White paper

A Connected Conservation Collaboration

Can satellite imagery and artificial intelligence modernise wildlife surveys in Africa?

November 2021–December 2022. This study investigates the feasibility of using artificial intelligence (AI) techniques and satellite images in wildlife surveys.

Three approaches to detect and identify animals on satellite images are examined: the first approach uses only the human eye, the second uses computer vision and open-source (pre-trained) neural network models, and the third uses a combination of computer vision and a bespoke convolutional neural network model.

The project aimed to test all three approaches, with wildlife sightings reported by field teams around the time of satellite pass-over. Although the results show promise, significant challenges in using these methods remain. This paper validates their ongoing value, particularly in more heterogeneous environments, and gives suggestions for future research.

Authors

Sophie Maxwell,

Dimension Data Applications,

Leslie Polizoti

Acknowledgements

We thank the Connected Conservation Foundation (CCF) for leading, coordinating, evaluating, and fostering the collaborative partnerships that made this project possible and providing resources for conducting the study.

The Airbus Foundation donated satellite imagery from Pléiades Neo, the new 30cm-resolution optical constellation, as well as Pléiades 50cm-resolution images directly acquired from the satellite and from the Airbus archive. Microsoft funding and computational resources were made available through the AI for Earth programme. Northern Rangelands Trust (NRT) and Madikwe Futures Company NPC (Madikwe) staff provided field-team and ranger expertise to collect GPS wildlife-sighting data from the field for ground truthing and to validate results. NTT and Dimension Data engineers developed wildlife detection algorithms and a pipeline to process imagery.

We are grateful to WildCru for sharing information, and to the teams at NRT (Kieran Avery, Samson Kuraru, Kasaine Letiktik, Antony Wandera) and Madikwe (Leslie Polizoti, Koos Potgieter) for their participation in discovery workshops to explore use cases, conservation objectives and validate the feasibility of developing a solution.

This white paper focuses on one of many use cases that have been explored with other partners. Thanks to the efforts of the CCF team, research and support in this space are ongoing.

Abstract

As species numbers decline faster than ever before, there is an urgent need to improve our understanding of population numbers. For conservationists to protect species that are under pressure, it is necessary for them to understand current populations, including their welfare and behaviour. This study investigates the feasibility of using AI techniques and satellite images in ecology studies and wildlife surveys. Three approaches for the detection and identification of animals on satellite images are examined. The first approach uses only the human eye to find and classify different species using field-based expertise and insight. The second approach uses computer vision and open-source (pretrained) neural network models to detect and identify animals. The third approach uses a combination of computer vision and a bespoke convolutional neural network model, trained on a set of synthetic animal images, to detect and identify animals. The project aimed to test all three approaches, by evaluating the results of the algorithm detections and those of the human detections on the same satellite images alongside ground sightings reported by field teams around the time of satellite pass-over. Although the results show promise, challenges remain, as neither the algorithm detections nor visual human inspection of the same satellite imagery gave conservation managers the level of accuracy in species classification required for wildlife surveys. This paper describes the methods of the study and challenges presented, and provides suggestions for future research.

Madikwe
GAME RESERVE • SOUTH AFRICA

NRT
NORTHERN RANGELANDS TRUST

Contents

1. Introduction

There is growing interest within the conservation community about harnessing ever-increasing high-resolution satellite images[1] and AI technologies to monitor large animals (over 1m in size) across vast and hard-to-reach areas.

Numerous studies[2] have shown sub-meter satellite imagery and deep learning as a viable monitoring technique in constrained use cases with set conditions, especially over homogenous landscapes. These include the detection and counting of polar bears, penguins, yaks, elk, wildebeest, zebra, whales, albatross, elephants, and livestock. However, animal detection and identification on satellite images of heterogeneous environments has proven to be more challenging.

This study investigated the viability of expert human-eyeballing and AI methods on satellite imagery in animal identification and classification in both heterogenous and homogenous savannah landscapes.

Various challenges uncovered through this investigation include:

- The lack of known animal sightings in the imagery to function as labelled training data. To address this, the teams explored the use of synthetic and/or adapted images that were representative of real sightings to act as training data for the approaches considered.
- The complexity in dealing with heterogeneous environments – landscapes that have complex and non-uniform terrain, flora and lighting. The complexity of these environments uncovered unique challenges that oppose typical image processing methodologies. Complexity in an image is synonymous with 'noise' in the image or the variation of brightness or colour that make it extremely difficult to isolate one region of interest from the background. To combat this, the AI teams explored unique-image processing techniques.
- The difficulty of both human and AI model analysis to discern between different species of large mammals, in certain situations, in 30cm-resolution imagery. These situations include when different species are together in the same area (elephants/rhinos, wildebeest/buffalo) and when juvenile species are mixing with other adult species (for example, it is difficult to discern the difference between a juvenile elephant and an adult buffalo or rhino). Additionally, when species such as elephants' bunch close together, the shadows of several animals merge into one, resembling a 'blob', and it is difficult to count exact numbers.

This study explores the challenges present in using AI for satellite image classification, highlights learnings, and validates the ongoing value of expert eyeballing, particularly in more heterogeneous environments.

Multiple **aerial counts are made over a protected area or region,** with a high degree of variance in results. These **surveys can take months to complete** and are expensive in terms aircraft hire, fuel and human resources.

2. Problem statement

As species numbers decline faster than ever before, there is an urgent need to improve our understanding of population numbers, how species are faring and where they are moving, so conservationists can better protect them.

Wildlife surveys in savannah environments are typically done using aerial counts from small aircrafts. Multiple aerial counts are made over a protected area or region, with a high degree of variance in results. These surveys can take months to complete and are expensive in terms aircraft hire, fuel and human resources.

CCF, NTT, the Airbus Foundation, Microsoft, Madikwe Futures Company NPC and NRT sought to answer the question:

In what situations can AI and satellite imagery become a viable method for modernising wildlife surveys and what additional value can it bring to biodiversity conservation efforts?

1. A summary of related work is available in Section 7: Related work

2. A summary of related work is available in Section 7: Related work

3. Approach

High-resolution satellite imagery and AI have the potential to take a snapshot in time and show the number and location of endangered species across hard-to-reach areas, within a short timeframe, thereby reducing the time and cost of the survey.

This study set out to move beyond constrained use cases and understand the feasibility of using these techniques in the field to modernise wildlife surveys and create production solutions which can be applied in real conditions for ecology studies.

3.1. Satellite imagery

Certain conditions were considered when looking at the Airbus satellite images. Through Airbus archive data and new satellite tasking from the Airbus satellites Pléiades[3] and Pléiades Neo[4] , this study endeavoured to build training data and test the feasibility of the three detection approaches (as described in Section 4: Methods).

Satellite tasking is the process followed to request a satellite to capture a designated area of interest (AOI). Areas of high wildlife density was the criterion used to select areas of interest in this study. The two sites selected, as shown in Figure 1, were Madikwe Game Reserve in South Africa and Sera Wildlife Conservancy, managed by NRT, in Kenya.

Madikwe Game Reserve, South Africa, 750km^2

Sera Wildlife Sanctuary, Kenya, 100km^2

Figure 1: Satellite tasking areas of interest (map data: Google, Airbus/Maxar Technologies)

3. For more information on the Airbus Pléiades satellite, see https://www.intelligence-airbusds.com/en/8692-pleiades.

4. For more information on the Airbus Pléiades Neo satellite, see https://www.airbus.com/en/products-services/space/earth-observation/earth-observation-portfolio/pleiades-neo.

3.2. Taskings and captures

In support of well-coordinated taskings and optimum captures, the following tasking requirements were targeted:

- An incident angle of below 10°, for maximum image resolution and a top-down view of the animals.
- Weather conditions with low cloud coverage (<10%), for minimal obstruction for potential sighting.
- Images were panchromatic and orthorectified in natural colour, and pansharpened on four-band format processing through the Airbus pipeline.

In September 2021, satellite images of Madikwe Game Reserve and Sera Wildlife Sanctuary NRT were acquired from the Airbus Pléiades satellite (50cm resolution at 15° incidence angle). Sample imagery at the same resolution (with differing incidence angles) was also acquired from the Airbus archive. This was used to compare the landscape and set up a data-processing pipeline.

In late November 2021, the Airbus Pléiades Neo constellation was launched, with the option of twice daily 30cm resolution acquisition available globally. Pléiades Neo tasking capability became available to CCF in January 2022 via the Airbus OneAtlas[6] platform.

In February and March 2022, the project acquired high-resolution 30cm Pléiades Neo imagery of high-density wildlife areas in both reserves. Details of each capture are listed in Table 1.

Date of acquisition	Satellite	Location	Angle	Object of interest sightings
13/02/2022	Pléiades Neo	Sera (NRT)	2.7	100
13/02/2022	Pléiades Neo	Il Ngwesi (NRT)	13	2
02/03/2022	Pléiades Neo	Madikwe (North)	13	55
02/03/2022	Pléiades Neo	Madikwe (South)	18	50
28/03/2022	Pléiades Neo	Madikwe (North)	25	10 High cloud cover
03/09/2022	Pléiades Neo	Madikwe (North)	39	5
26/09/2021	Pléiades	Sera (NRT)	5	54
02/10/2021	Pléiades	Sera (NRT)	5	High cloud cover
28/08/2021	Pléiades	Madikwe	15	121

Table 1: Satellite capture data set

5. The incidence angle is the angle from the target point of view. It represents the angle between the ground normal and look direction from the satellite, combining the pitch and roll angles. Source: https://www.intelligence-airbusds.com/en/8719-angle-conversion#:~:text=The%20incidence%20angle%20is%20the,the%20instrument%20point%20of%20view.

