



OPEN

## Automated rhinoceros detection in satellite imagery using deep learning

Isla Duporge<sup>1</sup>✉, Xiaomin Lin<sup>2</sup>, Aadi Palnitkar<sup>3,6</sup>, Arjun Suresh<sup>3,6</sup>, Olga Isupova<sup>4</sup>, Daniel Rubenstein<sup>1</sup> & John Yiannis Aloimonos<sup>3,5</sup>

Rhinoceroses face severe threats from poaching, habitat fragmentation, and ongoing habitat degradation. Monitoring rhinoceros across the vast, often inaccessible landscapes they inhabit is challenging. In this study, we assess the feasibility of detecting white rhinoceroses using very high-resolution (33–36 cm) satellite imagery acquired over the world's largest private rhinoceros reserve in South Africa using a YOLO-based object detection model (YOLOv12x). We test whether synthetic imagery enhances model performance, whether rhinoceroses can be reliably distinguished from elephants in satellite imagery, and whether synthetically generated rhinoceroses are visually distinguishable from real ones by human annotators. We achieve an average precision (AP) of 0.65 in detection accuracy with synthetic augmentation yielding a marginal improvement. This study provides a demonstration of monitoring rhinos using this approach and introduces an open-access dataset to support the development and testing of new models. The aim is to facilitate effective monitoring of rhinos across the vast landscapes they inhabit. Developing new detection techniques can strengthen conservation and recovery initiatives, including translocations, assessment of breeding program success, and evaluation of anti-poaching efforts.

**Keywords** Deep Learning, Object Detection, Wildlife Monitoring, Population Estimation, Conservation Technology, Satellite Remote Sensing, Earth Observation

Reliable methods to accurately count wildlife are critical to detect changes in population sizes and document demographic trends that can be used to guide the conservation and management of populations. For many species, obtaining reliable data on population numbers is challenging, as species often occupy remote or inaccessible habitats, are nocturnal, or display cryptic behaviors that hinder detection. Wildlife survey methods vary in their strengths, with the optimal choice dependent on the monitoring objective. For example, foot patrols are necessary for close-range assessments of body condition, while estimating survival rates necessitates individual identification. The use of UAVs for monitoring wildlife is promising<sup>1</sup>, but the scale of surveys is currently limited by UAV battery life, and data can be compromised by the acoustic disturbance UAVs cause to animals<sup>2</sup>. To date, Earth observation satellite imagery has been used primarily for species detection, but this approach also provides broader ecological insights into habitat selection and migration dynamics<sup>3</sup>. Several comprehensive reviews have examined the use of satellite imagery for wildlife monitoring<sup>4–6</sup>, however, empirical studies remain limited relative to the number of species for which this monitoring approach is now technically viable. The high cost of sub-meter imagery remains a major challenge; costs depend on the provider, order volume, and image specifications. Freely available imagery from satellite programs such as NASA's Landsat Series (30 m resolution, 16-day revisit) and ESA's optical sensors under the Copernicus program (10–20 m resolution, 5-day revisit) offers valuable broad-scale coverage for habitat assessment; however, their coarse spatial resolution prevents detection of individual animals. In contrast, several commercial providers offer sub-meter imagery suitable for detailed object detection. For example, Maxar Technologies' WorldView-3 satellite—used in this study—and the WorldView Legion constellation provide panchromatic imagery at resolutions as fine as 0.29 m, with up to 15 revisits per day, enabling the detection of individual animals from low Earth orbit<sup>7</sup>. As an unprecedented number of high spatial- and temporal-resolution commercial constellations are launched<sup>8</sup>, and

<sup>1</sup>Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA. <sup>2</sup>Department of Electrical Engineering, University of South Florida, Tampa, USA. <sup>3</sup>Maryland Robotics Center, School of Engineering, University of Maryland, MD, USA. <sup>4</sup>Leverhulme Centre for Nature Recovery, The School of Geography and the Environment, University of Oxford, Oxford, UK. <sup>5</sup>Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, MD, USA. <sup>6</sup>These authors contributed equally to this work: Aadi Palnitkar and Arjun Suresh. ✉email: Isla.duporge@princeton.edu

companies seek to remain competitive, costs are expected to fall. In addition to high costs, manual analysis of large satellite images remains labor-intensive, time-consuming, and tedious. Deep-learning approaches allow for automation of this process, provided sufficient labeled training data are available. Both point- and object-based deep-learning approaches have proven effective for detecting individual animals in satellite imagery, as successfully demonstrated in the case of whales<sup>9</sup>, elephants<sup>10</sup>, and wildebeest<sup>3</sup>. For wildebeest, smaller body size is offset by their tendency to form large, dense aggregations, which in the case of migratory wildebeest number in the hundreds of thousands, which increases ease of detectability<sup>11</sup>. In contrast, white rhinoceroses typically occur in much smaller crashes or around five and rarely exceeding 10 individuals<sup>12</sup>, making them harder to detect despite having a larger individual body size.

