The Effectiveness of Using Drones and Convolutional Neural Networks to Monitor Populations of Nile Crocodile and Common Hippopotami

by

A.B. Woolcock

MSc. Advanced Wildlife Conservation in Practice

2019

The Effectiveness of Using drones and Convolutional Neural Networks to Monitor Populations of Nile Crocodile and Common Hippopotami

Word Count: 6463

by

Alexander Blazeby Woolcock

(Student no: 15011450)

This dissertation is submitted in partial fulfilment of the requirements of the University of the West of England, Bristol, for the degree of Master of Science.

This research was carried out in collaboration with Ezemvelo KZN Wildlife, South Africa.

December 2019

1. Abstract

Aquatic megafauna are difficult to observe and count due to the issues of detectability and inaccessibility. Traditional methods have proven to be an effective way of obtaining population estimates but have limitations. The recent proliferation of drone technology has offered an innovative way of surveying animal populations at high spatio-temporal resolutions. However, drones are hindered by a data analysis bottleneck that increases the time taken to process data. Convolutional Neural Networks (CNNs) are an emerging category of deep learning that can automate the data analysis process. This research compares traditional survey methods (transect and helicopter counts), alongside drones in order to detect and enumerate Nile crocodiles (*Crocodylus niloticus*) and common hippopotami (*Hippopotamus amphibious*) in two reserves in South Africa. This study also analyses the feasibility of using CNNs to more efficiently analyse drone imagery. The results reveal that the drone counts were more accurate than traditional methods; counting 21.1% more crocodiles. Where vegetation was open, counting hippos with a drone showed a similar pattern (counting 43.4% more than the traditional methods). When vegetation was dense the drone produced less accurate population estimates than traditional methods, due to the limited disturbance on the hippos. This disturbance was not sufficient to make the hippos move into open water, making them clearly visible. As a new analytical tool, the accuracy of CNNs was limited (85%) due to the reduced data set to train the CNN. However, with an expanded data set, object detection would be more accurate, making it more applicable for expedited and automated data analysis.

Keywords: convolutional neural networks, crocodile, drones, deep learning for ecology, hippopotamus, population monitoring.

2. Introduction

In many cases wildlife populations are undergoing significant decline due to increasing anthropogenic pressures, such as habitat degradation and intensive poaching (Lhoest *et al*., 2015). This has led to extinction rates that are a hundredfold greater than the background extinction rates (Hodgson *et al*. 2018). Accurate and efficient population estimates are crucial for ecological studies and wildlife management (Elsey & Trosclair, 2016; Hodgson *et al*., 2018; Gray *et al*., 2019). The effectiveness of management decision making is often dependent upon the accuracy and efficiency of the ecological data upon which decisions are based. This means that improved data collection methods may herald improved ecological outcomes (Hodgson *et al*., 2018).

Traditional methods such as drive counts and aerial counts have proven to be an effective way of obtaining population estimates of animals (Jachmann, 2002). These techniques have regularly been used in Tembe Elephant Park and Ndumo Game Reserve since 1971 (Calverley, 2014). However, problems with safety, cost, visibility, statistical integrity and logistics have hindered in the application of these methodologies (Jachmann, 2002; Jones, 2003). The recent proliferation of non-military applications of drones over the last decade has been of growing interest to the scientific community (Wich & Pin Koh, 2018). Drones offer a relatively risk-free and low-cost method of rapidly and systematically observing natural phenomena at high spatio-temporal resolution (Jiménez López & Mulero-Pázmány, 2019). The use of drones is an increasing facet of conservation ecology, with the potential to revolutionise the way in which animals and their habitats are monitored (Elsey & Trosclair, 2016; Longmore *et al*., 2017; Witczuk *et al*., 2018). This methodology is increasingly viewed as a supplement to, or a replacement of, traditional methods of surveying flora and fauna

(Christie *et al*., 2016). Within the conservation community, drones have mainly been used in three areas: (1) To determine the distribution and density of animal species (Vermeulen *et al*., 2013; Hodgson *et al*., 2018; Burke *et al*., 2019); (2) The mapping of land cover and land cover changes (Mulero-Pázmány *et al*., 2015; Fernández-Guisuraga *et al*., 2018; Moudrý *et al*., 2019; Beyer *et al*., 2019); (3) Anti-poaching activities (Mulero-Pázmány *et al*., 2014; Hambrecht *et al*., 2019; Bondi *et al*., 2019).

Traditional methods, such as counts from a vehicle or helicopter raise several constraints; high cost, safety, and inaccurate population estimates (Lhoest *et al*., 2015). Observer efficiency is an important source of bias leading to underestimates, especially where animals live in large congregations (Linchant *et al*., 2018). Therefore, these methods produce highly variable results. UAV's allow for very high-resolution data acquisition of both the spatial and temporal domain and may overcome the constraints of terrestrial and occupied aircraft counts (Seymour *et al*., 2017; Thapa *et al*., 2018).

The behaviour and habitat of aquatic mammals make collecting data on marine fauna challenging. This is due to the fact that they spend much of their time underwater, move rapidly over large areas and occupy remote habitats (Gray *et al*., 2019). As a result, aerial surveys are commonly used to collect population data on these inaccessible species (Chrétien, Théau & Ménard, 2016). UAV's are able to carry out similar tasks as helicopters, often more reliably, at lower costs and inducing less disturbance on the target species (Linchant *et al*., 2018; Ezat, Fritsch & Downs, 2018). Hodgson *et al*., (2016) found that UAV derived estimates of population size, resulted in smaller cumulative variances than other methods.

However, while UAVs collect detailed information rapidly, they do not overcome an existing data analysis bottleneck (Seymour *et al*., 2017; Nguyen *et al*., 2017). Specifically; manual counting animals in imagery is time consuming and inefficient. One possible method for overcoming this constraint is automated methods for detection, localisation and enumeration of target animals. Convolutional Neural Networks (CNNs) are a prominent category of deep learning classifier, inspired by the neural connections in the brain (Lee *et al*., 2015). These allow efficient discrimination of objects in noisy and complex environments (Gray *et al*., 2019) and may expedite the data analysis process (Brodrick, Davies & Asner, 2019). CNNs have already been implemented to advance ecological efforts such as; extinction risk (Wearn, Freeman & Jacoby, 2019), individual animal recognition (Devost, Lai & Berteaux, 2019), species recognition (Buehler *et al*., 2019) and enumeration of flora and fauna (Guirado *et al*., 2019; Mubin *et al*., 2019). CNNs are however prohibitively complex to implement, computationally intensive and may require more data than is practicable for most ecological studies (Gray *et al*., 2019).

This study will ascertain the possible variances of the three most common methodologies for enumerating Nile crocodiles (*Crocodylus niloticus*) and hippopotami (*Hippopotamus amphibious*). As different reserves use a plethora of survey methods, and emerging technologies offer new and innovative census techniques, it is essential to identify the differences between these methods, so as to maximise the efficiency of the identification of population changes over spatial and temporal periods.

In reviewing the relevant literature, it is apparent that this is the first research conducted which directly compares helicopter surveys, drive transect and drone surveys. As a new and innovative solution to the existing data analysis bottleneck, CNNs have had few real-world applications. There have been a small number of studies that have applied machine learning to analyse aerial imagery in Africa (Kellenberger, Marcos & Tuia, 2018; Eikelboom *et al.*, 2019). These studies have shown some success in machine learning applications; however, these methods have been semi-automatous and still require large amounts of data analysis. No research has been found to date, that use machine learning to fully autonomously analyse drone imagery in Africa. This research thereby acts as baseline data to inform and advice future studies.

3. Methodology

3.1 Study Sites

The research was conducted in Tembe Elephant Park (TEP) and Ndumo Game Reserve (NGR), which are both located in the KwaZulu-Natal Province of South Africa and managed by Ezemvelo KZN Wildlife (EKZNW) (Figure 1).

