

VISION-BASED HUMAN DETECTION FOR MARITIME SEARCH AND RESCUE UNMANNED SURFACE VESSELS

Liam Allen¹, Patrick Miller², Griffin Hevesy¹, Jakub Gelvanovski²,
Vijay Jetton¹, Kush Patel³, Aaron Webb¹, Mats De Lausnay^{2,3},
Harold Park², Akshith Kantareddy⁴, Aditya Prashanth⁴

Faculty Mentor: Dr. Michael Otte¹

Research Librarian: Nevenka Zdravkovska⁵

¹Department of Aerospace Engineering

²Department of Electrical & Computer Engineering

³Department of Mechanical Engineering

⁴Department of Computer Science

⁵University of Maryland Libraries

May 10, 2026

Thesis submitted in partial fulfillment of the requirements of the Gemstone Honors Program,
University of Maryland, 2026

Abstract

Maritime search and rescue (SAR) operations rely heavily on manned crews, and the time required to organize and deploy personnel can significantly delay the location and rescue of individuals in distress. Unmanned Surface Vessels (USVs) are robots that operate autonomously or semi-autonomously on the surface of the water. They have been applied to an expanding range of scenarios, including mine countermeasures, water surveying, oceanographic research, and SAR. However, limitations in water-level vision systems, autonomous navigation, and reliable data transfer complicate the effectiveness of a SAR USV. For example, the lack of representative datasets that feature people in water further impedes the development and training of SAR-applicable machine vision. To fill a gap in the research, we created applicable datasets, developed, tested, and trained a vision system, and integrated that vision system into a fully autonomous USV.

Acknowledgements

First and foremost, we would like to thank our mentor, Dr. Michael Otte, for his support, guidance, and feedback throughout the Gemstone program. We thank our panel discussants, Mr. Timothy O'Connor, Dr. Romel Gomez, Mr. Joseph Schwartz, and Dr. Loy McGuire. We thank Mr. Ivan Penskiy for providing access to the UMD Maryland Robotics Center during the initial construction stages of our project, and Dr. David Akin for providing access to the Neutral Buoyancy Research Facility during the integration and testing stages of our project. We thank Nevenka Zdravkovska for all her help as our team librarian and for the generous Library Awards grant that helped fund our project. Other individuals and organizations we would like to thank who made the research process go smoothly include: UMD Facilities, UMD RecWell, BlackSea Technologies, the Marine Technology Society, the IEEE Oceanic Engineering Society, NAVSEA PMS 406, Matthew Zavalij, and all of our LaunchUMD Donors. Lastly, we would like to thank Dr. David Lovell, Dr. Allison Lansverk, Leslie Lizama, Brianna Lucas, Jannelle Dang, and everybody involved in running the Gemstone Program for making our research possible.

Table of Contents

Acknowledgements	i
1 Introduction	1
1.1 Problem Statement	1
1.2 Novelty and Contributions	2
1.3 Outline of Thesis	3
1.4 Background Information	3
1.4.1 Maritime Search and Rescue (SAR)	3
1.4.2 Search Patterns	5
2 Related Work	6
2.1 Unmanned Surface Vessels	7
2.2 Obstacle Avoidance	9
2.3 Vision and Detection	11
3 Methodology	13
3.1 USV Construction	13
3.1.1 Structural	13
3.1.2 Electrical	15
3.1.3 Software	19
3.2 Testing	20
3.2.1 USV (without cameras) Testing	20
3.2.2 Vision Model Data Collection	22
3.2.3 Vision Model Training and Testing	26
3.2.4 USV-Vision System Integration	29

3.3	Vision System-USV Integration Testing	30
4	Results	32
4.1	Vision System	32
4.1.1	Comparison of Tested Models	32
4.1.2	Performance of Vision Models	34
4.2	USV Navigation	40
4.3	USV-Vision System Integration	41
5	Discussion	44
5.1	Analysis of Results	44
5.1.1	Vision System	44
5.1.2	USV-Vision System Integration	46
5.2	Future Work	49
6	Conclusion	50

Abbreviations

APF - Artificial Potential Field

AUV - Autonomous Underwater Vehicle

COLREGs - Convention on the International Regulations for Preventing Collisions at Sea

DRL - Deep Reinforcement Learning

DWA - Dynamic Window Approach

GPS - Global Positioning System

IAMSAR - International Aeronautical and Maritime Search and Rescue

ICAO - International Civil Aviation Organization

IMO - International Maritime Organization

IP65 - Ingress Protection, 6-Rating Solid Protection, 5-Rating Liquid Protection

LiDAR - Light Detection and Ranging

LDPE - Low-Density Polyethylene

LLA - Lives Lost After Notification

LLB - Lives Before Notification

LS - Lives Saved

LUF - Lives Unaccounted For

MOB - Man Overboard

MRCC - Maritime Rescue Co-ordination Center

OA - Obstacle Avoidance

RAM - Random Access Memory

RC - Remote Control

RCC - Rescue Co-ordination Center

RGB-D - Red, Green, Blue and Depth

SAR - Search and Rescue

SRR - Search and Rescue Region

SSD - Solid State Drive

SSH - Secure Shell

SSL - Space Systems Laboratory

TCP - Transmission Control Protocol

UAV - Unmanned Aerial Vehicle

UMD - University of Maryland

USV - Unmanned Surface Vessel

UUV - Unmanned Underwater Vehicle

VPN - Virtual Private Network

Section 1: Introduction

1.1 Problem Statement

In 2015, maritime search and rescue (SAR) operations in the US saved nearly 17,000 lives with approximately 34,000 responses to distress signals [1]. Unfortunately, even with improvements in SAR technology, more than 4000 people die from drowning every year in the United States alone [2]. According to Mario Vittone, as of 2019, an SAR operation takes 4.5 hours following a distress signal, on average [3]. Unfortunately, the chances of rescuing survivors decrease with time [4], and weather conditions, limited daylight, or currents can drastically increase the time until rescue [3]. At the same time, harmful weather conditions or low water temperatures can decrease survival time [3]. This poses a clear challenge because the SAR response time can often exceed the time available to save a drowning victim. Therefore, the goal of our research is to explore autonomous maritime SAR methods to reduce response times.

Detecting individuals in water poses a unique challenge for autonomous vision systems because victims are often obstructed from view (via water, glare, or poor lighting). Publicly available image datasets of people treading or floating in water are hard to find, placing limitations on developing vision models specifically trained around such images. Fortunately, thermal vision systems can provide more data for a decision model. For example, the heat signature of a person’s head is distinct within a cold environment (e.g., a body of water). However, thermal cameras come with their own issues, including surface glare and temperature distortion below the waterline. These facts pose difficult challenges for SAR operations and for designing a reliable vision system. Our research aims to address these issues.

1.2 Novelty and Contributions

In this project, we developed a proof of concept for an unmanned surface vessel (USV) to detect people in the water using machine vision. The main novelty of our project is our vision system, which combines visual and thermal imaging to locate humans in marine settings. Our discussions with stakeholders involved in maritime systems and autonomy reinforced this focus. These stakeholders included Navy engineers working in autonomous systems and other industry professionals in the USV space. In this dissertation, we compare the effectiveness of visual and thermal imaging techniques, both individually and in conjunction, to detect a person in marine settings.

We hypothesized that combining visual image data with infrared image data could improve searches, as one data source may pick up a person when the other cannot, helping to inform the movement of the USV towards that position of interest. Using both vision methods and comparing their data will allow our vision system to make more informed decisions. Our project seeks to explore the computational capability to support both visual and thermal computer vision systems.

The decision to use a USV instead of any other autonomous platform sought to provide a slight niche for this project. This contrasts with the existing literature, which is largely focused on Unmanned Aerial Vehicles (UAVs) and some Unmanned Underwater Vehicles (UUVs).

The main contributions of this project are as follows:

- Comparison of strictly visual, strictly infrared, and combined systems and their capability of detecting a human in water
- Design and creation of a prototype USV capable of supporting these vision systems
- Combined testing of the vision system and the USV
- A review of the existing literature on maritime SAR, USVs, search patterns, obstacle

avoidance, and vision and detection

1.3 Outline of Thesis

The thesis is broken down into 5 sections following the introduction. Section 2 is for related work, where we will go through existing literature related to our project. Section 3 is where we outline the steps we took to test our vision system and USV. Section 4 compiles our results from the tests in the previous section. Section 5 analyzes those results to highlight the takeaways from our work. Finally, Section 6 wraps up and summarizes the thesis. The appendix parses the budget and spotlights the ethics and impact considerations of the project.

1.4 Background Information

The following subsection contains different background information on maritime search and rescue (SAR) and search patterns. This section helps to contextualize our research and current related work.

1.4.1 Maritime Search and Rescue (SAR)

SAR consists of two coordinated operations: *search*, defined as the use of personnel and resources to “locate persons in distress,” and *rescue*, defined as the retrieval of those persons, the provision of medical care and transport to safety [5]. Maritime SAR involves additional successive phases, including: “patrolling, travel to the event, directly collecting people at sea, collecting people by other means (floating devices or smaller boats), and safely returning people to land” [6]. Since 1999, SAR operations worldwide have been conducted within the framework of the International Aeronautical and Maritime Search and Rescue (IAMSAR) Manual, published by the International Maritime Organization (IMO) and the International Civil Aviation Organization (ICAO) [7]. The manual is divided into three volumes: the first

covers the organizational and management structure of SAR operations; the second provides assistance to those planning and coordinating SAR operations; the third acts as a guide consisting of procedures to be carried out on board SAR vessels [7].

A general framework for a maritime SAR mission is a five-stage process. The first stage, often referred to as the awareness stage, involves the reception of a rescue call or an alert [8]. This usually involves information such as the Last Known Position (LKP) of the victim(s), the number of victims needing rescue, the type of incident, and what kind of help is required. The alert and all accompanying information are then referred to the relevant Maritime Rescue Coordination Center (MRCC) or Rescue Sub-Center (RSC) [9]. The next step, the initial action stage, is where evaluation and decision-making based on the rescue call occur [8]. Personnel collect additional information concerning the relative location of the incident regarding the weather, tidal conditions, current trajectories, etc. Resources necessary for the operation are identified and prepared. The next stage is the planning stage, which involves formulating a response through risk assessment, identification of optimal search patterns and entry points, and request for any additional necessary resources. The operations/execution stage involves the physical execution of the SAR. The resources and personnel of the responding MRCC(s) are used to search the SAR Region (SRR), assist and retrieve the victims in a state of emergency, and transport the victims back to safety [8]. The conclusion stage is where operations end after the “determination that the subjects of the SAR incident are either not in distress or have been rescued, or if it is assessed that further efforts will not increase the likelihood of locating any additional survivors” [8].

The US Coast Guard Office of Search and Rescue (CG-SAR) oversees maritime SAR and addresses each operation with four primary objectives: minimizing loss of life, injury and property damage; minimizing risk to rescuers; optimizing the use of available resources; and maintaining a global leadership role in maritime SAR [10]. The performance metrics used by CG-SAR are measured following the benchmark goals within the Marine Information for Safety and Law Enforcement (MISLE) database. Variables such as *LS* represent “lives

saved,” *LLB* and *LLA* indicate “lives lost before/after notification”, respectively, and *LUF* means “lives unaccounted for” as input into MISLE [10], yielding a metric defined by

$$\frac{LS}{LS + (LLB + LLA + LUF)} \quad (1.1)$$

A secondary benchmark is used analogously, but with variables regarding property loss, instead of loss of life. These metrics are not universal and may be different depending on the complexity of a maritime SAR operation and the composition of the nations involved.

The benchmarks used to measure the success of SAR missions do not yield a perfect rate, highlighting various challenges and limitations that prevent personnel from achieving the desired outcomes. For example, a maritime environment is highly dynamic; changing weather conditions, currents, and tides [8] can drastically move survivors and make it significantly more difficult for personnel to locate them. The structure of maritime SAR protocol is also very hierarchical and rigid, which can be limiting in regards to adaptability: something vital in dynamic environments such as the ocean. Additionally, SAR operations are also often based on vague or limited information (e.g., LKP), making it extremely difficult to estimate the initial location of an incident [8] and where survivors have been displaced due to winds or tides. There are also huge limitations when it comes to resources, whether it is time, money, personnel, training, or assets [8]. Together, these limitations highlight the need for more adaptable and scalable solutions, which motivates our research.

1.4.2 Search Patterns

The United States Coast Guard (USCG) employs a variety of search patterns to look for people in distress. Different patterns are used according to the size of the search area and how much is known about the location of the search target [11]. These patterns are also specifically developed to handle the potential for leeway (movement due to wind) and drift (movement due to current).

