心理與神經資訊學
(Psychoinformatics & Neuroinformatics)

課號: Psy5261

識別碼: 227U9340

教室:彷彿在雲端

時間: 一789

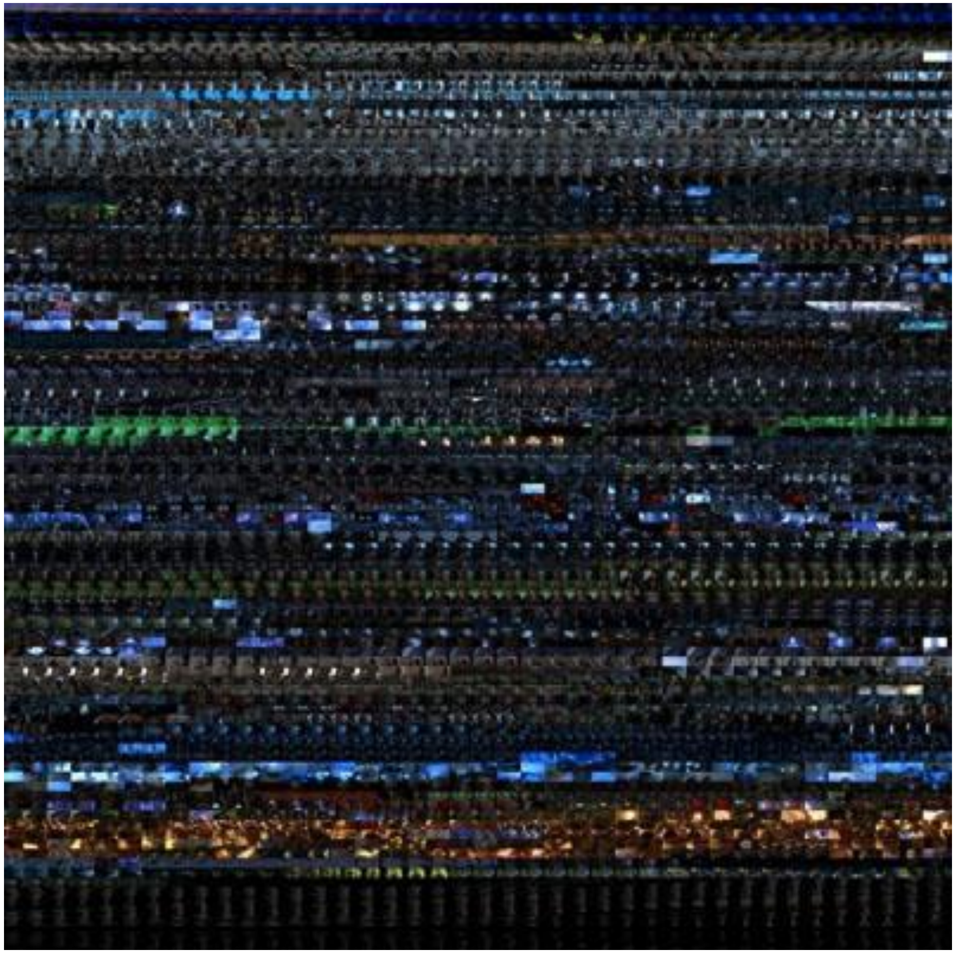
黃從仁

請pip install opencv-python, dlib

!!

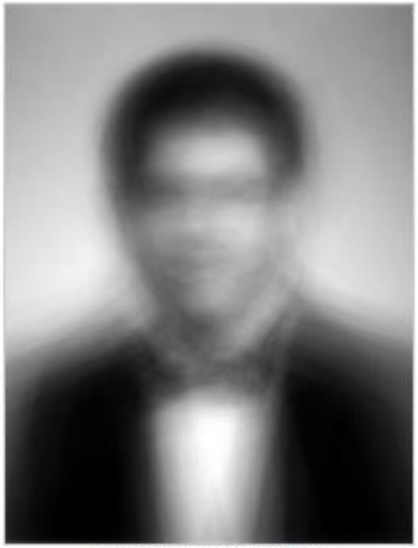
視覺化幫助洞察資料(1/2)

不同類別影片的主色調相異

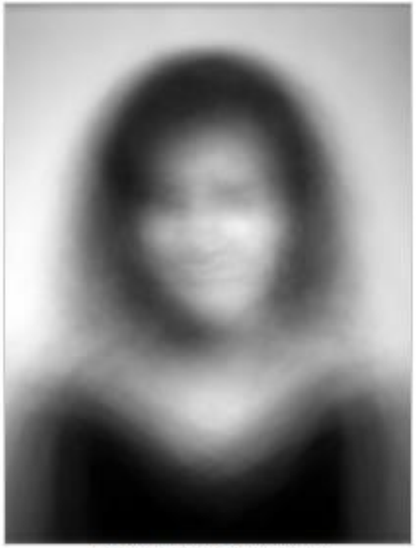


視覺化幫助洞察資料(2/2)

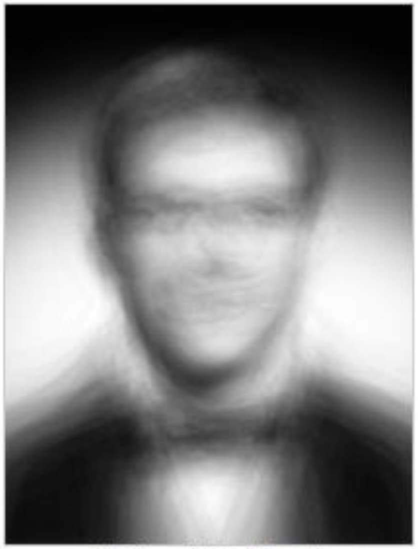
不同年代的男女外貌有何不同?



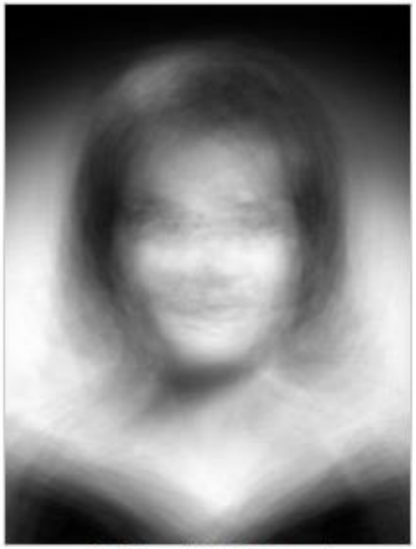
The Class of 1988 (left panel)



The Class of 1988 (right panel)



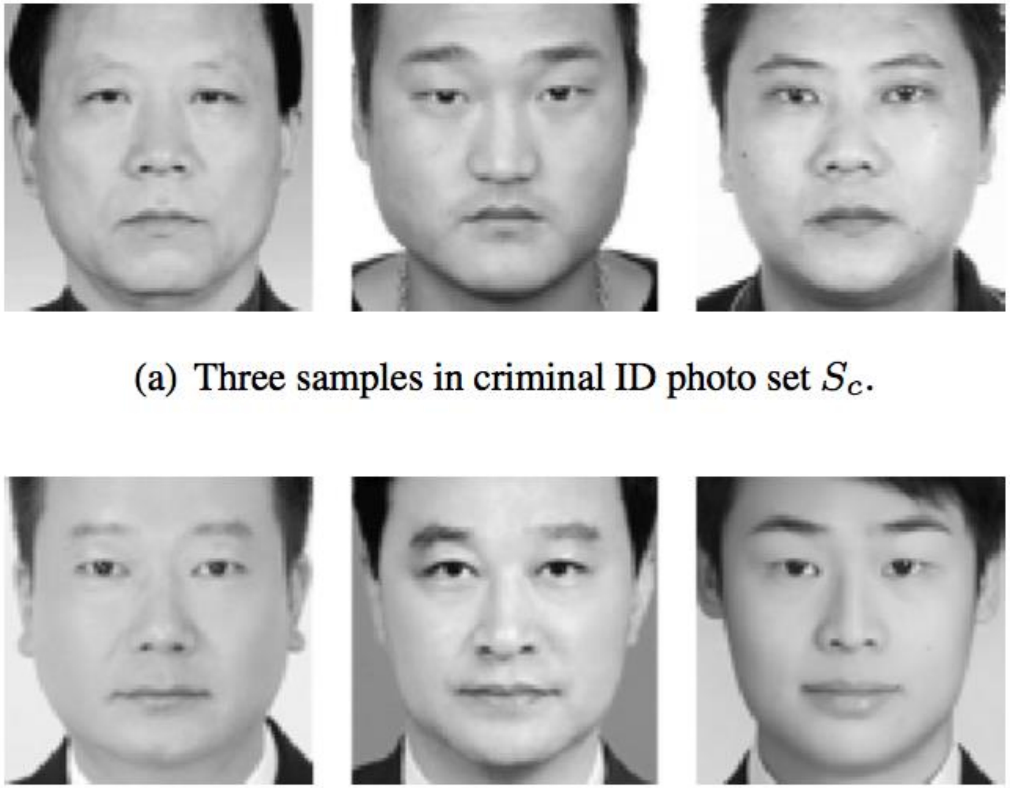
The Class of 1967 (left panel)



The Class of 1967 (right panel)

CNN的應用(1/2)

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.

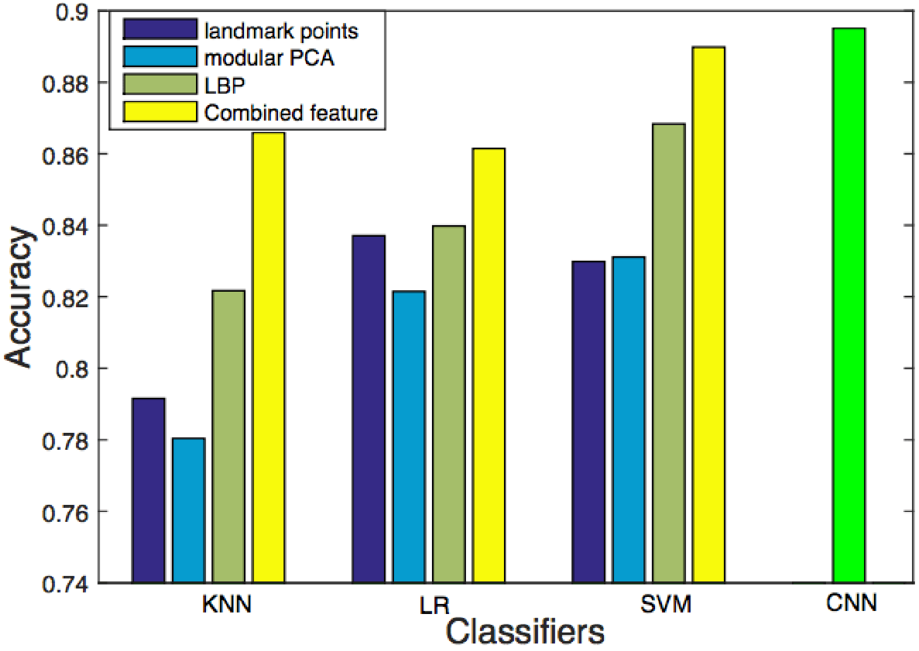


Figure 2. Accuracy of all four classifiers in all thirteen cases.

CNN的應用(2/2)

判斷誰比較美(regression)

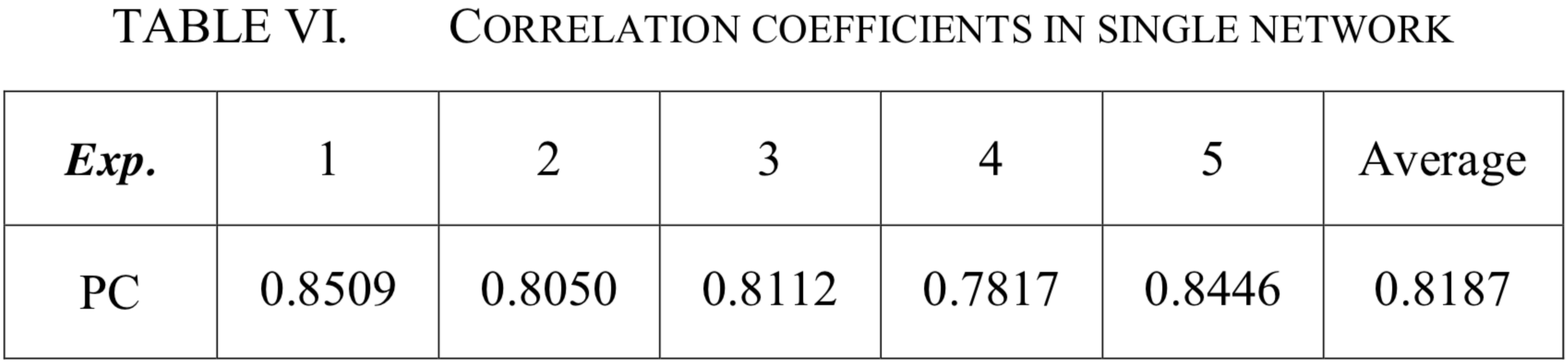
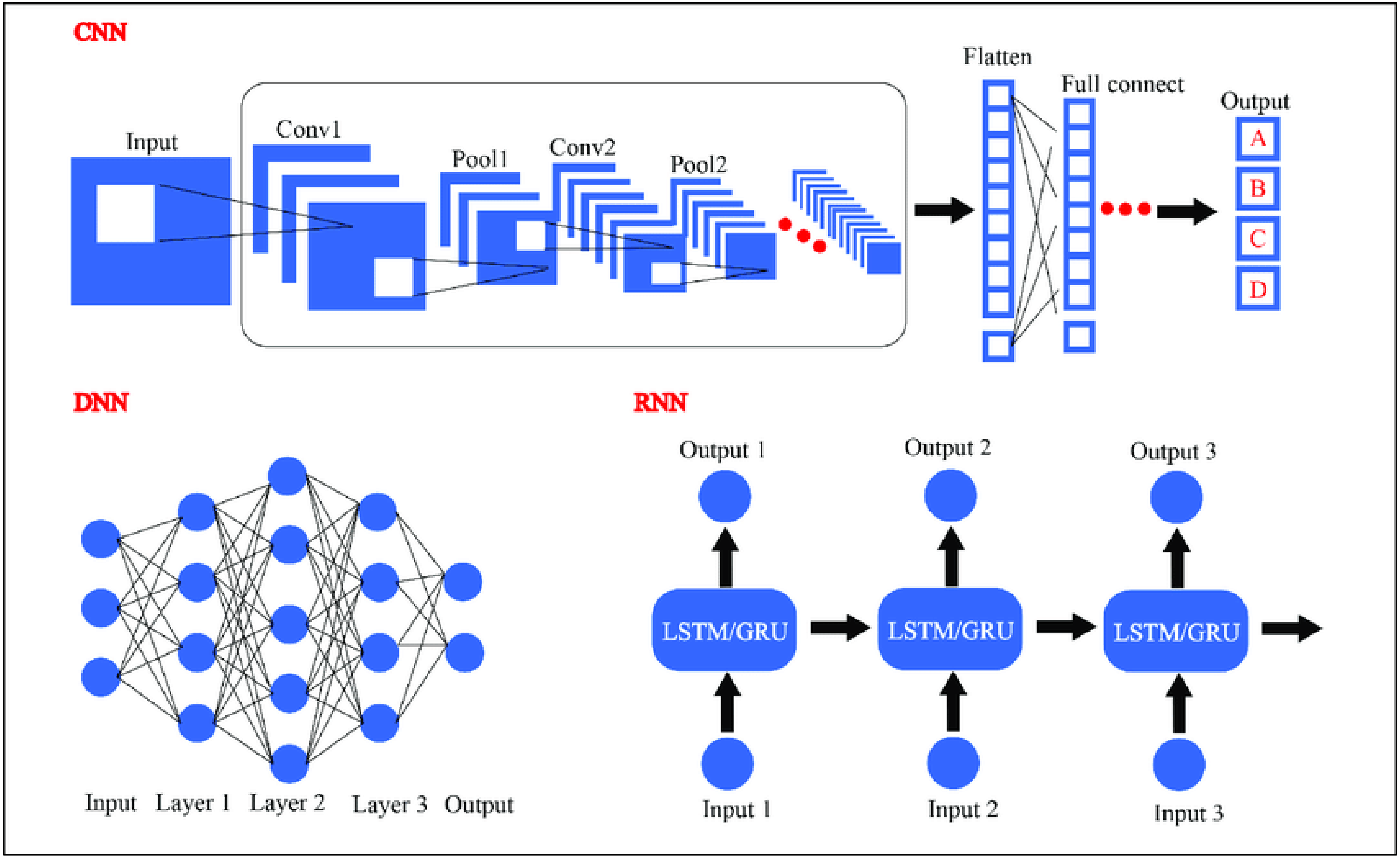