Satellite-based monitoring is suited to species detection<sup>13</sup> and population counting<sup>14</sup>—tasks for which aerial surveys are the main alternative, with each method subject to distinct biases that influence survey accuracy. Both techniques are subject to availability or concealment bias, where animals present in the survey area may remain undetected when hidden—such as rhinos resting under trees during the hottest parts of the day<sup>15</sup>. This bias has spatial and temporal components: satellite imagery can reduce spatial bias by capturing much larger areas in a single acquisition, but temporal bias remains due to fixed orbital schedules. Aerial surveys provide greater flexibility in timing to avoid periods of low visibility, whereas the revisit limitations of satellites are expected to diminish with the launch of new constellations, which will offer higher revisit frequencies and improved temporal resolution.

Observer bias and detectability bias—where object visibility decreases with distance from the observer—are mitigated in satellite surveys; however, classification error, i.e., the incorrect assignment of the wrong species category (e.g., mistaking a rhino for an elephant or vice versa), affects both techniques and is particularly challenging for satellite surveys due to the lower resolution of imagery. This form of sampling error can skew population estimates, particularly in areas where multiple large-bodied species co-occur, such as rhinos, elephants, and hippos. Involving species experts familiar with diurnal movement patterns can reduce this detection bias by improving image interpretation and minimizing classification error.

Given the critical conservation status of rhino species, improving the accuracy and scalability of surveying methods is essential for tracking population trends and informing protection efforts. As umbrella species, rhinos provide protection for many co-occurring species; conserving them also safeguards the broader ecosystems and landscapes they inhabit. Rhinos remain critically threatened throughout much of their range. Globally, five rhinoceros species and eleven subspecies have been documented, of which four subspecies have recently gone extinct. The Northern White Rhino (*Ceratotherium simum cottoni*) is now functionally extinct, while the Southern White Rhino (*Ceratotherium simum simum*), the most numerous remaining species, is classified as Near Threatened by the IUCN and is most populous in South Africa<sup>16</sup>. Poaching, primarily driven by demand for rhino horn as a status symbol and in traditional medicine, remains a significant threat with over 1,000 individuals killed annually between 2013 and 2016<sup>17</sup>. In recent years, there has been a reported decline in mortality<sup>18</sup>, while the reduction is in part due to a reduction in poaching effort, the disruption to rhino demography has been significant<sup>19</sup>.

Protection efforts for rhinos extend beyond biological management to include law enforcement initiatives to disrupt organized crime, strategies for trade and demand reduction, and the development of alternative livelihoods for local communities. Conservation strategies for rhino range from breeding programs and translocations<sup>20,21</sup>, to demand-reduction campaigns<sup>22</sup>, the creation of synthetic horn alternatives<sup>23</sup>, dehorning<sup>24</sup>, and proposals to legalize the horn trade<sup>25</sup>. Despite all these initiatives, the future of rhinos remains uncertain. Currently, around half of the global white rhino population is privately managed, while efforts continue to seek long-term solutions for protecting populations in the wild<sup>26</sup>.

In this study, we evaluate the feasibility of detecting rhinoceroses using high-resolution imagery (33–36 cm) acquired from the commercial WorldView-3 satellite (Maxar Technologies) over the world's largest private rhino reserve in South Africa, home to approximately 2,000 individual white rhino (*Ceratotherium simum simum*). We use temporal image differencing as shown in Fig. 1 to identify rhinos. We then automate the detection process using the YOLOv12x object detection model<sup>27</sup>—detections from this model are shown in (Fig. 2).

