443 Final Presentation

Andrew, Lyla, Trina

Introduction to Multi-Object Tracking (MOT)

Multi-Object Tracking (MOT) involves identifying specific objects across multiple frames of a video

- Consistent classification across entire video
- Re-ID (re-identification) is performed as objects exit and re-enter the field of view
- We are using a single camera tracking

Dataset

- Our dataset consists of animated videos depicting people walking through a room.

- Videos serve as a baseline(training) for our Multi-Object Tracking task.

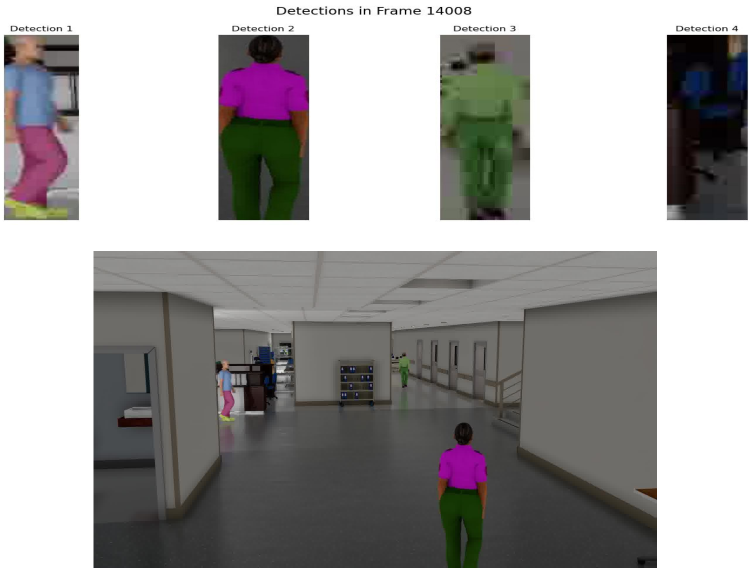
- The dataset allows us to explore and evaluate different tracking algorithms and techniques in a controlled environment.

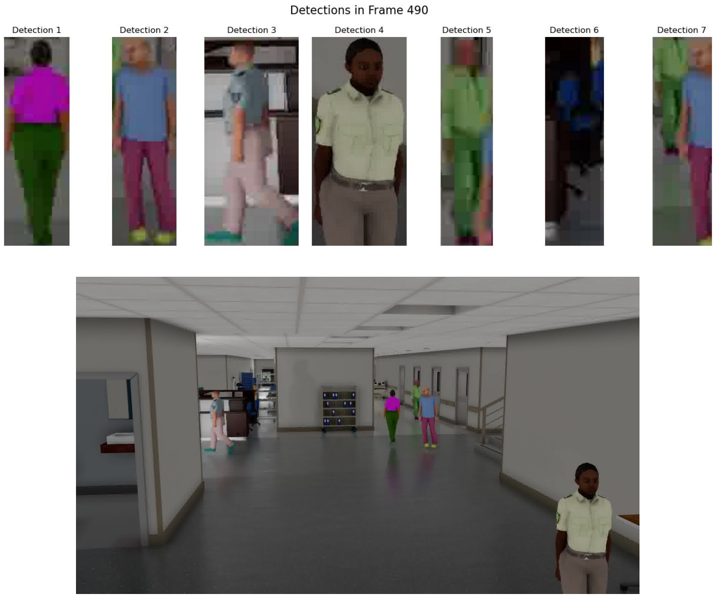
Problem Statement

- Goal to develop a multi-object tracking system for tracking objects in video sequences
- Focus on tracking desired objects, specifically people
- **Deliverable**: Achieve a high testing accuracy with our test dataset

Detection Challenges

- Challenges in object detection: obstructions, scale variations, cluttered backgrounds
- Inaccurate detections can impact tracking performance
- Overcoming challenges using involves tweaking algorithms
- Applying pre-processing techniques to improve detection quality

