



QBE DATA SCIENCE WORKSHOP

Introduction to Computer Vision (CV)

How CV is leveraged in expanding business

Table of Contents

POINTS FOR DISCUSSION:

- What is Computer Vision?
- Framework and infrastructures in building a system
- Challenges in Computer Vision
- Applications
- Is remote work here to stay?
- What's next?





computer vision

Computer vision is the field of computer science that focuses on replicating parts of the complexity of the human vision system and enabling computers to identify and process objects in images and videos in the same way that humans do.

[From https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision](https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision)

"A branch of artificial intelligence and image processing concerned with computer processing of images from the real world. Computer vision typically requires a combination of low level image processing to enhance image quality (e.g. remove noise, increase contrast) and higher level pattern recognition and image understanding to recognize features present in the image"

Image: as seen by computer

Computer: array of brightness values vs Human: perceive structure

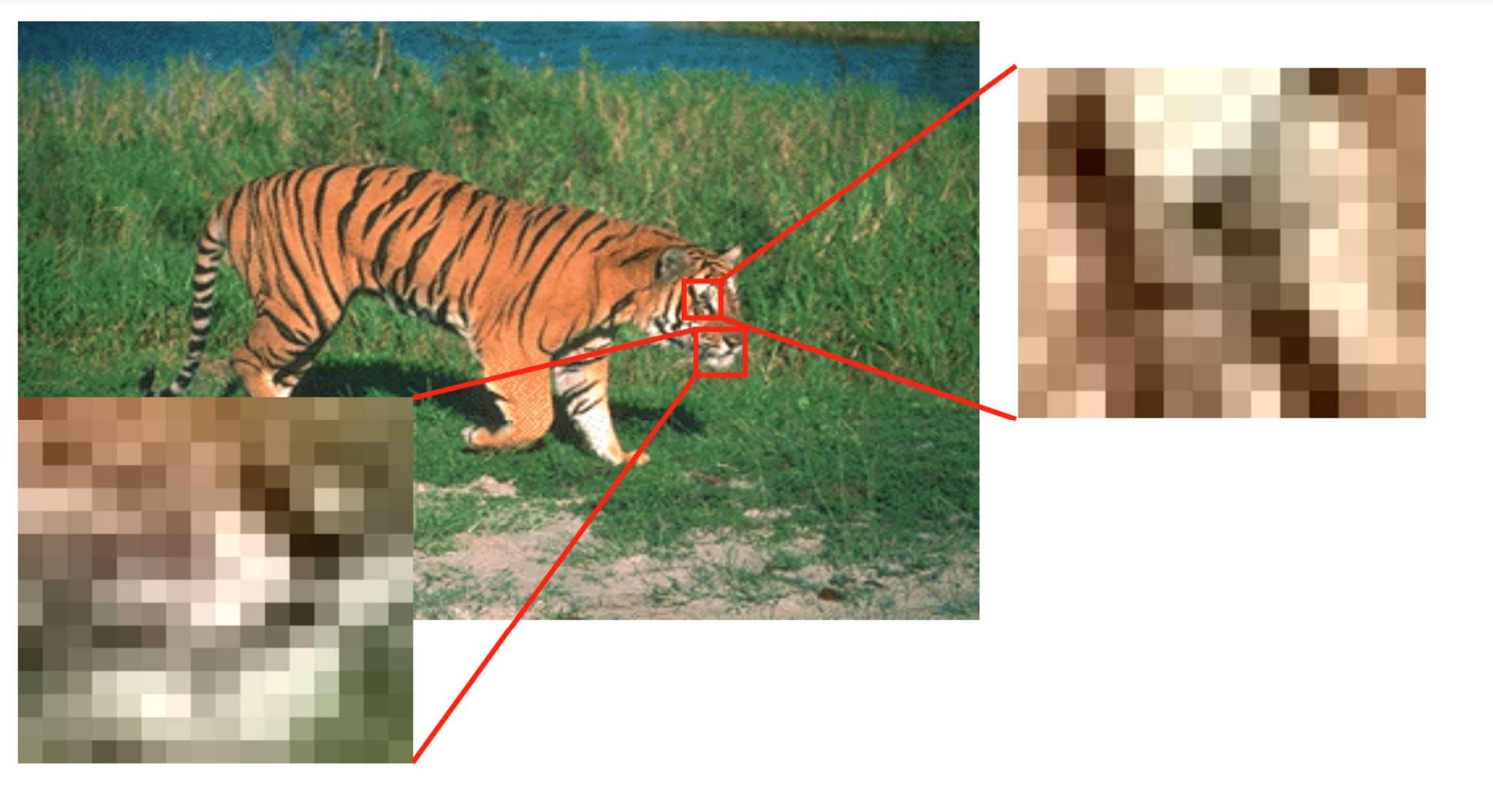
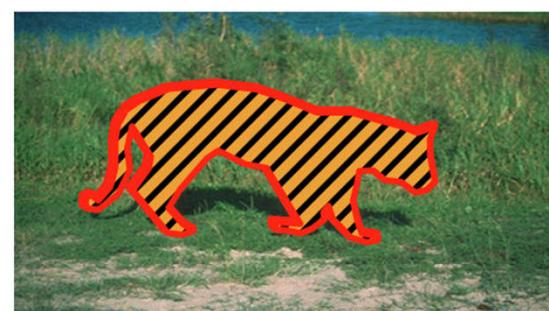


Image: as seen by humans

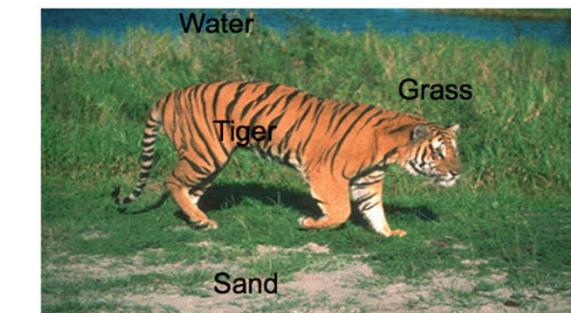
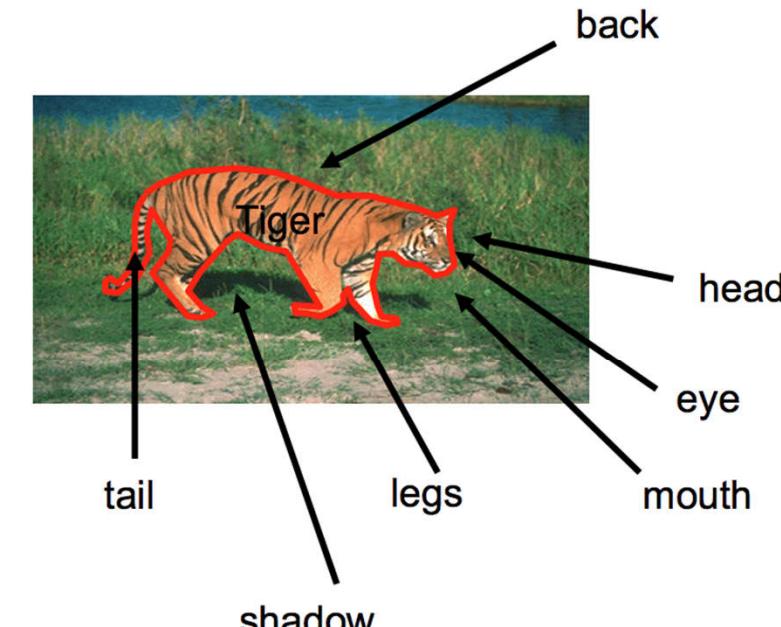
Goal of computer vision (CV): equip computers with perception



pixels



regions



Semantic labels

Goal of computer vision (CV)

understand beauty (combination of color, shape, texture, form, ...?)



(c) Isabel Saludares

Every image tells a story



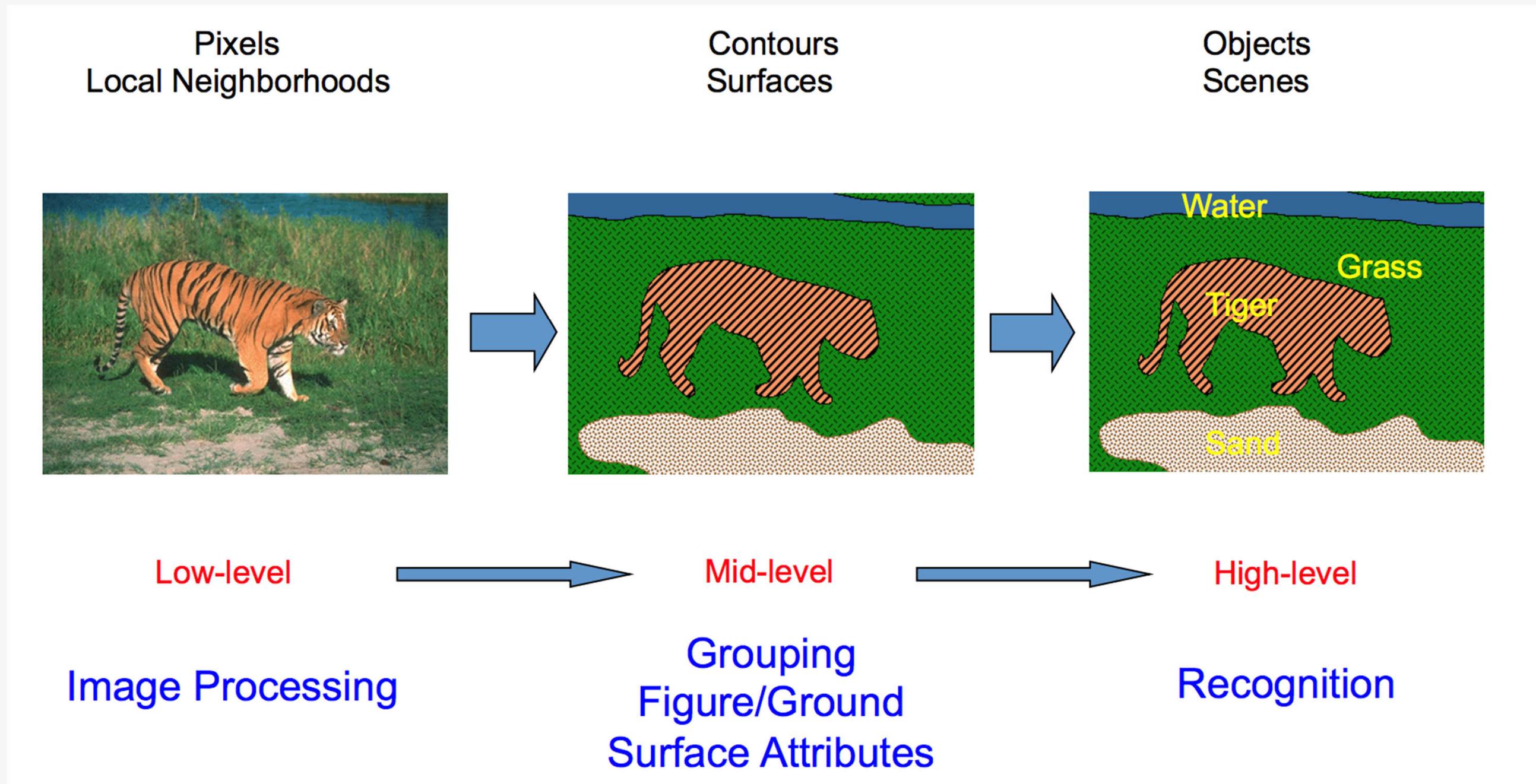
La Gare Montparnasse



(c) Isabel Saludares

- Computer vision goal: know the 'story' behind the picture
- Computer properties of the world (which was captured in an image)
 - 3D shape
 - names of people or objects
 - what is happening (action)

Computer vision approach (prior 2000)



Feed forward approach



Low Level Processing

simple operations such as image processing to reduce noise, contrast enhancement, and image sharpening.

*input: images
output: images*



Mid Level Processing

segmentation, description of those objects to reduce them to a form suitable for computer processing, and classification of individual objects.

*input: images
output: features extracted*



High Level Processing

How to make sense of a collection of recognized objects, mimic cognitive processes normally associated with vision.

*input: images
output: description, identified objects, etc.*

Low Level Processing

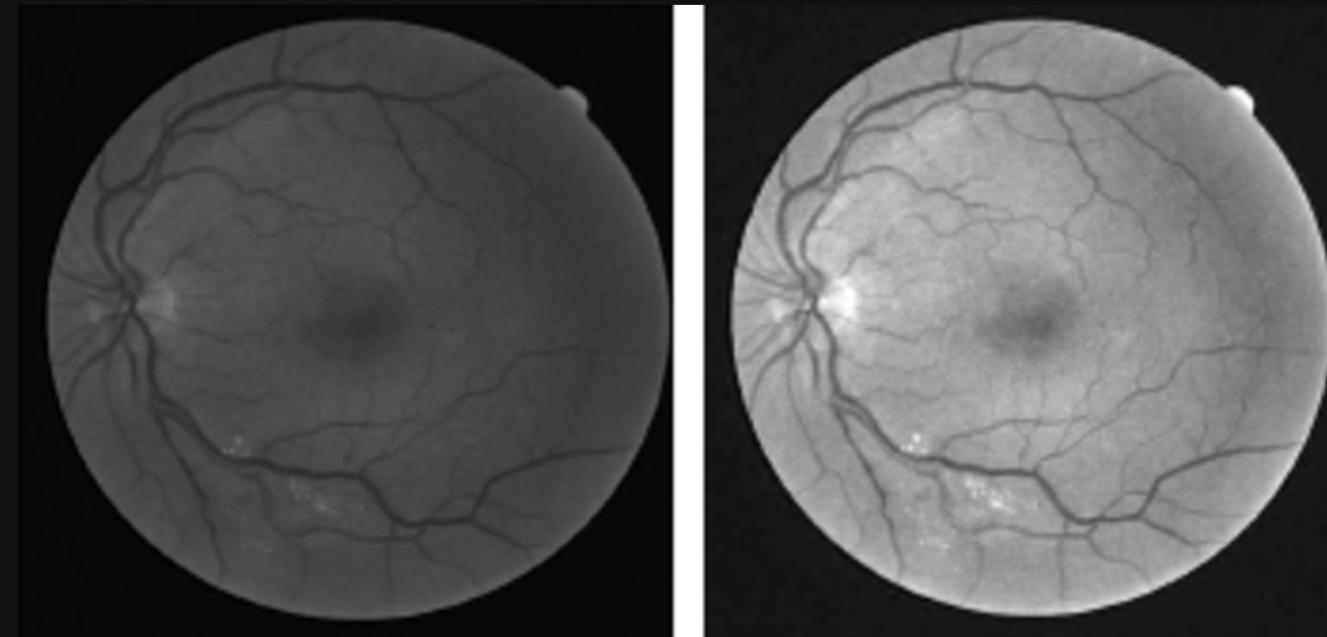


preprocessing for noise reduction [2]

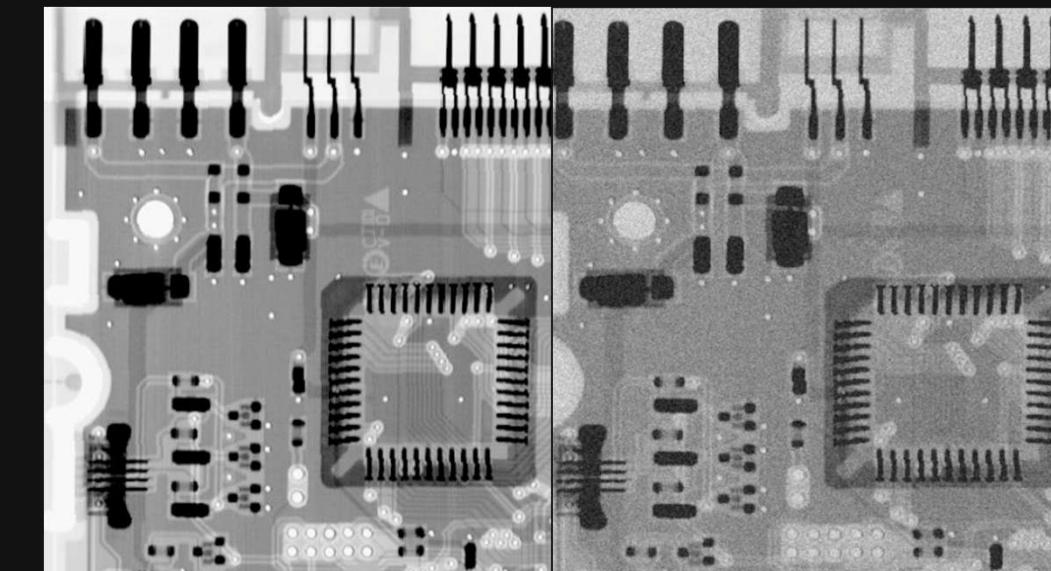
[1] Gonzalez and Woods 2002

[2] https://www.youtube.com/watch?v=A_2sk7QfH_Y

[3] Sagar, Anantha & Balasubramanian, Senthilkumar & Chandrasekaran, Venkatachalam. (2007). Automatic Detection of Anatomical Structures in Digital Fundus Retinal Images.. Proceedings of IAPR Conference on Machine Vision Applications, MVA 2007. 483-486.



