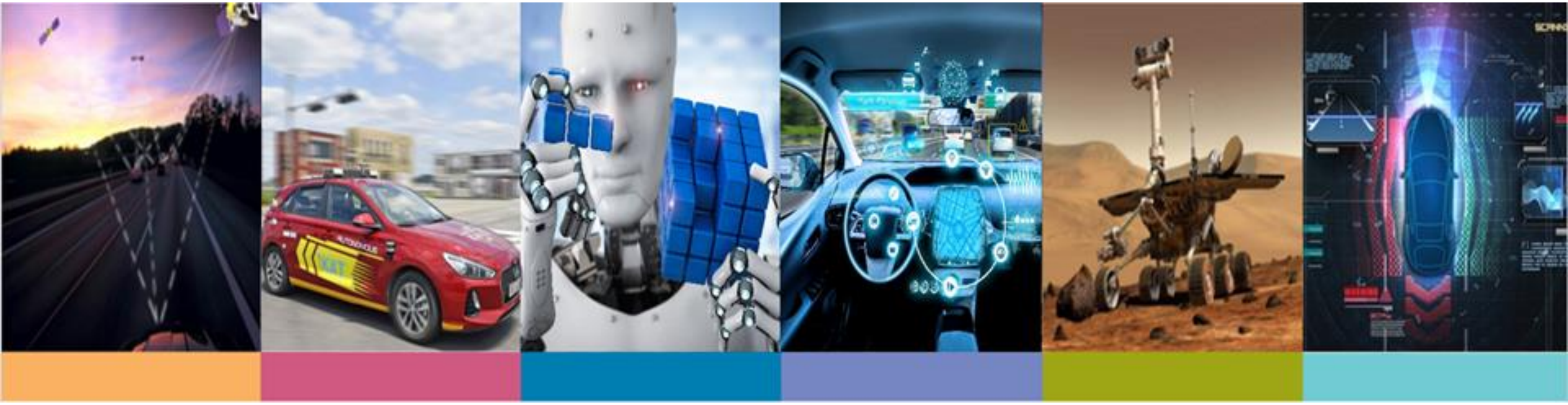


# Accurate 3D Multi Object Tracking for Autonomous Vehicles using Long-term and Short-term Tendency through IMM Kalman Filter

Christofel Rio Goenawan - 20233970

Final Project Presentation
MO560
June 14th 2024

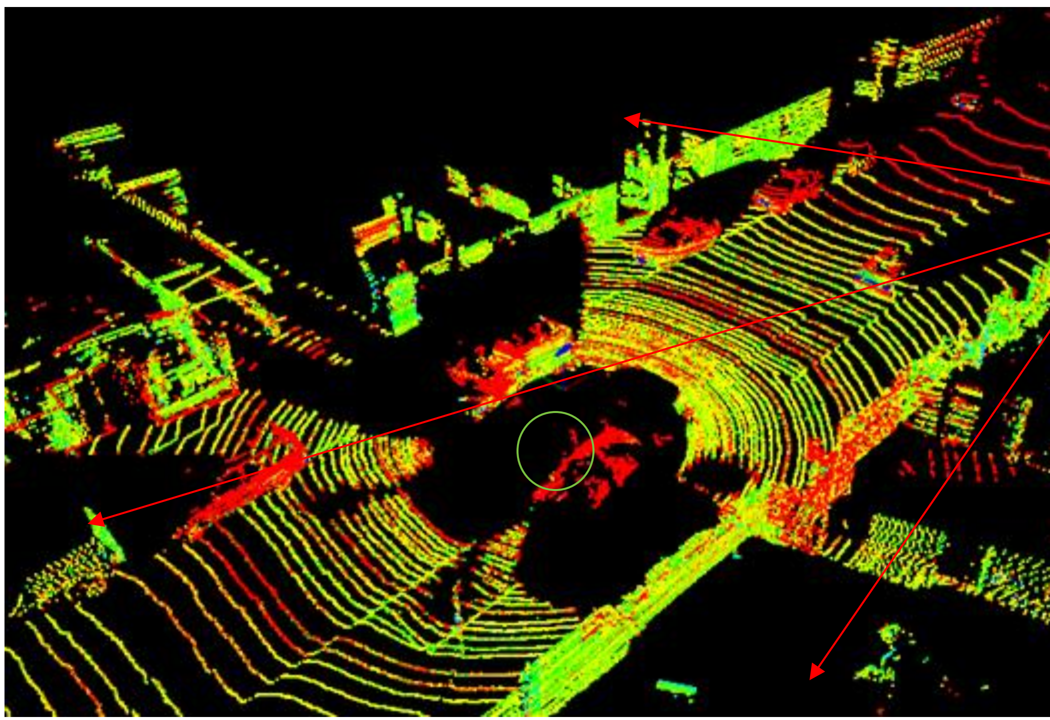
# Contents

# Contents

# What is the Problem ?

**<u>LiDAR Point Cloud</u>**

- LiDAR data is **sparse** and **unordered** [2] → **Inaccurate detection on far points and occlusions**
- Set of all points (x,y,z in sensor frame + intensity)→ LiDAR Point Cloud
- How to represent LiDAR point for 3D object detection? → **Voxel based representation** / **Point based representation**



Sparse : Lot of area doesn't contain points

: LiDAR sensor

Figure 3. Visualized LiDAR Point Cloud. Source : Talha Muhammad, Revolutionizing Autonomous Vehicle : Unveiling the Power of Point Clouds. LinkedIn. 2023.

