SAIR

Spatial AI & Robotics Lab

CSE 473/573

COMPUTER VISION & IMAGE PROCESSING

Chen Wang

Spatial AI & Robotics Lab

Department of Computer Science and Engineering

University at Buffalo The State University of New York

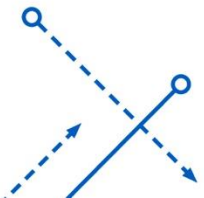
Introduction

University at Buffalo
Department of Computer Science and Engineering

CSE 473/573 - Computer Vision and Image Processing

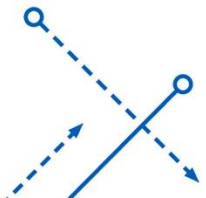
Spring 2026

TuTh 9:30AM - 10:50AM
Location: Knox 109
Jan 21, 2026 - May 5, 2026



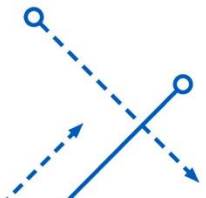
Instructor Information

Instructor:	**Chen Wang, PhD. (https://sairlab.org)**
Office:	304 Davis Hall
Email:	Contacted through Piazza



Special Considerations

- Prefers to be contacted through Piazza
  - Special Questions On demand
    - Contact: cse4573ta@sairlab.org
    - Office Hours: Davis 337, **Fri**. 2pm-3pm, Jan 30-May 8
- Piazza page
  - https://piazza.com/buffalo/spring2026/cse4573
  - Access code: sairlab
- Slides will be put up regularly.
- Instructor and TAs will monitor Piazza questions



Course Information

- Lectures, Quizzes, Projects, and Mid and Final Exam.
- This course is an introduction to those areas of Artificial Intelligence that deal with fundamental issues and techniques of **computer vision** and **image processing**.
- The emphasis is on **physical, mathematical, and information-processing aspects** of the vision.
- Topics to be covered include **image formation, edge detection and segmentation, convolution, image enhancement techniques, extraction of features such as color, texture, and shape, object detection, 3-D vision, and their applications.**
- The material is based on graduate-level texts augmented with research papers.

Overview

- The course will move faster than F25.
- **Self-discipline** is important.
- The emphasis of the course is to develop practical skills for solving **Computer Vision** and **Image Processing** problems
- **Fair evaluations: undergraduate and graduate students will be scored separately**
- Academic Integrity (AI) will be taken seriously and zero tolerance to any cheating
  - Please work on projects independently.
- ChatGPT and Copilot are not allowed in quizzes, projects, exams.

Prerequisites:

- CSE 203
  - Data Structure
- Strong Knowledge of Linear Algebra
- Strong Programming Experience in Python
- Git & GitHub:
  - https://www.coursera.org/learn/introduction-git-github

CSE4/573 F26 v.s. F25

- Students Feedback:
  - Too many assignments.
  - Expect Mid-exam.
  - Expect project solutions.
- We will make the following changes for SP26.
  - Replace 3 homework by
    - 1 midterm exam
    - Student presentation for sharing project solutions.
  - Remove less important contents for lectures.
    - For example, CVIP History.

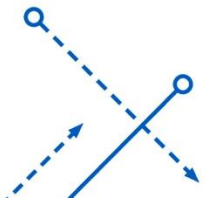
Grading

Weighting	Assessment	Number
10%	Midterm Exam	1
20%	Quizzes	5
30%	Projects	3
10%	Presentation (Project Solution)	1
30%	**Final Exam:** 5/11/2026, 8:00AM - 11:00AM, Cooke 121	1

Today: Quiz 0

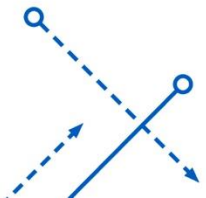
Other Information

- Class attendance and participation is expected.
  - Random Quizzes.
  - Quizzes will be given
    - **Class time only; online** through **UB Learns**.
  - **You'll need a laptop.**



Presentation

- 5-minute group presentation
- Teams of 2 students, self-organized
- One presenter per team
  - Choose best solution
- Both members receive the same score
- Present 1 of 3 projects
- Projects remain individual
- Teams may be formed after project submission



Presentation Scoring Criteria

- Clearly showing your unique contributions/solutions
- Logical flow with polished, visually consistent slides
- Strong motivation and problem framing
- Efficient and well justified solutions
- High quality Q&A responses
- **Remember to answer:**
  - What did I do well?
  - What can I improve?
- **Scoring Structure**
  - **Peer evaluation** and TA/Instructor evaluation

Grading Table

1. Undergraduate Version

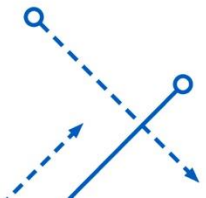
Score	Letter Grade	Score	Letter Grade
[85,100]	A	[52,58)	C+
[80,85)	A-	[46,52)	C
[73,80)	B+	[41,46)	C-
[65,73)	B	[40,41)	D
[58,65)	B-	[0,40)	F

2. Graduate Version

Score	Letter Grade	Score	Letter Grade
[87,100]	A	[55,60)	C+
[82,87)	A-	[50,55)	C
[75,82)	B+	[46,50)	C-
[67,75)	B	[45,46)	D
[60,67)	B-	[0,45)	F

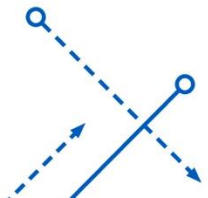
Grading and Submissions

- All assignments will be graded out of 100 points
- **Failure Tolerance**: We will drop **ONE Quiz** Grade, whichever results in a **higher** overall grade.
- All projects will be turned in via UB Learns.
- **GitHub Classroom** will be used to record coding history.
  - Used to detect AI issues.



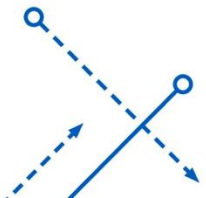
Option B

- Feedback about no compulsory attendance.
  - Release videos regularly
  - Turn lectures to office hours.
  - Only require class time on quiz days.
    - Teaching the most important topics in final exam.
  - The others are the same.



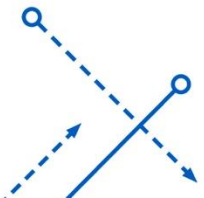
Exam Policy

- No makeup quizzes (no exceptional cases).
  - You may use the “failure tolerance” policy.
- No makeup exams will be given except in **extreme circumstances** and when consistent with University Policy.
- Notify your instructor & TA **1 month prior to the exam** via **piazza** if you are going to miss it. If it is medically impossible for you to give prior notice, please obtain **a note from a physician detailing the period** (with reasons) you were medically incapable of communicating with the instructor.
- You are responsible for knowing about the **exam date**. Please plan your travel and other activities accordingly.



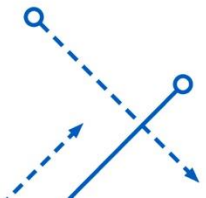
Late Submission Policy

- Completed project deliverables are to be submitted by their deadline (11:59pm).
- Grace days: You will be allowed a total of **3 grace days** throughout the semester
- Submissions beyond the 3 grace days will reduce your grade by 50%;
- No individual project will be accepted after 3 days late. No additional late days allowed.



Regrading for Errors

- Quizzes and exams may be submitted for regrading to correct grading errors.
- Regrade requests should be submitted to Piazza, and are due **no later than 2 days** after the scores are posted.
- When work is submitted for regrade, the entire work may be regraded, **which may result in a lower grade.**
- Work done in pencil may not be considered for regrading.



Disabilities

- If you have a diagnosed disability (physical, learning, or psychological) that will make it difficult for you to carry out the course work as outlined, or that requires accommodations such as recruiting note-takers, readers, or extended time on exams or assignments, please **advise the instructor during the first two weeks of the course** so that we may review possible arrangements for reasonable accommodations.
- In addition, if you have not yet done so, contact the **Office of Disability Services**.

Academic Integrity

- All work must be your own
  - Do not take the answers, words, ideas or research findings of other people as yours; cite and acknowledge properly, and develop your own ideas.
  - No cheating
  - According to departmental policy, any violation of academic integrity will result in an “F” for the course, and termination of departmental financial scholarship.
  - Tools will be used to check similarity. **Similar submissions will result in “F” for all involved parties.**
- Use of code from online, e.g., GitHub, must include a proper and clearly visible attribution in your report.

How to Fail the Course

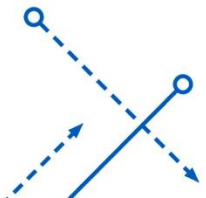
- Don't meet the prerequisites
- Don't pay attention to information on Piazza
- Start your project at the last minute
- Wait until the deadline to submit for the first time
- Don't read the syllabus carefully (Academic Integrity violation and late submission policy etc.)
- **Cheat (please check AI policy of the university and department if you are new)**
- https://www.buffalo.edu/academic-integrity/policies.html

Questions?

- **Instructors/TA/Graders will stay for a few minutes after each lecture. Simply ask!**
- Syllabus and Slides are on Piazza.

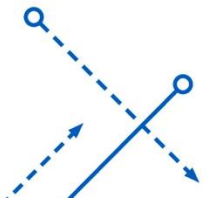
Content

- Definition
- Research Topics
- Applications
- Quiz 0
  - UB Learn, Academic Integrity Letter

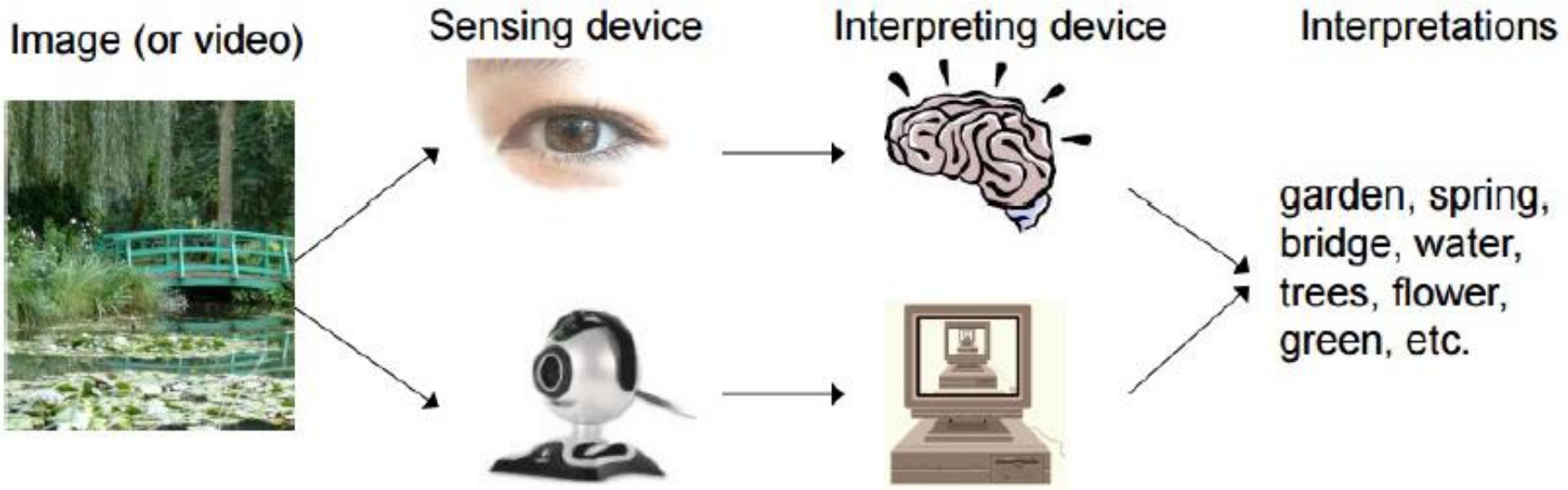


What is Computer Vision

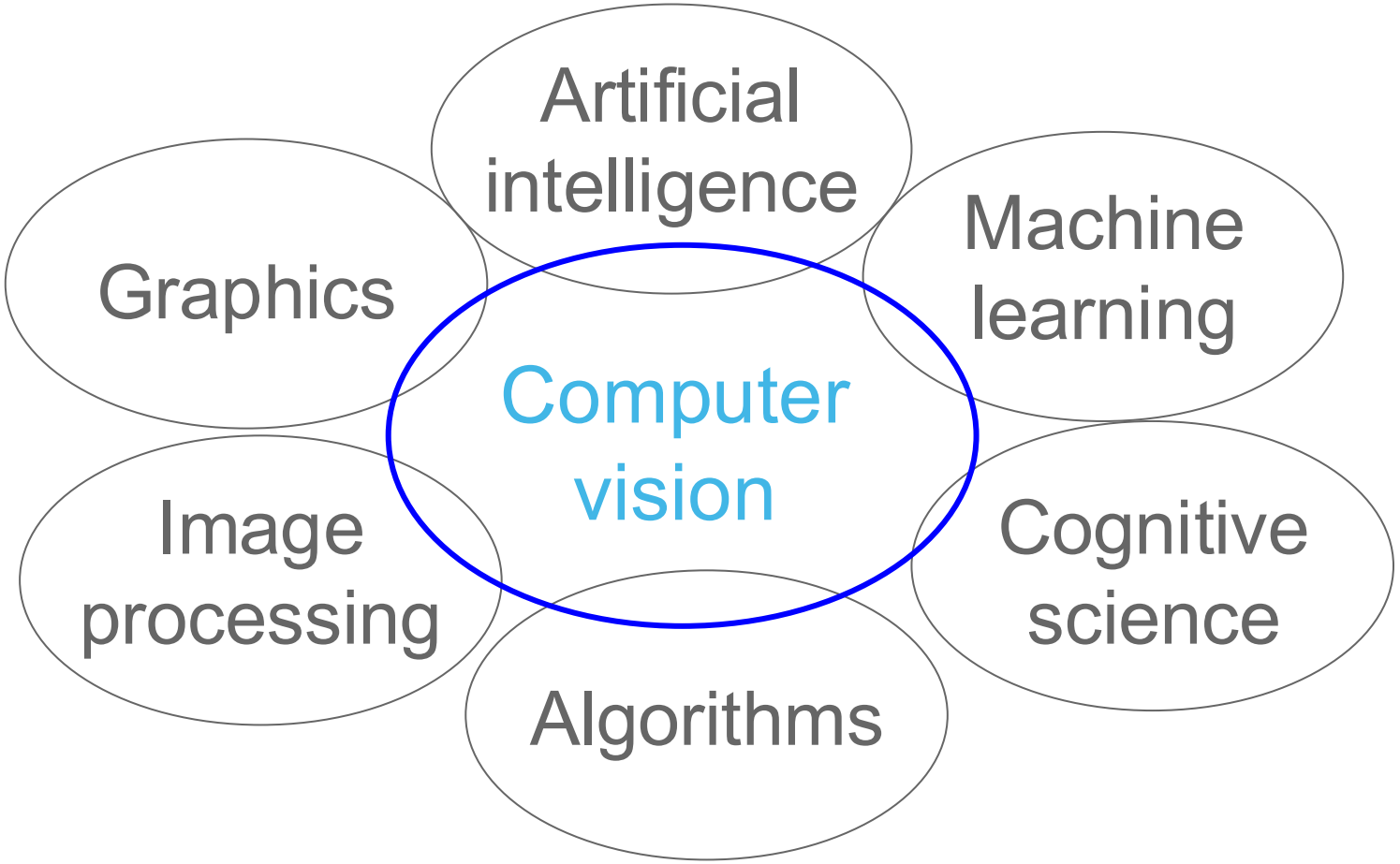
- **Computer vision** is a field of **computer science**
  - works on enabling **computers** to see,
  - identify and process images in the same way that human **vision** does, and
  - then provide proper output.
- It is like imparting human intelligence and instincts about vision to a **computer**.