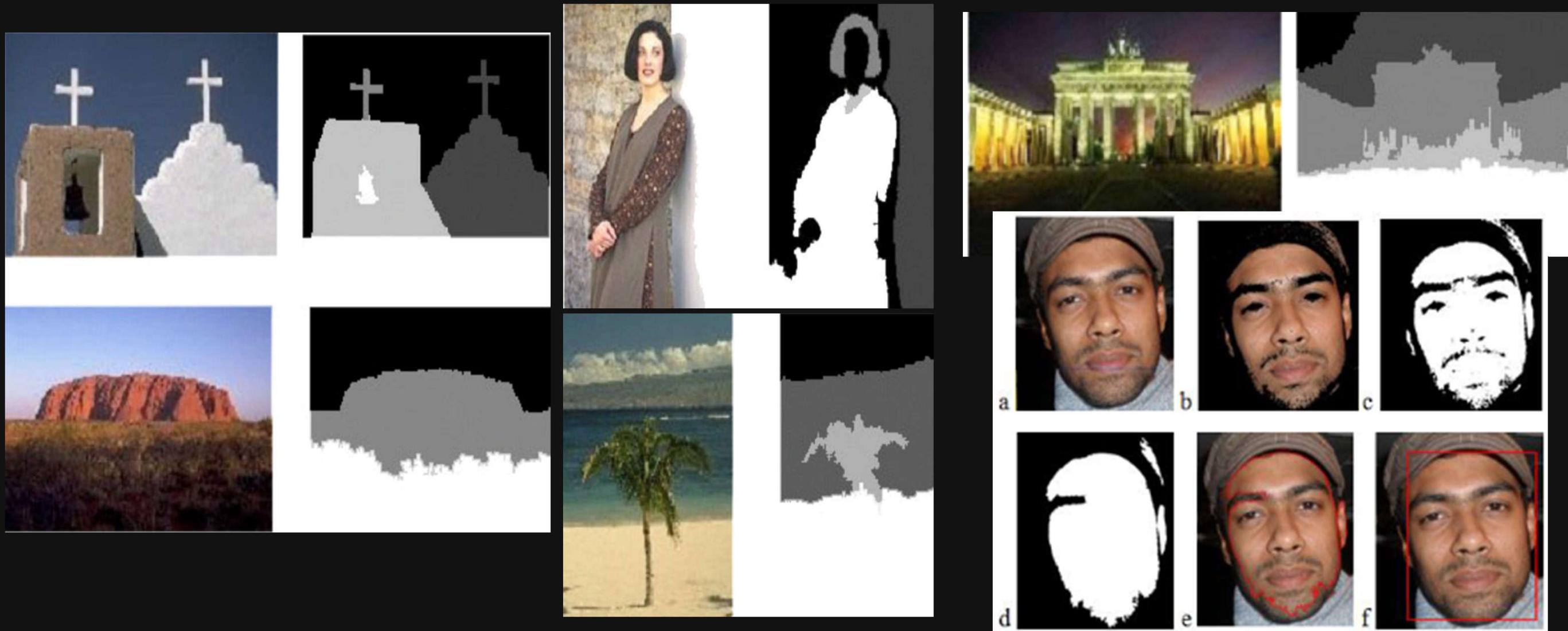
FGLG for contrast enhancement: (a) green channel of a test image. (b) image after contrast enhancement [3]



Noise reduction and contrast enhancement [1]

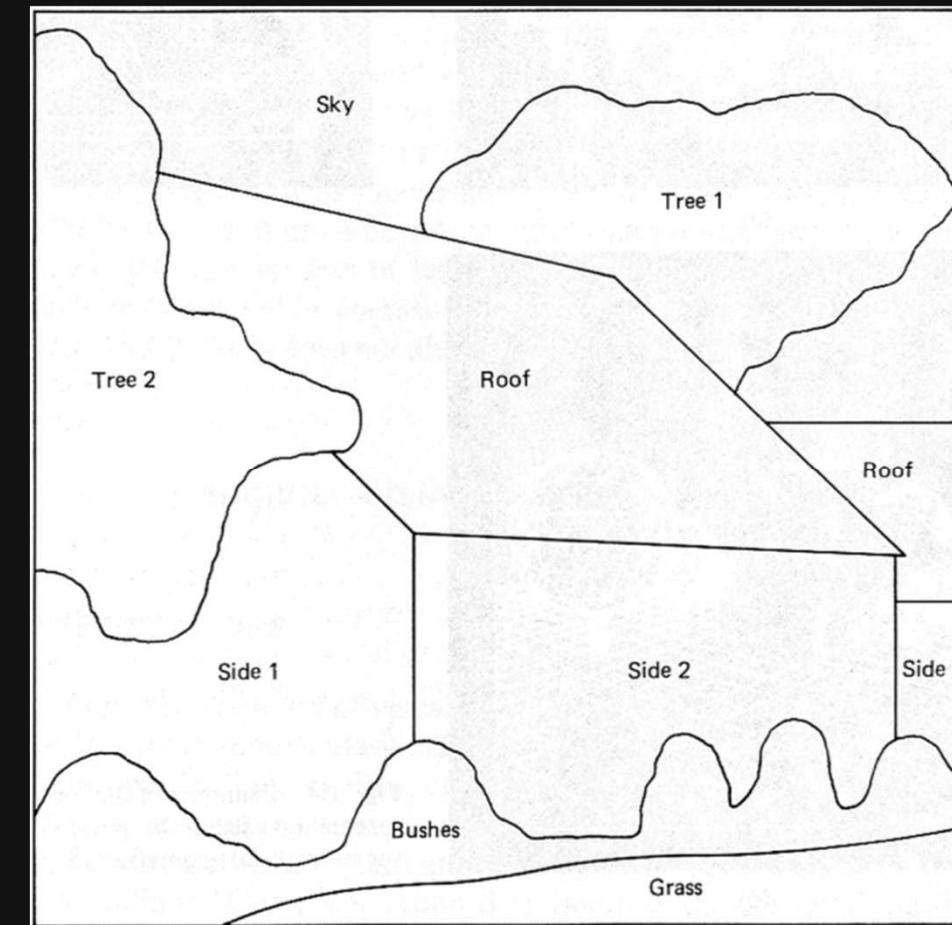
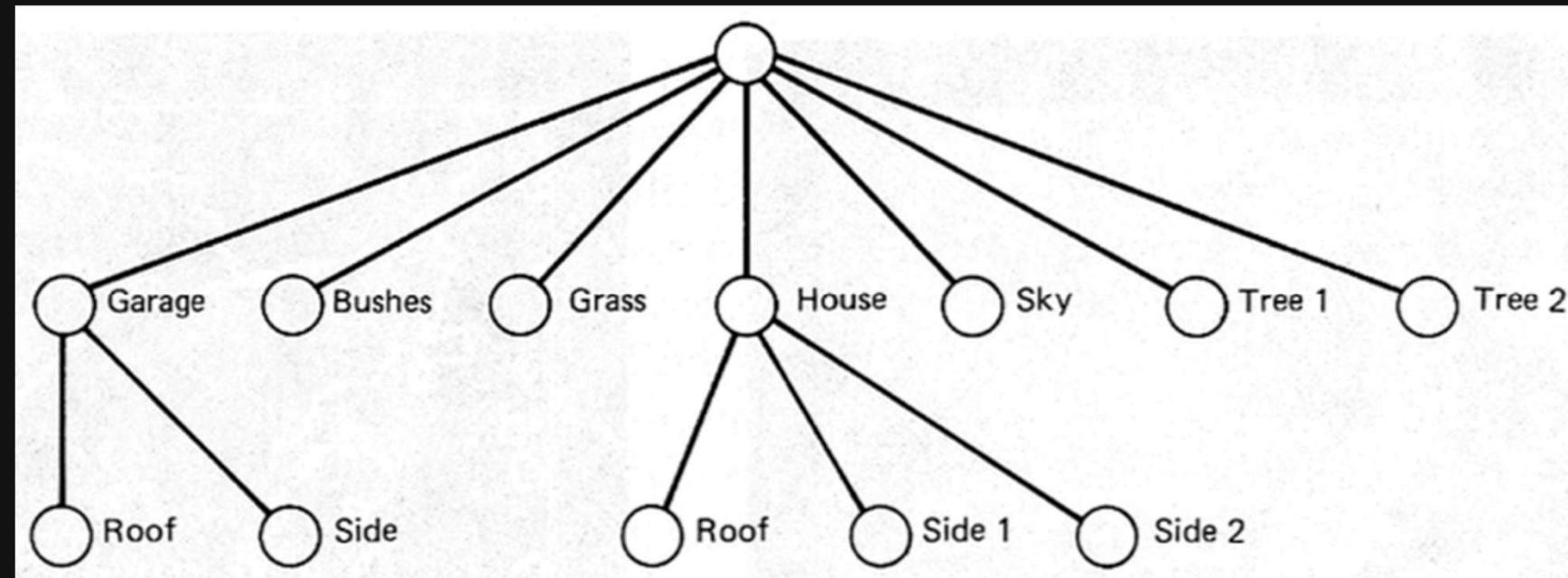
Mid Level Processing

- ♦ Input: image, Output: edges, contours, identity of objects
- ♦ Describe objects in image in a form suitable for automatic interpretation by computer
- ♦ segmentation: partitioning of image into regions or objects



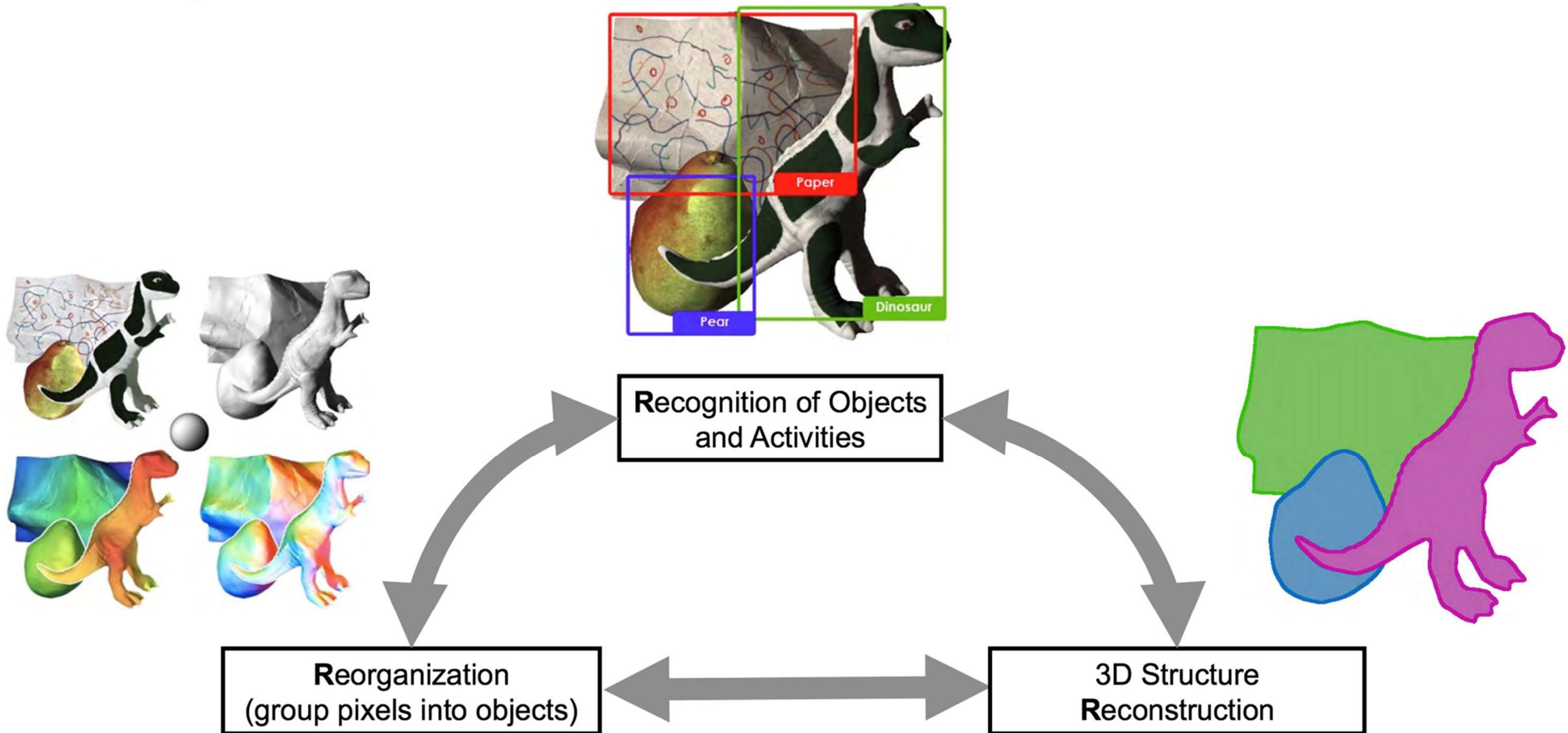
High Level Processing

- ♦ Input: output of mid-level vision, Output: task dependent
- ♦ "making sense" out of the ensemble of objects
- ♦ corresponds to the cognitive functions normally associated with vision



Recognition, Reconstruction, and Reorganization

Interacting CV Problems



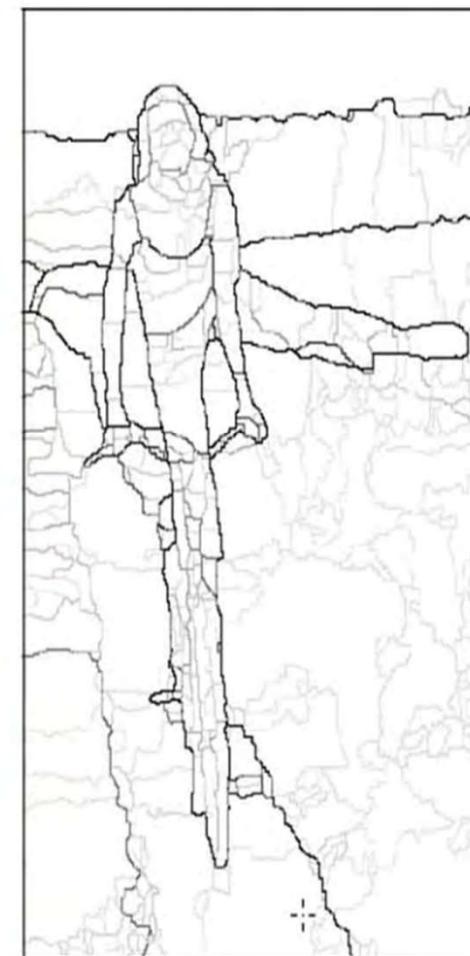
Reorganization -> Recognition

Interacting CV Problems

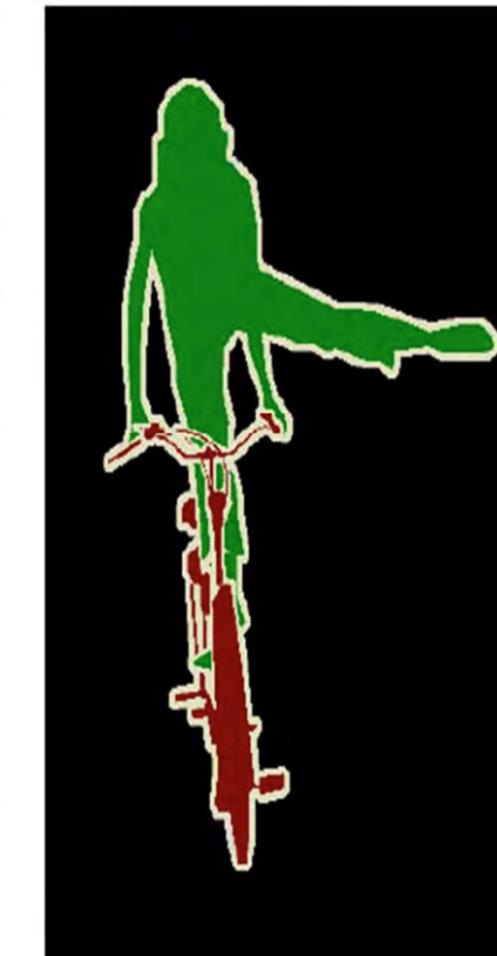
Original Image



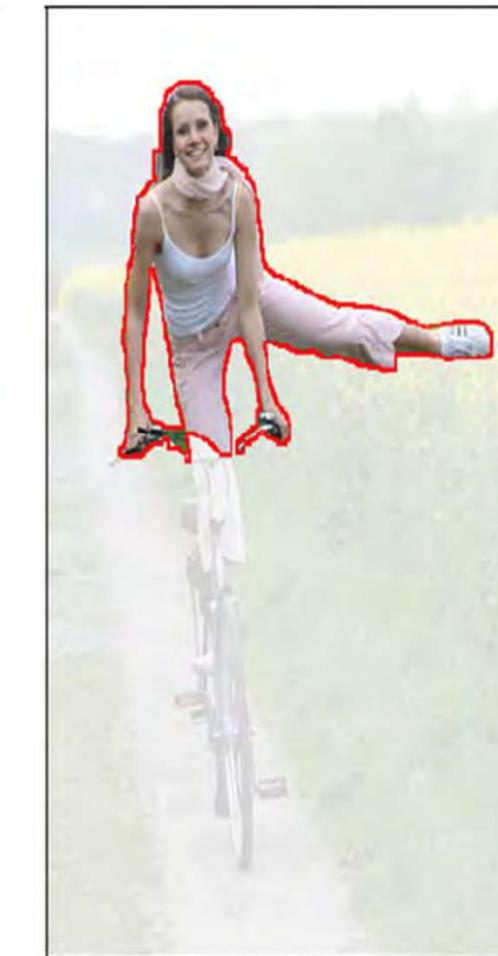
Multiscale hier.



