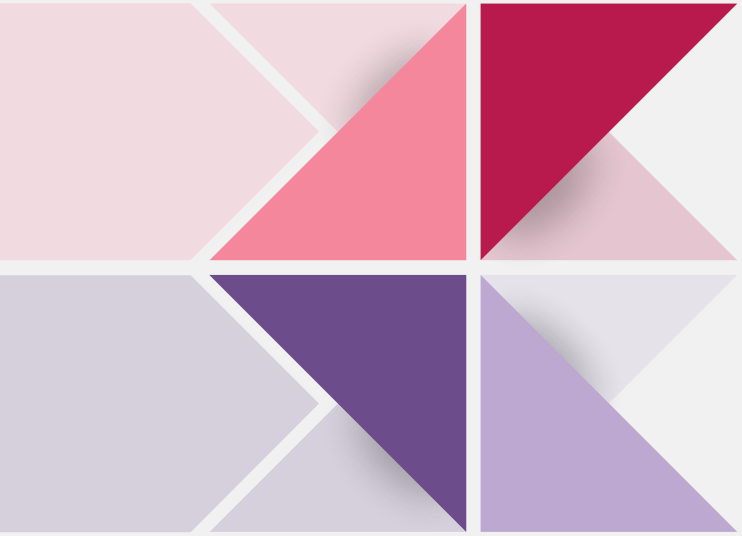
深度學習

Deep Learning



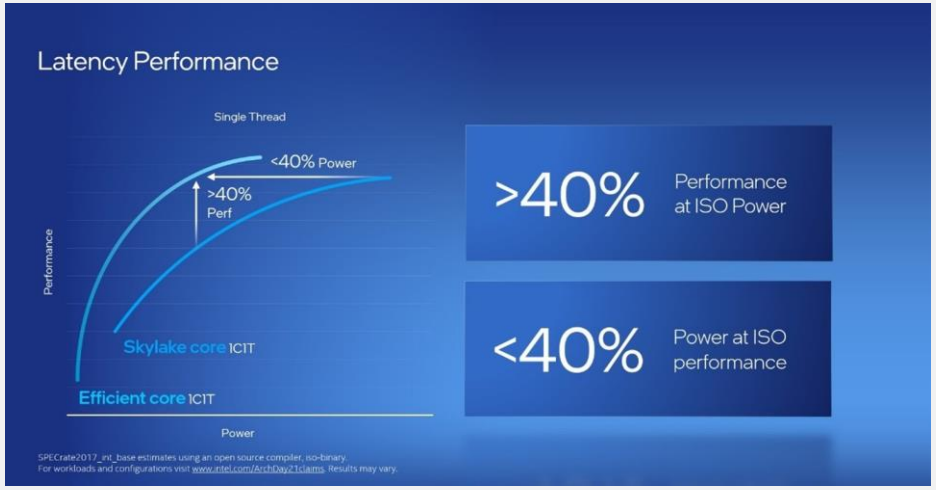
1

為什麼深度學習會成為主流?

為什麼深度學習會成為主流?

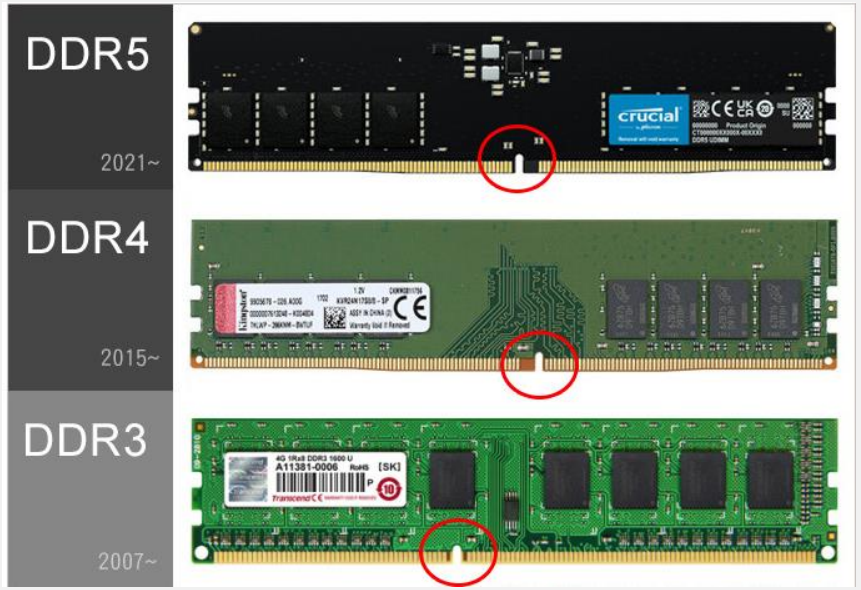
1. 硬體的進步

CPU



資料來源

RAM



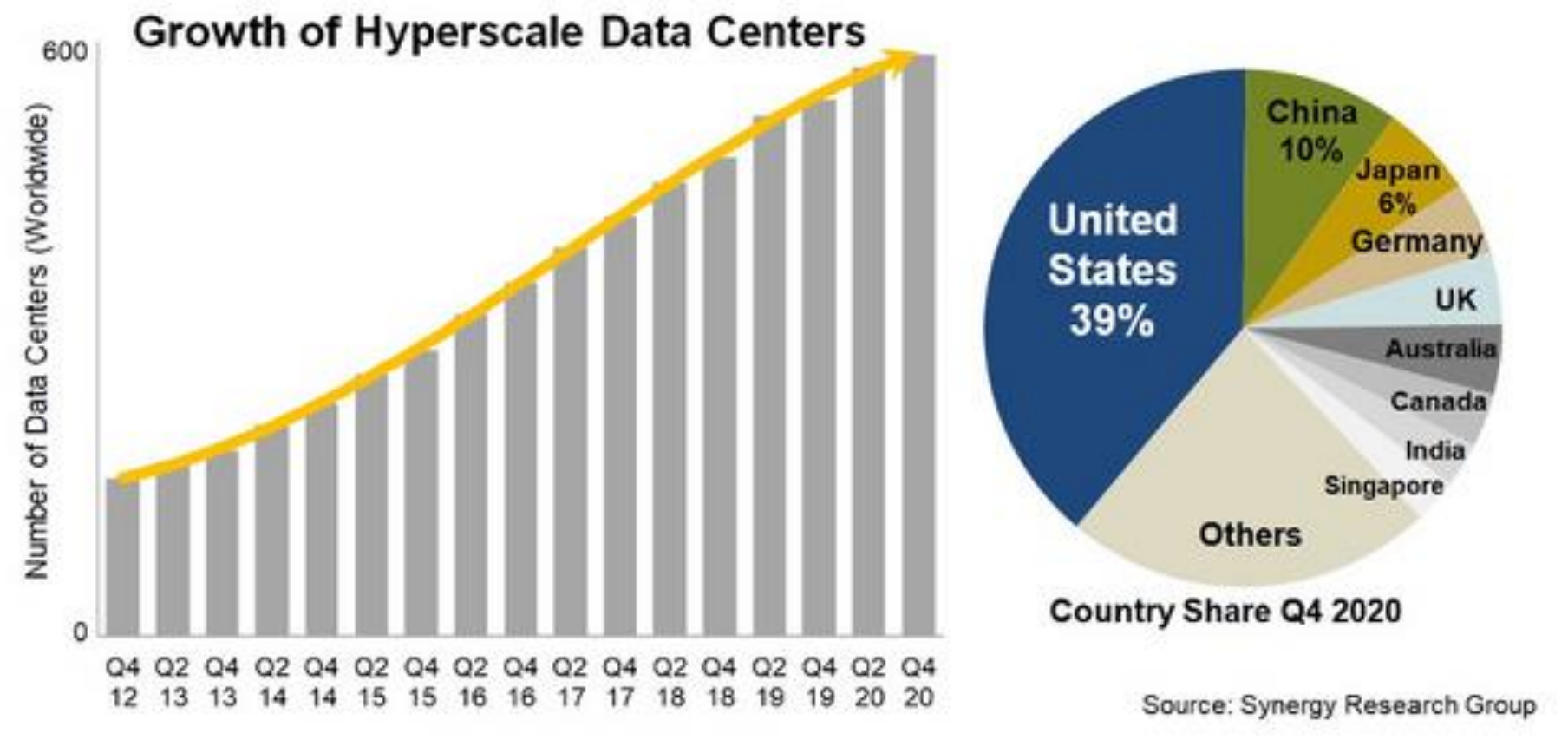
資料來源

GPU

資料來源

為什麼深度學習會成為主流?

