# OPTICAL COMMUNICATION AND DEEP LEARNING-BASED OBJECT DETECTION FOR UNDERWATER SWARM ROBOTICS

by

Clifford Oppong Boakye-Mensah

B.Sc., July 2019, Kwame Nkrumah University of Science and Technology

A Thesis Submitted to the Faculty of
Norfolk State University in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

ELECTRONICS ENGINEERING

NORFOLK STATE UNIVERSITY

Approved by:

---

Dr. Hongzhi Guo (Advisor)

---

Dr. Makarand Deo (Member)

---

Dr. Prathap Basappa(Member)

# ABSTRACT

SOFTWARE-DEFINED OPTICAL COMMUNICATION AND IMAGE PROCESSING FOR UNDERWATER ROBOT SWARMS

Clifford Oppong Boakye-Mensah
Norfolk State University, 2022
Advisor: Dr. Hongzhi Guo

Underwater Wireless Communications have facilitated huge progress in the field of oceanography. Due to the nature of the underwater environment, the employment of unmanned vehicles (robots) has become highly popular, usually requiring highly reliable communication networks in order to facilitate real-time high-quality image processing and video streaming. While Underwater Optical Communication Systems have the potential to assist in these tasks, they are also accompanied by prominent challenges. Problems arising from the underwater communication channel, robot trajectory control and scheduling, and reliability of the optical communication link threaten to neutralize the high-bandwidth, low-latency nature of underwater optical communication. In this thesis, a high performance remotely operated vehicle is assembled and deployed for exploration. An object detection model is also implemented to process the videos obtained from the ROV. A low-cost, software-defined, optical communication system is proposed to tackle the challenges mentioned above. A test-bed is set up and configured to ensure the reliability of the communication link. Adaptive algorithms based on the characteristics of the underwater communication channel are proposed to control robot trajectory. Results from various simulations show that highly accurate object detection models can be implemented and deployed. Experiment results also show that a highly reliable yet low-cost communication system can be configured. The proposed distributed algorithm can also reliably obtain optimal trajectories and locations for underwater robots without knowing global environmental information.

# ACKNOWLEDGEMENTS

I would like to thank God Almighty for His never-ending goodness since I began this program. My sincerest gratitude goes out to Dr. Hongzhi Guo, my advisor, for his constant support and guidance, and my thesis committee members Dr. Makarand Deo and Dr. Prathap Basappa for their guidance throughout my studies, and during the process of my thesis.

Finally, I would like to acknowledge Mr Matthew Mensah and the late Sophia Inkebil Mensah, to whom I owe a huge part of my successes.

This study was supported by the The Thomas F. and Kate Miller Jeffress Memorial Trust, Bank of America, Trustee.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| **UWC** | **U**nderwater **W**ireless **C**ommunication |
| **UOWC** | **U**nderwater **O**ptical **W**ireless **C**ommunication |
| **UOWN** | **U**nderwater **O**ptical **W**ireless **N**etwork |
| **VLC** | **V**isible **L**ight **C**ommunication |
| **AUV** | **A**utonomous **U**nderwater exploration **V**ehicle |
| **BER** | **B**it **E**rror **R**ate |
| **PPM** | **P**ulse **P**osition **M**odulation |
| **OPPM** | **O**verlapping **P**ulse **P**osition **M**odulation |

# 1 INTRODUCTION

## 1.1 Background

The vast nature of the oceans mean that there are limitless research opportunities in oceanography. Aside commercial purposes, the exploration of the underwater environment has borne improvements in areas such as climate change monitoring, marine species migration patterns, monitoring of marine litter and other targets, and the exploration of deep-sea wreckage. Surveillance applications such as the aforementioned have increased the need and demand for swarm robotics and high-speed image processing and data communication. (Guo and Boakye-Mensah, 2022).

Generally, the exploration of extreme environments, especially the underwater environment, has seen significant improvement which can be attributed largely to two main factors.

Firstly, the introduction of *AUVs* has made the exploration of harsher regions in the underwater environment more feasible. Explorations such as the one in (Bingham et al., 2010) off the shores of Greece in 2005 to explore deep-sea wreckage, and by the Japan Agency for Marine Earth Science and Technology (JAMSTEC) in 2018 to find marine litter have all extensively employed AUVs.

That aside, wireless communication technologies, by virtue of the elimination of wired connections and the improved mobility and range of operation they provide, have helped the implementation of swarm robotics in underwater environments. These technologies have vastly improved the quality and reach of industrial, military, and surveillance exploration in the underwater environment (Saeed et al., 2019).

Existing wireless communication technologies include Underwater Acoustic and

RF Communication Systems. Acoustic communication, which is extensively employed in underwater communication, enables the implementation of long-range communication over several kilometers. However, systems based on underwater acoustic communication suffer from high propagation delays, and communication is affected when environmental factors such as temperature, multipath fading, pressure, and water salinity also change. Another huge challenge these systems face is the inability to provide a high data rate due to the limited bandwidth. The data rates possible with such systems generally lie on the order of several *kilobits per second* (Akyildiz, Wang, and Sun, 2015).

RF communication on the other hand provides relatively higher data transmission and is more tolerant to changes in the underwater environment itself. However, these systems are accompanied by bulky, costly, and often high energy-consuming transceivers in order to facilitate reliable communication (Zeng et al., 2016). Also, despite the fact that RF systems demonstrate smaller propagation delay at lower frequencies, the communication range and data rate cannot satisfy high-quality video transmission (Guo, Sun, and Wang, 2021; Guo, Sun, and Wang, 2017). This makes it only suitable for moderately data-rate-intensive communication over shorter distances. Another point worth noting is that RF signals experience very high attenuation losses as the frequency increases (Kaushal and Kaddoum, 2016).

Among the available technologies, wireless optical communication (visible light communication) has demonstrated the potential to provide the best performance. Some optical systems have been know to obtain several *Gbps* data rates (Lu et al., 2016a; Sun et al., 2020) within several tens of meters. As the distance increases, wireless optical communication can achieve a rates of *5Mbps* at distances of about *200 m* (Johnson, Green, and Leeson, 2013). Nevertheless, the communication range is still limited to allow direct peer-to-peer communication from deep underwater to the water surface. This presents a challenge which this work will seek to address by employing a relay network which will extend the communication range, as illustrated in Fig. 1.1. Such a relay network, as shall be illustrated in subsequent

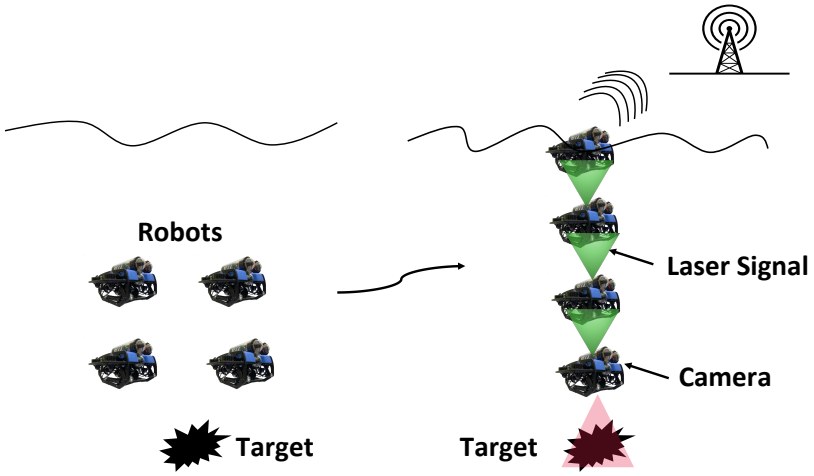


FIGURE 1.1: Illustration of the underwater multi-hop robotic network. A swarm of multiple robots cooperatively search for underwater targets. Once a target is identified, robots form a multi-hop network using optical communication to relay real-time data to terrestrial/surface communication networks.

chapters require scheduling and control algorithms in order to maintain an organized and reliable communication system.

There are other challenges that are faced during VLC, most notable the spatial-temporal dynamics which affect the transmitted signal (Saeed et al., 2019; Johnson, Green, and Leeson, 2013).Despite these challenges communication under these conditions have proven to be feasible. The attenuation of light lies around around 0.39dB/m in ocean water (Kaushal and Kaddoum, 2016). Under such situations, reliable, power-efficient optical communication systems can be implemented.

However, existing underwater communication research are mainly based on theoretical analysis. A few testbeds have been developed using high-cost equipment that cannot be easily installed on sensors or robots. In order for such systems to be fully exploited on a large scale, it is expedient to have low-cost implementations that can still provide reliable communication.

Another technology that has been applied to improve oceanography applications is Deep Learning. Areas such as target detection and monitoring, and image enhancement have found extensive use for Deep Learning, specifically Computer Vision.

Considering the different motivations discussed, the aim of this thesis is to present a robot-based system which incorporates a target detection model into the AUVs and communicates over a relay network using a low-cost VLC system. The contributions of this system include:

- The assembly and deployment of two underwater robots for high-speed video streaming and data gathering in the underwater environment.
- The implementation of a Deep Learning based object detection model to support the identification of specific targets under water.
- The design of a distributed trajectory design algorithm for underwater robots to form a relay network, which can route data from deep underwater to water surfaces.
- The usage of a low-power low-cost laser diode to facilitate visible light communication over air/water media along with a low-cost Silicon Photo Multiplier (SiPM) as the receiver.

## 1.2 Research Challenges

### 1.2.1 Visible Light Communication

Visible light communication in the underwater environment is generally challenging due to the following reasons:

- **Spatial-Temporal Dynamics:** The underwater wireless optical channel experiences spatial-temporal dynamics due to the turbulence and chlorophyll concentration variations under water (Saeed et al., 2019) (Johnson, Green, and Leeson, 2013). This has implications at the various layers of communication. One major effect of these dynamics is that they result in a depth-dependent attenuation. This has the potential to affect not only the power levels of the

received signal but also the integrity of the bits transmitted. This makes it expedient to attain a profile of the attenuation and its variation in order to facilitate reliable communication.

Again, these dynamics generate problems when it comes to scheduling and localization of swarm robots. The instability of the wireless link between any two robots due to turbulence-induced fading leads to high levels of randomness during communication especially with the level of received power. In scenarios where the optimal localization of the robots depend on the received power, this randomness may mislead the robot to move to a sub-optimal location.

- **Environmental Information:** As mentioned earlier, the wireless optical channel is affected by the environmental factors such as chlorophyll concentration variation and turbulence. Information about these environmental factors however, is not always available. The underwater environment can be classified into different types based on the concentration of chlorophyll (Johnson, Green, and Leeson, 2013). This makes it impractical for robots to first identify the exact water-type and conditions and then design their trajectory accordingly.

- **Control Algorithms:** Robots are generally widely spread in the underwater environment during swarming for the purposes of efficiently monitoring a specific target or phenomenon. It is challenging to use centralized control algorithms to plan the trajectory of all the robots, mainly because such algorithms require collecting information from each robot. This leads to excessive delay in robot control and high resource consumption.

### 1.2.2 Object Detection

The use of AUVs in the monitoring and detection of targets in the underwater environment is challenging due to the following reasons:

- **Image Distortion:** Changes in the light and turbidity conditions in the underwater environment increases the difficulty of vision in such environments (Lu et al., 2016b). This makes the monitoring of targets and objects very difficult. That aside, the light absorption and scattering characteristics of the environments generally means that images that will be obtained in the environment will be low-contrast, blurry and have much of the colors degraded hence leaving a hued representation of the actual image (Islam, Xia, and Sattar, 2020).

- **Target Distortion:** Over time, many objects to be identified within the underwater environment go through a lot of changes. These objects already are of various shapes colors and sizes. The huge challenge that arises is that all these different objects are to be identified regardless of these variables (Fulton et al., 2019).

There have been several studies which have addressed and approached the above challenges from different perspectives.

# 2 RELATED WORK

## 2.1 Related Works

### 2.1.1 Visible Light Communication and Robotic Networks

There have been numerous bodies of work aimed at exploring the feasibility of visible light communication not only in the underwater environment, but also through the air-water media boundary. In (Carver et al., 2020a) a bi-directional air-water communication system is presented. The aim of the system is to implement a high-bandwidth communication link that is reliable after traversing the air/water boundary. Two transceiver nodes are placed on either side of the boundary, with one on land and the other one underwater. The modulator circuit comprises an Field-Programmable Gate Array (FPGA), an Arduino Due, voltage converters and a laser diode driver. These collectively work to generate a signal which is used to modulate the light. The system employs OPPM as the modulation scheme. The laser diode is interfaced with a MEMS mirror, triplet lens and a fisheye lens in order to steer the light over a wider angle to address the spatial-temporal dynamics mentioned earlier. That aside, ultrasonic sensors are used to predict the depth and nature of the underwater medium in order to correctly steer the light in the direction of the receiver. Both the transmitter and the receiver are interfaced with an optical filter to mitigate the influence of ambient light on the communication link. The procedure yields an average throughput of ***5.04 Mbps*** with BERs on the order of ***0.01***. The study in (Chen et al., 2017) covers air-water communication using both red laser and blue-green laser. While the red laser provides bit rates less than ***110 kbps*** over 1 m, the 520 nm blue-green laser provides up to ***5.5 Gbps*** data rates over a 5 m air channel and 21 m water channel. This work uses Quadrature Amplitude Modulation (QAM) and Orthogonal Frequency Division Multiplexing (OFDM) to

modulate a pseudo-random sequence and steps are taken to mitigate inter-symbol interference (ISI). At the receiver end, an Avalanche Photo Diode (APD) is interfaced with a lens as the front-end. A mixed signal oscilloscope is used to sample the signal after which the signal is transmitted to a computer to be demodulated.

