

TRACKING RECOVERY WITH RE-IDENTIFICATION

**Khanh Nguyen
Master of Computer Science
Semester Project
12 credits**

Supervisor: Taylor Mordan

Problems



OpenPifPaf assign IDs based on linking two continuous frames

⇒ Problem: A unique identity might be assigned different IDs across frames if it is not continuously detected

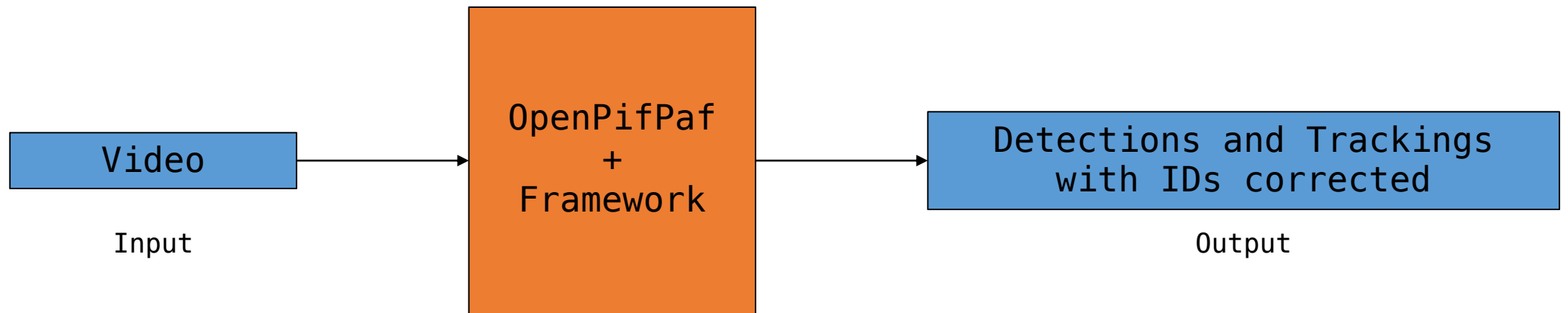
Potential reasons:

- Occlusions
- Objects going out-of-view
- Missed tracking

The 2nd girl got a new ID (2 → 3) since she is not detected in a certain frame

What we want to improve

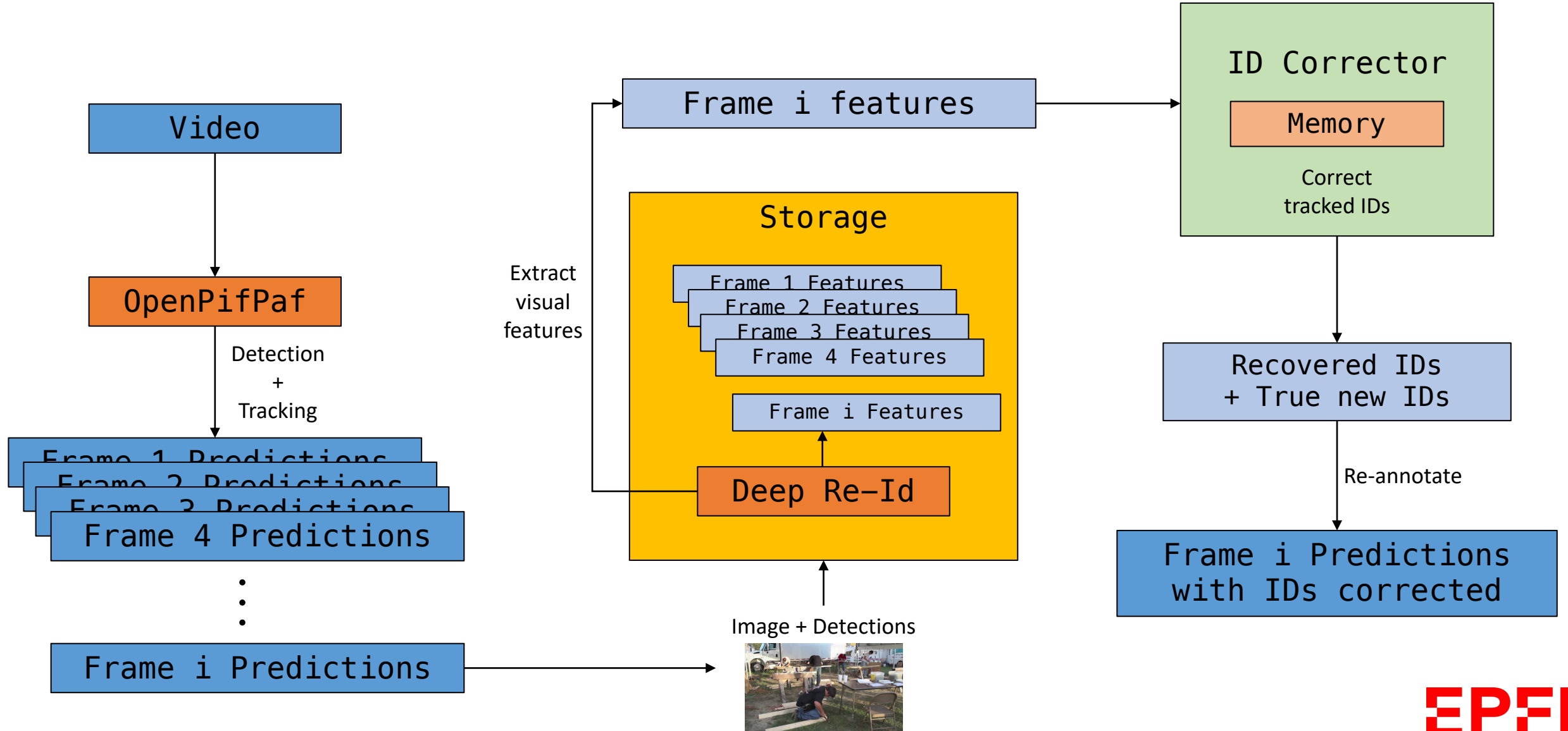
We want to improve OpenPifPaf to increase the performance of the assignments of tracking ID that would assign a single ID to a unique identity across different frames



