Project 3: Autonomous Vehicle Dashboard

Jesulona Akinyele
Masters of Science in Robotics Engineering
Worcester Polytechnic Institute
Worcester, Massachusetts 01609
Email: jfakinyele@wpi.edu

Krunal M. Bhatt
Masters of Science in Robotics Engineering
Worcester Polytechnic Institute
Worcester, Massachusetts 01609
Email: kmbhatt@wpi.edu

***Abstract*—This document outlines a project aimed at developing advanced visualizations for human-robot interaction (HRI), specifically focusing on the enhancement of visual feedback systems in autonomous vehicles. Recognizing the pivotal role of intuitive and efficient visualizations in building user trust and facilitating machine understanding, we propose to design a visualization system inspired by, yet superior to, Tesla's latest dashboard designs. Utilizing videos captured from a 2023 Tesla Model S, the project encourages the use of both deep learning and classical approaches to render comprehensive visual insights into the vehicle's environmental perception and decision-making processes. The outcome will be evaluated based on creativity, effectiveness, and the aesthetic appeal of the visualizations, employing Blender for rendering. This endeavor not only aims to address the shortcomings of current HRI visual feedback but also to set a new benchmark in the domain.**

I. Phase 1

In Phase 1, our primary objective was to lay down the foundation for our visualization system by implementing basic yet critical features essential for a self-driving car's navigation and interaction with its environment. These features encompass lane detection, vehicle detection, pedestrian identification, traffic light recognition, and road sign detection.

A. *Basic Object Detection*

We embarked on this journey with the implementation of lane detection using YoloPV2, successfully identifying lanes but without differentiating between lane types. For vehicle detection, we employed Detic to identify different types of vehicles. Traffic lights, street signs, and pedestrian detection were also conducted using Detic, allowing for a broad spectrum of object recognition.

From the detected objects, we calculated the centroid and used Zoedepth for depth estimation, followed by a 3D reprojection to integrate these objects into our 3D scene. This multi-step process was aimed at creating a realistic and intuitive visualization of the car's environment in Blender.

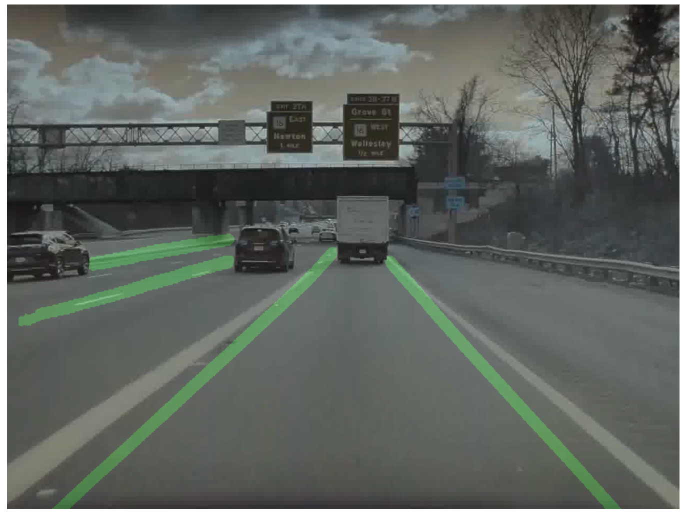

Fig. 1. Lane Detection

Fig. 2. Detected Cars and People

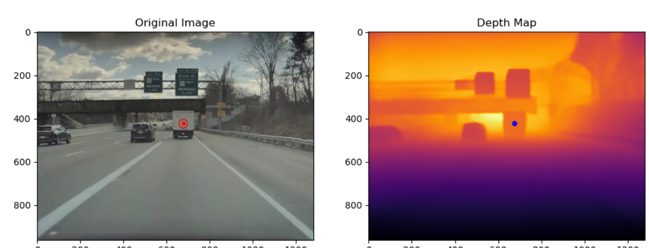

Fig. 3. Depth Map

B. Phase 1 Issues

Despite the successful implementation of object detection algorithms and depth estimation, we encountered several pitfalls that affected the project's progression:

- **3D Reprojection Challenges:** Although all components were functional, incorrect 3D reprojection techniques hindered our ability to accurately place and render objects within the virtual environment. The mismatch between the object's estimated position and its actual placement in the 3D scene affected the realism and utility of the visualizations.
- **Lane Rendering:** Rendering lanes in Blender proved challenging. Our initial attempts did not successfully differentiate between various lane types, such as dashed or solid, nor did we incorporate color variations (white and yellow)
- **Vehicle, Traffic Light, and Pedestrian Representation:** While we could detect different entities within the scene, our approach to representing these detections in Blender was basic. All vehicles were represented by a generic car shape, and pedestrians were depicted in a uniform manner, facing the same direction. This simplification was a placeholder for more sophisticated representations in future phases.
- **Texture Application for Road Signs:** Applying textures to road signs, other than stop signs, presented difficulties.