There have also been works aimed at improving the relay and trajectory control methods in robotic networks. The topology employed to define the relay network in this thesis is a *1-Dimensional (1D)* topology. The 1D relay scheduling problem has been studied in (Chattopadhyay et al., 2016). The study employs a theoretical knowledge of the wireless channel to deploy relay nodes on-the-go along a given path. The study also proves that, given an adequate number of nodes along the said path, the path loss effects of the wireless channel could be circumvented. A study exploring a distributed network connectivity-preserving trajectory control algorithm for swarm robotics has been proposed in (Zavlanos, Ribeiro, and Pappas, 2012). In order to plan robots' motion to maximize the throughput and maintain network connectivity, the study considers the terrestrial wireless channel in tandem with the communication rates available.

### 2.1.2 Image Processing

In (Islam, Xia, and Sattar, 2020), a generative adversarial network (GAN) is designed to provide real-time enhancement to the images taken during underwater exploration. The study makes use of two different datasets; one dataset contains a collection of distorted images and the other contains enhanced versions of the images in the first dataset. The GAN is trained to learn a mapping from the distorted images to the enhanced images by using an encoder-decoder network. The study in (Fulton et al., 2019) investigates the implementation of different object detection architectures towards the real-time identification/detection of plastic litter in the underwater environment. The *YOLOv2, Tiny-YOLO, Faster R-CNN and Single Shot*

*Detection (SSD)* architectures are compared in terms of detection ability and processing times. The YOLO architectures were found to have lower processing times; however their detection performances were generally inferior to the Faster R-CNN and SSD models.

# 3 ROBOT ASSEMBLY AND DEPLOYMENT

As mentioned in earlier chapters, the usage of AUVs have borne significant improvement in the field of oceanography. Being largely extreme environments, the ocean and other large water bodies very often present circumstances that make exploration by humans very difficult. Again, due to the usually vast surface areas covered by these bodies, the usage of AUVs make it more efficient in the patrolling of as compared to manned explorations. The main aim of this chapter is summarized below:

- Assemble a robot that has the ability to cover relatively more ground within a short while in the underwater environment, and is designed to with stand some level of unpredictability in these environments, for example regarding ocean/river currents.
- Establish a means of control and communication with the robot during deployment.
- Monitor video feed from robot and leverage that to collect data for image processing and object detection.

The AUV employed in this study is the BlueROV2 by Blue Robotics as it satisfies the aims of this section. The BlueROV2 has been used in a number of explorations in the past and has been known to to deployed up to 130$m$ beneath sea surface. The work in (Carver et al., 2020a) employs the use of the BlueROV2. In this chapter, the assembly, software setup and operation, and deployment results of the BlueROV2 will be discussed.

## 3.1 Assembly

The robot was delivered in its constituent parts and assembled on delivery with assistance from the Blue Robotics online guide. The frame of the robot is made of High Density Polyethylene (HDPE), and it comes with separate acrylic tubes and end caps which house the electronics and the battery. The robot is equipped with thrusters which control the direction, velocity and acceleration of motion under water. There are lumen lights which are attached to the sides of the frames and are used in cases of hindered visibility due to low-light conditions. These lights have adjustable brightness and can move over a 135° angle. The robot is also equipped with a 1080p digital low-light USB camera which has an 80° view horizontally and spans 180° vertically.

FIGURE 3.1: Assembly of the BlueROV2

The robot also circuitry also comprises a depth and temperature sensor, a 3-Degrees-Of-Freedom (3-DOF) accelerometer, gyroscope and magnetometer, an internal barometer and a leak detection system which alerts the pilot when any of the air-tight enclosures is compromised. A portable battery is a part of the architecture and the robot is connected to a Fathom-X tether interface via a tether which is

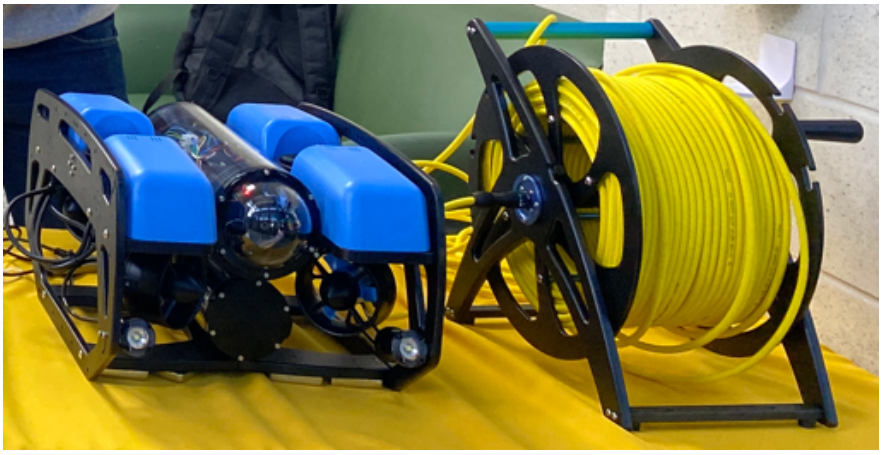

FIGURE 3.2: A completely assembled BlueROV2 connected to a spool via tether.

wounded on a spool. The fully assembled BlueROV2 is shown in Figure 3.2.

## 3.2 Software Setup and Operation

All the components mentioned in subsection 3.1 are connected via an electronic interface. On the board of the robot, a Pixhawk flight controller is attached and this controls the lumen lights, thrusters, camera tilt servo, power distribution, and all other sensors. The flight controller is connected to the tether interface via a Raspberry Pi companion computer. The electronic circuitry on the board is illustrated in Figure 3.3

Communication with the robot is set-up and maintained via a software interface known as QGroundControl and instructions are passed using a game controller/joystick. The software requires the robot to be connected to the tether interface and the tether interface then connected to a computer via a Local Area Network(LAN). This can be setup in the network settings of the Operating Software (OS) being used. Through QGroundControl, details such as the altitude, ground speed and flight time of the robot can be monitored. The video feed is also observed through the software. The firmware on the companion computer can also be upgraded through the software.

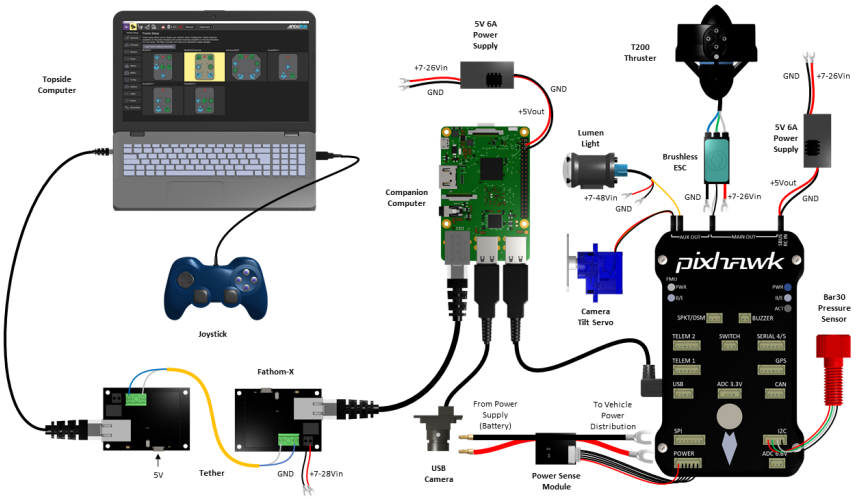


FIGURE 3.3: Electronic circuitry of the BlueROV2(BlueRobotics, n.d.)

Before the very first flight, all the sensors and thrusters are calibrated. The pitch and yaw mechanism of the robot is also configured and through the software as well as the assignment of buttons on the controller to specific actions. The software is available on Windows, Mac and Linux platforms.

## 3.3 Deployment Results

The robots have been deployed twice for underwater exploration. The first deployment was in February 2022 and the second one in April 2022. Both instances have been in different water bodies located in Chesapeake VA and Virginia Beach VA.

The main aim of these flights was to collect data to aid image processing and the training of the object detection model. As discussed in Chapter 2, the lighting conditions in the environment hugely impacted the quality of imagery that was observed. In Figure 3.4, sample frames obtained from the both explorations are displayed. The two frames above were obtained from the February flight while the last two frames were from the flight in April.

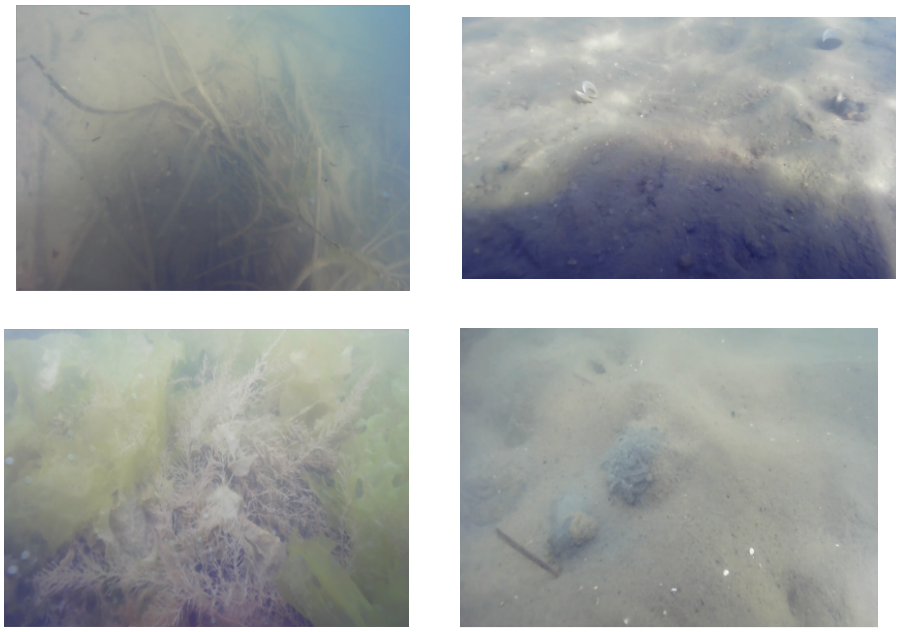

FIGURE 3.4: Frames obtained from the explorations in February 2022 and April 2022

The main challenge during the exploration was the difficulty in identifying different objects clearly due to the blurry nature of the images. Also, another difficulty arose due to the absence of plastics in the water bodies which have been explored so far. Future explorations will target different water bodies in order to encounter relatively more plastic waste in these environments.

# 4 OBJECT DETECTION

This chapter outlines the steps taken to perform object detection in the underwater environment. The chapter begins with a look at the dataset used in training the object detection model. An overview of neural networks and Convolutional Neural Networks is also. The Resnet-50 Faster R-CNN model, which presents some modifications to a typical neural network architecture will be reviewed along with the training process and the results of training.

## 4.1 Dataset

The dataset, known as the Trash-ICRA dataset, was obtained from an exploration by the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) which was conducted from 2018-02-01 to 2018-09-15 (Fulton et al., 2019).


FIGURE 4.1: Examples of images taken from the dataset (Fulton, Hong, and Sattar, 2020).

The data, downloaded as a zip file, contained a training set, validation set and a test dataset. All the sets were accompanied by text files and xml files both of

which contained labelling and bounding box information for each respective image. The training set contains 5,700 images, the test set contains over a 1000 images and the validation set contains about 700 images. All these images were sampled from videos recorded during the exploration. The images in the dataset describe different cases in the underwater environment, some of which are repeated especially in the training set. There are occurrences of marine life, marine litter (plastic), ROVs, pieces of wood and some instances of metal. Specifically Table 4.1 outlines the distribution of the 11 various classes in the dataset. The varying lighting and color conditions makes this dataset a challenging one to work with, for the reasons discussed in Chapter 1.

| Class | Number |
|---|---|
| 'bio' | 1951 |
| 'plastic' | 4580 |
| 'rov' | 1799 |
| 'timestamp' | 8773 |
| 'unknown' | 147 |
| 'metal' | 48 |
| 'wood' | 54 |
| 'rubber' | 15 |
| 'cloth' | 5 |
| 'fishing' | 12 |
| 'paper' | 10 |
| 'papper' | 1 |

TABLE 4.1: Distribution of the various classes in the Trash ICRA dataset.

As expected, the images in the dataset generally are hued and blurry as shown in Figure 4.1. The images, taken from the training set, show different instances where there are pieces of marine litter (plastics) and some aquatic life.

## 4.2 CNNs and the Faster RCNN Architecture

### 4.2.1 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks have become a very important part of many Deep Learning applications. In math, convolution represents an operation between two functions in which another function is produced. This third function which is produced, usually quantifies the way in which one function modifies the other function(Choudhari, n.d.). In the most basic sense, a CNN is a network that has the ability to learn both simple and complex features from data provided to it. As a result of this ability, the field of Computer Vision has employed the use of CNNs to a very large extent. Aside Computer Vision, fields such as Natural Language Processing have employed the use of CNNs to learn concepts from text datasets. Provided a piece of data can be represented as a set of numbers/arrays, some level of information can be drawn from it by a CNN. Take an image for example. A typical coloured (RGB) image can be represented as 3 $n \times m$ array where the $n$ and $m$ represent the dimensions of each channel of the image. One of the basic units of a CNN is a convolutional layer. This layer is usually a $c \times c$ array of weights which are applied directly to the input image in question. Therefore, for the input image mentioned earlier, the layer (also known as a kernel or filter) is slid over the the image and a simple convolution is performed using matrix multiplication. An output image is obtained and this image contains characteristics of both the input image, and the kernel used at this stage. At the beginning of convolution, basic features such as lines, circles, and other rudimentary shapes are learnt from the image. As the kernels begin to stack up, more complex and heirarchical features from the data are learnt. Generally, a loss function, an optimizer, and a few other hyperparameters are employed to govern the learning process. The loss function, which we seek to minimize, provides a measure of how far away the model's predictions are from the ground truth. The optimizer, and hyperparameters such as the learning

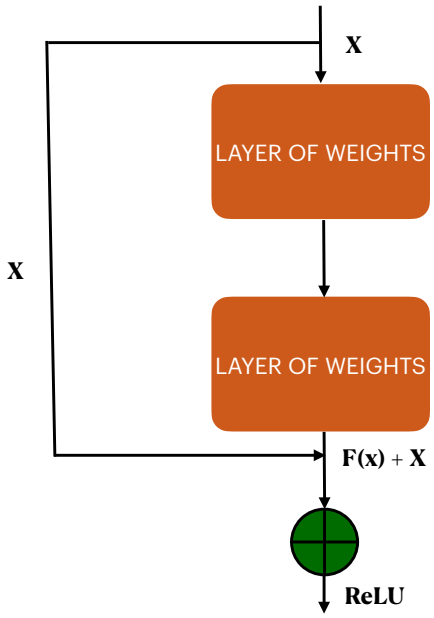


FIGURE 4.2: Resnet architecture with identity connections to solve the exploding and vanishing gradient problems.

rate then govern how fast a model learns, and steep the steps taken in the learning process are. This way, computers are able to obtain some semblance of vision and perform tasks such as image classification, image segmentation, object localization, object detection, and image captioning. Object detection as task combines two different tasks: classification and object localization.

