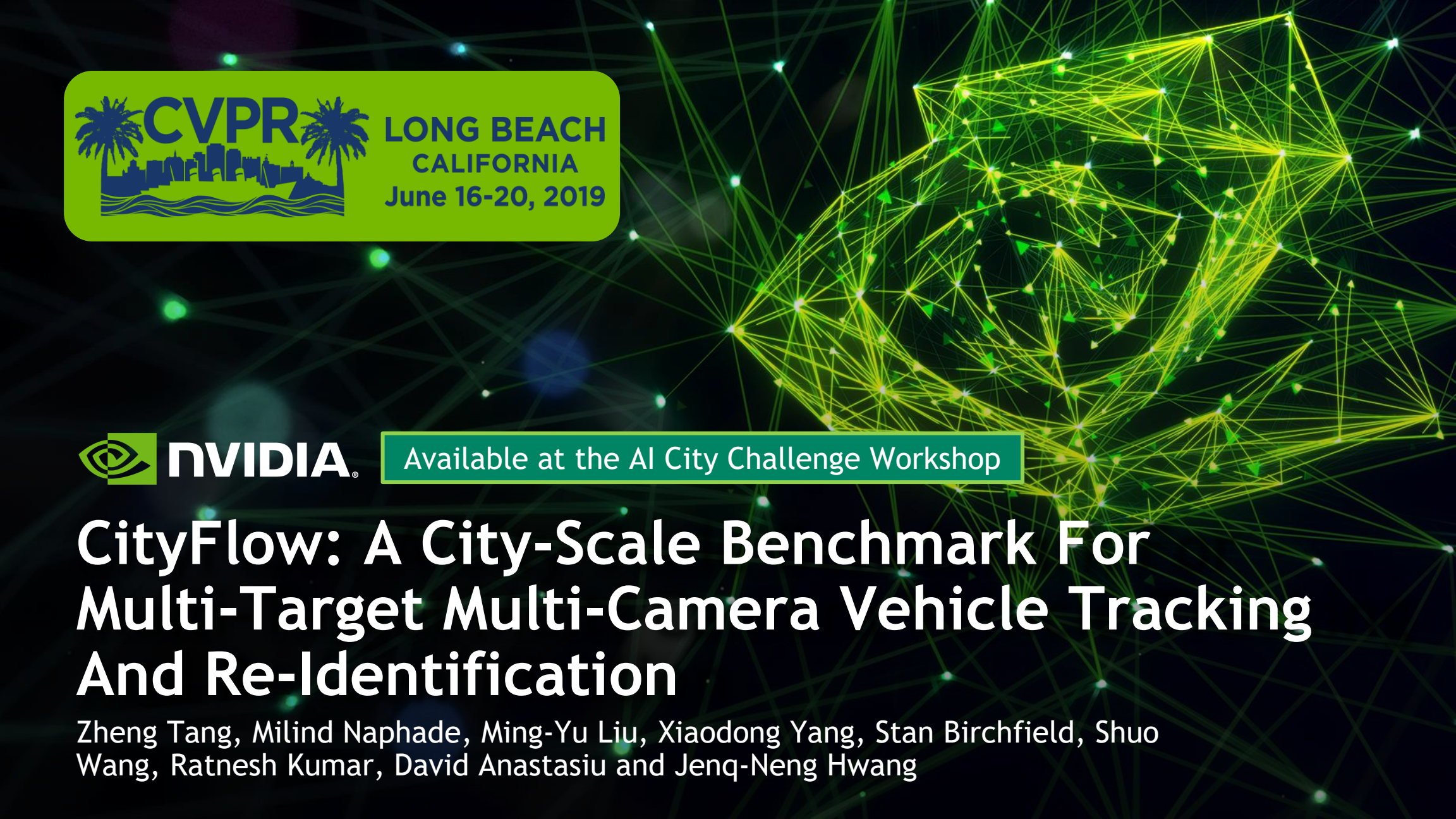


CityFlow: A City-Scale Benchmark For Multi-Target Multi-Camera Vehicle Tracking And Re-Identification

Zheng Tang, Milind Naphade, Ming-Yu Liu, Xiaodong Yang, Stan Birchfield, Shuo Wang, Ratnesh Kumar, David Anastasiu and Jenq-Neng Hwang

OBJECT REID vs. MTMC TRACKING

Person-based Problems

Object Re-identification (ReID)

Z. Zheng, L. Zheng, and Y. Yang. Unlabeled samples generated by GAN improve the person re-identification baseline in vitro. In *Proc. ICCV*, pages 3754–3762, 2017.

