Deep Learning for Image Data Processing

Instructor: Prof. Yi Fang

yfang@nyu.edu

Python tutorial: http://learnpython.org/

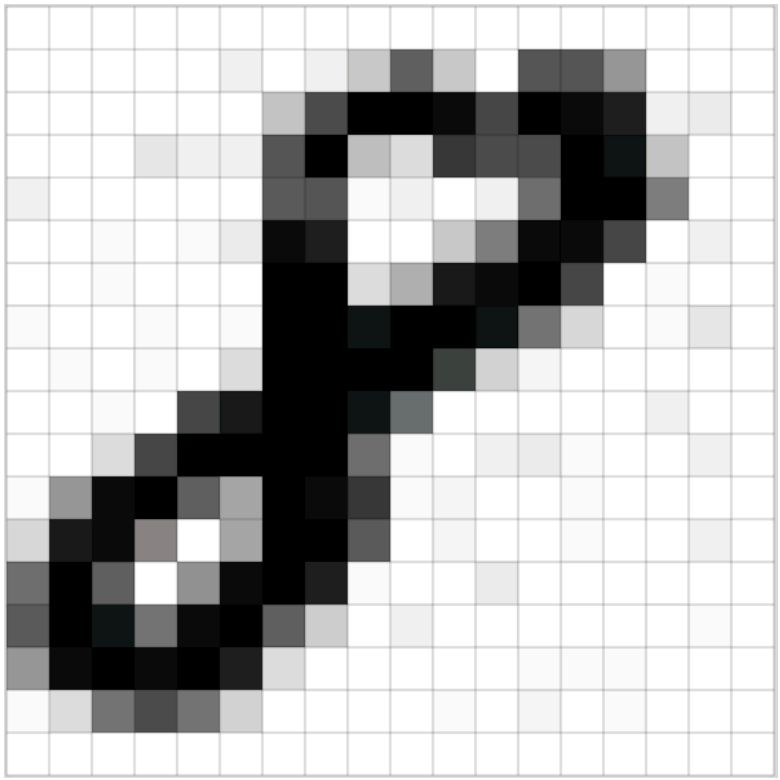
TensorFlow tutorial: https://www.tensorflow.org/tutorials/

PyTorch tutorial: https://pytorch.org/tutorials/

Acknowledge: The slides are partially referred to the online http:indoml.com (https://indoml.com/2018/03/07/student-notes-convolutional-neural-networks-cnn-introduction/) and Prof.

How Image looks like?

- An array of digital numbers.



0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	12	0	11	39	137	37	0	152	147	84	0	0
0	0	1	0	0	0	41	160	250	255	235	162	255	238	206	11	13
0	0	0	16	9	9	150	251	45	21	184	159	154	255	233	40	0
10	0	0	0	0	0	145	146	3	10	0	11	124	253	255	107	0
0	0	3	0	4	15	236	216	0	0	38	109	247	240	169	0	11
1	0	2	0	0	0	253	253	23	62	224	241	255	164	0	5	0
6	0	0	4	0	3	252	250	228	255	255	234	112	28	0	2	17
0	2	1	4	0	21	255	253	251	255	172	31	8	0	1	0	0
0	0	4	0	163	225	251	255	229	120	0	0	0	0	0	11	0
0	0	21	162	255	255	254	255	126	6	0	10	14	6	0	0	9
3	79	242	255	141	66	255	245	189	7	8	0	0	5	0	0	0
26	221	237	98	0	67	251	255	144	0	8	0	0	7	0	0	11
125	255	141	0	87	244	255	208	3	0	0	13	0	1	0	1	0
145	248	228	116	235	255	141	34	0	11	0	1	0	0	0	1	3
85	237	253	246	255	210	21	1	0	1	0	0	6	2	4	0	0
6	23	112	157	114	32	0	0	0	0	2	0	8	0	7	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

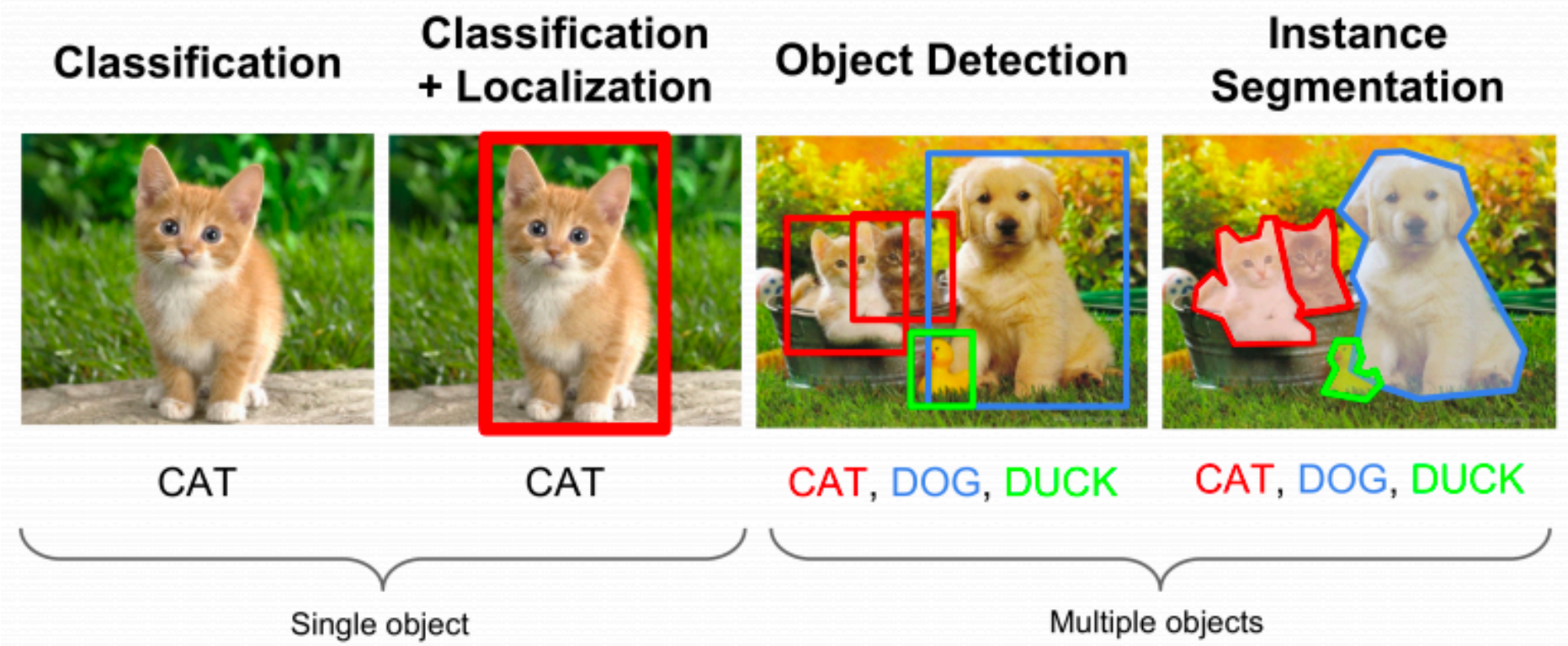
Image Processing

- What is image processing?
  - Human vision - perceive and understand world
  - Computer vision, Image Understanding / Interpretation, Image processing.
    - Real world environment -> sensors (Lidar or cameras) -> 2D/3D images -> Image analysis

Applications

Src: https://www.picswe.com/pics/vision-applications-32.html

Image Processing Tasks



Src: Stanford course cs224d course):

Image Processing Pipeline

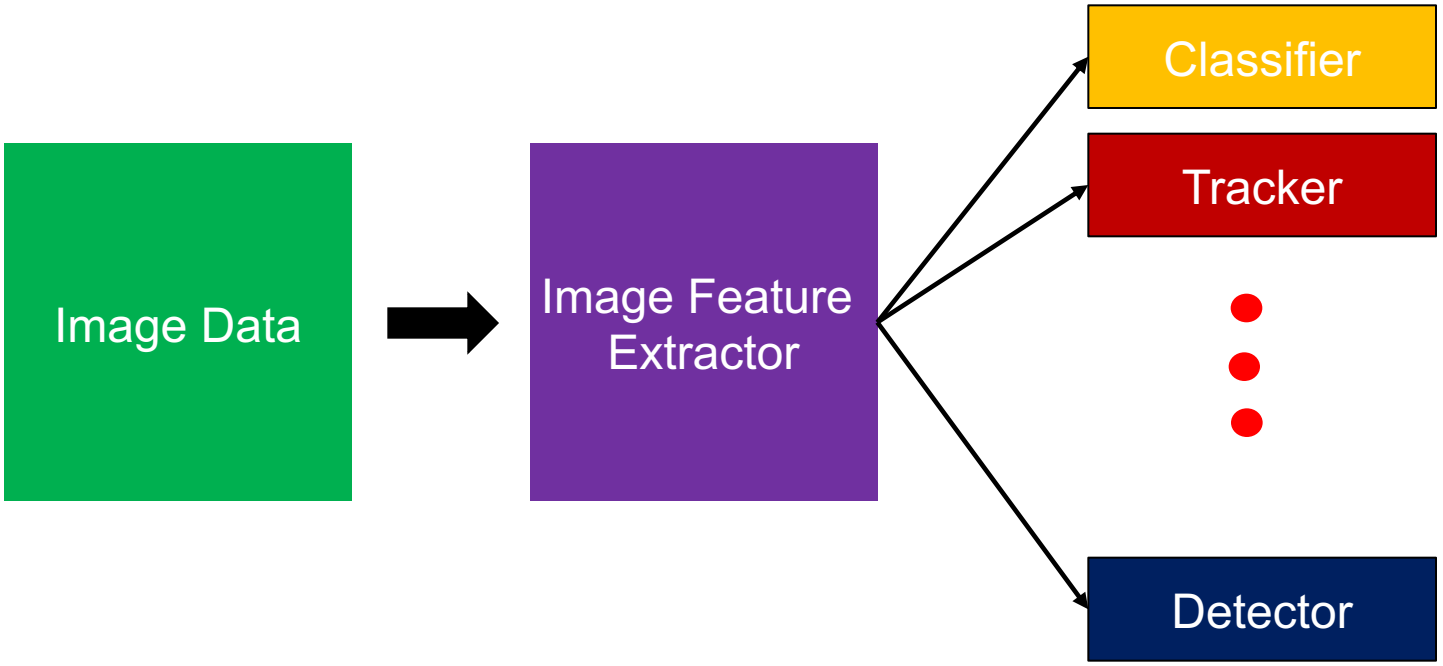
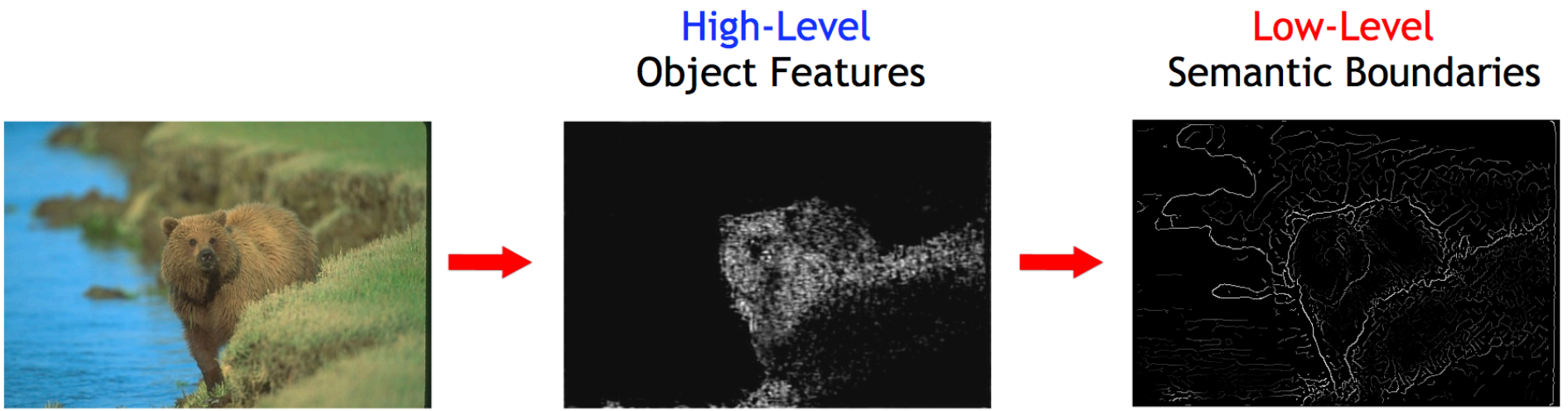


