THE UNIVERSITY OF MELBOURNE

Introduction

Semester 2, 2025

Kris Ehinger

Outline

- What is computer vision?
- Welcome to COMP90086
- Overview of computer vision

What is computer vision?

Demo

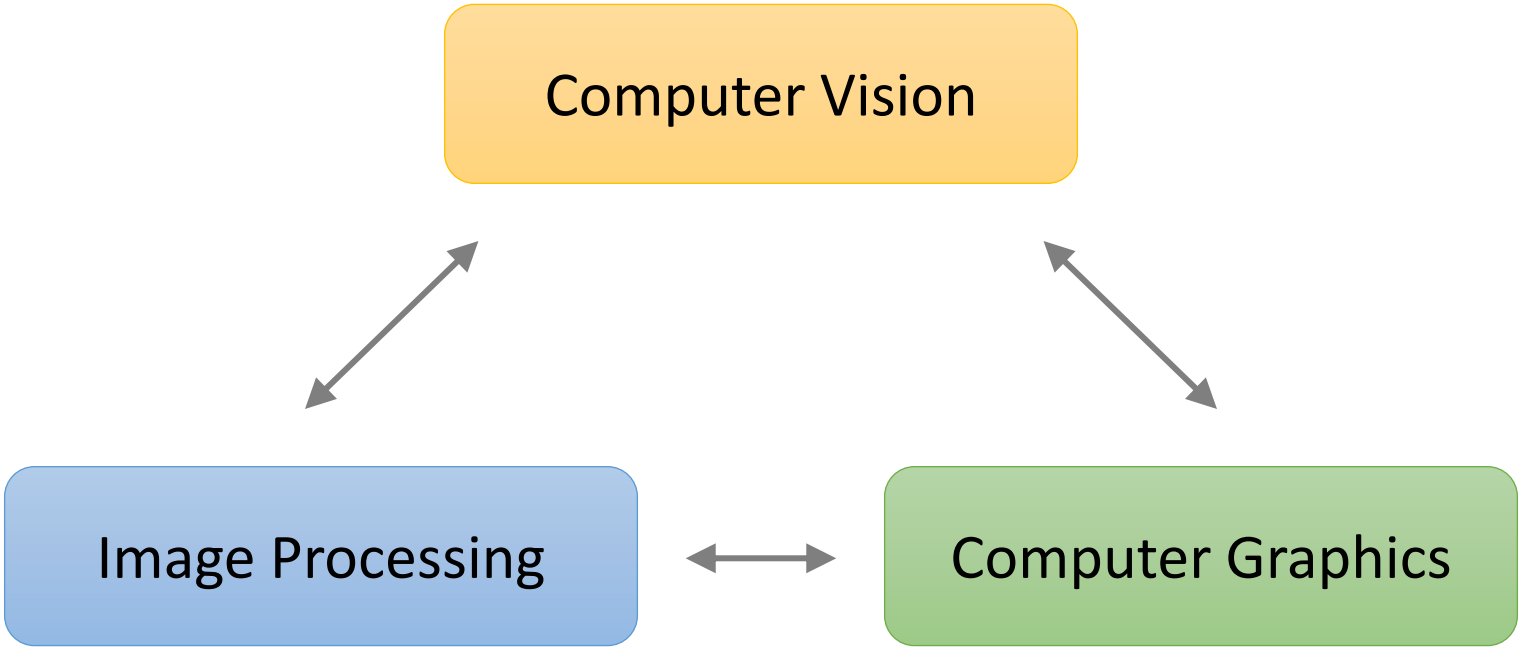
Computer vision

- Algorithms for high-level understanding of digital images or videos
- Simulate or replicate processes that a biological visual system can do
- Often interdisciplinary – AI, machine learning, physics (optics), neuroscience, psychology, art

Computer vision tasks

- Recognize images
- Localise and identify objects
- Segment image regions
- Model relations between images
- Recover 3D structure
- Perform visual navigation
- Perform visually-guided actions (e.g., grasping objects)

Related fields



Welcome to COMP90086

Lecturers

- Kris Ehinger (subject co-ordinator)
- kris.ehinger@unimelb.edu.au

- Tom Drummond
- tom.drummond@unimelb.edu.au

2D image -> 3D structure
Visual problem-solving
Human prediction / interaction

Tutors

- May Le (head tutor)

