ARTICLE | Article in Press

Accelerated development and deployment of computer vision models for invasive aquatic pests

Rose A. Pearson · Gareth Preston · Jeremy Bulleid · Svenja David · Felix Zareie-Vaux · Daniel Clements · Leigh W. Tait

Received: 11 September 2025 / Accepted: 15 September 2026
© The Author(s) 2026

Abstract
Aquatic non-indigenous species (NIS) incur significant cultural, environmental and economic costs worldwide, and human surveillance is expensive and impractical at large scales. Computer vision models (CVMs) deployed on remote and autonomous vehicles can reduce the burden on trained human observers, but aquatic environments present unique challenges and lack frameworks for biosecurity-focused development and deployment. We present a practical framework for developing and deploying new species-specific CVMs for real-time use across surface vessel and remote vehicle platforms. We demonstrate the utility of our framework through its application to several taxonomically and ecologically diverse NIS in New Zealand: Mediterranean fanworm *Sabella spallanzanii*, South African oxygen weed *Lagarosiphon major,* and exotic *Caulerpa* (*Caulerpa brachypus* and *C. parvifolia*). For *S. spallanzanii*, we demonstrate efficient training from small datasets and deployment for real-time detection. Using exotic *Caulerpa*, we show how CVMs can be rapidly improved during an early incursion response. For *L. major*, standardised field validation methods enable the comparison of CVM and human detection rates across diverse locations and operating conditions. Collectively, these case studies demonstrate that our framework enables accurate detection and the robust assessment of model effectiveness under realistic field conditions and can be effectively applied to imagery from multiple platforms.

Keywords Artificial intelligence · Biosecurity · Invasion ecology · Machine learning · Remote operated vehicle · Surveillance

Introduction

Non-indigenous species (NIS) incur significant cultural, environmental and economic costs to marine and freshwater environments worldwide[1,2]. The rate of marine and freshwater NIS incursions continues to rise with the globalisation of people and goods[3], with a projected rise of 36% by 2050 across marine regions and habitat types[4]. Surveillance of aquatic pests has traditionally been a challenging and expensive operation. Improving the coverage and efficiency of aquatic pest surveillance has the potential to translate to effective management and intervention, with tangible economic and environmental benefits to aquatic ecosystems and related industries.

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.



https://doi.org/10.1038/s41598-026-72212-8

Published online: 23 September 2026

scientific reports

The early detection and delimitation of NIS incursions is critical as control and eradication is often only possible when an NIS population is small and spatially restricted[5–8]. Human divers with a high level of taxonomic expertise are typically considered to be the 'gold standard' for detecting NIS incursions, due to their fast pace and ability to detect a wide range of target species[9]. However, human surveillance is expensive and labour intensive, requiring extensive technical and taxonomic training, and divers can only investigate a small proportion of the aquatic environment[3], which makes it unfeasible for the current and growing scale of aquatic NIS incursions.

Camera systems such as remotely operated vehicles (ROVs), autonomous underwater vehicles (AUVs), unmanned surface vehicles (USV), and towed or fixed cameras are increasingly used to capture images or video footage for aquatic biosecurity monitoring[3]. Notably, computer vision models (CVMs) deployed on commercially-available ROVs and USVs integrated with camera systems have the potential to empower amateur pilots and citizen scientists interested in NIS surveillance. Tait, et al.[9] investigated the application of ROVs for marine NIS surveillance using both human and CVM detection methods. They found ROVs with either human review in post or CVM based detection to be equally accurate but less time efficient (excluding CVM model development time) than human divers due to ROV piloting time and image assessment time in the case of human review[9]. The use of CVMs eliminated the imagery assessment time, and greatly reduced the costs of remote methods[9]. While AUVs can circumvent human piloting[10,11], they require reliable CVMs to manage the scale of image data gathered[12]. The development and deployment of CVMs remains a challenge for improving the efficiency of aquatic biosecurity surveillance[1,13].

Bounding box-based object detection is commonly used to identify aquatic species[9,14,15]. Although there are a myriad of different architectures dedicated to performing object detection, these are typically extensions of a convolutional neural network (CNN) architecture[16]. Popular CNN architectures are commonly trained for CV across large, labelled image datasets (e.g. ImageNet, Caltech 101) and then used as a backbone model to develop new two-stage models via 'transfer learning'. This approach results in improved performance for small image datasets[16–18]. Data augmentation, where training datasets are enlarged and diversified by randomly manipulating image samples (e.g. flipping, rotation, hue variation, or the addition of noise), is another technique used to improve the performance of CV models trained on small datasets.

Transfer learning and data augmentation are important for the development of aquatic NIS CVMs, given the difficulty and cost of acquiring underwater imagery, as well as factors such as turbidity, variable lighting, and complex backgrounds leading to images with noise, distortion, uneven illumination, low contrast, and limited visibility[11,19,20]. Jian et al.[19] reviewed models for underwater object detection, noting several key challenges—including poor quality imagery compounded by outdated equipment, a lack of sufficient quality, underwater image datasets, and architectural limitations in small object detection and distinguishing foreground and background distortions. Jeong, et al.[21] found that some of these limitations, specifically dataset size, turbidity and lighting, can be mitigated by the choice of utilised backbone model (in their case, EfficientNet). However, Jeong, et al.[21] also noted that the accuracy and precision of object detection models is significantly impacted by the quality of the training and validation imagery (e.g. resolution, sharpness, noise, contrast, colour accuracy, artefacts, brightness and exposure, motion blur, occlusion, and clutter), with higher quality training images generally leading to better performance and poor quality training images degrading performance, which highlights the importance of training imagery quality guidelines in an aquatic context.

Field deployment is a significant expense in aquatic biosecurity and so the concurrent detection and removal of pest species is optimal to reduce the overall expense of a biosecurity programme. This goal requires deployed CVMs to be capable of achieving near or actual real-time inferencing. In practical terms, for a model to be considered real-time, the frames per second (FPS) achieved by a CVM during inferencing must match the video rate (typically 24FPS or greater). Various model architectures exist that are capable of real-time detection[22,23], however Jeong et al.[21] struggled to achieve real-time inferencing with an architecture that is suitable for real-time detection (EfficientNet) at the standard video resolutions of 1024 × 576 (PAL) and 1280 × 736 (HD) reporting FPS

of only 0.7 to 11FPS. Depending on the architecture, various techniques (e.g. quantization aware training) exist for reducing run-time memory overhead and increasing inference speed.

Although there are increasing examples of CVMs used for detecting and delimiting aquatic NIS[9,13,16,21], existing examples have lacked a clear approach or standardised assessment framework. This makes comparisons across species and biosecurity programmes challenging. We present a biosecurity specific framework to improve the quality and efficiency of CVM development, deployment and evaluation for use in aquatic biosecurity surveillance (Fig. 1). We demonstrate the utility of our framework through its application to several aquatic NIS in New Zealand: the invasive marine species Mediterranean fanworm *Sabella spallanzanii*, and exotic *Caulerpa* (*Caulerpa brachypus* and *C. parvifolia*), and the invasive submerged freshwater macrophyte South African oxygen weed *Lagarosiphon major*. Following these case studies, we highlight challenges and avenues for further improvement.

Methods

We present a generalised framework for CVM development and deployment suitable for use in a range of aquatic NIS (Fig. 1). This provides a standardised approach for efficiently developing, evaluating and deploying CVMs in aquatic biosecurity programmes. We outline our approach in three main stages: (1) data collection, (2) model development, and (3) deployment.