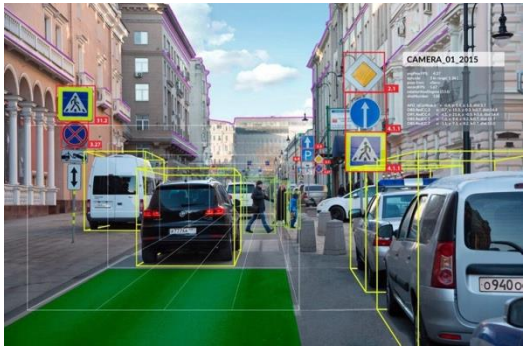
Computer Vision vs Human Vision



Related disciplines



Computer Vision vs Computer Graphics



Digital Image Processing
Computational Photography

Computer Vision

Images (2D)

Geometry (3D)
Shape

\+

Photometry
Appearance

Computer Graphics

SAIR
Spatial AI & Robotics Lab

Example: Visual Effects need CV + CG

Image Processing vs. Computer Vision

- Image Processing
  - Research area within electrical engineering/signal processing
  - Focus on syntax, low-level features

image → image (Denoising/inpainting)

- Computer Vision
  - Research area within computer science/artificial intelligence
  - Focus on semantics, symbolic or geometric descriptions

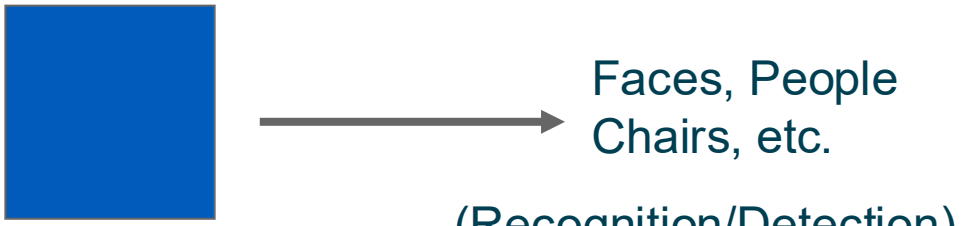


image → Faces, People, Chairs, etc. (Recognition/Detection)

Why learn Computer Vision?

- One third of human brain devoted to vision.
- 80% information is from vision.

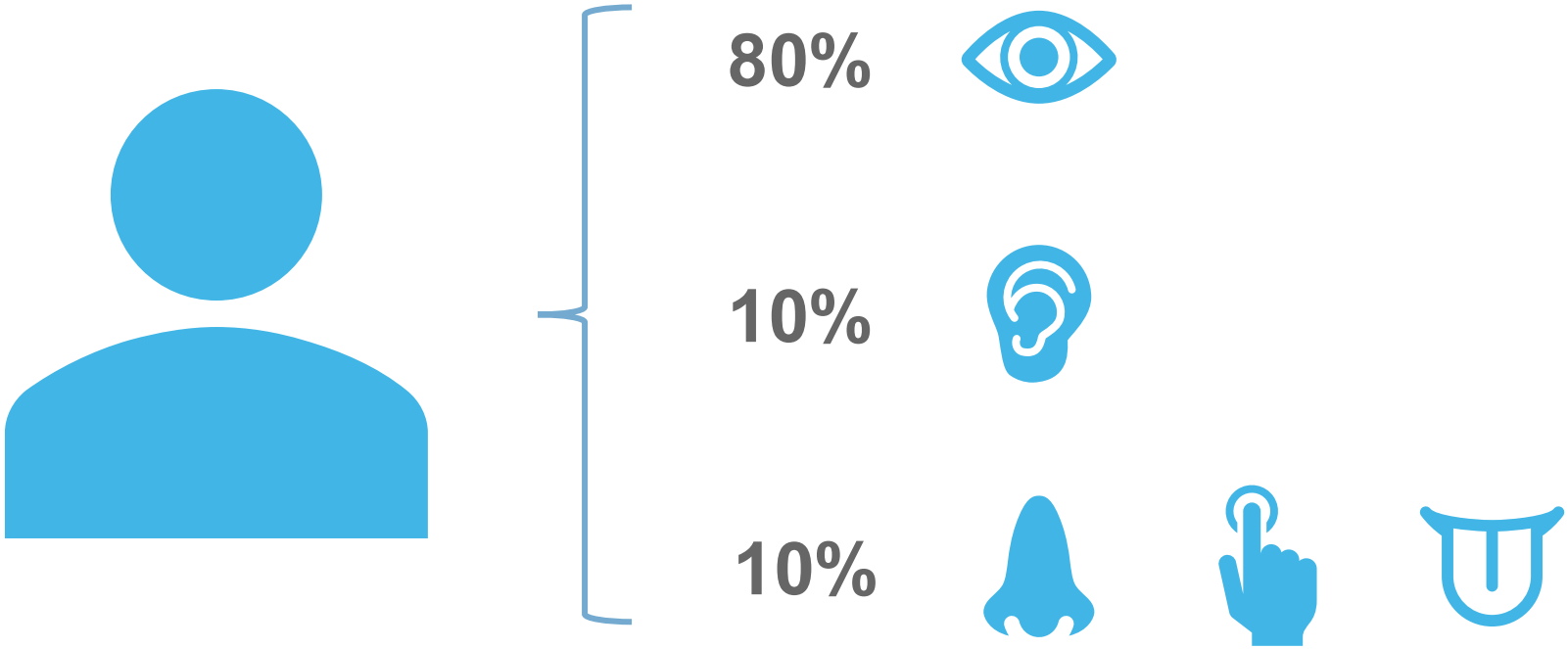


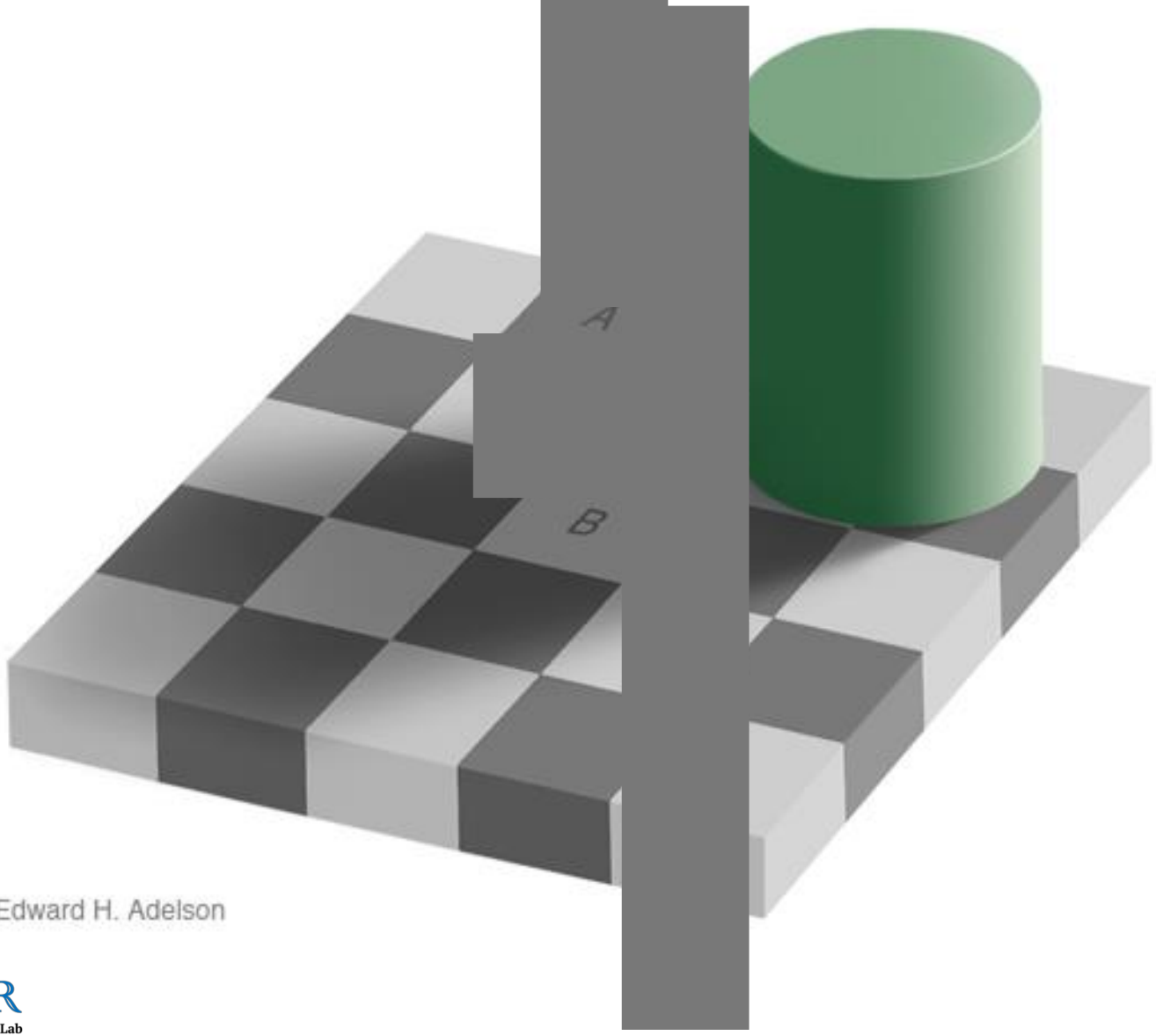
Image & Goal

An image is an array of numbers (pixels).

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

- To Bridge the Gap between Pixels and Meaning

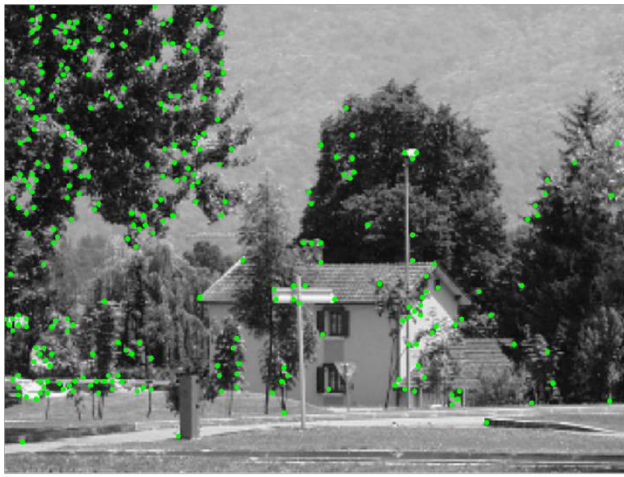
Which one is brig r? A or B?