6. For more information on the OneAtlas platform, see https://oneatlas.airbus.com/home

3.3. Capture learnings

3.3.1. Pass-over times

This study found that the time of satellite pass-over is a major constraint to performing wildlife surveys in areas near the equator. The capture times for these latitudes are consistently around midday, but African wildlife often venture out into the open in the early morning to feed and drink, then retreat to the shade to avoid the heat of the midday sun. This is the case in both wet and dry seasons. In NRT Kenya, traditional aerial census typically starts at 7.30am and is completed by 9am.

Pléiades Neo satellites are on an orbit path, which means images of Madikwe in South Africa can be acquired only between 10:00am and 11:00am SAST, and images of NRT conservancies in Kenya can be acquired between 11.30am and 12.30pm EAT. Whilst satellites maximise daylight pass-through, the lower latitude in South Africa and tasking that started an hour earlier allowed for better conditions for wildlife sightings and approach testing. The project analysis therefore continued with a focus on Madikwe Game Reserve, while sharing and reviewing results with NRT throughout.

3.3.2. Cloud cover

Another major hurdle in using satellite imagery to accurately identify and count animals is cloud cover. The Madikwe captures were tasked on optimal days, with excellent visibility and clear skies, with the image recording only 5% cloud cover.

A field team sighting identified 2 rhino at a known location, however on the satellite image, this sighting was concealed by the only cloud in the sky that morning. If this had been an annual wildlife survey, these two important sightings of critically endangered species would have been missed. The example in Image 1 has been included to show the risk cloud cover will pose if this method replaces traditional wildlife surveys. There is a high probability that some small cloud will obscure important sightings, when surveying such large areas even on a largely cloud free day.

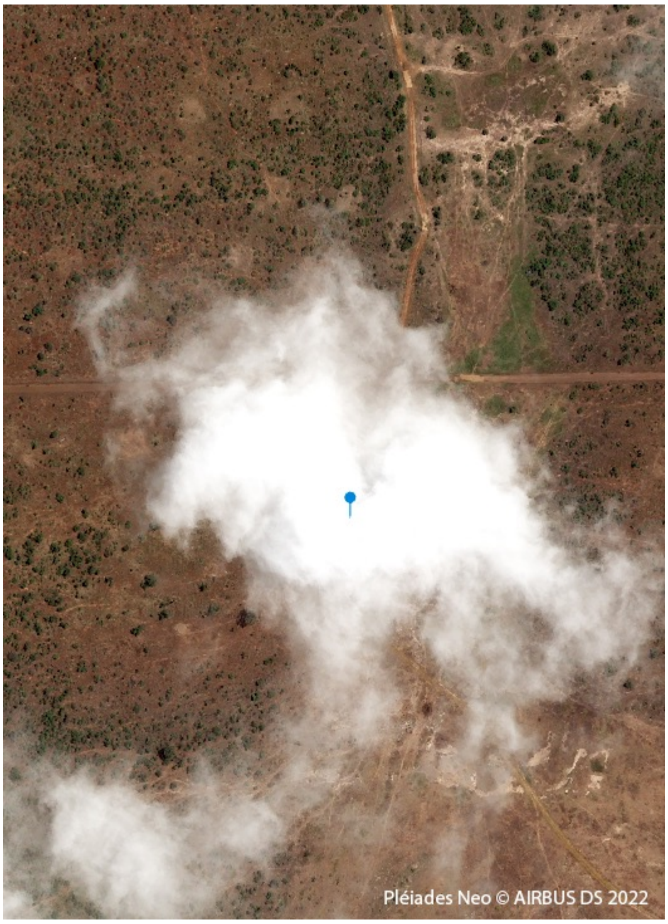


Image 1: Satellite image showing cloud cover, Madikwe Game Reserve

3.3.3. Shadows

Image 2: Satellite image of group of elephants under trees with shadows at a dam in Madikwe

Shadows in remotely sensed imagery are produced by objects such as clouds, trees and mountains. Shadows have the potential to impact the accuracy of information extraction. A challenge experienced in this study is when the shadows of groups of animals merge into one block, particularly when the animals are around or near vegetation or large features like rocks. Image 2 shows elephants clustered together at a dam in Madikwe. In this image, it is difficult to accurately count how many elephants are in the yellow box. Independent human detection of this image identified 5 elephants, but this is debated.

3.3.3. Incident angles

This study aimed to capture satellite images at the best resolution, with new acquisitions requested as close to the nadir line as possible, at the lowest incident angle. (Nadir line is the point on the Earth's surface directly below the satellite) However, at points during the study, images with a high incident angle were delivered instead.

Image 3: Pléiades Neo 30cm image at incident angle 25°, Madikwe Game Reserve

Image 3 shows the Pléiades Neo 30cm image captured at Madikwe Game Reserve at an incident angle of 25°. Through the clouds, the image shows 9 elephants in an open area, and with a more defined outline than that of the elephants shown Image 2, where the image was captured at a top-down angle below 15° degrees.

On a second occasion, again the satellite's priority was diverted, so the project received a Pléiades Neo 30cm image, shown in Image 4, with an incident angle of 39°. At the exact time of this capture, Madikwe's field teams sighted 5 elephants at these location coordinates.

Image 4: Pléiades Neo 30cm image at incident angle 39°, Madikwe Game Reserve

At 39°, the orthorectification resulted in considerable image distortion and lower resolution, making it difficult to spot the sighted elephants. Zooming in to the sighting location, it is exceedingly difficult for the human eye to identify the 5 elephants in the image without knowing the location coordinates (which were provided).

When comparing the above oblique images, it was noted that the definition of the elephants' outlines was clearer in the image taken on the nadir line, as shown in Image 3. Therefore, for future work (Section 6), it would be useful to explore a more optimum incident angle for spotting elephants, specifically, between 18°and 27°.

4. Methods

This study examined three approaches for the detection and identification of species on satellite images:

- The first approach, referred to as the Human Identification and Ground Truthing approach, uses local, field expert human-eyeballing of the satellite image and on-the-ground wildlife sightings from field teams.
- The second approach, referred to as the Spot and Label Object Detection approach, uses publicly available training data and three pre-trained models – RetinaNet, EfficientDet and YOLOv5 – to detect animal species on a selection of satellite images.
- The third approach, referred to as the Binary Large Object Detection and Classification approach, uses a set of synthetically created animal images which are pre-processed by computer vision models and then used to train a convolutional neural network (CNN) machine learning model.

The approaches were evaluated by comparing the second and third AI detection and identification results with those of the first expert human identification of animals on the same image. In addition, teams used best efforts to compare all the results with on-the-ground field sightings made at the time of the satellite image acquisition.

4.1. Manual detection methods

A powerful tool for analysing satellite data is specialised knowledge of the target landscape as well as the behaviours and characteristics of the objects of interest. A key aspect of the success of this study is field data validation and insight to reconcile and enhance detection models. Manual identification leverages conservation and field-team expertise by using humans with expertise to eyeball satellite imagery. Although it is incredibly time-consuming to scan satellite images, and exceptional attention to detail and insight are required for this process, it proved invaluable in probing the value and challenges of satellite data. This process, in combination with on-the-ground field sightings made at the time of satellite image acquisition, was used to deduce and interrogate species identification.

4.1.1. Approach 1: Human Identification and Ground Truthing

To coordinate this approach, two areas in Madikwe were targeted, totalling 200km2 of Madikwe Game Reserve. Both were areas known for wildlife density. The capture was scheduled for 10:33am South African Standard Time (SAST) on 2 March 2022, at an incident angle of <15°. On the same day, three teams collected field data on wildlife positions between 10:30am and 12:30pm.

A similar process was undertaken for Sera Wildlife Sanctuary and Il Ngwesi Sanctuary in collaboration with NRT's GIS and field teams. At 11:30am East Africa Time (EAT) on 13 February 2022, 30cm Pléiades Neo imagery was captured of 100km2 of Sera Wildlife Sanctuary at an incident angle of 1.64° with <5% cloud cover. Only a handful of target species were sighted at this time, so the focus moved to Madikwe.

As shown in Figure 2, field teams took different routes through Madikwe's tasked areas, specifically near water points where wildlife (particularly elephants and rhinos) is frequently seen, either at the water point itself or travelling to/from it.

During this tasking, 57 animals were identified on the ground and in the field during the satellite pass-over and were plotted by location (as seen in Table 2) with grid references to an area of approximately 500mx500m. Sightings were plotted in the middle of the grid, or where the vehicle spotting the animals was likely positioned when an identification was made. These are also shown in Table 2.

The most relevant sightings were between 10:44am and 10:59am, with many animals also spotted around 11:30am to 11:40am. Given their speed of travel over an hour, these herds were expected to be well within the tasking area at 10:33am, which was when the satellite would pass, and were expected to be seen on the satellite imagery.

