

心理與神經資訊學

(Psychoinformatics & Neuroinformatics)

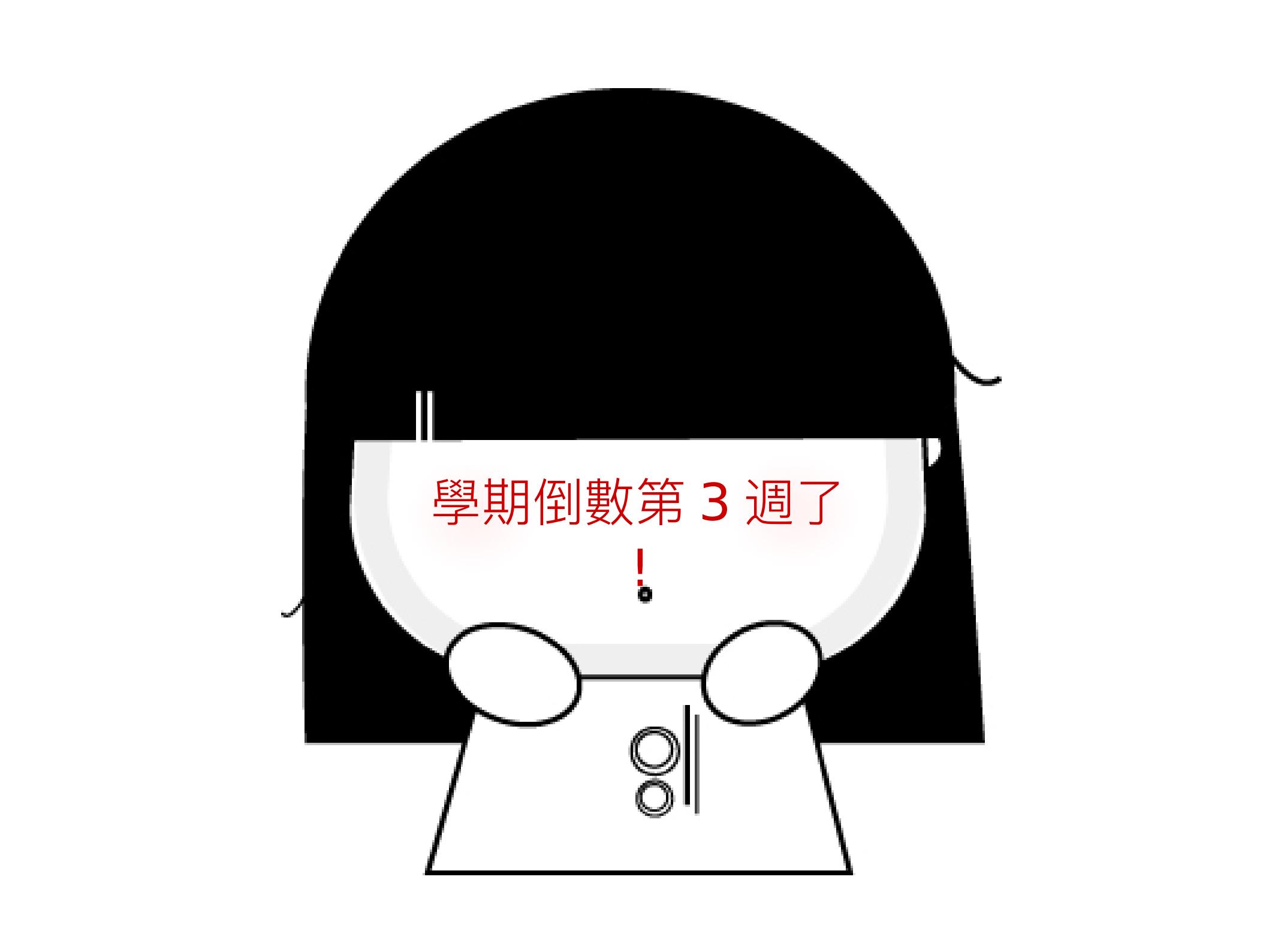
課號：Psy5261

識別碼：227U9340

教室：博雅 101

時間：四 234





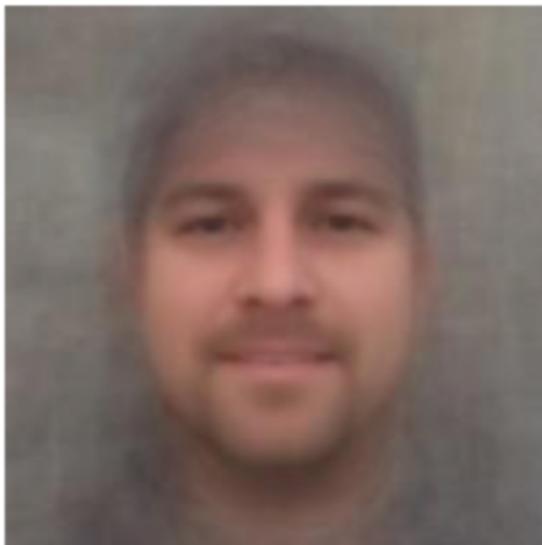
學期倒數第 3 週了

!

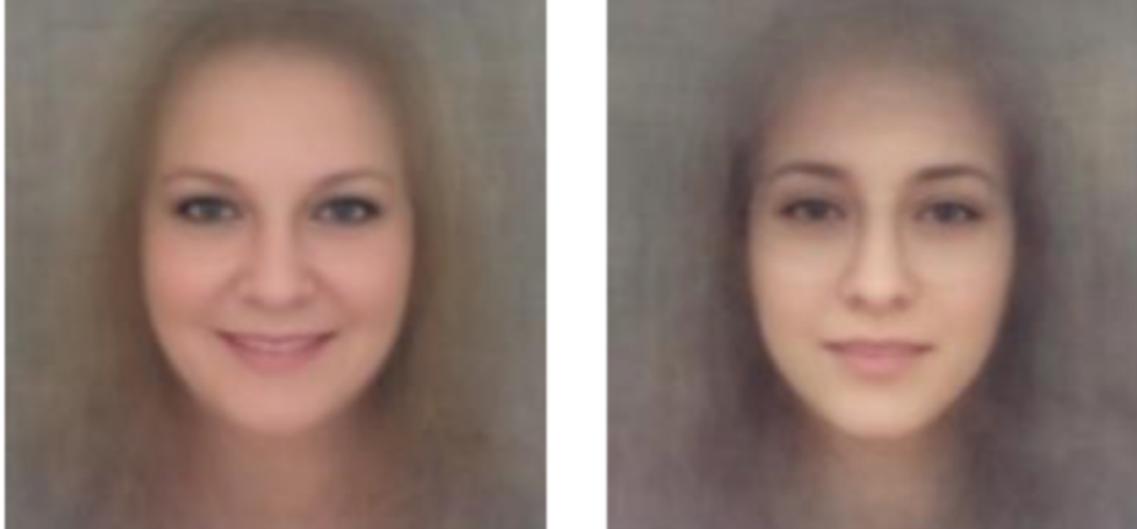
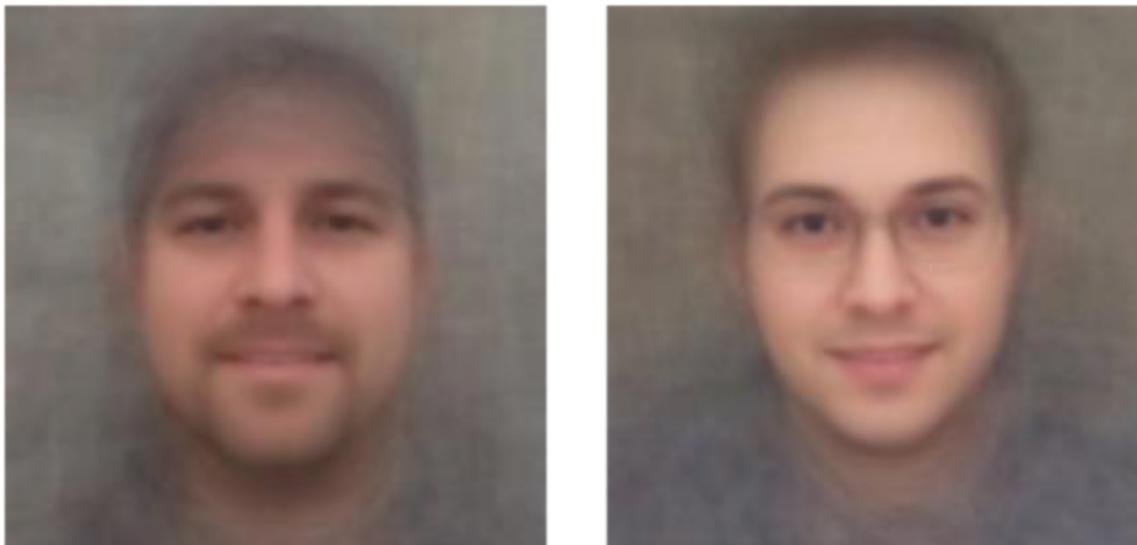
CNN 的應用 (1/3)

利用 VGG-Face 來判斷是否為同志 (classification)

Composite heterosexual faces



Composite gay faces



CNN 的應用 (2/3)

判斷誰可能犯罪 (classification)



(a) Three samples in criminal ID photo set S_c .



(b) Three samples in non-criminal ID photo set S_n

Figure 1. Sample ID photos in our data set.

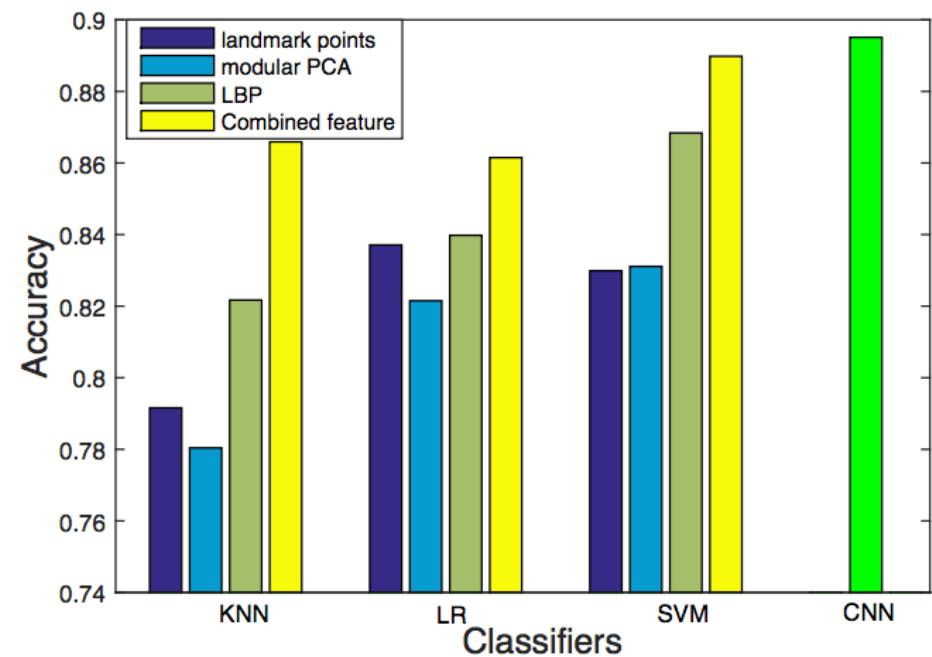


Figure 2. Accuracy of all four classifiers in all thirteen cases.

CNN 的應用 (3/3)

判斷誰比較美 (regression)



TABLE VI. CORRELATION COEFFICIENTS IN SINGLE NETWORK

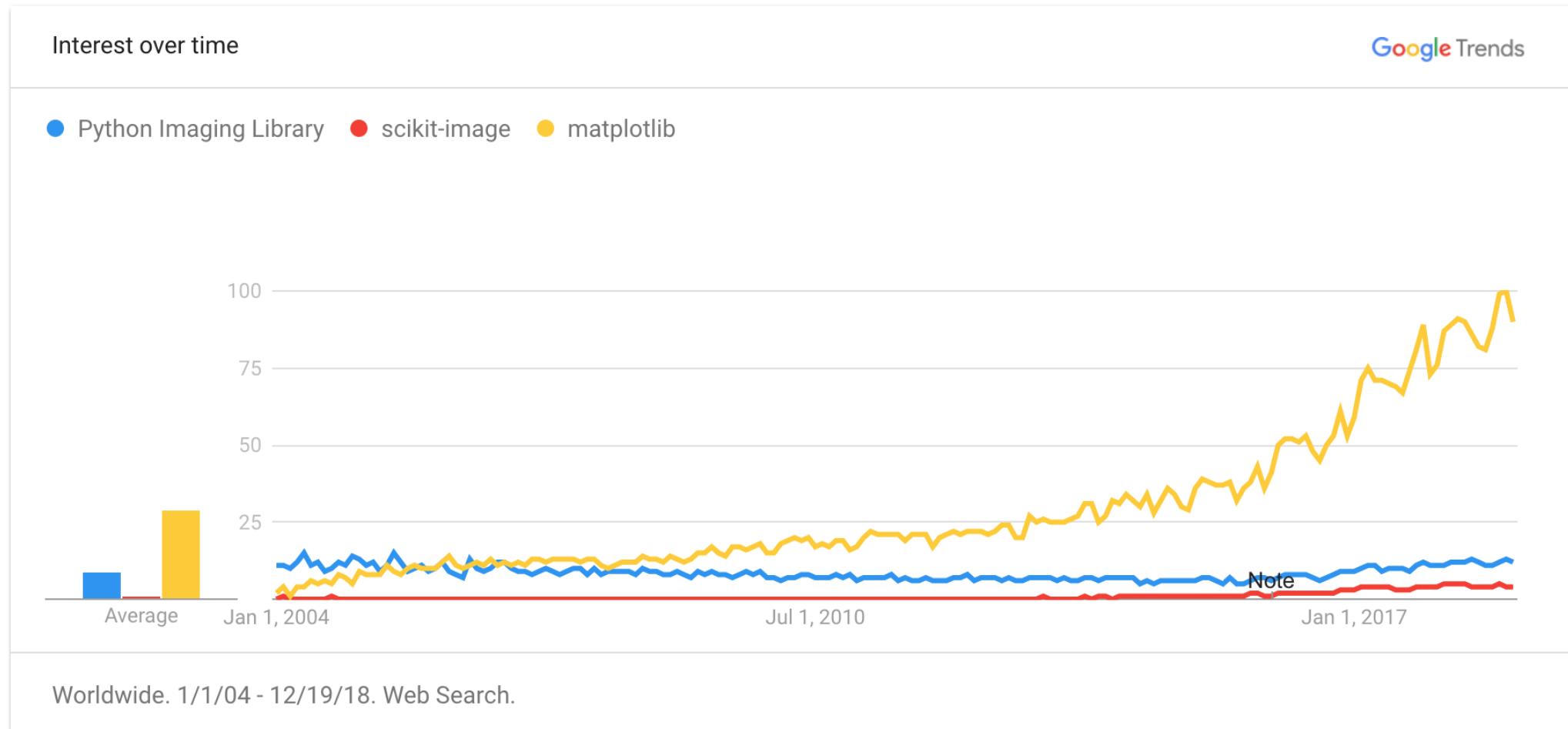
<i>Exp.</i>	1	2	3	4	5	Average
PC	0.8509	0.8050	0.8112	0.7817	0.8446	0.8187

基本影像處理

(Imaging Processing)

影像處理的套件

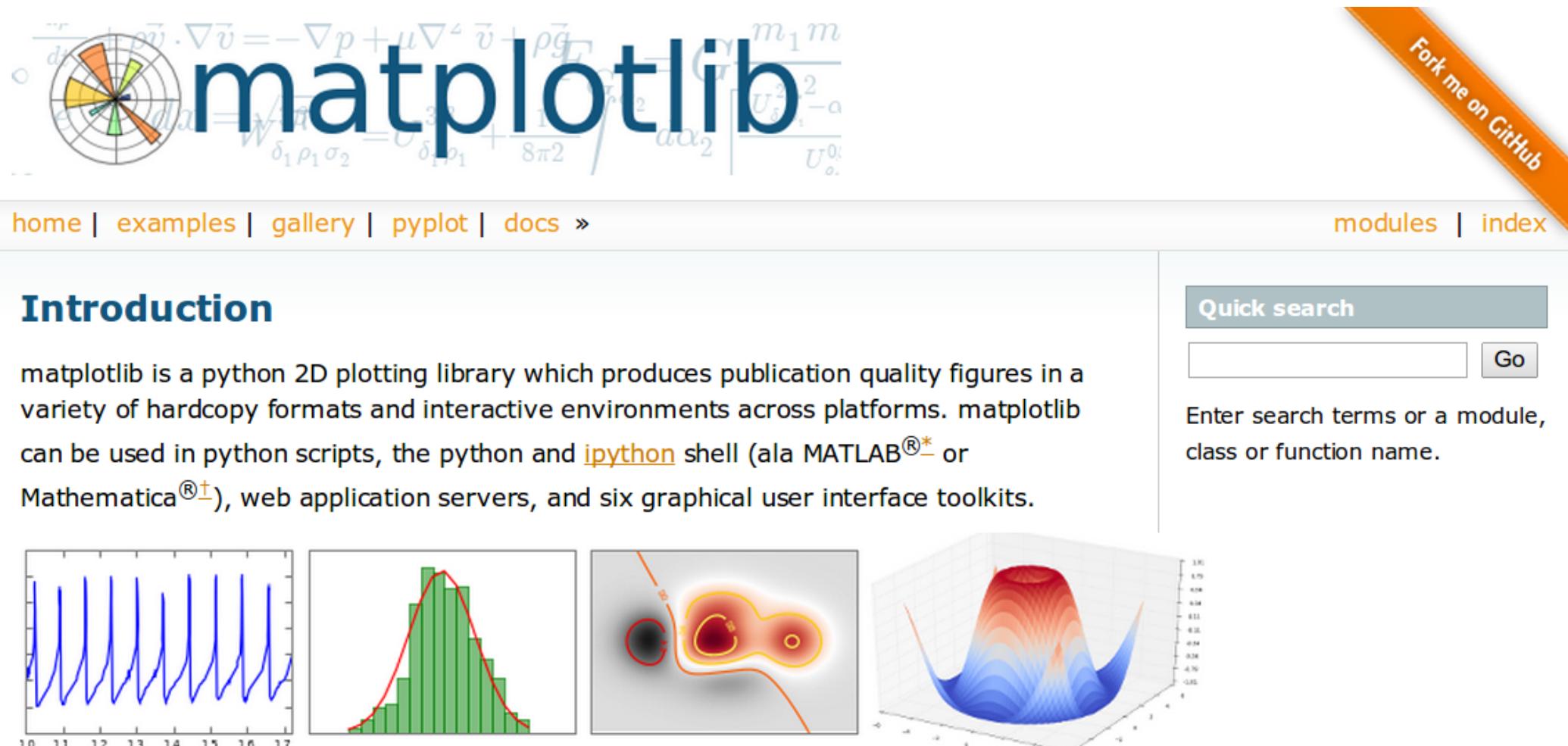
有不同的選擇



簡單處理用 matplotlib; 複雜情況用 PIL/Pillow

Matplotlib 簡介

Matplotlib 其實是模仿 Matlab 的繪圖指令



The screenshot shows the official matplotlib website. At the top, there's a large logo with the word "matplotlib" in blue. To the left of the logo is a circular graphic containing various colored arrows and mathematical symbols. To the right is a diagonal orange bar with the text "Fork me on GitHub". Below the logo, there's a navigation bar with links: "home | examples | gallery | pyplot | docs »" on the left and "modules | index" on the right. A "Quick search" input field is located on the right side of the page. The main content area features four small plots: a line plot with vertical spikes, a histogram with a red normal distribution curve, a contour plot of a complex shape, and a 3D surface plot of a bell-shaped curve.