Edward H. Adelson

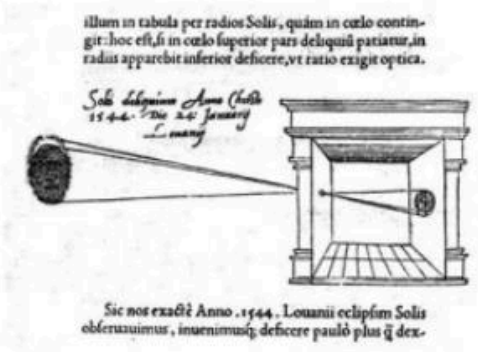
Feature Extraction

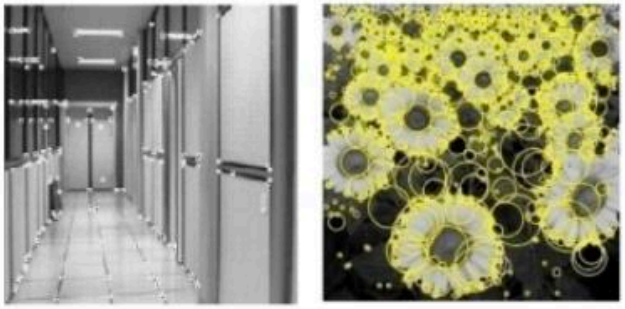




- Basic image processing and image formation

Filtering, edge detection





Feature extraction

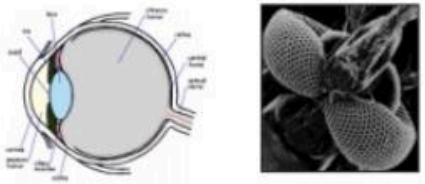
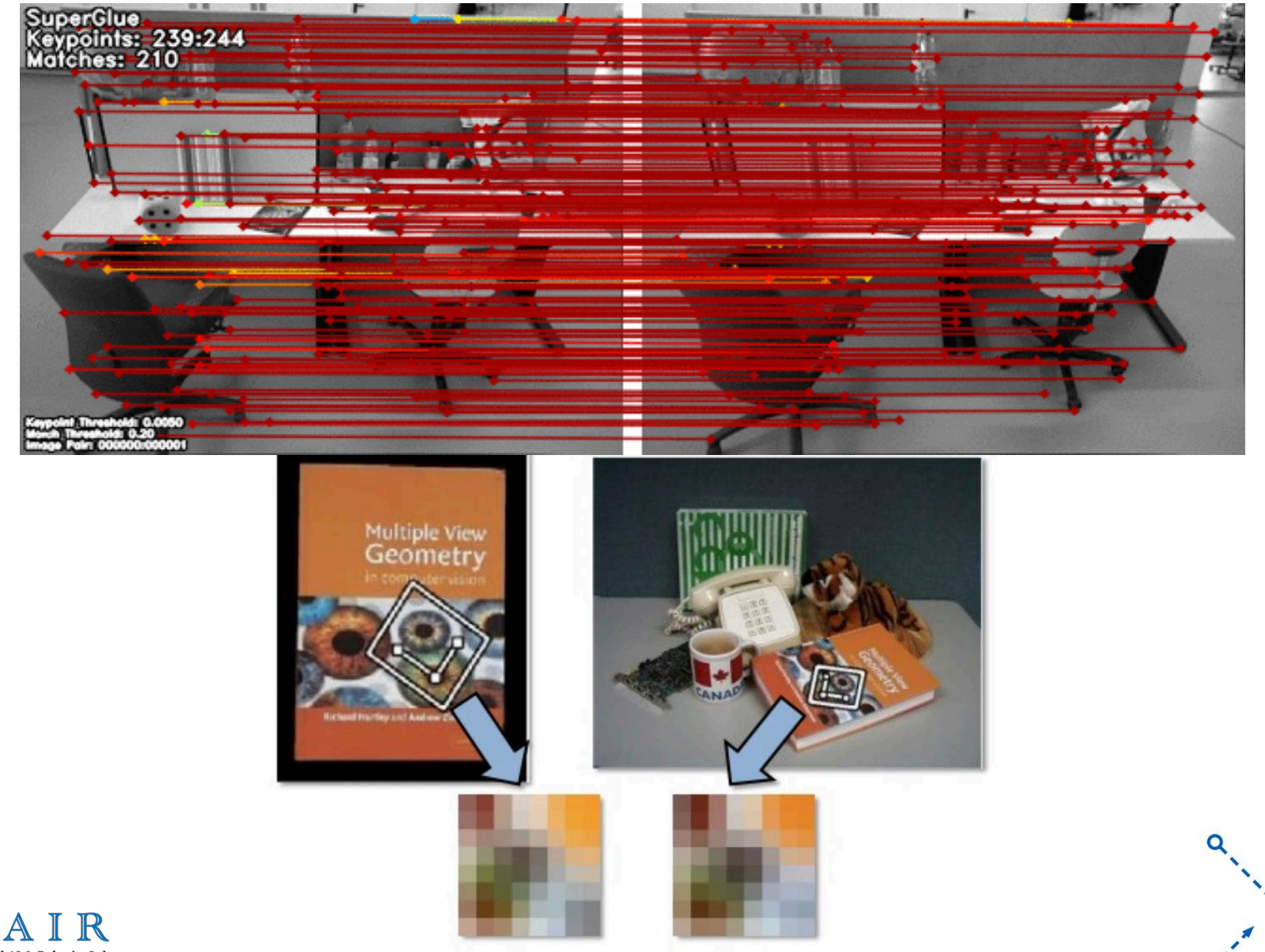
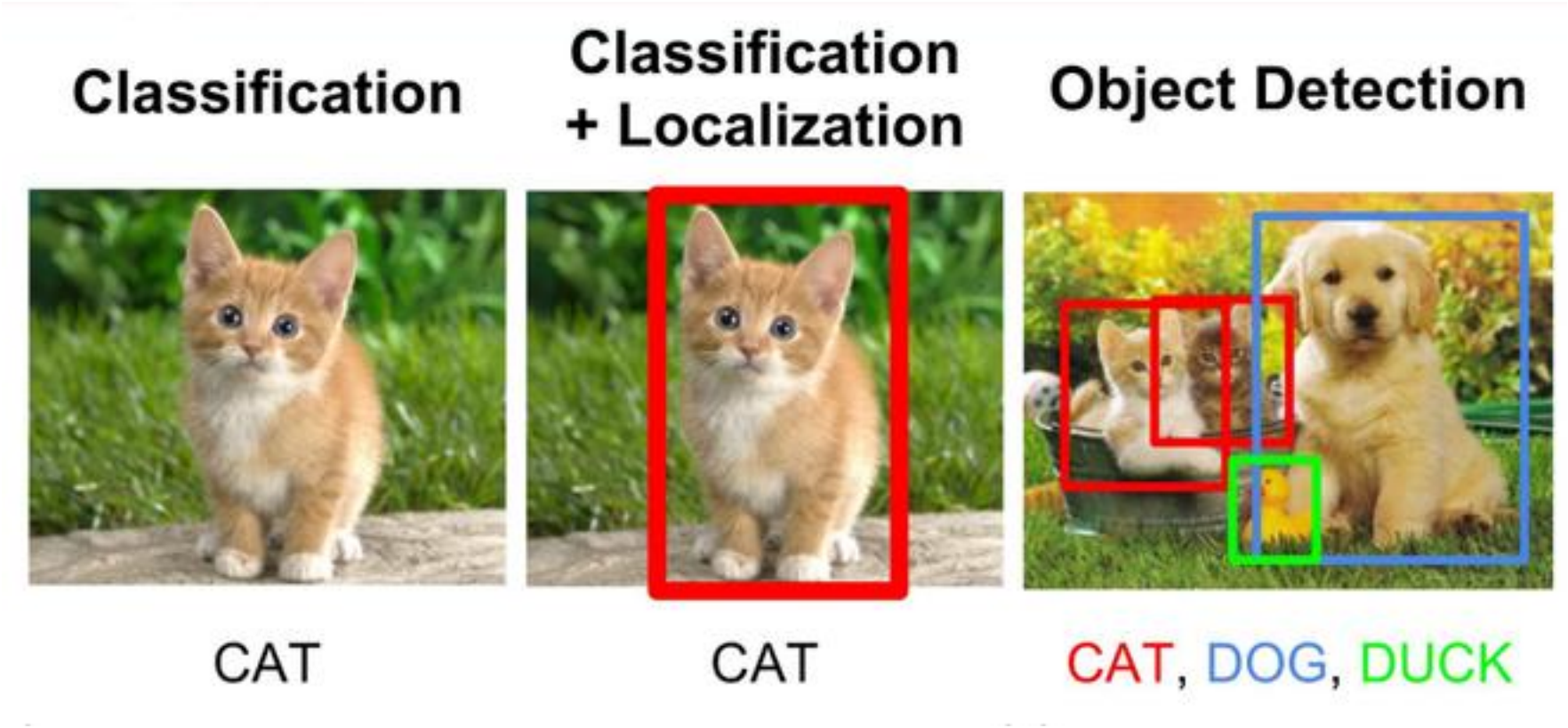


Image formation

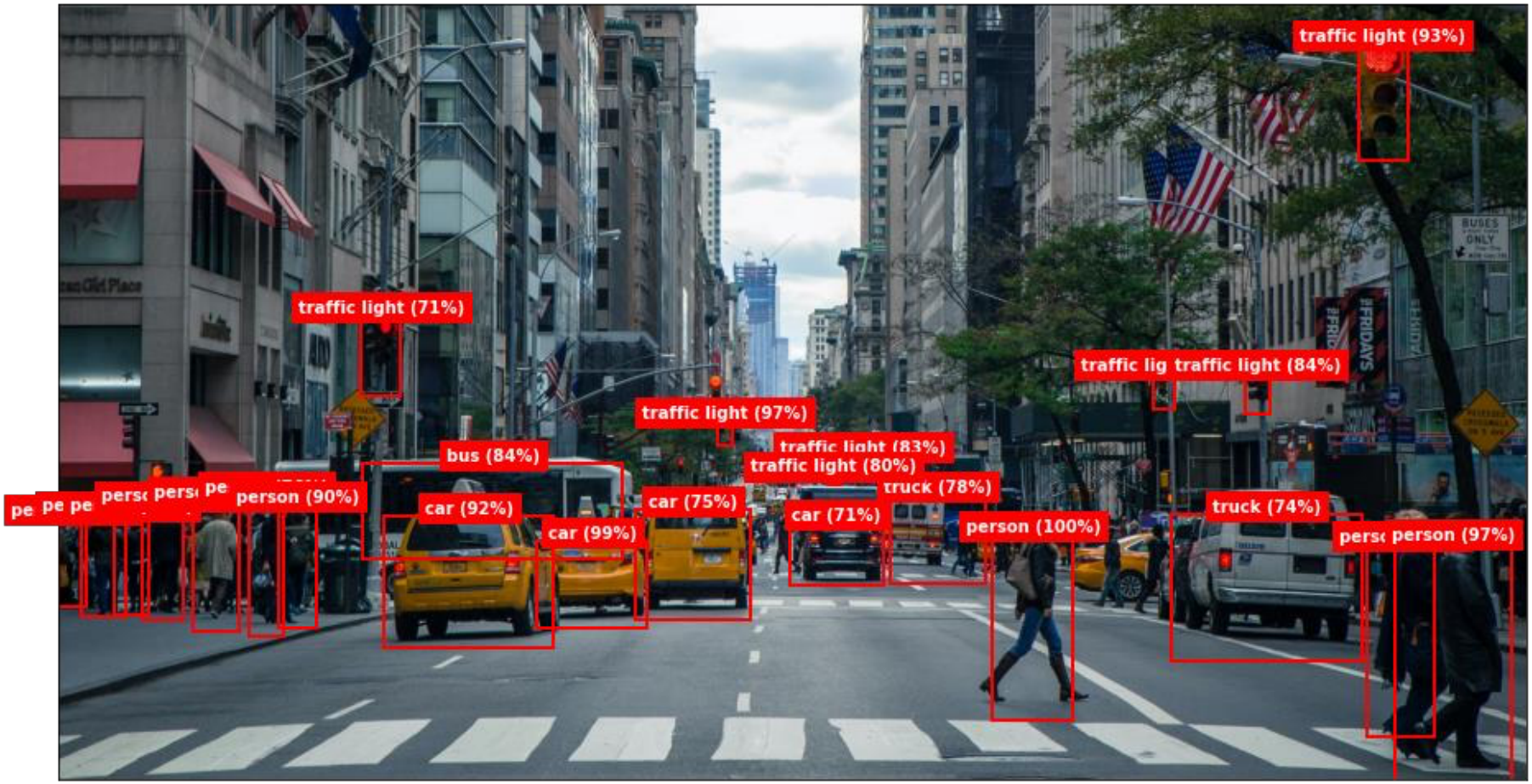
Feature Matching



Recognition and Detection

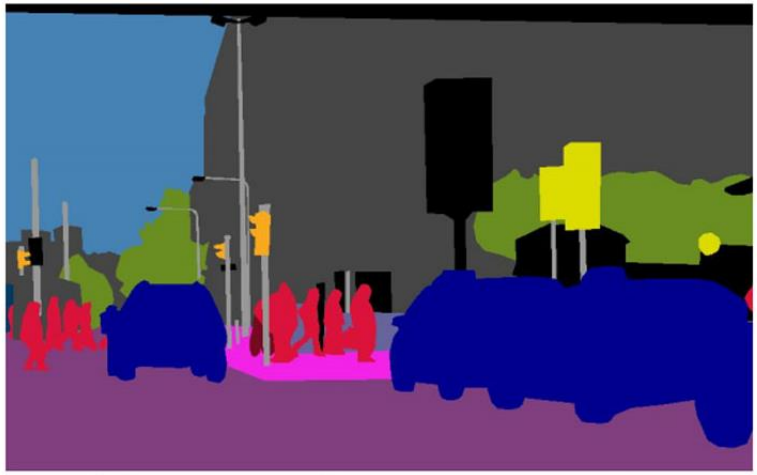


Detection

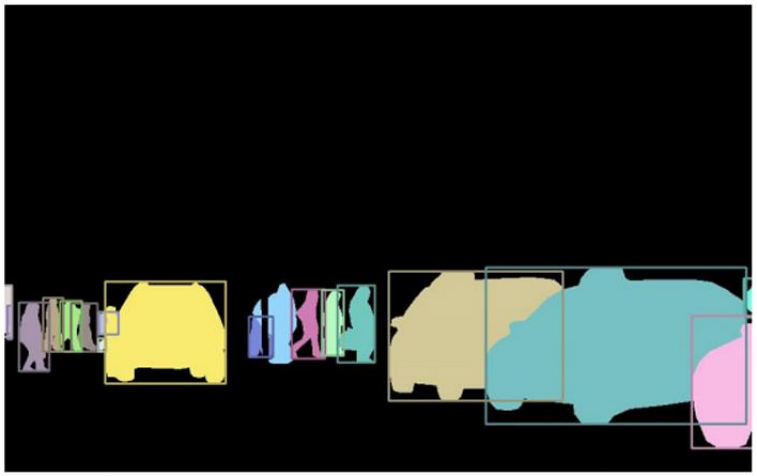

traffic light (93%)
traffic light (71%)
traffic lig traffic light (84%)
traffic light (97%)
traffic light (83%)
traffic light (80%)
bus (84%)
truck (78%)
car (92%)
car (75%)
car (71%)
car (99%)
person (100%)
truck (74%)
person (90%)
person (97%)

Segmentation

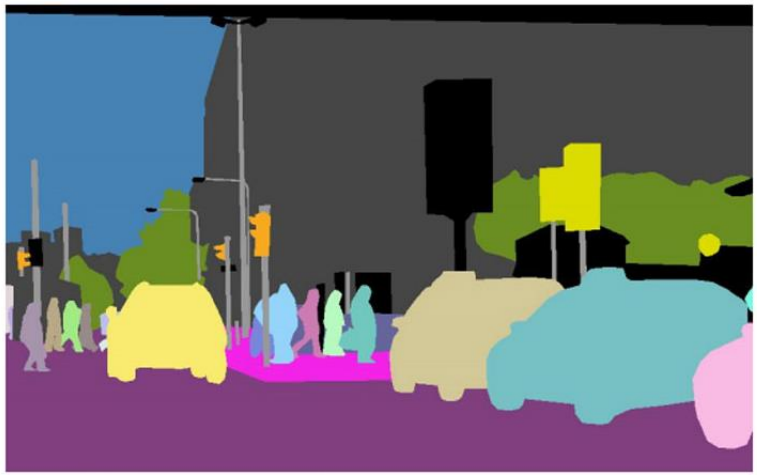
(a) image



(b) semantic segmentation

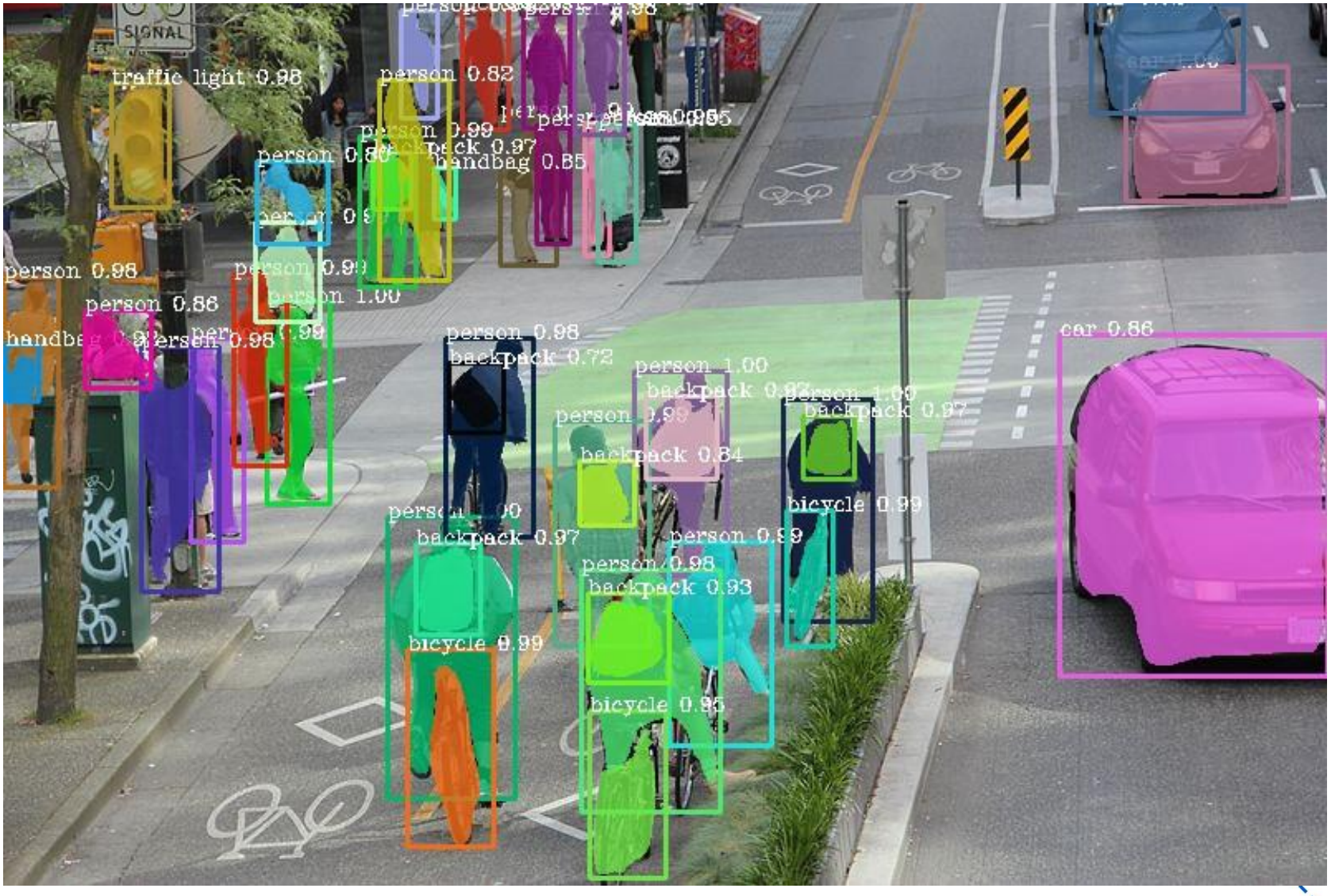


(c) instance segmentation



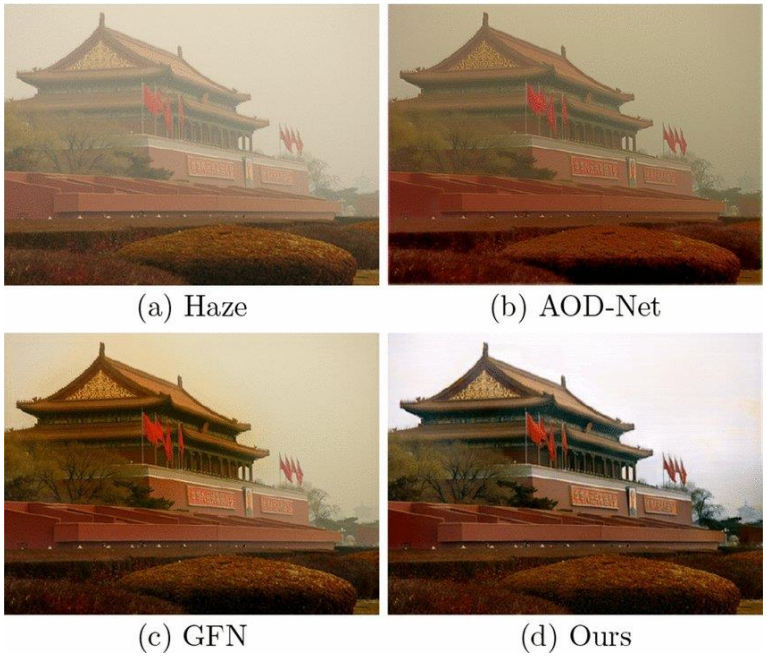
(d) panoptic segmentation

Detection + Segmentation



Inpainting

Denoising (Dehazing, Deblurring)


