COTS Control Innovation Program TECHNICAL REPORT

# The COTS Surveillance System (CSS): end-to-end technology for the detection of reef pests

Scott Bainbridge, Mohammad Ali Armin, Geoff Page, Lachlan Tychsen-Smith, Greg Coleman, Zeeshan Hayder, Jeremy Oorloff, De'vereux Harvey, Brendan Do, Ben Marsh, Emma Lawrence and Brano Kusy


Australian Government
REEF TRUST
Great Barrier Reef Foundation
Australian Institute of Marine Science
CSIRO
James Cook University Australia
The University of Queensland Australia

# The COTS Surveillance System (CSS): end-to-end technology for the detection of reef pests

Scott Bainbridge[1], Mohammad Ali Armin[2], Geoff Page[1], Lachlan Tychsen-Smith[2], Greg Coleman[1], Zeeshan Hayder[2], Jeremy Oorloff[2], De'vereux Harvey[1], Brendan Do[2], Ben Marsh[1], Emma Lawrence[2] and Brano Kusy[2]

1. Australian Institute of Marine Science (AIMS)
2. Commonwealth Scientific and Industrial Research Organisation (CSIRO)


**COTS Control Innovation Program** | A research and development partnership to better predict, detect and respond to crown-of-thorns starfish outbreaks

Australian Government | REEF TRUST
Great Barrier Reef Foundation
Australian Institute of Marine Science
CSIRO
James Cook University Australia
The University of Queensland Australia

## Inquiries should be addressed to:

Scott Bainbridge
Australian Institute of Marine Science
S.Bainbridge@aims.gov.au

## This report should be cited as

Bainbridge S, Armin MA, Page G, Tychsen-Smith L, Coleman G, Oorloff J, Harvey D, Do B, Marsh B, Lawrence E, Kusy B. (2025) The COTS Surveillance System (CSS): end-to-end technology for the detection of reef pests. A report to the Australian Government by the COTS Control Innovation Program (94 pp).

## Funding Acknowledgement

The COTS Control Innovation Program aims to accelerate the development of innovative surveillance and control methods to manage outbreaks of coral-eating starfish on the Great Barrier Reef. The Program is a collaboration between the Great Barrier Reef Foundation, Australian Institute of Marine Science, Commonwealth Scientific and Industrial Research Organisation, James Cook University and The University of Queensland. The Program is funded by the partnership between the Australian Government's Reef Trust and the Great Barrier Reef Foundation.

## Traditional Owner Acknowledgement

The COTS Control Innovation Program extends its deepest respect and recognition to all Traditional Owners of the Great Barrier Reef and its Catchments, as First Nations Peoples holding the hopes, dreams, traditions and cultures of the Reef.

## Disclaimer



## Use of Trademarks and Company Names



## Cover Image

ReefScan-Deep platform undergoing testing, Davies Reef. Credit: D. Harvey, AIMS 2024.

# Contents

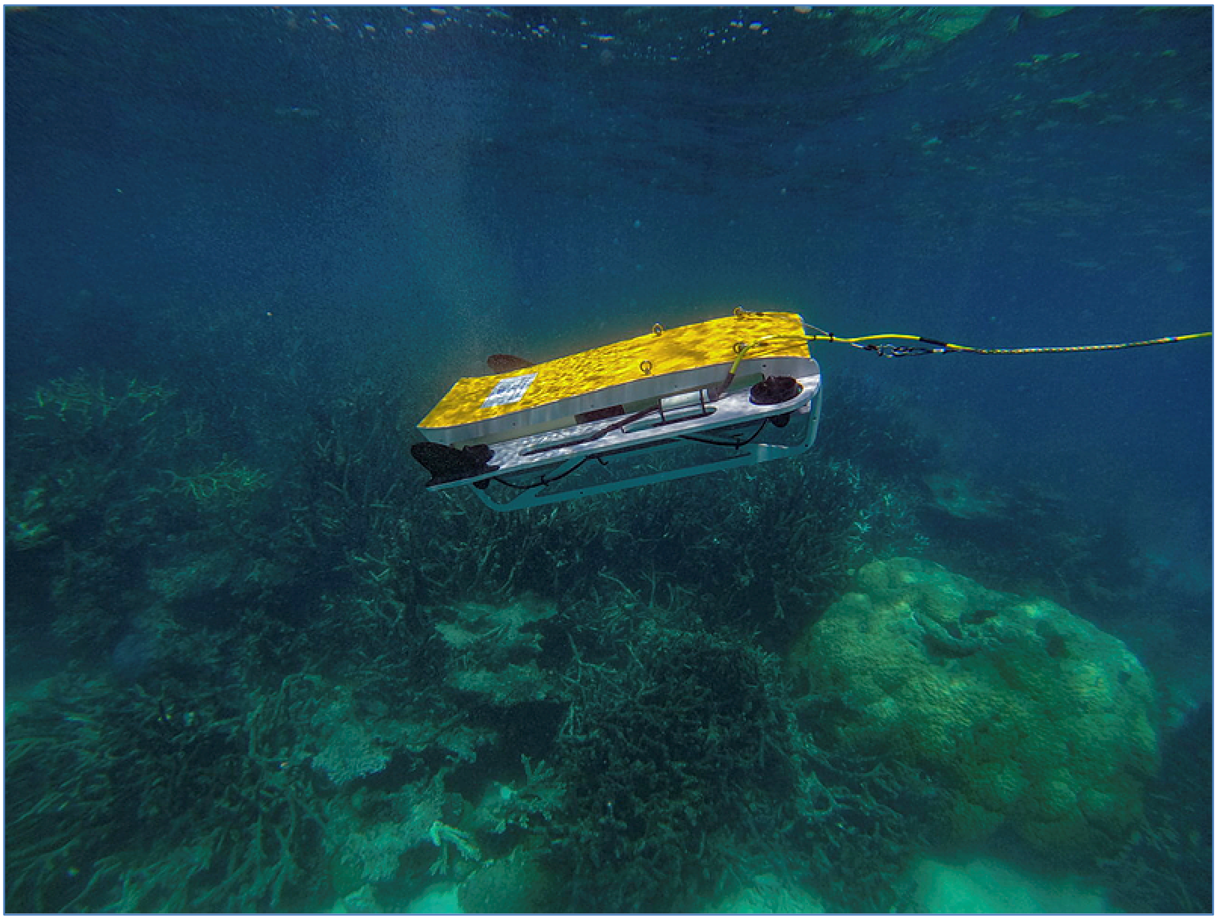

Figure 1. ReefScan-Deep platform being tested (Protype-I) (credit: M. Roman, AIMS 2023)

# List of Figures

# List of Tables



Figure 2. ReefScan undergoing testing (Operational Platforms) (S. Bainbridge, AIMS 2025)

# Acronyms and Abbreviations

| | |
|---|---|
| AI | Artificial Intelligence - computer systems able to perform tasks normally requiring human intelligence |
| AIMS | Australian Institute of Marine Science (www.aims.gov.au) |
| bbox | Bounding Box - a rectangle in image coordinates that encompasses the area of interest or detected feature. |
| CCIP | Crown-of-thorns starfish Control Innovation Program |
| COTS | Crown-of-thorns starfish |
| CPUE | Catch per Unit Effort – an indirect measure of the abundance of a target species |
| CSIRO | Commonwealth Scientific and Industrial Research Organisation (www.csiro.au) |
| CSV | Comma Separated Values – file format for data |
| CVAT | Computer Vision Annotation Tool – open-source tool for annotating images to train Machine Learning models: [https://www.cvat.ai/](https://www.cvat.ai/) |
| EXIF | Exchangeable Image File Format – format for storing image metadata |
| FPS / fps | Frames per Second, the number of sequential images collected per second |
| GBR | Great Barrier Reef – coral reef system located off the north-east Australian coast |
| GBRMPA | Great Barrier Reef Marine Park Authority (now the Reef Authority) |
| GCS | Ground Control Station |
| GPS | Global Positioning System – satellite-based positioning system |
| KML | Keyhole Markup Language - a file format used to display geographic data |
| LIDAR | Light Detection and Ranging - a sensing method that uses light in the form of a pulsed laser to measure distances (ranges) |
| ML | Machine Learning, a form of Artificial Intelligence |
| MQTT | Message Queuing Telemetry Transport – a lightweight messaging protocol |
| QC | Quality Control, the process of checking the quality of the collected data |
| QPWS | Queensland Parks and Wildlife Service |
| ROS | Robot Operating System - an open-source operating system used for robotics |
| ROV | Remotely Operated Vehicle |
| SBAS | Satellite Based Augmentation System – a system to increase GPS accuracy |
| USB | Universal Serial Bus - a form of computer data connection |
| UTC | Coordinated Universal Time - the primary time standard globally used to regulate clocks and time. Queensland time is UTC time + 10 hours. |
| VDSL | Very High-speed Digital Subscriber Line - a type of communication technology that uses copper lines to transmit data. |

# EXECUTIVE SUMMARY


A fundamental part of controlling Crown-of-Thorns Starfish (COTS) is first being able to find them. Detection of starfish is therefore a fundamental component of starfish control and so increasing the effectiveness of detection leads to a direct increase in the effectiveness of the control work. The CCIP-D-04 Project, *The COTS Surveillance System (CSS): end-to-end technology for the detection of reef pests*, was funded to develop and apply new technologies, as end-to-end solutions, to detect starfish while simultaneously providing high resolution estimates of benthic cover including percent hard coral cover.

The project goal was to utilise advances in technology to develop and deliver an end-to-end solution that removed or reduced the known issues with the current methods and which opened up new approaches to the detection of starfish. The work needed to deliver a solution that at least matched the current manual methods in terms of area of reef surveyed, number of starfish detected, as well as being delivered within the current resource and operational envelopes. Where possible, the system should enable new approaches to the overall control work, deliver new capacity to the CCIP project and the COTS community, and empower new partners in the COTS control space.

The current primary method for detecting COTS uses human based manta tow surveys (Matthews et al, 2024) which have several limitations that reduce the effectiveness of the detection component, and in turn, of the control work. These include the need to extensively train observers to remove observer bias and so deliver consistent results, the inability to record detailed information about the benthic cover (such as being able to estimate change in coral cover), and the need to have people in the water with potential marine pests.

The CSS looked to combine cutting edge Artificial Intelligence and marine imaging platforms to deliver a step change, both in the detection of COTS, but also as a new capacity for COTS and reef surveillance and monitoring.

The developed system was delivered via three components: a new underwater towed imaging platform, a set of real time and delayed mode Machine Learning models for COTS, COTS Scars and Benthic Habitat, and a series of workflows that join these together as an end-to-end solution.

The platform consists of a small (1.1 m (l) x 60 cm (w) x 50 cm (h)) sled that is towed behind the survey vessel on a combined data and towing tether, and which uses battery powered thrusters to maintain a set attitude (position in the water) and depth (typically 3-5 m off the bottom). The platform has two high resolution cameras that take images of the bottom over a 10 m wide swath as well as a forward-facing navigation video camera. The platform is controlled from the surface over the tether using a small Ground Control Station, much like an aerial drone control system, with the operator setting the required height above the bottom and adjusting this as needed to match the topography of the reef being surveyed.

The platform cameras image the reef, surveying a 10m wide swath, as it is towed around the perimeter of the reef. The images are fed in real time to an on-board computer running a Machine Learning model that analyses each image for starfish, the scars they leave and the benthic habitat the platform has passed over. The real time results are displayed on a field tablet giving instant feedback to the field team about the number and location of starfish.

The platform was developed over two prototypes into a final Operational Platform. Field trials were undertaken with the Control Teams with the platform being deployed in parallel with the existing human-based survey methods. The platform development included adding a second camera to increase the field of view or surveyed swath, refinements to increase the strength and robustness of the system and re-engineering to make the system easier to use, especially, making it lighter and easier to deploy. A final engineering review was done in October 2024 with the various stakeholders, including the Control Team representatives, to drive the final platform design and delivery.

Machine Learning models were developed for COTS (CSIRO) and Benthic Life Form (AIMS). It was found that, at low COTS densities, COTS scars were a better diagnostic of COTS presence than COTS themselves. As a result, the COTS Model was extended to include Scars. As there are many causes of coral scars, including COTS, *Drupella* and coral bleaching, the model developed for Scars did not perform at a level considered to be fully operational and so is considered to be developmental at this stage. Collecting and training the model on more images of differing types of scars will resolve this.

The final models gave an accuracy score of 87% for detecting COTS, 91% for tracking COTS from image to image, 40% for detecting Scars and 73% for mapping Hard Corals (where an accuracy of 80% or better is considered to be operational). The accuracy scores improved as the new images of starfish and scars were collected and used to train the models, with the COTS model going from an initial accuracy of 64% to a final accuracy of 87%.

Workflows were developed, as software routines, that moved the collected data through a series of processing steps to deliver a set of data products that drive actionable outcomes. As the images are collected, a real-time workflow uses Machine Learning models to give immediate feedback about the reef just surveyed with the field operator notified of any starfish detected. This allows the field personnel to respond such as by doing a quick snorkel survey to confirm the results or to call in the Control Team to action that part of the reef.

End of day workflows use the greater processing of an on-board laptop to do more detailed analysis with the goal of informing the next day’s work (such as stay at this reef or move onto the next one). End of trip workflows push the collected images and extracted data to the AIMS ReefCloud system where it is analysed for benthic cover and then pushed to institutional and project data systems.

Four trips were undertaken with the COTS Control Teams and another four trips with the Queensland Parks and Wildlife Service (QPWS) to collect images to train the models as well as get feedback on the use and performance of the platform. A total of over 640,000 mages were collected representing some 71 surveys and 50 hours of towing.

The project successfully developed and delivered three complete systems including hardware, software, Machine Learning Models and training material to the Reef Authority in mid-2025. The platforms provide a robust easy to use method to collect geo-tagged high-resolution images of the reef while the Machine Learning models give operational level detections of COTS (87%), Scars (40%) and corals (73%) within workflows that deliver actionable outcomes including in real-time.

The project encountered challenges in several key areas, specifically:

- The Scars model currently performs below what is considered to be an operational level (40% accuracy versus 80% for a typical operational model) and does not track scars over frames, so does not provide counts. This is due to the difficulty in distinguishing between COTS related scars and other sources of scars or white coral, something that divers find difficult. This can be addressed through training the model with more images.

- The level of platform autonomy, especially for collision detection and avoidance, was not developed to the level that was anticipated as this proved to be more complex than initially thought (reefs are complex topographically). The system operator still needs to continuously monitor the platform where it was hoped that the operator would only need to deal with very sudden changes in reef topography.

- Technology transfer to the Control Teams fell short of expectations, as we had anticipated conducting more extensive field-based implementation with the operational teams. This was due to delays in developing the system and the general logistical difficulties in getting people and equipment in the field.

Even with these issues, the project delivery new capability to the Control Teams and the CCIP project with work planned for the systems to be utilised in a range of upcoming detection activities. Importantly, the project provides a pathway for future application of emerging technologies as well as a series of partnerships, including organisational partnerships, that empowers future work in this space.

The CSS represents a new tool for detecting COTS that has application within the current control work but also as a new capacity within the COTS community. The project successfully integrated a set of novel technologies into an end-to-end system that has application across a wide range of ecological areas including understanding deep water COTS, supporting surveys of areas currently not visited, improving our understanding of initiation zones, and the response of COTS to coral bleaching.

The CSS delivers to a larger vision of reducing observer bias from manta tow surveys (e.g., due to different level of training and field experience) and improving resolution and accuracy of survey data based on an image-based data workflow. Once fully validated and calibrated, the platform will provide data with equivalent value in making and supporting management decisions, irrespective of who operated the platform. This changes the logistical models that can be used to implement control and monitoring work and opens new pathways to increase effectiveness and to scale the work currently being done.

# 1. INTRODUCTION

## 1.1 Introduction

The Crown of Thorns Starfish (COTS, *Acanthaster* cf. *solaris*) is a naturally occurring predator of coral on the Great Barrier Reef (GBR) and across the Indo-Pacific. While a natural part of these systems, the starfish can undergo dramatic surges in population numbers, termed outbreaks, that can cause large loss of coral (Babcock et al 2016). Various studies (e.g. Bozec et al 2022, Vercelloni et al 2017, De'ath et al 2012) found that COTS accounted for a significant loss of coral cover over the extent of the Great Barrier Reef since monitoring began in the mid 1980's, with COTS outbreaks seen as a major stressor of coral reefs (Emslie et al 2024).

In 2012, a COTS Control Program was established on the GBR to action control as a response to reduce the impact of COTS outbreaks. An Integrated Pest Management (IPM) framework was subsequently adopted (Westcott et al 2016, Fletcher et al 2020) and a number of Control Teams were funded to remove COTS at key sites, to both reduce the immediate impact on coral at those sites, and to reduce the spread of COTS from reef to reef by reducing the number of adults available for spawning.

The Control Teams undertake initial surveys using the manta tow method (Miller et al 2018) to locate areas of the reef where COTS are present (above an ecological threshold, typically one starfish seen per two-minute manta tow) and then use follow up dive teams to manually cull the COTS using an injection of bile salts or vinegar (Reef Authority 2023a).

The manta tow detection component of the work, while only typically taking up 10% of field time (Matthews et al 2024), is critical in being able to deploy the control measures to the correct place at the correct time. Being able to accurately locate large aggregations of starfish that are amenable to culling is core to the effectiveness and efficiency of the Control Teams work. Simply put: getting the 10% right ensures the best use of the remaining 90%.

Detecting starfish is therefore a critical component of the control work and one that has the potential to be a limiter on the effectiveness of the work. Not being able to find starfish limits the ability to control them in the same way that missing or not detecting a large outbreak can result in future loss of coral, and so mitigate the overall effectiveness of the control program.

The manta tow method has a long history of use for monitoring coral cover and for detecting COTS (Miller et al 2018). While it remains a widely used method, as one of the few large-scale rapid-survey techniques available, it has a number of drawbacks. Studies, such as Fernandes et al (1990), show that typically a manta tow only picks up around 5% of the COTS identified using other diver-based measures, while Lawrence et al. (2024b) showed that manta tow derived coral-cover estimates showed a large degree of variability and error when compared to other methods of estimating coral cover.

As a result, manta tow is a data collection method that needs significant investment in training and standardisation to deliver consistent results between observers. This poses a challenge for the COTS Control Program, which relies on a large number of people from differing backgrounds, across different vessels, collecting reliable consistent data.

## 1.2 Project Goals

The COTS Surveillance System (CSS) project was initiated to research and develop the application of new technologies to expand the COTS Control surveillance and detection toolbox beyond the manta tow method, offering a new detection and monitoring tool that is more efficient (faster, fewer people), safer (less people in the water), more flexible and which creates a multi-use dataset. The objective was to develop a solution that could deliver consistent results across users and areas so that the methodology and system can be used both to scale up control-based detection work and to support other monitoring and detection applications such as long-term monitoring of COTS and coral.

The specific goals of the project were:

- Identify a set of technologies to enable automatic Artificial Intelligence based identification of COTS in collected benthic images as a pathway to increase the effectiveness of the detection component of the COTS control work.
- Design an end-to-end system to deliver a solution that would meet the identified set of needs as determined by the end users (Control Teams) using technology that was mature in other domains even if novel in the marine domain.
- Development of a novel underwater towed platform that could house high resolution cameras to image the benthos of the reef at a scale equivalent to what a human on a manta tow would experience.
- Develop, train and tune a series of Machine Learning models for COTS, Scars and Benthic life forms to enable real time or equivalent identification of COTS and Scars in the collected images and to characterise the benthos surveyed.
- Develop a set of workflows, as software, that extracted actionable information from the collected data to support on-water (real-time), in-field (next-day) and in-office (end of trip) decision making and response.
- Undertake technology transfer of the project outcomes with the Control Teams to ensure that the systems work within the required operational constraints, add value to the work of the field teams, and are designed to meet the specific needs of the users including engineering certification, manuals and training material.

The project looked to bring together components that had been proven in either other projects or in other domains, delivered as a novel solution. The linkage components were the workflows that defined how the systems deliver data to specific users at specific times to support the control work. This includes real-time alerts for detections of COTS through to end of day analysis to support the next day's work and end of trip analysis that supports larger scale understandings of the status of COTS and corals over the Great Barrier Reef.

## 1.3 Impact pathway within the CCIP Portfolio

The project sits within the CCIP **Detection** Sub-Program (as D-04) but has direct input into the D-02 *Tool Comparison* and D-01 *Monitoring Design* projects, also in the Detection Sub-Program (see Figure 3 below). While the project delivers to the **Expanded Toolbox** Output and the higher-level **Improved Detection and Monitoring** Outcome, it also contributes directly to the **Operational Response** set of Outcomes by improving the quality and extent of data available for decision making.

The system capability opens opportunities for testing new control and detection methodologies (e.g. night surveys, deep reef surveys, automated culling methods) and so may in the future contribute to the **Prediction** and **Response** Outcomes or their future equivalents.

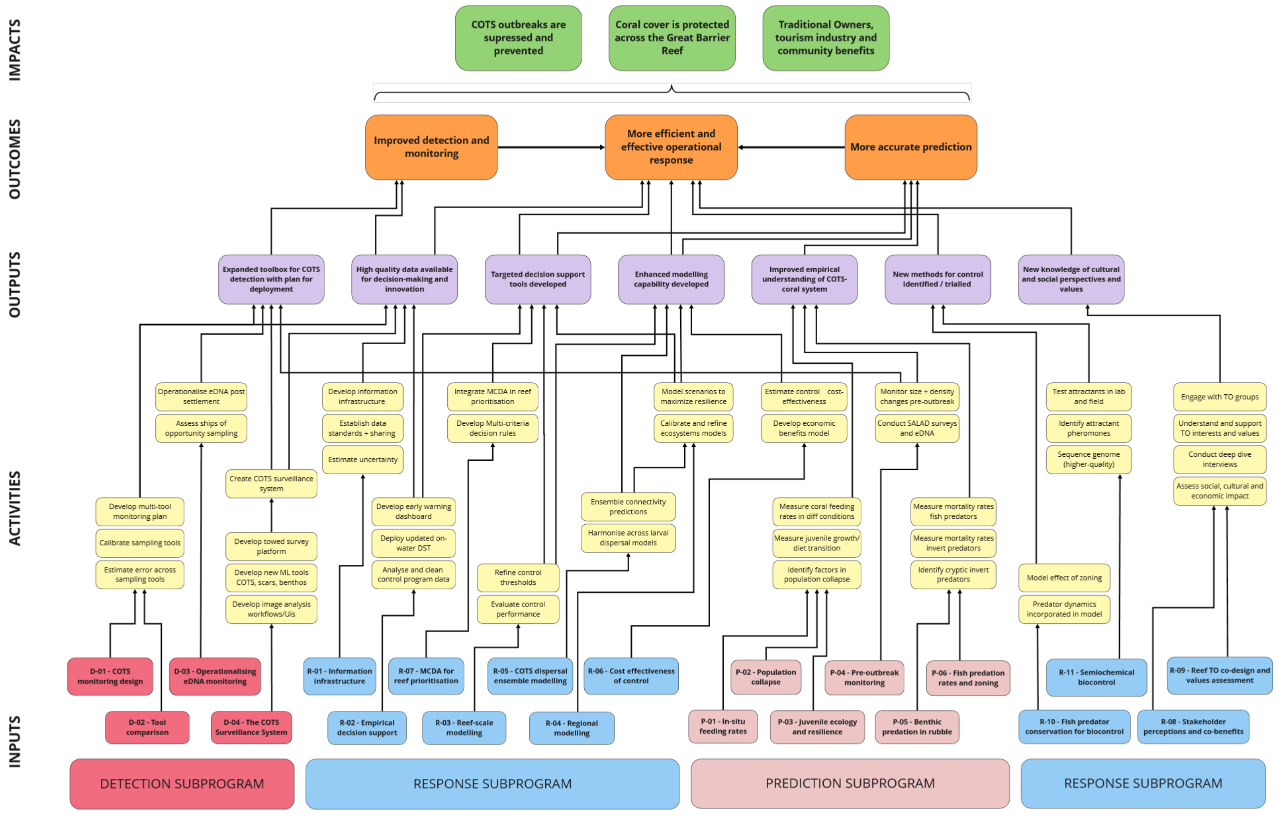


Figure 3. CCIP Project Logic Model.

# 2. METHODS

## 2.1 System Overview

The overall design of the COTS Surveillance System consists of three elements: a novel towed underwater imaging platform (ReefScan-Deep), a series of real time and delayed mode Machine Learning models for the detection of COTS, the scars they leave, and the benthic habitat the platform surveys, and a series of workflows that join these together to deliver an end-to-end solution.

The system is designed to be modular in that differing data capture platforms can be used, new and updated models can be integrated into the workflows, and new workflows can be easily implemented. This allows for future developments and advances to be readily integrated and for a number of partners to work on differing parts of the system without the need to re-engineer the existing components. Where possible open-source or industry standard components have been used, especially ones in use in the robotics and AI space.

The systems were developed to meet the needs of the COTS Control Teams who currently undertake the majority of starfish control efforts along the GBR. Other users, such as Park Rangers, Tourism Operators and Traditional Owners, were also considered with a view that building capability in these sectors may be one path to scale the current control operations.

## 2.2 System Design

### 2.2.1 Platform

The platform was developed based on an existing open-source ROV (the Blue Robotics Blue-ROV) merged with the AIMS ReefScan control architecture, itself based on the open-source ROS environment. Using the existing manta tow method (Miller et al. 2019) as the base use case, an engineering study was done to design a towed platform that had the required stability to allow for the capture of high-resolution images.

The engineering design was developed into the first prototype (Prototype-I) which was field tested on several AIMS trips and on other vessels including two trips with the COTS Control Teams. Several issues were found with the prototype, including issues with water ingress, that ultimately lead to the failure of the system. As a result of the testing, the system was re-designed as a second-generation design in early 2024 (Prototype-II), with the main changes being the use of an all-metal frame to increase robustness, an aluminium electronics housing to reduce the risk of leaks, and twin payload cameras instead of the initial single camera.

The Prototype-II system was tested on several field trips including a further two trips with the COTS Control Team as well as trips with QPWS and on AIMS vessels. No issues were found with the system although Control Team feedback for the Prototype-II system was that making systems lighter and smaller made them easier to deploy and use in the field. An engineering review of the second-generation platform (Prototype-II) was completed in October 2024, with a range of stakeholders, to develop the final engineering design for the delivered system. From this the final system (Operational Platform) was developed.

