

Machine Learning Techniques

Shyi-Chyi Cheng (鄭錫齊)

Email: csc@mail.ntou.edu.tw

Tel: 02-24622192-6653

章節目錄

- ❖ 第一章 簡介
- ❖ 第二章 Python入門
- ❖ 第三章 貝氏定理回顧
- ❖ 第四章 線性分類器
- ❖ 第五章 非線性分類器
- ❖ 第六章 誤差反向傳播法
- ❖ 第七章 與學習有關的技巧
- ❖ 第八章 卷積神經網路
- ❖ 第九章 深度學習

深度學習

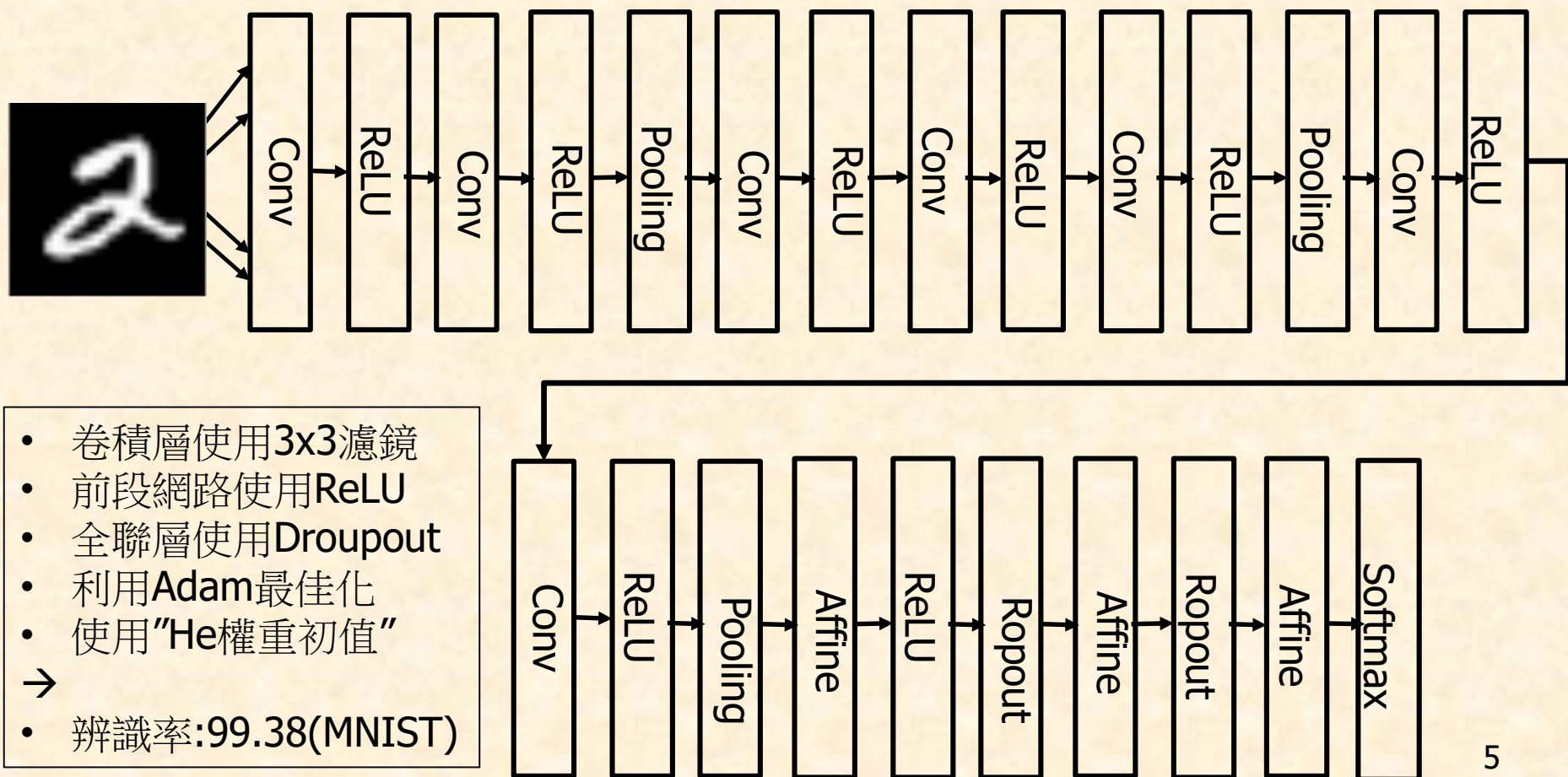
Deep Learning

學習重點

- ❖ 很多問題在加深網路的層數之後，都可以達到提高效能的效果。
- ❖ **ILSVRC**影像辨識競賽最近的趨勢是，前幾名都是使用深度學習的手法，而且使用的網路也深度化。
- ❖ 知名的深度學習網路包括**VGG**、**GoogleNet**、**ResNet**。
- ❖ 利用**GPU**、分散式學習、位元精準度縮減，可以提升深度學習的速度。
- ❖ 深度學習(神經網路)除了辨識物體，也可以用來偵測物體或切割影像。
- ❖ 使用深度學習的應用程式，包括產生圖說、產生影像、強化學習等。
- ❖ 深度學習使得自動駕駛的可行性備受期待。

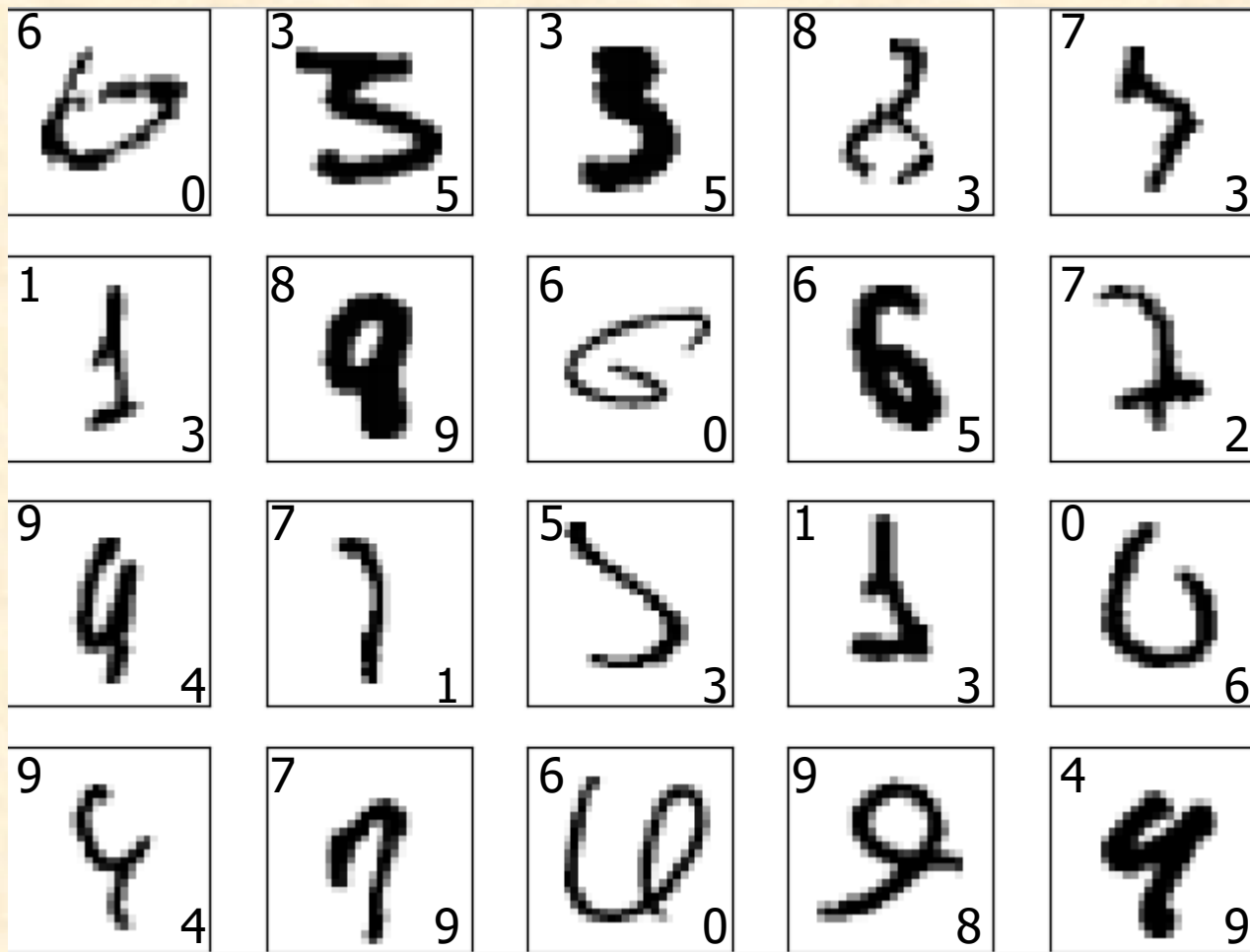
增加網路的層數

- ❖ 深度學習是加了很多層的多層神經網路。
- ❖ CNN的參數最佳化技術是深度學習成功的重要根據。



增加網路的層數

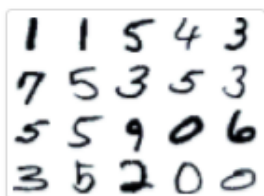
❖ 錯誤分類範例(MNIST Dataset)，錯誤率: 0.61%



進一步提升辨識率

❖ “What is the class of this image?”











http://rodrigob.github.io/are_we_there_yet/build/classification_datasets_results.html



MNIST 50 results collected

Units: error %

[Classify handwritten digits](#). Some additional results are available on the [original dataset page](#).