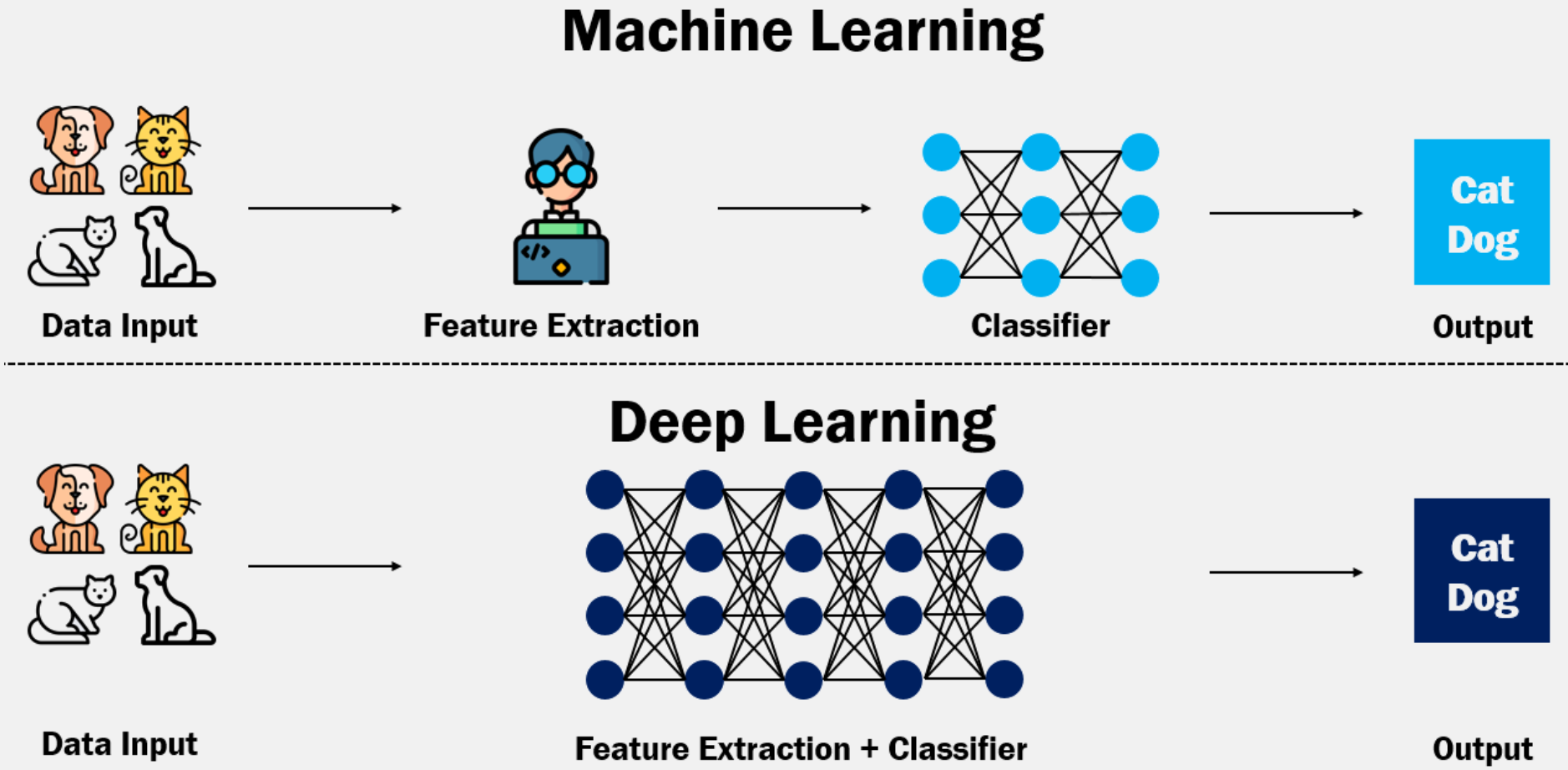
2. 資料的成長



資料來源

為什麼深度學習會成為主流?

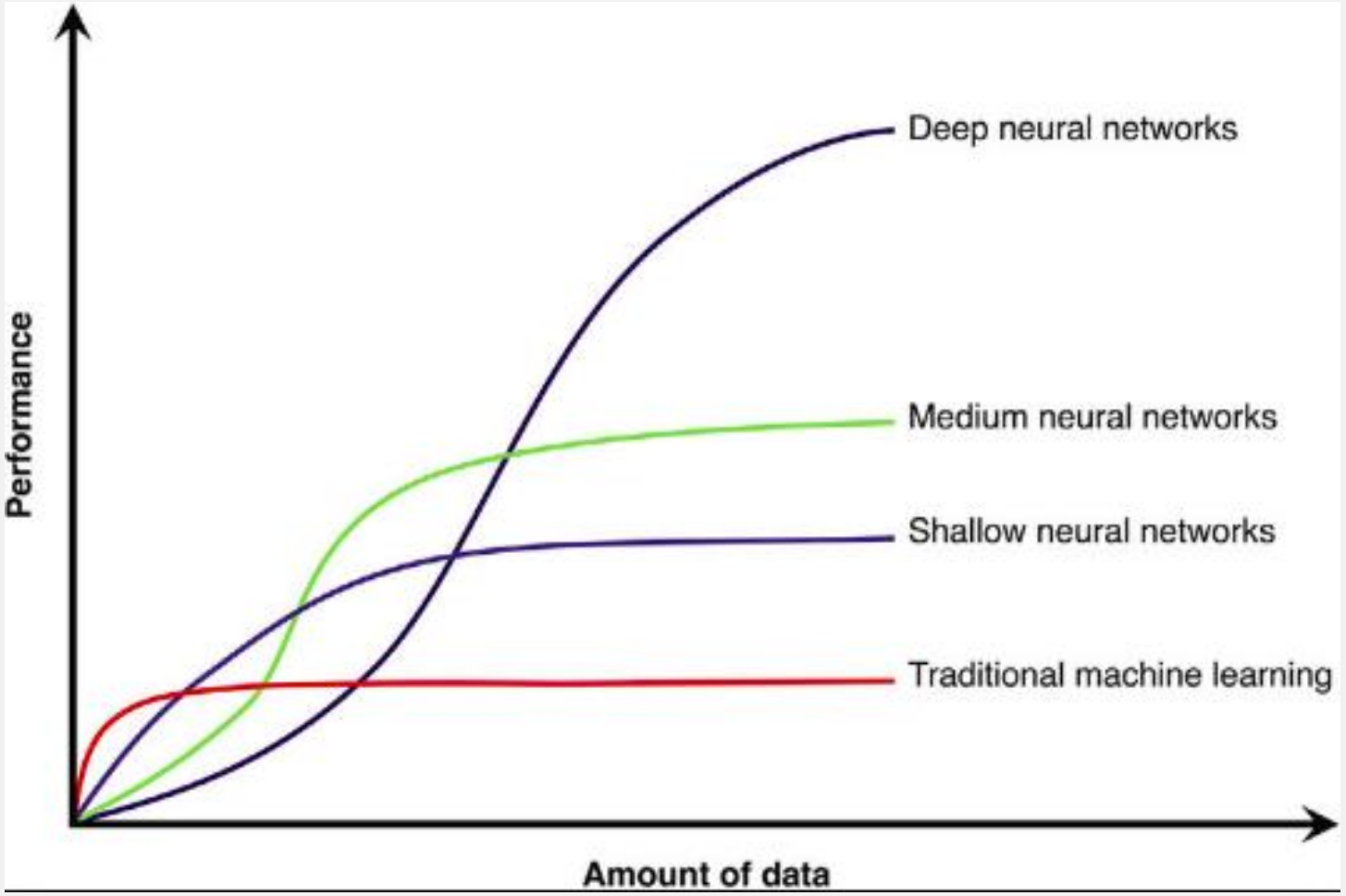
3. 技術的變化

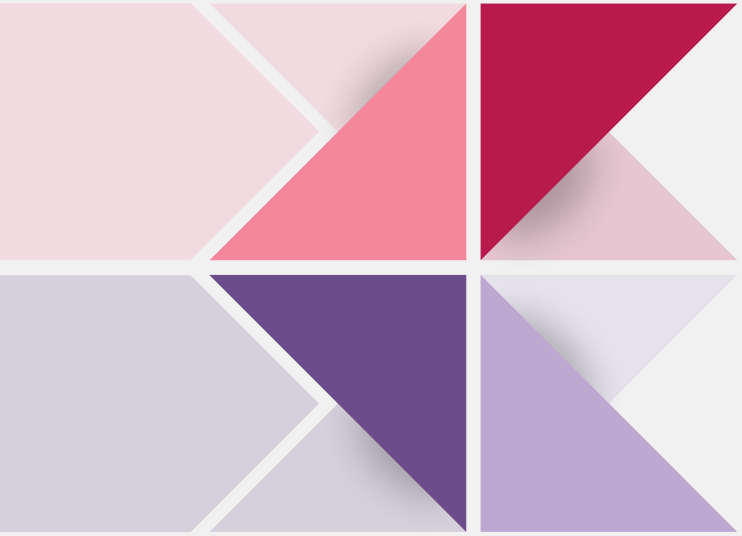


為什麼深度學習會成為主流?

