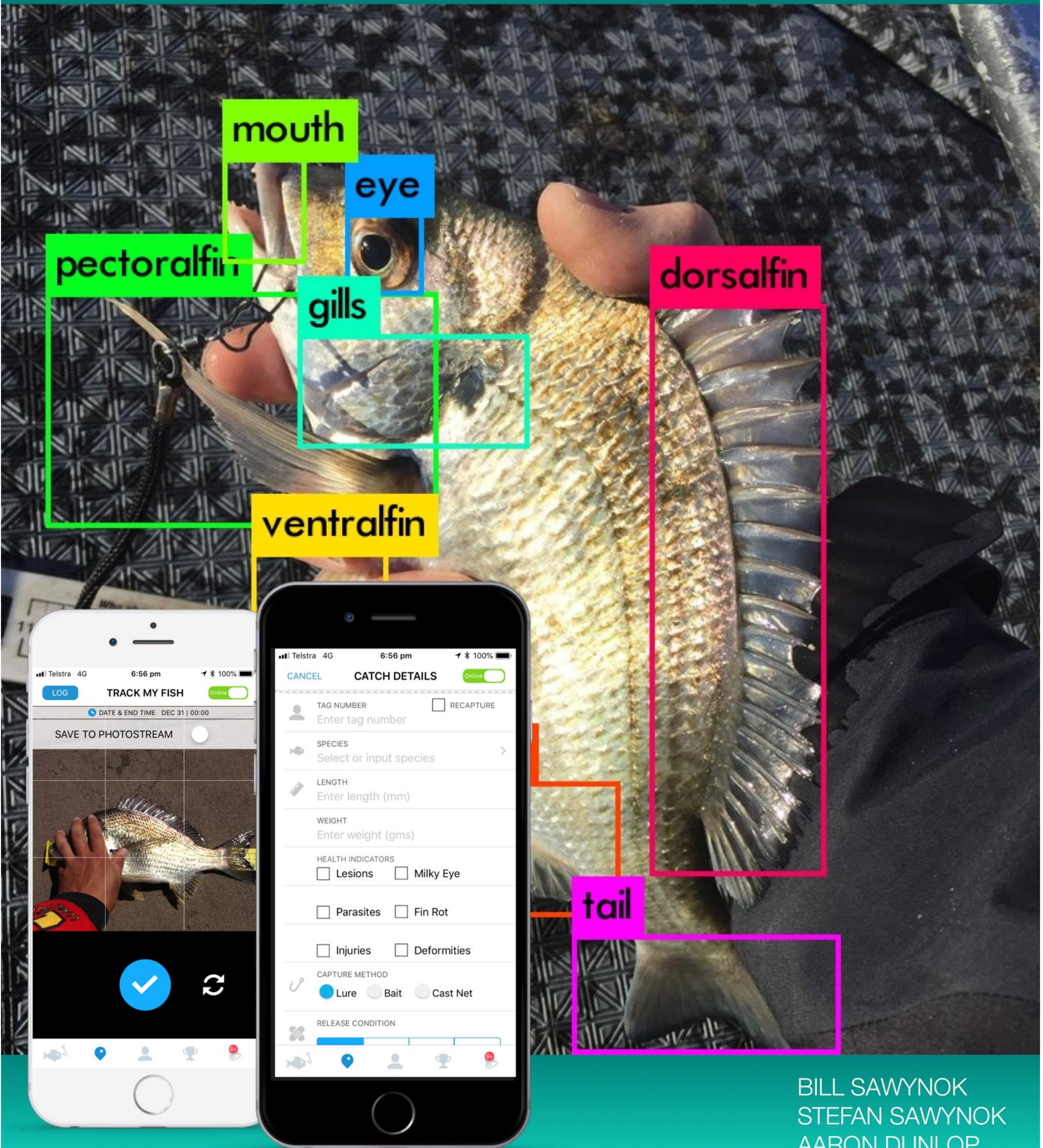
NEW TOOLS TO ASSESS VISUAL FISH HEALTH



BILL SAWYNOK
STEFAN SAWYNOK
AARON DUNLOP

FRDC
Gladstone Healthy Harbour Partnership
infofish AUSTRALIA

ISBN 978-0-6484281-0-7

New tools to assess visual fish health
FRDC project 2017-141
2018

Disclaimer
The authors do not warrant that the information in this document is free from errors or omissions. The authors do not accept any form of liability, be it contractual, tortious, or otherwise, for the contents of this document or for any consequences arising from its use or any reliance placed upon it. The information, opinions and advice contained in this document may not relate, or be relevant, to a reader's particular circumstances. Opinions expressed by the authors are the individual opinions expressed by those persons and are not necessarily those of the publisher, research provider or the FRDC.

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: Bill Sawynok
Address: PO Box 9793
Frenchville Qld 4701
Phone:
Fax: 07 4928 6133
Email: bill@info-fish.net

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

Contents

Figures

Tables

Acknowledgements

Funding for this project was provided by the Fisheries Research and Development Corporation, Gladstone Healthy Harbour Partnership and Infofish Australia Pty Ltd.

In undertaking the project, the following contributions are acknowledged.

- The review of the project milestone reports and the draft final report by the Gladstone Healthy Harbour Partnership's Steering Committee and its Independent Scientific Panel
- The provision of photos by the Central Queensland University Fisheries Team
- The provision of photos from tournaments around Australia by the ABT Fishing Tournament circuit
- The perseverance of Phoenix Sawynok in reviewing thousands of fish images
- The assistance of Johnny Mitchell and Daniel Powell in the collection of field photos

Executive Summary

In 2018 Infofish Australia Pty Ltd undertook a trial in Gladstone Harbour using machine learning tools to assess photos for fish health issues. The project was commissioned by the Fisheries Research and Development Corporation (FRDC) and the Gladstone Healthy Harbour Partnership (GHHP) to evaluate a number of tools for the visual assessment of fish health.

The objectives of the project were:

1. To deploy tools to automate data collection and assessment of fish health using data collected in Gladstone Harbour as a trial.
2. To undertake structured data collection of fish samples using Gladstone Harbour Partnership's reporting zones and the Boyne Tannum HookUp fishing competition.
3. To evaluate the potential to adapt the methods developed to monitor fish health in other estuaries and ports in Australia.

The Infofish Trackmyfish phone app was used to take photos of both sides of a fish with samples collected from 13 GHHP reporting zones, at the Boyne Tannum HookUp (BTHU) fishing competition, by the Central Queensland University in Gladstone Harbour and from ABT fishing tournaments around Australia. The target species were Yellowfin and Pikey Bream however samples were collected from a range of other species. The fishing competition version of the Trackmyfish phone app has had a high rate of uptake with around 40 competitions around Australia having used or will use the app in the coming months. This provides new opportunities to collect fish images to assess health issues.

Following a worldwide search of Object Detection algorithms TensorFlow Object Detection API and Yolo version 2 and 3 (You only look once) were selected for further evaluation. The technologies were evaluated based on their training ability, accuracy and scalability, and useability in the real world. In the final evaluation Yolo, with its guided process to training, ease of additional human validation and built in statistical assessment of trained models was found to be superior.

The training of the machine learning models was focused on Bream and carried out in two parts. Initial training was to recognise fish parts such as fins, tail, gills, eyes and mouth and fish health issues such as fin and tail damage, wounds and "redness" (eg lesions, scale damage). Images were assessed using a batch process and all health factors were assessed simultaneously. Results were then converted to a value of 0 (not detected) or 1 (detected).

There was a total of 1,242 images assessed. Machine and human agreement levels ranged from 50%-86% for fin splitting, 60-93% for tail splitting, 78-93% for tail damage, while the wound model was unsuccessful in all instances. The CQU samples were also assessed for "redness". There were 58 images assessed with an 86% agreement between machine and human assessment. Images continue to be collected and those results will improve with more images for the training models, particularly for species other than Bream and for health issues where there were few images eg milky eyes.

Based on the training models developed to date the results from the human and machine assessments were acceptable. With more images and further development of training models the results will continue to improve and be applied to an ever-growing number of species and issues.

While the objectives were to assess new tools some additional activities were undertaken to provide a more complete picture of factors around health issues and provide guidance to GHHP for the development of fish health indicators. Length-weight data were collected at the BTHU from 2003-2018 and this provided the opportunity to assess fish condition for Bream. Mean and median relative fish condition was calculated for each Bream species for each of the years.

A relative fish condition factor of 1 is considered to be good for an individual fish or a sample from the population with less than 1 being poor to moderate and more than 1 being good condition. The median values for both species ranged from 0.99-1.03 for all years from 2003-2018 indicating relatively stable fish conditions over those years.

Fish sampled in the GHHP zones were tagged. Recapture data from those tagged fish and historic Suntag tagging and recapture data were assessed for fish movement to provide an insight into a potential vector for propagation of health issues.

The movement of fish recaptured in the first year after tagging was calculated as this was the most likely time that fish health issues could be propagated. Fish that were recaptured within 2km of where tagged were considered to be resident with 58-89% being recaptured locally and 0-9% recaptured over 20km from where tagged. This suggests that health issues are more likely to be propagated locally initially but with sufficient numbers moving to eventually spread any health issue that may be transferable.

The project has demonstrated that machine learning technology can be applied to assess visual health issues. The tools that have been used will continue to evolve and improve and are likely to be used in other fisheries areas beside health and in other disciplines.

Using fishing competitions, for the first time, there is the possibility of collecting large amounts of data on visual fish health and injury issues in a cost effective way from around Australia. It is likely that a broad range of fishing competitions will see the benefit of having this data from their events and initial discussions with event organisers have confirmed that.

Taking this approach is also much more likely to be successful as it is a bottom up approach that is already gaining acceptance from fishing competitions. A top down approach through the various fishing bureaucracies is unlikely to work as it would be hard to get commitment, take a long time to implement, be constrained by a range of protocols and would be inordinately costly.

For GHHP it is recommended that it consider using relative fish condition as one of its measures of fish health. While the incidence of lesions in the fish sampled was not high it would be an effective measure as historically there have been years when “redspot” lesions were common. Continuing fish deaths and injuries from fish spilling from Lake Awoonga into the Boyne River also need to be considered.

For FRDC it is recommended that there be further development of the use of this technology, initially in developing more robust models and extending the range of species and issues that can be identified. Infofish is already using this in conjunction with other new technology such as a BioSonics echo sounder that can count the number of fish and there are significant other opportunities to be explored.

Keywords: Fish health, machine learning, Gladstone Harbour, Boyne Tannum HookUp (BTHU), Gladstone Healthy Harbour Partnership (GHHP), Yellowfin Bream, Pikey Bream

1. Introduction

In early 2011 major flooding affected the Gladstone area resulting in the spilling of Lake Awoonga on the Boyne River. That resulted in the release of over 30,000 Barramundi into the river and subsequently into Gladstone Harbour and adjoining waterways (Dennis et al 2016).

Also in 2011 there were major industrial developments around the harbour including the construction of 3 LNG plants on Curtis Island on the northern side of the harbour. These developments resulted in a dredging program for the western basin with the removal of 22 million cubic metres of dredged material (GPC 2017).

Coinciding with those events in 2011, a range of fish health issues emerged that raised public concern and resulted in the closure of Gladstone Harbour to all fishing in September-October 2011.

As a result, the Gladstone Healthy Harbour Partnership (GHHP) was established in 2012 to assess the health of Gladstone Harbour. The GHHP is responsible for producing an annual report card on the health of the harbour that includes environmental, social, cultural and economic health indicators. Fish recruitment and health were identified as important environmental health indicators.

The initial report cards focused on fish recruitment however in 2017 GHHP decided to include fish health. GHHP called for proposals to assess fish health following some preliminary review work. While a proposal submitted by Infofish Australia was unsuccessful it was invited to submit a revised proposal that focused on better automated data collection (using photographs and video) and assessment (artificial intelligence and machine learning algorithms).

It was recognised that these innovations had the potential to introduce new technologies into the visual assessment of fish health and other fisheries areas. It was considered that these technologies needed to be evaluated and this project was developed in response to that need.

While not part of the objectives of this project, in the broader context of fish health issues in the Gladstone area, particularly from a public perspective, the continued incidence of dead and sick Barramundi in the Boyne River needs to be considered in the development of fish health indicators for the report card.

Barramundi stocks in the Boyne River increased on a number of occasions since 2011 through fish that spilled from Lake Awoonga. Many of the fish that spilled died or were injured, some severely, from going over the dam spillway. Many of the injuries were scale loss from abrasion from the concrete spillway and from the rocks at the bottom of the spillway. The scale loss often resulted in subsequent infection. Most years there have been reports of dead Barramundi, mainly larger fish over 800mm. Apart from 2011 the incidences have been largely confined to the Boyne River.

Infofish received and documented reports from fishers of dead Barramundi in the Boyne River each year and in Jan 2014 (Sawynok et al 2014) undertook a survey in the river to document the numbers of dead fish. Many of those reports included photos of the dead or dying fish.

Based on documented reports by fishers of dead, dying or sick Barramundi the following estimates in Table 1 were made of the numbers in the Boyne River each year since 2011. These estimates are crude however they do provide some sense of the scale of the issue.

Figure 1 shows 2 recently dead Barramundi recorded in a survey in Jan 2014. These fish had only died in the previous 24 hours. Both looked to be healthy with no scale loss or lesions. The only health issue they exhibited were milky eyes.