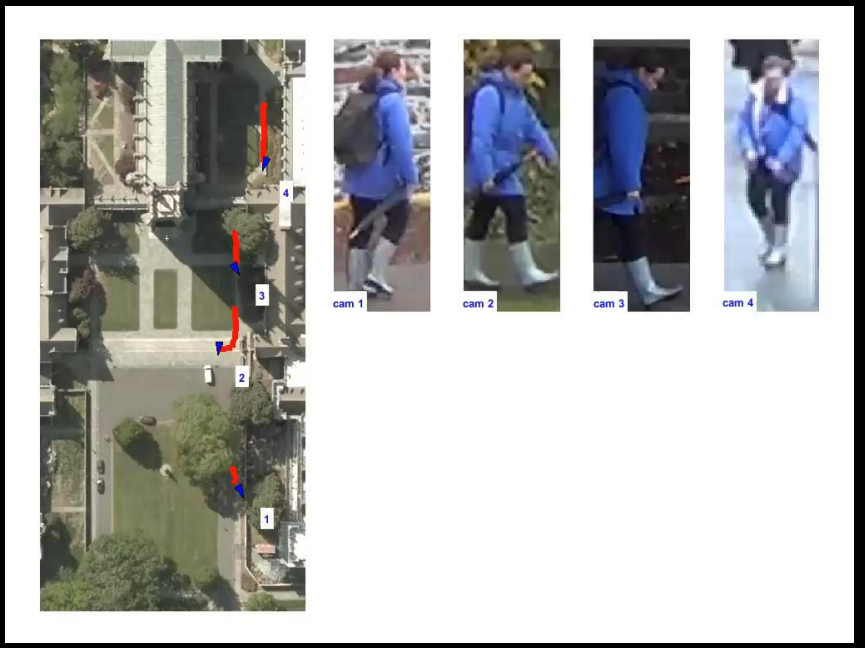
OBJECT REID vs. MTMC TRACKING

Person-based Problems

Object Re-identification (ReID)

Z. Zheng, L. Zheng, and Y. Yang. Unlabeled samples generated by GAN improve the person re-identification baseline in vitro. In *Proc. ICCV*, pages 3754–3762, 2017.

Multi-target Multi-camera (MTMC) Tracking



E. Ristani, F. Solera, R. Zou, R. Cucchiara, and C. Tomasi. Online-learning-based human tracking across non-overlapping cameras. In *Proc. ICCV*, pages 17-35, 2016.

OBJECT REID vs. MTMC TRACKING

Vehicle-based Problems

Challenges: High intra-class variability & inter-class similarity

Object Re-identification (ReID)

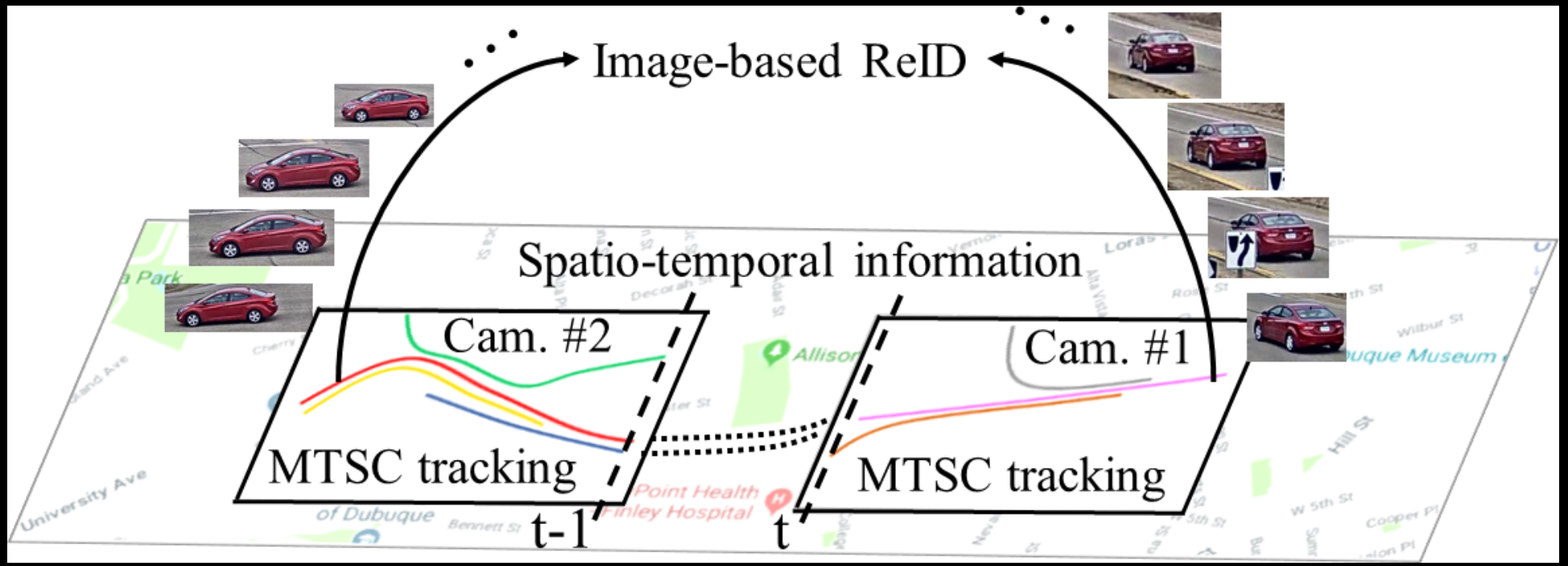
X. Liu, W. Liu, T. Mei, and H. Ma. PROVID: Progressive and multimodal vehicle reidentification for large-scale urban surveillance. *IEEE TMM*, 20(3):645–658, 2017.

Multi-target Multi-camera (MTMC) Tracking

Z. Tang, G. Wang, H. Xiao, A. Zheng, and J. N. Hwang. Single-camera and inter-camera vehicle tracking and 3D speed estimation based on fusion of visual and semantic features. In *Proc. CVPR Workshops*, pages 108–115, 2018.

