ECE 445

SENIOR DESIGN LABORATORY

PROJECT PROPOSAL

Semantic Communications for Unmanned Aerial Vehicles

Team #25

YU LIU (yul9@illinois.edu)
CHANG SU (changsu4@illinis.edu)
CHENHAO LI (cl89@illinis.edu)
TIANZE DU (tianzed2@illinis.edu)

Sponsor: Meng Zhang
TA: Xiaoyue Li

April 3, 2023

Contents

1 What Modifications We Made **1**

2 Introduction **2**
2.1 Problem . . . 2
2.2 Solution and Visual Aid . . . 2
2.3 High-level requirements list . . . 3

3 Design **4**
3.1 Block Diagram . . . 4
3.2 Subsystem Overview . . . 4
3.2.1 UAV mechanical, balance and dynamic Subsystem (UAVS) . . . 4
3.2.2 Lighting Semantic Extraction Subsystems (LSES) . . . 5
3.2.3 Mutual Communication Subsystem (MCS) . . . 5
3.3 Subsystem Requirements . . . 5
3.3.1 UAV mechanical, balance and dynamic Subsystem (UAVS) . . . 5
3.3.2 Lighting Semantic Extraction Subsystems (LSES) . . . 6
3.3.3 Mutual Communication Subsystem (MCS) . . . 6
3.4 Tolerance Analysis . . . 6
3.4.1 UAV mechanical, balance and dynamic Subsystem (UAVS) . . . 6
3.4.2 Lighting Semantic Extraction Subsystem (LSES) . . . 6
3.4.3 Mutual communication subsystem (MCS) . . . 7

4 Ethics and Safety **8**
4.1 Ethics . . . 8
4.2 Safety . . . 8

References **9**

1 What Modifications We Made

In this part, we'll give you an overview of what feedback we received and what modifications we've made to the previous proposal.

Feedback 1: For subsystem requirements, the comment said that we should add more detailed requirements with quantitative measurable values.

Modification 1: To achieve it, we look up some relevant hardware parameters and add many specific subsystem criteria. At the same time, the requirement for off-the-shelf components has been removed.

Feedback 2: For tolerance analysis part, according to feedback, we lacked mathematical analysis or simulation.

Modification 2: After communicating with our TA, we do simulations and run the YOLO network on our own computer to detect objects to confirm the feasibility of our project.

Feedback 3: For ethics and safety part, the one TA's feedback thought that we do not have references for this part and one point was deducted.

Modification 3: We did have references about ethics and safety. To make this clearer, we add more references.

Feedback 4: We also received feedback from the writing TA. We were not reduced points but TA gave us some useful advice.

Modification 4: Based on the advice, we add our sponsor on the cover page, split the long paragraph of the introduction section into several short paragraphs, and write full words in small letters and bracket abbreviations for all abbreviation words we used.

Modification 5: For subsystem overview, though we did not receive any feedback, we modify this part according to our understanding of the rubric. We add some details, clarify the composition of the subsystem, and determine the scenarios, training networks, and datasets for which the drone is applicable.

The main changes we made are bolded in the text.

2 Introduction

2.1 Problem

Existing communication systems are mainly based on Shannon's information theory, and they are mostly developed to maximize data-oriented performance indicators, such as communication data rate, while ignoring content-related information or considering only upper-level information [1]. In the process of transmission, once the noise makes some bits in the process of transmission wrong, the meaning of the transmission result will become ambiguous [2]. This not only increases the demand for communication resources but also limits the transmission rate of information.

In this case, people start to think about using semantic communication. Semantic communication breaks through the traditional theoretical framework of Shannon's information theory, making great breakthroughs in reducing communication loss, improving transmission rate and accuracy, and transforming the content of communication into the meaning of information more valuable to human beings, thus fundamentally transforming the existing communication architecture into a more universal intelligent and human-oriented system [3]. In semantic communication, the meaning of the result can be roughly predicted even if some bits are disturbed by interference.

And now, most of the existing image semantic communication technologies rely on the direct transmission of the whole image between the sender and the receiver [4]. The transmission process rather than the semantic understanding algorithm is the bottleneck of its performance. In real life, there are many scenarios that require the unmanned aerial vehicle (UAV)'s overlooking function and the UAV's direct real-time communication with other intelligent devices.

