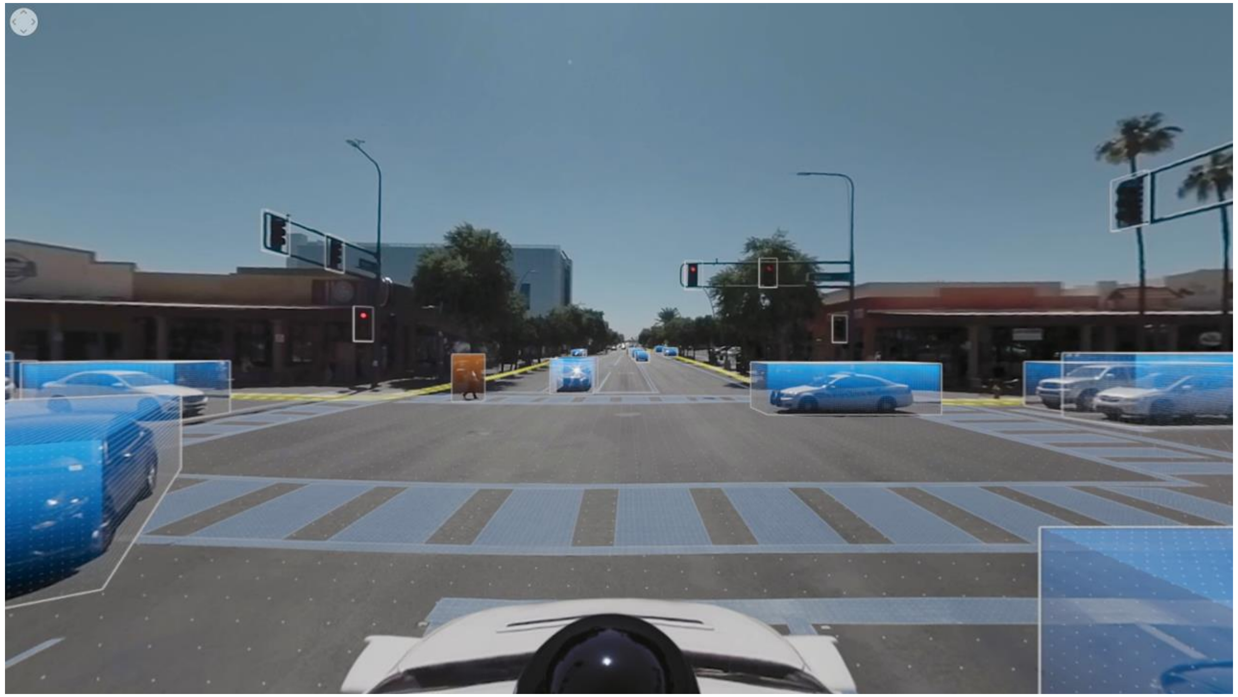
Instance Completion and Motion Estimation with Deep Shape Priors for Autonomous Driving

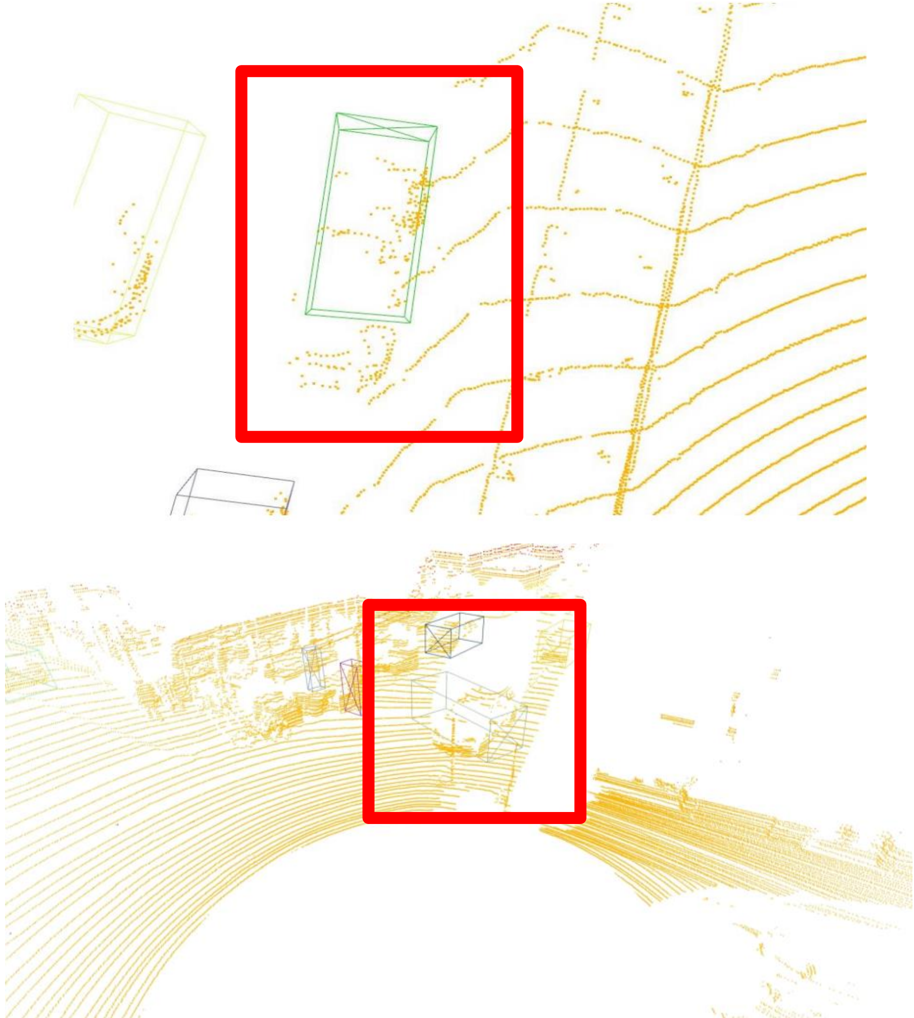
Team members
Panyawat (Ohm) Rattana
Shashank (Sai) Dammalapati

Supervisors
Xingguang (Starry) Zhong
Yue Pan

Motivation



[Courtesy of Waymo]