SUMMARY OF MTMC TRACKING

Image-based ReID + MTSC Tracking + Spatio-temporal Association



MTMC TRACKING & REID BENCHMARKS

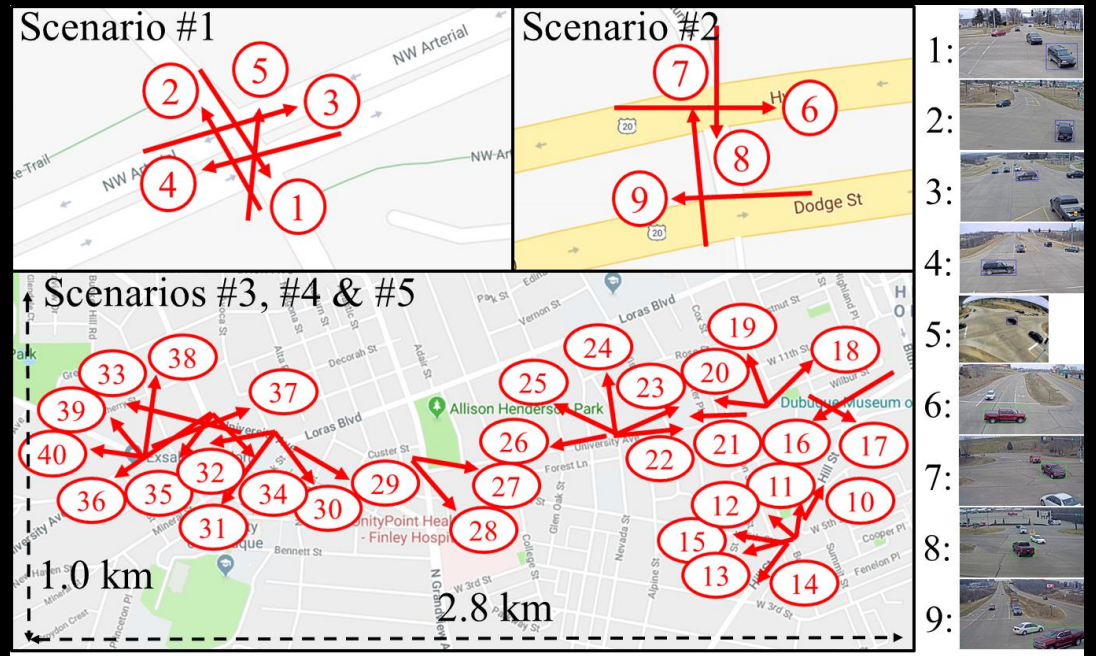
		Benchmark	# cameras	# boxes	# boxes/ID	Video	Geom.	Multiview
person	ReID	Market1501 [59]	6	32,668	30.8	✗	✗	✓
		DukeMTMC-reID [35, 63]	8	36,411	20.1	✗	✗	✓
		MSMT17 [47]	15	126,441	21.8	✗	✗	✓
		CUHK03 [23]	2	13,164	19.3	✗	✗	✗
		CUHK01 [22]	2	3,884	4.0	✗	✗	✗
		VIPeR [12]	2	1,264	2.0	✗	✗	✗
		PRID [15]	2	1,134	1.2	✗	✗	✗
		CAVIAR [9]	2	610	8.5	✗	✗	✗
	MTMC	MARS [58]	6	1,191,003	944.5	✗	✗	✓
		DukeMTMC [35, 53]	8	4,077,132	571.2	✓	✓	✓
		NLPR_MCT [8]	12	36,411	65.8	✓	✓	✓

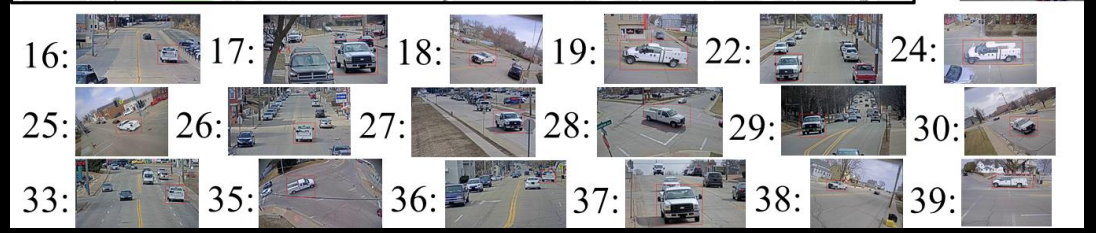
MTMC TRACKING & REID BENCHMARKS

		Benchmark	# cameras	# boxes	# boxes/ID	Video	Geom.	Multiview
person	ReID	Market1501 [59]	6	32,668	30.8	✗	✗	✓
		DukeMTMC-reID [35, 63]	8	36,411	20.1	✗	✗	✓
		MSMT17 [47]	15	126,441	21.8	✗	✗	✓
		CUHK03 [23]	2	13,164	19.3	✗	✗	✗
		CUHK01 [22]	2	3,884	4.0	✗	✗	✗
		VIPeR [12]	2	1,264	2.0	✗	✗	✗
		PRID [15]	2	1,134	1.2	✗	✗	✗
		CAVIAR [9]	2	610	8.5	✗	✗	✗
	MTMC	MARS [58]	6	1,191,003	944.5	✗	✗	✓
		DukeMTMC [35, 53]	8	4,077,132	571.2	✓	✓	✓
		NLPR_MCT [8]	12	36,411	65.8	✓	✓	✓
vehicle	ReID	VeRi-776 [29]	20	49,357	63.6	✗	✓	✓
		VehicleID [27]	2	221,763	8.4	✗	✗	✗
		PKU-VD1 [55]	-	846,358	6.0	✗	✗	✗
		PKU-VD2 [55]	-	807,260	10.1	✗	✗	✗
	MTMC	**CityFlow (proposed)**	**40**	**229,680**	**344.9**	✓	✓	✓