Ground truth



MCG best candidates among 400



Recognition -> Reorganization

Interacting CV Problems



Recognition -> Reconstruction

Interacting CV Problems

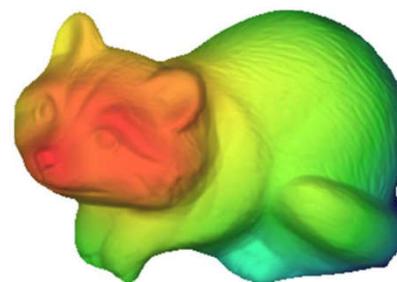
Shape illumination reflectance from shading (SIRFS)

Input:



Image

Output:



Shape



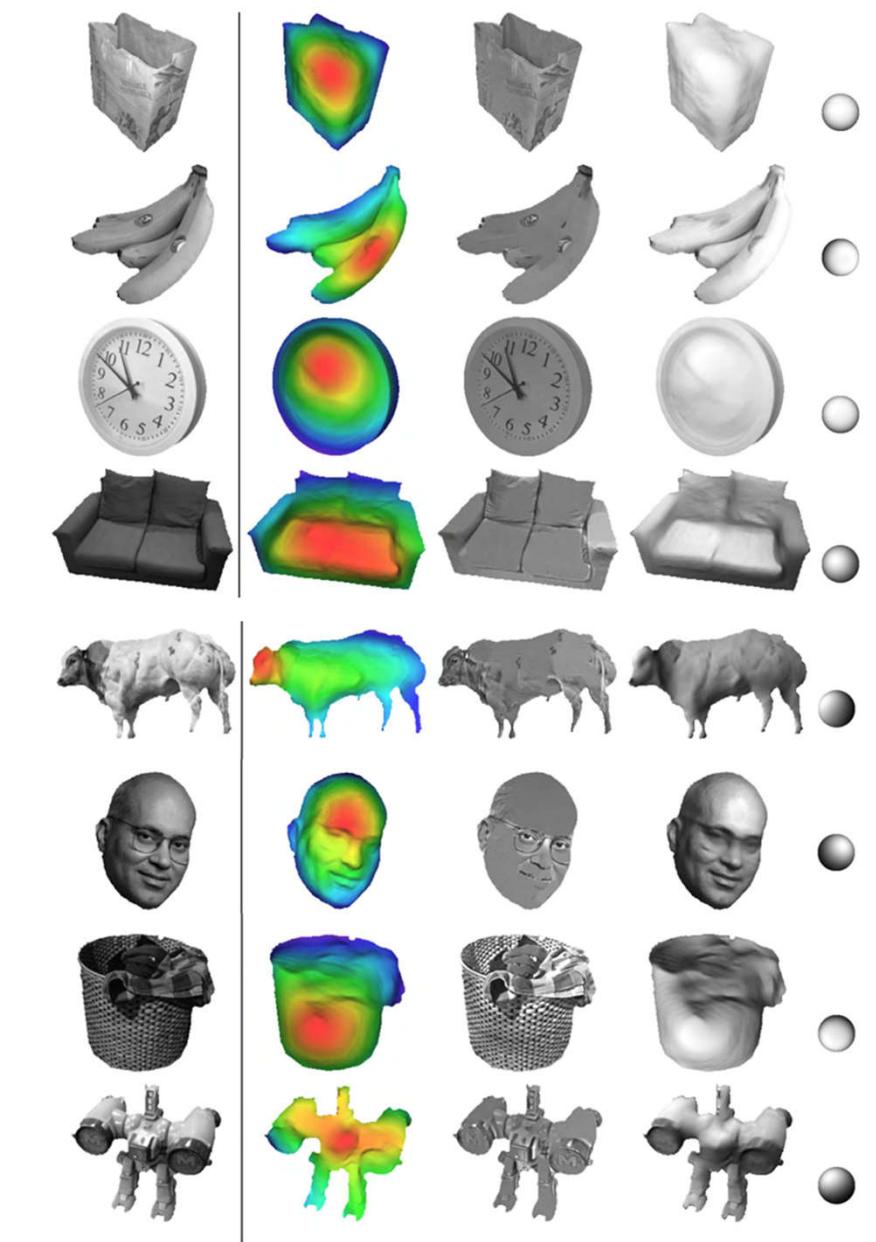
Reflectance



Shading



Illumination

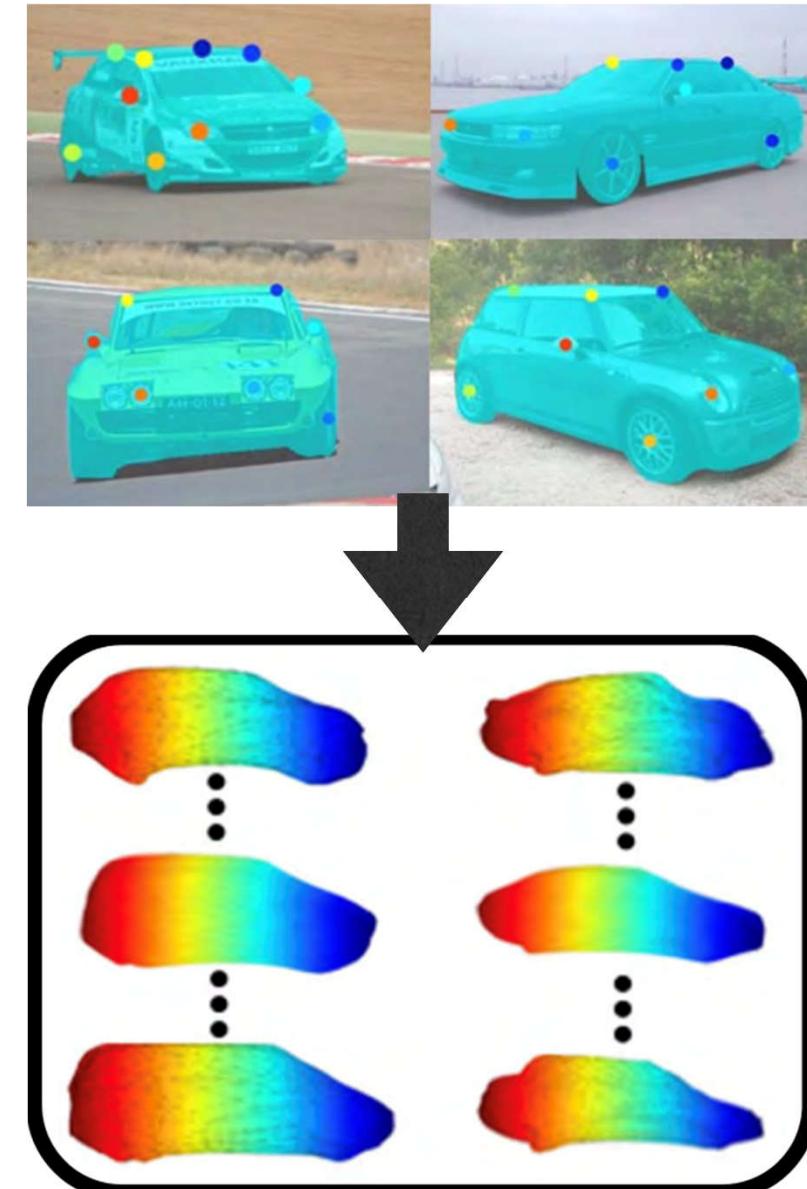
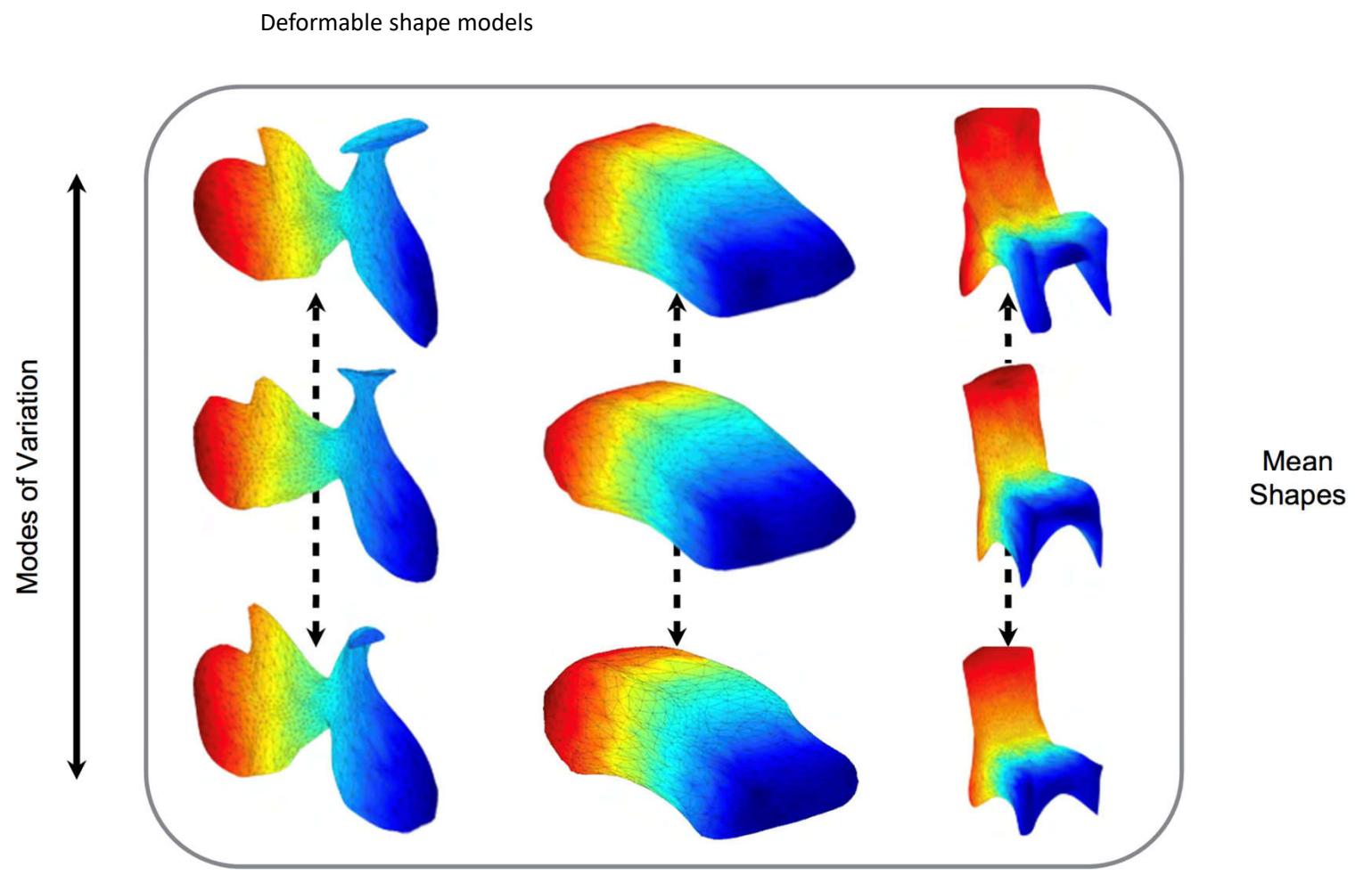


Barron & Malik, CVPR 2011, CVPR 2012, ECCV 2012, PAMI 2015

Recognition -> Reconstruction

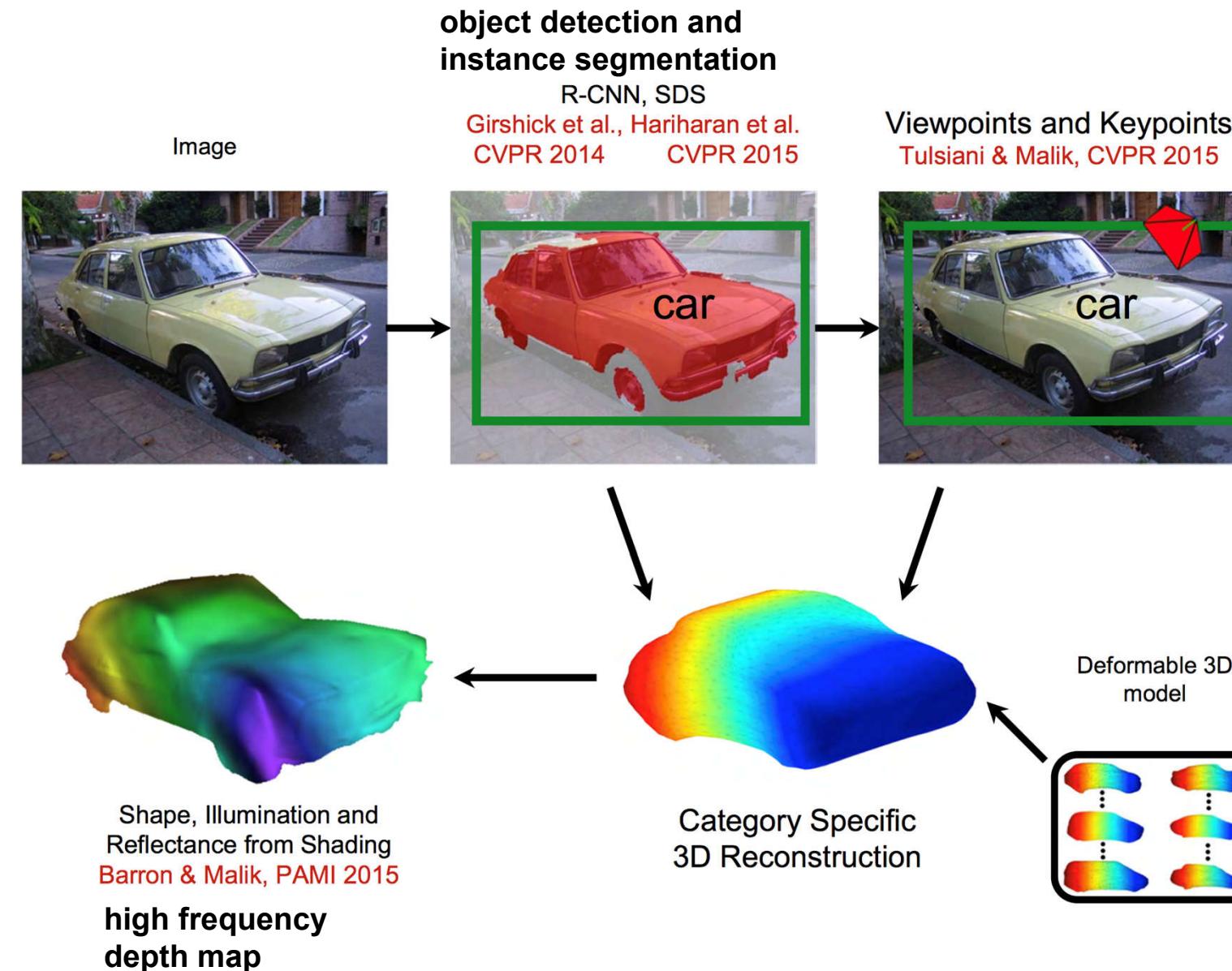
Interacting CV Problems

Learn category specific 3D models from 2D images of objects



Recognition -> Reconstruction

3D reconstruction of objects from a single image of a scene

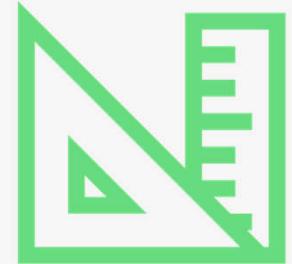


Building intelligent systems

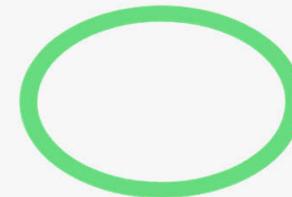
General framework when building IS for institutions

01 Designing

Understanding the users and the needs through designing the expected use and identifying the technologies to be used

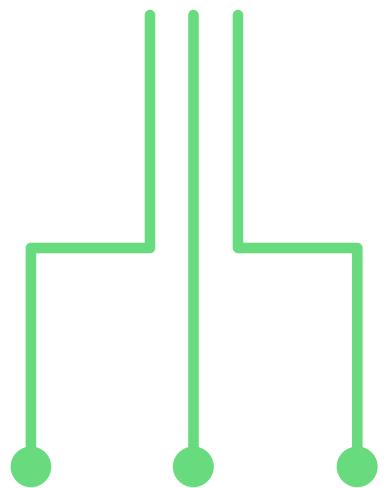


02 Proof-of-Concept (PoC)

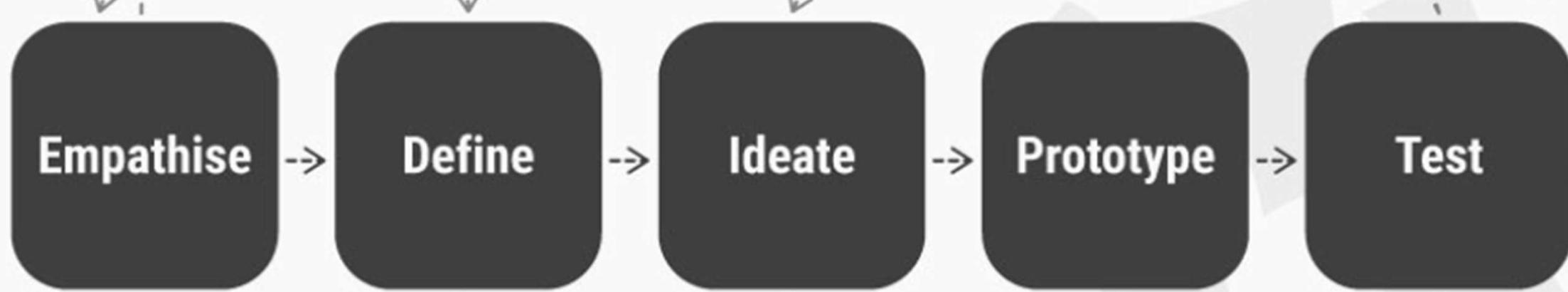


03 Integrate and Scale





Design Thinking: a non-linear process



Learn about users
through testing

Empathise to help
define the problem

Tests create new
ideas for the project

Ideate

Prototype

Test

Learn from
prototypes to spark
new ideas

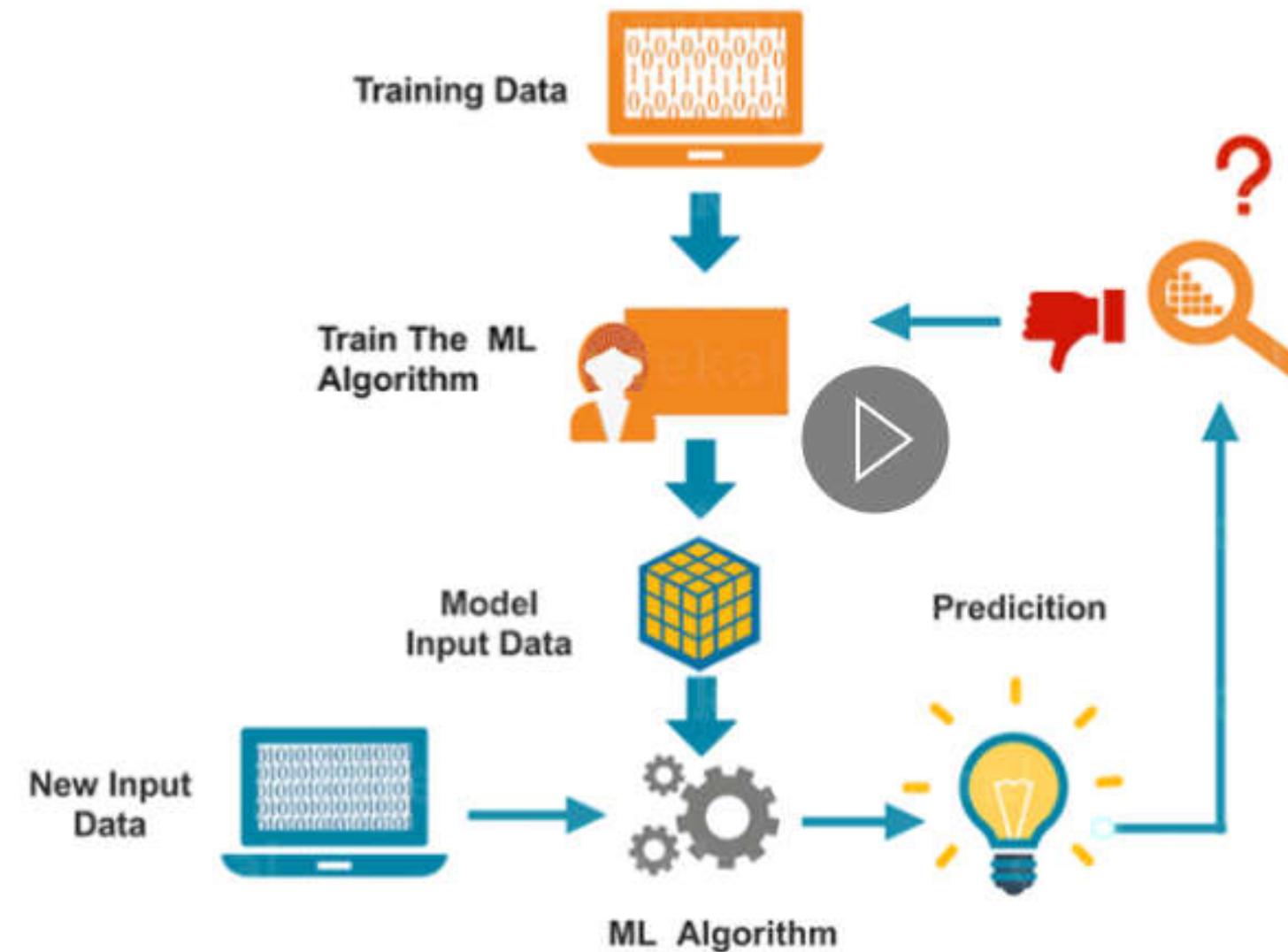
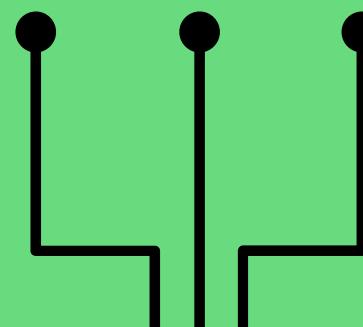
Tests reveal insights that
redefine the problem



