ECE 445

SENIOR DESIGN LABORATORY

DESIGN DOCUMENT

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 21, 2023

Contents

1 Introduction

1.1 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 a 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.

1.2 Solution and Visual Aid

Firstly, 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.

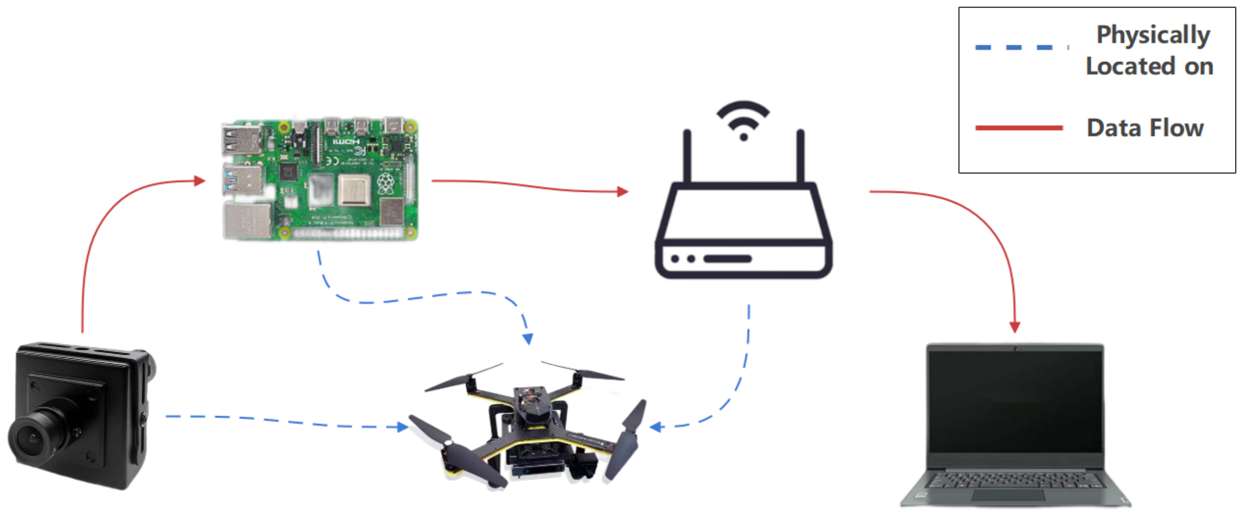

Figure 1: Visual Aid: Data Flow and Physical Location

1.3 High-level Requirements

- 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 the 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.

