CSC8499 - Project and Dissertation for MSc in Advanced Computer Science

Autonomous driving: unknown object detection

Interim Report

Author:
Jiawen Yao
Advanced Computer Science

Introduction

The environmental perception ability of autonomous driving technology is greatly dependent on the environmental information acquired by various sensors, which include cameras, radars, and lidars. Among these sensors, cameras are one of the primary visual sensors. They are widely used in target recognition, lane marking detection, traffic signal recognition, etc. However, camera perception sensors are significantly affected by external physical interference factors. The occlusion of surface stains (such as raindrops, water mist, mud, frost, etc.) on the camera are directly impactful on the clarity of the image and the accuracy of the judgment by the perception system.

Previous studies have demonstrated that when there is foreign matter attached on the camera lens, the autonomous driving system is likely to miss or misidentify targets or trigger misclassification events of harmless objects as dangerous obstacles (Bijelic et al., 2019). That is, the autonomous vehicle may take incorrect acceleration, steering, or braking actions, and in the worst case, cause traffic accidents that threaten the lives of passengers. In addition, the application of nuScenes and other multimodal datasets also show that using more sensors and diversifying data can solve environmental problems (Caesar et al., 2020). Therefore, it is not only very important to recognize and model the camera lens stains to enhance the robustness of perception but also a challenging research problem to solve the autonomous driving environmental problems.

In recent years, real-time detection algorithms represented by the YOLO (You Only Look Once) series have made significant progress in the field of object detection. As the new version of this series, YOLOv8 has the advantages of high accuracy, fast speed and lightweight application, and has been verified in multiple autonomous driving subtasks (Jocher, G. et al., 2023). However, traditional target detection algorithms rarely consider the robustness problems in sensor failure or blur vision situation, especially in the "unknown type of stains" situation. Therefore, some scholars construct abnormal camera input as open set recognition (Open Set Recognition) or OOD (Out-of-Distribution) detection problems and obtain risk warning by designing a special binary classifier to judge unknown stains (Liang, S., Li, Y. and Srikant, R., 2017).

Aims and Objectives

Obtaining a single detection and classification system in which different OOD detection methods are integrated to distinguish different types of known and unknown stains on the lens by using YOLOv8 method for different types of known unknown traffic objects detection. Test the behavior of the system in known and unknown stain conditions. The system is expected to be efficient, accurate and deal with real world visual challenges in autonomous vehicle’s environment. To accomplish this aim the following objectives are defined:

- Conduct background research
  - Existing sensor stain recognition technology;
  - Existing OOD technology;
- To prepare the necessary dataset;
- To design and train a unified model based on the YOLOv8 framework capable of detecting traffic targets and identifying stains;
- Formulate stain recognition as OOD classification problems and combine mainstream OOD detection methods, such as Maximum Softmax Probability (MSP), ODIN, Energy based methods, etc.
- Design an interactive interface based on Gradio for model inference and results display.
- Analyze the robustness of the detection system in realistic autonomous driving scene by using synthetic and real contaminated data.
- Evaluation the effectiveness of OOD classifiers to distinguish in-distribution (ID) traffic targets and OOD stains anomalies.

Overview of Progress

In the first stage of the project, we have chosen the Woodscape dataset (Yogamani et al., 2019) as the basis for traffic object detection and stain recognition. The dataset offers many multi-weathers, multi-sensor images taken in real-world driving conditions and is a suitable source for training perception models for autonomous driving.

Since the YOLOv8 training pipeline requires the labels to be in YOLO format (normalized bounding box coordinates and corresponding class label), we have first converted the original Woodscape annotations to YOLO format. This pre-processing allowed us to train a single object detection model to identify common traffic targets (vehicles, pedestrians, bicycles, traffic lights) as well as stains (transparent, semi-transparent, opaque) on the lens of the camera.