- Kazi Adnan
- Mashnoon Islam
- Yanbei Jiang
- Chao Lei
- Suhail Najeeb
- Rinu Sebastian
- Jinrui Yang
- Cipher Zhang

Student representatives

- Want to be a student representative for COMP90086?
  - Collect student opinions
  - Provide feedback to teaching staff
  - Meet with SSLC once per semester
- Email us with CV by start of Week 2 (next Monday)

Contacting us

- General inquiries: Ed forum on LMS
  - We encourage all students to join in discussions – answering other students' questions is one of the best ways to improve your own understanding
  - Please do not post sections of your code or reports publicly! If you must include these, private-message the instructors
- Personal/private concerns: Email the instructors
  - If you email us about a general inquiry, we may ask you to re-post your question in the forum
- Please include COMP90086 in email subject

Workshops (starting week 2)

Day	Start	End	Location	Tutor
Monday	11:00	12:00	379-B1-B113	Cipher
Monday	12:00	13:00	379-B1-B113	Suhail
Monday	14:00	15:00	379-B1-B113	Suhail
Tuesday	11:00	12:00	379-B1-B113	Mashnoon
Wednesday	10:00	11:00	379-B1-B117	May
Wednesday	16:00	17:00	379-B1-B114	Rinu
Wednesday	17:00	18:00	379-B1-B113	Kazi
Wednesday	18:00	19:00	379-B1-B114	Kazi
Thursday	14:00	15:00	379-B1-B132	Mashnoon
Friday	13:00	14:00	379-B1-B116	Cipher
Friday	14:00	15:00	379-B1-B116	May

Workshops (starting week 2)

Day	Start	End	Location	Tutor
Monday	11:00	12:00	379-B1-B113	Cipher
Monday	12:00	13:00	379-B1-B113	Suhail
Monday	14:00	15:00	379-B1-B113	Suhail
Tuesday	11:00	12:00	379-B1-B113	Mashnoon
Wednesday	10:00	11:00	379-B1-B117	May
Wednesday	16:00	17:00	379-B1-B114	Rinu
Wednesday	17:00	18:00	379-B1-B113	Kazi
Wednesday	18:00	19:00	379-B1-B114	Kazi
Thursday	14:00	15:00	379-B1-B132	Mashnoon
Friday	13:00	14:00	379-B1-B116	Cipher
Friday	14:00	15:00	379-B1-B116	May

Subject material

- LMS is the primary portal for the subject
  - Lecture schedule, tutorial/practical schedule
  - Content page for each week
- Lecture content
  - Handouts will be posted before lecture
  - Slides and lecture capture available after lecture
- Workshops
  - Cover content from previous week's lecture
  - Handouts posted before the first workshop
  - Solutions posted after the last workshop

Assessment

- Assignments 1-3 (20%, weeks 3-10)
  - Implement algorithms, experiment on provided data sets, and answer questions
  - Individual work
- Final project (30%, week 12)
  - Design a method to solve an open-ended computer vision problem, present algorithm and experiments in a written report
  - Work in groups of 2
- Final exam (50%, during exam period)

Late submissions / extensions

- Late policy / extension instructions are given on each assignment spec
- Extensions of up to 3 calendar days:
  - Submit documentation to us using the appropriate link on Canvas LMS (e.g., “Assignment 1 extension request”)
- Extensions > 3 calendar days:
  - Apply via Special Consideration

Academic integrity

- Assignments and exams are expected to be your **own original work**
- It's **not** your original work if it's:
  - Copied from another student
  - Taken from a website, textbook, article, github, etc.
  - Generated by an AI, chatGPT, etc.
- Any code or text that's not your own original work must be cited (including AI sources like chatGPT)
- https://academicintegrity.unimelb.edu.au/