2 Individual Design Work

2.1 Block Diagram

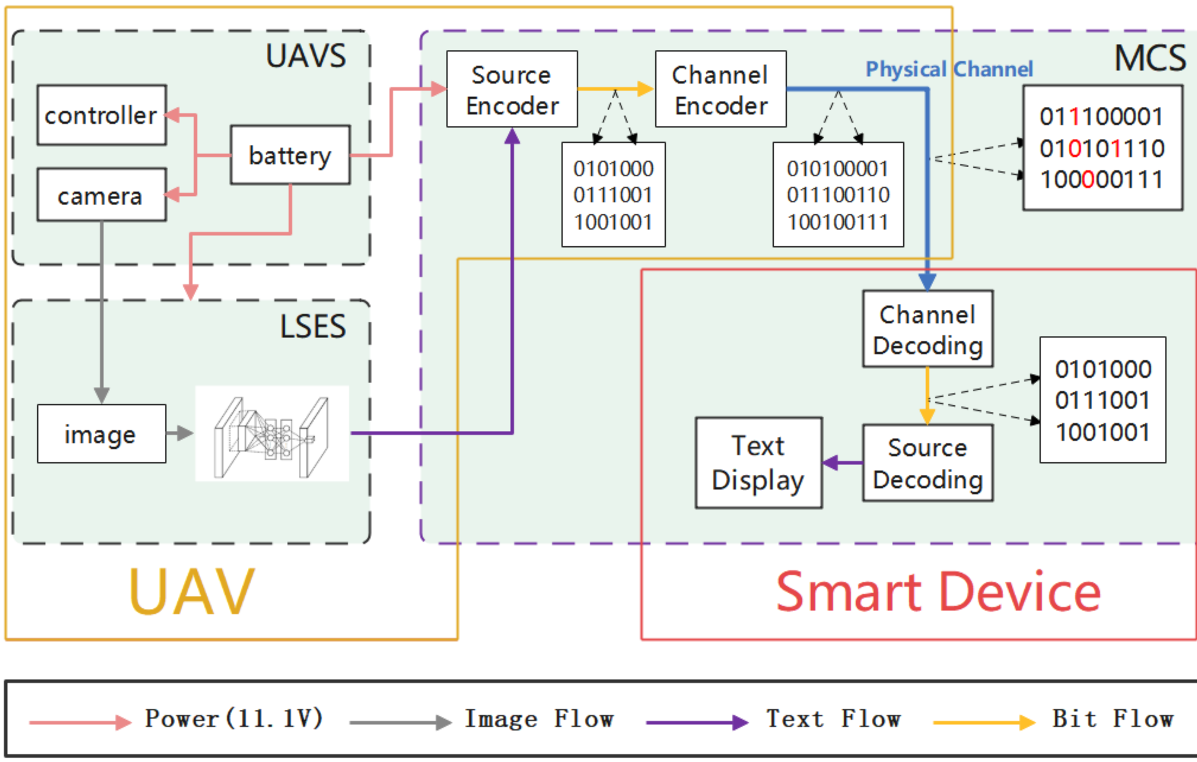

Figure 2: System Block Diagram

2.2 UAV mechanical, balance and dynamic Subsystem (UAVS)

2.2.1 Descriptions

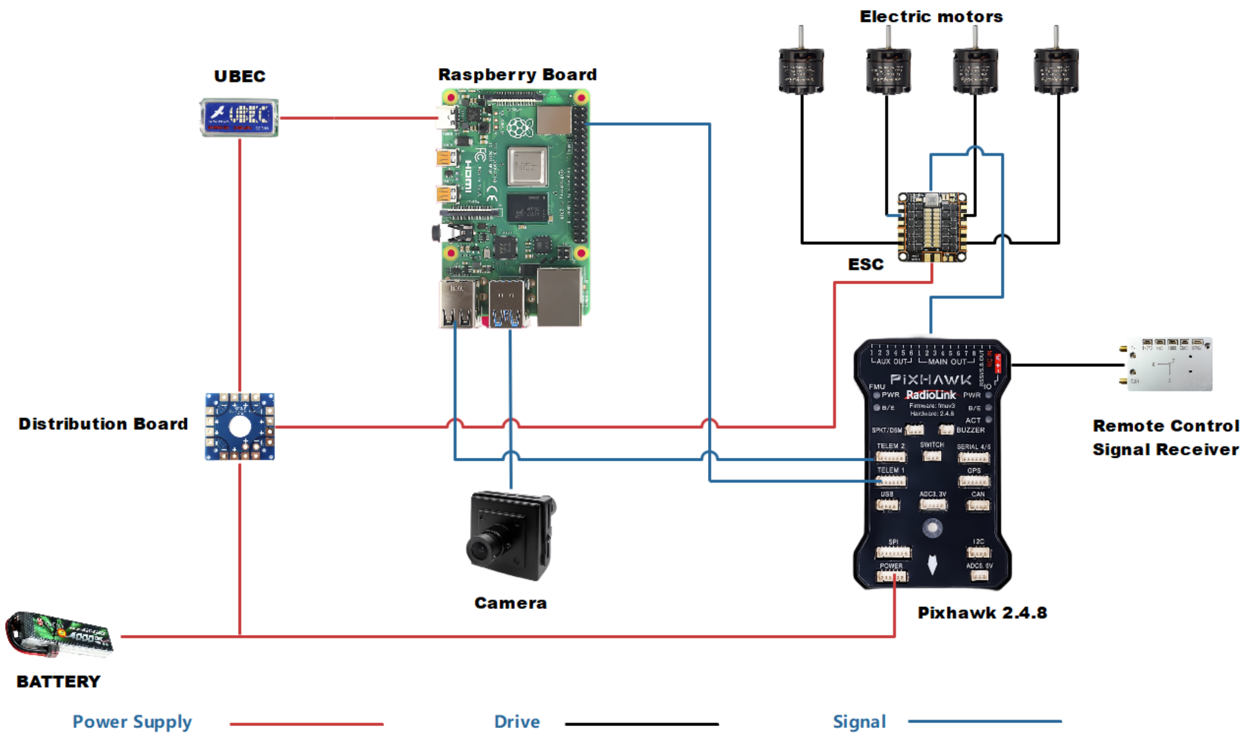

Figure 3: Structure of UAV

The UAV subsystem includes a power module, a controller module, and a camera as shown in figure 3. 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, PIXHAWK, 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.