Figure 2: Planned field team routes, Madikwe Game Reserve (map data: Google, Airbus/ Maxar Technologies)

Species	Number	Time (SAST)	Location
Elephant	9	10:44am	Location 1
Giraffe	6	10:59am	Location 2
Giraffe	3	11:30am	Location 3
Elephant	9	11:30am	Location 3
Kudu	2	11:30am	Location 3
Elephant	4	11:34am	Location 4
Elephant	15	11:38am	Location 5
Elephant	9	12:31pm	Location 6

Table 2: Extract of Field sighting recordings, Madikwe Game Reserve

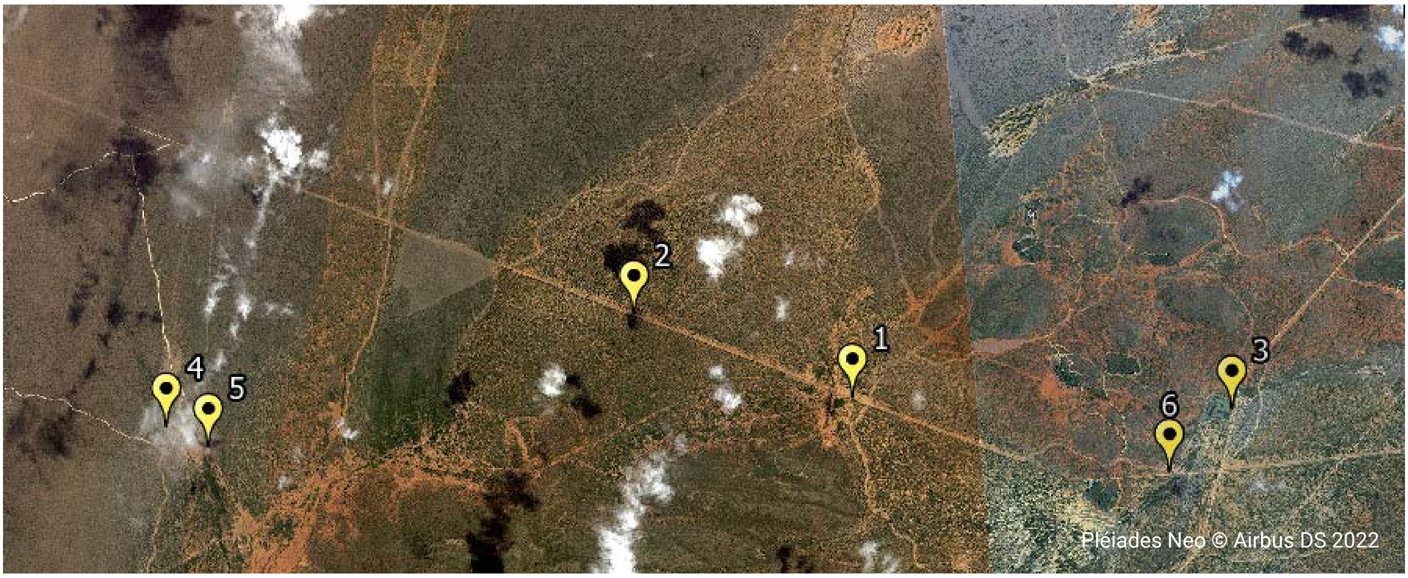


Figure 3: Field sightings locations, Madikwe Game Reserve

Figure 4 shows extracts from the satellite image shown in

Figure 3 indicate sightings of elephants. Also included is one cropped image of the southern tasking area that captured and sighted rhinos during this exercise. All exported raster images were taken at a 1:100 scale. Using field data, it was concluded that the bottom left image contains only rhinos, and the animals in the rest of the image are elephants. This example indicates how difficult it is to differentiate between rhinos and elephants in 30cm imagery that is taken close to the nadir line.

Figure 4: Animal sightings – extracts from images, Madikwe Game Reserve

Conservation experts with substantial field experience then manually scanned the satellite imagery with the human eye, identifying 110 large mammals across the 200km2 area of interest. Conservation experts used their in-depth knowledge of the reserve and the behaviours of different species groups to record their understanding of the sightings, assigning a confidence value (%) to the certainty of their identification. Examples of conservationist and field knowledge enhancing this process include the knowledge of species moving in herds, movement along known game paths, standing in ways indicative of specific species groups, and the known locations frequented by of popular species.

Here are two further examples showing the challenges in identifying elephants and determining accurate elephant numbers in 30cm satellite imagery with the human eye.

Example 1

At 10.44am, 10 minutes after satellite image capture, field teams sighted a herd of 9 elephants at location 1.When the imagery was later analysed by local conservation experts, who know the terrain and animals, they looked for 9 elephants in the vicinity. The red pins in Image 5: Zoom 1:983, Madikwe Game Reserve mark the potential location of elephants identified using just the human eye. In this image, some elephants are out in the open by the road, others are in the bushes near the pan. From this zoom level (1:983), the naked eye can tell that most of the pins mark animals. If there is some scale – for example, if the naked eye has been calibrated by looking at 30cm satellite imagery – and some on-the-ground knowledge of the terrain, one can determine that the animals are elephants, because they are larger than the other animals that appear in other areas of the satellite imagery. Some knowledge of elephant behaviour – that is, that there are often in herds in this location – supports this conclusion.

Image 5: Zoom 1:983, Madikwe Game Reserve

Image 6 is the upper half of Image 5, zoomed to 1:629. The

elephants marked with the red pins are in relatively open terrain. These elephants have a fairly clear elephant shape from the top-down: a roundish head with a larger, oblong body.

The elephant marked with the blue pin is a bit trickier. Although it does not have exactly the same shape as the elephants with the red pins, the shape is similar enough and the size is consistent with the size of the other elephants. With knowledge of the terrain, there is nothing else that it can be: there is no physical object like a rock or bush in that location.

Now look at where the blue and pink arrows are pointing, towards the bushes. It seems like there might be other elephant(s) next to the bush. It is more difficult to tell from the image only because the bush is partially obscuring the shape of the elephant. Determining that these are not bushes or rocks requires on-the-ground confirmation.

Image 6: Zoom 1:629, Madikwe Game Reserve

Zooming in further does not help – in fact, it makes the image worse. Image 5 is the part of Image 4 with the blue-pinned elephant zoomed in to 1:165. Even at 30cm, the resolution is insufficient to confirm what the shape is.

Image 7: Zoom 1:165, Madikwe Game Reserve

Example 2

A second example that illustrates the challenges in identifying and counting elephants in 30cm imagery relates to the size of elephants and various wildlife. Image 8 is at 1:629. The naked eye can see two objects inside the red circle. The two objects look much smaller than the surrounding elephants, and it would not be typical elephant behaviour for two young elephants to be away from their mothers, let alone separated from the herd by bushes. We can conclude that these two objects are likely not elephants, but affirmatively identifying what they are is not possible from the image.

Image 6: Zoom 1:629, Madikwe Game Reserve

Adult elephants are the largest animals in the study areas, and there are numerous challenges to identifying them and counting them in heterogeneous environments using 30cm satellite imagery, as shown above. These challenges are even greater for animals that are smaller than adult elephants: young elephants, buffalo, rhino and plains game like wildebeest, zebra, impala, and so on. It is also difficult to differentiate between different species when game is mixed, for example when congregating around watering holes.

4.2. Artificial intelligence detection methods

To create a species-detection model, labelled imagery of the objects of interest is needed to train an intelligent model. While there is availability of free and open Earth observation data at the medium to low spatial resolution, image licence agreements currently restrict the sharing of high-resolution imagery. This limits the ability to build a central repository of training data that can be shared by independent projects.

Additionally, manually screening satellite data for species-labelling purposes is not a viable method of building large training data sets. This is because volumes of aerial perspectives of savannah animals across available Airbus high-resolution imagery are low.

Objects within these satellite images can often be too small or obscure for the human eye to notice, and sometimes too small for methodologies to accurately identify (as noted in Section 4.1).

Working with satellite data also creates unique challenges namely:

- Imagery is not easily accessible because the scale of the image is too large.
- Satellite images have several features which prove challenging when it comes to classifying objects (incident angles, shadows, clouds, noise, lighting, radiation, atmosphere).
- There are limited verification methods to confirm the accuracy of object classification.
- Changing terrain and background conditions as well as insufficient resolution of satellite images are additional obstacles.

Noting these challenges, the team set about exploring alternative methods to create synthetic training data and test sets.

After an in-depth discovery phase, separate teams defined two AI methodologies to address the problem statement at hand: Spot and Label Object Detection and Binary Large Object Detection and Classification.

Creating synthetic training data and using AI techniques, the team from NTT and Dimension Data addressed the following research questions:

- Can we use synthetic data to train models?
- Can we detect and classify large animals (over 1m in size) in homogeneous environments?
- Can we detect and classify these animals in heterogeneous environments?
- Can we improve the degree of accuracy of detection and identification by augmenting inference with additional knowledge?

4.2.1. Synthetic image test sets

Because training imagery was lacking, both approaches explored the use of synthetic test sets inserted on satellite images (referred to throughout this paper as 'stitching'). Synthetic image stitching sought to create a set of labelled imagery using synthetic species samples stitched into the original satellite image.