matplotlib

Fork me on GitHub

home | examples | gallery | pyplot | docs » modules | index

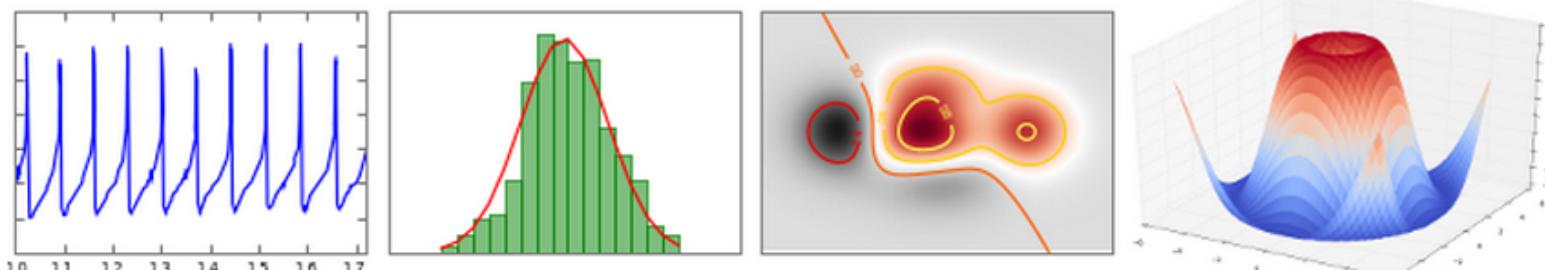
Quick search

Enter search terms or a module, class or function name.

Go

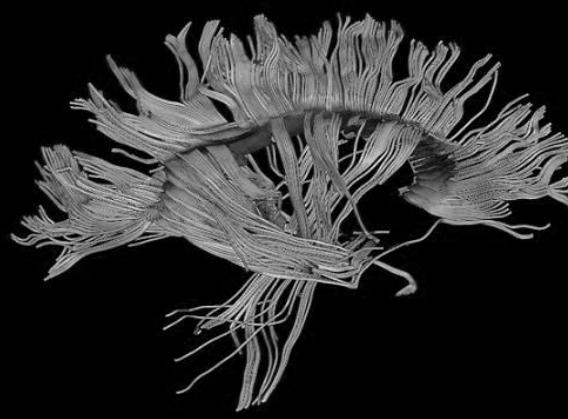
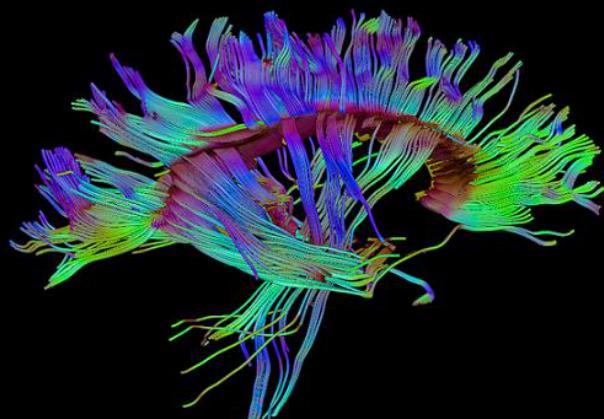
Introduction

matplotlib is a python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. matplotlib can be used in python scripts, the python and [ipython](#) shell (ala MATLAB®* or Mathematica®†), web application servers, and six graphical user interface toolkits.



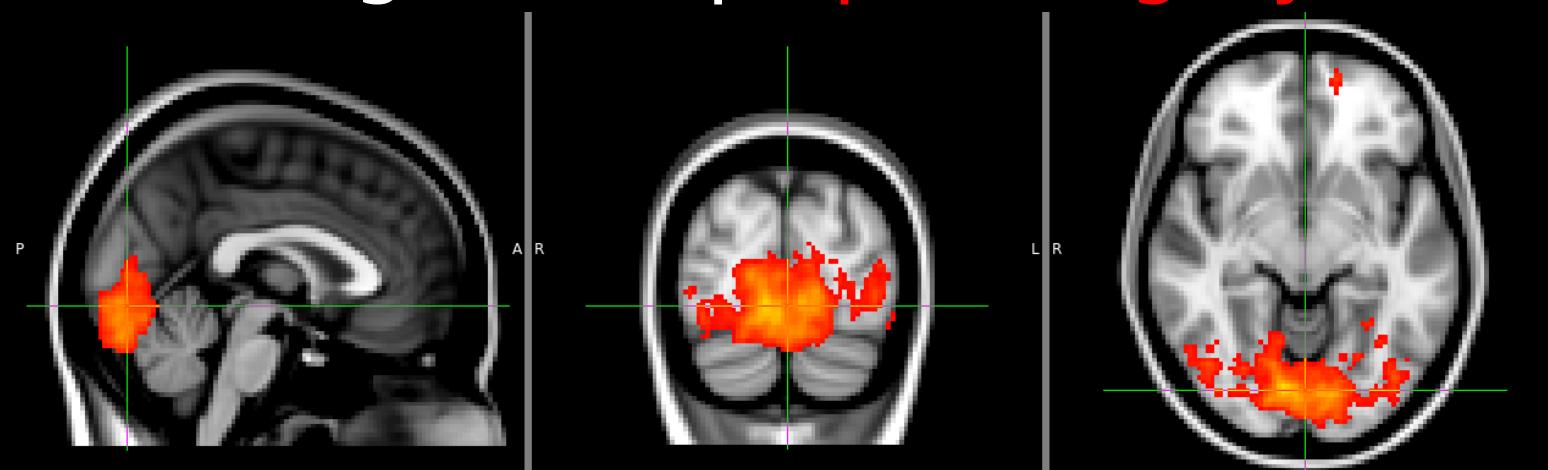
Matplotlib 的基本影像處理 (1/3)

```
import matplotlib.pyplot as plt  
Import numpy as np, scipy.misc  
plt.close('all'); img=plt.imread('DTI.jpg')  
print(img.shape,type(img),img.dtype)  
plt.imshow(img); plt.figure()  
img2=np.mean(img,2); print(img2.shape)  
plt.imshow(img2,cmap=plt.cm.gray)  
scipy.misc.imsave('DTI_gray.jpg',img2)
```



Matplotlib 的基本影像處理 (2/3)

```
import numpy as np,matplotlib.pyplot as plt  
img=[]  
img.append(np.float64(plt.imread('MRI1.jpg')))  
img.append(np.float64(plt.imread('MRI2.jpg')))  
img.append(img[1]-img[0]) #contrast  
for i in range(3):  
    plt.subplot(1,3,i+1); plt.axis('off')  
    plt.imshow(img[i],cmap=plt.cm.gray)
```



Matplotlib 的基本影像處理 (3/3)

```
Import numpy as np, matplotlib.pyplot as plt  
plt.close('all'); img=[]  
img.append(plt.imread('face.jpg')) # for FFA  
img.append(plt.imread('house.jpg')) # for PPA  
k=np.arange(1,10,2)/10.0  
for i in range(5):  
    plt.subplot(1,5,i+1);plt.axis('off')  
    hybrid=k[i]*img[0]+(1-k[i])*img[1]  
    plt.imshow(hybrid/255.0)
```



恰當的視覺化幫助洞察資料 (1/2)

不同年代的男女外貌有何不同？



The Class of 1988 (left panel)



The Class of 1988 (right panel)



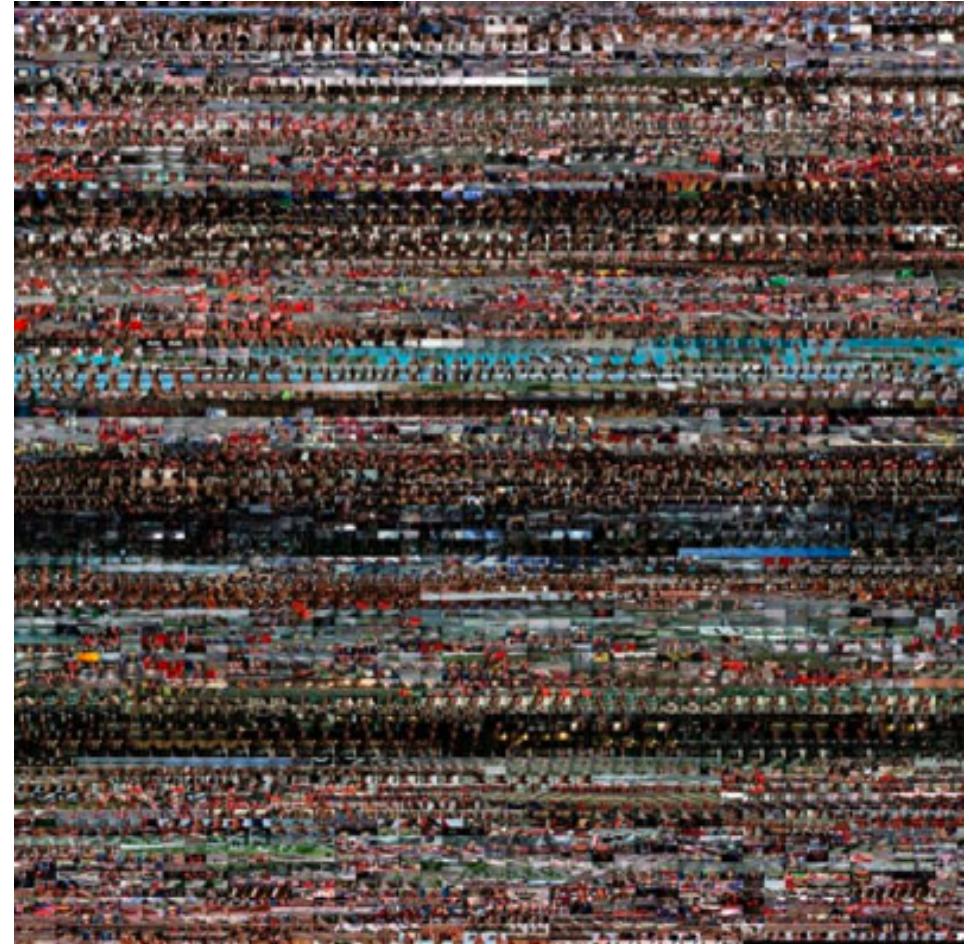
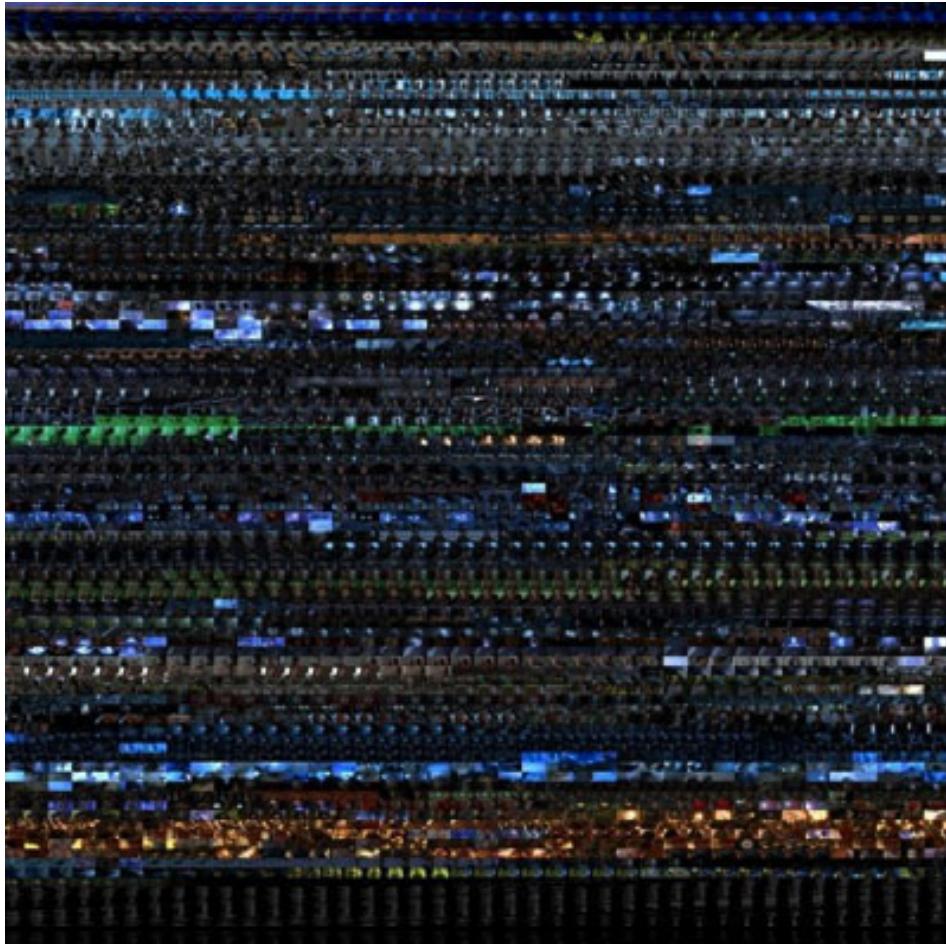
The Class of 1967 (left panel)



The Class of 1967 (right panel)

恰當的視覺化幫助洞察資料 (2/2)

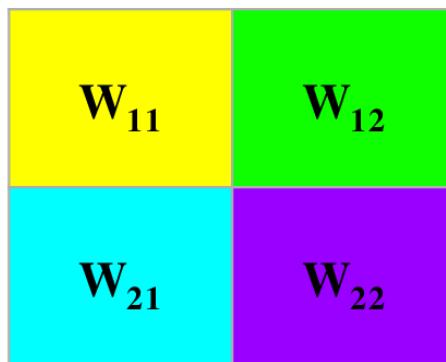
不同類別影片的主色調相異



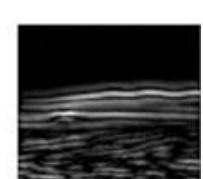
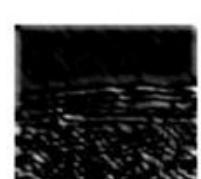
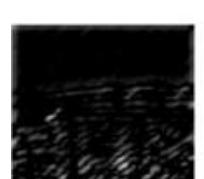
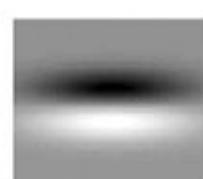
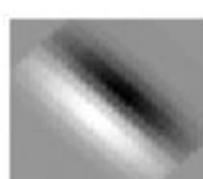
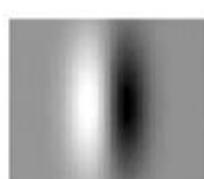
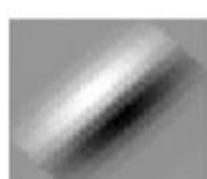
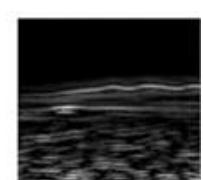
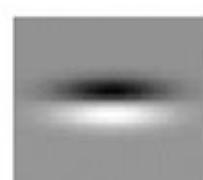
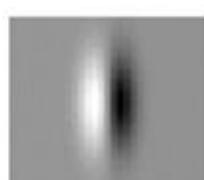
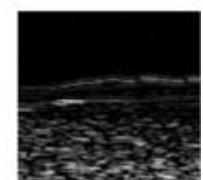
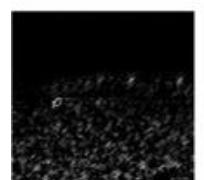
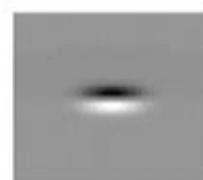
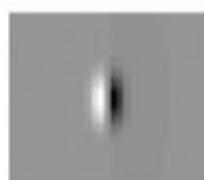
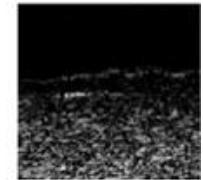
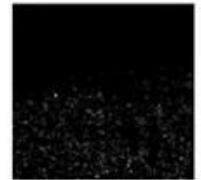
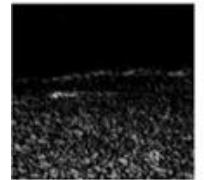
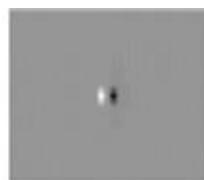
2D Spatial Convolution

