Computers and Electronics in Agriculture 207 (2023) 107744

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Using mobile-based augmented reality and object detection for real-time Abalone growth monitoring

Thomas Napier *, Ickjai Lee

Information Technology Academy, College of Science & Engineering, James Cook University, Townsville, QLD 4811, Australia

A R T I C L E I N F O

Keywords:
Deep learning
Greenlip Abalone
Augmented reality
Object detection
Aquaculture

A B S T R A C T

Abalone are becoming increasingly popular for human consumption. Whilst their popularity has risen, measuring the number and size distribution of Abalone at various stages of growth in existing farms remains a significant challenge. Current Abalone stock management techniques rely on manual inspection which is time consuming, causes stress to the animal, and results in mediocre data quality. To rectify this, we propose a novel mobile-based tool which combines object detection and augmented reality for the real-time counting and measuring of Abalone, that is both network and location independent. We applied our portable handset tool to both measure and count Abalone at various growth stages, and performed extended measuring evaluation to assess the robustness of our proposed approach. Our experimental results revealed that the proposed tool greatly outperforms traditional approaches and was able to successfully count up to 15 Abalone at various life stages with above 95% accuracy, as well as significantly decrease the time taken to measure Abalone while still maintaining an accuracy within a maximum error range of 2.5% of the Abalone's actual size.

1. Introduction

Aquaculture is one of the world's most rapidly growing food industries, increasing at a rate of approximately 8% annually, due to escalating human consumption and declines in wild commercial yield (Garlock et al., 2020). Of all aquacultural creatures, one of the most highly sought after is the Abalone. Being one of the rarest and most expensive of any seafood, Abalone are a type of single-shelled marine snail that are found in very few parts of the world. As such, the commercial Abalone aquaculture sectors in Australia, particularly that of the Greenlip Abalone (*Haliotis Laevigata*) have gained significant popularity in particular because of their high demand and export potential (Hart et al., 2013). Consequently, this uptake in interest has resulted in a proportional increase in price, and thus, abalone production and growth mechanisms have become areas of high interest (Venter et al., 2016). Despite the high potential and increasing size of the industry, the methods of abalone farming have remained mostly unchanged, and still require comprehensive support from trained technical staff from the abalone's nursery stages through to when they are market-ready. Typically this process involves approximately 4 years of close observation and manual labour which attests to the need of optimising production management for consistently high yield rates (Fallu, 1991; Troell et al., 2006).

Traditionally, farmers have used a combination of manual inspection and approximation techniques to track stock and ensure their populations remain healthy (Hong et al., 2014; França Albuquerque et al., 2019). However, as the scale of production increases based on market demand, the inaccuracies inherent within manual stock tracking techniques become magnified. For abalone in the nursery stages, farmers are required to retrieve a sample of juvenile abalone growth plates, perform a manual count, and then use averaging techniques to approximate their population (Heasman and Savva, 2007). Similar counting methods are applied for the later stages of growth, but in addition, the abalone are weighed, and their lengths are individually hand-measured which results in significant disturbance and stress of the animals when they are removed from their substrate (Reaburn and Edwards, 2003). There are clear issues with such techniques, as these practices are highly time consuming, inaccurate, susceptible to human error and there is no definite assurance in the data quality, which further demonstrates the need for a superior system.

Thus, automated abalone measuring and counting techniques have become areas of high potential. Techniques which have been developed for other aquatic creatures possess inherent problems and complications when applied to an abalone context. This is largely due to the operating environment of onshore and offshore abalone farms, which can be potentially located in rural areas with unpredictable environmental conditions, weather, and network coverage (Heasman and Savva, 2007). Previous approaches seen in related works have been observed to have some common factors which limit their usability

* Corresponding author.
E-mail addresses: Thomas.Napier@jcu.edu.au (T. Napier), Ickjai.Lee@jcu.edu.au (I. Lee).

https://doi.org/10.1016/j.compag.2023.107744
Received 25 May 2022; Received in revised form 5 February 2023; Accepted 26 February 2023
Available online 2 March 2023
0168-1699/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

under the same uncertain conditions such as location dependence, expensive setup and maintenance costs, network reliance and non-real time capability (Koprowski et al., 2013; Pinkiewicz et al., 2011; Lee et al., 2015). A simple solution is to use a tool which most farmers will already possess – a smartphone – which enables for the capturing of high quality images, without the need for any specific equipment training. This presents a unique opportunity for an automated mobile-based solution that is capable of counting and measuring abalone in real-time, which allows for freedom in location without the reliance of network or cloud connectivity. With these features, farmers can have more regularly available data and accurate population counts, enabling them to enhance their current monitoring and stocktaking practices, and enable for rapid decision making without the overhead and additional running expenses from complex server–client based configurations.

Recently, considerable strides have already been made in the application of vision-based techniques to other aquaculture areas such as fish, however, little work has been conducted in the domain of automating abalone farming (Cook, 2016). While research in automating fish counting and measuring have used several techniques ranging from image-based machine vision systems, hydroacoustical estimations, and video-based analysis techniques, currently very few of the existing approaches have been developed, or are applicable specifically to abalone. There have been some sensor-based approaches such as in a study conducted in 2015 for the mechanical grading and weight estimation of market-ready samples (Lee et al., 2015). However, the system is incapable of handling juvenile abalone, and has several drawbacks including a complex setup, and mandatory network connectivity, among others.

The primary aim of this study will be to integrate two distinct mobile technologies, being object detection and Augmented Reality (AR), in order to develop a fully mobile-based, real-time, network independent, non-invasive abalone measuring and counting tool to assist farmers' decision making. This will be achieved by first determining the most suitable vision-based techniques for counting abalone on smartphones in real time without the need for a network connection. The chosen techniques will need to be able to detect abalone throughout its various life stages; be efficient and portable such that it can be used in real time on a mobile device and be fully self-contained to where it requires no communication with any external connections so that it can be used under any conditions. It is further aimed that the most suitable AR method for measuring abalone will be determined based on a similar set of metrics, with the added goal of being performant in a variety of environmental and lightning conditions. The use of AR technology in real-time presents a great practical benefit of reducing labour intensity whilst also opening up the data collection process to a broader user group than just experienced operators. Moreover, such technology also ensures that the animals are minimally disturbed which is critically important in maintaining healthy stock.

In summary, the main contributions of this paper are:

- to identify key unique characteristics in Abalone counting and measuring;
- to propose a mobile-based AR structure to assist farmers in measuring Abalone with limited internet connections;
- to investigate a mobile-based object detection structure for assisting Abalone counting in remote farms with no internet connections;
- to integrate mobile-based AR and object detection into a self-contained tool for improving the Abalone counting and measuring processes in real-time;
- to provide practical experimental results in various settings using live Abalone, demonstrating the efficiency and effectiveness of our proposed framework;
- to evaluate any gaps that need to be addressed.

2. Related works

2.1. Counting and measuring in abalone farming

As it has been briefly discussed already, the current state of abalone farming, and by large, the aquacultural industry have been limited to manual detection, counting and grading systems. For example, in most fisheries, where these manual forms of stock taking are highly prevalent, such tasks can take multiple people several hours to individually screen tens of thousands of fishes, with no definite assurance that their numbers are accurate or true (França Albuquerque et al., 2019). In the case of abalone, manual inspection for counting and measuring is typically performed by hand by trained staff at the juvenile and mature stages. Here, often any mode of disturbance decreases growth rates or can even cause the abalone to die if enough damage to the soft underlying body is done (Reaburn and Edwards, 2003).

At the nursery stages, juvenile abalone are highly fragile and susceptible to death if handled incorrectly. For farmers to achieve high nursery plate yield rates, trained staff must follow a strict procedure for any out-of-water counting, measuring or health inspections. When abalone are at this very delicate stage, they cannot be exposed to air for more than 1 min at any one time, and only then, when air temperatures are between 10–25 °C (Heasman and Savva, 2007). Such conditions require workers to be careful and precise, which can significantly increase the time taken to count and measure abalone stock. As abalone grow only, on average, two to three centimetres annually, it is imperative for farmers to see a return on investment (Venter et al., 2016). Thus, ensuring their survival and constant growth at their mature stages is also highly important.

Currently, matured grow-out and weaner stage abalone are counted and weighed also using manual methods (Hong et al., 2014; Reaburn and Edwards, 2003; França Albuquerque et al., 2019). At these stages, it is less likely that they are as delicate as they were as juveniles, however, monitoring stock growth statistics are still necessary for identifying changes in population size. In addition, close monitoring of the population can also assist farms in determining optimal selling points and can aid in the performance assessment of different feeding strategies. However, due to the current methods of counting and measuring stock, there are still issues with accuracy and scalability which only become more apparent as operation sizes expand, and stock volume increases.

2.2. Deep-learning for abalone detection and counting

Previously proposed methods of counting abalone typically involved data capturing either manually or using commercially available optical detection systems and using approximation techniques to estimate farm populations. In the literature, we observed four main groups of approaches for counting and measuring aquatic products observed throughout the literature are sensor-based methods, acoustic based methods, video analysis methods and image processing methods (Li et al., 2020). These techniques have all demonstrated high potential across a multitude of tasks for other aquatic products such as fish, oysters, scallop, and prawns but there is a notable lack in the domain of abalone aquaculture.

Sensor-based methods are approaches which typically employ the use of electronic and mechanical devices including optical sensors, infra-red and electrode resistivity for counting aquatic life. For abalone specifically, there was study conducted in 2015 for the mechanical grading and weight estimation of market-ready samples (Lee et al., 2015). Able to achieve a weight estimation accuracy to within 8 g, the proposed capturing system uses a complex, rail-guided, LED infrared backlit data collection method, where images are taken and then transmitted over an Ethernet network to a nearby storage device for analysis. Such a system has several drawbacks, the first of which is that it is limited to adult-sized abalone only, as juveniles would be too fragile for such a system. Secondly, an implementation like this, and others

(Baumgartner et al., 2012), has multiple points of failure due to its complex setup and reliance on network connectivity to the attached computer.

