An Autonomous Indoor Personal Robot with Real-Time Object Detection

Robotic Vision Lab

Department of Computer Science and Engineering

K. Nguyen and W.J. Beksi

Abstract

In this project, we present an autonomous indoor personal robot that can detect everyday indoor objects in real-time. Our robot is customized from a TurtleBot3 robot, with the original single-board computer (SBC) replaced by an NVIDIA Jetson Nano 2GB in order to deploy an object detector machine learning model in real-time on an onboard NVIDIA GPU. We then integrate the Intel RealSense D455 RGB-D camera along with a 3D-printed mount into the SBC for RGB and depth images acquisition during operation. We also deploy the MobileNet SSD v2 for real-time object detection with a rate of 10 FPS. To facilitate code reusability for educational purposes, we also publicize our source code and hardware design as an end-to-end tutorial. The source code and tutorial are publicly available at https://github.com/mkhangg/turtlebot3.

Introduction & Overview

- As the demand for indoor surveying robots is significantly growing, there is a pressing need for a miniature robot that can detect objects in real-time in indoor environment. To meet these needs and demands, we have developed a highly customizable indoor robot that can fulfill users' requirements based on their specific wishes.
- The overall robotic system is shown in *Fig. 1*. The robot is equipped with the NVIDIA Jetson Nano 2GB, the Intel RS RGB-D camera D455 with a 3D-printed mount, and the Slamtec RPLIDAR sensors.

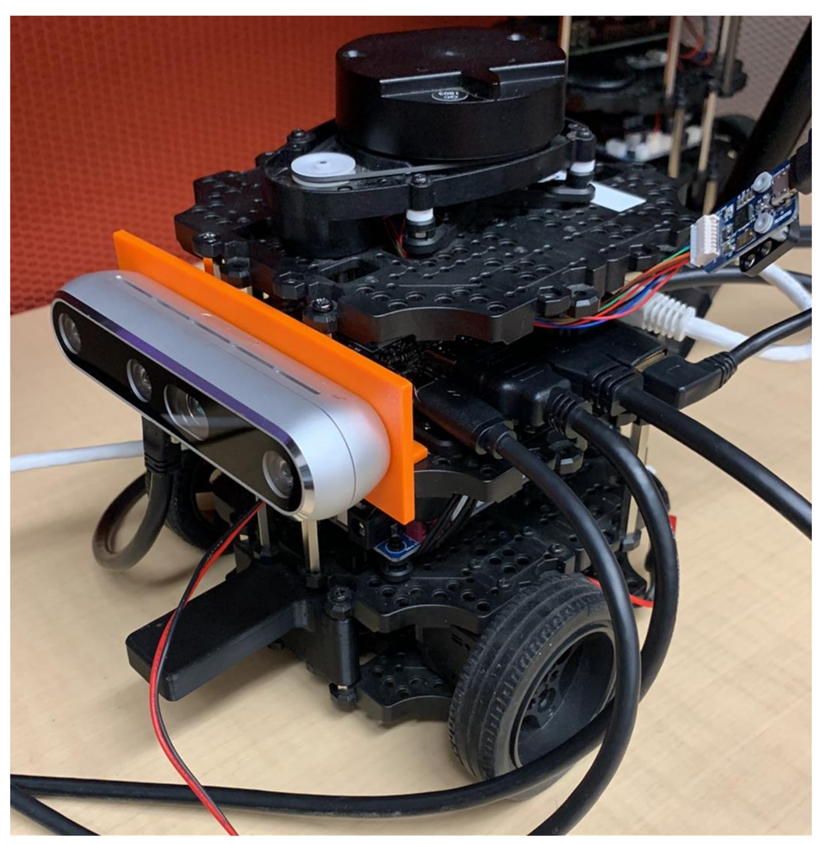

Figure 1. *Autonomous indoor personal robot.*

Hardware Architecture

- The customized robot's hardware architecture contains three layers (hardware layer, OS layer, and middle layer) with an object detector *(Fig. 2)*.
- The hardware layer consists of an Intel RealSense D455 RGB-D camera, Slamtec LiDAR sensors, a Wi-Fi dongle, and a Jetson Nano 2GB SBC.
- The OS layer consists of Ubuntu 18.04 and drivers for Intel cameras.
- The middleware layer only contains ROS Melodic.

