3D Multi-Object Detection And Tracking

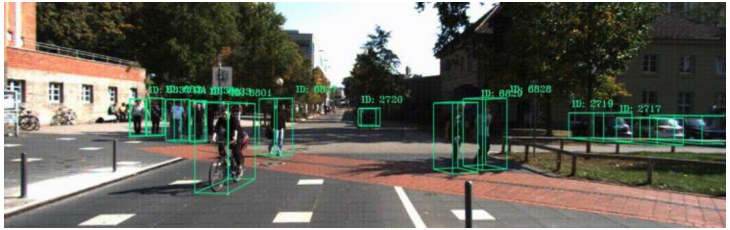
Team - 30

Harnish Dave Ezhilan Veluchami Shiva sam kumar
Mohanraj Babu Yogesh Kumar

3D Multi-object detection

- Multi-object tracking (MOT) monitors multiple moving objects over time using video data.
- The MOT is to detect objects in each frame of the video.
- For applications like autonomous driving or surveillance, MOT needs to be performed in real-time, which demands efficient algorithms to quickly process high volumes of data.

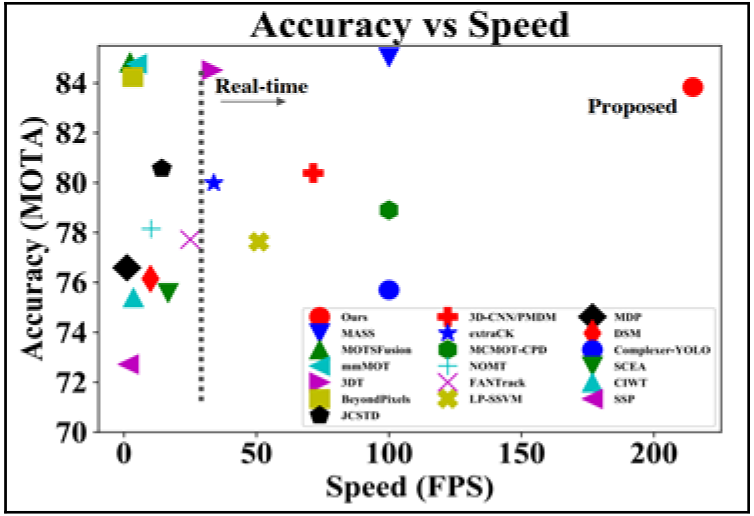
Bounding Box for detected objects in the environment



Related works

Baseline, Similar Papers and improvement on related papers

	Related Papers **FANTrack[2] & mmMOT[3]**	Baseline Paper **AB3DMOT[1]**
Complexity	• Focused on improving accuracy with complexity	• Focused on simpler implementation for real applications
Speed vs Accuracy	• High MOT Accuracy, slower real-time MOT	• Faster real-time 3D MOT for similar accuracy
3D Evaluation	• Projection of 3D on 2D plane compared with 2D ground truth	• 3D IoU cost functions compared with 3D ground truth
MOT Metrics	• Uses MOTA, MOTP metrics	• Integral metrics sAMOTA, AMOTP metrics



[1] X. Weng, J. Wang, D. Held and K. Kitani, "3D Multi-Object Tracking: A Baseline and New Evaluation Metrics," *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, NV, USA, 2020, pp. 10359-10366, doi: 10.1109/IROS45743.2020.9341164.

[2] Erkan Baser, Venkateshwaran Balasubramanian, Prarthana Bhattacharyya, Krzysztof Czarnecki, "FANTrack: 3D Multi-Object Tracking with Feature Association Network," arXiv:1905.02843.