Acoustic-based methods are approaches that use sound waves to overcome issues with underwater image capturing such as in murky conditions and low light levels. For approaches which use sonar there are studies where the researchers have used a multi-beam mobile sailing robot to automatically estimate fish density in open waters (Koprowski et al., 2013). While this method was efficient at counting and measuring aquatic life in areas of uncertain water conditions with a measurement error of less than 8%, the solution is still relatively expensive to produce at around 2500 Euros per unit, and the images captured are not available in real-time.

Similar techniques have been used to analyse fish behaviour, and growth rate within fish tanks such as in a study conducted in 2005 (Conti et al., 2006). Using remotely-activatable acoustical emitters and receivers to monitor fish within a cylindrical tank, researchers were able to harmlessly record the reverberation time series in the tank with swimming fish to estimate their density and growth rates, without the need for human intervention. However, the system was limited in that measurements are only taken every 10 min, thus results were not in real-time and sudden changes in population or behaviour would not be discovered until results were collated and analysed.

For video-based analysis methods, typically multiple cameras, and other sensors are employed for data capturing, and are often restricted to controlled environments. In one study, video recording sequences are captured from aquaculture tanks using multi-camera systems for crab and fish behaviour analysis (Pinkiewicz et al., 2011). These methods generally do not include the integration of AR, unlike many similar methods seen outside of aquaculture. Furthermore, these types of systems are generally not real-time, and videos are recorded for post-analysis, which also requires the retention of large volumes of footage, and thus, server hosting costs and storage considerations become additional limiting factors to these types of approaches.

Other methods which incorporate the use of video-analysis suffer from network-reliance as footage is typically required to be transferred from the capture system to the web-hosted tool for processing, resulting in latency and susceptibility to poor signal conditions in remote areas (Cao et al., 2020; Spampinato. et al., 2008). In addition, these systems also suffer from complex, often immovable configurations. Such approaches would not be as suitable for abalone, as the smallest changes in population or growth patterns could be potentially costly for farmers due to the slow growing rates.

Finally, for embedded image processing, typically these methods are non-invasive, location-independent, and highly efficient because they have power and computation restrictions. For one such method, applied to the task of scallop detection, researchers used image processing methods embedded in an Autonomous Underwater Vehicle (AUV) system (Rasmussen et al., 2017). This method has the reoccurring issue, that in order to offset processing due to restrictions in hardware, it requires connectivity to a network to transmit data to the system for analysis, meaning that it is also not in real-time.

Further approaches such as one proposed in 2018, focus on using low-cost image processing methods embedded on a Raspberry Pi 2 for the counting of fish with an average accuracy equal to 96.64% (Hernández-Ontiveros et al., 2018). Their approach is both real-time, low cost and easily portable. However, in a real-life environment, the method requires a specialised setup where fish are fed through a two-tank interconnected system which forces fish to travel from one tank, past the counting system into a collection tank. While this method may be applicable to fish, abalone typically live a sedentary lifestyle and remain in the same general area for all of its life, which would likely be problematic for such a system.

2.3. Augmented reality for real-time measuring

With the ubiquity of mobile devices in modern society, AR has seen an increase in usage in mobile applications. Fundamentally, it is a virtual experience whereby real-time, interactable objects such as three-dimensional (3D) models, text and images are superimposed onto the physical world and are made viewable through a digital medium, such as a smartphone or tablet. Such technology can be useful for projecting virtual content and manipulating it within a 3D environment, in real-time. Objects that in an AR scene are strictly digital, and as such, the boundaries of the physical reality can be extended through technology in such a way that it perceptually enriches a user's actual environment. We need only look back to 2016 with the recent example of Pokémon GO app produced by Niantic, which quickly became the top mobile game in the US at the time to see the potential that AR-enabled applications can achieve (Eklind and Stark, 2018). However the practical benefits of AR enables users to perform a variety of tasks by providing real-time, on-the-spot, interactable augmented information. As such, the technology itself has already been implemented in a variety of industries (Orciuoli et al., 2020; Eklind and Stark, 2018; Dash et al., 2018; Lovreglio and Kinateder, 2020).

One successful example of the technology being used outside of the aquaculture domain, is in a 2018 study that focuses on applying AR to create visual aids through display for early childhood training. The study employed a convolutional neural network image classification system in combination with AR to detect markers (letters of the alphabet) and motion track relevant objects to the markers (Dash et al., 2018). This approach has a strong focus on using marker-based AR, with an integration of deep learning image classification network. This highlights how the integration of these technologies can combine to optimise or improve a task, in this case, making learning more engaging for children. However, the design is limited to only very visually distinct alphabetical letters and it requires camera calibration for motion tracking which constrains it, making it unsuitable for mobile devices.

For approaches which employ the use a mobile device, there have been medical-based tools that use image classification and AR to detect and measure medical bedsores such as a study proposed in 2020 (Orciuoli et al., 2020). To measure these physical ailments, researchers used a web hosted server which classifies the bedsore image using PyTorch,[1] in conjunction with OpenCV[2] to measure bedsore (marked out beforehand) which is then shown graphically with EasyAR[3] once processed through the measurement tool (Orciuoli et al., 2020). This highlights that the combination of these technologies can be useful, for the detection and measurement of objects within an image. However, the authors make use of web services to achieve this which creates unnecessary overhead. In addition, the tool is overly dependent on user interaction for the initial measurement line drawing, and the solution is not in real-time.

Another developing area is the application of AR technology in agriculture and aquaculture. AR in these fields has been used to overcome inefficient data collection processes and network connectivity issues in rural farmland areas, and provide great practical benefit. Examples include where AR has been used to handle the delayed in-situ water quality data collection (Xi et al., 2019), and also it has been used for improved data capture in prawn farm management (Rahman et al., 2021). In addition, wearable smartglasses for AR (Caria et al., 2020; Phupattanasilp and Tong, 2019) have been applied for precision farming, and AR has been used for monitoring multirobot system in agricultural field operation (Huuskonen and Oksanen, 2019). Please refer to Anastasiou et al. (2023) and de Oliveira and Corrêa (2020)

[1] https://pytorch.org/.

[2] https://opencv.org/.

[3] https://www.easyar.com/.

Table 1
Literature review summary.

	Sensor-based	Acoustic-based	Video analysis	Embedded apps	This research
Mobile	No	No	No	Yes	Yes
Object measurement	Yes	No	Yes	Sometimes	Yes
Object counting	Yes	Yes	Yes	Yes	Yes
Real-time	Sometimes	No	No	No	Yes
Network independent	No	No	No	Sometimes	Yes
Location independent	Sometimes	Sometimes	No	Yes	Yes

for a comprehensive coverage of applications of AR in agriculture and aquaculture.

Studies into the applications of AR in aquaculture are extremely underdeveloped, despite the potential impact that it could have on increasing efficiency in farm management. AR has the potential to solve a variety of problems experienced by farmers. As identified in a recent survey, several key issues were identified within an aquaculture farm management. The most important issue is that decision making in feeding and growth strategies are made difficult, because existing systems lack the ability to provide timely results. In addition, the lack of efficient integration with existing systems and procedures, as well as difficulties with training new staff with the complexities of farm management, mean that existing solutions often require extensive instruction from skilled farmers (Xi et al., 2018). An AR-based approach is capable of offsetting most of these issues by providing real-time, localised information for rapid decision making, in an easily learnable manner.

While research in this area is limited, a recent study confirmed through the use of a machine learning enabled AR workspace system, that these technologies positively facilitate real-time prawn pond monitoring (Rahman et al., 2021). However, such a system was built around wearable technology, such as the Google Glass and Microsoft HoloLens, which allows for farmers to operate in a hands-free way, but the upfront costs are considerably higher than a mobile-based implementation. Despite this however, this study illustrates how the combination of AR, and machine learning can be used to improve the speed of decision making, made possible through the availability of real-time data.

2.4. Summary of literature gap

The deficiencies in the literature are mainly that the areas of Deep Learning (DL) and AR have not been applied in the real-time measuring and counting of abalone, leaving a considerable gap. The most prevalent limitations of related approaches are that they are network-reliant and/or not real-time, requiring constant connectivity for the system to work. Such a solution not only requires the solution to be pre-arranged to function like many other approaches (Parr et al., 1995; White et al., 2006; Baumgartner et al., 2012; Koprowski et al., 2013), but it also creates unnecessary overhead from back-and-forth communication between the system and the client.

As summarised from reviewing the advantages and disadvantages of the approaches as per Table 1, it can be deduced from the identified thematic groups that no existing aquacultural approaches are entirely suitable for the given task. In addition, few of the review approaches are transferrable for counting and measuring abalone in real-time, where the approach uses a mobile device, and is both network and location independent.

3. Overview of proposed method

The system designed for counting and measuring abalone in real-time, without a network connection has been defined as beginning with the on-device camera, as image capturing is at the core of the instrument. With image capturing handled, it is up to the decision of the user and the nature of the task they are trying to complete. In the context of this project, the decision must occur at the start of usage, as simultaneous operation was outside of the project time and resource scope. As outlined in Fig. 1, the proposed framework consists of two main conceptual branches contained within a single Android application, with no external connections required during its execution or operation.

For counting, images are taken from the camera, and morphed into a usable 416 × 416 pixel size in preparation for input into the embedded TensorFlow Lite[4] model pre-trained on custom datasets using the process of transfer learning. Two TensorFlow Lite models were trained on two sets of training data, one for nursery abalone and the other for weaner/grow-out stage abalone. The output of these models – once trained – were two self-contained .tflite files each with an associated class label text file, both of which were transferred to the local storage on the mobile device and embedded such that it could be accessed and used during the execution of the mobile application. Images captured by the system would then be passed through these models and the output results, including the prediction scores, object locational coordinates within the image and number of objects would be returned. From here, the bounding box coordinates are stretched to match the original image size and are drawn on-screen to indicate the position of the object along with the accuracy from prediction scores, class of object, and overall number of objects within the image.

