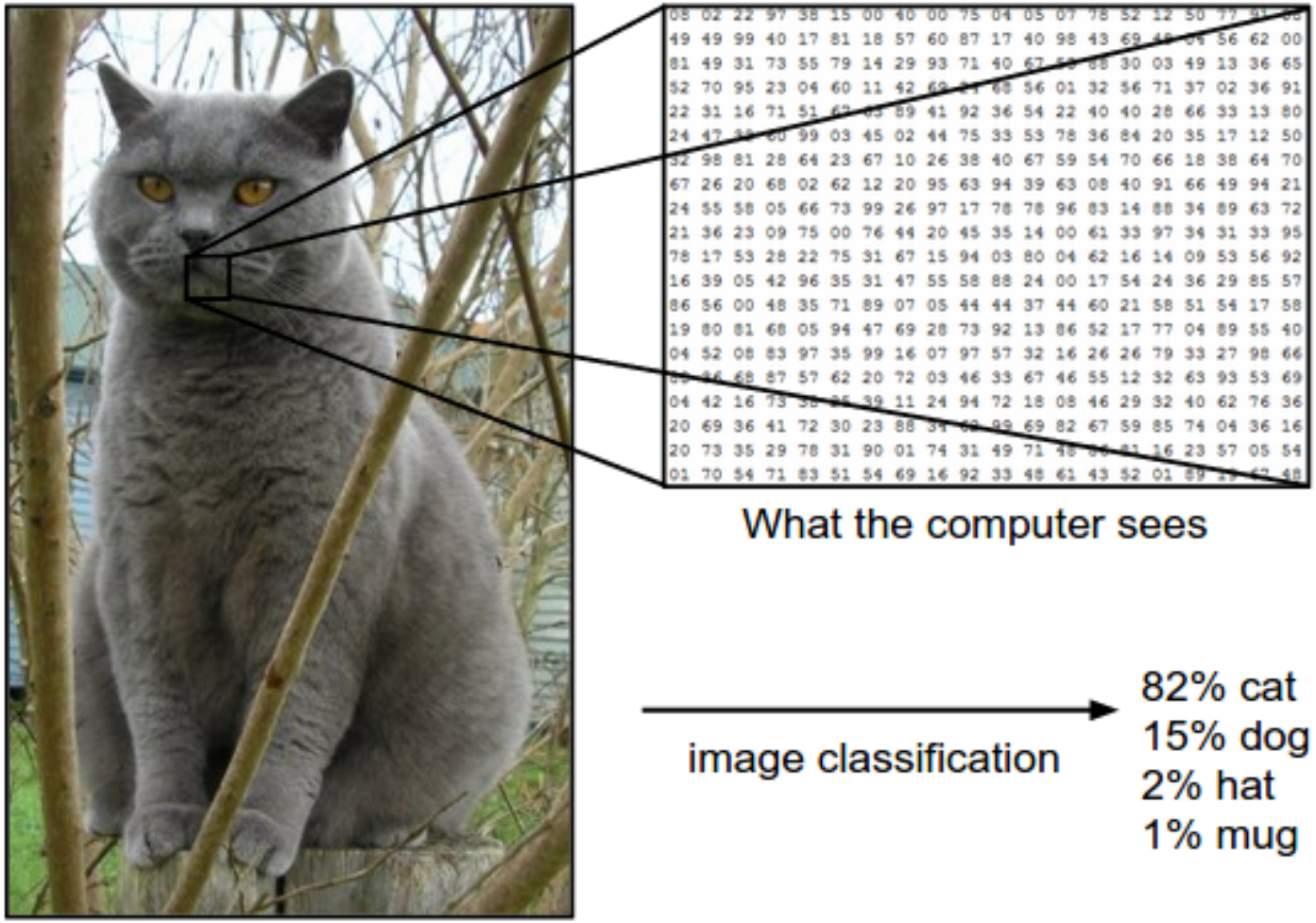
NARLabs 國家實驗研究院

國家高速網路與計算中心

影像瑕疵檢測

Image Classification



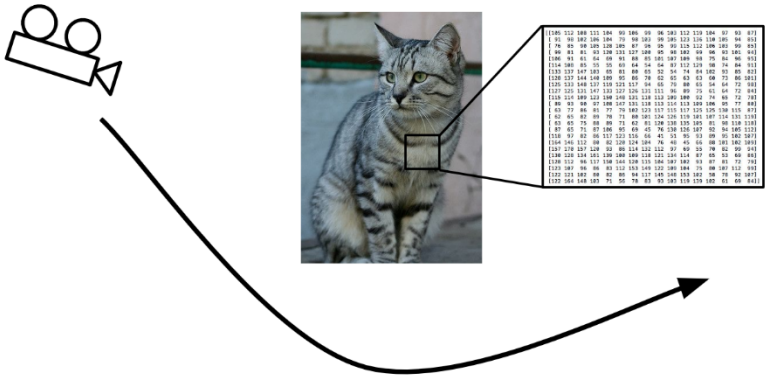
- 影像只是個3維陣列，其值為整數介於0~255之間
- e.g. 800 x 600 x 3 (3 channels RGB)

