

## 蘭花辨識器-利用深度學習模型對台灣野生蘭花進行物種分類

## An Orchid Identifier: Using Deep Learning to Identify Wild Orchids of Taiwan

## 一、前言

## 1. 題目動機與目的

雖然自己於大學的主修為機械工程，但因著自己從小對自然生態的熱愛，而雙主修了森林系，因此自己目前的專業領域算是涵蓋機械工程、森林環境及生物等不同方面。

一開始在構思本次作業的題目時，由於自己在機械工程領域求學時比較少接觸與大數據分析的相關經驗，比較著重在如機構分析、機器人控制等偏硬體的部分，目前想到與機器學習或深度學習相關的部分大概是利用電腦視覺結合機器人控制，來對周圍障礙物做偵測，讓機器人能做出相對應的動作以閃避障礙物。然而此題目需應用到許多艱深的技術，包括機器人機構設計、電路板控制等，此外若要應用深度學習技術，亦須事先收集足夠的數據(如：記錄不同障礙物情境的影片等)，以上種種均難以在短時間內完成，故決定先暫緩此項題目的製作。

而在森林領域專業部分，由於在系上有修習統計相關課程，加上不論是在課程上或是出野外調查也常常會需要做數據分析，因此在許多面向都能運用機器學習的技術來幫助完成相關的分析作業。因自己對於植物有些研究，平時除了翻閱書籍、上網查詢資料外，偶爾會出野外拍攝，也為此至今累積了不少植物照片在手。然而在後續作業時可能會遇到一些問題或困難：首先若是要對在手的大量植物照片做物種的分類與整理，即需要花費大量時間；再來若是有拍攝到不會辨認的植物時，更需翻書、上網或是詢問同學朋友，此過程中又需要相當多的時間與精力，且有時無法立即得到正確解答。因此若是能開發出一套自動辨識植物種類的軟體，讓使用者不論是在野外調查還是在家中做資料整理時，只要輸入一張照片至模型，模型即能輸出該照片中的植物種類，就能節省人工辨識所花費的時間，而若是模型本身的正確率夠高，更能進一步避免一些人工辨識上的錯誤與疏失。

目前坊間雖然已有類似的生物辨識軟體問世，例如：形色、inaturalist等等，使用上也相當方便，但自己實際上在使用時，卻常發現若是照片特徵不夠清楚、或是幾種型態較為類似的種類，往往有判斷失準的情形出現，此外，為了提升軟體被使用的機率與普及度，這些軟體很多是針對地球上所有生物作辨識(如：inaturalist)，因此在對於某些較細的分類群，例如在某幾類特定的植物上，就會有蠻高的出錯率。因此我想如何更精準的

提取這些相似物種的特徵，將是改善此類辨識器的一大課題。

台灣有 4000 多種原生植物，因此若要做出一套準確率夠高的植物辨識器，需要 4000 多個 label，每種植物更是需要大量的照片數據才能夠完成。礙於期末時間有限，加上自己最近主要是對於台灣的野生蘭花有研究，市面上亦沒有專門辨識蘭花的軟體或分類器出現，也因此在這次的期末作業當中，決定將範圍濃縮至以台灣野生蘭種分類為主，透過深度學習模型技術，來對不同的野生蘭做種類的辨識，期望能夠提升機器對於蘭花辨識的準確性。

## 二、資料收集

### 1. 照片取得

因為去年才開始出野外拍攝蘭花，因此這次作業所使用的圖片基本上均是這一年來自己拍攝取得。在蒐集資料的過程花費了此次作業絕大部分的時間。因為自己出野外的次數不多，故所拍攝過的蘭種有限(台灣約有 400 種野生蘭，自身拍過的僅 120 種)，再扣除僅見過一兩次的蘭種(照片量過少無法做訓練)，以及一些過於模糊或是涵蓋兩種蘭花以上的照片之後，總共蒐集到僅約 935 張照片，涵蓋了 23 種蘭花種類。