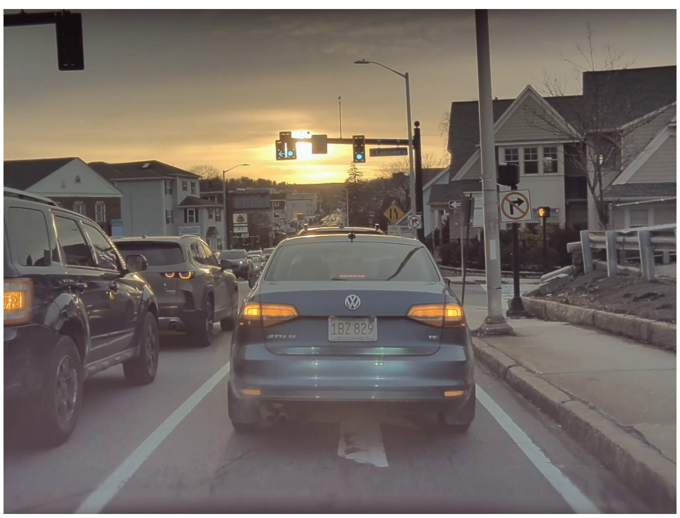

Fig. 4. Phase 1 Example Input

We plotted the lanes using "Splines" in Blender and used a mesh to highlight the area where the lanes are plotted.

Phase 1 laid the groundwork for our visualization system, with significant achievements in basic object detection and depth estimation. However, the challenges encountered in 3D reprojection, lane rendering, and the simplistic representation of detected objects have provided valuable lessons.

II. Phase 2

Building upon the foundational work of Phase 1, Phase 2 aims to significantly enhance the visualization system's granularity and functionality. This phase introduces advanced features aimed at refining the vision system's ability to aid in the autonomous vehicle's navigation decisions through more detailed and accurate object detection and classification.

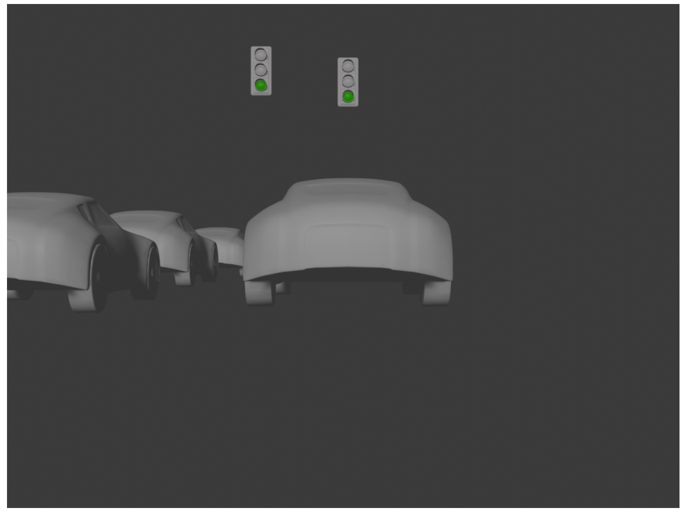

Fig. 5. Phase 1 Example Output

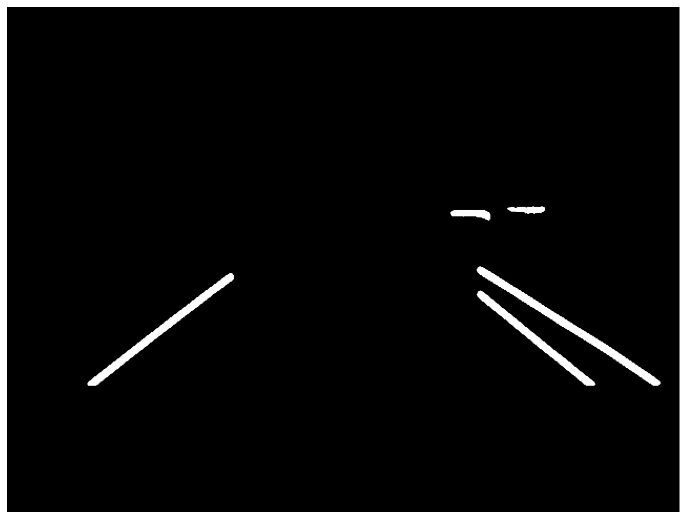

Fig. 6. Phase 1 Example Detected Lanes

A. Advanced Features

A critical advancement in this phase was the extensive focus on capturing orientation data using YOLO3D, enabling a more detailed analysis and visualization of vehicles, traffic lights, street signs, pedestrians, and other objects within the vehicle's environment.