PROPOSED BENCHMARK DATASET

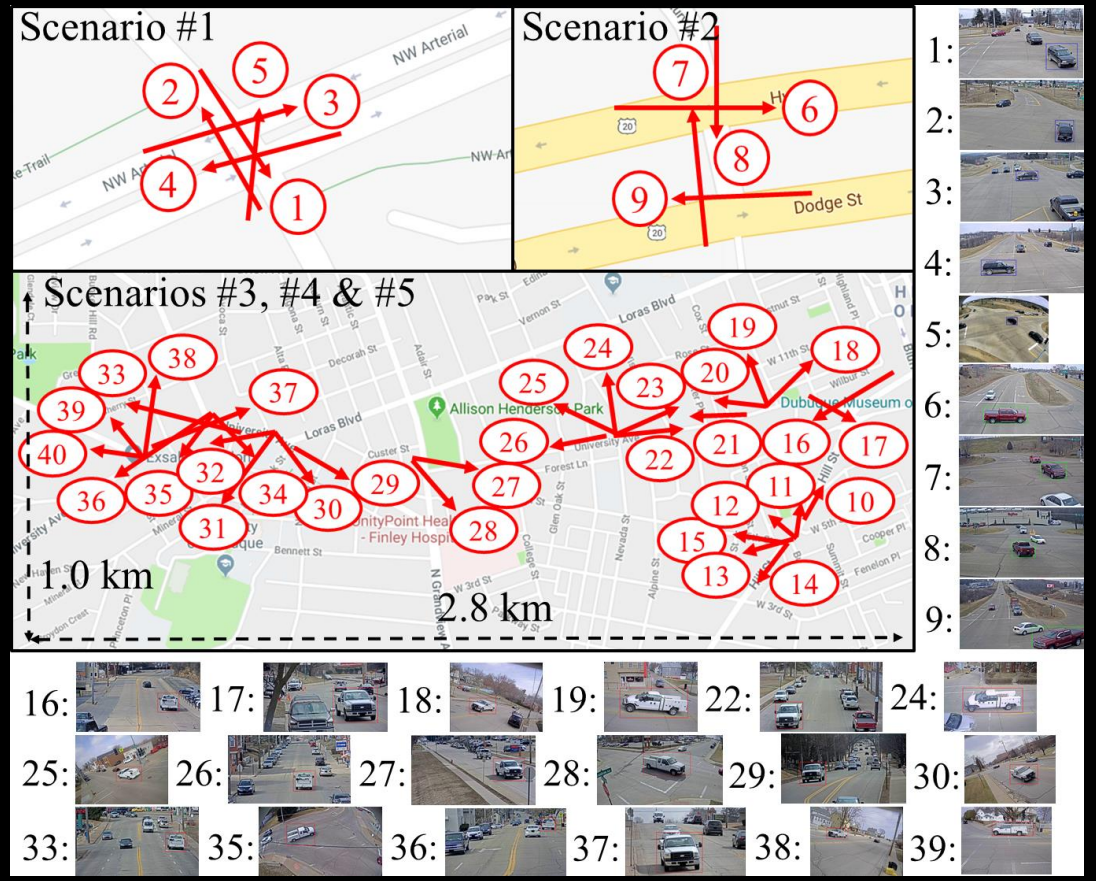
Videos Captured by 40 Cameras across 10 Intersections





PROPOSED BENCHMARK DATASET

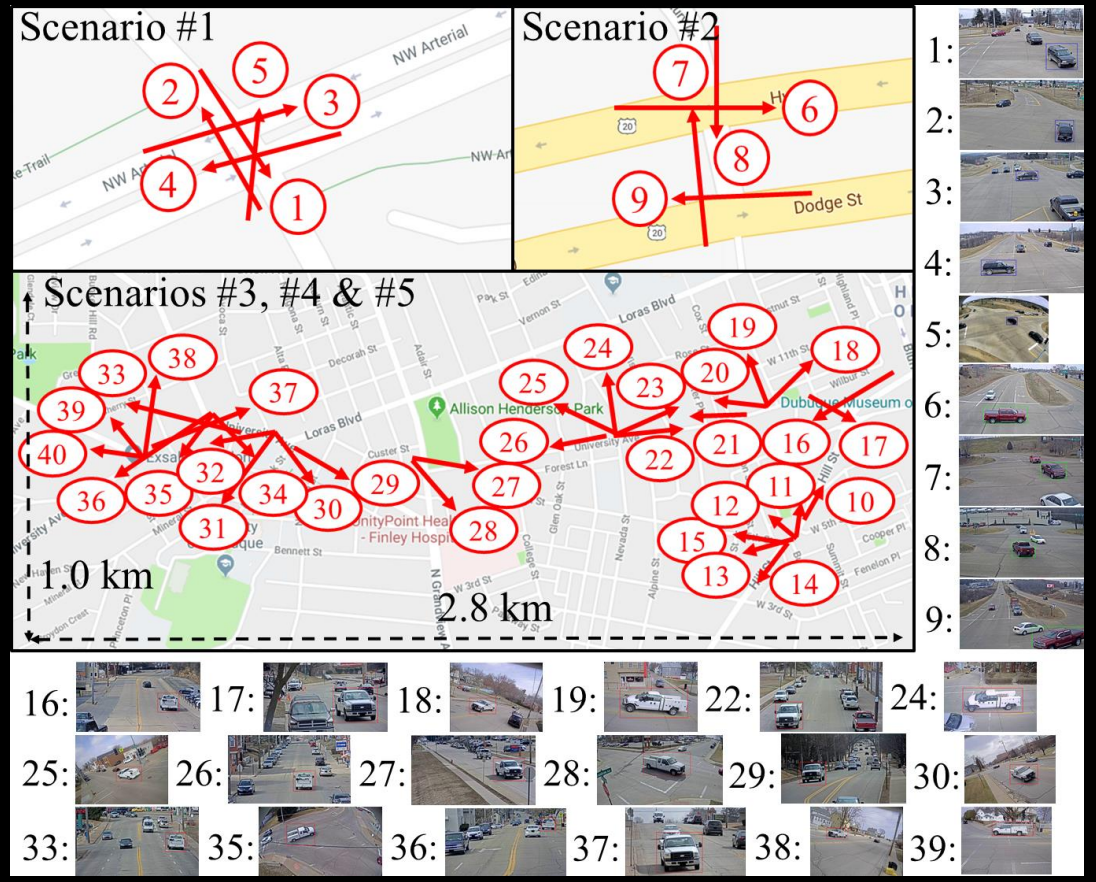
Videos Captured by 40 Cameras across 10 Intersections



