

# 影像處理概論

## 作業二：影像處理與顯著性偵測實習

### 原理介紹

#### Deep learning

深度學習是機器學習的一種，以深度神經網路為架構，對於資料的特徵進行學習。

神經網路將數以千計的神經元串連起來，形成一個網路。在此網路中，每一層的各個神經元都有一或多個輸入值，透過不同的權重和簡單的函式進行運算，產生輸出值，並連接到其他下一層的神經元當作輸入，模擬神經元傳遞訊號的架構。當神經網路的層數甚多時，即可稱為深度神經網路。

當深度神經網路的一個輸入，在經過層層運算之後，所產生的結論與實際答案有誤差時，利用此誤差來回朔並修正各個神經元中的權重。透過各種輸入，不斷的重複上述動作，使得誤差逐漸變小，即為深度神經網路學習。

#### Salient object detection

顯著性偵測的主要目的在於將影像中最具吸引力或人類感興趣的部分標示出來。影像經過深度神經網路後，提取一系列不同程度的感知特徵，其中越高層的具有更多語意訊息，越低層的具有更多細節訊息，因此可以透過高層的語意理解掌握顯著物在全局的位置，搭配低層的細節特徵來偵測影像的顯著性。

#### Super pixel and super pixel segmentation

超像素的主要目的在於將具有相似特性的像素合併起來，形成一個具有代表性的大像素。超像素將像素級的影像，透過基於圖形或梯度上升的算法，劃分成區域級的影像。

基於圖形的超像素生成方法將每個像素視為圖中的節點，兩個節點之間的邊權重與相鄰像素之間的相似性成比例，並透過最小化圖中定義的成本函數來形成超像素。

基於梯度上升的方法從粗略的像素初始聚類開始，梯度上升法迭代的修改聚類，直到滿足一些收斂標準以形成超像素。

#### SLIC Super pixel segmentation

SLIC 利用聚類算法，在一開始時，每一個聚類的中心平均的分佈在原始影像上，之後每進行一次迭代，種子像素合併周圍的像素，逐漸形成超像素。

## 實作方法

### Labelme

1. 在終端機輸入 labelme 開啟 Labelme 圖形介面。
2. 打開圖片。
3. 點擊 Create Polygons。
4. 用滑鼠標記 Salient objects。
5. 儲存標籤。
6. 執行提供的程式產生所需檔案。

### 深度學習網路

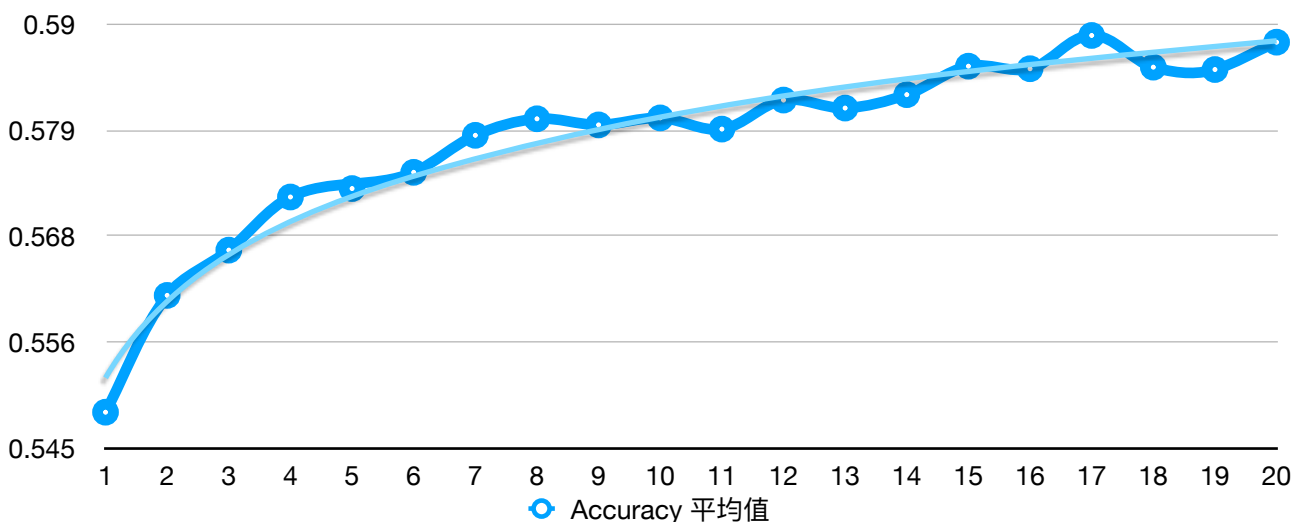
1. 使用 Matlab 為 Training data 和 Testing data 產生 all\_Q.mat 以及 all\_superpixel\_label.mat。
2. 下載並修改 train\_superCNN.py 中的網路架構及路徑後執行，進行訓練。
3. 下載 eval.py 並修改其中的路徑，儲存成 test\_superCNN.py 後執行，進行測試。