Table 1: Estimated numbers of dead, dying or sick Barramundi reported by fishers from 2011-2018

Year	Number of fish
2011	2,000+
2012	160+
2013	40+
2014	40+
2015	5+
2016	none reported
2017	400+
2018	5+

Figure 1: Two recently dead Barramundi from a survey in Jan 2014

2. Objectives

The objectives of the project were:

1. To deploy tools to automate data collection and assessment of fish health using data collected in Gladstone Harbour as a trial.
2. To undertake structured data collection of fish samples using Gladstone Harbour Partnership's reporting zones and the Boyne Tannum HookUp fishing competition.
3. To evaluate the potential to adapt the methods developed to monitor fish health in other estuaries and ports in Australia.

3. Methods

Figure 2 is a simplified flow chart of the process from collecting data in the field from Gladstone Harbour to the results of machine and human assessment. The following sections 3.1-3.7 provide details of the methods used in the various steps including the process for selecting the Object Detection algorithms used.

Section 3.8 provides the methods used for a standalone assessment of fish condition using data from the BTHU and section 3.9 provides the methods used for assessing a potential vector for the propagation of health issues.

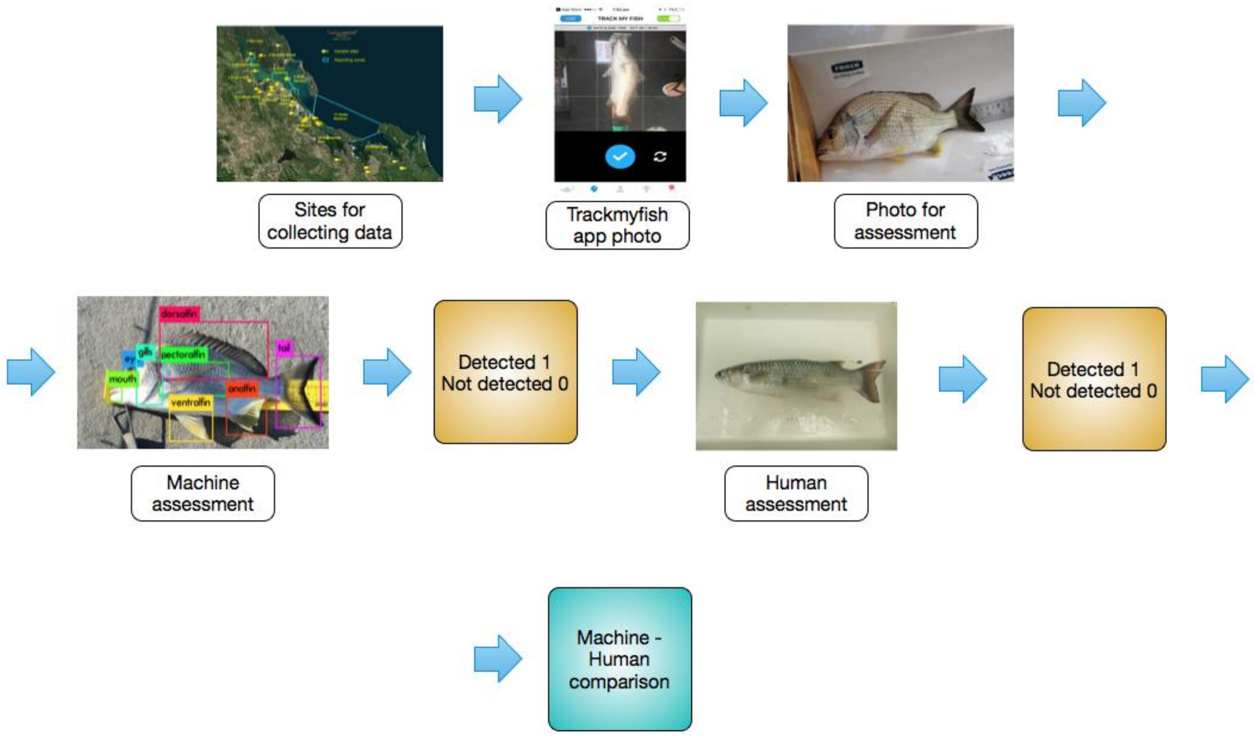


Figure 2: Simplified flow chart of the process from field collection of data to the comparison of the machine and human assessments

3.1 Location of project

The area of interest extends from the Narrows in the north to Rodds Bay in the south. The GHHP divided its area of interest into 13 reporting zones as shown in Figure 3. Fish samples for this project were collected from each of the 13 reporting zones.

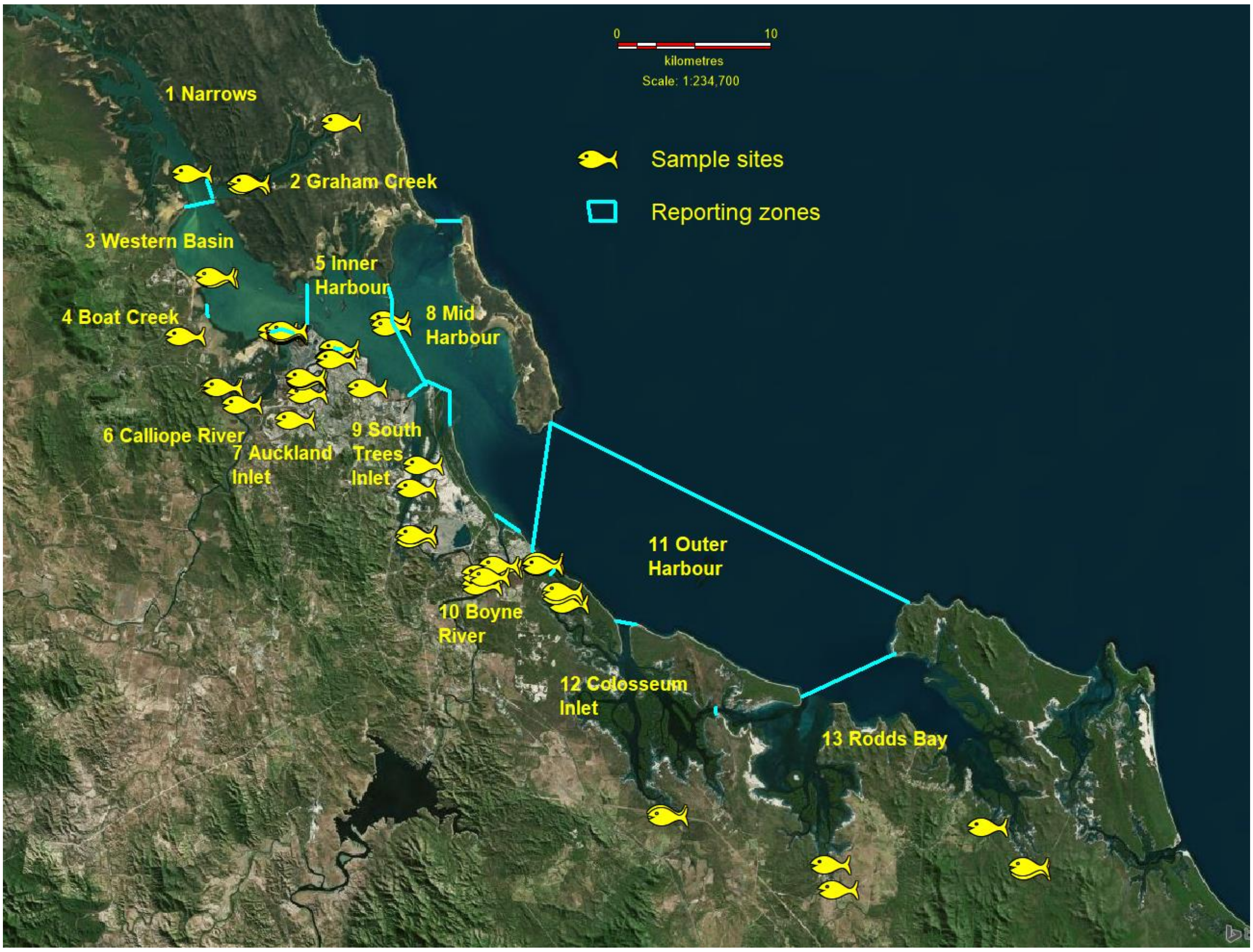


Figure 3: Gladstone Healthy Harbour Partnership reporting zones and sampling sites

3.2 Data collection through Trackmyfish

Infofish had developed a suite of phone apps based around a parent app called Trackmyfish. The base app was modified to collect data for this project. A version of the app suitable for iPhone was available from the App Store and an Android version was available from Google Play. The data collected through the app was:

- Photos of each side of the fish on a measuring ruler
- Tag number for fish that were tagged
- Fork length of the fish in millimetres
- Weight of the fish in grams
- Check boxes to record visual health issues (lesions, milky eyes, parasites, fin damage, injuries and deformities) (added for this project)
- Automatically records the GPS location if that feature is turned on

Figure 4 shows the screen used to capture the image of a fish and the screen used to record the fish details. Fish were measured to the nearest mm on a rigid ruler or brag mat that had been checked for accuracy. Most fish were tagged using Hallprint anchor tags (35mm) with a unique identifying number. Fish were tagged to allow movement to be assessed as a vector for the spreading of fish health issues. Fish were weighed on Ohaus Valor 4000W scales that record weight to the nearest gram up to 6kg.

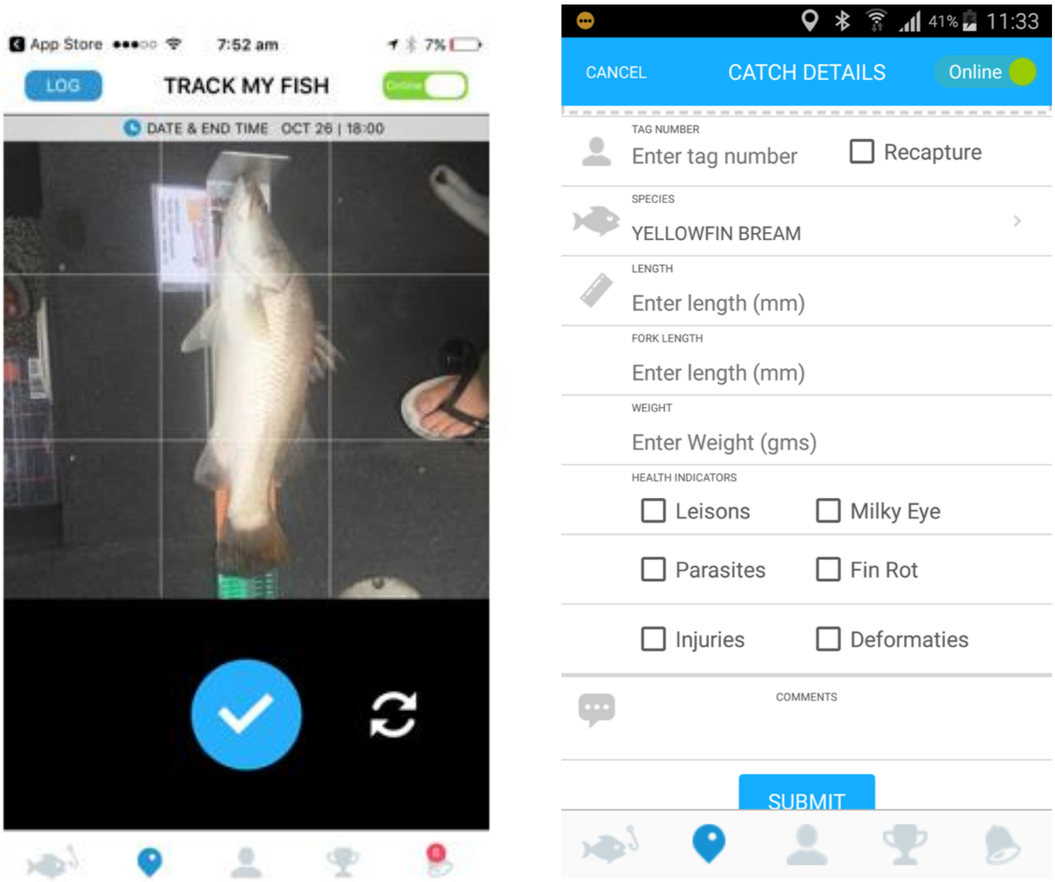


Figure 4: Screens to capture the fish image and collect the details of the fish

3.3 Collecting field samples

Field data collection was originally aimed at collecting around 25 photographic samples from each of the 13 GHHP reporting zones and from the BTHU fishing competition.

Data collection was extended to include the following to provide a greater range of samples. Numbers from each source are in Table 3:

- Photo samples of fish from the 13 GHHP reporting regions
- Photo samples of fish from the BTHU
- Video samples from the BTHU
- Central Queensland University samples collected in the Gladstone area from their health project
- ABT Tournament Bream samples from Australia wide tournaments
- Barramundi samples from the Boyne River

Target species were Yellowfin Bream (*Acanthopagrus australis*), Pikey Bream (*Acanthopagrus berda*), Flattail Mullet (*Liza dussumieri*) and Sea Mullet (*Mugil cephalus*) however a number of other species were included with some exhibiting health issues. Bream were selected as they were the most recreationally caught species in Gladstone and are found around Australia. Mullet were selected as they are an important commercial species as well as a popular bait species for recreational fishers and also found around Australia. However, there were insufficient images of Mullet to include them in the final assessment.

Field collection of samples in the 13 reporting zones was undertaken in a structured way with samples collected from each zone in a single day during March-April 2018. Samples were collected by line fishing and castnet only. The target was 25 samples in each zone. In some zones this was achieved at a single location while in others a number of locations needed to be fish to get to the required numbers.

The Boyne Tannum Hookup (BTHU) was held from 4-6 May 2018. The BTHU is one of the largest family-oriented fishing competitions in Australia with over 3,000 entrants. As part of the event there is a live weigh-in section where fish of a number of species can be brought in alive. These fish are measured, weighed and tagged and then placed in a holding tank for public viewing before eventually being released into the Boyne River during the competition.

