Received: 12 February 2021 | Revised: 7 October 2021 | Accepted: 25 November 2021

DOI: 10.1111/aje.12950

ARTICLE

African Journal of Ecology WILEY

Effectiveness of using drones and convolutional neural networks to monitor aquatic megafauna

Alexander B. Woolcock[1] | Sam Cotton[2] | Alison J. Cotton[2]

[1]University of the West of England, Frenchay Campus, Bristol, UK

[2]Department of Field Conservation and Science, Bristol Zoological Society, Clifton, UK

Correspondence
Sam Cotton, Department of Field Conservation and Science, Bristol Zoological Society, Clifton, Bristol, BS8 3HA, UK.
Email: scotton@bristolzoo.org.uk

Abstract

Aquatic megafauna are difficult to observe and count due to the inaccessibility and issues of detectability. Traditional transect and helicopter counts are useful for obtaining population estimates, but they often have logistical and cost limitations. The recent proliferation of drone technology offers an innovative way of surveying animal populations. However, data collected from drones are hindered by an analysis bottleneck that increases the time needed to process them. Convolutional Neural Networks (CNNs) are an emerging category of deep learning that can automate this data analysis process. Here, we compare traditional methods with drone surveys, by detecting and counting Nile crocodiles (*Crocodylus niloticus*) and common hippopotami (*Hippopotamus amphibious*). We evaluate the utility of CNNs for object detection and quantification in complex environments. Drone counts were more accurate than traditional methods; identifying 21% more crocodiles. Where vegetation was open, hippo counts with a drone showed a similar pattern (identifying 43% more). When vegetation was dense the drone produced less-accurate population estimates than traditional methods. CNN accuracy was limited (85%) due to the reduced training dataset available for the CNN. However, with an expanded data set, object detection is likely to be more accurate, making it more applicable for expedited and automated data analysis.

KEYWORDS
aquatic surveying, Convolutional neural networks, crocodiles, deep learning for ecology, hippopotami, population monitoring

Résumé

La mégafaune aquatique est difficile à observer et à dénombrer en raison de son inaccessibilité et des problèmes liés à sa détectabilité. Les dénombrements traditionnels par transect et par hélicoptère sont utiles afin d'obtenir des estimations de la population, mais ils comportent souvent des limites logistiques et financières. La prolifération récente de la technologie des drones offre un moyen novateur de recenser les populations animales. Cependant, l'efficacité du recueil de données par les drones est entravée par un goulot d'étranglement causé par le processus d'analyse, qui augmente le temps de traitement nécessaire de ces mêmes données. Les réseaux neuronaux convolutifs (RNC) sont une catégorie émergente d'apprentissage approfondi qui peut automatiser ce processus d'analyse de données. Notre objectif est ici de comparer les méthodes traditionnelles avec les relevés effectués par des drones, en détectant et en comptant les crocodiles du Nil (*Crocodylus niloticus*) et les hippopotames communs (*Hippopotamus amphibious*). Nous cherchons à évaluer l'utilité des RNC dans la détection et la quantification d'objets dans des environnements complexes. Le dénombrement effectué au moyen de drones était plus précis que celui obtenu par des méthodes traditionnelles; nous avons identifié 21 % de crocodiles supplémentaires. Les dénombrements d'hippopotames effectués avec un drone montraient une tendance similaire (43 % de plus) dans les zones à végétation ouverte. Dans les zones à végétation dense, les estimations de population effectuées à l'aide du drone étaient moins précises que celles réalisées par des méthodes traditionnelles. La précision des RCN était limitée (85 %) en raison de la quantité limitée des ensembles de données d'apprentissage disponible pour ces derniers. Cependant, la détection d'éléments est susceptible d'être plus précise avec des ensembles de données plus étendus, ce qui rendrait son utilisation plus adéquate aux fins d'accélération et d'automatisation des analyses de données.

1 | INTRODUCTION

Many wildlife populations are undergoing significant decline due to an increasing anthropogenic pressures, such as habitat degradation and intensive poaching (Lhoest et al., 2015). This has led to extinction rates that are a hundred-fold 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; Gray et al., 2019; Hodgson et al., 2018). The effectiveness of management decision-making is often dependent upon the quality and quantity of the ecological data upon which decisions are based. This means that improved data collection methods may herald more effective ecological outcomes (Dunstan et al., 2020; Hodgson et al., 2018).

Traditional methods of surveying large animals, such as drive counts and aerial counts, have proven to be a popular and successful way of obtaining population estimates (Jachmann, 2002). However, problems with safety, cost, visibility, statistical integrity and logistics have hindered the application of these methodologies (Jachmann, 2002; Jones, 2003). Observer efficiency is an important source of bias leading to underestimates of population size, especially where animals live in large congregations (Linchant et al., 2018). Therefore, these methods produce highly variable results (Combrink et al., 2011; Jachmann, 2002).

Drones (also known as unmanned aerial vehicles (or UAVs)) allow for high-resolution data acquisition in both the spatial and temporal domain and may overcome the constraints of terrestrial and occupied aircraft counts (Jiménez López & Mulero-Pázmány, 2019; Seymour et al., 2017; Thapa et al., 2018). 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). The use of drones for surveying is an increasing facet of conservation ecology and has 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).

The behaviour and habitat of aquatic mammals make collecting survey data particularly challenging. This is due to the fact that they spend much of their time underwater, move rapidly over large areas and occupy remote habitats (Anderson & Gaston, 2013; Gray et al., 2019). As a result, aerial surveys are typically used to collect population data on these species (Chrétien et al., 2016). Drones are able to carry out similar tasks as helicopters, often more reliably, at lower costs, and inducing less disturbance on the target species (Ezat et al., 2018; Linchant et al., 2018; Raoult et al., 2020; Schofield et al., 2019). Hodgson et al. (2016) found that drone derived estimates of population size, resulted in smaller cumulative variances than other methods. However, drones are restricted by a shorter flight time than conventional helicopters, which drastically reduces the area that can be surveyed.

