AMRITA
VISHWA VIDYAPEETHAM
DEEMED TO BE UNIVERSITY

श्रद्धावान् लभते ज्ञानम्

21AI637 Deep Learning 3-0-2-4 (S2 MTech AI)

21AM645 Deep Learning 3-0-2-4 (S2 M.Tech CS(AI/ML)

PhD Coursework

Class Timings

- Tuesday 11 am to 11 : 40 am
- Wednesday 11 am to 11 : 40 am
- Thursday 8:50 am to 9:40 am
- Lab – Tuesday(3.00PM to 4:50PM)

.

Course Objectives

- To introduce to students,
  - Different deep neural network architectures,
  - Training strategies/algorithms, possible challenges, tools and techniques available in designing and deploying solutions to different practical/ Engineering problems.

Course Outcomes

- After the course completion, the students will be able to,
- **CO1:** Be able to design, train, deploy neural networks for solving different practical/engineering problems and analyse and report its efficacy
- **CO2** : Have a good level of knowledge (Both Conceptual and Mathematical) on different neural network settings to pursue Research in this Field
- **CO3**: Build skills in using established ML tools/libraries and in building self-learning skills in the field

- Prerequisites • Linear Algebra and Probability • Computational Methods for Optimisation

Syllabus

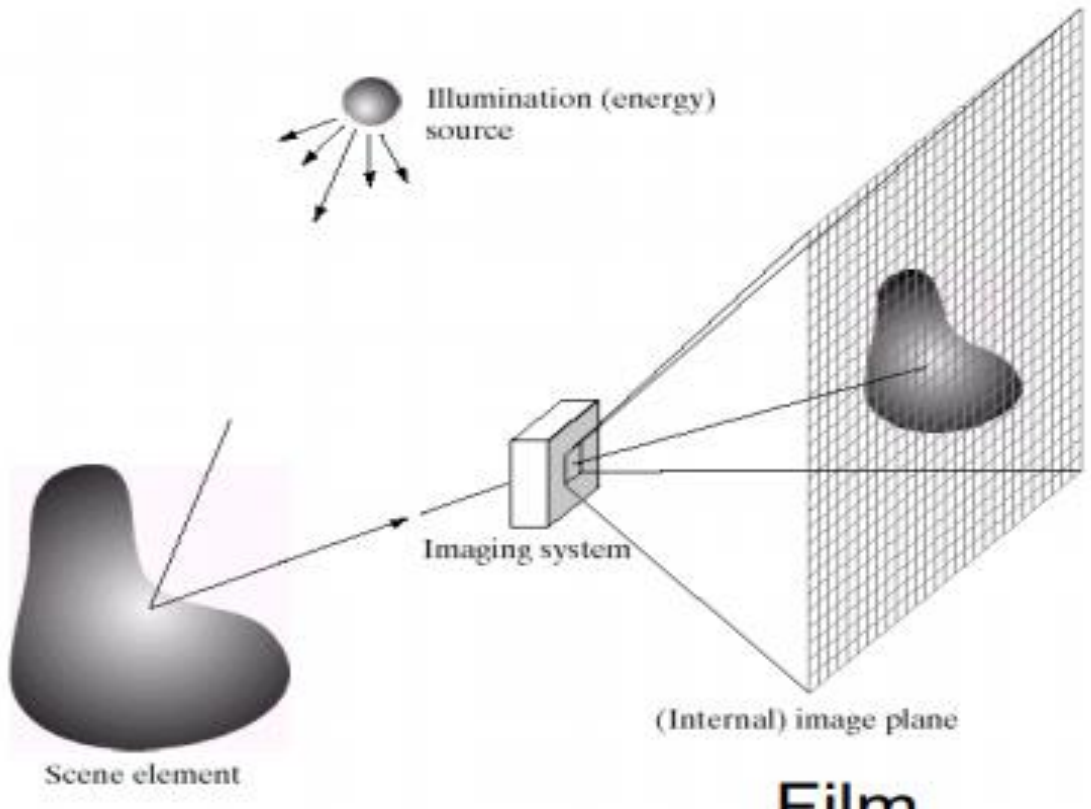
- Neural Networks basics – Linear Separable Problems and Perceptron –
- Multi layer neural networks and Back Propagation, Practical aspects of Deep Learning: Train/ Dev / Test sets, Bias/variance, Vanishing/exploding gradients, Gradient checking, Hyper Parameter Tuning
- Convolutional Neural Networks – Basics and Evolution of Popular CNN architectures – Transfer Learning– Applications : Object Detection and Localization, Face Recognition,
- Neural Style Transfer Recurrent Neural Networks – GRU – LSTM – NLP – Word Embeddings – Transfer Learning –
- Attention Models – Applications : Sentinel Classification, Speech Recognition, Action Recognition
- Restricted Boltzmann Machine – Deep Belief Network –
- Auto Encoders – Applications: SemiSupervised classification, Noise Reduction, Non-linear Dimensionality Reduction Goal Oriented Decision Making – Policy and Target Networks –
- **Deep Quality Network for Reinforcement Learning**
- Introduction to GAN – Encoder/Decoder, Generator/Discriminator architectures Challenges in NN training –
- Data Augmentation – Hyper parameter Settings – Transfer Learning– Developing and Deploying ML Models (e.g., Matlab/Tensor Flow/PyTorch)

Deep Learning Models (Big Picture)

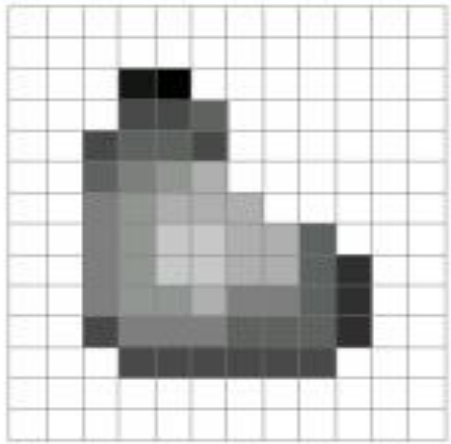
- **Deep Neural Networks**
  - Neural Networks
  - Deep Neural Networks/Feed forward Networks
- **Convolutional Neural Networks (CNN)**
  - LeNet
  - AlexNet
  - VGG
  - ResNet
  - GoogleNet
- **Object Detection models**
  - RCNN
  - Yolo
- **Generative models**
  - Generative Adversarial Networks,(GAN)
- **Sequential Models**
  - RNN
  - LSTM
  - Encoder – Decoder models
  - Attention mechanism
  - **Transformers**
  - **Transformers, Transformer-XL, Masked Language Modelling**
  - **Generative Pre-trained Transformer 3 (GPT-3)**
- **Autoregressive models**
  - Time series forecasting
  - **Generative Pre-trained Transformer 3 (GPT-3)**
  - **NADE, MADE(Masked Autoencoder Density Estimator), PixelRNN**

Introduction to Computer Vision

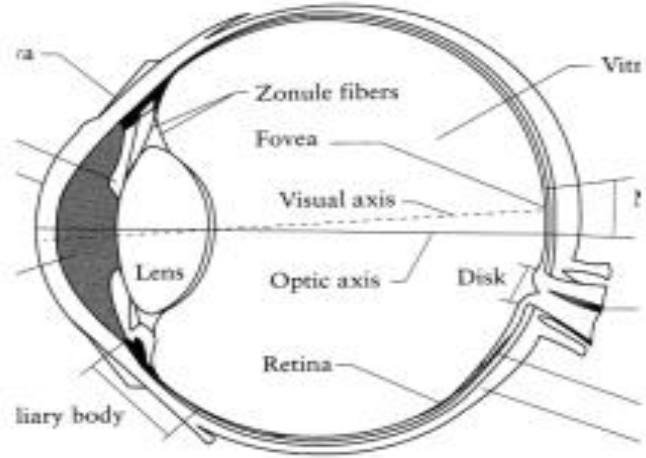
Image Formation: Simple Model



Film



Digital Camera



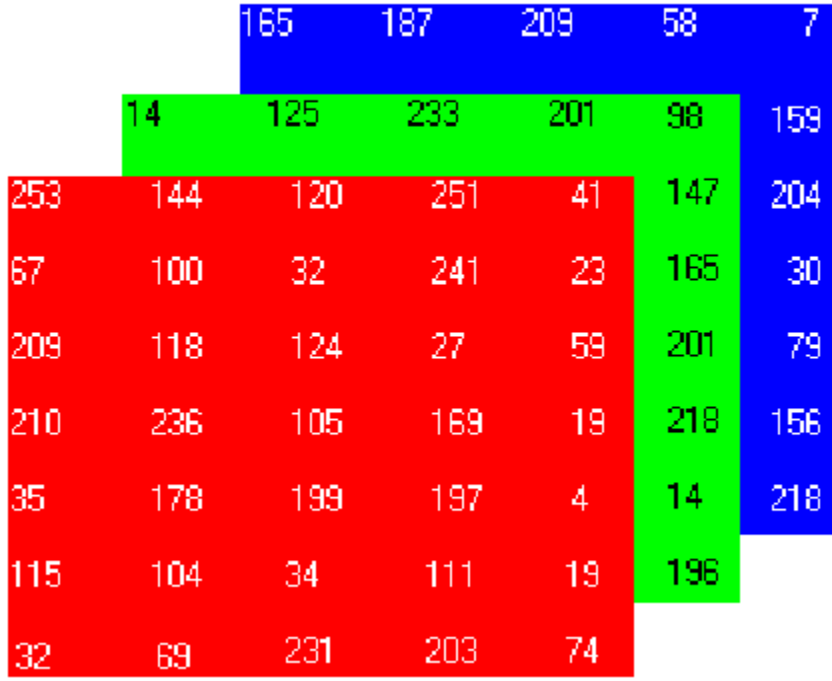
The Eye

Image Matrix- pixel,resolution(spatial, graylevel)

$F(x,y)$

148	123	52	107	123	162	172	123	64	89	···
147	130	92	95	98	130	171	155	169	163	···
141	118	121	148	117	107	144	137	136	134	···
82	106	93	172	149	131	138	114	113	129	···
57	101	72	54	109	111	104	135	106	125	···
138	135	114	82	121	110	34	76	101	111	···
138	102	128	159	168	147	116	129	124	117	···
113	89	89	109	106	126	114	150	164	145	···
120	121	123	87	85	70	119	64	79	127	···
145	141	143	134	111	124	117	113	64	112	···
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	

