

積水容器影像物件辨識

藍孟彬 國立清華大學

許健平教授-高速通訊與計算實驗室

Outline

- 目標
- 解題方法
 - 資料
 - 模型與超參數
- 結果
- 延伸應用

目標

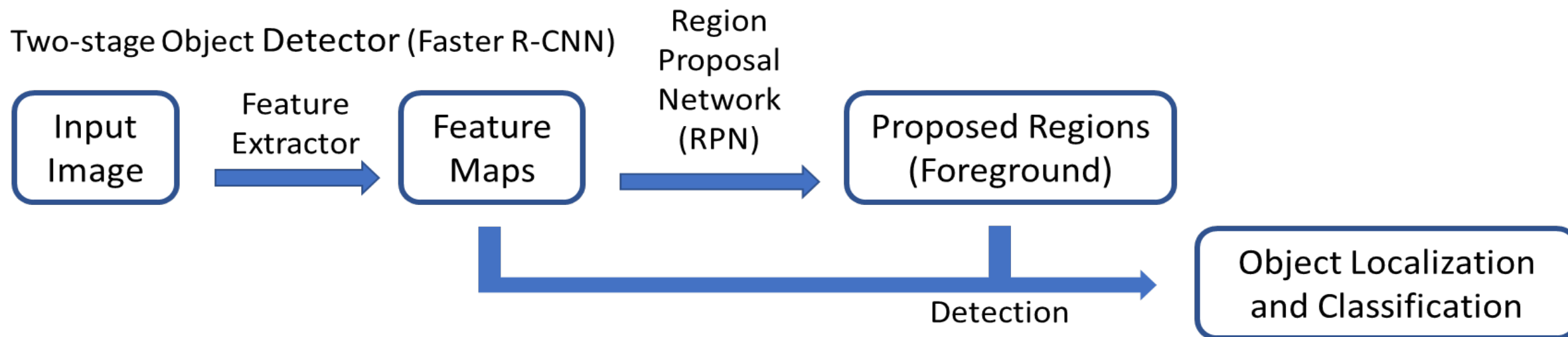
- 目標: 為了防治登革熱疫情，建立一個積水容器圖片分類器，來自動化尋找積水容器。
- 方案:

Object Detection using Deep Learning



主流的物件偵測演算法

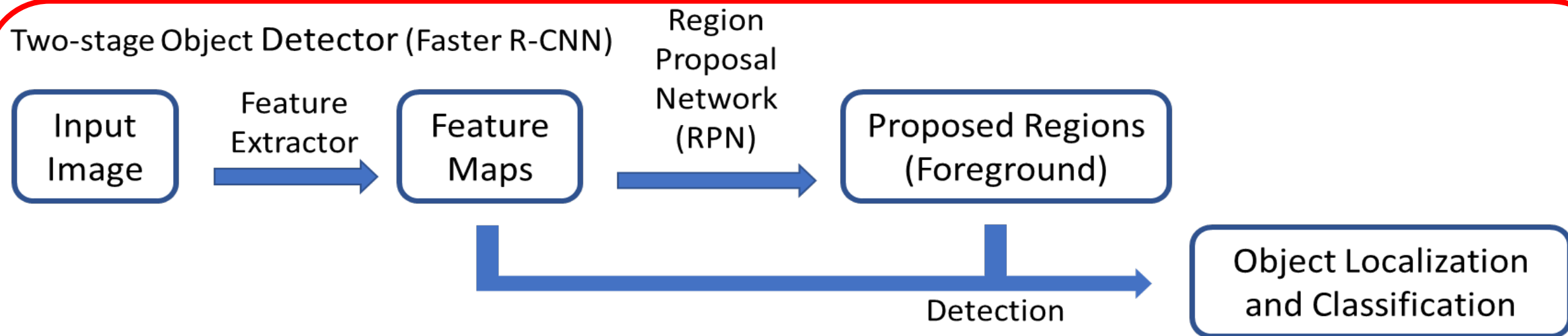
Two-stage Object Detector (Faster R-CNN)



One-stage Object Detector



主流的物件偵測演算法



One-stage Object Detector



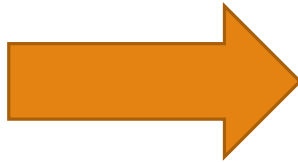
Two-stage object detector(1/2)



Two-stage object detector(2/2)



Region
Proposal
Network
(RPN)



