CSC 422/522
Computer Vision and Pattern Recognition

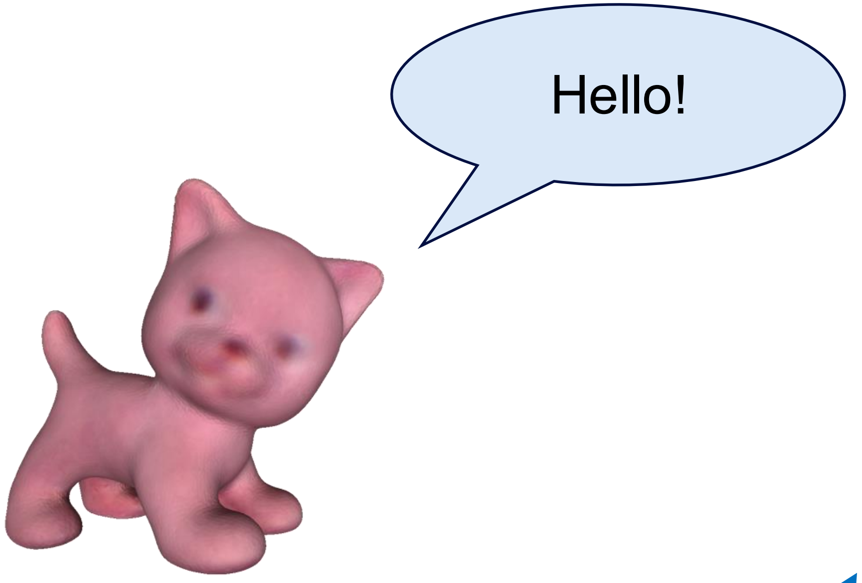
Introduction and Logistics

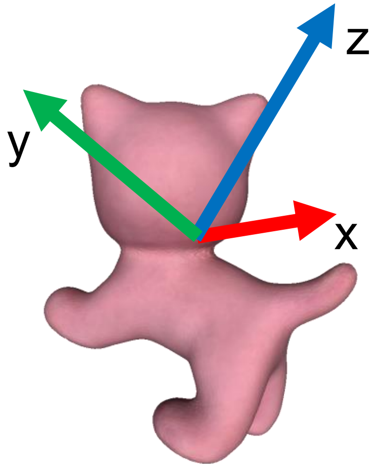
2026 Spring

South Dakota State University

Today

- Introduction to computer vision
- Course logistics

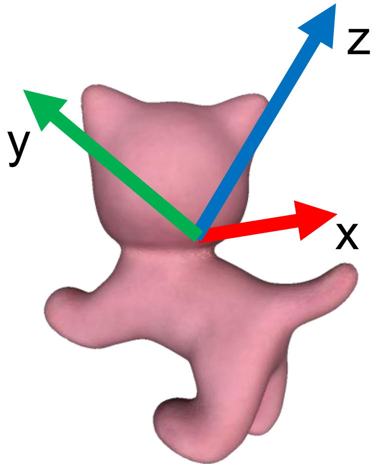




Instructor Information

- Name: Ruyi Lian
- Email: Ruyi.Lian@sdstate.edu
- Office: DEH 117
- Office Hours: Tuesday, Wednesday, and Thursday: 2:30 – 3:30 PM, or by appointment

- Education: Ph.D. in Computer Science, Stony Brook University, 2025
- Research areas:
  - **3D Computer Vision**
  - Computer Graphics
  - Robotics
  - Computational Biology
  - AI for Science



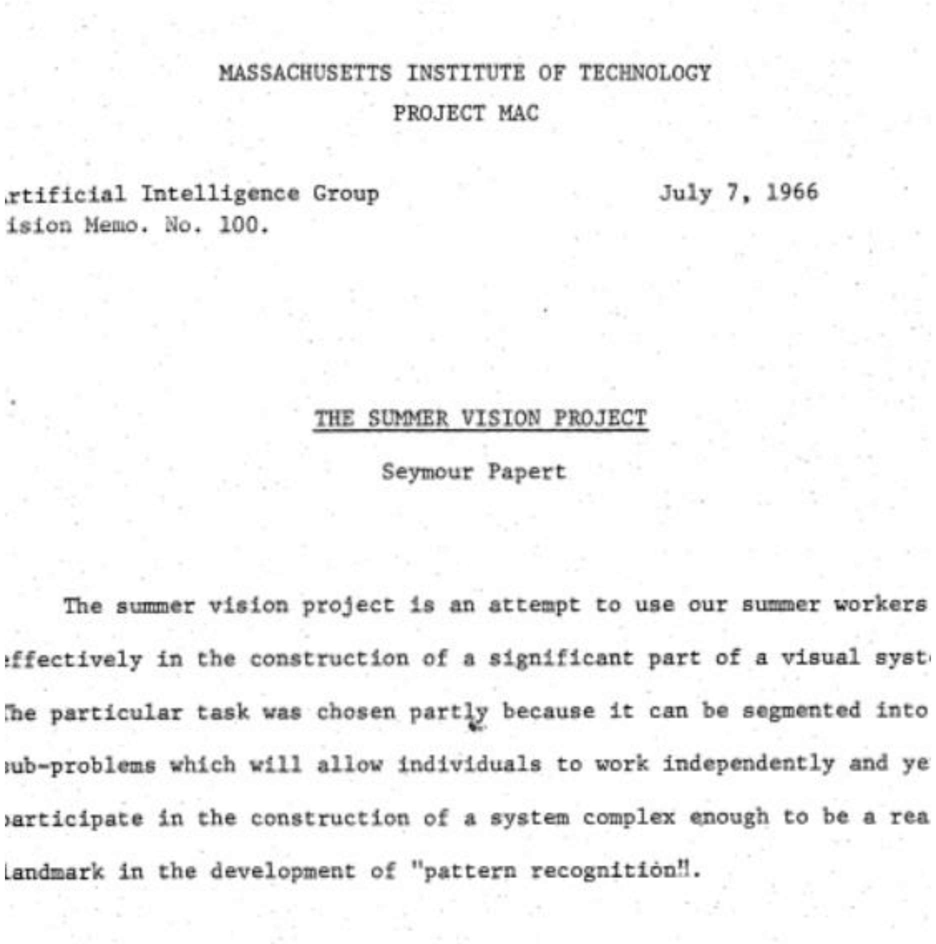
Introduction

Image source: https://medium.com/@fraidoonomarzai99/introduction-to-computer-vision-cv-in-depth-81a9ffc1b9f1

What is Computer Vision (CV)

- Computer vision is an interdisciplinary field that deals with how computers can be made to gain high-level understanding from digital images or videos.
- Official birth of CV: In 1966, Seymour Papert, an MIT professor working at the AI lab, set out on an ambitious venture known as the Summer Vision Project. He aimed to address the challenge of machine vision and achieve a solution within a short span of a few months.