3. 技術的變化

資料量增加時，不同技術的效能比較
(Machine vs Deep Learning)

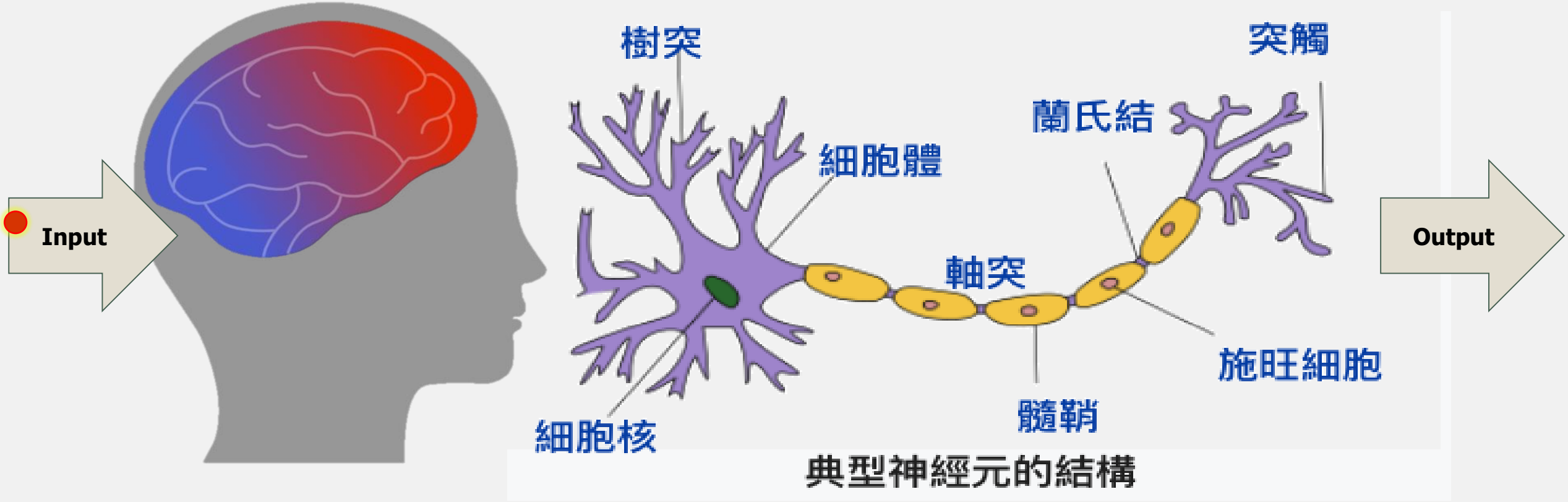




2

深度學習是什麼技術?

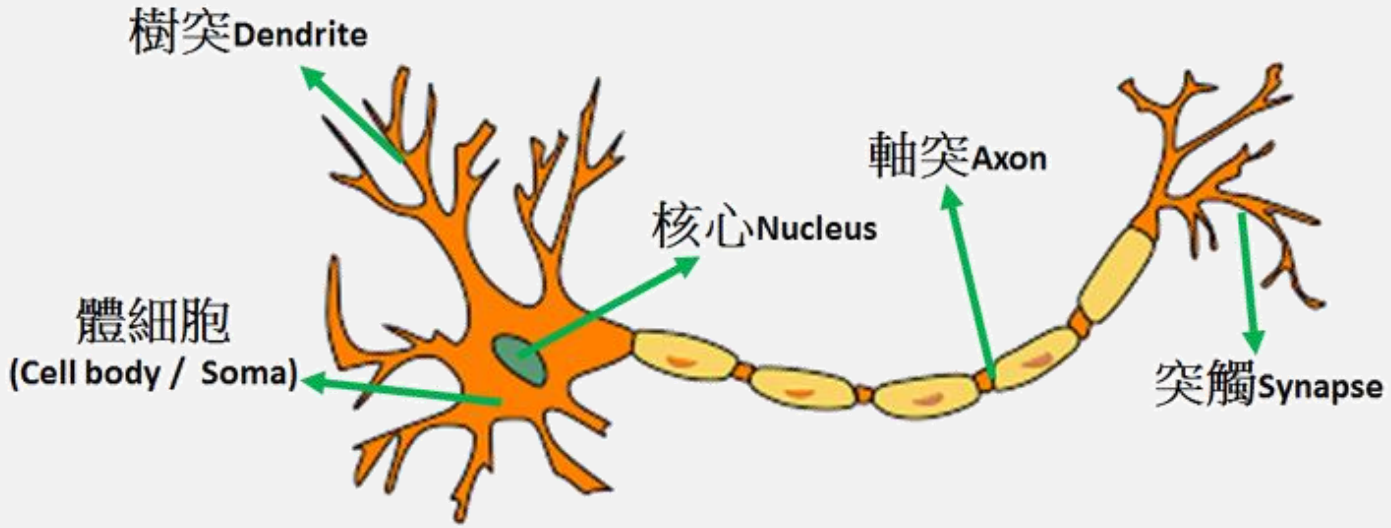
深度學習是什麼技術?

