CVPR 2023

[Submitted on 2 Aug 2022 (v1), last revised 19 Jun 2023 (this version, v3)]

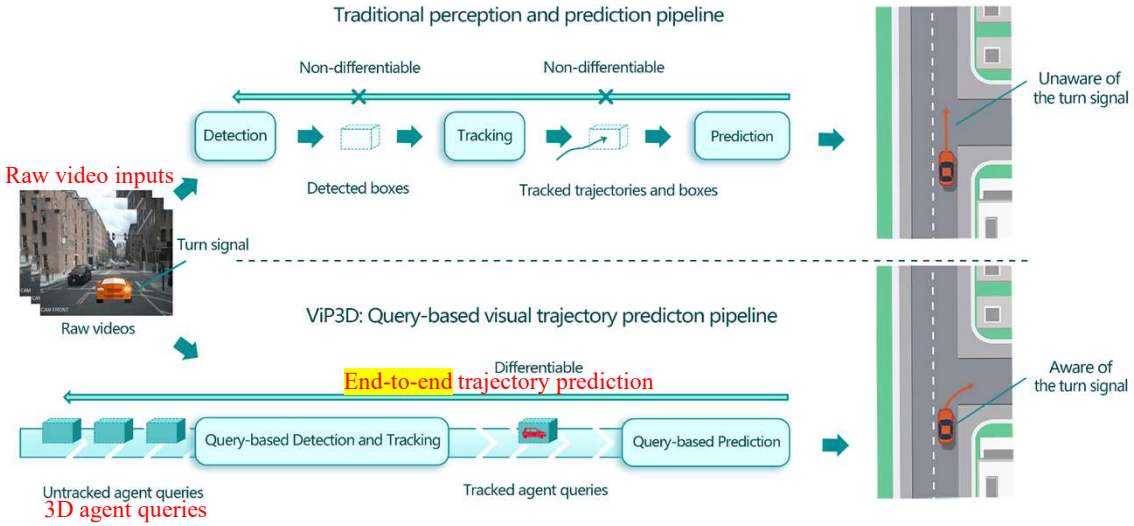
Visual Trajectory Prediction via 3D Agent Queries

Junru Gu[1*] Chenxu Hu[1*] Tianyuan Zhang[2,3] Xuanyao Chen[2,4]
Yilun Wang[5] Yue Wang[6] Hang Zhao[1,2†]

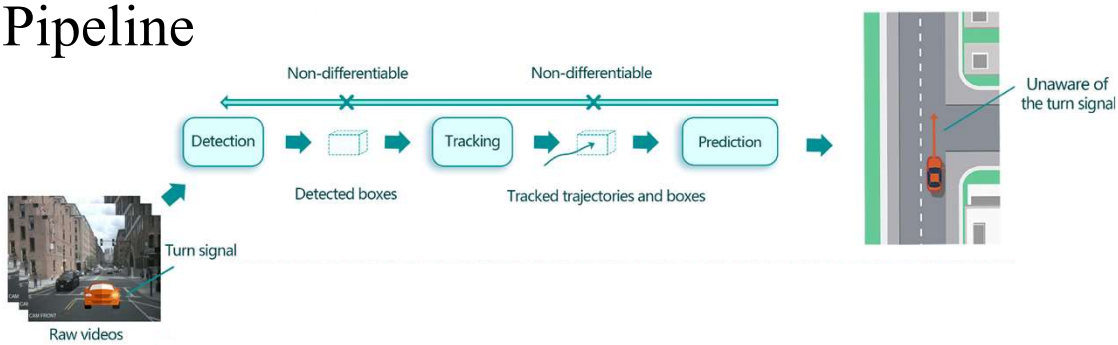
[1]IIIS, Tsinghua University [2]Shanghai Qi Zhi Institute
[3]CMU [4]Fudan University [5]Li Auto [6]MIT