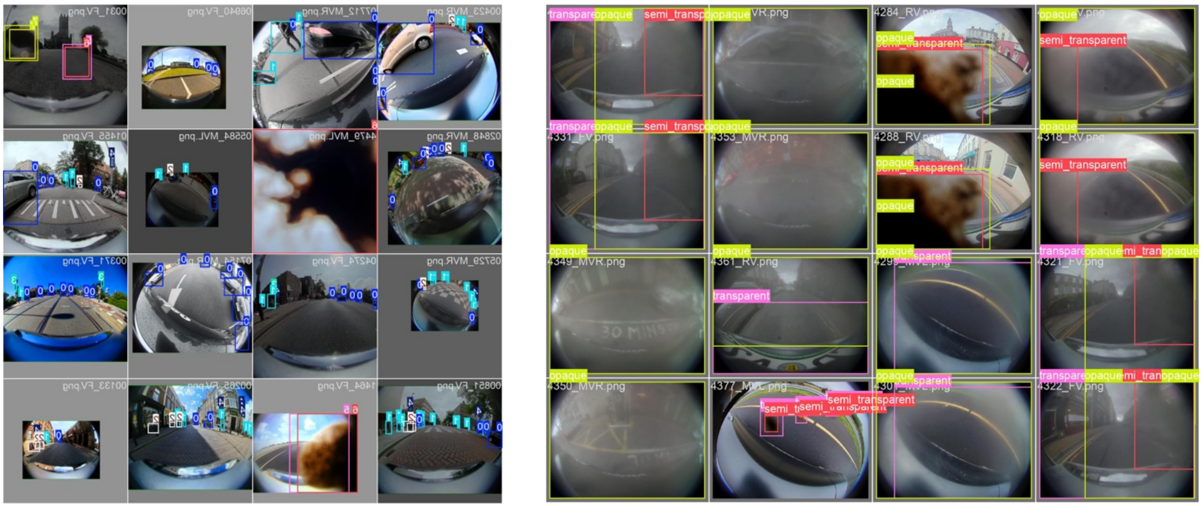

Image 1. Sample images for batch visualization

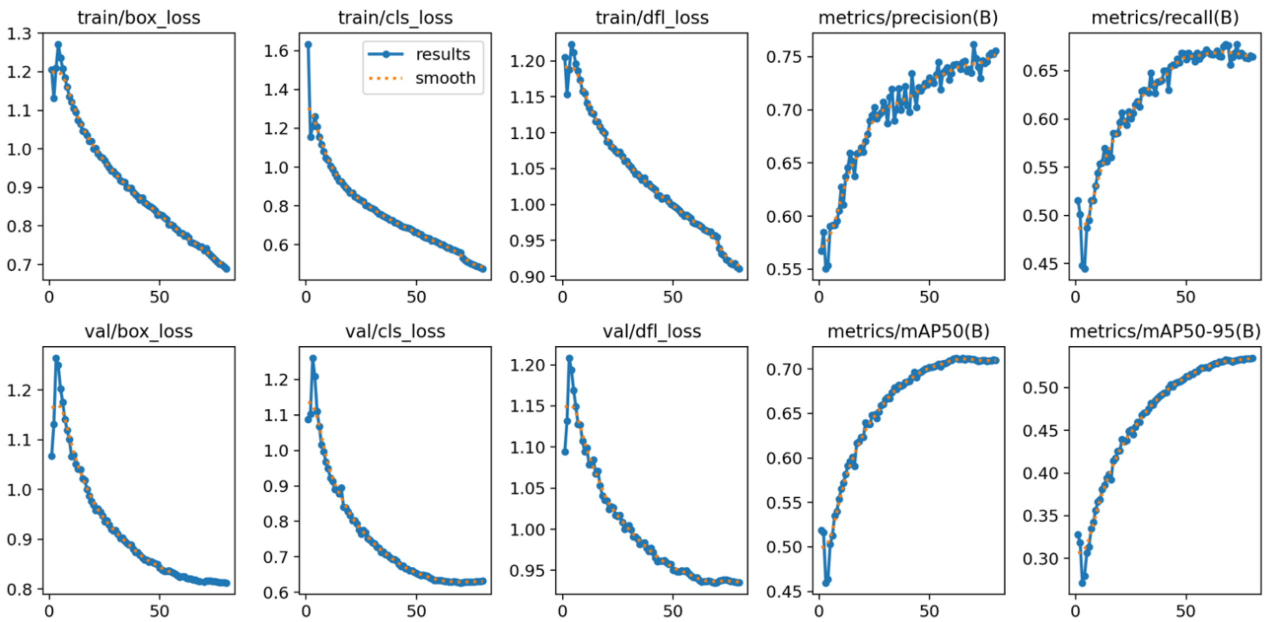

Image 2. Visual summary of the training process

Using the YOLOv8 framework, a high-performance detection model was trained, demonstrating initial success in simultaneously detecting traffic objects and camera surface stains.