TABLE VI. CORRELATION COEFFICIENTS IN SINGLE NETWORK

Exp.	1	2	3	4	5	Average
PC	0.8509	0.8050	0.8112	0.7817	0.8446	0.8187

三種基本網路架構：今天主角是CNN

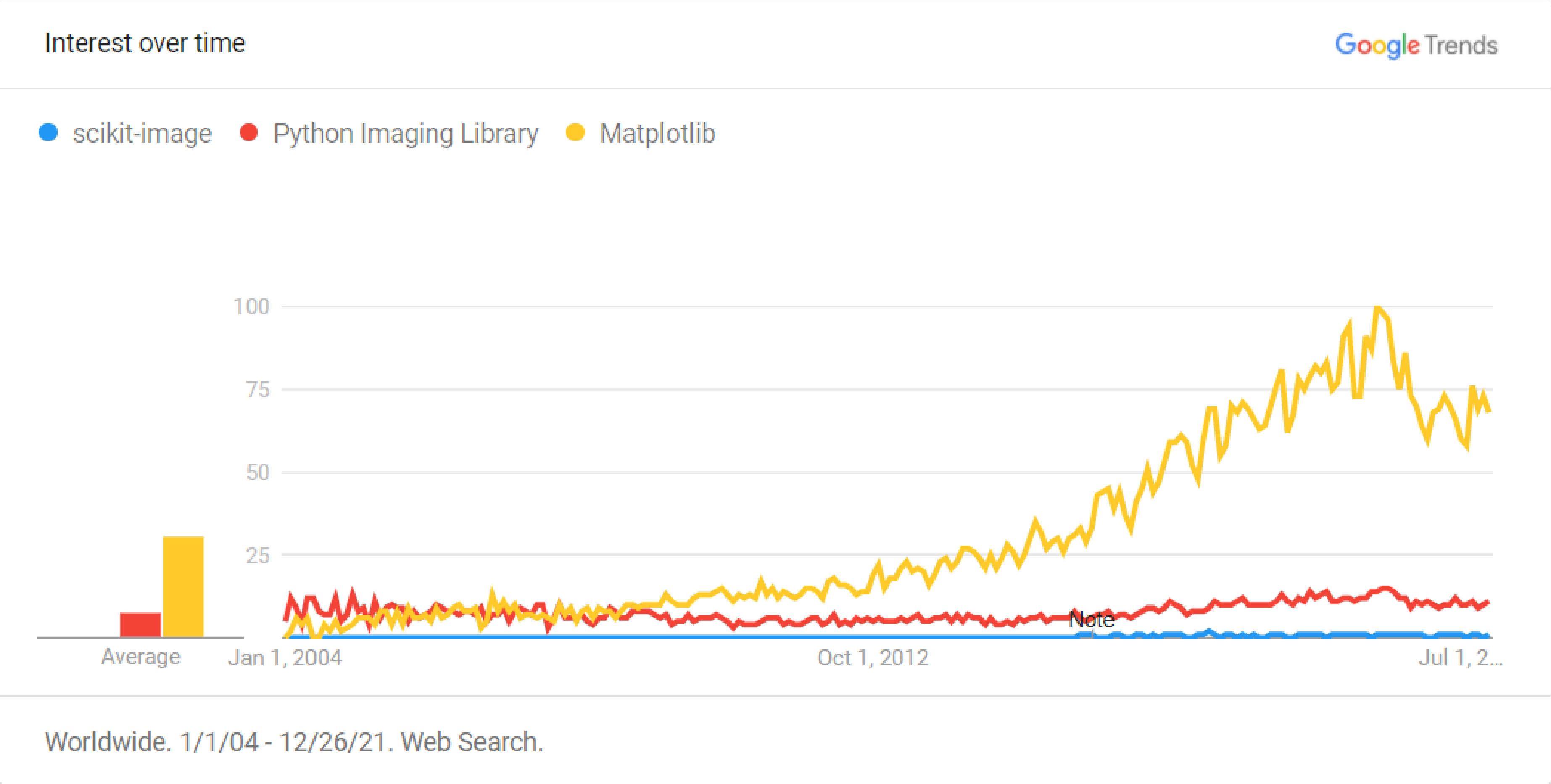
CNN通常處理影像資料;RNN通常處理語言資料



基本影像處理
(Imaging Processing)

影像處理的套件

有不同的選擇

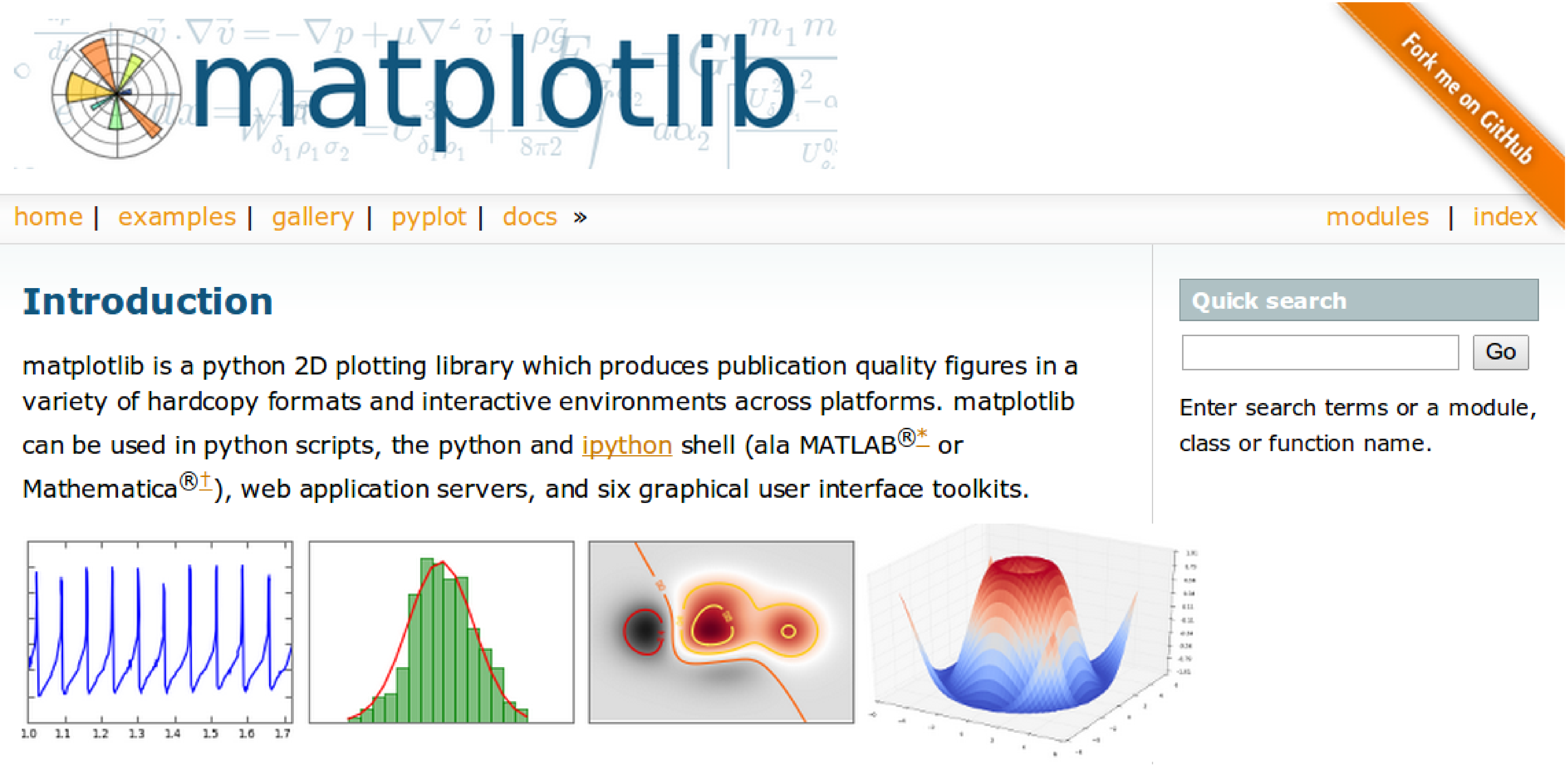


Worldwide. 1/1/04 - 12/26/21. Web Search.

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



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

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.

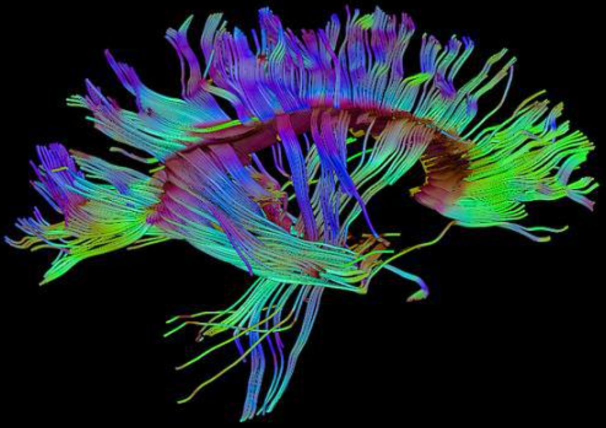
Quick search

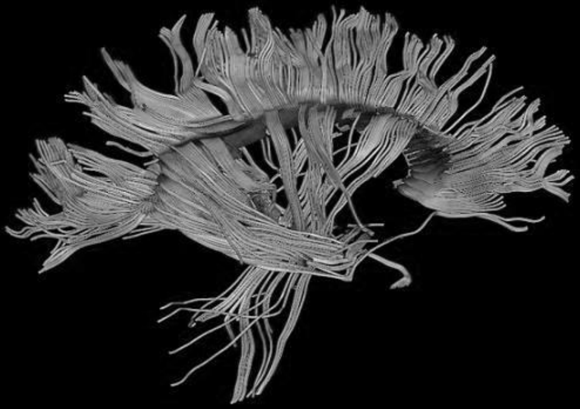
Go

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

Matplotlib的基本影像處理(1/3)

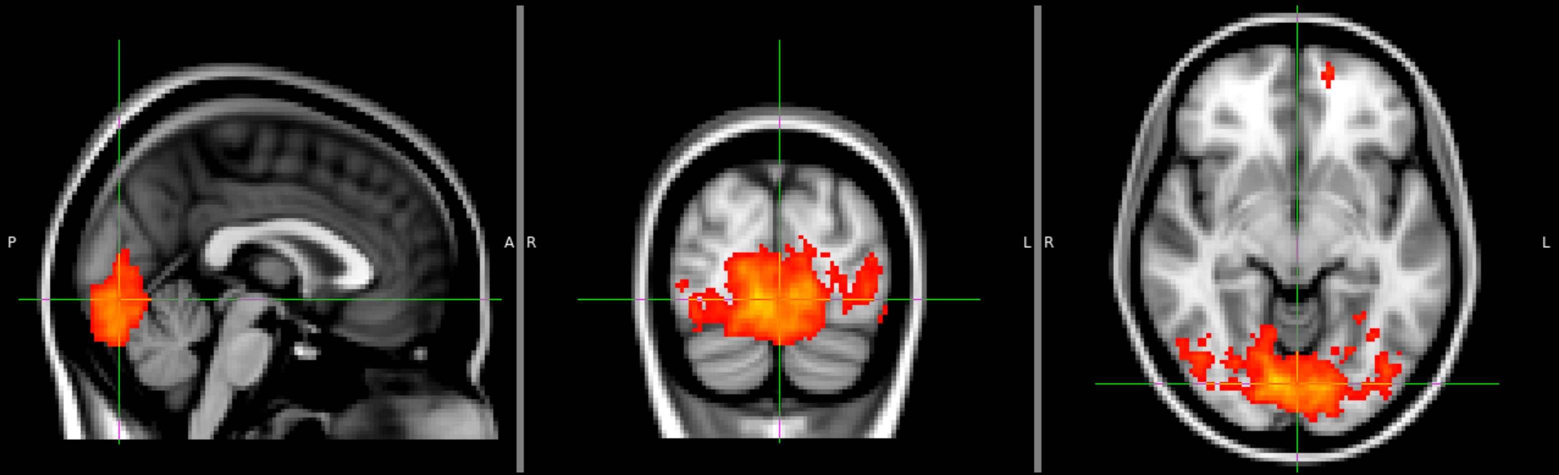
```
import matplotlib.pyplot as plt
Import numpy as np
plt.close('all'); img=plt.imread('DTI.jpg')
print(img.shape,type(img),img.dtype)
plt.imshow(img); plt.figure()
img2=np.mean(img,2); print(img2.shape)
plt.imshow(img2,cmap=plt.cm.gray)
```





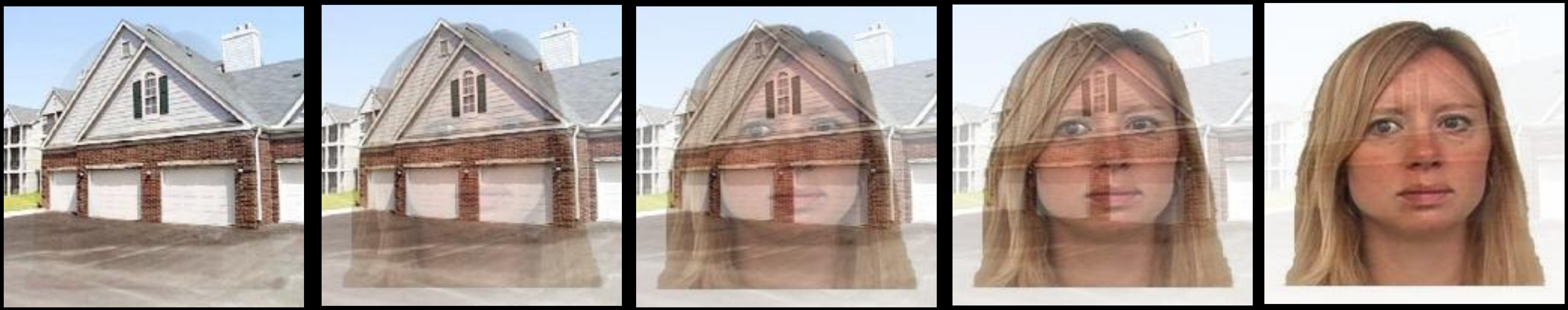
Matplotlib的基本影像處理(2/3)

```
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)
```