$I(u,v)$



Color image

Gray Scale image 8 bit gray scale image 0-255 0-black, 255- white

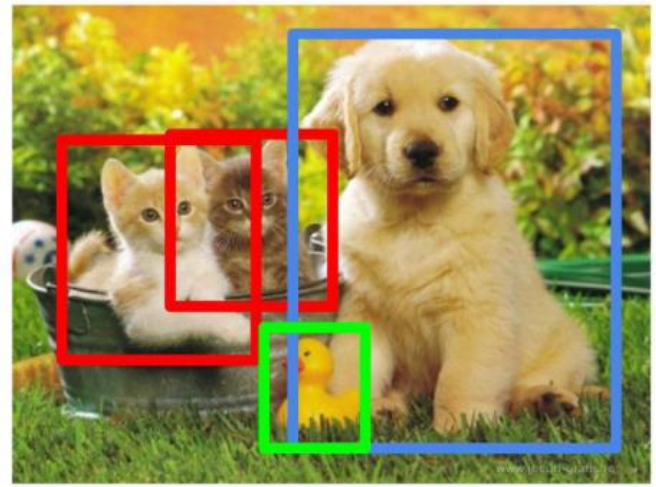
Classification

CAT

Classification + Localization

CAT

Object Detection



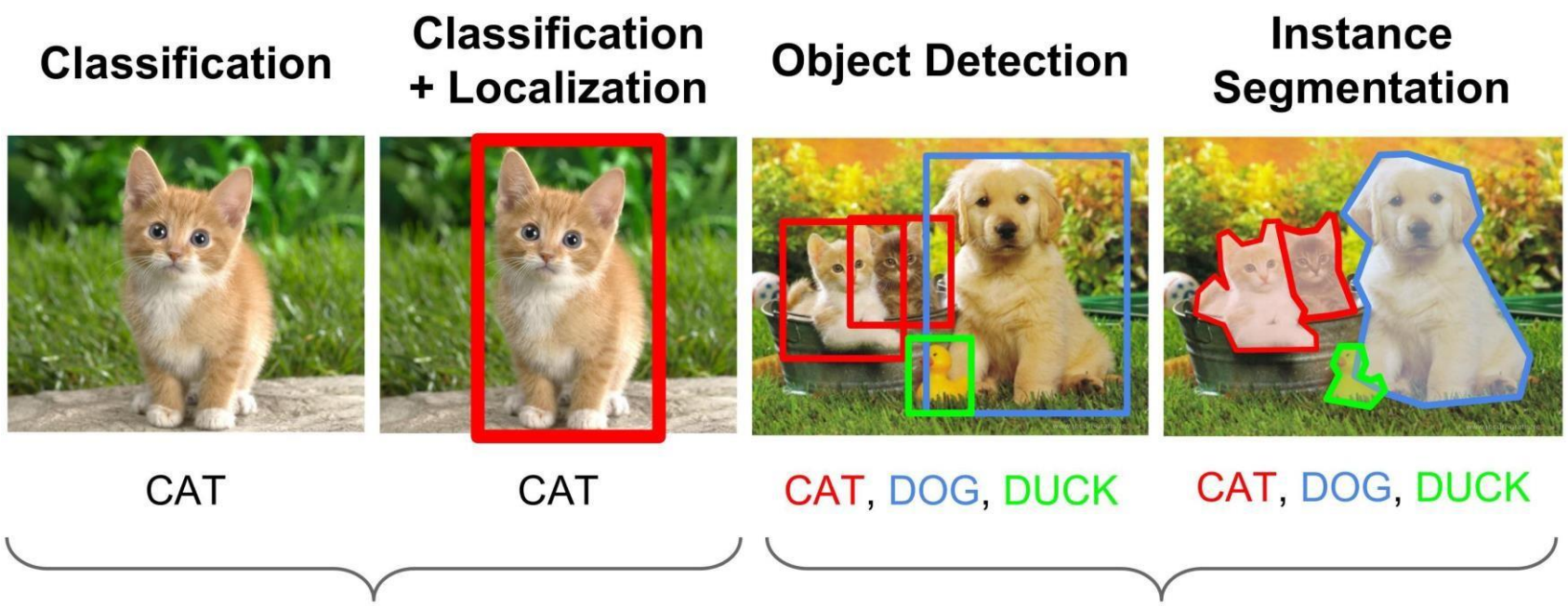
CAT, DOG, DUCK

Instance Segmentation

CAT, DOG, DUCK

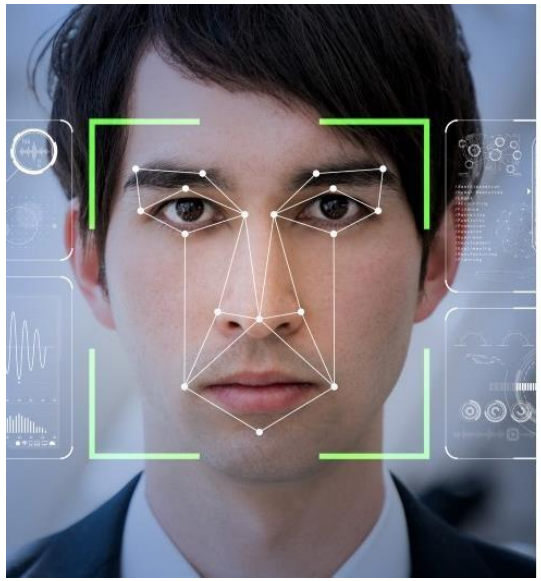
Single object

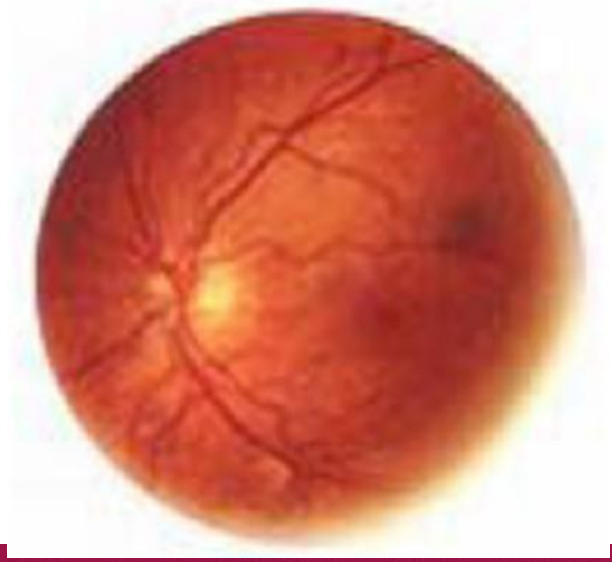
Multiple objects


Classification
Classification + Localization
Object Detection
Instance Segmentation
CAT
CAT
CAT, DOG, DUCK
CAT, DOG, DUCK

