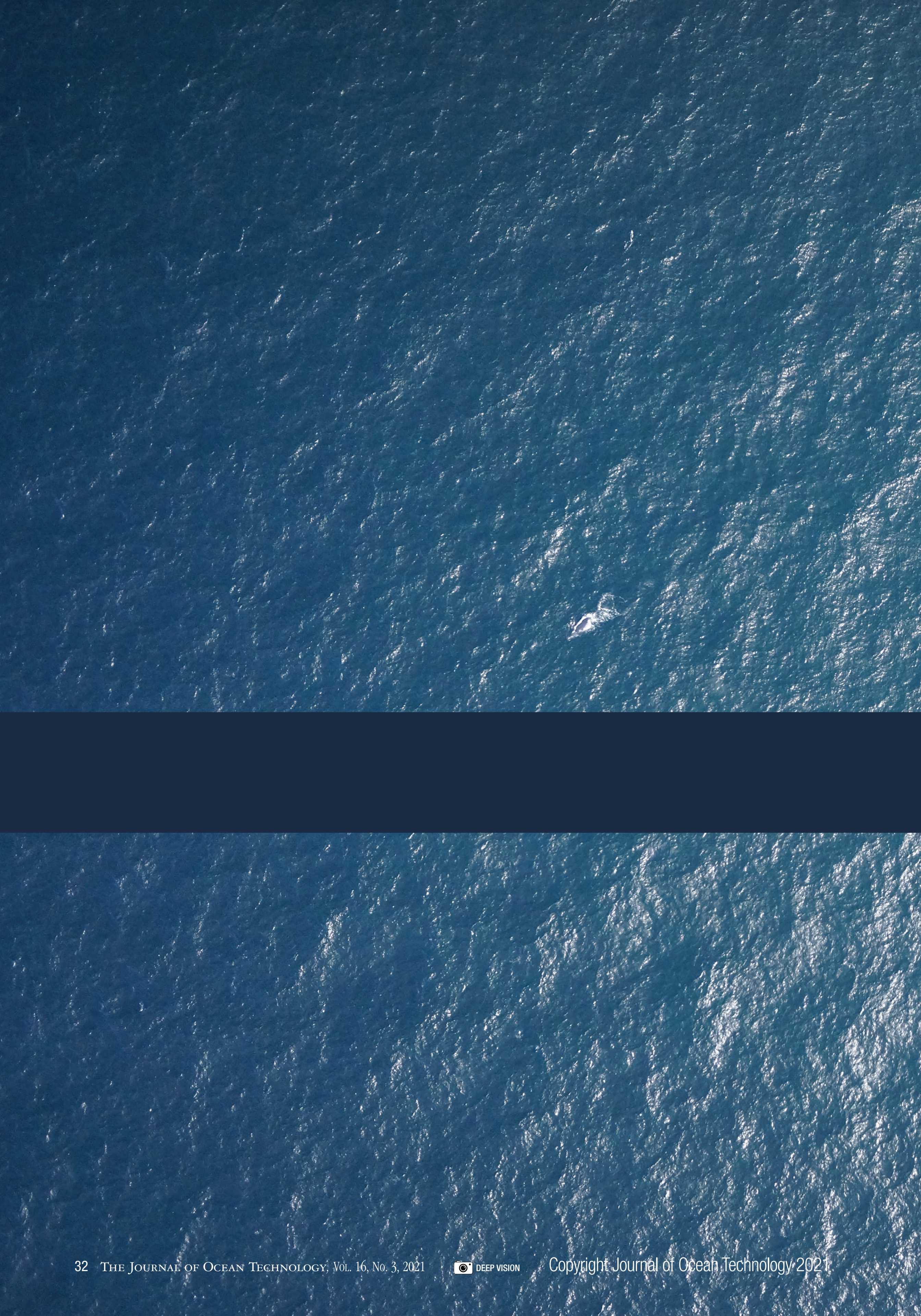



Copyright Journal of Ocean Technology 2021

# Whale Watching

## Automating Aerial and Surface Level Cetacean Monitoring for Improved Population Surveys

by Michael Outhouse, Alan Parslow, and Alvin Beach

Copyright Journal of Ocean Technology 2021

*Conservation efforts for at-risk marine species is a multi-disciplinary problem, and one spanning sub-surface, surface, and aerial spaces. This essay discusses Deep Vision's contribution to improved population surveys of North Atlantic right whales through the development of artificial intelligence (AI) to automatically detect, track, and geotag this endangered species using commercial off-the-shelf (COTS) electro-optical sensors. The scalability of the technology, including its resilience under all weather conditions and its application both as a surface level, mast mounted monitoring solution for ships, and as an aerial solution for uninhabited aerial vehicles and crewed surveillance aircraft is outlined.*

