

以深度學習方法進行人體骨架之偵測

系所／電子工程學系

指導老師／黃炳森

組員／李承翰、梁生瑀

隨著電腦和手機的快速發展，人工智慧也逐漸受到重視。其中，以深度學習為大宗。本次專題透過深度學習，搭配桌上型電腦和 Nvidia GTX1080 Ti 圖形顯示卡。並使用 Python 程式語言，讓電腦先使用 MSCOCO 的數據集 (dataset) 進行訓練，之後進行尋找關節點和關節點連線的骨架偵測，以呈現圖片不同人物的骨架。

這裡我們使用 bottom-up 結構的“分數-關聯表示法”(representation of association scores via Part Affinity Fields)，同時進行關節點偵測和關聯性偵測。

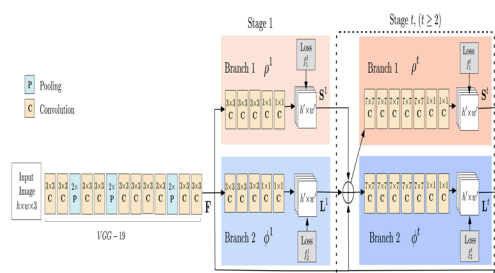


圖 1：神經網路架構圖

測量數據方面，使用已標注關節點的數據集 (MSCOCO)，構建三個訓練用的 data：歸一化的圖片、S_Label(檢測關節點的 Label)、L_Label(關節點之間關聯性的 Label)。將圖片送入預訓練

過的 VGG-19 的前 10 層得到一組 F，將 F 輸入到 STAGE 中，生成兩部分 output($S_j^t(p)$, $L_j^t(p)$)，對 output 分別與 $S_Label(S_j^*)$ 、 $L_Label(L_j^*)$ 相減生成兩部分，分別做 L2 norm 後分別得到 $S\ loss(f_{S^t})$ 和 $L\ loss(f_{L^t})$ ，取 summation 即為總 $loss(f)$ 。並使用 Part Confidence Maps 與 Part Affinity Fields(PAFs) 找出圖片中的關節點和兩個關節點之間的連線。

訓練部分，使用到如下訓練技術：

1. Transform Learning 使用預分類訓練完成的模型，選取其特徵提取層，降低複雜任務的訓練難度。
2. SGD(Stochastic Gradient Descent) With Momentum：對搖擺幅度較大的 feature 進行抑制。
3. Weight decay: 減小過擬合。
4. 在 loss function 內對 output 做類 dropout 提升訓練效果。



圖 2：結果圖

實驗證明，通過 bottom-up 結構同時進行關節點偵測和關聯性偵測，能夠以較少的計算時間，得到較佳的結果。