Fish brought to the weigh-in were sampled for health issues. Photos were taken of both sides of the fish in the measuring cradle as well as being videoed by a camera mounted above the measuring table. That provided two different datasets for the same fish. All of the fish brought into the live weigh-in were photographed although not all fish were captured on video.

Figure 5: Fish being measured, photographed and videoed at the Boyne Tannum HookUp

Figure 5 shows the measuring table with the measuring cradle and fish with the video camera directly above and the fish being photographed using an IPad. Figure 6 show a view of a fish from the video camera while Figure 7 and Figure 7Figure 8 show typical photos collected of fish presented at the weigh-in.

Figure 6: View of fish (Barred Javelin) in measuring cradle from the video camera

Figure 7: BTHU Yellowfin Bream with small lesion on side

Figure 8: BTHU typical Pikey Bream

The Central Queensland University also collected fish samples in the Gladstone area as part of the GHHP assessment of fish health. They also collected photos as part of that project and provided those for machine assessment.

ABT Tournament Bream samples were collected through the Trackmyfish phone app from tournaments around Australia. They were collected for tournament purposes, not collected for this purpose. A sub sample of around 600 fish were machine assessed to examine how useful the system was in using Australia wide tournament data. Given that an objective of the project was to evaluate the potential to adapt the methods to other parts of Australia, this assisted in that evaluation.

There were also a number of photos obtained from the Boyne River of Barramundi with specific health issues that were relevant to fish health issues in the Gladstone area. These were also machine assessed. These included dead fish and fish with abrasions, lesions and milky eyes.

3.4 Post field machine assessment steps

The development of the machine learning process was in several steps:

- Evaluate Convolutional Neural Network (CNN) Object Detection algorithms to determine the most appropriate for detecting fish features and the various health issues
- Selecting the most appropriate algorithms for each health issue being assessed
- Training the algorithms to identify a range of features on fish (eg fins, gills, eyes) and health issues (eg lesions, fin damage, cloudy eyes, wounds)
- Applying the algorithms to the photos and videos and determining if fish health issues were detected

3.5 Selecting appropriate Object Detection algorithms

Following a worldwide search of Object Detection algorithms there were 2 that were selected for further evaluation. These were:

- TensorFlow Object Detection API
- Yolo v2 and Yolo v3 Object Detection algorithm

3.5.1 TensorFlow Object Detection API vs Darknet Yolo (You only look once)

TensorFlow is an Open Source Machine Learning platform developed in Python. TensorFlow was developed by the Google Brain team for internal Google use. It was released under the Apache 2.0 open source license on November 9, 2015.

Yolo is a computer vision focused software system developed by Joseph Redmond at the University of Washington with assistance from Ali Farhadi (Yolo v3). Initially Version 2 was trialed, then updated to Version 3 in the final assessment.

Both TensorFlow and Yolo use the same set of pre-trained models (VGG, Googlenet, Resnet) as a foundation, though each provides a specific model that works best for their platform.

In evaluating the technology 3 key factors were considered:

- Training
    - What tools are provided to code training data
    - What level of independent assessment of trained models is available
    - What skill levels are required for human coders to complete training
    - Can training rules for human coders be established
- Accuracy and Scalability
    - Accuracy of the model (assessing ability to predict rather than % accuracy of the prediction)
    - Ability to service many users
    - Ability to work with differing quality of fish data
    - Ability to be automated
- Usability in the real world
    - Easy maintenance of code
    - Ability to be installed and used in a production environment
    - Ability to integrate with image capture processes

3.5.2 Training

Note – for assessment a test model of fish body parts (tail, anal fin, ventral fin, pectoral fin, dorsal fin, gills, eyes) was used. This model was used as if the test couldn't detect these common features, there was a low likelihood of success with smaller health features).

TensorFlow

The training process is the largest point of difference between TensorFlow and Yolo. TensorFlow uses a "whole of image" approach to training, so training images need to include the whole of the object. This means images need to be cropped and saved as a new file for each factor being assessed. With the test model (body parts of the fish) extending to 7 different sections, this was a very time-consuming process and thus only a smaller number of images could realistically be coded. TensorFlow comes with a training graph that tracks the level of error in the training process but does not come with any tools to do post model assessment. Trained models (an output file that can be used for detection) are only output on completed training, so only a completely trained model can be assessed. The TensorFlow Object API does not have any useful post training assessment to determine if the model is viable or not, so the only real test is to detect objects.

Yolo

Yolo takes a different approach to training using a guided training process where end users use a software tool to mark out the areas on the image where an object (eg tail) can be spotted. This process is much faster to perform with operators with skill in identifying the object of desire requiring only a 10 minute training course to be highly productive in creating training datasets. As the coding process is non-destructive there is no need to keep separate copies of cropped images and review by a second party for accuracy is similarly a simple process.

The second advantage of this coding process is the Yolo tool can "validate itself" by checking a set of images coded by a human and comparing its' results both in terms of overlap of the computer generated with the human's bounding box (mean average precision or mAP) and the ratio of the overlap to the non-overlap portion (intersection of union or IoU). The average of these two statistics over a set of images provide a reliable indicator of how successfully the machine learning model would work in real examples, thus it saved a lot of time in determining whether to use a trained model or not. A low score on these indicators always resulted in a trained model that failed to detect anything, a high score resulted in a model that was more likely to successfully detect items.

3.5.3 Accuracy and Scalability

TensorFlow

While TensorFlow training performed at similar levels to Yolo when it came to detection on the test model, there was variable success in training on health factors. Easily identified issues such as lesions trained fine but small items such as fin damage were very difficult to train on. The resulting cropped training images were either too small to be used, or confusing to the training as it couldn't differentiate between the larger structure (fin) and the actual issue (splitting or other damage). As a result, TensorFlow was unable to detect these issues.

On scalability measures, TensorFlow proved capable of handling large batches of images and larger user bases through upgrading hardware. Context of images was also harder to assess but on available evidence, should handle photos of fish in different contexts, though there is no way to assess what combination of contexts would work beyond trial and error.

Yolo

With its guided process to identifying which part of an image should be coded to a class, Yolo provided superior detection of small health issues such as fin damage on the same dataset in that, unlike TensorFlow where none of the issues were detected, Yolo provided many successful detections.

On scalability measures, Yolo proved capable of handling large batches of images and larger user bases through upgrading hardware. Context of images was easier to assess as models could be trained using batches with the same context (eg all fish on the ruler) and batches of variable context (eg fisher holding the fish) and compare the mAP, IoU and final detection on fresh images.

Note - The ratio of contexts used in training images has to be carefully managed, such that random selection of images to run statistical validation will end up with the same ratio as the training data.

3.5.4 Usability in the real world

TensorFlow

TensorFlow is a general-purpose environment for a range of machine learning and computational purposes and as such, code has to be developed to make it work. In order to expedite that process pre-existing code developed by the Google team was sourced for the purposes of trialing. This presented a number of challenges across the trial as TensorFlow, Python and its libraries being open source, are being updated all the time and by the end of the trial the latest version, while still working, would need updating to remain compatible with TensorFlow. More than once the test environment had to be rebuilt because of conflicts in code libraries.

Overall, TensorFlow worked fine but presented key risks in terms of maintainability due to the decentralised nature of the management of code base. Nonetheless there is a clear (if complex) process to install on another environment. Integration with image capture is straight forward as detection processes can be run both individually or in batch mode.

Yolo

Yolo was developed in C and has to be compiled in the target environment. This proved relatively straight forward and once compiled the executables could be installed on other environments with a simple script without need for dependencies other than CUDA, the standard library for GPU processing. As Yolo is a dedicated system there is support for changes to the code for many common tasks, though thus far no changes have been required as all required functionality was ready to go.
Overall, C is more robust as a language and movement of Yolo between environments is simpler than TensorFlow. Overall in terms of performance, the C code base makes Yolo significantly faster (better than 2X faster than TensorFlow on the same environment) especially in assessing video. Yolo provides a straight forward migration process and can be run individually on images or in batches making for equivalent integration to TensorFlow.

3.5.5 Final evaluation

Regardless of the accuracy or speed of the image assessment process, training is the biggest differentiator of success – garbage in garbage out. The other variables in this decision and statistics attached to them can be seen one way or the other based on the requirements of the implementor. The Yolo system with its guided process to training, ease of additional human validation and built in statistical assessment of trained models was found to be superior. Yolo demonstrably reduces the time taken to create and validate working trained models.

As the training environment could be tailored to the specific needs of fishing and provide a guided training process for human image coders – Yolo was ultimately adjudged to be the better choice of the two for the health assessment process.

3.5.6 TensorFlow and "redness"

While Yolo was found to be superior in many instances there was a role for TensorFlow. In this case the Google Inception model has been used as this has a very high level of accuracy after retraining. Unlike Yolo, TensorFlow has greater ability to detect less well defined patterns such as small lesions, fin infection, tail infection or "redness".

Detection is made possible through slicing of the main image into smaller images (eg 200px*200px) which removes most of the context of the fish and allows the machine learning to focus on the key differences between small slices, comparing slices with no issues with slices with a range of "redness" issues. While this method is less efficient, requiring potentially a hundred or more assessments per fish, it offers the maximum chance of finding an issue.

In terms of final reporting of issues, a result of one slice reporting a positive result is recorded as a positive result for the whole fish. Like Yolo, this method does provide some context of where on the fish the issue occurred as positive slices are recorded, but there is no clean "box" around the issue in question.

Longer term this process will help with assessment of issues where less training data is available.

3.5.7 Use of UNet

One model that hasn't been used but may be important in the future is UNet which is used for assessment of medical conditions examining x-rays, MRI and photos. UNet requires a specific issue that can be cleanly defined (eg skin cancer) but has the advantage of processing at much greater accuracy.

UNet uses "masks", that is, images where all data is redacted other than the issue, which is further processed until a negative shape is produced. The masked images are then used for training data. UNet has significant training overheads to conventional Convolutional Networks but may be important in the future for specific issues or use of radiographic images to assess fish health.

3.6 Post field machine assessment

3.6.1 Training the Object Recognition Algorithms

To undertake the training of the machine learning models to recognise the various parts of a fish and health issues ideally requires 1,000+ images however training can be completed on fewer images. In this case we aimed for 700+ images however acceptable results were obtained for some issues with fewer images. Acceptable results were where a 95% accuracy was achieved in line with machine learning algorithm practices. Some health/injury factors such as wounds and milky eyes will need additional datasets as there were insufficient images available to complete the training. These will be added over time when a sufficient number of images are available.

In training machine learning models, the training is for robustness, that is the ability for the model to cope with new sources of data rather than just the conditions in the training set.

Where possible the species balance needed to be maintained (ie roughly equal number of images)

- Balance the number of items identified in training coding per factor being assessed - if too many images have one issue only, then the models will overfit
- Show fish in as many conditions as possible (held, etc) and not just on a ruler

When training models, if one condition is identified more often (eg fin damage) those will have a higher prominence in the model. This ratio will improve over time as additional data becomes available, but in order to not over bias the trained models two separate models were established:

- Fin damage (which is more common)
- Wounds/disease factors

The training process was as follows:

- Images were selected for training
- Training has focused on Bream as that species had sufficient images to assess however training will now be extended to other species
- Images were coded using a boxing tool that allows an association of a region on the image with an issue (eg boxing of fins, eyes, gills etc)
- The coded training data was converted into machine learning format using a script
- A list of images was generated for use in training and a separate list of images was compiled for use in validation by the machine learning model by script that randomly selected 10% of the images for validation
- The data was then moved into the machine learning environment
- The optimal parameters for training were calculated and configuration files were set up
- The machine learning process then commenced
- Where there were less than 500 images run a minimum of 5,000 training cycles for the testing model, 10,000 for a production model where there were over 500 images (note training process saves the model every 100 cycles)
- A batch process was run to assess each of the output trained models for two parameters, which were used as a standard for assessment:
    - Intersection of Union (IoU) – this compared the generated bounding boxes on the test images with the manually provided boxes. IoU – Sum (Area of Overlap/Area of Union))/number of images assessed
    - Mean Average Precision (mAP) – this looked only at the intersection of the detection box verses the human generated box. mAP – Sum (% overall intersect per tested image)/number of images assessed
- Models that had the highest IoU and mAP were accepted
- Backup was made of any previously trained model (if present)
- The accepted trained model was moved into the live model folder and renamed to the final model

Each element was trained for 6,000+ cycles using 2 models with the best mAP selected with different levels of training, so that it could be assessed if additional training cycles improved accuracy. Tail damage was split up into two categories (tailsplit – splits and damage – less regular type damage). Using 2 models worked better as between the two the models picked up damage better than both models lumped together.

3.6.2 Assessing collected images

Assessment of images was run in a batch process for all photos and videos based on the following:

- Each photo/video was assessed individually
- All health factors were detected simultaneously each time a photo/video was assessed

- An output photo/video was generated that highlighted all detected issues
- An output text file was generated with all items predicted with the level of certainty (0-100%)

Results were loaded into a datafile through a script that read all the output predictions and converted them to single rows of data with each health item given a 0/1 value.

0 = not detected, 1 = detected

The assessment process was as follows:

- Separate folders were created for images from the different sources – GHHP, BTHU, CQU and ABT
- Separate output folders were created for the assessed images – GHHP, BTHU, CQU and ABT
- Separate folders were created for the videos – BTHU Video
- Separate folders were created for the assessed videos – BTHU Video
- All images from these batches were placed into the appropriate folder
- Scripts were generated to run an assessment against each image/video
- Assessment scripts were run to read in all detected factors and compiled into a CSV file
- Output images were manually checked if required
- An output dataset was created for analysis

Figure 9 is a Flattail Mullet from the images provided by CQU and used to assess “redness”.

