

卷積神經網路

Convolutional Neural Network

國立東華大學電機工程學系 楊哲旻

Outline



1 模型架構

2 特點

3 卷積層

4 卷積與填充

5 自動提取特徵

6 池化層

7 全連接層

8 超參數

8 卷積神經網路實作

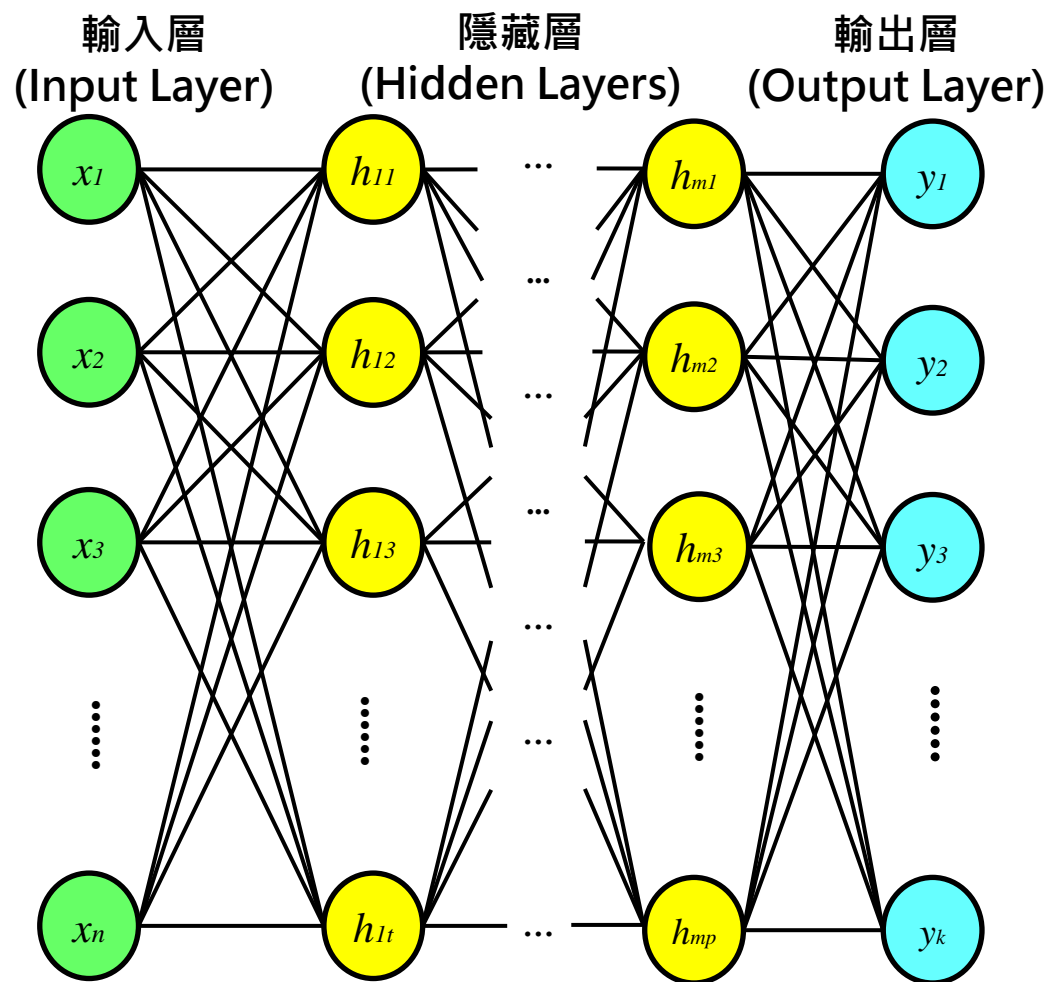
機器學習 – 卷積神經網路

01. 模型架構

2

卷積神經網路為深度學習的分類模型，架構為一輸入層、數個隱藏層與一輸出層所組成，其中隱藏層包含有卷積層、池化層與全連接層。它與多層感知器不同點是並非所有神經元會全部連結。