Result	Method	Venue	Details
0.21%	Regularization of Neural Networks using DropConnect 	ICML 2013	
0.23%	Multi-column Deep Neural Networks for Image Classification 	CVPR 2012	
0.23%	APAC: Augmented PAttern Classification with Neural Networks 	arXiv 2015	
0.24%	Batch-normalized Maxout Network in Network 	arXiv 2015	Details
0.29%	Generalizing Pooling Functions in Convolutional Neural Networks: Mixed, Gated, and Tree 	AISTATS 2016	Details
0.31%	Recurrent Convolutional Neural Network for Object Recognition 	CVPR 2015	
0.31%	On the Importance of Normalisation Layers in Deep Learning with Piecewise Linear Activation Units 	arXiv 2015	
0.32%	Fractional Max-Pooling 	arXiv 2015	Details
0.33%	Competitive Multi-scale Convolution 	arXiv 2015	
0.35%	Deep Big Simple Neural Nets Excel on Handwritten Digit Recognition 	Neural Computation 2010	Details

資料擴增技巧

- ❖ 資料擴增(data augmentation)技巧是公認可進一步提升辨識率的有效方法。
- ❖ 資料擴增是利用演算法，“人工”擴增訓練影像，擴增方式包括“旋轉變形”、“移動變形”、“crop處理”、“flip處理”、調整明暗、放大縮小等等。

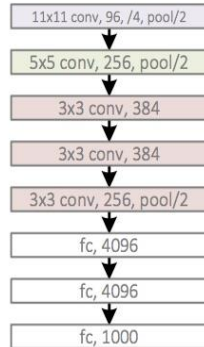


加深層數有甚麼好處？

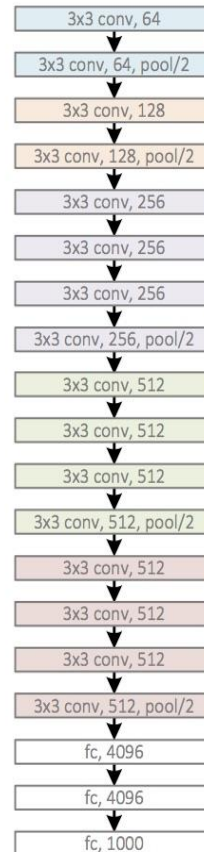
- ❖ 目前缺乏理論證明，但在實務上，加深層數有助提升辨識正確性。

Revolution of Depth

AlexNet, 8 layers
(ILSVRC 2012)



VGG, 19 layers
(ILSVRC 2014)

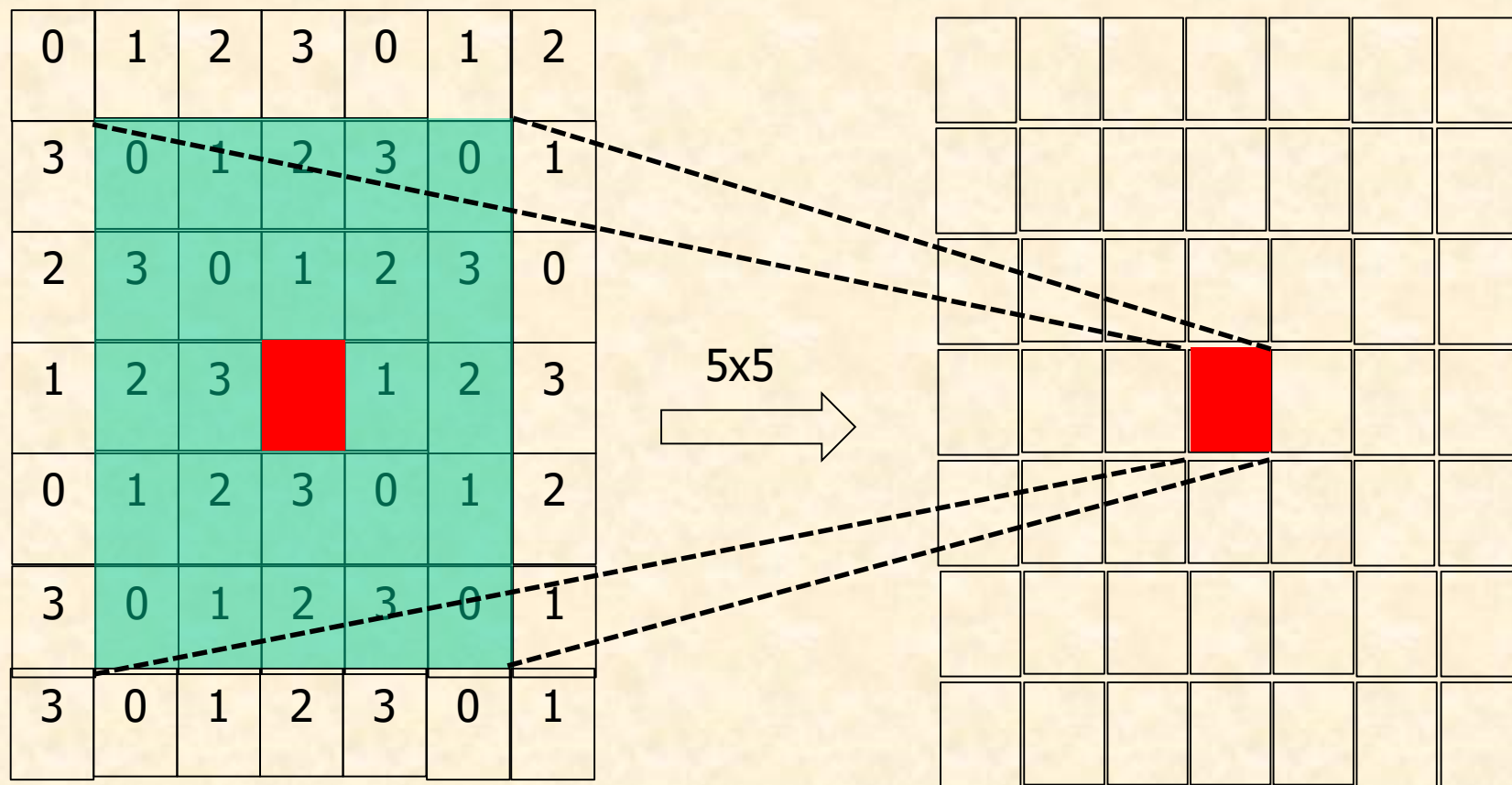


GoogleNet, 22 layers
(ILSVRC 2014)



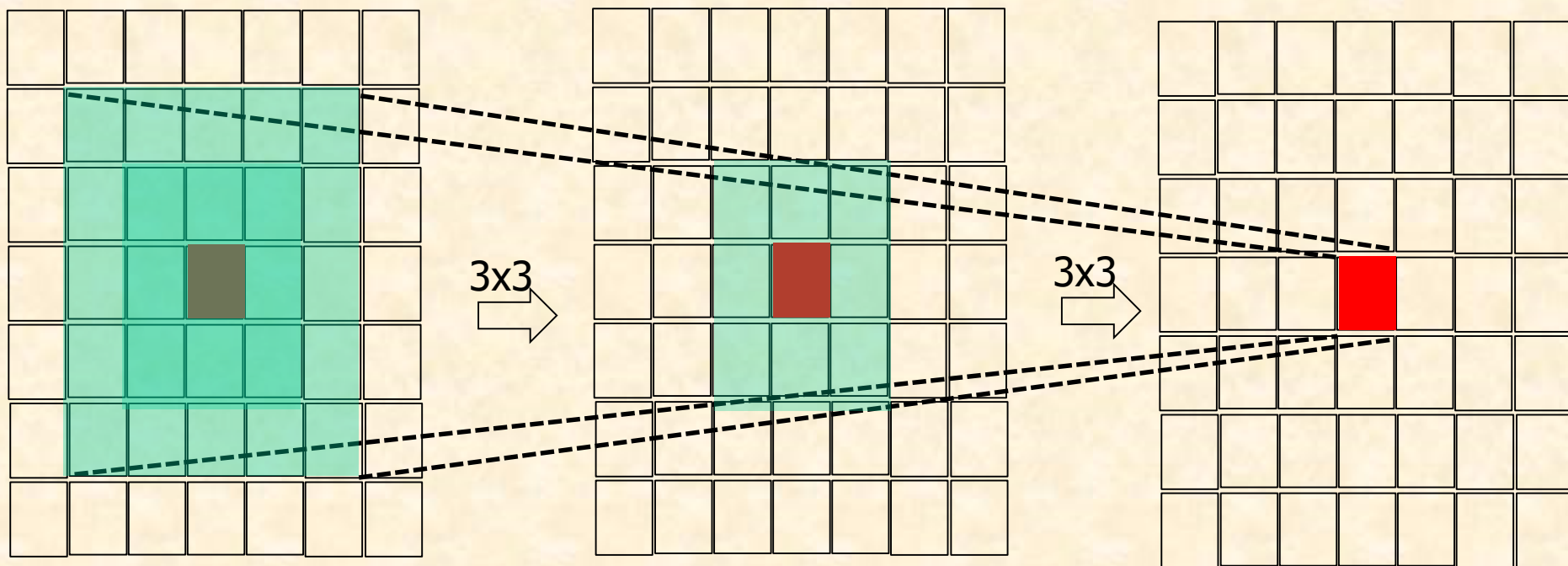
加深層數有甚麼好處？

❖ 觀察CNN的濾鏡大小的意義:5x5濾鏡



加深層數有甚麼好處？

❖ 觀察CNN的濾鏡大小的意義:5x5濾鏡=兩層3x3



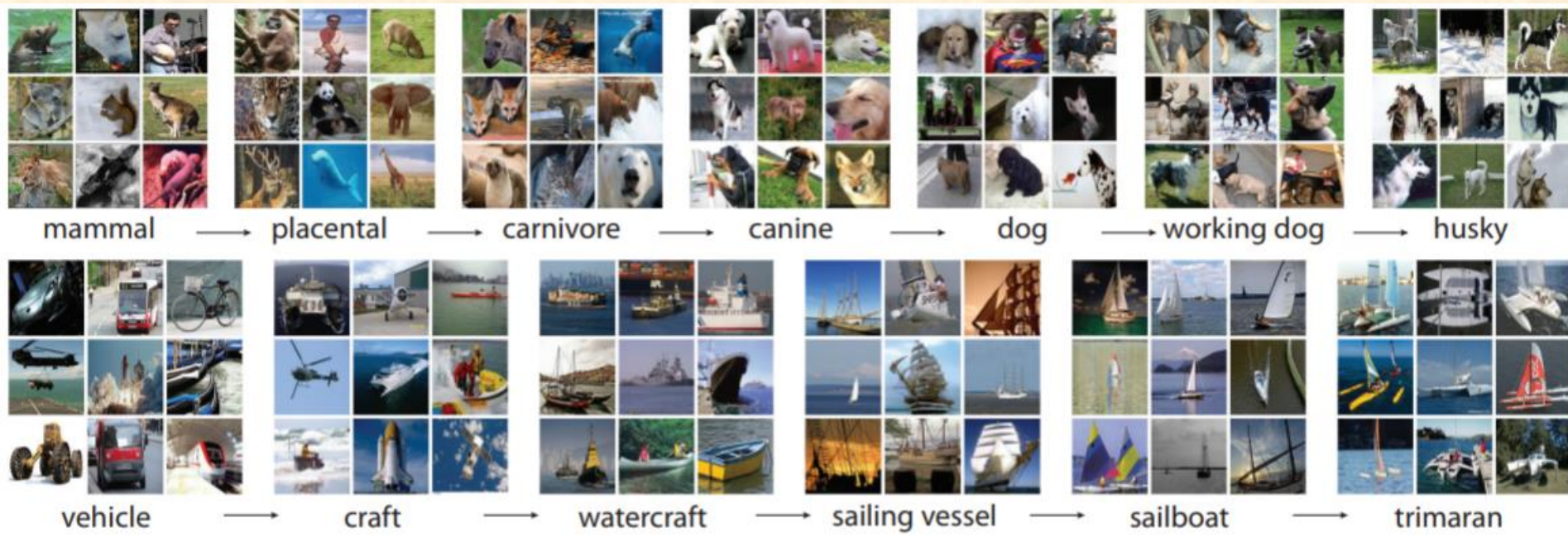
用3x3濾鏡製作實際上7x7濾鏡，需要3卷積層

加深層數有甚麼好處？

- ❖ 使用較小的濾鏡，參數較少，學習速度高。
- ❖ 多層參數可以較少參數建構廣泛的感受視野 (receptive field)
 - 感受視野(receptive field)指讓神經元產生變化的局部空間區域。
- ❖ 多層ReLU等活化函數可以增加神經網路“非線性”的表達能力。
- ❖ 重疊非線性函數，能達到更複雜的表現力。

