

ML 2019 SPRING FINAL PROPOSAL

隊名：只有一個人會睡覺

隊員：b05901170 陳柏志 b06901096 黃昱翰 b06901004 劉穎立

所選題目：第三題(Medical Image Detection)

1. problem study

(1) RetinaNet

RetinaNet RetinaNet 是 one stage detector 為了避免 foreground-background class imbalance，使用了 FPN 架構及 Focal Loss 所設計出的 model。因為一張圖的物體相較於背景的比例相當低，所以 one stage detector 看到的大多為背景，而導致有物體的 bbox 也被一同歸類為背景而分辨不出來，所以為了解決這個問題，並且避免運算速度和 RCNN 一樣慢，所以採用新的 loss 算法：Focal Loss，這個 Loss 會把判斷正確的 box 的 Loss 權重調高，藉以讓訓練時較少的 object 對於 loss 會有比較大的影響，就可避免 background 太多的問題了，同時他是 classification 和 box regressor 平行 train。

(2) RCNN

RCNN 也是常見的一個圖像切割辨識的作法。傳統的 RCNN 是先用一些 image segmentation 法將圖片切塊，再過一些 pretrain 好的 model，便是該區塊是什麼物體。後來為了增加算速度等目的，又有 fast RCNN、faster RCNN 等類似概念的模型出現，本組目前是嘗試用 RCNN 的概念，加入一些我們的想法，將來也會嘗試用 fast RCNN 和 faster RCNN，並觀察比較他們的表現，尋找最好的方法。

2. proposed method

(1) 增加 FPN 層數

我們有去看 rsna pneumonia detection challenge 的 kaggle 討論區，發現第一名和第二名的作法都是用 RetinaNet，然後他們的作法除了 ensemble 不同的 backbone 和其他的 net 以外，還有增加 fpn 層數來增強對於比較小的 box 的 detection，還有 resize image 讓他有適量的 feature，然後我們 train 的 epoch 數太少，因為時間有限，所以我們應該會多 train 幾個 model，同時應該會想辦法讓 data augmentation 更有效。

(2) 先用 Selective Search 找出可能的框框，在用 CNN 決定他是腫瘤的機率

- a. 我們將 training data 分成兩部分，將原本的 1024*1024 的大圖，依照 csv 檔中的框框分成框框內跟框框外，分別隨機切出好幾塊 128*128 的小圖，並標上 label(框框內是 1、框框外是 0)，拿這些小圖跟 label 去 train 一個 CNN 的 classification model。

- b. 將 test 的原圖用類似 convolution 的方法，掃出很多 128×128 的小圖，拿去過 a. 中 train 好的 CNN，並將結果做成一個由 1, 0 組成的 map
- c. 先將 b. 中作出的 map 用 Selective Search 切出可能有東西的框框，再去算此範圍中單位面積的 1 的數量，作為他的分數，最後選擇左右半平面分數最高，且大於我們從 training data 中觀察而得的閾值的框框作為結果。

(3) 資料預處理

我們未來預計嘗試增加資料對比度、事先將肺部切割出來等方法，來讓 model 夠容易 train 起來。