Title: Planning-oriented Autonomous Driving

Reviewer:
Raman Kumar Jha
rj2712

Summary:
Traditional autonomous driving systems use modular pipelines that suffer from error accumulation, feature misalignment, and poor task coordination. This leads to suboptimal safety and planning efficiency. UniAD introduces a unified, end-to-end framework integrating five key tasks (detection, tracking, mapping, motion forecasting, and occupancy prediction) into a single network optimized for planning. It uses a query-based transformer architecture to model interactions between agents, maps, and the ego vehicle across all stages. Unlike historical approaches that either deploy standalone models or multi-task learning with separate heads, UniAD adopts a planning-oriented philosophy. It prioritizes task coordination over isolated optimization, ensuring all modules contribute synergistically to the final planning objective.
The innovation makes sense, as autonomous driving requires holistic scene understanding. UniAD's unified design minimizes cascading errors and enables richer agent-environment interactions. UniAD achieves state-of-the-art performance on nuScenes, which makes it an important contribution to planning research.

Strengths:

1. UniAD is the first framework to unify perception, prediction, and planning in a single network. The query-based design enables seamless cross-task communication, addressing historical issues of error propagation in modular systems.
2. Using track queries (for temporal consistency in detection/tracking) and ego-vehicle queries (explicitly modeling the SDV's interactions) provides a scalable way to handle dynamic interactions.

3. The paper provides extensive ablation studies and outperforms LiDAR-based methods proving the viability of vision-centric systems. The non-linear trajectory optimization and collision-aware planning demonstrate practical safety enhancements.

The paper has solved several earlier existing problems by proposing the first framework to unify perception, prediction, and planning in a single network. Yes, the paper takes a new angle of looking at the problem, as it is not focused on just one unified network for a single task, but instead, it is solving many problems in a single network. The model and the experiment are not totally ingenuine, as it addresses multiple tasks in one network, so it uses a transformer, motionformer, mapformer, occformer, and planner for this. Yes, the performance stands out, as it performed all the tasks beating all SOTAs while maintaining a unified network.

Weaknesses:

1. While UniAD achieves superior performance, its transformer-heavy architecture incurs high FLOPs and low FPS, limiting real-time deployment. The two-stage training further increases development overhead.
2. Experiments are confined to the nuScenes dataset. Real-world testing in edge cases like adverse weather, and erratic pedestrians is absent. Failures in long-tail scenarios highlight generalization gaps.
3. UniAD focuses on motion/occupancy prediction but omits behavior prediction such as intent recognition. This limits its ability to model high-level agent strategies critical for urban driving.

Yes, the premise of the paper definitely makes sense, as it is solving various tasks in one unified model. The methodology is a little complex, as it has many modules for different tasks, but it is also comprehensive. It could have been compared to other datasets and could have been tested on real-world scenarios, which is currently missing. The limitations have been discussed above.

Possible Future Extensions:

1. Lightweight deployment is an aspect of this paper, which can be used for future extensions. It will reduce computational costs via model compression (e.g., distillation, pruning) or hybrid architectures. It might sacrifice performance for efficiency.
2. Integration of behavior prediction can help the network, perform more efficiently. It will enhance planning by predicting agent intents (e.g., lane changes, turns) using language-conditioned or reinforcement learning frameworks. It can Increase model complexity and training instability.

Conclusion:

UniAD advances autonomous driving by unifying critical tasks into a planning-oriented framework, validated through rigorous experimentation. Its query-based architecture sets a new benchmark for end-to-end systems. Despite computational overhead and limited real-world testing, the methodological novelty and empirical results provide a positive review. It should get accepted with minor revisions addressing deployment feasibility and edge-case analysis. I would give this paper a strong acceptance. The strengths in task integration and performance outweigh the weaknesses, which are common in early-stage frameworks.