Incorrect Bounding Box Predictions

Motivation – Shape Priors

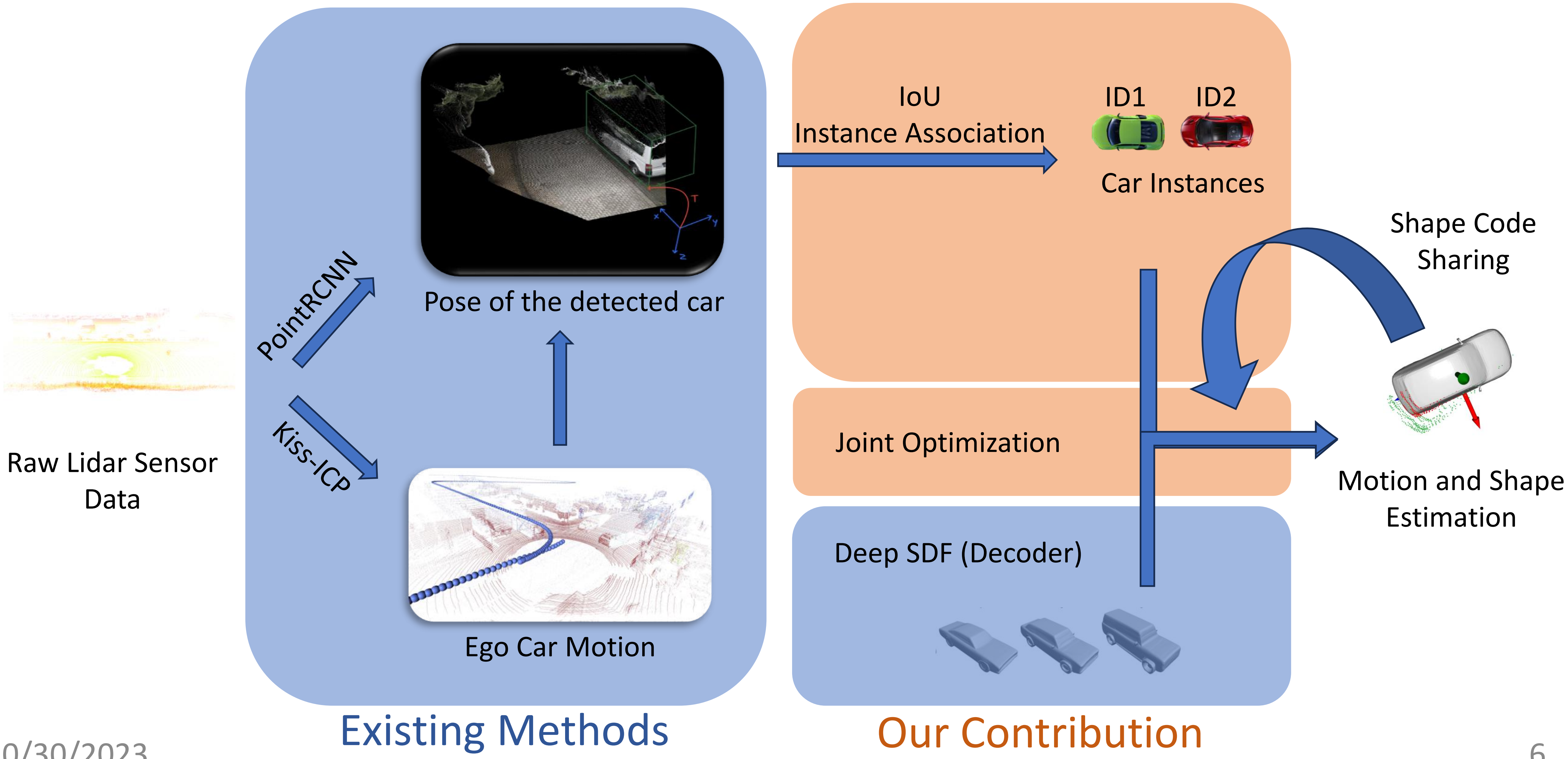
VW Up Car

Motivation – Shape Priors

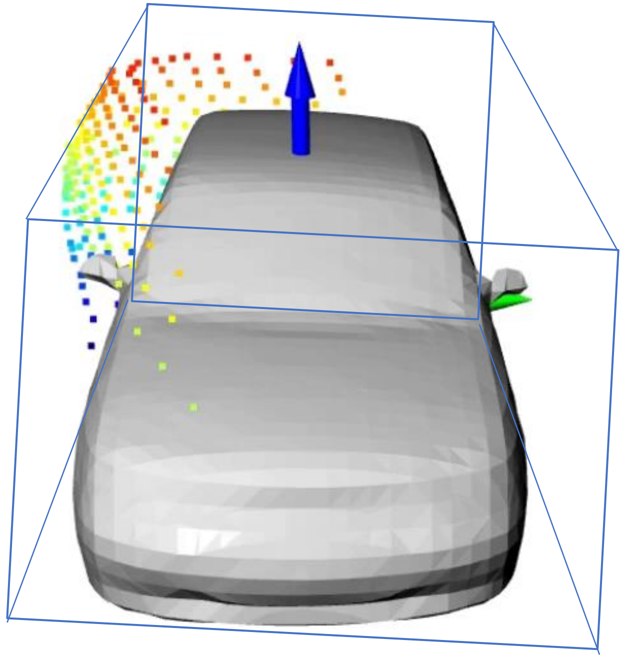
Car in this scene is VW UP

Prior Work

Pipeline

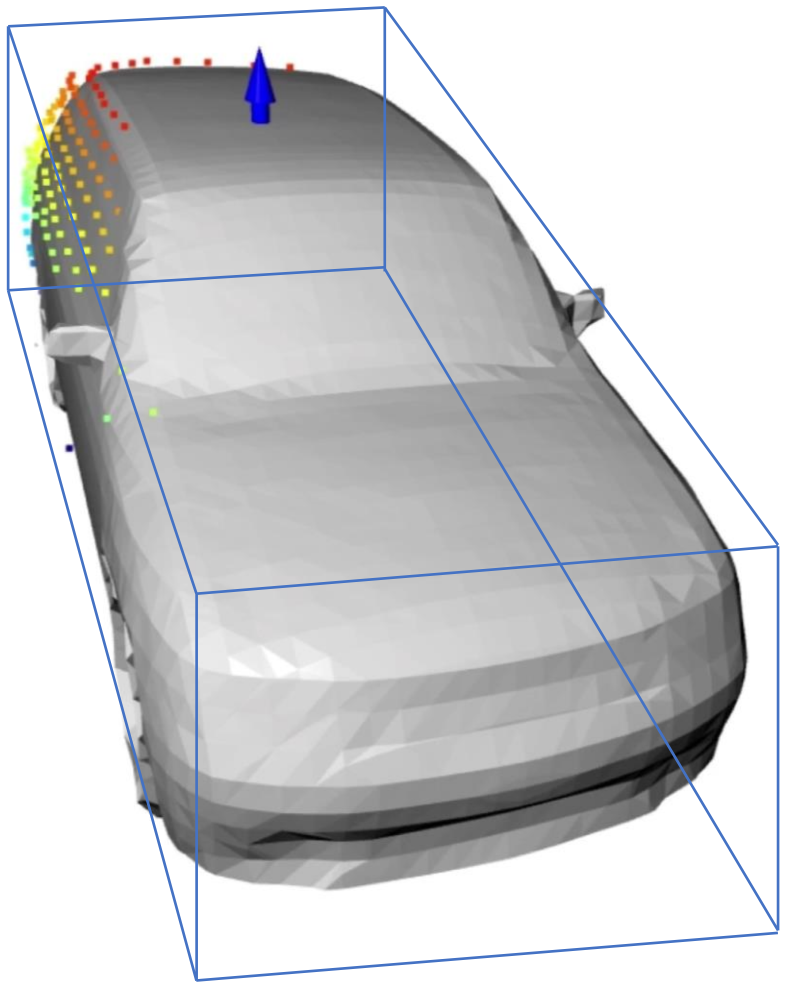

Raw Lidar Sensor Data
PointRCNN
Kiss-ICP
Pose of the detected car
Ego Car Motion
Existing Methods
IoU
Instance Association
