Application of a machine learning approach for effective stock management of abalone

Kurt Schoenhoff, Kyungmi Lee, Jason Holdsworth, Hemmaphan Suwanwiwat, Jan Strugnell, Ickjai Lee

December 2022

FRDC Project No 2019/151

ISBN 978-0-6457422-6-8

Application of a machine learning approach for effective stock management of abalone
2019/151

2022

Ownership of Intellectual property rights

Creative Commons licence

Disclaimer

The Fisheries Research and Development Corporation plans, invests in and manages fisheries research and development throughout Australia. It is a statutory authority within the portfolio of the federal Minister for Agriculture, Fisheries and Forestry, jointly funded by the Australian Government and the fishing industry.

Researcher Contact Details

Name: Prof. Jan Strugnell
Address: James Cook University
Townsville 4811
Phone: 07 47816357
Fax: -
Email: jan.strugnell@jcu.edu.au

FRDC Contact Details

Address: 25 Geils Court
Deakin ACT 2600
Phone: 02 6285 0400
Fax: 02 6285 0499
Email: frdc@frdc.com.au
Web: www.frdc.com.au

In submitting this report, the researcher has agreed to FRDC publishing this material in its edited form.

Contents

Figures

Acknowledgments

This project was funded by The Australian Abalone Growers Association and the Fisheries Research and Development Corporation on behalf of the Australian Government. Southern Ocean Mariculture and Yumbah provided images for this study and also access to these farms in order to obtain images. This project also was carried out in collaboration with Phoebe Arbon (James Cook University), who was awarded a 2020 Science and Innovation Award for young people in agriculture, fisheries and forestry. The purpose of that project was to develop a machine learning model which would allow for the quantitative assessment of abalone in growout production systems, specifically slab tanks, through analysis of images.

Executive Summary

This report provides detail on the development of a machine learning tool as a method for counting and measuring abalone at various stages of production. The study was carried out on hybrid abalone with ~2000 images (nursery, weaner and growout stage) collected from Southern Ocean Mariculture and Yumbah. A deep learning based method for counting abalone (nursery stage) and counting and measuring abalone (weaner and growout stages) was successfully developed and trained. A user-friendly cross platform software application was developed to enable use of the tool by abalone farmer. The tool should make stock assessment faster, more accurate and provide less stress to farmers. The project was carried out from 2019-2022 by the James Cook University Information Technology team in Cairns (Kyungmi Lee, Ickjai Lee, Jason Holdsworth, Hemmaphan Suwanwiwat, Kurt Schoenhoff) in with contributions from Phoebe Arbon (who held a 2020 Science and Innovation Award for young people in agriculture, fisheries and forestry with a similar time frame) and Jan Strugnell.

Background

Effective stock management of abalone relies upon the collection of the number and size distribution of abalone at various stages of production. Currently this information is collected by hand which is expensive, limited in scope and causes stress to animals which can compromise growth and survival. Machine learning is now at a stage that this approach can be applied to various stages of abalone production to enable automated counting and measuring of abalone.

Aims/objectives

The aim of this project was to develop and implement artificial intelligence as a method for accurately measuring and counting abalone at nursery, weaning and growout.

Methodology

Images were collected from Southern Ocean Mariculture and Yumbah from all stages of production (nursery, weaner and growout). Subsets of these images were annotated and used to develop and train deep learning based models for different stages of the abalone lifecycle. An agile and iterative co-development processes was employed to ensure efficiency in annotating images and model development.

Results/key findings

A highly accurate machine learning based method for counting abalone (nursery stage) and counting and measuring abalone (weaner and growout stages) was successfully developed and trained. A user-friendly cross platform software application was developed to enable use of the tool by abalone farmer. This will enable fast and effective counting and average length estimation by farmers utilising commonly available mid-tier camera equipment for data acquisition and commonly available computing hardware. A *.csv file can be downloaded from the application for subsequent use by the farmer.

Implications for relevant stakeholders

The main outcome of this project is vastly increased efficiency and improved accuracy in counting and measuring abalone in production systems which will be of benefit for the Australian abalone aquaculture industry. The tool will provide improvements in efficiency and productivity in the short term and an advanced platform for research and development (e.g. feed trials etc) going forward.

Recommendations

Recommended computational hardware to run the software application includes: CPU: 6+ Core 3.5GHz (AMD or Intel), GPU: Nvidia RTX3060, RAM: +16GB DDR4 non-ECC, Storage: 128GB SSD/NVME, Operating System: Ubuntu Server 20.04LTS. In April 2022 this would cost ~$2,000. To collect suitable images a +20 Effective mega pixel camera on a fixed frame is recommended.

There is considerable scope for further development of this tool for other abalone farming applications. These include, but are not limited to, possibilities for direct weight estimation, whole tank counts and identification of food and faeces. In addition, automated integration with a database system, automatic tank detection, and improvements in speed of the system are all possible areas for future development.

Keywords

Abalone, artificial intelligence, *Haliotis*, machine learning

Introduction

In many fisheries and aquaculture management applications the ability to gain reliable data in a timely manner has always been a limiting factor. Until the last few years, machine learning was not able to reliably learn autonomously from the target imagery without substantial human oversight during the training process. This was mostly due to limitations in inefficiencies and ineffectiveness in learning models and processing power, however this has recently changed, and this field is now rapidly expanding and improving with advances in deep learning being applied to provide efficiency and effectiveness for industry.

At James Cook University we are solving bottlenecks in data collection and interpretation for fisheries and aquaculture industries by implementing machine learning techniques that utilise varying levels of human supervision. The projects include the identification of freshly caught fish on board commercial fishing vessels and also the assessment of the composition of benthic fauna to better manage fisheries habitat. For recent applications assessing fish moving against artificial marine structures, we have been able to repurpose machine learning source code that had previously been developed for a different purpose meaning that no extensive re-training of the artificial intelligence was required.

Abalone aquaculture is well-suited to a machine learning approach. Abalone are largely stationary, relatively uniform in shape and tend to be easily detected against the background substrate. Application of machine learning will provide significant efficiencies and improved accuracies in counting and measuring abalone and will remove handling induced stress.

Determining the number and size distribution of abalone present at various stages of production is critical information for effective stock management. Currently the Australian abalone aquaculture industry gathers this information through time consuming manual measurements taken periodically. However, the resulting data is of mediocre quality, is limited in its scope, and collecting the data causes stress to the animals (as it is removed from the water) which can compromise growth and survival. Automated counting and measuring of abalone will increase the amount of data and frequency of data collection leading to gains infarm efficiency and productivity in the short term and, in the longer term, will provide an advanced platform for further R & D improvements including accurate data collection during experimental trials (e.g. feeds, temperature). Artificial intelligence and machine learning has now matured to a point that accurately counting and measuring abalone is possible using this approach, however specific application to the abalone industry is yet to be achieved.

This project involves the development, training and validation of a machine learning model to identify, segment and measure quantitative abalone traits in production systems and, render the product data to be accessible and applicable for farmers.

Objectives

The objective of this study was to develop and implement artificial intelligence as a method for accurately measuring and counting abalone at nursery, weaning and growout.

Method

In recent years Artificial Intelligence (AI) using Artificial Neural Networks, often referred to as “Networks” or “Models”, has excelled at image processing. These networks are often comprised of many layers and rather than the task being “programmed” into them, many examples of the desired outcome are used to “train” the network to perform the required task. This technique is generally known as Machine Learning (ML) and when done utilising deep networks is referred to as Deep Learning (DL).

DL based learning differs from human learning in that a DL model is ineffective at gaining a generalised effectiveness for a task be one or few examples, to be effective DL networks generally require many examples with which to build up a capability for the generalised task. To be able to develop an effective DL solution for counting and measuring abalone we first require images and annotations to describe the task we wish to perform (counting and measuring).

During two field trips to abalone farms in Victoria (Dec 2020 and April 2021) images were captured of abalone throughout the nursery, hides and slab growout stages. The April 2021 field trip was combined with a field trip for a 2020 Science Innovation Award awarded to Phoebe Arbon (Aquaculture Department, James Cook University), that was focussed on the slab growout of Abalone. These two field trips resulted in ~2,000 images being taken of abalone.

Subsets of these images were annotated and used to develop and train DL based modes for the tasks required for different stages of the abalone lifecycle. To ensure the most effective use of time in annotating and model development the general approach taken was an agile / iterative process. In this process the networks and datasets are co-developed with a series of annotations performed then training is performed on the annotations. This ensured that problem areas could be identified and rectified promptly. A generalised framework for the project is shown in Figure 1.

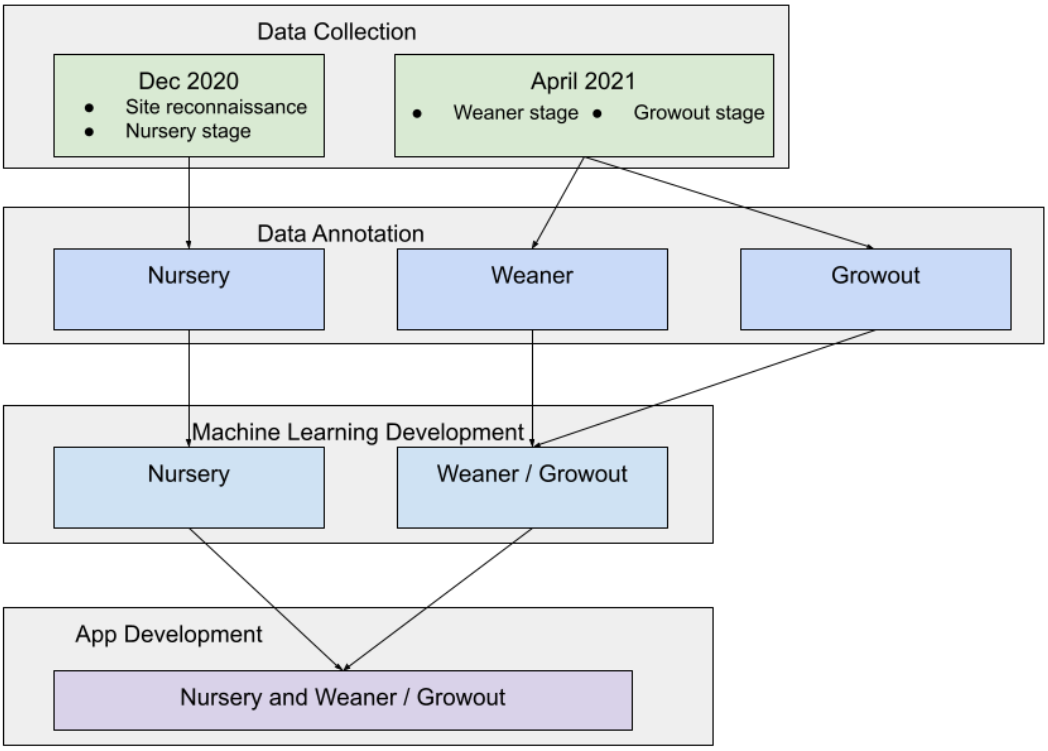


Figure 1. A generalised framework for this project.

After models were suitably developed and trained a host application needed to be developed to allow an end user without programming experience to utilise the trained models. This application also housed the necessary pre and post processing as part of a data pipeline to negate the need for user involvement throughout the process.

An early version of the application (nursery stage - counting only) was supplied to Southern Ocean Mariculture (SOM) and was trialled for usability and effectiveness. During this trial it was found that the system was easy to use however it was not effective at counting abalone at around two months old, this age group was missing from the dataset. This was rectified by annotating a number of images at this age, including in the dataset and retraining.

Nursery Stage

During this stage the application was developed for counting the youngest abalone (<20 mm) that have shells that have become opaque and are growing attached to a clear/opaque sheet of plastic referred to as a “slide”. These slides are photographed with a high-resolution camera (20+MP) and the images are processed by a DL model resulting in a count of abalone per image (1 side of a slide).

Initially, several methods were examined for suitability, of these the most suitable candidates for further exploration were based upon the following methods:

- Contour / Edge Detection based Computer Vision (CV):

  In this method the image is converted into a binary mask and edges, or series of edges (contours) are used to detect the objects of interest.

- DL based Object Detection:

  In this method a deep Artificial Neural Network is used to perform the object detection after being trained on a set of example images.

After initial experimentation on some provided example images, it was determined that the DL based approach would be the most invariant to changes in image composition (exposure, colour temperature, etc). Figure 2 displays nursery stage dataset and model development.