Ideally, we would have the same input/output pipeline of OpenPifPaf and only change the IDs assigned to object tracked

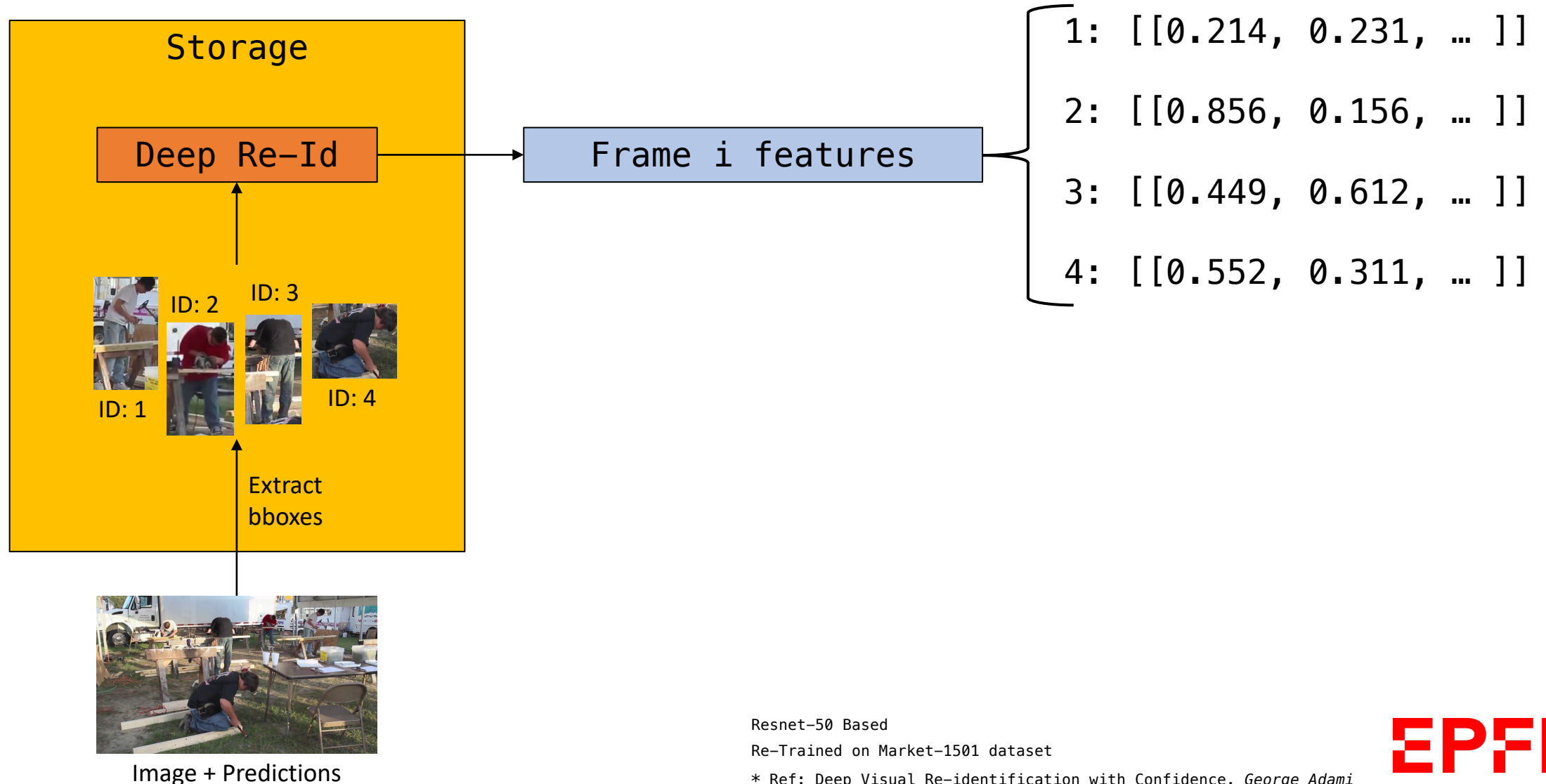
Method

Online Identity Recovery



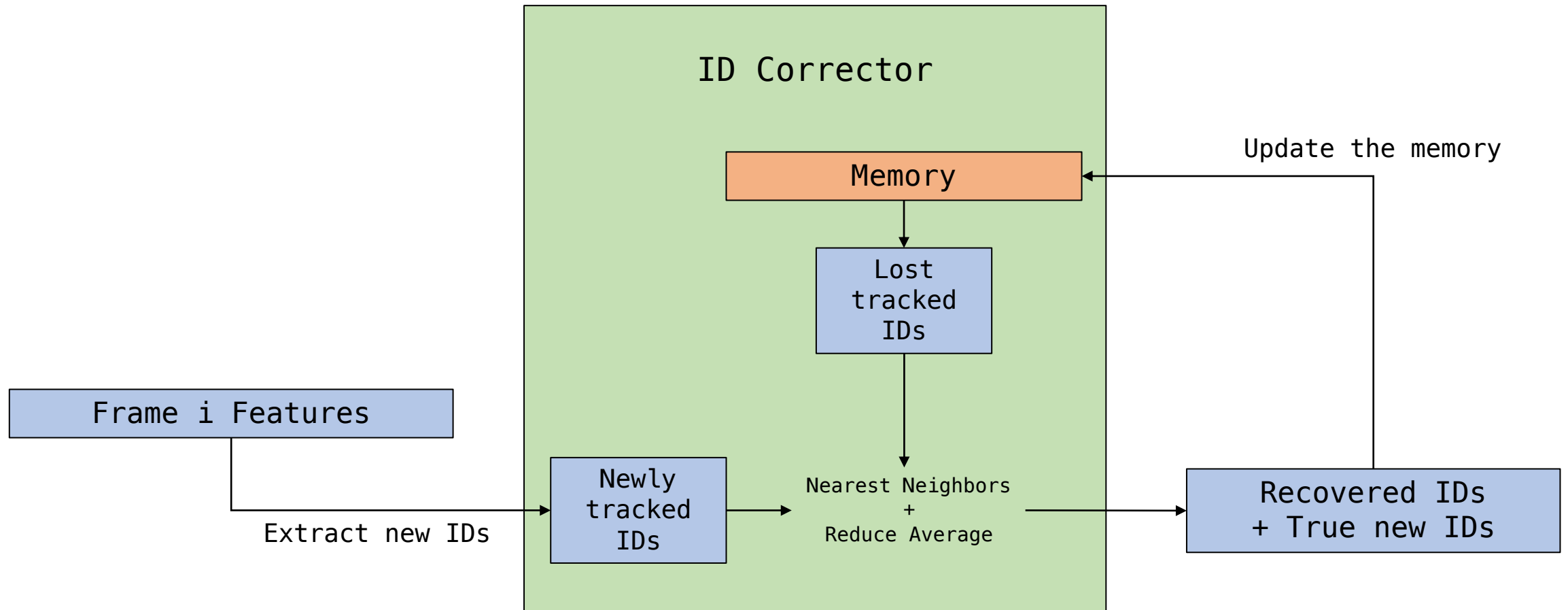
Deep Re-ID Model

Extract features vectors from visual clues



ID Tracker

Keep hold of memory and matched old IDs with newly tracked IDs
Can be customized with hyperparameters – memory length, no. candidates
Different memory mode: 'recent', 'sparse', 'first'



Reduce Average?

Query \rightarrow [*ID 1: 0.1*, ID 2: 0.2, ID 2: 0.3, ID 3: 0.4, ID 1: 0.5]



Query \rightarrow [*ID 2: 0.25*, ID 1: 0.3, ID 3: 0.4]

Why?

- *Intuitive: Make the retrieval more robust*
- *Experimental: Improve accuracy of matching IDs to the ground truth*

Evaluation Metric

Evaluation metric: Multiple Object Tracking Accuracy (MOTA)
Dataset: Posetrack2018