### 4.2.2 Resnet-50 Architecture

As neural networks become bigger in order to handle more complex tasks such as object detection, some problems naturally arise within architectures. Due to the high number of layers stacked onto a neural network, problems such as vanishing gradient and exploding gradients tend to occur. The result of these problems manifests in the saturation and eventual degradation of the accuracy of such networks. Adding more layers at this stage serves a disadvantage as the training error will keep increasing. The ResNet architecture tries to solve the gradient problem by introducing a shortcut into the network (Mukti and Biswas, 2019). As illustrated in the Figure 4.2 below, introducing an identity at various points in the networks aids the prevention of the vanishing and exploding gradient problems.

### 4.2.3 Resnet-50 Faster RCNN Model

The Faster RCNN architecture presents an improvement on the already existing R-CNN and Fast R-CNN architectures in order to perform the classification and localization tasks. Region-based Convolutional Neural Networks (R-CNNs) were introduced in (Girshick et al., 2014) to aid object detection tasks. Given an input image, the network proposed about 2,000 various regions within the image where an object may be located, and the CNN generated a 4096-dimensional feature map. These maps were classified by a pre-trained Support Vector Machine (SVM). The main challenges with the R-CNN architecture was the fact that each stage of the network was independent, and hence the network could not be trained end-to-end. Also, the region proposals were fed separately to the network, making it resource intensive and almost impossible to implement in real time applications(Gad, n.d.). This led to the introduction of the Fast R-CNN architecture in (Girshick, 2015). The R-CNN architecture was a single-stage architecture which introduced a region of interest (ROI) pooling layer and used a multi-task loss to train the model end-to-end, making it faster than the R-CNN architecture.

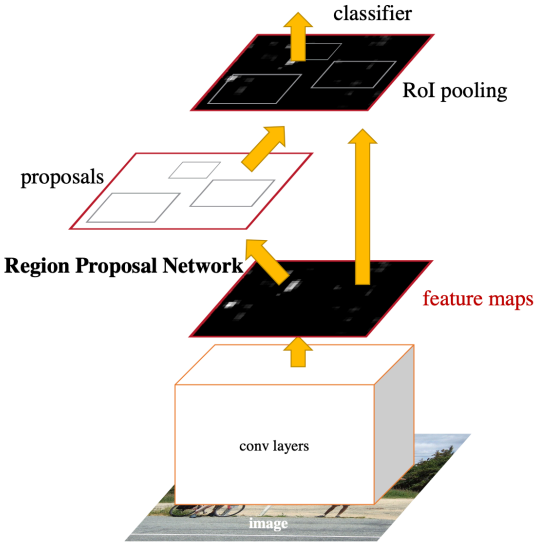


FIGURE 4.3: Faster RCNN architecture with the Region Proposal Network and RoI pooling (Ren et al., 2015)

The Faster R-CNN architecture (Ren et al., 2015) improves on this architecture

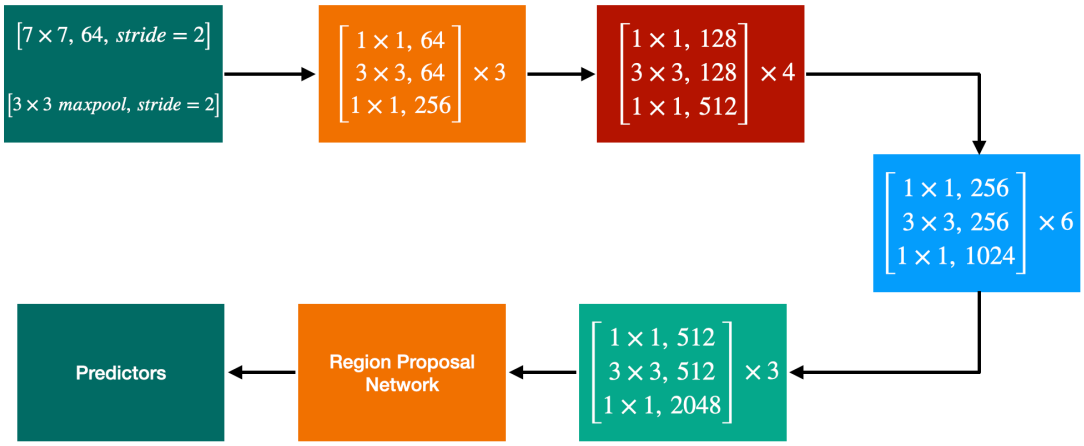


FIGURE 4.4: Illustration of the Resnet-50 Faster RCNN architecture provided by Pytorch

by introducing a Region Proposal Network, a convolutional neural network which is trained together with the CNN to find the best locations within a given image to propose for classification. This is illustrated in Figure 4.3.

The Resnet-50 Faster RCNN model is a 50-layer architecture which follows the Faster RCNN architecture. These layers are divided into 5 different blocks which can be frozen during training and their weights preserved or unfrozen and assigned new weights during training. The arrangement of these layers are described in Figure 4.4.

### 4.2.4 Training and Evaluation of Model Performance

In order to simplify the model and improve performance, the classes in the dataset were summarized into 2 main classes. The first was 'plastics', which covered all instances of plastics and the second class was named 'other' and this comprised all

other classes except the 'timestamp' class. The timestamp class was not considered due to the fact that future explorations would not have any of such inscriptions in the frame.

**Model Construction Parameters:**

The *torchvision.models.detection.fasterrcnn_resnet50_fpn* model provided by Pytorch received certain parameters under different circumstances (Pytorch, n.d.). During the construction of the model for training purposes, the parameters passed are:

- pre-trained: This specifies whether or not the model is pre-trained on the Common Objects in Context (COCO) 2017 dataset. Since we did not required only a model with a pre-trained backbone, this was set to 'false'.
- num_classes: This specifies the number of output classes including the background. This was set to 3.
- pretrained_backbone: This indicates whether a model with a backbone trained on the Imagenet dataset is desired. This was set to be 'true'.
- trainable_backbone_layers: This indicates the number of blocks which will be unfrozen updated during training.

**Input and Ouput Parameters:**

The model, when constructed, expects all inputs to be tensors whose values lie between 0 - 1. The model takes as inputs the images, bounding box coordinates, and label for each bounding box coordinate. The model returns the losses (classification and regression) for both the RPN and RCNN after training. During inference however, the model receives only the input image as an input. It performs calculations and returns a dictionary containing the bounding boxes, labels, and scores for each of the labels assigned to a bounding box.

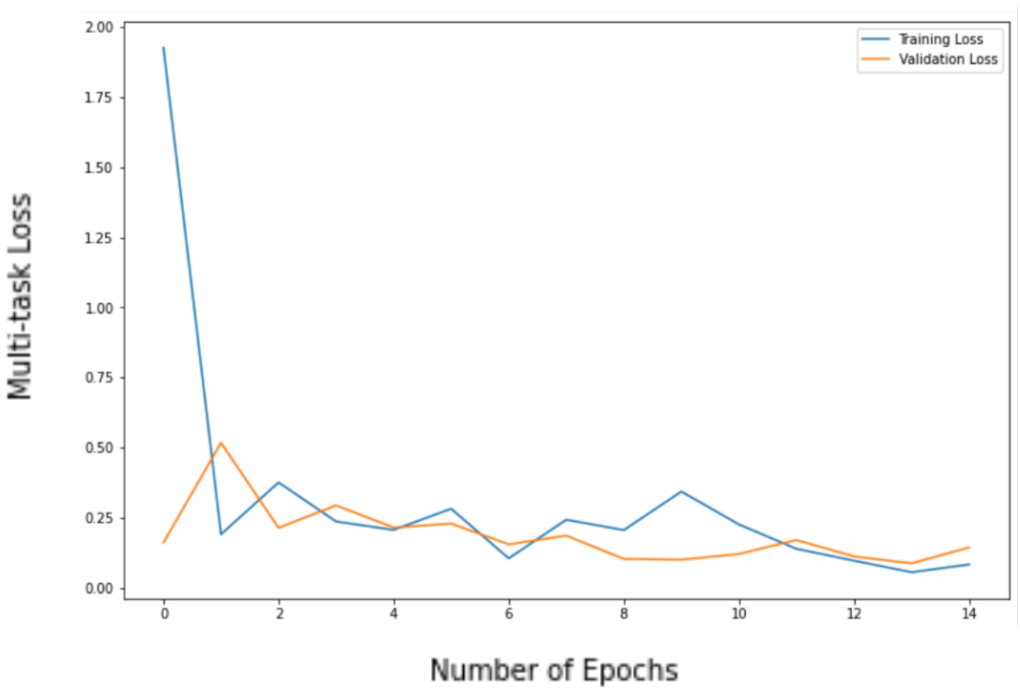


FIGURE 4.5: Multi-task loss over 30 epochs of training

**Training Process and Results**

During the training, Stochastic Gradient Descent is used as the optimizer with a learning rate of 0.003. A learning rate scheduler is used to adjust the learning rate as the training progresses and the model is trained for 40 epochs. After each epoch, the model returns a summary of various metrics such as the multi-task loss, the average precision and average recalls at various Intersection over Union (IoU) values. The multi-task loss comprises the classifier loss, the bounding box regressor loss, the objectness loss and the loss from the RPN. The loss is plotted as shown in Figure 4.5.

The training classifier loss at the beginning of the first epoch is 1.881, and this forms the biggest part of the multi-task loss which is a total of 1.9624. The classifier loss reduces over the epochs and plateaus to 0.0223 at the thirtieth epoch. The bounding box regressor loss begins at a value of 0.0207 and slightly increases to 0.0436. The objectness loss at the last epoch is at a value of 0.0002, falling from an initial value of 0.6879. The RPN loss was at an initial value of 0.0298 and falls to a value of 0.0014. The validation loss values follow similar trajectories.

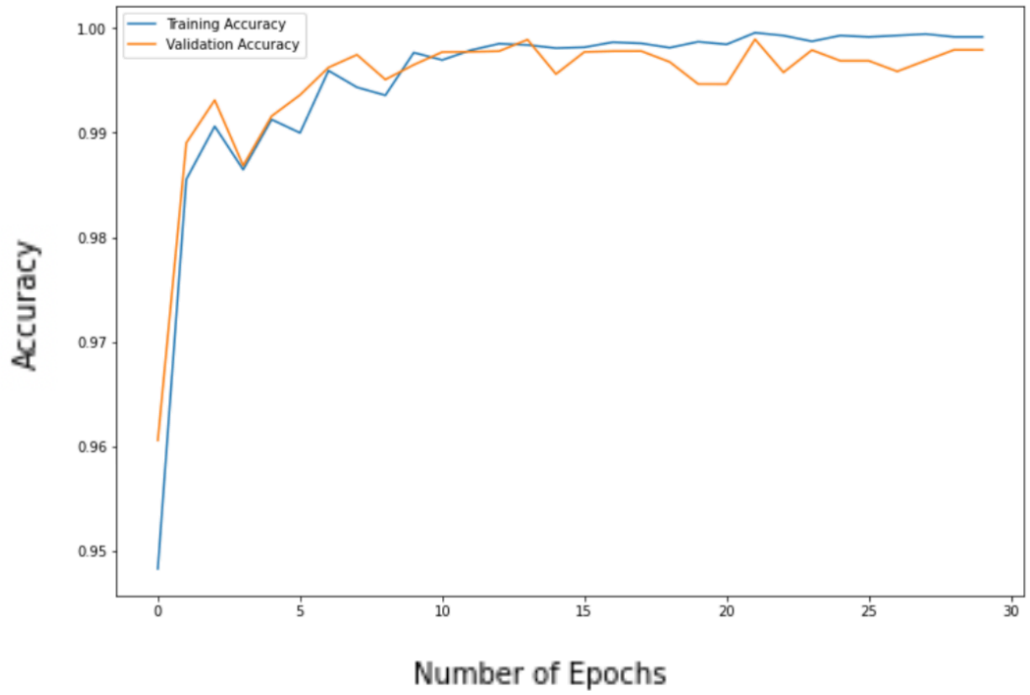


FIGURE 4.6: Model accuracy over 30 epochs of training

The accuracy plot is illustrated in Figure 4.6.

The model's accuracy on the training set has an initial value of 0.95 and increases steeply and plateaus after the 15th epoch. The increase in accuracy after the 15*th* epoch is slight. The validation accuracy is similar in nature to the training accuracy. The accuracy increases until the 10*th* epoch and oscillates between 0.995 and 0.998.

**Predictions**

The test dataset was isolated from the training process. After training the model was used to make predictions on some images from the test set. Figure 4.7 shows the predictions from the model on the images along with their probability scores as compared to their ground truths.