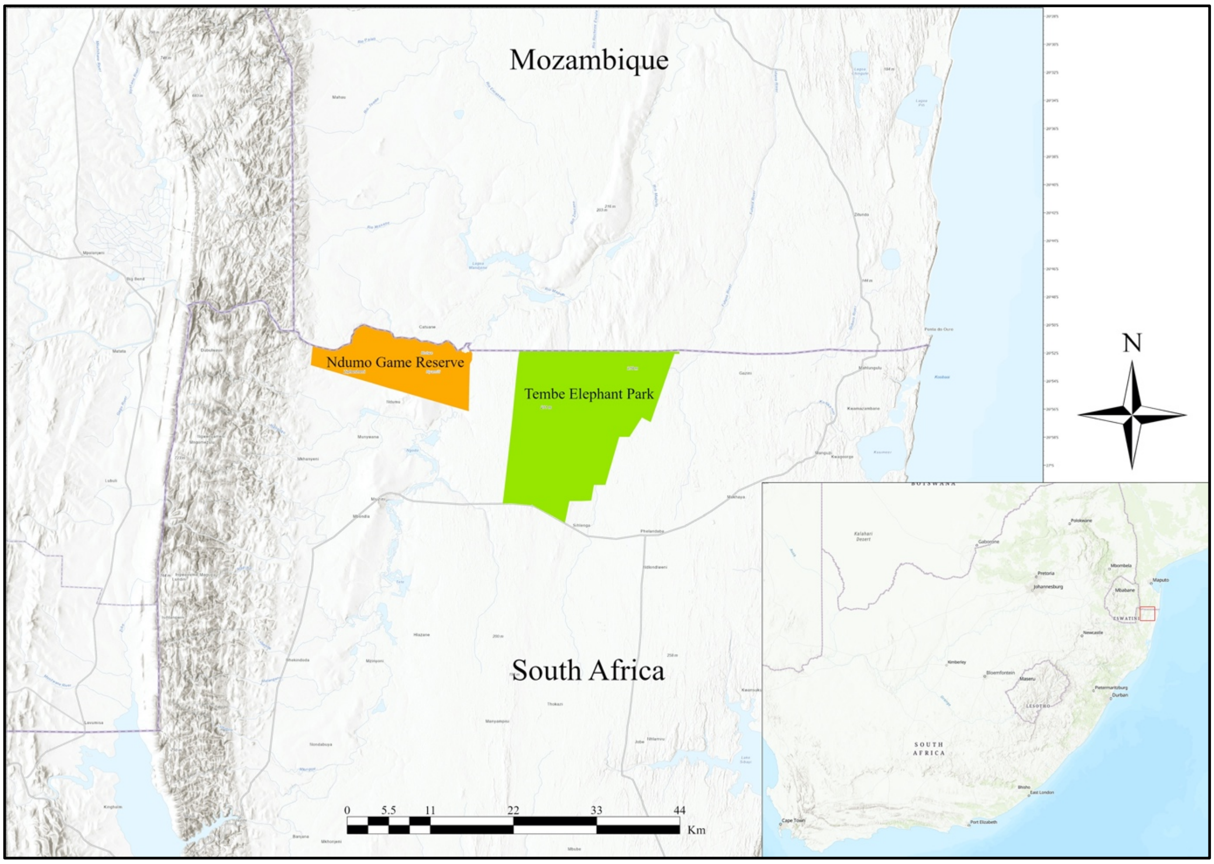


Figure 1 Map of Tembe Elephant Park and Ndumo Game Reserve

TEP (30,000-ha) was proclaimed in 1983, with the northern limit being the international boundary between South Africa and Mozambique (-26.902848, 32.561041). The topography is defined by high, linear, north-south oriented dune cordons (Matthews *et al*, 2001). The lowest-lying area of the park is the Muzi Swamp, which is an extensive system covering about 560-ha, stretching for 25km from north to south along the eastern edge of the reserve (Van Eeden, 2007; Gaugris & Van Rooyen, 2010; Van Rooyen *et al*., 2012). The Muzi Swamp consists of dense reedbeds, with stands of *Phragmites australis* as the most abundant species (Van Rooyen *et al*., 2012). The Muzi swamp represents the only natural source of permanent water within the reserve and thus possesses high densities of animals, including the only populations of hippos and crocodiles within the reserve. These water bodies will act as a focal point for this study. The Mbangweni corridor, found on TEP's eastern boundary, is a margin of approximately 5.5km which separates TEP and NGR.

NGR (10,000-ha) is one of the oldest game reserves in South Africa. Established in 1924 as a sanctuary for hippopotami, NGR is a reserve situated on the Mozambique Coastal Plain (-26.890264, 32.300418) and supports the third largest crocodile population in South Africa (Combrink, 2004; Calverley & Downs, 2014). The reserve is inherently patchy and characterised by a mosaic of permanent and ephemeral pans, streams and rivers with varying degrees of connectivity which fluctuate on a seasonal basis (Calverley & Downs, 2014). The reserve conserves two major river systems: the Usuthu River, which forms the Northern boundary of the reserve against Mozambique, while the Pongola River and floodplain run along the Eastern boundary of NGR (Calverley & Downs, 2015). The largest permanent pan in the Pongola floodplain and NGR is Lake Nyamithi (55ha), which possesses the highest densities of crocodiles and hippos in the reserve (Calverley, 2013; Calverley & Downs, 2014; Calverley & Downs, 2015). This lake will act as the focal point for research within this reserve. Lake Nyamithi is an irregular oval shape. During the study the lake had a 1.8km longitudinal axis running East to West and a maximum width of 531m. The Northern shore is characterised by a fringe forest of fever trees (*Acacia xanthophloea*) with gently sloping lawns of Couch grass (*Cynodon dactylon*). The Southern shore is rockier and possess a steeper slope (Calverley, 2013; Calverley & Downs, 2015).

3.2 General surveys

The field work was conducted from 30th July to 18th September 2019. All of the different surveys were carried out at the same time of day (11am - 2pm), and under similar climatic conditions (temp = 23°C ±3; clear sky; wind 5mph ±2). This reduced the chances of external variables influencing the counts. The different methods were carried out on separate days, so as to reduce the disturbance bias from the other counts.

3.3 Ground surveys

In NGR Lake Nyamithi is circumnavigated by a road which is on average 30m from the shoreline and offers a clear field of view across the lake. A vehicle was driven at low speeds (10km/hour) and the crocodiles and hippos were counted by two observers. The data was collected using ArcGIS Collector (ESRI, 2019); the species and GPS location was recorded. 10 X 42 binoculars were used to count individuals. A total of three repeats were carried out. In the case of TEP there is no road which runs along the Muzi Swamp where the water bodies were located. This made it necessary to carry out the ground counts on foot. As was the case with NGR, two observers counted crocodiles and hippos, while a third recorded the information using ArcGIS collector. Due to the logistics and risks associated with walking in a “Big 5” reserve, in an area that had the highest densities of animals, only one count was possible.

3.4 Helicopter counts

A Long Ranger helicopter (at a cost of £417 per hour) containing four people was flown on a predetermined north-south oriented axis over the reserve. The transects were situated 1km apart and arranged systematically to cover the entire reserve. The helicopter was flown at a height of 90m above the ground and at speeds of approximately 30-40 knots. Where hippos and crocodiles were spotted the helicopter deviated from the transect, and a total count was undertaken. The data was captured on a notebook computer using Cartalinx v1.2 (Hagan & Eastman, 1999) which, when connected to a GPS, allowed simultaneous collection of flight path information, animal number and species. Mapping the distribution of hippos and crocodiles was done by importing the data into ArcGIS Pro (ESRI, 2019a).

3.5 Drone counts

The drone surveys were carried out with a standard DJI Mavic Pro 2 quadcopter. This drone was chosen because of the 20 MP Hasselblad camera with a 1" sensor, it's maximum flight time of 31 min and adjustable aperture of f/2.8-f/11. This drone is also one of the more affordable commercial drones, retailing at £1349. A Polar Pro (Polar Pro, 2019) ND4 polarising filter was used to reduce the glare from the surface of the water bodies.

Pix4D mapper (Pix4D, 2019) is a software package used to transfer drone imagery into a georeferenced digital spatial model. For this study Pix4D mapper was used to create a predetermined flight plan over the focal area, and to autonomously fly the drone whilst capturing aerial imagery. The Pix4D Capture mobile app was used to determine the flight settings and to design a flight plan. The drone was flown in parallel lines at an altitude of 100m above ground level, at a speed of 12 km/h. The drone camera was set to 90° and the white balance set to "sunny". The ground sample distance (GSD), being the distance between adjacent pixel centres on the ground, was 0.02m. In the case of Lake Nyamithi the number of images taken was around 1200 and Muzi swamp 550 images. Due to the drones limited battery life, it was necessary to map the target area in separate instalments. This involved landing the drone during the flight plan and exchanging the battery before continuing with the flight plan. There were no obvious negative interactions with the crocodiles and hippos; further animal ethical considerations can be found in Mulero-Pázmány *et al*. (2017) and a more detailed protocol for minimising drone disturbance to wildlife can be found in Hodgson & Koh (2016).

3.6 Analysis

The aerial drone images were mosaicked using the Pix4D mapper software in order to obtain a single multiband RGB image of the lake/swamp, known as an ortho-mosaic. The software identifies points in the overlaying images to generate a 3D point cloud model of the landscape. An image overlap of 80% front and 70% side, is needed to generate an ortho-mosaic map of sufficient quality (Ezat, Fritsch & Downs, 2018). The image stitching process was made up of three stages: initial processing, point cloud and densification, and finally, DSM and ortho-mosaic generation, which are carried out automatically. The ortho-mosaics were corrected for inconsistencies, for example animal movements between adjacent images. A total of seven ortho-mosaics were created: four of Nyamithi and three of the Muzi Swamp.

The process of developing CNNs was carried out using an ESRI (2019b) workflow as a template and the algorithm and methodology were adapted to better suit crocodiles. In the first instance the crocodiles were manually and systematically counted in the ortho-mosaic. Then using ArcGIS Pro's training samples tool, individual crocodiles were digitally selected and exported as a Tiff file, which ensured that the training data retained RGB bands. Each training sample comprised a 448 x 448 x 3 (RGB) tensor and accompanying label classifying the tensor as positive or negative (1 or 0, respectively). To further augment the training set, the images underwent random horizontal flipping. The training samples had crocodiles centred randomly within the image window, which was done to simulate the random positioning of crocodiles in the feature map.

In order to train the object detector, a Single Shot MultiBox Detector (SSD) (Liu *et al*., 2016) was implemented using Jupyter Lab (Anaconda, 2019). The SSD model is a feed forward convolutional network that produces a fixed number of bounding boxes and scores for the presence of object class instances in those boxes. The scoring system is based on the presence of key features, such as edges, curves or colour gradient. This feature creates a deeper layer in the bounding boxes and aggregates the features from the previous layers, combining them into groups of curves and edges that may indicate a crocodile tail or head. This step is followed by a non-maximum suppression step to produce the final detections. The early network layers are based on standard SGG-16 architecture developed by Simonyan & Zisserman (2014), which has shown high quality image classification and superiority over other networks (Russakovsky *et al*., 2015; Canziani, Paszke & Culurciello, 2016; Liu *et al*., 2016). This architecture will act as a base for the CNN, and auxiliary structures will be used in order to produce detections highlighting a number of key features. A convolutional feature layer was added to the end of the base network to allow for predictions at multiple scales. Every added feature layer can produce a fixed set of detection predictions using convolutional fillers. For a feature layer the basic element for predicting parameters for a potential detection is a kernel, which produces either a score for a category or shape offset, relative to the default training scheme. At each of the locations where the kernel is applied, it employs a binary normalised exponential function, which the final fully connected layer and its learnt combinations of features; are valued between 0 or 1, with high values signalling high confidence of the crocodile in the image window. Finally, the ortho-mosaic is divided into a set of default bounding boxes. The default boxes tile the ortho-mosaic in a convolutional manner, so the position of each box relative to its corresponding cell is fixed. At each cell, the offsets are predicted relative to the box shape and colour gradient. At each box, at a given spatial location, the value and the offset relative to the original box shape is recorded. This is extrapolated across the entire ortho-mosaic, yielding numbers of observed crocodiles in the survey area.

