

FORT

Fisheye Online Realtime Tracking with an improved Kalman Filter

Nathan Odic, Benoît Faure, and Baptiste Magnier



IEEE MMSP 2023

September 27-29, 2023 | Poitiers, France



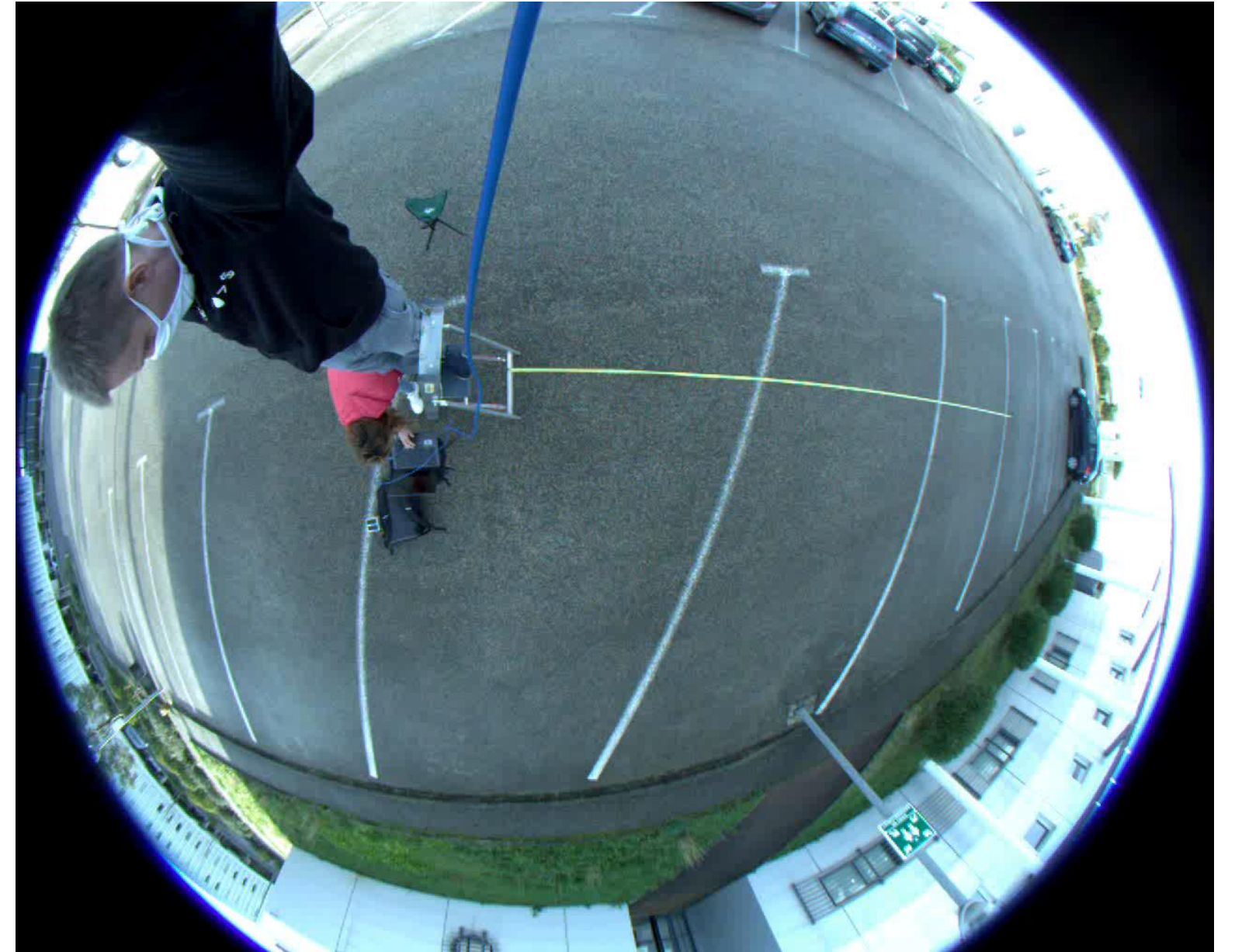
IMT Mines Alès
École Mines-Télécom



Table of Contents

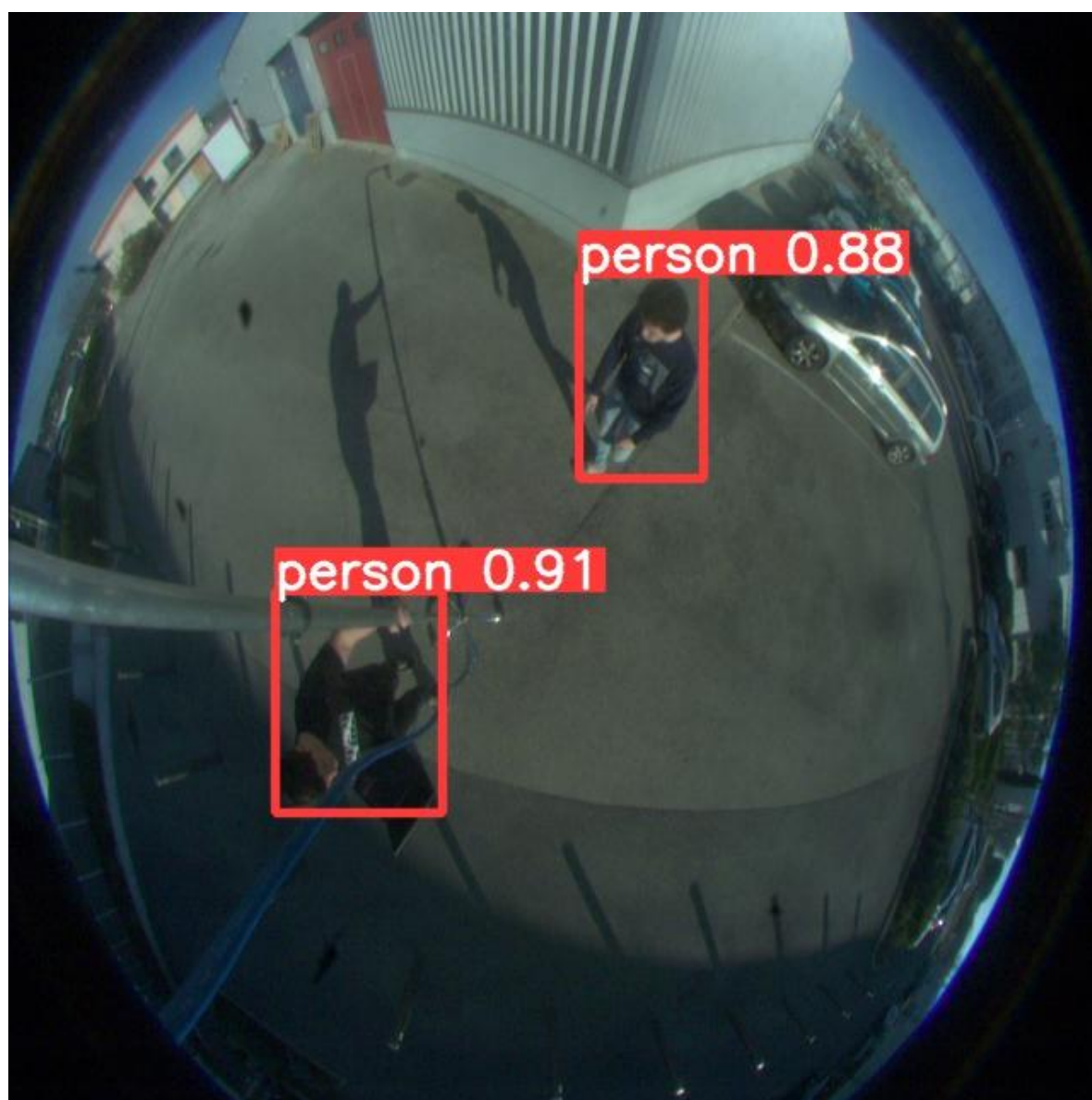
1. What is a fisheye lens