In the second component, for the task of measuring objects, Google's ARCore[5] was used. For this task, several layers of user interaction are used in conjunction with sensory data made available from on-device components including the camera, Inertial Measurement Unit (IMU), accelerometer and gyroscope. Initially a plane must be established within the operating user's environment which is achieved by using the sensory data available to locate and track feature points at areas of visual distinction. Using clusters of feature points, ARCore can then overlay and establish the boundaries of a two-dimensional (2D) plane which enables for the accurate positioning of anchor points and other digital objects relative to real surfaces, but within the augmented environment. With a plane established, the object to be measured must be situated close to it, and using the on-screen positional markers, users simply need to align the dots with either side of the object that they are trying to measure.

Once aligned, a user can request to measure the distance using a single button click. ARCore uses the pixel axis coordinates from the 2D screen and casts a straight line – or ray – from both on-screen marker points and projects them into the virtual world space within the camera's view. If either of these rays intersect the established plane, an anchor point is created at the point of intersection. From this, we can attach viewable 3D models and render them to the screen as visual indicators of where the anchor points were placed. It is then a matter of calculating the distance between the two established anchor points using the Euclidean Distance formula, and then relaying that to the user.

Both component functionalities allow for the user of the application to rapidly measure singular abalone or larger objects at once or count multiple abalone in real time. Such a design means that users are not forced to use additional devices or change the main screen of the

[4] https://www.tensorflow.org/lite.
[5] https://developers.google.com/ar.

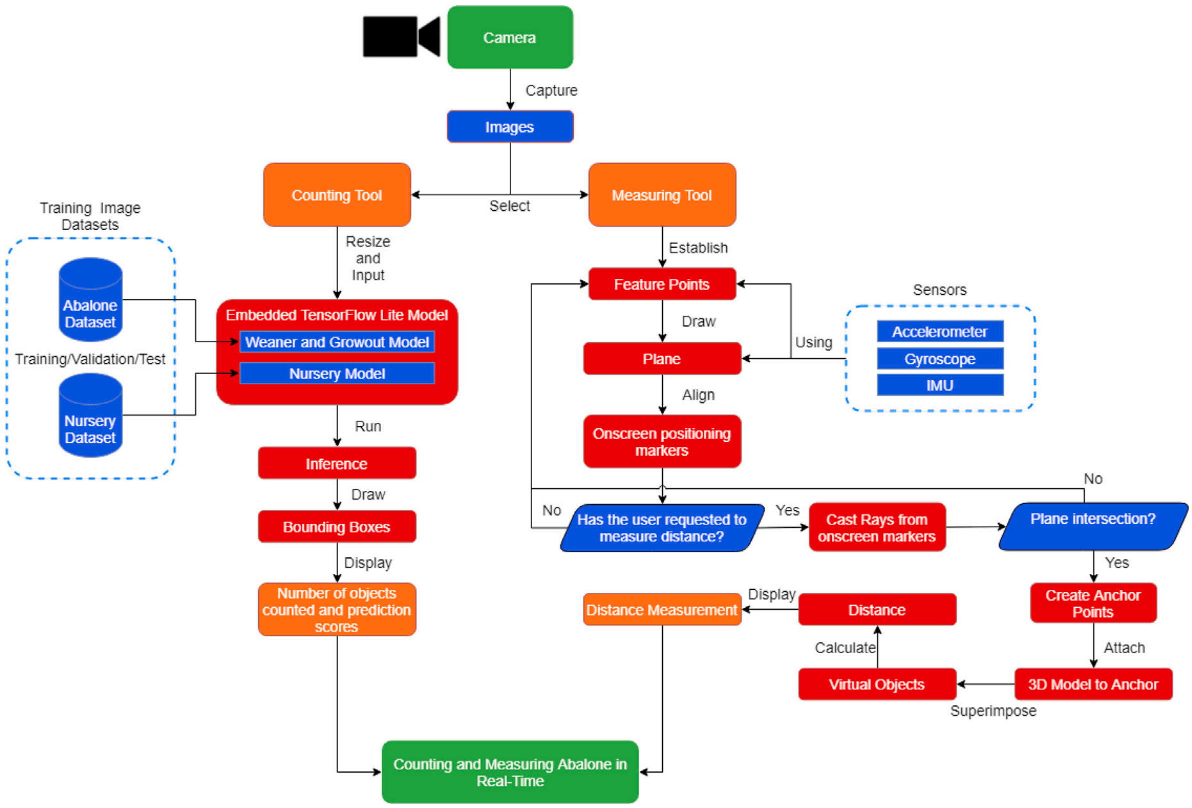


Fig. 1. Conceptual framework for a mobile-based abalone counting and measuring tool.

application as both tasks can be easily swapped to. The ease of resetting the application, and additional advantage of real-time enabled features supports the operator in making one or more measurements without the need to take extensive time to reset the observation. In both cases, the focal objects do not need to be manually handled except to lie the creature flat, and it allows operators to easily collect objective data for growth inspections and population monitoring.

We designed this tool with the goal of improving the data collection process, and thus, we did not extend the application into an all-in-one data capture, management, and storage solution. Images and all associated metadata, captured by the system, are saved on-device which can be subsequently transferred to a central system for historical record-keeping.

4. A lightweight object detection method for abalone counting

Object recognition fundamentally is a visual detection problem in computer vision and image processing. The term object recognition is a generalisation that is used to describe the combination of both image classification and object localisation (Xiao et al., 2020). Image classification relates to a process of prediction, whereby a deep neural network is trained to recognise the characteristics of an object within a labelled example image using repeated exposure, so that it can be applied to new, unseen examples. Moreover, object localisation is used to locate the presence of objects within an image and to indicate their position using bounding boxes. This technique allows for low-cost cameras, such as those found on smartphones, security cameras, etc., to be used to extract and identify one or more objects within an image or video and draw a bounding box around their extent. With the ubiquity of smartphones, locally embedding a lightweight object detection method can enable for the detection and counting of abalone to occur anywhere without the need for a network.

4.1. Datasets and detection network training

A sample collection of around $n = 105$ images of grow-out, weaner and $n = 130$ nursery abalone obtained by during a James Cook University (JCU) research field trip in December 2020 from a farm in Victoria, Australia. Images would consist of either nursery growth plates with between 0 to upwards of 20 juvenile abalone per image, or for grow out and weaner stages, which had from to 10 to 50 per plate. With these images collected, a series of pre-processing steps as outlined in Fig. 2, were used to ensure that the data was of quality and acceptable for model training.

First, all images were resized to 416 × 416 pixels as this is the required input for the YOLOv4 network architecture (Bochkovskiy et al., 2020). Once resized, and due to the limited amount of quality images available for model training, some data augmentation steps were applied such as mirroring, rotation, brightness, and noise changes. The aim of this technique was to generate more variety and conditions from a single image, to simulate the image capturing under different circumstances. This was performed to simultaneously test the robustness of the implementation process, to increase the amount of training data, and to prevent issues with overfitting, where the model learns too much detail and noise in the dataset resulting in reduced performance on the test dataset, or underfitting, where the model fails to capture sufficient detail.

Each image in both datasets were then manually annotated using the graphical interface, LabelImg,[6] which produced a series of images and Extensible Markup Language (XML) annotation files. The data was then subsequently split into training, test, and validation subsets of roughly a 70/15/15% split, respectively. This is done so that once the

[6] https://github.com/tzutalin/labelImg.

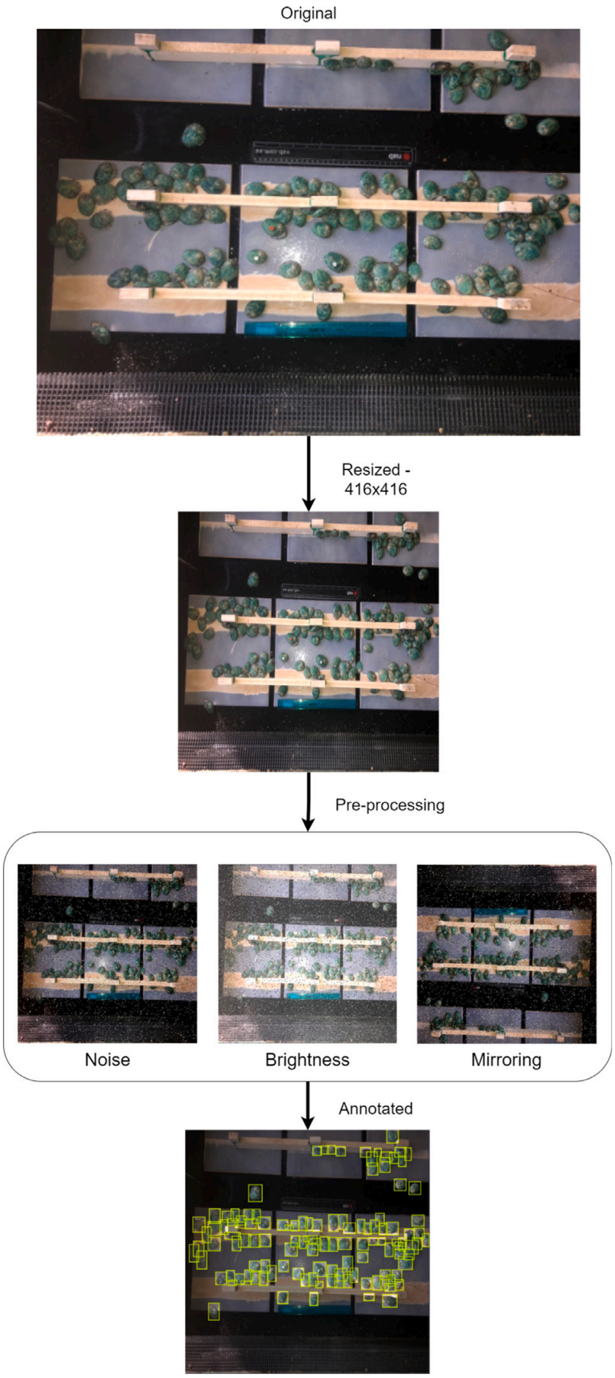


Fig. 2. Data pre-processing and annotation workflow for training a deep-learning network model.

model has been re-trained to classify objects based on a specific set of data, the model can be evaluated on new data to assess its accuracy.

