

# Ultrasonic activated energy-efficient autonomous camera system for activity observation

Sifiso Sechemane<sup>1\*</sup>, Anthony Barnard<sup>3</sup>, Andrew Young<sup>2</sup>, and Theo van Niekerk<sup>1</sup>

<sup>1</sup>Department of Mechatronics, Nelson Mandela University, Gqeberha, South Africa

<sup>2</sup>eNtsa Engineering, Nelson Mandela University, Gqeberha, South Africa

<sup>3</sup>Marine Remote Imagery Platform, South African Institute for Aquatic Biodiversity, Gqeberha, South Africa

**Abstract.** This Paper presents a proof-of-concept system for autonomous object detection and tracking, employing a pan-tilt mechanism and ultrasonic sensors for energy-efficient operation. The system's design aims to reduce unnecessary power consumption by activating energy-intensive components, such as the camera and LED lights, only when nearby objects are detected. Key functionalities, including distance sensing and object detection, are evaluated across multiple tests to assess system performance. While the system demonstrated consistent distance sensing and power savings, object detection results were inconsistent, primarily due to the limited training dataset and hardware constraints. Recommendations for future improvements include expanding the dataset, exploring advanced object detection models, and upgrading the camera to enhance detection consistency and overall performance. This prototype serves as a stepping stone toward developing robust, autonomous underwater photography systems for marine research.

## 1 Introduction

The vast depths of the world's oceans are teeming with mysteries, includes the elusive coelacanth, a nocturnal fish species discovered off the South African coast. Despite its intrigue, this deep-sea dweller presents considerable challenges for marine research due to its habitat in caves located hundreds of meters below sea level.

Studying these rare creatures often necessitates extensive and costly marine expeditions, as their habitats are far from shorelines. Such expeditions typically require large vessels with various resources, including fuel, sustenance, personnel, and specialised equipment, leading to substantial overall expenses [1]. Traditional methods of underwater exploration and documentation, such as diving to investigate topics of interest, while exhaustive, can be resource-intensive, risky, and sometimes intrusive, especially in deeper, more sensitive marine environments [2]. Recognising these challenges, innovative solutions are needed to advance research in this field.

---

\* Corresponding author: [sechmanes@gmail.com](mailto:sechmanes@gmail.com)

This paper presents a proof-of-concept autonomous camera system designed to track and photograph objects of interest within their natural habitats. The proposed system is housed within a see-through dome equipped with a pan-tilt mechanism to ensure dynamic and focused imaging. Although parts of the design incorporate waterproofing to withstand underwater conditions, this study emphasises other aspects of the system, such as the performance of the camera's object detection capabilities, rather than subjecting its waterproof capabilities to testing.

A core feature of the design is its energy-efficient operation, allowing prolonged deployments with minimal energy wastage. The system is designed to allow the more energy-draining components, such as the camera and lighting, to remain dormant, in an off mode, becoming active only when the system detects nearby entities. In addition to this energy-conscious approach, the design includes a self-cleaning mechanism to address challenges posed by the likelihood of the see-through dome accumulating debris over time. LED illumination is implemented as lighting to allow for the system to have vision in low-light environments.

The primary objectives of this proof-of-concept are to showcase the potential of an autonomous system in capturing marine data with minimal human intervention and to establish a foundation for design advancements addressing similar problems. This report delves into the design nuances, focusing on the system's object detection capabilities and strategic energy conservation methods. Through this exploration, the goal is to gather invaluable insights that will shape the evolution of a more refined and efficient final system.

The remainder of this paper is structured as follows: Section 2 presents a comprehensive literature review, covering relevant research in underwater photography, object detection, and power conservation in autonomous systems. Section 3 details the methodology and design of our proof-of-concept system, explaining how we integrated various components including off-the-shelf ultrasonic sensors, a custom-designed pan-tilt camera mechanism, and an original dome cleaning system. Section 4 describes the testing procedures and presents the results of our experiments, focusing on sensor calibration, object detection accuracy, and power consumption. Finally, Section 5 concludes the paper, summarizing our findings, discussing the limitations of the current system, and proposing directions for future research and development in autonomous underwater photography systems.

## 2 Literature review

Capturing visual data in underwater environments poses unique challenges. Traditional methods, such as diving expeditions, can be resource-intensive and intrusive and may disturb the sensitive marine ecosystem [2]. To mitigate these issues, various approaches have been proposed, including the use of remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs) equipped with cameras [3,4]. However, these solutions often require continuous human intervention or are limited by tethered operations.

As an alternative, stationary camera systems offer the advantage of targeted, non-intrusive monitoring within a fixed vicinity. The proposed system in this study is designed as a pan-tilt camera with an object detection algorithm, aiming to track autonomously and photograph objects of interest in their natural habitat.