(a) Haze (b) AOD-Net

(c) GFN (d) Ours

Super Resolution

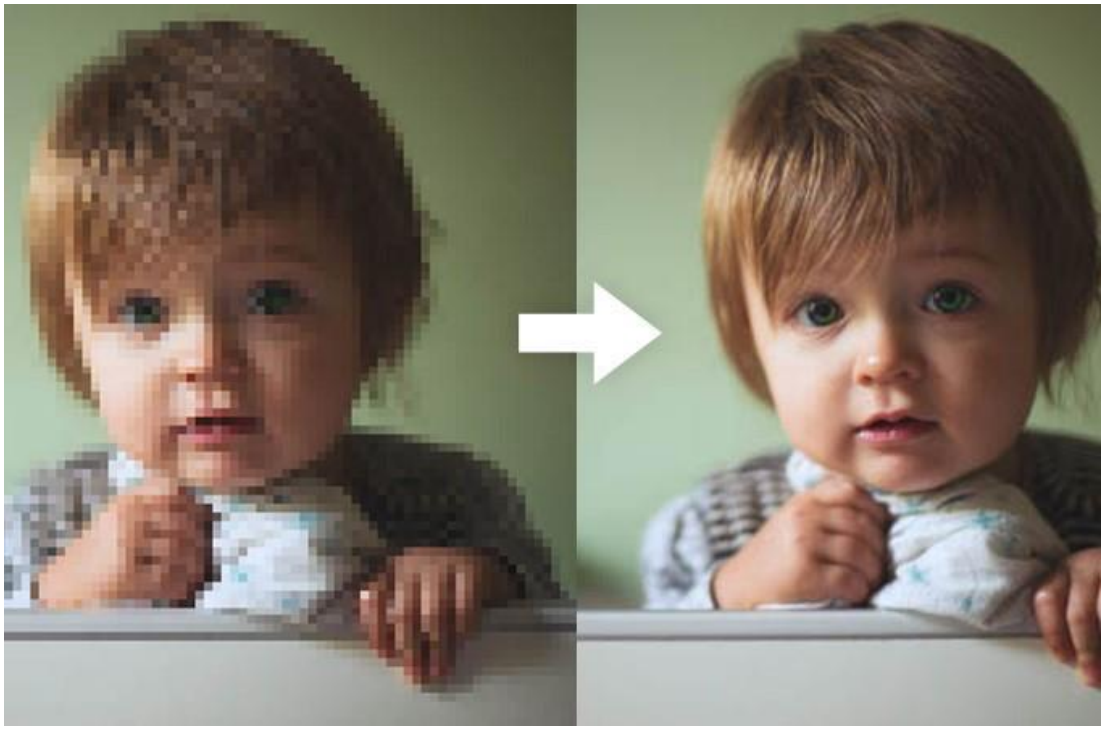
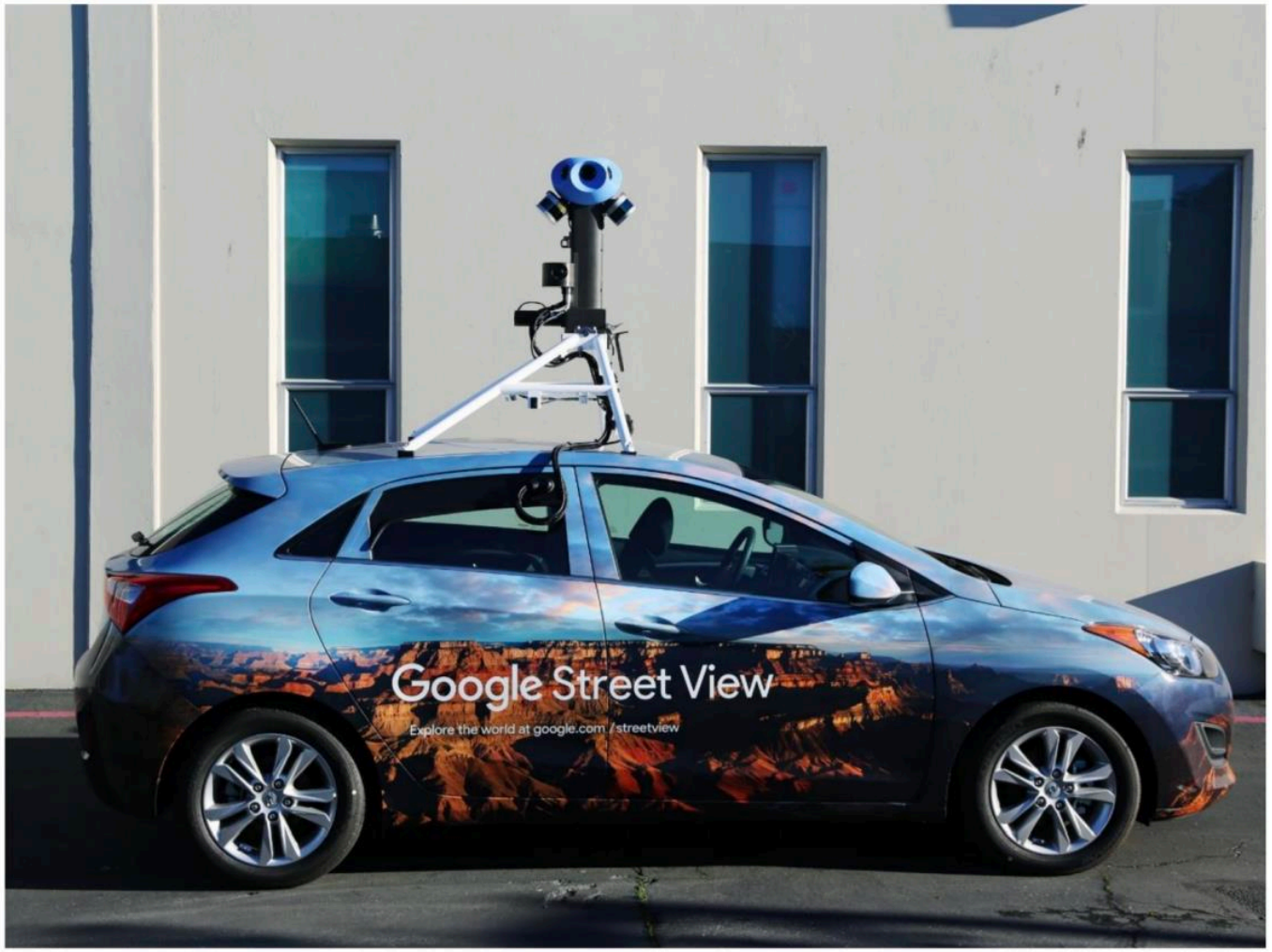


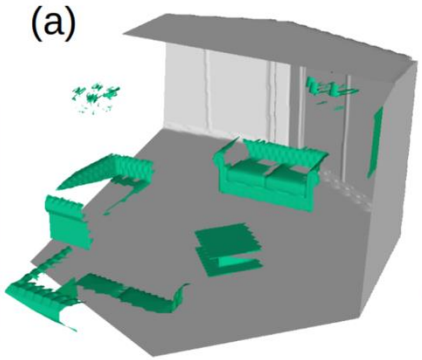
Image Stitching: Panorama

Example: Google Street View

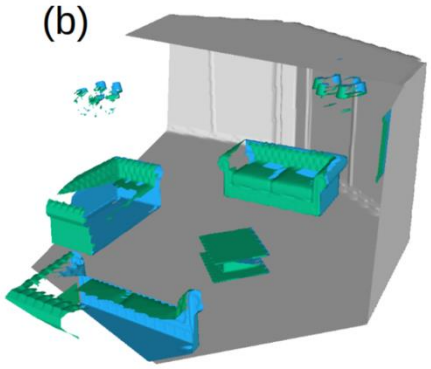


3D Reconstruction (RGB/RGBD)

