

課程名稱：Video streaming and tracking

作業名稱：Homework 1 – Classification

系級：資訊科學與工程研究所二年級

學號：310551135

姓名：張宇辰

## 1. Environment

- OS: Ubuntu 20.04
- GPU: GTX 1080Ti
- Packages

matplotlib == 3.5.3	numpy == 1.22.4
pandas == 1.4.4	scikit-image == 0.19.3
torch == 1.9.1	Anaconda (conda env name “video_hw1_env”)
Python == 3.8.1	

## 2. Reproduce my result

- Hyperparameters

Name	Value
Epoch	100
Learning Rate	0.008
Batch Size	64
Loss	CrossEntropy
Optimizer	SGD

- 執行訓練：

1. 在專案根目錄下，將所有資料都放入，如：HW1\_310551135/train, HW1\_310551135/train.csv, ...
2. 在專案根目錄下輸入：`python train.py`  
(若訓練完要觀看 loss 圖表，請拿掉 train.py 最下方的註解再訓練)

```
118 # epoch_count = range(1, len(history["train_loss"])+1)
119 # plt.plot(epoch_count, history["train_loss"])
120 # plt.plot(epoch_count, history["valid_loss"])
121 # plt.legend(['train', 'valid'])
122 # plt.xlabel('Epoch')
123 # plt.ylabel('Loss')
124 # plt.show()
125
126 # epoch_count = range(1, len(train_acc)+1)
127 # plt.plot(epoch_count, train_acc)
128 # plt.plot(epoch_count, valid_acc)
129 # plt.legend(['train_acc', 'valid_acc'])
130 # plt.xlabel('Epoch')
131 # plt.ylabel('Accuracy')
132 # plt.show()
```

- 執行測試：

1. 在專案根目錄下，將所有資料都放入，如：HW1\_310551135/test
2. 在專案根目錄底下輸入：`python test.py`  
(執行完畢後，預測結果將儲存成 HW1\_310551135.csv)

### 3. Number of Model parameters

- 參數計算方式 (助教提供)

```
number_of_params = sum(p.numel() for p in model.parameters() if p.requires_grad)
```

- 模型參數總量：413,170

### 4. Explain my structure

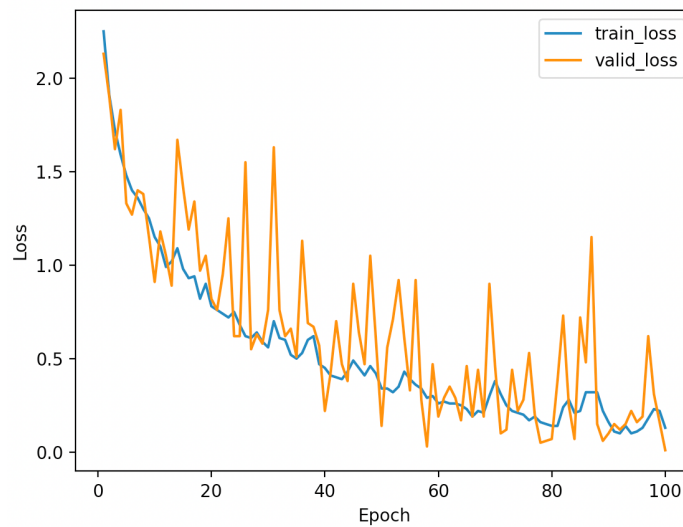
- Input image size: 128 x 128 x 3
- Model structure

Input (128x128x3)
CONV3-8
CONV3-8
POOL (2x2)
CONV3-16
CONV3-16
POOL (2x2)
CONV3-32
CONV3-32
POOL (2x2)
CONV3-64
CONV3-64
POOL (2x2)
CONV3-64
CONV3-64
POOL (2x2)
FC-256
FC-10

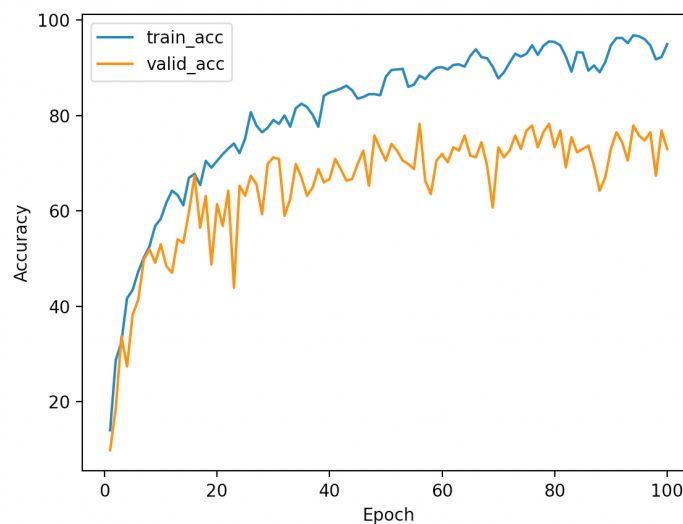
因為原本的 VGG 網路在訓練完便可獲得 70%以上的 Accuracy，因此將重點放在減少模型的參數量，同時又能維持 65%~75%左右的 Accuracy。首先，嘗試將各層的參數量降低，再試著減少訓練的層數，過程中以少量資料訓練觀察模型更動後的結果。

## 5. Results

- Training and Validation loss curve



- Training and Validation accuracy curve



## 6. Problems encountered and discussion

- 由於較缺乏訓練模型的經驗，剛開始一直遇到 image dimension 不符合模型輸入。透過上網查詢解法，得知 Pytorch 接受的參數維度依序為[N, C, H, W]，而起初 input 的維度為[N, H, W, C]，因此利用 *permute* 函式將維度順序重新排序即可。此外，在 transform 階段直接將 image 轉成 tensor 格式亦可解決維度順序問題。
- 起初訓練的 loss curve 雖然整體趨勢往下遞減，但局部區域的起伏非常大。多次嘗試後，試著將 batch size 調大、learning rate 逐步調小到適合區間，雖然局部區域的起伏確實有改善；但仍有一定程度的起伏，此問題調整許多沒有獲得解決，希望老師或助教能幫忙解說問題原因。