Object detection and tracking have been extensively studied in various domains, including human and vehicle identification [5,6]. However, animal identification and re-identification from images and videos have received comparatively less attention. Niu et al. [7] proposed an object-tracking method using an IP pan-tilt-zoom (PTZ) camera for human upper body tracking in online applications. Their method detects candidate targets by extracting moving objects using optical flow and sampling around the image centre, with the target being detected among candidate samples using a fuzzy classifier.

Zhang et al. [8] presented an effective approach for active tracking with a PTZ camera. They adopted background subtraction with a Gaussian Mixture Model for object detection and region covariance for object description. They employed a local search method with motion compensation for active tracking and acceleration, demonstrating the feasibility and efficiency of their system through indoor and outdoor video sequences.

In the context of underwater fish detection, Muksit et al. [9] introduced YOLO-Fish, a deep learning-based fish detection model. They proposed two models, YOLO-Fish-1 and YOLO-Fish-2, which enhance YOLOv3 by addressing issues related to the misdetection of tiny fish and the capability to detect fish appearance in dynamic environments. The authors also introduced two datasets, DeepFish and OzFish, to test their models, achieving average precision of 76.56% and 75.70%, respectively, for fish detection in unconstrained real-world marine environments.

In the context of marine life detection and tracking, Leira et al. [10] developed a vision-based system for detecting and tracking fish from a moving underwater vehicle. Their evaluation metrics focused on detection rate and tracking consistency, emphasizing the importance of maintaining continuous detection in dynamic underwater environments. This approach underscores the relevance of consistency-based metrics in real-world applications where maintaining object tracking over time is crucial, particularly in challenging conditions such as underwater environments.

Williams et al. [11] compared two state-of-the-art methods for animal identification: the MMDetector, which outputs bounding boxes, and the UniTrack video tracker. They proposed a combination method to fuse the outputs of both methods, demonstrating its capability to outperform the individual techniques.

Chavda and Dhamecha [12] surveyed various object detection and tracking techniques using PTZ cameras, highlighting the challenges associated with dynamic backgrounds and the need for robust algorithms to handle occlusions and complex scenes.

Patil et al. [13] developed a machine vision-enabled bot for object tracking using a Raspberry Pi, a pan-tilt mechanism, and a web interface. Their system utilised TensorFlow for object detection and OpenCV for image processing, demonstrating the potential for remote object tracking and control. While their system focused on above-ground applications, it showcases the integration of object detection algorithms and pan-tilt mechanisms, which is relevant to the present study.

The present study builds upon these existing works by proposing an autonomous camera system optimised for power conservation, incorporating a pan-tilt mechanism and object detection algorithms. The system is designed to remain dormant until triggered by nearby entities, enabling prolonged deployments in underwater environments. Furthermore, including a self-cleaning mechanism and LED illumination addresses the challenges posed by low-light conditions and the accumulation of debris on the camera dome.

## **3 Methodology and design**

### **3.1 Application requirements**

The primary objective of this study is to develop a prototype of an autonomous camera system capable of detecting and tracking objects of interest in various environments, including underwater and in the wild. Although waterproofing is not a strict requirement for this prototype, the design considers the potential need for water resistance and protection from environmental factors such as rain, moisture, and dirt.

The system incorporates a see-through dome enclosure to protect the camera. Additionally, an automatic cleaning mechanism maintains the dome's clarity in case debris

accumulates over time. LED lighting is also included to illuminate the camera in low-light conditions.

Power conservation is a crucial consideration in the system's design. To minimise battery consumption, the camera and LED lights remain offline during the initial startup. Instead, ultrasonic sensors, which consume less power, are used to detect nearby objects. The microprocessor continuously monitors the readings from these sensors, and once an object comes within a specified range, it triggers the activation of the camera and LED lights.

Upon activation, the camera works with an object detection algorithm running on the microprocessor. The algorithm collaborates with the motors of the pan-tilt mechanism to orient the camera and search for objects of interest, such as fish, even in non-aquatic environments. When an object of interest is detected, the pan-tilt motors adjust the camera's position to keep the object centred within the frame. If the object remains centred for a sufficient duration, the camera captures and saves images.