*[2] Qian, Rui et al. 2022. 3D Object Detection for Autonomous Driving: A Survey . Elsevier Pattern Recognition (2022).*

# What is the Problem : Voxel- based 3D Object Detection

**Voxel- based Representation**

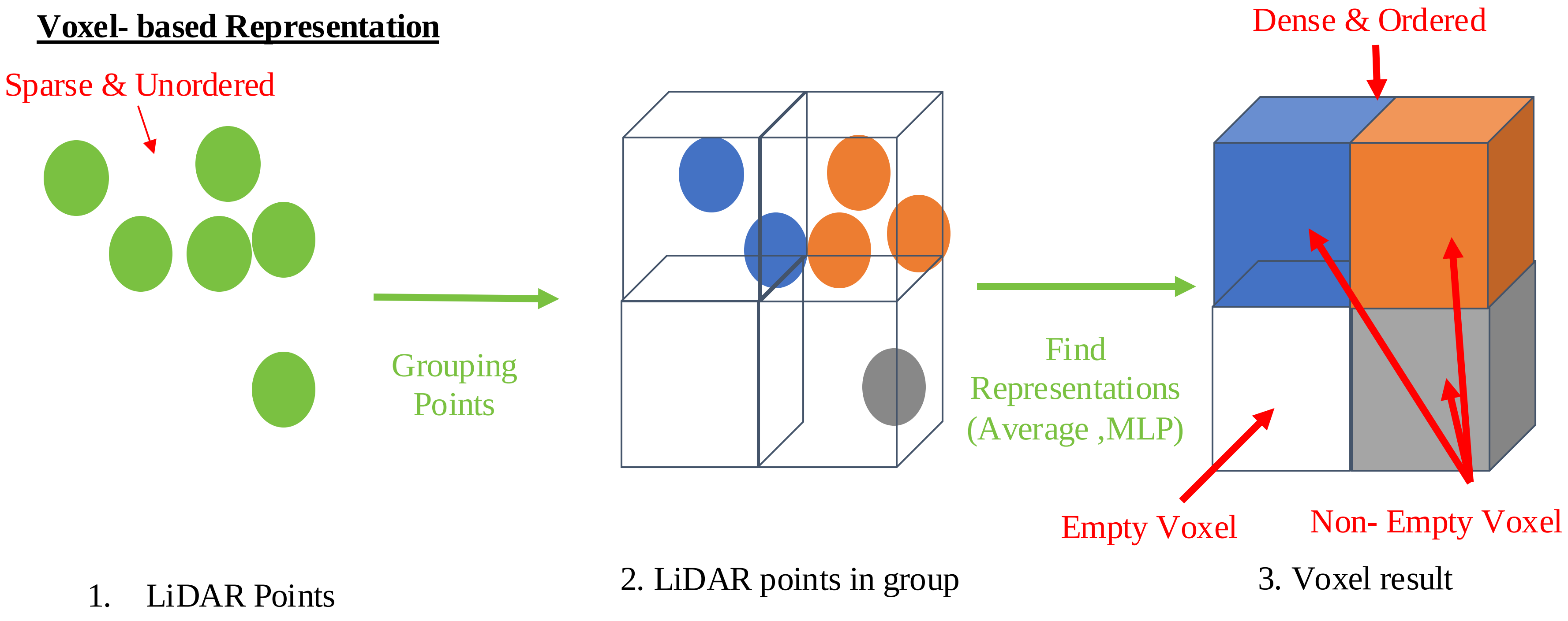



Figure 4. Process of making LiDAR point cloud into Voxel.

# What is the Problem : Point- based 3D Object Detection

**Point- Based Representation**

- Process directly as **unordered** and **sparse** point set using Set Abstraction [3]