Reference: https://en.wikipedia.org/wiki/Computer_vision#Definition


MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

rtificial Intelligence Group July 7, 1966
ision Memo. No. 100.

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual syst The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and ye participate in the construction of a system complex enough to be a rea landmark in the development of "pattern recognition".

Reference: https://dspace.mit.edu/handle/1721.1/6125

CV Task Example: Camera Calibration

- *Geometric camera calibration*, also referred to as *camera resectioning*, estimates the parameters of a lens and image sensor of an image or video camera.

- You can use these parameters to correct for lens distortion, measure the size of an object in world units, or determine the location of the camera in the scene.

Reference: https://www.mathworks.com/help/vision/ug/camera-calibration.html

CV Task Example: Image Enhancement

- Image enhancement refers to the process of improving the visual quality of an image through various techniques.
  - Denoising
  - Deblurring
  - Super-resolution
  - Contrast enhancement
  - …

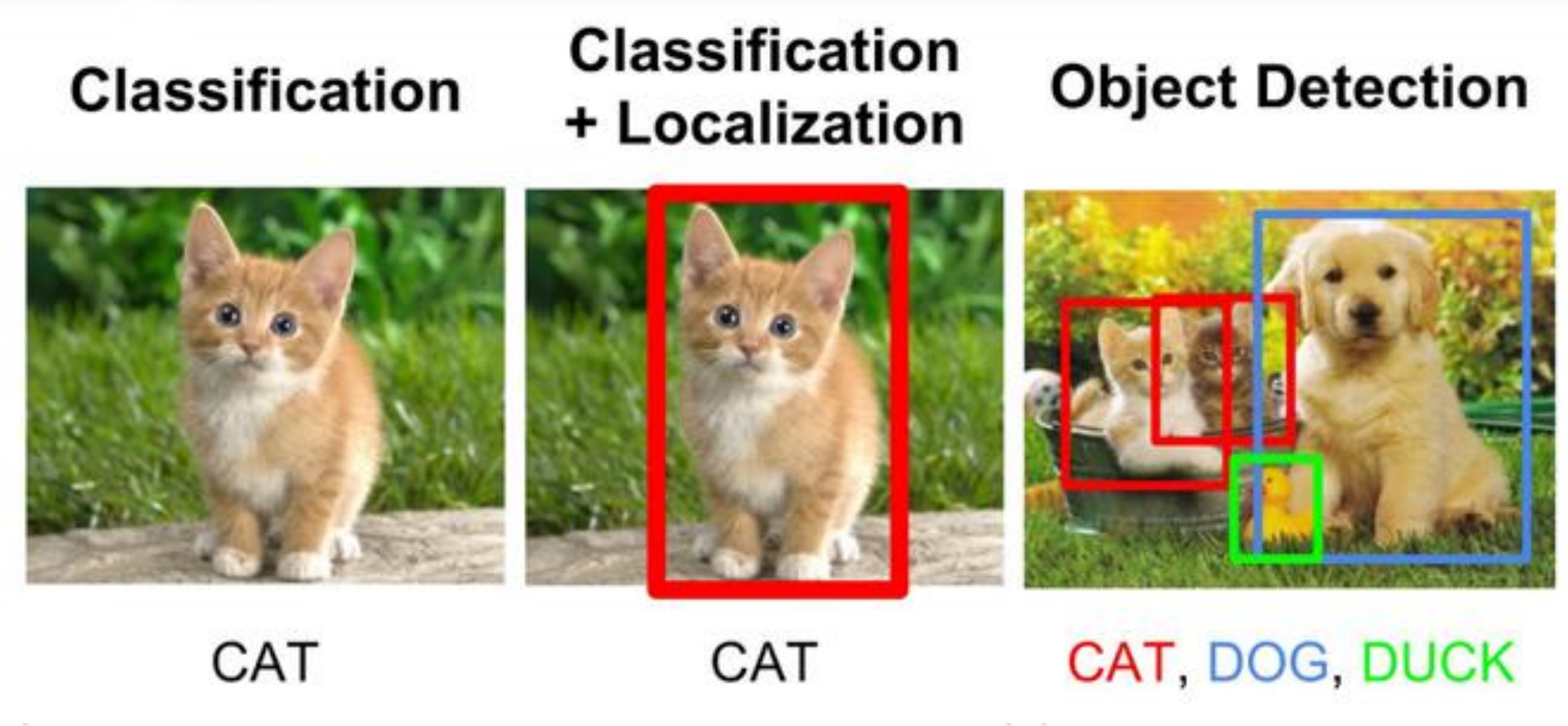
Source: https://www.mathworks.com/discovery/image-enhancement.html

CV Task Example: Image Stitching

- Image stitching is the process of combining multiple photographic images with overlapping fields of view to produce a segmented panorama or high-resolution image.