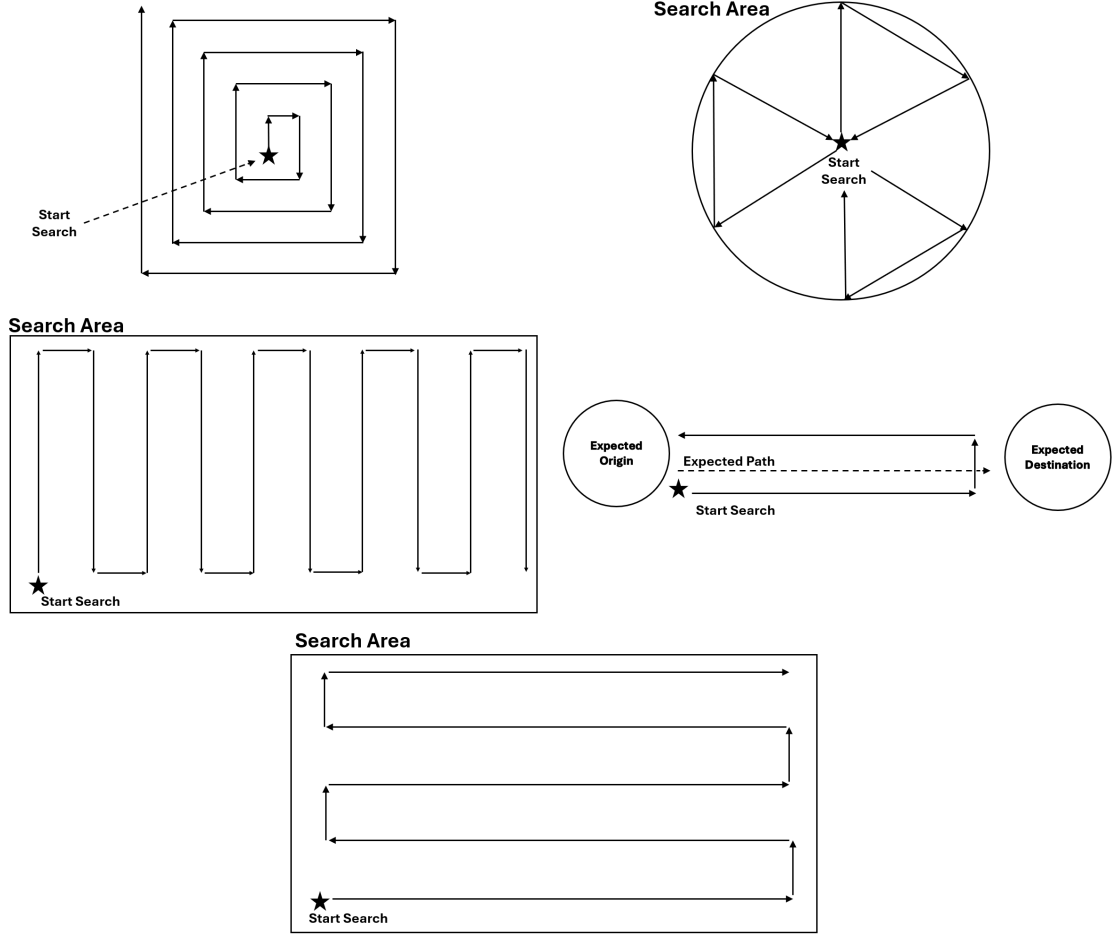


Figure 1.1: Above are images of the various search patterns currently in use by the USCG. Depicted are *Expanding Square*, *Sector*, *Creeping Line*, *Trackline*, *Single Unit*, *Return*, and *Parallel, Single Unit*. [11]

Each search pattern depicted in Figure 1.1 has its own use case based on the search conditions. *Expanding Square* and *Sector* are most often used for a smaller search area while *Creeping Line* and *Parallel, Single Unit* are used for larger search areas [11].

Section 2: Related Work

In this section, we discuss research on USVs, obstacle avoidance, and autonomous vision-detection systems.

2.1 Unmanned Surface Vessels

Maritime SAR is based on crewed vessels and aircraft. These limitations have driven interest in unmanned surface vessels (USVs) as a platform for autonomous search, detection, and response. USVs, also called autonomous surface vehicles (ASVs), operate on the water’s surface without a crew onboard and can be controlled remotely or programmed “to conduct pre-planned missions and make decisions with minimal real-time human input” [12]. Although preliminary forms of the USV have been used as early as World War II, modern iterations have significantly increased in prominence since the 1990s [13] with military, commercial, and scientific communities recognizing their potential for a variety of tasks. Such tasks include oceanographic surveying [14], water quality testing [15], environmental monitoring, and mine countermeasures [16], [17], [18]. Their relevance to maritime SAR is significant: operations are difficult to organize, costly, require significant personnel, and can put more people at risk [19]. USVs mitigate many of these concerns by reducing manpower requirements, the number of lives at risk, and the time and resources required to organize a rescue operation. In addition, USVs have low maintenance costs, compact size [16], and several advantages over similar counterparts. These advantages include greater resistance to hazardous weather conditions compared to unmanned aerial vehicles (UAVs) [19] and the lack of a required submersion system, needed by autonomous underwater vehicles (AUVs).

Previous groups have studied the application of USVs to SAR missions. An early system, EMILY, was a remote-controlled robotic rescue buoy developed by the Center for Robot-Assisted SAR (CRASAR) and Hydronalix in 2010 [19]. Although not fully autonomous, EMILY demonstrated the potential of small surface vehicles in SAR scenarios, as both the Hellenic Coast Guard and the Los Angeles County Fire Department successfully used it to recover drowning victims [20]. However, EMILY relied heavily on human operation and lacked an onboard vision system, requiring UAV assistance for effective victim detection. These limitations motivated the development of SmartEMILY, which incorporated the Global Po-

sitioning System (GPS) and a flight controller to enable partially autonomous navigation. Despite these improvements, SmartEMILY continued to face challenges in the design of the operator interface, leading to unclear communication and limited environmental awareness [20].

Although EMILY and SmartEMILY are more widely known and studied, additional SAR-focused USVs have since been constructed and researched. The SWIFT USV, presented at the OCEANS conference in 2014, was specifically developed for surf and intertidal environments and demonstrated the ability to transport flotation devices, including a four-person liferaft [21]. Its waterjet propulsion system provided improved speed and maneuverability, while its compact size allowed wider production but indicated limited performance against strong waves or wind [21]. Another prototype, presented at the 5th International Undergraduate Research Conference (IUGRC), integrated solar and wind power to extend the operational duration and successfully detected a simulated drowning victim, but required prior coordinate input for accurate navigation [22]. The research presented at the 2023 International Conference on Advanced Robotics showcased a USV navigation system capable of operating in calm water, light waves, and even large waves, though only at very low speeds for the latter (<1 knot) [23]. A recent simulation study of a coupled USV-UAV demonstrated that the USV alone was able to detect 60% of drowning victims and was much better at predicting the change in the location of the victim than the UAV [24]. However, the USV remained significantly worse than the UAV in locating victims in the first place, indicating limitations in its vision system capabilities [24].

Despite significant progress in the application of USVs to SAR, several technical constraints remain in the way of implementation. With ideal conditions and resources, USVs could theoretically satisfy the requirements of SAR without human operation [19]. However, most existing systems are teleoperated or only partially autonomous due to limitations in machine vision, autonomous navigation, and reliable data transfer [16]. USVs have very limited and sensitive vision capabilities, which are worsened by poor weather conditions, glare,

lighting, and obstructions. Even with aircraft or UAV assistance, USV vision capabilities suffer from discrepancies with displacement, angle, and lighting in relation to the potential victim [25]. Challenges also arise in navigation, as most USV GPS-based systems lack precise victim location information. In addition, poor vision, obstacles, and harsh weather further complicate path-planning [25]. Slow or unreliable data transmission can reduce autonomy, and unclear or poorly designed user interfaces (as seen before with EMILY) introduce additional setbacks in USV control and monitoring [20]. The continued need for teleoperation reduces the number of human responders who could be helping in other urgent matters [25], undermining one of the main benefits of a SAR USV.

These shortcomings indicate a need for USVs capable of reliable onboard human detection and autonomous execution of search patterns, without dependence on external operators. A recent analysis [26] conducted for the 2020 Zhuhai Wanshan International Intelligent Vessel Competition highlights that effective USV autonomy requires appropriate system architecture, open-source software integration, and, most importantly, reliable onboard vision and detection systems. These identified requirements are often unfulfilled as indicated above, particularly with regard to reliable vision and navigation system. Collectively, this pattern motivates the current development of an integrated vision system and autonomous navigation framework for maritime SAR applications.

2.2 Obstacle Avoidance

Deep Reinforcement Learning (DRL) provides a solution to obstacle avoidance and difficult maritime navigation. When utilizing DRL, USV control algorithms are extensively trained in simulation to maximize a cumulative reward function. This system gives rewards based on proximity to the goal and assigns penalties based on adherence to regulation or proximity to obstacles. A terminal state, or large penalty, is assigned to collisions. DRL methods have proven effective for COLREGs (The International Regulations for Preventing Collisions at Sea)-compliant collision avoidance, with Deep Q-Network (DQN) algorithms successfully

training optimal parameters for real-time decision making in multi-vessel encounters [27]. Through training the network, the USV learns to adapt to all of the scenarios accounted for in simulation to effectively handle unforeseen obstacles. This method’s accuracy can be further improved using vision detection; vision systems using the YOLOv5 vision model enabled type-specific avoidance strategies [28]. For example, identifying a moving body as a ship has different requirements to that of a log, meaning vision detection allows for a more targeted approach [28].

An immediate and less computational approach to obstacle avoidance uses an Artificial Potential Field (APF) that moves purely on the vector sum of an attractive force to the goal and a strong repulsive force from detected obstacles. However, traditional APF methods suffer from inherent disadvantages, including local minima traps and unreachable targets in congested environments [29]. This occurs because strong opposing forces might freeze the USV, requiring augmentations such as additional lateral force methods that detect potential local minima in advance and guide the USV to escape by applying controlled angular acceleration [29], or simple DRL policies to account for these traps where DQN training can optimize the weight coefficients of the APF objective function [27]. The improved APF approaches incorporating waypoint generation algorithms with radius of gyration restrictions ensure planned paths satisfy the USV’s minimum turning radius constraints while maintaining smooth trajectories [29]. Relying solely on the obstacle avoidance (OA) algorithm often leads to low path-planning efficiency, unreasonable routes, and easy trapping [27]. However, the combined method integrating improved A-star global path planning with DRL-enhanced Dynamic Window Approach (DWA) algorithms shows significant improvement in both single USV and multi-USV systems, achieving COLREGs-compliant collision avoidance with shorter paths and reduced sailing times compared to traditional methods [27].

In the case of maritime SAR with multiple USVs, the team hopes to utilize existing obstacle avoidance strategies, using a hybrid static and dynamic pathfinding approach to account for the case of obstacles as a safety measure, where a lightweight solution with

sufficient accuracy suffices for the sparse environment of our applications.

2.3 Vision and Detection

Vision systems, in an SAR context, can enable autonomous platforms to detect the presence of people and distinguish them from potential obstacles. A standard land-based SAR vision system is an array of RGB-D (Red, Green, Blue, and Depth) and thermal cameras, as shown by previous work at the University of Moratuwa [30]. The research describes the development of a deep neural network that processes real-time imaging, identifies potential targets for a remote operator, and supports adaptive navigation to ensure efficient operation. However, such capabilities require advanced imaging systems and significant computational power to achieve reliable human detection [31], and glare alongside visual occlusion in the water creates more difficulties.

A UAV’s higher vantage point provides a better viewing angle for a mounted vision system. However, the background of water makes the detection of humans less accurate. Due to this, issues have arisen with Man Overboard (MOB) systems designed to deploy UAVs off ships to locate potential victims. These AI-based systems require training data to create a heuristic of what a person in the water looks like. Research published in Future Generation Computer Systems found that most existing datasets are created under calm ocean conditions, which poses a problem for training human detection under rough conditions [32]. Such conditions, such as choppy seas or strong currents, introduce high levels of disturbance to an image, making reliable detection difficult. This group, along with Rodin and Johansen [33], evaluated the detectability of objects in water.

Algorithm development is one method to improve autonomous vehicles, but there are other approaches to improve detection accuracy. One potential solution [34] is to expand the machine vision database with more images of humans in water. This allows machine learning vision systems to have more cases to compare a scenario against, making results more accurate. Other researchers have used computer simulations to run MOB programs

to improve the machine-learning aspects of marine vision systems [35]. In our research, we adapt this idea to create a simulation that accurately mimics what a USV’s camera would see on the water. This system allows us to simulate a wide range of ocean conditions to prepare our USVs effectively.

Vision systems designed for human detection typically have two components in their camera array: RGB-D and thermal cameras. Depending on the requirements of the system (field of view, power limitations, etc.), there can be either one camera or multiple cameras integrated together. Although high-quality cameras tend to be expensive, low-cost alternatives exist, as explored by Heinrich Ruser [36]. Guided by this research, our limited budget, and our goal of building an affordable and scalable USV, we opted for lower-priced but reliable components for our vision system. Due to the marine setting, our vision system must be waterproof. Many camera arrays used for AUVs and USVs are viable for waterproofing. Examples of underwater use of thermal imaging cameras can be seen in research on deep-sea thermal hotspots [37] and even infrared detection systems used for fishing [38]. However, even with adequate camera systems, a USV-based platform requires a significant influx of data since what the camera sees at the water level would look different from what a UAV sees. The limited vision due to the low vantage point makes detection difficult. Having more than one USV in the water at a time would allow an integrated system of shared data to get a better idea of a scenario in rough conditions.

USV vision systems intended for SAR applications must be able to correctly distinguish a person from their surroundings. A primary concern is minimizing the occurrence of false positives and false negatives. This is something we would need to avoid as much as possible because mislabeling could lead to the dismissal or misplacement of life-saving resources. By contrast, the vision system mislabeling an animal or flotsam as a human will waste valuable time that could be used to locate the actual person. The core principle of avoiding false positives and negatives is important in our research and heavily influences our methodology.

Section 3: Methodology

Our research covers our vision system, USV, and the integration of the two. The following section is broken up to isolate these project components while highlighting their construction and testing procedures.

3.1 USV Construction

The development of our USV was divided into three main subsystems: structural, electrical, and software. The structural components make up the hull itself and any brackets or housings that we fabricated. The electrical subsystem encapsulates any electrically powered components and the wires that connect them. The software is characterized by programs that run the USV through the electrical components.

3.1.1 Structural

The hull is a high-capacity catamaran with dimensions of 47" x 36.5" x 18". The majority of the hardware was purchased as spare parts from Blue Robotics' BlueBoat[39]. Our goal was to create a BlueBoat with a computer system that is more suited for our uses.

The two hulls of the catamaran are made from Low-Density Polyethylene (LDPE), and each carries their own payload in water-tight compartments. This cargo space is accessible through a hatch in the top of each hull. Each hull has a mount on the stern for an M200 motor. The two hulls are connected using a series of anodized aluminum crossbeams, and a through tube is attached to the interior of the two hulls via the back beam. The existing cargo space was insufficient to house many of our larger electronic components. To address this, we bolted a $\frac{1}{4}$ " plywood board across the aluminum crossbeams to create a stable platform for more cargo space. We applied 3 sheets of fiberglass to the plywood to strengthen it and

minimize any adverse effects that could be caused by contact with the water. This platform is home to a waterproof Pelican case and our cameras in their housings.

