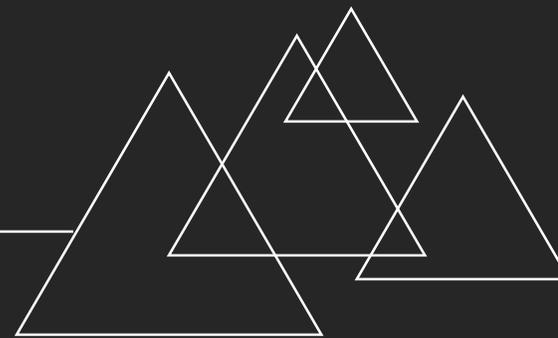
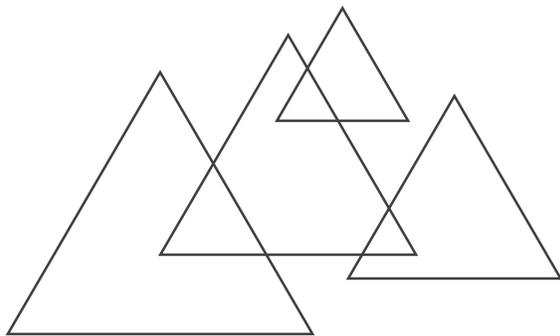


# 鳥類百科-鳥類圖像分類器

---

109034551-林奕辰





# 目錄

---

1. 背景

2. 研究方法

3. 模型訓練

4. 結果與討論

# 背景

鳥類為日常生活中最常見的脊索動物之一，無論在公園中或校園中都可以輕易的發現鳥類的蹤跡，據調查台灣總共記錄到87科674種鳥類，但是大多數民眾對於野鳥認識甚少，本研究希望藉由建立鳥類百科增進民眾對鳥類的認識。

除了增進一般民眾對於野鳥的知識外，本研究成果也可以使用於學術單位或是與野鳥保育的相關單位，可以串聯野外攝影機的影像資料，即時且全天候的分析獲得的影像資料。

## 夜鷺vs.企鵝

頭部延伸到背部



胖胖的身軀  
灰白色肚子  
尖尖的嘴吧  
黑灰色羽毛



相似8成



# 背景

---

## 5W1H

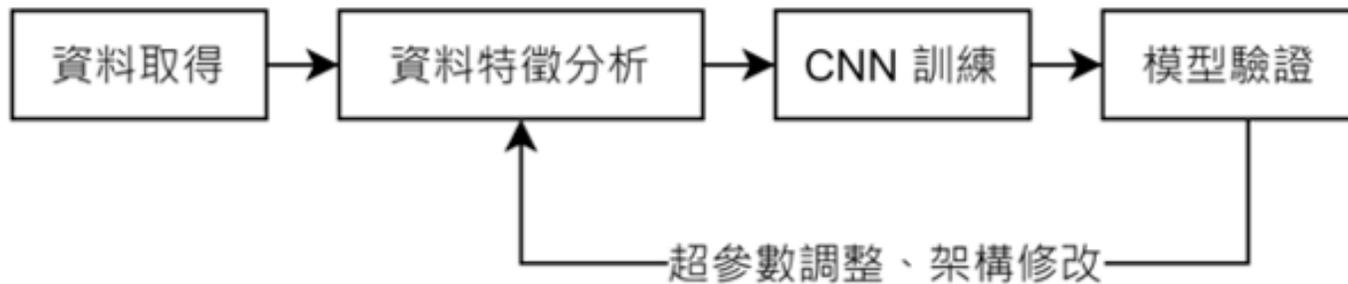
- Who 一般民眾、學術單位。
- What 鳥類品種之鑑定。
- When 需要辨識鳥類品種時。
- Why 鳥類種類繁多而且特徵難以辨識。
- How 電腦視覺、機器學習。



