卷積神經網路 Convolutional Neural Network
&
電腦視覺 Computer Vision
Part4

林彥宇&教研處

「版權聲明頁」

本投影片已經獲得作者授權台灣人工智慧學校得以使用於教學用途，如需取得重製權以及公開傳輸權需要透過台灣人工智慧學校取得著作人同意；如果需要修改本投影片著作，則需要取得改作權；另外，如果有需要以光碟或紙本等實體的方式傳播，則需要取得人工智慧學校散佈權。

課程內容

本日課程：

1. CNN Renowned Models

延伸閱讀 (Optional)：

1. A Brief History of CNN Models
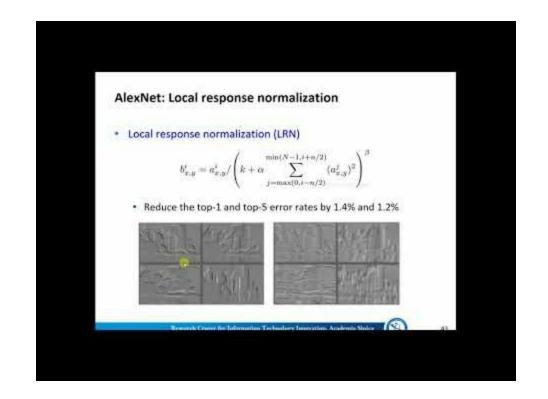
台灣人工智慧學校

本次課程結束後你 (妳) 應該會什麼?

- **軟實力**
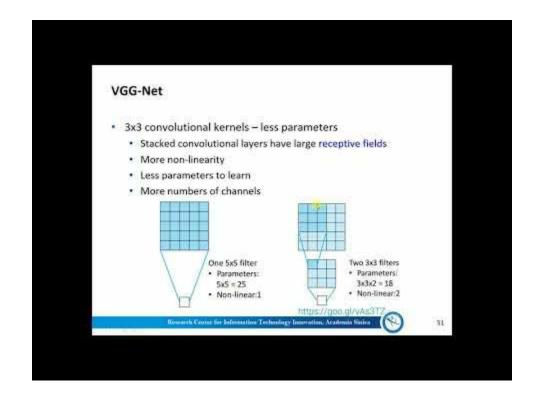 - 了解 CNN模型的發展歷史以及各大經典模型
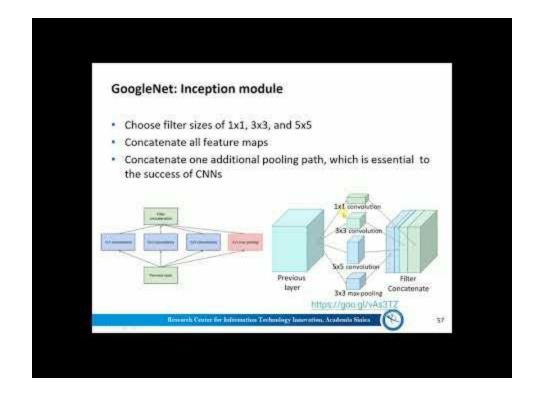 - 比較各種模型的不同

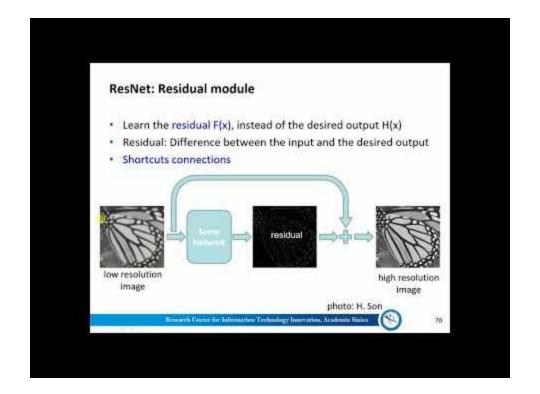# CNN Renowned Models
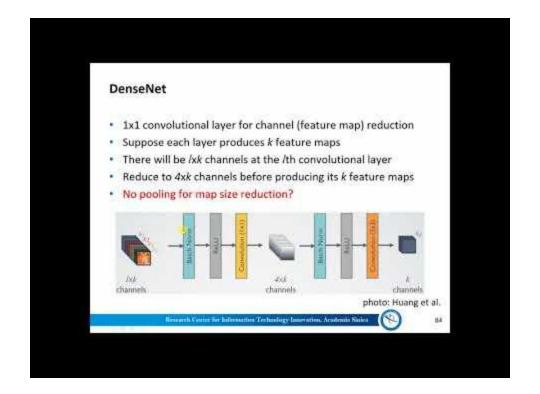
理論講授 - AlexNet