Moreover, for the nursery stage abalone dataset, since we were unable to obtain specimens for the experimentation phase, a decision was made to introduce a subset of images of groups of small rounded grey pebbles ranging in lengths from 10–20 mm to simulate juvenile abalone shells. A small dataset of these pebbles, annotated and labelled as nursery abalone was included in the final dataset for training. Such an approach would give us the flexibility to test the counting implementation on weaner/grow-out, and nursery stage abalone without hugely impacting the overall dataset accuracy.

Table 2
Comparison of training performance metrics for both the nursery and mature abalone datasets.

Metric	Nursery dataset	Weaner/Growout dataset
Average Precision (AP)	85.12	82.46
Precision	0.93	0.9
Recall	0.87	0.84
F1-Score	0.9	0.87
Mean intersection over union	65.49	84.17
Mean AP (mAP)	0.85	0.83

Model training took place using the process of transfer learning on a base model checkpoint pre-trained on the Common Objects in Context (COCO) dataset[7] - a collection of 330,000 images with over 80 object classes, to enable for the transference of knowledge to be applied to detecting abalone without the need to fully train a model from scratch. YOLOv4 was specifically chosen as our network architecture because it has the nearest real-time speeds, is lightweight, and still maintains a high level of accuracy compared to others which makes it ideal for mobile-based abalone detection.

As part of the training process, the model was fine-tuned part-way through to ensure that feature extraction was performant and that its associations are adjusted according to the new dataset. The result of this was that the new training rate was significantly lower. Each model, for both datasets, was subsequently trained for 2000 epochs on custom data using YOLOv4 within a Darknet[8] neural network framework. The resulting training performance metrics based on validation data can be seen in Table 2.

4.2. Model integration and validation

Following model training, a process of conversion from a standard model, into a lightweight, mobile compatible format was conducted using conversion libraries. The library chosen for this was TensorFlow Lite,[9] and this was because it possessed the qualities which supported our project, including it being quite a mature library with extensive documentation which assisted development. In addition, the library also possessed integration capabilities with other AR frameworks already such as ARCore, it was free to use, and it supported Python and Java, two well-known languages to the primary researcher, making it an ideal candidate.

Using TensorFlow Lite, a simple labels text file is required in the same order that the model was trained in so that the objects can be correctly classified. The two residual files after the retraining process were then embedded in the project structure of the mobile app and the pathing variables were updated to ensure the new files associations were correctly connected. This enables for images to be passed through the model in real-time, and using the output given from the model, bounding boxes could be drawn with the corresponding confidence scores and class label names.

To validate our models, a series of metrics will provide the necessary objectivity to assess whether it is accurate. Firstly, the count error will enable us to assess how accurate the implementation is versus the existing manual techniques. Counting accuracy is calculated by recording the average number of objects counted and the ground truth amount across a series of tests and cross referencing the results. By taking the ratio of correctly and incorrectly predicted instances to the total instances can be used to further analyse the reliability and robustness of the model.

$$\text{Counting Accuracy} = \left(\frac{\text{No. of Correctly Counted Objects}}{\text{Total No. of Counted Objects}}\right) * 100.$$

7 https://cocodataset.org/.
8 https://pjreddie.com/darknet/.
9 https://www.tensorflow.org/lite.

There are also classification-based metrics based upon the confusion matrix by which we can evaluate the accuracy of the model. Firstly, during training the model can be tested against the 15% validation data subset for performance analysis, the results of which can be found in Table 2. Here, metrics such as recall, F1-score, and Mean Intersection over Union (mIoU) are calculated. Recall or sensitivity is used to calculate the true predictions from all correctly predicted data, and it involves taking the number of correctly counted objects, divided by the total number of actual, relevant objects, where:

$$\text{Recall} = \frac{\text{No. of Correctly Counted Objects}}{\text{Number of Actual Objects}}.$$

F1-score is the weighted average of precision and recall. The F1 score takes the precision and recall into account which is typically a superior metric for evaluating a model than accuracy as long as false positives and false negatives have a similar cost or weighting. The F1-score is calculated using the equation:

$$\text{F1-score} = 2 * (\frac{\text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}).$$

Furthermore, the commonly adopted IoU metric is used specifically for visual detection tasks because it computes the difference between ground truth annotations and predicted bounding boxes produced by the model. The output from an object detection model is the prediction confidence score, and bounding box coordinates for each object within the image. Based on the scores of each box, unnecessary boxes are removed based on an established threshold value, whereby any prediction scores that fall below the threshold are not used:

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}.$$

5. A mobile AR method for abalone measuring

There were a multitude of steps used to implement an ARCore enabled Android application, which can be encapsulated into a series of key steps, starting with the AR plane establishment, through to the enabling of user interaction for anchor point placement. The steps outlined below in Fig. 3, displays the series of steps we used for the implementation of a network independent, marker-based ARCore system that is capable of enabling users to measure Abalone in real-time. Prior to using the application however, it is vital to check the user's device for AR 'required' availability upon start-up, as this determines whether the user's device and Android Software Development Kit (SDK) satisfy the minimum version of 24 (Android 7.0 (Nougat)) or later. If it meets the requirements, the user is then prompted to allow the app to access the device's camera before the ARCore session is established.

5.1. Development framework

Once camera permission is provided from the user after they launch the app, the initialisation of the virtual scene is handled through an inflated SceneForm fragment, whereby any 3D assets are rendered within the camera's view. At this stage, the fragment also handles camera initialisation and permission handling, and once successful, a shared camera feed from the on-board camera interlinked with ARCore. After a camera uplink has been established, the next stage is to find a plane within the environment using feature points. A user will do this by moving their device throughout their environment, on any flat surfaces. In most cases, the more varied the surface is, the easier it is for ARCore to establish a plane. This is due to its innate environmental understanding feature-set, whereby contrasts in texture, colour and shape enable ARCore to find feature points easier.

Once a plane has been found within the environment, the user can then request to measure the object, and through the press of a button, ARCore will project a ray from that coordinate point on the two-dimensional screen into the three-dimensional virtual environment. If it is found that it intersects the plane, or anchor, data about the objects found is made available, allowing for user interaction to be managed.

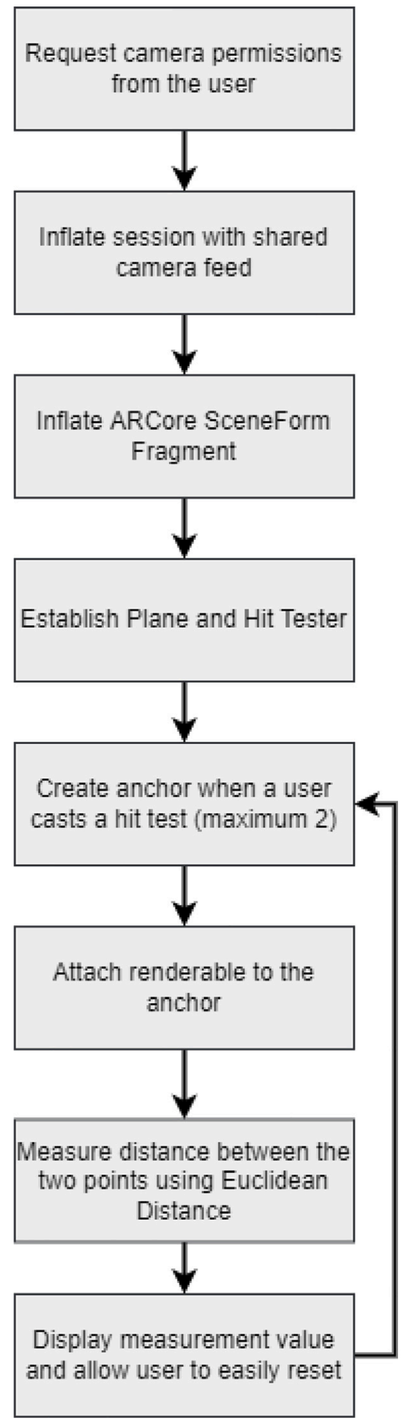


Fig. 3. ARCore implementation workflow for measuring the distance between two points.

An anchor is created at the point of intersection on an established plane, after the initial input from the user's mobile device's screen. Anchors define a three-dimensional pose in world space, that is then pinned in the context of the trackable, in this case, the plane. This means that the anchor point itself can be properly tracked during motion, irrespective of where the object is placed in world space.

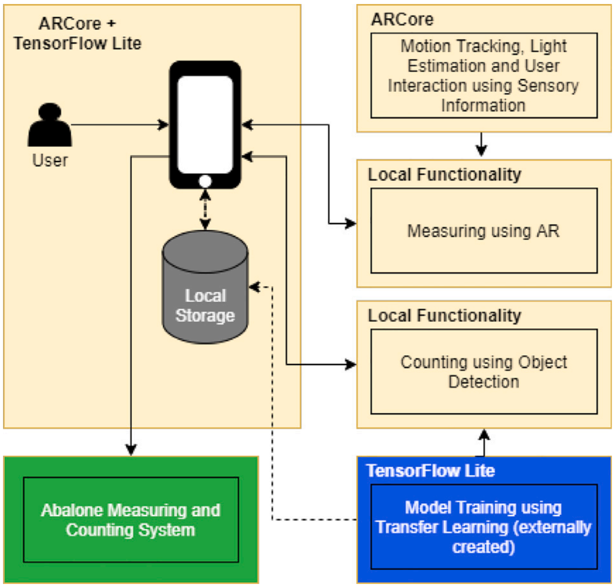
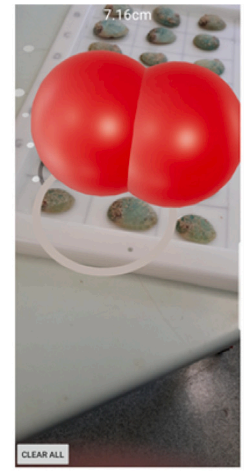


Fig. 4. Design architecture of proposed method.

