

人工智慧晶片設計與應用

AI-ON-CHIP FOR MACHINE LEARNING AND INFERENCE

Lab 1

E14096724

鄭喆嚴

1. 假設有一個 size 為 $32 \times 32 \times 1$ input feature map，若要用 convolution 來產生一個 $28 \times 28 \times 1$ output feature map，請比較使用以下兩種 Kernel size 計算時，所需的 parameter 數量 和 MAC 數量，須包含計算過程才予以計分。
(16 points)(p.s no padding & stride=1)

[5x5x1 CONV] (p.s 一個 convolution layer)

$$\text{parameters} = 1 * 5 * 5 * 1 + 1 = 26$$

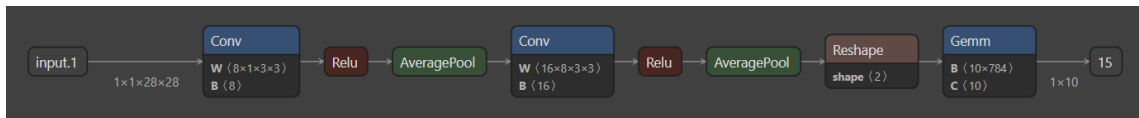
$$\text{MACs} = 1 * 5 * 5 * 1 * 28 * 28 = 19600$$

[3x3x1 CONV] + [3x3x1 CONV] (p.s 兩個 convolution layer 串聯)

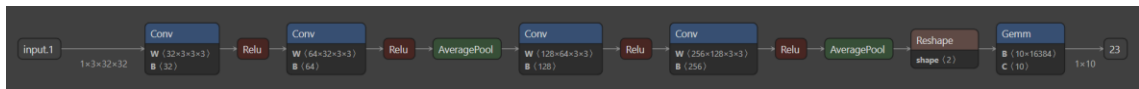
$$\text{parameters} = 1 * 3 * 3 * 1 + 1 + 1 * 3 * 3 * 1 + 1 = 20$$

$$\text{MACs} = 1 * 3 * 3 * 1 * 30 * 30 + 1 * 3 * 3 * 1 * 28 * 28 = 15156$$

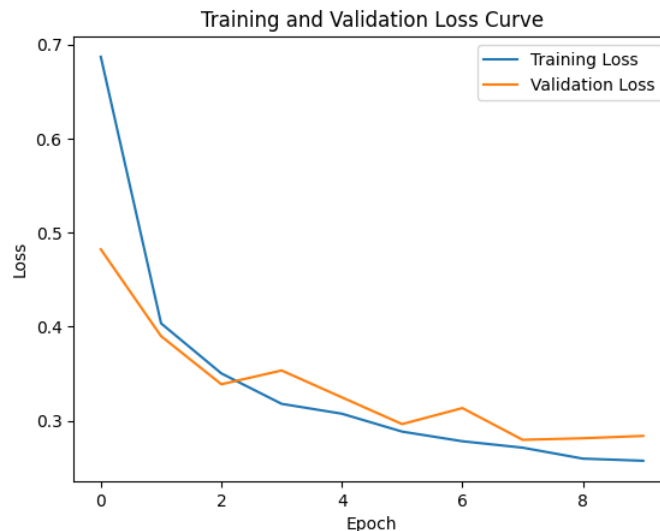
2. 藉由 Netron 將你所設計的兩個模型視覺化並截圖 (5 points)
model (fashion mnist):



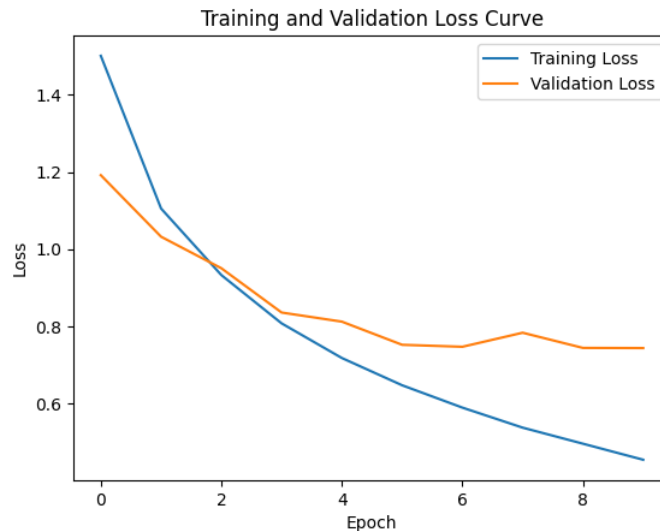
model (cifar10):



3. 畫出 train/val loss curve，並判斷是否 overfitting? (10 points)



上圖是 model (fashion mnist)訓練時的 train/val loss curve，可以看出訓練和驗證損失之間的差異並沒有很大，損失曲線的走勢是訓練損失逐漸下降，驗證損失也跟隨訓練損失下降，所以可以判斷沒有 overfitting。



上圖是 model (cifar10) 訓練時的 train/val loss curve，可以看出訓練和驗證損失之間的差異越來越大，損失曲線的走勢是訓練損失逐漸下降，驗證損失一開始也跟隨訓練損失下降，後來越來越平緩，所以可以判斷可能有一點 overfitting。

4. 你覺得這個 lab 有什麼可以改進的地方以及你的心得? (4 points 認真表達心得一律滿分)

我覺得這次的 lab 已經設計的很棒，如果可以更加詳細的教學在設計模型時的一些方向和設計的優缺點，我覺得這樣調整會有助於學生的實作。

這次的 lab 讓我更深入了解了有關神經網絡模型的訓練和優化，通過設計自己的神經網絡模型並且調整各種參數來達到要求的正確率，在做這個 lab 的過程中讓我感覺到很有趣，並且有學習到調整模型或是調整各種參數會對整體造成什麼影響。

5. 實作上你做了甚麼調整 (learning rate, image augmentation 等等) 維持精準度, 減少 MAC, parameters, FLOP 和避免模型 overfitting? 加不同的 data augmentation 會在 testing data 精準度上面有甚麼影響? 請具體以文字和數據描述。(10 points)

我調整模型的卷積層和全連階層的數量和通道數來提升精準度，並且在達到目標的精確度後漸漸地減少 MAC 和 parameters；我調整 learning rate 來維持精準度同時避免模型 overfitting，我先從較小的 learning rate 漸漸增大來調整到較高的精準度，並且同時注意 train/val loss curve 來避免模型 overfitting。

通常在加上 data augmentation 會提升 testing data 精準度，像是旋轉、縮放、平移等簡單的幾何變換通常會提升模型的性能和精準度，但是如果過度或不適當的 data augmentation 反而可能會讓模型的性能和精準度下降。