[3] Wenwei Zhang, Hui Zhou, Shuyang Sun, Zhe Wang, Jianping Shi, Chen Change Loy; "Robust Multi-Modality Multi-Object Tracking," Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 2365-2374

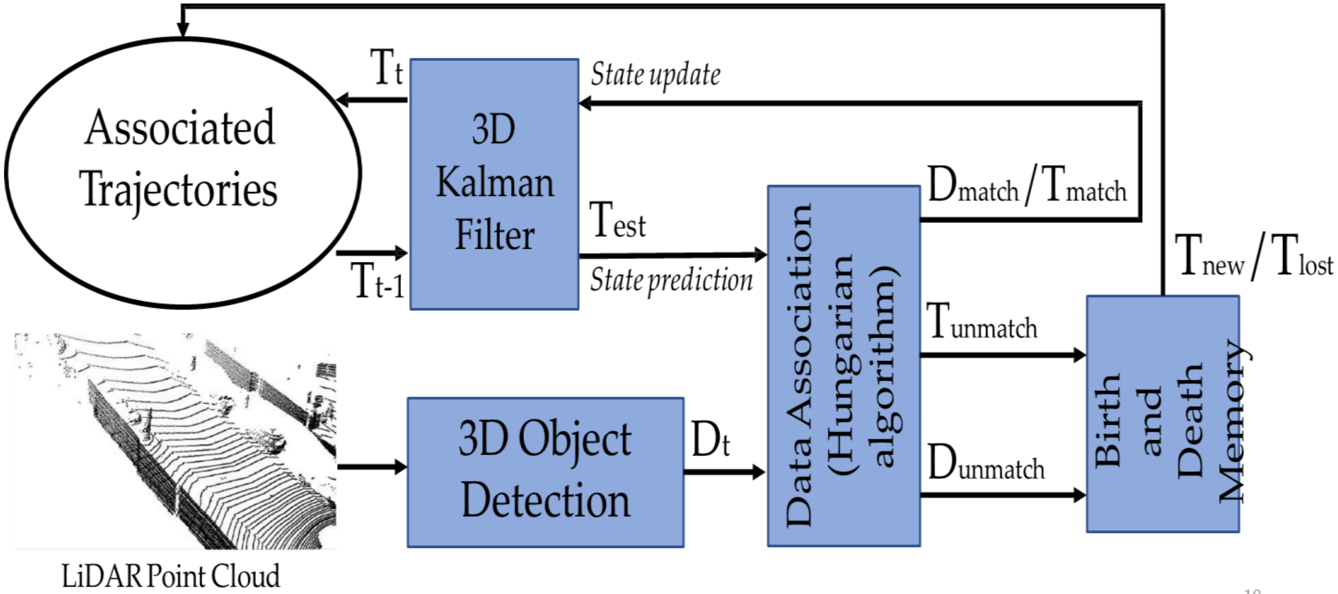
Baseline

Background

3D MOT Pipe Line

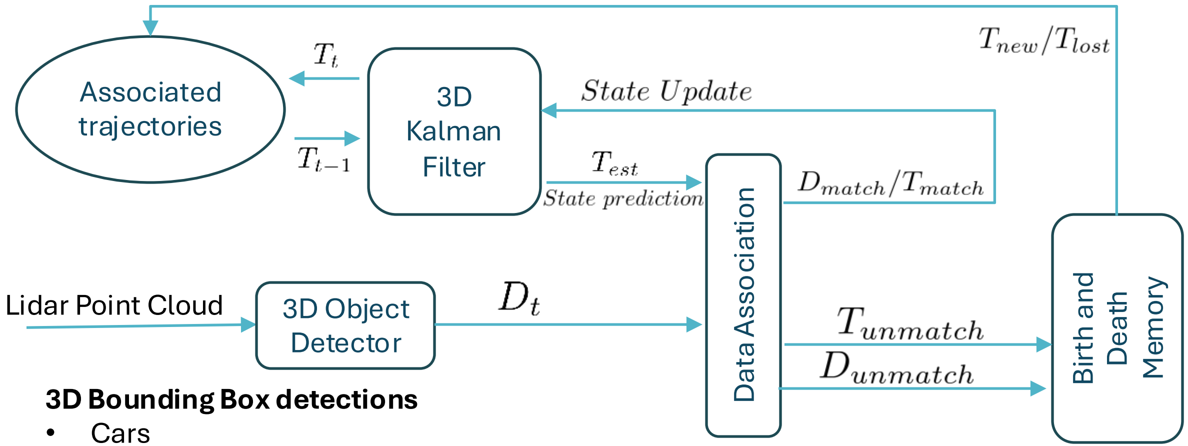
Sensor data
↓
3D object Detection
↓
Data Association
↓
Evaluation