Prerequisites

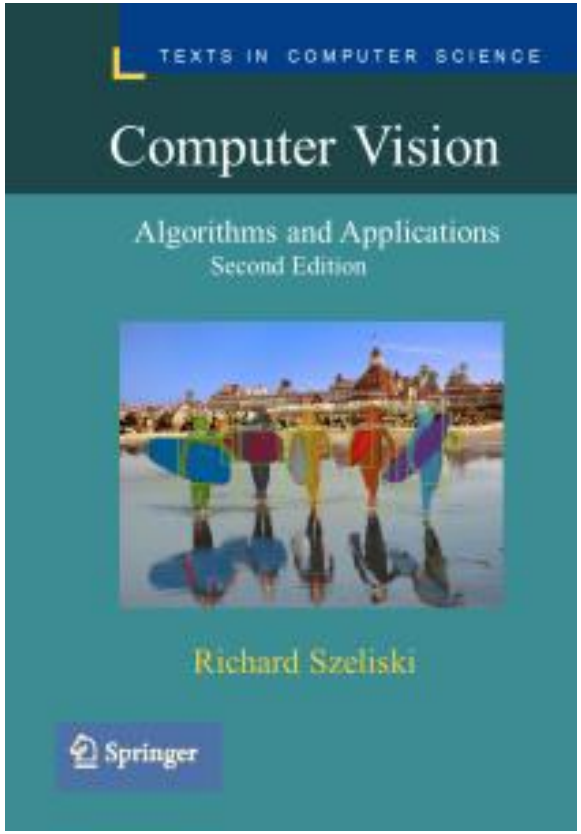
- Machine learning background
  - Training, testing, evaluating machine learning algorithms
  - Common methods like clustering, KNN, neural networks
- Programming skills
  - Workshops and assignments are in Python
- Mathematical skills
  - Basic familiarity with probability, geometry, linear algebra

Textbooks

- Suggested links and readings will be posted on LMS each week
- Readings are not required – optional links to expand your knowledge of the week's topics if you are interested

Recommended textbook

- *Computer Vision: Algorithms and Applications 2nd Ed*, Richard Szeliski (2021)
- Electronic copy available at https://szeliski.org/Book/



Other textbooks

- *Pattern Recognition and Machine Learning* by Christopher Bishop
- *Computer Vision: Models, Learning, and Inference* by Simon J.D. Prince

Content of this subject

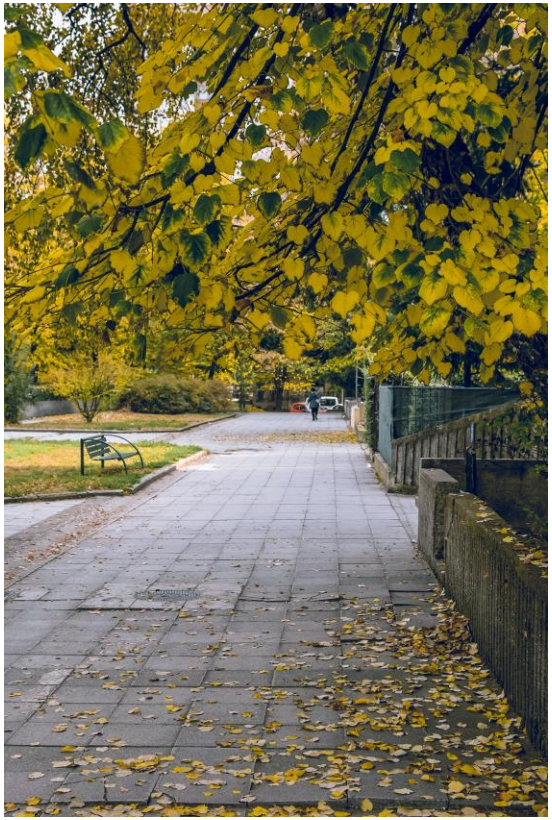
Week 1	Image formation, basics	
Week 2	Filtering	
Week 3	Recognition models and representation learning	
Week 4		Asst 1 due
Week 5		
Week 6	3D perception	
Week 7		Asst 2 due
Week 8	Generative models	
Week 9		
Week 10	Segmentation	Asst 3 due
Week 11	Object detection	
Week 12	Revision & guest lecture	Final project due