The synthetic test set comprised a mix of 10,000 species targets. Some images were synthetic, from animal figures, some were adapted from drone or close-range aerial views, and some were computer-generated. Using these different sources of imagery of animals, the samples were scaled to match the required satellite resolution. The scaled-down samples were then inserted into the satellite imagery on a pixel-by-pixel basis. Using this artificial set of labelled data, the model was trained to identify and classify potential objects of interest (which varied in size, colour and shadow variation).

Down sampling of other aerial data followed a similar approach but focused on using drone and aerial data that was downscaled and masked to match the resolution of the relevant satellite data. The sample dataset focused on synthetic variations of size, colour and shadow.

4.2.2 Approach 2: Spot and Label Object Detection

The object detectors proposed in this approach look at supervised learning models with labelled datasets. The approach focused on a single species for validation. The rhino was selected for this approach, based on the novelty of the problem statement for rhino conservation and the limited exploration of rhino modelling using satellite data.

4.2.2.1. Methodology

Three state-of-the art object detector models were identified as being best suited to rhino classification: RetinaNet (using the ResNet50 backbone), EfficientDet (EfficientNet backbone) and YOLOv5 (using CSPDarkNet and a PANet neck). The models were selected for their varying levels of success when being applied to other datasets.

Research in support of this approach is a CowNet study (Robinson et al., 2021) in which semi-consistent results were achieved in localising and detecting large hooved animals (cows and elks) in open fields on 30cm:1pix images. The study had over 10,000 labelled animals in its dataset, making it conducive to model training.

For this study, metrics of interest were precision and recall. A metric that would be a good indicator of model performance was detections on unreferenced data (i.e., an image with no confirmed rhinos). The intent was to evaluate model performance by testing the number of positive rhino identifications.

7. Ray Harris and Ingo Baumann. 2015. Open data policies and satellite Earth observation. Space Policy, Volume 32, pages 44–53. Available from: https://www.sciencedirect.com/science/article/abs/pii/s0265964615000028

4.2.2.2. Synthetic training data preparation

For this study, no satellite training data was available for rhinos. Object detection approaches require an incredibly large amount of training data to ensure that the model can reasonably learn differentiating features.

To address this species-specific challenge, a synthetic training data set was explored. This involved cropping rhino images taken from aerial imagery of varying resolution into bounding boxes. Each bounding box was a masked rhino, which was then downscaled and resized to the resolution of the satellite sub-image in the training set. For 50cm, the longest side of the bounding box was mapped to 9 pixels (i.e., 4.5m on the ground). The shorter side was mapped to nine*boxMinLength/boxMaxLength (which ranges from 4 to 7 pixels). For 30cm, the longest side was mapped to 15 pixels (i.e., 4.5m on the ground).

The key challenges for both 50cm and 30cm satellite data were the consistency in the conversion quality of 16-bit source images to 8-bit images required by the model, the handling of large image sizes and the amount of signal in the image due to splitting and upscaling (50cm images were split at 96x96 pixels, then upscaled to 384x384 pixels; 30cm imagery was cut at 768x768 pixels and downscaled to 384x384 pixels).

A hold-out method was applied by splitting the dataset into a 'train' set and a 'test' set, with precision and recall metrics evaluated for model performance. When running the model on the held-out dataset, it was expected there would be a high number of true positive detections and a lower number of false positives and negatives. False negatives are concerning, as they mean the detector is missing positive rhino sightings. However, the lack of labelled training datasets posed two issues with this approach: there was not enough training data or validation data to measure the accuracy of the model's performance.

4.2.2.3. Model technical description

The RetinaNet is a 1-stage detector with two subnetworks for classification and regression. It uses focal loss to address the imbalance between foreground and background. Examples such as backgrounds are down-weighed, which enables the model to focus training on the harder examples. One of the RetinaNet backbone configurations is the resnet_50_fpn. The intermittent assimilation of the identity function in the Resnet design makes the network capable of going deeper, thus increasing the capacity for learning. The Feature Pyramid Network (FPN) design connects various stages in the architecture, making it capable of detecting objects at different scales. These variations enable the RetinaNet/Resnet_50/FPN combination model to be both accurate and fast.

The 'You Only Look Once' (YOLO) is a set that performs detection in a single stage. The model does not utilise Region Proposal Networks (RPN); instead, it divides the image into grid boxes and each grid box is responsible for detecting species within its borders. Information is processed as a regression problem for both the bounding boxes and the class probabilities. The model preserves spatial relationships and can encode and recognise contextual information. One of the backbone configurations available is 'small,' i.e., YOLOv5s. Being small, it has the advantage of fast processing speed.

EfficientDet is a 1-stage design model; it uses multiple bidirectional FPNs to refine the scalability of features for depth, width and resolution.

4.2.2.4. Training, testing and model results

Models were initially trained on 50cm Madikwe images. Figure 5 shows the comparison of model performance results. It is noted that, due to the complexity of the image background, many detections were in fact not rhinos but consistent misclassifications of other objects of interest.

This exercise was of limited value to addressing the problem statement; however, it confirmed the assumption of YOLOv5 being a preferred model for the classic approach to explore further satellite scenarios (preferably at a higher resolution with larger test datasets).

RetinaNet | Models: 1 | Epochs: 50 | 50cm images | Fixed placement | Colour-balanced |

- Detections
  - True Positives: 1410
  - False Positives: 795
  - False Negatives: 463
- F2 Score: 0.727
- Recall: 0.7528
- Precision: 0.6395

EffDet | Models: 1 | Epochs: 50 | 50cm images | Fixed placement | Colour-balanced |

- Detections
  - True Positives: 871
  - False Positives: 21
  - False Negatives: 1002
- F2 Score: 0.5194
- Recall: 0.4650
- Precision: 0.9765

YoloV5 | Models: 1 | Epochs: 50 | 50cm images | Fixed placement | Colour-balanced |

- Detections
  - True Positives: 1692
  - False Positives: 59
  - False Negatives: 181
- F2 Score: 0.9153
- Recall: 0.9034
- Precision: 0.9663

Figure 5: 50cm synthetic model results for Spot and Label Object Detection model

The YOLOv5 model was evaluated on 30cm imagery. Again, the results show a low detection threshold (0.2) with a high number of false positives. Figure 6 is an example of this test result, with 47 true positives (based on synthetic stitching), 905 false positives and 279 false negatives (F16 Score: 0.14310254952137239; recall: 0.1441717791411043).

The results for 30cm were still unsatisfactory because of the high number or false positives provided by the model. However, the spread of false positives compared to the synthetic ground truth indicates that a heatmap of potential detections, with custom model refinement and a larger labelled training dataset, may be a viable option. (See section 6 Future work)

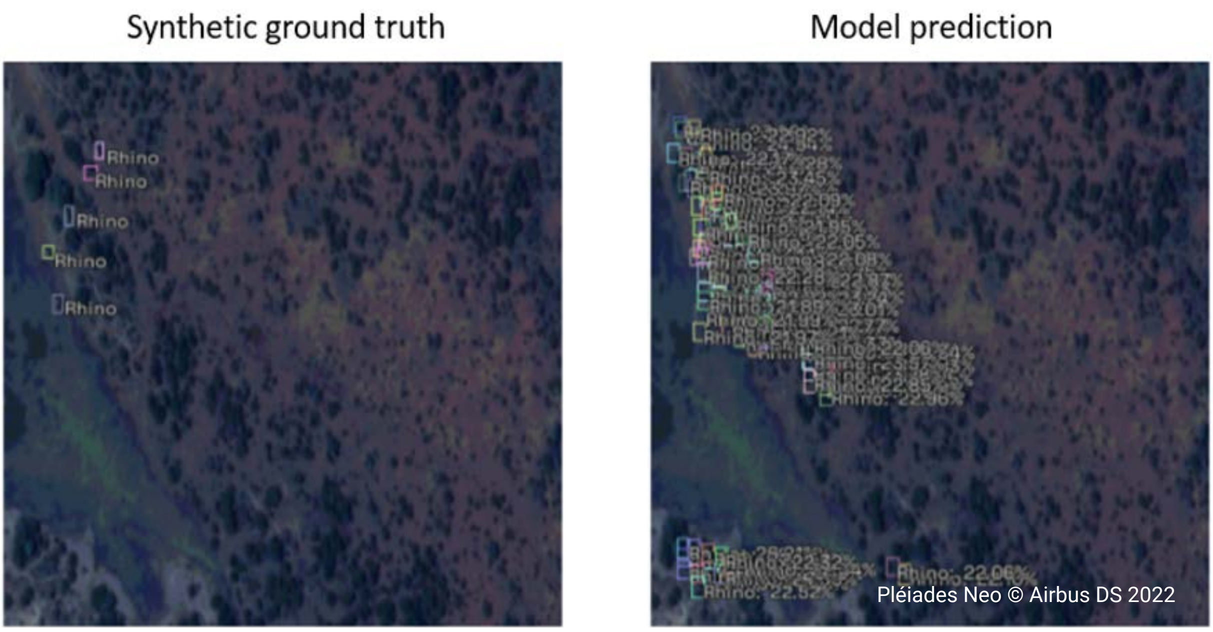


Figure 7: Process flow for Binary Large Object Detection and Classification approach

Based on the low and inconsistent performance of industry-standard supervised learning models with limited ground-truth training data, this approach was deemed insufficient in terms of field-team accuracy for wildlife monitoring.