## 深度學習模型訓練與評估結果

### Training

#### Accuracy 表現結果

將全部 Training 結果的 Accuracy 做平均，產生 1 至 20 個 Epochs 的 Accuracy 表現，以其平均 Accuracy 做圖，其結果如下圖。



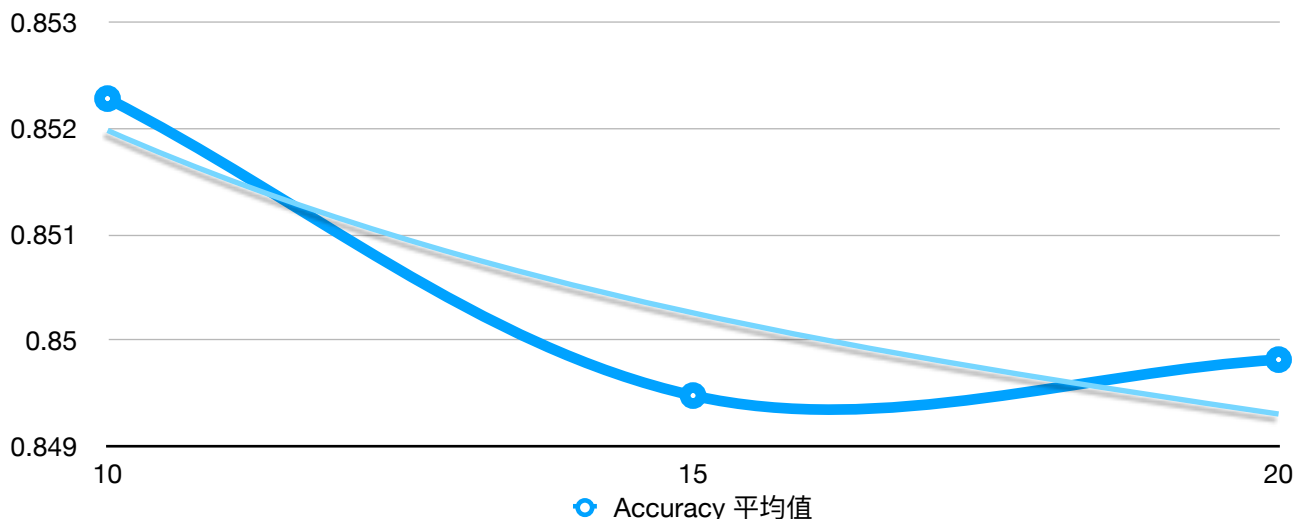
## 觀察心得

圖中的淺藍色曲線為對數趨勢線，可以看出在 10 個 Epochs 之前，訓練出來的 Accuracy 有明顯的在增加，之後的增加幅度開始慢慢趨緩，我認為這表示模型能夠再從 Testing data 中再學習到的內容已經有限，因此未來在訓練模型的時候，不能一味地追求更多的 Epochs，也需要去思考影響結果的原因，才能避免訓練結果的 Accuracy 達到瓶頸。

