Tracking Moving Object in Different Scenes

Cheng An Hsieh (chengan2)

Ben Chiang (benchian)

18752 Project

Content

- Problem statement (1 slide)
- Data collection and pre-processing (up to 2 slides)
- Feature extraction (3 slides)
  - Including Data visualization (1 slide)
- Regression/classification/time series prediction (4 slides)
  - One of these methods should meet the performance specification
- Picture of the software code (up to 2 slides)
  - Explain each section of the code
- Slides explaining methods that are unfamiliar to the class (up to 4 slides) because they were not taught in the lectures
- You can also include additional slides if they help explain the project better

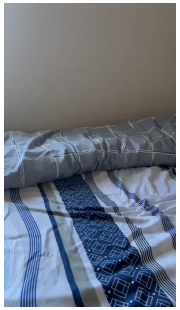
Problem Statement

- This project aims to solve the problem of tracking moving object in different scenes. Data will be in the form of a video captured by a stationary camera. We will hand label scene classes and object-of-interest's position. This project is particularly interesting and useful to solve because accurate tracking of an object-of-interest in wide variety of scenes is critical for many downstream applications, such as self-driving, path planning, and scene understanding where there might be some important objects that algorithms should attend to at all times. For instance, pedestrians in autonomous vehicles.

Data collection

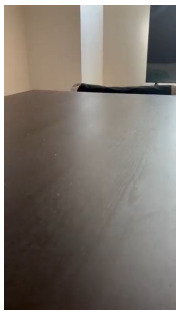
- Device:
  - Iphone 15
  - HD mode(1920x1080), 30 FPS
- 4 Scenes:
  - Living Room
  - Bedroom
  - Bathroom
  - Dining room
- Moving object:
  - CMU water bottle