Based on the low and inconsistent performance of industry-standard supervised learning models with limited ground-truth training data, this approach was deemed insufficient in terms of field-team accuracy for wildlife monitoring. However a more refined heatmapping approach of findings, as depicted in Figure 6, remains a candidate for future exploration.

For further research and learning, the codebase for Approach 2: Spot and Label Object Detection is available at the Connected Conservation Foundation GitHub. The repo includes the Spot and Label Object Detection approach pipeline: training, extracting animal sightings, creating synthetic data, splitting the satellite images for inference and, most valuably, the heatmap.

[Access the codebase for Approach 2: Spot and Label Object Detection]

Please email [info@connectedconservation.foundation] if you'd like to further research using these source materials

4.2.3. Approach 3: Binary Large Object Detection and Classification

This approach sought to create a set of labelled imagery using synthetic 3D animal model samples stitched into the satellite imagery. Using this artificial set of labelled data, a detection model was trained to identify and classify potential objects of interest and a process was devised to run a new satellite image through a species detector.

The actual species-detection aspect was split into two events: object detection and object classification. Object detection uses computer vision techniques to identify regions of interest, while object classification makes use of a bespoke convolution neural network model to perform predictions.

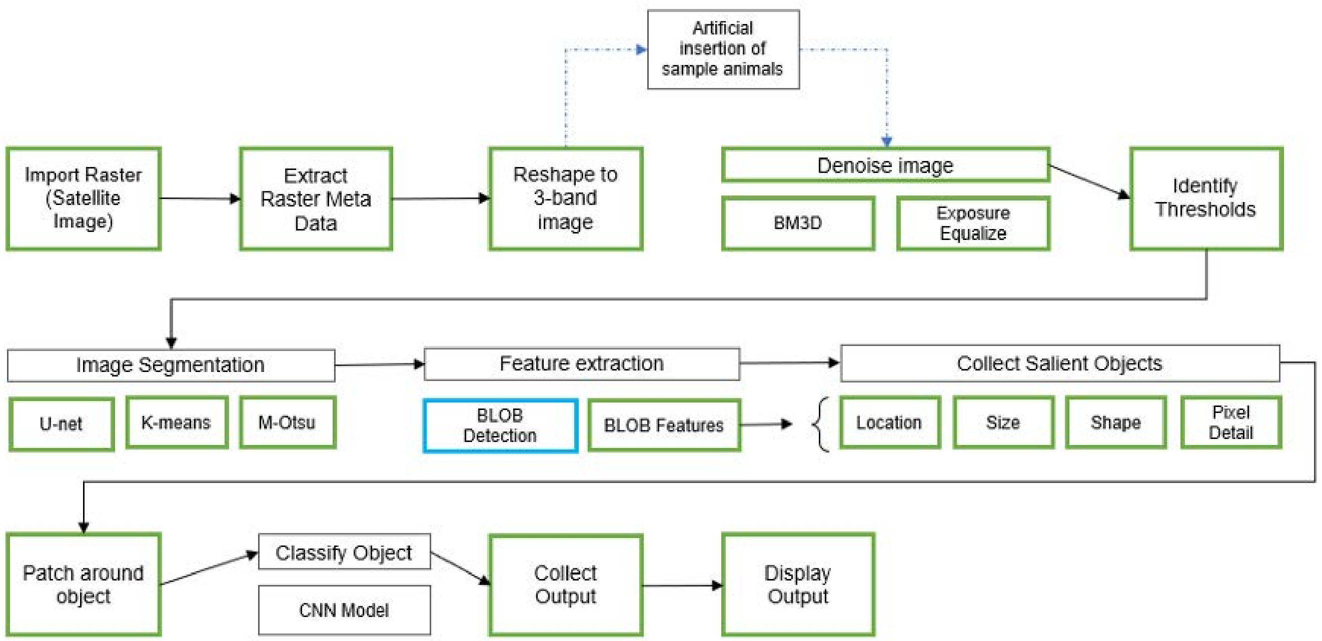


Figure 7: Process flow for Binary Large Object Detection and Classification approach

Object detection takes an input and tries to isolate regions of interest. If a region of interest is identified, the region is passed to the classification model for a prediction on what class (animal) the object may be, along with a likelihood score. The pixel coordinates of every detection as well as the classification of the animal are recorded and converted into map coordinates that can be used to find and identify animals.

The pixel coordinates of every detection as well as the classification of the animal are recorded and converted into map coordinates that can be used to find and identify animals.

4.2.3.1. Methodology

A supervised learning approach, CNN, was selected to detect and identify species on satellite images. With this approach, images are mapped to a set of specified labels. The convolutional neural network learns the visual features contained in the training images associated with each label and classifies unlabelled images.

The images passed to the model were at 30cm resolution, with a height and width of 18cm. These images were split into a training and validation set: 80% of the images were used as training images and the remaining 20% were used for validation. All these images had labels linked to them. These labels corresponded to the following classes on which the prediction could be made:

- Elephant
- Elephant carcass
- Giraffe
- Hippo
- Wildebeest
- Hyena
- Hyena carcass
- Rhino
- Rhino carcass

4.2.3.2. Synthetic training data preparation

As detailed in Section 4.2.1, the creation of synthetic data is vital to training a CNN model for animal classification. For this approach, the team used open-sourced computer-generated imagery of the animals in question to create samples. Samples are seen top-down, in 3D, and the outline and size of the animals, as seen from the nadir line. The samples were downscaled by a factor, using inter-area interpolation – a method of interpolating pixel values when downscaling to preserve as much of the original information as possible. The images were downscaled to match the satellite imagery resolution.

It was also important that all backgrounds of the sample images were created with an empty (zero) pixel value, as this ensures that only the animal is inserted into the satellite image, and no background of the sample image is included.

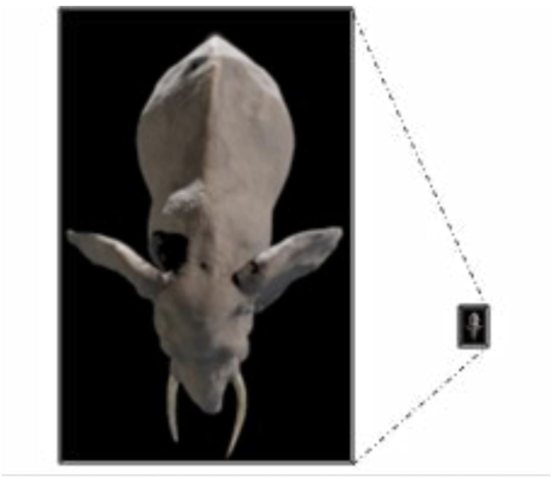


Figure 8: Downscaled elephant sample (to actual size)

For each animal class, more training data can be created by rotating the original sample (through 360 degrees) as well as capturing colour variations. By augmenting the dataset in this way, many sample images were generated for training. A single sample was repeated for each animal class, with a focus on top-down views with minimal lighting variation.

To create a training set, labelled animals on satellite imagery were needed. This was achieved by inserting each animal sample into a known location on the satellite image, on a pixel-by-pixel basis. The algorithm selected a random location on the satellite image and ingested a single animal sample. The 'stitching' of the animal sample into the satellite image was done by looking at the pixel values of the sample. At the location selected in the satellite image, the algorithm inserted each pixel from the sample that was not zero into the satellite image. If the sample pixel value was zero, either the satellite background or the sample pixel value was kept, thereby ensuring a perfect insertion of only the animal into the satellite.

With defined samples at known locations, training data could be extracted from the now augmented satellite image.

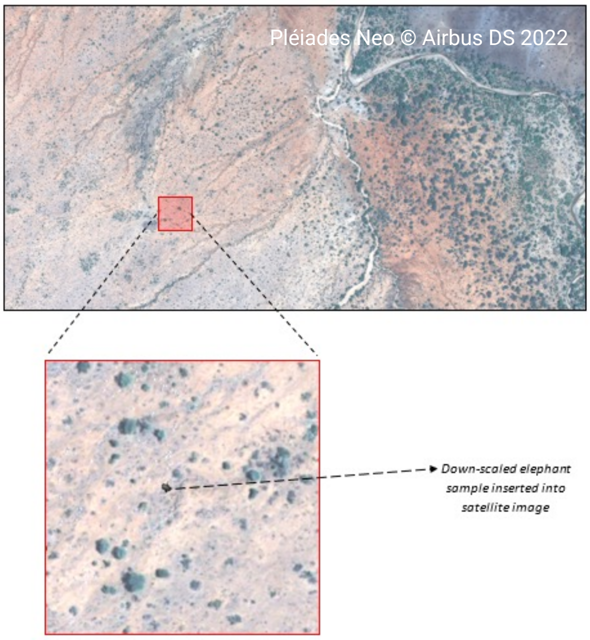


Figure 9: Subsection of satellite image with elephant sample inserted

4.2.3.3. Model architecture

The convolutional neural network utilised for this use case consisted of six convolutional layers with an input shape of 18x18x3. The third dimension in the input shape refers to the red, green and blue colour channel. There was also a maximum pooling layer: a pooling operation that calculates the largest value in each patch of each feature map. This resulted in a down-sampled or pooled feature map that highlighted the most present feature in the patch.