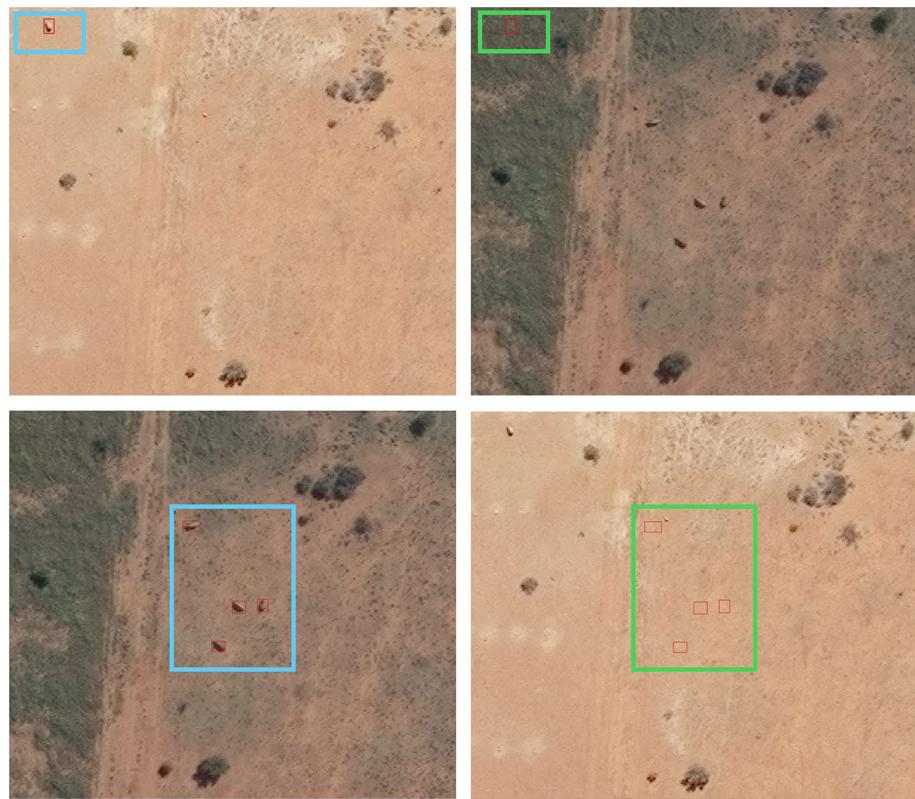
The site is selected to minimize potential confusion with other African megafauna and to provide consistent background conditions for the detection task. It is not fully representative of the broader range of habitats occupied by rhinos, and other locations may present either greater or lesser challenges for detection than those examined here. We assess whether augmenting the training dataset with synthetically generated rhino imagery improves model performance. We also investigate whether human evaluators can distinguish synthetic from real rhinos, and whether the model can reliably differentiate rhinos (*Ceratotherium simum simum*) from African elephants (*Loxodonta africana*) using satellite imagery from Addo Elephant Park, South Africa.

## Results

The YOLOv12x object detection model, trained on 1,700 manually annotated sub-images and evaluated on a test set of 300 images, achieved an average precision (AP) of 0.6308, with a precision of 65.2%, recall of 61.4%, and an F1 score of 63.3%. This demonstrates the feasibility of detecting rhinoceroses in satellite imagery despite their cryptic appearance, small pixel footprint, and the complexity of heterogeneous landscapes.

A separate model trained solely on 1,000 synthetic images achieved a lower AP of 0.26 (precision: 40.5%, recall: 22.3%, F1 score: 28.6%) when evaluated on real-world imagery. However, augmenting the training set with synthetic data led to a modest improvement in detection performance, with the combined model reaching a peak AP of 0.65 (precision: 66.7%, recall: 63.1%, F1 score: 64.8%).

To assess the visual realism of synthetic images, human annotators were asked to distinguish between real and synthetic examples. For real images, annotators correctly identified 140 out of 270 as genuine (true positives),



**Fig. 1.** Temporal comparison of WorldView-3 satellite imagery illustrating rhino movement. A blue box highlights the presence of a rhino, while a green box shows its absence in the same location on a later acquisition, confirming the object as a moving animal rather than a static feature. The red box similarly indicates a rhino detection and its subsequent absence in the same location at a later time. Images acquired on 3 September 2019 and 13 September 2022. Imagery from Maxar Technologies.

yielding a recall of 53%. For synthetic images, 130 out of 171 were correctly identified as fake (true negatives), while 41 were misclassified as real (false positives), resulting in a specificity of 76%.

When trained on a combined dataset of 7,532 individually annotated instances of African elephants (*Loxodonta africana*) and white rhinoceroses (*Ceratotherium simum simum*), the model failed to reliably distinguish between the two species. We tested three two-class training configurations (rhino vs. elephant), although no image contained both species together. On the held-out test set of 324 images (1,158 annotated instances), the model achieved an overall AP of 0.247, with 32.8% precision, 28.7% recall, and an F1 score of 30.6%. Performance was notably stronger for rhinos (AP 0.350, precision 42.1%, recall 39.0%, F1 40.5%) than for elephants (AP 0.143, precision 23.5%, recall 18.3%, F1 20.6%).

## Discussion

The results of this study demonstrate that white rhinoceroses (*Ceratotherium simum simum*) can be identified and detected in satellite imagery Fig. 3 with manual screening and temporal image differencing. The 33–36 cm resolution imagery used in this study, combined with YOLOv12x, achieved only a moderate average precision (AP) of 0.63 without any synthetic data included in training. Future research could build on the freely available training dataset provided here to improve detection accuracy by testing alternative models, such as U-Net<sup>28</sup> or HerdNet<sup>29</sup>. We encourage others to build on this dataset by incorporating imagery of additional rhino species from the diverse ecological contexts in which they occur across Asia and Africa. The challenge of distinguishing between species at this spatial resolution is evident from the poor performance in separating elephants from rhinos. We encourage future research to address this limitation, which will be surmountable as satellite imagery with higher spatial resolution becomes available. The capacity to detect animals directly, without relying on temporal image differencing, will be enabled by advances in satellite video acquisition<sup>30</sup>. The issue of double-counting is minimized in satellite surveys because very large areas upwards of 4000 km<sup>2</sup><sup>3</sup> can be captured in a single image at less than 50 cm resolution. When multiple images from different timestamps are used and individuals are treated as independent observations, the same problem of double-counting that occurs in traditional aerial surveys can arise. In such cases, established principles for estimating detection probability and abundance can be applied to account for availability bias<sup>31</sup>.

