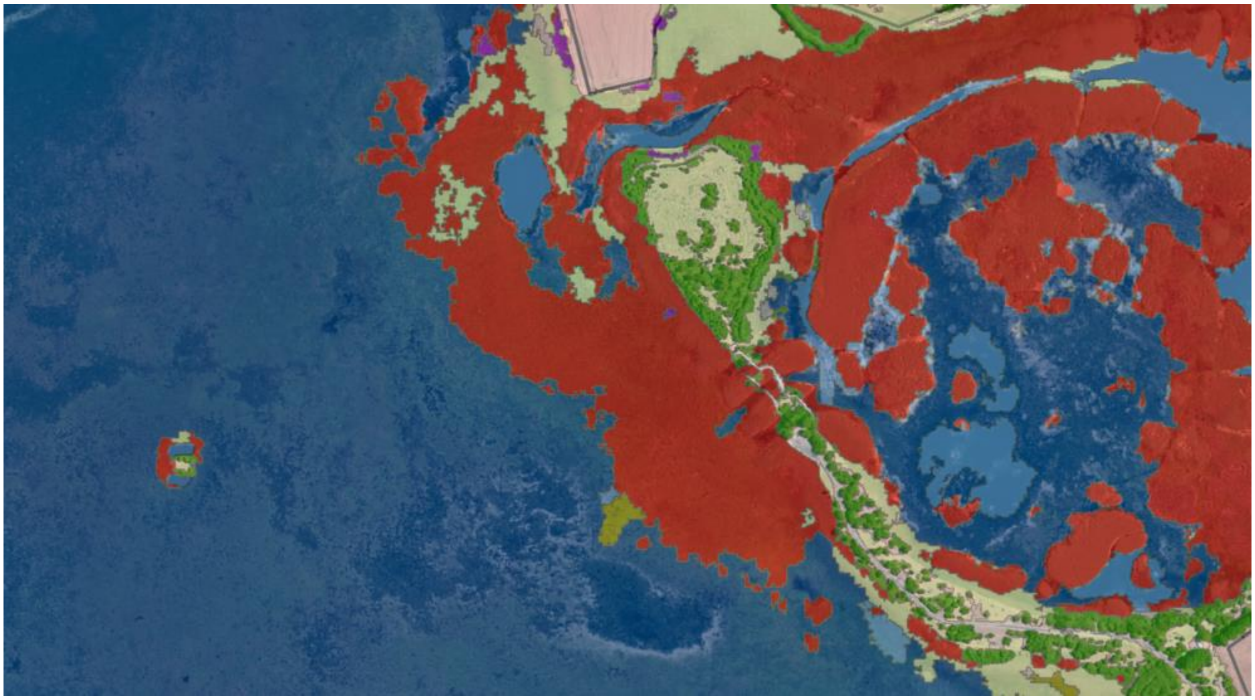
Bools and Hacks Lagoon
Model and Data Quality Report

Spatial Solutions

prepared by Lynker Analytics



Contents

Document Status

Version	Author	Edit Date	Reviewed By	Review Date
Final	Michael Stanley	11th May 2023	David Knox	11th May 2023

Overview

The Friends of Bool and Hacks Lagoon group and BirdLife Australia provided Lynker Analytics with aerial photography for the Bools and Hacks lagoon. The imagery consisted of 66 ECW files which covered the Bool lagoon at a resolution of 0.106m. In addition to the imagery, twenty-seven ground truth points were also provided to assist in the correct annotation of the eight target classes.

Lynker then manually annotated these images into a polyline annotation dataset. The classes followed by their class id are:

- Tussock 1
- Tree 2
- Sedge 3
- Reed 4
- Grasses 5
- Open Water 6
- Ground 7
- Aquatic Floating 8

Lynker used a machine learning training process called supervised learning, whereby a machine learning model is trained using example image and annotation pairs to learn the same decision outcomes on new or previously unseen images.

Machine Learning is notoriously data-hungry and model accuracy is sensitive to the quality and quantity of input data. An annotation process that used polylines to quickly develop a large dataset of positively annotated pixels was used to develop the dataset of target classes to train the supervised model.