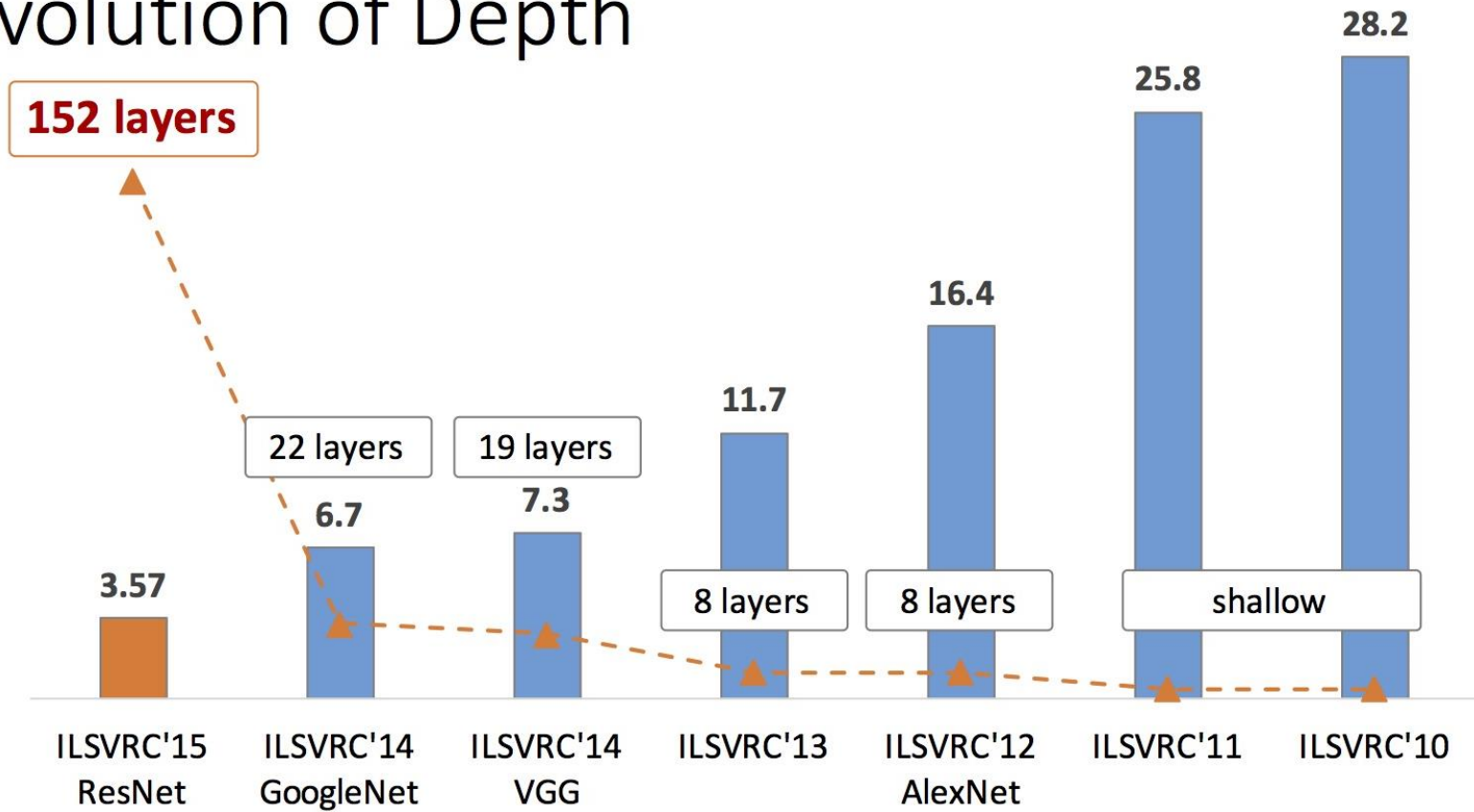
深度學習的歷史

- ❖ 2012年大型視覺辨識競賽ILSVRC(ImageNet Large Scale Visual Recognition Challenge)開啟深度學習的紀元
- ❖ ImageNet是超過100萬張RGB(224x224x3)影像的資料集，每張影像均包含標籤(類別名稱)。
 - 類別個數超過1000個
 - 從2012年之AlexNet起，深度學習方法都是表現最佳的方法。



ILSVRC近幾年結果

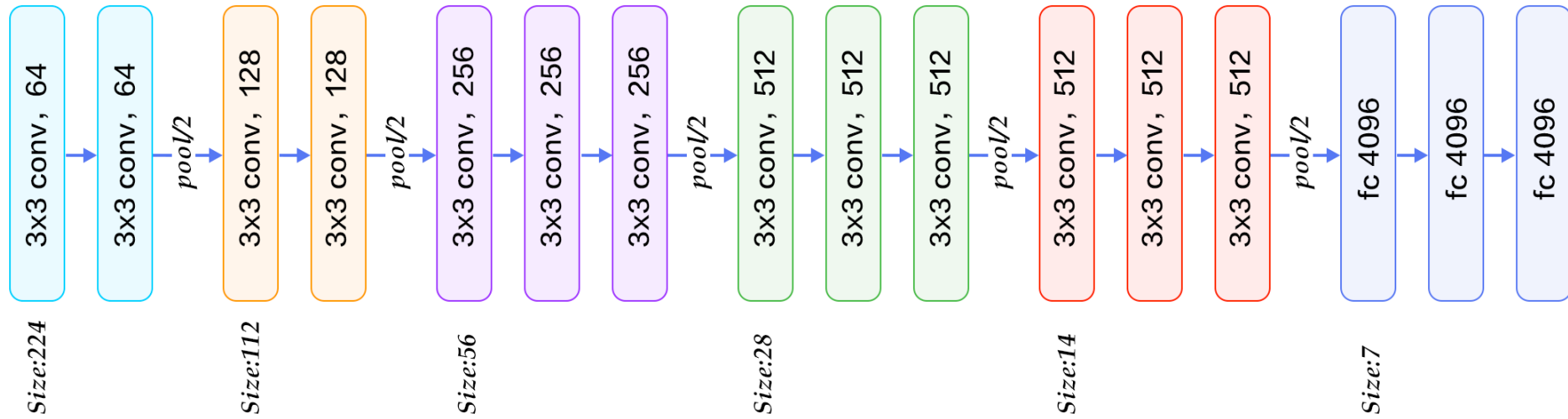
Revolution of Depth



ImageNet Classification top-5 error (%)

VGG

- ❖ VGG是由卷積層及池化層構成的基本CNN。
- ❖ VGG16: 16層；VGG19: 19層



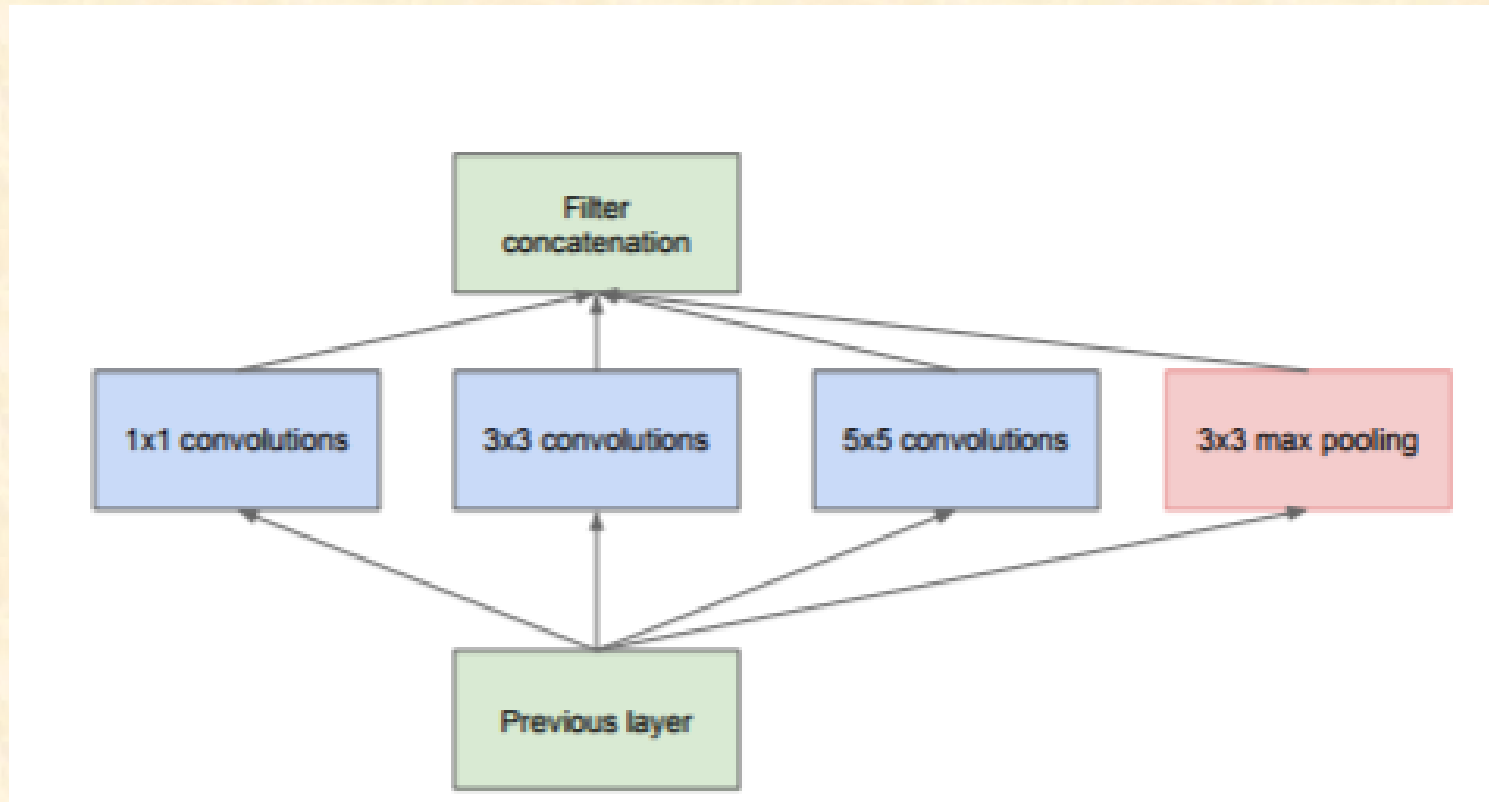
GoogleNet

- ❖ GoogleNet: 22 layers (ILSVRC 2014)
- ❖ 特色
 - 不僅會往垂直方向加深網路，也會往水平方向加深(擴增)