Source:http://cs231n.github.io/classification/

NARLabs

Image Classification Challenges

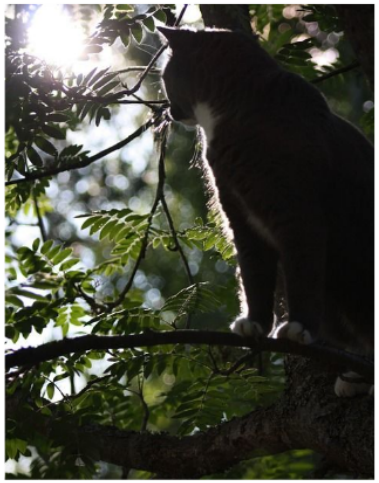
視角變化



背景干擾

This image is CC0 1.0 public domain

照明



This image is CC0 1.0 public domain

遮蔽

This image is CC0 1.0 public domain

變形

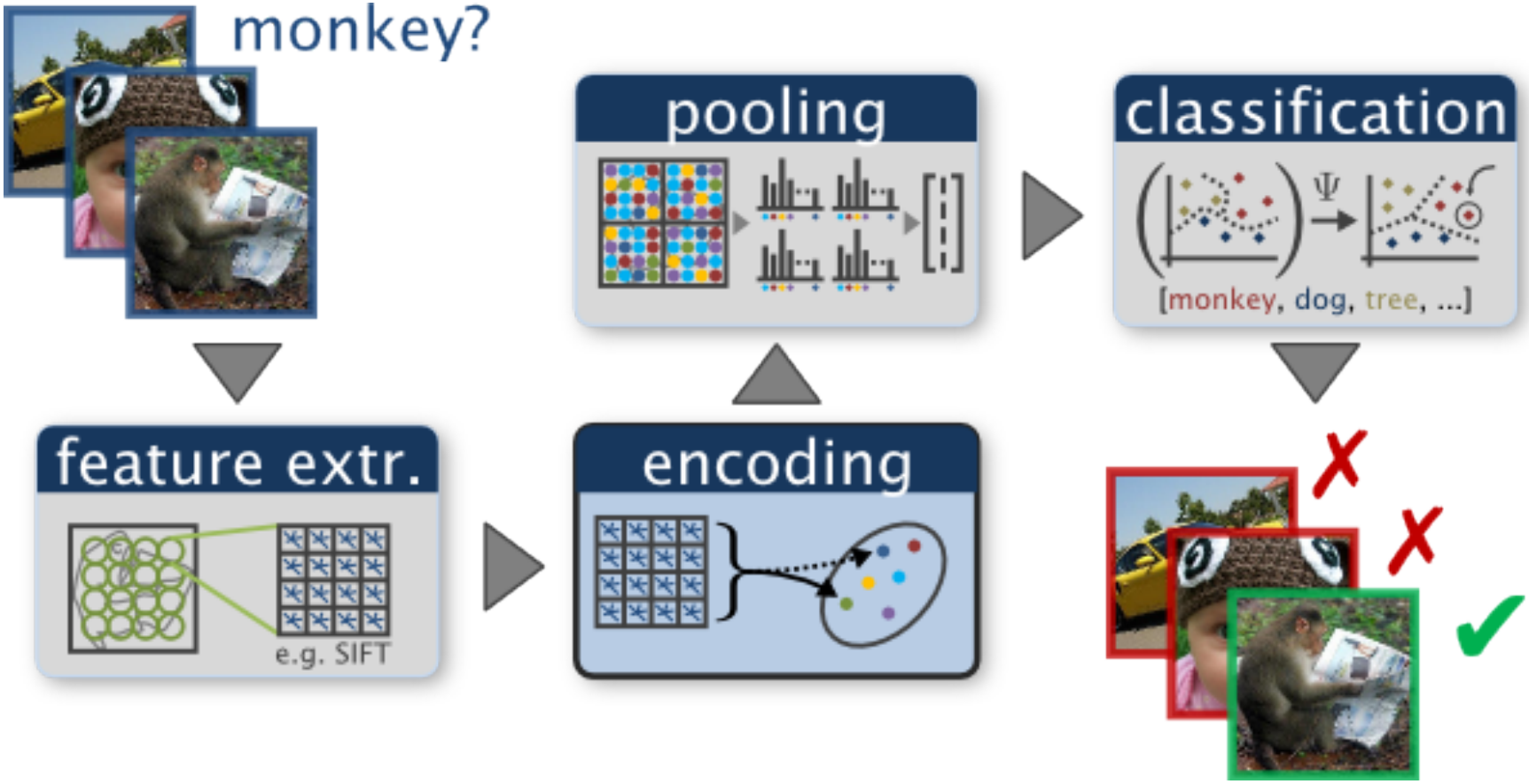
This image by Umberto Salvagnin
is licensed under CC-BY 2.0

同類多樣性

This image is CC0 1.0 public domain

NARLabs

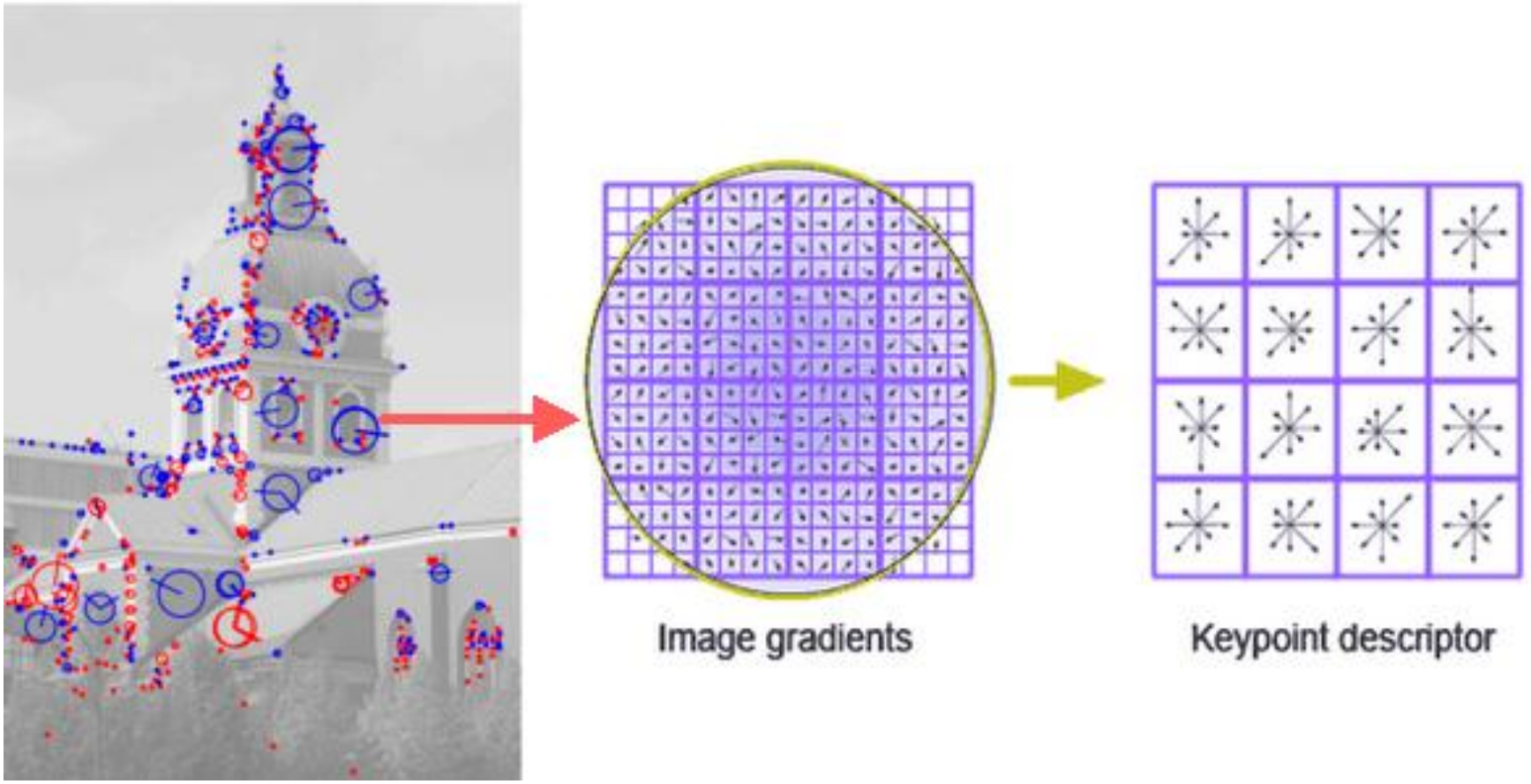
Image Classification Pipeline



Source: https://tanmaybinaykiya.github.io/scene-

Feature extraction

Scale Invariant Feature Transform(SIFT)

