Video Surveillance for Road Traffic Monitoring

Module C6 - Team 4 - Week 4

Lucía Arribas
Laura Salort
Miren Samaniego
Júlia Bujosa

OVERVIEW

1.

OBJECT DETECTION

Faster R-CNN

2.

SIMPLE OBJECT TRACKING

Use SORT

2b.

IMPROVEMENT

Use DeepSORT

3.

BASE MULTI-CAMERA

Reprojection to GPS coordinates

3b.

IMPROVEMENT

+ Appearance based matching

OBJECT DETECTION - TRAINED MODEL

DATASET :

- Train sequences: S01 and S04
- Test sequences: S03

FRAMEWORK : Detectron2

- DefaultPredictor → Model Inference
- DefaultTrainer → Training
- COCOEvaluator → Evaluation

MODEL USED

- Backbone: ResNet-50 + FPN
- Architecture: Faster R-CNN

TRAINING HYPERPARAMETERS

- Number of classes: 1 (cars only)
- Batch size: 4
- Initial learning rate (LR): 0.00025
- Total iterations: 10,000
- LR decay:a at 6k and 8k iterations
- Image size: 512

OBJECT DETECTION - TRAINED MODEL RESULTS on test set

Metric used: COCOEvaluator

mAP: 86.33

Pretrained

Finetuned

OBJECT TRACKING - SIMPLE METHOD

Detectron2 detects cars

Filtering of detections using ROI

SORT assigns IDs and adjusts positions

MULTICAMERA TRACKING - Projection to GPS coordinates

1 Load initial information

- Load video start times
- Read the homography and distortion matrices for each camera.
- Load GT annotations and predictions from each camera.

2 Transform coordinates

- Align frames between cameras.
- For each pair of frames:
  - Undistort pixel coordinates in case of camera distortion.
  - Projecting pixel coordinates of the center of the bounding boxes to GPS coordinates using the homographies.

MULTICAMERA TRACKING - Projection to GPS coordinates

3 Matching between coordinates

- Calculate haversine distance between projected GPS points.
- Match detections between cameras using thresholded distance.

$$D(x,y) = 2\arcsin[\sqrt{\sin^2((x_{lat} - y_{lat})/2) + \cos(x_{lat})\cos(y_{lat})\ sin^2((x_{lon} - y_{lon})/2)}]$$

4 Association of IDs between cameras (Re-ID)

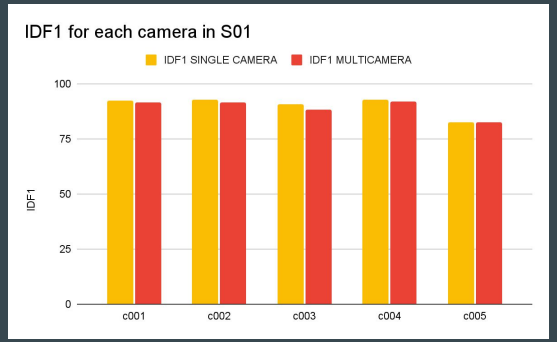
- Assign the same ID to matched objects between frames, and remove those that are not in at least 2 cameras.

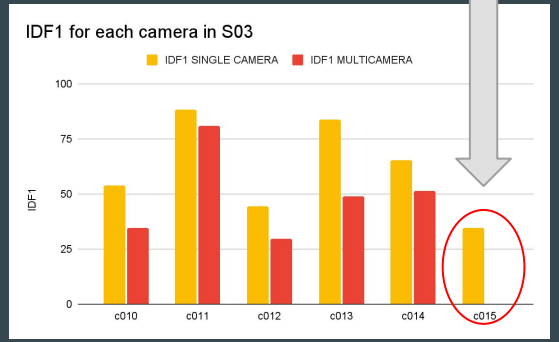
MULTICAMERA TRACKING - First results

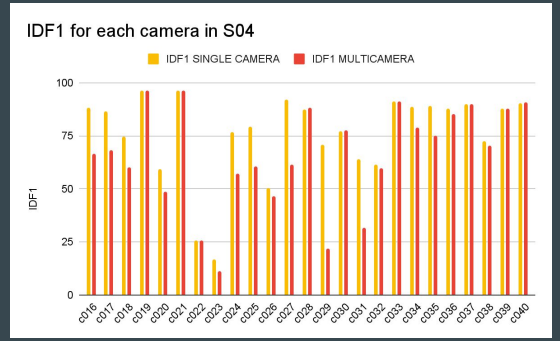
We obtain better results for S01 and S04, given that we trained the object detector on those sequences. We see that it does not generalize as well with sequence S03. Additionally, in general the IDF1 results are worse when using this multi-camera tracking approach.