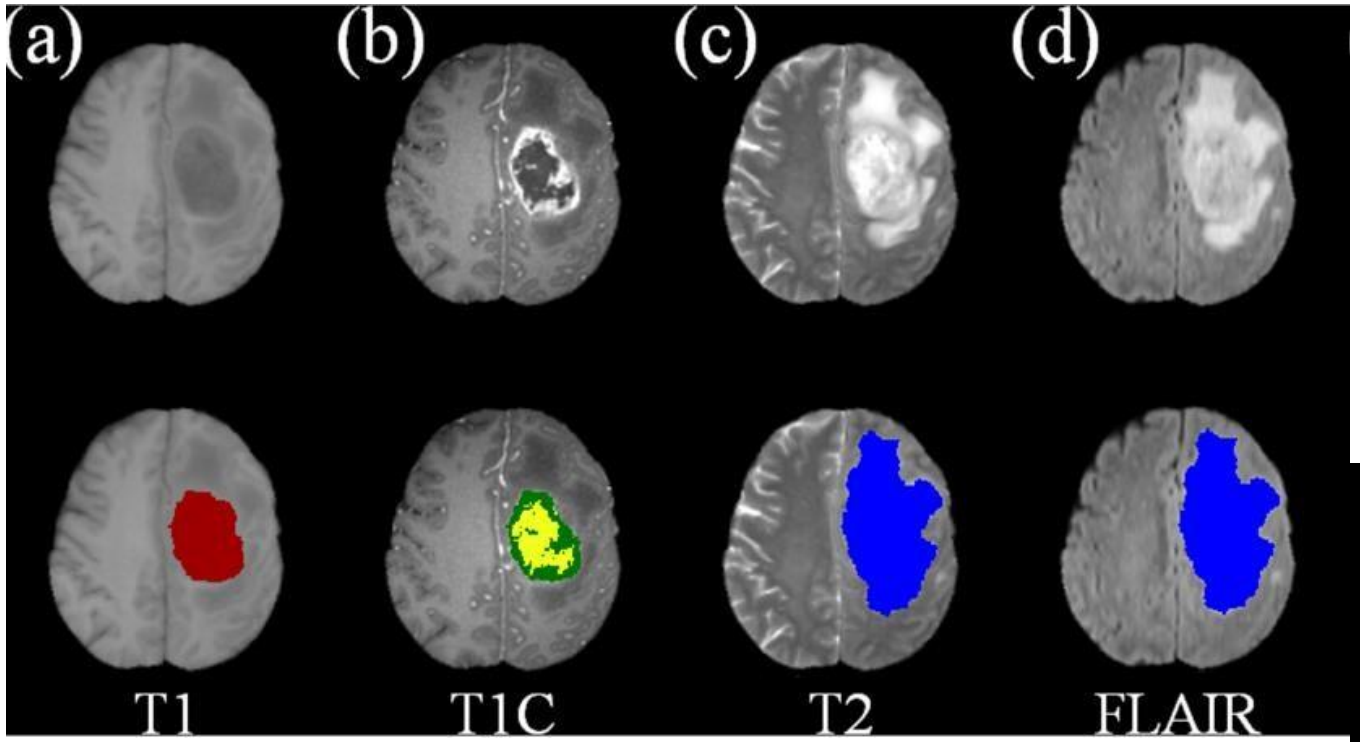

PERSON, CAT, DOG
(A) Classification
(B) Detection
(C) Segmention

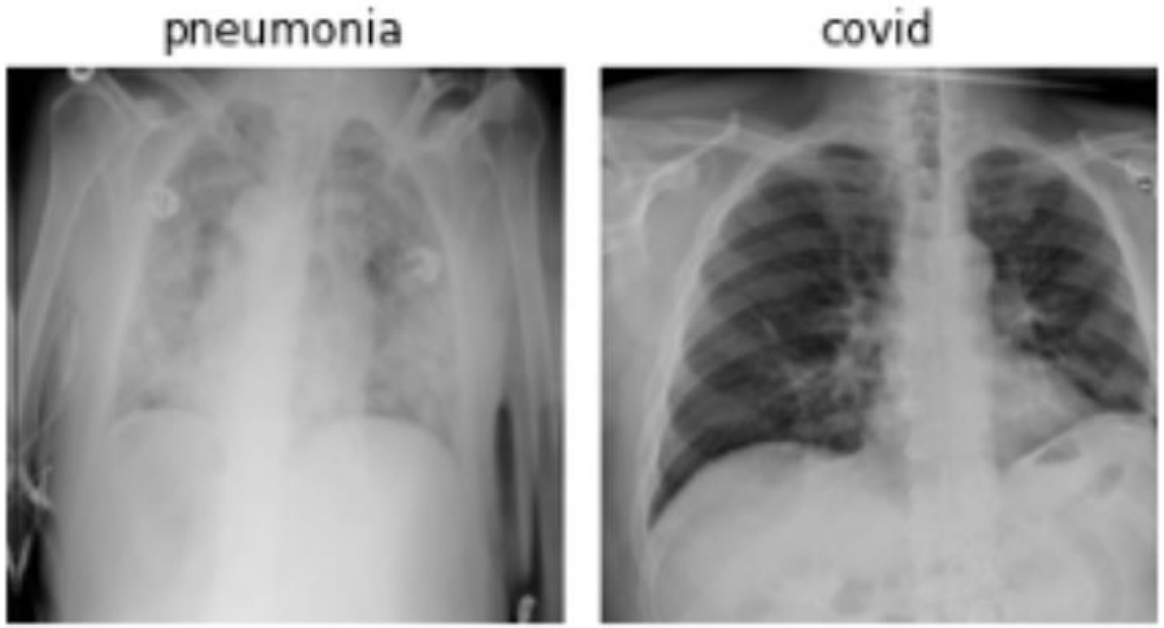
Biometric Applications

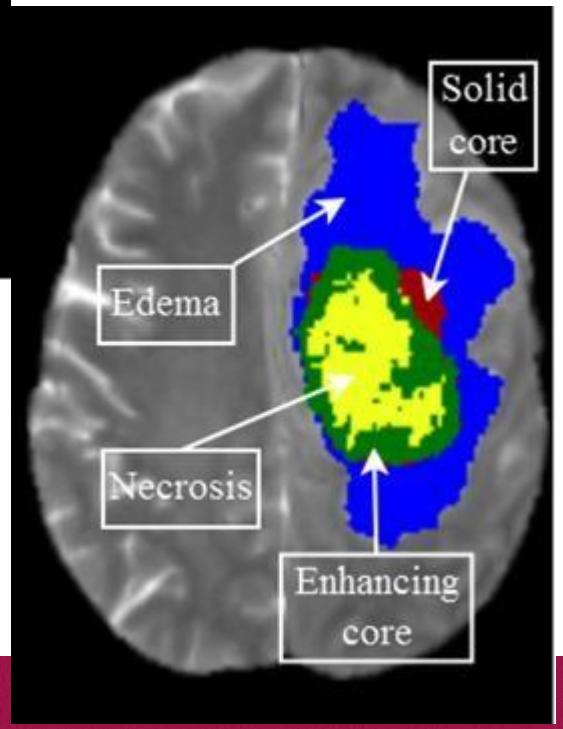




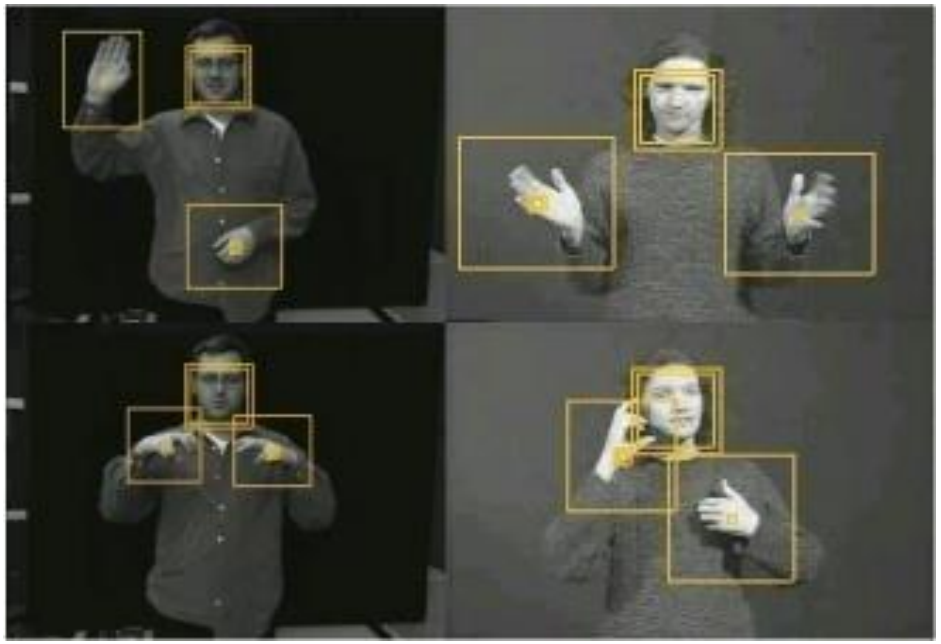
Medical Image Analysis

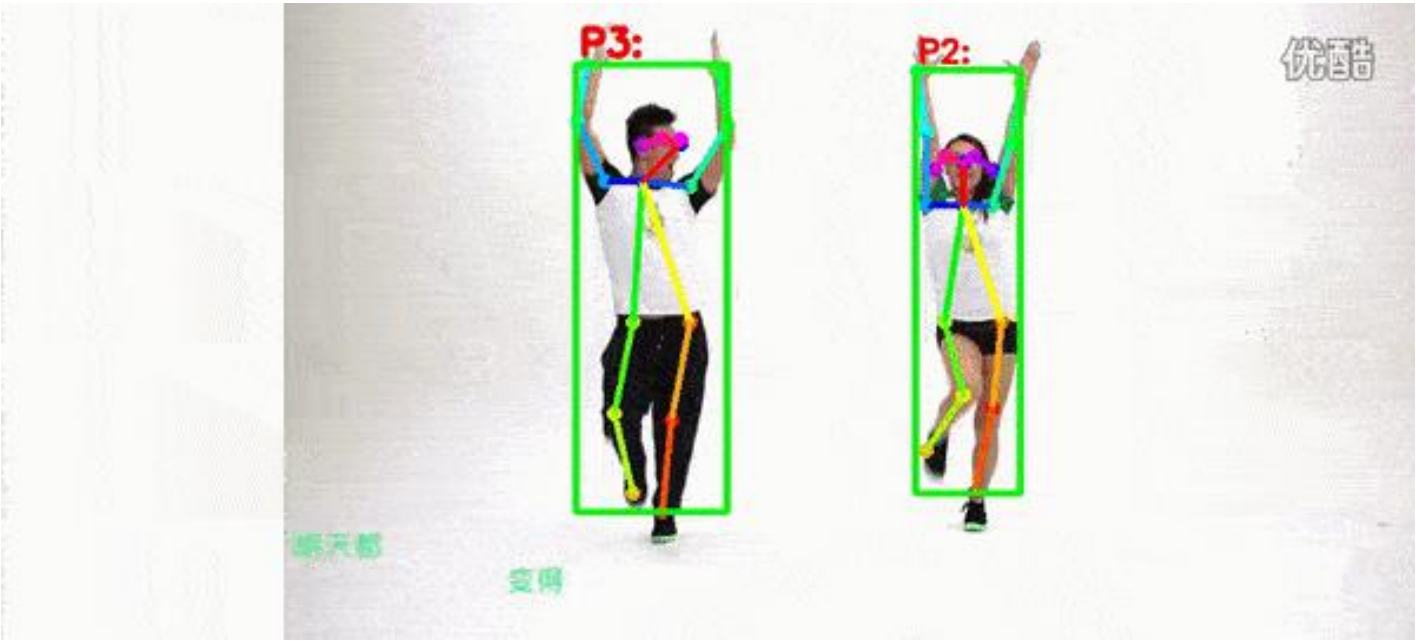




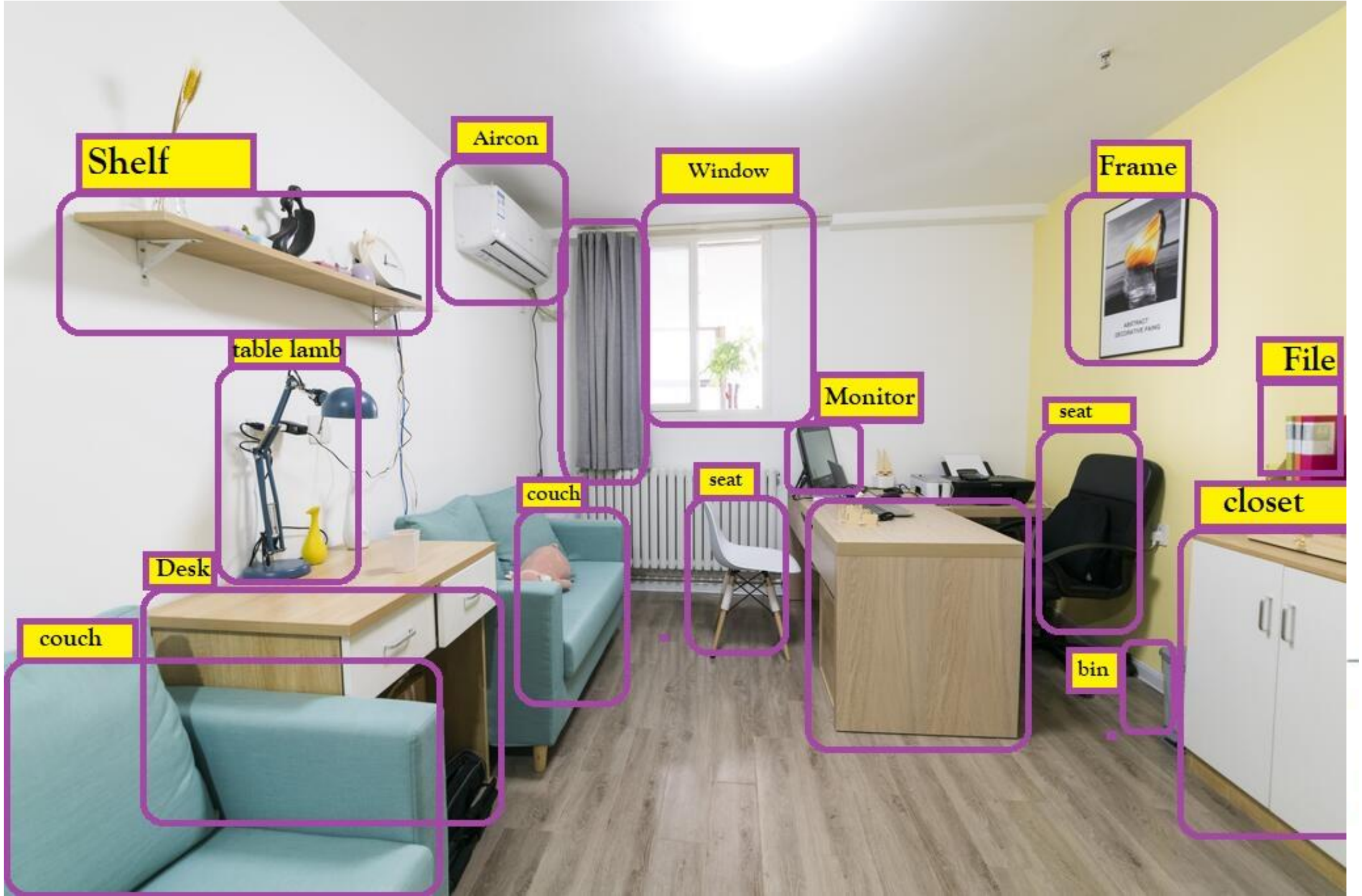


Video Processing - Gesture Recognition/Action Recognition





Scene Understanding



Foil

Yes

No

What is the sandwich laying on?

What is the food on?

Is this a sandwich?

Has the sandwich been cut?