Whilst drones can collect detailed information rapidly, they do not overcome an existing data analysis bottleneck (Nguyen et al., 2017; Seymour et al., 2017). Specifically; manually counting animals in resultant imagery is time consuming and inefficient. One possible method for overcoming this constraint is to employ automated methods for detection, localisation and enumeration of target animals (Corcoran et al., 2021). A range of techniques can be incorporated with varying degrees of accuracy (Kalantar et al., 2016; Ventura et al., 2018). Convolutional Neural Networks (CNNs) are a prominent and rapidly expanding category of deep learning classifier, inspired by the neural connections in the brain (Lee et al., 2017). These allow efficient discrimination of objects in noisy and complex environments (Dujon & Schofield, 2019; Gray et al., 2019) and may expedite the data analysis process (Brodrick et al., 2019). CNNs are, however, complex to implement, computationally intensive, and may require more data than is practicable for most ecological studies (Dujon et al., 2021; Gray et al., 2019).

In this study, we evaluate the comparative effectiveness of using ground, helicopter and drone counts for surveying populations of Nile crocodiles (*Crocodylus niloticus*) and hippopotami (*Hippopotamus amphibious*) in two South African protected areas. As different reserves use a variety of survey methods, and emerging technologies offer new and innovative census techniques, it is essential to differentiate between these methods, so as to maximise the efficiency of the identification of population changes over spatial and temporal periods. As a new and innovative solution to the existing data analysis bottleneck, machine learning has had few real-world applications. There have been a small number of studies that have applied machine learning to analyse aerial imagery in Africa (Eikelboom et al., 2019; Kellenberger et al., 2018); however, these methods have been semi-automatous and still require large amounts of data analysis. Given the continued ecological decline, innovative technologies such as CNNs and drones may offer scientists more accurate and efficient methodologies, thereby aiding conservation initiatives. The paucity of research data on the subject means that our research can act as baseline data to inform and advise future studies in this promising method for population monitoring.

2 | METHODOLOGY

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). The Muzi Swamp, is an extensive system covering about 560-ha, stretching for 25km from north to south along the eastern edge of the reserve (Gaugris & Van Rooyen, 2010; Van Eeden, 2007; Van Rooyen et al., 2004) consisting of dense reedbeds, with stands of *Phragmites australis* as the most abundant species (Van Rooyen et al., 2004). 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 act as a focal point for this study.

NGR (10,000-ha) is one of the oldest game reserves in South Africa. Established in 1924 as a sanctuary for hippopotami. Situated on the Mozambique Coastal Plain (−26.890264, 32.300418) it supports the third largest crocodile population in South Africa (Calverley & Downs, 2014; Combrink, 2004). The reserve is inherently patchy and characterised by a mosaic of permanent and ephemeral pans, streams and rivers with varying degrees of connectivity, which

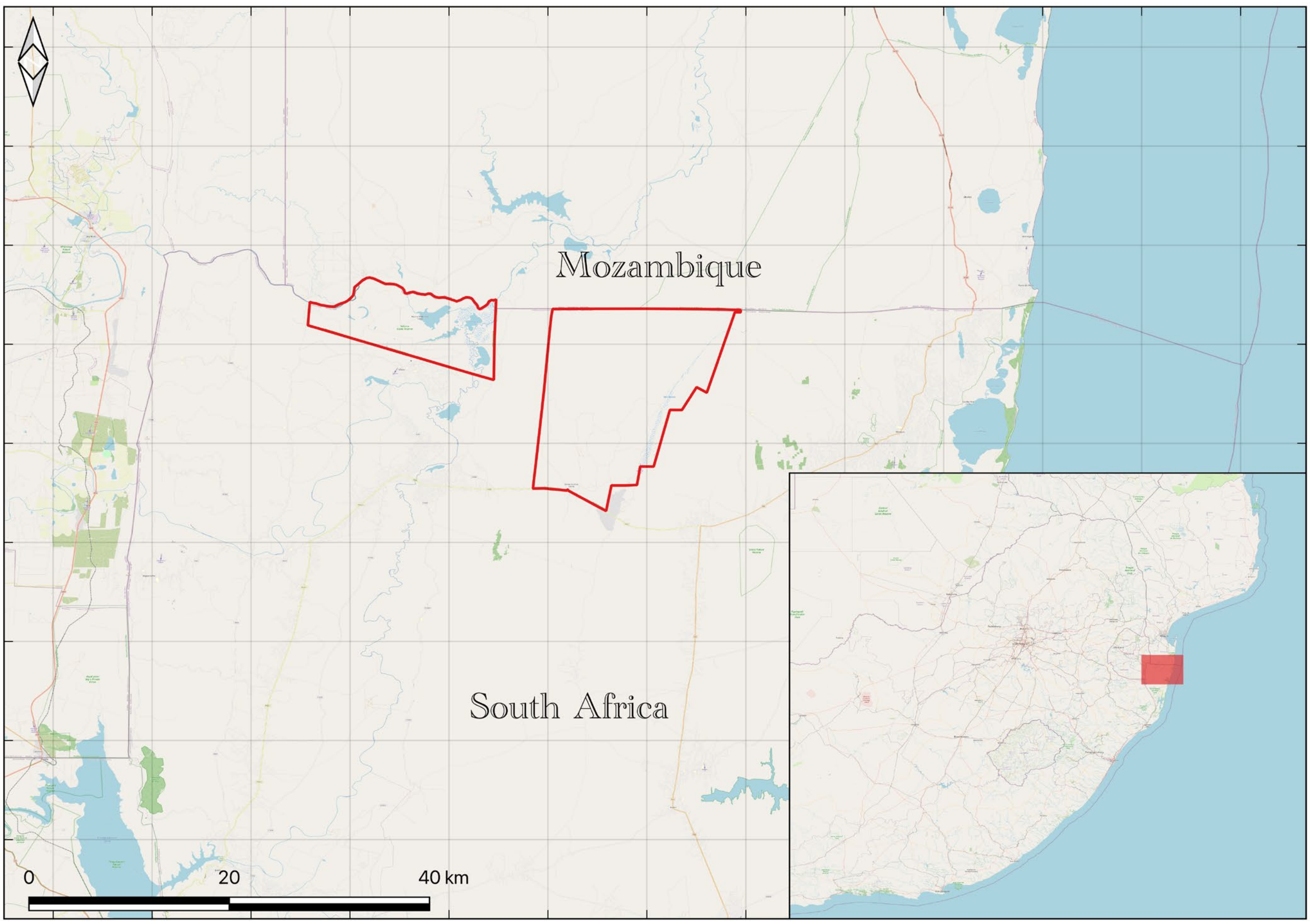

FIGURE 1 Map of Tembe Elephant Park & Ndumo Game Reserve

fluctuate on a seasonal basis (Calverley & Downs, 2014). The largest permanent pan in NGR is Lake Nyamathi (55ha), which possesses the highest densities of crocodiles and hippos in the reserve (Calverley, 2013; Calverley & Downs, 2014, 2015). This lake will act as the focal point for research within this reserve. Lake Nyamathi 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).

