XJTLU | SCHOOL OF ROBOTICS 智能机器人学院

# Case Studies towards the Deep Learning Autonomous Driving AI

Team Member: Ruiying Ma, Xueting Han, Haoyuan Shan, Jiadong Yu, Haoran Zhao, Zhinuo Zhang

Supervisor: Dr. Fan Zhang, School of Robotics, Xi'an Jiaotong-Liverpool University

SURF-2024-0102

## Abstract

It is an emerging research question in AI: how to enable a vehicle to self-drive safely?

Contemporary Automated Vehicles (AVs) or Autonomous Driving applications may outperform human in driving tasks, yet they are far from perfect especially when complex attentional and action selections are required for safety.

In the current work, we performed literature review on the state-of-the-art AI developments in automated vehicles, and further conducted case studies to test open-sourced data and models, especially in perception, decision-making, and planning.

Moreover, the findings will help understanding the learning & developing process in automated vehicles and further benefit relevant research.

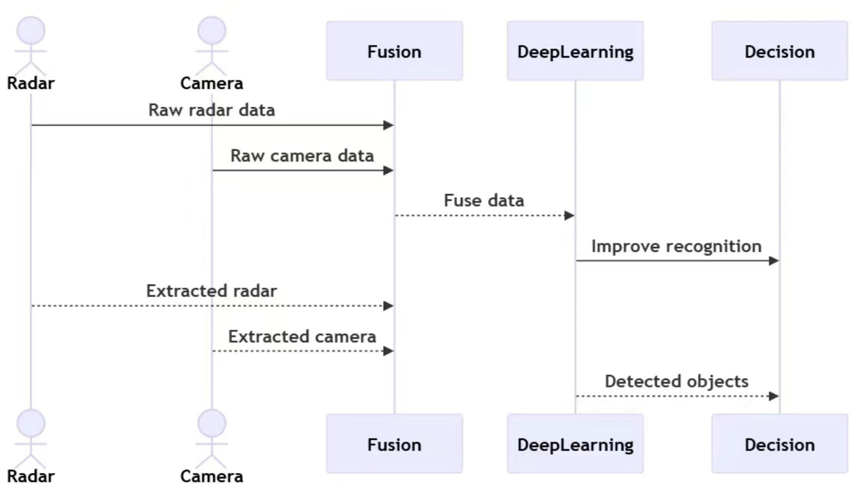
## Study 1: Integrating Perception, Fusion, and Computing Architectures for Autonomous Driving: A Comprehensive Review

**Introduction:**

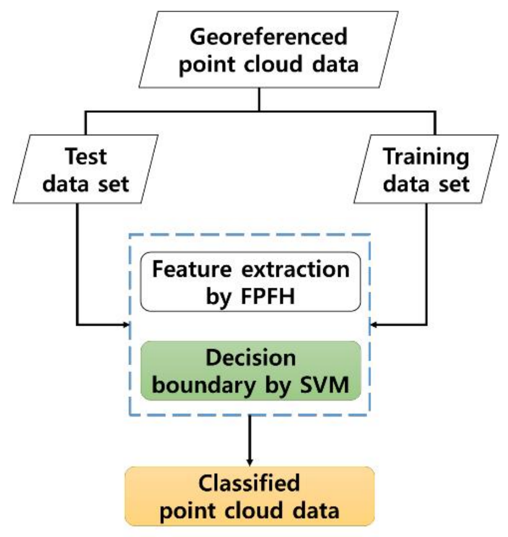
- Here we examined the critical technologies and architectural frameworks involved in the development of autonomous driving systems.
- The study also briefly touches upon the testing and verification methods employed to ensure system safety and reliability.

**Perception & Data Fusion:**

- Perception: integration of radar and camera data through various fusion techniques
- Data Fusion: multi-modal deep learning[1] framework for RGB-D object recognition, combine RGB with depth to enhance recognition accuracy