The UAVs currently on the market can only fly and take pictures, and transmit the pictures to the receiver using traditional means of communication [5]. But in many cases, the direct transmission of images from UAV is a huge waste of power and transmission. So, our goal is to develop an unmanned aerial vehicle (UAV) technology that allows the UAV to transmit images using semantic communication. More specifically, our UAV can build on the capabilities of existing UAV to process a sample of the image taken, extract specific semantics, and convey its symbolic representation to the target receiver (for example, another UAV or smart device). In this way, all we need to transmit is a sentence instead of a whole picture. We hope this technique will be much faster than transmitting each complete image directly.

2.2 Solution and Visual Aid

First, we'll assemble a powered UAV, complete with batteries and controls, as well as a camera, a transmission module, and Raspberry PI. The controller will control the four propellers used to control the UAV's movement. The controller ensures that the UAV has enough power to carry all of its equipment and is always balanced. During the flight of the UAV, we use the camera of the UAV to shoot images, use the Raspberry PI to process images, and extract the semantic features in the images. Then the encoder encodes these features into digital signals and transmits them to the receiver. The receiver is another smart device, such as a smartphone or computer. The digital signal is transmitted through a physical channel, mainly through a WiFi transmission module. Finally, the receiver's decoder can translate the bits back into semantic information to make sense of the message. We will use a display screen to display the message received by the receiver.

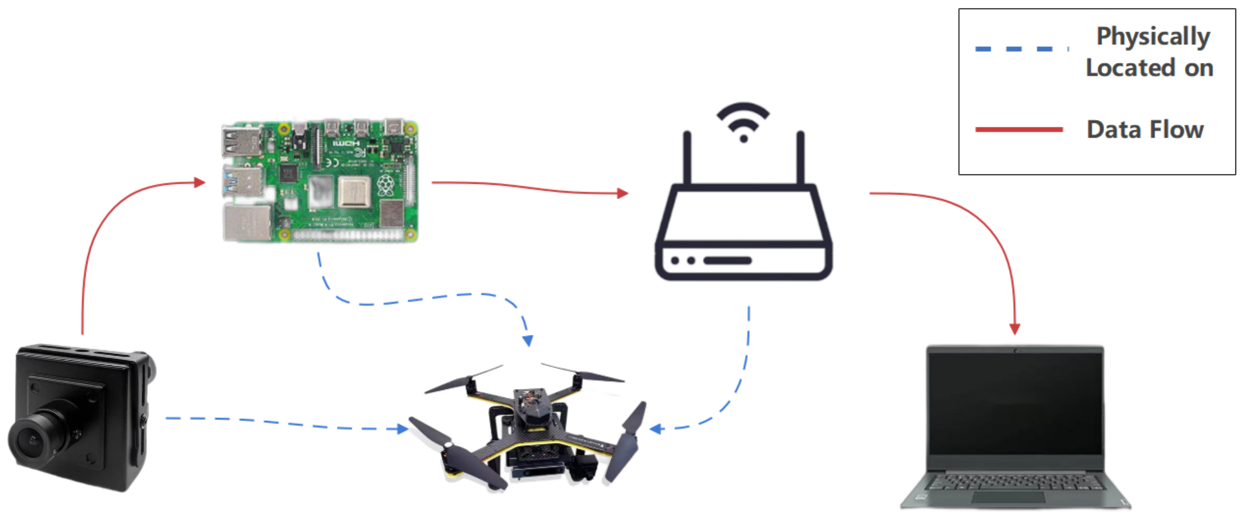

Figure 1: Visual Aid: Data Flow and Physical Location

2.3 High-level requirements list

- The UAV must be able to carry cameras and microcomputers, such as the Raspberry PI, to move around. The UAV can hover up to at least 5 meters in the air and take sample images with a resolution of 1920 by 1080.
- The UAV must understand its images and extract useful information (semantics), especially the number of people and even the types of their behavior. The UAV should be able to predict the number of people with 80% accuracy, and predict the types of behavior with 70% accuracy.
- The UAV must transmit semantic information to receiver successfully. The faster transmission speed than the traditional communication method and stronger anti-interference ability is highly appreciated. The time required to transfer each picture should be less than 0.5s.