$$MOTA = 1 - \frac{\sum_t (m_t + fp_t + mme_t)}{\sum_t g_t}$$

Objects misses (m): Objects in ground truth but not in detected

False positives (fp): Objects detected and not in ground truth

Mismatch errors (mme): the number of times there is a switch in id-mappings between the objects in ground truth and tracking

We only evaluate MME

- MME is dominated by objects misses and false positives => overall MOTA improvement is minimal (maximum 1% from 65.1 -> 66.1)
- MME is directly related to the problem statement

Evaluation metrics

Posetrack MME is not enough

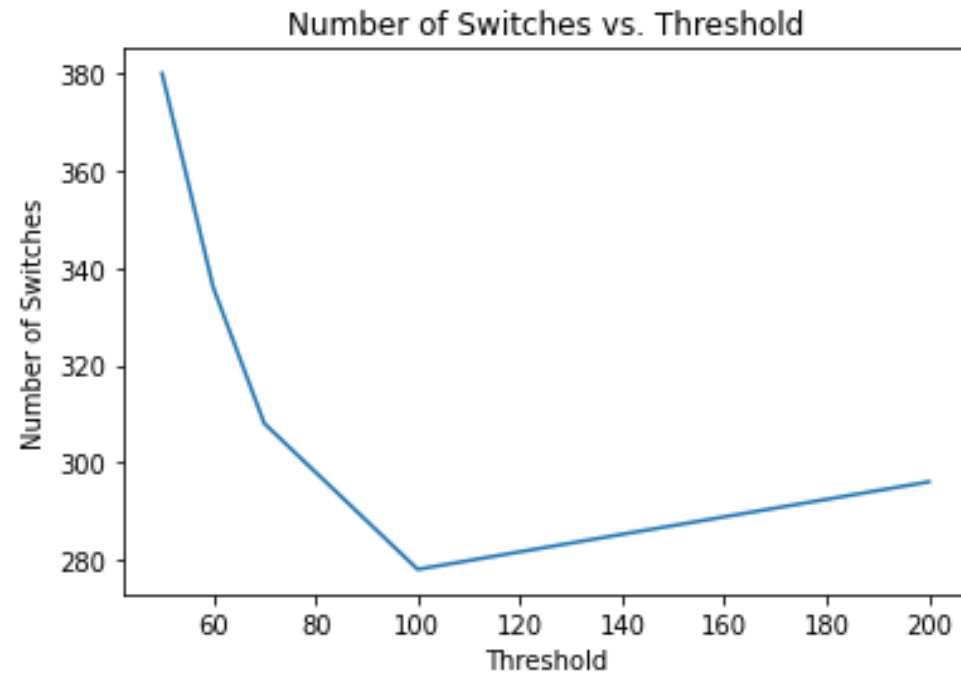
- Posetrack MME evaluate the switching on IDs of **Joints** (not **Person**), which might be relevant since the matching of joints are based on distance, not based on ground truth person identity
- MME does not consider a new person got assigned to the ID of someone who already existed