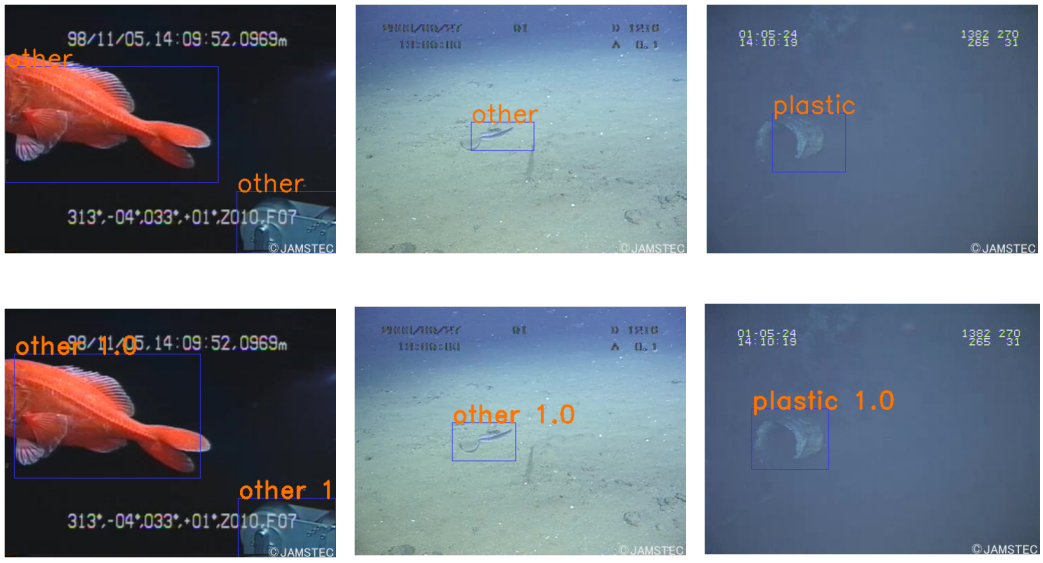


FIGURE 4.7: Predictions from the test dataset compared to their ground truth

# 5 VLC AND TRAJECTORY CONTROL

In this chapter, the communication system design is considered in its constituent parts.

- A low-cost software defined testbed is proposed to ascertain the feasibility of optical communication between robots at the physical layer in the absence of some high-cost elements. The testbed construction will be analysed and the purpose of each part will be examined.
- When a reliable communication link is set up, the trajectory of the robots can be planned. Two control algorithms (centralized and distributed) will be proposed to solve the scheduling and relay problem. The spatial-temporal dynamics of the underwater environment are analysed and considered in the design of the algorithms.

## 5.1 VLC between Transmitter and Receiver

To simulate the physical communication between two robots, a testbed which is a low-cost implementation of the work in (Carver et al., 2020a), is set up. The main difference lies in the addition of re-configurable peripherals which can be employed interchangeably depending on the scenario. The logical flow of data from the transmission to reception is illustrated in Figure 5.1.

**Transmitter**

The transmitted signal is generated using an FPGA, which is used interchangeably with a function generator. The FPGA is used to generate a random sequence and modulate it using Pulse Position Modulation (PPM). The function generator on the

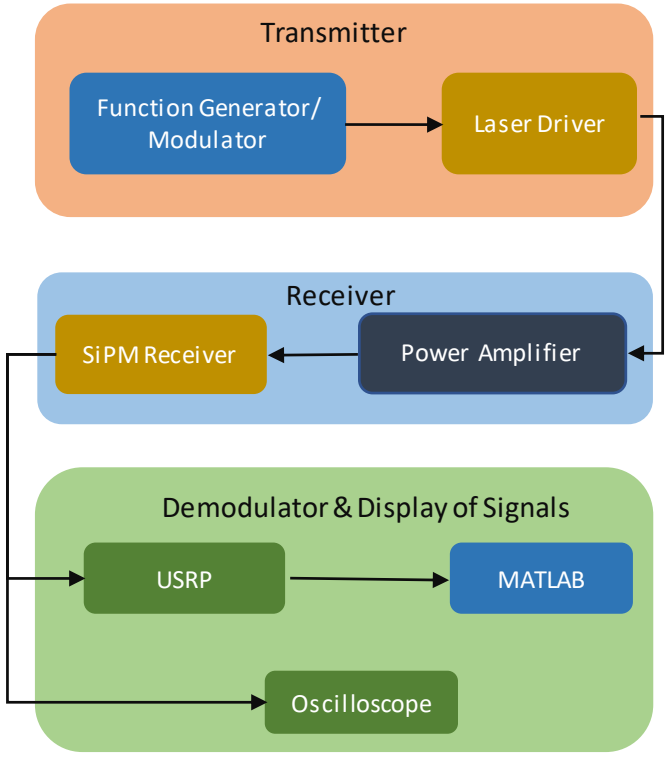


FIGURE 5.1: Summary of transmitter-receiver system design.

other hand is used to generate a simple periodic square wave signal for transmission. This data is fed to the laser diode in order to modulate the light signal.

### Transmission Medium and Receiver

The signal is transmitted over two boundaries: the air/glass boundary and the glass/water boundary. Despite potential concerns about the reflective properties of both boundaries, the signal does not experience strong reflections. The signal is received using a SiPM which is connected to a power amplifier.

### Demodulation and Data Processing

For demodulation, two stations are used interchangeably. The first is a simple oscilloscope with which the signal from the power amplifier is observed. For the analysis of the data, this is carried out using a USRP N210 with a LFRX daughter board which is connected to a computer via a Gigabit ethernet cable and a router. Using Simulink, a virtual model of the USRP is created and employed to read the

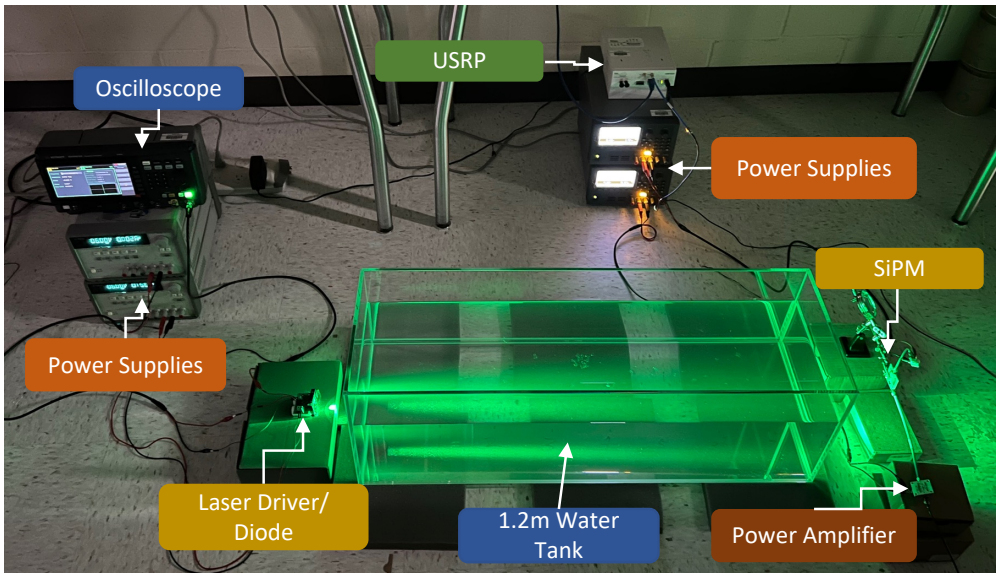


FIGURE 5.2: Setup of the developed VLC testbed.

signal being received. During data processing, parameters such as the USRP gain and threshold of detection have to be adjusted in order to maintain reliable communication.

During transmission, the function generator is tuned to produce a square wave signal at 500 kHz, 1 MHz, and 2 MHz, all with a 50% duty cycle. This signal is connected to the LMG1020-EVM Nano-second laser driver which is populated with a laser diode which outputs a 530 nm laser signal. The laser driver has 3 inputs: a DC bias input that ranges from 5.5 V to 15 V, a DC input voltage directly applied to the laser diode which ranges from 0 V to 75 V and the input port which receives the input signal from the function generator. Both the laser diode and SiPM are elevated slightly above ground level and are placed in a directly line of sight as shown in Fig. 5.2.

To test the robustness of the transmission, as will be seen in Section 4, communication is attempted with the receiver shifted slightly out of the direct line of sight of the transmitter. The positions of the receiver relative to the transmitter are illustrated in Fig. 5.3.

At the receiver end, the SiPM is biased using a voltage of 28 V with the maximum recommended voltage being 30 V. The SiPM had a spectral range ranging

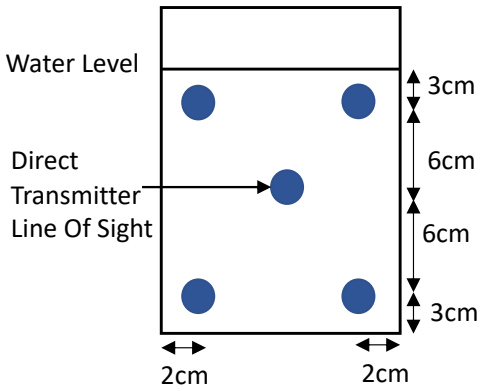


FIGURE 5.3: Receiver positions at the other end of the tank (right-hand side of the tank). The transmitter is on the left-hand side of the tank.

from 300 nm to 950 nm and provided satisfactory performance at 530 nm. The fast output of the SiPM was connected to a power amplifier which provided necessary amplification of the signal from around 500 mV to as much as 1.97 V. The USRP receives the signal from the power amplifier and is connected to a computer for signal demodulation. The connection is facilitated using a NETGEAR router and two Gigabit ethernet cables. Each data point in all the results to be discussed was repeated 5 times under the same conditions to ensure consistency and investigate the variability of the property under consideration. It is worth mentioning that these experiments were performed generally under low ambient light conditions.

## 5.2 Comparison of Air/Water Communication

The output was first observed using an oscilloscope to verify the presence of the signal at the output and the correct working of both the transmitter and receiver. As illustrated in Fig. 5.4, the topmost signal is the signal obtained from the function generator. The laser driver shortens the pulse width of the signal and hence, the output signal records relatively thin spikes in the presence of a 1 bit. As seen in the second row of the diagram, the received signal follows that characteristic. The third row shows a signal captured from a different moment in time through MATLAB when the output is connected through the USRP. It is noticeable that the MATLAB

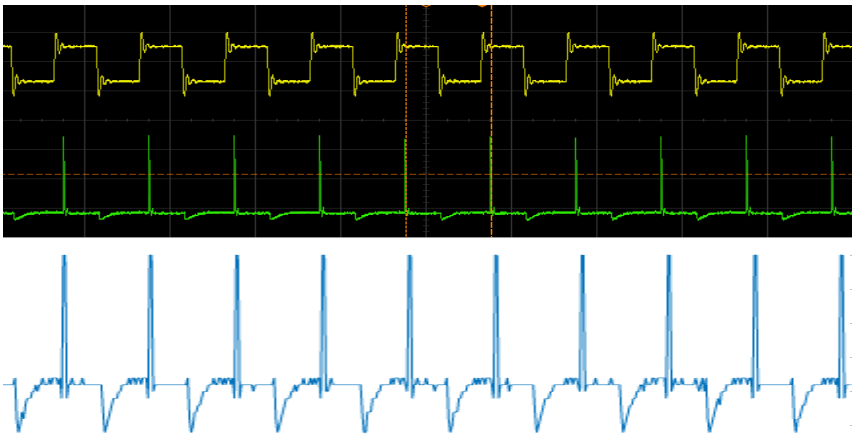

FIGURE 5.4: Screenshots of received signals from the oscilloscope and MATLAB sampled at different instances

signal shows more pronounced spikes in the negative direction compared to the signal directly from the power amplifier. This is due to the internal gain of the USRP which was increased manually to enable better analysis of the waveforms.

The USRP model in Simulink was employed in the sampling of the received data. The model, which originally had sampling rate of 100 MSamples/second was configured to obtain 2000 samples per data frame. This configuration provided a data typically of size $50001 \times 2000$. This data comprised some leading zeros which were filtered out of the data set during analysis. The SiPM, by nature, comes with a recovery time which follows a rise in current triggered by a photon interacting with the photo multiplier. Thus, some bits were lost in transmission and were only partially received. Such instances were also filtered out in order to ascertain the true "fully on" performance of the SiPM. The detection of a 1 *bit* or a 0 *bit* was determined by the presence or absence of a rising edge within a given cycle. This threshold, alongside the number of 1 and 0 to be expected per frame, were chosen from continual observation of the signal waveforms in MATLAB. Using this threshold, the number of bits received were calculated and the bit error rate was calculated by comparing a current bit to the next incoming bit. For each frame containing 2000 samples, the BER is calculated and the average BER is taken for over all the available frames.

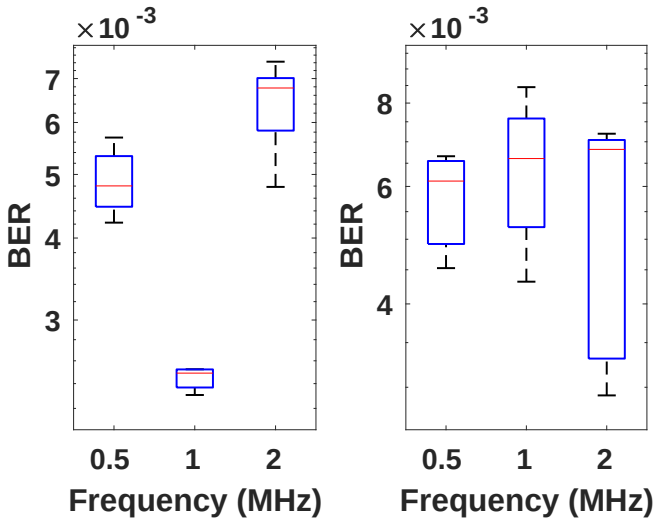


FIGURE 5.5: BER performance through air (Left) and through water (Right)

### 5.2.1 BER: Air vs Water

The performance of the set up in through-air communication was first investigated. The receiver was placed in direct line-of-sight (LOS) at a distance of $1m$ away from the transmitter. In Fig. 5.6, the BER is analysed at different frequencies. The set up was repeated with transmission being carried out through water. The BERs in both scenarios were on the order of $10^{-3}$. In air, communication at 1 MHz recorded the lowest average BER of the three transmissions, with little variation over the transmissions. During transmission through water at 2MHz, 2 bits were transmitted during each cycle thus providing a bit rate of 4 Mbps with an average BER of $5.5 \times 10^{-3}$.