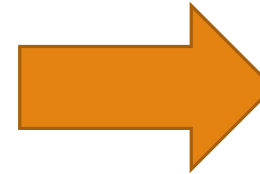
Foregrounds



Backgrounds



Detection



Bucket



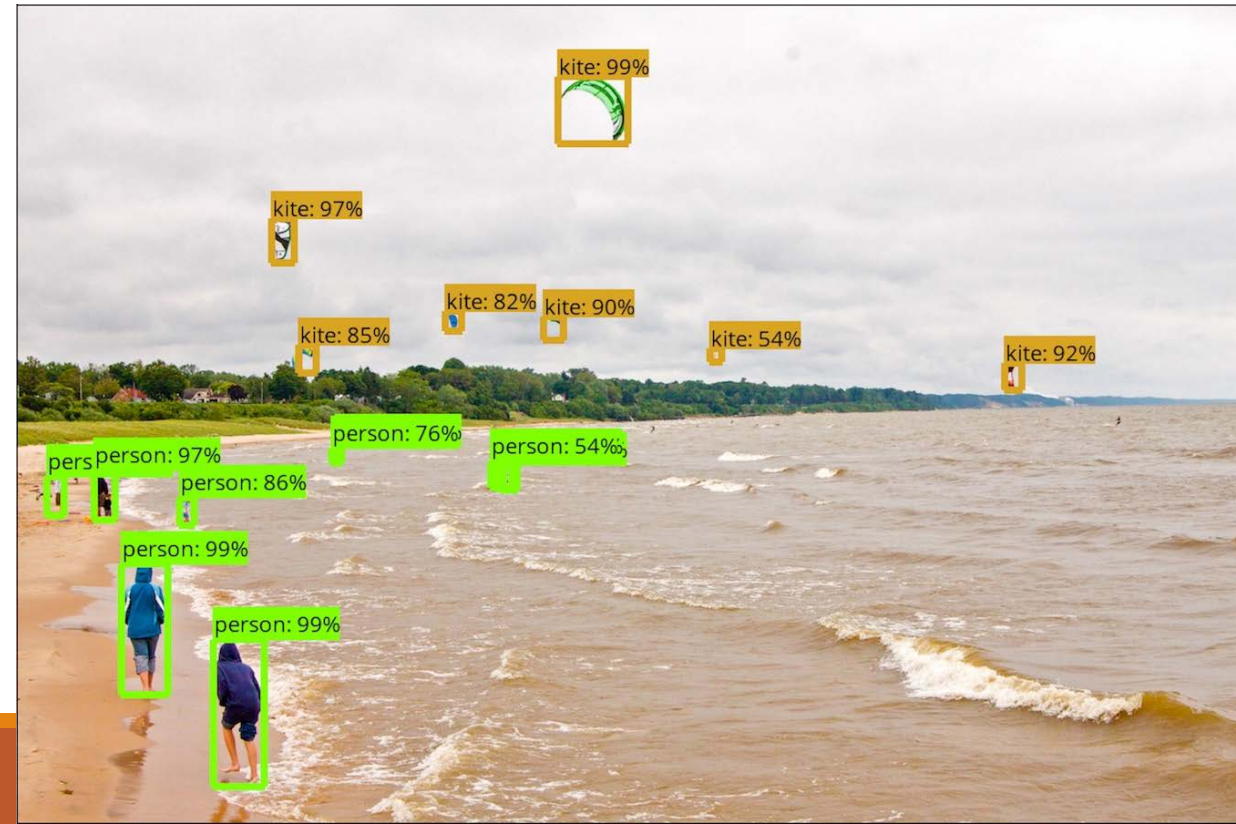
Plate



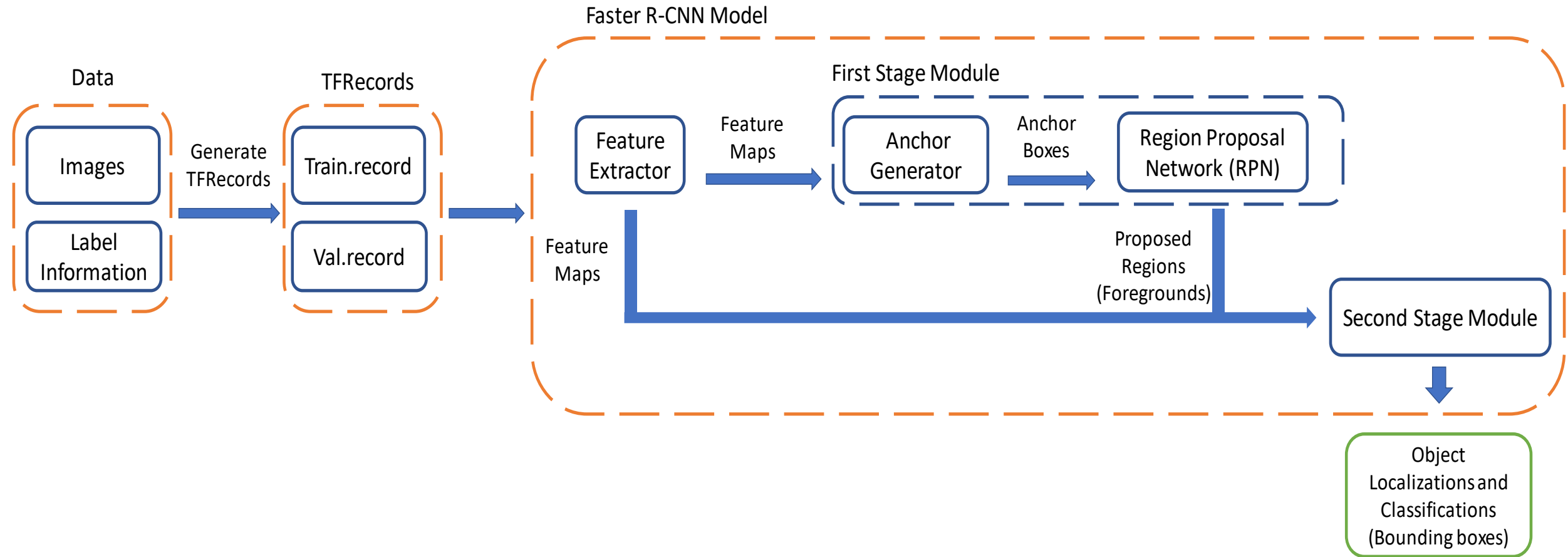
