

INSTITUTO TECNOLÓGICO DE BUENOS AIRES (ITBA)
ESCUELA DE GESTIÓN Y TECNOLOGÍA

**Rewilding: Aplicación de Inteligencia
Artificial para la Conservación de la
Biodiversidad**

AUTORES:

Noguera, Abril (61541)
Arbués, Lucas (61890)
de Achával, Ignacio (60696)

DOCENTES:

Juan Pablo Rodriguez Varela

**TRABAJO FINAL PRESENTADO PARA LA OBTENCIÓN DEL TÍTULO
DE LICENCIADO EN ANALITICA EMPRESARIAL Y SOCIAL**

Buenos Aires

18 de Diciembre 2023

Índice:

Definición del Problema	2
Definición del Alcance	4
Plan de Mejora y KPIs	4
Estado del Arte	6
Abordaje del Problema	7
Outputs del Proyecto	9
Plan de Trabajo	10
Etapas del Proyecto	10
Metodología de Trabajo Ágil	12
Análisis de Riesgos	13
Datos a Utilizar	15
Herramientas a Utilizar	16
Análisis Exploratorio de Datos (EDA)	17
Disparadores de Exploración	17
Limpieza y Preprocesamiento	18
Análisis Descriptivo	20
Análisis Exploratorio	24
Exploración de Información no Procesada	29
Conclusiones del Análisis Exploratorio	32
Business Case	34
Hipótesis de Trabajo	34
Análisis de Situación Actual	34
Justificación del Proyecto	38
Escenario Automatizado	39
Escenario Semi-Automatizado	40
Resultados Potenciales	42
Enfoque de Solución	42
Metodologías a Implementar	45
Seteo de Experimentación	45
Desarrollo de la Solución	48
Tarea 1: Detección de Imágenes con Animales.	48
Tarea 2: Detección de Especies.	49
Tarea 3: Recuento de Cantidad.	51
Proceso Semi-Automatizado	54
Resultados y Conclusiones del Desarrollo	55
Tiempo Total del Proceso de Etiquetado	55
Cantidad Promedio de Imágenes Analizadas por Hora	56
Precisión de Etiquetado	57
Costo del Proceso	57
Ratio de Imágenes Útiles	58
Conclusiones	59
Futuras Direcciones	60
Anexo	62
Bibliografía	70

Definición del Problema

La *Fundación Rewilding Argentina* constituye una comunidad de conservacionistas que trabaja con el objetivo de revertir la crisis de extinción de especies. La fundación se enfoca en el concepto de "rewild", que implica la **restauración de ecosistemas naturales y la reintroducción de especies nativas** para restablecer los procesos ecológicos y la biodiversidad.

Mientras esta organización pone en marcha iniciativas de conservación, un panorama desafiante se revela en el horizonte global. Un informe presentado por la ONU en 2020 (Convenio sobre la Diversidad Biológica (CBD), 2020) postula la inminencia de una sexta extinción, proyectando así una advertencia crítica sobre la futura biodiversidad del planeta. Indicios actuales advierten una **reducción masiva de la biodiversidad**, culminando en la desaparición del 10% al 20% de todas las formas de vida en el planeta.

Rewilding tiene proyectos de restauración que apuntan a recuperar la funcionalidad de los ecosistemas naturales. Estas iniciativas se detallan en la **Figura 1**. Las acciones de restauración y conservación incluyen translocaciones de individuos autóctonos, manejo de pastizales mediante quemas prescriptas, enriquecimiento de bosques degradados, extracción de cercos y cierre de caminos innecesarios, control de especies exóticas y remoción de ganado doméstico. Actualmente, la fundación cuenta con proyectos en Chubut, Corrientes, Chaco y Santa Cruz, los mismos tienen alcance terrestre y marítimo.

Con estas acciones en curso, es crucial medir la efectividad de las intervenciones **estimando la abundancia de las especies** en cada zona de acción para observar si la vida silvestre regional está presente en cantidades suficientes para cumplir sus roles ecológicos, lo cual servirá como base para futuras acciones conservacionistas.



Figura 1: Diagrama de Proyectos de Rewilding Argentina.

Las *transectas* utilizadas para contabilizar la fauna en un ecosistema natural constan de rutas predefinidas que se recorren minuciosamente para registrar y cuantificar la presencia de diferentes especies animales en un área determinada. Este método de muestreo es ampliamente utilizado en estudios de biodiversidad y ecología para

obtener datos sobre la distribución y abundancia de la fauna. Durante el recorrido de la transecta, los investigadores anotan observaciones detalladas de los animales que encuentran, registrando información como especie, cantidad, ubicación y comportamiento.

"Cuando entras en un ecosistema natural, básicamente **te pierdes el 90% de los animales que allí hay**" dice Jorge Ahumada, conservacionista de vida silvestre. Pero incluso si no se pueden ver, los animales están allí, permaneciendo tranquilos y escondidos hasta que los humanos se alejen para poder viajar y cazar. Estos factores pueden introducir cierto grado de sesgo o limitaciones en la recopilación de información.

Gracias a los avances tecnológicos de la modernidad, con la ayuda de cámaras trampa, los expertos en conservación pueden observar la vida silvestre para medir y estimar la abundancia de las especies sin la necesidad de estar presentes, mitigando así el factor de ahuyentamiento de los animales. *Rewilding Argentina* realiza anualmente proyectos de captura de imágenes con cámaras trampa con una duración de 30 días en cada una de sus regiones de presencia. Tan solo en una de las zonas de captura se pueden tomar 120,000 imágenes. Aquí es donde se presentan barreras al trabajo humano.

En cualquier proyecto de conservación que involucra imágenes de cámaras trampa, gran parte del trabajo proviene de **examinar manualmente miles de imágenes e identificar especies**. Es posible que hasta el 80% de las fotografías no contengan animales dado a que la cámara fue activada por elementos como el movimiento de la flora o el viento. Aparecen, además, diversos factores que potencian el error humano, como la oscuridad de ciertas imágenes, la superposición de árboles o arbustos por delante de los animales o la fatiga del repetitivo trabajo de observar una y otra vez imágenes similares. A modo de resumen, el procesamiento de la información actualmente ocupa semanas de trabajo y además está expuesto a errores humanos implícitos en la clasificación manual.

El presente proyecto nace como respuesta a la necesidad de optimizar dicho proceso, combinando la precisión y velocidad de la inteligencia artificial con la agudeza analítica humana. Se propone la implementación de una plataforma integrada con modelos de inteligencia artificial que permite a la organización cargar, clasificar y compartir los datos obtenidos por las cámaras. Por lo que el agregado de valor se ve reflejado en la **aceleración drástica del proceso** de clasificación de las imágenes y en la mejora de la representación de las mismas dando un salto de **calidad en la precisión** del etiquetado.

Tener información precisa, detallada y actualizada sobre la ubicación y comportamiento de los animales en su hábitat natural potencia significativamente la capacidad para estudiar y conservar ecosistemas. Los investigadores de vida silvestre pueden evaluar mejor qué tan abundantes son las especies, dónde están ubicadas y cómo se mueven de un lugar a otro. Los responsables de la formulación de políticas pueden utilizar conocimientos para elegir dónde y cómo cambiar los límites de las áreas protegidas. Basar sus decisiones acerca de los próximos proyectos de restauración en información confiable para así lograr los objetivos de la

fundación. **Usar la inteligencia artificial para recuperar poblaciones de animales en situación crítica.**

Definición del Alcance

El análisis y la herramienta final están desarrollados en base a los datos recibidos que refieren a la región de Patagonia Azul, en Camarones, Chubut. En esta región existen dos zonas: Sauce e Ibérica. Los resultados obtenidos en este área servirán como base para evaluar la viabilidad de una posible expansión del proyecto a todas las zonas abarcadas por Rewilding Argentina, en respuesta al interés expresado por la organización.

Plan de Mejora y KPIs

El proceso de conservación se enfrenta al desafío de filtrar y analizar eficazmente grandes volúmenes de datos visuales, un procedimiento tradicionalmente lento y propenso a errores.

En respuesta a esta situación, se implementarán Indicadores Clave de Desempeño (KPIs) que permitirán evaluar y comparar el estado actual del proceso con las propuestas de mejora proyectadas. Estos KPIs se establecerán como referencias esenciales para la optimización continua y el aumento de la exactitud operativa en el proyecto.

Los Indicadores de Desempeño (KPIs) son:

- **Tiempo Total del Proceso de Etiquetado:** Este parámetro mide el tiempo requerido para procesar las imágenes, expresado en jornadas laborales ocupadas. Actualmente, el tiempo dedicado al proceso de clasificación de las fotografías abarca 13 jornadas laborales según evaluaciones históricas, en las zonas denominadas como "Sauce" e "Ibérica". Se busca implementar estrategias que permitan reducir significativamente este período. El objetivo es liberar recursos humanos especializados, optimizando así su labor, dada la asignación a otras funciones críticas dentro de la organización.
- **Cantidad Promedio de Imágenes Analizadas por Hora:** Estudios relacionados (Fennell et al., 2022) indican que la tasa de clasificación manual de un humano oscila entre 300 y 1000 imágenes por hora. La velocidad de 300 imágenes por hora se observa en escenarios donde cada imagen exhibe fauna, demandando un análisis meticuloso. En contraste, en contextos donde las fotografías no exhiben animales, la velocidad puede ascender hasta 1000 imágenes por hora, dado que se obvia el requerimiento de una clasificación rigurosa. En base a esto se determinará cuántas veces más rápida es la inteligencia artificial en realizar este proceso.
- **Precisión en la Clasificación de Imágenes:** La clasificación errónea, como se ilustra en la **Figura 2**, puede originarse debido a dos fenómenos principales: el primero es la identificación incorrecta de animales donde no existen, un fenómeno conocido como 'Falso Positivo'. El segundo es la no detección de la

presencia de animales cuando en realidad están presentes, lo cual se denomina 'Falso Negativo'.



Figura 2: Matriz de Confusión.

La eficacia de los modelos de clasificación se evaluará considerando ambos tipos de errores potenciales. Este enfoque tiene como objetivo minimizar la tasa de error inherente a la clasificación realizada por humanos, mejorando así la precisión global del modelo.

Para determinar la tasa de error humano en el proceso de clasificación manual actual, se recurrirá al análisis de estudios existentes que abordan las dificultades inherentes a la clasificación de especímenes mediante cámaras trampa.

- En análisis realizados (Norouzzadeh et al., 2018) se establece una tasa de error específica para cada aspecto del proceso de clasificación: detección de presencia animal, identificación de especies y cuantificación. Este estudio asigna un valor de precisión del 96.6% a la clasificación de presencia y especies. En cuanto al conteo, la precisión se determina en un 90%.
- Investigaciones adicionales (Johansson et al., 2020) estiman que existe un 12.5% de error en la clasificación de la presencia animal.
- El trabajo (Chesser M., 2012) se centra específicamente en los errores asociados con falsos negativos y falsos positivos. Este estudio concluye que los falsos negativos presentan una mayor probabilidad de ocurrencia, estimada en un 16.7%. Por otro lado, la incidencia de falsos positivos se evalúa en un nivel significativamente menor, con una probabilidad estimada de solo 0.1%.

A lo largo de este proyecto, se efectuará un seguimiento meticoloso de la tasa de error asociada al proceso de etiquetado utilizado por *Rewilding*, con el fin de establecer un benchmark preciso para la optimización continua de la precisión en la clasificación.

- **Costo del Proceso:** El valor de la jornada laboral del empleado encargado del proceso de clasificación es de \$10,730 (Noviembre, 2023). Además, se asocia un valor al error de clasificación (**Tabla 4**). Estos costos serán contrastados con los que conlleva la solución proyectada. Se define un paquete que involucra el mantenimiento, los recursos de profesionales, comunicación periódica y capacitación del personal por \$1,000,000 (Noviembre, 2023) anual. Además, se asigna un 10% adicional del presupuesto total como reserva para contingencias. Se debe tener en cuenta que si en un futuro se expandiera la solución a otras regiones estos costos serían aprovechados por todas, por lo que estarían justificados.
- **Ratio de Imágenes Útiles:** Las imágenes que no generan valor para *Rewilding* son aquellas donde no se presentan animales y aquellas que presentan un animal repetido en una secuencia, cómo se expresa en la **Figura 3**.



Figura 3: Representación de Imágenes Útiles

Se realizará un seguimiento para definir el benchmark de la utilidad de las imágenes según el proceso actual de *Rewilding*. La implementación del procesamiento semi-automatizado propuesto no solo aumentará el porcentaje de imágenes valiosas al potenciar la precisión de los modelos predictivos, sino que también convertirá la totalidad de las imágenes en un recurso valioso. Esto no solo beneficiará las áreas de estudio actuales, sino que también proporcionará una base sólida para la expansión de las actividades de conservación de *Rewilding* a nuevas regiones.

Estado del Arte

Existen métodos, modelos e investigaciones de vanguardia con respecto a la problemática de clasificación de imágenes de cámaras trampa, entre ellos se destacan:

- **Modelo “Megadetector” de Microsoft**

El análisis automatizado de imágenes de cámaras trampa representa un campo de investigación emergente y vital para la biología de la conservación. Dentro de este contexto, el modelo MegaDetector se posiciona como un benchmark significativo. Desarrollado por Microsoft AI for Earth, este modelo de detección de objetos basado en aprendizaje profundo está entrenado en un vasto conjunto de imágenes de cámaras trampa, lo que lo capacita para identificar de manera efectiva la presencia de animales, personas y vehículos en diversas condiciones ambientales y geográficas.

MegaDetector utiliza una red neuronal convolucional (CNN) y ha sido ampliamente validado en múltiples estudios. El modelo ha logrado una

precisión de detección de animales con una tasa de hasta el 99% y un recall del 92% a un umbral de confianza del 90%, según pruebas realizadas en diversos contextos ecológicos.

- **Uso de Detección de Objetos para la Clasificación de Imágenes en Cámaras Trampa (Fennell et. al. 2022)**

Este estudio presenta una metodología innovadora aplicando el modelo MegaDetector para clasificar automáticamente imágenes de cámaras trampa. El modelo demostró una precisión y recall destacables en la identificación de humanos y animales, con una precisión del 99% y un recall del 95% para la detección de humanos, y un 82% de precisión y 92% de recall para la detección de animales, a un umbral de confianza del 90%.

- **Uso del Aprendizaje Profundo para la Identificación, Conteo y Descripción Automáticos de Animales Salvajes en Imágenes de Cámaras Trampa (Norouzzadeh et al. 2018)**

Este estudio desarrolló un modelo innovador basado en redes neuronales convolucionales (CNNs) para analizar imágenes de cámaras trampa. Su enfoque utiliza aprendizaje profundo para identificar, contar y describir automáticamente animales salvajes capturados en estas imágenes. El estudio destaca la capacidad del modelo para procesar grandes volúmenes de datos y clasificar diferentes especies animales con alta precisión, ofreciendo una solución eficaz al desafío tradicional de clasificación manual que es tanto laborioso como propenso a errores. Los resultados obtenidos mostraron que el modelo podría detectar animales con una precisión de hasta el 96.6%.