3 Design

3.1 Block Diagram

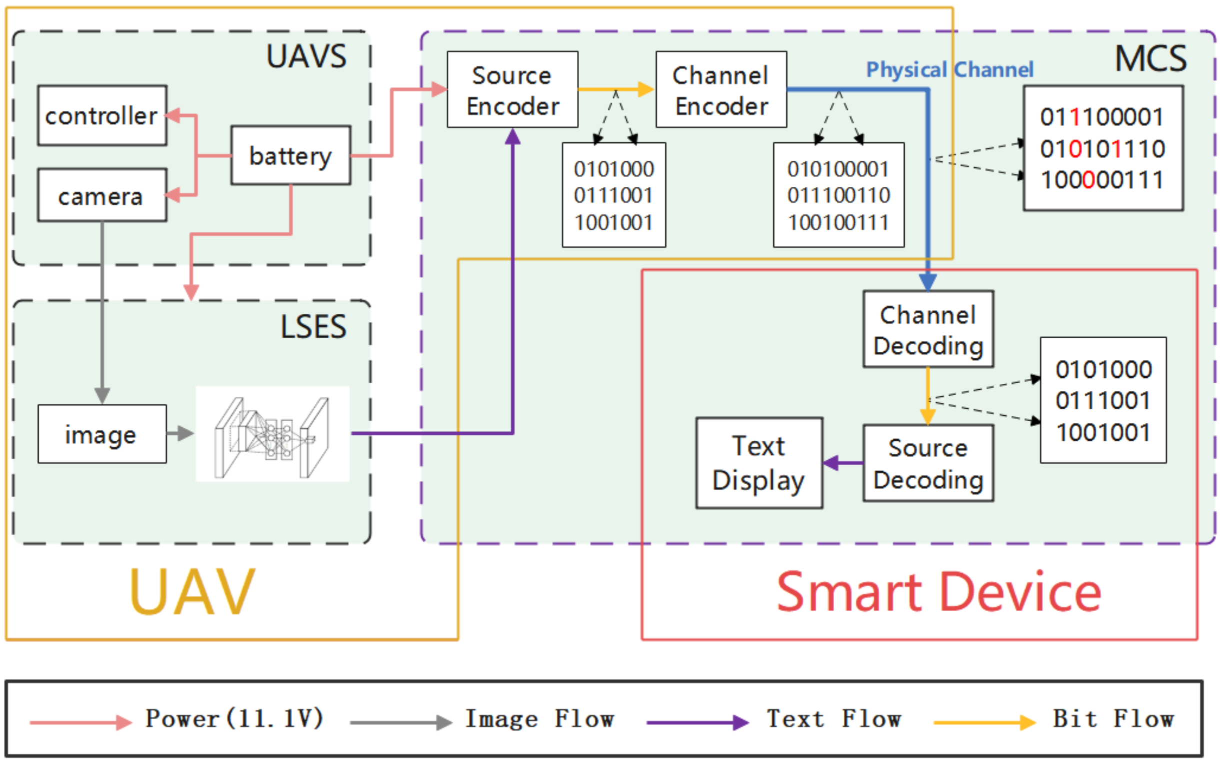

Figure 2: System Block Diagram

3.2 Subsystem Overview

3.2.1 UAV mechanical, balance and dynamic Subsystem (UAVS)

The UAV subsystem includes a power module, a controller module, and a camera. The power module, which **includes a distributor board and a lithium battery**, is to supply stable voltage for every device and other subsystems on UAV. **The controller module will receive signal and control the four propellers, which are used to control the UAV's movement.** It is also important to ensure that the UAV has enough power to carry all the devices and keep them balanced all the time. The camera will take images with high resolution, which will be used as input data for lighting semantic extraction subsystems (LSES). We will modify a dumb UAV. Equipment likes, a camera, Raspberry Pi, and a communication module needs to be added.

3.2.2 Lighting Semantic Extraction Subsystems (LSES)

Given an image as input, LSES can be designed to analyze the image and generate a descriptive sentence that includes the estimated number of people present in the image through the network, as well as identifying the actions or activities being performed by each individual within the scene. LSES could utilize advanced computer vision algorithms, for example, YOLOv5s [6], and machine learning techniques to accurately detect and identify the people and their actions in the image, providing valuable insights for a range of applications. **Our project will focus on some specific domains, especially some common school sports, such as running, cycling, basketball, football, badminton, and so on. We will use some public datasets, for example, UCF101 [7], to train and optimize our network. UCF101 is a motion recognition dataset of realistic action videos, providing 13,320 videos from 101 action categories.**

To be useful in real-world scenarios, the semantic extraction system should be designed to be robust enough to tolerate noise and other types of visual interference that can be present in images. To achieve this, the system could incorporate various techniques, such as image denoising, feature detection. It can provide more reliable and accurate estimates of the number of people present in the image and their actions. The semantic information extracted by LSES will serve as the input of the mutual communication subsystem (MCS).