Figure 9: CQU image of Flattail Mullet used in assessing “redness”

3.7 Post field independent visual human assessment

Post field human assessment of health issues was completed for the samples collected in the 13 GHHP reporting zones.

The master dataset was the reporting zones structured sample. In the field assessment, only health issues were recorded while in the post field assessment injuries were included as health issues were minor while injuries were common.

In the initial assessment table only one column was provided for injury however this was broken out into multiple factors (wounds, scale damage, tail and fin damage) to better identify the different types of injuries.

For the remaining datasets BTHU photos, BTHU video, CQU photos and ABT photos these were machine assessed first and then by human.

The purpose of human assessment was to provide a useful classification of issues detected without providing diagnosis. A secondary process provided the most useful and accurate classification in the most cost-effective manner, so that the process could be continued until machine learning models are self-sustaining.

The human assessment process was developed around two human assessors, one for assessing the bulk of the images for reporting of any issues, then a more expert user to provide a more appropriate classification.

1. Human assessor was trained on a large number of healthy fish of the species in question
2. Human assessor was not trained on health issues but was trained on identifying fin issues and tail issues
3. Human assessor recorded all fin and tail issues
4. Human assessor marked anything that was out of the ordinary from a healthy fish for post assessment
5. A selection of Issues recorded were passed to veterinarian with expertise in fish for comments and advise on assessment
6. Images in post assessment were passed to environmental scientist on the team for classification incorporating feedback from veterinarian.

3.8 Assessing Fish Condition

While health of fish can be assessed based on visible health issues it can also be assessed by condition (the relationship between length and weight). Length and weight data have been collected from 2003 to 2017, except 2009 and 2011, in the BTHU. Length-weight data was collected in 2018 again in the BTHU and in the structured surveys in the GHHP reporting zones. The growth type and relative condition of Bream (Yellowfin and Pikey) in 2018 was compared to the historical BTHU datasets.

Length and weight data were plotted in RStudio and the numerical relationships were calculated for each Bream species. The length-weight relationship for each species was established using a power curve function, which generates the formula $W={}_{a}L^{b}$.

In this formula:

- W is the weight of the fish
- 'a' is the constant (or exponent) describing the rate of change of weight with respect to length
- L is the total or fork length of the fish, and
- 'b' denotes the weight at sample length

Values of 'b' also provide information on fish growth (Table 2) (LeCren 1951).

Table 2: Description of 'b' values used in length-weight analyses

b	Fish Growth	Description
=1	isometric	growth is uniform for length and weight
>1	positively allometric	fish get rounder as they grow in length
<1	negatively allometric	fish get longer as they grow, relative to roundness

Mean and median relative condition (Kn) (LeCren 1951) was calculated for each Bream species. Kn is useful for detecting prolonged physiological stress on a fish population and shows the natural variation of condition eg pre and post spawning (Swingle and Shell 1971, Peig and Green 2010, Guidelli et al 2011). Relative condition is the division of the actual weight of each fish in the sample by the mean weight of the total sample.

The relative condition factor is calculated with the formula Kn = W/W' where:

- W is the actual weight of an individual fish, and
- W' is the predicted length–specific mean weight for the population under study, which is calculated from the power curve function 'a' and 'b' value output in the equation $W=aL^b$.

A relative condition factor of 1 is expected for an individual or a sample/sub-population. An individual or sample of fish with a relative condition factor < 1 is considered to be in poor to moderate condition, while a relative condition factor > 1 is considered to be in good condition (LeCren 1951). Median values have been used in this study as the ranges of weight for Bream species at larger sizes can be highly variable.

3.9 Potential vector for propagation of health issues

Fish sampled in the 13 reporting zones were tagged with Hallprint T-bar anchor tags to assess fish movement based on recaptures. Historic tag data from 2000/01—2017/18 collected through Suntag were also assessed in relation to fish movement patterns.

Tag locations were recorded based on Suntag grid maps with grids being $1km^2$. Locations of recaptures were also recorded at the grid map level where sufficient details were provided of the recapture location. Recaptures of fish provide an insight into residence in a location or movement of fish with the potential to propagate health issues.

The 13 reporting zones are covered by the following Suntag grid maps (available from www.suntag.org.au):

- Curtis Island Gladstone (CISG)
- Gladstone Harbour (GLD)

- Calliope River (CR02)
- Boyne River (BRG)
- Rodds Bay (RBT)

Fish recaptures were assessed based on the time at liberty from tagging to recapture and the distance and direction the fish moved between tagging and recapture. Distance moved was based on the shortest distance by waterway between locations.

For the historic tagging data it was considered that fish recaptured in the first year after tagging were most likely to provide information on movement that could be associated with the propagation of health issues. Movement of fish tagged during the collection of samples was also assessed on the same basis. The distance moved from the tag to the recapture location was based on the following:

- 0-2km were fish that were considered to have been recaptured in the same area
- 3-10km
- 11-20km
- 20+km

Fish submitted to the BTHU live weigh in since 1999 were tagged and released at the Bray Park boat ramp on the Boyne River. All fish were released at the same location, so it provides a view of fish dispersal that could potentially result in propagation of health issues.

Fish health issues emerged in 2011 after the spilling of Lake Awoonga and releasing an estimated 30,000+ Barramundi into the Boyne River, Gladstone Harbour and beyond. Fish tagged in Lake Awoonga were part of the spill and recaptures of these fish, and fish from a subsequent spill in 2017 were assessed for movement and potential propagation of fish health issues. Days out were calculated from the time the spill occurred and not from the date of tagging.

4. Results

4.1 Fish samples collected

Table 3shows the number of fish samples from the various sources including the number of Bream (both species). For the assessment, all fish in the GHHP and CQU samples were used, while for the BTHU and ABT tournaments only Bream were assessed.

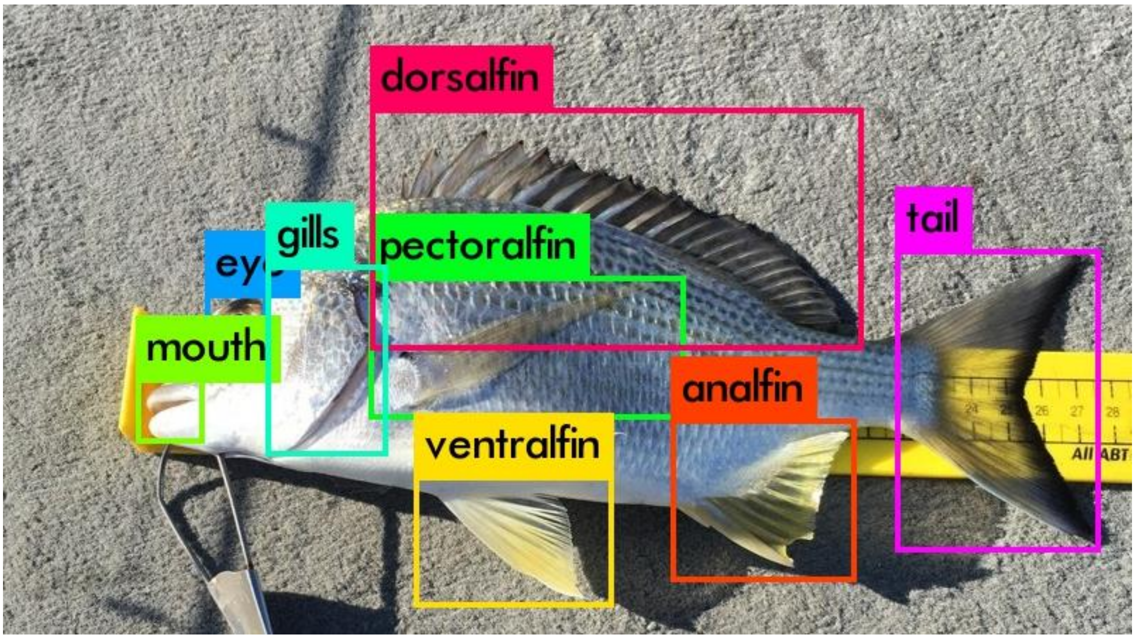
Table 4 shows the number of selected training cycles and the resulting maP (mean average Precision) and IoU (intersection of Union). Figure 10 shows the result of the output of the training model using a ‘standard’ Image and an image of a fish being held. Figure 11 shows the detection of a small tail split with statistics available on the bounding box. The dog images provide the user with examples of how to use the boxing tool.

Table 3: Number of photos and videos obtained from the various sources

Source	Photos (fish)	Photos (Bream)	Videos (fish)
GHHP ZONES	345	233	
BOYNE TANNUM HOOKUP	508	240	500
CQU HEALTH SAMPLES	58	2	
ABT TOURNAMENTS	599	599	
BOYNE RIVER	10	0	
TOTAL	1,520	1,074	500

Table 4: Models selected and the number of selected training cycles and the resulting maP and ioU

	Model	selected training cycles	mAP	IoU
1	finsplit.model.1	2288	24.98	24.98
2	finsplit.model.2	4472	18.08	18.08
3	tailsplit.model.1	936	19.74	14.85
4	tailsplit.model.2	5824	9.51	3.33
5	taildamage.model.1	5200	51.62	52.21
6	taildamage.model.2	6800	52.17	51.17
7	wound.model.1	2288	26.62	65.34
8	wound.model.2	4160	17.05	65.52



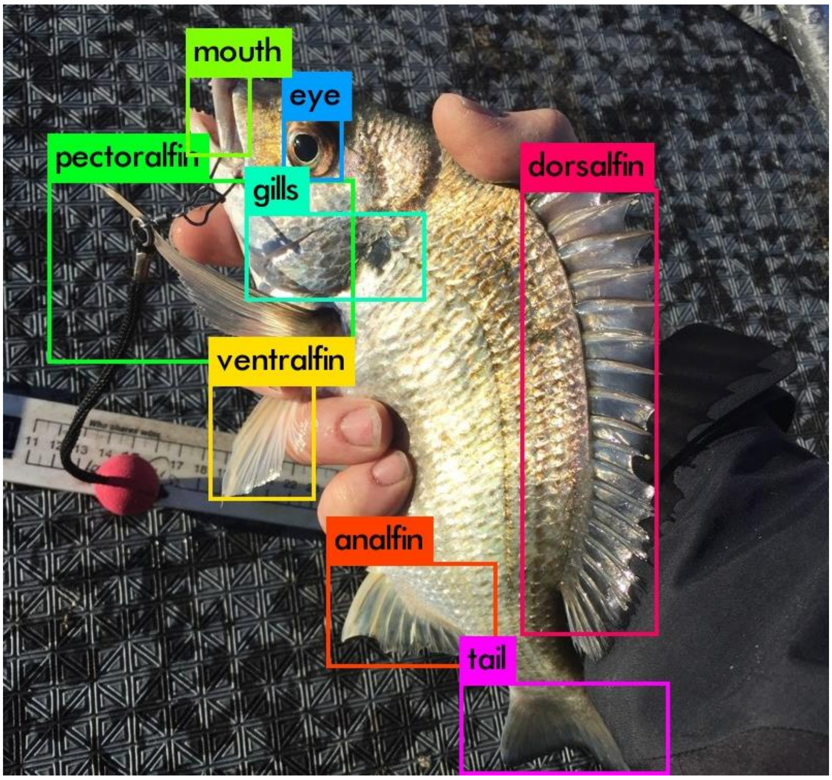


Figure 10: Output images showing the elements of the fish as recognised by the model

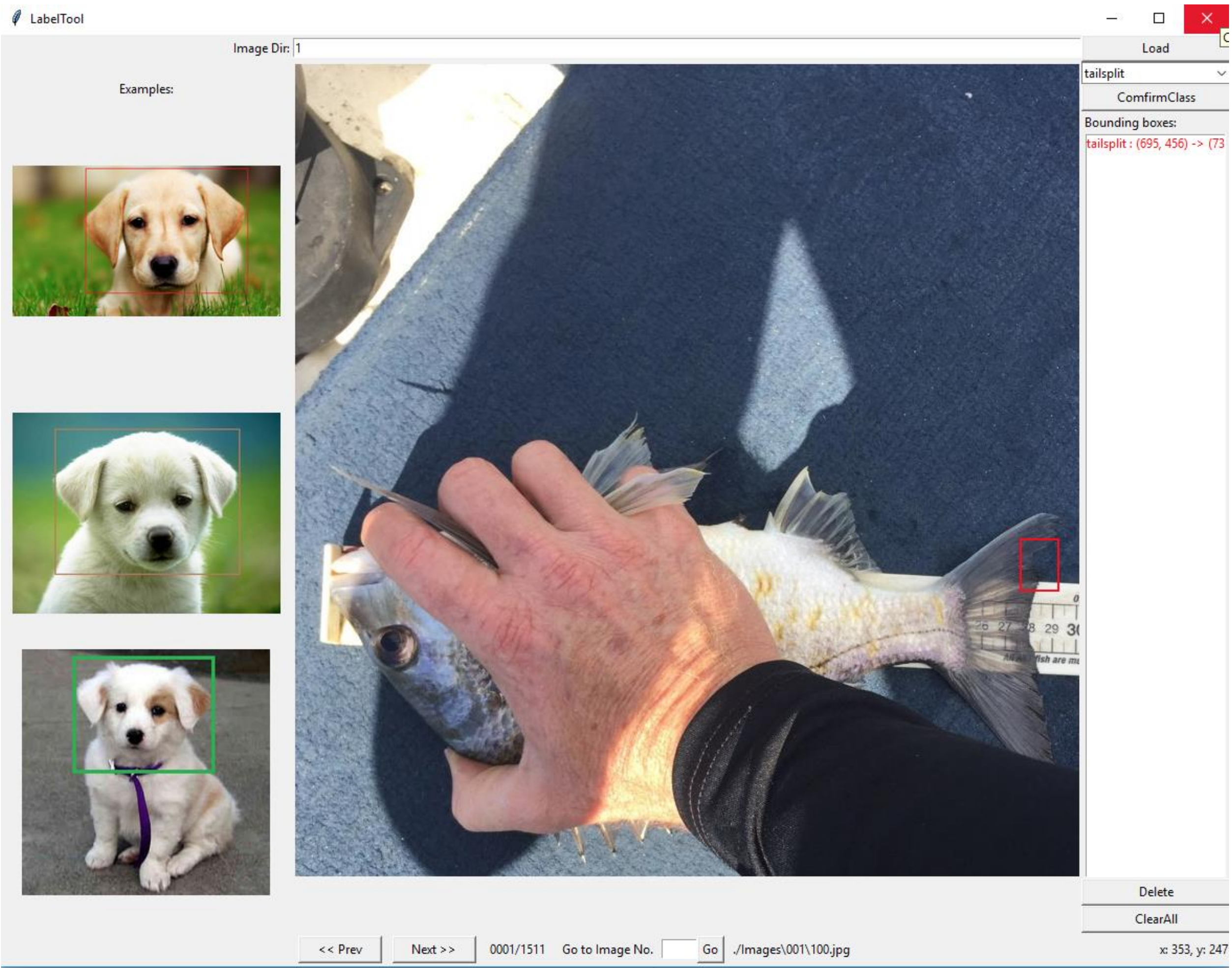


Figure 11: Machine learning training with target object tail split identified (dog samples show how the boxing tool is used)