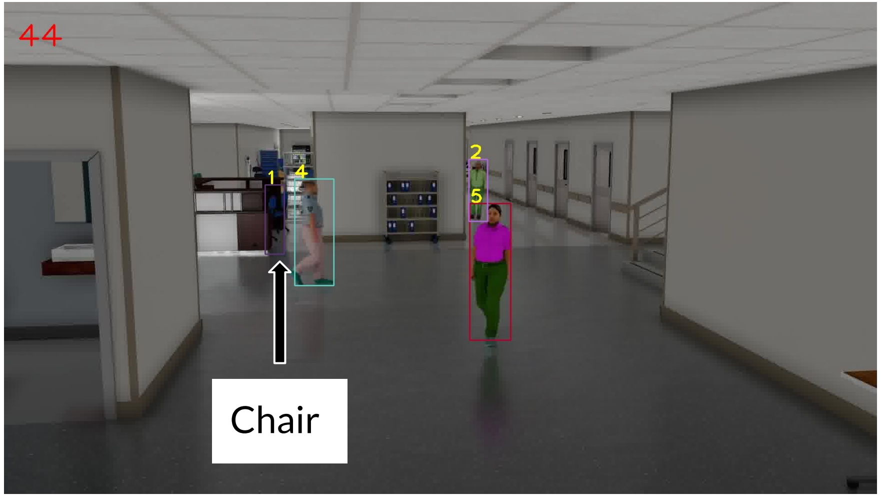


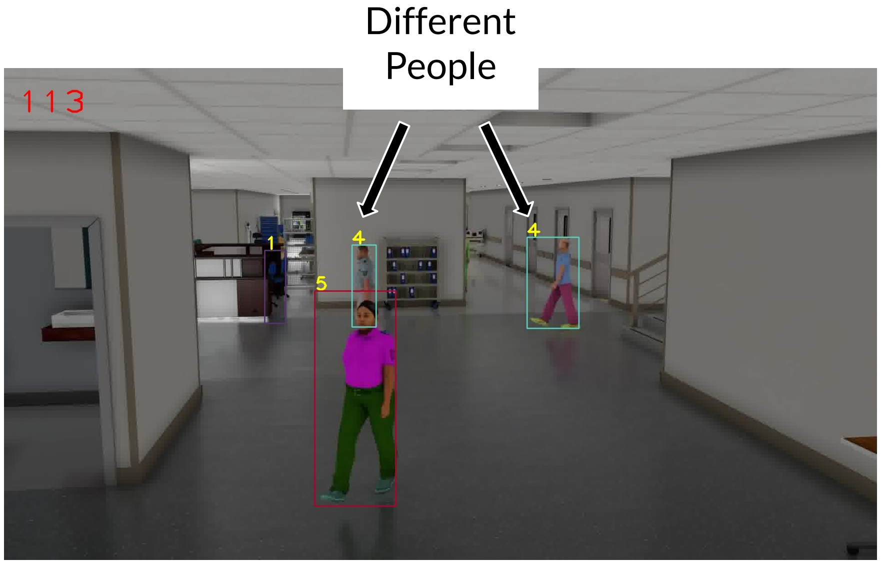


Detection Challenges and Obstacles

Visualized results from running test with original (unchanged) baseline code

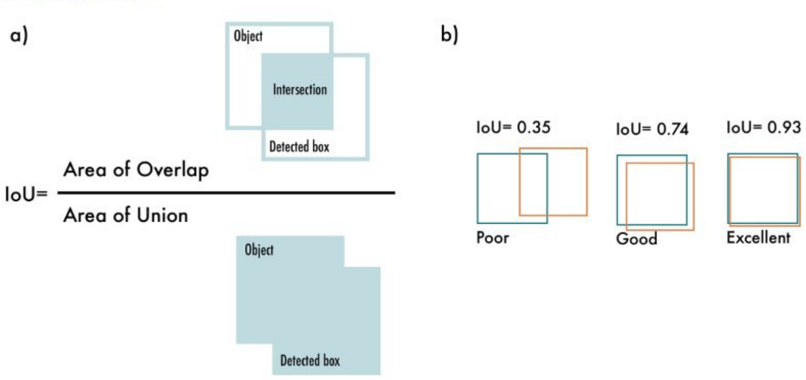
- “Recognizing” objects that are not people
- Classifying two people as the same person