Living Room



Bedroom

Bathroom



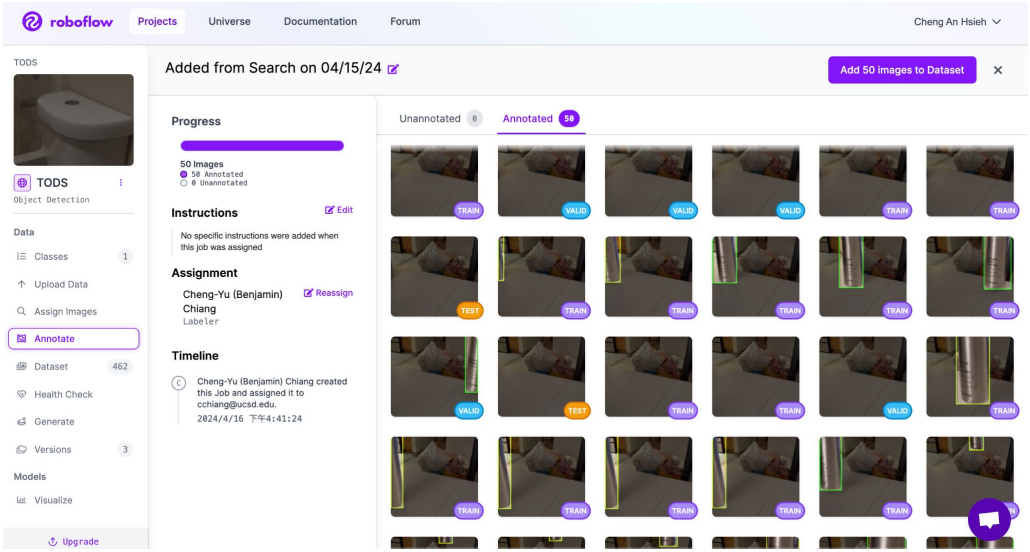
Dining Room

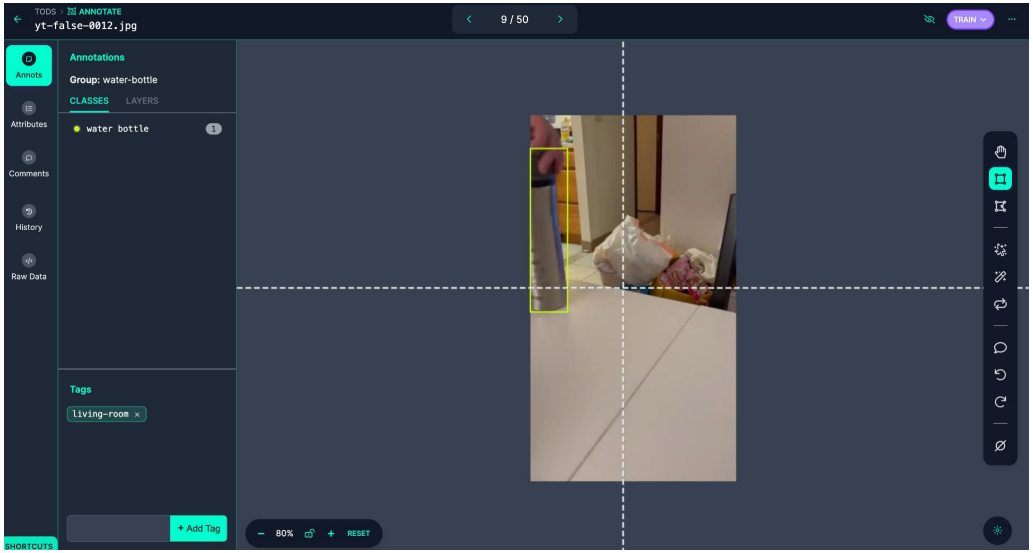
Image Samples of All Scenarios

Different scenes with and without water bottle

Data Pre-processing

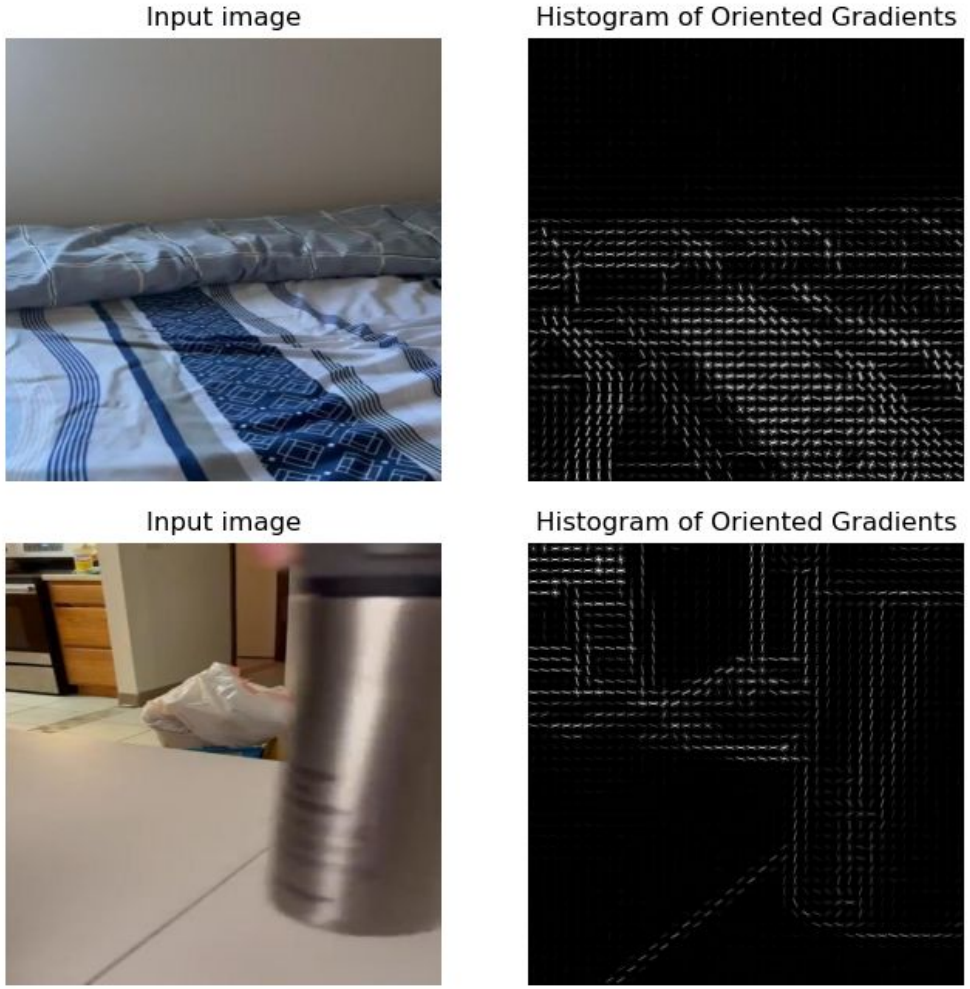
- Annotation:
  - human-labeling:
    - object bounding box
    - scene category
  - tool: Roboflow
- Preprocessing:
  - Resize to 640x640
- Dataset format:
  - Coco
- Data amount:
  - 4 videos(30s) in different scenes
  - 512 images in total