ID1
ID2
Car Instances
Shape Code Sharing
Joint Optimization
Deep SDF (Decoder)
Motion and Shape Estimation
Our Contribution

Result - Joint Optimization

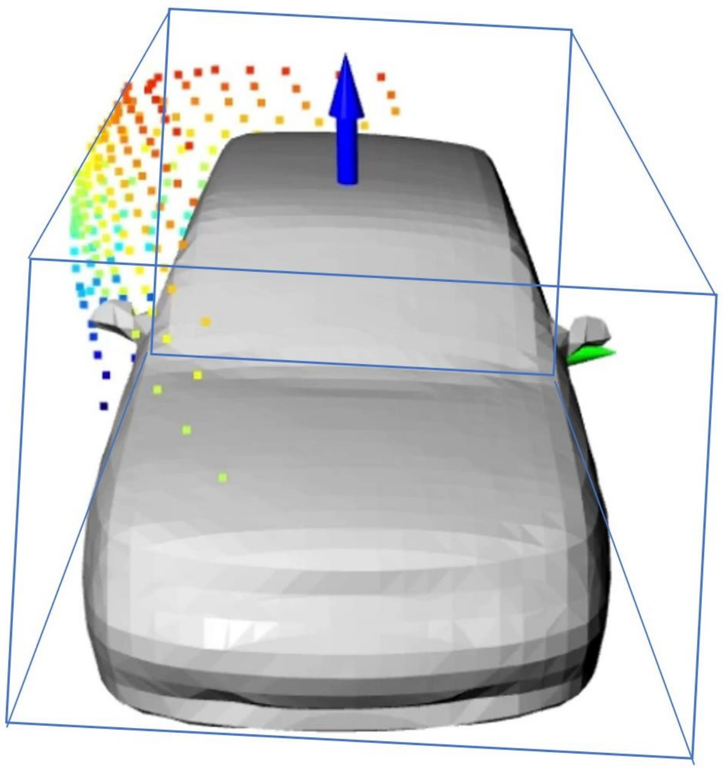


Visualization shows
1. Point Cloud – belonging to a detected car surface
2. Bounding Box is initialized with PointRCNN
3. Over the process of optimization we get improved bounding box and shape code