Convolution 與 Correlation 傻傻分不清楚

X_{11}	X_{12}	X_{13}
X_{21}	X_{22}	X_{23}
X_{31}	X_{32}	X_{33}



2D Spatial Filtering



Padding & Stride

Input 與 Output 的維度關係一定要會算

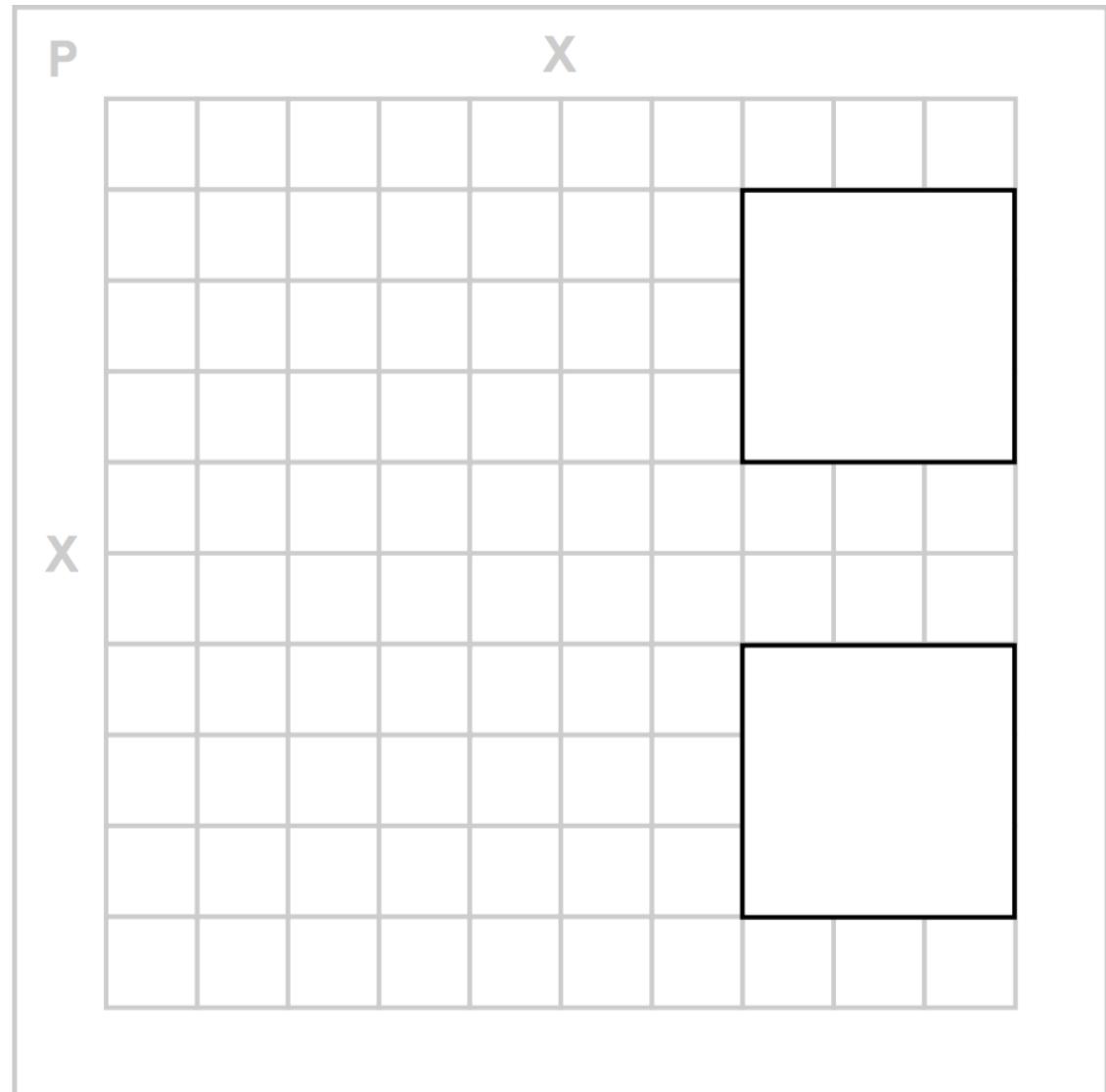
$$1 + \frac{X - F + 2P}{S}$$

X = image size

F = filter size

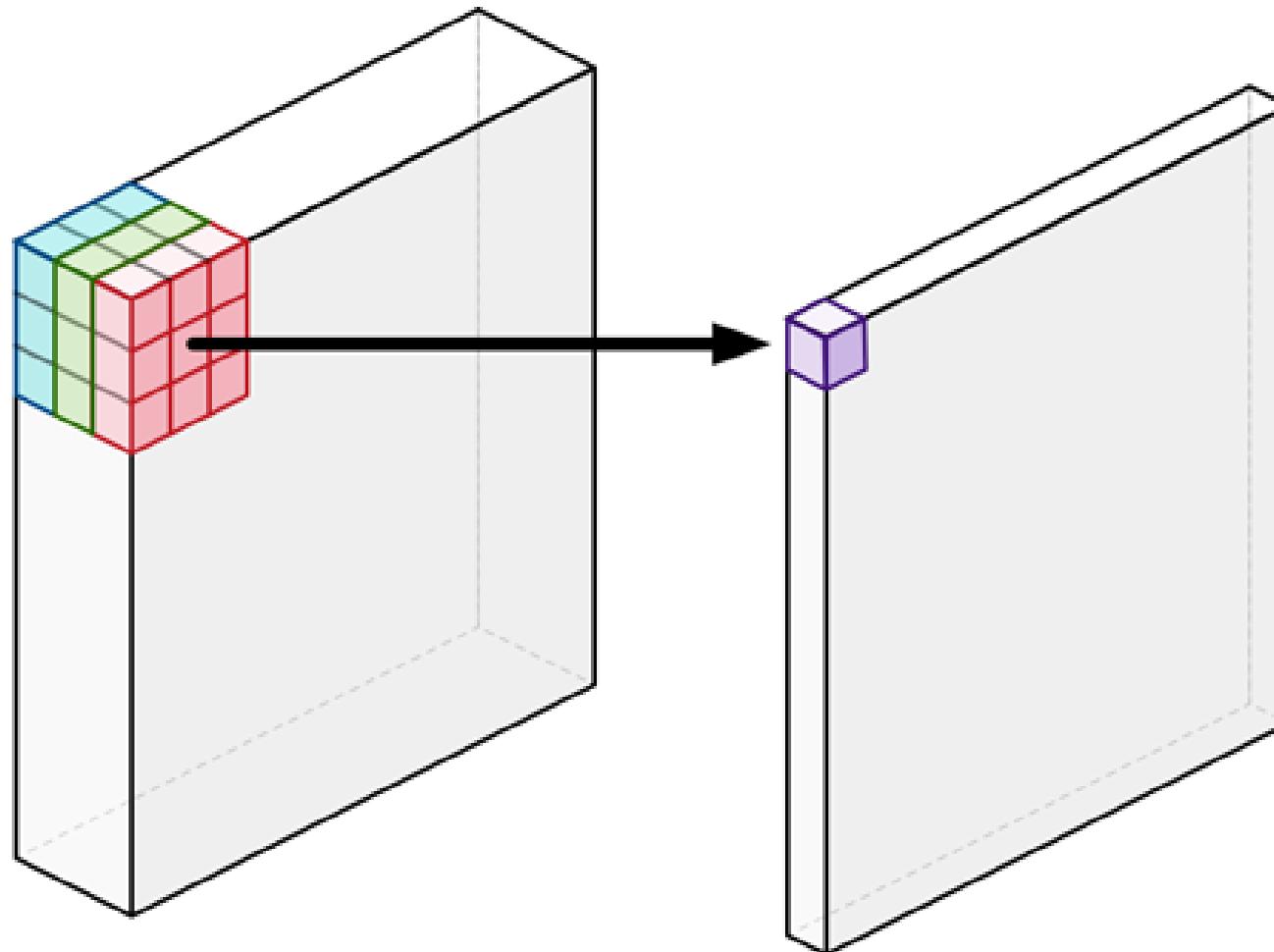
P = padding

S = stride



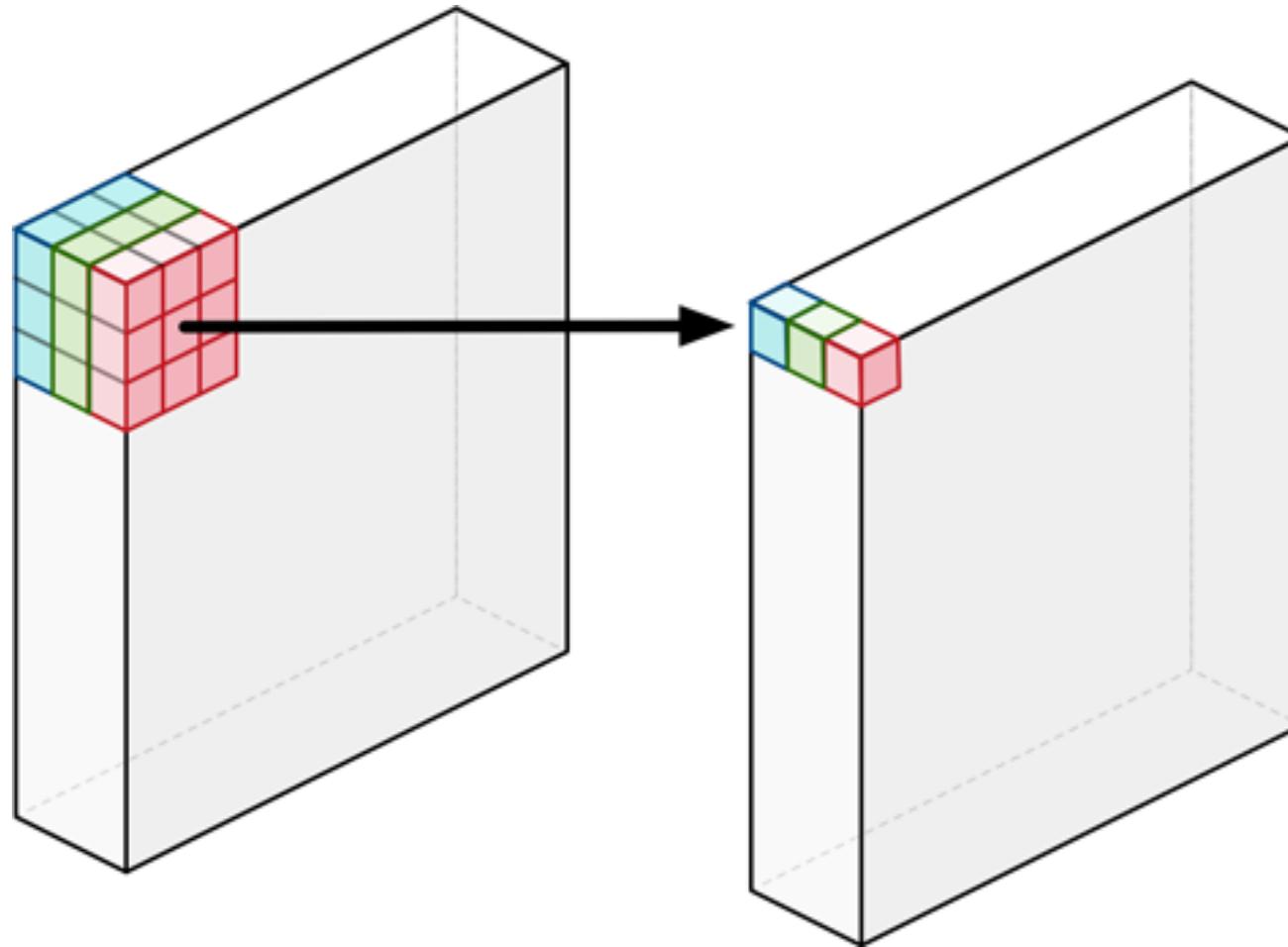
3D Spatial Convolution

和 2D 一樣就是 template matching



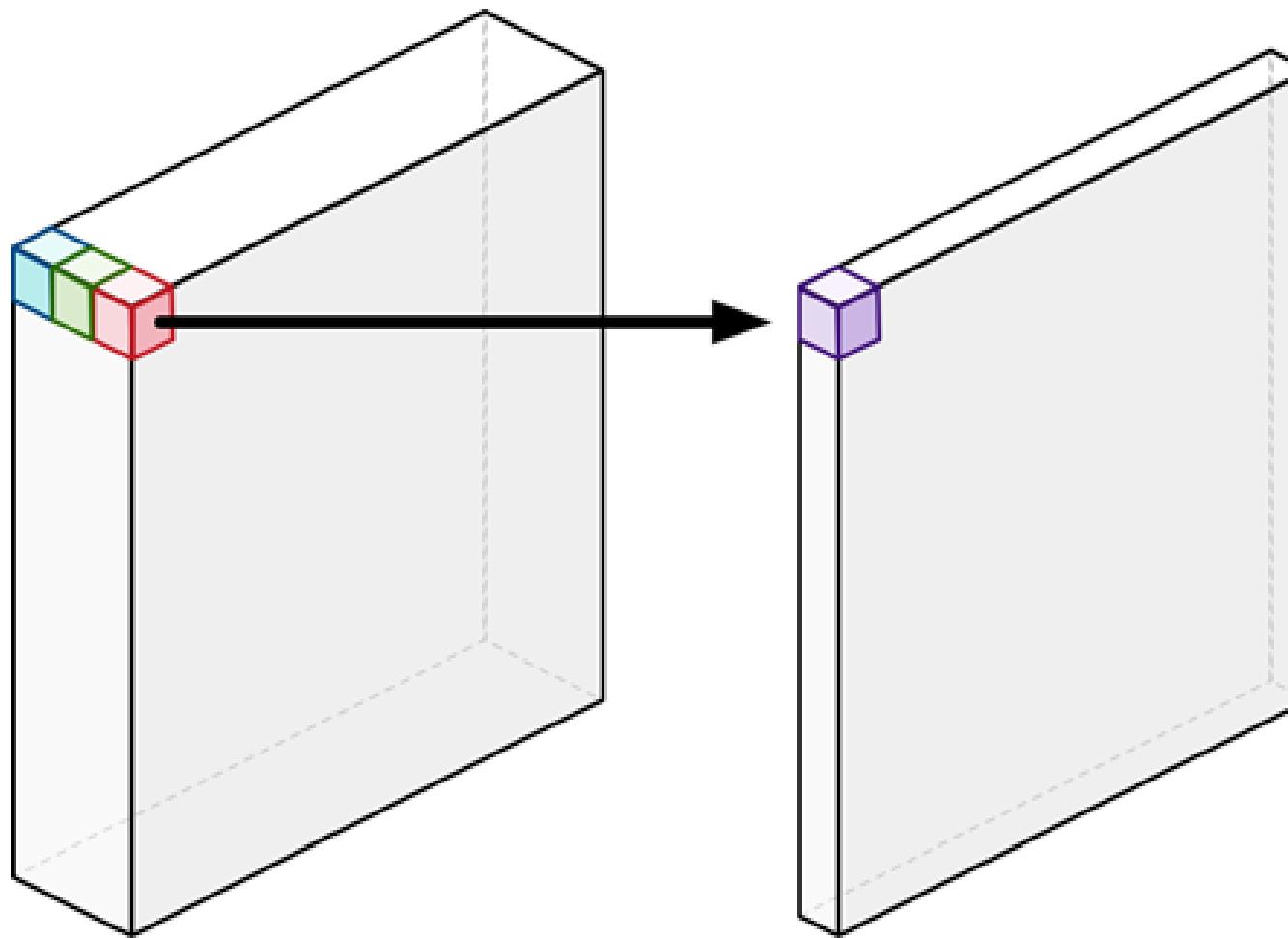
Depthwise Separable Conv.

