Funded through the
*Murray–Darling Water and Environment Research Program*

# Nesting waterbird colony population size monitoring using automated counts on drone imagery


Prepared for the Murray–Darling Basin Authority
by
Ramesh Raja Segaran, Dillon Campbell, Molly Ellis Hennekam, Marnie Denlay

Unmanned Research Aircraft Facility

The University of Adelaide
Adelaide, South Australia

June 2022


make history.

# Table of Contents

List of Figures 2

List of Tables 2

Executive Summary 3

Acknowledgements 4

1. Background 5

2. Scope 5

3. Methods 6

3.1. Data collection 6

3.1.1. Study location 6

3.1.2. Equipment 7

3.1.3. Image capture 7

3.2. Automated detection tool 8

3.2.1. Tool functionality 8

3.2.2. Training tool for Mullins Swamp imagery 9

4. Model results and evaluation 11

5. Discussion 13

5.1. Model accuracy, error and uncertainty 13

5.1.1. Data capture 13

5.1.2. Model performance 13

5.2. From demonstration to implementation 14

6. Data capture guidelines 16

6.1. Civil aviation regulations 16

6.2. Flight planning considerations 16

6.2.1. Imaging objectives 16

6.2.2. Defining the imaging extent 17

6.2.3. Automated flight mission parameters 17

6.2.4. Nominal spatial resolution 18

6.3. Wildlife disturbance risk 18

6.4. Imaging procedure 19

6.4.1. Mission planning 19

6.4.2. Performing flights 20

**7. Model training guidelines 21**

7.1. Data preparation 21

7.1.1. Setting parameters 21

7.1.2. DotDotGoose 21

7.1.3. Labelling 21

7.2. Training the model 22

7.3. Implementing a trained detector model 23

7.4. Evaluating model performance 23

**References 24**

# List of Figures

Figure 1: Mullins Swamp survey site in South Australia ....................6

Figure 2: Diagram displaying the training and detection processing pipelines for the tool ....................9

Figure 3: Manual vs Predicted bird counts for images in the hold-out evaluation set ....................11

Figure 4: Examples of image slices with predicted bounding boxes and associate confidence value .....11

Figure 5. Predicted vs Manual bird counts for the hold-out set captured at the Mullins Swamp Northern Lake site. ....................12

Figure 6. An example of overcounting within a hold-out set from a separate colony and location to the training set. The gaps in the substrate floating on the water surface display features similar enough to be counted as a detection. ....................12

Figure 7: Line of sight regulatory classifications ....................16

# List of Tables

Table 1: Imaging parameters used to capture dataset used for detection tool development ....................7

Table 2: Compatibility of DJI flight planning apps with aircraft series ....................19

Table 3: Flight and image parameters for all image datasets collected during drone imagery field campaign ....................25

# Executive Summary


Project Objectives

The Murray-Darling Basin Authority (MDBA) requested the development of an automated tool for counting nesting waterbirds in the Murray Darling Basin to be delivered in a useable format for implementation across various locations and target species. To address these needs, this project undertook the development of a tool for a nominal species (Straw-necked Ibis) at a known breeding site (Mullins Swamp, near Beachport South Australia). In the Spring of 2021, a field campaign was undertaken by the Unmanned Research Aircraft Facility (URAF) to collect drone imagery of the colony. The aim was to establish the flight parameters and image characteristics necessary for automated detection of individuals. In collaboration with the Australian Institute of Machine Learning (AIML), an automated detection tool using machine learning was developed. The drone imagery collected was used to train, test and evaluate a model (in the detection of Straw-necked Ibis) to demonstrate its potential.

Outcomes

The considerable potential for accessible data collection and analysis techniques was demonstrated across the workflow.

- off-the-shelf drone-mounted cameras used to capture extremely high resolution images (~ 6 mm) at scales suitable for imaging extensive colonies
- development of machine learning software suitable for users with minimal expertise on a desktop computer
- utilisation of images collected in combination with machine learning software to train a detector model that estimates the Straw-necked Ibis colony population size with >90% accuracy

Deliverables

- raw drone imagery of Straw-necked Ibis colonies collected at Mullins Swamp captured with varied parameters, under a range of conditions
- image analysis software for operation in-house at the MDBA
- a demonstration Straw-necked Ibis detector model trained with the Mullins Swamp drone imagery
- operational guidelines for subsequent drone image capture and use of automated image analysis tool

# About this report

This work was undertaken as a part of the Murray–Darling Basin Water and Environment Research Program (MD–WERP).

MD–WERP is a $20M Australian Government initiative to strengthen scientific knowledge of the Murray–Darling Basin. It is managed through a partnership between the Department of Climate Change, Energy, the Environment and Water, the Commonwealth Environmental Water Holder and the Murray–Darling Basin Authority.

For more information visit: www.getinvolved.mdba.gov.au/md-werp

# Acknowledgements

We would like to thank John Mullins and Moira Neagle for the permission and facilitation of the drone image capture campaign at their property, and their generosity in time and knowledge during our trip to Berri.

Steven Humphris for field assistance during drone image capture campaign.

Milad Abou Dakka and Aaron Lane at the Australian Institute of Machine Learning (AIML) for machine learning methods and software development.

Jarrod Hodgson and Rowan Mott for sharing their invaluable experience and expertise in project ideation.

Gabrielle Hunt for project management and a collaborative approach to ongoing developments through the course of this project.

Cite as:

Raja Segaran, R., Campbell, D., Hennekam, M.E., Denlay, M. (2022) *Nesting waterbird colony population size monitoring using automated counts on drone imagery*. Prepared for the Murray Darling Basin Authority, Canberra, Australia

# 1. Background

The *Basin Plan 2012* outlines several commitments to maintaining and improving water-dependent ecosystems including bird breeding and recruitment. Ensuring waterbird breeding events occur is also critical to maintaining the ecological character of Ramsar sites and other wetlands. To help assess the success of water management initiatives there is need to accurately and routinely monitor waterbird numbers and breeding events, particularly to:

- determine whether environmental watering events successfully supported and / or triggered a breeding event, and
- regularly assess colonial breeding success numerically.

Furthermore, there is a lack of systematic continuous surveys of breeding, making Basin-wide analysis and trend analysis difficult.

The Murray-Darling Basin Authority (MDBA) contracted the development and demonstration of a tool to count colonially nesting waterbirds automatically using off-the-shelf drones and cameras.

# 2. Scope

The MDBA sought to have a tool developed to automate counts of nesting waterbirds in the Murray Darling Basin. The tool was to consist of two key components.