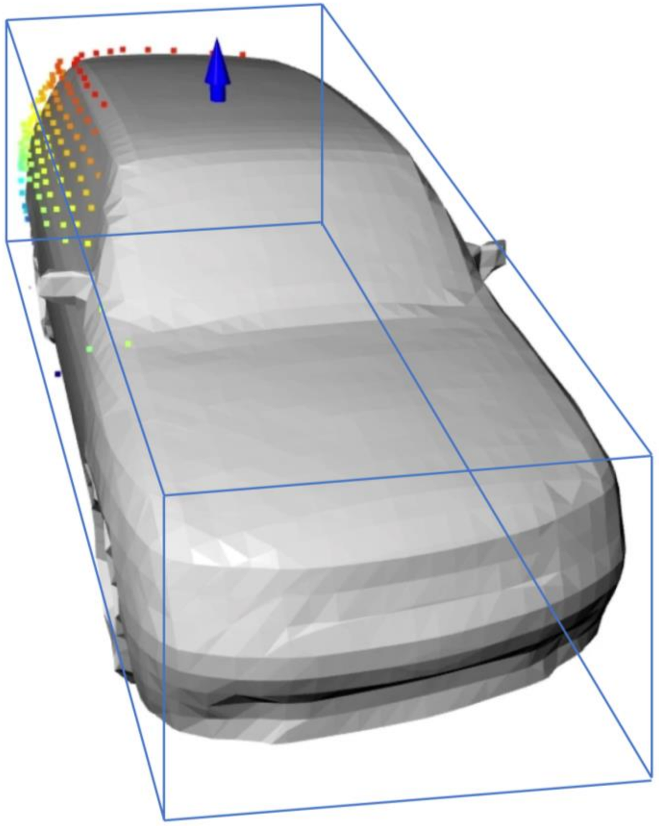
Result - Joint Optimization



Result - Joint Optimization

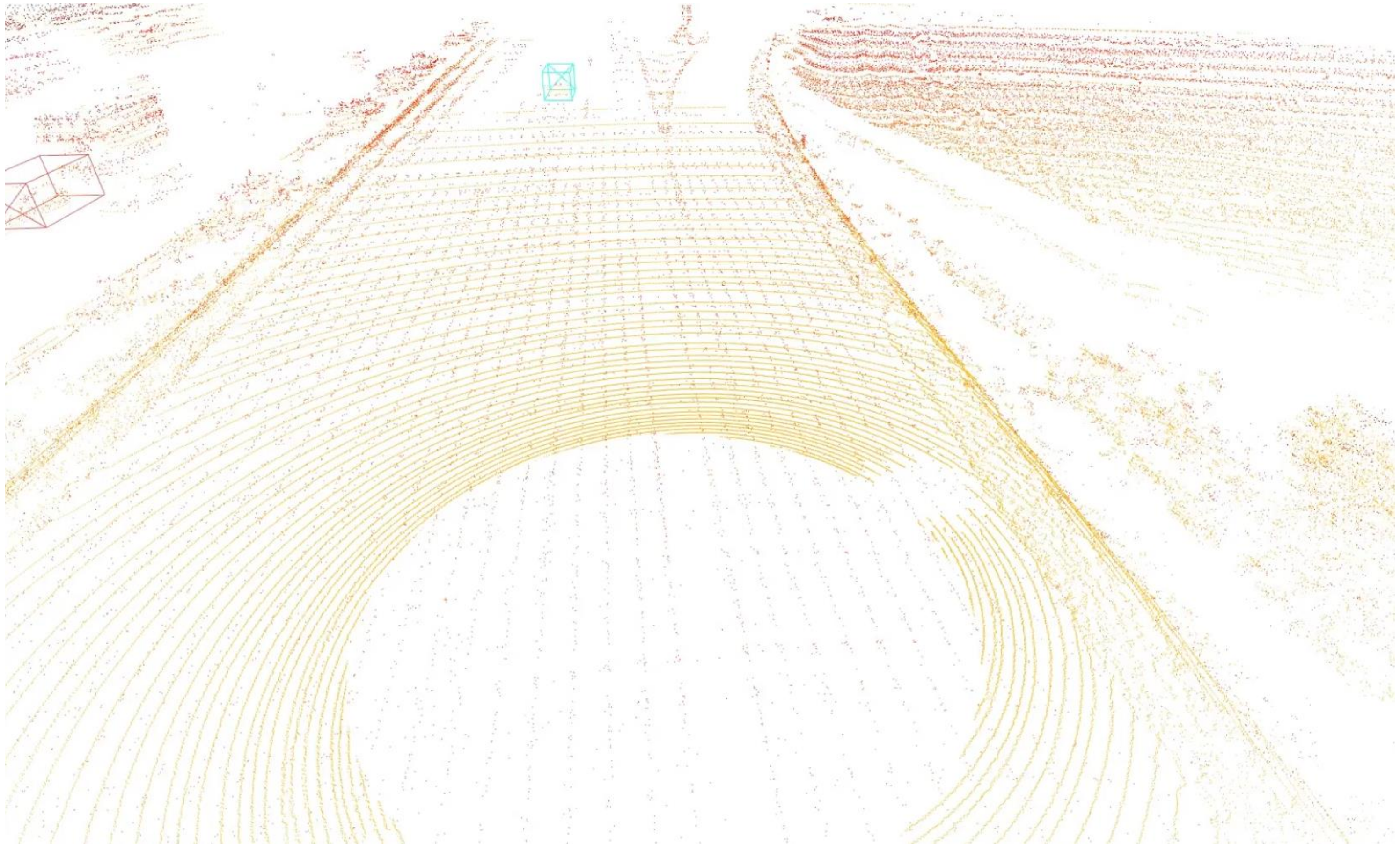


Bounding Box predicted by PointRCNN



Optimized Bounding Box

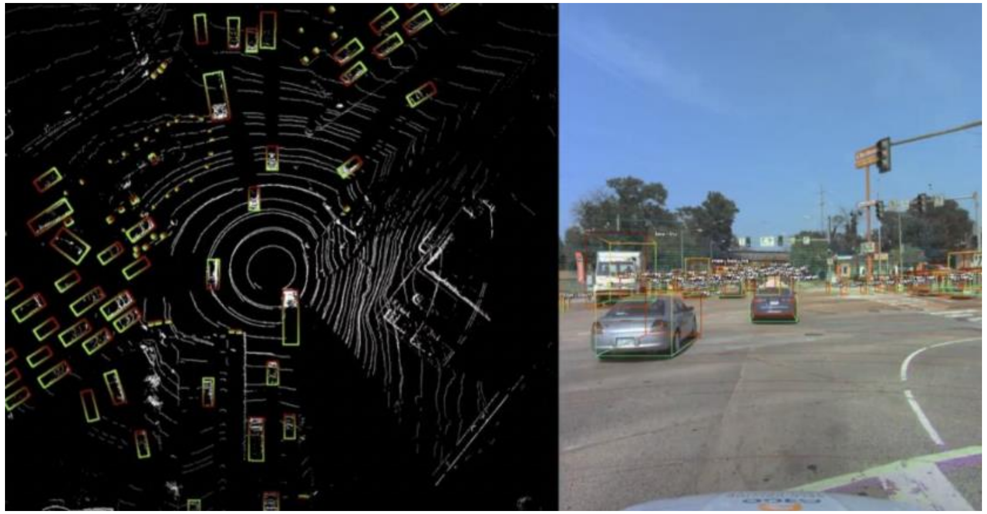
Result



Instance Completion and Motion Estimation for multiple cars

What is missing ?

Improve Instance Association
by using a SOTA MOT method



Evaluation using ground truth

1. Evaluation metric – MOT, Detection Metrics
2. Dataset with ground truth – Waymo, Nuscenes

Improvements

Improvement 1 : Multi Object Tracking (1)

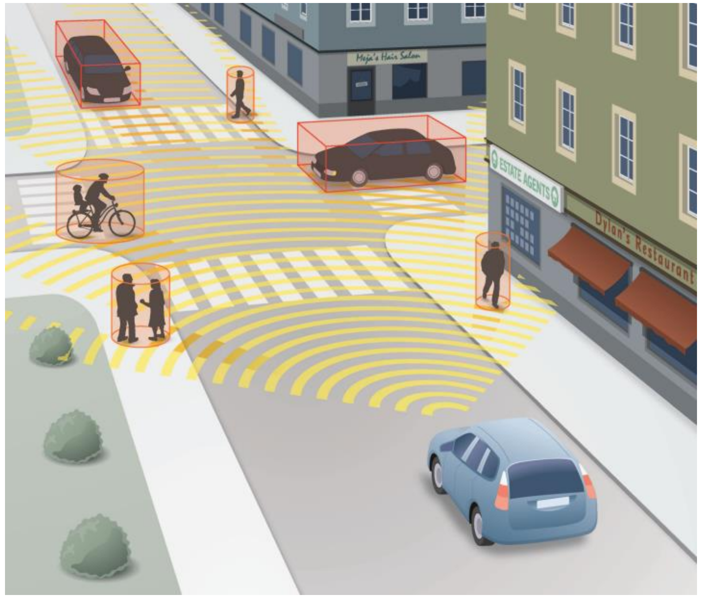


Illustration by Per Thorneus, reproduced from K. Granstrom et al, "Random Set Methods: Estimation of Multiple Extended Objects," in IEEE Robotics & Automation Magazine, June 2014