出現在 Xception Net 及其後繼的 Mobile Net 中



Pointwise Convolution

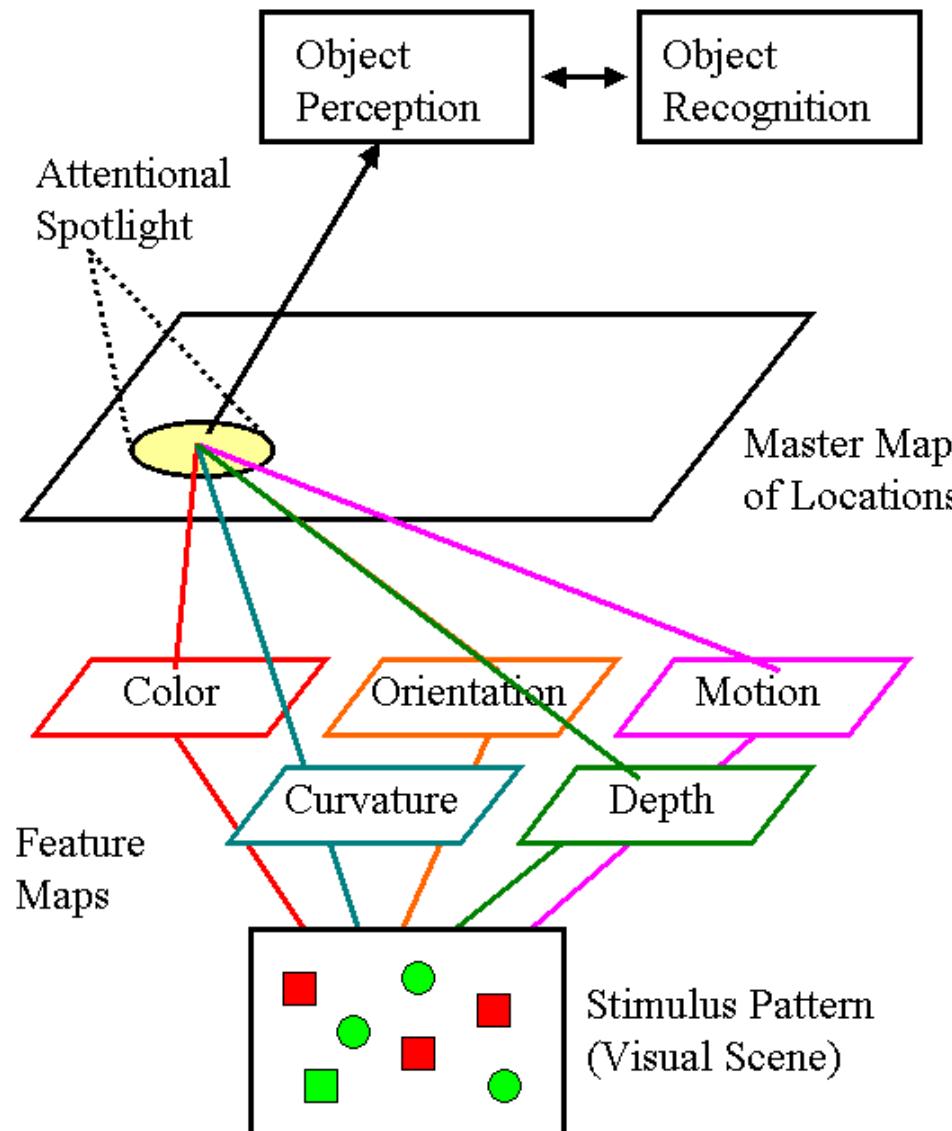
Inception, Xception, Squeeze, Mobile Nets 皆有



Feature Integration Theory

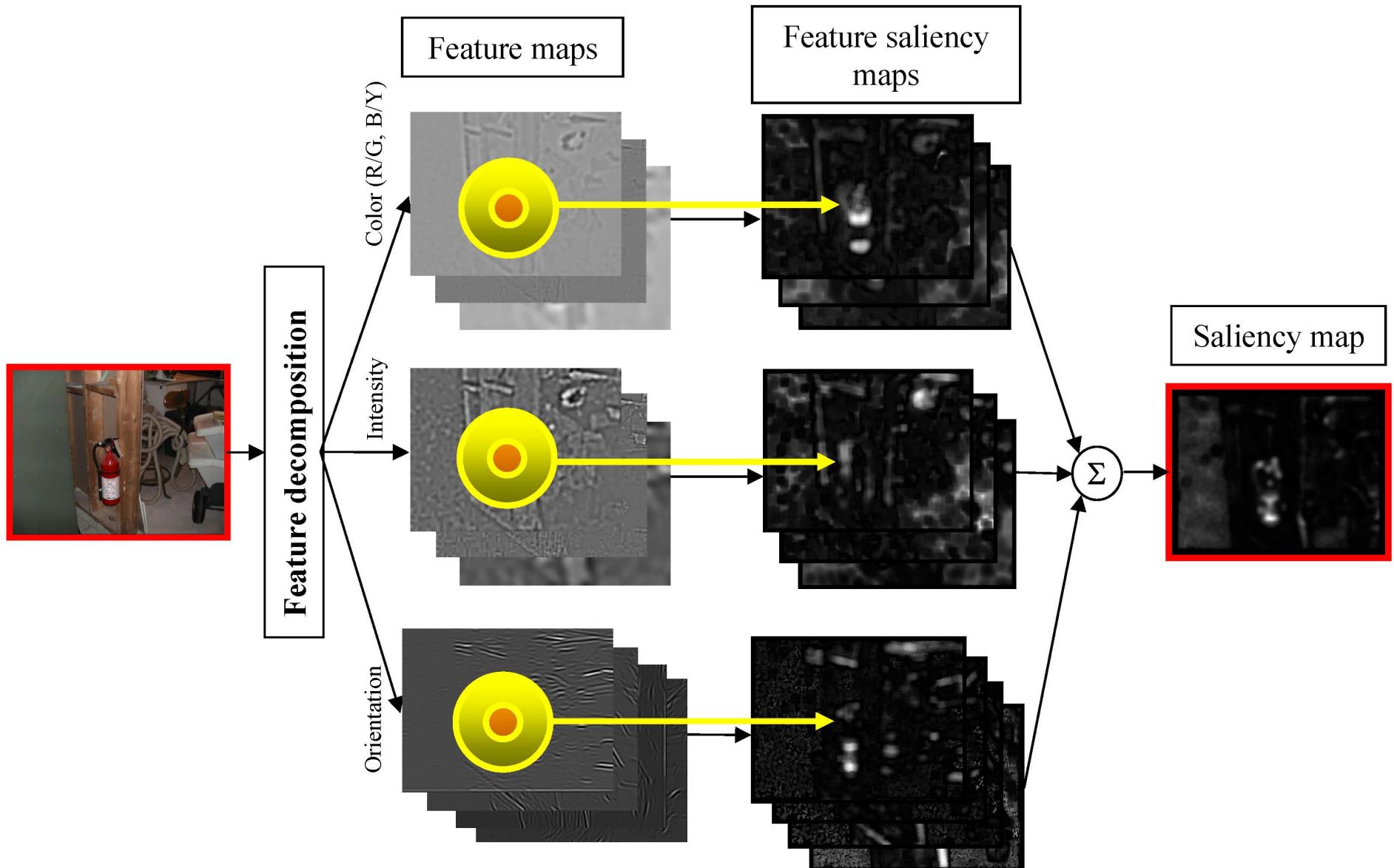
就是 Features 分開處理再整合

Feature Integration Theory (Treisman)



Saliency Map

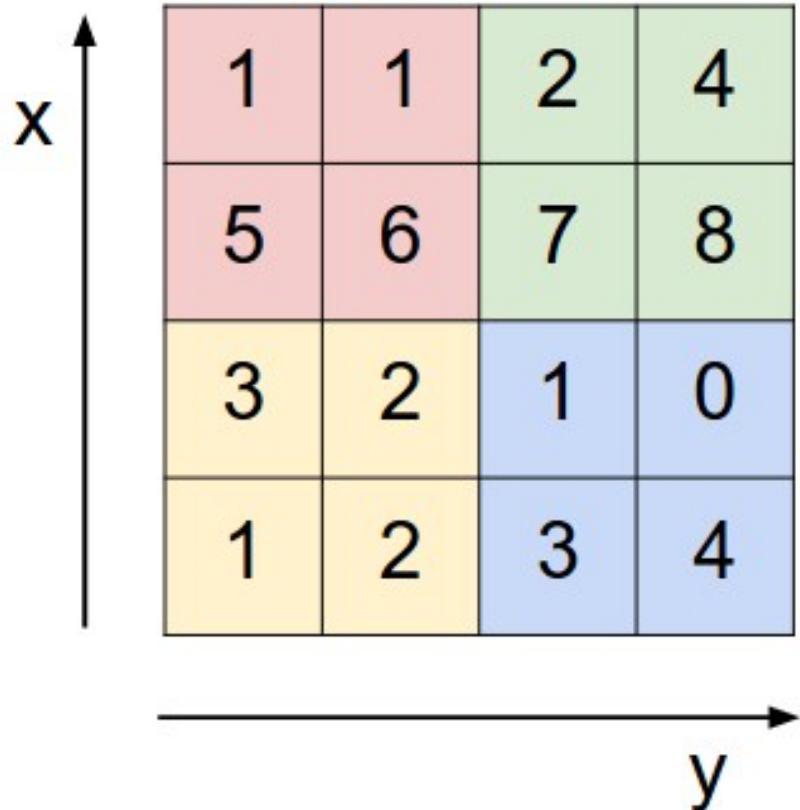
只是 Feature Integration Theory 的實境版



(Max) Pooling

就是 downsampling

Single depth slice



max pool with 2x2 filters
and stride 2

A 2x2 grid representing the output of max pooling. It contains four cells: top-left (6) in pink, top-right (8) in light green, bottom-left (3) in yellow, and bottom-right (4) in medium blue. An arrow points from the input grid to this output grid.

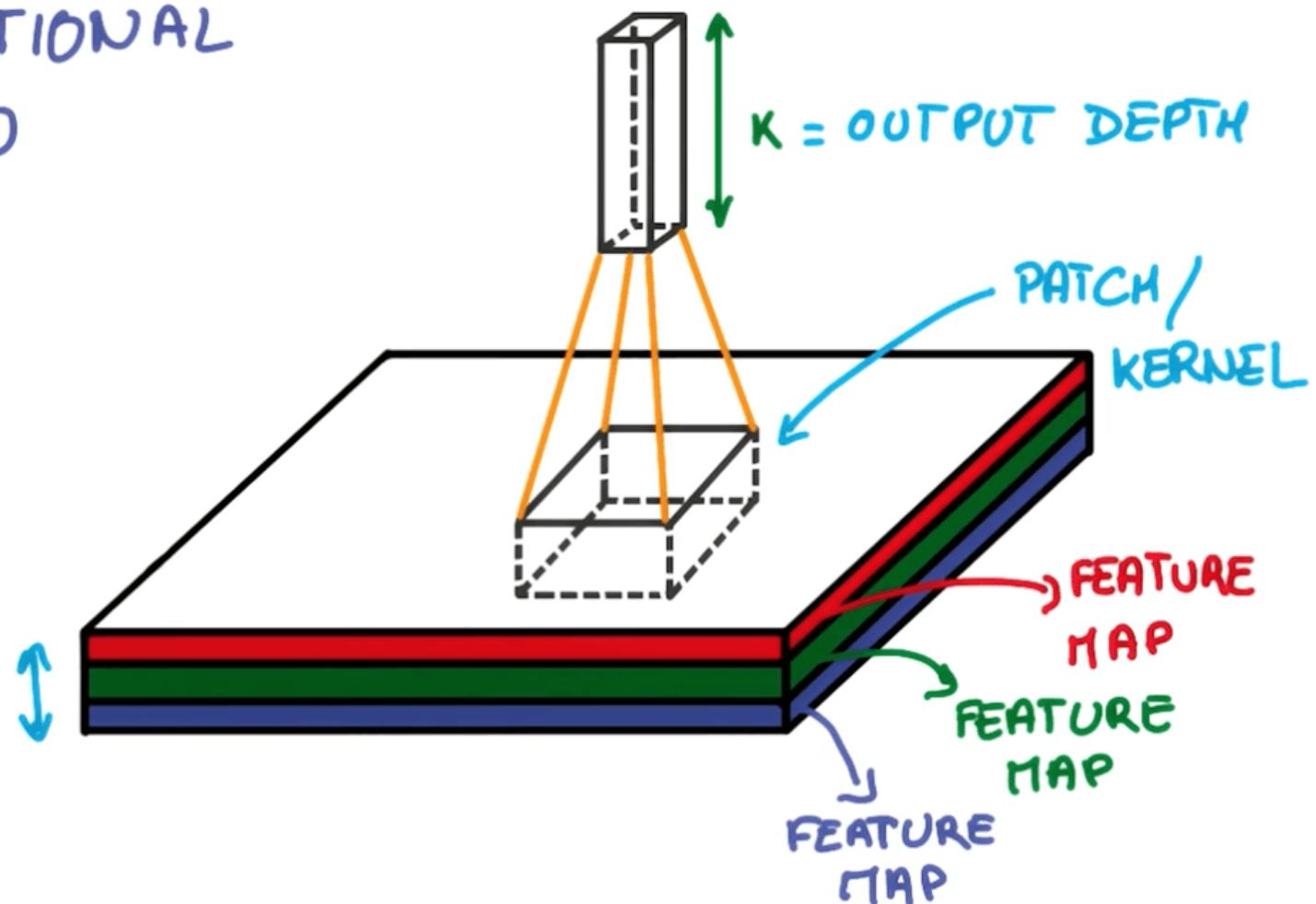
6	8
3	4

CNN Terms

Kernel 對應到的就是生物上的 Receptive Field

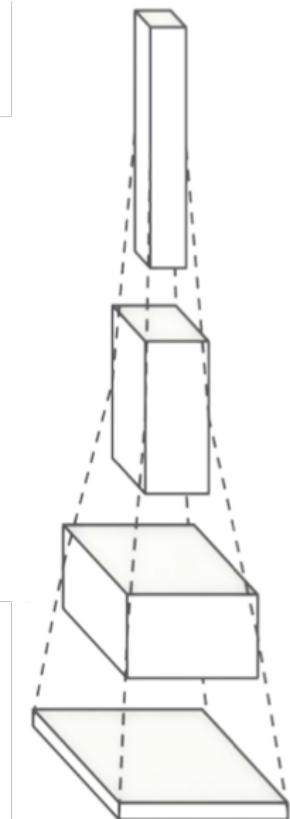
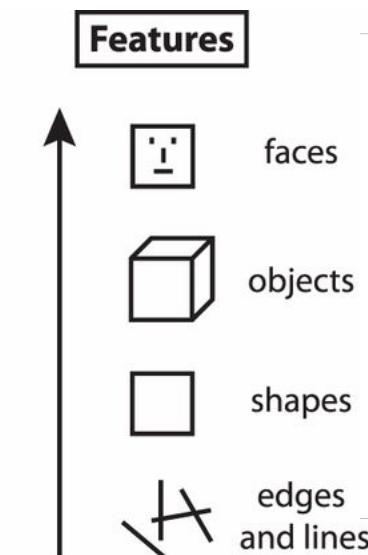
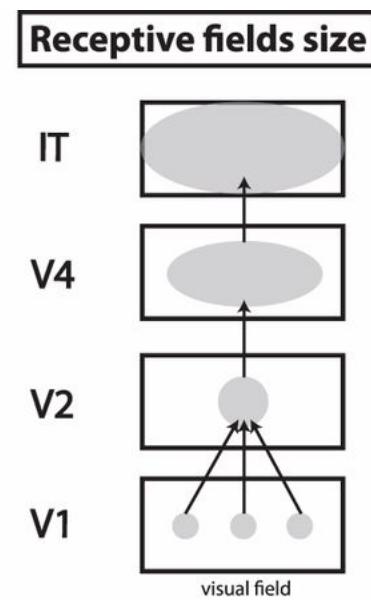
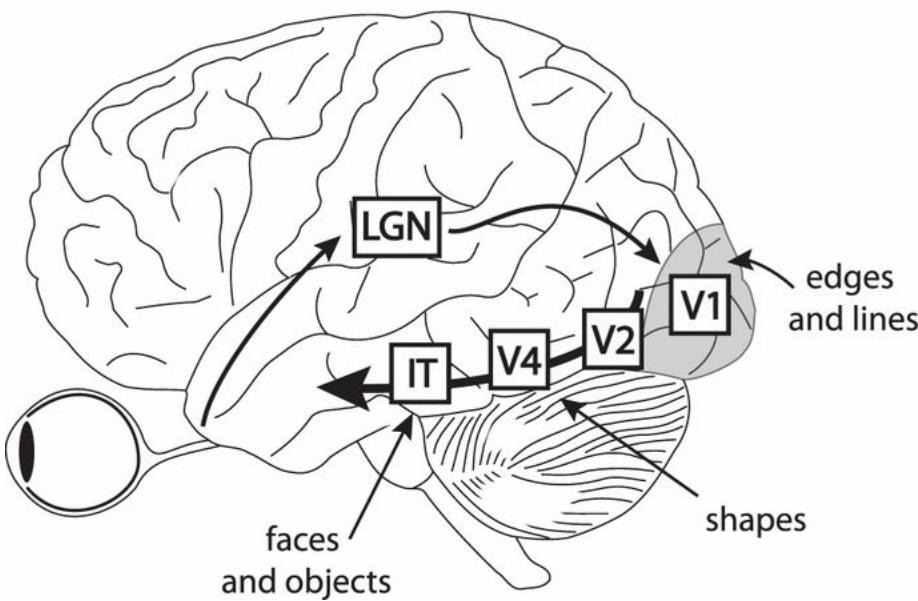
CONVOLUTIONAL
LINGO

INPUT
DEPTH



CNN 的金字塔結構

是在仿 ventral visual pathway in the brain



基本電腦視覺 (Computer Vision)

電腦視覺處理的問題 (1/2)