Sequence	IDF1
S01	89,16
S03	40,90
S04	65,89

Video c015 has 8 fps!

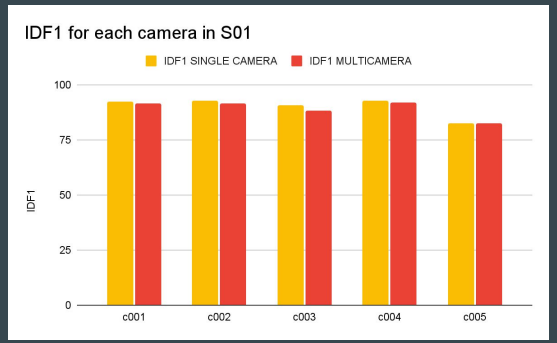


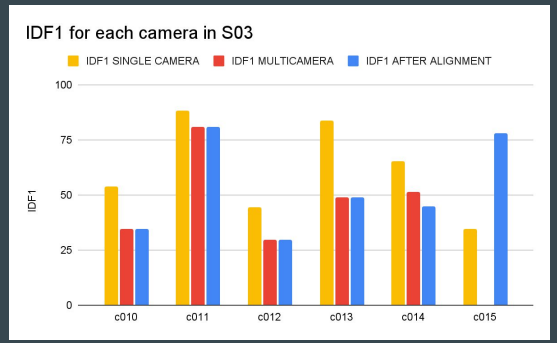


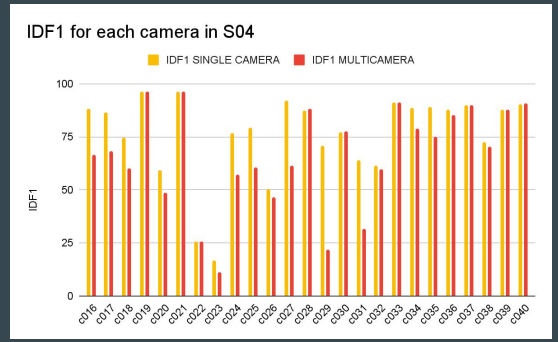


MULTICAMERA TRACKING - After correct alignment

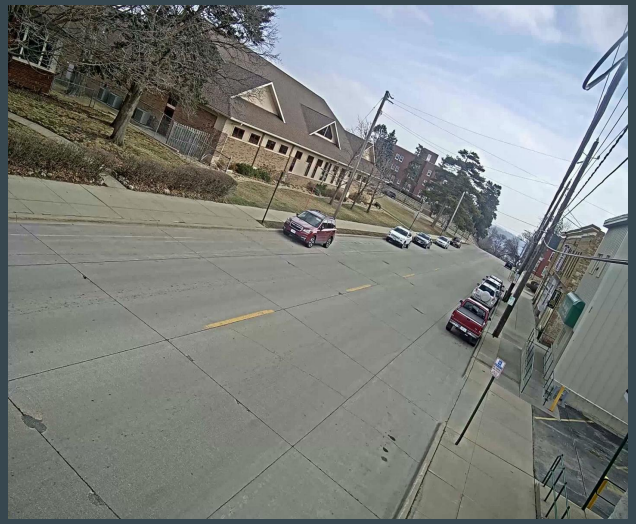
Sequence	IDF1
S01	89,16
S03	52,84
S04	65,89