2. How tracking algorithm works
3. State of the Art
4. Our Algorithm, FORT
 1. FORT Mechanism
 2. YOLO Detection
 3. YOLO Comparison
 4. Association
 5. FORT Global Architecture
5. Comparison with SOTA tracking algorithms
6. Conclusion

1. What is a fisheye camera?



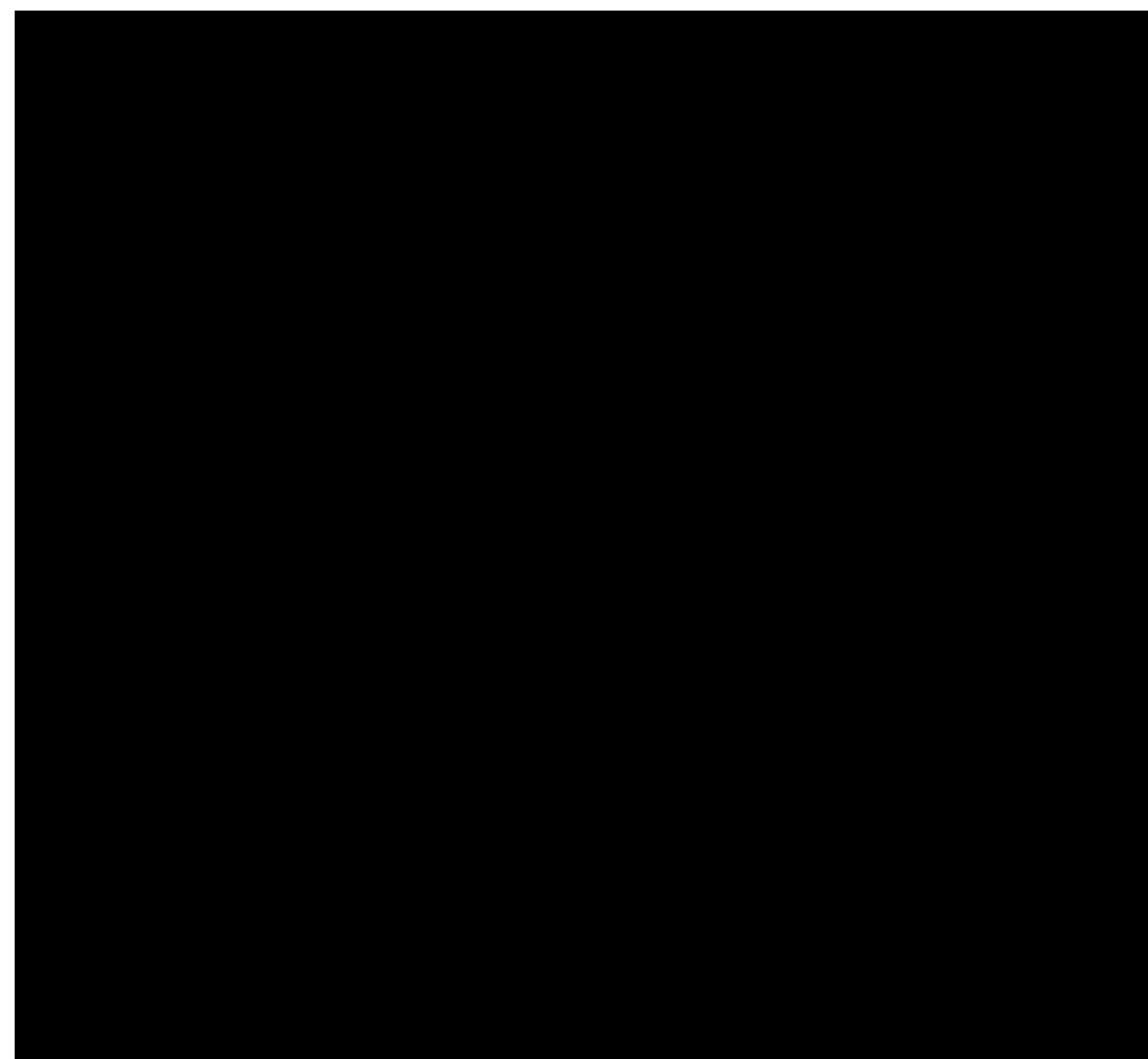
2. What is tracking?