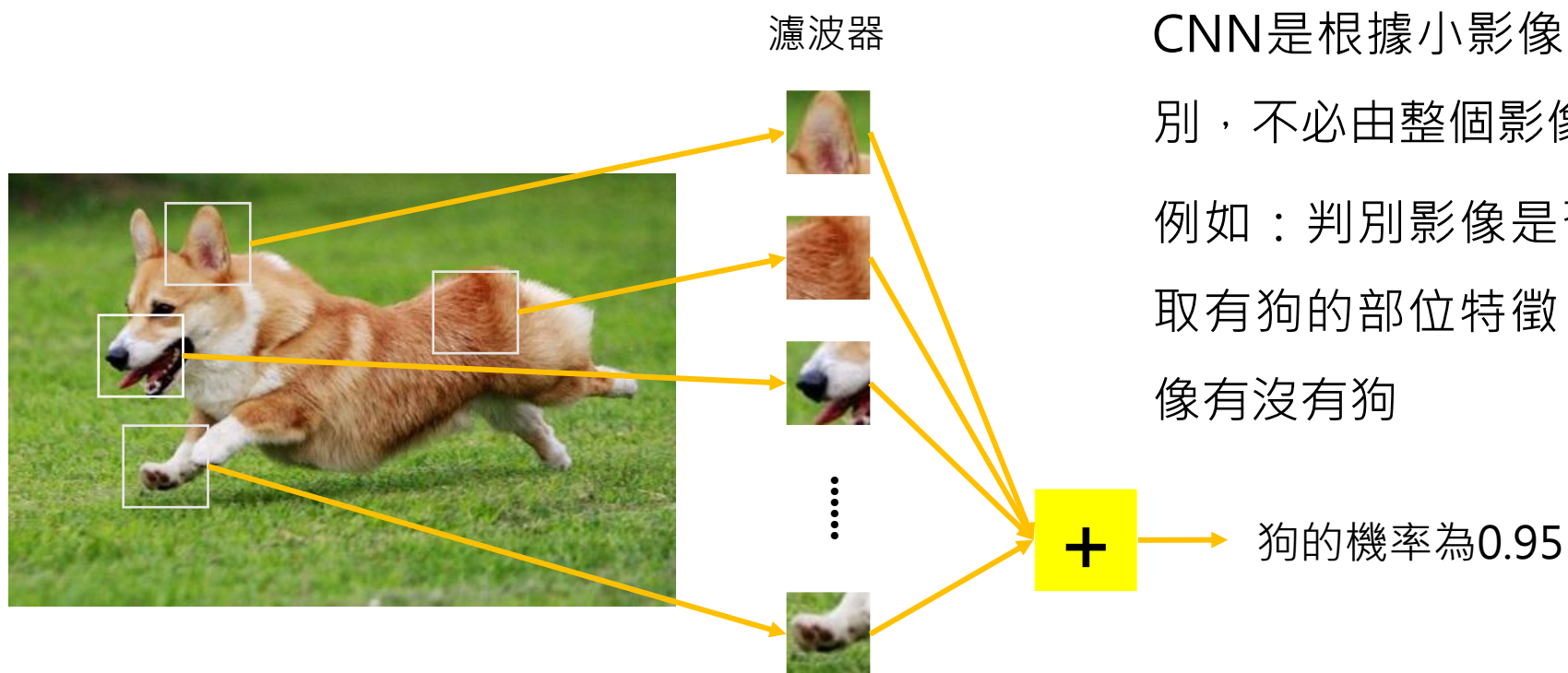
訓練過程是從訓練集中以梯度下降法來確定權重與偏差。



機器學習 – 卷積神經網路

02. 特點

卷積神經網路概述：卷積層(Convolution layer)