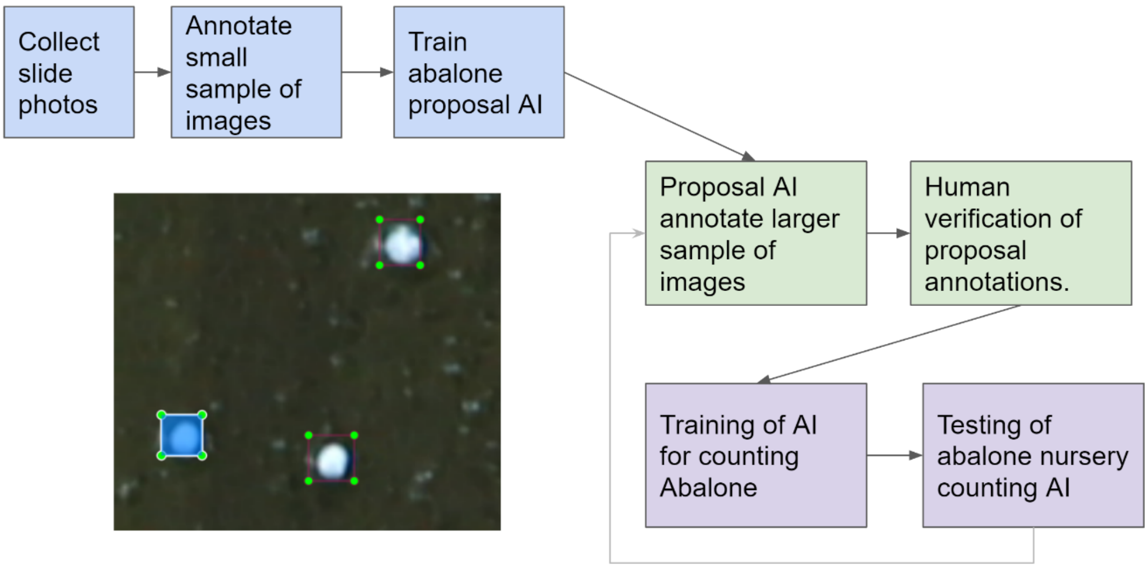


Figure 2. Nursery stage dataset and model development. Note the iterative approach where AI proposals are verified by humans before being used to (re)train. Individual abalone are <20mm.

After each field trip to collect images of the nursery stage a subset of the images was annotated manually and processed to create an initial dataset of ~475 images with 1,793 annotated abalone. This dataset was used to train a DL based object detection network (Faster R-CNN (Ren et al., 2015)). This trained network was used to annotate more of the collected images to produce a dataset that contained over 800 images with over 2500 annotated abalone. The automatically annotated images were verified by experts (JCU & SOM), then this dataset was used for experimentation in determining the best model and training hyperparameters for the final model.

Data (Image) Collection

During December 2020, two members of the JCU Information Technology group travelled to Victoria to photograph abalone for use in creating an abalone counting system during the nursery stage. During this field trip approx. 1000 images were taken of abalone during the nursery stage at SOM. The abalone photographed were aged between 3 – 7 weeks old. The minimum of 3-5 weeks old was used as before this time abalone do not have a visibly calcified shell and are transparent. The oldest nursery stage abalone onsite were 5-7 weeks old. Figure 3 depicts Apparatus used to capture images of nursey stage abalone.

Figure 3. Apparatus used to capture images of nursey stage abalone. Black background and framing to hold the abalone slides allows fast and effective image capture.

Images were taken using a 20MP camera utilising a tripod and frame to hold the slides. Due to the high temperatures during the visit, care was taken to minimise the time the abalone were out of the water to prevent stress and potential mortality. With correct placement of the frame and camera an average time of 20-30 seconds was required to photograph both sides of the slide. Slide dimensions were 300 x 600 (Height x Width) and the frame used to photograph the slides had slits in the sides that were able to hold the slide in a consistent manner without the back of the slide contacting the frame.

Due to the distance between camera and slide being fixed, manual adjustment of the camera settings was used to ensure maximum clarity and image quality. Due to the size of the smallest abalone and the size of the slide a minimum of 30-50 pixel (px) length of the abalone was able to be maintained for the abalone to be detected (counted).

Pre-processing (Data Annotation)

DL based approaches generally require many examples of the task that is to be conducted to perform well. Annotating large datasets for object detection can be a laborious process so an approach was taken to minimise the time spent manually annotating images and maximising the accuracy of the count. Annotation and data preparation for the images was performed in iterative process.

1. A small dataset was manually annotated (475 images), and this was used to train a small capacity variant of the Faster R-CNN object detector utilising a process called "transfer learning". In this process a pretrained (on a similar task) version of the network is used as the initial set of weights. This network is then trained (shown examples) of the specific task required until the network converges to a satisfactory level.
2. The trained network was used to automatically annotate another approx. 400 images. This process produces annotation proposals which were then verified by human experts (JCU & SOM).
3. The verified annotations were used to train a more capable variant of Faster R-CNN in a series of experiments to determine the best model and training hyperparameters.

At the end of this process, we had 800 images with over 2,500 instances of annotated abalone. A "training" dataset comprising of 70% of the images and annotations and a "test" dataset comprising of the remaining images and annotations. This dataset performed well above the required 70-80% accuracy required for this stage and development was halted at this stage. Figure 4 illustrates an overview of the dataset generation and DL process showing the inclusion of slab growout stage annotations.

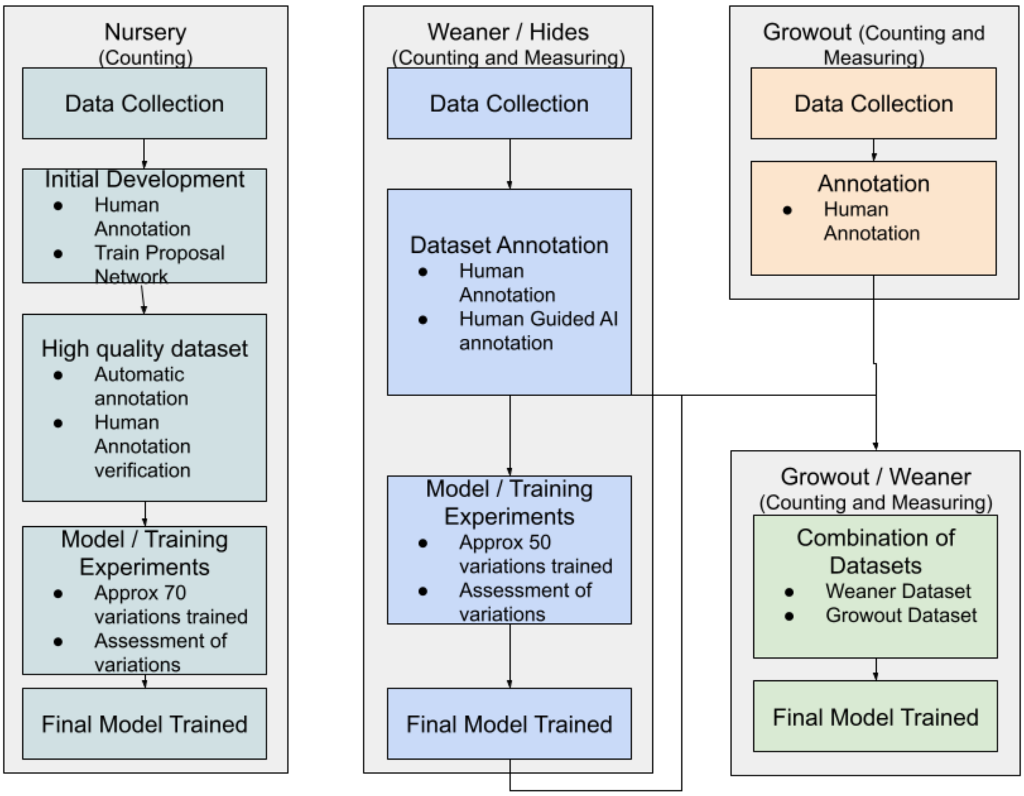


Figure 4. Overview of the dataset generation and DL process showing the inclusion of slab growout stage annotations.

Building the AI (DL) model

The base model selected to perform the task of object detection (for counting) as a Faster R-CNN. The main reason for choosing this architecture was that it is a proven method for accurate fast object detection. It uses standard Residual Neural Network (ResNet) backbone architecture (He et al., 2016) and is available for most

DL frameworks. Utilising standard backbones allows for quick experimentation with different network capabilities to find the best balance between network capacity, network throughput speed, and accuracy.

To minimise training times and therefore the number of different experiments able to be performed in a given time an initial backbone was chosen to be a 26 layer ResNet. Initial training rounds showed that this model was incapable of reliable detections at full (24MP).

Due to the relatively small size of the abalone in the images (10-30px across) downscaling images was not considered. To enable efficient training an image tiling regime was experimented with. The final tiling scheme utilised for processing images was 5x5 with a 20% overlap.

A set of experiments was designed to cover a broad range of training hyper parameters. Included in this hyper parameter search (but not limited to) was learning rate, dropout rate, optimiser types and scheduling strategies, backbone depth, and layer freezing strategies, etc.

The best performing model from these experiments was found to be using a MobileNet (Howard et al., 2017)) based backbone (MobileNet-Large-100) and this model and trained weights were utilised to auto-annotate tiled images to create a more substantial dataset.

For the final model to be used in the application the human verified auto-annotated images were used as the training dataset. Due to having a substantially larger training dataset a similar set of experiments as with the smaller dataset was performed to determine the optimal backbone and training procedures. Once these experiments were performed the final results and training data were analysed to look for trends in model architecture and training regime.

A smaller but more specific and fine-grained set of experiments were performed to produce a more optimal model for deployment. After a trial deployment it was found that the model was not performing well on abalone over 2 months of age. To rectify this, images of the 2+ month old abalone were supplied by SOM and manually annotated. These ~100 images were added to the dataset and the model was retrained. The model was then capable of detecting the older abalone successfully.

Weaner Stage

For the weaner stage (including hides) both count and length of the abalone are important. (Weaner stage abalone range from 13-40mm.) Matching this task to commonly specified image ML tasks we chose instance segmentation where each instance has a mask inferred from the image. To utilise this common task in measuring length we look for the furthest distance between two points (pixels) in the mask.

To collect the images for the creation of a dataset for this task visited onsite for five days in April 2021 and collected ~300 images of abalone from the underside of hides (shown in Figure 5). Also collected during this field trip were ~900 images of abalone in growout slabs. A selection of images was annotated to create a training dataset comprising of >1000 images with over 1600 abalone mask annotations, then series of experiments were performed to optimise model and training regime.

Figure 5. Fast and effective image capture onsite minimise the potential for harming stock.

Pre-processing (Data Annotation)

Due to the significant time involved with manual annotation we utilised a neural network to perform user guided annotations (Deep Extreme Cut) (shown in Figure 6). This method allowed for faster and more accurate generation of the segmentation masks for each abalone.

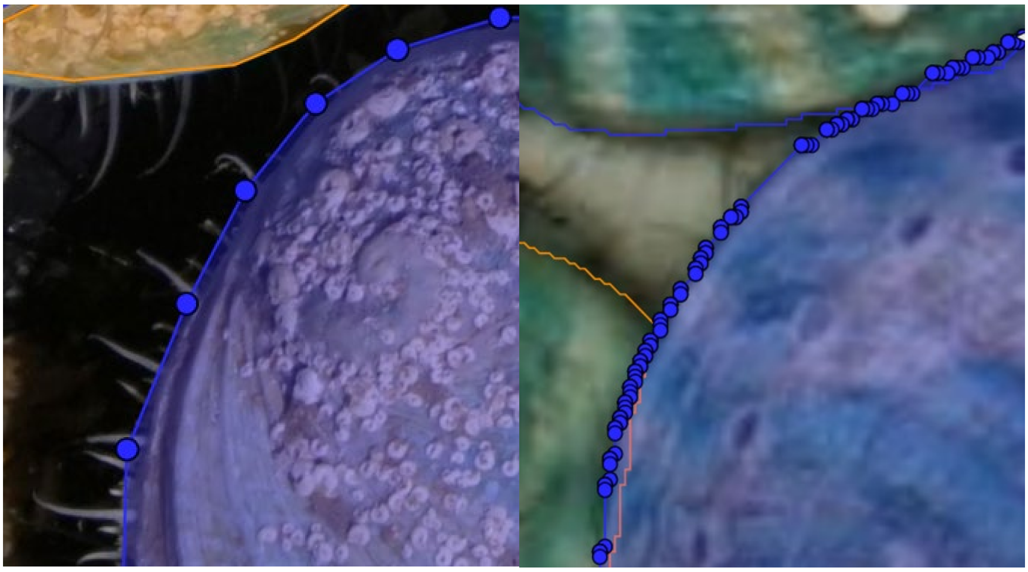


Figure 6. Left – Manual annotation, note few points and mask boundary extends outside the abalone shell. Right – Deep Extreme Cut mask outline is comprised of many points that closely fits the edge of the abalone even at much higher zoom.

Due to the large number of obscured abalone that would be unsuitable for gaining a length measurement. Each of the abalone masks was classified as either "whole" or suitable for length measurements or "obscured" and suitable for counting only.

At the end of this process, we had over 1000 image tiles of annotated abalone with over 16,000 individual abalone masks. Each of the images and associated annotations was then randomly split into two datasets. A "training" dataset comprising of 70% of the images and annotations and a "test" dataset comprising of the remaining images and annotations.

Building AI (DL) model

For instance, for segmentation experimentation, a Mask-R-CNN based segmentation head was used on standard pre-trained backbones available from PyTorch Torchvision. The development followed a similar pathway to that used for the nursery stage. Firstly, a set of experiments was designed to ascertain general trends as to backbone depth and training regime. In this set of experiments, it was mainly combinations of batch size, backbone depth and learning rate that were performed. Firstly, we wanted to get an idea as to what depth of backbone was required for this combination of dataset and task. These experiments showed that the shallower backbones (MobileNets and < 50 layer ResNets) were a bottle neck for information traversing the network for this task and limited overall accuracy of the masks produced. Also discovered was that larger backbones (ResNet101) did not produce substantially better results however they were still included on the next rounds of experimentation to ensure that the best result for the project.