Detection



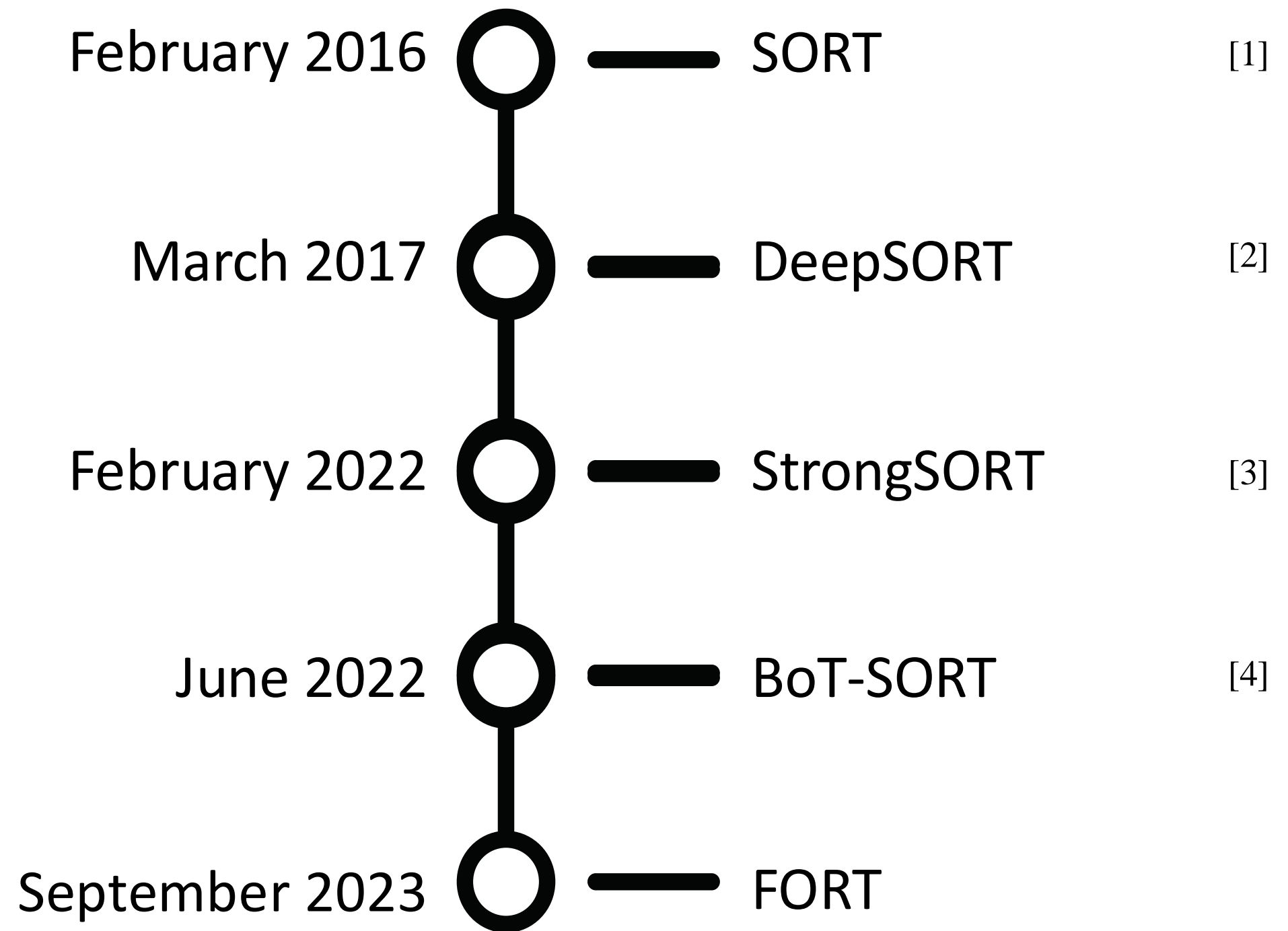
Trained on more than 2000 images

Tracking



Faure, Benoit, et al. "Performance of recent tiny/small yolo versions in the context of top-view fisheye images." *International Conference on Image Analysis and Processing*. Cham: Springer International Publishing, 2022.

3. State of the art



[1] Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). **Simple online and realtime tracking**. In *IEEE ICIP* (pp. 3464–3468).

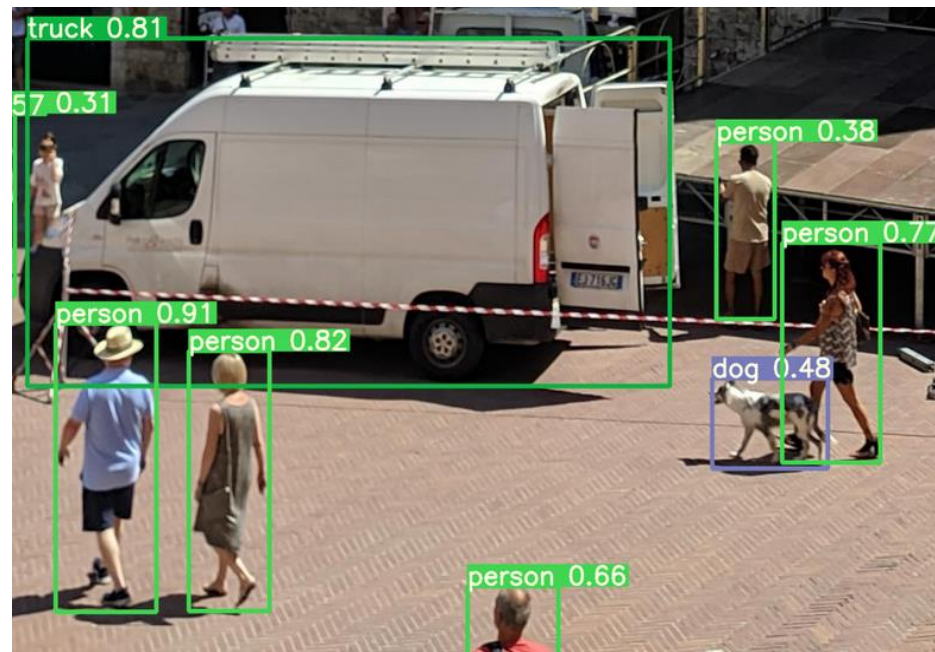
[2] Wojke, N., Bewley, A., & Paulus, D. (2017). **Simple online and realtime tracking with a deep association metric**. In *IEEE ICIP* (pp. 3645–3649).