The model's performance on holdout data was shown to have a classification accuracy of 0.965 and mean F1 score also of 0.965. Sedge was the lowest performing class often instead being predicted to be grasses or ground. The aquatic floating class was the highest performing class in the holdout set, every pixel of this class in the holdout set was correctly predicted and no other classes were incorrectly predicted to belong to the aquatic floating class.

Data Exploration

Lagoons in South Australia represent an essential component of the region's biodiversity, and their preservation is crucial to maintaining a healthy and sustainable environment. Due to the diverse nature of vegetation found in lagoons and their complex web of ecological interactions with the surrounding environment accurately modelling this complexity is a unique challenge. Figure 1 highlights some of this complexity displaying how the transition of tussock to reed may be difficult to delineate (bottom right of the image), as well as where trees grow over Reed (bottom left of the image).

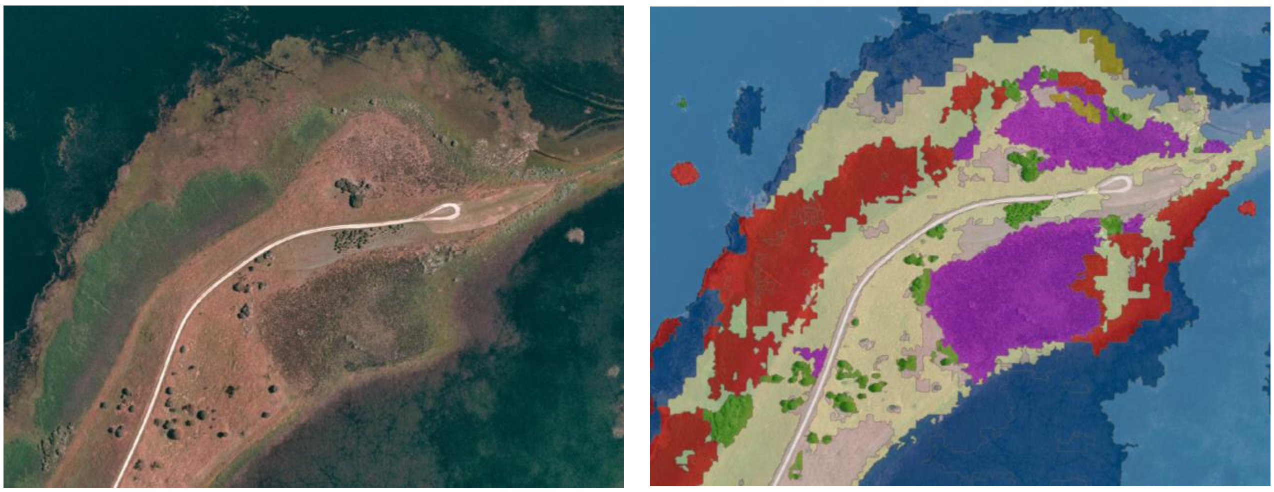


Figure 1. comparison of imagery vs inference overlaying the imagery

The provided imagery had a size of 9423x9423 pixels and a ground sample distance (resolution) of approximately 0.106m (10.6cm pixel width) and used the WGS 1984 UTM Zone 54S projected coordinate system. We note that in the northwest of the bool lagoon the imagery did not extend to the boundary of the SA Ramsar Reserves, shown in figure 2.

Figure 2. The northwest extent of the Bool Lagoon

Machine Learning Method

8 distinct classes were selected to classify the entire area of the Bools and Hacks lagoon. These classes were chosen as they allow us to accurately segment the majority of the land cover of the region. Despite this, we are aware that some vegetation classes do exist in the area that cannot be accounted for with an 8 class approach.

One such example was tea tree recruits. Training data for this additional class were included during the initial model development, but was later excluded due to its perceived similarities to the tussock class, which ultimately led to a decreased classification accuracy overall.

Annotated data were split into three non-intersecting categories, train, validation, and a holdout set. These were split with a 70,15,15 percent split respectively.

The model uses the training data to learn from. The validation data is not used for training but is used to measure progress. When the validation loss (not shown here but this measure correlates with model accuracy) is at its best, the model is saved. The holdout data is both unseen by the model but is from separate source images (975 images were held out from the training and validation sets) and so is the most independent indication of accuracy. We use this holdout set to measure the stated model accuracy used in this report.