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 (temperature (±SD) =23(±3)°C; clear sky; windspeed (±SD) 5(±2)mph). 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 counting methods.

In NGR Lake Nyamathi is circumnavigated by a road, which is approximately 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 (Figure 2). The data were collected using ArcGIS Collector (ESRI, 2018-2019); the species and GPS location was recorded. 10 X 42 binoculars were used to assist the counting of individuals. A total of three repeats were carried out. In the case of TEP there is no road, which runs along with the Muzi Swamp where the water bodies were located. This made it necessary to carry out the ground counts on foot, which involved walking a transect of 7.2km along the southern side of the Muzi Swamp (Figure 2). As was the case with NGR, two observers counted crocodiles and hippos, whilst 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.

A Long Ranger helicopter containing four people was flown on a predetermined north-south oriented axis over TEP and NGR (Figure 2). 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 (~55–75 km/h). When hippos and crocodiles were observed the helicopter deviated from the transect, and a total count was undertaken with a camera to validate the sightings. The data were 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).

The drone surveys were carried out with a standard DJI Mavic Pro 2 quadcopter with a 77 degree field of view. 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 in grid pattern whilst capturing aerial imagery (Figure 2). 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. A total of 1335 images were taken of Lake Nyamathi and 550 images were taken from the Muzi swamp. Due to the drone's 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. The study was approved by UWE's Animal Ethics and Welfare Committee (UWE Bristol, 2020) prior to study commencing. There were no obvious negative interactions with the crocodiles and hippos; further animal ethical considerations can be found in Mulero-Pázmány et al. (2017).

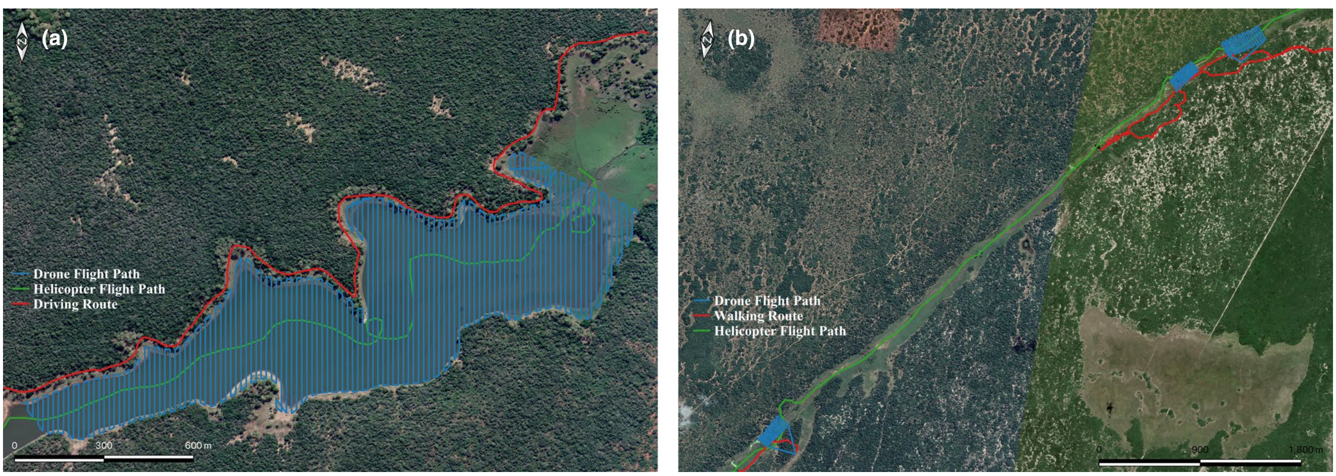

FIGURE 2 Survey Routes in Tembe Elephant Park & Ndumo Game Reserve

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 ortho-mosaics were manually checked for inconsistencies, for example animal movements between adjacent images. However, no inconsistencies were identified. A total of seven ortho-mosaics were created: four of Nyamathi 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 the target species. Due to the limited number of hippos identified in the drone imagery, crocodiles were chosen as the target species. Crocodiles were present in larger number allowing the larger training set to be created, improving the accuracy of the algorithm. 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 selected and exported as a Tiff file, which ensured that the training data retained RGB bands. The training samples were taken from a range of specifications, for example basking, partially submerged, overlapping crocodiles etc., this allowed the CNN to identify crocodiles in the broadest range of environments/ behaviours. Each training sample comprised a 448 × 448 × 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 and Zisserman (2014), which has shown high quality image classification and superiority over other networks (Canziani et al., 2016; Liu et al., 2016; Russakovsky et al., 2015). This architecture acts 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 and 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. 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, whilst 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 was constructed, it was trained over several epochs, with an epoch being the number of complete passes over the training dataset the algorithm has completed. 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 3).

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

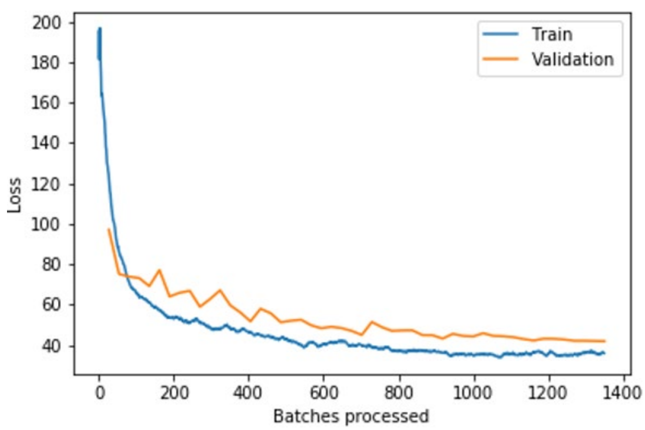

FIGURE 3 Loss graph outlining the training of the CNN

was then imported into ArcGIS and run over a trimmed portion of the ortho-mosaic, which included Lake Nyamathi and a 4m margin, which consisted of open ground adjacent to the water's edge where the crocodiles would bask in the sun. This decreased the running time of the model and also decreased the objects (e.g. trees), which could give false positives. Refer to Figure A1, which shows a flow chart of the CNN analysis, and Figure A2, 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 (1) 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.

$$\text{Margin of error} = \frac{(D - N)}{N} \times 100. \tag{1}$$

As was carried out by Li et al. (2019) the precision, recall and *F1-score* are calculated via equations (2)-(4), 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.