GoogleNet的Inception 結構

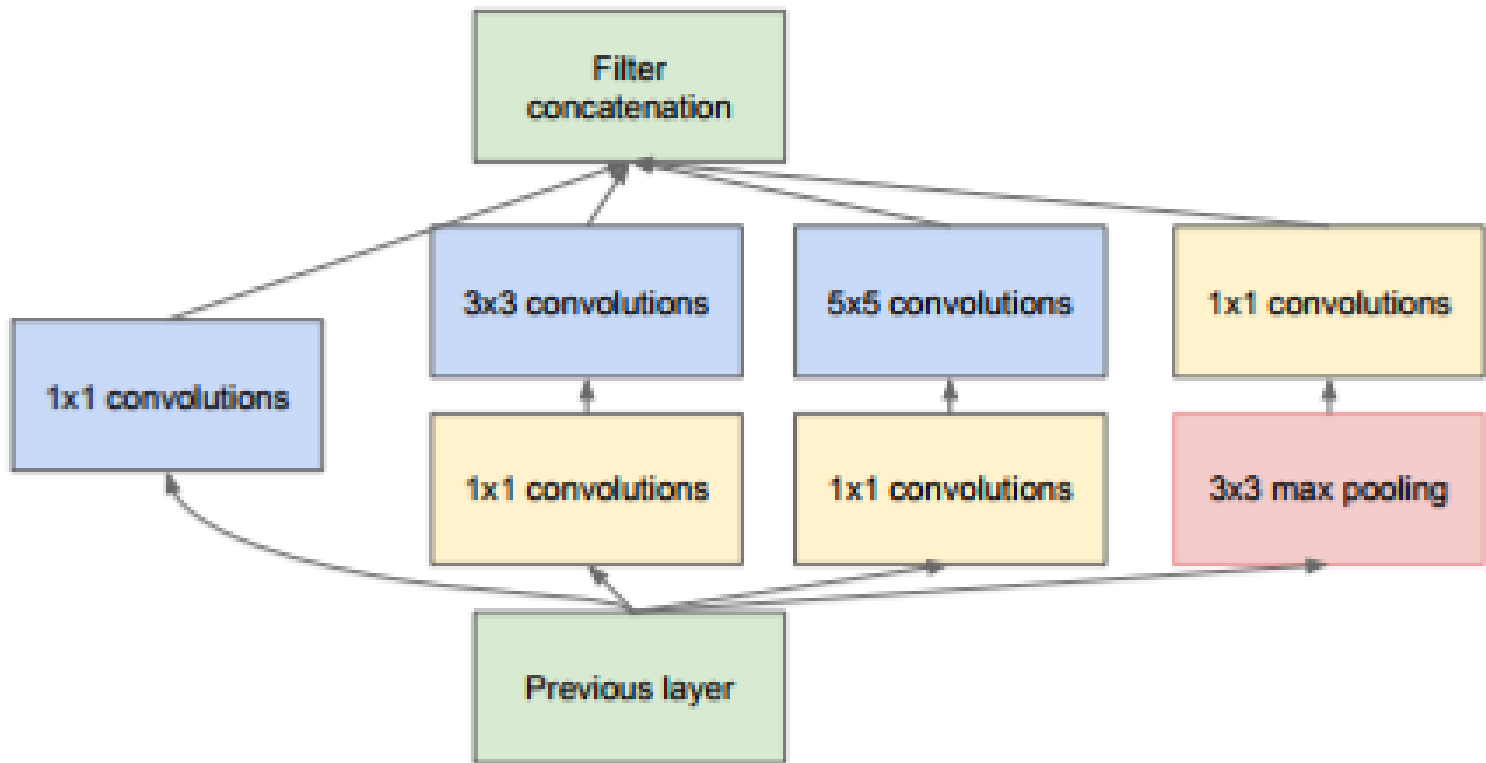
❖ Inception with naïve version



1x1 convolution 用以縮小色板方向的尺寸，能減少參數及提高處理速度 17

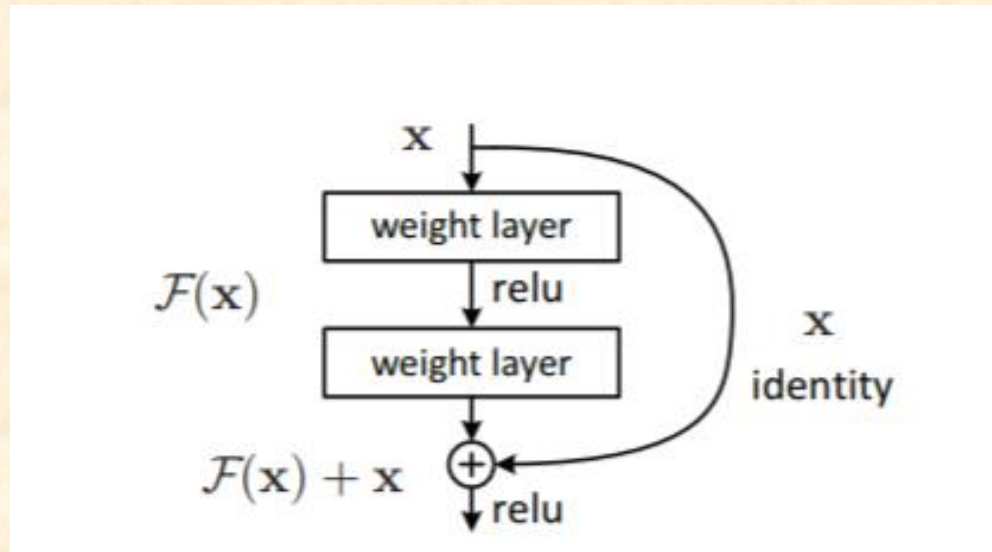
GoogleNet的Inception 結構

❖ Inception with dimension reduction



ResNet

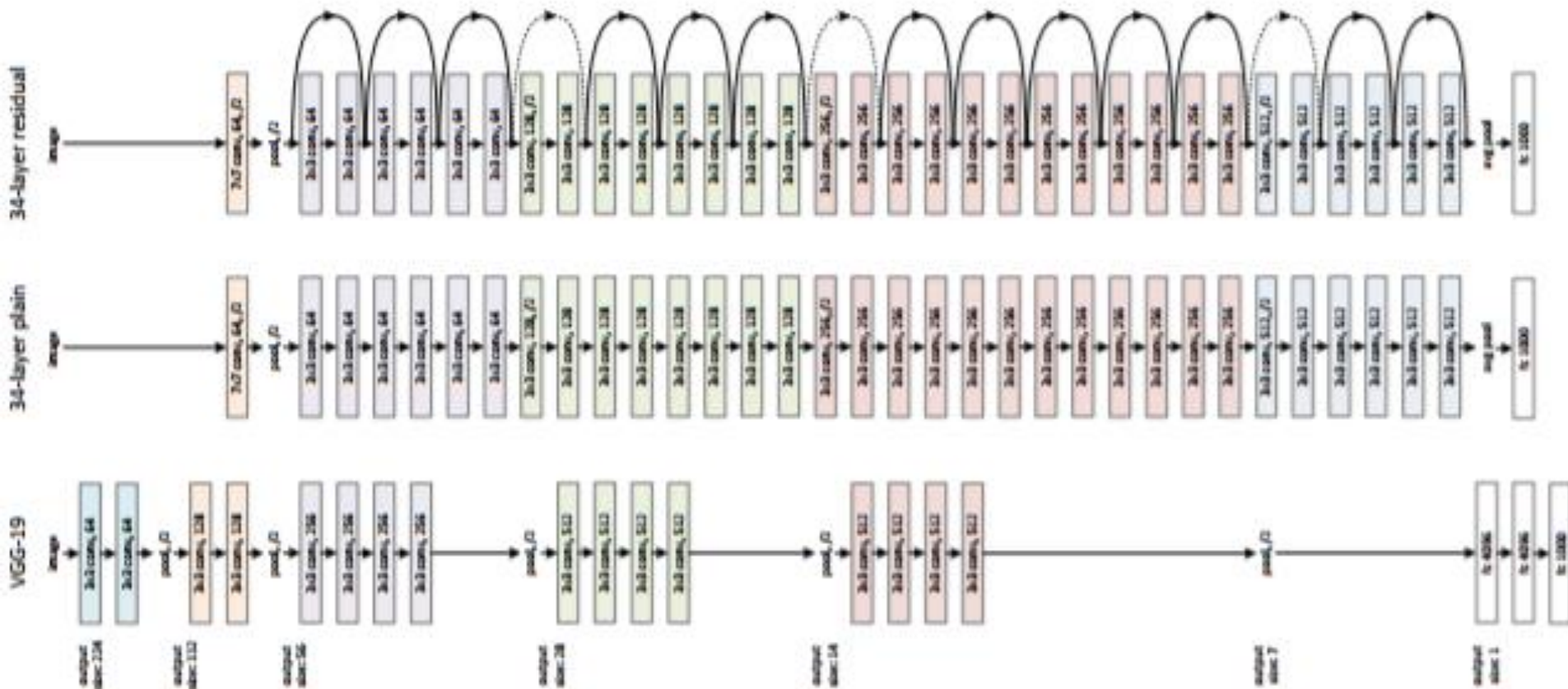
- ❖ ResNet是由Microsoft團隊開發的網路，特色是具有能加深比過去更多層的架構。
- ❖ 加深層數的問題
 - 無法順利學習→效能變差。
- ❖ ResNet使用“跳躍架構”解決這個問題
 - 利用“跳躍架構”能讓訊號在不衰減的狀態下傳遞下去。
 - 反向梯度傳遞時亦然。