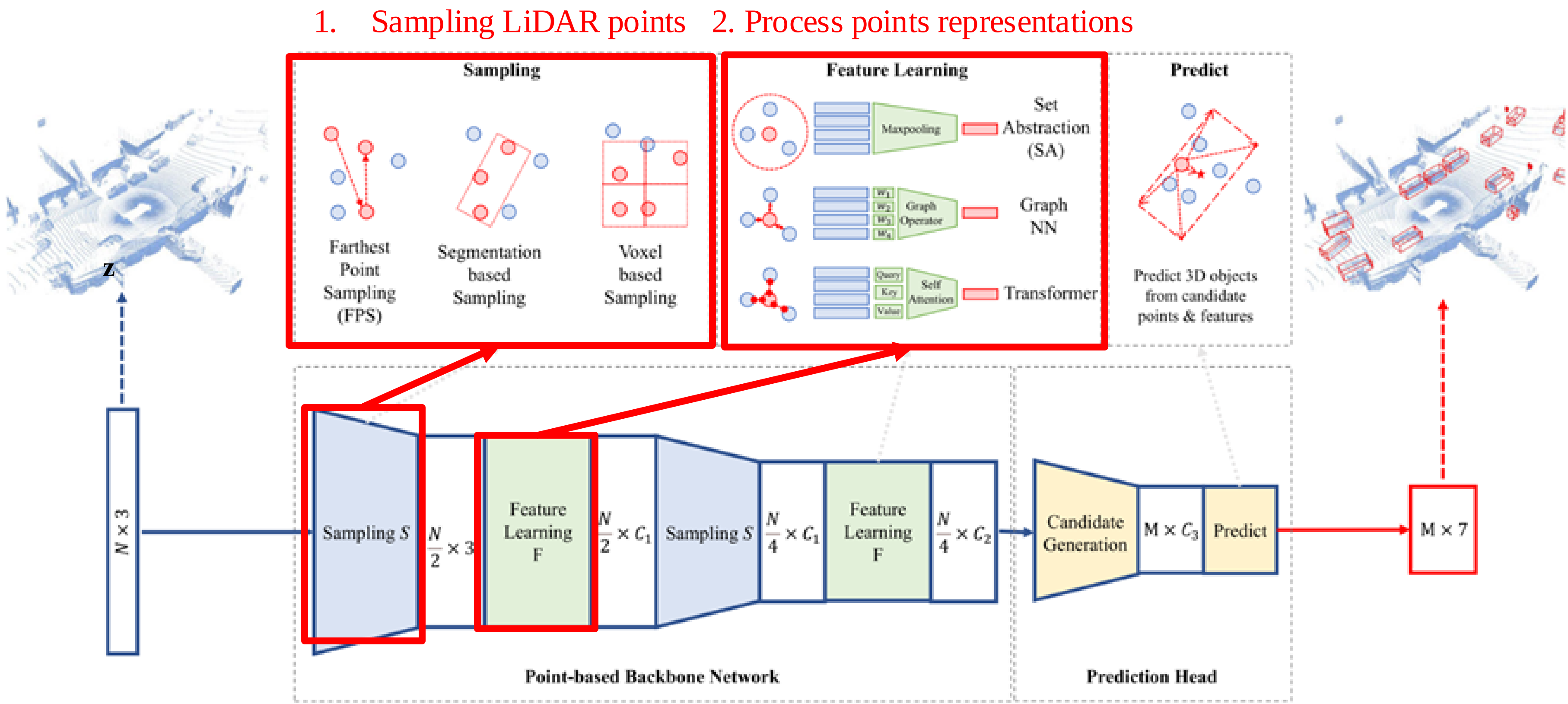



Figure 5. Object Detection process using Point- Based Representation. Source : [2]

*[3] Qi, Charles R., et al. "Pointnet: Deep learning on point sets for 3d classification and segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.*

# Review : Voxel vs Point Based

**Voxel- Based Representation**

+ Capture global correlation of 3D Objects better

- Lack of detailed geometric and semantic information [3]

**Point-Based Representation**

+ Contain fined- grain location and semantic information

- Doesn't perform well in sparse and occluded points [4]
- Not robust to outlier points

*We can combine prediction of Voxel- Based and Point- Based for better prediction & tracking*

# Multi Model Kalman Filter

**Since the training process is long and I didn't found point-based model with good performance then**

**change the Final Project Topic**

[4] Qian, Rui, Xin Lai, and Xirong Li. "3D object detection for autonomous driving: A survey." Pattern Recognition 130 (2022): 108796.

# Contents

# What is the Problem ?

**<u>3D Object Detection & Tracking for Autonomous Vehicle</u>**

- 3D object detection + tracking is important task in Autonomous Vehicle
- 3D object detection in autonomous vehicle processed LiDAR point cloud using deep learning
- Currently SOTA of 3D multi object tracking using **tracking by detection method [1]**

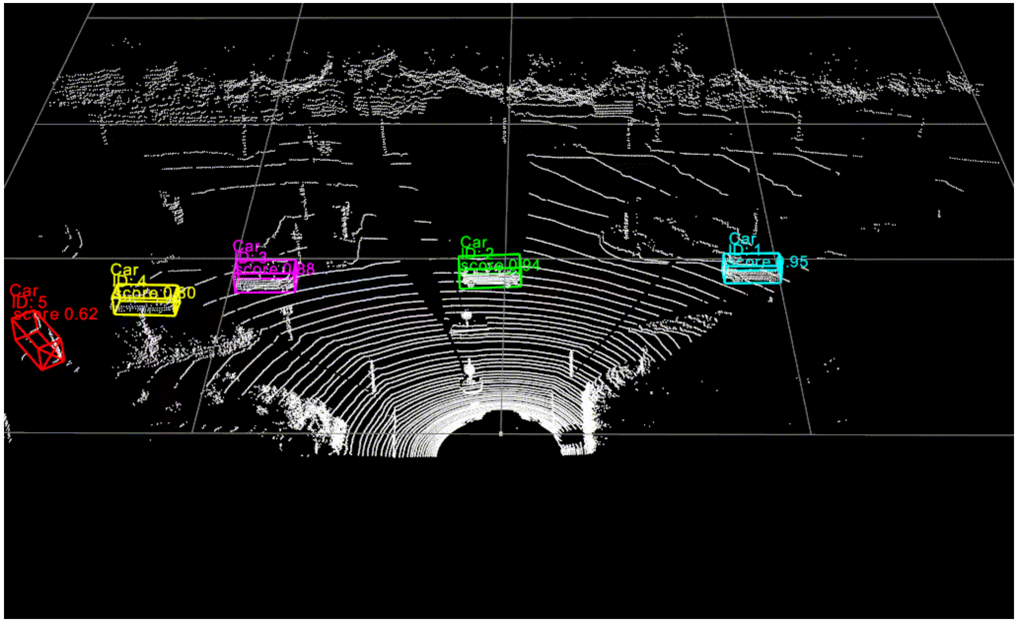



Figure 1. Result of 3D Multi Object Tracking is Tracking ID, Position, Size and Orientation of Objects

*[1]* Zhang, Peng, et al. "3D Multiple Object Tracking on Autonomous Driving: A Literature Review." arXiv preprint arXiv:2309.15411 (2023).

# What is the Problem ?

**<u>3D Object Detection & Tracking for Autonomous Vehicle</u>**

- 3D object detection + tracking is important task in Autonomous Vehicle
- 3D object detection in autonomous vehicle processed LiDAR point cloud using deep learning
- Currently SOTA of 3D multi object tracking using **tracking by detection method [1]**

Usually update final prediction and predict position next step using Kalman Filter

