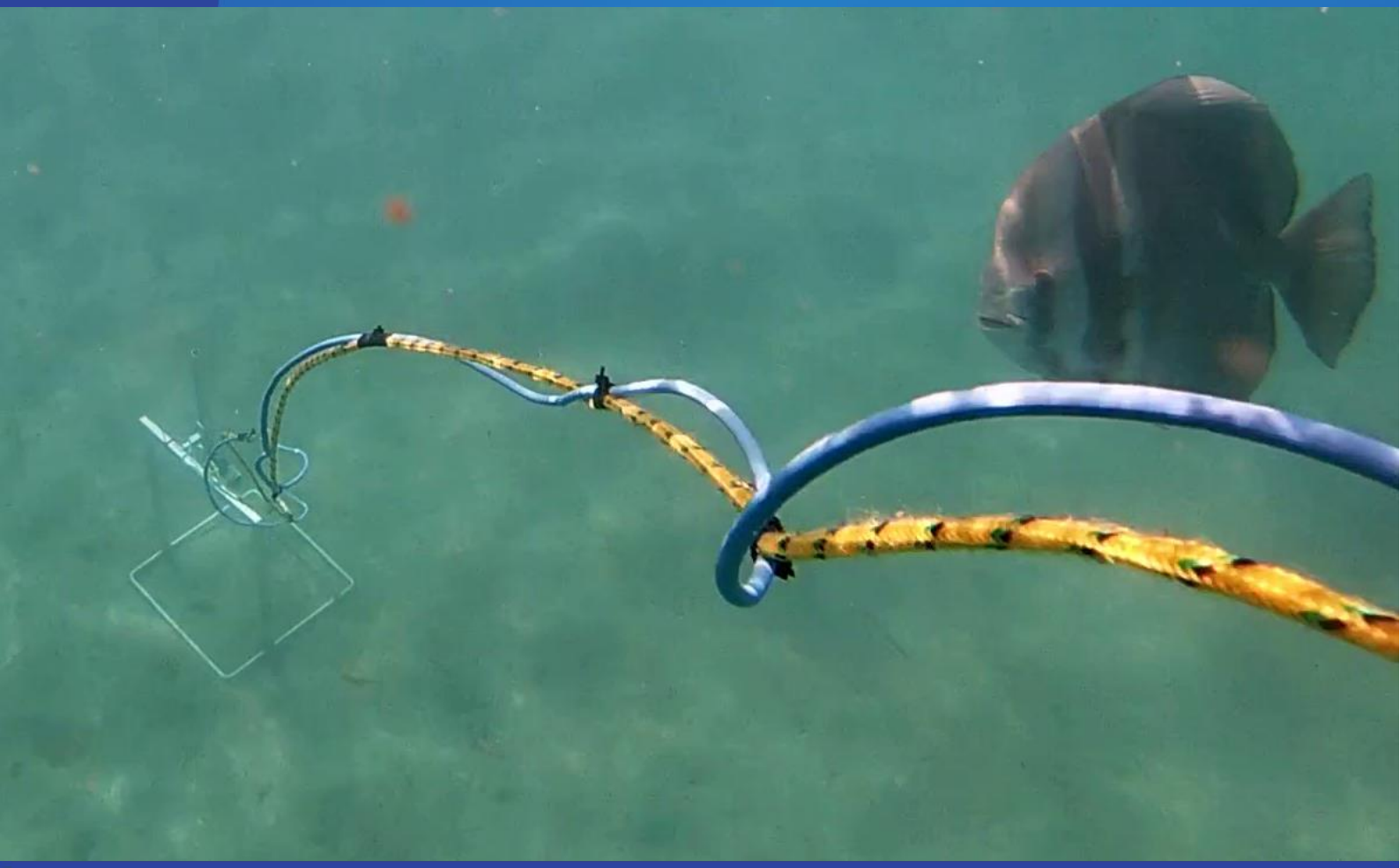
TropWATER
Centre for Tropical Water and Aquatic Ecosystem Research

JCU

Subtidal Seagrass Detector

July 2022 | Report No. 22/39



Authored by: Langlois, L.A., Collier, C.J. and McKenzie, L.J.

Subtidal Seagrass Detector:

Development and preliminary validation

Centre for Tropical Water & Aquatic Ecosystem Research (TropWATER) James Cook University

Townsville Phone : (07) 4781 4262 Email: TropWATER@jcu.edu.au

Web: www.jcu.edu.au/tropwater/

The report may be cited as
Langlois, Lucas A., Collier, Catherine J. and McKenzie, Len J. (2022) Subtidal Seagrass Detector: Development and preliminary validation. TropWATER Publication 22/39, Centre for Tropical Water & Aquatic Ecosystem Research, James Cook University, Cairns, 24 pp.

Contacts
For more information contact: Lucas Langlois, lucas.langlois@jcu.edu.au and +61 7 4232 1857

Contents

Executive Summary

This report presents the development and evaluation of a Subtidal Seagrass Detector (the Detector). Deep learning models were used to detect most forms of seagrass occurring in the northeast Australian seascape from underwater images and classify them based on how much seagrass was present. Images were collected by scientists and trained citizen scientists undertaking routine monitoring using drop-cameras mounted over a 50 x 50 cm quadrat. The Detector is composed of three separate models able to perform the specific tasks of: detecting the presence of seagrass (*Model #1*); classify the seagrass present into three broad cover classes (low, medium, high) (*Model #2*); and classify the substrate or image complexity (simple of complex) (*Model #3*). We were able to successfully train the three models to achieve high level accuracies with 97%, 80.7% and 97.9%, respectively. With the ability to further refine and train these models with newly acquired images from different location and from different sources (e.g. ROV), we are confident that our ability to detect seagrass will improve over time. With this tool we will be able rapidly assess a large number of images collected by various contributors, such as citizen scientists, QPWS Rangers and Indigenous rangers that frequently access the Reef and seagrass habitats of northern Australia. This would provide invaluable insights about the extent and condition of subtidal seagrass in currently data-poor areas.

1. Introduction

Seagrasses are one of the most valuable marine ecosystems on the planet, and an integral component of the northeast Australian seascape that includes: the Great Barrier Reef, Torres Strait, and the Great Sandy Marine Park. Seagrass ecosystems in these marine domains are ecologically, socially and culturally connected and contain values of national and international significance (Johnson *et al.*, 2018).

The Great Barrier Reef (the Reef) is the most extensive reef system in the world, in which seagrass is estimated to cover approximately 35,679 km^2 (McKenzie *et al.*, 2022b). Over 90% of the Reef's seagrass meadows occur in subtidal waters, with the deepest record to 76 m (Carter *et al.*, 2021b), although most field surveys are in depths shallower than 15 m (McKenzie *et al.*, 2022b). There are 15 seagrass species reported within the Reef, occurring in estuaries, coastal, reef and deep water habitats and forming meadows comprised of different mixes of species (Carter *et al.*, 2021b). Seagrass ecosystems of the Reef support a range of goods and benefits to species of conservation interest and society. The seagrass habitats of Torres Strait to the north are also of national significance due to their large extent, diversity and the vital role they play to ecology and the cultural economy of the region (Carter *et al.*, 2021a). Similarly, the seagrasses within the Great Sandy Marine Park to the south support internationally important wetlands, highly valued fisheries and the extensive subtidal meadows in Hervey Bay are critical for marine turtles and the second largest dugong population in eastern Australia (Preen *et al.*, 1995; McKenzie *et al.*, 2000). Catchment and coastal development, climate change and extreme weather events threaten seagrass ecosystem resilience and drive periodic decline. Maintaining up-to-date information on the distribution and condition of seagrass meadows is needed to protect and restore seagrass ecosystems.

A wide range of methods have been applied to assess and monitor changes in subtidal seagrass, including free-diving, SCUBA diving, towed camera, towed sled, grabs or drop–camera (McKenzie *et al.*, 2022b). Most of these techniques rely on trained scientists to visually confirm, quantify and identify the presence of seagrass in situ. This labour-intensive work, combined with the tremendously large area of the Reef, makes assessing the state (extent and condition) of subtidal seagrass prohibitively time consuming and expensive.

In recent years, the use of digital cameras and autonomous underwater vehicles (AUVs) has led to an exponential increase in availability of underwater imagery. When this imagery is geotagged or geolocated, it provides an invaluable resource for spatial assessments, and when collected by a range of providers and the wider community who are accessing the Reef for a range of other activities (tourism, Reef management), is highly cost effective. For example, the Queensland Parks and Wildlife Service uses drop cameras to collect photoquadrats of the benthos within seagrass habitats for processing by and inclusion in the Inshore Seagrass component of the GBR Marine Monitoring Program (MMP). Recent projects such as The Great Reef Census aim at tapping into the power of citizen science to collect images and provide new sources of information about the Reef. A similar approach could be applied to seagrass. This digital data can be analysed automatically if the workflows are in place to deal with structured big data streams.