Weight layer使用卷積層製作

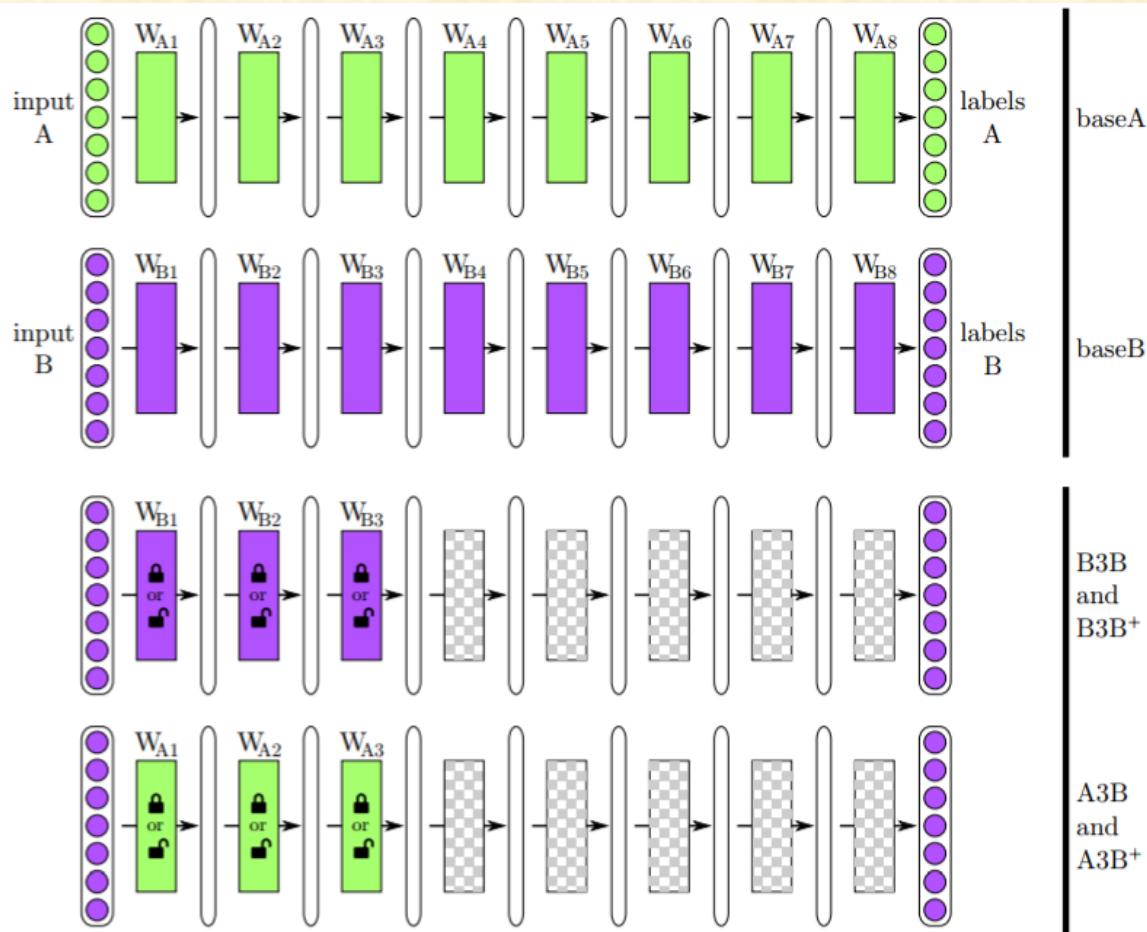
ResNet架構

❖ ResNet架構: VGG+跳躍架構。



遷移學習 (Transfer Learning)

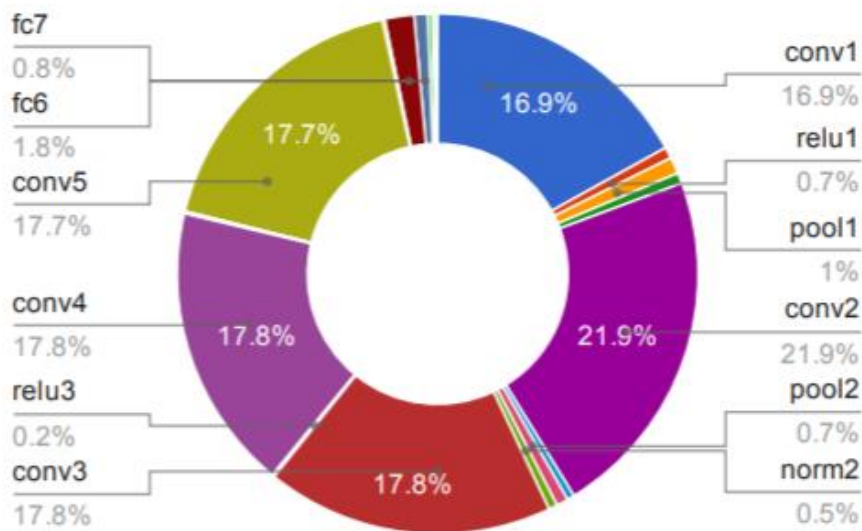
- ❖ 把從ImageNet的龐大資料集中學習到的權重資料，運用在其他辨識問題，稱作遷移學習
- ❖ 主要作法是，把學習完畢的權重(一部分)拷貝到其他神經網路，再輸入新的訓練樣本，微調權重。
- ❖ 學習遷移對於訓練資料較少的情況，尤其能發揮效果。



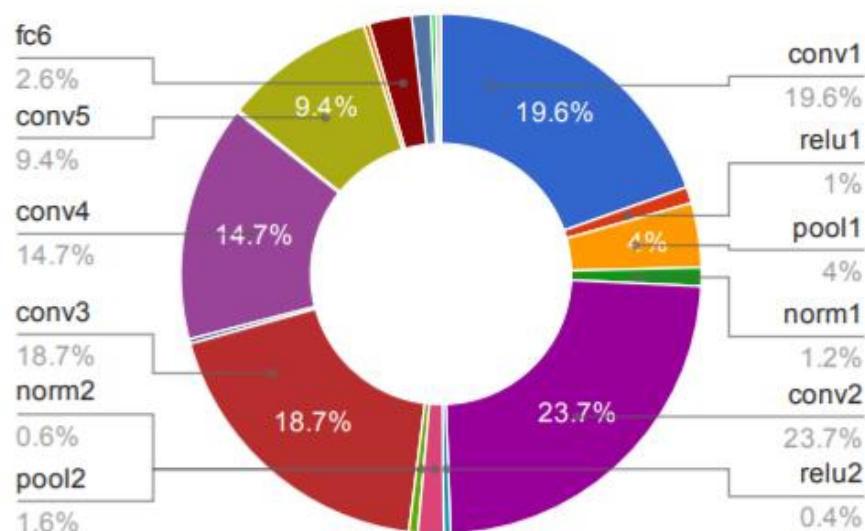
深度學習的高速化

- ❖ 單一CPU的效能難以支撐深部學習的計算需求。
- ❖ 目前的多數深度學習的框架支援GPU(Graphics Processing Unit)，來處理大量的計算。
- ❖ 最新的深度學習的框架支援多個GPU及多台裝置進行分算式計算。