TensorFlow Object Detection API

- Faster R-CNN

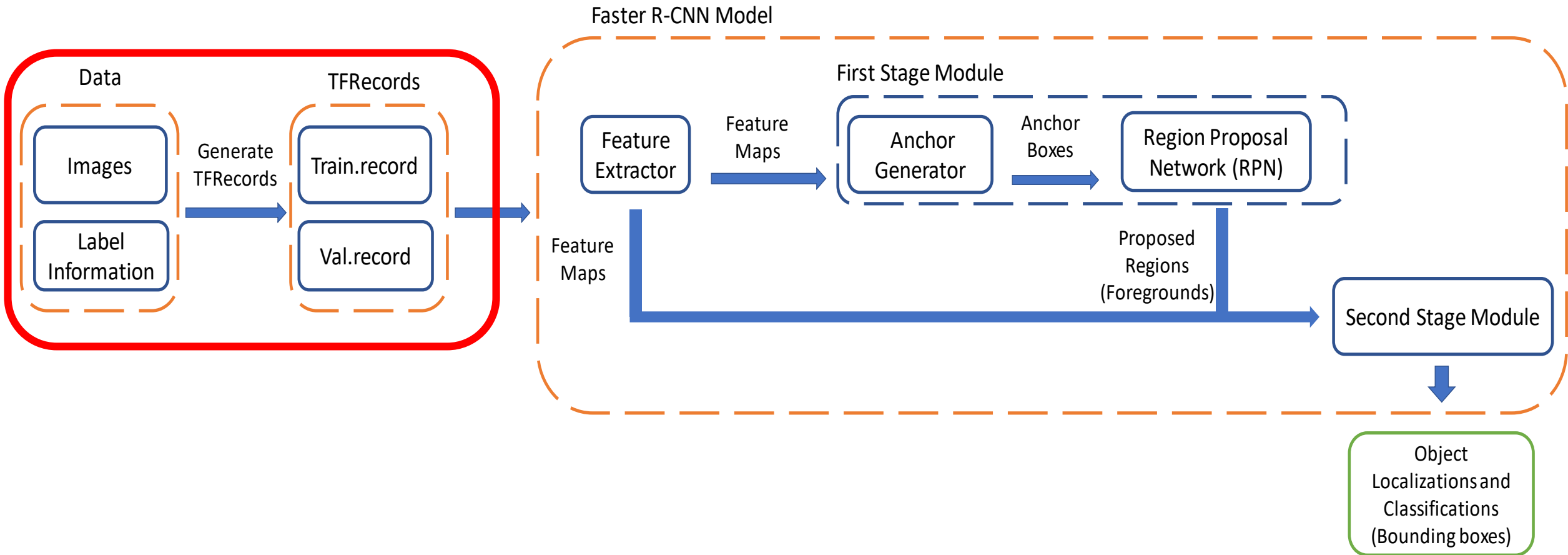
- 由Google創建，基於TensorFlow的物件偵測開源專案
- 結合了最新神經網路架構的Faster R-CNN 模型



