

1. Participantes:

- Alejandro Aristizabal
- Juan David Rico
- Alexander Hernandez
- Juan Felipe Jiménez

2. Título del Proyecto.

Aerial Animal Detector – identificación de especies y conteo de animales

3. Organización/Grupo de investigación.

Proyecto Guacamaya, CINFONIA: haciendo uso de inteligencia artificial, para estudiar la diversidad amazónica, ha transformado la investigación y conservación ambiental. Los modelos desarrollados por el equipo están disponibles para descarga y son utilizados para estudiar una amplia variedad de especies y ecosistemas. De acuerdo con el profesor Pablo Arbeláez, estos modelos pueden adaptarse a otras regiones y aplicarse a problemas similares en lugares como África o Asia.

Experto: Isai Daniel Chacon Silva. Magister en Ingeniería Biomédica, Universidad de los Andes. Investigador, Laboratorio de Microsoft Al for Good.

Nuestro proyecto se enmarca en el área de biodiversidad y conservación del medio ambiente, específicamente en la gestión de conflictos entre la vida silvestre y el ganado en áreas protegidas de África subsahariana.

4. Contexto, problema y Justificación. Impacto esperado.

Problema y contexto



La conservación de la biodiversidad en África subsahariana es un reto crucial para garantizar la sostenibilidad ecológica y el bienestar de las comunidades locales. Los conflictos entre la fauna silvestre y el ganado han aumentado en las últimas décadas debido a la expansión de la población humana y la creciente demanda de tierras para la ganadería. Estos conflictos generan una competencia directa por recursos como pastizales y fuentes de agua, afectando tanto a la vida silvestre como a los medios de subsistencia de las comunidades rurales. La falta de un monitoreo preciso y continuo de la densidad de animales en estas regiones impide tomar decisiones informadas para equilibrar las necesidades de conservación y desarrollo económico local.

Uno de los principales desafíos en este contexto es la detección y conteo de ganado y fauna silvestre en imágenes aéreas. Los métodos tradicionales, como la observación manual desde aeronaves o el análisis visual de fotografías, presentan serias limitaciones, incluyendo errores en la identificación de especies, dificultades en la detección de animales en manadas densas y altos costos operativos. Además, las técnicas basadas en redes neuronales convolucionales (CNN) han mostrado limitaciones en estos escenarios debido a factores como la oclusión de individuos, la variabilidad en la iluminación, la complejidad de los fondos y las diferencias de escala.

Este proyecto aborda el problema del conteo automático de animales en imágenes aéreas mediante un enfoque basado en aprendizaje profundo. El objetivo principal es diseñar y entrenar un modelo capaz de detectar y contar con precisión animales en manadas densas, superando las limitaciones de los métodos tradicionales. Para ello, se utilizará una arquitectura especializada en tareas de conteo, optimizando la capacidad del modelo para manejar los desafíos inherentes al problema, como la oclusión parcial de los animales y la distribución no uniforme de los mismos en las imágenes.

El impacto de este proyecto será significativo en tres dimensiones clave:

 Ambiental: Contribuirá a la conservación de la biodiversidad al permitir un monitoreo más preciso y frecuente de las poblaciones de fauna silvestre y ganado, facilitando la implementación de estrategias de manejo sostenible.



- Social: Apoyará a las comunidades rurales al proporcionar herramientas para optimizar el uso de los recursos naturales y minimizar los conflictos entre ganaderos y la vida silvestre.
- Tecnológico: Representará un avance en el desarrollo de modelos de aprendizaje profundo adaptados al conteo de animales en entornos complejos, con potenciales aplicaciones en otras áreas de monitoreo ecológico y conservación.

5. Objetivos

Objetivo General. Desarrollar un modelos de aprendizaje profundo para la identificación de especies y el conteo automático de animales en imágenes aéreas, con impactos ambientales, sociales y tecnológicos.

Objetivos Específicos.

- 1. Identificar modelos existentes especializados en la problemática planteada.
- 2. Elegir el modelo que se utilizará como la línea base para este proyecto.
- 3. Definir las métricas para evaluar nuestro modelo frente a los modelos existentes.
- 4. Construir una interfaz de usuario intuitiva para facilitar el uso de nuestro modelo.
- 5. Evaluar resultados de nuestro modelo.
- 6. Documentar las conclusiones identificadas.

