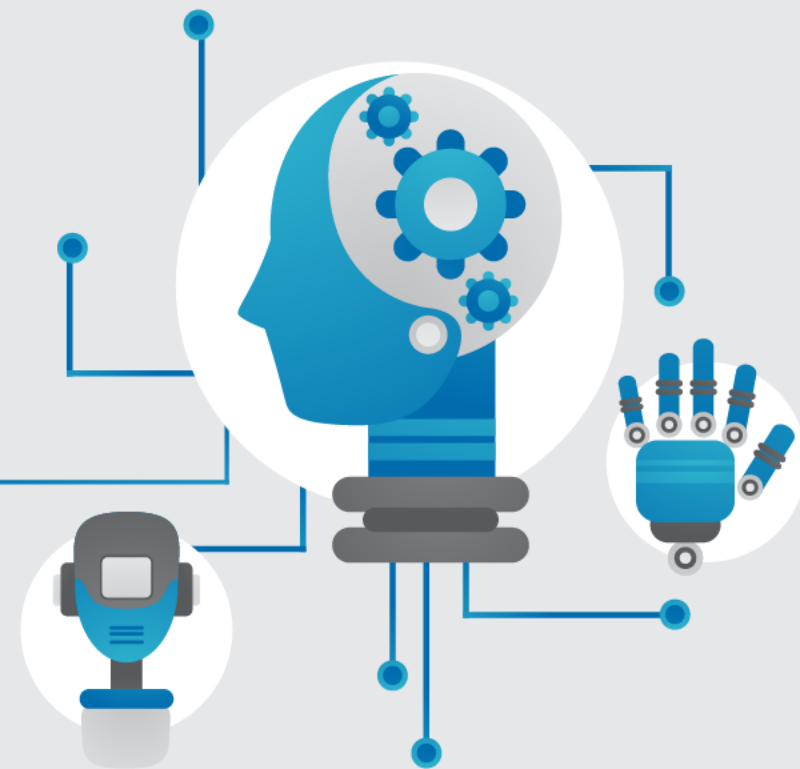


卷積式神經網路概念





MLP模型訓練之缺點



- › 輸入的圖像檔需轉換成一維陣列，
輸入層和第一個隱藏層之間的權重數量非常大。

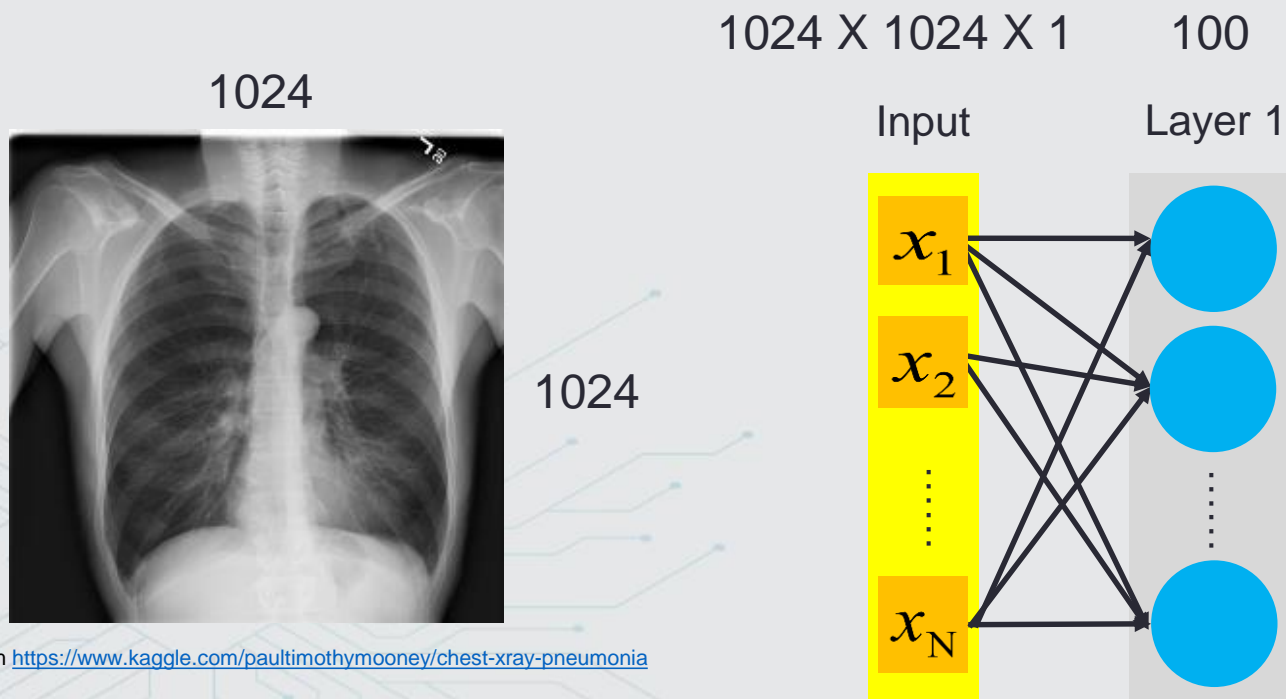


Photo from <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>