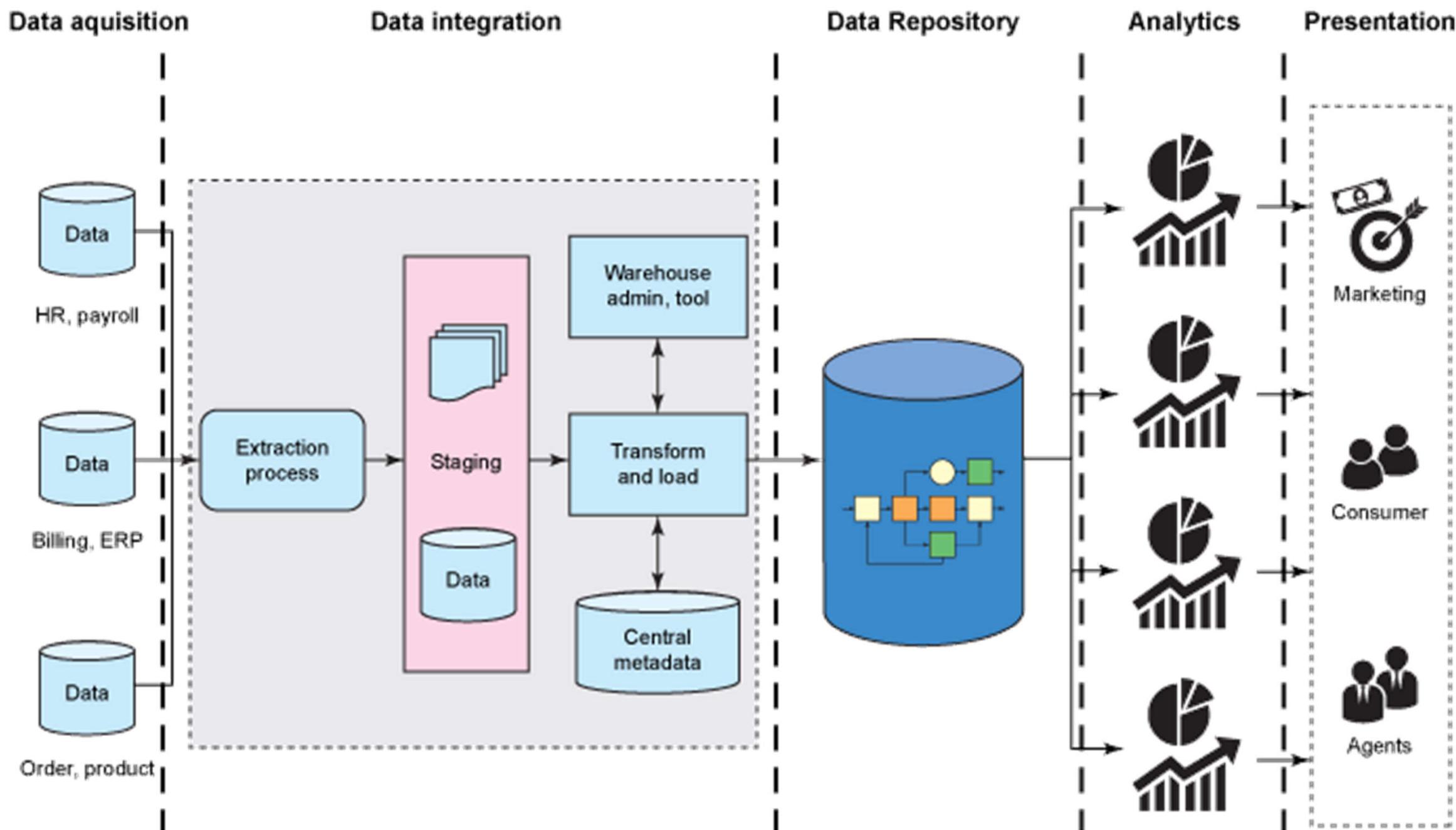
INTERACTION DESIGN
FOUNDATION

INTERACTION-DESIGN.ORG

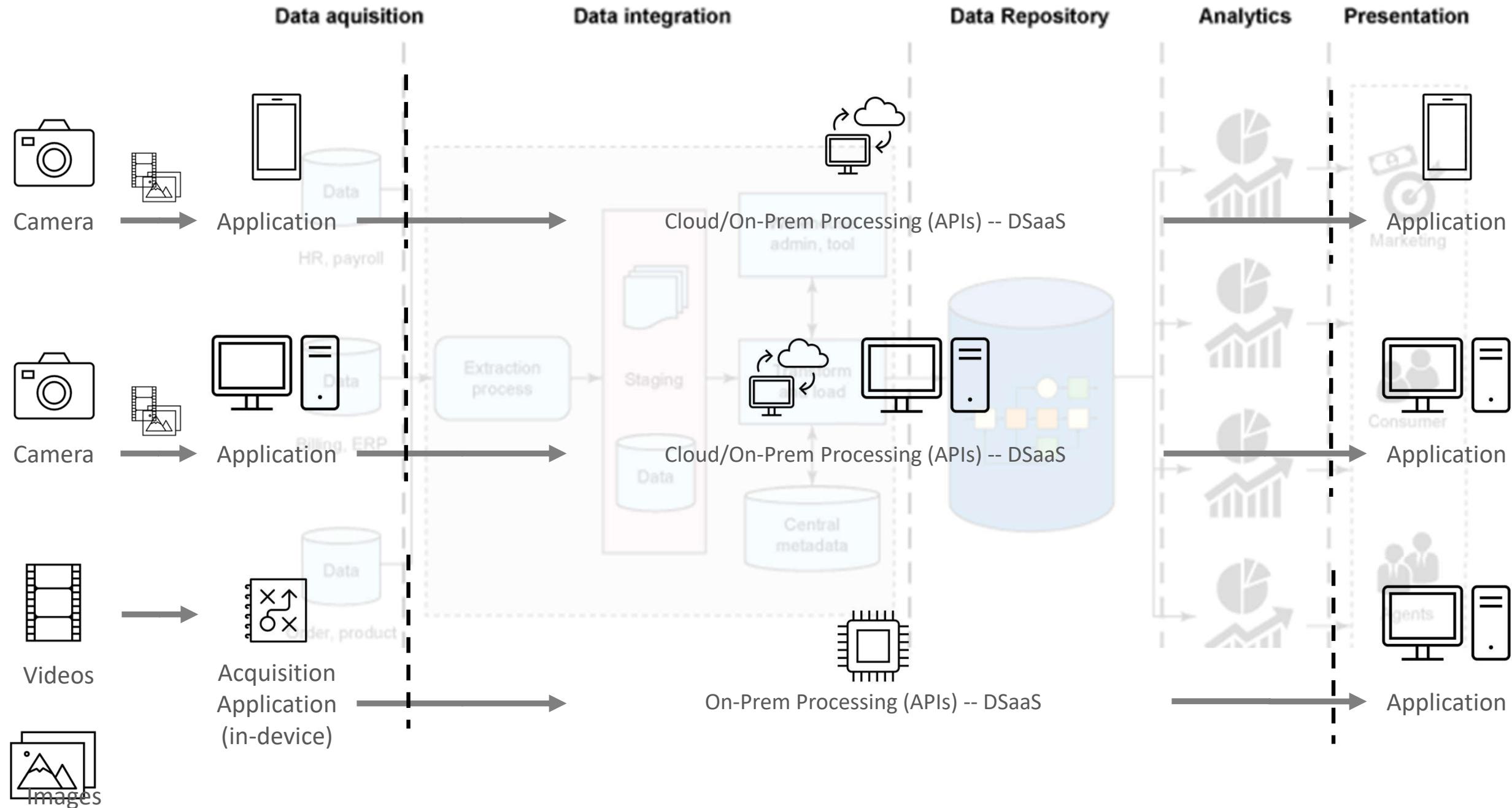
Proof of Concept (PoC)



General Framework



General Framework



Components of a standard CV system

Acquisition

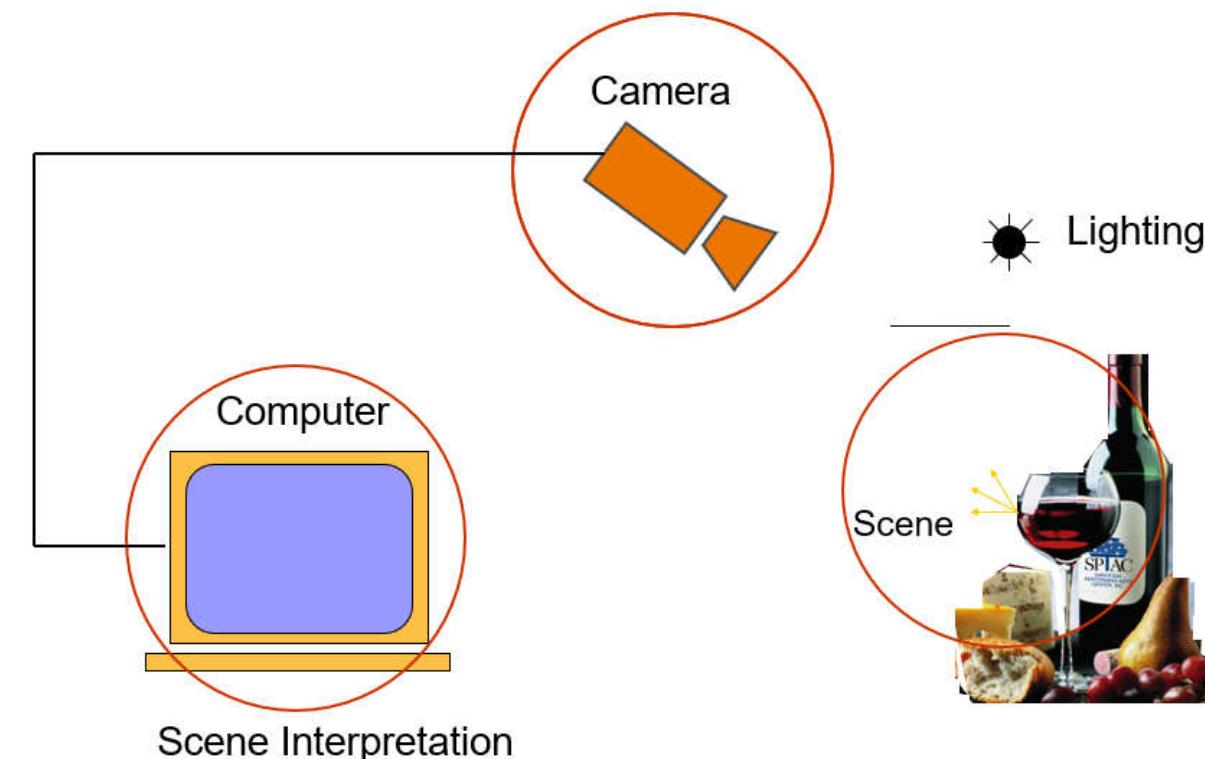
- Camera
- Lighting devices
- Lens + Frame grabber

Data Integration and Processing

- Processing Software

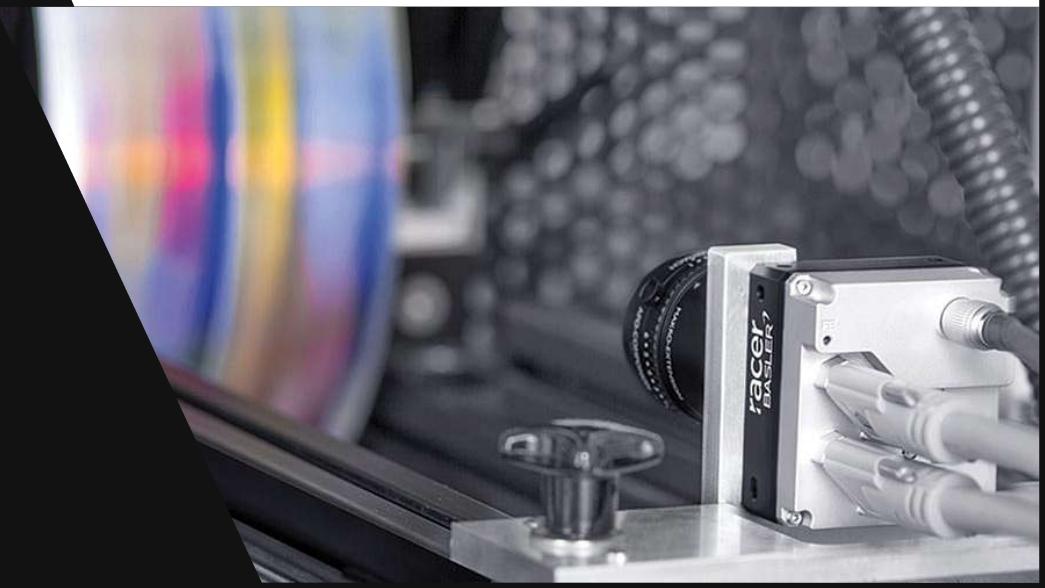
Presentation

- Application (display screen, user-interface)



Acquisition Camera

Machine Vision vs Surveillance
Area scan or Line scan
Monochrome vs Color
Global or rolling shutter
Frame rate
Resolution
Size



Acquisition Camera

Machine Vision vs Surveillance

Area scan or Line scan

Monochrome vs Color

Global or rolling shutter

Frame rate

Resolution

Size



Acquisition Camera

Machine Vision vs Surveillance

Area scan or Line scan

Monochrome vs Color

Global or rolling shutter

Frame rate

Resolution

Size



Acquisition Camera

Machine Vision vs Surveillance

Area scan or Line scan

Monochrome vs Color

Global or rolling shutter

Frame rate

Resolution

Size

Acquisition Camera

Machine Vision vs Surveillance

Area scan or Line scan

Monochrome vs Color

Global or rolling shutter

Frame rate

Resolution

Size



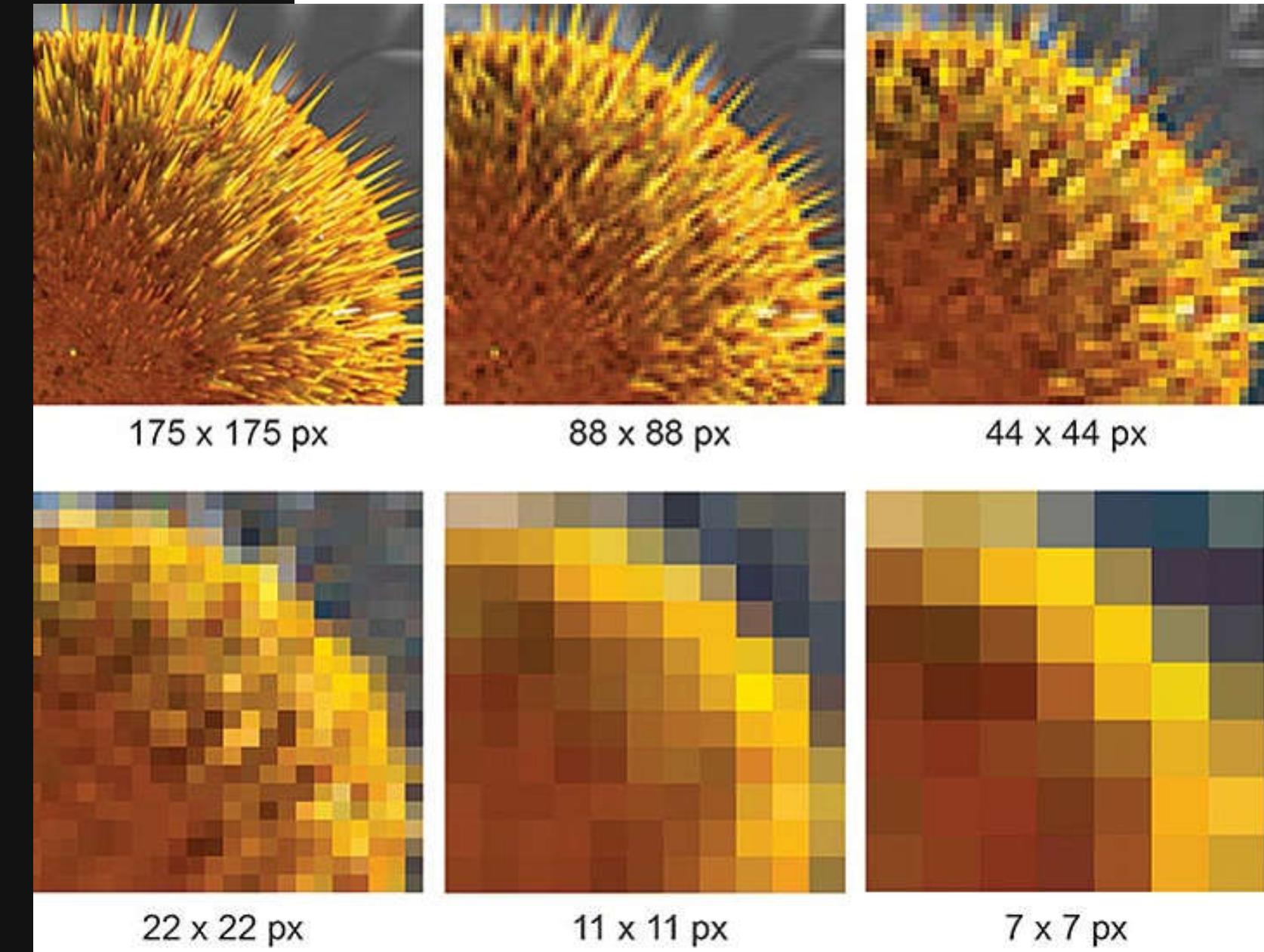
Acquisition Camera

Machine Vision vs Surveillance
Area scan or Line scan
Monochrome vs Color
Global or rolling shutter
Frame rate
Resolution
Size