Faster R-CNN Model



Faster R-CNN Model



Data Augmentation

- Horizontal flip
- Image scale
- Crop image
- Pad image



Original



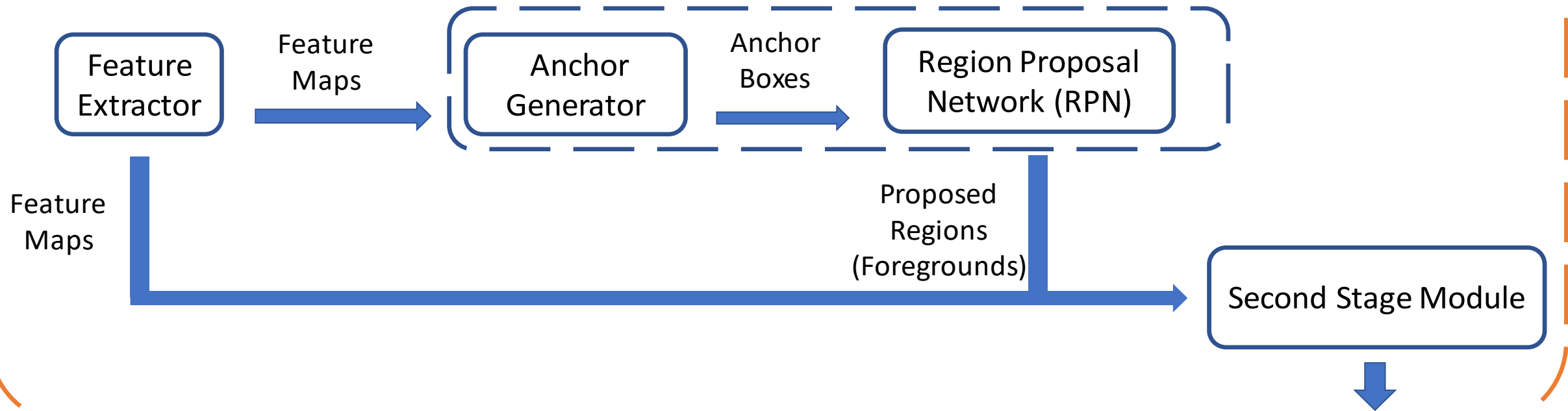
Horizontal Flip



Pad & Crop

Faster R-CNN Model

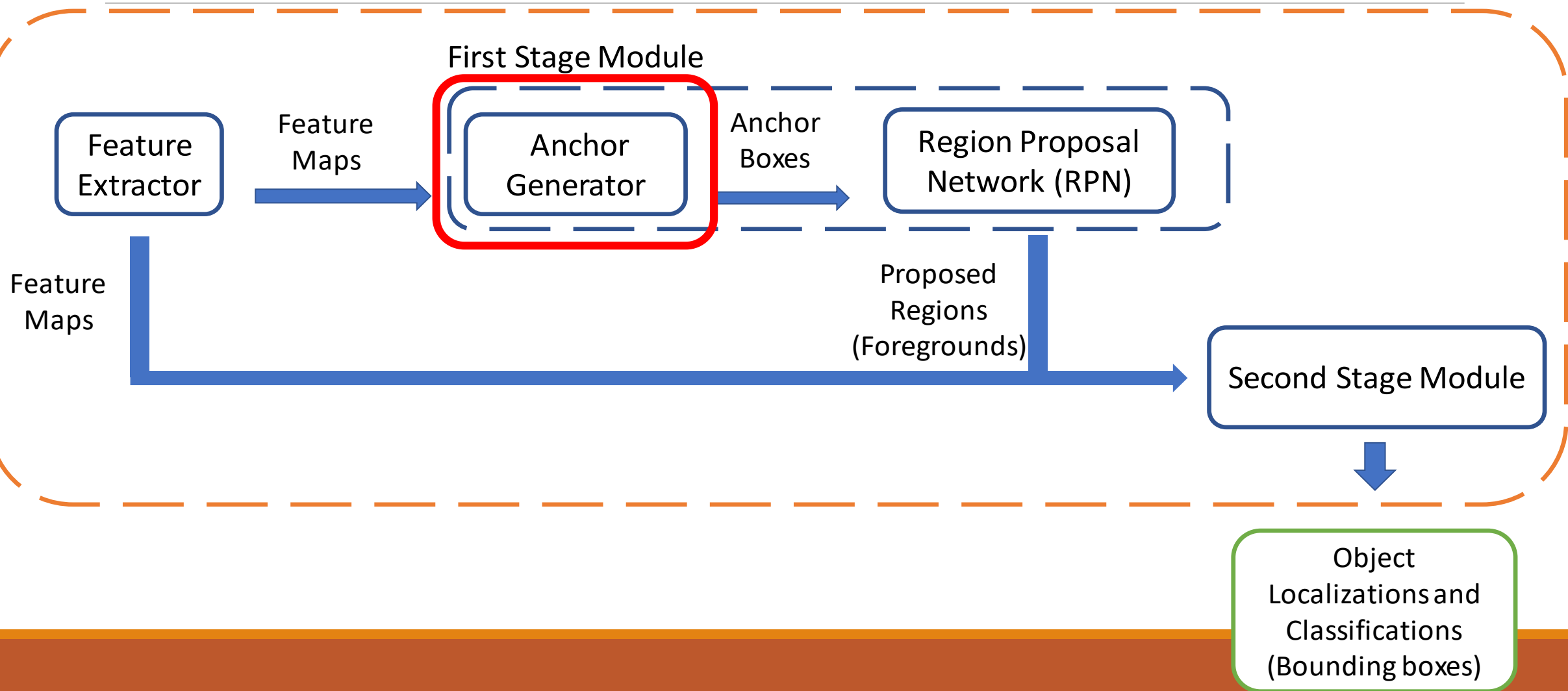
First Stage Module



Second Stage Module

Object
Localizations and
Classifications
(Bounding boxes)

Faster R-CNN Model



First Stage Module - Anchor Generator

- Hyperparameter: **scale** and **aspect ratio**

Image size = 800x600

$\text{height} = \text{base anchor size} * \text{scale} / \sqrt{\text{aspect ratio}}$

$\text{weight} = \text{base anchor size} * \text{scale} * \sqrt{\text{aspect ratio}}$