In the long term, assembling a spectrally and contextually diverse library of wildlife examples in satellite imagery across multiple habitats will advance the field, enabling the refinement of robust automated detection methods and benchmarks and reducing reliance on manual annotation. Large-scale datasets have been compiled



**Fig. 2.** Top row: original unlabeled satellite imagery. Bottom row: rhinoceros detections generated by the YOLOv12 model. WorldView-3 image acquired on 13 September 2022. Imagery from Maxar Technologies.



**Fig. 3.** Zoomed-in visualization of two rhinoceroses within the reserve, illustrating that each individual occupies fewer than 30 pixels, including body shadow. WorldView-3 image acquired on 3 September 2019. Imagery from Maxar Technologies.

for camera-trap studies<sup>32,33</sup>, but are lacking for satellite imagery due to the high cost and proprietary nature of imagery. The cost of acquiring sub-half-meter resolution satellite imagery is substantial; however, we are at an unprecedented point in history with a record number of Earth observation satellites being launched<sup>8</sup>. As the market for very high-resolution satellite imagery grows and competition increases, acquisition costs are likely to decrease over time, making this monitoring technique more accessible. In addition, several satellite providers offer imagery at no cost for projects with ecological<sup>34</sup> or societal benefits<sup>35</sup>.

At present, satellite-based surveying is competitive in cost to manned aerial surveys in remote or logistically challenging regions, such as Arctic environments, while in areas with established aviation infrastructure, conventional aerial surveys may remain more cost-effective in the near term. Different survey methods serve different purposes, and the optimal approach depends on the wildlife monitoring objective. At present, the

primary applications of satellite-based monitoring are species detection and counting; however, beyond counting, satellites can be used to examine how groups of animals move together across vast areas, providing a better understanding of collective behavior dynamics<sup>3</sup>. While the cost of census counts is an important consideration, accuracy and bias quantification are equally critical. Conducting aerial and satellite surveys simultaneously across representative sites would enable direct evaluation of detection accuracy and quantification of relative bias errors. Such cross-method comparisons are well established in ecology, for example in contrasts between mark–recapture and block-count estimation<sup>36</sup>. Embracing methodological pluralism is essential to validate monitoring approaches rather than relying on a single technique.

In our study, incorporating synthetic images led to only a modest improvement in model performance (peak AP of 0.65). However, this limited gain may be attributed more to the challenge of dissociating pixel clusters due to current spatial resolution constraints of satellite imagery than to the effectiveness of synthetic augmentation itself. As wildlife satellite monitoring advances, continued development of processing pipelines that leverage synthetic data, as piloted here, will be essential to overcome sparse training data and unlock the full potential of satellite-based biodiversity monitoring.

While not directly examined in this study, the integration of AI-generated imagery into wildlife detection pipelines raises a potential future concern: fabricated or manipulated imagery could distort population estimates. Although synthetic data can enhance model training and address class imbalance, a common limitation in wildlife datasets<sup>37</sup>, it also introduces the risk of misuse. This risk becomes particularly salient if species population estimates are incorporated into financial mechanisms, such as biodiversity credit markets—satellite-based detection is already used to quantify the benefits of forest carbon offset projects<sup>38</sup>, and similar integration for species accounting is conceivable. Should this occur, ensuring rigorous data provenance capable of distinguishing genuine from synthetic imagery will be essential<sup>39</sup>, paralleling safeguards now being implemented in other fields vulnerable to deepfake manipulation<sup>40</sup>. This study demonstrates the feasibility of satellite-based monitoring for surveying rhino populations and provides a foundation for a scalable monitoring approach in support of their conservation.

## Methods

### Satellite image acquisition

We used six cloud-free, high-resolution WorldView-3 satellite images (Maxar Technologies) acquired between 2015 and 2022 (Table 1). Each image had a ground sample distance of 33–36 cm and covered a substantial portion of the farm. Across all six dates, the total area surveyed was 188.6 km<sup>2</sup>, with locations imaged more than once across the six images. The farm is located in the North West province of South Africa, the climate is characterized by a wet summer season (November–March) and a dry winter season (April–October); all satellite imagery in this study was acquired during the dry season, spanning early (May), mid (July), and late (September) phases, when reduced vegetation cover and clearer atmospheric conditions generally enhance detectability. This site was selected to minimize potential confusion with other African megafauna and to ensure consistent background conditions for the detection task. It is not representative of the broader range of habitats occupied by rhinos, and other locations may present either greater or lesser challenges for detection than those examined here. To produce high-quality labeled training data, annotations were carried out by a professional annotation service (<https://labelyourdata.com/>). Each image was independently annotated by two trained annotators, with all instances of disagreement adjudicated by a senior reviewer to verify the presence or absence of rhinoceros. Any annotations unique to a single annotator were subjected to an additional verification step to confirm accuracy before inclusion in the final dataset. In several instances, small trees and shrubs were initially misidentified as rhinoceroses; however, temporal image differencing revealed that these objects were stationary, and they were subsequently excluded from the training dataset.