3.2.3 Mutual Communication Subsystem (MCS)

MCS accepts the text information extracted from images by LSES. This subsystem converts text into a bits signal and transmits it to another smart device over a physical channel. MCS consists of two separate parts: the transmitter on UAV and the receivers on smart devices, which are connected by the physical channel. The transmitter includes a source encoder which transfers text into a bit sequence and a channel encoder which adds another few bits to avoid transmission mistakes. Correspondingly, the receiver has two decoders, a channel decoder and a source decoder, which will convert the bits signal back to text. In the process of transmission, some bits may become wrong due to interference, which is represented by red bits in the block diagram. The channel encoder and decoder can correct these mistakes automatically. And finally, the text information will be displayed on a screen. The text example could be like "There are 15 people playing basketball."

3.3 Subsystem Requirements

3.3.1 UAV mechanical, balance and dynamic Subsystem (UAVS)

The dumb UAV needs to be modified. Equipment likes, a camera, Raspberry Pi, and a Communication Module needs to be added. We need to make sure that the UAV can still fly smoothly and quickly after adding these devices. **The UAV is expected to fly in a circle with a radius of 5m through the air and quickly rise into the air and hover.**

3.3.2 Lighting Semantic Extraction Subsystems (LSES)

LSES needs to understand image information, for example, the number of people and their behaviors on images **at a high accuracy of at least 70%**. And finally, summarize information into text, for example, "15 people are playing basketball", and send it to the mutual communication subsystem (MCS). **The LSES should be small but efficient, which can be carried on UAVs and extract information quickly for at most 0.5s for each photo.**

3.3.3 Mutual Communication Subsystem (MCS)

MCS needs to successfully transfer text information from UAV to smart devices. The text should have similar semantic information but not necessarily the same words. This text will finally display on a display screen. The text could be like "There are 15 people playing basketball." **The communication should be quick and without losing semantic information. The time for transmission should be less than 0.5s for each sentence with about 10 words.**

3.4 Tolerance Analysis

3.4.1 UAV mechanical, balance and dynamic Subsystem (UAVS)

The most important part of this subsystem is the battery. The insufficient power supply of the battery will lead to slow processing of all parts, and too large power supply voltage will make the UAV's power too high, or even spontaneous combustion. Faced with this problem, we decided to set the voltage at 11.1V, so that the power of the drone would not be too high, and timely alarm when the battery level was below 50%, so that the drone would not crash due to loss of power.

3.4.2 Lighting Semantic Extraction Subsystem (LSES)

As for the LSES, motion blur may occur when there is a relative movement between the camera and the object being captured. This problem can be reduced using techniques such as deconvolution [8], which attempts to reverse the blurring effect by estimating the motion blur kernel and applying an inverse filter to the image. Also, lens distortion may occur when the lens of the camera does not produce a perfect image, resulting in image distortion or aberrations. This problem can be corrected using techniques such as barrel distortion correction [9] or perspective correction [10]. In this task, We have less space and computing power to process images since we have to place the processor on the drone. In this regard, we build models that are as efficient as possible and focus on specific application scenarios. If necessary, we may need to purchase more advanced processors.

To show the feasibility of using YOLOv5s model on semantic tasks, we first ran YOLO model on simple baseline tasks like object detection (OD) using MIO-TCD Dataset[11].

The confusion matrix obtained below clearly demonstrate the excellent performance YOLO has on classifying objects with similar scenarios.