- **Wildlife Insights**

Esta es una plataforma tecnológica avanzada y colaborativa destinada a transformar la manera en que se recopilan, analizan e interpretan los datos provenientes de cámaras trampa en estudios ecológicos y de conservación.

Su principal innovación radica en la aplicación del aprendizaje automático para la identificación de especies, lo que reduce drásticamente el tiempo y esfuerzo tradicionalmente requeridos en el proceso de clasificación manual. Además, al ofrecer una base de datos global y accesible, Wildlife Insights promueve la colaboración y el intercambio de información entre investigadores, fortaleciendo así los esfuerzos colectivos de conservación.

Las herramientas integradas de análisis y visualización facilitan la interpretación de tendencias y distribuciones de especies, convirtiendo a esta plataforma en un referente moderno en la investigación y conservación de la biodiversidad.

Abordaje del Problema

El proyecto se centra en la implementación de metodologías de Redes Neuronales Convolucionales para enfrentar el desafío de clasificar y analizar grandes volúmenes

de imágenes de cámaras trampa. Con ese objetivo, es necesario un enfoque metódico desde la adquisición inicial de datos hasta la validación y confirmación de los resultados obtenidos.

El banco de imágenes y los datos clasificados por *Rewilding* son almacenados en la infraestructura de *Amazon Web Services*, garantizando accesibilidad y colaboración eficiente entre los miembros del equipo del proyecto. Se efectúa una meticulosa limpieza de datos para asegurar la coherencia y precisión en la correspondencia entre imágenes y registros. Además, se lleva a cabo un análisis descriptivo y exploratorio profundo de la información mediante las herramientas *Python* y *Power BI* para extraer perspectivas relevantes.

A través de consultas con el equipo de *Rewilding* y un análisis exhaustivo de los registros existentes, se han identificado tres procesos operativos fundamentales susceptibles de ser mejorados a través de la automatización:

- Clasificación precisa de imágenes determinando la presencia o ausencia de animales.
- Para aquellas imágenes que previamente han sido identificadas con la presencia de un animal:
 - Identificación de la especie del animal capturado en la imagen.
 - Cuantificación de los animales presentes en dichas imágenes.

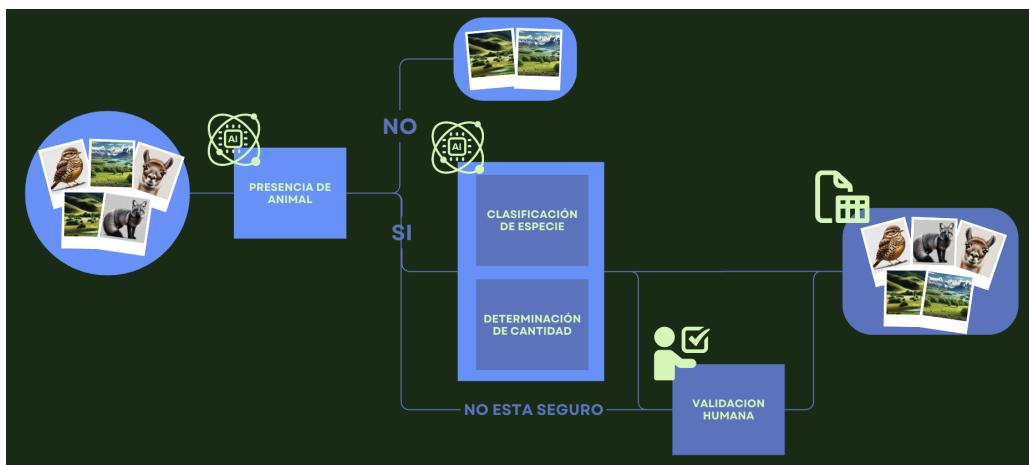


Figura 4: Diagrama de Flujo del Proceso

La **Figura 4** ilustra el proceso de clasificación de imágenes propuesto. Inicialmente, cada imagen es analizada por un avanzado modelo de Inteligencia Artificial, diseñado para detectar la presencia de fauna. En caso de identificar un animal, la imagen se procesa mediante modelos específicos que determinan la especie y cuentan el número de individuos presentes.

Para aquellas imágenes en las que el modelo no logre una confianza adecuada en su clasificación, se procederá a una revisión humana. Este enfoque reduce significativamente la carga laboral sin comprometer la precisión.

Este flujo de trabajo es facilitado al usuario a través de una interfaz intuitiva encargada de hacer la carga de los datos clasificados a la planilla y mover las imágenes sin animales a una carpeta de imágenes "vacías".

La fase final del proceso involucra la validación y comparación de resultados, donde se calculan Indicadores Clave de Desempeño (KPIs) y se comparan con la situación actual, con la posterior preparación detallada de la presentación para los clientes, asegurándose del cumplimiento de estándares concretos.

El éxito del proyecto se fundamenta en la colaboración sinérgica de expertos en el dominio de datos, cada uno aportando habilidades esenciales. Es necesario el perfil de Ingeniero de Datos para la gestión, preparación y limpieza de datos. El Analista de Datos realizará la interpretación de la información para extraer tendencias y patrones, y evaluará los KPIs comparando los resultados obtenidos. El desarrollo de algoritmos avanzados de aprendizaje automático para la clasificación precisa de imágenes es otorgado al Científico de Datos.

Outputs del Proyecto

El proyecto genera una serie de entregables claves diseñados para garantizar la aplicabilidad de la solución propuesta. Estos entregables incluyen:

- **Interfaz de Usuario:** Mediante una interfaz sencilla el empleado vincula la base de imágenes recolectadas luego de cada periodo de muestra. Esta incorpora un modelo de clasificación avanzado para procesar y categorizar las imágenes, proporcionando una hoja de cálculo con imágenes etiquetadas listas para el análisis estadístico y la estimación de la abundancia de especies.
- **Interfaz de Validación:** Esta herramienta está diseñada para que los usuarios realicen una revisión y validación manual en casos donde la confianza en el etiquetado automático no sea suficiente. La operativa y funcionalidades de las interfaces se describen con mayor detalle en el **Anexo 1**.
- **Paquete de Documentación:**
 - **Resumen Ejecutivo:** Una panorámica de los objetivos, metodologías y principales hallazgos del proyecto.
 - **Descripción de Metodologías:** Profunda exposición sobre las técnicas empleadas, modelos usados y KPIs evaluados.
 - **Instrucciones de Uso:** Guía detallada, coherente y acompañada de ilustraciones pertinentes para facilitar una adecuada comprensión y navegación de la interfaz visual.
- **Piloto 2023:** Se efectuará una demostración del sistema utilizando el conjunto de datos del año 2023 aún sin procesar, para confirmar la funcionalidad del sistema y detallar los resultados obtenidos.

La interfaz visual que permitirá los beneficios planteados en el proceso de validación es expresada en detalle en el **Anexo 1**. El proceso de trabajo propuesto concluirá con la creación de una base de datos integral, la cual consolidará y albergará de manera

sistemática los datos derivados de la clasificación. Esta base de datos no solo servirá como un repositorio de información, sino que también estará diseñada para facilitar el análisis y la extracción de insights valiosos.

Plan de Trabajo

Etapas del Proyecto

El desarrollo del proyecto se divide en cinco etapas fundamentales:

1. Obtención y Preparación de Datos:

La Fundación suministra un repositorio integral de imágenes obtenidas por cámaras trampa, que incluye datos valiosos como la identificación del dispositivo, coordenadas geográficas precisas, así como la fecha y hora exactas de cada captura. Adicionalmente, proporcionan un documento estructurado que correlaciona las imágenes con sus respectivas clasificación de especies y el número de individuos detectados.

Los datos recopilados pueden contener inconsistencias, valores atípicos o datos faltantes. Para abordar estos problemas, se llevará a cabo un proceso de limpieza de datos. Esto implica la identificación y corrección de errores, la eliminación de duplicados y la estandarización de formatos para garantizar la coherencia en los datos.

Se realizará un análisis descriptivo detallado de la base de datos para obtener una comprensión más profunda de su contenido. Esto incluirá estadísticas resumidas, visualizaciones y exploración de tendencias y patrones presentes en los datos.

Para preparar los datos para los modelos de detección de presencia de individuos se realizará un proceso de generación de imágenes etiquetadas sin animales. Esto se debe a que actualmente no hay imágenes etiquetadas sin animales, ya que no son del interés del cliente. Para realizar esto se utilizará el modelo *Megadetector* con una precisión del 95% para identificar las imágenes que no presenten fauna.

2. Entrenamiento de Algoritmos Diversos:

Se desempeñará un papel fundamental en la selección, entrenamiento y optimización de los algoritmos de Deep Learning para abordar las tres problemáticas identificadas: detección de presencias, clasificación de especies y conteo de individuos.

Previo a la aplicación de los modelos se preparan los datos para poder ser procesados por ellos. Se dividen los datos en entrenamiento y prueba usando técnicas de estratificación, que son efectivas en casos de desbalanceamiento de las clases. Para evaluar la capacidad de generalización de los modelos y prevenir el sobreajuste, se realizará una validación cruzada utilizando varios subconjuntos de datos de validación.

Se realizará un entrenamiento inicial con el conjunto de datos de entrenamiento. Durante esta fase, los modelos aprenderán a identificar patrones en las imágenes y

realizarán tareas de clasificación y detección. Para optimizar el rendimiento de los modelos, se estudiará la efectividad de utilizar modelos previamente entrenados, la utilización de metodologías de Transfer Learning con ImageNet.

Se realizará una búsqueda de hiperparámetros, esto implica ajustar valores como la tasa de aprendizaje, la cantidad de capas en la red neuronal, las funciones de activación y otros parámetros clave para obtener el mejor rendimiento posible. Esto garantiza que los modelos funcionen bien en datos nuevos y no solo en los datos de entrenamiento. A medida que los modelos se entrena y ajustan, se realizará una evaluación continua de su desempeño. Esto incluye la monitorización de métricas de rendimiento: el accuracy, el recall y la precision.

Con base en los resultados de la evaluación continua, se realizarán ajustes adicionales en los modelos para mejorar su precisión y eficiencia. Todos los detalles relacionados con la selección, entrenamiento y ajuste de los algoritmos, así como los resultados obtenidos, se documentarán minuciosamente. Esto permitirá una comprensión completa de los enfoques y resultados para su posterior comunicación a la fundación *Rewilding Argentina*.

3. Validación y Comparación de Resultados:

Una vez que los modelos de Inteligencia Artificial estén implementados y en funcionamiento, es esencial llevar a cabo una evaluación exhaustiva de su desempeño. Esta fase tiene como objetivo medir la eficacia y eficiencia de los modelos en comparación con la situación actual, que implica la clasificación manual de imágenes.

Se realizará una comparación detallada entre los resultados de los modelos y la clasificación manual utilizando los Indicadores Clave de Desempeño. Se visualizará la información para un claro y preciso entendimiento de los resultados y se analizarán posibles sesgos y limitaciones de los modelos.

Basado en los resultados de esta fase, se determinarán los modelos que maximicen los resultados y se seleccionarán para implementar la solución.

4. Desarrollo de la Interfaz al Usuario:

Se implementará una interfaz de usuario que operará de acuerdo con el flujo de trabajo ilustrado en la **Figura 4**. Dicha interfaz facilitará la integración y gestión de la base de imágenes, permitiendo la ejecución efectiva de las predicciones por parte de los modelos de inteligencia artificial subsecuentes.

Cuando la presencia de animales resulte negativa, se trasladarán las imágenes a una carpeta que reúne todas las imágenes con esta característica. Esto se hace con el objetivo de no interferir con las necesidades del cliente y poder reutilizarlas para futuro re-entrenamiento del modelo.

Si se identificara la presencia de un animal la imagen será procesada por los modelos de clasificación de especie y determinación de cantidad.

En instancias donde el modelo presente un grado de confianza subóptimo para determinar la clase de la imagen, dichas imágenes serán sometidas a una revisión

manual detallada. Para tal propósito, se diseñará una interfaz gráfica intuitiva y de alto estándar que facilite el proceso de etiquetado al usuario.

Esta interfaz integrará mecanismos que garantizan la uniformidad de la información, restringiendo opciones para evitar la introducción de inconsistencias, a diferencia de los métodos actuales. La interfaz desplegará las imágenes pertenecientes a una misma ráfaga y solicitará al usuario la clasificación de las mismas. Un contador indicará el número de imágenes pendientes, brindando una perspectiva clara del progreso. Además, se permitirá pausar y reanudar el proceso según la conveniencia del usuario, evitando así la fatiga y monotonía.

Al concluir el proceso, se garantiza que el usuario dispondrá de información íntegra y coherente en una base de datos meticulosamente estructurada.

5. Preparación de la Presentación para los Clientes:

La etapa final se centra en la detallada documentación del proyecto y una revisión meticulosa de sus componentes previos a la presentación. Esta documentación comprenderá manuales que cubrirán todos los aspectos de la interfaz, desde la subida inicial de imágenes hasta la interpretación de los resultados generados por la inteligencia artificial.

El objetivo es que la presentación cumpla con los más altos estándares profesionales y esté lista para su entrega a *Rewilding Argentina*.

La fecha de presentación ha sido establecida en concordancia con los plazos de la fundación para el análisis de datos del 2023. Se llevará a cabo una demostración del proyecto piloto utilizando los datos recientes para ilustrar de manera efectiva las capacidades y beneficios de la propuesta. Esto proporcionará a la fundación la oportunidad de realizar sus análisis y proyecciones basados en la información y resultados obtenidos.

Metodología de Trabajo Ágil

El desarrollo del proyecto adoptará la metodología Agile para garantizar la flexibilidad y adaptabilidad frente a los desafíos y requerimientos emergentes. Este enfoque iterativo y colaborativo posibilita ajustes rápidos a las fluctuaciones del proyecto y a las expectativas cambiantes. La estrecha colaboración con los equipos de Rewilding Argentina asegura que las soluciones propuestas permanezcan en sintonía con sus metas y requisitos dinámicos. Cada etapa concluida se materializará en un entregable que encapsula los avances logrados, una estrategia que se visualiza en el diagrama de Gantt (**Figura 5**) elaborado para la planificación del proyecto.

