深度學習理論與實作

期末專題報告書

What the bird

組員：

C109151104 鄭政文

C109151150 吳俊傑

C109151162 許哲晟

C109151163 丁襄龍

研究動機

近年來，隨著科技的不斷進步，深度學習技術在各個領域都展現出了巨大的應用價值。雖然已經取得了一定成果，但在生態監控方面仍然存在著較大的挑戰。傳統的生態監控方法主要依賴人工處理，例如通過架設攝影機來捕捉生物的影像，然後依靠人類觀察影片以判斷其中的生物種類。這種方法存在著效率低下、人力時間成本高昂，並且容易受到主觀因素的影響，導致辨識結果可能不夠一致。透過機器辨識，我們可以節省這些成本，並提高效率。機器辨識的準確性也較高，因為機器不受情緒和疲勞的影響，而且可以實現大範圍、長時間的監測，即使在偏遠或困難進入的地區也能進行監測，這樣就可以獲取更全面、準確的鳥類資料。 其次，大量的鳥類數據需要處理和分析，這需要耗費大量時間和精力。透過機器學習，可以加速這一過程，提高處理效率。這樣我們就能更快地獲取、分析和應用鳥類數據，做出相應的決策和行動，對於保護生態環境也是極為重要的。 鳥類在生態系統中扮演著重要的角色，它們的分佈和行為反映了生態系統的變化和健康狀況。通過監測鳥類，我們可以了解生態系統的健康狀況，進而保護重要的生態環境和物種。

研究目的

我們的研究旨在開發一套鳥類辨識系統，以提高鳥類監測效率。通過增加訓練數據和引入深度學習技術，我們致力於提高系統的準確性和可靠性，確保監測結果的準確性。我們希望將這項技術應用於更廣泛的場景，包括自然保護區的監測、外來種控制和農業害鳥防治等，提升其社會價值。通過更準確的鳥類監測，我們可以更好地感知生態環境變化，為生態保育提供更有力的支持，並提供相應的保育建議，增加對生態保育的關注度。 我們的目標是建立長期的鳥類監測體系，持續追蹤鳥類群落變化，為生態保育政策提供科學依據，並進行生態系統健康狀態的長期監測。透過此次專案，我們希望開發易於操作且具有教育價值的鳥類辨識工具，普及鳥類知識、引導環境保護觀念，增加公眾對自然生態和生物多樣性的關注和了解。

技術探討

YOLO（You Only Look Once）是一種單次前向傳播的卷積神經網路，它以高效率和快速性著稱。與傳統的目標檢測方法相比，如滑動視窗或候選區域的生成（如R-CNN系列），YOLO將整個目標檢測過程簡化為一個單一的回歸問題。

YOLO將圖像分成多個網格單元，每個單元負責預測對應的目標框和框中的目標類別。這意味著在一次前向傳播中，YOLO可以直接預測圖像中的所有目標框和對應的類別概率。由於這種端到端的訓練方式，YOLO具有非常高的處理速度，使其能夠實時進行目標檢測。然而，雖然YOLO在處理速度和簡化流程方面表現出色，但在檢測小目標和密集目標時效果不如其他算法。對於特別小或特別大的目標，可能存在準確性問題。

YOLOv8的結構(圖2)由三個部分組成，其中骨幹網路(Backbone)負責多尺度卷積特徵擷取，主要作用是捕捉圖像中的細節，其stage 2會輸出低階特徵、stage 3會輸出中階特徵、stage 4中會經過SPPF後輸出高階特徵；頸部(Neck)目的是多尺度卷積特徵融合，呈現FPN和PAN結構，其中FPN （feature pyramid networks）即是特徵金字塔網路，採用多尺度來對不同大小的目標進行檢測；PAN為自底向上的特徵金字塔。這樣結合操作，FPN層自頂向下傳達強語意特徵，而特徵金字塔則自底向上傳達強定位特徵，從不同的主幹層對不同的檢測層進行特徵融合；頸部連接到頭部(Head)共有三個路徑，分別藉由不同階層的特徵來偵測圖片中的小、中、大型物件，而在v8版本中有用解耦合頭(decoupled-head)來個別偵測物件的位置和物件類別，來增加偵測的準確性。