A simple way to detect objects in an image is to divide the image into a grid and the SSD is responsible for identification within each cell. The SSD also adds convolutional layers to the architecture so as to ensure that the spatial size of the final layer is the same size of the grid. This allows the SSD to be fast and efficient, while taking advantage of the grids within each image window. For this project a grid of 4x4 was implemented, which divided each training sample into 16 sections, making it effective for crocodiles.

Once the appropriate model has been constructed, it needs to be trained over several epochs. The accuracy and loss were monitored during the CNN training. The training process was stopped when the accuracy and loss didn’t improve over consecutive epochs (Figure 2). The final measure of the model’s accuracy was 0.66357.

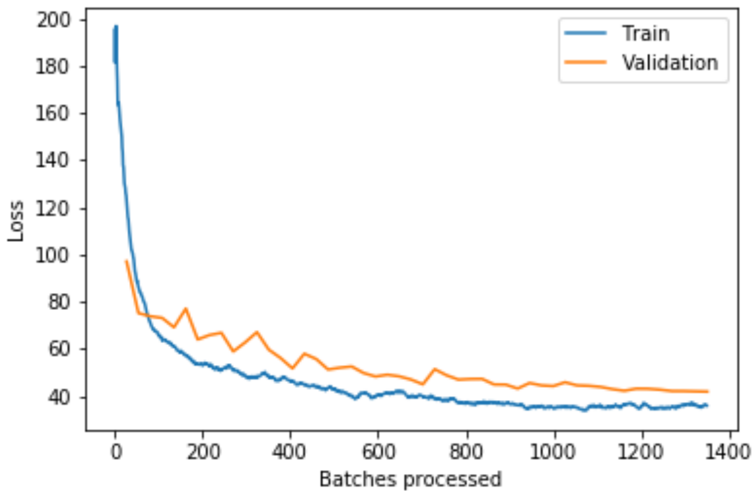


Figure 2 Loss graph outlining the training of the CNN

In this case 50 epochs were used to train the model. During the training, the model compares the default boxes to the corresponding ground truth (training data). Initially the model began by matching each ground truth to the default box, with the best overlap higher than a threshold of 0.5. This allows the network to predict high scores for multiple overlapping default boxes rather than the model picking only one that has the maximum overlap (Liu *et al*., 2016). At each epoch, the loss (error rate) and validation set (indication of model learning) for the training data was reported. The model is continuously trained until the validation loss begins to increase. This is an indication that the model is beginning to overfit to the training data. The optimum model hyper-parameters were selected to minimise the loss on the validation set. For the CNN, the sizes of convolutional kernel and batch are 4 and 16 respectively. The model was then imported into ArcGIS and run over a trimmed portion of the ortho-mosaic, which included Lake Nyamithi and 4m of the bank extending from the water's edge. This decreased the running time of the model and also decreased the objects (e.g. trees), which could give false positives. Refer to Appendix A which shows a flow chart of the CNN analysis, and Appendix B which displays an overview of the CNN architecture.

The model was validated by reviewing the detections manually and systematically using ArcGIS Pro and comparing them to the manual counts. The instances of duplicate detections, false positives or any false negatives encountered were recorded. The evaluation metric used in this experiment is the margin of error (a) for the number of crocodiles detected after applying the CNN, where D is the number of detected crocodiles in the image after application of the CNN and N is the actual number of crocodiles in the image.

$$margin\ of\ error\ = \frac{(D-N)}{N}\ X\ 100 \tag{a}$$

As was carried out by Li *et al*. (2019) the precision, recall and F1-score are calculated via equations (b)-(d) respectively, where True Positive (TP) indicates the number of correctly identified crocodiles, the False Positive (FP) indicates the number of other objects that are incorrectly identified as crocodiles, and the False Negative (FN) indicates the number of crocodiles not detected.

$$Precision = \frac{TP}{TP+FP} \tag{b}$$

$$Recall = \frac{TP}{TP+FN} \tag{c}$$

$$F1 - Score = \frac{2 \times\ Precision\ \times\ Recall}{(Precision\ +\ Recall)} \tag{d}$$

4. Results

4.1 Tembe Elephant Park

Results of ground/ helicopter/ drone comparative surveys for TEP are summarised in Table 1. The ground counts carried out on foot in TEP involved walking a transect of 7.2km along the southern side of the Muzi Swamp. No crocodiles were counted in this transect. A total of 16 hippos were counted (Figure 3A). These were in two separate pods: one of which was basking in the middle of a pan (N=13), the other was seen in a northern pan surrounded by reed beds (N=6). The helicopter counts recorded a total of 29 hippos, also in two separate pods. Both pods were seen in the same location as was observed from the ground counts (Southern pod N=18, northern pod N=11). A total of 4 crocodiles were counted at two separate locations (Figure 3B). The ortho-mosaic obtained from the aerial imagery, shows the water bodies in the Muzi Swamp that have not yet dried up (Figure 3C). A total area of 82.7 hectares were mapped in three separate battery instalments, on three separate occasions (Table 1) and the hippos and crocodiles were counted manually. The mean number of hippos counted was 16 (SD=0.66). Only hippos in one pod were seen from the drone imagery, and these were basking in the middle of a pan (figure 3F). In the three repetitions, only one pod was located, which was in the same pan on each repetition. The mean number of crocodiles counted was 8 (SD=1.41). The ground and helicopter surveys observed a separate pod of hippos in a northern pan which were not detected in the drone survey (Figure 3A, B&C). These hippos were observed in both instances emerging from reed beds into open water when the counters were in close proximity (<30m) or flew near the general area.

Table 1 Number of crocodiles and hippopotami counted via the different methodologies in Tembe Elephant Park

Tembe		
Walking Counts		
Date	**Crocodiles Counted**	**Hippos Counted**
18/09/2019	0	19
Helicopter Counts		
Date	**Crocodiles Counted**	**Hippos Counted**
29/07/2019	4	29
Drone Counts		
Date	**Crocodiles Counted**	**Hippos Counted**
04/08/2019	7	17
11/08/2019	10	16
02/09/2019	7	15
Mean	8	16
Standard Deviation (SD)	1.414	0.816

