Vision Language Navigation Drone

Initial Project and Group Identification Document

Senior Design
Group Three

Jazmine Roman
Electrical Engineering
Cannen Carpenter
Computer Engineering

Kevin Nilsen
Photonics Sciences and Engineering
Derek Murdza
Computer Engineering

Sponsor: **Dr. Ying Ma**

Table of Contents

1.0 Introduction

The purpose of this project is to build a lightweight autonomous drone that can listen to natural language voice commands, translate them into instructions that the system can carry out to maneuver around a specified area while, using a sensor system, determining which objects within the area to fly to as well as rotate, change altitude, and land at ground level. The group will need to construct their own drone from scratch, and integrate all of its imaging systems with their artificial intelligence(AI) they develop. Several algorithms and machine learning software will either be open source or provided by the sponsor.

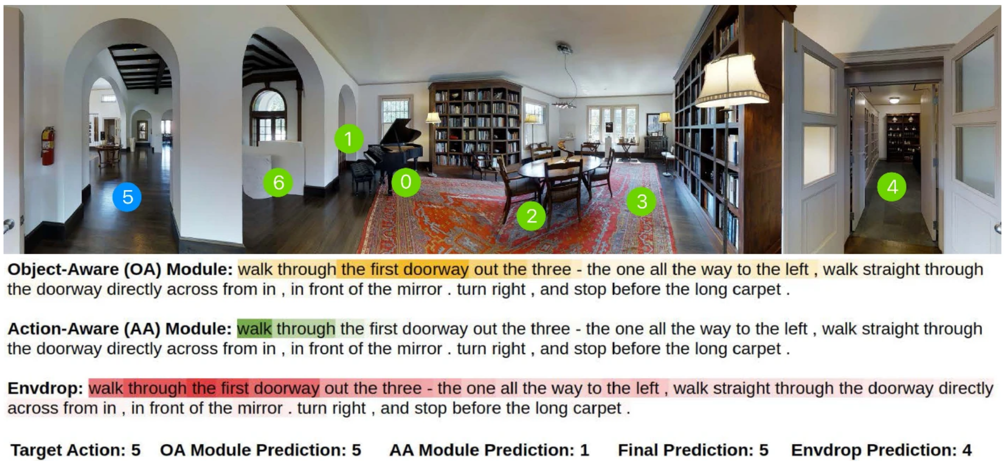


Figure 1.1: Example of logic for processing instructions from "Object-and-action aware model for visual language navigation."[1]

We have shown interest in this project as a group because of the growing field of machine learning. Machine learning has been shown as an impressive and efficient way to train AI to complete tasks of varying difficulties. Due to the complexities of directly coding object detection into a computer software program, it is not a very efficient process, and it often requires more time and effort than one gets out of it. The group has learned that machine learning algorithms can be used to, “train,” a program to do a certain objective. Through many trials, a program can more or less, “code itself,” using different connections and methods to reach a goal, like identifying a piece of furniture or vocal pattern. For every iteration, the machine learning software may have many different programs at once and test them for accuracy. The program that is the, “best,” according to the inputs one gives it is saved. The software then uses this best program, adapts it, and tries to make even more connections. Eventually, the program will be able to do the specific task it was assigned very well. The group saw this project as a great opportunity that would challenge us to learn new skills and collaborate with a diverse group to attain a satisfactory response from the sponsor.

The group had to choose a mode of transportation for the system. The sponsor gave this group a choice between a car and a quadcopter drone. Despite the drone being more complicated to construct and there being the possibility of more government regulations being a hindrance, the group chose to make the quadcopter. It should be noted that a drone requires far more time for building, testing, and adjusting than a car would, leaving the group less time to use the drone to train the AI for navigation. Since the time is only limited to two semesters, the sponsor recommended that the navigation requirements were reduced. Instead of having to navigate an entire room or hallway with a large number of potential objects and obstacles to identify, the group could instead focus on navigation around several specific objects that they choose.