### 5.2.2 BER: Communication over various distances and relative positioning.

The communication link through air was set up over varying distances. The BER performance was analysed when the distance between the transmitter and receiver were 1 m, 2 m, and 4 m, respectively. At a distance of 2 m, the gain variable in the Simulink model was increased in order to improve the detection performance.

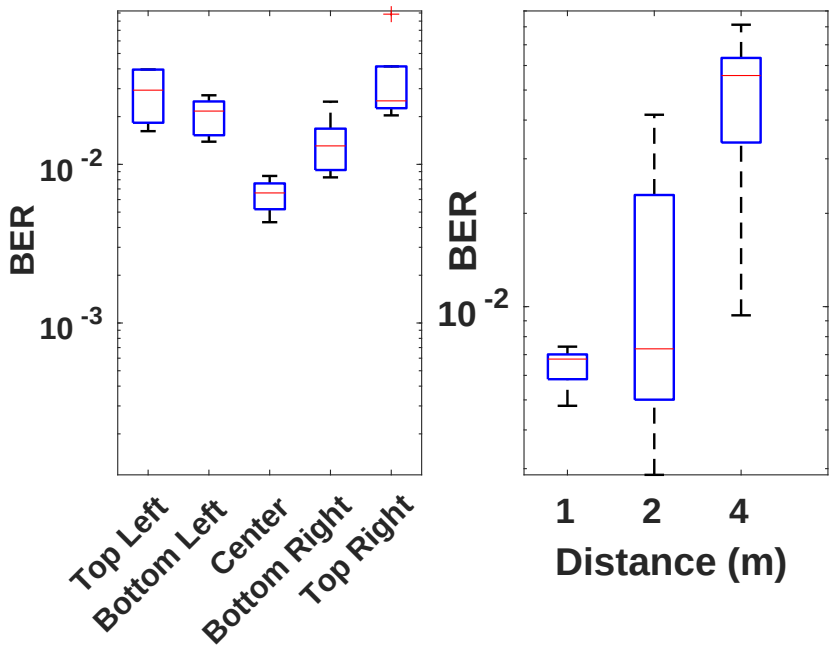


FIGURE 5.6: BER performance over varying transmitter-receiver alignments (Left) and over variable distances (Right)

When the distance was 4 m, not only was the gain increased, but the detection threshold was reduced in order to maintain reliable communication.

As seen in Fig. 5.6, the BER slightly rises and also experiences more variance with increasing distances. Generally, optical communication heavily depends on the maintenance and integrity of a direct line-of-sight between transmitter and receiver. During the communication through water, to ascertain the behavior of the system under non-optimal LOS conditions, the receiver was shifted away from the direct LOS of the transmitter as illustrated in Fig. **??**. The results show a slight raise in the average BER as the receiver deviates from the direct LOS of the transmitter. However, the system still maintains a high level of reliability under those conditions.

## 5.3 Relay and Trajectory Control

Communication using the testbed in Section 5.1 is a very simple implementation of the case scenario. In the ocean, the problem becomes more complex. Before

the trajectory control problem is addressed, this study introduces the underwater optical channel with turbulence-induced fading and the spatial heterogeneity of the attenuation loss. These factors make the underwater wireless channel a more complex problem to analyse. A centralized trajectory control algorithm is designed to obtain robots' optimal locations to maximize the network throughput by having global knowledge of all the robots within the given network. The inefficiencies of the centralized algorithm leads to the proposal of a simpler and more efficient distributed trajectory control algorithm. The key feature of the developed algorithm is that it is adaptive to changes in the environment and it does not require any global information about the robots within the network or the environment itself.

### 5.3.1 Underwater Optical Communication Channel

In a typical scenario as illustrated in Figure 1.1, data communication starts from the anchor robot, i.e., the robot at the lowest depth, monitoring the target. There are $n_r$ robots, and for the $i$th robot its location is $\boldsymbol{x}_i \in \mathbb{R}^3$. Robots communicate via the underwater optical communication channel to set up the network and transmit pertinent information. The $i$th robot sends data to the $(i+1)$th robot using optical communication, as shown in Fig. 1.1.

In this thesis, it is considered that robots can exchange their location information with their neighbors using acoustic-based communication. Although high-speed data communication uses optical communications, robots can use acoustic signals for motion coordination and other critical low-speed data communication.

The wireless optical communication channel is affected by multiple complicated factors. Some works (Johnson, Green, and Leeson, 2013) have been dedicated to studying the relationship between water (ocean) composition and the attenuation that underwater optical links encounter. Aside attenuation, optical communication is also affected by turbulence-induced fading. In order to correctly address the challenges and make the best use of the underwater channel, it becomes necessary to

model the underwater environment.

**Attenuation**

The attenuation loss in the underwater environment comprises two wavelength-dependent factors: absorption and scattering. Absorption occurs when a part of the light energy is changed into another form as a result of its interaction with the environment. Scattering on the other hand occurs when a part of the light beam is redirected due to reflection, refraction, or diffraction of the light beam. Aside the wavelength dependence, the attenuation loss can be modeled as a function of the depth $-z$. The relationship between attenuation, absorption, and scattering can be expressed as (Johnson, Green, and Leeson, 2013)

$$c(\lambda, z) = a(\lambda) + b(\lambda, z), \tag{5.1}$$

where $c(\lambda)$ represents the attenuation coefficient, $a(\lambda)$ is the absorption coefficient, $b(\lambda)$ represents the scattering coefficient, and $\lambda$ is the wavelength.

**Absorption**

As mentioned earlier, absorption is dependent on the wavelength of operation. Aside that, environmental conditions such as the type of water, the amount of chlorophyll, and the concentrations of fulvic and humic acids play an important role in determining the absorption coefficient. The absorption coefficient can be written as (Haltrin, 1999)

$$\begin{aligned} a(\lambda) = a_w(\lambda) + a_f^0 C_f e^{-k_f \lambda} \\ + a_h^0 C_h e^{-k_h \lambda} + a_c^0(\lambda) \left( C_c / C_c^0 \right)^{0.602}, \end{aligned} \tag{5.2}$$

where $a_w(\lambda)$ is the absorption coefficient of pure water in $\mathrm{m}^{-1}$, $a_f^0$ is the absorption coefficient of fulvic acid, $a_h^0$ is the absorption coefficient of humic acid, constants $k_f$

and $k_h$ are the exponential coefficients of fulvic and humic acids, respectively, and $a_c^0(\lambda)$ represents the chlorophyll-a absorption coefficient of living phytoplankton. The value of $a_c^0(\lambda)$ can be derived from the expression (Johnson, Green, and Leeson, 2013)

$$a_c^0(\lambda) = A(\lambda){C_c}^{-B(\lambda)}, \tag{5.3}$$

where $A$ and $B$ are wavelength-dependent constants. At $530nm$, $A = 0.0117$ and $B = 0.139$ (Bricaud et al., 1995). $C_c$ is the concentration of chlorophyll-a and ${C_c}^0 = 1$ mg/m$^3$. $C_f$ and $C_h$ are the concentrations of fulvic acid and humic acids respectively and are computed as below (Johnson, Green, and Leeson, 2013):

$$C_f = 1.74098 C_c e^{0.12327 C_c} \tag{5.4}$$

$$C_h = 0.19334 C_c e^{0.12343 C_c} \tag{5.5}$$

**Scattering**

Generally, scattering occurs due to the particles in water. The scattering coefficient is determined according to (Haltrin, 1999):

$$b(\lambda, z) = b_w(\lambda) + b_s^0(\lambda) C_s + b_l^0(\lambda) C_l \tag{5.6}$$

The pure water scattering coefficient is $b_w(\lambda)$, and the variables ${b_s}^0(\lambda)$ and ${b_l}^0(\lambda)$ represent the scattering coefficients for small and large particles respectively. $C_s$ and $C_l$ are the concentrations of small and large particles within the underwater environment, respectively. The values for the variables can be calculated according to the following expressions (Johnson, Green, and Leeson, 2013)

$$b_w(\lambda) = 0.005826(400/\lambda)^{4.322} \tag{5.7}$$

$${b_s}^0(\lambda) = 1.1513(400/\lambda)^{1.7} \tag{5.8}$$

$$b_l{}^0(\lambda) = 0.3411(400/\lambda)^{0.3} \tag{5.9}$$

$$C_s = 0.01739 C_c e^{0.11631 C_c} \tag{5.10}$$

$$C_l = 0.76284 C_c e^{0.0309 C_c} \tag{5.11}$$

The chlorophyll concentration $C_c$ can be defined over depth $z$ and as a function of the background chlorophyll concentration $B_0$, the vertical gradient of concentration $S$, and the Deep Chlorophyll Maximum (DCM) $z_{max}$. The DCM is the subsurface region, generally, between 20 and 120 m (Johnson, Green, and Leeson, 2013), where the chlorophyll concentration is at its maximum due to a balance of light and nutrients. $\sigma_{chl}$ is the standard deviation of the chlorophyll concentration. They are related by (Kameda and Matsumura, 1998):

$$C_c(z) = B_0 + Sz + \frac{h}{\sigma_{chl}\sqrt{2\pi}} exp\left[\frac{-(z - z_{max})^2}{2{\sigma_{chl}}^2}\right] \tag{5.12}$$

The standard deviation of the chlorophyll is a function of the total chlorophyll above the background levels $h$, the depth of the DCM, and the chlorophyll level $C_{chl}$ at the DCM, which is

$$\sigma_{chl} = \frac{h}{\sqrt{2\pi\left[C_{chl}(z_{max}) - B_0 - Sz_{max}\right]}} \tag{5.13}$$

Considering the definition of the attenuation coefficient $c(\lambda, z)$ as a function of the wavelength and depth beneath the water surface, given a wavelength $\lambda$ of 530 nm, the variation of the attenuation against the depth can be plotted. Assuming that the change in the absorption and scattering coefficients over a horizontal distance in the open ocean environment is negligible (Johnson, Green, and Leeson, 2013), the attenuation profiles for four different chlorophyll vertical gradients S1, S2, S3, and S4 are shown in Fig. 5.7. The vertical gradients represent different chlorophyll profiles with regard to their concentration. S1 represents a region where the chlorophyll concentration is generally less than 0.04 $mg/m^3$ and S2 represents the profile

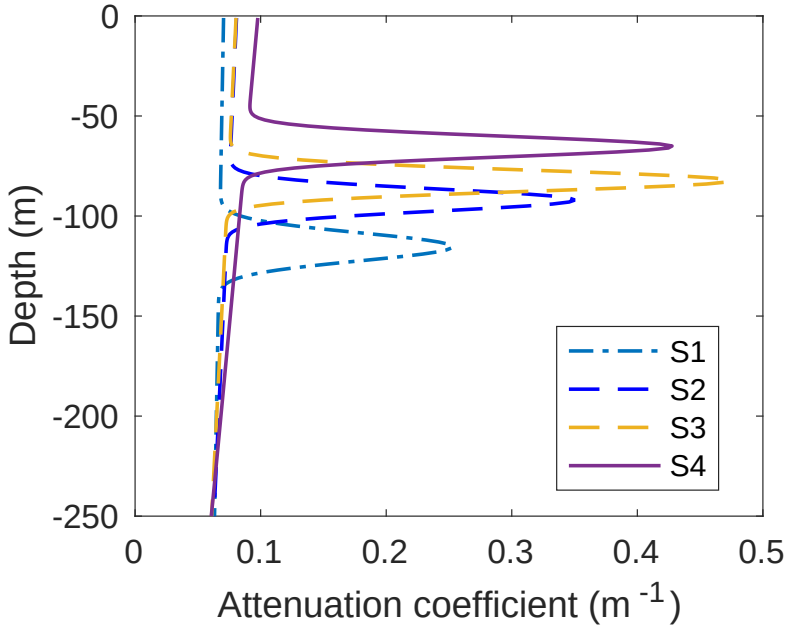


FIGURE 5.7: Variation of attenuation coefficient with increasing depth over 250 m, calculated over four vertical gradients of chlorophyll. The environmental parameters for S1, S2, S3, and S4 are from (Johnson, Green, and Leeson, 2013).

within which the chlorophyll concentration lies between $0.04 - 0.08$ mg/m$^3$. The detailed parameters of S1 and S4 are from (Johnson, Green, and Leeson, 2013). The attenuation coefficient is calculated over a depth of 250 m below the surface of the ocean. From Fig. 5.7, the attenuation coefficient gradually increases within the first 50 m below the surface of the ocean. As the depth increases, the attenuation coefficient experiences a sharp increase in value, followed by a sharp decrease after it reaches its peak value. For different underwater environments, the largest attenuation coefficient appears at different depths. Not having thorough knowledge of all the possible factors makes it challenging to obtain an accurate channel model.

**Water Turbulence**

The water turbulence is modeled as stochastic fading, which is another environmental effect on signal propagation. During propagation, the optical signal experiences intensity fluctuations, known as scintillation, which eventually results in optical turbulence. These fluctuations are best characterized using the scintillation index $\sigma$ which is proportional to the Rytov's variance ${\sigma_r}^2$. The Rytov's variance is

given by

$$\sigma_r{}^2 = 37.3K_3 \left(\frac{2\pi}{\lambda}\right)^{7/6} L^{11/6}, \tag{5.14}$$

where $K_3$ is the constant that determines the turbulence strength and $L$ is the migration length (distance) of the light beam (Liu, Xu, and Yang, 2015). The scintillation index can also be derived from the expression

$$\sigma^2 = exp\left[\frac{0.49\sigma_r{}^2}{(1+1.11\sigma_r{}^{12/5})^{7/6}} + \frac{0.51\sigma_r{}^2}{(1+0.69\sigma_r{}^{12/5})^{5/6}}\right] - 1 \tag{5.15}$$

The fading variable is dependent on the scintillation index, and it satisfies the log-normal probability distribution function (Hill and Frehlich, 1997; Liu, Xu, and Yang, 2015)

$$f(I) = \frac{1}{I\sigma\sqrt{2\pi}/I_0} exp\left(-\frac{(ln(I/I_0)-u)^2}{2\sigma^2}\right), \tag{5.16}$$

where $I$ is the received light intensity, $I_0$ is the mean received light intensity, and $u = -\sigma^2/2$. The light intensity is defined as the power per unit area of the optical transmitter or receiver.