從 What 到 What-Where

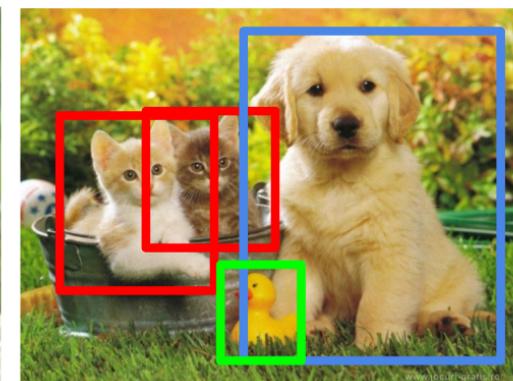
Classification



Classification + Localization



Object Detection



Instance Segmentation



CAT

CAT

CAT, DOG, DUCK

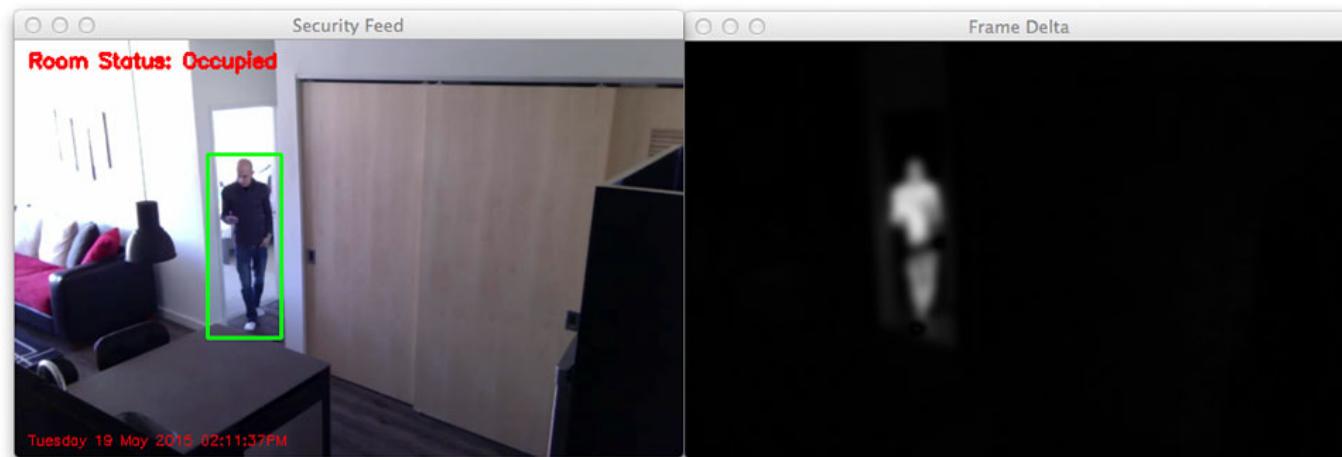
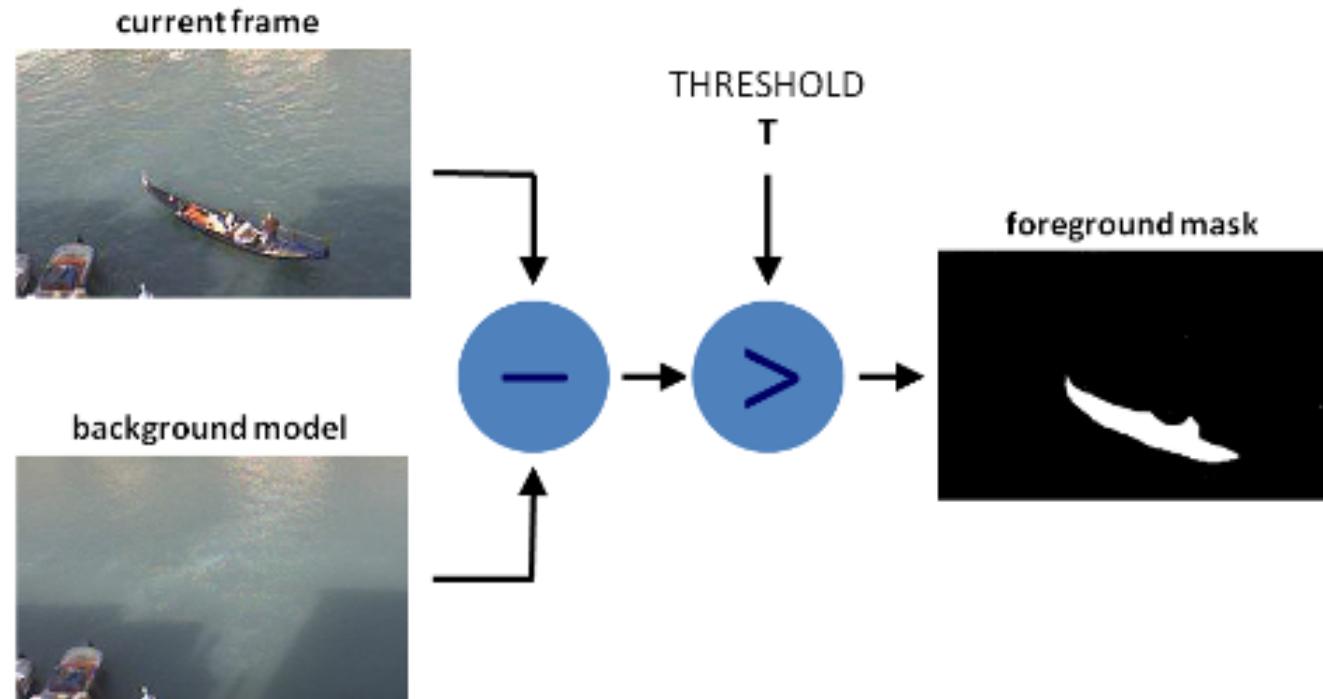
CAT, DOG, DUCK

Single object

Multiple objects

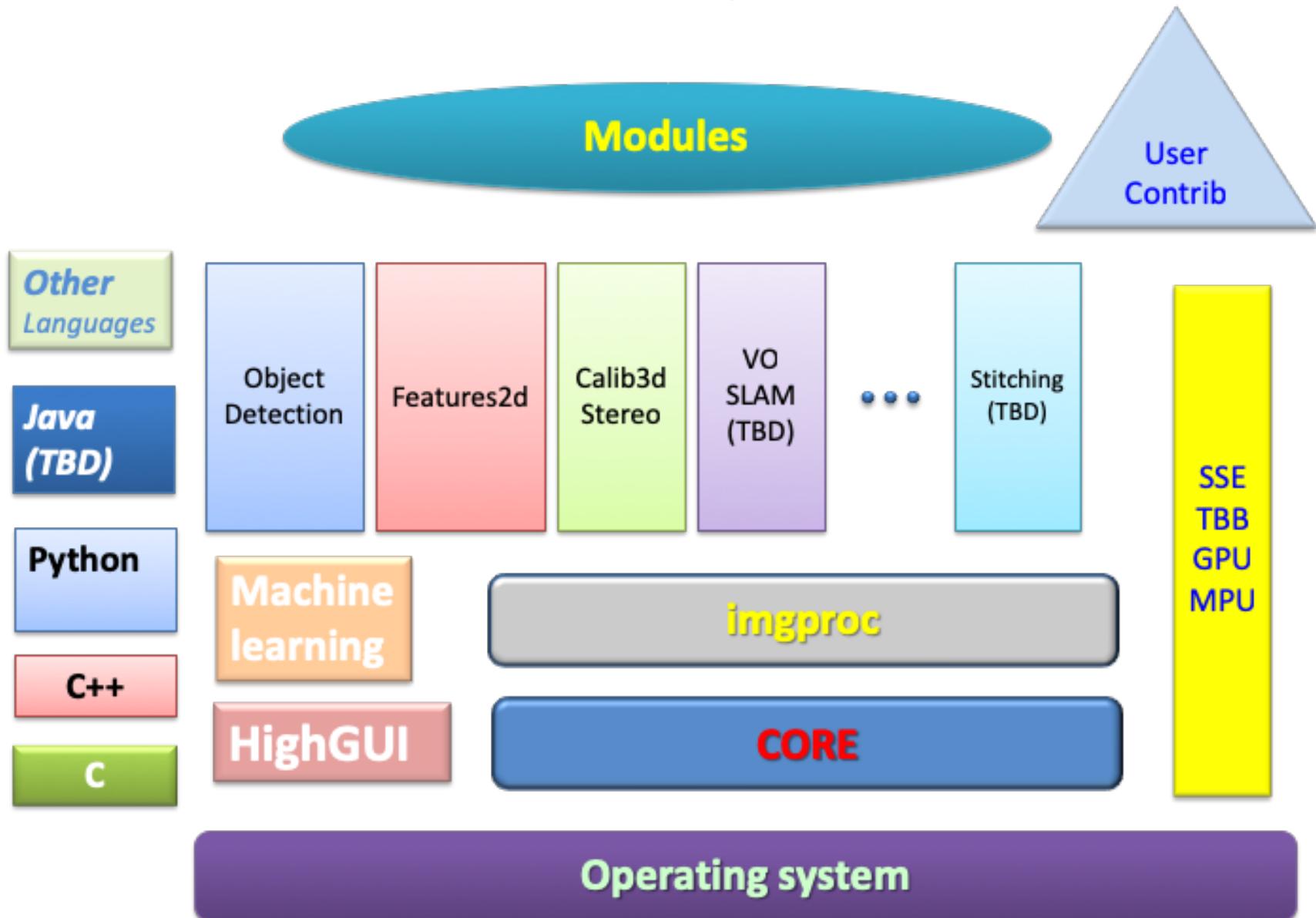
電腦視覺處理的問題 (2/2)

從 images 到 videos



OpenCV (1/2)

有影像處理能力 & 不同語言的 APIs



OpenCV (2/2)

有超過 2,500 種演算法

General Image Processing Functions

Image Erosion Dilation Closing Morphology

Image Pyramids

Coarse-to-fine optical flow estimation

Segmentation

Foreground Background

Geometric descriptors

Image Boundary

Features

Features extracted from a textured surface.

Transforms

Affine (2x2) Perspective (3x3)

Machine Learning:

- Detection,
- Recognition

Optical Flow in 1D

Tracking

Optical Flow in 1D

Matrix Math

Gary Bradski

Camera calibration, Stereo, 3D

Tracking a calibration object

Utilities and Data Structures

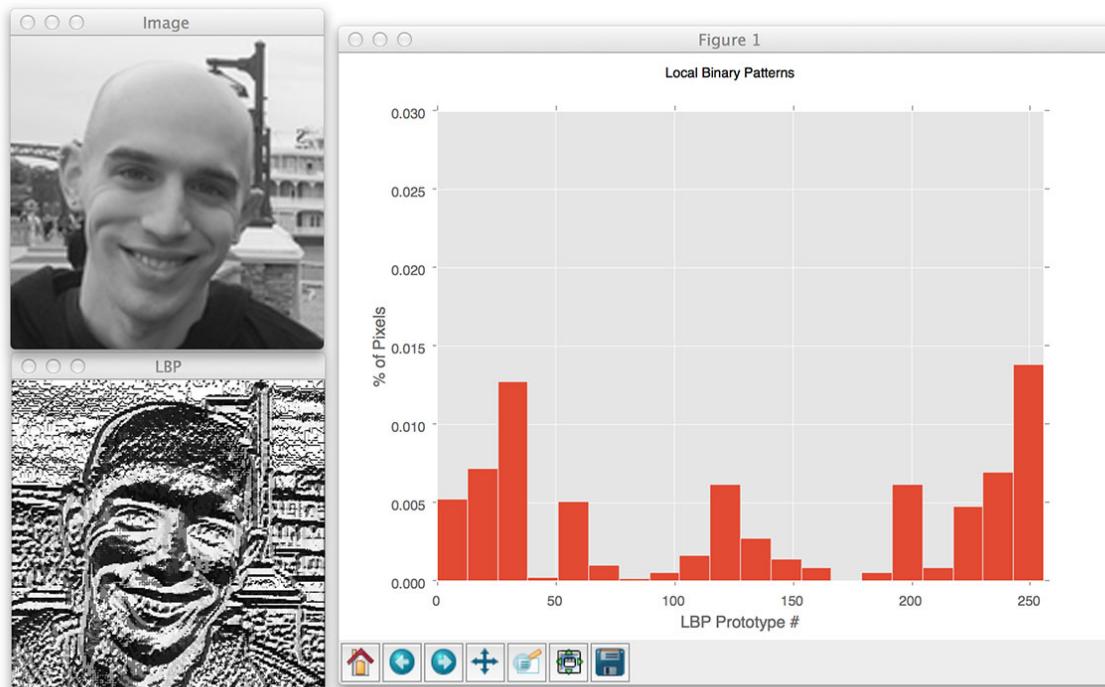
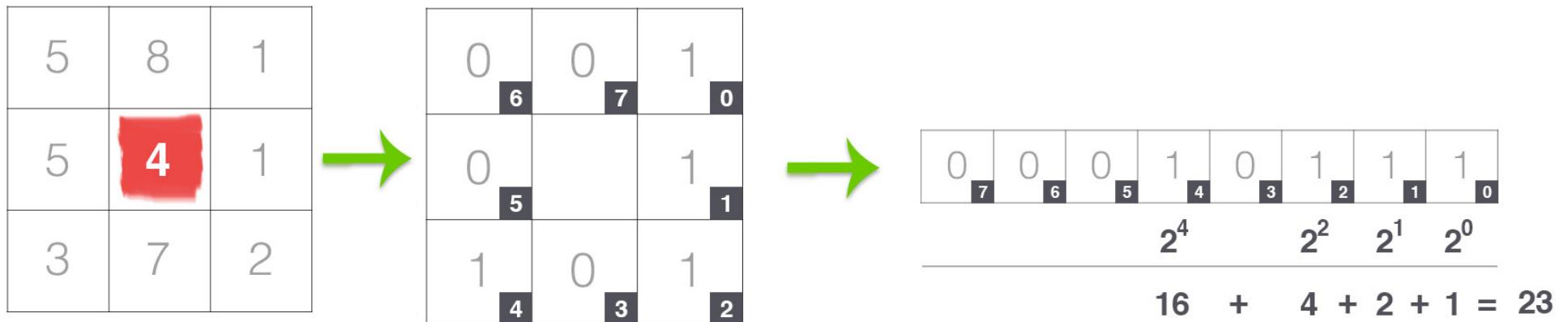
OpenCV