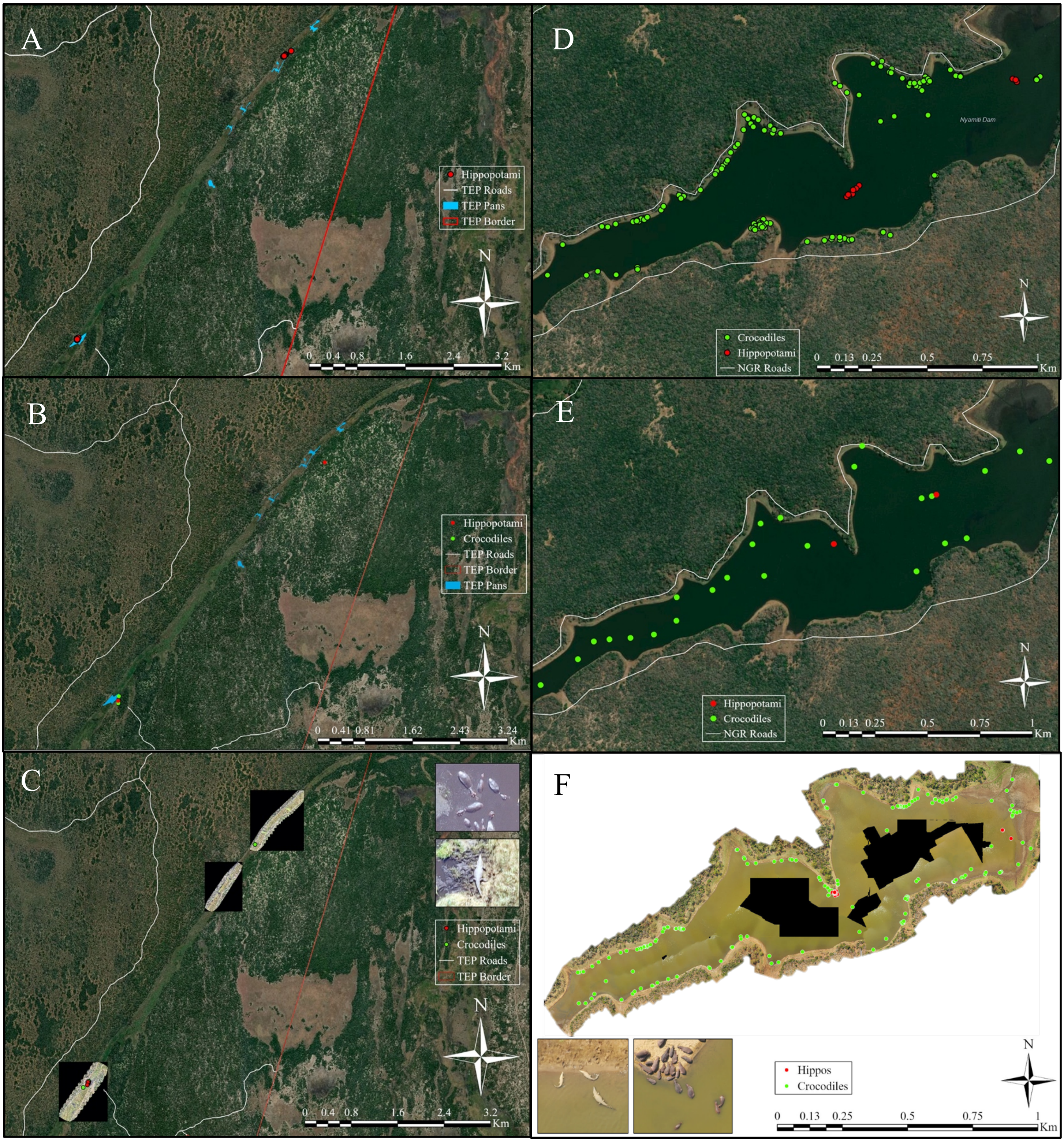


Figure 3 Comparative map of methodologies conducted in TEP and NGR. A) Drive counts TEP; B) Helicopter counts TEP; C) Drone counts TEP on 30/8/2019; D) Walking counts NGR; E) Helicopter counts NGR; F) Drone counts NGR on 4/8/2019. The drone counts give an image of both crocodiles and hippos, showing the resolution whilst flying at a height 100m

4.2 Ndumo Game Reserve

Results of ground/ helicopter/ drone comparative surveys for TEP are summarised in Table 2. The mean number of hippos counted was 15.67 (SD=4.11). These were in two separate pods that were either basking on the edge of the lake or wallowing in the water. The mean number of crocodiles counted was 126 (SD= 9.9). These crocodiles were relatively evenly distributed around the edge of the lake, although a higher density was found on the southern edge of the lake (Figure 3D). The helicopter counts counted 15 hippos in the same two pods as did the drive counts. A total of 152 crocodiles were counted (Figure 3E). However, the GPS location of the individual crocodile was taken as the location of the helicopter. As the helicopter was not always directly above the crocodile when the GPS point was taken, it was difficult to plot an accurate location from the helicopter data.

A total of four drone censuses were conducted over lake Nyamithi. This involved mapping a total area of 57 hectares, which was carried out using four battery instalments. Via manual counting, the mean number of hippos counted was 26.5 (SD=2.6). As was the case with the ground and helicopter surveys, the hippos were in two distinctive pods: the first basking on the southern bank of the lake, and the second; wallowing in the water (Figure 3F). The mean number of crocodiles was 192.8 (SD=2.3). The drone counts again showed similar spatial data as the ground counts. The crocodiles were evenly distributed with a higher density on the southern shore.

Table 2 Number of crocodiles and hippopotami counted via the different methodologies in Ndumo Game Reserve

Ndumo Game Reserve		
Drive Counts		
Date	**Crocodiles Counted**	**Hippos Counted**
30/07/2019	119	15
15/08/2019	140	21
10/09/2019	119	11
Mean	126	15.667
Standard Deviation (SD)	9.899	4.12
Helicopter Counts		
Date	**Crocodiles Counted**	**Hippos Counted**
30/07/2019	152	15
Drone Counts		
Date	**Crocodiles Counted**	**Hippos Counted**
22/08/2019	189	23
25/08/2019	195	25
30/08/2019	193	29
13/09/2019	194	29
Mean	192.75	26.5
Standard deviation (SD)	2.278	2.598

4.3 Convolutional Neural Network Detections

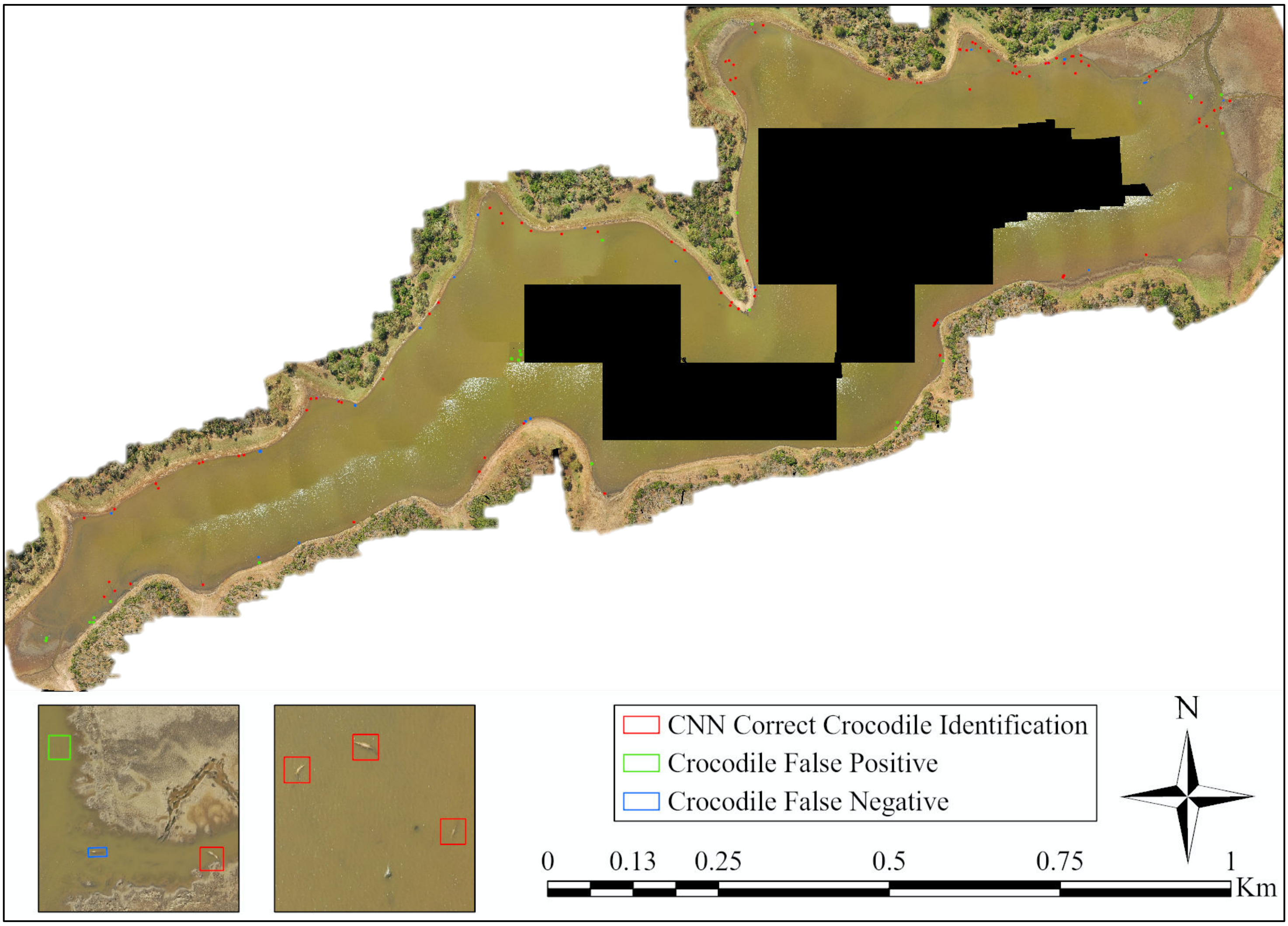


Figure 4 CNN detections of Crocodiles in NGR 13/09/2019

Figure 4 shows the CNN applied to an ortho-mosaic conducted in Nyamithi. The overall F1-Score (accuracy) of the model was 84% (Table 3). Analysis of the results obtained by the CNN found that the outputs (precision, recall and F1-score) were relatively consistent and ranged from 84-85.6%. False positives were generally due to either, sections of wood or occasionally birds. Manual counts were seen to have better detection capability, on average counting +14.43% more crocodiles than the CNN.

Table 3 Statistical analysis of the Convolutional Neural Network, applied to the 13/09/2019 drone ortho-mosaic. Analysis outlines the different measures of accuracy of the model.

CNN Detection from ortho-mosaic								
Date	**Actual Number of crocodiles**	**True Positives**	**False Positives**	**False Negatives**	**Margin of error (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
13/09/2019	194	166	31	28	-14.43	84.2639594	85.5670103	84.9104859

5. Discussion

This study examines the changes in relative abundance between drone, helicopter and ground surveys. Previous research has indicated that UAV derived estimates of population size result in smaller cumulative variances than other methods (Hodgeson *et al*., 2016). This suggests that drone estimates may be more accurate than the traditional methods, such as ground and helicopter surveys. However, no research directly compares these three methodologies. The research conducted, is essential to ensure that reserves have consistent population estimates. Without this it would be impossible to derive patterns within populations. This research also addresses the issue of a data analysis bottleneck that arises alongside new technological advances. As an emerging technology CNNs have had few real-life applications. There have been limited previous studies that have applied machine learning to analyse aerial imagery in Africa (Kellenberger, Marcos & Tuia, 2018; Eikelboom *et al*., 2019). These studies have shown some success in machine learning applications; however, these methods have been semi-automatous and still require large amounts of data analysis. In reviewing the relevant literature, it is apparent that no research has been found to date, that uses machine learning to autonomously analyse, drone imagery in Africa. This research serves to demonstrate the effectiveness of drone technology in conjunction with deep learning, as a method of conducting animal censuses to a higher degree of accuracy than has hitherto been possible.