The next set of experiments was for fine tuning network architecture late in the backbone utilising structures available in the literature. The following variants were assessed Feature Pyramid Network (FPN) (Lin et al, 2016), ResNet Conv-4 backbone with Conv5 head (C4) and a ResNet conv-5 backbone with dilations in Conv5 as per Deformable ConvNet (DC5). The most accurate variants were generally the FPN and so we only used these variants for further experiments.

The penultimate experimentation was performed with ResNet50 and Resnt101 FPN variants over a range of learning rates and batch sizes. The final selected model was a ResNet-50 variant as this achieved within 1% of the accuracy of the Resnet-101 and was significantly more efficient.

Late in the application development stage a dataset from the growout stage (20-100 mm) was included in the training data, the best performing models and training regimes were selected for retraining with this data and assessed. Experimentally we found that inclusion of this data did not warrant changing the backbone to a Resnet101 variant as the difference was not substantial enough to make up for the processing increase.

Prototype Software Framework for hosting the DL models

Developing an AI based application often requires the use of AI specific libraries and frameworks. These utilities are generally aimed at ease of experimentation and model development rather than deployment. To reduce development time, aid in cross platform compatibility and reduce programming language complexity we trialled the hosting of a generic classification model in a Django based web framework. After the success of this trial this framework was further developed for the final application. To simplify installation and aid in transportability we placed all of these software components into a Docker container.

Following generally accepted best programming practice, an agile based code development was utilised as the core development methodology. This was closely realised with a simple Kanban style issue/feature on hosted Trello closely coupled with a full Git-Flow git branching methodology using a private GitHub repository for versioning and sharing code.

To streamline development and maintain modularity a generalised data pipeline was proposed. Firstly, the application requires image(s) from the user, these images are then pre-processed, the ML processing takes place (inference), the output of this is taken a processed into the required final form and this is returned to the user. To aid in confidence of the system and to give the user more understanding of the process we annotate (a subset) the original images with the processed ML outputs. Whilst there are some differences in pre/post processing between abalone growth stages a significant amount is the same or similar. Any significant differences will be noted in the process description below, otherwise it should be assumed that the process is shared between growth stages. Figure 7 illustrates a high level diagram showing the breakdown of the application processes.

When opening the app in the browser, the user selects which part of the abalone growing stage they wish to process images for. Nursery stage allows the user to select a different tiling strategy (see pre-processing - tiling) as well as upload a single or selection of images (up to 250mb) by dragging and dropping them on the provided box (as depicted in Figure 8). The Slab/Hides stage is similar however it also allows for the setting of the scale object size and using a pre-set known pixels/mm. Once the “process images” button is pressed the images are uploaded from the browser to the application and the pre-processing stage is commenced.

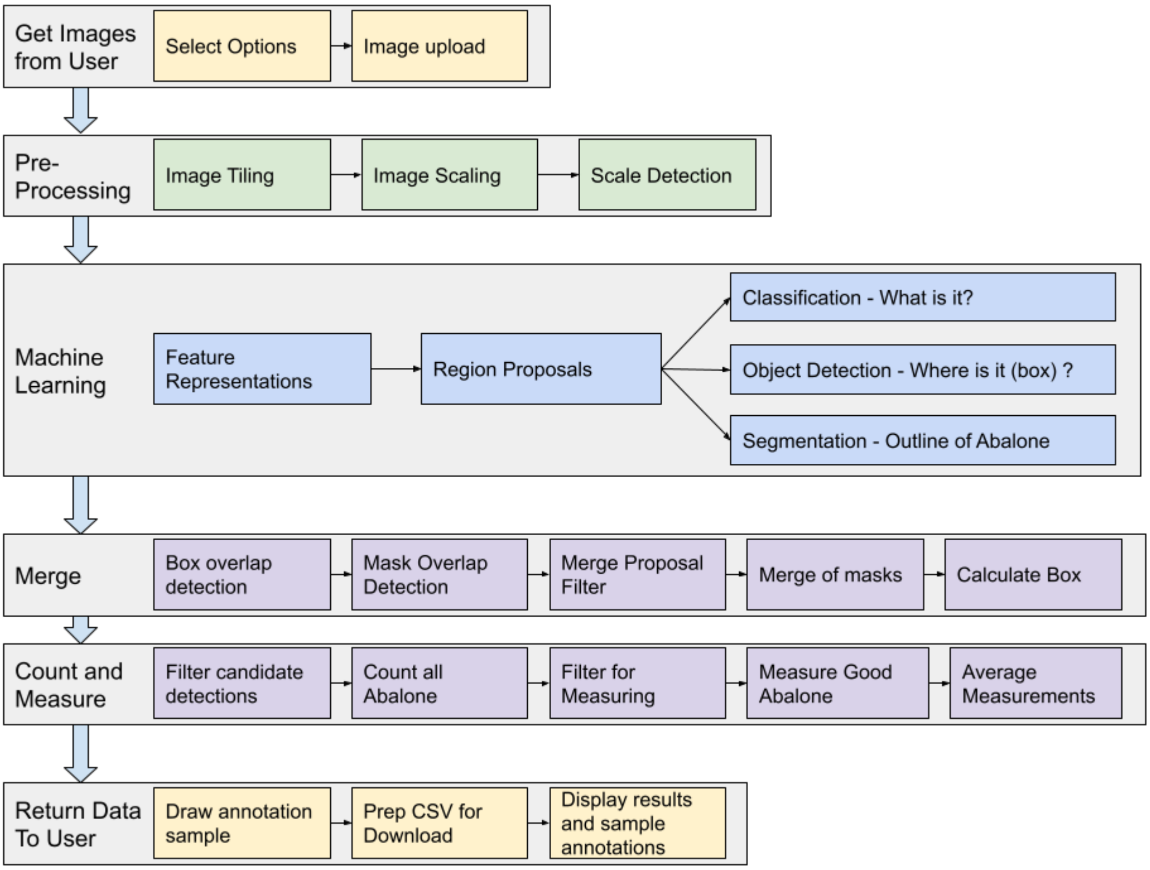


Figure 7. High level diagram showing the breakdown of the application processes. For nursery segmentations (mask) processing and processing specifically for length is omitted.

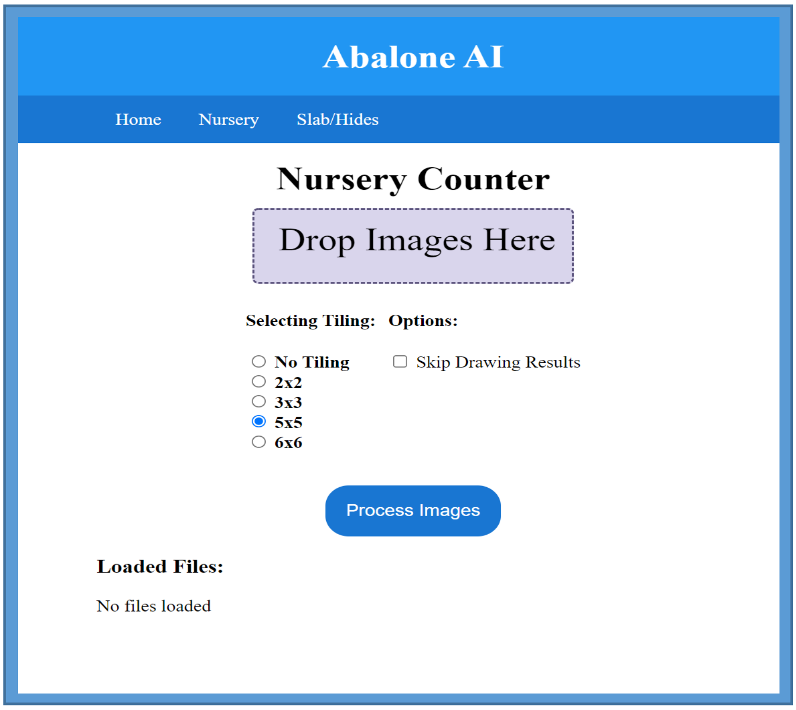


Figure 8. Image input page of application where the user can enter nursery stage impages for counting.

Pre-processing

At this stage the input images and settings were process in way that the ML model is able to be used effectively for the task required.

a. Automatic scale detection

To be able to take measurements from the image domain (pixels) and convert back into real world measurements (mm) a scale (px/mm) is required. For a fixed, rigid photography setup it is trivial (potentially more accurate) to manually measure this and use this as a conversion. For situations where there may be some variance in distances between the lens and the abalone, we included an automatic method that may help in reducing these errors provided by Phoebe Arbon, developed as part of her award for Young People in Forestry, Fisheries and Agriculture (2020) (illustrated in Figure 9).

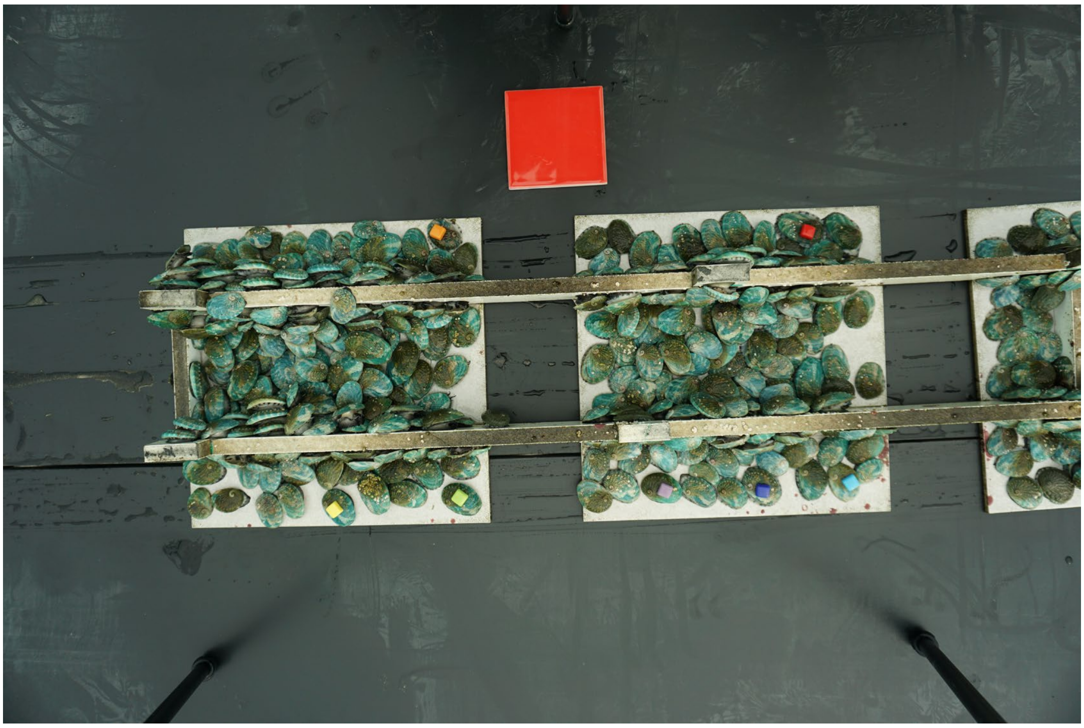


Figure 9. Example of correct scale object placement, measurement plane is the same as the abalone to be measured and is perpendicular to the camera lens.

Note: This method still requires the use of a camera mounted perpendicular to the plane in which the abalone to be measured are on. It also requires the measurement object (red square) to be accurately measured and placed on the same plane. Additionally, the measurement object should be within the centre 80% of the image. Deviations from these requirements will lead to incorrect scale detection which will lead to inaccurate length measurements for all abalone in the image and potentially skew the measurements for the whole batch of images.

b. Tiling

Due to the scale invariance of the Convolutional Neural Networks (CNN) utilised for image based DL tasks, coupled with the requirement to detect abalone over a range of sizes in an image. We utilised a variable image tiling with overlap scheme that allows optimum detection over the required ranges expected in use. This system gives the end user control over image tiling for both nursery and hide/slab stages. We also included a visual guide for end users to help with determining the best tiling scheme for the image(s) (shown in Figure 10).