To classify the Bools and Hacks lagoon an 8-class segmentation model was trained to be able to infer the class of unseen imagery across the entire lagoon. 1862 polylines were drawn over examples from the eight classes across the entire lagoon to create a dataset suitable for training and evaluating the machine learning (ML) model.

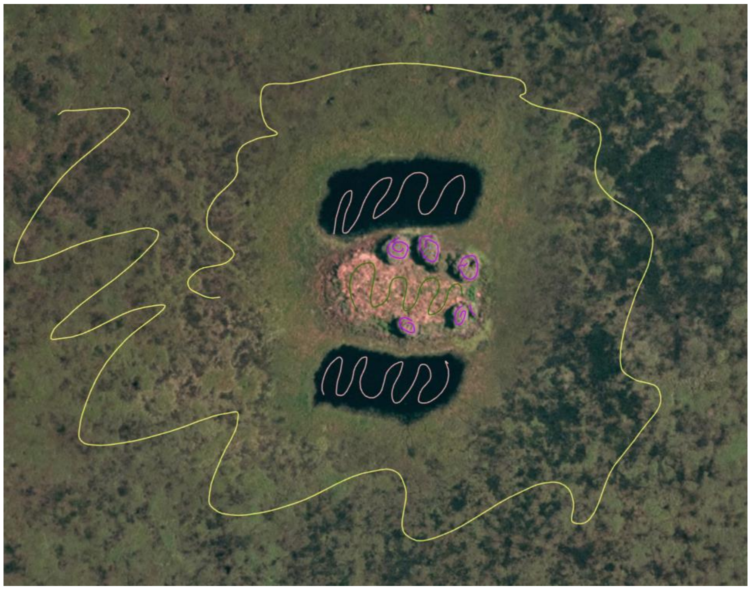


Figure 3. Example of polyline annotations

Figure 3 shows an area that was annotated with polylines for four different classes. Polylines were buffered with a padding of 0.1m and were exported in tiles of 256x256m to create the dataset of imagery used in training, validating, and testing the model that was developed during this project.

The data were split as follows:

- 4453 images were used to train the model.
- 1076 images were used when validating the model's performance at each training step.
- 975 images consisting of 75581 annotated pixels were included in a holdout set, this was used to test the final model's performance across the eight classes. Table 2 displays the results of the inference from this data.

To infer on the imagery tiles were exported with a size of 10240mx10240m, a stride of 9984 was used to allow for the tiles to have an overlap of 256. This was done to allow for the edges of inferred tiles to be removed without resulting in missing data in the final mosaic.

Post processing techniques

Inference tiles were resampled with a majority filter to be 1m x 1m. To remove missing data and tidy up the inference we then applied the "RegionGroup" function to group neighboring cells with the same value into unique regions or zones. This helped to segment the larger study area into smaller regions. The "SetNull" function was then applied to all region groups with a 'count' value of less than 18, this helped to remove unwanted data and isolate specific noisy inferences. Finally, the Nibble function helped to correct this noise and create a more continuous representation of the study area.

Rasters were then clipped to remove pixels within 12.8m of the border of the inference tile, this removed poor inferences at the edge of our tiles from being included in the output dataset.

An eliminate was used to further remove polygons below a certain size (specific to each class) to reduce additional noise present from the raw output of the inference. The shape area was determined by using a sample of twenty instances for each class and first calculating the area of both true and false positives and selecting the value that removed most of the false positives while retaining the true positives.

This was not done for the ground class as misclassifications of the ground class were large so using an eliminate would have also removed many true positive ground inferences. These areas of misclassification were known and so were tidied up manually after inference.

The rules for the eliminate function were:

gridcode = 1 AND Shape_Area < 260

gridcode = 2 AND Shape_Area < 24

gridcode = 3 AND Shape_Area < 400

gridcode = 4 AND Shape_Area < 150

gridcode =5 AND Shape_Area < 400

gridcode = 6 AND Shape_Area < 600

gridcode = 8 AND Shape_Area < 1100

Where gridcode represents the class id

Figure 4. Sedge areas are often classified as ground.

The automated post processing techniques reduced the number of polygons in the vectorized raw ML inference from 185,177 polygons to 15,594 polygons.