### 2. 物種篩選及照片選擇

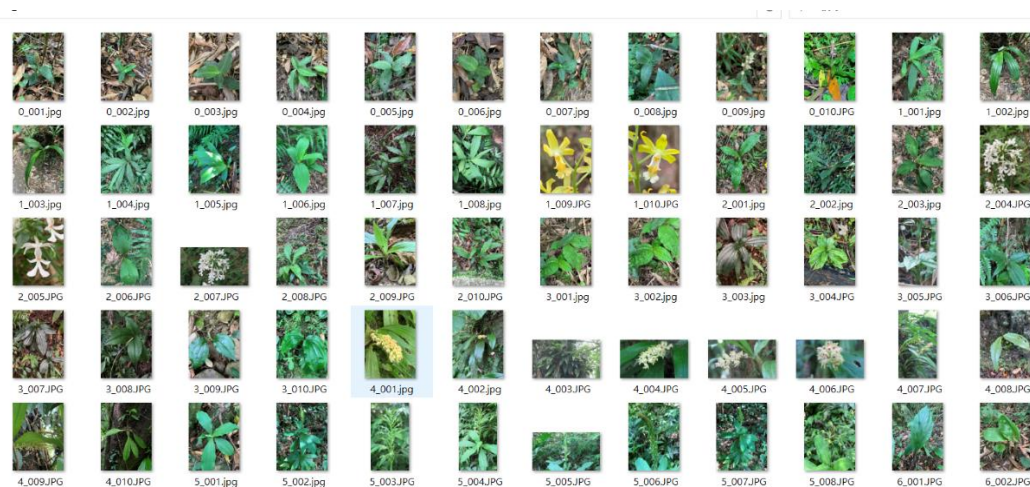
呈上點所述，在這次的作業中，為了提升使用上的實用性，而選用了 23 種台灣山區常見的野生蘭花，包括：高嶺斑葉蘭、綠花肖頭蕊蘭、白鶴蘭、大花羊耳蒜、黃絨蘭、冠毛玉鳳蘭、一葉蟬花蘭、長葉羊耳蒜、扁球羊耳蒜、長苞斑葉蘭、翹距根節蘭、台灣根節蘭、黃吊蘭、長葉杜鵑蘭、反捲根節蘭、鳳蘭、長距根節蘭、紋星蘭、大腳筒蘭、烏來石仙桃、黃萼捲瓣蘭、白石斛以及阿里山根節蘭，加上這些蘭種自己有較多拍攝經驗，故有相對較多的照片量；而在照片的選擇上，除了選擇自己已經確定種類的照片來訓練之外，在同一種蘭花當中也盡量選擇不同族群的照片，以及不同特徵的照片，包括：整體植株生態照、葉片特寫照、花或果實的特寫照等，讓模型能去學習這些重要特徵；此外也刪除過於模糊的照片。處理完之後，平均下來每種蘭花有 40 張照片(最多 79 張，最少 8 張，算是一極不平衡的數據集)，故自己認為此次訓練上具有一定的難度。

## 三、資料前處理

### 1. 資料夾結構

首先將所有 935 張圖片，在確保每個 label 於 training set 與 testing set 中均有照片的前提下，隨意分成 769 張 training set 及 166 張 testing test(validation test)，分別存於 train 和 test 兩個資料夾中。而在兩個資料夾中，照片檔案命名的方式，以 label\_編號.jpg 的方式來命名，即 0\_001.jpg、0\_002.jpg.....以此類推，在每張照片名稱中底線前方的數字極

為該張照片的 label(圖一)。label 與蘭花種類名稱的對照詳見所附 excel 檔 label\_versus\_Orchid\_name.xlsx。



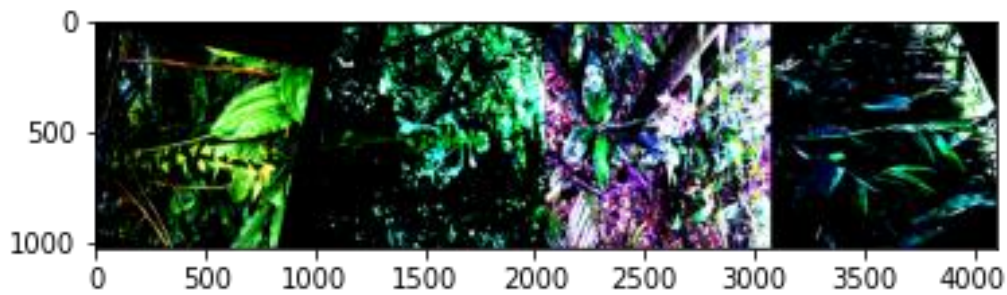
圖一 testing dataset(共有 166 張圖片)