https://tsinghua-mars-lab.github.io/ViP3D/

nuScenes



Traditional Perception and Prediction Pipeline



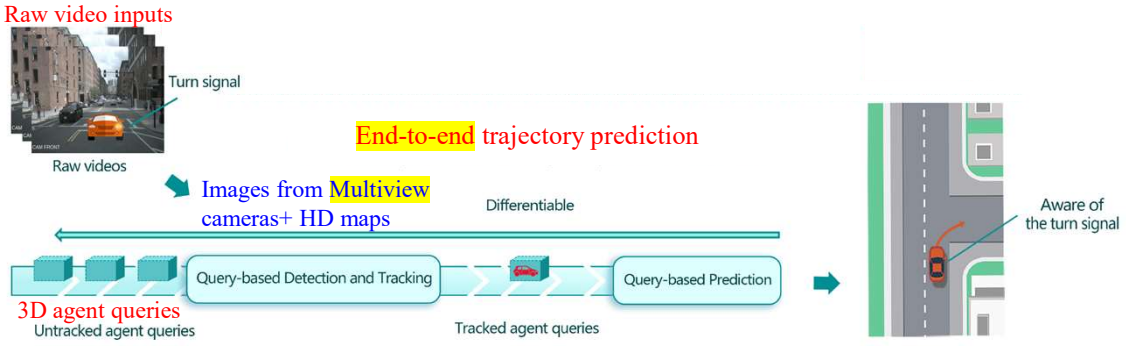
- Perceive agents and predict their behaviors.
- Separate modules
  - Lost useful information
  - Error accumulation

End-to-end models

LiDAR-based: IntenNet / FaF / PnPNet

Drawbacks:
(1) no rich information from cameras
(2) convolutional feature map, suffering from non-differentiable operations (ex: NMS in detection, object association in tracking)

Visual Trajectory Prediction via 3D Agent Queries



- Vision-based, fully differentiable
- Interpretable and High performance

nuScenes

Related Work: Trajectory Prediction

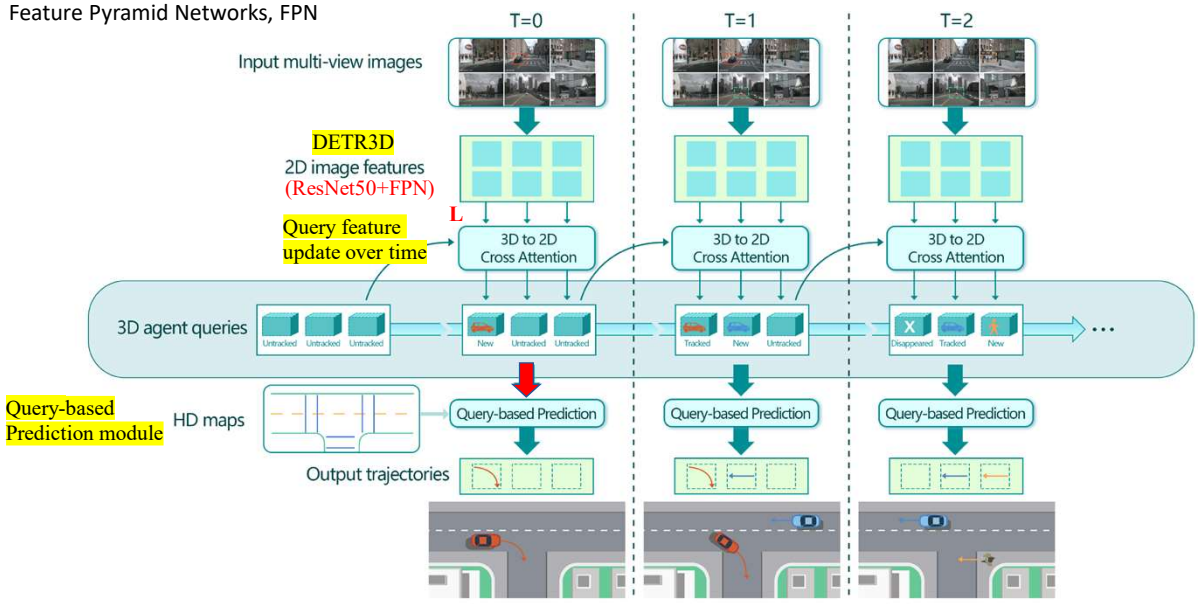
- GNNs
- Transformers
- Modeling uncertainty using latent variables
- Goal-based methods: predicting the intention first
  - Predict trajectory conditioning on these goals.