This consensus-based procedure yielded 4,223 validated rhinoceros annotations. To further improve annotation reliability and reduce the likelihood of false positives arising from visually similar static features (e.g., rocks, shrubs, shadows), a temporal image differencing method was employed. This approach involved confirming the presence of a suspected rhinoceros by examining its location in images acquired at different times as shown in Fig. 2 below. Only those objects that exhibited positional displacement consistent with rhino movement were retained.

Satellite Image ID	Date	Training Labels	Test Labels	Validation Labels	Total Labels
104001000B0A7D00	2015-05-08	91	12	6	109
104001001E10CF00	2016-07-12	227	44	22	293
104001001B4BBC00	2016-05-22	941	117	58	1116
1040010051205100	2019-09-03	434	51	26	511
104001005AC36B00	2020-05-07	709	156	74	939
104001007BC2F800	2022-09-13	73	1	0	74
104001007BC2F800	2022-09-13	754	284	143	1181

**Table 1.** Summary of Worldview-3 Satellite Images Used in Training (3229 unique rhinos), Validation (513 unique rhinos), and Test Datasets (481 unique rhinos), each label contains between 0–64 rhino.

## Preprocessing and model selection

Given the large spatial dimensions of the WorldView-3 satellite imagery and the irregular extents of the georeferenced TIFF files (imagery in which each pixel is assigned precise geographic coordinates), we adopted an image tiling strategy to facilitate efficient training. The area surveyed covers 188.6 km, which converts to 1.73 billion pixels at 33 cm resolution. Each image was subdivided into uniformly sized tiles of  $512 \times 512$  pixels. This patch size was selected to meet the input requirements of deep learning architectures while preserving the fine spatial resolution necessary to detect small, visually cryptic targets such as rhinoceros. Care was taken during the tiling process to preserve spatial accuracy and prevent the truncation of rhinos at tile edges, there was no overlap across the images. Subsequent preprocessing and quality control were conducted using QGIS (Quantum GIS)<sup>41</sup>, an open-source geographic information system. These steps included manual inspection to ensure annotation consistency, correction of any labeling discrepancies i.e., tightening labels around the rhino, and conversion of polygon shapefiles (.shp) into the bounding box format compatible with the YOLO training framework. The final dataset was partitioned into 1,700 training images, 150 validation images, and 300 test images. Table 2 shows the training hyperparameters used during training.

These numbers refer to images that comprise different numbers of rhino per image, while Table 1 refers to the number of unique rhino “objects” in the imagery. For object detection, we employed YOLOv12x<sup>27</sup>, the largest variant of the recently released YOLOv12 model family. We selected YOLOv12x for its high-resolution feature extraction and strong multi-scale detection performance, which enables identification of small, visually cryptic targets such as rhino.

## Synthetic data generation

To explore the potential of synthetic data in augmenting model training, we constructed a 3D scene in Blender, to emulate the satellite observation perspective. Blender is an open-source 3D computer graphics software used for creating 3D models, and simulations<sup>42</sup>. To generate synthetic imagery, a virtual camera was positioned directly overhead of a flat plane representing the ground. The plane was textured using patches cut from real satellite images to create realistic backgrounds. For each synthetic image, a random patch was selected, with all patches manually screened to remove those containing cloud cover, visible reserve infrastructure, or real rhinoceroses. Three-dimensional rhinoceros models were randomly distributed across the scene with variable scaling and rotation parameters, simulating natural variation in animal size and orientation As shown in Fig. 4. The models were placed sparsely to reflect realistic densities observed in the original dataset.

Synthetic images were rendered at a resolution of  $512 \times 512$  pixels to match the dimensions used for training with real image tiles. While this approach enabled the rapid generation of labeled training data, initial training runs using YOLOv12x on the purely synthetic dataset yielded low performance, with a baseline average precision (AP) of 0.26 substantially below that of the real-image model. The model exhibited difficulty in generalizing from synthetic scenes, likely due to differences in texture fidelity and visual complexity, despite extensive hyperparameter tuning using grid search. In this process, key hyperparameters such as the learning rate (0.001), momentum (0.937), weight decay (0.0005), optimizer (SGD), image size (512), batch size (8), epochs (500), and training options (augmentation, multi-scale input, mixed-precision, checkpoint saving) were systematically varied across predefined values to identify the best-performing configuration.

The synthetic data went through several iterations, focusing on altering the size, distribution, and number of 3D model placements to see what the model would respond best to.