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{2}$$

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{3}$$

$$F1 - \text{Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{(\text{Precision} + \text{Recall})}. \tag{4}$$

3 | RESULTS

Results of ground, helicopter and drone comparative surveys for TEP are summarised in Table 1. The drone data allowed precise mapping of the TEP floodplain with extensive detail allowing easy identification of individual crocodiles and hippopotami. The ortho-mosaic obtained from the aerial imagery, shows the water bodies in the Muzi Swamp that have not yet dried up (Figure 4C). 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. Only hippos in one pod were seen from the drone imagery, and these were basking in the middle of a pan in open water (Figure 4F). In the three repetitions, only one pod was located, which was in the same pan on each repetition. The ground and helicopter surveys observed a separate pod of hippos in a northern pan, which were not detected in the drone survey (Figure 4A, 4B & 4C). These hippos were observed in both instances emerging from hydrophytic 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

Tembe	Crocodiles counted	Hippos counted
Walking Counts		
18/09/2019	0	19
Helicopter Counts		
29/07/2019	4	29
Drone Counts		
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
Standard Error (SE)	0.816	0.408

Results of ground, helicopter and drone comparative surveys for NGR are summarised in Table 2. The hippos were found 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). A slightly higher density of crocodiles was found on the northern edge of the lake (Figure 4D, 4E) where the ground was less rocky. The helicopter counts counted 15 hippos in the same two pods as did the drive counts. A total of four drone censuses were conducted over lake Nyamathi. This involved mapping a total area of 57 hectares. As was the case with the ground and helicopter surveys, the hippos were in two distinctive pods: the first basking on the northern bank of the lake, and the second; wallowing in the water (Figure 4F). The drone counts again showed similar spatial data as the ground counts, with a higher density of crocodiles on the northern shore.

Figure 5 shows the CNN applied to an ortho-mosaic conducted in Nyamathi. 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 downed tree trunks or occasionally birds. Manual counts were seen to have better detection capability, on average counting +14.43% more crocodiles than the CNN.

4 | DISCUSSION

The ability of the drone 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 drone yielded higher population estimates. In NGR the drone counted 21.6% more crocodiles than ground surveys and 21.1% more than helicopter surveys. In TEP results showed a similar pattern (drone was 100%

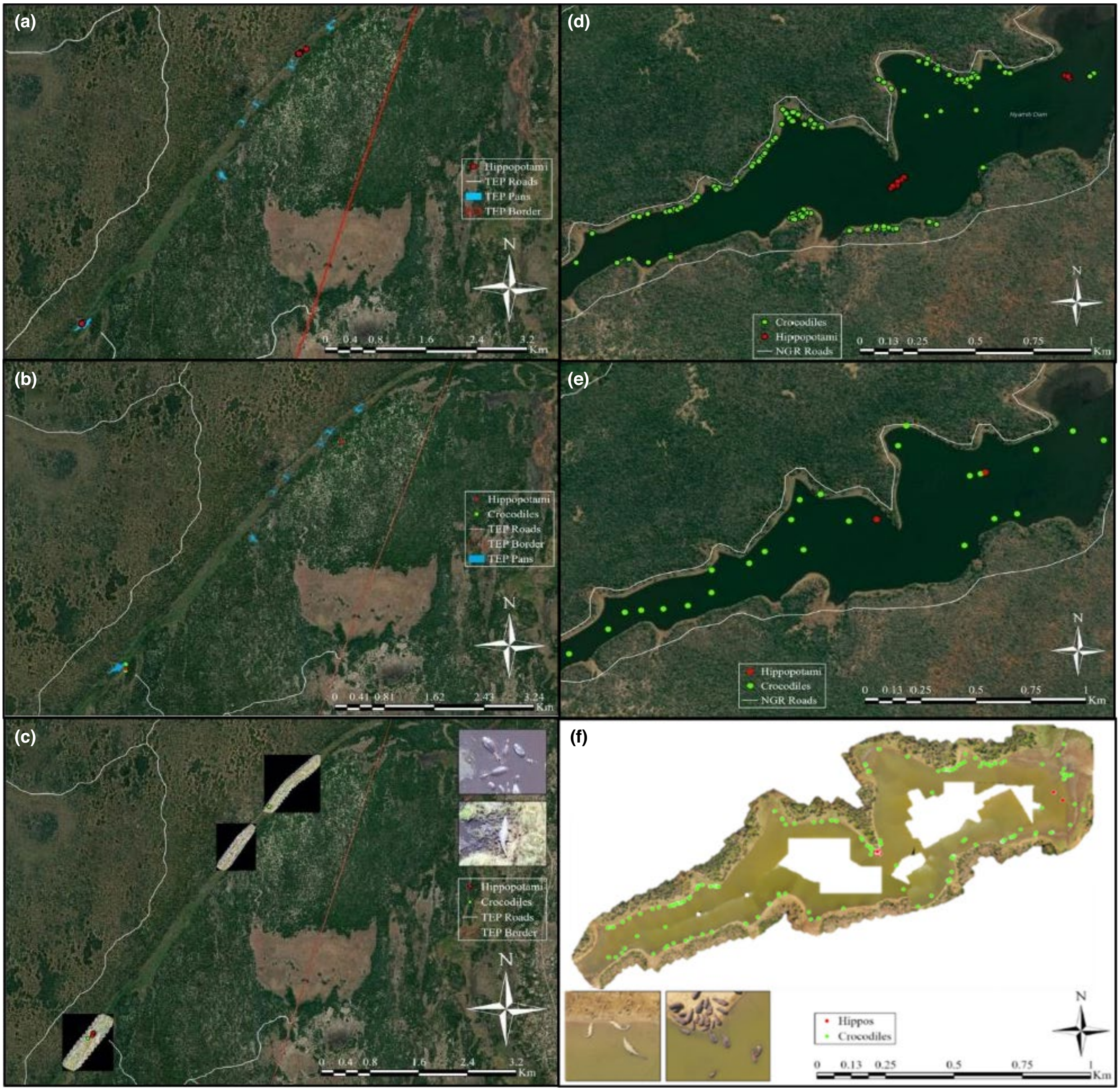

FIGURE 4 Comparative map of methodologies conducted in TEP and NGR. (a) Drive counts TEP; (b) Helicopter counts TEP; (c) Drone counts TEP on 30/8/2019; (d) Walking counts NGR; (e) Helicopter counts NGR; (f) Drone counts NGR on 4/8/2019. The drone counts give an image of both crocodiles and hippos, showing the resolution whilst flying at a height 100 m

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. Drone counts produce the highest figure, and with little variance in repeated surveys. Hodgson et al. (2016) found that drone derived estimates of population size resulted in smaller cumulative variances than drive counts. This was also found with this research at NGR. 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. When counting crocodiles, it is, therefore, much more efficient and accurate than the other two methods.