**Introduction**

The North Atlantic right whale (NARW) is an endangered species that is closely monitored in Canadian waters. The whales calve off the coast of Florida and then migrate north each year, arriving in Canadian waters in mid-summer and travelling through the Cabot Strait to rich feeding grounds in the Gulf of St. Lawrence.

The survival of the species is threatened by human marine activity (see Figure 1). To prevent species collapse, the Canadian government implemented vessel speed restrictions during the NARW season to mitigate collision risk and mandated temporary fishery closures to reduce the risk of entanglement in fishing gear.

Maximizing conservation efforts while minimizing fishery and shipping disruptions requires robust monitoring efforts. Monitoring initiatives track NARW behaviour and habitat. These initiatives employ crewed and uncrewed aircraft, reported sightings from surface vessels, and, increasingly, the use of artificial intelligence to supplement the efforts of operators.

In 2017, 17 North Atlantic right whale incidents, resulting in 12 deaths, were attributed to human marine activity in the

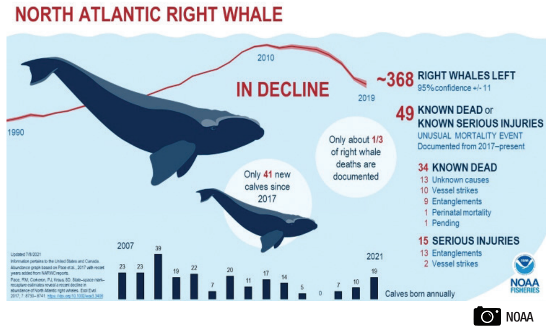



Figure 1: The year 2017 marked a critical year for the North Atlantic right whale population, signalling a population in decline.

Gulf of St. Lawrence. For a species with an estimated population at that time of 411, this represented a significant impact to a highly endangered species. Emergency measures were enacted in the later part of the 2017 NARW season to limit risk exposure. Leading into the 2018 season, a wide range of measures was adopted to limit human activity in areas where NARW were known to be present. Measures included limitations on the amount of rope in the water for some commercial fishing licences, static and dynamic fishery closures, and speed limitations. The result of these measures was clearly seen in the data, and no NARW deaths were recorded in 2018.

While the measures proved successful, the basis for closures depended largely on real-time monitoring of the whale population. For instance, the confirmed presence of NARW in a specific fishing zone can trigger the temporary closure of that zone to certain licence holders. It is documented that in 2017 and 2018, only 100 NARW were observed in the Gulf of St. Lawrence; with a population of slightly more than 400, this means that over 300 NARW were potentially exposed to risk. Increasing monitoring efforts provides an opportunity to observe more of the species and provide a better view of where activity and risk may be. This has a clear benefit for conservation efforts and mitigates the economic impact of fishery closures or limitations by enabling a region to open sooner once the risk is adequately reduced. The dependency is on observation.