[3] Du, Y., Song, Y., Yang, B., & Zhao, Y. (2022). **StrongSORT: Make deepsort great again**. *arXiv:2202.13514*.

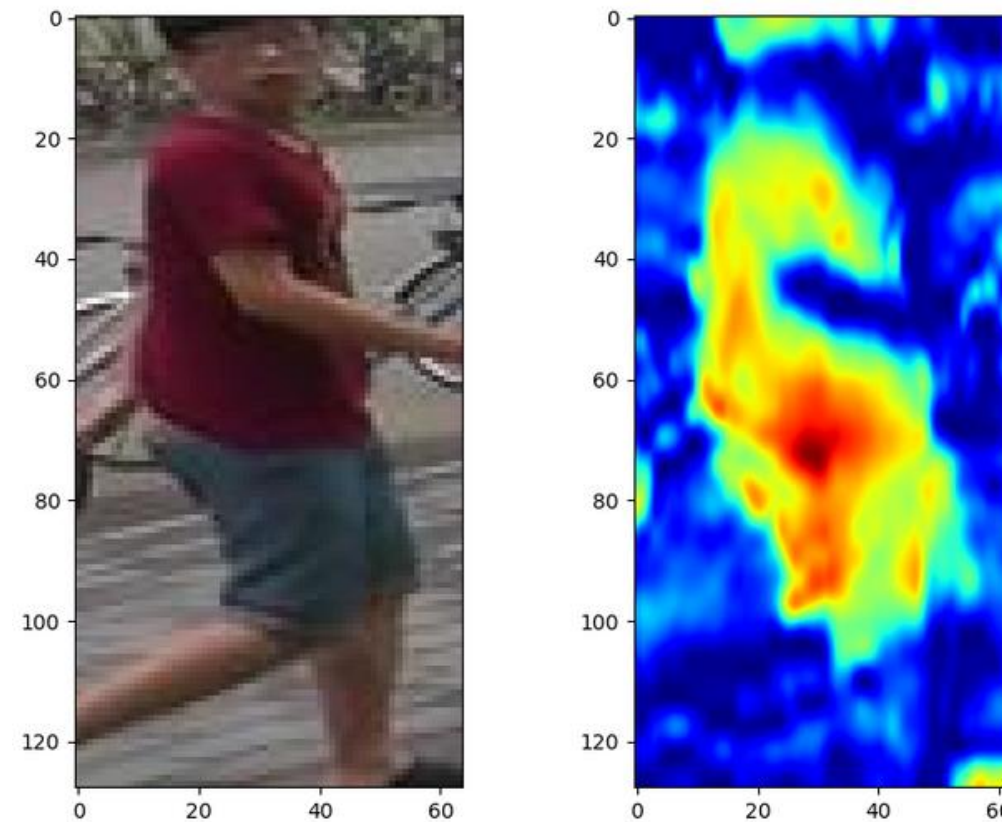
[4] Aharon, N., Orfaig, R., & Bobrovsky, B.Z. (2022). **BoT-SORT: Robust Associations Multi-Pedestrian Tracking**. *arXiv:2206.14651*.