(a) Mobile device mounted in stationary gimble arm positioned for Abalone measuring experiments.



(b) Mobile (Android) application screen as seen when measuring Abalone after markers have been placed.

Fig. 5. Abalone measuring experimental setup using a mobile device placed in a 32 cm tall camera gimbal angled downwards for testing AR method A.

5.2. Distance

Once an anchor is established, a renderable marker, such as a three-dimensional (3D) sphere, is then programmatically attached to the anchor to visually denote the anchor placement. From there, ARCore's SceneForm API handles interaction-based events that rely on further hit tests when using an object type of TransformableNode. These nodes allow the world position of anchors to be moved within the virtual environment if it remains on an established plane. Once two anchor points in the environment have been established, the distance between the spheres can be measured using the Euclidean distance formula, where

$$distance = \sqrt{[(x1-x2)^2+(y1-y2)^2+(z1-z2)^2]}.$$

Here, distance will derive a measurement value in centimetres which equates to the length of the true straight line between the world position coordinates of the first anchor point x, y and z, and $x1$, $y1$, and $z1$ respectively and similarly for the second point. The resulting number is then returned and can be parsed to any layout object which uses string text. It is important to note that ARCore is able to calculate this distance without a known reference object that must always be in the scene. It does this by building its own environmental understanding which allows the smartphone to detect the size and location of surfaces in reference to the AR markers in the virtual world space.

6. Application and experiments

6.1. Prototype system development

The framework model described previously has been implemented as an Android AR and object detection-based application, the architecture of which is visualised in Fig. 4. The figure shows that ARCore and TensorFlow Lite are the chosen technologies that will be combined and constructed using local storage, sensory information and externally created TensorFlow Lite models re-trained using transfer learning. TensorFlow Lite is shown to provide the local counting functionality using object detection, and likewise, ARCore is shown to facilitate measuring using AR. The layout of the diagram shows how the user interacts with the system, and how each component with the architecture relates and ultimately how abalone measuring, and counting are coupled together.

6.2. Experiments with AR measuring tool

For this research, an experimental design will be used where the independent variable will be the method of measuring and counting abalone — between ground truth, manual estimation to simulate performance from traditional techniques used by farmers and comparing it to the automated system. For dependent variables, we will be evaluating the change in the speed and accuracy of the measuring process when using both a manual approach compared against the proposed system, controlling for the species and life stage of abalone being evaluated — the Greenlip Abalone.

We tested the prototype system on 24 live mature-sized abalone to evaluate both the counting and measuring method and validate the results. We followed a general process for testing and collecting data using both methods, in a way that was parallel to each other. Each method started by collecting data on the speed and accuracy of manual methods, before we used estimative counting and measuring techniques supported by abalone farming literature, before testing our methods.

As observed in Fig. 5, experiments with the 24 live Greenlip Abalone were conducted within a typical laboratory setup with overhead lighting. The 24 abalone were placed on a gridded plastic tray and were roughly spaced apart. The measuring device was a Samsung Galaxy S8 Plus which is capable of running AR-enabled applications. The phone was seated in a 32 cm long camera gimbal with approximately a 45 degree tilt. The camera gimbal was attached at a 90 degree angle to a tripod set at a height of 40 cm tall. The tray of abalone was then placed within the camera's view. The measuring tool was selected from with the android application and the tripod was moved for approximately 11 s before a plane was established from environmental feature points.

For data collection, the first step was to measure each of the abalone individually by hand with a ruler, record their lengths for use as the ground truth values while simultaneously counting the time taken to measure all of them using a stopwatch. The resulting Table 3 was constructed from the collection of ground truth measurements which took 148 s to record.

For approximation techniques such as those used by farmers, a similar approach to those detailed in a manual for hatchery blacklip abalone production was used (Heasman and Savva, 2007). Using pseudorandom number generation, we randomly selected 20% of our abalone and used their measurements to approximate the lengths of the population. As such, 5 abalone were selected, and were hand measured for example,

Table 3
Abalone ground truth measurements for AR method A.

	Column 1 (mm)	Column 2 (mm)	Column 3 (mm)
A	62	70	49
B	72	60	54
C	64	53	62
D	66	61	60
E	70	45	60
F	63	56	51
G	32	60	50
H	32	46	52

Table 4
Abalone measurements using AR method A.

	Column 1 (mm)	Column 2 (mm)	Column 3 (mm)
A	51.7	65.2	56
B	78	68	79
C	68	51	51
D	69	72	55
E	68	49	67
F	67	60.7	56.9
G	45	71	62
H	43	56	67

Table 5
Abalone shell ground truth for AR method B.

	Column 1 (mm)	Column 2 (mm)	Column 3 (mm)
A	52	47	41
B	40	40	37
C	39	79	43
D	40	43	44
E	41	46	50
F	47	81	49
G	37	42	45
H	35	43	45

Table 6
Abalone shell measurements using AR method B.

	Column 1 (mm)	Column 2 (mm)	Column 3 (mm)
A	54.3	47.7	42
B	42	43	34.9
C	41	77.5	43.5
D	39.8	41	42.7
E	40.1	47.4	48
F	44.1	80.4	48.6
G	35	43.5	41
H	32	44.4	42

in one instance, our random numbers are 24, 11, 6, 13 and 9, where Column 1A = 62 mm, Column 2A = 70 mm, etc. Keeping the ordering of abalone intact, we then measured each abalone by hand such that we ended up with a series of lengths corresponding to their position in Table 3. In our case, the lengths recorded were 52 mm, 61 mm, 54 mm, 70 mm, and 62 mm. When averaged, it resulted with a length of 59.8 mm which was then applied to the remainder of the abalone. We repeated this process several times, and recorded the time taken with each iteration.

We then measured using the proposed instrument mounted in the gimble arm, with a plane already established. Initially, it was observed to be near-impossible to measure the abalone object using a stationary camera position, because the size of the visual marker would block the other side of the object, and thus, a second marker could not be placed. In response, the phone was removed from the gimble and moved around from side to side at an arm's length away which allowed for the second marker to be consistently placed. The results of these measurements are placed in Table 4.

6.2.1. Initial problems with AR method A

Early on in our experimentation of the AR measuring tool, we found that results were mostly underwhelming and comparable to manual methods. A large contributing factor to the loss of accuracy in our initial tests may have been due to the surface in which they were conducted on. ARCore has a difficulty placing feature points on surfaces with visual uniformity, such as the white tray used to conduct these tests on, which can lead to a decrease in average precision. However, a more likely explanation is that the methods used to place the measurement markers can be incorrect and imprecise because it is entirely dependent on where the user places the anchoring positions. Due to this reliance on user input, many factors can alter the end results such as the viewing angle of the device itself, and the precision of where the anchors are positioned relative to the edges of the object. As most users would likely be using their finger for touch interaction, it was found that reliable and precise anchor placement was difficult to achieve.

Thus, to reduce the dependence on user interaction, AR method A was modified the include two permanent on-screen markers, denoted by red squares. With these defined 2D coordinates on screen, the user simply then needs to align the dots with the edges of the object they wish to measure and tap the single button at the bottom of the screen, and two rays will be cast from either marker. Rather than relying on the user to point and tap to create an anchor inaccurately, this method of placement allows for higher precision.

6.2.2. AR method B

The implementation of method B was after the initial experiments were conducted, and the live abalone were no longer available for testing. As a substitute, several Greenlip Abalone shells were obtained and tested on in a similar fashion to the previous method with both their ground truth lengths, as observed in Table 5 and measurements using the instrument were recorded as seen in Table 6. However, the estimation techniques were not applied here and instead, a normalised accuracy was calculated between all experiments.

An interesting observation was made when conducting the experiments with method B. It was discovered that sometimes, particularly with smaller shells, the augmented dots which signify where ARCore has established a plane can overlap and interfere with the measurement process, as it can obscure an edge. However, the issue can be easily resolved either by physically moving the object or re-establishing the plane by moving the device around.

To further analyse method B, experiments were extended to other aquatic products to validate robustness in the technique. In these experiments, two reef fish were obtained — one Black Pomfret Trevally and a Red-Throat Emperor. Like previous measuring experiments, the ground truth length values were recorded by hand measuring each fish using a tape measure.

Both fish were kept frozen and were thawed until use to ensure accurate measuring. Each fish was placed onto a flat horizontal surface and measured using the same phone, with a bright overhead light. Estimation could not be conducted in the previous way because only two fish were available for the experiment. As such, an assumed estimation accuracy based on multiple abalone experiments was substituted and applied to the fish experiments.

Unlike method A's experiments for measuring abalone, no tripod was used, instead, a handheld method was employed, whereby the device was held at an arm's length away and moved from side to side for plane establishment. The use of real fish for testing meant that the markers had to be oriented to be at the furthest distant points at either side of each fish, a demonstration of which can be observed in Figs. 6 and 7.

6.3. Experiments with abalone counting tool

As seen in Fig. 8, the counting of weaner and grow out abalone was achieved using a tripod mounted-camera setup within a laboratory environment. Sitting flat and parallel to the 24 live abalone, the phone

(a) Mobile application when orienting on-screen markers for the measurement of travelly.

(b) Mobile application screen when AR markers have been placed after initial plane establishment.