Copyright Journal of Ocean Technology 2021

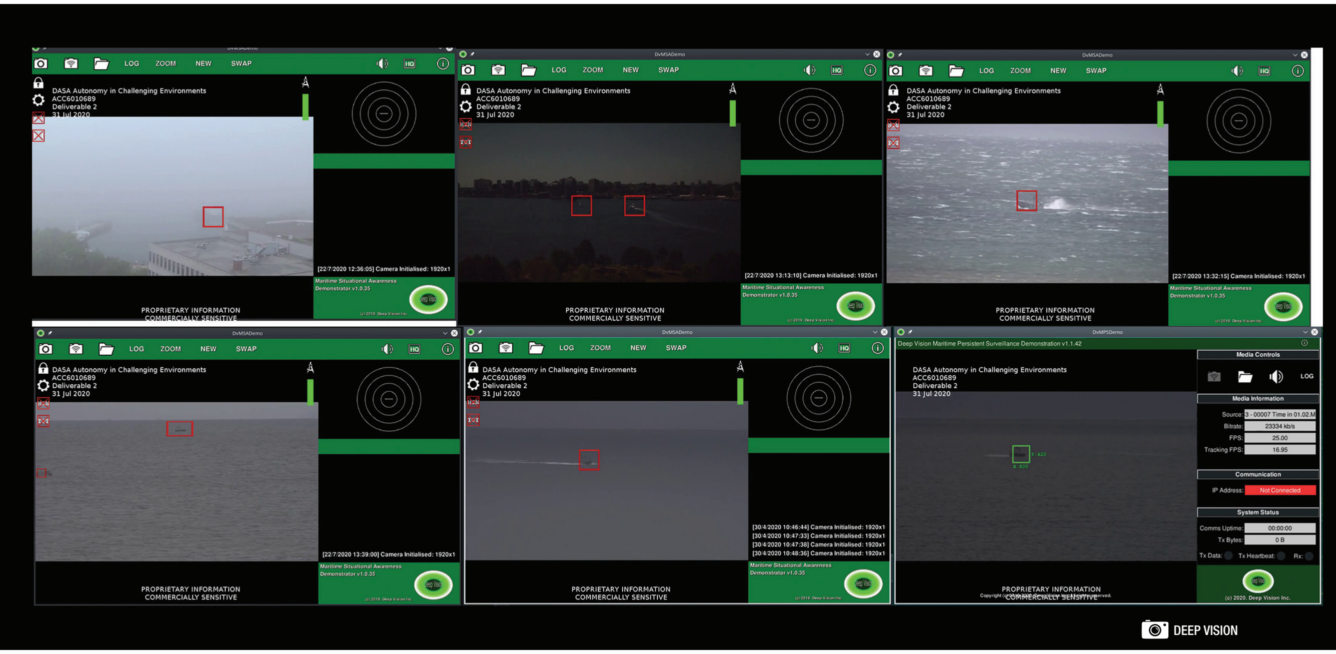



Figure 2: Deep Vision's maritime persistent surveillance technology finds objects near the surface of the water in all conditions. This capability was leveraged for the detection of cetaceans from aerial or surface platforms.

Deep Vision has historically applied its technology to the defence domain. With a long history of developing novel intelligent vision technology for military applications, Deep Vision saw an opportunity to leverage its expertise in the critical area of NARW conservation. As a Nova Scotian company, Deep Vision also has a vested interest in supporting the Atlantic Canadian industry. The same vision technology developed for autonomous maritime persistent surveillance in the naval domain can be used to find and track NARW, protecting the species while minimizing the duration of closed zones.

**Automating Maritime Surveillance**

The maritime domain is unique for machine vision. The environment, both atmosphere and ocean, is highly dynamic. Waves, reduced visibility due to fog or cloud cover, poor or variable lighting, and precipitation are common factors that impact the visibility of objects on or near the surface of the water. This can impact the resolution with which an object is imaged and it can impact the duration of the imaging. For instance, a buoy from a lobster trap may be visible only periodically and incompletely due to wave action; when the buoy appears, it may be for very short periods of time and it may present as highly occluded and deformed due to waves and reduced visibility. Identifying where objects on the surface of the water are must be as dynamic as the water itself, and must scale with the conditions as the conditions change.

Deep Vision has positioned its technology and underpinning research in precisely this domain: providing automatic surface object detection in all-weather maritime conditions. At its core, what Deep Vision has developed is a detector that takes passive camera data (visible or infrared) and reports to the operator or system the location of all objects on or near the surface of the water (Figure 2). The nature of those objects (whether they be highly anticipated structures such as navigation buoys or as unexpected as a drifting log) does not factor into the detection. Anything near the surface of the water that is not itself the water or a consequence of the environment is detected.

From a maritime surveillance perspective, this distinction that any and all objects are

Copyright Journal of Ocean Technology 2021

detected is important. For instance, when navigating an uncrewed surface vessel, understanding that there is an obstacle in the way of a planned route is critical, independent of what that obstacle is. To achieve this capability, Deep Vision has leveraged the principles of unsupervised machine learning and applied them to what is classically referred to as the pattern-of-life.