After post processing some manual cleanup was also done, to remove false positive inferences. The most common occurrence of these was for the tree class, this was often found on aquatic floating vegetation or tussock. We believe this is due to the similarities between tea tree recruits and the tussock class.

To tidy up such regions a lasso selection was done over known areas where no true trees existed, for example over large bodies of deep open water, care was taken to ensure that anomalous trees were not included in this selection. From the lasso select a subset of the selection where the gridcode (class id) was equal to two (class id for tree) to select all the tree inferences within the selected region and exclude other classes. These were then assigned to the correct class for the area. An example of the inference before and after this process is shown in Figure 5.

After the manual tidy-up, the features were converted to a raster dataset to remove the excess adjacent polygons of the same class, the process for removing missing data and tidying up the inference was conducted once more after resampling the raster with a majority filter of 2.5m. The number of polygons after this process had been conducted and the inference was clipped to the reserve region was 6,065.

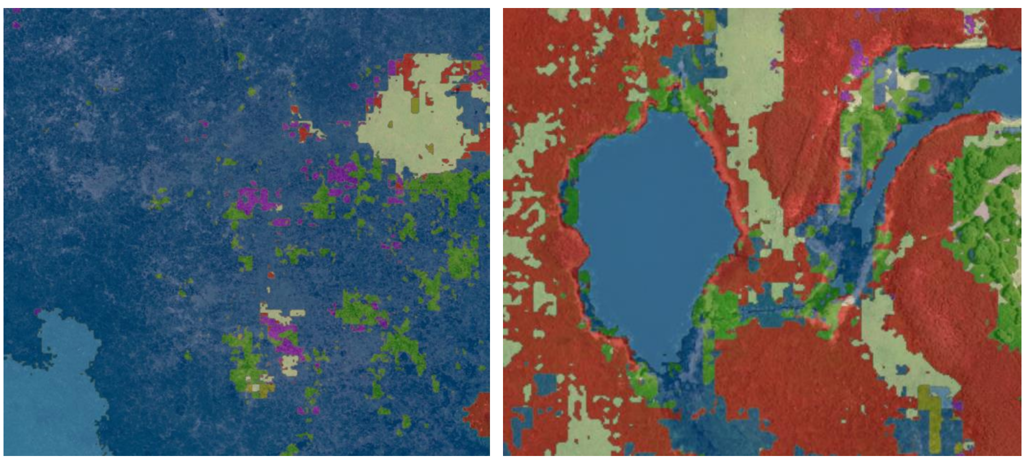


Figure 5.1 False positive inferences before post processing and manual tidy up

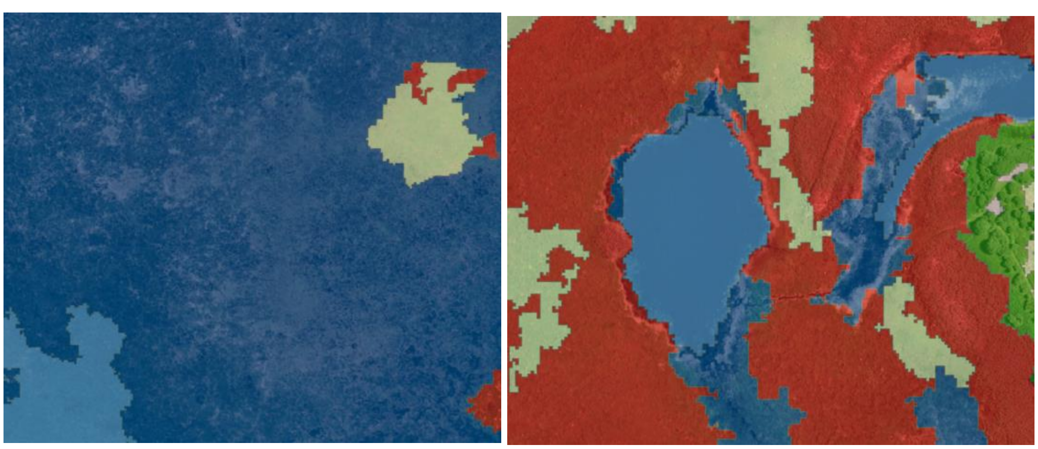


Figure 5.2 False positive inferences after post processing and manual tidy up

Results

Overall, the model performed very well on our holdout dataset, on average the model identified 96.5% of the pixels in each class correctly (recall). Despite this high score due to the high level of biodiversity in this wetland our holdout points do not fully explain the complexity of this dataset so these figures may be slightly inflated for certain areas of the Bools lagoon.