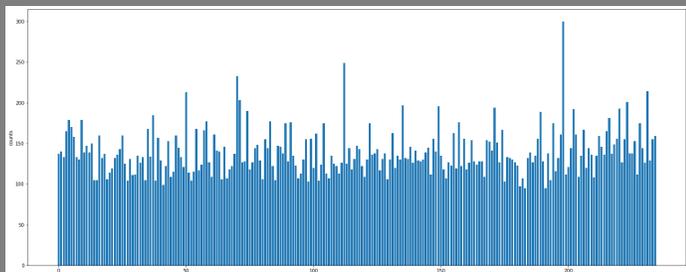
# 背景

## 研究流程

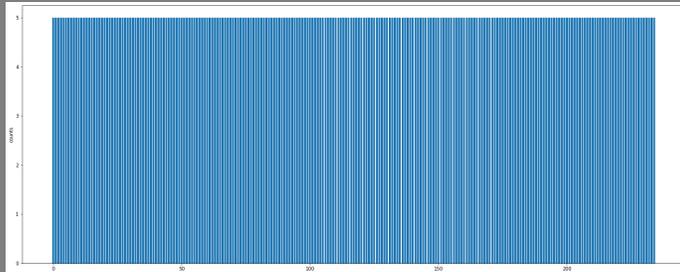
本研究將自 **Kaggle** 取得資料集，並對資料進行初步分析。再使用資料進行模型的訓練以及驗證，並以測試集評估模型的績效。藉此流程不斷修正模型以達到參數優化、提升準度的目標。



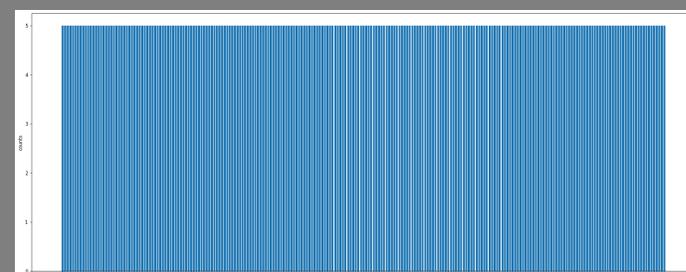
# 研究方法-資料特徵



訓練集資料分布



驗證集資料分布



測試集資料分布

## 資料特徵

本資料集使用 Kaggle 之 250 Bird Species data set，資料集被劃分為訓練集、驗證集以及測試集。資料集最新版本中包含 235 種鳥類的彩色圖片。訓練集中有 32672 張訓練用圖片，驗證集和測試集則各有 1175 張影像。

其中測試集和驗證集中各鳥類圖片都有 5 張，而訓練集中的各種別圖片數量最少有 95 張，最多的有 300 張，數量之間的差異並不極端，人工檢查後資料集中圖片大致正常，故直接使用本資料訓練模型。

```
count    235.000000
mean     139.029787
std      27.789255
min      95.000000
25%     122.000000
50%     133.000000
75%     153.500000
```

訓練集統計資料



# 研究方法-資料前處理



處理前影像

```
train_datagen=ImageDataGenerator(rescale=1./225)
val_datagen=ImageDataGenerator(rescale=1/255)
test_datagen=ImageDataGenerator(rescale=1/255)

train_generator=train_datagen.flow_from_directory(train_dir,
                                                  target_size=(200,200),
                                                  color_mode='rgb',
                                                  class_mode='sparse',batch_size=32,shuffle=True)

val_generator=val_datagen.flow_from_directory(validation_dir,
                                              target_size=(200,200),
                                              color_mode='rgb',
                                              class_mode='sparse',batch_size=32,shuffle=True)

test_generator=test_datagen.flow_from_directory(test_dir,
                                                target_size=(200,200),
                                                color_mode='rgb',
                                                class_mode='sparse',batch_size=32)

train_generator.class_indices
```



處理後影像

資料前處理

先將所需資料放置到所屬資料夾，使用 Keras 的功能將資料從資料夾中讀出，將圖片轉為(200,200,3)尺寸的tensor，並將色階值均一化，影像色階均一化可以使模型因為輸入值浮動較小，可以更準確的預測影像類別。類別encoding 使用sparse。此方式會將類別對應為一整數，如{'AFRICAN CROWNED CRANE': 0,'AFRICAN FIREFINCH': 1,'ALBATROSS': 2,}



# 研究方法-模型建立

選用Xception net 的架構，並在輸出的部分加入自定義的 layers 並調整超參數以達到提高模型準度的目標。

下表為該模型的基本資料，準確度為使用 ImageNet validation dataset 測試之結果。



Model	size	Top-1 acc	Top-5 acc	Parameters	depth
Xception	88MB	0.79	0.945	22910480	126