2.2.2 Requirements and Verifications

Requirements	Verifications
The ultra battery elimination circuit(UBEC) could provide 5.8V from an 11.1V source.	We will measure the output voltage from UBEC using an oscilloscope, ensuring that the output voltage stays within 5% of 5.8V.
UAV can still fly smoothly and quickly after adding devices.	We will do some fly tests. Such as 1. Let the UAV fly in a circle with a radius of 5m through the air. 2. Let the UAV quickly rise into the air and hover. The success criterion is that the UAV can move and hover smoothly in three dimensions.
UAV must be able to hover up to at least 5 meters in the air and take clear photos.	We will do flying tests and take some simple photos. Such as 1. Let the drone fly to an altitude of 5m and circle around to take pictures. The success criterion is that it can hover up steadily at the required height, and the picture is clear and not blurred by movement.

2.3 Lighting Semantic Extraction Subsystems (LSES)

2.3.1 Descriptions

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] which is shown in figure 4, 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, as shown in figure 2.5.3 is a motion recognition dataset of realistic action videos, providing 13,320 videos from 101 action categories. The semantic information extracted by LSES will serve as the input of the mutual communication subsystem (MCS).

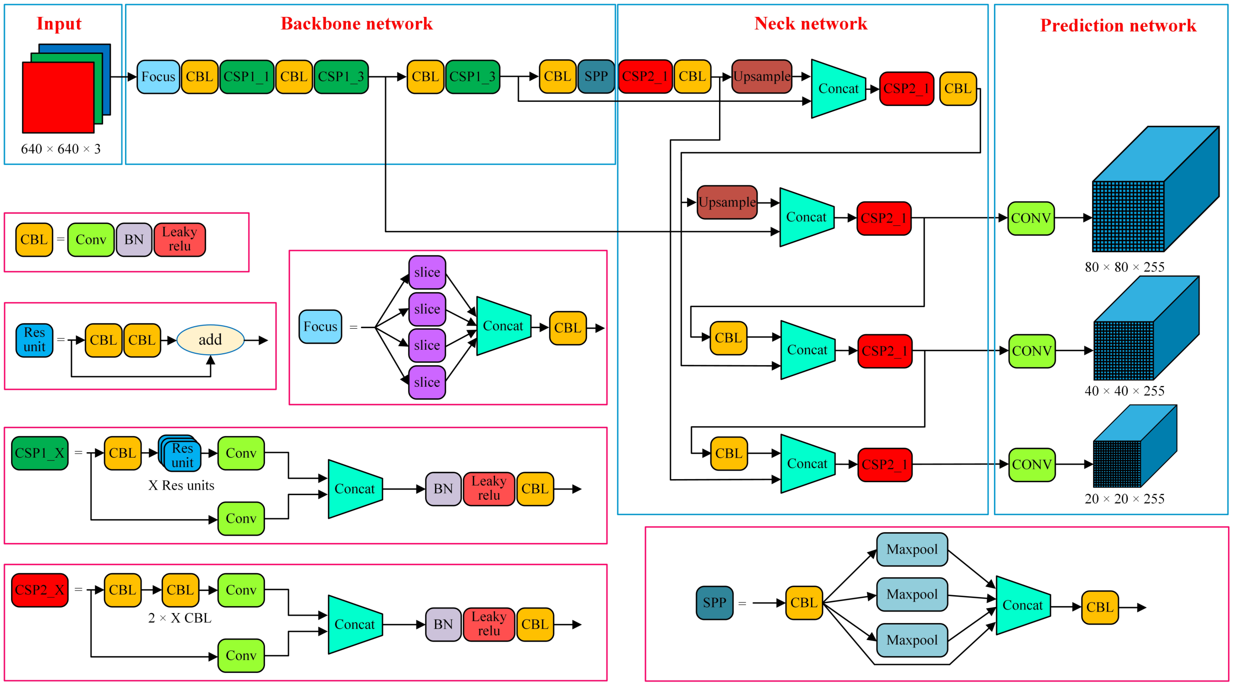

Figure 4: Structure of YOLOv5s Model[6]