$$1024 \times 1024 \times 100 \cong 10^8 \text{ weights}$$



降低圖像訓練的複雜度



- › 僅萃取圖像中重要特徵 (feature) 進行訓練
- › 特徵是由相鄰像素 (pixel) 所構成
- › 使用卷積運算萃取特徵

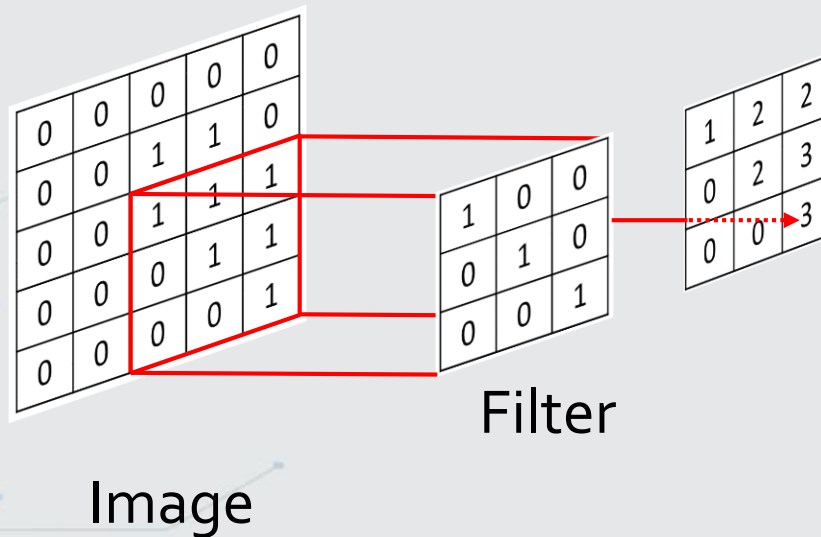
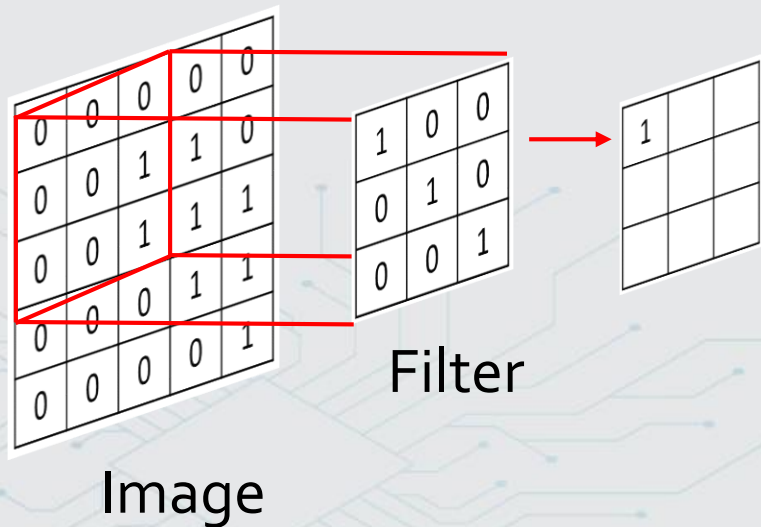