*1) **Vehicles**:* We developed a system capable of not just identifying different vehicles but subclassifying them into categories such as Sedans, SUVs, Hatchbacks, Pickup Trucks, and Motorcycles. This was complemented by accurately determining their orientation, significantly improving the render's realism.

*2) **Traffic Lights and Road Signs**:* While the previous phase addressed basic traffic light and stop sign detection, Phase 2 aimed to classify traffic light arrows and indicate road signs on the ground, including arrows and speed limit signs. However, challenges were faced in displaying traffic light colors, arrows, and road signage due to limitations in processing and time constraints.

*3) **Additional Objects**:* Further granularity was achieved by identifying and rendering additional objects such as dustbins, traffic poles, and traffic cylinders as their respective 3D models in the scene.

*4) **Pedestrian Pose**:* A novel feature introduced in this phase was the detection of pedestrian poses, moving beyond simple classification to provide frame-by-frame pose analysis.

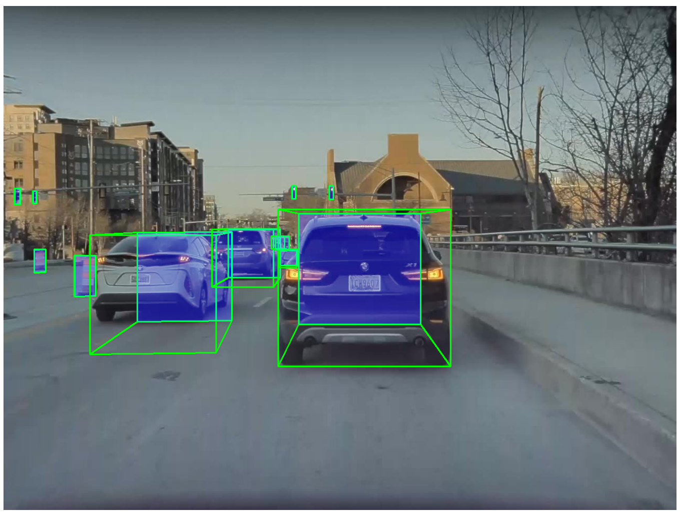

Fig. 7. Phase 2 Correctly projected Features

B. Challenges and Solutions

Despite significant advancements, Phase 2 encountered several challenges:

- **3D Bounding Box Reprojection:** Initially, the reprojection of 3D bounding boxes using the stock YOLO3D code was inaccurate. This issue was resolved by incorporating our camera matrix into the code and adjusting values in the *classaverages.txt* file to correctly bound the features of interest.
- **Rendering Issues:** Even after adjustments, rendering faced obstacles. Accurate depiction of traffic light colors, arrows, and road signage remained unresolved within the project's timeframe.

III. Phase 3

Phase 3 of the project aimed at integrating additional cognitive abilities to enhance the decision-making capabilities of the autonomous driving system. This phase focused on refining object detection to include dynamic attributes of vehicles, such as brake lights and indicators, and distinguishing between parked and moving vehicles to aid in navigation and planning.

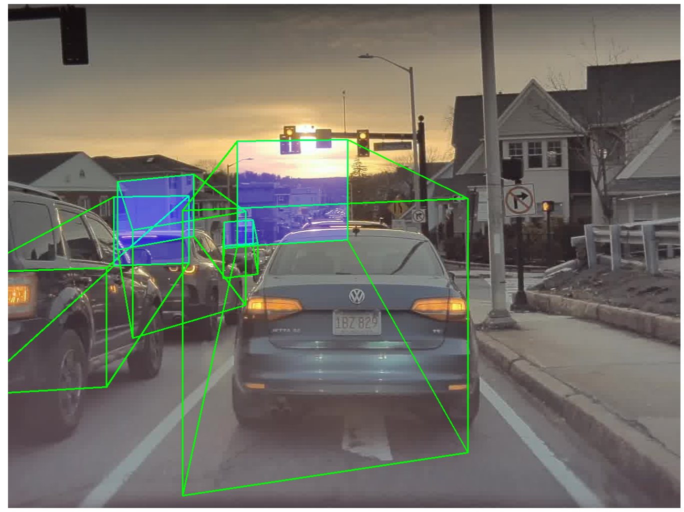

Fig. 8. Phase 2 Wrong Projection Example

A. Bells and Whistles

*1) **Improved Object Positioning**:* A significant breakthrough was achieved by enhancing object positioning accuracy. This improvement came from integrating Detic for object detection in place of the 2D detection phase in YOLO3D. This change enabled more precise identification and localization of objects within the scene.