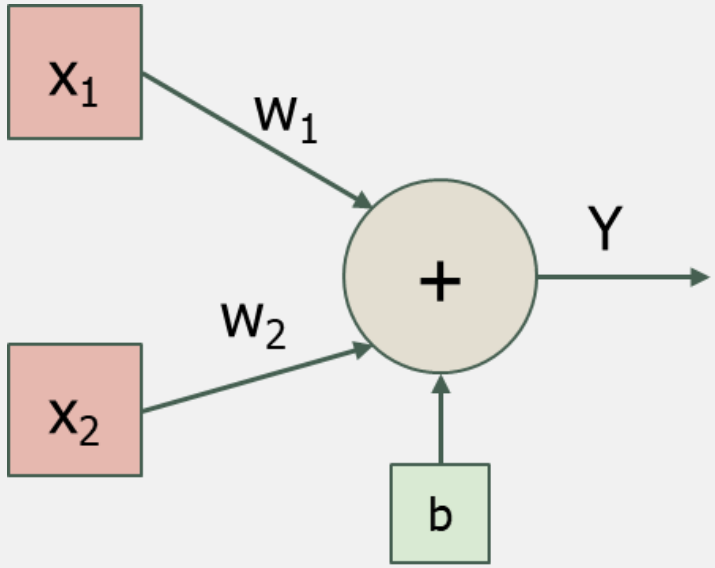


典型神經元的結構

資料來源

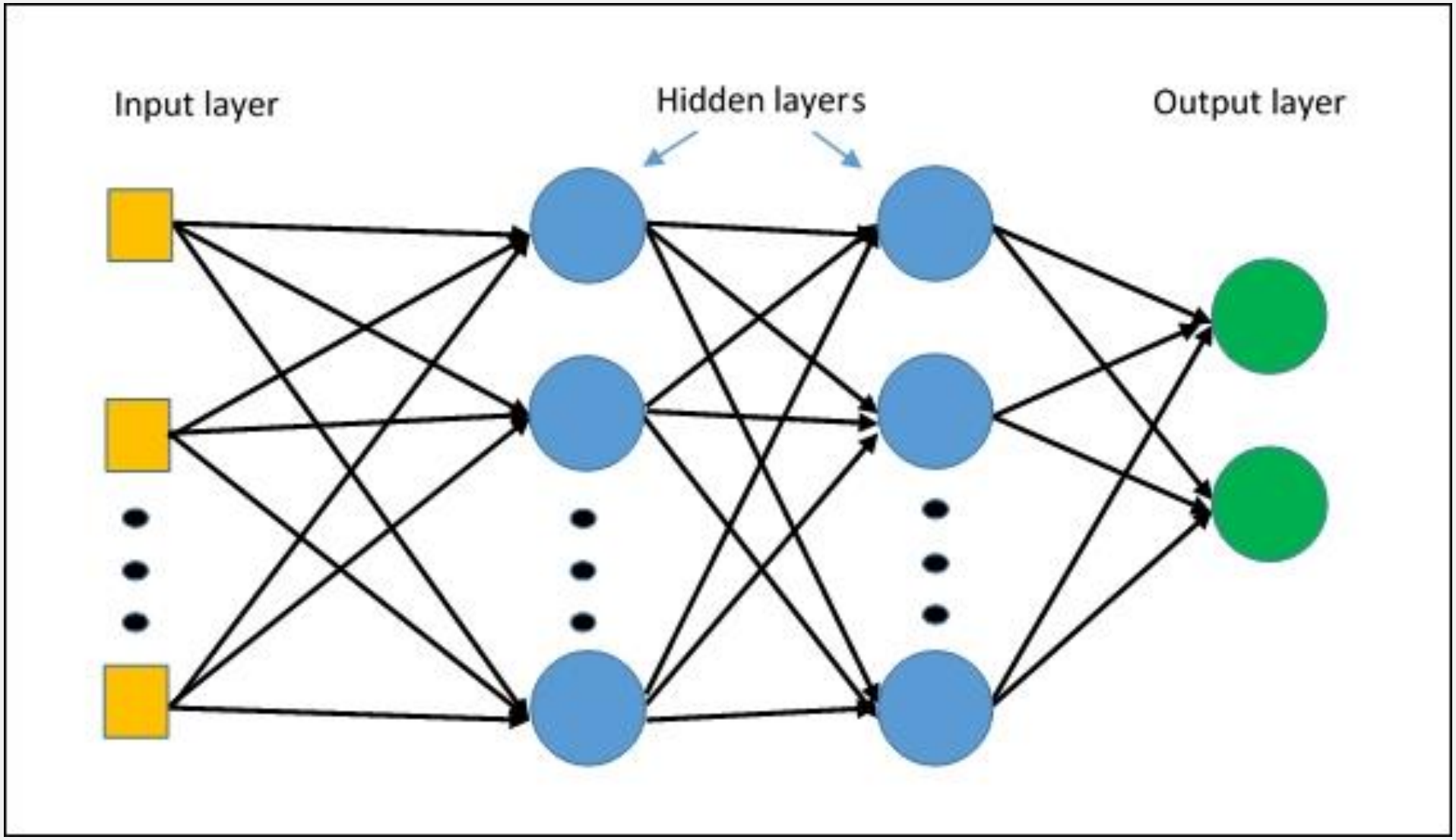
深度學習是什麼技術?





- Neurons:
  - Have *K* inputs (*dendrites*).
  - Have 1 output (*axon*).
  - If the sum of the input signals surpasses a *threshold*, sends an *action potential* to the axon.
- Synapses
  - Transmit electrical signals between neurons.

深度學習是什麼技術?

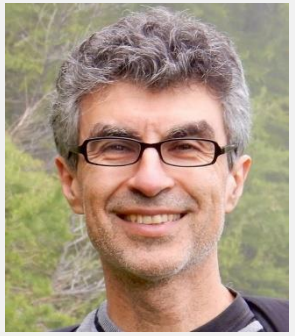


深度學習是什麼技術?

LeCun、Bengio、Hinton, Deep Learning, Nature, 2015

- **https://www.cs.toronto.edu/~hinton/absps/NatureDeepReview.pdf**
- **Yann LeCun, 紐約大學、Facebook人工智慧**
- **Yoshua Bengio, 蒙特婁大學**
- **Geoffrey Hinton, 多倫多大學、Google Brain**

'Godfathers of AI' honored with 2019 Turing Award



REVIEW

doi:10.1038/nature14539

Deep learning

Yann LeCun[1,2], Yoshua Bengio[3] & Geoffrey Hinton[4,5]

卷積類神經網路

ImageNet (ILSVRC)

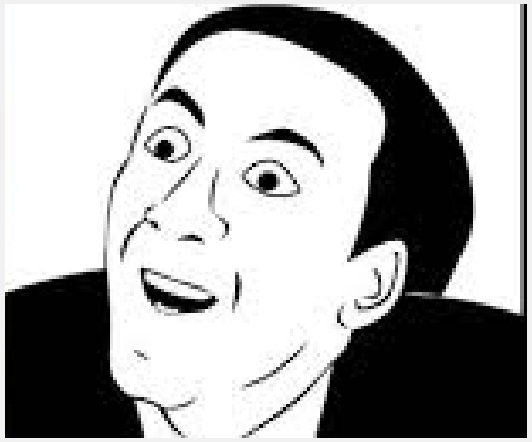
- 李飛飛 等人
- **1000** 個類別 以及 **1.2M** 張影像
- **~2017**