We also evaluate Person MME & accuracy

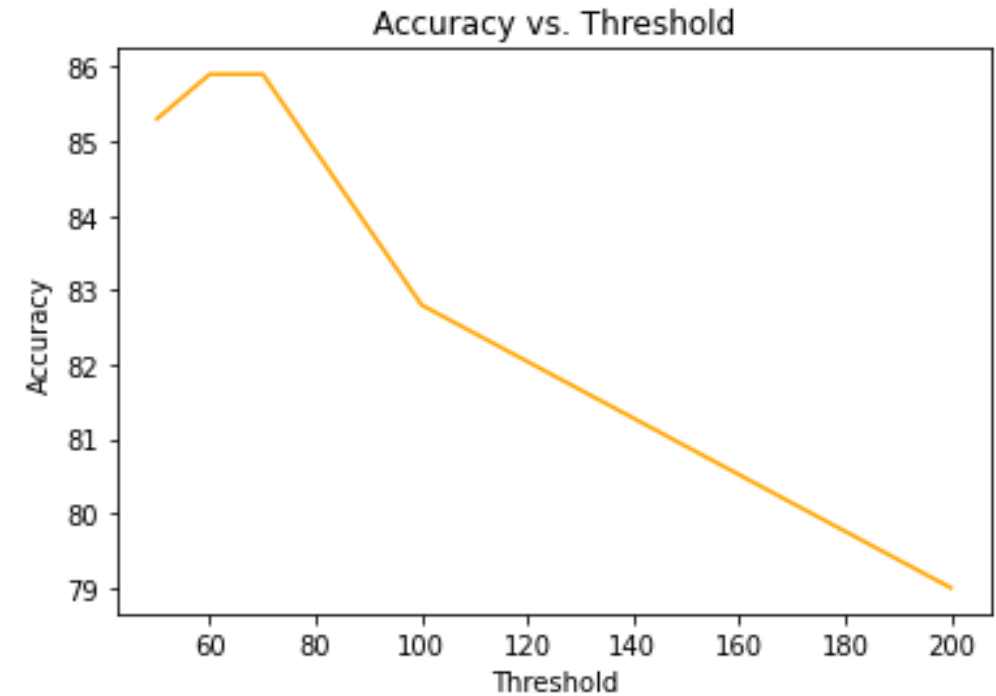
- ⇒ True Accuracy evaluate the assigned ID to the original ID (first match)
- ⇒ Recovered Accuracy evaluate the assigned ID against any IDs has been assigned in the past

Evaluation metrics

Less number of IDs switches ...



... does not mean a better IDs accuracy



Results of the best method

	OpenPifPaf Trackings	Modified Trackings
Joints MME	5420	3658
Person MME	498	310
True Accuracy	83.9	86.0
Recovered Accuracy	—	93.3
Posetrack MOTA	65.069	65.415

=> ~40% reduction in identity switches

Sample results

(Sample are picked from a pool that have at least one id switches removed)