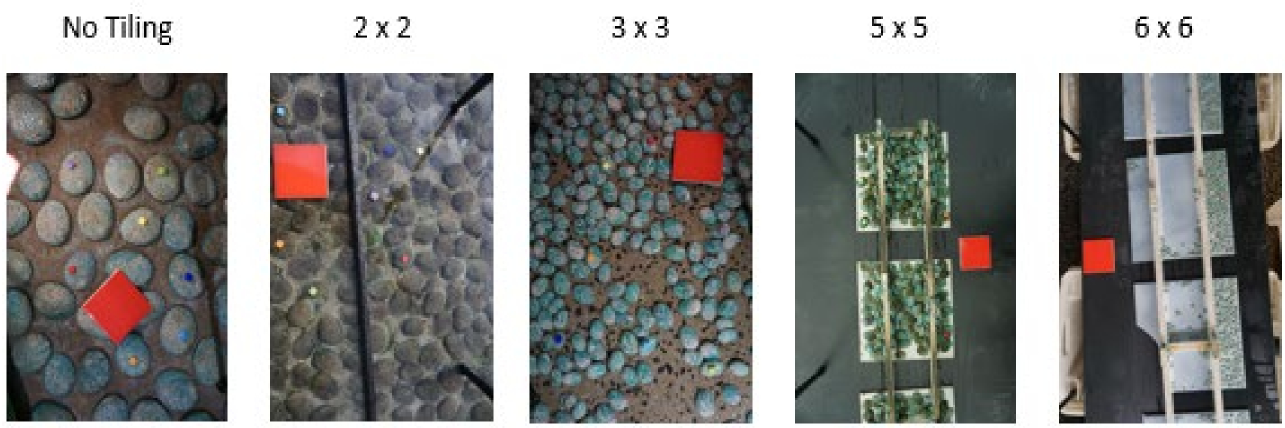


Figure 10. Example images provided as a guide to tiling strategy selection.

Machine learning inference

The ML processing is handled by two separate models. The nursery stage is conducted by a Faster R-CNN that uses a Resnet50 backbone. The weaner and growout stages are undertaken utilising a Mask R-CNN with a Resnet50 based backbone. Both of these networks take image data in the form of a tensor with shape (*b*, 3, *w*, *h*) where *b* is the batch, and *w*, *h* being width and height respectively. Processing is performed on a CUDA GPU if available (recommended) with ability to use the CPU as a fall back for processing.

Post processing

A significant portion of the processing requirements of the application is required for post processing the results from the ML processing. Firstly, results are limited to a maximum of 200 per tile/scaled image (top 200 taken). These are the filtered based upon confidence level from the classification, detections below 50% confidence are dropped. Figure 11 depicts how masks from titles overlaid upon abalone are merged together.

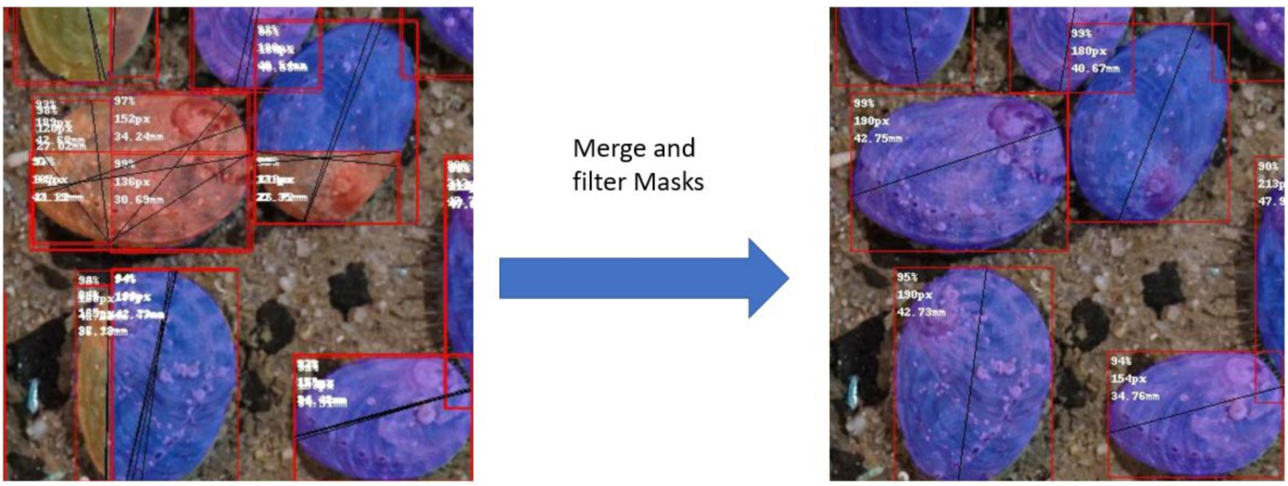


Figure 11. Left – Example of masks from tiles overlaid upon abalone before merge process. Right – Example of successfully merged masks.To reduce rounding errors masks and boxes are then rescaled back to the original image dimensions and checked for suitability for merging. Merging suitability first checks for overlap with the box co-ordinates (this is where it stops for nursery), secondly If the boxes overlap the masks are checked for overlap via an Intersection over Union (Jaccard Index) if above 0.04 then the masks are merged and the boxes are recalculated. If any of the proposal masks are a whole abalone, then this label will be preserved for the merged mask.

As a final filter, all whole abalone masks are checked to see if the lie within 1% of the edge of the image, if so the label is changed so that the abalone is not used for measuring length.

Whole abalone are then processed to get the maximum distance between points, this measurement is scaled from pixels to mm using the automatically detected or pre-set scale. This measurement is used to determine the average length of abalone for an image.

Return data to user

As the final stage in the pipeline, the inferred information is made available to the user (Figure 12). This is in the form of a table displayed on the screen and a "*.CSV" file that is made available for download. Also available for the user to download and inspect is an overlay of the annotations over the input image. This feature is designed to give confidence in the application and allow quick identification of issues that may arise from such potential issues as poor image quality, incorrect tiling strategy, etc.

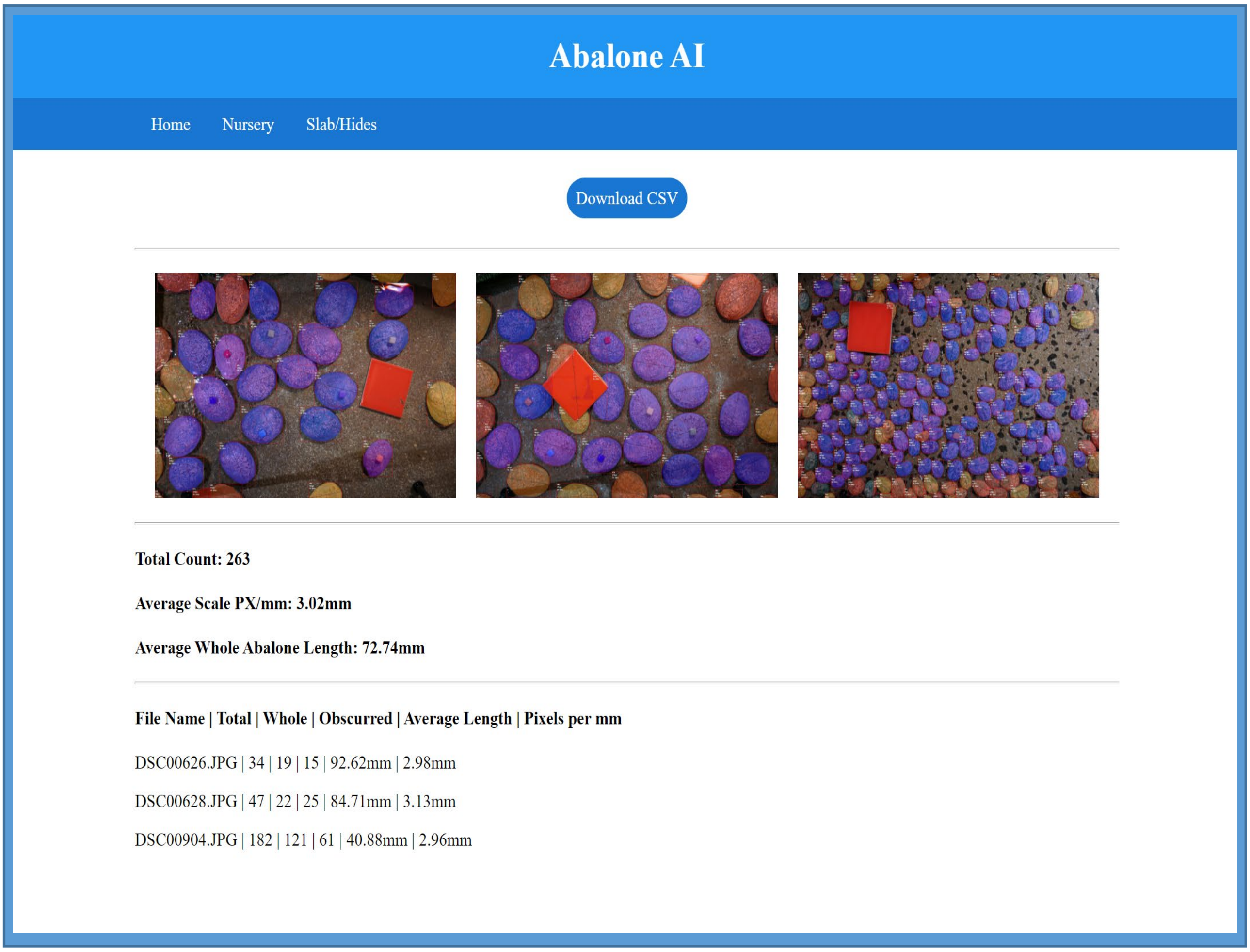


Figure 12. Example of returned information to the user. Note: Blue masks indicate abalone suitable for length measurement and red denotes counting only.

Results / Discussion

During this project we were able to successfully develop and train ML based methods for counting abalone in the nursery stage as well as count, determine suitable abalone for measurement, and measure abalone being grown in the hides and slab tanks. We also developed a cross platform application that should allow farmers to simply utilise these technologies to allow fast and effective counting and average length estimation utilising commonly available mid-tier camera equipment for data acquisition and commonly available computing hardware.

Nursery

In this task we are counting detected abalone with an Intersection over Union (IoU) over 50%, utilising the MS-COCO evaluation metrics for objected detection on images the system had not seen during training. Average Precision (AP) serves as a measure to evaluate the performance of object detectors, it is a single number metric that encapsulates both precision and recall and summarizes the Precision-Recall curve by averaging precision across recall values from 0 to 1 with 1 being a perfect score. Our network achieved an AP score of .96 indicating it is capable of detecting and localising abalone with a high level of accuracy (Figure 13).

When inspecting images qualitatively we notice that the system is very good at not detecting false positives, however it may fail to find some of the abalone in the image and may merge some overlapping abalone into a single detection. We feel that this is important as it shows the system can discern between the target species, abalone, and incorrect species (copepods) and image artefacts such as reflections very well.

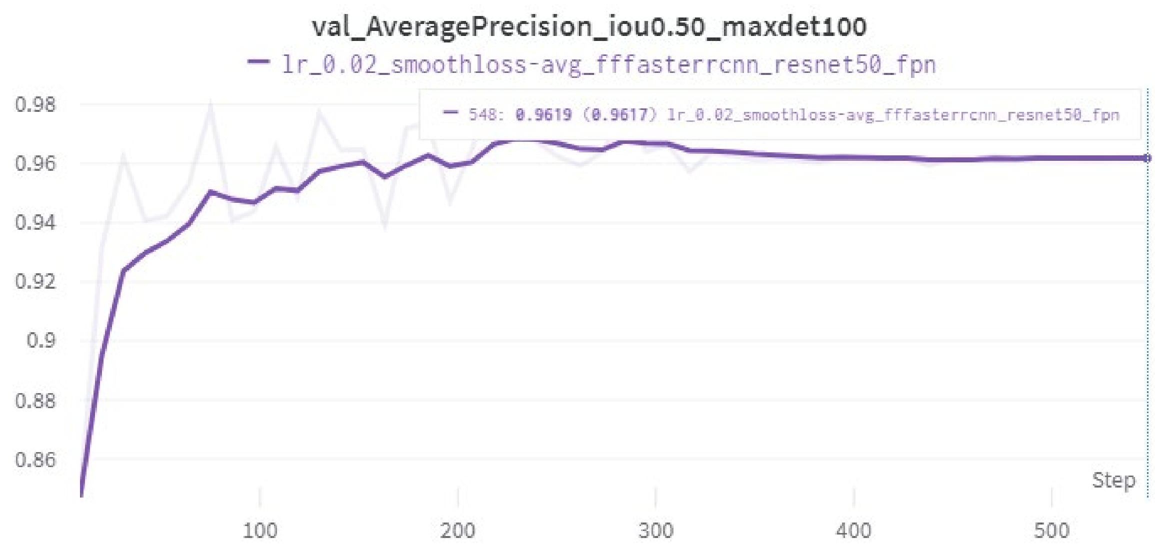


Figure 13. Depicts average precision (Y-axis) as training progressed (X-axis: steps). For task of counting a detection with over 50% overlap is classed as a good detection.

Hides / Slabs

In this stage we utilise instance segmentation metrics to measure the accuracy of the mask being inferred and accuracy of the class in a combined metric, AP which is calculated via commonly utilised code sourced from MS-COCO instance segmentation dataset. Our trained model achieved a score of 0.95 for detections with an IoU over 50% showing that not only can the model detect abalone with a high level of confidence, but it can accurately draw a mask to fit the abalone in an image (Figure 14). Figure 15 illustrates false positive rates whilst Figure 16 shows corresponding false negative rates.