Overview



Methodology

Our approach

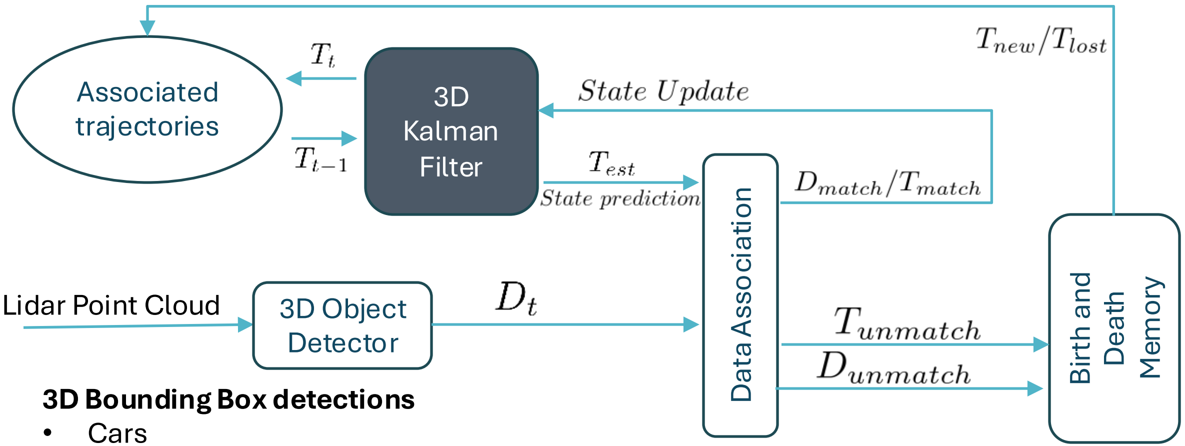


3D Bounding Box detections

- Cars
- Pedestrian
- Bicycle

Methodology

Our approach



3D Bounding Box detections

- Cars
- Pedestrian
- Bicycle

$$[x, y, z, theta, l, w, h, dx, dy, dz]$$

+

$$d\theta \quad / \quad \theta + d\theta$$

Can we use a better prediction model or EKF?

State (dimension 10)