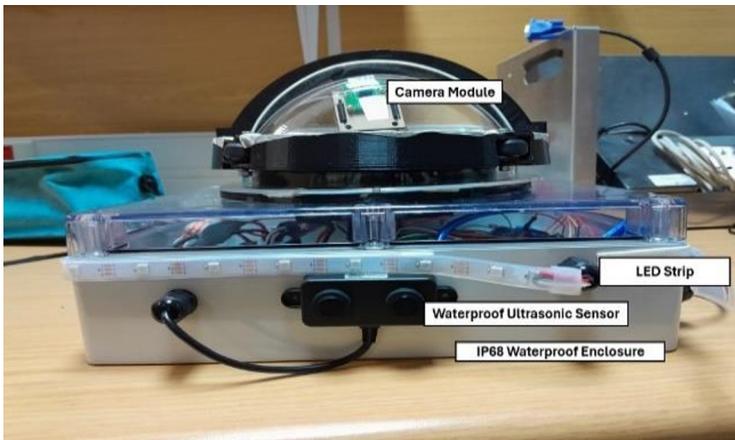
The system continues to track and photograph the object of interest until a specified number of images are taken or if the object is no longer detected for a certain period. At this point, the camera and LED lights are switched off, and the system reverts to relying on the ultrasonic sensors for object detection.

To maintain optimal performance, the system's dome undergoes periodic cleaning by the integrated cleaning mechanism at predetermined intervals.

These application requirements guide the development of the prototype, focusing on object detection, tracking, power efficiency, and adaptability to various environments while considering the potential for future waterproofing and wild deployment.

### 3.2 System overview

The prototype design, as shown in Figure 1, is built around an IP68 polycarbonate enclosure. An acrylic dome, housing a Raspberry Pi Camera Module v2 and a pan-tilt mechanism, is mounted on the enclosure's lid, allowing the camera an unobstructed view of at least 180 degrees.



**Fig. 1.** Overall system design.

To address the potential accumulation of dirt on the acrylic dome, a 3D-printed cleaning system is integrated into the design. This system features an arc structure with a sponge attached underneath, rotating around the dome's surface to wipe away debris. The mechanism behind this autonomous cleaning process is detailed in Section 3.4.

Various essential electrical components are housed inside the IP68 enclosure. A WS2812B individually addressable LED strip is wrapped around the enclosure's surface,

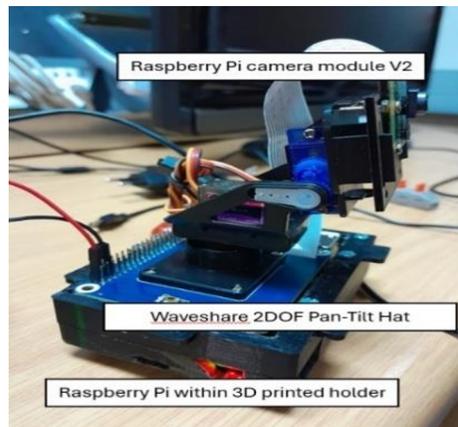
providing programmable lighting for the camera. The microprocessor controls the LED strip's activation and deactivation to conserve power.

Three AO2YYUW waterproof ultrasonic sensors are positioned on three of the four sides of the enclosure. These sensors were selected for their IP67 rating and ability to function reliably in the system's intended environment. Although the sensors' 60-degree beam pattern does not offer full coverage around the enclosure, their placement was sufficient for the prototype's goal: detecting nearby objects to trigger the activation of the LED lights and camera system. This configuration optimises power consumption by using the low-power ultrasonic sensors, which collectively draw only 0.021A of current, to keep the more power-intensive components, such as the camera and LED lights (drawing 1.404A in total), dormant until needed.

Further details on the system's current draw can be found in Section 4.4. Given that the system is a proof of concept, full environmental coverage was not prioritized. The primary focus was on achieving functional performance and power efficiency with the available components.

### 3.3 Pan-tilt mechanism

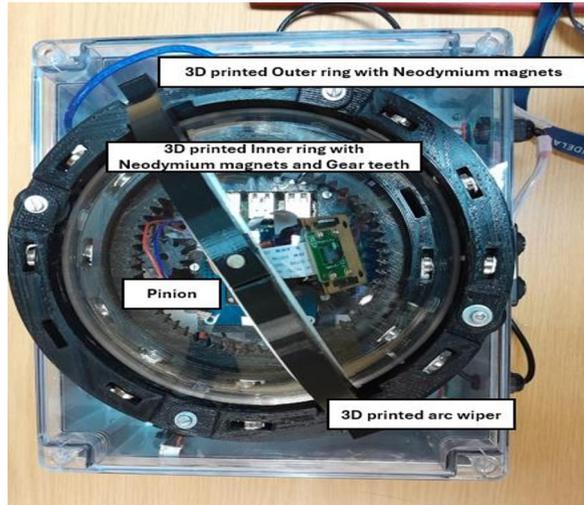
Figure 2 illustrates the pan-tilt camera setup, which consists of a Raspberry Pi Camera Module mounted on a Waveshare 2-DOF Pan-Tilt hardware attached on top (HAT). The HAT includes a control board, a pan-tilt bracket, and two SG90 servo motors. The Raspberry Pi, situated beneath the Pan-Tilt HAT and inside a 3D-printed holder, controls the camera module, pan-tilt mechanism, object detection algorithm, and image capture. The 3D-printed holder is responsible for securing the camera system, elevating it from the hole cut out on the enclosure lid.