Data collection

Data collection includes data acquisition, review and annotation; we consider image preprocessing, data augmentation and dataset splitting in the model development stage. Data acquisition is the most challenging of these steps given the inherent difficulties in underwater image acquisition compounded by the potential for NIS specimen scarcity early in a new incursion.

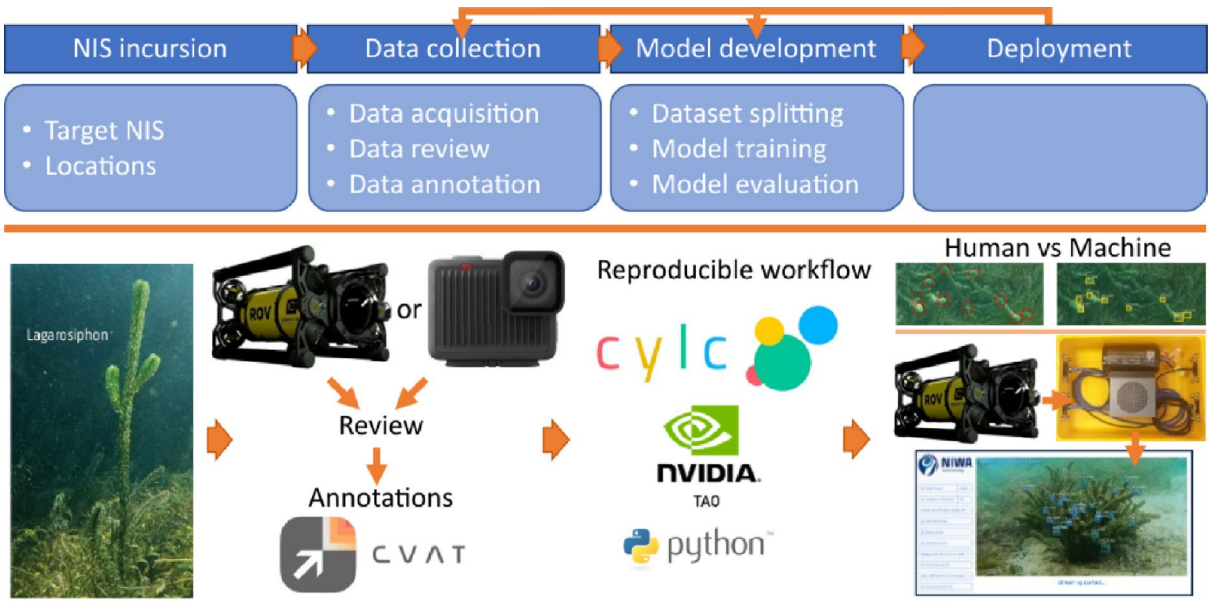


Fig. 1 Top: A general framework for the development and deployment of aquatic NIS detectors for biosecurity applications. Data acquisition, and deployment require biosecurity and aquatic specific considerations. Bottom: an indication of the shared hardware, software and processes used to develop and deploy the NIS CVMs presented in the case studies.

Data acquisition

Given the importance of training image quality and dataset size on CVM performance[21], we recommend the following guidelines to improve the quality of acquired imagery. (1) Use a high-quality camera sensor and lens with imagery captured at high resolution. (2) Minimise turbidity, bubbles, distortions, blur and control lighting, (3) Collect images with the NIS as the main subject, without blurred or partial instances of the NIS in the foreground or background. (4) Collect images of the NIS from a range of angles and positions, (5) Collect images of many different individuals of the target NIS organism, (6) focus on specimens of characteristic appearance before collecting footage of atypical individuals. (7) Collect imagery of the wider aquatic ecosystems where the CVM will be deployed without the NIS present.

Imagery can be collected by a wide variety of cameras and deployment platforms. Across our case studies, we used a BoxFish ROV, a Blue Robotics ROV, snorkeler-held and surface vessel-mounted GoPro cameras, and a diver-held DSLR camera to acquire imagery of the target NIS and wider aquatic ecosystem. Imagery can be collected as video or still photos by ecologists with taxonomic expertise performing biosecurity surveillance and control fieldwork. Labelled image repositories (e.g. iNaturalist, FishNet) are another potential source of imagery for CVMs[13], although these were not used for this study.

Data review

The acquired NIS and aquatic ecosystem imagery must be reviewed to select individual images for annotation and model training. We organised unlabelled imagery by location and date and then triaged files into potential imagery of the target NIS or the wider aquatic ecosystem. Organising unlabelled imagery is particularly important as the scale of imagery increases. While automated clustering approaches can provide order to an unorganised image dataset[12], in our mid-size organisation we found direct consultation with taxonomists involved in pest control to be an efficient way to both narrow down the imagery to screen, and to discuss the characteristic features that may best be used to identify the target organism.

Potential imagery should then be reviewed using the image acquisition guidelines to select images and video frames for annotation. We used the OpenCV Python library or FFmpeg command line tool to extract relevant frames from video, while preserving their frame size and resolution. Specifically, we thinned video to between one to five FPS depending on any movement of the pest and the relative camera to pest movement to reduce interframe duplication of target angle and position.

Data annotation

The selected images must be annotated to define bounding boxes around all target NIS (or other defining NIS characteristics) in each frame. We used CVAT.ai to streamline the annotation process. Given the small size and fine detail of some of our NIS, we found it important to preserve the full resolution of the images during visualisation for annotation. When video frames were annotated, we used hotkeys to streamline the annotation process, and bounding box tracking between annotated frames extracted from a video.

The initial annotation of each NIS was performed by a single investigator. These annotations were reviewed with another team member to discuss the strategy and how to treat any images where the extent or presence of NIS was ambiguous. The agreed strategy was documented to allow subsequent annotations by the future personnel.

Model development

We used a Cylc[24] scheduler to create a fully integrated workflow for reducibly and efficiently performing model development across different NIS. The workflow behaviour was controlled by a configuration file to enhance reproducibility. The configuration file specified the selected datasets for training and validation, model size,

training epochs, model architecture, pretrained classifier, and the full set of datasets for testing. This approach facilitated iterative CVM improvement through repeated model training as the training data, model architecture, and the pre-trained classifier were varied. Inference with video reconstruction of unseen footage was also implemented to facilitate qualitative review prior to deployment. All model development was performed on the New Zealand eScience Infrastructure high-performance computer.

Dataset splitting

We used Python scripts within our workflow to randomly subdivide the annotated imagery into validation, training and test datasets based on the split ratios specified in the configuration file. Each annotated dataset was subdivided separately so that previously trained models can be reevaluated across all test data as more data is collected, streamlining iterative CVM development. A split ratio of 8:1:1 for train, validation and test was applied across all annotated datasets for each NIS.

Model training

We used NVIDIA's TAO toolkit to perform model training as it integrates with the hardware (Jetson Orin) used for deployment. Specifically, our workflow used Python scripts to perform data preprocessing and to create TAO compliant configuration files and bash scripts before performing model training using TAO. The main preprocessing steps included resizing the split datasets to match the CVM resolution, running k-means clustering to characterise the default annotation box sizes of the dataset, and populating the portions of the TAO configuration file that were varied between runs. The varied portions of the TAO configuration file included the default annotation box sizes, the model architecture, the pre-trained classifier, the training dataset, the model resolution, and the number of training epochs to run. We used the NVIDIA recommended data augmentation steps (e.g. jitter, hue, exposure, random mosaic). Model training was performed on a NVIDIA A100 GPU. Upon training completion, the trained model with the largest mean average prevision (mAP) was selected for evaluation and export. Our workflow also supported 'hot starts' for partially completed training runs that were interrupted before completion to reduce overall time and compute during completion of training.

