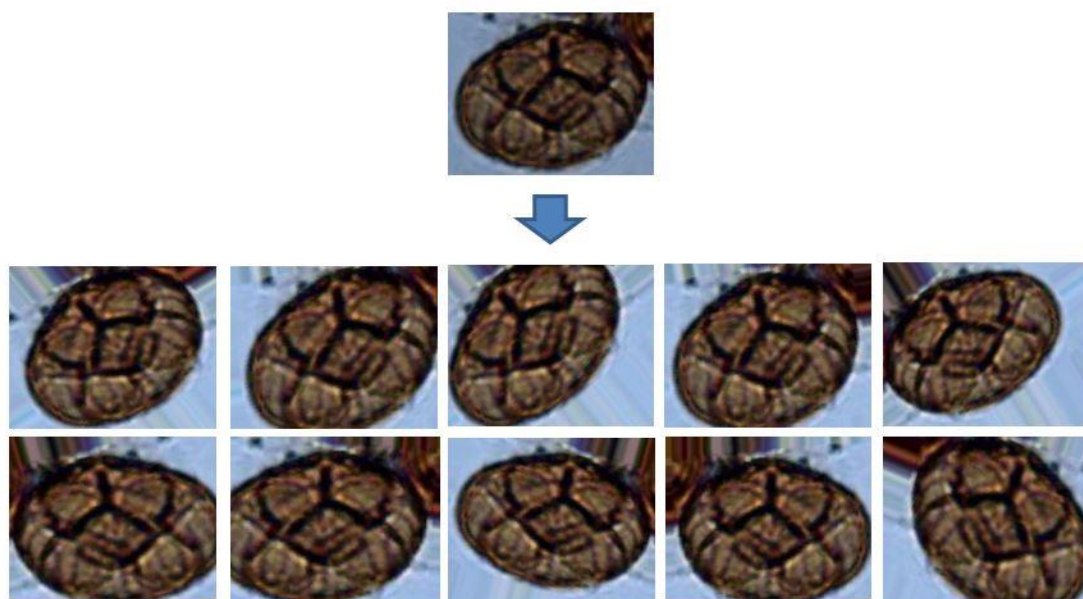


圖 4.17 資料擴充範例

4.3.4 第二階段實驗結果

在我們第一階段的實驗結果以 RGB 圖片使用 Simple CNN 搭配交叉驗證後得到 81.55%的準確率為最佳，因此第二階段已採用 Simple CNN 架構為主，由於

認為準確率還有提升的空間，所以在第二階段的實驗結果透過資料擴充的方式，將圖片擴充後再以 Simple CNN 進行訓練，觀察是否能再提高準確率。在第二段資料擴充的操作我們分為以下兩種：

1.將 805 張的圖片全部進行擴充，得到 8714 張，與原本以 805 張搭配 10 折交叉驗證的結果作比較。所得實驗結果見表 4.4

資料以 Simple CNN 搭配 10 折交叉驗證訓練		
10 折交叉驗證	擴張前 805 張(%)	擴張後 8714 張(%)
1	78.26	95.56
2	80.43	94.98
3	80.43	95.65
4	85.86	96.21
5	84.78	96.56
6	84.05	95.18
7	78.26	96.44
8	84.05	96.20
9	84.05	96.17
10	75.36	96.51
平均	81.55 ± 3.32	95.95 ± 0.54

表 4.4 資料以交叉驗證分割並以資料擴充前後作比較

2.手動將總共 805 張的图片分為 575 張訓練資料集以及 230 張的測試資料集去訓練得到一個結果後，將訓練資料集進行擴充，擴充到 6256 張，測試資料集維持原本的 230 張，再去訓練得到一個資料擴充後的結果，最後將兩結果進行比較來看資料擴充的效益，其中擴充前的參數上有稍作調整，學習率訂為 0.0001，Epoch 訂為 500 次。所得實驗結果見表 4.5

資料分訓練及測試以 Simple CNN 訓練	
擴充前	擴充後
訓練資料集 575 張/ 測試資料集 230 張	訓練資料集 6256 張/ 測試資料集 230 張
訓練準確率:82.09% 測試準確率:78.69%	訓練準確率:91.38% 測試準確率:87.82%

表 4.5 使用 Holdout 驗證並以資料擴充前後作比較

從表 4.4 與表 4.5 的實驗結果來看，透過資料擴充方法所帶來的效益很高，分訓練及測試的部分測試準確率提高了大約 9%，而 10 折交叉驗證則是提高了有 14.4%之多。在第二階段我們也將混淆矩陣繪出，如圖 4.18，與第一部分做比較，從圖中可以觀察到原本分類情形較差的第 2、8、16、21 類依比例來說分類的情形都有獲得不小的改善，如第 2 類原本 35 張有 14 張錯誤分類到其他類別，錯誤率為 40%，而在圖 23 中總共 379 張圖有 25 張錯誤分類到其他類別，錯誤率約為 6.6%，獲得不小的改善，說明了資料擴充的可行性。