End-to-End Perception and Prediction

- Jointly optimizing detection, tracking, and prediction
- FaF, IntentNet, PnPNet
- Rely on BEV feature maps or heatmaps, leading to unavoidable non-differentiable operation

Tracked agent queries

Feature Pyramid Networks, FPN



3DETR: An End-to-End Transformer Model for 3D Object Detection

DETR3D: 3D Object Detection from Multi-view Images via 3D-to-2D Queries

Yue Wang
Massachusetts Institute of Technology
yuewang@csail.mit.edu

Vitor Guizilini*
Toyota Research Institute
vitor.guizilini@tri.global

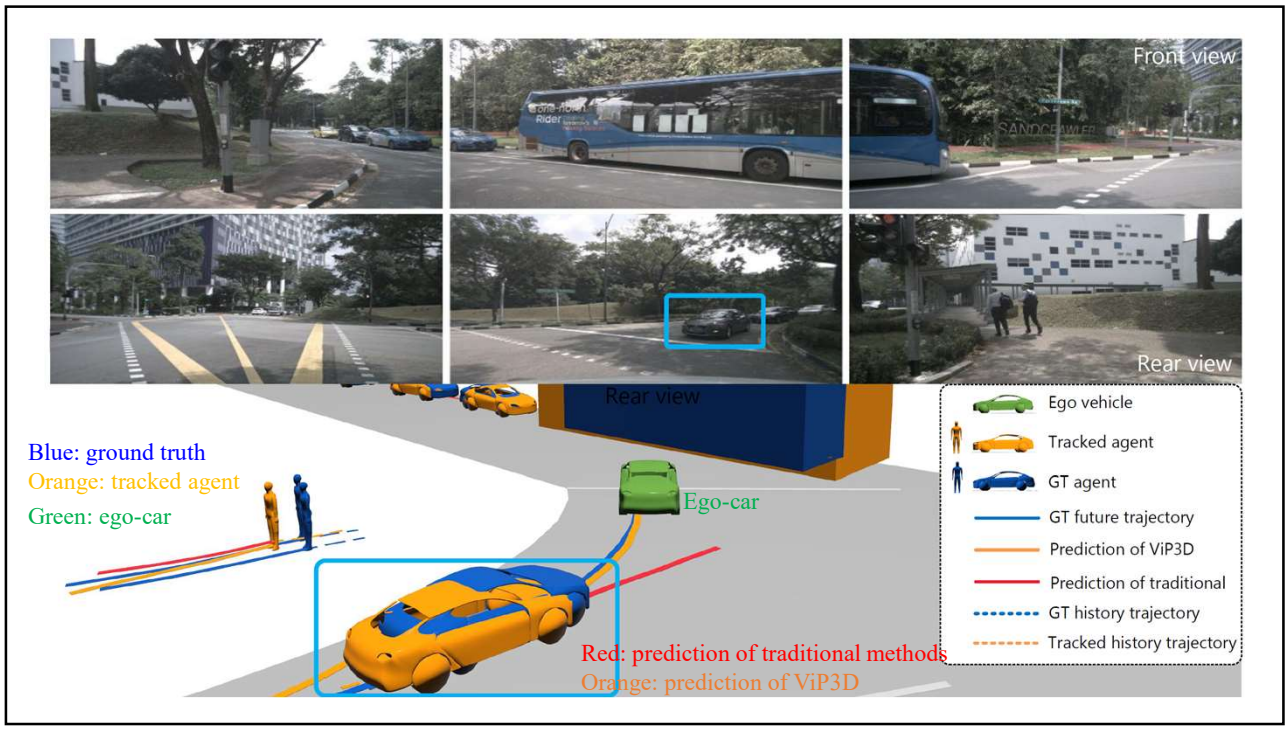
Tianyuan Zhang*
Carnegie Mellon University
tianyuaz@andrew.cmu.edu