Fitting

Optimized low-level functions

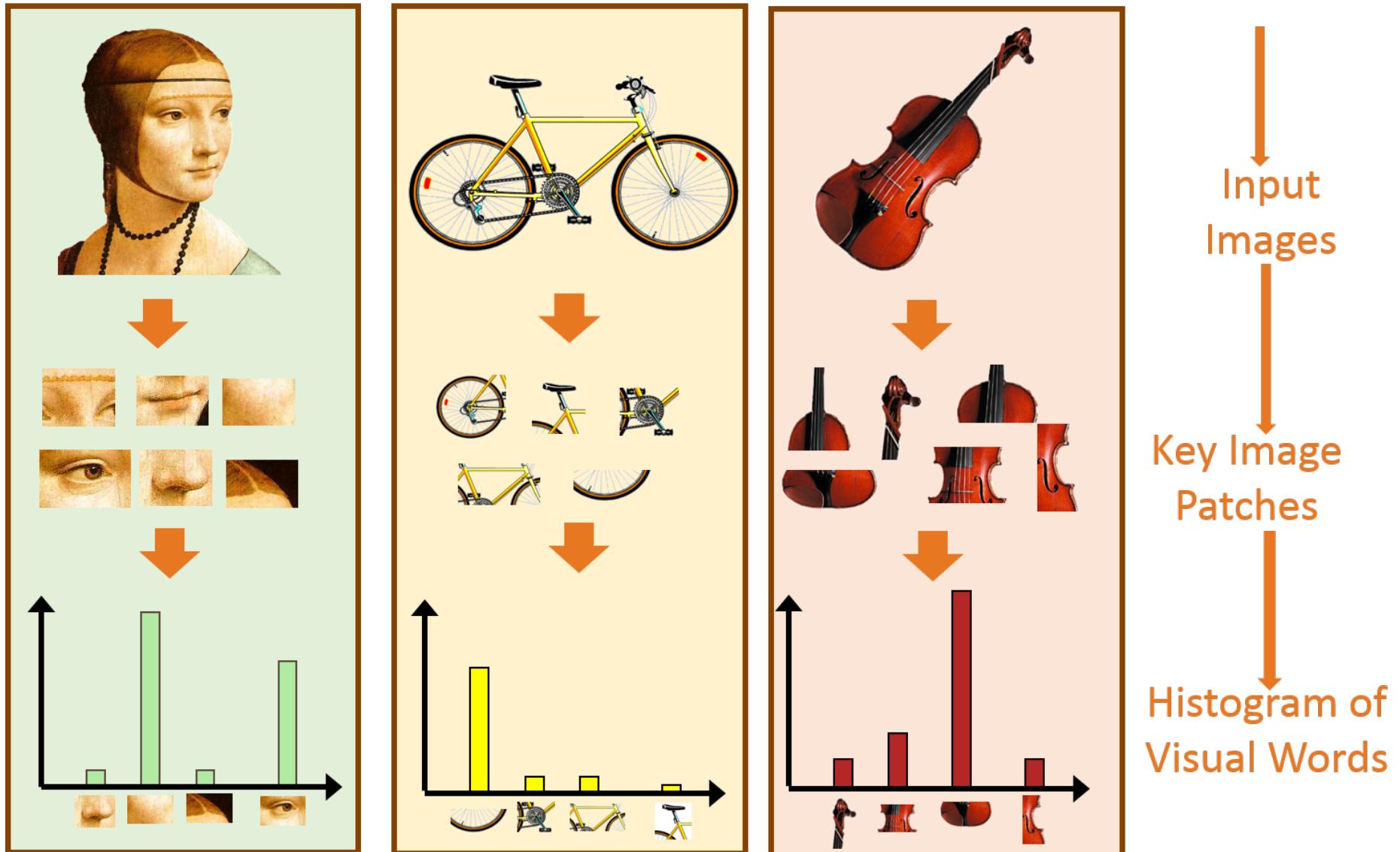
Visual Descriptors: LBP

Features: Local Bin Patterns



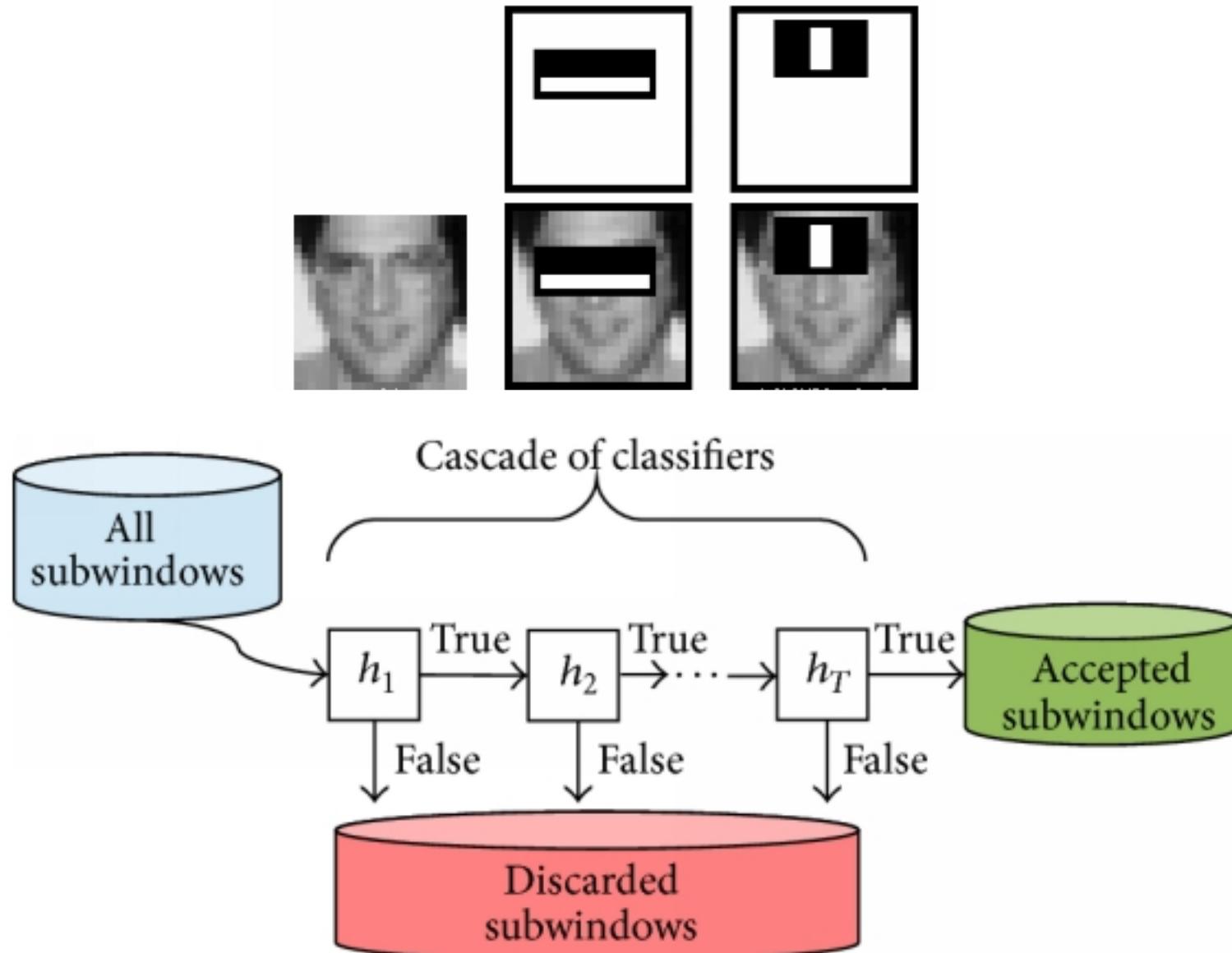
Visual Descriptors: VBOW

Features: Visual Bag of Words



Haar Cascades

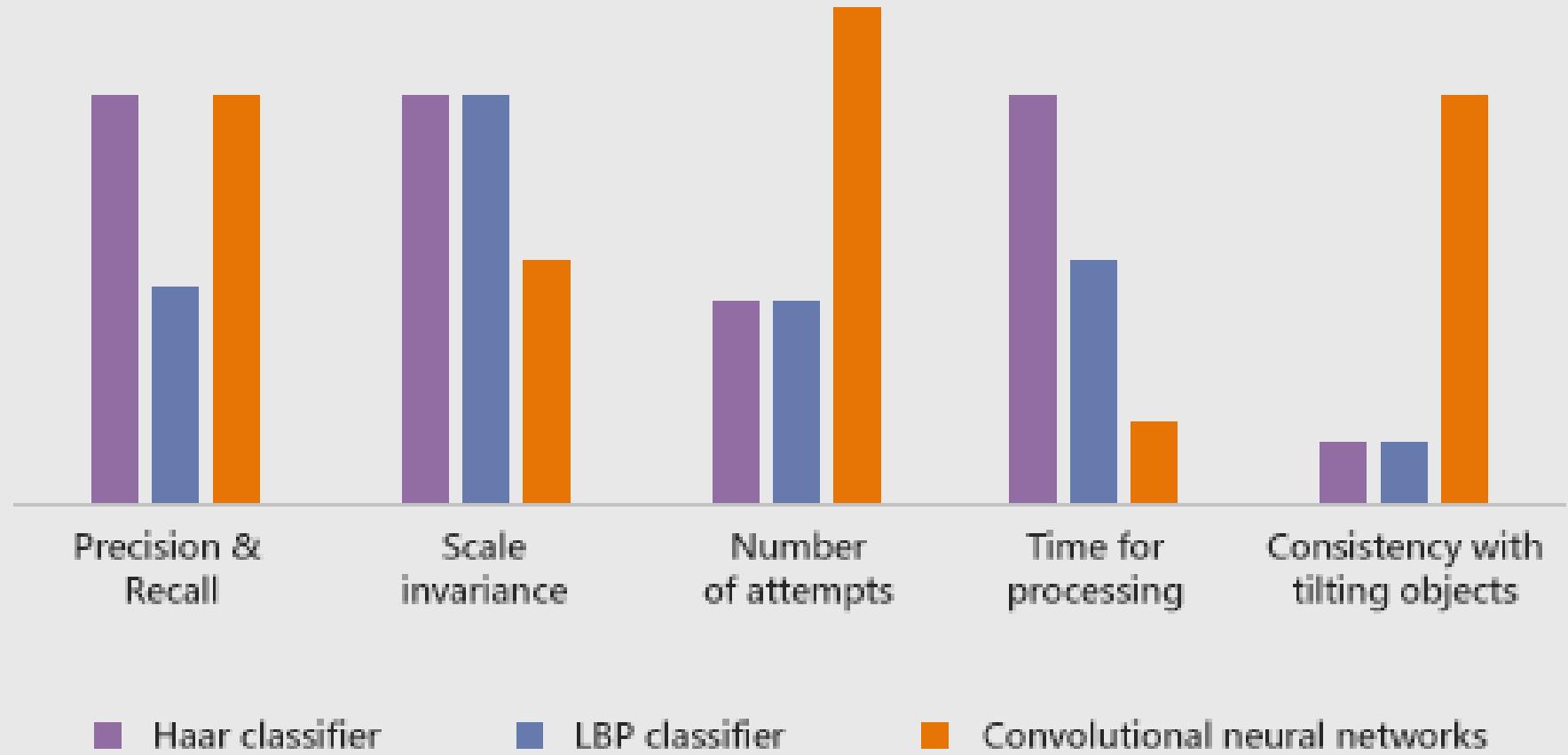
AdaBoost on basic edge/line features



Haar Cascades vs. CNN

各有優劣

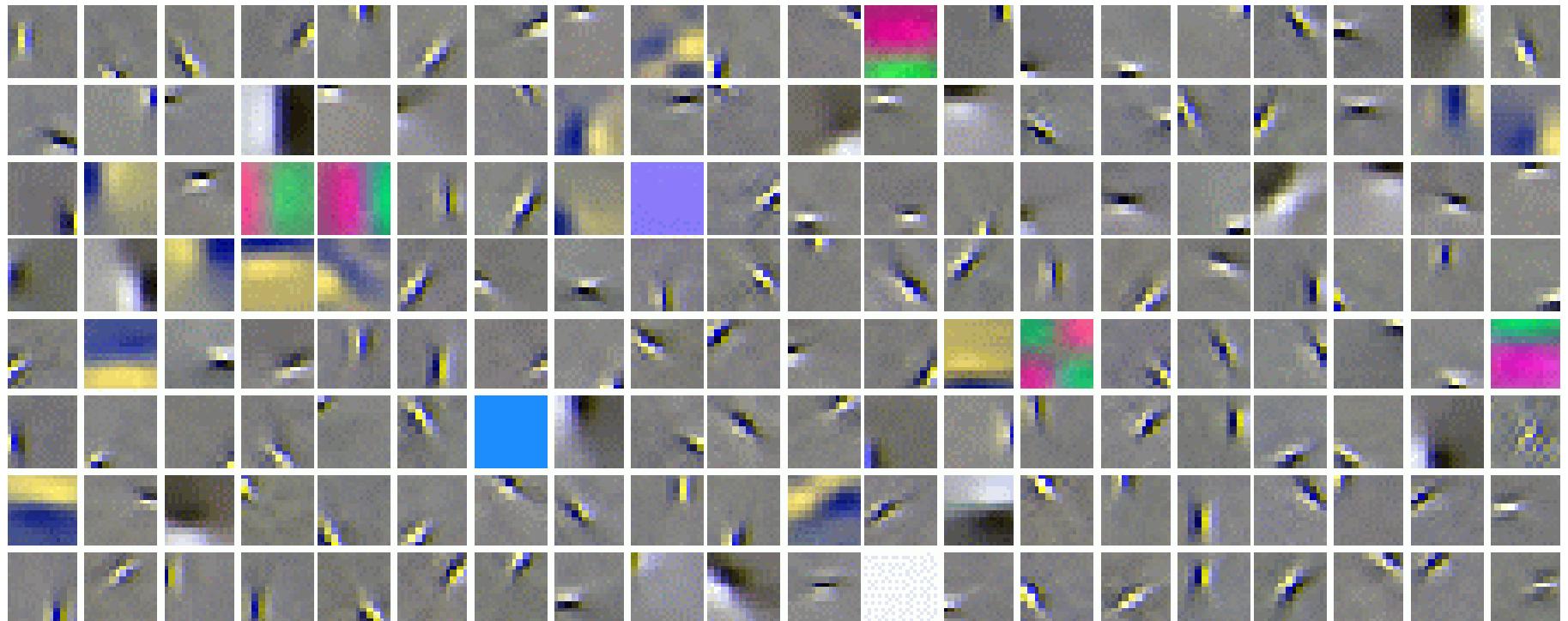
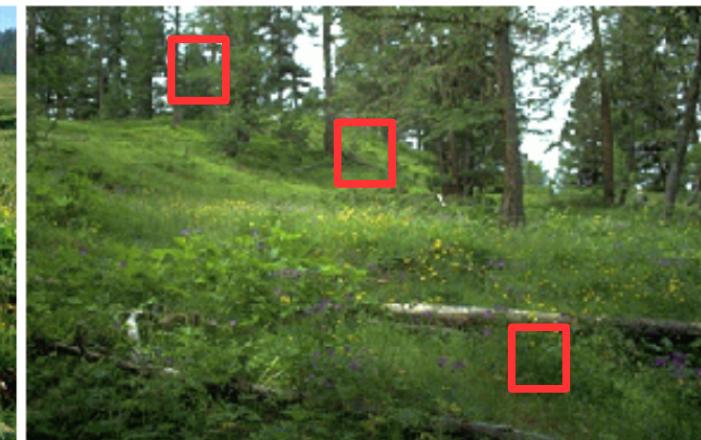
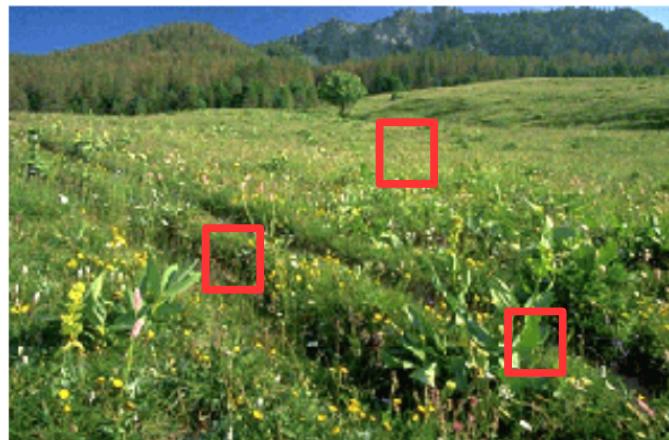
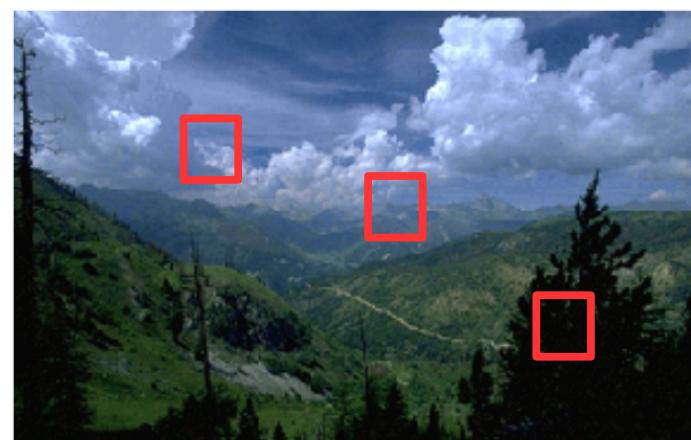
Characteristics of methods for object detection



卷積神經網路

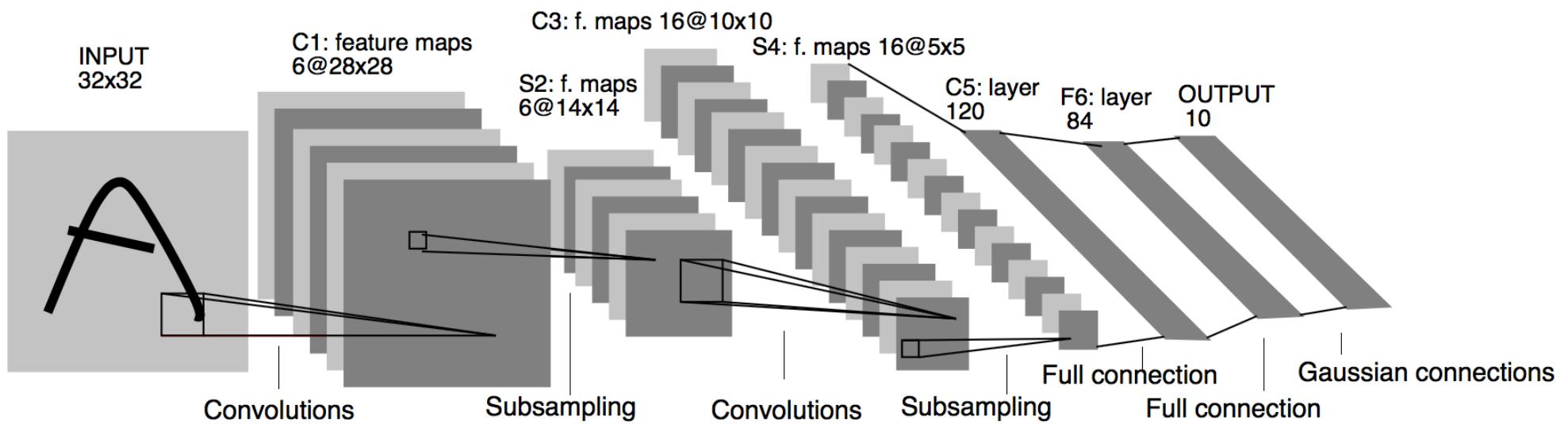
(Convolutional Neural Networks)

Patterns shared by Patches



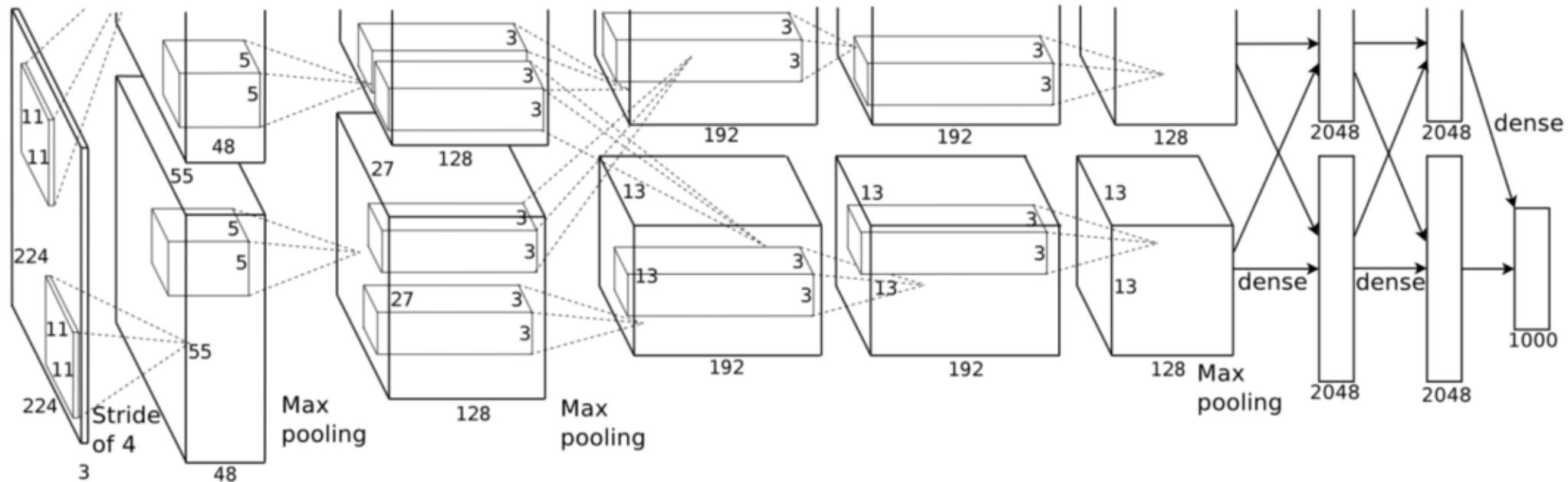
LeNet (1998)

有 convolution 有 pooling

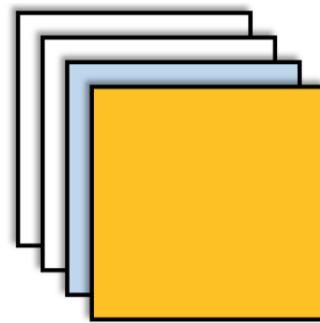


AlexNet (2012)