4. Our Algorithm, FORT

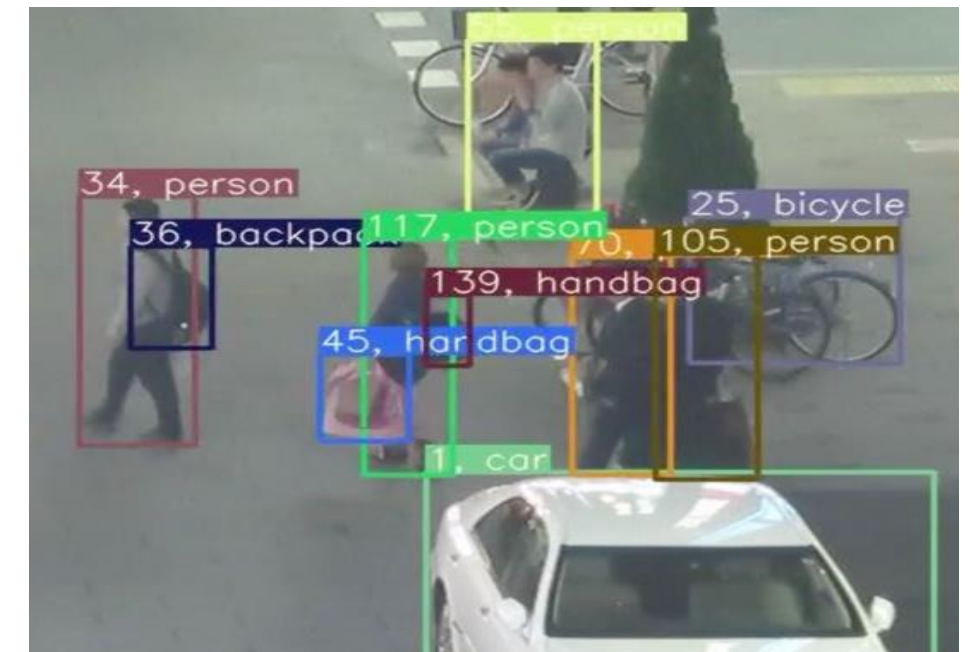
4.1 FORT Mechanism



Detection

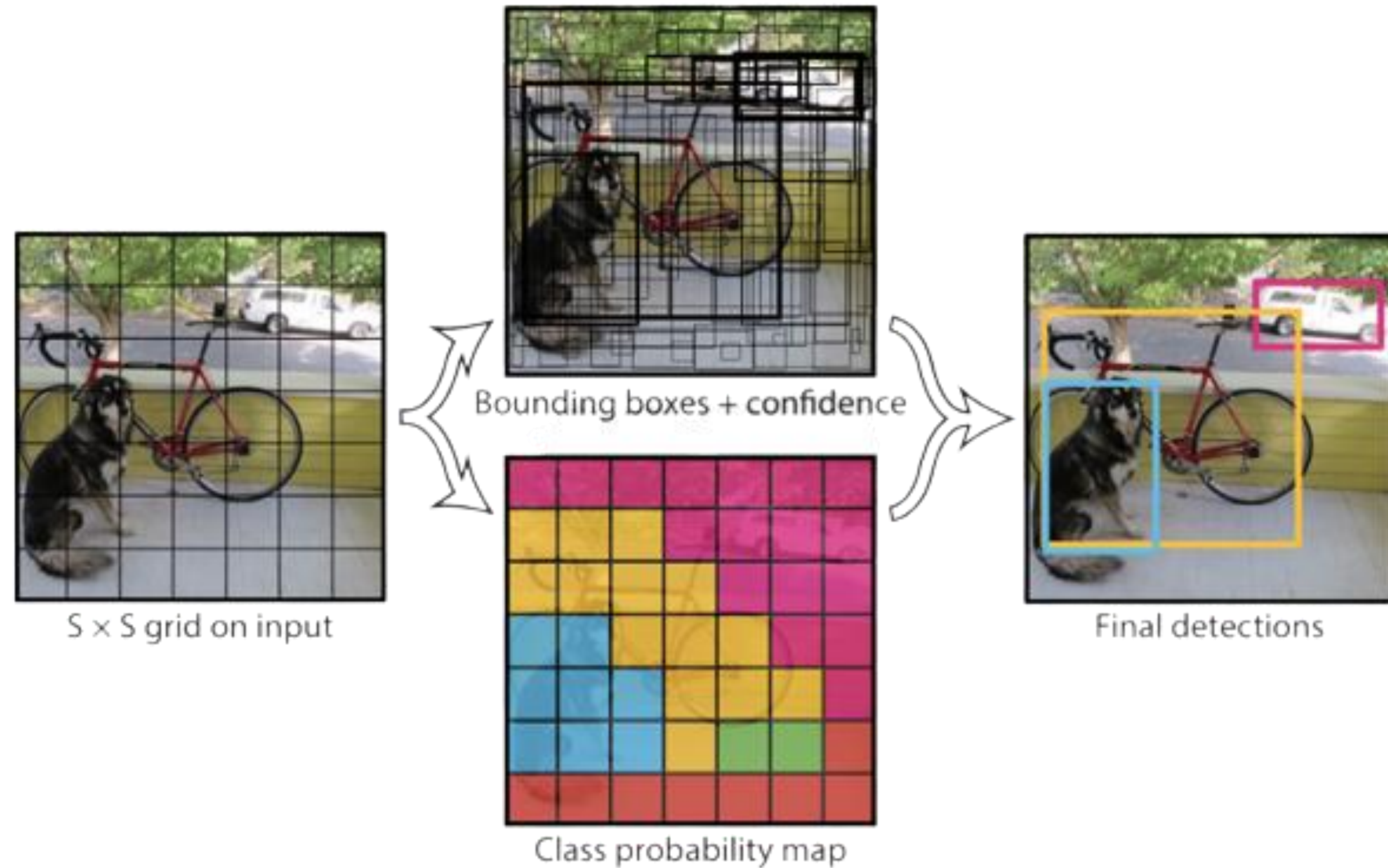


Feature Extraction
and IoU



Association

4.2 YOLO Detection



Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016).
You only look once: Unified, real-time object detection. In *Proceedings of the IEEE CVPR*.

4.3 YOLO Comparison

Comparaison 1

Version	Precision	Recall	$mAP_{0.5}$	$mAP_{0.5:0.95}$	Detection (ms)
YOLOv3-t	0.988	0.961	0.984	0.644	20.1
YOLOv4-t	0.976	0.936	0.976	0.626	15.2
→ YOLOv5-s	0.987	0.991	0.993	0.724	25.8

Dataset: 2313 train + 581 test Source: CERIS Lab + Mirror World

B. Faure, N. Odic, O. Haggui, and B. Magnier, “**Performance of recent tiny/small YOLO versions in the context of top-view fisheye images,**” in ISHAPE 2022