$$[x, y, z, theta, l, w, h, dx, dy, dz]$$

Constant velocity model

$$x = x + dx$$
$$y = y + dy$$
$$z = z + dz$$

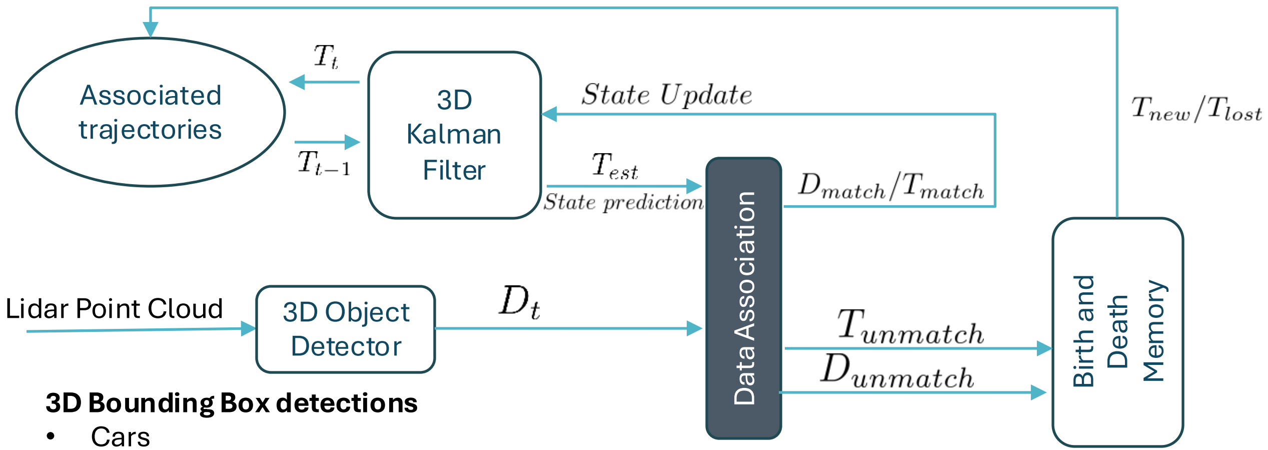
Measured states

$$[\boxed{x, y, z, theta, l, w, h,} dx, dy, dz]$$

- Linear measurement model
- Less noisy

Methodology

Our approach



3D Bounding Box detections

- Cars
- Pedestrian
- Bicycle

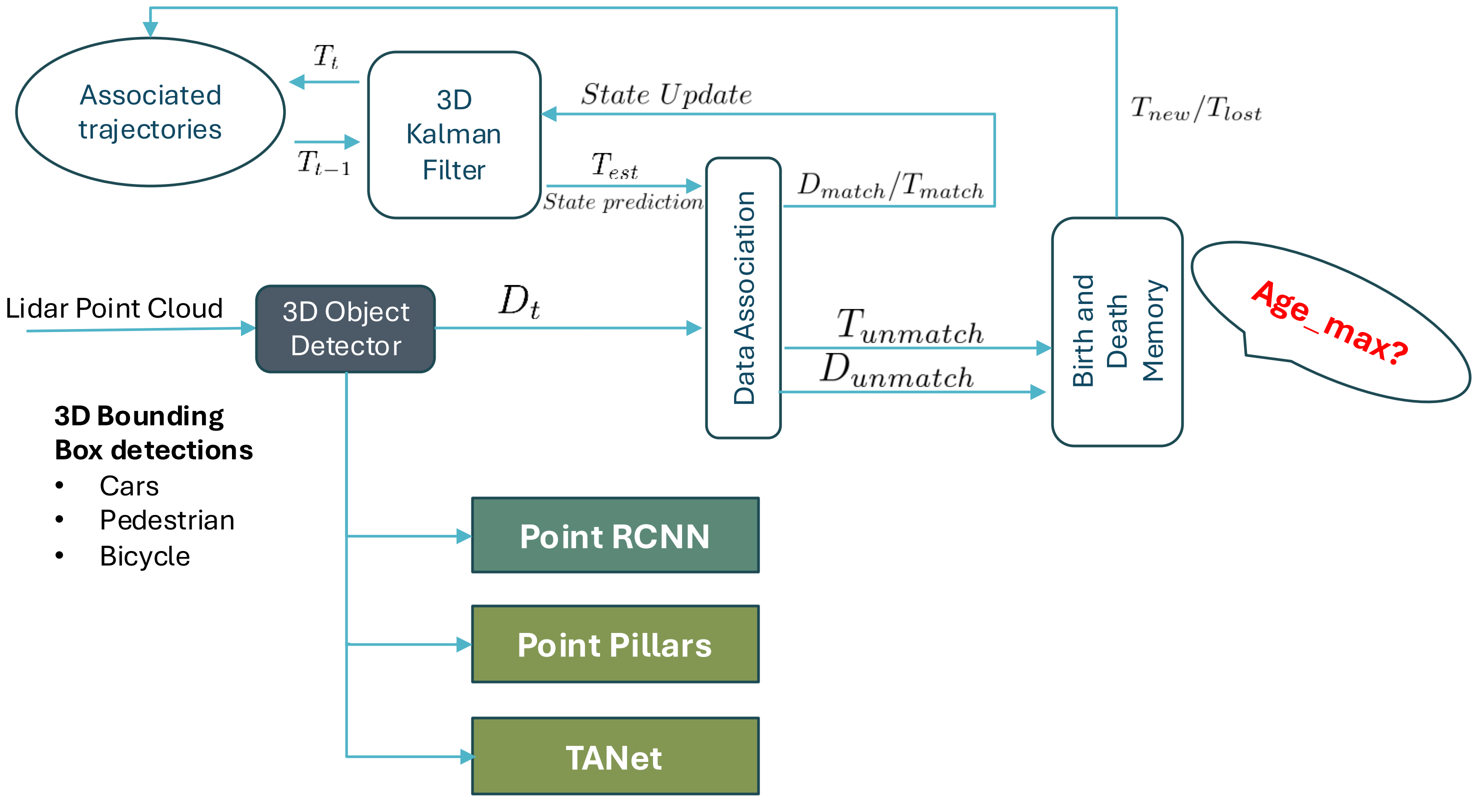
'Hungarian' vs 'greedy'