## 2. Colab 下載檔案

由於此次作業一樣是針對圖片做深度學習訓練，流程與期末專題的內容類似，因此程式原則上拿期末專題所使用的程式碼來做修改。一樣使用 Pytorch 框架來做訓練，定義好下載 dataset 的 class，包括初始化資料、取得資料(檔名和 label)和長度回傳(圖片數量)，再將包好的 dataset 用 DataLoader 函式下載下來。

## 3. Data augmentation

與期末專題一樣，透過 data augmentation 來增加 training set 的資料量。先將 training set 的圖片，統一大小為 1024\*1024、RGB 格式，並水平翻轉、仿射縮放及調整對比度等方式來擴增資料集容量，並將所有圖片轉換為 tesnor 的形式儲存。根據以上的這些處理方式，依照不同的參數，如：進行水平翻轉的機率、旋轉的角度(15~30 度)、縮放比例(0.8~1.2 倍)等，總共設置 5 種變化組合，故資料集的容量經過處理後放大了五倍，最後用來訓練的圖片數量共有  $769 * 5 = 3845$  張，用來 validation 的圖片則一樣為 166 張。值得一提的事，若是使用 normalization，會發現雖然照片的顏色會稍微失真，但有些照片卻能夠藉由將背景顏色加深，而成功把蘭花植株的部分凸顯出來(圖二)；而沒有做 normalization 的圖片看起來相對較無變化。實際訓練完之後，兩者準確度未有太大差異(詳見後方說明)。



圖二 normalization 後的蘭花圖片

#### 四、模型訓練

##### 1. Pretrained model

根據 final project 的經驗，使用 efficientnet 具有較佳的訓練準確度，因此在本次作業中也是使用 efficientnet 來訓練。由於此次訓練所使用圖片解析度較高，為避免 Colab 記憶體爆炸，因此使用 efficientnet\_b2 作為此次訓練的 pretrained model，下載欲訓練好的 weight，並更改輸出的 feature 數為 23。

##### 2. Hyperparameters

在超參數的使用方面，Batch size 設定為 4，Optimizer 使用 Adam、learning rate 設定 0.0001、weight decay  $1e-05$ ，epoch 設 3~7，1 個 epoch 的訓練時間大約 15 分鐘。

#### 五、訓練結果及討論

如先前所述，由於此次訓練的資料量較少，加上要預測的 label 亦比期末專題的 4 個 label 來得多上許多，因此本來就十分擔心模型會訓練不起來。實際上訓練時，即便是使用 pretrained model，跑完一個 epoch 後的訓練準確度仍僅有 0.43~0.5，不過驗證準確度卻已有 0.70~0.80，而在訓練多一些 epoch 之後，發現訓練準確度及驗證準確度均有所提升，最後兩者均能夠達到 0.9 以上的準確度，其中圖片未做 normalize 的準確度約 0.92，做過 normalize 的準確度約 0.94，兩者均有不錯的 accuracy。訓練過程見下方表(一)所示。

表(一) 模型訓練準確度變化

Epoch	Training Accuracy	Validation(Test) Accuracy
1	0.43061	0.75595
2	0.75104	0.85119
3	0.88098	0.90476
4	0.92256	0.95238
5	0.95478	0.93452
6	0.96206	0.94048
7	0.96544	0.94048



原本擔心此次作業的資料量較少，加上資料量 label 不平衡，可能會難以訓練，實際上實作下來卻還是有不錯的 accuracy，令我感到有些驚艷。推測能有如此不錯的結果之可能原因，首先是自己所拍攝的照片解析度夠高，因此模型能擷取更為細部的特徵；再來是 label 量明顯較坊間的物種辨識器少，所以訓練較為容易；此外在所選的蘭花種類中，部分蘭種的彼此差異也較大，加上在某一些蘭花的圖片中，幾乎都是花的特寫(如：阿里山根節蘭)或是葉的特寫，也為此機器就容易去汲取每種蘭花所獨有的特徵出來訓練。