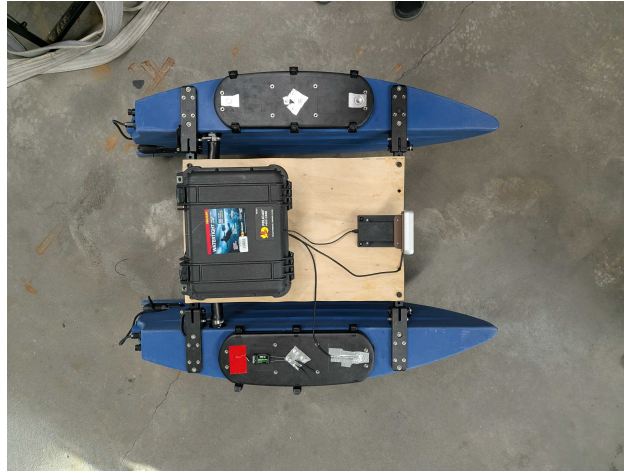


Figure 3.1: This is the top view of the USV. The vision system is towards the bow (right), and the pelican case is towards the stern (left)

The Pelican case is a 1400 Protector Case and is mounted with industrial Velcro on the stern side of the plywood platform. The case houses our computer and the power supply for the computer. A hole was drilled on the stern side of the case to feed wires to the computer.

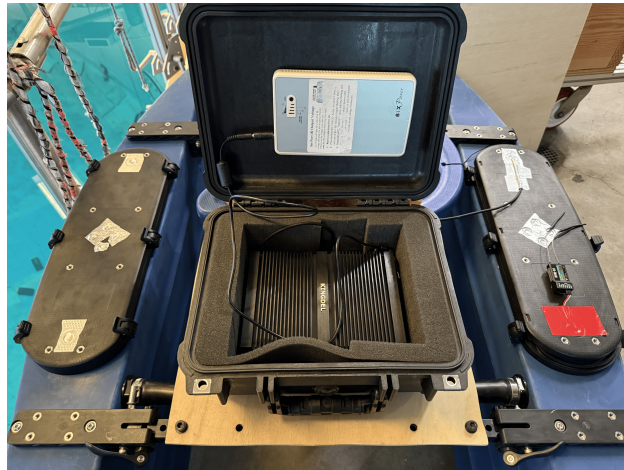


Figure 3.2: This image shows the internals of the waterproof Pelican case. The KINGDEL computer is in the bottom of the case while the power supply is attached to the inside of the lid.

On the section of the plywood platform closest to the bow, we bolted our vision system. At the base of our physical vision system is a 3D-printed waterproof housing for the thermal camera, shown below in Figure 3.3. On top of the housing, we mounted our visual camera. This camera is IP65 rated, so it did not require a housing of its own to protect it from splashes.

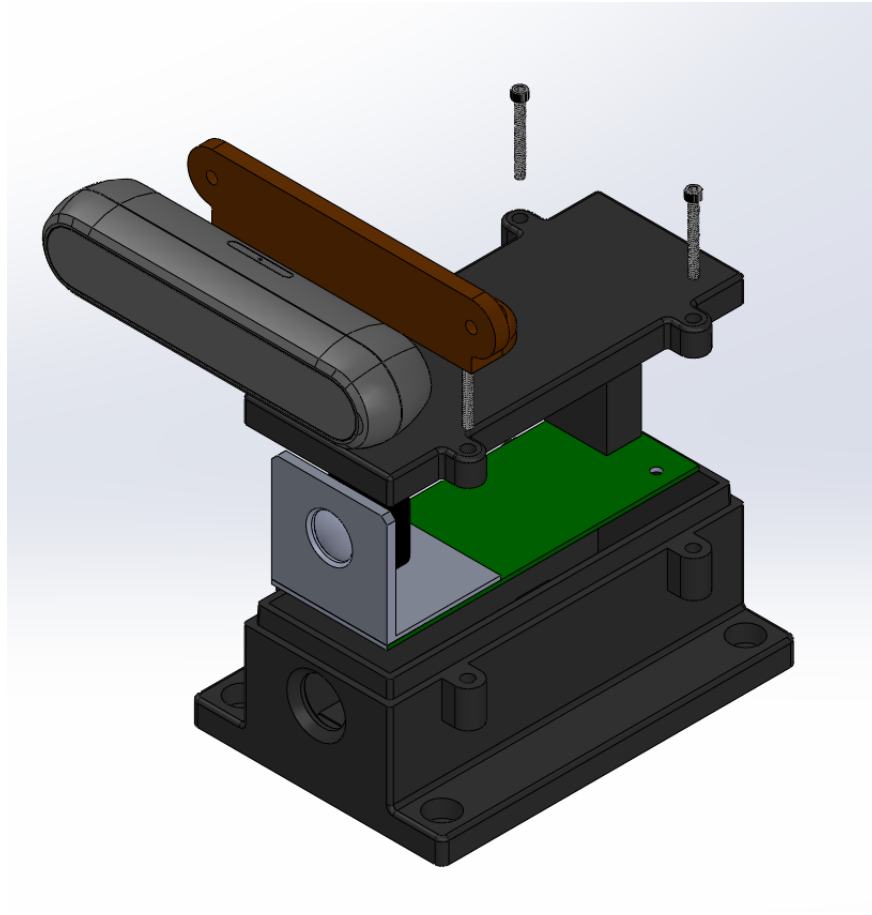


Figure 3.3: This screenshot shows an exploded view of our vision system, corresponding housing and mounts. The visual camera mount is shown in brown and the thermal camera housing is shown in black

3.1.2 Electrical

The hull and its accessories are designed to safely house the electronics onboard. The storage of these electronics can be split into three distinct sections: the Pelican case, camera setup, and port and starboard hulls. The corresponding electronic element within each section can

be seen in Table 3.1.

Electronic Housing	Pelican Case	Camera Setup	Port Hull	Starboard Hull
Contents	KINGDEL NC3000, Bix Power BP90	Intel RealSense Depth Camera D456, SeekThermal SD314SPX Drone Core Starter Kit	BlueRobotics Basic ESC 500	ArduPilot, GPS, RC Control Unit, 14.8V lithium-ion battery, M200 motors

Table 3.1: This table shows all four places that electronics can be held and a list of what electronics are in there.

3.1.2.1 Pelican Case

The Pelican case was used to house and protect the central computing system and its power source. The computer was a KINGDEL NC3000, an x86-based passively-cooled computer running Ubuntu 24.04. The computer has 32GB of RAM and 512GB of SSD storage. The BlueBoat uses an onboard Raspberry Pi 4 with the Blue Robotics BlueOS operating system, which was not easily adaptable to our needs. The selection of an x86 computer allowed maximum compatibility with the software necessary to connect the cameras with the navigational systems.

3.1.2.2 Camera Setup

Our thermal camera is stored in a waterproof housing on the front of the boat, with the visual camera mounted on top of it. We considered using Light Detection and Ranging (LiDAR) vision but decided against it due to our limited funds. Additional reasons for choosing our camera setup can be explained by this excerpt from our conference paper: [40]

“Our vision system combines standard and thermal vision systems to hopefully create better outcomes in detecting people. By combining both types of cameras, the system can

detect a potential person in distress from a distance or in poor conditions using the thermal camera and then confirm the detection with the standard camera. The machine vision camera that we selected is the Intel RealSense Depth Camera D456. We ultimately selected the camera because it was inexpensive and offered an $85^{\circ} \times 57^{\circ}$ field of view, an RGB sensor, and Python support. There were many alternatives, including cameras with built-in machine vision capabilities, image stabilization, and specifically designed for marine use. However, our goal is to develop a cost effective vision system for a USV, and those cameras were much more expensive than the selected camera. The USV vision system uses the SeekThermal SD314SPX Drone Core Starter Kit as its thermal camera. The primary selection criteria were performance, affordability, and modularity, making a camera component more feasible than a complete camera system like those offered by other companies. The SeekThermal kit proved ideal due to its affordability, lightweight design (which enabled a custom external casing for seamless USV integration) and its relatively straightforward software integration with the overall vision system.”

3.1.2.3 Hulls

Inside the port hull, we keep the ESC to filter the power to the M200 motor (shown in Figure 3.4). Both hulls have an M200 motor mounted on the stern for propulsion, which is the sole propulsion source for our USV.

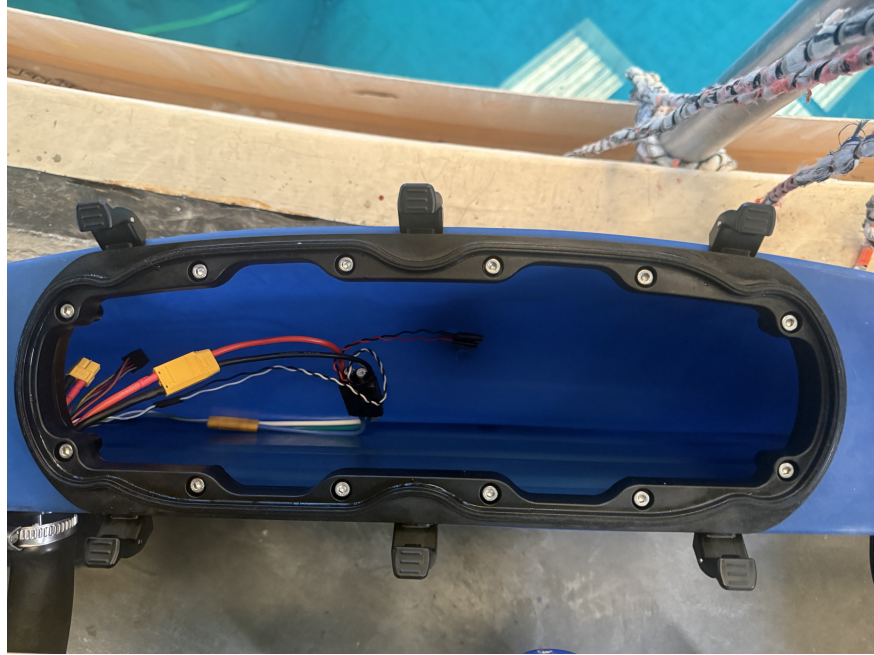


Figure 3.4: This image shows the internals of the port hull of the USV.

In the starboard hull, we keep the majority of our automation electronics. On a platform with standoffs from the hatch, we keep our ArduPilot and GPS/Compass (shown in Figure 3.5).

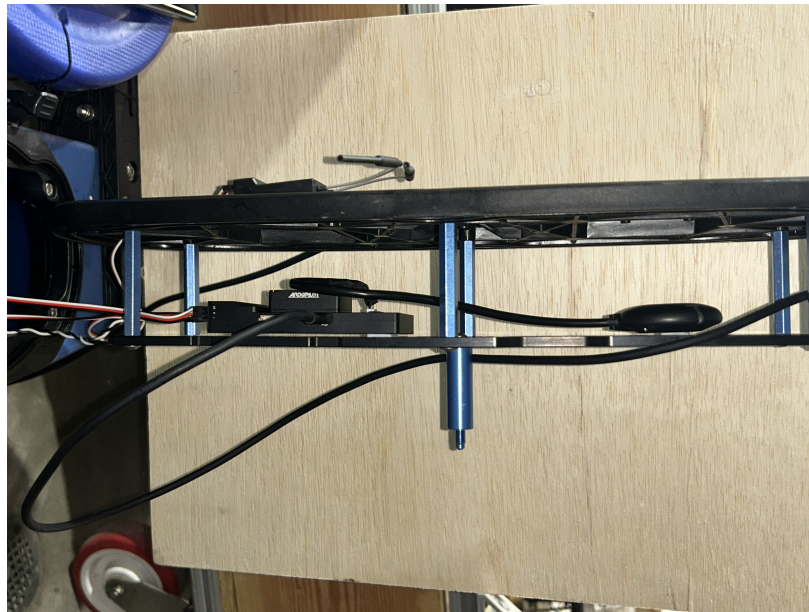


Figure 3.5: This image shows the internals of the standoffed board. This board contains the ardupilot (left), the GPS/compass (right). The RC unit is also visible on top of the hatch.

The ArduPilot manages the control of the entire USV by taking in commands from the computer and converting them to navigational control. The GPS sits towards the bow side of the platform, and provides accurate heading and location information. On top of the hatch, we keep our FS-I6B RC control unit, which connects to our Flysky FS-I6X remote controller, which can take over boat control if needed. A 14.8V lithium-ion battery is plugged in at the bottom of the starboard hull and provides power to both M200 motors. All of this can be seen below in Figure 3.6.

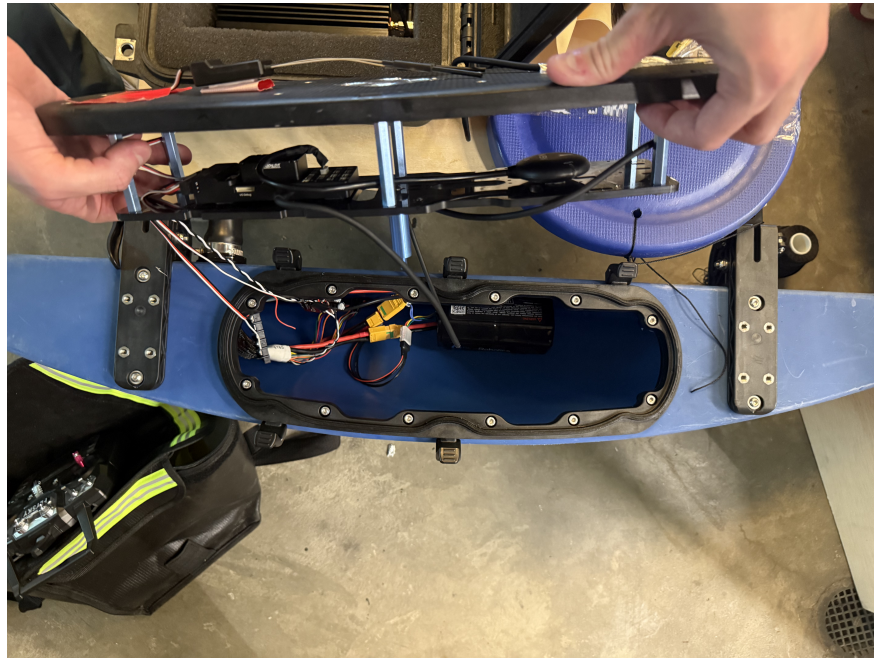


Figure 3.6: This image shows the internals of the starboard hull. The battery can be seen housed in the bottom of the hull.