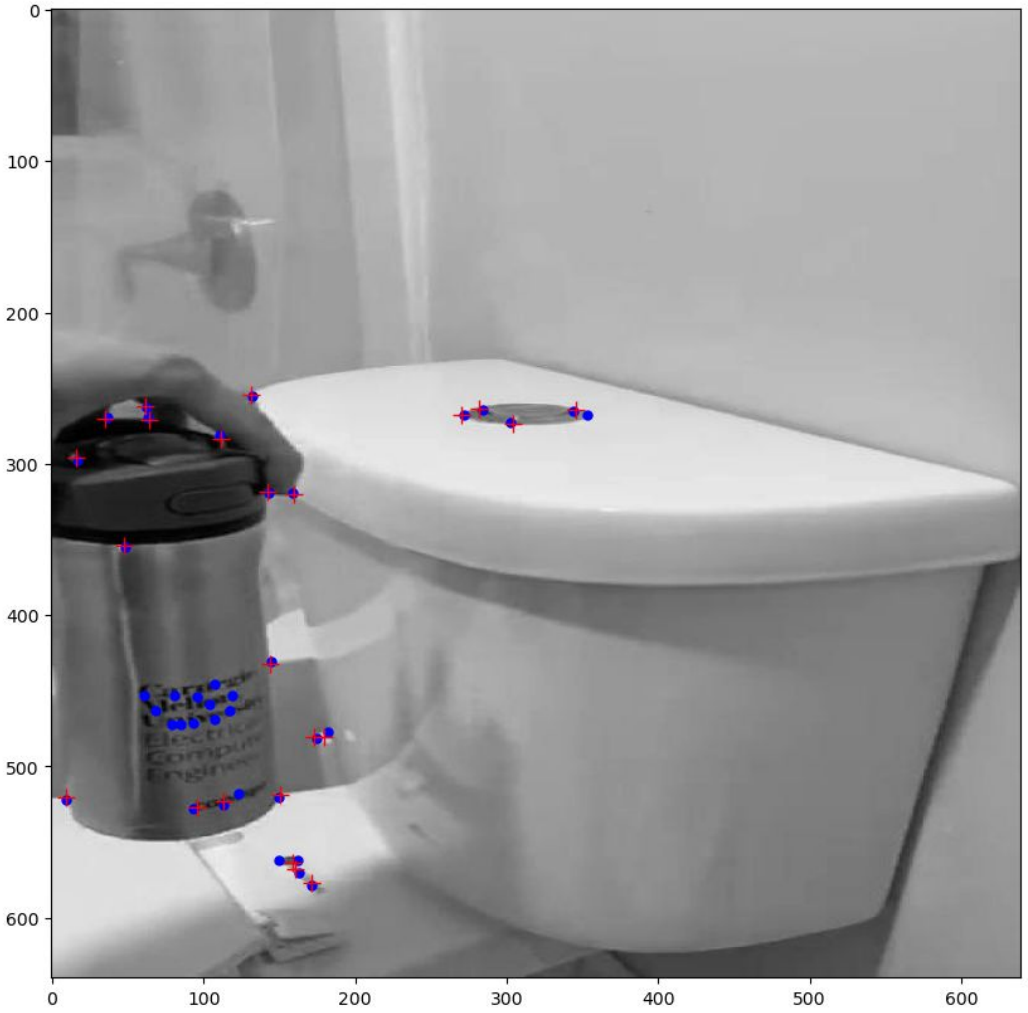
Feature extraction 1

- Histogram of Gradients(HOG)



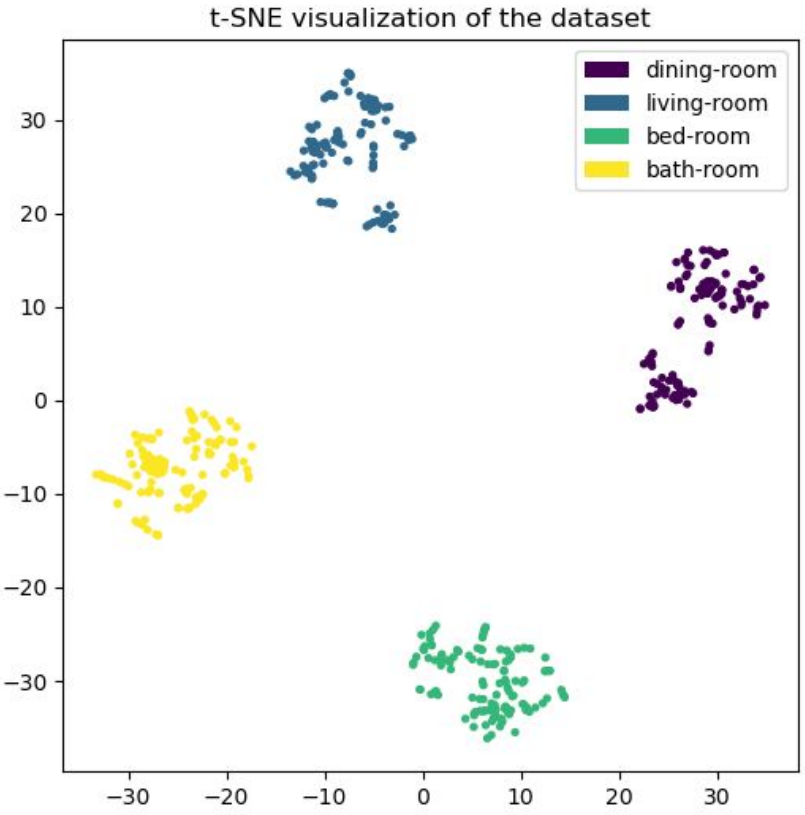
Feature extraction 2

- List of keypoints(corners)



Feature extraction

- t-SNE



Clear separation of different classes in embedding space!

Shows that the task of classifying different scenes is feasible