Color Code

Green: prediction ID matching groundtruth ID

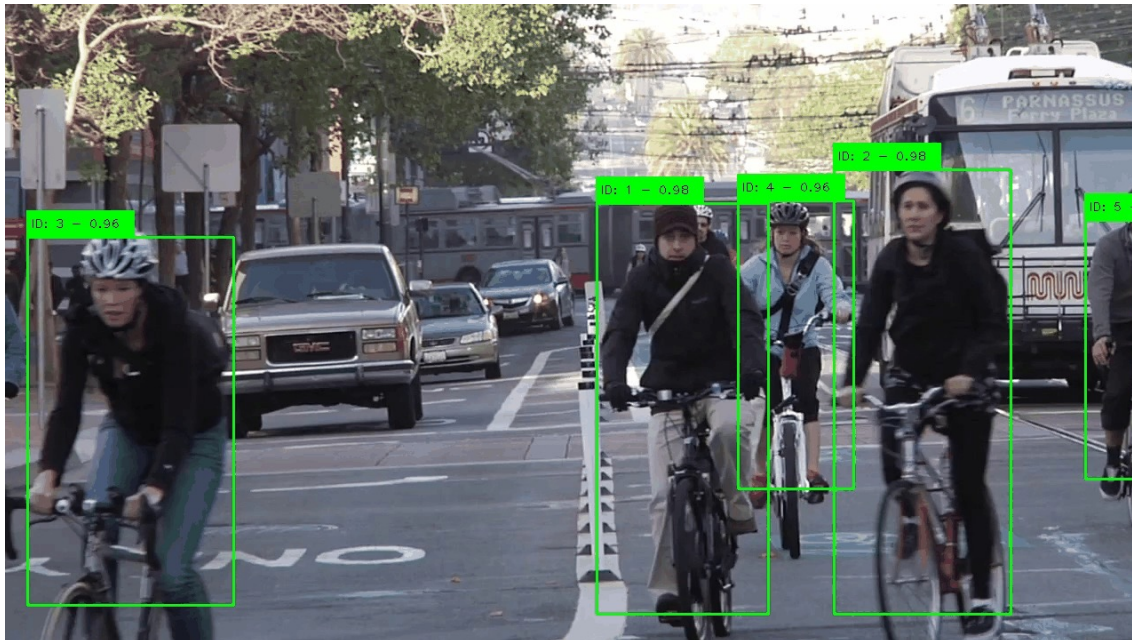
Red: prediction ID not matching groundtruth ID

Dark Green: ID that has been recovered and match groundtruth ID

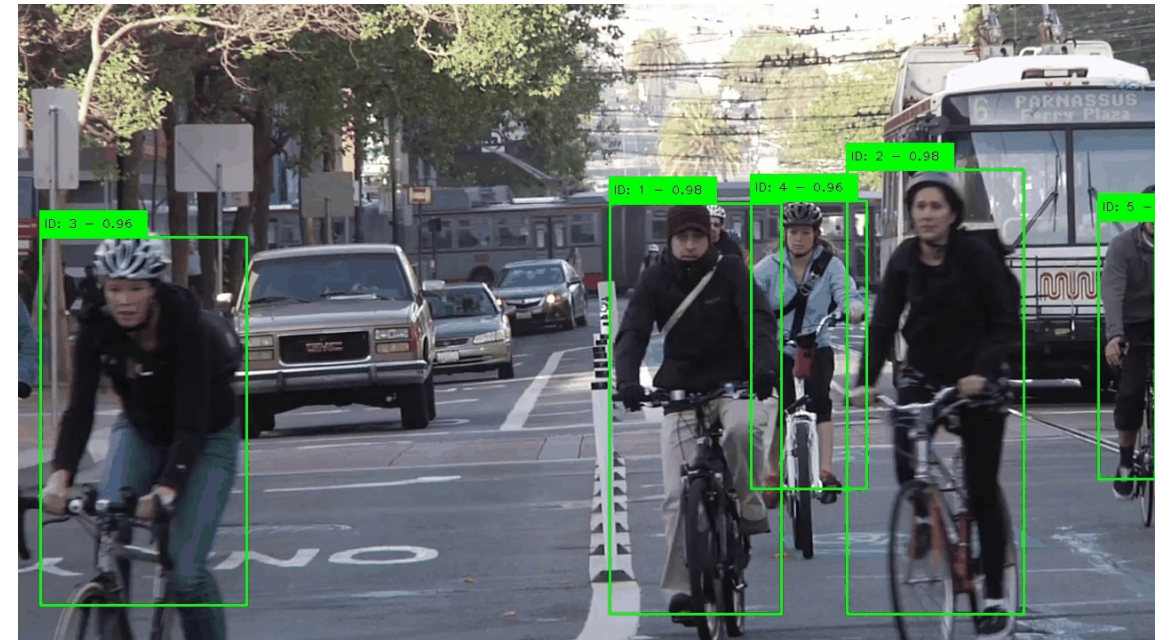
Orange: ID that has been recovered but does not match groundtruth ID

Grey: False Positive

Before



After



Sample results

(Sample are picked from a pool that have at least one id switches removed)

Color Code

Green: prediction ID matching groundtruth ID

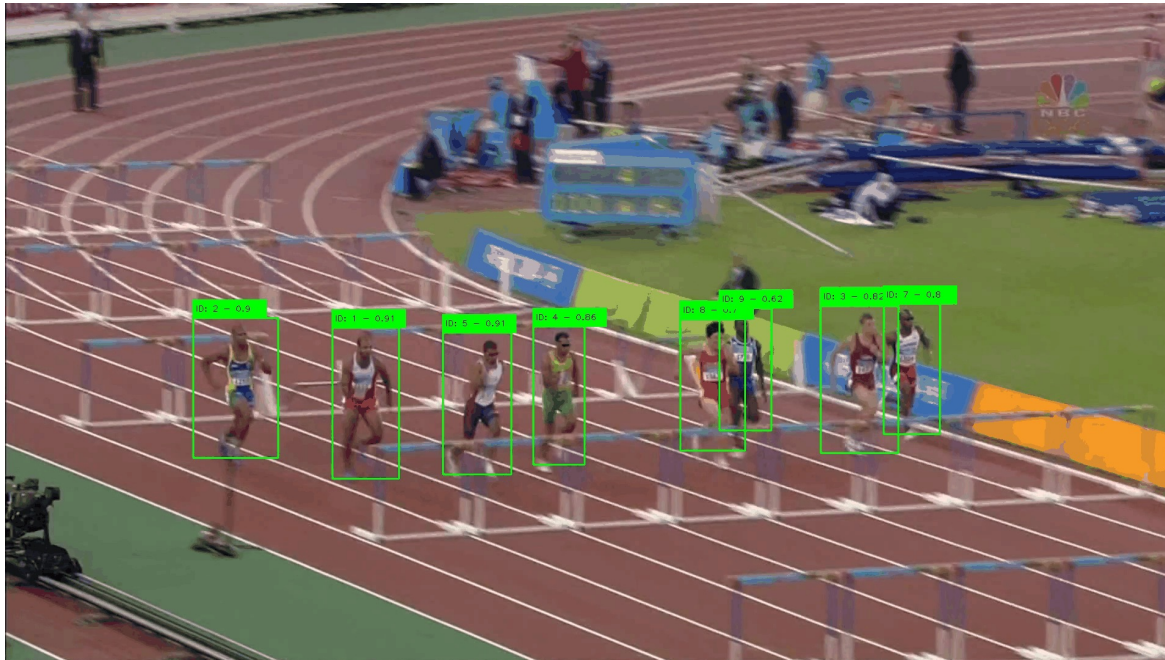
Red: prediction ID not matching groundtruth ID