3.1.3 Software

We relied on Tailscale to establish a virtual private network (VPN) connection from a laptop to the onboard computer, as University of Maryland (UMD) WiFi would not allow us to connect directly through secure shell (SSH). We used a transmission control protocol (TCP) configuration to communicate from our laptop to the boat. Once the connection was established, the ArduPilot Mission Planner software was used to control the settings on the

boat. This software provided several configurations for running the boat, including manual control for using the remote control, auto for running pre-defined patterns, and loiter to stay in place. This software allowed us to capture the path of the boat during runs using the onboard GPS, providing us with greater data on the movement of the boat and its ability to perform search patterns.

3.2 Testing

We used multiple pools and bodies of water as testing locations with varying dimensions depending on availability. This paper covers five of those locations, which we will refer to as Testing Areas A-E, which are described in Table 3.2.

Testing Area	A	B	C	D	E
Dimensions	50 ft diameter	75 ft x 60 ft	500 ft x 500 ft	150 ft diameter	500 ft x 1000 ft
What was tested?	Visual Camera	Visual and Thermal Cameras	USV	USV & Vision System	USV & Vision System

Table 3.2: This table shows a list of our 5 testing areas with their approximate dimensions

We tested all individual components of our overall system, starting with cameras and the USV (without cameras) separately. For our final testing, we integrated the two systems together. This testing system allowed us to work out any issues on a smaller scale before bringing them to the full testing scale.

3.2.1 USV (without cameras) Testing

To initially test the USV, we used a large indoor pool in the Neutral Buoyancy Research Facility (Testing Area A) at UMD. The first tests ensured that the motors and batteries onboard the boat were functional. To accomplish this, the USV was placed in the pool with the vision and detection system disconnected. Movement was remotely controlled using the

Flysky FS-I6X remote controller. This test also allowed us to confirm the integrity of the boat, which required small repairs due to missing parts. In the first build iteration, the boat was physically limited in maneuverability, including being unable to move backward. This was resolved by limiting the power to the motors to allow for finer movements and setting the configuration in ArduPilot to allow for driving the motors in reverse.

Next, we started with the boat in the middle of Testing Area A and had it spin about its central axis, checking the accuracy of the GPS tracker. This was limited due to the spotty nature of the GPS tracker indoors. The ArduPilot Mission Planner software would not allow us to complete search patterns without the accurate GPS data available.

To run a more intensive GPS test, we took the boat to Testing Area C, Lake Minsi. This allowed us to test the GPS without signal interference and see how accurately the boat could follow positional commands. We marked out a 500' x 500' square and had the USV run an automated search pattern using ArduPilot. The projected search pattern is shown in Figure 3.6. During this test, the USV experienced heavy winds but was still able to stick to its search pattern.

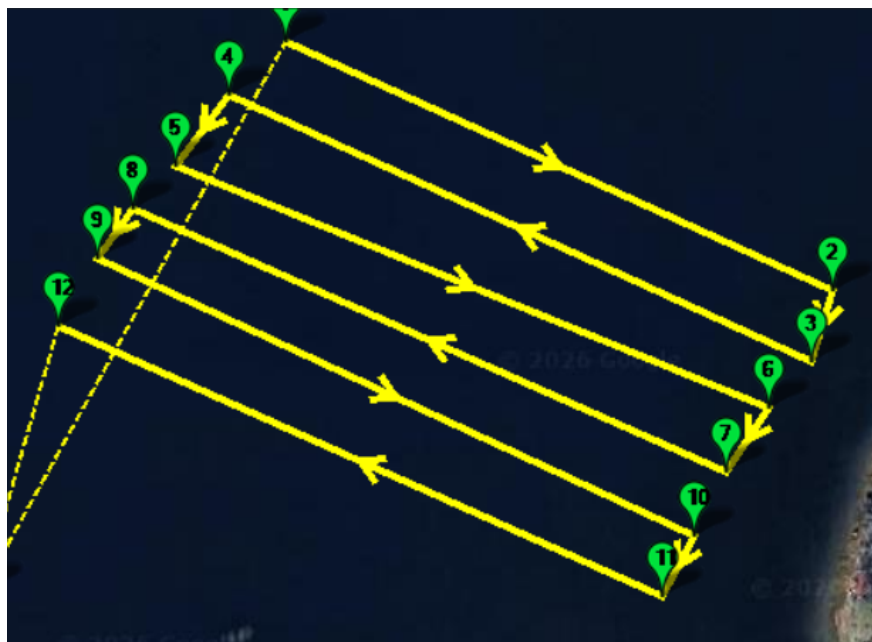


Figure 3.7: This screenshot shows the planned search pattern in Lake Minsi on the ArduPilot software.

3.2.2 Vision Model Data Collection

We used a variety of methods to collect the image data to be used for the model, since there are few publicly available datasets of people in water from the point of view of a USV. The first method was for our team members to volunteer to enter the pool in testing area A and simulate the positions that a marine SAR victim would be in. We then used our machine vision camera to capture images of the volunteers from four different angles to get image data of their front, back, and sides. The second method was similar, but with an expanded scope, with more volunteers from outside the team for greater image diversity. The third method was to obtain access to the University of Ljubljana LaRS image dataset [41], which fits our use case as many of its images include people in and near bodies of water. The limitation of this data for our research was that the pictures were not taken from the perspective of a USV.

The rest of this section is an excerpt from our conference paper, which outlines the methodology with respect to our computer vision model [40].

3.2.2.1 Photo Acquisition Process

“When designing testing procedures, we kept a number of priorities in mind. The presence of life vests was viewed as an independent variable that was changed throughout testing. Special attention was paid to subject diversity and thorough subject capture. Current computer imaging technologies have difficulty in identifying subjects with darker skin tones. This is mainly due to less contrast from the environment and within the skin itself. Our testing procedures prioritized getting subjects of multiple races and skin tones to try and combat this complication. In addition, current photo libraries don’t capture a full range of human movement and positions. A thorough library was captured of subjects from multiple angles, lightings, positions, and postures in the water” [40].

3.2.2.2 Subject Protocol

“Four main postures were utilized during testing that would best represent a life-saving situation. These postures are referred to as: stationary, swimming, hands up, and drowning. Stationary and swimming represent the more typical actions of water activity.

- The stationary posture, as shown in Figure [3.8], is characterized by a subject that has their head above water and doesn’t change location. This represents the basic appearance of treading water, which is a staple action in open water.



Figure 3.8: The subject is in the stationary posture. This posture exposes the top of the subject’s body to the camera. The entire head is visible along with the top of the torso and the top of life jacket that the subject is wearing [40].

- The swimming posture (Figure 3.9) includes all kinds of swimming strokes. The main characteristic of this posture is the fact that the body is tilted on an angle, providing a different appearance to a camera than just treading water. Because all strokes are included, the subject could be facing the water, facing upwards, or even on their side. However, the presence of a tilt remains unchanged while they are swimming.



Figure 3.9: Subjects are in in swimming posture in Pool A. Both subjects are exhibiting the angle that is used to characterize that posture. The subject on the left is angled so their back is up and facing the water. The subject on the right is facing up with their back in the water [40].

- The hands up posture is characterized by the subject raising their arms above head height, as can be seen in Figure [3.10]. This posture is used to represent a swimmer who is waving their arms above their head to get the attention of a lifesaver. By raising the arms above the head, the body drops further down into the water, giving less of a target for the vision system to identify.
- The drowning posture consists of the subject placing their head under the water and going limp, so he floats with only their back exposed to the surface, as shown in Figure [3.11]. This posture is used to represent an unconscious or drowned swimmer. Having only the back of a swimmer exposed presents a difficult challenge for a computer vision system which can typically use human shape and color disparity to identify a human” [40].



Figure 3.10: A subject is doing the hands up posture during testing in Pool A. Due to the raised arms and lack of a life jacket, the subjects torso is deeper in the water than it would be if the subject were treading water [40].



Figure 3.11: A subject is in drowning posture. The only part of the subject visible to the camera is the back of their head and part of the life jacket [40].

3.2.3 Vision Model Training and Testing

All of our computer models (both visual and thermal) were trained and tested using the Roboflow platform [42]. This allowed us to upload images to the site, box images where the human is, and train them based on an existing vision model and those boxed images. We based our initial models on YOLO, then our final visual and thermal models were built on the RF-DETR model [43].

The remainder of this section is an excerpt from our conference paper [40].

“Testing procedures were carried out in three main formats that span two different pools, which will be referred to as Pool A and Pool B. Pool A is a circular tank with a diameter of 50 feet. Pool B is a rectangular pool with 75 feet by 60 feet. Although their depths are different (Pool A: 25 feet deep, Pool B: 13 feet deep), the only requirement is that a subject cannot stand at the bottom, so depth is irrelevant. The testing procedures were designed with the pools in mind.

During testing in circular pool A, the Intel RealSense Depth Camera D456 was positioned at various accessible points around the edge of the pool for the photo shoot process. During testing, subjects were instructed to move around the pool and perform the four key postures. The photos were taken manually through a laptop connected to the camera. These photos were taken in a rapid fashion in order to capture as many angles and scenarios as possible. The subject(s) would then move to various positions in the pool, varying distance from the camera and location in the frame of the camera. Once a suitable number of photographs were captured, the camera operator would adjust their angle around the pool, and the subject(s) would repeat the process.

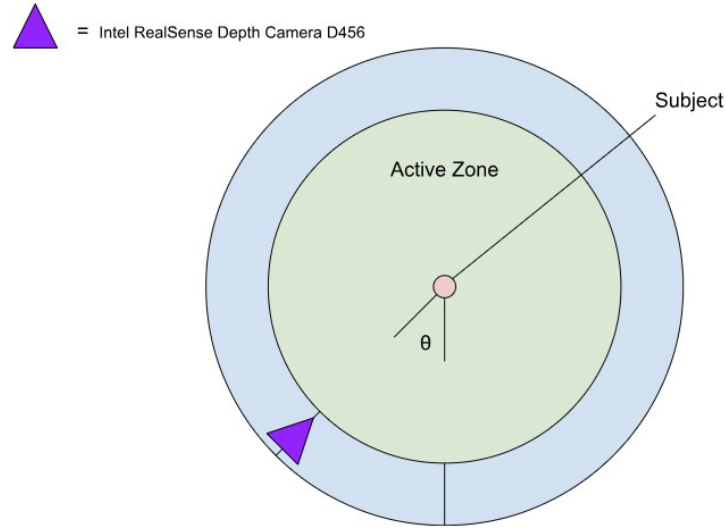


Figure 3.12: This figure shows a basic diagram of the testing process in Pool A. The Intel RealSense Depth Camera D456 takes photos of the subject from the edge of the pool. The subject is free to move around the Active Zone. The camera is moved periodically throughout the testing to vary the photo angle [40].

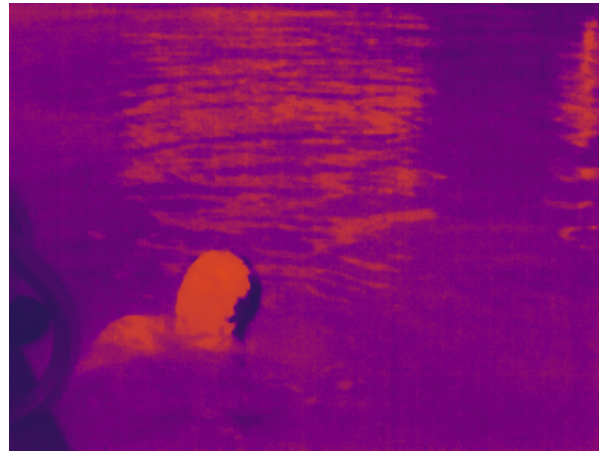


Figure 3.13: This thermal image shows a subject in stationary posture from the point of view of our SeekThermal SD314SPX thermal imaging camera. The subject's head and shoulders appear above the water, but nothing can be seen beyond the surface.

The pool B test photography process was completed at the same rate, although fewer angles were used to capture the photographs. During this testing process, the SeekThermal SD314SPX thermal imaging camera was also used to capture images of the subjects. The two cameras were operated separately and sometimes captured different angles.

The number of subjects varied throughout the tests. Having multiple subjects during testing causes complications to arise when one subject is blocking another subject. By gradually increasing the number of subjects in each successive test, we were able to train the model progressively. Test 1 consists of images of only one subject. Test 2 consisted of two subjects in the pool at the same time. Test 3 consisted of a varying number of subjects, up to 15 subjects at a time.

	Test 1	Test 2	Test 3
Location	Pool A	Pool A	Pool B
Subjects	1	2	Up to 15
Camera Method	Visual	Visual	Visual & Thermal
Photos Taken	Visual: 1,526	Visual: 827	Visual: 4,701
	Thermal: 0	Thermal: 0	Thermal: 5,066

Table 3.3: Testing Variables and Photos Collected from Test 1, Test 2, and Test 3

Once the photo libraries were collected, Test 1 and Test 2 were sampled into two groups: training and testing. These groups were roughly equal, and each made up a significant portion of those photos” [40].

	Test 1	Test 2	Test 3
Training	438	2,353	998
Testing	1,088	4,701	998

Table 3.4: Photo Allocation from Test 1, Test 2, and Test 3

3.2.4 USV-Vision System Integration

Throughout testing of the USV with its vision system, we tested with the visual camera first before testing with both the visual and thermal cameras.