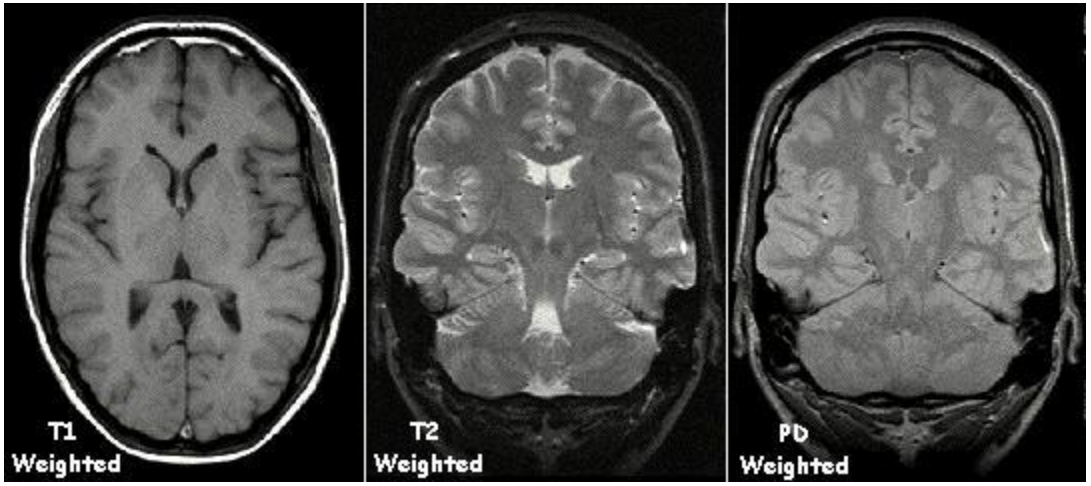
To do (this week!)

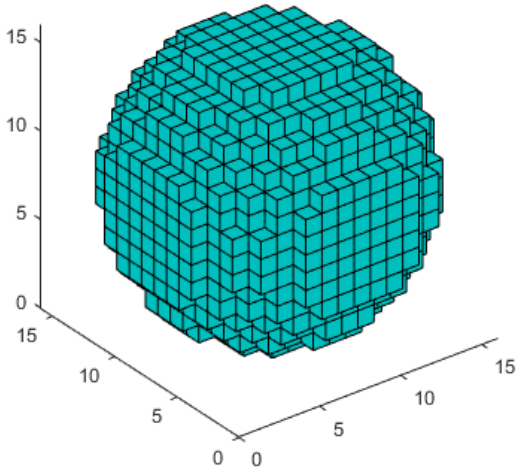
- Install Jupyter Lab / Jupyter Notebook

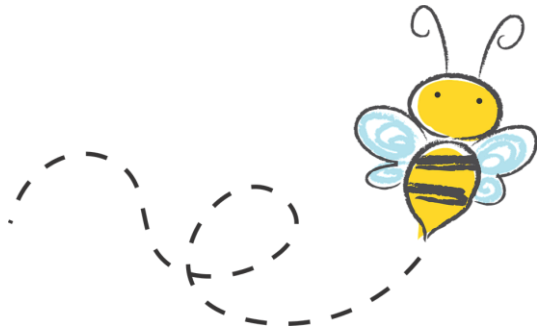
Introduction to computer vision

What is an image?