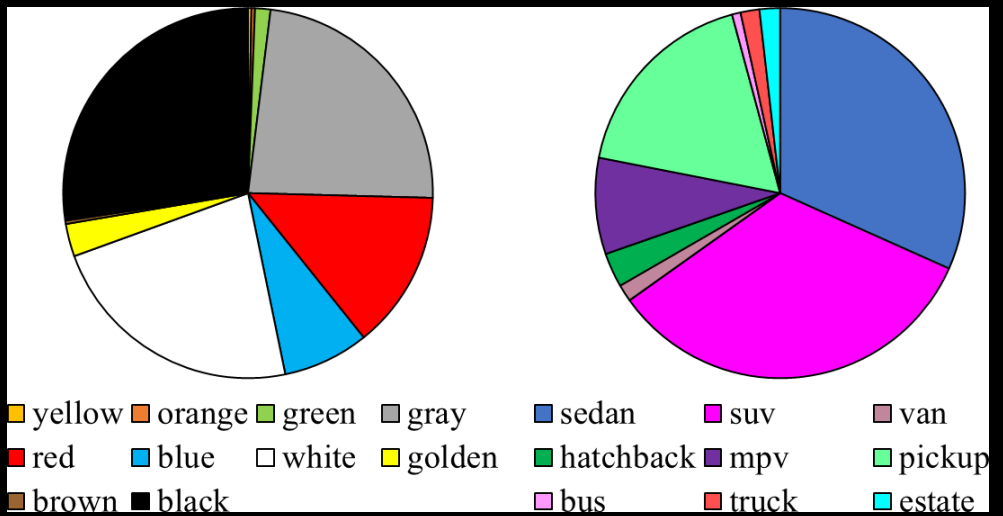
	Time (min.)	# cam.	# boxes	# IDs	Scene type	LOS
1	17.13	5	20,772	95	highway	A
2	13.52	4	20,956	145	highway	B
3	23.33	6	6,174	18	residential	A
4	17.97	25	17,302	71	residential	A
5	123.08	19	164,476	337	residential	B
total	195.03	40	229,680	666		