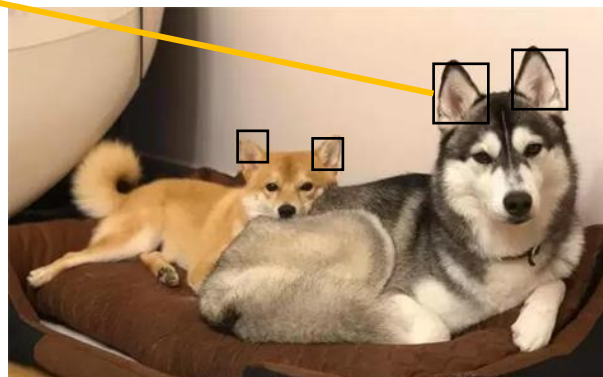
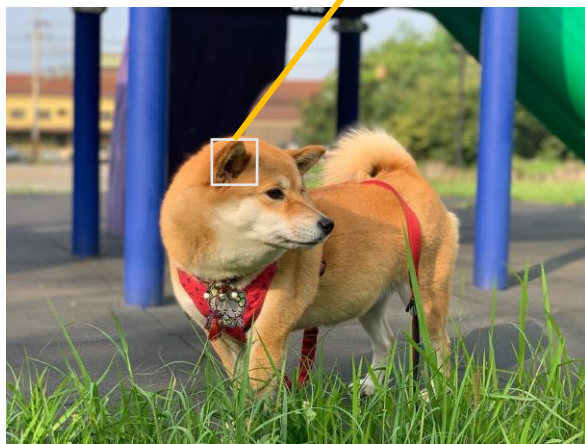
機器學習－卷積神經網路

02. 特點

卷積神經網路概述：卷積層(Convolution layer)

對不同影像上，若相同的特徵出現在不同的區域依然可以找到，因共用相同的參數

權重共享



機器學習 – 卷積神經網路

02. 特點

卷積神經網路概述：池化層(Pooling layer)