1. Guidelines for image characteristics and data collection requirements for sub-contractors engaged by the MDBA as needed
2. Data processing software to be run in-house for the MDBA
    a. to detect and count a target species of nesting water bird colonies from appropriately collected drone imagery
    b. with capability to train additional detectors for other species in-house with appropriate data
    c. able to be run with minimal expertise
    d. suitable documentation to run the tool

A demonstration detector was trained to detect and count a selected species (nominally the Straw-necked Ibis (*Threskiornis spinicollis*) in Mullins Swamp. A drone imagery collection campaign was designed and performed to identify the imaging characteristics necessary to train a model for the target species. The field campaign informed the data collection protocol to specify imaging characteristics such as camera specifications, resolution, altitude and coverage. The detection tool pipeline and protocol enable training additional detectors for other target species or for other geographies / contexts with in-house image labelling.

# 3. Methods

The imagery for this project was collected by the Unmanned Research Aircraft Facility (URAF). The detection model software was developed by URAF and the Australian Institute for Machine Learning (AIML).

The methods sections describe the data collection procedures, followed by the process of training the detector.

## 3.1. Data collection

This section describes the drone imagery capture methods to generate data of indicative quality to successfully train the object detection model.

Drone flights were conducted by pilots holding Remotely Piloted Licences (RePL) under the University of Adelaide Remotely Piloted Operators Certificate (ReOC) in accordance with Civil Aviation Safety Authority (CASA) regulations. Animal ethics approval was sought prior to the commencement of the field campaign (University of Adelaide Ethics Approval 35375).

### 3.1.1. Study location

The original study site was planned to occur at a test site within the Murray-Darling Basin. However, COVID-19 interstate travel restrictions led to relocating the study site to Mullins Swamp, near Beachport, South Australia (-37.5099, 140.1482). This site was host to a Straw-necked Ibis breeding colony with an estimated 20,000-30,000 birds nesting at the site in the Spring.

**Figure 1: Mullins Swamp survey site in South Australia**

### 3.1.2. Equipment

Image acquisition was performed with widely available off-the-shelf equipment to enable future acquisitions to be performed by third party drone operators.

Drone

DJI Matrice M300: Approx. weight 6.3 kg (without payload); Flight time: approx. 40 minutes. Manual and automated flights controlled using DJI Pilot ground station software.

The aircraft requires that the operator hold a Multirotor <25kg Remote Pilot Licence (RePL) operating under a Remotely Piloted Aircraft Operators Certificate (ReOC) (see CASA for further regulation details).

Camera

A Zenmuse P1 RGB camera with a 45 megapixel full-frame sensor. The camera is able to accommodate the DL / DL-S lenses ranging from 16mm to 50 mm. The gimballed camera native to the DJI Matrice M300, with real-time video feed to the ground-station. Each image is embedded with a geotag of the aircraft coordinates at the time of its capture.

### 3.1.3. Image capture

Imagery was captured at Mullins Swamp in mid-Spring on the 27th and 28th of September 2021 to coincide with a known breeding event for approximately 20,000 Straw-necked Ibis. Imagery was primarily collected over the central area of Mullins Swamp, referred to as Mullins Swamp Mid-lake. Some opportunistic additional data was collected at a breeding site at the Mullins Swamp Northern Lake, 1 km north of the Mid-Lake.

Mission planning

Manual reconnaissance flights with live video from the aircraft were used to establish the extent of the breeding colony. The 25 ha extent provided a conservative 50 m buffer to ensure that the whole colony was captured.

The boundary of the extent was entered into the flight planning software DJI Pilot. The software is primarily used for mapping purposes to capture imagery for photogrammetry. It creates a flight plan following parallel transects to cover the desired imaging extent. The user specified altitude, overlap and sidelap are used by the flight planning software to compute nominal spatial resolution, flight speed, as well as capture intervals along and between transects. The parameters also account for the camera specifications such as sensor resolution and size, lens focal length and capture rate.

The parameters were selected to capture imagery of:

- adequate spatial resolution to identify discriminable features of adult individuals such as shape, size and colour
- each image is adjacent to one another with no overlap or gaps between them
- maintain as much altitude as practicable to avoid distressing or harming wildlife

**Table 1: Imaging parameters used to capture dataset used for detection tool development**

| Date and time | Area | Overlap and sidelap | Altitude | Lens focal length | Image footprint | # of images | Data volume | Nominal spatial resolution | Flight duration |
|---|---|---|---|---|---|---|---|---|---|
| 28/09/2021 1015 | 25 ha | 10% | 70 m | 50 mm | 50 x 33 m | 205 | 3.53 GB | 6.2 mm | 6 mins |

Additional data

Numerous image datasets were collected during the campaign, trialling various capture parameters including altitude, lens focal length and time of day. These data were not analysed, but are included in the project data deliverables package (Appendix A). The highest nominal spatial resolution data is used throughout this report, however there is considerable opportunity to explore the model sensitivity to data captured under those variations in capture parameters. If it could be established that lower resolution image data yields comparable / adequate population estimation precision, the outcome would result in more efficient field costs, data volume and processing times.

As part of the data collection campaign, imaging was performed in accordance with standard photogrammetric practice involving 70% overlap and sidelap. They were used to generate a georectified orthomosaic. The orthomosaic was used as a visualisation basemap. Although beyond the scope of this project, the orthomosaic has potential for comparative analysis of population estimates between the low overlap and orthorectified datasets. Although photogrammetric methods involve considerably greater time, data and compute costs, they can present opportunities for spatially explicit analyses of potentially ecologically meaningful monitoring methods such as colony shape, size, location and patterns across space and time. The automated detection tool does not currently support the orthomosaic data structure, but it could be enabled with minor modification to the code. The current functionality of the tool uses individual images which avoids the prohibitive barriers that are inherent in the use of orthomosaics.

## 3.2. Automated detection tool

The machine learning automated detection tool is software collaboratively built by AIML and URAF. The primary functions of the tool allow users with minimal programming expertise to train a model capable of detecting target species and to then use that model to automatically count the target species. A demonstrative trained model was built to count Straw-necked Ibis in a breeding colony in Mullins Swamp Mid-Lake using the imagery described in Section 3.1.3.

### 3.2.1. Tool functionality

The workflow used for automated object detection involves three relatively discreet components: training a detector model, implementing the trained detector model, and detector model performance evaluation.

The initial component involves training a detector by providing it with examples of the object to be detected. Detector model training is performed by a user systematically identifying every individual in the images used for training. The identified individuals are labelled by drawing boxes over the individual as precisely as possible to encompass all pixels occupied by the bird, avoiding pixels that do not.

Within the training component of the tool, sub-modules allow the user to:

- organise and filter images before labelling
- label the images
- prepare the labelled datasets in a format which can be ingested by the model training component

The training component contains a module to initialise an instance of Label-Studio, an open-source data annotation application for machine learning. The application includes features for manually interpreting images and annotating the target species with bounding boxes. The downstream Python modules ingest the annotations exported by Label-Studio.

The labelled images and labels are used to train a detector, which generates a model that can be used to identify similar objects in other imagery. The trained detector model ingests images where it detects the objects, referred to as predictions. The output is a sum of the number of predicted detections it has made for each image.