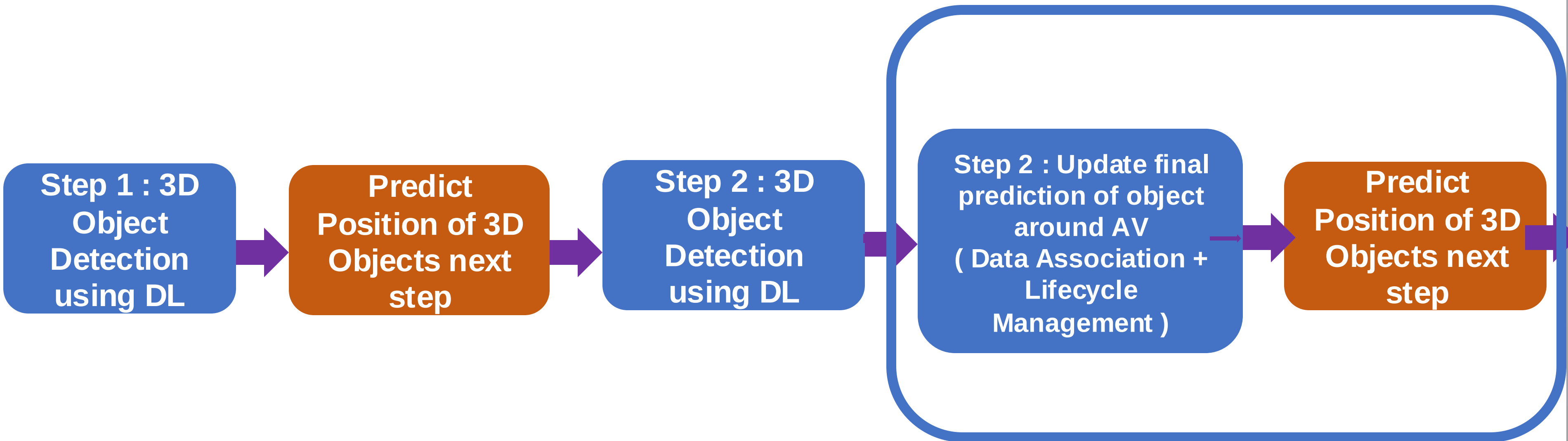
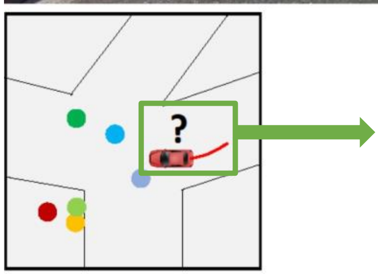



Figure 2. Process of 3D Multi Object Tracking by using Tracking by Detection Method

*[1]* Zhang, Peng, et al. "3D Multiple Object Tracking on Autonomous Driving: A Literature Review." arXiv preprint arXiv:2309.15411 (2023).

# What is the Problem ?

- But in complex situation **objects will move in different movement**, ex: in intersection, crash etc hence Kalman Filter will predict object next state in wrong state
- Predict wrong state will decrease tracking localization accuracy and wrong object recognition


(a)

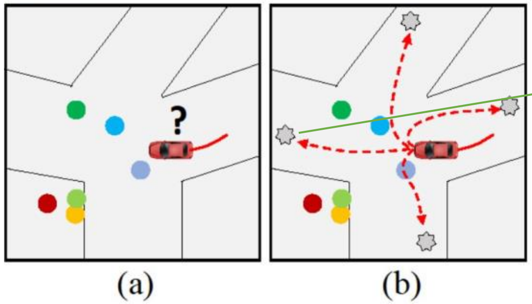
Because red car driver  have cycle in front of him he will stop in intersection

Figure 3. Diagram of Car Movement in Intersection


*[1]* Zhang, Peng, et al. "3D Multiple Object Tracking on Autonomous Driving: A Literature Review." arXiv preprint arXiv:2309.15411 (2023).

# What is the Problem ?

- But in complex situation **objects will move in different movement**, ex: in intersection, crash etc hence Kalman Filter will predict object next state in wrong state
- Predict wrong state will decrease tracking localization accuracy and wrong object recognition


If we see the red car movement in long term we know he will go to left side of intersection

Figure 3. Process of 3D Multi Object Tracking by using Tracking by Detection Method

**How if we combine prediction of short-term tracking and long-term tracking to have accurate 3D Multi Object Tracking using Multi Model Kalman Filter ?**

*[1]* Zhang, Peng, et al. "3D Multiple Object Tracking on Autonomous Driving: A Literature Review." arXiv preprint arXiv:2309.15411 (2023).

# Contents

# What is the Problem ?

**State of the Art Method**

| Publication | Method |
| --- | --- |
| Cho, M.; Kim, E.<br>3D LiDAR Multi-Object Tracking with Short-Term and Long-Term Multi-Level Associations.<br>Remote Sens. 2023, 15, 5486. | Recognize object using long- term object and short-term object |
| Qin, Zheng, et al.<br>"Motiontrack: Learning robust short-term and long-term motions for multi-object tracking."<br>Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023. | Proposed End- to end object detection and tracking in video using long-term attention and short-term attention |
| Emami, A., Sarvi, M. & Asadi Bagloee, S.<br>Using Kalman filter algorithm for short-term traffic flow prediction in a connected vehicle environment. J. Mod. Transport. 27, 222–232 (2019). https://doi.org/10.1007/s40534-019-0193-2 | Proposed multi- object tracking using short- term model |

Figure 2. Process of 3D Multi Object Tracking by using Tracking by Detection Method

**Hence there is no state of the art method that combine Short- Term prediction and Long-Term prediction for 3D Multi Object Tracking using Multi Model Kalman Filter**

*[1]* Zhang, Peng, et al. "3D Multiple Object Tracking on Autonomous Driving: A Literature Review." arXiv preprint arXiv:2309.15411 (2023).