	Precision	Recall	F1-score	Support
NoData_0	0	0	0	0
Tussock_1	0.961	0.991	0.976	17443
Tree_2	0.981	0.993	0.987	9218
Sedge_3	0.976	0.842	0.904	7751
Reed_4	1	0.911	0.953	7111
Grasses_5	0.913	0.974	0.942	10395
Water_6	0.992	1	0.996	10172
Ground_7	0.944	0.969	0.956	9428
Aquatic_floating_8	1	1	1	4063
micro avg	0.965	0.965	0.965	75581
macro avg	0.863	0.853	0.857	75581
weighted avg	0.966	0.965	0.965	75581

Table 1. Evaluation of model performance across each class

Precision = true positives / (true positives + false positives)

This measures the number of correct predictions as a percentage of everything predicted to be in the class. A value of one means that no other classes were incorrectly predicted to be part of the class.

Recall = true positives / (true positives + false negatives)

Recall measures the number of correct predictions as a percentage of the total number of instances of that class. A value of one means that our inference caught every instance of the class that was available in the holdout set.

Predicted as ->	Tussock_1	Tree_2	Sedge_3	Reed_4	Grasses_5	Water_6	Ground_7	Aquatic_floating_8	Support
Tussock_1	17284	25	0	0	27	85	22	0	17443
Tree_2	0	9156	0	0	61	0	1	0	9218
Sedge_3	0	0	6525	0	741	0	485	0	7751
Reed_4	576	57	0	6478	0	0	0	0	7111
Grasses_5	0	96	143	0	10120	0	36	0	10395
Water_6	0	0	0	0	0	10172	0	0	10172
Ground_7	133	0	19	0	140	0	9136	0	9428
Aquatic_floating_8	0	0	0	0	0	0	0	4063	4063

Table 2. Confusion matrix for the holdout set across the eight classes.

Table 2 shows the breakdown of the pixel predictions for each of the holdout images. Predicted class is shown on the x axis while the true class is shown on the Y axis. For example, 6478 reed pixels were correctly predicted. No other classes were predicted to be “reeds”, resulting in a precision value of 1, while 57 reed pixels were predicted to be trees and 576 “reeds” were predicted to be “tussock”.

Incorrect predictions of “reed” as “tussock” often occurred in areas in which the two different types of vegetation were growing in close proximity. This was more prominent in dry basins where smaller clusters of reed were growing among larger groups of tussock.

The “sedge” class was most susceptible to false negatives, often being misclassified as either “grass” or “ground”. This was likely a result of the “sedge” in dry basins being difficult to delineate from “ground” in the imagery, as shown in Figure 4.

Shadows

No special handling of shadows has been performed in this project. By inspection of images and the classified outputs, it rarely results in the shadows of trees being classified as open water but does not appear to be a significant problem, however, this has not been rigorously evaluated.

Pixel brightness could be considered and those pixels falling below a brightness threshold could be considered as “No Data” areas and excluded from further analysis. This has not been done in the current project.

Conclusions

The model trained during this project has shown good results across all classes on the holdout dataset. Most false positive inferences were limited to the “tussock” and “tree” classes, though we also noticed some confusion delineating the “ground”, “grasses” and “sedge” classes.

Key areas that result in confusion both in the model and during human annotation were tea tree recruits and tussock as these are difficult to distinguish from arial photography. Vegetation below the surface of the water also resulted in confusion and was a common cause of reed being misclassified as the “tree” class.

The process and models developed here may be applied to similar imagery without modification if classes are consistent with those present in the Bools and Hacks lagoon.

Acknowledgements

We would like to express our gratitude to the Minister for Environment and Water for the introduction of the Friends of Parks Partnership Grants Program. The conservation of biodiversity, Aboriginal cultural heritage, and European heritage activities across Australian national parks and reserves is of international importance and we appreciate the program’s flexibility in supporting this wide range of conservation activities.

We extend our sincere appreciation to all those involved in the creation and implementation of this program, as well as the dedicated FoP groups whose passion and dedication drive positive change.