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	371	0	0	0	0	0	0	1	0	1	0	0	0	0	0	5	0	0	0	0	0	0	0
2	0	354	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	18	0	4
3	1	0	363	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	1	0	0	6	0
4	0	0	0	365	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0
5	0	0	0	0	375	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
6	0	0	1	0	0	375	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	371	2	0	0	0	4	1	0	0	0	0	0	0	0	0	0	0
8	0	2	7	0	0	0	0	360	6	0	0	0	0	0	0	0	0	0	1	0	0	3	0
9	0	0	0	0	0	0	0	2	371	0	0	0	7	0	0	0	0	0	0	0	0	0	1
10	0	0	0	1	0	1	0	3	0	357	0	0	10	0	4	0	0	2	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	373	4	0	0	0	0	0	0	1	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	6	374	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	4	0	0	0	9	5	1	0	359	1	0	0	0	0	0	0	1	0	0
14	0	0	0	1	0	0	0	0	1	0	0	0	0	371	0	0	0	0	0	0	0	0	0
15	0	0	0	4	0	0	0	0	0	1	0	0	1	0	359	0	0	1	0	12	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	343	0	34	0	0	0	0	0
17	0	0	2	0	1	0	0	3	1	0	0	0	6	0	0	0	367	0	1	0	0	1	0
18	0	0	0	0	0	0	0	0	0	3	0	0	0	0	10	19	0	346	0	1	0	0	0
19	0	0	0	0	0	0	0	1	1	0	2	0	0	0	0	0	0	0	378	0	0	0	0
20	0	0	1	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	372	0	0	0
21	0	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	349	0	7
22	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	370	0
23	0	7	0	0	0	0	0	0	6	0	0	0	1	0	13	0	0	0	2	0	13	0	338

圖 4.18 資料擴充後的混淆矩陣

藉由前面第二階段的實驗證明了對於花粉粒資料集使用資料擴充方法的可行性，然而此卷積神經網路的架構是否泛用於其他花粉粒資料集仍未可知，因此本研究又額外從 The Global Pollen Project 網站當中整理了 9 類的花粉粒資料，如圖 4.19，將此 9 類與原本的 23 類資料集合併成一個 32 類，共 1173 張圖片的花粉粒資料集，觀察在資料集更大、種類更多的情況下，是否還能夠獲得理想的分類結果。實驗的部分採用與前面同樣的方式，將圖片經前處理後，以 Simple CNN 搭配 10 折交叉驗證分類，接著再以資料擴充的方式將圖片擴充到 12759 張圖片，同樣以 Simple CNN 搭配 10 折交叉驗證分類，兩階段所得到的準確率見表 4.6，從表 4.6 可看到所獲得的平均準確率與 23 類資料的沒有太大差距，證明 Simple CNN 此架構在對於花粉粒的分類有一定的泛化性。

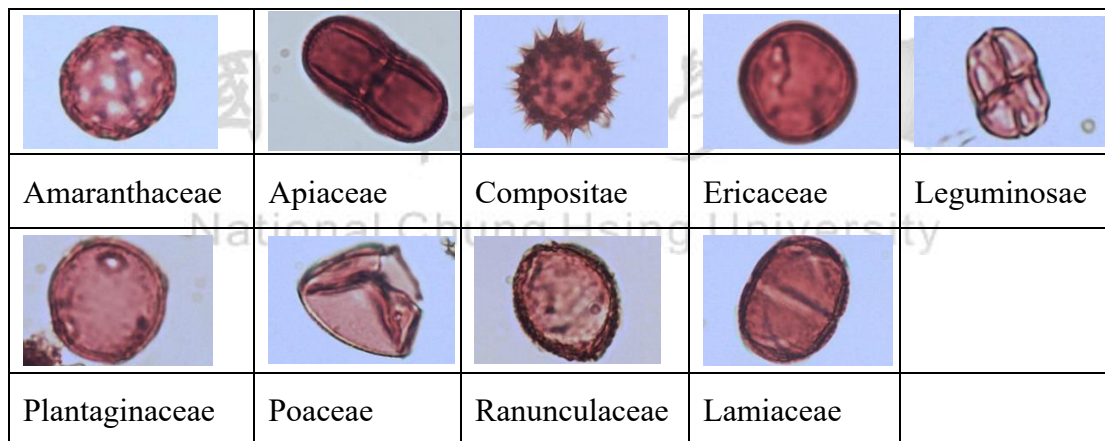


圖 4.19 從 The Global Pollen Project 所收集而來的 9 個類別的花粉粒圖像資料集

Simple CNN with 10-cross validation		
10-fold cross validation	1179 images(%)	12759 images(%)
1	73.13	95.80
2	78.78	95.09
3	80.30	95.16
4	84.61	95.93
5	74.61	96.40
6	78.30	96.63
7	82.85	94.74
8	81.73	94.81
9	71.84	96.51
10	76.69	95.63
Average	78.28±4.01	95.67±0.67