Image Feature

- Low-Level Image Feature
- High-Level Image Feature

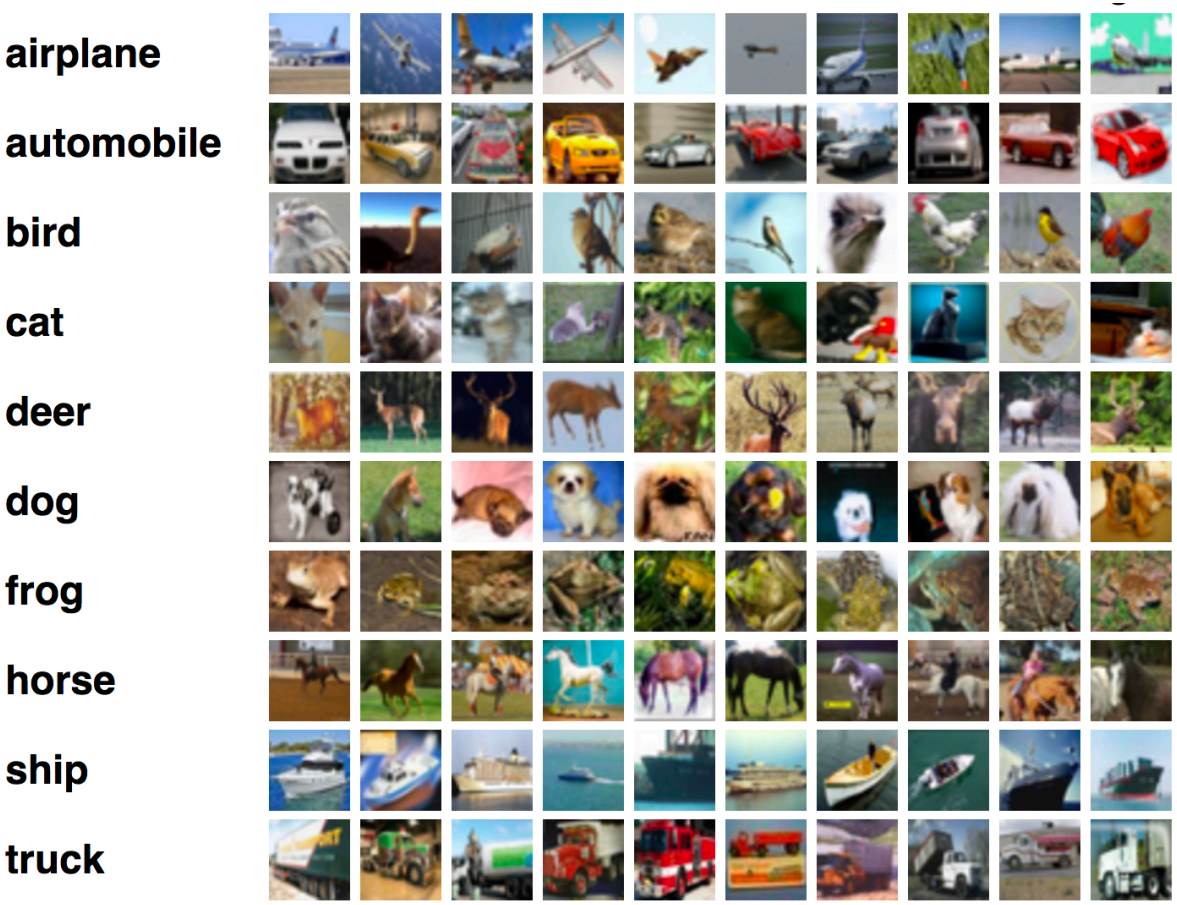


Src: https://www.seas.upenn.edu/~jshi/

High level image processing

- **Image Classification**
- **Image Captioning**
- **Image retrieval**
- **Global Image Feature Descriptor**

Image Classification



Src: https://medium.com Link

Image Captioning

"man in black shirt is playing guitar."

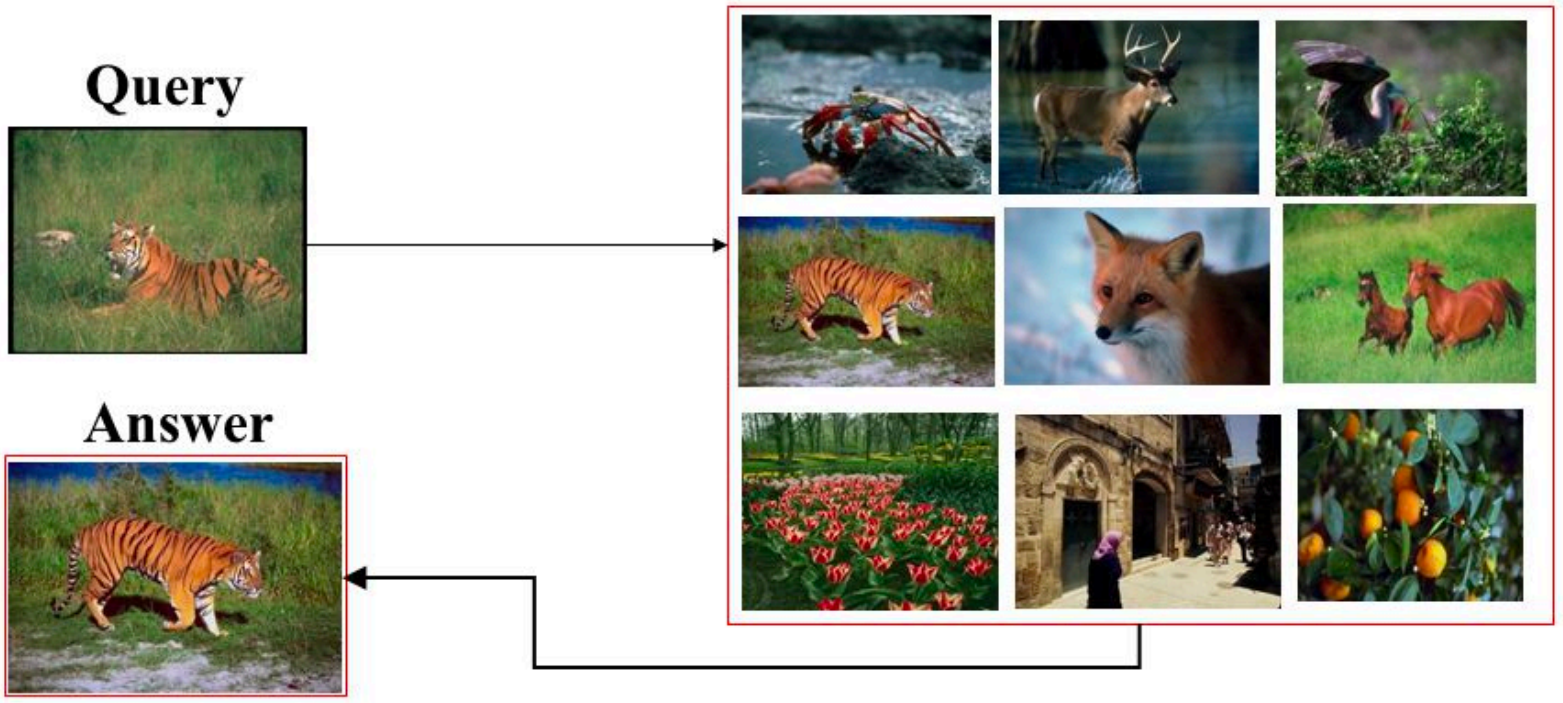
"construction worker in orange safety vest is working on road."

"girl in pink dress is jumping in air."

"black and white dog jumps over bar."

Src: https://daniel.lasiman.com/post/image-captioning/

Image Retrieval

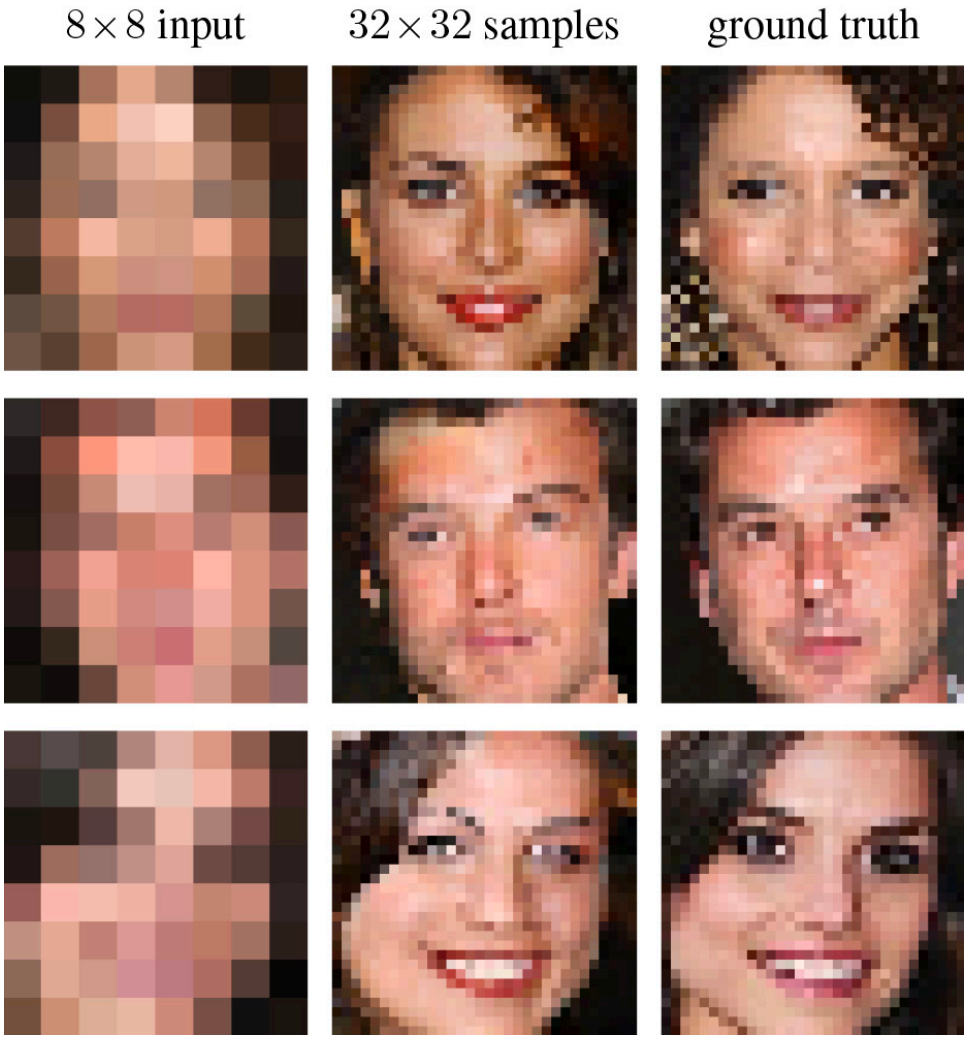


Src: https://www.analyticsvidhya.com/blog/2017/11/information-retrieval-using-kdtree/