Detector model performance assessment requires that an independent evaluation set of labelled images, unseen by the trained model, are run through the trained detector model. The sum of labelled individuals in each image in the evaluation set is compared to the sum of predicted detections for the same image. The level of agreement between those counts is represented by a line of best fit in a linear regression.

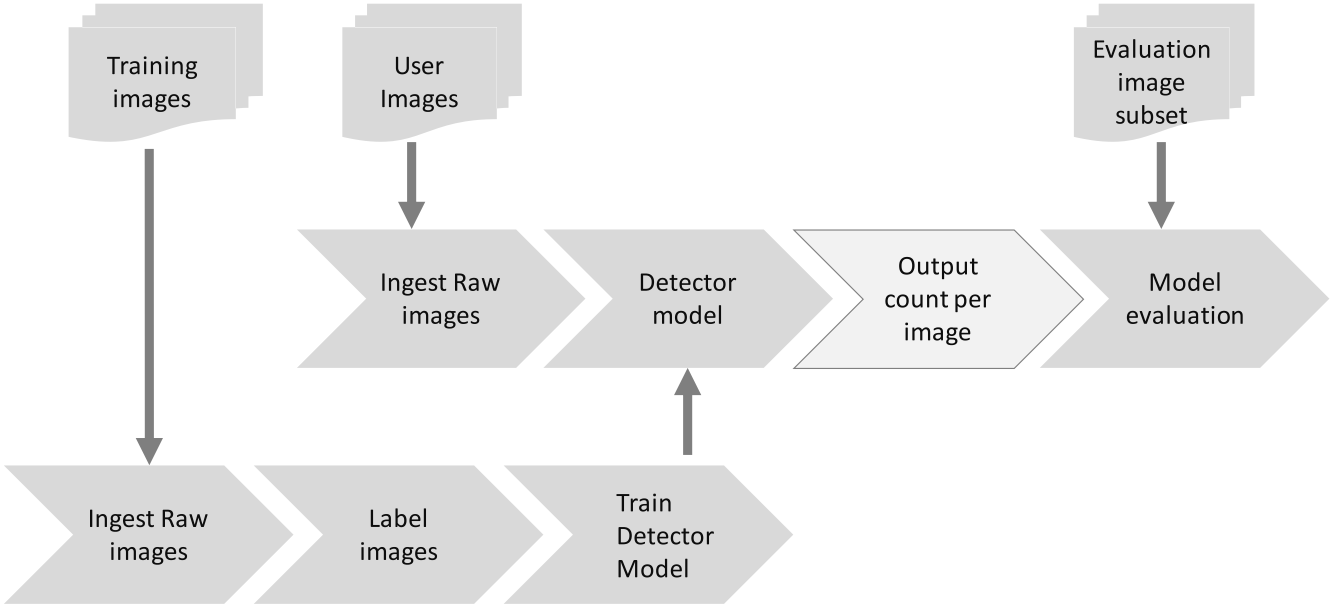


**Figure 2: Diagram displaying the training and detection processing pipelines for the tool**

### 3.2.2. Training tool for Mullins Swamp imagery

A model was trained to detect Straw-necked Ibis using the images collected at Mullin's Swamp on 28th September 2021. The following describes the workflow used to train the model.

1. 205 raw images captured (data capture as per Table 1)
2. Filter out the images with no birds
    a. Mark every bird in all images with points using DotDotGoose
    b. 88 images with no birds marked discarded
    c. 117 images containing birds left to use to train the model
3. Create calibration and evaluation subsets
    a. 17 images randomly selected for a hold-out evaluation subset
    b. 100 images used for training
4. Each of the 100 images were split into 160 (16 columns x 10 rows) smaller non-overlapping slices as required by the FasterRCNN base network
5. The slices were filtered using the DotDotGoose point file to exclude slices without birds from the labelling set
6. The slices containing birds were labelled using Label-Studio

7. A Region Proposal Network (RPN) was used with a FasterRCNN base network to train the object detector, ingesting the labelled images
8. Empty slices were added into the training set to reduce false positive detection bias and successfully reduced overcounting
9. To avoid overfitting, early stopping patience was employed to reduce the number of training epochs
10. Performance between models was compared and the selected model was chosen based on the highest Mean Average Precision (mAP) on the validation set

# 4. Model results and evaluation

The counts in the 17 hold-out evaluation images were compared against the predicted counts of the detector model. There is agreement between the two counts (Figure 3), with a strong linear slope (y=1.009) and correlation $R^2 = 0.997$, resulting in an overall error of 3 %. Figure 4 illustrates the outputs of the predicted bounding boxes for the individuals detected and the annotated confidence that the predicted object matches the trained object.

