

深度學習基礎概論

0325

目錄

- CNN model
- Code :

CNN Model

Why use CNN model?

在進行影像辨識時，將所有圖片壓成一維，再使用 fully connected 可能會產生下列問題

- pattern 會出現在圖片中的某個位置，且 pattern 是由相鄰的 pixels 所構成，若壓成一維可能會破壞 pattern 的結構
- 面對高 pixel 的圖片，會大幅增加計算成本
- 事實上，CNN model 是 fully connected 的子集，有較少的參數，因此有較大的 model bias

CNN model

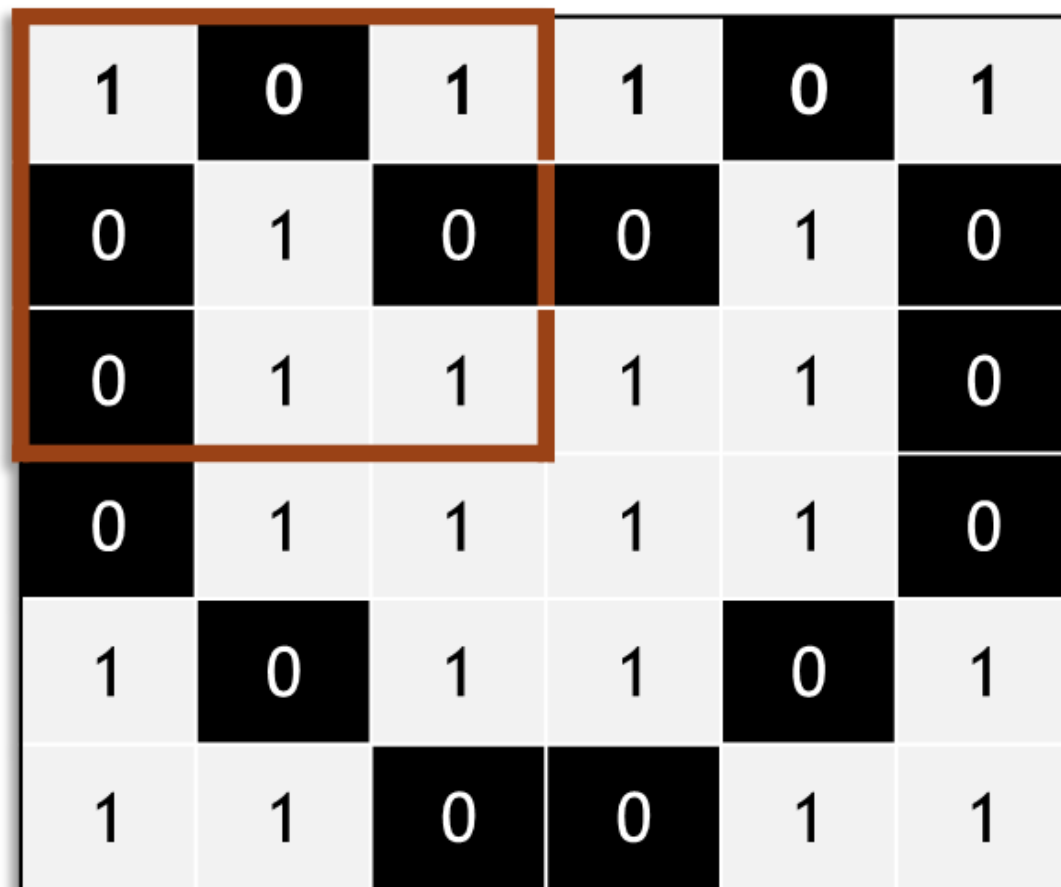
第一個CNN model Lenet-5([1998](#))

整個 CNN model 主要可以分為三個部分

- 卷積層(Convolutional layer)
- 池化層(Pooling layer)
- 全連接層(fully connected layer)