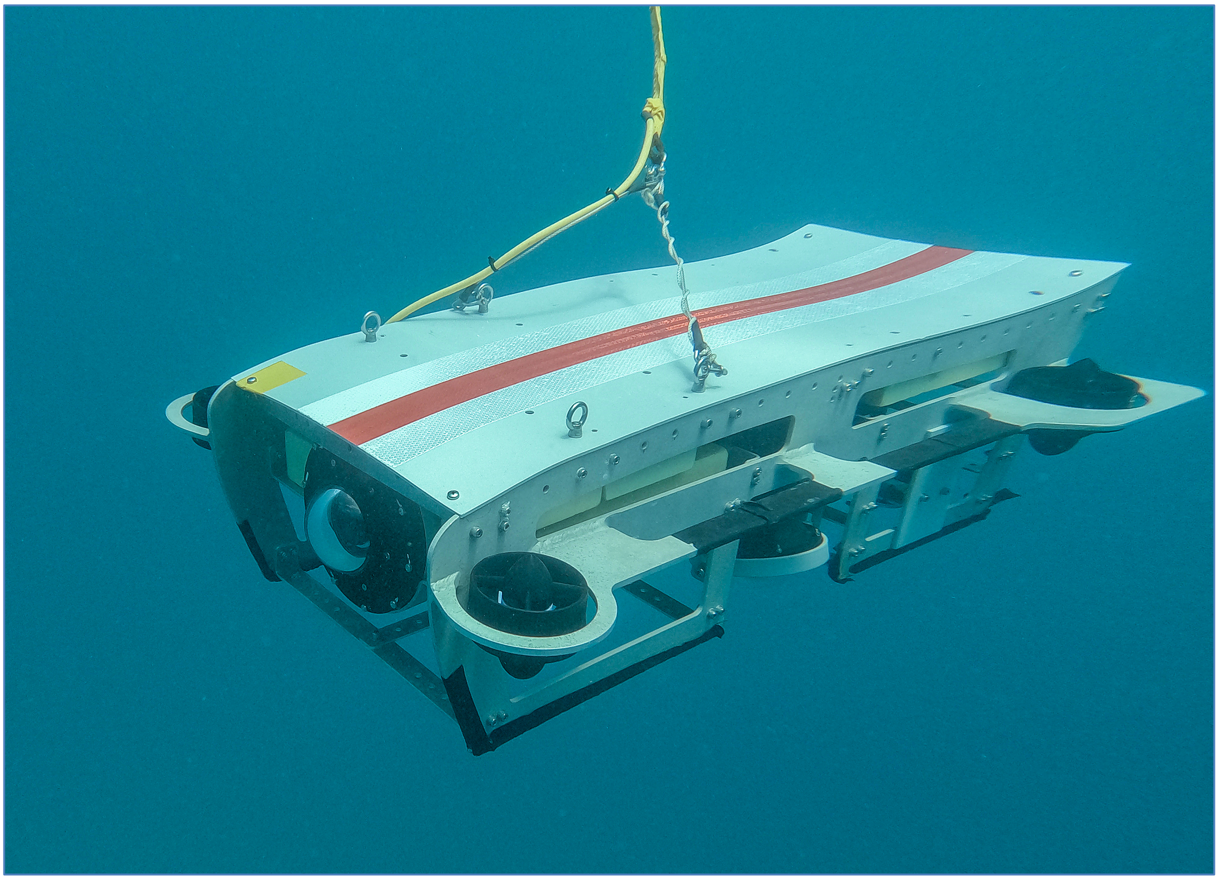

Figure 4. ReefScan-Deep Platform (Prototype-II) in use (credit: D. Harvey, AIMS 2024)

### 2.2.2 Machine Learning Models

A core part of the system was the development and application of a series of Machine Learning (ML) models that enable the system to analyse the collected images to detect COTS and Scars and to estimate the cover of the main benthic forms, such as corals. The models were integrated into the various workflows to give information to the operators to drive responses that optimises the Control work of the Teams.

The initial approach was to develop a model for COTS detection (CSIRO) and a model for benthic classification (AIMS) and to integrate these into the real-time workflow, end of day, and end of trip workflows. The real-time detection would be used to drive on-water responses (such as identifying areas of interest via COTS detection, deciding where and when to deploy Control resources and so on), while the end of day COTS detections and benthic cover estimates would be used to plan and optimise the next day's work (such stay at this reef and continue the Control work or move to the next reef). Finally, the end of trip model outcomes would use increased computing resources, via cloud-based platforms, to do more detailed analysis and feed data into reporting and dashboard systems.

As the project developed it was realised that at low COTS densities the presence of COTS feeding scars is diagnostic for the presence of COTS and so the COTS model was altered to include the detection of COTS Scars (referred to as the Scars model). This gives three models in operation: a combined COTS and Scars model from CSIRO, and the Benthic model from AIMS.

The COTS model implemented tracking to ensure that a detected COTS was not double counted across the series of images that it was identified in, ensuring that multiple occurrences of the starfish were only counted once. This was not done for Scars as the Scars model was only developed late in the project and project resources did not allow for a new tracking algorithm to be implemented for Scars.

The limitations of the computing power that is available in the field for the real-time models meant that the benthic cover model was moved to the end of day workflow (using a provided laptop) as well as the end of trip workflow, via the AIMS ReefCloud platform, where a more detailed analysis of the benthic data was undertaken.

### 2.2.3 Workflows

The system is based on a series of workflows, that is the flow of data from data collection through to final analysis and delivery. The overall system workflow is shown in Figure 5 with more detailed flowcharts for the individual workflows shown in Appendix B.

The workflows are designed to deliver actionable information at various stages of the Control process. The real time workflows give information about the area of reef just surveyed and can be used to identify clusters of COTS and Scars. This allows for diver-based teams to be allocated to investigate further as the detection component of the work is being undertaken.

The end of day workflow is designed to give more detailed spatial representations of the data collected, such as heat-maps and Google Earth overlays, that can inform the next day's work. The goal is to provide an overview of the day's work to allow for the logistics of the subsequent work to be optimised based on the results. This could be decisions to target certain cull zones, to move to the next reef, or to focus on current activities. This feeds into the existing field planning mechanisms used by the Control Teams.

The end of trip workflow uses the AIMS ReefCloud system where the images are uploaded, stored, and a more complex benthic cover model run. The outputs of this model, along with the end of day results for COTS and Scars, are the information that is generated from the workflows. These are designed to feed into existing information systems such as the CCIP Information System and the Reef Knowledge system.

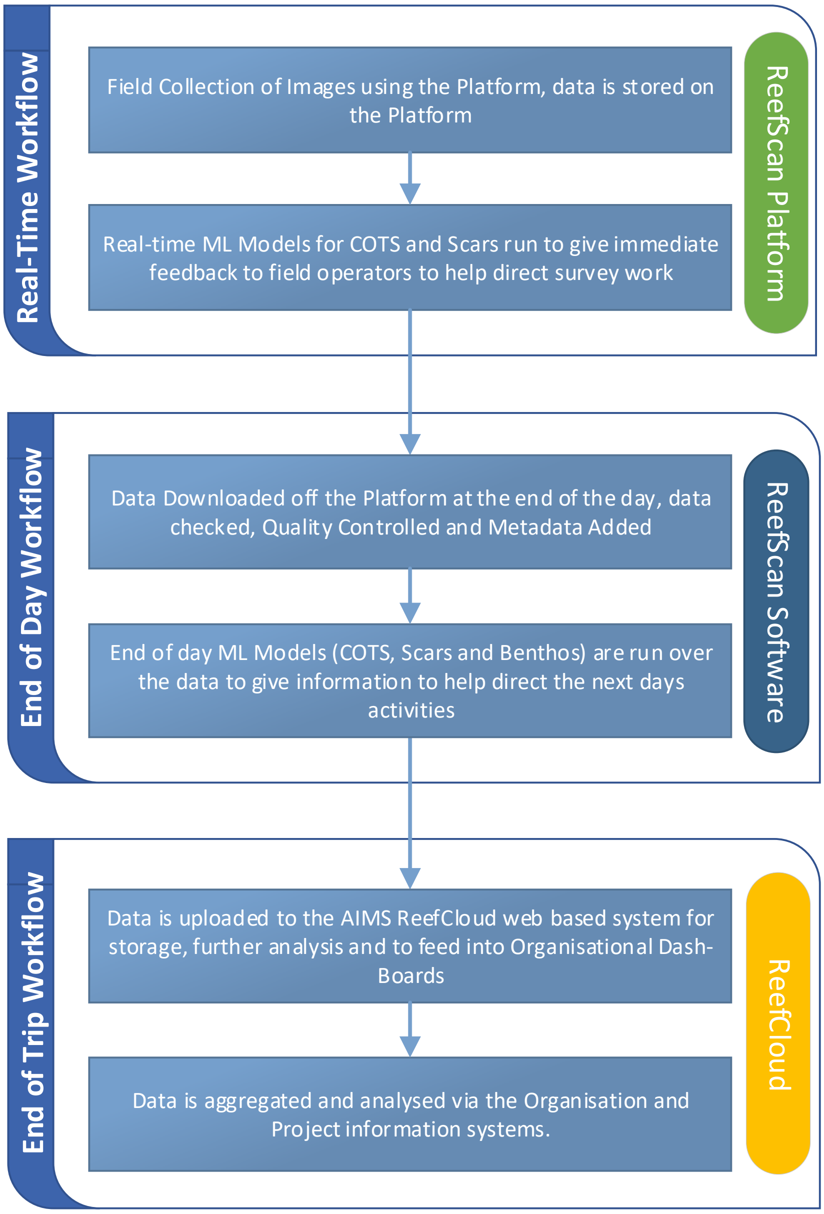


Figure 5. Workflows that underpin the COTS Surveillance System.

## 2.3 Platform Development

### 2.3.1 Platform Design Decisions

Some of the initial design decisions were tested and validated during the platform development. In all three generations were developed: an initial prototype (Prototype-I), a development from this (Prototype-II), and the final delivered platform (Operational Platform).

*Still images versus video*

A decision was made to capture still images over video as these give better resolution (16MP versus 8MP for 4K equivalent video), could be captured as non-compressed images (as TIFF format), and, with a frame rate of 3–4 frames per second (fps), to give complete along-track coverage of the bottom. Testing showed that the better resolution gives more information to the Machine Learning models and so increased accuracy from the model over that derived from reduced resolution of video.

While collecting still images gives better quality data to the ML, the reduced frame rate of 3–4 frames per second (fps) means that object tracking accuracy between images is reduced over what could be achieved in a video of 24–30 fps. This is because the increased frame rate for the video provides more views of the object being tracked and so there is less change from frame to frame to interpret.

The trade-off is between an increased frame rate but reduced image quality or a reduced frame rate but increased image quality. The decision was made to prioritise image quality for this system with the understanding that future variants, with more compute power, will allow for higher frame rates and that accuracy in detecting a starfish is more important than accuracy in tracking it once detected.

*Single camera versus stereo cameras*

The original design had a single camera with a wide field of view lens to capture a swath width of about the same side-to-side distance as the height of the camera above the bottom. This means that in five metres of water, the camera sees about five metres side to side across track. Similarly in 10 m of water, the side to side field of view is about 10 m. Testing of this in comparison to a human undertaking a manta tow survey showed that a single camera system (field of view of 86°) did not match the field of view of the human surveyor (typically 120-140°, Bainbridge and Gardner 2016), especially in areas with coral walls or in complex topography, as the person could move their head side to side to better capture the terrain while the camera was fixed.

Based on this comparison, a second camera was added (cameras are located beside each other, one to port the other to starboard) to give a synced stereo view of the bottom, increasing the field of view side-to-side. The cameras are mounted so that they can be moved outwards from vertical and so can be splayed to further increase the field of view. The movement can be set before the survey, from vertical through to 30 degrees from vertical, allowing the system to increase the swath width surveyed and to better see side features such as bommie walls and reef slopes. This adjustment can be done in the tender before the survey, or the platform retrieved and adjusted as needed.

The decision to move from a single-camera to a two-camera system impacted the detection tracking components. If using data from two side-by-side cameras, the detected starfish need to be tracked not only *along* the track from frame to frame but also *across* track from camera to camera. This significantly complicates the tracking required by the ML models, developing this capability was not possible given the limited project time frame. However, the current two-camera system can be configured with either an active single or dual camera setup, making it flexible to operate using the single-camera ML model developed in this project, while also ensuring the platform can adapt in future as the ML model capability is advanced to support tracking across cameras.

### 2.3.2 Design and Function

The platform (ReefScan-Deep) was designed to be able to collect image data from the reef that gave similar detecting power as that of a person using the current Manta Tow method.

This led to the following design criteria:

- The unit should be designed to operate over reef slopes from around 3 m down to 20 m which encompasses the area typically targeted by the Control Teams (see Matthews et al, 2024).
- The system should have optical characteristics (cameras, lenses, domes, etc) and resulting field of view and resolution to be able to detect COTS based on what a person on Manta Tow could record (as an initial starting point).
- The operational parameters (depth, towing speed, time of day, etc) need to be selected to optimise image quality but again should reflect those of the Manta Tow method.
- The system should be able to operate and be certified down to 50 m to support surveys of deep-water COTS in the future.
- That the unit had to be small and light enough for a single or double person lift and needed to fit in the types of vessels currently used for COTS surveys.
- The system needed to be operational with the current sets of available resources, especially field personnel, and work within existing logistics and constraints.

Given the overall design parameters, the system was specified based on available solutions and previous experience by the project partners. The final system specifications are shown in Table 1.

### 2.3.3 Platform Specification

The platform system consists of three components: the actual platform, which is towed underwater, the surface-vessel based Ground Control Station which used to control or ‘fly’ the underwater platform, and a tether or combined communication and towing cable, that electrically and physically joins the two together. This is shown in Figure 6.

The tether is physically connected to the underwater platform via a stainless-steel ‘sock’ on the tether and a bridle on the platform. At the vessel end, it is attached to a bungy cord that is in turn attached to the vessel via a cleat or similar. The bungy smooths out the towing force on the platform to allow it to tow smoothly. The tether length can be adjusted by simply paying out more tether from the tow vessel with the principle that the tether length should be two to three times the depth of the towed platform. If the platform is at 6m of depth, then the tether should be adjusted to give 12-18m of length between the towing vessel and the underwater platform. The tether has marks at every 5 m to aid setting the tether length.

The tether uses waterproof connectors to provide a data connection between the underwater platform and the Ground Control Station allowing the operator to control the platform while the underwater platform images and navigation camera are visible to the surface-based operator.

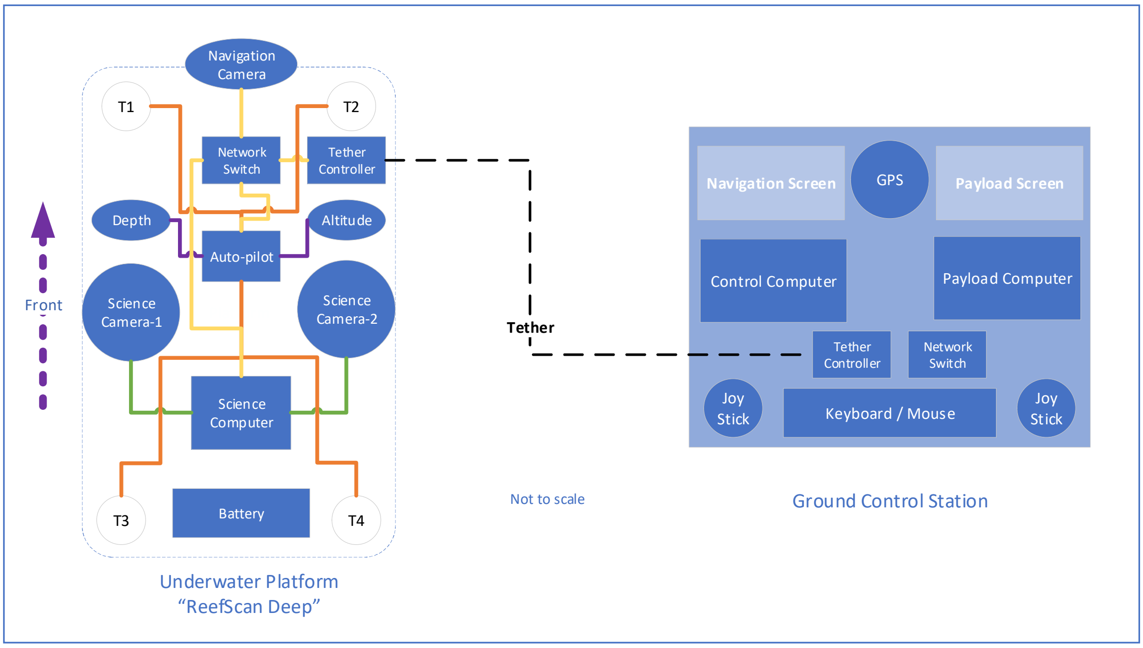


Figure 6. Schematic of the ReefScan-Deep platform.

Table 1. Platform specifications

| Parameter | Value |
|---|---|
| Type | Towed underwater imaging platform (“ReefScan-Deep”) |
| Tow Speed | 2-3 knots |
| Weight | Approximately 30 kg (two person lift) |
| Tether: | 20-50 m copper / Kevlar tether, length being 2:1 to 3:1 of depth (e.g. for running the platform at 8 m depth around 16-20 m of deployed tether is required) |
| Deployment Vessel: | Small (4-6 m) RIB or equivalent |
| Deployment Team: | 3 People - Driver, Operator, Observer |
| Platform Size: | 1.1 m x 0.6 m x 0.5 m (L x W x H) |
| Depth Rating: | Rated / Certified to 50 m, designed and tested to 70 m |
| Forward Propulsion: | Via the tow vessel |
| Vertical Propulsion: | 4 x thrusters |
| Operational Modes: | Manual operation, stabilised (roll + pitch) operation, depth hold operation (operational), altitude hold (developmental) |
| Operational Depth: | 3-18 m for Control work, 20-45 m for deep COTS surveys. |
| Operational Height: | 3-5 m above the bottom (2-8m operational limits) |
| Cameras: | Navigation: 1 x forward navigation video camera (4K video)<br>Science: 2 x downwards payload still cameras (16MP stills images, 3-4 frames per second) |
| Sensors: | Pressure senor (depth), sonar altimeter (height above bottom) |
| Location: | SBAS based GPS located on the ground control station |
| Run Time: | Platform: 2-3 hours in internal batteries (field wet-swappable, run time set by the number of available batteries)<br>GCS: 2 hours with option of external power supply (indefinite run time if run off external power) |
| Survey Duration: | Limited by the number of available batteries for the platform |
| Data Storage: | 2 TB NV-RAM on board storage |
| Science Payload: | 2 x 16MP machine vision planar still image cameras operating at 3-4 fps, full camera synchronisation for true stereo imagery, 8mm equivalent rectilinear lenses with 4” BK7 glass domes. |
| Science Capture: | Continuous along track with overlap dependant on height and speed, across-track side coverage equivalent to 2 x height above the bottom. |

The main design outcome was that the system needed to be relatively small so that it can be used in a small survey vessel, needed to operate at similar operational logistics as the current methods but needed to deliver high quality imagery that was suitable for machine and human based analysis.

After the initial design was completed a series of prototypes were developed and tested including operating these with the Control Teams as part of their field work. This allowed the design to be adapted based on feedback and operational experience from the Control Teams. Two major prototypes (Prototype-I and Prototype-II) were developed before the final design was fixed and the operational units built.

The design focus was in making the units lighter and so easier to use in the field, adding handles and lift points to facilitate manual handling, making them easier to maintain and clean (fresh water flushing of thrusters), moving to an aluminium frame to increase robustness, and hardware updates to utilise the latest available components. The transition from the final prototype to operational unit included an engineering certification to ensure that the design met all required standards and that the unit was overall fit for purpose from an engineering standpoint.


Figure 7. ReefScan-Deep platform (Operational Platform) (credit: S. Bainbridge, AIMS 2025)

### 2.3.4 Platform Components

The platform consists of a metal box frame with four thrusters (one at each corner) that control the orientation of the platform and can be used to change and set the depth. Within the frame is the main electronics enclosure that houses the navigation and camera computers as well as forward facing navigation camera. Behind the enclosure sits the battery that powers the platform with the two downward facing payload cameras located between the battery and the enclosure. The unit in operation is shown in Figure 4, with more detail shown in Appendix A.

The platform has four thrusters (Blue Robotics T200 thrusters) set at the periphery of the platform that allow for vertical movement (dive or come to the surface) as well as giving pitch and roll correction with yaw correction done through the action of the towing craft. Control is via the open-source ArduSub controller implemented via a Pixhawk control unit.

The platform internal navigation controller is able to maintain the platform at a set depth in level or stabilised flight where the platform automatically adjusts its position in the water, via the thrusters, to maintain a stable level operation at the set depth. The role of the operator is therefore to set the height or altitude of the platform above the bottom and to adjust this as required, based on the up-coming terrain.

The optimal survey altitude is between 3-5 m above the bottom. Going closer than 3 m reduces the field of view and increases the potential for a collision, going higher than 5 m can reduce the resolution and image quality making accurate detections difficult. The optimal height above the bottom will depend on the water clarity (turbid water may mean going closer to the bottom to get good imagery) and the complexity of the terrain (highly variable terrain may require a slightly higher flight altitude to reduce the chance of impact).

The platform carries a forward-facing high resolution (4K) navigation camera that assists with operating the platform and two downward facing science cameras that capture the benthic imagery. The platform has a pressure sensor that gives the depth of the platform below the surface and a sonar 'pinger' (Blue Robotics Ping Sonar) that gives the altitude of the platform above the bottom, the sum of these two gives the water depth.

The control system implements depth hold capability out of the box and a prototype altitude hold (terrain following) capability was added. This is still in beta as the implementation of this capability has proven to be problematic given the irregularity and sudden changes in terrain typical of coral reefs and the limitation of a sonar-based altitude sensor (which has a limited field of view and is not particularly accurate).

The underwater platform is self-powered via internal wet-swappable batteries with a run time of 2–3 hours depending on the load on the thrusters. As the power drain from the thrusters can be very large (over 70 amps at 14v), it is not practical to feed power from the surface via the tether as this would mean a significant amount of power going over the wire which in turn raises safety issues for use in a marine environment.

### 2.3.5 Ground Control Station

The Ground Control Station (GCS) is the surface component that controls the underwater platform. The station currently uses the QGroundControl open-source software which allows the platform to control its movement either in depth/altitude hold or through direct input from the operator using a small on-board Windows based computer. The GCS contains a screen to see the underwater navigation camera, controls to operate the platform, and a computer that runs the control software, housed in a waterproof Pelican case (see Figure 8).


Figure 8. Ground Control Station (GCS) in use (credit: S. Bainbridge, AIMS 2024)

The GCS has a Garmin GPS which receives Satellite Based Augmentation System (SBAS) corrections allowing the unit to deliver spatial accuracy to 0.5 to 0.7 m for most of the time. The GPS location and time are written into the header of each image collected by the scientific camera, allowing each image to be accurately geo-located (although to the location is of the GCS NOT the towed platform).

### 2.3.6 Camera Payload (scientific imaging system)

The scientific imaging system or payload consists of the two downward facing cameras and a dedicate computer that processes the images, including running the real time machine learning models. The system is designed so that this is modular and so the payload can be updated without impacting the control or navigation components. The imaging and control systems are connected by a network connection so that the control system can respond to input from the imaging system and vice versa.

The cameras used are industrial Machine Vision cameras that are solid state (no moving shutter or other components that you find in a traditional digital SLR camera) and so are highly reliable and able to work at high frame rates for extended periods. The cameras are powered off the main platform battery and deliver data over a USB connection to the payload computer where the images can be run against the COTS and Scars models. The control and payload computers are connected by a network with the overall control software located on both computers to ensure that data can flow between the control and payload computers.

For example, the GPS location is collected by the computer in the Ground Control Station and then passed down the tether to the Payload Computer where it embeds the location data into the header of each image file collected.

The cameras use high quality lenses and a glass dome port to ensure the highest image quality. Each camera is synchronised giving true stereo images of the bottom that can be used for photogrammetry style analysis if needed. Each image has the GPS date and time, the GPS location and the estimated altitude written into the image header along with the survey metadata (which is entered via the tablet), so forming a permanent record of the exact time and position of every image collected.

### 2.3.7 Platform Testing and Validation

The initial prototypes were field tested on AIMS, COTS Control and Queensland Parks and Wildlife Service (QPWS) trips (a total of four with the Control Teams, four trips on the QPWS vessels and multiple trips on the AIMS vessels). These trips were used to refine the prototype systems and define the form of the operational system. In October 2024, an engineering review was held with key stakeholders, including reef managers and the Control Teams, to finalise the design and specifications for the final delivered operational system (see Appendix F).

The main points raised in the review and with the field testing were:

- Making the units as small and light as possible to make deployment easier.
- Additional handles and lifting points to improve manual handling
- Removal of sharp corners and other surfaces that may impact users, use of soft materials.
- Change to magnetic lens caps that are easier to use than the previous Velcro ones
- Certification of lifting points as per the appropriate standards
- Pressure testing to the appropriate standards to ensure depth rating and certification
- Updates to the software to increase usability
- Better ability to view the detected COTS in real time using the tablet, ability to filter out any false detections

The engineering review looked to ensure that the underlying engineering is appropriate for the intended use of the system and so focused on the strength of the system, the loads that the system may be exposed to, lifting points and so on. This is important and ensures that the systems are fit for use and comply with all appropriate standards and regulations.

## 2.4 Workflows

### 2.4.1 Real-Time Workflow

The real-time workflow looks to identify the presence of COTS and Scars in the collected images (remembering that the benthic model is currently run in the end-of-day workflow) as they are collected so that the operator can respond while the vessel is at the reef. Responses may vary from doing a quick snorkel reconnaissance, radioing in for the Control Team to get ready to target the area, or just to note areas of the reef for follow-up activity. The goal is to provide instant feedback and so give the operator an overview of the reef status so that reactive monitoring and control activities can be implemented as required.

A simplified diagram of the real-time workflow is shown in Figure 10, along with the end of day workflow, a more detailed workflow is shown in Appendix B.

An initial part of the workflow is the collection of metadata to describe the survey. This is entered in the field at the start of the survey using the tablet, and includes information about the location, the people undertaking the survey, and the environmental conditions such as sea state, water visibility and so on. The metadata can additionally be edited and added to as part of the end of day workflow using a laptop.

Images are captured by the Capture Node on the ReefScan system, they are then passed to a node that adds in the date/time, GPS location and transect identifier information to the image header, the image is named based on the transect identifier and the date time (as UTC, including fractions of seconds) and stored in the transect folder.