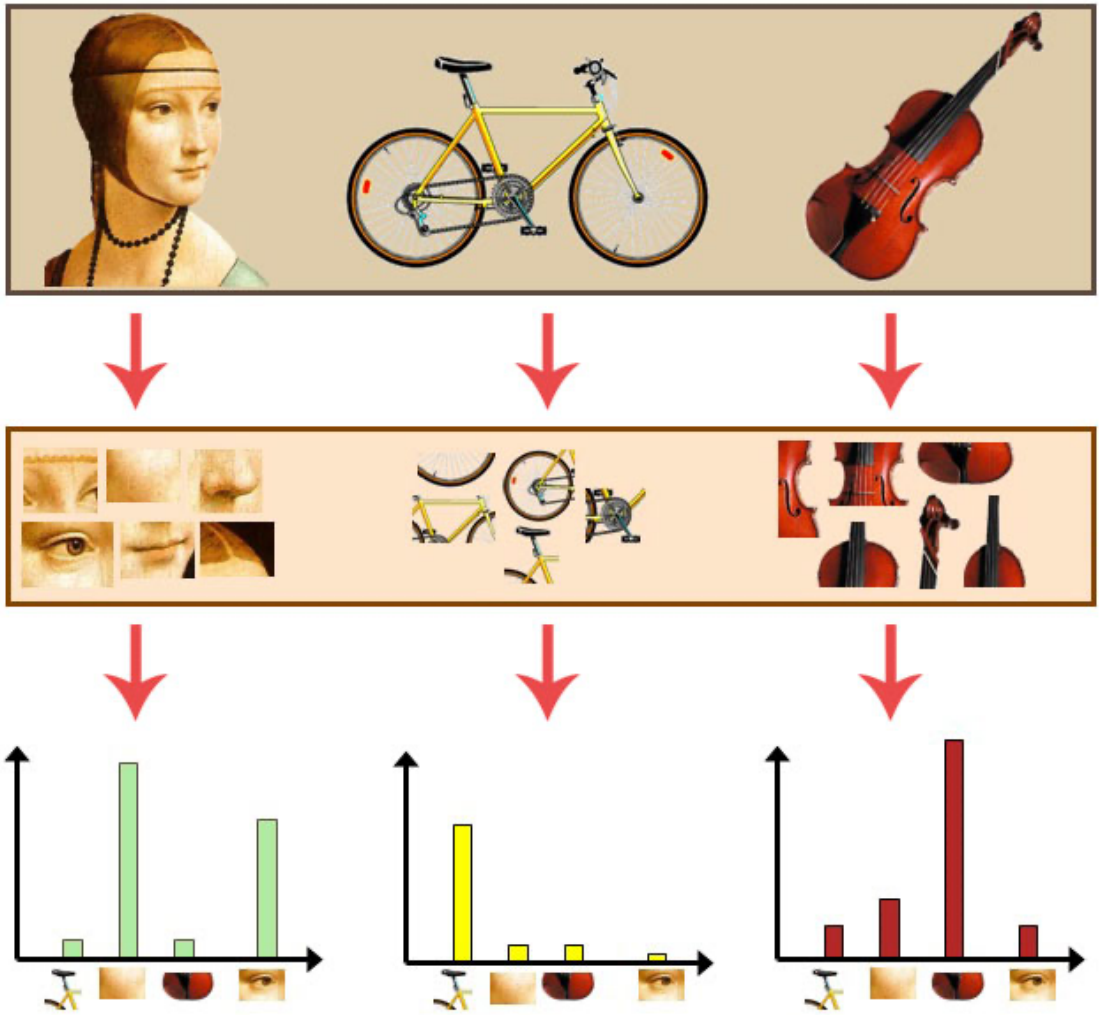


Source: https://xmfbit.github.io/2017/01/30/cs131-sift/

NARLabs

Feature encoding

Bag of visual words

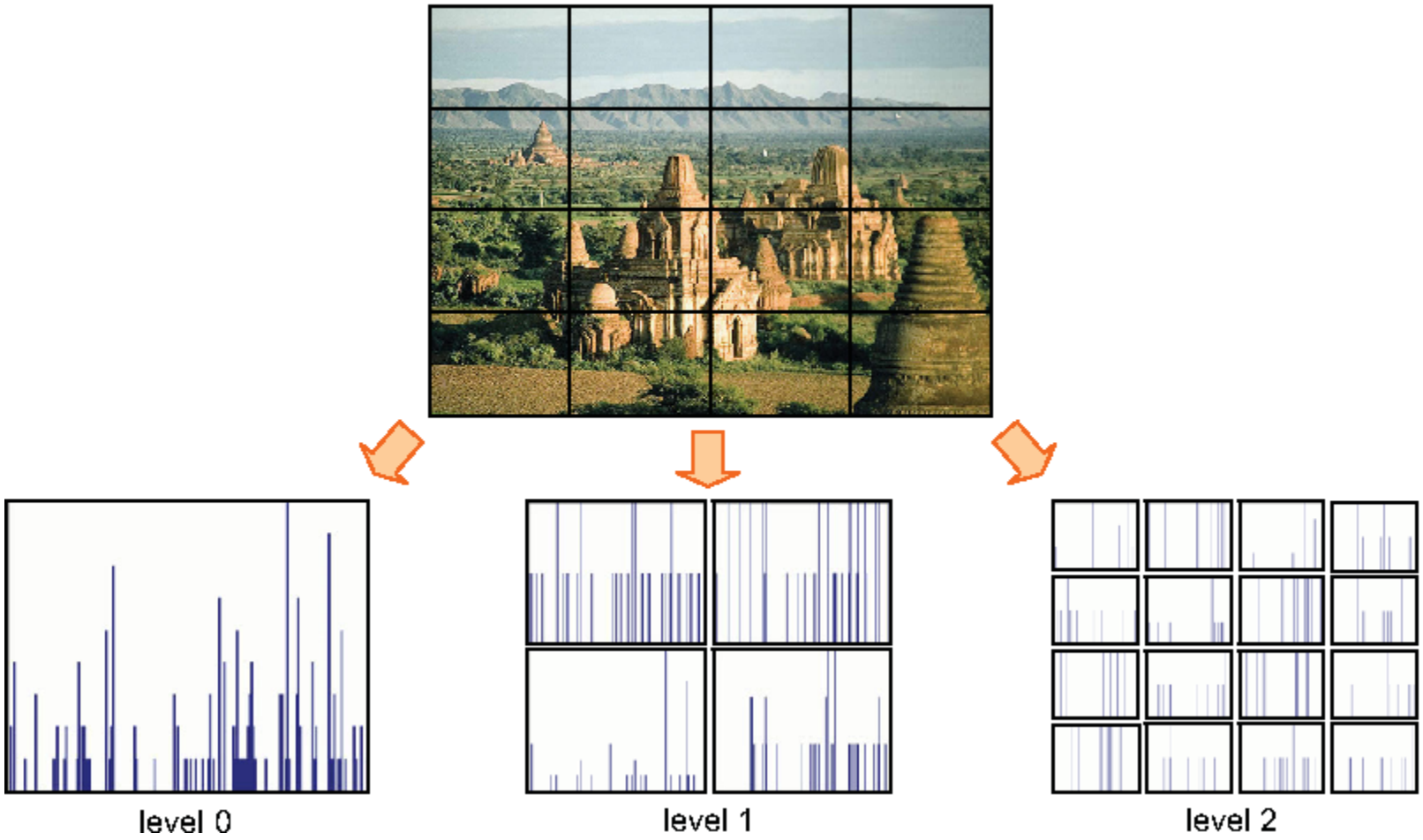


Source: https://www.researchgate.net/figure/An-illustration-of-the-visual-Bag-of-Words-BoW-representation-for-three-input-images_fig1_278827397

Pooling

Spatial Pyramid Matching

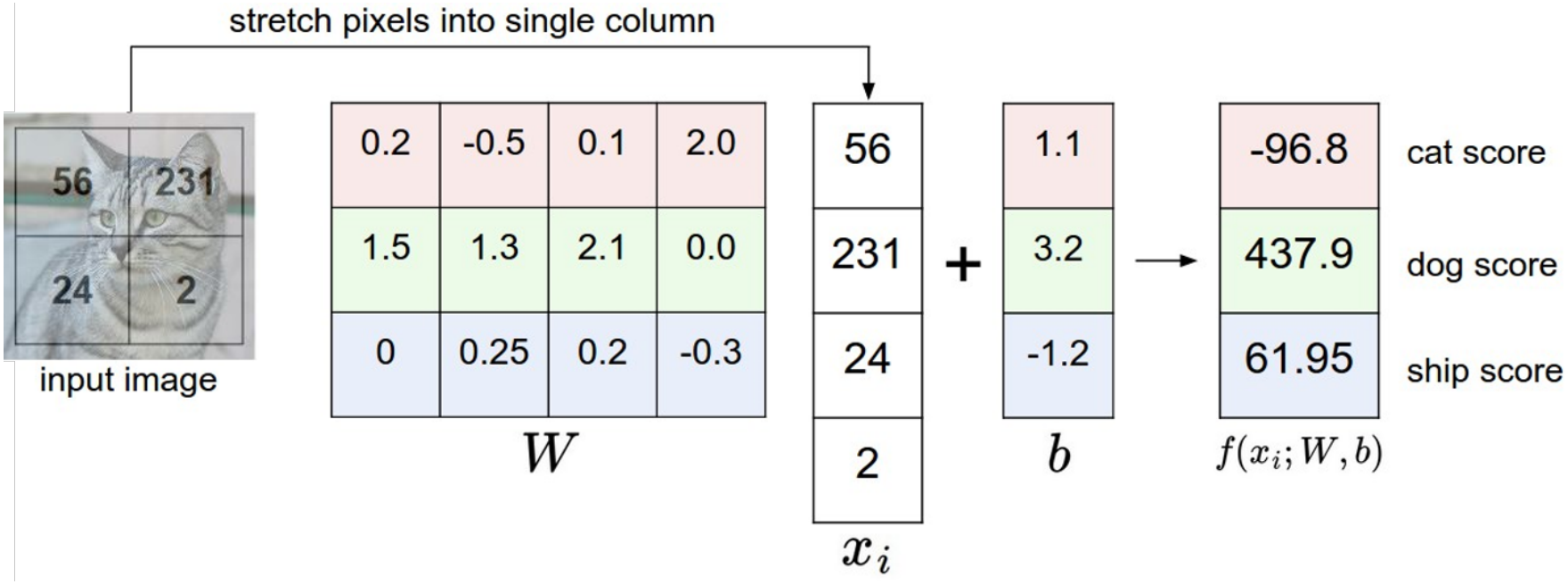
Lazebnik et al.



Source: https://pdfs.semanticscholar.org/d55b/f8b8eba06866f7e34731319bde3800971d7d.pdf