## Differentiating elephant and rhino

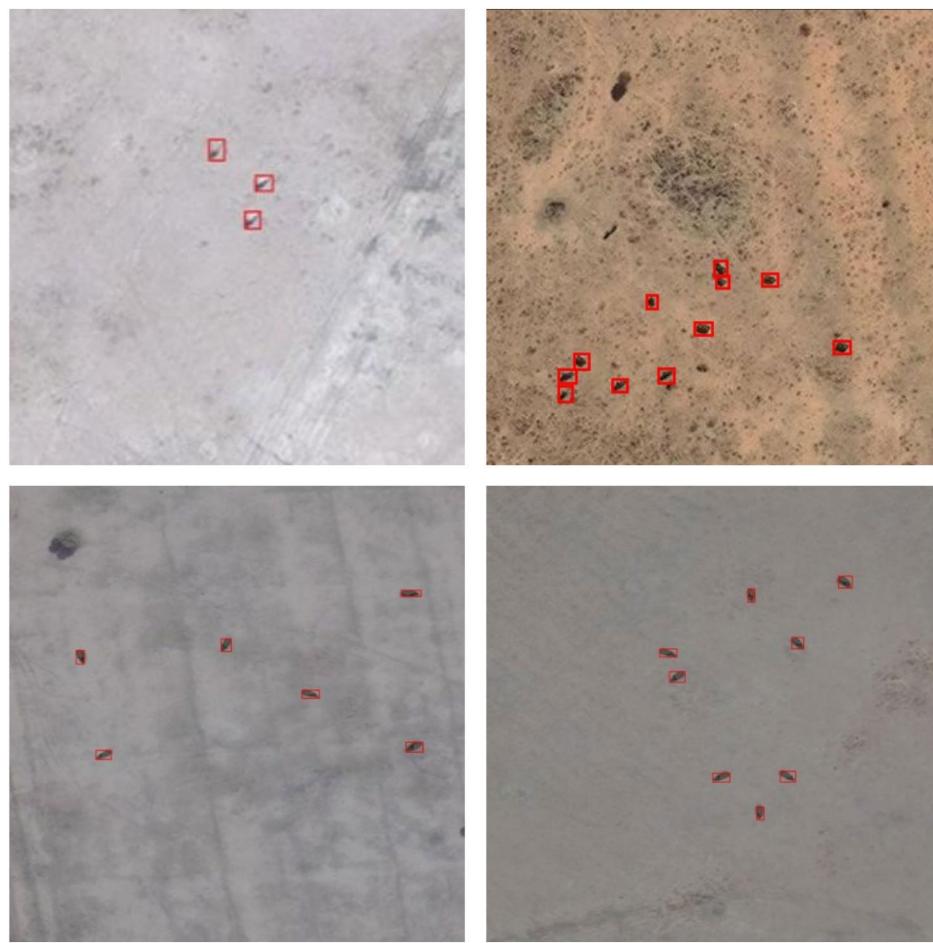
To assess whether African elephants (*Loxodonta africana*) and white rhinoceroses (*Ceratotherium simum*) can be distinguished in satellite imagery, we trained a YOLOv12 object detection model on a combined dataset. Elephant training data were sourced from Duporge et al. (2021)<sup>10,43</sup> combined with newly annotated rhino imagery. All images were acquired from very high-resolution commercial satellites (30–50 cm per pixel), and annotations were manually generated by trained observers using bounding boxes to delineate individual animals. The final dataset comprised 7,532 annotations within the train set, 324 test images, and 125 validation images, containing 1,126 elephant and 4,223 rhino annotations with varying numbers of each species in each image. We used an 80–10–10 split for training, validation, and testing.

## Human evaluation of synthetic data

To assess the visual realism of our synthetic data, we conducted a blinded classification task involving human annotators (Princeton University Freshman students). Eight participants—none of whom were involved in the project and with little to no prior exposure to satellite wildlife imagery—were asked to classify images as either real or synthetic. The dataset comprised an equal mix of real and AI-generated high-resolution satellite images of rhinoceroses. The 108 images (54 real, 54 synthetic) were divided into two labeling tasks. Annotators were instructed to categorize each image as “real” or “synthetic” without any prior training or guidance. Results revealed that participants frequently misclassified synthetic images as real, underscoring the high visual fidelity of the generated data (Fig. 4).

Learning rate	Momentum	Weight decay	Optimizer	Batch size	Number of epochs
0.001	0.937	0.0005	SGD	8	500

**Table 2.** Hyperparameter values for YOLOv12 training.



**Fig. 4.** Top row: examples of real satellite detections of rhinos. Bottom row: synthetic rhino images generated in Blender using realistic 3D models overlaid on actual satellite backgrounds. WorldView-3 imagery acquired on 3 September 2019. Imagery from Maxar Technologies.