Low level image processing

- **Pre-processing**
  - suppresses noise (image pre-processing)
  - enhances some object features - relevant to understanding the image
  - edge extraction, smoothing, thresholding etc.
- **Image segmentation**
  - separate objects from the image background
  - colour segmentation, region growing, edge linking etc
- **Image pixel level feature description**

Image Super-resolution

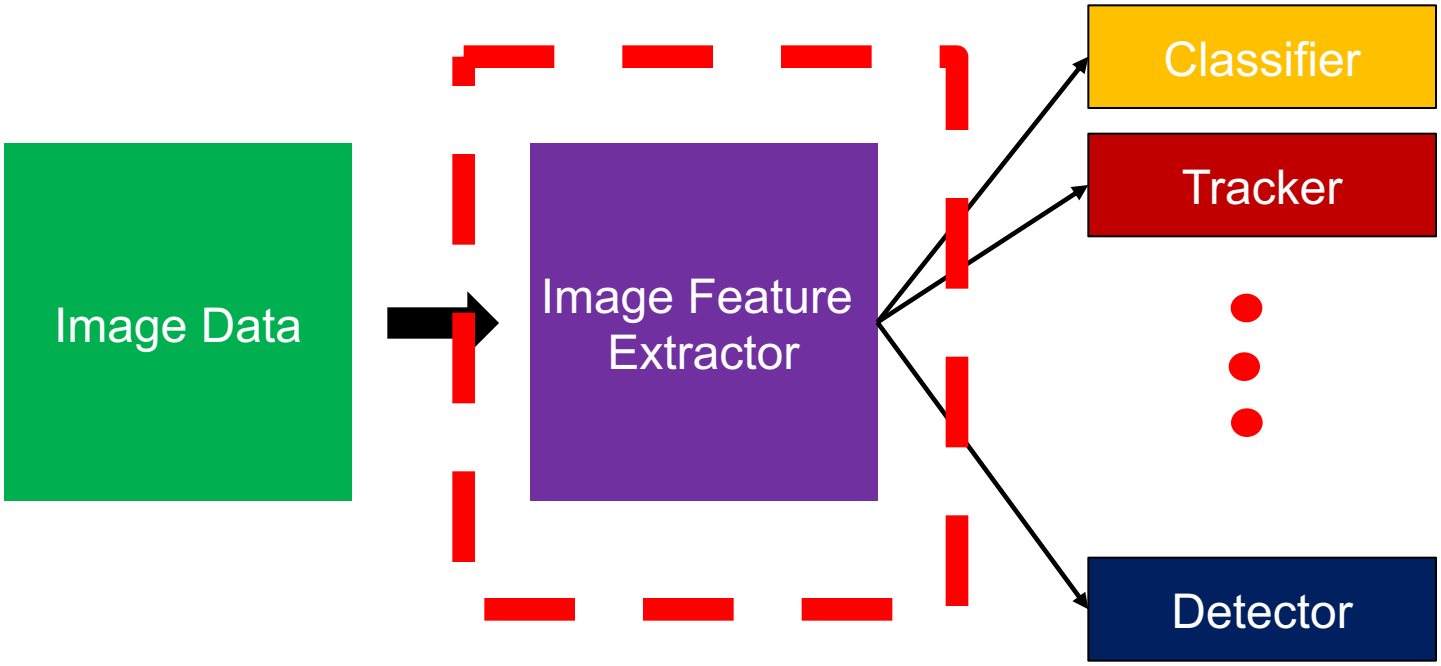


Src: Google Brain

Image Segmentation

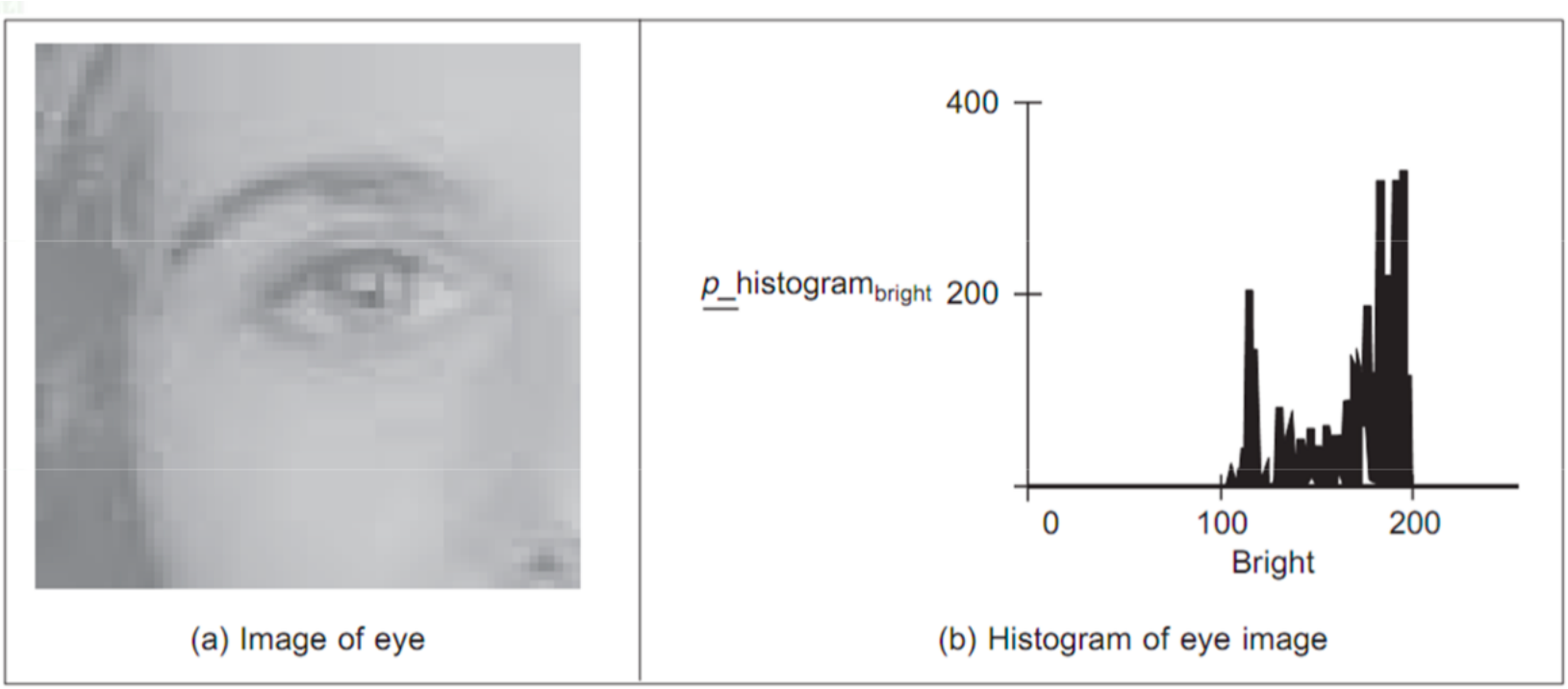
Src: Nvidia Link

Image Processing Pipeline



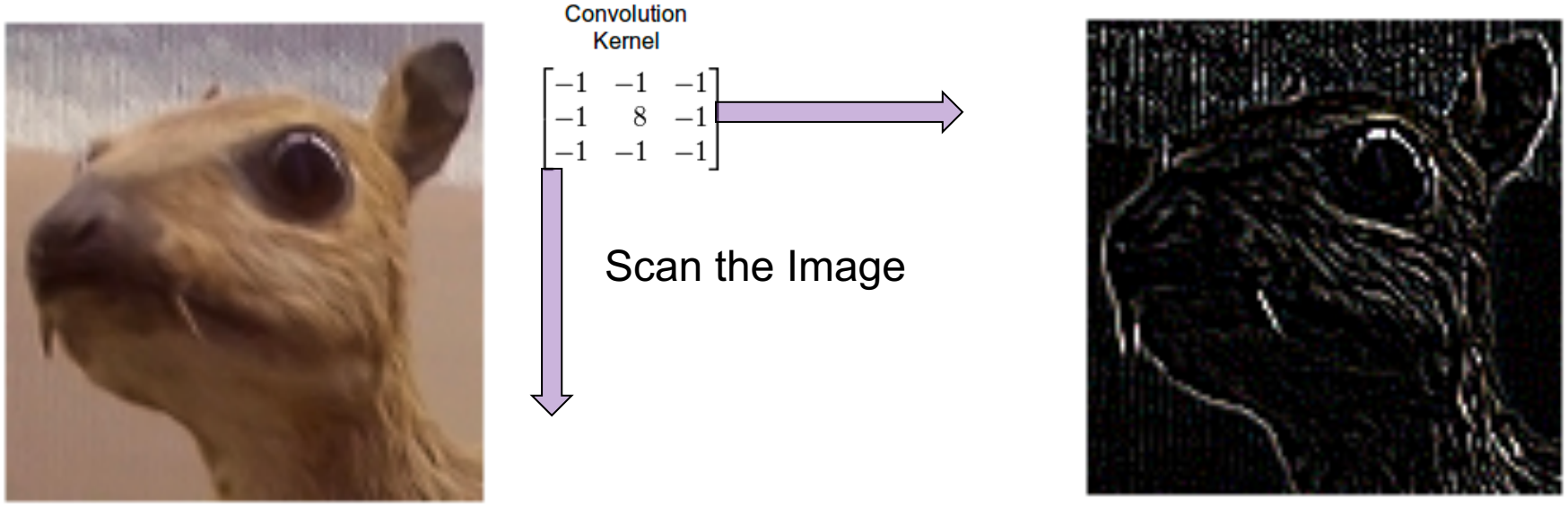
Global Feature Extractor

- Extractor: Statistics



Local Feature Extractor

- Extractor: Feature extraction by convolution



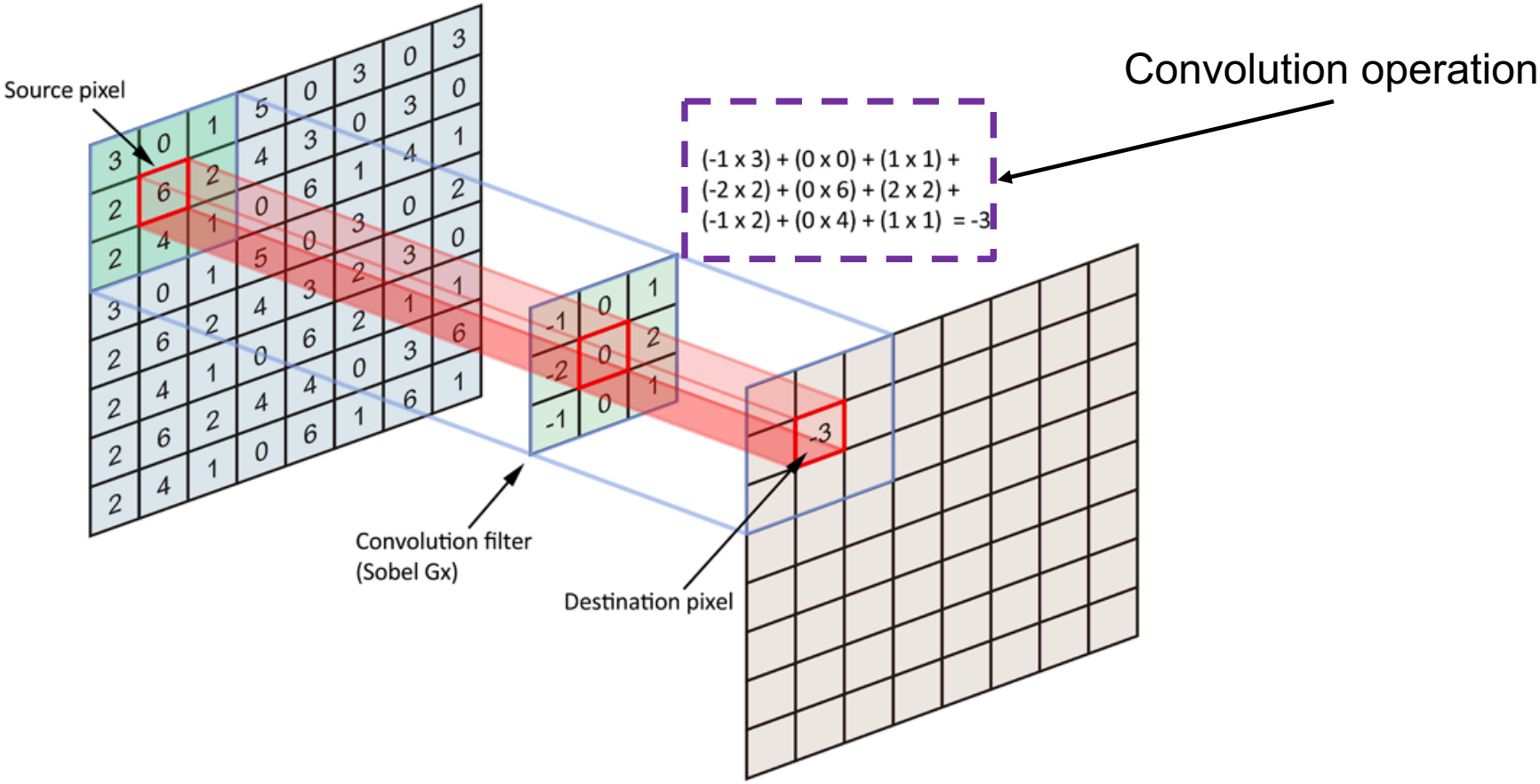