理論講授 - VGG Net

理論講授 - GoogLeNet

理論講授 - ResNet

ResNet: Residual module

- Learn the residual F(x), instead of the desired output H(x)
- Residual: Difference between the input and the desired output
- Shortcuts connections

Some Network → residual

low resolution image → high resolution image

photo: H. Son

Research Center for Information Technology Innovation, Academia Sinica

70

DenseNet

- 1x1 convolutional layer for channel (feature map) reduction
- Suppose each layer produces k feature maps
- There will be $l \times k$ channels at the lth convolutional layer
- Reduce to $4 \times k$ channels before producing its k feature maps
- No pooling for map size reduction?

$l \times k$ channels

$4 \times k$ channels

k channels

photo: Huang et al.

理論講授 - Network Comparison

[小結]Modern CNN model (optional)

GoogLeNet

(b) Inception module with dimension reductions
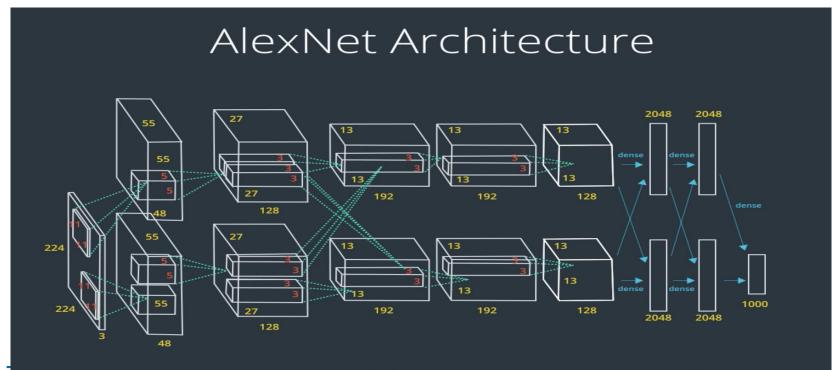
強大的模型 from FB, Google...

- ImageNet 的比賽開辦以來, 所有實驗室與企業都希望能夠找出最好的 CNN 架構來提高準確率

Model	Alexnet	VGG	GoogleNet	ResNet
Time	2012	2013	2014	2015
Layer	8	19	22	152
Error	16.4%	7.3%	6.7%	3.57%
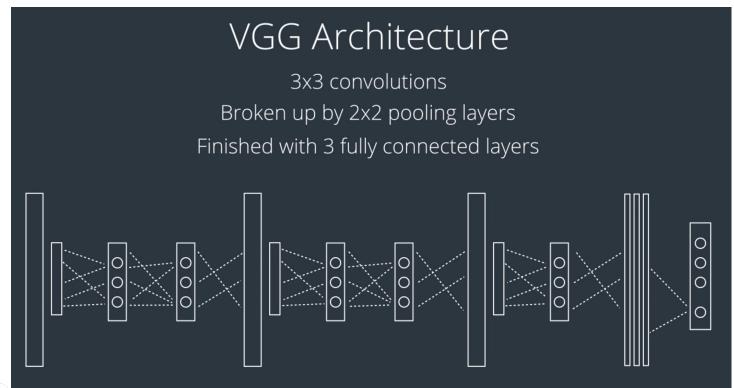
AlexNet paper

VGGNet [paper](#)