Classification

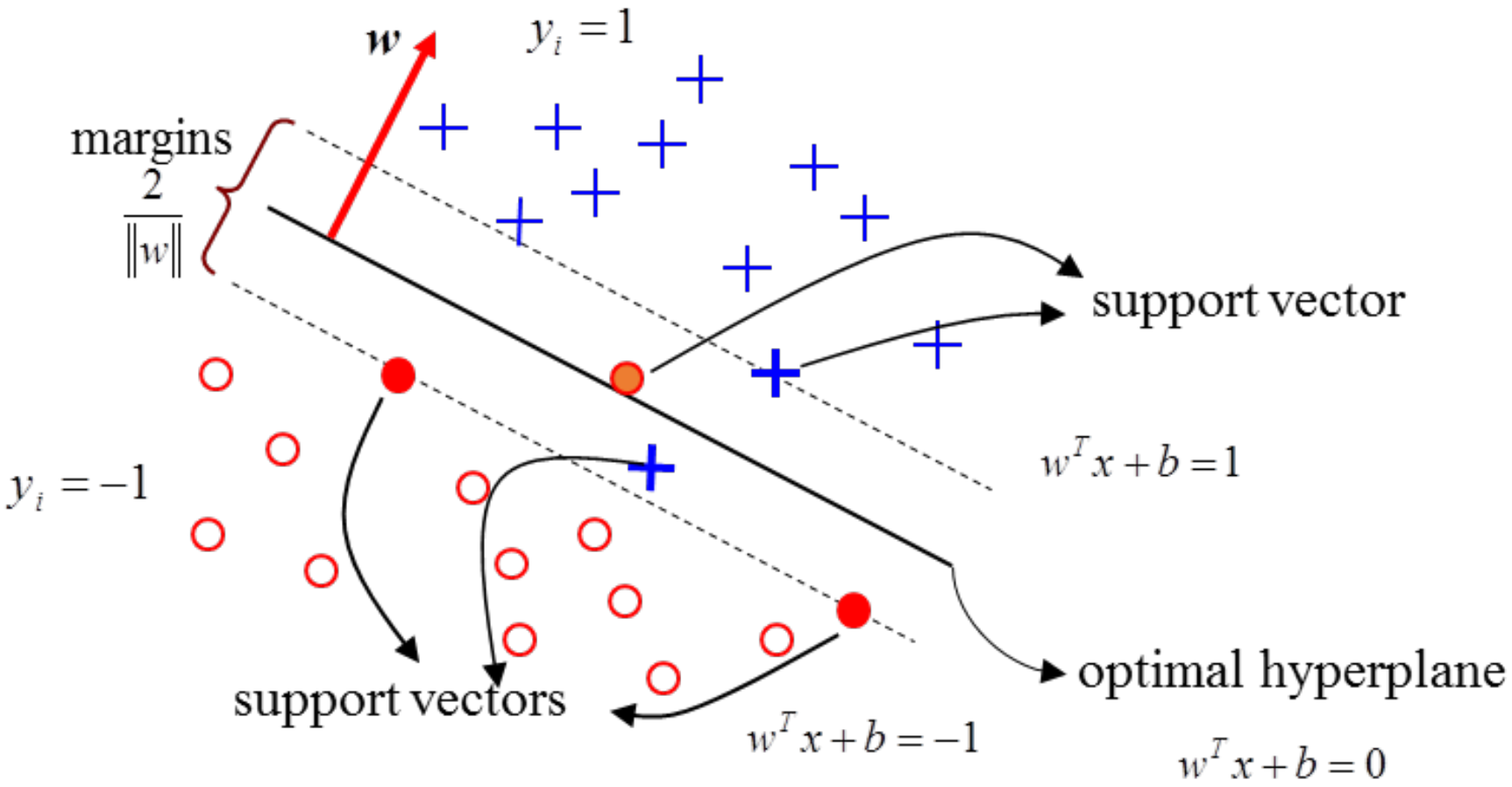
Example with an image with 4 pixels, and 3 classes (cat/dog/ship)



NARLabs

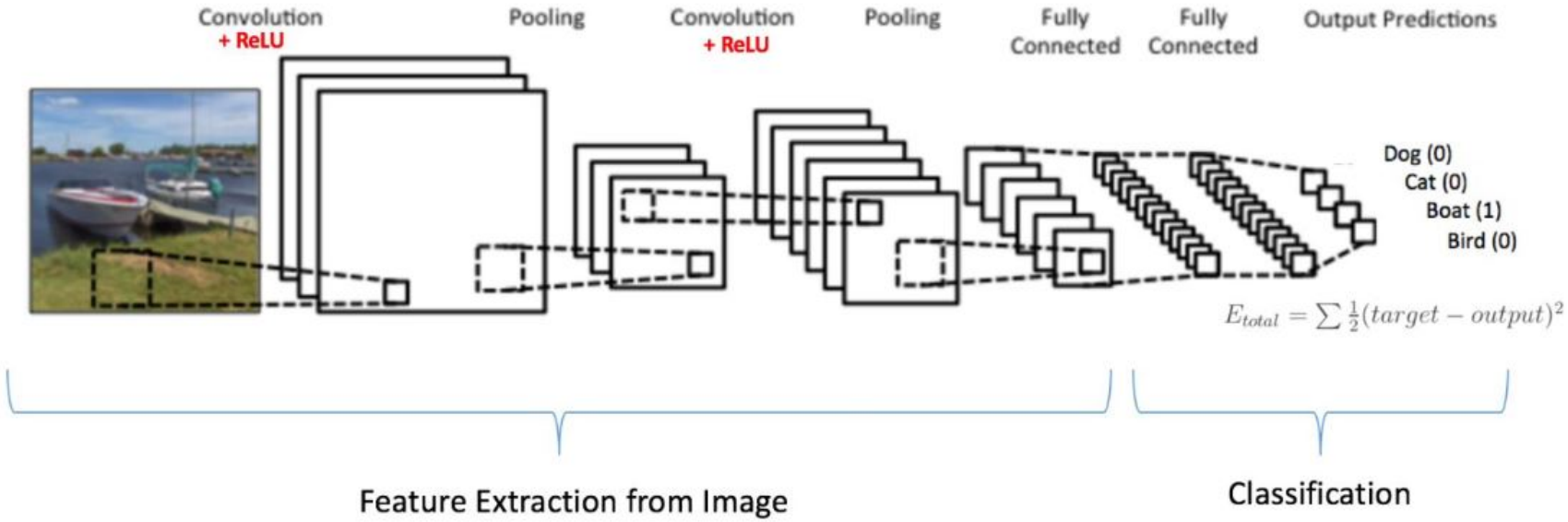
Classification

Support Vector Machine(SVM)



Source: https://miro.medium.com/max/762/1*aqNgiEu0ZBeB-ojoWyryZA.png

Convolutional Neural Networks (CNN)