Method Overview

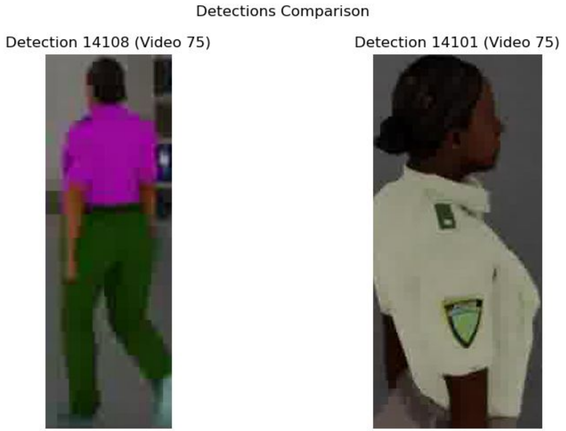
- Two main components:
    a. single-camera tracking using IoU(Intersection over union)
    b. Post-processing K-means clustering based on appearance
- Added confidence interval filtering to improve accuracy
- Combined approach improves tracking accuracy and reliability



Cosine Differences for Appearance Comparison

- Measure the dissimilarity between appearance features using cosine difference
- Lower cosine difference indicates higher similarity in appearance
- Cosine differences used for clustering and association of objects

Cosine Difference between the Two: 0.0010015368461608887



Cosine Difference between the two detections: 0.765363335609436

Object Pair Tracking

- Track object pairs across frames to establish object associations
- Use motion and appearance cues for pairing objects
- Predict object locations based on previous motion patterns
- Object pair tracking ensures continuity in object trajectories

Original

Next Frame

Removing Short Tracklet Merging

- This approach help group tracklets with similar appearances together
- Removing short tracklet merging improves the reliability of object trajectories in multi-object tracking.
- Tracklets are short track segments that capture the movement of objects over a brief period.
- Short tracklets, being brief segments, can introduce noise and inconsistencies to the tracking results.
- Merging short tracklets into longer ones enhances tracking accuracy by creating more complete and continuous trajectories.
- Two common clustering algorithms used for merging tracklets are
    a. Agglomerative Clustering-recursively merge clusters of sample data by linkage distance
    b. K-Means Clustering-recursively group data into K groups by distance from centroids

Filtering Low-Confidence Detections

- Remove low-confidence detections to improve tracking accuracy
- Set confidence threshold for detection filtering
- Eliminate detections with low confidence scores
- Focus on high-confidence detections for reliable tracking

Observations:

- MOTA (accuracy) highest at 0.3
- MOTP (precision) highest at 0.1
- IDF1 decreases as confidence increases
- Recall increases as confidence increases
- Overall, somewhere between 0.3 and 0.6 is a happy medium

	IDF1	IDP	IDR	idtp	idfp	idfn	MOTA	MOTP	Rcll	Prcn	IDs	MT	PT	ML
MultiCam	93.80	94.36	89.14	31039.0	1854.0	3782.0	84.71	10.87	88.33	98.08	658	3	1	1
	IDF1	IDP	IDR	idtp	idfp	idfn	MOTA	MOTP	Rcll	Prcn	IDs	MT	PT	ML
MultiCam	90.81	86.84	89.61	28864.0	4376.0	3348.0	85.51	10.74	93.77	96.32	1506	3	1	1
	IDF1	IDP	IDR	idtp	idfp	idfn	MOTA	MOTP	Rcll	Prcn	IDs	MT	PT	ML
MultiCam	90.07	79.14	92.48	27525.0	7257.0	2238.0	76.96	10.45	96.73	91.80	3313	3	0	2
	IDF1	IDP	IDR	idtp	idfp	idfn	MOTA	MOTP	Rcll	Prcn	IDs	MT	PT	ML
MultiCam	60.32	51.80	68.65	16870.0	15697.0	7703.0	63.11	9.74	97.99	76.78	1290	4	1	0