GoogLeNet paper
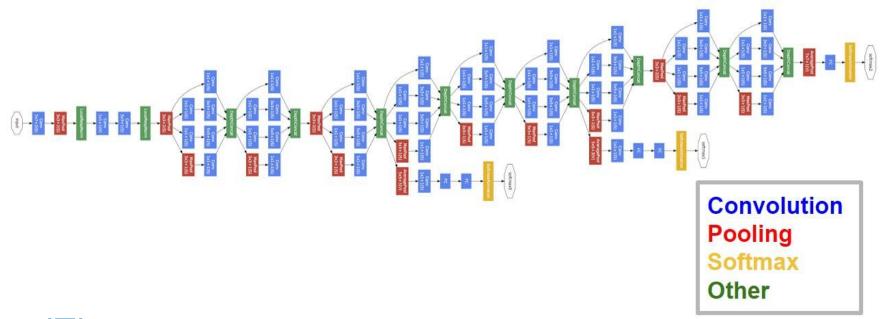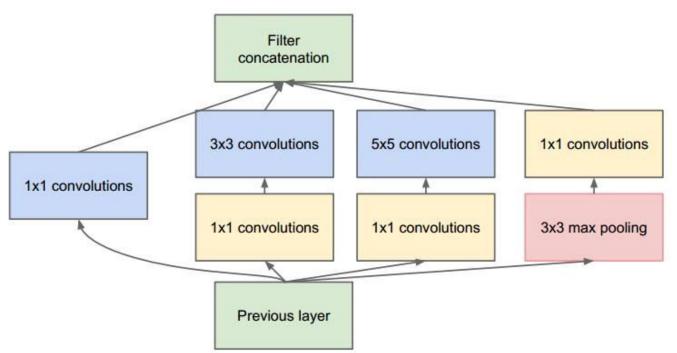
Convolution
Pooling
Softmax
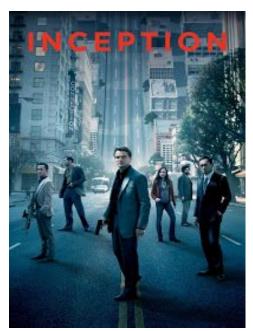**Other**

GoogLeNet
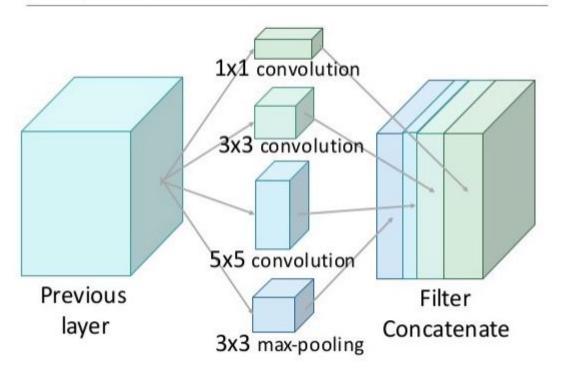
(b) Inception module with dimension reductions

How Inception work

Inception Module

1x1 convolution
3x3 convolution
5x5 convolution
3x3 max-pooling
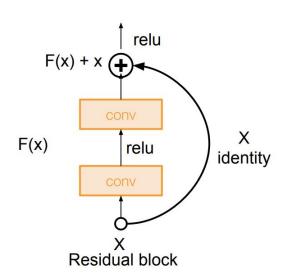Previous layer
Filter Concatenate
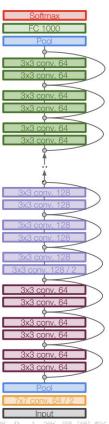
ResNet [paper](#)

Case Study: ResNet

[He et al., 2015]

Very deep networks using residual connections

- 152-layer model for ImageNet
- ILSVRC'15 classification winner (3.57% top 5 error)
- Swept all classification and detection competitions in ILSVRC'15 and COCO'15!