PAN-FPN：CV中物件辨識的神經網路架構，它將特徵金字塔網路(FPN)與路徑聚合網路(PAN)結合，以提高物件辨識的準確性和效率；FPN用於不同比例的影像特徵提取，PAN用於跨網路的不同層聚合特徵。這允許網路檢測不同大小和分辨率的對象，並能處理具有更多對象的複雜場景

CSPLayer\_2Conv：簡稱C2F，透過跨層部分連接機制，在不同層之間建立直接連接，即是在卷積特徵始過ConVModule之後會被分成一半，一部分經過n個DarknetBottleneck進行深度的CNN運算，另一部分會跳過，並在最後會融合起來再進經過ConVModule，這樣可提高網路資訊流動性和特徵表達能力。

DarknetBottleneck：該模塊包含兩個ConVModule，並且有兩種類型，在add = True時會使用殘差連接，使其解決深度神經網路訓練過程中的梯度消失和梯度爆炸等問題，從而實現更深的網絡架構。

ConVModule：其包含一個二維的卷積層(Conv2d)、一個批次正規化、一個SiLU激勵函數，即是一個卷積塊。

SPPF：其前身為SPP，即空間金字塔池化，SPP的主要思想是在不同大小的子區域上應用池化操作，然後將所有子區域的池化結果結合起來。這使得模型能夠處理任意大小的圖像，而不需要將它們調整為相同大小，可以避免發生固定池化層大小使其偵測尺度不容易調整的問題。SPP的一個主要缺點是計算量較大，因為它需要在不同大小的子區域上應用池化操作。而SPPF的理念是藉由小型池化層來取代大型池化層，像是(圖3) K=13可用三個K=5池化層取代，這樣可以減少運算量

Upsample：用於在特徵金字塔網絡（FPN）或解碼器部分增加特徵圖的尺寸，從而實現對不同尺度目標的檢測。常用得的方法有：最近鄰插值(Nearest neighbor interpolation)、雙線性插值(Bi-Linear interpolation)、雙立方插值(Bi-Cubic interpolation)、三線性插值(Trilinear Interpolation)、反池化(UnPooling