EntleBucher

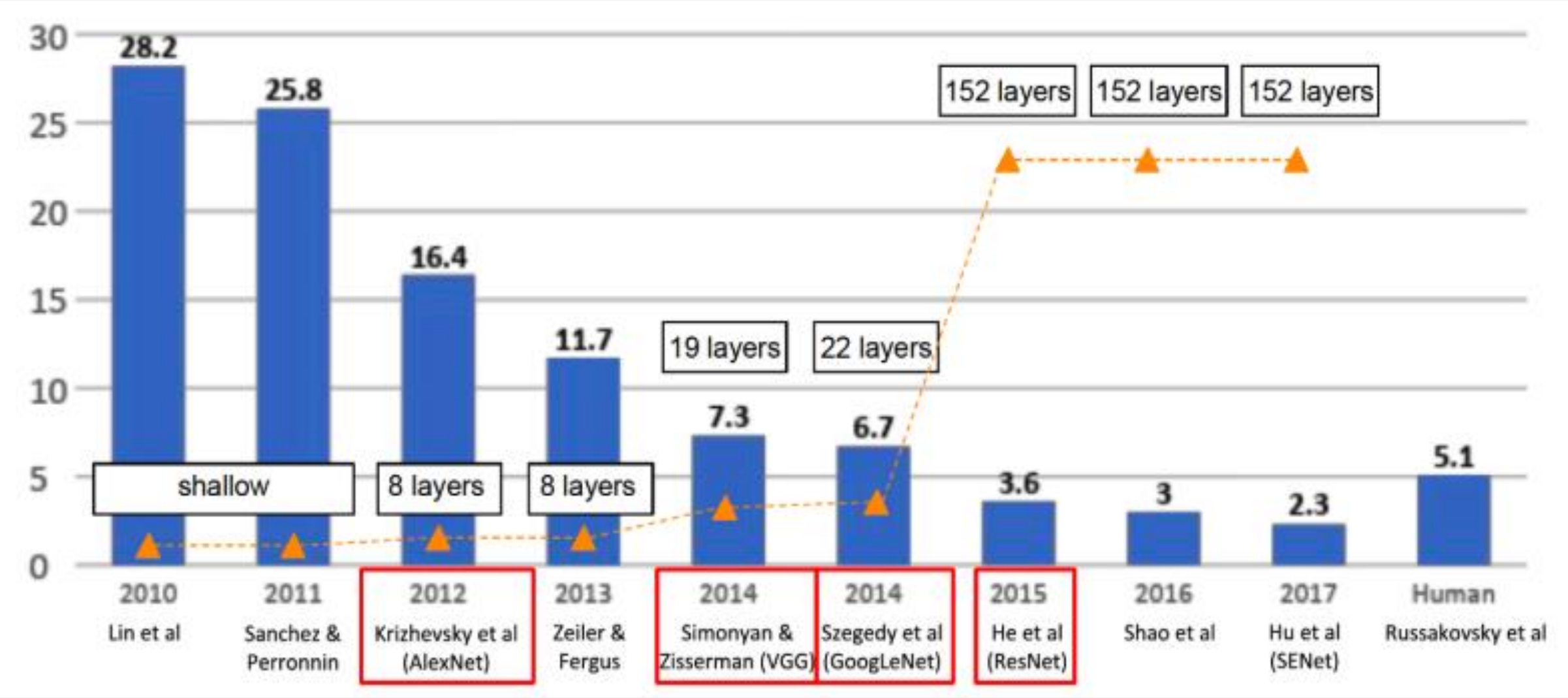
恩特布山犬

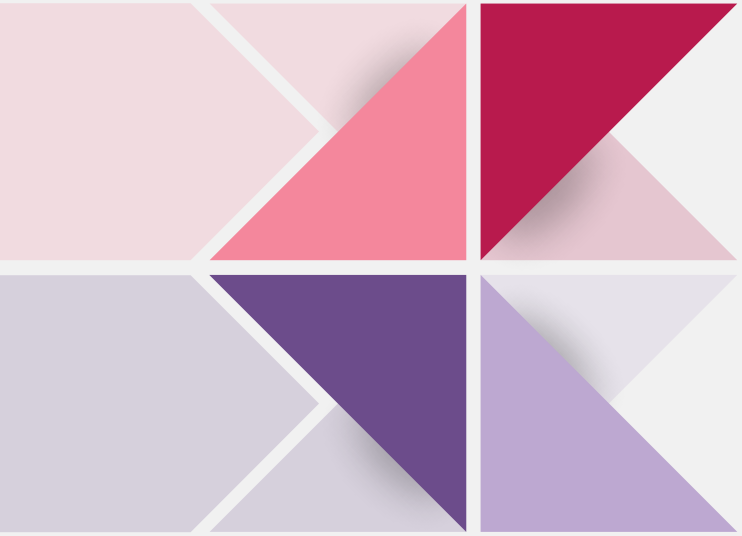
Appenzeller

阿彭策爾山犬



卷積類神經網路


30
25
20
15
10
5
0
28.2
25.8
16.4
11.7
7.3
6.7
3.6
3
2.3
5.1
shallow
8 layers
8 layers
19 layers
22 layers
152 layers
152 layers
152 layers
2010
Lin et al
2011
Sanchez & Perronnin
2012
Krizhevsky et al (AlexNet)
2013
Zeiler & Fergus
2014
Simonyan & Zisserman (VGG)
2014
Szegedy et al (GoogLeNet)
2015
He et al (ResNet)
2016
Shao et al
2017
Hu et al (SENet)
Human
Russakovsky et al