Deep–learning technology provides potentially unprecedented opportunities to increase efficiency for the analyses of underwater images. Deep–learning models are being used for counting fish (Sheaves *et al.*, 2020), identifying species of plankton (Schröder *et al.*, 2020) and estimating coral cover (Beijbom *et al.*, 2015). Few studies showed the potential of its application for coverage estimation (Reus *et al.*, 2018) as well as detection and classification (Moniruzzaman *et al.*, 2019; Raine *et al.*, 2020). While these showed interesting technical methods, they were not developed for operation applications. An

operational model that can detect seagrass within the Reef will expand the ability to rapidly assess and easily provide data critical for large scale assessments. In particular, a model that can detect seagrass presence even with diverse physical appearances among the 15 species in the Reef, and in a range of habitat types with variable benthic substrates. As seagrass can also be very sparse in the Reef, with an historic baseline of 22.6 ±1.2% cover (McKenzie *et al.*, 2015) and subtidal percent covers frequently less than 10%, a detector is needed to cope with such circumstances.

In this report we will detail the development of a Subtidal Seagrass Detector (the Detector) using a deep neural network (DNN) to analyse underwater images to detect and classify seagrasses. This will enable rapid processing of many images. It will form an integral step in workflow from image capture to provision of rapidly and easily accessed information. Up to date information on the extent and condition of seagrass is required for marine spatial planning and for the implementation of other management responses to protect Reef ecosystems.

2. Methods

2.1. Dataset

Our subtidal image dataset was composed of 7440 photoquadrats collected by drop-camera and SCUBA divers as part as the Great Barrier Reef Marine Monitoring Program (MMP) (McKenzie *et al.*, 2022a), the Seagrass-Watch Global Seagrass Observing Network (Seagrass-Watch, 2022) and the Torres Strait Ranger Subtidal Monitoring Program (Carter *et al.*, 2021a). Images were captured between 2014 and 2021 from 28 sites across 18 unique locations within the coastal and reef subtidal habitats from Torres Strait to Hervey Bay (Figure 1, Table 1). Images were annotated by assessing: (1) the percent cover of seagrass (McKenzie *et al.*, 2003), (2) the seagrass morphology of the dominant species based on largest percent cover (straplike, oval–shaped or fernlike), (3) percent cover of algae, (4) substrate complexity (simple or complex), and (5) quality of the photo (0=photo unusable, 1=photo clear with more than 90% of quadrat in the frame, 2=photo with bad visibility with more than 90% of quadrat in the frame, 3= photo clear with quadrat partially not visible, 4= photo oblique with quadrat not totally on the bottom). Only photos with a rating of 1 (5782 in total) were retained to ensure optimal performance. All images were cropped to the outer boundary of the quadrat and standardised to a 1024 × 1024 pixel size.

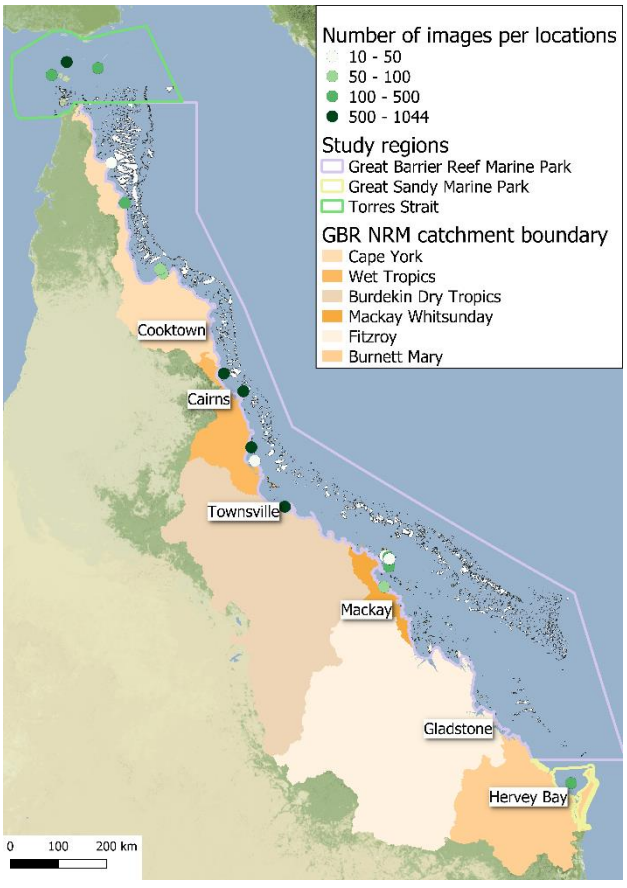


Figure 1. Map showing the location and number of images used for the Subtidal Seagrass Detector in the Torres Strait, the Great Barrier Reef World Heritage Area and Great Sandy Marine Park.

2.2. Detector models

2.2.1. Seagrass presence detector *(Model #1)*

We defined seagrass presence as an area of benthos spatially dominated by seagrass, which we classed as ≥3% cover (*sensu* Mount *et al.*, 2007). Images with seagrass cover less than 3% were excluded, resulting in the removal of an additional 819 images from the analysis. This maximised the power of detection to levels where seagrass was clearly visible. There were 1727 images with seagrass absent and 3236 with seagrass present. To ensure a balance dataset of the two classes 1727 images were chosen at random out of the 3236 while ensuring the inclusion of all images from the minor seagrass morphology classes oval-shaped (522) and fernlike (165). The remaining images with seagrass present (1509) were retained for further testing.

2.2.2. Seagrass cover category classifier *(Model #2)*

Cover categories were first established based on four cover quantiles, which were equivalent to seagrass percent cover categories of ≥3 <9%, ≥9 <15%, ≥15 <30% and ≥30%. However, the resulting model did not adequately distinguish between the two middle categories (less than 60% accuracy). Therefore, those two classes were merged resulting in three main classes used in *Model #2*: (1) low seagrass cover (≥3 <10%), (2) medium seagrass cover (≥10 <30%), and (3) high seagrass cover (≥30%) (Figure 2). The classes were somewhat unbalanced with 1082, 1509 and 644 images respectively. However, a few more images in the medium class can be beneficial as it can help improve accuracy for that class which is the most commonly occurring at MMP sites (long term mean of 14% seagrass cover for coastal and reef subtidal sites).

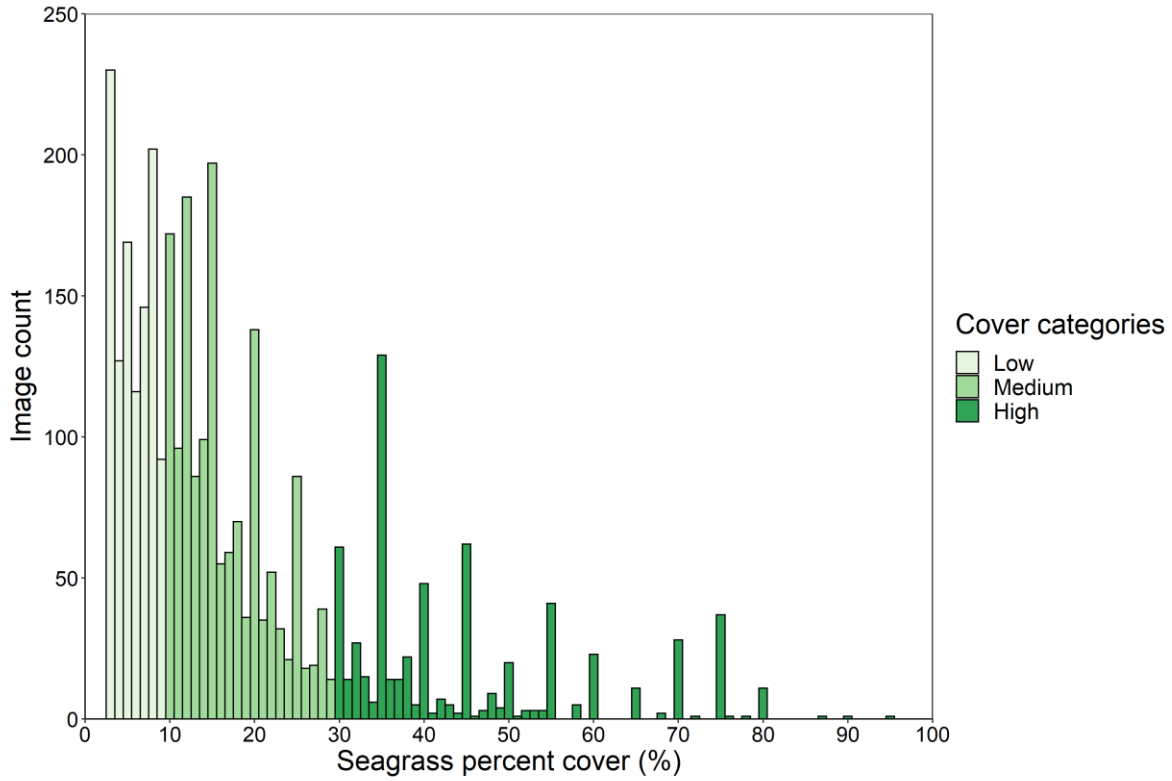


Figure 2. Distribution of seagrass percent cover in the image dataset used for *Model #2*.