Our study indicated that when counting hippos in TEP, the helicopter counts yielded more accurate population estimates than walking counts (34.5%) and drone counts (44.8%). In the case of TEP the encounter rate of hippos, and clearly their detectability recorded with the drone, was lower than that of the helicopter surveys, resulting lower density estimates. Many studies have highlighted the fact that manned aircraft can disturb wildlife (Mulero-Pázmány et al., 2017), resulting in altered and sometimes illusive behaviour that can make the study and enumeration of animals problematic (Chrétien et al., 2016; Christie et al., 2016; Gentle et al., 2018; Jones et al., 2006).

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

Ndumo Game Reserve	Crocodiles counted	Hippos counted
Drive Counts		
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
Standard Error (SE)	5.715	2.373
Helicopter Counts		
30/07/2019	152	15
Drone Counts		
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
Standard Error (SE)	1.139	1.299

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). 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 et al., 2015; Tablado & Jenni, 2017), as often the animal does not exhibit any symptoms (Ditmer et al., 2015). However, Mulero-Pázmány et al. (2017) has found that noise is the most prominent cause of disturbance. This theory is supported by our research.

During the three repeats of the drone surveys in TEP, only one pod of hippos was observed. The northern pod was not detected using drones (it was using helicopter surveys). We can, therefore, 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 (Chrétien et al.,

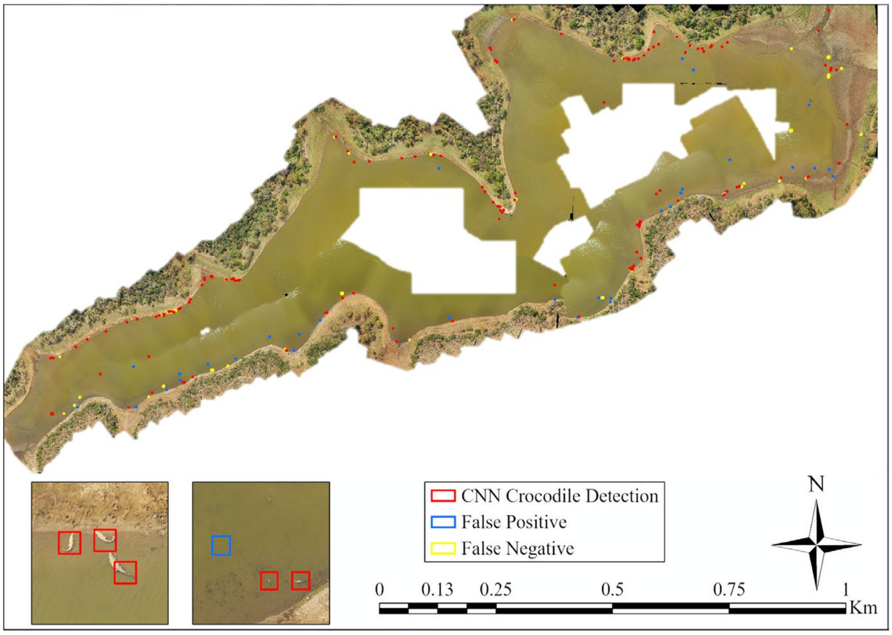

FIGURE 5 CNN detections of Crocodiles in NGR 13/09/2019

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

Actual Number of crocodiles	True Positives	False Positives	False Negatives	Margin of error (%)	Precision (%)	Recall (%)	F1-Score (%)
194	166	31	28	−14.43	84.26	85.57	84.91

2016; Christie et al., 2016; Gentle et al., 2018; Jones et al., 2006). 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? Raoult et al. (2020) gives further guidelines on conducting research with drones on specific aquatic species. An alternative method 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 (Kays et al., 2019; Longmore et al., 2017; Witczuk et al., 2018), and could dramatically improve the detection of hippos located in dense vegetation, whilst posing little or no disturbance. However, in NGR the drone counted 43.4% more hippos than helicopter and 41.1% more than driving counts. 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. Conversely Nyamathi, is made up of open water with the closest vegetation being <4 m from the edge of the water. This makes aquatic megafauna far easier to observe, allowing more accurate population estimates.

Given that the three survey methods were not carried out simultaneously, some temporal variability is possible in the densities found due to animal 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 Nyamathi 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. As this study was conducted during the dry season, there were no flooding events and extensive grazing was available surrounding the Nyamathi and Muzi Swamps. We can assume, therefore, that the movement of hippos was minimal and had little impact on our comparison of the methods.

Previous work has shown CNNs to be an accurate and time efficient analysis technique to enumerate animals in drone imagery (Barry, 2018; Corcoran et al., 2019; Kellenberger et al., 2018; Rivas et al., 2018). Gray 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; Willi et al., 2019; Yousif 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 (Brodrick et al., 2019; Guirado et al., 2019; Li et al., 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 to compete with Chabot et al. (2018)—85,267 training images; Chew et al. (2018)—1500 training images; Gray et al. (2019)—467 training images; Mubin et al. (2019)—260 training images; Cheang et al. (2017)—300 training images; Eikelboom et al. (2019)—516 training images; you would expect to see similar levels of accuracies as Chew et al. (2018) (96.4%). Gathering training sets as large as the aforementioned can be a time-consuming and costly process. Papakonstantinou et al. (2021) has implemented citizen based imagery analysis to improve detection results, which may help to overcome the challenges with obtaining large training schemes. The more inherently complex the data, the larger the training scheme needs to be (Wearn et al., 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.

5 | CONCLUSION

This research compared three 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. 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. 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.

