Multi-Object Tracking Challenge

CV3DST Lecture Exercises

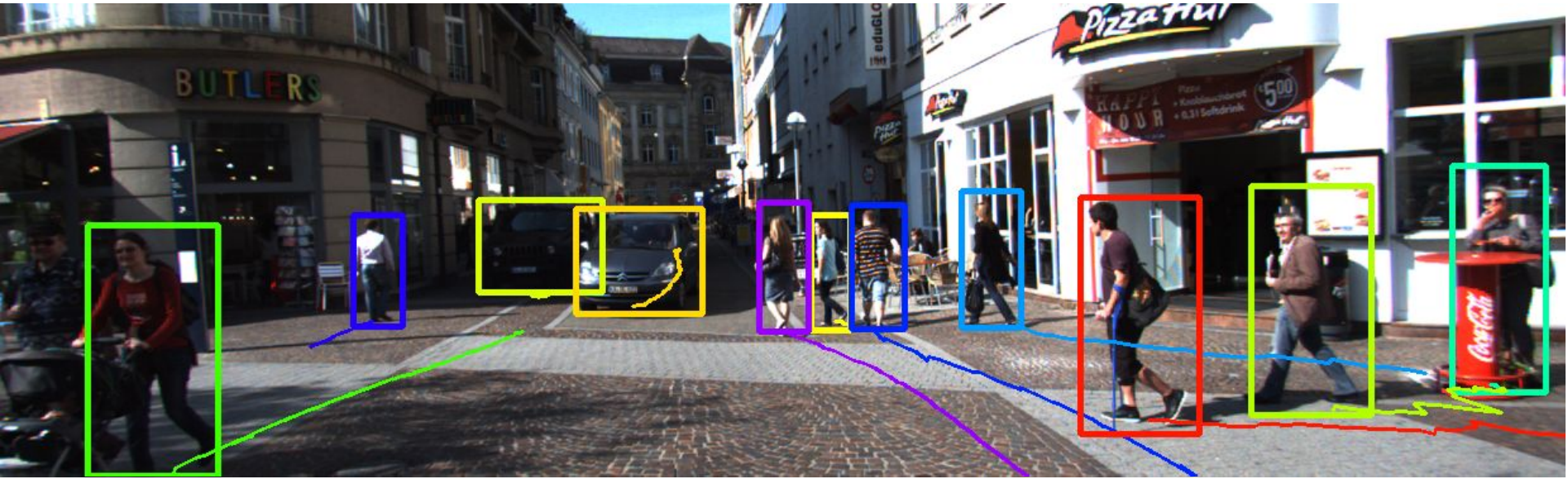
Multi-Object Tracking

Multi-Object Tracking

- Origins
  - SONAR, RADAR

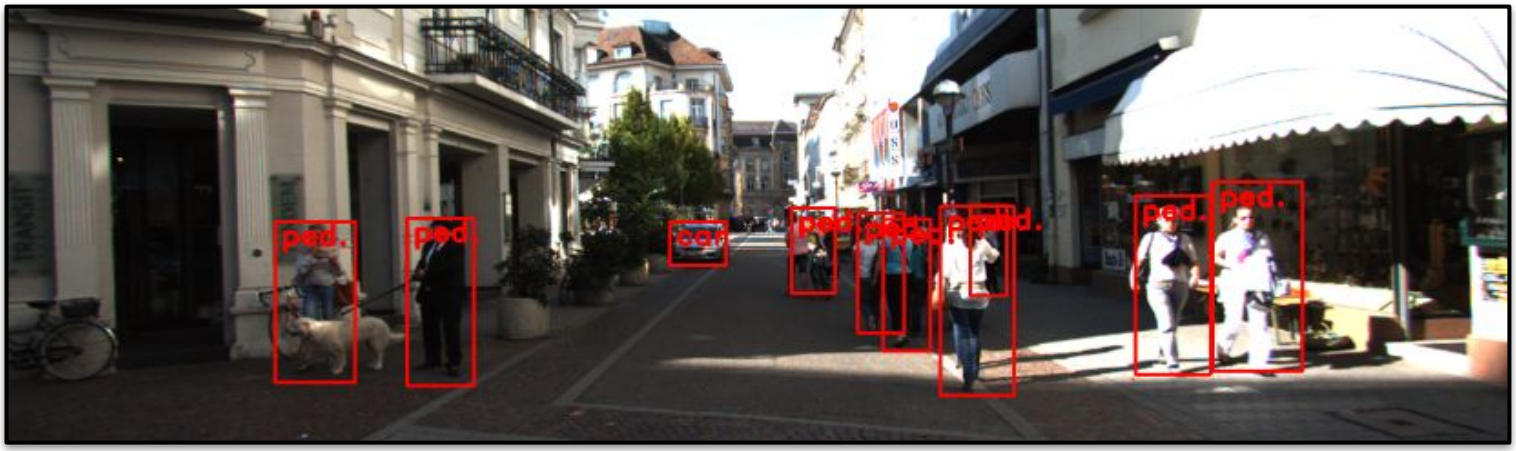
- Given a raw stream of sensory data:
  - Localize objects
  - Estimate object identities over time
  - Estimate when objects enter and leave sensing area

Vision-based Multi-Object Tracking



Vision-based Multi-Object Tracking

- Vision-based tracking
  - Sensor: camera
  - How to obtain the evidence for the presence of objects?
  - Tracking-by-detection



Challenge

Challenge

- **Given:** a baseline multi-object tracker
- **Task:** improve its tracking performance by applying different techniques from the lecture
- **Tracking-by-detection** paradigm
  - Apply object detector to each frame independently
  - Data association
- **The challenge:** connect the detections of the same object and produce identity preserving tracks