2.2.3. Substrate complexity classifier *(Model #3)*

The substrate complexity classifier was applied to all images without any seagrass present. Those images were labelled either as 'simple substrate' or as 'complex substrate'. The 'simple' category was assigned to clear images with mostly sandy bottoms while the 'complex' category was assigned to images that met at least one of the following conditions:

- had unconsolidated substrates, such as rock, live coral or coral rubble
- had a visually significant amount of macroalgae
- where labelling was difficult (e.g. poor visibility, small seagrass species, poor image contrast).

Out of the 1727 images without seagrass, 1129 had simple substrate and 598 had complex substrate. Similar to *Model #1*, a random 531 simple substrate images were excluded and retained for further testing to unsure a balance dataset during training. This classifier can provide a potential reason for the absence of seagrass as well as highlighting potential shortfall in the seagrass detection from *Model #1*. In complex substrate habitats, seagrass could be present, however, percent cover is most likely to be low (<10%) and particularly difficult to detect by the model. Images predicted into this category can be later manually inspected to confirm the absence of seagrass.

All three final datasets were split 60-20-20 into a training, validation and test set.

2.3. Deep Neural Network (DNN) modelling

2.3.1. Image classification workflow

Our aim for this study was to develop a Detector that would be able to achieve three separate classification tasks: (1) detect the presence/absence of seagrass, (2) estimate the seagrass cover (low, medium or high), and (3) identify the level of complexity of the substrate (simple or complex). Separate deep–learning models were developed to execute each of these tasks independently which maximised model accuracy and reduced category imbalance (Figure 3).

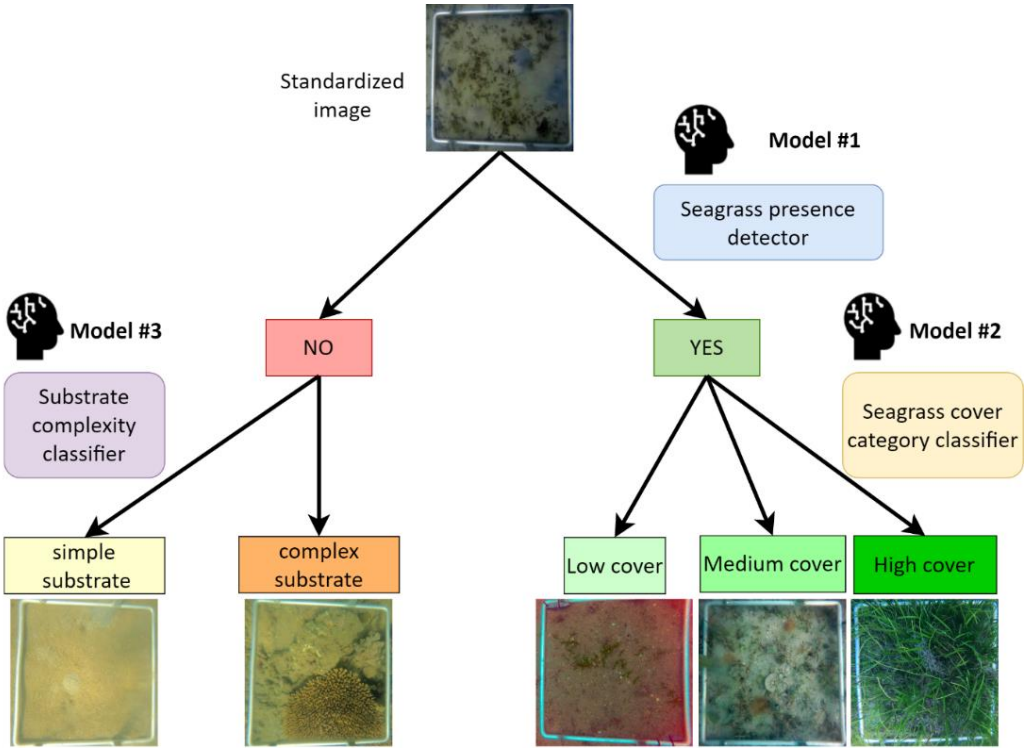


Figure 3. Diagram detailing the image classification workflow process of the Detector with the three deep–learning models involved

2.3.2. Model architecture

The classification models were composed of a binary classification model for *Model #1* and *Model #3* and multiclass classification for *Model #2*. The classification employed deep–learning and more specifically deep neural networks (DNNs). Training a neural network can be a protracted process and requires a large number of images to achieve satisfactory results. Transfer learning has been developed where an already successfully trained network such as VGG16 can be used as a feature extractor and coupled with a new classifier trained for the new specific task Tammina, 2019. Our network was composed of a VGG16 model pre-trained on the ImageNet classification tasks Zhang *et al.*, 2015. Instead of the final dense layer from the original VGG16 model, we created our own custom classier composed of a sequence of two fully connected layers (with 512 nodes and ReLU Agarap, 2018 activation), two consecutive dropout Srivastava *et al.*, 2014 (probability of 0.05 and 0.5) to prevent overfitting and a final dense layer with one node for each of predicted class activated by either the Sigmoid or Softmax function (Figure 4).

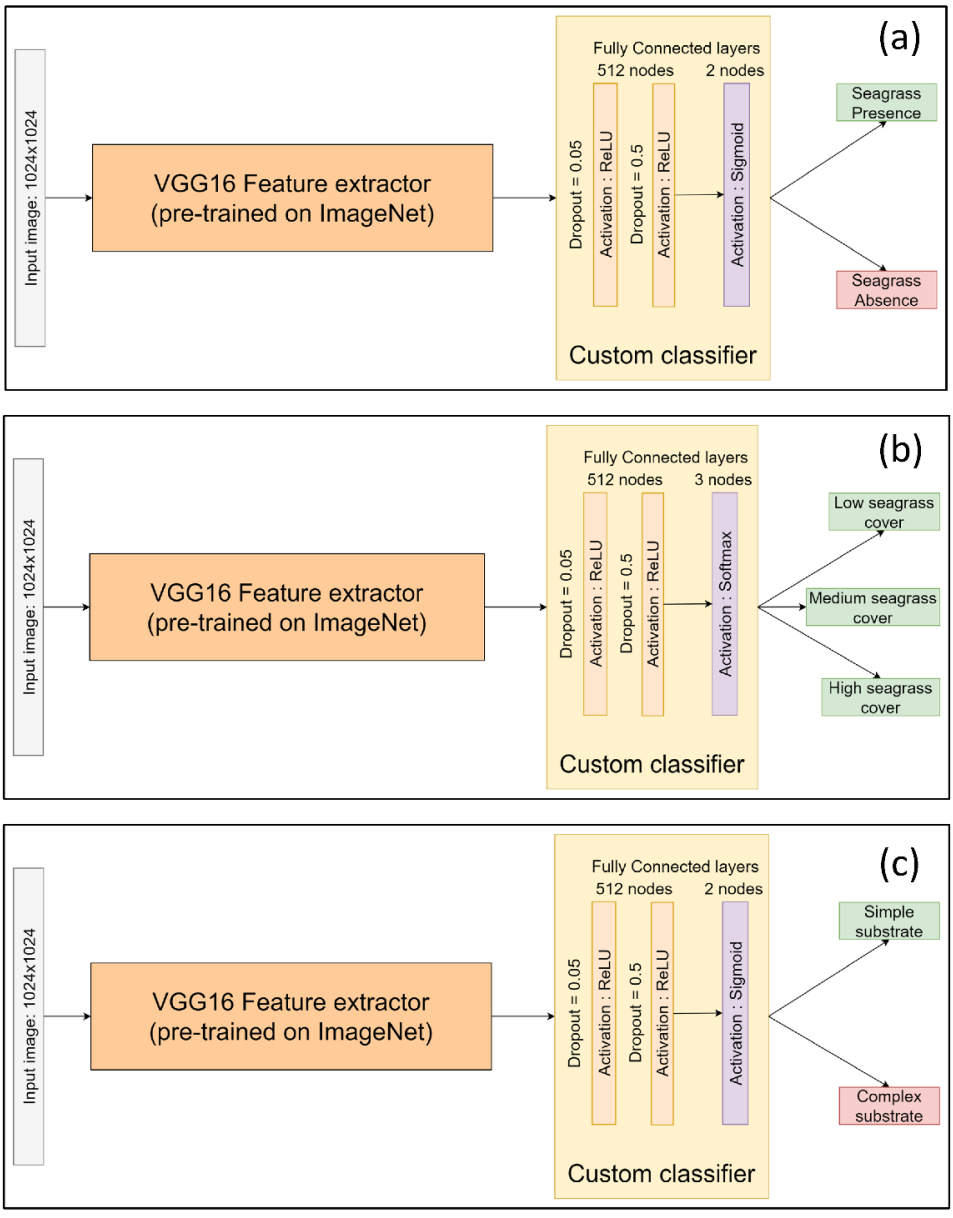


Figure 4. Convolution neural network architecture of: (a) *Model #1*, (b) *Model #2* and (c) *Model #3*.