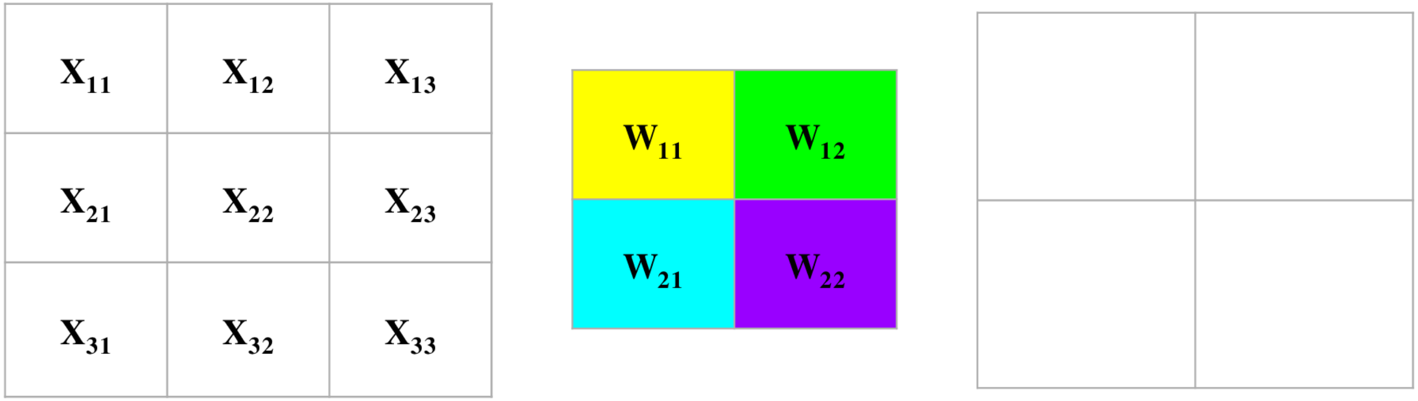
Matplotlib的基本影像處理(3/3)

```
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)
```