Dark Green: ID that has been recovered and match groundtruth ID

Orange: ID that has been recovered but does not match groundtruth ID

Grey: False Positive

Before



After



Integrated with OpenPifPaf
... and can be utilized in other detection systems

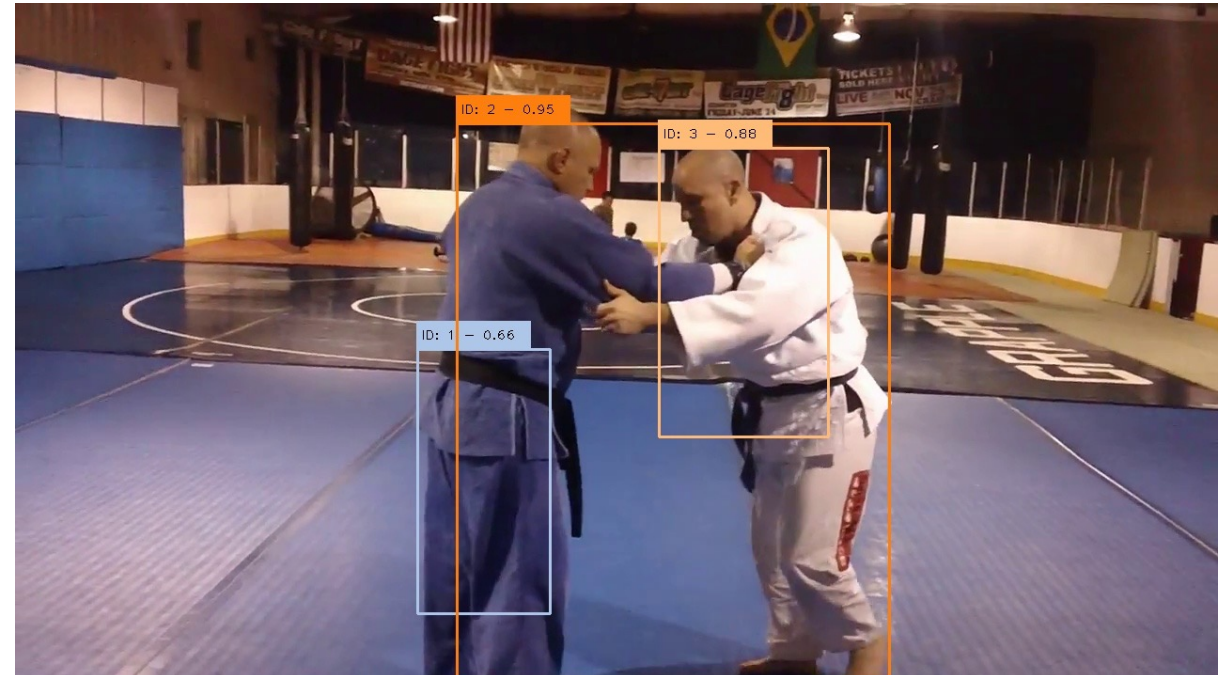
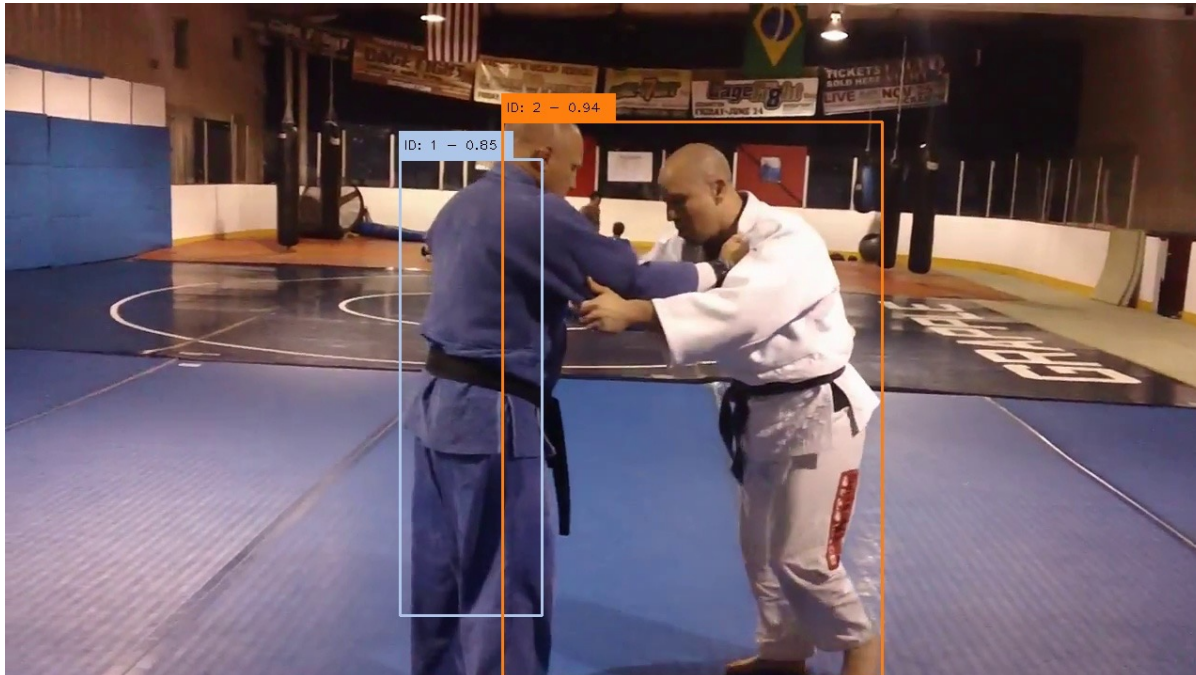