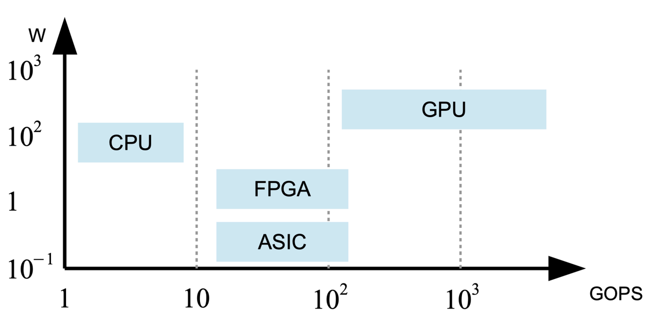

*Fusion of Radar and Camera Sensing Modalities*



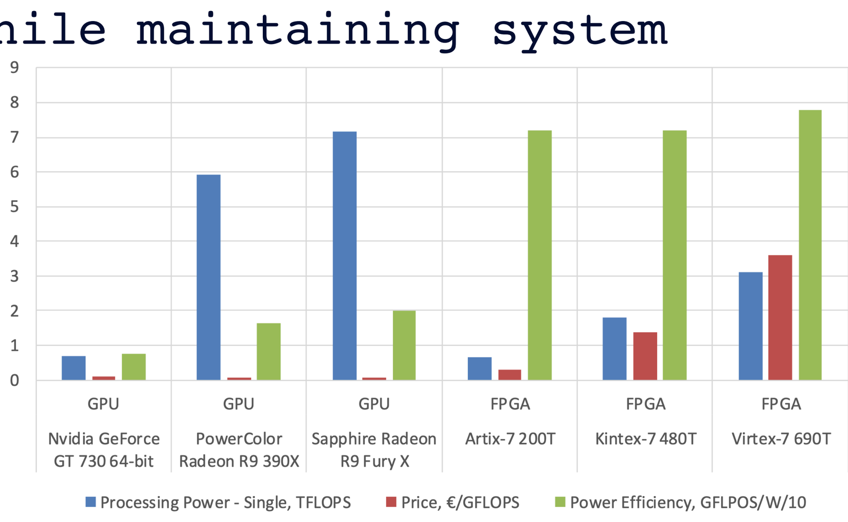

*Flowchart of processing LiDAR data*

**SW and HW Architecture:**

- A hybrid GPU and FPGA computing system where the GPU serves as the primary computing platform
- The hybrid computing architecture has also proven effective in optimising performance and energy efficiency while maintaining system safety [2,3].




*GPU vs FPGA Performance Comparison [2]*




*GPU vs FPGA Performance Comparison [3]*

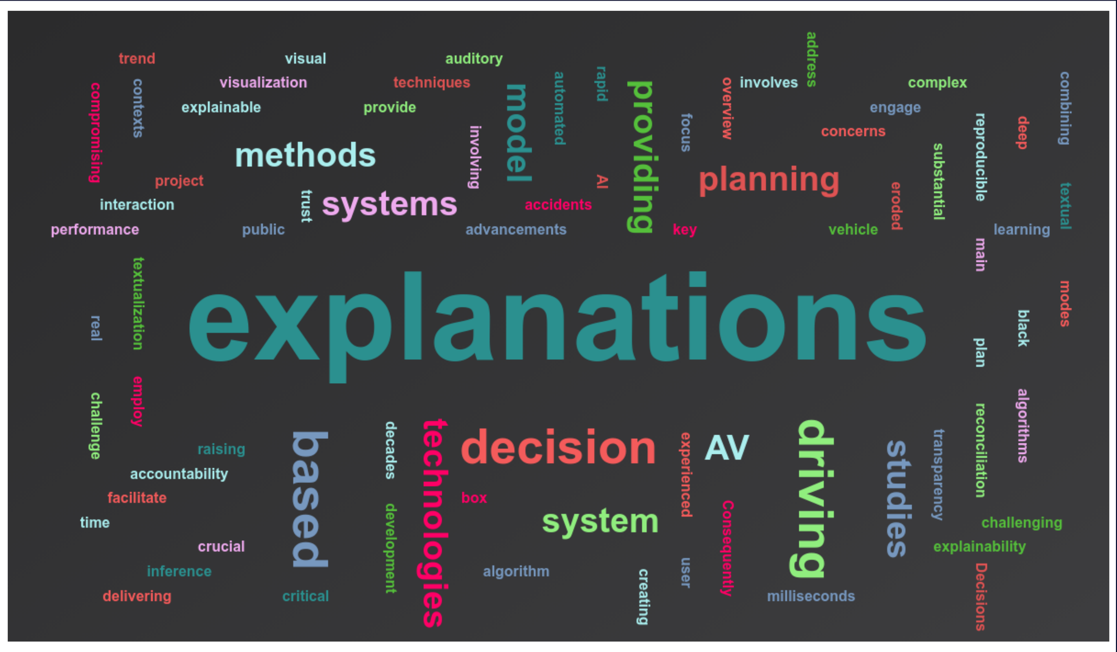
**Discussion:**

- The study underscores the importance of integrating multiple technologies to develop a robust autonomous driving system.
- The combination of novel perception, data fusion, and SW/HW architectures contributes to the overall reliability and safety of AVs.
- Future challenges remain, particularly in refining these technologies to handle more complex driving scenarios.