**Fig. 2.** Pan-tilt camera setup.

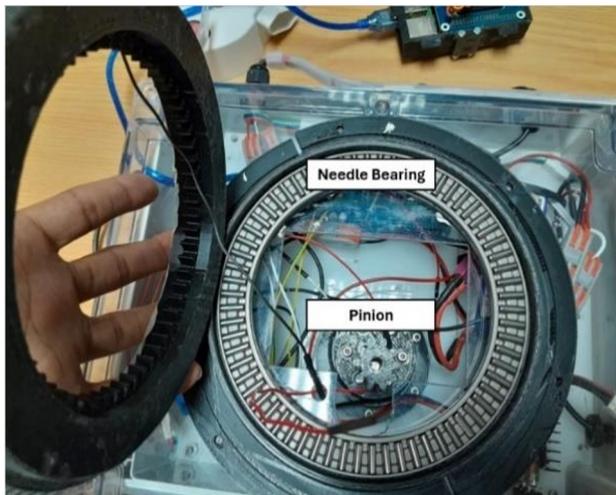
### 3.4 Cleaning subsystem

Figure 3 presents a top view of the system, highlighting the 3D-printed cleaning mechanism. The outer ring, which carries neodymium magnets resembling the size of triple-stacked 2 rand coins, has wheels attached for smooth rotation. This outer ring is connected to a bracket with a sponge material that wipes the dome's surface. The movement of the outer ring is driven by the magnetic attraction between its magnets and those in the inner ring, which remains inside the dome. The neodymium magnets have a pull strength of approximately 3.3kg, providing sufficient coupling to ensure reliable movement of the outer ring under normal use conditions.



**Fig. 3.** Top view of system.

The inner ring is rotated by a pinion gear, driven by a geared DC brush motor (12VDC 0.5A 13RPM) inside the enclosure. Figure 4 provides a top-down view of the system with the camera, wiper, and dome removed, revealing the inner ring's gear teeth.

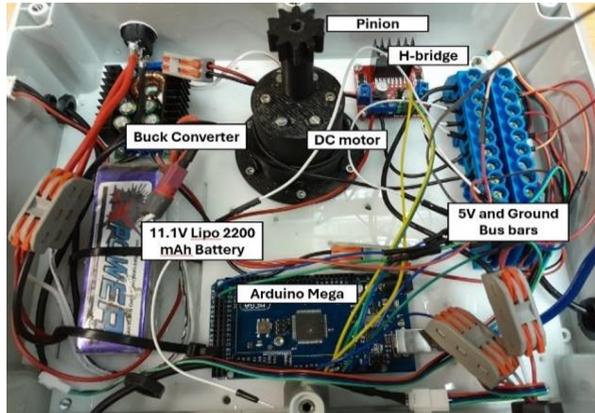


**Fig. 4.** Inner ring and bearing on top of enclosure.

A small axial needle bearing (inner diameter: 120mm, outer diameter: 155mm) is placed between the inner ring and the enclosure lid to facilitate smooth rotation. As the inner ring turns, the outer ring follows due to the magnetic coupling, consequently rotating the wiper to clean the dome's surface.

### 3.5 Electrical elements

Figure 5 depicts the electrical circuitry inside the enclosure box, excluding the LED lights, Raspberry Pi, servo motors, camera module, and ultrasonic sensors. The components include an Arduino Mega, an 11.1V LiPo 2200mAh battery, two bus bars for power distribution, a 12V 0.5A geared DC motor with a torque rating of 10.7kg.cm, a buck converter capable of handling 9A current, and an L298N H-bridge DC motor driver.



**Fig. 5.** Electrical setup inside the enclosure.

### 3.6 Software and integration

The system employs Arduino code to manage components connected to the Arduino Mega, while Python code handles processes for components linked to the Raspberry Pi. The Arduino Mega heavily utilises digital pins to receive and process data from sensors and PWM signals to control the rotational speed of the DC motor driving the cleaning system.

For object detection, a TensorFlow Lite framework with the EfficientDet D0 model architecture is used. The primary model, Aquarium.tflite, is trained on the Aquarium COCO dataset, which consists of 638 images featuring fish, stingrays, jellyfish, penguins, sharks, puffins, and starfish. The chosen model and dataset provided a functional baseline for object detection within the proof of concept, which allowed for the testing of basic object detection functionality.

The Python code on the Raspberry Pi manages servo motor control based on detected objects in the camera's view. When an object of interest, such as a fish, is detected, the system draws a bounding box around it. The Python code instructs the servo motors to adjust the camera's position, centring the object within the frame. Once the object is near the centre, the Raspberry Pi begins saving pictures.