表 4.6 資料以交叉驗證分割並以資料擴充前後作比較(32 類)

	資料集	準確率
Gonçalves et al. [1]	805張圖片共23類	64%
以Simple CNN分類	805張圖片共23類	81.55%

表 4.7 與使用一般機器學習方法比較

	資料集	準確率
A. Daood et al. [2]	1000 LM 圖片擴充至14000 張圖片, 30 類	89.95%
A. Daood et al. [2]	1000 LM圖片擴充至15000 張圖片, 30 類	93.99%
本篇方法(Simple CNN)	805 LM 圖片擴充至8714 張圖片, 23 類	95.95%
本篇方法(Simple CNN)	1179 LM 圖片擴充至12759張圖片, 32 類	95.67%

表 4.8 與使用深度學習的論文比較

表 4.7 與表 4.8 為本研究與過去研究的比較，在使用同資料集的情況下，本研究相較於使用機器學習方法來分類的[1]在準確率上有明顯的改善，另外比較同樣使用深度學習架構的[2]，雖然所使用的資料集以及卷積神經網路的架構不同，不過相比於 30 類的資料集，本研究對於 32 類的資料集可得到 95.67%的準確率。

第五章 結論與建議

5.1 結論

本研究使用了一個公開的花粉粒資料集，包含了 805 張圖片共 23 類的圖片，其中種類在過往研究中為第二多的，代表了其分類的難度。

第一階段使用了灰階與彩色圖以三種不同的卷積神經網路架構包括 LeNet-5、AlexNet 及 Simple CNN 針對 23 種類別的花粉粒分類，分類的結果顯示出彩色圖以 Simple CNN 搭配 10 折交叉驗證所得的分類結果最好，達到了 81.55% 的準確率，說明顏色特徵對於花粉粒的重要性，而我們也畫出混淆矩陣來觀察每一類花粉粒的分類情形，其中對於分類效果較差的種類以花粉粒的表面特徵去推測其較差的原因。

第二階段為了讓第一階段的結果更進一步提升，採用了資料擴充的技術，將圖片擴充至大約 10 倍的張數再以 Simple CNN 搭配 10 折交叉驗證去訓練，藉由此方法我們的準確率從原本的 81.55% 提升到 95.95%，而原本分類情況較差的類別也獲得了一些改善，本研究也進一步從 The Global Pollen Project 網站上收集了 9 個類別的花粉粒與原本 23 類的資料集合併為一個 32 類的資料集，使用資料擴充方法後以 Simple CNN 搭配 10 折交叉驗證同樣能得到高分類準確率，證明了 Simple CNN 對於花粉粒分類有一定的泛化性。除此之外，本研究也透過特徵可視化來觀察到同類別的花粉粒相似的特徵，如：顏色相近、形狀相似、紋理的方向性類似...等，而不同類別的花粉粒以人眼或許難以分辨，但機器可以藉由所卷積出的特徵數量以及過濾器的位置的差異，去判定其為不同類別，雖然有些類別仍存在較多錯誤分類的情形，不過這些情形在經過資料擴充後錯誤分類的比例也得到了很大的改善。

5.2 未來展望

未來對於花粉粒的研究最主要還是資料集的收集，如果有足夠大的資料集就不需額外做資料擴充，並且可以對不同資料集使用相同的研究方法做比較。而在卷積神經網路的泛化性也是一個重點，如何視圖片尺寸大小選擇合適的卷積神經網路架構，以及簡單架構的神經網路架構能否達到與深層架構同樣的分類效果。而資料擴充大大提升了準確率，不過當擴充更多張時是否能夠再提升，或者擴張成幾倍時所得準確率最高。除此之外，在特徵可視化的部分所提取的特徵個數對於分類所帶來的影響有哪些關聯性也可以去研究並探討。