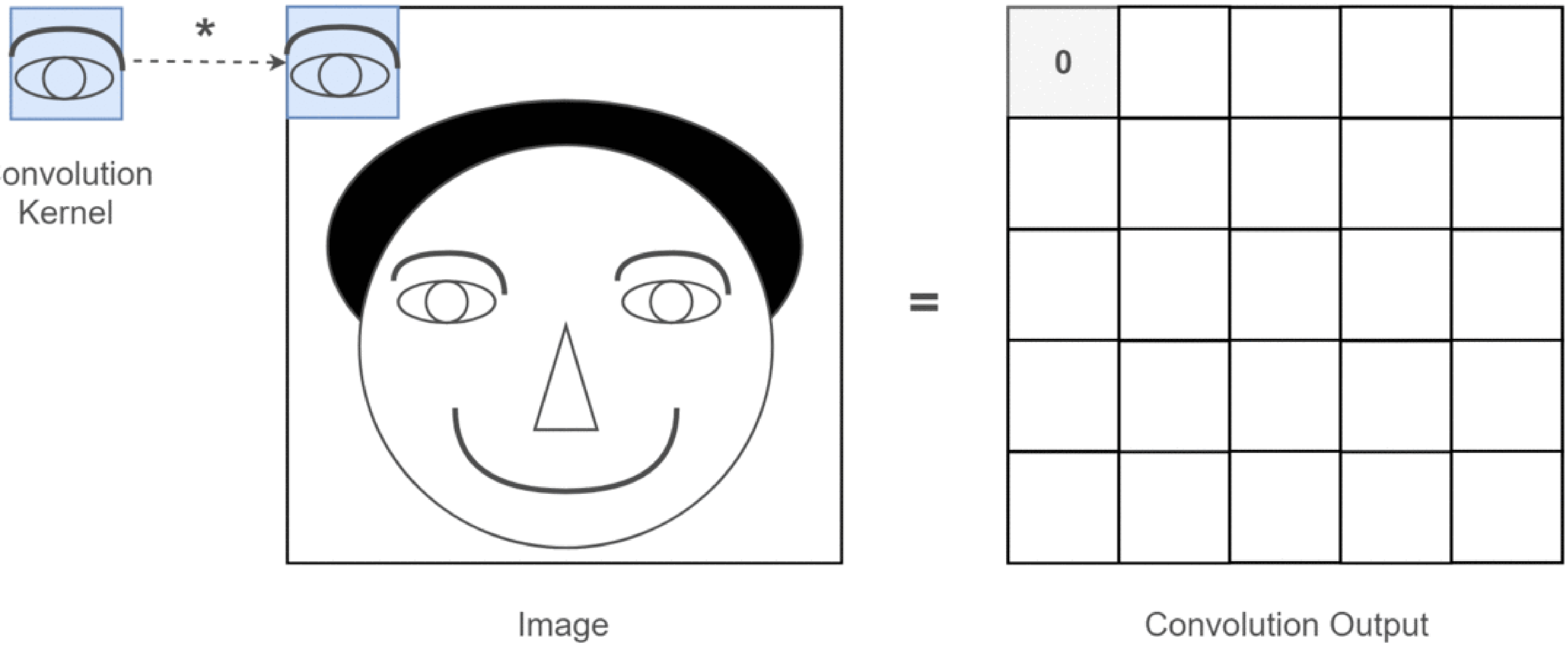
卷積 = 在空間滑動偵測

因為相同的特徵有可能在不同的位置出現



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

W_{11}	W_{12}
W_{21}	W_{22}



Padding & Stride

Input與Output的維度關係一定要會算

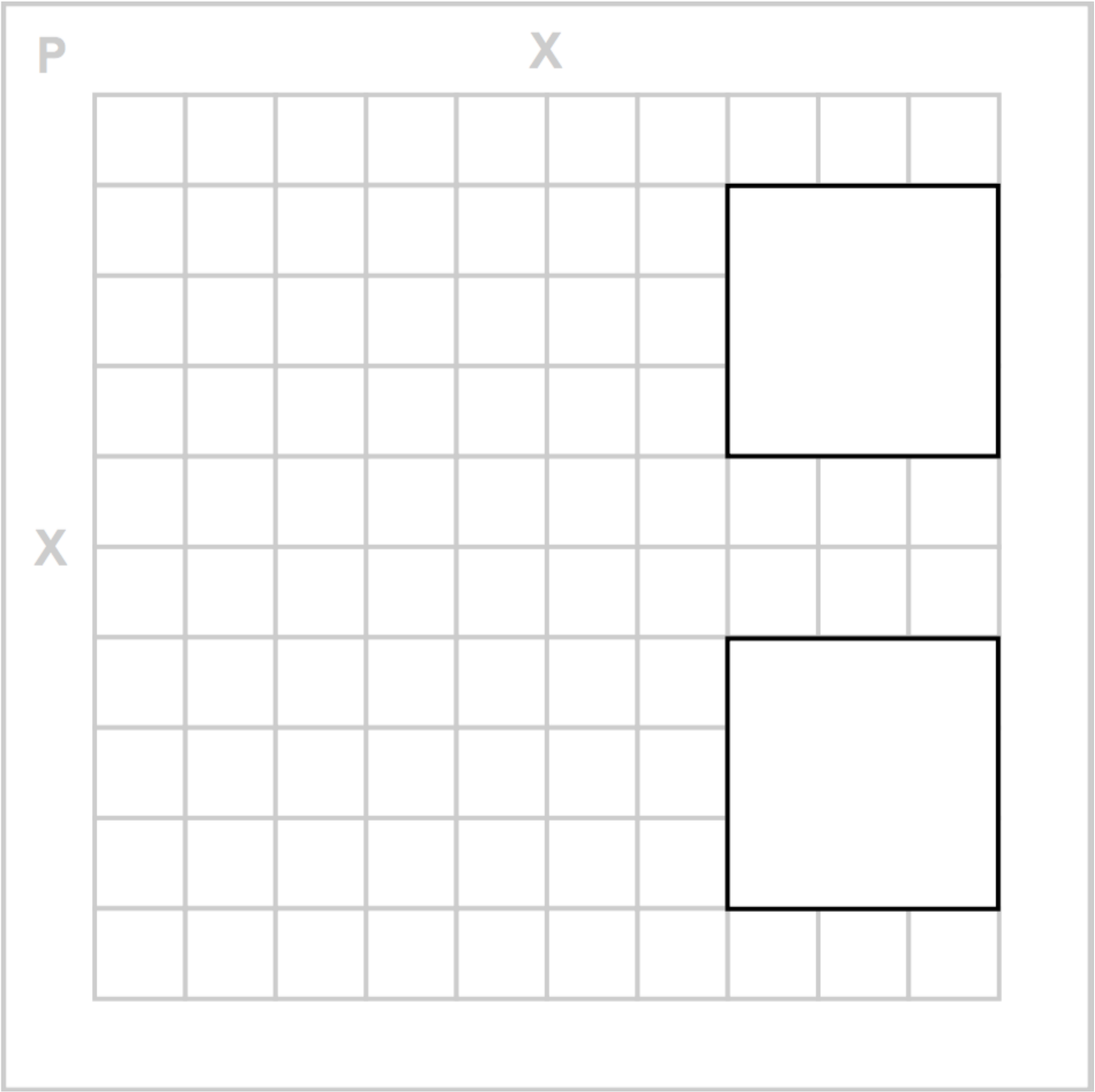
$$1 + \frac{X - F + 2P}{S}$$

X = image size

F = filter size

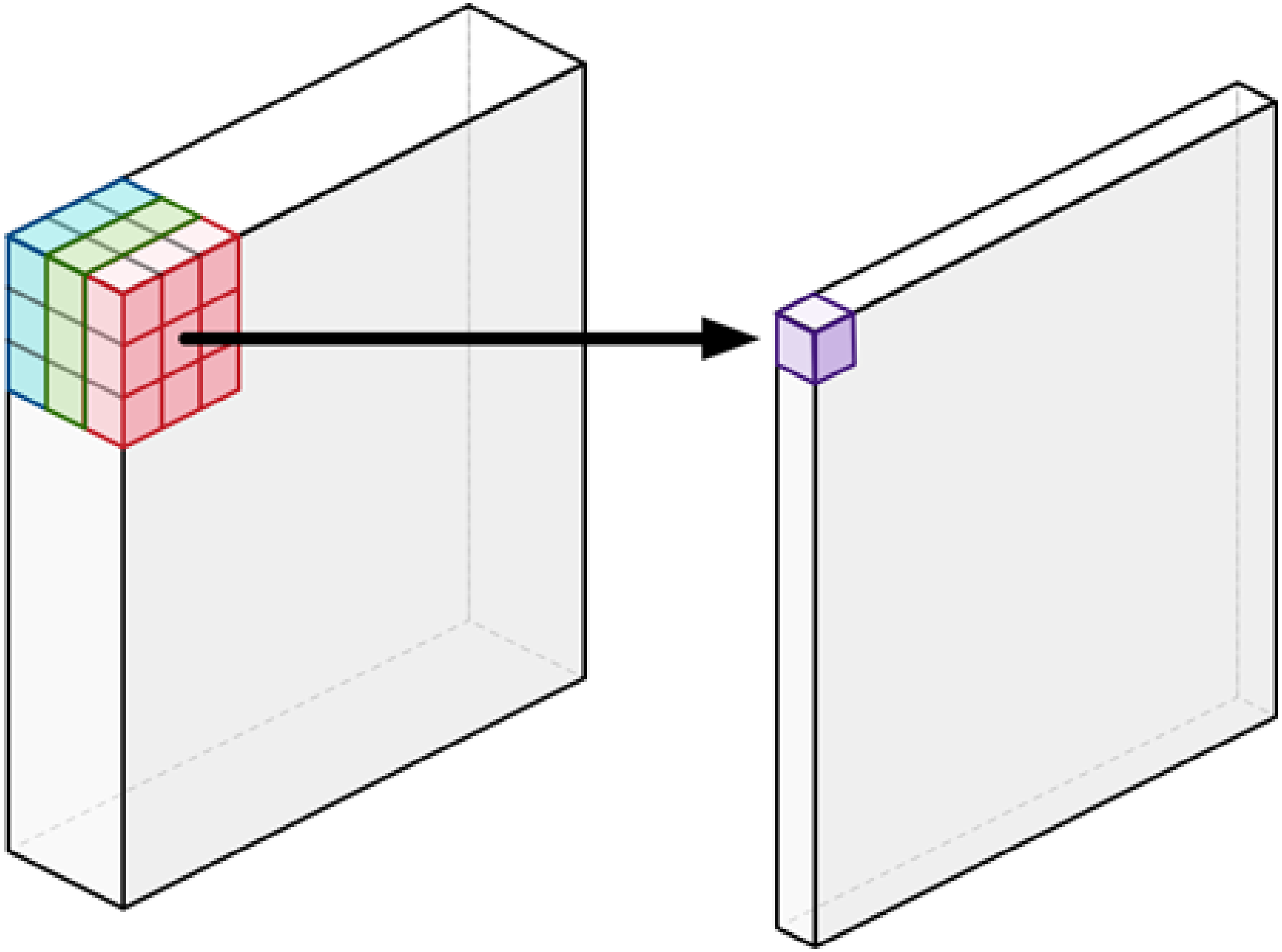
P = padding

S = stride



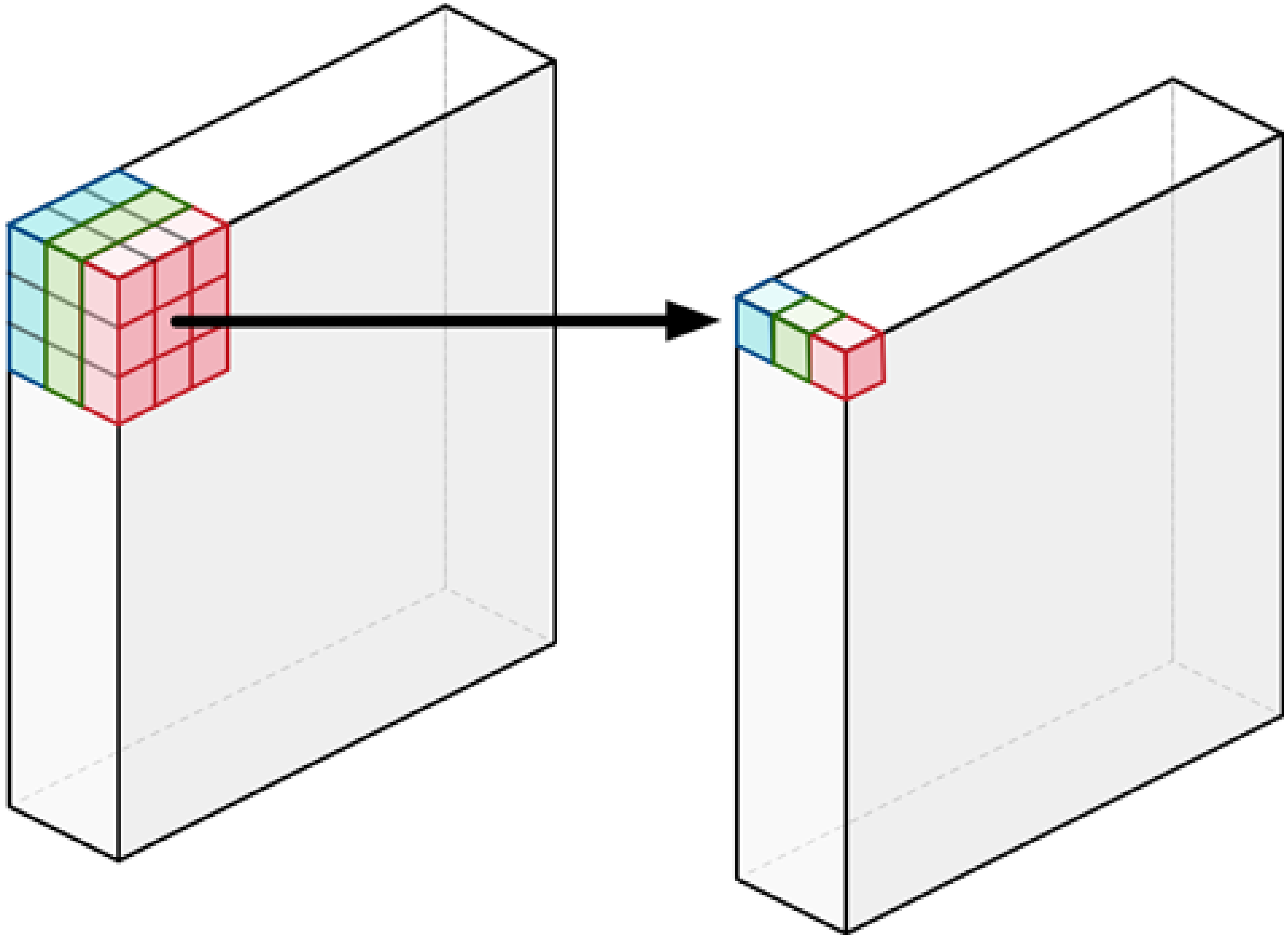
3D Spatial Convolution

和2D一樣就是template matching



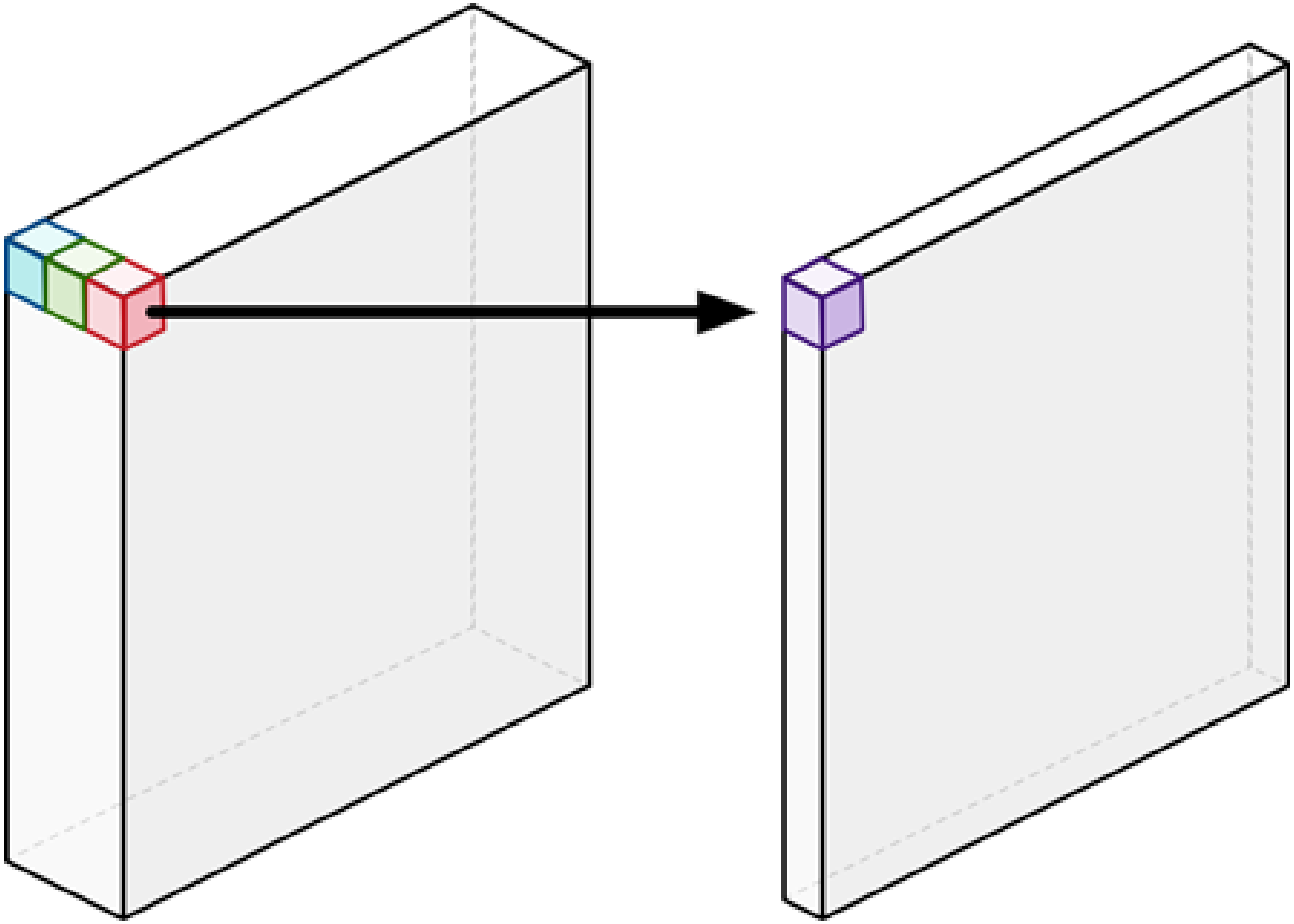
Depthwise Separable Conv.

各層的規律性分開找計算複雜度較低(如2^{27}→$3*2^{9}$)



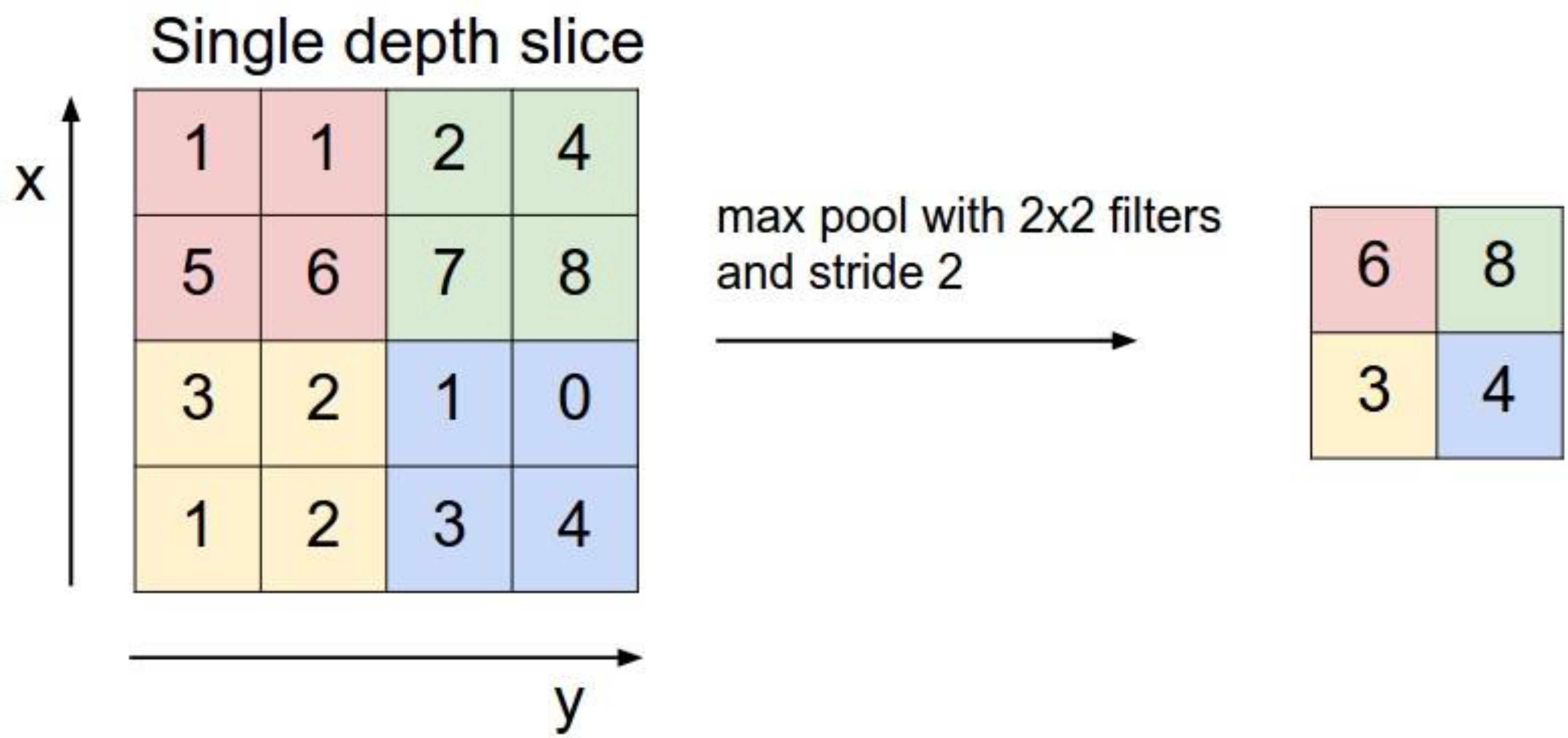
Pointwise Convolution

就是去把不同層做加權平均



(Max) Pooling

就是downsampling



Single depth slice

1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

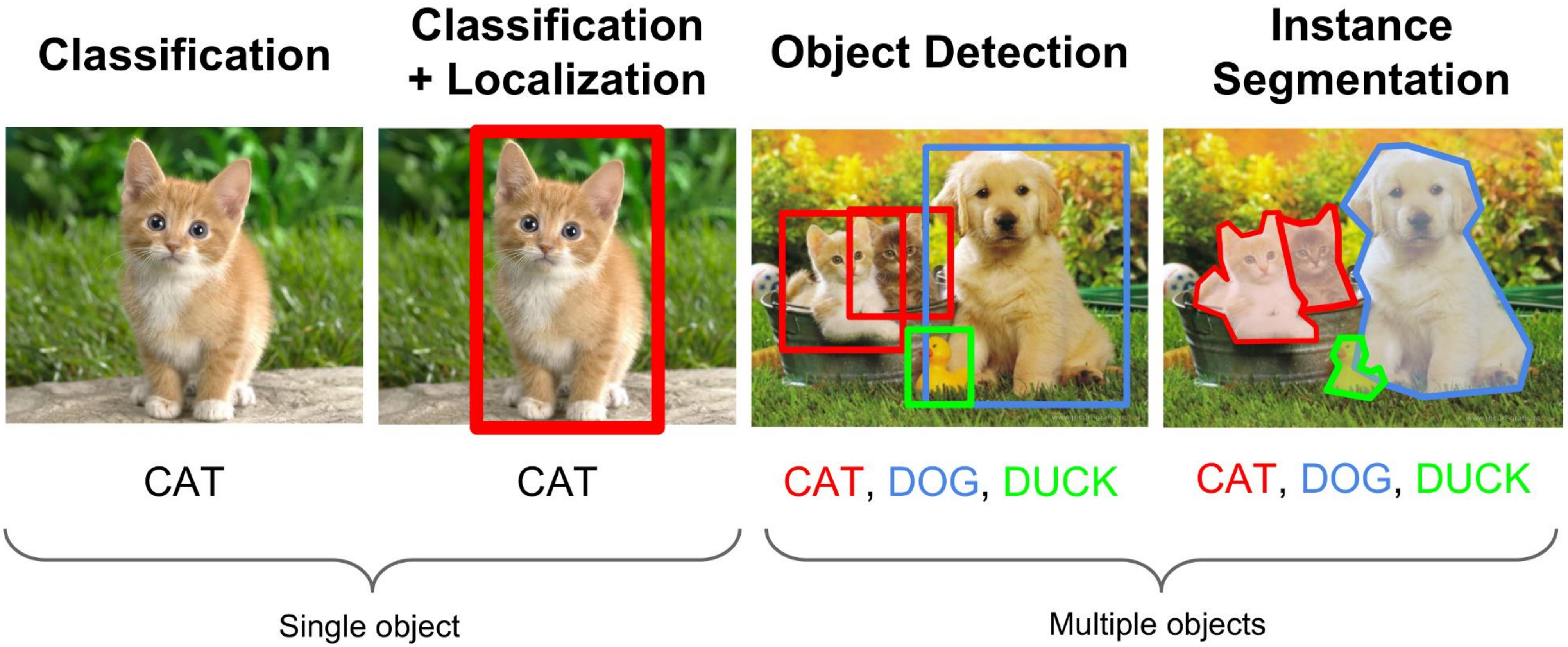
max pool with 2x2 filters and stride 2

6	8
3	4

基本電腦視覺
(Computer Vision)

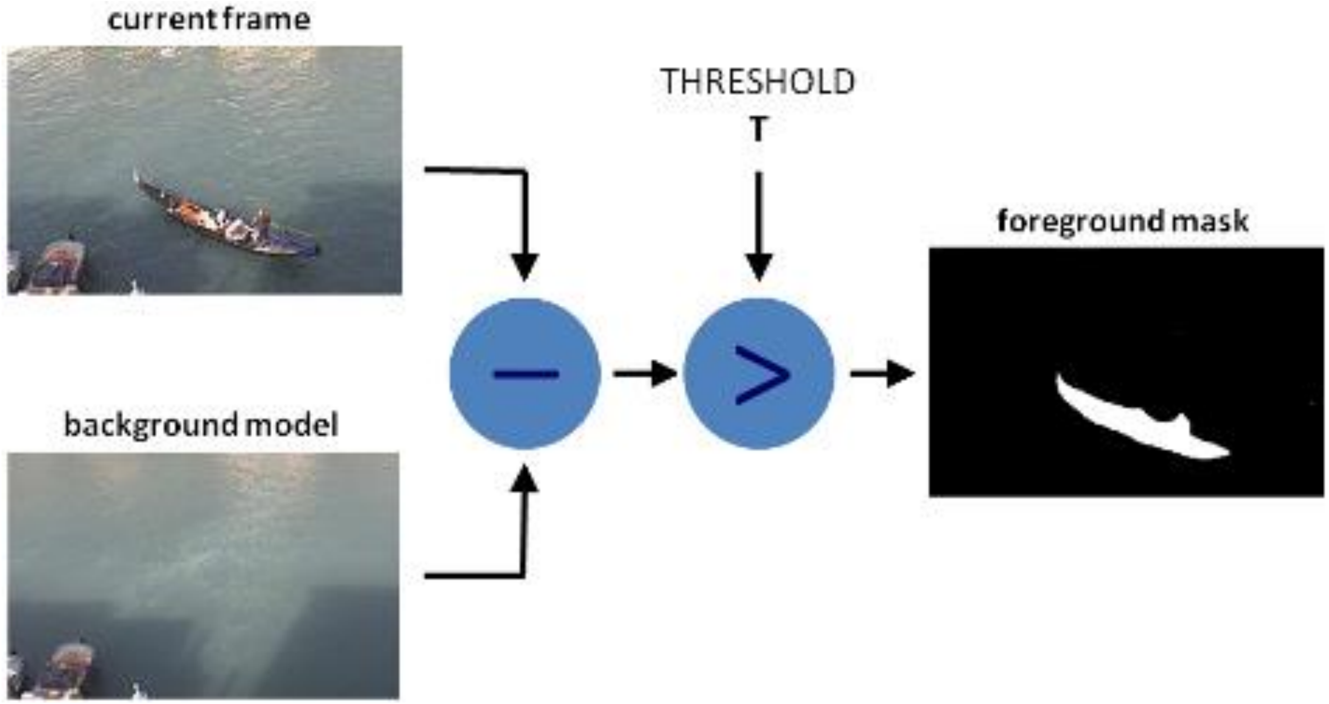
電腦視覺處理的問題(1/2)