Original video: <https://www.pexels.com/video/men-playing-tennis-at-daylight-992695/>

Known limitations

Vulnerable to inaccurate detections:

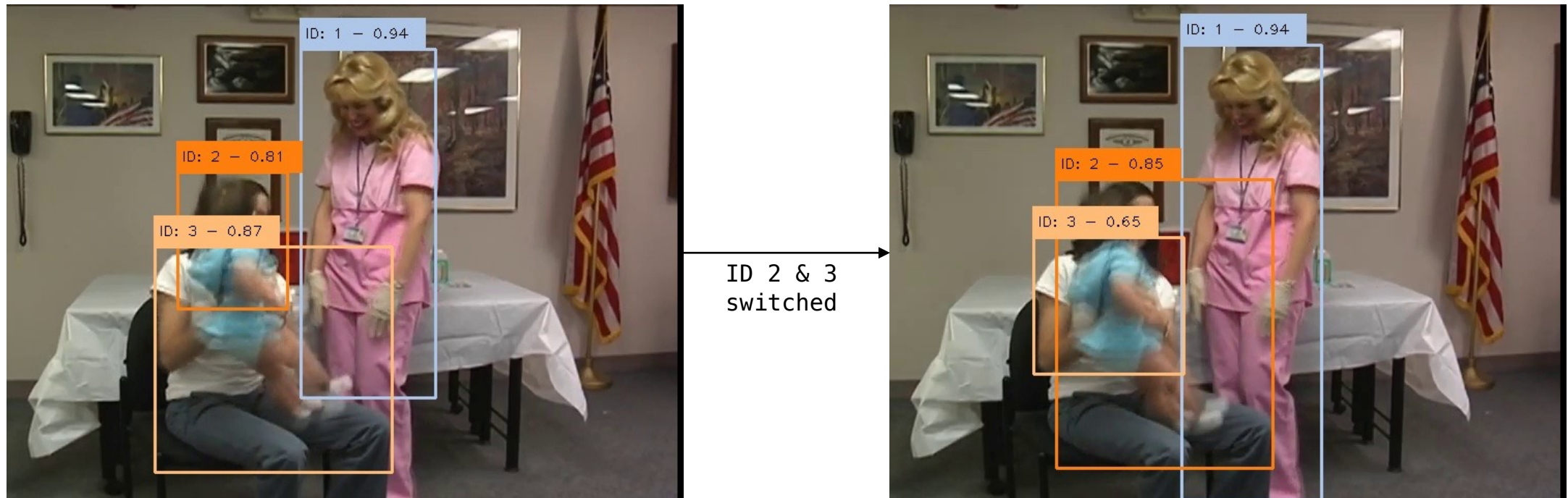
- Since the memory only consists of IDs that does not exist in the current frame
- Memory can contain existing IDs, but needs a lot of optimization to make it robust



Known limitations

Did not handle cross-persons ID switches:

- The query only contains new tracking IDs
- Query can extend to existing IDs, but also needs optimization to be robust
- This kind of unexpected swapping also impact the short-term memory structure