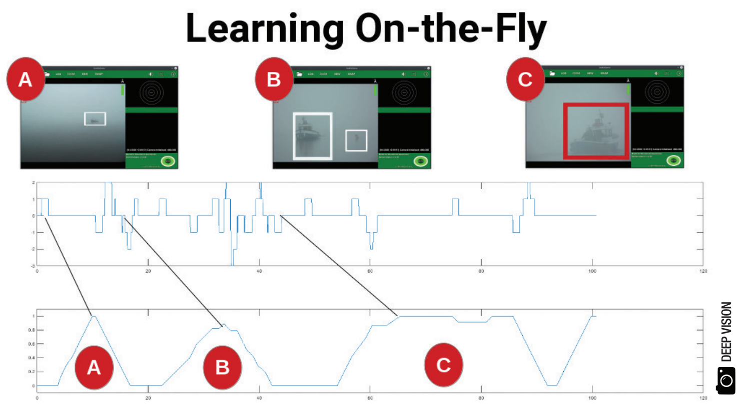



Figure 3: Information about objects and the environment is learned in real time while the system is in operation. The system adjusts itself continually to optimize operation in whatever conditions are present. Peaks in the graph correspond to periods of real-time learning.

**Pattern-of-Life**

Maritime object detection operates by understanding which visual features and behavioural characteristics are expected: the pattern-of-life (PoL). The expected characteristics vary spatially and temporally. What may be unique in one region may be very typical in another and, likewise, what may represent uniquely at one time may be expected at another.

The PoL is the set of real-time learned characteristics (visual and behavioural) that characterize the expected appearance and behaviour in a region over time. To assess the PoL, the input needs to reflect the visual appearance and the behaviour characteristics to the extent that trends can be ascertained. The framework used to achieve this is the unsupervised clustering of visual and behavioural properties via the k-means clustering algorithm. As observations are made during run time, trends emerge in the k-means clustering, which results in an effective partition of background features from those likely associated with an object.

**Detecting Objects**

Regions of interest are locally identified in the native coordinate space of the sensor. A similar learning phase is applied to the regions of interest to identify trends. Environmental features, such as crashing waves or sun speckle, emerge from this second phase of learning and are, thereby, filtered from the detection list. Resulting from this process is a set of detected surface-level maritime objects. A tracking process is inherent to the methodology, with each unique object having a position in sensor coordinates and a unique identifier that allows it to be differentiated from other objects that are imaged (Figure 3).

**Whale Watching**

With the news of increased human-related incidents involving North Atlantic right whales in 2017, Deep Vision objectively evaluated the potential to refine its generic maritime surveillance technology into a tool for automatic NARW detection. A rapid pace study was done using video data freely available in the public domain, such as drone clips from whale watching adventures. The purpose of the study was to evaluate a series of videos from aerial and surface perspectives using Deep Vision's technology and identify if the whales were being detected along with the other surface-level objects in the scene.

The short case study proved conclusively that there was merit in moving forward. Figure 4 shows a frame capture from an early experiment. In this video, a whale surfaced near the vessel on a clear weather day. Both the whale and vessel were detected, with the technology demonstrating promise for detection both when the whale is visible subsurface and when it breaches the water's surface.