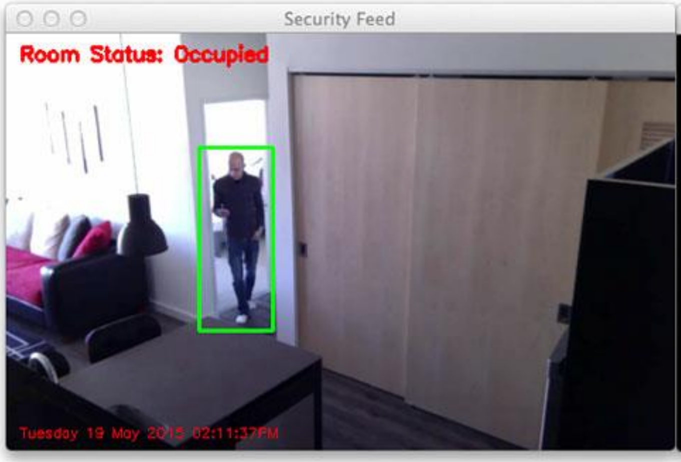
從What到What-Where



電腦視覺處理的問題(2/2)

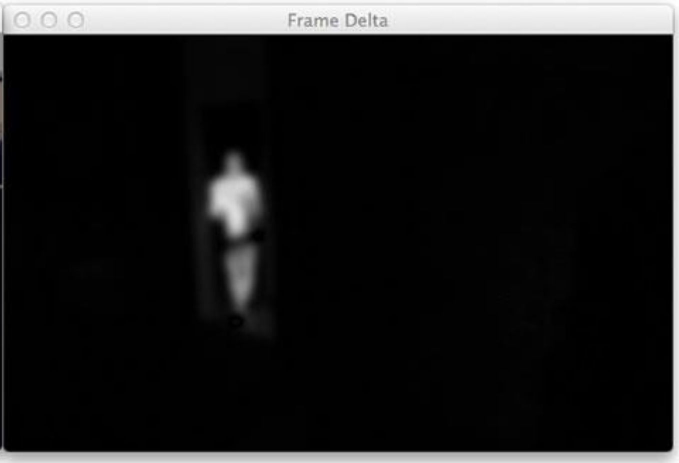
從images到videos





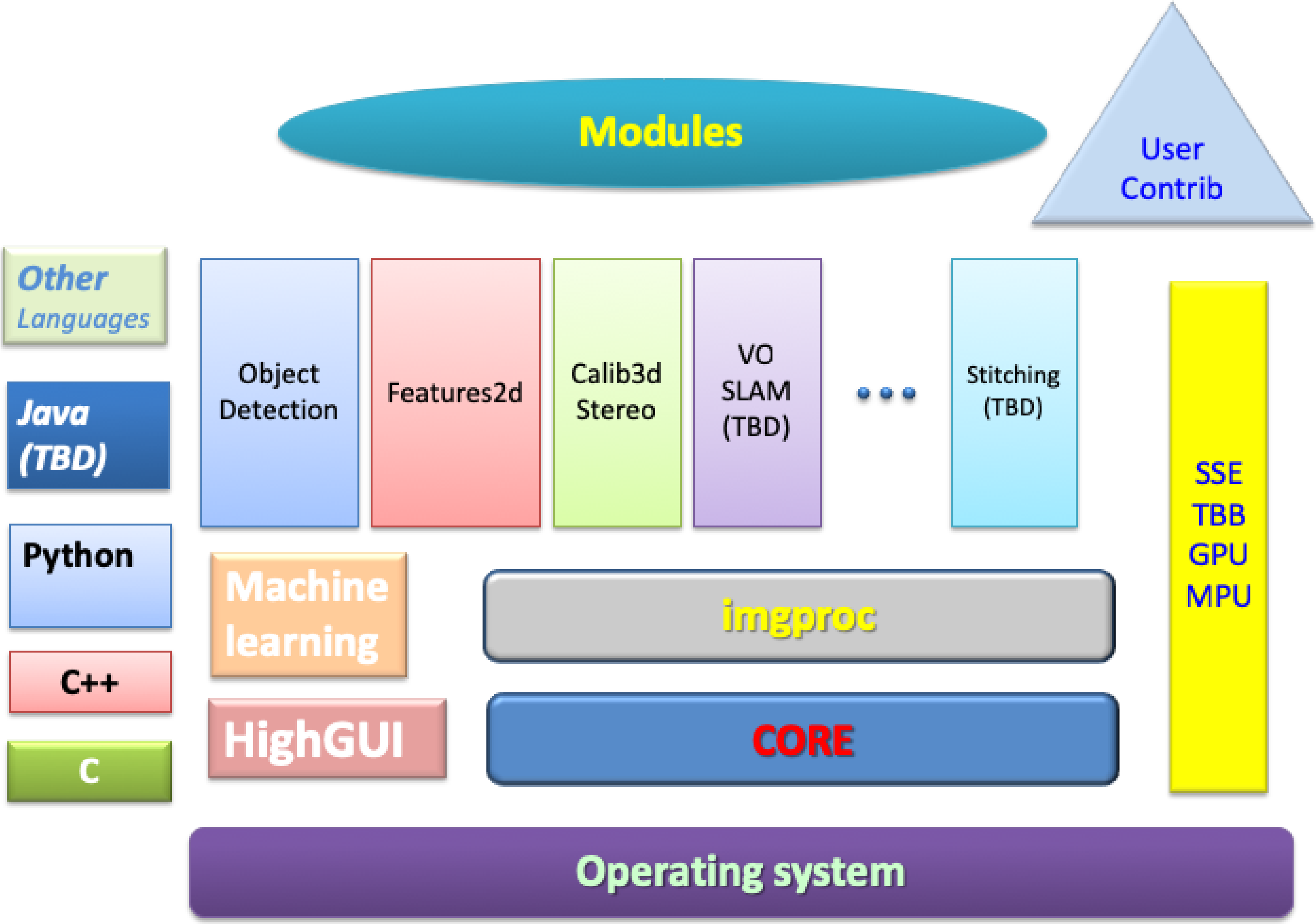
Room Status: Occupied

Tuesday 19 May 2015 02:11:37PM



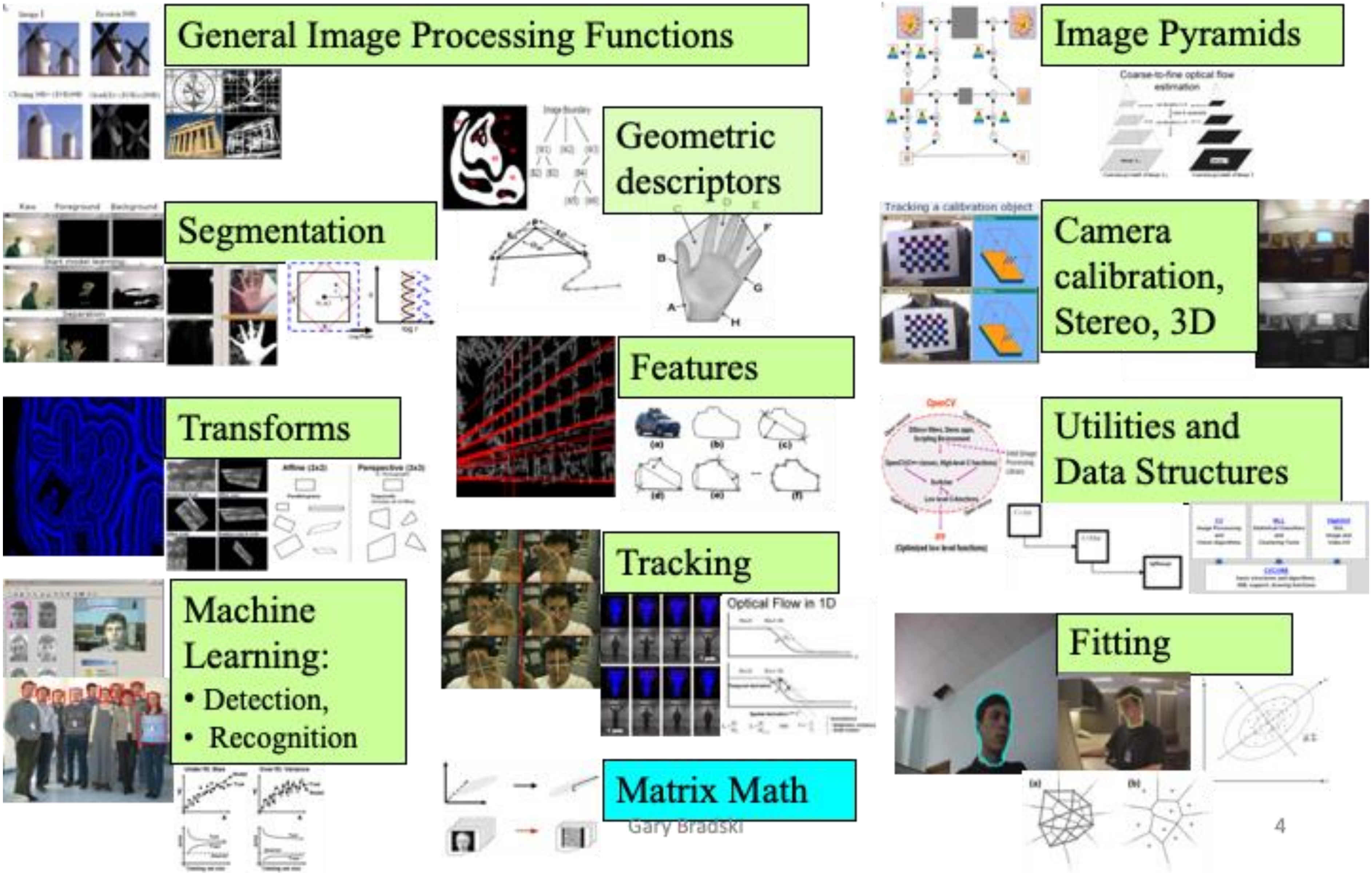
OpenCV (1/2)

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

Segmentation

Transforms

Machine Learning:
- Detection,
- Recognition

Geometric descriptors

Features

Tracking

Matrix Math

Image Pyramids

Camera calibration, Stereo, 3D

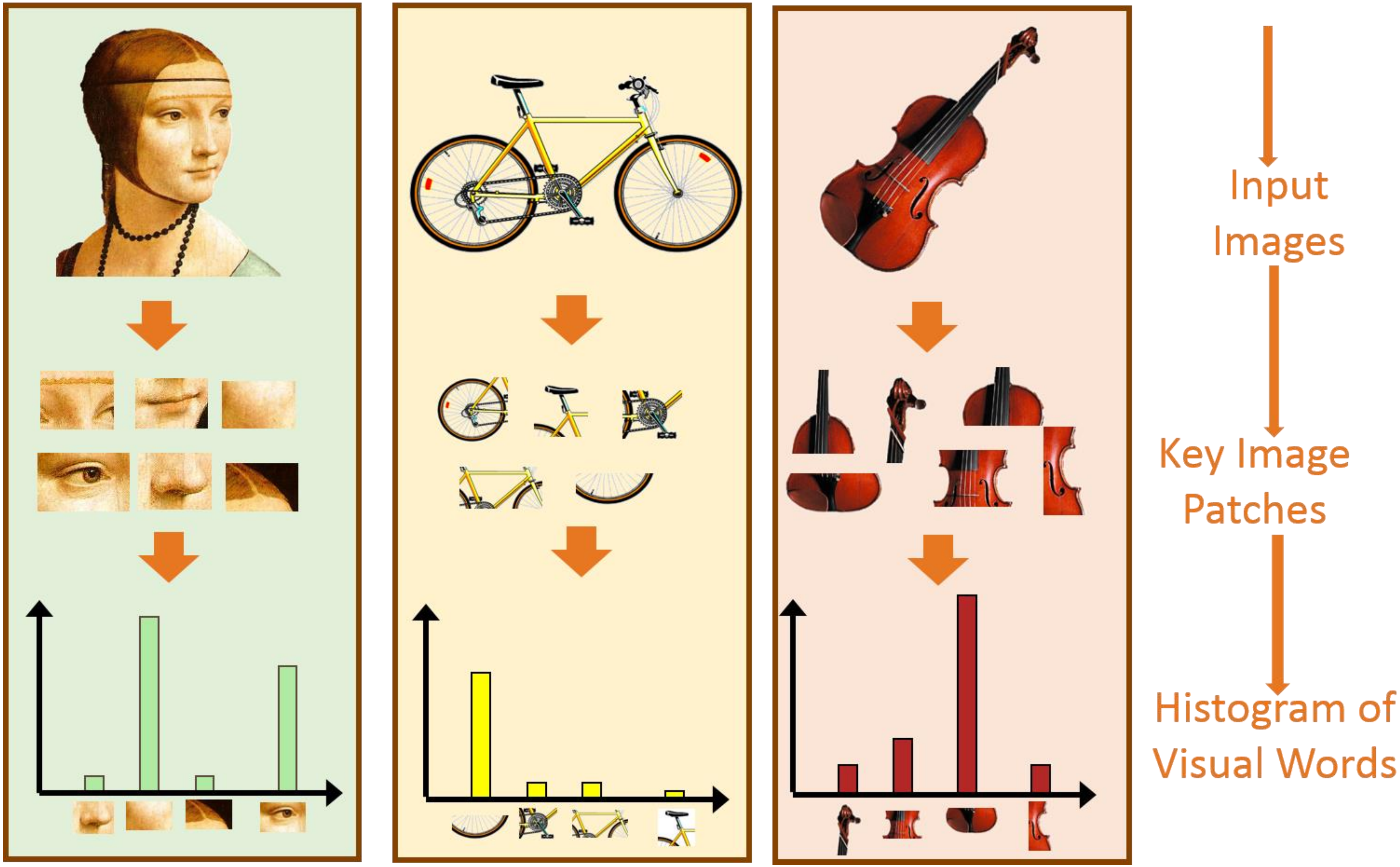
Utilities and Data Structures

Fitting

Tracking a calibration object
Optical Flow in 1D
Coarse-to-fine optical flow estimation
OpenCV
Gary Bradski
4