# Contents

# Proposed Method

**Accurate 3D Multi Object Tracking using Long- Term and Short- Term Tendency using Multi Model Kalman Filter**

- Proposed predicting 3D object using Deep Learning then predict trajectory using Long- Term and Short- Term Tendency using Multi Model Kalman Filter

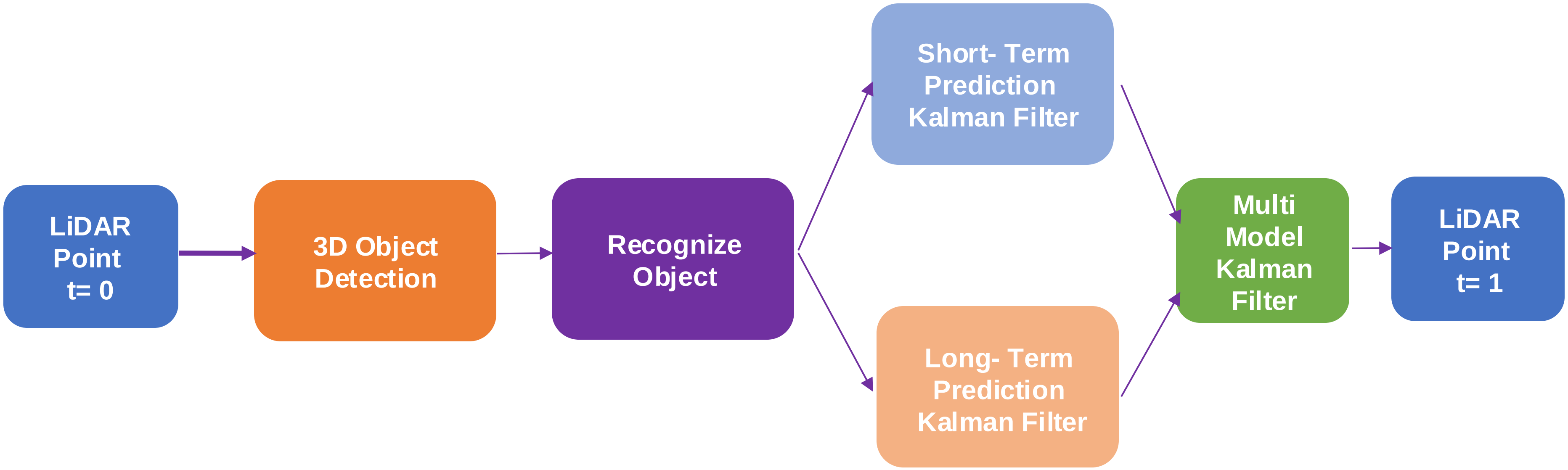
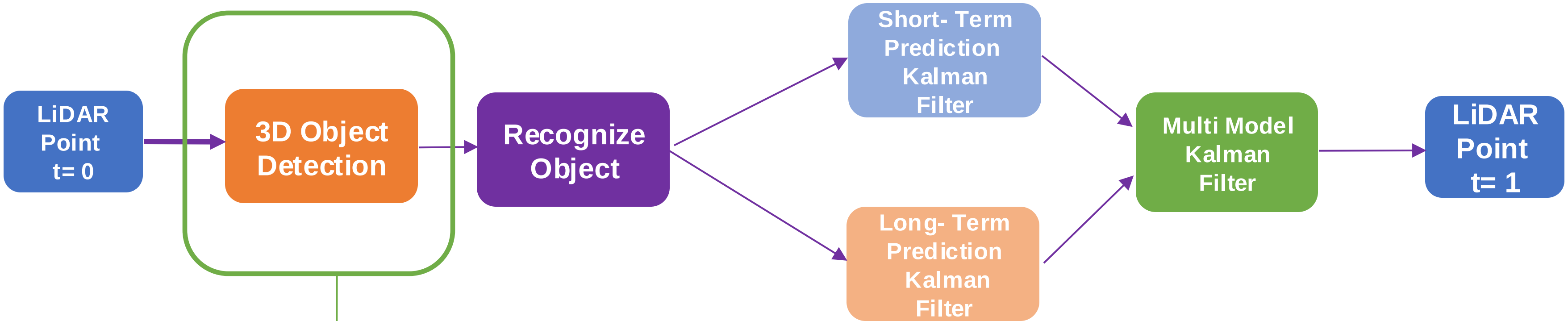



Figure 6. Process of Point- Voxel RCNN 3D Object Detection.


*[7] Mao, Jiageng, et al. "Voxel transformer for 3d object detection." Proceedings of the IEEE/CVF international conference on computer vision. 2021.*
*[8] Pan, Xuran, et al. "3d object detection with pointformer." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021.*

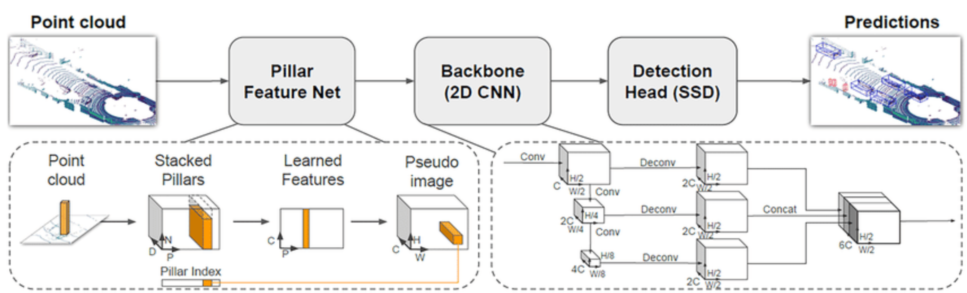
# Proposed Method

**Accurate 3D Multi Object Tracking using Long- Term and Short- Term Tendency using Multi Model Kalman Filter**




**3D Object Detection**

- Using state of the Art **Point Pillar** to detect cars object
- Predict location, size, orientation and confidence score of bounding box