)、反卷積(Deconvolution)。

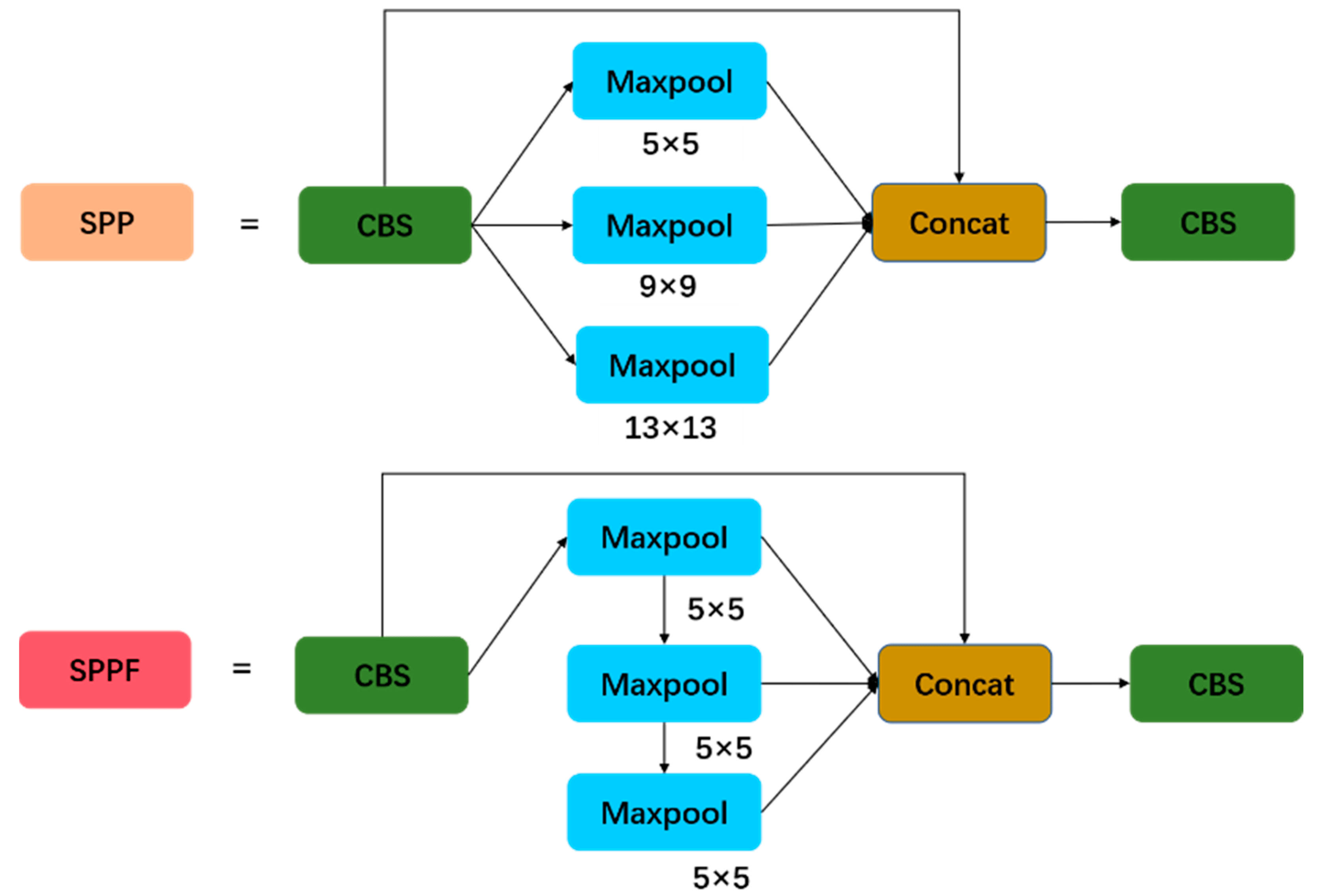