Therefore, for a transmit robot at $\boldsymbol{x}_i$ and a receive robot at $\boldsymbol{x}_j$, the mean received optical power is $\tilde{P}_r(\boldsymbol{x}_i, \boldsymbol{x}_j)$, which can be written as (Saeed, Al-Naffouri, and Alouini, 2018; Arnon and Kedar, 2009; Saeed et al., 2019):

$$\tilde{P}_r(\boldsymbol{x}_i, \boldsymbol{x}_j) = P_t\rho_i\rho_j e^{\left(\frac{-\int_{z_i}^{z_j} c(\lambda,z)zdz}{cos\theta_{ij}}\right)} \cdot \frac{B_r cos\theta_{ij}}{2\pi d_{ij}^2(1-cos\theta_0)} \tag{5.17}$$

$$= P_t\rho_i\rho_j \prod_{z_i+n_d\Delta z}^{z_j} e^{\left(\frac{-c(\lambda,z_i+n_d\Delta z)\Delta z}{cos\theta_{ij}}\right)} \cdot \frac{B_r cos\theta_{ij}}{2\pi d_{ij}^2(1-cos\theta_0)} \tag{5.18}$$

where , $n_d \in \mathbb{Z}$, $P_t$ is the optical transmit power, which is a constant for all the robots, and $\rho_i$ and $\rho_j$ are the optical efficiencies of the transmitter and receiver, respectively. $c(\lambda)$ represents the total attenuation and $d_{ij}$ is the vertical distance between the nodes. $\theta_{ij}$ is the trajectory angle between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$. $\theta_{ij}$ is the angle between the perpendicular to the receiver plane and the transmitter-receiver trajectory. $\theta_0$ is

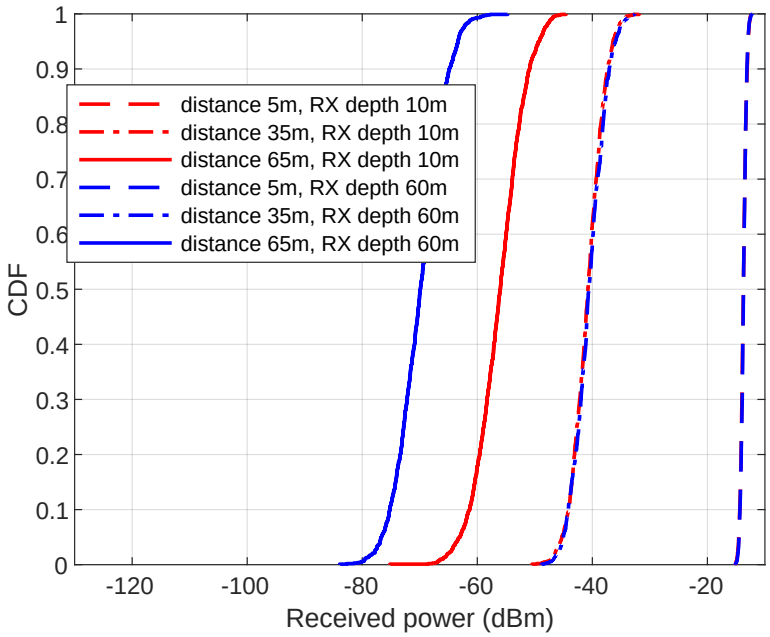


FIGURE 5.8: CDF of received power in presence of water turbulence. The considered water type is S1. The depth of the receiver is 10 m and 60 m and the transmitter is placed at 5 m, 35 m and 65 m under the receiver.

the divergence angle of the transmitter, and $B_r$ is the receiver aperture area.

To obtain the received power with turbulence fading, we use a lognormal random number generator with $\mu$ and $\sigma$ to generate a random variable $F$. Then, the received power with fading can be written as

$$P_r(\boldsymbol{x}_i, \boldsymbol{x}_j) = F \cdot \tilde{P}_r(\boldsymbol{x}_i, \boldsymbol{x}_j) \tag{5.19}$$

In Fig. 5.8, we numerically evaluate the impact of distance and depth on the received power. The receiver depth is 10 m and 60 m. The transmitter is under the receiver with distances 5 m, 35 m, and 65 m. As shown in the figure, when distance is small, the variance of the received power is small, which indicates a reliable wireless optical channel. However, as the distance increases, the variance of the received power increases and the channel becomes more unpredictable. Despite the level of randomness that accompanies the turbulence-induced fading, the plot clearly shows that the received power can be expected to be smaller in magnitude

with increasing transmission distance. Also, at 60 m, the attenuation loss is large, as shown in Fig. 5.7, and the received power is small. This is not obvious when the distance is small. It is because within a short range the variation of attenuation loss is small and the exponential attenuation loss is small.

The results indicate that as robots' distance increases, the wireless optical channel becomes unreliable. This now presents a challenge when coordinating the motion of the robots because the received power do not accurately reflect their distance from each other. Note that, data-driven approaches, such as the reinforcement learning and graph neural networks (Ruiz, Gama, and Ribeiro, 2021; Gama, Bruna, and Ribeiro, 2020) may not perform well in such environments since the data (received power) are highly stochastic and noisy. It is challenging to learn a control policy based on collected data.

**Multi-Robot Relay Network Modeling**

Wireless optical communication using lasers or LEDs is highly directional and there exists the chance that the signal may be blocked by objects. The 1D relay scheduling in (Chattopadhyay et al., 2016) uses RF signals and due to its broadcasting nature, the interference was considered therein. As shown in Fig. 5.9, due to the blockage, the interference in the considered relay network can be neglected provided that robots only move in the interference-free cone. Despite no interference, the problem is still challenging due to the spatial change of the attenuation loss. Next, we first introduce a centralized control policy with knowledge of the wireless optical channel model. Then, we design a distributed control policy, which only uses robots' local observations.

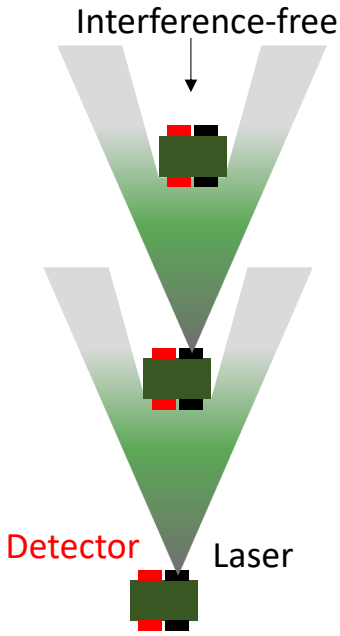


FIGURE 5.9: Illustration of the interference-free region between different robots with optical communication

### 5.3.2 Optimal Trajectory Design

#### Centralized Trajectory Design

In this section, an optimization-based solution is introduced, where each robot has perfect knowledge of the wireless optical channel. The optimal controller has rich information to plan the trajectory of each robot.

To obtain the trajectory of each robot, we solve the following problem

$$\max \min\{P_r(z_i, z_{i+1})\} \text{ for } i = 1, \cdots, n_r - 1 \tag{5.20}$$

$$\text{s.t. } z_1 < z_2 < \cdots < z_{n_r} = 0 \tag{5.21}$$

$$\text{Equation (5.19).} \tag{5.22}$$

Note that, since robots only move vertically, we only use their vertical location $z_i$ instead of the 3D location $\boldsymbol{x}_i$, i.e., $P_r(\boldsymbol{x}_i, \boldsymbol{x}_j)$ is replaced with $P_r(z_i, z_j)$. Also, for the centralized trajectory control, we consider the accurate channel model is available. Thus, the fading is not considered; only the mean received power is used. The objective of the above problem is to maximize the minimum received power between

any two adjacent robots. For a relay network without any network congestion, the overall data rate is limited by the link with the minimum data rate. Also, the data rate is determined by the received power and noise power. In this paper, we consider the noise power is uniform which is not affected by the spatial change. Although underwater ambient noise has different power at different depths, the bandpass filter can effectively reduce the ambient noise (Carver et al., 2020b). Thus, the dominant factor is the received power. By maximizing the minimum received power, we can obtain the maximum data rate of the relay network.

The constraints in (5.21) ensures that the $n_r$th robot will move to the surface. The robot at $z_1$ does not move and keeps monitoring the target. All other robots relay the data to the surface robot. The received power is obtained by using (5.19), which is complicated due to the spatial change of the attenuation coefficient.

To solve the max-min problem, we equivalently solve the following problem

$$\begin{aligned} &\min -P_0 \\ \text{s.t. } &P_0 - P_r(z_i, z_{i+1}) \leq 0, \text{ for } i = 1, \cdots, n_r - 1 \\ &z_1 < z_2 < \cdots < z_{n_r} = 0, \text{ Equation (5.19)} \end{aligned}$$

By minimizing $-P_0$, we equivalently obtain the maximized minimum received power. To find the optimal $\boldsymbol{z}^* = [z_2^*, \cdots, z_{n_r-1}^*]^t$, we use the primal-dual approach to find the saddle point of the Lagrangian associated with the above problem. Since $z_{n_r} = 0$ in the end, we do not optimize $z_{nr}$, but gradually increase it. In other word, the lead robot gradually moves to the water surface. The Lagrangian is

$$\begin{aligned} \mathcal{L}(\boldsymbol{z}, \boldsymbol{\epsilon}, \boldsymbol{\mu}) = &- P_0 + \sum_{i=1}^{n_r-1} \epsilon_i (P_0 - P_r(z_i, z_{i+1})) \\ &+ \sum_{i=1}^{n_r-1} \mu_i (z_i - z_{i+1}), \end{aligned} \tag{5.23}$$

where $\boldsymbol{\epsilon} = [\epsilon_1, \cdots, \epsilon_{n_r-1}]^t$ and $\boldsymbol{\mu} = [\mu_1, \cdots, \mu_{n_r-1}]^t$ are Lagrangian multipliers,

which are nonnegative numbers. The saddle point $(\boldsymbol{z}^*, \boldsymbol{\epsilon}^*, \boldsymbol{\mu}^*)$ satisfies the following relation

$$\mathcal{L}(\boldsymbol{z}, \boldsymbol{\epsilon}^*, \boldsymbol{\mu}^*) \leq \mathcal{L}(\boldsymbol{z}^*, \boldsymbol{\epsilon}^*, \boldsymbol{\mu}^*) \leq \mathcal{L}(\boldsymbol{z}^*, \boldsymbol{\epsilon}, \boldsymbol{\mu}). \tag{5.24}$$

The optimal primal and dual variables can be obtained by alternating the maximization of $\mathcal{L}(\boldsymbol{z}, \boldsymbol{\epsilon}^*, \boldsymbol{\mu}^*)$ and minimization of $\mathcal{L}(\boldsymbol{z}^*, \boldsymbol{\epsilon}, \boldsymbol{\mu})$. Next, we use gradient ascent and gradient decent methods to obtain the optimal values iteratively.

The derivative of the dual variables are

$$\begin{aligned}
\frac{d\mathcal{L}(\boldsymbol{z}, \boldsymbol{\epsilon}, \boldsymbol{\mu})}{d\epsilon_i} &= P_0 - P_r(z_i, z_{i+1}), \text{ for } i = 1, 2, \cdots, n_r - 1 \\
\frac{d\mathcal{L}(\boldsymbol{z}, \boldsymbol{\epsilon}, \boldsymbol{\mu})}{d\mu_i} &= z_i - z_{i+1}, \text{ for } i = 1, 2, \cdots, n_r - 1
\end{aligned} \tag{5.25}$$

Then, we can iteratively update $\epsilon_i$ and $\mu_i$ by

$$\begin{aligned}
\epsilon_i^{k+1} &= \epsilon_i^k - (P_0 - P_r(z_i, z_{i+1}))\Delta\epsilon, \text{ for } i = 1, 2, \cdots, n_r - 1 \\
\mu_i^{k+1} &= \mu_i^k - (z_i - z_{i+1})\Delta\mu, \text{ for } i = 1, 2, \cdots, n_r - 1
\end{aligned} \tag{5.26}$$

where $\Delta\epsilon$ and $\Delta\mu$ are step values, the superscript $k$ indicates the iteration step number. Next, by using updated dual variables, the robot location will be updated.

Note that, $P_0 = \min(\{P_r(z_i, z_{i+1}), \text{ for } i = 1, 2, \cdots, n_r - 1\})$, which is a function of $z_i$. Thus, the derivative of $z_i$ is

$$\begin{aligned}
\frac{d\mathcal{L}(\boldsymbol{z}, \boldsymbol{\epsilon}, \boldsymbol{\mu})}{dz_i} &= (\epsilon_i - 1)\left[1_{P_0 = P_r(z_i, z_{i+1})}\frac{dP_r(z_i, z_{i+1})}{dz_i}\right. \\
&\left. + 1_{P_0 = P_r(z_{i-1}, z_i)}\frac{dP_r(z_{i-1}, z_i)}{dz_i}\right] - \epsilon_i\frac{dP_r(z_i, z_{i+1})}{dz_i} \\
&- \epsilon_{i+1}\frac{dP_r(z_{i-1}, z_i)}{dz_i} + \mu_i - \mu_{i+1}, \text{ for } i = 2, \cdots, n_r - 1
\end{aligned} \tag{5.27}$$

where $1_{x_a = x_b}$ is an indicator function which is 1 if $x_a = x_b$ is satisfied. The derivative of the received power cannot be analytically obtained due to the spatial dependent

attenuation coefficient. This study employs a numerical approach to obtain approximations of the derivative. The derivative can be extremely large or small and the robot cannot update its location simply based on derivative. We consider at each time step, the robot move $\Delta z$ and the direction is $\frac{d\mathcal{L}(z,\epsilon,\mu)}{dz_i} / \| \frac{d\mathcal{L}(z,\epsilon,\mu)}{dz_i} \|$. Thus, the location of each robot can be updated as

$$z_i^{k+1} = z_i^k + \frac{d\mathcal{L}(\boldsymbol{z}, \boldsymbol{\epsilon}, \boldsymbol{\mu})}{dz_i} / \| \frac{d\mathcal{L}(\boldsymbol{z}, \boldsymbol{\epsilon}, \boldsymbol{\mu})}{dz_i} \| \Delta z. \tag{5.28}$$

Once robots obtain new locations, the received power and $P_0$ will be updated and a new round starts by updating $\boldsymbol{\epsilon}$ and $\boldsymbol{\mu}$. This process continues until the variance of robots' locations changes or the received power variance is smaller than a threshold.