卷積運算



- › 將每個像素及其鄰近像素由一個濾波器 filter (kernel) 加權後加總起來。
- › 對每個像素執行此卷積過程

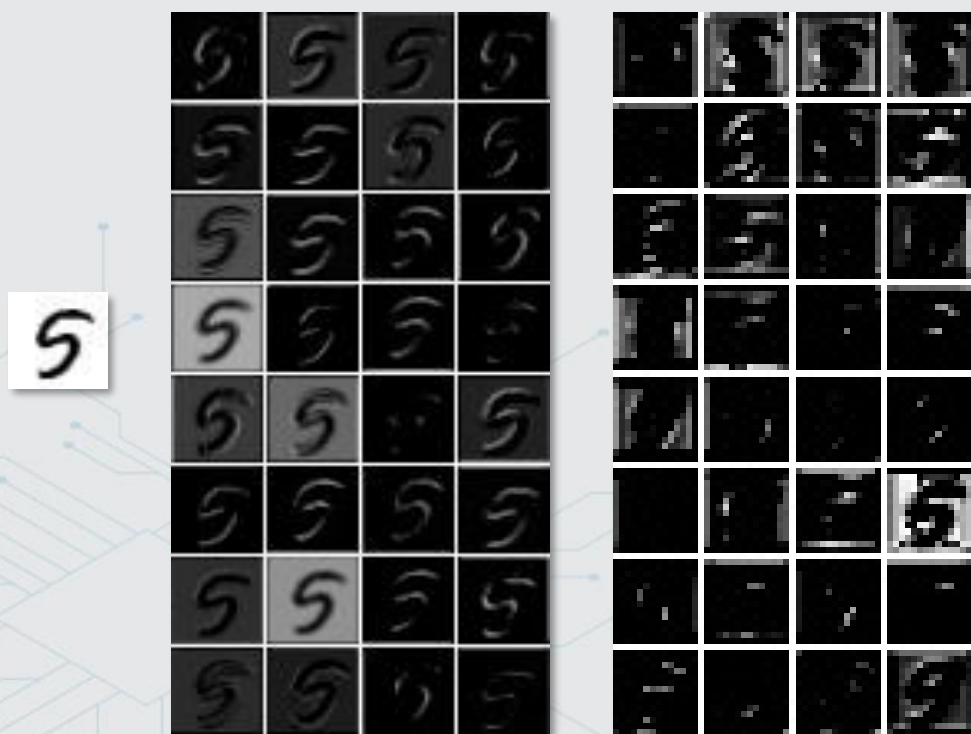




卷積運算



› 兩層卷積的feature maps



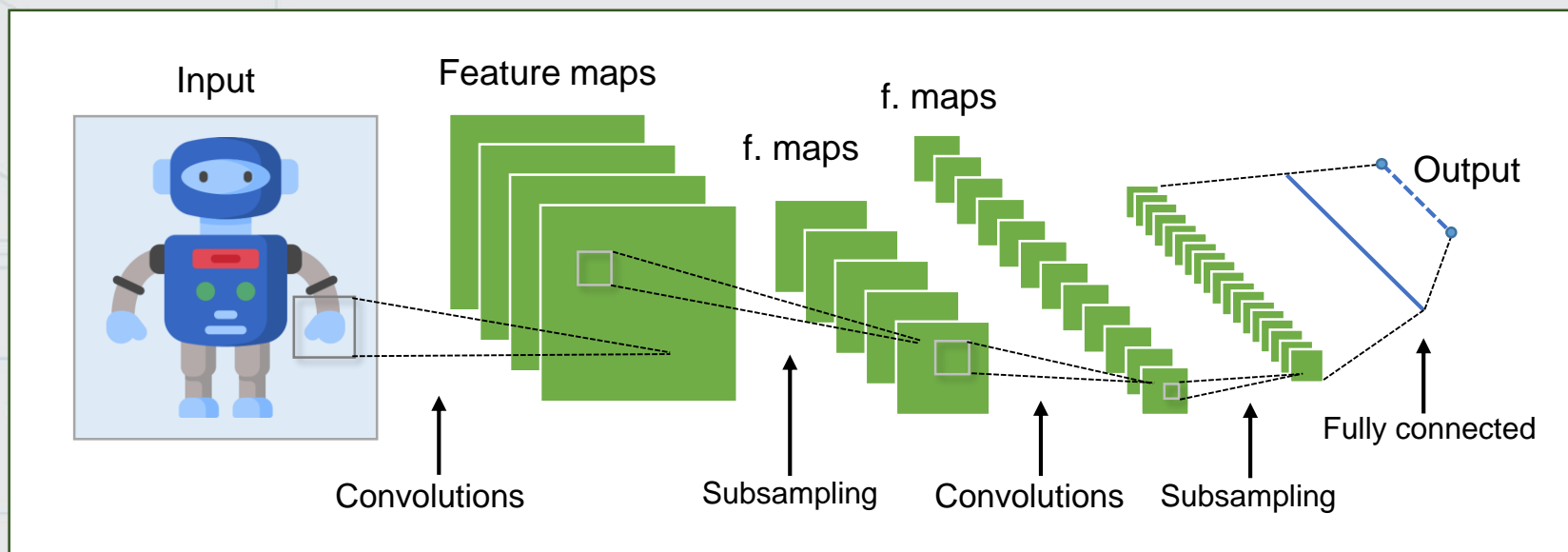


卷積神經網路



- › 卷積神經網路 (Convolutional Neural Network)
由輸入和輸出層以及多個隱藏層組成。

CNN的隱藏層通常由卷積層、池化層、激活層、平坦層和完全連接層組成。





卷積層的超參數

機器學習實務



Zero-padding (0-填充)

Stride (跨步)

Depth (深度)