GPU Forward Time Distribution



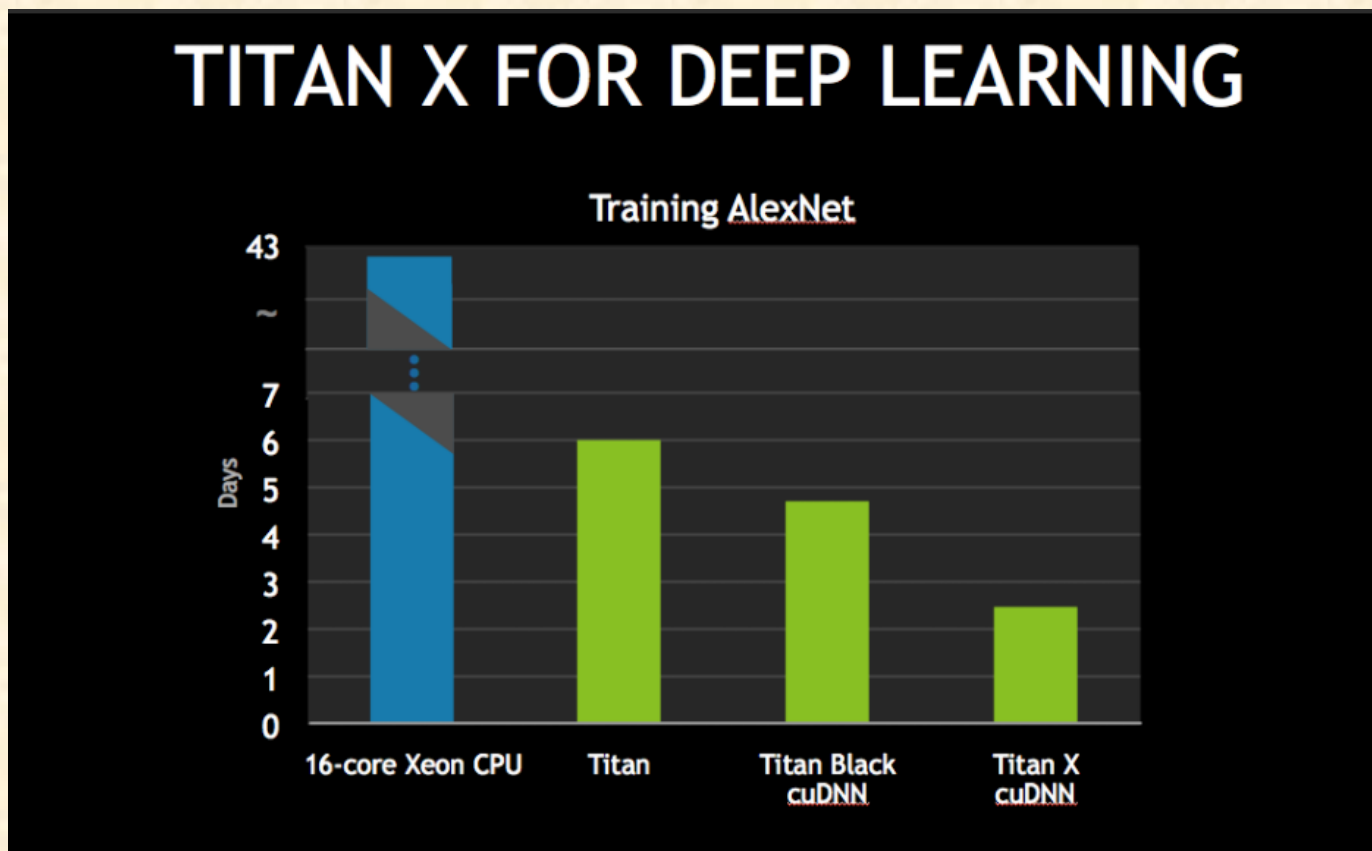
CPU Forward Time Distribution



以AlexNet的forward處理對象，分析時間複雜度:卷積層耗費大量運算資源

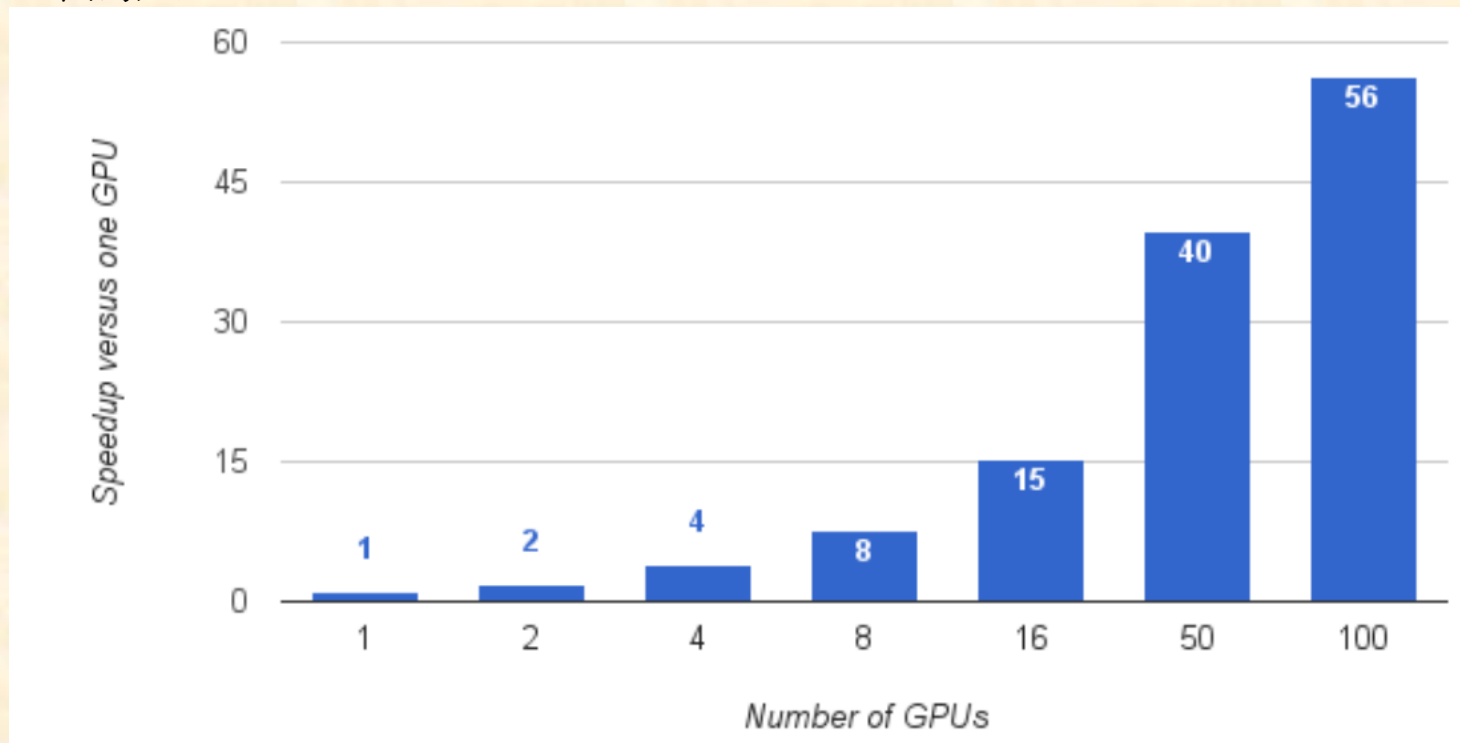
利用**GPU**加速

- ❖ GPU本來用來加速電腦繪圖。
- ❖ GPU現在被廣泛用來快速處理平行運算。
- ❖ 卷基層使用大量的乘積及加總運算，**GPU**的架構比起**CPU**容易平行處理。



分散式學習

- ❖ 為了進一步提高深度學習的計算速度，越來越多的深度學習的框架支援多個GPU及多台裝置進行分散式計算
 - Google的TensorFlow
 - Microsoft 的CNTK(Computational Network Toolkit)
- ❖ 分散式學習在幫助大型資料中心建立低延遲、高吞吐量的網路應用，效果驚人。



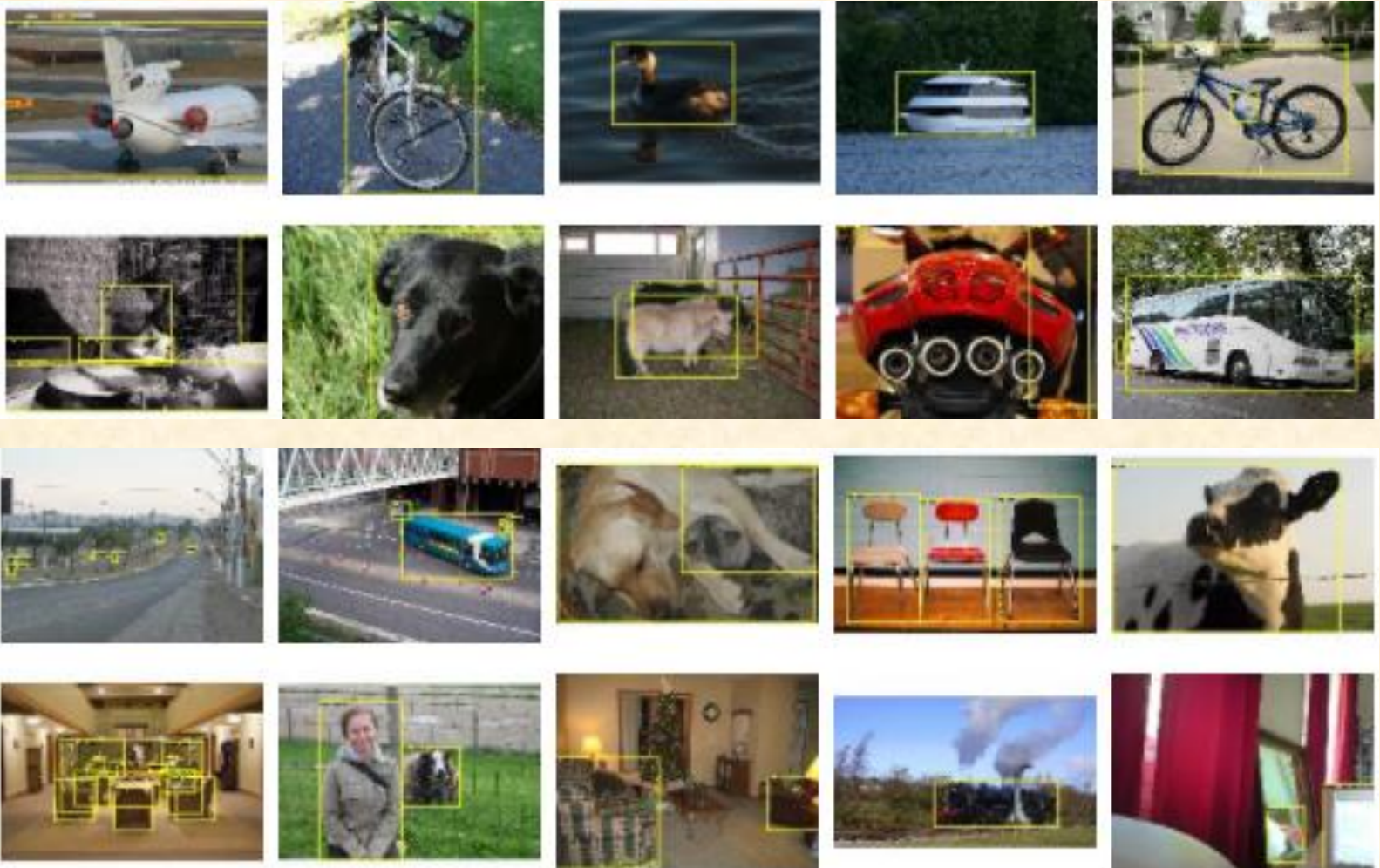
縮減運算精確度的位元

- ❖ 記憶體容量及匯流排頻寬也影響深度學習的速度
- ❖ 電腦主要是利用**64位元**或**32位元**的浮點數來表示數值。
- ❖ 位元數高，運算誤差較小，但是運算及記憶體成本加大
- ❖ 深度學習不需要非常多的數值精確度位元數，原因是神經網路具容錯能力的特性
- ❖ 深度學習使用**16位元**的半精確度浮點數(**half float**)，即可順利學習
 - **NVIDIA** 新一代的**GPU-Pascal**架構，也支援半精確度浮點數運算，比起前一代**NVIDIA** 的**Maxwell GPU**(把半精確度浮點數當作**storage**單位，但非運算單位)，速度提升**2倍**。
- ❖ **Python**的**NumPy**支援半精確度浮點數。

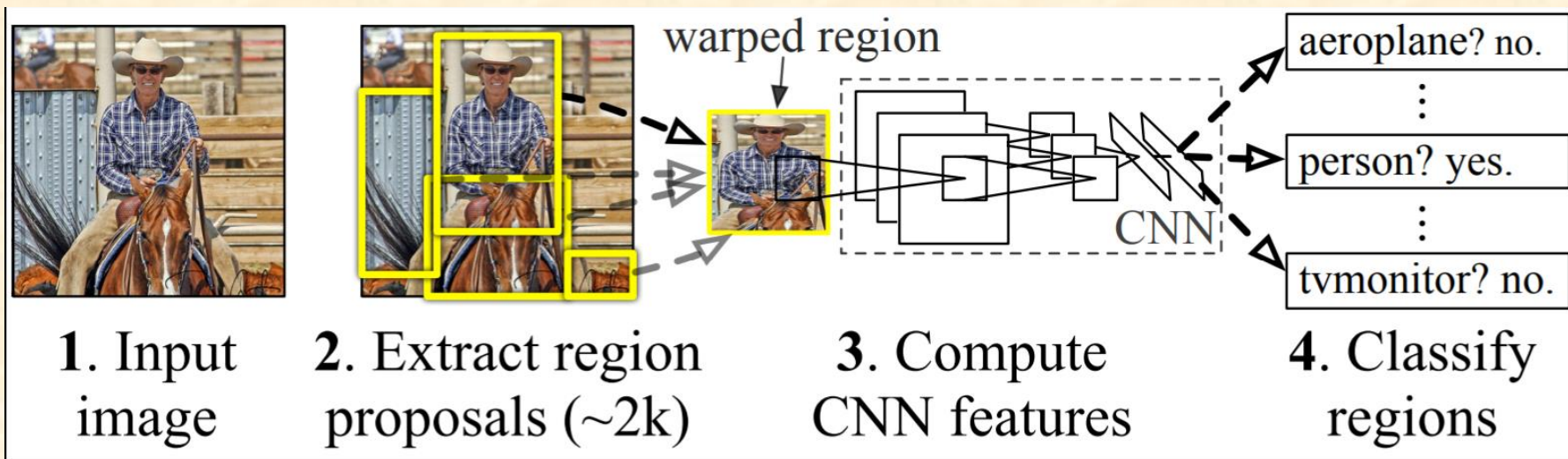
Binarized Neural Networks 提出使用**1位元**表現權重及中間資料的方法
<https://arxiv.org/pdf/1602.02830.pdf>