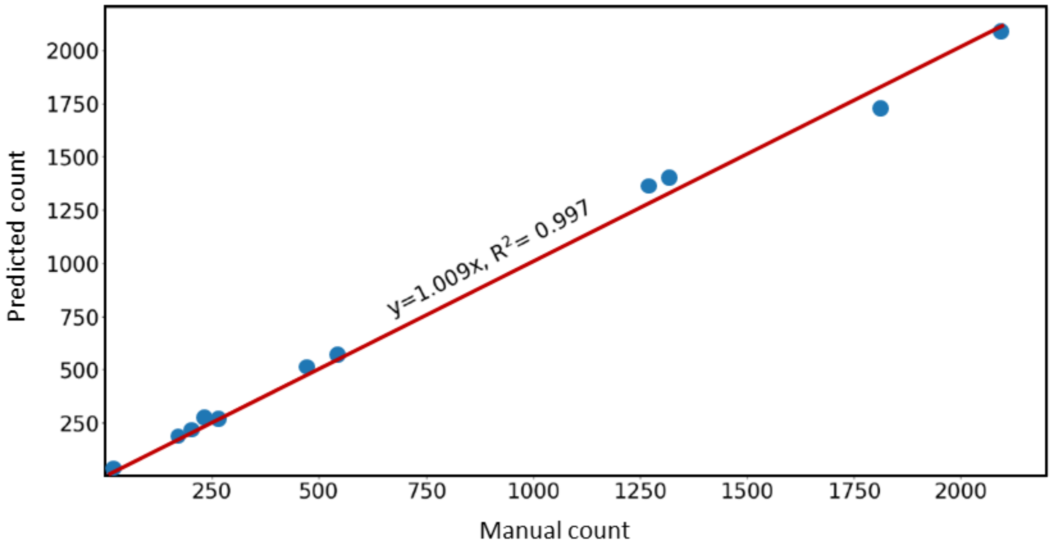


**Figure 3: Manual vs Predicted bird counts for images in the hold-out evaluation set**

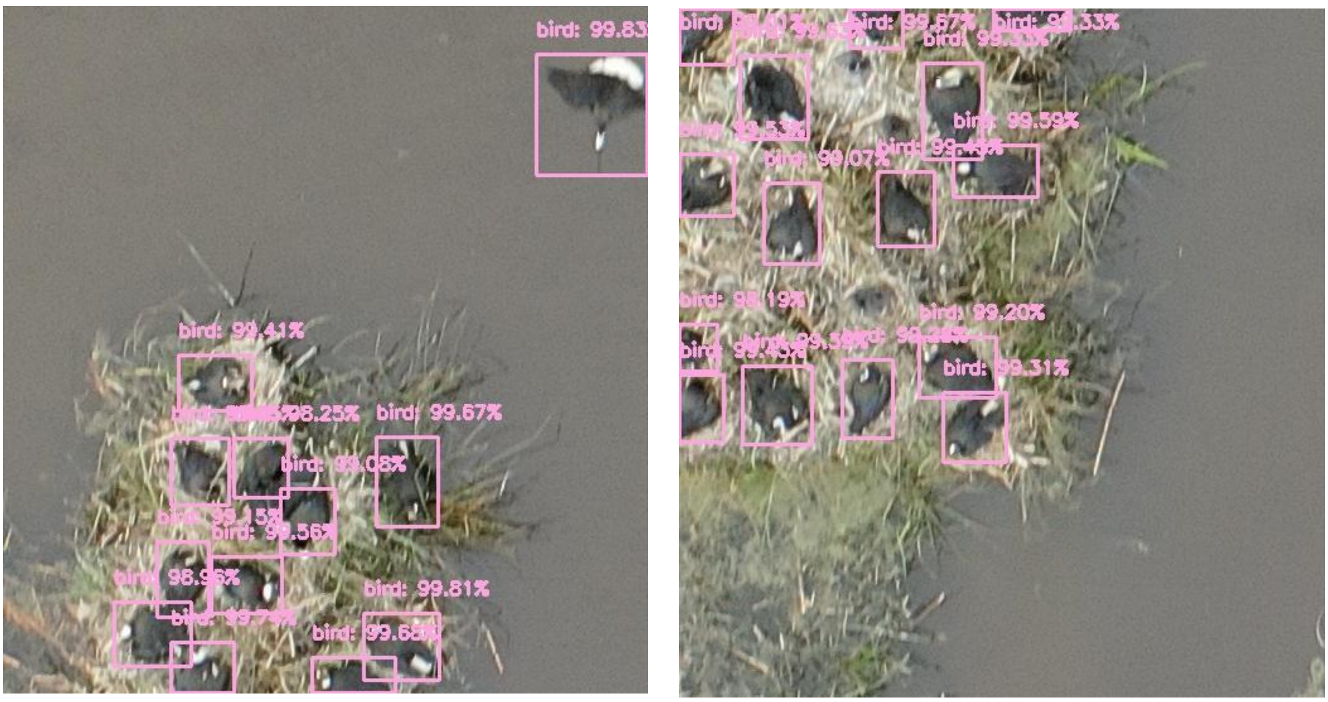


**Figure 4: Examples of image slices with predicted bounding boxes and associate confidence value**

An opportunistic dataset was collected at a separate Straw-necked Ibis breeding colony north of the original site. It allowed a test of the model performance under slightly different environmental conditions. The seven images containing birds were run through the detector model. The number of birds predicted by the model show strong linearity between the manual counts and 8.2% error on the predicted counts (Figure 5). A large proportion of the error was due to false positive detections, most likely due to differences in colour and texture in the substrate (Figure 6). This form of overcount may be reduced by retraining the model with additional blank slices from this site in the training phase or by adjusting the detection confidence threshold.

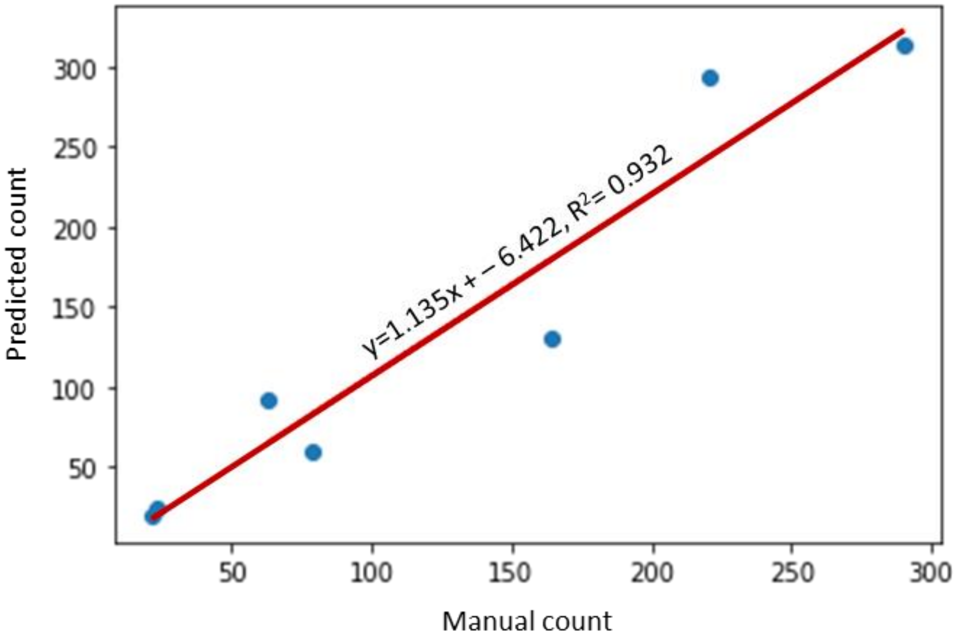


**Figure 5. Predicted vs Manual bird counts for the hold-out set captured at the Mullins Swamp Northern Lake site.**

**Figure 6. An example of overcounting within a hold-out set from a separate colony and location to the training set. The gaps in the substrate floating on the water surface display features similar enough to be counted as a detection.**

# 5. Discussion

The trained detector model outputs demonstrate the capacity for the tool to detect individual Straw-necked Ibis in the drone images with a high level of accuracy compared to those manually identified by a person. Numerous factors contributed to the accuracy. These include factors relating to image capture, the target species, site specific factors and model training. The imaging and model training guidelines in Sections 6 and 7 respectively outline key workflow considerations for future projects.

## 5.1. Model accuracy, error and uncertainty

The reported high accuracy metrics are related to the sum of individuals present in each image. The imagery was captured with adequate resolution to identify features of the target species and discriminate them from other landscape features present in the imagery. Lighting conditions during capture were sufficient to capture sharp, low noise images. However, the high detection accuracy metrics do not account for various sources of error and uncertainty in the data capture and tool implementation components of the workflow. Some of these uncertainties can be mitigated, while others are currently unavoidable and should be considered when interpreting the results.

### 5.1.1. Data capture

The high level of agreement between human and model predictions are related to the individuals present in the images. Factors contributing to the encouraging results include the nominal spatial resolution of the imagery relative to the Straw-necked Ibis, features in the appearance of the species that are distinctive relative to the background substrate and their congregation in an open landscape with no occluding vegetation.

However, some disparity exists between the ‘true’ colony population size and what is represented in the imagery as a proportion of the colony in flight / transience were not captured.