Receptive fields

- Model 每次看多大的區域
- 以 3×3 為例，在這個 size 下，共會有 3×3 個 weights 可以 train



1	0	1	1	0	1
0	1	0	0	1	0
0	1	1	1	1	0
0	1	1	1	1	0
1	0	1	1	0	1
1	1	0	0	1	1

Filter

- 每個 receptive field 都有一組參數可以去訓練
- 同一個 receptive field 可以有多個 filter 去監看

Convolutional layer

- Convolution 是由兩個步驟組組成的運算：滑動(stride) + 內積
- Stride：每次 filter 移動距離
- Convolution 的結果成為 feature map

1 _{x1}	1 _{x0}	1 _{x1}	0	0
0 _{x0}	1 _{x1}	1 _{x0}	1	0
0 _{x1}	0 _{x0}	1 _{x1}	1	1
0	0	1	1	0
0	1	1	0	0

Image

4		

Convolved
Feature

Zero Padding

- 從右下角的圖可以發現 **feature map** 較原本圖小了一圈
- 透過在最外圈向外補一圈 0 的方式使做完 **convolution** 後有和原圖相同大小的 **feature map**

原始影像

1	0	1	1	0	1
0	1	0	0	1	0
0	1	1	1	1	0
0	1	1	1	1	0
1	0	1	1	0	1
1	1	0	0	1	1

零填充

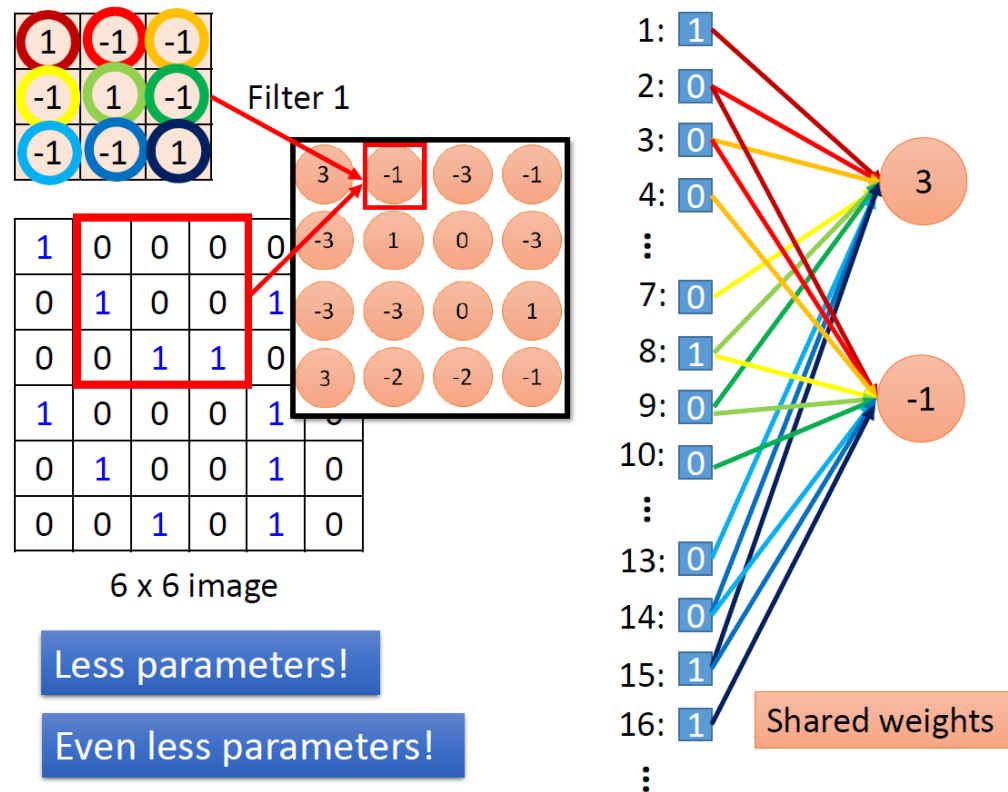
0	0	0	0	0	0	0	0
0	1	0	1	1	0	1	0
0	0	1	0	0	1	0	0
0	0	1	1	1	1	0	0
0	0	1	1	1	1	0	0
0	1	0	1	1	0	1	0
0	1	1	0	0	1	1	0
0	0	0	0	0	0	0	0

