

國立臺灣師範大學附屬高級中學  
第十一屆 1509 科學班專題研究報告

預處理對超深型圖像辨識模型  
泛化能力的影響

專題組別：資訊科

作 者：曹禕中、張柏豐

指導教授：國立臺灣大學陳中平教授

指導老師：洪允東、陳智勝老師

# 預處理對超深型圖像辨識模型泛化能力的影響

作者：曹禕中、張柏豐

指導教授：陳中平教授

助教：吳思穎助教

## 摘要

由於已有大量文獻探討影像辨識的模型，但同樣重要的預處理對模型泛化的影響，卻不見擁有同樣人氣。因此我們以 Cifar10 著手進行訓練並探討不同預處理函式的配置對模型泛化能力的影響。我們以五種不同的預處理函式對模型進行訓練，測試其對不同於 Cifar10 的資料集的正确率，並希望由此了解可以預處理對超深型圖像辨識模型泛化能力的影響。

## 壹、研究動機

本研究目的是探討不同的預處理對超深型圖像辨識模型泛化能力的影響，在科技日新月異的進步下，各種利用深度學習辨識圖片的研究比比皆是，利用深度學習辨識影像的範例已不勝枚舉。而每次訓練的資料若是不經處理，可能使得圖片特徵不利擷取，或是資料集數目過少，因此通常需要在訓練前對圖片作預處理，方便模型學習、辨識。而泛化能力亦是考量模型實用程度之一大指標，因此我們將目的轉移為更原始的「預處理對超深型圖像辨識模型泛化能力的影響」。我們選擇 Cifar10作為我們的資料集，並且研究在幾種不同種類的預處理下，能對模型的泛化能力提升多少。

## 貳、研究目的

探討訓練時預處理函式的不同對模型泛化能力的影響。

## 參、研究設備及器材

### 一、訓練資料

#### (一)、cifar10

為訓練模型的主體，即改變預處理的對象

### 二、顯卡

#### (一)、GeForce RTX 2070

### 四、開發環境

#### (一)、Jupyter Lab、Keras

本研究使用的開發環境為自架的 Jupyter Lab，並且搭配一張 RTX2070顯示卡，因為有圖像輸出的介面，方便圖表等數據的即時展現，且便於操控進程，對開發人工智慧十分友善。

## 肆、研究過程或方法



### 一、建立適合的影像處理模型

我們使用 Keras 建立並且訓練模型，使用的模型以 VGG-16為基礎，在針對我們的需求作更進一步的修改。同時，由於 Cifar10的資料比較小（32x32像素），但是 VGG-16是設計給224x224像素的圖片，為了使模型能更有效率的學習，我們降低了每個卷積層 Kernal 的數量，也減少了最後兩層全連接層神經元的數量。