影像作再取樣（縮小或放大）不會改變物件。我們可以對像素進行欠取樣(Undersampling)使影像大小更小，讓運算數據量更少



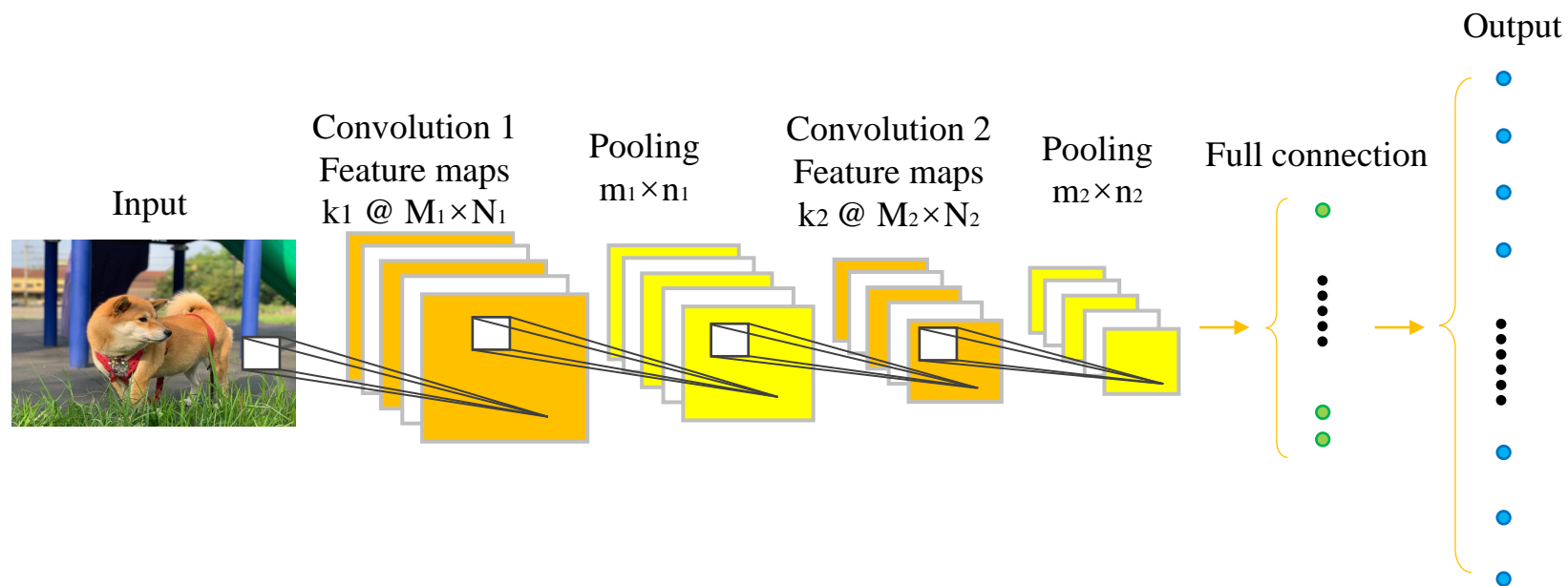
Undersampling
→



機器學習 – 卷積神經網路

02. 特點

卷積層與池化層可以多層，最終提取出來的特徵皆由全連接層連接至輸出層



特點 1

- 尋找局部特徵

特點 2

- 相同的特徵能在不同的區域尋找

特點 3

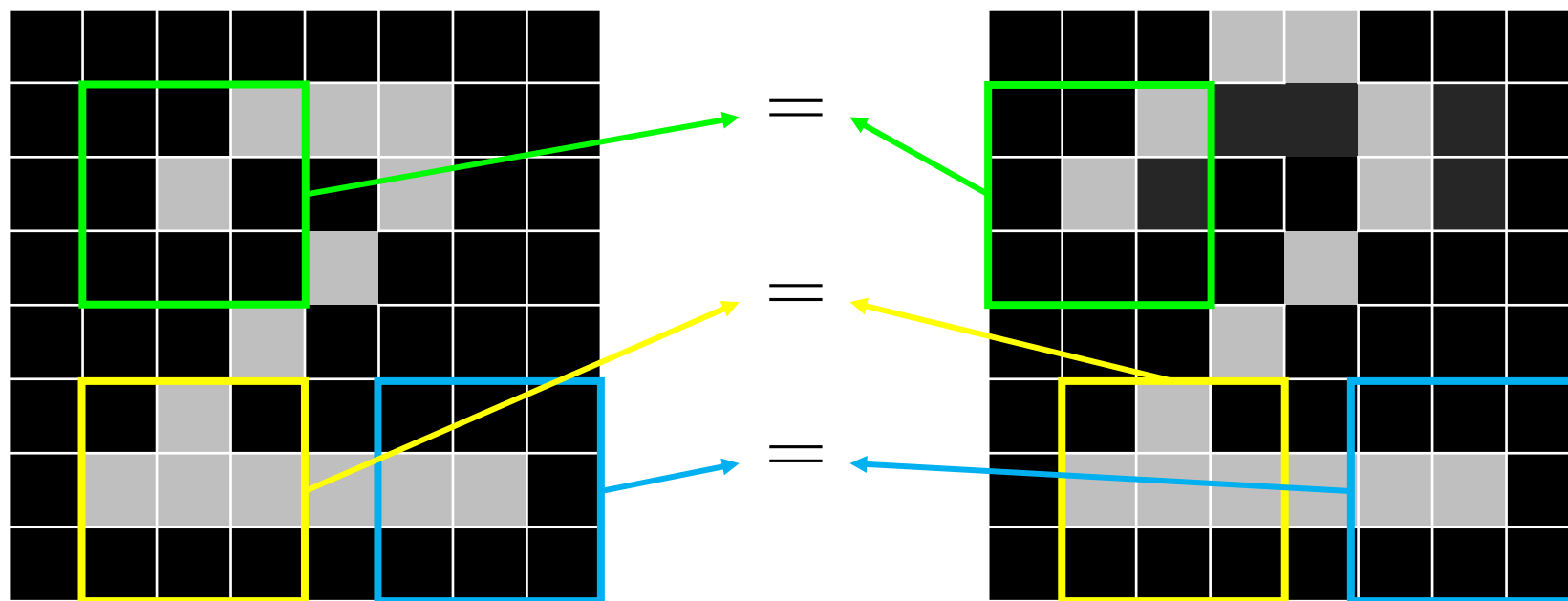
- 對像素進行再取樣不會改變物件(縮小)，減少計算量

機器學習 – 卷積神經網路

03. 卷積層

卷積層(Convolution layer)

卷積層如何利用特徵來判斷？



機器學習 – 卷積神經網路

04. 卷積與填充

卷積層(Convolution layer)

-1	-1	-1
-1	-1	1
-1	-1	-1

濾波器

-1	-1	-1
-1	-1	1
-1	-1	-1

影像局部區域

$$\begin{aligned} & [(-1) \times (-1) + (-1) \times (-1) + (-1) \times (-1) + (-1) \times (-1) + (-1) \\ & \times (-1) + 1 \times 1 + (-1) \times (-1) + (-1) \times (-1) + (-1) \times (-1)] / 9 = 1 \end{aligned}$$