During the project only a very limited number of fish were detected with health issues and even then issues fell into a wide variety of classifications. In order to focus on usable practical training, 3 key issues were selected – tail damage, fin damage and wounds. Additionally, milky eye was trained for but as no instances were encountered during sampling, milky eye was excluded from the assessment.

Results were better on square images over widescreen images as training was completed using square images and the image recognition is optimal for square images.

There was a total of 1,242 images assessed. Machine and human agreement levels ranged from 50%-86% for fin splitting, 60-93% for tail splitting, 78-93% for tail damage, while the wound model was unsuccessful in all instances.

Table 5 shows the comparison of machine and human assessments with the total number of images assessed, the number of images where the machine and human assessments agreed and the number of machine and human positive assessments for a particular element.

The CQU samples were also assessed for "redness". There were 58 images assessed with an 86% agreement between machine and human assessment as shown in Table 6.

Table 5: Comparison of results of machine and human assessment for each element assessed

FIN SPLIT						
	assessment	**total**	**machine/ human agree**	**machine positive**	**human positive**	**percentage agreement**
1	abt.model.1	599	493	110	184	82.3
2	abt.model.2	599	515	106	184	86.0
3	bth.model.1	240	168	12	80	70.0
4	bth.model.2	240	183	23	80	76.3
5	cqu.model.1	58	32	9	35	55.2
6	cqu.model.2	58	29	6	35	50.0
7	ghhp.model.1	345	254	129	166	73.6
8	ghhp.model.2	345	273	106	166	79.1

TAIL SPLIT						
	assessment	total	machine/ human agree	machine positive	human positive	percentage agreement
1	abt.model.1	599	480	140	184	80.1
2	abt.model.2	599	555	205	184	92.7
3	bth.model.1	240	186	71	80	77.5
4	bth.model.2	240	217	64	80	90.4
5	cqu.model.1	58	38	29	35	65.5
6	cqu.model.2	58	35	26	35	60.3
7	ghhp.model.1	345	215	284	166	62.3
8	ghhp.model.2	345	244	243	166	70.7

TAIL DAMAGE						
	assessment	total	machine/ human agree	machine positive	human positive	percentage agreement
1	abt.model.1	599	530	89	70	88.5
2	abt.model.2	599	573	66	70	95.7
3	bth.model.1	240	196	56	60	81.7
4	bth.model.2	240	223	57	60	92.9
5	cqu.model.1	58	46	10	14	79.3
6	cqu.model.2	58	49	9	14	84.5
7	ghhp.model.1	345	269	188	160	78.0
8	ghhp.model.2	345	303	168	160	87.8

WOUND						
	assessment	total	machine/ human agree	machine positive	human positive	percentage agreement
1	abt.model.1	599	0	6	1	0
2	abt.model.2	599	0	2	1	0
3	bth.model.1	240	0	37	2	0
4	bth.model.2	240	0	1	2	0
5	cqu.model.1	58	0	2	7	0
6	cqu.model.2	58	0	1	7	0
7	ghhp.model.1	345	0	3	11	0
8	ghhp.model.2	345	0	1	11	0

Table 6: Comparison of results of machine and human assessment for "redness"

ORIGINAL MODEL - REDNESS						
	assessment	**total**	**machine/ human agree**	**machine positive**	**human positive**	**percentage agreement**
8	cqu	58	49	34	41	86.0

4.2 Relative fish condition

4.2.1 BTHU and GHHP Data

As an additional standalone assessment, the growth type and relative condition for Yellowfin and Pikey Bream species were calculated from the 2018 BTHU and GHHP data. A combined growth type and relative condition was calculated also.

The length-weight curve for Yellowfin Bream from BTHU and GHHP datasets (combined) is presented in Figure 12 noting a division between datasets at approximately 250mm. The length-weight relationship of 139 Yellowfin Bream in the BTHU dataset showed greater variation in weight at a given length ($R^2 = 0.90$) than the 65 Yellowfin Bream in the GHHP dataset ($R^2 = 0.97$) (Table 7). It should be noted that the GHHP dataset comprised smaller Bream (<250mm) than the BTHU dataset (no fish <250mm) and variation in weight is more common in fish at larger sizes. Negative allometric growth was displayed by Yellowfin Bream in both datasets. The range of relative condition factors for Yellowfin Bream was narrower in the BTHU sample and the median relative condition factor calculated for Yellowfin Bream in the BTHU and GHHP datasets were 1.01 and 1.00, respectively.

The length-weight curve for Pikey Bream from BTHU and GHHP datasets (combined) is presented in Figure 13 noting a division between datasets at approximately 250mm. The length-weight relationship of 98 Pikey Bream in the BTHU dataset showed slightly greater variation in weight at a given length ($R^2 = 0.93$) than the 193 Pikey Bream in the GHHP dataset ($R^2 = 0.95$) (Table 8). Similar to the Yellowfin Bream samples, the GHHP dataset comprised more smaller Bream (<250mm) than the BTHU dataset (no fish <250mm) and variation in weight is more common in fish at larger sizes.

Positive allometric growth was displayed by Pikey Bream in the BTHU dataset while negative allometric growth was displayed in the GHHP dataset. The range of relative condition factors for Pikey Bream was much narrower in the BTHU sample and the median relative condition factor calculated for Pikey Bream in the BTHU and GHHP datasets were 0.99 and 1.00, respectively.

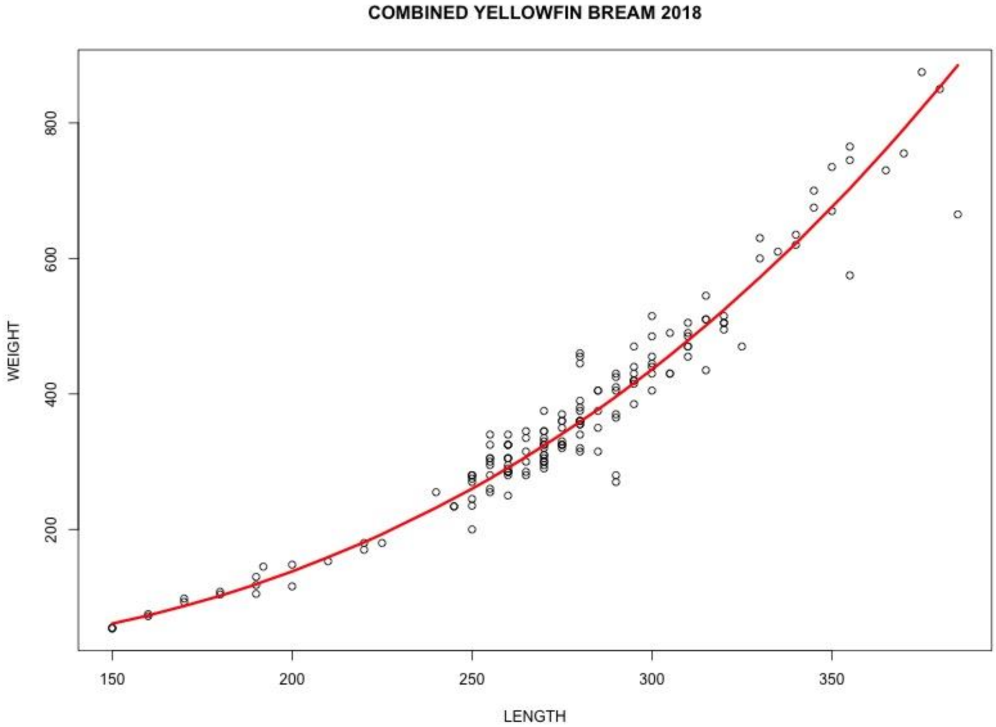


Figure 12: Length (mm) and weight(g) curves for Yellowfin Bream in BTHU and GHHP samples.

Table 7: Growth type and relative condition factor for Yellowfin Bream in BTHU and GHHP samples.

Dataset	n	b value	Growth Type	R^2	Relative Condition (Kn)			
					Min.	Max.	Mean	Median
BTHU	139	2.69	allometric (-)	0.90	0.68	1.27	1.01	1.01
GHHP	65	2.93	allometric (-)	0.97	0.73	1.35	1.01	1.00

Given that each 2018 dataset comprised Yellowfin and Pikey Bream of different size ranges, eg strictly >250mm in the BTHU dataset (legal fish measured only) and mostly <250mm in the GHHP dataset, length-weight data for each species was combined and growth type and relative condition calculated to determine whether the results from either dataset were potentially unreliable. The results indicate that length-weight relationships, growth types and relative condition summary statistics of the individual 2018 datasets are comparable to the more complete combined dataset (Table 9).

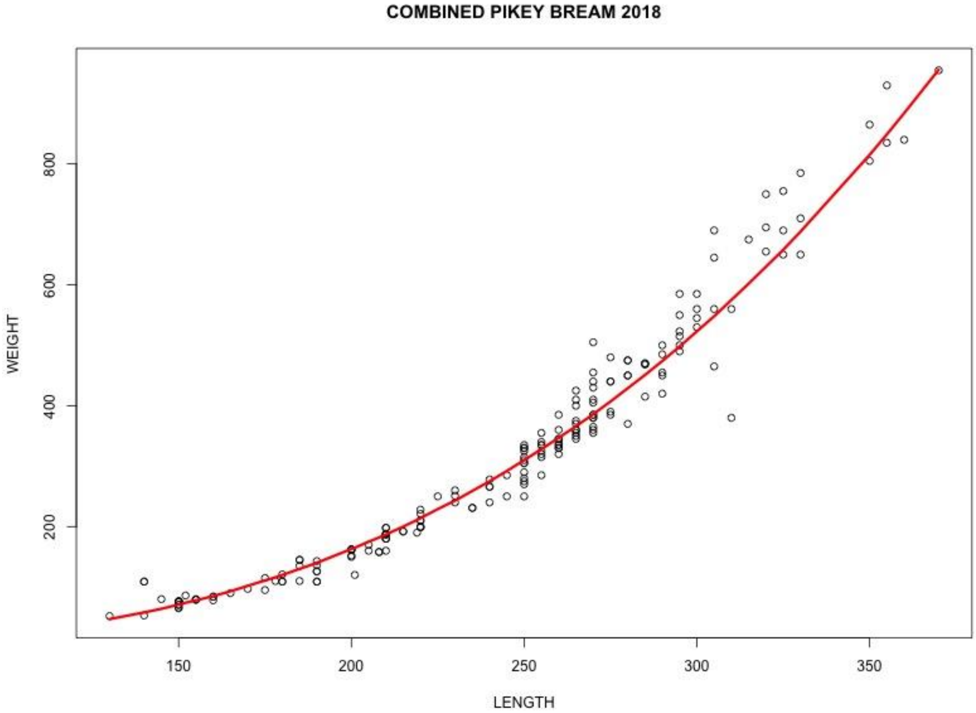


Figure 13: Length (mm) and weight (g) curves for Pikey Bream in BTHU and GHHP samples.

Table 8: Growth type and relative condition factor for Pikey Bream in BTHU and GHHP samples.