3

卷積類神經網路

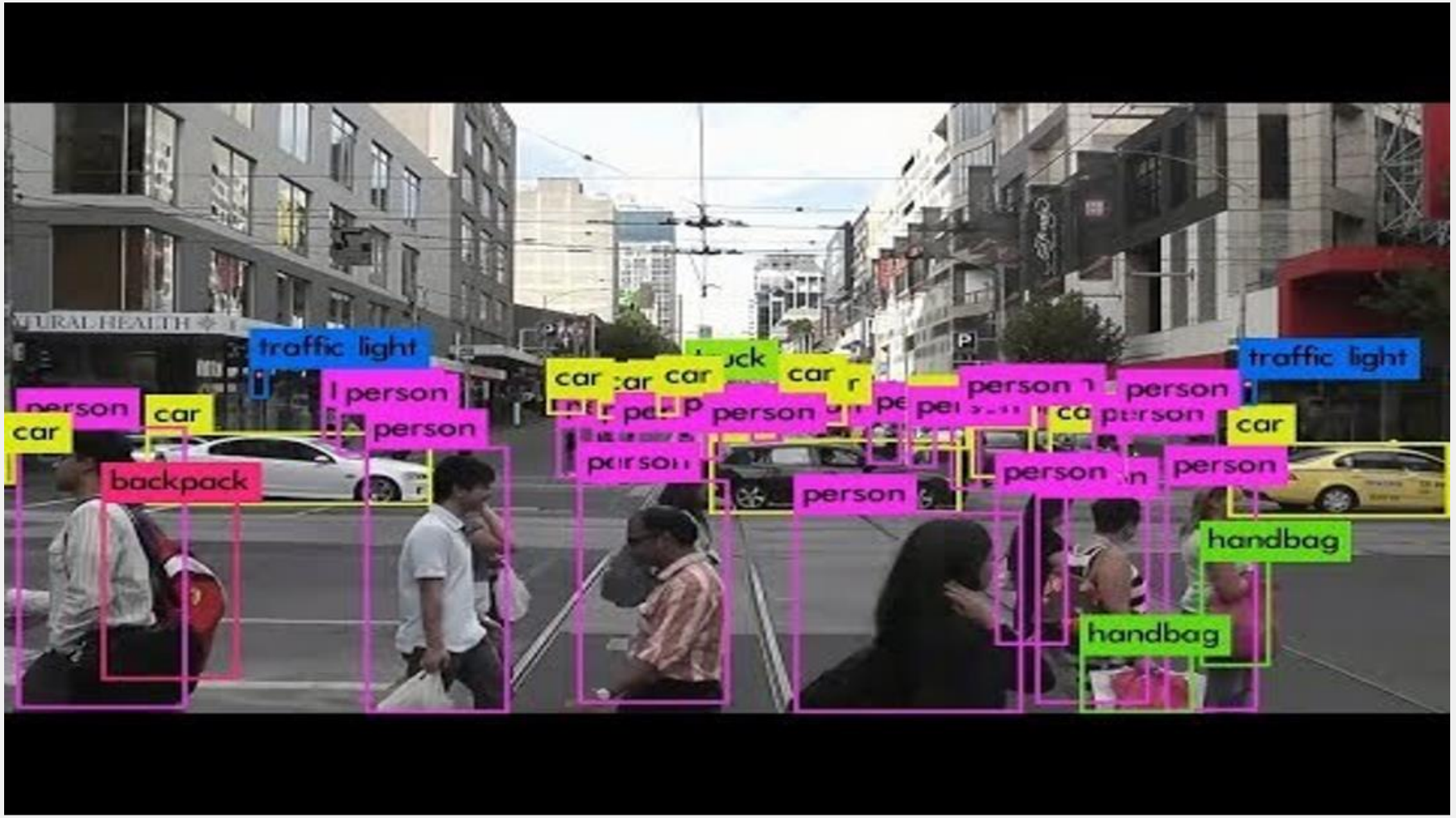
卷積類神經網路

數字辨識

資料來源

卷積類神經網路

道路即時偵測



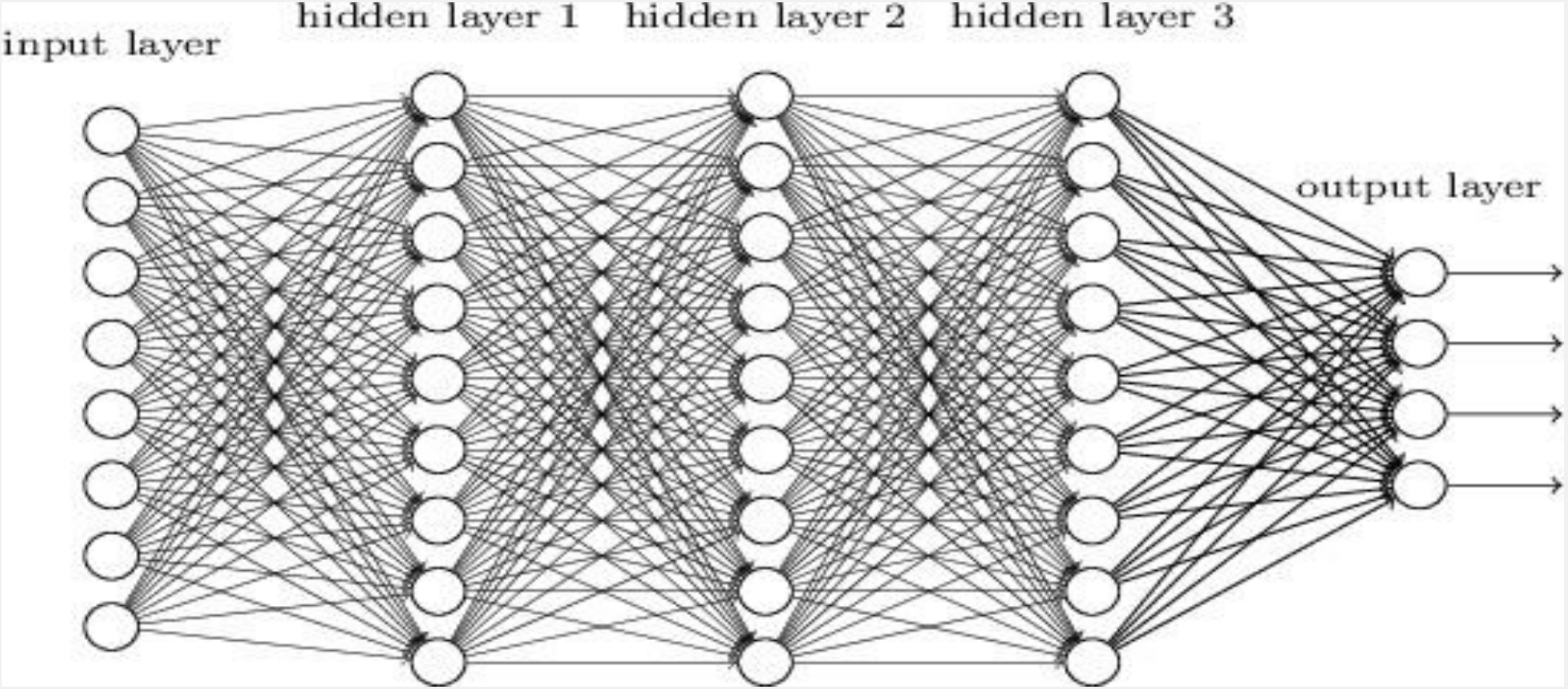
資料來源

卷積類神經網路

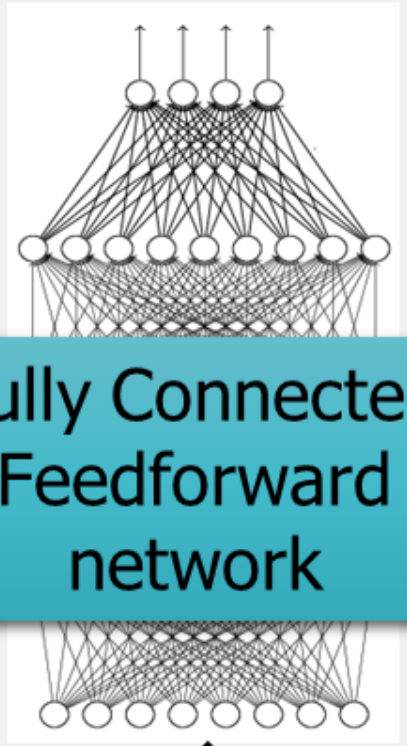
大腦腫瘤切割

資料來源

卷積類神經網路


input layer
hidden layer 1
hidden layer 2
hidden layer 3
output layer

卷積類神經網路



cat / dog

Fully Connected Feedforward network

Convolution

Max Pooling

Convolution

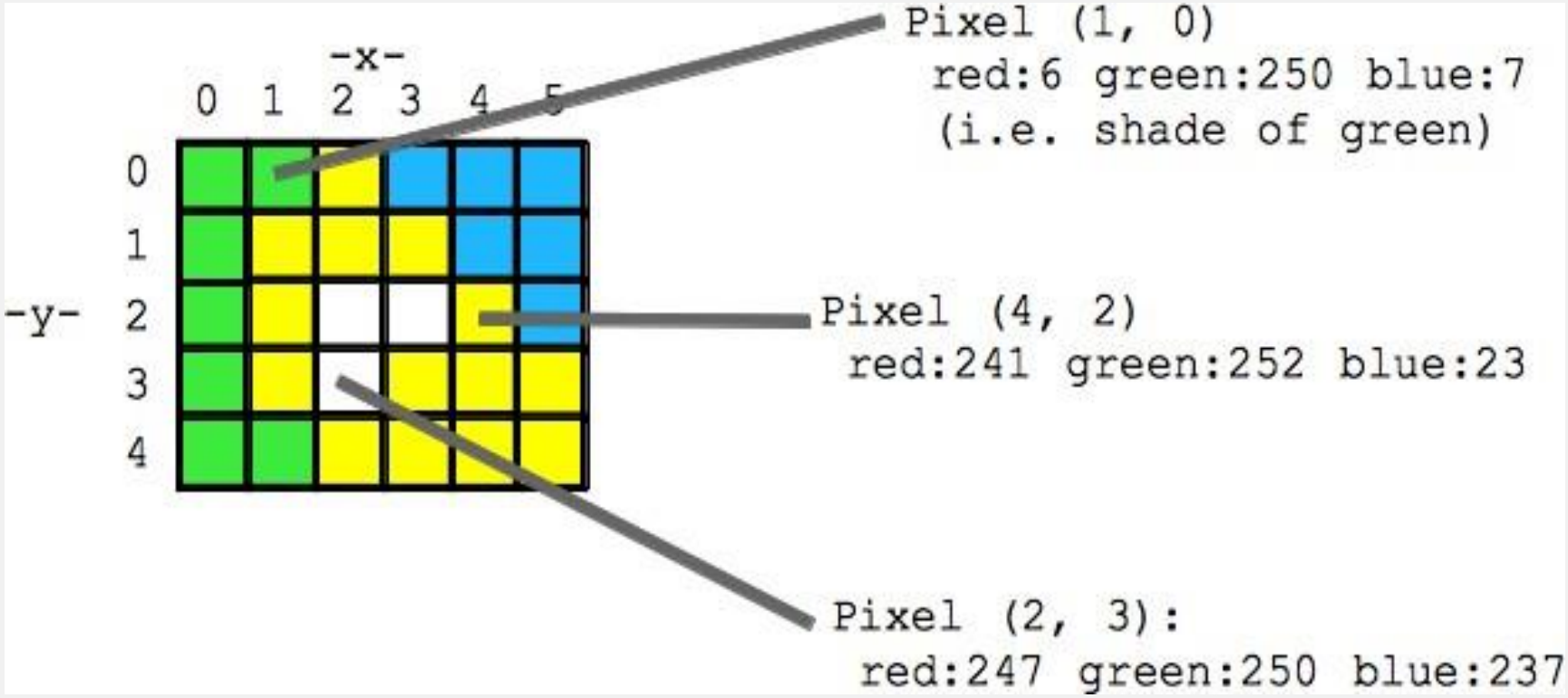
Max Pooling

Flattened

卷積類神經網路

Each image contain many pixels and each pixels compose red, green, blue (RGB)

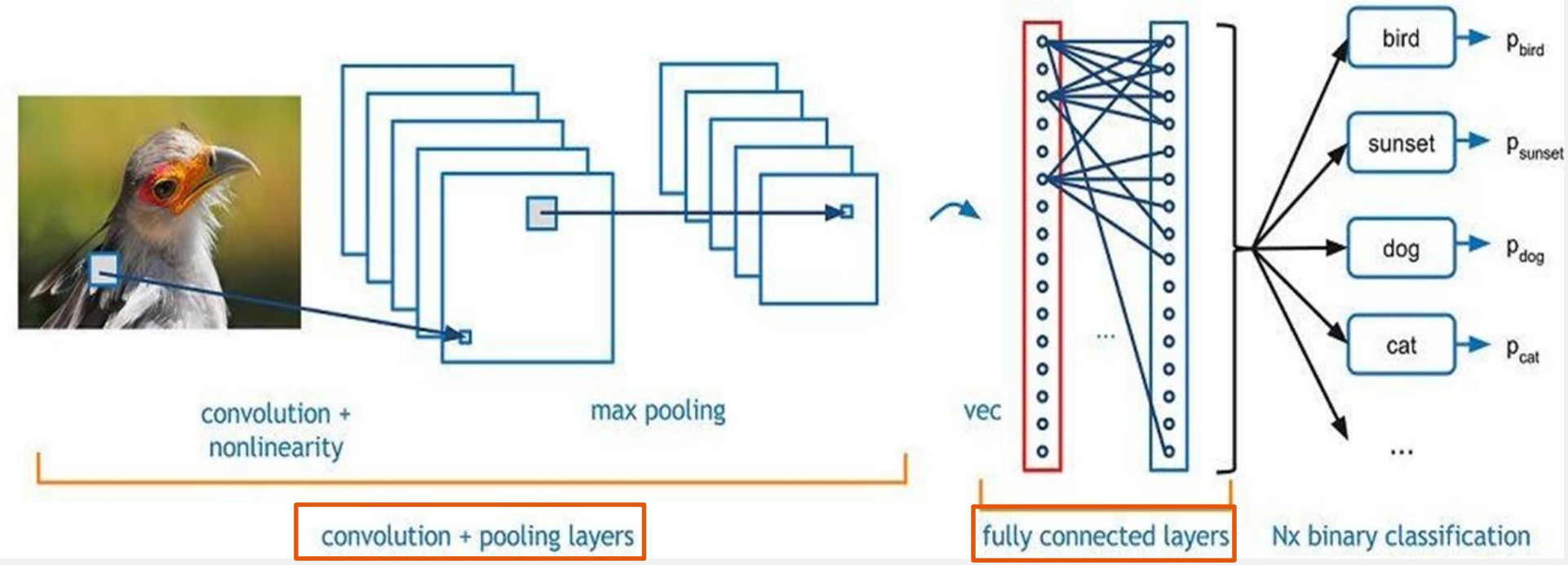
Each channel have brightness levels between 0~255