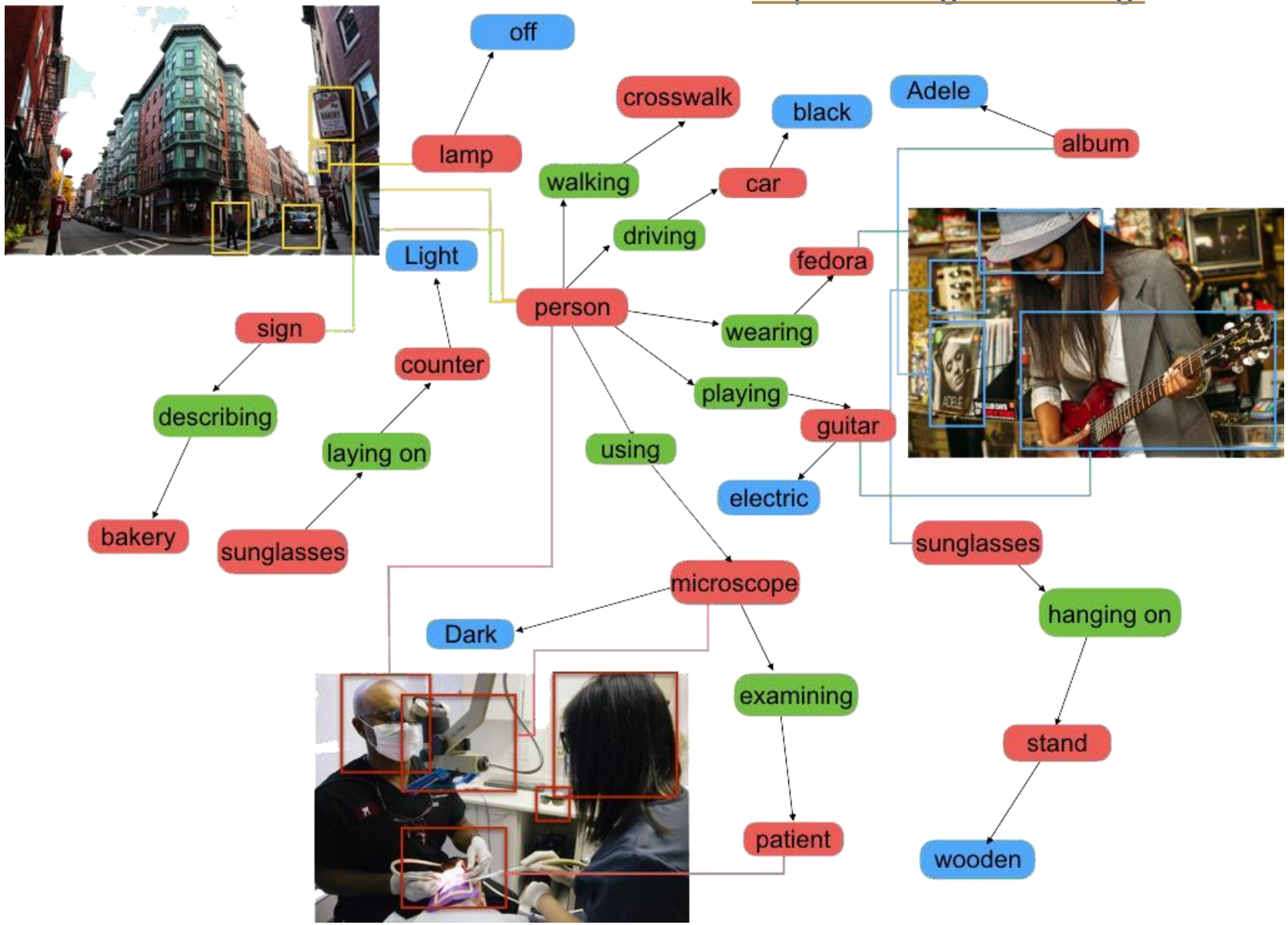
Is the boy wearing a hat?

Is the boy playing baseball?

Is this a professional baseball player?

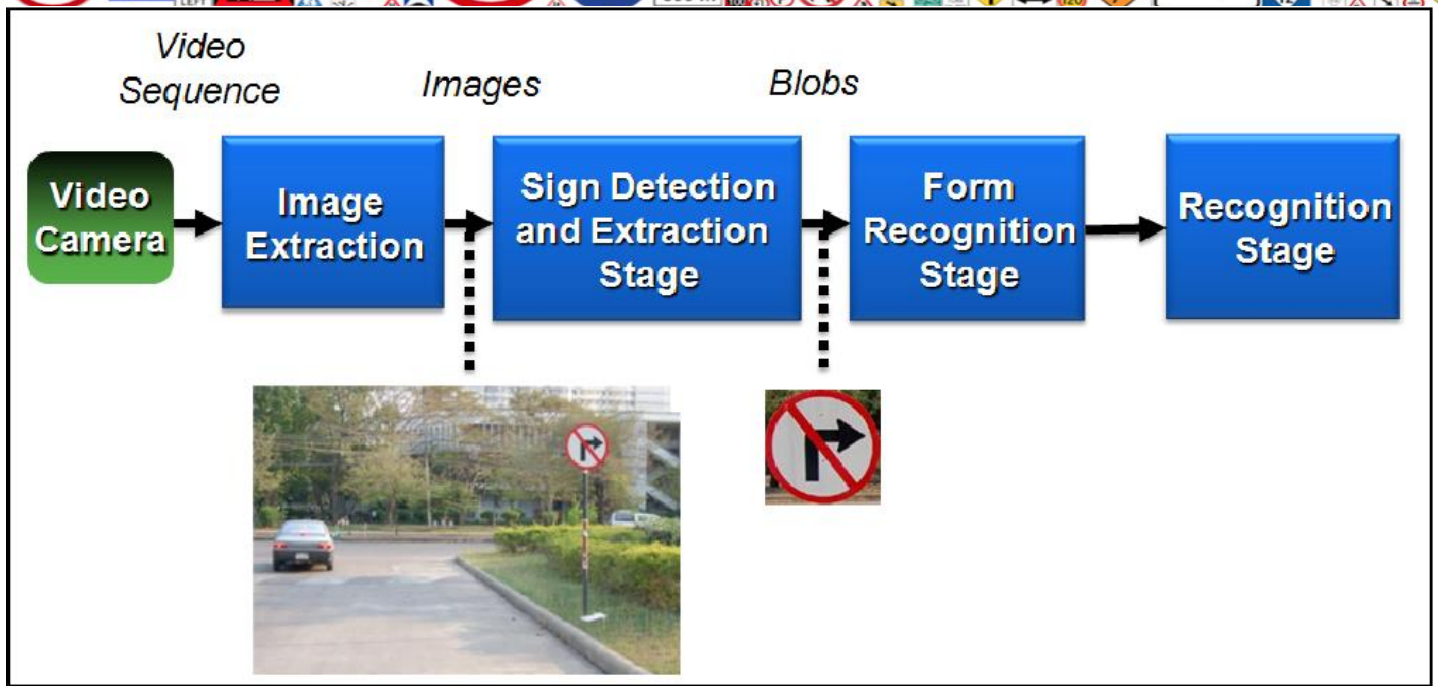
Did the boy hit the ball?