# 研究方法-模型建立

為了優化模型本研究測試了不同超參數以及激活函數

的組合，其組合如表所示。

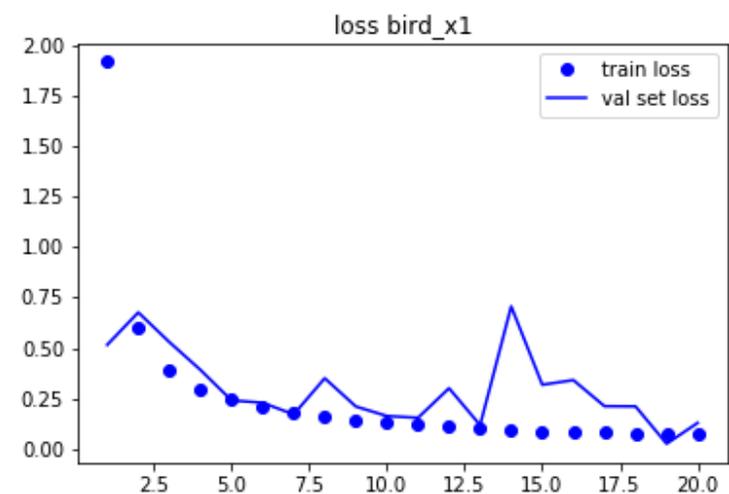
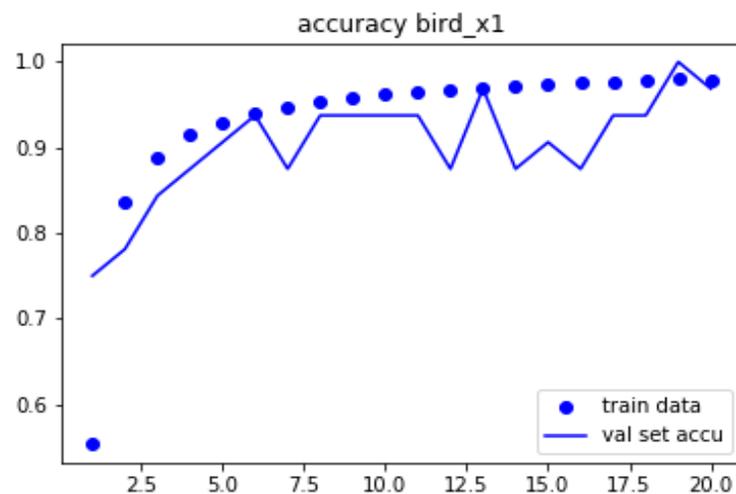
model	x1	x2	x3	x4	x5a
base	xception				xception
epoch	20	20	20	5	20
0	global_max_pooling2d				global_max_pooling3d
1	flatten				flatten
2	Dense 235 relu	Dense 235 sigmoid	Dense 1024 relu	Dense 1024 elu	Dense 1024 elu
3	Dense 235 softmax				Dense 235 softmax
loss	sparse_categorical_crossentropy				sparse_categorical_crossentropy
optimizer	adam	adam	adam	adagrad	adagrad
test acc	0.9523	0.9549	0.9609	0.9787	0.9898