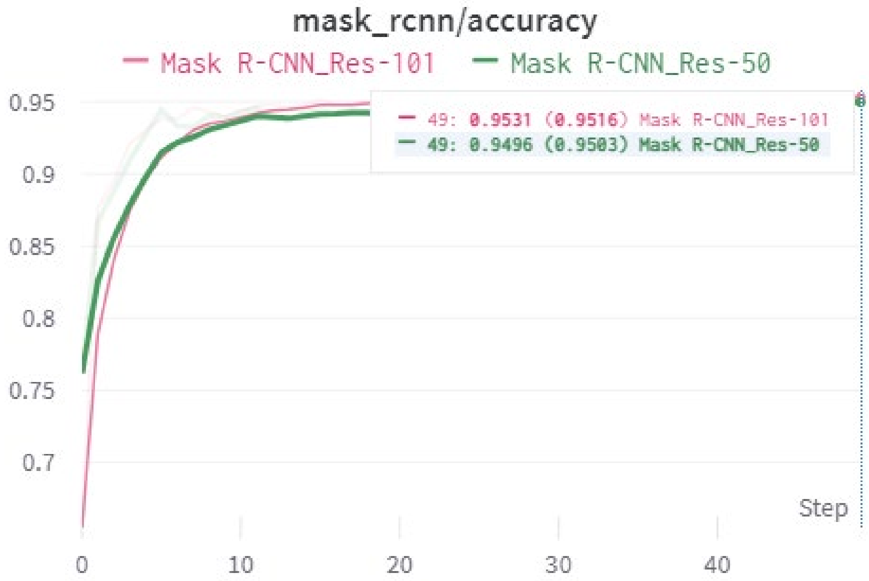


Figure 14. Training accuracy (y axis) for ResNet-50 versus Resnet-101 backbone variations (X-axis: steps). ResNet -50 was later deployed in the application.

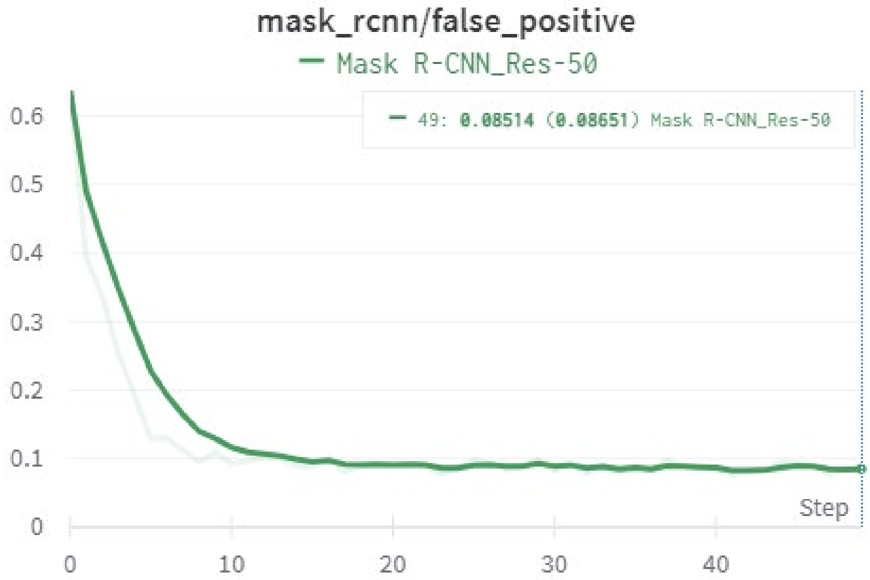


Figure 15. False positive rate (Y-axis) as training progresses (X-axis: steps).

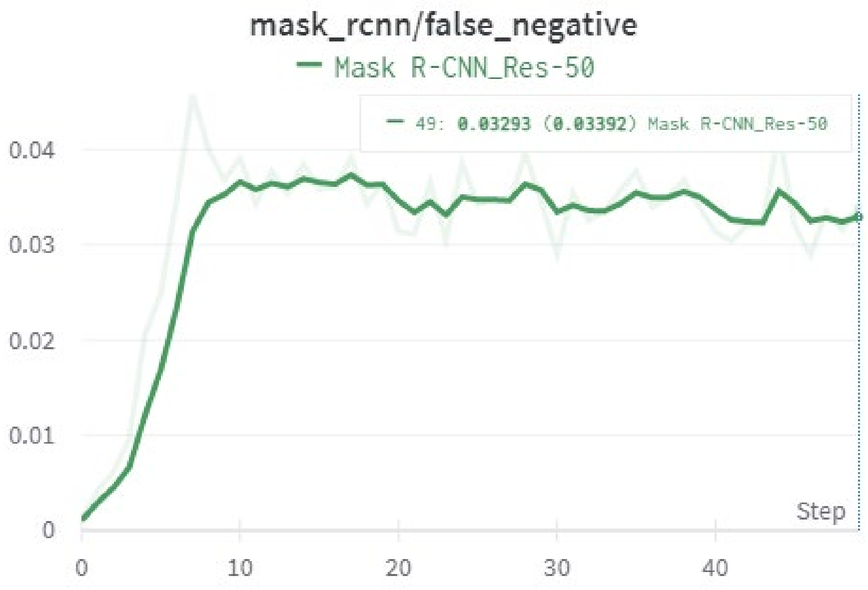


Figure 16. False negatives (Y-axis) as training progresses (X-axis: steps).

Practical whole system accuracy

For the nursery stage our ML results are directly transferrable to real world results as there is no necessity for scaling. Provided reasonable quality images can be used for input the system, it will perform well for early stage nursery abalone.

Where mm measurements are required for later conversion into a weight estimate there is a number of places where errors can be introduced that can be reduced provided the correct image capture methods are developed and adhered to (Standard Operating Procedures). Potential image capture errors include:

1. Image quality: Generally high resolution sharp and adequately lit images are required. Mid-low tier camera equipment and above is usually capable of taking suitable images for this system provided attention is taken to attend to these qualities. We found that suitable camera mounting (tripod) and fixed focus returned repeatable high quality images even with simple hand held, battery powered lighting.
2. Camera is perpendicular to the plane the abalone are on: This is especially important for accurately measuring abalone in different parts of the image. Angled images will skew measurements of both the automatic measurement object and individual Abalone. Related to this is distortions introduced by camera lens, this can be measured and accounted for by taking reference images with a grid pattern and running an algorithm to calculate the distortion, the inverse of this is then applied to the source images.
3. Abalone take up a suitable amount of the image: Extreme close ups or abalone too small/distant in an image will affect detection of abalone and measurement calculations. Some sample images along with suggested tiling selection has been provided in a help page in the program.

Implications

The main outcome of this project is vastly increased efficiency and improved accuracy in counting and measuring abalone in production systems which will be of benefit for the Australian abalone aquaculture industry.

Recommendations

Computer Hardware/Software:

Recommended minimum hardware would include the following and have been given a current cost estimate (April 2022) of $2,000. This is recommended for a single farm intermittent usage, for multiple sites it is recommended that CPU core count and RAM are increased. Best performance for GPU based deployments currently requires one of the more common Linux operating systems. It may be possible to achieve this in Windows however in our experience support for this is still too cutting edge for an onsite deployment.

- CPU: 6+ Core 3.5GHz (AMD or Intel)
- GPU: Nvidia RTX3060
- RAM: +16GB DDR4 non-ECC
- Storage: 128GB SSD/NVME
- Operating System: Ubuntu Server 20.04LTS

Image acquisition hardware:

We found a +20 Effective mega pixel camera to be more than suitable for both tasks. We utilised a 24-105mm lens and found this effective for all stages. If a camera is to be fixed to a frame for counting nursery stage (recommended) abalone, ensure that the lens will allow full framing of the slide in the image to be captured. We also recommend a fixed frame from measuring and counting abalone on the underside of hides, this will ensure accurate and repeatable measurements can be performed. We were able to acquire usable images with a tripod for the growout stage and would recommend a suitably rigid tripod system or perhaps there may be some better type of framing / gimbaling system that could be created for fast and accurate images for this stage.

Further development

A number of areas for further develop of this application and the technologies being harnessed by it have been identified, including:

AI Development: The most significant technology in use for this project is based upon Artificial Neural Networks which are an area of active research that is currently undergoing very rapid development. It should be expected that new research in this field will likely yield improvements in accuracy and efficiency. We also wanted to specifically point out the following areas as inspiration for further developing this system.

1. Utilisation of ML techniques for direct weight estimation: Currently the system is only inferring a max length of an abalone, which is later used for weight estimate. There has been research into using DL networks for directly inferring attributes such as orientation, occlusion and estimation of unseen parts of objects. The success in these areas shows potential for application of these ideas and techniques in developing technologies that could accurately infer an Abalone individual's weight in an image.
2. Utilisation of ML techniques for whole tank counts and estimated biomass. There have been some significant developments in 2D to 3D computation utilising ML based techniques. These technologies can infer a faces 3D contours from a single 2D image. Another technique is able to utilise multiple images of a scene to create a 3D version of the scene and localise objects within that scene. These technologies may have significant value in creating accurate whole tank and individual counts and weights.
3. Identification of food and scat for automated cleaning and feeding. Utilising the same technologies as this project has used, coverage of food and scat could be detected in images and used as input for determination of cleaning and feeding cycles. These could also be used as indicators of metabolic rate for populations of abalone and potential early warnings for disease / health issues in a population.

4. Real time distance and measurement plane detection, utilising technologies such as AR-Core real time measurement plane and distance could be saved at the time of image capture allowing for faster mobile image capture and potentially a more accurate end to end solution.
5. Increase accuracy and robustness of current system. Retraining the current networks with more images of more diverse growing conditions will make the system more robust to changes in lighting and appearance of the Abalone throughout their lifecycle. The current system was trained using images from two field trips and some additional images. Where the appearance of the abalone is different from those images that were captured the system will drop in accuracy.

Development in usability and utility: Whist the current system is expected to be robust and useful in its current state, there are a number of areas in where improvements would benefit the end user. Some of these are outlined below and it is expected that more suggestions will come from the adopters of this technology.

1. Making the system perform faster: Development of the current codebase with efficiency in mind should yield at least a 10x increase in speed. Whilst significant development time was taken to ensure this system operated in an acceptable timeframe, much more can be done.
2. Ability to load several tanks worth of images at once. Changes to the UI of this project could allow users to enter sets of images for tanks and have the averages calculated for each tank. This would allow the user to set up these after taking a few tanks worth of images and head back out to take more images.
3. Automatic tank detection. We could link a batching system to a set of QR codes or similar technology, this would allow the user to upload a large number of images and the system detect the tank for each set of images and format the output accordingly.
4. Automated integration with database system. This would allow the system to store the inferred information directly to a database system, cutting down on manually handling data.

Extension and Adoption

The project was communicated to the industry through this report. Periodic updates were provided to AAGA (via zoom) members throughout the project. A presentation explaining project development and a working demonstration was provided to AAGA members on April 4th, 2022. Further presentations will be given to AAGA at their invitation.

Going forward remote assistance will be provided with installation of the system on suitable computers or hardware in Cairns. In addition, supply of code or a docker container will be provided for self-installation.

It is likely that other aquaculture and fishing industries could also benefit with a similar approach for assessing stock of other species. Early stage discussions will benefit optimal tool development.

Project coverage

1. The future of abalone farming, 'This is Uni', JCU publication, Tianna Killoran, 27 July 2021.

https://www.jcu.edu.au/this-is-uni/science-and-technology/articles/the-future-of-abalone-farming

2. AI advancing for abalone. FRDC, Michelle Daw, 24 Nov 2022

https://www.frdc.com.au/ai-advancing-abalone

Project materials developed

As detailed above, a computer program was developed as part of this project. Screen grabs and detail are highlighted above.

Appendices

Guidelines for operating the AI system for counting and measuring Abalone in a farm environment

1. Introduction

This software utilises Artificial Intelligence (AI) to count and measure abalone from images. It is capable of counting abalone from when they first become opaque (approx 5 weeks) right through to harvesting, it is also able to measure the maximum (shell) length of individual abalone in images where a scale is supplied or a reference object is used. It is designed to be used to aid in the counting and measuring of abalone throughout their lifecycle in an aquaculture based setting.

This document is designed as a guide to help understand and therefore utilise this technology to better count and measure abalone. Due to variances in farm locations, farming techniques and reporting requirements, a single, detailed, precise set of Standard Operating Procedures cannot be determined. To better help you determine the best set of SOPs for your requirements we provide information on what we have found to be important for system accuracy and give examples of how we have captured images, used the system, and what we believe could be useful information based on our experience. **Ultimately to determine your SOPs you should start from what we suggest and experiment to determine what is most effective for your needs.**

We would also like to acknowledge that applying these cutting edge technologies in this application is a new and developing field, and there are still some areas where there is room for improvement. It should be pointed out that it may not always yield results that are expected under all circumstances and in applying this technology to your situation it should be assessed across the range of your usages as to suitability and accuracy.

Consistency over time is key to being able to build better farm information that can then be utilised not only for the abalone/tanks in the images but also for future reference to measure different combinations of genetics and growth factors. Proper determination of the number, location, and frequency of image capture is yet to be determined and we would expect that the fastest way to arrive at the best determination of this would be **to initiate the use of this technology by tasking capturing as much of the growth cycle of the abalone as possible, as often as possible, and then scaling back to a regime that provides what is required.**

Due to the speed at which this technology allows the capture of data, we would expect that this technology would allow the capture of far more information than is currently possible due to logistical, labour and financial restrictions, the better accumulation of farm data should allow for the development of better farming practices based upon better information.