While this tracking and re-identification feature is important for maintaining object visibility in the camera's view, detailed performance metrics for the tracking functionality were not a primary focus of this system's testing phase. The initial scope of testing centred on object detection performance rather than tracking efficiency. Future testing could include a more thorough evaluation of the tracking algorithm's performance, particularly in dynamic environments or with multiple objects.

Serial communication facilitates communication between the Raspberry Pi 4B and Arduino Mega. When the sensors connected to the Arduino detect an object, the Arduino notifies the Raspberry Pi, prompting the activation of the LED lights and camera. When the camera and LED lights are off, the Raspberry Pi focuses on processing distance values received from the sensors via the Arduino Mega, which also controls the programmable LED lights.

## 4 Testing and results

Four tests were conducted to evaluate the system's performance and capabilities, focusing on essential functions and goals. The first test assessed the calibration of the three ultrasonic

sensors to ensure they were functioning correctly and providing reliable distance measurements. The system was positioned perpendicular to a ruler, with sound waves reflecting off a straight-walled box. This test confirmed that the sensors were accurate at shorter distances, validating their use in triggering system components based on proximity. While coverage across the entire beam pattern was not evaluated, the sensors were deemed reliable for the prototype's functional goals.

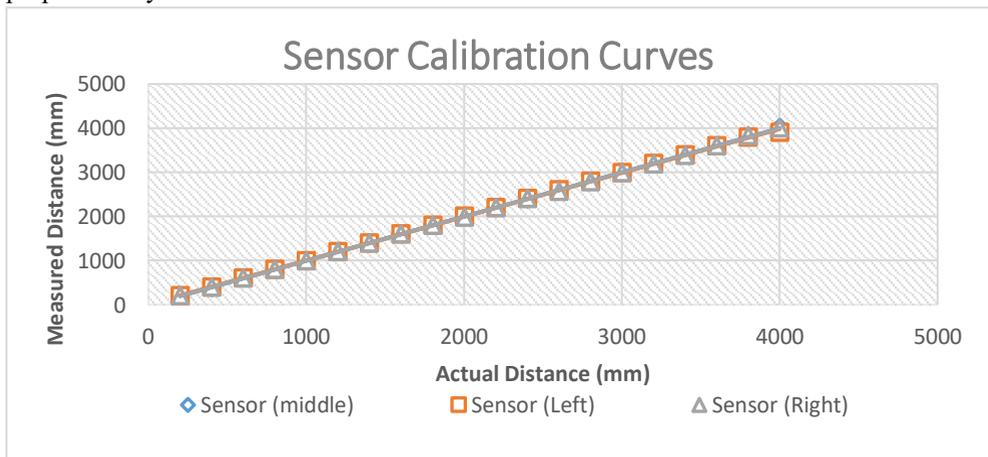
The second test evaluated the system's detection consistency at various distances, rather than traditional object detection accuracy metrics like IoU, precision, recall, or F1-score. A printed photograph of a coelacanth, sized similarly to the smallest known coelacanth, was pasted on a white wall. The system was placed at different distances from the photo, and detection consistency was measured by calculating the percentage of frames (out of 100) in which the system successfully detected or identified the coelacanth within a 10-second time frame at each distance. This test focused on the system's ability to consistently detect the object across various ranges.

The third test investigated whether the object detection model used (EfficientDet D0), which was trained on the Aquarium COCO dataset found on Roboflow, could mistakenly identify a non-fish creature as a fish. This false positive test provides an important intermediate result in evaluating the reliability and precision of the detection model in differentiating between target and non-target objects.

The fourth test measured the system's power consumption under different operating conditions. Using a multimeter, the current draw was recorded when the camera and LED lights were both on and off. This test aimed to evaluate the power efficiency of the system by comparing the energy usage of the camera and lights to that of the low-power ultrasonic sensors. The results help determine whether relying on the sensors as the primary always-on component leads to significant power savings, validating the system's energy optimisation strategy for prolonged deployments.

#### 4.1 Ultrasonic sensor calibration test

Figure 6 presents the results of the sensor calibration. All three sensors demonstrated good accuracy at shorter distances, with sensed distances closely matching the actual distances. The data from the middle, left, and right sensors were plotted on the same axis to highlight their overall alignment and consistency. This was done to avoid unnecessary duplication of graphs and to show that, for most distances, the sensors tracked real distances reliably and proportionally.