http://visualgenome.org/



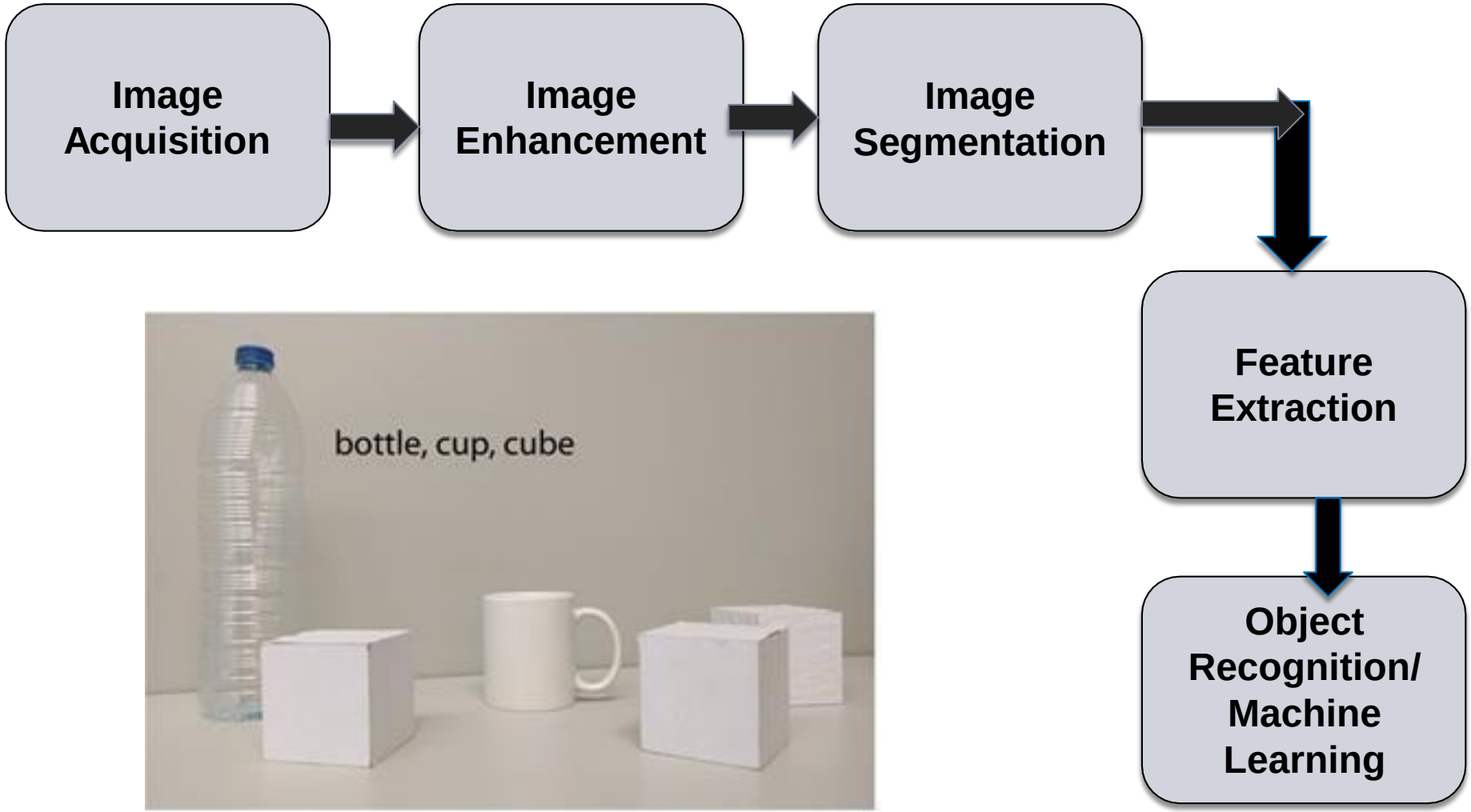
Traffic Sign Recognition

Traffic Sign Recognition



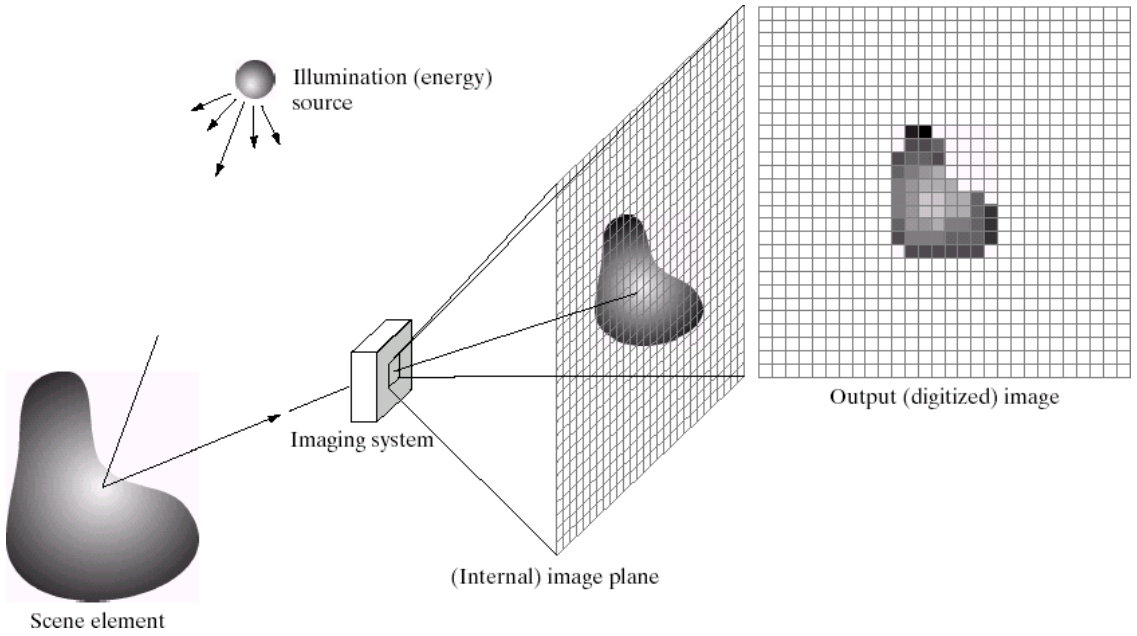
Static Sign Language Recognition

KEY STEPS-IMAGE ANALYSIS

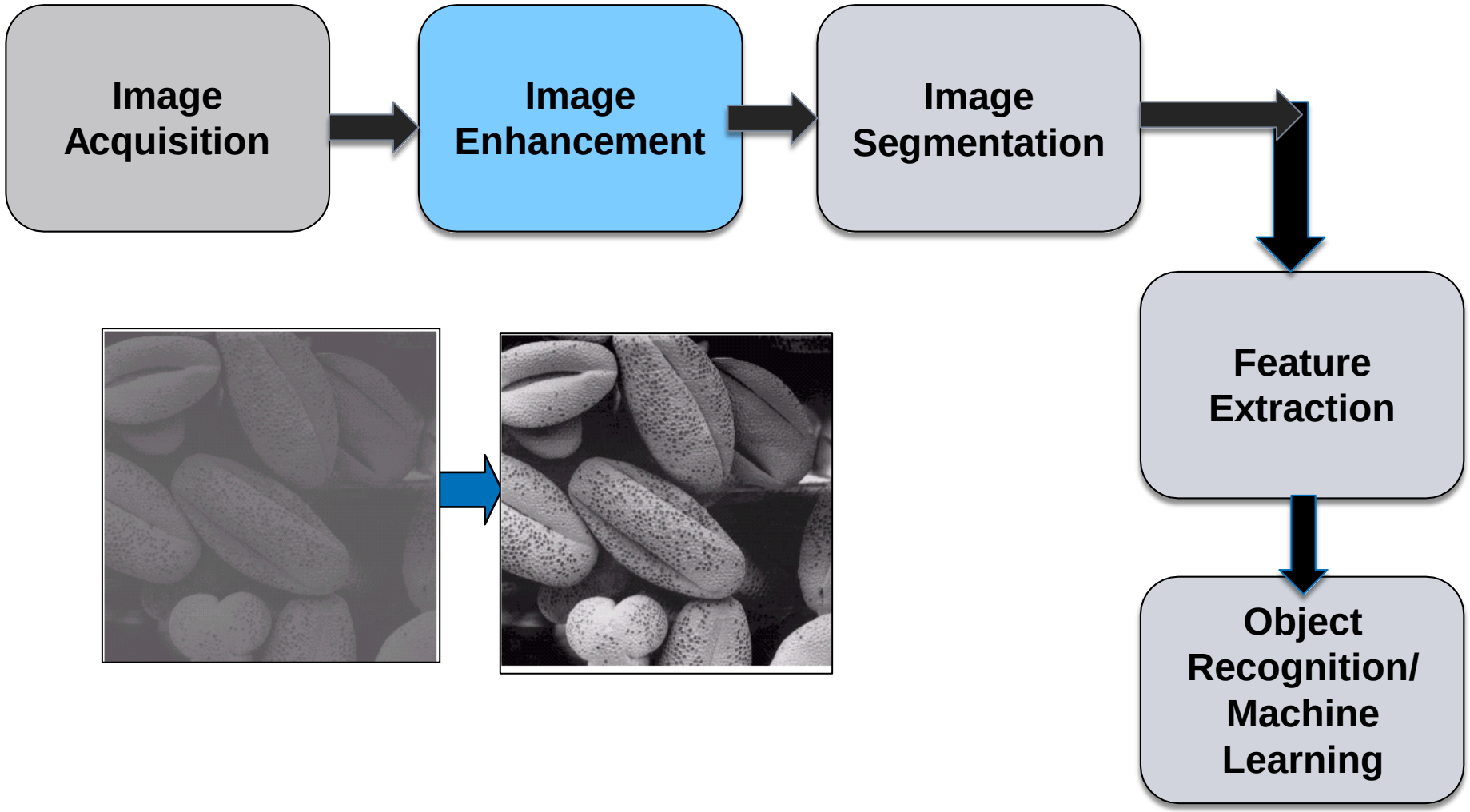


KEY STEPS-IMAGE ANALYSIS

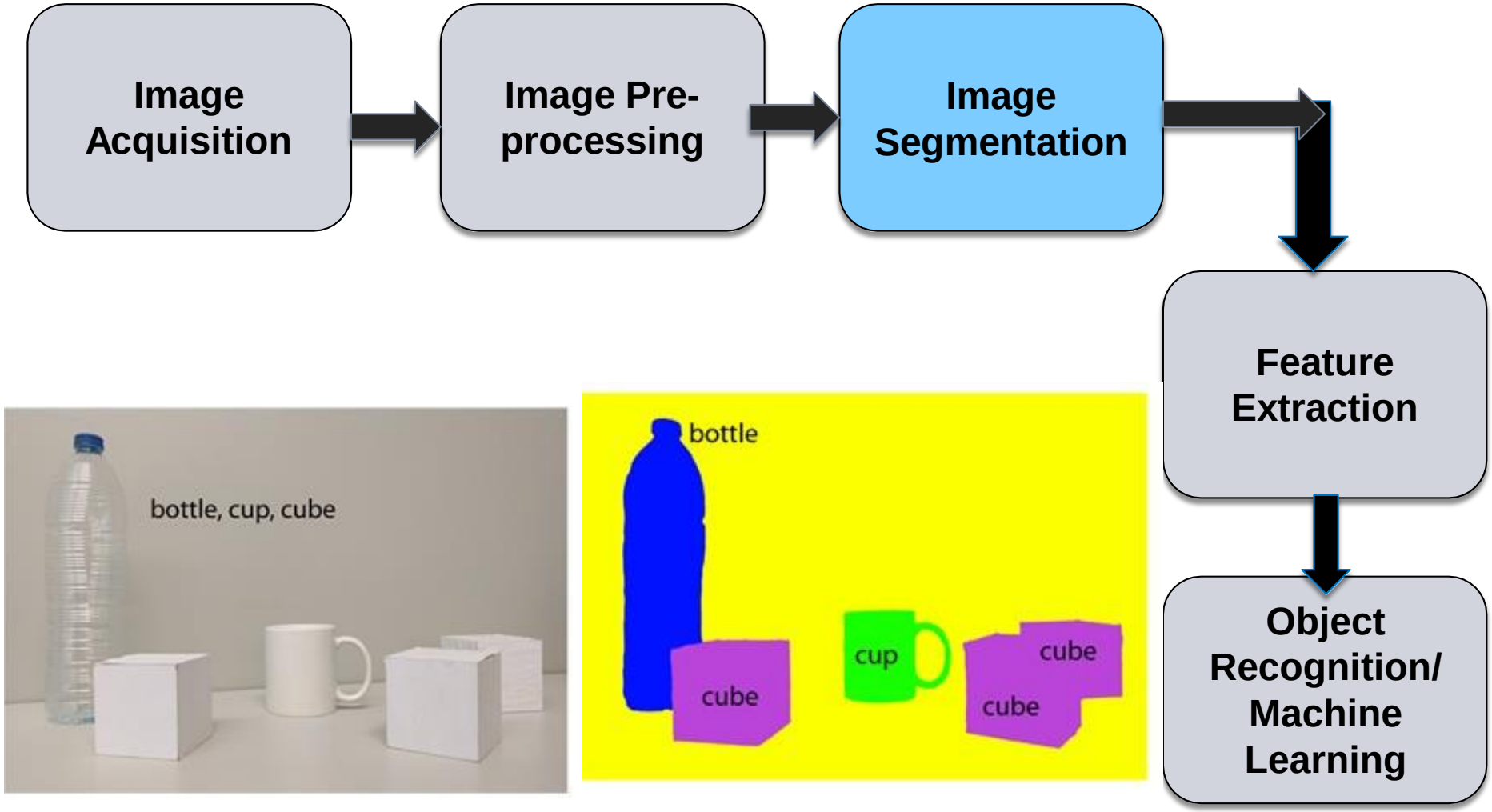
Image Acquisition → Image Enhancement → Image Segmentation → Feature Extraction → Object Recognition/ Machine Learning