The image is then passed to the ML model which detects any COTS and Scars in the image. If a COTS is found, a message is generated with an ID allocated to the detection, along with the bounding box that encompasses the starfish in the image. For COTS, a tracking program is run to determine if the detection is one previously seen, in which case this detection is allocated the ID of the previously seen detection so that the same COTS is not counted twice. The detection is passed to the tablet for confirmation by the field team (typically by the Observer, see Section 2.6 for the designated roles of the field personnel) who can mark the detection as true (it is a starfish = true-positive) or false (it is not a starfish = false-positive).

If a Scar is detected in the image, then a mask that encompasses pixels in the scar is generated along with a probability for each included pixel in the mask. Detections are passed as messages to the rest of the system and then picked up by the tablet display node that shows the detection as an image which the user can then examine to confirm or reject the detection. Detection confirmations can be done in real time via the tablet or at the end of the day via the desktop software as part of the end-of-day workflow. Any confirmations of both COTS and Scars are logged for future use.

At the end of the survey the system generates a map of the detections, also called a heat-map (Figure 9), that allows the field team to correlate the survey area with the detected starfish. This has value in understanding where on the reef the starfish have been seen.

### 2.4.2 End-of-Day Workflow

Once the surveys for the day have been completed the survey unit is brought back to the main vessel and, using the provided ReefScan software, the data is downloaded off the platform and the platform batteries recharged. Any missing metadata is added and the images and surveys are quality controlled by the user using the provided software. The data is then run through local versions of the COTS and Scars models as well as the Benthic model. The data from these can then be exported as Keyhole Markup Language (KML) files, a file format used to display geographic data, to aid in planning the activities for the next day.

The data is backed up to two external hard drives as part of the download process. A simplified view of the end of day workflow is shown in Figure 10, with the more detailed workflow shown in Appendix B.

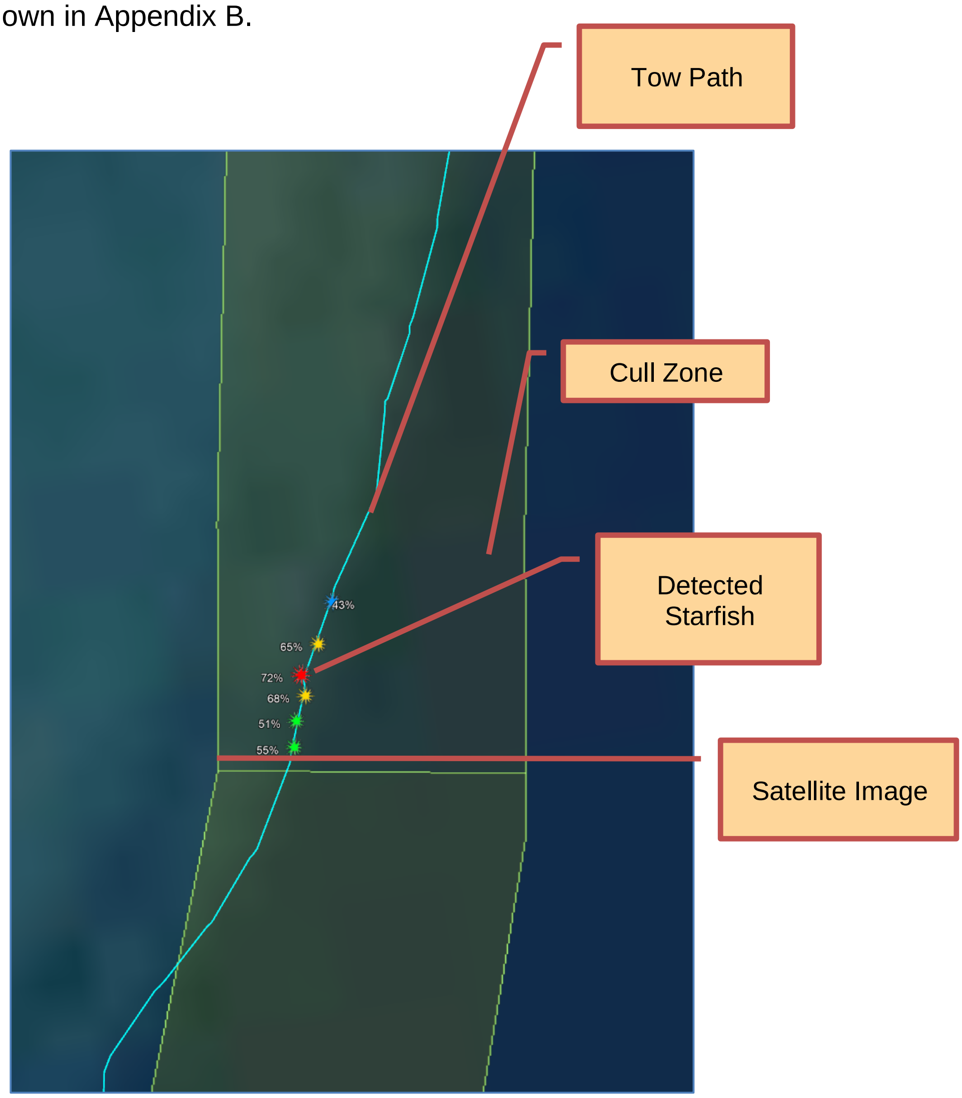


Figure 9. Example of a real-time detection Heat-Map, Snake Reef (Google Earth). The colours represent the confidence value from the detected starfish. Green indicates 50-60%, Yellow indicates 60-70% and Red indicates greater than 70% detection confidence.

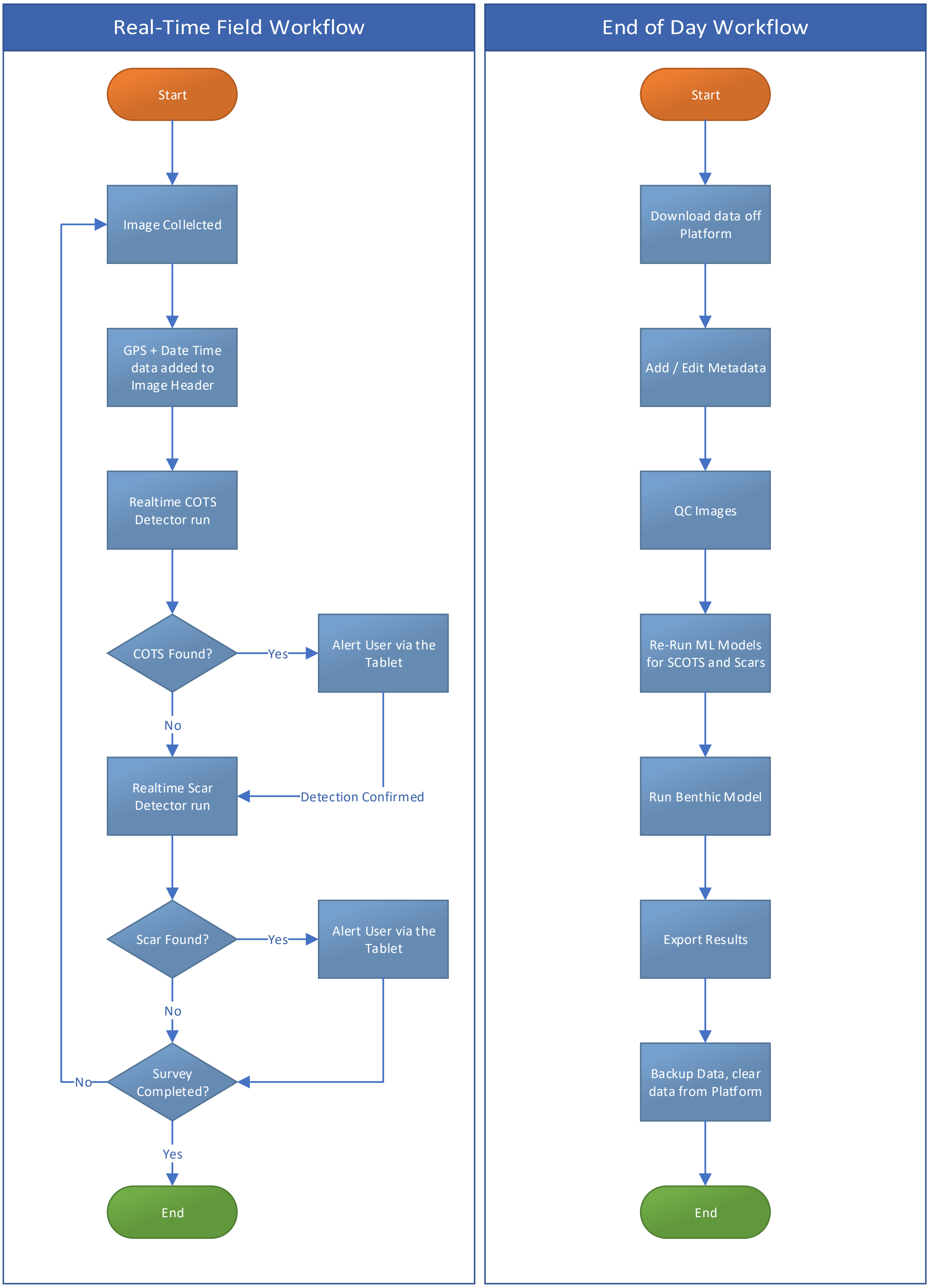


Figure 10. Real-time and end-of-day workflows (see also Appendix B).

### 2.4.3 End-of-Trip Workflow

The end of trip workflow uses the AIMS ReefScan desktop software to do a number of tasks:

1. Any non-downloaded data is downloaded off the platform and backed up onto two hard drives.

2. The COTS and Scars detection data, along with the metadata, is exported into an Excel compatible format (CSV file).

3. The images are uploaded to the AIMS ReefCloud system for further analysis and model training for the benthic identifications.

4. The ReefScan platform memory is then cleared to ensure enough hard disk memory for future surveys.

The AIMS ReefCloud system is a cloud-based computing platform for the automated analysis of benthic images for coral cover and benthic forms. It has an equivalent ML model to the desktop software (the desktop benthic model is derived from the ReefCloud model) but adds additional functionality in that images can be manually annotated and the model re-trained on the new data. This allows for the benthic model to be re-trained to increase the accuracy of the model.

The end of day benthic model, while derived from the ReefCloud model, runs with fewer analytical layers and so has less performance than the ReefCloud model to accommodate the more limited computing resources available at sea. The desktop benthic model, for example, may miss-identify some classes more often than the cloud-based one. The ReefCloud model can be re-trained on the collected data and then that model can be exported back into the desktop software so that as the model improves the core functionality is made available on all platforms.

At this time, while the system collects information about correct and incorrect COTS and Scars identifications (via the Observer confirming detections using the tablet) there is no pathway to automatically train the model on this new data. This may be a feature of future systems.

The end of trip workflow also allows for data to be exported to other systems. For the COTS and Scars data this is an Excel compatible file (CSV file), for the benthic data this is an export from the ReefCloud system. Future work will look to streamline this so that data directly follows to information systems, such as the Reef Knowledge system, that is being developed by the Reef Authority and partners.

The end-of-trip workflow is shown in Appendix B.

## 2.5 Model Development

Initially the focus of the Machine Learning models was the detection of COTS. In low COTS density scenarios, finding COTS has a high degree of variability – the starfish has to be out in the open during day-time hours and the platform needs to go directly over it. This variability has been documented for other survey methods, such as the manta tow (Lawrence et al. 2024a). For example, the threshold value for the COTS Control Teams to cull an area is one COTS per manta tow, which means that the detection variability can have a large impact on the activities of the Control Teams.

In low COTS density scenarios, a better strategy is to rely on detecting COTS scars as a proxy for COTS, rather than on direct COTS detections. COTS scars are static, cover larger geographic areas, are visible for several days, and thus provide a good proxy of recent COTS activity (Chandler et al. 2023). As a result, the ML work was extended to include detecting scars.

This was a major change for the project with additional work being done to include scars. This became complex as there are multiple sources of scars, only some of which are created by COTS. As a result, the model needed to have some ability to distinguish between COTS induced scarring and scars produced by other means and other events, such as coral bleaching, that can look similar. The final model results show an accuracy of around 40% for Scars detection indicating both the level of maturity of the model and the difficultly in distinguishing visually between COTS Scars and scars from other sources (80% or better accuracy is considered to be an operational model so a result of 40 % is well below a level where the model is considered to be operational).

Most of the issues were with false-positives, that is detecting other sources of white coral (such as coral disease, grazing by fish and so on) as COTS scars. The model therefore needs to be trained on more COTS scar images so that it can better distinguish between scars from COTS and scars from other sources. As such the scars detection work was moved to developmental and so the scars model is still a work in progress.

### 2.5.1 Model Design

The project initially started with the idea of developing two ML models, one for COTS detection that would be undertaken by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) (COTS Model), and one for benthic composition that would be undertaken by AIMS (Benthic Model). Following the initial work, it was decided to extend the CSIRO model to include COTS Scars (Scars Model) as an extension to the COTS model. The combined COTS and Scars model was designed to be part of the real-time workflow while the Benthic model would be run in real-time if the field systems had enough compute power or, if not, would be run as part of the end-of-day workflow.

The models were developed using the following methodology:

- Several field trips were conducted to generate training dataset for the ML models. Images that contained objects of interest, such as COTS or scars, were quality controlled and annotated by a human.

- Next, the dataset was divided into training images, against which the model was developed, and test images, against which the model performance was assessed. Typically, 20% of images are held back as test images with the remaining 80% used to train the model (Sivakumar et al, 2024).
- Images that didn't contain the object, but which were typical of the images to be collected were also included in the training data.
- Where test images were not detected correctly, additional annotated images were used to train the model until the model's performance reached sufficient accuracy.
- To ensure models performed as required, on-going annotation effort will likely be required to address instances where the model performs poorly (e.g. deep water, dark images).

What makes a model useful in someways depends on the use. If the impact of the model getting things wrong is high, then the model must be trained and optimised to deliver the required level of utility.

Model results can be characterized as belonging to three outcomes:

1. The model correctly identifies the object in an image or correctly identifies that the object is not in the image (= true-positive or true-negative, respectively)
2. The model identifies an object in the image that is not in the image (= false-positive)
3. The model fails to identify an object in the image (= false-negative)

These outcomes can be used to develop various measures of the model performance. Commonly used measures include (Goutte and Gaussier, 2005):

### *Recall:*

Recall measures the ability of the model to correctly identify all positive instances (true-positives) out of all actual positive instances. This is a measure of how many times the model gets it right for images that do contain the object.

### *Precision:*

Precision measures the ability of the model to correctly identify positive instances (true-positives) out of all instances it predicted as positive. This is the ratio of correct positive decisions over the number of positive detections (which includes false-positives)

### *F1 Score:*

The F1 score is a metric that combines both precision and recall. It is calculated as the harmonic mean of precision and recall, meaning it gives equal weight to both. A perfect F1 score is 1.0, while 0.0 indicates the worst possible performance.

The F1 score helps you find a balance between precision and recall. If you prioritize precision (minimizing false-positives), you might sacrifice some recall (identifying true positives). Conversely, if you prioritize recall (identifying all true-positives), you might sacrifice some precision (leading to more false-positives).

While there's no universal benchmark, an F1 score of 0.9 or higher is generally considered excellent, indicating a model that effectively balances precision and recall. Scores between 0.7 and 0.9 are typically considered good for most applications. Scores below 0.5 suggest the model needs significant improvement (Goutte and Gaussier, 2005). As such, a model that has an F1 score above 0.8 (80%) can be considered as operational, those below 0.5 (50%) as in development (beta) and those in between as operational with caveats (for example use results with caution or contextualise with other data).

The performance of the model (as per the F1 score) is influenced by how visually unique the object looks or, conversely, how easy it is to confuse with other objects. For COTS, we expect high scores as COTS are visually distinct, although there are other reef objects that share similar shape and colour. For Scars, we expect lower scores as even a trained observer can have trouble identifying a COTS Scar versus other sources of white coral such as disease, storm damage and so on (Morgan Pratchett, pers comms).

While the F1 score is a measure of overall model performance, each detection also gets a detection score between 0 and 1 where a higher score is where the model is more confident of the detection. To remove false-positives the detections need to be filtered, for example all detections with a detection score of greater than 50% maybe said to be an actual detected starfish and anything below as a false-positive. This means that work needs to be done to understand the score the model delivers at the detection level with the filtering based on the impact of missed detections (filter too high) or many false-positives (filter too low).

### 2.5.2 Development of the COTS Model

CSIRO implemented an ML model, DeNet (Tychsen-Smith and Petersson, 2017), for COTS detection, initially training it on publicly available data for COTS detection (see Table 5). The annotated dataset plays a critical role in the development of the ML models.

#### *Dataset Annotation*

An annotated dataset is a collection of data that has been labelled with meaningful information to help a computer understand what the data represents. For example, an image might be labelled to show it contains Crown-of-Thorns Starfish (COTS) or indicate the presence of feeding scars. These annotations act as examples that teach machine learning models how to recognize patterns or make decisions. By training on many such labelled examples, the model learns to make predictions or classifications on new, unlabelled data.

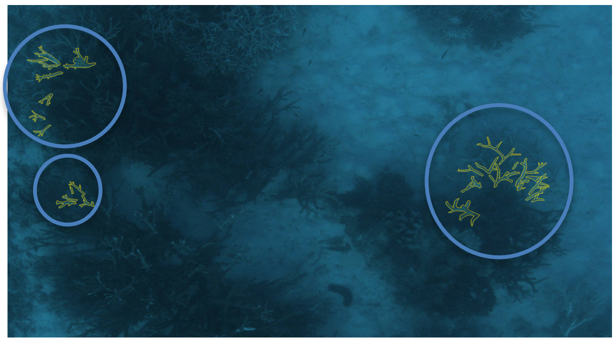

Figure 11. A sample image with scars labelling

### *Model Development Phases*

The model performance is dependent on the nature of the images themselves, such as scale, resolution and even lighting. As a result, the initial model was re-trained on operational images from the platform. This can be shown in Table 6 where the initial model performance had an accuracy score of 64% but jumped to over 87% once re-trained further on operational images.

For this project, the task of real-time and end-of-day models for COTS instance detection is broken down into two phases applied sequentially:

- **Object Detection:** The detection phase takes the current image as input and generates a set of bounding boxes (x, y, width, height) and associated normalised scores, with each score indicating the likelihood that a COTS is contained within the bounding box.

- **Object Tracking:** The tracking phase then combines these bounding boxes with the set of previously generated bounding boxes (from the previous images) to calculate the likelihood that this detection is the same as the previous one, just moved forward in space and time, or if it is a new starfish detection. This allows the system to identify the same starfish over subsequent fames as distinct from treating each detection as a new starfish.

### *Object Detection*

The DeNet Model consists of an object detector, implemented in C++ using TensorRT, that first applies a set of large bounding boxes to divide the image into sections and then a classifier that looks in each box to see if it contains a COTS. Finally, the boxes are compared to ensure that any COTS that occurs in more than one box is not the same animal and so not counted twice. This is all done in real time using the platform’s Nvidia Jetson Graphics Processing Unit (GPU).

The detection process can be broken down into two neural networks followed by a Non-Max Suppression method:

1. **Region Proposal Network:** The first neural network generates a set of coarse bounding boxes proposals or regions-of-interest which may contain an object.

2. **Classifier Head:** The second neural network is responsible for taking the regions-of-interest and classifying whether they contain a COTS and, if so, updating the bounding box to better fit the object within.

3. **Non-Max Suppression:** This method is applied to remove duplicate bounding boxes and ensure that each object instance has only a single associated bounding box.

The two neural networks are jointly trained using the annotated dataset described in Section 3.4.1. Briefly, the machine learning training method works by constructing an appropriate error function, and then attempting to minimise it via GPU accelerated stochastic gradient descent methods.

*Object Tracking*

Given the hardware constraints (the real time computer on the platform has limited time (the collection image frame rate is 3-4 frames per second leaving only ⅓ to ¼ of a second to process each image before the next image is collected) and processing power (constrained by the space and power that can be provided on the platform)), a custom method utilising traditional optical flow methods was employed for object tracking.

This process can be broken into two steps:

- **Update:** Given the existing bounding boxes, estimate the new bounding boxes for the current frame. An optical-flow based method was applied which compares the previous image with the current image to estimate the new location of the pixels contained within each bounding box.

- **Merge:** Combine the existing updated bounding boxes with new bounding boxes generated by the detector, removing duplicates.

### 2.5.3 Development of the Scars Model

Initially it was planned to implement two DeNet models in sequence, a COTS model and then a separate Scars Model. Work showed however that a single model was more effective, both computationally and in terms of ease of deployment. As such a single combined COTS and Scars DeNet model was developed and used operationally.

Note that while a single model was used for both COTS and Scars, the level of reliability of each is different. While cryptic, COTS have specific features that make them visually different from other benthic organisms. The performance of the COTS model is thus reasonably robust (see the Results section). Multiple causes, such as bleaching, disease, or other pests, can create coral scarring and COTS scars are more challenging to identify visually, especially when combined with the high diversity of coral morphologies on the GBR. \

The COTS scar model is thus less robust and performs at a level below what would be considered operational (F1 scores below 50%, see the Results). While increasing the size of training data may help improve the scar model, it is likely that more sophisticated ML modelling approaches will be needed to achieve high operational accuracy.

The end-of-day and real-time DeNet models were redesigned to output both bounding box detections for COTS and pixel-wise semantic segmentation for COTS Scars. These tasks are performed within the same model with negligible impact on runtime and memory consumption. In its current configuration, the model takes a 1344x768 pixel image as input and generates a normalised scar score for each 8x8 pixel region. By observing this scar score the system can apply further processing to estimate the likelihood that the image contains any coral scars and the percentage of the image with scarring.

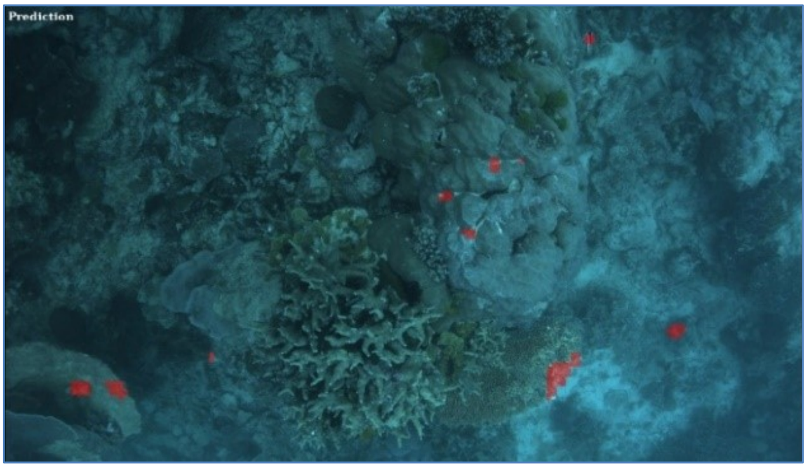


Figure 12. Scars model output with pixel-wise semantic segmentation of the image as scars

### 2.5.4 Development of the Benthic Model

AIMS has previously developed a Machine Learning (ML) model for benthic life forms using data collected by the AIMS Long-Term Monitoring Program (LTMP). The LTMP model comprises a total of 56 classes arranged into 14 hierarchical groups.

For this project, the LTMP model was simplified slightly to a of total 50 classes in seven hierarchical groups removing some items that were of less relevance (ones at Genus or even Species level noting that the original LTMP images were taken by a diver 50cm off the bottom and so give more taxonomic resolution than the ReefScan images). The pruned set of classes better reflects benthic life forms over the LTMP classes that include more taxonomic forms reflecting the nature of the ReefScan images (larger field of view, less resolution) and the information needs of the project (percent cover estimates of the main benthic life-forms).

The total set of used classes is shown in Appendix D with the higher-level functional grouping shown in Table 2.

Table 2. Benthic model functional groups

| Code | Description |
|---|---|
| A | Algae |
| AB | Abiotic |
| HC | Hard Coral |
| IN | Indeterminate |
| OT | Other |
| SC | Soft Coral |
| SP | Sponges |

The model was initially trained using diver-based imagery using the back-catalogue of hundreds of thousands of manually annotated images collected by AIMS. The model was then re-trained via manual annotation of the ReefScan images with an ecologist manually identifying and allocating areas on the images to benthic classes.

The model was developed using a Convoluted Neural Network approach using the TensorFlow framework and was implemented both as a real-time model within in the real-time workflow and as part of the end-of-day workflow. The real-time benthic model is currently not used operationally due to prioritising the real-time COTS and Scars models within the limited computing resources available to the real-time system (currently an Nvidia Orin Jetson compute platform).

### 2.5.5 Training image collection and annotation

The platform was used on six field trips (three with prototype-I and three with prototype-II) undertaking some 71 surveys collecting over six hundred thousand images (see Table 3). Note that the system was used on other AIMS developmental trips, but these images were not included as they were collected under test conditions and not under conditions typical of that of a Control Team (differing parts of the reef, differing altitudes, tow speeds, and so on).

A map showing the survey trip locations is shown in Figure 13. Note that in Table 3 the term “survey” just refers to a continuous run with the platform with no specific length or time value. As such these do not relate to manta tow transects (which are 200 m / two minutes) or other formally defined (distance or time) surveys. Two additional trips were done with the Operational Platform in early 2025 on the QPWS vessels with an additional 26 surveys undertaken but, as the project had completed, these were not used for further development of the models.

The COTS and Scars ML models are trained by giving them images with COTS or Scars where the location of the COTS or Scars has been manually identified by a trained observer. This allows the ML model to understand the shape, texture and features that comprise a starfish as distinct from other similar looking objects. The Benthic model was developed from the existing LTMP model but re-trained for the platform images (which have differing scale and resolution) using annotations by an AIMS trained ecologist using the ReefCloud system.

The collected images were manually annotated by CSIRO (images from prototype-I) and AIMS (images from prototype-II) using the Computer Vision Annotation Tool (CVAT) software as bounding boxes for the annotated COTS and as pixel masks (polygon) for the annotated Scars. Annotation involves annually identifying the location and extent of the starfish or scar (typically as a bounding box that encompasses the object) using a human expert. These annotations represent actual starfish and scars and are used to train the model.

The total number of annotated COTS and COTS scars and their final split for training and test / evaluation of the COTS and COTS scar detection and segmentation are presented in Table 4.

Table 3. Details of field work undertaken to test prototype systems.

| Date | Agency | Vessel | Unit deployed | Number of surveys | Number of images collected |
|---|---|---|---|---|---|
| Apr-23 | CSIRO / QPWS | *Reef Ranger* | Prototype-I | 22 | 57,357 |
| Oct-23 | Pacific Marine Group | *Odyssey* | Prototype-I | 7 | 37,574 |
| Dec-23 | Blue Planet Marine | *Infamis* | Prototype-I | 9 | 80,308 |
| Jan-24 | QPWS | *Reef Resilience* | Prototype-II | 11 | 93,624 |
| Sep-24 | Blue Planet Marine | *Flying Fish* | Prototype-II | 16 | 288,212 |
| Oct-24 | Pacific Marine Group | *Odyssey* | Prototype-II | 6 | 83,060 |
| **Total** | | | | **71** | **640,135** |

Table 4. List of COTS and COTS Scars annotations from manually reviewed images.

| Annotated data | Number of images | Images (COTS) | COTS bounding boxes | Images (Scars) | Scars polygons |
|---|---|---|---|---|---|
| Training set | 9,744 | 6,152 | 8,204 | 4,157 | 16,321 |
| Test set | 1,499 | 918 | 1,004 | 639 | 1,821 |
| **Total** | **11,243** | **7,070** | **9,208** | **4,796** | **18,142** |

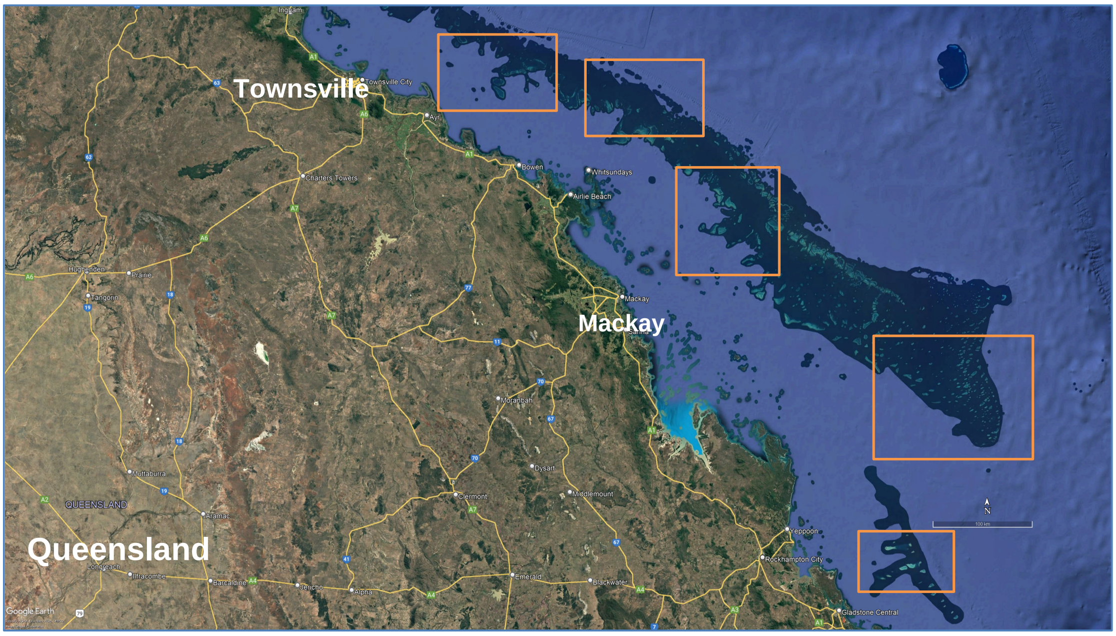


Figure 13. Locations of the field work (map source: Google Earth 2024).

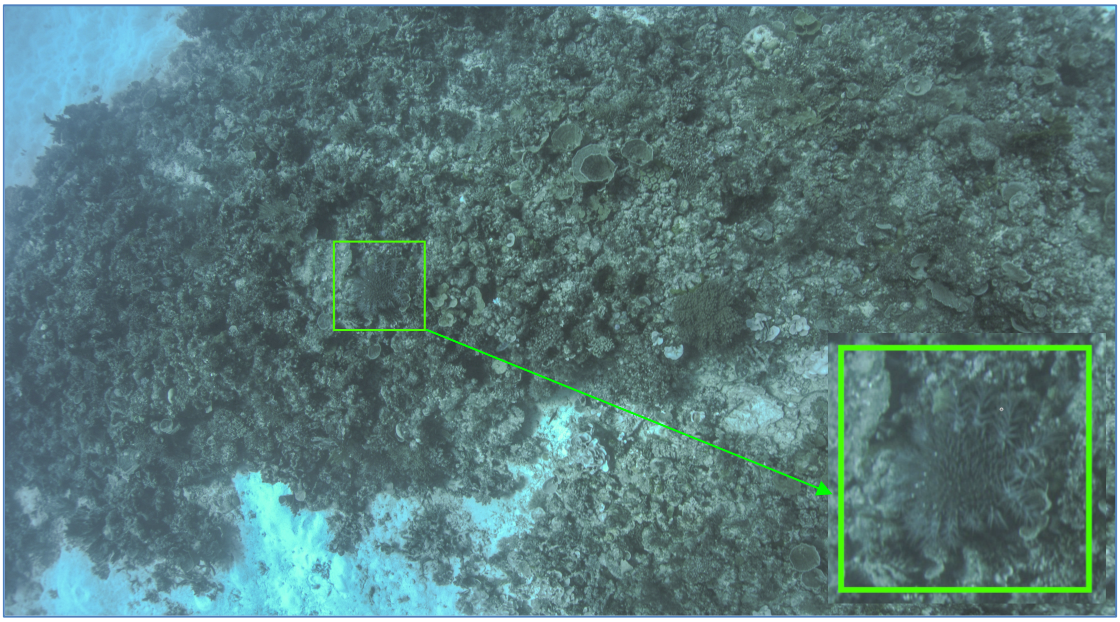

Figure 14. Example image showing a detected COTS (Snake Reef, enhanced for publication).

## 2.5.6 Annotated Training Dataset (COTS and Scars)

The data was annotated for COTS and Scars using the CVAT software (www.cvat.ai). Due to the large number of images, the images were first visually scanned to identify images with COTS and Scars, and these were fed into the CVAT software. For COTS, a bounding box was manually placed around any identified COTS, for Scars a number of pixels that made up the scar were identified in the CVAT software which then used image segmentation to find the pixels that represented that scar.

Once the annotation was finalised, the dataset was quality controlled and then divided into training and test sets to evaluate the model's performance at various stages of the project. The test set includes 1004 instances of COTS and 1,821 Scar polygons, enabling the evaluation of the model's performance in both COTS detection and tracking as well as scar segmentation. Table 5 summarises the data collected for model development.

Note that for the initial COTS model, training was done via a previously developed crowd-sourced platform called Kaggle (www.kaggle.com) where teams compete to develop the best ML model to answer the problem put forward, in this case identifying COTS in images. Details of the challenge can be found at: https://www.kaggle.com/competitions/tensorflow-great-barrier-reef/discussion/290062.

Table 5. Summary of annotated data used for model development.

| Data type | Collection | Number of images | COTS bounding boxes | Scars polygons |
|---|---|---|---|---|
| Train (initial) | Kaggle (train& val.) | 10,285 | 26,014 | 0 |
| Train (Jan22) | Kaggle + CCIP | 13,492 | 30,219 | 757 |
| Train (Jun23) | Kaggle + CCIP | 17,285 | 31,757 | 13,281 |
| Train Final (Nov24) | Kaggle + CCIP | 20,029 | 34,218 | 16,321 |
| **Test** | **CCIP** | **1,499** | **1,004** | **1,821** |

## 2.6 System Operation

The system is designed to be operated by a team of three noting that this reflects the current Manta Tow method where a team of three is also used (driver, manta tow person and observer).

The team consists of the following roles:

Driver: Keeps the vessel along a set path and depth contour at 2-3 knots and is responsible for overall safety of the vessel and crew.

Operator: Uses the Ground Control Station to 'fly' the Platform 3-5 metres above the bottom using the forward-facing Navigation camera and ensures that the Platform does not strike any upcoming topography by communicating to the Driver if they need to stop or deviate.

Observer: Manages the tether letting more tether out as the platform dives and reeling in excess tether as the platform rises, noting that the length deployed should be two (2) times the platform depth (tether lengths are marked on the cable). The tether should be kept tensioned so that it does not go near the motor.

The Observer scans the topography of the reef that the Platform will go over to let the Operator know what is coming up so that they can better deal with any sudden changes in topography.

The system allows for real time detections of COTS and Scars to be reviewed using the tablet, typically this is done by the Observer. The detections can be filtered to reduce the number of false positives where each detection is presented as an image with either a bounding box around the detected object for COTS or as coloured pixels for Scars. The user can zoom and pan into the images to evaluate the detection and confirm if it is correct or not. Confirmations are logged and used later to help train the models.

When the survey is complete a ‘heat-map’ of the detections is presented to the Observer via the tablet. This allows them to see spatially the distribution of the detections and gives them a way to contextualise the results based on the reef just surveyed.

## 2.7 Survey Metadata

Metadata for each survey can be entered using the tablet based ReefScan App or using the ReefScan desktop software at the end of the day. The optimum approach is to enter the metadata via the tablet while undertaking the survey and then to edit or add to this as needed later.

The metadata looks to record information about the survey to help with the storage and use of the data. The desktop software allows for the metadata to be edited to ensure that this is complete and correct.

The following metadata is collected:

- The Reef Name and Reef-Id (selected from the Reef Authority Gazetteer)
- Date and time (auto filled from the GPS)
- The Operator, Driver and Observer names
- A name for the survey
- The environmental conditions:
    - The cloud cover in octas (0 = no cloud, 8 = full cloud)
    - The wind state (Calm 0-5 | 5-10 | 10-15 | 15-20 | 20-25 | > 25 knots)
    - The water visibility (an estimate of distance you can see the bottom down to in metres)
    - The tide state (Rising | Falling | High | Low)
    - The sea condition (Calm | Slight | Moderate | Rough)

## 2.8 System Certification and Validation

### 2.8.1 System Certification and Permits

An external company, EDMS Australia, was engaged to undertake the certification of the final operational platform. This was done to ensure the platform was fit for purpose noting that there are few formal standards that relate to low-voltage small underwater systems.

The certification looked at:

| | |
|---|---|
| AS4991 | Lifting Devices (lifting points, loads, etc.) |
| AS1200 | Pressure Equipment (housings, etc.) |
| AS1201 | Pressure Vessels (housings) |
| AS1664 | Aluminium Structures (frames, structures) |
| AS4024 | Safety of Machinery (mechanism of use) |

The design and construction of the platforms and equipment was done in light of the relevant standards and to the operational conditions with regard to lifting and pressure rating of housings and enclosures. An engineering review (see Appendix F) was done with the Control Teams and other stakeholders to drive the final design of the delivered units.

The units were deployed under the AIMS Great Barrier Reef Marine Park Authority (Reef Authority) General Permit (Permit G21/38062.1) which covers the use of towed platforms within the GBR Marine Park. No samples were taken, nor experiments performed, and so ethics approval was not required.

### 2.8.2 Stakeholder Engagement

The use and benefits of the developed systems were communicated via a series of technology transfer / stakeholder engagement events. Mostly these were done by one or two AIMS staff taking the system on an operational COTS Control trip using the new system in parallel with the existing human-based surveillance methods. This allowed the Control Teams to understand how the unit is to be deployed, the limitations of the unit, the data and information workflows, and to learn how the system operates.

A series of shore-based stakeholder engagement workshops were also held to involve a larger number of people, including the engineering review (see Appendix F). The engineering review gave the users a chance to comment on the current system (final prototype) and suggest improvements based on their experience. As this was done on the final prototype the scope of major changes was limited but a number of changes were made based on feedback (see the list of items identified in Appendix F).

The review included the Reef Authority who will assume operational management of the new surveillance systems in their role as managers of the COTS Control Program and so they will have overall control of the delivered capacity in terms of how it is deployed within the Control Program.

# 3. RESULTS

## 3.1 Platform Development

Three generations of the system were developed over the life of the project: an initial prototype based on previous platforms used for towed video (Prototype-I), a second prototype that looked to address structural and reliability issues with the first prototype (Prototype-II), and the final delivered Operational Platform.

The main change from Protype-I to Prototype-II was replacing the plastic frame with an alloy frame, the move to two cameras over the previous single one and swapping out some of the lower grade components with higher specified industrial components (mostly ones rated to higher operational temperatures). The Ground Control Station was moved from two separate computers into a single integrated unit more suitable for use in small vessels. The second prototype was developed enough to test with the Control Teams and was used extensively for the collection of images for the project.

User feedback and an engineering review led to the finessing of the second prototype into the final delivered platform. The electronics canister was changed from plastic (Delrin) to alloy to allow for the system to be rated deeper (from 40 m for Prototype-II to 50 m, tested to 70 m, for the Operational Platform) to allow for future deeper reef surveys. Some of the electronics components, such as the forward-facing navigation camera, were upgraded to more industrial designs to ensure reliability. The ground Control Station was reduced in size and weight to facilitate field use, and supporting materials such as Standard Operating Procedures (SOP's), maintenance guides and certification where developed.

The main changes from Prototype I to Prototype II were:

- Move to an alloy frame over a plastic frame to increase rigidity and field robustness.
- Move to two cameras over the original single camera
- Development of a transport case to make transport and storage easier
- Move to industrial level componentry over the initial prosumer level, for example using an industrial controller (Orange Co-Pilot) over a Raspberry-Pi
- Development of an integrated top-box or Ground Control Station to replace the two laptops used previously

Prototype II to the Operational Platform:

- Addition of additional handles for ease of transport
- Reduction in weight for the frame to reduce overall weight to near 28 kgs
- Move to an alloy electronics canister over a Delrin one, increased depth rating and better heat dissipation
- Move to a higher resolution forward facing navigation camera
- Total re-design of the top box to be smaller and lighter

- Development of SOPs to guide the best practice use of the system
- Certification of the platform.

The final operational system was used in two field trips in early 2025 on the QPWS vessels and then delivered as final complete units to the Reef Authority in July 2025.

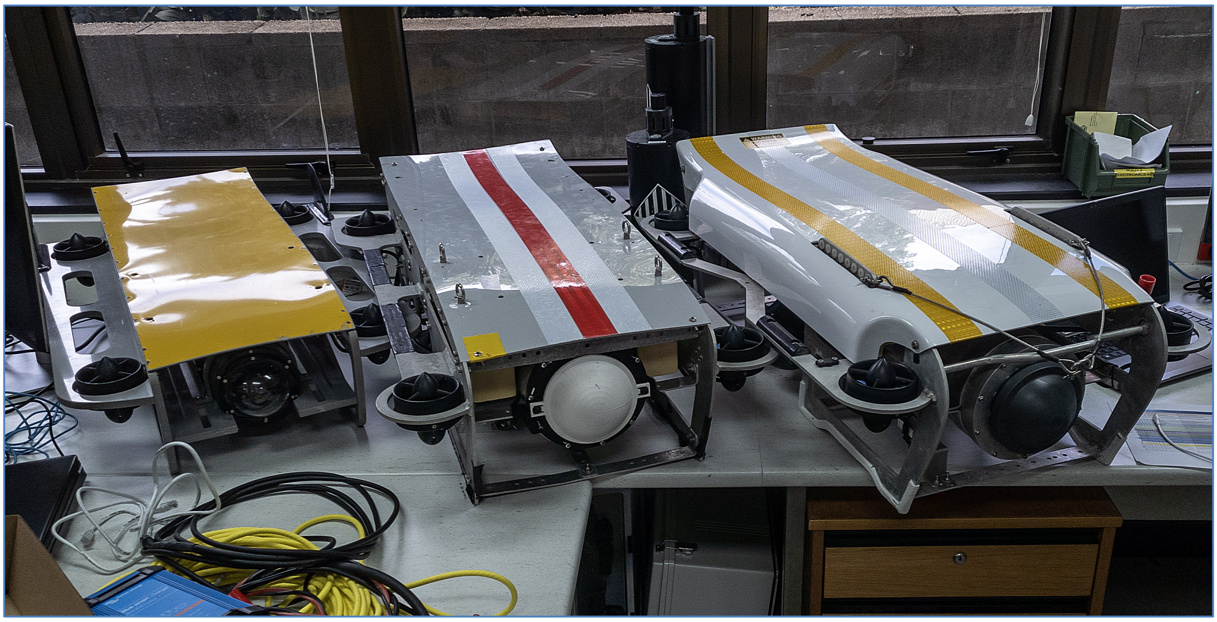

Figure 15. Platform development from Prototype-I (left) to the Operational Platform (right)

## 3.2 System Application and Use

The platform, as Prototype-I and Prototype-II, was used on a total of six control style field trips with either the COTS Control Teams or, for other CCIP projects, via QPWS, with a total of 71 surveys (where surveys were defined as a distinct application of the platform not a deployment of specific duration or length) with 640,135 images collected. This represents some 50.8 hours of operation in total.

The initial prototype platform (Prototype-I) was damaged on one trip on a QPWS vessel when the internal electronics enclosure leaked resulting in the development of the second prototype. This second platform was re-designed and tested by the Control Teams with the outcome from this informing the final operational design via the engineering review.

The Prototype-II platform was used on four COTS Control trips: two with Blue Planet Marine (BPM) and two with Pacific Marine Group (PMG). Due to logistics, it was not possible to do a trip with the third Control Team, Inloc, during the project though this is planned in future.

An additional two trips were done as part of a separate CCIP validation study using the Operational Platform on the QPWS vessel the R.V. *Reef Ranger* in early 2025. For these trips an additional 26 surveys were completed and 46,000 images collected. As this was done at the very end of the project it was not possible to annotate these and use them for model development.

## 3.3 Workflow Development

The real-time workflows were adapted from existing workflows to include the application of the COTS and Scars model. The combined COTS and Scars model is run in its own container wrapped as a ROS node allowing for updates and changes to the model to be made independently from the main architecture.

With the addition of the model node, the architecture was extended to produce detection messages for both COTS and Scars from the detection node that could then be acted upon by other nodes. This included one node that allowed the tablet to notify the operator of a detection, one that produced a quick-look of the detection allowing the user to confirm the detection in real-time, and finally, a mapping node for mapping the detections on a map as a tablet-based app to produce a heat-map of the surveyed area.

The models were also integrated into the end-of-day workflow that runs under custom software in a Windows based laptop or desktop. The software downloads the images off the platform creating a backup on two differing drives for redundancy. From there the user can enter and update the survey metadata, quality check the images and run the benthic model to give benthic percent cover estimates for the surveys completed. The COTS and Scars models can be re-run to detect starfish and scars. Any detections, either from the real time work using the tablet in the field, or the detections from the end of day workflow, can then be confirmed.

The reason for doing the confirmations in the field and again at the end of the day is that often there is not enough time in the field to do a more detailed confirmation of the detection (in a moving vessel) and so being able to review these using a laptop and monitor gives more consistent results. The core data can then be displayed as a map or exported as a KML file for overlay in Google Earth.

The end-of-day workflow is targeted at getting images off the platform and securely onto local hard drives and then to generate information that can be used to plan subsequent field activities. The work done was designed to integrate into the planning processes of the COTS Control Teams as understood from the field trips undertaken on the Control vessels. At the time of reporting this integration was mostly by providing Google Earth overlays that work with the existing planning maps (such as the cull zone maps). Future work will look to tighten this integration so that a more standardised workflow that works with the Control Team workflows can be developed.

The end-of-trip workflow pushes the collected images to the AIMS ReefCloud analysis system for further analysis. While the ML model for benthos used in the end of day workflow is based on the LTMP one in ReefCloud, the ReefCloud model has access to more computing resources and so can apply more complex models (models that have more convolutional layers) which provides increased accuracy.

ReefCloud also allows for manual annotation of new images and so the model can be easily re-trained on the images collected by the user. ReefCloud therefore provides a more complete environment for running and training models and, while the desktop models used for the end of day workflows are driven off the ReefCloud ones, they are limited in capacity by the available amount of computing power available on the vessel (typically just a laptop).

The software developed for the end-of-day workflows had an additional module that allows synchronisation between the collected images and a project within ReefCloud. The software allows the user to match the surveyed data with a ReefCloud project with the data then being automatically pushed to ReefCloud.

As the frame rate of the ReefScan Deep platforms is 3–4 frames per second the upload software removes what are effectively duplicates and only uploads the unique images to ensure detections are not duplicated (the real-time and end-of-day workflows use object tracking to identify the same COTS in image sequences to stop double counting starfish).

## 3.4 Models

### 3.4.1 COTS Model (detection and tracking)

The results evaluating the COTS model are reported based on COTS detection (per-image metric) and COTS tracking (per-instance metrics).

*COTS detection (per-image metric)*

For object detection without tracking, evaluation metrics are typically applied on a per-frame basis. In such cases, the predicted and the annotated bounding boxes for each frame are compared, generating the true positive (TP), false positive (FP), false negative (FN), and true negative (TN) statistics. These metrics are commonly summarised using precision (the percentage of positive predictions that are correct) and recall (the percentage of ground truth instances that are correctly identified). These are combined into a Peak F1 score which is a measure of the harmonic mean of precision and recall (Goutte and Gaussier 2005).

To clarify the performance metrics used, the F1 score measures the balance between precision (the proportion of detected instances that are correct) and recall (the proportion of actual instances that were correctly detected) and is calculated at a fixed decision threshold (e.g. 0.5 or 50%).

The Peak F1 score, in contrast, represents the maximum F1 score achieved across all thresholds, reflecting the model's best possible balance of precision and recall. In our model development, we report average precision, average recall, and peak F1 score, all computed using the annotated training dataset described in Table 6. The corresponding precision-recall curve, which shows how precision and recall vary with different thresholds, is shown in the left panel of Figure 17. These metrics are specifically used to evaluate detection performance, and the percentages shown reflect how accurately the model identifies COTS or scars across varying levels of confidence.

Table 6. Model performance for COTS Detection (per-image metrics).

| Model | Average Precision | Average Recall | Peak F1-Score |
|---|---|---|---|
| Initial | 50.3% | 49.9% | 64.2% |
| Train (Jan22) | 76.1% | 76.4% | 79.4% |
| Train (Jun23) | 84.7% | 84.3% | 84.7% |
| Train Final (Nov24) | 87.5% | 87.4% | 87.2% |

Using a target precision of 80% as an example, the initial model, trained on data from a different camera, platform, or similar configurations, exhibited relatively poor performance, detecting only about 64% of the ground truth COTS. With the addition of approximately 4.2K COTS annotations from the ReefScan platform, the January 2022 model showed a significant improvement, detecting around 79% of the ground truth COTS. Further enhancement was observed with the June 2023 model, which incorporated an additional ~1.5K annotations and achieved a detection rate of approximately 84%. The final model, which included ~2.5K of additional annotations, demonstrated a modest improvement, detecting around 87% of the ground truth COTS (see Table 6).

### *COTS detection with tracking (per-instance metrics)*

When object detection with tracking is required, the approach must account for linked sequences of bounding boxes across frames, enabling sequence-to-sequence comparison. The tracker's goal is to identify each starfish on the reef for mapping and inclusion in COTS density estimates. A true positive is defined when any predicted bounding box in a sequence sufficiently overlaps a ground truth bounding box, with this method, the predicted sequence is not required to overlap or detect every ground truth bounding box. Instead, the sequence score is assigned as the maximum detector score over the sequence, reflecting the need to observe each COTS instance only once, as multiple observations across frames provide no additional value, as can be seen in the following image with two subsequent frames.

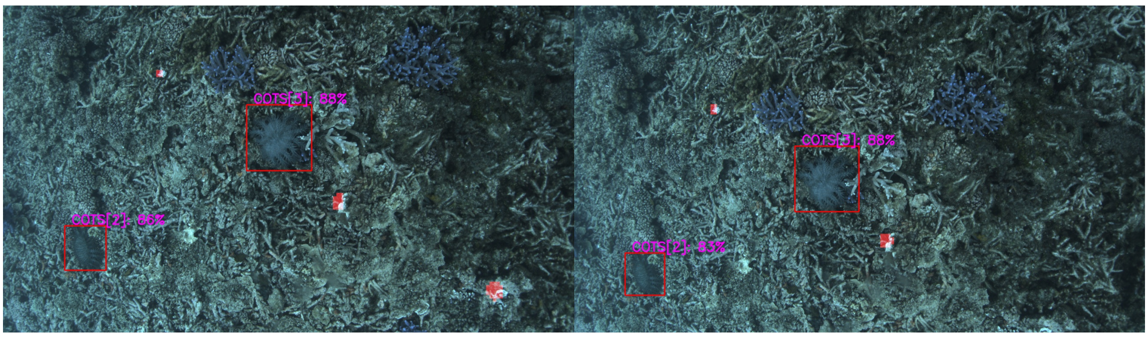


Figure 16. Example of COTS being tracked across sequential image frames.

Table 7 presents the results of COTS tracking for 61 COTS instances in the test set with each instance observed in an average of 16.5 frames. The right panel of Figure 17 illustrates the corresponding precision-recall graph for COTS tracking (per-instance metrics).

Table 7. Model performance for COTS tracking (per-instance metrics).

| Model | Average Precision | Average Recall | Peak F1-Score |
|---|---|---|---|
| Initial | 79.7% | 80.2% | 84.1% |
| Train (Jan22) | 81.6% | 81.7% | 82.6% |
| Train (Jun23) | 90.5% | 90.8% | 90.8% |
| Train Final (Nov24) | 92.7% | 92.9% | 91.7% |

The values in Table 7 are higher than the per-image metric (Table 6) because the model needs to detect each annotated COTS instance only once across all images where it appears. A similar trend is observed as with the per-image statistics, though the comparison is noisier due to the limited number of ground truth instances (~61 only) and challenges in selecting optimal parameters for the object tracker. The final detector and tracker achieve approximately 92% detection of COTS instances with 93% precision which are both well above a level that is considered to be operational (around 80% detection and precision).

