AI on Chip 2024

LAB I AI Model

REPORT

Student name: \_\_胡家豪\_\_\_\_\_\_\_\_\_

Student ID: \_\_N26122246\_\_\_\_\_\_\_\_

目錄

[一. Question List 3](#_Toc161219679)

[1. Caculation 3](#_Toc161219680)

[2. Model visualization 3](#_Toc161219681)

[a. Fashion-MNIST 3](#_Toc161219682)

[b. Cifar10 3](#_Toc161219683)

[3. Loss curve plot and overfitting judgement 3](#_Toc161219684)

[a. Fashion-MNIST 3](#_Toc161219685)

[b. Cifar10 3](#_Toc161219686)

[4. My opinion 4](#_Toc161219687)

[5. Model modify 4](#_Toc161219688)

1. Question List
2. Caculation
3. 5\*5 Convolution：

Parameter：用 input channel\*output channel\*kernal size得到1\*1\*5\*5=25

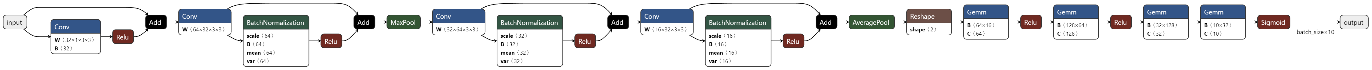
MAC：用Parameter\*output feature map = 25\*28\*28=19600

1. 兩個 3\*3 Convolution：

Parameter：同上但計算兩個 3\*3\*1\*2 = 18

MAC：3\*3\*30\*30+3\*3\*28\*28=15156

1. Model visualization
2. Fashion-MNIST



1. Cifar10



1. Loss curve plot and overfitting judgement
2. Fashion-MNIST

雖然看起來有些差異，但是縱軸的Loss都差不多，所以沒有overfitting

一張含有 繪圖, 螢幕擷取畫面, 文字, 行 的圖片

自動產生的描述

1. Cifar10

Valid loss 明顯並沒有隨著epoch 下降，但是 Train loss 有，已經overfitting了

一張含有 繪圖, 文字, 圖表, 螢幕擷取畫面 的圖片

自動產生的描述

1. My opinion

心得：我覺得這次作業讓人學會到如何使用pytorch搭模型，雖然因為我自己的知識有限，所以搭出來的模型都是隨便亂試碰碰運氣，但是看著自己的模型有隨著自己的更動而準確度有上升的時候也挺有成就感的。

改進：我覺得可以花一些篇幅解釋code在幹嘛，像是範例的resnet我一直很好奇是怎麼用迴圈建構模型的。

或是本次只改config與model的部分，在訓練地方的code我有些不太了解。不過大體而言我覺得學到很多。

1. Model modifications

模型中我使用SGD作為optimization的方法，但是前期一直train不起來，後來改用Adam才能夠train起來。

在避免overfitting方面，我在cifar10的模型有加了dropout layer，但是一開始dropout設定太大(0.5)，導致模型準確度卡在30幾趴就上不去了，後來調低dropout(0.3)才將模型train上去。

而在MNIST的資料集，我嘗試手動進行Shortcut(像resnet)，後來發現效果很不錯，比起只有一層一層的疊，精確度上升的比較快也比較高。