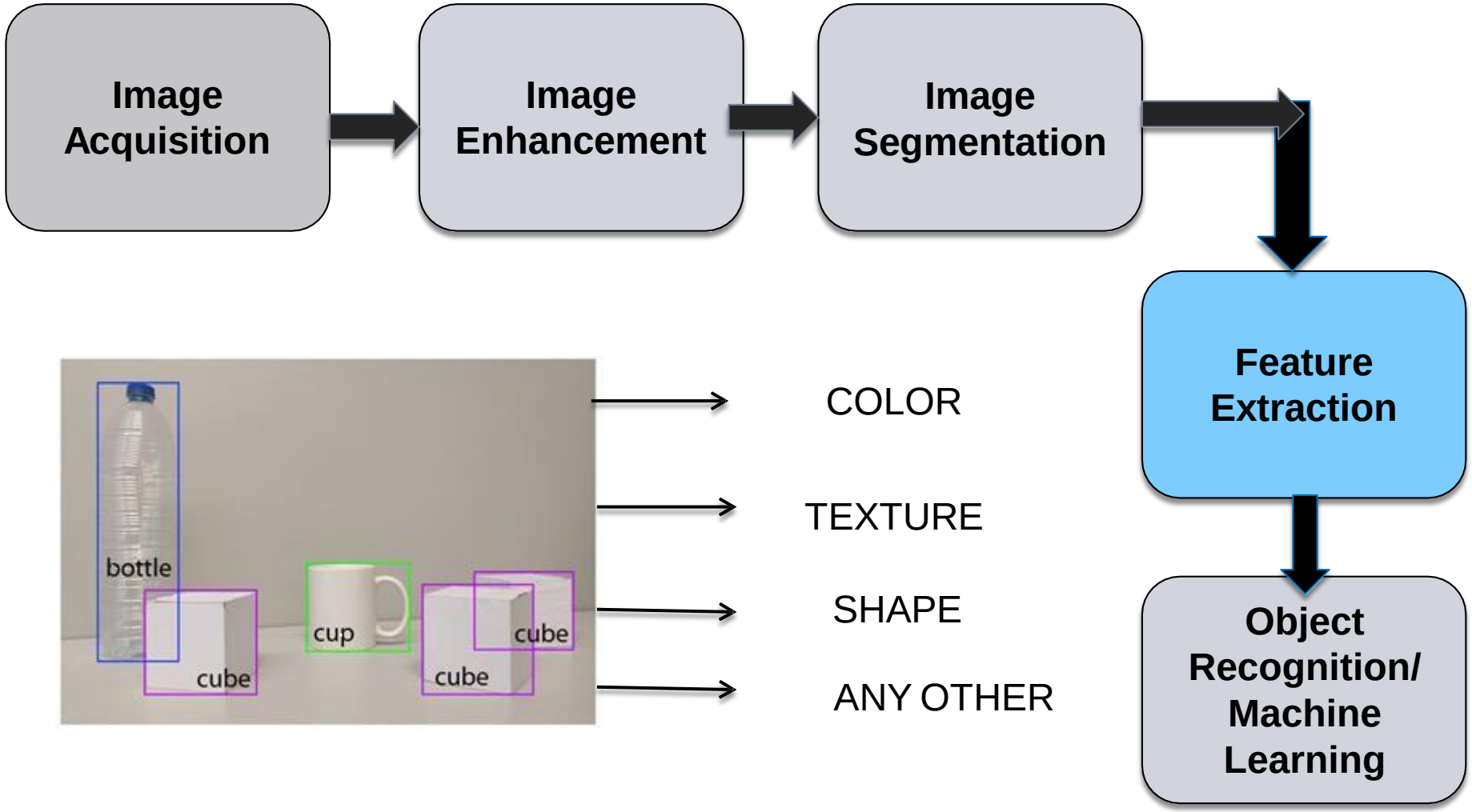
KEY STEPS-IMAGE ANALYSIS



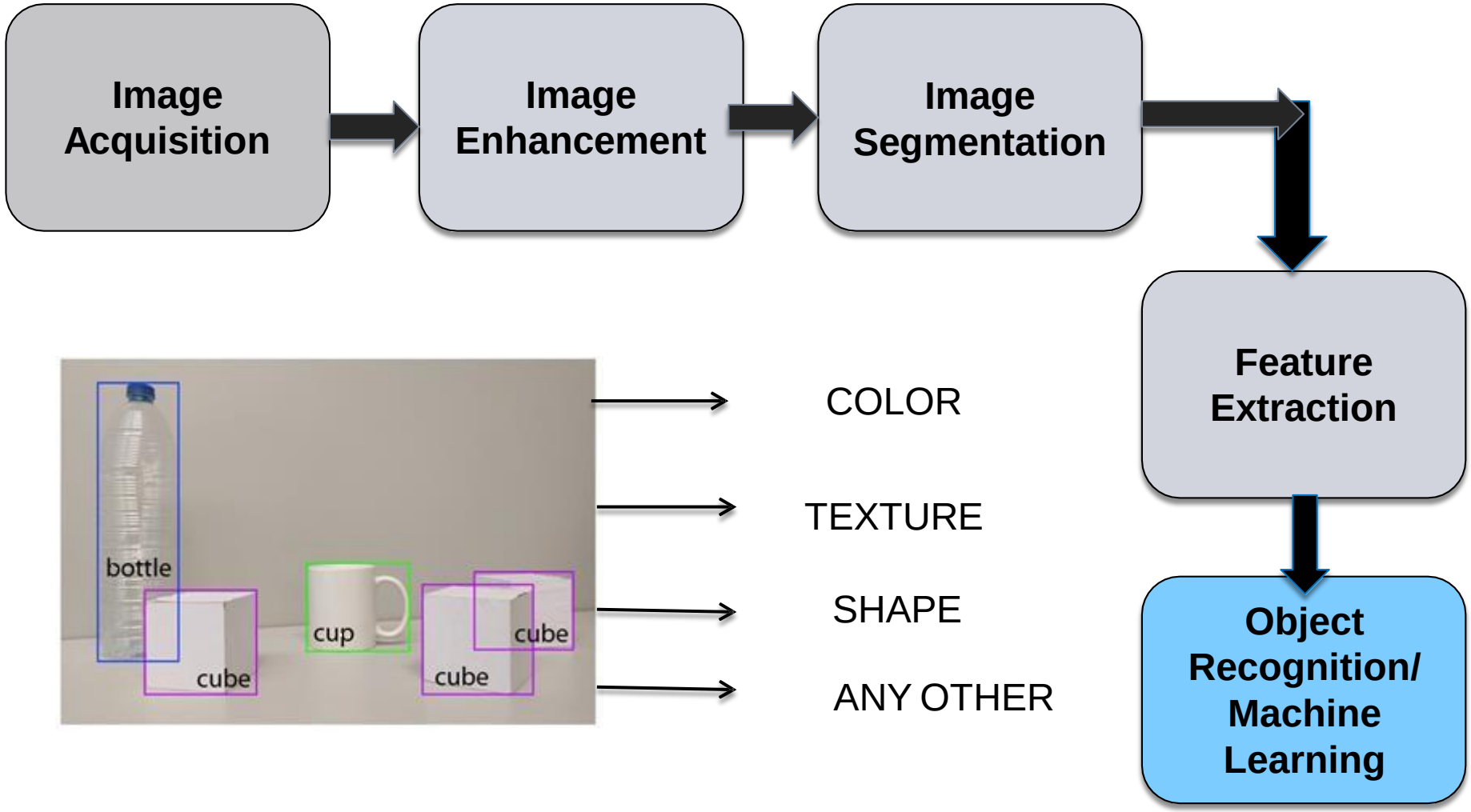
KEY STEPS-IMAGE ANALYSIS



KEY STEPS-IMAGE ANALYSIS



KEY STEPS-IMAGE ANALYSIS



KEY STEPS-IMAGE ANALYSIS

Image Acquisition → Image Enhancement → Image Segmentation → Feature Extraction → Object Recognition/ Machine Learning

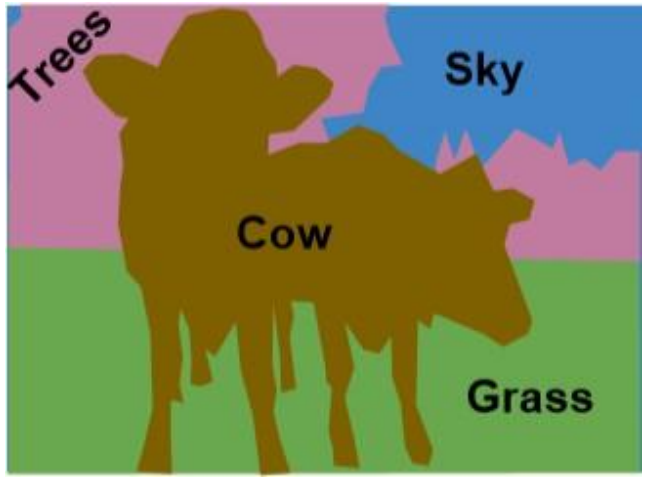
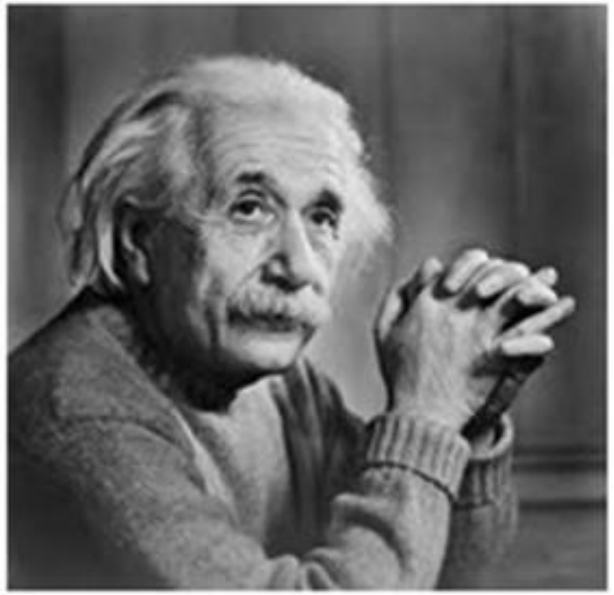


Image Negative



- Spatial domain enhancement methods can be generalized as
  - $g(x,y) = T[f(x,y)]$

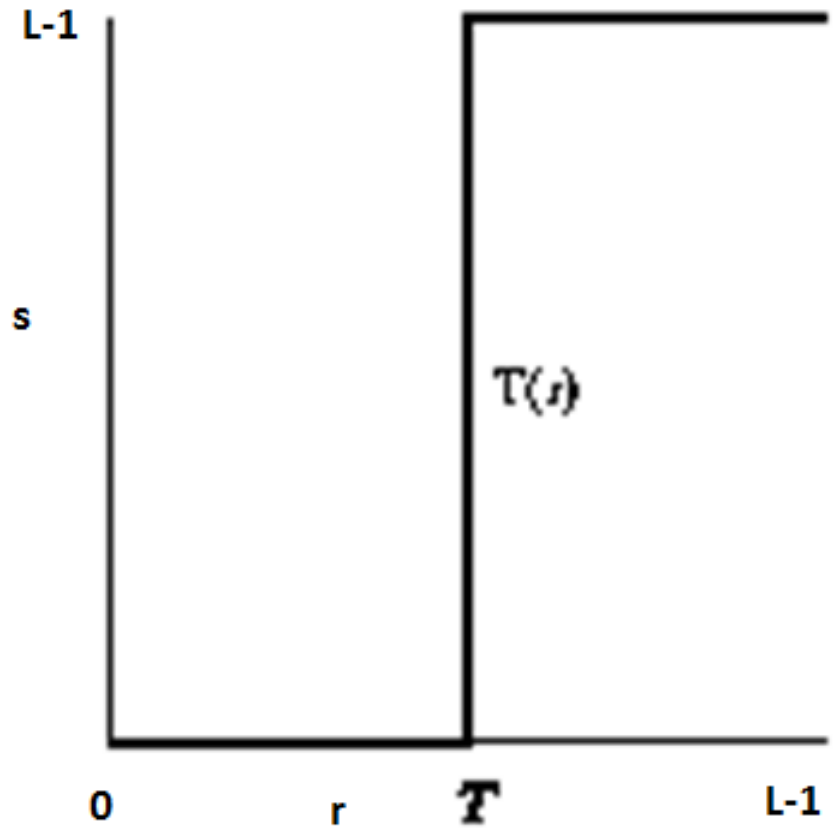
$f(x,y)$: input image

$g(x,y)$: processed (output) image

$T[*]$: an operator on f (or a set of input images), defined over neighborhood of (x,y)