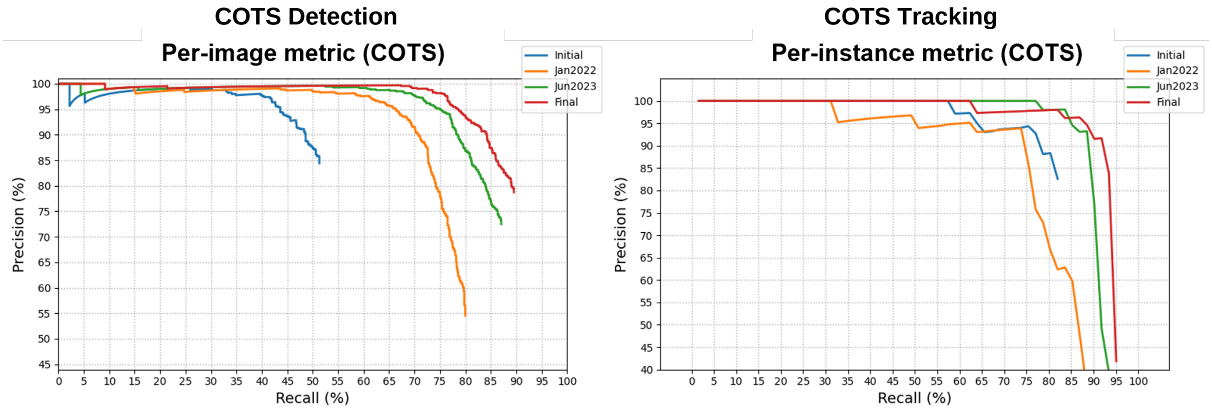


Figure 17. Comparison of model evolution for COTS Detection (left) and COTS Tracking (right)

### 3.4.2 Scars Model

Similar to COTS detection and tracking, we evaluated the scars segmentation model's performance on the test data (see Table 6) using pixel-wise precision, recall, and peak F1-Score. Additionally, we reported the peak Intersection over Union (IoU), a widely used metric in image segmentation, calculated as the ratio of the true positive area to the union of true positive, false positive, and false negative areas. This metric reflects the degree of overlap between the segmented area and the ground truth mask, relative to their combined area.

The results presented in Table 8 were obtained by running the scars segmentation model on the test images. Although the model's performance is lower compared to COTS detection and tracking, it shows a significant improvement, increasing from 20% to 40% Peak F1-Score between the January 2022 and the final training sets. This highlights the evolution of the segmentation model for scar identification.

It is important to note that factors such as water quality, image characteristics, and the quality of ground truth data may contribute to the overall lower performance. Examples of both COTS detection and Scar segmentation are provided in Appendix C.

Table 8. Comparison of segmentation masks within each image independently.

| Model | Average Precision | Average Recall | Peak F1-Score | Peak IoU |
|---|---|---|---|---|
| Initial | -- | -- | -- | -- |
| Train (Jan22) | 11.6% | 11.5% | 20% | 11.5% |
| Train (Jun23) | 35.5% | 35.6% | 40% | 25.2% |
| Train Final (Nov24) | 35.3% | 35.3% | 40% | 25.3% |

### 3.4.3 Benthic Model

Images from the ReefScan platform were manually annotated by a trained ecologist using the 50 classes (Appendix D) with the base model being that developed by the AIMS Long Term Monitoring Program (LTMP) as implemented in the ReefCloud system. The LTMP model is based on over a hundred thousand expert annotated images taken from diver surveys. An 'Unknown' category was used where the images were too deep or where turbidity or other reasons meant that an accurate identification was not possible. This category was amalgamated into the 'Other' functional group.

The F1 scores for the functional groups are shown in Table 9 below. For Hard Coral this is around 73%, which is somewhat low (where above 80% is considered to be in the range of a good operational level model), indicating that further annotation and model development is required. Overall, the model has an F1 score of around 77%. Full details of the F1 scores for each class are shown in Appendix E.

Table 9. F1 scores for the Benthic Functional groups.

| Functional Group | F1 Score (%) |
|---|---|
| Algae | 76 |
| Abiotic | 69 |
| **Hard Coral** | **73** |
| Invertebrates | 66 |
| Other | 85 |
| Soft Coral | 71 |
| Sponges | 91 |
| **Overall** | **77** |

The accuracy of the model can also be assessed by comparing the human-derived benthic scores with the scores from the model test / evaluation data. This data was manually scored to allow for a direct comparison between the two methods. For the main functional groups this is shown in Figure 18 as absolute percent cover values rather than as the percent difference between the observed and modelled. Looking at this metric, the overall percent cover value for the human-based scores for Hard Coral is 19.5%, for the Model this is 17.7%.

This shows that at an absolute level the model and human produce very similar scores when analysing images collected using the towed platform, well within the required level of accuracy. When considering that human-derived manta tow scores for coral cover are collected as percent cover categories (Cat-0 (0% coral cover) | Cat-1-minus (1-5%) | Cat-1-plus (5-10%) | Cat-2 (10-30%) | Cat-3 (30-50%) | Cat-4 (50-75%) | Cat-5 (75-100%); Miller et al. 2018), the results of the benthic ML model are well within the level of accuracy produced using the current manta tow method. For example, for the overall average coral cover both the human and machine methods would give a result of Category-2 (10-30% cover).

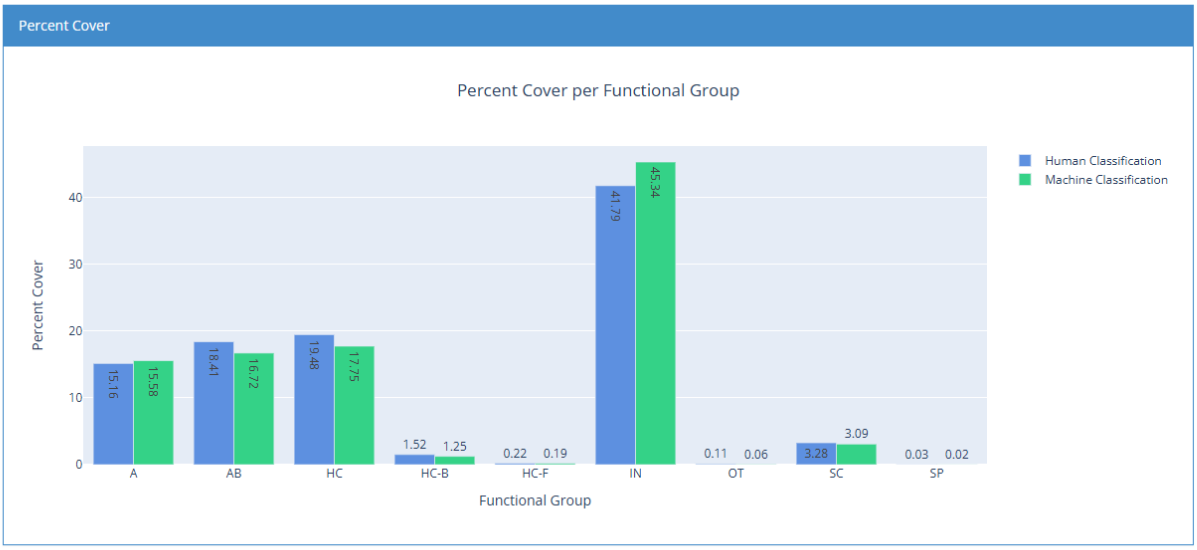


Figure 18. Comparison between human and machine-derived benthic cover scores.

It is also possible to analyse the data by looking at the miss-identifications to see where the model is miss-identifying particular classes. This can be used to target model training to train it to better discriminate between functional groups. One way to do this is to look at the how often each class is miss-identified as another class with this data presented as a matrix of each class against each other class, called a confusion matrix. A confusion matrix is therefore a visual representation of a classification model's performance, comparing predicted outcomes to actual outcomes. The functional level group confusion matrix along with the complete confusion matrixes are shown as Figures 31 and 32 in Appendix E.

The matrix shows that there are misidentifications between Hard Coral and Soft Corals (8%) and Hard Corals and Sponges (16%), so these are places that the model can be actively improved in future work.

# 4. DISCUSSION AND OUTPUTS

## 4.1 Project Outcomes

The project successfully delivered three CSS operational units consisting of the new underwater platform, Machine Larning models, and the associated workflows (as custom software) to the Reef Authority as the new operator of the COTS Control Program, with a system nominally allocated to each Control Operator.

The platform evolved through two protype stages via extensive field trials with the Control Teams and other operators to give the final operational platform. The focus of this development was to improve the reliability and usability of the system as well as some key performance upgrades, including the transition from a single camera to dual cameras to increase the field of view surveyed.

Three Machine Learning models were delivered including a COTS model that had a detection rate of over 90% (over 80% is considered to be operationally effective) as well as an ability to track COTS from image to image of over 93%. The Scars model, developed later into the project, performed at a level of around 40% accuracy reflecting the difficultly in separating Scars from COTS and scars from other sources, a problem that even trained divers have. The benthic model was developed from the AIMS LTMP model re-trained for differing resolution and scale with the model performing just under the operational threshold at around 77% accuracy. While this is less than desired, the current manta tow method bins percent cover values into categories and so the ML model output for percent coral cover, as categories, is very close or equivalent to that of the human observers.

Workflows were developed and implemented to allow for critical data to be delivered to the field teams at points where the information could be actioned. This included a real-time workflow that gives instant feedback while the survey is being undertaken, an end-of-day workflow that informs the next day's work, and an end-of-trip workflow that feeds data into other Institutional data systems.

As with any project there were some objectives that were not fully completed. The terrain following behaviour of the platform has only been implemented in beta form as this proved to be more technically challenging than expected (reefs are complex topographically). The Scars model is still under development which limits its operational use (it tends to deliver too many false-positives due to COTS Scars being confused with other sources of white coral) and the Benthic model still needs work to refine some classifications (such as distinguishing between hard and soft corals). Finally, the workflow component that links to existing data systems is still under development reflecting that these systems themselves are still in development.

The project has successfully delivered new capability and has demonstrated both the application and value of new technologies to existing problems. The work delivers a pathway towards the larger vision of a unified survey method based on images that form a point of truth and that can reduce issues with observer bias so that all collected data contributes equality to decision making and supports actions to sustain and protect the Reef.

## 4.2 Platform Development

The platform developed through two prototypes (Prototype-I and Prototype-II) to the final operational design (Operational Platform). Three units of the Operational Platform were built and delivered to the Reef Authority.

The main development components to the platform included:

### *Move to dual cameras*

The move to Prototype-II included the move from a single planar (directly pointing down) camera with a field of view of around 86 degrees to dual cameras set up as side-by-side (port and starboard) stereo cameras. The driver for this change was to increase the field of view across track to better reflect what a human observer can achieve. The manta tow method (Miller et al, 2018) has a field of nominal view of 10 m (independent of depth) across track which the two-camera system achieves in around 6 m of water.

The set up also allows for the capture of true stereo images in that the cameras are synchronised to capture images at exactly the same time, and which have a degree of overlap along the centre of the track, and so some objects will be visible in both cameras. This allows for stereo photogrammetry to be done where the overlap is used to align the port and starboard images allowing the processing of ortho-mosaic output images.

The level of overlap depends on the height of the cameras above the bottom and how they are positioned. If the intent is to capture as wide a swath width as possible then the cameras can be angled outwards from vertical, up to 30 degrees each side, to increase the across track field of view to around 150 degrees. If the intent is to capture stereo imagery for photogrammetry, then the cameras can be positioned to be both planar with a across track field of view of 100-120 degrees.

Having two cameras introduces issues with the COTS tracking in that the same COTS can not only be seen in a sequence of images *along* the track but can, in some circumstances, be seen in each camera *across* track. This means that the tracking now needs to estimate the future position along and across track to ensure that the same COTS is not double counted. While this is possible, it is yet to be implemented and so the COTS detection is currently only done off one camera.

The field use of the two-camera system is still to be fully operationalised in terms of what types of camera orientations are best in observing particular topography. This will be the focus of more work as dual and multi-camera designs are developed in other related projects.

### *Work in to increase the platform Autonomy*

One key to increasing the effectiveness of the platform is to increase the level of autonomy, especially around collision detection and avoidance. Currently the Operator needs to be vigilant to ensure that any rapid changes in topography are dealt with by manually guiding the platform up and over features. This requires a dedicated person to ensure the safety of the platform. It may be possible to automate the platform to an extent where the Operator only needs to deal with exceptions, where the platform cannot itself take corrective action, and so, if these were rare (for example only requiring the Operator to intervene once or twice a survey), then the roles of the Operator and Observer could be combined. In this scenario a significant resource saving can be achieved as well as making the platform simpler to use.

The project did look to implement a terrain following function to complement the existing depth hold capability, but this was not developed to an operational level in the project. The main issues in operationalising this are around the accuracy of the current altitude sensor, which is sonar based and only perceives a small part of the upcoming terrain, and the need to tune the system so that it is neither too responsive nor not responsive enough. This will take extensive field testing to get it tuned so that it responds appropriately and unfortunately this time was not available in the project as scoped.

One important aspect of developing autonomy capability is that it must add value to the users, that is make deployments easier not harder. It is possible to imagine a system that is too sensitive to potential obstacles and so over-reacts and the reverse, where the platform fails or is inconsistent in its response, making the operation actually more human intensive. Any autonomy therefore has to add functionality to the operators and needs to be driven by what works operationally and so needs to be both appropriate and reliable. This is not novel and reflects work around other autonomous platforms (e.g. self-driving cars) but, if achieved, has the potential to dramatically increase the usability of the platforms and systems.

### *Increase the robustness of the platform while reducing weight and size*

Field use is tough on equipment and people and requires equipment to be robust enough to survive use on moving vessels, often under poor weather conditions, but also light and small enough to be easily handled. These requirements are often in conflict: making systems more robust often makes them larger and heavier which in turn makes them more difficult to use in the confined spaces of small vessels and in rough weather. The engineering challenge was to find a middle ground where systems were as light and easy to handle as possible, but robust enough to survive.

The systems were designed using Computational Fluid Dynamics (CFD) and other modelling approaches to optimise the design to give a balance between strength and size / weight. For example, the plastic frame used in Prototype-I was replaced with a stronger aluminium frame for Prototype-II but, through modelling of operational stress, it was possible to design the new frame to be only marginally heavier than the plastic one (an increase of less than one kg). In the same way the plastic electronics enclosure of Prototype-II was upgraded to alloy for the Operational Platform to increase strength and heat dissipation but, through optimising the design, the increase in weight was only marginal. Overall, the Operational Platform is two kilograms lighter than the Prototype-II platform.

Other approaches were made to increase the utility of the platform. Observations were made of the platform in use by the Control Teams in the field which lead to a re-design of the frame to include front and rear handles, as well as the side handles to make it easier to pass from a tender onto the main vessel. The lens caps were moved to magnetic attachment over the existing Velcro system which is easier to use in a moving vessel. Finally, the attachment points for the tether and data cables were moved to be easier to access as previously they were blocked by the battery, moving the cable points to the side of the platform makes it much easier to attach the tether and data cables.

## 4.3 Comparison with Manta Tow

The primary method currently used to detect COTS is the manta tow method. From Matthews et al (2024) the method can be described as:

> *Reefs are monitored using manta tow surveys, whereby a trained, experienced observer is towed behind a small vessel (5-6m) around the entire reef perimeter in a series of two-minute tows. Each tow is approximately 200m in length with a swathe width of ~10m (~2000 m2 survey area). At the end of each tow, observers record: 1) the estimated categorical coral cover, and 2) the number of COTS observed.* (Matthews et al 2024).

As this method is used extensively in the COTS Control Program it is important to look at how the CSS directly compares. There have been formal comparisons of the number of starfish detected with each method, done as part of the larger CCIP portfolio (see Lawrence et al 2024b), and so here the focus is on the operational components.

Both manta tow and the towed platform methods need COTS to be visible and not hidden for a detection to occur. They both are path dependent in that COTS that are off the path will not be detected but, as the manta tow person can move their head, the current method can potentially pick up some off-track COTS, especially if there are scars or other indications of COTS in their field of view.

To help address this, the platform was changed from a single camera to a stereo dual camera set up where the cameras can be moved or 'splayed' so that the field of view or swath width can be increased. This capability was introduced relatively late in the development, so the difference this makes to the detection of COTS is yet to be tested. As the ML models also track each COTS along-track to avoid counting the same animal multiple times, there is a need to now do this across-track as well, so that if a starfish is seen in one camera it can be tracked to the second camera and again not counted twice. This is yet to be developed.

The benefits of using the ReefScan platform are:

- The system can operate faster that the current manta tow method and without breaks so a greater area can be surveyed with the same resources.
- The personnel involved don't need to be trained ecologists and so this increases the pool of people that can be utilised, reduces the need for specific training (for example taxonomic training) and so better utilises the resources available.
- The method reduces the observer bias and other issues with the manta tow method (for example see Miller and De'ath, 1996) and so produces data that is more consistent across teams, days and locations.
- The data collected form a permanent record of the benthos that can be analysed later to address other questions.
- The collected data gives more precise measures of benthic cover than manual manta tow methods and can be analysed to give other measures such as structural complexity, cover of particular life forms, along with measures of change for repeat surveys.
- As the time spent manta towing is added to the bottom time of divers, the current method can reduce the allowable dive hours of the field team which in turn may reduce the time that can be allocated to Control work, while the use of the Platform does not add to diver bottom time.
- The methods can survey areas deeper than observed by a person undertaking manta tow, and even below what divers can observe, and so maybe useful for investigating deep water COTS or for surveying areas that are still in diver range, but which are not easily visible on a surface manta survey.
- The method removes people from being in the water which may have safety implications in terms of working in areas with known marine pests (although the Control work requires people to dive so this maybe limited).
- The image data forms a robust defendable baseline against which change can be measured as well as a data source for future investigation and analysis.

In terms of operation the two methods are very similar in that they typically use a team of three people (driver + manta-tow person + observer for manta tow and driver + operator + observer for platform tows), tow at similar speeds, and work in similar conditions (sea state, light, etc.). While the logistical benefits are small, the quality and reliability of the data are much greater with the platform as it removes the need for trained observers, removes any subsequent observer bias, and delivers safety improvements. With further development it is hoped to reduce the number of people involved to just two (Driver and Operator) which will result in logistical efficiencies.

## 4.4 Operational models for COTS Detection

The main opportunity that the platform brings to the COTS Control Program is that the logistics and operational envelope for the towed platform (how it is deployed, where it is deployed and under what environmental conditions) can be very different to that of the human manta tow method. Currently COTS surveillance is carried out as part of the Control Teams culling work with around 10% of time spent on determining where to cull and the rest spent on the culling process itself (Reef Authority 2023a, Matthews et al, 2024).

Only those reefs identified by the Reef Authority as Target Reefs (approximately 200 per year) are candidates for culling and the surveillance work is only undertaken where it supports the culling work. This means that adjacent reefs and reefs not under consideration for control work are typically not surveyed.

The use of a platform that does not require a trained ecologist to interpret the data means that the surveillance work can be undertaken by other partners and for other uses, such as long-term monitoring and more generalised COTS detection outside the COTS Control Program's priority reefs (such as adjacent reefs or reefs with high connectivity to outbreak reefs). This introduces the potential for differing logistics and scalability. Monitoring and surveillance can be de-coupled from culling enabling other partners to be involved in the surveillance work and they could therefore expand monitoring and surveillance operations to detect and estimate COTS densities in new locations. Addressing other ecological questions that have been to date challenging, such as deep-water COTS populations, can also be supported using the platform.

## 4.5 Model Development

Training machine learning models to recognize objects works best when there are lots of images of the object. For this project, the object was starfish. It was important to collect images from different angles, with different lighting, and in different settings. One of the main goals of the project was to gather as many useful starfish images as possible to help train the model well.

This proved to be problematic as, while COTS are abundant in some regions of the GBR (such as in the Swains), locally they are often rare or highly aggregated. As a result, despite extensive field work, the project was unable to get the number of images of starfish hoped for. Typically, you would want over 10,000 images to train the model, but less than 1,000 unique COTS were captured and, even with the same COTS in multiple images, this was less than optimum.

A number of actions were taken to remedy this. This included prioritising field work in areas with known outbreaks, targeting parts of reefs where COTS had been previously detected, and undertaking more field work especially in the latter part of the project when more developed platforms were available (such as the Prototype-II platform).

## 4.6 COTS and Scars

The overall performance of the COTS model is good with an F1 score of 92% which is well above the threshold of around 80% for a usable operational model. The main observed issue is with false-positives as, while whole starfish are very distinct, partially obscured starfish tend to look very much like other reef organisms such as sponges and so on. This reflects that the central disk of the starfish is very diagnostic in identification while partial starfish, such as where the starfish is partially obscured and only a few legs are visible, are harder to distinguish from other similar shaped corals and sponges.

The model however rarely misses a starfish (false-negative) and therefore performs well at detecting whole starfish but tends to be overly sensitive (producing false-positives) for partial starfish. As such the model is tuned to not miss starfish even if that results in some false-positives. The model's performance has improved significantly with each training set of images and so it is expected that with more training images, especially of partially obscured starfish, that this would reduce the number of false-positives detected.

For Scars the model performance is lower at around 40% reflecting that Scars caused by COTS, as distinct from other causes of localised white coral, are hard to distinguish. This reflects that, even for trained divers, this is difficult as there are many processes and activities that lead to patches of white coral that look like a scar. These include *Drupella* feeding scars, coral bleaching, natural coral mortality, feedings scars from fish and so on. Scars are a more difficult to visually identify by humans than COTS and this is reflected in the relative scores of the Machine Learning models.

Field operation of the Scar model showed a large number of false-positive detections for Scars. Even after filtering out low confidence detections, a significant number of items remained to review and so, at this stage of the model development, Scars detection can really be thought of as providing a result of none | some | many.

This outcome was shared with the Control Teams in that the model was providing information about the relative density of detected Scars and that this should be used with other information to decide the appropriate response. In many ways this is not that different to the manual method where the real information is whether there are scars present, or not, as a diagnostic for the presence of COTS, rather than having any predefined number of scars or threshold that in itself had meaning.

## 4.7 Corals and Benthos

The Benthic model was developed from the LTMP model used in ReefCloud but re-trained to deal with the images collected by ReefScan versus the diver-based images collected by LTMP (Jonker et al, 2020). The main change is that the images are collected from further away (3-5 m for ReefScan versus 50 cm for divers) under a greater variability in lighting.

The LTMP model is more focused on the information that a diver-based image can resolve, especially around taxonomic information, but uses higher level codes that allow the more detailed information to collapse to a small number of codes (those shown in Table 2).

This level is still more detailed than the standard manta tow category which just includes live coral as one entity (along with dead-coral and soft-coral) (Miller et al. 2018).

The model performance of 73% for hard coral is just under what would be considered good for an operational model (80%) but as the coral cover field estimates are binned into relatively large chunks, the model does not need to be as accurate to still give the same information.

To date the training process has been to identify what is under a fixed set of five points superimposed on each training image with these identifications used to re-train the model. This provides a somewhat random selection of benthic types that get annotated and so does not target rare or easily confused forms. These can therefore be under-represented in the training set used to train the model and hence the model performs poorly for these. The solution is to select points based on targeted life forms and so increase the representation of these in the training data to remove the bias against rare forms.

As there have been a large number of benthic images already collected, an approach of identifying rare and confused forms and targeting these with more intense sampling to increase the training set will result in the model performing better for these forms. In this way the model can be re-trained to address observed deficiencies based on the data already collected. This was not done during the project simply due to time constraints.

## 4.8 Outputs

The project has delivered the following outcomes against the overall portfolio goals:

- Delivery of three certified end-to-end systems to COTS Control Program for operational testing
  - ➢ Delivered as three operational, fully certified systems to the COTS Control Program (nominally one to each Operator) in July 2025, via the Reef Authority, as the Control Program management agency.
- ML model for COTS detection, validated in operational environment.
  - ➢ Delivered and used operationally, running as part of the real-time and end-of-day workflows. Current F1 score is 87% accuracy for COTS detection and 91% accuracy for tracking COTS from image to image.
- ML model for COTS Scars (proof of concept / prototype).
  - ➢ Delivered as an extension to the COTS model, also running in real-time and in delayed mode, currently with an F1 score of 40% indicating that further development is required and reflecting that the source of scars can be complex.

- ML model for benthic habitat, validated in operational environment.
    - ➢ Delivered with an F1 score of 77% for all benthic classes and 73% for hard corals. Currently implemented as a delayed mode model due to limited real-time computing resources and the need to prioritise the COTS and Scars models.
- Development of user-interfaces and operational workflows to support real-time and end-of-day decision making.
    - ➢ Mostly delivered. A tablet-based control interface was developed for the real time system operation and to review potential COTS detections. Custom software was developed for the end-of-day workflows and to link the field data to the ReefCloud system. Further work is required to feed the data from the system into the CCIP Information System and to other reporting dashboards (currently unfunded).
- Integration of system components, validated in operational environment.
    - ➢ Delivered with four trips being undertaken with the Control Teams to integrate the system into their logistics and operational methods and a further four trips on the QPWS vessels including two utilising the final Operational Platform.
- Training of COTS Control Teams in operation of system, including provision of SOP/manual.
    - ➢ Partially delivered. Four technology transfer trips have been done with two of the providers (two each), manuals and Standard Operating Procedures (SOPs) have been developed and will be delivered as part of the final systems. It is planned to continue the joint work with the Control Teams to complete the training and technology transfer.
- Annotated image library.
    - ➢ Delivered, these are available as open access files from the CSIRO repository, see the link in Section 8.2.
- Knowledge and recommendations on use of technology, including ML model accuracy, use of system in operational environment.
    - ➢ Partially delivered. The use of the model outputs, especially given that the detection levels that trigger responses are so low, have been socialised with the COTS Control Teams to ensure they understand the current limitations of the models and how the data can be used operationally to drive the COTS Control work.
- Data from field surveys.
    - ➢ Delivered, these are available from AIMS, the metadata record that describes the data including how to access it is given in Section 8.1.

## 4.9 Control Team feedback

The Control Teams were exposed to the CSS at a number of points in the development. A total of four trips were done at sea with the Control Teams with the various prototype systems to allow them both to use the systems and to give direct feedback about how they can be better optimised for the work being undertaken. A final formal review was done with the Control Teams to finalise the design of the final Operational Platform that was delivered.

The general feedback included:

- The users were generally happy with the system and the capability it delivered and there were no real issues that they saw would prevent them from using the equipment (no 'show-stoppers').
- They wanted the equipment to be smaller, especially the Ground Control Station, and for the platforms and equipment to have more handles to facilitate safe handling at sea.
- The Scars model produced a lot of false-positives which they found difficult to deal with in the field (often too many to review in the time available or it became a distraction to the operation of the equipment), and so filtering was added where the user could effectively not see detections that had a low confidence score (typically less than 50%).
- They found the end of day software to be somewhat confusing and so this was re-designed to be more straight forward.
- They wanted better transport cases and an easier way to clean the platform at the end of each day. A Pelican transport case was purchased for each system to facilitate safe transport and a method of washing the platform developed and implemented.
- Due to turn over of staff they need training materials that they could use on each trip to bring new staff up to speed with the system. These have been developed and will be refined with the Control Teams outside of this project.

The final feedback was around field strategies to best to use the platforms and the new capacities that they bring. The questions were around if the CSS should be deployed in the same way as the existing methods or if there are other deployment strategies that may make more sense given the capability of the system. For example, is it best to schedule the detection work in time currently allocated to diver downtime (as the CSS doesn't involve any in-water activity it does not add to the diver's bottom time) or is it best to allocate a block of time and use that to do the detection work and then come in afterwards with the Control work.