Given that all three of the survey methods were not carried out simultaneously, it is possible that there could be some temporal variability in the densities found due to movements or behavioural differences. As NGR and TEP are open systems, limiting these variabilities brings forth its own set of challenges. Calverley and Downs (2015) observed an outflux of crocodiles leaving the NGR into the Rio Maputu, however, this took place from early November onwards. They concluded that large-scale seasonal movement or migration in reptiles is uncommon and therefore variability between the different methods would be negligible at the scale of Nyamithi and the Muzi Swamp. Scotcher *et al*. (1978) found that in NGR large scale movement of hippos was attributed to insufficient grazing, which was due to high and extended periods of flooding. In these instances, hippos moved to the eastern hills for grazing. As this study was conducted during the dry season, there were no flooding events and extensive grazing was available surrounding the Nyamithi and Muzi Swamps. We can assume that the movement of hippos was minimal and had little impact on the comparison of the methods. Further research will be useful to confirm these assumptions.

As the counts performed were minimum counts, and the movement of animals had little impact, it can be assumed that the highest number of sightings is the most accurate estimate of populations size within a particular body of water.

Crocodiles

The ability for the UAV to complete unmanned flights on this scale whilst surveying wildlife is incredibly promising and highlights the logistical potential for this technique in undertaking future surveys. The results show that when counting crocodiles, the UAV yielded significantly higher population estimates. In NGR the drone counted 21.6% more than ground surveys and 21.1% more than helicopter surveys. In TEP results showed a similar pattern (UAV was 100% more accurate than ground and 50% more accurate than helicopter counts). In the case of counting crocodiles with drones under these conditions it is clear that drones are the most effective method of obtaining population estimates for crocodiles. Drone counts produce the highest figure, and with little variance in their repeated surveys. Helicopter and ground surveys are dependent upon the ability of the observers. Highly trained observers will produce more accurate counts (Linchant *et al*., 2018). This increases the degree of bias to which these methodologies may be prone. Drones are less influenced by this bias and thereby have advantages over the other two methods. The data collected by the drone surveys, albeit more time consuming to analyse, even with the CNNs; has many other applications. It is therefore, much more time efficient than the other two methods.

Hippos

The research indicated that when counting hippos in TEP, the helicopter counts ascertained more accurate population estimates (helicopter counted 34.5% more hippos than walking and 44.8% more than drone) However, in NGR the drone counted 43.4% more hippos than helicopter and 41.1% more than driving counts.

In the case of NGR the encounter rate of hippos, and clearly their detectability recorded with the drone was lower than that of the helicopter surveys, resulting low density estimates. Many studies have highlighted the fact that manned aircraft can disturb wildlife (Muelero-Pázmány *et al*., 2017), resulting in altered and sometimes illusive behaviour that can make the study and enumeration of animals problematic (IV, Pearlstine & Percival, 2006; Christie *et al*., 2016; Chrétein, Théau & Ménard, 2016; Gentle *et al*., 2018). However, in this case, the opposite seems to be the case. During the survey, the helicopter caused enough disturbance to make the hippos leave the reed beds and move into open water for safety where they were easily visible. This is also the case, albeit to a lesser degree, with the walking counts. When the observers were within 30m of the hippos within the reed beds, they exhibited the same behaviour as the helicopter counts (running from reeds into open water), leading to a lower population estimate. A similar situation has been identified by Gentle *et al*. (2018). They recorded implausibly low densities of macropods from drone surveys, compared to helicopter surveys. This was attributed to a lack of "flushing", coupled with a suboptimal camera, which made detection challenging and thus leading-to inaccurate population estimates. It is difficult to quantify the disturbance a particular method is causing an animal (Pomeroy, O'connor & Davies, 2015; Tablado & Jenni, 2017), as often the animal does not exhibit any symptoms (Ditmer *et al*., 2015). However, Muelero-Pázmány *et al*. (2017) has found that noise is the most prominent cause of disturbance. This theory is supported by this research.

During the three repeats of the drone surveys in TEP only one pod of hippos was observed. The northern pod was not detected. We can therefor assume that the drone flying at 100m did not pose a high enough level of disturbance to flush the hippos from the reed beds, where they could be observed and counted. This is contrary to many studies that have used drones, which have suggested that their low acoustic signature, is advantageous for reducing disturbance and evasive behaviour, in turn making population estimates more accurate (IV, Pearlstine & Percival, 2006; Christie *et al*., 2016; Chrétein, Théau & Ménard, 2016; Gentle *et al*., 2018). In the case of hippo counting in TEP, it is evident that the methodology producing the highest levels of disturbance results in the more accurate population estimates. This raises many questions surrounding the ethics of data collection. Should a method be chosen if it elicits-a larger wildlife reaction and therefore a more accurate population estimate. An alternative could be to improve the drone sensor, for example by incorporating a thermal imaging camera which would drastically improve the detection of animals through vegetation without causing significant disturbance to the animal thereby inducing a behavioural change. This technique has already been successfully trialled on a number of species (Longmore *et al*., 2017; Witczuk *et al*., 2018; Kays *et al.*, 2019), and could dramatically improve the detection of hippos located in dense vegetation, whilst posing little or no disturbance.

In the case of Nyamithi, the drone counted significantly more hippos than any other method. This can be explained by the type of vegetation in the study area; as the Muzi Swamp is made up of dense reed beds where the detection of hippos can be an issue. Nyamithi on the other hand, is made up of open water with the closest vegetation being <4m from the edge of the water. This makes water dependant animals far easier to observe, allowing more accurate estimate of population size. Hodgson *et al*., (2016) found that UAV derived estimates of population size resulted in smaller cumulative variances than other methods. This was also found with this research at both TEP and NGR. However, due to the inability to perform repetitions on helicopter counts, it was not possible to ascertain levels of variance between repetitions. This made only the variance between ground and drone counts comparable. In NGR the ground counts showed an increase in SD of 336% when counting crocodiles, and an increase of 59% when counting hippos. As the drive counts show a higher variance in population estimations and the movement of animals within the population having little effect, it would suggest that highly varied results were less accurate and less reliable than drone counts.

Within the context of this study, drones, with the exception of one instance, were the optimum method for enumerating hippos and crocodiles. However, they were limited when it came to surveying water that was surrounded by thick and obscuring vegetation. Drones did not pose enough disturbance to the hippos to force them into open water. This is both a positive and a negative. Drones offer a method of surveying these species without imposing significant stress and disturbance, which could act as a significant bias on the results and may lead to inaccurate findings. On the other hand, drones can underestimate animal populations due to the limited disturbance factor, making detections more challenging.

Previous work has shown CNNs to be an accurate and time efficient analysis technique to enumerate animals in drone imagery (Barry, 2018; Kellenberger, Marcos & Tuia, 2018; Rivas *et al*., 2018; Corcoran *et al*., 2019). Grey *et al*., (2019) found that although promising, CNNs can be inherently complex to implement. Moreover, they are computationally intensive and complex, and may require more data than is practicable for most ecological studies. Similar conclusions were drawn from this study. To develop effective CNNs, large volumes of training data are required to learn suitable parameter values; sometimes over 1000 images (Moya *et al*. 2015; Sacchi *et al*. 2016; Yousif *et al*. 2019; Willi *et al*. 2019). As the drone imagery of hippos was limited, it was not possible to build up a significant enough training scheme to implement the identification of hippos using CNNs. This was more feasible with crocodiles due to the larger number of images. However, at the time of this study it was not possible to merge different training sets, created over different ortho-mosaics. Therefore, the size of the training scheme was limited by the number of crocodiles in a single ortho-mosaic (N=180). A comprehensive training and validation data set was critical for developing accurate CNNs (Guirado *et al*., 2019; Li *et al*., 2019; Brodrick, Davies & Asner, 2019).

The performance of the CNN was limited (85% accuracy), however, when the restricted volume of training data is taken into account; these low levels of accuracy are not surprising. With a larger collection of training samples (N=450), to compete with Grey *et al*. (2019) – 467 training images; Mubin et al. (2019) - 260 training images; Cheang, Cheang & Tay (2017) – 300 training images; Eikelboom *et al*. (2019) - 516 training images; Chew *et al*. (2018) – 1500 training images, you would expect to see similar levels of accuracies (97%). The more inherently complex the data, the larger the training scheme needs to be (Wearn, Freeman & Jacoby, 2019). As crocodile images vary considerably depending on their environment (basking on bank, partially submerged, partially obscured etc), they require a larger data set to account for this variation.

6. Conclusion

This research compared a number of survey and analysis methods to enumerate crocodiles and hippos in TEP and NGR. The results of this study should be considered in the context of the spatial location and the survey methods chosen. This study expands upon the limited research available and suggests the optimum census methods for crocodiles and hippos, within the specific habitat types. Drones were seen to be the most effective method when it came to enumerating crocodiles and hippos in an open habitat type. Helicopters were seen to be most effective when flushing was necessary to detect the individuals in dense vegetation. The CNNs were successfully applied to the drone imagery and show great potential in identifying crocodiles. They do, however, have limitations when only a small batch of training data is used; reducing their effectiveness for studies such as this. This is the first research of its kind that compares the two most common methods alongside the new and innovative method of drone surveys and CNNs. This will allow managers and ecologists to choose the most effective methods in the context of their reserve or to apply correctional factors where a range of methods have been incorporated.

Varying methodologies need to be incorporated at the reserve level, rather than a species level. Whilst drones may be the most effective method in this instance, few studies have looked at their feasibility for studies over a larger spacial scale (Linchant *et al*., 2015; Gentle *et al*., 2018). At present technology appears to be the limiting factor. This issue is likely to be less critical in the future as drone batteries and sensors improve. Research into drone applications over a wider area would be essential to effectively implement the use of drones at the reserve level.