Convolution Math

The convolution of f and g is written f∗g, using an asterisk or star. It is defined as the integral of the product of the two functions after one is reversed and shifted. As such, it is a particular kind of integral transform:

$$(f * g)(t) \triangleq \int_{-\infty}^{\infty} f(\tau) g(t - \tau)\, d\tau.$$

Src: Wikipedia

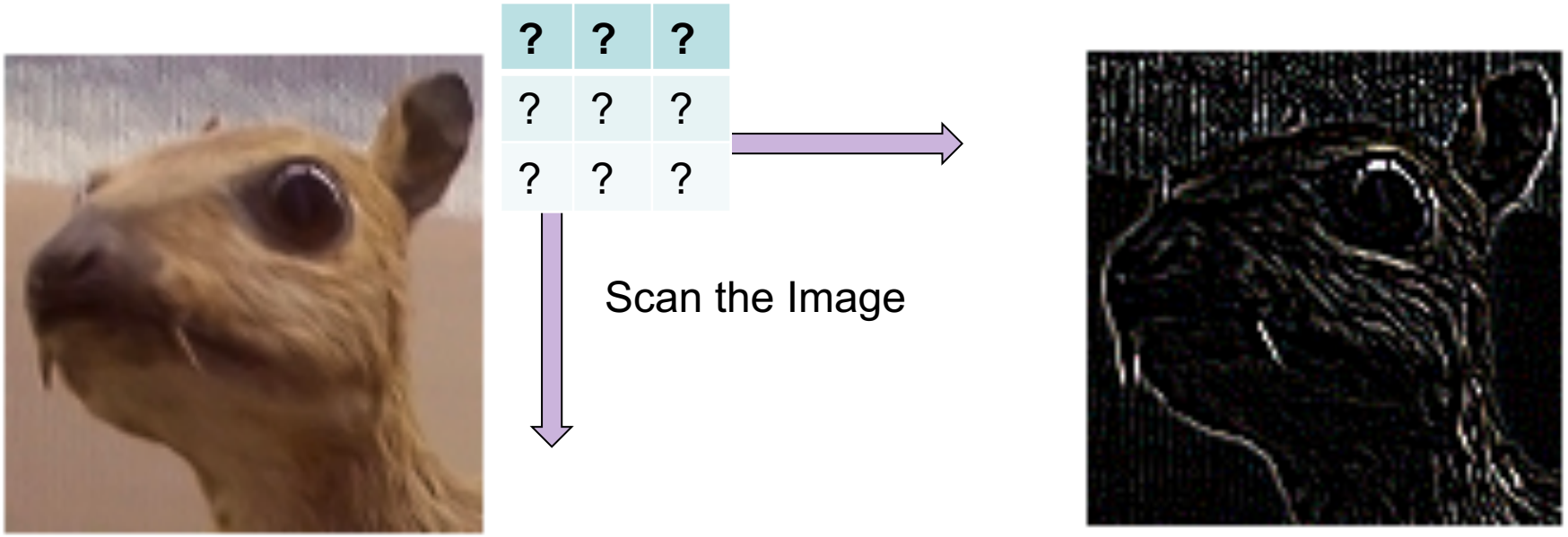
Convolution in Image



Src: Link

The weight in Convolution

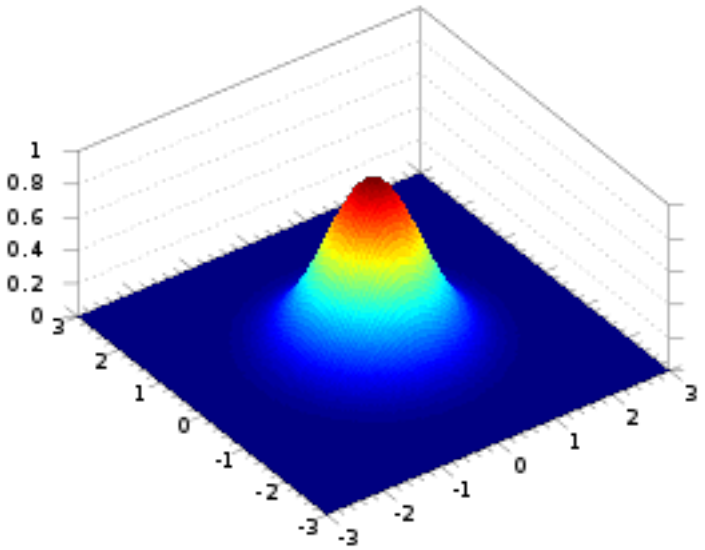
How to determine the weights?



- Hand craft weights
- Learned weights

Hand Craft Operators

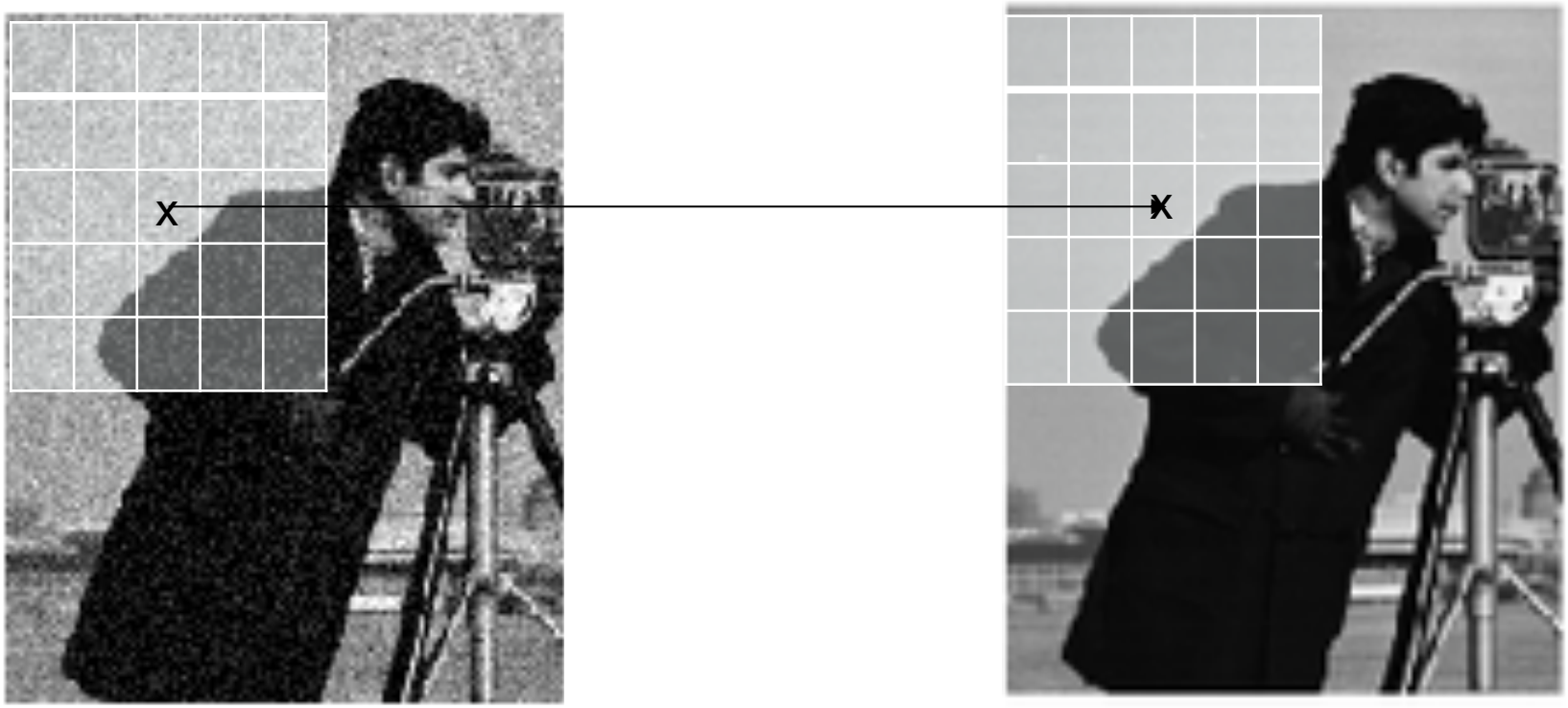
- Manually defined weights



$$\frac{1}{273}\begin{bmatrix} 1 & 4 & 7 & 4 & 1 \\ 4 & 16 & 26 & 16 & 4 \\ 7 & 26 & 41 & 26 & 7 \\ 4 & 16 & 26 & 16 & 4 \\ 1 & 4 & 7 & 4 & 1 \end{bmatrix}$$