# 模型訓練-X1

X1	超參數
epoch	20
2	Dense 235 relu
3	Dense 235 softmax
optimizer	adam

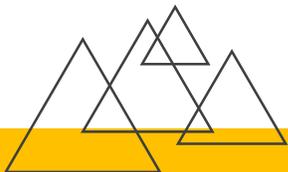
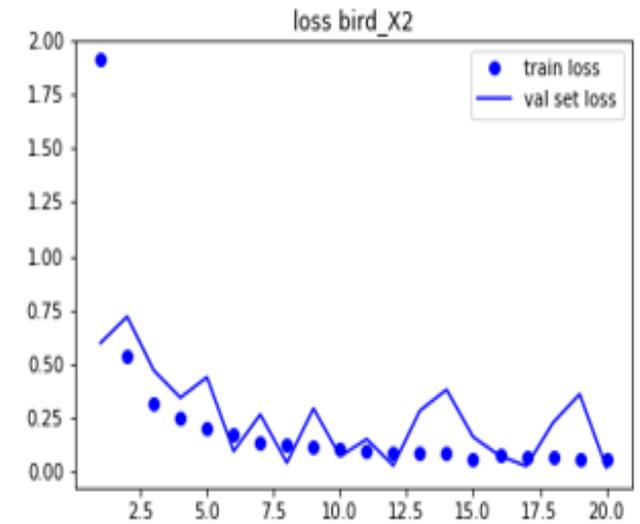
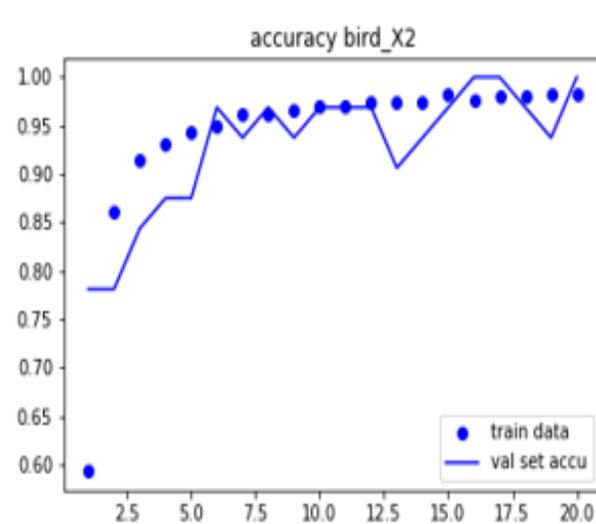
本模型的測試集準確度為**0.9523%**



# 模型訓練-X2

X1	超參數
epoch	20
2	Dense 235 relu
3	Dense 235 softmax
optimizer	adam

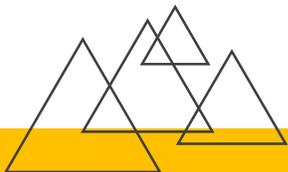
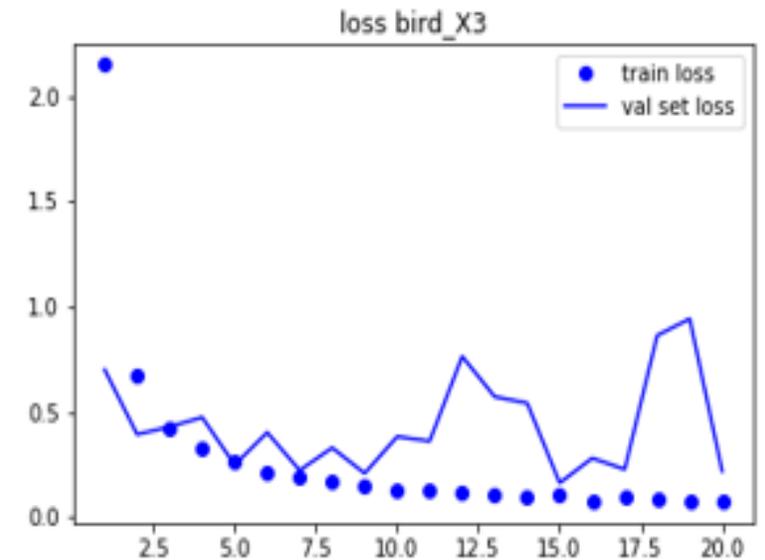
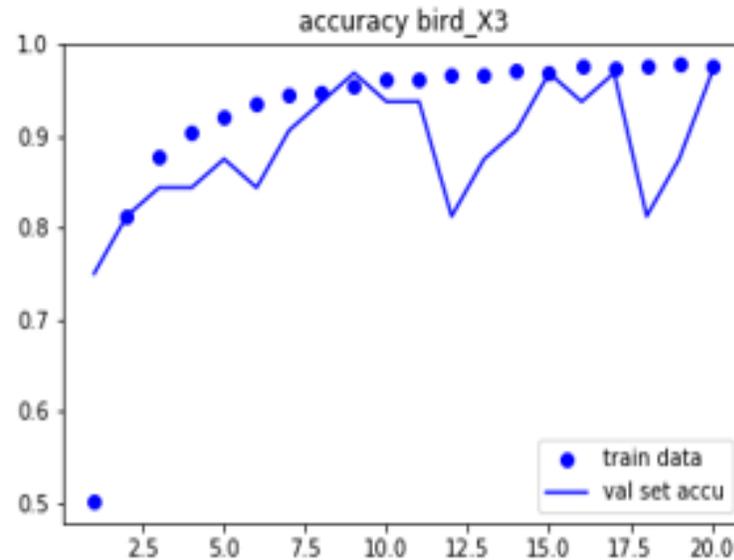
本模型的測試集準確度為**0.9549%**。在更換激活函數後準度並無明顯提高，推測是 `flatten` 層後的全連結層寬度不足造成資訊瓶頸，使得重要參數遺失，若想提升準度可以加寬此層。