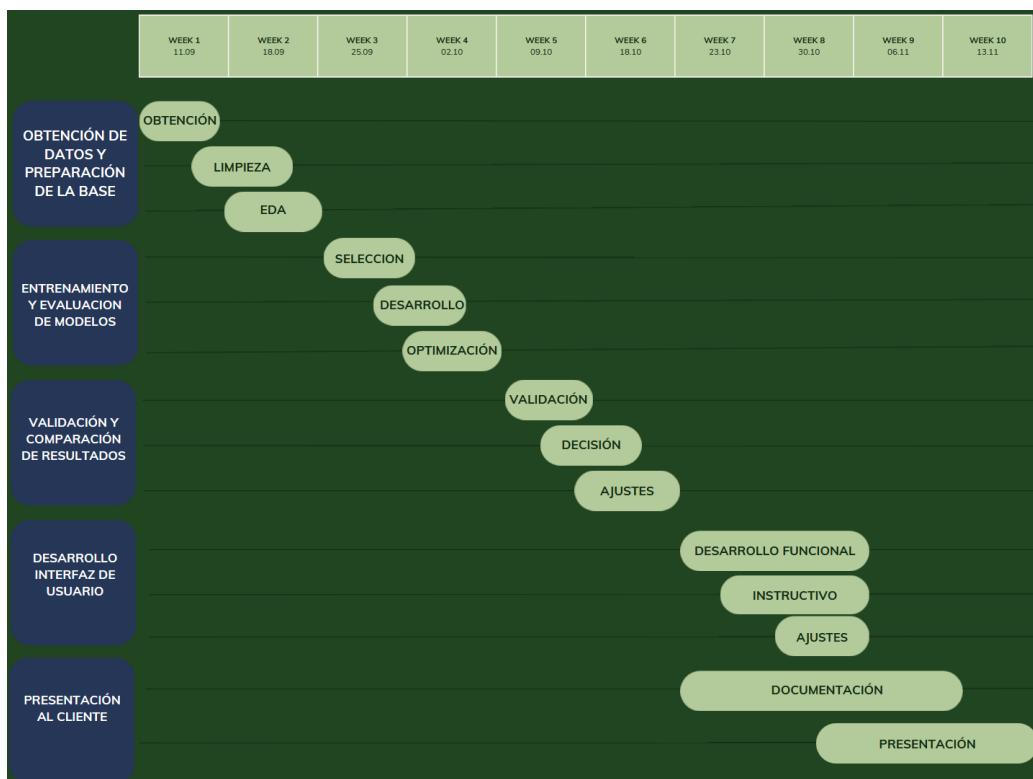


Figura 5: Diagrama de Gantt

Se realizará una reunión semanal sobre el estatus de las tareas planificadas, con el objetivo de generar iteraciones en las definiciones que se mantengan actualizadas y claras.

Análisis de Riesgos

Construimos las siguiente matriz de riesgos:

Riesgo	Probabilidad de ocurrencia	Gravedad del impacto	Plan de mitigación
Insuficiencia de datos por especie para la eficacia del modelo de clasificación	Alta	Media	<ul style="list-style-type: none"> Adoptar técnicas de Data Augmentation. Adoptar modelos pre-entrenados utilizando Transfer Learning. Explorar colaboraciones con otras zonas de conservación para fortalecer el conjunto de datos. Aplicar metodología de Milestone, fijando objetivos

			intermedios dentro de los resultados esperados.
Obstáculos en la presentación de resultados para el uso eficiente por parte del cliente.	Baja	Alta	<ul style="list-style-type: none"> • Desarrollar un protocolo de comunicación estructurado que permita la transmisión clara y detallada de la funcionalidad y los resultados de la herramienta. • Implementar sesiones de formación personalizadas y continuas para los usuarios, asegurando su familiarización y competencia en la operación de la interfaz. • Crear documentación exhaustiva y recursos de ayuda en línea accesibles, como tutoriales en video.
Pérdida de datos críticos debido a fallas en el almacenamiento o respaldo.	Baja	Alta	<ul style="list-style-type: none"> • Implementar sistemas de respaldo de datos robustos y regulares. • Realizar auditorías de seguridad de datos para evitar pérdidas.
El modelo de inteligencia artificial propuesto pueda cometer errores significativos al analizar una gran cantidad de imágenes, lo que podría resultar en datos inexactos y sesgados.	Media	Alta	<ul style="list-style-type: none"> • Dividir los datos en conjuntos de entrenamiento y prueba para evaluar la precisión del modelo antes de su implementación. • Implementar un sistema de monitoreo continuo para detectar y corregir errores a medida que surgen. • Asegurar la calidad de los datos de entrenamiento y prueba, y considerar la inclusión de imágenes de diversas condiciones y ubicaciones.

Tabla 1: Análisis de Riesgos

Datos a Utilizar

Se adquirió la base de imágenes de los últimos procesos de captura con cámaras trampa. Esta base se genera año a año para realizar una estimación del recuento de las especies, con el objetivo de evaluar el avance de los proyectos de restauración de los ecosistemas.

La base se encuentra segmentada en carpetas dada la zona, el año y la cámara. Se cuenta con información de los últimos cuatro años, desde el proceso realizado en 2020 hasta el realizado en 2023. La zona Ibérica cuenta con 8 cámaras mientras que la zona Sauce cuenta con 19.

Además, cada imagen presenta, dentro de la misma, como se observa en la **Figura 6**, la información de la fecha y hora en la que se tomó y las condiciones climáticas. Esta información puede ser accedida mediante la metadata del archivo.



Figura 6: Imagen de Muestra.

En total, la base posee 165083 imágenes. La mayor proporción de la información está en el sitio 'Sauce', este comprende el 90% de la base. Además, la mayor cantidad de información es del año 2023, esto se debe a que esta información hasta el día de hoy no ha sido procesada y no se eliminaron las tomas en las que no se presentan animales, como se hizo en años anteriores. Cómo las imágenes vacías no son de interés de la fundación estas son eliminadas cuando se realiza el proceso de etiquetado.

En conjunto a las imágenes, se dispone de una hoja de cálculo con la información de las imágenes que fueron etiquetadas manualmente para los años 2020, 2021 y 2022. En la misma se describe:

- **Sitio:** Refiere a la zona, es decir, puede ser 'Sauce' o 'Ibérica'.
- **Año:** Años entre el 2020 y el 2022. Las imágenes del 2023 aún no han sido procesadas.
- **Cámara:** Código único de la cámara, compuesto por la zona y el número de cámara, por ejemplo, 'IB4' para la cuarta cámara de la zona Ibérica.
- **Nombre:** Presenta el nombre del archivo.

- **Especie:** Determina la especie presente en la imagen.
- **Cantidad:** Determina la cantidad de individuos presentes en la imagen.
- **Fecha:** Se extrae la fecha de la imagen.
- **Hora:** Se extrae la hora de la imagen.

La cantidad de imágenes etiquetadas en total, en los tres años de muestreo, es de 2648. Por lo que la proporción de imágenes con animales etiquetados sobre el total de la base es muy baja (5.9%). Esto se debe, en parte, a que en las secuencias donde aparece uno o varios animales en un período breve de tiempo solo se etiqueta una imagen, para evitar etiquetar el mismo individuo múltiples veces.

Herramientas a Utilizar

Para el desarrollo del proyecto se utilizarán diversas herramientas, plataformas y técnicas que permitirán la construcción y aplicación de los modelos de clasificación previamente descritos.

La base de imágenes fue proporcionada mediante un disco rígido con un tamaño de aproximadamente 550 GB, la herramienta resultante permitirá que el cliente continúe con este método de almacenamiento para no incurrir en nuevos costos. Con efectos de tener mayor accesibilidad durante el trabajo las capturas y la base serán almacenadas en la nube de *Amazon Web Services* utilizando el servicio de buckets S3. Es importante aclarar que esto fue realizado para el análisis y el desarrollo del presente proyecto con el objetivo de almacenar una copia de las imágenes recibidas, el trabajo no contempla la migración de los datos de *Rewilding* a una plataforma de Cloud.

Tanto para el acceso a AWS como para el desarrollo del programa se optó por el lenguaje de programación *Python*, dada su amplia comunidad, las librerías disponibles especializadas para los fines de interés y los frameworks de Aprendizaje Profundo como *Pytorch* y *TensorFlow*.

En cuanto a las técnicas específicas para la construcción del modelo, se evaluarán aquellas referidas a:

- **Redes Neuronales Convolucionales (CNN):** Técnica más comúnmente utilizada para la clasificación de imágenes. Estas redes están diseñadas específicamente para procesar datos de imágenes y son capaces de aprender jerarquías de características a diferentes niveles de abstracción. Las CNN han demostrado un rendimiento excepcional en una amplia variedad de problemas de clasificación de imágenes.
- **Data Augmentation:** Implica crear versiones ligeramente modificadas de las imágenes de entrenamiento mediante rotaciones, reflejos, recortes y otros cambios. Es de utilidad para expandir la cantidad de imágenes sin animales que están etiquetadas, dado que la mayoría de esta índole fueron eliminadas.
- **Upsampling/Downsampling:** Dado el desbalance del dataset de entrenamiento entre las imágenes que efectivamente poseen animales y

aquellas que no, se deberá utilizar alguna de estas técnicas para el correcto entrenamiento.

- **Transfer Learning:** La transferencia de aprendizaje implica tomar una red neuronal pre-entrenada en un conjunto de datos grande y ajustarla para un problema específico. Para el tipo de problema en cuestión, se suele utilizar modelos entrenados en base a ImageNet, una vasta base de datos de imágenes que abarca múltiples categorías. ImageNet ha sido fundamental en los avances de la visión por computadora al proporcionar un conjunto de datos estándar para entrenar modelos profundos. Al aprovechar los pesos y características aprendidas en ImageNet, se puede acelerar y mejorar el entrenamiento para tareas específicas, reduciendo la necesidad de grandes cantidades de datos propios.
- **Arquitecturas Conocidas:** En el desarrollo del modelo se seleccionarán arquitecturas de redes neuronales avanzadas como VGG16, ResNet e InceptionV3, apreciadas por su eficiencia y precisión en la clasificación de imágenes. Estas arquitecturas se adaptarán cuidadosamente a las necesidades del proyecto, considerando la naturaleza y complejidad de los datos. La decisión sobre qué arquitectura utilizar se basará en pruebas de rendimiento y su compatibilidad con la infraestructura de procesamiento de datos existente.
- **Uso de Modelos Externos:** Además de las técnicas mencionadas, se contempla la integración de modelos externos pre-entrenados para alguna tarea específica, como el modelo *MegaDetector* de Microsoft, que han demostrado ser eficientes en la detección y clasificación de animales en imágenes de trampas-cámara. Estos modelos, al estar ya entrenados en grandes conjuntos de datos, pueden ofrecer una base sólida y ahorrar tiempo y recursos en el proceso de entrenamiento, si bien la robustez de los mismos sugiere que se puede lograr más precisión con un entrenamiento especializado a un bioma específico.

Análisis Exploratorio de Datos (EDA)

Disparadores de Exploración

Para iniciar el análisis descriptivo y exploratorio de los datos, se han planteado preguntas clave y enfoques que respaldaron el proyecto y serán fundamentales para su resolución. Estas áreas de interés son las siguientes:

Distribución de Especies: Se examinará la distribución de especies en el conjunto de imágenes capturadas. Permitiendo identificar las especies más y menos comunes en las imágenes.

Tendencias Temporales: Aprovechando la información temporal presente en las imágenes, se analizará la variación de la presencia de diferentes especies a lo largo del tiempo. Identificar patrones estacionales o tendencias de comportamiento.

Además, se identificarán, si existen, los patrones de actividad de ciertas especies en momentos específicos del día o de la temporada.

Calidad de Imágenes: Se analizará la calidad de las imágenes. Se mostrará la distribución de imágenes oscuras o imágenes con baja visibilidad. Esto ayudará a identificar dificultades presentadas al encargado de clasificación.

Consecución de Apariciones: Se realizará una investigación detallada sobre la frecuencia de aparición del mismo individuo en secuencias de imágenes consecutivas. El objetivo es explorar y comprender mejor el fenómeno de persistencia de un individuo frente a la cámara, analizando los factores que podrían explicar esta continuidad en la presencia.

Estos enfoques y análisis son fundamentales para comprender mejor los datos recopilados y extraer información relevante para el proyecto. Permitirán la comprensión de la fauna, su comportamiento y la calidad de las imágenes capturadas, lo que permitirá respaldar las decisiones y conclusiones del proyecto.

Limpieza y Preprocesamiento

La fase de limpieza y preprocesamiento de datos constituye un elemento crítico del proyecto, ya que sienta las bases para un análisis robusto y preciso. Este proceso meticuloso permite la extracción de información significativa y fiable a partir de los datos suministrados por la fundación Rewilding Argentina.

La base proporcionada no presenta valores nulos en ella. Pero no hace referencia al total de las imágenes obtenidas. Sólo fueron etiquetadas aquellas imágenes en las que existe presencia animal y, aún así, no todas ellas, ya que hay ocasiones en las que los animales se mantienen en el rango de captura por un tiempo prolongado y la sensibilidad de la cámara trampa hace que se capturen más de una vez. En estas ocasiones descritas, el equipo de Rewilding definió clasificar una sola imagen que represente al conjunto de las imágenes consecutivas. Puesto esto, se calcula que la **proporción de imágenes efectivamente etiquetadas, sobre el total de imágenes con animales, es de un 5.9%**. Cabe destacar que esta proporción sería aún menor si no se eliminaran las imágenes en las que no se presentan animales.

Sitio	Año	Cámara	Especie	Cantidad	Nombre	Fecha	Hora
Sauce	2022	S1	Guanaco	1	SAUCE 2022 .1 (7).JPG	19/02/2022	14:03:44
Sauce	2022	S1	Guanaco	3	SAUCE 2022 .1 (12).JPG	22/02/2022	13:47:24
Sauce	2022	S1	Guanaco	1	SAUCE 2022 .1 (15).JPG	22/02/2022	13:47:30
Sauce	2022	S1	Guanaco	1	SAUCE 2022 .1 (32).JPG	23/02/2022	14:01:46
Sauce	2022	S1	Guanaco	1	SAUCE 2022 .1 (41).JPG	23/02/2022	14:02:06
Sauce	2022	S1	Guanaco	1	SAUCE 2022 .1 (72).JPG	24/02/2022	7:04:06
Sauce	2022	S1	Guanaco	2	SAUCE 2022 .1 (82).JPG	25/02/2022	12:18:10
Sauce	2022	S1	Guanaco	11	SAUCE 2022 .1 (97).JPG	27/02/2022	13:46:56
Sauce	2022	S1	Guanaco	1	SAUCE 2022 .1 (132).JPG	28/02/2022	18:26:36
Sauce	2022	S1	Guanaco	1	SAUCE 2022 .1 (152).JPG	28/02/2022	19:27:58
Sauce	2022	S1	Guanaco	1	SAUCE 2022 .1 (160).JPG	28/02/2022	19:30:26
Sauce	2022	S1	Bandurrita	1	SAUCE 2022 .1 (212).JPG	1/03/2022	1:44:34
Sauce	2022	S1	Liebre	1	SAUCE 2022 .1 (217).JPG	1/03/2022	6:26:10
Sauce	2022	S1	Guanaco	1	SAUCE 2022 .1 (222).JPG	1/03/2022	7:52:52
Sauce	2022	S1	Guanaco	1	SAUCE 2022 .1 (226).JPG	1/03/2022	7:53:02
Sauce	2022	S1	Guanaco	1	SAUCE 2022 .1 (292).JPG	3/03/2022	9:18:44
Sauce	2022	S1	Zorro Gris	1	SAUCE 2022 .1 (375).JPG	4/03/2022	9:08:24
Sauce	2022	S1	Guanaco	1	SAUCE 2022 .1 (387).JPG	5/03/2022	8:44:32
Ibérica	2021	IB7	Aguilucho Común	1	1-IB7- (231).JPG	2/05/2021	12:21:34
Ibérica	2021	IB7	Aguilucho Común	1	1-IB7- (236).JPG	6/05/2021	11:30:10
Ibérica	2021	IB7	Cormoran Imperial	12	1-IB7- (246).JPG	8/05/2021	11:53:04
Ibérica	2021	IB7	Carancho	1	1-IB7- (251).JPG	13/05/2021	8:53:46
Ibérica	2021	IB7	Carancho	1	1-IB7- (257).JPG	13/05/2021	8:55:22
Ibérica	2021	IB7	Tucuquere	1	1-IB7- (261).JPG	15/05/2021	22:49:58
Ibérica	2021	IB7	Guanaco	2	1-IB7- (266).JPG	16/05/2021	23:57:40
Ibérica	2021	IB8	Zorro Gris	1	0-IB8- (66).JPG	14/04/2021	1:18:26
Ibérica	2021	IB8	Carancho	1	0-IB8- (108).JPG	19/04/2021	12:45:26

Figura 7: Inconsistencias en la base de datos original

La base de datos, creada manualmente por el personal de *Rewilding*, presenta inconsistencias en las variables de texto. Esto se puede observar en la **Figura 7**. Para abordar este problema, se ha realizado un preprocesamiento de la base de datos. Este proceso incluyó la eliminación de espacios antes y después de los textos, la conversión de todos los textos a minúsculas, la eliminación de espacios dobles y la eliminación de tildes. Como resultado de este preprocesamiento se logró estandarizar las columnas categóricas al eliminar duplicados que, debido a variaciones simples en los textos, se consideran como valores distintos. Este paso es fundamental para **garantizar la coherencia y la integridad de los datos**, lo que facilita el análisis subsiguiente y asegura la precisión en los resultados.