而觀察分類錯誤的照片中(以有做過 normalization，準確度較高的預測結果為主)，錯誤率較高的蘭花有綠花肖頭蕊蘭、白鶴蘭、反捲根節蘭、白石斛(以上各 2 張判斷錯誤)、烏來石仙桃及翹距根節蘭(以上各 1 張判斷錯誤)，其中白鶴蘭、反捲根節蘭與翹距根節蘭三者均同為根節蘭屬植物，外型較為相近，因而有分類錯誤情形：其中一張為葉緣有稍微波浪狀的白鶴蘭，被模型判斷為葉緣固定會有波浪捲的反捲根節蘭(圖三、四)；兩張有花的反捲根節蘭，被判斷成長距根節蘭(兩種的花均為紫色，但花形差異大，圖五~七)；此外，一張遠景的烏來石仙桃，因為背景環境相似而被判斷成大腳筒蘭(圖八、九)、另有一張綠花肖頭蕊蘭，因拍攝角度關係而被錯判成長葉杜鵑蘭(圖十、十一)、此外還有一張花部特寫的翹距根節蘭(圖十二、十三)，卻被判斷成白鶴蘭……。

然而另一方面，容易被登山客搞混的翹距根節蘭與阿里山根節蘭(圖十四、十五)，在此模型卻沒有彼此判斷錯誤的情況發生，阿里山根節蘭甚至全部判斷正確，在此推測可能是因為兩種蘭花的照片情況(如：環境等)差異較大，因此模型能夠成功判斷，實際上的情況可能需要再輸入兩種蘭花拍攝角度相似且背景類似的照片才能判定模型的好壞。



圖三、四 此株白鶴蘭(左)，因葉緣波浪狀被判斷為反捲根節蘭(右)





圖五~七 反捲根節蘭(左中)，因花色相同且背景相似而被判斷為長距根節蘭(右)



圖八、九 烏來石仙桃(左)，兩張生態環境相似而被判斷為大腳筒蘭(右)



圖十、十一 綠花肖頭蕊蘭(左)，因拍攝角度關係被判斷為長葉杜鵑蘭(右)





圖十二、十三 翹距根節蘭花部特寫(左)，被判斷成白鶴蘭(右)，表示模型對於細部特徵汲取仍有改進空間



圖十四、十五 外型相似易被混淆的翹距根節蘭(左)與阿里山根節蘭(右)，對此二種模型有不錯的判斷結果

有鑑於以上情形可得知，此模型雖然具有蠻高的準確度，但是對於汲取相似物種的細部特徵方面，仍有很大的改善空間；此外模型亦會將背景資訊如樹

枝、雜草等一起訓練，導致背景相似的照片會有判斷錯誤情形，因此如何降低背景雜訊，讓模型只提取蘭花植株的特徵，也是此模型可以再改善的一個面向。

## 六、結論

雖然整體來說這次訓練的蘭花辨識器準確度有 94 % 還算不錯，但這僅僅是透過少量的訓練資料及蘭花種類所做出來的結果，若是要讓此模型能有實際層面上應用，成為登山客能方便隨手使用的蘭花辨識器，仍有很長一段距離要走：第一，台灣的野生蘭有 400 多種，其中近半數為分布地十分狹隘的稀有蘭種，因此蒐集足夠照片具有相當大的難度；再來許多蘭花的特徵相當接近，甚至連專家都難以正確識別，因此若要讓機器能正確辨識出這些蘭種，想必將是一大挑戰。

當然希望自己在本課程結束之後若有餘力，能夠將此模型做更進一步的改良：透過多次的野外調查，擴大自己現有的 dataset、並增加訓練的蘭花種數 (label 數量)；並進一步了解 efficientnet 或是其他 pretrain model 的演算法，看看是否能提高模型提取蘭花植株特徵之機率。

## 七、參考資料

1. 課程講義

2. 期末專題報告-恩至割和他的鴨子

[file:///C:/Users/leo44/Desktop/ML\\_final/%E6%81%A9%E8%87%B3%E5%89%B2%E5%92%8C%E4%BB%96%E7%9A%84%E9%B4%A8%E5%AD%90.pdf](file:///C:/Users/leo44/Desktop/ML_final/%E6%81%A9%E8%87%B3%E5%89%B2%E5%92%8C%E4%BB%96%E7%9A%84%E9%B4%A8%E5%AD%90.pdf)