The imaging mission was performed in a pattern of parallel transects over 6 minutes. During the mission it is likely that individuals moved during the mission resulting in either being imaged twice or not at all. This is expected to have negligible impacts on the overall count in this case but could present greater issues for colonies with different behavioural patterns.

During the imaging missions a considerable proportion of the colony was also in flight, engaging in thermal gliding in a column rising to ~200 m above the colony (>100 m above the drone during capture). Our approximate estimation of birds in flight is >1,000, almost all of which will have avoided being imaged, while some may have been imaged as birds in flight. The consequence is an underestimation of final population count and is considered an unavoidable source of error.

Colony behaviour is expected to vary between habitats and monitoring epochs as a response to environmental conditions such as temperature, vegetation type etc. These unaccounted variations will affect the comparability between monitoring events. To mitigate these impacts, observations of colony behaviour during data capture should be included in metadata at the time of data capture.

### 5.1.2. Model performance

The model evaluation metrics indicate high detection accuracy for the dataset the model was trained on in the Mid-Lake. While the accuracy remained relatively high for the Northern Lake dataset, greater error was observed. The sources of increased error are expected consequences of variation in habitat structure and lighting conditions, such as the false positive detection of shadow in Figure 6. These errors are analogous to what is to be expected for subsequent monitoring events. Although not addressed in this study, colonies comprised of a species mix are likely to impact detection accuracy and may benefit from the following actions: increase the nominal spatial resolution of the imagery to improve discrimination between the features; append additional training data from the site or increase overall model training sample size. It is recommended that for any monitoring event data, a subset of the data is labelled for use in model evaluation for quality assurance of that dataset.

If the error increase exceeds the desired limits (subject to the specific monitoring objectives), the following actions are expected to improve detection accuracy:

- Update the established detector model with a subset of data representative of variations in environmental conditions.
- *Data augmentation* in machine learning is a process that can reduce detector model sensitivity to variations in the target, image quality and environment. Data augmentation increases training data sample size and variability by artificially distorting the input training data. Although not implemented in the tool delivered, data augmentation can be accessed with some additional software code.

## 5.2. From demonstration to implementation

The workflow developed to image and detect Straw-necked Ibis at Mullins Swamp yielded promising results due to careful consideration of methods for image capture, species and habitat selection and the development of the automated detection software and model. Similar results can be expected for subsequent monitoring projects that meet comparable target species and habitat. Sections 6 and 7 below sets out initial guidelines for data capture and model training to adapt the workflow for candidate monitoring projects.

This project aimed to demonstrate and deliver a methodology for mature off-the-shelf drones, cameras and machine learning software to improve nesting waterbird colony population size estimations. The workflow presented is accessible to non-expert users. While the technical barriers to adopting these technologies have been lowered considerably, some domain knowledge and experience in the design, implementation and interpretation is critical for meaningful and reliable monitoring projects.

The design phase should determine what level of accuracy, error and uncertainty are acceptable for the specific monitoring objectives. The variation of all proposed locations and times need to be accounted for at this stage. Those factors will influence the data capture parameters such as the nominal spatial resolution as it relates to the target species, habitat and environmental factors. Conservatively high resolution requirements will inflate the costs of fieldwork, data storage and data processing, while inadequate resolution will compromise the reliability of detection and counts. Critically, it should be established very early in planning whether these imaging and processing methods are suitable for the target species and habitat.

Personnel performing the data capture will require a combination of technical and ecological knowledge on site as immediate problem solving will almost certainly be necessary at the time of capture. Technical expertise is needed for sound imaging parameters choices to accommodate the site and environmental contexts. Ecological knowledge is needed to determine the colony and imaging extents, to avoid introducing unnecessary distress or harm to wildlife, and to take note of behaviours that may have consequences on the count accuracy.

The data processing phase requires decisions on adequate sample sizes for model training and evaluation. Both samples must adequately represent the colony, habitat and environmental variation. Interpretation of results needs an appreciation of the accuracy and uncertainties inherent in the data quality, field conditions and the model performance to derive meaningful information.

The personnel involved in designing and implementing the entire workflow to suit the monitoring objectives, target species and habitats will require some capacity building. Many of the nuances across the workflow can only be gained through trial, implementation and inevitable error. However, the learning process could support key monitoring personnel obtaining a high level understanding of the principles underlying the components in the workflow. Expert led workshops can be developed to cover the fundamentals of drone operation, image capture and the machine learning methods used in the tool.

We recommend that protocols are developed across the workflow. Protocols for data capture and image analysis streamlines planning and implementation, produces consistent and comparable outcomes, and optimises adapting to lessons learnt. Protocols setting standard procedures and parameters for typical environments and target species need to be co-designed by ecologists, analysts and decision makers.

# 6. Data capture guidelines

The objective of the drone imaging campaign is to capture images to be used for the automated detection and counting of the target bird species. Satisfying the objective encompasses the consideration of factors including the visual and behavioural characteristics of the target species, responding to dynamic environmental conditions and site accessibility, while complying with relevant civil aviation regulations.

## 6.1. Civil aviation regulations

An essential initial consideration in the planning phase is to establish that contracted third party drone operators have the capabilities to perform the operations to a high standard. The operator must possess the necessary certification, insurance, equipment and skills to perform the intended operations. To operate in accordance with Civil Aviation Safety Authority (CASA) regulations, it is most likely that the operator must hold a Remote Pilot License, a valid Remote Operator's Certificate (ReOC) and adequate aviation, liability and indemnity insurance. Operators with inadequate certification and insurance cover present a potential for reputational risk by association. Additionally, flying drones over publicly held land typically require the above certification to obtain permits to operate.

A regulation that may present a barrier to operations is the requirement to maintain visual line of sight (VLOS) of the drone. VLOS regulations are likely to impact imaging large areas or sites with access limitation that force setting a take-off / landing base location a great distance from the imaging extent. An increasing number of drone operators can operate with extended visual line of sight (EVLOS) and beyond visual line of sight (BVLOS) capabilities (Figure 7).