La base de datos carece de una estructura que precise la ruta exacta a la imagen, esta información se recopila a partir de la combinación de otras columnas. Este proceso resulta menos intuitivo de lo que podría parecer en primera instancia, debido a que en la anotación manual no se siguieron los estándares de precisión necesarios. En este contexto, al referenciar las rutas construidas en base a las columnas proporcionadas, solamente el 14.7% de las imágenes coincidieron con la ruta generada, lo que significa que una proporción significativa de imágenes no concuerda con la información suministrada. Las razones principales detrás de esta discrepancia incluyen:

- Diferentes convenciones de nomenclatura de archivos.
- Inclusión de carpetas adicionales para el almacenamiento de archivos.
- Variaciones en la numeración de los nombres de los archivos.
- Variaciones en las extracciones de datos temporales (Fecha y Hora).
- Situaciones en las que se presenta un archivo, pero solo se almacena la imagen siguiente.

Después de un proceso exhaustivo, se logró reducir esta inconsistencia al **referenciar correctamente el 99.4% de los datos**. Para validar la integridad del cruce de información, se extrajo la fecha y hora de captura de la foto directamente desde la metadata del archivo. La estandarización y corrección de las rutas de los archivos en la base de datos genera un valor esencial para el proyecto al garantizar la consistencia y precisión de los datos. Esta **mejora en la integridad de los datos**, conseguida a través de la estandarización y corrección de las rutas de los archivos en la base de datos, es de gran aporte de valor al proyecto. Asegura la consistencia y precisión en el manejo de la información, lo que resulta fundamental para el análisis eficaz y la recuperación precisa de datos. Además, minimiza los errores potenciales y optimiza significativamente el procesamiento de datos. La base final con rutas funcionales tiene un tamaño de 2628 imágenes.

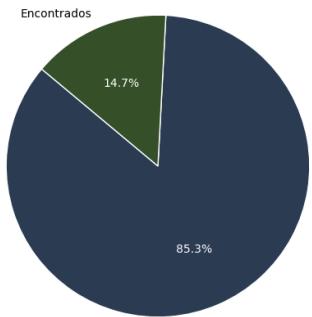


Figura 8: Proporción de Archivos que Coincidan pre-procesamiento.

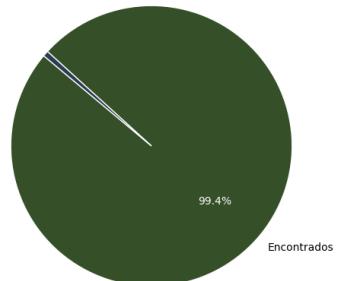


Figura 9: Proporción de Archivos que Coincidan post-procesamiento.

Aunque estos registros se consideran no aptos para el análisis predictivo, debido a la falta de correlación con las imágenes, serán de gran valor para el análisis descriptivo de la base de datos.

Análisis Descriptivo

La base de datos a utilizar para el aprendizaje supervisado tiene 2645 filas y 8 columnas. Cada uno de los registros representa una de las imágenes en las que se encontraron animales. Las variables descriptivas son categóricas, numéricas y temporales. Nuevamente, las variables presentadas son sitio, cámara, nombre, especie, cantidad, fecha y hora.

El **sitio** se ve representado por una variable categórica que indica la zona geográfica de la cámara que capturó la imagen, con dos valores posibles: 'Sauce' e 'Ibérica'. Una notable 86.2% de las imágenes provienen de la zona de Sauce. Esta predominancia se justifica por el hecho de que el área de Sauce tiene una mayor cantidad de cámaras, representando aproximadamente el 70% del total de dispositivos desplegados. Además, se ejecutó un proyecto adicional en Sauce en el año 2020, mientras que las cámaras se implementaron en Ibérica en 2021. Es relevante señalar que la duración de los proyectos de captura de imágenes es uniforme para ambos sitios. Las zonas y cámaras fueron representadas en la **Figura 10**.



Figura 10: Mapa Representativo de los Sitios.

Hasta la fecha, se tienen registros desde el **año** 2020 hasta el 2022. Donde en Sauce se realizan tres campañas de captura anuales y en Ibérica dos. Es importante señalar que las imágenes correspondientes al año 2023 aún no han sido sometidas a procesamiento o análisis.

Las **cámaras** utilizadas en todos los proyectos son consistentes y se colocan en ubicaciones predefinidas cada año para facilitar la comparación de resultados anuales. En el área de Sauce, se disponen de 19 cámaras, mientras que en la región Ibérica, se utilizan 8 cámaras. La selección de ubicaciones para estas cámaras se basa en un enfoque lógico y matemático, con distancias adecuadas entre ellas para garantizar que los mismos individuos no aparezcan en múltiples cámaras en el mismo día. Como se puede apreciar en la **Figura 11**, la distribución de las apariciones de la fauna por cámara no muestra una uniformidad constante a lo largo de los años. Aunque existen áreas con una mayor densidad de animales en comparación con otras, no se observa una estabilidad en las cantidades registradas por la misma cámara. Esto indica que la fauna no permanece en una ubicación fija a lo largo de los años, sino que está en constante movimiento. En consecuencia, no se identifica la necesidad de asignar un enfoque especial en el análisis a cámaras específicas, ya que la distribución de las apariciones varía con el tiempo y no se evidencia una cámara que requiera atención prioritaria.

Sin embargo, si consideramos la escasa actividad detectada por cámaras específicas, como S7, S8, S9, S10, IB3 y IB8, como se observa en la **Figura 11**, podría ser estratégicamente ventajoso para la iniciativa de Rewilding considerar la desinstalación o la reubicación de estas unidades para optimizar la eficacia del monitoreo.

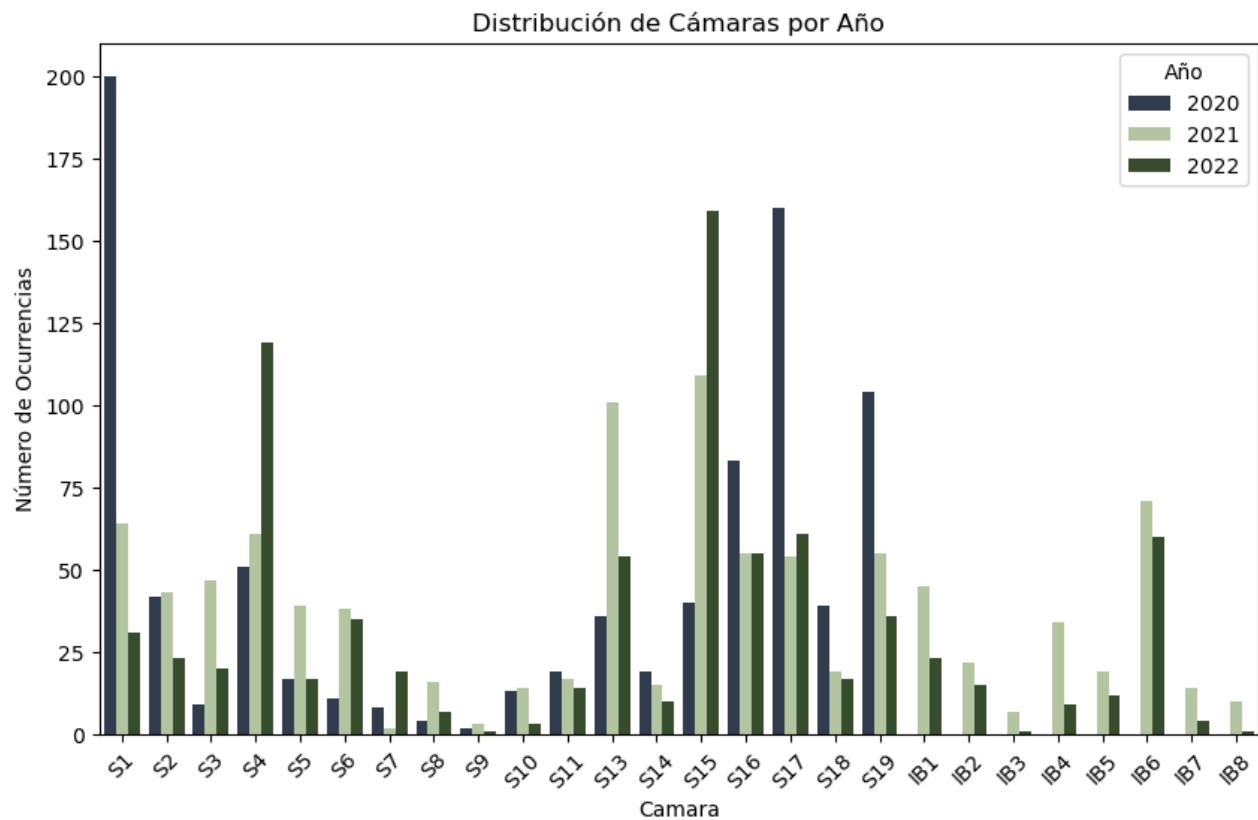


Figura 11: Distribución de apariciones por Cámara y Año.

Los **nombres** de los archivos se han diseñado con el propósito de permitir su identificación. No obstante, en ocasiones, estos nombres ofrecen información adicional sobre la cámara, el año o el lugar, pero no proporcionan datos completos que permitan ser completamente distintivos. En otros casos, se emplea una denominación genérica, como 'IMG_0001.JPG'. En resumen, el nombre de los archivos, por si solo, carece de la capacidad descriptiva necesaria para una identificación precisa de la imagen, siendo necesario contar con información adicional y relevante para llevar a cabo dicho proceso de identificación de manera efectiva.

Se han identificado un total de 37 **especies** clasificadas en el estudio. En su mayoría, las especies identificadas son comunes en ambos sitios.

La variable **cantidad** refleja el número de animales presentes en cada imagen. Este valor oscila en un rango de 1 a 24, con una media de 1.57. Estos datos indican que, en general, la mayoría de las imágenes contienen un número reducido de animales, aunque existen casos excepcionales en los que la cantidad es más elevada. Concretamente, el 72.1% de las imágenes únicamente muestra la presencia de un animal.

No solo, en la mayor cantidad de casos los animales se presentan solos, sino que también son pocas las especies que se presentan en manada. Las especies que se presentan en manada son los guanacos y las aves. Esta tendencia puede apreciarse detalladamente en la **Figura 12**.

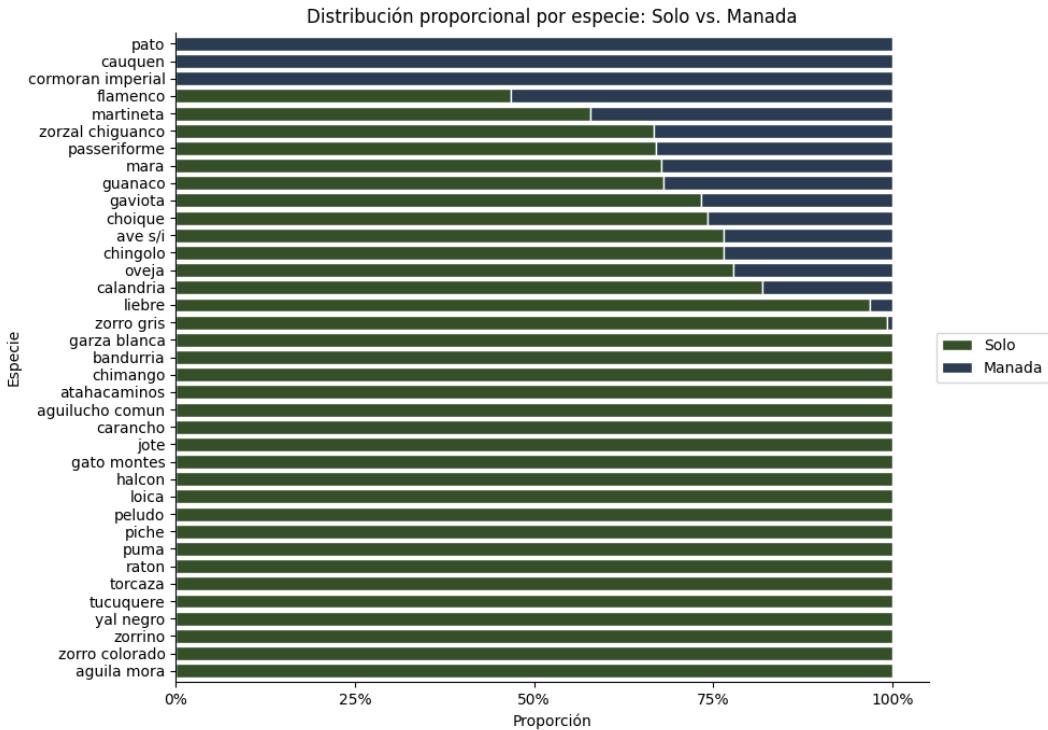


Figura 12: Distribución Proporcional por Especie de los Animales que se presentan Solos y en Manada.

Es importante señalar que, en el contexto de la estimación de la cantidad, no consideraremos como valores atípicos los casos extremadamente altos. En lugar de eliminar los outliers (**Figura 13**), se realizan análisis que sean robustos a la presencia de valores atípicos, como el uso de estadísticas de tendencia central robustas o pruebas estadísticas no paramétricas. Esta decisión se fundamenta en la necesidad de lograr la máxima precisión en nuestras mediciones, evitando descartar información relevante para el análisis.

