UC San Diego

Deep Learning Lab on Computer Vision Problems

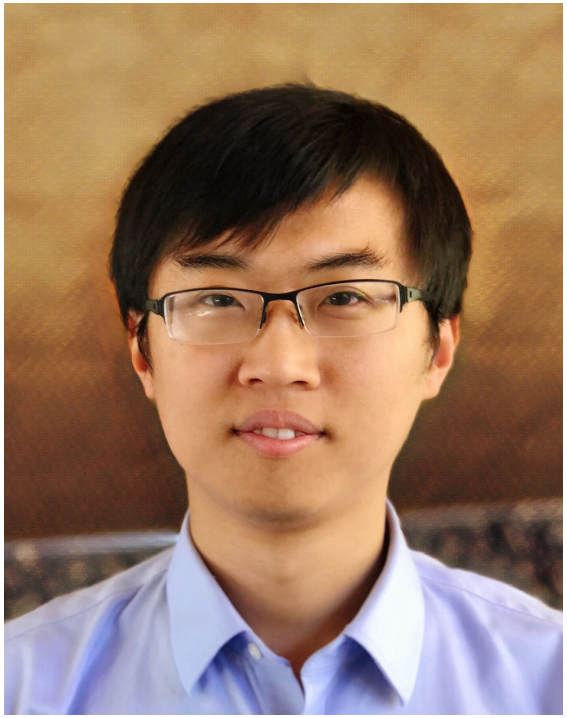
Hao Su

Course website

- https://ucsd-cse-dllab.github.io/FA20/index.html

Who are we?

Instructor: Hao Su



Teaching Assistant: Jiayuan Gu

Consultant Volunteer: Tongzhou Mu

Goal

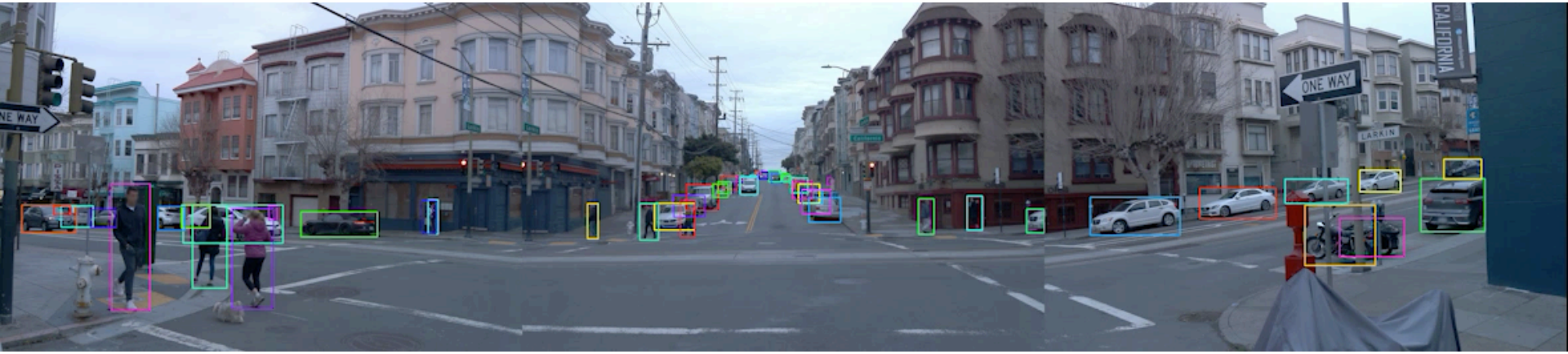
- To build strong machine learning engineers
  - who can solve real problems, but not who can get A+ on paper sheet
- I assume that you already know what DL is and what CV is
- Learning by doing! Get your hands dirty!
- We give you personalized feedback through project meetings

Method: Win A Game

- You will team up and work on a public vision challenge
- You goal is to climb to the top of the leaderboard
- Clear objective, clear metric, clear literature
- And you have to do things CORRECTLY

Waymo Open Dataset Challenge

- **Dataset overview**
  - 1,950 segments of 20s each, collected in diverse conditions
  - Sensor data
    - from lidar and camera
  - Labeled data
    - Labels for 4 object classes - Vehicles, Pedestrians, Cyclists, Signs
    - 12.6M 3D bounding box labels with tracking IDs on lidar data
    - 11.8M 2D bounding box labels with tracking IDs on camera data