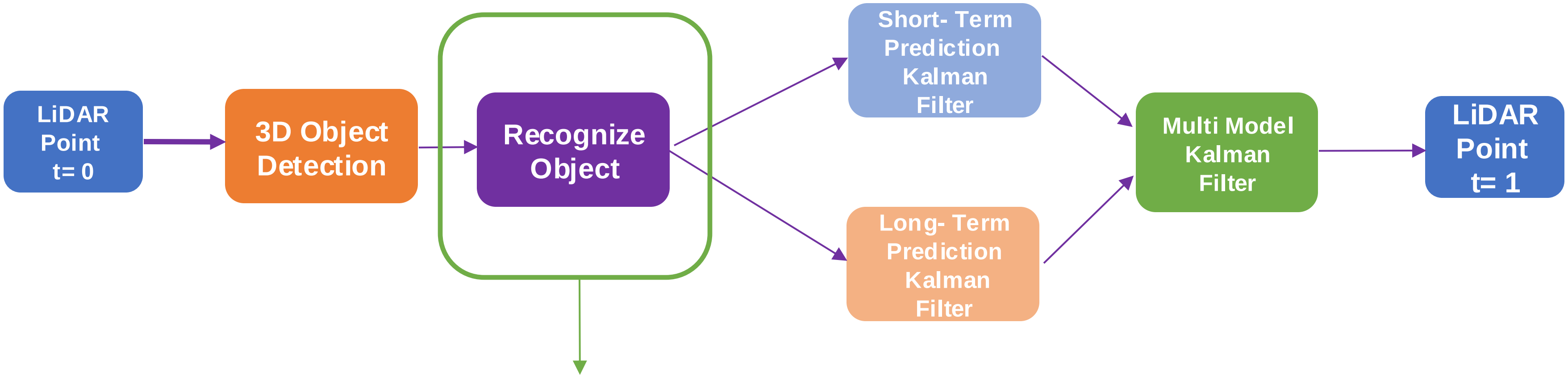


*[7] Mao, Jiageng, et al. "Voxel transformer for 3d object detection." Proceedings of the IEEE/CVF international conference on computer vision. 2021.*

*[8] Pan, Xuran, et al. "3d object detection with pointformer." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021.*

# Proposed Method

**Accurate 3D Multi Object Tracking using Long- Term and Short- Term Tendency using Multi Model Kalman Filter**



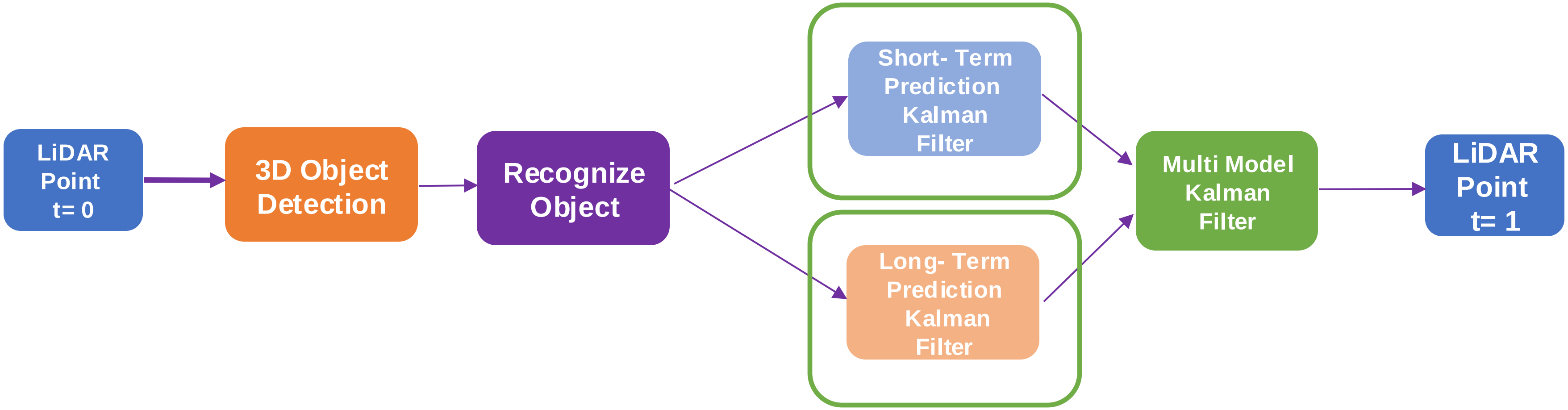

**Recognize Object (Data Association)**

- Using **Hungarian Algorithm** to recognize object detected by Point Pillar using predicted object state
- Using proposed **confidence-score aware cost function** IOU of Bounding Box weighted by detection confidence score

*Hungarian Algorithm Cost Function* = ***IoU of Detected Bounding Box and Predicted Bounding Box * 1/ Detected Bounding Box Confidence Score***

# Proposed Method

**Accurate 3D Multi Object Tracking using Long- Term and Short- Term Tendency using Multi Model Kalman Filter**




**Short- Term Prediction Kalman Filter**

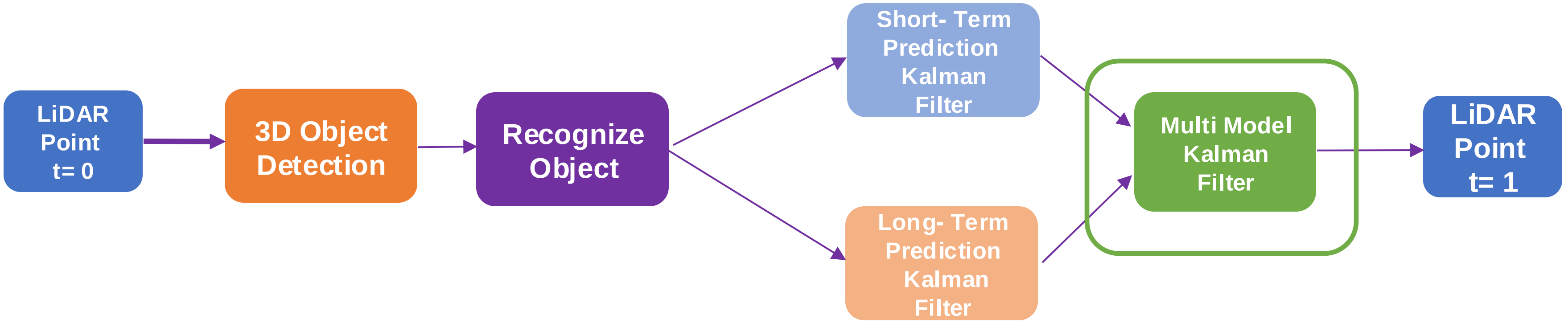
- Using Uscented Kalman Filter (UKF) to tracking object (position x , y , and orientation θ based on the ID
- Update state using **previous prediction *(t-1)***