At this stage the response was to do the surveys as per the existing logistics, but it would be interesting to model this and see the relevant returns for differing deployment strategies.

# 5. RESEARCH SYNERGIES AND NEXT STEPS

## 5.1 Research Synergies

Within the CCIP portfolio the CCIP-D-04 COTS Surveillance System project delivers to the following (see Figure 3):

- Outputs: Expanded toolbox for COTS detection with plans for deployment
    - Outcomes: Improved detection and monitoring

It also directly delivers to the D-01 COTS Monitoring Design and D-02 Tool Comparison projects in the Detection sub-program with two field trips being undertaken as part of the D-02 project. The technology developed in this project is key to delivering the baseline and early warning detection strategy that has been developed under CCIP-D-01. Critically, it will significantly enhance the monitoring capacity on reefs by allowing the concurrent estimation of COTS and coral on transects and remove any concerns around observer bias.

Outside the CCIP portfolio, the work has application across a range of other reef related projects and programs. As a source of monitoring data, the project naturally fits within the Reef Knowledge program and the Reef 2050 Integrated Monitoring and Reporting Program (RIMReP) information systems with the potential to provide a new uniform set of benthic composition and COTS / Scars occurrence data across more reefs and regions. The data, once validated, could contribute to the AIMS Long Term Monitoring Program and to the QPWS survey work, especially around reactionary or response surveys.

The collected data forms a source of information that can be used to address identified needs in the Reef 2050 Plan, which, released in 2015, is a combined State and Federal Government strategy to improve the Reef's health and resilience (https://www.dcceew.gov.au/parks-heritage/great-barrier-reef/protecting/reef-2050-plan). Appendix F shows a mapping of the coral reef parameters and how the project work could potentially contribute. Some of these align with the COTS work (e.g. coral cover and starfish numbers) while others are more aligned with other management objectives.

The larger vision is that unifying the survey method around an image-based data workflow means that issues with user training and observer bias are reduced to the point that, once validated and calibrated, the platform data collected by all parties has equivalent value in making and supporting management decisions. This means that multiple contributors can work across the GBR region to collect data that has utility in delivering management decisions. Scaling the number of data sources is the only way to ensure that areas currently not surveyed, such as the far north, are included in our understanding of the reef.

## 5.2 Next Steps

The project work acts as a demonstrator of new capacity in monitoring, detection and surveillance. The next step is to refine the work undertaken to make it easier to use for the COTS Control Teams, especially focusing on making the platform more autonomous, and then to drive activities to increase the uptake and use of the system via a knowledge transfer process.

The next steps are all about increasing the value from the system:

1. Continue to develop the ML models to increase the robustness of the automated analysis and therefore the value these bring to the Control Teams.
2. Refine the platform to increase the level of autonomy so that it becomes easier to use and to reduce the operational overhead.
3. Model new logistics to see if there are new modes of operation that may increase the effectiveness of the Control Teams (desktop study).
4. Adapt the system for de-coupled monitoring and surveillance, targeting areas such as deep reefs, night-time surveys, surveys in currently un-monitored areas and surveys using new participants such as Tourism Operators and Traditional Owners.
5. Work with the other parts of CCIP, and the general reef community, to adapt and apply the capability to deliver capacity and data to support improved management.
6. Work with the various information systems to ingest the information from the automated analysis as well as help develop systems to deal with the core image data being collected, so these become an enduring resource for future work.

The initial work is refining the systems to make them more effective in the field. This includes the physical design of the systems (making them lighter, more robust, easier to deploy, easier to service and longer battery duration), support of the equipment via training and materials, refinement of the ML models to increase their accuracy, and general support of the field teams in delivering data and information to the required end points. The goal at this level is to improve the effectiveness of the field component of the Control Teams.

The next level is to apply the developed capability to other areas of the portfolio around the idea of delivery monitoring and surveillance capacity de-coupled from the Control work. This maybe via other partners, such as QPWS and Traditional Owners, and could be targeted at more general scenarios (such as getting data from areas currently not surveyed) or more specific questions such as deep-water COTS, ecological responses to environmental changes, and so on.

The final package is around data management and in ensuring there is infrastructure in place to facilitate access and use of the data, from the analysed outputs to the raw image data.

The project has a number of elements that need additional work and so form an obvious set of work packages for future funding.

### *Developing the ML Models*

The first is to continue the development of the ML Models through additional model training with new images collected by the delivered platforms. This is a low-risk high-outcome activity but one that will be critical in operationalising the work. The aim would be to get the Benthic models closer to the 80% level, especially for hard corals, to get the Scars models over 60% and on their way to 80% (this could be ambitious given the known issues in identifying Scars) and to target partial or obscured COTS to reduce the number of false positives generated.

As surveys can occur at different times of the day, under varying weather conditions, or across diverse reef locations, it is important that the model has enough images of COTS and Scars to reflect this level of variability. The current models may not fully encompass this level of variability and so expanding the dataset with annotations from diverse conditions can help refine the models and enhance their performance. Another part of the model development and validation is un-picking the scar detection work to be better able to distinguish between COTS Scars and other sources of scars.

### *Increasing the level of platform autonomy*

The second work package is to refine the platforms so that they are easier to use and deliver more value to the Control Teams. The current platforms are operational but any work to increase the level of autonomy and functionality will deliver benefits to the Control Teams.

The main work will be to increase the level of autonomy by implementing altitude hold in an operational manner (currently it is in development mode only) so that the platform is able to better follow the terrain and to include more forward-facing sensors (currently only a single sonar sensor but with the potential to include forward-facing stereo navigation cameras), to better anticipate the terrain that is upcoming. This increased situational awareness may give the operator more information via warnings of impeding issues but may also allow for the platform to take independent action to avoid a collision or to maintain the correct altitude. This will reduce the demands on the Operator and so may allow for the current separate Observer and Operator roles to be combined.

### *Increase the uptake and use of the data*

The third area is to complete the data workflow work to deliver the derived information to reporting and other information systems. Many of these are still under development and so, as this capability developed in the host agencies, the system can be updated to directly deliver to these.

A direct flow on from this is the need to store, manage and value-add to the images that will be collected by the platforms. For many reefs the initial surveys will be the first permanent image-record of the reef and so the images form an important record or baseline. The ability to process images in the future using new ML models and approaches means that the raw data forms an important resource for understanding and documenting change.

*Delivery of capacity to the coral reef community*

Looking at the portfolio of projects and work undertaken in the COTS Control space there are a number of other opportunities or synergies. These include deep reef surveys to understand COTS at depth, understanding the response of COTS to warm water events (such as do COTS go deeper during such events), to do surveys at night and so understand night-time numbers and locations, to survey in areas currently not surveyed, and to survey in key locations, such as the initiation zone, that may inform other control decisions. The core capacity that the CSS delivers has application in collecting data to answer many of these questions.

The capability the system provides has a number of potential uses outside COTS detection including the development of Machine Learning based visual surveys for other target species, testing new methods for detection, such as UV light surveys, collection of oceanographic and other data via sensors on the platforms, and the collection of baseline image data to document reef state.

# 6. MANAGEMENT IMPLICATIONS AND IMPACT

Currently a number of agencies and programs undertake surveys for corals and starfish along the GBR, but the data collected often differ, lack standardisation, or are designed to address specific questions, making them unavailable for broader use. At the same time, large parts of the reef, such as the far north and the Torres Strait, have sparse or little routine monitoring or surveillance (mostly due to cost and logistical reasons).

The CSS, by collecting high quality images, effectively allows for the separation of the data collection and analysis steps allowing each step to be done by separate resources at separate times. This reduces the need for multi-skilled personnel (such as field personnel who also have higher level ecological / taxonomic skills) increasing the pool of people that can undertake the work. The use of Machine Learning, both in real time and as part of post-survey workflows, allows for the taxonomic and ecological knowledge to be encapsulated in a model that can be made available to everyone.

The project looks to deliver a capability that, once validated against existing methods, will provide a standardised way of surveying reefs that ensures the data from all surveys, no matter who collected it, has value in supporting management decisions. The method reduces the number of people in the water, is suitable for areas with marine pests, requires technically trained but not scientifically trained personnel, is suitable for use by Traditional Owners and Tourism Operators, and delivers a permanent record of the reef at that time suitable for a range of current and future analysis methods.

The direct application of this will be to:

- Improve on-water operations and ensure that all surveys completed contribute to the pool of information available to make management decisions.
- Increase the set of tools that can be used to monitor corals and starfish especially in areas with marine pests or where human-based surveys are problematic.
- Allow new partners to contribute to the set of surveys completed, with the goal of increasing the spatial and ecological footprint of the reefs surveyed each year.
- Provide the data to facilitate a more informed reef prioritisation process.
- Provide data from currently unsurveyed habitats to address ecological questions that in turn inform the Integrated Pest Management framework and provide input to the COTS Strategic Management framework.
- Provide an information resource that can contribute to future understanding about the status and trends in reef health and function.

The project work provides a capability across the portfolio that can be used to improve monitoring, both in documenting the status and trends of reefs and to track the impact of the control work. In particular it allows for other partners to contribute standardised high-quality data to information systems to support better management and to support new ecological studies.

The outcome will be an increased ability to detect and control / supress COTS, to monitor new areas and so detect incipient outbreaks, to deliver capacity to answer ecological questions about starfish responses and behaviour, to involve new partners, such as Traditional Owners and Tourism Operators, and to provide data to support future studies into the patterns of outbreaks and recovery.

# 7. ACKNOWLEDGEMENTS

The authors and Project Team would like to acknowledge the help and enthusiasm of the COTS Control Teams, especially the teams of the RV *Odyssey*, RV *Flying Fish*, RV *Infamis* as well as the AIMS crew on the RV *Cape Ferguson*. We would also like to thank the QPWS crew and staff of the RV *Reef Ranger* and RV *Reef Resilience* for help with the field work and for their enthusiasm for the work.

We would also like to thank Eric Fisher of GBR Biology (https://www.gbrbiology.com) for access to the Reef Magic pontoon on Moore Reef and his help in running the initial testing and trials.

Thanks also go to Dr Russ Babcock, Melissa Boer, and Kylie Maguire for annotating and checking the COTS and Scars images and to Sina Strahl and Nico Briggs for their work in annotating the benthic images.

# 8. DATA ACCESSIBILITY

## 8.1 Raw / Source Data

A metadata record for the raw imagery collected as part of the project is available at:

https://tsv-apps.aims.gov.au/metadata/view/ea928345-e456-4b0e-90c1-8806ae32c822

The raw data is held on the AIMS internal data servers (as detailed in the metadata record) and can be accessed by contacting the AIMS Data Centre at: adc@aims.gov.au.

The raw data consists of a number of surveys or sequences and then the images collected as part of that survey. The raw images have the GPS location (decimal degrees, WGS84) and date/time (in UTC) along with the depth and altitude of the camera embedded in the header of the images as JPEG EXIF data (V2.3) which can be read by any image display program.

Where stereo transects were conducted the files are split into directories for Camera-1 (Cam1) and Camera-2 (Cam2) with Cam1 being the port camera and Cam2 being the starboard one.

A full log of the survey with the details of each image collected is included as a CSV file in the image directory.

## 8.2 Annotated Data / Annotations

The annotated data for COTS and COTS scars is held by CSIRO and can be accessed at:

https://doi.org/10.25919/03a7-hn83

For COTS detections, these are stored as the image name, the bounding box in the image where experts have manually annotated the starfish. For COTS Scars, the annotation is done using the Segment Anything Model (SAM) embedded in CVAT to generate a polygon mask describing the pixels that belong to the scar. The available annotated data is in COCO format, a universal format used for annotations and model development in computer vision.

The annotations and the image library used to train the models are available from CSIRO at the link above. For the benthic model the images are available from AIMS as per the metadata record.

# 9. REFERENCES

Babcock R, Dambacher J, Morello E, Plagányi É, Hayes K, Sweatman H, Pratchett M. (2016) Assessing Different Causes of Crown-of-Thorns Starfish Outbreaks and Appropriate Responses for Management on the Great Barrier Reef. PLOS ONE 11: e0169048.

Bainbridge S and Gardner S. (2016) Comparison of Human and Camera Visual Acuity - Setting the Benchmark for Shallow Water Autonomous Imaging Platforms. Journal of Marine Science and Engineering 4, no. 1: 17. https://doi.org/10.3390/jmse4010017.

Bozec Y-M, Hock K, Mason R, Baird M, Castro-Sanguino C, Condie S, Puotinen M, Thompson A, Mumby PJ. (2022) Cumulative impacts across Australia's Great Barrier Reef: a mechanistic evaluation. Ecological Monographs 92(1): e01494. https://doi.org/10.1002/ecm.1494

Chandler J, Burn D, Caballes C, Doll P, Kwong S, Lang B, Pacey K, Pratchett M. (2023) Increasing densities of Pacific crown-of-thorns starfish (*Acanthaster* cf. *solaris*) at Lizard Island, northern Great Barrier Reef, resolved using a novel survey method. Sci Rep 13, 19306 https://doi.org/10.1038/s41598-023-46749-x.

De'ath G, Fabricius K, Sweatman H, Puotinen M. (2012) The 27–year decline of coral cover on the Great Barrier Reef and its causes. Proceedings of the National Academy of Sciences,109(44), 17995-17999.

Emslie M, Ceccarelli D, Logan M, Blandford M, Bray P, Campili A, Jonker M, Parker J, Prenzlau T, Sinclair-Taylor T. (2024) Changing dynamics of Great Barrier Reef hard coral cover in the Anthropocene. Coral Reefs 43:747-762.

Fernandes L, Marsh H, Moran P, Sinclair D. (1990) Bias in manta tow surveys of Acanthaster planci. Coral Reefs 9, 155–160. https://doi.org/10.1007/BF00258228.

Fletcher C, Bonin M, Westcott D. (2020) An ecologically-based operational strategy for COTS Control: Integrated decision making from the site to the regional scale. Reef and Rainforest Research Centre Limited, Cairns (65pp).

Goutte C and Gaussier E. (2005) A Probabilistic Interpretation of Precision, Recall and *F*-Score, with Implication for Evaluation. In: Losada DE, Fernández-Luna JM. (eds) Advances in Information Retrieval. ECIR 2005. Lecture Notes in Computer Science, vol 3408. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-31865-1_25.

Jonker M, Bray P, Johns K, Osborne K. (2020) Surveys of benthic reef communities using underwater digital photography and counts of juvenile corals. Long Term Monitoring of the Great Barrier Reef - Standard Operational Procedure Number 10, 2nd ed. Australian Institute of Marine Science, Townsville, Australia. (36pp) https://doi.org/10.25845/jjzj-0v14.

Lawrence E, Foster S, Gladish D, Matthews S, Williamson D, Uthicke S, Doyle J, Pratchett M, Bainbridge S, Armin MA, Crosswell J. (2024a) COTS Monitoring Design: sample design for science and management decisions. A report to the Australian Government by the COTS Control Innovation Program (59pp).

Lawrence E, Foster S, Williamson D, Matthews S, Schultz D, Pratchett M, Uthicke S, Bainbridge S, Doyle J, Doll P, Armin MA, Ewels C. (2024b) A comparison and calibration of COTS and coral monitoring tools. A report to the Australian Government by the COTS Control Innovation Program.

Matthews S, Williamson D, Beeden R, Emslie M, Abom R, Beard D, Bonin M, Bray P, Campili A, Ceccarelli D, Fernandes L, Fletcher C, Godoy D, Hemingson C, Jonker M, Lang B, Morris S, Mosquera E, Phillips G, Sinclair-Taylor T, Taylor S, Tracey D, Wilmes J, Quincey R. (2024) Protecting Great Barrier Reef resilience through effective management of crown-of-thorns starfish outbreaks. PLOS ONE 19(4): e0298073.https://doi.org/10.1371/journal.pone.0298073.

Miller I and De'ath G. (1996) Effects of training on observer performance in assessing benthic cover by means of the manta tow technique. Marine and Freshwater Research 47, 19-26. https://doi.org/10.1071/MF9960019.

Miller I, Jonker M, Coleman G. (2018) Crown-of-thorns starfish and coral surveys using the manta tow technique. Australian Institute of Marine Science, Townsville.

Reef Authority (2023a) Culling Guidelines - Crown-of-Thorns Starfish Control Program. Great Barrier Reef Marine Park Authority, Townsville. (17pp).

Reef Authority (2023b) Integrated Pest Management Guidelines - Crown-of-Thorns Starfish Control Program. Great Barrier Reef Marine Park Authority, Townsville. (19pp).

Sivakumar M, Parthasarathy S, Padmapriya T. (2024) Trade-off between training and testing ratio in machine learning for medical image processing. Peer J Computer Science 10: e2245 https://doi.org/10.7717/peerj-cs.2245.

Tychsen-Smith L and Petersson L. (2017) DeNet: Scalable real-time object detection with directed sparse sampling. Proceedings of the IEEE international conference on computer vision, pp 428-436.

Vercelloni J, Mengersen K, Ruggeri F, Caley J. (2017) Improved Coral Population Estimation Reveals Trends at Multiple Scales on Australia's Great Barrier Reef. Ecosystems 20, 1337–1350 (2017). https://doi.org/10.1007/s10021-017-0115-2

Westcott D, Fletcher C, Babcock R, Plaganyi-Lloyd E. (2016) A Strategy to Link Research and Management of Crown-of-Thorns Starfish on the Great Barrier Reef: An Integrated Pest Management Approach. Report to the National Environmental Science Programme. Reef and Rainforest Research Centre Limited, Cairns (80pp).

# 10. CONTACTS:

## 10.1 AIMS Contacts:


| | | |
|---|---|---|
| Scott Bainbridge | S.Bainbridge@aims.gov.au | 07 4753 4444 |
| Geoff Page | G.Page@aims.gov.au | 07 4753 4444 |
| Greg Coleman | G.Coleman@aims.gov.au | 07 4753 4444 |
| De'vereux Harvey | D.Harvey@aims.gov.au | 07 4753 4444 |


## 10.2 CSIRO Contacts:


| | | |
|---|---|---|
| Mohammad Ali Armin | Ali.Armin@data61.csiro.au | 02 6246 4001 |
| Jeremy Oorloff | Jeremy.Oorloff@data61.csiro.au | 02 6246 4001 |
| Emma Lawrence | Emma.Lawrence@data61.csiro.au | 1300 363 400 |
| Brano Kusy | Brano.Kusy@data61.csiro.au | 07 3327 4444 |

# APPENDIX A – PLATFORM DESIGN AND DEVELOPMENT

## *Platform Design (final Operational Platform)*

The platform is a small, towed sled measuring around 1.1 m long by 60 cm wide by 50 cm high (see Figure 19). The top of the unit has four thrusters, one at each corner, that control the platform attitude (roll and pitch) and depth, a set of two side handles per side with additional handles front and rear. On the top are side-rails for attaching the bridle that provides the physical connection to the tether, as well as the connector for the communication component of the tether. Blue arrow shows the direction of towing.

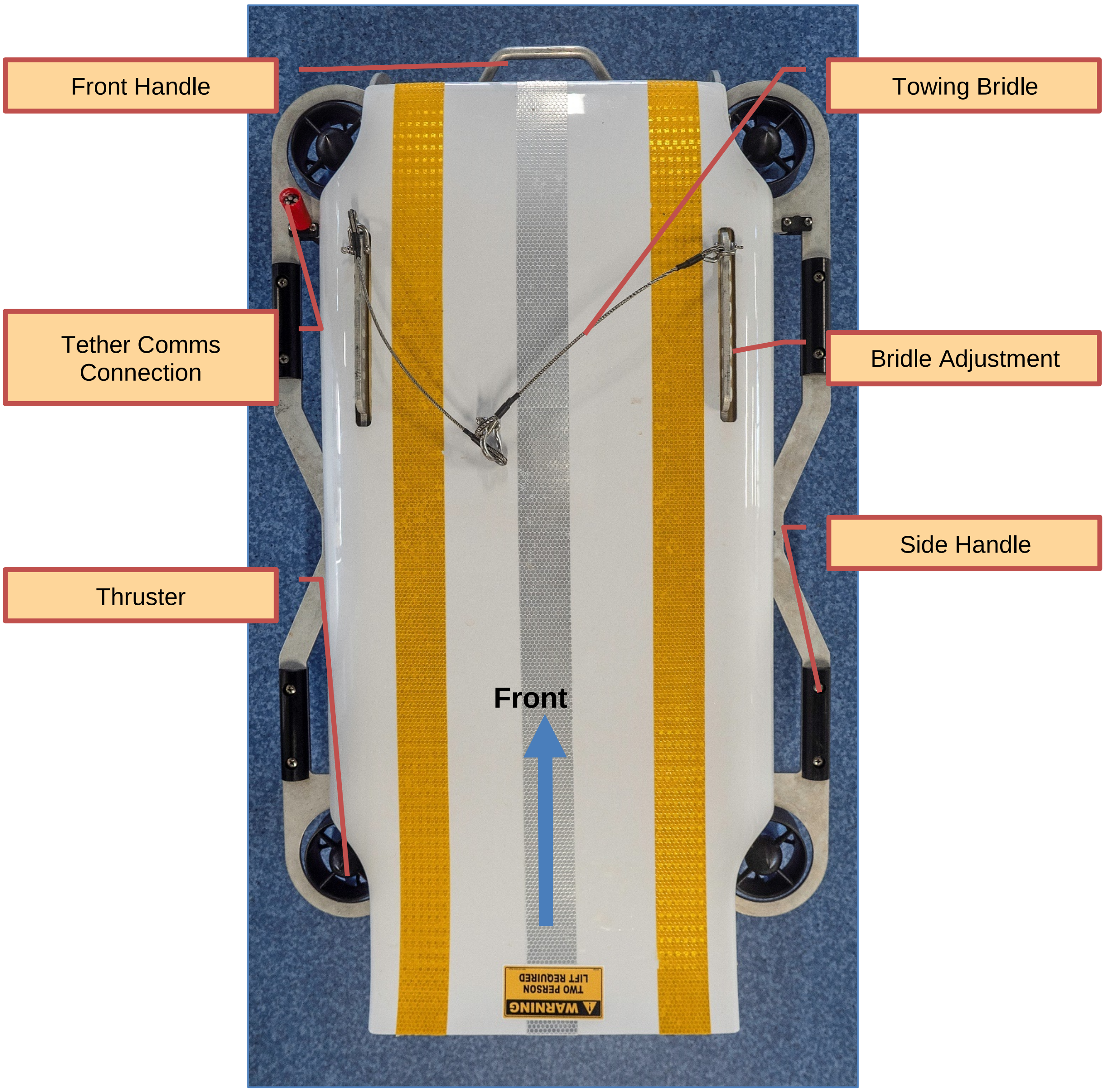


Figure 19. ReefScan Deep Platform (Operational Platform) – Top View

On the bottom of the platform is a large electronics enclosure that includes the navigation camera, lens and port at the front as well as the platform computer and other electronics. The battery that sits behind the enclosure with the thrusters at each corner, an altitude sensor near the front and the two scientific cameras located in the middle of the platform.

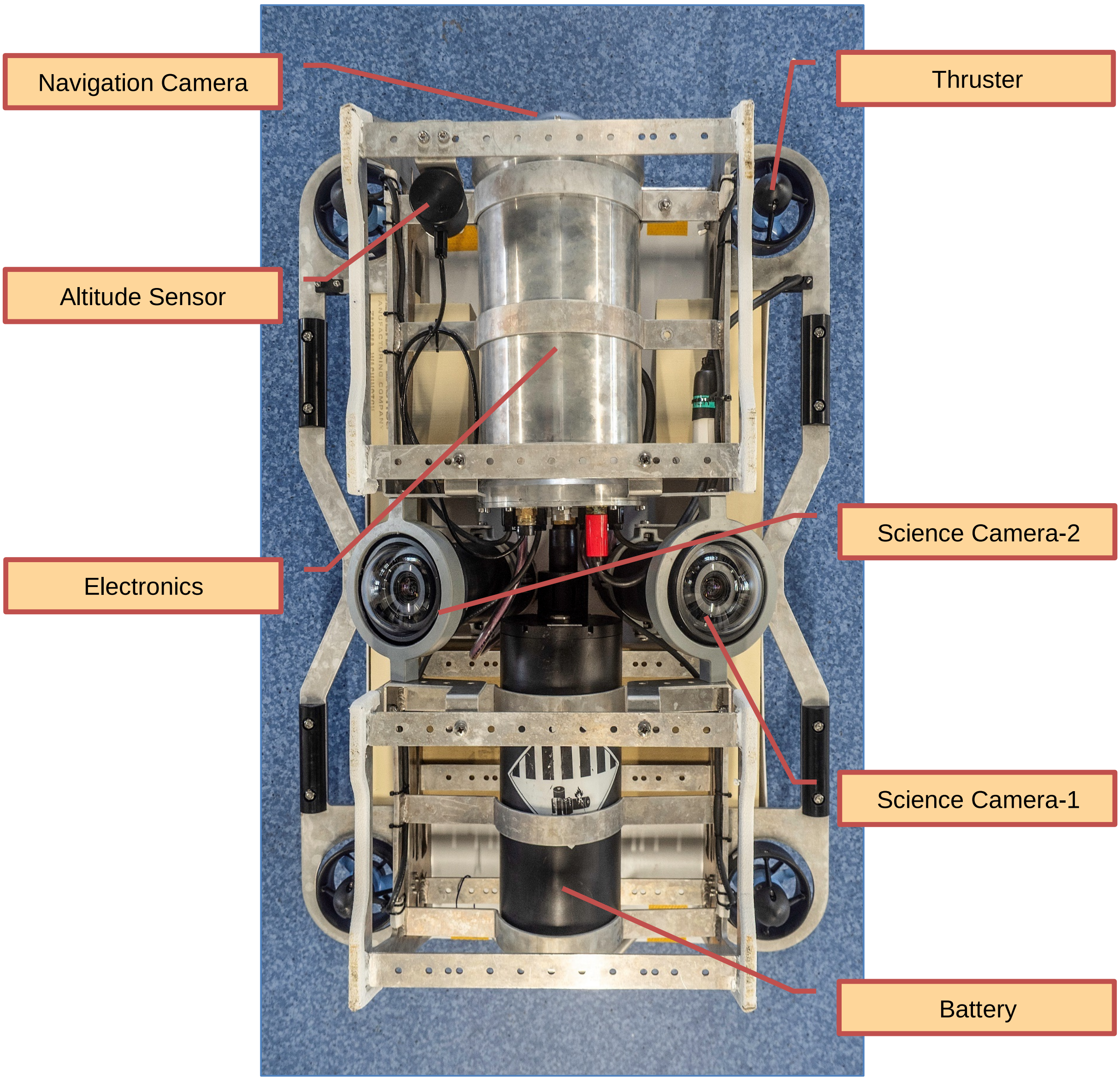


Figure 20. ReefScan Deep Platform (Operational Platform) – Bottom View

### *Software Architecture*

The control computer is based on the QGroundControl open-source software running on the surface computer in the GCS, which communicates with the platform-based autopilot over the tether. This is currently run as is with the only customisation being the mapping of button functions in the GCS.