Source: https://en.wikipedia.org/wiki/Image_stitching

CV Task Example: Classification, Localization, and Object Detection



Source: https://www.geeksforgeeks.org/computer-vision/what-is-image-classification/

CV Task Example: Object Tracking

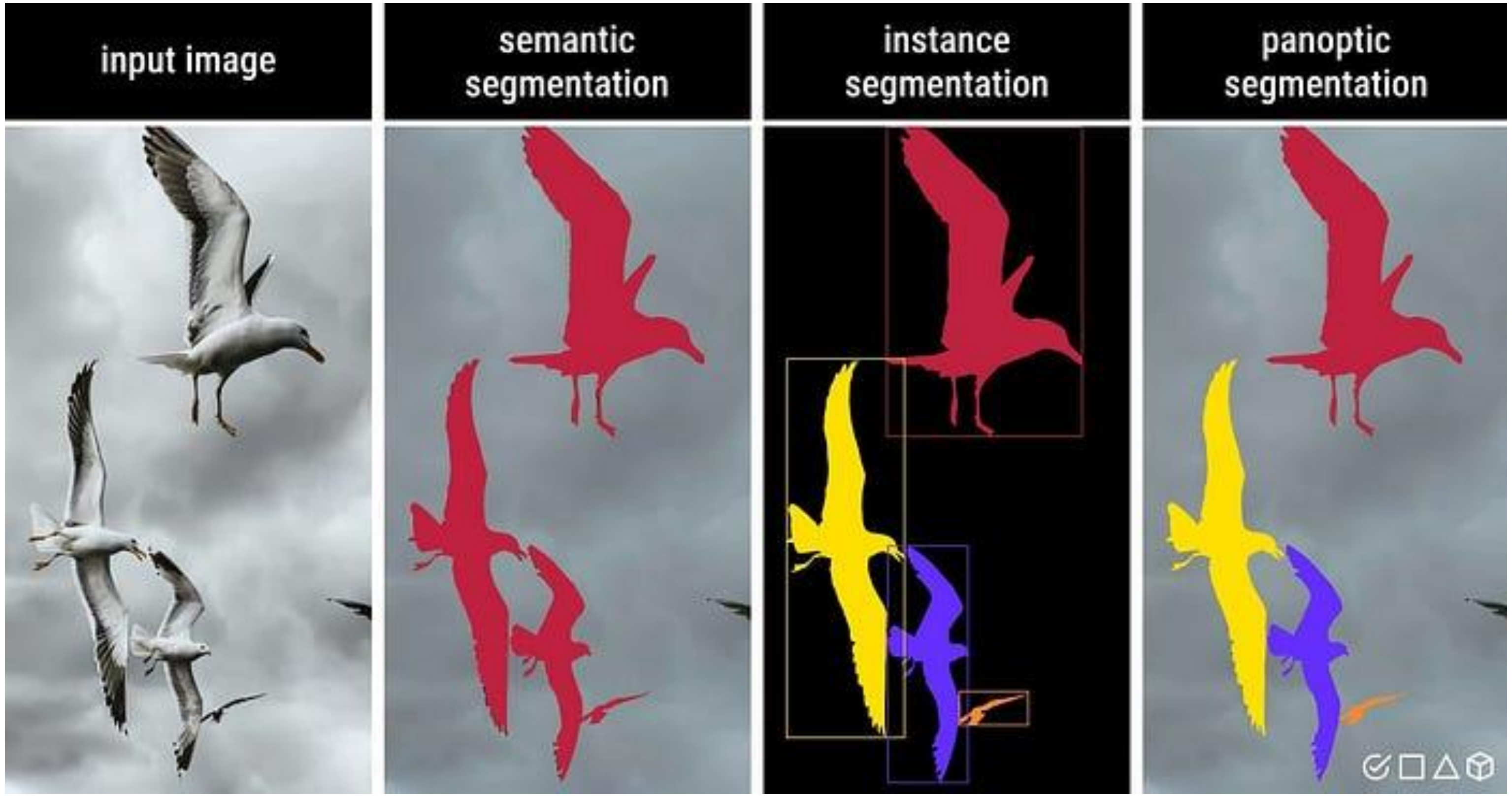
Source: https://hengfan2010.github.io/projects/TOTB/

CV Task Example: Image Segmentation

- **Semantic segmentation** classifies each pixel based on its semantic class. All the birds belong to the same class.

- **Instance segmentation** assigns unique labels to different instances, even if they are of the same semantic class. Each bird belongs to a different class.

- **Panoptic segmentation** combines the two, providing both class-level and instance-level labels. Each bird has its own class, but they are all identified as a "bird".



Source: https://medium.com/@raj.pulapakura/image-segmentation-a-beginners-guide-0ede91052db7

CV Task Example: Depth Estimation

- Depth estimation is the task of predicting depth of the objects present in an image.

Source: https://huggingface.co/tasks/depth-estimation

CV Task Example: Optical Flow

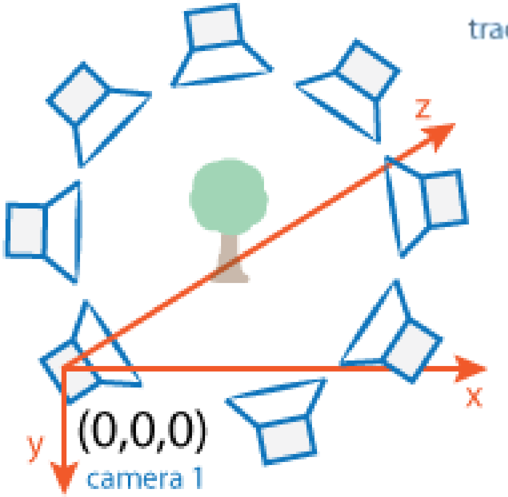
- Optical flow is the pattern of apparent motion of image objects between two consecutive frames caused by the movement of object or camera.
- It is 2D vector field where each vector is a displacement vector showing the movement of points from first frame to second.