PROPOSED BENCHMARK DATASET

Videos Captured by 40 Cameras across 10 Intersections

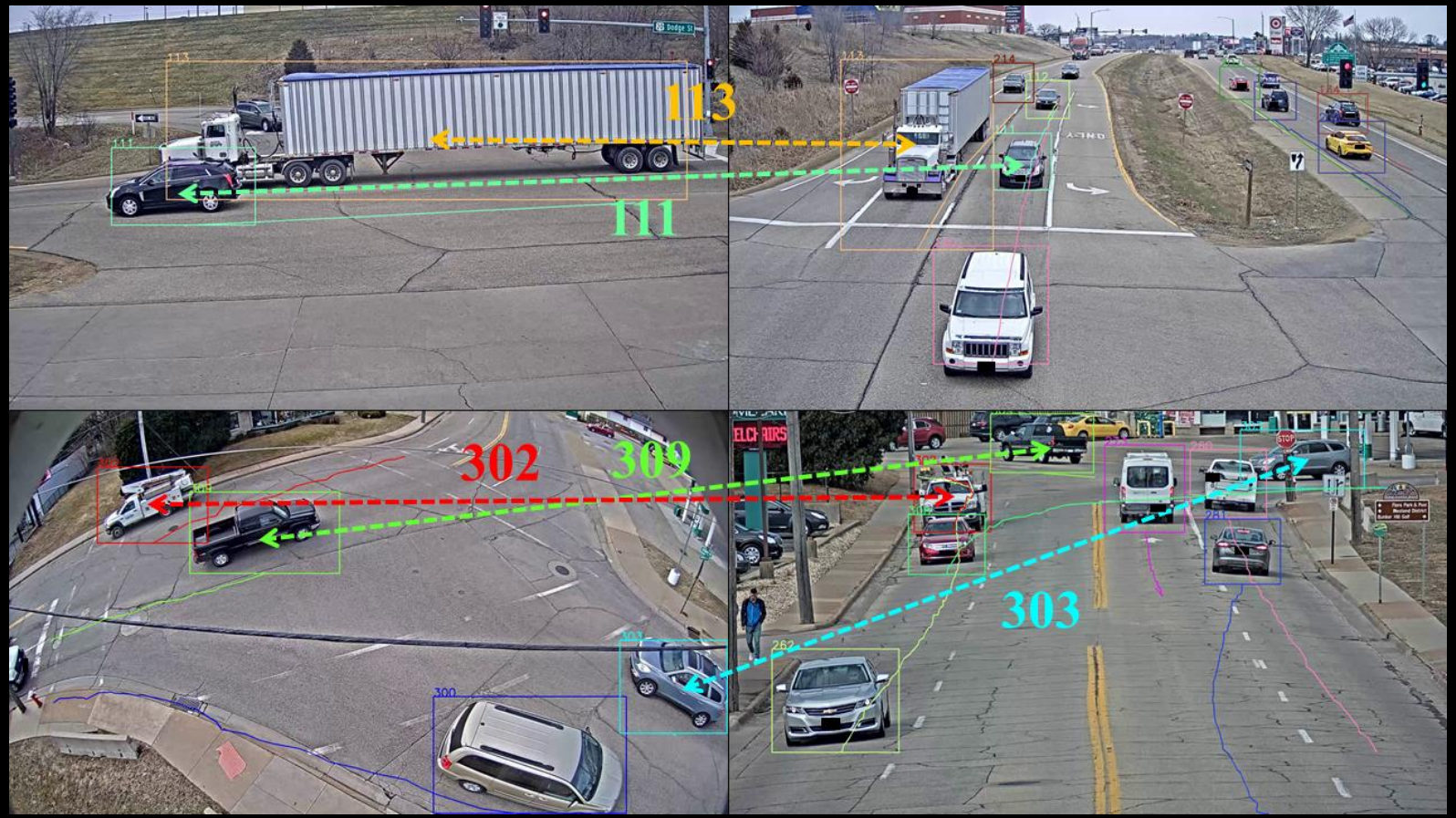


	Time (min.)	# cam.	# boxes	# IDs	Scene type	LOS
1	17.13	5	20,772	95	highway	A
2	13.52	4	20,956	145	highway	B
3	23.33	6	6,174	18	residential	A
4	17.97	25	17,302	71	residential	A
5	123.08	19	164,476	337	residential	B
total	195.03	40	229,680	666		