同質填充

1	1	0	1	1	0	1	1
1	1	0	1	1	0	1	1
0	0	1	0	0	1	0	0
0	0	1	1	1	1	0	0
0	0	1	1	1	1	0	0
1	1	0	1	1	0	1	1
1	1	1	0	0	1	1	1
1	1	1	0	0	1	1	1

Parameter sharing

- 當 filter 滑動的時候，也表示同一組 filter 的權重正在不同區域裡面運作，這樣的過程中 filter 沒有改變，也就是權重是不會改變的

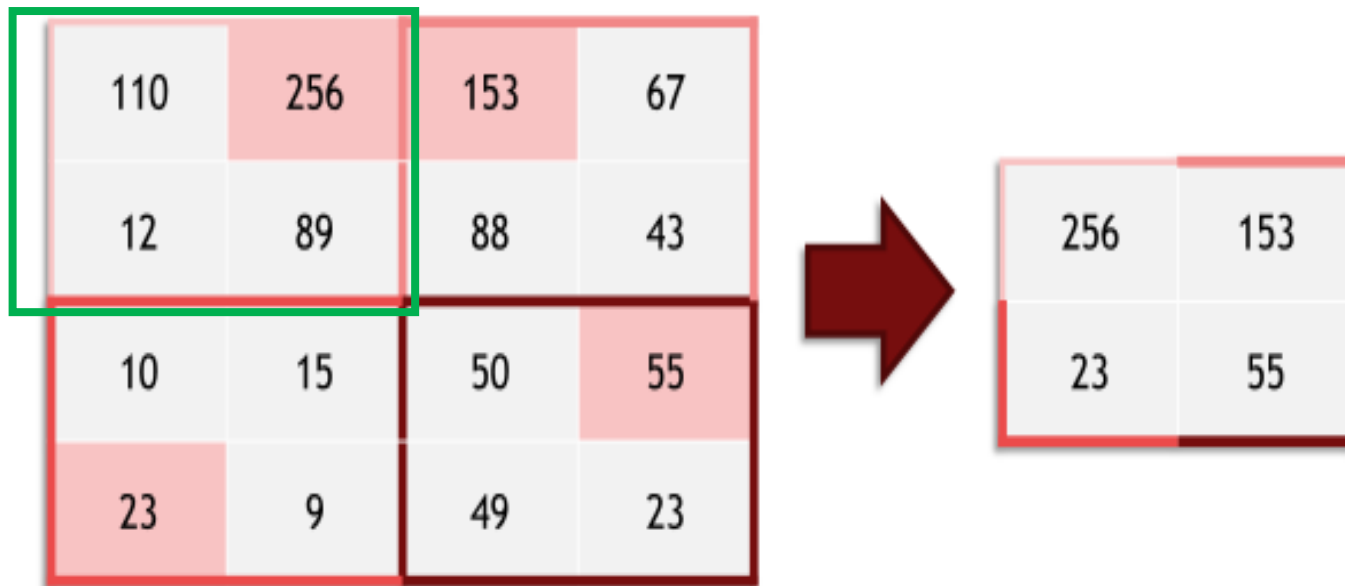


一次 convolution 後有幾張圖片？

- 以 LeNet-5 model 為例，單看第一層 input 32×32 的圖片，經過 5 個 filter，形成五張 feature maps，但其實還必須將這五張 feature maps 疊合成為一張圖片，才會是第一層的 output
- 可以想成輸出一張圖片，但變厚了，或是深度變深了