The ReefScan system software is based on the ROS system, a message-based system where a set of self-contained nodes undertake activities based on the messages they receive. One node is the Control Node which has responsibility for overall control, but nodes operate independently based on what messages they receive.

This makes the system modular so that new capability can be added without needing to refactor or re-code the other nodes. It also allows for other contributors to develop and test nodes independently of the ReefScan system with the ROS system allowing for full capture of previous sessions (via bag files) which can be played back to test new capacity and nodes. The basic node structure is shown in Figure 21.

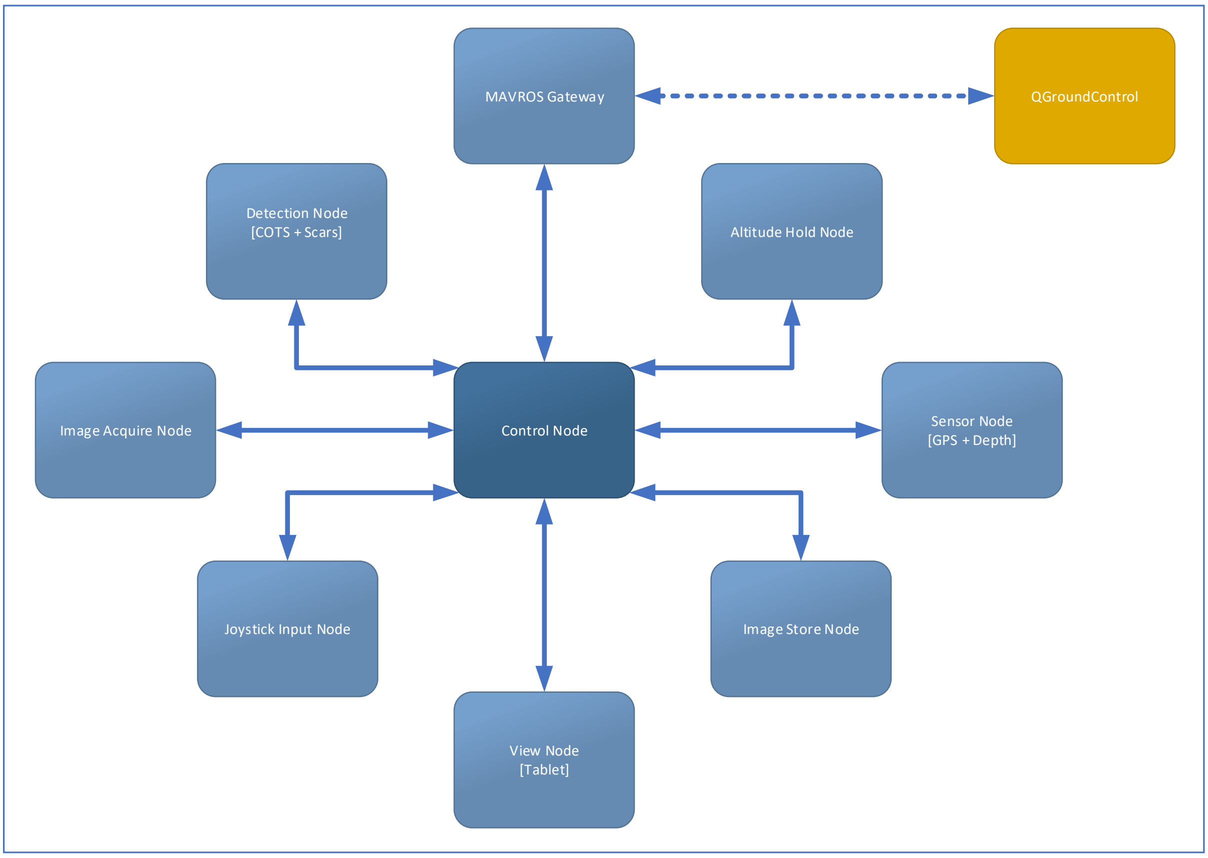


Figure 21. ReefScan-Deep ROS Node structure

The QGroundControl Station control system and the ReefScan analysis systems communicate via a Mav-ROS gateway over an ethernet network connection allowing for the QGroundControl station to send data to the ReefScan system (such as date/time and location information from the GPS and data from the depth and altitude sensors) and for the ReefScan system to send information to the QGroundControl system, such as input to the auto-pilot from the ROS based altitude hold node.

The Control Tablet communicates to the rest of the system via an MQTT server located on the QGroundControl computer in the surface GCS system, via a wireless hot spot. Through the MQTT server, the tablet can subscribe to messages from the ROS system via a ROS MQTT_bridge Node and so can respond directly to detections processed in the detections ROS node (Figure 21).

### *Power Systems*

The underwater platform is powered off a local bespoke 14 Volt 17.5 Amp Hour Lithium Iron Phosphate LiFePO4 battery system provided by Master Instruments in an AIMS designed housing that includes a vent plug to ensure no build-up of gases on charge and discharge. The batteries are “wet-swappable”, in that they use Sub Conn wet-mate connectors and so are safe to change in a wet environment, such as in a small tender (in fact they can be swapped completely submerged).

The batteries are rated to 50 m depth, tested to 70 m, and provide up to 17.5 Amp Hours of power which is enough to run the platform for 2-3 hours. The batteries can be charged via a commercially available charger with a charge time to full of around four hours. Two batteries are provided with each platform giving a run time of four to six hours per day.

The surface GCS is powered off an internal battery that provides 2-4 hours of run time, but which can be powered off the tender motor / alternator or off an external battery. This arrangement has been selected so that the weight of the GCS is reduced while allowing a range of external power sources to be used to supplement the internal battery.

Note that the tether does not provide power to the underwater platform, the underwater platform is powered off its local battery only.

### *Tether*

The tether is a 50 m copper-based cable from Blue-Robotics and is rated to 175 kg breaking strain (typical pull load of the ReefScan Deep is 30 kg) and has four twisted pairs of copper connectors along with a Kevlar structural component. A VDSL unit is used to run ethernet over this link with the link going into an ethernet switch at both ends into which both the control and analysis systems connect.

### *Control Systems*

The system uses the open source ArduSub auto-pilot system that controls the attitude of the platform with a number of pre-set modes including full manual control (including for pitch and roll) or stabilised mode where roll and pitch are automatically corrected for. This leaves the operator only needing to set the depth via the joy-stick controllers on the GCS. In depth-hold mode the platform maintains the current depth using data from the pressure sensor.

A custom altitude hold, where the unit tracks a consistent altitude above the bottom (or terrain hold capability), has been added but this is currently in beta form. The altitude pinger only samples a small area of the upcoming bottom so an amount of pre-processing is required to extract the altitude of the upcoming benthos and then to react. The altitude hold is implemented as a ROS node in the ReefScan part of the platform; this communicates via the Mav-ROS bridge to the autopilot.

The system currently needs an operator who monitors the system making sure that it remains at the required height and deals with any sudden changes in topography as the unit is towed over the reef. The altitude hold is currently not able to deal with very sudden topography changes, such as coming into a wall, and so the operator needs to be monitoring the system to adjust the depth as needed. In a similar way, sudden drop-offs also need to be manually driven although the altitude hold system will catch up and correct the altitude. This will be an area of future work to better sense the upcoming environment and adjust before the platform is at risk of hitting the bottom.

### *Ground Control Station (GCS)*

The platform is controlled via a Ground Control Station (GCS). The GCS consists of separate but linked systems for the control of the platform and control of the payload. The platform has a number of operational modes from fully manual, stabilised flight (where the auto-pilot controls roll and pitch) to depth hold and the developmental altitude hold. The operator can see the upcoming terrain via a forward-facing high-definition camera and use this, with information from the observer, to ensure that the platform avoids obstacles and navigates any sudden changes in topography. See Section 2.6 for the anticipated operational roles required to deploy the system.

The GCS has its own battery supply with feeds for powering off the tender outboard or from external batteries or solar panels. It has displays for both the navigation and payload cameras, a keyboard to interact with the control computer (which runs Windows 11) and joy-stick controls to control the platform including a gain control that increases or decreases the reaction of the system to input, and an up and down joystick control for adjusting the altitude / depth.

### *Imaging Systems*

There are two imaging systems on the platform. The first is a forward-facing navigation video camera that is used to assist the operator to drive the platform, in particular to see upcoming terrain so that the height above the bottom can be adjusted to avoid collisions.

The second is the main science imaging payload. These consist of a separate pair of downward facing machine vision (Allied Vision Alvium 1800 U-1620 with Sony IMX542 sensor) USB connected still cameras that capture 16MP colour images at a frame rate of 3-4 frames per second. This ensures that at a forward speed of 2-3 knots complete coverage is recorded along track with the cameras arranged to give around 30% side to side overlap.

The cameras are paired to Fujinon (Fujifilm) 8 mm equivalent lenses (model CF8ZA-1S, 86° side-to-side field of view) and are mounted behind 4" BK7 optical glass domes which ensures the images are sharp from corner to corner. The lenses are manually set to a focus distance of 5 m with an aperture of f8 to give a depth of focus from 3 m through to 8 m.

The general field of view for each camera is that the swath wide or side-to-side field of view is the same as the camera altitude or distance between the camera and the bottom. If the platform is 5 m off the bottom, then the field of view for each camera is 5 m wide. Depending on the side overlap of the cameras, the total combined field of view across the two cameras for 5 m of altitude can be up to 10 m wide.

The use of Machine Vision cameras means that they will last effectively indefinitely (they have no moving parts), and they are able to be fully integrated into the ROS architecture. The downside is that the images are raw files from the sensor and so need to be pre-processed to bring out the image detail for human based analysis, although the ML models are trained on the raw images.

### *Analysis Systems*

The analysis system sits separately to the control systems and consists of two computers; one located in the platform and one in the surface Ground Control Station. Both computers run Nvidia Jetson GPU cards allowing for image processing and for the Machine Learning to occur on either computer. The distributed nature of the ROS system, where messages are visible to all nodes that subscribe to them, means that the processing can be easily split between the platform and GCS computers as required. The Node structure is shown in Figure 21.

The platform computer is responsible for the initial processing and storage of the images and sends a scaled down version of the collected image to the top computer over the tether. The top computer then implements the Machine Learning models, as a ROS Node, and then sends the model outputs as messages. The View Node picks up these messages and passes them to an MQTT server running on the GCS which makes these messages available to the Tablet-App that subscribes to the same server, via a wireless hot spot running on the GCS.

The tablet is then able to display the detection and the original images via the MQTT server bridge to the platform ROS system and then feed-back any user responses to confirm the validity of the detection. The tablet also allows for full control of the payload camera via the MQTT server link. This is shown in Figure 22 below.

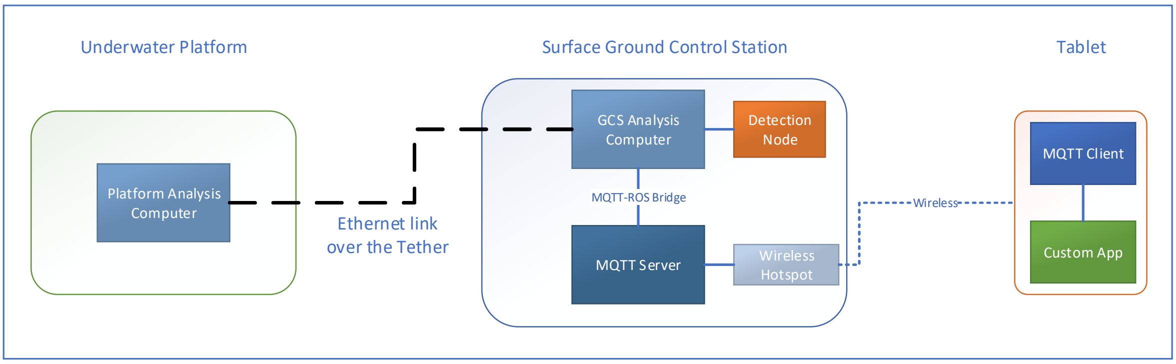


Figure 22. Schematic layout of the analysis system

The Detection Node implements the CSIRO COTS and Scars model. This analyses each image as it is collected and identifies the presence and location of any COTS (as a bounding box) and Scars (as an image mask), with the outcomes delivered as a message for other parts of the system to act upon.

Nodes are developed in Python or C++.

# APPENDIX B – WORKFLOW DIAGRAMS

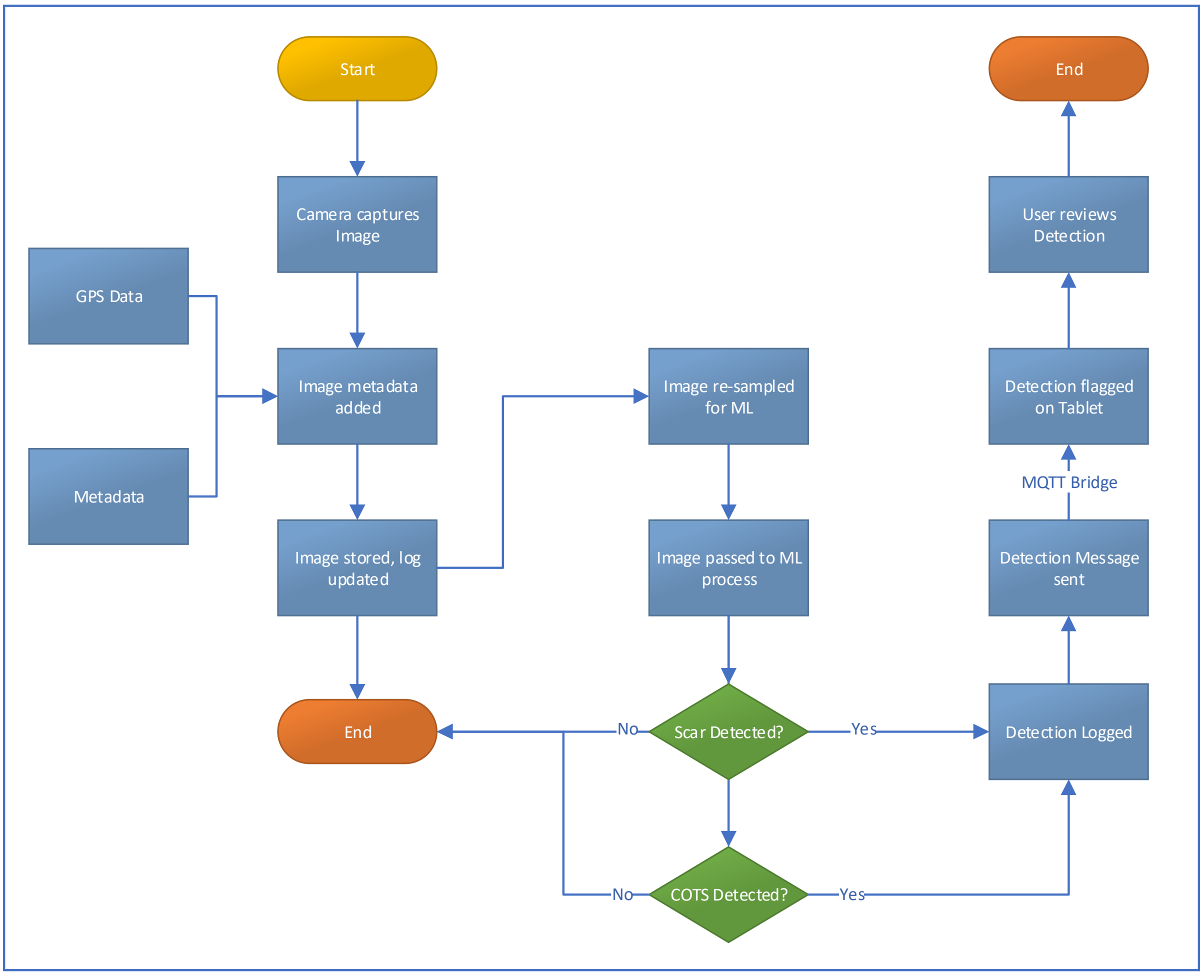


Figure 23. Real-time workflow showing how images are captured and processed.



Australian Government | Reef Trust
Great Barrier Reef Foundation
Australian Government | Australian Institute of Marine Science
CSIRO
James Cook University Australia
The University of Queensland Australia

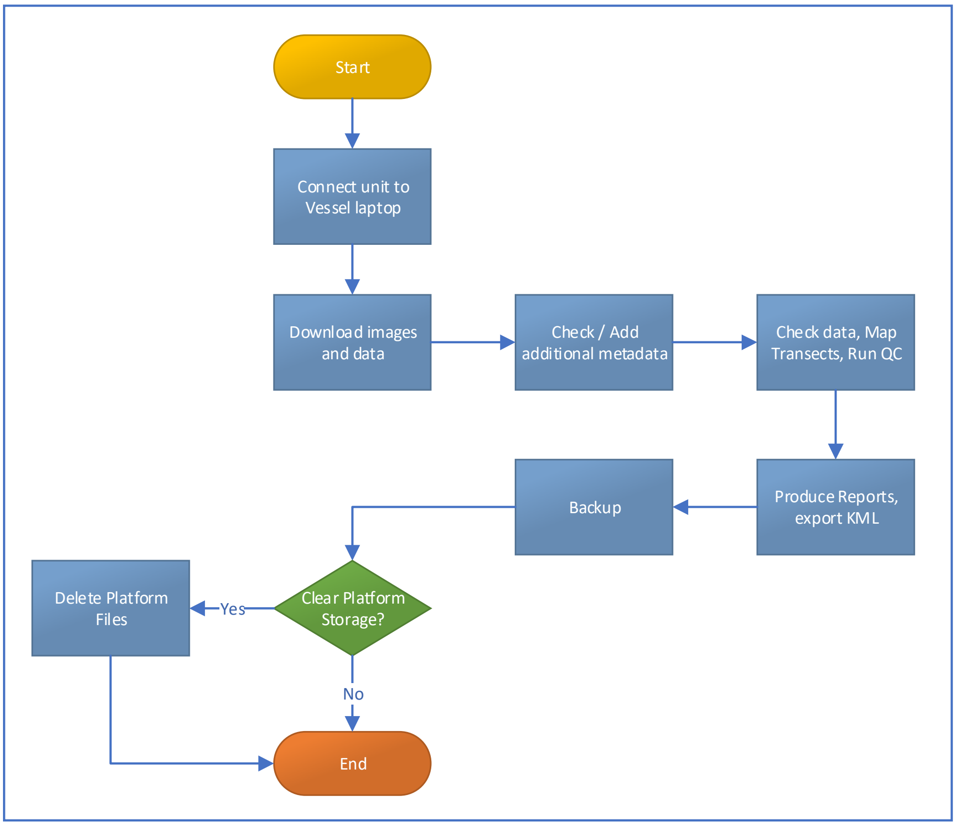


Figure 24. End-of-day workflow

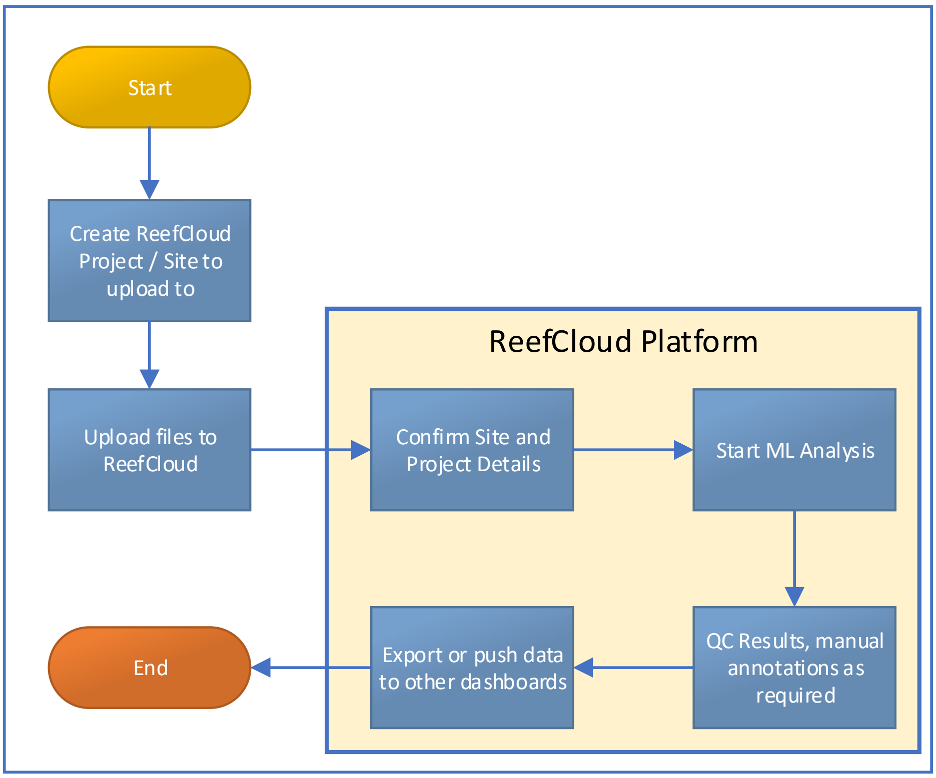


Figure 25. End-of-trip workflow showing use of the ReefCloud Platform

## APPENDIX C – COTS AND COTS SCARS ML MODEL RESULTS

Some examples of COTS detection are shown in Figure 26 below. Figures 27 and 28 represent some examples of COTS Scar segmentation by the DeNet model. For visualisation purposes, we assign every region with a scar score above 0.05 as red. Images in the left column are the model prediction, and the right is the ground truth as determined by the annotators (domain experts).