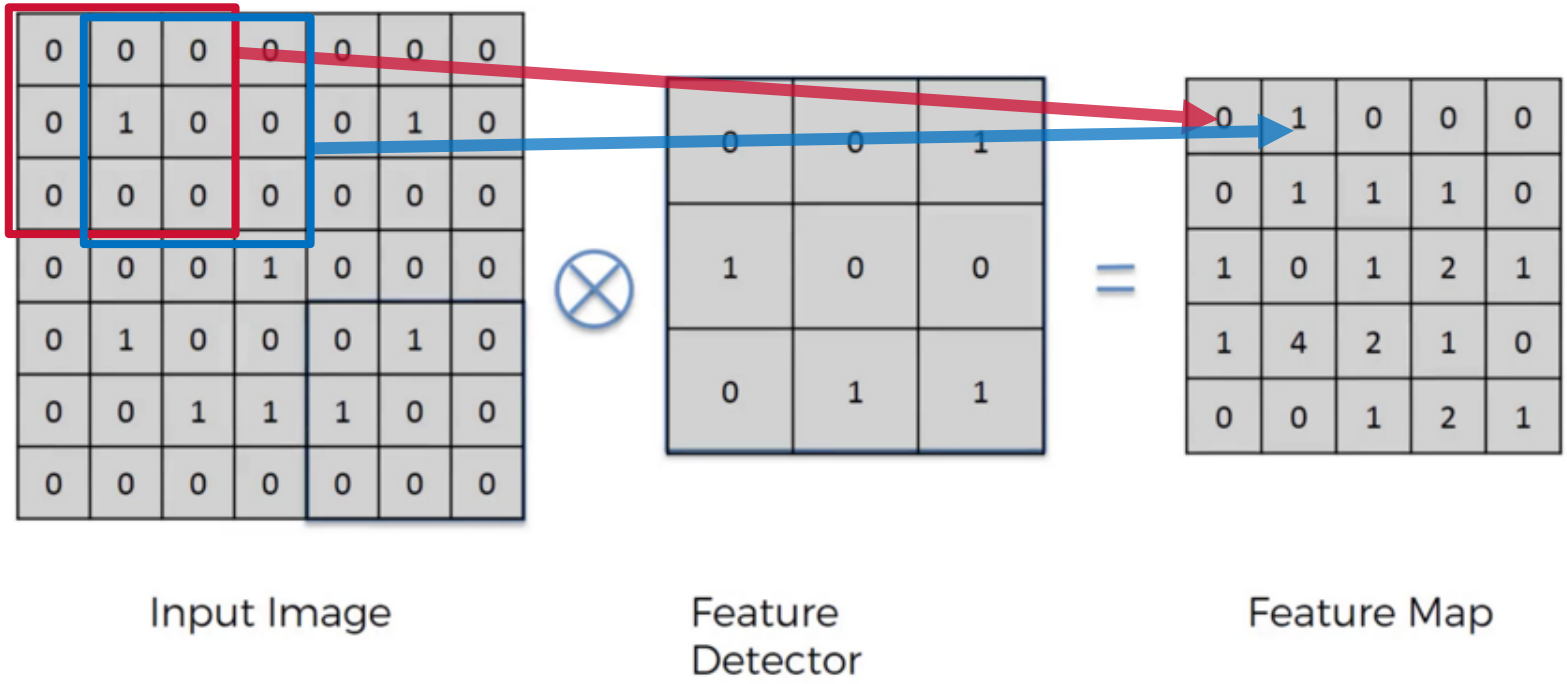


- CNN 通常由以下4種元件構成:
  - Convolution
  - Non Linearity (ReLU)
  - Pooling or Subsampling
  - Classification (Fully Connected Layers)

Source: https://courses.cs.washington.edu/courses/cse484/19sp/files/cse484-guest-lecture-ivan-sp19.pdf

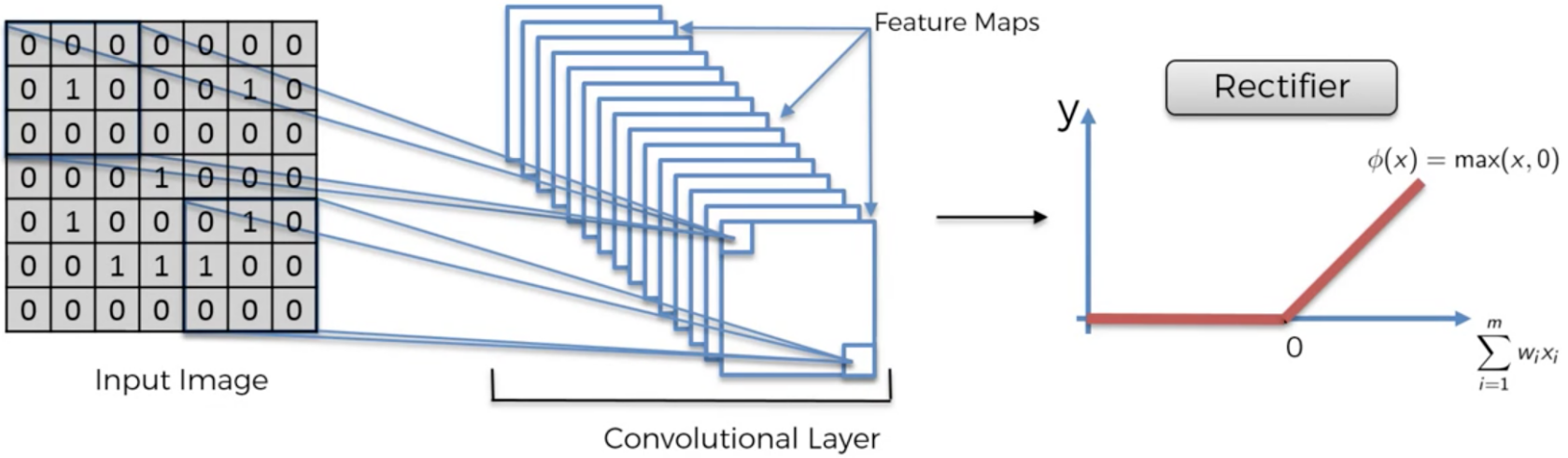
Convolution卷積層

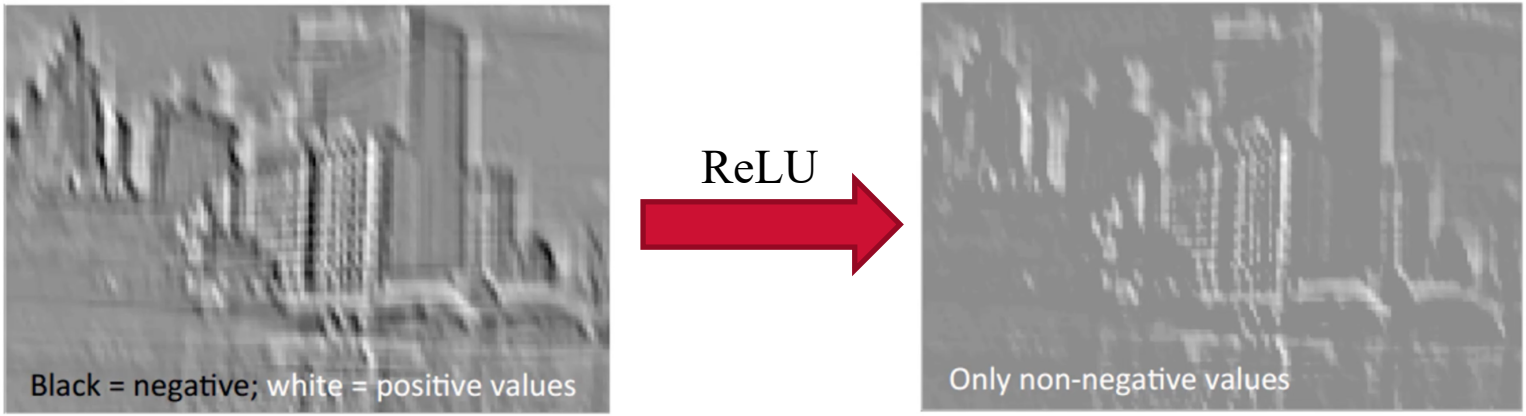
- 將原始圖片的與特定的Feature Detector(filter)做卷積運算
- 卷積運算就是將下圖兩個3x3的矩陣作相乘後再相加
- Feature Detector(Filter)會隨機產生好幾種，其目的就是萃取出影像當中的一些特徵