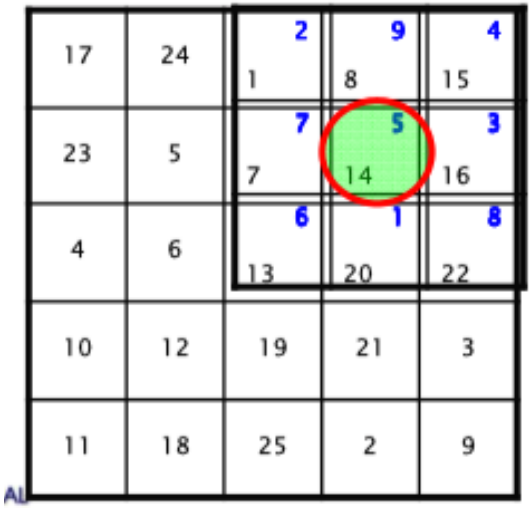
Gaussian Template

Multiplication of a Template



Convolution Operator on Image

2-D convolution

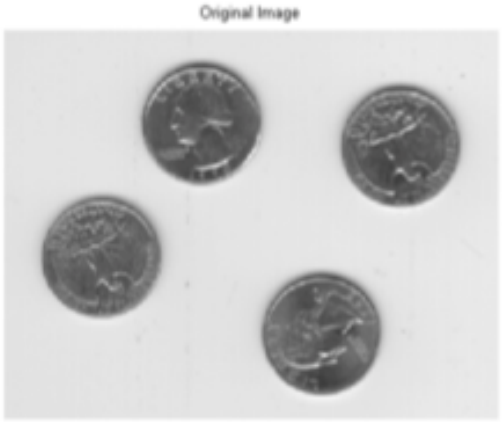


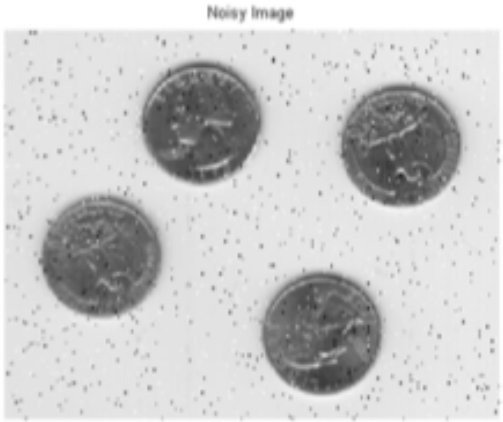
$$1x2+8x9+15x4+7x7+14x5+16x3+13x6+20x1+22x8=575$$

Denoising Operator

- Uniform averaging

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

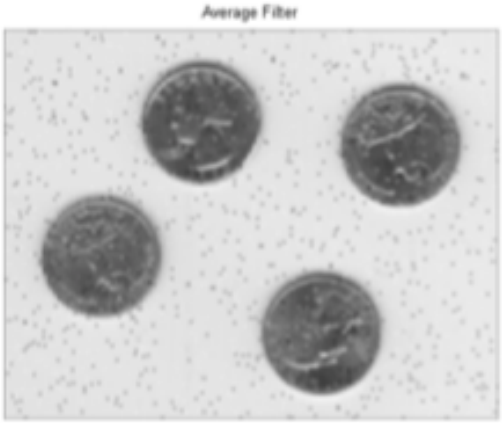


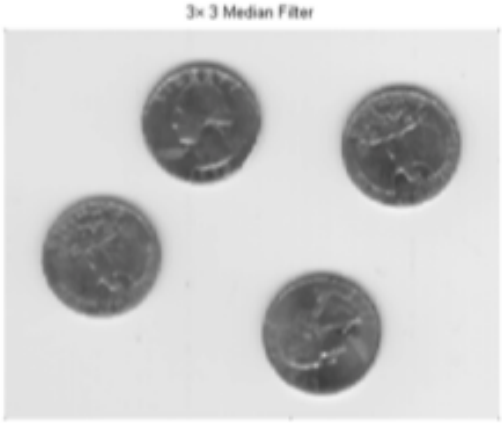


- Median filter

20	5	43
78	3	22
115	189	200

43

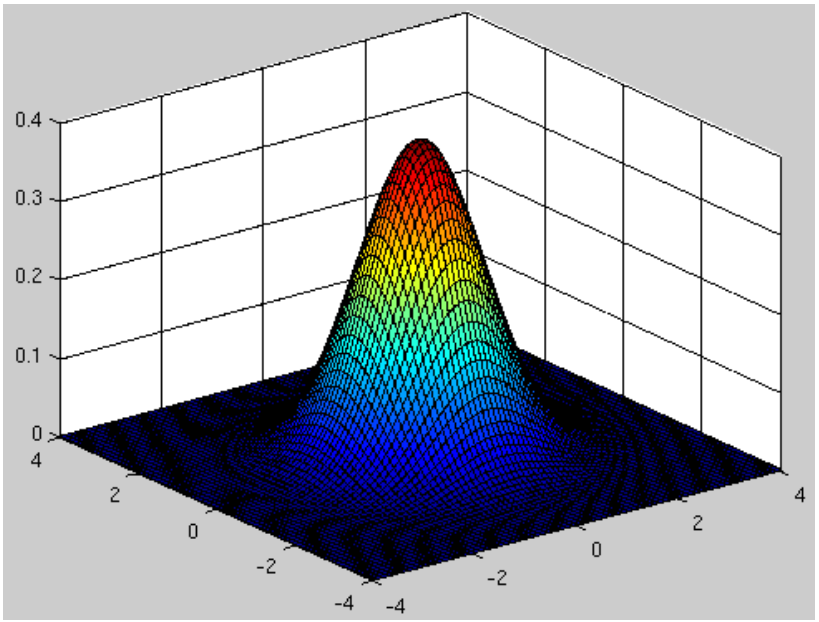




Gaussian Operator

Math:

$$G(x,y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp^{\frac{-(x^2+y^2)}{2\sigma^2}}$$



Gaussian with $\sigma = 1$

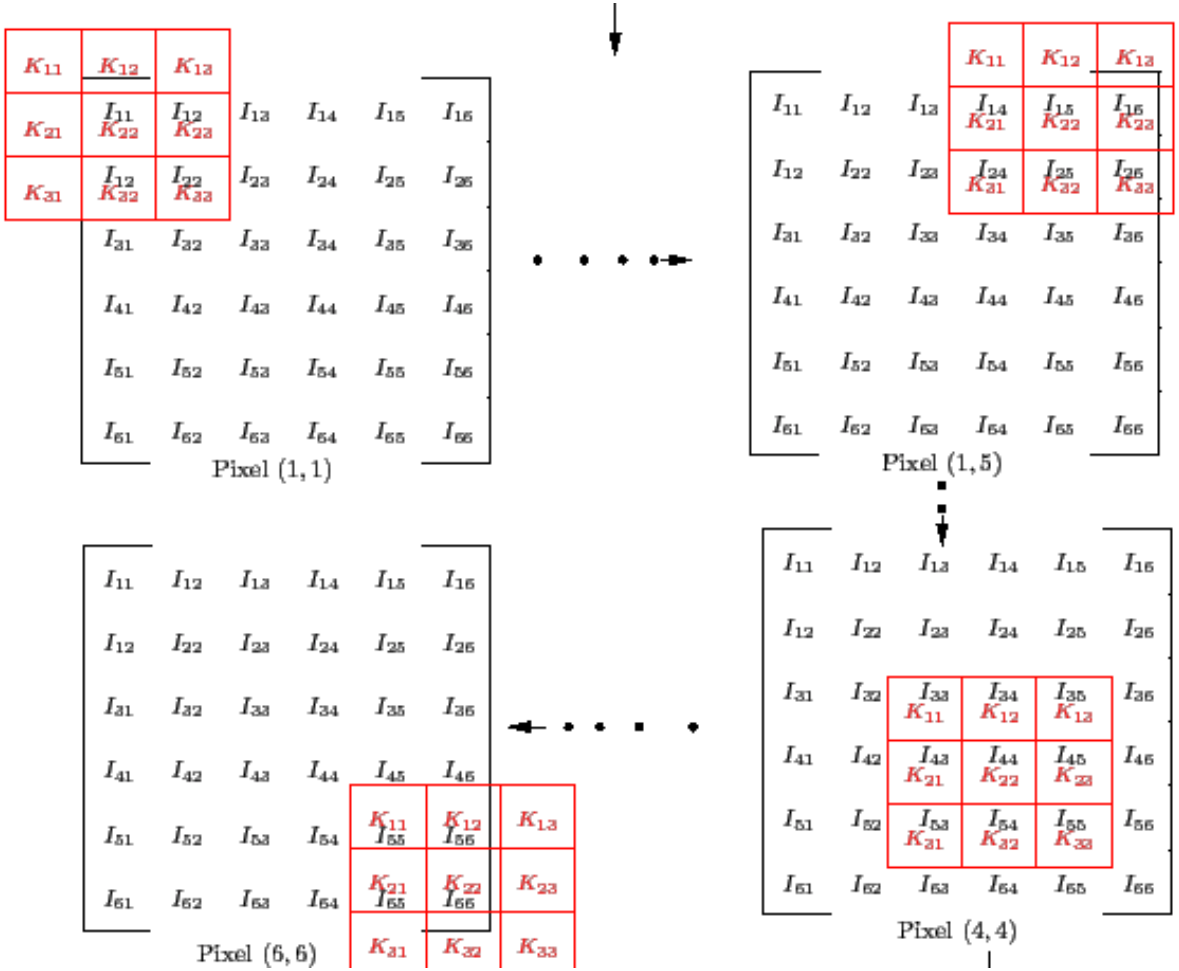
Gaussian Template

0.003	0.0133	0.0219	0.0133	0.003
0.0133	0.0596	0.0983	0.0596	0.0133
0.0219	0.0983	0.1621	0.0983	0.0219
0.0133	0.0596	0.0983	0.0596	0.0133
0.003	0.0133	0.0219	0.0133	0.003

Gaussian with $\sigma = 1$

Src: http://ugweb.cs.ualberta.ca/~vis/courses/cde/assign/diffusion/

Convolution steps



Convolved Image

$$\begin{bmatrix} I^*_{11} & I^*_{12} & I^*_{13} & I^*_{14} & I^*_{15} & I^*_{16} \\ I^*_{12} & I^*_{22} & I^*_{23} & I^*_{24} & I^*_{25} & I^*_{26} \\ I^*_{31} & I^*_{32} & I^*_{33} & I^*_{34} & I^*_{35} & I^*_{36} \\ I^*_{41} & I^*_{42} & I^*_{43} & I^*_{44} & I^*_{45} & I^*_{46} \\ I^*_{51} & I^*_{52} & I^*_{53} & I^*_{54} & I^*_{55} & I^*_{56} \\ I^*_{61} & I^*_{62} & I^*_{63} & I^*_{64} & I^*_{65} & I^*_{66} \end{bmatrix}$$

Examples of Gaussian Averaging

(a) 3 × 3

(a) 5 × 5

(a) 7 × 7

- The noise is reduced by convolving the intensity image with a gaussian kernel (linear filtering)
- The amount of filtering can be controlled by changing the coefficients of the convolution mask

Noisy image (Gaussian noise $\sigma = 0.01$)

Filtering

Filtered image (5x5 Gaussian filter $\sigma = 1$)

Edge Detector

Edge detection includes a variety of mathematical methods that aim at identifying points in a digital image at which the image brightness changes sharply or, more formally, has discontinuities.

