

Aprendizaje profundo aplicado a datos de Biodiversidad: Reconociendo mamíferos africanos a partir de imágenes de distintas fuentes.

Mayra L. Maldonado

Proyectos Educativos Regionales de Autoayuda. 6 avenida 8-25, zona 2. Ciudad Guatemala. Guatemala, 01002
mayrasmaldonadom@gmail.com

El análisis exploratorio de datos (EDA, por sus siglas en inglés) se utiliza para examinar y visualizar datos para entender sus características más prominentes, identificar patrones, descubrir anomalías, y probar hipótesis. Con estos programas buscamos inspeccionar visualmente las imágenes, evaluar la eficacia de los sistemas de reconocimiento, así como evaluar y revisar rápidamente imágenes nuevas (Greenberg 2025).

En este ejercicio usamos RetinaNet, un algoritmo avanzado de detección de objetos, que consigue extraordinarias mejoras respecto al algoritmo de golpe YOLO (you only look once), tanto en el tiempo de ejecución como en ahorro de poder computacional en la identificación y clasificación de animales en imágenes de cámara trampa (Vecvanags *et al.* 2022). Se compara con Timelapse que es un software para etiquetar, analizar y manejar imágenes, integrando revisión por humanos y reconocimiento por inteligencia artificial eficientemente (van Lunteren 2023, Vecvanags *et al.* 2022).

Metodología

Se utiliza el grupo de imágenes "African Wildlife" de Ultralytics para realizar el análisis exploratorio. Representa en su mayoría cuatro especies encontradas en África subsahariana: búfalos, cebras, elefantes y rinocerontes (ver gráfica de pastel). Comparando el éxito de identificación entre RetinaNet y Timelapse.

Se utiliza el tutorial de Gosh (2025, Vecvanags *et al.* 2022) para refinar las detecciones de animales silvestres con RetinaNet (RN) usando PyTorch, replicando la implementación en un notebook de Colab.

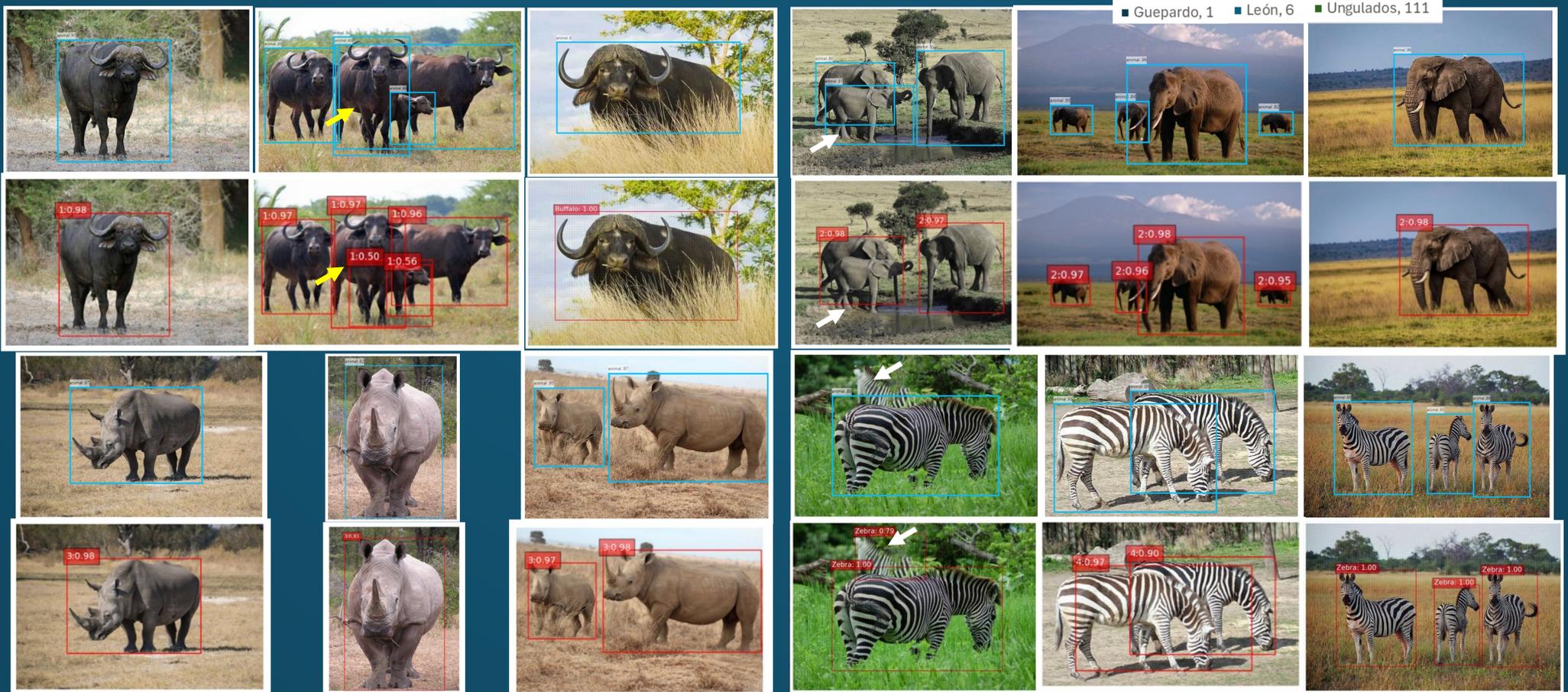
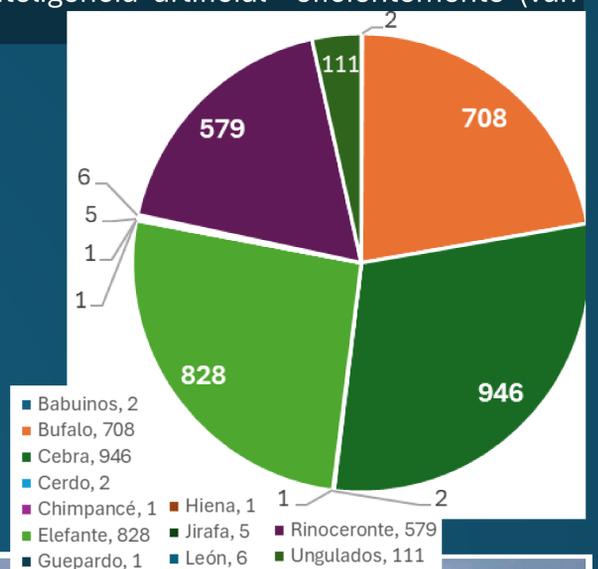
Para Timelapse (TL) se utilizó metodología de Greenberg (2025, van Lunteren 2023).