## Study 2: Enhancing Explainability on Automated Planning and Decision-Making

**Introduction:**

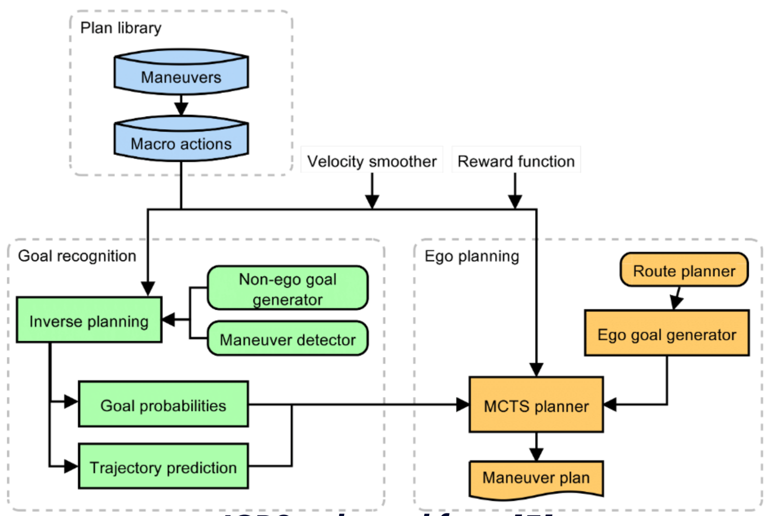
- This study addresses the critical need for transparency and accountability in AVs
- Focus: explainable planning and decision-making
- Emphasise the importance of clear explanations in fostering public trust and ensuring safety (e.g. [4])



*Top Keywords in XAI of Planning and Decision Making*

**Methods:**

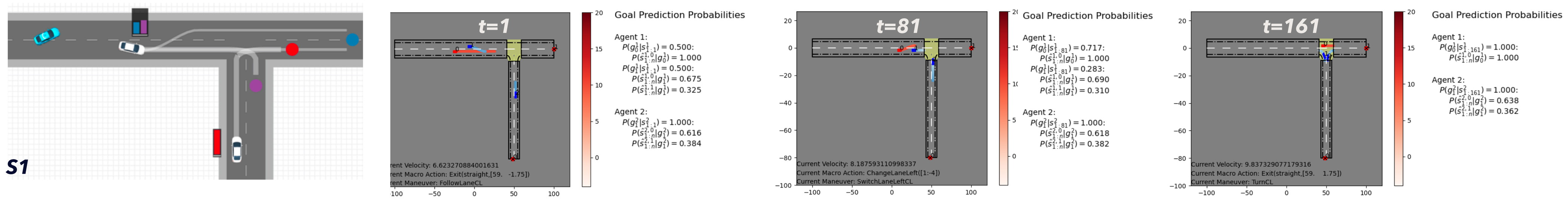
- Reproduce IPG2: Interpretable Goal-based Prediction and Planning [5]
- Implementation: rational inverse planning to recognise other vehicles' goals and integrates information into long-term planning and prediction.
- We evaluated the system performance in simulations of realistic driving scenarios




*IGP2, adapted from [5]*

**Results:**

- We tested S1 (t-junction) from the proposed testing scenarios, configuring parameters include the type, location, speed, goals, and other behaviours of the agents
- Our results were consistent with the expectations outlined in [5], demonstrating the system's effectiveness.




*Schematics of testing scenario S1 (adapted from [4]) and our testing results (at t=1, t=81, t=161)*