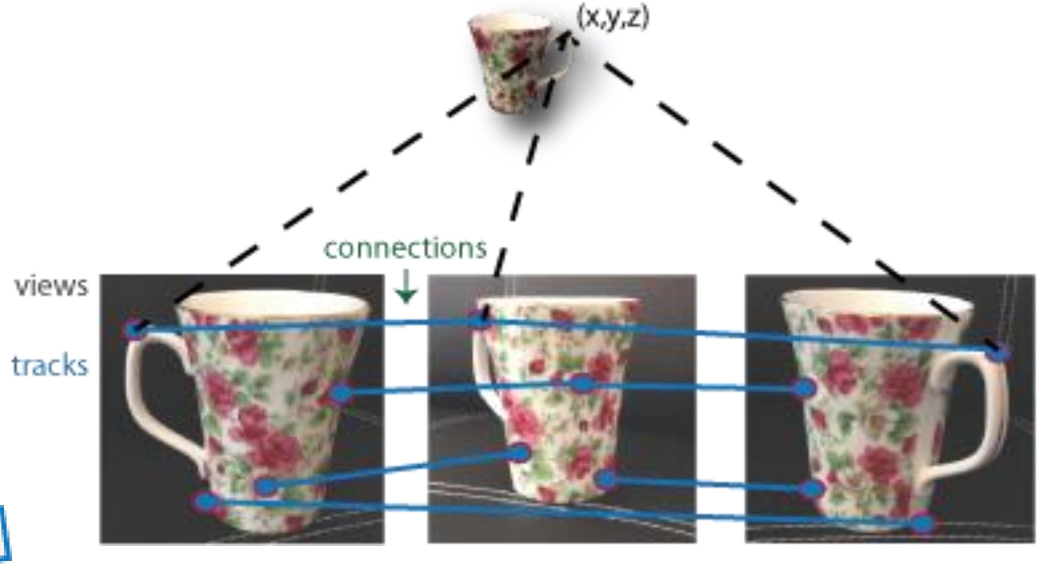
Reference: https://docs.opencv.org/3.4/d4/dee/tutorial_optical_flow.html

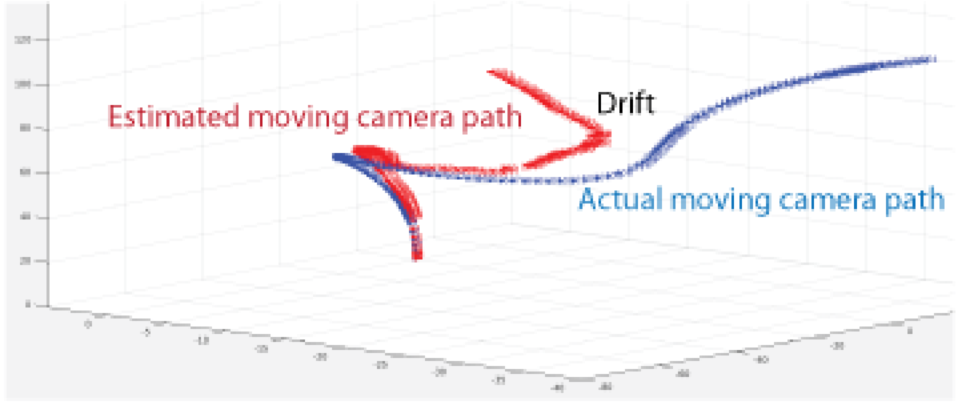
CV Task Example: Structure from Motion

- Structure from motion (SfM) is the process of estimating the 3D structure of a scene from a set of 2D images.

- SfM is used in many applications, such as 3D reconstruction, 3D scanning, augmented reality, and visual simultaneous localization and mapping (vSLAM).



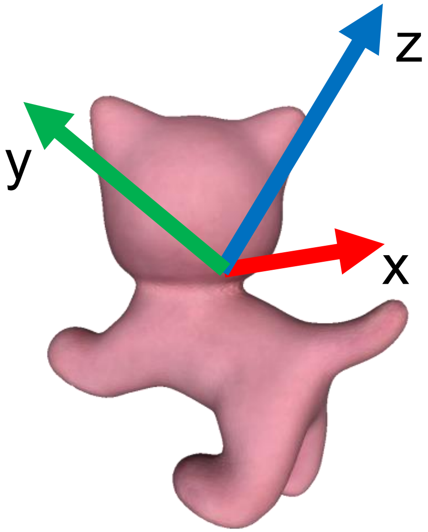




Reference: https://www.mathworks.com/help/vision/ug/what-is-structure-from-motion.html

CV Task Example: Pose Estimation

Input: Image



Output: Pose

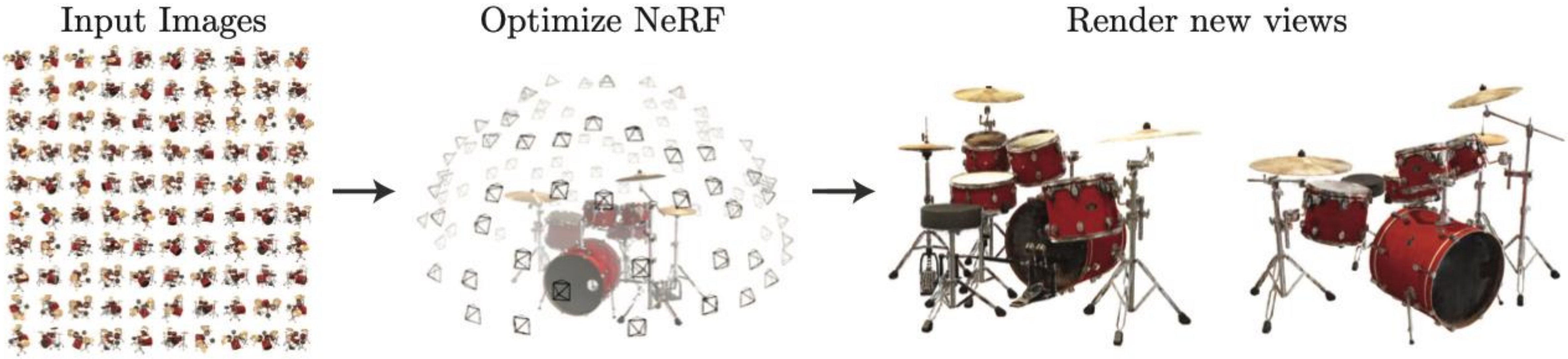
(Rotation + Translation)

The target object (toy cat) is from Linemod dataset.