6. Estado del arte.

El avance en el monitoreo de fauna silvestre mediante imágenes aéreas ha avanzado significativamente, en gran medida debido al desarrollo de técnicas de aprendizaje profundo. A continuación presentamos algunas de las principales tecnologías y modelos más relevantes:

Redes Neuronales Convolucionales (CNN): Las CNN han demostrado ser altamente efectivas en tareas de detección y clasificación de objetos en imágenes. Su capacidad para aprender características espaciales las hace ideales para el reconocimiento de patrones en



datos visuales. Estas redes han sido aplicadas exitosamente en la identificación y conteo automático de animales silvestres, alcanzando precisiones superiores al 93% en algunos casos.

Modelos de Detección de Objetos: Uno de los modelos más avanzados es el de Faster R-CNN, que ha sido reconocido por su eficacia en la detección de objetos en tiempo real. Combina propuestas de regiones con redes neuronales convolucionales para mejorar la precisión y la velocidad de la detección.

Aprendizaje por transferencia: Esta técnica consiste en adaptar modelos previamente entrenados en grandes conjuntos de datos a tareas específicas con menos datos disponibles. Por ejemplo, se ha utilizado el aprendizaje por transferencia para la clasificación de animales en imágenes, mejorando el rendimiento en escenarios con datos limitados.

Redes Neuronales para Datos 3D: Con el uso de sensores como LiDAR y Radar, es posible obtener nubes de puntos que representan escenas en tres dimensiones. Para procesar estos datos, se han desarrollado arquitecturas especializadas que permiten la detección y clasificación de objetos en entornos tridimensionales, ampliando las posibilidades de monitoreo de fauna en hábitats complejos.

Modelos Preentrenados: El uso de modelos de aprendizaje profundo previamente entrenados facilita la implementación de soluciones sin requerir grandes volúmenes de datos de entrenamiento. Estos modelos pueden reconocer formas, patrones y texturas complejas, acelerando los flujos de trabajo geoespaciales y la extracción de características de imágenes.

Desarrollo de aplicaciones específicas: Se han desarrollado sistemas especializados, como el "Rastreador de Gorilas", que emplea inteligencia artificial para monitorear y proteger especies en peligro de extinción. Este proyecto utiliza cámaras y redes neuronales para analizar el comportamiento y la salud de los gorilas, permitiendo la detección temprana de enfermedades y contribuyendo a su conservación.

Referencia 1: Montañez Sogamoso, Ó., Roa Martín, C., & Avendaño Fernández, E. (2023). Estado de arte y aplicación de técnicas de aprendizaje profundo desde imágenes



reconstruidas a partir de nubes de puntos de sensores de radar y Lidar. Revista de Ingeniería.

Referencia 2: SIGSA. (s.f.). *Modelos de aprendizaje profundo previamente entrenados*. Recuperado de https://www.sigsa.info/es-mx/arcgis/deep-learning-models

Referencia 3: Deleuze, I. (2024, septiembre 5). *Un proyecto de IA pionero en el zoo de Berlín busca proteger a los gorilas, en peligro de extinción*. Cadena SER. Recuperado de https://cadenaser.com/nacional/2024/09/05/un-proyecto-de-ia-pionero-en-el-zoo-de-berlin-busca-proteger-a-los-gorilas-en-peligro-de-extincion-cadena-ser/

Principales desafíos: A pesar que se han dado avances importantes, persisten algunos desafíos que dificultan la precisión de los modelos, algunos de los más importantes son:

- Oclusión de animales: cuando uno o más animales están parcial o totalmente bloqueados por otros animales, objetos o elementos del entorno en la imagen. Las oclusiones dificultan la identificación y el conteo precisos.
- Fondos complejos: Las imágenes aéreas suelen capturar entornos con una variedad de elementos como vegetación, sombras, cuerpos de agua y otros objetos naturales. Estos fondos heterogéneos pueden confundir a los algoritmos de detección, que podrían tener dificultades para distinguir entre los animales y el entorno circulante.
- Variaciones de escala: Los animales pueden aparecer en diferentes tamaños en las imágenes debido a factores como la altitud de la toma, la posición relativa al sensor y las diferencias intrínsecas de tamaño entre especies. Estas variaciones de escala representan un reto para los modelos de detección, que deben ser capaces de identificar animales tanto grandes como pequeños con igual precisión.

7. Metodología y enfoque propuesto

Enfoque técnico.