Technology is permeating almost all aspects of society. As this pace of innovation continues to accelerate, conservationists need to be aggressive in adopting, adapting and channelling these advances into positive outcomes for world biodiversity. This study succeeds in its examination of drones and CNNs to better understand complex ecological patterns. Further research, however, is needed to increase accessibility and improve implementation.

ACKNOWLEDGEMENT

We thank Ezemvelo KZN Wildlife for their assistance and logistical support, with particular reference to Catariné Hanekom and Len Gunter. We also thank E. Woolcock; M. Sibiya; Corporal Massimara and M. Siengya for their fieldwork assistance.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Alexander B. Woolcock https://orcid.org/0000-0003-3683-9230

REFERENCES

Anderson, K., & Gaston, K. J. (2013). Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Frontiers in Ecology and the Environment*, *11*(3), 138–146. https://doi.org/10.1890/120150

Barry, J. (2018). Applications of deep learning in ornithology. *Biodiversity Information Science and Standards*, *2*, e27251.

Bristol, U. W. E. (2020). *Policy statement and the use of animals in research*. https://www1.uwe.ac.uk/research/researchgovernance/animalsinresearchstatement.aspx. Accessed 20 August 2020

Brodrick, P. G., Davies, A. B., & Asner, G. P. (2019). Uncovering ecological patterns with convolutional neural networks. *Trends in Ecology & Evolution*, *34*(8), 734–745.

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*. University of KwaZulu-Natal. (Doctoral dissertation.

Calverley, P. M., & Downs, C. T. (2014). Habitat use by Nile crocodiles in Ndumo Game Reserve, South Africa: A naturally patchy environment. *Herpetologica*, *70*(4), 426–438. https://doi.org/10.1655/HERPETOLOGICA-D-13-00088

Calverley, P. M., & Downs, C. T. (2015). Movement and home range of Nile crocodiles in Ndumo game reserve. *South Africa. Koedoe*, *57*(1), 1–13. https://doi.org/10.4102/koedoe.v57i1.1234

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

Chabot, D., Dillon, C., & Francis, C. (2018). An approach for using off-the-shelf object-based image analysis software to detect and count birds in large volumes of aerial imagery. *Avian Conservation and Ecology*, *13*(1). https://doi.org/10.5751/ACE-01205-130115

Cheang, E. K., Cheang, T. K., & 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., & 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), 12. https://doi.org/10.1186/s12942-018-0132-1

Chrétien, L. P., Théau, J., & 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), 181–191. https://doi.org/10.1002/wsb.629

Christie, K. S., Gilbert, S. L., Brown, C. L., Hatfield, M., & 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), 241–251. https://doi.org/10.1002/fee.1281

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

Combrink, X., Korrûbel, J. L., Kyle, R., Taylor, R., & Ross, P. (2011). Evidence of a declining Nile crocodile (*Crocodylus niloticus*) population at Lake Sibaya, South Africa. *African Journal of Wildlife Research*, *41*(2), 145–157.

Corcoran, E., Denman, S., Hanger, J., Wilson, B., & Hamilton, G. (2019). Automated detection of koalas using low-level aerial surveillance and machine learning. *Scientific Reports*, *9*(1), 3208. https://doi.org/10.1038/s41598-019-39917-5

Corcoran, E., Winsen, M., Sudholz, A., & Hamilton, G. (2021). Automated detection of wildlife using drones: Synthesis, opportunities and constraints. *Methods in Ecology and Evolution*, *12*(6), 1103–1114. https://doi.org/10.1111/2041-210X.13581

Ditmer, M. A., Vincent, J. B., Werden, L. K., Tanner, J. C., Laske, T. G., Iaizzo, P. A., Garshelis, D. L., & Fieberg, J. R. (2015). Bears show a physiological but limited behavioral response to unmanned aerial vehicles. *Current Biology*, *25*(17), 2278–2283. https://doi.org/10.1016/j.cub.2015.07.024

Dujon, A. M., Ierodiaconou, D., Geeson, J. J., Arnould, J. P., Allan, B. M., Katselidis, K. A., & Schofield, G. (2021). Machine learning to detect marine animals in UAV imagery: Effect of morphology, spacing, behaviour and habitat. *Remote Sensing in Ecology and, Conservation*, *7*(3), 341–354. https://doi.org/10.1002/rse2.205

Dujon, A. M., & Schofield, G. (2019). Importance of machine learning for enhancing ecological studies using information-rich imagery. *Endangered Species Research*, *39*, 91–104. https://doi.org/10.3354/esr00958

Dunstan, A., Robertson, K., Fitzpatrick, R., Pickford, J., & Meager, J. (2020). Use of unmanned aerial vehicles (UAVs) for mark-resight nesting population estimation of adult female green sea turtles at Raine Island. *PLoS One*, *15*(6), e0228524. https://doi.org/10.1371/journal.pone.0228524

Eikelboom, J. A., Wind, J., van de Ven, E., Kenana, L. M., Schroder, B., de Knegt, H. J., van Langevelde, F., & Prins, H. H. (2019). Improving the precision and accuracy of animal population estimates with aerial image object detection. *Methods in Ecology and Evolution*, *10*(11), 1875–1887.

Elsey, R. M., & Trosclair, P. L. (2016). The use of an unmanned aerial vehicle to locate alligator nests. *Southeastern Naturalist*, *15*(1), 76–83. https://doi.org/10.1656/058.015.0106

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

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

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

Ezat, M. A., Fritsch, C. J., & 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*, 76–81. https://doi.org/10.1016/j.biocon.2018.04.032

Gaugris, J. Y., & 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), 126–134. https://doi.org/10.1111/j.1365-2028.2009.01095.x

Gentle, M., Finch, N., Speed, J., & Pople, A. (2018). A comparison of unmanned aerial vehicles (drones) and manned helicopters for monitoring macropod populations. *Wildlife Research*, *45*(7), 586–594. https://doi.org/10.1071/WR18034

Gray, P. C., Fleishman, A. B., Klein, D. J., McKown, M. W., Bézy, V. S., Lohmann, K. J., & Johnston, D. W. (2019). A convolutional neural network for detecting sea turtles in drone imagery. *Methods in Ecology and Evolution*, *10*(3), 345–355. https://doi.org/10.1111/2041-210X.13132