Model evaluation

Our workflow used the mAP, Recall (R), and false negative rate (FNR) to evaluate each trained model across the test datasets specified in the workflow configuration file. These test datasets can be updated to include more test images than those used in training to allow for comparison between CVMs as the amount of available training imagery grows. mAP is the mean across all classes of the model's precision across a range of different intersection over union's (IOU) and thresholds. R is the true positives over actual instances at a fixed IOU and detection threshold. The FNR is the false negatives over true positives and false negatives at a fixed IOU and detection threshold[21,25]. We used a 50% IOU and detection threshold. Further implementation specific details of how our workflow calculated the CVM mAP, R and FNR can be found in the Zenodo repository: https://doi.org/10.5281/zenodo.22103946.

Deployment

The field deployment of any NIS detector is critical given the operational focus of biosecurity. For our case studies, we developed a system for deploying the CVMs for aquatic biosecurity surveillance on an NVIDIA Jetson AGX Orin to perform real-time NIS detection. This system has flexibility to receive video from a range of different ROV, surface vessel and handheld camera systems. Using this method, NIS detections are integrated alongside timestamp and GPS information, which is displayed in real-time through a webpage interface. The

webpage interface supports CVM selection, the selection of the detection confidence thresholds, starting or stopping inferencing, and the ability to download inferenced video, still images and detection location files. The still images allow for validation of target detections.

Hardware configuration

Since the hardware for deployment differs from that used in CVM training, our NVIDIA A100 trained CVMs were reconfigured and compiled for deployment on the NVIDIA Jetson AGX Orin using NVIDIA's TensorRT SDK. Similarly, the smaller memory and processing footprint of the NVIDIA Jetson AGX Orin, required Quantization Aware Training (QAT) to allow for the optimised selection of FP32, FP16 or INT8 precision for each model layer to achieve real-time inference while maintaining model accuracy during conversion (Table 2; explored further in the case studies below). Our system also processed the video feed of the attached camera system to match it to the expected video format, resolution and frame rate of the CVM. We integrated our hardware with BoxFish ROV and GoPro camera platforms.

Field validation

Field validation provides an assessment of the CVM performance in the operational environment. This can be performed live or with prerecorded footage. Qualitative assessment may be sufficient early in the development of a CVM for a new NIS as instances of poor performance are identified and used to target the collection of new training imagery. However, the quantitative assessment through the calculation of model R and false positive count (FPC) compared to a human expert is important for any operationalised CVM. We calculated these estimates relative to each organism encountered across the video feed. This meant that each organism only needed to be detected once within a single frame to be counted. R and FPC can be filtered to consider only detections across multiple frames (e.g. two frames) if detection persistence is a desired feature of a CVM. These formulations of R and FP differ to the frame-by-frame metrics used during model development to ensure that these estimates align with practical aims of a biosecurity surveillance programme.

While quantitative CVM assessment can be applied to ad hoc video of variable duration, flight path, and speed, such variation makes it difficult to compare R and FPC between location or different operational settings (e.g. speed or distance to target). Therefore, field validation ideally involves multiple transects representing different pest abundance levels across several locations at several fixed camera speeds and heights. Such standardisation allows for the assessment of current and future CVMs at different operational settings and pest abundances, prior to potential operational deployment.

Case studies

We used our framework to develop and deploy CVMs for the surveillance of three aquatic pest in New Zealand marine and freshwater environments. For *S. spallanzanii* we explored the impact of CVM architecture on performance, and training configuration on inference speed. For the exotic *Caulerpa* case study we demonstrate rapid CVM improvement through successive field deployments. For the *Lagarosiphon major* case study, we demonstrate how field validation allows for comparison of model performance between locations and other operational settings.

Case study 1: *Sabella spallanzanii*

The Mediterranean fanworm *Sabella spallanzanii* is a highly invasive marine pest first recorded in New Zealand in March 2008. It is currently present at multiple locations across both the North and South islands[26,27]. The

species is demonstrably harmful to benthic ecosystems and shellfish aquaculture[28–31], and is a target for ongoing monitoring within the long-running, six monthly New Zealand Marine High Risk Site Surveillance (NMHRSS) programme[32]. *S. spallanzanii* tube length typically ranges from 10-50 cm with a tentacle crown width of 15 cm. We developed an ROV-deployed CVM for *S. spallanzanii* and evaluated against a human across two videos from two different locations (Fig. 2). Full methods and results are provided in the Supplement.

Data collection

A Boxfish ROV was utilised alongside divers during the 2019 survey season to give approximately 12 h 15 min of 4 k video footage. The video footage was reviewed over approximately 18 h 45 min to identify 19 twenty-second video clips of suitable quality for training. This footage contained 11 *S. spallanzanii* organisms from the Auckland Viaduct, and representative marine ecosystems from the Auckland Viaduct, Nelson Marina, and Whakaraupō Lyttleton Harbour. The video clips were thinned to approximately five FPS, yielding 100 frames per video clip and a total of 1900 frames. These video clips were annotated and reviewed within the CVAT.ai platform. A total of 43 h was spent in video review and annotation.

Model development

The model development workflow was used to explore the impact of dataset size, diversity, model resolution and CVM architecture on model performance (Supplement 1), before the model architectures (YOLOv4 ResNet50 and GoogLeNet) and resolutions (PAL (1024 × 576) and FHD (1920 × 1080) with the best mAP, R and FNR were selected for deployment.

Deployment

We used QAT to improve the inference speed of the deployed models. QAT allows for the optimised selection of FP32, FP16 or INT8 precision for each model layer during CVM compilation to the deployment hardware to reduce the model size and memory requirement. The CVM throughput was reported (Table 1) by the on-device queries per second (QPS). This showed that with FP32-16 or INT8 optimised layer weights provided the fastest QPS. Multiple detections in a frame will result in a lower inference FPS than the reported QPS. The inference FPS of the FP32-16 or INT8 optimised CVMs was recorded (Table 1) with both CVMs achieving greater than real time inference rates. The configured camera frame rate of 30 FPS provided an upper limit on the overall system FPS.

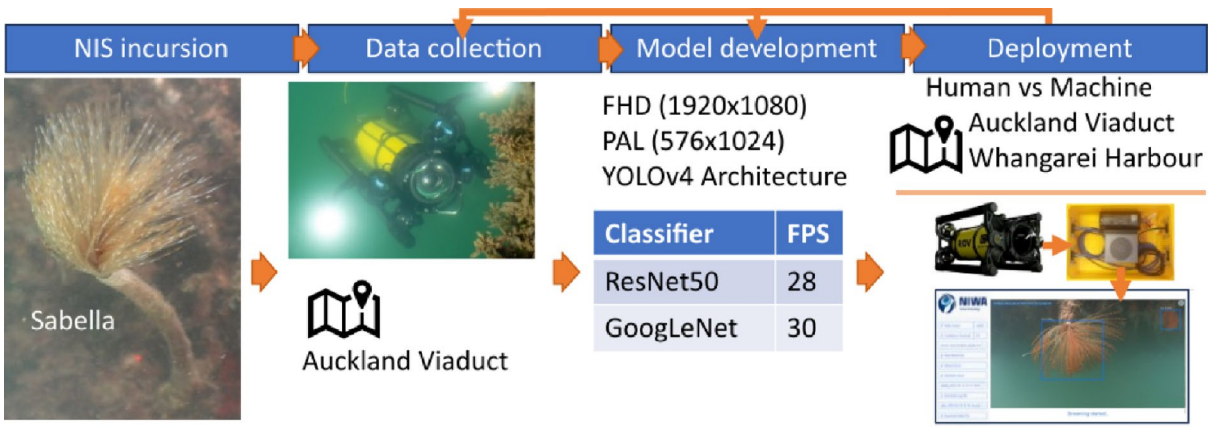


Fig. 2 The workflow to CVM development and deployment as applied to *S. spallanzanii*.