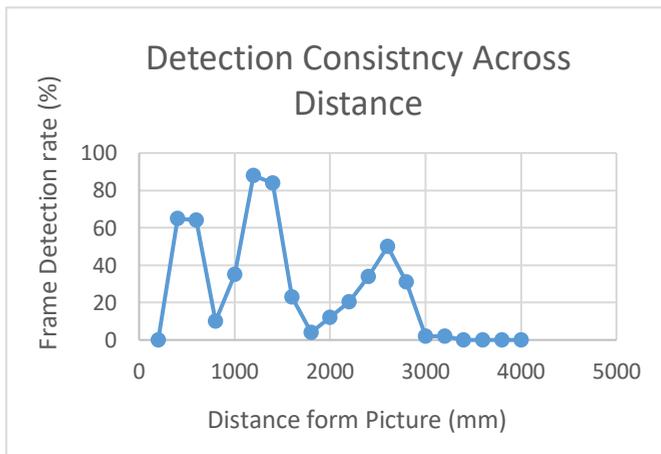
**Fig. 6.** Sensor calibration curves.

However, as the distance increased, very slight discrepancies between the actual and sensed distances became slightly apparent, particularly beyond 2600mm. At distances like 2600mm, 3800mm, and 4000mm, larger variances were observed across the sensors, which may have resulted from environmental factors or minor measurement errors during testing. These outliers were consistent across repeated tests, suggesting that while the sensors performed well, accuracy decreased at greater distances. The right sensor, for example, maintained good consistency up to 3800mm, beyond which variances increased, especially at 4000mm where the standard deviation was the highest.

Despite these outliers, the overall performance of the sensors remained dependable, particularly in controlled conditions and at shorter distances, validating their functionality for the intended proximity-based activation of the system's components.

## 4.2 Evaluation of detection consistency across distance

This test evaluated the system's detection consistency at various distances from a coelacanth image, focusing on practical system responsiveness rather than standard object detection accuracy metrics such as Intersection over Union (IoU), precision, or recall. For this proof-of-concept system, detection consistency—measured as the percentage of frames successfully detecting the object over time—was prioritized as it reflects the system's ability to maintain continuous detection across varying conditions. This approach is particularly relevant for real-world scenarios where stable, continuous detection is more critical than raw accuracy alone. However, in future work, conventional object detection metrics such as IoU and precision could be employed to further validate and benchmark the system's detection capabilities against standard models.



**Fig. 7.** Detection consistency across distance.

As illustrated in Figure 7, the system exhibited optimal detection consistency of 88% when positioned 1200mm from the image. However, both closer and farther distances saw significant declines in consistency, with detection dropping to 0% at 200mm and beyond 3400mm. These results suggest that the system performs most effectively when the object in view closely resembles the size and characteristics of the fish images in the training dataset.

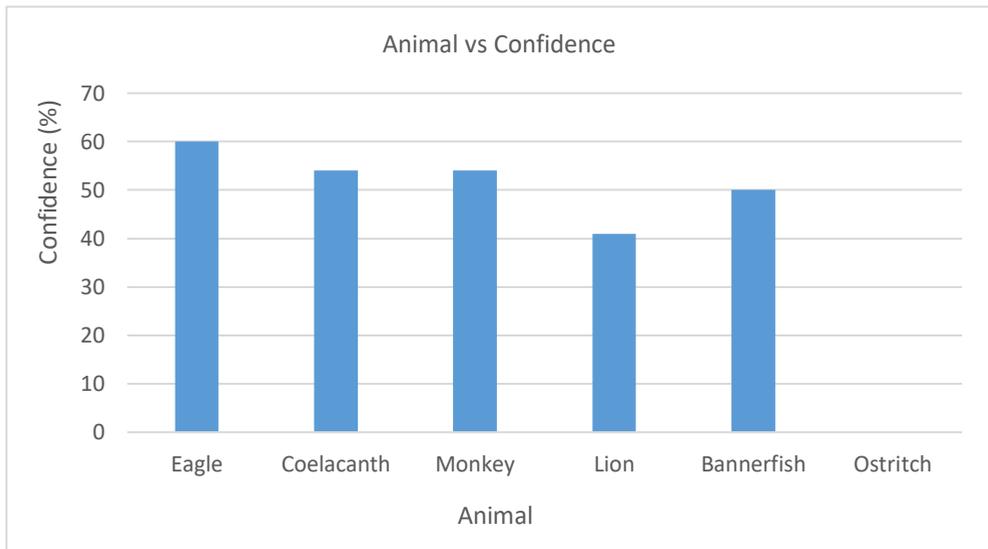
The detection consistency graph deviates from an exponential decay pattern, instead peaking at approximately 1200mm. This optimal performance distance aligns with where the image size and appearance most closely match the examples in the training dataset. At closer distances, such as 200mm, the object may occupy too much of the frame, resulting in poor

detection performance. Conversely, at greater distances beyond 2600mm, the object becomes too small or unclear for reliable recognition.

Interestingly, at intermediate distances where the detection rate is low, the system struggles because the image proportions and details don't correspond well to those it was trained on. This observation, along with the performance decline at extreme distances, indicates that the system's ability to generalize across various object scales and perspectives is limited.