7. Acknowledgements

I would like to thank my supervisor, Alison Cotton for the help and advice given throughout my project. I am also grateful for the help, assistance and logistical support from Ezemvelo KZN Wildlife, with particular reference to Catheriné Hanekom (district ecologist) and Len Gunter (Manager of Tembe Elephant Park). I would like to thank E. Woolcock; M. Sibiya; Corporal Massimara and M. Siengya for their help and fieldwork assistance. I am appreciative to the University of the West of England who have partially funded this research. Finally, I am grateful for the constructive comments of the reviewers.

7. Conflict of Interest

The author declares no conflicts of interest

8. References

Anaconda (2019) *Jupyter Lab*. Available from: https://www.anaconda.com.

Barry, J. (2018) Applications of Deep Learning in Ornithology. *Biodiversity Information Science and Standards*, *2*, p.e27251.

Beyer, F., Jurasinski, G., Couwenberg, J. and Grenzdörffer, G. (2019) Multisensor data to derive peatland vegetation communities using a fixed-wing unmanned aerial vehicle. *International Journal of Remote Sensing*. pp.1-23.

Bolger, D.T., Morrison, T.A., Vance, B., Lee, D. and Farid, H. (2012) A computer-assisted system for photographic mark–recapture analysis. *Methods in Ecology and Evolution*. *3*(5), pp.813-822.

Bondi, E., Fang, F., Hamilton, M., Kar, D., Dmello, D., Noronha, V., Choi, J., Hannaford, R., Iyer, A., Joppa, L. and Tambe, M. (2019) Automatic Detection of Poachers and Wildlife with UAVs. *Artificial Intelligence and Conservation*. p.77.

Brodrick, P.G., Davies, A.B. and Asner, G.P. (2019) Uncovering ecological patterns with convolutional neural networks. *Trends in ecology & evolution*.

Buehler, P., Carroll, B., Bhatia, A., Gupta, V. and Lee, D.E. (2019) An automated program to find animals and crop photographs for individual recognition. *Ecological Informatics*. *50*, pp.191-196.

Burke, C., Rashman, M., Wich, S., Symons, A., Theron, C. and Longmore, S. (2019) Optimizing observing strategies for monitoring animals using drone-mounted thermal infrared cameras. *International Journal of Remote Sensing*. *40*(2), pp.439-467.

Calverley, P. (2013) *The conservation ecology of the Nile crocodile (Crocodylus niloticus) at Ndumo Game Reserve in north eastern KwaZulu-Natal and the Rio Maputo Floodplain in south eastern Mozambique* (Doctoral dissertation, University of KwaZulu-Natal, Pietermaritzburg).

Calverley, P. (2014) *The conservation ecology of the Nile crocodile (Crocodylus niloticus) at Ndumo Game Reserve in north eastern KwaZulu-Natal and the Rio Maputo Floodplain in south eastern Mozambique* (Doctoral dissertation).

Calverley, P.M. and Downs, C.T. (2014) Habitat use by Nile crocodiles in Ndumo Game Reserve, South Africa: a naturally patchy environment. *Herpetologica*, *70*(4), pp.426-438.

Calverley, P.M. and Downs, C.T. (2015) Movement and Home Range of Nile Crocodiles in Ndumo Game Reserve, South Africa. *Koedoe*, *57*(1), pp.1-13.

Canziani, A., Paszke, A. and Culurciello, E. (2016) An analysis of deep neural network models for practical applications. *arXiv preprint arXiv:1605.07678*.

Cheang, E.K., Cheang, T.K. and Tay, Y.H. (2017) Using convolutional neural networks to count palm trees in satellite images. *arXiv preprint arXiv:1701.06462*.

Chew, R.F., Amer, S., Jones, K., Unangst, J., Cajka, J., Allpress, J. and Bruhn, M. (2018) Residential scene classification for gridded population sampling in developing countries using deep convolutional neural networks on satellite imagery. *International journal of health geographics*, *17*(1), p.12.

Chrétien, L.P., Théau, J. and Ménard, P. (2016) Visible and thermal infrared remote sensing for the detection of white-tailed deer using an unmanned aerial system. *Wildlife Society Bulletin*. *40*(1), pp.181-191.

Christie, K.S., Gilbert, S.L., Brown, C.L., Hatfield, M. and Hanson, L. (2016) Unmanned aircraft systems in wildlife research: current and future applications of a transformative technology. *Frontiers in Ecology and the Environment*, *14*(5), pp.241-251.

Combrink, A.S (2004) Population Status of Crocodylus niloticus (Nile Crocodile) at Lake Sibaya, Republic of South Africa. M.Sc. thesis, University of KwaZulu- Natal, South Africa.

Corcoran, E., Denman, S., Hanger, J., Wilson, B. and Hamilton, G. (2019) Automated detection of koalas using low-level aerial surveillance and machine learning. *Scientific reports*. *9*(1), p.3208.

Devost, E., Lai, S., Casajus, N. and Berteaux, D. (2019) FoxMask: a new automated tool for animal detection in camera trap images. *BioRxiv*, p.640037.

Ditmer, M.A., Vincent, J.B., Werden, L.K., Tanner, J.C., Laske, T.G., Iaizzo, P.A., Garshelis, D.L. and Fieberg, J.R. (2015) Bears show a physiological but limited behavioral response to unmanned aerial vehicles. *Current Biology*, *25*(17), pp.2278-2283.

Eikelboom, J.A., Wind, J., van de Ven, E., Kenana, L.M., Schroder, B., de Knegt, H.J., van Langevelde, F. and Prins, H.H. (2019) Improving the precision and accuracy of animal population estimates with aerial image object detection. *Methods in Ecology and Evolution*.

Elsey, R.M. and Trosclair, P.L. (2016) The use of an unmanned aerial vehicle to locate alligator nests. *Southeastern Naturalist*. *15*(1), pp.76-83.

ESRI (2018-2019) *ArcGIS Collector version 19.2.0* [mobile app]. Available from: https://www.esri.com/en-us/arcgis/products/collector-for-arcgis/resources. [Accessed01/03/2019].

ESRI (2019a) *ArcGIS Pro version 2.4.0.* [Programme]. Available from: https://www.esri.com/en-us/arcgis/products/arcgis-pro/resources. [Accessed 03/05/2019].

ESRI (2019b) *Use Deep Learning to Assess Palm Tree Health*. Available from: https://learn.arcgis.com/en/projects/use-deep-learning-to-assess-palm-tree-health/ [Accessed 04/04/2019].

Ezat, M.A., Fritsch, C.J. and Downs, C.T. (2018) Use of an unmanned aerial vehicle (drone) to survey Nile crocodile populations: a case study at Lake Nyamithi, Ndumo game reserve, South Africa. *Biological conservation*. *223*, pp.76-81.

Fernández-Guisuraga, J., Sanz-Ablanedo, E., Suárez-Seoane, S. and Calvo, L. (2018) Using unmanned aerial vehicles in postfire vegetation survey campaigns through large and heterogeneous areas: Opportunities and challenges. *Sensors*. *18*(2), p.586.

Gaugris, J.Y. and Van Rooyen, M.W. (2010) Effects of water dependence on the utilization pattern of woody vegetation by elephants in the Tembe Elephant Park, Maputaland, South Africa. *African journal of ecology*, *48*(1), pp.126-134.

Gentle, M., Finch, N., Speed, J. and Pople, A., 2018. A comparison of unmanned aerial vehicles (drones) and manned helicopters for monitoring macropod populations. *Wildlife research*, *45*(7), pp.586-594.

Gray, P.C., Fleishman, A.B., Klein, D.J., McKown, M.W., Bézy, V.S., Lohmann, K.J. and Johnston, D.W. (2019) A convolutional neural network for detecting sea turtles in drone imagery. *Methods in Ecology and Evolution*. *10*(3), pp.345-355.

Guirado, E., Tabik, S., Rivas, M.L., Alcaraz-Segura, D. and Herrera, F. (2019) Whale counting in satellite and aerial images with deep learning. *Scientific reports*, *9*(1), pp.1-12.

Hagan, J.E., Eastman, J.R. (1999) Cartalinx: The Spatial Data Builder, version 1.2. Clark University, Worcester, MA, USA.

Hambrecht, L., Brown, R.P., Piel, A.K. and Wich, S.A. (2019) Detecting ‘poachers’ with drones: Factors influencing the probability of detection with TIR and RGB imaging in miombo woodlands, Tanzania. *Biological Conservation*. *233*, pp.109-117.

Hodgson, J.C., Baylis, S.M., Mott, R., Herrod, A. and Clarke, R.H. (2016) Precision wildlife monitoring using unmanned aerial vehicles. *Scientific reports*. *6*, p.22574.

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

Hodgson, J.C., Mott, R., Baylis, S.M., Pham, T.T., Wotherspoon, S., Kilpatrick, A.D., Raja Segaran, R., Reid, I., Terauds, A. and Koh, L.P. (2018) Drones count wildlife more accurately and precisely than humans. *Methods in Ecology and Evolution*. *9*(5), pp.1160-1167.

IV, G.P.J., Pearlstine, L.G. and Percival, H.F., 2006. An assessment of small unmanned aerial vehicles for wildlife research. *Wildlife Society Bulletin*, *34*(3), pp.750-758.

Jachmann, H. (2002) Comparison of aerial counts with ground counts for large African herbivores. *Journal of Applied Ecology*. *39*(5), pp.841-852.

Jackson, R.M., Roe, J.D., Wangchuk, R. and Hunter, D.O. (2006) Estimating snow leopard population abundance using photography and capture-recapture techniques. *Wildlife Society Bulletin*. *34*(3), pp.772-781.