Image Negative: $s = L - 1 - r$

Thresholding Function

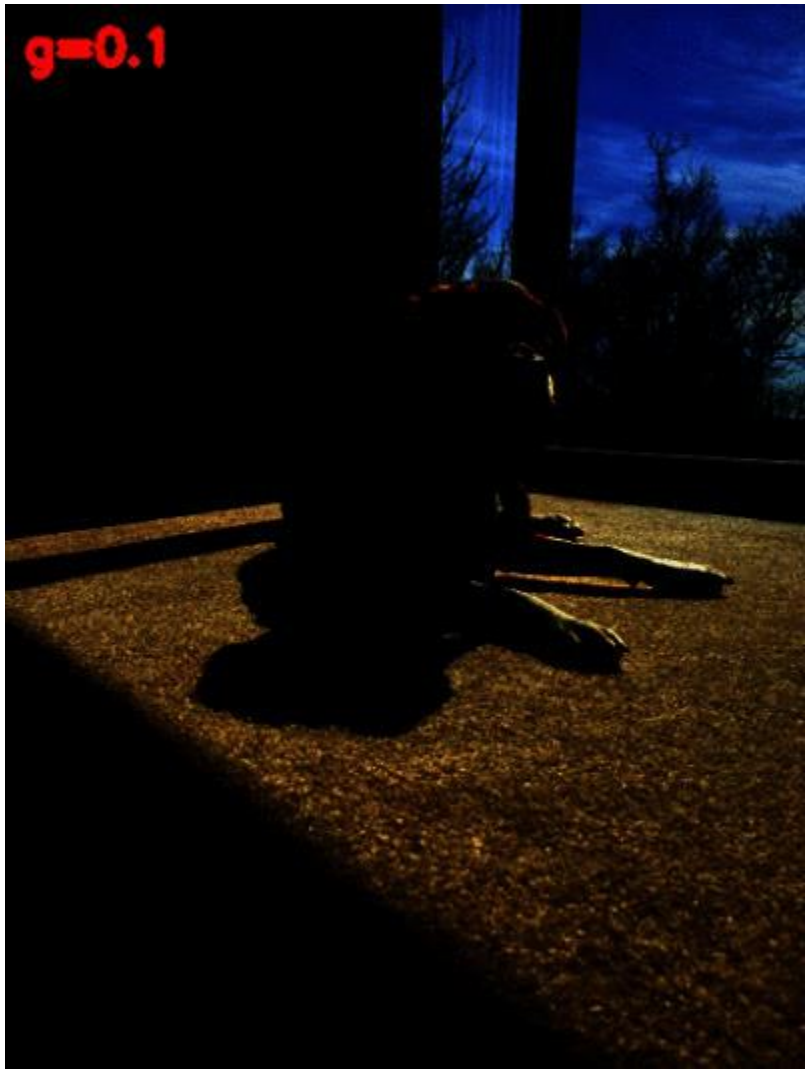


Thresholding Example

Original Image

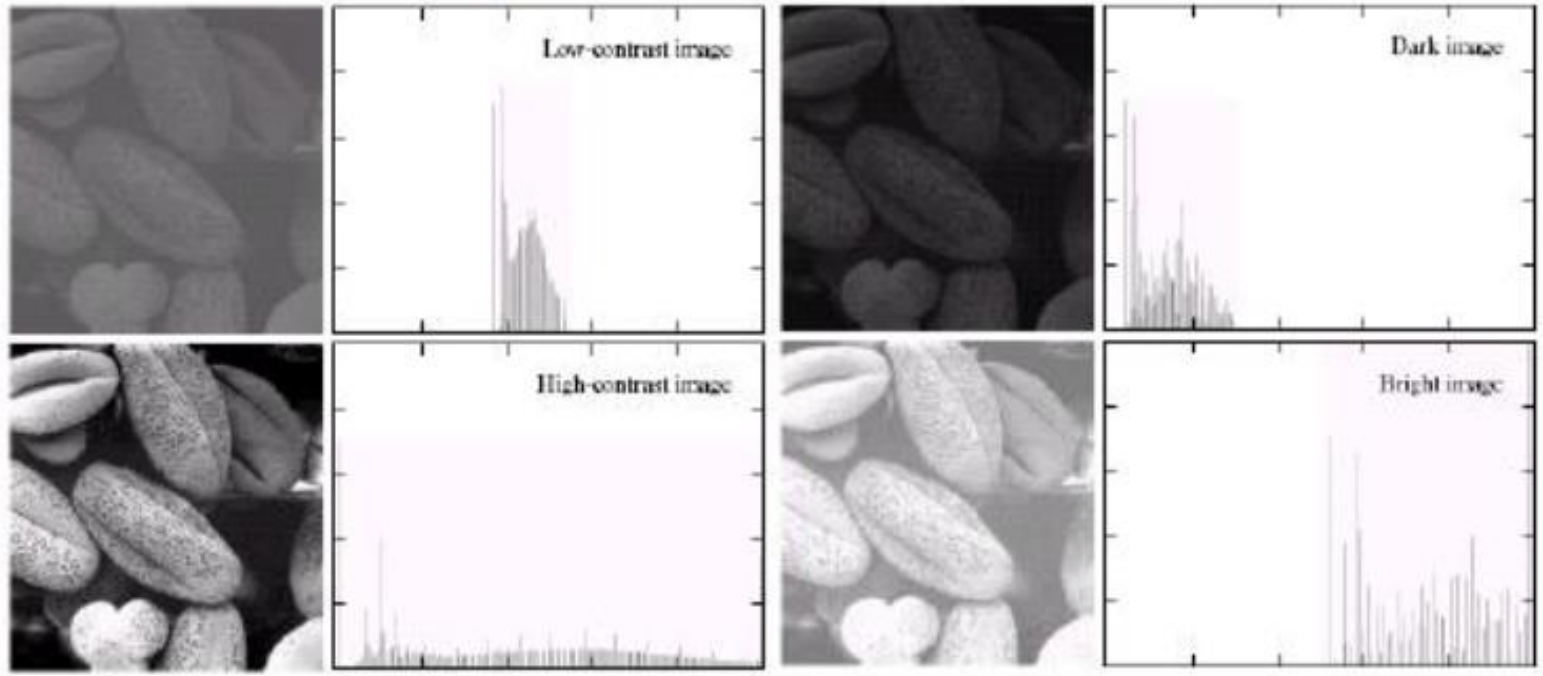
Thresholded Image

Gama Correction



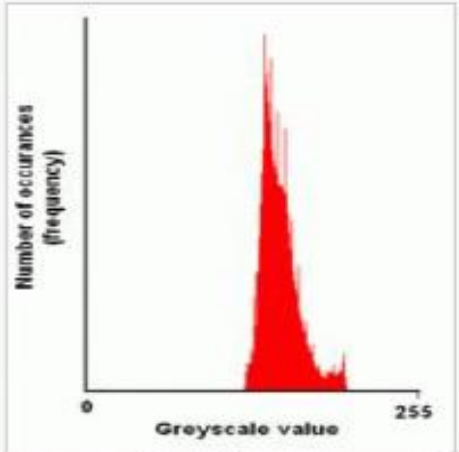
Histogram : Example

- A selection of images and their Histograms
- Note that the high contrast image has the most evenly spaced histogram
- Histograms of low contrast images are located in certain portions and not in the entire gray scale range



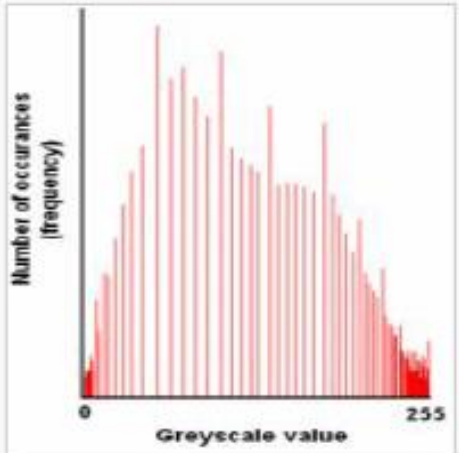
Histogram Equalization: Example

An unequalized image



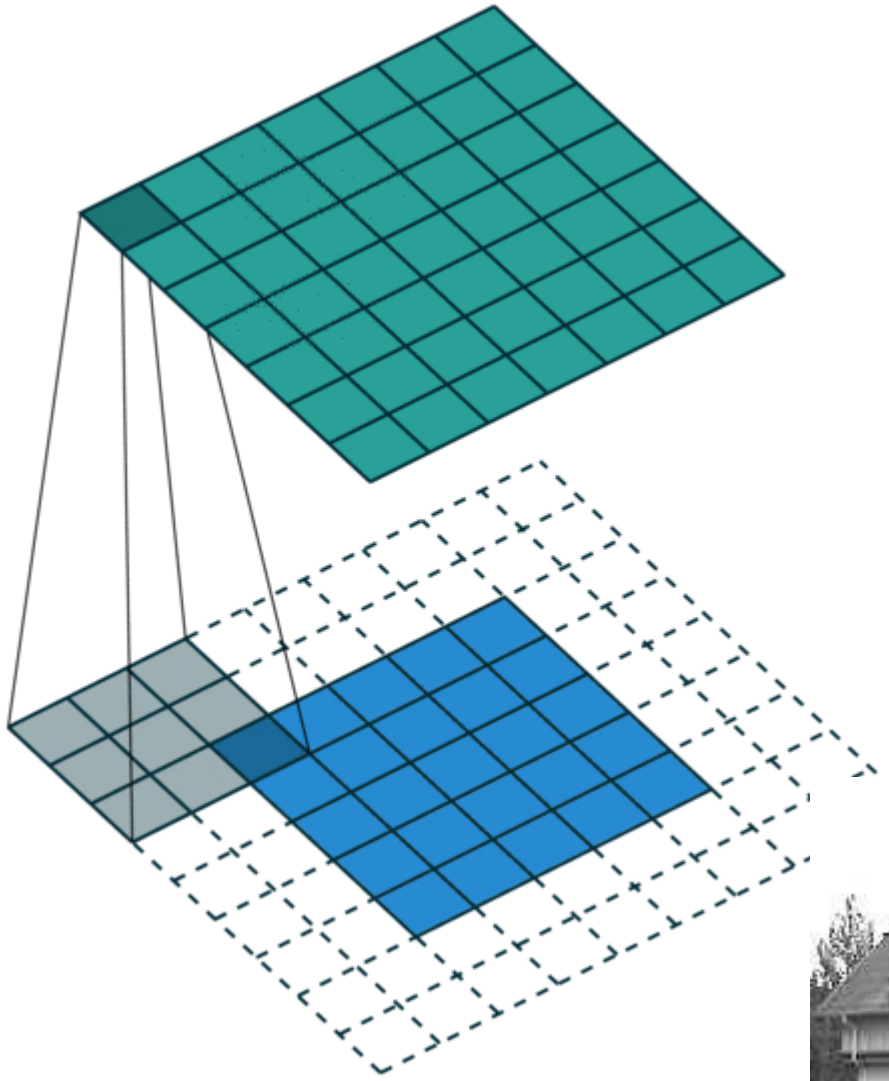
Corresponding histogram

The same image after histogram equalization



Corresponding histogram

Convolution operation in images



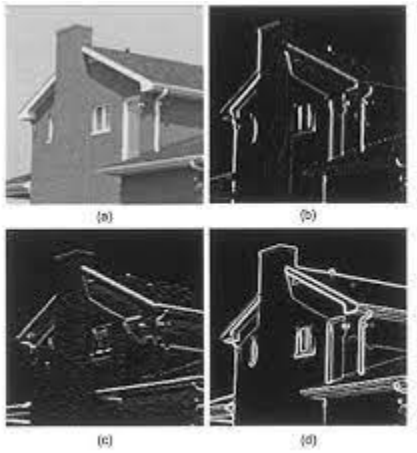
Image

100	100	200	200
100	100	200	200
100	100	200	200
100	100	200	200

Kernel/Filter

-1	0	1
-2	0	2
-1	0	1

-100
-200
-100
200
400
+200
=400



Edge detection using sobel operator

Smoothing Spatial Filters

- One of the simplest spatial filtering operations we can perform is a smoothing operation
  - Simply average all of the pixels in a neighbourhood around a central value
  - Especially useful in removing noise from images

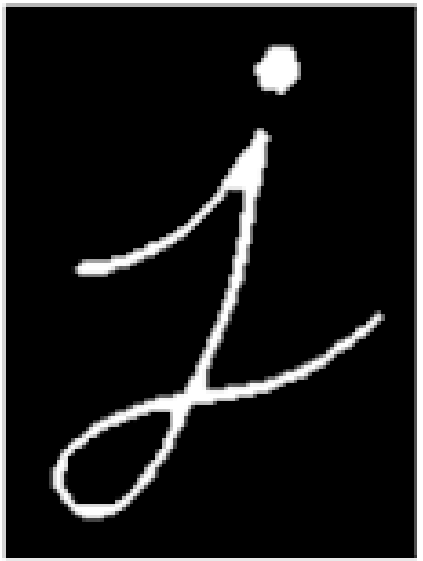
$^1/_9$	$^1/_9$	$^1/_9$
$^1/_9$	$^1/_9$	$^1/_9$
$^1/_9$	$^1/_9$	$^1/_9$

Simple Averaging Filter

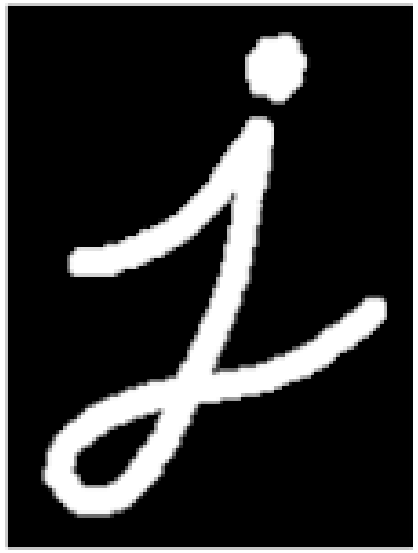
or

$$\frac{1}{9}\times\begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

Image Preprocessing- Erosion and dilation



erosion



original

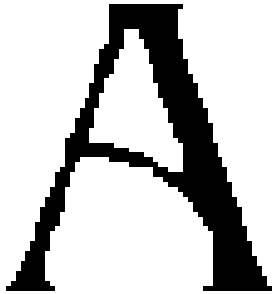
dilation

字母"j":(左）侵蚀，（中）原始图像，（右）扩张

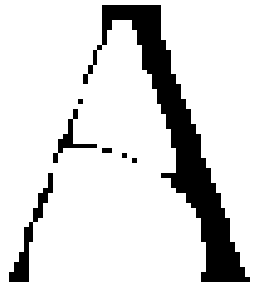
Erosion Example 1

Watch out: In these examples a 1 refers to a black pixel!