Acquisition Camera

Machine Vision vs Surveillance
Area scan or Line scan
Monochrome vs Color
Global or rolling shutter
Frame rate
Resolution
Size



Acquisition Camera

Machine Vision vs Surveillance

Area scan or Line scan

Monochrome vs Color

Global or rolling shutter

Frame rate

Resolution

Size



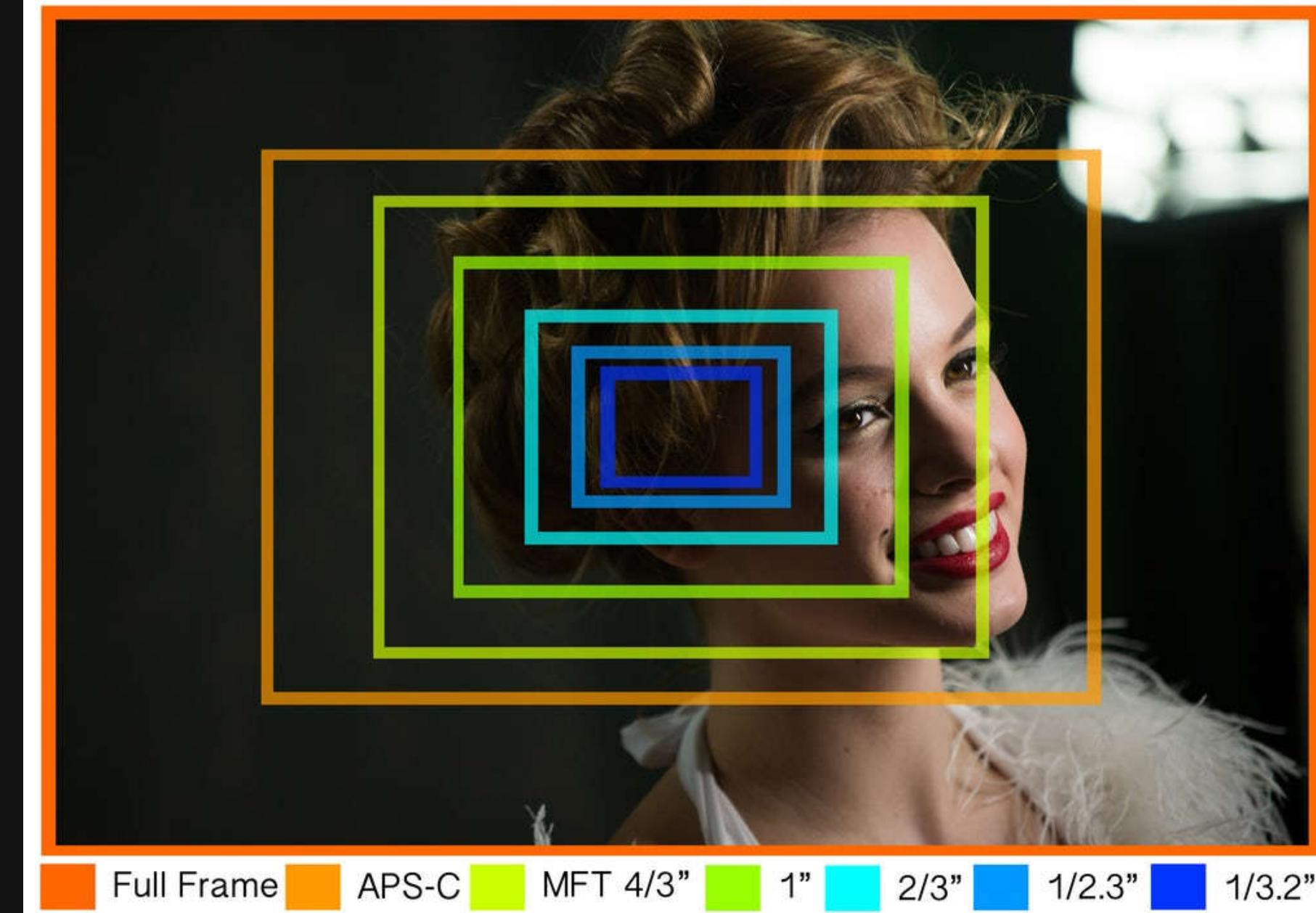
Acquisition

Lens

Lens Size

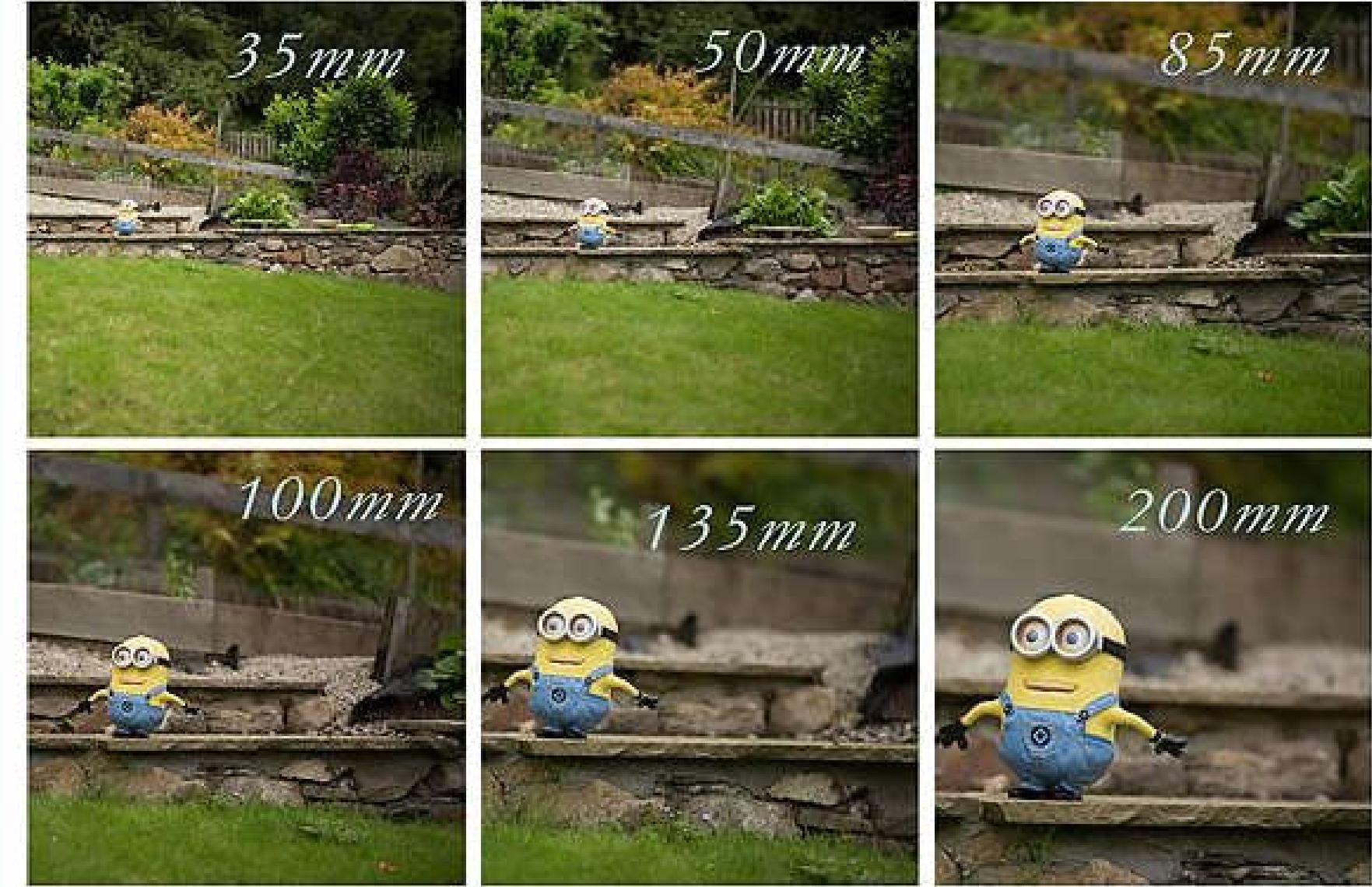
Focal Length

Wide Angle or not



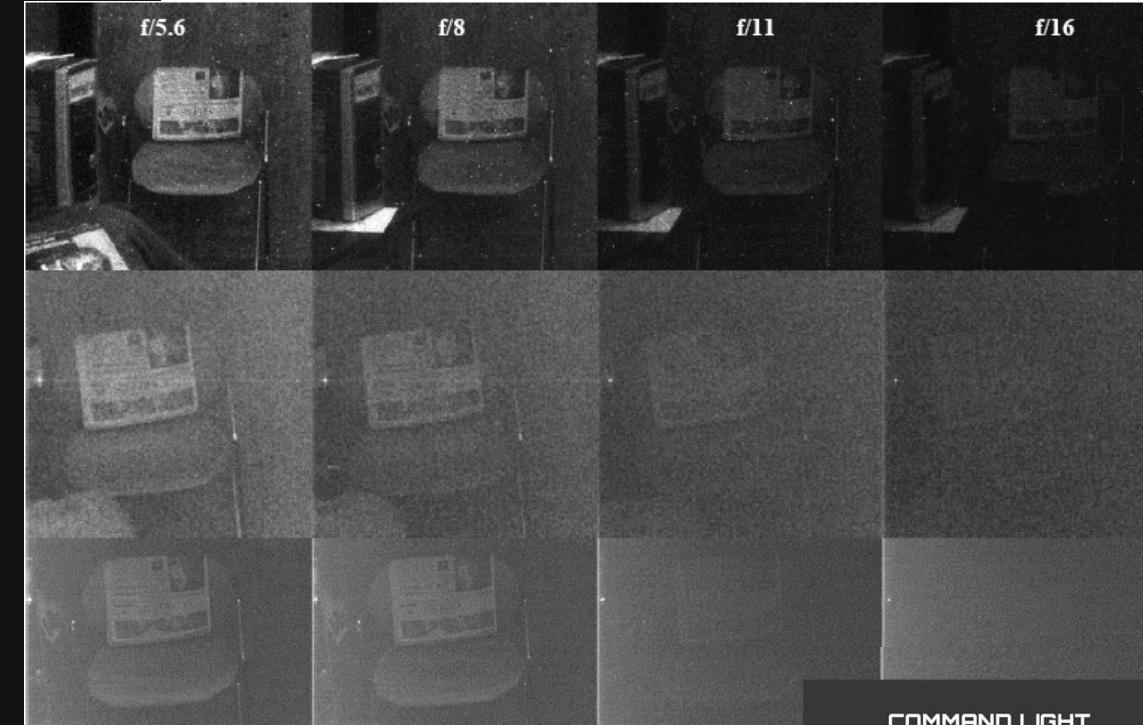
Acquisition Lens

Lens Size
Focal Length
Wide Angle or not

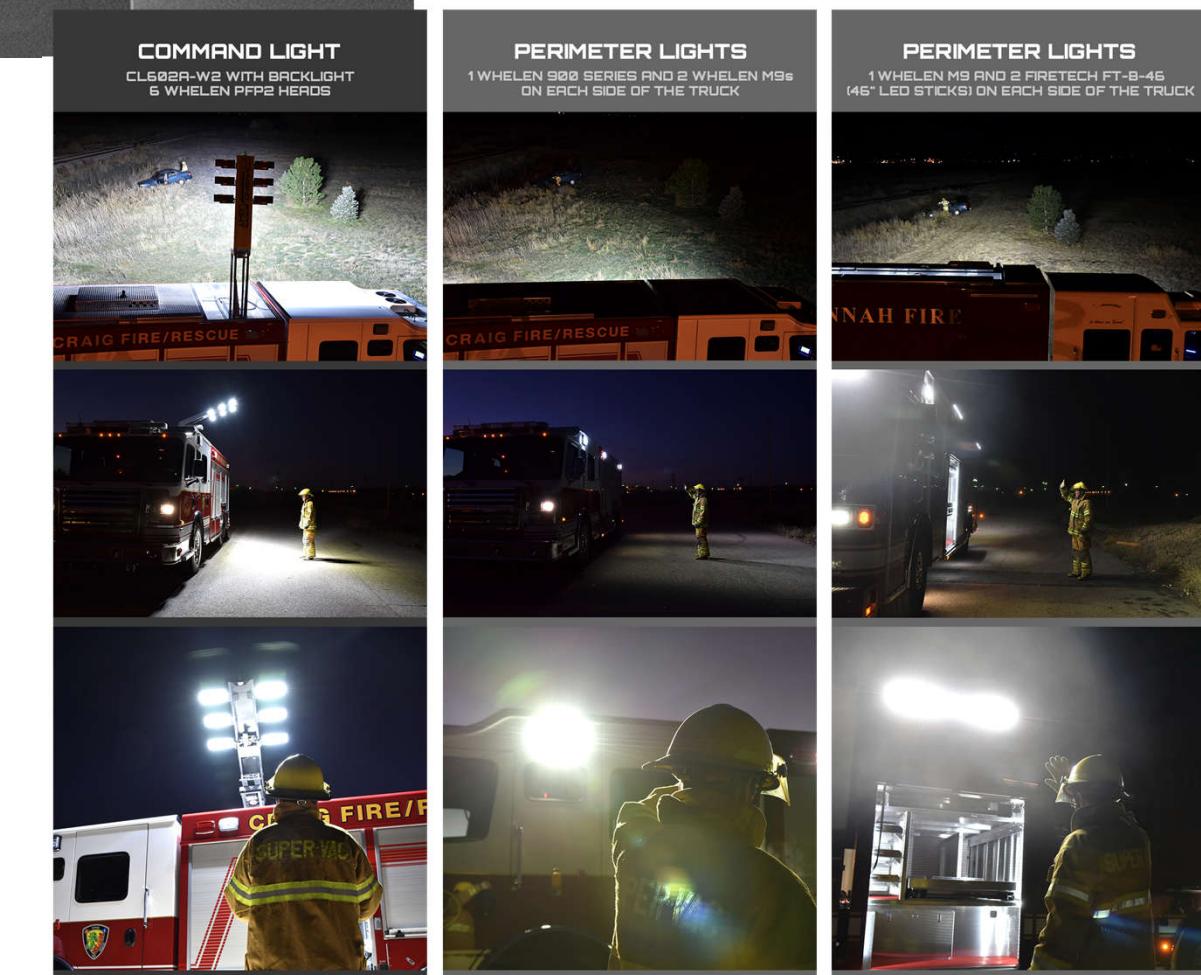


Acquisition Lens

Lens Size
Focal Length
Wide Angle or not

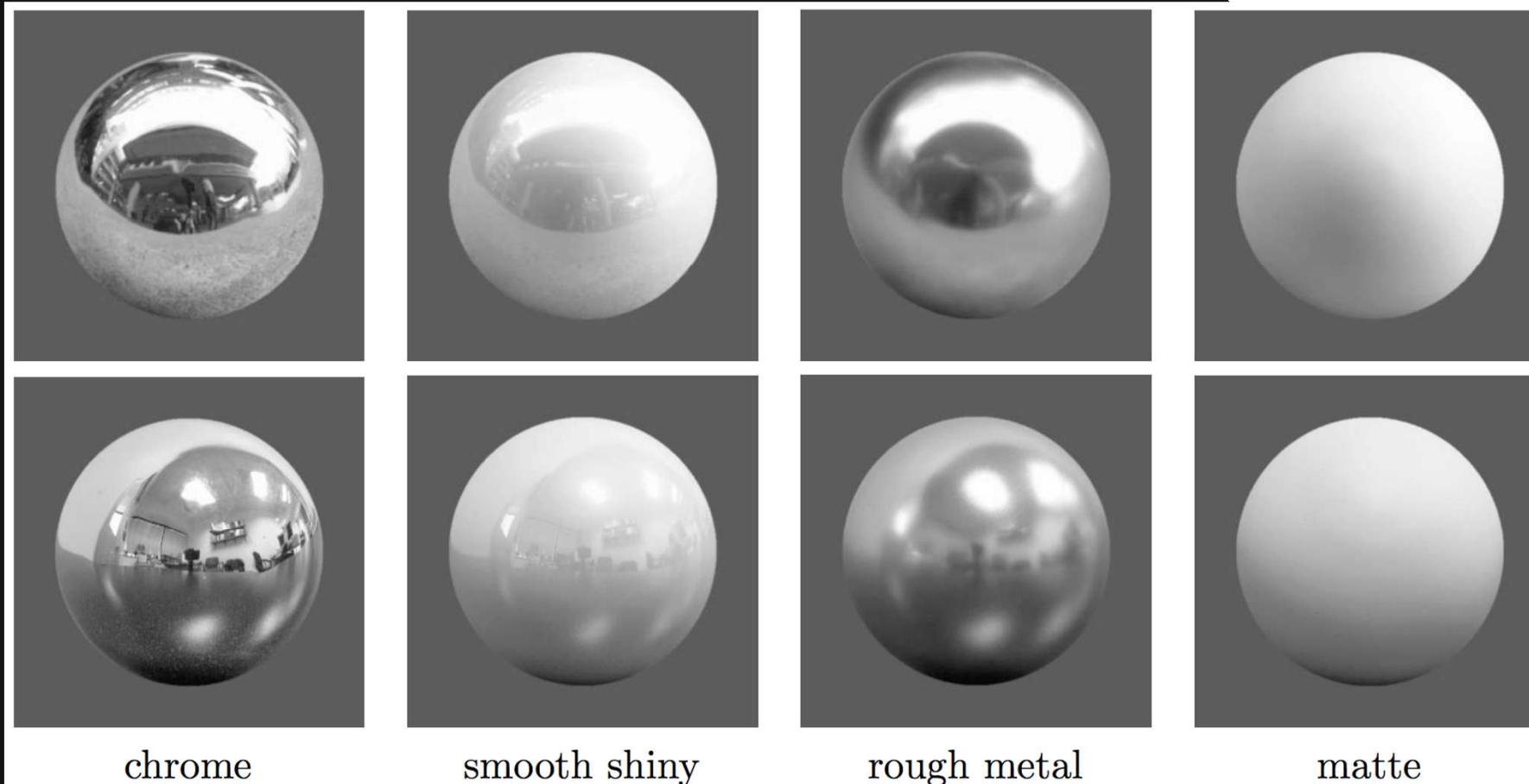


EMCCD and ICCD Camera
comparison in low ambient
light conditions



Lighting Devices

Lens



The two images in each column are photographs of the same sphere, shown over a standard gray background. The images in each row were photographed in the same location, under the same illumination.