Para abordar el problema de detección y conteo de animales en manadas densas a partir de imágenes aéreas, hemos adoptado un enfoque basado en el modelo preentrenado



HerdNet, una arquitectura optimizada para la detección de fauna en imágenes de alta densidad.

Técnicas de IA y su justificación:

A. HerdNet como modelo base:

- HerdNet ha demostrado un rendimiento robusto en tareas de conteo y clasificación de animales en imágenes aéreas, con resultados comparables a los obtenidos por métodos manuales.
- Se basa en redes neuronales convolucionales (CNN) especializadas en la segmentación y detección de objetos, lo que permite manejar desafíos como oclusiones, fondos complejos y variaciones de escala.

B. Transformaciones y preprocesamiento de datos:

- Se aplican técnicas de normalización, redimensionado y aumentación de datos (como rotaciones aleatorias y ajustes de brillo/contraste) para mejorar la capacidad generalizable del modelo.
- Uso de la herramienta Patcher para segmentar las imágenes en fragmentos más pequeños, facilitando el procesamiento y reduciendo el impacto de los fondos complejos.

C. Conversión de etiquetas de bounding boxes a puntos:

• Se optó por marcar a los animales mediante puntos en lugar de cajas delimitadoras, dado que esta técnica ha demostrado ser más efectiva en escenarios con alta densidad de individuos y resolución limitada.

D. Función de pérdida:

• Se emplean Focal Loss y Cross Entropy Loss, con un ajuste de pesos para abordar el desbalance en la distribución de clases de especies de animales.

E. Registro de experimentos y gestión de versiones en MLFlow:



 Se almacena y rastrea cada versión del modelo entrenado, permitiendo la comparación de resultados y la selección de la mejor configuración basada en métricas de desempeño.

Metodología de Desarrollo:

El desarrollo de la solución sigue un enfoque iterativo basado en experimentación y mejora continua:

A. Recolección y Preprocesamiento de Datos:

- Se descarga el dataset de imágenes público.
- Se aplicarán transformaciones y segmentación de imágenes.
- Conversión de anotaciones de bounding boxes a puntos de referencia.

B. Entrenamiento del Modelo:

- Ajuste inicial del modelo **HerdNet**, manteniendo la arquitectura base.
- Optimización con Adam y ajuste de la tasa de aprendizaje.

C. Validación y Evaluación:

- Medición de desempeño en el conjunto de validación mediante métricas F1-Score, Precisión y Recall.
- Comparación con los resultados de HerdNet para evaluar mejoras.

D. Ajuste y Experimentación:

- Implementaremos diferentes estrategias de fine-tuning.
- Analizaremos el impacto de la segmentación de imágenes en el rendimiento del modelo.
- Refinaremos los hiperparámetros para mejorar la capacidad de generalización del modelo.

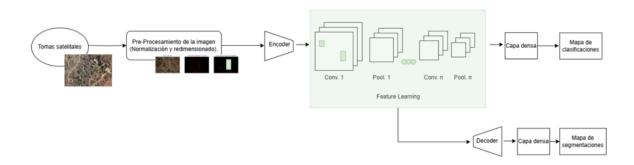
E. Implementación y Despliegue:





- Integraremos el modelo con una interfaz de usuario para la carga y análisis de imágenes aéreas.
- Utilizaremos MLFlow y Github para almacenar las mejores versiones del modelo y facilitar su reproducción

Maqueta de la solución propuesta:



Herramientas.

Para el desarrollo de esta solución contemplamos el uso de las siguientes herramientas que podrán ser modificadas, en caso que encontremos algunas más adecuadas durante el refinamiento de los parámetros del modelo:

- Lenguaje de programación: Python.
- Framework de aprendizaje profundo: PyTorch.
- Bibliotecas de procesamiento de imágenes: Albumentations, OpenCV, Matplotlib.
- Herramientas de manejo de datos: Pandas, Numpy.
- Modelo preentrenado: HerdNet.
- Plataforma de gestión de experimentos: MLFlow.
- Plataforma de almacenamiento de versiones de modelos: AWS S3
- Repositorio de código: GitHub.
- Herramientas de evaluación de modelos: Animaloc (HerdNet Evaluator)