遮罩(Mask)或稱窗口 在原始影像上移動，濾波器(Filter)或稱卷積核內的值對遮罩內的值作內積計算，這過程稱為卷積(Convolution)或稱旋積

- 深度(Depth)：k個濾波器作卷積，深度即為k
- 步長(Stride)：每次移動遮罩的距離
- 零填充(Zero-padding)：原始影像周圍補上零

-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	1	1	1	-1	-1
-1	-1	1	-1	-1	1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1
-1	-1	-1	1	-1	-1	-1	-1
-1	1	1	1	1	1	1	-1
-1	-1	-1	-1	-1	-1	-1	-1

濾波器

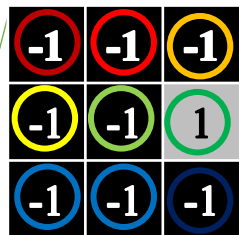
影像

- 濾波器大小：3×3 卷積後得到的圖像大小為：3×3
- 步長為 1
- 沒有做填充

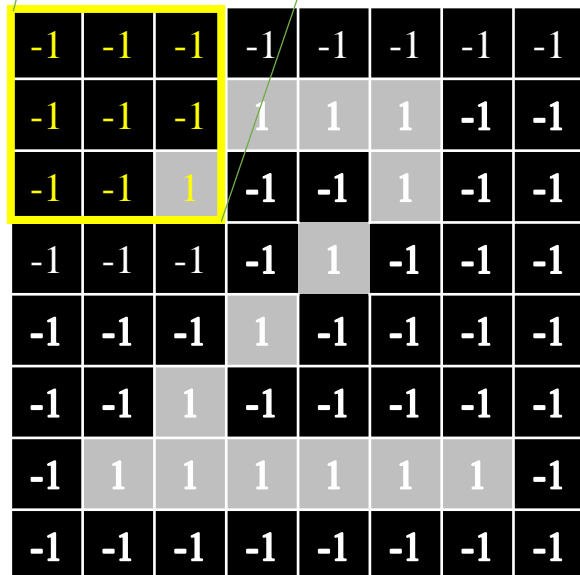
$$W_{output} = (W_{input} - F + 2P)/S + 1$$

$$= (8 - 3 + 2 \times 0)/1 + 1 = 6$$

- W_{output} 是卷積後特徵圖的寬(高)度
- W_{input} 是卷積前圖像寬(高)度
- F是濾波器的寬(高)度
- P是填充補幾圈
- S是步長

