深度學習課程報告 –

Object Detection

**學生：吳玫萱 RE6091054**

Github：https://github.com/meihsuan0301/Deep-learning-class

參考資料：<https://github.com/sgrvinod/a-PyTorch-Tutorial-to-Object-Detection>

最後成績(上傳截止)

**Public set：**



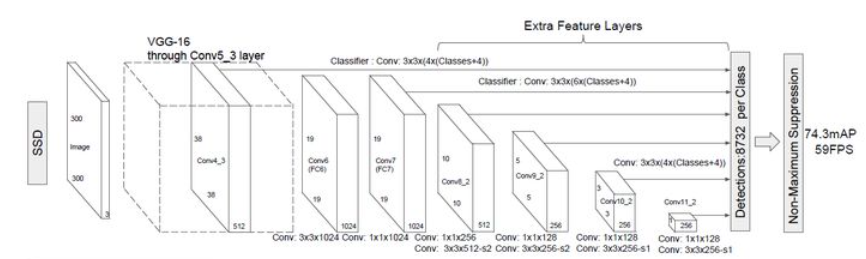
一、大綱

每年登革熱的流行季節，衛生局需派員稽查各住家、社區及髒亂地點是否有積水容器，然而容器種類眾多容易被忽略，且以人工辨識往往耗費太多人力與時間成本。因此本議題提供各種積水容器之標註資料，希望能透過影像物件偵測技術，能讓稽查人員藉由影像或是視訊提醒其積水容器之物件位置，本報告採用Single Shot MultiBox Detector(SSD)方法進行任務。

二、SSD物件徵測方法介紹

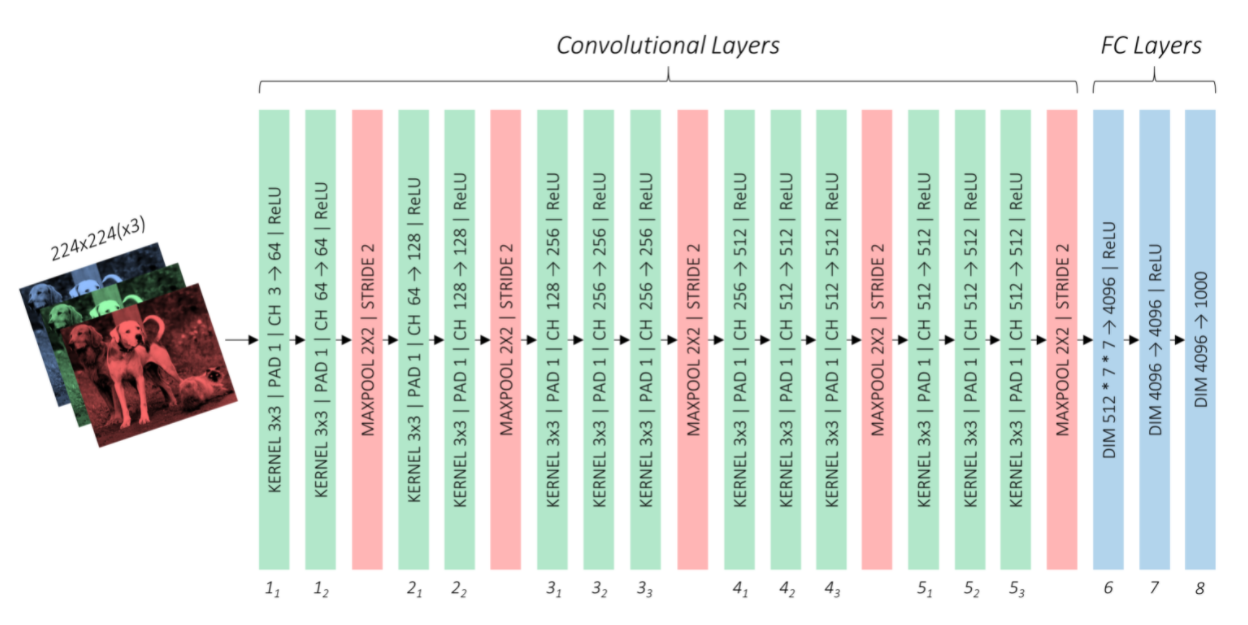
**(一) 網路結構**

SSD採用VGG16作為基礎模型，然後在VGG16的基礎上新增了卷積層來獲得更多的特徵圖以用於檢測。SSD的網路結構如下圖所示。可以看到SSD利用了多尺度的特徵圖做檢測。原論文模型的輸入圖片大小是。