Contrary to other studies (Raine *et al.*, 2020) we chose not to split our original images as it would have meant having to create new labels for thousands of sub-images. Instead, the input image size was increased. After multiple trials we found that optimal results were achieved for the input size of 1024x1024 pixels. We also tried more complex networks such as Resnet50, Xception and EfficientNet but they did not perform nearly as well and were much slower to train.

2.3.3. Model training

The DNNs were all trained on batches of eight random images per training iteration. When the DNN has gone through as many iterations as needed to process the full training image set, this constitutes an epoch. Throughout the whole training process, the progress of the learning is monitored by evaluating the model performance on the validation image set.

We started with an initial training phase where only the final classification layers (custom classifier part) were trainable and the rest of the VGG16 layers were frozen. During this phase the Adam optimizer Kingma and Ba, 2014 was used with an initial learning rate of 0.001. If the loss on the validation image set did not improve after 10 epochs the learning rate was reduced by half up to four times after which the training was stopped. That process lasted 60 to 68 epochs. A fine-tuning training phase followed, where the VGG16 layers were unfrozen and set as trainable. This was done over 100 epochs and with the RMSprop optimizer Tieleman and Hinton, 2014 and a much slower learning rate of 0.00001. The fine-tuning is meant to ensure the feature extraction is optimised for our input size as well as increasing performance of the models.

To further prevent overfitting and best capture, the potential illumination and turbidity variations of underwater images, colour-based data augmentation was applied where brightness (-70 to 70), contrast (0.1 to 0.3), blur (sigma 0 to 0.5) and the red channel (-50 to 50) were randomly altered at each training iteration.

2.3.4. Model evaluation (testing)

The training process stopped once all the DNNs have reached a plateau where further training did note further improve performances on the validation set.

We then conducted final evaluation of the model performance on the test image set (20% of the total) where a detailed accuracy assessment was done. For *Model #1* and *Model #3*, further testing was conducted by running the model on the remaining images not included in the training, validation and test sets.

3. Results

3.1. Model #1

Model #1 achieved 97.0% accuracy (Table 2) on the test image set (691). We had 3 false positive and 18 false negative classifications (Figure 5, Table 3a). The false positives were all images from Low Isles and taken on SCUBA. We suspect that the presence of turf algae and the low image quality could be the source of the misclassification. The small number of false positives suggests the model was not overestimating seagrass presence.

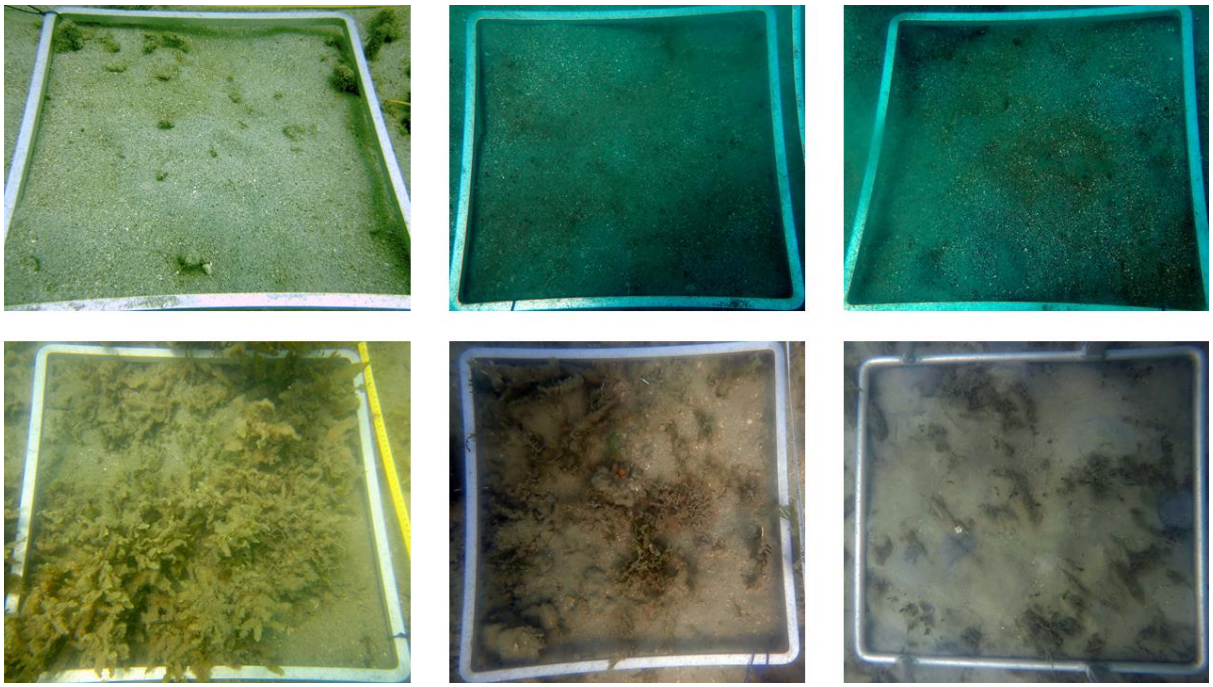


Figure 5. Examples of images misclassified by Model #1 with false positives on top row and false negative on the bottom row.

Of the false negative images, 16 had a percent cover lower than 10% and in nine of these percent cover was lower than 5% (Figure 6). In addition, 14 of the false negative images had a complex substrate with seven having more than 15% algae cover. This was further confirmed by running the model on the remaining seagrass photos not included in the training, validation and test sets. The model failed to detect seagrass in 38 out of 1509 images, achieving 97.4% accuracy. A similar pattern was observed where 31 of the misclassified images had less than 10% seagrass cover and 33 had complex substrate (Figure 6).

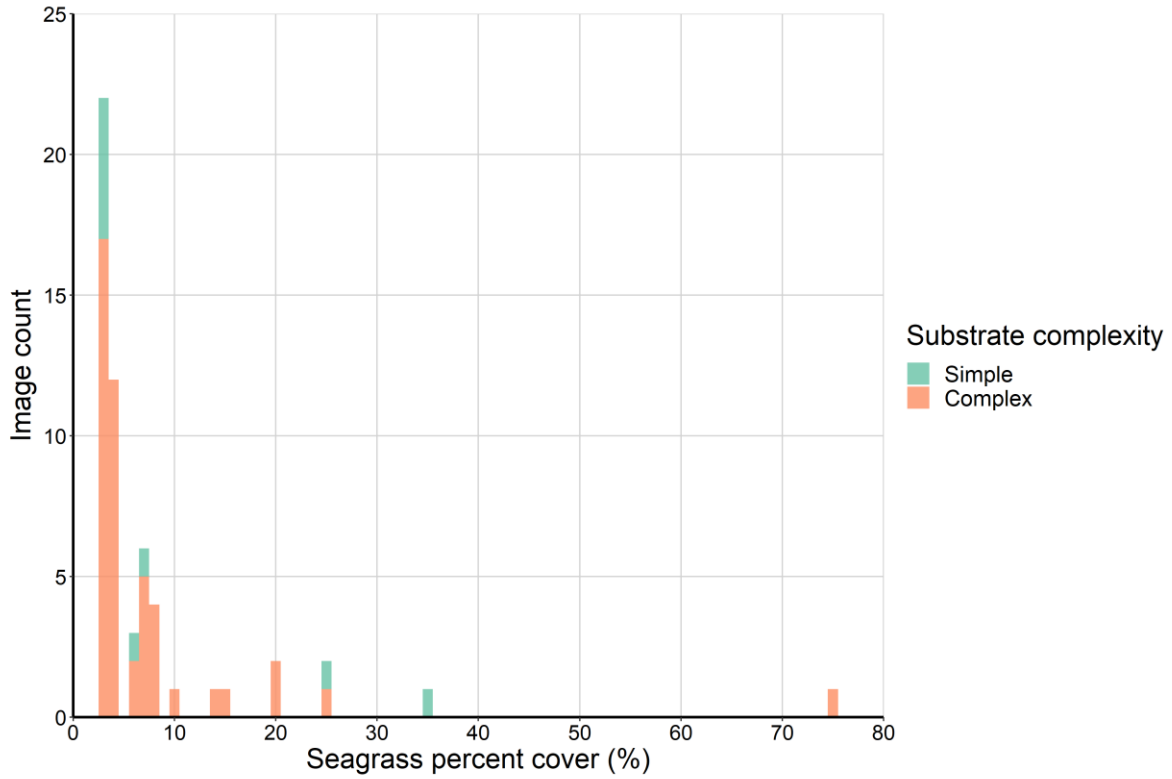


Figure 6. Histogram of the distribution of the seagrass percent cover and substrate complexity present in the 56 images misclassified (false negative) in *Model #1*.