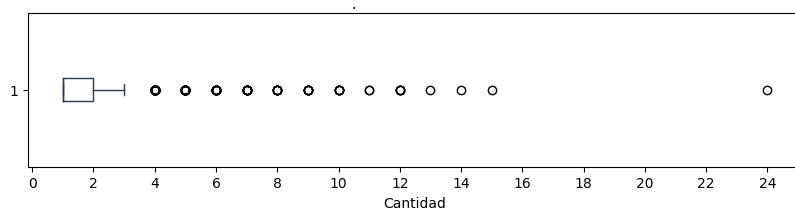


Figura 13: Boxplot de Cantidad de Apariciones de Animales

La **fecha** y la **hora** se derivan de los datos incrustados en las imágenes capturadas por las cámaras. Sin embargo, se ha observado que en ciertos casos estas marcas temporales presentan configuraciones incorrectas, lo que resulta en información inexacta. Se han identificado instancias en las que las fechas reflejadas en las imágenes son del año 1998, lo cual no coincide con los años en que se llevaron a cabo los proyectos de registro. Esta discrepancia temporal plantea desafíos significativos en la correcta interpretación de los datos recopilados. La confiabilidad

de la información temporal es afectada, por lo que se tendrá en cuenta la variable año para análisis futuros.

Se hallaron presentes en la base imágenes rotas (**Figura 14**), es decir, imágenes con píxeles quemados que no mostraban la totalidad de la información. En la exploración aleatoria se encontraron 16 de ellas. Se analizó si las mismas tenían relación con una cámara en específico pero no hubo relación, cómo tampoco lo hubo con el año de captura.



Figura 14: Ejemplo Imagen Rota.

Análisis Exploratorio

Tras completar un riguroso análisis descriptivo, el proyecto avanza hacia una fase de análisis exploratorio más profundo. Se abordarán aspectos específicos que evalúen los disparadores de exploración propuestos. Este análisis permitirá no solo profundizar en la comprensión de los patrones observados, sino también en la identificación de correlaciones y comportamientos significativos en los datos. Este paso es esencial para obtener insights detallados y para formular estrategias efectivas en respuesta a los hallazgos del estudio.

Distribución de Especies:

Luego de la normalización de los valores de las variables, se identificaron 37 especies. De todas formas, estas no siguen una distribución uniforme. Los guanacos son, con diferencia, el animal más presente representando el 68.84%. Los más interesantes para el análisis son aquellos con menor distribución, que se deberán incrementar con el tiempo aplicando los proyectos de restauración de *Rewilding*.

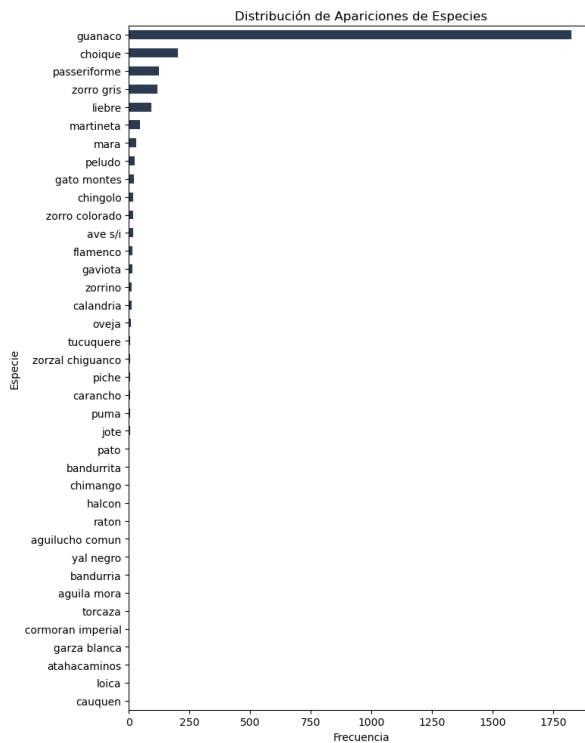


Figura 15: Distribución de Apariciones de Especies

Dado el enfoque en el análisis predictivo que se llevará a cabo en fases posteriores, es importante considerar que la clasificación por especie podría verse dificultada debido a la disponibilidad limitada de datos de entrenamiento para cada especie individual. Esta limitación es particularmente relevante en el caso de las aves, ya que en las imágenes, su presencia puede estar acompañada de otros animales, como se expone en la **Figura 17**. Por lo tanto, se propone una estrategia de clasificación que agrupe a los animales en categorías distintas, a saber: aves, mamíferos herbívoros y mamíferos carnívoros. En la **Figura 16** se observa la distribución de cada una de las categorías, que aunque sigan desbalanceada por la presencia de los guanacos, permiten un análisis más preciso.

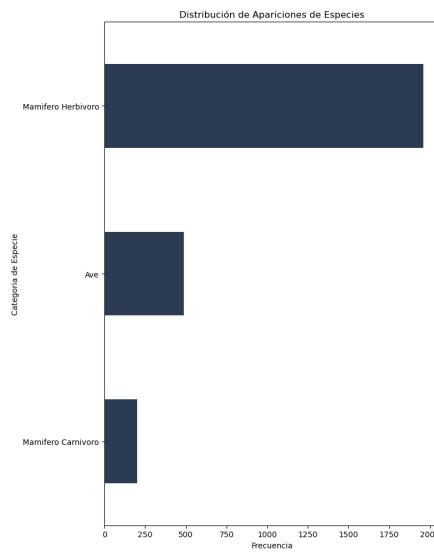


Figura 16: Distribución de Apariciones por Categorías de Especies



Figura 17: Ejemplo de Imagen donde se presenta un Ave y Guanacos

La **Figura 17** ilustra un ejemplo concreto de los desafíos asociados al etiquetado manual de datos. En este caso particular, una misma imagen se describió en tres registros diferentes. En uno de ellos, se etiquetó como "passeriforme" (un tipo de ave) con una cantidad de 1. En los otros dos registros, se identificó como "guanacos", uno con una cantidad de 11 y el otro con una cantidad de 2. Esta situación plantea un potencial problema cuando se utiliza esta información para entrenar modelos de inteligencia artificial, ya que puede **llover a interpretaciones inconsistentes de las cantidades reales**, en este caso, 13 guanacos.

Tendencias Temporales:

A diferencia de las expectativas de *Rewilding*, se ha registrado una disminución constante en la diversidad de especies observadas en la región de Sauce e Ibérica desde el año 2020. Esta tendencia a la baja podría sugerir que en el año 2023 se espera una reducción adicional en el número de animales avistados. Sin embargo, es importante destacar que la información actual no proporciona suficiente claridad para concluir de manera definitiva que esta disminución se mantendrá en el futuro.

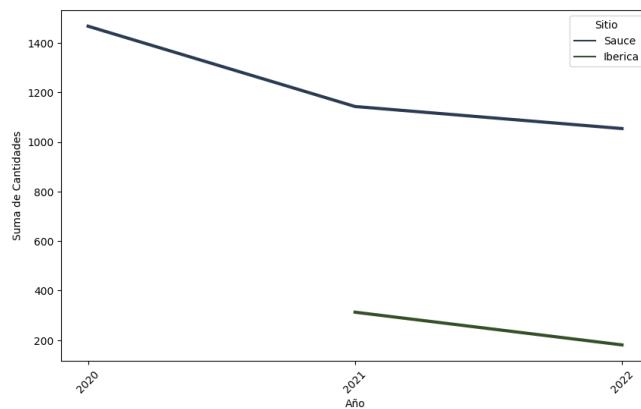


Figura 18: Tendencia de Cantidades Totales por Año y Sitio

En el año 2021, se observó un fenómeno particular en el cual la población de especies depredadoras (mamíferos carnívoros) experimentó un aumento, mientras que la población de mamíferos herbívoros disminuyó proporcionalmente. Este patrón revela un proceso de caza natural entre las especies animales, como se observa en la **Figura 19** y **Figura 20**, un aspecto de especial interés para la Fundación y que requiere una acción estratégica. Esto será utilizado para ver si los resultados de los modelos cumplen con las tendencias históricas.

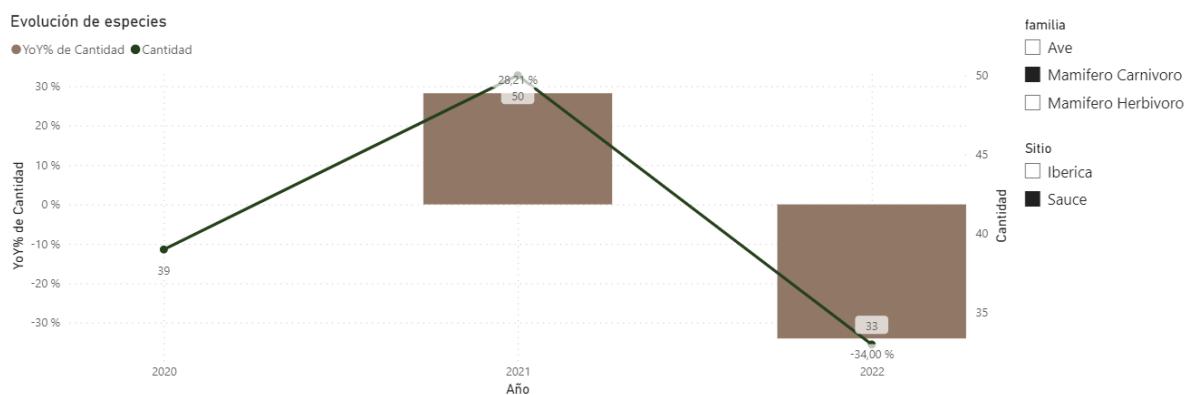


Figura 19: Evolución de Mamíferos Carnívoros en Sauce 2021.

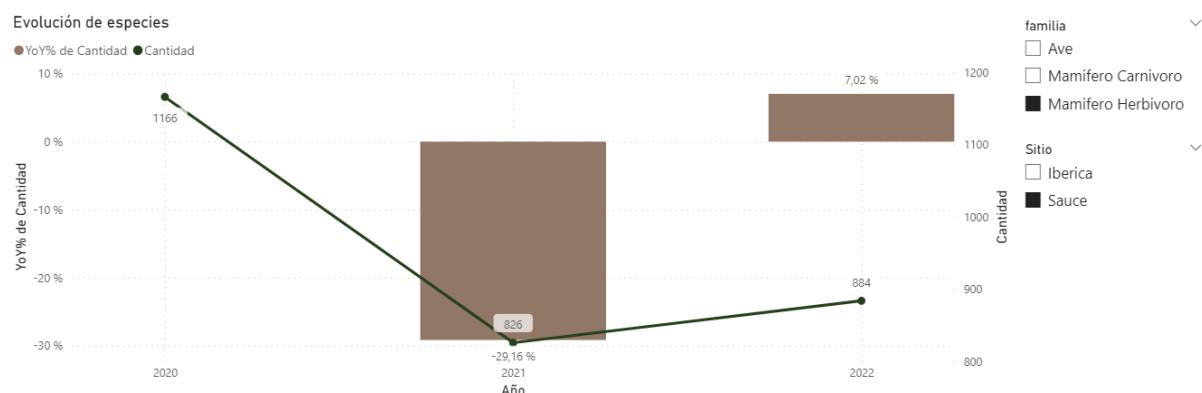


Figura 20: Evolución de Mamíferos Herbívoros en Sauce 2021.

Calidad de Imágenes:

En base a la información etiquetada se hizo un análisis de distribución de calidad de las imágenes que demuestre la visibilidad presentada. Se toma la luminosidad promedio presente en cada una de las fotos y se determina si este promedio está sobre o bajo el umbral de oscuridad. Para determinar si la imagen está oscura, la luminosidad promedio debe ser menor a 100. El 8.1% de las imágenes que fueron etiquetadas ocurrieron de noche.

Para descartar que la proporción de imágenes nocturnas etiquetadas sean por la falta de visión del etiquetador se analizó si la distribución de luminosidad presente

en las imágenes etiquetadas cumple la misma distribución del estado natural de la base. Se realizó la prueba de Kolmogorov-Smirnov para determinar:

H_0 : Las muestras (M_a y M_b) siguen la misma distribución.

H_1 : Las muestras (M_a y M_b) no siguen la misma distribución.

Siendo las muestras:

M_a : Muestra de imágenes etiquetadas del 2020 a 2022. $n = 2648$

M_b : Muestra de imágenes sin procesar del 2023. $n = 10000$

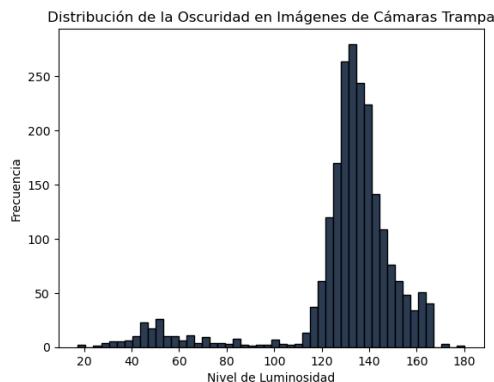


Figura 21: Distribución de Luminosidad de Imágenes Etiquetadas.

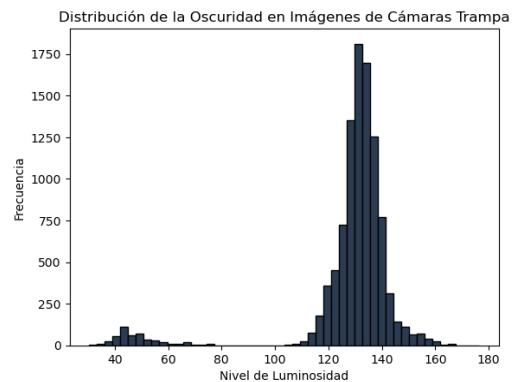


Figura 22: Distribución de Luminosidad de Imágenes sin Procesar.

En el caso de M_b se elige una muestra de 10000 unidades porque cumple con la precisión ($HW = 3$ puntos de luminosidad) buscada. Los resultados del análisis determinan que no se puede rechazar la H_0 con un nivel de confianza $\alpha = 0.05$, por lo que no se puede demostrar que las muestras no sigan la misma distribución. Estos hallazgos contribuyen a una mejor comprensión del impacto de las imágenes de baja visibilidad en el proceso de clasificación. Al identificar que no existen diferencias significativas entre la distribución de luminosidad en las imágenes naturales y las procesadas, se llega a la conclusión de que los movimientos de los animales durante la noche son correctamente clasificados.

Además, se puede inferir que las cámaras no se activan con tanta frecuencia durante la noche, lo cual es coherente con la expectativa de un menor movimiento animal en ese período.