Fig. 6. Measuring trevally using AR method B.

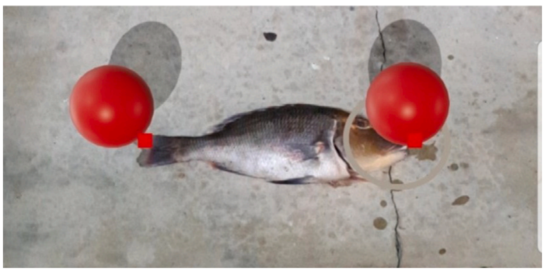
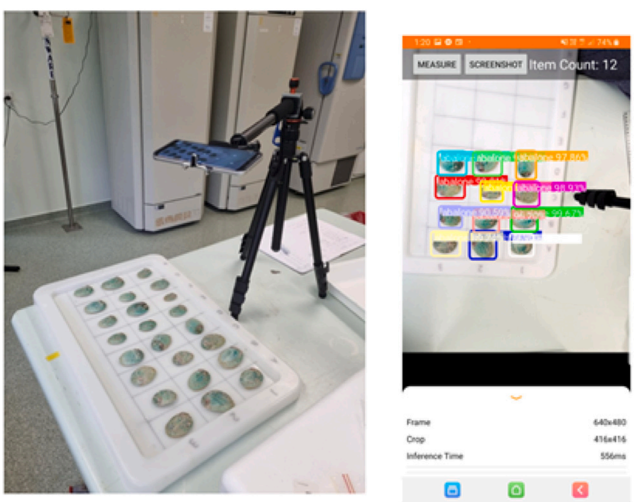


Fig. 7. Measuring red-throat emperor using AR method B.



(a) Mobile device mounted in stationary gimble arm positioned above tray for Abalone counting experiments.

(b) Mobile (Android) application screen as seen when counting Abalone.

Fig. 8. Abalone counting experimental setup using a mobile device placed in a camera-mounted tripod positioned parallel 63 cm above the testing surface.

camera was situated 63 cm above the tray, ensuring that the full tray was visible during the counting process.

Data collection was achieved by taking continuous screen captures of the counting process. Once 20 iterations of object counting were complete, abalone were slowly, individually removed from the tray, and again, 20 iterations of data were captured. In each iteration, the amount of objects correctly counted, incorrectly counted, bounding boxes and the inference time were collected. This incrementally continued up until the dataset was halved to 12 abalone. This alteration was conducted to see if changing the amount of objects being detected would affect accuracy results.

For counting juvenile abalone, we were unable to obtain any real specimens. Due to their fragile nature, young abalone must be kept in specialised conditions, and any testing we would have conducted would have likely caused damage. Thus, experimentation was conducted through the use of mimic objects, which meant that testing could be conducted outside of the laboratory. As such, 24 mimic objects were placed ontop of a glossy, reflective black surface in an attempt to replicate similar conditions to the surface of water and the nursery abalone growth plates. The counting instrument was subsequently suspended using a 32 cm gimble such that the camera was 60 cm and parallel above the round grey pebbles imitating juvenile abalone.

A process similar to one previously described for counting was conducted, whereby the correct and incorrect number of objects counted was recorded for 20 iterations per number of objects. Starting at 24 juvenile abalone, the count was slowly decreased until there were only 12 remaining.

7. Results and discussion

7.1. AR measuring performance

From our measurement experiments on both mature abalone and juveniles, we have found that through the usage of AR, our approach is successfully able to outperform manual and estimation techniques used by farmers. In addition, the use of AR technology has ensured that the Abalone are minimally disturbed and enables non-experts to conduct measurement assessments, while assisting in the reduction of image-based data collection issues such as narrow visual angles, or unwanted scaling variability. In terms of accuracy, which is calculated by comparing the total average ground truth length for abalone versus the total average measured length using our approach, method A showed that the tool initially was underperforming with a 6.8% loss in average accuracy across all measurements taken. This was later identified to be a result of over-reliance on user interaction, and thus method B was developed which saw the use of on-screen markers to eliminate these imprecisions. Method B had superior performance during the experiments, achieving an average precision score equal to 99.08%. In comparison to other aquatic animals, our method achieves similar results in terms of measuring performance while also being more practical and portable than existing approaches, such as the interconnected tank fish counting system (Hernández-Ontiveros et al., 2018). Fig. 9 highlights and compares each of the four methods of measuring.

However, in terms of speed, the initial developed method did slightly improve upon hand measured systems, reducing the time taken from 150 to 125 s, and greatly improved on estimation techniques. Yet while method A did increase efficiency, it still suffered in situations where small objects were being measured, because often the first marker placement would visually obstruct the second one from being placed. This issue was resolved with method B, however, as both AR markers would be placed concurrently, thus avoiding the situation where the first marker visually obstructed the second. This also had the secondary benefit of hastening measurements, as users would only need to simply align the markers and tap once, whereas prior, they

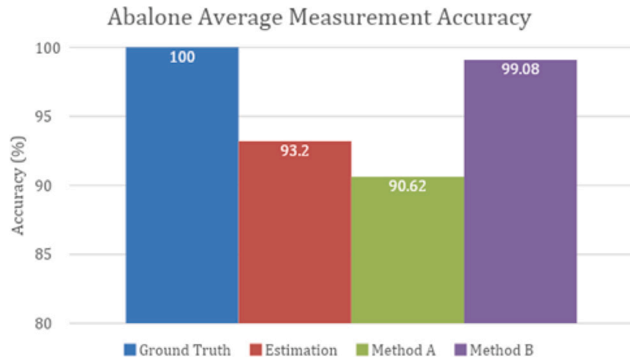


Fig. 9. Average percent accuracy over ground truth, estimation, AR method A and AR method B Abalone measurement experiments.

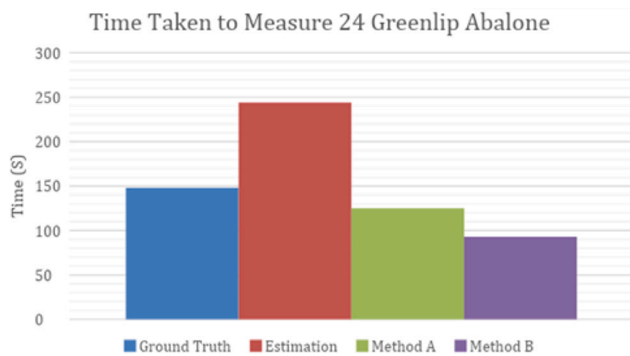


Fig. 10. Comparison of the average time taken in seconds to measure 24 greenlip abalone using manual, estimation, AR method A and AR method B.

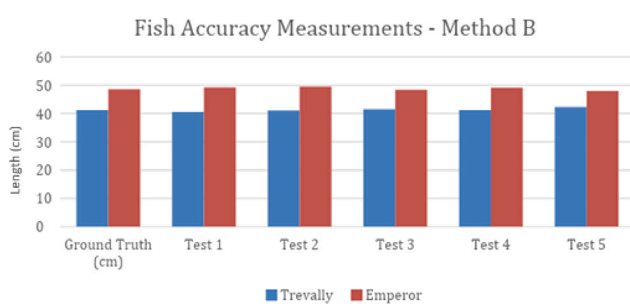


Fig. 11. Comparison of both fish species accuracy measurements across 5 tests using AR method B.

needed to align, place one marker, then physically move to place the second, which increased the time for counting. As seen in Fig. 10, measuring using method B was able to greatly outperform all other methods, decreasing the time taken to measure 24 Abalone to 93 s.

For reasons of robustness, method B was also tested on two species of fish, and as observed below in Fig. 11, the technique was able to produce comparable, high levels of accuracy despite the object being considerably larger.

Taking a closer look at the percentage error calculated from the fish measurement tests in Fig. 12, it can be observed that the technique maintains accuracy to within a maximum of 2.5% of the object's actual size.

Through the comparison of average measurement accuracies across all of the measured objects as visualised in Fig. 13, it becomes apparent that AR method A performed poorly compared to all other methods. However, with small tweaks in framework design and the decision to restrict user input, AR method B performs consistently well in all experimental situations with accuracies between 98%–100%.

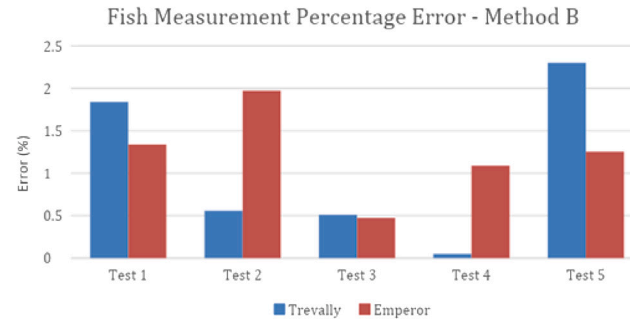


Fig. 12. Comparison of the percentage error calculated from both fish species measurement tests using AR method B.

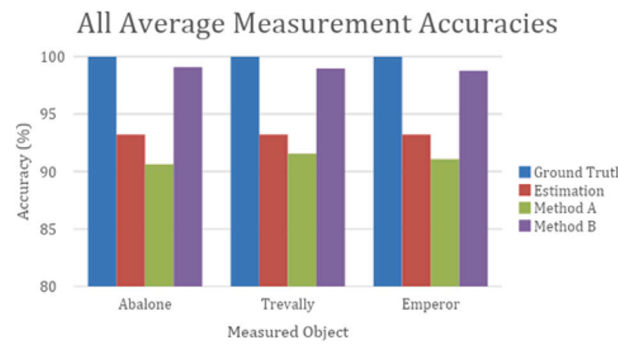


Fig. 13. Average measurement accuracies for all experimented methods.

