

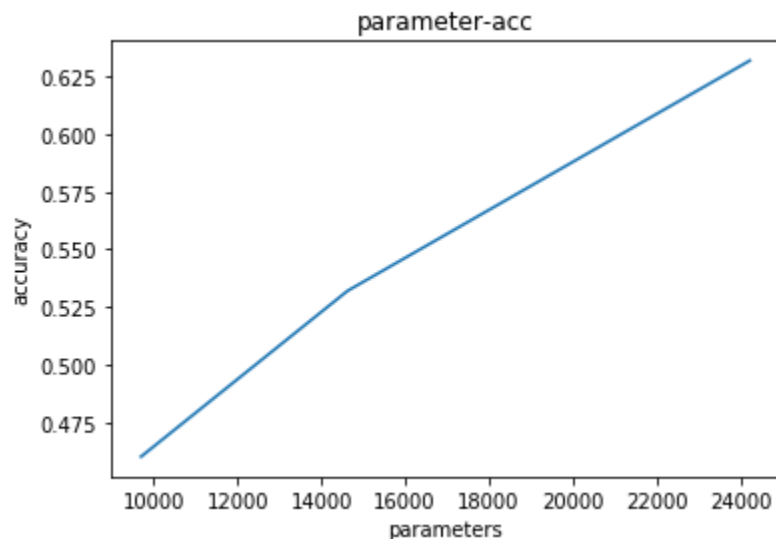
學號：B06507007 系級：材料二 姓名：王致雄

1. 請比較你本次作業的架構，參數量、結果和原 **HW3** 作業架構、參數量、結果做比較。(1%)

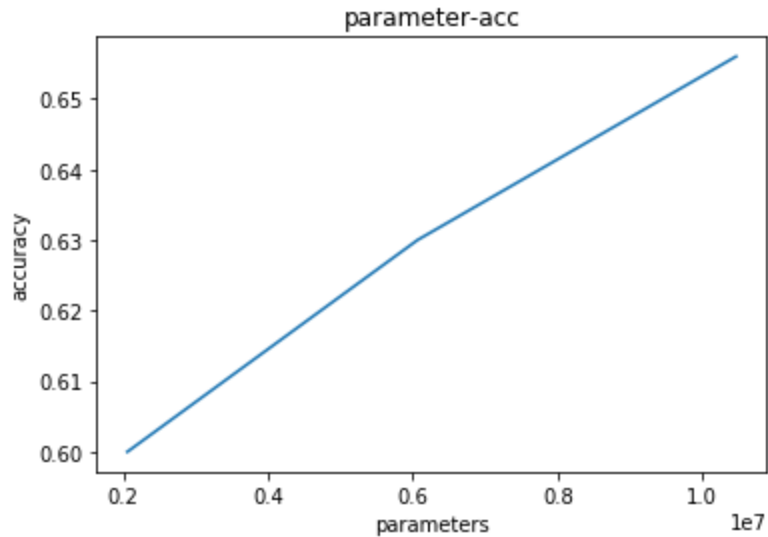
HW3: 共四層 convolution，採用 CCP(C 為 convolution(3x3),P 為 max pooling)，並且每一層都有作 normalization，以及兩層的 feed forward network；上述每一層都有做 dropout(drop rate=0.3)。單一模型的 validation set 準確率多在 62~62.5%收斂，而有些則到達 64~65%。此模型有 10464455 個可訓練的參數。

本次:使用 MobileNet 架構，一共 7 層，將 channel 逐漸放大，最後再用 feed forward network 產生 7 個 output；上述每一層都有做 dropout(drop rate=0.3)。最終的結果為 private:0.63193 public:0.63388。此模型有 24231 個參數，大小為 143kb。

2. 請使用 **MobileNet** 的架構，畫出參數量-acc 的散布圖（橫軸為參數量，縱軸為 **accuracy**，且至少 3 個點，參數量選擇時儘量不要離的太近，結果選擇只要大致收斂，不用 **train** 到最好沒關係。）(1%)



3. 請使用一般 **CNN** 的架構，畫出參數量-acc 的散布圖（橫軸為參數量，縱軸為 **accuracy**，且至少 3 個點，參數量選擇時儘量不要離的太近，結果選擇只要大致收斂，不用 **train** 到最好沒關係。）(1%)



4. 請你比較題 2 和題 3 的結果，並請針對當參數量相當少的時候，如果兩者參數量相當，兩者的差異，以及你認為為什麼會造成這個原因。(2%)
5. 我將我 hw3 的 CCP 結構只保留一層，**feed forward network** 也只保留一層的結果是可以把 **CNN** 的參數量壓到 3 萬左右，與 **MobileNet** 約略相等。而結果是約在 50%收斂(validation)。我認為的原因是 **MobileNet** 的架構可以更有效的利用/共用參數，使 **MobileNet** 即使使用較少參數也能達成較好結果。另外還有一般人為模型的深度的影響會比廣度來的大，這個 **CNN** 僅有 2 層，而 **MobileNet** 有 7 層，雖然兩者參數量相當，但是較深的模型表現較好是可以預期的。