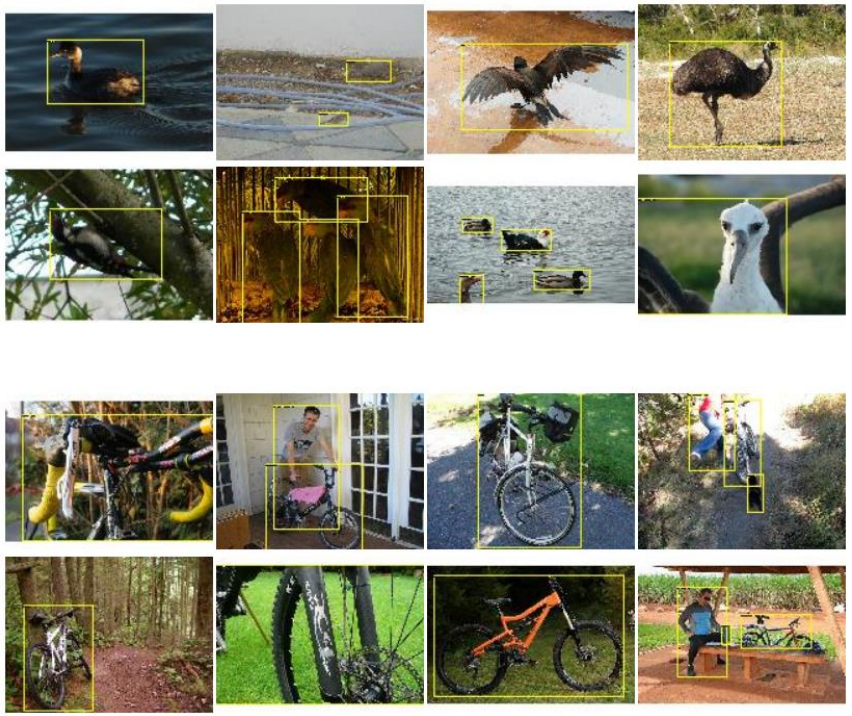
Machine learning problem

- Example: object recognition task
  - Input: image
  - Label: object class (e.g., “bicycle”)
  - Attributes: ?