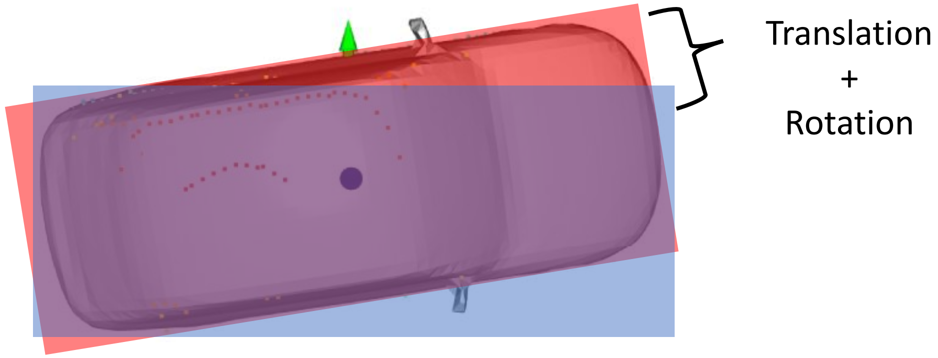
[1] Multi-Object Tracking (MOT) is the task of detecting the presence of multiple objects in video or a lidar sequence, and associating these detections over time according to object identities

Improvement 1 : Multi Object Tracking (2)

We will replace our IoU-based data association with a SOTA multi object tracking method.

1. This will improve the data association.
2. The task of this project is to verify if our method based on shape priors improve the SOTA MOT.

Improvement2 : loss function



Initial Bounding Box Prediction
Optimized Bounding Box

$$L = \lambda_s L_s + \lambda_R L_R$$

Latent Code Regularisation

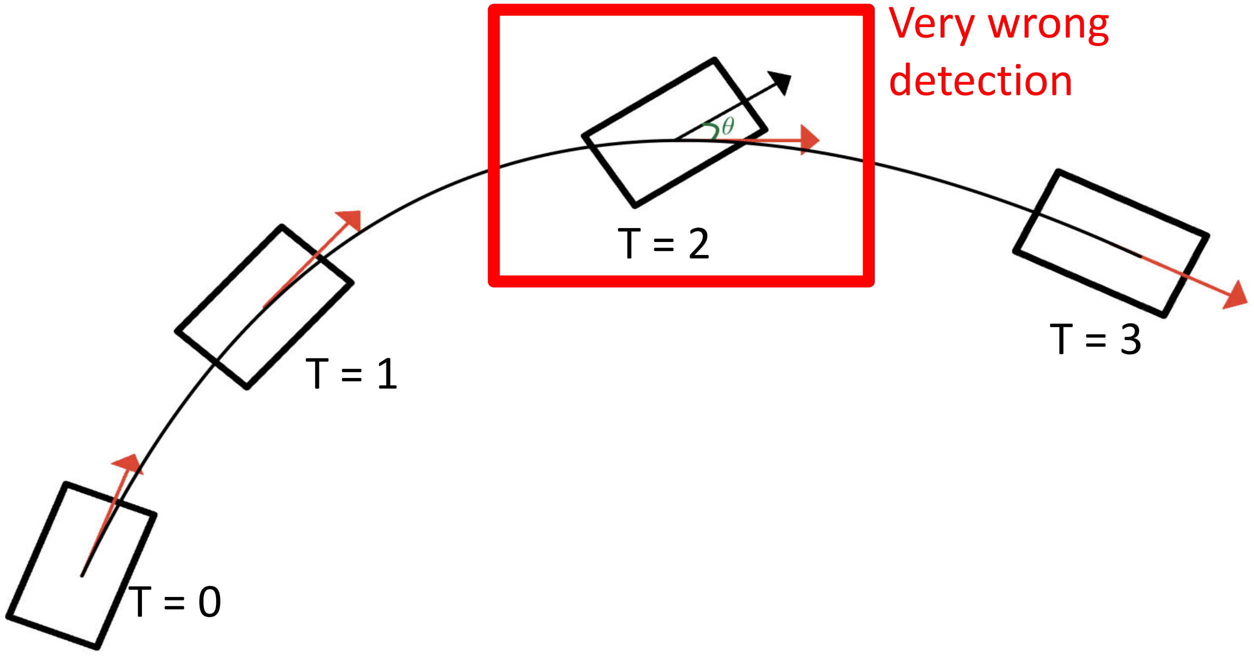
$$L_R = |1 - ||z|||$$

SDF Loss

$$L_s = |s - \hat{s}|$$

Limitations of the current method:
If the estimated bounding box is far from ground truth, the current method is unable to converge to correct solution

Improvement2 : Optimization with prior motion



- **Black Bounding Box** is the Predicted Detection
- **Red Arrow** is the motion prior
- **Black Arrow** shows the bounding box's orientation

We can leverage motion priors, such as velocity and angular velocity, as an additional component in the loss function to rectify occasional inaccuracies in bounding box predictions