Copyright Journal of Ocean Technology 2021

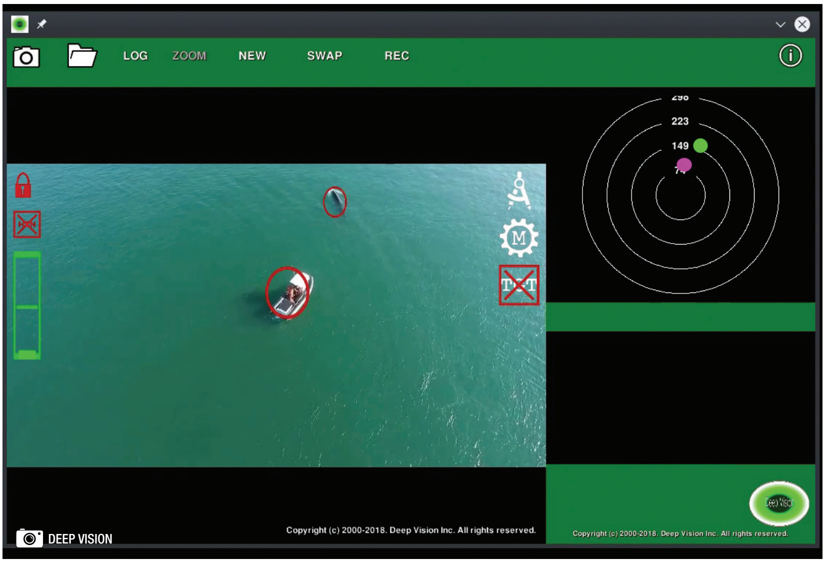



Figure 4: A rapid study was done using Deep Vision's maritime persistent surveillance technology to evaluate the likelihood of cetacean detection. This study was focused on determining if cetaceans were in the full set of surface-level detected objects.

Work began in late 2018 on an ambitious plan to develop an automated cetacean detector that works from surface and aerial platforms, and from fixed and moving installations. The focus was on exploiting COTS visible light cameras, as they are readily available on most existing vessels and aircraft (crewed and uncrewed). By developing an unconstrained system, the potential to interface with a wide variety of platforms increases the survey potential and area covered, and thus has a clear conservation advantage.

Through the later part of 2018 and the first quarter of 2019, an accelerated program of work formed the foundation for Deep Vision's cetacean detection capability. While the basic principles extended the core maritime surveillance technology, considerable new and innovative work was developed to refine the technology into a cetacean detector. In addition to furthering the scientific and mathematical principles underlying the technology, a range of issues was solved that furthered the objective of providing an unconstrained system with easy integration options.

**Cetacean Classification**

The classification model employed for monitoring of the cetaceans was a hybrid of visual and contextual classification. The core detection as described isolates an object as something that appears or behaves differently to the water, and is repeatedly visible with similar characteristics over a short period of time. This provides the set of all surface-level maritime objects, including the subset of cetaceans. Specific information unique to cetaceans was considered in filtering on this set: in particular, physical limitations and behavioural characteristics.

Physical dimension was used to limit the set of detected objects to those which may possibly have been a cetacean. No lower limit was applied, while the upper limit was capped at the maximum length of cetacean that would be considered. Through the exploitation of platform and camera data, real world measurements were generated for candidate cetaceans in real time.

Behavioural characteristics required extensive literature review. Unlike a boat that is continuously visible (notwithstanding environmental considerations), a cetacean is only visible in limited intervals. The duration of visibility, the dynamics of how the visual presentation changes, and the rate at which the visible signature dissipates all impact the behavioural metric and contribute to positive detection of a cetacean.

Copyright Journal of Ocean Technology 2021

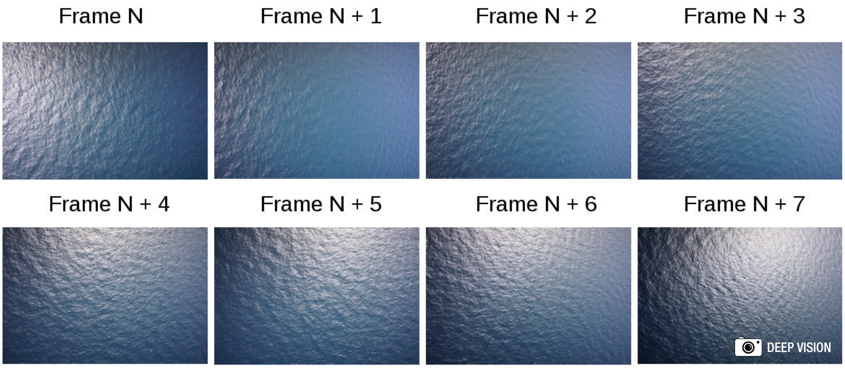



Figure 5: Sample of consecutive frames from a typical aerial survey. The images are 4 K (6000x4000 pixels) and captured at a rate of 0.25 Hz.

It was assumed that cetaceans are largely visible through their effect on the environment rather than on visibility of their bodies directly. Building a classification schema requiring visibility of the body was deemed limiting. The optical thickness of the water confounds imaging of subsurface whales due to a range of factors including camera, altitude, viewing angle, weather conditions, etc. Whales breaching the surface to breathe or to feed do so with a driving force that disturbs the water around them, resulting in the visual signature being predominately the effect of the whale on the water rather than the whale itself. Deep Vision's expertise in unsupervised machine learning made it possible to differentiate between disturbed water from wave action versus disturbed water from a cetacean.