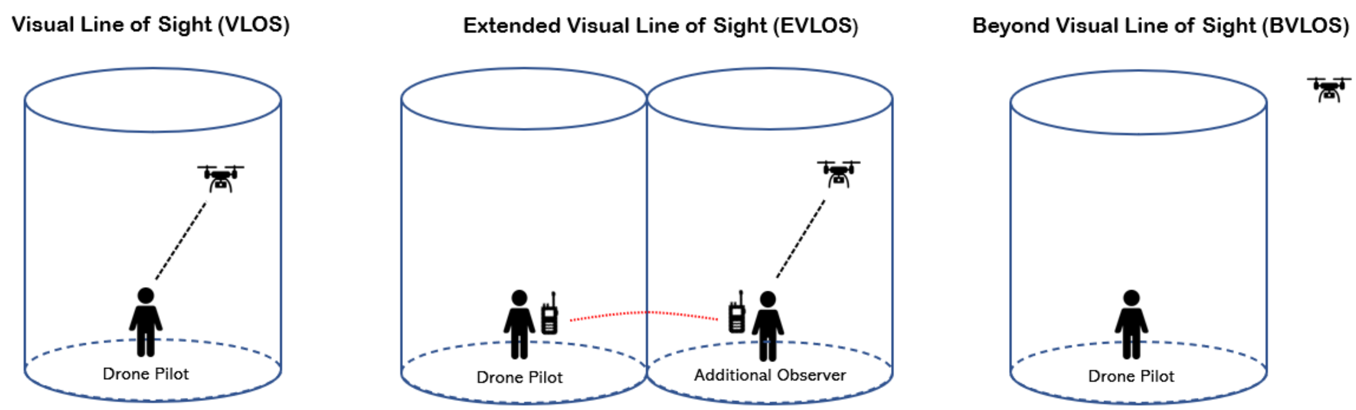


**Figure 7: Line of sight regulatory classifications**

## 6.2. Flight planning considerations

### 6.2.1. Imaging objectives

An appropriate imaging campaign includes choosing imaging parameters needed to have enough, but not too much, visual information to discriminate the individuals of the target species. Adequate image quality allows for the clear discrimination of individuals in the target species against all objects in the image. Some of the factors that influence detection accuracy include:

- Physical size of the target species
- Variability in the appearance and size of the target species
- Mix of species within the colony
- Texture, shape and markings of target species that discriminate them from non-target species
- Contrast between the target species and the background substrate (soil, nesting vegetation, water)
- Clear visibility of the whole individual bird with no or minimal obstruction to visibility from overhanging vegetation
- Solar illumination conditions for each monitoring epoch and variability between them

Adapting the imaging parameters described below can be tailored to the conditions relating to the above factors for specific monitoring species and habitats. However, there are real limits to the conditions which, once exceeded, would lead to error, costs or harm that exceed acceptable levels for the specific monitoring objectives.

### 6.2.2. Defining the imaging extent

The imaging extent will need to be defined after the shape and distribution of the target colony has been established leading up to or immediately preceding the monitoring event. The definition of the imaging extent should consider the nesting characteristics of the species for the specific colony, habitat structure and the monitoring objectives which will require some familiarity with the target species.

Defining the imaging extent for relatively dense populations in discrete and easily identifiable colonies is straightforward. The extents can be defined by ground survey where the necessary coordinates can be recorded prior to the drone operator departing to the site. Alternatively, the extent can be established on a reconnaissance manual drone flight using the live video feed to record the necessary coordinates.

More complex scenarios where variable density, species mix and colony location over large extents will require greater consideration in the definition of the modelling objectives. Defining these extents may require that the drone operator has access to the necessary expertise to assess the trade-offs between the drone and camera capabilities, regulatory compliance, ecological characteristics, and monitoring objectives. Some cases may be require considering alternatives to achieving coverage of the entire colony, such as robust sampling design.

Although the majority of these issues should be discussed in the planning phase, it is impractical to account for all of the realities present onsite that require *ad hoc* decision making at the time of capture.

### 6.2.3. Automated flight mission parameters

The relevant parameters for specifying the automated mission parameters for image capture are ubiquitous for all mapping mission planning software packages. The mission planning software computes the flight parameters for the mapping extent and the drone / camera capabilities based on user specified inputs relating to of the following parameters:

- Capture altitude which determines the nominal spatial resolution for the camera and lens focal length specifications
- Image overlap and sidelap

### 6.2.4. Nominal spatial resolution

Spatial resolution refers to the unit of area covered by a square pixel in an image represented by the width or length of the area. In photogrammetry, the orthorectification process ensures that the spatial resolution of the imagery is consistent across the orthomosaic. As the images are not orthorectified, their spatial resolution is not consistent, varying within and between each image due to geometric distortions from the camera and variable terrain. The spatial resolution, also referred to as ground sample distance, reported in the automated flight planning parameters are only indicative and are therefore referred to as nominal spatial resolution.

The nominal spatial resolution of the images is a function of the camera resolution, lens focal length and the aircraft altitude (further information in this link).

The appropriate nominal spatial resolution for the target species and habitat will vary as a consequence of the factors listed above in Section 6.2.1. No established protocols currently exist to determine the necessary resolution for the target species size and discernible features. However, based on the success of the Mullins Swamp Straw-necked Ibis demonstration detector model, it can be used as baseline for the nominal spatial resolution for the target species. The mean dimensions of the bounding boxes used to train the Straw-necked Ibis detector model was ~58 x 40 pixels, or 2,500 pixels for imagery with a nominal spatial resolution of 6.2 mm. Consequently, a smaller target species may be expected to perform comparably with higher resolution by specifying a lower image capture altitude.

Overlap and sidelap

Overlap and sidelap refer to imaging parameters necessary for photogrammetry applications. Overlap and sidelap are expressed as the percent of common area captured by adjacent images, which is necessary to generate orthomosaics. For this application, overlap and sidelap would ideally be set to 0% to capture images contiguously with no overlap or gaps to avoid double imaging or missed imaging of individuals. However, the DJI Pilot automated mission planning software minimum is 10% overlap and sidelap. Consequently, the processing software developed for this project allows for the cropping of overlapping portions of the imagery.

If aircraft Realtime Kinematic (RTK) Positioning is available, we recommend its use to allow for precise relative positioning of imaging locations to avoid over / under capture of overlap and sidelap.

## 6.3. Wildlife disturbance risk

The risks of disturbance to the target species and the habitat it occupies are critical considerations that must be addressed in the planning phase of monitoring for each species and habitat. Drone operations pose the real risk of stressing breeding and nesting species that may result in considerable adverse impacts on the success of the breeding event. The most obvious consequence is the drone colliding with birds in flight, either fleeing the site or during normal activity, resulting in harm to the birds and damaging the drone.

Stressing breeding colonies can result in reduced sexual activity and the abandonment of nests during incubation. The impacts may range from short lived evasive behaviours that quickly resume normal activity to the fleeing the habitat for periods that are long enough to compromise the breeding event (Lyons et al, 2018). Moreover, the response behaviour may result in collision with the drone, posing damage to the birds and the aircraft.

Knowledge and evidence the behavioural responses of wildlife to drone activity is sparse and therefore difficult to assess. Responses to stress from drone activity can range from immediately fleeing the site to inducing acute stress that have no visibly observable changes to behaviour (Weimerskirch, Prudor & Schull, 2018; Lyons et al, 2018).

While some risks of disturbance can be evaluated in the planning phase, due to the lack of knowledge on behavioural responses for each species, the assessment of disturbance must be monitored during the monitoring event by personnel with adequate knowledge. This will include assessing colony behaviour prior to, during and following the drone operations. Importantly, the risks should extend the consideration of risks beyond the target species to include other wildlife present at the site. It is recommended that an evidence-based protocol of behavioural responses to be observed and thresholds that trigger the abortion of the drone operation are developed. Additionally, the behavioural response of the species present should be included as metadata and published to contribute to the research and monitoring community.