## Data availability

The Python and Blender-based framework used for generating the synthetic rhinoceros dataset is available at: Github Repository : <https://github.com/sat-rhino/sat-rhino>. All associated data supporting the results of this manuscript are publicly accessible at: <https://doi.org/10.5281/zenodo.17360632>.

Received: 20 May 2025; Accepted: 10 October 2025

Published online: 10 November 2025

## References

1. Kholiavchenko, M. et al. Deep dive into KABR: a dataset for understanding ungulate behavior from in-situ drone video. *Multimed. Tools Appl.* **84**, 24563–24582. <https://doi.org/10.1007/s11042-024-19258-1> (2024).
2. Duporge, I. et al. Determination of optimal flight altitude to minimise acoustic drone disturbance to wildlife using species audiograms. *Methods Ecol. Evol.* **12**, 2196–2207 (2021).
3. Duporge, I. et al. AI-based satellite survey offers independent assessment of migratory wildebeest numbers in the serengeti. *PNAS Nexus* **4**, gaf264 (2025).
4. Larue, M., Stapleton, S. & Anderson, M. L. Feasibility of using high-resolution satellite imagery to assess vertebrate wildlife populations. *Conserv. Biol.* **31**, 213–220 (2017).
5. Wang, D., Shao, Q. & Yue, H. Surveying wild animals from satellites, manned aircraft and unmanned aerial systems (UASs): A review. *Remote. Sens.* **11**, 1308 (2019).
6. Delplanque, A. et al. Wildlife detection, counting and survey using satellite imagery: are we there yet?. *GISci. Remote Sens.* **61**, 2348863 (2024).
7. Maxar technologies. <https://www.maxar.com/worldview-legion>. Accessed: 2025-08-09.
8. European Space Agency. We're launching more than ever. [https://www.esa.int/ESA\\_Multimedia/Images/2021/03/We\\_re\\_launching\\_g\\_more\\_than\\_ever](https://www.esa.int/ESA_Multimedia/Images/2021/03/We_re_launching_g_more_than_ever) (2021). Accessed: 2025-08-07.
9. Gaur, A., Liu, C., Lin, X., Karapetyan, N. & Aloimonos, Y. Whale detection enhancement through synthetic satellite images. In *OCEANS 2023-MTS/IEEE US Gulf Coast 1–7*. <https://doi.org/10.48550/arXiv.2308.07766> (IEEE, 2023).
10. Duporge, I., Isupova, O., Reece, S., Macdonald, D. W. & Wang, T. Using very-high-resolution satellite imagery and deep learning to detect and count african elephants in heterogeneous landscapes. *Remote. Sens. Ecol. Conserv.* **7**, 369–381 (2021).