### 二、預處理結果測試

我們分別測試兩種類型的預處理對驗證集正確率和 loss 的影響，分別是統一，和加入雜訊這兩種策略。

#### 第一類：統一

只改變原始資料的單一項目，而我們選擇的項目有兩項，分別是：

##### （一）、亮度統一

將所有的照片亮度統一成一半的亮度值。

##### （二）、統一對比度

將每張照片的對比度統一設定為-1 到 1 之間。

## 第二類：加入雜訊（綜合項目的變化）

同時改變原始資料的不同項目，並且我們對不同程度的預處理函數進行級。共分三級，排序由影響最小到最大。

### （一）、第一級

不做任何修改。

示例圖：這是一隻青蛙。


### （二）、第二級

最多寬度縮放5%，高度縮放5%，銳利化邊緣5%，縮放5%，色調轉換10%。

示例圖：，這是一架飛機。

### （三）、第三級

對多旋轉45度，寬度縮放20%，高度縮放20%，銳利化邊緣20%，縮放20%，顏色可能完全轉換，有可能水平垂直翻轉。

示例圖：，這是一隻貓。

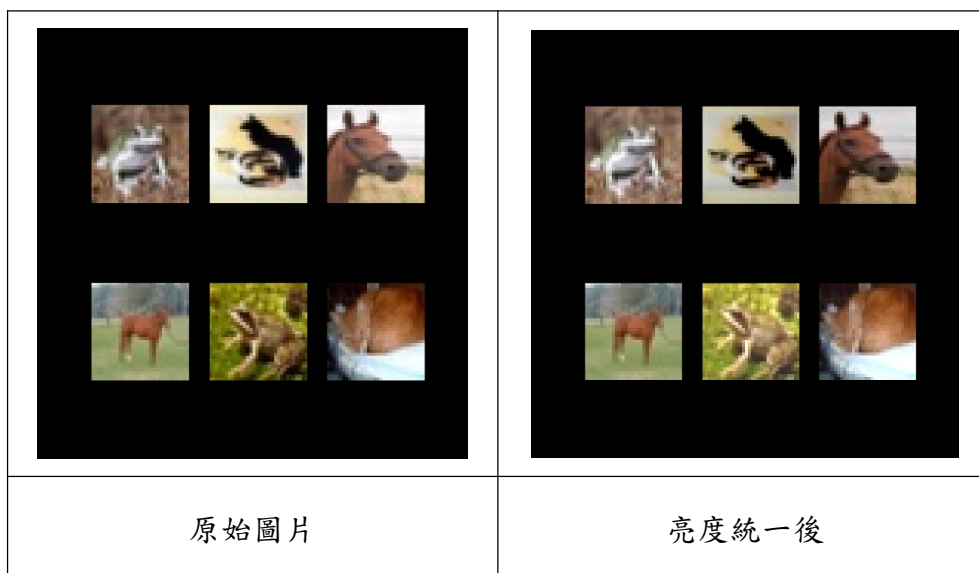
## 伍、研究結果與討論

一、統一策略的亮度統一示範圖例如下：

將亮度統一成50%

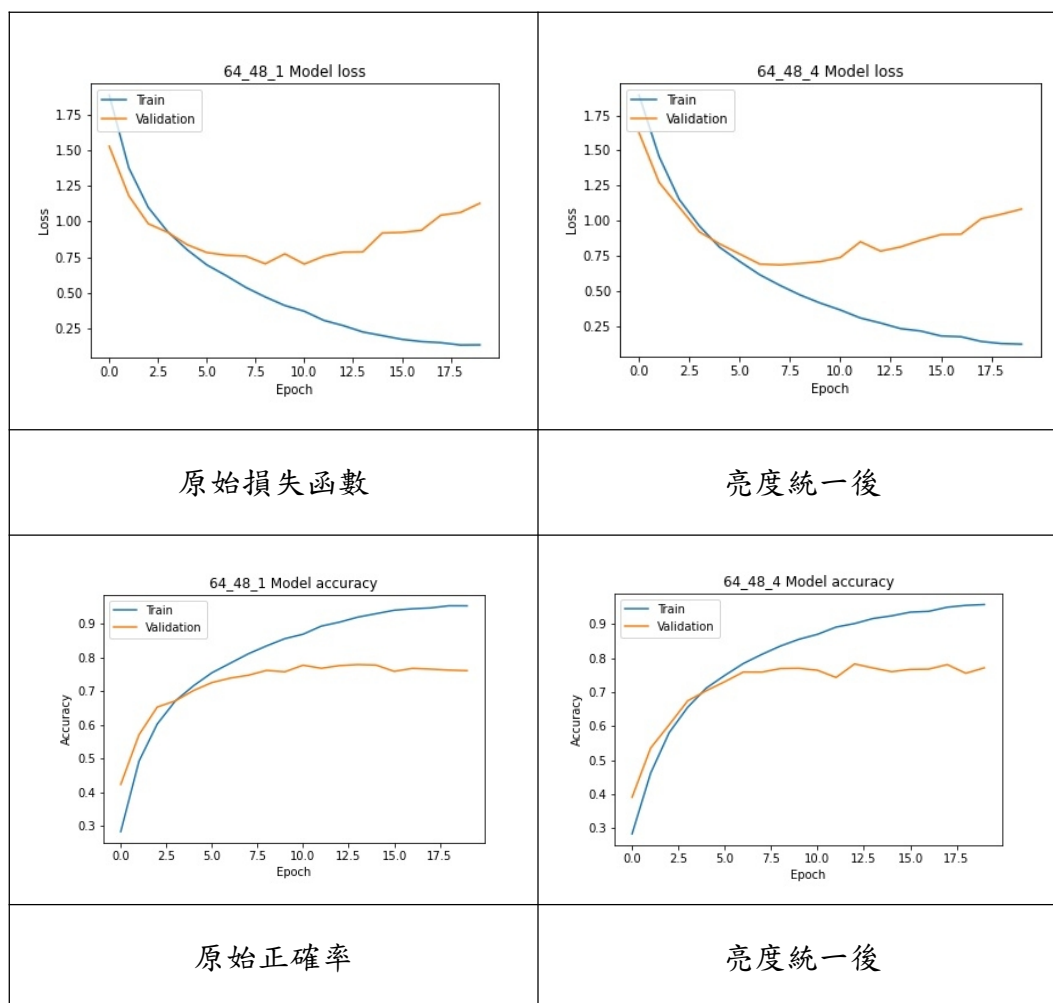
由-1~1改變成1~0

照片成果對照結果如下：