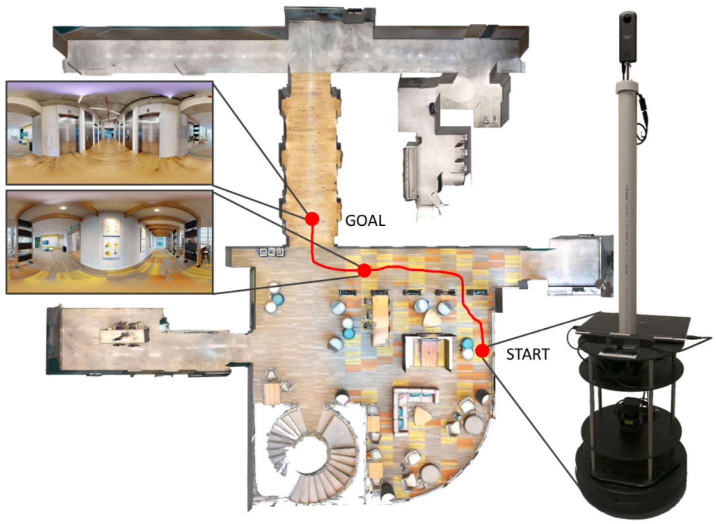


Figure 1.2: An example of the “map” to be created by the drone from "Sim-to-real transfer for vision-and-language navigation."[2] Note: The environment may look very different because our team is not required to navigate a residential environment, only one with specific objects.

2.0 Requirement Specifications

The following table presents the specifications that will be met once the drone is completed in terms of measurements, weight, capabilities, and actions:

2.0	The system should be able to autonomously fly
2.1	The system should be able to be controlled by an external computer if necessary
2.2	The system should be able to hover and maintain a single position in air for 5 minutes or more
2.3	The system should be able to navigate to a specific point within 1 meter of intended target and hover above it

2.4	The system should have sensors that will be able to detect objects 8-10m away (26 - 33ft)
2.5	The system should be able to accurately and consistently identify objects within an 8 meter radius of its sensors
2.6	The system should not have to account for new objects at runtime, as there is a designated course with obstacles that will not be altered
2.7	The system should have an audio component to send/receive voice commands
2.8	The system should be able to maneuver the pre-mapped course with a 5% error, and the un-mapped course with a 20% error
2.9	The system should have an automatic landing sequence that will be initiated if a serious error occurs
2.10	The system should only land in safe, flat areas where it will not damage neither the environment nor itself

3.0 Project Constraints and Standards

Below is a list of all of the basic project constraints

3.0	The system must have a weight of less than 3lbs
3.1	The system must have the correct permits to fly legally
3.2	The system must be able to remain stable during small gusts of wind
3.3	The system must comply with communication standards for remote microphone (i.e. bluetooth)
3.4	The system must follow the rules of the selected programming language(s) used throughout the software

3.5	The system must contain parts that are within a reasonable budget
3.6	The system must avoid putting humans in harm's way

The below list the parts need to build the drone with the parts' approximate weight:

1.0	Drone Frame	60g-140g
1.1	Four Brushless Motors	25g each 100g total
1.2	Eight Propellers (four for backup)	17.3g each set of four
1.3	Propeller protector rings	12g
1.3	Four Electric Speed Controller	4-6g
1.4	Power Distribution Board	25g
1.5	Flight Controller	20g
1.6	Battery	170g-270g
1.7	Battery Charger	30g (will not be attached to drone)
1.8	Battery Fire-proof case	40g
1.8	RC Receiver	0.47g
1.9	SD card and SD Reader	315g
2.0	Camera	734g
2.1	RC Video transmitter	453g=1lb (not attached to drone)
2.2	Sensors	5.8g
2.3	Wires	7.5g
2.4	Raspberry Pi 4B	46g
2.5	Vibration Dampening Plate (3D printed)	unknown
2.6	GPS and Compass Module	18g
2.7	GPS mount	20g

2.8	Screws	9.4g
2.9	Screw driver	N/A
3.0	Bluetooth microphone	N/A

The below list the parts needed for the training zone that will be set up in Professors Ying Ma's lab:

1.0	Green Trable
1.1	Netting
1.2	Tubes to hold up netting
1.2	City 3D model

4.0 Budget and Financing

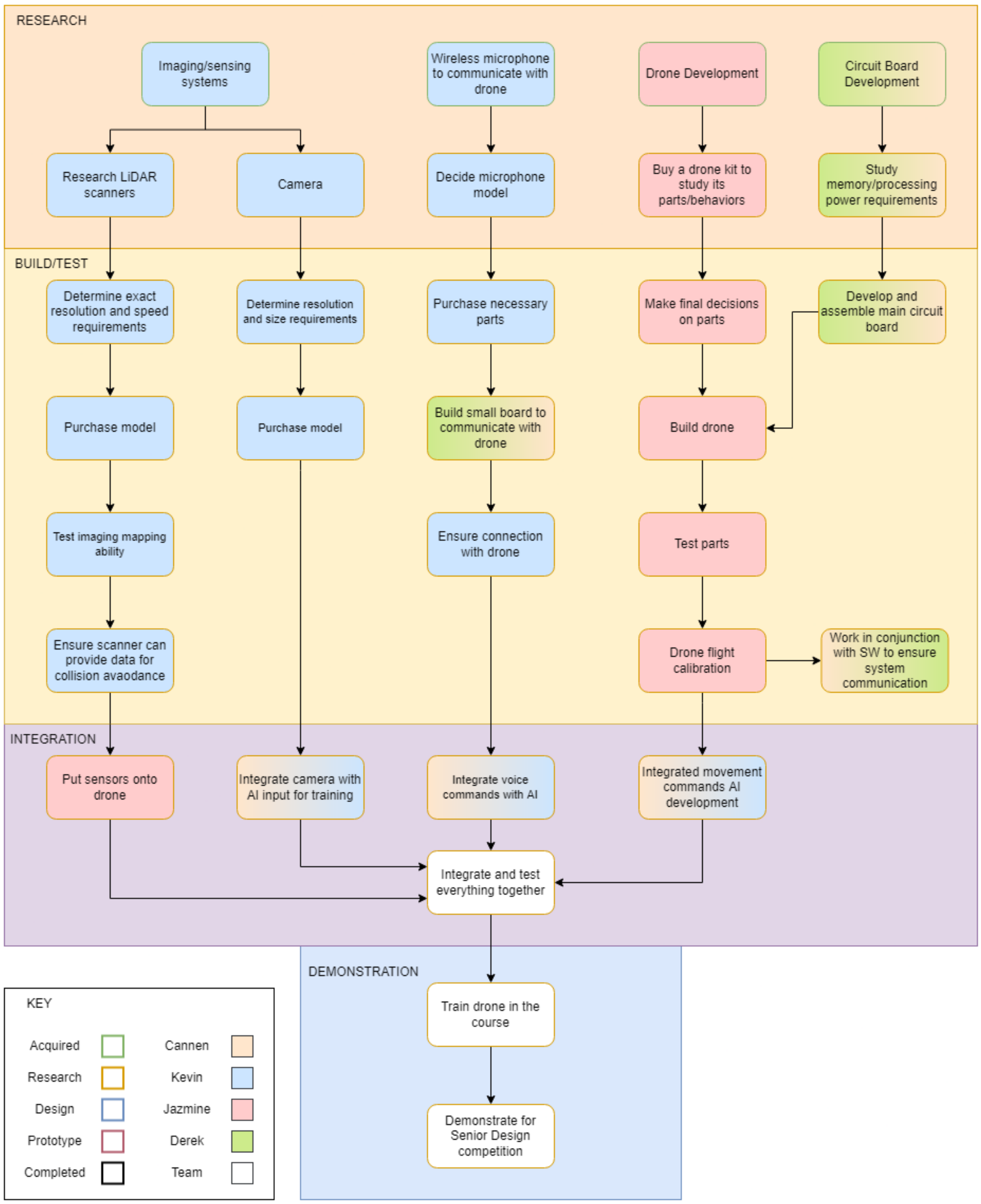
After looking through several different kits for inspiration on how to build their own, the team identified some basic components that they would need in order to build the drone itself: a lightweight frame, four brushless motors, four propellers, an electric speed controller, a power distribution board, a flight controller, and a rechargeable lithium ion battery of sufficient capacity. There will also need to be several other pieces of equipment that will be used for the drone's sensing and computing abilities. These other pieces include a receiver (for the wireless microphone signal), a camera (for object identification), some type of video transmitter (to send the video data back to a computer on the ground), and an on-board processing unit that can collect data from all the devices,, and be able to store that data as well as the AI's navigation program.