This was followed by a dropout layer, which was used to prevent the model from overfitting. Dropout works by randomly setting the outgoing edges of hidden units (neurons that make up hidden layers) to zero at each update of the training phase.

The SoftMax function was used as the activation function in the output layer of the convolutional neural network to predict the multinomial probability distribution. The SoftMax activation function is used for multi-class classification problems where class membership is required on more than two class labels. In this use case, there were nine possible classes, each representing a different species.

4.2.3.4. Training model

The model was trained through an 80/20 split: 80% of the available training data was used to train the convolutional neural network while 20% was used to test validation accuracy. The model was trained on the above animal classes and can easily be adjusted, depending on the sample images provided. The team aimed to provide training images that covered small (hyena), medium (wildebeest) and large (rhino, elephant) animal sizes.

The convolutional neural network was trained over 10 epochs with a batch size of 32, meaning that the machine learning algorithm would pass over the entire training dataset 10 times, working through samples of size of 32 each time, before updating the internal model parameters.

Figure 10 shows the training and validation accuracy plot of the model: while validation accuracy stops improving and flattens out at around 10 epochs, training accuracy will continue to improve beyond 10 epochs as the model seeks to find the best fit for the training data. The plot indicates that the model was not overfitting over 10 epochs.

This was also evident in the training and validation loss plot, where the validation loss is on a steady downward path throughout the 10 epochs without any sudden increase in loss.

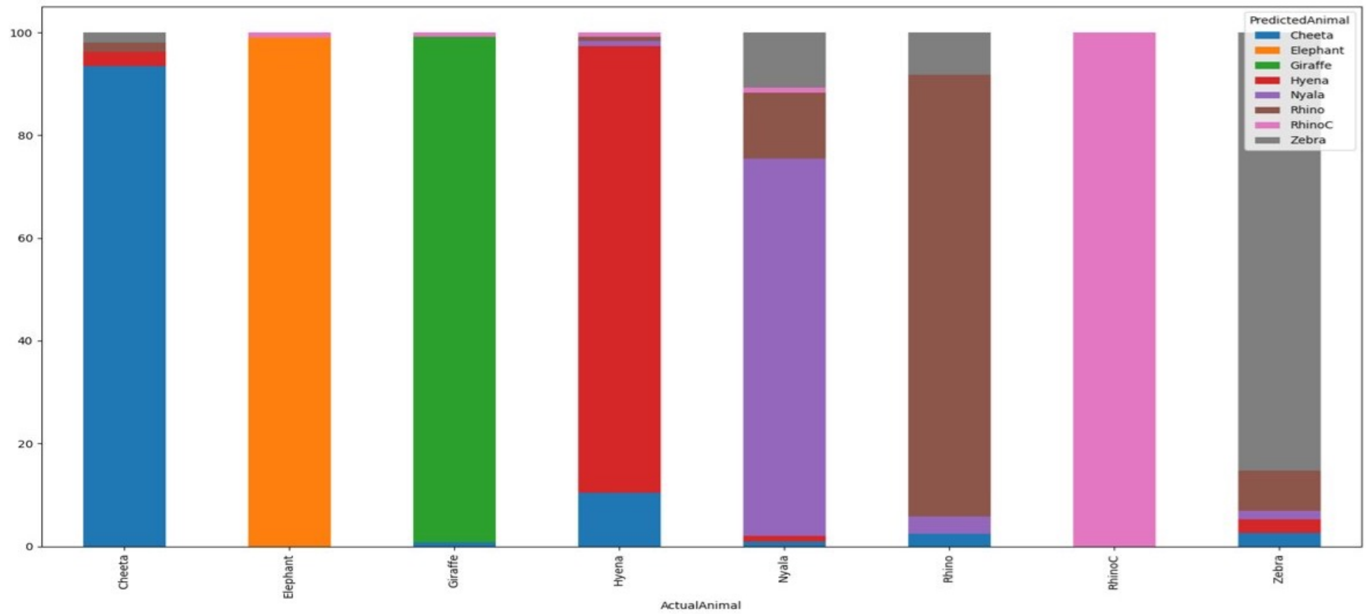


Figure 10: Convolutional neural network model training accuracy

4.2.3.5. Testing and model results

The current iteration of the convolutional neural network was tested on an unseen, augmented synthetic dataset consisting of the animals depicted in Figure 10. The synthetic animals were stitched into satellite imagery background and passed into the model for prediction.

The X-axis shows the actual synthetic animals that were fed into the model and the colour corresponding to the key represents the percentage of the respective animal predicted.

The model's accuracy was tested in various ways. Although the training of the model itself is a supervised machine learning problem, the entire solution scope had to be treated as unsupervised, as there was no way to accurately validate the prediction labels at the current satellite image resolution. In terms of defining the accuracy of the model, validation accuracy was used to test the model against unseen images for testing.

This was sufficient for the trained model to be used as a classifier in the overall solution.

4.2.3.5.1. Computer vision

Before any image is sent for classification, a region of interest must be identified. For this model, a region of interest was described as a 16-pixel-by-16-pixel region that contains a potential Binary Large Object (BLOB). A BLOB is a feature in the region that possesses qualities such as connected pixel mass, consistent pixel values and circularity, among others. Only if the region contains a BLOB, is it sent for classification.

Computer vision techniques are key in the BLOB detection process. The two main steps of the process are described below.

4.2.3.5.1.1. Image patching

This is the process of iterating through the image, splitting the satellite image into blocks of 128 pixels by 128 pixels. The blocks overlap by 30% to ensure no potential BLOB is lost at the image border. The process of splitting the satellite image is done by recursively reducing the satellite image matrix to the

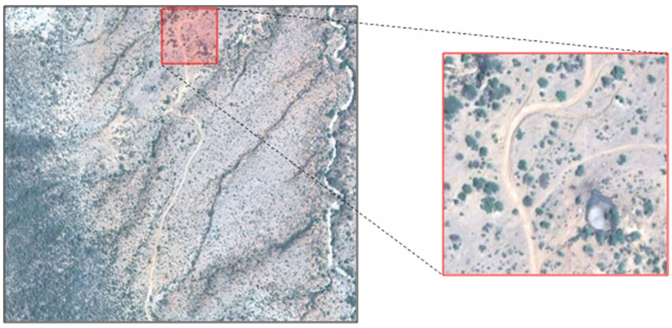


Figure 11: Representation of an individual image 'patch'

bounds defined by the sub-image.

Each image patch is then passed for further processing according to the rest of the algorithm.

4.2.3.5.1.2. Image processing

Each image undergoes a series of processing steps (as described earlier in this paper). Several kernel convolution methods (such as Gaussian and 2D filtering) are applied to each image patch. The purpose of this is to smooth the image, eliminating the large amount of noise present in the landscape.

The goal of the image processing step is to eliminate unwanted image noise and to further isolate and highlight regions of interest. The image groups in Figure 12 show the histogram

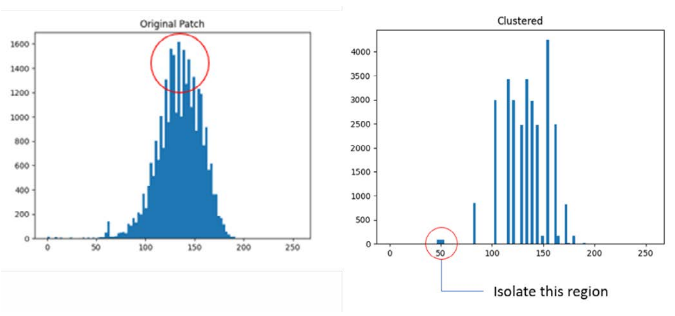


Figure 12: Results of image processing

of pixel values of a random image patch containing a known animal sample before and after image processing has been applied.

Morphological operations such as erosion and dilation were performed using kernel convolution. These operations have the goal of joining partially connected pixel groups while separating sparsely connected pixel groups. The result emphasises differentiation between objects on an image.

Image clustering was then performed to map similar areas to a single pixel value. An example of this would be for all trees in an area (with varying shades of green) to be mapped to a single shade of green. This is done to emphasise any objects not conforming to the background landscape, in other words, a potential animal. Once this object (if any) is emphasised, it can be isolated using BLOB detection. This is done by isolating a collection of pixels determined to be an object by a set of predetermined criteria such as circularity, pixel connectedness and size.

Once a BLOB is detected, the CV process is complete, and the sample is passed for classification.

The team building the CNN detection process had some positive results on various imagery using synthetic images. When the model was run on artificially inserted samples (such as the image group shown in Figure 13), results were consistent over largely heterogeneous environments. The number of positive sightings decreases over more complex terrain.