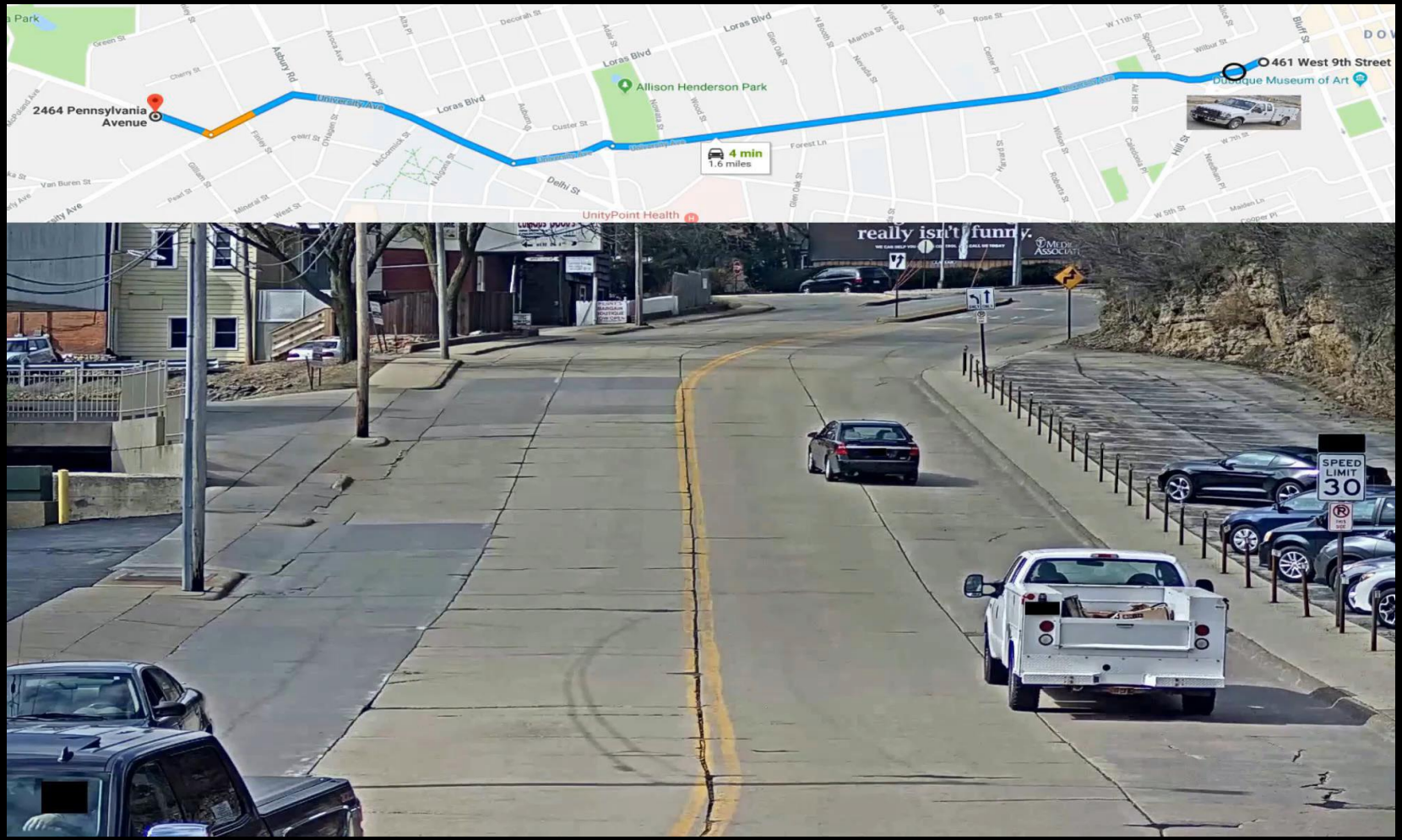


ACCELERATED MTMC ANNOTATION

Assisted by Automated Single-camera Tracking

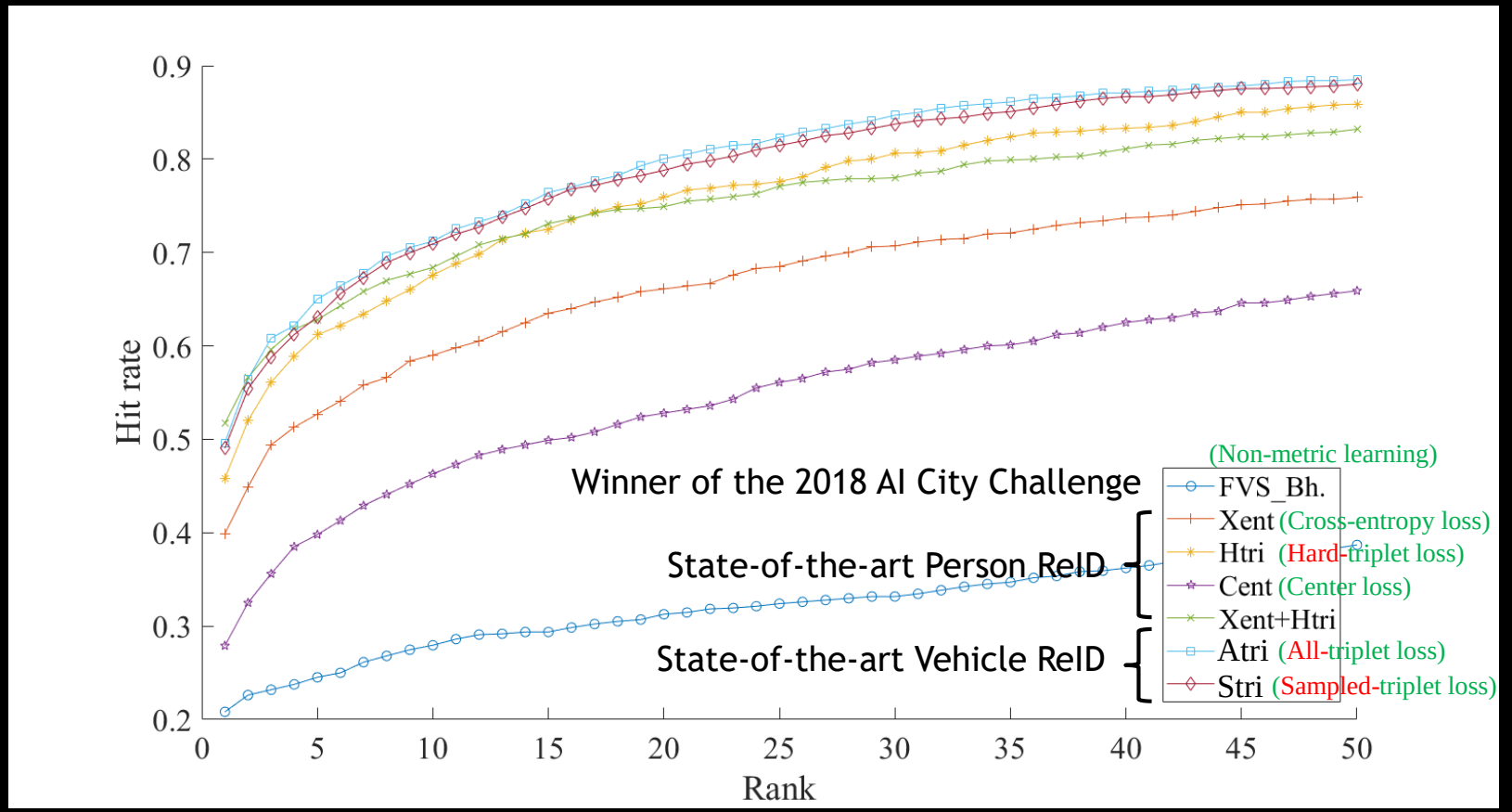


DEMONSTRATION OF CAMERA VIEWS



EVALUATION OF IMAGE-BASED REID

Cumulative Matching Characteristic (CMC) Curves



- Z. Tang, G. Wang, H. Xiao, A. Zheng, and J. N. Hwang. Single-camera and inter-camera vehicle tracking and 3D speed estimation based on fusion of visual and semantic features. In *Proc. CVPR Wokshops*, pages 108-115, 2018.
- K. Zhou. deep-person-reid. https://github.com/KaiyangZhou/deep-person-reid, 2018.
- R. Kumar, E. Weill, F. Aghdasi, and P. Sriram. Vehicle reidentification: an efficient baseline using triplet embedding. In *Proc. IJCNN*, 2019.