Table 1 The inference rate in QPS as measured on device with all model layer numeric precision at FP32, optionally FP32 or FP16, optionally FP32, FP16 or INT8; and the inference rate in FPS, the R, and the FPC when applied to 60 s videos from two field deployments of four *S. spallanzanii* CVMs deployed on a Jetson Orin.

Model	Input size	Jetson throughput (QPS)			Auckland Viaduct			Whangarei Harbour		
	Height × width	FP32	FP32-16	FP32-16, INT8	FPS	FPC	R (%)	FPS	FPC	R (%)
ResNet50	576 × 1024	32.2	62.3	136.4	30.0	0	66.7	28.2	0	16.6
GoogLeNet	576 × 1024	45.5	88.6	105.9	30.0	1	73.3	30	18	16.6
ResNet50	1920 × 1080	10.3	19.7	32.5	28.3	0	66.7	30	0	16.6
GoogLeNet	1920 × 1080	14.9	28.5	41.5	30.0	12	53.3	30	3	16.6

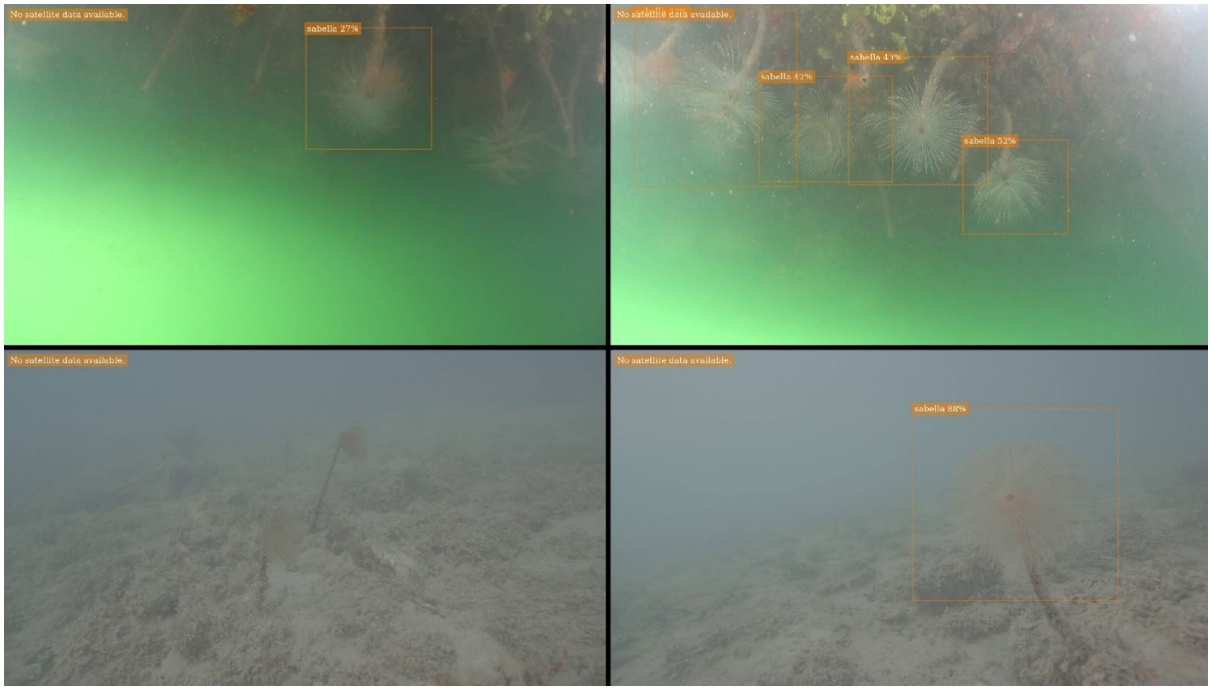


Fig. 3 Frames illustrating the relative performance of the *S. spallanzanii* ResNet50 YOLOv4 FHD CVM collected at Auckland Viaduct (top), and Whangarei Harbour (bottom) showing instances of false negatives (left), and successful detections (right).

We assessed the performance of the deployed models across two videos of *S. spallanzanii* that was not used in CVM development. The R and FPC we calculated across each video filtered to consider detections across at least two frames (Table 1). The Auckland Viaduct video was collected on the same day and location as the imagery used to train the models, while the Whangarei Harbour video was collected in 2025 at a different location. Both videos include a mix of front and side-on *S. spallanzanii* individuals with differing levels of blur, turbidity, and size (Fig. 3).

Results (Table 1) demonstrated the importance of exploring different approaches for improving the inference rate during model deployment, and field validation in CVM evaluation. Comparing the GoogLeNet and ResNet50 CVMs showed that the GoogLeNet produced far more false detections, despite similar performance during model development (Supplement). Finally, a comparison of model performance across the two sites showed a substantial degradation in performance at Whangarei Harbour (Table 1, Fig. 3). This video was both from a different site and included a greater proportion of side-on and distant imagery of *S. spallanzanii*.

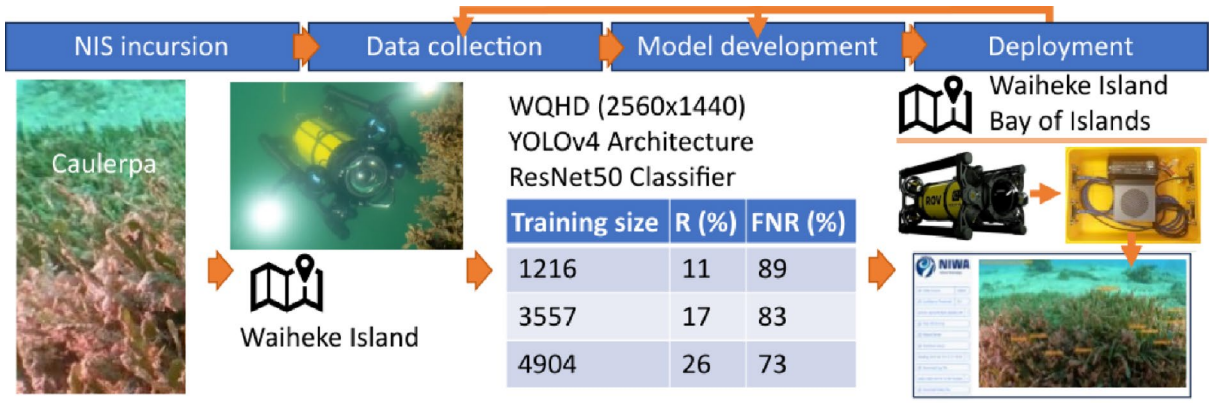


Fig. 4 The workflow to CVM development and deployment as applied to exotic *Caulerpa*.

Table 2 Evaluation results during the exotic *Caulerpa* CVM development.

Training dataset		Test imagery of all iterations (105 frames, 499 instances)		
Iteration	*Caulerpa* instances	mAP (%)	R (%)	FNR (%)
1	1216	26	11	89
2	3557	51	17	83
3	4904	59	26	68

The evaluation metrics were rerun over the complete test dataset from all three iteration of training data to ensure a fair comparison.