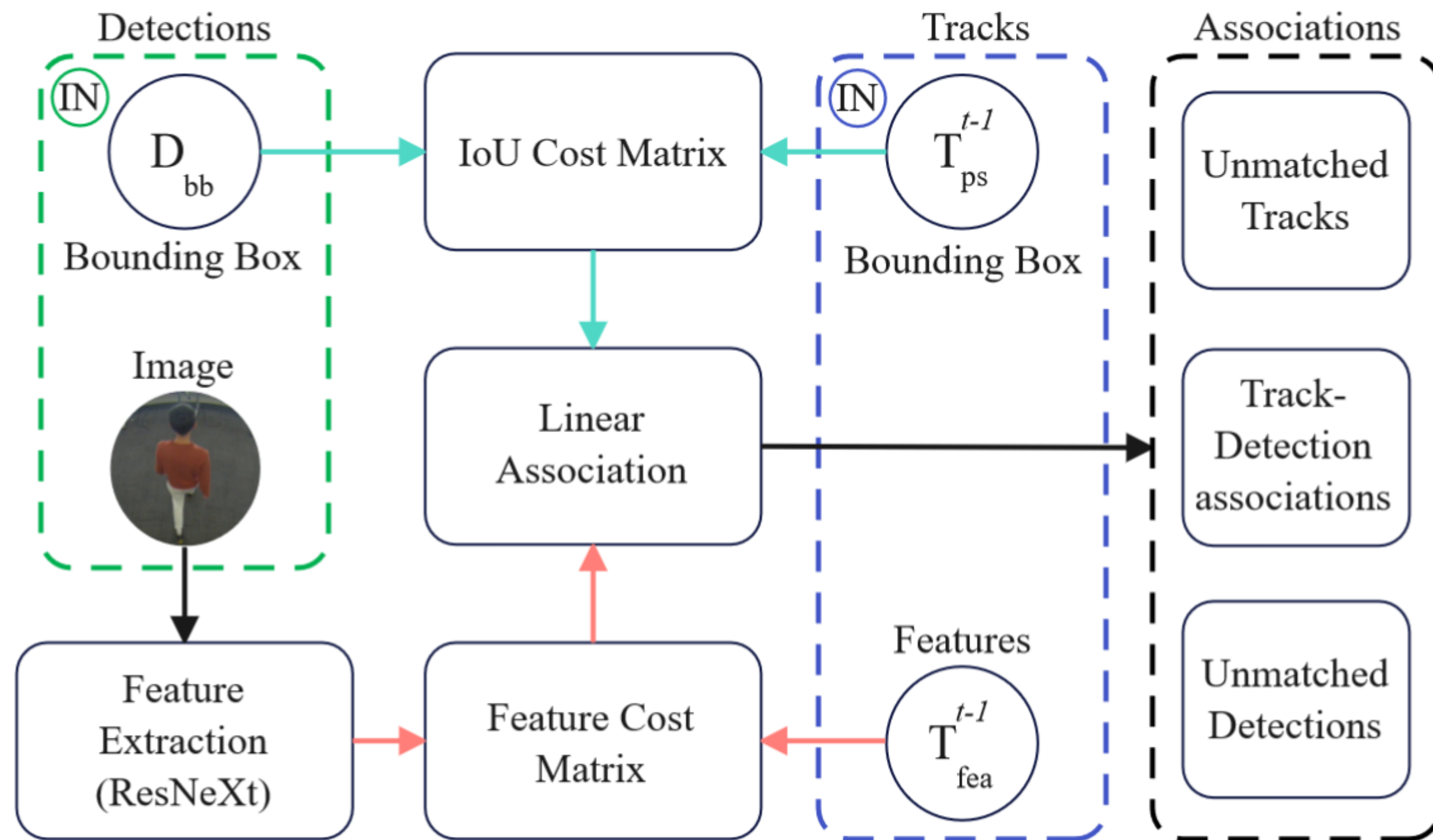
Comparaison 2

Version	$mAP_{0.5}$	$mAP_{0.5:0.95}$	Detection (ms)	Weights Size
YOLOv5l	0.716	0.421	31.3	147 MB
YOLOX	0.636	0.341	36.6	194 MB
→ YOLOv7x	0.789	0.393	28.5	137 MB

Dataset: 2828 train + 491 test

Source: CERIS Lab + Mirror World

4.4 Association



Association between selected detections and tracks.

4.5 FORT Global Architecture

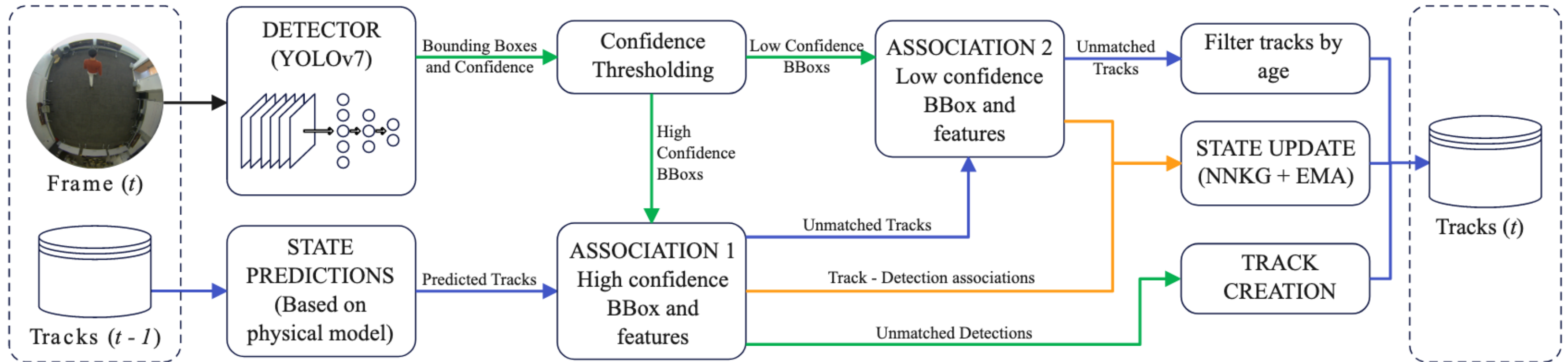


Fig. 1. Global architecture of the FORT algorithm for a frame at time t .

