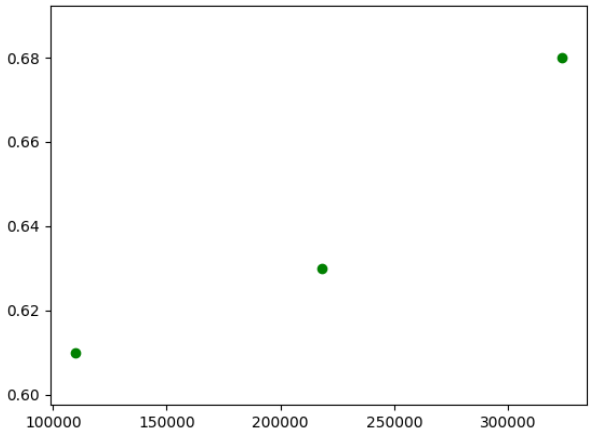
學號：B05901170 系級：電機三 姓名：陳柏志

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%)

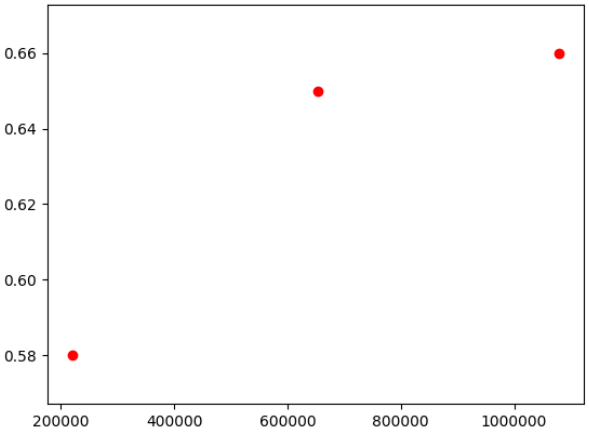
	HW3 Model	HW8 Model
模型架構	========== Conv2d-1 [-1, 64, 48, 48] BatchNorm2d-2 [-1, 64, 48, 48] 128 LeakyReLU-3 [-1, 64, 48, 48] 0 Conv2d-4 [-1, 64, 48, 48] 36,928 BatchNorm2d-5 [-1, 64, 48, 48] 128 LeakyReLU-6 [-1, 64, 48, 48] 0 MaxPool2d-7 [-1, 64, 24, 24] 0 Conv2d-8 [-1, 64, 24, 24] 36,928 BatchNorm2d-9 [-1, 64, 24, 24] 128 LeakyReLU-10 [-1, 64, 24, 24] 0 Conv2d-11 [-1, 64, 24, 24] 36,928 BatchNorm2d-12 [-1, 64, 24, 24] 128 LeakyReLU-13 [-1, 64, 24, 24] 0 MaxPool2d-14 [-1, 64, 12, 12] 0 Conv2d-15 [-1, 128, 12, 12] 73,856 BatchNorm2d-16 [-1, 128, 12, 12] 256 LeakyReLU-17 [-1, 128, 12, 12] 0 Conv2d-18 [-1, 128, 12, 12] 147,584 BatchNorm2d-19 [-1, 128, 12, 12] 256 LeakyReLU-20 [-1, 128, 12, 12] 0 MaxPool2d-21 [-1, 128, 6, 6] 0 Linear-22 [-1, 512] 2,359,808 LeakyReLU-23 [-1, 512] 0 Dropout-24 [-1, 512] 0 Linear-25 [-1, 512] 262,656 LeakyReLU-26 [-1, 512] 0 Dropout-27 [-1, 512] 0 Linear-28 [-1, 7] 3,591 ========== Total params: 2,959,943	========== Conv2d-1 [-1, 16, 24, 24] 144 ReLU-2 [-1, 16, 24, 24] 0 BatchNorm2d-3 [-1, 16, 24, 24] 32 Dropout-4 [-1, 16, 24, 24] 0 Conv2d-5 [-1, 16, 24, 24] 144 ReLU-6 [-1, 16, 24, 24] 0 BatchNorm2d-7 [-1, 16, 24, 24] 32 Conv2d-8 [-1, 32, 24, 24] 512 ReLU-9 [-1, 32, 24, 24] 0 BatchNorm2d-10 [-1, 32, 24, 24] 64 Dropout-11 [-1, 32, 24, 24] 0 Conv2d-12 [-1, 32, 24, 24] 288 ReLU-13 [-1, 32, 24, 24] 0 BatchNorm2d-14 [-1, 32, 24, 24] 64 Conv2d-15 [-1, 32, 24, 24] 1,024 ReLU-16 [-1, 32, 24, 24] 0 BatchNorm2d-17 [-1, 32, 24, 24] 64 Dropout-18 [-1, 32, 24, 24] 0 Conv2d-19 [-1, 32, 12, 12] 288 ReLU-20 [-1, 32, 12, 12] 0 BatchNorm2d-21 [-1, 32, 12, 12] 64 Conv2d-22 [-1, 64, 12, 12] 2,048 ReLU-23 [-1, 64, 12, 12] 0 BatchNorm2d-24 [-1, 64, 12, 12] 128 Dropout-25 [-1, 64, 12, 12] 0 Conv2d-26 [-1, 64, 12, 12] 576 ReLU-27 [-1, 64, 12, 12] 0 BatchNorm2d-28 [-1, 64, 12, 12] 128 Conv2d-29 [-1, 64, 12, 12] 4,096 ReLU-30 [-1, 64, 12, 12] 0 BatchNorm2d-31 [-1, 64, 12, 12] 128 Dropout-32 [-1, 64, 12, 12] 0 Conv2d-33 [-1, 64, 12, 12] 576 ReLU-34 [-1, 64, 12, 12] 0 BatchNorm2d-35 [-1, 64, 12, 12] 128 Conv2d-36 [-1, 128, 12, 12] 8,192 ReLU-37 [-1, 128, 12, 12] 0 BatchNorm2d-38 [-1, 128, 12, 12] 256 Dropout-39 [-1, 128, 12, 12] 0 Conv2d-40 [-1, 128, 12, 12] 1,152