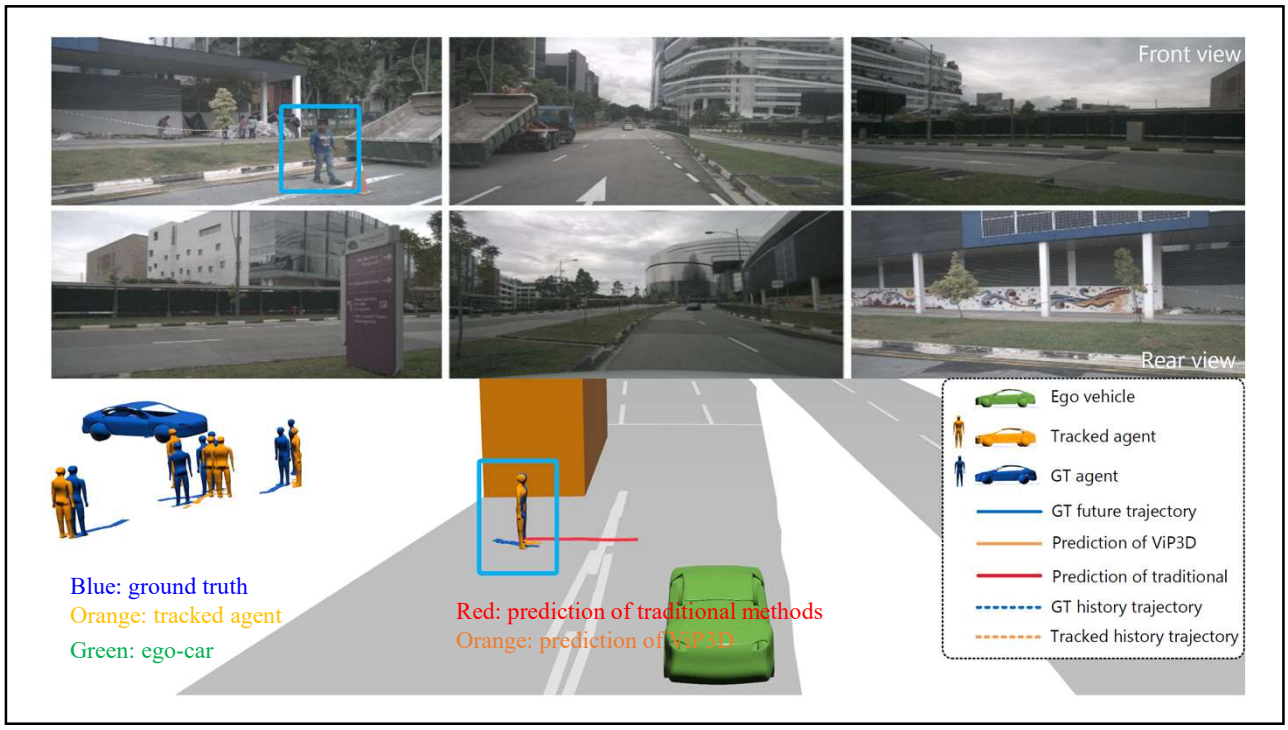
Yilun Wang
Li Auto
yilunw@cs.stanford.edu

Hang Zhao ¶
Tsinghua University
hangzhao@mail.tsinghua.edu.cn

Justin Solomon ¶
Massachusetts Institute of Technology
jsolomon@mit.edu

https://github.com/WangYueFt/detr3d


Front view
Rear view
Rear view
Ego-car
Blue: ground truth
Orange: tracked agent
Green: ego-car
Red: prediction of traditional methods
Orange: prediction of ViP3D
Ego vehicle
Tracked agent
GT agent
GT future trajectory
Prediction of ViP3D
Prediction of traditional
GT history trajectory
Tracked history trajectory


Front view
Rear view
Blue: ground truth
Orange: tracked agent
Green: ego-car
Red: prediction of traditional methods
Orange: prediction of ViP3D
Ego vehicle
Tracked agent
GT agent
GT future trajectory
Prediction of ViP3D
Prediction of traditional
GT history trajectory
Tracked history trajectory