Frame grabber

Lens

Single vs Multiple Cameras

Cable vs wireless

Interface	Cable Lengths	Bandwidth max. in MB/s.	Multi-Camera	Cable Costs	“Real-time”	“Plug & Play”
 GiG Vision	100 m	100	■	■	■	■
 USB Vision	8 m	350	■	■	■	■
 Camera Link	10 m	850	■	■	■	■

Processing Software-Hardware

HOW TECH HELPS COMPANIES EVOLVE AND INNOVATE



PC-Based

- **Flexibility:** number of options: line scan versus an area scan camera with the PC., use third party software packages vs smart cameras tend to be single source software
- **Power:** greater power and speed -- Intel processors used internally, used to handle the 'tougher' applications



Embedded

- **Stand-alone box:** limited flexibility
- application-specific routines
- camera is **tethered** to unit rather than self-contained
- have the ability to handle **multiple cameras**



Smart Camera

- **Cost:** include the camera, lenses, lighting (sometimes), cabling and processing.
- **Simplicity:**
- **Integration**
- **Reliability**

Processing Software-Hardware

HOW TECH HELPS COMPANIES EVOLVE AND INNOVATE

	<i>PC-based</i>	<i>Smart Camera</i>
<i>Flexibility</i>	Excellent	Poor
<i>Ruggedness</i>	Poor	Excellent
<i>Size</i>	Multiple-box system Imaging head can be very small	All-in-one box Not necessarily very small
<i>Functionality</i>	Expandable	Limited
<i>Performance</i>	Expandable	Limited
<i>Ease of use</i>	Needs computer skill	No computer skill needed

Processing Software



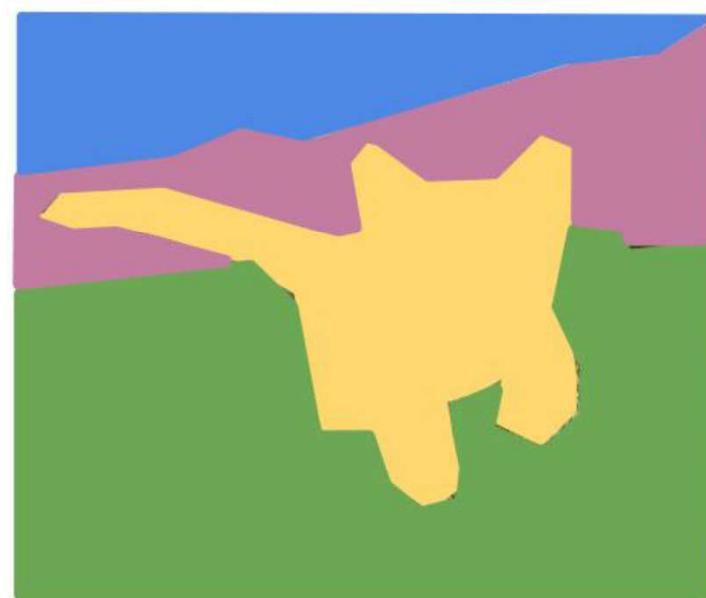


Pattern Recognition

Machine learning Algorithms for Pattern Recognition

Avoid compromising the security of the platforms you use within the company by investing heavily in cybersecurity. Secure your company's data by looking into the best platforms out there that can protect against cybercriminals, hackers, and other fraudulent acts.

Recognition: Object detection



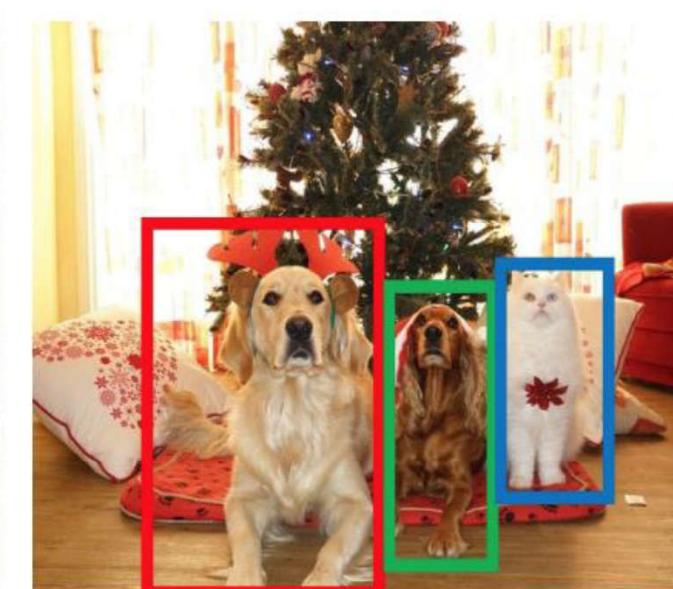
GRASS, CAT,
TREE, SKY

No objects, just pixels



CAT

Single Object



DOG, DOG, CAT

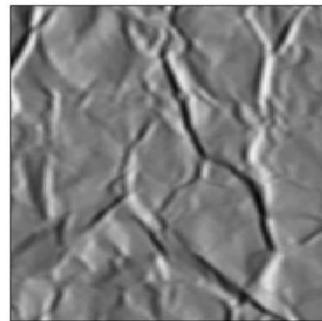
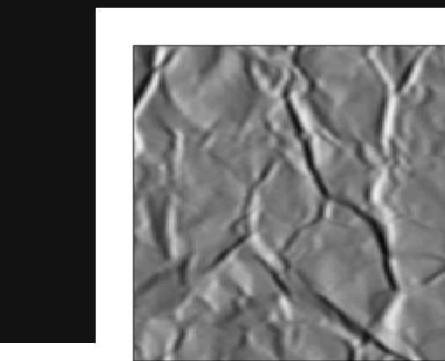
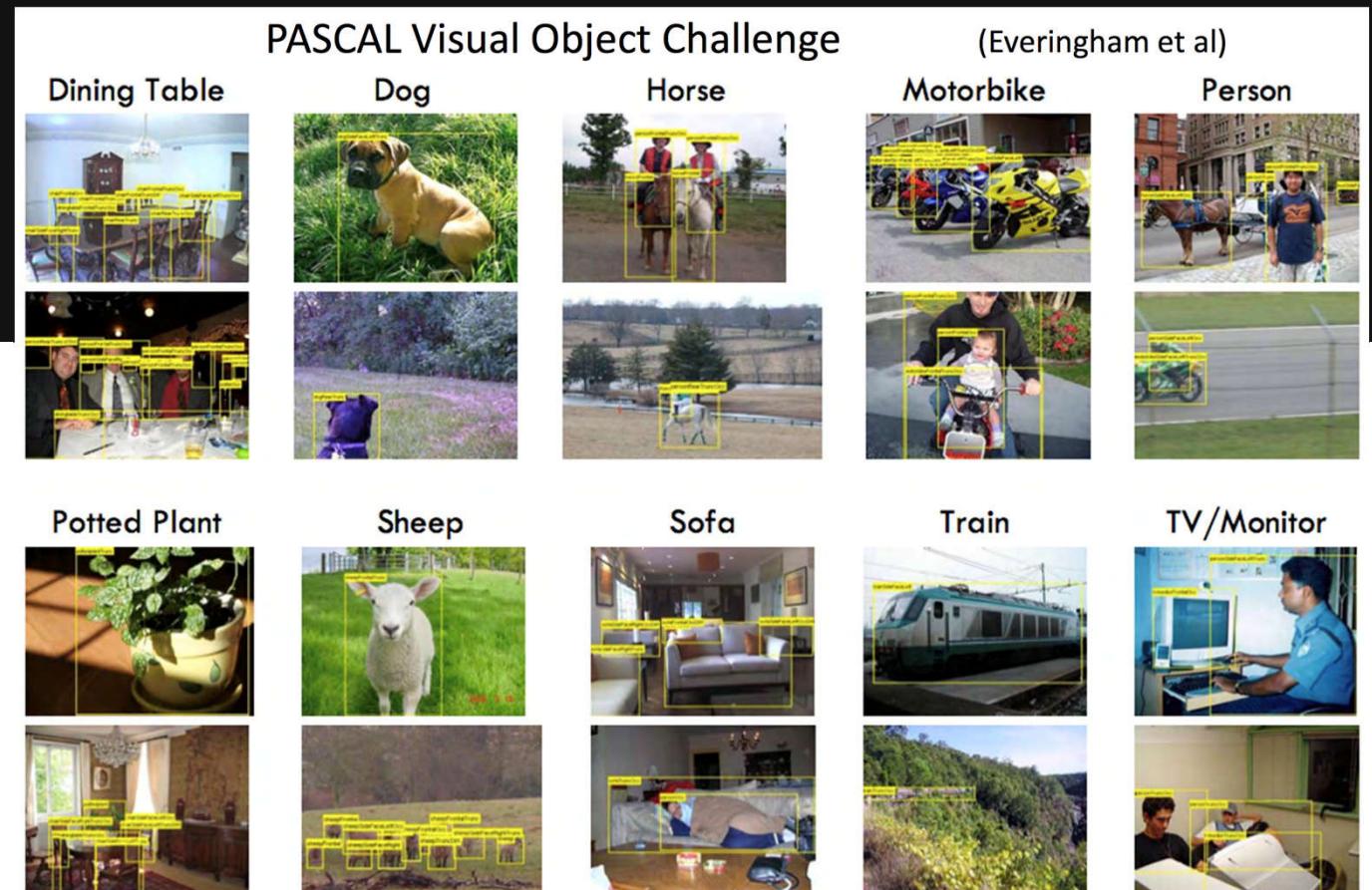
Multiple Object



DOG, DOG, CAT

[This image is CC0 public domain](#)

Recognition: Object detection



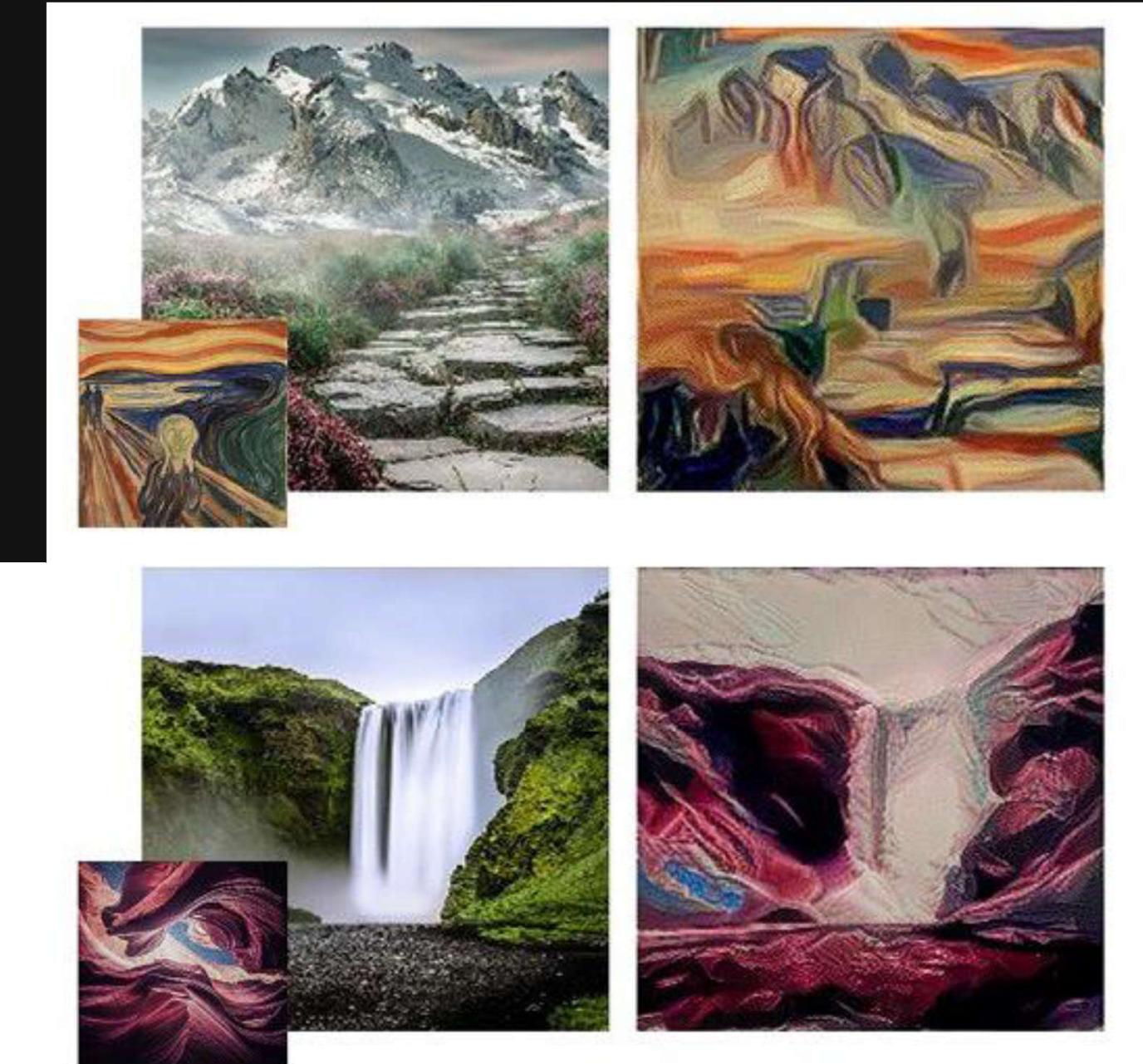
Felt?
Polyester?
Terrycloth?
Rough Plaster?
Leather?
Plaster?
Concrete?
Crumpled Paper?
Sponge?
Limestone?
Brick?



⋮



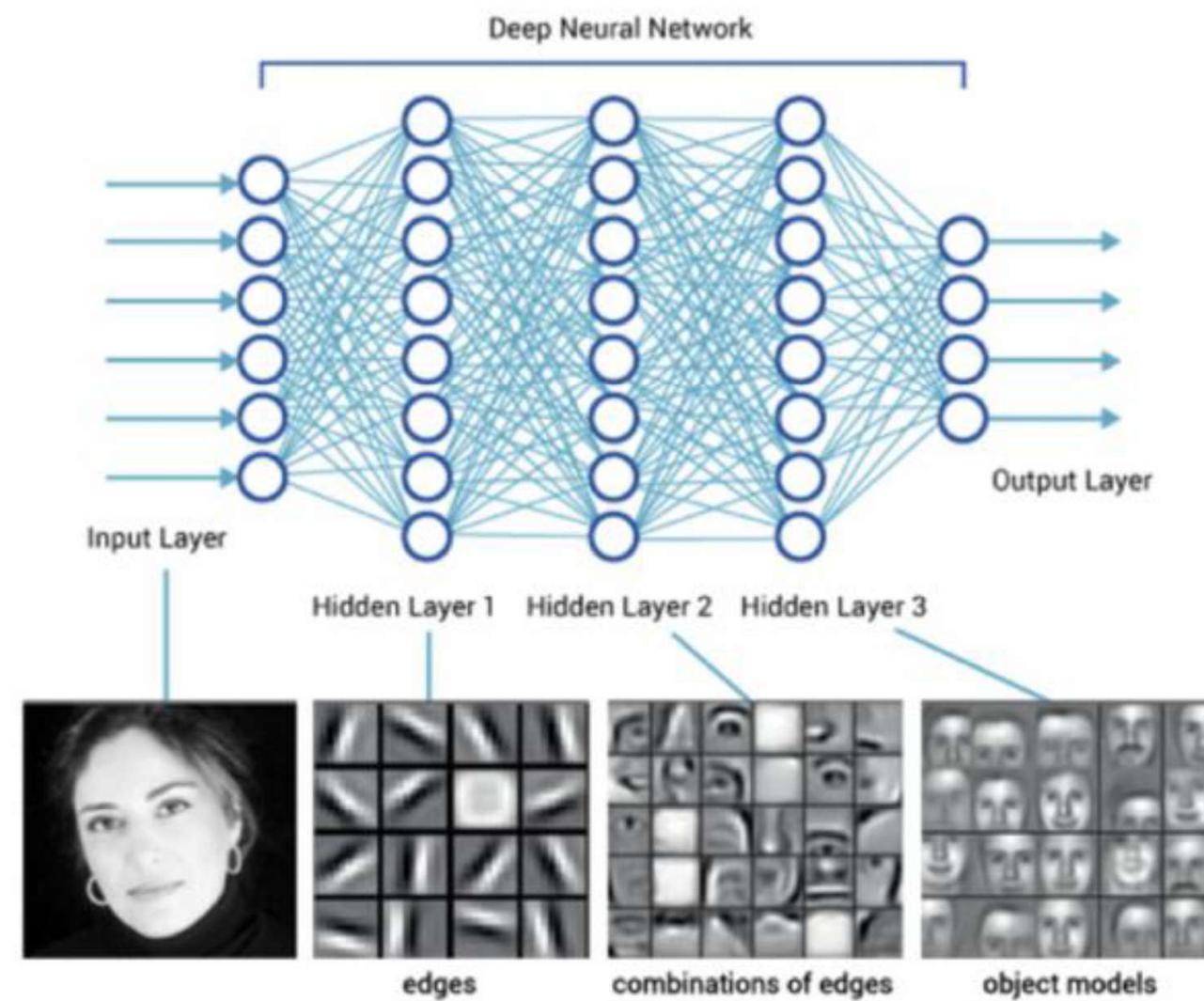
discrimination



Style transfer

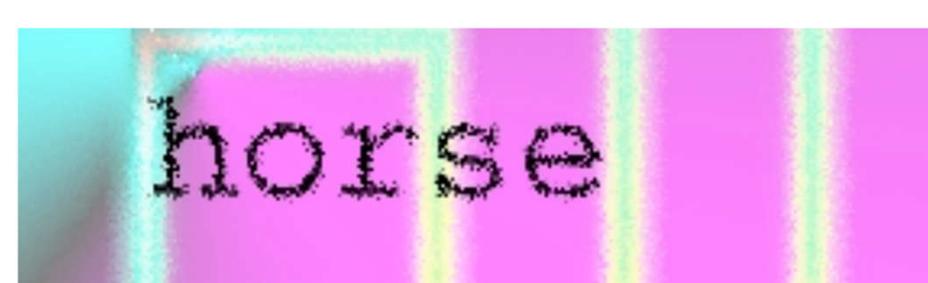
Recognition: Face

0	1	4
3	5	8
1	7	3



Recognition: Text

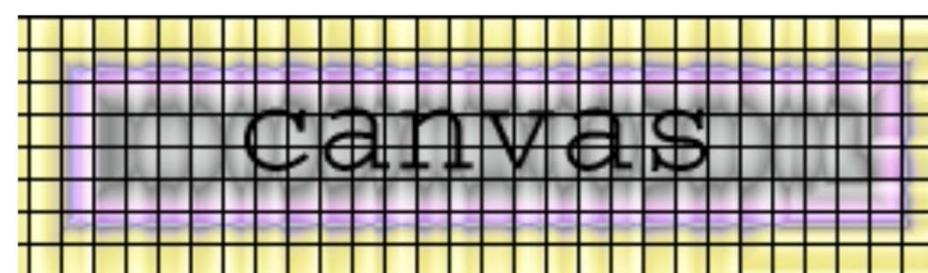
171 of 192



horse



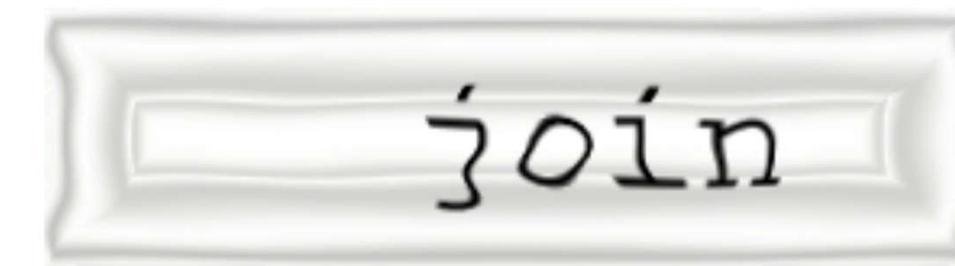
smile



canvas



spade



join



here

Yahoo's EZ-Gimpy Captcha -- 171 of 192 images correctly identified: 92 %

Recognition: Action



phoning

running



walking

ridinghorse

Display and User-Interface



Display and User-Interface

Display screen or a robotic arm to carry out an instruction obtained from image interpretation.