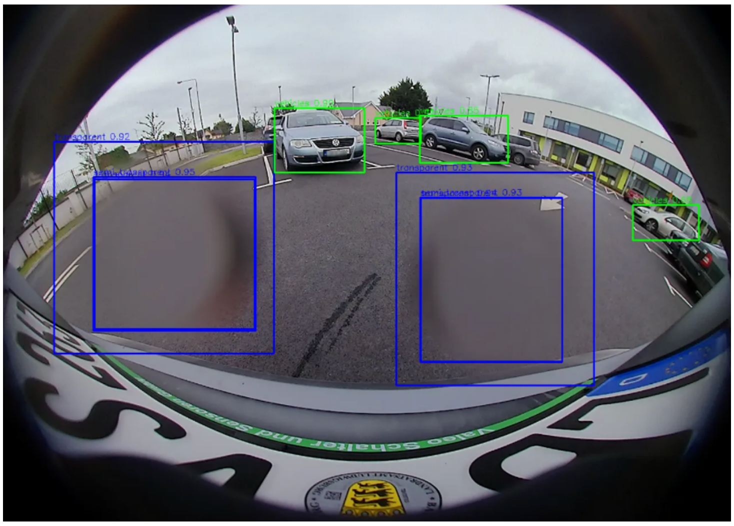

Image 3. Result of object detection

In the next project stage stain recognition will be redefined as out-of-distribution (OOD) detection task. Therefore, we train binary classifier, which can distinguish between in distribution (ID) traffic objects and out-of-distribution (OOD) stain regions. The whole approach can be represented as cropping each YOLOv8-detected object and passing it through OOD classifier for additional filtering. This two-stage pipeline raises confidence of decisions of the downstream modules and enables system to explicitly detect uncertain/anomalous inputs.

Recently, we discovered that modeling Out-of-Distribution (OOD) classification is a suitable way for stain detection due to the nature of the occlusion problem. That is, uncertainty of real-world occlusions. For example, Müller et al. (2020) modeled unknown object soiling on cameras of autonomous vehicles as an OOD problem. They found that common classification networks are not able to generalize to unseen occlusions. Therefore, using general OOD techniques is crucial in safety-critical perception tasks. By modeling stain detection in this way, the system gains the ability to explicitly detect and separate previously unseen contaminants, which would affect the overall vision-based modules. Therefore, this research direction enables the system to not only increase robustness to variable environmental conditions, but also provides a clearer uncertainty estimate for the downstream modules to make safer decisions.

To increase robustness and accuracy, the OOD classification model will implement and compare the following popular OOD detection techniques:

- Maximum Softmax Probability (MSP) (Hendrycks & Gimpel, 2017),
- ODIN (Liang et al., 2018), which uses temperature scaling and input perturbations,
- Energy-Based Models (Liu et al., 2020), which estimate uncertainty via energy scores from the model logits.logits.

These techniques will be implemented using the OpenOOD framework (Yang J et al., 2022) - a unified benchmarking and deployment suite for OOD detection methods. Evaluation metrics like AUROC, AUPR, FPR@95 will be used to evaluate effectiveness of the techniques in case of stain-related anomalies.

By using YOLOv8 and OOD classification, the project aims to create an interpretable, reliable and extensible perception system that maintains robustness under adverse conditions.

Project Plan

My project started in April 2025 and will end in August 2025. At the milestone level, the overall flow of the project will follow a waterfall structure to provide a systematic structure for the dataset preparation, model development and evaluation, while the software development and OOD classification model training are components using an agile methodology to allow for iterative training based on intermediate results and validation. This way, the project allows for systematic planning while facilitating adaptation based on experimental results.

The Gantt chart below shows the different phases and activities within the project.