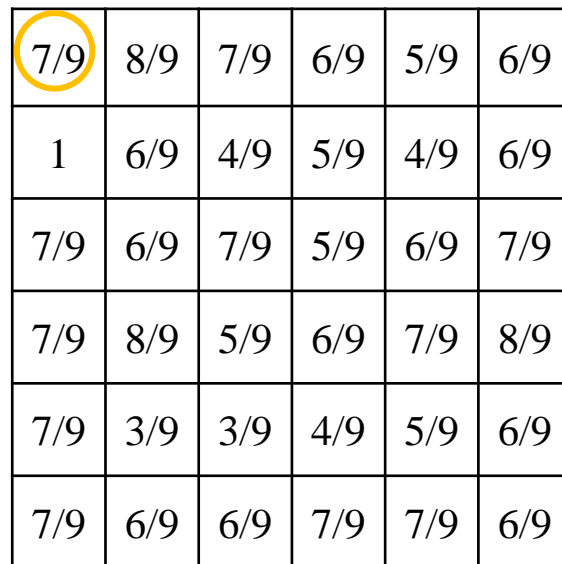


濾波器

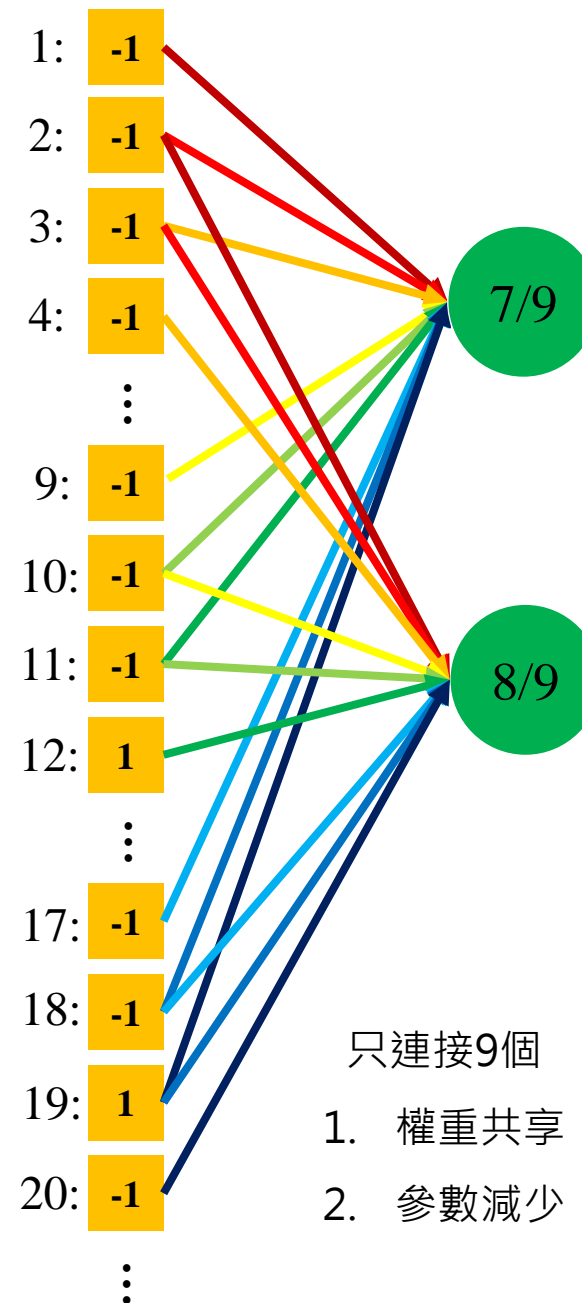


影像

未做填充的影像其特徵圖將變小



特徵圖



只連接9個

1. 權重共享
2. 參數減少

機器學習 – 卷積神經網路

05. 自動提取特徵

濾波器



⋮



-1	-1	-1
-1	-1	1
-1	-1	-1

具有特徵的濾波器內的權重其數值是如何得到？

其實一開始都設置初始權重，因此一開始特徵並不是最好的特徵，因此模型預測與標籤所計算的誤差會非常高。但透過梯度下降法更新濾波器的權重，隨著訓練多次迭代，它會自動趨近於它認為最好的特徵，這CNN該特點稱為自動提取特徵

機器學習 – 卷積神經網路

06. 池化層

池化層(Pooling layer)

- 局部最大池化(Max Pooling)
- 局部平均池化(Average Pooling)
- 局部隨機池化(Stochastic Pooling)
- 全域最大池化(Global Max Pooling)
- 全域平均池化(Global Average Pooling) [1]

- 池化後的深度沒變
- 全域池化滑窗大小與整張特徵圖的大小一樣大

[1] Min Lin, Qiang Chen, Shuicheng Yan, “Network In Network,” arXiv:1312.4400, 2014.

特徵圖

0.5	0.9	0.1	0.4	0	1
1	0	0.4	0	0.5	0
0.4	0.8	0.3	1	0.7	0.9
0.2	0.6	0.5	0.2	0.1	0.3
1	0.7	1	0.1	0.4	0.8
0	0.8	0.2	0	0.3	1

Max Pooling
(Size: 2×2, Stride: 2)

1.0	0.4	1.0
0.8	1.0	0.9
1.0	1.0	1.0

Average Pooling
(Size: 2×2, Stride: 2)