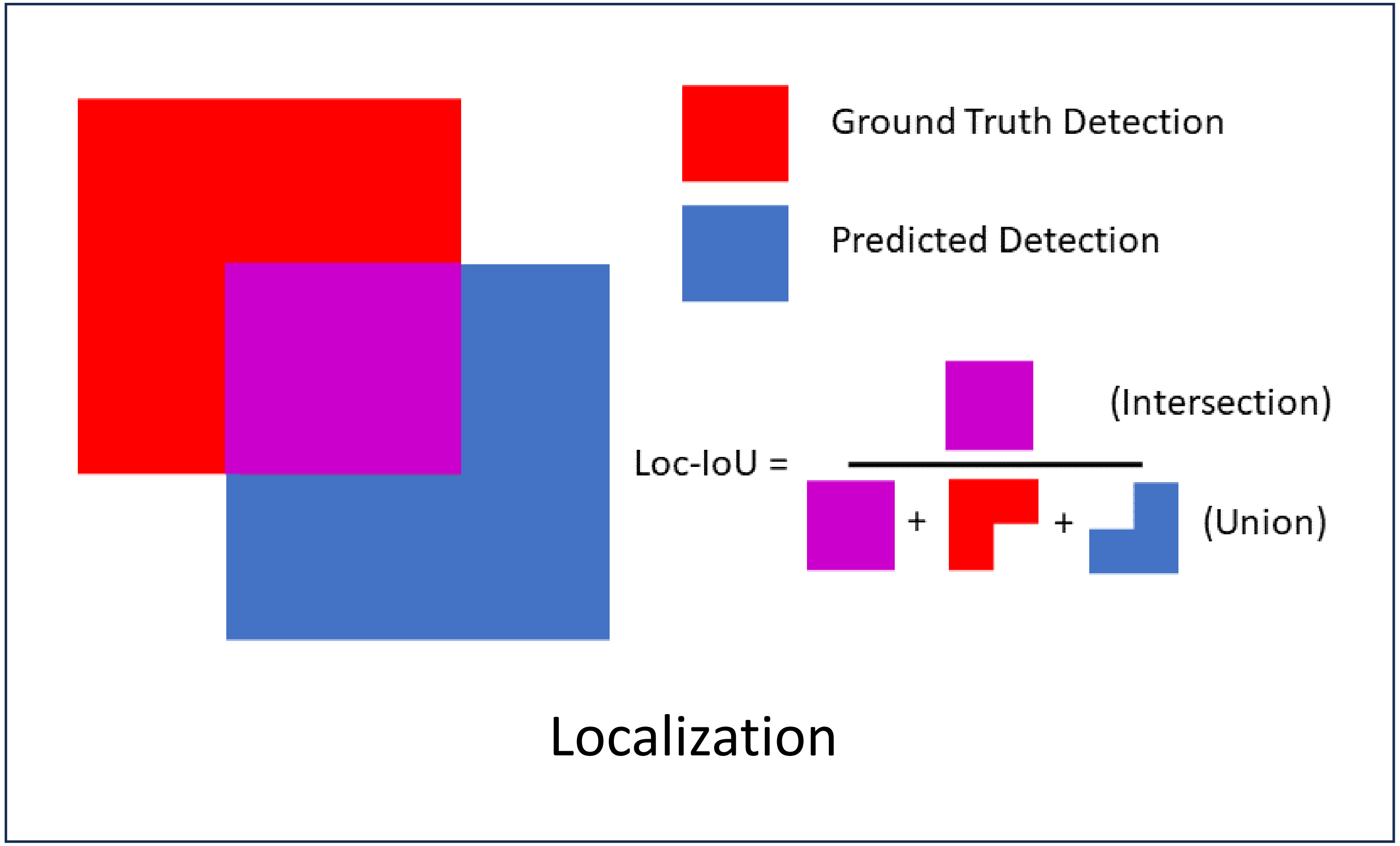
METRICS

HOTA Metric

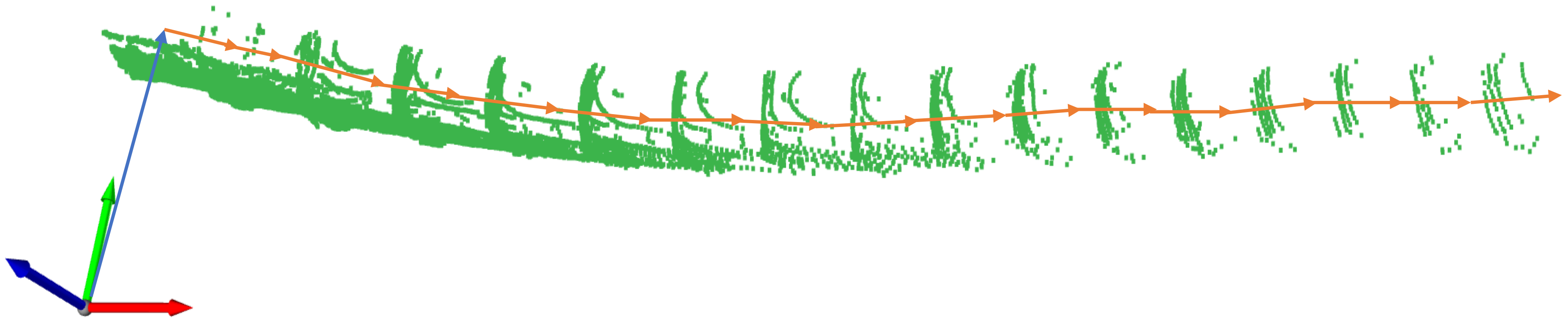
HOTA: A Higher Order Metric for Evaluating Multi-Object Tracking.

This metric consists of:

1. Localization
2. Association
3. Detection



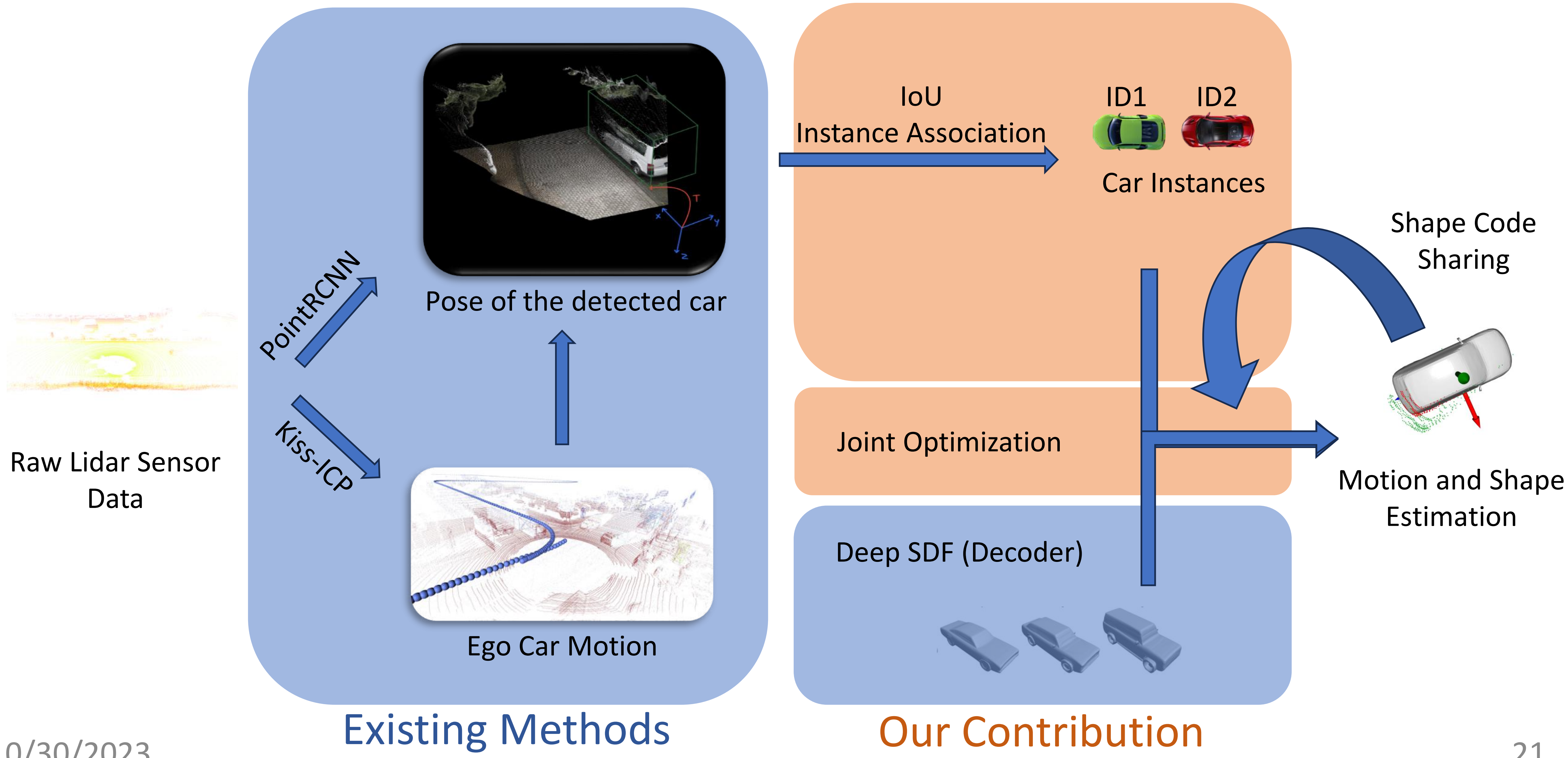
Evaluating with Baseline Methods



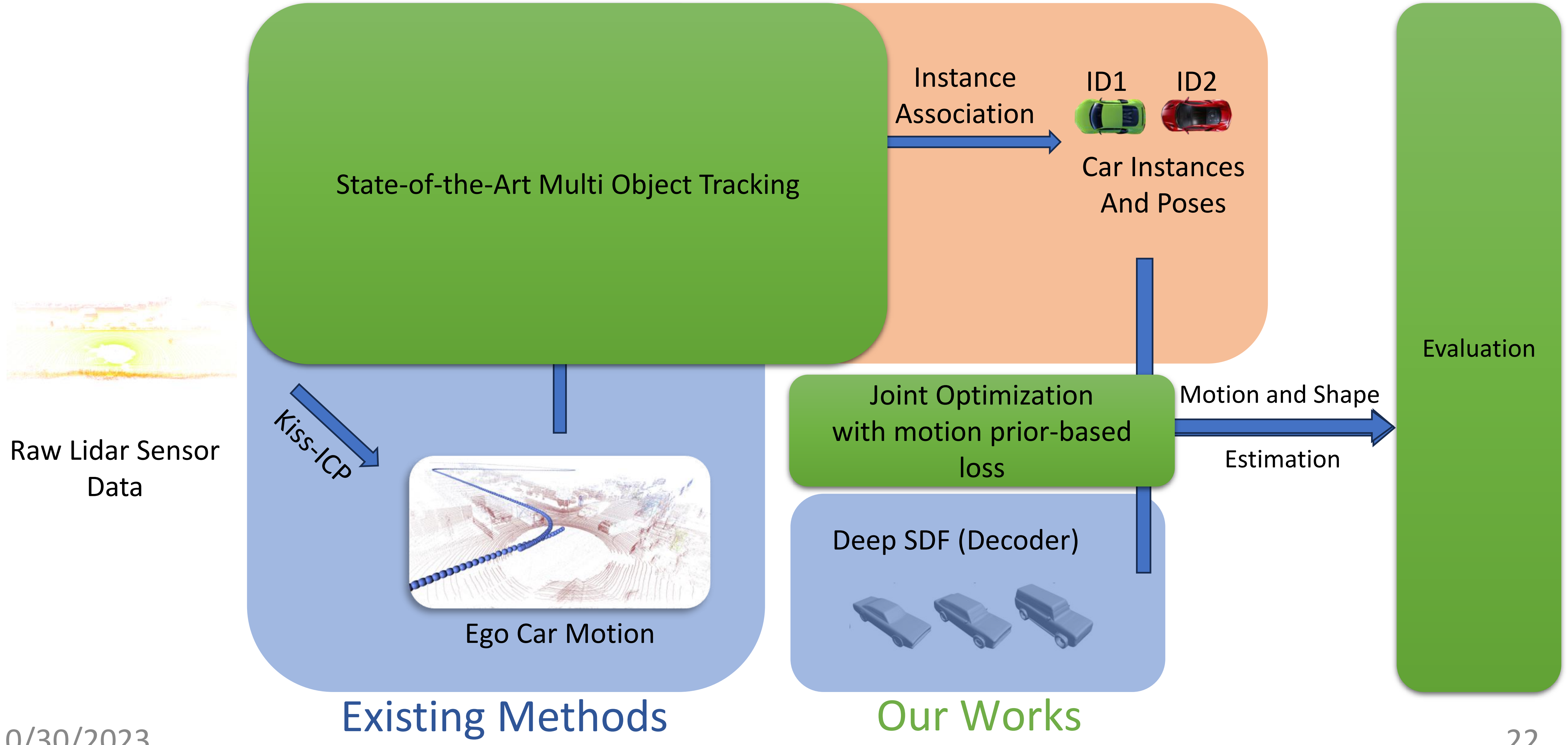
Method	Remarks
MOT	We will use an existing SOTA MOT
MOT + ICP	ICP requires correct Point-to-Point data association
MOT + Ours	We expect higher score in the metrics because of shape prior

Conclusion

Pipeline


Raw Lidar Sensor Data
PointRCNN
Kiss-ICP
Pose of the detected car
Ego Car Motion
Existing Methods
IoU
Instance Association
ID1
ID2
Car Instances
Joint Optimization
Deep SDF (Decoder)
Shape Code Sharing
Motion and Shape Estimation
Our Contribution

Pipeline



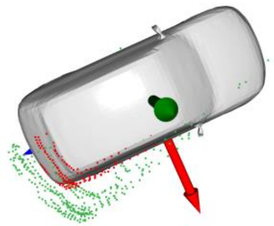
State-of-the-Art Multi Object Tracking
Instance Association
ID1
ID2
Car Instances And Poses
Raw Lidar Sensor Data
Kiss-ICP
Ego Car Motion
Joint Optimization with motion prior-based loss
Deep SDF (Decoder)
Motion and Shape Estimation
Evaluation
Existing Methods
Our Works