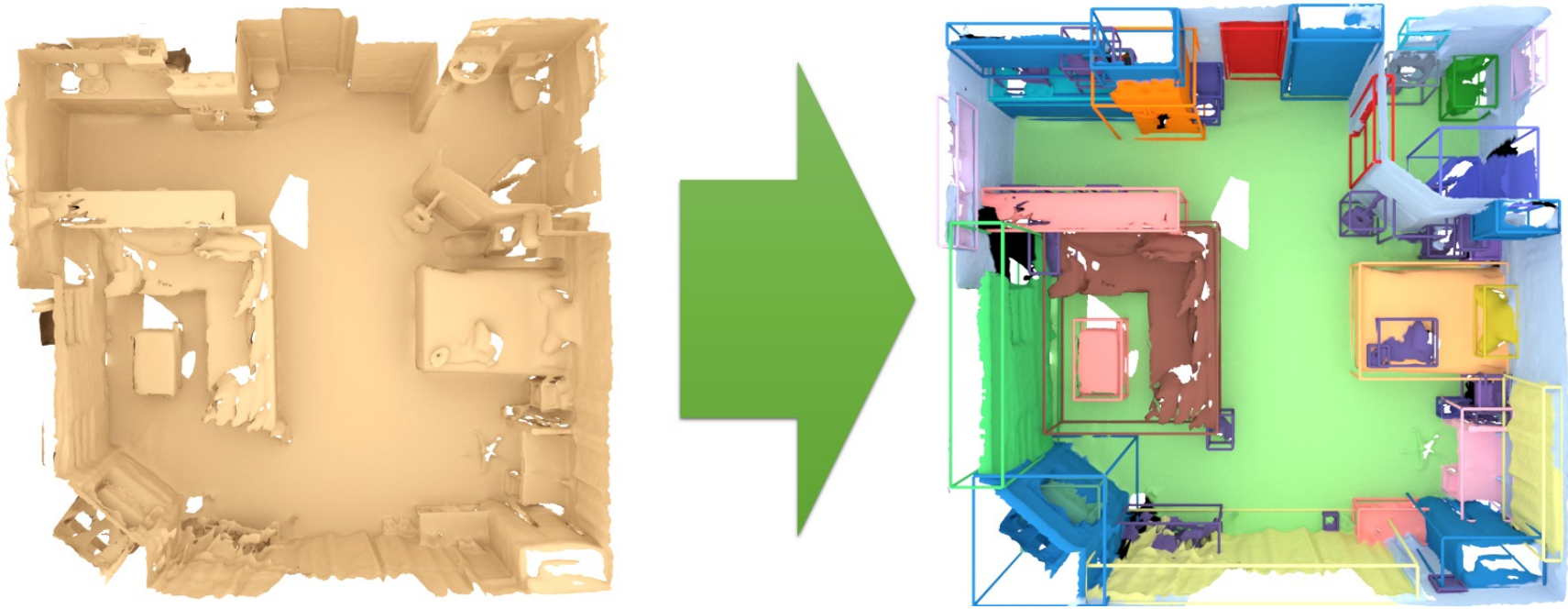


Waymo Open Dataset Challenge

- **3D Tracking**
  - Task
    - Given a temporal sequence (prior frames) of lidar and camera data, produce a set of 3D upright boxes and the correspondences between boxes across frames.
  - Metric
    - Multiple Object Tracking Accuracy (MOTA): Accounts for all object configuration errors made by the tracker, false positives, misses, mismatches, over all frames.

ScanNet Benchmark

- **Dataset overview**
  - Mesh
  - 1513 scans of indoor scenes
  - 3D data is provided with the RGB-D video sequences (depth-color aligned) as well as reconstructed meshes as .ply files.