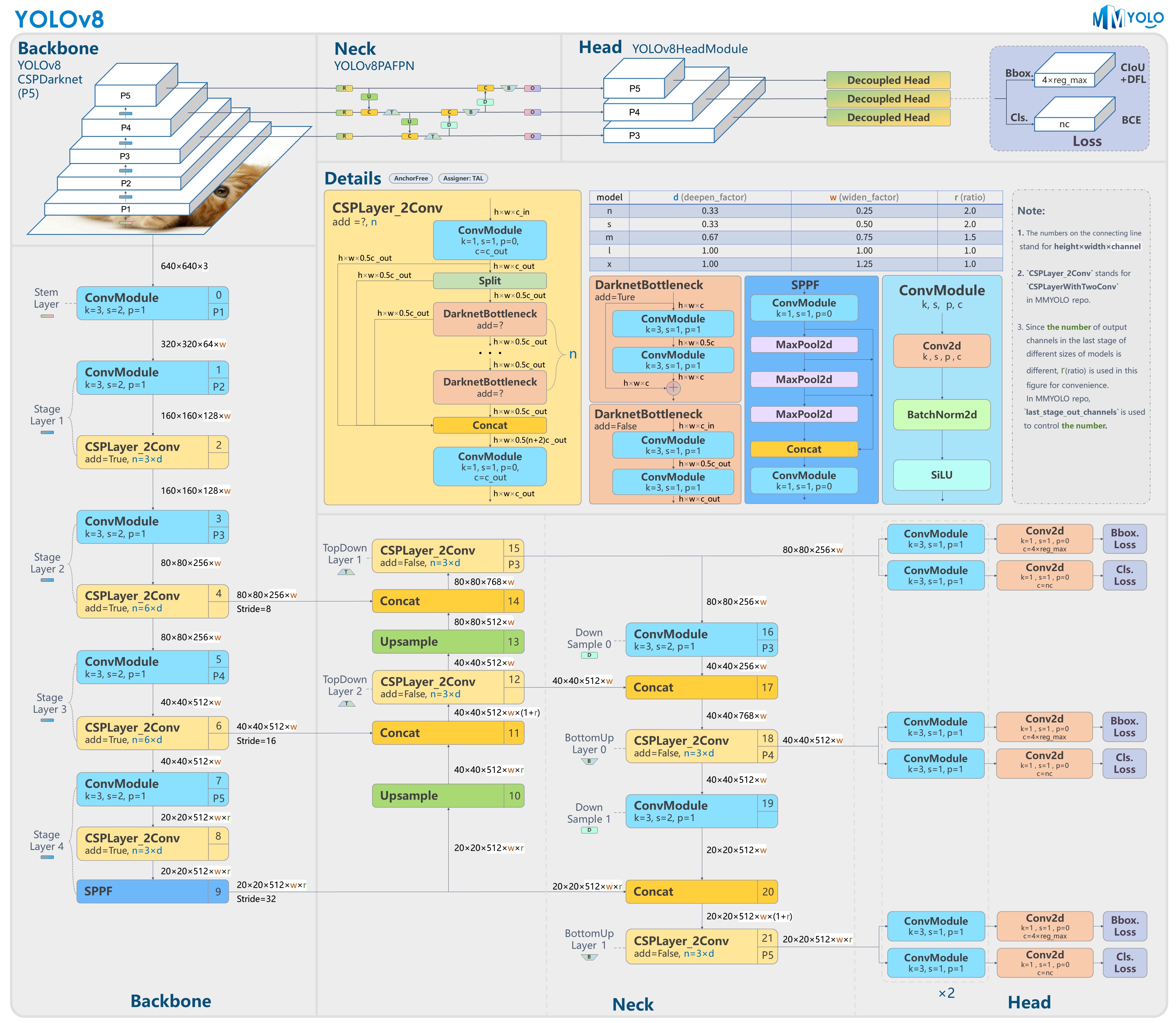
圖2，YOLOv8結構圖