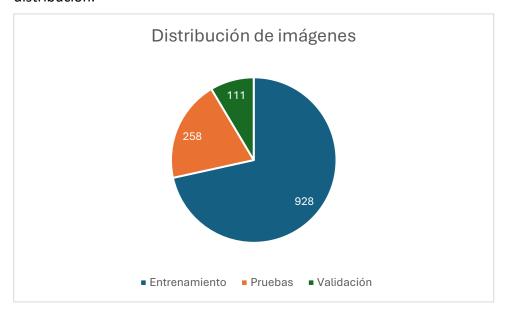
8. Exploración y descripción de los datos



Características de los datos

Los datos que utilizaremos provienen de la siguiente fuente: https://dataverse.uliege.be/file.xhtml?fileId=11098&version=1.0

Los datos tomados de dicha fuente están diseñados específicamente para tareas de conteo de animales en manadas densas utilizando imágenes aéreas y ya se encuentran organizadas en tres conjuntos: entrenamiento, validación y prueba con la siguiente distribución:



Las imágenes están distribuidas en los siguientes porcentajes:

Entrenamiento: 71%, Pruebas: 20% y Validación: 9%

Especies presentes en los datos:

1. Topis(Damaliscus lunatus jimela)



Fuente: https://es.wikipedia.org/wiki/Damaliscus_lunatus_jimela

2. Buffalos (Syncerus caffer)



Fuente: https://www.naturepl.com/stock-photo-african-buffalo-nature-image01320863.html

3. Kobs (kobus kob)





 $\textbf{Fuente:}\ \underline{https://www.ugandarwandagorillatours.com/safaris-blog/the-kob.html}$

4. Warthogs (Phacochoerus africanus)



Fuente: https://www.jungledragon.com/specie/367/common_warthog.html

5. Waterbucks (kobus ellipsiprymnus)



Universidad de los Andes

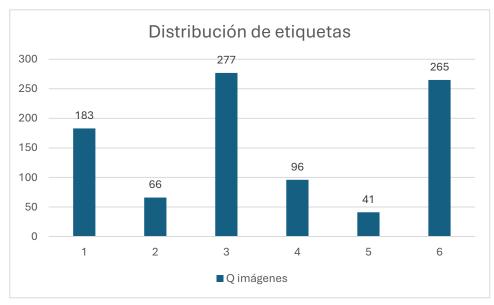
6. Elephants (Loxodontaafricana)



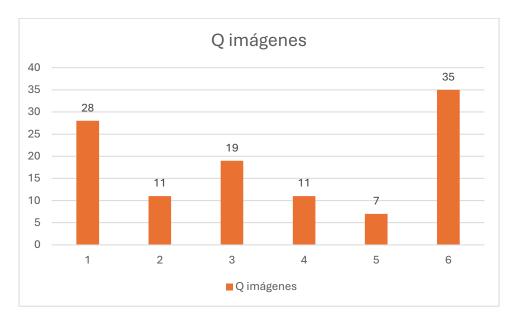
Fuente: https://www.shutterstock.com/es/search/3-elephants

Distribución de imágenes:

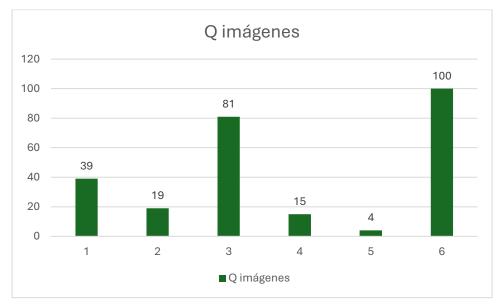
1. Entrenamiento: 928 imágenes



2. Validación: 111 imágenes



3. Prueba: 258 imágenes

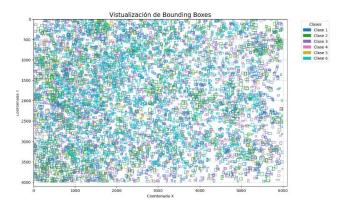


Despliegue de imágenes:

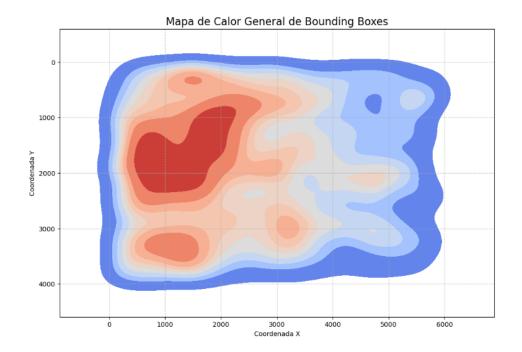


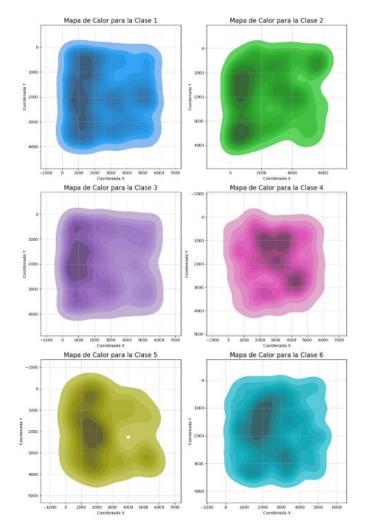
Universidad de los Andes

Complementando el ejercicio descriptivo de los datos, el equipo de trabajo hizo un estudio del despliegue que tendría la ubicación de los animales en las imágenes de manera acumulada. Se hicieron cálculos adicionales para calcular los rectángulos, o bounding boxes, que rodean al animal según el archivo .cvs de la fuente de datos, y que contiene las coordenadas de cada animal por imagen. Este ejercicio se hizo con el propósito de identificar si existía alguna zona de las imágenes en la que de manera consistente se ubicaran más animales, sin embargo el resultado fue negativo ya que se evidencia una alta dispersión de animales en todas las imágenes y no es posible definir una única zona de interés como se evidencia a continuación:



Adicionalmente, y con el fin de soportar la verificación de algún patrón de ubicación que pueda incidir en las imágenes, se hizo un mapa de calor para verificar las mayores concentraciones de ubicaciones de animales según las coordenadas, el resultado fue el siguiente:



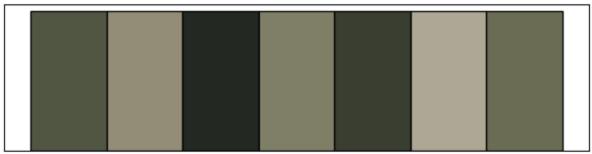


Teniendo en cuenta el objetivo del modelo de identificar la especie del animal y contar la cantidad de animales en la imagen, la alta dispersión de animales en todas las imágenes, no permite definir una única zona de interés por lo que el modelo deberá analizar las imágenes completas.

Paleta de colores:

A continuación se contempla la paleta de los 7 colores más representativos de las imágenes de validación en acumulado:

Paleta de los 7 colores más representativos



Este resultado se obtuvo luego de realizar un ejercicio de *K-Means*, agrupando los colores más frecuentes en las imágenes. Se tomó el grupo de imágenes de validación para el análisis para controlar la necesidad de capacidad de cómputo.

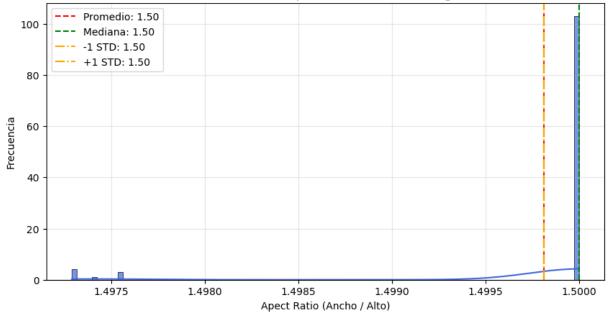
Se pueden observar principalmente colores verdes oscuros, marrones y grises, lo que coincide con que el entorno de las imágenes contiene mucha vegetación, suelo y sombras, y que al tiempo reflejan la apariencia de los paisajes donde se capturan las imágenes (bosques, zonas semidesérticas, etc). Estos colores serán de una alta relevancia para el entrenamiento del modelo ya que serán los que permitirán identificar los animales, clasificarlos en la clase correspondiente y diferenciarlos de los colores de fondo.