物體辨識: Visual Bag of Words

不考慮順序，只考慮物體特徵有無出現



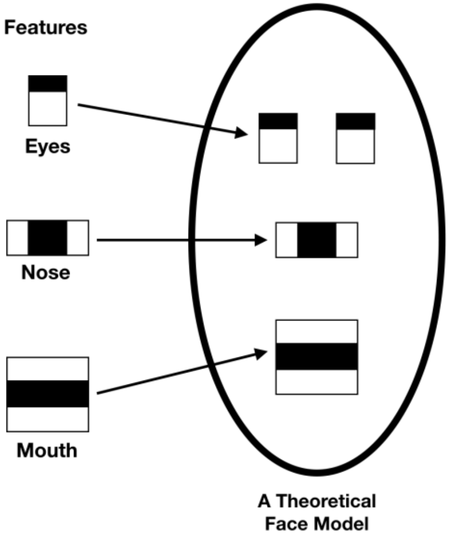
Input Images

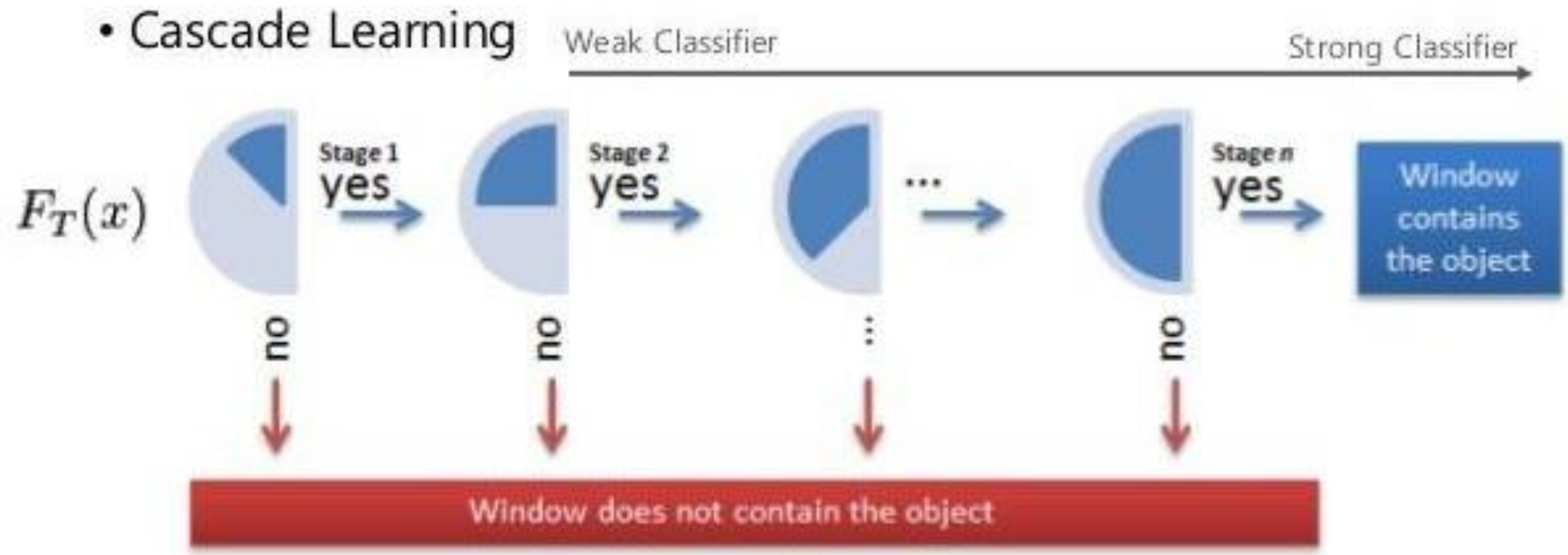
Key Image Patches

Histogram of Visual Words

人臉辨識: Haar Cascades

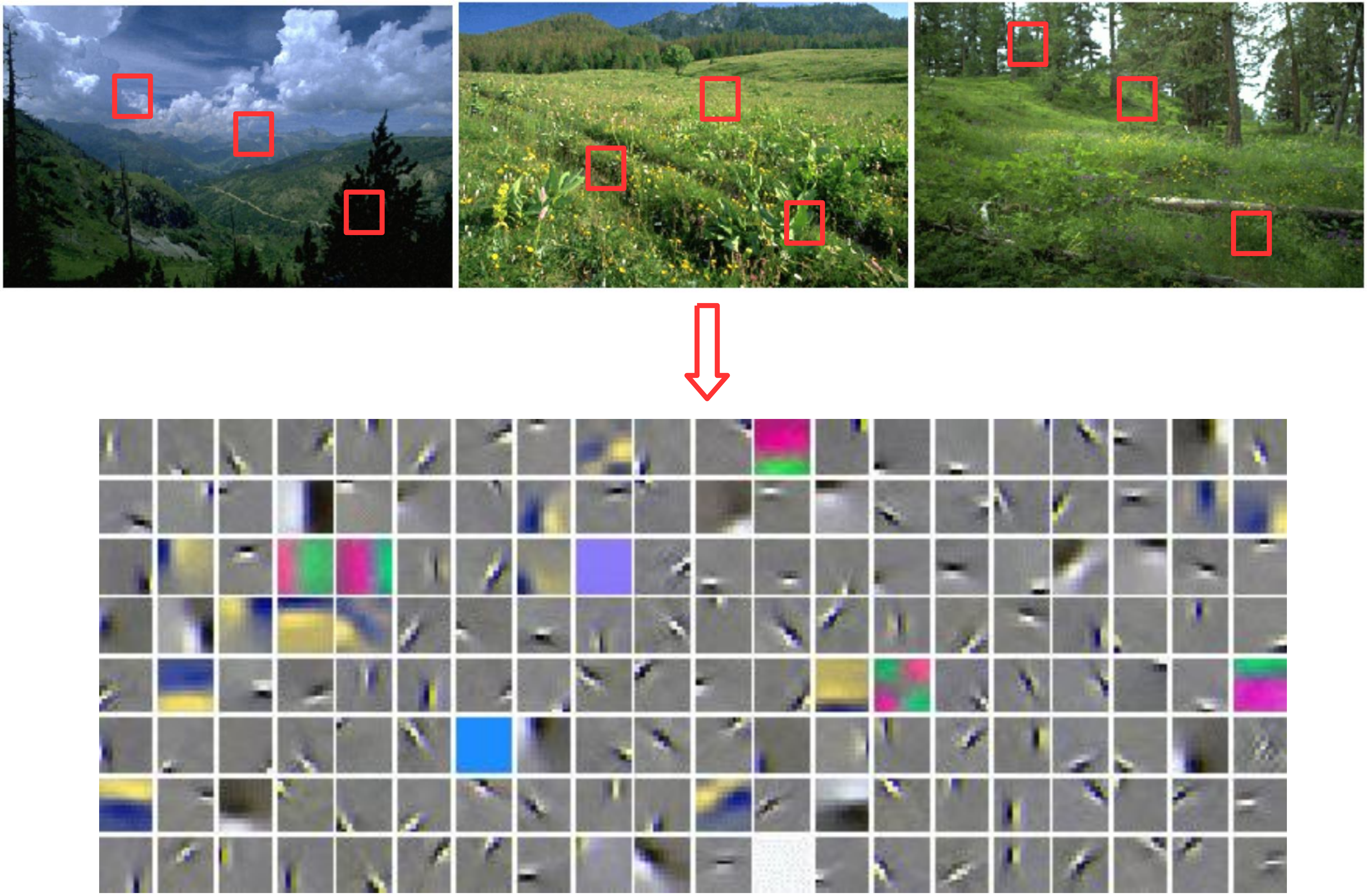
AdaBoost on basic edge/line features





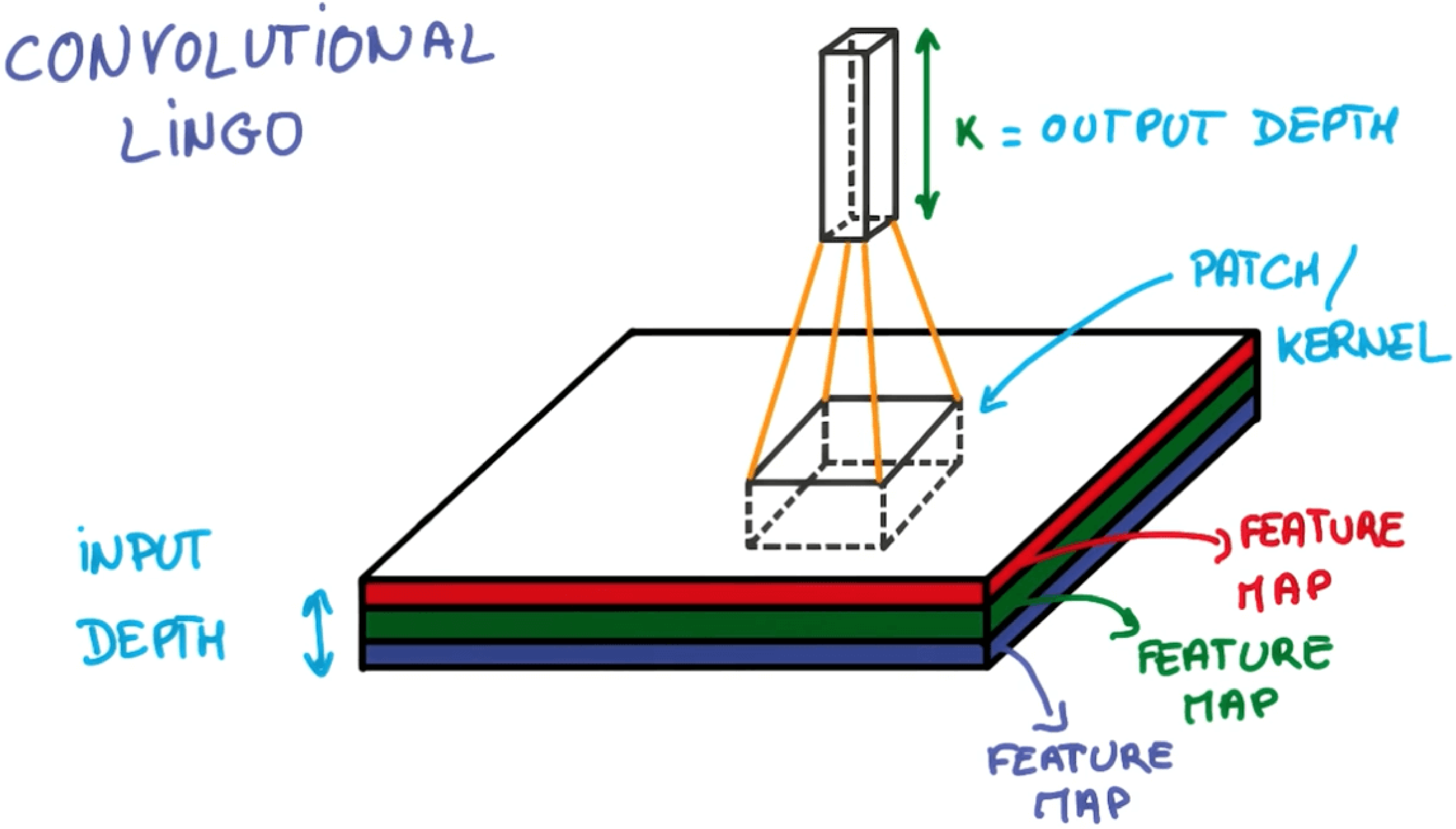
卷積神經網路
(Convolutional Neural Networks)

Shared Patterns of Patches



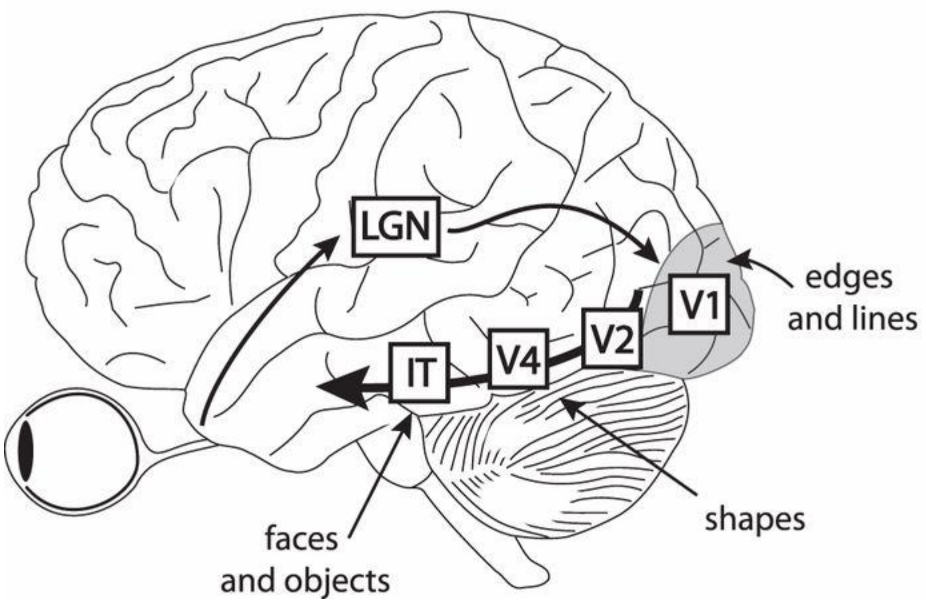
CNN Terms

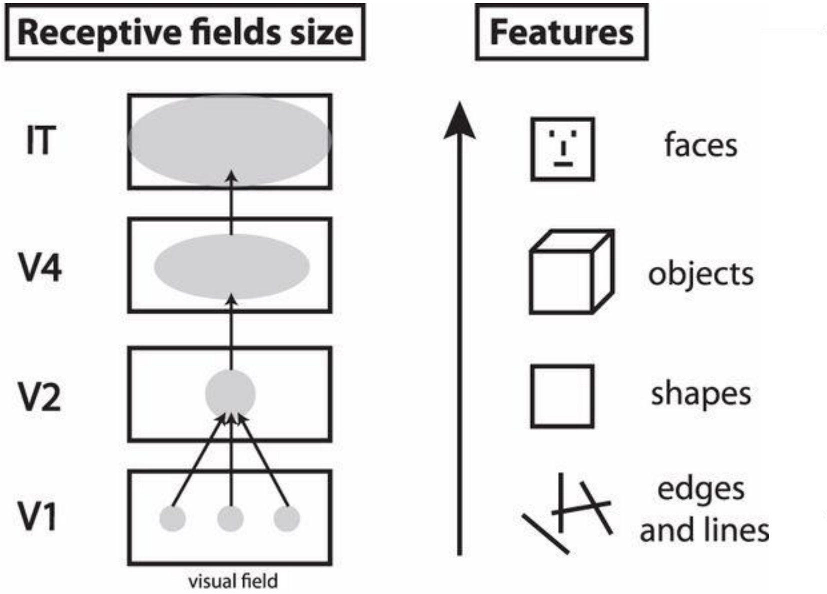
Kernel是template，也是shared network weights