The team estimated the cost of their own drone to be similar to that of a mid range consumer drone kit, which can be several hundred dollars.The team has also analyzed lower budget mapping LiDAR scanners, whose costs can come as low as $100. With all of this in mind and with consultation from the sponsor, the team came up with a total estimated budget of less than $1000.

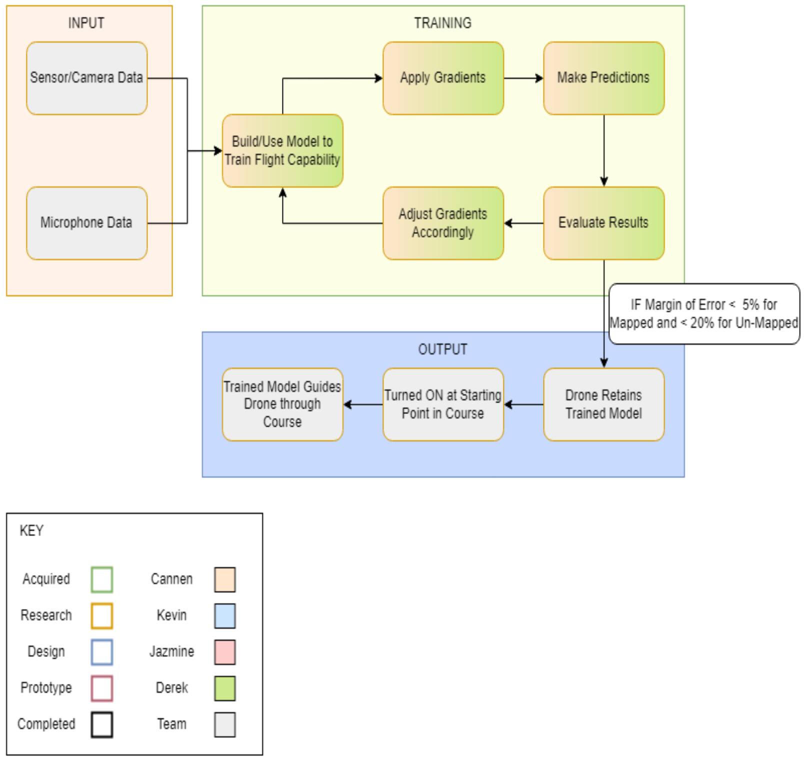
The team's generous sponsor, Dr. Ying Ma, has confirmed that they will completely finance any of the group's hardware needs within reason, They have defined a specific process to acquire any funding they may need. According to the sponsor's terms, in order to get funding for specific parts, the group must first submit a proposal to the sponsor. This should include the price of the part, the date that the part is needed by, and most importantly, detailed reasoning and justifications as to why the team needs the part specified. The sponsor will then review the proposal and promptly respond to the team with whether or not they will choose to allow or deny funding for the part. The team has decided that no further research into additional sponsorships is needed.

5.0 Project Block Diagrams

5.1 Hardware of VLN Drone



5.2 Software of VLN Drone



6.0 Decision Matrices

Primary Programming Language Selection					
1 = worst 5 = best	Difficulty	Familiarity with Language	Motivation/ Excitement	Score	Comments
Python	4	5	5	14	The machine learning algorithm primarily being used is 'PyTorch,' which is based in python
C/C++	2	3	3	8	Somewhat familiar, and has more compatibility/similarity with major microcontrollers

Drone Flight Object Detection System						
1 = worst 5 = best	Cost	Difficulty	Familiarity with Technology	Motivation/ Excitement	Score	Comments
Simple Laser Distance Detection	4	4	4	1	13	Only detects how far the drone is from an object in front of it. No mapping capability.
360 Degree LiDAR Scanner	1	3	4	5	13	More expensive, but allows us to create a comprehensive map of the area the drone will be moving around. Would lead to a higher success rate.
Camera Image Analysis	3	3	4	2	12	More difficult mapping. Also requires more AI training to be able to "see" when objects are getting larger and therefore closer.