3.2. Model #2

Model #2 had an overall accuracy of 80.7% (Table 2) on the test image set (647). The highest accuracy was achieved for the medium cover class (84.3%), followed by the low cover class (78.5%) and the high cover class (75.9%). However, these differences in accuracies were marginal and most likely a consequence of the unbalanced nature of the cover classes image dataset (Figure 7, Table 3c).

All the misclassified images of the low cover classes (45) were incorrectly predicted to be in the medium cover category. Misclassification occurred for images with percent cover between 7 and 9% (31) (Figure 8). Furthermore, 32 of which also had a complex substrate, further highlighting the difficulty categorising images close to the threshold of 10%, especially for complex substrates where algae for example could be biasing the predictions.

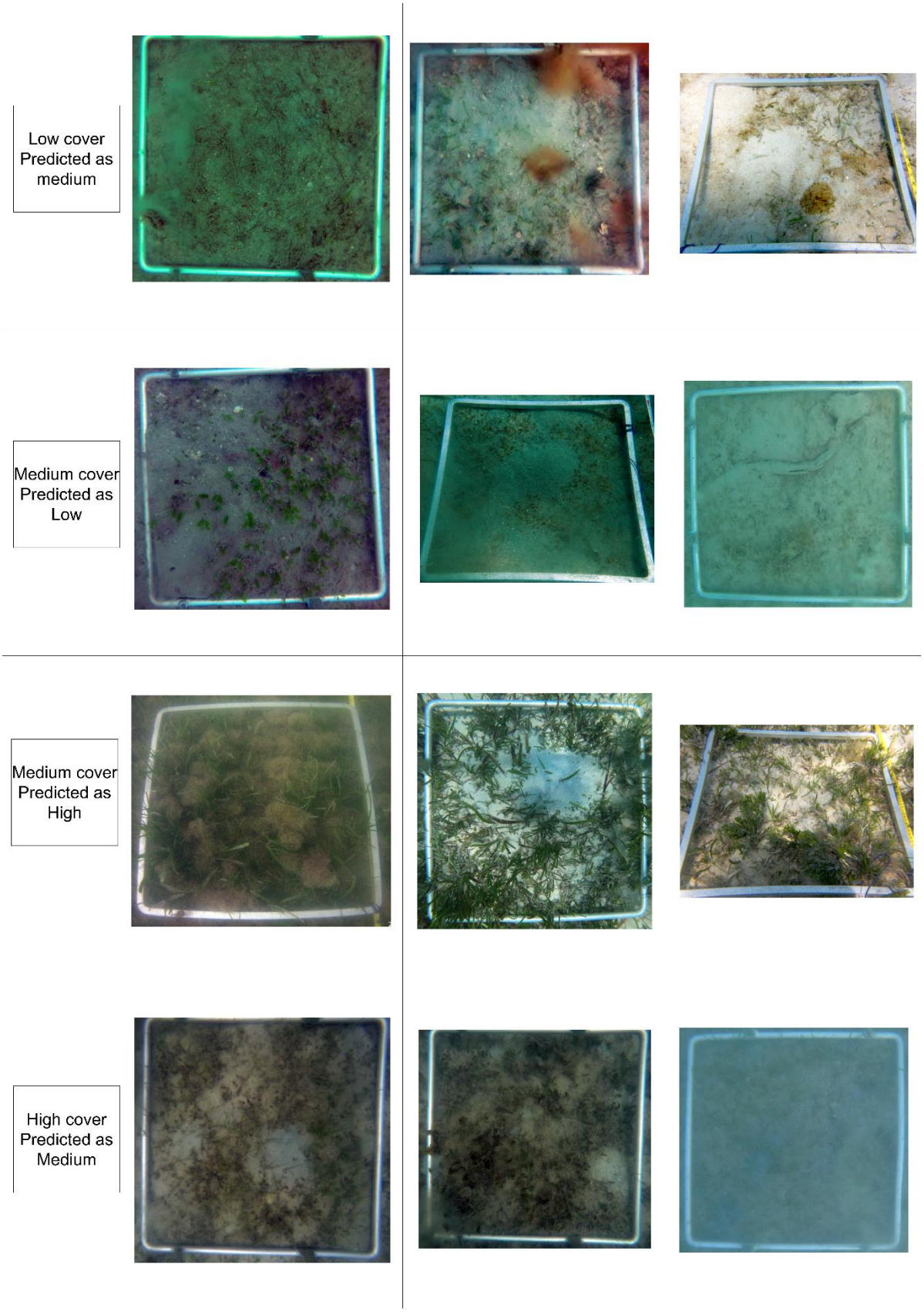


Figure 7. Examples of images misclassified by Model #2 from the low, medium and high cover categories.

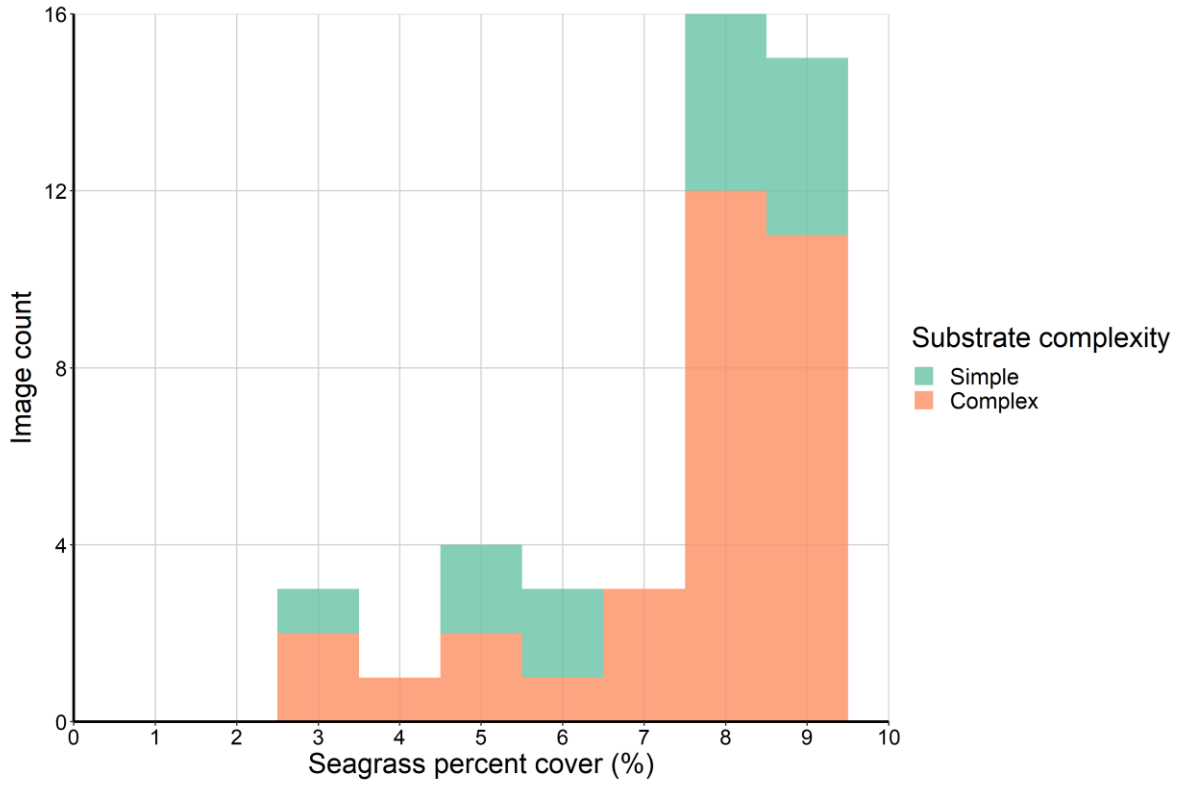


Figure 8. Histogram of the distribution of the seagrass percent cover and substrate complexity present in the 45 misclassified images of the low cover category (false medium) by *Model #2*.