	Madrugada	Mañana	Mediodía	Tarde-noche
Ave	16,82%	34,55%	22,88%	25,76%
Mamífero carnívoro	34,43%	24,59%	17,21%	23,77%
Mamífero herbívoro	10,67%	34,08%	37,52%	17,73%
	12,58%	33,84%	34,20%	19,38%

Tabla 2: Porcentaje de Apariciones por Categoría y Momento del Día

En relación con los patrones de actividad diaria de las especies, se dividió el día en cuatro períodos de seis horas cada uno: Madrugada (de 00:01 a 06:00), Mañana (de 06:01 a 12:00), Mediodía (de 12:01 a 18:00) y Tarde-noche (de 18:01 a 00:00). Se observa que los depredadores suelen registrarse principalmente en el rango horario de 18:00 a 06:00, lo que concuerda con sus comportamientos habituales. En contraste, los mamíferos herbívoros muestran una mayor actividad durante las horas de luz. Este patrón se repite en ambas regiones analizadas y podría estar influenciado por las condiciones de iluminación al momento de la captura de las imágenes.

Exploración de Información no Procesada

El procedimiento de clasificación ejecutado por *Rewilding* descarta las imágenes sin presencia de vida animal. Consecuentemente, para examinar la esencia de los datos originales, es imperativo analizar el conjunto de imágenes del año 2023 que aún no ha sido objeto de procesamiento.

Para llevar a cabo esta tarea, se realizaron pruebas utilizando el modelo *Megadetector*, ampliamente reconocido y utilizado en organizaciones líderes en la conservación de especies. Para llevar a cabo este análisis, se ha tomado como muestra las imágenes capturadas por las cámaras S4 (un total de 6,690 imágenes), IB3 (966 imágenes), IB4 (165 imágenes) e IB7 (180 imágenes) en el año 2023.

El modelo demostró su capacidad para clasificar imágenes de animales no solo en situaciones relativamente sencillas de identificar para un observador humano, sino también en imágenes que presentaban desafíos, como la presencia de animales poco visibles, distantes, superpuestos, desenfocados debido a movimientos rápidos, con obstrucciones parciales y en condiciones nocturnas.

Además, el modelo exhibió una notable habilidad para evitar confusiones entre vehículos o seres humanos y animales. En el caso de los humanos, logró identificarlos con alta confianza incluso cuando obstruyen parcialmente la vista de la cámara, posiblemente durante la configuración de la misma.

A continuación (**Figura 23** a **Figura 30**), se ejemplifican estas observaciones con outputs del modelo.



Figura 23: Ejemplo Animal Visible



Figura 24: Ejemplo Animal Alejado.



Figura 25: Ejemplo Animal Camuflado.



Figura 26: Ejemplo Animal en Condiciones Nocturnas.

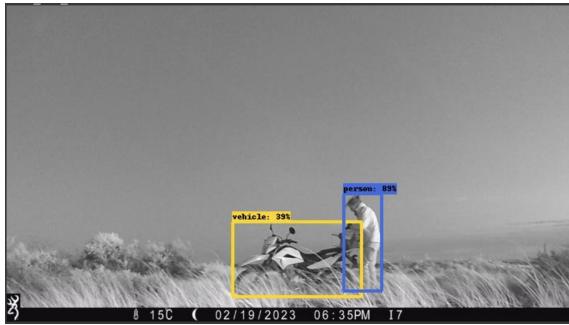


Figura 27: Ejemplo Hombre y Vehículo.



Figura 28: Ejemplo Animal Alejado en Condiciones Nocturnas



Figura 29: Ejemplo Animal con Obstrucción Visual.



Figura 30: Ejemplo Animal en Pasaje Veloz.

Aporte Pre-Procesamiento:

Luego de observar los resultados del modelo con imágenes de la base de datos de manera no supervisada se ha tomado la decisión de emplearlo para identificar posibles falsos positivos en el conjunto de datos de imágenes clasificadas por *Rewilding* de manera manual. Este paso es crucial para establecer una base confiable sobre la cual se llevará a cabo el proceso de entrenamiento posterior del modelo.

Para llevar a cabo esta tarea, se ha aplicado un filtro a las imágenes que forman parte del conjunto de datos clasificado. Se descartaron aquellas detecciones cuya confianza sea inferior al 20%. De las 2,485 imágenes analizadas, en el 93.5% de ellas se confirma la presencia animal. Pero esto determina que un 6.5% es susceptible a errores en la clasificación. Por lo que se deben revisar estas 161 imágenes.

Estas capturas fueron sometidas a una revisión manual, durante la cual se identificaron 15 casos de errores humanos. Este análisis permitió determinar **la tasa de Falsos Positivos asociada al proyecto, establecida en un 0.6%**. En virtud de garantizar un proceso de entrenamiento eficiente y confiable, se ha tomado la decisión de descartar estas imágenes erróneas. Como resultado, se dispone de una base de entrenamiento conformada por 2,470 imágenes.

Consecución de Apariciones:

Puesto todo lo evaluado previamente, se ha llegado a la conclusión de que, cuando se detecta la presencia de un animal en una imagen, existe una probabilidad significativa de que se generen varias imágenes consecutivas con el mismo animal presentes en ellas. Este fenómeno es de gran relevancia en el contexto de la automatización, ya que puede dar lugar a la contabilización múltiple de un mismo individuo.

A partir del análisis realizado con *Megadetector*, se ha registrado cuántas imágenes consecutivas muestran la presencia de animales, en función del número de imágenes en la secuencia (**Figura 31**).

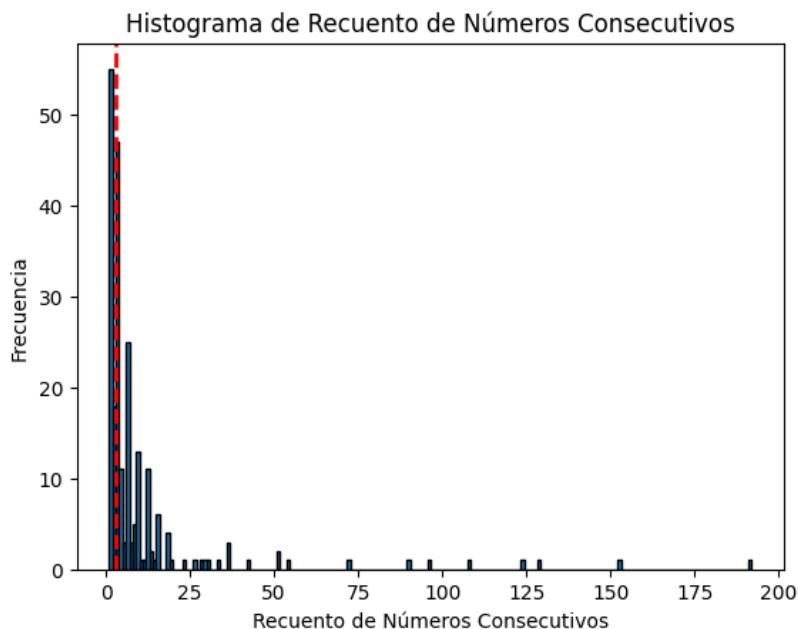


Figura 31: Recuento de Cantidad de Imágenes Consecutivas con Detecciones de Animales.

Este análisis revela que la mediana de imágenes consecutivas con presencia de individuos detectada es de 3 (señalada en rojo en la **Figura 31**), con una media de 10.66. Estos hallazgos pueden estar relacionados con la velocidad a la que se desplazan las diferentes especies de animales. Esta información, que puede

obtenerse directamente de las imágenes, proporcionará una **visión más completa de los patrones de movimiento de las especies y su comportamiento**.

Las conclusiones preliminares o supuestos previstos son que los animales pueden permanecer en el radar de la cámara por suficiente tiempo haciendo que esta se active ante pequeños movimientos. Es importante considerar que la duración de la presencia de un animal frente a la cámara puede variar considerablemente según la especie. Por ejemplo, la permanencia de un ave frente a la cámara tiende a ser mucho más breve en comparación con la de un mamífero. Estas diferencias en los patrones de comportamiento animal deben ser tomadas en cuenta en el análisis y la interpretación de los datos capturados.

Imágenes sin Animales:

Con el propósito de evaluar la viabilidad de etiquetar imágenes que no contienen animales para su uso futuro en el entrenamiento del modelo, se han empleado los resultados de los análisis de *Megadetector*. En este proceso, se han seleccionado exclusivamente aquellas imágenes que presentan una confianza nula en presentar animales (es decir, que esté 100% seguro de que no los haya), con el fin de minimizar la probabilidad de incluir imágenes que realmente contienen animales. En este enfoque, la prioridad no ha sido lograr una clasificación binaria precisa, sino más bien obtener un etiquetado correcto y confiable para las imágenes que carecen de animales.

Cámara	Imágenes sin Presencia de Animales	Proporción sobre el Total de Imágenes
IB3	445	46.07%
IB4	43	26.06%
IB7	47	26.11%
S4	2206	32.97%

Tabla 3: Imágenes con Alta Confianza de no Poseer Animales

Conclusiones del Análisis Exploratorio

Las conclusiones derivadas de la exploración del proyecto responden a los disparadores esenciales planteados en las fases iniciales del análisis.

El estudio sobre la **distribución de especies** reveló una distribución de datos desequilibrada, con una notable preponderancia de guanacos. Esta disparidad representa un desafío que se abordará mediante estrategias de balanceo de datos. Se efectuarán divisiones estratificadas del conjunto de datos para garantizar una representación equitativa de cada especie. Adicionalmente, se implementarán métodos de Upsampling y Data Augmentation para incrementar la representación de especies con menor frecuencia en la muestra, optimizando así la efectividad del entrenamiento de los modelos.

Ante la previsión de que el volumen de imágenes disponibles para el entrenamiento multiclase no alcanzará niveles óptimos, se adoptará una estrategia basada en hitos intermedios (*Milestones*). Esta aproximación permitirá establecer objetivos progresivos y realistas, focalizándose en la mejora incremental de la precisión, aun cuando ello implique una reducción de la complejidad en la predicción.

Se identificó un criterio poco amigable de etiquetado de imágenes donde se existe más de una especie. La presencia de múltiples especies en una única imagen puede inducir a confusión en los modelos de inteligencia artificial, comprometiendo la precisión de la clasificación de especies. Para mitigar este impacto, se excluirán del conjunto de entrenamiento aquellos casos que contengan imágenes con especies múltiples, con el fin de preservar la integridad y la eficacia del aprendizaje automático.

Los análisis de **tendencias temporales** sugieren una inclinación hacia una disminución en el número de individuos. Asimismo, se observa que un incremento en la población de mamíferos suele coincidir con una disminución en la cantidad de herbívoros. Estos patrones serán empleados como referencia para corroborar la coherencia de los resultados obtenidos en la prueba piloto con los datos históricos de campañas previas.

En relación con la **calidad de las imágenes**, se ha establecido que, si bien la presencia de animales durante la noche es reducida, la luminosidad de las imágenes nocturnas no constituye un obstáculo significativo para la clasificación.

La revisión detallada de las **secuencias de captura** reveló que, en promedio, cada avistamiento individual se documenta en una media de diez fotografías. La fundación ha optado por la clasificación integral de todas las imágenes para minimizar posibles errores asociados con la eliminación de secuencias por métodos estadísticos.

Para optimizar la eficiencia del proceso de validación sin comprometer la integridad de los datos, se incorporará en la interfaz una funcionalidad que permita agrupar imágenes sucesivas dentro de un margen temporal personalizable por el usuario. Este enfoque racionaliza la tarea de clasificación, liberando al personal de la necesidad de evaluar cada imagen individualmente y facilitando un enfoque más dinámico y ágil en la validación de las secuencias de captura.

El análisis meticoloso del proceso reveló la existencia de divergencias metodológicas, atribuibles a las variaciones temporales en la recolección de datos y a las diferencias en los enfoques de procesamiento por parte de distintos encargados de la tarea. Tales divergencias se manifiestan en asignaciones erróneas de etiquetas y recuentos en las imágenes, induciendo a una distorsión de los datos. Además, se han identificado errores de clasificación que comprometen la precisión y la trazabilidad del proceso analítico. A partir de este estudio se identifica el benchmark correspondiente al proceso de la **tasa de falsos positivos como el 0.6%**.

Estas constataciones subrayan la necesidad de estandarizar las prácticas de muestreo y análisis para asegurar la integridad y la fiabilidad de los datos recopilados. Es ambicioso pensar que todo el proceso puede ser automatizado, pero la cooperación humano-máquina puede lograr mejores resultados en menor tiempo.

La interfaz propuesta está diseñada para restringir las elecciones del usuario, mitigando así el riesgo de introducir discrepancias en el conjunto de datos.

Business Case

El propósito del business case es estudiar los posibles outcomes de los escenarios posibles, automatización total o parcial del proceso de clasificación de imágenes para identificar individuos en capturas de cámaras trampa.

Hipótesis de Trabajo

Hipótesis Principal: El uso de algoritmos de aprendizaje automático en la detección y clasificación de especies animales a partir de imágenes de cámaras trampa, **mejorará significativamente la precisión** en la categorización de datos de biodiversidad en áreas de conservación de Rewilding Argentina.

Hipótesis Secundaria: La implementación de una interfaz visual con integración de inteligencia artificial permitirá **reducir de manera sustancial los tiempos** de procesamiento de datos, tanto totales como de trabajo humano, mediante la minimización de la dependencia de la revisión manual de decenas de miles de imágenes.

Análisis de Situación Actual

Con el propósito de determinar las oportunidades de mejora del proyecto, se han identificado variables clave que reflejan la situación operativa actual, en referencia a la precisión y tiempo de etiquetado.

Como se explicitó previamente, el proceso actual es realizado manualmente por empleados de la organización, mediante una desagregación del proceso se pueden identificar tres tareas: clasificación de presencia de animal, de especie del animal en cuestión y de cantidad de animales presentes. Los datos necesarios para cuantificar la tasa de error humano en cada una de las tareas involucradas incluyen:

- **Clasificación de Presencias:**
 - **Cantidad de Imágenes Totales (N):** Se tuvo en cuenta la cantidad de imágenes obtenidas en el año 2023, ya que aún no están procesadas y, por ende, la base se encuentra completa. Se estima una actividad similar en cada campaña realizada. La base presenta 118,769 imágenes.
 - **Porcentaje de Imágenes con Presencia Animal (n):** Detallado su cálculo en el **Anexo 2**, se estima que el 31.88% de las imágenes obtenidas en el proceso de captura muestran presencia de individuos.
 - **Porcentaje de Imágenes sin Presencia Animal:** Consecuentemente, se estima que el 68.12% de las imágenes no muestran presencia de fauna y se consideran vacías.