Although this centralized optimization approach is rigorous and comprehensive, it demonstrates the following drawbacks for real implementation.

- First, the solution may be trapped in a local optimal value and a globally optimal solution cannot be guaranteed. The obtained solution is suboptimal.
- The derivative is based on numerical approximations, which may have approximation errors. Also, without the knowledge of the underwater optical channel model, we cannot obtain it. As discussed in the preceding sections, the underwater environment is highly dynamic and an accurate channel model usually is not available.
- The scalability of the approach is not guaranteed, i.e., as the number of robots increases, the dimension of the solution increases, and the solution may not converge.

In general, the centralized optimization method requires the knowledge of the wireless optical channel and the solution is suboptimal due to the complexity of the wireless optical channel. Next, we provide a distributed solution based on (5.27) and (5.28) to address the above issues.

**Distributed Trajectory Design**

The distributed trajectory design for multi-robot relay has been studied in (Zavlanos, Ribeiro, and Pappas, 2012; Jackson et al., 2020). The difference compared to this study is that the wireless channel is modeled based on distance without considering fading issues, and the wireless channel model is also available. Our distributed control solution is different from existing works because it only uses the received power from neighbors to adjust robot locations without global knowledge. Also, we do not require a wireless channel model to plan robot trajectories. Since the robot does not have an accurate channel model for distributed trajectory control, we will consider turbulence and fading in this subsection.

From (5.28), we learn that the robot updates its location based on the derivative of the Lagrangian function. In (5.27), if $\frac{dP_r(z_i,z_{i+1})}{dz_i} > 0$ due to the increase of $z_i$, the $i$th robot and the $(i+1)$th robot get close and the received power by the $(i+1)$th robot increases. If this term is dominant, i.e., the received power by the $(i+1)$th robot increases dramatically, $\frac{d\mathcal{L}(\boldsymbol{z},\boldsymbol{\epsilon},\boldsymbol{\mu})}{dz_i}$ is negative since $\epsilon_i \geq 0$. Then, according to (5.28), the robot will move downwards and get away from the $(i+1)$th robot. Similarly, if $\frac{dP_r(z_{i-1},z_i)}{dz_i} < 0$ due to the increase of $z_i$, the received power of the $i$th robot decreases. If $\frac{dP_r(z_{i-1},z_i)}{dz_i}$ is a dominant term, $\frac{d\mathcal{L}(\boldsymbol{z},\boldsymbol{\epsilon},\boldsymbol{\mu})}{dz_i}$ is nonnegative. According to (5.28), the robot moves upward to get away from the $(i-1)$th robot. Finally, when $z_i$ converges to the optimal location $z_i^*$, $\frac{d\mathcal{L}(\boldsymbol{z},\boldsymbol{\epsilon},\boldsymbol{\mu})}{dz_i} = 0$. Under the optimal condition, all the received power are equivalent. This is trivial to prove because increase any of the received power will decrease another one which reduces $P_0$. Thus, in (5.27) both of the two indicator functions are valid and we need

$$\frac{dP_r(z_i,z_{i+1})}{dz_i} + \frac{dP_r(z_{i-1},z_i)}{dz_i} = 0, \tag{5.29}$$

where we implicitly assume $\mu_i \ll \epsilon_i$. Generally, if the received power of a wireless link increases, the transmitter moves downward and the receiver moves upward. When we consider fading, $\frac{dP_r(z_i,z_{i+1})}{dz_i}$ and $\frac{dP_r(z_{i-1},z_i)}{dz_i}$ are random numbers, we can use

their expectations.

Our distributed solution is based on this observation: the robot moves towards the opposite direction of the change of received power. The challenge is how to analytically develop the relationship between the robot motion and the change of received power to ensure that robots can move to the optimal locations.

Consider that a robot is at $z_i$ and its upper neighbor is at $z_{i+1}$ and lower neighbor is at $z_{i-1}$. We assume the $i$th robot has the information about $z_{i+1}$ and $z_{i-1}$ by using underwater localization, which is usually obtained by using acoustic signals. The received power by the $i$th robot is $P_r(z_{i-1}, z_i)$, and the received power by the $(i+1)$th robot is $P_r(z_i, z_{i+1})$, which will be feedback to the $i$th robot. Then, the $i$th robot uses a simple channel model to estimate the attenuation loss, which is

$$\exp(-c_1(z_{i+1} - z_i)) = \mathbb{E}\left(P_r(z_i, z_{i+1})\right) \tag{5.30}$$

$$\exp(-c_2(z_i - z_{i-1})) = \mathbb{E}\left(P_r(z_{i-1}, z_i)\right), \tag{5.31}$$

where $\mathbb{E}(\cdot)$ is the expectation which can be obtained by collecting multiple samples and find the mean received power. The estimated attenuation loss is

$$c_1 = \frac{\log \mathbb{E}\left(P_r(z_i, z_{i+1})\right)}{z_i - z_{i+1}} \quad c_2 = \frac{\log \mathbb{E}\left(P_r(z_{i-1}, z_i)\right)}{z_{i-1} - z_i}. \tag{5.32}$$

Next, using $c_1$ and $c_2$, we need to obtain an updated $z_i$ to satisfy the requirement in (5.29). Therefore, we have

$$\frac{d\mathbb{E}\left(P_r(z_i, z_{i+1})\right)}{dz_i} + \frac{d\mathbb{E}\left(P_r(z_{i-1}, z_i)\right)}{dz_i} = c_1(z_{i+1} - z_i)$$
$$- c_2(z_i - z_{i-1}) = 0. \tag{5.33}$$

The optimal $z_i$ is

$$z_i^* = \frac{c_1 z_{i+1} + c_2 z_{i-1}}{c_1 + c_2}. \tag{5.34}$$

If the current location is $z_i$ and the next time step location is $z_i^*$, then the $i$th robot updates its location based on

$$z_i^{k+1} = z_i^k + \frac{\log \mathbb{E}\left(P_r(z_{i-1}, z_i)\right) - \log \mathbb{E}\left(P_r(z_i, z_{i+1})\right)}{c_1 + c_2}. \tag{5.35}$$

The above design can converge to the optimal location since when the two links have the same received power, the numerator in the second term of (5.35) becomes zero and the robot does not move if its neighbors do not move. Since $c_1$ and $c_2$ are positive numbers, if the lower received power is higher, the robot moves upward, and if the upper received power is higher, the robot moves downward.

Compared to the centralized solution, the above approach only requires local observation of the received power and the robot locations without knowing any global channel model or robot status. Since the algorithm is fully distributed, it is scalable.

## 5.4 Trajectory Control Simulation

In this section, we numerically evaluate the performance of the proposed centralized and distributed algorithms. We mainly focus on two aspects: the improvement of data rate by using multiple robots and the adaptability of the algorithms in different underwater environments.

### 5.4.1 Number of Robots

We consider the water type as S1, and the parameters are given in (Johnson, Green, and Leeson, 2013). The depth of the target is $d_t = 150$ m. First, we consider 5 robots and compare the performance of the centralized and distributed algorithms. The initial location of the $i$th robot is $[-d_t + 5(i-1)]$. The 1st robot is the anchor robot, which does not move, and the last robot is the lead robot, which moves toward the surface with a constant step $z_{step}$. The laser transmission power is 0.1 W, the

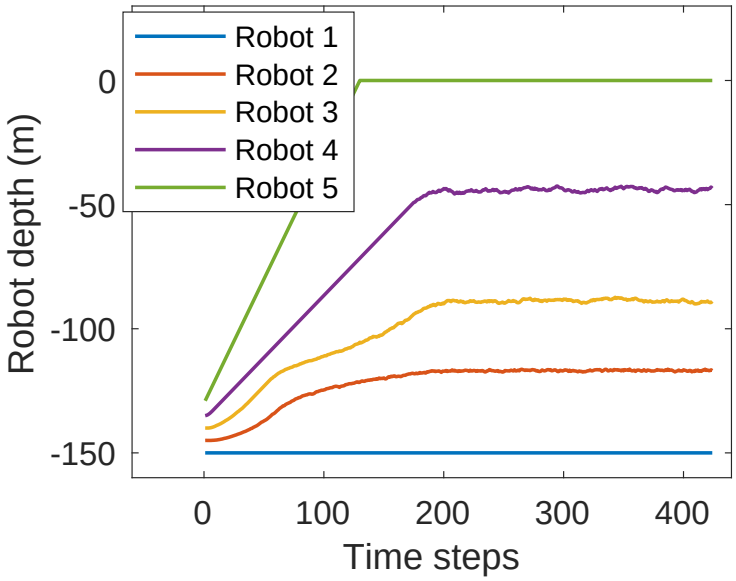


FIGURE 5.10: Robot vertical trajectory using distributed algorithm. The final locations of robots are [-150.0, -116.8, -89.0, -43.4, 0.0] m. The received power is [-44.8, -43.6, -46.9, -46.2] dBm.

divergence angle is $\pi/10$, and the receiver aperture area is 0.005 $m^2$. Note that, robots only move upward and downward to find their optimal depth. Moving towards any other direction may increase interference.

First, we consider $z_{step} = 0.5$ m. The distributed algorithm is shown in Fig. 5.10 and the centralized algorithm is shown in Fig. 5.11. As shown in the figures, both the distributed and the centralized algorithms can guide robots to locations with similar received power for each relay link. The variance of the received power using the distributed algorithm is smaller than that using the centralized algorithm. This is mainly due to the value of $z_{step}$ and the approximation of the received power derivative.

In Fig. 5.12, the $z_{step}$ is reduced from 0.5 m to 0.1 m. As we can see from the figure, the received power variance is slightly reduced with minor location changes.

In Fig. 5.13, we increase the robot number from 5 to 15 with a step of 5. We run each scenario 10 times and show the mean and standard deviation in the figure. Since the distributed algorithm performs better in terms of computing efficiency and results, we only evaluate it here. As we can see, as the robot number increases the received power also increases and the variance of the received power becomes

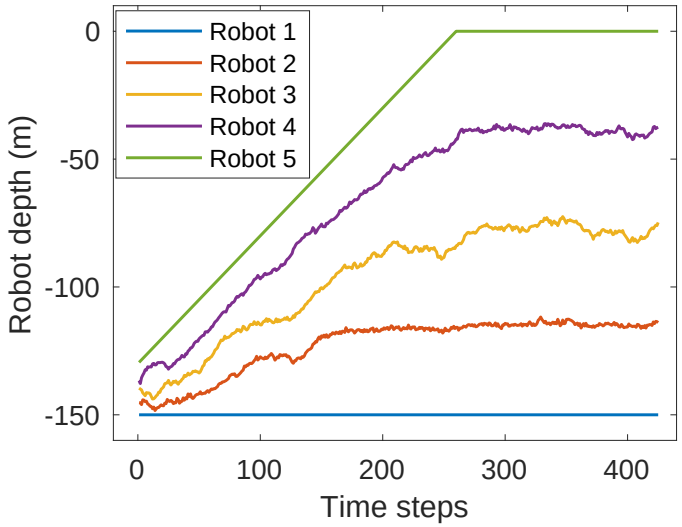


FIGURE 5.11: Robot vertical trajectory using centralized algorithm with step 0.5 m. The final locations of robots are [-150.0, -114.0, -75.5, -38.2, 0.0] m. The received power is [-49.1, -48.0, -41.0, -41.7] dBm.

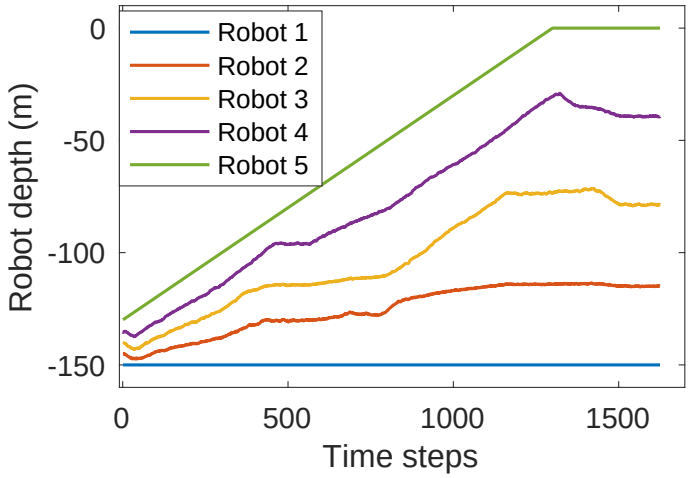


FIGURE 5.12: Robot vertical trajectory using centralized algorithm with step 0.1 m. The final locations of robots are [-150.0, -114.8, -78.6, -39.4, 0.0] m. The received power is [-47.9, -47.5, -42.0, -42.4] dBm.

