Developing Leader Follower Autonomy for Self-driving Vehicles Using Computer Vision and Deep Learning

Course-based Undergraduate Research Experience

Joseph Schulte*, Justin Dombecki, Mitchell Pleune,
Giuseppe DeRose, Nicholas Paul, CJ Chung**

Department of Math and Computer Science Lawrence Technological University, Southfield, MI 48075
* jschulte@LTU.edu ** cchung@LTU.edu

Introduction

The purpose of our project is to add humans as a participants in Leader Follower Autonomy to multi robot systems (MRS) with leader-follower and follower-follower functions [2, 3]. Little prior research & development has been accomplished in adding humans to the MRS. One of the focused goals of this project is to facilitate more versatile interactions between people and autonomous robot vehicles. A self-driving car will, at the baremost necessity, be able to see people as obstacles, taking them into account when deciding the safest actions. But the future of intelligent machinery isn't one where humans and machines operate separately, it's one where machines integrate naturally into human activity. A conceptual model of this project is shown in Figure 1 below.

Figure 1: Concept model of the project using ArUco markers

Design & Approach

At the core of this system is vision processing. The research vehicle called ACTor (Autonomous Campus TranspORt) [1] is equipped with a camera and a powerful computer with a GPU. The computer interprets what it sees using an extremely efficient deep neural network architecture called YOLO [4]. Using this model, it can quickly and accurately identify any person in its vision. This is sufficient to estimate their location and enable simple interactions, like following and stopping.

For simple behaviors, there are pre-trained YOLO networks freely available based on common datasets. These allow us to detect people, various vehicles, stop signs, animals, and several other categories of objects without any special preparation. For more advanced behavior, such as hand gesture control, we can easily train our own networks based on YOLO and integrate the new object types into our system.

To allow the vehicle to follow other objects, such as other vehicles, we use a more traditional marker-based approach. The lead vehicle is given an ArUco marker [5], a square tile with a special pattern printed on it, which the follower vehicle can recognize (also using the camera) and locate. Once it sees the marker, the rest of the following process is more or less the same as with a human. A human leader is also given the marker to uniquely identify a person to follow.

Figure 2 shows the software architecture and interactions between processes via topics based on ROS (Robot Operating System).

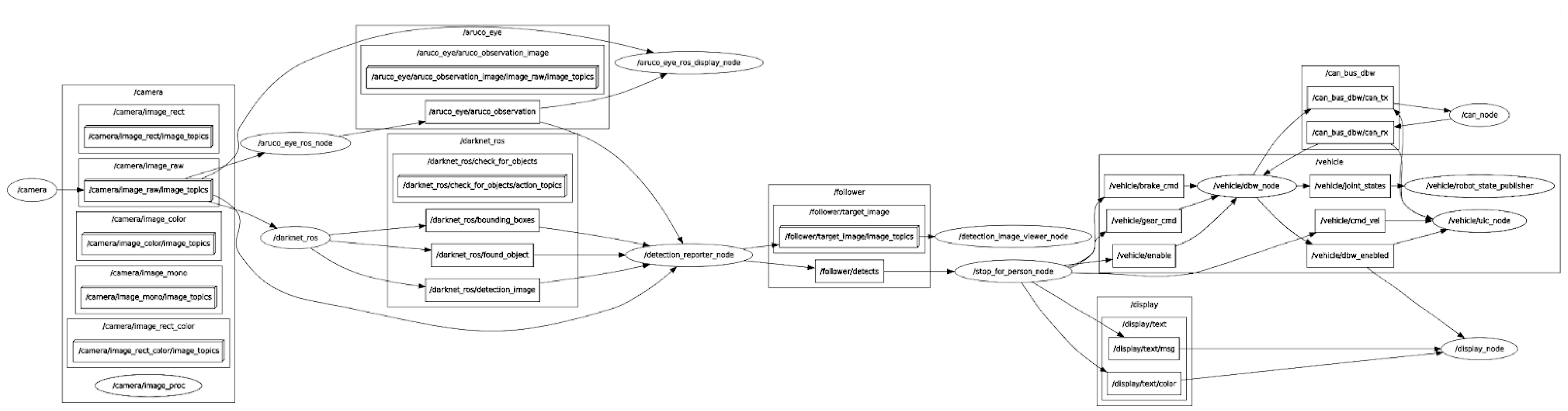

Figure 2: System software architecture - ROS node (process) computation graph.

Testing

The testing was done in the Atrium and hyper links of the three videos are: video1, video2, and video3. Figure 3 shows how we teach the system to identify a person to follow in the beginning. Figure 4 is a screenshot from our system showing how the trained deep neural network recognizes people and identifies the target person with the ArUco marker.