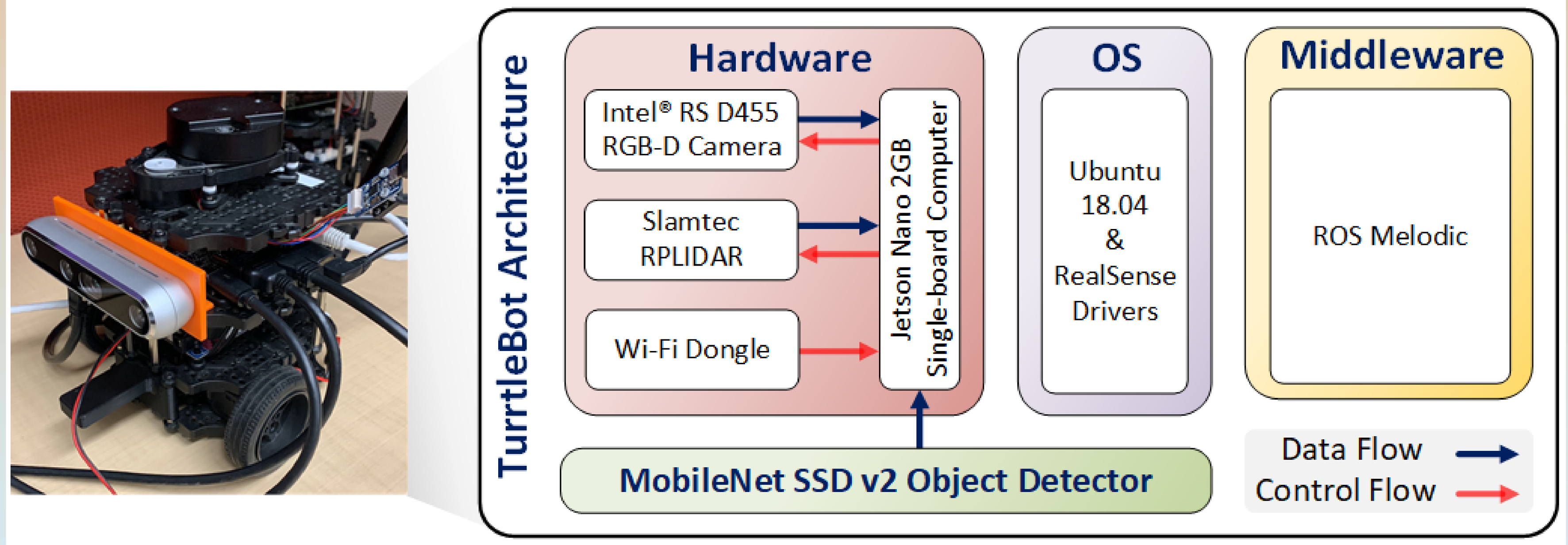

Figure 2. *TurtleBot's architecture includes a hardware layer, OS layer, and middleware layer, along with an object detector.*

RGB-Depth Images and Point Cloud Acquisition

- The robot can perform real time acquisition of RGD-D images (*Fig. 3a*) with a rate of 10 fps on the SBC's CPU, and it can also perform acquisition of point clouds (*Fig. 3b*) with a rate of 10 FPS with the same hardware configuration.
- The acquisition data is stored in the data storage of the SBC.