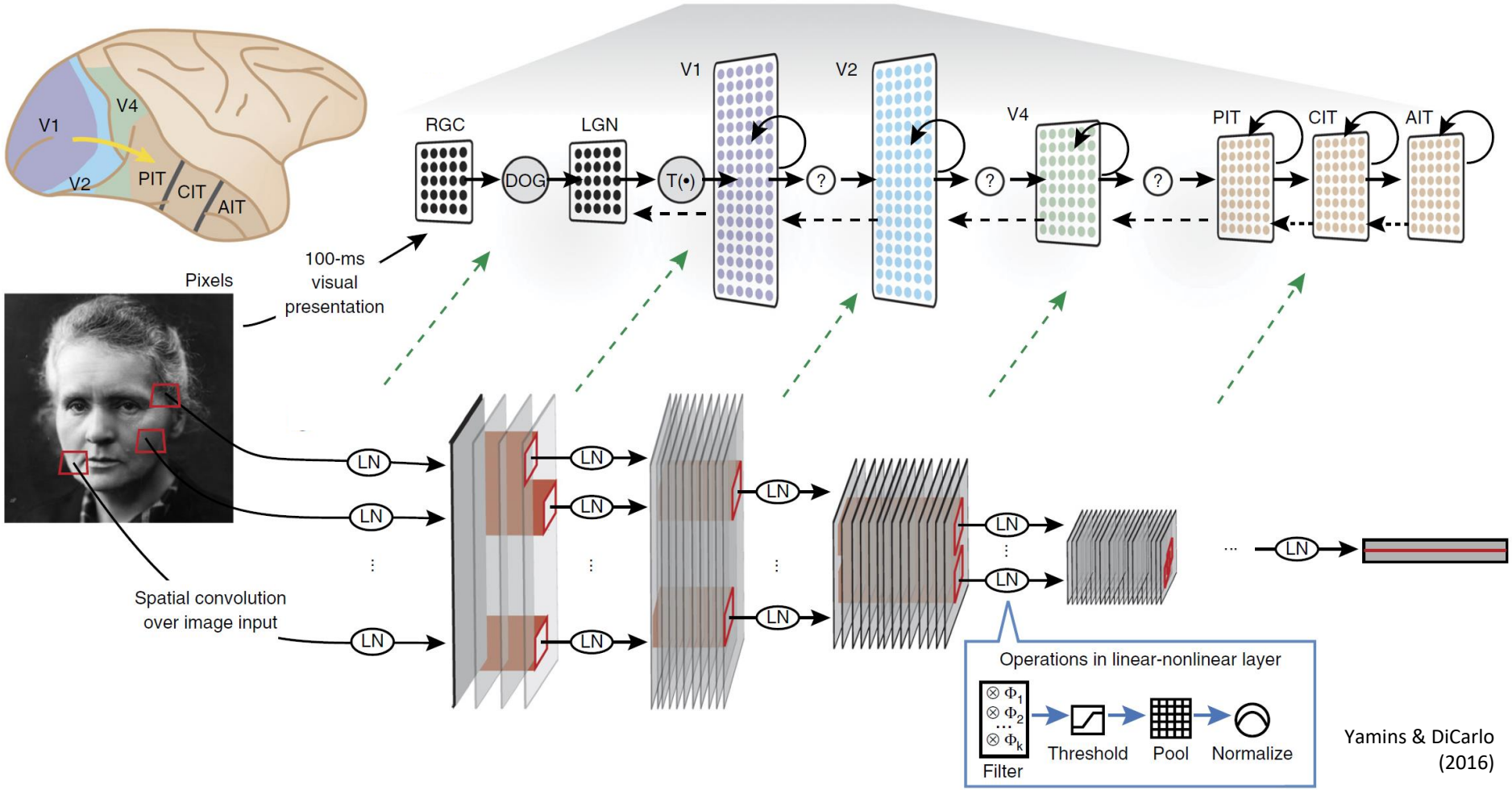
Difficulties

- Class similarity
- Within-class variation
  - Exemplars
  - Size
  - Position
  - Lighting
- Background clutter
- Occlusion



Images: PASCAL VOC2011

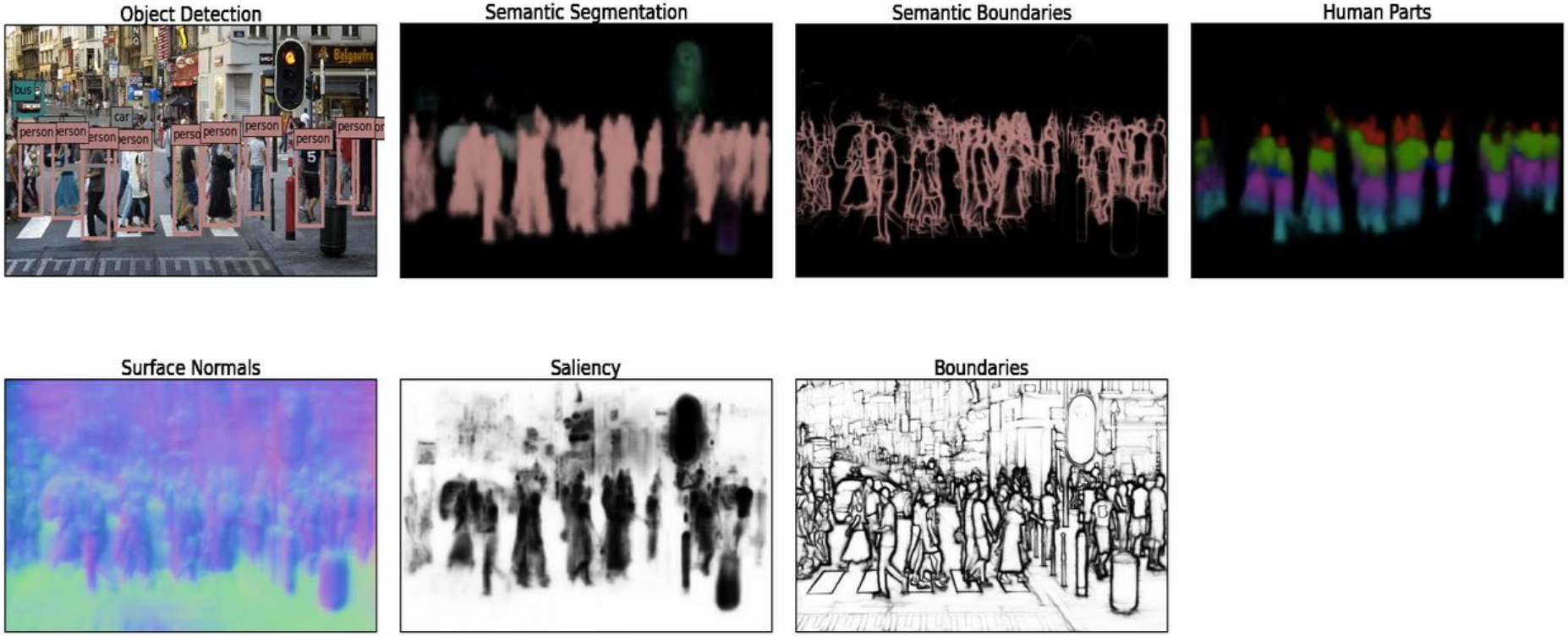
Visual encoding



Yamins & DiCarlo (2016)

Machine learning problem

- Is computer vision just classification?



Kokkinos (2016)