

# Lab5

## Low Complexity Model

313510164 陳緯亭 電子碩一

November 25, 2024

### Contents

|          |                                                                              |          |
|----------|------------------------------------------------------------------------------|----------|
| <b>1</b> | <b>Which methods I apply to lower complexity and bridge the accuracy gap</b> | <b>1</b> |
| <b>2</b> | <b>Why do I apply the methods</b>                                            | <b>1</b> |
| <b>3</b> | <b>How do I implement</b>                                                    | <b>2</b> |
| <b>4</b> | <b>Screenshot</b>                                                            | <b>3</b> |

# 1 Which methods I apply to lower complexity and bridge the accuracy gap

1. 減少 Channel 數量。
2. 某些卷基層對性能提升有限，去掉部分卷基層。
3. 使用更小的 Kernel，進行通道的壓縮。
4. 降低 FLOPS，減少每層的運算。
5. 減少中間層的節點數量。
6. Dropout 層的比例調整，減少過擬合。

## 2 Why do I apply the methods

原本採用 ShuffleNet 作為架構，雖然它的 FLOPS 和參數量都挺小的，但是模型的準確度超低，因此改用其他架構。

```
flops = 3671096  
Model parameter size = 7211.203 kB  
Accuracy = 68.19 %
```

Fig. 1: ShuffleNet Testing Results

最後採用一個自定義的卷積神經網絡 (CNN) 模型，適用於像 CIFAR-100 這樣的數據集，該數據集有 100 個類別。這個模型的準確度很高，但是 FLOPS 和參數量都很大，因此需要進行模型簡化。

主要有：卷積層 (conv1, conv2, conv3, conv4) 這些層從輸入圖像中提取層次化的空間特徵。

每個卷積區塊包括：

1. 卷積層 (nn.Conv2d): 將過濾器應用於輸入圖像，用於檢測空間模式。
2. 批量正則化 (nn.BatchNorm2d): 正則化激活值，以穩定訓練並加速收斂。
3. ReLU 激活 (nn.ReLU): 引入非線性，使網絡能學習更複雜的模式。

4. 最大池化 (nn.MaxPool2d，在 conv2 和 conv4 中：下採樣特徵圖以減少空間維度，從而提取重要特徵。
5. 隨著層數的加深，特徵圖的通道數逐漸增長 (從 3  $\rightarrow$  256)，使模型能學習到更複雜的特徵。

### 3 How do I implement

#### 1. 建立高準確率模型

- 一開始先選擇一個成熟且表現良好的模型架構 (例如 ResNet、VGG16 或自定義模型)，在這裡我自定義模型。
- 使用完整的通道數訓練模型，確保模型能夠捕捉足夠的特徵。
- 針對數據集進行充分訓練，達到高準確率。

#### 2. 漸進式減少通道數

- 確定哪些卷積層的通道數可以減少。
- 使用剪枝準則篩選需移除的通道，例如：
  - L1 範數：移除權重值最小的卷積核。
  - 激活值重要性：移除對前向傳播貢獻最小的卷積核。
- 每次減少通道數後，對模型進行微調 (Fine-tune)。

#### 3. 減少通道後的微調

- 每次減少通道數後，進行模型微調以恢復準確率：
  - 使用較小的學習率。
  - 訓練較少的 epoch，避免過擬合。

#### 4. 評估與迭代

#### 5. 其他方法 (此次未採用)

- 可使用工具如 Torch-Pruning 來實現自動化的通道剪枝。
- 使用知識蒸餾: 訓練一個較小的模型 (學生模型) 來模仿高準確率模型 (教師模型) 的輸出。

## 4 Screenshot

```
flops = 31269888  
Model parameter size = 8677.812 kB  
Accuracy = 90.39 %
```

Fig. 2: Testing Results