以下出現的圖表其橫軸的 Epoch 為訓練次數，縱軸為 Loss 值或 Accuracy（正確率），而 Validation 為驗證集，Train 為訓練集。

以圖表呈現的損失函數以及正確率如下：



對照組 Accuracy：0.7781

實驗組 Accuracy：0.7723

增長幅度：-0.75%

對照組 Loss：1.0229

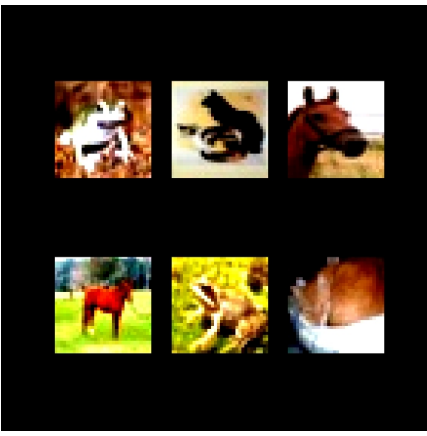
實驗組 Loss：0.9591

增長幅度：-6.24%

二、統一策略的對比度統一示範圖例如下：

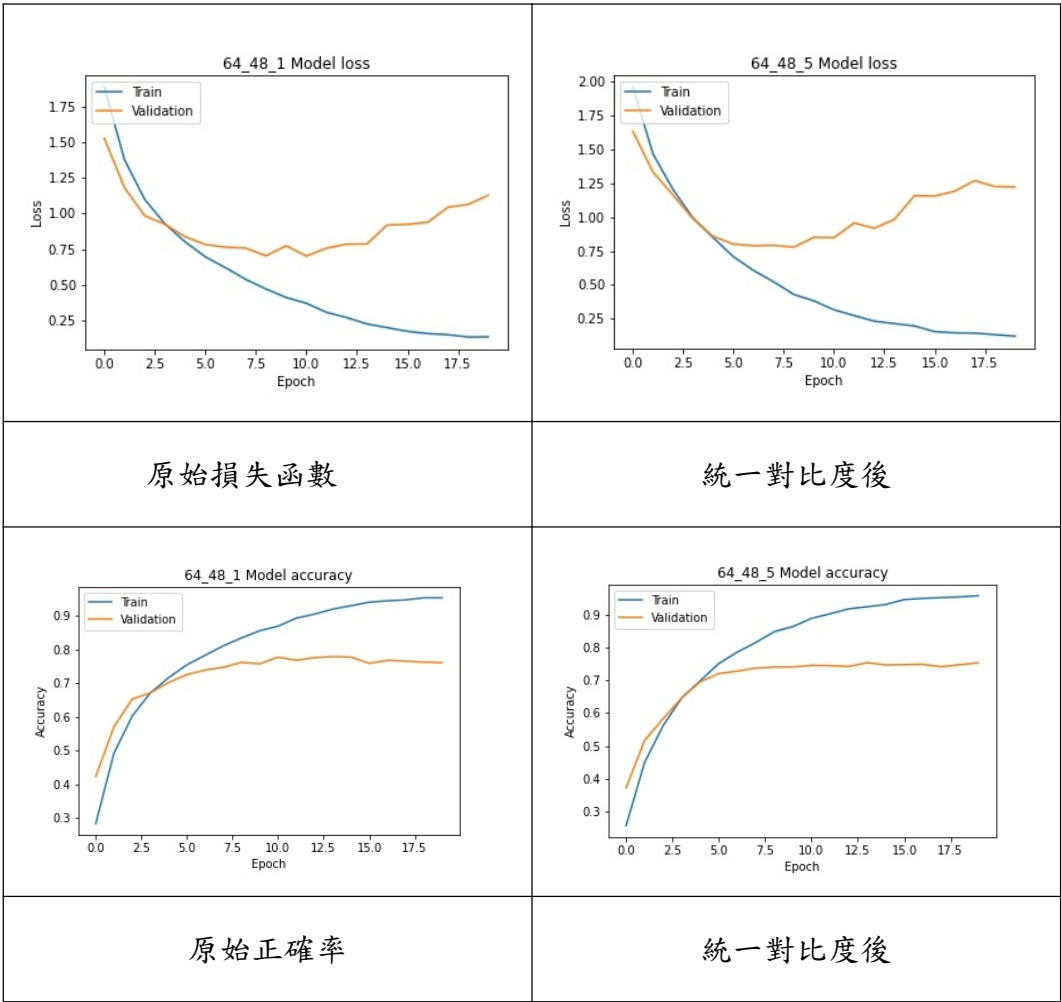
將對比度的總平均調整為0且範圍為0至-1

照片成果對照結果如下：

	
原始圖片	統一對比度後

以下出現的圖表其橫軸的 Epoch 為訓練次數，縱軸為 Loss 值或 Accuracy（正確率），而 Validation 為驗證集，Train 為訓練集。

以圖表呈現的損失函數以及正確率如下：



對照組 Accuracy：0.7781

實驗組 Accuracy：0.7123

增長幅度：-8.46%

對照組 Loss：1.0229

實驗組 Loss：0.0826

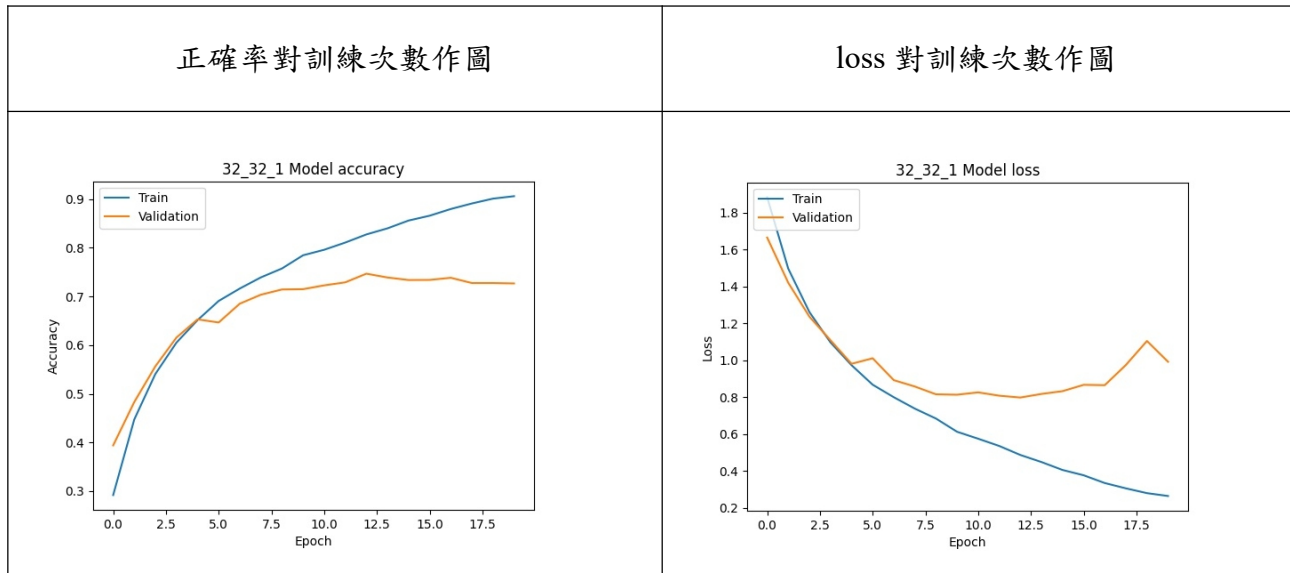


增長幅度：+5.84%

三、綜合項目的第一級結果如下（即對照組）：