4.5 FORT Global Architecture

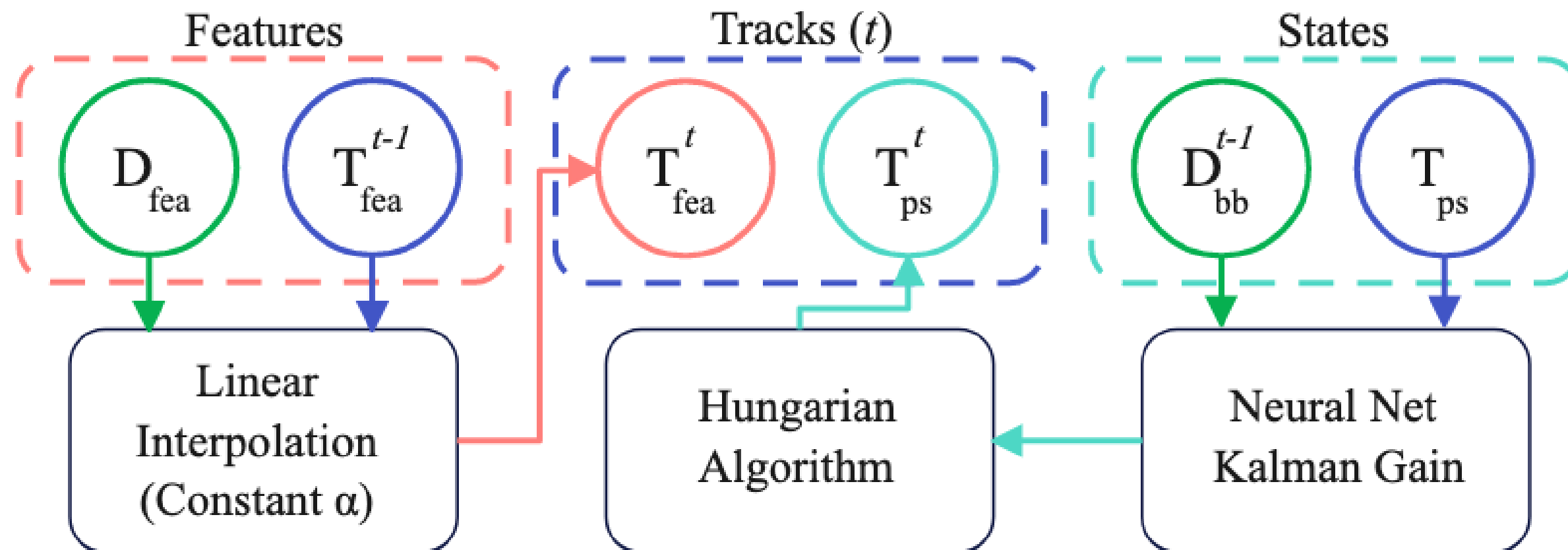


Fig. 3. Update process for a track associated with a detection

5. Comparison

Metrics

[1]
$$\text{MOTA} = 1 - \frac{|\text{FN}| + |\text{FP}| + |\text{IDSW}|}{|\text{gtDet}|}$$
 Did the algorithm often get it right?

[2]
$$\text{IDF1} = \frac{|\text{IDTP}|}{|\text{IDTP}| + 0.5 |\text{IDFN}| + 0.5 |\text{IDFP}|}$$
 Was it the right person?

[3]
$$\text{HOTA}$$
 Was it in the right place?
(And all of the above)

Dataset

20 videos / 9750 frames
Source: CERIS Lab + Mirror World

[1] Bernardin, K., & Stiefelhagen, R. (2008). Evaluating multiple object tracking performance: the clear mot metrics. *EURASIP JIVP*, 2008, 1–10.

[2] Ristani, E., Solera, F., Zou, R., Cucchiara, R., & Tomasi, C. (2016). Performance measures and a data set for multi-target, multi-camera tracking. In *ECCV* (pp. 17–35).

[3] Jonathon Luiten, Aljosa Os\uep, Patrick Dendorfer, Philip Torr, Andreas Geiger, Laura Leal-Taixé, & Bastian Leibe (2020). HOTA: A Higher Order Metric for Evaluating Multi-object Tracking. *International Journal of Computer Vision*, 129(2), 548–578.

5. Comparison

Comparison of the metrics obtained with the different trackers on the entire dataset (Tesla T4 GPU).