There were 48 misclassified images of the medium cover classes, of which 31 were predicted as low cover and 17 as high cover. The false low cover images were mostly close to the 10% threshold with 27 if these images being between 10 and 15% seagrass cover (Figure 9). Images dominated by smaller seagrass species with rounded and fernlike morphology were also a source of misclassification. The false high classifications were solely dominated by straplike species and 10 images had a seagrass cover between 20 and 30%.

There were 32 misclassified images of the high cover class, which were all predicted as a medium cover. Similarly, to the previous classes, a vast majority of these were close to the adjacent cover category threshold with 28 of these images having less than 38% seagrass cover (Figure 10). Straplike morphology dominated in 27 of the misclassified images except for those with percent cover of more than 40% which were dominated by rounded and fernlike morphology.

The type of substrate was not a significant driver of prediction errors for the medium and high cover class.

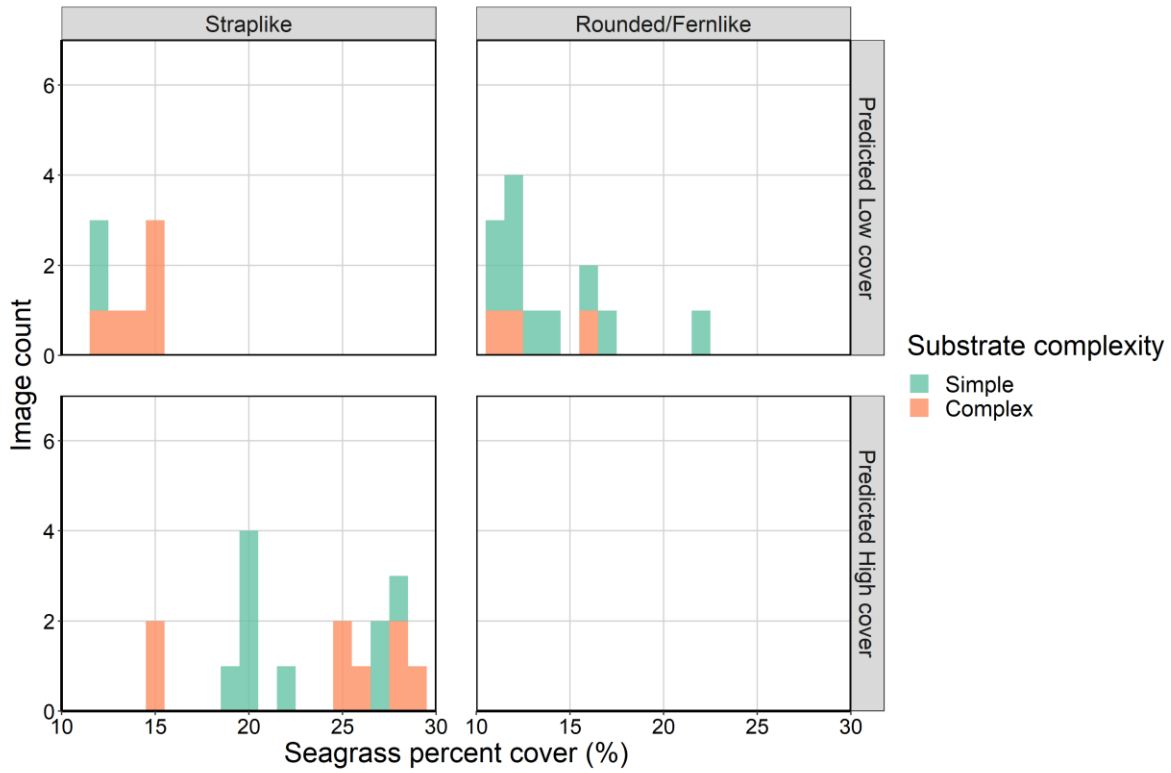


Figure 9. Histogram of the distribution of the seagrass percent cover and substrate complexity present in the 45 misclassified images of the medium cover category using *Model #2*).

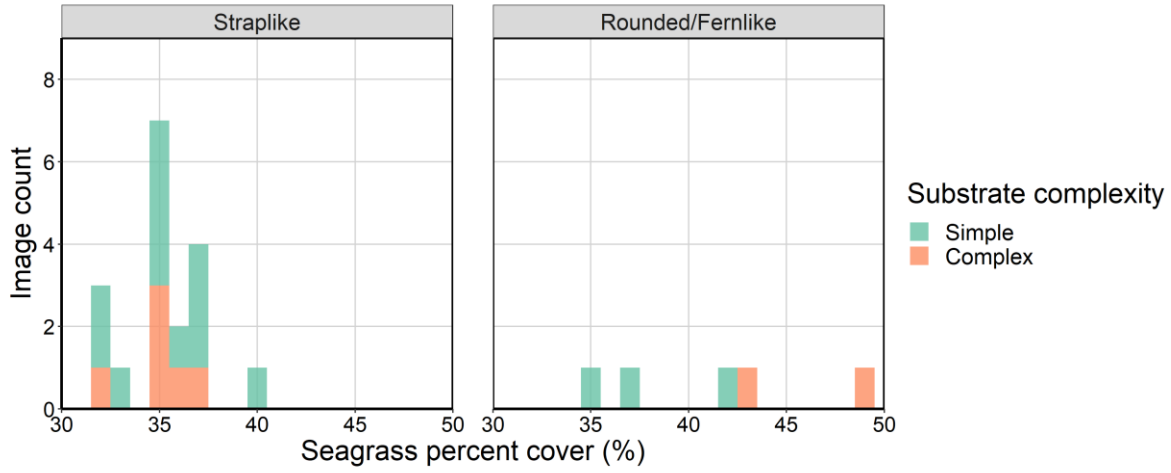


Figure 10. Histogram of the distribution of the seagrass percent cover and substrate complexity present in the 45 misclassified images of the medium cover category using *Model #2*).

3.3. Model #3

Our subtidal substrate complexity classifier (*Model #3*) achieved an accuracy of 97.9% (Table 2) on the test image set (240) and on the simple substrate only images remaining (531). There were two images misclassified as complex and three images were misclassified as simple instead of complex out of the test image set (Figure 11, Table 3d). These images were also difficult to manually classify because they were mostly composed of a simple sandy substrate with some additional features such as algae or soft coral, or have poor visibility.

There were 11 images misclassified as complex instead of simple out of the simple substrate images remaining. These had 7% algae cover on average and 10 had more than 3%. This may be a consequence of the arbitrary binary classification used during the labelling process. It is very difficult to establish a clear difference between a quadrat with a simple sandy substrate with some algae or other feature like coral and a complex substrate. These instances are uncommon within the dataset, with 82 images labelled as simple substrate and more than 3% algae cover and occurred mainly only at the Dunk Island and Low isles sites (36 and 30 images respectively). This could be easily refined further by increasing the image dataset and by setting clearer thresholds or rules to define the substrate complexity classes.

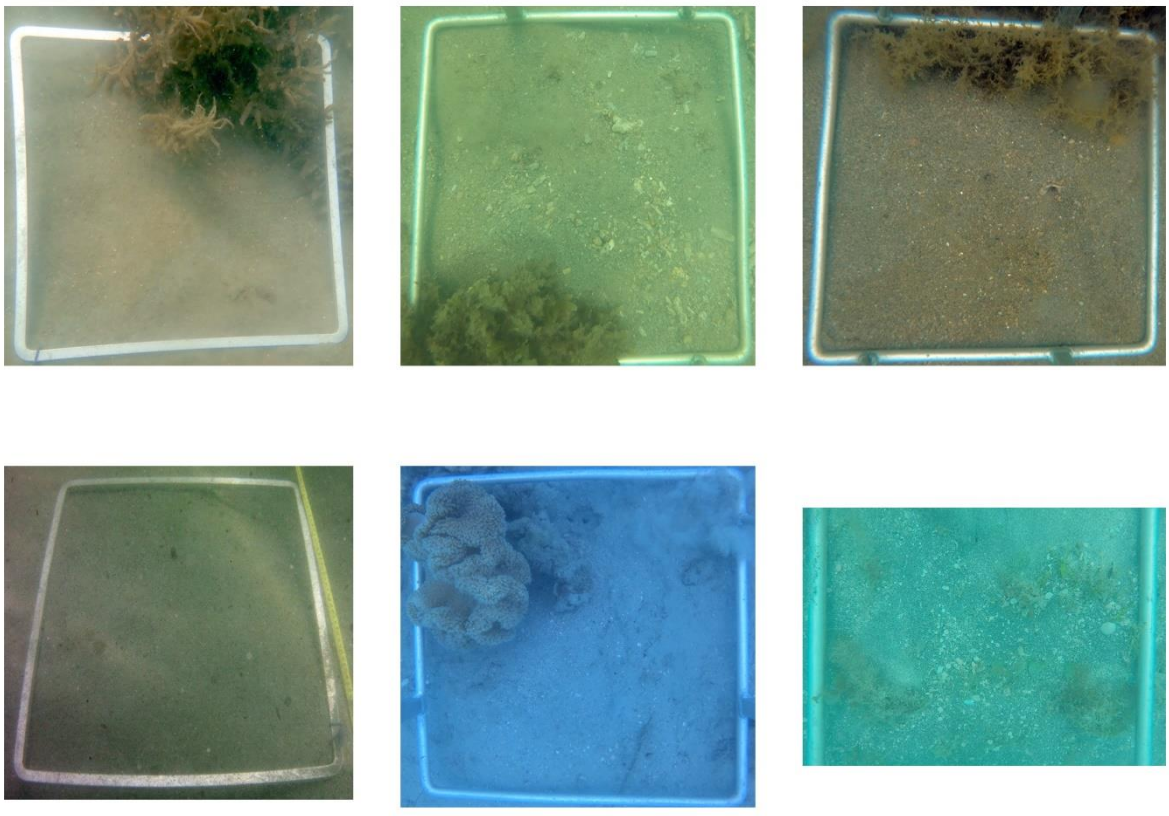


Figure 11. Examples of images misclassified by *Model #3* with false complex on top row and false simple on the bottom row.