0.60	0.23	0.38
0.50	0.50	0.50
0.63	0.33	0.63

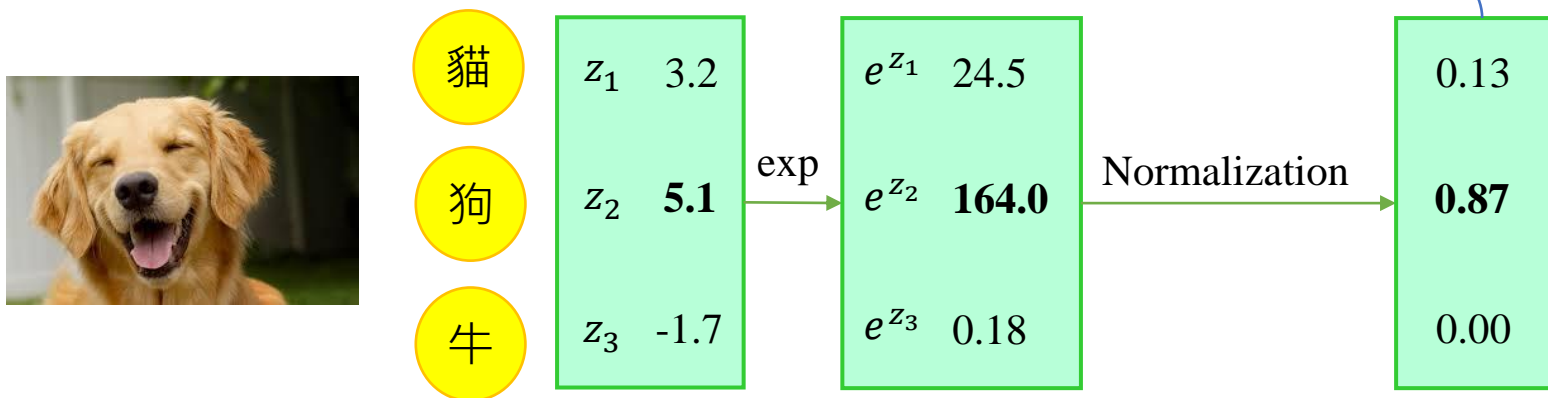
機器學習 – 卷積神經網路

07. 全連接層

全連接層(Fully Connected Layer)

將所有特徵圖變為一維度，並組合在一起來進行輸出分類。通常在全連接層與輸出層間會使用 Softmax 函數來輸出機率，使所有類別的機率和為 1

$$y_j = f_{Softmax}(z_j) = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad \text{for } j = 1, \dots, K$$



$$\frac{e^{z_1}}{e^{z_1} + e^{z_2} + e^{z_3}} = \frac{24.5}{24.5 + 164.0 + 0.15} = 0.13$$
$$\frac{e^{z_2}}{e^{z_1} + e^{z_2} + e^{z_3}} = \frac{164.0}{24.5 + 164.0 + 0.15} = 0.87$$
$$\frac{e^{z_{35}}}{e^{z_1} + e^{z_2} + e^{z_3}} = \frac{0.15}{24.5 + 164.0 + 0.15} = 0.00$$

機器學習 – 卷積神經網路

08. 超參數



卷積神經網路常用的超參數

- 學習速率 (Learning Rate)
- 批量 (Batch)
- 迭代次數 (Number of iterations)
- 正則化懲罰係數 (Regularization)
- 隱藏層層數與神經元數量(卷積層、池化層與全連接層)
- 丟棄法 (Dropout)
- 卷積層的濾波器大小與數量
- 池化層的濾波器與其大小