Some of the possible risk mitigation actions include minimising the operational time to reduce the potential for disturbance and minimising the proximity of the aircraft to the colony by increasing altitude. These can be addressed by selecting appropriate drone / camera equipment. For example, the equipment used to capture data for the Straw-necked Ibis (Section 0) can accommodate the use of lenses with greater focal length which allows for greater capture altitudes to achieve adequate spatial resolution; has camera capture rates of up to three times faster than alternatives; and greater flight time to reduce the length of time in flight and the number of flights into and out of the monitoring area.

Due to the idiosyncratic response of wildlife to drone disturbance and lack of knowledge regarding the potential impacts of disturbance, it is strongly recommended that any operations are conducted in accordance with the best practices guidelines defined by Hodgson and Koh (2016). These guidelines are designed to minimise disturbance of drone flights on wildlife when there is insufficient knowledge to make evidence-based decisions.

## 6.4. Imaging procedure

### 6.4.1. Mission planning

The method for flight planning will depend on the aircraft used. DJI are currently market leaders for consumer and commercial grade drones and support ‘native’ integration of sensors from a range of manufacturers.

DJI currently provide flight planning software that are supported by specific aircraft (Table 2).

**Table 2: Compatibility of DJI flight planning apps with aircraft series**

| DJI Flight planning apps | Aircraft series compatibility | | | | |
|---|---|---|---|---|---|
| | DJI Matrice 600 | DJI Matrice 300 | DJI Matrice 200 | DJI Inspire 2 | DJI Phantom 4 |
| DJI Pilot | Yes | Yes | Yes | No | Yes* |
| DJI GS Pro | Yes | No | Yes | Yes | Yes |

* DJI Phantom 4 RTK only

The DJI Matrice M300 equipped with a Zenmuse P1 camera is recommended due to its technological maturity, accessibility and advanced capabilities. However, most DJI aircraft (excluding the Mavic 3 series) are compatible with flight planning / mapping software. The major drawback of using other aircraft is the requirement to fly lower and longer to achieve the same coverage and resolution. Coverage and resolution estimates for different platforms, sensors and lenses can be simply calculated using free flight planning software to assist with decision making.

Planning flights on the M300 with the DJI Pilot app is straightforward, requiring user parameters that are standard across any drone mapping mission planning software. DJI Pilot allows for an imaging extent to be created by interacting with the flight controller touchscreen to draw the boundary. Alternatively, pre-fight survey delineating the colony extent can be imported as a KML file. Parameters for flight altitude, overlap and transect orientation, among others, can be set in DJI Pilot.

A copy of Table 1 presents the imaging parameters used to capture the data for the Straw-necked Ibis demonstration model.

| Area | Overlap and sidelap | Altitude* | Lens focal length | Image footprint | # of images | Data volume | Nominal spatial resolution | Flight duration |
|---|---|---|---|---|---|---|---|---|
| 25 ha | 10% | 70 m | 50 mm | 50 x 33 m | 205 | 3.53 GB | 6.2 mm | 6 mins |

*Altitude refers to the height of the aircraft above ground level (AGL) over the imaging extent. The height of the take-off location relative to the imaging extent must be accounted for in the mission plan.

### 6.4.2. Performing flights

As mentioned in Section 6.1, it is essential that all flights are performed in accordance with CASA regulations. Compliance considerations and exemption submission processing times must be identified in the initial planning phase.

Maintaining situational awareness in surrounding airspace and also on the ground is important to complete the flight in a safe manner. Maintaining awareness of birds in the airspace, ensuring pilot is able to respond quickly to birds that may be entering into nearby airspace. Having a spotter to monitor the airspace for raptors / birds of prey as well as manned aircraft is highly recommended. DJI Pilot and DJI Ground Station Pro both contain the ability to pause and resume flight missions in the case of another aircraft or wildlife entering the nearby airspace.

# 7. Model training guidelines

A step-by-step instruction guide to run the object detection pipeline from start to end can be found in the 'readme.pdf' provided with the toolkit.

The following components of these guidelines provide additional information to plan for and perform the analyses.

## 7.1. Data preparation

### 7.1.1. Setting parameters

The parameters in the params.yaml file included in the toolkit must be set for each project in order for training and inference modules to run correctly. As a starting point, most of the parameters can be left as default, however explicit definition of some parameters is required. Each object detection model training requires the raw image height and width in pixels be explicitly set according to the sensor used. The default image height and width (in pixel number) provided in the params.yaml template are based on the DJI Zenmuse P1 camera and will only work for that that sensor. Additionally, the name of the directory containing raw images ('project_name' in params.yaml) must be specified. It is also recommended that the slice height and width are specified as integer fractions of the raw image dimensions.

From lines 30-42 of the params.yaml, hyperparameters can be altered, although it is recommended that they are left on default for at least the first pass through the training module and can be 'fine-tuned' if necessary. A hyperparameter refers to any of the tuning parameters which influence the way the model learns.

**Note:** The default location of for input images is relative to the working directory and params.yaml file. Absolute paths can also be provided in params.yaml, but require additional manual intervention.

More details regarding the functionality of parameters can be found inside the params.yaml file.

### 7.1.2. DotDotGoose

DotDotGoose is a free, open source application with graphical user interface used to initially identify image portions that contain the target species. It is a required step to filter the image training slices that contain any individuals to be labelled and to omit slices absent of individuals. The software enables the user to place points on individuals present in the raw input images. Subsequently, DotDotGoose exports those points and the position of the individuals relative to the origin of the raw image. The points are used by the **'split_raw_dataset'** module to filter out slices absent of individuals for labelling. This process saves a significant amount of user time by avoiding searching through empty slices during the labelling phase.

Before proceeding to the next step, ensure the path of the points (.pnt) file is specified in 'params.yaml'.

Running the **'split_raw_dataset'** module will prepare the raw images into slices ready to be labelled.

### 7.1.3. Labelling

The toolkit relies on the use of Label-Studio to perform the labelling individuals of the target species used in the training component. Using alternative image labelling software requires code modification to suit its format. Step by step instructions on using Label-Studio for labelling can be found within the 'readme.pdf' instruction guide.

Multiple species within a single training set can be labelled using Label-Studio by adding additional classes to the labelling menu. **Please note:** In the params.yaml the value input for num_classes should equal the number of species + 1. The additional class represents the 'absent' class to teach the model to discriminate the background from the target species. The background information represented by the absent class is a default and does not need to be explicitly labelled.

The quality of the labelling will have a significant impact on performance of the object detector. If objects are poorly labelled, e.g. bounding boxes being too large or too small, they may include or exclude information necessary for the model to correctly learn. The most important rule is to ensure that label bounding boxes fit tightly around the target individual.