Jiménez López, J. and Mulero-Pázmány, M. (2019) Drones for conservation in protected areas: Present and future. *Drones*. *3*(1), p.10.

Jones, G.P. (2003) *The feasibility of using small unmanned aerial vehicles for wildlife research* (Doctoral dissertation, University of Florida).

Kays, R., Sheppard, J., Mclean, K., Welch, C., Paunescu, C., Wang, V., Kravit, G. and Crofoot, M., (2019) Hot monkey, cold reality: surveying rainforest canopy mammals using drone-mounted thermal infrared sensors. *International journal of remote sensing*, *40*(2), pp.407-419.

Kellenberger, B., Marcos, D. and Tuia, D. (2018) Detecting mammals in UAV images: Best practices to address a substantially imbalanced dataset with deep learning. *Remote sensing of environment*, *216*, pp.139-153.

Lee, J.G., Jun, S., Cho, Y.W., Lee, H., Kim, G.B., Seo, J.B. and Kim, N. (2017) Deep learning in medical imaging: general overview. *Korean journal of radiology*. *18*(4), pp.570-584.

Lhoest, S., Linchant, J., Quevauvillers, S., Vermeulen, C. and Lejeune, P. (2015) How many hippos (HOMHIP): algorithm for automatic counts of animals with infra-red thermal imagery from UAV. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. *40*.

Li, W., Dong, R., Fu, H. and Yu, L. (2019) Large-scale oil palm tree detection from high-resolution satellite images using two-stage convolutional neural networks. *Remote Sensing*, *11*(1), p.11.

Linchant, J., Lhoest, S., Quevauvillers, S., Lejeune, P., Vermeulen, C., Ngabinzeke, J.S., Belanganayi, B.L., Delvingt, W. and Bouché, P. (2018) UAS imagery reveals new survey opportunities for counting hippos. *PloS one*. *13*(11), p.e0206413.

Linchant, J., Lisein, J., Semeki, J., Lejeune, P. and Vermeulen, C. (2015) Are unmanned aircraft systems (UAS s) the future of wildlife monitoring? A review of accomplishments and challenges. *Mammal Review*, *45*(4), pp.239-252.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y. and Berg, A.C. (2016) October. Ssd: Single shot multibox detector. In *European conference on computer vision* (pp. 21-37). Springer, Cham.

Longmore, S.N., Collins, R.P., Pfeifer, S., Fox, S.E., Mulero-Pázmány, M., Bezombes, F., Goodwin, A., De Juan Ovelar, M., Knapen, J.H. and Wich, S.A. (2017) Adapting astronomical source detection software to help detect animals in thermal images obtained by unmanned aerial systems. *International Journal of Remote Sensing*, *38*(8-10), pp.2623-2638.

Lyons, M.B., Brandis, K.J., Murray, N.J., Wilshire, J.H., McCann, J.A., Kingsford, R.T. and Callaghan, C.T. (2019) Monitoring large and complex wildlife aggregations with drones. *Methods in Ecology and Evolution*.

Marini, S., Fanelli, E., Sbragaglia, V., Azzurro, E., Fernandez, J.D.R. and Aguzzi, J. (2018) Tracking fish abundance by underwater image recognition. *Scientific reports*, *8*(1), p.13748.

Matthews, W.S., Van Wyk, A.E., Van Rooyen, N. and Botha, G.A. (2001) Vegetation of the Tembe Elephant Park, Maputaland, South Africa. *South African Journal of Botany*, *67*(4), pp.573-594.

Moudrý, V., Gdulová, K., Fogl, M., Klápště, P., Urban, R., Komárek, J., Moudrá, L., Štroner, M., Barták, V. and Solský, M. (2019) Comparison of leaf-off and leaf-on combined UAV imagery and airborne LiDAR for assessment of a post-mining site terrain and vegetation structure: Prospects for monitoring hazards and restoration success. *Applied Geography*. *104*, pp.32-41.

Moya, Ó., Mansilla, P.L., Madrazo, S., Igual, J.M., Rotger, A., Romano, A. and Tavecchia, G. (2015) APHIS: a new software for photo-matching in ecological studies. *Ecological informatics*, *27*, pp.64-70.

Mubin, N.A., Nadarajoo, E., Shafri, H.Z.M. and Hamedianfar, A. (2019) Young and mature oil palm tree detection and counting using convolutional neural network deep learning method. *International Journal of Remote Sensing*, *40*(19), pp.7500-7515.

Mulero-Pázmány, M., Jenni-Eiermann, S., Strebel, N., Sattler, T., Negro, J.J. and Tablado, Z. (2017) Unmanned aircraft systems as a new source of disturbance for wildlife: A systematic review. *PloS one*, *12*(6), p.e0178448.

Mulero-Pázmány, M., Stolper, R., Van Essen, L.D., Negro, J.J. and Sassen, T. (2014) Remotely piloted aircraft systems as a rhinoceros anti-poaching tool in Africa. *PloS one*. *9*(1), p.e83873.

Mulero-Pázmány, M., Barasona, J.Á., Acevedo, P., Vicente, J. and Negro, J.J. (2015) Unmanned aircraft systems complement biologging in spatial ecology studies. *Ecology and evolution*. *5*(21), pp.4808-4818.

Nguyen, H., Maclagan, S.J., Nguyen, T.D., Nguyen, T., Flemons, P., Andrews, K., Ritchie, E.G. and Phung, D. (2017) October. Animal recognition and identification with deep convolutional neural networks for automated wildlife monitoring. In *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 40-49).

Norouzzadeh, M.S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M.S., Packer, C. and Clune, J. (2018) Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences*. *115*(25), pp. E5716-E5725.

Phung, D. (2017) Animal recognition and identification with deep convolutional neural networks for automated wildlife monitoring. In *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA).* (pp. 40-49).

Pix4D (2019) *Pix4D Mapper*. [Programme] Available from: https://www.pix4d.com/product/pix4dmapper-photogrammetry-software

Polar Pro (2019) *Cinema Series*. Available from: https://www.polarprofilters.com/collections/mavic-2-custom-filter-sets/products/custom-6-pack-cinema-series-mavic-2-pro

Pomeroy, P., O'connor, L. and Davies, P. (2015) Assessing use of and reaction to unmanned aerial systems in gray and harbor seals during breeding and molt in the UK. *Journal of Unmanned Vehicle Systems*, *3*(3), pp.102-113.

Popat, P., Sheth, P. and Jain, S. (2019) Animal/Object Identification Using Deep Learning on Raspberry Pi. In *Information and Communication Technology for Intelligent Systems* (pp. 319-327). Springer, Singapore.

Rivas, A., Chamoso, P., González-Briones, A. and Corchado, J. (2018) Detection of cattle using drones and convolutional neural networks. *Sensors*, *18*(7), p.2048.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M. and Berg, A.C. (2015) Imagenet large scale visual recognition challenge. *International journal of computer vision*, *115*(3), pp.211-252.

Sacchi, R., Scali, S., Mangiacotti, M., Sannolo, M. and Zuffi, M.A. (2016) Digital identification and analysis. *Reptile Ecology and Conservation A Handbook of Techniques*, pp.59-72.

Schneider, S., Taylor, G.W., Linquist, S. and Kremer, S.C. (2019) Similarity Learning Networks for Animal Individual Re-Identification-Beyond the Capabilities of a Human Observer. *arXiv preprint arXiv:1902.09324*.

Scotcher, J.S.B., Stewart, D.R.M. and Breen, C.M. (1978) The diet of the hippopotamus in Ndumu Game Reserve Natal, as determined by faecal analysis. *South African Journal of Wildlife Research-24-month delayed open access*, *8*(1), pp.1-11.

Seymour, A.C., Dale, J., Hammill, M., Halpin, P.N. and Johnston, D.W. (2017) Automated detection and enumeration of marine wildlife using unmanned aircraft systems (UAS) and thermal imagery. *Scientific Reports*. *7*, p.45127.

Simonyan, K. and Zisserman, A. (2014) Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Tablado, Z. and Jenni, L. (2017) Determinants of uncertainty in wildlife responses to human disturbance. *Biological Reviews*, *92*(1), pp.216-233.

Thapa, G.J., Thapa, K., Thapa, R., Jnawali, S.R., Wich, S.A., Poudyal, L.P. and Karki, S. (2018) Counting crocodiles from the sky: monitoring the critically endangered gharial (Gavialis gangeticus) population with an unmanned aerial vehicle (UAV). *Journal of Unmanned Vehicle Systems*. *6*(2), pp.71-82.

Trnovszký, T., Kamencay, P., Orješek, R., Benčo, M. and Sýkora, P. (2017) Animal recognition system based on convolutional neural network. *Digital Image Processing and Computer Graphics*. 15(3). pp.517-525.

Van Eeden, D.G. (2007) *Aspects of the ecology and conservation status of selected wildlife in and around Tembe Elephant Park, KwaZulu-Natal, South Africa* (Doctoral dissertation, University of Pretoria).

Van Rooyen, M.W., Tosh, C.A., Van Rooyen, N., Matthews, W.S. and Kellerman, M.J.S. (2004) Impact of harvesting and fire on Phragmites australis reed quality in Tembe Elephant Park, Maputaland. *Koedoe*, *47*(1), pp.31-40.

Vermeulen, C., Lejeune, P., Lisein, J., Sawadogo, P. and Bouché, P. (2013) Unmanned aerial survey of elephants. *PloS one*. *8*(2), p.e54700.

Wearn, O.R., Freeman, R. and Jacoby, D.M. (2019) Responsible AI for conservation. *Nature Machine Intelligence*, *1*(2), p.72.