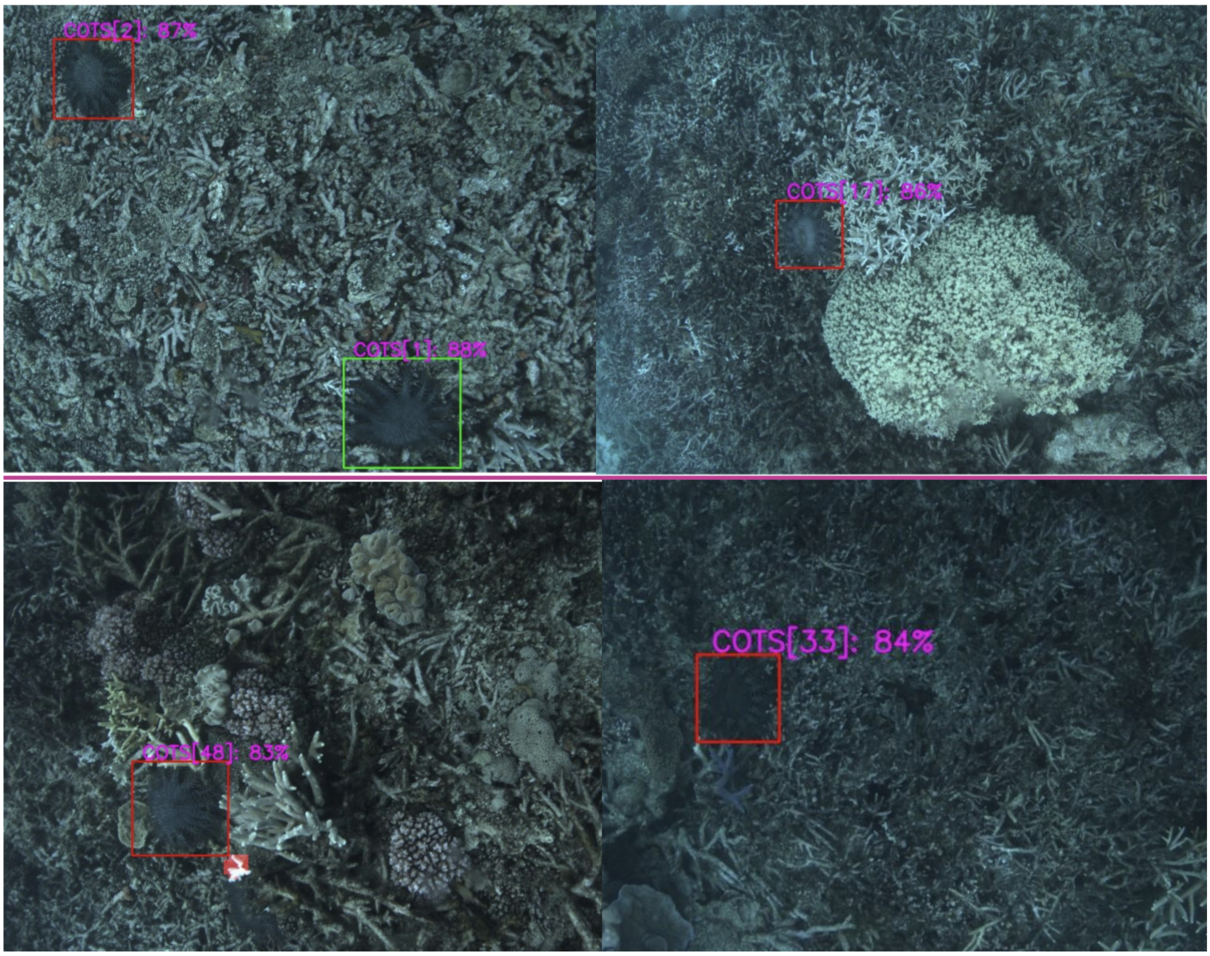


Figure 26. Examples of COTS detected by the DeNet Model

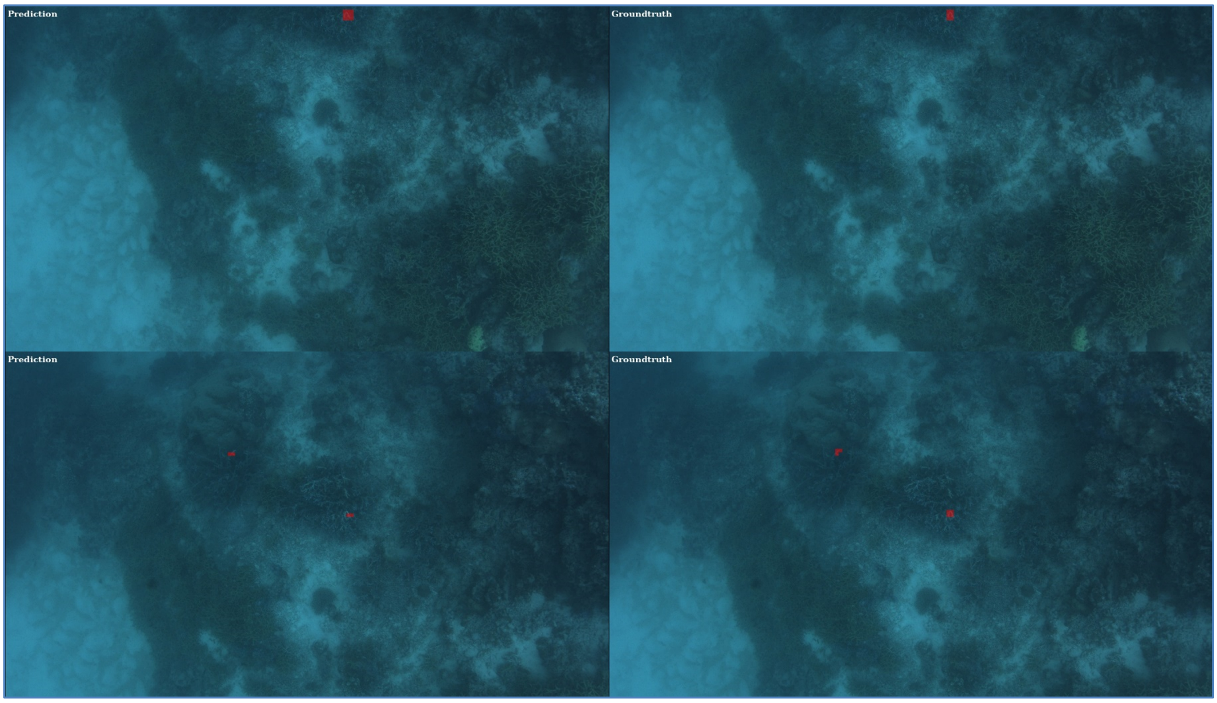


Figure 27. Examples of good identification of Scars by the ML model

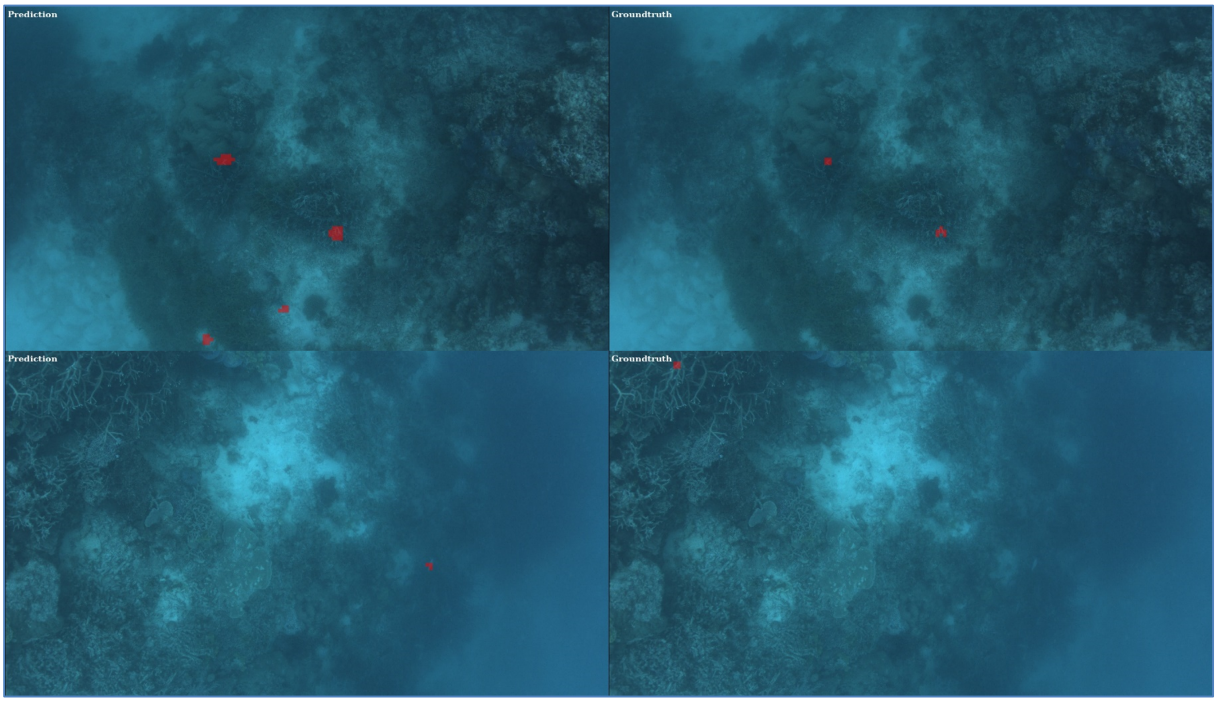


Figure 28. Examples of false-positive Scars identified by the ML model

Examples of when DeNet detects both COTS and COTS scars in a sequence of image frames are illustrated in Figure 29.

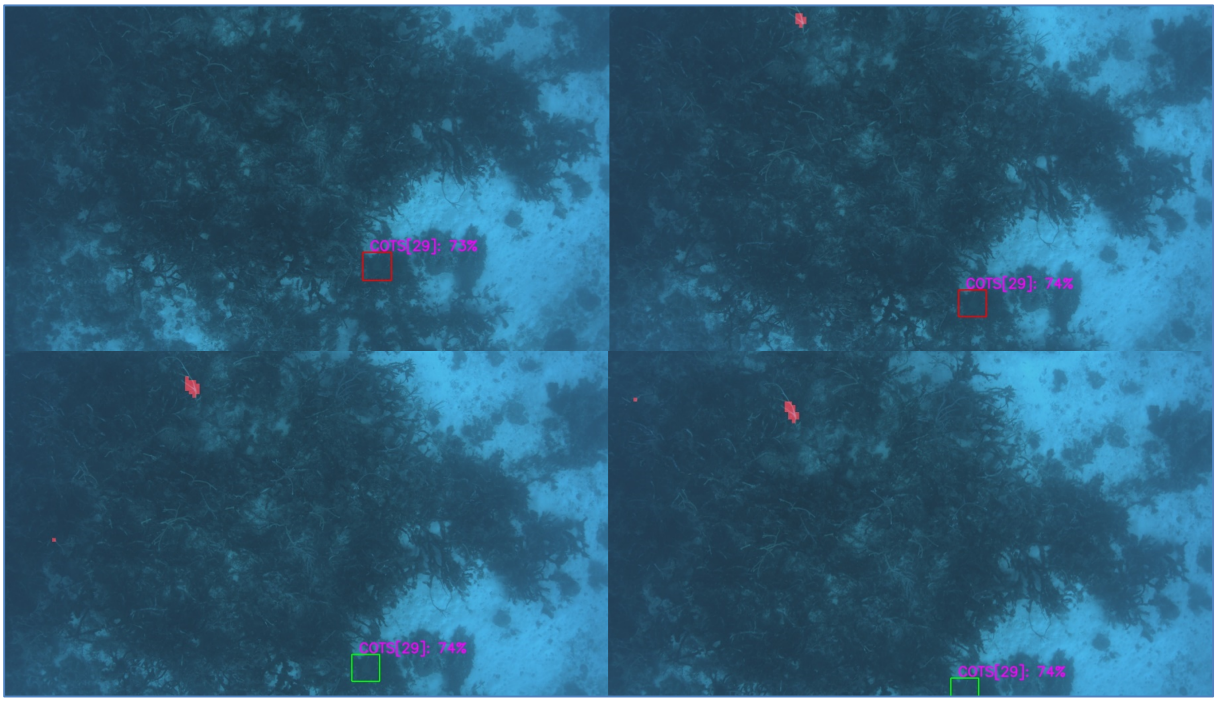


Figure 29. Examples of co-identified COTS (green) and Scars (red) over a series of images

# APPENDIX D – BENTHIC ML CLASSES

Table 10. Benthic Classes used for the Benthic Machine Learning model

| Benthic Class | Functional Group* |
|---|---|
| Algae Other | A |
| Macroalgae | A |
| Turf Algae | A |
| Dead coral (recent) | AB |
| Reefal substrate | AB |
| Rubble | AB |
| Sand | AB |
| Bottlebrush Acropora | HC |
| Bottlebrush Acropora - Bleached | HC-B |
| Branching Acropora | HC |
| Branching Acropora - Bleached | HC-B |
| Branching Acropora - Fluorescing | HC-F |
| Branching non-Acropora | HC |
| Branching non-Acropora - Bleached | HC-B |
| Branching non-Acropora - Fluorescing | HC-F |
| Corymbose Acropora | HC |
| Corymbose Acropora - Bleached | HC-B |
| Corymbose Acropora - Fluorescing | HC-F |
| Digital Acropora - Fluorescing | HC-F |
| Digitate Acropora | HC |
| Digitate Acropora - Bleached | HC-B |
| Encrusting Acropora | HC |
| Encrusting Acropora - Bleached | HC-B |
| Encrusting non-Acropora | HC |
| Encrusting non-Acropora - Bleached | HC-B |
| Encrusting non-Acropora - Fluorescing | HC-F |
| Foliose non-Acropora | HC |
| Massive non-Acropora | HC |

| | |
|---|---|
| Massive non-Acropora - Bleached | HC-B |
| Massive non-Acropora - Fluorescing | HC-F |
| Mushroom coral | HC |
| Solitary Coral | HC |
| Sub-massive non-Acropora | HC |
| Sub-massive non-Acropora - Bleached | HC-B |
| Tabulate Acropora | HC |
| Tabulate Acropora - Bleached | HC-B |
| Tabulate Acropora - Fluorescing | HC-F |
| Tabulate non-Acropora | HC |
| Tabulate non-Acropora - Bleached | HC-B |
| Unknown | IN |
| Water | IN |
| Heliopora | OT |
| Millepora | OT |
| Other Organisms | OT |
| Zoanthids | OT |
| Encrusting Soft Coral | SC |
| Erect Soft Coral | SC |
| Encrusting Soft Coral - Bleached | SC-B |
| Erect Soft Coral - Bleached | SC-B |
| Sponge | SP |

*The functional groups are detailed in Table 2.

# APPENDIX E – BENTHIC ML MODEL RESULTS

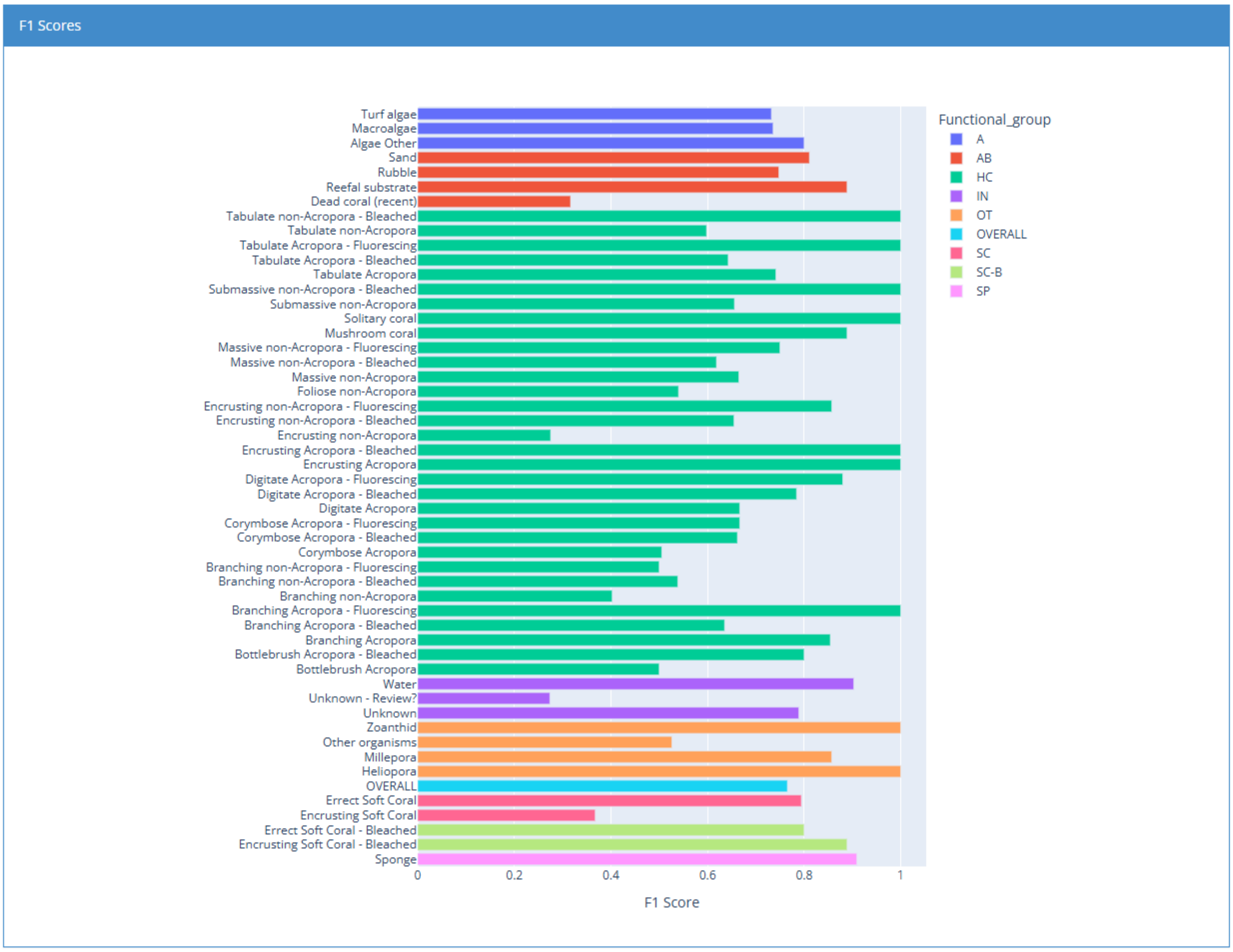


Figure 30. F1 Scores for the complete set of Benthic ML Classes

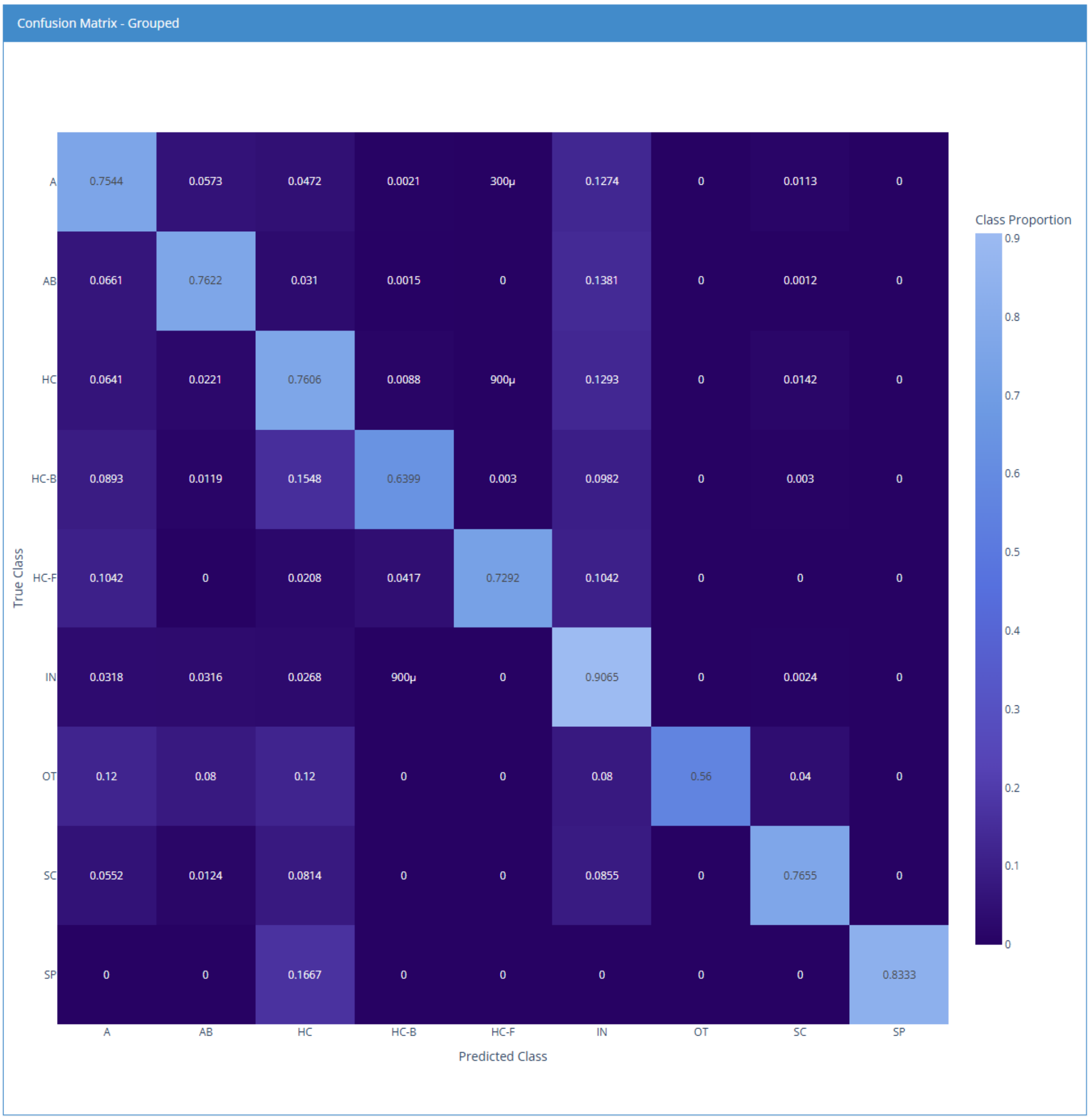


Figure 31. Confusion matrix for the aggregated set of Benthic ML Classes

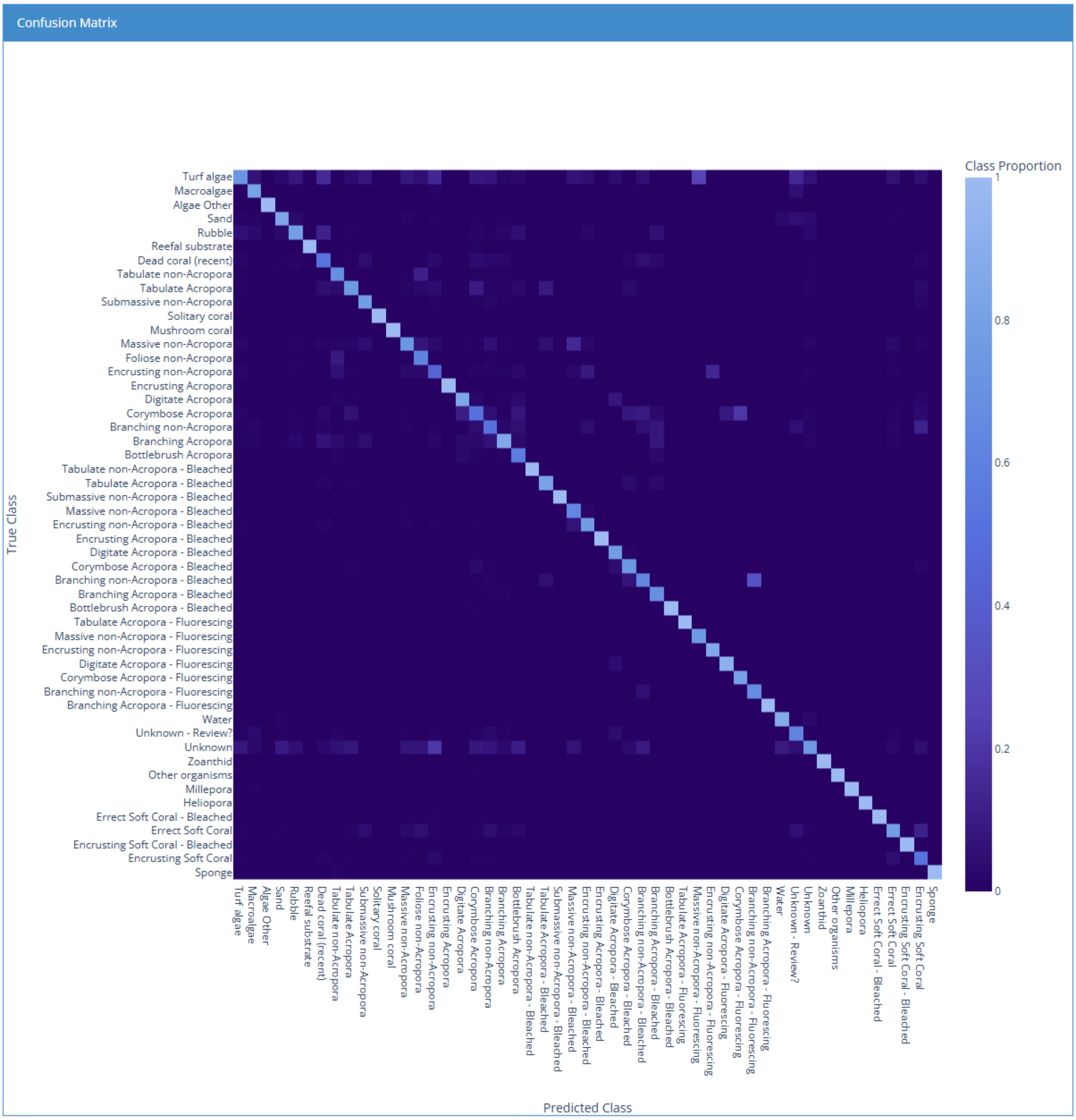


Figure 32. Confusion matrix for the complete set of Benthic ML Classes


Figure 33. Error Estimates for each of the Benthic ML Classes

# APPENDIX F – DESIGN REVIEW MEETING ACTION ITEMS, 15-OCT-2024

| Task Number | Description | Action |
|---|---|---|
| 1. | Update the TRA to include the risk that an entangled platform may impact the tender so the idea of a manual quick release or cutting device may be one response | Dev |
| 2. | Clear labelling the lift points, that they are rated as well as maybe warnings around the thrusters so that it is clear where you pick up the device, where to lift it, etc. | Dev |
| 3. | Label on the top hydrofoil about storing it in the shade when not in use or another warning about over heating | Nathan |
| 4. | Looking at deployment and retrieval options, concept for November trip | Mech |
| 5. | Fabric sun-shade for the top box, main concern is glare on screens | Mech |
| 6. | Update design to include rear handle for lifting in and out of the water | Dev |
| 7. | Markings on the camera mount for the angle of the camera so that they can be reset to the same point, maybe 0 \| 5 \| 15 \| 25 \| 35 \| 45 degrees | Dev |
| 8. | Identify field of view (meters) at each angle | Testing |
| 9. | Float on the tether so that if it is thrown over the side it will float | Nathan |
| 10. | Procedure for hot swapping batteries | Geoff |
| 11. | Quick release for tether (removing for ground control is acceptable) | Nathan |
| 12. | Looking at deeper reef visits (maybe greater than 40m). Scott to advise. | Scott |
| 13. | Switching cameras on the fly, but only recording with a single camera to reduce data upload | Geoff |
| 14. | Scoping number of batteries needed for each system | Geoff |

| | | |
|---|---|---|
| 15. | Charging strategy - bank charging of batteries or one charger per battery? | Geoff |
| 16. | Confirm charging times for the batteries | Nathan |
| 17. | Introduce ergonomic grips for handling | Dev |
| 18. | User Manual – trouble shooting | Nathan |
| 19. | Software camera identification of images (port/starboard) | Greg |
| 20. | Software reporting of locations of separate application | Greg |
| 21. | Software memory storage alert | Geoff |
| 22. | Software battery low voltage alert | Geoff |
| 23. | Software split ReefScan camera views, changing port starboard, both at once | Greg |
| 24. | Software larger display of vehicle parameters (Voltage, depth, etc.) | Geoff |
| 25. | Cable management – measure the feed out length of cable | Mech |
| 26. | SOP developed including how-to videos | Scott |
| 27 | Maintenance manual and videos developed | Scott / Dev |
| 28. | Download/Upload software allows the user to select which image set are needed for download. E.g. if the user was going around a wall and only images from camera 1 are needed, only download from camera 1 and remove camera 2 images. | Greg |

# APPENDIX G – REEF 2050 PLAN INDICATORS

Extracted from: https://www.dcceew.gov.au/sites/default/files/documents/reef-2050-objectives-goals-2021-2025.pdf

Table 11. Potential contribution of the work towards the Reef 2050 Plan indicators.

| Management Objective | Indicators | Delivers Against |
|---|---|---|
| Coral reef habitats maintain good condition and resilience | Percentage hard coral cover | Yes |
| | Coral disease per unit of coral cover | With Development |
| | Benthic Algae - Proportion of macroalgal cover | With Development |
| | Benthic Algae - Algal turf height | No / unlikely |
| | Benthic Algae - Percentage of crustose coralline algae (CCA) cover | Potentially |
| | Microbial community composition | No |
| | Herbivore biomass | Potentially |
| | Mitigated crown-of-thorns starfish damage through maintenance of populations at below outbreak densities | Yes |
| | Capacity of individual reefs to recover post disturbance as measured through state and trend of: – Hard coral community composition • Density of juvenile corals – Post disturbance coral size class distribution | Yes |
| | Capacity for sustained functioning of the Reef ecosystem, as indicated by – measuring the current carbonate budget of individual reefs and – predicting the future carbonate budget | No |
| | An integrative analysis framework is being developed to support reporting against this objective, with a focus on the condition and recovery capacity of coral reef habitat | Yes |

| | | |
|---|---|---|
| **Resilient seagrass meadows that maintain condition** | **Seagrass spatial distribution** | **With development** |
| | Seagrass biomass / cover | With development |
| | Community composition / types | With development |
| | Trend in asexual / sexual reproductive capability | Potentially |
| **Governance systems are inclusive, coherent and adaptive:** | Support for decision-making improves, including: integrated monitoring and reporting, data management and decision support. | Yes |
| | Capacity for adaptive and anticipatory management increases | Yes |
| | Planning, management and decision making is more inclusive of rights and interests of stakeholders, Traditional owners and communities | Potentially |
| | Policy and program coherence between tiers of government and portfolio areas is improved | Yes |
| | Satisfaction with governance and management increases | Uncertain |
| | Co-management with Traditional Owners increases. | Yes |