3.2.4.1 USV with Visual Camera Only

Our first testing of our full USV with its vision system tested the ability of the boat to find a person. This first test only had the visual camera mounted on the boat, not the thermal camera. The boat was placed in Testing Area A with the visual camera enabled, with a person standing at the edge of the pool. The objective of this test was to see whether the visual camera could locate the person at the edge of the pool and move in their direction. The visual camera had some difficulty recognizing the subject at the edge of the pool due to its limited range. We then tested a subject in water with the USV. With this shorter range between the USV and the subject, the USV was far more successful in locating the subject. The range of detection of the visual camera was limited to approximately 20 feet. These tests further emphasized the need to improve the view of the visual camera to allow more accurate detection at greater distance. The addition of the thermal camera also provided a significant increase in range and accuracy.

3.2.4.2 USV with Full Vision System

We had significant difficulty in finding a testing location that was suitable for a human subject test. Most bodies of water in close proximity to UMD do not allow swimming, so we came up with a testing apparatus that would act as our test subject. *Craig* (as this apparatus was named) is a rubber mask of a human male that is mounted on top of a flotation device. This is supposed to simulate the head of a test subject above the water. The flotation device was then towed using a remote control boat. It was designed to be towed to a location and then left stationary at the time of testing to represent a subject treading water. The human

shape of the mask was enough to be recognized by the visual camera as seen in Figure 3.14.

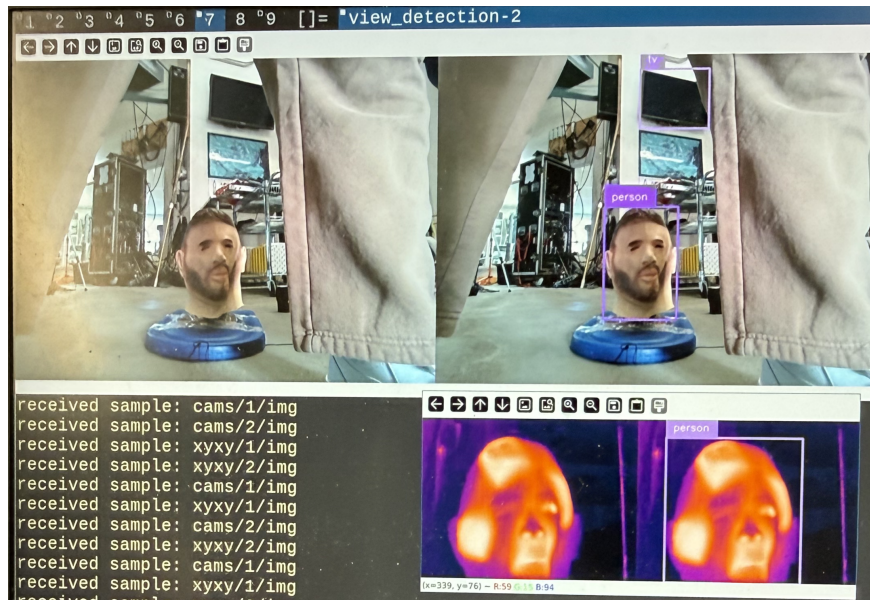


Figure 3.14: This screenshot shows both the thermal and visual cameras identifying *Craig*.

We inserted heating elements on the inside of the mask to increase the surface temperature. To test whether our vision system could effectively detect *Craig*, we placed *Craig* at one end of Testing Area A and positioned the cameras at the opposite end. This provided a 50-foot line of sight from the camera to the dummy head. We ensured that the thermal signature and visual camera were able to detect *Craig*. Once we confirmed the camera's ability to see *Craig*, we took the camera setup to Testing Area D. We placed the head in one corner of a roughly 150-foot diameter retention pond. From the USV's vision system, we were able to detect and distinguish the shape of a human head from opposite corners of the pond.

3.3 Vision System-USV Integration Testing

Our final tests were carried out to analyze the way our vision system and path-planning software interacted. We traveled to McHenry, Maryland, and conducted human testing on Deep Creek Lake. Our first test revolved around the ArduPilot 'Return to Land' (RTL)

feature for a fail safe in case of any issues during testing. From there, we tested ArduPilot's polygon features to set up simple search patterns to use in our human trials. Finally, human testing was performed to evaluate the cooperation of our USV and vision system. Throughout these tests, we used our USV, a human subject, and a team member in a kayak in the waters for safety, as can be seen in Figure 3.15. We ran several search patterns in a quadrilateral zone and allowed the computer to process images while driving along a desired path. After processing images directly on the onboard computer, we switched to relaying image data back to shore and computing there to compare latency and processing capabilities. Using Mission Planner in tandem with our own thermal and visual machine vision model, we obtained qualitative data regarding the USV's position tracking and camera detection. Towards the end of testing, we switched off the visual camera to focus on the thermal camera in an attempt to reduce the computational latency observed initially during the tests.



Figure 3.15: This image shows all in-water aspects of testing done in Deep Creek Lake

The main testing procedure consisted of having the boat spin around with our test subject in the immediate area. If the boat noticed the subject in its field of view, it would turn to face the subject and then follow them for as long as the subject was in frame. Our system constantly received image frames from the cameras and heading estimates from the ArduPilot. For each image frame, we paired it with the closest in time heading sample. After running the object detection algorithm on the image, we calculated the yaw angle at

which the person is located in the image frame. We then added the heading estimate to the yaw to obtain the person-centered heading. Finally, we commanded the ArduPilot to face in that direction.

We completed repetitions of this test over the span of two hours, with breaks to reset or debug. We tracked the number of times the USV located the subject and how long it stayed tracking the subject each time. At the end of testing, we ran a few search patterns using the Mission Planner polygon feature to see how an actual search pattern may impact detection capabilities.

Section 4: Results

In the following section, we discuss results from our completed tests. We examine the results of testing the visual and thermal models. In addition, we discuss results from testing the accuracy of the ArduPilot GPS in a large test area. At this time, we have limited results from testing our entire system, so part of this section is dedicated to future testing that we will do.

4.1 Vision System

To begin our discussion, we analyze our results from the thermal and visual cameras. We discuss tests of different YOLO models against each other to provide a benchmark for our own testing. From our conference paper, we then show the results for our visual and thermal cameras along with a comparison of our visual model and YOLO.

4.1.1 Comparison of Tested Models

In order to have a point of reference for our visual model, we analyzed the performance of multiple YOLOv8 models. To accomplish this, we randomly selected approximately 300 images from one of our datasets and ran 5 different YOLOv8 models (YOLOv8x, YOLOv8s,

YOLOv8l, YOLOv8m, and YOLOv8n) on that dataset to analyze execution time and confidence levels. We found that the YOLOv8x model had an average confidence level of 65.008%, which was the highest of the five, but had the longest average execution time, with it taking about 4.19 seconds for each image. On the other hand, the YOLOv8n model had the shortest execution time of 0.29 seconds and had a lower average confidence level of 59.895%. We found that larger models tended to have longer execution times and higher confidence levels compared to smaller models. Our model achieved a class-specific mAP@50 of 81.6% and a precision of 86.1%. mAP@50 (Mean Average Precision at 0.5 IoU) measures the correctness of the model by counting a prediction correct if it overlaps with at least 50% of the indicated correct outline. Precision is the ratio of true positives to total positives (true and false). The baseline YOLOv8 models showed lower confidence and precision in the specific context of marine human detection. Although the YOLOv8x model provided the highest baseline confidence level, our final model surpassed this, likely because it was specifically trained on datasets like LaRS and our internal pool testing.

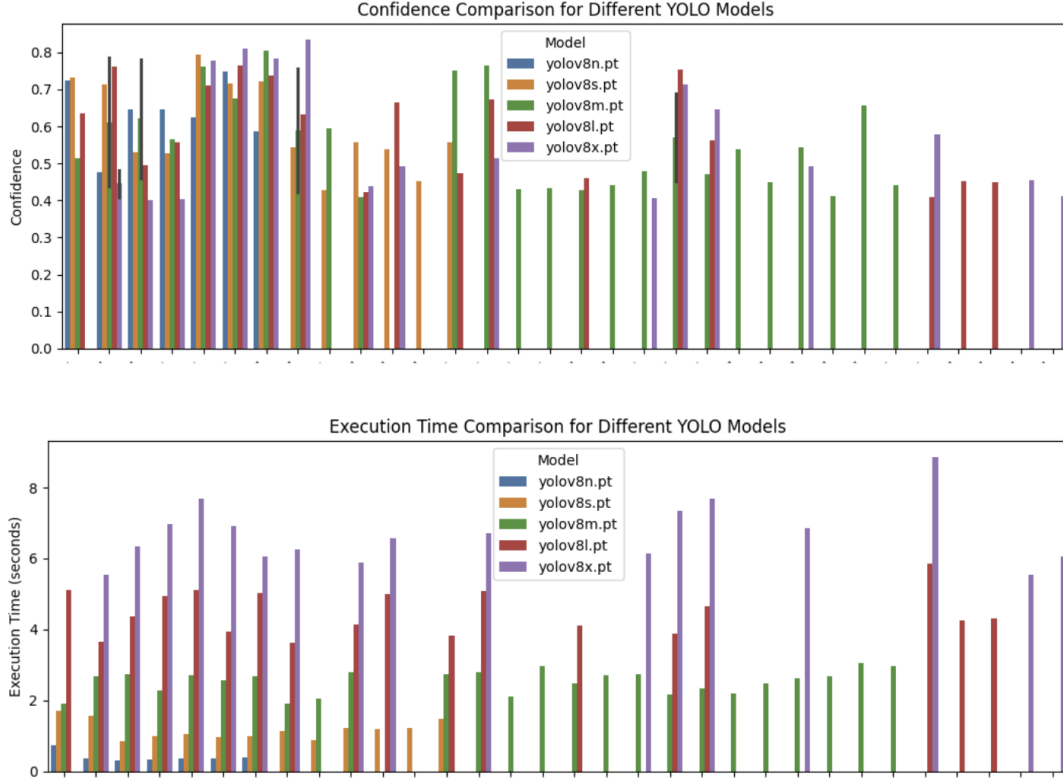


Figure 4.1: Comparison testing of different YOLO computer vision models with respect to confidence rating on pictures of people in water from our datasets as well as the execution time needed to generate that result.

4.1.2 Performance of Vision Models

This section begins with an excerpt from our conference paper to highlight the initial results of the vision system [40].

“We decided to base our results on the confidence levels of the subjects located by the model being tested. This method raises a concern of the presence of false positives in our data. To address this concern, we implemented countermeasures. Firstly, all training data was carefully annotated to make sure the model would never be trained on incorrect data. As an additional fail-safe, we decided to implement a filter on our model so that it only considers a confidence level of 0.4 or higher. As a result, the model must be significantly

confident that a target is a subject for it to be considered a positive identification. For one more layer of safety, we went through all the tested photos to make sure no false positives existed. Therefore, we can be confident that all confidence levels included are representative of our model and not false positives.

These YOLO models can be used as a basis to train new computer vision models from a given dataset of images. The data found YOLOv8x to be the most performant in general, but requiring the highest execution time, introducing a tradeoff that is seen with most models, where higher performance has higher execution time. This testing also provided us with a baseline of what to expect in terms of confidence intervals and execution time when we used our own model for later applications.

4.1.2.1 Visual Results

The visual model test on 1,088 photos of the Test 1 data yielded a strongly left-skewed distribution of confidence values (Figure 4.2) with a mean of 0.7169 and a mode of 0.80-0.85.

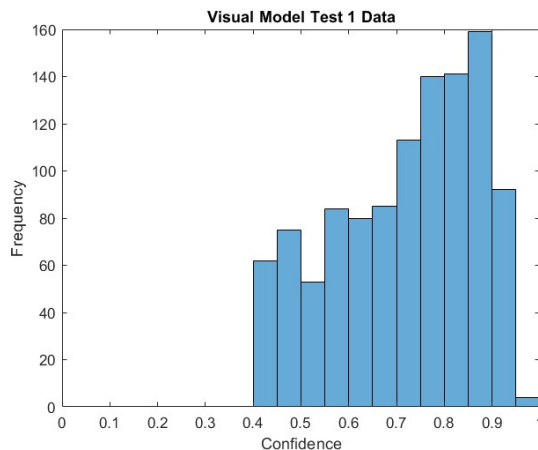


Figure 4.2: Histogram of the confidence levels collected by running our model on the Test 1 data. Test 1 took place in Pool A with one subject being photographed. This histogram is skewed left, with a mean of 0.7169 and a mode of 0.85-0.90.

	Data Points	Mean	Mode	σ
Test 1 Data	1088	0.7169	0.80-0.85	0.1496

Table 4.1: Characteristics of Data Acquired by Testing our Visual Model on Test 1 Data

When dealing with data from Test 2 or Test 3, we needed to take into account that our model can identify multiple subjects in the same photo. Therefore, there are more data points than there are original photographs. By using this system, we were able to find the number of subjects identified in total and use that as a measure of the accuracy of our model.

As the model progressed, it became more confident in its distinctions. Mean and mode rose while standard deviation dropped, making a tighter and higher confidence distribution. Testing the model on Test 2 data yielded a distribution (Figure 8) with a mean of 0.7673 and a mode of 0.84-0.87.

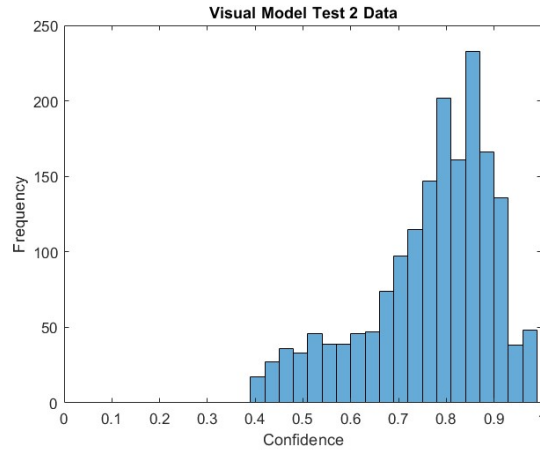


Figure 4.3: This histogram displays the confidence levels of our visual model produced from testing on data from Test 2. Test 2 took place in Pool A with two subjects. The distribution consists of 1747 confidence levels and is skewed left with a mean of 0.7673 and a mode at 0.84-0.87.