- **Tasa de Falsos Positivos (FP):** En base a las conclusiones del análisis exploratorio realizado, se encontró que en años anteriores existió un 0.6% de incidencia de falsos positivos sobre las imágenes etiquetadas.
- **Tasa de Falsos Negativos (FN):** Dado que no es factible analizar directamente los errores de clasificación en base a los trabajos de años anteriores debido a la eliminación previa de imágenes sin animales, se ha recurrido a literatura especializada para esta estimación. Un estudio relevante (Chesser, 2012) sugiere que la tasa de error humano en el etiquetado de falsos negativos en proyectos similares es del 16.7% .
- **Tasa de Error:** Ponderando el error obtenido en falsos positivos y negativos se obtiene la tasa de error.

$$\text{Tasa de Error Humano} = 0.6\% \times 68.12\% + 16.7\% \times 31.88\% = 5.7\%$$

- **Costo de Falso Negativo (C_{FN}):** La detección de una cantidad menor de individuos de los realmente presentes tiene implicaciones financieras y operativas significativas. Este error de clasificación lleva a una subestimación de la presencia animal, lo cual podría resultar en una asignación de recursos excesiva para el siguiente periodo de restauración. Además, el esfuerzo invertido en la gestión y conservación basada en datos incorrectos es en vano, ya que no refleja la realidad en el terreno, lo que puede llevar a tomar decisiones de manejo inadecuadas. El valor que se destina a cada falso negativo es de \$300.
- **Costo de Falso Positivo (C_{FP}):** Por otro lado, si la cantidad real de animales es más baja de lo que los datos sugieren, =se subestiman los esfuerzos necesarios para las campañas de restauración y conservación, lo cual puede ser más perjudicial a largo plazo. Si no se reconoce adecuadamente la densidad de la fauna, las acciones de conservación podrían no ser suficientes para preservar las especies en peligro, lo que podría llevar a una disminución en su población e incluso a la extinción. El fracaso en la conservación no solo afecta a la biodiversidad sino que también puede comprometer la financiación futura y la confianza de los interesados en el éxito de las iniciativas de conservación. El valor que se destina a cada falso positivo es de \$1000.

- **Clasificación de Especies:**

- **Tasa de Error en el Etiquetado de Especies (E_E):** Para la estimación de este error, se tomará como referencia el cálculo realizado en un estudio de similares características y cantidad de especies involucradas, realizado en el parque nacional de Serengeti en Tanzania, el cual sugiere que los voluntarios que realizan clasificación de especies fallan en un 3.4% de las ocasiones. (Norouzzadeh et al., 2018).

- **Costo de Error en el Etiquetado de Especies (C_E):** Se corresponde con la suma de los costos asociados con un falso positivo y un falso negativo, ya que estos errores de clasificación afectan directamente a más de una especie. Cuando una especie es identificada incorrectamente (falso positivo), no solo se registra indebidamente su presencia, sino que también se omite la presencia real de otra especie (falso negativo). Por lo que el valor asignado será de \$1300.
- **Clasificación de Cantidades:**
 - **Tasa de Error en el Etiquetado de Cantidad (E_C):** Basándose en el mismo estudio mencionado en el punto anterior, la tasa de error estimada será del 10% de las imágenes observadas.
 - **Costo de Error en el Etiquetado de Cantidad (C_C):** La cuantificación se justificará con la media de los costos asociados con los falsos positivos y negativos, dado que el error de conteo puede manifestarse tanto en la sobreestimación como en la subestimación de individuos. La incertidumbre inherente a la dirección del error justifica la adopción de un enfoque promediado para calcular el impacto financiero y operacional. Por lo que el valor asignado será de \$650.

Se puede observar un resumen de estas variables en la **Tabla 4**.

	Cantidad de Imagenes	Proporcion de Presencia Animal	Tasa de Error	Costo Percibido
Clasificación de Presencia Animal	118,769	31.88%	Falso Positivo: 0.6% Falso Negativo: 16.7% Error: 5.7%	Falso Positivo: \$1000 Falso Negativo: \$300
			3.4%	\$1300
			10%	\$650

Tabla 4: Datos para Cuantificación del Error Humano

Para el cálculo de tiempo de procesamiento de cada imagen son necesarios los siguientes datos, resumidos en la **Tabla 5**, además de los previamente nombrados:

- **Cantidad de Imágenes Etiquetadas por Hora:** Conforme a lo reportado (Fennell et al. 2022), la clasificación manual se lleva a cabo a una velocidad de 300 imágenes por hora para aquellas que contienen presencia animal. En contraste, las imágenes que no presentan animales se procesan a una tasa más elevada de 1000 imágenes por hora.

- **Costo por Jornada Laboral:** El valor de la jornada laboral individual del empleado encargado del proceso de clasificación es de \$10,730 (Noviembre, 2023).

Cantidad de Imágenes Etiquetadas por Hora	Costo Jornada Laboral
Con Presencia Animal: 300 Sin Presencia Animal: 1000	\$10730

Tabla 5: Datos para Cuantificación del Tiempo de Procesamiento.

Basándose en los datos proporcionados, se procederá a calcular la tasa de error inherente al proceso de clasificación actual. Esto se realiza ponderando los errores con la proporción de base con animales y sin animales.

$$\text{Error Humano Estimado} = (FN + E_E + E_c) \frac{n}{N} + FP \times \frac{1-n}{N}$$

$$\text{Error Humano Estimado} = (16.7\% + 3.4\% + 10\%) \times 31.88\% + 0.6\% \times 68.12\% = 10\%$$

$$\text{Precisión Humana} = 1 - 10\% = 90\%$$

Como se puede observar en la suma de probabilidades, los eventos de error no tienen solapamiento, es decir son independientes, esto se explica ya que si una imagen con animal resulta en falso negativo no podrá ocurrir en los errores de clasificación de las tareas posteriores. Además, se conoce que los errores de identificación de especies se dan casi exclusivamente en aquellos animales que aparecen sólos y por una duración corta de tiempo, mientras que los errores de clasificación de cantidad lo hacen para aquellas especies que se mueven en manada y aparecen en una cantidad prolongada de imágenes, ya que la especie se hace evidente pero la cantidad se vuelve difícil de interpretar.

De manera similar, se detalla el cálculo del tiempo promedio que necesita cada imagen para su clasificación:

$$\text{Cantidad de Imágenes por Hora} = 1000 \times 68.12\% + 300 \times 31.88\% = 777$$

$$\text{Cantidad de Segundos por Imagen} = \frac{3600}{777} = 4.63 \text{ segundos}$$

Adicionalmente, es posible calcular el costo implícito asociado con la clasificación de cada imagen en el proceso vigente. Este cálculo permitirá, posteriormente, proyectar el gasto total correspondiente a cada campaña de captura de imágenes y compararlo con aquel que resultaría de los cambios propuestos.

$$\text{Costo Error Manual por Imagen} = \frac{1-n}{N} \times FP \times C_{FP} + \frac{n}{N} \times (FN \times C_{FN} + E_E \times C_E + E_c \times C_c)$$

$$\text{Costo Error por Imagen} = 68.12\% \times 0.6\% \times \$1000$$

$$+ 31.88\% \times (16.7\% \times \$300 + 3.4\% \times \$1300 + 10\% \times \$650) = \$54.87$$

$$\text{Costo Tiempo por Imagen} = 4.63 \text{ seg} \times \frac{\$10,730}{8 \times 60 \times 60} = \$1.73$$

$$\text{Costo Total por Imagen} = \$54.87 + \$1.73 = \$56.6$$

Al proyectar estos costos al volumen total de imágenes presentes en la base de datos, se obtiene una cifra representativa del costo total acumulado:

$$\text{Costo Total Actual} = 118,769 * \$56.6 = \$6,722,325.4$$

Justificación del Proyecto

El análisis financiero del proyecto revela que, con la implementación de las estrategias propuestas, el beneficio potencial máximo se estima en \$6,722,325.4. Esta cifra representa el punto óptimo de eficiencia económica que se podría alcanzar al optimizar los procesos actuales y mitigar los costos asociados con las tasas de error en la clasificación. Se anticipa una reducción significativa en la tasa de error inherente al procesamiento manual, actualmente estimada en un 10%. Paralelamente, se busca optimizar el tiempo de procesamiento, con el objetivo de superar la capacidad actual de 777 imágenes por hora. Estas mejoras serán clave para aumentar la eficiencia y precisión del sistema.

Tras un meticuloso análisis de la situación vigente, la ejecución de este proyecto se justifica por las significativas mejoras que promete ofrecer en distintos ámbitos clave:

- **Optimización del Tiempo Operativo y Precisión de Etiquetado:** El tratamiento actual de las imágenes para derivar conclusiones sobre la condición de las poblaciones del ecosistema requiere extensos períodos de labor y está sujeto a errores intrínsecos a la clasificación manual. La adopción de una solución superadora promete reducir la necesidad de intervención humana, reservándose para instancias que requieran validación específica, disminuyendo así la frecuencia de errores. Se justifica el proyecto en caso de que la solución planteada logre reducir el tiempo de procesamiento mejorando la precisión de etiquetado actual definida como 90% por el análisis.
- **Refuerzo de la Comunicación Estratégica:** Los hallazgos derivados del proyecto son fundamentales para la comunicación con los patrocinadores de la fundación. El incremento en la precisión no solo asegura la continuación del soporte financiero actual sino que también potencia la captación de nuevos benefactores interesados en apoyar las iniciativas de conservación. La transparencia y la demostración del impacto tangible de los esfuerzos son esenciales para el mantenimiento de la financiación de las operaciones.
- **Estandarización del Etiquetado:** Se ha detectado una falta de uniformidad y estandarización en el proceso de etiquetado de imágenes, lo que puede conducir a inconsistencias y complicaciones en la interpretación de datos. La incorporación de un sistema semi-automatizado estandariza la salida de datos, mejorando considerablemente la coherencia y fiabilidad de la información procesada.
- **Potencial de Ampliación y Replicabilidad:** Es importante destacar que esta solución no se limita a una región específica de *Rewilding*, sino que puede

replicarse fácilmente en otras regiones donde se lleven a cabo proyectos similares. Además, existe la posibilidad de compartir esta solución con otras fundaciones de conservación en todo el mundo, lo que contribuiría a mejorar la eficacia y la calidad de los esfuerzos de conservación a nivel global. Esta escalabilidad y potencial de exportación hacen que este proyecto sea valioso no sólo para nuestra fundación, sino también para la comunidad internacional de conservación de la vida silvestre. Además, la acumulación y análisis de un volumen mayor de datos fiables no sólo robustece la fiabilidad de los esfuerzos actuales sino que también incrementa considerablemente la precisión de los modelos predictivos empleados. Por ende, este enriquecimiento de la base de datos subyace en la mejora continua de los resultados obtenidos por los modelos, optimizando las estrategias de conservación a largo plazo

Escenario Automatizado

Los resultados esperados se presentan con base en un análisis comparativo y benchmarking de estudios relevantes, particularmente las investigaciones llevadas a cabo (Norouzzadeh et al. 2018) y (Fennell et al. 2022). Este análisis incluye las tasas de error y los tiempos de procesamiento asociados con la implementación de soluciones de inteligencia artificial para abordar la clasificación automática en imágenes de cámaras trampa (**Tabla 6**).

	Tasa de Error en Presencia	Tasa de Error en Especie	Tasa de Error en Cantidad	Imagenes por Hora
Clasificación Manual	Falso Positivo: 0.6% Falso Negativo: 16.7% Error: 5.7%	3.4%	10%	Con Presencia Animal: 300 Sin Presencia Animal: 1000
Clasificación Automática	Falso Positivo: 2.49% Falso Negativo: 3.8% Error = 2.91%	5.1%	36.9%	6480

Tabla 6: Datos para Cálculo de Indicadores en Clasificación Automática

En un análisis preliminar, es evidente que la frecuencia de errores de clasificación por falsos negativos es considerablemente más baja en los sistemas automatizados en comparación con la evaluación humana. Sin embargo, en otros aspectos de modelado, la Inteligencia Artificial aún no supera el desempeño humano. No obstante, la superioridad de los modelos se manifiesta en su capacidad para procesar un volumen significativamente mayor de imágenes por hora.

Los datos presentados se emplearon para la cuantificación conforme al método establecido en el análisis de la situación vigente, presentado en el **Anexo 3**. De este proceso, emergen los indicadores que se exponen en la **Tabla 7**.

	Tasa de Error	Tiempo por Imagen	Costo Total
Clasificación Manual	10%	4.63 seg	\$56.5
Clasificación Automática	16.3%  62.97%	0.56 seg  -87.91%	\$115.59  104.58%

Tabla 7: Resultados de los Indicadores para la Clasificación Manual y Automática

Los resultados reflejan un aumento significativo en la tasa de error al pasar de la clasificación manual a la automática, con un incremento del 62.97%. Sin embargo, la clasificación automática ofrece una mejora sustancial en la eficiencia, reduciendo el tiempo de procesamiento por imagen en un 87.91%. A pesar de estos avances en rapidez, el costo total por imagen estimado de la clasificación automática es mayor, con un incremento del 104.59% comparado con el método manual.

Estos hallazgos sugieren que un enfoque completamente automatizado puede no ser la solución más adecuada para abordar el problema en cuestión de manera efectiva.

Escenario Semi-Automatizado

Un sistema semi-automatizado combinaría la velocidad y eficiencia del procesamiento automático para las imágenes con alta confianza de clasificación, mientras que las imágenes con menor confianza serían abordadas mediante validación humana, con una sugerencia de clasificación. Este método híbrido pretende asegurar la precisión sin sacrificar la eficiencia, optimizando los costos totales y manteniendo la tasa de error dentro de límites aceptables. La validación humana en puntos críticos del proceso garantizará la calidad de los datos, lo que es esencial para la fiabilidad y aplicabilidad de los resultados en iniciativas de conservación y estudios de biodiversidad.

Se plantea el escenario de validación, al que se le han asociado indicadores clave que incluyen la tasa de error, el tiempo y el costo, cada uno con justificaciones específicas y valores asignados:

- **Tasa de Error:** La tasa de error propuesta se alinea con la asignada al proceso de clasificación manual con una reducción de un 50%, dado que la naturaleza del entorno de validación es más amena para el usuario reduciendo la posibilidad de confusión. Partiendo de una tasa de error base de 10%, se anticipa que la tasa de error en la validación sea del 5%, ofreciendo una mejora tangible en la calidad de la clasificación.
- **Tiempo por Imagen:** El tiempo estimado para la validación de imágenes se mejorará significativamente mediante la implementación de la interfaz visual detallada en el **Anexo 1**. Esta interfaz facilita una revisión y clasificación ágil por parte del especialista, gracias a la disposición estratégica de las imágenes y las opciones de etiquetado en un único espacio de trabajo. Además, la capacidad de etiquetar secuencias de imágenes simultáneamente reduce el volumen total de imágenes a validar individualmente. El tiempo que

se estima que un usuario dedica a validar una imagen en la interfaz es de 3 segundos, como propuesta de reducción frente al tiempo estimado previamente. Según el análisis de consecución de apariciones desarrollado en la exploración de datos los animales detectados permanecen en el radar de la cámara por una media de 11 imágenes, se estima que el usuario de la interfaz precisará un esfuerzo de detención para identificar la clasificación de dos de las imágenes: la primera y la última. Al resto de las imágenes en la rafaga se les destina un segundo en ver que correspondan a la etiqueta. Por lo tanto, el tiempo asignado para la validación de cada imagen se ajusta en función del promedio de imágenes consecutivas detectadas o “ráfagas”, resultando en un tiempo de validación de 1.36 segundos por imagen, de la siguiente forma.

$$Tiempo\ Validación\ por\ Imagen = \frac{9 \times 1 + 2 \times 3\ segundos\ por\ Rafaga}{11\ imágenes\ en\ Rafaga} = 1.36\ segundos$$

- **Costo de Error:** En términos de costos se mantiene la lógica descrita en el análisis de la sección “situación actual”.
- **Costo de Validación:** Cada imagen que se tome para la validación tiene implícita un costo operativo. Sabiendo que un experto cobra \$10730 por jornada laboral y necesita, en promedio, 1.36 segundos para validar una imagen, el costo por imagen será de \$0.51.