引入 Relu, LRN, Data Augmentation, Dropout
並用兩顆 Nvidia GTX580 訓練 6 天



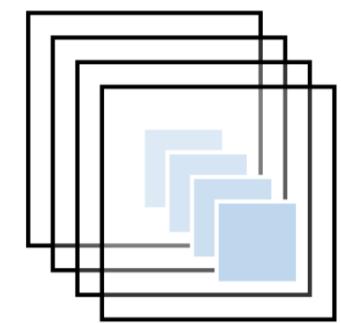
Normalization



(a) Batch-Norm



(b) Layer-Norm

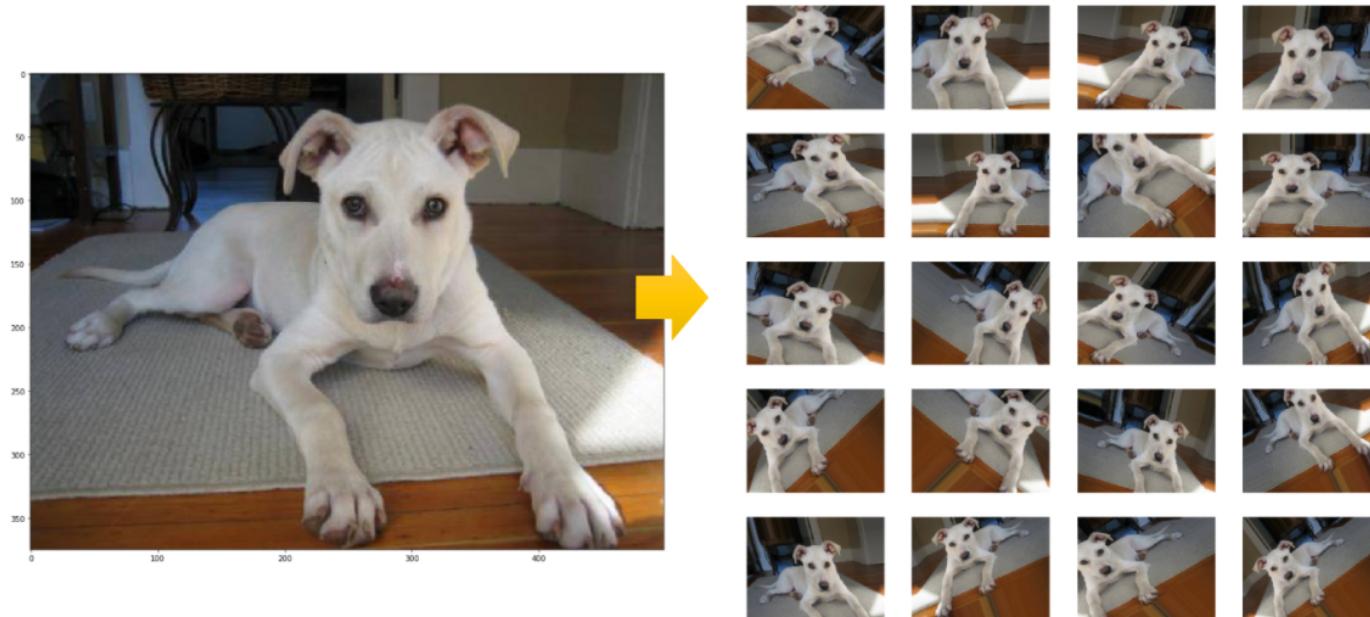


(c) Div-Norm

Model	CIFAR-10 Acc.	CIFAR-100 Acc.
Baseline	0.7565	0.4409
Baseline +WD +Dropout	0.7795	0.4179
BN	0.7807	0.4814
LN	0.7211	0.4249
BN*	0.8179	0.5156
LN*	0.8091	0.4957
DN*	0.8122	0.5066

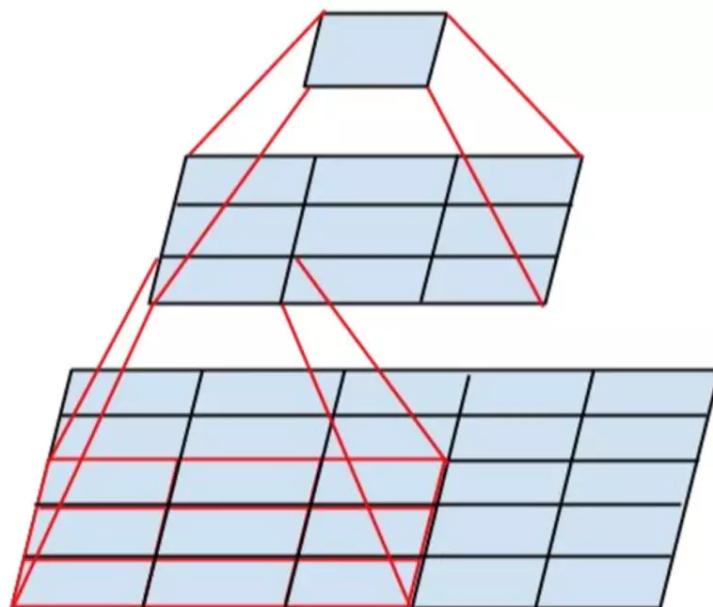
Data Augmentation

把資料集變大：仿造視覺的 tolerance (“invariance”)



VGG Net (2014)

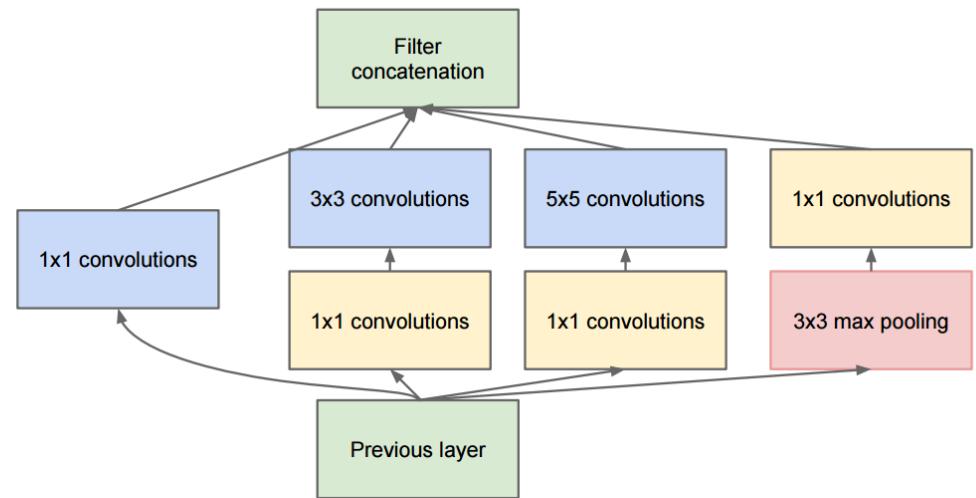
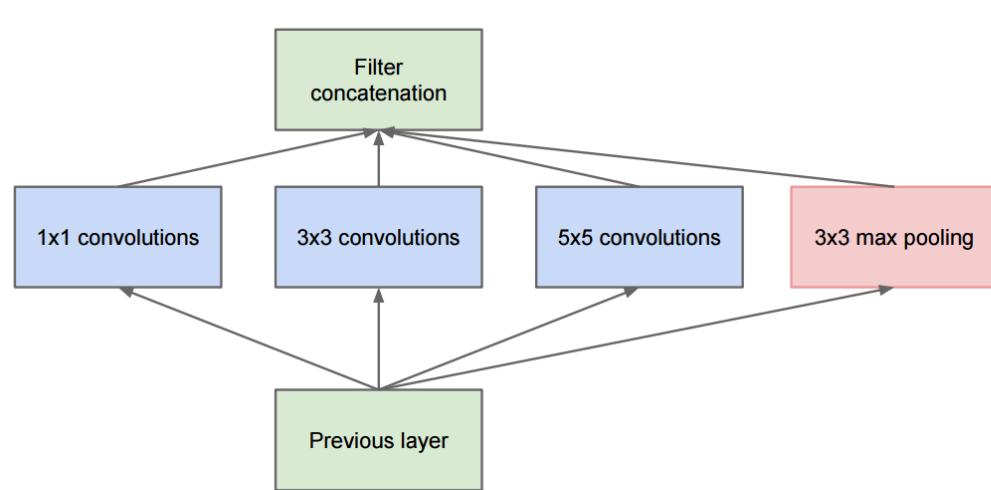
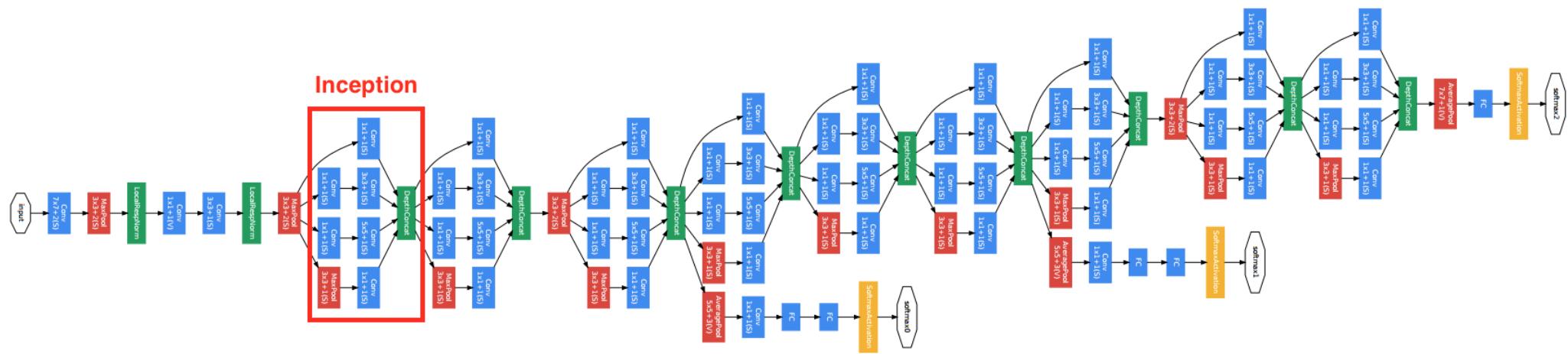
闡明下游 convolution kernels 只要 3×3



ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224×224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

GoogLeNet (2015)

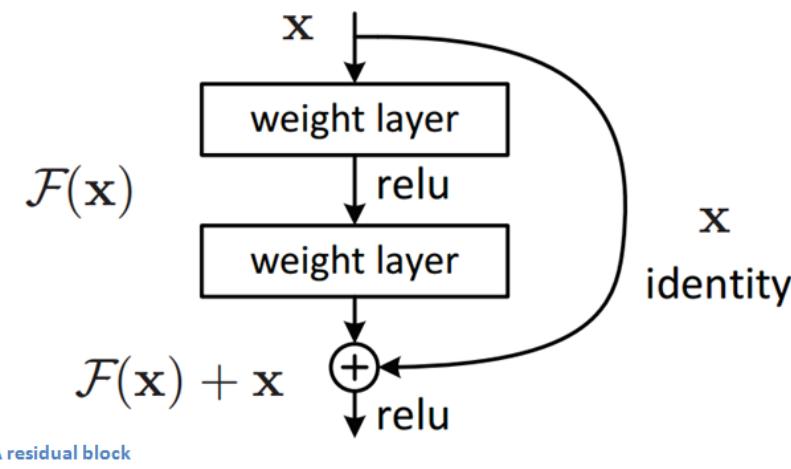
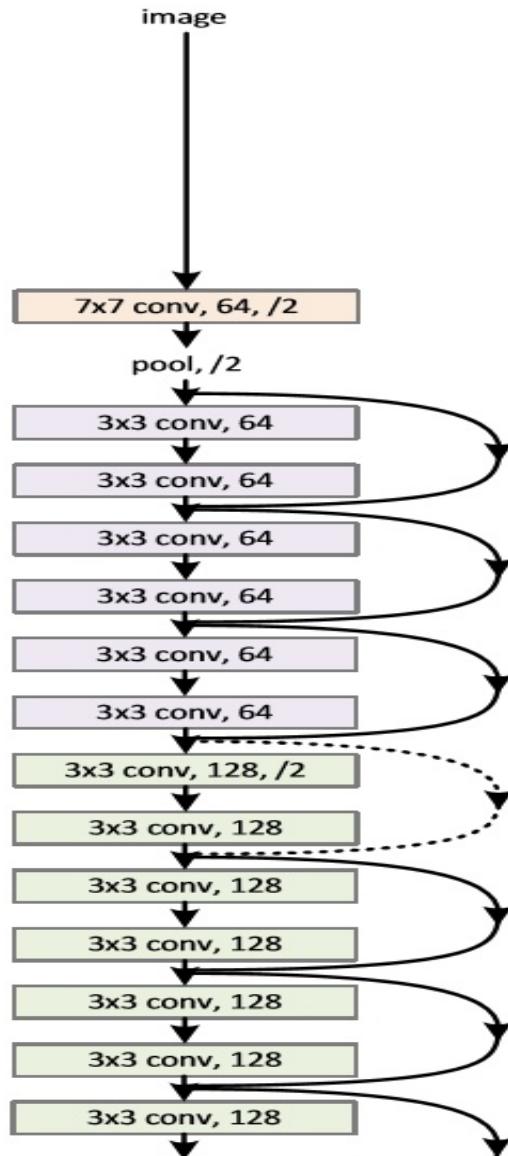
Inception module 主要仿製 multiscale vision



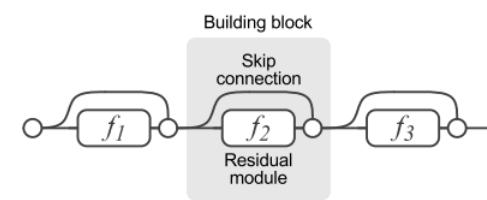
ResNet (2015)

更有效率的學習：改學自己還不足之處

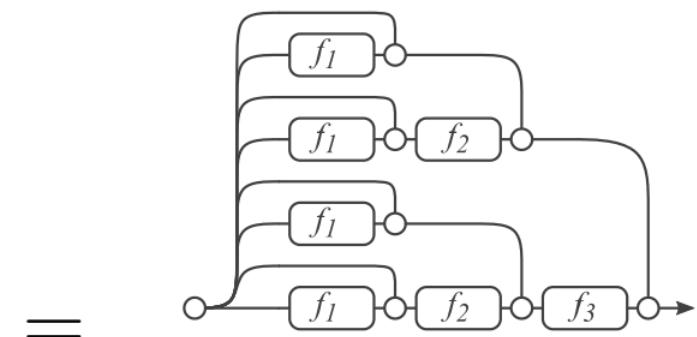
34-layer residual



A residual block



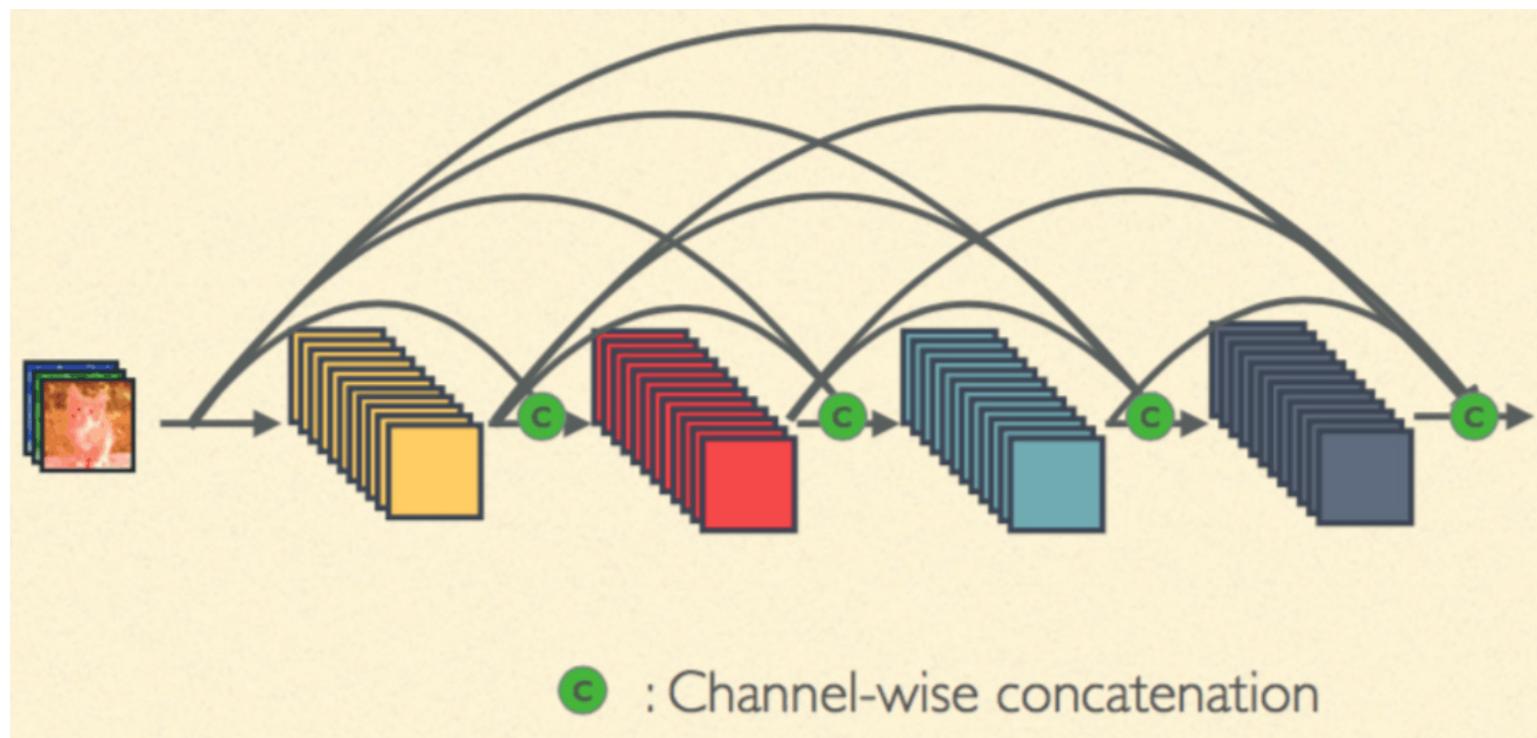
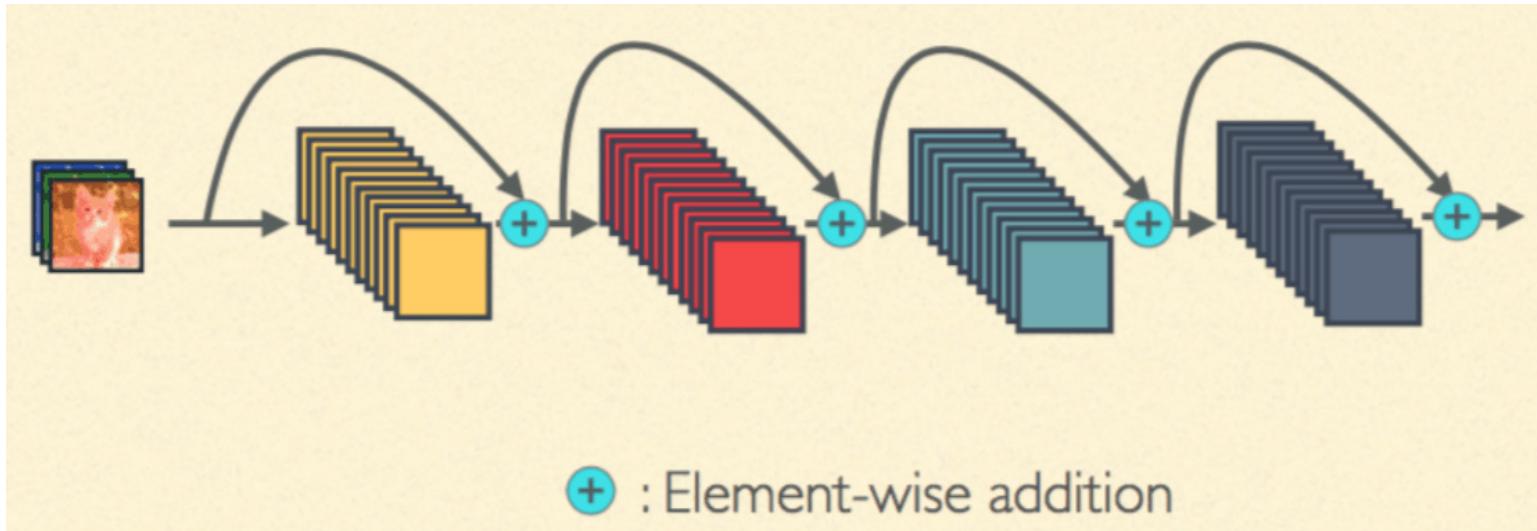
(a) Conventional 3-block residual network