**Long- Term Prediction Kalman Filter**

- Using Uscented Kalman Filter (UKF) to tracking object (position x, y , and orientation θ ) based on the ID
- Update state using **long-time prediction *(t-8)***

# Proposed Method

**Accurate 3D Multi Object Tracking using Long- Term and Short- Term Tendency using Multi Model Kalman Filter**




**Multi Model Kalman Filter**

- Calculate probability of Short- Term Kalman Filter and Long- Term Kalman Filter based on the detected object

$$\mathrm{pdf}(y_k|p_j) \approx \frac{\exp(-r_k^T S_k^{-1} r_k/2)}{(2\pi)^{q/2}|S_k|^{1/2}}$$

- Update weighted average of Short- Term and Long- Term Kalman Filter based on detected object

$$\Pr(p_j|y_{k-1}) = \frac{\Pr(y_{k-1}|p_j)\Pr(p_j)}{\Pr(y_{k-1})}$$

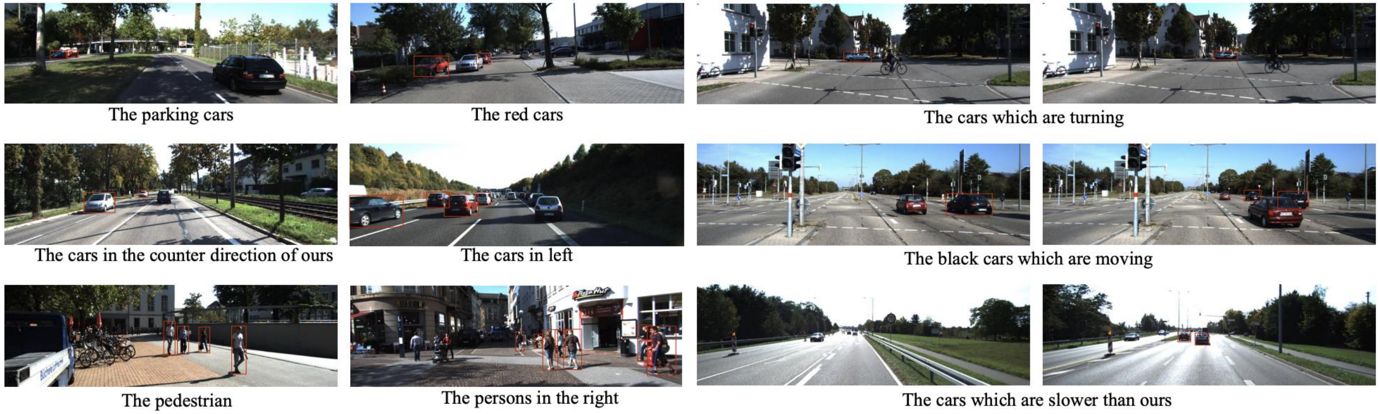
- Predict current object location and orientation using average Short- Term and Long- Term Kalman Filter

# Contents

# Experiment Setting

## Experiment Setting

- Using **KITTI 3D Object Tracking** for 3D Object Detection and Multi Object Tracking
- Consisted of more than 18 scenes with 50 tracked object per scenes



The parking cars | The red cars | The cars which are turning
The cars in the counter direction of ours | The cars in left | The black cars which are moving
The pedestrian | The persons in the right | The cars which are slower than ours

## Evaluating Metrics

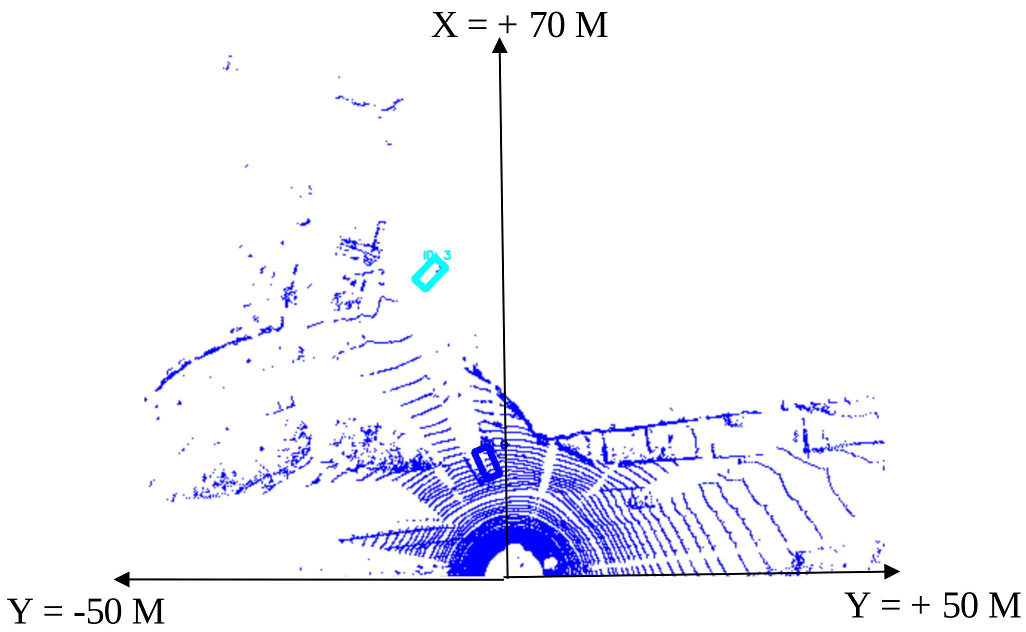
- Using current evaluating metrics for 3D object tracking :