# 模型訓練-X3

X1	超參數
epoch	20
2	Dense 1024 relu
3	Dense 235 softmax
optimizer	adam

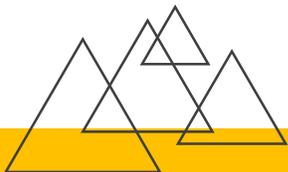
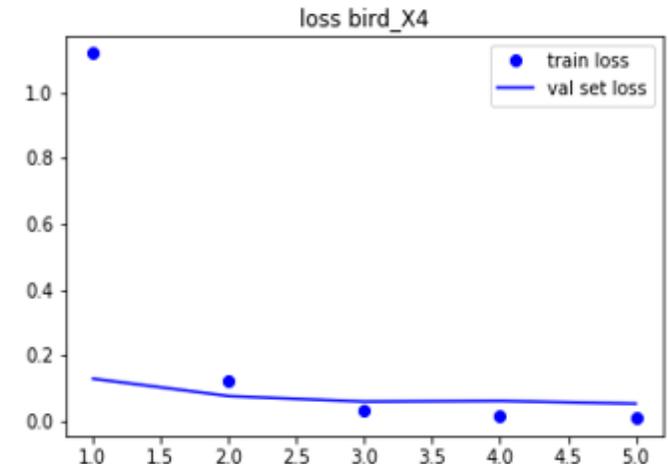
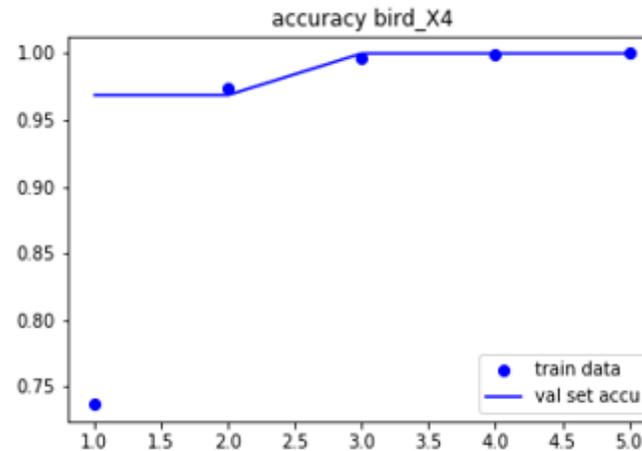
本模型的測試集準確度為**0.9609%**。在加寬全連結層之後準度有略微提升。



