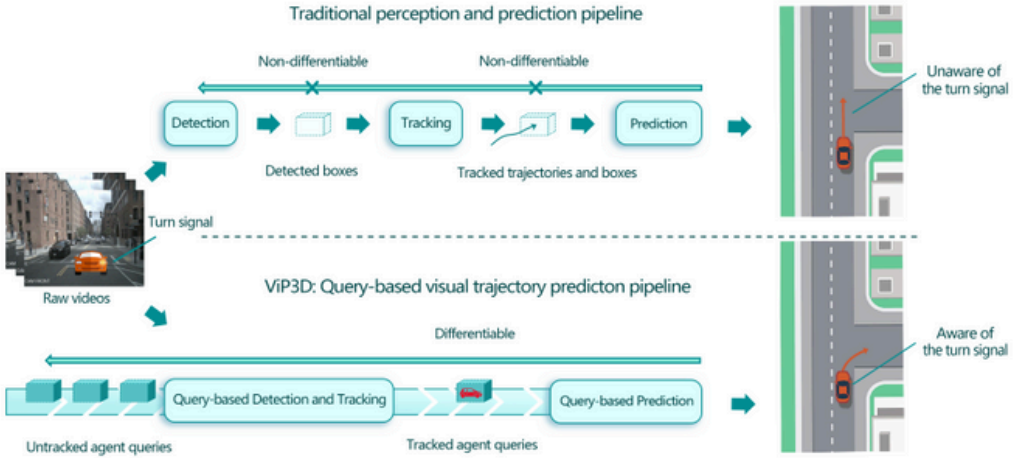
3D Agent-Based Visual Trajectory Prediction

What is the problem?

- The problem in existing autonomous driving systems is that perception (detecting objects) and prediction (forecasting future trajectories) are handled separately.
- The perception module passes hand-picked features (like agent bounding boxes and trajectories) to the prediction module.
- This separation restricts the prediction module to limited information, and errors in perception can propagate and worsen prediction accuracy



What has been done earlier?

- Earlier work on trajectory prediction, like IntentNet and FaF, has explored end-to-end models using LiDAR data.
- These models, however, fail to effectively use fine-grained visual information from cameras, and rely on non-differentiable operations such as non-maximum suppression, making them less efficient and harder to optimize

Swastik Pradhan B421061

What are the remaining challenges? What novel solution proposed by the authors to solve the problem?

- One of the remaining challenges is that existing models cannot fully leverage fine-grained visual cues, such as turn signals, brake lights, or pedestrians' body poses, which provide important context for predicting movements. Additionally, error accumulation from the multi-stage pipeline continues to degrade prediction performance, and existing methods struggle to generalize well to real-time dynamic environments
- The authors propose a novel solution called ViP3D, which introduces an end-to-end pipeline based on 3D agent queries. This fully differentiable model can directly process raw video data from cameras, integrating detection, tracking, and prediction in a unified system. By leveraging multi-view visual data and maintaining information across time steps through 3D agent queries, ViP3D can produce more accurate and robust trajectory predictions, eliminating the need for manually selected features. This improves prediction performance by avoiding the error accumulation that plagued traditional approaches.

Swastik Pradhan B421061