參考文獻

- [1] A. B. Gonçalves, J. S. Souza, G. G. da Silva, M. P. Cereda, A. Pott, M. H. Naka, *et al.*, "Feature Extraction and Machine Learning for the Classification of Brazilian Savannah Pollen Grains," *PloS one*, vol. 11, p. e0157044, 2016.
- [2] A. Daoood, E. Ribeiro, and M. Bush, "Pollen Grain Recognition Using Deep Learning," in *International Symposium on Visual Computing*, 2016, pp. 321-330.
- [3] P. Li and J. R. Flenley, "Pollen texture identification using neural networks," *Grana*, vol. 38, pp. 59-64, 1999.
- [4] E. Stillman and J. R. Flenley, "The needs and prospects for automation in palynology," *Quaternary Science Reviews*, vol. 15, pp. 1-5, 1996.
- [5] I. France, A. Duller, G. Duller, and H. Lamb, "A new approach to automated pollen analysis," *Quaternary Science Reviews*, vol. 19, pp. 537-546, 2000.
- [6] M. Rodriguez-Damian, E. Cernadas, A. Formella, and R. Sa-Otero, "Pollen classification using brightness-based and shape-based descriptors," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, 2004, pp. 212-215.
- [7] A. Corbi, C. Cortes, J. Bousquet, A. Basomba, A. Cistero, J. Garcia-Selles, *et al.*, "Allergenic cross-reactivity among pollens of Urticaceae," *International Archives of Allergy and Immunology*, vol. 77, pp. 377-383, 1985.
- [8] S. H. Landsmeer, E. A. Hendriks, L. A. De Weger, J. H. Reiber, and B. C. Stoel, "Detection of pollen grains in multifocal optical microscopy images of air samples," *Microscopy research and technique*, vol. 72, pp. 424-430, 2009.
- [9] W. Treloar, G. Taylor, and J. Flenley, "Towards automation of palynology 1: analysis of pollen shape and ornamentation using simple geometric measures, derived from scanning electron microscope images," *Journal of quaternary science*, vol. 19, pp. 745-754, 2004.
- [10] N. M. García, V. A. E. Chaves, J. C. Briceño, and C. M. Travieso, "Pollen grains contour analysis on verification approach," in *International Conference on Hybrid Artificial Intelligence Systems*, 2012, pp. 521-532.
- [11] M. del Pozo-Baños, J. R. Ticay-Rivas, J. B. Alonso, and C. M. Travieso, "Features extraction techniques for pollen grain classification," *Neurocomputing*, vol. 150, pp. 377-391, 2015.
- [12] D. S. da Silva, L. N. B. Quinta, A. B. Gonçalves, H. Pistori, and M. R. Borth, "Application of wavelet transform in the classification of pollen grains," *African Journal of Agricultural Research*, vol. 9, pp. 908-913, 2014.

- [13] M. Langford, G. Taylor, and J. Flenley, "Computerized identification of pollen grains by texture analysis," *Review of Palaeobotany and Palynology*, vol. 64, pp. 197-203, 1990.
- [14] M. del Pozo-Baños, J. R. Ticay-Rivas, J. Cabrera-Falcón, J. Arroyo, C. M. Travieso-González, L. Sánchez-Chavez, *et al.*, "Image processing for pollen classification," in *Biodiversity Enrichment in a Diverse World*, ed: Intech, 2012.
- [15] Y. Zhang, D. Fountain, R. Hodgson, J. Flenley, and S. Gunetileke, "Towards automation of palynology 3: pollen pattern recognition using Gabor transforms and digital moments," *Journal of quaternary science*, vol. 19, pp. 763-768, 2004.
- [16] S. W. Punyasena, D. K. Tcheng, C. Wesseln, and P. G. Mueller, "Classifying black and white spruce pollen using layered machine learning," *New Phytologist*, vol. 196, pp. 937-944, 2012.
- [17] M. Rodriguez-Damian, E. Cernadas, A. Formella, M. Fernández-Delgado, and P. De Sa-Otero, "Automatic detection and classification of grains of pollen based on shape and texture," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 36, pp. 531-542, 2006.
- [18] O. Ronneberger, H. Burkhardt, and E. Schultz, "General-purpose object recognition in 3D volume data sets using gray-scale invariants-classification of airborne pollen-grains recorded with a confocal laser scanning microscope," in *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, 2002, pp. 290-295.
- [19] M. Reisert and H. Burkhardt, "Invariant features for 3d-data based on group integration using directional information and spherical harmonic expansion," in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, 2006, pp. 206-209.
- [20] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, pp. 2278-2324, 1998.
- [21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.
- [22] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, pp. 1929-1958, 2014.
- [23] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [24] D. Klep, "Data augmentation of a handwritten character dataset for a

- Convolutional Neural Network and integration into a Bayesian Linear Framework," 2016.
- [25] L. Perez and J. Wang, "The Effectiveness of Data Augmentation in Image Classification using Deep Learning," *arXiv preprint arXiv:1712.04621*, 2017.
 - [26] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *European conference on computer vision*, 2014, pp. 818-833.
 - [27] L. M. Zintgraf, T. S. Cohen, T. Adel, and M. Welling, "Visualizing deep neural network decisions: Prediction difference analysis," *arXiv preprint arXiv:1702.04595*, 2017.