圖3、SPP和SPPF比較圖

研究方法

資料收集

我們在一個有著台灣所有鳥類的資料庫網站「eBird」進行資料收集，使用了Google的One click Image Downloader網頁插件來快速下載(表1)列出的鳥類圖片作為本次研究的資料集，總共有3475張鳥類照片，並將資料集採7:3的方式切割為如表2所示的數量作為訓練集與測試集。

|  |  |  |
| --- | --- | --- |
| 鳥類名稱 | 數量 | 比例 |
| 岩鷚 | 130 | 3.74% |
| 黑冠麻鷺 | 388 | 11.16% |
| 紋翼畫眉 | 151 | 4.34% |
| 紅尾伯勞 | 300 | 8.63% |
| 山麻雀 | 180 | 5.17% |
| 環頸雉 | 556 | 16% |
| 臺灣白喉噪眉 | 122 | 3.51% |
| 董雞 | 92 | 2.64% |
| 煤山雀 | 290 | 8.34% |
| 麻雀 | 361 | 10.38% |
| 野鴿 | 282 | 8.11% |
| 鉛色水鶇 | 239 | 6.87% |
| 八哥 | 384 | 11.05% |

表1 資料集中各鳥類的數量與比例

|  |  |  |
| --- | --- | --- |
|  | 訓練集 | 測試集 |
| 圖片數量 | 2433 | 1042 |

表2 訓練集與測試集數量

資料前處理

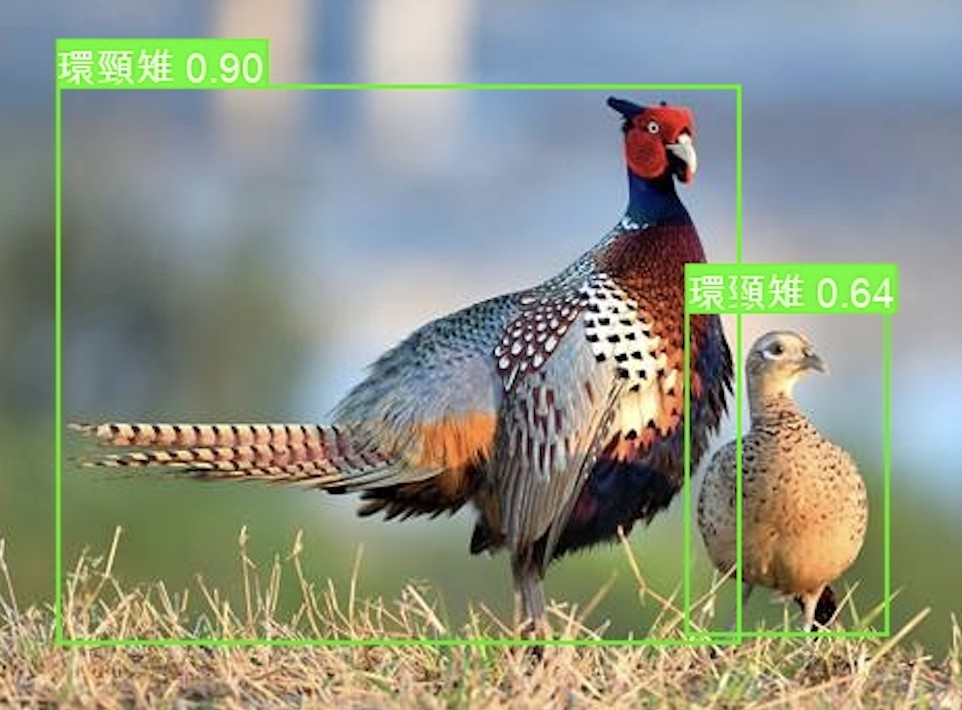
使用了makesense網站，對所有鳥類圖片進行框選並標記出圖片中所有鳥的種類（圖1）。



圖1 左為標記前右為標記後

預期結果

使用者提供一張包含鳥類的圖像，例如一張自然風景圖中有多隻鳥飛行或停棲在樹枝上的圖片。使用預先訓練好的 YOLO 模型對輸入圖像進行物件檢測。該模型已經學習了鳥類的外觀特徵，能夠有效地檢測並定位圖像中的鳥類。系統將在輸出圖像上標記出檢測到的鳥類的位置和類別。每個檢測結果包括一個邊界框，表示鳥類的位置，以及一個標籤，表示鳥類的類別， 如果圖像中有多種不同種類的鳥類，系統將檢測並標記所有不同類別的鳥類。並確保檢測結果的準確性和效率。它將快速地且準確地辨識圖像中的鳥類，並且能夠處理各種不同場景下的鳥類檢測任務。



引用

[1]<https://ebird.org/region/TW>

[2]<https://github.com/ultralytics/ultralytics>

[3]<https://www.makesense.ai/>

[4][https://mmyolo.readthedocs.io/en/latest/recommended\_topics/algorithm\_descriptions/yolov8\_description.html](https://mmyolo.readthedocs.io/en/latest/recommended_topics/algorithm_descriptions/yolov8_description.html%5C)

[5] [https://arxiv.org/abs/1803.01534v4](https://arxiv.org/abs/1803.01534v4%5C)

[6] <https://chromewebstore.google.com/detail/one-click-image-downloade/djcobamaplcmhmaocomnkfdbcoiggepo>

[7]生態辨識相關報導