Src: Wikipedia

Prewitt Operator

$$\mathbf{G_x} = \begin{bmatrix} +1 & 0 & -1 \\ +1 & 0 & -1 \\ +1 & 0 & -1 \end{bmatrix} * \mathbf{A} \quad \text{and} \quad \mathbf{G_y} = \begin{bmatrix} +1 & +1 & +1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} * \mathbf{A}$$

where $*$ here denotes the 1-dimensional convolution operation.

Result

Grayscale image of a brick wall & a bike rack

Gradient with Prewitt operator of grayscale image of a brick wall & a bike rack

Sobel Operator

$$\mathbf{G}_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * \mathbf{A} \quad \text{and} \quad \mathbf{G}_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * \mathbf{A}$$

where $*$ here denotes the 2-dimensional signal processing convolution operation.

Result

Grayscale test image of brick wall and bike rack

Normalized gradient magnitude from Sobel–Feldman operator

Many More …

- Laplace operator
- Roberts Cross
- Gabor filter

Filter Bank

- Gaussian pyramid: A multi-resolution representation of an image formed by several images, each one a subsampled and Gaussian smoothed version of the original one at increasing standard deviation.

Src: http://homepages.inf.ed.ac.uk/rbf/CVDICT/cvg.htm

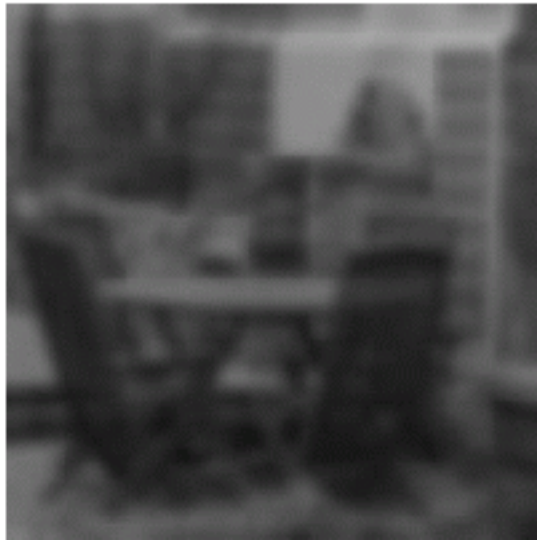
Gaussian image pyramid

Gaussian Smoothed Images

Original Image

sigma = 1.0

sigma = 3.0



Scale-invariant Feature

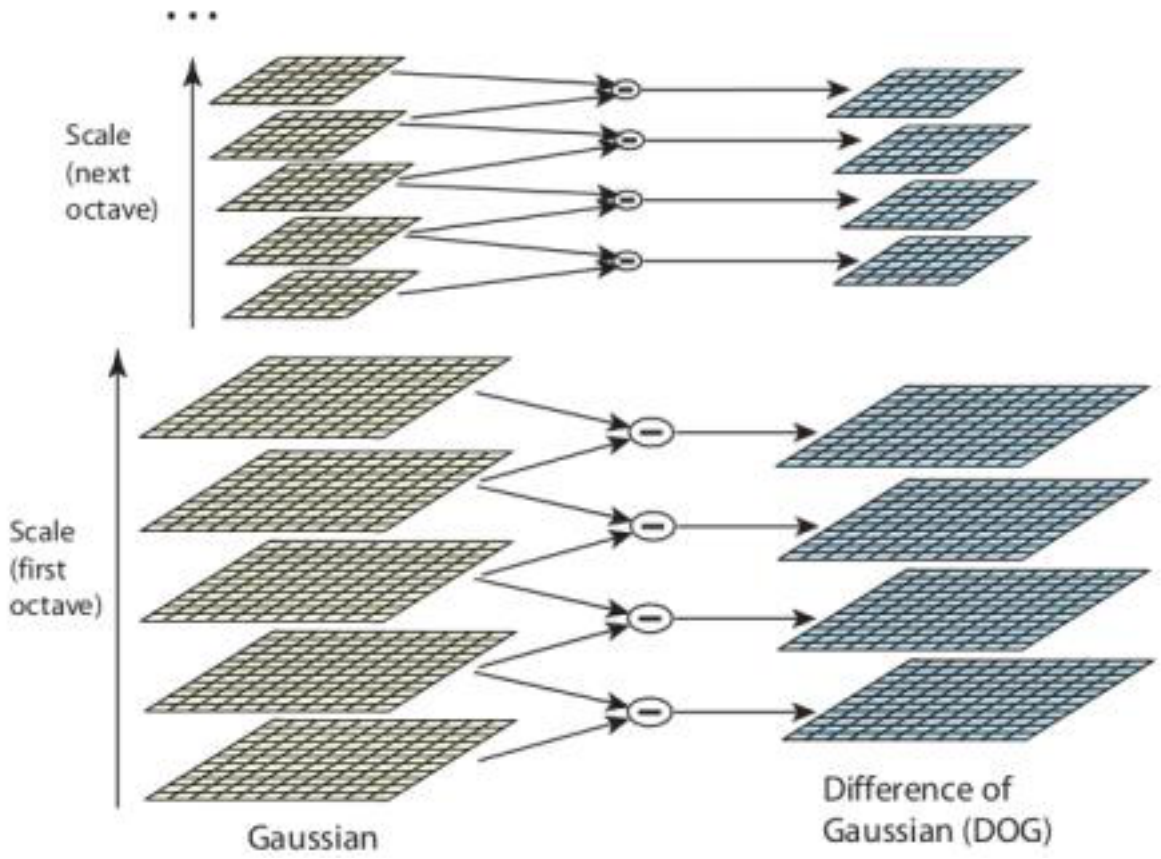
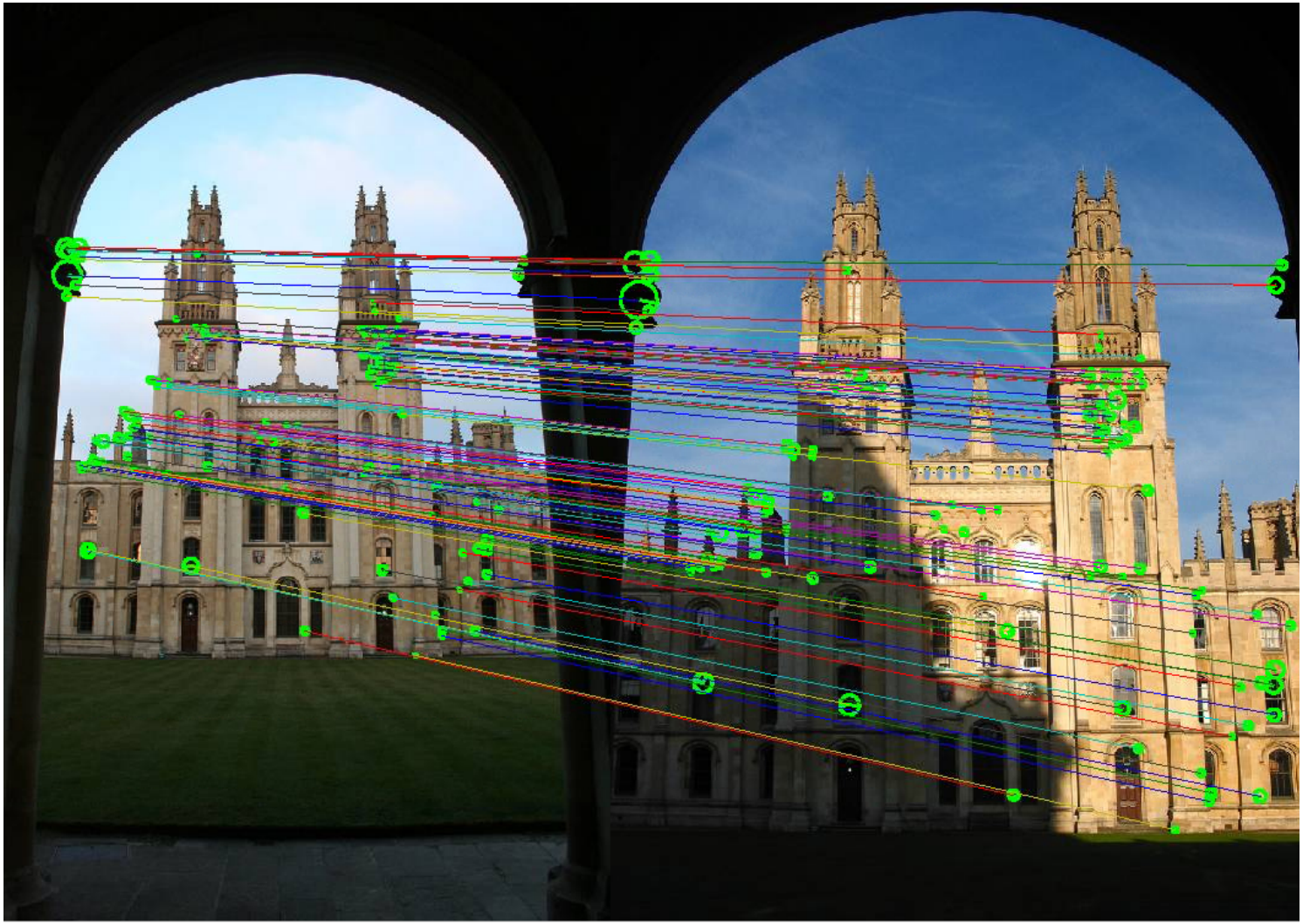
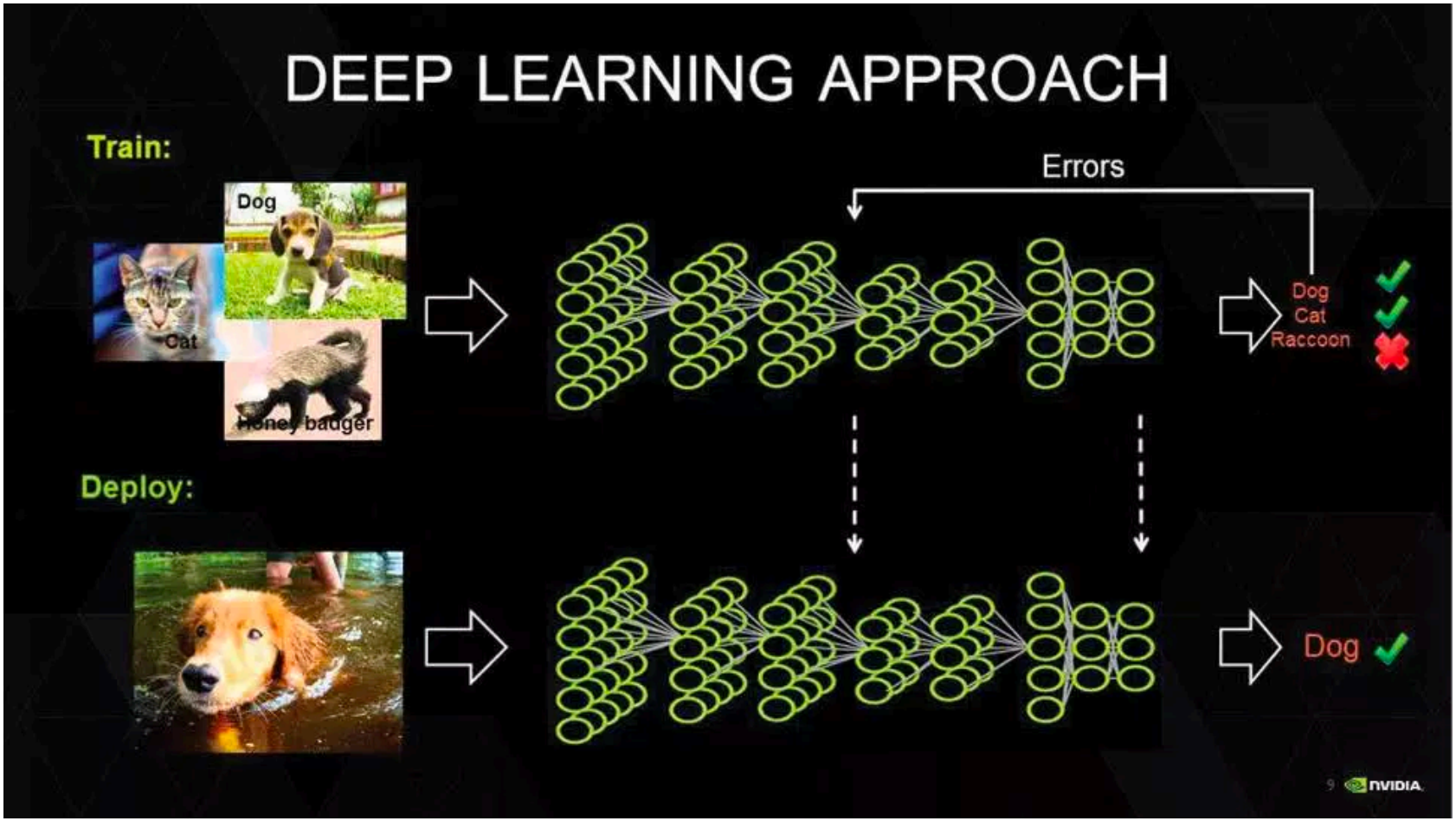