CV Task Example: 3D Reconstruction

Source: https://3dprintingindustry.com/news/nvidia-ai-produces-3d-models-from-2d-videos-3d-printing-applications-forthcoming-222532/

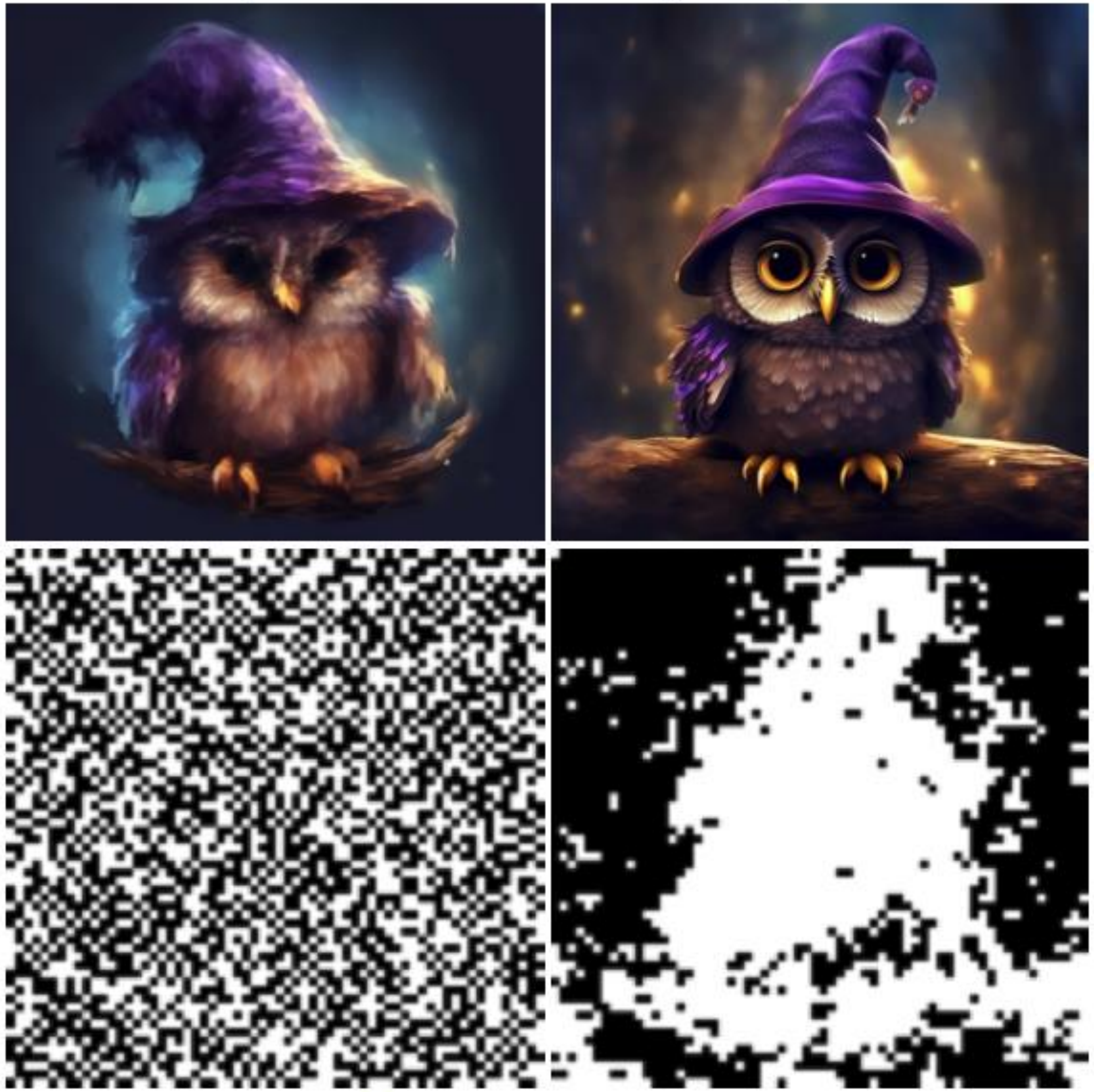
CV Task Example: Neural Rendering and Novel View Synthesis



Mildenhall, Ben, et al. "Nerf: Representing scenes as neural radiance fields for view synthesis." *Communications of the ACM* 65.1 (2021): 99-106.

CV Task Example: Image Generation

Prompt: “*A cute owl wearing a wizard hat.*”



(a) Spatial [3] (b) Importance-based (Ours)

Wu, Haoyu, et al. "Importance-based token merging for efficient image and video generation." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2025.

Computer Vision Is a Rapidly Expanding Field

- Computer vision extends well beyond a fixed set of canonical tasks.
- The research community continuously redefines the field's scope.
- For example, in the latest CVPR *(the top conference in computer vision field)*, topics of interest cover all aspects of computer vision and pattern recognition including, but **not limited to the topics**:
  - 3D from multi-view and sensors
  - 3D from single images
  - Adversarial attack and defense
  - …