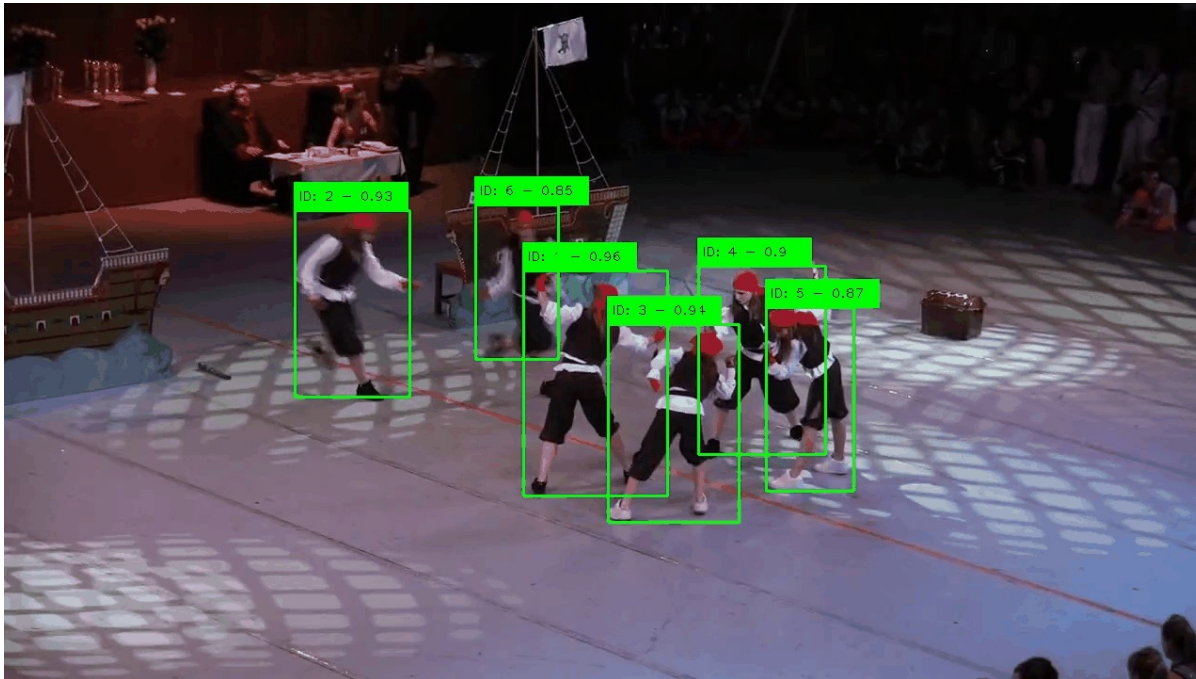


Known limitations

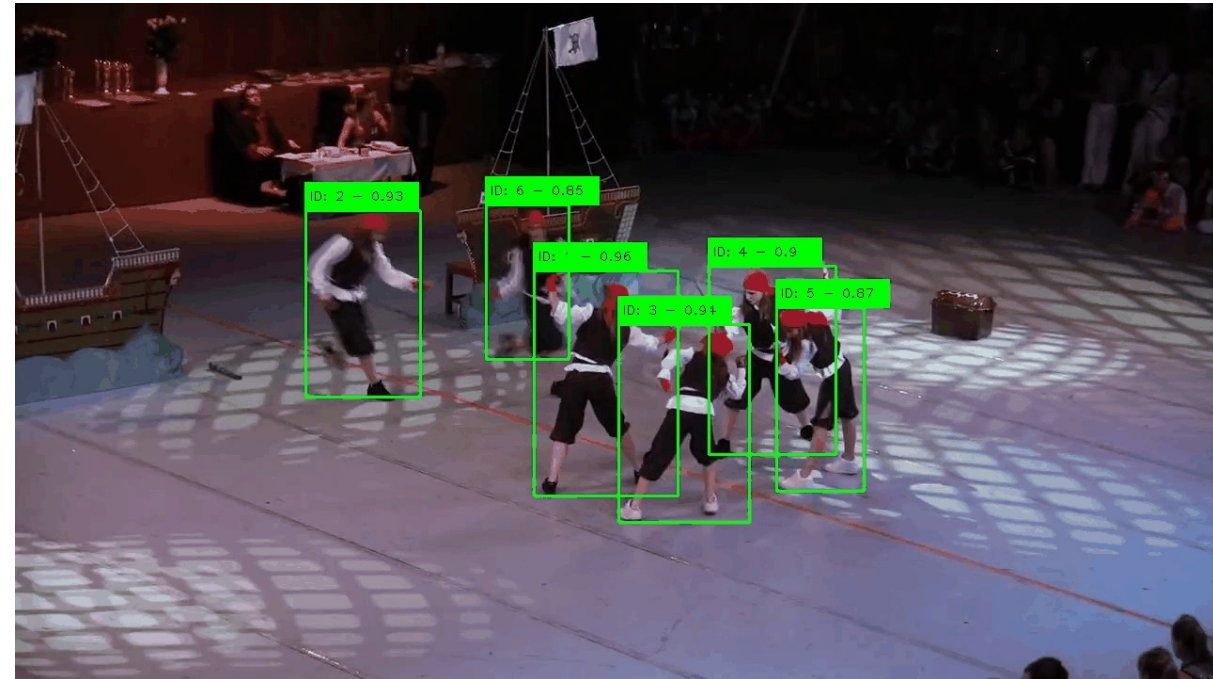
Reliance on visual cues (deep-reid model):

- Cannot handle well inputs with visually similar identities

Before



After



What have been done since midterm

- Developed the Storage system to extract features (and metadata) from annotation in an online way
- Developed the ID Corrector
- Experimentation and extension to the naïve ID Corrector
- Extended Posetrack poseval to consider Persons MOTA and Accuracy
- Developed as a standalone package and integrated into OpenPifPaf
- Visualizing tools

Thank you!