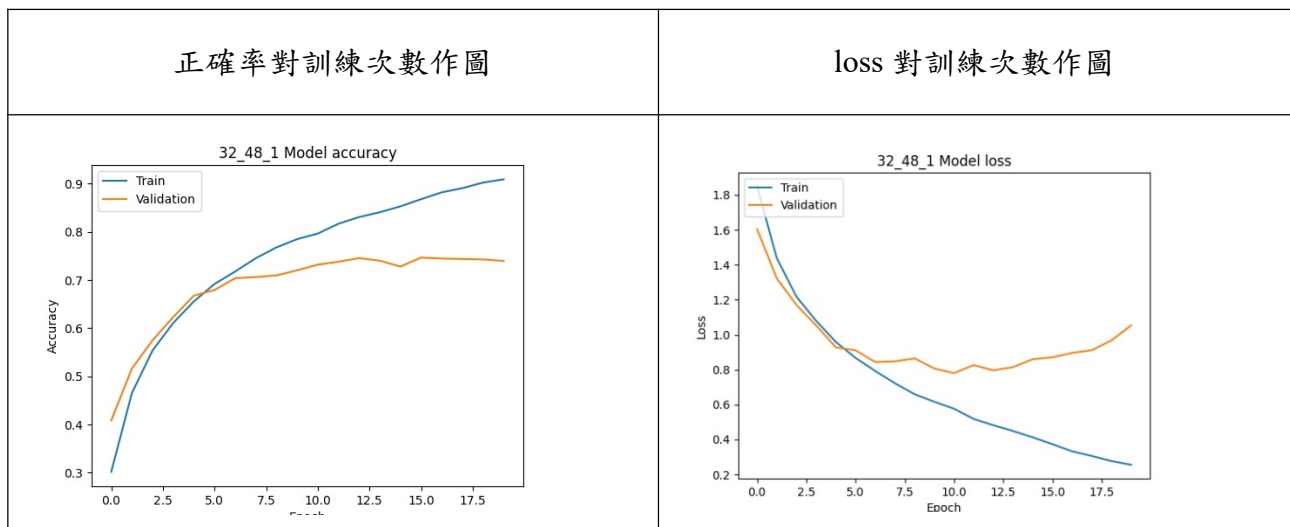
A 模型配合第一級預處理函式：

（對照組）驗證集正確率：0.7484



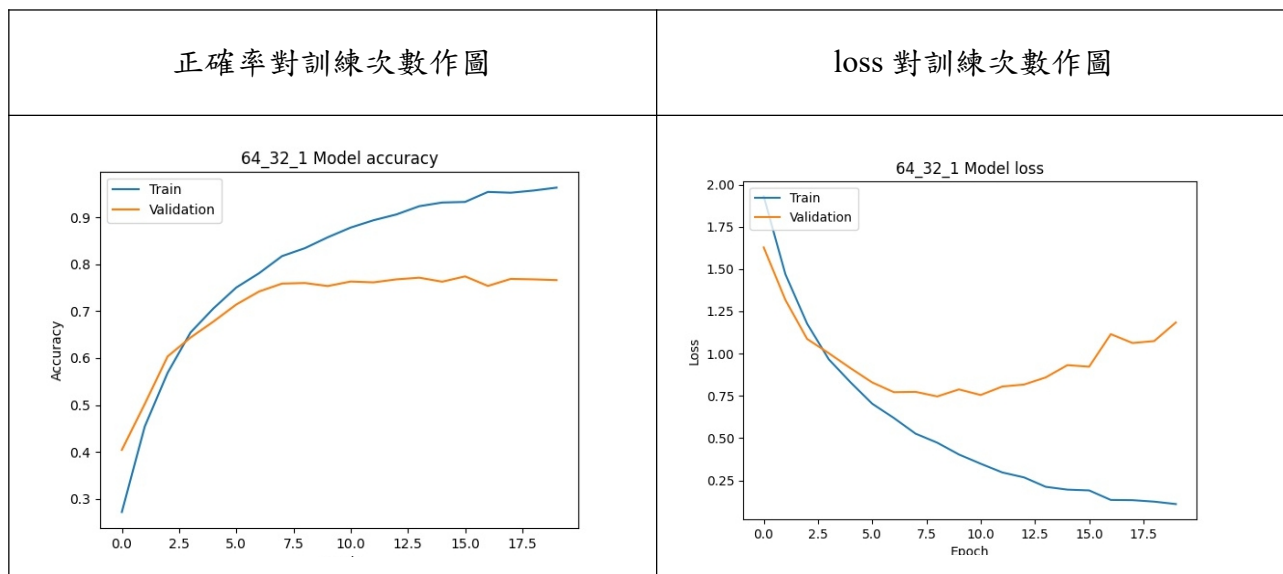
B 模型配合第一級預處理函式：

（對照組）驗證集正確率：0.7394



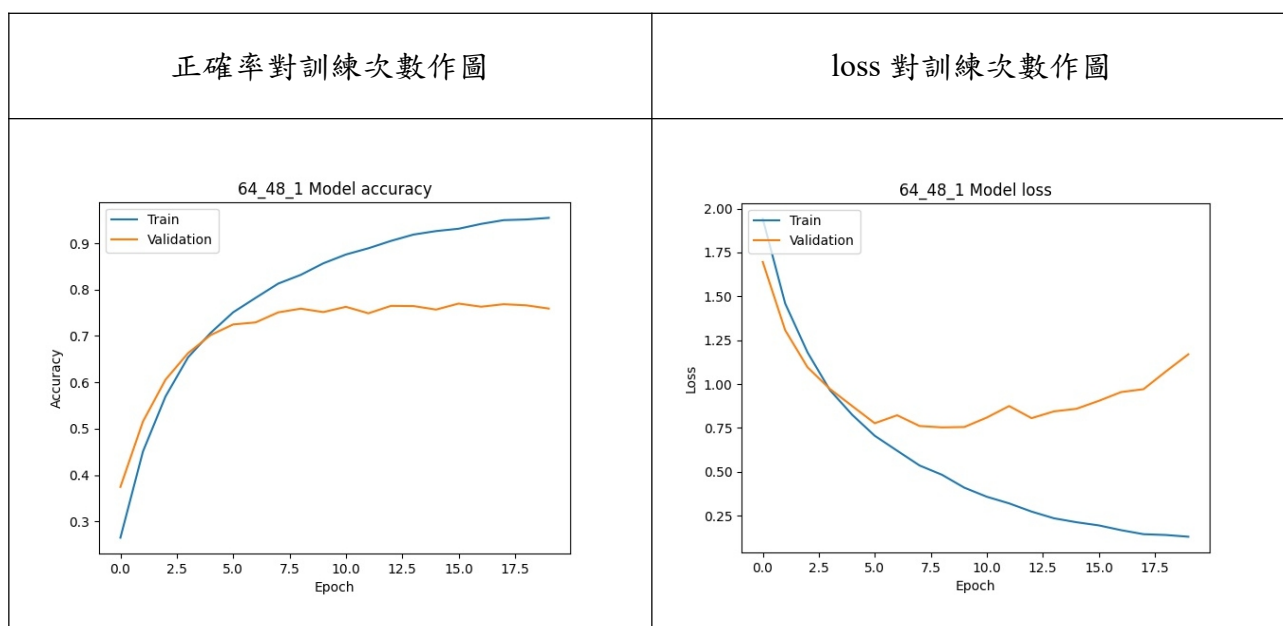
C 模型配合第一級預處理函式：

(對照組) 驗證集正確率：0.7572



D 模型配合第一級預處理函式：

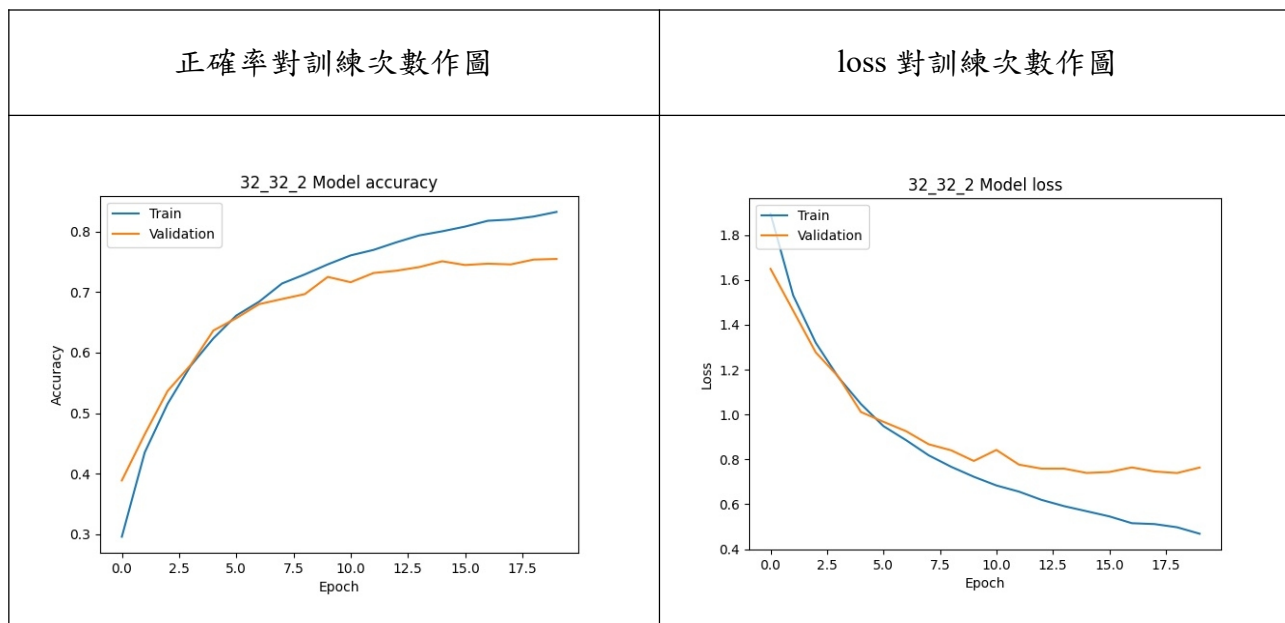
(對照組) 驗證集正確率：0.7781



四、綜合項目的第二級調整結果如下：

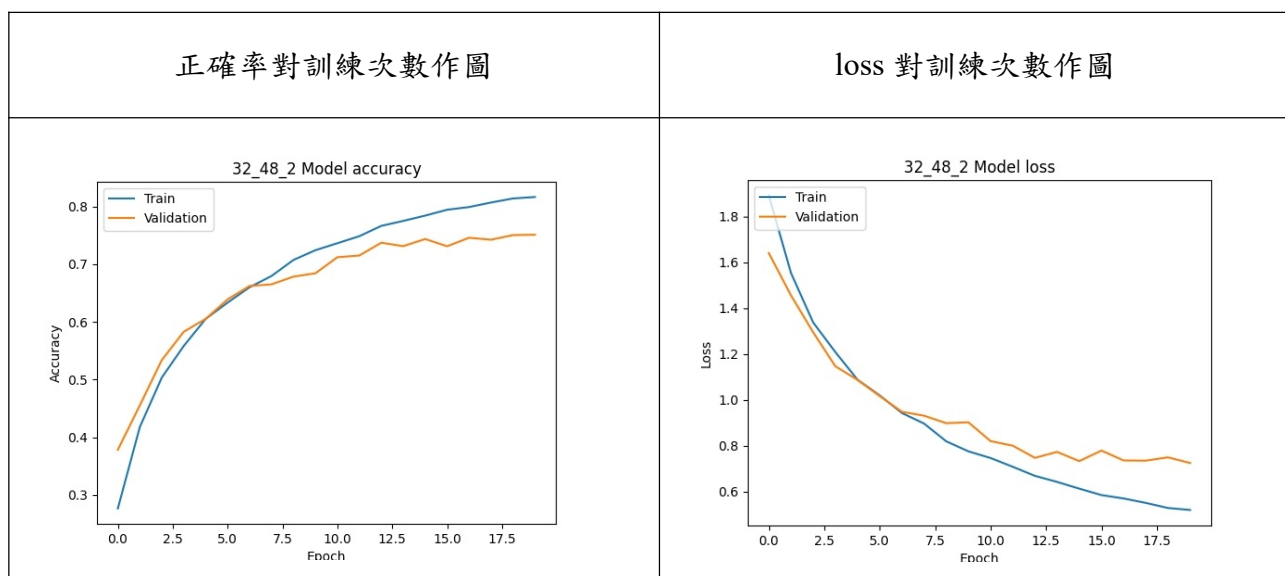
A 模型配合第二級預處理函式：

(實驗組) 驗證集正確率：0.7685



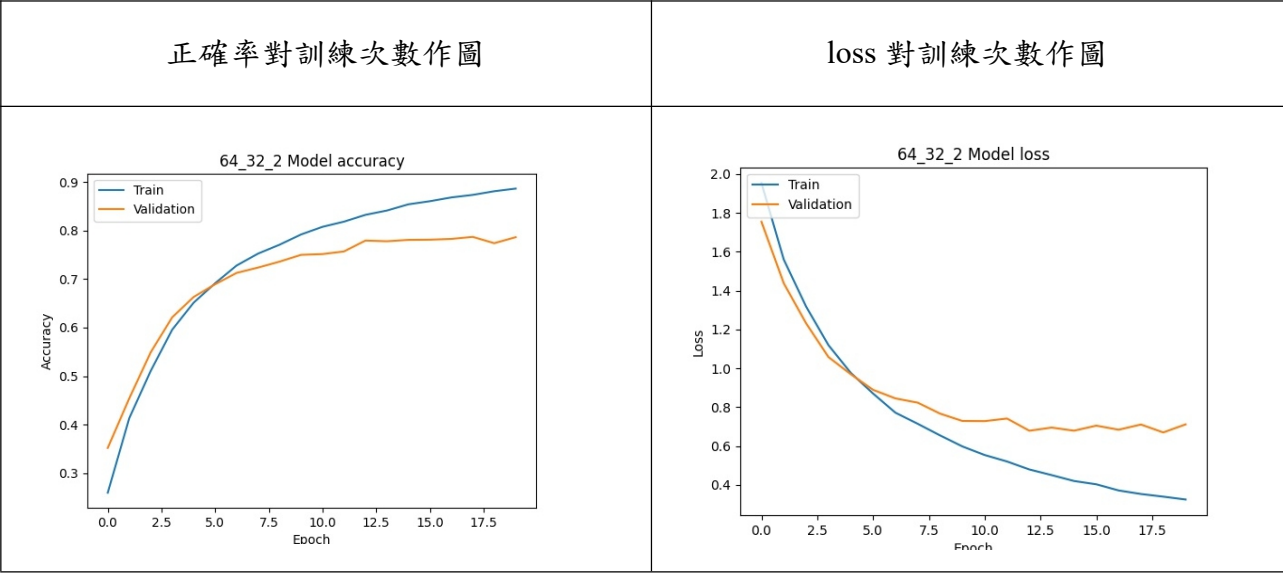
B 模型配合第二級預處理函式：

(實驗組) 驗證集正確率：0.7663



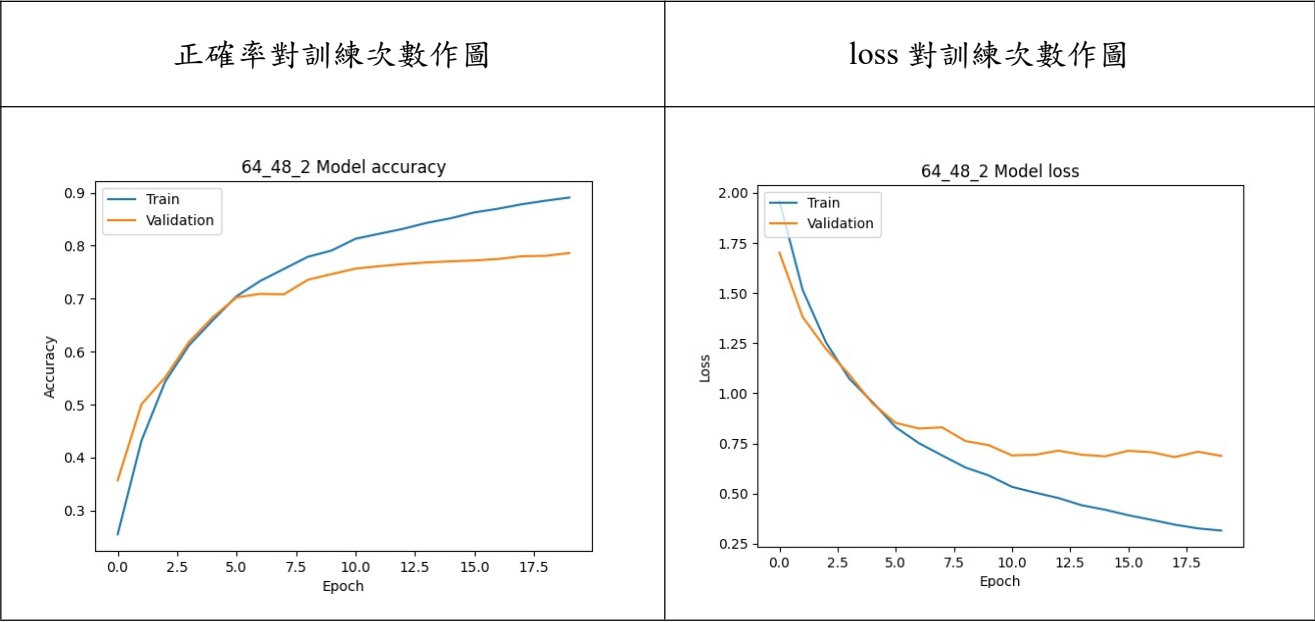
C 模型配合第二級預處理函式：

(實驗組) 驗證集正確率：0.7794



D 模型配合第二級預處理函式：

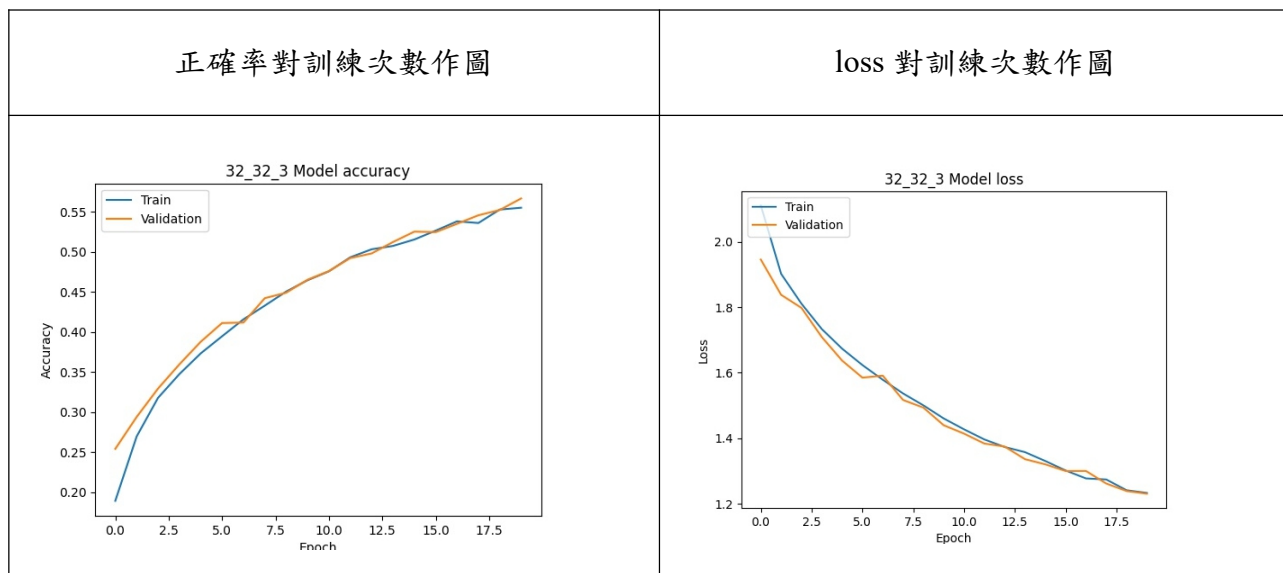