7.2. Abalone counting performance

From our experiments counting weaner/grow-out and nursery abalone, we have confirmed that object detection is capable of counting objects in real-time, where results are location and network independent, to an extent. When counting up to 15 objects, the current framework is capable of producing results for both juvenile and adult abalone with between 95%–100% accuracy. However, as seen in Fig. 14, once the count of objects exceeds 15, the accuracy decreases linearly as the technique struggles to recognise the presence of an additional abalone. Such a result is to be reasonably expected given the computational limitations of smartphones, which become increasingly apparent when many objects are present within a single image. Object detection is a computationally expensive task that is typically reserved for high performance systems. The task of simultaneously classifying, and drawing a bounding box around each object in an image in real-time can cause significant computational demands, in an already constrained environment, which consequently leads to the decreases in counting performance (Martinez-Alpiste et al., 2022). While the body of work available for comparison is limited, our method was able to unintensively count both matured and juvenile abalone which could not have been conducted using existing mechanical systems for Abalone whilst still achieving comparable results (Lee et al., 2015). However, in comparison to some remotely activatable acoustic approaches, our proposed method requires a higher degree of human intervention for both the measuring and counting process. While this does mean that human involvement is necessary, it also means results are in real-time and can be captured in rapid sequence rather than 10-min intervals, thus sudden changes in the health or behaviour of the Abalone can be more closely observed (Conti et al., 2006).

In terms of time taken, as highlighted in Fig. 15, both models were able to achieve a modest average inference time for 24 objects equal to 606 ms for weaner/grow-out and 394 ms for nursery stages. This confirms that the implemented object detectors are capable of

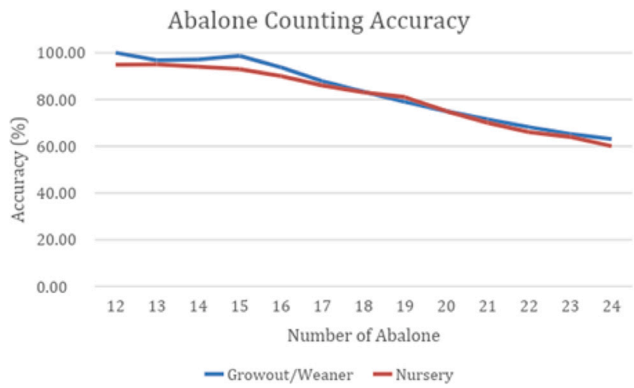


Fig. 14. Percentage accuracy performance of Abalone counting method at increasing numbers of Abalone.

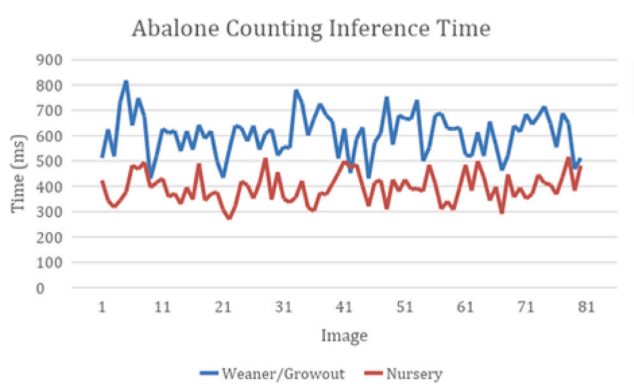


Fig. 15. Graph of abalone counting inference time.

achieving real-time inference while still maintaining a high base level of accuracy for up until 15 objects at once.

Manual counting methods are difficult to quantify because it is highly dependent on the individual doing the counting. Assuming the average human response time is 231 ms (Woods et al., 2015), it would take several seconds for an average person to count 24 objects with a high probability of achieving 100% accuracy, thus it was not considered in these results, as an image processing technique will nearly always be significantly faster.

7.3. Discussion

From the experimental results, we can confidently conclude that our methods are both more efficient and accurate than traditional and estimation techniques for counting and measuring abalone. While our methods cannot be used simultaneously as we originally intended, the tools by themselves are useful and provide a degree of improvement over traditional methods that warrant its field usage. For counting, our method was able to effectively up to 15 abalone at once in real-time with an accuracy between 98%–100%. For more than 15 objects, the accuracy slowly decreases linearly. We found that this seemed odd and contradictory based on our initial research and experiments, and as such would like to discuss this further.

One seemingly explanatory cause is due to the limitations smartphones, as discussed earlier. However, another potential cause of performance loss is due to the use of the YOLOv4 architecture which has historically had issues counting small objects, evident through its efficacy on other datasets with smaller image sizes (Bochkovskiy et al., 2020). We believe, due to its single-shot detection architecture, it is not as suitable for capturing finer details present in smaller objects and 416 $\times$ 416 images. The model possesses a tradeoff of speed over accuracy, which was valued for our purpose, but in doing so, it is unable to fully detect all objects in an image to a value greater than the minimum specified accuracy threshold value.

Another explanation is our method of experimentation, where the objects may have been too close together, or camera too distant away from the subjects themselves. Due to the scope of our research and time restrictions, we were unable to perform the diverse range of experiments to obtain further results for analysis. A potential future direction and solution may be to split the images up into smaller regions in real-time and have that be fed into the model for potentially increased counting performance.

From our testing, measuring using mobile-based AR has proven to offer major speed advantages when it comes to measuring smaller objects, such as weaner abalone shells, or larger creatures such as reef fish. Our method of measuring demonstrated robustness in either case, where the accuracy remained consistent across different sizes of objects, indicating that the method can be extended to other applications and fields such as agriculture, on-site training, or education. However, in the case of very small objects, AR may not be ideal solution as results from experiments revealed that sometimes the augmented information can obscure the edges of small objects, such that they are hidden from the user's view.

With these results, it indicates that our approach can assist farmers in nearly doubling their efficiency when stock taking without risking harm or shock to the animal, and without them requiring the assistance of trained technical staff who are forced to precisely, and time consumingly, remove abalone in a manner to prevent death. While this technique will never fully replace the technical expertise required to grow abalone sustainably and efficiently, it can serve as an optimisation tool, to go alongside farmers for increased productivity and objectivity whilst monitoring stock growth.

8. Conclusion

Object detection is a rapidly evolving technology area that has attracted many researchers, yet little work has been conducted in the aquaculture sector, despite its high applicability. Through this paper, we have demonstrated a prototype approach to solving the issues experienced by abalone farmers when it comes to monitoring the growth of their populations. This research also has the potential the impact other areas of study, particularly fields in which more effective, image-based automation is required in unpredictable conditions or environments.

Overall, we have shown that our framework is indeed more efficient and effective than manual and estimation monitoring methods. This, in turn, alleviates the problems of inaccuracy and inefficiency identified within manual methods. Our proposed method is real-time, location and network independent and it does improve the accuracy and speed of counting and measuring of both juvenile and matured abalone. Object recognition has proven to be a promising approach to automate the processes of counting abalone at various stages of their lifecycle and can be a good foundation for further automation research. Likewise, AR has shown that it is a fast and robust way of measuring objects and can be applied many different types of objects outside of abalone.

Our framework could be further extended to recognise multiple, overlapping, moving abalone at various orientations and simultaneously measure their lengths. Our implementation also still does not fully overcome the issue of the need for trained technical staff, as it still requires manual operation by a user. Additional modifications to the approach would revolve around further removing the need for user-interaction such as auto-calibration and event triggering, which, in turn, would help improve measuring accuracy and reduce data capturing times. Moreover, our framework only operates on the Android platform, so future work could look to expand this to a multi-platform version.

The method proposed was entirely tested in laboratory conditions, where lighting was consistent, and conditions were calm. As a result, more testing of counting is required, with additional variations in conditions, such as moving objects, motion blur and ideally in-situ validation. We believe the next step in furthering this work would be to test this framework in the field to validate whether the resources proposed contribute to reducing handling time, and to a faster and more effective decision-making process. Moreover, further directions include performance experiments with occluded and overlapping Abalone. as well as measuring the power consumption and storage requirements during data collection. Finally, it is worth noting that the combination of these technologies could yield valuable results that are widely applicable to areas within and outside of aquaculture. The gap identified from our literature review still exists and may be deserving of future research with more automated methods, such as the employment of autonomous drones for total human-less monitoring.

CRediT authorship contribution statement

Thomas Napier: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualisation. **Ickjai Lee:** Writing – review & editing, Formal analysis, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Anastasiou, E., Balafoutis, A.T., Fountas, S., 2023. Applications of extended reality (XR) in agriculture, livestock farming, and aquaculture: A review. Smart Agric. Technol. 3, 100105. http://dx.doi.org/10.1016/j.atech.2022.100105, URL: https://www.sciencedirect.com/science/article/pii/S2772375522000703.

Baumgartner, L.J., Bettanin, M., McPherson, J., Jones, M., Zampatti, B., Beyer, K., 2012. Influence of turbidity and passage rate on the efficiency of an infrared counter to enumerate and measure riverine fish. J. Appl. Ichthyol. 28, 531–536. http://dx.doi.org/10.1111/j.1439-0426.2012.01947.x.

Bochkovskiy, A., Wang, C.-Y., Liao, H.-Y.M., 2020. YOLOv4: Optimal speed and accuracy of object detection. arXiv:2004.10934.

Cao, S., Zhao, D., Liu, X., Sun, Y., 2020. Real-time robust detector for underwater live crabs based on deep learning. Comput. Electron. Agric. 172, 105339. http://dx.doi.org/10.1016/j.compag.2020.105339.

Caria, M., Todde, G., Sara, G., Piras, M., Pazzona, A., 2020. Performance and usability of smartglasses for augmented reality in precision livestock farming operations. Appl. Sci. 10 (7), http://dx.doi.org/10.3390/app10072318, URL: https://www.mdpi.com/2076-3417/10/7/2318.

Conti, S.G., Roux, P., Fauvel, C., Maurer, B.D., Demer, D.A., 2006. Acoustical monitoring of fish density, behavior, and growth rate in a tank. Aquaculture 251, 314–323. http://dx.doi.org/10.1016/j.aquaculture.2005.06.018.