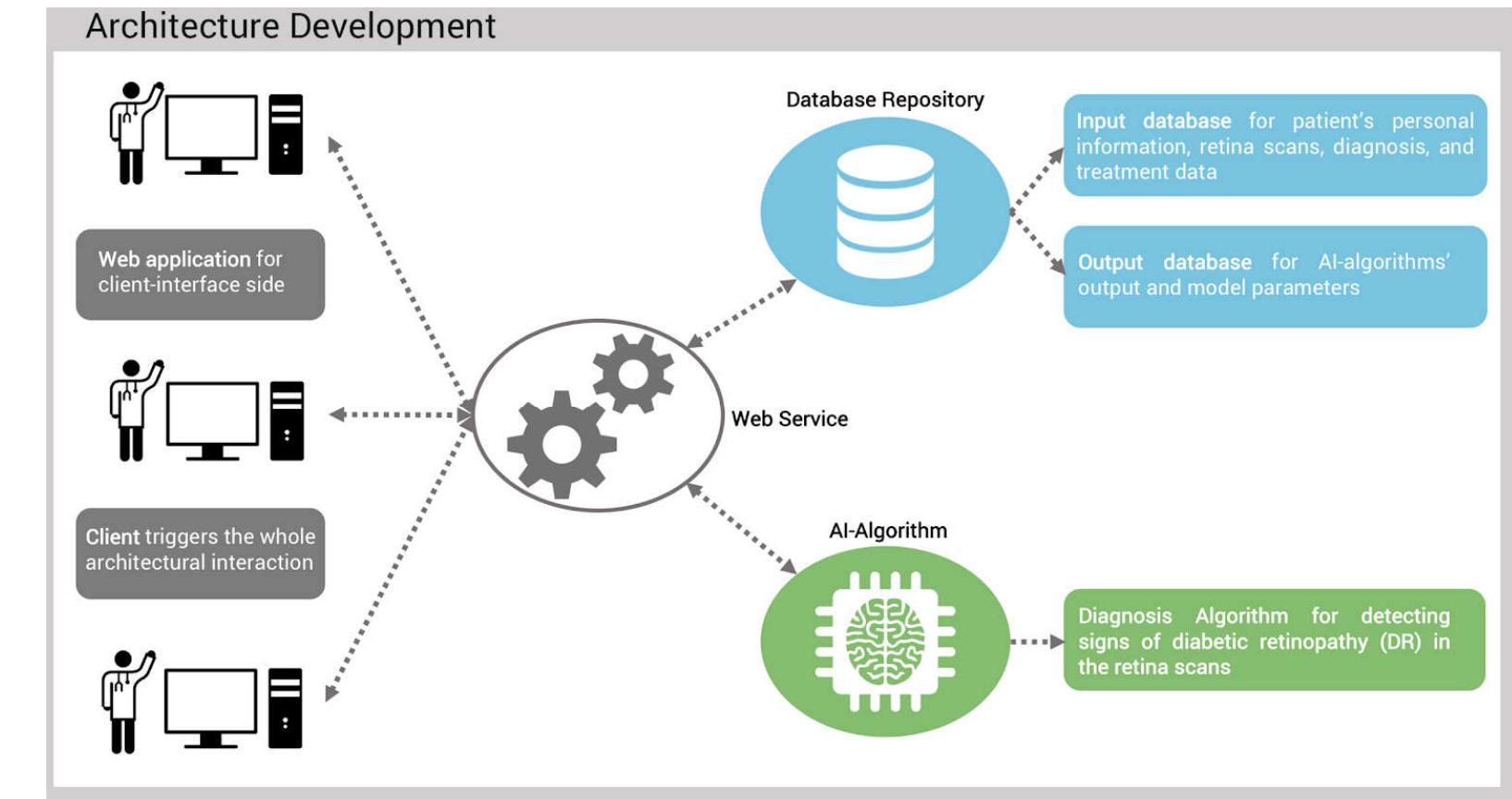
Tesla's driverless system detecting objects in a foggy scenario

Display and User-Interface



Display and User-Interface

Dashboards: summary of findings



A PILOT STUDY ON ARTIFICIAL INTELLIGENCE-POWERED MEDICAL IMAGE ANALYSIS FOR BREAST CANCER DETECTION

Medical Image Analysis





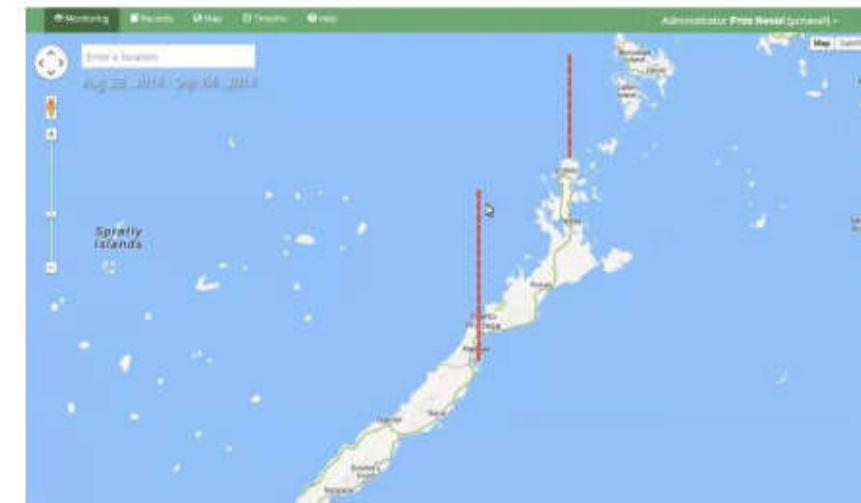
Highly Secured Encryption
AES-128 & RSA-256



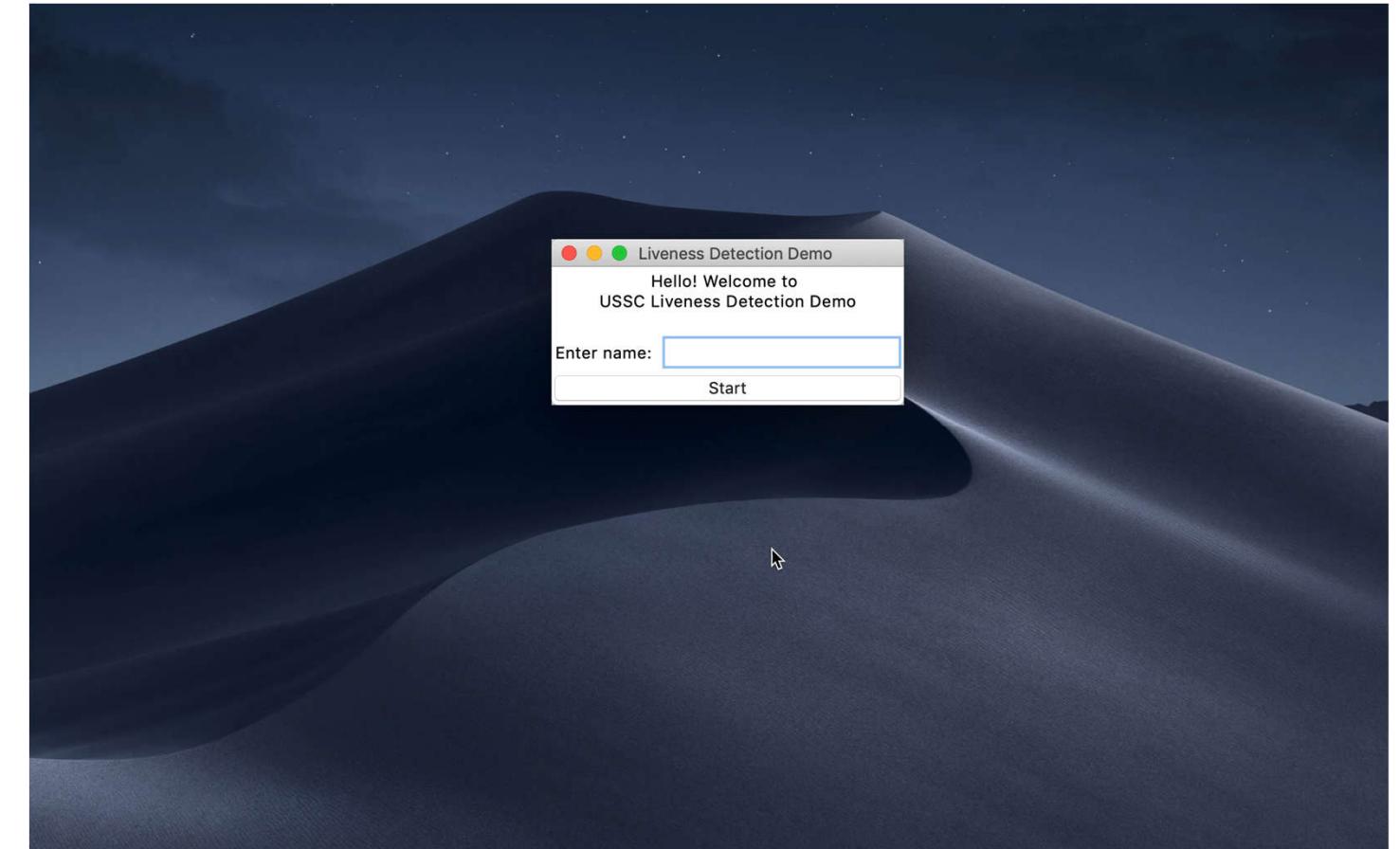
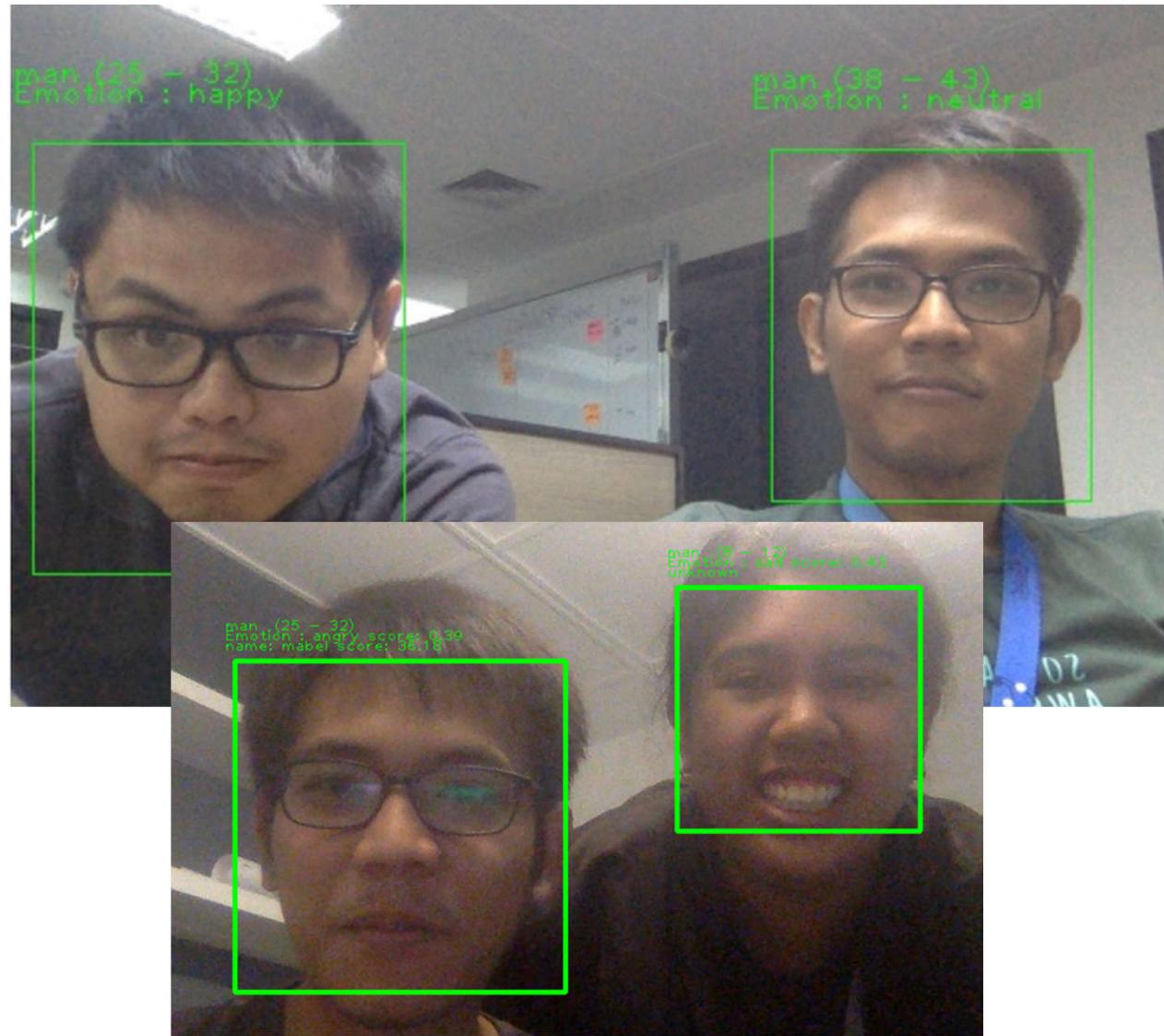
**Data Integrity over Poor
Connectivity**



"Fire and Forget"



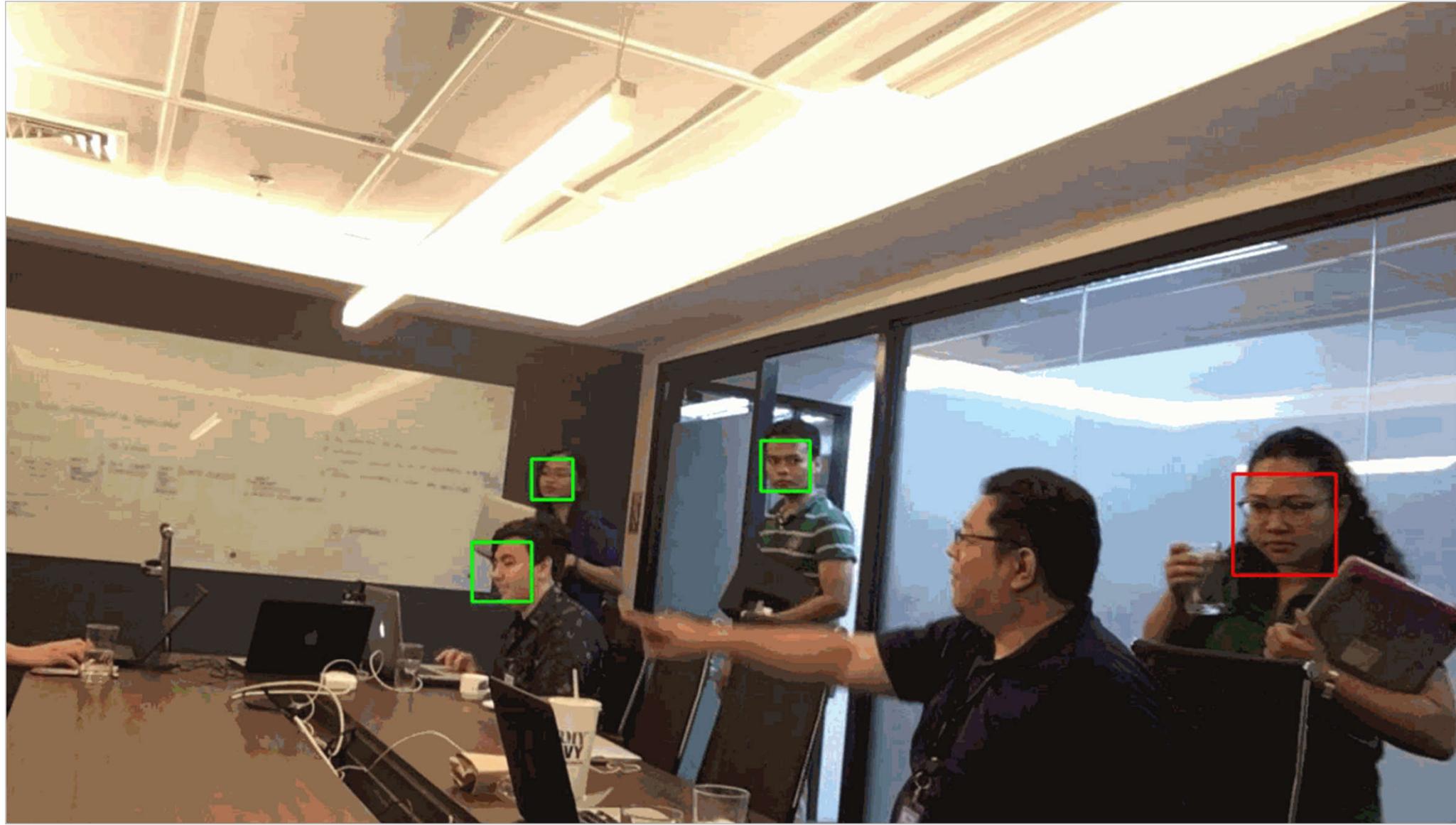
Medical Image Analysis



KYC: Face detection and Recognition

AGE, SEX, EMOTION, BLINK, POSE


Neural
MECHANICS



Intrusion Detection: face recognition



<https://www.analyticsindiamag.com/what-is-the-difference-between-computer-vision-and-image-processing/>