(實驗組) 驗證集正確率：0.7933



五、綜合項目的第三級調整結果如下：

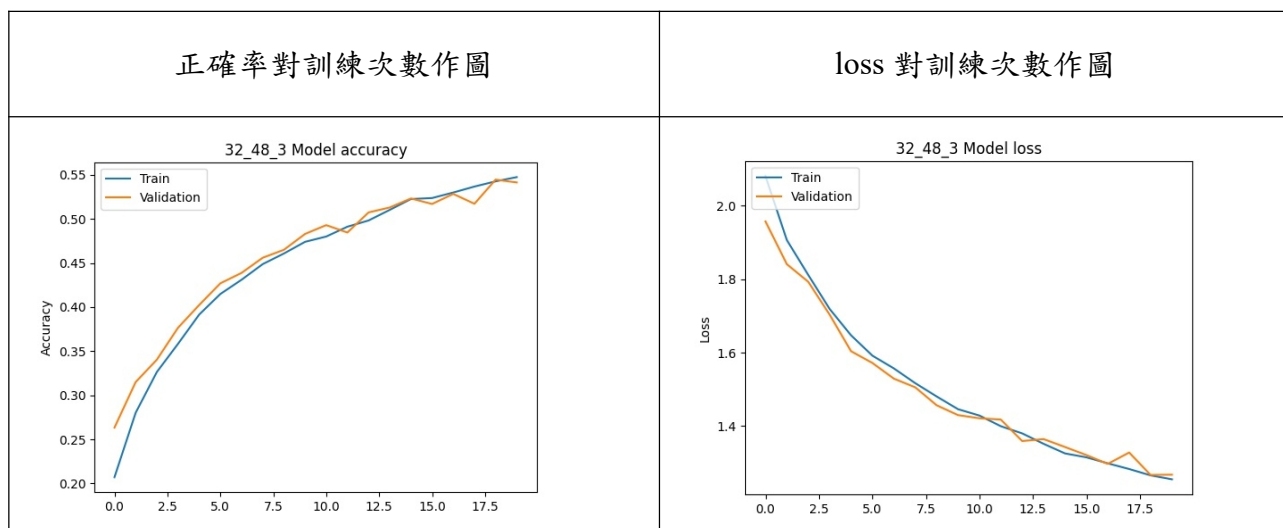
A 模型配合第三級預處理函式：

(實驗組) 驗證集正確率：0.5544



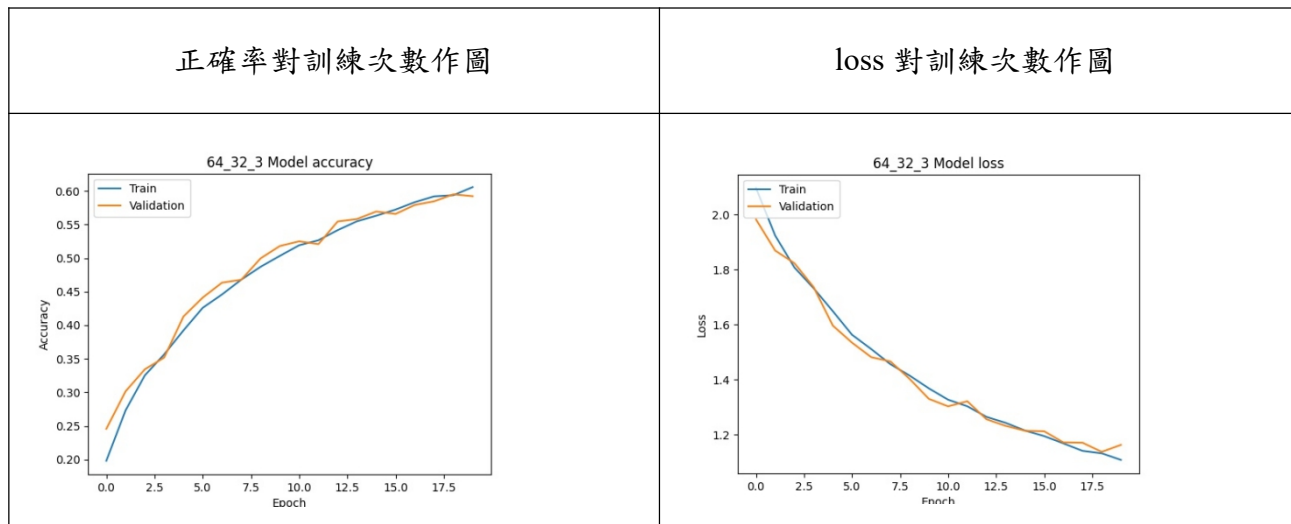
B 模型配合第三級預處理函式：

(實驗組) 驗證集正確率：0.5535



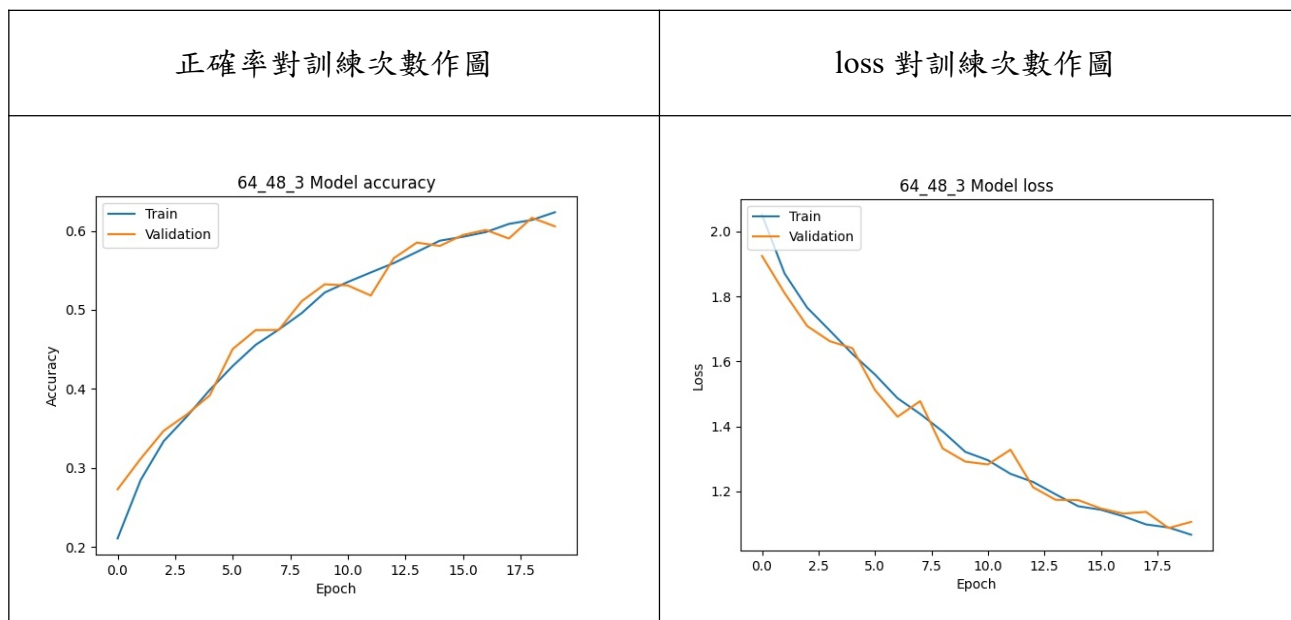
C 模型配合第三級預處理函式：

(實驗組) 驗證集正確率：0.5787



D 模型配合第三級預處理函式：

(實驗組) 驗證集正確率：0.5968

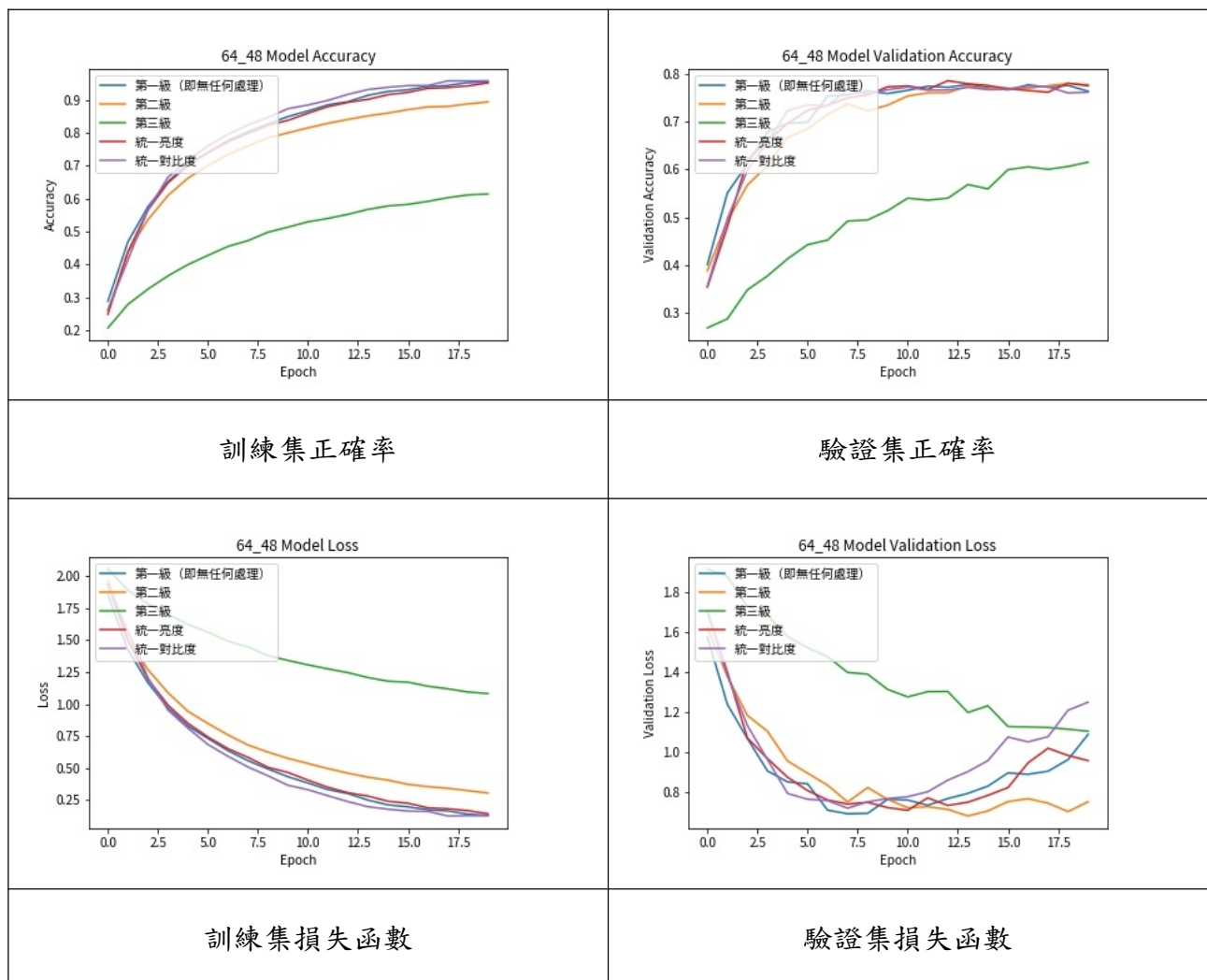


其中我們使用的訓練參數如下，驗證集分割10%，批次大小為250（一個批次計算一次 loss ），並且保存在訓練過程中驗證集正確率最高的那個時間下的模型，訓練回數200次。B 模型(32\_48)代表第一層捲積層 kernal 數量為32個，全連接層一層有48個神經元。其中表格內容總整理如下：

預處理函式等級		A 模型(32_32)	B 模型(32_48)	C 模型(64_32)	D 模型(64_48)
第一級	Loss	0.8812	0.8853	1.0806	1.0229