Image Matching

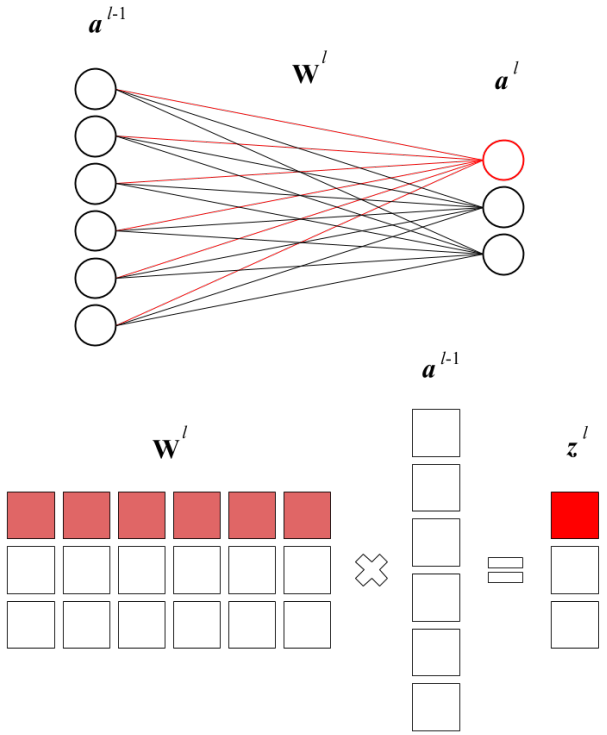


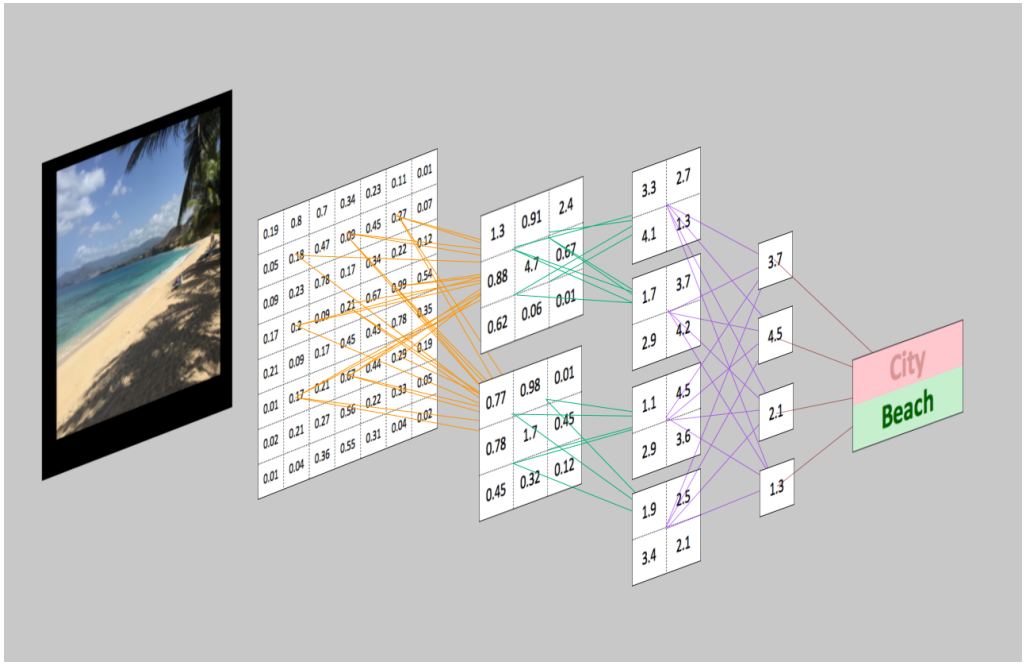
Src: Link

Learned Operators

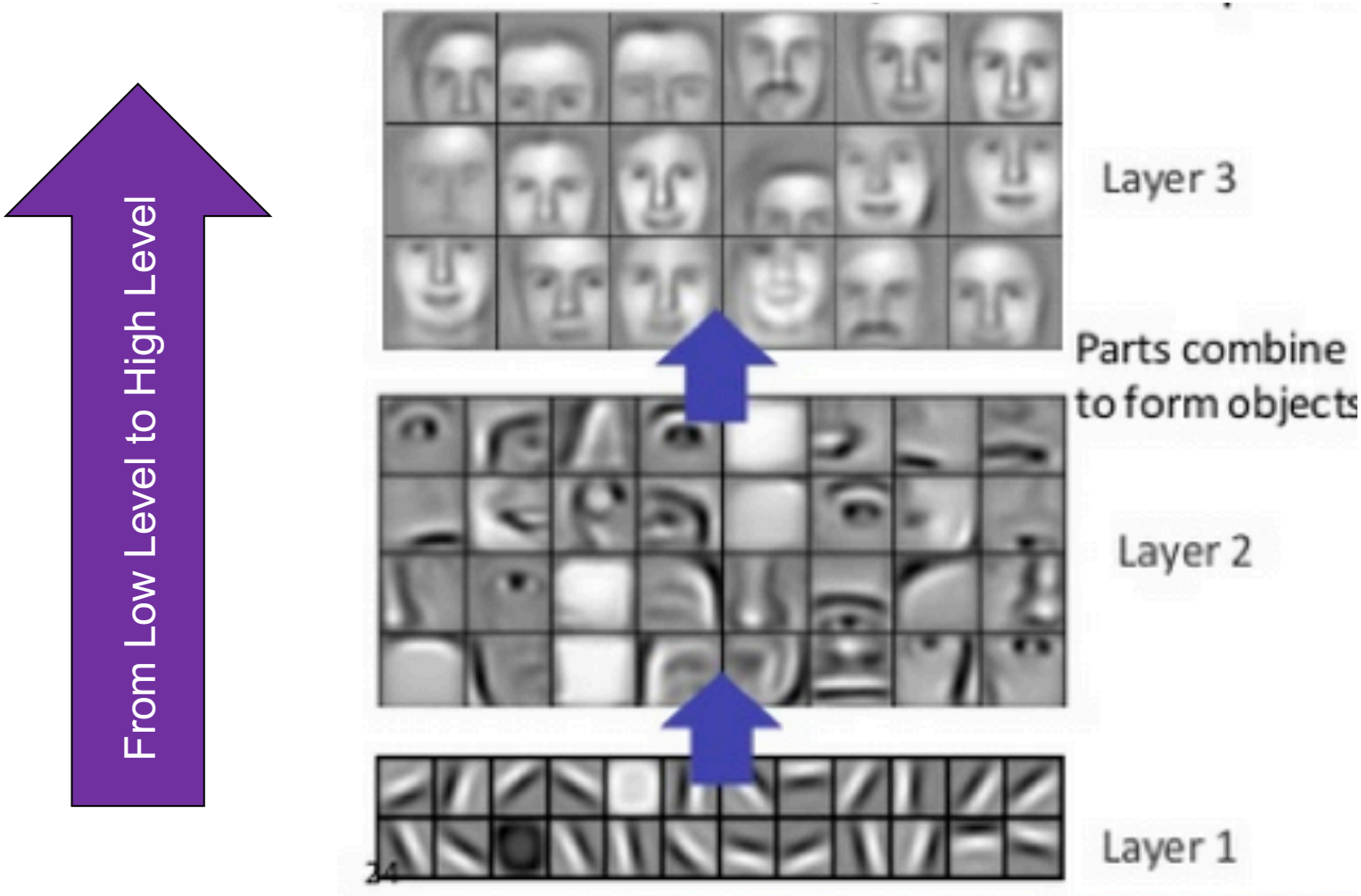


Learned Operators

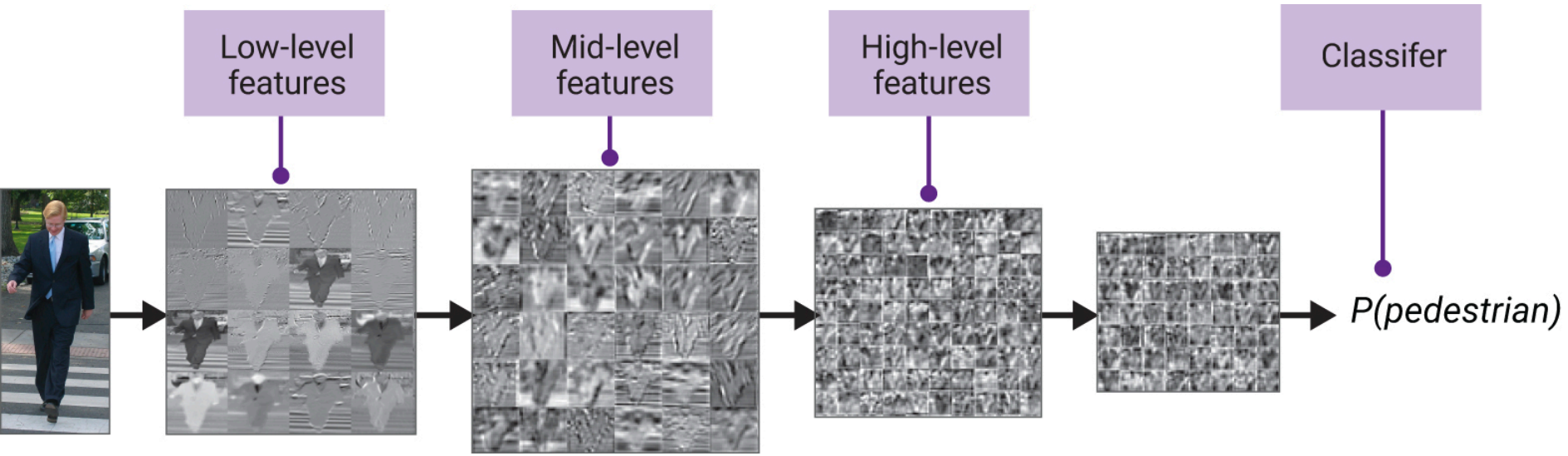




Src: https://www.oreilly.com/library/view/deep-learning/9781491924570/ch04.html
https://medium.com/@erikhallstrm/backpropagation-from-the-beginning-77356edf427d

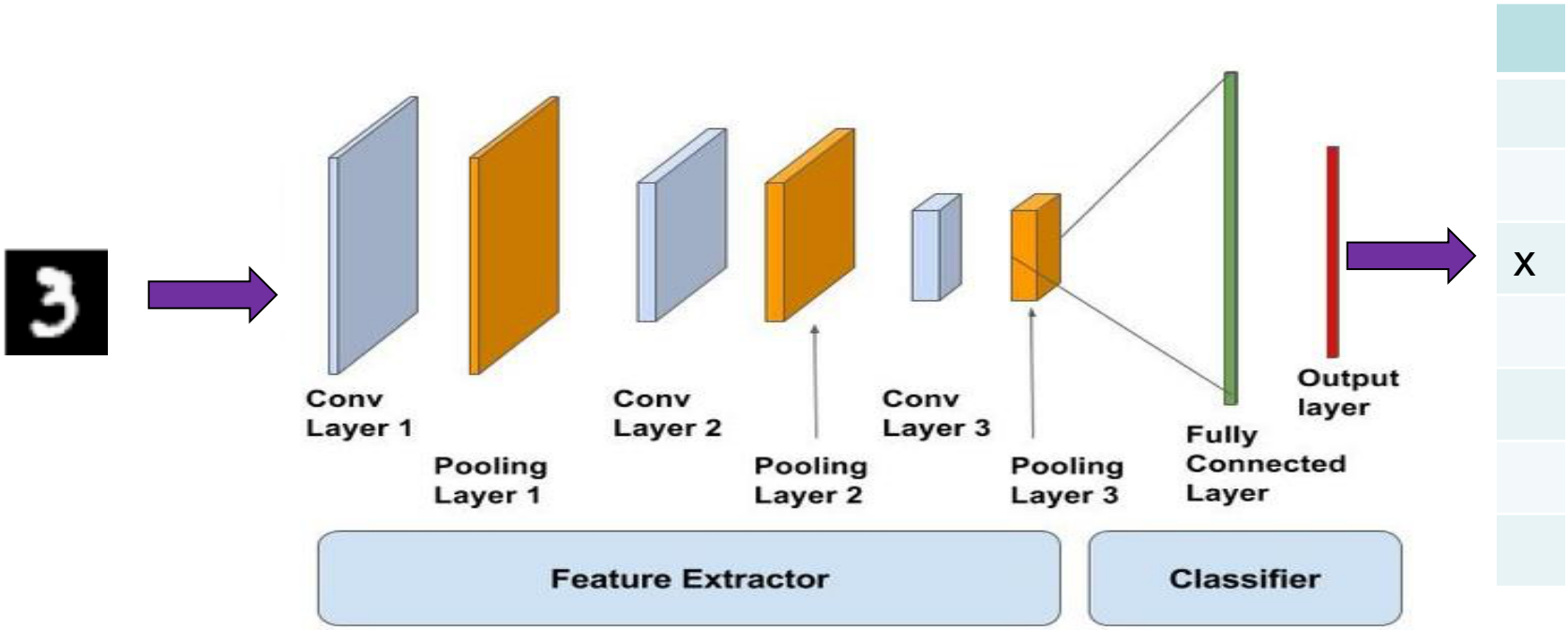


Src: https://skymind.ai/wiki/neural-network

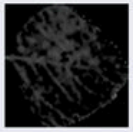
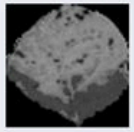
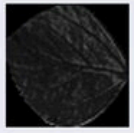
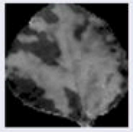
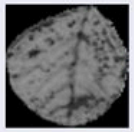
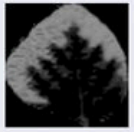