The creation of a labelling reference guide prior to fully committing to the labelling process is strongly recommended. The reference guide establishes the standards for labelling personnel in their decision making, e.g. whether birds in flight should be included. The reference should include example image subsets indicating the extent of the bounding box, objects to include and acceptable variations of the object (e.g. juveniles, the level of occlusion by vegetation). Conversely, objects to exclude should also be exemplified, e.g. other species. The document is critical to ensure consistency between labelling personnel and between monitoring events, especially for labelling hold out data subsets used for performance evaluation.

Once image labelling is complete, the **'prepare_training_set'** module can be run. This script scans through the labels created in Label-Studio and prepares a training set that conforms to the COCO format required to train the model. The reformatted labels will then be saved under a new directory (coco-detection-dataset, as default in the params.yaml file) as well as the image slices. Within this directory, the **'prepare_training_set'** module also includes a sample of empty image slices in the training set to reduce the rate of false positive detections.

## 7.2. Training the model

Once a training set has been prepared following the guidelines in Section 7.1, the model can be trained. Before running the **'train_mdba'** module, confirm that params.yaml includes the correct paths and the hyperparameter values are filled. Using the default hyperparameters should yield good results, however these can be refined if necessary.

It is possible to change the ratio of training and validation images. The default value (0.2) results in 20% of the input dataset being used for validation and 80% for training. The code currently does not allocate any of the raw images to a holdout set, so it is recommended to keep some from being ingested into the pipeline to evaluate model performance.

By default, the number of epochs is set to 10. As the model 'learns', a loss value is reported after each epoch. If the loss value continues to decrease significantly after the 10$^{th}$ epoch, consider increasing the number of epochs. The number of epochs determines how many times the training dataset is 'looked at' by the learning algorithm and can improve model performance. However, if the number of epochs is too high, the learning algorithm is susceptible to overfitting. To avoid this, a parameter in params.yaml called 'early stopping patience' has been included to limit the number of epochs if the algorithm stops learning (i.e. when the loss function plateaus).

The output of the **'train_mdba'** module is a .pth file. The .pth file is a common format used to save trained models using the python library [PyTorch](). The name and location of the model can be defined in the params.yaml file on lines 46 and 47. Changing the name of the model may be useful if the user would like to keep multiple versions of the model to evaluate performance rather than overwriting the existing model. The specified .pth file is called upon when implementing a trained detector model.

## 7.3. Implementing a trained detector model

A trained detector model can be used to ingest images 'unseen' by the model in the training process, e.g. images from subsequent monitoring events. This is referred to as *inferencing* in the field of machine learning. The **'bird_count'** module implements inferencing on new images by providing it with paths to a directory of input images and the trained detector model.

The image overlap and sidelap used in the data capture campaign are required parameters to omit detections from being included in the detection count output for each image. The parameter is specified in the params.yaml file. The proportion of overlap and sidelap between adjacent images is specified as a decimal value (e.g. 10% = 0.1).

In the process of inferencing, each detection of an object is attributed with a confidence metric of the object's similarity to what it was trained to detect. The lower threshold of the confidence metric is set with the 'confidence threshold' parameter in the params.yaml file. The default confidence threshold is set to 0.96. The parameter can be refined to tune the sensitivity to improve the performance, especially for models detecting excessive false positives. Refining the parameter is most likely necessary for imagery containing landscape or lighting differences. Higher confidence thresholds reduce the number of detections and increases the likelihood of false negative detection resulting in under-counting. Lower confidence thresholds will increase the number of detections and increases the likelihood of false positives in background features resulting in over-counting.

Data augmentation can improve the robustness of the model when subject to variability in image characteristics. Common forms of data augmentation include resampling, warping, rotating and modifying colours of the original image slices. The desired outcomes from this process improve the detectability of target species under different environmental conditions such as variations in lighting or background substrate / vegetation. Although the tool delivered in its current state does not support augmentation, the pipeline does contain a template in which common augmentation processes from the *torchvision* library can be implemented with minimal expertise (see https://pytorch.org/vision/stable/transforms.html).

## 7.4. Evaluating model performance

Model performance evaluation can be performed by running the detector model on a hold-out subset of the data. A hold-out dataset can be any images of the target species that the model has not 'seen' during the training process. For example, images from the same colony, containing the same target species, but not included in the training set. It is recommended that model evaluation is performed for subsequent monitoring data that is input into a previously trained detector model.

The toolkit includes an evaluation module **'eval_mdba'** which tests for linearity between the human user identified counts and predicted counts. The module exports a graph in PNG format with the linear regression model and R-squared value for human user vs predicted counts (e.g. Figure 3).

# References

Hodgson, J. C., & Koh, L. P. (2016). Best practice for minimising unmanned aerial vehicle disturbance to wildlife in biological field research. *Current Biology*, *26*(10), R404-R405.

Lyons M, Brandis K., Callaghan C., McCann J., Mills C., Ryall S., Kingsford R. T. 2018, Bird interactions with drones, from individuals to large colonies, *Australian Field Ornithology*, vol. 35, pp. 51 – 56.

McEvoy, J. F., Hall, G. P., & McDonald, P. G. (2016). Evaluation of unmanned aerial vehicle shape, flight path and camera type for waterfowl surveys: disturbance effects and species recognition. *PeerJ*, *4*, e1831.

Weimerskirch, H., Prudor, A., Schull, Q., 2018. Flights of drones over sub-Antarctic seabirds show species- and status-specific behavioural and physiological responses. *Polar Biol*. 41, 259–266.

# Appendix A

**Table 3: Flight and image parameters for all image datasets collected during drone imagery field campaign**

| Date and time (year 2021) | Site | Overlap and sidelap (%) | Altitude (m AGL) | Lens focal length (mm) | # of images | Data volume (GB) | Nominal spatial resolution (mm) | Flight duration (mins) |
|---|---|---|---|---|---|---|---|---|
| 27/09 1418 | Mid Lake | 10 | 70 | 35 | 59 | 1.29 | 8.8 | 7 |
| 27/09 1427 | Mid Lake | 10 / 10 | 100 | 35 | 52 | 1.09 GB | 12.6 | 3 |
| 27/09 1456 | Mid Lake | 10 / 10 | 70 | 50 | 141 | 2.85 GB | 6.2 | 7 |
| 27/09 1554 | Mid Lake | 10 / 10 | 70 | 35 | 164 | 3.78 GB | 8.8 | 5 |
| 27/09 1615 | Mid Lake | 10 / 10 | 70 | 50 | 257 | 5.52 GB | 6.2 | 8 |
| 28/09 0930 | Mid Lake | 70 / 70 | 70 | 35 | 986 | 22.0 GB | 8.8 | 23 |
| 28/09 1016* | Mid Lake | 10 / 10 | 70 | 50 | 205 | 3.53 GB | 6.2 | 6 |
| 28/09 1139 | Northern Lake | 10 / 10 | 70 | 50 | 233 | 4.49 GB | 6.2 | 7 |

*Image dataset utilised for detection tool development.