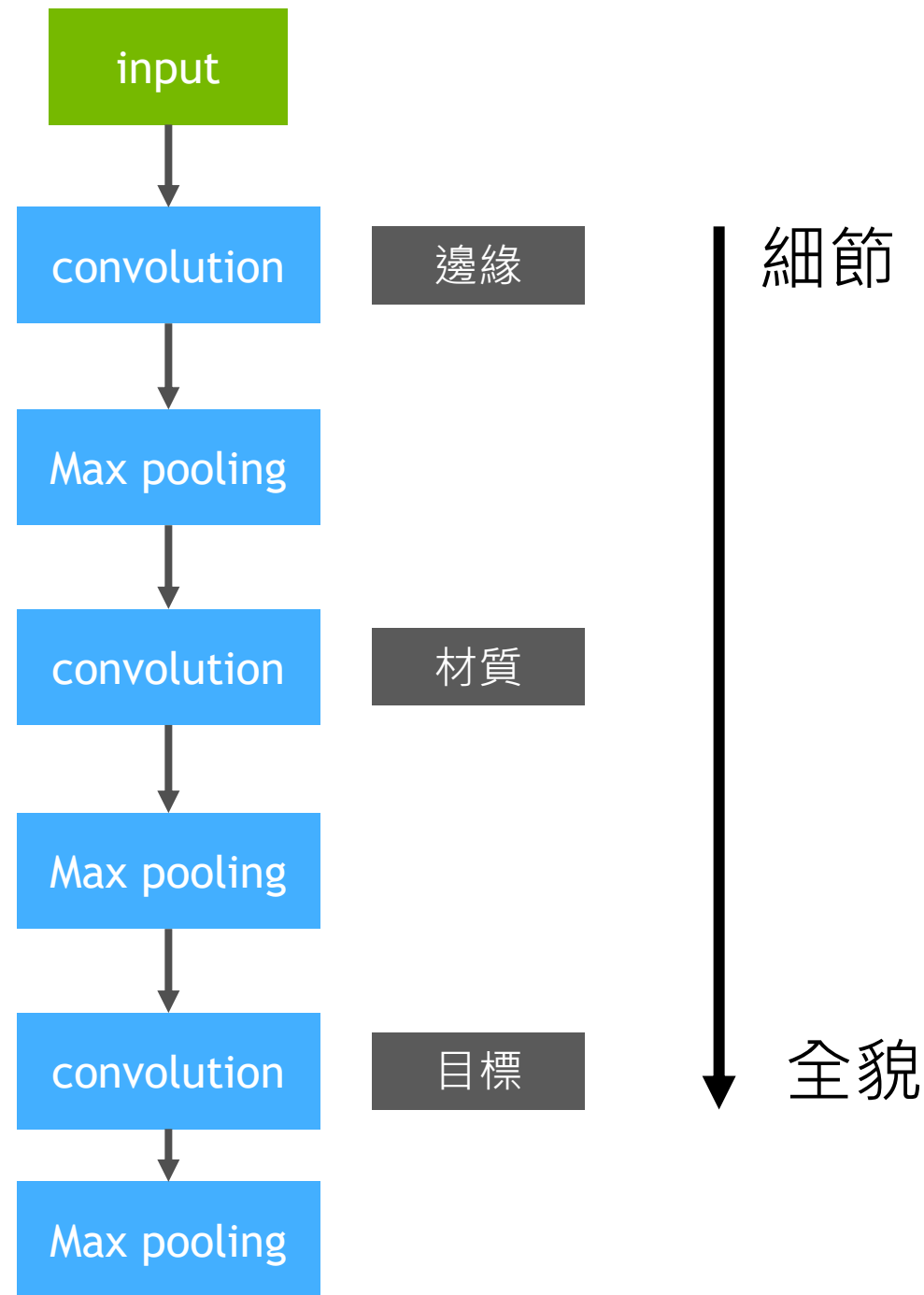
Max Pooling

- 和 convolution layer 相同也是有一個 filter 在圖片上滑動運算，不同的點在於，滑動的區域不重複
- 以左上角為例
- $\text{Max}(110, 256, 12, 89) = 256$
- 做完 max Pooling 後圖片會長寬會縮成一半