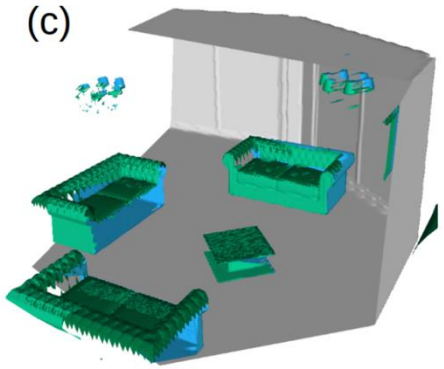
RGB Image



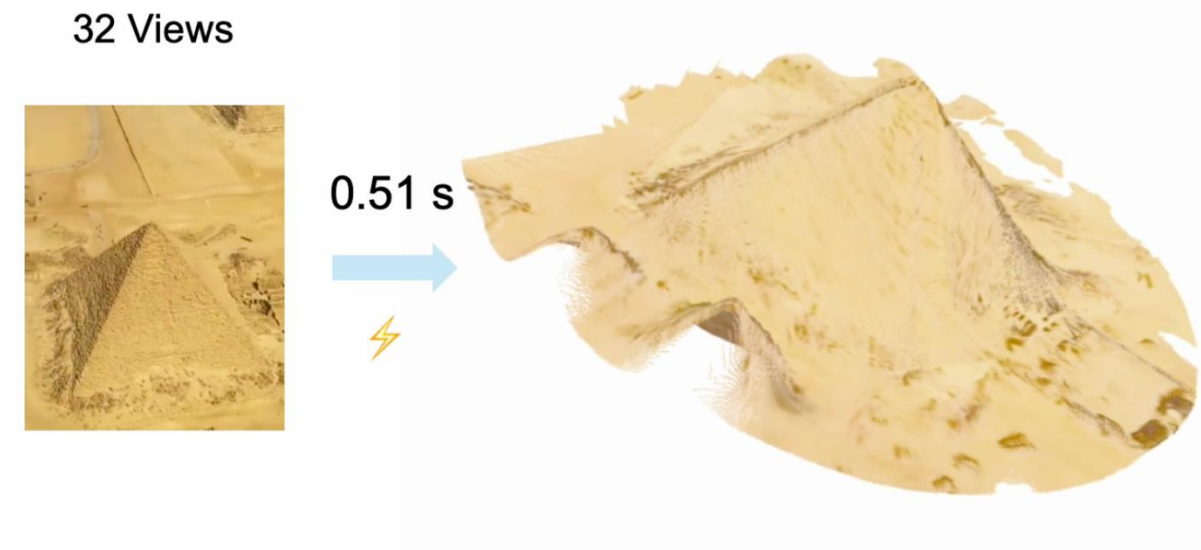
2.5D Object Surfaces



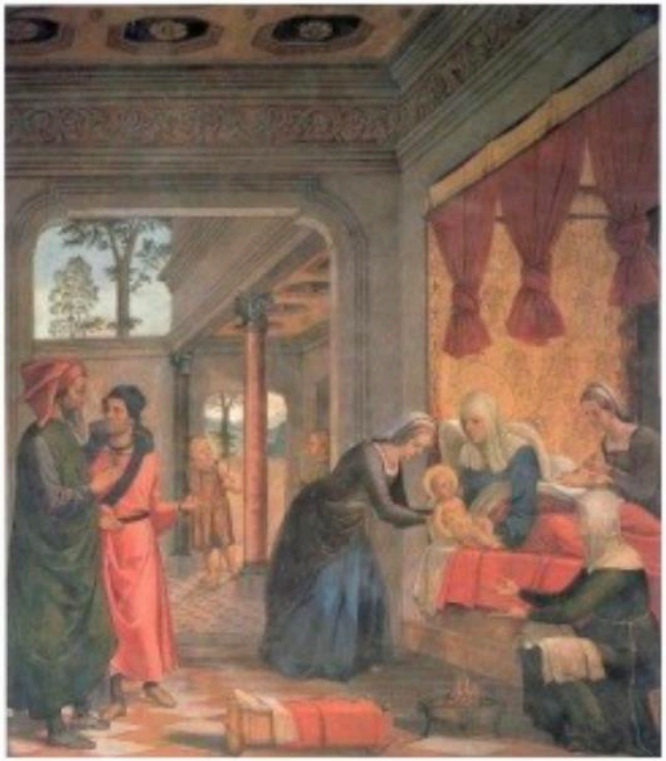
Multi-layer Surfaces



Multi-layer and
Virtual-view Surfaces



Single View Modeling

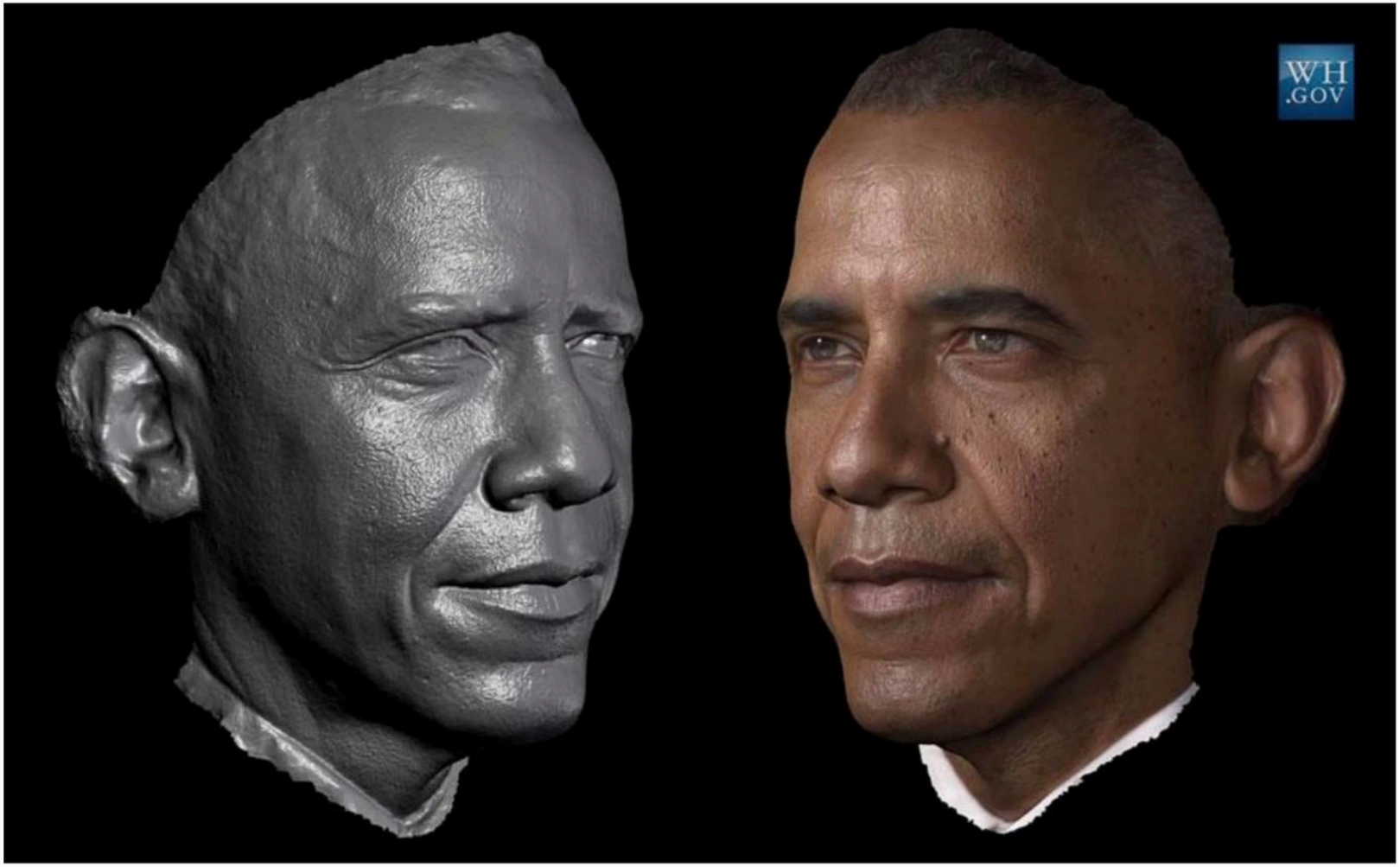


Can you do single view modeling?

Can you do single view modeling?

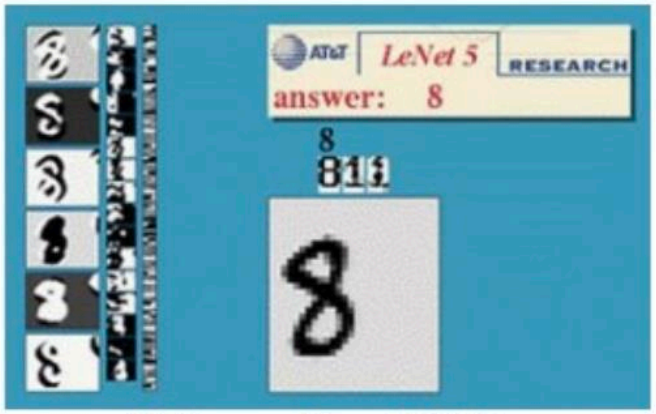
Multi-Camera

3D Scanning



Everyday Applications - OCR

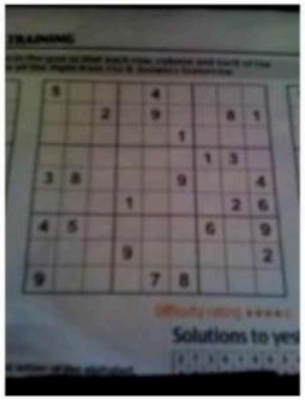
- If you have a scanner, it probably came with OCR software



Digit recognition, AT&T labs
http://www.research.att.com/~yann/

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Automatic check processing

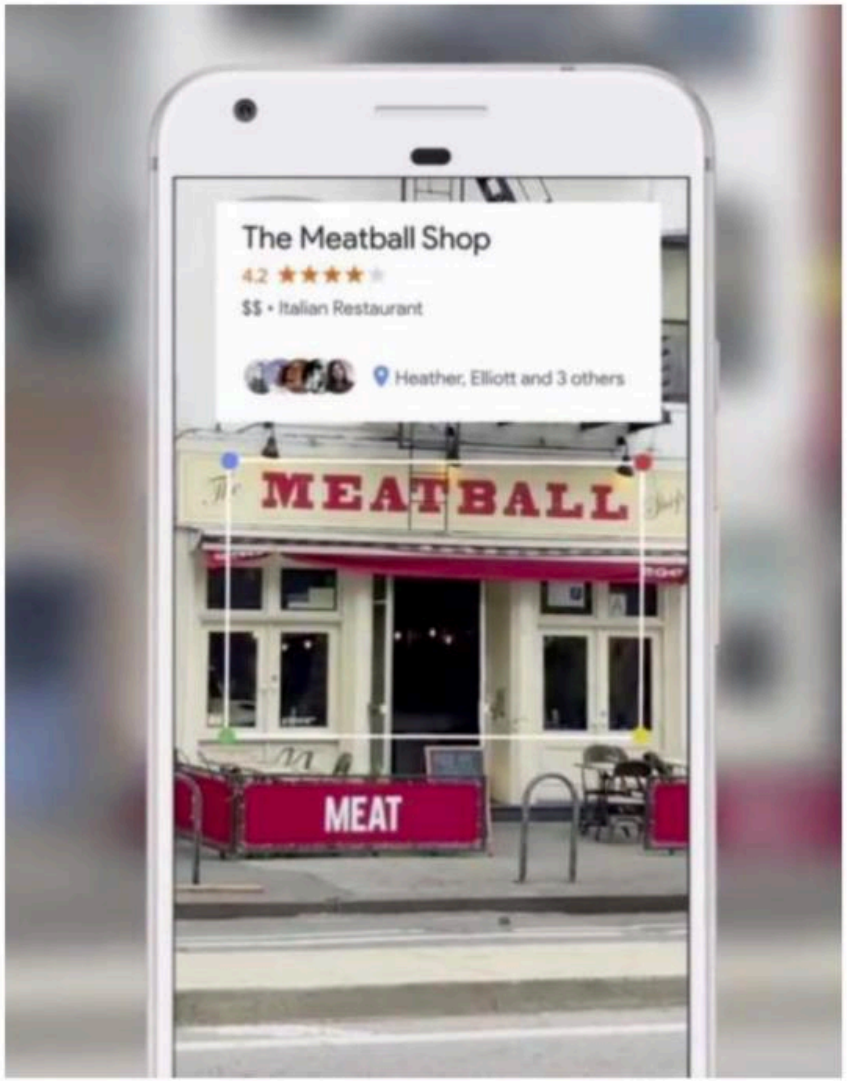


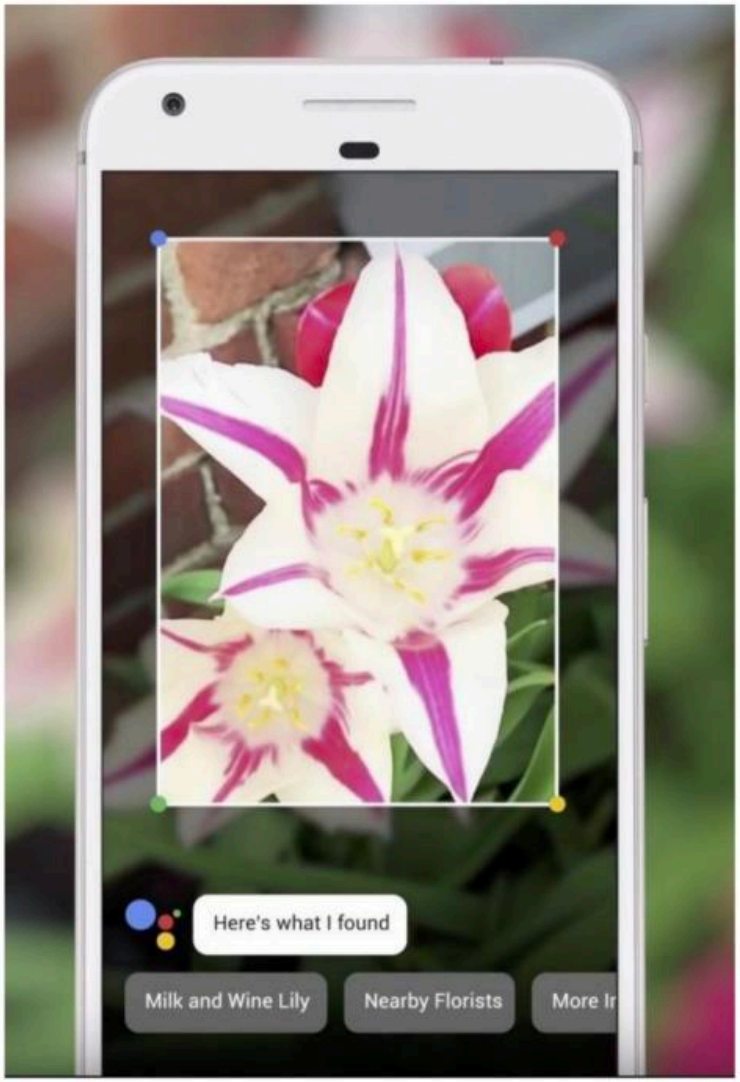
Sudoku grabber
http://sudokugrab.blogspot.com/

Source: S. Seitz

Object Recognition

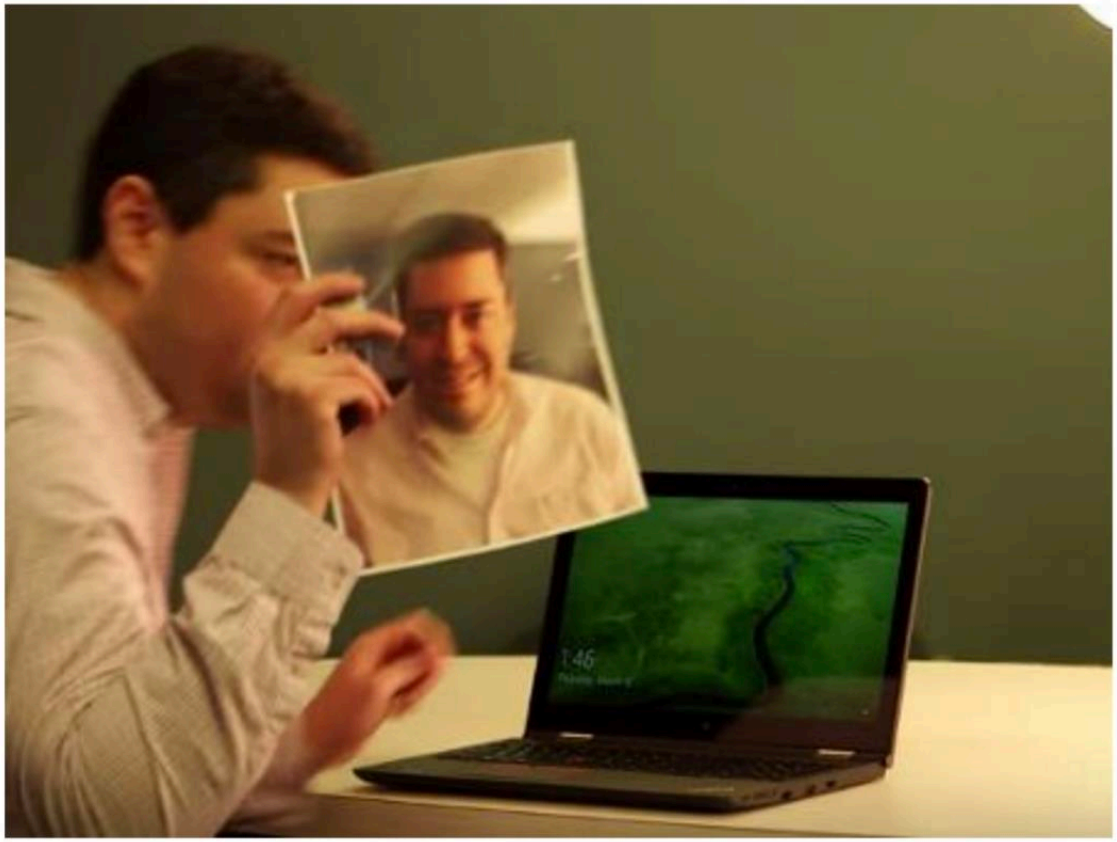
Visual Search: Google Lens





Face Detection

Face Recognition

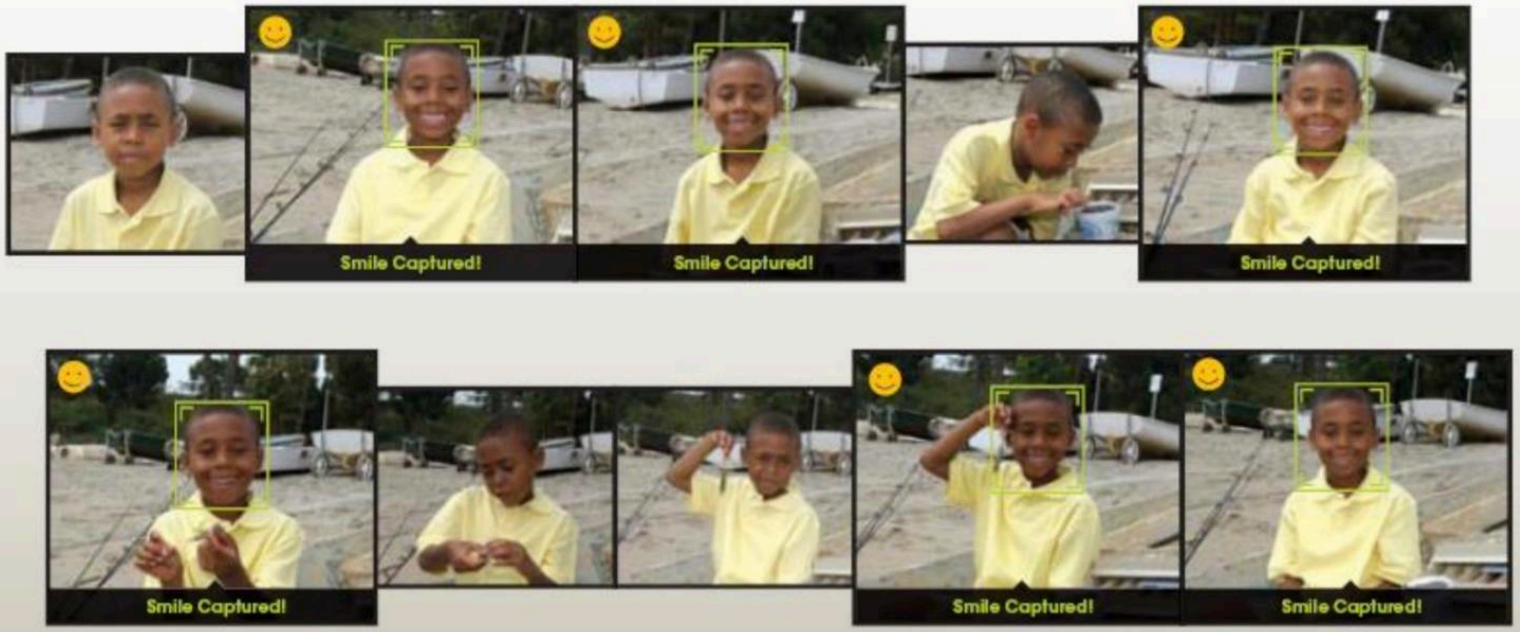


How to solve this problem?