Full list: https://cvpr.thecvf.com/Conferences/2026/CallForPapers

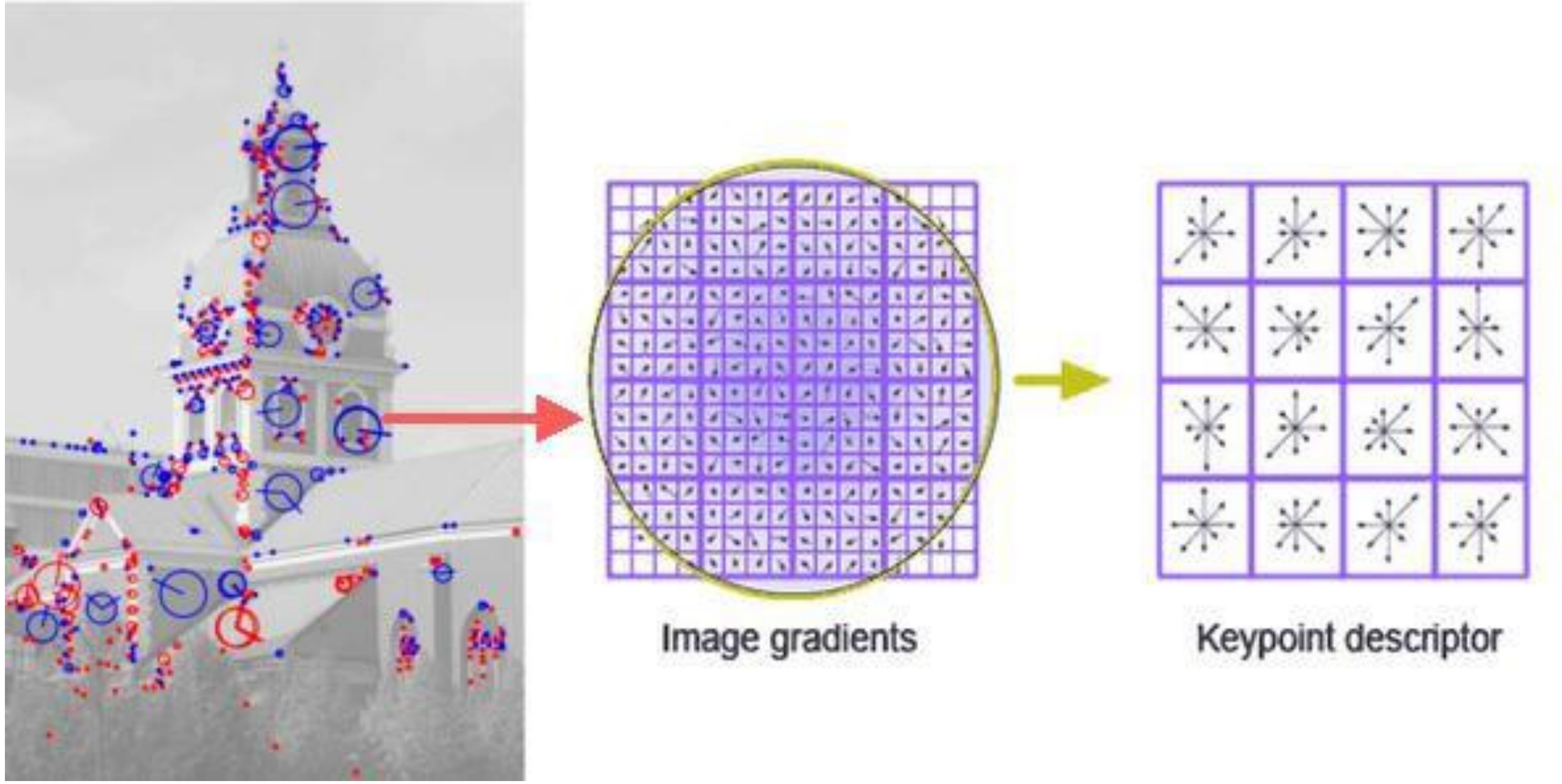
CVPR 2026 Call for Papers

Papers in the main technical program must describe high-quality, original research. Topics of interest cover all aspects of computer vision and pattern recognition including, but not limited to:

- 3D from multi-view and sensors
- 3D from single images
- Adversarial attack and defense
- Autonomous driving
- Biometrics
- Computational imaging
- Computer vision for social good
- Computer vision theory
- Datasets and evaluation
- Deep learning architectures and techniques
- Document analysis and understanding
- Efficient and scalable vision
- Embodied vision: Active agents, simulation
- Event-based cameras
- Explainable computer vision
- Humans: Face, body, pose, gesture, movement
- Image and video synthesis and generation
- Low-level vision
- Machine learning (other than deep learning)
- Medical and biological vision, cell microscopy
- Multimodal learning
- Optimization methods (other than deep learning)
- Photogrammetry and remote sensing
- Physics-based vision and shape-from-X
- Recognition: Categorization, detection, retrieval
- Representation learning
- Computer Vision for Robotics
- Scene analysis and understanding
- Segmentation, grouping and shape analysis
- Self-, semi-, meta- and unsupervised learning
- Transfer/ low-shot/ continual/ long-tail learning
- Transparency, fairness, accountability, privacy and ethics in vision
- Video: Action and event understanding
- Video: Low-level analysis, motion, and tracking
- Vision + graphics
- Vision, language, and reasoning
- Vision applications and systems

Full list: https://cvpr.thecvf.com/Conferences/2026/CallForPapers

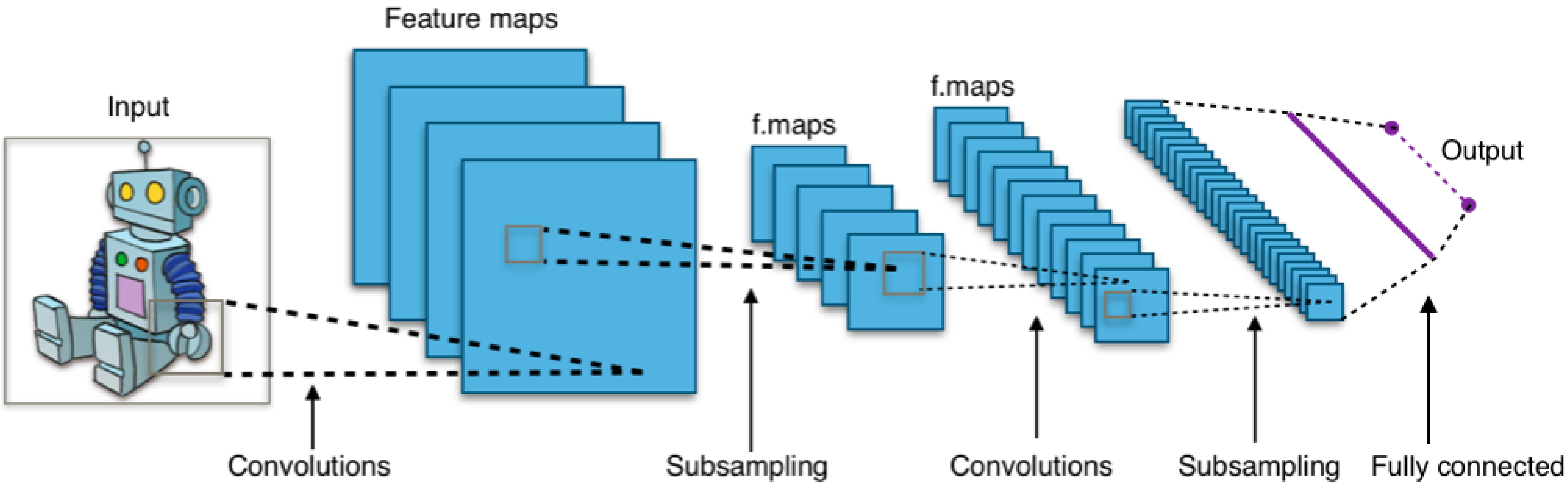
How: Classical Vision vs. Deep Learning



One traditional way of feature extraction

Source: https://naadispeaks.blog/2018/08/12/deep-learning-vs-traditional-computer-vision/