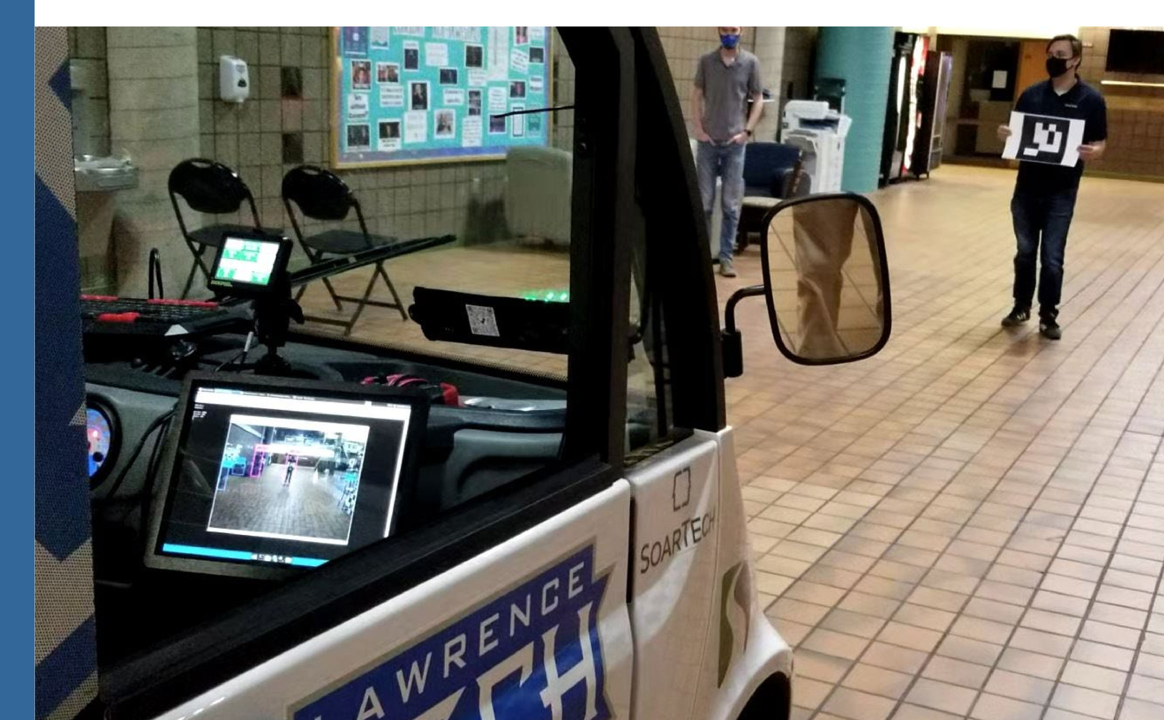

Figure 3: Identifying a person with ArUco maker

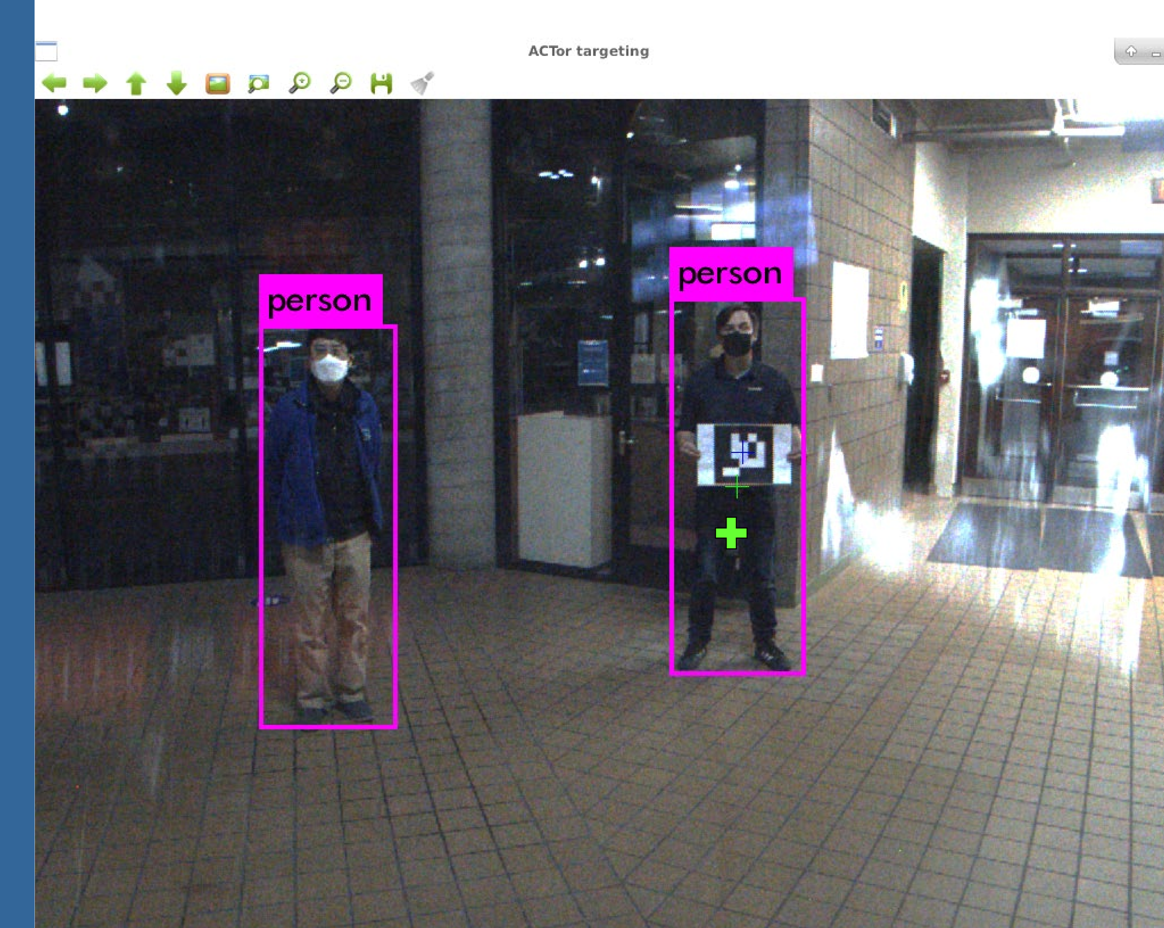

Figure 4: A screenshot from the system. A person with ArUco marker is identified with a green cross sign as a target to follow.

Acknowledgements

This project is supported by the US Army GVSC, GLS&T, NDIA-MI, Dataspeed Inc., DENSO, SoarTech, RealTime Technologies, and LTU reserve funds.

Findings and Results

The first obstacle we encountered was computer performance. YOLO is an extremely efficient architecture, but it still needs to be run on a GPU to have the throughput and latency we need. We equipped the vehicle with a dedicated GPU computer for this purpose.

The second obstacle was consistency and reliability. Lighting conditions vary wildly and there can be many people in the vehicle's vision, making it difficult to maintain consistent behavior.

To accommodate different lighting conditions we manually adjust the camera before testing. However, in the future, this can be automated. The camera setup is one of the more significant unexpected hurdles of this system.

When confronted with multiple people, the issue becomes deciding which one is the target. There are a variety of ways to accomplish this, including color traces, feature recognition, and markers. Our solution for maintaining a single target was to compute the overlap of the bounding boxes between the newest and the last known detection. The area with the greatest overlap is the target.