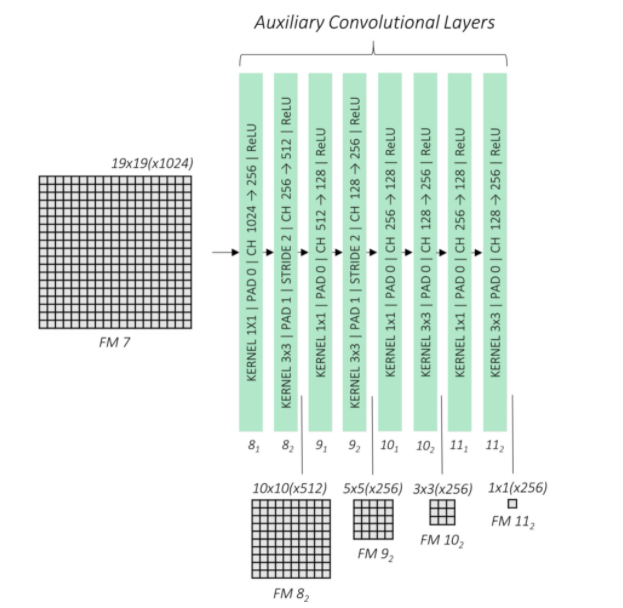


可以將模型組織成三個部分：

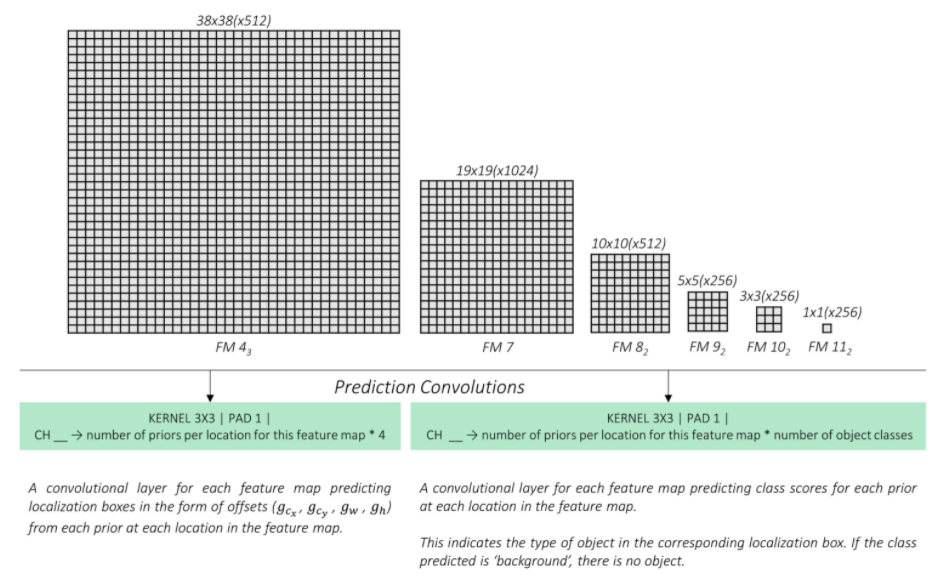
* 源自現有圖像分類架構的**基礎卷積**將提供較低級別的特徵圖。



* 添加在基礎網絡之上的**輔助卷積**將提供更高級別的特徵圖。



* 將在這些特徵圖中定位和識別對象的**預測卷積**。



**(二) 概念點簡述**

* **Single-Shot Detection**

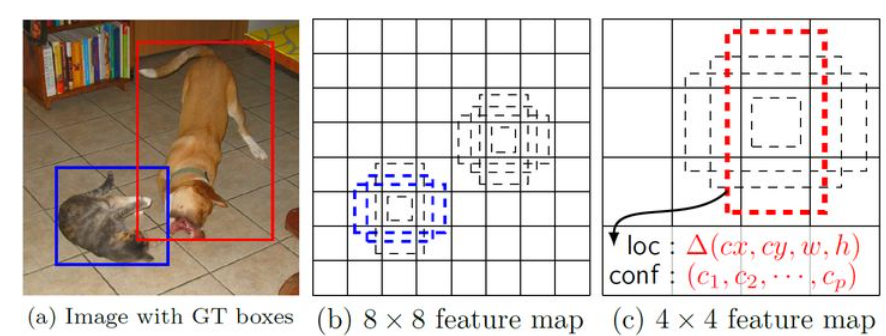
**早**期的對象檢測架構由兩個不同的階段組成—執行對象定位的區域提議網絡和用於檢測提議區域中對像類型的分類器。在計算上，這些可能非常昂貴，因此不適合現實世界的實時應用程序。單次模型將定位和檢測任務封裝在網絡的單次前向掃描中，從而顯著加快檢測速度，同時可部署在更輕的硬件上。