Autonomous Vehicles

Object detection and recognition, navigation and planning

Feb 16, 2012, 11:02am EST

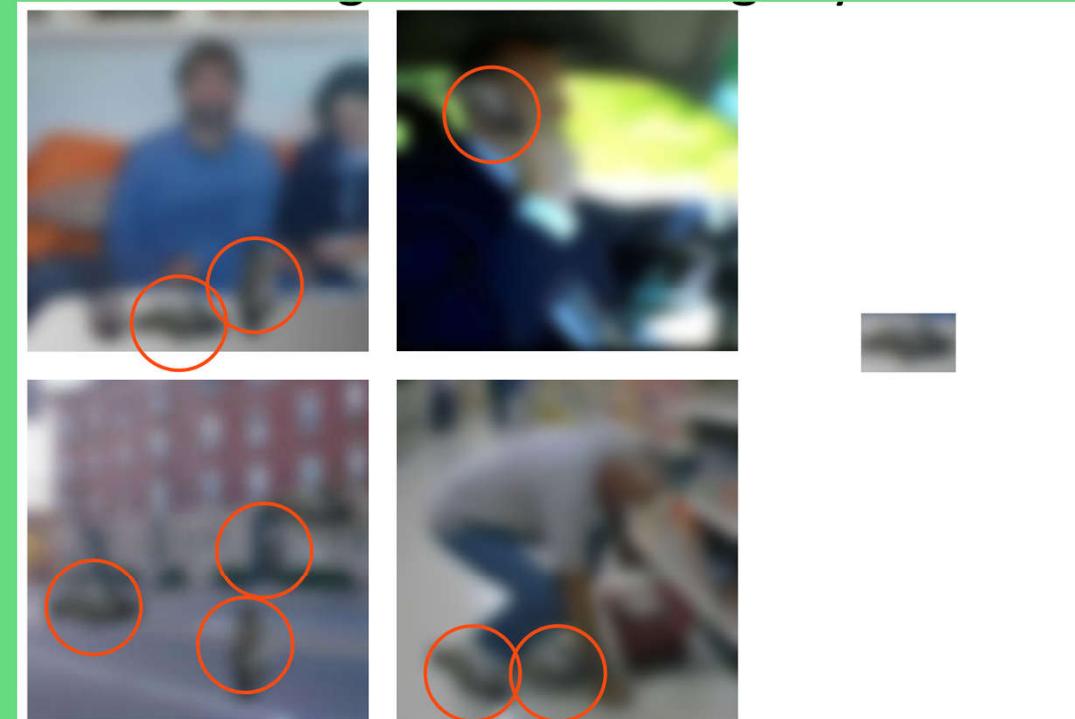
How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did



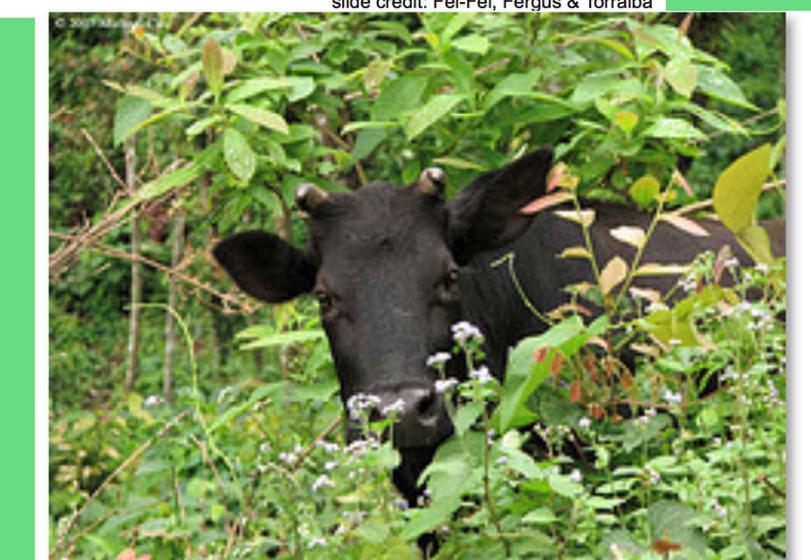
Retail and Media

Recommendation systems

Why is computer vision hard?



slide credit: Fei-Fei, Fergus & Torralba



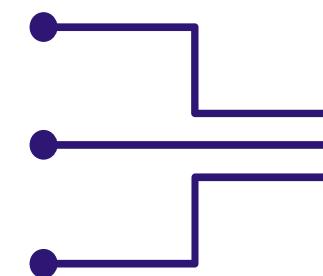
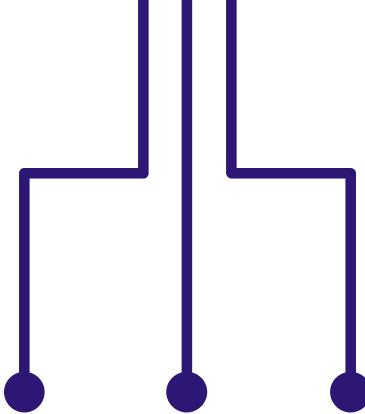
Why is computer vision hard?

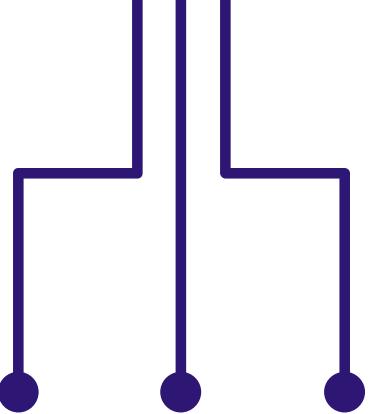
Perception -- inherently ambiguous problem

- Many different 3D scenes could have given rise to a particular 2D picture
- We often need to use prior knowledge about the structure of the world

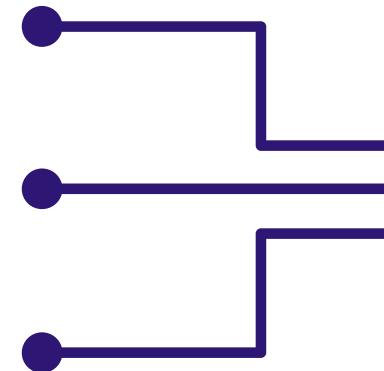


DESIGN SPRINT



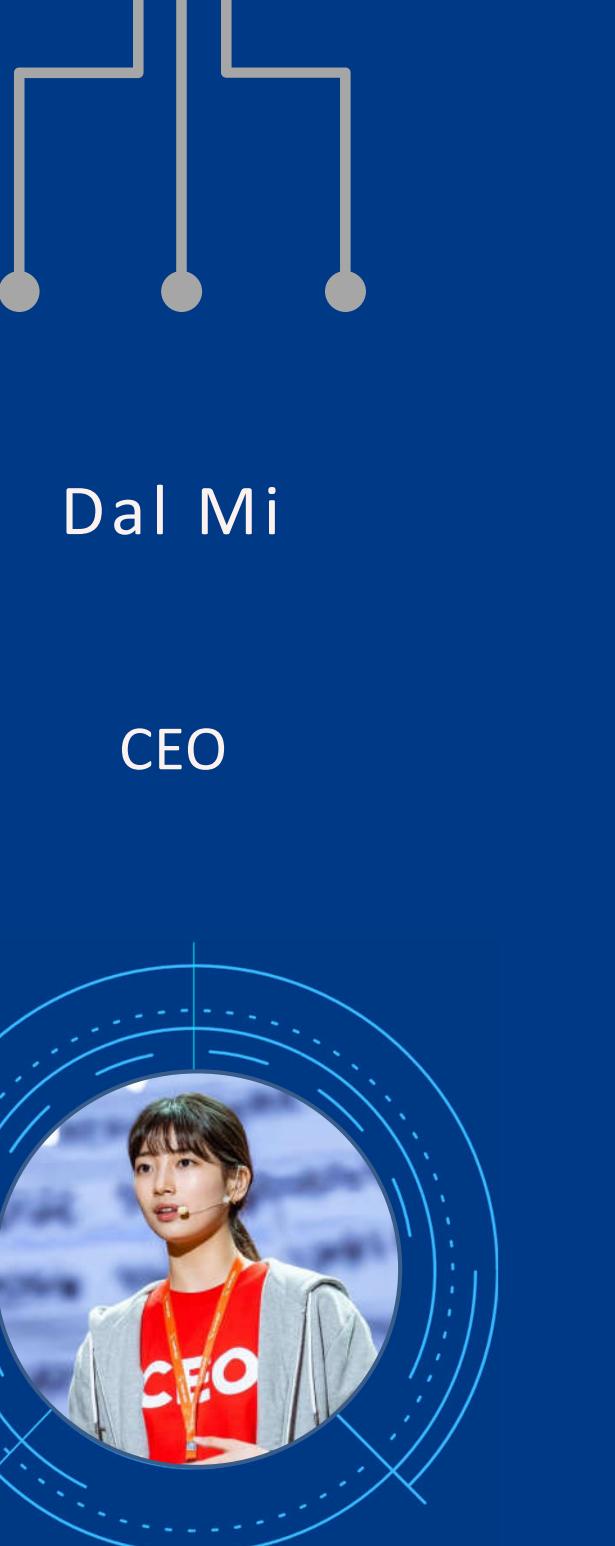
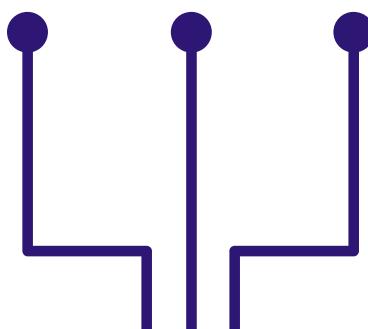


PROOF-OF-CONCEPT





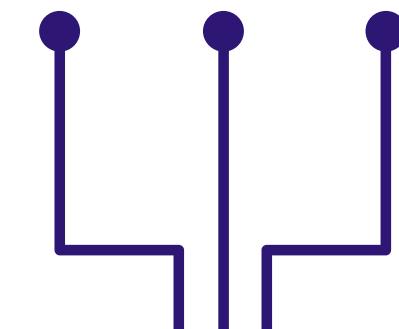
Do-San, Chul-San,
Yong-san
CTO and
Software Engineers

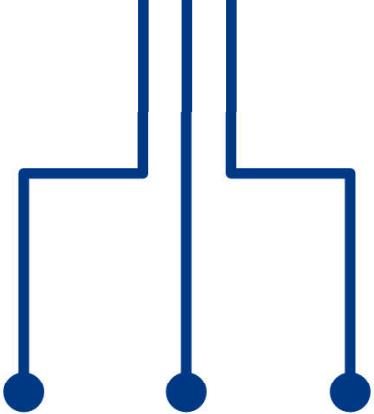


Dal Mi
CEO



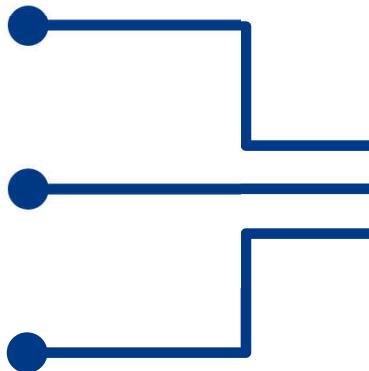
Jung Saha
Designer

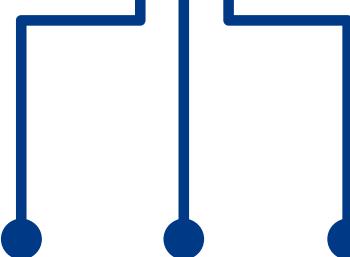




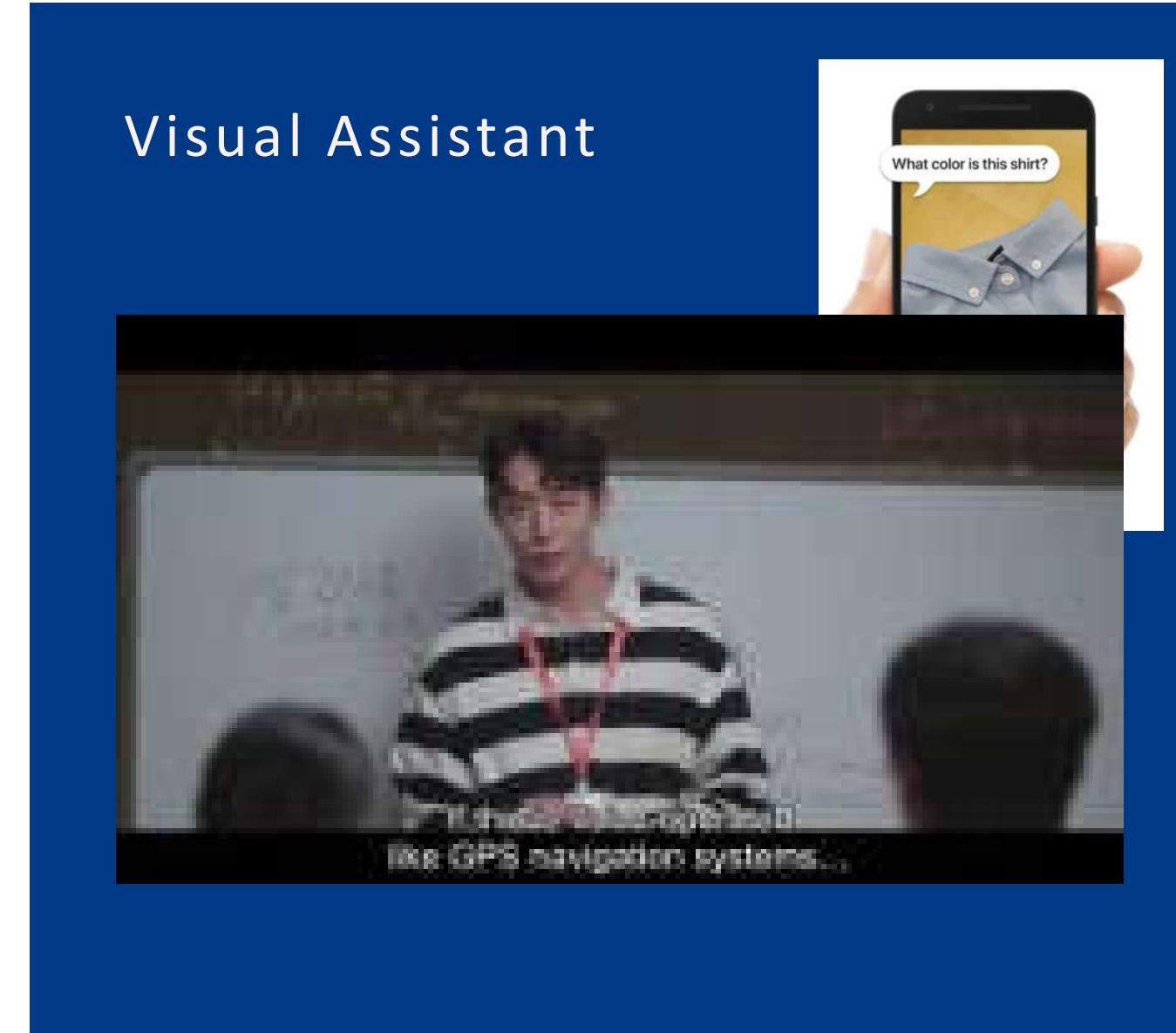
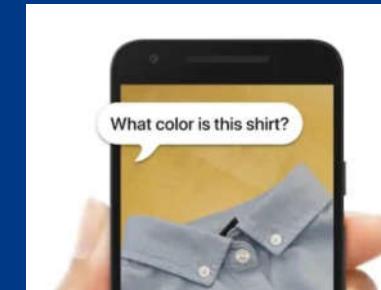
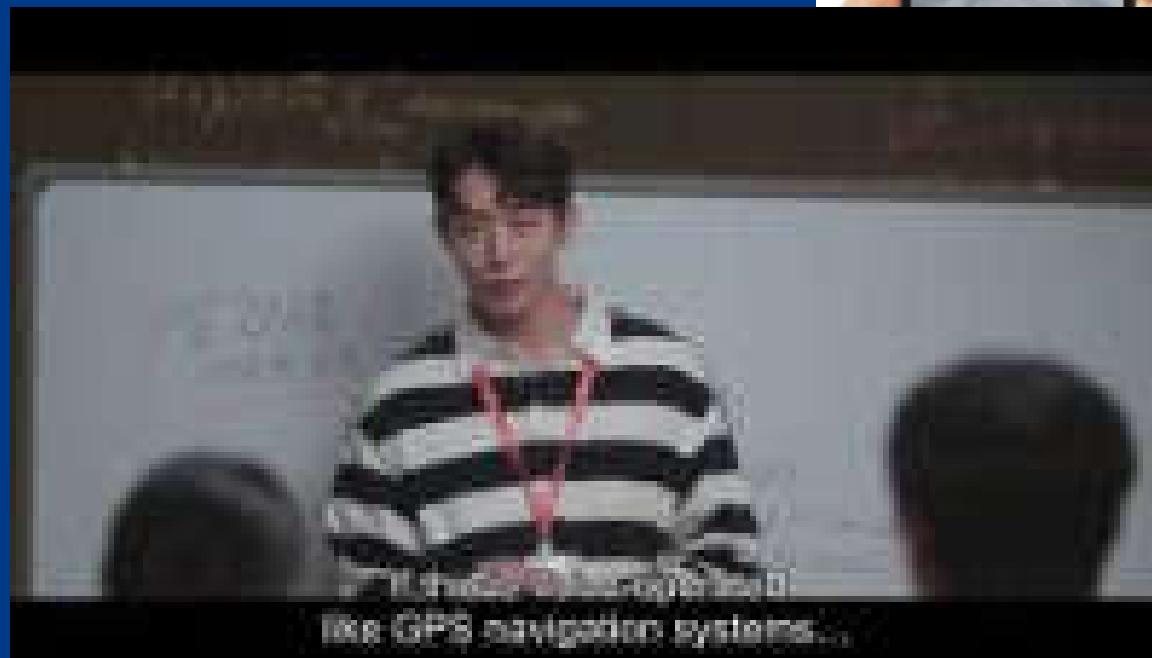
Object Detection: detect people (e.g.YOLO v3, Faster R-CNN)
Deployment: Raspberry Pi

DEMO





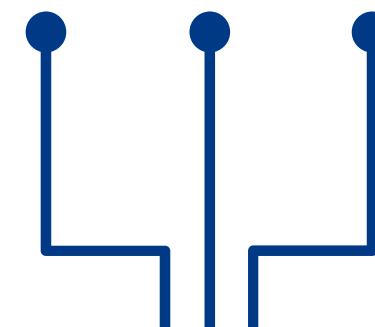
Visual Assistant



Autonomous Vehicles



Oh, there's a cat.



Use case

Optical Character Recognition (OCR)

Framework, Prototype, Deployment