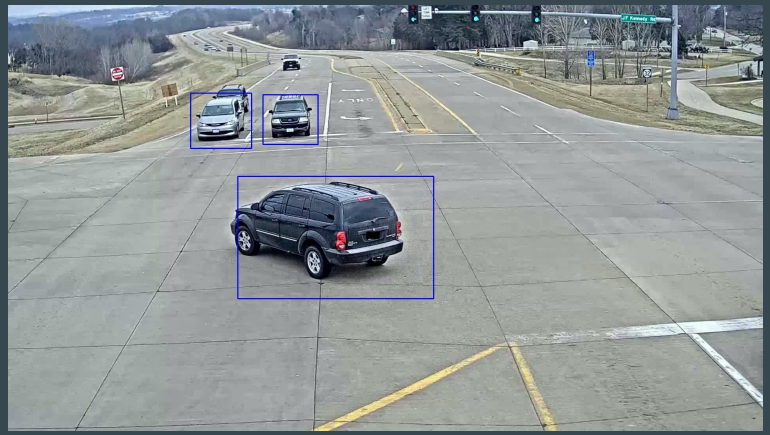


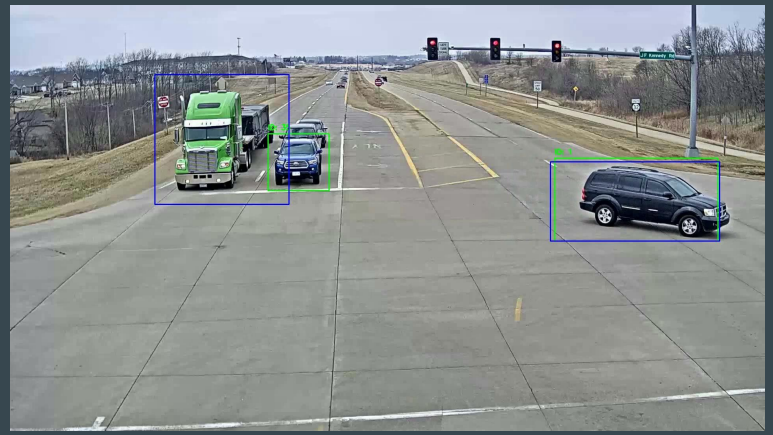
MULTICAMERA TRACKING - Qualitative results



We can see a clear failure case, the same car is not detected in either frames. We have observed that this is caused by the single camera tracking failing to detect these cars, and not the multicamera tracking.

MULTICAMERA TRACKING - Qualitative results

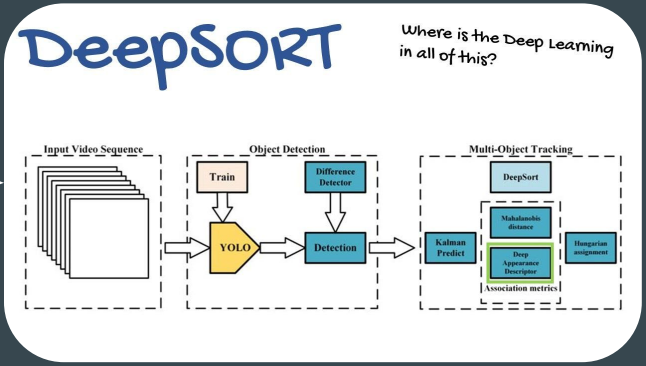




We see here another case where the tracking results are better, it assigns correctly the same ID to the same car in different cameras. Sometimes it struggles with certain car detections, but we believe is caused by the camera tracking done before multi-camera.

(SINGLE CAM) OBJECT TRACKING - TRYING TO IMPROVE

- **Poor results** obtained in multi-camera tracking (specially s03 and s04)
- Try to **improve** method from **lower layers** .



Custom embedder: FastReID

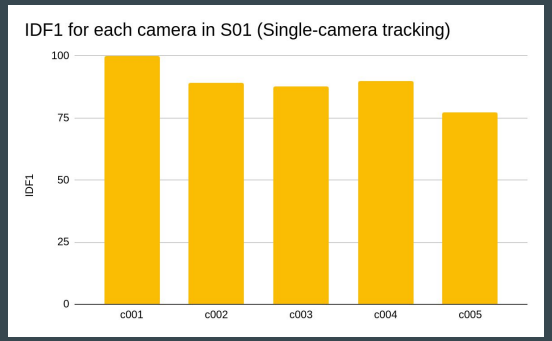
Model trained on **VehicleID** and **VeriWild** datasets, specialized in doing vehicle re-identification.