# 模型訓練-X4

X1	超參數
epoch	5
2	Dense 1024 elu
3	Dense 235 softmax
optimizer	adagrad

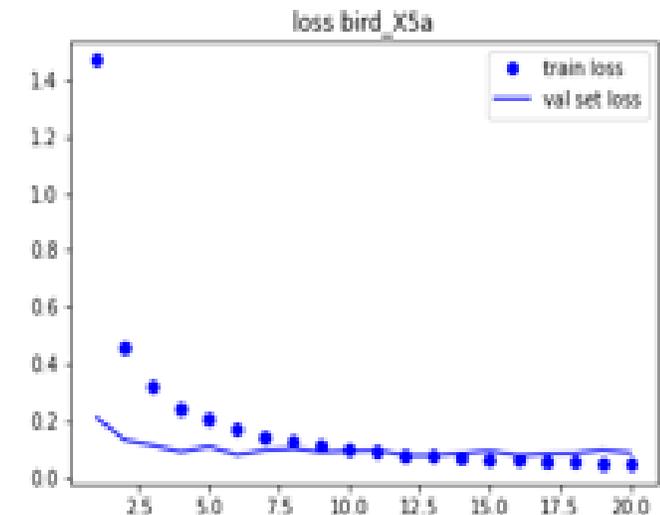
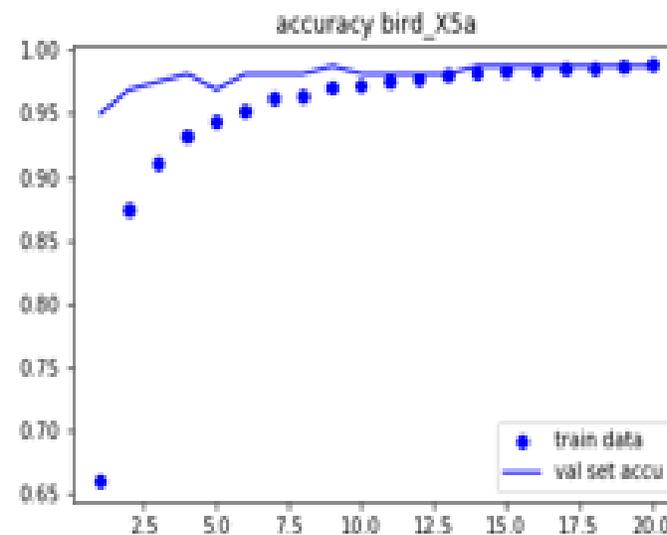
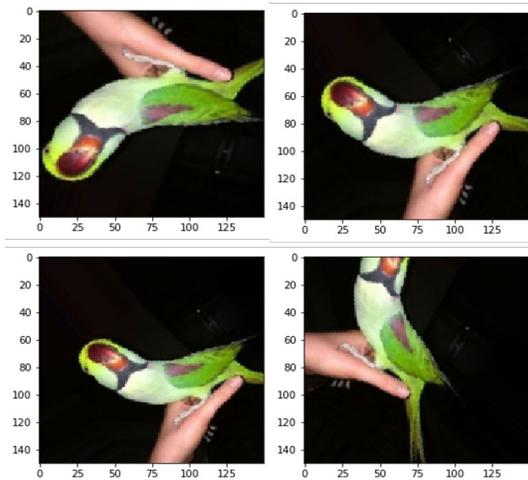
X4模型在flatten層後連接 1024個節點的全連結層激活函數為 elu (Dense 1024 elu)，elu可以避免梯度為0導致神經元死亡情形，並可以更快的訓練，之後連接 235個節點的全連結層激活函數為 softmax (Dense 235 softmax)，optimizer為adagrad。本次訓練時由於模型收斂速度快，為了避免過度擬合只訓練5個世代就結束訓練。本模型的測試集準確度為**0.9787%**，較前述參數組合高出不少。



# 模型訓練-X5a

X1	超參數
epoch	20
2	Dense 1024 elu
3	Dense 235 softmax
optimizer	adagrad

本模型沿用X4模型的架構但是為了避免過擬，在資料輸入時會對影像進行擴增，隨機對輸入影像進行旋轉、位移、縮放或是翻轉，擴增後結果如圖。使用擴增後的資料訓練可以避免過度擬合，並可以增加模型的泛化能力，在測試集中可以有較佳表現。由於可以避免過度擬合本次實驗訓練20個世代。**本模型的測試集準確度為0.9898%**，為所有模型中準度最佳的。



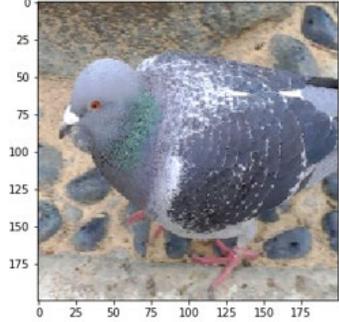


# 結果與討論

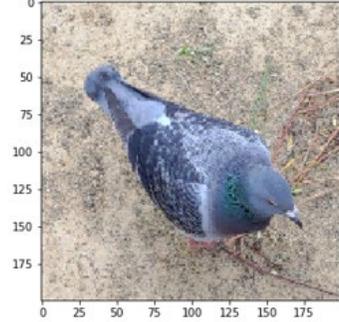
# 結果與討論-實際測試

以下圖片為在清大校園內拍攝的鴿子，使用績效最好的模型進行判斷，模型將以下照片都判斷為鴿子。因為本次使用資料集包含多國鳥類，所以只有少部分類別是在清大校園內找到，故本次測試時只有使用鴿子的圖像。由測試結果中可以看出本模型有實用化的潛力。