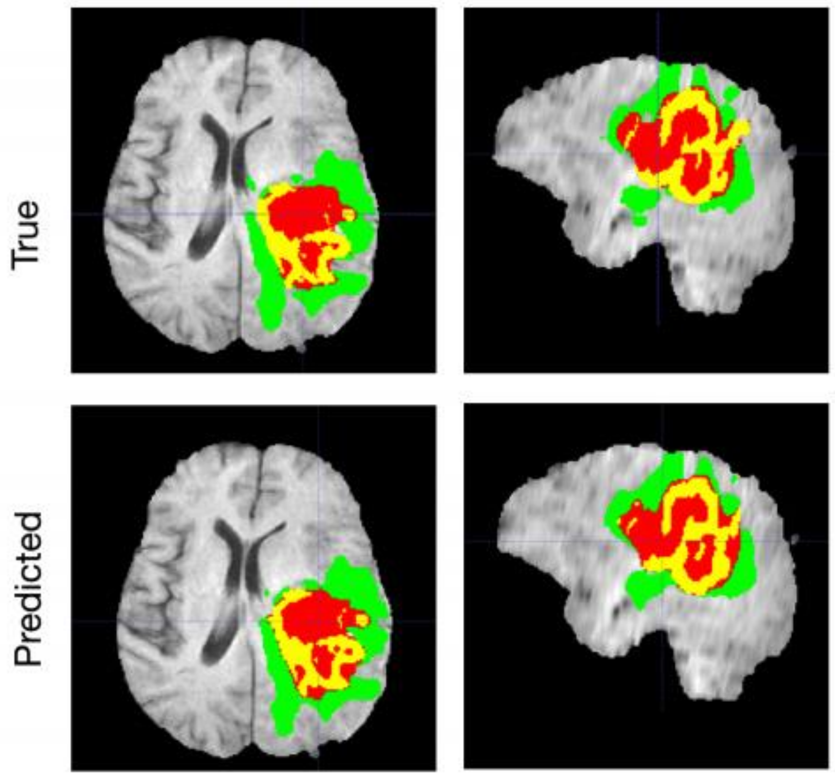
How: Classical Vision vs. Deep Learning



Convolutional Neural Network (CNN)
Source: https://en.wikipedia.org/wiki/Convolutional_neural_network

Why Computer Vision Matters

- Real-world applications, including but not limited to:
  - Healthcare: medical imaging
  - Autonomous systems: driving, drone, robotics
  - AR, VR
  - Scientific discoveries via analysis of visual data
- Learning how computer vision systems work help us understand the current limitations and risks (e.g., bias, privacy, brittleness, black-box behavior, etc.).



Tumor Segmentation in MRI

Image source: https://developer.nvidia.com/blog/automatically-segmenting-brain-tumors-with-ai/

What Are Your Motivations?

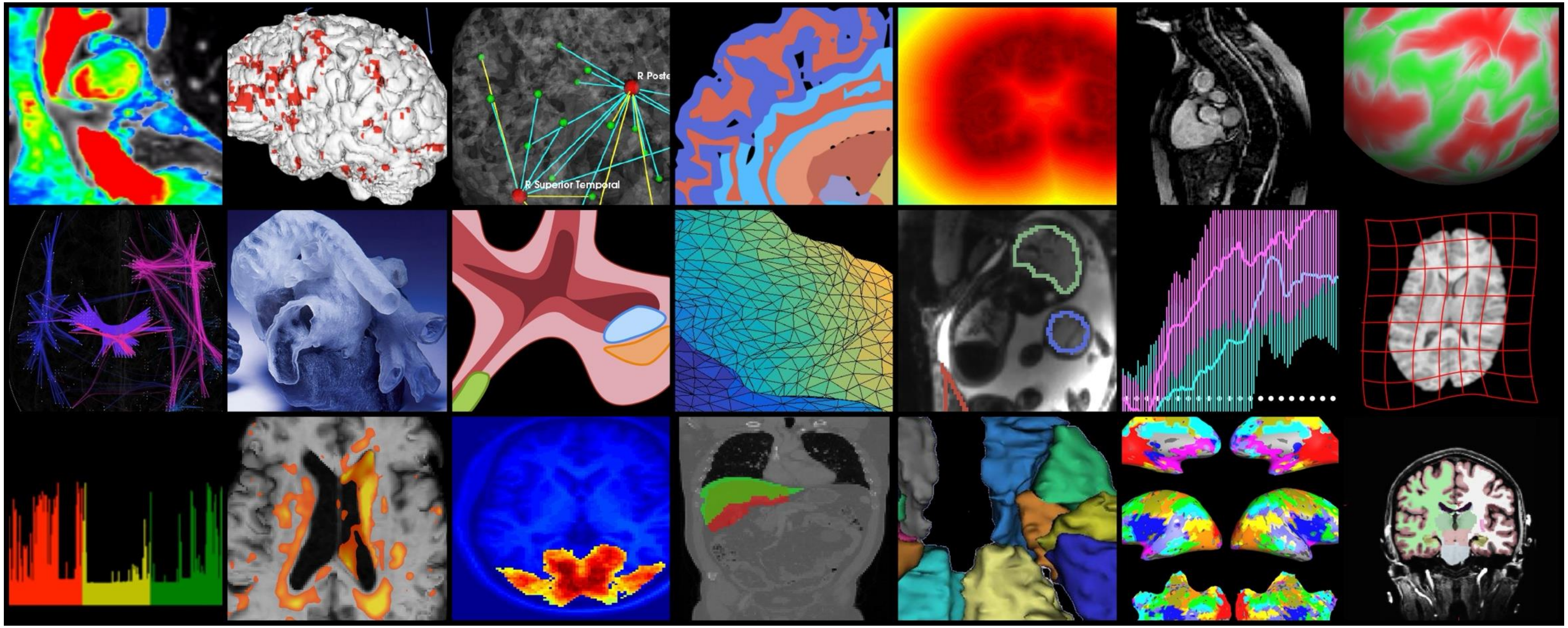


Image source: https://www.csail.mit.edu/research/medical-vision-group

Logistics

Image source: https://onlinedegrees.sandiego.edu/wp-content/uploads/2023/04/What-Is-Computer-Vision_An-Introduction.jpg