(b) Unraveled view of (a)

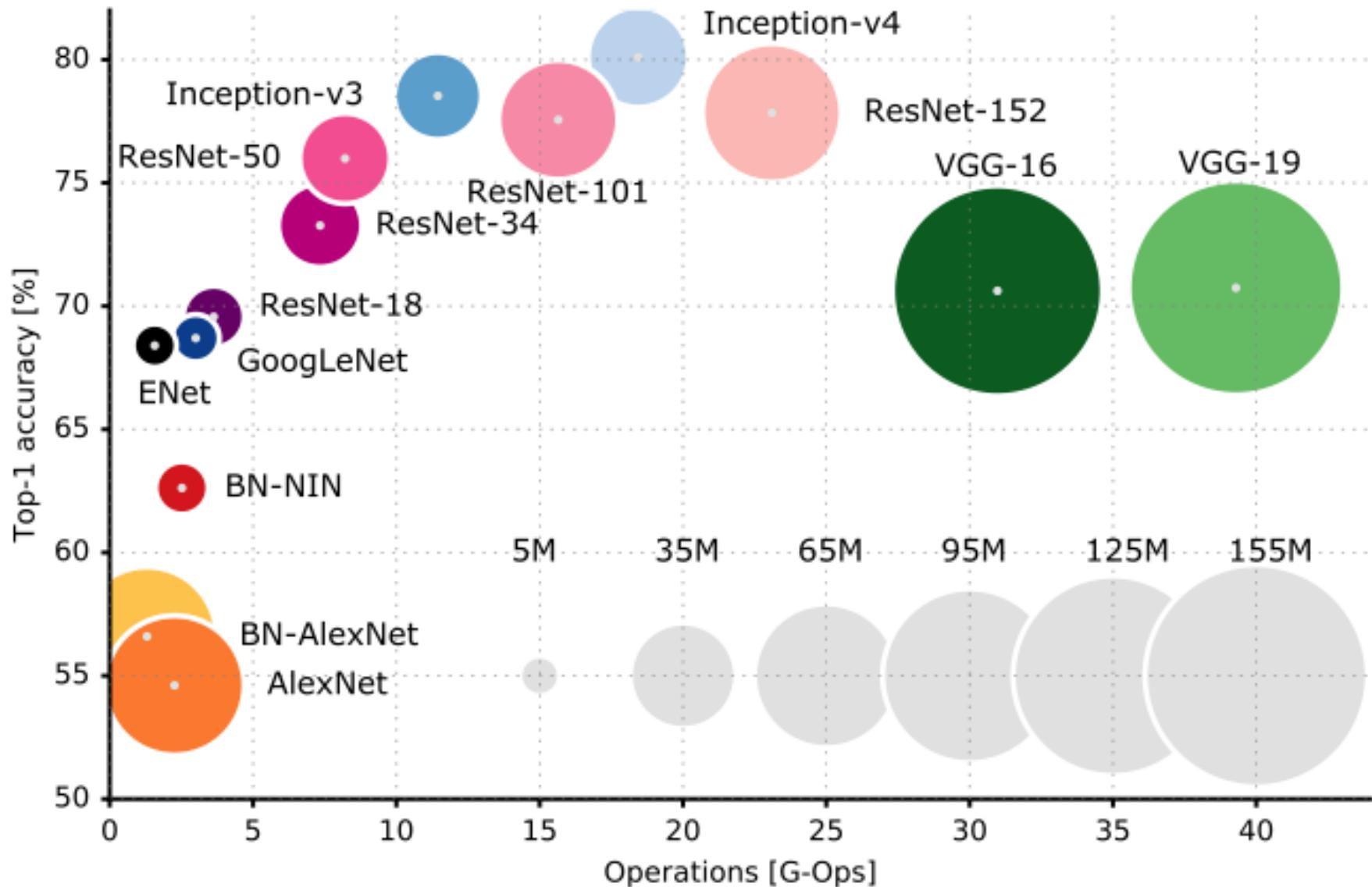
DenseNet (2017)

比 ResNet 加入更多跳接



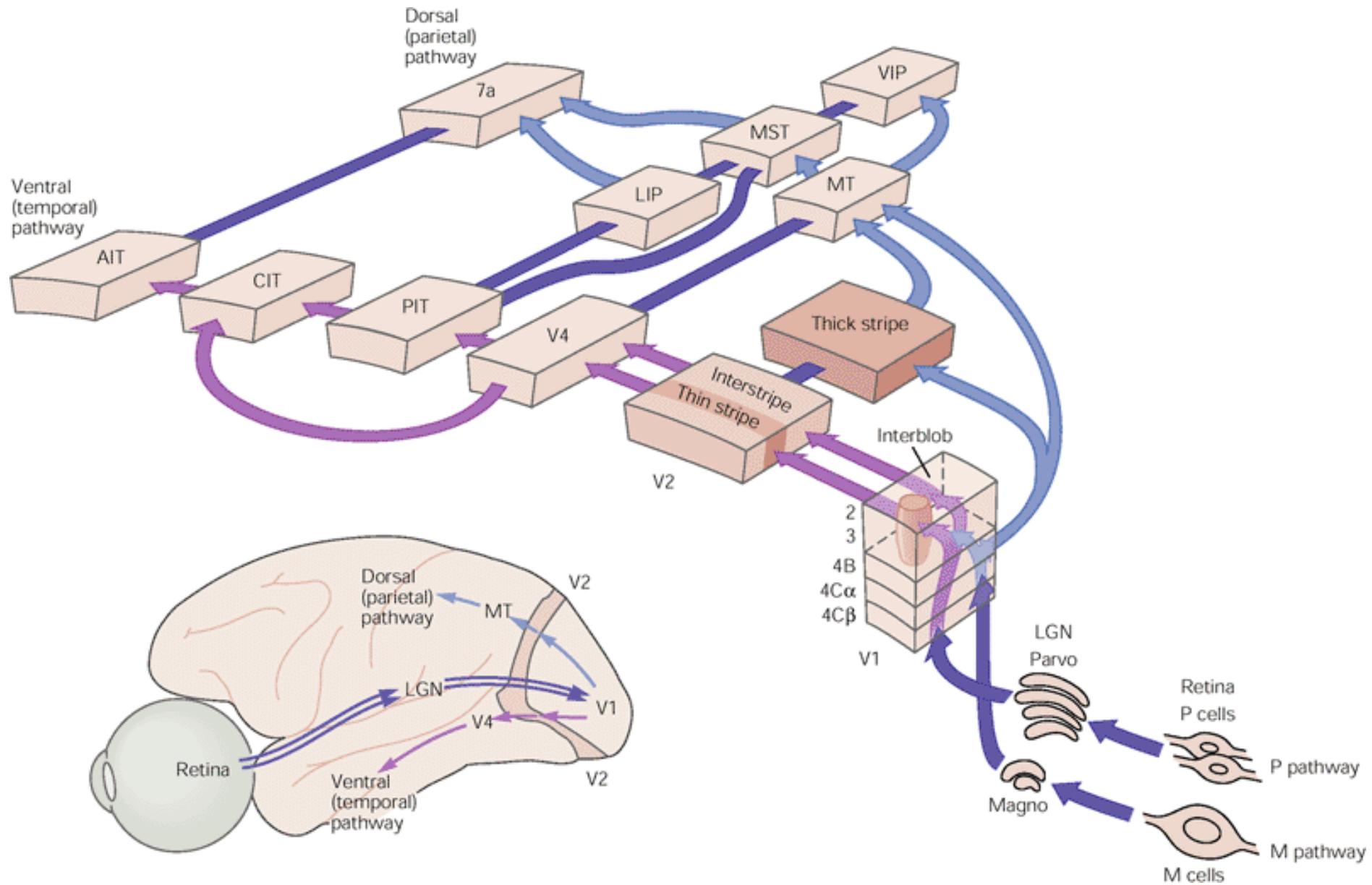
CNN 模型比較

有些雖強大但計算太重



CNN 只模仿了 What Pathway

但還有 Where Pathway

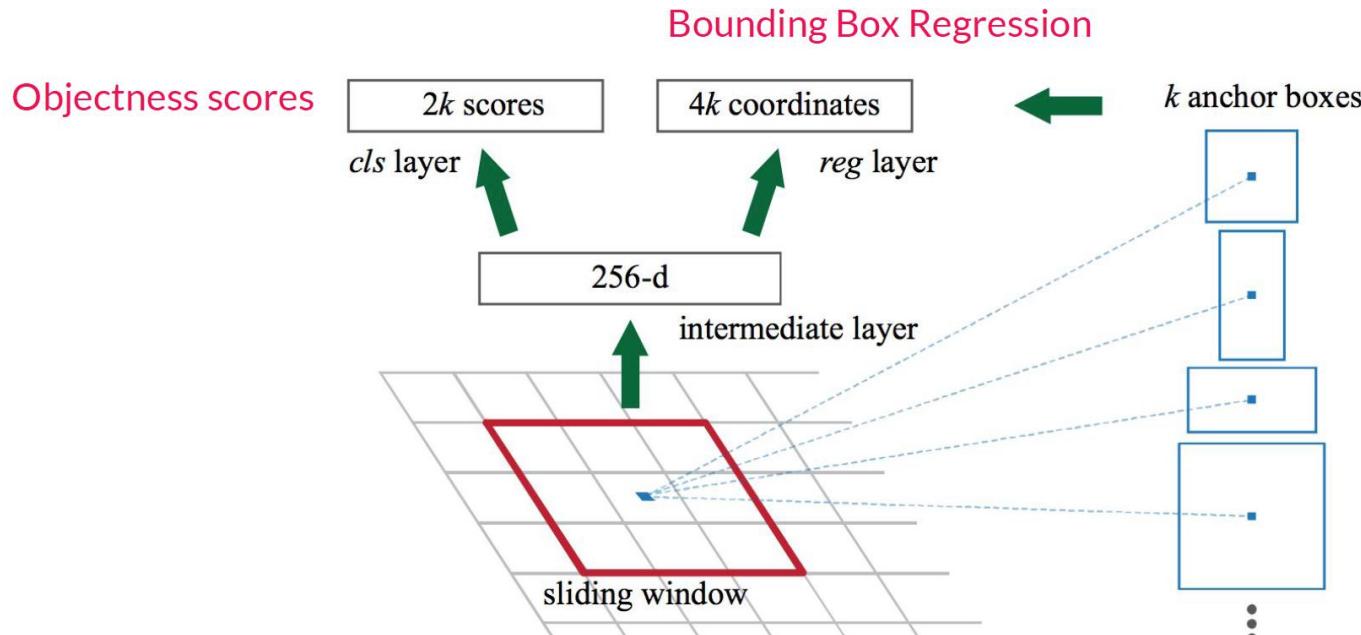


只模仿 What Pathway 的不足

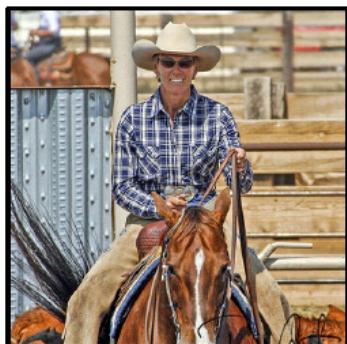
無法分解問題來各個擊破



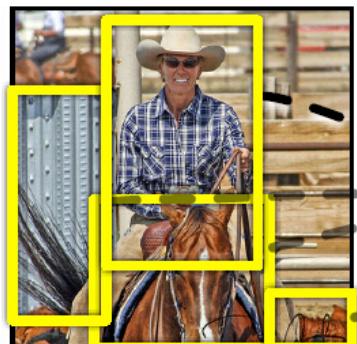
Region-based CNN (R-CNN)



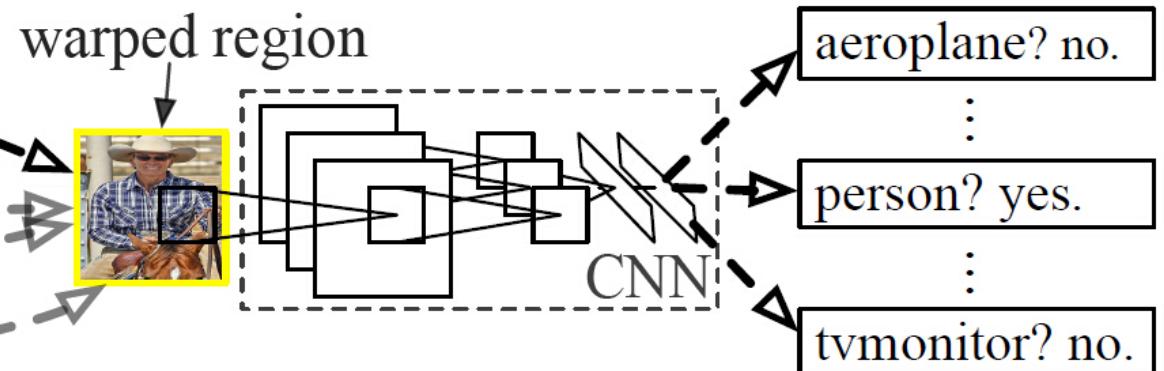
R-CNN: *Regions with CNN features*



1. Input image



2. Extract region proposals ($\sim 2k$)

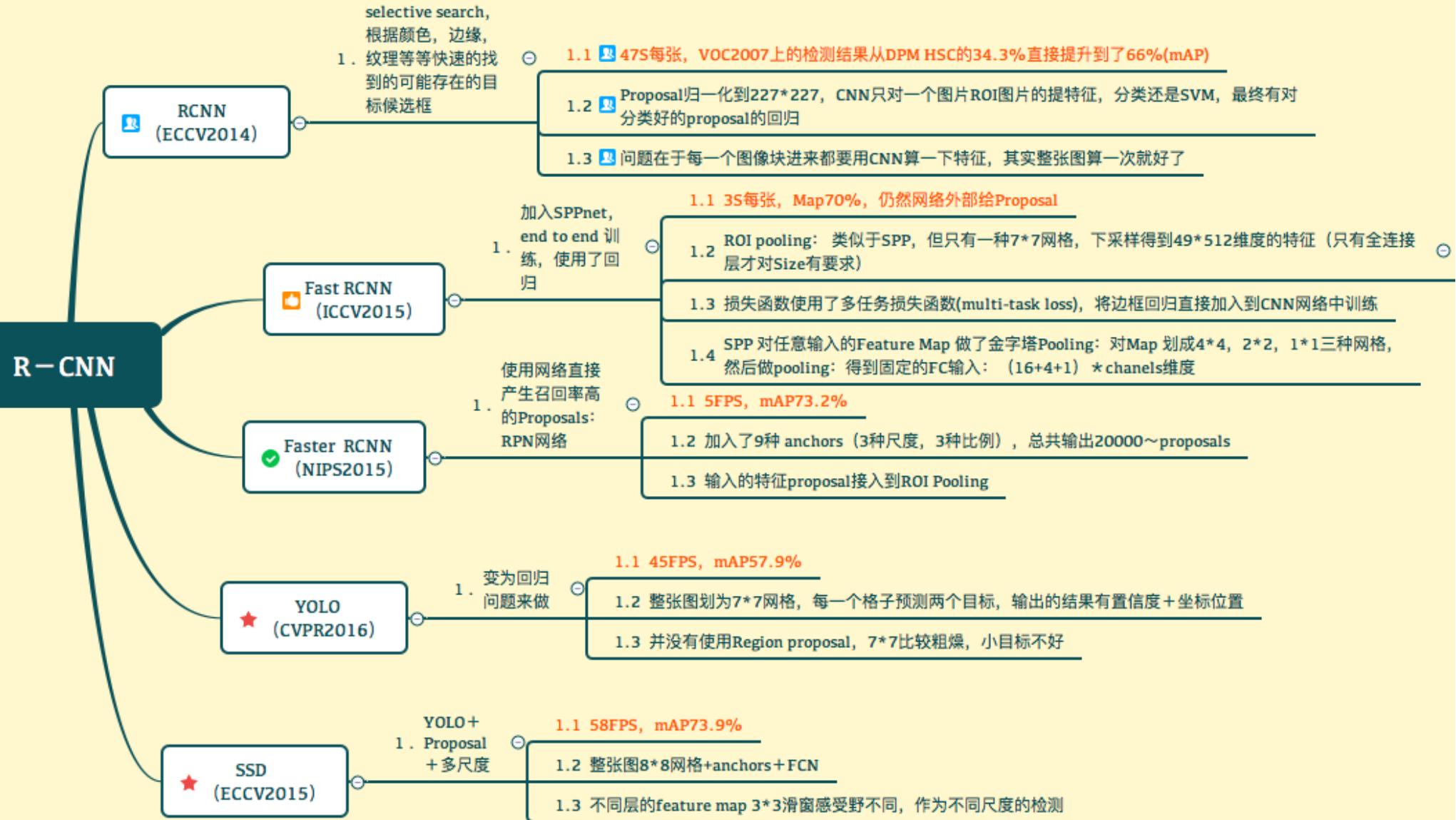


3. Compute CNN features

4. Classify regions

R-CNN 的徒子徒孫

族繁不及備載



Game Over