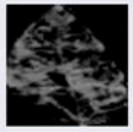
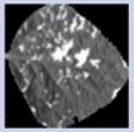

Low-level features
Mid-level features
High-level features
Classifer
P(pedestrian)

MNIST Dataset Classification

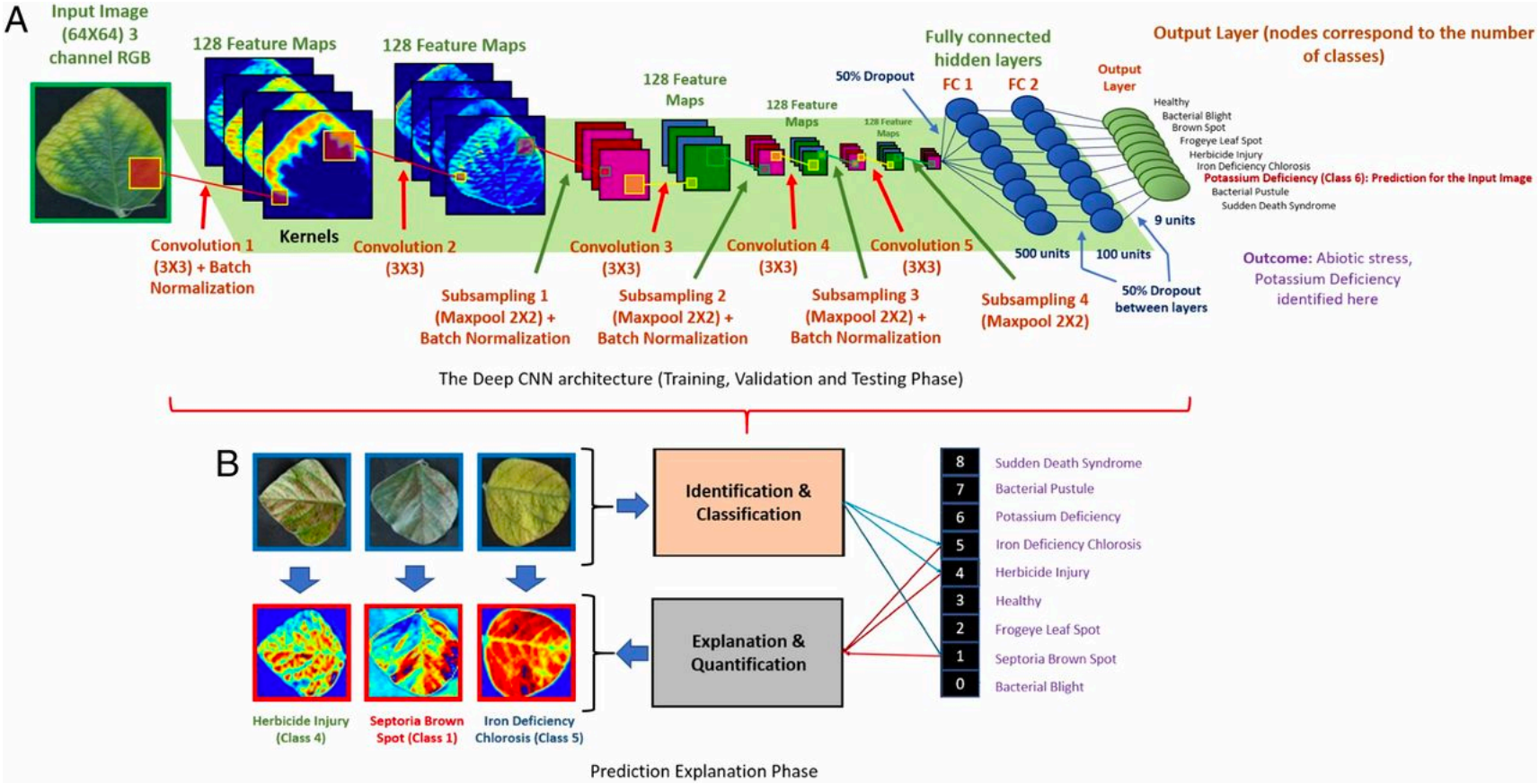
Network Structure



Example

Soybean Leaves	Disease Names and Labels	Bacterial Blight (Class 0)	Septoria Brown Spot (Class 1)	Frogeye Leaf Spot (Class 2)	Healthy (Class 3)	Herbicide Injury (Class 4)	Iron Deficiency Chlorosis (Class 5)	Potassium Deficiency (Class 6)	Bacterial Pustule (Class 7)	Sudden Death Syndrome (Class 8)
	Original Image									
	Prediction Explanation Output									
	Rater marked Image									

Src:https://www.pnas.org/content/115/18/4613



Src: https://www.pnas.org/content/115/18/4613