**Staring at Water**

The trial data used in support of this study was provided in kind by Transport Canada. The data consisted of unmanned aerial vehicle-captured imagery spread over seven survey days in August 2018 (Figure 5). Each set was comprised of high resolution (4 K, 6000x4000) images and associated platform metadata. All images were captured at a rate of approximately once every four seconds (0.25 frames per second).

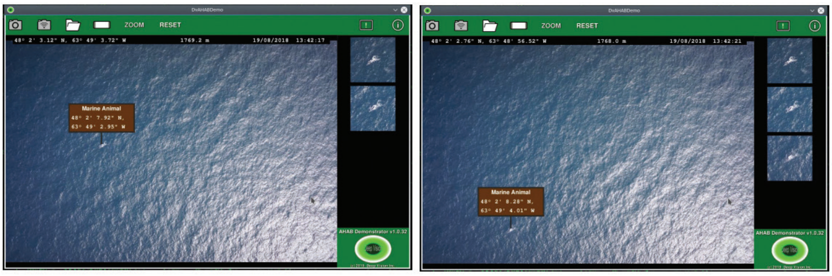
The survey data spanned a wide range of operational conditions, including contrasting lighting (bright day, dark day); highly focused sunlight producing wide area sun glare; calm seas; and heavier, white cap marked seas.

A typical survey consisted of a large number of images with very few examples of cetaceans or any feature aside from the water. For example, a survey in the Gaspé region on August 19, 2018, consisted of 2,464 4 K images, representing approximately a three-hour survey. The survey data was processed by the cetacean detector with the following results (see Figures 6 and 7):

- A total of two distinct cetaceans were detected, providing a total of six geotagged snapshots/crops of the observations.
- No false detections (non-cetacean classifications/detections) were produced.
- A total of three surface anomalies were missed. These are comprised of surface features that visually appear distinct from waves or sun glare but did not conform to the expected physical characteristics of disturbed water induced by a cetacean.

Processing power required for cetacean surveys of this nature was evaluated. On a standard workstation (Intel i7-2600 K CPU @ 3.40 GHz), processing the entire three-hour survey took two minutes twenty seconds. This equates to a processing rate of 17-18 Hz on 4 K imagery. On a low powered ARM-based device (ODROID-XU4), processing the entire three-hour survey took eight minutes sixteen seconds. This equates to a processing rate of 5 Hz on 4 K imagery. The latter is equivalent to ~60 Hz at 1080p resolution, more than enough to enable most Edge AI applications.


Copyright Journal of Ocean Technology 2021



DEEP VISION

Figure 6: Detection of a cetacean in consecutive frames from aerial survey data.

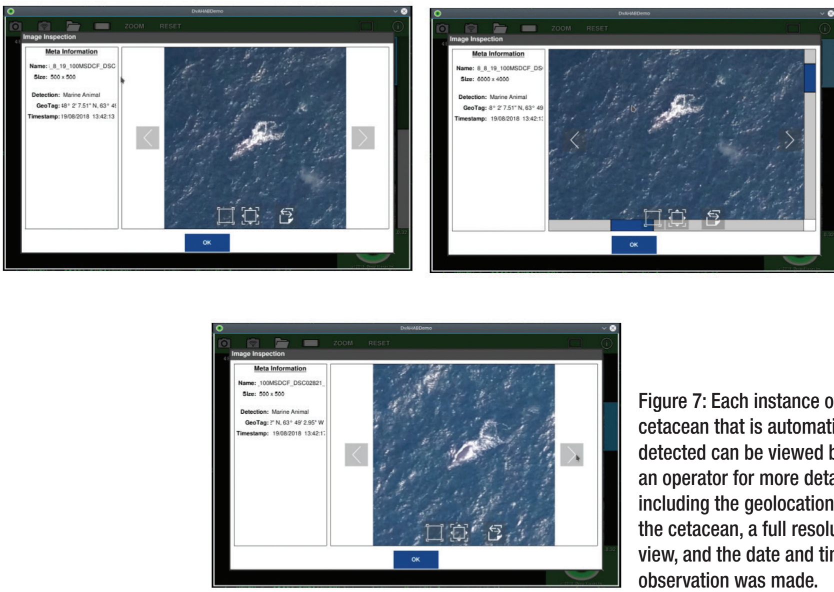


Figure 7: Each instance of a cetacean that is automatically detected can be viewed by an operator for more detail, including the geolocation of the cetacean, a full resolution view, and the date and time the observation was made.