Dataset	n	b value	Growth Type	R^2	Relative Condition (Kn)			
					Min.	Max.	Mean	Median
BTHU	98	3.06	allometric (+)	0.93	0.82	1.29	1.00	0.99
GHHP	193	2.76	allometric (-)	0.95	0.71	1.82	1.01	1.00

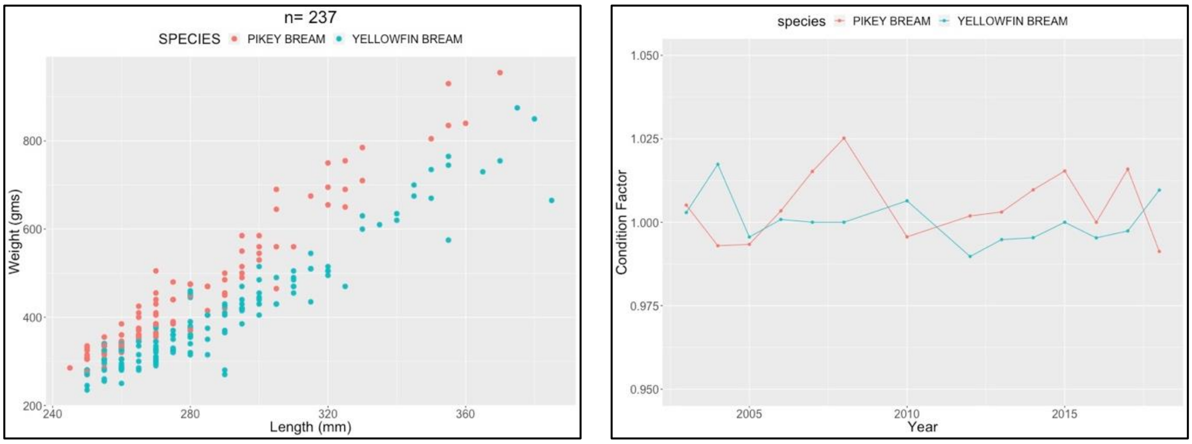
Table 9: Relative condition factor for the BTHU and GHHP samples combined

Species	n	b value	Growth Type	R^2	Relative Condition (Kn)			
					Min.	Max.	Mean	Median
Yellowfin Bream	204	2.81	allometric (-)	0.98	0.68	1.40	1.01	1.00
Pikey Bream	291	2.88	allometric (-)	0.97	0.66	1.88	1.01	1.00

4.2.2 Historical Data

The length weight curves for Pikey and Yellowfin Bream caught in the BTHU in 2018 and the mean and median ranges of relative condition for Yellowfin and Pikey Bream, calculated for all years (2003 to 2018) are displayed in Figure 14, a and b. The summary statistics for Yellowfin and Pikey Bream for all years are presented in Table 11 and Table 12Table 10. Yellowfin Bream displayed positive allometric growth in 1 of 13 years and Pikey Bream displayed positive allometric growth in 4 of 13 years. Summarised relative condition for each species for all years are presented in Table 10.

The range of historical median relative condition factors for Yellowfin and Pikey Bream was 0.99 to 1.02 and 0.99 to 1.03, respectively. Median relative condition of Yellowfin and Pikey Bream from the 2018 datasets were within the range calculated for each species from the historical datasets.



a) b)

Figure 14: Length weight curve for Bream species 2018 (a) and median (b) condition factors for Yellowfin and Pikey Bream from the BTHU

Table 10: Relative condition factor from BTHU from 2003-2017 and results from 2018

Species	Relative Condition (Kn) (Historical range)		Kn (BTHU 2018)	Kn (GHHP 2018)	Kn (BTHU and GHHP combined)
	Min	Max			
Yellowfin Bream	0.99	1.02	1.01	1.00	1.00
Pikey Bream	0.99	1.03	0.99	1.00	1.00

4.3 Potential vector for propagation of health issues

4.3.1 Current

There were 289 fish (all species) tagged from 9/3/2018-20/4/2018 during the collection of samples in the 13 GHHP reporting regions. Of those fish there have been 5 (1.7%) recaptured through to 31/7/2018. Of those fish 3 were recaptured in the same area as tagged (0-2km) and the other 2 were recaptured from 3-5km from where tagged. It is expected that there will be further recaptures over time.

4.3.2 Historical

From 2000/01-2017/18 there were 24,395 fish (all species) tagged in the reporting regions (excluding fish tagged at the BTHU) with 2,185 recaptures (9.0%). There were 1,489 fish recaptured in the first year after tagging of which 1,433 had sufficient data to determine movement.

Figure 15 shows the percentage of recaptures based on the distance moved and based on the grid maps. Overall 75.1% were caught in the same area as tagged (0-2km) while 4.8% were recaptured 20+km from where tagged.

From 2000-2018 there were 6,465 fish (all species tagged at the BTHU with 330 recaptures (5.1%). As all fish were released at the same location, the Bray Park boat ramp, this provided the opportunity to review the distance moved as well as the direction.

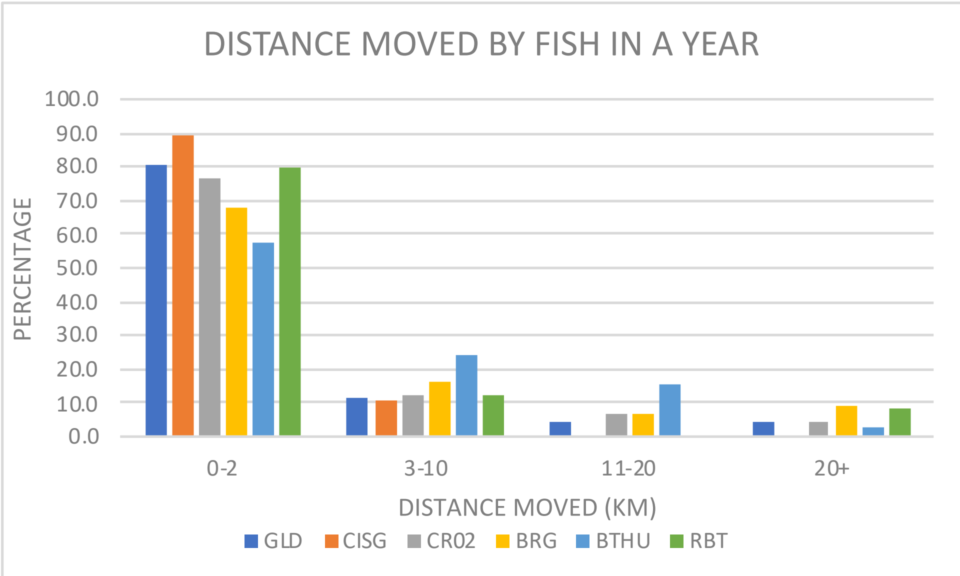


Figure 15: Distance from tagging to recapture location for fish recaptured within a year of release

Figure 16 shows the distance moved (kms) and the direction for 4 key species released at the BTHU. Species are Yellowfin Bream, Pikey Bream, Dusky Flathead and Barred Javelin.

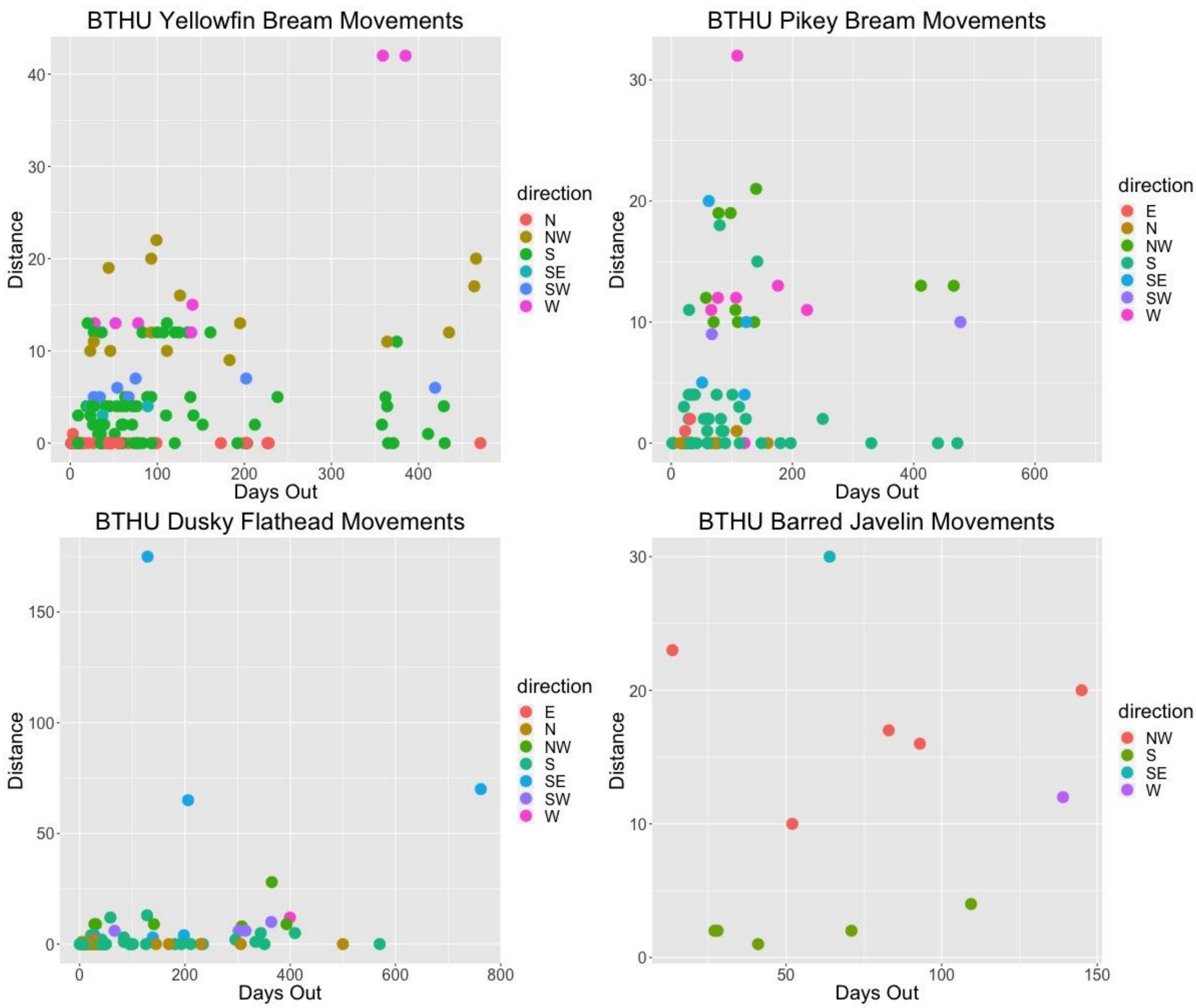


Figure 16: Distance (kms) moved and direction for 4 species released at the BTHU

Of the fish that spilled from Lake Awoonga in 2011, 27 fish were recaptured within 12 months of the spill occurring. Of those fish, 18 (66.7%) were recaptured in the Boyne River while 9 (33.3%) were recaptured in Gladstone Harbour or beyond.

The Barramundi that were recaptured less than 80km from Lake Awoonga in 2011 are displayed in Figure 17. Barramundi that moved further than 80km are not included as that is deemed beyond the Gladstone area.

There were further spills from 2012-2015 with most of those fish being recaptured in the Boyne River. There was no spill in 2016. A further spill of fish occurred in 2017. There were 38 recaptures through to May 2018 and 36 of those were in the year after the spill. All these fish were recaptured in the Boyne River as shown in Figure 18.

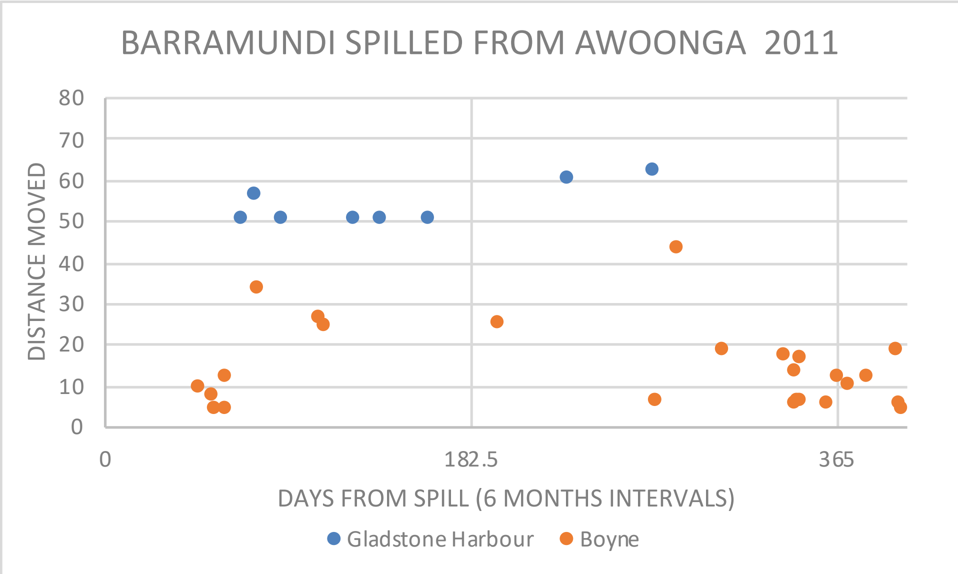


Figure 17: Distance moved (kms) by Barramundi that spilled from Lake Awoonga in 2011 and recaptured in the Boyne River or Gladstone Harbour

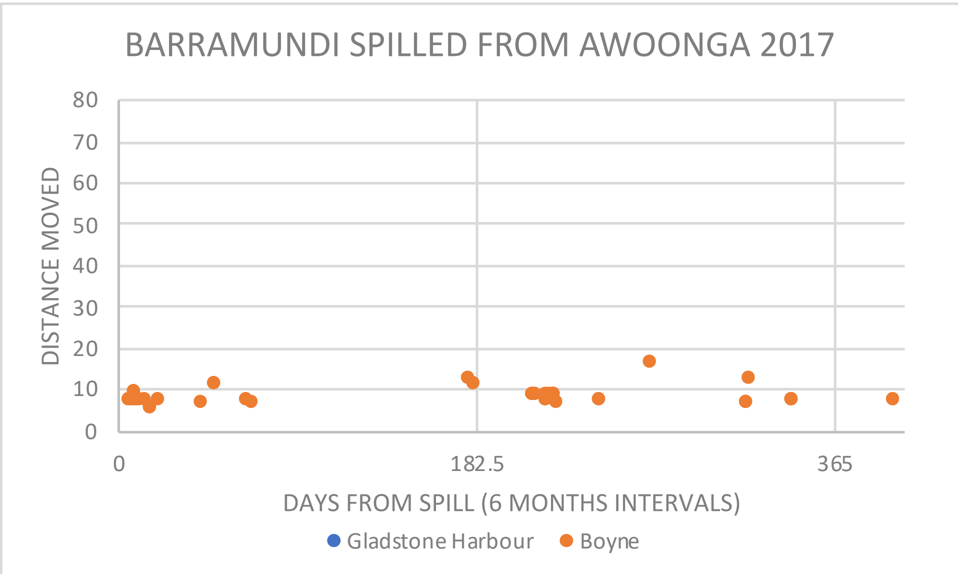


Figure 18: Distance moved (kms) by Barramundi that spilled from Lake Awoonga in 2017 and recaptured in the Boyne River or Gladstone Harbour

5. Discussion

There is a difference between developing a process to work in a laboratory environment and one to work in a real world context. In developing machine learning a number of additional considerations were integrated into developing the assessment process.