Aspect ratio:

Como parte del ejercicio de la exploración de los datos, el equipo realizó un análisis sobre la distribución del aspect ratio para verificar cómo se distribuyen las proporciones de ancho y alto de las imágenes. Como se evidencia en la gráfica a continuación, hay una baja variabilidad en el aspect ratio de las imágenes:



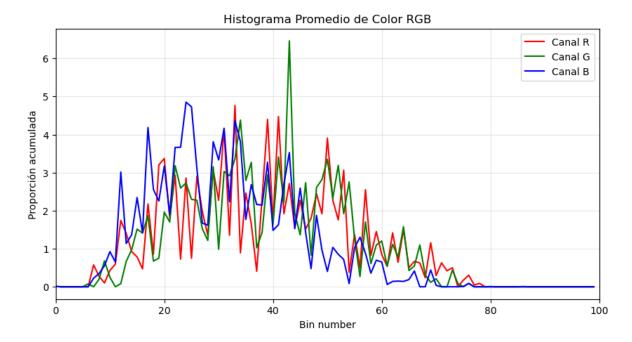




El resultado de un aspect ratio relativamente homogéneo, minimiza la distorsión causada por diferencias de escala o forma, disminuyendo la necesidad de normalización de los datos y disminuyendo la posibilidad de generar sesgos en el modelo a causa de los ratios de las imágenes.

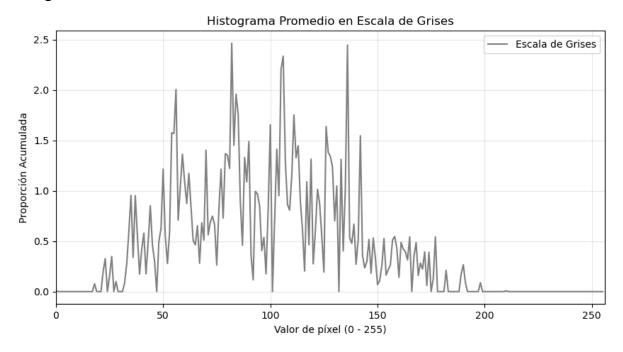
Histograma Promedio de los Canales RGB:





Con la gráfica anterior podemos concluir que ninguno de los colores parece ser completamente dominante en toda la imagen lo cual es positivo ya que ayuda a disminuir sesgos por estos motivos, sin embargo hay picos en donde un canal se eleva por encima de los demás, indicando regiones de la imágen donde ese color es más fuerte. Hay un pico grande alrededor de los bins 40-50 en el canal verde, lo cual es consistente con una alta presencia de vegetación o fondos naturales en las imágenes.

Histograma Promedio de Grises:



Se observa que la mayoría de los píxeles tienen valores entre 50 y 150, lo que sugiere que el dataset no es ni demasiado oscuro ni demasiado claro. Esto indica una buena distribución de luminosidad en las imágenes.

Por otra parte, hay pocos valores en los extremos, lo que significa que hay pocos píxeles completamente negros o blancos. Esto es bueno porque evita la sobreexposición o subexposición en las imágenes.

La presencia de picos irregulares entre 50 y 150 sugiere que hay contrastes y variaciones en la iluminación de las imágenes, lo que indica diferentes tipos de escenarios (ej. sombras, iluminación natural, fondos mixtos), esto evitará sesgos hacia imágenes demasiado oscuras o demasiado claras.

Preparación de datos

Para asegurar un entrenamiento y validación eficiente del modelo se contemplan las siguientes técnicas de limpieza y preparación de los datos:

A. Exploración y Análisis de Calidad:

- Se analizaron las distribuciones de especies en los conjuntos de entrenamiento, validación y prueba.
- Se validó la dispersión de los animales en las imágenes, concluyendo que el modelo debe analizar las imágenes completas debido a la ausencia de zonas de alta concentración.

B. Transformaciones propuestas:

- Segmentación de imágenes: utilizando la herramienta Patcher para dividir las imágenes en fragmentos más pequeños y optimizar el procesamiento.
- Se hará uso de rotaciones, ajustes de brillo/contraste y normalización para mejorar la robustez del modelo.
- Transformación de bounding boxes en puntos centrales dado que se ha identificado que funcionan mejor para el conteo y la segmentación, dadas las condiciones de oclusión y variabilidad en las imágenes de animales.

9. Consideraciones Éticas

Impacto social y ético

El modelo tiene el potencial de mejorar significativamente la gestión de la biodiversidad en África al proporcionar herramientas más precisas para el monitoreo de fauna silvestre.

- Teniendo en cuenta que el dataset proviene de fuentes públicas, se deben respetar los términos de su licencia, asegurando que el modelo y sus resultados no se utilicen con fines comerciales sin la debida autorización.
- Es importante que los organismos responsables de la conservación interpreten correctamente los resultados y los utilicen como una herramienta de apoyo en lugar de una solución automatizada definitiva.
- Existe el riesgo que tecnologías avanzadas de monitoreo sean utilizadas con propósitos no éticos como la caza furtiva.