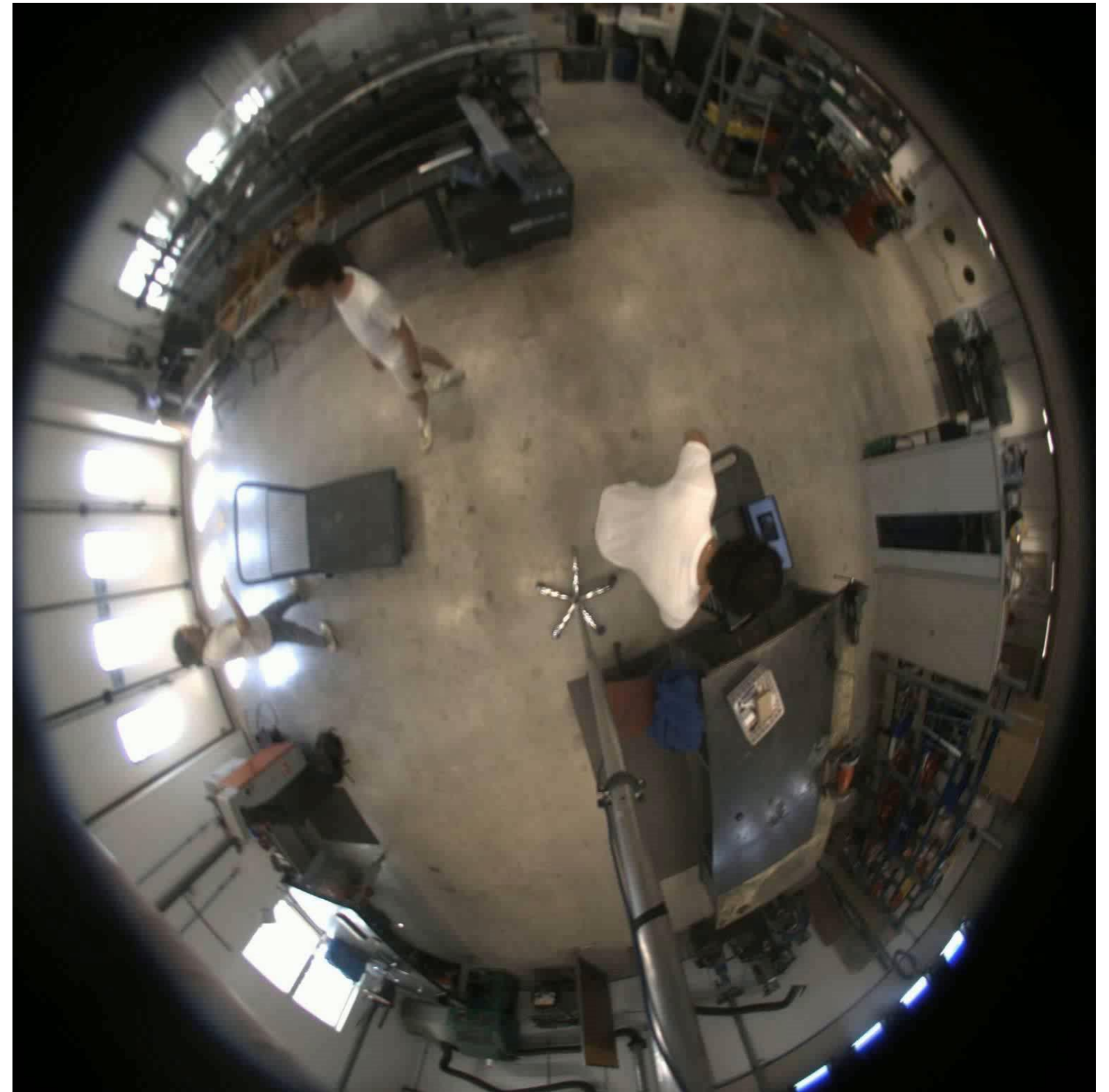
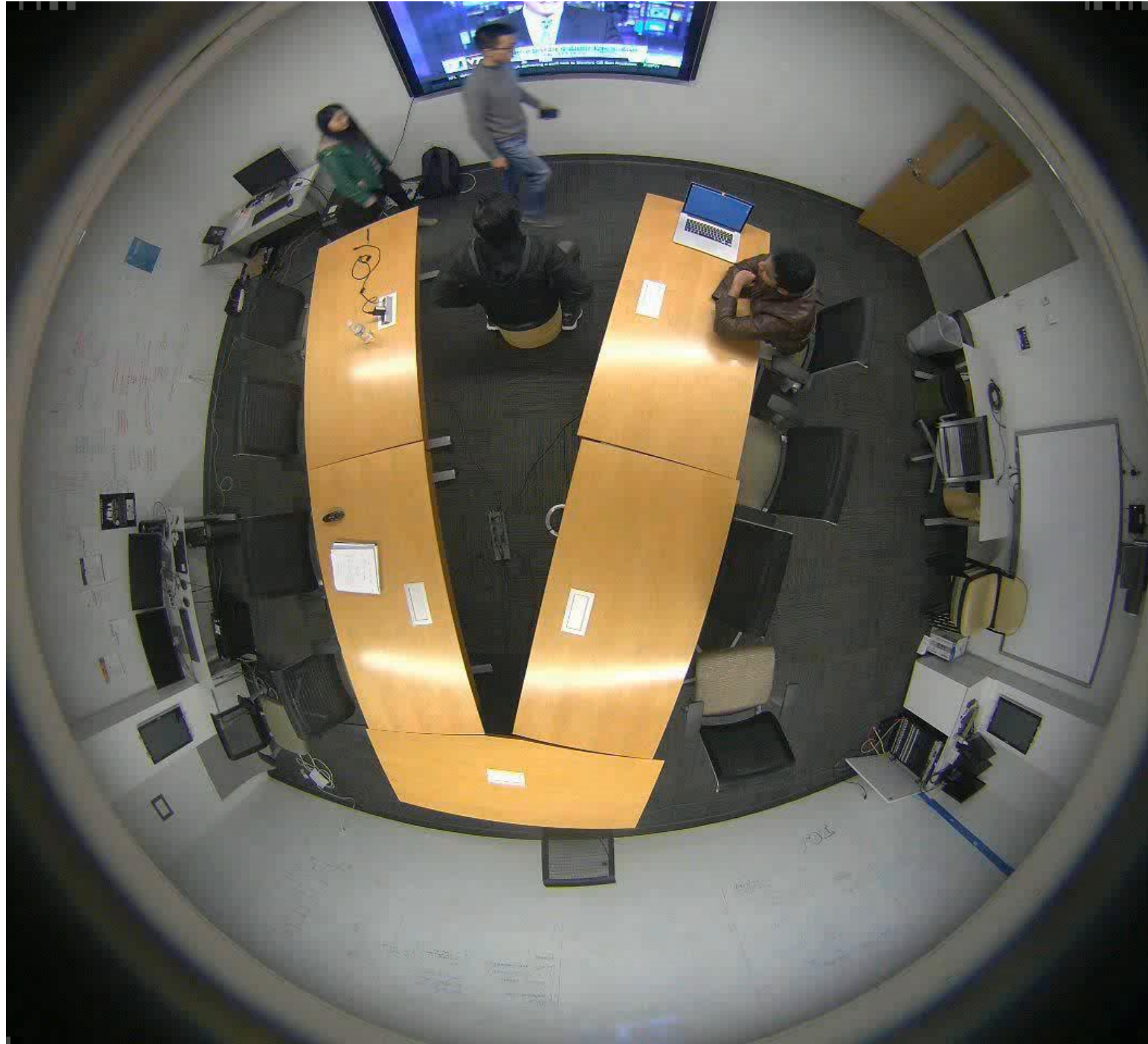
Tracking Algorithm	Metrics			Interference time (ms)
	HOTA	MOTA	IDF1	
StrongSORT	46,2	46,1	64,1	81,3
BoT-SORT	45,6	53,9	66,9	99,1
FORT	47,6	56,9	66,7	52,6

$$\text{MOTA} = 1 - \frac{|\text{FN}| + |\text{FP}| + |\text{IDSW}|}{|\text{gtDet}|} \quad \text{IDF1} = \frac{|\text{IDTP}|}{|\text{IDTP}| + 0.5 |\text{IDFN}| + 0.5 |\text{IDFP}|}$$

Dataset: 20 videos + 9750 frames Source: CERIS Lab + Mirror World

5. Comparison

orange: StrongSORT,
green: BoT-SORT,
blue: FORT (our)



6. Conclusion

Source code :

<https://github.com/BenoitFaureIMT/FORT>

Visual results (20 videos):

<https://partage.imt.fr/index.php/s/kkdckzCJgT9MFDR>

orange: StrongSORT,
green: BoT-SORT,
blue: FORT (our)