DEEP VISION

The maritime domain presents its challenges with weather heavily influencing the visibility of cetaceans. With increased wind comes increased wave action and, as a result, increasingly disturbed water due to waves. When a cetacean breaches the surface, the combination of wave dynamics and atmospheric dynamics can distort, minimize, or amplify the visible signature of the cetacean. Trials were conducted to evaluate the impact of stormy conditions on the automatic detection of cetaceans.

A survey from August 11, 2018, in the Gaspé region was used as representative of more challenging conditions. The survey consisted of 2,097 4 K images sampled at approximately 0.25 Hz. Figure 8 shows an example of the conditions surveyed on the day. Due to the challenging nature of the survey, a careful manual inspection of the imagery was done to identify what could likely be a cetacean. The manual survey was unable to identify any instances of cetaceans.

The cetacean detector was used to process the survey data and resulted in 17 detections (Figure 9). Each detection was reviewed. While 16 detections were concluded to be induced by wave action and not of interest, one detection clearly showed a cetacean submerged near disturbed water from a wave. To place into perspective the achievement in this result, it took a manual observer approximately three hours to review the survey and conclude that no cetaceans were present. The cetacean detector processed the same set in approximately two minutes thirty seconds and produced

Copyright Journal of Ocean Technology 2021

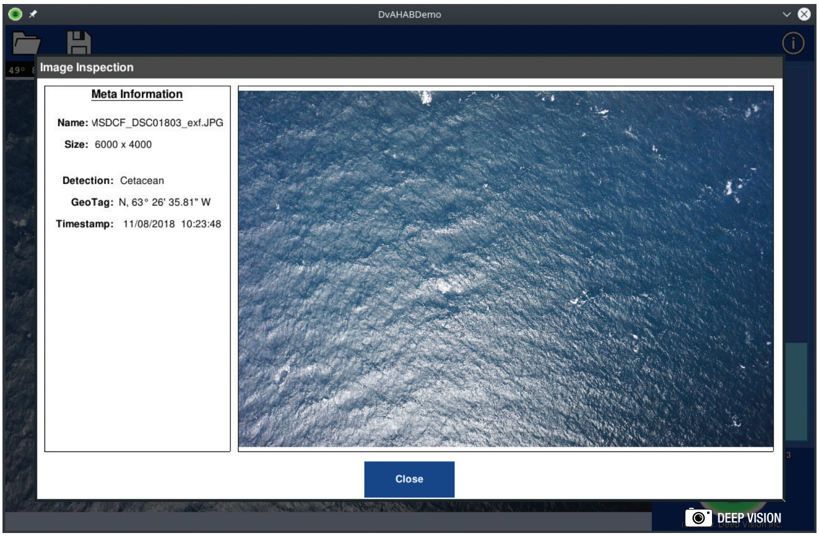



Figure 8: An example of an aerial survey on a day with heavier seas and sun glare.