Smile Detection

The Smile Shutter flow

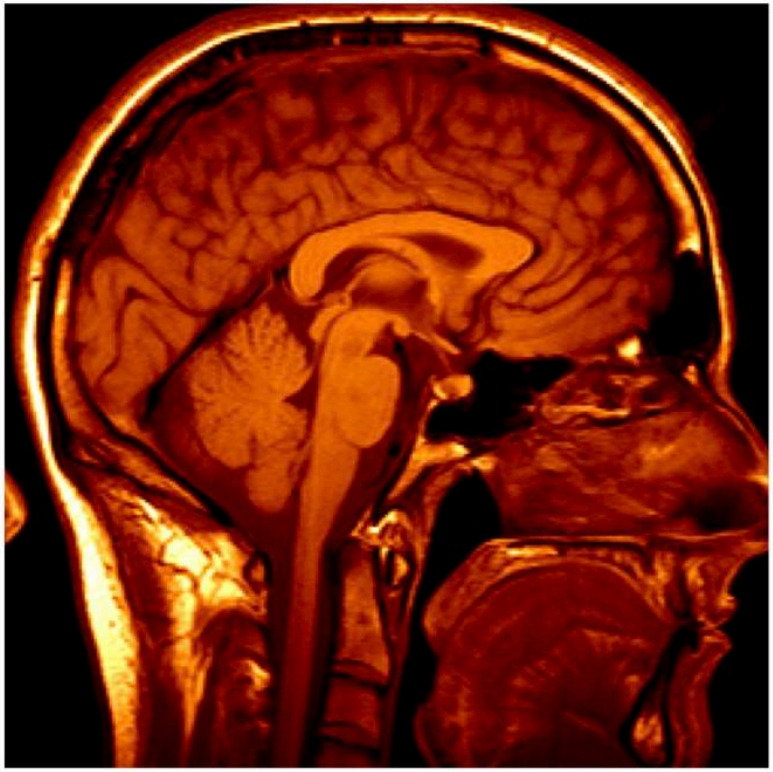
Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



Entertainment

Video Surveillance

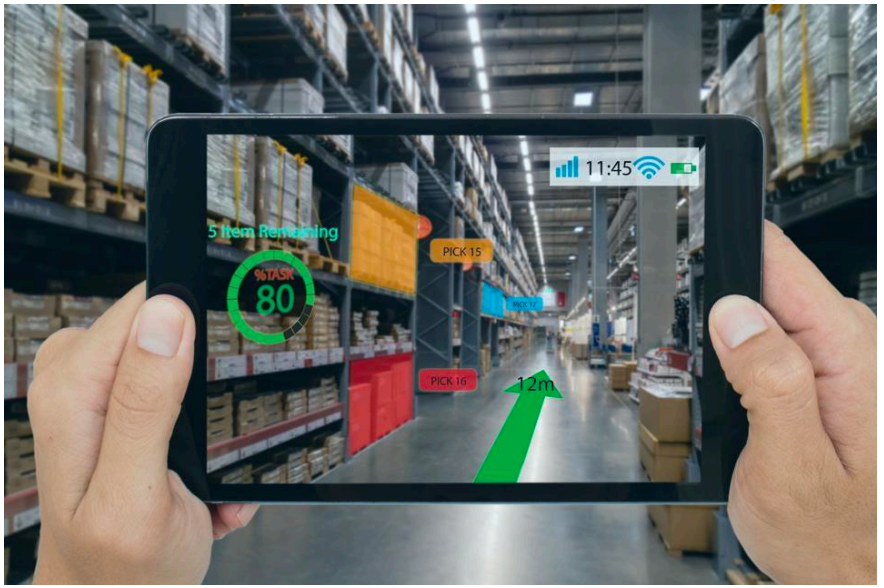
Medical Imaging



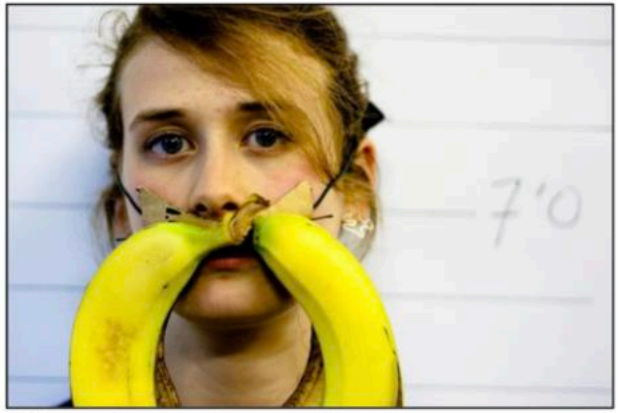
3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Virtual/Augmented/Mixed Reality

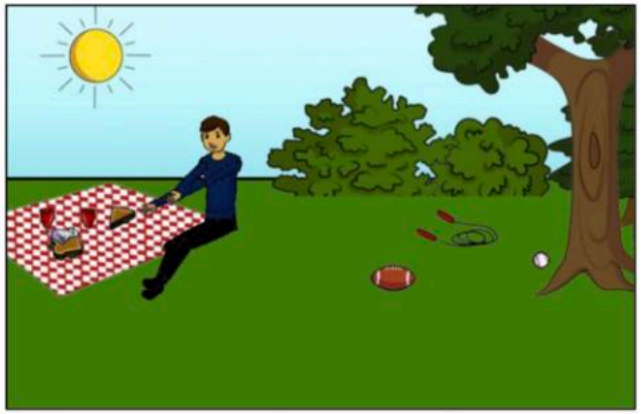


Visual Question and Answering (VQA)

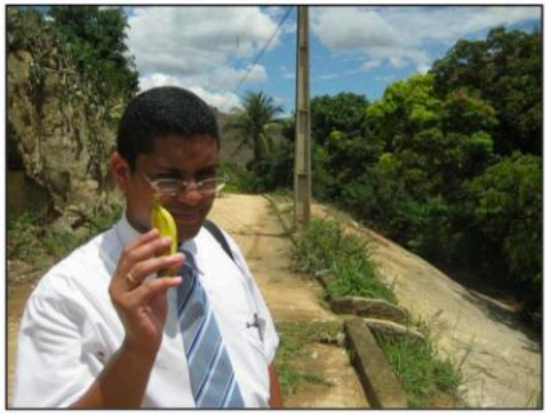


What color are her eyes?
What is the mustache made of?

How many slices of pizza are there?
Is this a vegetarian pizza?



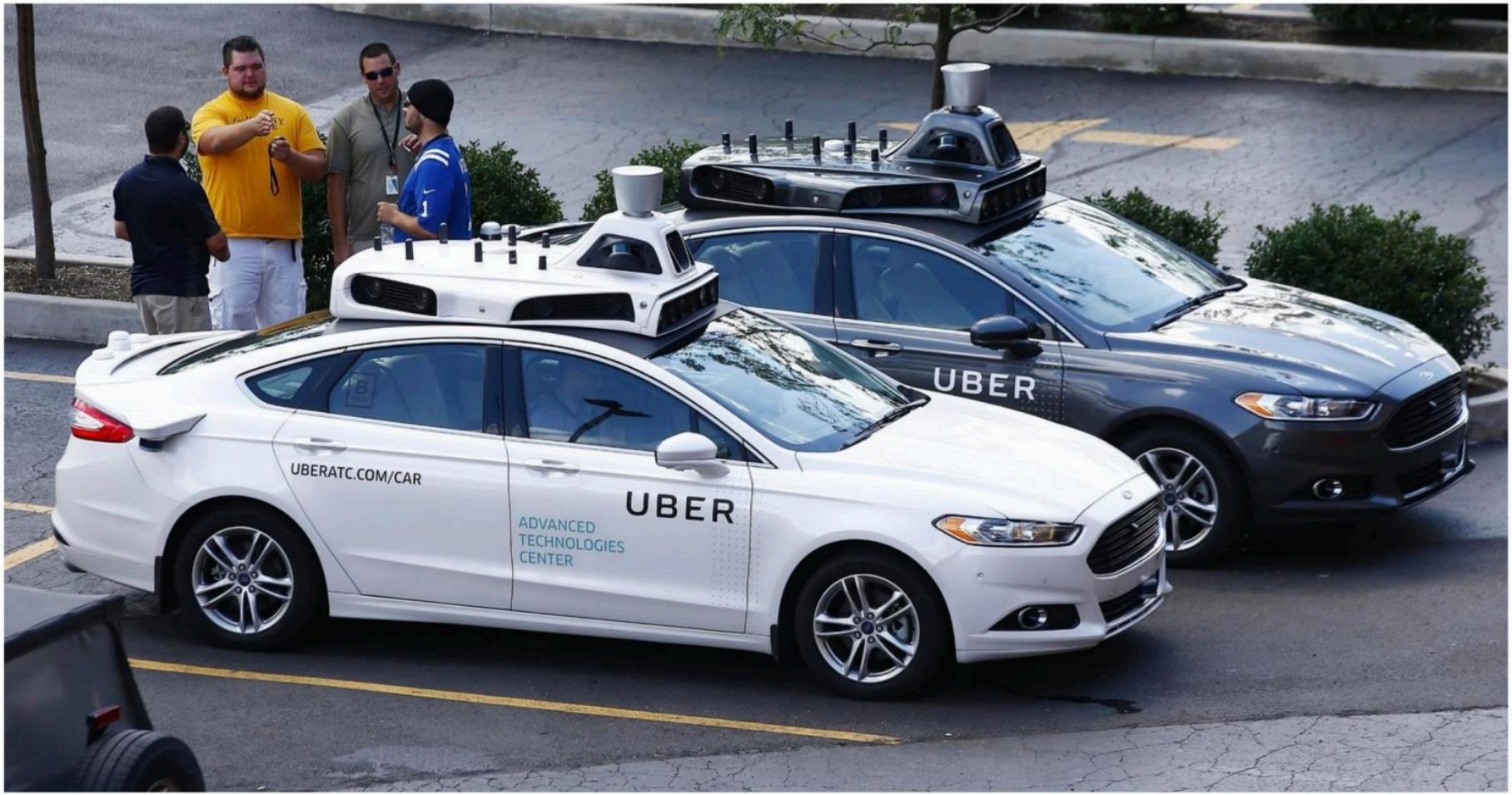
Is this person expecting company?
What is just under the tree?



Does it appear to be rainy?
Does this person have 20/20 vision?

https://dac.cs.vt.edu/research-project/visual-question-answering-vqa/

Autonomous Vehicle Navigation



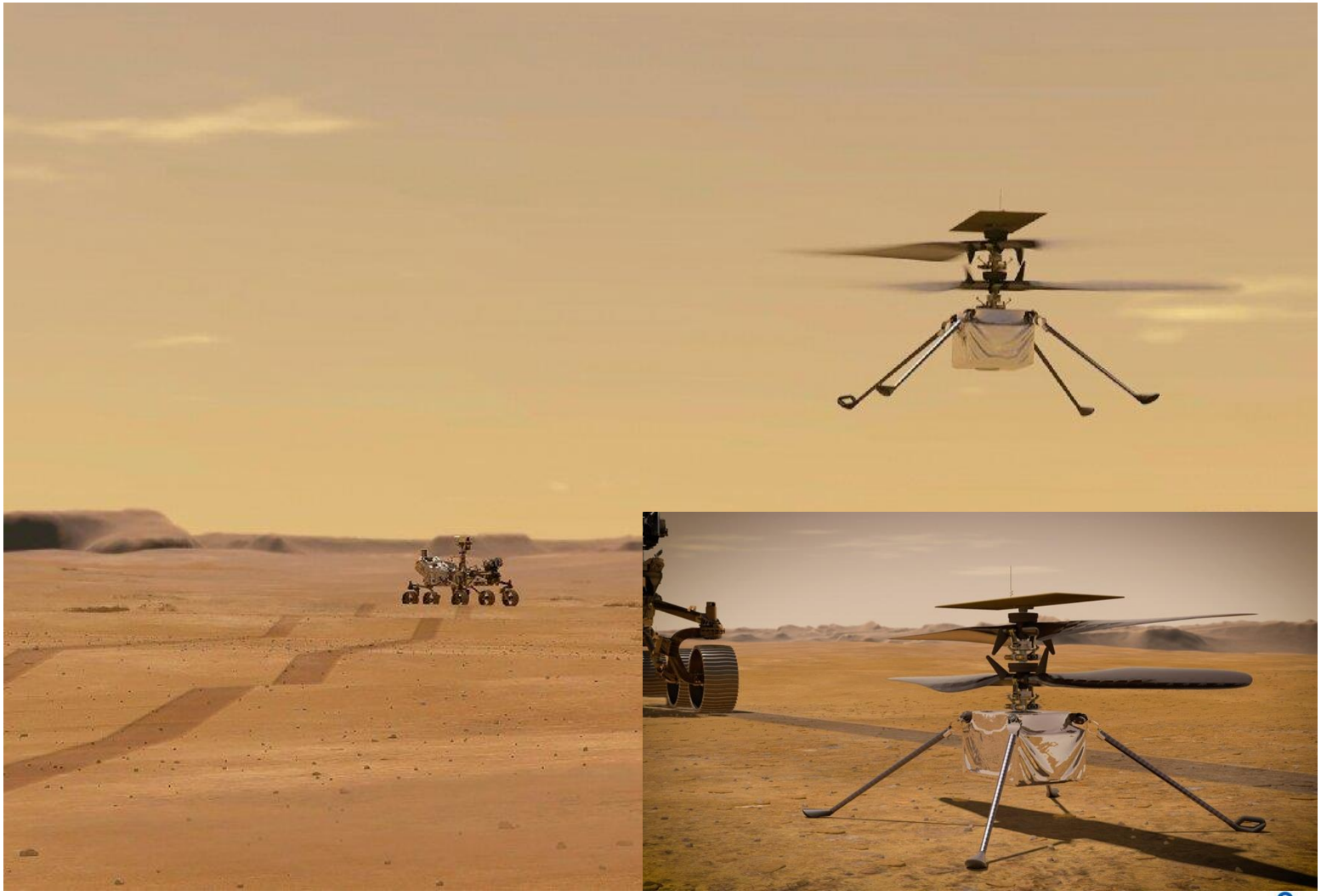
Vision in Space

NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read “Computer Vision on Mars” by Matthies et al.