Another difficulty arises from trying to derive location information from the 2D camera image. YOLO gives us information on where the person or object is in the image, and using this we can find direction quite reliably but we can only roughly estimate distance. When restricted to only the camera, we use the lateral size of the person as a measure of distance.

ACTor 1's LIDAR is of particular value for this problem. While the camera provides highly detailed visual information to enable neural network-based object detection, the LIDAR provides exceedingly accurate depth information to locate objects in space relative to the vehicle. Using the direction indicated by the camera detection, we can isolate a section of the LIDAR's depth map and resolve the location to within 3cm accuracy.

References

1. Paul, N., Pleune, M., Chung, C., Faulkner, C., Warrick, B., Bleicher, S., A Practical, Modular, and Adaptable Autonomous Vehicle Research Platform, IEEE International Conference on Electro Information Technology 2018
2. https://www.roboticresearch.com/leader-follower/ (accessed on March 30, 2021)
3. B. Madhevan, M. Sreekumar, Tracking Algorithm Using Leader Follower Approach for Multi Robots, Procedia Engineering, Volume 64, 2013, Pages 1426-1435, ISSN 1877-7058, https://doi.org/10.1016/j.proeng.2013.09.224.
4. Redmon, Joseph et al. "YOLOv3: An Incremental Improvement". arXiv. (2018).
5. ArUco: https://docs.opencv.org/master/d5/dae/tutorial_aruco_detection.html (accessed on March 30, 2021)