	Data Points	Mean	Mode	σ
Test 2 Data	1747	0.7673	0.84-0.87	0.1318

Table 4.2: Characteristics of Data Collected by Testing the Visual Model on Test 2 Data

4.1.2.2 Thermal Results

After training a separate model on the thermal data from Test 3 and testing on another portion of that same data, a confidence distribution (Figure 9) was created from 1198 confidence levels with a mean of 0.8349, a mode of 0.84-0.87, and a standard deviation of 0.0829.

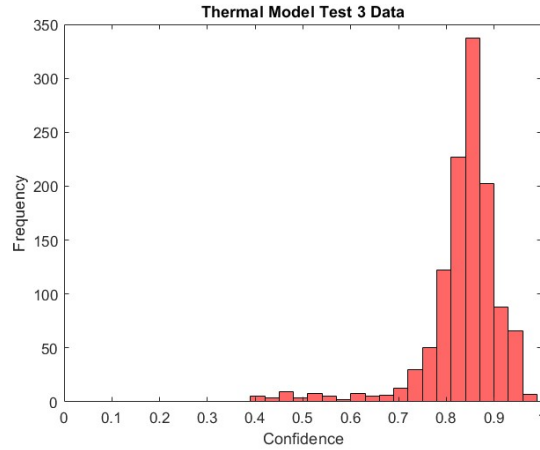


Figure 4.4: The histogram displays the confidence levels acquired from running our thermal model on the thermal data from Test 3. Test 3 was run in Pool B with up to 15 subjects in the frame at any given time. The distribution is created from 1198 confidence levels, and is skewed left with a mean of 0.8349 and a mode of 0.84-0.87.

	Data Points	Mean	Mode	σ
Test 3 Thermal Data	1198	0.8349	0.84-0.87	0.0829

Table 4.3: Characteristics of Data Acquired by Testing Our Thermal Model on Test 3 Thermal Data

4.1.2.3 Comparison

When assembling our testing data for comparison with the data from Nano, we adjusted our filter to have a confidence level of 0.5 or higher in order to be included. Although this is not as representative of our model, Nano uses a threshold of 0.5, so matching it creates more even testing criteria.

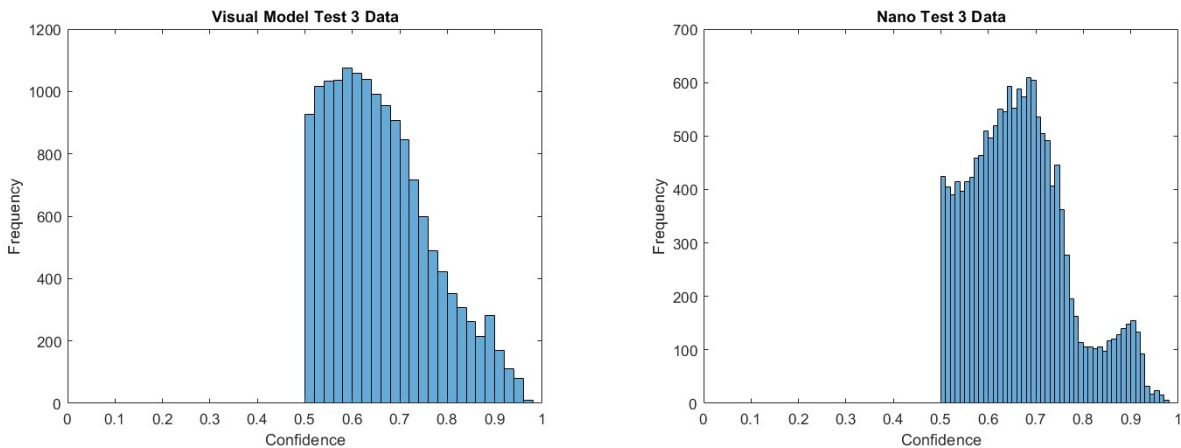


Figure 4.5: Left: Histogram produced with 14,900 confidence levels acquired from testing our visual model on Test 3 data. This distribution exhibits a mean of 0.6605 and a relatively constant decrease in frequency as the confidence level increases from its mode at 0.58-0.60. Right: Histogram produced with confidence levels acquired from testing RF-DETR Nano on the same Test 3 data. This distribution exhibits a mean of 0.6674 and a bimodal distribution. The frequency constantly decreases from its mode at 0.68-0.69 until it picks up again to account for a local mode at 0.90-0.91.

	Our Model	Nano	% Difference
Number of Subject Locations	14,900	15,072	1.14%
Mean Confidence	0.6605	0.6674	1.03%
σ	0.1068	0.1027	3.99%
Skewness	0.6135	0.5788	6.34%

Table 4.4: Characteristics of Data and Percent Differences from Comparison Tests Between Our Visual Model and the Visual Model RF-DETR Nano”

Beyond our conference paper, we developed a final visual and thermal model that was used in our eventual combined tests. For our final visual model, we used 1599 images to train, 525 to validate, and 229 to test, all from the internal image collection. We expanded the model with images from an indoor pool data collection event, where we had other undergraduate students provide a wider scope. This model was built upon the RF-DETR (Nano) model. The model showed a mAP@50 of 85.4%, a precision = 86.6% and a recall of 85%, a clear improvement over the second model in each category. Recall is the fraction of true positives to true positives plus false negatives. The final vision model was intended to balance high-accuracy detection with lower latency requirements of autonomous marine SAR. The model showed a particularly strong performance for the “person” class, achieving a class-specific mAP50 of 81.6%. Furthermore, it also had a precision of 86.1%, highlighting the ratio of true positives to total positives. More importantly, the recall of 83% indicates that potential detections are unlikely to be missed. After testing the visual camera on its own, we switched our focus to the thermal camera. Our final thermal model uses a data set of 998 images from the thermal camera during tests with people in an indoor pool. It was trained on the RF-DETR model, specifically on the ‘Nano’ version. The model had a mAP@50 of 93.3%, a precision of 91.9% and a recall of 91%.

4.2 USV Navigation

The Lake Minsi test was the most in-depth testing we conducted to evaluate the GPS and navigation system. The winds during the test made it difficult for the USV to remain on its exact line, but also allowed us to see how accurate our correction system is. As we can observe, the boat's heading was repeatedly corrected to follow the intended search pattern. One main flaw we found with this testing is the USV's inability to turn in place given its current output configuration. At the moment, the boat makes small 3-point turns, but it would be faster and more efficient to implement pivot turning in the future.

The results of these two tests are shown in Figures 4.6 and 4.7. The average displacement between the boat and the exact search pattern was 7 centimeters. The maximum displacement was 3.11 meters, which was observed during the turning stage of the search pattern. While the boat was in the straight portions of the pattern, the oscillations in displacement had a much smaller amplitude, generally less than 1 meter. These oscillations were much larger in the turning parts of the pattern, getting up to 3 meters, and had a less regular shape than the straight parts of the pattern. The USV took approximately 9 minutes to complete the search pattern in the 500' x 500' area.

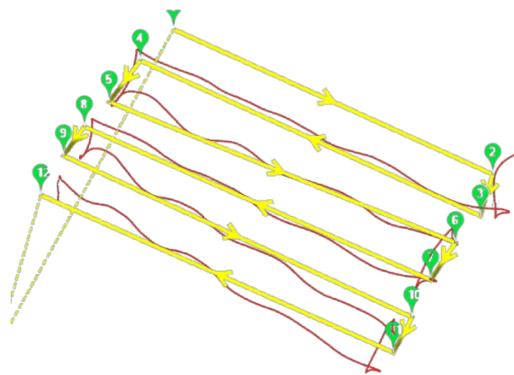


Figure 4.6: This screenshot shows the original search pattern (yellow) with the actual USV movement (red) overlaid on top.

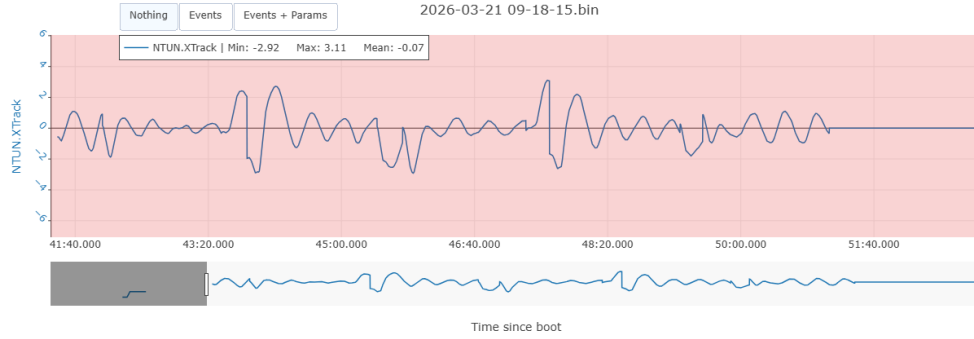


Figure 4.7: This screenshot shows the distance of the boat from the line it followed (the yellow path provided in Figure 4.6)

4.3 USV-Vision System Integration

The first test we conducted on the integrated system was on USV error handling to ensure the safety of the USV and any subjects nearby. With ArduPilot’s RTL feature, we were able to set a return point for the boat to go to whenever there were any errors with the connection or software. In addition, we could put the boat in RTL mode using the ArduPilot UI to have it return to the same place. This was a useful functionality for our tests as it provided us with a fail-safe in the event of any problems during testing.

Beyond failure modes, we were successfully able to run multiple search patterns using the survey grid feature in the ArduPilot software. Specifically, we drew multiple quadrilaterals around an area off the coast and used the simple grid feature to complete creeping line patterns around those polygons. The boat followed these patterns with relative accuracy. These polygons were used throughout the rest of our tests at Deep Creek.

For our first human trial, our program commanded the ArduPilot to rotate the boat clockwise until a person was detected within the camera frame. Once a person was sighted, the program would center the boat on the person and keep that subject in frame. During the test, strong currents and winds proved difficult, yielding the trajectory shown in Figure 4.8.

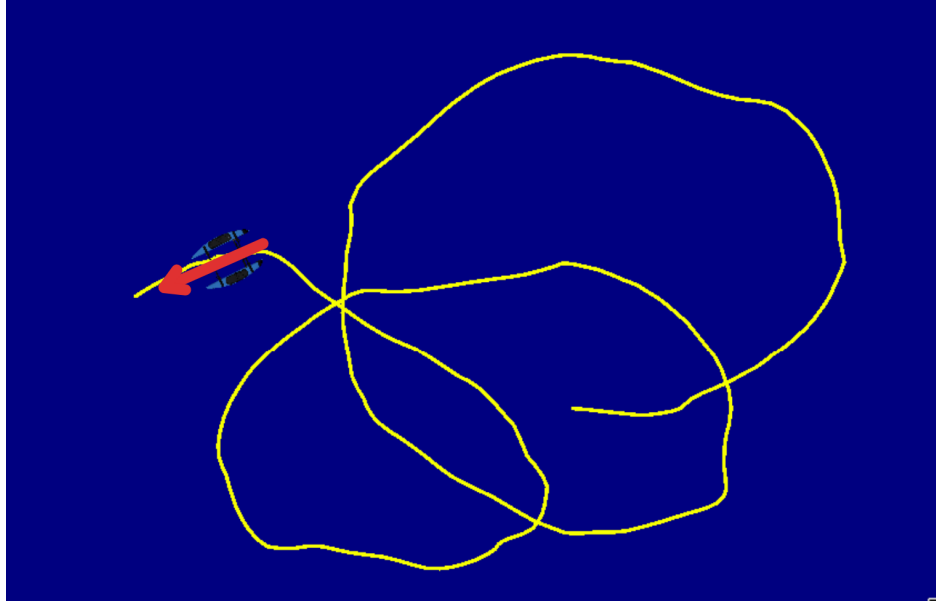


Figure 4.8: In real world testing, the ArduPilot moved the boat in a backwards spiral when commanded to rotate in-place (red arrow indicates vehicle front).

Regarding vision tracking abilities, we found that the system was able to keep track of a person only for brief periods of time. The 63 instances of the system tracking a person produced a median time of around 1.9 seconds, as indicated in Figure 4.9. This short time of tracking the individual resulted from the boat continuing its circular path away from the person, even when the vision system correctly detected and indicated a person. This was due to latency between the boat control system and the vision system, causing the boat to overturn.

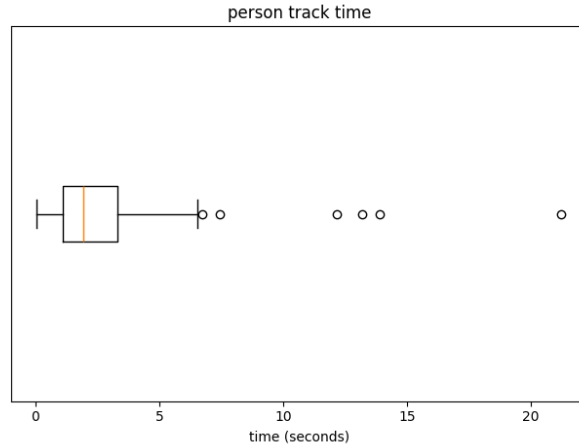


Figure 4.9: This image shows the distribution of successful human tracking time. Denoted are the median (1.9 seconds) and the standard deviation.

Figure 4.10 goes further to show the entire path the boat took while tracking a subject. This is the case where a person has been detected, and the boat tries to stay with them. The subject in the water made small movements as the boat approached; however, they stayed in the same general location. Therefore, the trajectory shown at the bottom of the figure highlights the severity of the leeway and drift during testing. The top of the figure shows the difference between the desired and actual compass direction of the USV, reinforcing the issue of drift and leeway.