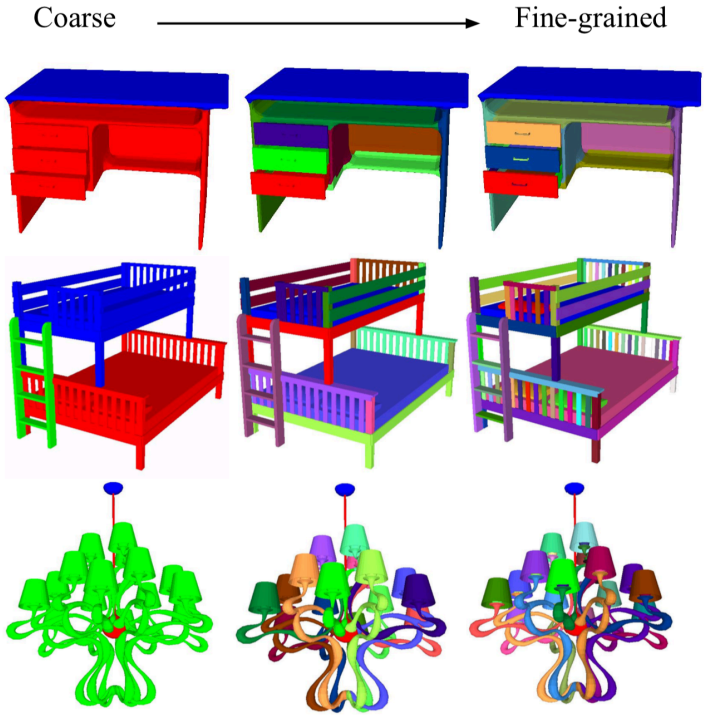


ScanNet Benchmark

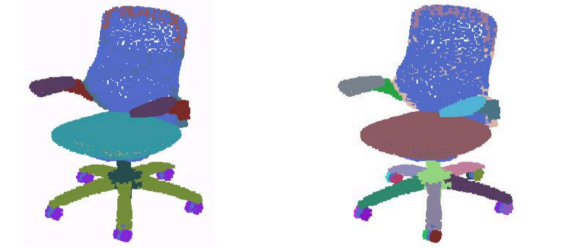
- **Instance Segmentation**
  - Task
    - Detecting and segmenting the object in an 3D scan mesh.
  - Metric
    - Average precision for each class.
    - Report the mean average precision AP at overlap 0.25 (AP 25%), overlap 0.5 (AP 50%), and over overlaps in the range [0.5:0.95:0.05] (AP).

PartNet

- **Dataset overview**
  - 3D objects annotated with fine-grained, instance-level, and hierarchical 3D part information
  - 573,585 part instances over 26,671 3D models covering 24 object categories

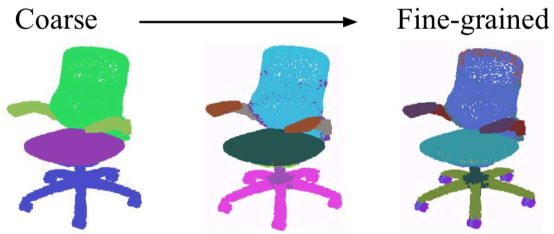


PartNet Dataset



Fine-grained Semantic Segmentation

Instance Segmentation



Hierarchical Semantic Segmentation

PartNet

- **Part Instance Segmentation**
  - Task
    - Detect every individual part instance and segment it out from the context of the shape
    - Input: a shape point cloud
    - Output: several disjoint masks over the entire point cloud, each of which corresponds to an individual part instance on the object
    - The detected masks should have no overlaps, but they together do not necessarily cover the entire point cloud, as some points may not belong to any part of interests
  - Metric
    - Per-category mean Average Precision (mAP)

LVIS

- **Dataset overview**
  - For long tail object recognition
  - ~2 million high-quality instance segmentation masks for over 1000 entry-level object categories in 164k images.
  - Due to the Zipfian distribution of categories in natural images, LVIS naturally has a long tail of categories with few training samples.

LVIS

- **Long-tail Object Detection**
  - Task
    - 2D object detection
    - A large number of categories and where per-category data is sometimes scarce
  - Metric
    - Overall average precision

Course Organization

- We take a flipped class format:
  - You learn by reading papers, writing codes, and discuss with teammates **offline**
  - Class hours are used for teacher-student discussions

- I expect that you are **highly motivated and collaborative**

Grading (tentative)

- Participation of discussions in course 40%
- Mid-term milestone report 30%
- Final report 30%

Schedule

- W0 (this lecture): class opening, ice-break, leave students papers to read
- W1: team forming due, professor's guidelines for doing projects, intro to computing resources, paper reading
- W2: presentation of comparative reading of important literature
- W3-W10: project meetings
  - Tu: Team 1 & Team 2
  - Th: Team 3 & Team 4
  - 40 min for each team. You have to show up in your 40 min. You can serve as an observer when other teams discuss.

How Do Teams Work?

- You will form teams of 4-5 (a self-intro session later)
- Each week there is a rolling leader (Project Manager), starting from W2 (this is W0)
- Duty of team leader:
  - Present on behalf of the team
  - Coordinate with team members
- Due date for teaming: Oct 6 (next Tue)
  - Decide team members and submit the rolling order
  - Decide the (tentative) project topic

Now, write a short bio of you (10 minutes)

Self introduction

Dues to Come

- Oct 2: bio paragraph (email TA)
- Oct 6: team construction
- Oct 13:
  - Comparative reading report of your topic (everyone has to write)
  - Presentation of comparative reading results (team leader's job)