Why Can’t Spot Fetch?

Colby College | Department of Computer Science

Junnan Shimizu (jshimi25@Colby.edu), Stacy Doore (sadoore@Colby.edu)

Introduction

Agile quadruped robots are becoming commercially available and used in applications from public safety to entertainment. Boston Dynamics’ Spot navigates its environment with five stereo cameras primarily used for object avoidance, but it can utilize these same sensors to recognize and classify objects in the environment through machine-learning object detection models. In our research, we are working on a proof of concept to use a Boston Dynamics Spot Explorer robot as a navigation assistant for navigators who are visually impaired.

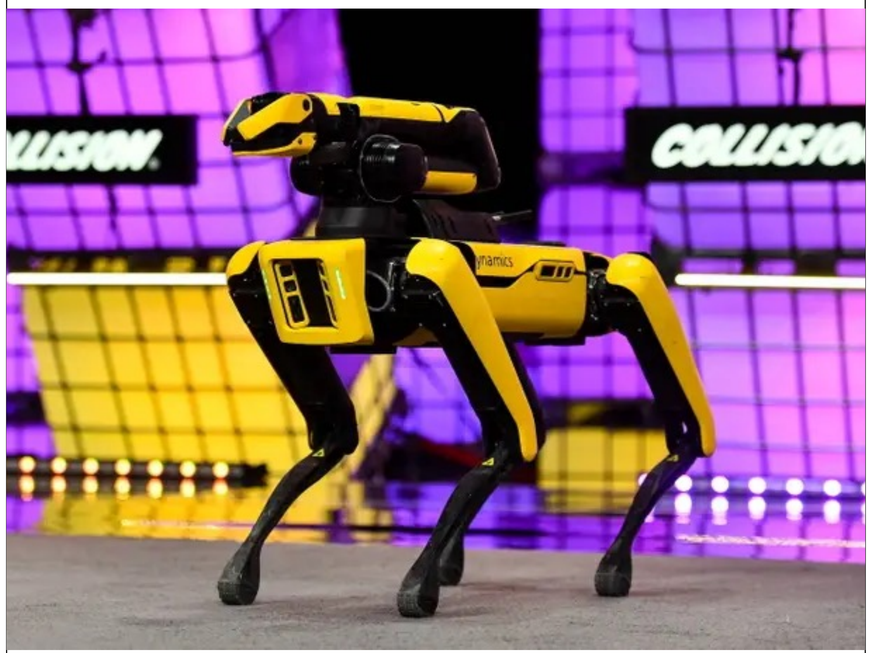
The goal of this project is to create a set of object detection/classification models that can be communicated to the robot handler, providing valuable spatial information for traversing environments including the location of stairs, doors, and emergency exits. The model for this project will also be trained to detect empty chairs, stair railings, and personal objects, depending on the individualized needs of the robot handler. This object detection component of the larger project, will be combined with a dual channel voice-based interface to communicate the robot’s actions or the presence of obstacles to the user. The primary activities for this summer research have been developing training datasets, creating object detection models, and testing different approaches for communicating spatial information collected from the robot’s sensors.

Background

Boston Dynamics’ Spot was released to the general public in June of 2020, opening the door of possibility. Early adopters used Spot in construction sites and research facilities, very quickly finding an increase in efficiency.

The Process

When Spot was released, the Spot Software Development Kit (SDK) was released alongside it, containing files to move the quadruped, control the arm, and even make it dance. The network compute bridge file was what we were after, allowing us to offload power-hungry processes to a computer on the same network. This computer is where the object detection model would run in real-time, helping dictate the decision-making process of the robot.



Roadblocks

The network compute bridge is where we ran into the first major roadblock. Although working 3 years ago, Boston Dynamics failed to maintain these example files, leaving most of the documentation and code deprecated and unusable with the newer versions of Spot and supporting programs like TensorFlow. As a result, an alternative needed to be found. We contacted Boston Dynamics about the issue, and they referred us to a fire extinguisher detector example which they claimed was kept up to date and working. Unfortunately, this program, too, had issues. Using Boston Dynamics’ own files did not seem like a viable option.

The Solution

Rather than going through Boston Dynamics’ system of sensors and software designed for Spot, the alternative would be to use an external camera and stream the information through the network to a computer on the same network. To do this, we are using a webcam and a Raspberry Pi. The video is streamed from the Raspberry Pi to a computer that is running an object detection model. Depending on the objects detected within the video, we can control Spot accordingly. For example, if an empty chair is detected, Spot can sit to signal to the user that there is an empty chair.

This solution would solve a handful of problems and limitations of Spot. For one, Spot’s five stereo cameras are black and white and low resolution, making any object detection models less accurate and in some cases impossible. Exit signs, for example, were unable to be seen, either over exposed or completely of view.

Streaming the video through a Raspberry Pi also solves the issues associated with the Boston Dynamics’ Spot SDK files. By creating our own files, we can avoid the using the included tablet entirely, simplifying the data collection and processing pipeline, in addition to allowing modern and up-to-date tools and software, future proofing our product.

Future Steps

Combining the functionality of this project with another project working with Spot and large language models would further push Spot’s functionality.

Rather than streaming the data to a computer, attaching a powerful computer like the Nvidia Jenson Orin to Spot would remove the streaming, simplifying the processing pipeline, leaving all the processing on the robot itself