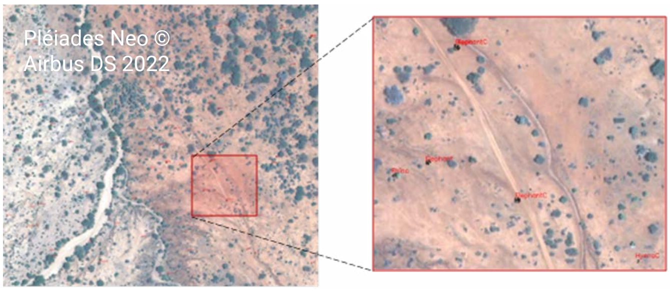


Figure 13: Classification output from artificially inserted samples

The animal classification process is heavily dependent on the input sample from the animal-detection process. While the performance of the animal-detection process is robust, the ability of the current model to isolate only the binary large object in question was limited, and this had a knock-on performance effect on the classification of this sample. Again, running the process in more complex environments compounds the isolation issue.

4.2.3.6. Results

The image detection (the ability of the model to detect Binary Large Objects) process was tested over actual sightings of large mammals in 50cm and 30cm imagery with some success in homogenous imagery and inconsistent results in heterogeneous data. Examples of the results are detailed below.

As depicted in Image 9, the convolutional neural network model predicted 5 elephants. When validating the findings with field teams who are familiar with the area and herd behaviour, it was presumed that the model prediction was accurate, based on the relative size, number and location of elephants. The animals also appear to be walking in a line on a game path with one or two off the path, likely eating. However, without on-the-ground field confirmation at the time of the pass-over, it is not possible to accurately determine if this is correct.

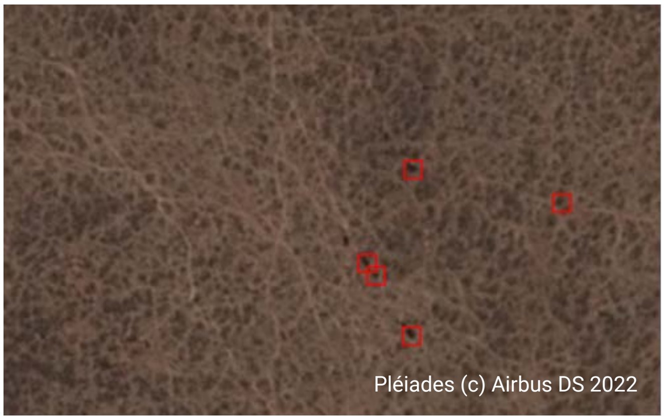


Image 9: Pléiades 50cm image, Madikwe Game Reserve

As shown in Image 10, the model detected small animals with low levels of accuracy. The model predicted 18 animals while human-eyeballing indicated over 150 small animals, with no verifiable ground-truth comparison. Attempts to refine the model to count and/or classify individual small species remain a challenge. The output is another viable candidate for heatmapping species or using models as an indicator of where to look for small, medium and large species.

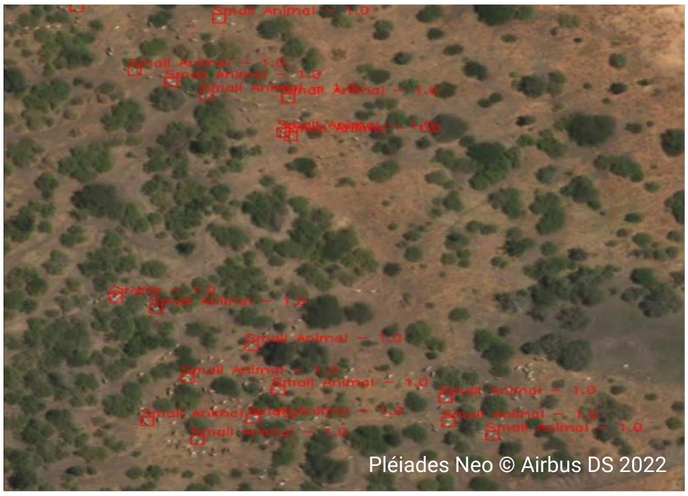


Image 10: Pléiades Neo 30cm image, Sera Wildlife Sanctuary, NRT

Another good test scenario occurred during a capture at a popular waterhole, as shown in Image 11. The model predicted 55 objects of interest in this image, with 4 elephant detections and 51 null counts (i.e., unable to classify species). Expert eyeball classification noted 23 elephants. The ground-truth counts at time of the satellite pass-over indicated 21 elephants. Based on field-team ground validation, the rate of false positives in this image is extremely high, making the model an inconsistent tool for accurate species modelling at scale.

Image 11: Pléiades 30cm image, Madikwe Game Reserve

4.2.3.7. Model output

There was some success with this model, in that it can accurately iterate through the satellite image and classify the artificially inserted animal samples. However, when tested against real wildlife sightings in 30cm satellite data, findings were inconsistent. Overall model performance increased significantly in homogeneous landscapes compared to more heterogeneous (bushy, trees etc.) landscapes. However, for this study, homogenous landscapes with target species present were limited at the identified test sites.

The accuracy of the model tested against unseen labelled imagery – that is, the validation accuracy of the convolutional neural network – was 99.1%. This is a reflection on the model's ability to learn based on the training data it was provided with, not a true reflection of the entire approach. Therefore, if more realistic labelled training data becomes available in the future, this model is a viable candidate for further exploration.

The predictive performance of the entire solution is dependent on three factors (as seen in Figure 14):

- **Model training: how accurately the synthetic training data represents the real-world targets (animals):** In size, shape and lighting, the training data accurately reflects real-world conditions, but might not capture varying lighting conditions in the background environment (shadows, over-exposure, etc.) as accurately.
- **Object detection: how well the computer vision (object detection) aspect performs:** The object detection process can work well in comprehensively defined circumstances. That is, the object detection performs well when it is configured for the landscape it is processing. Satellite imagery inherently possesses a large amount of variation, from incident angle and exposure to shadows and changing landscape seasonality. The scale of variation is extremely difficult to capture dynamically.
- **Object classification: how well the neural network is performing (object classification):** Assuming a potential object of interest is correctly identified and isolated, it is found that the classification algorithm is able to classify the correct size category of species targets.

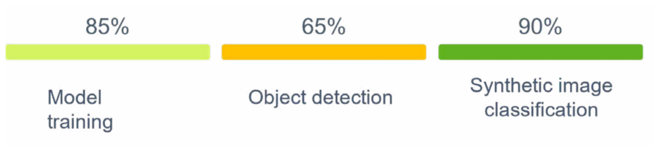


Figure 14: Predictive performance of the Binary Large Object Detection and Classification model

The methodology itself holds promise: when the computer vision and computational intelligence components have been combined, the model can process satellite imagery of a massive area of land (over 400 million pixels) in an iterative manner.

The computer vision component can identify objects in a satellite image of any size and pass the area of interest over to the convolutional neural network to make a prediction. The procedural ability alone – to handle and process such large imagery with a confident degree of accuracy – proves that the method has viability for surveying species in large homogenous land areas. As it is not feasible to validate the entire solution against absolute ground-truth data at the scale required, any performance measure is derived from observable factors from subsets of the satellite capture and not the solution holistically. The recommendation is therefore for the model to be trained and tested on a greater set of satellite data that contains confirmed labelled species to improve and test overall model performance.

For further research and learning, the codebase for Approach 3: Binary Large Object Detection and Classification is available the Connected Conservation Foundation GitHub.

[Access the codebase for Approach 3: Binary Large Object Detection and Classification]

Please email [info@connectedconservation.foundation] if you'd like to further conduct research using these source materials

5. Conclusion

This field-based study set out to compare the accuracy of different methods of identifying large mammals in the highest-resolution satellite imagery available from Airbus (from the Pléiades and Pléiades Neo satellites). These methods included detections made by two AI approaches and the human eye, which were compared with on-the-ground field sightings for a small set of locations made at the same time as the satellite image acquisition. The study found that both AI-based and human detection of animals on satellite imagery of heterogeneous landscape presents many opportunities for error and variable results.

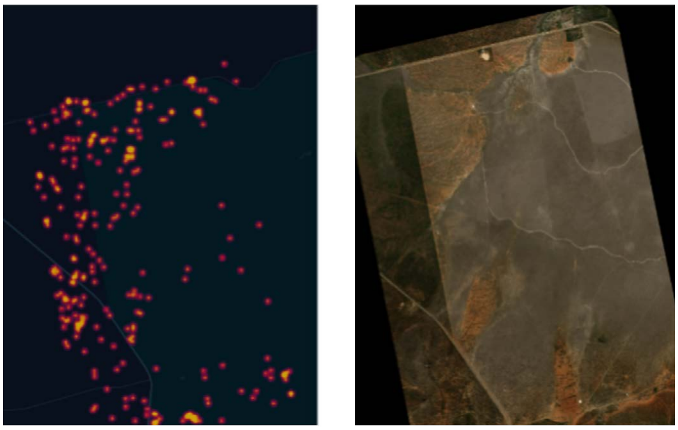
In reviewing the practicalities described in Approach 1 and results from Approaches 2 and 3, and comparing these to traditional wildlife survey techniques, conservation teams at NRT and Madikwe Game Reserve concluded that the accuracy of both AI-based and human detection of species on 30cm satellite imagery was not satisfactory for their heterogenous landscapes and therefore could not replace traditional survey techniques. Even when applying additional information on animal behaviours and known movements during Approach 1, field teams were unable to validate species classifications with satisfactory certainty for a wildlife survey.