EVALUATION OF IMAGE-BASED REID

Qualitative Performance Comparison

(Non-metric learning) FVS_Bh.

(Cross-entropy) Xent

(Hard-triplet) Htri

(Center) Cent

Xent+Htri

(All-triplet) Atri

(Sampled-triplet) Stri

EVALUATION OF MTSC TRACKING

Evaluation Metrics Adopted from MOTChallenge

Method	IDF1	Recall	FAR	MT	MOTA	MOTP
DS+YOLO	78.9%	67.6%	8.6	778	67.4%	65.8%
DS+SSD	79.5%	69.2%	8.3	756	68.9%	65.5%
DS+FRCNN	78.9%	66.9%	15.3	761	66.7%	65.5%
TC+YOLO	79.1%	68.1%	8.5	871	68.0%	**66.0%**
TC+SSD	**79.7%**	**70.4%**	7.4	895	**70.3%**	65.6%
TC+FRCNN	78.7%	68.5%	12.0	957	68.4%	65.9%
MO+YOLO	77.8%	69.0%	8.5	965	68.6%	**66.0%**
MO+SSD	72.8%	68.0%	**6.3**	980	67.0%	65.9%
MO+FRCNN	75.6%	69.5%	10.8	**1094**	68.6%	**66.0%**

- L. Leal-Taixé, A. Milan, I. Reid, S. Roth, and K. Schindler. MOTChallenge 2015: Towards a benchmark for multi-target tracking. *arXiv:1504.01942*, 2015.
N. Wojke, A. Bewley, and D. Paulus. Simple online and realtime tracking with a deep association metric. In *Proc. ICIP*, pages 3645-3649, 2017.
- Z. Tang, G. Wang, H. Xiao, A. Zheng, and J. N. Hwang. Single-camera and inter-camera vehicle tracking and 3D speed estimation based on fusion of visual and semantic features. In *Proc. CVPR Workshops*, pages 108-115, 2018.
- Z. Tang and J. N. Hwang. MOANA: An online learned adaptive appearance model for robust multiple object tracking in3D. *IEEE Access*, 7(1):31934-31945, 2019
- J. Redmon and A. Farhadi. YOLOv3: An incremental improvement. *arXiv:1804.02767*, 2018.
- W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg. SSD: Single shot multibox detector. In *Proc. ECCV*, pages 21-37, 2016.
- S. Ren, K. He, R. Girshick, and J. Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. In *Proc. NeurIPS*, pages 91-99, 2015.

- MTSC tracking baselines: **DS**: Deep SORT; **TC**: Tracklet clustering; **MO**: MOANA
- Object detection baselines: **YOLO**: YOLOv3; **SSD**: SSD512; **FRCNN**: Faster R-CNN

Note: False positives are not considered as only vehicles passing through >1 camera are annotated.

EVALUATION OF MTMC TRACKING

Image-based ReID + MTSC Tracking + Spatio-temporal Association

Spatio-temporal association	MTSC tracking	Image-based ReID						
		FVS_Bh.	Xent	Htri	Cent	Xent+Htri	BA	BS
PROVID [29]	DeepSORT [50]	21.5%	31.3%	35.3%	27.6%	34.5%	**35.6%**	33.6%
	TC [43]	22.1%	35.2%	39.4%	32.7%	39.9%	**40.6%**	39.0%
	MOANA [40]	21.7%	29.1%	33.0%	26.1%	31.9%	**34.4%**	31.8%
2WGMMF [20]	DeepSORT [50]	25.0%	35.3%	38.4%	31.2%	37.5%	**40.3%**	39.8%
	TC [43]	**27.6%**	39.5%	41.7%	34.7%	**43.3%**	44.1%	**45.1%**
	MOANA [40]	20.2%	32.2%	35.9%	28.2%	36.5%	**38.1%**	37.7%
FVS [43]	DeepSORT [50]	24.9%	36.4%	40.0%	30.8%	39.0%	41.3%	**41.4%**
	TC [43]	**27.6%**	**40.5%**	**42.7%**	**36.6%**	42.4%	**46.3%**	**46.0%**
	MOANA [40]	21.2%	32.7%	36.4%	29.2%	37.5%	**39.5%**	36.9%

- Spatio-temporal association schemes: **PROVID**: Spatio-temporal re-ranking;

  2WGMMF: Learning transition time; **FVS**: Manually set transition time

- X. Liu, W. Liu, T. Mei, and H. Ma. PROVID: Progressive and multimodal vehicle reidentification for large-scale urban surveillance. *IEEE TMM*, 20(3):645-658, 2017.
- Y. G. Lee, Z. Tang, J. N. Hwang, and Z. Fang. Online-learning-based human tracking across non-overlapping cameras. *IEEE TCSVT*, 28(10):2870-2883, 2018.
- Z. Tang, G. Wang, H. Xiao, A. Zheng, and J. N. Hwang. Single-camera and inter-camera vehicle tracking and 3D speed estimation based on fusion of visual and semantic features. In *Proc. CVPR Workshop*, pages 108-115, 2018.

Benchmark datasets and evaluation server available at the

AI City Challenge Workshop

- 3 challenge tracks
- 334 participating teams from 44 countries
- 1,337 successful submissions

Poster 3.1 #90