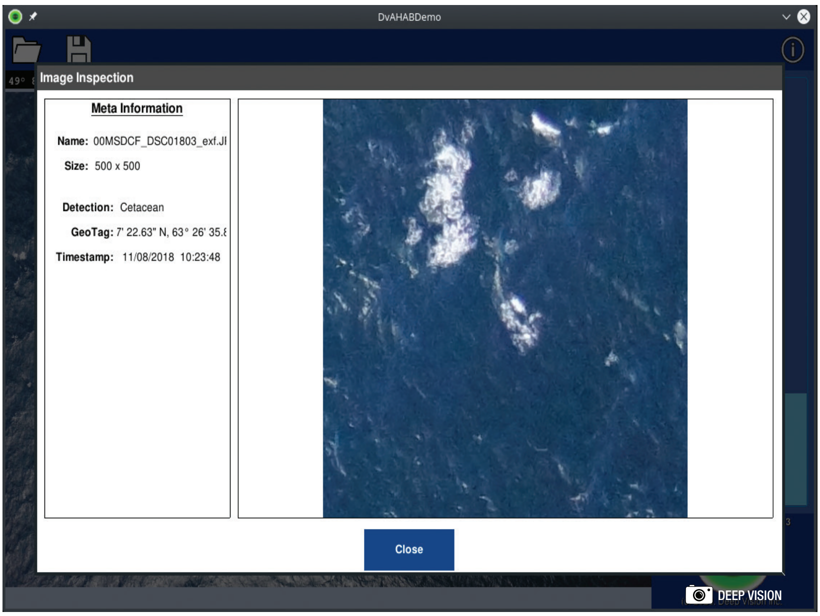



Figure 9: A cetacean detected amid heavy seas. This instance was not immediately visible on manual inspection, but detected by Deep Vision's technology.

17 possible cetaceans. It then took under one minute for a person to manually review the detections, discount 16, and identify the presence of a cetacean. The dataset to review was reduced by 99.18% and a cetacean was found that would otherwise have been missed.

## Moving Forward

Conservation of North Atlantic right whale and other at-risk species is a multi-disciplinary problem with a wide spectrum of stakeholders. Key to this effort is monitoring and observation. Policies that have proven successful in averting human-caused mortality of the NARW rely on understanding where the whales are. Modification of human activities that jeopardize NARWs rely on knowing when NARWs enter and depart particular zones. This enables balance between conservation and the economic interests of the region.


Copyright Journal of Ocean Technology 2021

Monitoring and observation fall in the hands of many parties, and are being improved through contributions from government, academia, and industry. Just as the problem has a diversity of contributors, so too are the solutions encompassing. Advancements in underwater listening technology have enabled solutions for subsea detection, with the technology fast becoming more reliable and cost effective. Crewed surface and aerial systems have proven extremely effective, yet are neither cost nor resource efficient. Crewed aerial patrols must be strategically delivered to make the most of the resources available, missing opportunities if NARW habitat varies significantly from expectation. Trained observers, known as marine mammal observers, must review the data for positive identification, an extremely time-consuming task.

Efforts like those taken by Deep Vision to automate surface and aerial detection make the most of the resources available. They allow marine mammal observers to review a select set of potential cetacean detections and confirm or disregard. More focused use of human resources will enable monitoring of larger geographic areas, faster turnaround in zone closure decision-making (with less unnecessary disruption of economic activity), and more efficient and effective preservation efforts.

Deep Vision is positioning its work in cetacean detection to play a vital role in conservation and mitigating economic losses for the region. A range of key strategic partnerships has been established between Deep Vision and industry with the goal of bringing this technology to market. This collaboration includes a highly respected marine environmental technology company and a not-for-profit marine science research institute.

Conservation of at-risk species is a problem that requires cooperation, collaboration, and innovation. The problem is bigger than one company, bigger than one government agency. As skills and innovations are tested, matured, and proven, the risk to species will continue to be reduced and the activities we depend on for our economic livelihood can co-exist with the many species that share the same waters.

Michael Outhouse, principal scientist, is a research scientist with 17 years of experience in the fields of machine vision, machine learning, data science/analytics, and autonomous systems. With a strong academic background in oceanography and marine science, he has focused his research on autonomous maritime systems, with an emphasis on leveraging and further developing unsupervised machine learning techniques to support robust operation of autonomous systems in unpredictable maritime conditions. Mr. Outhouse has been with Deep Vision since 2004, first serving as a scientific consultant, and later moving into the position of principal scientist.

Alan Parslow, CEO, is the driving force behind Deep Vision since its inception in 2000. His work has focused on the mathematical and philosophical concepts of machine vision and autonomous systems. He developed the foundational machine vision and data abstraction methodology on which Deep Vision's technology and development programs depend. Mr. Parslow has focused his research interests on autonomous systems for challenging maritime tasks, such as search and rescue, marine mammal spotting, and beyond-line-of-sight applications.

Alvin Beach, CTO, is a highly experienced software engineer in the areas of research and product development of innovative, disruptive technologies and solutions. He has a long history of developing cutting edge software and hardware solutions for next generation autonomous technologies. His focus has been on low-SWAP and limited bandwidth systems. Mr. Beach began his career at Deep Vision as a software engineer in 2003, and has since become the CTO. In addition to his role as principal engineer, Mr. Beach is an accomplished project manager and quality manager.

Copyright Journal of Ocean Technology 2021