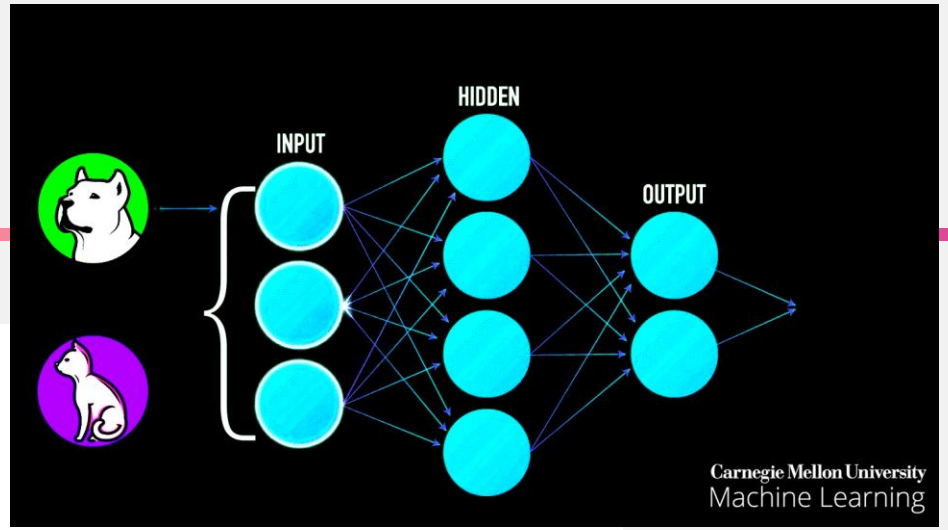
卷積類神經網路

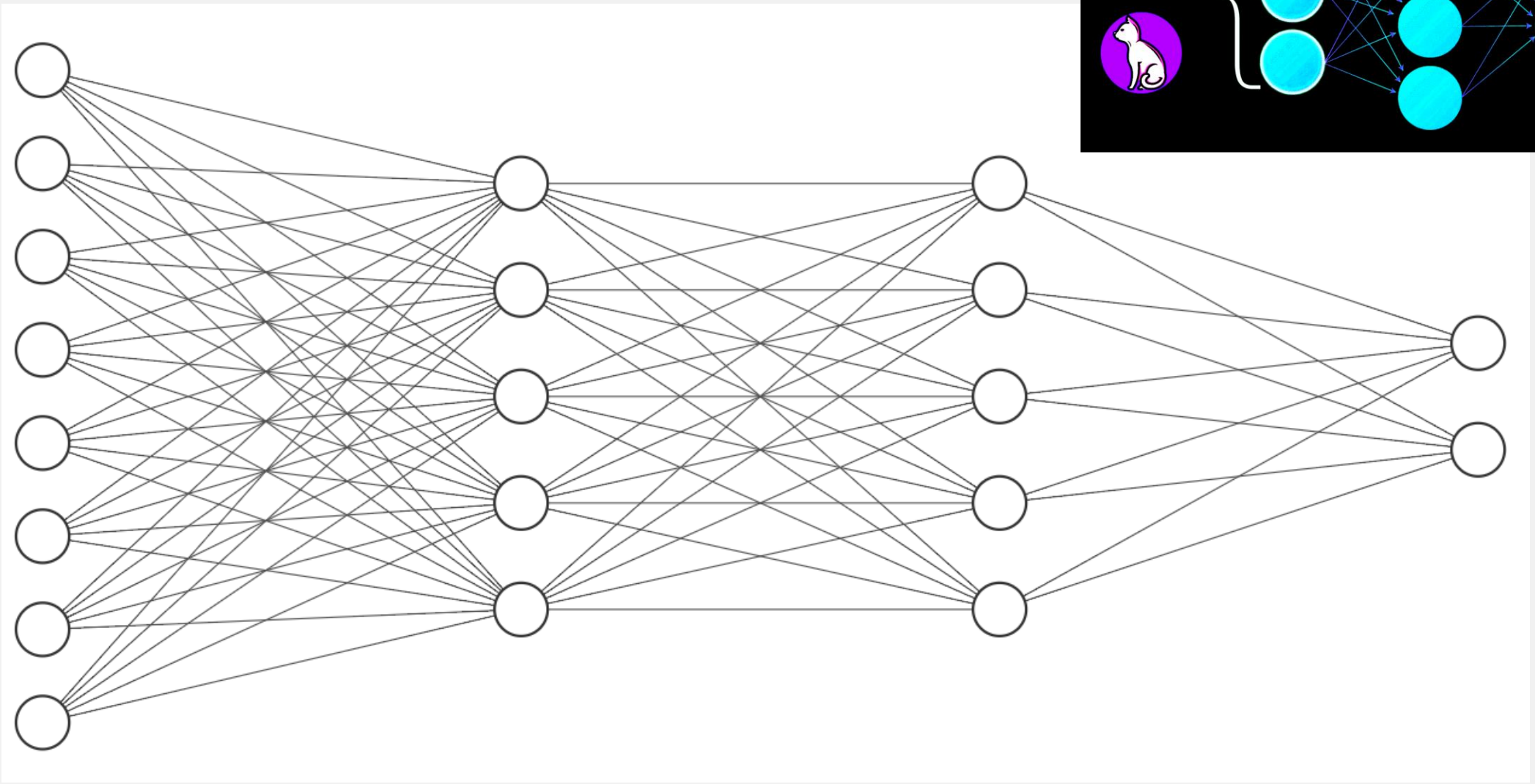
Convolution Neural Network (CNN)

- 卷積層 (Convolution layer)
- 池化層 (Max-pooling layer)
- 全連通層 (Fully connected layer)