Case study 2: exotic *Caulerpa*

Caulerpa brachypus and *C. parvifolia* are two species of non-indigenous marine algae (collectively referred to as exotic *Caulerpa*) that were first detected at Aotea Great Barrier Island in the Hauraki Gulf in July 2021[33–35]. Biosecurity New Zealand subsequently initiated a biosecurity response to delimit exotic *Caulerpa* distribution, manage its spread, and eliminate local populations at several locations. As of 2026, exotic *Caulerpa* is present at multiple locations around the Hauraki Gulf, the Coromandel, and Northland in northern New Zealand. Surveillance has primarily been undertaken using spot-dives, transect surveys, and radial mapping. Exotic *Caulerpa* fronds typically range from one to 10 cm tall and 0.5 to 2 cm wide. We developed successive ROV-deployed CVM for exotic *Caulerpa* for field campaigns (Fig. 4). Full methods and results are provided in Tait et al.[34].

Data collection

Since *C. brachypus* and *C. parvifolia* are morphologically indistinguishable in the field[35], the species were treated as a single NIS for CVM development. During our initial model development, we only had access to 2 h 20 min video footage of exotic *Caulerpa* video from a single location (Waiheke Island) collected over a three-day period in June 2024. For our first CVM, we identified a 15 s section of suitable *Caulerpa* (close up, in focus, low turbidity, a range of angles) in 30 min of video review, which was thinned to five FPS to give 75 frames which were annotated over an 8-h period to produce 1216 instances of *Caulerpa*.

Model development and model iteration

The annotated imagery was used alongside the representative marine ecosystem imagery identified during the *S. spallanzanii* CVM development to produce a WQHD (2560 × 1440) model (Table 2). A WQHD resolution was used given the small size of the exotic *Caulerpa* fronds especially during emergence of new fronds to ensure a reasonable number of pixels were retained for each frond. This was qualitatively accessed on the deployment

hardware using another segment of the prerecorded exotic *Caulerpa* video over a 30-min period. Two snippets (25 s and 20 s) of video with high FNR were identified and thinned to two FPS. The first was annotated over an additional 8-h period to give an additional 2341 (total of 3557) instances of *Caulerpa*. This was combined with the previous training data and used to produce another CVM (Table 2).

Deployment and model iteration

This second CVM was deployed during survey at Waiheke Island in March 2025 at several sites with depths of 4-18 m. This showed the CVM was able to identify individual fronds but struggled in densely overlapping patches of *Caulerpa* and instances of small *Caulerpa* frond size. Prior to the next field campaign, the second of the previously identified video snippets with high FNR which had more instances of both dense and small *Caulerpa* was annotated over another 8-h period and used to train a third model (Table 2).

In late March 2025, the third CVM was deployed during a field campaign in response to the emerging *Caulerpa* invasion in the Bay of Islands at several sites with depths of 20-35 m. This showed the latest CVM could both detect more *Caulerpa* fronds in an area of dense infestation and also from a greater distance.

Exotic *Caulerpa* is morphologically variable and can be found in dense or sparse patches, and individual fronds often vary in length and shape. It can also be interspersed with other organisms (Fig. 5a), or covered in sediments (Fig. 5b). Despite this variability, the CVM was able to identify individual fronds within every imaged clump, and in some cases at a reasonable distance (Fig. 5c). A small number of false-positives were noted when samples were interspersed with seagrass (*Zostera muelleri*), although detection confidence in such cases was very low (Fig. 5d). The reduced time to produce three models (a total of 25 h) demonstrates the increased efficiency through repeated use of the framework.

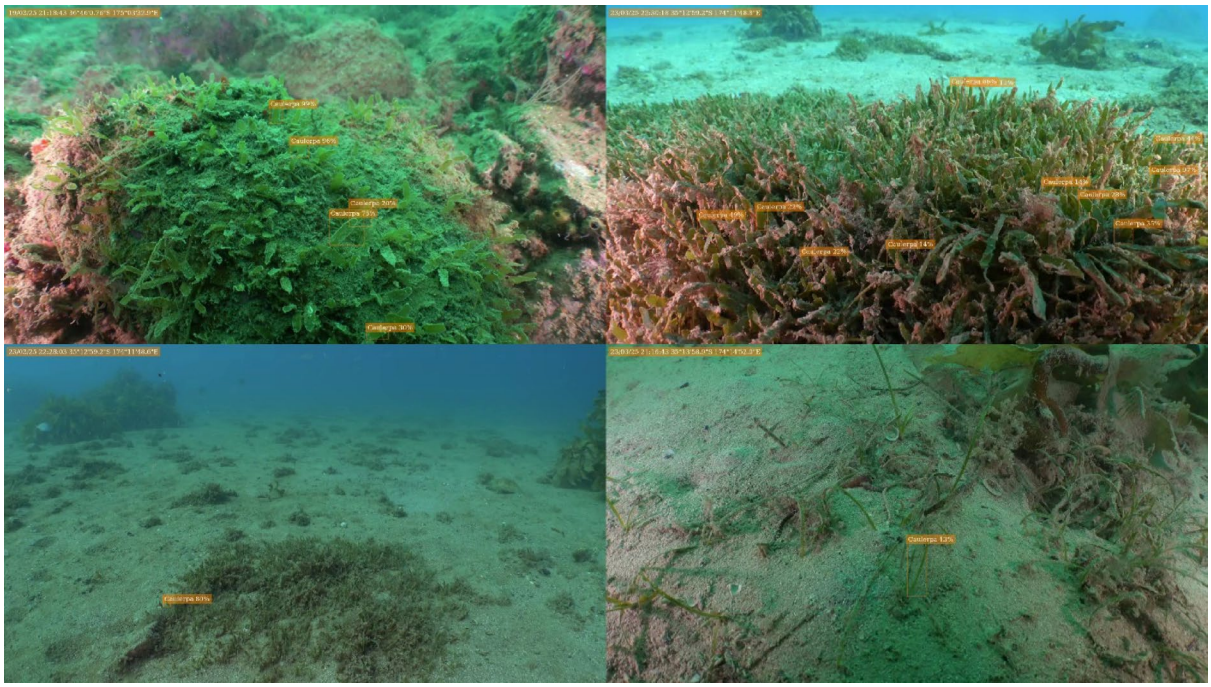


Fig. 5 Detection of exotic *Caulerpa* by the ResNet50 YOLOv4 WQHD CVM. (top left) Exotic *Caulerpa* on rocky reef habitat, (top right) Exotic *Caulerpa* on soft sediment habitat, (bottom left) Exotic *Caulerpa* detected at distance, (bottom right) False detection of seagrass as exotic *Caulerpa*.