Until new aerial drone technologies enable resolution to move beyond 30cm resolutions to < 20cm, these methods should not be pursued as an individual-species-count survey technique for large mammals in variable landscapes.

The next phase of the project will go on to identify the species and situations where 50cm and 30cm resolution satellite imagery can be of value in locating other animals in homogenous, hard-to-reach environments where greater variability in accuracy can be tolerated. For example, it was reported that more accurate results can be expected when isolating scenarios to areas with only one known species found in a homogenous, open landscapes or seascapes, especially when the species of interest stand out in strong contrast to their background environment.

A notable outcome and learning of this study considers the error of the AI models. Rather than giving precise locations, it provides a heatmap of potential sightings as an indicator of 'where to look' for human-eyeballing of an image. As it is incredibly time-consuming to scan satellite images with the human eye – an exercise that also requires exceptional diligence – consulting an AI model is valuable in that it highlights areas of potential detection. This 'where to look' heatmap of sightings offers suggestions for performing a first pass of the imagery for human identification and validation. Field teams believe this would be useful for certain use cases, including locating hard-to-find colonies or herds of certain species in remote areas.

An example of the heatmap output produced in Approach 2 for field teams to review is shown in Figure 15.



Pléiades Neo © Airbus DS 2022

Figure 15: 'Where to look' heatmapping example

6. Future Work

Efforts directed at homogenous use cases and species heatmapping may prove favourable at scale. A next step is to explore the species and situations where AI and 30cm imagery can produce more accurate results. Future exploration and studies should look at situations when species are in groups, not mixed with other animals, and stand out against homogenous backgrounds in terms of colour and limited background variability, such as water. Examples include albatross, dugong, penguins, desert oryx species and hippos.

Additionally, having multiple images of the same landscape presents opportunities to explore the benefits of change detection: identifying permanent or moving objects in the landscape to extract objects of interest and create more definitive labelled training sets for further model training. The project team did raise questions about whether partners could afford the high expense of taking multiple images of the same area of interest. In addition, from the initial comparisons made during this project it is clear that there are significant challenges presented by the variability of lighting, shadows and incident angles on images of the same area of interest taken at different times of the day or year.

In support of future and further conservation research and value, the codebase for the approaches covered in this study have been made available to be trialled and improved upon via the Connected Conservation Foundation GitHub repository.

Access the codebase for all approaches covered in this study.

Please email **info@connectedconservation.foundation** if you'd like to further conduct research using these source materials

7. Related Work

Table 3 gives a summary of past and current research in animal detection and identification. To inform this study on the varying levels of success using high-resolution satellite imagery for detection and identification, research related to the following species was explored: polar bears, penguins, yaks, elk, wildebeest, zebra, whales, albatross, elephants and livestock.

Species	Research and summary of findings
Elephants	**Automatic detection of elephants in South Africa** Researchers were able to automate the counting and detection of elephants in Addo Elephant Park with a higher accuracy level than that achieved by humans detecting species within the satellite imagery. This was done using a neural network algorithm and high-resolution satellite imagery. The detections were not compared to field data on the ground. Reference: Duporge, I., Isupova, O., Reece, S., Macdonald, D. and Wang, T., 2020. Using very-high-resolution satellite imagery and deep learning to detect and count African elephants in heterogeneous landscapes. Remote Sensing in Ecology and Conservation, 7(3), pp.369–381.
Large mammals	**Automatic counting large mammals and discerning mammals from their background** In this study, researchers were able to automate the counting of large mammals like zebra and wildebeests using a neural network algorithm and high-resolution satellite imagery. Researchers also attempted distinguishing animals from various backgrounds. Reference: Xue, Y., Wang, T. and Skidmore, A., 2017. Automatic counting of large mammals from very high-resolution panchromatic satellite imagery. Remote Sensing, 9(9), p.878.
Albatross	**Manual counting of individual albatrosses in remote areas** Researchers were able to accurately count a population of albatrosses in a remote, inaccessible area using extremely high-resolution satellite imagery (0.3m/pixel). Reference: Bowler, E., Fretwell, P., French, G. and Mackiewicz, M., 2020. Using deep learning to count albatrosses from space: assessing results considering ground-truth uncertainty. Remote Sensing, 12(12), p.2026.
Elk and livestock	**Detecting cattle and elk in the wild from space** This study focused on a baseline method, CowNet, that simultaneously estimates the number of animals in an image (counts) and predicts their location at a pixel level (localises). Reference: Robinson, C., Ortiz, A., Hughey, L., Stabach, J.A. and Ferres, J.M.L., 2021. Detecting cattle and elk in the wild from space. arXiv preprint arXiv:2106.15448. **Counting cows: Tracking illegal cattle ranching from high-resolution satellite imagery** Using a neural network algorithm and high-resolution satellite images, researchers were able to automate the counting of cattle in the Amazon to monitor illegal cattle farming. More images and work are recommended for scaling this approach. Reference: Laradji, I., Rodriguez, P., Kalaitzis, F., Vazquez, D., Young, R., Davey, E. and Lacoste, A., 2020. Counting cows: Tracking illegal cattle ranching from high-resolution satellite imagery. arXiv preprint arXiv:2011.07369.
Livestock	**Counting cows: Tracking illegal cattle ranching from high-resolution satellite imagery** Using a neural network algorithm and high-resolution satellite images, researchers were able to automate the counting of cattle in the Amazon to monitor illegal cattle farming. More images and work are recommended for scaling this approach. Reference: Laradji, I., Rodriguez, P., Kalaitzis, F., Vazquez, D., Young, R., Davey, E. and Lacoste, A., 2020. Counting cows: Tracking illegal cattle ranching from high-resolution satellite imagery. arXiv preprint arXiv:2011.07369.

Species	Research and summary of findings
Livestock	**How do you find the green sheep? A critical review of the use of remotely sensed imagery to detect and count animals** This paper reviews recent studies on methods, data requirements and the practical and operational considerations of using remote sensing technologies to estimate animal abundance. It also considers the advantages of better resolution and lower cost imagery, greater computing power, and more advanced statistical algorithms and programming. Reference: Hollings, T., Burgman, M., van Andel, M., Gilbert, M., Robinson, T. and Robinson, A., 2018. How do you find the green sheep? A critical review of the use of remotely sensed imagery to detect and count animals. Methods in Ecology and Evolution, 9(4), pp.881–892.
Multiple species	**Surveying wild animals from satellites, manned aircraft and unmanned aerial systems (UASs) – a review ("Surveying Wild Animals from Satellites, Manned Aircraft and Unmanned ...")** This paper looks at numerous studies on wild-animal surveys and focuses on the data used, animal-detection methods, and their accuracies. The proposed use of UAS imagery in combination with very high-resolution (VHR) satellite imagery would produce critical population data for large wild-animal species and colonies over large areas. The development of software systems for automatically producing image mosaics and recognising wild animals will further improve survey efficiency. Reference: Wang, D., Shao, Q. and Yue, H., 2019. Surveying wild animals from satellites, manned aircraft, and unmanned aerial systems (UASs): A review. Remote Sensing, 11(11), p.1308.
Multiple species	**Spotting East African mammals in open savannah from space** Knowledge of population dynamics is essential for managing and conserving wildlife; however, traditional methods of counting come with challenges. This study explores the possibility of detecting large animals in open areas from VHR GeoEye-1 satellite images. A hybrid image classification method was employed by incorporating the advantages of both pixel-based and object-based image classification approaches. The results showed that, for the first time, it is feasible to perform automated detection and counting of large wild animals in open areas from space, and therefore provided a complementary and alternative approach to conventional wildlife counting methods. Reference: Yang, Z., Wang, T., Skidmore, A.K., De Leeuw, J., Said, M.Y. and Freer, J., 2014. Spotting East African mammals in open savannah from space. PloS one, 9(12), p.e115989.
Multiple species	**Satellite imagery for wildlife monitoring & tracking** High-resolution satellite imagery gives scientists and researchers increasingly up-to-date geospatial data. By using neural networks processing, reliable statistics are obtained for monitoring wildlife migrations, mapping habitats and tracking endangered species in remote areas of the world to assist in management and conservation activities. Reference: Satellite Imaging Corporation. 2022. Satellite imagery for wildlife monitoring & tracking. [online] Available at: <https://www.satimagingcorp.com/applications/environmental-impact-studies/wildlife-and-marine-conservation/wildlife-monitoring/> [Accessed 10 October 2022].
Polar bears	**Polar bears from space: assessing satellite imagery as a tool to track Arctic wildlife** High-resolution satellite imagery was evaluated as a tool to track the distribution and abundance of polar bears. Bears were distinguished from other light-coloured spots by comparing images collected on different dates. A sample of ground-truth points demonstrated that the bears were accurately classified. These findings suggest that satellite imagery is a promising tool for monitoring polar bears on land, with implications for use with other Arctic wildlife. Reference: Stapleton, S., LaRue, M., Lecomte, N., Atkinson, S., Garshelis, D., Porter, C. and Atwood, T., 2014. Polar bears from space: assessing satellite imagery as a tool to track Arctic wildlife. PLoS One, 9(7), p.e101513.