11. Wu, Z. et al. Deep learning enables satellite-based monitoring of large populations of terrestrial mammals across heterogeneous landscape. *Nat. Commun.* **14** <https://doi.org/10.1038/s41467-023-38819-x> (2023).
12. Owen-Smith, R. N. The social ethology of the white rhinoceros ceratotherium simum (burchell 1817). <https://doi.org/10.1111/j.1439-0310.1975.tb02010.x> (1975).
13. Xu, Z., Wang, T., Skidmore, A. K. & Lamprey, R. A review of deep learning techniques for detecting animals in aerial and satellite images. *Int. J. Appl. Earth Obs. Geoinformation* **128**, <https://doi.org/10.1016/j.jag.2024.103732> (2024).
14. Fretwell, P. T. & Trathan, P. N. Discovery of new colonies by sentinel2 reveals good and bad news for emperor penguins. *Remote. Sens. Ecol. Conserv.* **7**, 139–153. <https://doi.org/10.1002/rse2.176> (2021).
15. Allbrook, D.B., Harthoorn, A.M., Luck, C.P. Temperature regulation in the white rhinoceros. *J. Physiol.* **143** (1958).
16. Emslie, R. Iucn red list of threatened species: Ceratotherium simum. *IUCN Red List of Threatened Species* <https://doi.org/10.2305/IUCN.UK.2020-1.RLTS.T4185A45813880.en> (2020).
17. Litchfield, C. A. Rhino poaching: apply conservation psychology. *Science* **340**, 1168–1168. <https://doi.org/10.1126/science.340.6137.1168> (2013).
18. IUCN Species Survival Commission (IUCN SSC), African and Asian Rhino Specialist Groups and TRAFFIC. Rhinoceroses (rhinocerotidae spp.). Tech. Rep., Convention on International Trade in Endangered Species of Wild Fauna and Flora (2022).
19. Ferreira, S. M. et al. Disruption of rhino demography by poachers may lead to population declines in kruger national park, south africa. *PLoS One* **10**, e0127783 (2015).
20. Inganga, B. & Imray, G. Kenya embarks on its biggest rhino relocation project. a previous attempt was a disaster. *AP News* (2024). Accessed on 2025-05-18.
21. The Australian Rhino Project. <https://theaustralianrhinoproject.org/index.php/about> (2025). Accessed: 2025-05-01.
22. Macilwain, C. Biologists out of africa over rhino dispute. *Nature* **368**, 677. <https://doi.org/10.1038/368677a0> (1994).
23. Saving rhinos with 3d printed horns. *Nat. Biotechnol.* **33**, 683. <https://doi.org/10.1038/nbt.3236> (2015).
24. Kuiper, T. et al. Dehorning reduces rhino poaching. *Science* **388**, 1075–1081 (2025).
25. Wikelski, M. Protect white rhinos by legalizing horn trade. *Science* **386**, 157. <https://doi.org/10.1126/science.abp3195> (2024).
26. Tensen, L. Under what circumstances can wildlife farming benefit species conservation?. *Glob. Ecol. Conserv.* **6**, 286–298. <https://doi.org/10.1016/j.gecco.2016.03.008> (2016).
27. Tian, Y., Ye, Q. & Doermann, D. Yolov12: Attention-centric real-time object detectors. *arXiv preprint arXiv:2502.12524* (2025).
28. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI* **234–241**, 2015. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28) (Springer International Publishing 2015).
29. Delplanque, A., Foucher, S., Lejeune, P., Linchant, J. & Théau, J. From crowd to herd counting: How to precisely detect and count african mammals using aerial imagery and deep learning?. *ISPRS J. Photogramm. Remote. Sens.* **197**, 167–180. <https://doi.org/10.1016/j.isprsjprs.2023.01.025> (2023).
30. Li, S. et al. A multitask benchmark dataset for satellite video: Object detection, tracking, and segmentation. *IEEE Trans. Geosci. Remote Sens.* **61**, 1–21 (2023).
31. Nichols, J. D. et al. A double-observer approach for estimating detection probability and abundance from point counts. *Auk* **117**, 393–408 (2000).
32. Zhao, J., Institution, S., McShea, W. & Institution, S. Behind eMammal's success: A data curator with a data standard. *J. Escience Libr.* **7**, (2018).
33. Swanson, A. et al. Snapshot serengeti, high-frequency annotated camera trap images of 40 mammalian species in an african savanna. *Sci. Data* **2**, 150026 (2015).
34. Koola Digital. Technologies to protect nature for people and planet. <https://connectedconservation.foundation/>. Accessed: 2025-05-09.
35. Open data program. <https://www.maxar.com/open-data>. Accessed: 2025-05-09.
36. Ferreira, S. M., Le Roex, N., Greaver, C. & Dreyer, C. Comparison of mark-recapture and block-count-based estimation of black rhinoceros populations. *Afr. J. Wildl. Res.* **50** (2020).
37. Duporge, I. et al. BaboonLand dataset: Tracking primates in the wild and automating behaviour recognition from drone videos. *Int. J. Comput. Vis.* **1–12** (2025).
38. Coffield, S. R. et al. Using remote sensing to quantify the additional climate benefits of california forest carbon offset projects. *Glob. Chang. Biol.* **28**, 6789–6806 (2022).
39. Aide, T. M. The biodiversity credit market needs rigorous baseline, monitoring, and validation practices. *npj Biodiversity* **3**, 30. <https://doi.org/10.1038/s44185-024-00062-6> (2024).
40. Babaei, R., Cheng, S., Duan, R. & Zhao, S. Generative artificial intelligence and the evolving challenge of deepfake detection: A systematic analysis. *J. Sens. Actuator Netw.* **14**, 17 (2025).
41. QGIS Development Team. *QGIS Geographic Information System*. Open Source Geospatial Foundation. <https://qgis.org> (2024). Version 3.x.
42. Blender Foundation. Blender: a 3d modelling and rendering package (2025). Version 4.3.2. Blender Foundation, Amsterdam, The Netherlands. Available at: <https://www.blender.org>. Accessed: May 4 (2025).
43. Duporge, I., Isupova, O., Reece, S., Macdonald, D. W. & Wang, T. Using very-high-resolution satellite imagery and deep learning to detect and count African elephants in heterogeneous landscapes. Zenodo. <https://doi.org/10.5281/zenodo.1736076> (2025).

## Acknowledgements

We are grateful to Dr Stephen Lee for assistance in acquiring the imagery for this study, to Valeria Parfeniuk at Labelyourdata (<https://labelyourdata.com/>) for assisting in image annotation, and to Peter S. Goodman for information on the chosen study area.

## Author contributions

I.D. conceptualized the project, acquired and curated the satellite data, coordinated the object-based annotation process, and led the writing of the manuscript. A.P. and A.S. conducted the model training, A.S. and X.L. developed the synthetic data pipeline with input from J.Y.A. O.I. & D.R. provided comments on the manuscript.

## Funding

The research was carried out while ID was supported by the Research Associate Program of the National Academy of Science (NAS), and the image labeling was supported by the NSF OAC 2118240 Imageomics Institute award.

## Declarations

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to I.D.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025