深度學習的應用範例

❖ 物體偵測



R-CNN進行物件偵測

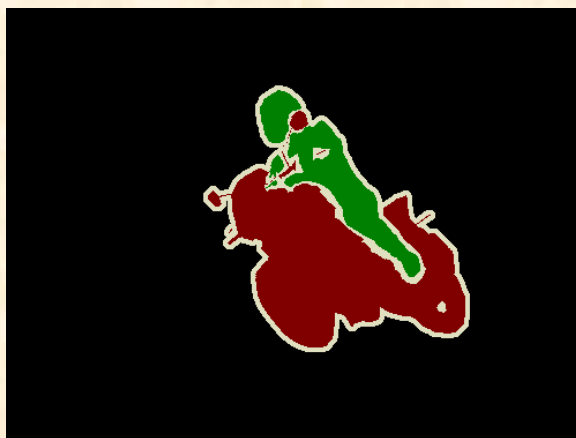


兩階段作法

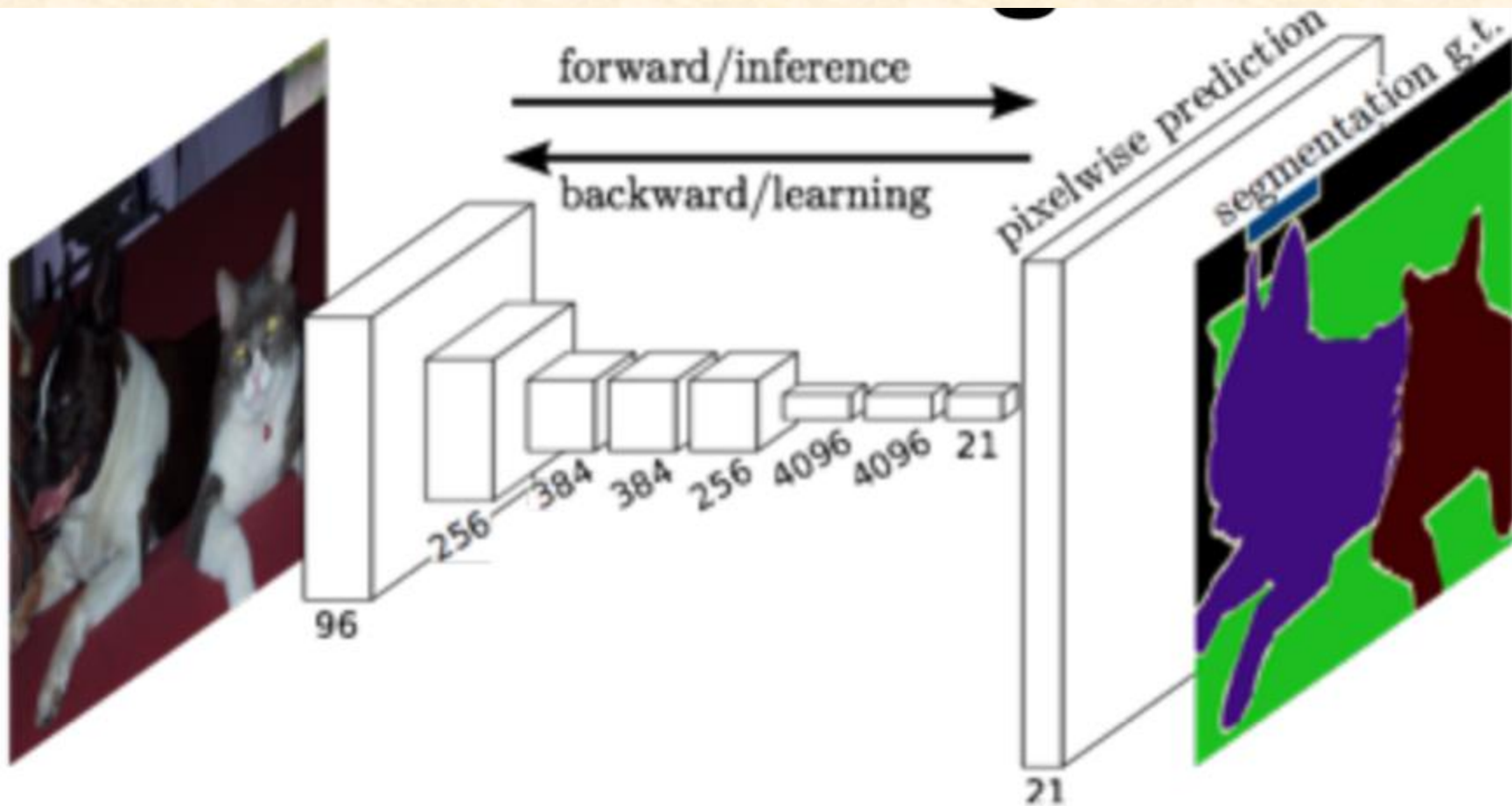
1. 先使用某種方法(Selective Search)抓取組成物件之候選區域
2. 使用CNN進行區域類別分類

影像分割

- ❖ 影像分割是指針對影像，以像素標籤進行類別分類
- ❖ 利用神經網路進行影像分割最直接的方法是以全部影像像素為對象，再依個別像素進行推論處理



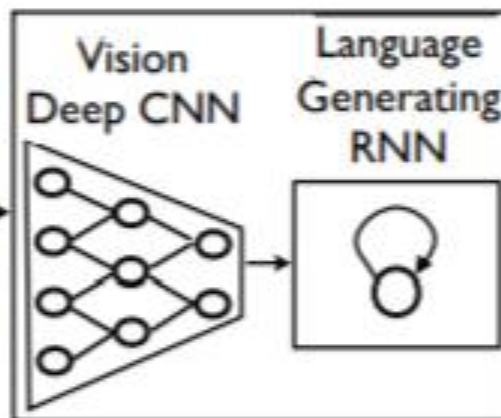
FCN for 影像分割



FCN for影像分割

- FCN的特色
 - 以單一像素進行分類
 - 全部由卷積層構成的神經網路
 - 改變CNN物件辨識用的網路權連接層中，中間資料的空間大小當作排列成1行節點來處理，用以保持空間大小，直到最後的輸出處理。
 - 導入縱橫增值放大法，擴大最後空間大小的處理
 - 把變小的中間資料放大成和輸入影像相同的大小
 - FCN利用反卷積運算(deconvolution)來完成縱橫增值放大法

NIC產生圖片說明文字



A group of people shopping at an outdoor market.

There are many vegetables at the fruit stand.

NIC產生圖片說明文字

- ❖ 輸入影像，自動產生圖片說明，是電腦視覺的重要研究。
- ❖ NIC (Neural Image Caption)結合處理圖片的CNN架構及處理自然語言的RNN(Recurrent Neural Network)，進行圖片說明系統研製。
 - RNN是指擁有遞迴功能的網路，常用於自然語言、時間序列資料等具連續性的資料上。
- ❖ NIC結合多個深度學習模型(CNN+RNN)
 - CNN擷取影像特徵，再把這些特徵傳遞給RNN。
 - RNN把 CNN擷取之影像特徵，遞迴地產生文字。
- ❖ 神經網路的遞迴(Recurrent)功能是指接受之前產生之資料的影響，產生現在的資料
 - 例如，受之前產生‘我’的影響，之後產生‘要’的資料；接著，受之前產生‘我要’的影響，之後產生‘睡覺’的機率大增。

A person riding a motorcycle on a dirt road.



Two dogs play in the grass.



A skateboarder does a trick on a ramp.



A dog is jumping to catch a frisbee.



A group of young people playing a game of frisbee.



Two hockey players are fighting over the puck.



A little girl in a pink hat is blowing bubbles.



A refrigerator filled with lots of food and drinks.



A herd of elephants walking across a dry grass field.



A close up of a cat laying on a couch.



A red motorcycle parked on the side of the road.



A yellow school bus parked in a parking lot.



Describes without errors

Describes with minor errors

Somewhat related to the image

Unrelated to the image

深度學習的未來

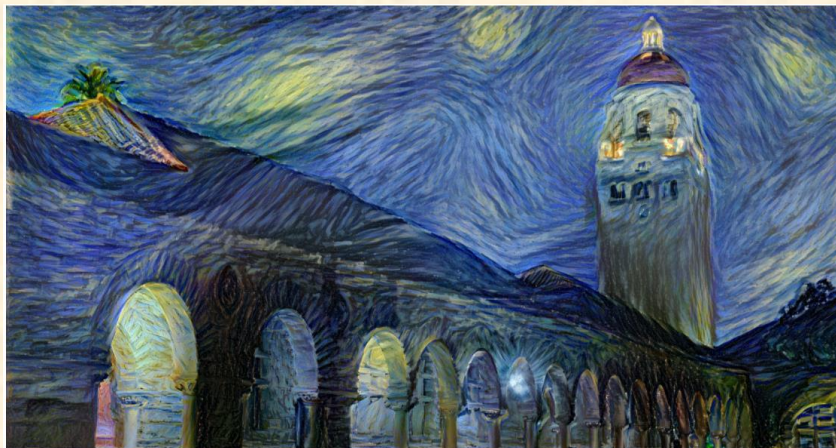
- ❖ 深度學習的成功開啟很多深具想像力的未來**AI**應用
- ❖ 應用範例
 - 轉換影像風格
 - 產生影像
 - 自動駕駛
 - 強化學習

深度學習轉換影像風格

- ❖ 輸入兩張影像，一張是“內容影像”，另一張是“風格影像”

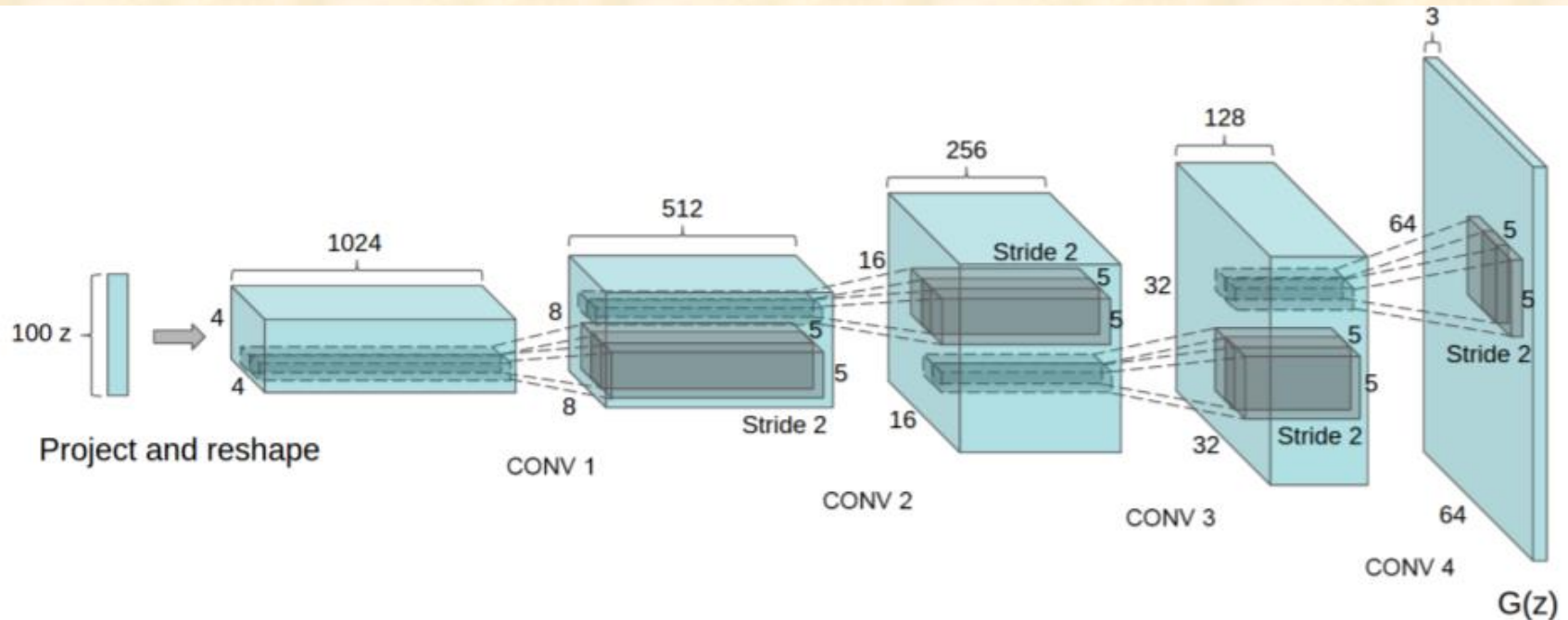


- ❖ 輸出一張描繪藝術風格的影像



產生影像

- ❖ DCGAN (Deep Convolutional Generative Adversarial Network)先使用大量影像進行學習，描繪新影像時不必再輸入任何影像