4. Discussion

4.1. Method performance and limitations

The main goal of this research was to determine the potential for deep–learning models to detect the presence of seagrass within underwater photos. Seagrass was identified in images containing a mix of seagrass species, seagrass morphologies and from a range of sites with a very high level of accuracy (97%). This was achieved using a simple neural network architecture. The performance of *Model #1* was higher than previously published deep–learning seagrass detection models (Raine *et al.*, 2020). However, a direct comparison between the accuracies is difficult due to differences in image dataset size and classifiers for seagrass morphology between studies.

We found that most of the misclassification occurred for images with complex substrate especially those with high algae percent cover. This is typical for deep–learning classification models that are still lacking the ability to apply extreme generalization the way humans do (Chollet, 2017). Differentiating among well-defined objects is usually straight forward with numerous documented examples on image datasets such as ImageNet (Krizhevsky *et al.*, 2017). The model outcomes for complex substrate, could possibly be improved by increasing the overall number of images, but also by having a balanced number of images with the same level of algae with and without seagrass. Indeed, deep–learning models are able to continue to "learn" with additional imagery. So as new images are being collected, our present models can be further trained which will lead to improved performance over time.

We also demonstrated it was possible to categorise seagrass cover into three broad classes with an accuracy of 80.7%. The choice of category boundaries was crucial in the model performance. Most of the classification errors happened around these boundaries and resulted in an image being placed into the adjacent category, rather than for example two categories away (i.e. a high being classed as low or vice versa). This needs to be considered when applying the model. For instance, the medium seagrass cover category was defined as ≥10 <30% during the labelling process, however the percent cover range of the images predicted in that class ranged from 7 to 35%.

Seagrass percent cover estimates can be difficult to assign for low densities. Except for a few structurally large species, individual seagrass leaves are very small and therefore may not be easy to identify. A study from Moniruzzaman et al. (2019) developed deep–learning models to detect single leaves of *Halophila ovalis*. This was effective for oblique close-up images with a sand background, but is likely to be less effective with nadir quadrat images as used in this study. Photoquadrats are used so that cover can be easily quantified in a standardised manner. Furthermore, it would require a significant effort to label a photoquadrat dataset where individual bounding boxes must be drawn for every single leaf in every image.

An alternative method to estimate percent cover of benthic taxa (e.g. coral, algae, seagrass) and substrate (e.g. sand, rock) is using a point annotation system. This method has been successfully used for coral reefs and invertebrate communities (González-Rivero *et al.*, 2016) and is publicly available through platforms such as CoralNet or ReefCloud. In seagrass habitats, the point annotation method is only able to detect seagrass when cover is above 25% (Kovacs *et al.*, 2022). This is because the method relies on classifying an area (224x224 pixels) around the annotated point. The dimension of the annotation area is not visible through the labelling interface and the person conducting the labelling is expected to label only what is directly under the point. This approach works for well–defined and larger objects like coral, however, it is not well adapted to scattered, low and sparse seagrass cover where there could be seagrass within the classifying area but not directly under the point, resulting in a high degree of misclassification. Furthermore, the labelling effort required to fully

label a single image with 10 to 30 points is significantly greater than conducting a single seagrass estimate for the entire image.

While we acknowledge the potential limitations of our models, especially *Model #2*, we believe the preliminary results are very exciting and encouraging. We are hoping to further develop our method towards highly accurate automatic seagrass detection that is suited for operational applications in a vast number of locations.

4.2. Operationalisation and Mainstreaming

This study was undertaken to demonstrate the feasibility of a subtidal seagrass detection model as a step towards operationalisation and mainstreaming of big data acquisition and analysis (e.g. Dalby *et al.*, 2021).

Traditional direct field observations provide instantaneous data, but need to be performed or overseen by formally trained scientists, and the data requires time consuming transcription into a database. Images (e.g. photoquadrats), however, can be collected by a variety of participants such as environmental practitioners, Indigenous ranger groups or members of the public without a formal scientific background (i.e. citizen scientists), requiring less capacity and resources. For example, rangers from the Queensland Park and Wildlife Services (QPWS) conduct subtidal seagrass monitoring using drop cameras that is currently integrated into the Great Barrier Reef Marine Monitoring Program (McKenzie *et al.*, 2021). Citizen scientists, QPWS Rangers and Indigenous rangers frequently access the Reef and seagrass habitats of northern Australia. Simplifying the methods and minimising the time required to capture data by using photoquadrats can vastly increase the volume, velocity, variety and geographic spread of image data collection. The models presented in this study facilitate the ability to mainstream data capture and increase the rate of image processing, enabling scientists to maximise Big Data analysis and reporting. With the addition of the deep–learning models, we can grow our capacity for image data handling. In parallel with the development of the models presented here we are also working on streamlining a higher efficiency image processing workflow. This includes handling either time-lapse or video (e.g. GoPro) input sources to generate deep–learning ready standardized quadrat images.

The operational applications for the subtidal seagrass detector are wide-ranging, including mapping and monitoring of the vast and remote northern Australian seagrass habitats. Image collection combined with a geotagging/geolocation, will enable the production of spatially explicit maps of subtidal areas. Our models are most adapted to this application as maps tends to only need simple information input like seagrass presence/absence. However, we have also shown possibilities for monitoring with the ability to detect broad seagrass cover categories which would enable temporal changes in seagrass abundance to be assessed.

4.3. Future directions

While the findings in this study are encouraging, we very much intend to further refine and improve those models and the associated data processing workflow over time. One of the main advantages of using Deep Neural Networks is their capacity to incrementally improve when additional training data is provided. Therefore, as more and more diverse images are supplied it will help us build more robust models and give greater confidence in the predictions. Our models are currently limited to be used on subtidal nadir photoquadrats captured using a drop camera. However, with the increasing popularity

of Autonomous Underwater Vehicles (AUVs), our DNNs would need to be trained to accept more versatile image inputs (e.g. oblique and without guiding bounds).

The other main area for future improvement is data handling. Our vision is to create a web portal where participants can upload images or videos and are able to visualise the raw outputs as predicted seagrass distribution maps.

5. Appendix

Table 1. Details about location, NRM region, source and number of images used in study. TSSMP= Torres Strait Seagrass Monitoring Program, QPWS = Queensland Parks & Wildlife Service, MMP = Great Barrier Reef Marine Monitoring Program, (MMP) = contributes to MMP.

Location	Natural Resource Management Region	Habitat	Image source	Number of images
Orman Rf	Torres Strait	reef	TSSMP	649
Dungeness Is.	Torres Strait	reef	TSSMP	333
Dugong Sanctuary	Torres Strait	reef	TSSMP	306
Margaret Bay	Cape York	coastal	QPWS (MMP)	20
Lloyd Bay	Cape York	coastal	QPWS (MMP)	108
Flinders Is.	Cape York	reef	QPWS (MMP)	98
Bathurst Bay	Cape York	reef	QPWS (MMP)	56
Low Isles	Wet Tropics	reef	MMP	631
Green Is.	Wet Tropics	reef	MMP	912
Dunk Island	Wet Tropics	reef	MMP	1044
Missionary Bay	Wet Tropics	coastal	QPWS (MMP)	42
Magnetic Is.	Burdekin	reef	MMP	710
Cid Harbour	Mackay Whitsunday	reef	QPWS (MMP)	14
Tongue Bay	Mackay Whitsunday	reef	QPWS (MMP)	119
White Heaven Beach	Mackay Whitsunday	reef	QPWS (MMP)	20
Lindeman Is.	Mackay Whitsunday	reef	MMP	485
Newry Bay	Mackay Whitsunday	coastal	QPWS (MMP)	96
Hervey Bay	Burnett Mary	coastal	Seagrass-Watch	151

Table 2. Deep Neural Network model summary outputs.