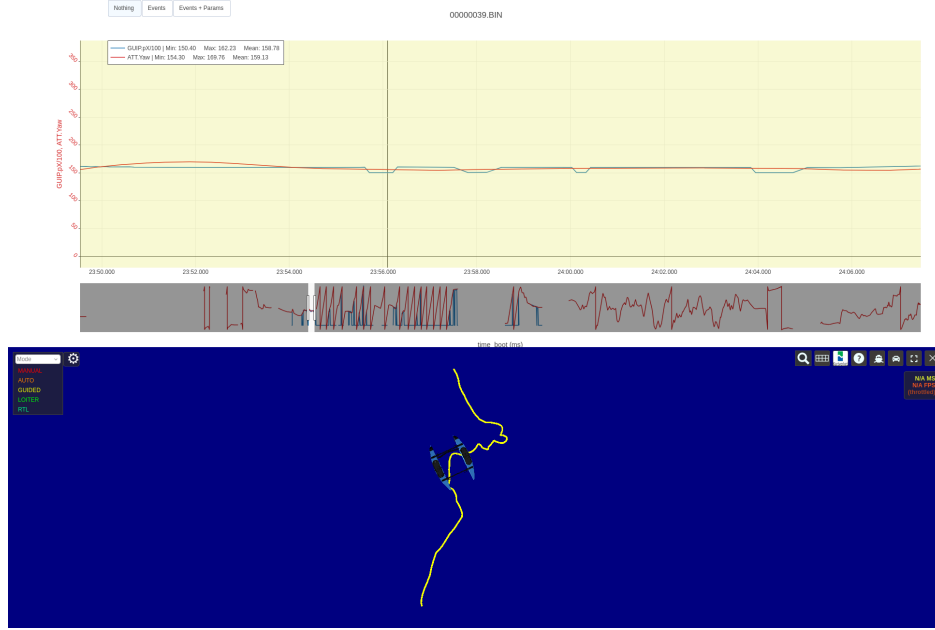


Figure 4.10: Boat tracking a subject. Top: Plot shows commanded heading (blue) versus actual heading (red). Bottom: trajectory of the boat over the course of 26 seconds.

Section 5: Discussion

In this section, we will analyze the results provided previously. We will then discuss potential future work related to our thesis.

5.1 Analysis of Results

Here, we discuss the results we collected, specifically regarding testing the vision system and USV.

5.1.1 Vision System

The analysis of our vision system comes from an excerpt of our conference paper [40].

“Both visual models provided similar results in terms of confidence levels, with the Nano model showing a slight increase around the 0.9 mark. This result indicates the need for

greater image supply in training and testing, and more metrics to determine accuracy. Qualitatively parsing the data, we found that our model handled difficult detection in many cases with higher accuracy, but was less confident in clearer scenarios. These results show that our vision system, although slightly more inaccurate, is capable of identifying subjects to a level similar to that of mass-trained computer vision systems. Throughout our testing, we have seen large improvements in the accuracy of our model with each training, so further improvements can be made by adding more data to our library. The improvement between Test 1 and Test 2 showed an increase of 0.0504 (6.57%) in the average confidence level. Our thermal model (with an average confidence level of 0.8349) outperformed the visual model significantly.”

Since this publication, we have developed final visual and thermal models and measured performance as shown in the results section. The thermal model generally showed higher performance in terms of the statistics provided by Roboflow. The higher performance of the thermal camera is accentuated by our observations when testing the two cameras on the USV itself. The thermal camera showed a significantly greater range of detection compared to the visual camera. This was complicated by the potential for reflections, similar to what we saw when testing in the direction of a building where the background temperature was higher in the camera’s view, as shown in Figure 5.1.



Figure 5.1: Testing in Area D. Observe thermal glare off the water.

5.1.2 USV-Vision System Integration

The software we developed showed mixed results during testing. With the initial tests looking around in a circle, the vision model was able to identify the person in the water from the live feed of the thermal camera. The controls of the boat did not respond in time, and the boat turned past the person. This issue persisted throughout multiple iterations.

These issues largely revolved around the computing setup we used to process images and the latency involved. We used a mobile phone as a hotspot to connect the boat via cellular data. The phone overheated in the waterproof Pelican case when the on-board computer handled boat controls and image processing, generating significant heat. This necessitated

an alternate setup in which we streamed the video feeds back to a computer on land for processing. This resolved the overheating issue since we were no longer running our object detection models on the onboard computer. This proved to be faster but still had a latency attached that was not resolvable during our tests. We attempted to correct for this latency by storing the exact timestamp and heading on the boat when a person was detected. However, this solution caused a further unresolved issue in our messaging system where image and heading data were not in order.

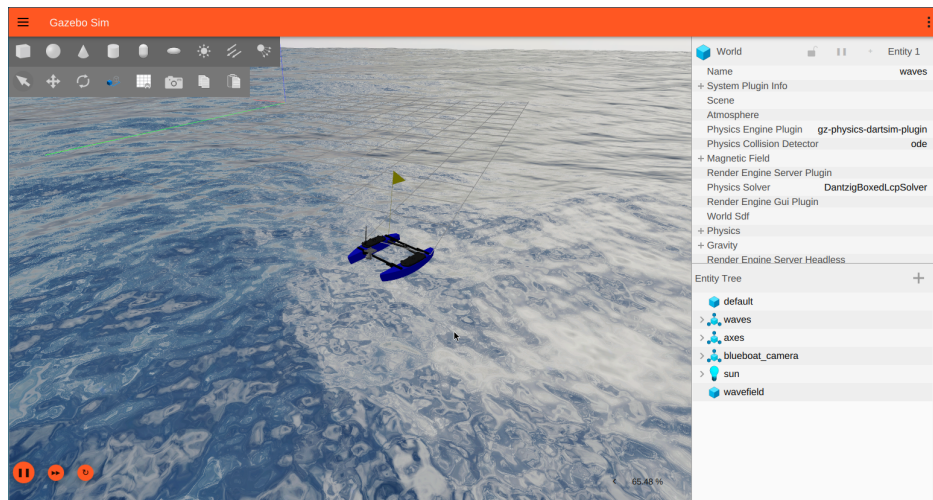


Figure 5.2: Our vehicle in a Gazebo simulation.

The vehicle’s performance in real-world testing was quite unlike what we had achieved in simulation using a Gazebo environment 5.2. When our program had the ArduPilot rotate the vehicle in place, the craft did so with relatively little drift. This behavior is entirely different from what we had experienced in the real world 4.8. This inconsistency possibly stems from the fact that we have not performed any tuning on the ArduPilot. Tuning the vehicle is important to get the best performance from the USV. Furthermore, it is possible that the ArduPilot settings were somehow malconfigured, causing the odd spiral motion. The earlier search pattern test in 4.6 also points to an issue with the vehicle. In simulation, a search pattern is able to be executed without the observed 3-point turns.



Figure 5.3: In the simulation environment, the ArduPilot properly rotates the boat in-place.

In the simulator, our vehicle properly fixated on a person for an indefinite amount of time. Several factors may be contributing to the tracking not working in the real world. Our cameras may have been improperly calibrated, resulting in us sending incorrect heading commands to the ArduPilot. Other logical errors could be present in our software that cause these issues.

One possible improvement to the latency issue is slowing down the rotation rate. With a slower rotation rate, the system would have more time to process and communicate, accounting for the latency in data streams. Secondly, we may simply reverse the rotation of the boat upon every detection. Currently, the vehicle always rotates clockwise until it finds a target. But, for example, if a target is just lost off our port side, it would be more reasonable to begin rotating counterclockwise.

Although many solutions are present, better performing hardware would solve most of these problems. This would also have to include thermal load considerations to ensure the computers do not overheat in a watertight setting. The limited range and resolution of our vision system make for a more difficult detection process.

5.2 Future Work

A major focus of this project was the combination of visual and thermal cameras for a combined vision system. As with any computer vision system, this could always be improved by having more data to train and test on. Most of our data was a collection of college students in water, as that was the population most available to us, so adding age variety would strengthen the models for wider application. In addition, all the pictures were taken in indoor pools with walls in the background, so taking more pictures outside may better narrow down the model for outdoor use. This is especially true in the thermal case, where it adds to the need to develop a model with pictures in a variety of ambient temperatures and different surroundings that may reflect heat in different ways.

With the combined vision system and USV, the greatest problem we faced was computing power and latency. Those could both be resolved in future work by having sufficient onboard computing power and cooling to handle the thermal load. For computing off the boat, this process could be improved by ensuring that data streams are relayed in order and by coupling the pictures with the heading and timestamp. Once the computer sees a person in the image, it just needs to notify the boat to turn back that way, avoiding the latency issues we observed.

In terms of testing our USV one of the main quality of life issues was trying to draw a search pattern around a person to test how well the boat detects during a search. It was hard for us to estimate exactly where the person was in the Mission Planner GPS map interface. To avoid this issue in future tests, it may be helpful for researchers to give the test subject a GPS tracker to hold to more easily draw a shape around them and test how well the boat searches for them in that pattern.

One of the interests in the future direction of this project is a USV system to be used for SAR purposes, mainly looking to reach a potential victim with faster response times. In order to build up to this system, future work would have to involve improvements in terms of obstacle avoidance, search algorithms, communications, and reliability testing for BlueBoat

or comparable USVs on the market. We did not employ any specific obstacle avoidance code in our USV, using only the ArduPilot to autonomously dictate the path of the boat using GPS data. The obstacle avoidance capabilities of the boat would need to be improved in order to handle a more volatile situation, as the boat is on the water with the potential for other boats, waves, and variable conditions. In terms of search algorithms, our project focused mainly on the simpler ones, either turning in a circle or a *Creeping Line* pattern, so more complex and optimized patterns could be tested for improved results in terms of time of detection. SAR operations would require long-range communication in order to relay information from the boat to rescuers on shore or on a boat. This would have the added difficulty of the maritime situation, which is less likely to have strong cellular service or any other internet connectivity. Finally, more reliability testing would be necessary to ensure that the USV could handle more adverse weather conditions, as our testing was completed, either in controlled indoor pools or generally calm outdoor waters. These topics were thought about in the ideation of our project, but were eventually decided as outside our scope in order to focus on the project.

Section 6: Conclusion

The focus of this project at the beginning was to investigate the feasibility of applying a USV for SAR operations. This involved building a platform for a USV that can be easily deployed and scaled. The scope of this project was narrowed to computer vision, as we were able to take all the images collected in bodies of water to train models to attempt to improve detection capabilities in a marine setting. The novelty of this project came with the combination of the thermal camera and the vision camera. The vision system built in detection redundancy in a case where a human operator is not present. Using the BlueBoat base provided a quick way to test on a vehicle to see how effective the combined vision system would be in a real-world application. Beyond testing the vision system, the materials and

resources purchased for this project were easily accessible and cost a few thousand dollars, making it possible for other researchers to continue the work and for public adoption in the future.

An avenue we had wished to investigate was the application of swarm robotics. We attempted to keep costs low so that multiple boats could potentially be used in tandem and fit within our limited budget. Taking this prototype and applying it to real SAR scenarios would be far more feasible if several USVs could be sweeping an area at the same time. Having multiple units will increase the complexity of this project. Communication systems will need to be set up ad-hoc in areas with more sporadic internet access, search algorithms will become more complex and may require greater computing resources, and movements will need to be coordinated to maximize the efficiency of searching an area. Efforts specific to boat design also quickly became out of scope for this project, but improvements can certainly be made in the area of the size of the boat and efficiency. This project highlights a minimum viable USV design that can be easily scaled in a research setting but is unlikely to be as efficient in a real-world scenario.

Overall, the hope of this project is to expand upon the literature on developing autonomous systems that could be applied to SAR scenarios, specifically marine SAR. Having autonomous systems in these cases could help protect rescuers and improve response times by having rapidly deployable autonomous vehicles. The combined vision system provides a potential path to improve accuracy. Furthermore, the system also reduces the risk of false negatives, which would be the preferred case to avoid in lifesaving scenarios.

Bibliography

- [1] Bureau of Transportation Statistics. *U.S. Coast Guard Search and Rescue Statistics, Fiscal Year*. Accessed: 2026-03-24. 2019. URL: <https://www.bts.gov/content/us-coast-guard-search-and-rescue-statistics-fiscal-year>.
- [2] Centers for Disease Control and Prevention. *Drowning Data*. Last updated: January 27, 2026. Accessed: 2026-03-24. 2026. URL: <https://www.cdc.gov/drowning/data-research/index.html>.
- [3] Moulton Avery and Mario Vittone. *How Long Until I'm Rescued? Cold Water Paddling Survival and Rescue Time*. National Center for Cold Water Safety. Accessed May 1, 2026. URL: <https://www.coldwatersafety.org/how-long-until-i-m-rescued>.
- [4] U.S. Coast Guard Auxiliary. *Search and Rescue*. Float Plan Central, U.S. Coast Guard Auxiliary. Accessed May 1, 2026. URL: <https://floatplancentral.cgaux.org/searchandrescue/SAR.htm>.
- [5] NASAR. *Fundamentals of search and rescue*. Sudbury, MA: Jones and Bartlett, Apr. 2004.
- [6] Fan Zhou, Hua Chen, and Peng Zhang. “Performance Evaluation of Maritime Search and Rescue Missions Using Automatic Identification System Data”. In: *Journal of Navigation* 73.6 (2020), pp. 1237–1246. DOI: 10.1017/S0373463320000223.
- [7] International Maritime Organization and International Civil Aviation Organization. *IAMSAR Manual: International Aeronautical and Maritime Search and Rescue Manual*. 2025th. Three volumes: I - Organization and Management, II - Mission Coordination, III - Mobile Facilities. IMO. London, 2025.