Guirado, E., Tabik, S., Rivas, M. L., Alcaraz-Segura, D., & Herrera, F. (2019). Whale counting in satellite and aerial images with deep learning. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-50795-9

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

Hodgson, J. C., Baylis, S. M., Mott, R., Herrod, A., & Clarke, R. H. (2016). Precision wildlife monitoring using unmanned aerial vehicles. *Scientific Reports*, *6*, 22574. https://doi.org/10.1038/srep22574

Hodgson, J. C., Mott, R., Baylis, S. M., Pham, T. T., Wotherspoon, S., Kilpatrick, A. D., Raja Segaran, R., Reid, I., Terauds, A., & Koh, L. P. (2018). Drones count wildlife more accurately and precisely than humans. *Methods in Ecology and Evolution*, *9*(5), 1160–1167. https://doi.org/10.1111/2041-210X.12974

Jachmann, H. (2002). Comparison of aerial counts with ground counts for large African herbivores. *Journal of Applied Ecology*, *39*(5), 841–852. https://doi.org/10.1046/j.1365-2664.2002.00752.x

Jiménez López, J., & Mulero-Pázmány, M. (2019). Drones for conservation in protected areas: Present and future. *Drones*, *3*(1), 10. https://doi.org/10.3390/drones3010010

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

Jones, G. P., Pearlstine, L. G., & Percival, H. F. (2006). An assessment of small unmanned aerial vehicles for wildlife research. *Wildlife Society Bulletin*, *34*(3), 750–758. https://doi.org/10.2193/0091-7648(2006)34[750:AAOSUA]2.0.CO;2

Kalantar, B., Mansor, S. B., Shafri, H. Z. M., & Halin, A. A. (2016). Integration of template matching and object-based image analysis for semi-automatic oil palm tree counting in UAV images. In: *Proceedings of the 37th Asian Conference on Remote Sensing*, (pp. 17-21).

Kays, R., Sheppard, J., Mclean, K., Welch, C., Paunescu, C., Wang, V., Kravit, G., & Crofoot, M. (2019). Hot monkey, cold reality: surveying rainforest canopy mammals using drone-mounted thermal infrared sensors. *International Journal of Remote Sensing*, *40*(2), 407–419. https://doi.org/10.1080/01431161.2018.1523580

Kellenberger, B., Marcos, D., & Tuia, D. (2018). Detecting mammals in UAV images: Best practices to address a substantially imbalanced dataset with deep learning. *Remote Sensing of Environment*, *216*, 139–153. https://doi.org/10.1016/j.rse.2018.06.028

Lee, J. G., Jun, S., Cho, Y. W., Lee, H., Kim, G. B., Seo, J. B., & Kim, N. (2017). Deep learning in medical imaging: general overview. *Korean Journal of Radiology*, *18*(4), 570–584. https://doi.org/10.3348/kjr.2017.18.4.570

Lhoest, S., Linchant, J., Quevauvillers, S., Vermeulen, C., & Lejeune, P. (2015). How many hippos (homhip): algorithm for automatic counts of animals with infra-red thermal imagery from uav. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XL-3/W3*, 355–362.

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

Linchant, J., Lhoest, S., Quevauvillers, S., Lejeune, P., Vermeulen, C., Ngabinzeke, J. S., Belanganayi, B. L., Delvingt, W., & Bouché, P. (2018). UAS imagery reveals new survey opportunities for counting hippos. *PLoS One*, *13*(11), e0206413. https://doi.org/10.1371/journal.pone.0206413

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

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., & 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), 2623–2638. https://doi.org/10.1080/01431161.2017.1280639

Moya, Ó., Mansilla, P. L., Madrazo, S., Igual, J. M., Rotger, A., Romano, A., & Tavecchia, G. (2015). APHIS: a new software for photo-matching in ecological studies. *Ecological Informatics*, *27*, 64–70. https://doi.org/10.1016/j.ecoinf.2015.03.003

Mubin, N. A., Nadarajoo, E., Shafri, H. Z. M., & 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), 7500–7515. https://doi.org/10.1080/01431161.2019.1569282

Mulero-Pázmány, M., Jenni-Eiermann, S., Strebel, N., Sattler, T., Negro, J. J., & Tablado, Z. (2017). Unmanned aircraft systems as a new source of disturbance for wildlife: A systematic review. *PLoS One*, *12*(6), e0178448. https://doi.org/10.1371/journal.pone.0178448

Nguyen, H., Maclagan, S. J., Nguyen, T. D., Nguyen, T., Flemons, P., Andrews, K., Ritchie, E. G., & 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).

Papakonstantinou, A., Batsaris, M., Spondylidis, S., & Topouzelis, K. (2021). A Citizen science unmanned aerial system data acquisition protocol and deep learning techniques for the automatic detection and mapping of marine litter concentrations in the coastal zone. *Drones*, *5*(1), 6. https://doi.org/10.3390/drones5010006

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

Pomeroy, P., O'connor, L., & 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), 102–113. https://doi.org/10.1139/juvs-2015-0013

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

Raoult, V., Colefax, A. P., Allan, B. M., Cagnazzi, D., Castelblanco-Martínez, N., Ierodiaconou, D., Johnston, D. W., Landeo-Yauri, S., Lyons, M., Pirotta, V., Schofield, G., & Butcher, P. A. (2020). Operational protocols for the use of drones in marine animal research. *Drones*, *4*(4), 64. https://doi.org/10.3390/drones4040064

Rivas, A., Chamoso, P., González-Briones, A., & Corchado, J. (2018). Detection of cattle using drones and convolutional neural networks. *Sensors*, *18*(7), 2048. https://doi.org/10.3390/s18072048

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. I. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, *115*(3), 211–252. https://doi.org/10.1007/s11263-015-0816-y

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

Schofield, G., Esteban, N., Katselidis, K. A., & Hays, G. C. (2019). Drones for research on sea turtles and other marine vertebrates – A review. *Biological Conservation*, *238*, 108214. https://doi.org/10.1016/j.biocon.2019.108214

Scotcher, J. S. B., Stewart, D. R. M., & 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), 1–11.