Dataset

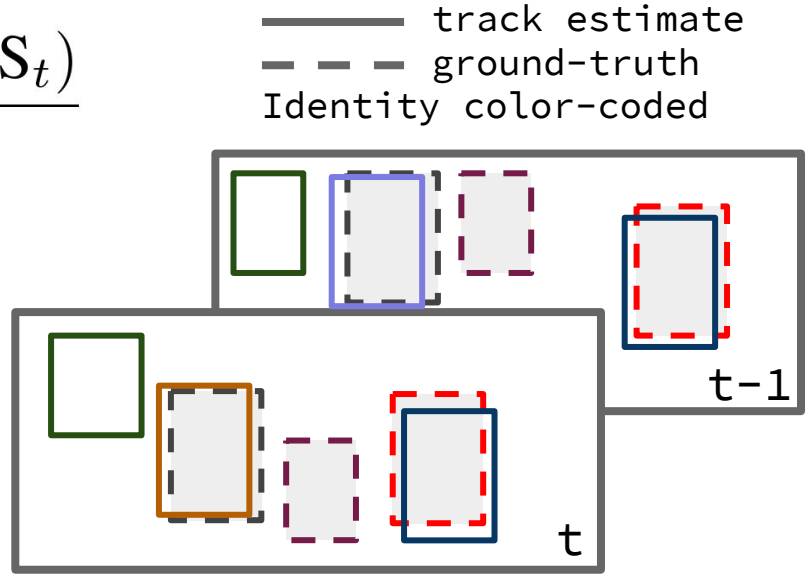
- MOTChallenge MOT16 dataset https://motchallenge.net/
- Define your own train/validation splits, on which you can validate your design decisions and hyper-parameters
- You will evaluate your final model on test sequences
- We will provide them at the end of the semester
  - You will not be given access to the ground-truth
  - You will upload your results to our evaluation server

Evaluation

- Multi-Object Tracking Accuracy and Precision

$$\mathrm{MOTA} = 1 - \frac{\sum_t (\mathrm{FP}_t + \mathrm{FN}_t + \mathrm{IDS}_t)}{\sum_t \mathrm{M}_t}$$

$$\mathrm{MOTP} = \frac{\sum_t \sum_{n,m} d_{tnm} a^*_{tnm}}{\sum_t |\,\mathrm{TP}_t\,|}$$



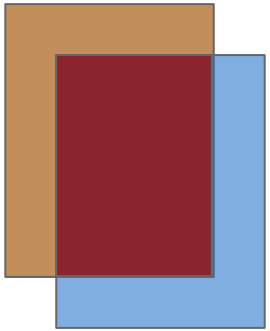
What Do We Provide?

- Google collab platform:
  - Dataset (MOT16 train split)
  - Object detector (Faster R-CNN, trained on our data)
  - Simple tracking baseline
  - Ground-truth tracks for supervision
  - Evaluation scripts
  - Instance segmentation masks for training

https://colab.research.google.com/drive/18uAKz1qMLvsr2h1w9tSk1zlMekhi-lUU

Baseline Tracker

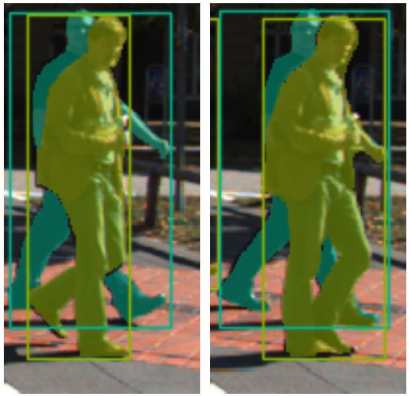
- Frame-by-frame detections (Faster R-CNN)
- **Association:** intersection-over-union (IoU)



- **Initialize** new tracks from non-associated detections
- **Remove** tracks that can not be extended with detections

Directions

- Object detection
  - Tracking performance depends on the detection quality
  - Detections provide signal for track initialization and termination
- Tracking
  - Assign correct identities to detected objects
  - Cope with occlusions, missing detections and false positives
- Leverage additional cues, e.g.,
  - Segmentation masks
  - Optical flow
  - Semantic segmentation



Rules and Timeline

Timeline

- Submission deadline: TBA

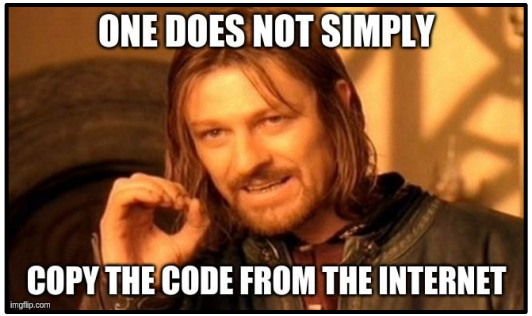
- Top 60% performers (based on MOTA) will get the bonus!

- Top K-performers will present their work in the week after the lectures (date: TBA, K: TBA)

Rules

- NOES
  - No teams!
  - Do not copy code from the internet!
  - You cannot use better of-the-shelf detectors!
  - You cannot use of-the-shelf trackers!

  Improvements on detection/tracking side you need to implement yourself. This is your individual work!
- YESES
  - Use any additional source of information:
    - Segmentation masks
    - Semantic segmentation, optical flow
    - … (see lectures!)



Feel free to use external code here.

THANK YOU FOR YOUR ATTENTION! HAVE FUN AND BE CREATIVE ;)