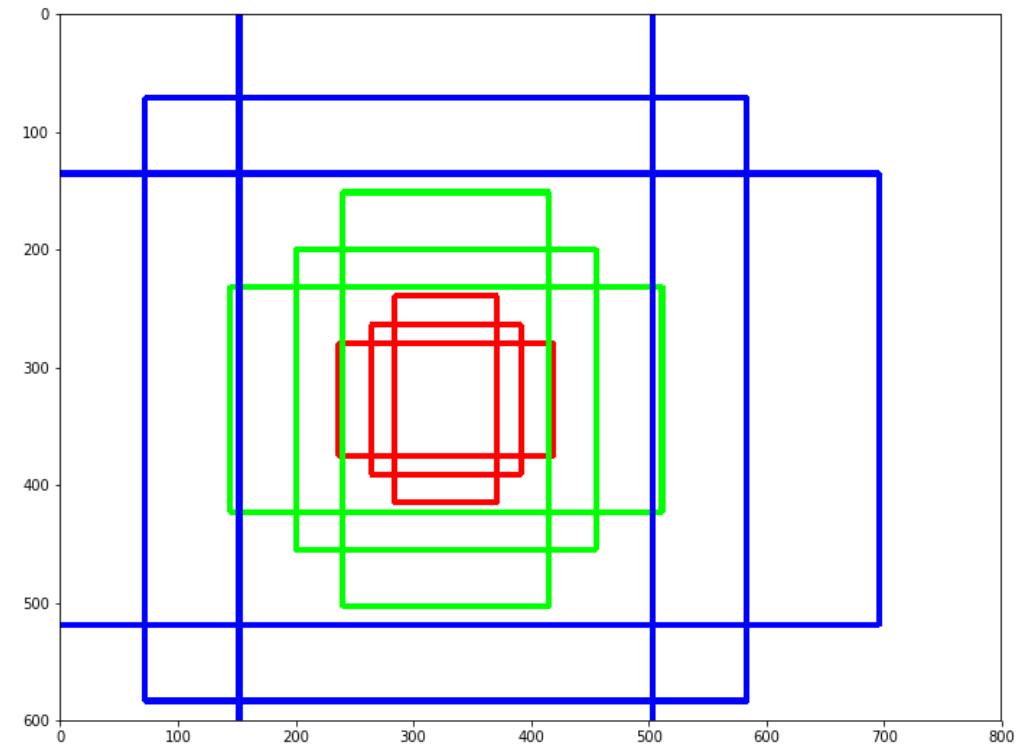
base anchor size = 256

The maximal square anchor box is

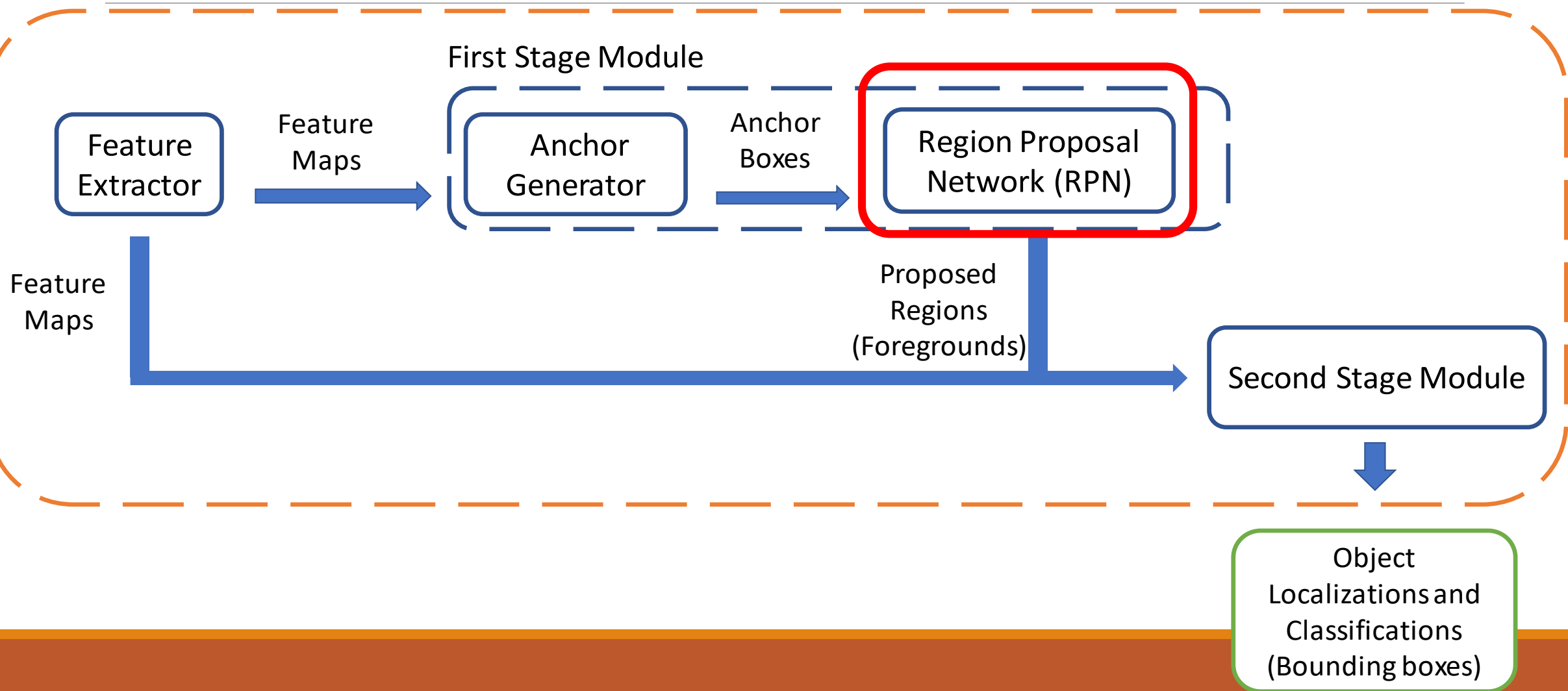
$256(\text{base anchor size}) * 2(\text{scale}) * 1(\text{aspect ratio}) = 512$

The minimal square anchor box is

$256(\text{base anchor size}) * 0.125(\text{scale}) * 1(\text{aspect ratio}) = 32$.