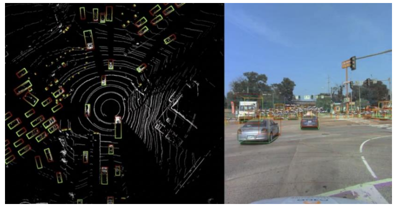
Citations

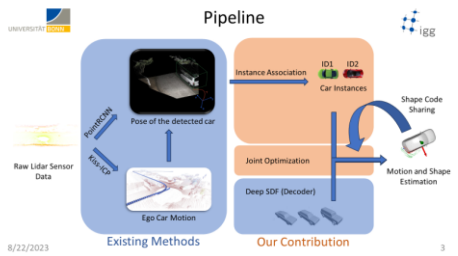
[1] Luiten, J., Ošep, A., Dendorfer, P. *et al.* HOTA: A Higher Order Metric for Evaluating Multi-object Tracking. *Int J Comput Vis* **129**, 548–578 (2021). https://doi.org/10.1007/s11263-020-01375-2

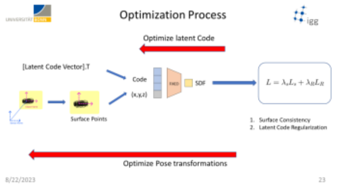
Any Questions?

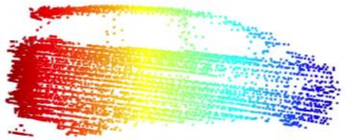
Any Questions?

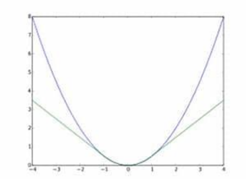


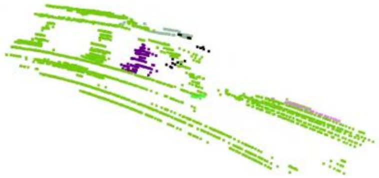






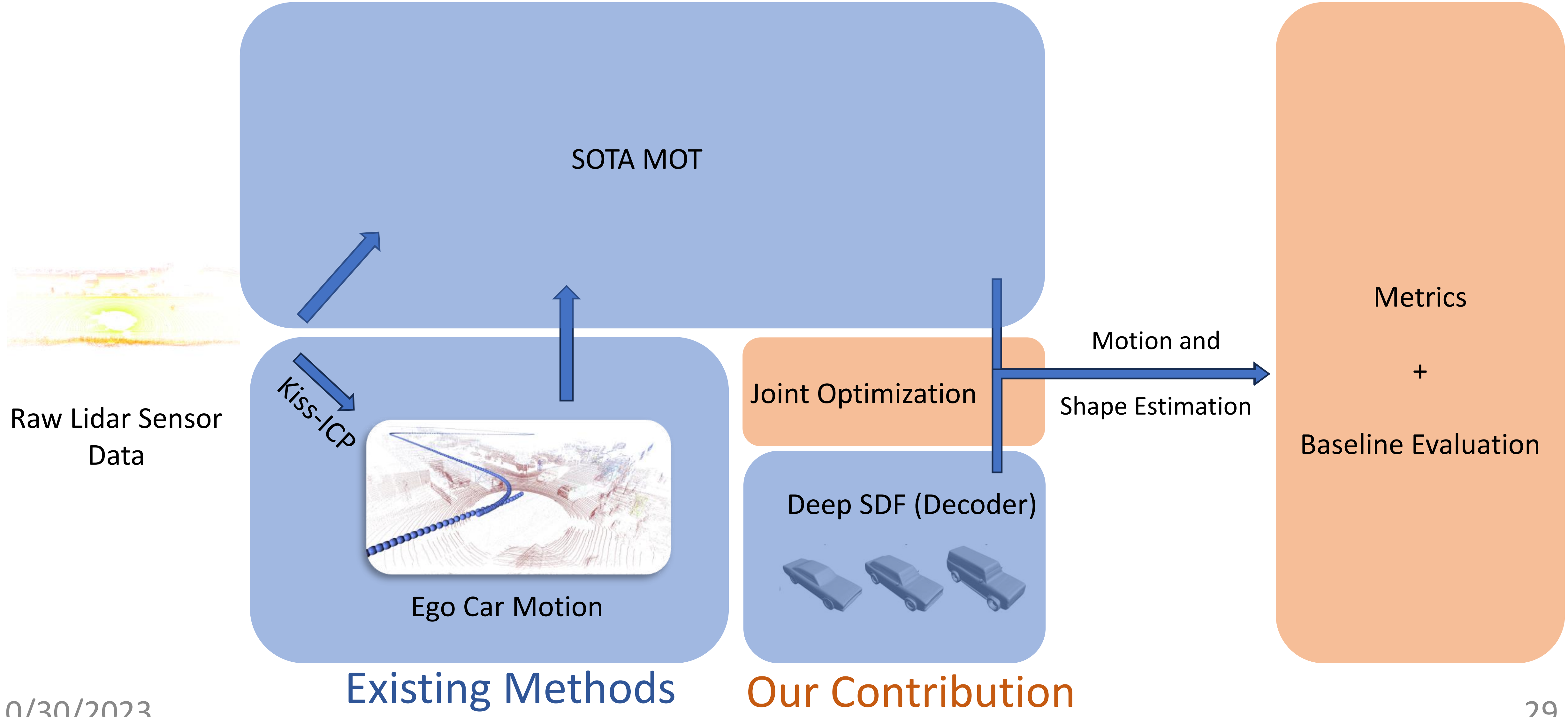






fps

Pipeline


SOTA MOT
Raw Lidar Sensor Data
Kiss-ICP
Ego Car Motion
Joint Optimization
Deep SDF (Decoder)
Motion and
Shape Estimation
Metrics
+
Baseline Evaluation
Existing Methods
Our Contribution

Detection Metrics

Average Precision Metric

mean Average Precision (mAP):
We use the well-known Average Precision metric, but define a match by considering the 2D center distance on the ground plane rather than intersection over union based affinities. Specifically, we match predictions with the ground truth objects that have the smallest center-distance up to a certain threshold. For a given match threshold we calculate average precision (AP) by integrating the recall vs precision curve for recalls and precisions > 0.1. We finally average over match thresholds of {0.5, 1, 2, 4} meters and compute the mean across classes.

True Positive Metric

Average Translation Error (ATE)
Average Scale Error (ASE)
Average Orientation Error (AOE).