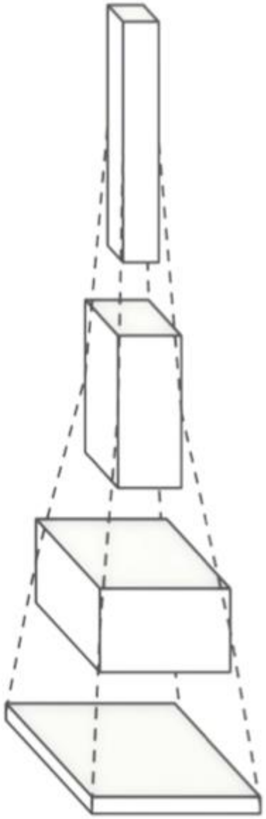


CNN的金字塔結構

是在仿ventral visual pathway in the brain

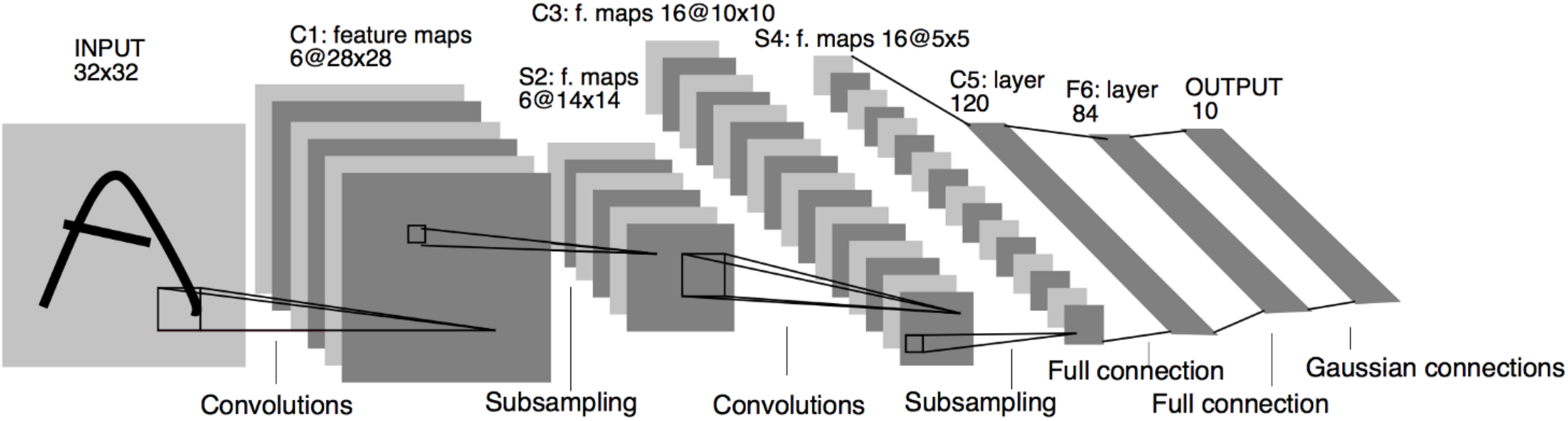


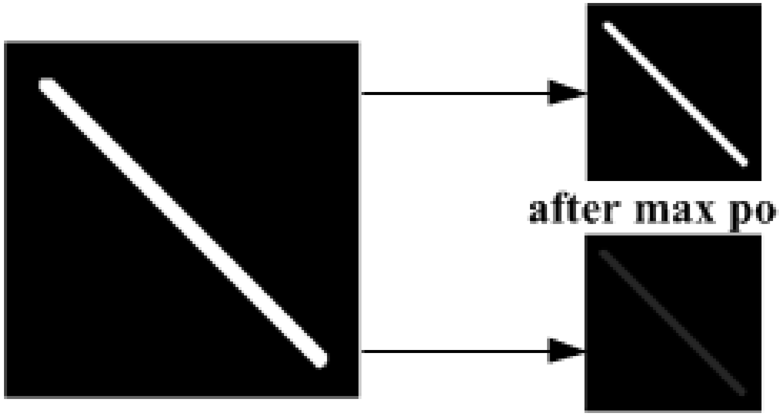




CNN的基本結構

卷積的 $O=1+(I-F) \Rightarrow F=I-O+1=32-28+1=5$

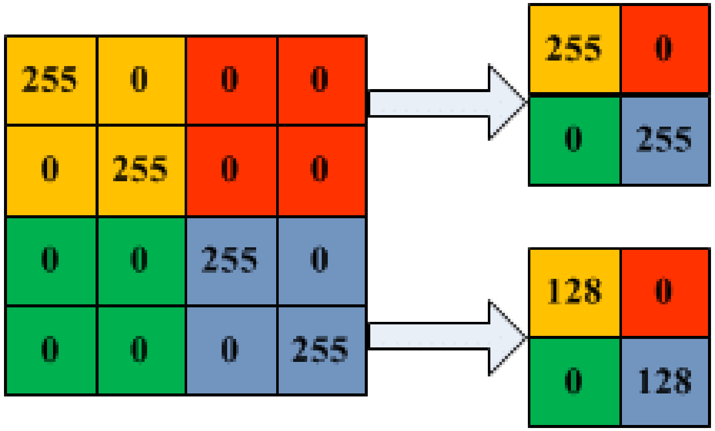




after max pooling

input feature map

after average pooling



255	0	0	0
0	255	0	0
0	0	255	0
0	0	0	255

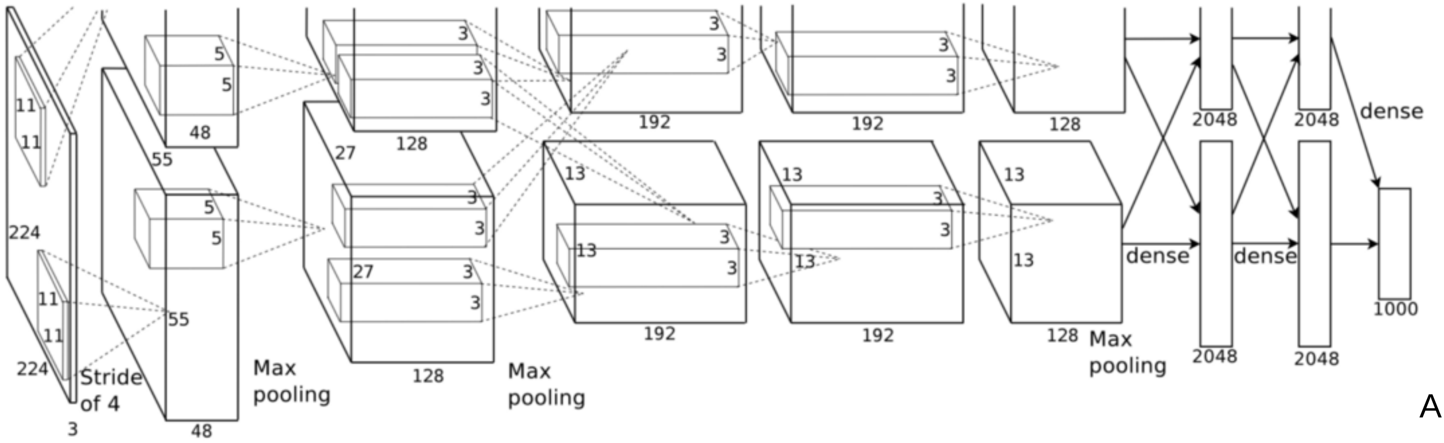
255	0
0	255

128	0
0	128

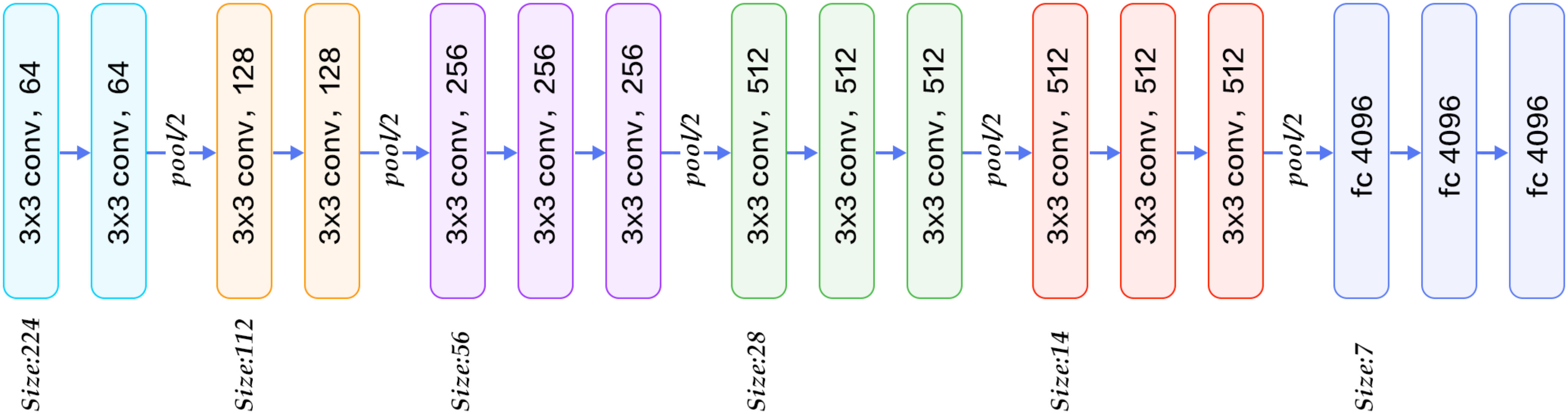
digit express of the pooling process

CNN架構的演進(1/3)

更深/廣→跳接



AlexNet (2012)



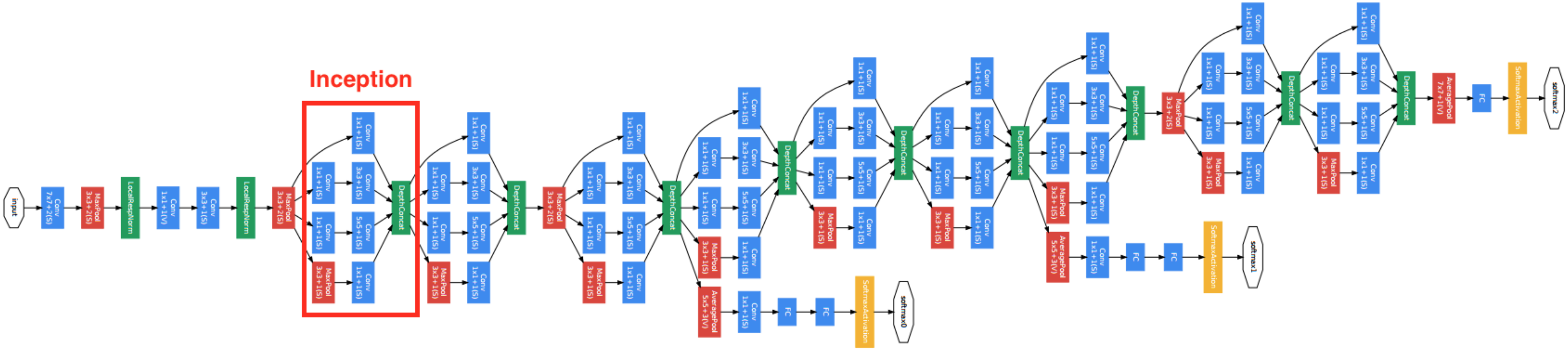
3x3 conv，64 → 3x3 conv，64 → *pool/2* → 3x3 conv，128 → 3x3 conv，128 → *pool/2* → 3x3 conv，256 → 3x3 conv，256 → 3x3 conv，256 → *pool/2* → 3x3 conv，512 → 3x3 conv，512 → 3x3 conv，512 → *pool/2* → 3x3 conv，512 → 3x3 conv，512 → 3x3 conv，512 → *pool/2* → fc 4096 → fc 4096 → fc 4096

Size:224
Size:112
Size:56
Size:28
Size:14
Size:7

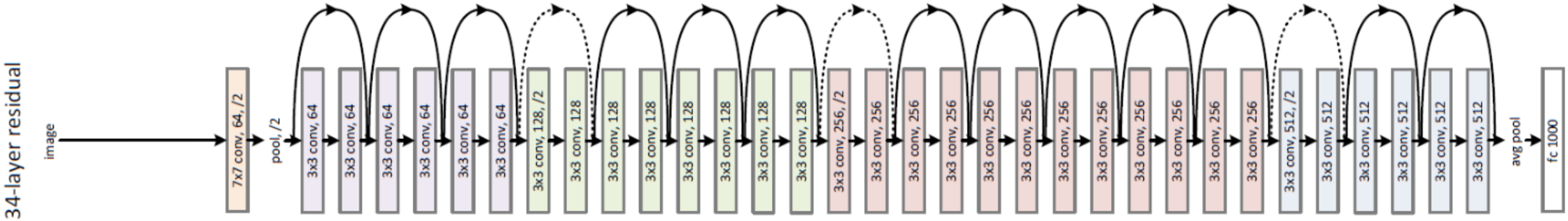
VGG (2014)

CNN架構的演進(2/3)

更深/廣→跳接



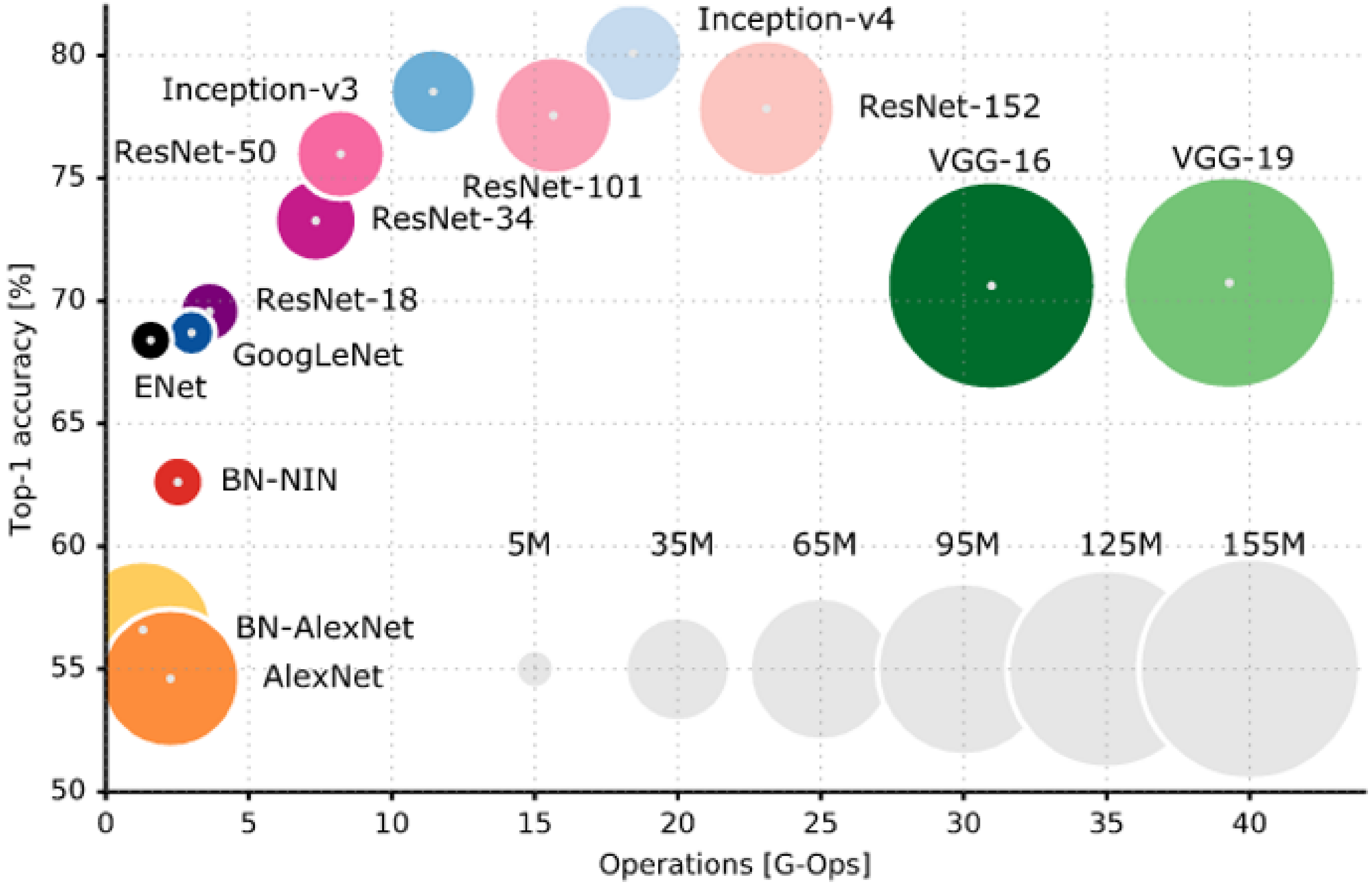
GoogLeNet (2014)