Camera 74 with confidence thresholds of 0.1, 0.3, 0.6, and 0.8

	IDF1	IDP	IDR	idtp	idfp	idfn	MOTA	MOTP	Rcll	Prcn	IDs	MT	PT	ML
MultiCam	100.16	94.14	97.12	52158.0	3246.0	1545.0	86.75	11.99	92.32	98.28	2119	4	1	0
	IDF1	IDP	IDR	idtp	idfp	idfn	MOTA	MOTP	Rcll	Prcn	IDs	MT	PT	ML
MultiCam	95.63	90.17	96.16	47981.0	5231.0	1917.0	90.01	11.77	97.82	96.75	2258	5	0	0
	IDF1	IDP	IDR	idtp	idfp	idfn	MOTA	MOTP	Rcll	Prcn	IDs	MT	PT	ML
MultiCam	81.77	72.22	85.98	39327.0	15127.0	6414.0	80.11	10.98	99.69	90.39	4110	5	0	0
	IDF1	IDP	IDR	idtp	idfp	idfn	MOTA	MOTP	Rcll	Prcn	IDs	MT	PT	ML
MultiCam	82.39	70.39	94.59	36819.0	15486.0	2106.0	65.04	10.35	99.98	77.14	2068	5	0	0

Camera 72 with confidence thresholds of 0.1, 0.3, 0.6, and 0.8

	IDF1	IDP	IDR	idtp	idfp	idfn	MOTA	MOTP	Rcll	Prcn	IDs	MT	PT	ML
MultiCam	88.08	75.04	92.80	52356.0	17411.0	4061.0	75.31	10.02	99.48	89.86	7303	5	0	0
	IDF1	IDP	IDR	idtp	idfp	idfn	MOTA	MOTP	Rcll	Prcn	IDs	MT	PT	ML
MultiCam	80.20	66.42	92.48	44219.0	22352.0	3596.0	60.26	9.18	99.99	76.54	4342	5	0	0

Camera 71 with confidence thresholds of 0.6 and 0.8

	IDF1	IDP	IDR	idtp	idfp	idfn	MOTA	MOTP	Rcll	Prcn	IDs	MT	PT	ML
MultiCam	88.71	76.34	88.50	47813.0	14818.0	6214.0	76.18	12.06	95.35	95.81	8105	3	1	1
	IDF1	IDP	IDR	idtp	idfp	idfn	MOTA	MOTP	Rcll	Prcn	IDs	MT	PT	ML
MultiCam	86.99	78.67	91.02	44782.0	12143.0	4416.0	82.13	11.63	99.04	90.62	3274	4	1	0
	IDF1	IDP	IDR	idtp	idfp	idfn	MOTA	MOTP	Rcll	Prcn	IDs	MT	PT	ML
MultiCam	79.66	65.38	93.63	37267.0	19733.0	2534.0	55.41	10.85	99.74	73.83	3571	5	0	0

Camera 73 with confidence thresholds of 0.3, 0.6, and 0.8

Results and Observations

Shown is Test case video 75 using 0.5 confidence interval.

Some error and mis-detection

Room for improvement

Conclusions

Challenge: implement single-camera MOT with reasonable success and consistency

We could improve through adding a kalman filter or other post processing algorithms

Results depends on how accurate you need to be

MOT is a valuable tool and has many applications

- Security (face, vehicle, object ID in cam footage)
- Social Media/Video Sharing (face tracking for filters)
- Scientific Research (faster classification of observations)