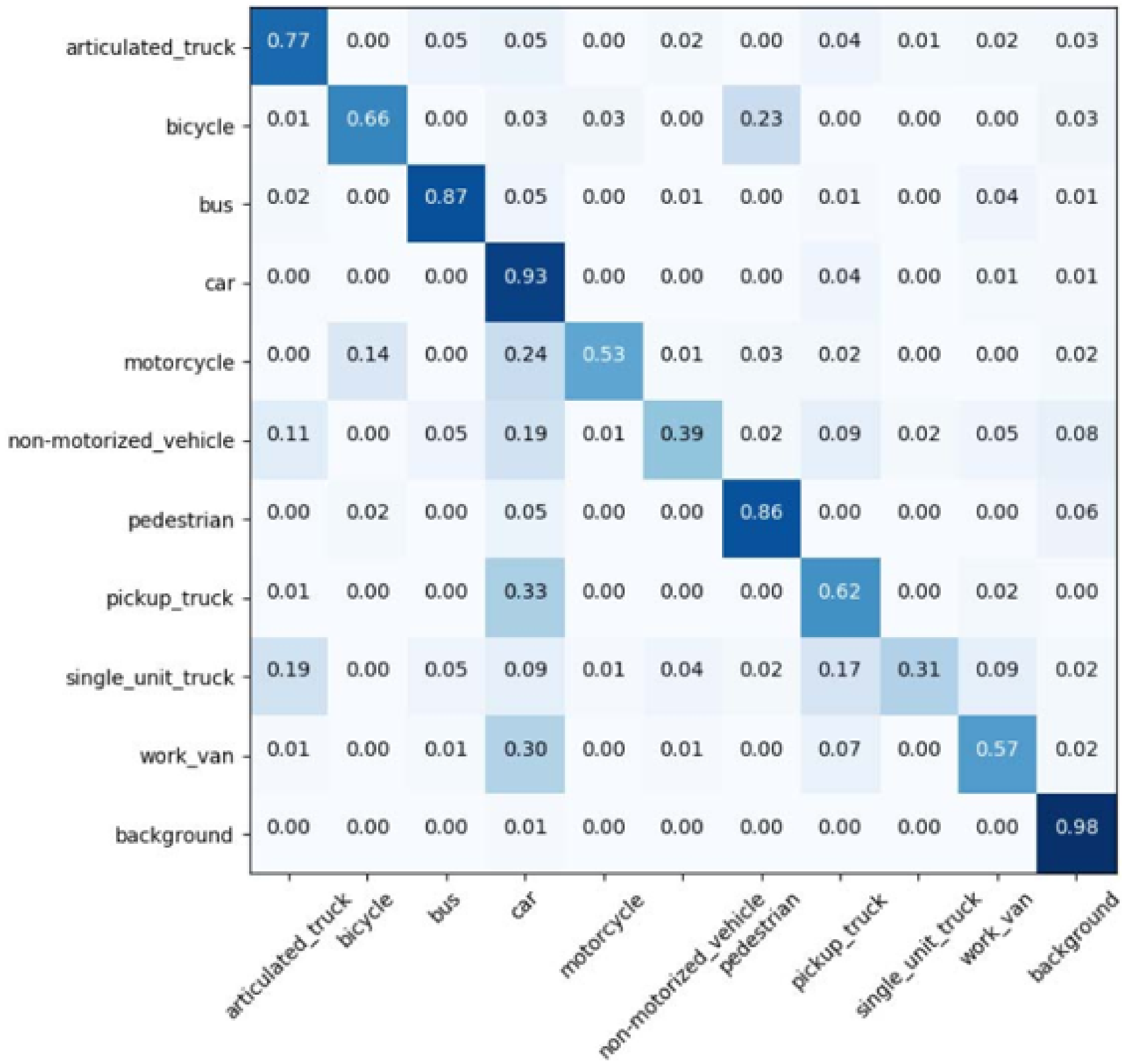

Figure 3: Confusion matrix obtained from MIO-TCD dataset with YOLOv5s model

3.4.3 Mutual communication subsystem (MCS)

For MCS, noise, like intermodulation distortion, may occur when multiple signals with different frequencies are combined, resulting in the creation of new signals at frequencies that were not present in the original signals. This problem could be solved by using the Channel Encoder Decoder and other filters to remove unwanted frequencies or other errors.

4 Ethics and Safety

4.1 Ethics

According to IEEE Code of Ethics 1 "to protect the privacy of others, and to disclose promptly fac Tors that might endanger the public or the environment;" [12], we promise that the data set used in the project will seek the permission of the owner, and the collected images will also be cleared after use in order to protect information security. Also, drones should conduct experiments in the permitted flight area to comply with relevant regulations. **Since our test and demonstration were carried out on the campus of Zhejiang University, according to the school's guidelines[13], we need to apply to the school in advance before the drone flies.** The final result of the project cannot be used in any scenario that infringes on public information and privacy.