Mars Exploration



Feb. 18, 2021, Jezero Crater, Mars

Perseverance rover and Ingenuity Helicopter on Mars

Credit: Håvard Fjær Grip

Techniques behind Mars Helicopter

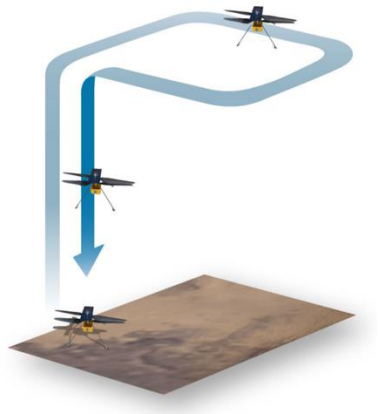
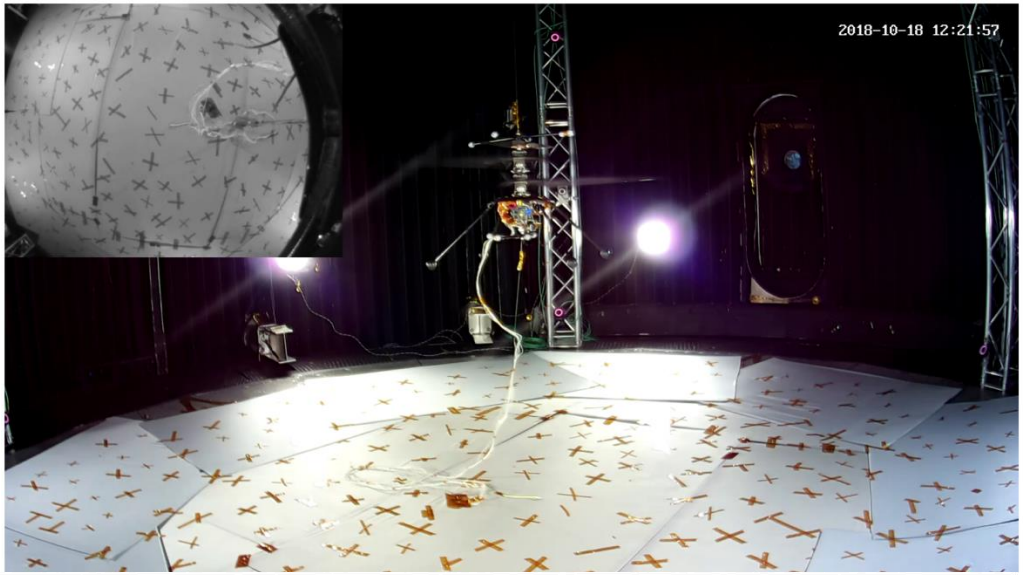
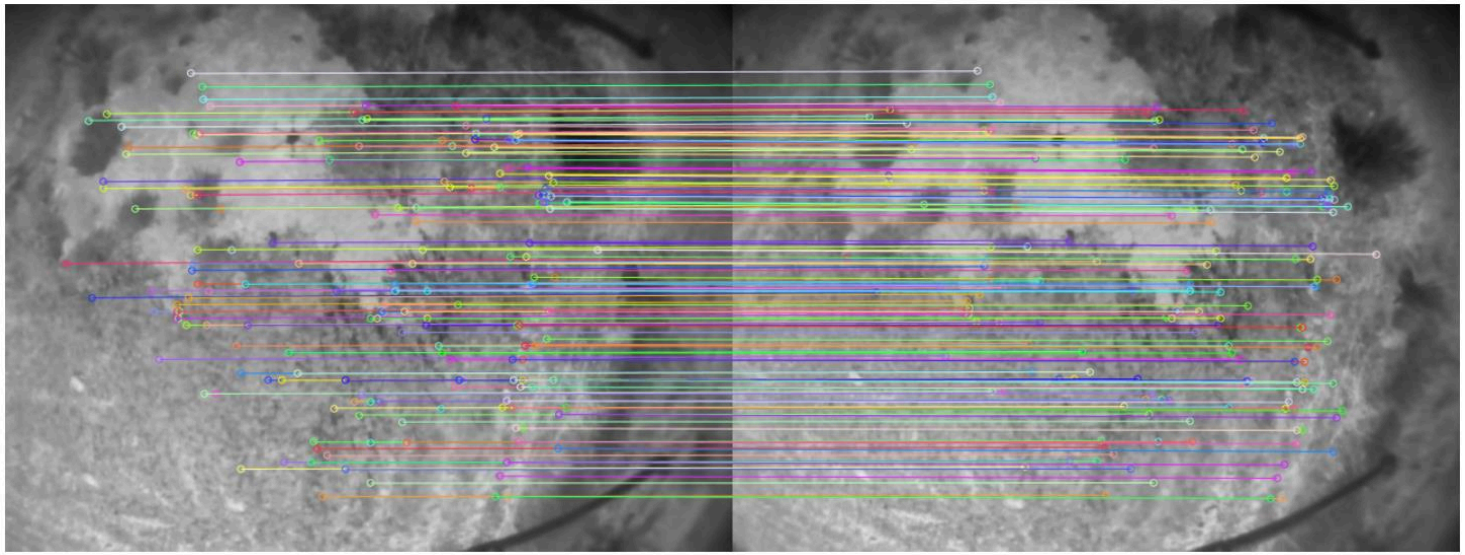


Fig. 2 Illustration of a Mars Helicopter flight, beginning and ending in the same pre-inspected safe area

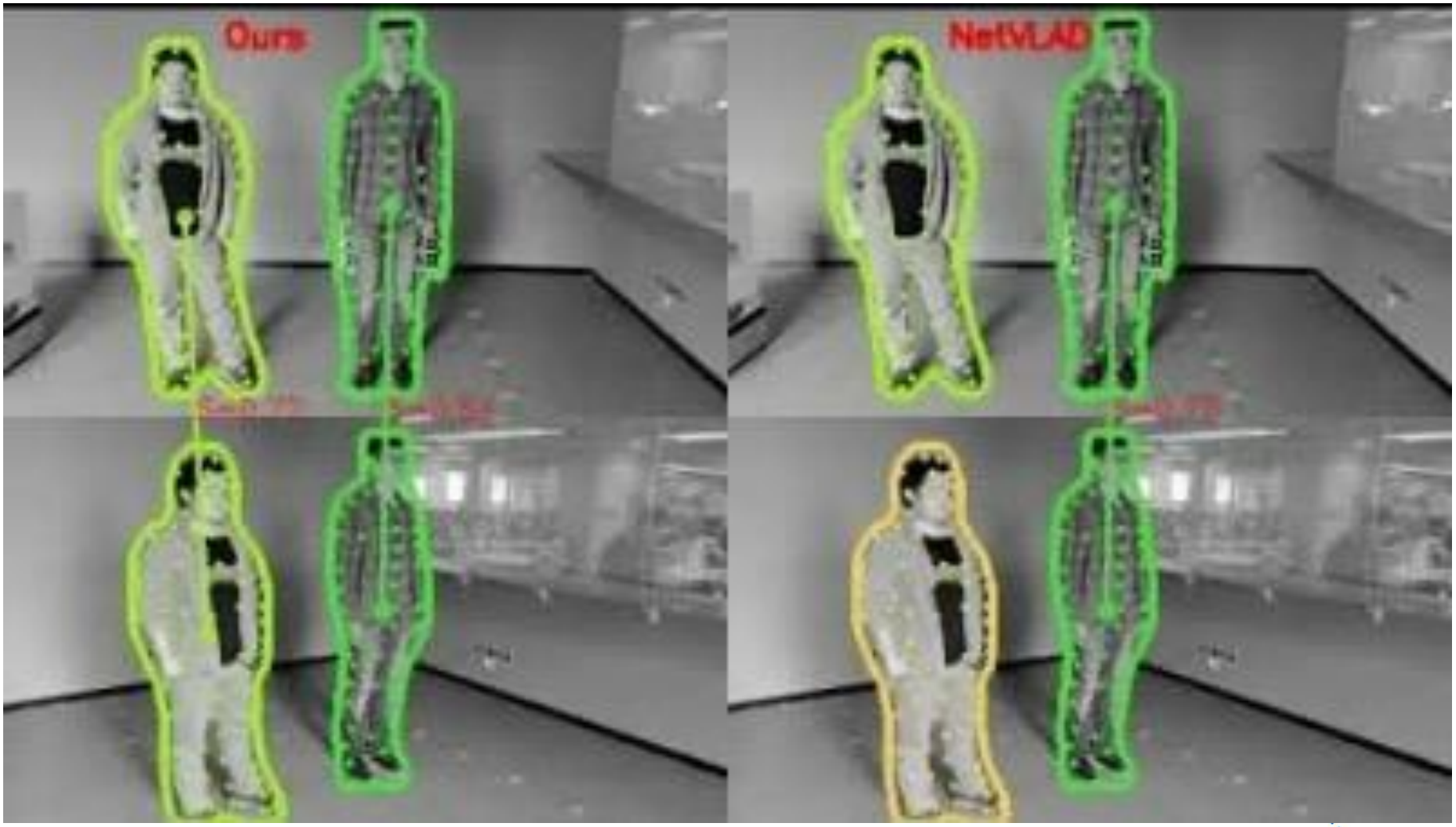




Top Venues You should know..

- Computer Vision
  - IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)
  - International Conf. on Computer Vision (ICCV)
  - European Conf. on Computer Vision (ECCV)
- Robotics (Check https://roboranking.org)
  - International Journal of Robotics Research (IJRR)
  - Transactions on Robotics (T-RO)
  - Science Robotics
  - Robotics: Science and Systems (RSS)
  - Robotics and Automation Letters (RA-L)
  - International Conf. on Robotics and Automation (ICRA)
  - International Conf. on Intelligent Robots and Systems (IROS)
- Machine Learning
  - Neural Information Processing Systems (NeurIPS)
  - International Conference on Machine Learning (ICML)
  - International Conference on Learning Representations (ICLR)

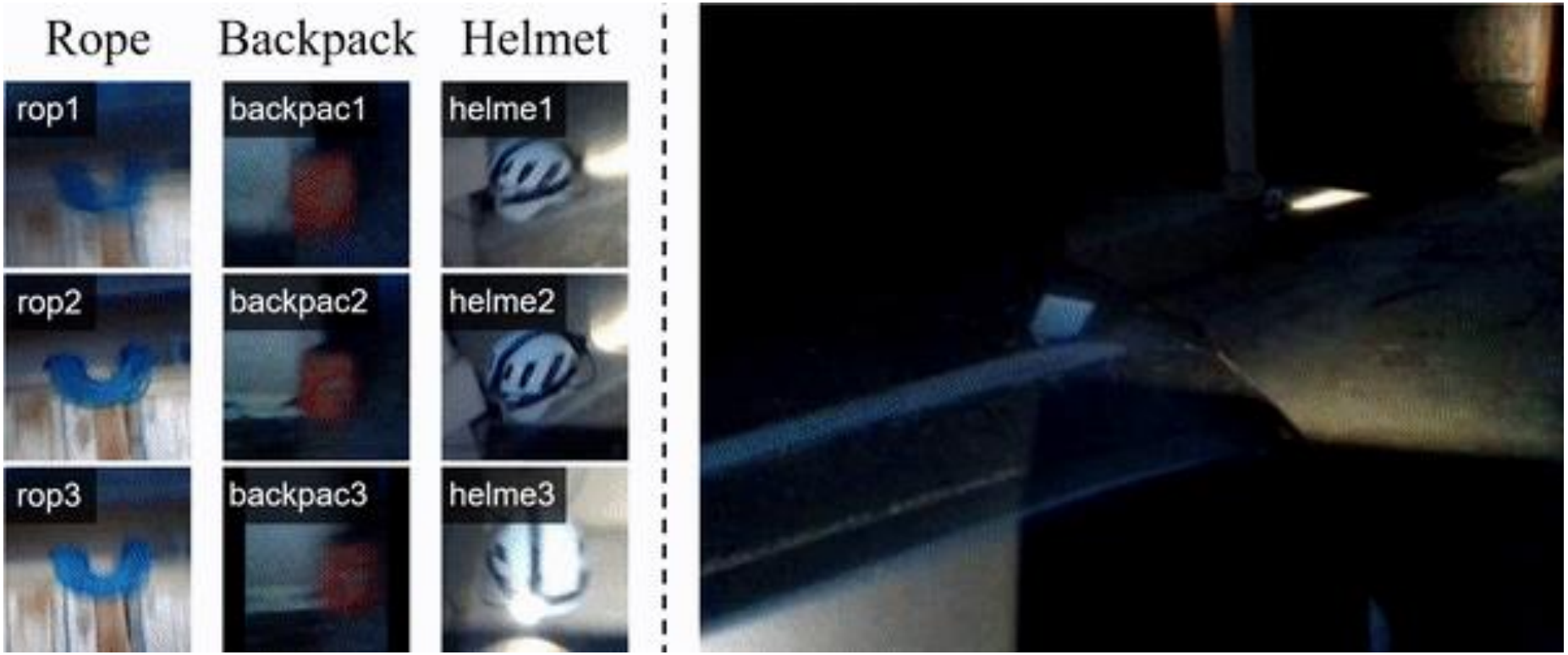
Person Re-identification



SAIR
Spatial AI & Robotics Lab

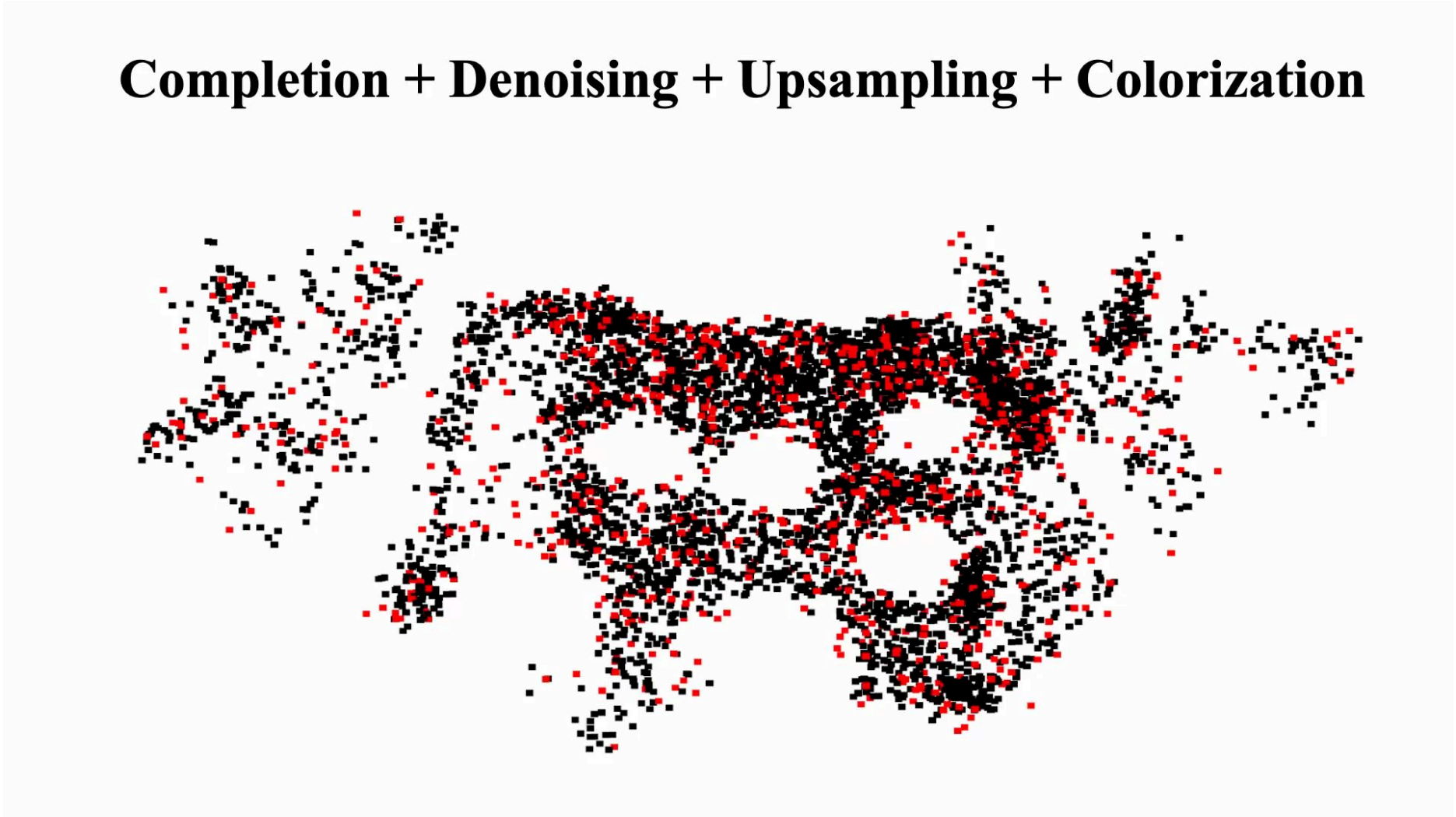
Credit: Kuan Xu (RA-L)