### 4.3 Object detection confidence test

This test evaluated the system's ability to differentiate between fish and non-fish animals. The system, running an EfficientDet-D0 object detection model, was coded to detect only the **fish** class from the Aquarium dataset, which included jellyfish, penguins, sharks, puffins, stingrays, starfish, and fish. For this test, printed images of six animals—an eagle, a coelacanth, a monkey, a lion, a bannerfish, and an ostrich—were used.



**Fig. 8.** Object detection animal vs confidence test.

Confidence values reflect how certain the system is that a detected object matches its training. Ideally, the confidence values for the coelacanth and bannerfish (both fish) should be high, while non-fish animals should result in low or zero confidence.

The system correctly identified the two fish (coelacanth and bannerfish) and excluded the ostrich with zero confidence, meaning it was correctly recognised as a non-fish object. However, the system incorrectly identified the eagle, monkey, and lion as fish, with relatively high confidence values of 60%, 54%, and 41%, respectively. These misclassifications indicate that the model struggled to differentiate between objects with similar shapes, likely due to the limited size and diversity of the fish dataset used for training.

While the system performed well with the actual fish images and excluded the ostrich, the high confidence scores for non-fish animals suggest that the model was overly reliant on shape similarities when classifying objects. This issue highlights the limitations of the dataset, which did not provide enough diverse examples for the model to learn to distinguish between fish and non-fish objects effectively. To improve the system's accuracy, a larger and

more varied dataset would be necessary, allowing the model to better generalize across different species and object types.

#### 4.4 Current draw test

The purpose of this test was to validate the use of ultrasonic sensors to assist in power conservation within the system. Since some form of sensing must always remain active for the system to be ready to photograph nearby objects once deployed, optimising power consumption is crucial. This test aimed to determine whether using the ultrasonic sensors as the primary always-on component would result in significant power savings compared to always having the camera and lights active. Table 1 presents the current draw of each electrical component in the system

**Table 1.** Current draw for certain system components.

Component	Current draw (A)
Arduino mega	0.038
H-bridge – with DC motor off	0.051
H-bridge – with DC motor on	0.09
3 x ultrasonic sensors	0.021
LED lights (Max Brightness)	1.224
Raspberry pi – camera off	0.423
Raspberry pi – camera on	0.603

Table 2 shows the overall current draw of the system when the camera and lights are both on and off. The measurements were taken using a multimeter, with the DC motor turned off during the readings.

**Table 2.** System current draw.

System state	Current draw (A)
Using camera and lights	1.889
Using ultrasonic sensors	0.665

Referring to Table 2, when the camera and lights are always on, the system draws 1.889A of current, which is significantly higher than the overall current draw of 0.665A when the camera and LED lights are turned off. This demonstrates the significant power savings achieved by using ultrasonic sensors to detect nearby objects, allowing the system to keep the camera and lights dormant when they are not needed.

By maintaining a lower current draw during periods of inactivity, the ultrasonic sensors contribute to more efficient power management, extending the system's operational life in energy-constrained environments. This test was crucial in confirming the effectiveness of this approach, as lighting for vision is often a power-consuming aspect of such systems. Validating the sensors' role in reducing unnecessary power usage emphasizes the importance of optimising components to achieve long-term energy efficiency in practical applications.

It is important to note that while the detection consistency test provides meaningful insights into system responsiveness, it is not intended to replace standard object detection metrics such as IoU, precision, or recall. These conventional metrics are more suited for evaluating object detection algorithms' accuracy on large, varied datasets. In contrast, this test was designed to measure the system's real-time responsiveness under controlled conditions, reflecting the system's practical use case rather than its benchmark performance against academic standards. Future tests can incorporate traditional object detection metrics to more thoroughly evaluate the model's accuracy, precision, and recall across a variety of datasets.

## 5 Recommendations for future improvements

Future iterations of the system could significantly improve by expanding the training dataset to include a broader range of fish species and non-fish animals. The current dataset, though containing multiple classes of marine life, was limited to detecting only fish, which likely contributed to the system's difficulty in distinguishing between fish and other animals. For instance, while the system correctly identified two fish (coelacanth and bannerfish) and excluded the ostrich, it incorrectly classified the eagle, monkey, and lion as fish with relatively high confidence. Expanding the dataset to include more diverse examples of both fish and non-fish objects would help the model learn more discriminative features, improving its classification accuracy.

Additionally, exploring more advanced object detection models, such as YOLO-Fish-1 and YOLO-Fish-2, could enhance the system's ability to detect fish in dynamic environments. A larger dataset, combined with more specialised models, would enable better generalization and detection performance, particularly in real-world conditions where object shapes and sizes vary.