2. Best practices

Due to this technology being based upon images of abalone, it is logical that better quality of images will provide better quality data. Of equal importance is ensuring that the data inferred from images is attributed to the correct abalone/tank/cohort, for this reason the system has been developed with a batch based approach. Our recommendations are to capture images in sets of batches that are of similar age and growth and use a similar regime of image capture. As an example capturing 5 images from the same locations, and in the same order from the same age group in a set of tanks allows the correct determination of which images are from which tanks by the simple annotation of the order of the tanks image capture. This also allows the person assessing the data returned from the system to be processed and interrogated in a meaningful manner.

2.1. Image Acquisition

The camera is at the heart of this technology and gaining a working knowledge of image capture and identification of causes of image issues is essential. Whilst we provide a camera recommendation later in this document (4.1) and ease of use was a determining factor in our choice, **we strongly encourage that time be set aside to become familiar with whatever device** is procured. A reasonable level of skill will allow the best images to be captured under a range of lighting and physical conditions as would naturally be present in an aquaculture based setting.

We strongly suggest that some time and materials be invested in the design and manufacture of custom frames/jigs or devices that allow the easier and more consistent capture of images as this will be a determining factor in the speed of image acquisition and the quality of images and therefore the data inferred by the system. Initially a tripod is probably the best way to become familiar with what is appropriate. It is our experience that hand held images are not suitable for the system.

Any device designed / manufactured for use should be designed with the following physical characteristics as well as mobility, ease of use and speed of use;

1. The device should be capable of ensuring the camera is perpendicular to the plane of measurement (slide, slab tank floor etc.)
2. The device should be capable of ensuring the camera is at a repeatable distance from the plane of measurement.
3. An image is able to be taken without disturbing or moving the camera.
4. Any vibration/swing or movement of the frame is quickly dampened or dissipated before the image is captured.

All images will require the entirety of the shot to be in sharp focus, ideally without strong reflection or distortions such as those that may be caused by disturbance of the surface of the water. Any movement of the camera will create blur and this is highly undesirable. Where height from image plane (abalone) is consistent it will be advantageous to set the focus thereby eliminating the camera having to autofocus for each image.

2.1.1. Nursery

Being able to take clear, high resolution images is essential for counting abalone, especially when counting abalone at their smallest. It is suggested that a frame that allows the easy placement and removal of the slides and the consistent placement of the camera be utilised. It may also be advantageous to have some sort of screen to reduce glare / reflections from the surface of the slide and provide some shelter for the abalone. We found that in a team of 2 we are able to capture approx 700 images (2 per slide) in a day including some movement between tanks and removal of slides from the baskets. At this speed we were able to photograph the abalone (5-7 weeks) with minimal observed stress even during hot weather.

For our camera to be able to fully frame the slide we required approximately 500mm from the slide to the camera mount. To aid in tank, basket, and slide identification we started from the water inlet end of the tanks, took images of every slide in every second basket starting with the slide closest to the water inlet and transferred the images after each tank. Where images are being taken at low focal length (close to camera) and the area is bright it may be advantageous to adjust camera settings for higher f/stop settings as this will allow a deeper portion of the image to be in focus, i.e. the edges and the centre are in focus.

2.1.2. Hides

Key to determination of the best technique for counting and measuring abalone in hides is the hide construction. Where hides are tiles on removable frames, it may be simple to set up a frame/tripod so that it is above a flat surface that the hide frame (upside down) can be slid under the camera's field of view to capture images of a few tiles at a time. Deep hides may require the removal or reduction of the water level. Obviously, the system can only count and measure abalone that are visible in the image. In situations where there is a constant ratio between visible and unseen abalone the system may still be able to be used to provide data before scaling the results.

Example of how weaner stage abalone on hides were photographed for this project.

Size of the abalone is the determining factor for how far away the camera should be from the tiles/measurement plane. The system works best at counting when individual abalone can be clearly identified in the image, for optimal measurements for the max length of the abalone shell it should be clearly visible where one abalone starts and another ends. A quick zoom in/out on an image and some common sense should get you in the ballpark of what is required. Further use of the system will allow you to fine-tune the distance.

Where images are being taken in a bright area a similar approach to f/stop and camera settings may be taken to the nursery stage to ensure all of the image is in focus (i.e. optimise for higher f/stop). Where image capture is in a dark area a balance between depth of field and the exposure length may need to be sought, however the image will be more susceptible to motion blur at longer exposure times. **Where lighting is required to capture a suitable image, it should be coming from a shallow angle to the side of the scene**, this will minimise reflections.

We found that with two people we could quickly and efficiently photograph frames, one person would be handling the frames and another operating the camera, a foot-switch operated camera may be able to reduce this to a single person operation, however using two people to handle the frames may be better for sustained image capture as this process places physical stress on the person moving the frames.

2.1.3. Grow-out

As with the hides, the physical setup of the tanks used for grow-out will determine the best technique and supporting equipment for capturing images. Our experience was mainly with modern, shallow, slab tanks under shade cloth or housed in dark sheds. We utilised a tripod to stabilise the camera and took a consistent number of images in each tank starting from the same locations in each tank in the same order. We utilised side-lighting to illuminate the scene and minimise reflections. Some of the issues we had were locating the tripod feet and some inconsistency in the camera height and angle, for these reasons we do not believe this is the best way to capture images for this stage. We would suggest some sort of portable, two person operated frame that has leg(s) to maintain a consistent height from the sides of the tank (and therefore the tank floor), with a camera on the centre of the frame. The frame could also be used to locate the side lighting allowing for the whole thing to be easily mobile with 2 people, and allow for fast and efficient, and consistent capture of images.

Example of how grow-out abalone in slab tanks were photographed for this project.

Where there is strong overhead lighting (such as thin shade cloth on a bright day), it may be best to drain the slab tank, experiment with glare removing lens filters or simply capture images at a time when the light is more suitable. For darker scenes some experimentation with aperture and exposure length may need to be done to determine the best balance between the available light, the depth of field required for the whole image to be in focus and potential motion blur. The best solution will take into account the sturyness of the frame and the cameras capabilities.

2.2. Use of software

The software developed to house the AI technology used to count and measure abalone has constraints that when properly met will allow the effective counting and measuring of abalone. There are some user adjustable settings that allow the system to perform under a range of sizes and numbers of abalone in an image. **When thinking in terms of the software it is best to think of the abalone size as being measured with reference to the size of the image and/or the number of pixels that are used to represent an abalone in the image.** The scale reference or scale factor is used to relate this back to real world measurements. There are many potential sources for error, however with the aforementioned image capture guide and proper usage of the software, these can be reduced to a manageable level.

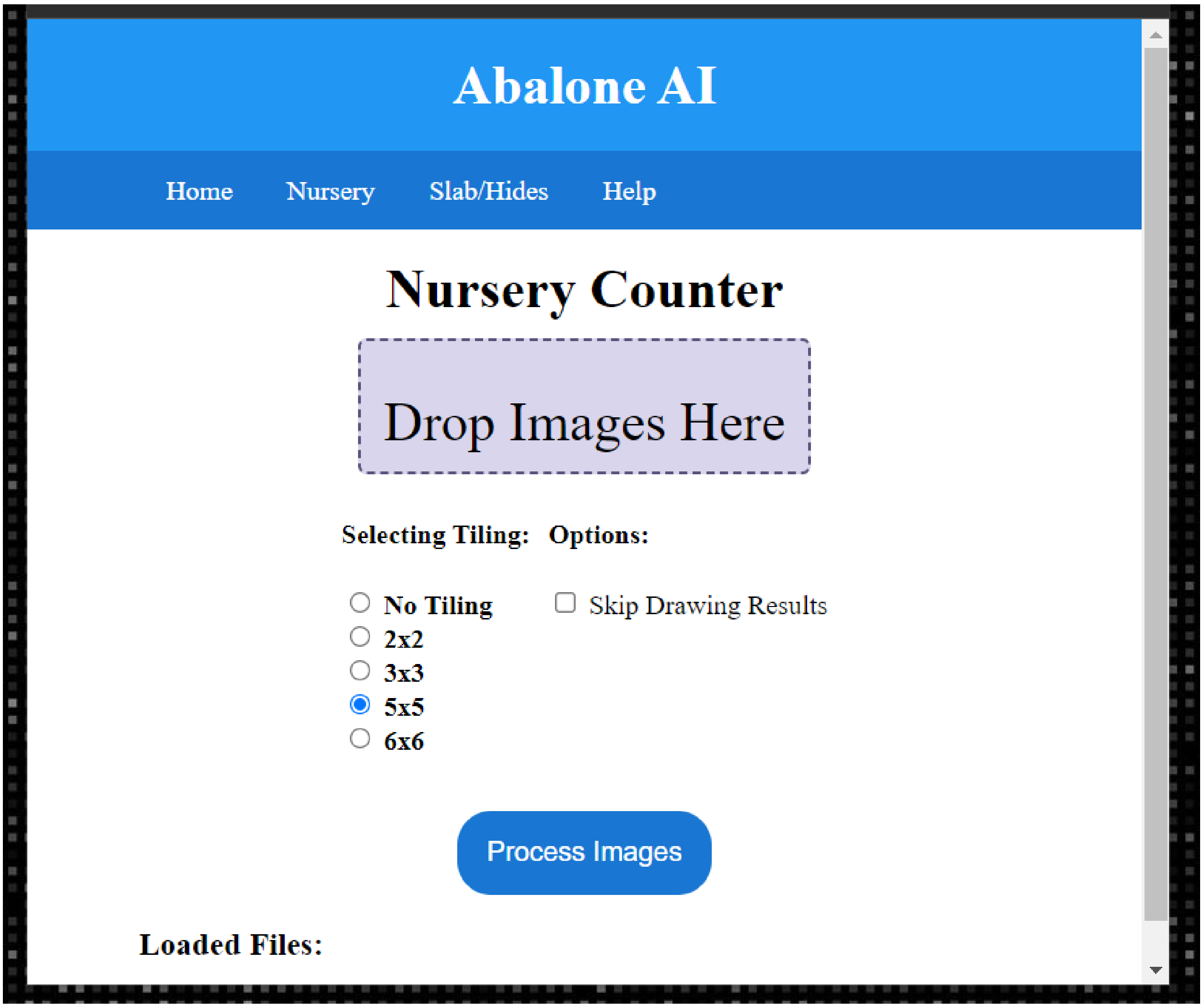
Both the AI and the pre/post processing requirements of the system are very large and the time taken for these to be performed are non-trivial. Don't expect the system to be able to provide instant inference of abalone counts and max length. In the hides/slab stages computation time is dependent on the tiling factor and the number of abalone in the image. **More abalone and higher tiling regimes will take significantly longer to process**. Generally, the larger the abalone in the image (the more pixels per abalone) the better the max shell measurement will be. As a general rule, **if you can see it well in an image the system should also be able to.**

With each abalone growth stage, we provide feedback to the user as to what the AI has found in the image, this is overlaid onto a set of the source images and the user is able to use this to gauge system performance. You should use this as a guide to what works with respect to image quality and tiling strategy, you should also check this frequently to ensure the system is providing quality data.Due to the significant time required to process images the system has been designed to support the usage of batched image processing, this significantly reduces the time required for the user to submit the images and store the returned data. To prevent excessive wait times and provide sane use of resources the largest batchsize has a soft limit of approx 1GB of images. This is over 100 high quality, high resolution(6000x4000px) JPEG images. Data is returned in the form of a ".CSV" file which is a standard open format and able to be utilised by almost all programs that use tabular data including Microsoft Excel and Google Sheets. Each line of data contains the image it was sourced from as well as the data inferred from the image as well as some debugging information. Batching methodology is left to the user to best determine and should depend on how fine grained the information for each batch is required. I.e. For normal farm size/count monitoring it may make sense to subsample a number of tanks in a cohort and process these in a batch (20 slabs at 5 images per slab) density, size and count can be inferred for the whole cohort. Should further information be required for the batch further processing of the data can easily yield the statistics on a per slab basis and these analysed to determine over/under performing slabs. In a different use case such as when performing feed trials, each slab or group of slabs with the same conditions can be photographed and processed as a separate batch, the batch statistics can then be compared to assess the feed trials and these can be compared over time to monitor the trials more intensively without physical disturbance.