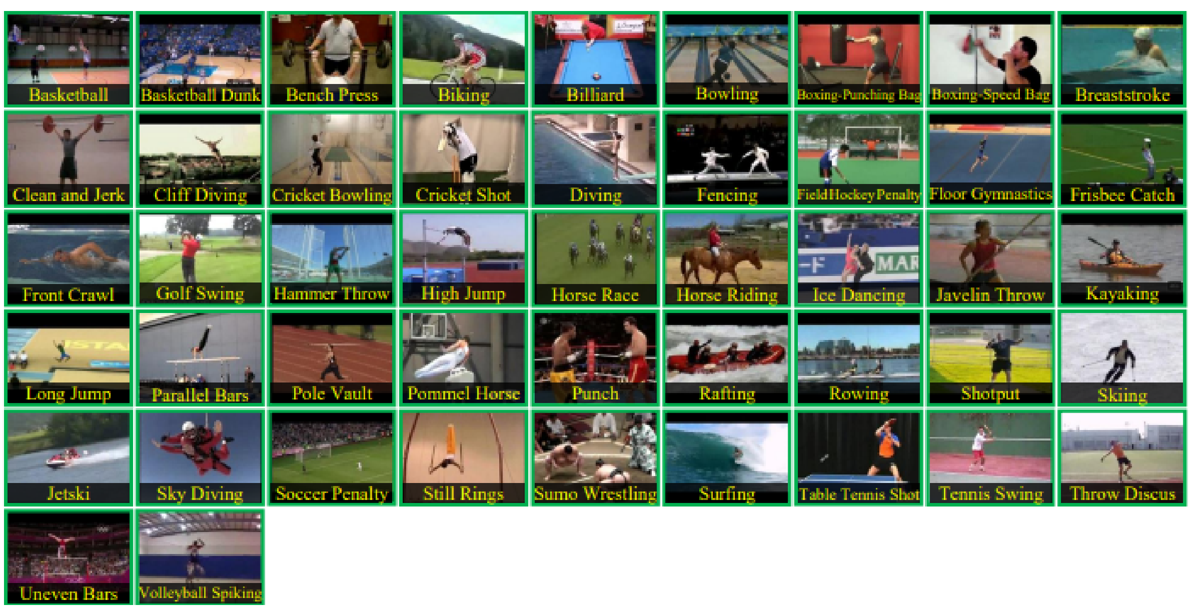

Figure 5: Dataset: UCF101[7]

2.3.2 Requirements and Verifications

Requirements	Verifications
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%	The first test is on the computer through elements like Raspberry Pi. We will pick some test images taken by our UAV or from the Internet and do a test on our algorithm. The expected accuracy is above 70%.
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.	The test on UAV will be given to check data transmission between subsystems and to test the accuracy in our campus scenarios. LSES will receive live pictures. The accuracy is expected to be at least 70%. The detection is expected to finish in average of 5s for 10 photos.

2.4 Mutual Communication Subsystem (MCS)

2.4.1 Descriptions

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 should have similar semantic information but not necessarily the same words. The text example could be like "There are 15 people waiting in a line." The communication should be quick and without losing semantic information.

2.4.2 Requirements and Verifications

Requirements	Verifications
MCS needs to successfully transmit text information without any error from UAV to smart devices, and display it on a screen.	The first test is to transmit some sample words from one smart device to another by the required time. Also, another test on UAV will be given to check data transmission between subsystems.
The time for transmission should be less than 0.5s for each sentence with about 10 words.	The delay due to the transmission should be less than 0.5s. Besides, **the transmission of text information needs to achieve 90% accuracy without affecting understanding.**

2.5 Tolerance Analysis

2.5.1 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. In order to obtain stable voltage, we also use an ultra battery elimination circuit (UBEC).

2.5.2 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]. We have successfully tested Yolov5s pre-trained model on our new dataset. The mAP on average is 50.2%. 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.

2.5.3 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.

Besides, the semantic channel capacity of a discrete memoryless channel [4] is expressed as

$$C_s = \sup_{p(Z|X)} \left\{ I\left(X;V\right) - H\left(Z|\, X\right) + \overline{H_S\left(V\right)} \right\},$$

Here $I(X;V)$ **is the mutual information between the source,** X **, and the transmission task,** V**. Here** $p(Z|X)$ **is the conditional probabilistic distribution that refers to a semantic coding strategy with the source,** X**, encoded into its semantic representation,** Z**, and** $H(Z|X)$ **means the semantic ambiguity of the coding.** $\overline{H_S(V)}$ **is the average logical information of the received messages for the task** V**. Then here we can see that if we could make** $\overline{H_S(V)}$ **be bigger than** $H(Z|X)$**, the semantic channel capacity could be always bigger than 0. That means the receiver can handle the semantic ambiguity. For our design, this is easy to accomplish.**

3 Cost and Schedule

3.1 Cost

First, for our labor cost, we assume everybody's hourly wage is ¥100/hour, and we need to work for 10 hours/week for all four people. And we need to do this for the following 10 weeks this semester. So for this part, our fixed development cost is :

$$4 \cdot \frac{20CNY}{hr} \cdot \frac{10hr}{wk} \cdot 10wks \cdot 2.5 = 20000CNY$$