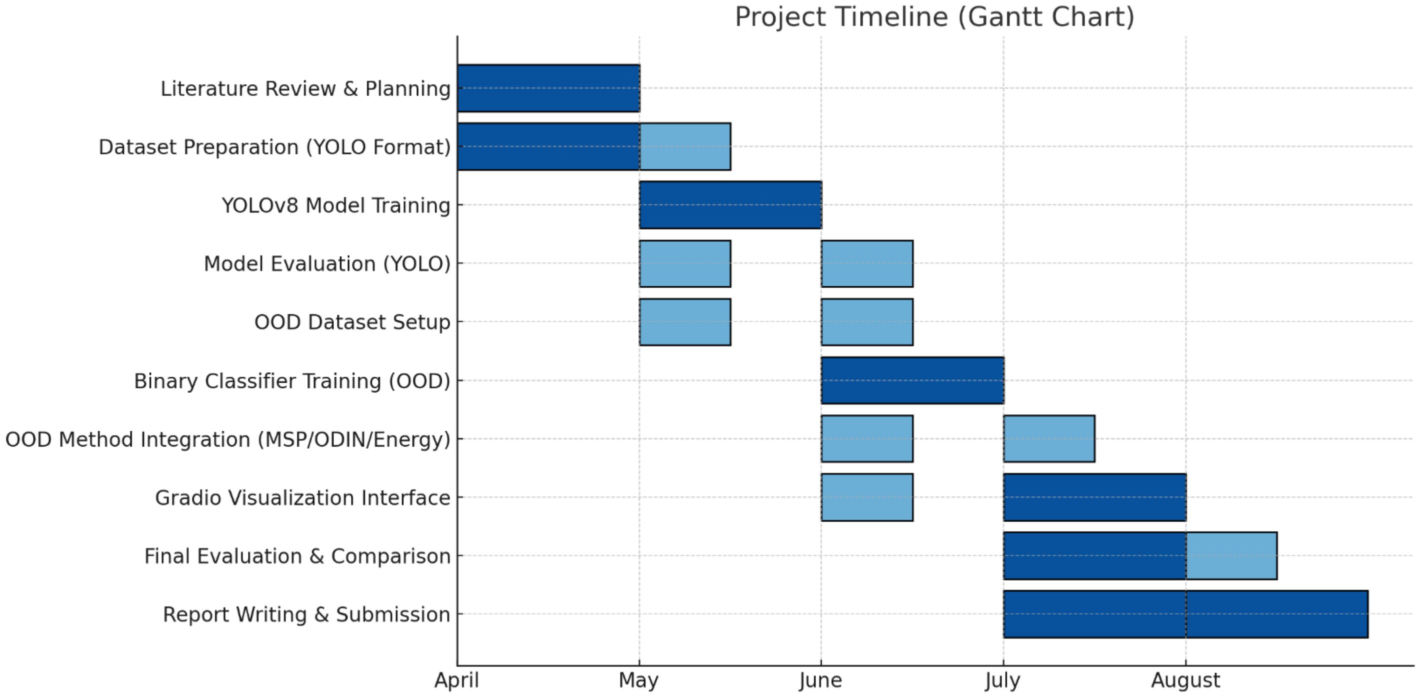

Image 3. Gantt Chart

References

Bijelic, M. et al. (2019) "Seeing Through Fog Without Seeing Fog: Deep Multimodal Sensor Fusion in Unseen Adverse Weather." Available at: https://doi.org/10.48550/arxiv.1902.08913.

Caesar, H. et al. (2020) "nuScenes: A Multimodal Dataset for Autonomous Driving," in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 11618–11628. Available at: https://doi.org/10.1109/CVPR42600.2020.01164.

Hendrycks, D. and Gimpel, K. (2016) "A Baseline for Detecting Misclassified and Out-of-Distribution Examples in Neural Networks." Available at: https://doi.org/10.48550/arxiv.1610.02136.

Kuutti, S. et al. (2021) "A Survey of Deep Learning Applications to Autonomous Vehicle Control," IEEE transactions on intelligent transportation systems, 22(2), pp. 712–733. Available at: https://doi.org/10.1109/TITS.2019.2962338.

Liang, S., Li, Y. and Srikant, R. (2017a) "Enhancing The Reliability of Out-of-distribution Image Detection in Neural Networks." Available at: https://doi.org/10.48550/arxiv.1706.02690.

Liang, S., Li, Y. and Srikant, R. (2017b) "Enhancing The Reliability of Out-of-distribution Image Detection in Neural Networks." Available at: https://doi.org/10.48550/arxiv.1706.02690.

Liu, W. et al. (2020) "Energy-based Out-of-distribution Detection." Available at: https://doi.org/10.48550/arxiv.2010.03759.

Yang, J. et al. (2022) "OpenOOD: Benchmarking Generalized Out-of-Distribution Detection." Available at: https://doi.org/10.48550/arxiv.2210.07242.

Yogamani, S. et al. (2019) "WoodScape: A multi-task, multi-camera fisheye dataset for autonomous driving." Available at: https://doi.org/10.48550/arxiv.1905.01489.