7.0 Project Milestones and Deadlines

September 2022 - At the beginning of the semester, the group must become familiar with methods of training a drone and come to a conclusion as to which tactics would be optimal for completing the project goal. The group must decide on specific aspects of the project to build on: which constraints, specifications, software, and machine learning algorithms to be used. They must use this information to complete the initial project documentation.

October 2022 - By the middle of the first semester, on October 11, the sensing and imaging systems should be in the prototype stage, enough for the mid semester demonstration of its optical components. This is a required task for the photonics sciences and engineering student in the group. An in-person demonstration must be done to show the work they have done so far in the semester. At this time, the team should also have purchased a drone kit to study and use to adapt their own drone design. Initial designs of the done should soon be completed.

November 2022 - Another in-person demo is required of the photonics sciences and engineering student in the group. This in-person demo will contain a more mature prototype of the systems demonstrated previously, and possibly another new system if one is created or ready to present. At this stage, the finalization of the parts list should be nearing completion. A proposal to the sponsor should be crafted to explain the necessary parts and final total cost. Also, the team should begin constructing the safety cage that the drone will be tested in. This cage will be located inside of a lab that the sponsor has provided them.

December 2022 - After the first semester is complete, the optical system should be prototyped and be functional enough to formally begin the training of the AI. Also, all of the parts necessary for the drone to fly should be identified, and a proposal should have already been submitted to the sponsor to ensure delivery of most of the parts before the next semester. Supply chain disruptions have been quite common in recent years, so the team should plan accordingly. They must act as if all of their parts may be delayed for a long time, so parts should be chosen well before they are needed. The group must begin assembling the drone no later than January 2023. There must be ample time to receive the parts and construct the drone to ensure there is plenty of time for the AI to be trained and integrated.

January 2023 - The final designs of the drone are reviewed and drone assembly should begin immediately with the parts that should have already arrived. The beginning of the on-board AI should also begin development using the imaging systems that had been prototyped in the previous semester.

March 2023 - By the middle of the second semester, the drone should be fully capable of hovering and maneuvering on its own without any external supports (i.e. ropes and/or remote controllers). The AI should have been well underway with its training as well, and steps should be taken to ensure everything can be integrated together.

May 2023 - By the end of the second semester, the onboard artificial intelligence should be fully trained using provided machine learning algorithms, and it should be fully integrated onto the drone. The drone should be able to easily recognize its predetermined objects, and take a person's speech and translate it into instructions that the drone can execute. A full demonstration should be ready for the senior design competition.

8.0 References

1. Qi, Yuankai, et al. "Object-and-action aware model for visual language navigation." *European Conference on Computer Vision*. Springer, Cham, 2020 https://link.springer.com/chapter/10.1007/978-3-030-58607-2_18#citeas.

2. Anderson, Peter, et al. "Sim-to-real transfer for vision-and-language navigation." *Conference on Robot Learning*. PMLR, 2021 https://arxiv.org/pdf/1904.04195.pdf.

3. Tan, Hao, Licheng Yu, and Mohit Bansal. "Learning to navigate unseen environments: Back translation with environmental dropout." *arXiv preprint arXiv:1904.04195* (2019) https://arxiv.org/pdf/1904.04195.pdf.

4. Zhu, Wanrong, et al. "Diagnosing vision-and-language navigation: What really matters." *arXiv preprint arXiv:2103.16561* (2021) https://arxiv.org/abs/2103.16561.

5. Wang, Hanqing, et al. "Active visual information gathering for vision-language navigation." *European Conference on Computer Vision*. Springer, Cham, 2020 https://link.springer.com/chapter/10.1007/978-3-030-58542-6_19.