學號:0716314 姓名:陳鎧勳

1. Reference:

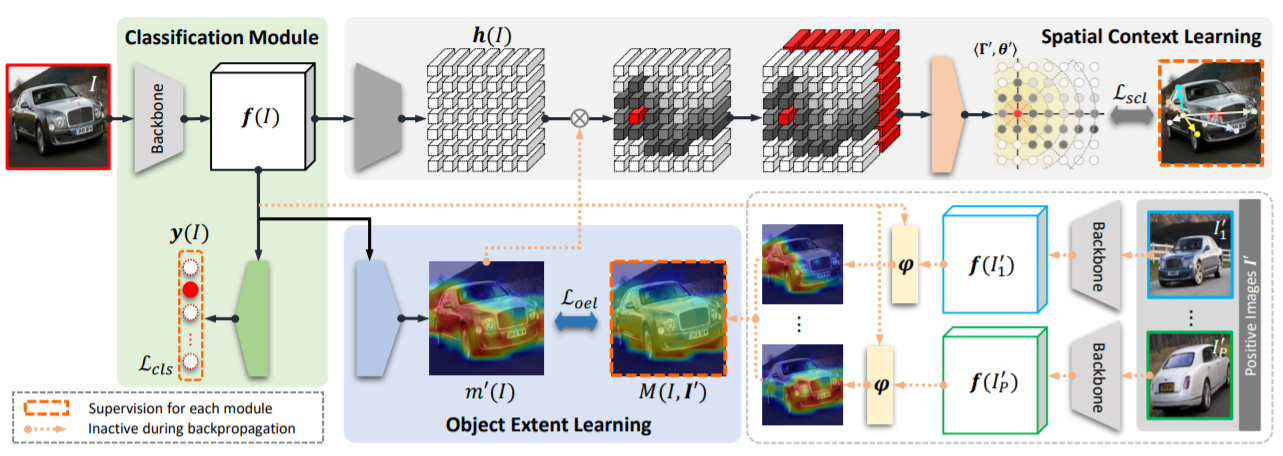
<https://github.com/JDAI-CV/LIO>

1. Brief introduction:

這個作業我所使用的技術是參考了Look-into-Object: Self-supervised Structure Modeling for Object Recognition這篇論文裡的方法，裡面所提到的OEL和SCL技術可以簡單快速地移植到任何的backbone上。

1. Methodology:
2. Data pre-process:

一開始拿到的raw data csv 是”id,label”，為了符合模型，所以我把csv格式改成”id.jpg,category(0~195)”，而我有一組dict紀錄每個category映射到哪個label方便之後轉回來。其他data augmentation會在下方Hyper-parameters介紹

1. Model architecture:

Backbone都是利用resnet-50，pipeline主要分為上面灰色的SCL和下面藍色的OEL以及右邊的綠色區塊。

OEL主要目的是分離前景和背景，方法是利用r個同種類的正樣本做類似attention的方法算出其相似度，並把結果當作ground truth和把經過1\*1卷積所得的m’(i)算，最後把m’(I)和灰色區塊的h(I)做fuse。

SCL則是因為物體的結構對於分類有重要意義，而找出物體空間上的相對關係有助於找出物體的結構，而SCL這邊的方法就是利用OEL所得到最有可能是物體的點來當作參考點，根據像素座標來預測像素與參考點的相對關係(用極座標來表示)，而ground truth就是實際算出來的極座標 ，這邊的loss為。

最後back propagation 時的，而在infer的時候，黑色和藍色區塊會被disable，只有綠色區塊有用。

1. Hyperparameter:

本次作業我的optimizer是選用SGD，而learning rate一開始是0.001而每40 epoch會變成原本的0.1倍，而momentum則是0.9，總共epoch數為200。

data augmentation則是會把每張圖片resize成512在random crop成448接著經過標準化(mean、std是整個dataset的)，之後隨機旋轉+-10度，最經過後random horizontal flip

最後，我根據不同的train/val切割法訓練了6個不同的模型，最終結果是由6個模型投票出來的

4. Summary:

這次作業學到很多，包括如何自己刻一個可以用的model還有設置hyperparameter像learning rate隨著epoch下降還有data augmentation，但只有單純的resnet50連過baseline都是一定的挑戰，所以去參閱了SOTA的文章，才順利的把準確度提高到超過baseline。