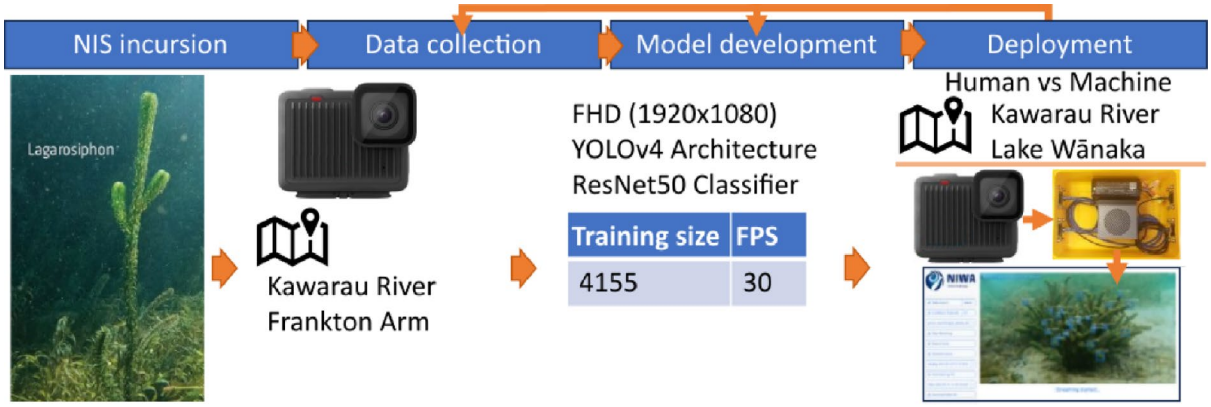


Fig. 6 The workflow to CVM development and deployment as applied to *L. major*.



Fig. 7 Frames illustrating the *L. major* detections by the CVM during field testing.

Case study 3: *Lagarosiphon major*

South African oxygen weed (*Lagarosiphon major*) is an invasive, submerged freshwater macrophyte first recorded in New Zealand in 1950[36]. It is widely naturalised in the North Island and locally abundant and actively spreading in South Island waterbodies. It impacts indigenous plant biodiversity, hydrogeneration, irrigation, flood protection and recreational values. Reproduction is vegetative by stem fragments and dispersal between waterbodies is human mediated (e.g. via watercraft and trailers, fishing nets or drainage machinery contaminated with *L. major* fragments). Only female plants are known to be present in New Zealand[36–38]. We developed and field tested a surface vessel-deployed CVM for *L. major* (Fig. 6). *L. major* is a bottom-rooting perennial with stems reaching lengths of up to 5 m. Leaves and shoots are dark green and leaves are alternate, arranged spirally around the stem and typically are curved downwards, 6–20 mm long[38]. Full results are provided in the Supplement.

Data collection

Video footage was collected with GoPro 7 and 9 cameras by divers from the Kawarau River and Frankton Arm of Lake Whakatipu in the South Island across multiple fieldwork trips between December 2021 and May 2023. All images had a resolution of 2704 × 1520. These were reviewed for low blur images containing *L. major* representing a range of abundances, and plant sediment cover. In total 657 frames were selected for annotation to give a total of 4155 instances of *L. major* (Fig. 7).

Table 3 The overall detection rates (%) and average number of false positive detections for the CVM across all sites at different camera-to-target distances (m) for the combined boat speeds and plant sedimentation runs.

Site		Number of plants	Abundance	Distance to target (m)	R		Average false positives per run	
					Mean (%)	SE	Mean	SE
Kawarau River	Site 1	1	Sparse	0.85	15.6	0.1	0	0
				1.25	0.0	0.1	0	0
	Site 2	7	Medium	0.2	61.5	0.1	0	0
				0.7	0.0	0.1	0	0
	Site 3	2	Sparse	0.3	61.5	0.2	1.4	1.0
				0.7	7.7	0.2	0.1	0.3
Lake Wānaka Wanaka	Site 4	2	Medium	0.3	81.3	0.2	2.1	0.9
				0.75	6.3	0.2	0	0
	Site 5	4	Medium	0.45	9.4	0.1	0.9	1.4
				0.85	0.0	0.1	0	0
	Site 6	4	Medium	0.35	70.0	0.1	1.1	1.4
				0.85	4.2	0.1	0	0

Model development

A single 1920 × 1080 resolution YOLOv4 ResNet50 pretrained classifier CVM was developed in 2024 using the workflow. The selected architecture was that of the best performing CVM in our *S. spallanzanii* and exotic *Caulerpa* case studies.

Deployment

The model was deployed on a small boat with a video captured with a GoPro7 or GoPro9 mounted on a submerged, 45° downward-facing pole. The depth of the pole could be controlled during each flight to maintain a constant distance above the lake or riverbed. The CVM was field tested along fixed transects at six sites in December 2024—three in the Kawarau River and three in Lake Wānaka. The sites were selected to evaluate detection of isolated infestations across a range of *L. major* plant abundance (sparse or medium, early stage plant detection). At each site the CVM was tested at two boat speeds, two camera distances to target, and pre- and post-sediment removal from the plants (dirty and clean plants, respectively).

Each transect was first surveyed by an ecologist to identify all *L. major* plants (ranging from a single shoot through to a dense clump with many shoots) and then surveyed two times by the boat-mounted CVM system for each speed (1kn, 3 kn), camera distance (Table 3) and plant sedimentation (dirty or clean). A detection was reported if at least one of the plants shoots were detected per individual plant, and a false positive was reported for a detection of anything aside from a surveyed *L. major* shoot. As no statistical difference was observed between the two boat speeds or plant sedimentation ($p > 0.05$, t-test, Supplement), average detection rates and false positive counts across the eight runs are combined (Table 3). No significant difference in performance between the Kawarau River and Lake Wānaka was found ($p > 0.05$, t-test) demonstrating good transferability of the CVM to other locations. The CVM showed significantly reduced performance at the greater distances to target in all but one site ($p < 0.05$, z-test) suggesting a higher resolution CVM may be beneficial.

Discussion

Here we present a biosecurity-focused framework for aquatic pest CVM development and deployment, which we applied to three aquatic NIS in New Zealand: Mediterranean fanworm *S. spallanzanii*, South African oxygen weed *L. major,* and exotic *Caulerpa* (*C. brachypus* and *C. parvifolia*). The framework was purposefully flexible

to allow seamless application to a wide range of aquatic pests in different aquatic environments across several imaging platforms. Our case studies span ROV-based deployment to depths of 35 m for exotic *Caulerpa* in marine environments, and surface vessel deployments to depths less than 3 m for *L. major* in freshwater lakes and rivers.

We investigated the CVM training and hardware configuration required to achieve real-time inferencing on a Jetson Orin in our *S. spallanzanii* case study, before we applied the same general configuration to the *L. major* and exotic *Caulerpa* case studies. We found our framework suitable for rapid CVM improvement through repeated CVM deployment to the Jetson Orin for inference of both prerecorded video, as well as live video captured in the field. This approach allowed for qualitative assessment of the CVM performance in a field setting and the rapid inclusion of footage targeting instances of poor performance in the next CVM. The exotic *Caulerpa* CVM was produced by the same team as the *S. spallanzanii* CVM in a substantially reduced timeframe from 45 h for one iteration of training imagery to 25 h for three iterations of training imagery.

Across our case studies we identified several key parameters that were essential to achieve accurate and repeatable detections. In the *L. major* case study, we identified that plant sedimentation, vessel speed and the field locations tested did not have a significant impact on the CVM performance, while increased distance to target resulted in reduced performance. Distance to target was a key contributor to model performance in the *Caulerpa* study. We show that the impact of an increased distance to target on the size (and number of pixels) of the image target can be offset by increasing model resolution. Increasing the resolution of inferencing is essential for small targets, or species where the discriminating characteristics require detailed information. In this regard, all CVM models must find a balance between the resolution required for feature preservation and constraints on model processing capabilities[39,40].

Across our different CVMs, we had different dataset sizes and diversity. While we cannot directly compare our CVMs given different levels of field assessment, we can make some broad observations. Unsurprisingly, greater dataset size and diversity (organisms, positions, camera orientation, location, etc.) resulted in greater CVM performance. We also identified that CVMs for more visually characteristic NIS can be produced from a smaller dataset size. This is suggested by our *S. spallanzanii* CVM compared to our *Caulerpa* and *L. major* CVMs. These results are corroborated by studies on highly characteristic insects[41], compared to more subtle tasks like plant disease detection[42]. Finally, we show that rapid CVM field deployment, qualitative review and identification or imagery for iterative improvement, and eventual field validation for a future CVM is an operationally efficient approach to aquatic NIS CVM development and deployment.