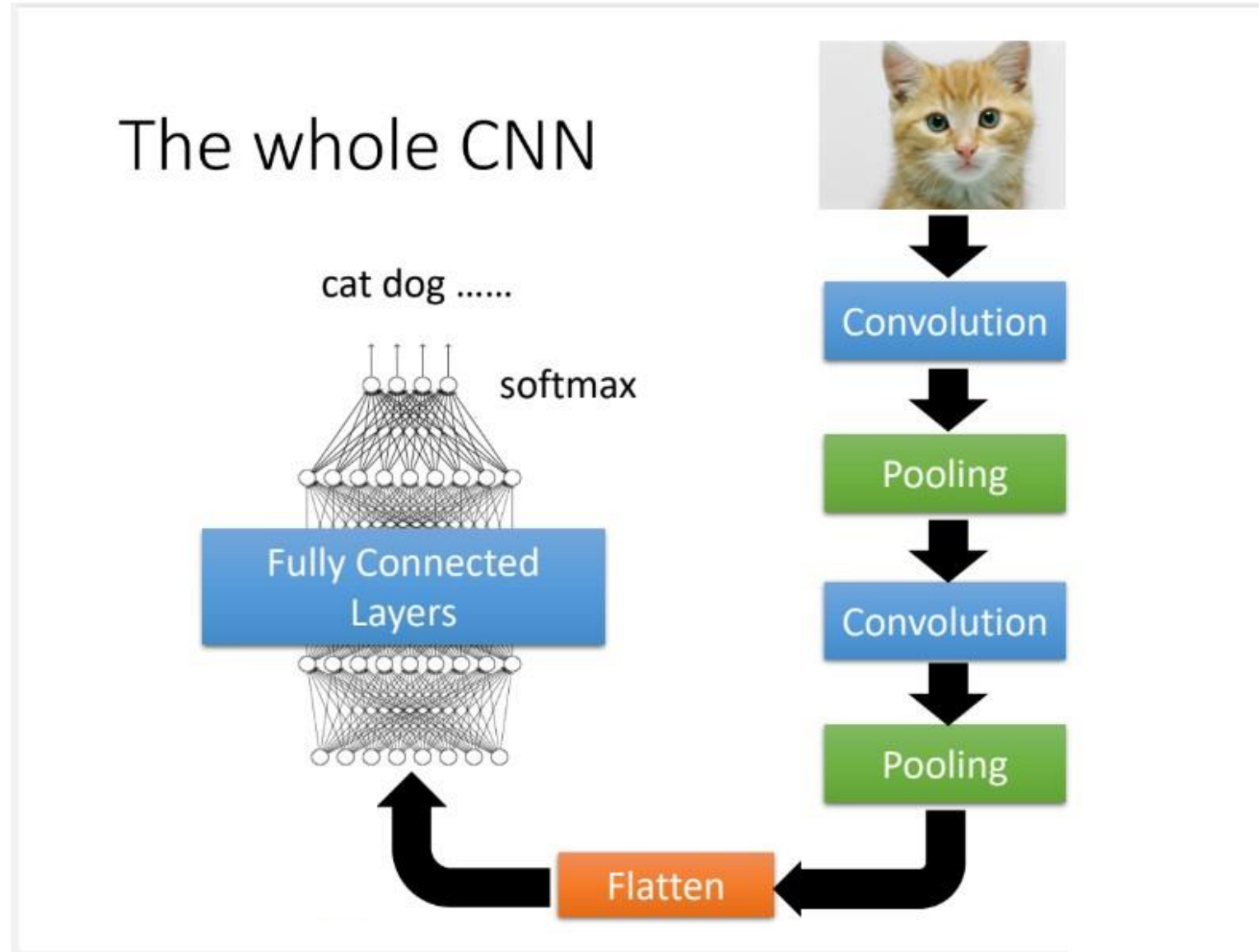


Max Pooling

- 也就是說在 filter 不變的情況下，透過 Max pooling，每個 filter 會從先看到圖片細節，在看到圖片的全貌
- Convolution 其實也有相同效果
- 使用 max pooling 效果比 average pooling 好
- 因為圖片尺寸減少 $\frac{1}{2}$ ，所以計算成本也減少 $\frac{1}{2}$



CNN Model



Code