		ReLU-41 [-1, 128, 12, 12] 0 BatchNorm2d-42 [-1, 128, 12, 12] 256 Conv2d-43 [-1, 128, 12, 12] 16,384 ReLU-44 [-1, 128, 12, 12] 0 BatchNorm2d-45 [-1, 128, 12, 12] 256 Dropout-46 [-1, 128, 12, 12] 0 Conv2d-47 [-1, 128, 6, 6] 1,152 ReLU-48 [-1, 128, 6, 6] 0 BatchNorm2d-49 [-1, 128, 6, 6] 256 Conv2d-50 [-1, 128, 6, 6] 16,384 ReLU-51 [-1, 128, 6, 6] 0 BatchNorm2d-52 [-1, 128, 6, 6] 256 Dropout-53 [-1, 128, 6, 6] 0 Conv2d-54 [-1, 128, 6, 6] 1,152 ReLU-55 [-1, 128, 6, 6] 0 BatchNorm2d-56 [-1, 128, 6, 6] 256 Conv2d-57 [-1, 128, 6, 6] 16,384 ReLU-58 [-1, 128, 6, 6] 0 BatchNorm2d-59 [-1, 128, 6, 6] 256 Dropout-60 [-1, 128, 6, 6] 0 Conv2d-61 [-1, 128, 6, 6] 1,152 ReLU-62 [-1, 128, 6, 6] 0 BatchNorm2d-63 [-1, 128, 6, 6] 256 Conv2d-64 [-1, 256, 6, 6] 32,768 ReLU-65 [-1, 256, 6, 6] 0 BatchNorm2d-66 [-1, 256, 6, 6] 512 Dropout-67 [-1, 256, 6, 6] 0 AvgPool2d-68 [-1, 256, 1, 1] 0 Linear-69 [-1, 7] 1,799 Dropout-70 [-1, 7] 0 ========== Total params: 109,607
訓練參數	Batch size = 256, Epoch = 50, Shuffle = True, No Data Augmentation, Optimizer = Adam with learning rate = 0.001	Batch size = 256, Epoch = 50, Shuffle = True, No Data Augmentation, Optimizer = Adam with learning rate = 0.001
準確率	Public Score: 0.72220 Private Score: 0.71412	Public Score: 0.65059 Private Score: 0.64558

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%)

   在參數較多的時候兩者的表現其實沒有差太多，但在參數少時，Mobile Net 明顯較優於普通 CNN Model。我認為這應該是因為 Mobile Net 中的 pointwise convolution 能以較少的參數量達到較好的模型表現。