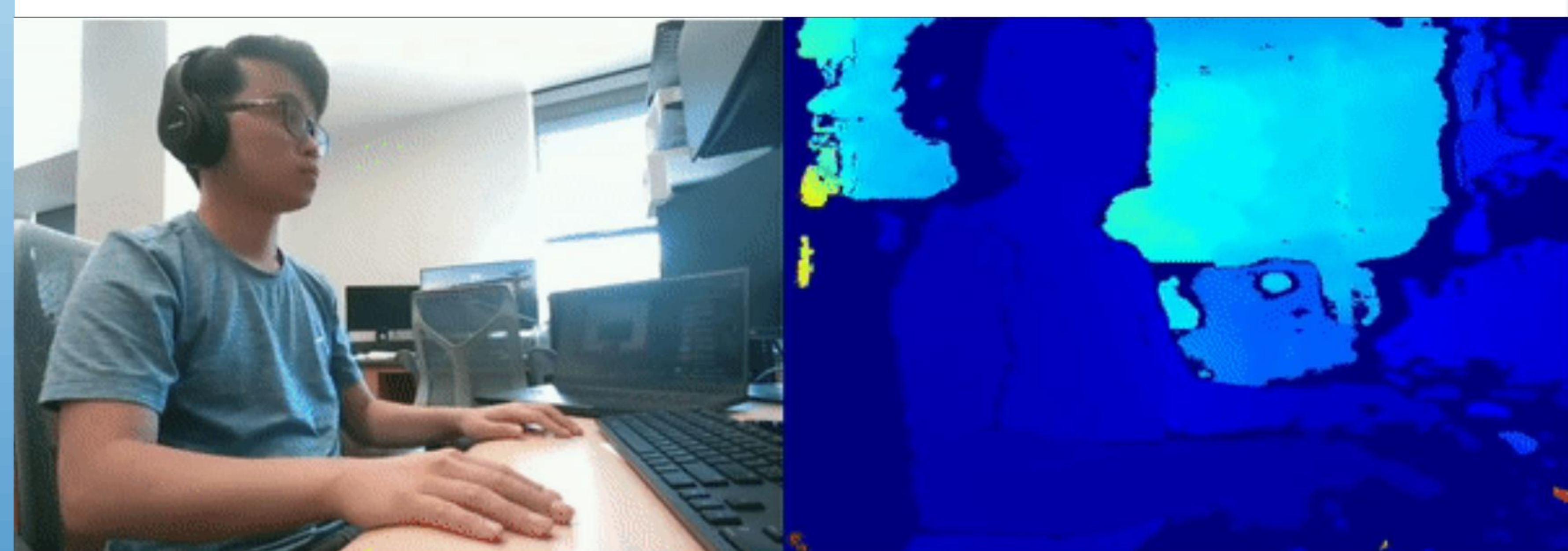

Figure 3a. *Real-time acquisition of RGB and depth images.*

Figure 3b. *Real-time acquisition of point cloud data.*

Real-time Object Detection Results

- The robot performs real-time object detection with pretrained MobileNet SSD v2 model on the onboard NVIDIA GPU with a rate of around 8-13 FPS.

Figure 4. *Real-time object detection with pretrained MobileNet SSD v2 model.*

Conclusions & Future Works

- In this project, we present an autonomous indoor personal robot that can detect everyday objects as well as humans in its field of view. The customized TurtleBot3 is equipped with an alternative NVIDIA Jetson Nano 2GB single-board computer and a mounted Intel RealSense D455 RGB-D camera. The results show that our robot can detect humans and conventional objects in real-time with a rate of 8-13 FPS. The hardware architecture for autonomous personal robots is also made public to users, together with the source code and design tutorials for designing and educational purposes.
- We reserve the tasks of 3D object detection and segmentation from the point cloud perspective [2, 3] for future work. More sophisticated hardware designs and devices, such as a custom PCB to substitute the current single-board computer or an integrated vision processing unit, are also considered [1]. Lastly, we also plan to integrate a robotic arm with a two-finger gripper that allows the robot to perform vision-based object manipulation in the future.

References

1. Dang, T., et al. “PerFC: An Efficient 2D and 3D Perception Software-Hardware Framework for Mobile Cobot,” Proceedings of the International Florida AI Research Society Conference. 2023.
2. Lang, Alex H., et al. "Pointpillars: Fast encoders for object detection from point clouds." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019.
3. Qi, Charles R., et al. "Pointnet: Deep learning on point sets for 3d classification and segmentation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2017.