Uso responsable de la IA



Posibles sesgos en los datos:

- Distribución desigual de especies, algunas especies están sobrerrepresentadas en los datos de entrenamiento, lo que afecta la capacidad del modelo para detectar especies menos comunes. –
 Contemplaremos técnicas de balanceo en los datos de entrenamiento para garantizar una distribución equitativa de las especies.
- El fondo y las condiciones de iluminación pueden hacer que ciertas especies sean más fáciles de detectar en comparación con otras. Aplicación de aumentación de datos, incluyendo modificaciones en el brillo y contraste, para mejorar la generalización del modelo.
- Diferencias en la calidad de las imágenes, pueden generar algunos sesgos en el modelo, si esto se presenta de manera repetitiva para alguna de las especies. – Posibilidad de explorar métricas separadas por especie para asegurar que el modelo tenga un rendimiento equitativo en todas las categorías.

Transparencia y explicabilidad

Para garantizar la transparencia del modelo y la explicabilidad de sus resultados, se han considerado las siguientes estrategias:

- A. Registro detallado de versiones: Utilizaremos MLFlow para documentar cada iteración del modelo, incluyendo hiperparámetros, métricas de desempeño y evaluaciones.
- B. Visualización de resultados: Se incluirán indicadores de confianza en cada predicción para comprender el nivel de certeza del modelo.
- C. Información para usuarios finales: Se contará con una interfaz de usuario que presentará los resultados más importantes de manera accesible para usuarios no técnicos.

Privacidad y seguridad

Para este proyecto utilizaremos datos de imágenes aéreas provenientes de fuentes públicas, no se manejan datos personales ni información confidencial. Sin embargo, se tendrán en cuenta las siguientes medidas para garantizar el uso responsable de los datos:



- A. Cumplimiento de licencias y restricciones de uso: Todos los datos utilizados provienen de fuentes abiertas y están sujetos a la licencia CC BY-NC-SA 4.0, es decir que no pueden usarse para fines comerciales y deben citarse correctamente.
- B. Almacenamiento seguro de datos: Se utilizarán servicios en la nube seguros (como AWS S3).

10. Evaluación y Métricas

De acuerdo con lo mencionado anteriormente, nuestro modelo toma como referencia el modelo creado por HerdNet por lo que nuestra propuesta de modelo busca igualar o superar las métricas alcanzadas por dicho modelo:

F1-Score (%)	MAE	RMSE	AC (Confusión Promedio) (%)
83.5	1.9	3.6	7.8

11. Cronograma

Este proyecto se empezó a desarrollar en la clase de Proyecto – Desarrollo de Soluciones por lo que hay algunas actividades que ya se han venido realizando y servirán de insumo para esta segunda fase, a continuación compartimos el cronograma tentativo:



11. Bibliografía

- Delplanque et al., 2022, Eikelboom et al., 2019, Kellenberger et al., 2017, Kellenberger et al., 2018, Kellenberger et al., 2019a, Naudé and Joubert, 2019, Peng et al., 2020, Torney et al., 2019
- Gao et al., 2020
- Lempitsky and Zisserman (2010)
- Li et al., 2021, Liu et al., 2018
- https://www.uniandes.edu.co/es/noticias/ingenieria/inteligencia-artificial-para-la-proteccion-del-amazonas
- Montañez Sogamoso, Ó., Roa Martín, C., & Avendaño Fernández, E. (2023). Estado de arte y aplicación de técnicas de aprendizaje profundo desde imágenes reconstruidas a partir de nubes de puntos de sensores de radar y Lidar. *Revista de Ingeniería*.
- SIGSA. (s.f.). *Modelos de aprendizaje profundo previamente entrenados*. Recuperado de https://www.sigsa.info/es-mx/arcgis/deep-learning-models



- Deleuze, I. (2024, septiembre 5). *Un proyecto de IA pionero en el zoo de Berlín busca proteger a los gorilas, en peligro de extinción*. Cadena SER. Recuperado de https://cadenaser.com/nacional/2024/09/05/un-proyecto-de-ia-pionero-en-el-zoo-de-berlin-busca-proteger-a-los-gorilas-en-peligro-de-extincion-cadena-ser/