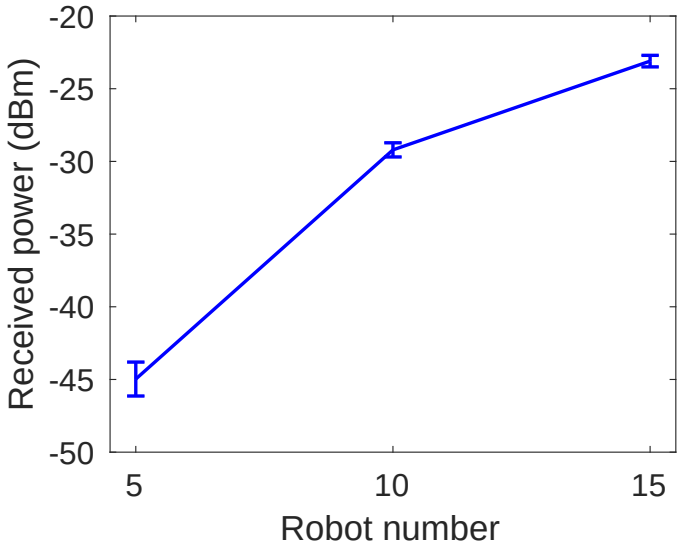


FIGURE 5.13: Impact of robot number on final received power when robots moved to their optimal locations.

smaller. More relay robots reduce the distance between two robots and increase the received power. However, the gain of using more robots gradually decreases, i.e., the gain is not linear.

### 5.4.2 Dynamic Underwater Environment

The main motivation to develop a simple distributed algorithm is to adapt to dynamic underwater environments. Next, we evaluated the distributed algorithm without environment knowledge and the centralized algorithm in different underwater environments, i.e., the S1, S2, S3, and S4 defined in (Johnson, Green, and Leeson, 2013). We obtain the mean value and the standard deviation of the received power of the 4 links between 5 robots. For the distributed algorithm in Fig. 5.14, we observe that the performances in different underwater environments are similar without dramatic change. Since the distributed algorithm only uses locale observations, it does not require global environmental information nor the channel model. Its performance is robust in dynamic underwater environments. For the centralized algorithm, we consider the S1 channel model is available, but the other three

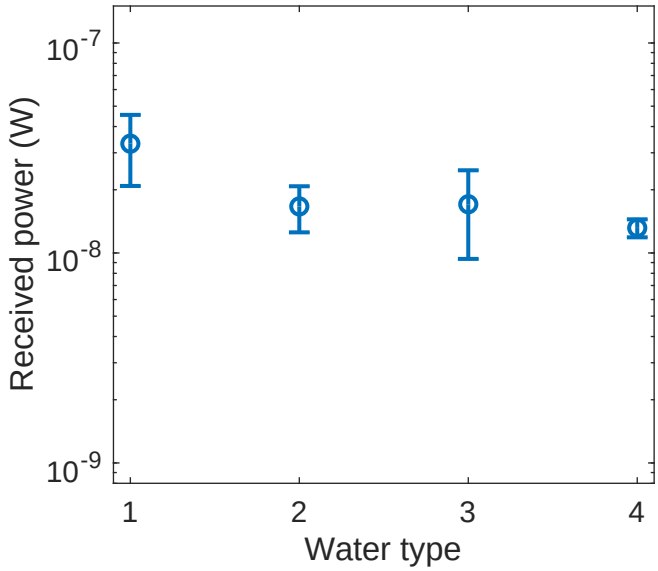


FIGURE 5.14: Final received power of the relay network with 5 robots using distributed algorithm in S1, S2, S3, and S4 underwater environment. The parameters of underwater environment are from (Johnson, Green, and Leeson, 2013).

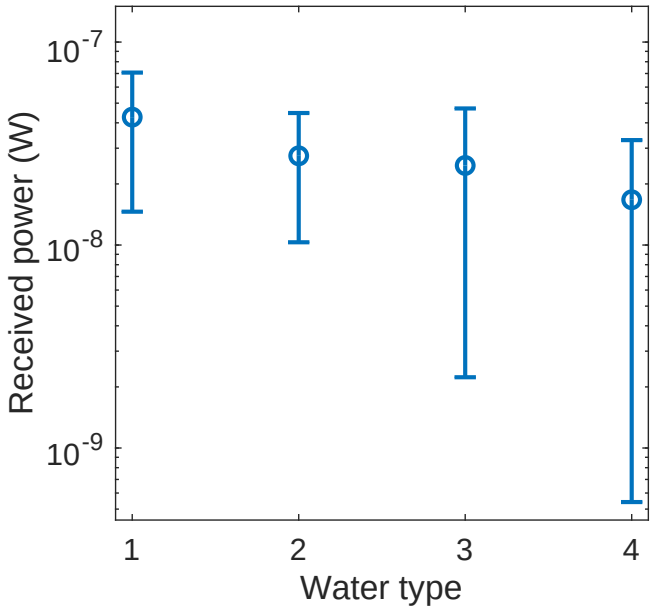


FIGURE 5.15: Final received power of the relay network with 5 robots using centralized algorithm in S1, S2, S3, and S4 underwater environment. The parameters of underwater environment are from (Johnson, Green, and Leeson, 2013).

environments are unknown, i.e., the centralized controller considers a static S1 environment. As shown in Fig. 5.15, the received power of all the relay links has a large variance in S2, S3, and S4 environments. Those links with a small received power will limit the overall data rate. Thus, without the channel and environmental information, the centralized algorithm cannot be adaptive to the dynamic underwater environment.

# 6 CONCLUSIONS AND FUTURE WORK

## 6.1 Conclusion

Applications such as climate change monitoring, target monitoring, and species migration monitoring can only benefit from the application of high-speed image processing and video streaming. This work considers a holistic approach to bridging these very important fields together. The study begins by considering the assembly and operation of an underwater robot which has the ability to travel up to $130m$ deep under water. The next section analyses the implementation of an underwater target detection model, followed by an optical testbed which has the potential to support high-speed video streaming. A distributed trajectory control algorithm is also proposed to organize and maintain the trajectory of robots during swarming.

## 6.2 Future Work

### 6.2.1 Further Exploration

One of the main challenges faced during the explorations was the absence of any plastics in all the underwater environments. Further exploration of different water bodies will be required in order to find more instances of plastics.

### 6.2.2 Object Detection and Image Enhancement

The nature of the dataset and the images collected during the exploration present a slightly more difficult learning process than many other object detection tasks. Using image enhancement techniques in MATLAB, the quality and nature of the images can be improved as well as the learning process. Again, metrics such as the

precision and recall of the model shall be investigated into greater detail. The aim of this process will be to improve the model's ability to more accurately detect plastics in the underwater environment.

### 6.2.3 Deployment of Model on Robot

Another step to be taken will be to deploy the object detection model on the companion computer embedded within the robot's architecture. This will serve the purpose of real-time object detection.

# BIBLIOGRAPHY

Akyildiz, Ian F, Pu Wang, and Zhi Sun (2015). "Realizing underwater communication through magnetic induction". In: *IEEE Communications Magazine* 53.11, pp. 42–48.

Arnon, Shlomi and Debbie Kedar (2009). "Non-line-of-sight underwater optical wireless communication network". In: *Journal of the Optical Society of America A* 26.3, pp. 530–539.

Bingham, Brian et al. (2010). "Robotic tools for deep water archaeology: Surveying an ancient shipwreck with an autonomous underwater vehicle". In: *Journal of Field Robotics* 27.6, pp. 702–717.

BlueRobotics (n.d.). *QGroundControl: Hardware Options*. URL: https://www.ardusub.com/images/hardware/Connection-Diagram-R1.png.

Bricaud, Annick et al. (1995). "Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: Analysis and parameterization". In: *Journal of Geophysical Research: Oceans* 100.C7, pp. 13321–13332.

Carver, Charles J. et al. (Feb. 2020a). "AmphiLight: Direct Air-Water Communication with Laser Light". In: *17th USENIX Symposium on Networked Systems Design and Implementation (NSDI 20)*. Santa Clara, CA: USENIX Association, pp. 373–388. ISBN: 978-1-939133-13-7. URL: https://www.usenix.org/conference/nsdi20/presentation/carver.

Carver, Charles J et al. (2020b). "Amphilight: Direct air-water communication with laser light". In: *17th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 20)*, pp. 373–388.

Chattopadhyay, Arpan et al. (2016). "Deploy-as-you-go wireless relay placement: An optimal sequential decision approach using the multi-relay channel model". In: *IEEE Transactions on Mobile Computing* 16.2, pp. 341–354.

Chen, Yifei et al. (2017). "26 m/5.5 Gbps air-water optical wireless communication based on an OFDM-modulated 520-nm laser diode". In: *Opt. Express* 25.13, pp. 14760–14765. DOI: 10.1364/OE.25.014760. URL: http://opg.optica.org/oe/abstract.cfm?URI=oe-25-13-14760.

Choudhari, Pratik (n.d.). *Understanding "convolution" operations in CNN*. URL: https://medium.com/analytics-vidhya/understanding-convolution-operations-in-cnn-1914045816d4.

Fulton, Michael et al. (2019). "Robotic detection of marine litter using deep visual detection models". In: *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 5752–5758.

Fulton, Michael S, Jungseok Hong, and Junaed Sattar (2020). *Trash-ICRA19: A Bounding Box Labeled Dataset of Underwater Trash*. URL: https://conservancy.umn.edu/handle/11299/214366.

Gad, Ahmed Fawzy (n.d.). *Faster R-CNN Explained for Object Detection Tasks*. URL: https://blog.paperspace.com/faster-r-cnn-explained-object-detection/.

Gama, Fernando, Joan Bruna, and Alejandro Ribeiro (2020). "Stability properties of graph neural networks". In: *IEEE Transactions on Signal Processing* 68, pp. 5680–5695.

Girshick, Ross (2015). "Fast r-cnn". In: *Proceedings of the IEEE international conference on computer vision*, pp. 1440–1448.

Girshick, Ross et al. (2014). "Rich feature hierarchies for accurate object detection and semantic segmentation". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 580–587.

Guo, Hongzhi and Clifford Boakye-Mensah (2022). "Distributed Trajectory Design for Underwater Multi-Robot Relay Networks". In: *2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, pp. 469–472.

Guo, Hongzhi, Zhi Sun, and Pu Wang (2017). "Multiple frequency band channel modeling and analysis for magnetic induction communication in practical underwater environments". In: *IEEE Transactions on Vehicular Technology* 66.8, pp. 6619–6632.

— (2021). "Joint Design of Communication, Wireless Energy Transfer, and Control for Swarm Autonomous Underwater Vehicles". In: *IEEE Transactions on Vehicular Technology* 70.2, pp. 1821–1835.

Haltrin, Vladimir I. (1999). "Chlorophyll-based model of seawater optical properties". In: *Appl. Opt.* 38.33, pp. 6826–6832. DOI: 10.1364/AO.38.006826.

Hill, Reginald J and Rod G Frehlich (1997). "Probability distribution of irradiance for the onset of strong scintillation". In: *Journal of the Optical Society of America A* 14.7, pp. 1530–1540.

Islam, Md Jahidul, Youya Xia, and Junaed Sattar (2020). "Fast underwater image enhancement for improved visual perception". In: *IEEE Robotics and Automation Letters* 5.2, pp. 3227–3234.

Jackson, Brian E et al. (2020). "Scalable cooperative transport of cable-suspended loads with UAVs using distributed trajectory optimization". In: *IEEE Robotics and Automation Letters* 5.2, pp. 3368–3374.

Johnson, Laura J, Roger J Green, and Mark S Leeson (2013). "Underwater optical wireless communications: depth dependent variations in attenuation". In: *Applied optics* 52.33, pp. 7867–7873.

Kameda, Takuhiko and Satsuki Matsumura (1998). "Chlorophyll biomass off Sanriku, northwestern Pacific, estimated by Ocean Color and Temperature Scanner (OCTS) and a vertical distribution model". In: *Journal of Oceanography* 54.5, pp. 509–516.

Kaushal, Hemani and Georges Kaddoum (2016). "Underwater Optical Wireless Communication". In: *IEEE Access* 4, pp. 1518–1547. DOI: 10.1109/ACCESS.2016.2552538.

Liu, Weihao, Zhengyuan Xu, and Liuqing Yang (2015). "SIMO detection schemes for underwater optical wireless communication under turbulence". In: *Photonics Research* 3.3, pp. 48–53.

Lu, Hai-Han et al. (2016a). "An 8 m/9.6 Gbps underwater wireless optical communication system". In: *IEEE Photonics Journal* 8.5, pp. 1–7.

Lu, Huimin et al. (2016b). "Underwater image descattering and quality assessment". In: *Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP)*. IEEE, pp. 1998–2002.

Mukti, Ishrat Zahan and Dipayan Biswas (2019). "Transfer learning based plant diseases detection using ResNet50". In: *2019 4th International Conference on Electrical Information and Communication Technology (EICT)*. IEEE, pp. 1–6.

Pytorch (n.d.). $FASTERRCNN_RESNET50_FPN$. URL: https://pytorch.org/vision/stable/generated/torchvision.models.detection.fasterrcnn_resnet50_fpn.html.

Ren, Shaoqing et al. (2015). "Faster r-cnn: Towards real-time object detection with region proposal networks". In: *Advances in neural information processing systems* 28.

Ruiz, Luana, Fernando Gama, and Alejandro Ribeiro (2021). "Graph neural networks: Architectures, stability, and transferability". In: *Proceedings of the IEEE* 109.5, pp. 660–682.

Saeed, Nasir, Tareq Y Al-Naffouri, and Mohamed-Slim Alouini (2018). "Outlier detection and optimal anchor placement for 3-D underwater optical wireless sensor network localization". In: *IEEE Transactions on Communications* 67.1, pp. 611–622.

Saeed, Nasir et al. (2019). "Underwater optical wireless communications, networking, and localization: A survey". In: *Ad Hoc Networks* 94, p. 101935.

Sun, Xiaobin et al. (2020). "A review on practical considerations and solutions in underwater wireless optical communication". In: *Journal of Lightwave Technology* 38.2, pp. 421–431.

Zavlanos, Michael M, Alejandro Ribeiro, and George J Pappas (2012). “Network integrity in mobile robotic networks”. In: *IEEE Transactions on Automatic Control* 58.1, pp. 3–18.

Zeng, Zhaoquan et al. (2016). “A survey of underwater optical wireless communications”. In: *IEEE communications surveys & tutorials* 19.1, pp. 204–238.