Cook, P.A., 2016. Recent trends in worldwide abalone production. J. Shellfish Res. 35, 581–583. http://dx.doi.org/10.2983/035.035.0302.

Dash, A.K., Behera, S.K., Dogra, D.P., Roy, P.P., 2018. Designing of marker-based augmented reality learning environment for kids using convolutional neural network architecture. Displays 55, 46–54. http://dx.doi.org/10.1016/j.displa.2018.10.003.

de Oliveira, M.E., Corrêa, C.G., 2020. Virtual Reality and Augmented reality applications in agriculture: a literature review. In: 2020 22nd Symposium on Virtual and Augmented Reality (SVR). pp. 1–9. http://dx.doi.org/10.1109/SVR51698.2020.00017.

Eklind, A., Stark, L., 2018. An exploratory research of ARcore's feature detection. https://pdfs.semanticscholar.org/9741/4cccbbf2e298b4e29ca3f9c830c6a69dca34.pdf.

Fallu, R., 1991. Abalone Farming. Fishing News Books, URL: https://www.cabdirect.org/cabdirect/abstract/19930104091.

França Albuquerque, P.L., Garcia, V., da Silva Oliveira, A., Lewandowski, T., Detweiler, C., Gonçalves, A.B., Costa, C.S., Naka, M.H., Pistori, H., 2019. Automatic live fingerlings counting using computer vision. Comput. Electron. Agric. 167, 105015. http://dx.doi.org/10.1016/j.compag.2019.105015, URL: https://www.sciencedirect.com/science/article/pii/S0168169919302881?dgcid=coauthor.

Garlock, T., Asche, F., Anderson, J., Bjørndal, T., Kumar, G., Lorenzen, K., Ropicki, A., Smith, M.D., Tveterås, R., 2020. A global blue revolution: Aquaculture Growth Across Regions, species, and countries. 28 (1), 107–116. http://dx.doi.org/10.1080/23308249.2019.1678111, Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/23308249.2019.1678111.

Hart, A.M., Strain, L.W.S., Hesp, S.A., 2013. Stock enhancement in greenlip abalone part III: Bioeconomic evaluation. 21 (3), 354–374. http://dx.doi.org/10.1080/10641262.2013.812506, Publisher: Taylor & Francis _eprint: http://dx.doi.org/10.1080/10641262.2013.812506.

Heasman, M., Savva, N., 2007. Manual for intensive hatchery production of abalone: Theory and practice for year-round, high density seed production of blacklip abalone (haliotis rubra). URL: https://www.lib.washington.edu/msd/norestriction/b67512665.pdf.

Hernández-Ontiveros, J., Inzunza-González, E., García-Guerrero, E., López-Bonilla, O., Infante-Prieto, S., Cárdenas-Valdez, J., Tlelo-Cuautle, E., 2018. Development and implementation of a fish counter by using an embedded system. Comput. Electron. Agric. 145, 53–62. http://dx.doi.org/10.1016/j.compag.2017.12.023.

Hong, H., Yang, X., You, Z., Cheng, F., 2014. Visual quality detection of aquatic products using machine vision. Aquac. Eng. 63, 62–71. http://dx.doi.org/10.1016/j.aquaeng.2014.10.003.

Huuskonen, J., Oksanen, T., 2019. Augmented reality for supervising multirobot system in agricultural field operation. IFAC-PapersOnLine 52 (30), 367–372. http://dx.doi.org/10.1016/j.ifacol.2019.12.568, URL: https://www.sciencedirect.com/science/article/pii/S2405896319324899, 6th IFAC Conference on Sensing, Control and Automation Technologies for Agriculture AGRICONTROL 2019.

Koprowski, R., Wróbel, Z., Kleszcz, A., Wilczyński, S., Woźnica, A., Łozowski, B., Pilarczyk, M., Karczewski, J., Migula, P., 2013. Mobile sailing robot for automatic estimation of fish density and monitoring water quality. BioMed. Eng. OnLine 12, 60. http://dx.doi.org/10.1186/1475-925x-12-60.

Lee, D., Lee, K., Kim, S., Yang, Y., 2015. Design of an optimum computer vision-based automatic abalone (haliotis discus hannai) grading algorithm. J. Food Sci. 80, E729–E733. http://dx.doi.org/10.1111/1750-3841.12799.

Li, D., Miao, Z., Peng, F., Wang, L., Hao, Y., Wang, Z., Chen, T., Li, H., Zheng, Y., 2020. Automatic counting methods in aquaculture: A review. J. World Aquac. Soc. 52, 269–283. http://dx.doi.org/10.1111/jwas.12745.

Lovreglio, R., Kinateder, M., 2020. Augmented reality for pedestrian evacuation research: Promises and limitations. Saf. Sci. 128, 104750. http://dx.doi.org/10.1016/j.ssci.2020.104750.

Martinez-Alpiste, I., Golcarenarenji, G., Wang, Q., Alcaraz-Calero, J.M., 2022. Smartphone-based real-time object recognition architecture for portable and constrained systems. J. Real-Time Image Proc. 19, 103–115. http://dx.doi.org/10.1007/s11554-021-01164-1.

Orciuoli, F., Orciuoli, F.J., Peduto, A., 2020. A Mobile Clinical DSS based on Augmented Reality and Deep Learning for the home cares of patients afflicted by bedsores. Procedia Comput. Sci. 175, 181–188. http://dx.doi.org/10.1016/j.procs.2020.07.028.

Parr, M.B., Byler, R.K., Diehl, K.C., Hackley, C.R., 1995. Machine vision based oyster meat grading and sorting machine. J. Aquatic Food Prod. Technol. 3, 5–24. http://dx.doi.org/10.1300/j030v03n04_02.

Phupattanasilp, P., Tong, S.-R., 2019. Augmented reality in the integrative internet of things (AR-IoT): Application for precision farming. Sustainability 11 (9), http://dx.doi.org/10.3390/su11092658, URL: https://www.mdpi.com/2071-1050/11/9/2658.

Pinkiewicz, T.H., Purser, G.J., Williams, R.N., 2011. A computer vision system to analyse the swimming behaviour of farmed fish in commercial aquaculture facilities: A case study using cage-held Atlantic salmon. Aquac. Eng. 45, 20–27. http://dx.doi.org/10.1016/j.aquaeng.2011.05.002, URL: https://www.sciencedirect.com/science/article/pii/S014486091100029X.

Rahman, A., Xi, M., Dabrowski, J.J., McCulloch, J., Arnold, S., Rana, M., George, A., Adcock, M., 2021. An integrated framework of sensing, machine learning, and augmented reality for aquaculture prawn farm management. Aquac. Eng. 95, 102192. http://dx.doi.org/10.1016/j.aquaeng.2021.102192.

Rasmussen, C., Zhao, J., Ferraro, D., Trembanis, A., 2017. Deep census: AUV-based scallop population monitoring. In: 2017 IEEE International Conference on Computer Vision Workshops (ICCVW). http://dx.doi.org/10.1109/iccvw.2017.338.

Reaburn, R., Edwards, S., 2003. Modelling the growth of cultured greenlip abalone (Haliotis laevigata leach): practical modes of measurement. Aquac. Res. 34, 253–259. http://dx.doi.org/10.1046/j.1365-2109.2003.00814.x.

Spampinato., C., Chen-Burger., Y., Nadarajan., G., B. Fisher., R., 2008. Detecting, tracking and counting fish in low quality unconstrained underwater videos. In: Proceedings of the Third International Conference on Computer Vision Theory and Applications - Volume 1: VISAPP, (VISIGRAPP 2008). SciTePress, INSTICC, pp. 514–519. http://dx.doi.org/10.5220/0001077705140519.

Troell, M., Robertson-Andersson, D., Anderson, R.J., Bolton, J.J., Maneveldt, G., Halling, C., Probyn, T., 2006. Abalone farming in South Africa: An overview with perspectives on kelp resources, abalone feed, potential for on-farm seaweed production and socio-economic importance. 257, 266–281. http://dx.doi.org/10.1016/j.aquaculture.2006.02.066, URL: https://www.sciencedirect.com/science/article/pii/S0044848606001906.

Venter, L., Loots, D.T., Vosloo, A., Jansen van Rensburg, P., Lindeque, J.Z., 2016. Abalone growth and associated aspects: now from a metabolic perspective. Rev. Aquac. 10, 451–473. http://dx.doi.org/10.1111/raq.12181.

White, D., Svellingen, C., Strachan, N., 2006. Automated measurement of species and length of fish by computer vision. Fish. Res. 80, 203–210. http://dx.doi.org/10.1016/j.fishres.2006.04.009.

Woods, D.L., Wyma, J.M., Yund, E.W., Herron, T.J., Reed, B., 2015. Factors influencing the latency of simple reaction time. Front. Human Neurosci. 9, http://dx.doi.org/10.3389/fnhum.2015.00131, URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4374455/.

Xi, M., Adcock, M., McCulloch, J., 2018. Future agriculture farm management using augmented reality. In: 2018 IEEE Workshop on Augmented and Virtual Realities for Good (VAR4Good). http://dx.doi.org/10.1109/var4good.2018.8576887.

Xi, M., Adcock, M., Mcculloch, J., 2019. An end-to-end augmented reality solution to support aquaculture farmers with data collection, storage, and analysis. In: Proceedings of the 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry. VRCAI '19, Association for Computing Machinery, New York, NY, USA, http://dx.doi.org/10.1145/3359997.3365721.

Xiao, Y., Tian, Z., Yu, J., Zhang, Y., Liu, S., Du, S., Lan, X., 2020. A review of object detection based on deep learning. Multimedia Tools Appl. 79, 23729–23791. http://dx.doi.org/10.1007/s11042-020-08976-6, URL: https://link.springer.com/article/10.1007%2Fs11042-020-08976-6.