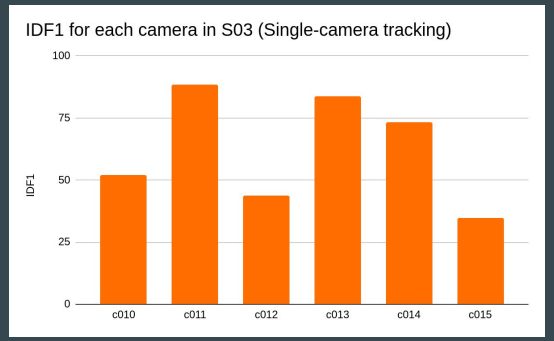
(SINGLE CAM) OBJECT TRACKING - TRYING TO IMPROVE

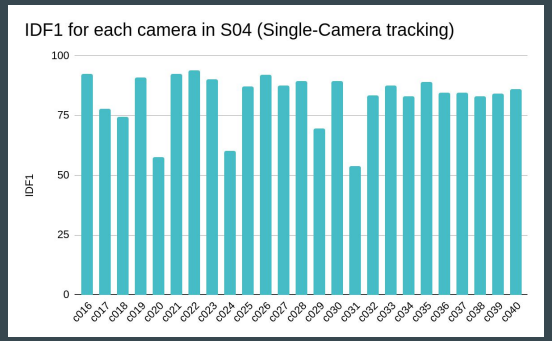
Detect objects → Filtering → DeepSORT tracking

Sequence	IDF1
S01	88.77
S03	62.61
S04	82.63

- Expected to work worse on S03 as object detector is trained in S01 and S04.
- It may obtain poorer results compared to using SORT, but in other cases is beneficial.







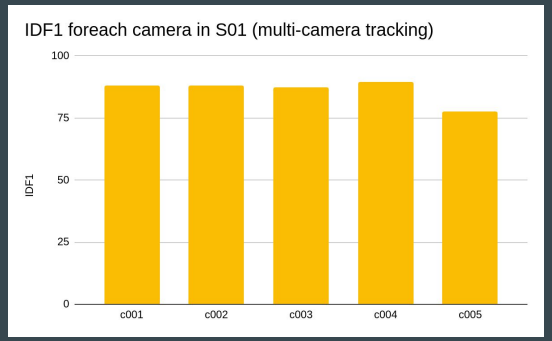
MULTI-CAMERA TRACKING AND DEEP SORT

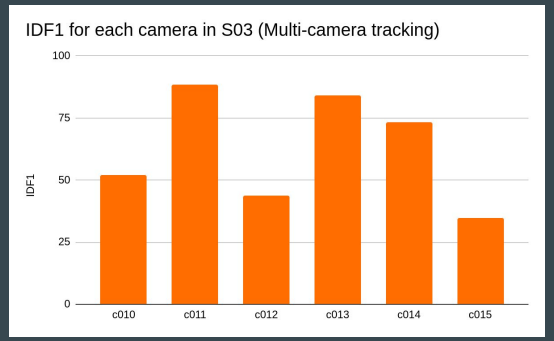
MULTI-CAMERA TRACKING AND DEEP SORT

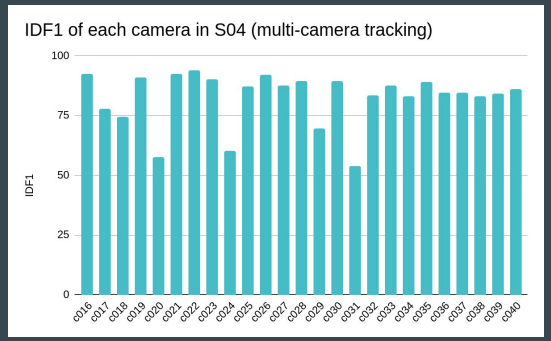
Detect objects → Filtering → DeepSORT tracking → Multi-camera

Sequence	IDF1
S01	**86.14**
S03	**62.68**
S04	82.63

- Camera handover complexity
- Consistency of the pipeline
- Challenges in S03 persist