*2) **Brake Light Monitoring**:* Through the innovative use of Detic for object detection and matching the centroid of a brake light with the bounding box of a car, we successfully implemented a system for brake light monitoring. This feature is critical for understanding the intentions of other vehicles on the road, particularly in situations requiring sudden stops or deceleration.

*3) **Lane Detection and Classification**:* Utilizing MaskR-CNN, we improved our lane detection and classification capabilities, extending our detection to include street markings. This enhancement not only contributes to the realism of our visualizations but also provides essential information for navigation and lane-keeping.

B. Challenges and Solutions

Despite the progress made, Phase 3 encountered notable challenges:

- **Parked vs. Moving Vehicles Detection:** While we aimed to distinguish between parked and moving vehicles, leveraging RAFT for optical flow analysis, time constraints prevented the completion of this feature. This remains an area for future development.
- **Indicator Detection:** Although we achieved brake light detection, Detic struggled with differentiating between indicators, brake lights, taillights, and reflectors. Given the commonality of brake lights, our system prioritizes their detection, treating other lighting features as secondary. This approach, while not perfect, allows us to capture essential vehicle behavior.

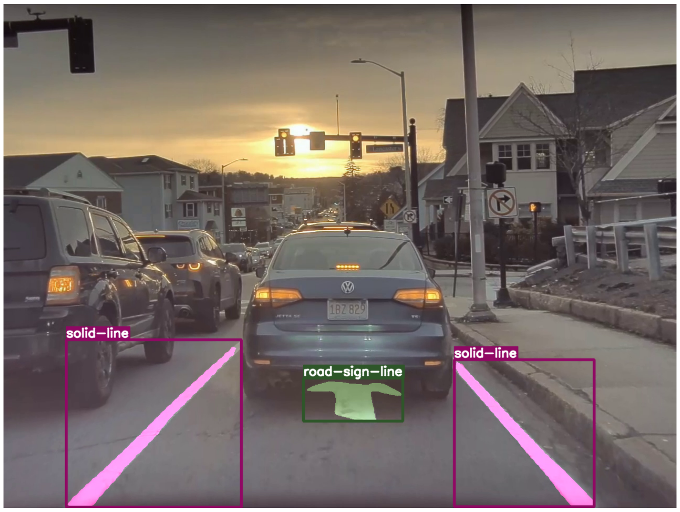

Fig. 9. Road Markings

IV. Conclusion

Phase 1 laid the groundwork for our visualization system, with significant achievements in basic object detection and depth estimation. However, the challenges encountered in 3D reprojection, lane rendering, and the simplistic representation of detected objects have provided valuable lessons. These issues will guide our improvements in subsequent phases, aiming for a more realistic and intuitive visualization experience.

Phase 2 marked a significant leap in the project's development, introducing a higher level of detail and accuracy in the visualization system. The enhancements in vehicle classification, orientation detection, and the inclusion of additional objects and pedestrian poses significantly improved the system's ability to provide realistic and informative visualizations. Despite the challenges encountered, the solutions found for 3D bounding box reprojection mark a promising direction for overcoming remaining obstacles in future work.

Phase 3 introduced significant enhancements to our visualization and detection system, particularly through the advanced detection of brake lights and improved object positioning. Despite facing challenges with the complete differentiation of vehicle dynamics and the full implementation of moving versus parked vehicle detection, the advancements made contribute to the overall goal of creating a more intelligent and responsive autonomous driving system. Future work will aim to address these limitations, further refining our system's ability to accurately interpret and visualize the complex dynamics of the driving environment.

References

[1] https://rbe549.github.io/spring2024/proj/p2/
[2] https://github.com/facebookresearch/Detic
[3] https://github.com/ruhyadi/YOLO3D/tree/main
[4] https://www.youtube.com/watch?v=Q1zqf5tfeJw
[5] https://www.it-jim.com/blog/nerf-in-2023-theory-and-practice/
[6] https://contra.com/p/L7evN73l-ne-rf-from-scratch
[7] Ben Mildenhall and Pratul P. Srinivasan and Matthew Tancik and Jonathan T. Barron and Ravi Ramamoorthi and Ren Ng,CVPR 2020, NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, 2003.08934, arXiv
[8] https://debuggercafe.com/lane-detection-using-mask-rcnn/