Search and Rescue (Few-shot Detection)



Credit: Bowen Li (ECCV)

Point Clouds

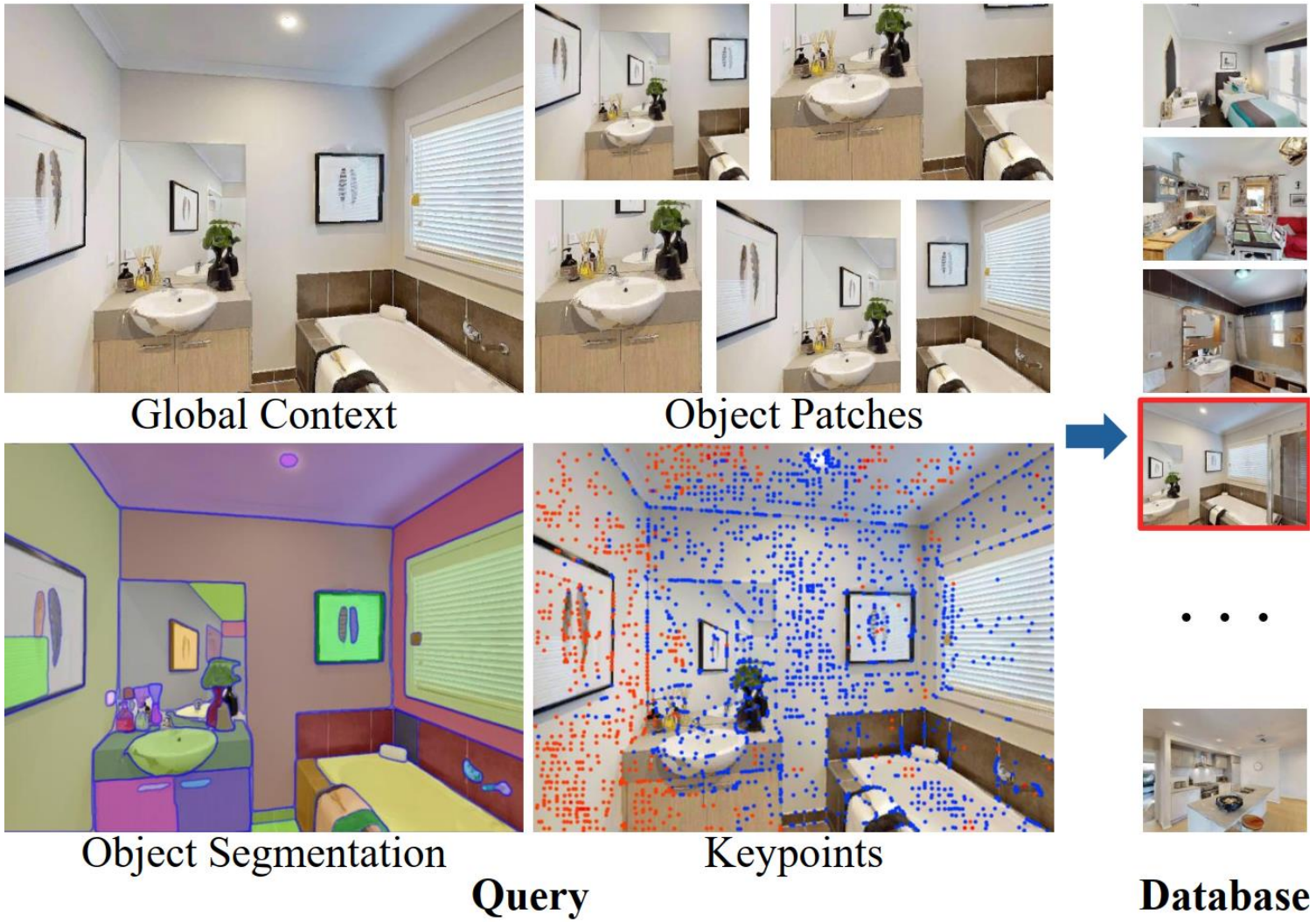


Credit: Yi Du (CVPR)

Simultaneously Localization and Mapping (SLAM)

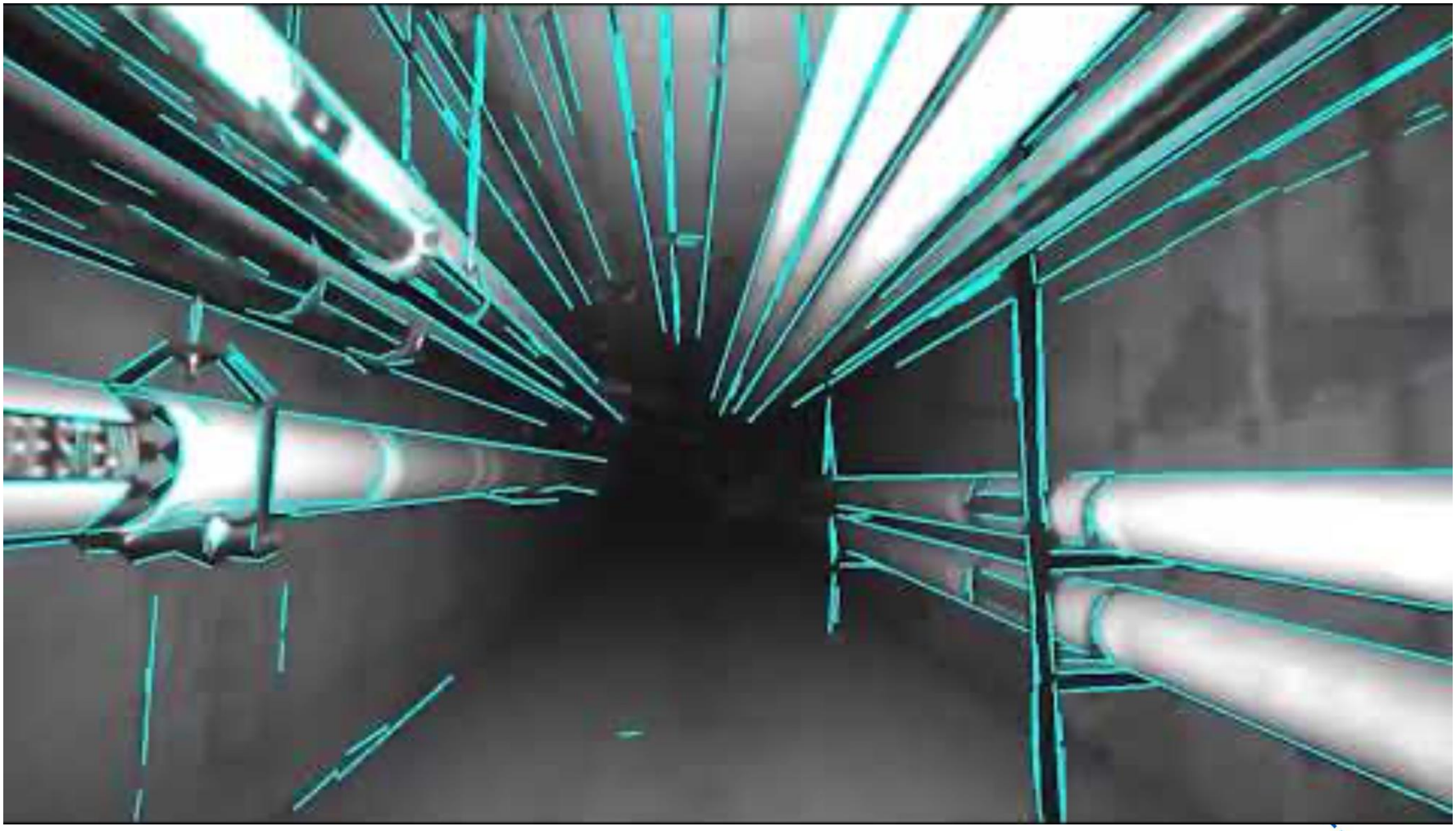
Credit: Shibo Zhao (Science Robotics)

Room Re-identification



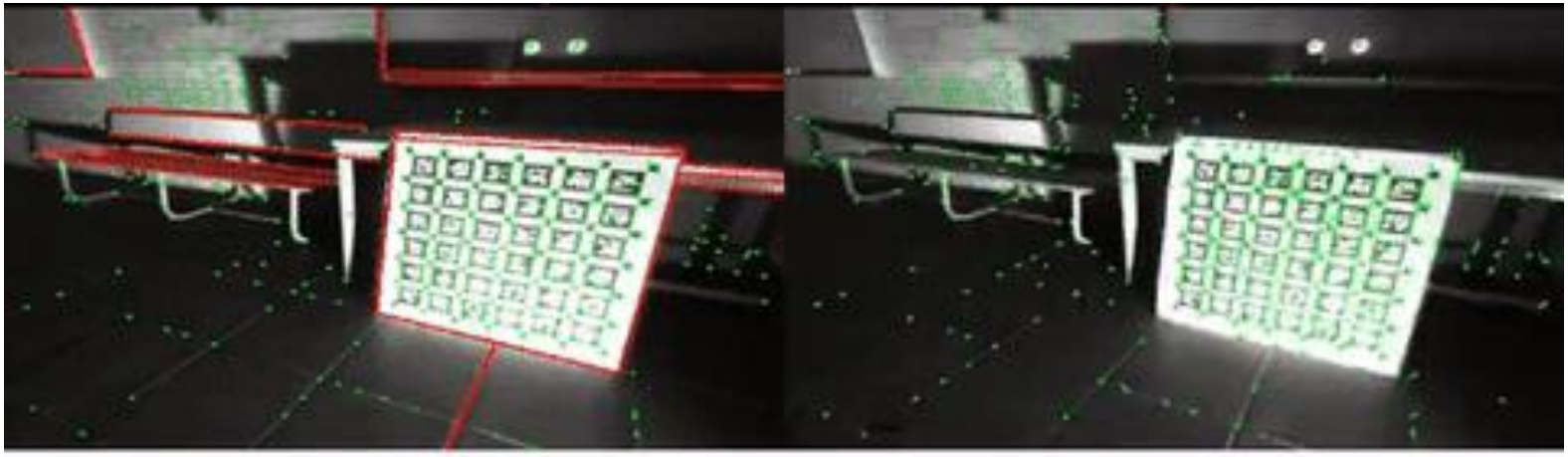
Credit: Runmao Yao (CVPR)

Line Detection



Credit: Xiao Lin (IROS)

Visual Odometry



Credit: Kuan Xu (TRO)

Warehouse Robot

Credit: Kuan Xu

Path Planning

Credit: Fan Yang (RSS)

Fast Task Planning

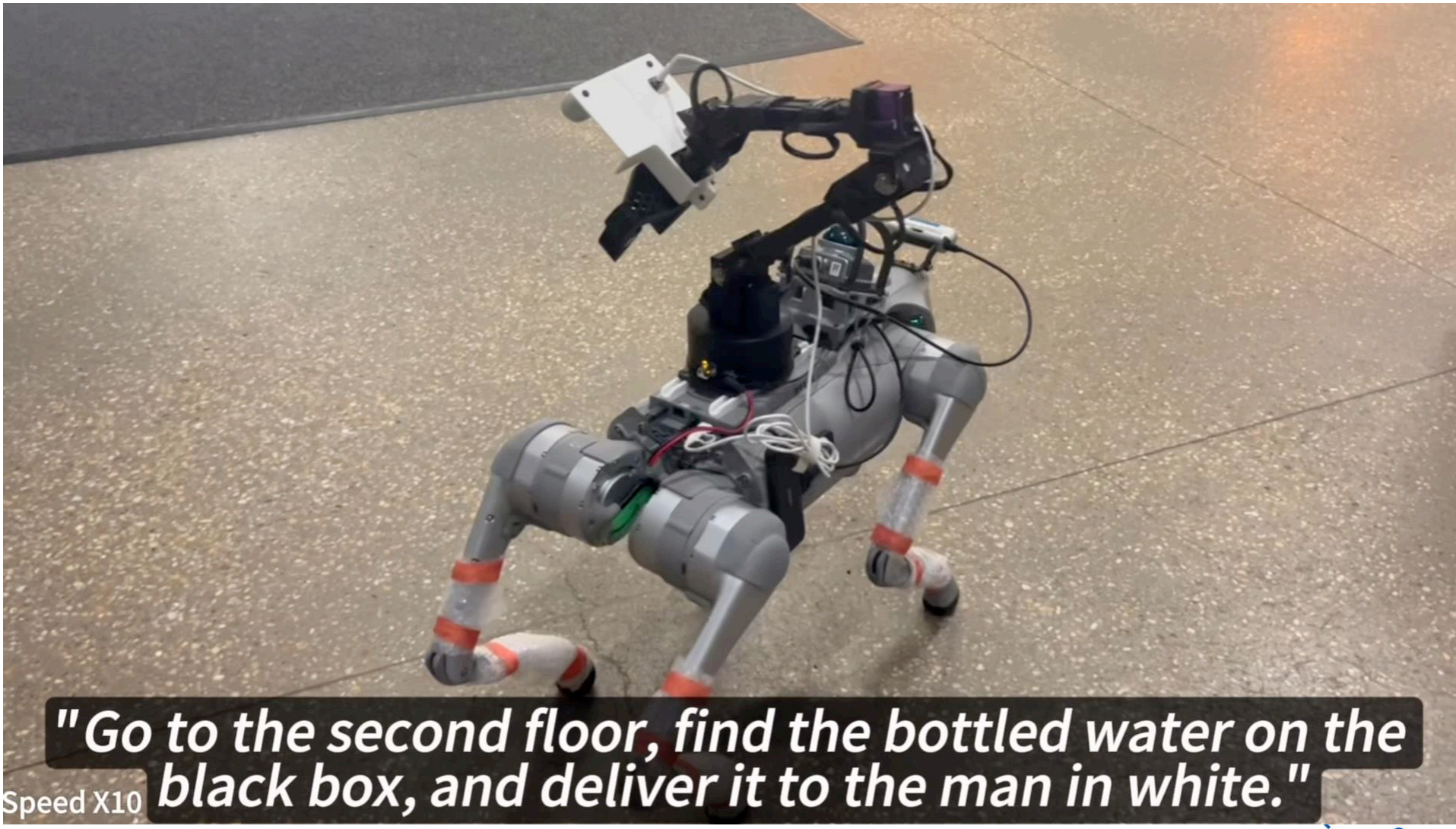
SAIR
Spatial AI & Robotics Lab

Credit: Qiwei Du

Vision-language Navigation

Credit: Yi Du

Vision-language Navigation



Credit: Yi Du

Last but not least

- Quiz 0
  - Help you familiar with UB learn system.
  - We'll take a picture for memories!