- Low ongoing costs
- Highest possible levels of automation with minimal human interactions
- Skills required to manage tasks such as training
- Scalability in terms of processing number of images
- Adaptability of the process to additional tasks

With this in mind, overall the project has taken longer than expected as the goal has been to develop a process that is ready to go live ongoing rather than being put into stasis as soon as the project is complete. Additional work has been undertaken to take the process to the point where it can be “live” in the sense of continuing training and developing better and better models. As a result, the formal outputs are not the finished products but rather indicative of the progress in training the models.

5.1 Training models

In each instance, a formal training strategy had to be developed. In a perfect world, more data should improve a model but this has not been the case. Hundreds of images taken in the same context improve the training for that context but produce less robust models overall when it comes to looking at new sources of images. Additionally, training in a single context reduces the level of robustness when minor changes occur in image data.

All of the models developed so far have been through 20+ iterations of development, improving and degrading as different training datasets are added and removed. Tail damage was initially treated as one model, whereas separating into two - one that looked at fin splits and one that looked at irregular damage, improved detection overall. A similar strategy will be deployed for fin damage.

For this reason, a longer term view has to be taken to developing a production training process to enable ongoing training.

5.2 Limitations on the model

The models used in this assessment have been exclusively trained on Yellowfin and Pikey Bream. This presents limitations when applying the model to other species, this was particularly noticeable on the CQU samples.

Simply adding additional species is not the sole answer as while a smaller number of additional species images will improve detection for non-target species, the model will be less robust overall and far less effective for Bream. For the model to be improved a similar number of images for each species to be assessed must be provided in order to maintain statistical stability of the input sample. Otherwise, sample biasing will occur, and training negatively affected.

Based on the existing library of images Barramundi, Australian Bass and King Threadfin are all available in sufficient numbers to provide a well-balanced sample. This should improve the model detection overall on additional species, though this is yet to be tested.

One important note – the more species included the less images are required per species. Overall a target sample of 2,000 images is ideal and achievable for fin and tail damage. Yolo trained models become increasingly reliable when trained with greater than 500 source images.

5.3 Model improvement

During the training process the importance of training cycles was assessed (how long the machine is trained for) verses input data. With relatively low numbers of images (<500) training time made little difference past 3,000 cycles as "variety" in the images was largely exhausted and the training forced to resample images in order to create new data. Past 1,000 source images longer training time does have an impact with a noticeable improvement in model performance with additional training.

5.4 A second method for "redness"

One of the key items to detect in the health assessment has been wounds, lesions and redspot or "redness". This worked well in the initial TensorFlow based model where limited available training data was supplemented by the ability to slice an image into smaller images to generate training data. As image recognition has advanced the increased resolution for training data has removed the ability to use this strategy.

In order to get around this limitation, "redness" images were combined with general wound images to increase the training data available. This resulted in an unstable model that performed very poorly. Several additional efforts were made to improve the model but the lack of training data remained an ongoing issue.

As a last effort, the original model was tested using the CQU data and this provided a reliable detector of "redness" items such as blood, redspot or lesions, while ignoring other red objects without differentiating between the source issue. Without any additional training above training conducted last year it achieved 86% agreement with human assessment that an issue was present.

As an interim step, this model is being updated to later versions of TensorFlow and provided with additional training to utilise the older model. This will provide a viable model for detecting a range of wound/"redness" issues.

This will enable the wounds training to be the focus of the more advanced Yolo models.

5.5 Improving fin split detection

Fin damage falls into two categories like tail damage. Overall detection of fin damage will be improved by splitting the model into two different models.

5.6 16:9 images, multi-detection and video

Yolo is capable of detecting multiple instances of an object in the same image or video. As the video resolution is 16:9, while square images are optimal for image recognition, this issue was also a factor in assessing images provided externally. This issue can be addressed through additional training data that incorporates images that are widescreen format, optimally in a separate model, or images provided for recognition can be cropped square. Which process will be better will be investigated.

5.7 Relative condition factor

The mean and median relative condition factors calculated for 2018 and the historic data showed some variation from year to year, as expected. The 2018 condition factors for Yellowfin and Pikey Bream were within the ranges calculated from the historic dataset suggesting that the condition of both species is acceptable. The growth of Pikey Bream was positively allometric (get rounder as they grow, in proportion to their length) on more occasions than Yellowfin Bream in the historic dataset and Pikey Bream are generally heavier than Yellowfin Bream at any given length. Considering both species occupy the same ecological niche, competition for resources may also result in one species condition being better than the other from year to year.

5.8 Potential vector for propagation of health issues

The recapture data suggests that the majority of fish do not move far in the first year at liberty with 68-89% of fish being recaptured within 2km of where tagged. This would suggest that any health issues are more likely to be spread within the local population initially. However there were sufficient numbers of fish 0-9% that moved greater than 20km. This would provide the opportunity for health issues to be more widely propagated.

Historically in 2011 Barramundi that spilled from Lake Awoonga spread rapidly beyond the Boyne River providing the opportunity for the spread of any health issues. However, based on recaptures, fish that spilled in 2017 up to a year later were only recaptured in the Boyne River. This would suggest that any health issues were most likely confined to the Boyne River. Following that spill the only reports of dead or dying fish have been from the Boyne River.

5.9 Potential to adapt methods to monitor fish health in other estuaries and ports of Australia

An objective of this project was to evaluate the potential to adapt the methods to other areas of Australia. Using a conventional research driven structured approach to collecting samples is unlikely to work as the protocols required would make it prohibitively expensive and would not likely be implemented in many, if any, jurisdictions.

A more innovative approach is required. Infofish considers that such a process has already begun even before the results of this project have been published.

Part of a suite of new technology tools that Infofish has developed is the Trackmyfish phone app that was used to collect the samples from GHHP zones and the BTHU. The phone app has been adapted to collect a broad range of data from catch and effort, tagging, fishing competitions etc. The fishing competition version can be tailored to any particular fishing competition format and has had a high rate of uptake with around 40 competitions having used or will use the app in the coming months. Competitions in all states except South Australia have already used the app. An essential feature of the app is that it is built around taking a photo of the fish.

This provides the opportunity to collect data from all around Australia at a very low cost. It also provides the opportunity to carry out a nationwide "audit" of fish health and fish handling issues that in turn can be used to educate fishers. Discussions with a number of competition organisers have indicated they are interested in including the collection of this data from their competitions. They see this as an important step in developing stewardship of fish resources through their competitions.

ABT tournaments for Bream are already conducted in all states except the Northern Territory and have shown interest in including collecting fish health data from their competitions. ABT is also considering establishing a Bream tournament in Gladstone with collecting health data as a specific objective.

This approach will also provide images at a low cost that can be used to improve the training models. To cover a broad range of species and issues will require a substantial library of images to be established and this is considered a cheap and quick method of improving that library.

Such an approach also has the potential to identify severe lesions or deformities that could be an indicator of more serious health issues that may require the use of pathology.

Applicability to other ports is an issue that required extension work. For the process to be "extendable" there needs to be some evidence that the process can be achieved in other locations using the same tools. Extension work has been conducted that demonstrates that the process can be replicated, thus the discussion needs to move to assessing need versus available resources.

Two audit tasks would need to be conducted:

- An audit of locations that might be needing assessment
- An audit of existing fishing activities that could be used for monitoring.

As part of the audit there should be a specific focus on catchments and ports that currently have or are developing environmental report cards. Most report cards will have fish indicators as part of their environmental assessment. While most of these will have different indicators, some will have included fish health. There is the potential to work with the managing organisations on the inclusion of visual fish health as an indicator.

Assuming an audit identified key locations and pilot sites some additional steps would be required.

- Activate the citizen science data collection process as the lowest cost process available
- A minimal useful sample would be at least 100 fish from a wide distribution geographically in the sample region (but preferably as large as possible)
- Detect if any issues are present in fish reported
- A trigger point needs to be defined on at whether this should lead to additional stratified sampling which requires many more sites to be assessed in extension. Regardless, any local scientist or other interested bodies would have an initial dataset to inform their approach or response, should any issues arise that cause concern.

In the short term it is unlikely that other organisations will want to develop their own capacity to assess fish health and will probably use the services of Infofish. As the process is developed and automated then other organisations may develop their own capacity if the demand for such a service is there.

6. Implications

For the first time there is the possibility of collecting large amounts of data on visual fish health and injury issues in a cost effective way. It is likely that a broad range of fishing competitions will see the benefit of having this data from their events. Initially the focus will likely be on injury as that is much more common than health issues. In turn having this data will allow competitions to improve their practices. For example, if fin or tail splitting is a common issue it is most likely the result of poor handling or use of inappropriate landing nets. Jaw injuries may be the result of poor use of fish grips or inappropriate hooks. This could provide the opportunity to “educate” competitors in relation to their fish handling practices.

Taking this approach is also much more likely to be successful as it is a bottom up approach that is already gaining acceptance from fishing competitions. A top down approach through the various fishing bureaucracies is unlikely to work as it would be hard to get commitment, take a long time to implement, be constrained by a range of protocols and would be inordinately costly.

However, the use of this technology will not be limited to health issues. Infofish has recently undertaken a project to assess fish resources in freshwater lagoons around Rockhampton. Conventional sounder technology was used to profile the bottom of the lagoons which showed up a plethora of Tilapia nests. The machine learning technology is now being used to identify and count the number of Tilapia nests. As well as providing an estimate of the size of the problem it may also be used in areas where Tilapia are less well established and improve our ability to estimate the size of the problem and its spread.

7. Recommendations

7.1 Developing Fish Health Indicators for the Gladstone Harbour Report Card

While this project was essentially about assessing machine learning tools in the identification of fish health issues the longer term objective for the GHHP was to identify appropriate measures that could be used for reporting on fish health in the report card for Gladstone Harbour.

There are a number of areas of fish health that could be considered for inclusion in the report card:

- The physical condition of fish using data from the BTHU
- The effect of visible heath issues, particularly lesions and fin damage using fish collected during recruitment surveys, at BTHU or other fishing events
- The incidence of dead and sick fish spilling from Lake Awoonga by an annual survey in February-March or after a spill of the lake and tagging of fish to track distribution of spilt fish
- The safety of fish for human consumption

This project potentially provides measures that can be used to obtain indicators addressing the first 2 dot points and Infofish has collected data to assist with the development of the third dot point.

7.2 Rationale for suggested Fish Health Indicators

The indicator that can be used for the physical condition of the fish is the relative fish condition. The condition of the fish will be related to such things as water quality, food supply and underlying health issues.

There are already established methodologies for assessing fish condition and this project has used the relative fish condition factor. Also the BTHU has collected the necessary data to calculate fish condition since 2003 so that there is already a chain of historic data that is available. The BTHU also provides a low cost avenue for obtaining the data on fish condition.

Also as fish that are presented at the BTHU come from all over the Gladstone area this would provide samples from most, if not all, zones however the number of samples would not be uniform from each zone. An assessment of the fish collected from GHHP zones and from the BTHU suggested that there were no anomalies between the datasets. It would be an advantage if the GHHP zones where BTHU fish come from is recorded to look at the difference in sample sizes from the zones.

While the incidence of lesions in the fish sampled was not high it would be a useful measure as historically there have been years when “redspot” lesions are common and affect a wide range of species. Lesions are also found on Barramundi that have lost scales from going over the spillway at Awoonga.

There is an opportunity to also collect fish with lesions at the BTHU however fishers would need to be encouraged to take photos or bring them in as often, if the lesions are severe, they would prefer to release the fish where caught. A supplementary method of collecting samples would be during recruitment surveys. There is no accepted methodology or protocols for assessing lesions or how an indicator could be derived so that would need to be developed.

Infofish is currently examining up to 14,000 photos for the whole range of fish health issues. These photos are from different locations around Australia and over different time scales. The aim is to compare the incidence of health issues, such as lesions, from all locations with the aim of establishing a baseline that can be used.

Assessing Barramundi in Lake Awoonga and fish that spill from the lake should be another area for consideration of the development of an indicator. Given the high level of fishing effort below the dam in the Boyne River after a spill this is very much in the spotlight. Infofish receives regular reports from fishers and the general public with many of those reports including photos.

Tagging of fish in Lake Awoonga and in the Boyne River below the dam has helped understand the distribution of fish that go over the dam spillway. Many of the fish that go over the spillway die or suffer injuries as a result of the trauma of negotiating the spillway and the rocks at the bottom. Much of the trauma is scale loss from abrasion from the concrete spillway and subsequent infection.