## 6 Conclusion

This proof-of-concept system, which integrates ultrasonic sensors for power conservation and a pan-tilt mechanism for autonomous object detection, demonstrates the potential for developing intelligent systems for underwater research. The system was coded to detect only fish from the Aquarium dataset, which contained a limited number of relevant training examples. This restriction likely contributed to the system's performance, where it correctly identified two fish (coelacanth and bannerfish) and excluded the ostrich as a non-fish object. However, it misclassified the eagle, monkey, and lion as fish with relatively high confidence values, indicating that the model struggled to differentiate between similar shapes. These results point to the need for a more comprehensive and diverse dataset to improve classification accuracy.

In conclusion, this system represents a significant step toward developing autonomous underwater imaging technology. The lessons learned from this study provide a strong foundation for future iterations, with improvements needed in dataset diversity, detection algorithms, and hardware upgrades. Expanding the dataset and leveraging more advanced object detection models will likely result in a more accurate and robust system, contributing to marine research and conservation efforts.

This work was financially supported by eNtsa at Nelson Mandela University. The author would like to thank Mr. A. Young for co-supervising this study and Ms. J. Dhlamini and Mr. K Kroutz for general assistance.

## References

1. M.R. Shortis, E.S. Harvey, and J.W. Seager, *A Review of the Status and Trends in Underwater Videometric Measurement*. In Proceedings of the SPIE Conference 6491, Videometrics IX, IS&T/SPIE Electronic Imaging, San Jose, CA, USA, 28 January–1 February (2007). Invited paper.
2. A.D. Rogers, A.S. Brierley, P.L. Croot, M.R. Cunha, R. Danovaro, C. Devery, A.H. Hoel, H.A. Ruhl, P. Sarradin, S. Trevisanut, S. van den Hove, H. Vieira, M. Visbeck, N. Mcdonough, K. Donaldson, K. Larkin, *Delving Deeper: Critical challenges for 21st century deep-sea research*. (2015) doi: 10.13140/RG.2.1.1868.2327
3. D.R. Yoerger, M. Jakuba, A.M. Bradley, and B. Bingham, *Int. J. Robotics Research*, vol. **26**, 1, (2007). doi: <https://doi.org/10.1177/0278364907073773>
4. A. D. Bowen et al., *Field trials of the Nereus hybrid underwater robotic vehicle in the Challenger Deep of the Mariana Trench*, in OCEANS 2009-EUROPE, Biloxi, MS, USA (2009). doi: 10.23919/OCEANS.2009.5422311
5. N. Dalal and B. Triggs, *Histograms of oriented gradients for human detection*, in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, (2005). doi: 10.1109/CVPR.2005.177
6. X. Mei and H. Ling, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. **33**, 11, (2011). doi: 10.1109/TPAMI.2011.66.
7. H. Niu, J. Liu, and Z. Yu, *Real-time object tracking system using PTZ camera*, in IEEE 4th International Conference on Civil Aviation Safety and Information Technology, Dali, China, (2022) doi: 10.1109/ICCASIT55263.2022.9986912
8. L. Zhang, K. Xu, S. Yu, R. Fu, and Y. Xu, *An effective approach to active tracking with a PTZ camera*, in Proceedings of the IEEE 2010 International Conference on Robotics and Biomimetics, IEEE, (2010), pp. 429-434. doi: 10.1109/ROBIO.2010.5723599
9. A.A. Muksit, F. Hasan, M.F.H.B. Emon, M.R. Haque, A.R. Anwary, and S. Shatabda, *YOLO-Fish: A robust fish detection model to detect fish in realistic underwater environment*, *Ecological Informatics*, vol. **72**, (2022). doi: 10.1016/j.ecoinf.2022.101847
10. F.S. Leira, T.A. Johansen, and T.I. Fossen, *Automatic detection, classification and tracking of objects in the ocean surface from UAVs using a thermal camera*, in 2015 IEEE Aerospace Conference, Big Sky, MT, USA, (2015), pp. 1-10. doi: 10.1109/AERO.2015.7119238
11. F. Williams, L.I. Kuncheva, J.J. Rodríguez, and S.L. Hennessey, *Combination of Object Tracking and Object Detection for Animal Recognition*, in IEEE 5th International Conference on Image Processing Applications and Systems, Genova, Italy (2022). doi: 10.1109/IPAS55744.2022.10053017
12. H.K. Chavda and M. Dhamecha, *Moving object tracking using PTZ camera in video surveillance system*, in International Conference on Energy, Communication, Data Analytics and Soft Computing, Chennai, India, (2017), doi: 10.1109/ICECDS.2017.8389917
13. R. Patil, A. Velingkar, M. N. Parmar, S. Khandhar, and B. Prajapati, *Journal of Information and Visualization*, vol **1**, 1, (2020) doi: 10.35877/454RI.jinav155