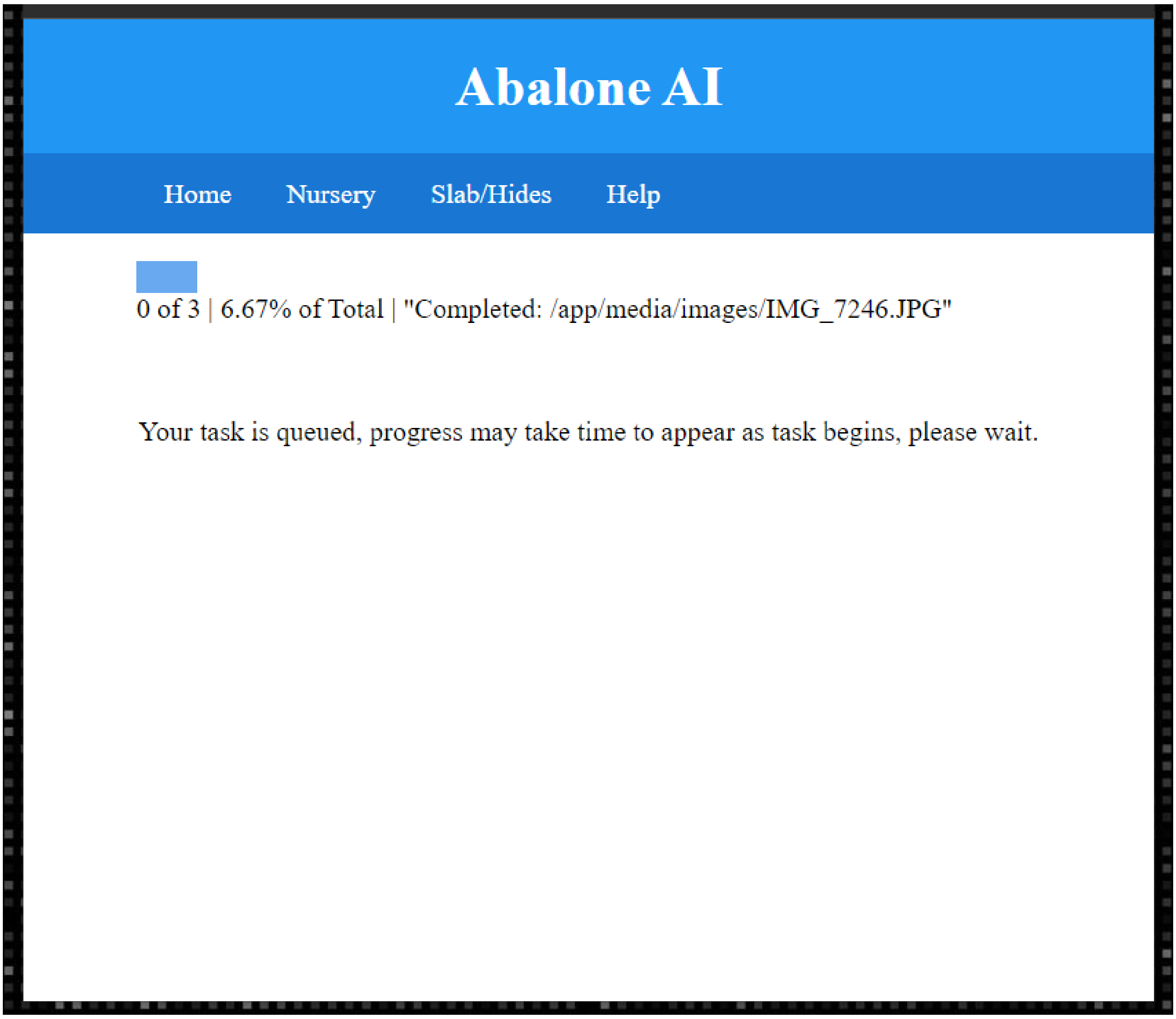
2.2.1. Nursery

The abalone spat need to be opaque enough to be visible in the image and not overly obscured by debris/weed. High levels of similar looking debris or species may confuse the system and lead to inaccurate results although the system has been found to perform well with rejecting non-target species such as Copepods.

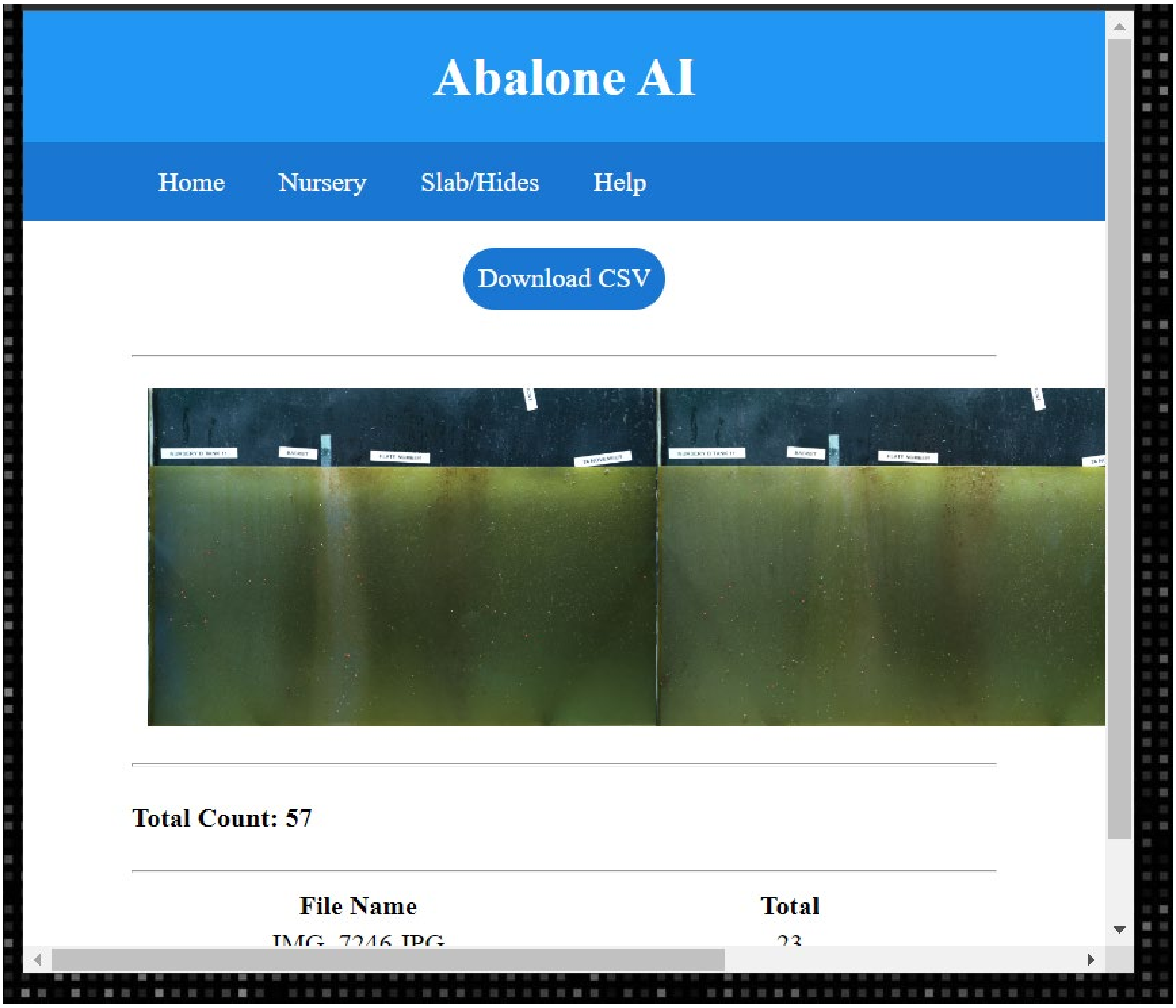
For this stage it is best if the **abalone are at least 30-50 pixels across and resemble a white disk**, for larger abalone better results may be obtained using the hides/slabs section. Unless you are using a very low(<5MP) or high(+40MP) resolution camera, you should use the preselected 5x5 tiling strategy, the system is untested with cameras outside the 20-40MP range and may perform badly without additional image processing.



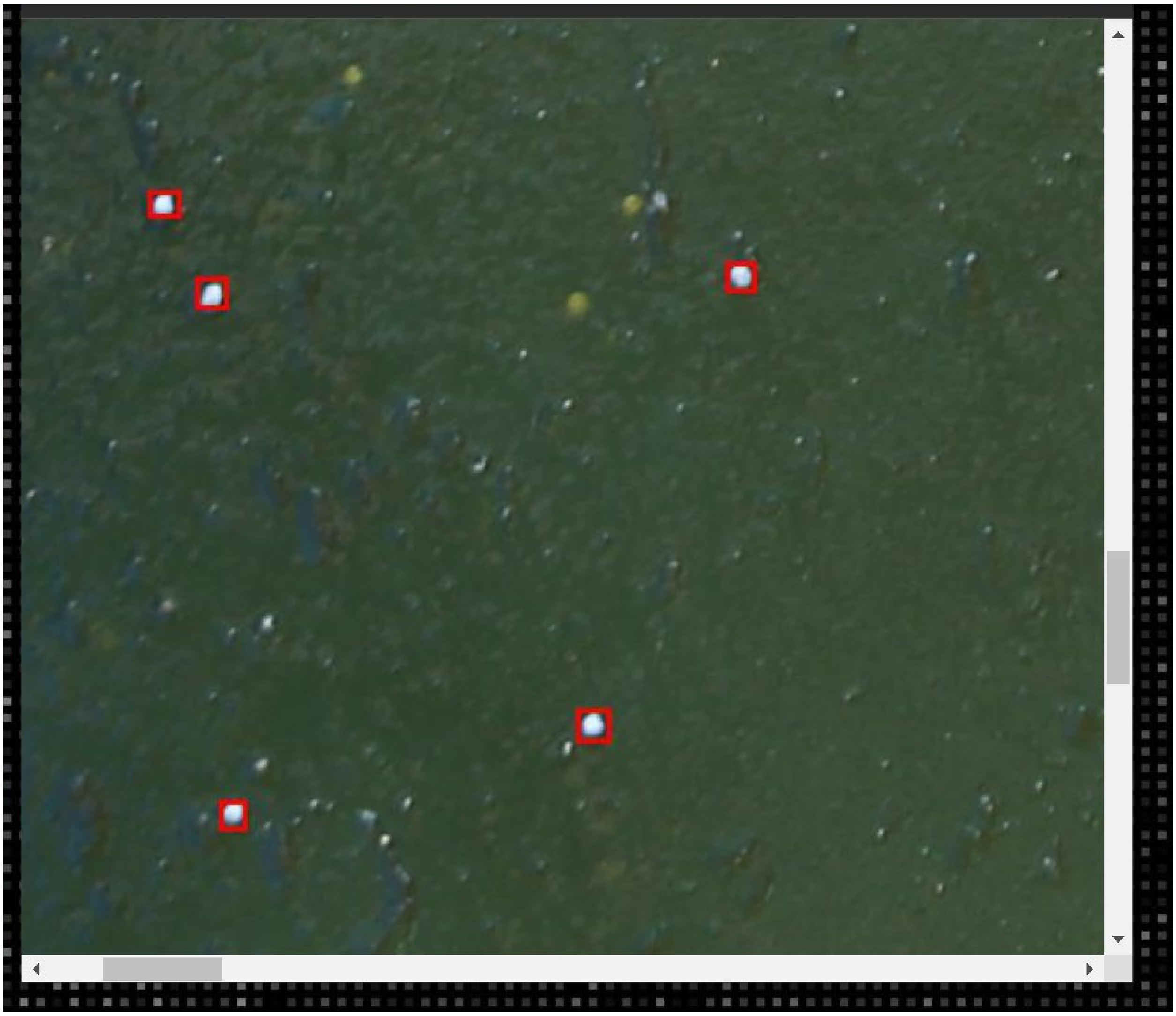
Usage is as simple as dragging and dropping the image you wish to process in a batch and pressing the "Process Images" button.



The processing page will keep the user notified of the current progress until the process has been completed.



The results are returned to the user in various forms, A summary is provided on the page as well as a link to the compiled results in a ".CSV" file. For the user to be able to assess how well the system has performed on the images supplied the detections are overlaid on the images and the user can inspect the images by clicking on them and then downloading or viewing in the browser.



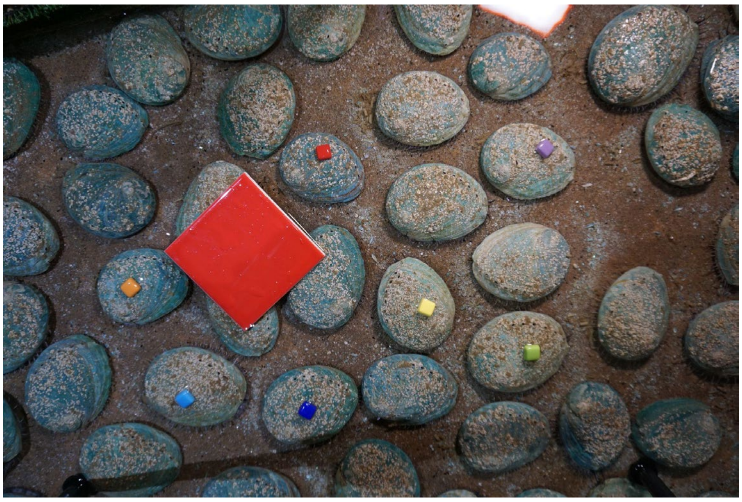
An example of a detected abalone on a nursery slide.