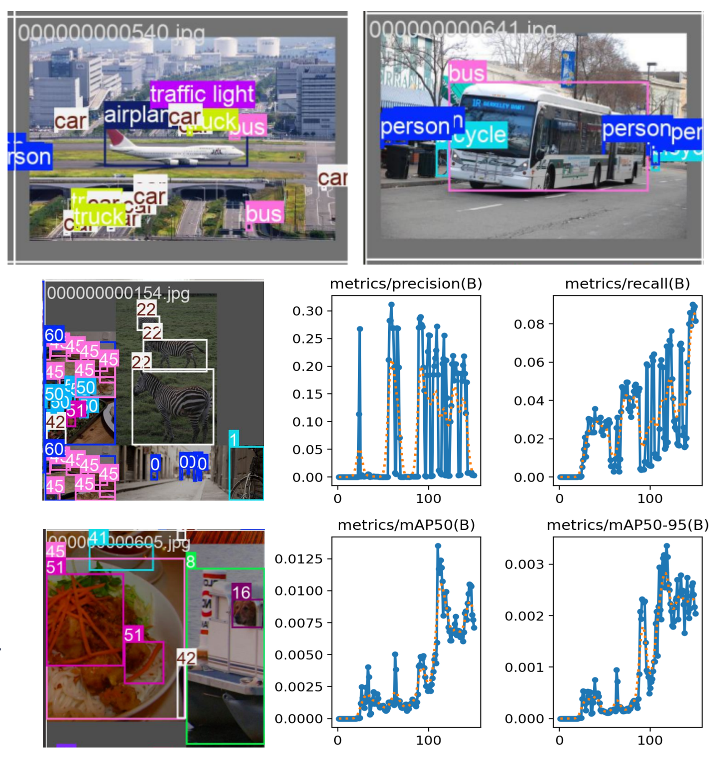
**Discussion:**

- This study underscores the critical role of explainability in the development of Automated Vehicles. By exploring and reproducing advanced methods like the IGP2 system, it contributes to the ongoing efforts to make AVs more transparent and trustworthy.
- Future research should continue to refine these techniques, focusing on further improving the balance between decision-making speed and explanation clarity.

## Study 3

**Introduction:**

- compare YOLOv5 with the FENet used in autonomous driving [6,7].
- Evaluate recognition and detection performance in driving scenarios.



*Example from COCO2017 and our results using YOLOv5*

**Methods & Results:**

- Dataset: COCO2017 datasets, simulation of multi-camera systems in CALRA
- Model Training: the adam optimiser unchanged & added a new ELA attention mechanism.
- YOLOv5 had low confidence and fewer identifications than FENet

**Conclusion:**

- Data augmentation based on collision detection and spatial context location expansion improved performance for small objects
- Consistent with [7]: external cameras improve vehicle safety and environmental awareness

*Simulation of multi-camera system in CARLA to train a YOLOv5 for fusion of 3D bounding boxes*

## Reference & Acknowledgements

This work was supported by XJTLU Summer Undergraduate Research Fellowship: SURF-2024-0102.

[1] Yao S, Guan R, Huang X, Li Z, Sha X, Yue Y, Lim EG, Seo H, Man KL, Zhu X, Yue Y. Radar-camera fusion for object detection and semantic segmentation in autonomous driving: A comprehensive review. IEEE Transactions on Intelligent Vehicles. 2023 Aug 22.
[2] Zhang, Q., Zhang, M., Chen, T., Sun, Z., Ma, Y., & Yu, B. (2019). Recent advances in convolutional neural network acceleration. Neurocomputing, 323, 37-51.
[3] BERTEN D. GPU vs FPGA performance comparison. InProceedings of the 2017 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays-FPGA 2016 May (Vol. 17, pp. 2-5).
[4] Xu Z, Zhang Y, Xie E, Zhao Z, Guo Y, Wong KY, Li Z, Zhao H. Drivegpt4: Interpretable end-to-end autonomous driving via large language model. IEEE Robotics and Automation Letters. 2024 Aug 7.
[5] Albrecht SV, Brewitt C, Wilhelm J, Gyevnar B, Eiras F, Dobre M, Ramamoorthy S. Interpretable goal-based prediction and planning for autonomous driving. In2021 IEEE International Conference on Robotics and Automation (ICRA) 2021 May 30 (pp. 1043-1049). IEEE.
[6] Jocher G, et al., ultralytics/yolov5: v6. 2-yolov5 classification models, apple m1, reproducibility, clearml and deci. ai integrations. Zenodo. 2022 Aug. doi: 10.5281/zenodo.4154370
[7] A. Hazarika, A. Vyas, M. Rahmati and Y. Wang, "Multi-Camera 3D Object Detection for Autonomous Driving Using Deep Learning and Self-Attention Mechanism," in IEEE Access, vol. 11, pp. 64608-64620, 2023, doi: 10.1109/ACCESS.2023.3288112

Xi'an Jiaotong-Liverpool University 西交利物浦大學

SURF Summer Undergraduate Research Fellowship