* **Multiscale Feature Maps**

在圖像分類任務中，我們的預測基於最終的卷積特徵圖—原始圖像的最小但最深的表示。在目標檢測中，來自中間卷積層的特徵圖也可以直接有用，因為它們代表不同尺度的原始圖像。因此，在不同特徵圖上操作的固定大小過濾器將能夠檢測各種大小的對象。

* **Priors**

這些是在特定特徵圖上的特定位置定義的預先計算的框，具有特定的縱橫比和比例。它們經過精心選擇，以匹配數據集中對象邊界框（即基本事實）的特徵。



* **Multibox**

這是一種將對象邊界框的預測公式化為回歸問題的技術，其中將檢測到的對象的坐標回歸到其真實坐標。此外，對於每個預測框，都會為各種對像類型生成分數。先驗作為預測的可行起點，因為它們是根據基本事實建模的。因此，將有與先驗一樣多的預測框，其中大多數將不包含任何對象。

* **Hard Negative Mining**

這是指明確選擇模型預測的最嚴重的誤報，並迫使它從這些示例中學習。換句話說，我們只挖掘模型發現最難正確識別的那些負面因素。在物體檢測的背景下，絕大多數預測框不包含物體，這也有助於減少負正不平衡。

* **Non-Maximum Suppression**

在任何給定位置，多個先驗可以顯著重疊。因此，由這些先驗產生的預測實際上可能是同一對象的重複。Non-Maximum Suppression (NMS) 是一種通過抑制除最高分數之外的所有預測來消除冗餘預測的方法。

三、實驗設置與實驗結果

**資料集**

由衛生福利部疾病管制署提供之蒐集自台灣各地具積水容器之場景，時間從 2008 年至 2018 年，其中 75% 的影像資料，長寬比為 400 x 300。訓練資料( train\_cdc.zip )共 2671 張，測試集資料共 2248 張。

每張影像提供其標註物件與物件 bounding box，標註積水容器物件類別有 aquarium、bottle、bowl、box、bucket、plastic\_bag、plate、styrofoam、tire、toilet、tub、washing\_machine、water\_tower，共13個類別。

**模型相關參數**

* 損失：MultiBoxLoss
* 優化器：SGD
* Batch size：8
* 學習率：初始學習律為0.001，並在多次迭代後衰減學習率，當iterations 到達80000, 100000時，學習率皆衰減致當下的0.1倍。
* 標準化方式：resize到，減平均除標準差。
* iterations = 500000
* 訓練使用機器：RTX 3090 24G x 4

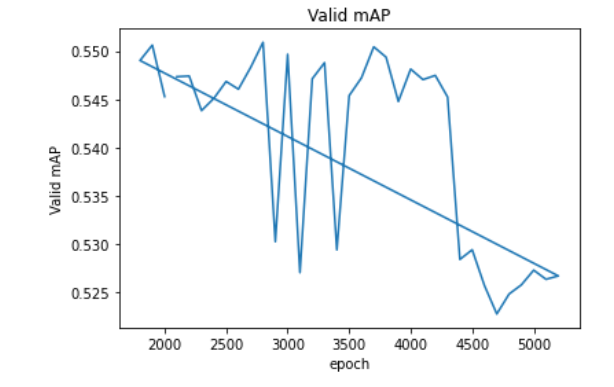
**評估標準**

官方評估方式採用mean Average Precision (mAP)於intersection over union (IoU) threshold 為 0.5。根據預測物件框與標註物件框IoU大於 0.5 時為 True Positive(TP)，反之為 False Positive(FP)，藉此得到 Precision。因此系統針對每一個物件評估其 AP 分數，再將 13 類積水容器物件 AP 進行平均計算，得到 mAP 評估數值。

**實驗結果**

1. 下表為所有實驗結果表，mAP部分展示的是test所得mAP結果，也就是上傳後所得結果。

|  |  |  |  |
| --- | --- | --- | --- |
| Epoch | min\_score | max\_overlap | mAP |
| 1895 | 0.2 | 0.5 | 0.3245256 |
| 1733 | 0.2 | 0.5 | 0.3986766 |
| 4804 | 0.2 | 0.5 | 0.3893562 |
| 5139 | 0.2 | 0.5 | 0.3767811 |
| 4804 | 0.15 | 0.5 | 0.4042980 |
| 4804 | 0.1 | 0.5 | 0.4125618 |
| 4804 | 0.01 | 0.5 | 0.4482412 |
| 1733 | 0.01 | 0.5 | 0.4468712 |
| 1733 | 0.01 | 0.45 | 0.4466649 |
| 1733 | 0.05 | 0.45 | 0.4235642 |
| 1733 | 0.005 | 0.45 | **0.4525664** |

1. 下表展示探討Epoch區間之valid mAP。
2. 可視化徵測結果範例圖