- Time complexity
- MOTA

- **CAR:** 'hungar'
- **Pedestrian:** 'greedy'
- **Cyclist:** 'hungar'

Methodology

Our approach



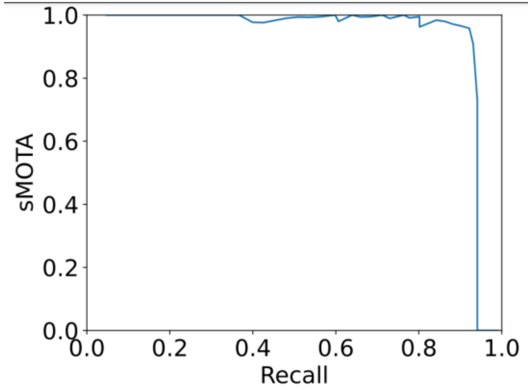
Results

- We ran the tracking algorithm with different 3D object detectors
- PointRCnn is Observed to have better precision of the three model
- PointPillar is Observed to be faster than PointCNN
- TANet has better performance in a clustered environment compared to PointPillars.
- Increasing the age_max parameter of the tracking algorithm resulted in less fragmentation in all the observed cases, but the FPS dropped so we can infer that it takes more time to process

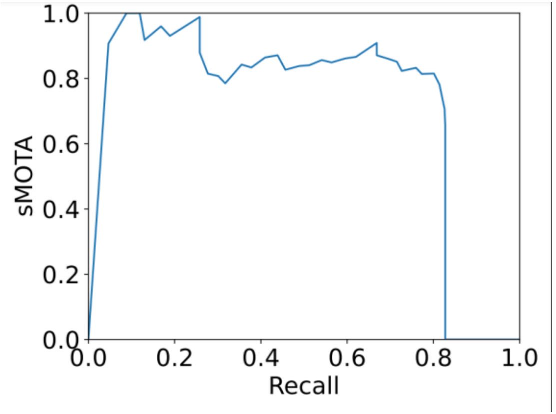
3D Mult object tracking Results
(iou threshold set to 0.7)

3D Object Detector	MOTA	MOTP	sAMOTA
PointRcnn (Baseline)	0.6248	0.8264	0.7496
PointPillars	0.6316	0.7509	0.7063
TANet	0.6418	0.7439	0.7064

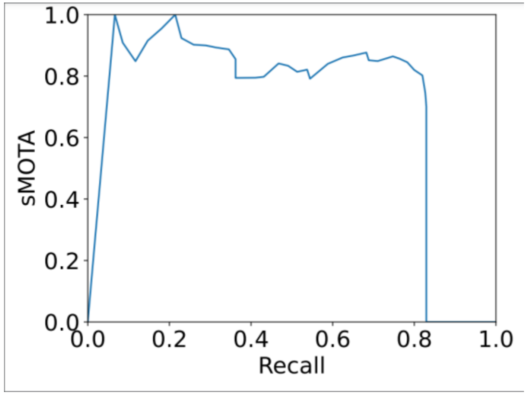
Results



PointRcnn



PointPillar



TANet

Questions???