Source: https://reurl.cc/ObLLzR

Rectified Linear Unit(ReLU)

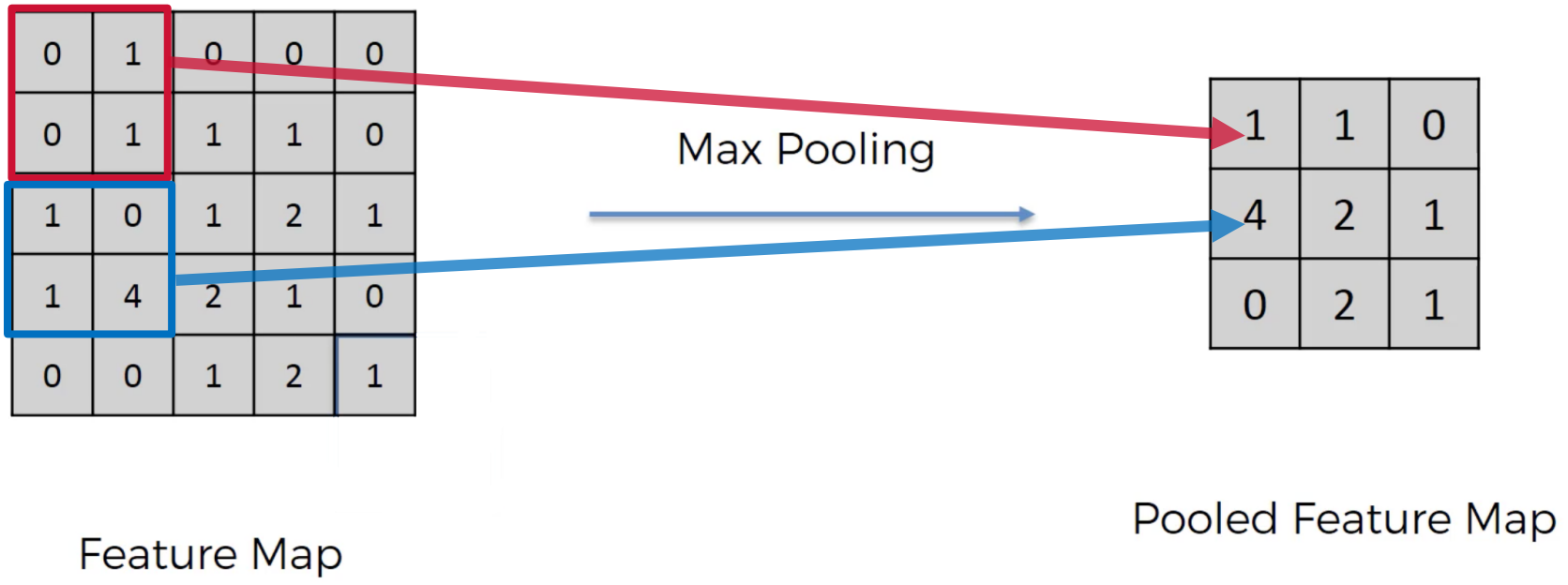




Source: https://reurl.cc/ObLLzR

Pooling 池化

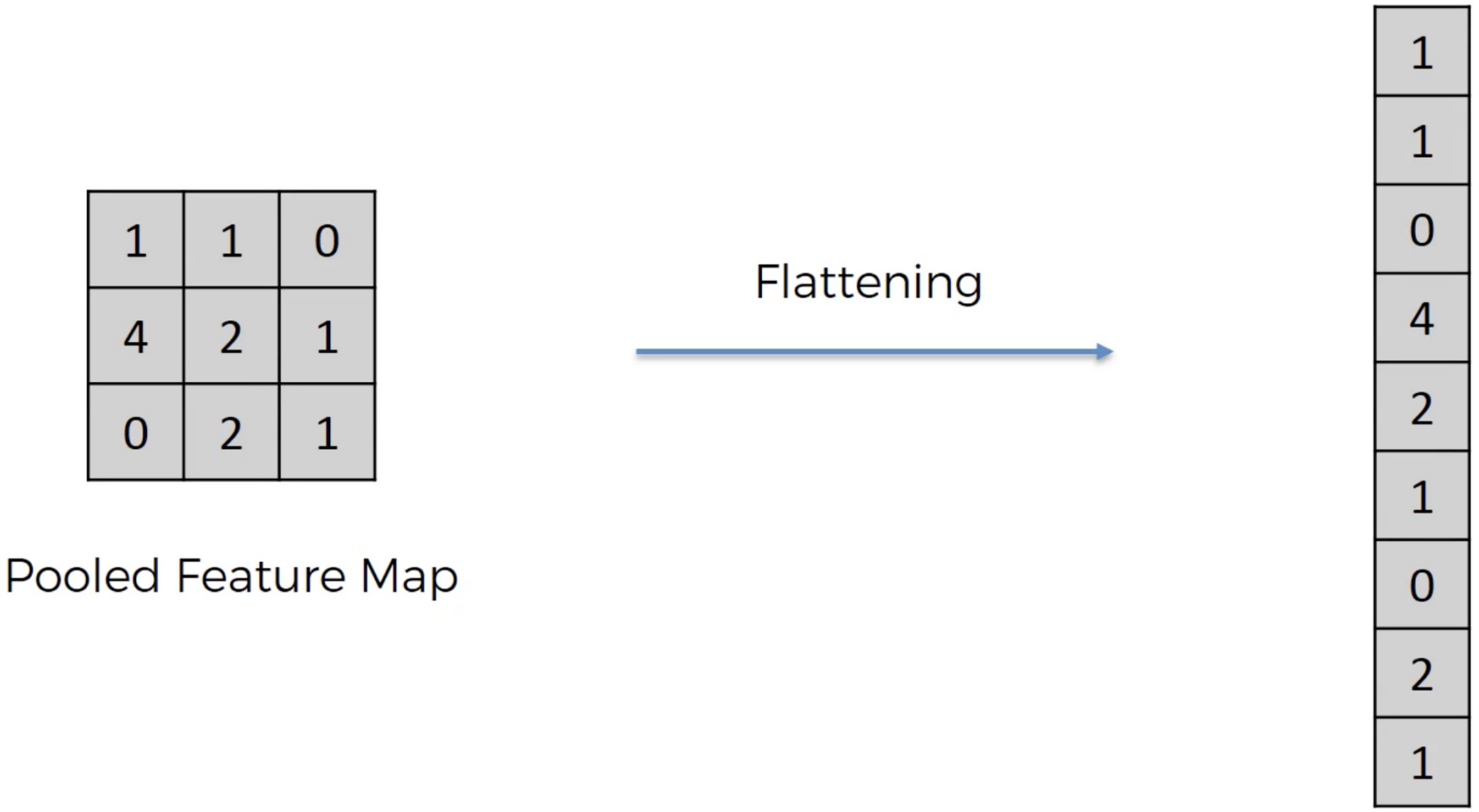
- 這邊主要是採用Max Pooling
- 概念很簡單只要挑出矩陣中的最大值就好
- 其目的是降低維度，並且保留重要特徵