第二級		0.7046	0.7410	0.8308	0.7259
第三級		1.2759	1.2395	1.2291	1.2172
統一 亮度		1.0511	0.9673	1.2324	0.9591
統一 對比度		0.9994	0.8715	0.9699	1.0826

第一級	Acc	0.7484	0.7394	0.7572	0.7781
第二級		0.7685	0.7663	0.7794	0.7933
第三級		0.5544	0.5535	0.5787	0.5968
統一 亮度		0.7198	0.7343	0.7549	0.7723
統一 對比度		0.6732	0.7155	0.7122	0.7123

## 六、不同預處理對正確率的影響





以上為 D 模型搭配五種預處理在訓練過程20個 epoch 中，訓練集、驗證集的正確率、損失函數對 epoch 作圖。可以發現不同預處理對訓練有著不同的影響，其中第三集預處理對訓練結果的影響最大。

#### 七、泛化結果（正確率）

	A	B	C	D
第一級	0.3966	0.3702	0.4249	0.4150
	0.7484	0.7394	0.7572	0.7781
第二級	0.4043	0.3592	0.4351	0.4673
	0.7685	0.7663	0.7794	0.7933
第三級	0.2542	0.1914	0.2124	0.2114
	0.5544	0.5535	0.5787	0.5968
統一亮度	0.3482	0.3959	0.3877	0.4187
	0.7198	0.7343	0.7549	0.7723
統一對比度	0.4142	0.4200	0.4446	0.4426
	0.6732	0.7155	0.7122	0.7123

灰色網底的是新資料集（另外找的資料庫）之泛化結果的正確率

白色網底的是原資料集（Cifar10）之訓練結果的正確率

我們取了一個模型對應各個預處理的訓練過程中，驗證集分數最高的 epoch 的模型。並且我們進行泛化測試的資料來源為8個不同於 Cifar10的資料集，飛機、鳥、鹿、青蛙、馬、船各一個，和一個分辨卡車和汽車的資料及和一個分辨貓或狗的資料集，恰對應原本 Cifar10

的資料庫。

在我們訓練的結果中可以發現，某些預處理會使辨識的正確率降低。但是這並不代表這會減損其泛化的能力。

且在我們的泛化結果數據中，也可以發現部分使驗證集正確率下降的組合，卻擁有比較好的泛化能力，並且差異頗大。雖然沒有明顯規律，但是在以 Cifar10 為資料集做圖片辨識並分類的任務中，可以看出不同的模型若以泛化為目的，似乎適用於不同的預處理函式。

而最後，在訓練結果中我們也發現，在統一對比度的正確率下降了，照理來說 loss 值應該要上升，不過在最後得到的結果卻是 loss 反而也下降了，因此我們推測應該也是訓練次數不夠多的關係，可能每次都差一點點就答對了，但正確率只看最後的結果，因此依舊是錯的，不過 loss 會因此而減少。因此若訓練次數放大，結果應該會比較平衡。

## 陸、結論

- 一、在適當的預處理下，能夠增加訓練集的多樣性，更好的利用每一張圖片所攜帶的資料。
- 二、預處理若使的圖片不易辨識，則會對訓練有不利的影響。
- 三、就算是會使得驗證集正確率下降的預處理函式，也有可能會使模型泛化能力增加。

## 柒、未來展望

- 一、探究不同策略之預處理對模型表現的影響。
- 二、用更多樣的控制變因深入測試各種預處理的影響。

## 捌、參考資料

- 一、Karen Simonyan & Andrew Zisserman. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. International Conference on Learning Representations
- 二、Y. Lecun, L. Bottou, Y. Bengio and P. Haffner. (1998). *Gradient-based learning applied to document recognition*, in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, doi: 10.1109/5.726791