Across our case studies, we have found that the involvement of trained taxonomists throughout the CVM development and deployment process reduces the development time and increases model performance through their in-depth knowledge of the operating environment and morphology of the target species[43]. This finding is especially true for field validation design, which was exemplified in the *L. major* case study.

In the interest of improving rapid surveillance of aquatic ecosystems for invasive species, it is essential that robust guidance for how to optimise imagery capture for training and implementation of models. While advances in terrestrial and agricultural landscapes have been rapid[44], far less has been achieved in aquatic ecosystems. We point to the unique challenges of visual detection in aquatic environments and identify key limitations relating to distance to targets and challenges with visibility due to water clarity[9]. Our studies provide a framework for model development and improvement unique to aquatic pests and will ideally pave the way for the use of imaging autonomous platforms.

We anticipate that future studies will further demonstrate the transferability of our framework by using it to develop CVMs for other aquatic NIS worldwide. We note, however, that not all NIS lend themselves to detection by CVMs. NIS that are visually indistinguishable from non-target species in operationally representative imagery will require other detection strategies. Here our framework could be extended to incorporate data from alternative sensors such as echosounders, hydrophones or eDNA sensors alongside updated data acquisition guidelines, model architectures, and deployment hardware[13,16]. Similarly, the framework is purposefully agnostic to model architecture so that it can easily be adapted to the latest CVM architecture suitable for aquatic imagery. While we

have focused on NIS detection, the framework presented here could also be used to develop aquatic CVMs for non-biosecurity applications such as fisheries or ecological research.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-72212-8.

Acknowledgements The authors thank Crispin Middleton, Oliver Evans, Matt Smith, Peter Marriott, Jules Burchardt, Richie Hughes and Martin Chew for their involvement in the *Caulerpa* surveillance campaigns where various iterations of the *Caulerpa* CVM was used. We also thank Hayden McDermott, Deborah Hofstra and Mary de Winton for their support for conducting the *L. major* trials. We thank Alan Tan for his thoughtful feedback on the manuscript. The authors wish to acknowledge the use of New Zealand eScience Infrastructure (NeSI) high performance computing facilities, consulting support and/or training services as part of this research. New Zealand's national facilities are provided by NeSI and funded jointly by NeSI's collaborator institutions and through the Ministry of Business, Innovation & Employment's Research Infrastructure programme. https://www.nesi.org.nz.

Author contributions R.A.P. led the development of the framework, and Sabella and Caulerpa computer vision models, with supervision from L.W.T. G.A.P and D.C. led the development of the Lagarosiphon computer vision model. G.A.P., J.B., D.C., S.D. and L.W.T. assisted with image acquisition, the development of the computer vision models, hardware testing, and field deployment. R.A.P. and F.Z. wrote the manuscript, which was reviewed by all authors. R.A.P. produced the figures, which were reviewed by all authors.

Funding This research was supported by the New Zealand Government's Strategic Science Investment Fund (SSIF; COBS2102, CEBS2402, CEBS2502, FWBS2501, FPBS2603, C0036821) to the National Institute of Water and Atmospheric Research (NIWA) and Earth Sciences New Zealand, as well as the Ministry for Primary Industries Marine High Risk Site Surveillance programme (SOW18048). Department of Conservation collaborated and contributed to the development and testing of the *L. major* detection system (DOC22213, DOC24801).

Data availability The annotated imagery used to develop the *S. spallanzanii* CVMs, as well as inferenced videos from two sites of the four field deployed models, are shared in a Zenodo repository: https://doi.org/10.5281/zenodo.22103946. The repository also includes a sterilised version of the Cylc workflow used to control the model development process. This has been sterilised to remove any path and configuration details specific to the New Zealand eScience Infrastructure high-performance computing environment. Finally, the configuration files used to produce the CVMs presented in the case study sections are included. These show the dataset, architecture, and backbone settings used to train each model.

Declarations

Competing interests We declare a competing interest, as defined by Nature Research, as this research was funded by two New Zealand government grants. This research was supported by the New Zealand Government's Strategic Science Investment Fund (SSIF; COBS2102, CEBS2402, CEBS2502, FWBS2501, FPBS2603, C0036821) to the National Institute of Water and Atmospheric Research (NIWA) and Earth Sciences New Zealand, as well as the Ministry for Primary Industries Marine High Risk Site Surveillance programme (SOW18048). Department of Conservation collaborated and contributed to the development and testing of the L. major detection system (DOC22213, DOC24801).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Cuthbert, R. N. et al. Biological invasion costs reveal insufficient proactive management worldwide. *Sci. Total Environ.* **819**, 153404. https://doi.org/10.1016/j.scitotenv.2022.153404 (2022).