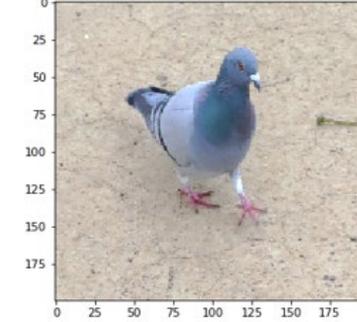
(99.99, '% chances are there that the Bird Is', 'ROCK DOVE', 'it is :', 'Rock dove')



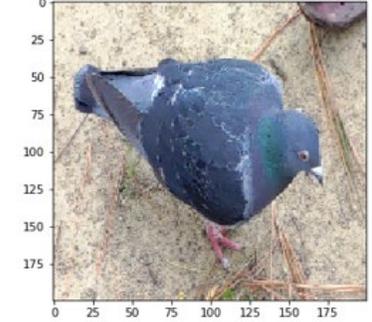
(98.78, '% chances are there that the Bird Is', 'ROCK DOVE', 'it is :', 'Rock dove')



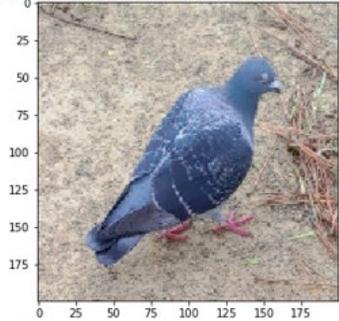
(99.61, '% chances are there that the Bird Is', 'ROCK DOVE', 'it is :', 'Rock dove')



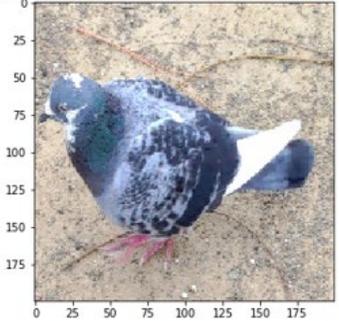
(99.41, '% chances are there that the Bird Is', 'ROCK DOVE', 'it is :', 'Rock dove')



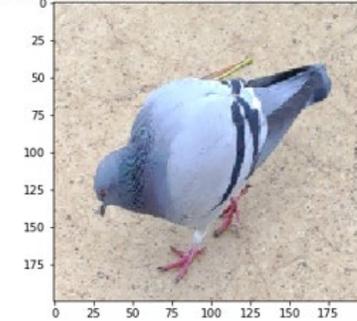
(100.0, '% chances are there that the Bird Is', 'ROCK DOVE', 'it is :', 'Rock dove')



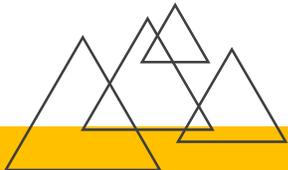
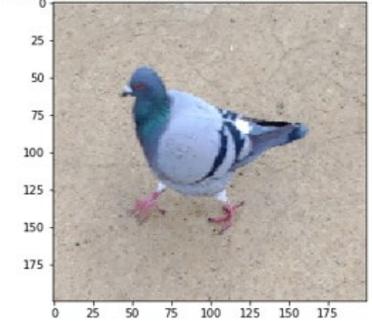
(99.98, '% chances are there that the Bird Is', 'ROCK DOVE', 'it is :', 'Rock dove')



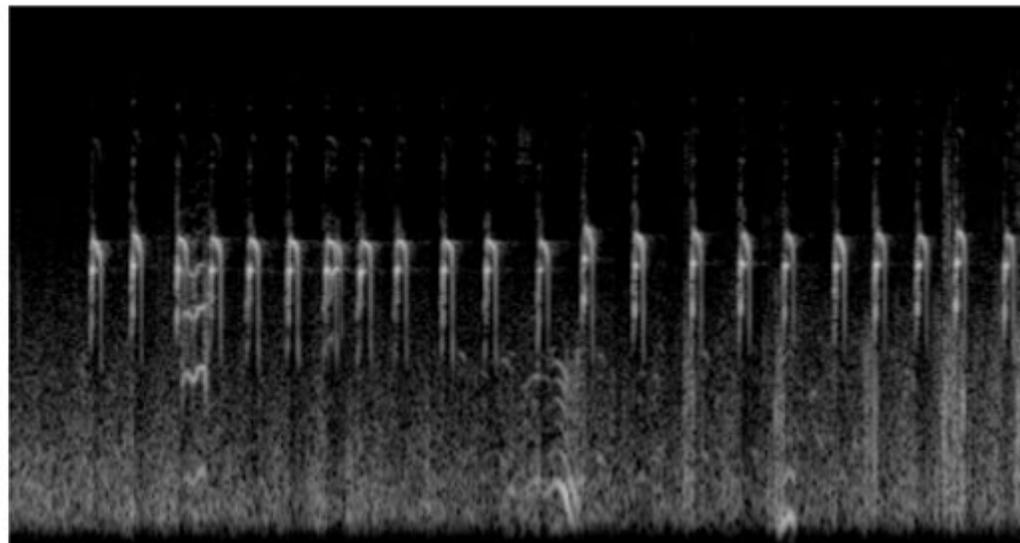
(67.65, '% chances are there that the Bird Is', 'ROCK DOVE', 'it is :', 'Rock dove')