2.2.2. Hides / Slabs

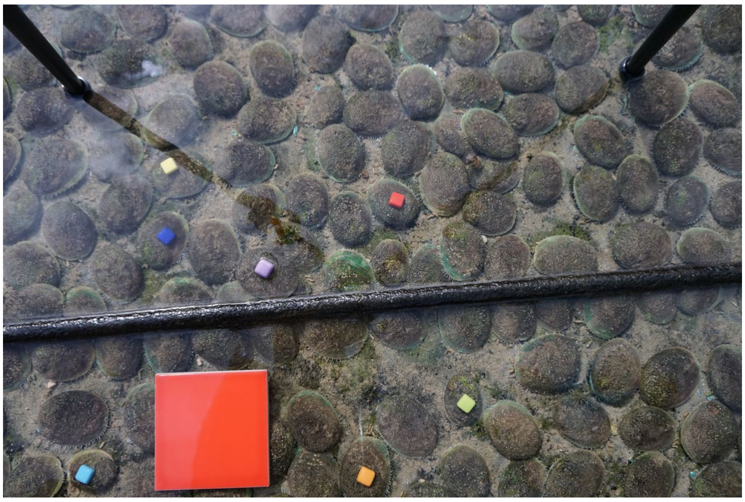
For abalone that have matured to have a typical oval shape and are no longer white, individual max shell length and total count for an image can be determined. There is no hard limit on the minimum or maximum physical shell size for this system. We have found that for reasonably densely packed abalone, **images containing approx 50-600 abalone will perform the best** provided the correct tiling strategy is selected.

2.2.2.1. Tiling strategy

By selecting the correct tiling strategy, the system can provide measurements and counts over a wide range of abalone images. Generally, abalone that are large in the image will give better max shell length measurements than those that cover a smaller number of pixels. Below is a set of example images to help select the correct tiling strategy.

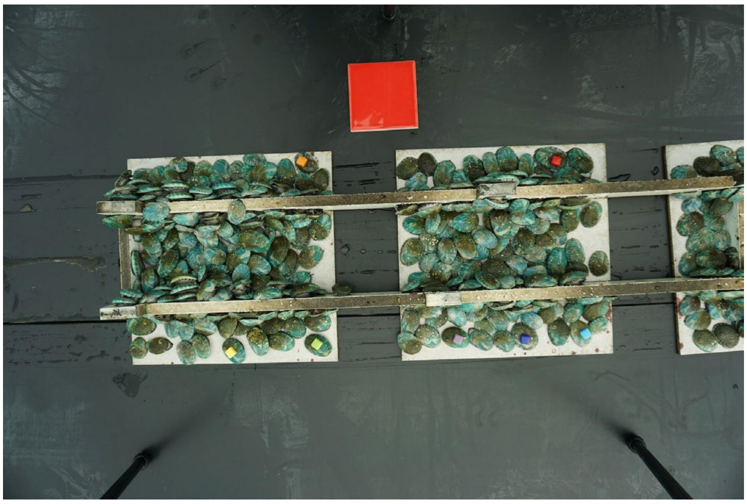


With few (approx. 30) large abalone a 1x1 tiling strategy can be employed.

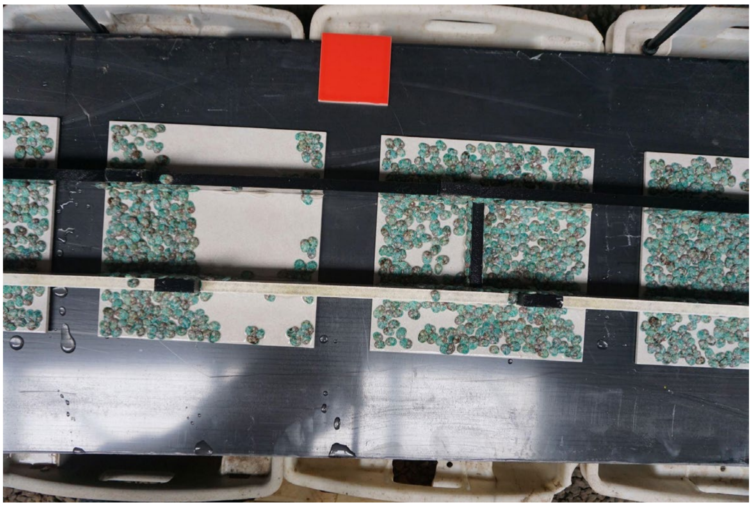


When the image contains approx. 100 densely packed abalone a 2x2 tiling strategy will give the best results.

Images where the abalone are smaller yet and may be approaching 400-500 densely packed in an image. For the most accurate max shell lengths 3x3 is the highest tiling strategy to be used on a 20mp camera.



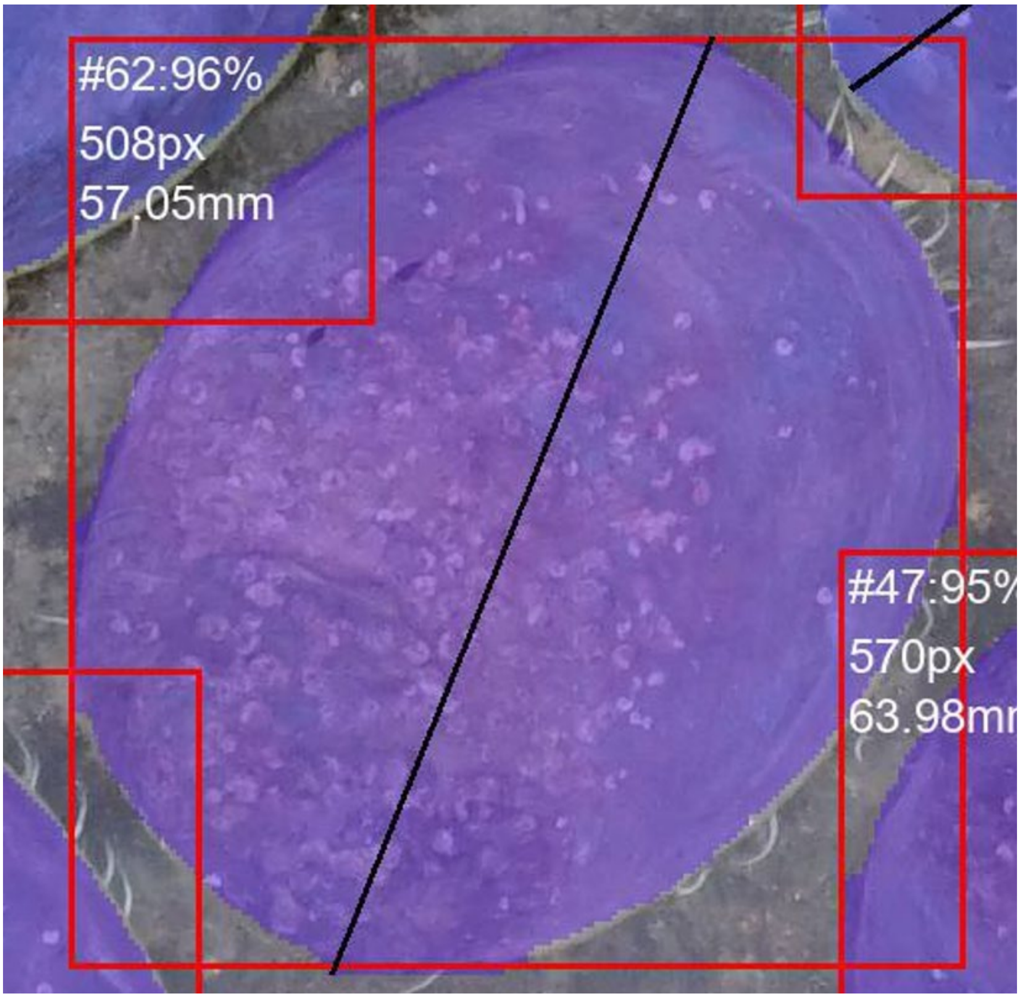
A 5x5 tiling strategy will allow the counting and approximate measuring of abalone in the 1000 densely packed count ranges. At this perspective reasonable attention needs to be taken to ensure sharp images are being taken if max shell length is to be measured.



6x6 tiling strategy is at the limits of what a 20MP camera can capture. For usable results to be returned by the system much care needs to be taken in capturing an image that is sharp across the entire image. At this tiling strategy a 50-100MP camera will be able to capture a full tank width at sufficient detail for abalone over 10mm to be counted, accurate max shell length measurements have not been tested at this resolution/tiling strategy.

Processing times for each image will increase with the increases in the number of abalone and the tiling strategy.

After inference the system will return a subset of the images with what the system has detected overlaid on the image. This will show all the abalone detected, whether the abalone max shell length was measured (obscured abalone are not used for length measurements). Each abalone will have a box drawn around it, its number (used in the ".CSV" file), its pixel length and the actual length (if scale reference is used). The mask of the outline of the abalone is overlaid, this is red if the abalone has been determined to be unsuitable for length measurement and blue if it is used for length measurement. A black line is drawn where the max length was determined.



An example of a detected and measured abalone.

A full image with detections overlaid.

2.2.2.2. *Scale detection*

To convert the image scale to a real-life scale there are two approaches.

1. If a stand or jig is used with consistent (<5% change) distance from camera to abalone the image scale can be calculated and entered into the system in the "Override Pixels per Metric (mm)" and the checkbox selected. This will allow for faster image capture and more consistent measurement of abalone provided a fixed distance from abalone.
2. The use of a red square (Bunnings tile pictured 97mm) can allow the system to estimate the scale of the image on the plane the square is placed. This system allows the camera height to be varied to best suit the abalone size in the image. Care should be taken to ensure the tile is flat on the plane to be measured, and that this plane is at the height of the shell max length. Usually this will mean the tile needs to be located on the floor of the slab for the most accurate measurements to be taken (i.e. the top of the tile is the height of the bottom of the shell).

Abalone counting and maximum shell length measurement image input page. Here tiling regime and scale overrides can be selected.

3. Hardware/OS for AI system

This system has been developed using cutting edge AI and requires substantial processing power to calculate the results. It has been designed to be accessible via a web browser on the local network. We strongly recommend a dedicated machine be used for this system as this will give the best user experience. To allow for more consistent deployment it has been designed to run inside a docker container and for best performance we specify a *nix operating system. We have found Ubuntu Server to be suitable.

3.1. Minimum hardware requirements

CPU: 4 core 3 GHz processor X86/64 (Intel/AMD)

RAM: 32 Gb

GPU: Nvidia 8GB

Storage: 250Gb SSD

3.2. Recommended Hardware (medium usage / single farm)

CPU: 8 core 3 GHz processor X86/64 (Intel/AMD)

RAM: 32Gb

GPU: Nvidia 8GB

Storage: 512Gb SSD (NVME M.2)

3.3. Recommended Hardware (Heavy usage / Multiple farm)

CPU: 12 core +4 GHz processor X86/64 (Intel/AMD)

RAM: 64Gb

GPU: Nvidia 12GB

Storage: 1Tb SSD (NVME M.2)

3.4. Prebuilt and installed systems

Turnkey, prebuilt systems are available from kurt@kurtsch.com.au. These systems can be built to usage requirements.

4. Notes on Camera selection

There are many variables when determining the best camera for the task of Counting and measuring abalone. We have a suggestion (4.1) for an entry level camera that would be suitable for testing the efficacy of the abalone AI system. Due to the harsh conditions on an abalone farm (salt water) the life expectancy of such devices may be shorter than usually expected. For longer term usage and more potential automation capabilities it is expected that industrial cameras and housings may be a better option.

In our experience there are a few qualities that we believe are most important for utilising images for AI based measuring and counting purposes.

1. Camera must be able to capture a sufficient number of pixels to be useful over a range of scales / distances. Especially important for the nursery counting if a whole slide is to be captured at once. Minimum 20 MP.
2. Camera / lens setup should be able to capture an image with minimum geometric distortion at minimum zoom(wide angle). As most lenses have the worst distortion effects at the widest angle settings it is advantageous for the camera to be able to correct for this itself or some other form of undistortion needs to be applied to each image as a part of the pre-processing. The recommended camera has approx 3.8% distortion at the widest angle that is corrected by the cameras software (JPEG) to less than 0.1%.
3. Camera / lens Low light capabilities are important especially for imaging abalone under shade cloth without additional lighting. The suggested camera has excellent low light response at wide angle with a maximum aperture of (f/1.8).

4.1. Recommended entry level camera

The recommended camera is a Sony Cyber-Shot DSC-RX100 II or higher, an overview of its performance can be found at Imaging-resources.com. We found this resource to be a good starting point for assessing some of the most important factors in camera performance.

References

He, K., Zhang, X., Ren, S. and Sun, J. (2016), 'Deep Residual Learning for Image Recognition', *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770-778, doi: 10.1109/CVPR.2016.90.

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. and Adam, H. (2017), 'MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications', *arxiv:1704.04861*.

Lin, T-Y., Dollár, P., Girshick, R., He, K., Hariharan, B. and Belongie, S. (2016), 'Feature Pyramid Networks for Object Detection', *arxiv.1612.03144.*

Ren, S., He, K., Girshick, R. and Sun, J. (2015), 'Faster R-CNN: towards real-time object detection with region proposal networks', In *Proceedings of the 28th International Conference on Neural Information Processing Systems – Volume 1*, pp. 91-99, MIT Press, Cambridge.