## Testing

### Accuracy 表現結果

將 Testing 結果的 Accuracy 做平均，分別用 10、15、20 個 Epochs 以其平均 Accuracy 做圖，其結果如下圖。



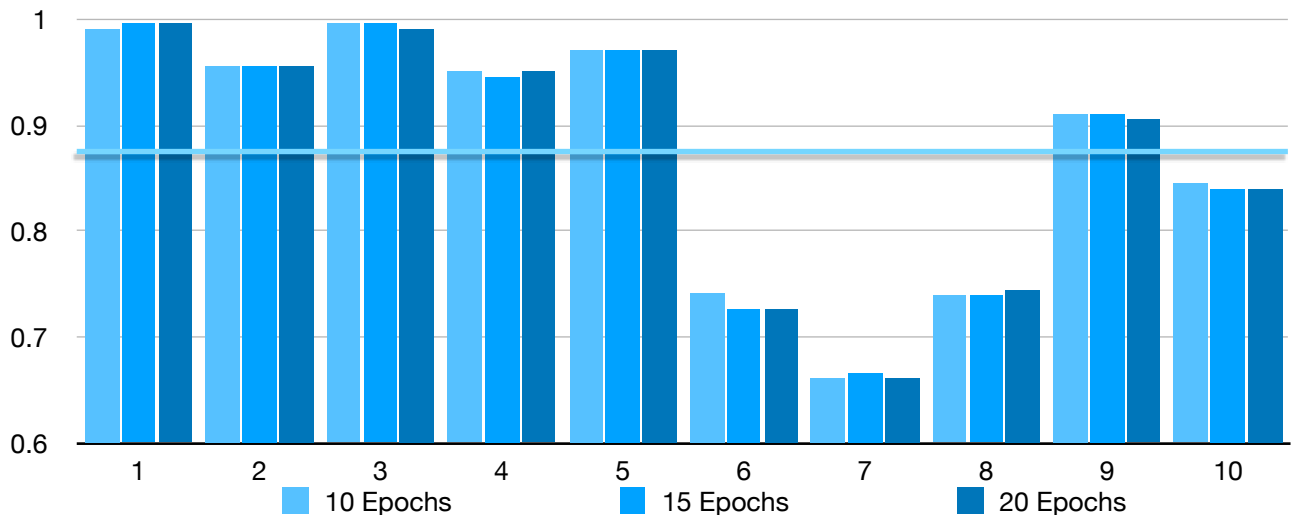
## 觀察心得

三者的表現都近似 0.85，其並無明顯的差異。圖中的淺藍色曲線為對數趨勢線，可以看出在 10 個 Epochs 時，其 Accuracy 是最高的，之後開始下降，因此可以很明顯地看出，訓練越多個 Epochs，其結果不一定會更好，如果有 Overfitting 的狀況發生，結果甚至可能會越來越差，也就是說，如果想要訓練出一個好的模型，絕不能只增加 Epochs，而是一定要理解模型背後的原理之後作出適當的調整，才有可能有效增進模型的表現。

除此之外，在相同 Epochs 的條件下，Testing 結果的平均 Accuracy 要明顯高於在 Training 時的平均 Accuracy，我並不是相當能理解這個性質，但我認為這可能是由於 Testing data 均為在校園內所拍攝，其相似性相當高的緣故所致。

另外，有些影像的測試結果準確率低至僅 0.2 左右，推測部分原因是由於 Salient object 的標示過於粗糙，或影像不夠清晰等等所造成的影響。

下圖為 Testing data 中的 10 張圖片，經由 10、15、20 個 Epochs 的訓練後測試之顯著性判斷結果，圖中淺藍色直線為其總平均，附上作為參考。



## 分析與討論

在作業二當中，經過從 Salient object 的標注，到產生 Super pixel 以及訓練深度神經網路的過程之後，我更加的瞭解影像處理與顯著性偵測的原理與方法，並學習建構深度學習之訓練與評估模型。同時，在作業二當中所使用到的 Keras 以及 Tensorflow 都是很好的工具，對於我未來在進行其他深度學習的實作時，相信可以很有效的派上用場。此外，將實驗結果整理、列表並使人易於閱讀也是重要的一環。

梯度消失問題：在神經網路中，每傳遞一層，梯度衰減。層數一多，梯度指數衰減後深層基本上接受不到有效的訓練訊號。但梯度消失問題可使用 ReLu 從根本上解決。

卷積神經網路的優點包括：共享的權重減少了訓練所需的引數，因此相同的權重可使濾波器不受訊號位置的影響來檢測訊號，使得模型的泛化能力更強；池化運算透過降低網路的空間解析度，進而消除訊號的微小偏移和扭曲，因此對輸入資料的平移不變性要求不高。

其缺點則是隨著深度的增加，模型的誤差從輸出層開始呈指數衰減，因此容易出現梯度消失問題。

Super pixel 的優點在於其大幅降低了影像的維度，同時剔除一些異常像素點，使得計算速度得已大大提升。

## 字數統計

2375 個字元（不含空格）

2505 個字元（含空格）

1754 個字