Course Information

Course Number:	CSC 422/522
Title:	Computer Vision and Pattern Recognition
Credits:	3
Academic Term/Year:	Spring, 2026
Course Description:	This course introduces fundamentals and applications, and recent advances in computer vision and pattern recognition. Topics covered in this course include single-view and two-view geometry, multi-view 3D reconstruction and registration, 2D/3D features and descriptors, object detection and tracking, and sematic scene understanding using deep convolutional neural networks (CNN). Students will have hands-on experience in implementing computer vision algorithms and deep CNN architectures in solving real-world problems.
Course Prerequisites:	MATH 250 (for CSC 422)
Instructional Mode:	Face-to-Face

Textbook

- No required textbook.
- Optional references:
  - "Computer Vision: Algorithms and Applications" by Richard Szeliski, Springer, 2nd Edition. (freely available online)
  - "Multi View Geometry in Computer Vision" by Richard Hartley and Andrew Zisserman, Cambridge, 2nd Edition.
- Given the rapid evolution of computer vision, you are encouraged to read recent papers and online blogs.

Grading Policy

CSC 422		CSC 522	
Exam #1	100 points	Exam #1	100 points
Exam #2	100 points	Exam #2	100 points
Final Project	150 points	Final Project	150 points
Assignments	200 points	Assignments	200 points
Participation	50 points	Participation	50 points
		Research Paper Review	100 points
Total	600 points	Total	700 points

<u>Grading Scale:</u>

90 – 100%	A
80 – 89%	B
70 – 79%	C
60 – 69%	D
0 – 59	F

Participation

- Pop quizzes will be given from time to time, which will be counted toward the participation grade.

Assignments

- There will be three coding assignments to practice non-deep-learning-based algorithms and train a simple CNN.
- They will be primarily scheduled in the first half of the semester, allowing the second half of the semester to be devoted to the final project.

Late Policy

- Assignments are due **at the beginning of class** on their respective due dates.
- Late assignments will have a reduction in points of 30% for the next class period and **absolutely** no late assignments will be accepted after the first reduction.

Programming Language

- The recommended programming language for this course is Python, particularly for implementing CNNs, is widely adopted in the computer vision community.
- Other programming languages, such as C++, MATLAB, etc., are permitted.
- We mainly use these Python libraries: NumPy, OpenCV, PyTorch.
- Please reach out to me if you need help getting started with Python.

Final Project

- The final project will be evaluated based on code implementation, written report, and in-class presentation.
- Collaborative projects are only permitted for CSC 422. Teams are limited to two members. A brief justification of workload distribution is required.
- In-class presentation dates: April 28 and April 30 (before the final week).

Exams

- The exam dates will be announced at least one week in advance. The tentative exam schedule is as follows:
  - Exam #1: March 12, class meeting time (the Thursday before the Spring Break)
  - Exam #2: May 8, 4-6PM (according to SDSU Spring Final Exam Schedule)

Exam Absence and Make-Up Policy

- Students who need to miss an exam are required to notify me in advance.
- Acceptable reasons for missing exams include business, medical, or school related circumstances. A written excuse **MUST** be provided.
- Emergency situations (e.g., medical or family emergency) should be handled by notifying the Dean of Student Affairs (688-4493).
- **Excused** make-up exams will only be given for (i) absences resulting from documented emergency situations or (ii) university excused trips; students must present the completed approved trip absence card to the faculty member prior to the trip or event.
- The make-up exam **MUST** be completed within one week of the original exam date; failure to do so will result in a grade of zero.

Research Paper Review (CSC 522 Only)

- Due with Final Project
- Format requirements:
  - 1 page
  - Single-spaced
  - 12 points
  - Times New Roman
- Choose one paper that you are interested in from: https://openaccess.thecvf.com/CVPR2025?day=all

Tentative Course Schedule, Part 1

13-Jan	Introduction and Logistics
15-Jan	Image Formation: Geometric Primitives and Transformations
20-Jan	Image Formation: Camera Geometry
22-Jan	Image Processing: Point Operators
27-Jan	Image Processing: Filtering
29-Jan	Image Processing: Fourier Transforms
3-Feb	Feature Detection, Matching, RANSAC
5-Feb	Image Alignment and Stitching
10-Feb	Motion Estimation (Emphasizing Optical Flow)
12-Feb	Structure from Motion and SLAM I
17-Feb	Structure from Motion and SLAM II

Note: subject to change (Please check the latest version posted in D2L)

Tentative Course Schedule, Part 2

19-Feb	Neural Networks: Basic Architecture
24-Feb	Neural Networks: Training and Inference
26-Feb	CNN: Basic Architecture
3-Mar	CNN: Backpropagation
5-Mar	CNN Based Classification
10-Mar	Review Topics for Exam 1
12-Mar	Exam 1
17-Mar	Spring Break
19-Mar	Spring Break
24-Mar	CNN Based Object Detection
26-Mar	CNN Based Object Tracking
31-Mar	CNN Based Segmentation

Note: for deep learning part, this course is more application driven and focus on implementation with PyTorch. If you want to learn comprehensive theories, please take the deep learning course :)

Note: subject to change (Please check the latest version posted in D2L)

Tentative Course Schedule, Part 3

2-Apr	Depth Estimation (emphasizing Epipolar Geometry)
7-Apr	temp: Pose Estimation, 3D Reconstruction
9-Apr	temp: Neural Rendering and Scene Undrestanding
14-Apr	Generative Models
16-Apr	Transformers and Foundation Models
21-Apr	Application of CNNs in Medical/Scientific Imaging
23-Apr	Review Topics for Exam 2
28-Apr	Final Project Presentation
30-Apr	Final Project Presentation
8-May	Exam 2

Note: subject to change (Please check the latest version posted in D2L)

Questions?

Image source: https://imgflip.com/i/44tk2h