2. Gallardo, B., Clavero, M., Sánchez, M. I. & Vilà, M. Global ecological impacts of invasive species in aquatic ecosystems. *Glob. Change Biol.* **22**, 151–163. https://doi.org/10.1111/gcb.13004 (2016).
3. Jurdak, R. et al. Autonomous surveillance for biosecurity. *Trends Biotechnol.* **33**, 201–207. https://doi.org/10.1016/j.tibtech.2015.01.003 (2015).
4. Seebens, H. et al. Projecting the continental accumulation of alien species through to 2050. *Glob. Change Biol.* https://doi.org/10.1111/gcb.15333 (2021).
5. Inglis, G. J. et al. Surveillance design for early detection of unwanted exotic marine organisms in New Zealand. Report No. 2005-17, 110 (Ministry of Agriculture and Forestry (MAF), Wellington, 2006).
6. Bax, N. et al. The control of biological invasions in the world's oceans. *Conserv. Biol.* **15**, 1234–1246. https://doi.org/10.1111/j.1523-1739.2001.99487.x (2008).
7. Myers, J. H., Simberloff, D., Kuris, A. M. & Carey, J. R. Eradication revisited: dealing with exotic species. *Trends Ecol. Evol.* **15**, 316–320. https://doi.org/10.1016/S0169-5347(00)01914-5 (2000).
8. Muller, C., Hofstra, D. & Champion, P. Eradication economics for invasive alien aquatic plants. *Manag. Biol. Invasions* **12**, 253–271. https://doi.org/10.3391/mbi.2021.12.2.04 (2021).
9. Tait, L. W. et al. Towards remote surveillance of marine pests: a comparison between remote operated vehicles and diver surveys. *Front. Mar. Sci.* **10**, 1102506. https://doi.org/10.3389/fmars.2023.1102506 (2023).
10. Bao, J., Li, D., Qiao, X. & Rauschenbach, T. Integrated navigation for autonomous underwater vehicles in aquaculture: a review. *Inf. Process. Agric.* **7**, 139–151. https://doi.org/10.1016/j.inpa.2019.04.003 (2020).
11. Hitchin, R., Turner, J. A. & Verling, E. Epibiota remote monitoring from digital imagery: Operational guidelines. 29 (The National Marine Biological Analytical Quality Control Scheme (NMBAQC), 2015).
12. Williams, S. B., Pizarro, O., Steinberg, D. M., Friedman, A. & Bryson, M. Reflections on a decade of autonomous underwater vehicles operations for marine survey at the Australian Centre for Field Robotics. *Annu. Rev. Control* **42**, 158–165. https://doi.org/10.1016/j.arcontrol.2016.09.010 (2016).
13. Campbell, M. L., Le, C. T. U., Abu-Khalaf, J. & Sherman, C. D. H. Leveraging artificial intelligence to enhance marine biosecurity. *Bioscience* https://doi.org/10.1093/biosci/biag049 (2026).
14. Knausgård, K. M. et al. Temperate fish detection and classification: A deep learning based approach. *Appl. Intell.* **52**, 6988–7001. https://doi.org/10.1007/s10489-020-02154-9 (2021).
15. Villon, S. et al. A new method to control error rates in automated species identification with deep learning algorithms. *Sci. Rep.* **10**, 10972. https://doi.org/10.1038/s41598-020-67573-7 (2020).
16. Goodwin, M. et al. Unlocking the potential of deep learning for marine ecology: Overview, applications, and outlook. *ICES J. Mar. Sci.* **79**, 319–336. https://doi.org/10.1093/icesjms/fsab255 (2022).
17. Hussain, M. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection. *Machines* https://doi.org/10.3390/machines11070677 (2023).
18. Ren, S., He, K., Girshick, R. & Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031 (2017).
19. Jian, M., Yang, N., Tao, C., Zhi, H. & Luo, H. Underwater object detection and datasets: A survey. *Intell. Mar. Technol. Syst.* **2**, 9. https://doi.org/10.1007/s44295-024-00023-6 (2024).
20. Gormley, K. et al. Automated image analysis of offshore infrastructure marine biofouling. *J. Mar. Sci. Eng.* https://doi.org/10.3390/jmse6010002 (2018).
21. Jeong, H. S. et al. Identifying *Hydrilla verticillata* in real time with a machine learning–based underwater object detection program. *Aquat. Conserv.* https://doi.org/10.1002/aqc.70054 (2025).
22. Jia, J., Fu, M., Liu, X. & Zheng, B. Underwater object detection based on improved EfficientDet. *Remote Sens.* https://doi.org/10.3390/rs14184487 (2022).
23. Zocco, F. et al. Towards more efficient EfficientDets and real-time marine debris detection. *IEEE Robot. Autom. Lett.* **8**, 2134–2141. https://doi.org/10.1109/LRA.2023.3245405 (2023).
24. Oliver, H. J., Shin, M. & Sanders, O. Cylc: A workflow engine for cycling systems. *J. Open Source Softw.* https://doi.org/10.21105/joss.00737 (2018).
25. Dai, L. et al. Edge-guided representation learning for underwater object detection. *CAAI Trans. Intell. Technol.* **9**, 1078–1091. https://doi.org/10.1049/cit2.12325 (2024).
26. Read, G., Inglis, G., Stratford, P. & Ahyong, S. Arrival of the alien fanworm *Sabella spallanzanii* (Gmelin, 1791) (Polychaeta: Sabellidae) in two New Zealand harbours. *Aquat. Invasions* **6**, 273–279. https://doi.org/10.3391/ai.2011.6.3.04 (2011).
27. Fletcher, L. Background information on the Mediterranean fanwrom *Sabella spallanzanii* to support regional response decisions. Report No. 2479A, 45 (Cawthron Institute, Nelson, 2014).
28. Tait, L. W. et al. Invasive ecosystem engineers threaten benthic nitrogen cycling by altering native infaunal and biofouling communities. *Sci. Rep.* **10**, 1581. https://doi.org/10.1038/s41598-020-58557-8 (2020).
29. Atalah, J. et al. The introduced fanworm, *Sabella spallanzanii*, alters soft sediment macrofauna and bacterial communities. *Front. Ecol. Evol.* **7**, 481. https://doi.org/10.3389/fevo.2019.00481 (2019).

30. Soliman, T. & Inglis, G. J. Forecasting the economic impacts of two biofouling invaders on aquaculture production of green-lipped mussels *Perna canaliculus* in New Zealand. *Aquac. Environ. Interact.* **10**, 1–12. https://doi.org/10.3354/aei00249 (2018).
31. Douglas, E. J. et al. *Sabella spallanzanii* and seafloor biodiversity enhancement in a marine soft-sediment system. *Diversity* https://doi.org/10.3390/d12060228 (2020).
32. Woods, C. et al. National Marine High Risk Site Surveillance: Annual synopsis report for all high risk sites 2023–24 (SOW23030). Report No. 2024/03, 324 (Ministry for Primary Industries, Wellington, 2024).
33. Middleton, I., Pande, A., D'Archino, R. & Middleton, C. Second delimiting survey of the non-native alga, *Caulerpa brachypus*, at Aotea/Great Barrier Island. Report No. 2021339WN, 33 (NIWA, Wellington, 2021).
34. Tait, L. W., Middleton, C., Pearson, R. A., Preston, G. & Bulleid, J. Advanced surveillance of exotic Caulerpa: Using computer vision for automated detection. Report No. 2025094CH, 59 (Earth Sciences New Zealand, Christchurch, 2025).
35. Response Planning Function. Technical Advisory Report: *Caulerpa* Great Barrier Island 2021 Response. 25 (Biosecurity New Zealand, 2022).
36. McGregor, P. & Gourlay, H. Assessing the prospects for biological control of lagarosiphon (*Lagarosiphon major* (Hydrocharitaceae)). Report No. 57, 13 (Department of Conservation, Wellington, 2002).
37. Paynter, Q. Feasibility of biocontrol of *Lagarosiphon major* in New Zealand. 25 (Landcare Research Manaaki Whenua, Palmerston North, 2013).
38. Champion, P. et al. Freshwater invasive species of New Zealand 2020. 68 (NIWA, Hamilton, 2020).
39. East, A. et al. Optimizing image capture for computer vision-powered taxonomic identification and trait recognition of biodiversity specimens. *Methods Ecol. Evol.* **16**, 2260–2275. https://doi.org/10.1111/2041-210x.70140 (2025).
40. Hudson, L. N. et al. Inselect: Automating the digitization of natural history collections. *PLoS ONE* **10**, e0143402. https://doi.org/10.1371/journal.pone.0143402 (2015).
41. Knyshov, A., Hoang, S., Weirauch, C. & Jockusch, E. Pretrained convolutional neural networks perform well in a challenging test case: Identification of plant bugs (Hemiptera: Miridae) using a small number of training images. *Insect Syst. Divers.* https://doi.org/10.1093/isd/ixab004 (2021).
42. Shoaib, M. et al. Deep learning for plant bioinformatics: An explainable gradient-based approach for disease detection. *Front. Plant Sci.* **14**, 1283235. https://doi.org/10.3389/fpls.2023.1283235 (2023).
43. Wäldchen, J. & Mäder, P. Plant species identification using computer vision techniques: A systematic literature teview. *Arch. Comput. Methods Eng.* **25**, 507–543. https://doi.org/10.1007/s11831-016-9206-z (2018).
44. Wang, S. et al. Advances in deep learning applications for plant disease and pest detection: A review. *Remote Sens.* https://doi.org/10.3390/rs17040698 (2025).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Rose A. Pearson[1] · Gareth Preston[1] · Jeremy Bulleid[1] · Svenja David[1] · Felix Zareie-Vaux[1] · Daniel Clements[1] · Leigh W. Tait[1,2]

✉ Rose A. Pearson
rose.pearson@earthsciences.nz

[1] Earth Sciences New Zealand, Auckland, New Zealand

[2] School of Biological Sciences, University of Canterbury, Christchurch, New Zealand