Original image



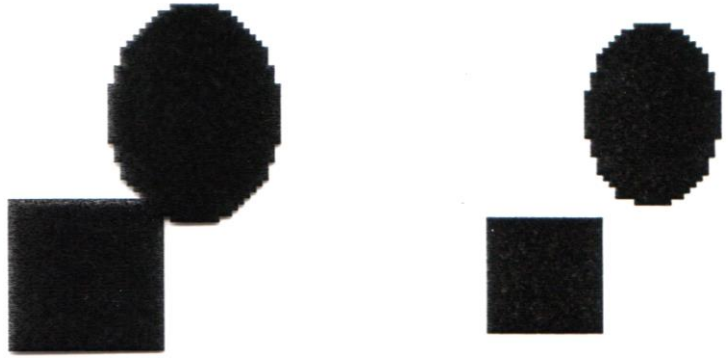
Erosion by 3*3 square structuring element

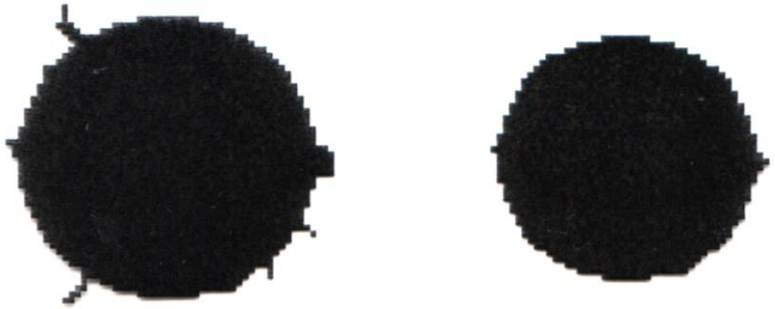


Erosion by 5*5 square structuring element

What Is Erosion For?

Erosion can split apart joined objects

Erosion can 

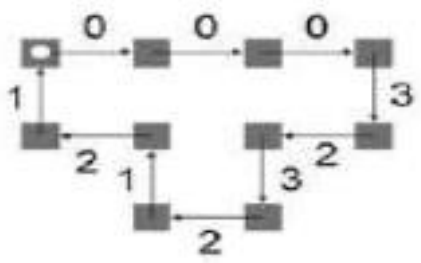
Watch out: 

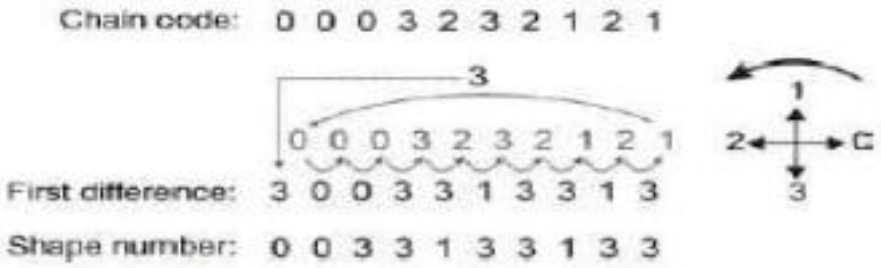
Feature Extraction and Classification

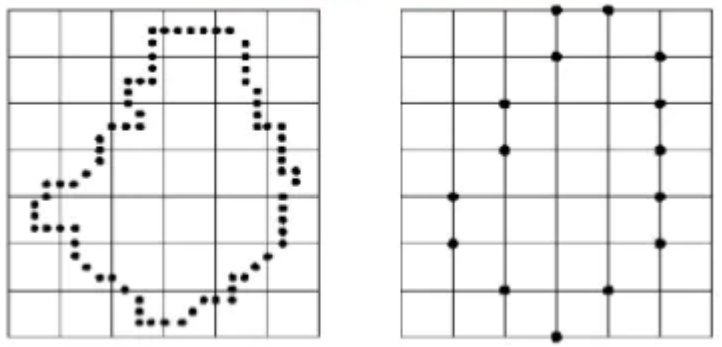
Feature Extraction- Hand Crafted

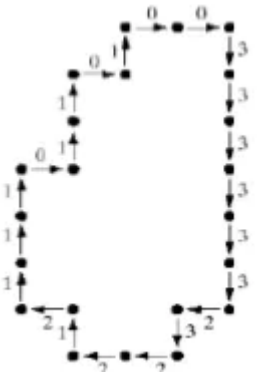
- Chaincode (shape)
- Fourier Descriptors (shape)
- Harris Corner Detection
- Gray Level Co-Occurrence Matrix (GLCM)- texture
- Histogram of Oriented Gradients (HOG)- shape
- Moment based feature (shape)
- Haar Cascades
- Color features- color spaces
- Scale-Invariant Feature Transform (SIFT)- keypoint based
- Speeded Up Robust Feature (SURF)- keypoint
- Lot more…..

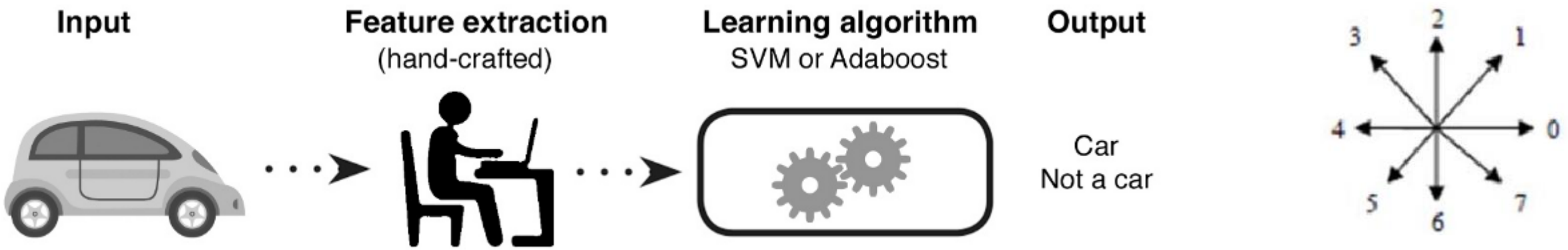
A simple example –chain code







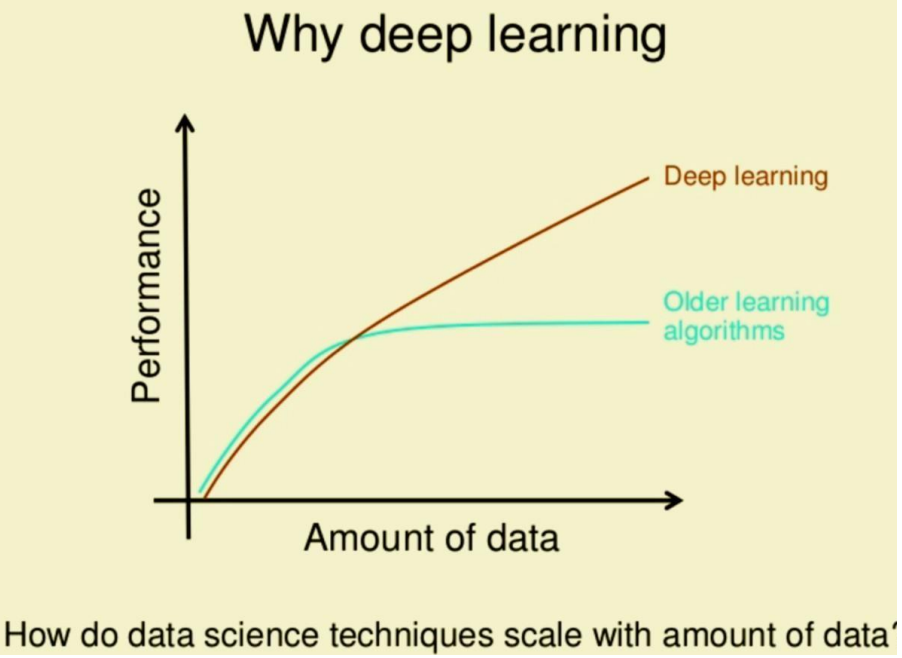


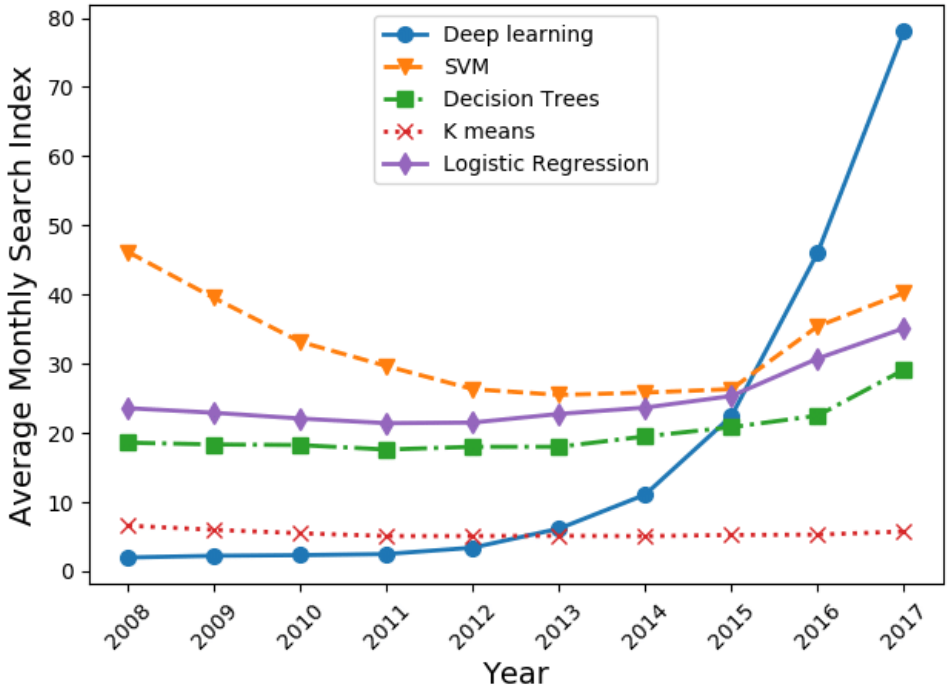


Deep Learning- A breakthrough

Given the Availability of Data, Deep Learning performance has surpassed all traditional algorithms

Artificial Intelligence Paradigmshift!-
Machine Learning and **Deep Learning?**





Promising results if trained with lot of data!

Thank you