ResNet (2015)

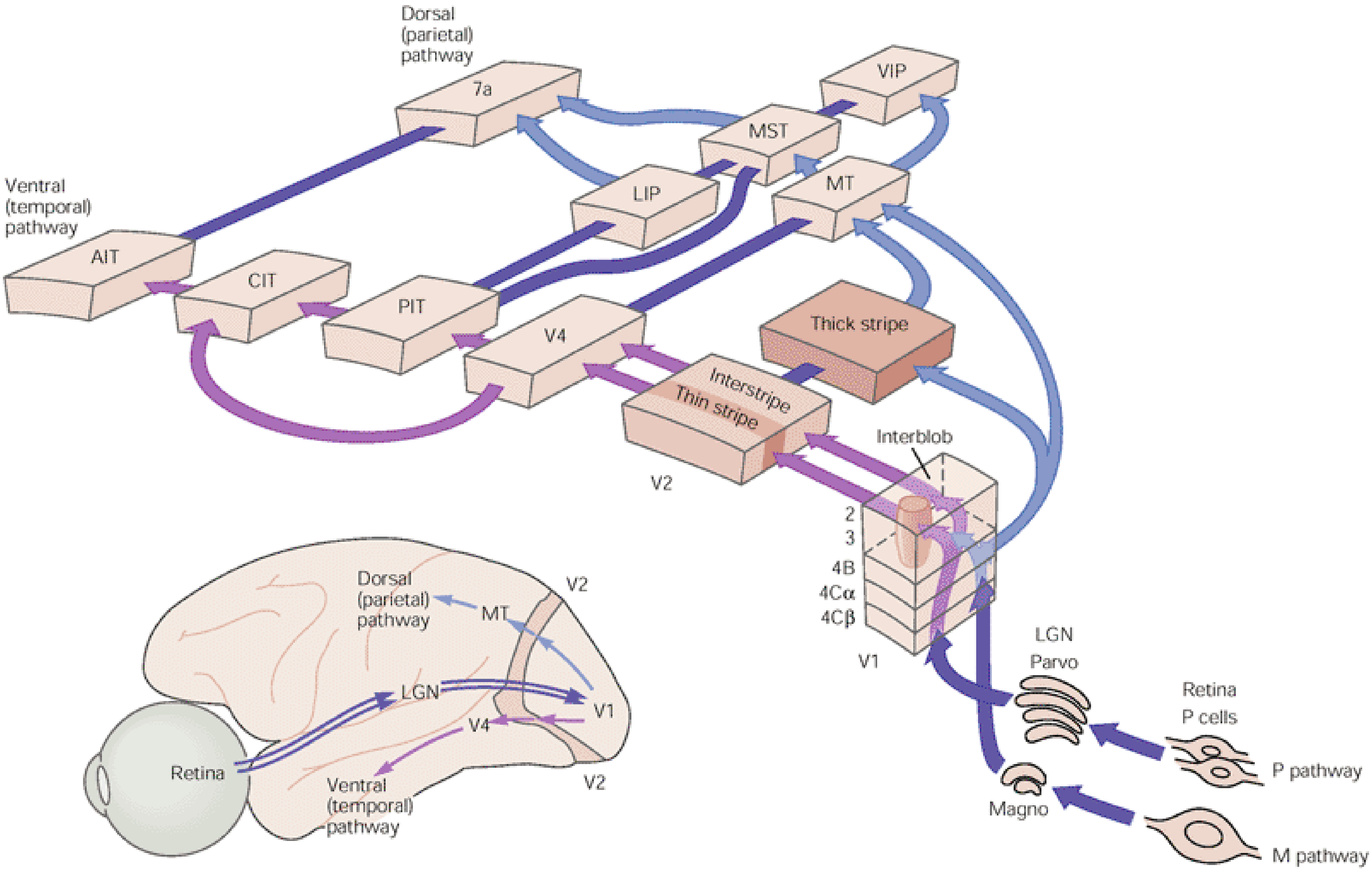
CNN架構的演進(3/3)

最複雜的模型≠表現最好的模型



CNN只模仿了What Pathway

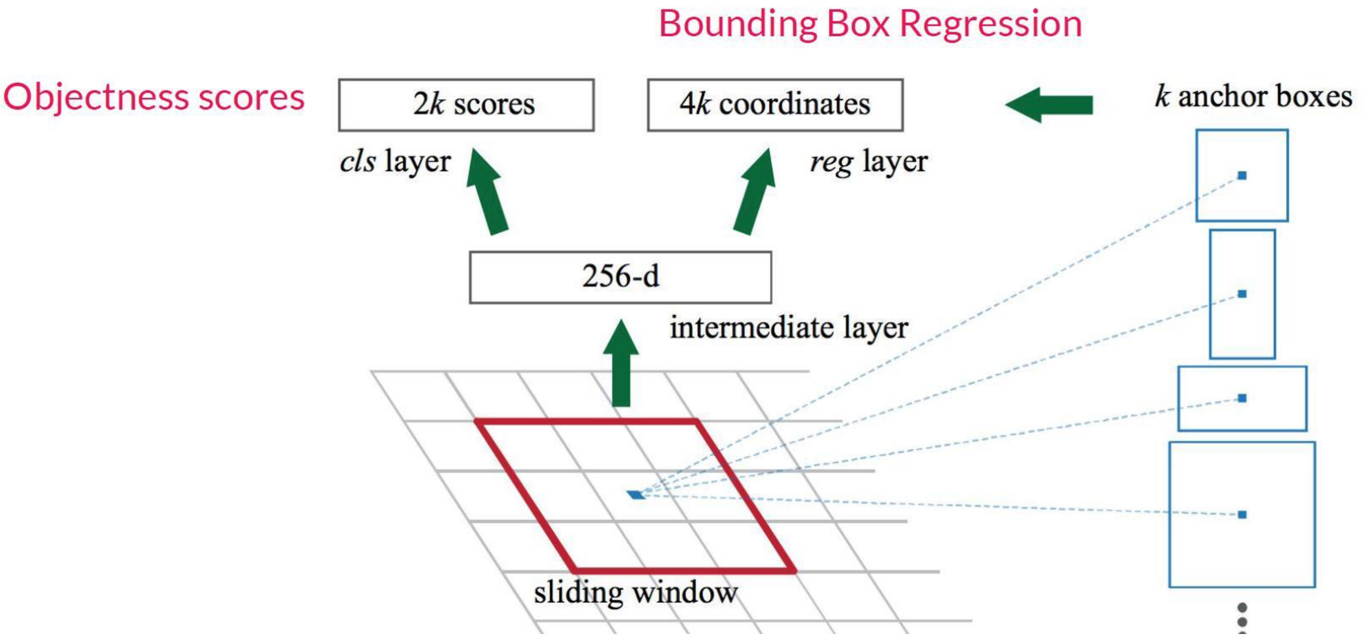
但還有Where Pathway



只模仿What Pathway的不足

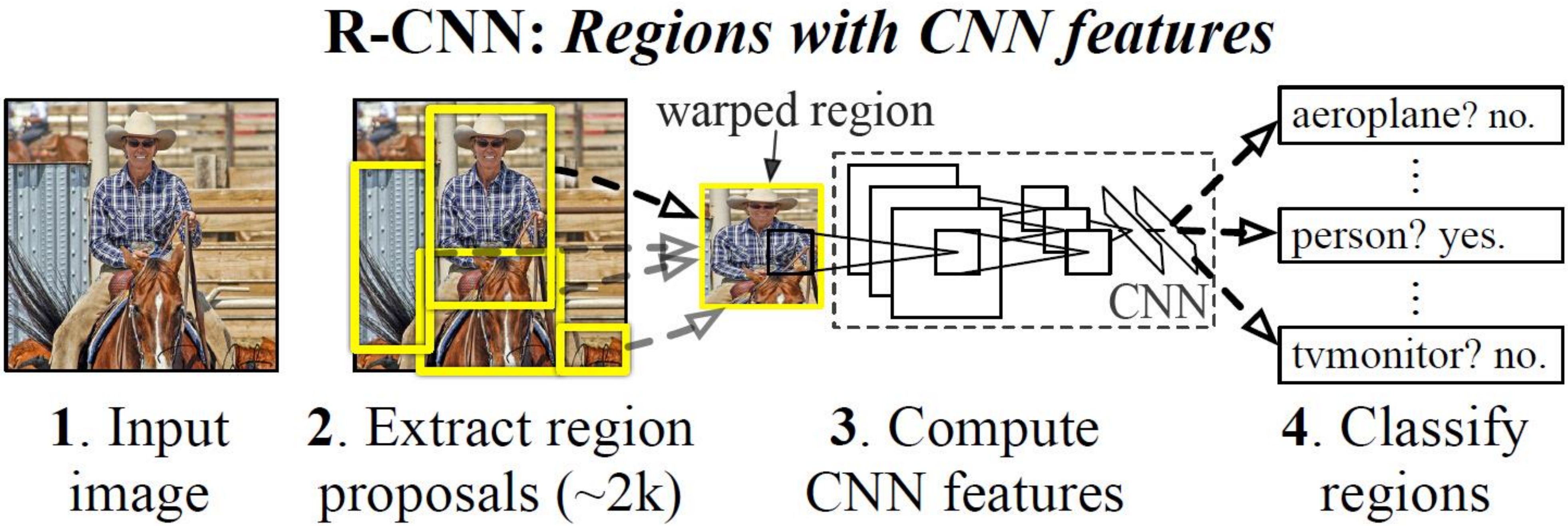
無法分解問題來各個擊破

Region-based CNN (R-CNN)



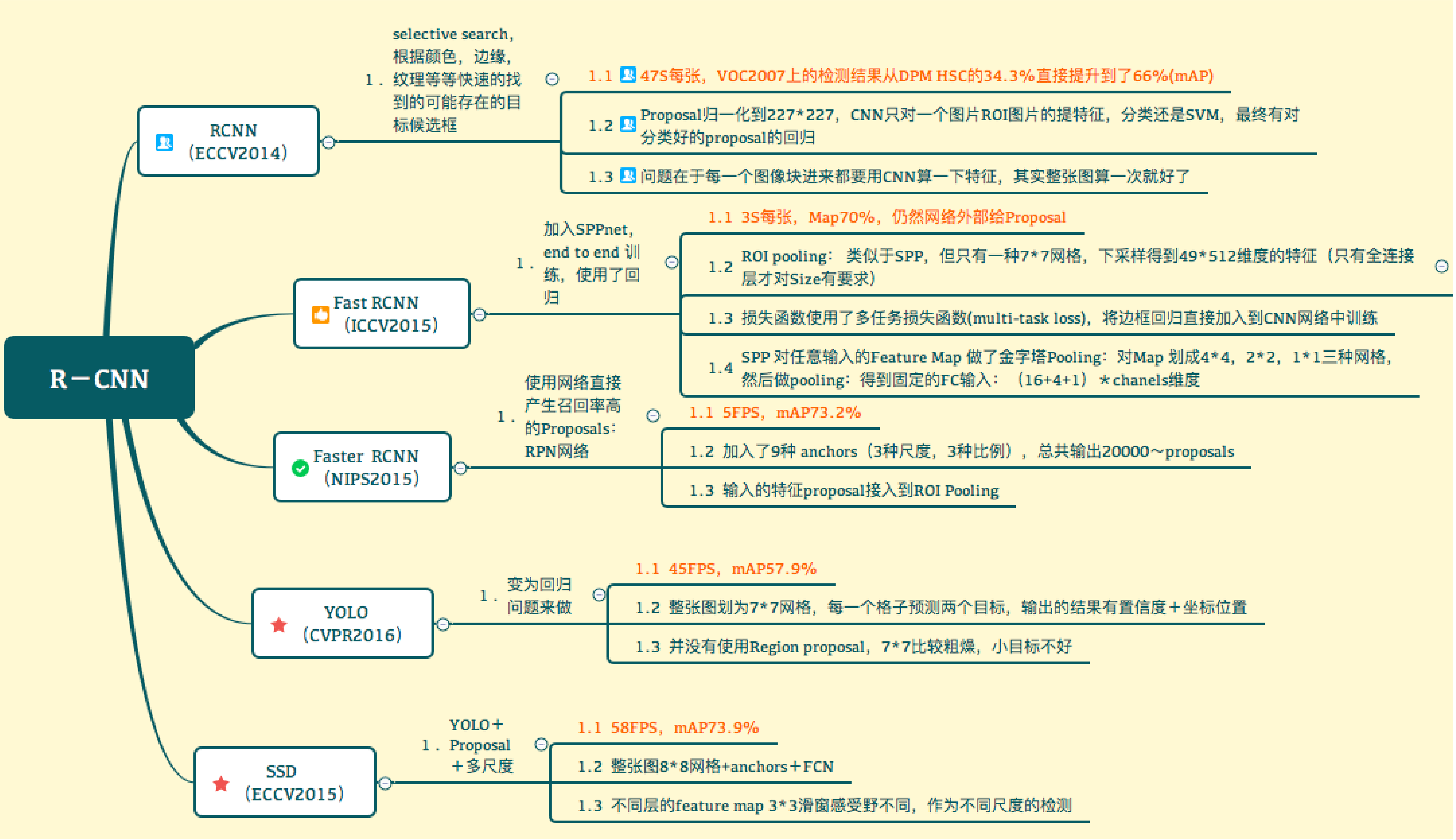
Bounding Box Regression

Objectness scores



R-CNN的徒子徒孫

族繁不及備載



R－CNN

RCNN（ECCV2014）

1． selective search，根据颜色，边缘，纹理等等快速的找到的可能存在的目标候选框

1.1 47S每张，VOC2007上的检测结果从DPM HSC的34.3%直接提升到了66%(mAP)

1.2 Proposal归一化到227*227，CNN只对一个图片ROI图片的提特征，分类还是SVM，最终有对分类好的proposal的回归

1.3 问题在于每一个图像块进来都要用CNN算一下特征，其实整张图算一次就好了

Fast RCNN（ICCV2015）

1． 加入SPPnet，end to end 训练，使用了回归

1.1 3S每张，Map70%，仍然网络外部给Proposal

1.2 ROI pooling： 类似于SPP，但只有一种7*7网格，下采样得到49*512维度的特征（只有全连接层才对Size有要求）

1.3 损失函数使用了多任务损失函数(multi-task loss)，将边框回归直接加入到CNN网络中训练

1.4 SPP 对任意输入的Feature Map 做了金字塔Pooling：对Map 划成4*4，2*2，1*1三种网格，然后做pooling：得到固定的FC输入：（16+4+1）＊chanels维度

Faster RCNN（NIPS2015）

1． 使用网络直接产生召回率高的Proposals：RPN网络

1.1 5FPS，mAP73.2%

1.2 加入了9种 anchors（3种尺度，3种比例），总共输出20000～proposals

1.3 输入的特征proposal接入到ROI Pooling

YOLO（CVPR2016）

1． 变为回归问题来做

1.1 45FPS，mAP57.9%

1.2 整张图划为7*7网格，每一个格子预测两个目标，输出的结果有置信度＋坐标位置

1.3 并没有使用Region proposal，7*7比较粗燥，小目标不好

SSD（ECCV2015）

1． YOLO＋Proposal＋多尺度

1.1 58FPS，mAP73.9%

1.2 整张图8*8网格+anchors＋FCN

1.3 不同层的feature map 3*3滑窗感受野不同，作为不同尺度的检测

Game Over