Then, since only one person is needed to operate the drone, we don't need a lot of bulk. For the parts and manufacturing prototype costs, it's estimated ts ¥2946 each:

Part	Vendor	Cost (prototype)	Cost (bulk)
Professional aerial photography UAV (CK10pro)	Taobao	¥1888	¥50
8GB Raspberry Pi (4B; generic)	Taobao	¥728	¥20
200W pixels Monocular Camera (Reshi Technology)	Taobao	¥150	¥20
WiFi module (Small R Technology; MT7620)	Taobao	¥50	¥40
Total	Taobao	¥2816	¥130

Then we add two parts together, our total development cost should be ¥22946.

3.2 Schedule

Week	Yu Liu	Chenhao Li	Chang Su	Tianze Du
3/20/23	Write Design Document 2.1 and 2.2	Write Design Document 2.3 and 2.4	Write Design Document Part 1 and 3	Write Design Document Part 4
3/27/23	Learn the use of the Raspberry Pi	Look for object detection algorithms	Find the appropriate dataset	Purchase the required parts
4/3/23	Simple programming on the Raspberry Pi	Run through object detection algorithms on the computer	Find the appropriate semantic segmentation algorithm	Add parts on the UAV
4/10/23	Run object detection algorithms on the Raspberry Pi	Find the right means of communication for UAV	Run the semantic segmentation algorithm on the computer	Design and construction of UAV balancing systems

4/17/23	Enable communication between drones and other smart devices	Enable communication between drones and other smart devices	Implement the semantic segmentation algorithm on the Raspberry Pi	Design and construction of drone power systems
4/24/23	Carry out the final inspection of the part for which he is responsible	Carry out the final inspection of the part for which he is responsible	Carry out the final inspection of the part for which he is responsible	Carry out the final inspection of the part for which he is responsible
5/1/23	Test flights of UAV, detection and analysis of errors	Test flights of UAV, detection and analysis of errors	Test flights of UAV, detection and analysis of errors	Test flights of UAV, detection and analysis of errors
5/8/23	Prepare for Mock demo	Prepare for Mock demo	Write the Final Report draft	Write the Final Report draft
5/15/23	Detect the overall effectiveness of the project	Detect Lighting Semantic Extraction Subsystems	Detect Lighting Semantic Extraction Subsystems	Detect UAV mechanical, balance and dynamic Subsystem
5/22/23	Write Final Report	Write Final Report	Prepare for Final Presentation	Prepare for Functionality Demonstration Video

4 Ethics and Safety

A number of potential ethical and safety issues had to be considered in our project. First of all, both the UAV and the Raspberry Pi board need to be powered by batteries, which cannot be replaced by other power sources. So the stability and safety of the batteries are an important part of ensuring the success of the project. According to the ECE445 battery safety document[11], we will understand the battery specifications before installing the battery, test the battery circuit packaging, charging and discharging, and the operating temperature, and pay attention to the isolation from other work areas such as the transmission module and the Raspberry Pi board to avoid impact.

In addition to this, as the drone is flying work outdoors, the weather will cause problems such as short circuits in our work area, which will affect our work. Therefore, we will follow IP66 standard to make the working area of the drone waterproof, so that it can operate normally even in the rain and snow.

Also, given that the drone needs to be manned, I need to make sure that the drone can fly safely and avoid collisions with people or other objects to reduce the risk of maneuvering. Currently, I decided to add sensors and GPS modules to determine the location of the drone. In addition, I will use the programming software to achieve special circumstances such as an automatic landing function. These measures maximize the safety of drone maneuvering.

In addition to the manipulation method, when choosing the flight area and time period, we also need to ensure that the drone will not cause threats and interference, and avoid flying in densely populated areas and flight-restricted areas. When flying, we need to comply with local regulations and rules to ensure that my project is operating within legal limits. Because our tests and demonstrations are conducted on the campus of Zhejiang University, according to the school's guidelines [12], we need to apply to the school in advance before the drone flight.

According to the Institute of Electrical and Electronics Engineers(IEEE) Code of Ethics 1 "to protect the privacy of others, and to disclose promptly factors that might endanger the public or the environment;"[13], 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. 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 acknowledge and correct errors" and "credit properly the contributions of others;" [13] We will take the problems that may occur in the whole process of the project seriously and correct it in time. For example, we need to adjust the index and model if the simulation result is not as expected. Also, the references involved and other open-source references will also be correctly referenced and identified with permission.

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] 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).

[12] 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).

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