產生影像

- ❖ DCGAN (Deep Convolutional Generative Adversarial Network)先使用大量影像進行學習，描繪新影像時不必再輸入任何影像

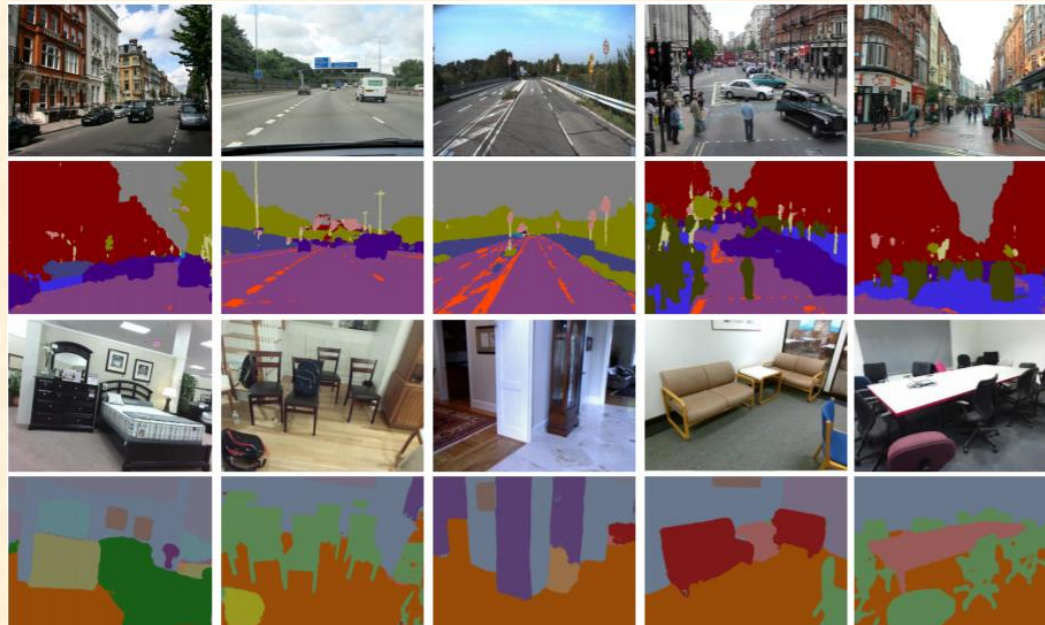


產生影像

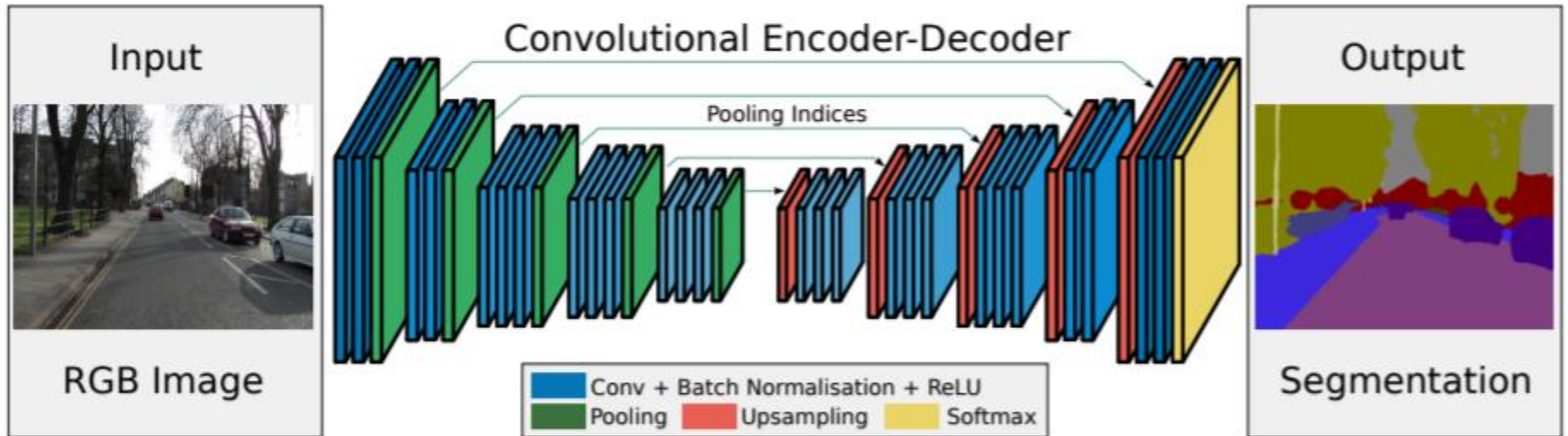
- ❖ DCGAN (Deep Convolutional Generative Adversarial Network)先使用大量影像進行學習，描繪新影像時不必再輸入任何影像
- ❖ 技術重點
 - 使用深度學習之兩種網路，兩者互相制衡，互相學習
 - Generator (生成者)
 - －產生與本尊相似的影像
 - －學習精巧的偽裝或與本尊依模樣的影像生成技術
 - Discriminator(鑑別者)
 - －判斷使否為本尊
 - －學習成為識破高超騙術的鑑定師
 - 屬於非監督式學習技術

自動駕駛

- ❖ 取代人類，由電腦駕駛汽車的“自動駕駛”技術，已經接近實現。
- ❖ 主要技術包括
 - 行車路徑規劃(path planning)
 - 攝影機或雷射等感測技術
 - 正確辨識周邊環境(人、車、地形)的電腦視覺技術
- ❖ 例子:以CNN為基礎的網路SegNet(A Deep Convolutional Encoder-Decoder Architecture for Robust Semantic Pixel-Wise Labelling)，可以精確辨識走路的环境

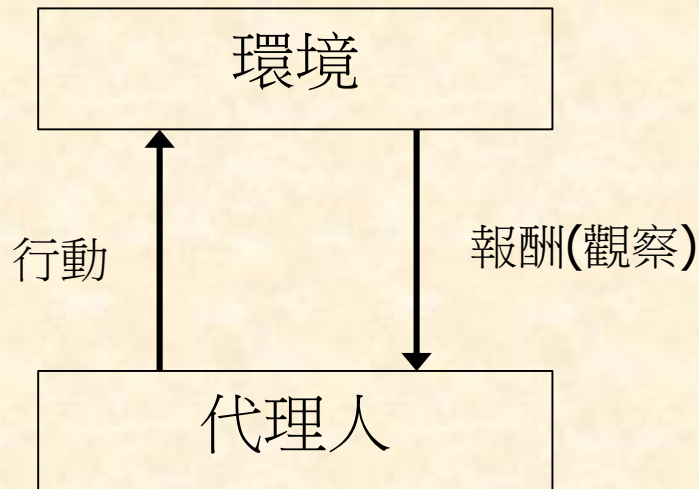


SegNet



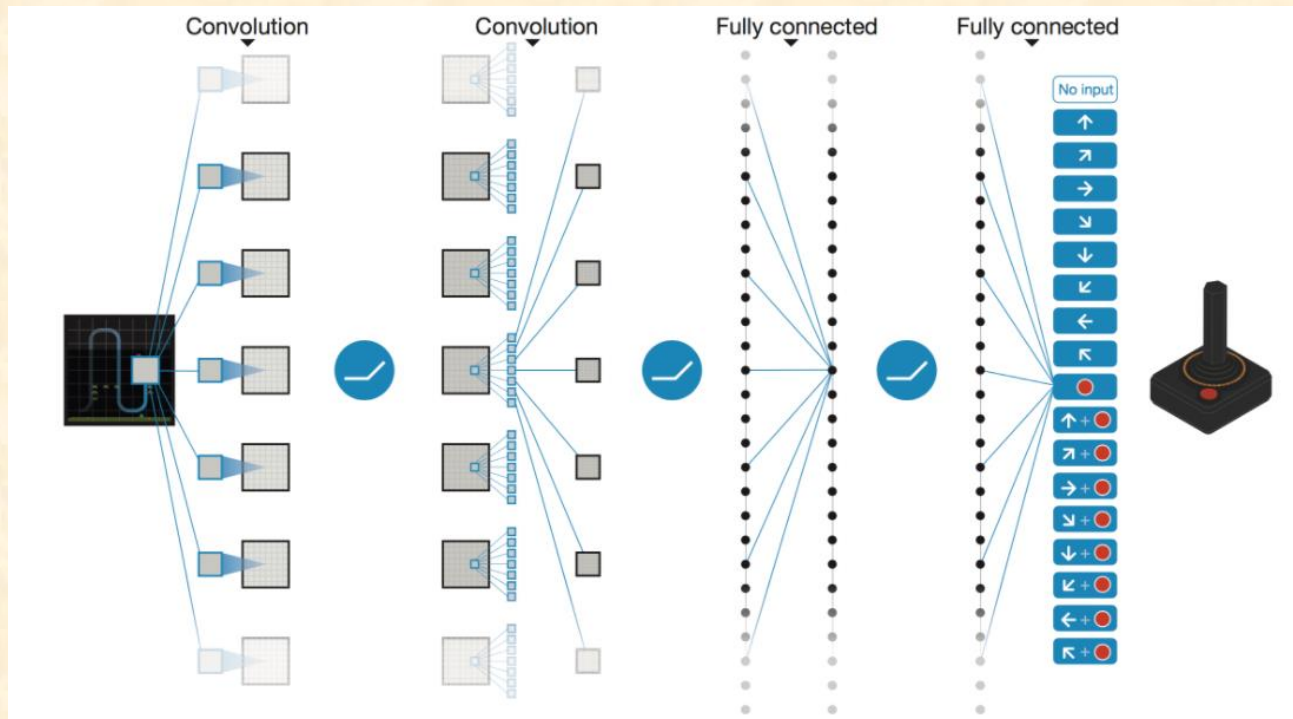
Deep Q-Network (強化學習)

- ❖ 強化學習(Reinforcement Learning)是一種從嘗試錯誤中，以獲得更好報酬為目的，進行自主式學習的方法。
- ❖ 強化學習主要由構成份子為
 - 代理人
 - 根據環境狀況來決定要採取的行動(Action)
 - 該行動會讓環境變化，進而獲得某些報酬(逞罰)



Deep Q-Network (強化學習)

- ❖ Deep Q-Network (DQN) 使用深度學習之Q學習的強化學習方法
- ❖ Q學習的特色
 - 強化學習演算法，用以找出最佳化行動價值函數，決定最佳行動
- ❖ DQN以CNN模擬Q學習的趨近函數
- ❖ DQN實驗讓電視遊戲自動學習，達到超越人類等級的超作手法



AlphaGo

- ❖ AI AlphaGo擊敗人類圍棋冠軍的新聞舉世矚目。
- ❖ AlphaGo的技術內部包括
 - 深度學習
 - 強化學習
 - 超過3,000萬個專業棋譜進行大數據學習
 - AlphaGo本身反覆與自己對戰，不斷學習
- ❖ AlphaGo與DQN都是Google Deep Mind公司的研究
- ❖ Deep Mind 正運用AlphaGo技術到各種AI問題，其舉動值得關注。

David Silver et al. (2016): Mastering the game of Go with deep neural networks and tree search. Nature 529, 7587 (2016), 484-489.

Any Questions?