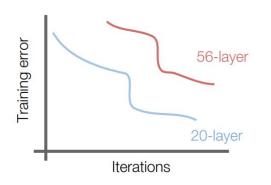Residual block

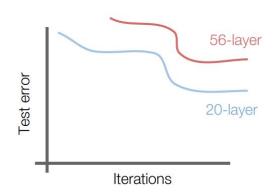$F(x) + x$, relu, X identity, F(x), X

ResNet

What happens when we continue stacking deeper layers on a "plain" convolutional neural network?
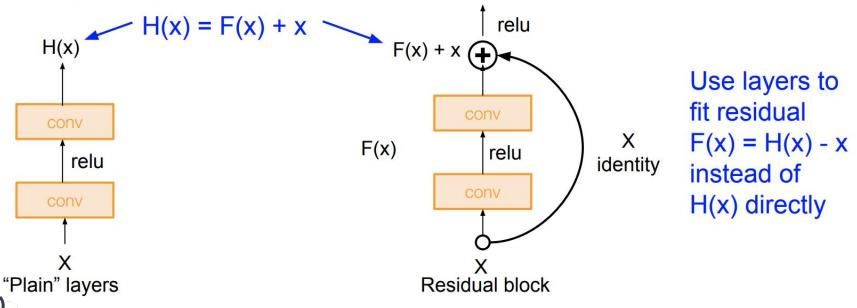

56-layer model performs worse on both training and test error
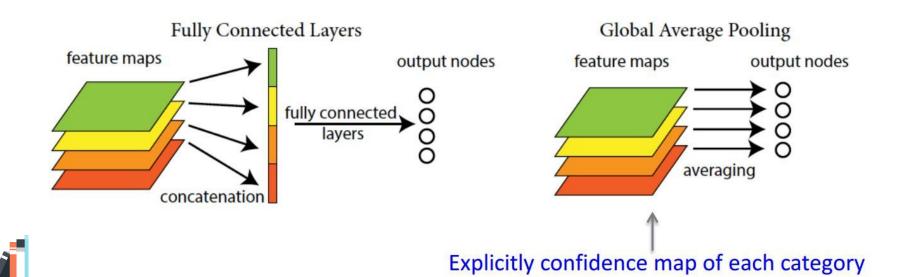-> The deeper model performs worse, but it's not caused by overfitting!

ResNet

Solution: Use network layers to fit a residual mapping instead of directly trying to fit a desired underlying mapping

$$H(x) = F(x) + x$$

H(x)

conv
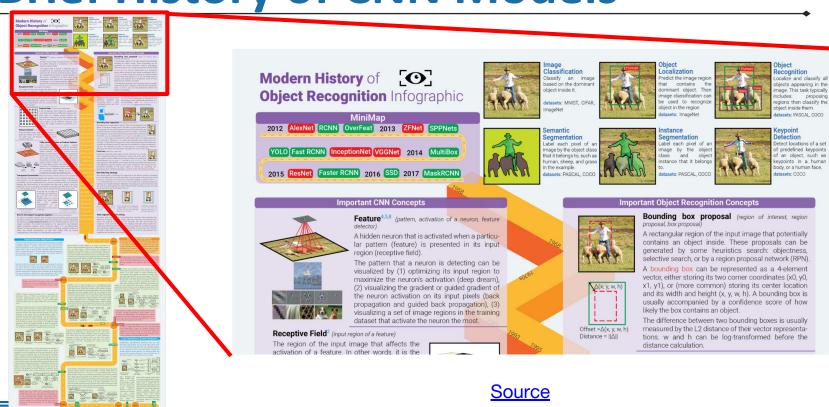
relu

conv

X
"Plain" layers

F(x) + x

relu

conv

relu

F(x)

conv

X
identity

X
Residual block

Use layers to fit residual $F(x) = H(x) - x$ instead of H(x) directly

ResNet

- 使用 Global Average Pooling

A Brief History of CNN Models

A Brief History of CNN Models

推薦閱讀

- [CNN 發展史](#)
- [CNN 經典模型整理](#)
- [ResNet 筆記](#)
- [ResNet & Inception](#)
- [The 9 Deep Learning Papers You Need To Know About](#)

台灣人工智慧學校