Zero-padding



› 在圖像的邊界添加額外的零

Image

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

Filter

1	0	0
0	1	0
0	0	1

New image

0	1	1	0	0
1	1	2	2	0
2	1	2	3	2
0	3	1	3	2
0	0	2	0	2



Stride



- › 窗格滑動的步距
- › 縮小卷積層的輸出

Image

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

Filter

1	0	0
0	1	0
0	0	1

New image

1	2
3	3

Set stride = 2



Stride



- › 滑動窗口的步距
- › 縮小卷積層的輸出

Image

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

Filter

1	0	0
0	1	0
0	0	1

New image

1	2
3	3

Set stride = 2



Stride



- › 滑動窗口的步距
- › 縮小卷積層的輸出

Image

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

Filter

1	0	0
0	1	0
0	0	1

New image

1	2
3	3

Set stride = 2



Stride



- › 滑動窗口的步距
- › 縮小卷積層的輸出

Image

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

Filter

1	0	0
0	1	0
0	0	1

New image

1	2
3	3

Set stride = 2



Stride



- › 滑動窗口的步距
- › 縮小卷積層的輸出

Image

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

Filter

1	0	0
0	1	0
0	0	1

New image

1	2
3	3

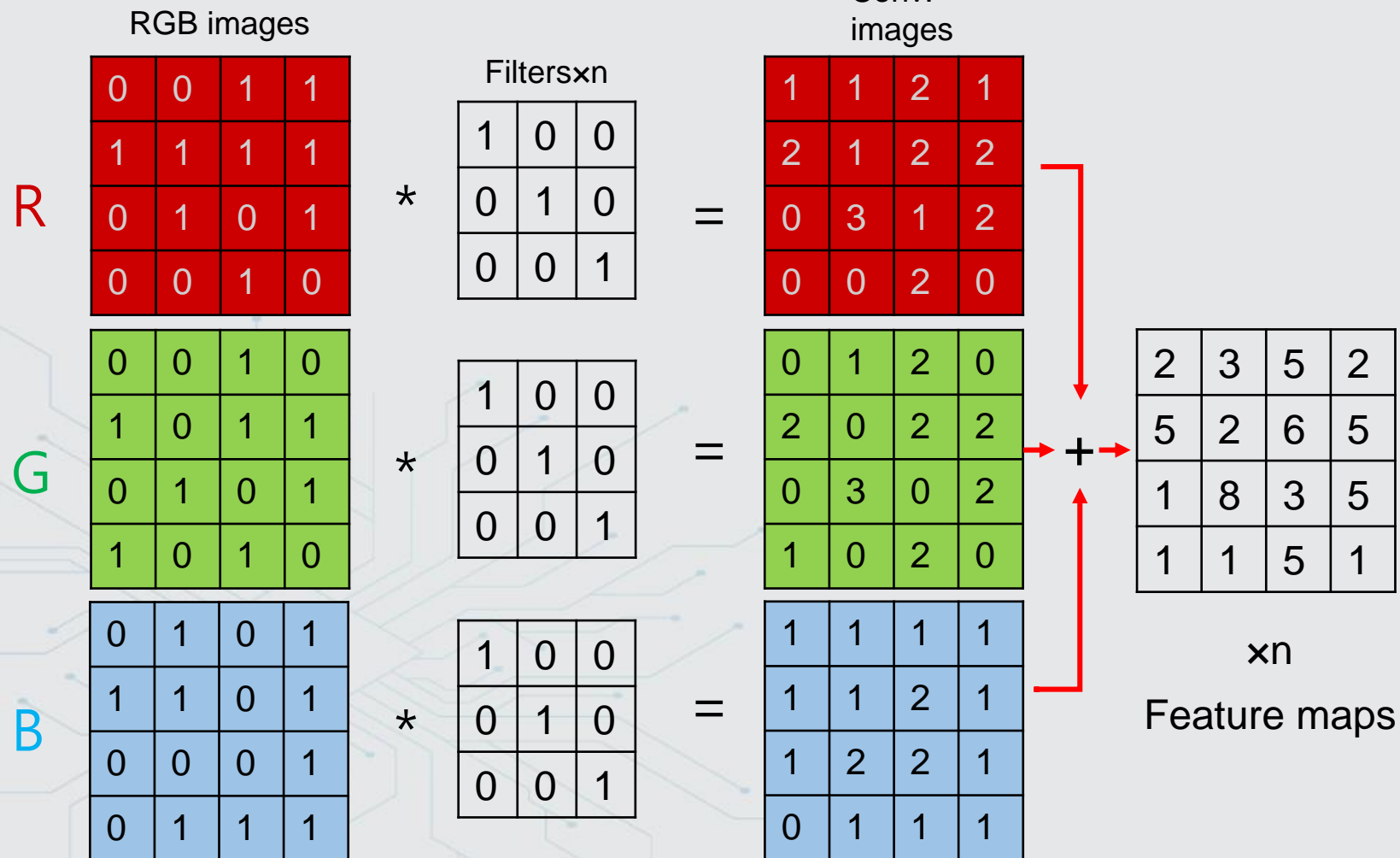
Set stride = 2



Depth



› Depth就是filter的個數





Pooling



› 池化層的優勢

- 減少權重
- 防止過度學習

› 最大池化

- 保留每個區域中之最大值

1	2	2	0
1	2	3	2
3	1	3	2
0	2	0	4

Max pooling



2	3
3	4



超參數的設置



- › 影像的大小須要能夠被 2 整除數次
 - 常見的影像 size 32, 64, 96, 224, 384, 512
- › 卷積層filter和stride的選擇
 - 比起使用一個 size 較大的 filter (7x7)，可先嘗試連續使用數個 size 小的 filter (3x3)
 - Stride 的值與 filter size 相關，通常 $stride \leq (W_f - 1) / 2$
- › 很深的CNN model (16+層) 多使用 3x3 filter 與 stride 1。



超參數的設置



- › Zero-padding 與 pooling layer 是選擇性的結構
- › Zero-padding 的使用取決於是否要保留邊界的資訊
- › Pooling layer 目的在避免overfitting與降低weights的數量，但也減少影像所包含資訊，一般不會大於3x3
- › 嘗試修改不錯效能的模型，會比建立一個全新的模型容易收斂，且模型的權重數量愈多，愈難調整出好的參數。