The difference in fish distribution from the 2011 and 2017 spills provides useful data on the possible propagation of fish health issues. However, capturing that into a fish health indicator will require further consideration.

7.3 Further development considerations for FRDC

The project has shown that the technology can be applied in assessing fish health issues and is ready to be used in a productions system.

There are a number of areas that FRDC may consider for investment in further development.

- Extending the training models to more species and issues such as wounds and milky eyes
- An assessment of potential end user requirements to focus further development
- Upscaling the technology and automating the process to meet the needs of end users
- Testing of the technology in an Australia wide context using data collected during fish competitions to identify areas where fish health issues may need further attention
- Evaluate whether visual fish health is an appropriate indicator for environment report cards where they are being used around Australia
- An assessment of other areas of fishery monitoring/data collection where the technology could be applied in conjunction with other new technologies

7.4 Rationale for further development

Infofish has already progressed well beyond the scope of this project having moved the process from an evaluation of technology to the development of a production system. That was always the aim as too many projects fail to make that transition.

The models to date have focussed on Bream. They are currently being expanded to a number of other species with sufficient images for Barramundi, King Threadfin and Australian Bass. The range of species where health issues are being assessed has also been expanded. The aim is to develop generalised models for each of the health issues where any species can be assessed.

For the next few months through to the end of the year machine and human assessment will be undertaken in parallel. By the end of the year the aim is to have the process fully automated and limit the human assessment to random audits.

An important outcome of the project will be the acceptance of fish health indicators for the report card for Gladstone Harbour. There are many areas around Australia where environmental report cards are being used. Most use some form of fish indicator and some have fish health as an indicator or potential indicator. However, it is likely that there is no consistent approach on the development of those indicators. There is an opportunity to provide a pathway for the inclusion of visual health as an indicator.

8. Conclusions

The emphasis has been on the development of a process that will work in the real world so that considerably more work has been undertaken than originally envisaged. That work will continue beyond this project.

The project has demonstrated that machine learning technology can be applied to assess visual health issues. The tools that have been used will continue to evolve and improve and are likely to be used in other fisheries areas beside health and in other disciplines. Infofish sees the value in further developing these technologies and will be using them in building improved business practices.

Based on the training models developed to date the results from the human and machine assessments were acceptable. With more images and further development of training models the results will continue to improve and be applied to an ever-growing number of species and issues.

While not a specific objective of the project it was considered necessary to provide additional information so that GHHP could better evaluate what it could use as indicators of fish health for its report card on Gladstone Harbour. The data presented on fish condition and a potential vector for propagation of health issues provides additional information that GHHP can use.

9. Extension and Adoption

As this project has dealt with new technology that has not had much exposure within the fishing industry the extension has initially been around testing the reaction of recreational fishers to the concept and assessing what sort of roles it can play in the real world.

With the uptake of the competition version of the Trackmyfish app it seemed logical to seek the views of competition organisers to see what level of interest there may be in collecting injury and health data in competitions. The reaction from competition organisers has been strongly positive. Data on injuries is seen to provide the opportunity to improve fish handling and practices eg split tails and split fins are generally the result of using knotted landing nets so data on this can be used to educate participants. This also provides the opportunity for competitions to demonstrate stewardship which is increasing seen as being important to maintain their social license.

It is considered important that there is initial “on the ground” uptake, primarily through competitions, before tackling the institutional level. At the institutional level, being new technology, it is likely to be greeted with at least scepticism, if not resistance. Also until FRDC and GHHP have evaluated the project there is little to be gained in moving forward with institutional extension as that will likely be influenced by the feedback received.

In terms of adoption Infofish has already decided to take the technology forward into a production mode. However it is likely that other businesses will also see the benefits of the technology, if not in the area of fish health, then in other areas of fisheries.

Figure 19 shows locations where over 100 images (red) and 50 images (blue) have been obtained through the Trackmyfish app and Suntag that are being and will be assessed for fish health. There is a total of 92 locations with approximately 9,200 images of which around 5,000 have been processed. Most images are from Queensland however images are also available from the Northern Territory, Western Australia and New South Wales.

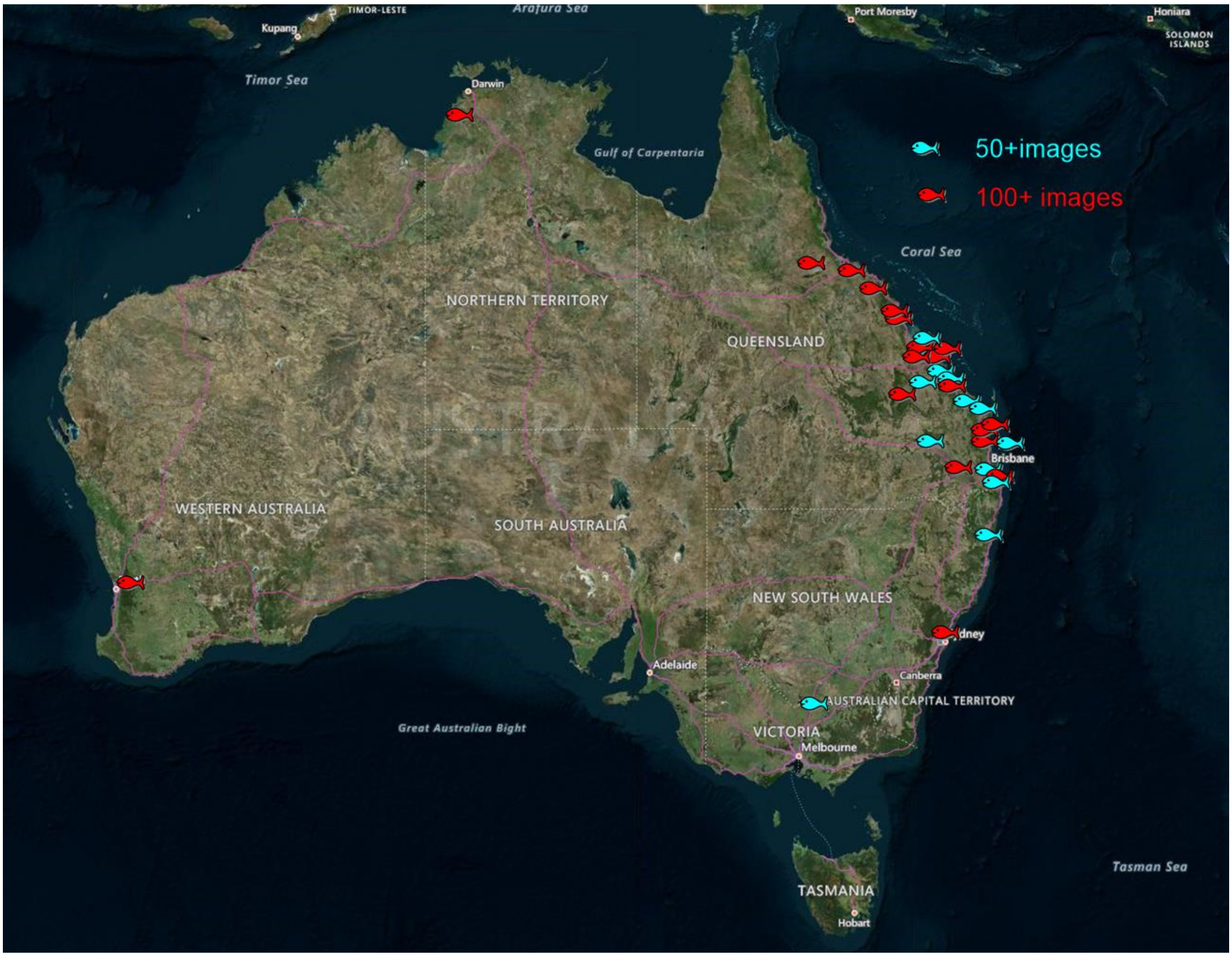


Figure 19: Sites around Australia where images have been collected through Trackmyfish and Suntag suitable for health assessment

9.1 Project coverage

There was an initial media release put out on the project titled "Innovation in delivering fisheries information" put out in March 2018. It was provided to all local television, radio and print outlets in Rockhampton and Gladstone however was only taken up by the Morning Bulletin newspaper in Rockhampton.

10. Project materials developed

The only product developed to date has been the fish health data collection version of the Trackmyfish phone app that was developed to collect the field samples in the GHHP reporting zones and at the Boyne Tannum HookUp. This version of the phone app was for internal use only. A public version of the app was developed that would allow the public to report fish health issues. However due to public perception concerns GHHP did not proceed with the release of that version.

11. Appendices

11.1 Historical condition data from the BTHU

Table 11: Yellowfin Bream BTHU historical data

Year	n	b Value	Growth Type	R2	Relative Condition (Kn) Summary Statistics			
					Min.	Max.	Mean	Median
2003	153	2.95	allometric (-)	0.95	0.76	1.23	1.00	1.00
2004	192	2.82	allometric (-)	0.89	0.41	1.46	1.01	1.02
2005	188	2.89	allometric (-)	0.85	0.42	1.62	1.01	1.00
2006	176	2.76	allometric (-)	0.94	0.77	1.40	1.01	1.00
2007	126	3.06	allometric (+)	0.96	0.81	1.21	1.00	1.00
2008	63	2.94	allometric (-)	0.94	0.81	1.58	1.01	1.00
2010	137	3.04	allometric (-)	0.94	0.71	1.23	1.01	1.01
2012	150	2.90	allometric (-)	0.92	0.71	1.35	1.01	0.99
2013	179	2.86	allometric (-)	0.92	0.81	1.70	1.01	0.99
2014	103	2.81	allometric (-)	0.92	0.74	1.39	1.01	1.00
2015	361	2.86	allometric (-)	0.92	0.58	1.61	1.01	1.00
2016	181	2.81	allometric (-)	0.92	0.82	1.37	1.01	1.00
2017	451	2.76	allometric (-)	0.88	0.61	1.58	1.01	1.00

Table 12: Pikey Bream BTHU historical data

Year	n	b Value	Growth Type	R2	Relative Condition (Kn) Summary Statistics			
					Min.	Max.	Mean	Median
2003	56	2.95	allometric (-)	0.95	0.78	1.17	1.00	1.01
2004	50	3.12	allometric (+)	0.94	0.77	1.22	1.01	0.99
2005	71	2.94	allometric (-)	0.89	0.73	1.34	1.01	0.99
2006	65	2.73	allometric (-)	0.81	0.75	1.33	1.01	1.00
2007	75	2.93	allometric (-)	0.94	0.72	1.26	1.01	1.02
2008	35	3.05	allometric (+)	0.89	0.62	1.19	1.01	1.03
2010	23	3.08	allometric (+)	0.88	0.84	1.25	1.00	1.00
2012	48	2.94	allometric (-)	0.91	0.77	1.30	1.01	1.00
2013	82	3.00	allometric (-)	0.86	0.76	1.43	1.01	1.00
2014	64	2.98	allometric (-)	0.90	0.65	1.39	1.01	1.01
2015	89	3.03	allometric (+)	0.90	0.72	1.28	1.01	1.02
2016	63	2.76	allometric (-)	0.85	0.74	1.20	1.01	1.00
2017	87	2.72	allometric (-)	0.86	0.72	1.26	1.01	1.02

11.2 11.2 Project staff

Principal Investigator: Bill Sawynok Infofish Australia Pty Ltd
Principal Technology Investigator: Stefan Sawynok Infofish Australia Pty Ltd
Fish Condition Investigator: Aaron Dunlop Infofish Australia Pty Ltd

Photo and Video Technician: Phoenix Sawynok Infofish Australia Pty Ltd
Finance Manager: Shirley Sawynok Infofish Australia Pty Ltd

11.3 11.3 Intellectual Property

Prior to the commencement of the project there was an Intellectual Property Agreement drawn up between Infofish Australia Pty Ltd, FRDC and GHHP.

12. References

1. Dennis M et al. *Pathology of finfish and mud crabs Scylla serrata during a mortality event associated with a harbour development project in Port Curtis, Australia.* Diseases of Aquatic Organisms, 2016. **121**(3): p. 173-188.

2. Gladstone Ports Corporation. *Western Basin Dredging and Disposal Project, Annual Compliance Report of Environment Protection and Biodiversity Conservation Act (EPBC) 2009/4904: For the attention of: Department of the Environment and Energy*. 2016 [cited 2017 5 June]; Available from: http://gpcl.com.au/EnvironmentDocuments/WBDDP_Annual_Compliance_Report_2016.PDF

3. Guidelli, G., Tavechio, W.L.G., Takemoto, R.M. & Pavanelli, G.C., 2011, 'Relative condition factor and parasitism in anostomid fishes from the floodplain of the Upper Parana River, Brazil', *Veterinary Parasitology* 177: 145-151.

4. LeCren, E.D., 1951, 'The length–weight relationship and seasonal cycle in gonad weight and condition in the perch Perca fluviatilis', *J. Anim. Ecol* 20: 201-219.

5. Peig, J. & Green, A.J., 2010, 'The paradigm of body condition: a critical reappraisal of current methods based on mass and length', *Functional Ecology* 24: 1323-1332.

6. Swingle, W.E. & Shell, E.W., 1971, 'Tables for computing relative conditions of some common freshwater fishes', *Alabama Agricultural Experiment Station, Auburn University, Circular* 183.

7. Sawynok B, Platten J, Parsons W, Sawynok S, 2014, *Gladfish 2014 Assessing Trends in Recreational Fishing in Gladstone Harbour and Adjacent Waterways*, Infofish Australia report: 75-77