Model / dataset	Class	Precision	Recall	F1-score	Images	Accuracy
Model 1 – Test set	Seagrass absent	0.95	0.99	0.97	346	0.97
	Seagrass present	0.99	0.95	0.97	345	
Model 1 – image remaining set	Seagrass absent	NA	NA	NA	NA	0.97
	Seagrass present	1	0.97	0.99	1509	
Model 2 – Test set	Low cover	0.84	0.78	0.81	209	
	Medium cover	0.77	0.84	0.8	305	0.81
	High cover	0.86	0.76	0.8	133	
Model 3 – Test set	Simple	0.98	0.98	0.98	120	0.98
	Complex	0.98	0.97	0.98	120	
Model 3 – image remaining set	simple	1	0.98	0.99	531	0.98
	complex	NA	NA	NA	NA	

Table 3. Deep Neural Network model confusion matrices.

a. *Model #1* – Test set

		Predicted	
		Seagrass absent	Seagrass present
Actual	Seagrass absent	343	3
	Seagrass present	18	327

b. *Model #1* – images remaining set

		Predicted	
		Seagrass absent	Seagrass present
Actual	Seagrass absent	0	0
	Seagrass present	38	1471

c. *Model #2* – Test set

		Predicted		
		Low cover	Medium cover	High cover
Actual	Low cover	164	45	0
	Medium cover	31	257	17
	High cover	0	32	101

d. *Model #3* – Test set

		Predicted	
		Simple	Complex
Actual	Simple	118	2
	Complex	3	117

e. *Model #3* – images remaining set

		Predicted	
		Simple	Complex
Actual	Simple	520	11
	Complex	0	0

References

Agarap, A.F., 2018, Deep Learning using Rectified Linear Units (ReLU). arXiv.

Beijbom, O., Edmunds, P.J., Roelfsema, C., Smith, J., Kline, D.I., Neal, B.P., Dunlap, M.J., Moriarty, V., Fan, T.-Y., Tan, C.-J., Chan, S., Treibitz, T., Gamst, A., Mitchell, B.G., Kriegman, D., 2015. Towards Automated Annotation of Benthic Survey Images: Variability of Human Experts and Operational Modes of Automation. *PLOS ONE* **10** (7): e0130312.

Carter, A., David, M., Whap, T., Hoffman, L., Scott, A., Rasheed, M., 2021a, Torres Strait seagrass 2021 report card. TropWATER Report No. 21/13. TropWATER, James Cook University, Cairns.

Carter, A.B., McKenna, S.A., Rasheed, M.A., Collier, C., McKenzie, L., Pitcher, R., Coles, R., 2021b. Synthesizing 35 years of seagrass spatial data from the Great Barrier Reef World Heritage Area, Queensland, Australia. *Limnology and Oceanography Letters* **6** (4): 216-226.

Chollet, F., 2017, Deep learning with python. Manning Publications.

Dalby, O., Sinha, I., Unsworth, R.K.F., McKenzie, L.J., Jones, B.L., Cullen-Unsworth, L.C., 2021. Citizen Science Driven Big Data Collection Requires Improved and Inclusive Societal Engagement. *Frontiers in Marine Science* **8** (432).

González-Rivero, M., Beijbom, O., Rodriguez-Ramirez, A., Holtrop, T., González-Marrero, Y., Ganase, A., Roelfsema, C., Phinn, S., Hoegh-Guldberg, O., 2016. Scaling up Ecological Measurements of Coral Reefs Using Semi-Automated Field Image Collection and Analysis. *Remote Sensing* **8** (1): 30.

Johnson, J.E., Welch, D.J., Marshall, P., Day, J., Marshall, N., Steinberg, C., Benthuysen, J., Sun, C., Brodie, J., Marsh, H., Hamann, M., Simpfendorfer, C.A., 2018, Characterising the values and connectivity of the northeast Australia seascape: Great Barrier Reef, Torres Strait, Coral Sea and Great Sandy Strait. Report to the National Environmental Science Program. Reef and Rainforest Research Centre Limited, Cairns, p. 81.

Kingma, D.P., Ba, J., 2014, Adam: A Method for Stochastic Optimization. arXiv.

Kovacs, E.M., Roelfsema, C., Udy, J., Baltais, S., Lyons, M., Phinn, S., 2022. Cloud Processing for Simultaneous Mapping of Seagrass Meadows in Optically Complex and Varied Water. *Remote Sensing* **14** (3): 609.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **60** (6): 84–90.

McKenzie, L.J., Campbell, S.J., Roder, C.A., 2003, Seagrass-Watch: Manual for Mapping & Monitoring Seagrass Resources. QFS, NFC, Cairns.

McKenzie, L.J., Collier, C., Langlois, L., Yoshida, R., Smith, N., Waycott, M., 2015, Reef Rescue Marine Monitoring Program - Inshore Seagrass, Annual Report for the sampling period 1st June 2013 – 31st May 2014. TropWATER, James Cook University, Cairns, p. 226.

McKenzie, L.J., Collier, C.J., Langlois, L.A., Yoshida, R.L., Uusitalo, J., Waycott, M., 2021, Marine Monitoring Program: Annual report for inshore seagrass monitoring 2019–20. James Cook University.

McKenzie, L.J., Collier, C.J., Langlois, L.A., Yoshida, R.L., Waycott, M., 2022a, Marine Monitoring Program: Annual Report for Inshore Seagrass Monitoring 2020–21. Report for the Great Barrier Reef Marine Park Authority. Great Barrier Reef Marine Park Authority, Townsville, p. 177.

McKenzie, L.J., Langlois, L.A., Roelfsema, C.M., 2022b. Improving Approaches to Mapping Seagrass within the Great Barrier Reef: From Field to Spaceborne Earth Observation. *Remote Sensing* **14** (11): 2604.

McKenzie, L.J., Roder, C.A., Roelofs, A.J., Lee Long, W.J., 2000, Post-flood monitoring of seagrasses in Hervey Bay and the Great Sandy Strait, 1999: Implications for dugong, turtle and fisheries management., Department of Primary Industries Information Series QI00059. Queensland Department of Primary Industries, NFC, Cairns, p. 46.

Moniruzzaman, M., Islam, S.M.S., Lavery, P., Bennamoun, M., 2019, Faster R-CNN Based Deep Learning for Seagrass Detection from Underwater Digital Images. 2019 Digital Image Computing: Techniques and Applications (DICTA), pp. 1-7.

Mount, R., Bricher, P., Newton, J., 2007, National Intertidal/Subtidal Benthic (NISB) Habitat Classification Scheme, Version 1.0, October 2007. National Land & Water Resources Audit & School of Geography and Environmental Studies, University of Tasmania, Hobart, Tasmania.

Preen, A.R., Lee Long, W.J., Coles, R.G., 1995. Flood and cyclone related loss, and partial recovery, of more than 1,000 km^2 of seagrass in Hervey Bay, Queensland, Australia. *Aquatic Botany* **52** 3-17.

Raine, S., Marchant, R., Moghadam, P., Maire, F., Kettle, B., Kusy, B., 2020. Multi-species Seagrass Detection and Classification from Underwater Images. *arXiv:2009.09924*8.

Reus, G., Möller, T., Jäger, J., Schultz, S.T., Kruschel, C., Hasenauer, J., Wolff, V., Fricke-Neuderth, K., 2018, Looking for Seagrass: Deep Learning for Visual Coverage Estimation. 2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO), pp. 1-6.

Schröder, S.-M., Kiko, R., Koch, R., 2020, MorphoCluster: Efficient Annotation of Plankton images by Clustering. arXiv.

Seagrass-Watch, 2022. Hervey Bay https://www.seagrasswatch.org/burnettmary/#HV. https://www.seagrasswatch.org/burnettmary/#HV, 22 June 2022

Sheaves, M., Bradley, M., Herrera, C., Mattone, C., Lennard, C., Sheaves, J., Konovalov, D.A., 2020. Optimizing video sampling for juvenile fish surveys: Using deep learning and evaluation of assumptions to produce critical fisheries parameters. *Fish and Fisheries* **21** (6): 1259-1276.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research* **15** (56): 1929-1958.

Tammina, S., 2019. Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images. *International Journal of Scientific and Research Publications (IJSRP)* **9** (10).

Tieleman, T., Hinton, G., 2014. RMSprop gradient optimization. *URL: http://www.cs.toronto.edu/tijmen/csc321/slides/lecture_slides_lec6.pdf*.

Zhang, X., Zou, J., He, K., Sun, J., 2015. Accelerating Very Deep Convolutional Networks for Classification and Detection.