According to IEEE Code of Ethics 5 "to seek, accept, and offer honest criticism of technical work, to acknowl Edge and correct errors" [12], and "credit properly the contributions of others;" [12] We will take the problems that may occur in the whole process of the project seriously and correct it in time. The references involved and other open-source references will also be correctly referenced and identified with permission.

4.2 Safety

After consideration, there are three main safety risks in our project: battery safety, UAV flight experimental safety, and code safety.

Battery safety: According to the ECE445 battery safety document [14], before installing the battery, we will understand the battery specifications and related data, and pay attention to isolation from other working areas such as drones to avoid impact.

UAV flight experimental safety: Given that the UAV we use is installed and built by us independently, we are responsible for everything that may occur in the UAV. Therefore, we should carefully check the UAV system, pay attention to the situation of components such as flight control modules and batteries in the UAV, and conduct experiments after being proficient in the flight control of the UAV in the unmanned area.

Code security: In our project, the LSES involves image recognition and semantic information extraction. This part of the code is an important part of the project, so we need to ensure that the relevant code we write is safe and will not be leaked.

References

[1] X. Luo, H.-H. Chen, and Q. Guo, "Semantic communications: Overview, open issues, and future research directions," *IEEE Wireless Communications*, vol. 29, no. 1, pp. 210–219, 2022.

[2] D. Deacon, M. Pickering, P. Golding, and G. Murdock, *Researching communications: A practical guide to methods in media and cultural analysis*. Bloomsbury Publishing USA, 2021.

[3] T. Han, Q. Yang, Z. Shi, S. He, and Z. Zhang, "Semantic-preserved communication system for highly efficient speech transmission," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 1, pp. 245–259, 2022.

[4] Z. Qin, X. Tao, J. Lu, and G. Y. Li, "Semantic communications: Principles and challenges," *arXiv preprint arXiv:2201.01389*, 2021.

[5] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Communications magazine*, vol. 54, no. 5, pp. 36–42, 2016.

[6] Y. Mao, "A pedestrian detection algorithm for low light and dense crowd based on improved yolo algorithm," *MATEC Web of Conferences*, vol. 355, p. 03 020, Jan. 2022. DOI: 10.1051/matecconf/202235503020.

[7] K. Soomro, A. Zamir, and M. Shah, "Ucf101: A dataset of 101 human actions classes from videos in the wild," *CoRR*, Dec. 2012.

[8] S. M. Riad, "The deconvolution problem: An overview," *Proceedings of the IEEE*, vol. 74, no. 1, pp. 82–85, 1986.

[9] M. Lee, H. Kim, and J. Paik, "Correction of barrel distortion in fisheye lens images using image-based estimation of distortion parameters," *IEEE ACCESS*, vol. 7, pp. 45 723–45 733, 2019.

[10] Z. Yahya, S. Imane, H. Hicham, A. Ghassane, and E. B.-I. Safia, "Applied imagery pattern recognition for photovoltaic modules' inspection: A review on methods, challenges and future development," *Sustainable Energy Technologies and Assessments*, vol. 52, p. 102 071, 2022.

[11] Z. Luo, F. Branchaud-Charron, C. Lemaire, *et al.*, "Mio-tcd: A new benchmark dataset for vehicle classification and localization," *IEEE Transactions on Image Processing*, vol. 27, no. 10, pp. 5129–5141, 2018. DOI: 10.1109/TIP.2018.2848705.

[12] IEEE. "Ieee code of ethics." (2020), [Online]. Available: https://www.ieee.org/about/corporate/governance/p7-8.html (visited on 03/07/2023).

[13] Z. University. "Drone school flight." (2020), [Online]. Available: http://xwfw.zju.edu.cn/wsbsdt.php?cmd=sx_content&blsxm=347300&dxlb=&fwxz=&bjlx=&fwms=&sldz=&fwzt=&zuixin_tuijian=&orderby=&jianhua_flag=&shoulibm=&one_run=&keyword=%E6%97%A0%E4%BA%BA%E6%9C%BA&p=1 (visited on 03/23/2023).

[14] U. of Illinois ECE445 Course Staff. "Safe practice for lead acid and lithium batteries." (2016), [Online]. Available: https://courses.engr.illinois.edu/ece445zjui/documents/GeneralBatterySafety.pdf (visited on 03/21/2023).