卷積神經網路－實作

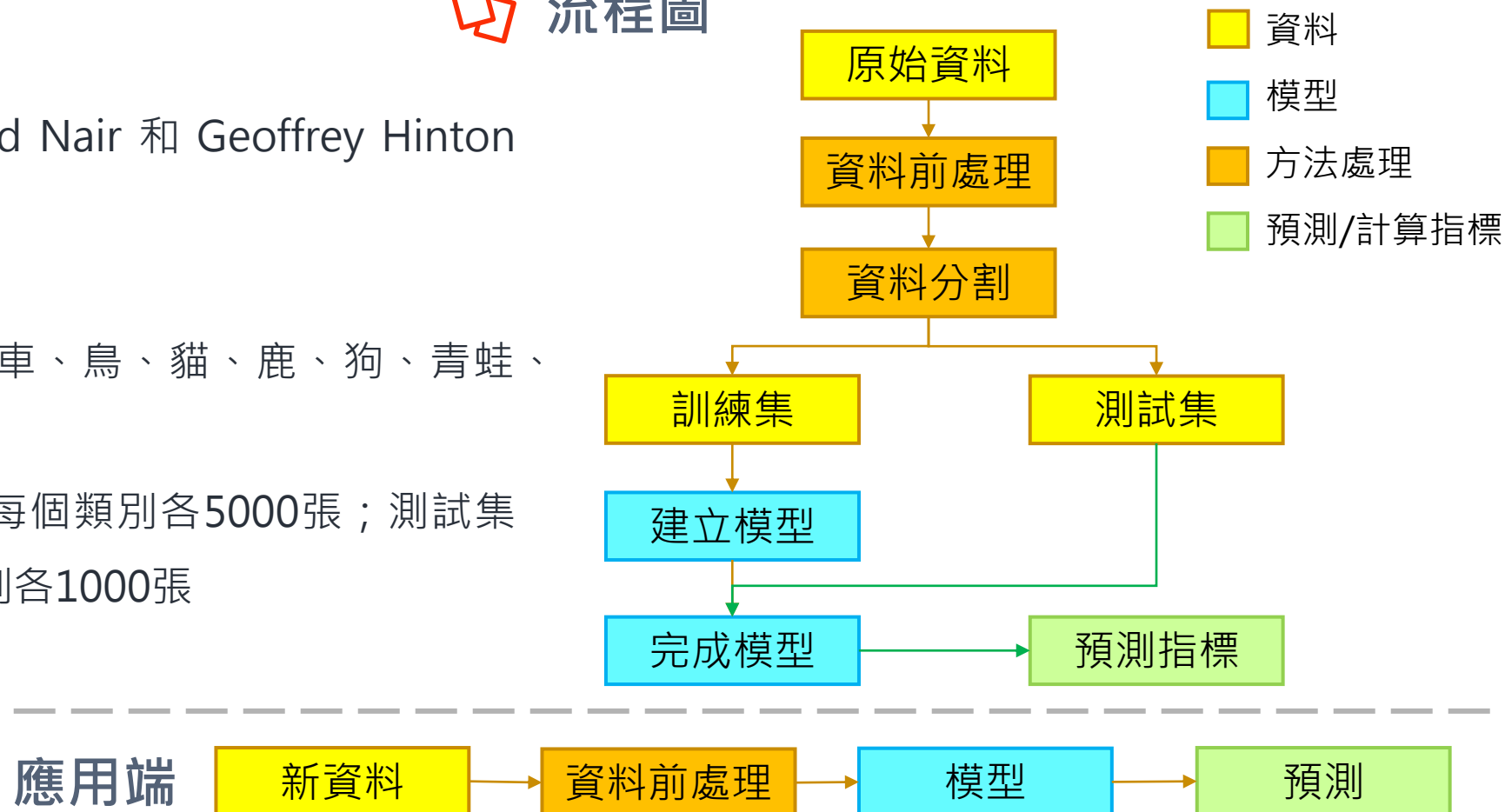
卷積神經網路 – 實作

Cifar10

Cifar10 Dataset

- 由 Alex Krizhevsky, Vinod Nair 和 Geoffrey Hinton 蒐集的資料集
- 60000張32x32彩色圖像
- 包含十種類別：飛機、汽車、鳥、貓、鹿、狗、青蛙、馬、船與卡車
- 訓練集為50000張，其中每個類別各5000張；測試集為10000張，其中每個類別各1000張

流程圖



卷積神經網路 – 實作

Cifar10

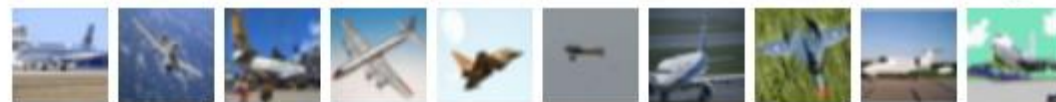
問題

1. 如果影像的資料集太小，該如何解決呢？
2. 超參數該如何選擇？

作業一

1. 請使用資料增量方式以提高模型的預測能力
2. 嘗試訓練Cifar100建立CNN分類器

airplane



automobile



bird



cat



deer



dog



frog



horse



ship



truck