Source: https://reurl.cc/ObLLzR

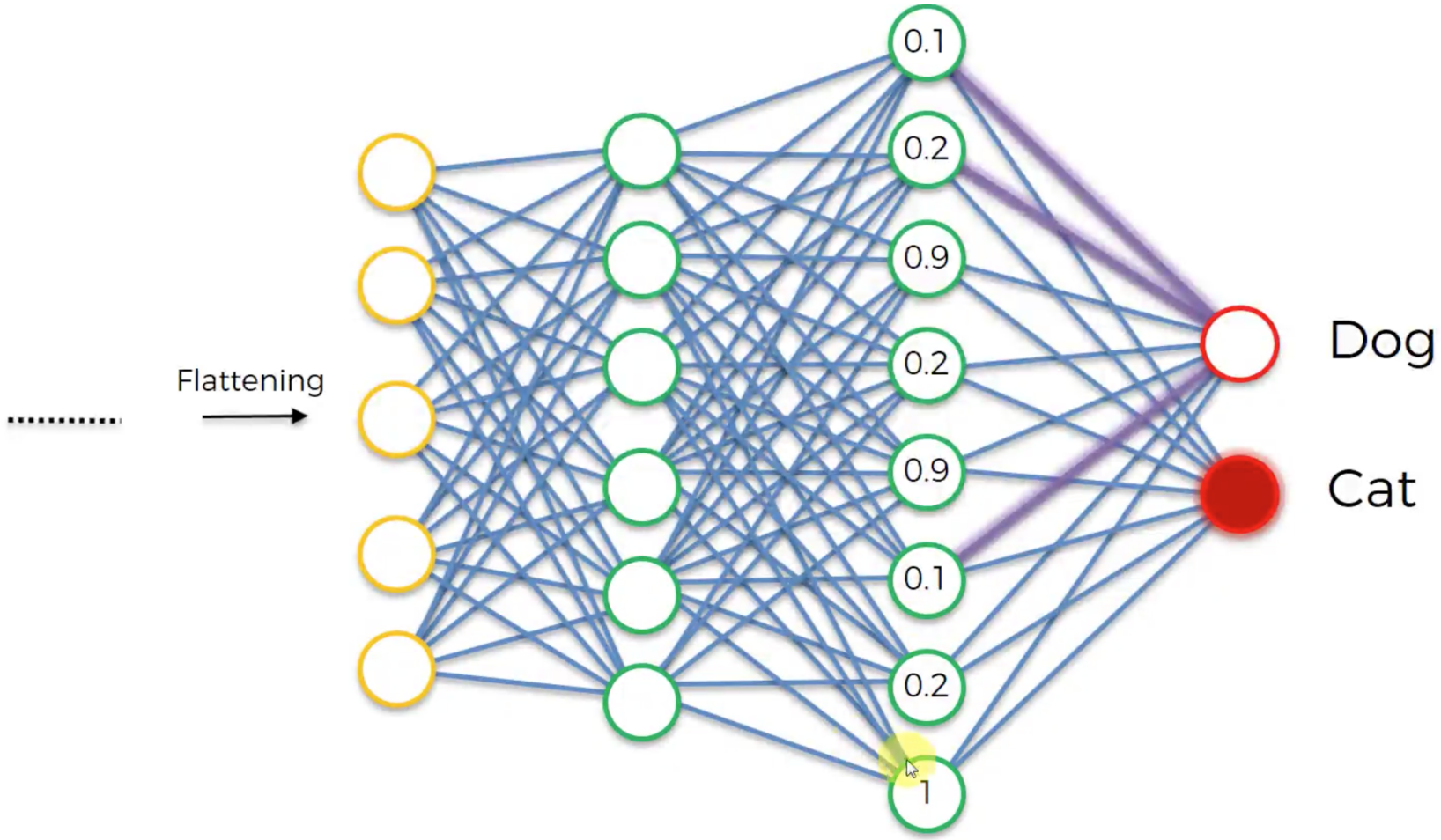
Fully Connected Layer 全連接層

- 全連接層是將前面運算的結果攤平後接到神經網絡



Source: https://reurl.cc/ObLLzR

Fully Connected Layer 全連接層



Source: https://reurl.cc/ObLLzR

IMAGENET Large Scale Visual Recognition Challenge

Year 2010

NEC-UIUC

Dense descriptor grid: HOG, LBP

Coding: local coordinate, super-vector

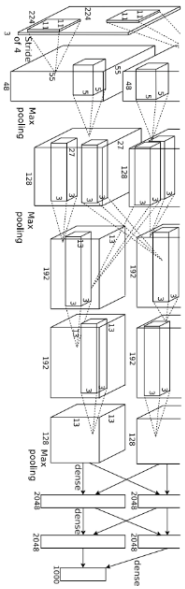
Pooling, SPM

Linear SVM

[Lin CVPR 2011]

Year 2012

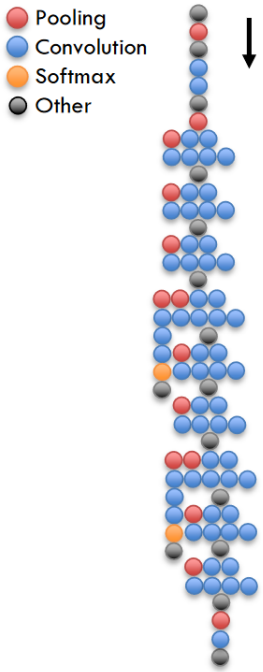
SuperVision



[Krizhevsky NIPS 2012]

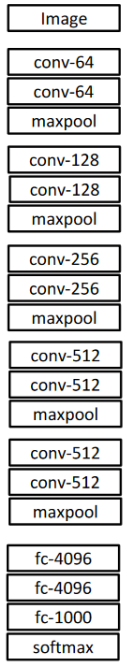
Year 2014

GoogLeNet



[Szegedy arxiv 2014]

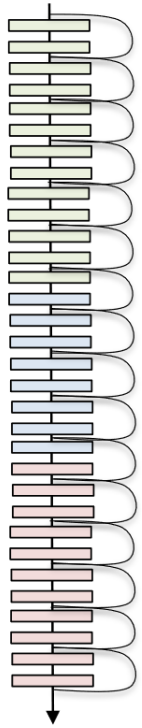
VGG



[Simonyan arxiv 2014]

Year 2015

MSRA



[He ICCV 2015]