**Average Multi Object Tracking Accuracy (AMOTA), False Positive, False Negative , and ID Switchs**

# Experiment Setting

**Experiment Parameter**

- Using pretrained **PointPillar 3D Object Detection** with Accuracy Detection 0.78 mAP on KITTI Dataset
- Using LiDAR point cloud with area [min x , min y , max x , max y] = ***[0 m , -50 m , +70 m , -50 m*]**
- Using initial probability of Short- term Kalman Filter and Long- Term Kalman Filter of **0.5** and **0.5**
- Detect and Tracking object class **"Car"** and **"Van"** only



# Contents

# Experiment Result- Quantitative

**Experiment Result**

| Tracking Model | AMOTA ↓ | False Positive↓ | False Negative ↓ | ID Switch ↓ |
|---|---|---|---|---|
| Short- Term Kalman Filter | 0.4211 | 10 % | 4.2 % | 128 |
| Long- Term Kalman Filter | 0.2137 | 25 % | 10.3 % | 299 |
| Long- Term and Short-Term using Multi Model Kalman Filter (Ours) | **0.4832** | **6 %** | **8.23 %** | **93** |

**Metrics with symbol ↓ means lower metrics is better*

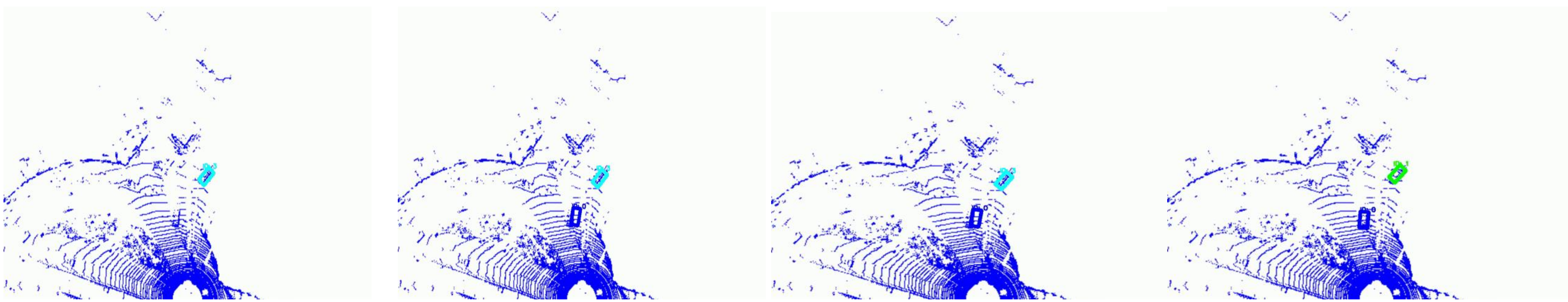
Figure 1. Result of 3D Multi Object Tracking is Tracking ID, Position, Size and Orientation of Objects

*[1]* Zhang, Peng, et al. "3D Multiple Object Tracking on Autonomous Driving: A Literature Review." arXiv preprint arXiv:2309.15411 (2023).

# Experiment Result- Qualitative

**Experiment Result**

Detection projected to Camera Image



Ground Truth

**Long- Term and Short- Term Multi Model KF (Ours)**

**Short- Term KF Model**

**Long- Term KF Model**

Figure 1. Result of 3D Multi Object Tracking is Tracking ID, Position, Size and Orientation of Objects

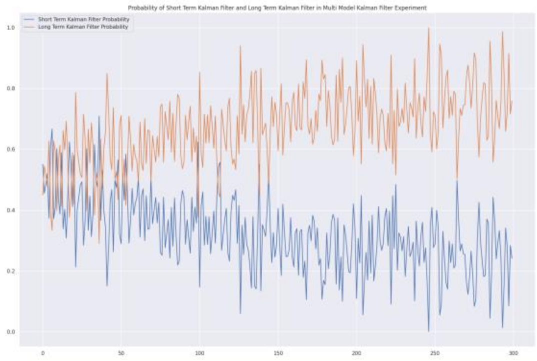
*[1]* Zhang, Peng, et al. "3D Multiple Object Tracking on Autonomous Driving: A Literature Review." arXiv preprint arXiv:2309.15411 (2023).

# Experiment Result- Analysis per Case

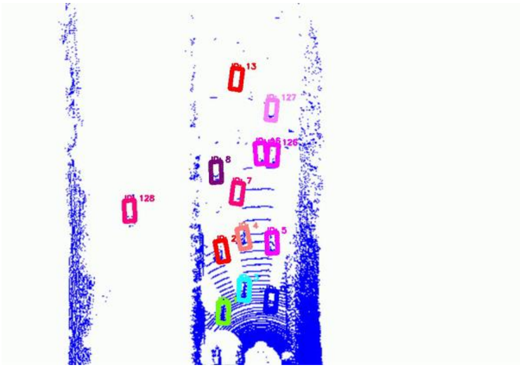
**Analysis per Case**

Detection projected to Camera Image

## Complex Scenario (U-Turn)




## Simple Scenario (Straigth Way)





*[1]* Zhang, Peng, et al. "3D Multiple Object Tracking on Autonomous Driving: A Literature Review." arXiv preprint arXiv:2309.15411 (2023).

# Contents

# Conclusion

**Conclusion :**

- We proposed using Long- Term Tendency and Short- Term Tendency for Accurate 3D Multi Object Tracking through Multi- Model Kalman Filter
- Proposed Long- term and Short-
- We will evaluate the Point-Voxel 3D object detection and MOT using Kalman Filter using Argoverse 1 Dataset
- Proposed Point-Voxel 3D object detection and tracking will have faster prediction and more explainable Point- based and Voxel- based 3D object detection & MOT combination than deep learning

# Thank You for Your Time

Do You Have Any Questions ?