Weimerskirch, H., Prudor, A. and Schull, Q. (2018) Flights of drones over sub-Antarctic seabirds show species-and status-specific behavioural and physiological responses. *Polar Biology*, *41*(2), pp.259-266.

Wich, S. and Pin Koh, L. (2018) Conservation Drones – Mapping and Monitoring Biodiversity. 1st ed. Oxford University Press.

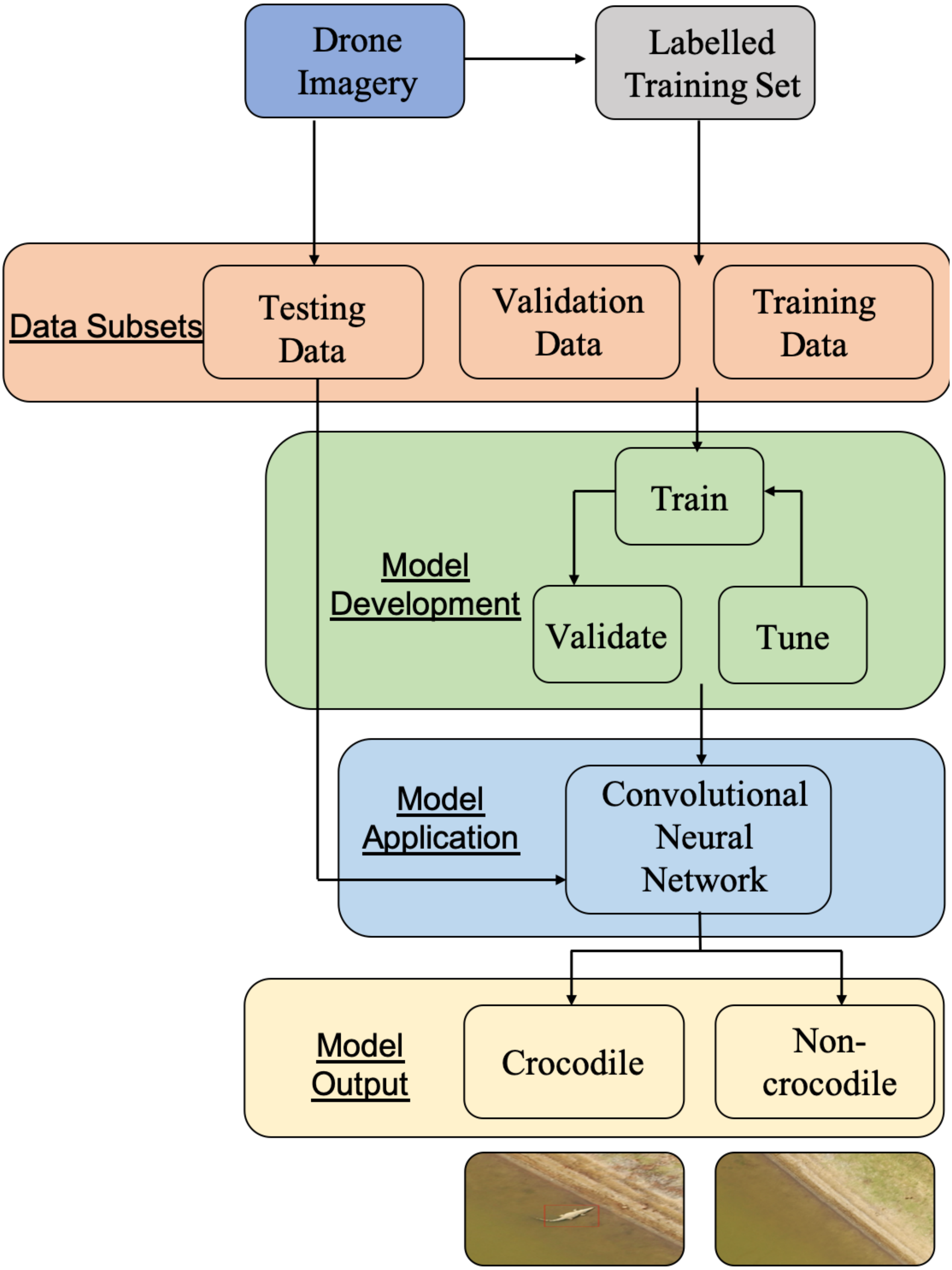
Willi, M., Pitman, R.T., Cardoso, A.W., Locke, C., Swanson, A., Boyer, A., Veldthuis, M. and Fortson, L. (2019) Identifying animal species in camera trap images using deep learning and citizen science. *Methods in Ecology and Evolution*. *10*(1), pp.80-91.

Witczuk, J., Pagacz, S., Zmarz, A. and Cypel, M. (2018) Exploring the feasibility of unmanned aerial vehicles and thermal imaging for ungulate surveys in forests-preliminary results. *International journal of remote sensing*, *39*(15-16), pp.5504-5521.

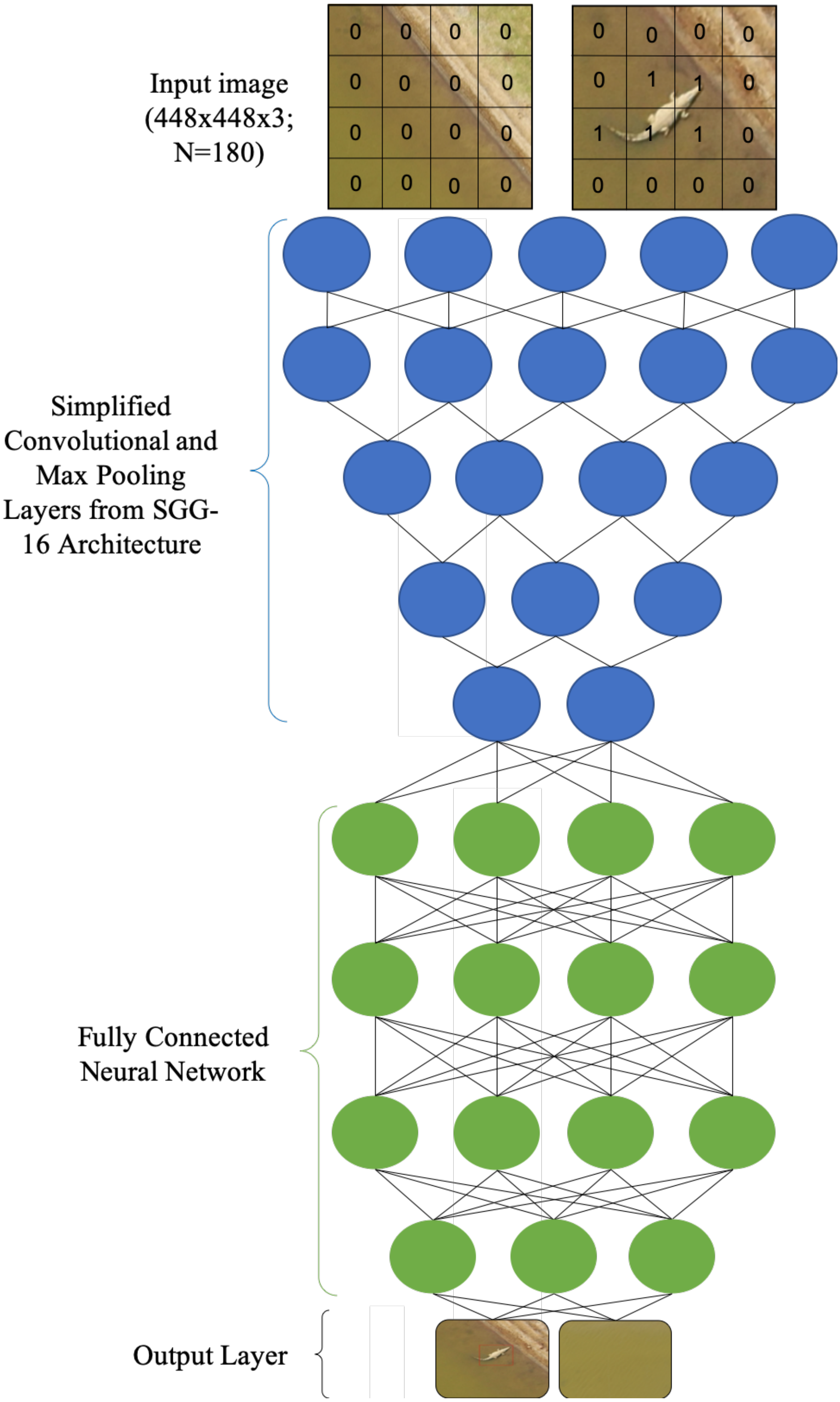
Xue, Y., Wang, T. and Skidmore, A.K. (2017) Automatic counting of large mammals from very high-resolution panchromatic satellite imagery. *Remote sensing*. *9*(9), p.878.

Yousif, H., Yuan, J., Kays, R. and He, Z. (2019) Animal Scanner: Software for classifying humans, animals, and empty frames in camera trap images. *Ecology and Evolution*.

9. Appendix



Appendix A Overview of Convolutional Neural Network analysis of crocodiles in Ndumo Game Reserve



Appendix B Overview of Convolutional Neural Network architecture. The input image is divided into 4 x 4 tensor. The convolutional layers perform the feature extraction for the CNN by scanning a few pixels at a time and creating a feature map. The max pooling layers reduce the amount of information, while maintaining the most important data. These layers were followed by fully connected layers which turns them into a single vector that can make the predictions for classification. The final layer employs a binary normalised exponential function which ingests the final fully connected layer and its learnt combinations of features, and returns a value between 1 and 0, with 1 signalling higher confidence of a present crocodile within that particular tile. This process is then replicated across the remaining tiles in the ortho-mosaic.