Faster R-CNN Model



Region Proposal Network (RPN)



Region
Proposal
Network
(RPN)



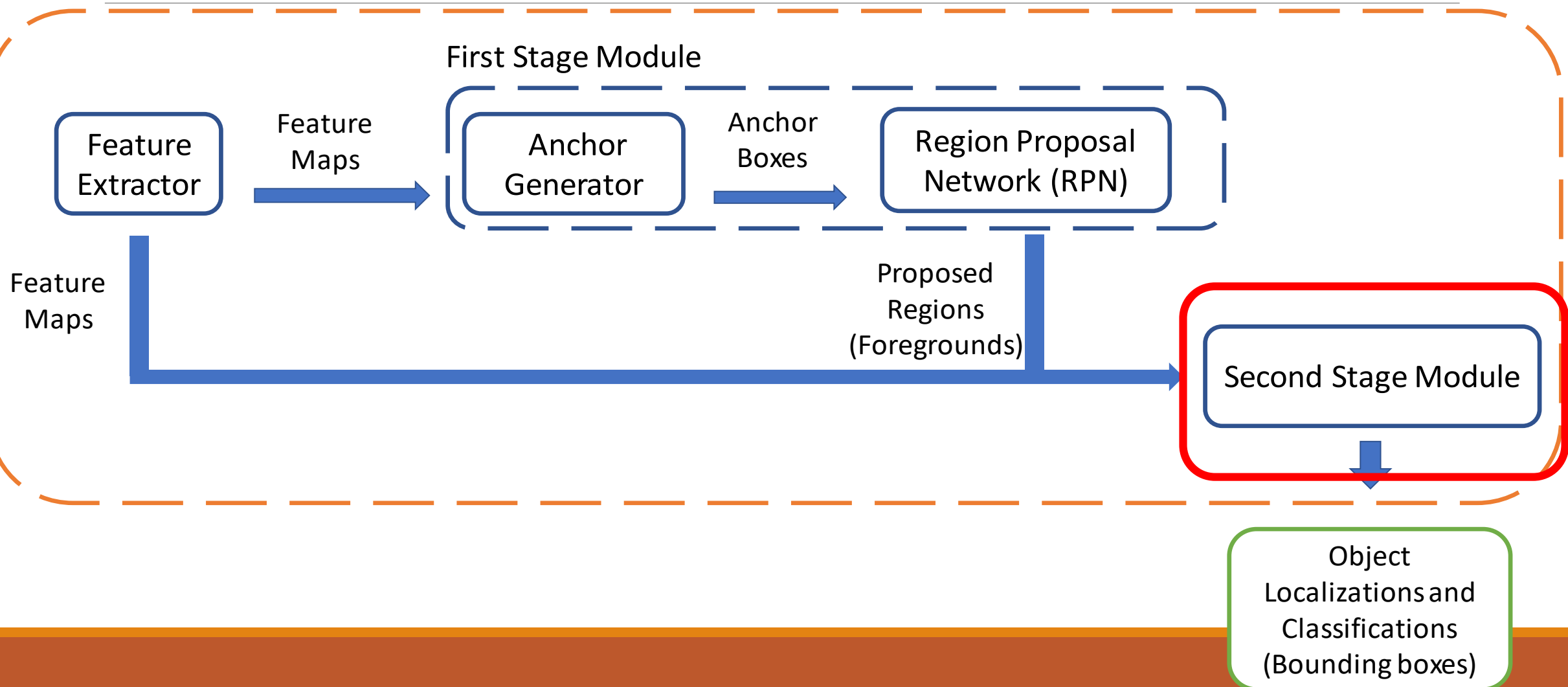
Foregrounds



Backgrounds



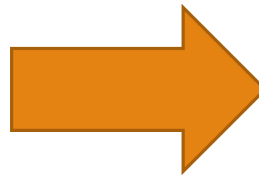
Faster R-CNN Model



Foregrounds



Detection



Bucket



Plate



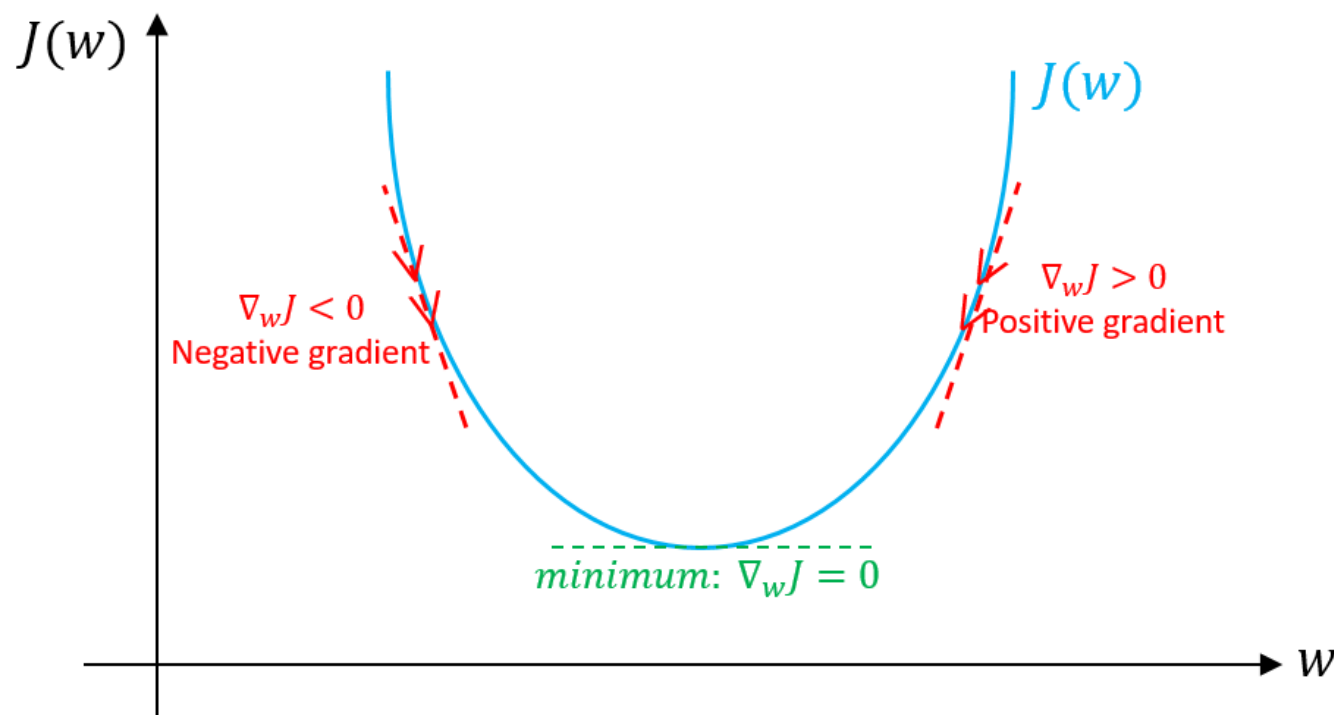
Gradient Descent Algorithm