Seymour, A. C., Dale, J., Hammill, M., Halpin, P. N., & Johnston, D. W. (2017). Automated detection and enumeration of marine wildlife using unmanned aircraft systems (UAS) and thermal imagery. *Scientific Reports*, *7*, 45127. https://doi.org/10.1038/srep45127

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

Tablado, Z., & Jenni, L. (2017). Determinants of uncertainty in wildlife responses to human disturbance. *Biological Reviews*, *92*(1), 216–233. https://doi.org/10.1111/brv.12224

Thapa, G. J., Thapa, K., Thapa, R., Jnawali, S. R., Wich, S. A., Poudyal, L. P., & 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), 71–82.

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*. University of Pretoria). (Doctoral dissertation).

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

Ventura, D., Bonifazi, A., Gravina, M. F., Belluscio, A., & Ardizzone, G. (2018). Mapping and classification of ecologically sensitive marine habitats using unmanned aerial vehicle (UAV) imagery and object-based image analysis (OBIA). *Remote Sensing*, *10*(9), 1331. https://doi.org/10.3390/rs10091331

Wearn, O. R., Freeman, R., & Jacoby, D. M. (2019). Responsible AI for conservation. *Nature Machine Intelligence*, *1*(2), 72. https://doi.org/10.1038/s42256-019-0022-7

Wich, S., & 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., & Fortson, L. (2019). Identifying animal species in camera trap images using deep learning and citizen science. *Methods in Ecology and Evolution*, *10*(1), 80–91. https://doi.org/10.1111/2041-210X.13099

Witczuk, J., Pagacz, S., Zmarz, A., & 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), 5504–5521. https://doi.org/10.1080/01431161.2017.1390621

Yousif, H., Yuan, J., Kays, R., & He, Z. (2019). Animal Scanner: Software for classifying humans, animals and empty frames in camera trap images. *Ecology and Evolution*. https://doi.org/10.1002/ece3.4747

How to cite this article: Woolcock, A. B., Cotton, S., & Cotton, A. J. (2021). Effectiveness of using drones and convolutional neural networks to monitor aquatic megafauna. *African Journal of Ecology*, 00, 1–13. https://doi.org/10.1111/aje.12950

APPENDIX

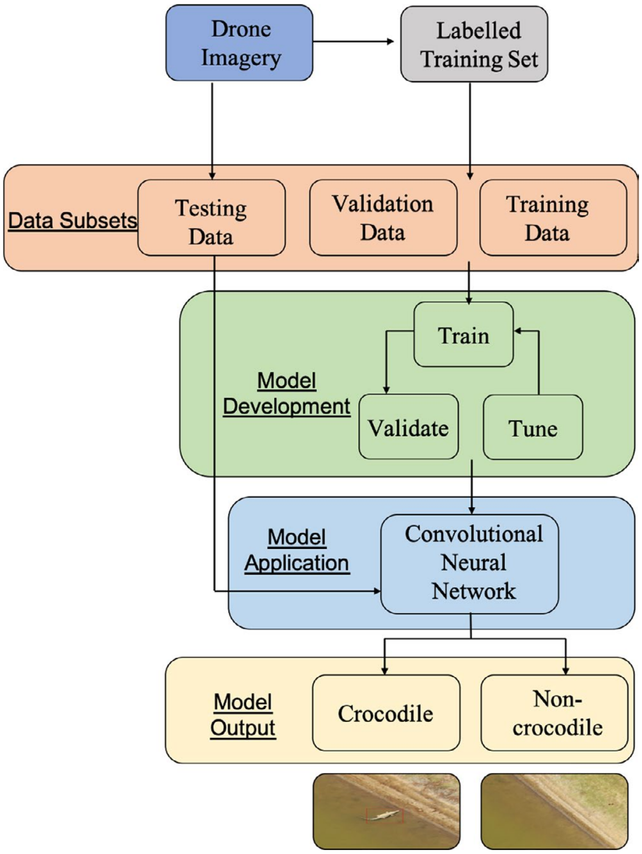

FIGURE A1 Overview of Convolutional Neural Network analysis of crocodiles in Ndumo Game

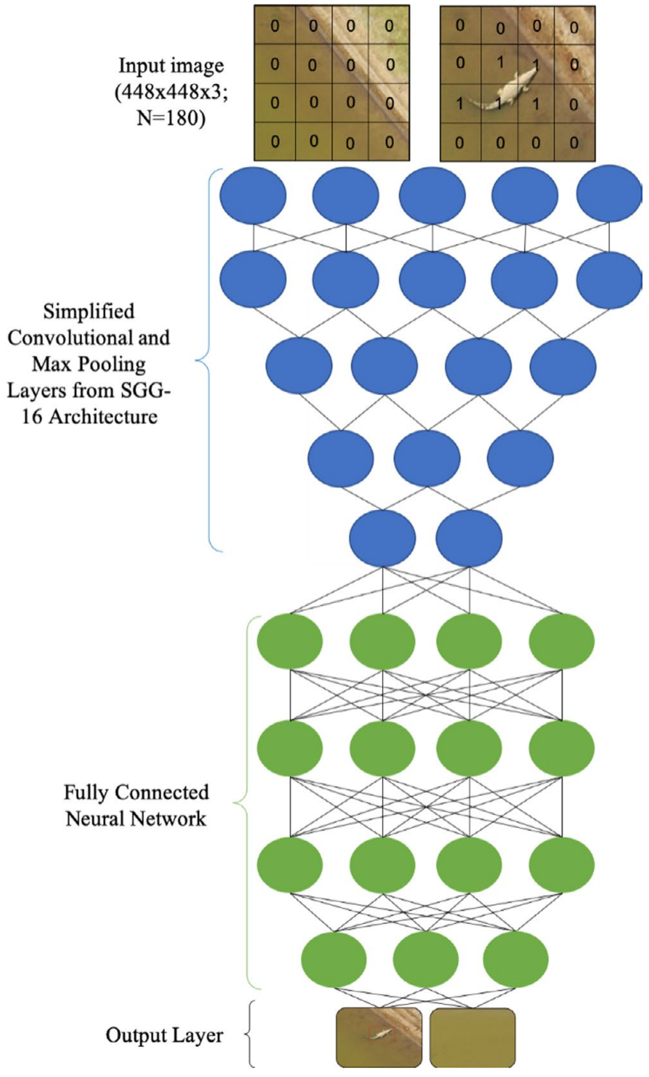

FIGURE A2 Overview of Convolutional Neural Network architecture. The input image is divided into 4 × 4 tensor. The convolutional layers perform the feature extraction for the CNN by scanning a few pixels at a time and creating a feature map. The max pooling layers reduce the amount of information, while 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