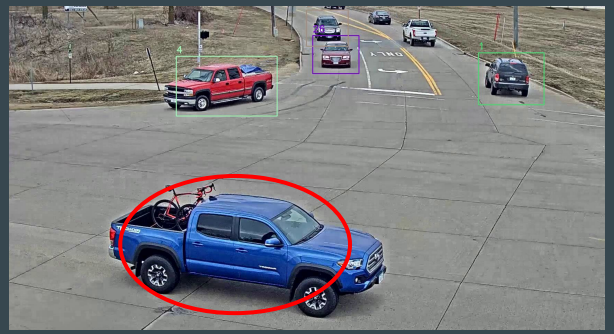


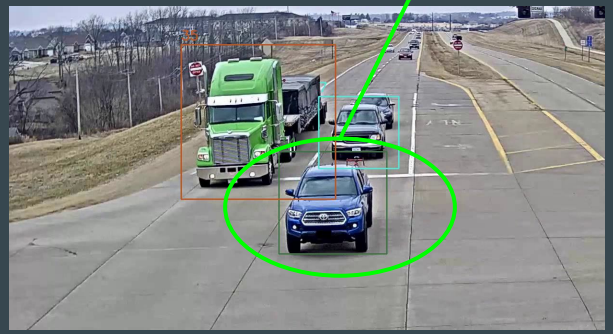




MULTI-CAMERA TRACKING AND DEEP SORT

Detect objects → Filtering → DeepSORT tracking → Multi-camera





MULTI-CAMERA TRACKING AND DEEP SORT

Detect objects → Filtering → DeepSORT tracking → Multi-camera

MULTI-CAMERA TRACKING AND APPEARANCE

Detect objects → Filtering → DeepSORT tracking → Multi-camera with appearence

1.Crop detected objects

↓

2. Deep Appearance Embeddings

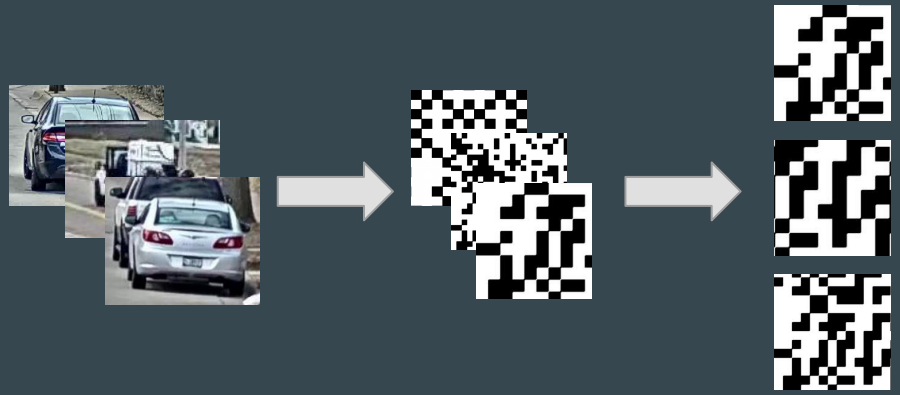
↓

3. Tracklet Level Averaging

↓

4. Multi-Camera association

FASTREID

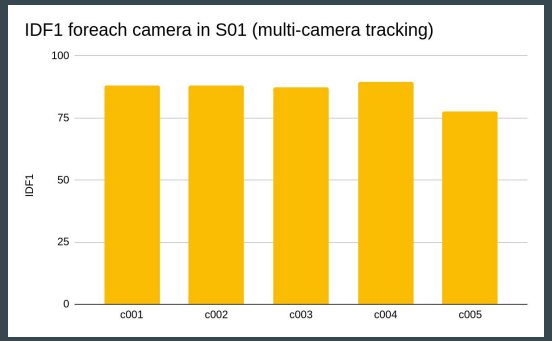


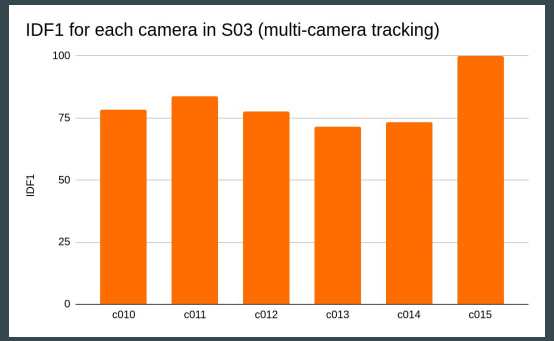
MULTI-CAMERA TRACKING AND APPEARANCE

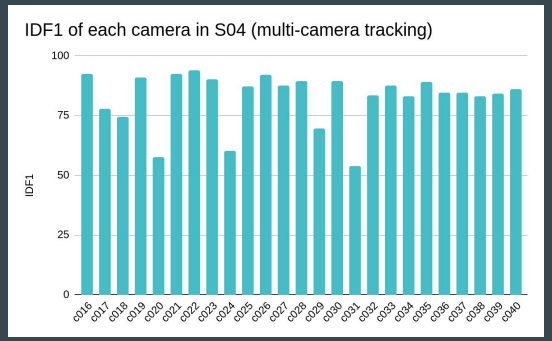
Detect objects → Filtering → DeepSORT tracking → Multi-camera with appearence

Sequence	IDF1
S01	86.14
S03	**80,76**
S04	82.63

- Added feature extractor helps distinguish individuals based on visual cues.
- Camera handover and cross-camera Re-ID
- No changes in S01 and S04. Why?