- Stochastic Gradient Descent with momentum

$$v_t = \gamma v_{t-1} + \eta \nabla_{\theta} J(\theta)$$

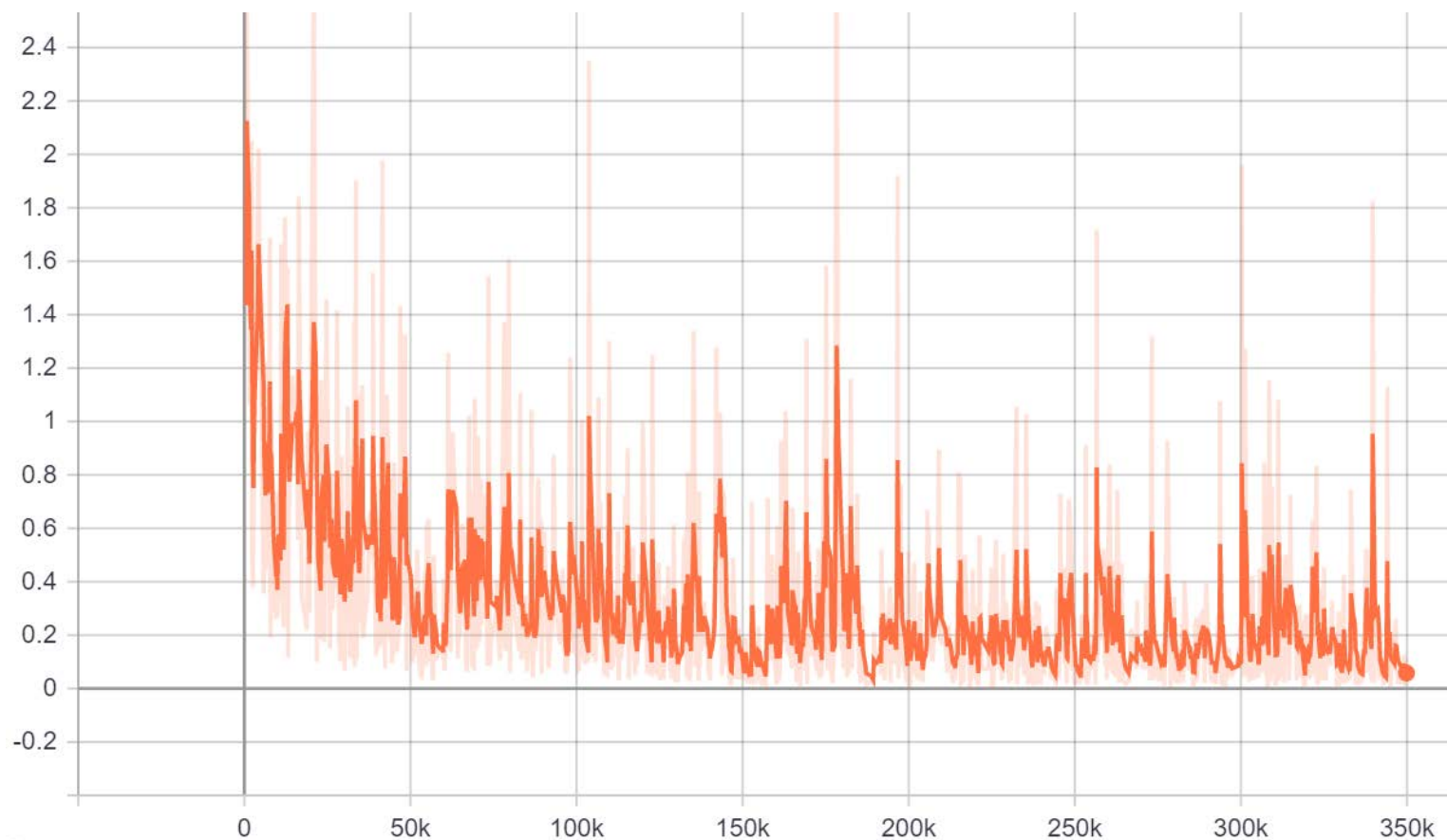
$$\theta_{t+1} = \theta_t - v_t$$

- η is a learning rate (學習率)

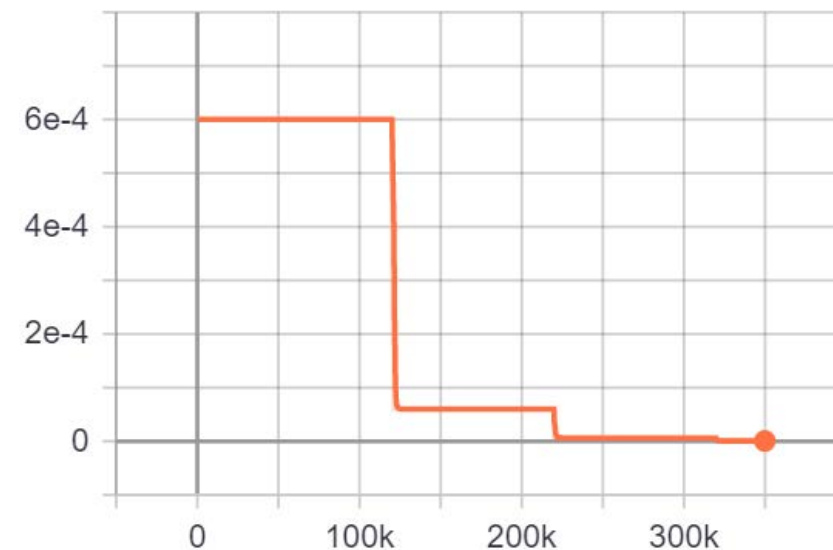


Total Loss and Learning Rate

Total Loss

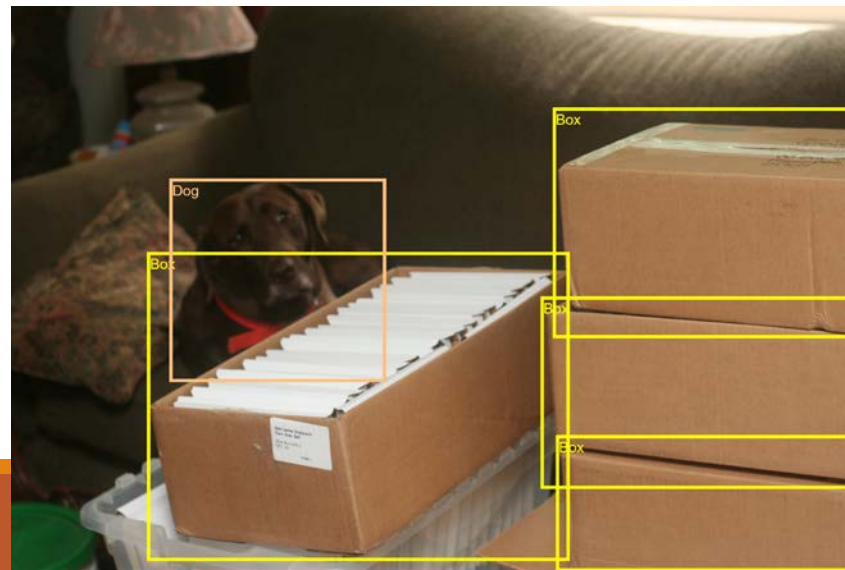


Learning Rate



Transfer Learning

- Open Image Dataset V4:
600 classes and 9,011,219 training images.
- 相關的類別: bottle, bowl, box, plastic bag, toilet, and washing machine.
- 準確度大幅提升



結果

Aldea	Score (mAP)
Public Test Dataset	0.641
Private Test Dataset	0.595



困難點

- 人造容器的形狀變化很大，
模型很難準確辨認確切是哪一類



延伸應用

- 無人機結合積水容器偵測