Resultados

El set completo de imágenes contiene un archivo de reconocimiento inicial (african-wildlife.yaml) y 3 carpetas, con subcarpetas de imágenes y etiquetas en archivo de texto:

- Entrenamiento (1052 imágenes)
- Validación (225 imágenes)
- Prueba (227 imágenes)

En el archivo de texto se da localización aproximada de los mamíferos. Ambos modelos evalúan las identificaciones a priori y convierten las cajas de YOLO en coordenadas absolutas en la fotografía para su caja de identificación. Abajo se muestran ejemplos de las fotografías generadas: caja azul para TL y roja para RN.



El tiempo necesario para trabajar RN fue significativamente más rápido que con TL. La interfaz de usuario es más amigable en TL; RN requiere más conocimiento de codificación para llevar a cabo la evaluación. El espacio de almacenamiento de TL es mayor que RN.

El trabajo previo para obtener el archivo *.yaml es muy importante, mientras más detalle se tenga, mejor será el resultado.

Las flechas blancas muestran subestimación en ambos modelos, NR en cebras y TL en elefantes. Las flechas amarillas muestran sobreestimación en NR en búfalos. Ningún programa es infalible, pero el utilizar dos métodos diferentes, puede servir para tener una mejor estimación de los individuos en estudio.

A pesar de que ambas metodologías buscan automatizar la identificación de objetos de interés en imágenes, se recomienda tener un humano confirmando las identificaciones durante el proceso.

Bibliografía

- Gosh, A. (2025). FineTuning RetinaNet for Wildlife Detection with PyTorch: A Step-by-Step Tutorial. LearnOpenCV. <https://learnopencv.com/finetuning-retinanet/>
- Greenberg, S. (2025). Timelapse Manual Series. Quick Start Guides. Last updated May 1, 2025, Timelapse Version 2.3.3.0. <https://timelapse.ucalgary.ca/Guides/>
- van Lunteren, P. (2023). AddaxAI: A no-code platform to train and deploy custom YOLOv5 object detection models. The Journal of Open Source Software, 8(88): 5581. <https://doi.org/10.21105/joss.05581>.
- Vecvanags, A.; Aktas, K.; Pavlovs, I.; Avots, E.; Filipovs, J.; Brauns, A.; Done, G.; Jakovels, D.; Anbarjafari, G. (2022). Ungulate Detection and Species Classification from Camera Trap Images Using RetinaNet and Faster R-CNN. Entropy, 24: 353. <https://doi.org/10.3390/e24030353>