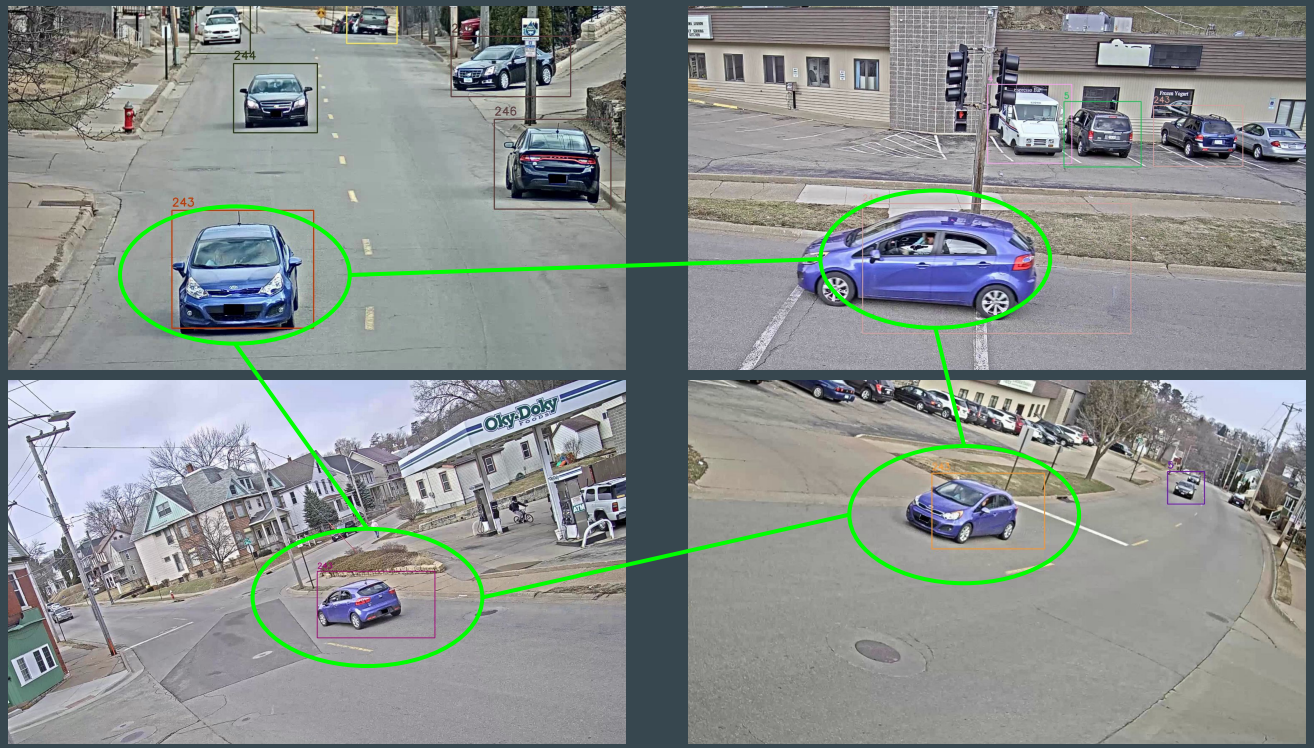






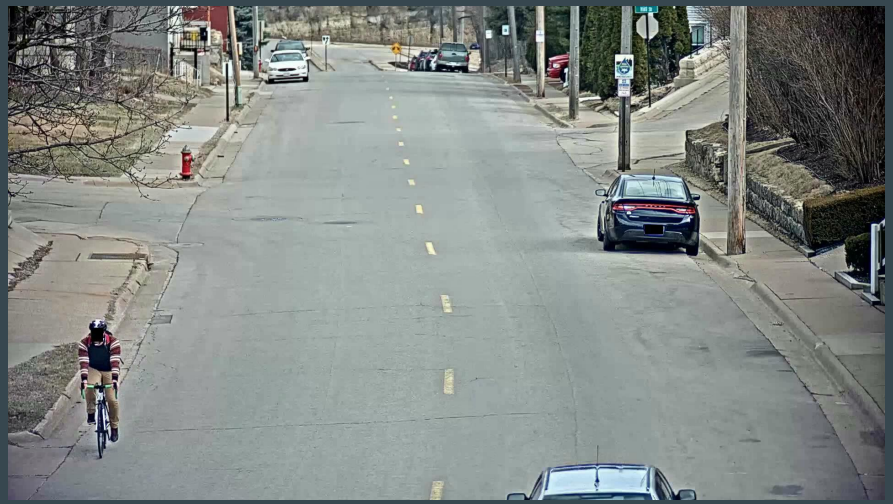
MULTI-CAMERA TRACKING AND APPEARANCE

Detect objects → Filtering → DeepSORT tracking → Multi-camera with appearence



MULTI-CAMERA TRACKING AND APPEARANCE

Detect objects → Filtering → DeepSORT tracking → Multi-camera with appearence



Strangely some parked cars are also detected, but they shouldn't because they do not appear in more than one camera. However we see that completely different parked cars are assigned the same ID in different cameras, possibly caused by projection errors and poor threshold selection.

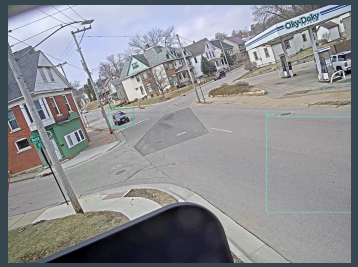
COMPARISON OF METHODS

Method	IDF1 S01	IDF1 S03	IDF1 SO4
Simple uni camera	39.4	28.84	19.76
Deep SORT uni camera	88.77	62.61	82.63
Simple multi camera	**89.16**	52.84	65.89
Deep SORT multi camera	86.14	62.68	82.63
Deep SORT multi camera + appearance	86.14	**80.76**	**82.63**

CHALLENGES, CONCLUSION & FUTURE IMPROVEMENTS

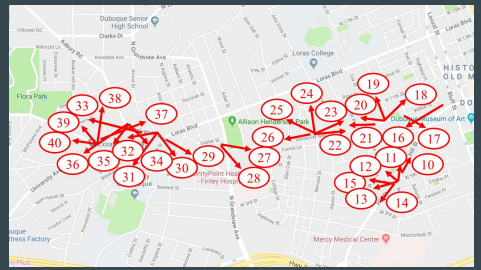
Challenges

- False detections
- Overfitting of the object detector to S01 and S04.
- Frame alignment is not perfect due to noise in video transmission
- Large differences in projected GPS coordinates due to poor calibration.



Conclusion

Although our model works well, we have not managed to make it as robust as we would like, mainly due to projection errors.

Improvements

- Find better threshold
- Improve appearance model
- For S04 consider only cameras pointing at the same intersections/streets