卷積類神經網路





Dog

Cat

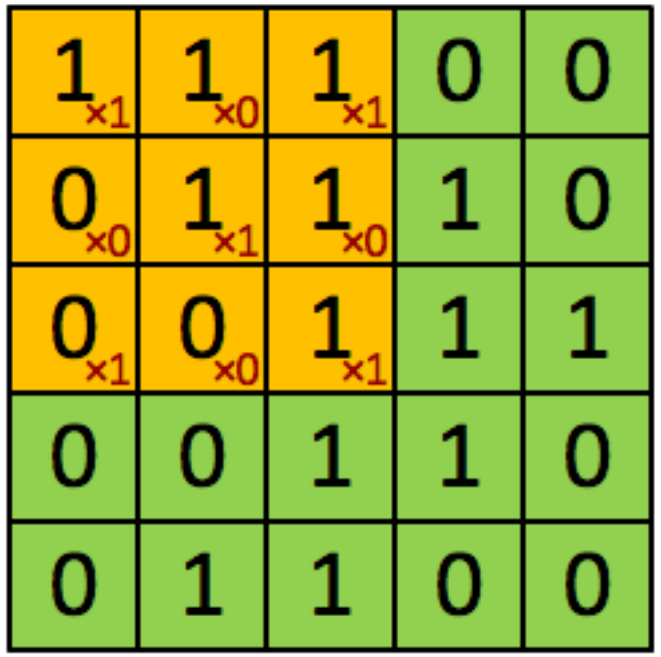
卷積類神經網路

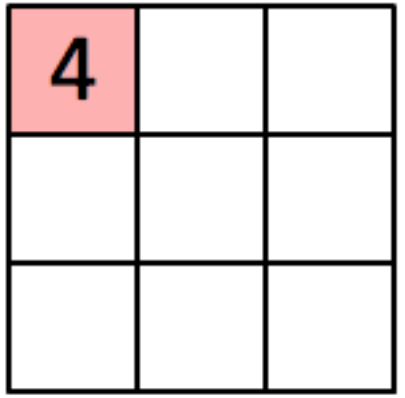
卷積層概念

Kernel = 3*3 , Padding = No, Stride = 1

Kernel

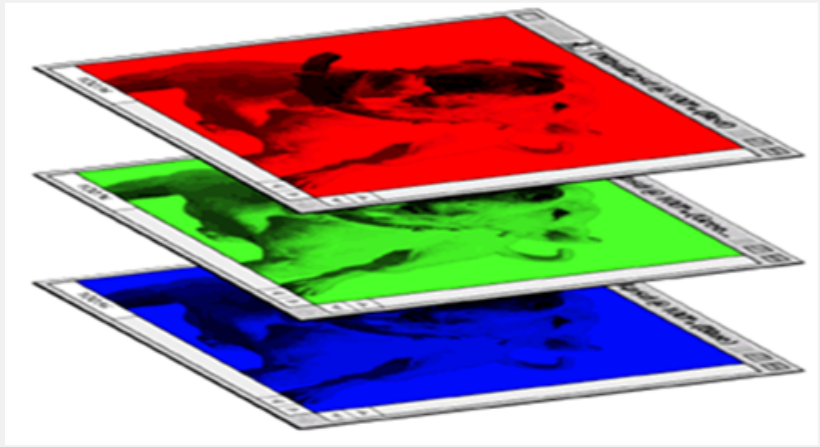
1	0	1
0	1	0
1	0	1

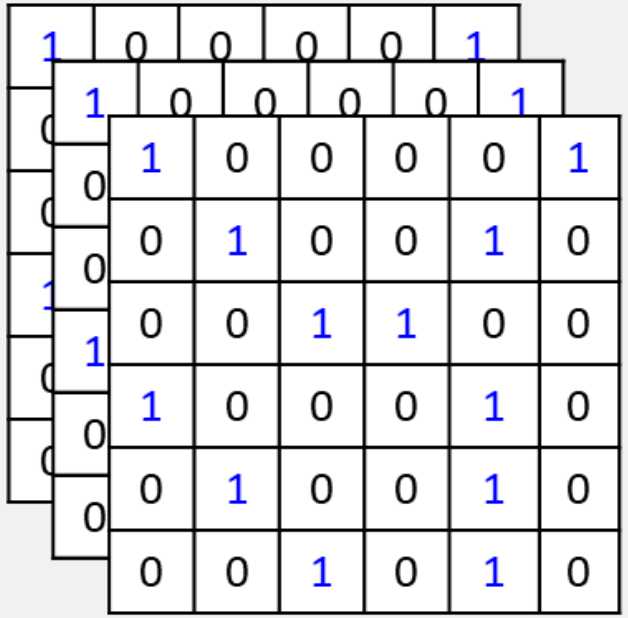


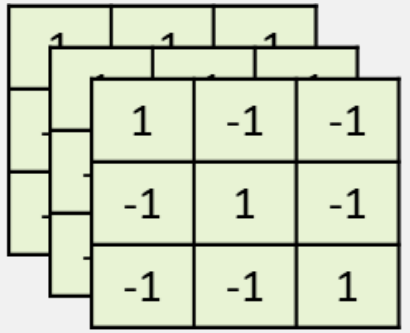


卷積類神經網路

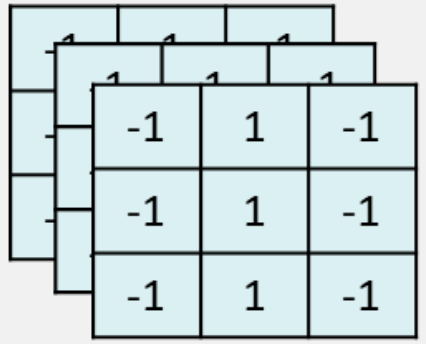
Color image







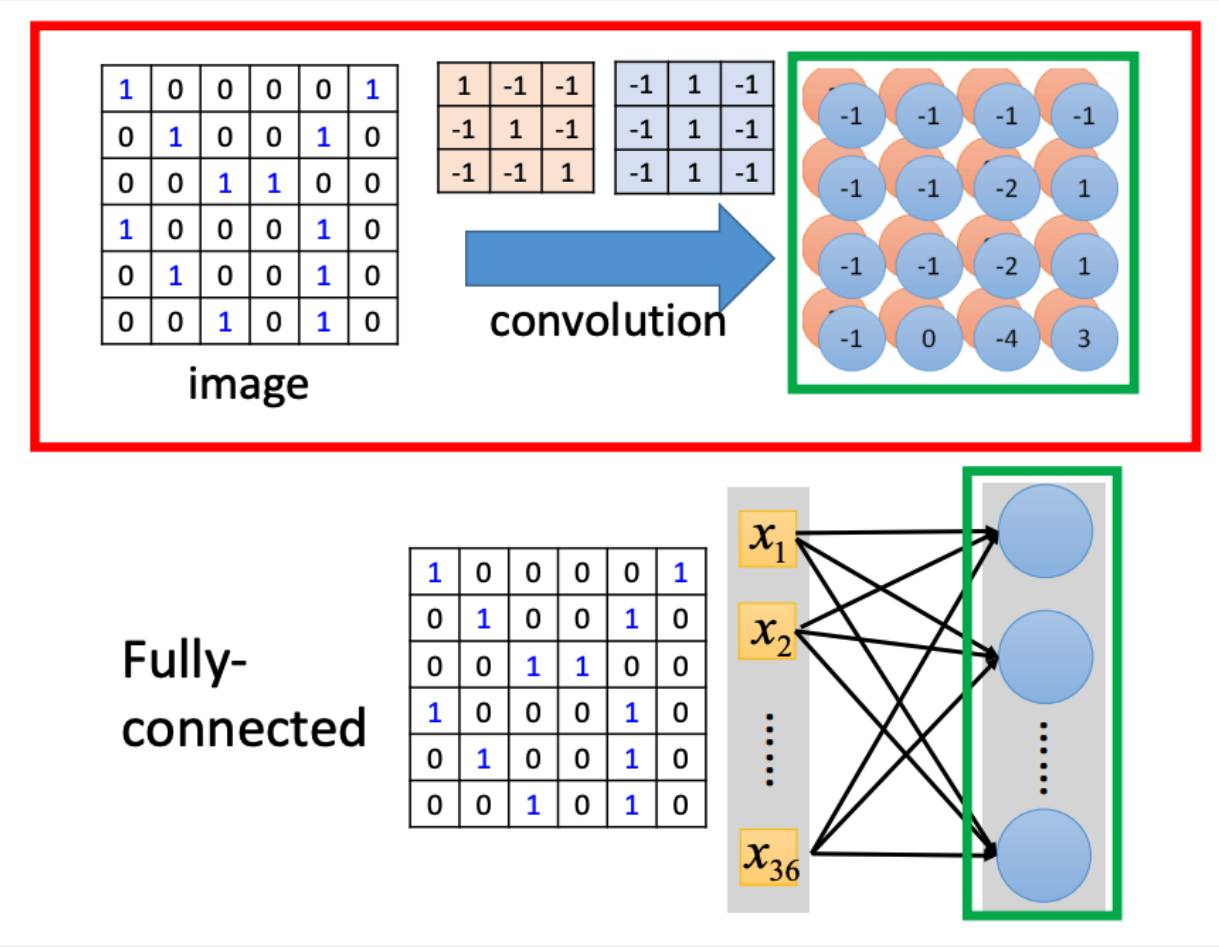
Filter 1



Filter 2

卷積類神經網路

卷積與全連通參數量差異



filter (neuron) count

$(3 \times 3 + 1) \times 3 = 30$

w x h　　bias

$(36 + 1) \times 3 = 111$

previous pixel count + bias

卷積類神經網路

池化層概念