(99.04, '% chances are there that the Bird Is', 'ROCK DOVE', 'it is :', 'Rock dove')



# 結果與討論-未來展望

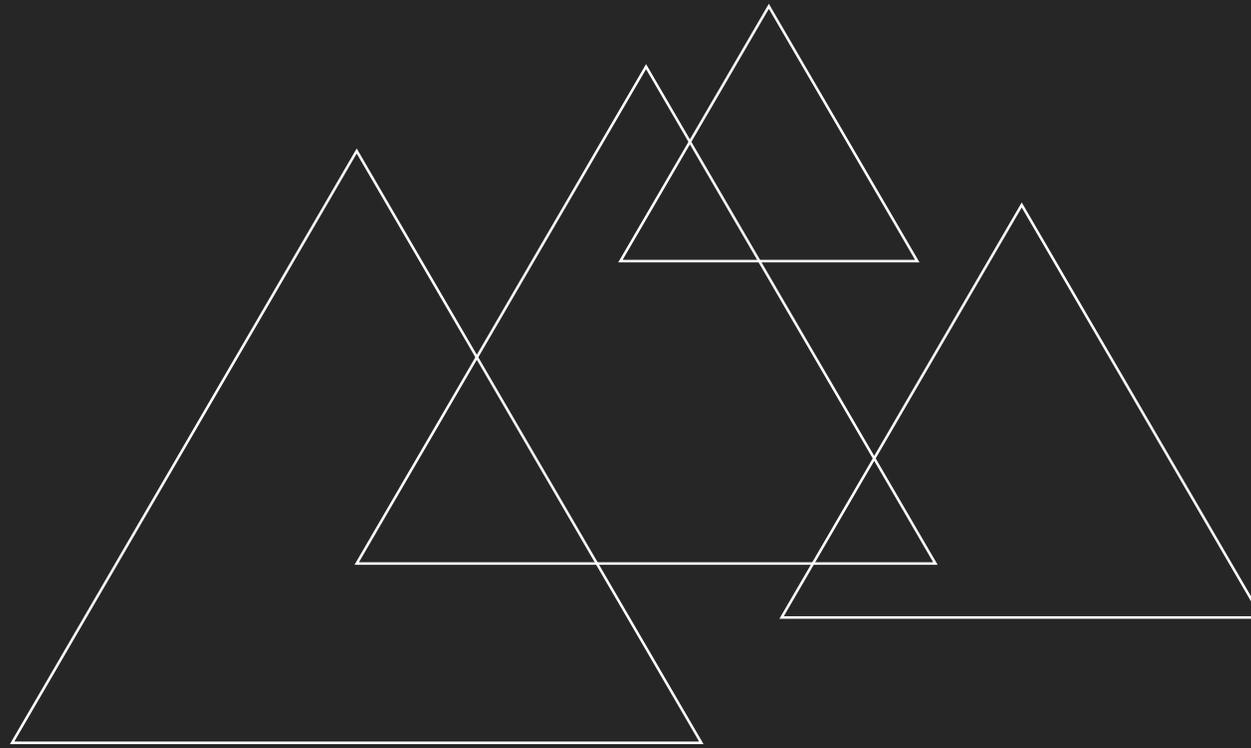


使用本土鳥類資料庫，讓模型

支援本土物種辨識

建立使用聲音辨識鳥類的模型，於多數時候在野外環境，因鳥類躲藏習性難以直接獲得鳥類照片，如可以藉由聲音辨識鳥類，則可以協助研究單位利用戶外收音設備長時間的觀察棲地中的鳥類。





**THANK YOU**