- [8] Kemal Ihsan Kilic et al. “Challenges and AI-driven solutions in maritime search and rescue planning: A comprehensive literature review”. In: *Marine Policy* 178 (2025), p. 106692. ISSN: 0308-597X. DOI: <https://doi.org/10.1016/j.marpol.2025.106692>.
- [9] “SAR Convention , 1979 : International Convention on Maritime Search and Rescue, 1979, as amended by resolution MSC.70(69).” In: (2000), v, 26 p. URL: <http://digitallibrary.un.org/record/417074>.
- [10] U.S. Coast Guard Office of Search and Rescue. *SAR Program Information*. Accessed: 2026-03-24. URL: <https://www.dco.uscg.mil/Our-Organization/Assistant-Commandant-for-Response-Policy-CG-5R/Office-of-Incident-Management-Preparedness-CG-5RI/US-Coast-Guard-Office-of-Search-and-Rescue-CG-SAR/CG-SAR-1/SAR-Program-Information/>.
- [11] *Coxswain SAR-Reference Guide*. U.S. Coast Guard. 2006.
- [12] *Uncrewed Surface Vessels - NOAA Ocean Exploration* — oceanexplorer.noaa.gov. [https://oceanexplorer.noaa.gov/technology/usv/#:~:text={U}ncrewed%20surface%20vessel%20\({U}{S}{V}\)%20is,%20published%20{A}ugust%2018%2{C}%2022](https://oceanexplorer.noaa.gov/technology/usv/#:~:text={U}ncrewed%20surface%20vessel%20({U}{S}{V})%20is,%20published%20{A}ugust%2018%2{C}%2022). [Accessed 24-03-2026].
- [13] James Hawkins-Dady et al. “Foiling Autonomous Surface Vessel: Project C-Flyer”. In: ().
- [14] Justin E Manley. “Unmanned surface vehicles, 15 years of development”. In: *OCEANS 2008*. Ieee. 2008, pp. 1–4.
- [15] Wonse Jo et al. “A low-cost and small USV platform for water quality monitoring”. In: *HardwareX* 6 (2019), e00076.
- [16] Zhixiang Liu et al. “Unmanned surface vehicles: An overview of developments and challenges”. In: *Annual Reviews in Control* 41 (2016), pp. 71–93.

- [17] Vladimir Djapic and Dula Nad. “Using collaborative autonomous vehicles in mine countermeasures”. In: *OCEANS’10 IEEE SYDNEY*. IEEE. 2010, pp. 1–7.
- [18] Gabriele Bruzzone et al. “Autonomous mine hunting mission for the Charlie USV”. In: *OCEANS 2011 IEEE-Spain*. IEEE. 2011, pp. 1–6.
- [19] Achmad Zidan Akbar, Chastine Fatichah, and Rudy Dikairono. “Autonomous surface vehicle in search and rescue process of marine casualty using computer vision based victims detection”. In: *2022 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*. IEEE. 2022, pp. 1–6.
- [20] Grant A Wilde and Robin R Murphy. “User interface for unmanned surface vehicles used to rescue drowning victims”. In: *2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. IEEE. 2018, pp. 1–8.
- [21] Diogo Machado et al. “Water jet based autonomous surface vehicle for coastal waters operations”. In: *2014 Oceans-St. John’s*. IEEE. 2014, pp. 1–8.
- [22] Dina Khaled et al. “Development of a sustainable unmanned surface vehicle (USV) for search and rescue operations”. In: *the international undergraduate research conference*. Vol. 5. 5. The Military Technical College. 2021, pp. 462–468.
- [23] Muhayy Ud Din et al. *MarineX: Design and Implementation of Unmanned Surface Vessel for Vision Guided Navigation*. 2023. arXiv: 2311.17197 [cs.R0]. URL: <https://arxiv.org/abs/2311.17197>.
- [24] Francisco Fernández Ramírez et al. “Coordinated sea rescue system based on unmanned air vehicles and surface vessels”. In: *OCEANS 2011 IEEE - Spain*. 2011, pp. 1–10. DOI: 10.1109/Oceans-Spain.2011.6003509.
- [25] Jan Dufek and Robin Murphy. “Visual pose estimation of USV from UAV to assist drowning victims recovery”. In: *2016 IEEE international symposium on safety, security, and rescue robotics (SSRR)*. IEEE. 2016, pp. 147–153.

- [26] Chuan Liu et al. “Development of usv autonomy: Architecture, implementation and sea trials”. In: *Brodogradnja* 73.1 (Mar. 2022), pp. 89–107.
- [27] Wei Guan and Kuo Wang. “Autonomous Collision Avoidance of Unmanned Surface Vehicles Based on Improved A-Star and Dynamic Window Approach Algorithms”. In: *IEEE Intelligent Transportation Systems Magazine* 15.3 (2023), pp. 36–50. DOI: 10.1109/MITS.2022.3229109.
- [28] Di Wang et al. “Real-Time Obstacle Detection and Localization for USV Using Stereo Vision and Edge Computing”. In: *2024 International Conference on Artificial Intelligence of Things and Systems (AIoTSys)*. 2024, pp. 1–8. DOI: 10.1109/AIoTSys63104.2024.10780548.
- [29] Yufei Zhuang et al. “Dynamic Path Planning of USV Based on Improved Artificial Potential Field Method in Harsh Environment”. In: *2022 China Automation Congress (CAC)*. 2022, pp. 5391–5396. DOI: 10.1109/CAC57257.2022.10054793.
- [30] K Jayalath and SR Munasinghe. “Drone-based autonomous human identification for search and rescue missions in real-time”. In: *2021 10th International Conference on Information and Automation for Sustainability (ICIAfS)*. IEEE. 2021, pp. 518–523.
- [31] Mostafa Rizk et al. “Optimization of deep-learning detection of humans in marine environment on edge devices”. In: *2022 29th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*. IEEE. 2022, pp. 1–4.
- [32] Mingkan Chen et al. “Weather-aware object detection method for maritime surveillance systems”. In: *Future Generation Computer Systems* 151 (2024), pp. 111–123.
- [33] Christopher Dahlin Rodin and Tor Arne Johansen. “Detectability of objects at the sea surface in visible light and thermal camera images”. In: *2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO)*. IEEE. 2018, pp. 1–10.

- [34] Donato Cafarelli et al. “MOBDrone: A drone video dataset for man overboard rescue”. In: *International Conference on Image Analysis and Processing*. Springer. 2022, pp. 633–644.
- [35] Valeria Alexandra Feraru, Rasmus Eckholdt Andersen, and Evangelos Boukas. “Towards an autonomous UAV-based system to assist search and rescue operations in man overboard incidents”. In: *2020 IEEE international symposium on safety, security, and rescue robotics (SSRR)*. IEEE. 2020, pp. 57–64.
- [36] Heinrich Ruser. “Object recognition with a smart low-cost active infrared sensor array”. In: *1st International Conference on Sensing Technology*. 2005, pp. 21–23.
- [37] Thomas R Walter et al. “Underwater and drone based photogrammetry reveals structural control at Geysir geothermal field in Iceland”. In: *Journal of Volcanology and Geothermal Research* 391 (2020), p. 106282.
- [38] Andreas Hermann, Jérôme Chladek, and Daniel Stepputtis. “iFO (infrared Fish Observation)—An open source low-cost infrared underwater video system”. In: *HardwareX* 8 (2020), e00149.
- [39] Blue Robotics. *BlueBoat – USV for Hydrographic Surveys and Robotics*. Accessed: 2026-03-27. 2025. URL: <https://bluerobotics.com/store/boat/blueboat/blueboat/>.
- [40] Liam Allen et al. “Vision System Model Development for Maritime Search and Rescue Operations”. In: *OCEANS 2025 - Great Lakes*. 2025, pp. 1–6. DOI: 10.23919/OCEANS59106.2025.11245034.
- [41] Lojze Žust, Janez Perš, and Matej Kristan. *LaRS: A Diverse Panoptic Maritime Obstacle Detection Dataset and Benchmark*. 2023. arXiv: 2308.09618 [cs.CV]. URL: <https://arxiv.org/abs/2308.09618>.
- [42] B. Dwyer, J. Nelson, T. Hansen, et al. *Roboflow (Version 1.0)*. Computer vision. Roboflow. 2026. URL: <https://roboflow.com>.

- [43] Isaac Robinson et al. “RF-DETR: Neural Architecture Search for Real-Time Detection Transformers”. In: *International Conference on Learning Representations (ICLR)*. Accepted to ICLR 2026; also available as arXiv:2511.09554. 2026. DOI: 10.48550/arXiv.2511.09554. URL: <https://arxiv.org/abs/2511.09554>.
- [44] University of Maryland. *About Launch UMD*. <https://launch.umd.edu/about>. Accessed: 2026-03-25. 2026.
- [45] Maryland Sea Grant. *Maryland Sea Grant*. <https://www.mdsg.umd.edu/>. Accessed: 2026-03-25. 2026.
- [46] David Leslie. “Understanding bias in facial recognition technologies”. In: *The Alan Turing Institute* 1 (Sept. 2020). DOI: <https://doi.org/10.5281/zenodo.4050457>. URL: <https://arxiv.org/ct%5Curl=https%3A%2F%2Fdx.doi.org%2F10.5281%2Fzenodo.4050457&v=e0b972fd>.

Appendix

Budget

The team developed a preliminary budget during a team meeting on March 28th, 2024. The two tables corresponding to the preliminary budget are shown below. The first table is the distribution of the proposed budget of \$6650 among the three subteams at the time: the vision, hull, and integration subteams. The second table delves into the categorized distribution of items that each subteam will purchase. The integration subteam, which received an overall budget of \$750, allocated its budget to an ArduPilot (\$600), a microcontroller (\$100), and any devices or hardware-related communication (\$50). The vision subteam had a budget of \$1,800 and allocated it across three categories: \$600 for a thermal camera, \$700 for a computer, and \$500 for a vision and range sensor. The hull subteam received a budget of \$4000, and the team would finalize the categories for specific items in the future.

Budget After March 28, 2024 Meeting				
Distribution by Subteam		Distribution by Item and Subteam		
Subteam	Amount	Item	Cost	Subteam
Vision	\$1800.00	Thermal Camera	\$600.00	Vision
Integration	\$750.00	Computer	\$100.00	Vision
Hull	\$4,000.00	Vision and Range Sensor	\$50.00	Vision
		ArduPilot	\$600.00	Integration
		Microcontroller	\$700.00	Integration
		Communication	\$500.00	Integration
Total	\$6,550.00	Estimation for Hull	\$4,000.00	Hull

Table 1: Team OARS Preliminary Budget as of March 28, 2024

The team's funding landscape changed after the Spring 2024 academic semester, with funds obtained from various sources, as shown in the table below. It is important to first note that the team raised funds in the Fall 2023 academic semester through LaunchUMD, a crowdfunding program run by the University of Maryland, where organizations around campus can raise money to fund their programs [44], earning \$3,505 during its month-long campaign. The team earned more funds after the fall semester. One of those funding sources came from the Library Awards during the Spring 2024 academic semester. The team also participated in a month-long LaunchUMD campaign, during which the team earned \$1,919 and an additional \$500, two \$250 bonuses, for having raised the most money and having the most donors among the Gemstone research teams participating in the LaunchUMD funding cycle. A third avenue in which the team gained funds was from the Gems Olympics hosted by the Gemstone Honors Program, where various research teams in the program competed against each other in different events. The team came victorious in the competition, and the team earned an additional \$500 in funds from the event. The team participated in the next two LaunchUMD campaigns during the Fall 2024 and Spring 2025 academic semesters to earn \$935 and \$50 from the two respective funding opportunities. The team has earned \$9,409 from the mentioned funding opportunities and events. A table listing all of the funding sources is provided below.

Team Funds Acquired		
Amount	Source	Semester
\$2,000	Library Awards	Spring 2024
\$3,505	LaunchUMD	Fall 2023
\$1,919	LaunchUMD	Spring 2024
\$500	LaunchUMD Bonus from GEMS	Spring 2024
\$500	Gems Olympics	Spring 2024
\$935	LaunchUMD	Fall 2024
\$50	Launch UMD	Spring 2025

Table 2: Team Funds Acquired

The table below summarizes the team’s expenses as of March 24th, 2026. The expenses are split into categories: vision, integration, hull, miscellaneous, and conference. The first three categories, vision, integration, and hull, correspond to the expenses made by the three subteams. The fourth category comprises miscellaneous charges made by the three subteams or for testing purposes at one of the team’s testing locations. The last category corresponds to conference expenses. Within those five categories, the team’s expenses were covered by the team’s finances or by funding provided by the Gemstone department through the Maryland Sea Grant fund [45].

The vision subteam has spent a total of \$1,413.70, with \$459.99 from the team’s funds and \$957.71 from the Maryland Sea Grant fund. The integration subteam has spent \$914.04, with \$400.09 coming from the team’s funds and \$513.95 provided by the Maryland Sea Grant fund. The hull subteam has spent a total of \$3952.78, with \$1649.59 from the team’s funds and \$2303.19 from the Maryland Sea Grant fund. The miscellaneous category has a total expenditure of \$556.75, with the team’s funds covering \$289.09 and the Maryland Sea Grant fund covering \$267.66. The conference charges are for four team members attending the Ocean 2025 Great Lakes Conference from September 29th, 2025, to October 2nd, 2025. The

team spent \$5196.43 attending the conference. The conference charges include transportation to and from the conference, per diem for the four team members, and accommodations while in Chicago.

Team OARS Expenses as of March 24th, 2026		
Category	Team's Funds	Maryland Sea Grant Fund
Vision	\$459.99	\$957.71
Integration	\$400.09	\$513.95
Hull	\$1,649.59	\$2,303.19
Miscellaneous	\$289.09	\$267.66
Conference	\$5,196.43	\$0.00

Table 3: Team OARS Expenses as of March 24th, 2026

Ethics and Equity-Impact Report

The diversity of our vision system database provides an ethical consideration for the project. Training a model requires a database of a wide variety of ages, genders, and races to handle all possible people detection cases. According to Dr. David Leslie from the Alan Turing Institute, facial recognition algorithms are more accurate with lighter skin tones due to the calibration of camera sensors to favor lighter skin tones. This was exemplified in a video in which a Hewlett Packard computer was able to accurately track the face of a Caucasian individual but not the face of a person with a darker complexion [46]. In order to account for these potential biases, we hosted many events to collect pictures from a variety of people in water. We reviewed existing visual databases to ensure significant representation of varied groups in our database. Our goal with these steps is to ensure that effort is made to account for visual differences in people. Furthermore, we focused on using our thermal vision system extensively, which does not rely on visible light, negating any potential skin tone biases.

Due to the nature of SAR, our USVs will often be deployed in extreme weather conditions. This raises the concern that some of our USVs could be lost at sea or rendered unusable during deployment. In order to prevent our USVs from being lost at sea, we implemented GPS tracking through ArduPilot, which would allow us to locate any lost USVs. We also encased our batteries in watertight compartments. This prevents water from reaching the battery and other electrical components so that no chemicals can leak out and pollute the surrounding water.