Dentro de este marco, se definirán umbrales de confianza para discernir que porcentaje de imágenes será procesado autónomamente por el modelo de inteligencia artificial y cual requerirá verificación humana. Aquellas predicciones del modelo que no alcancen el umbral establecido serán redirigidas a los especialistas para su confirmación mediante la interfaz especializada. Esta metodología busca optimizar la exactitud del modelo mientras se gestiona eficientemente la intervención manual, aplicando una heurística diseñada para minimizar costos asociados. El algoritmo y su implementación se explican con detalle en el **Anexo 4**, que incluye el código pertinente.

	Tasa de Error en Presencia	Tasa de Error en Especie	Tasa de Error en Cantidad	Imagenes por Hora
Clasificación Manual	Falso Positivo: 0.6% Falso Negativo: 16.7% Error: 5.7%	3.4%	10%	Con Presencia Animal: 300 Sin Presencia Animal: 1000
Clasificación Automática	Falso Positivo: 2.49% Falso Negativo: 3.8% Error = 2.91%	5.1%	36.9%	6480
Clasificación Semi-Automática	Validación: 5.79%			Validación con Ráfagas: 2647

Tabla 8: Datos para Cálculo de Indicadores en Clasificación Semi-Automática

Siguiendo un procedimiento similar al escenario previamente descrito, se han calculado los indicadores correspondientes al entorno semi-automatizado, tal y como se expone en el **Anexo 3**. Los resultados de este análisis se presentan de manera estructurada en la **Tabla 9**.

	Tasa de Error	Tiempo por Imagen	Costo Total
Clasificación Manual	10%	4.63 seg	\$56.5
Clasificación Automática	16.3%  62.97%	0.56 seg  -87.91%	\$115.59  104.58%
Clasificación Semi-Automática	6.13%  -38.7%	1.45 seg  -68.75%	\$22.69  -59.83%

Tabla 9: Resultados de los Indicadores para la Clasificación Manual, Automática y Semi-Automática

La situación semi-automatizada muestra una tasa de error del 6.13%, lo que representa una disminución del 38.7% en comparación con la clasificación manual que tenía un error del 10%. En términos de eficiencia temporal, el tiempo por imagen es de 1.45 segundos, lo que indica una reducción del 68.75% frente al tiempo de clasificación manual de 4.63 segundos. En cuanto al costo por imagen, la clasificación semi-automática incurre en un gasto de \$22.69, lo que refleja una disminución del 59.83% en comparación con el costo de la clasificación manual, que es de \$56.5.

Resultados Potenciales

Como punto principal, la implementación de un sistema de clasificación semi-automático en el escenario presentado se justifica dada la mejora significativa en la tasa de error de la clasificación, lo que permitirá a la organización poseer información más fidedigna para el análisis.

La reducción en el tiempo de procesamiento por imagen es sustancial, bajando de 4.63 segundos a solo 1.45 segundos, como resultado de la baja en la cantidad de imágenes a revisar y la posibilidad de clasificación de rafagas. Este incremento en la velocidad de procesamiento puede traducirse en un aumento notable de la productividad y la posibilidad de procesar un volumen más grande de imágenes en un período más corto.

Además, desde una perspectiva económica, la clasificación semi-automática resulta ser más rentable, reduciendo los costos en un 26.59% con respecto a la clasificación manual. Esto puede permitir la realocación de recursos financieros a otras áreas críticas del proyecto o investigación.

En resumen, la solución semi-automatizada repercute en una mejora en todos los indicadores analizados, lo que la convierte en la mejor opción frente a los métodos de clasificación manual o completamente automáticos, según los escenarios presentados.

Enfoque de Solución

Para implementar la solución propuesta, se dispone de una base de datos estructurada que asocia las imágenes con los siguientes atributos:

- **Presencia de Animal:** Variable booleana que determina si la imagen presenta o no un animal.
- **Especie:** En los casos donde se identifica un animal, este atributo clasifica la especie correspondiente.
- **Cantidad:** Cuando se detecta un animal, esta variable cuantifica el número de individuos presentes.

Como se señaló anteriormente, no se cuenta con un conjunto de datos etiquetado que carezca de animales. Por lo tanto, se utilizó el modelo *Megadetector* para identificar, de manera no supervisada, imágenes sin presencia animal. Se estableció como criterio que no se detectara ninguna entidad con 100% de confianza. Se seleccionó una muestra representativa de 250 imágenes de cada cámara para abarcar diversos contextos y evitar un entrenamiento sesgado. A través de este proceso, se identificaron un total de 2,834 imágenes con alta probabilidad de no contener animales, las cuales se utilizarán para entrenar los modelos.

La experimentación se basará en la metodología de Milestones, donde se fijarán objetivos intermedios dentro de los resultados esperados. Este enfoque metódico y estructurado asegura la integridad y la relevancia de los datos utilizados, y sienta las bases para el desarrollo de un modelo de Inteligencia Artificial robusto y preciso, que contribuirá significativamente a los esfuerzos de conservación y estudio de la biodiversidad. A continuación se presentan los hitos a perseguir para la resolución:

1. **Predicción de Presencia de Animales:** El objetivo primordial es afinar la precisión del modelo al determinar la presencia de animales en las imágenes capturadas.
2. **Predicción de Especies:** Dada la distribución desigual de especies, que ha sido previamente examinada, la tarea de clasificar las mismas plantea ciertos desafíos. Para abordar este reto, se ha optado por establecer una secuencia de cuatro objetivos estratégicos con el propósito de optimizar el rendimiento predictivo. Se continuará con el siguiente nivel de dificultad de predicción únicamente si se lograra con el anterior un nivel mínimo de Accuracy del 80%.
 - a. **Detección de Guanacos:** Dado que los guanacos constituyen la mayoría de las observaciones, con un 68.8% del total del conjunto de datos, se diseñará un modelo específico para determinar si el animal capturado en la imagen corresponde a un guanaco o no.
 - b. **Clasificación por Categorías:** Tras alcanzar un nivel de exactitud satisfactorio en la identificación de guanacos (con un umbral mínimo del 80%), se avanzará hacia la discriminación de especies basándonos en categorías más amplias: mamíferos herbívoros, mamíferos carnívoros y aves.
 - c. **Top 5 de Especies Probables:** En función de proporcionar un panorama más detallado y teniendo en cuenta las incertidumbres inherentes, el modelo ofrecerá las tres especies más probables con sus respectivas probabilidades. Esto brinda una visión con más matices, especialmente en casos donde el modelo puede no estar completamente seguro.
 - d. **Identificación Detallada de Especies:** Una vez que la clasificación por categorías demuestre ser robusta (superando el 80% de precisión), se

intentará discernir entre las distintas especies que habitan el ecosistema en cuestión.

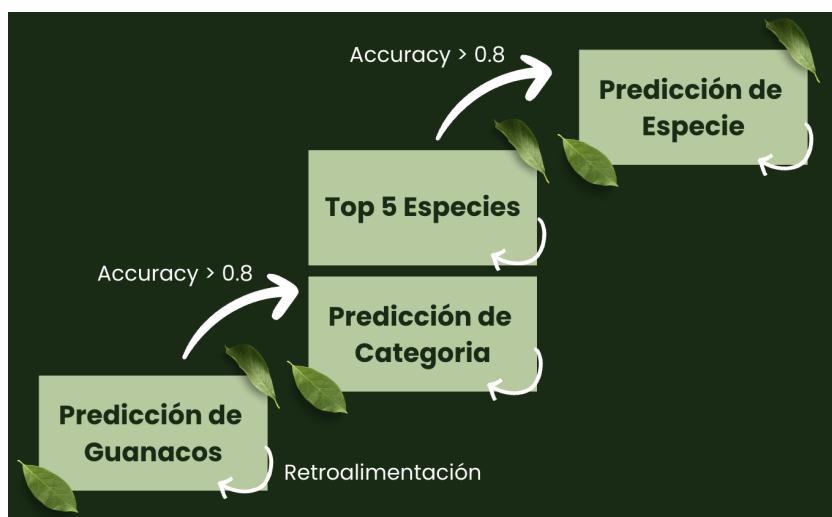


Figura 32: Diagrama de Framework Predicción de Especies.

3. **Predicción de Cantidades:** Uno de los retos más intrincados en el análisis de imágenes es la correcta cuantificación de animales presentes en una captura. Los modelos actuales que buscan contabilizar objetos o entidades en imágenes no han alcanzado un nivel óptimo de eficacia. En vista de este escenario, se han establecido tres objetivos estratégicos para enfrentar el desafío. Nuevamente, se avanzará con la dificultad de predicción únicamente si se alcanzara un mínimo de Accuracy del 80%.

- Discriminación entre Individuo Aislado y Grupo:** En base al análisis previo de la distribución de animales en las imágenes, se observa que el 72,1% de las capturas muestran animales en solitario. Es notable mencionar que las especies que tienden a aparecer en grupos son, en particular, los guanacos y las aves. Esta característica intrínseca del conjunto de datos puede ser aprovechada para diseñar un modelo capaz de distinguir entre la presencia de un único individuo y múltiples entidades.
- Detección de Cantidad:** Si la etapa anterior demuestra ser robusta, alcanzando un nivel de precisión superior al 80%, el siguiente paso será el desarrollo de un algoritmo especializado que estime de forma precisa el número de individuos presentes en cada imagen.

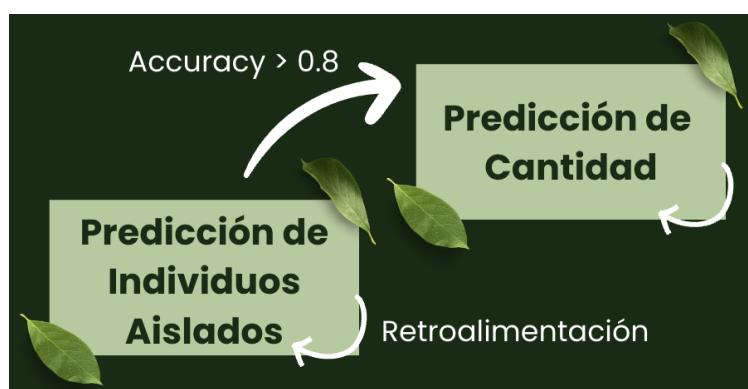


Figura 33: Diagrama de Framework Predicción de Cantidad.

Metodologías a Implementar

En el desarrollo del proyecto, se exploraron tres arquitecturas de vanguardia en el ámbito de las redes neuronales profundas: VGG16, InceptionV3 y ResNet.

- **VGG16:** Desarrollada por el Visual Graphics Group de la Universidad de Oxford, esta arquitectura se destaca por su simplicidad, utilizando exclusivamente capas convolucionales de 3x3 seguidas de capas de pooling para reducción de dimensionalidad. A pesar de su estructura aparentemente simple, VGG16 consiste en 16 capas y ha demostrado una eficacia excepcional en la clasificación y localización de imágenes.
- **InceptionV3:** Originaria de Google, la arquitectura Inception es conocida por su módulo "Inception", que permite a la red realizar convoluciones a diferentes escalas simultáneamente. Esta característica permite a InceptionV3 captar patrones en diferentes niveles de granularidad y ha mostrado ser particularmente eficiente en la gestión de recursos computacionales.
- **ResNet:** Desarrollada por Microsoft Research, ResNet introdujo la noción de "conexiones residuales". Estas conexiones permiten que las activaciones salten algunas capas y se sumen a las activaciones de capas posteriores, facilitando el entrenamiento de redes muy profundas al abordar el problema de desvanecimiento del gradiente. ResNet ha establecido varios récords en precisión de clasificación, siendo una elección preferida para tareas que requieren una profundidad considerable.

Para maximizar el potencial de estas arquitecturas en el contexto del proyecto, se implementó una estrategia de transferencia de aprendizaje utilizando modelos previamente entrenados en el dataset ImageNet, un vasto repositorio de imágenes etiquetadas que ha sido pilar en el avance del aprendizaje profundo en la visión por computadora. La idea detrás de este enfoque es beneficiarse de las características genéricas aprendidas por las redes en conjuntos de datos extensos y adaptarlas a las particularidades del dataset específico del proyecto. Esta combinación busca capitalizar la rica representación de ImageNet con la especificidad del conjunto de datos en estudio, con el objetivo de alcanzar una performance de clasificación óptima.

Seteo de Experimentación

Una vez los datos han sido depurados, estandarizados, validados y están en condiciones para su manipulación, se inicia la fase experimental del proyecto. Se cuenta con una base de datos etiquetada compuesta por 6049 registros verificables, aptos para su utilización en las fases de entrenamiento y validación.

En el contexto de aprendizaje automático y, en particular, en el ámbito del procesamiento de imágenes, se puede representar imágenes con tensores a través de sus dimensiones: alto, ancho y canales de color. Por ejemplo, una imagen a color de tamaño 150x150 pixeles se podría representar con un tensor de dimensiones y tres canales de color: rojo, verde y azul. Para la implementación y gestión eficiente del proyecto, la base de datos consolidada contendrá la ruta a la imagen, el tensor generado y las etiquetas respectivas de la imagen.

Dicha base es segmentada en conjuntos de entrenamiento y prueba, en una proporción del 80% y 20%, respectivamente. Dado que existe desigualdad en todos los casos de predicción, se implementarán mecanismos de balanceo en la separación en subconjuntos para asegurar una distribución equitativa de los datos durante la segmentación. Este mecanismo es conocido como Stratified Split sobre la clase objetivo. Además, esta partición se realiza con una semilla para que todos los modelos se entrenen y hagan pruebas sobre los mismos datos y que la aleatoriedad no favorezca a ninguno.

El conjunto de prueba se preservará inalterado hasta la conclusión de la fase experimental, momento en el que se utilizará por única vez para evaluar el rendimiento del modelo. Este procedimiento asegura que dicho conjunto refleje de manera fidedigna la realidad, al no estar sujeto a modificaciones ni haber sido empleado durante el entrenamiento.

En situaciones donde el conjunto de datos de entrenamiento exhibe un desbalance significativo entre las clases, se emplearán estrategias de Data Augmentation con el propósito de potenciar la capacidad de generalización del modelo y equilibrar la distribución de las clases. Estas técnicas involucran la aplicación de transformaciones sutiles a las imágenes, incluyendo ajustes en luminosidad y contraste, rotaciones, introducción de perturbaciones aleatorias y reflexiones. El objetivo de estas transformaciones es enriquecer el conjunto de datos con variantes únicas, previniendo así el overfitting al evitar la simple duplicación de ejemplares existentes.

En cuanto al conjunto de entrenamiento, este se subdivide en datos para entrenamiento y para validación, aplicando técnicas de Cross Validation. Con el objetivo de optimizar la capacidad de generalización del modelo, evitando el sobreajuste a los datos de entrenamiento, mediante la selección correcta de hiper-parámetros.

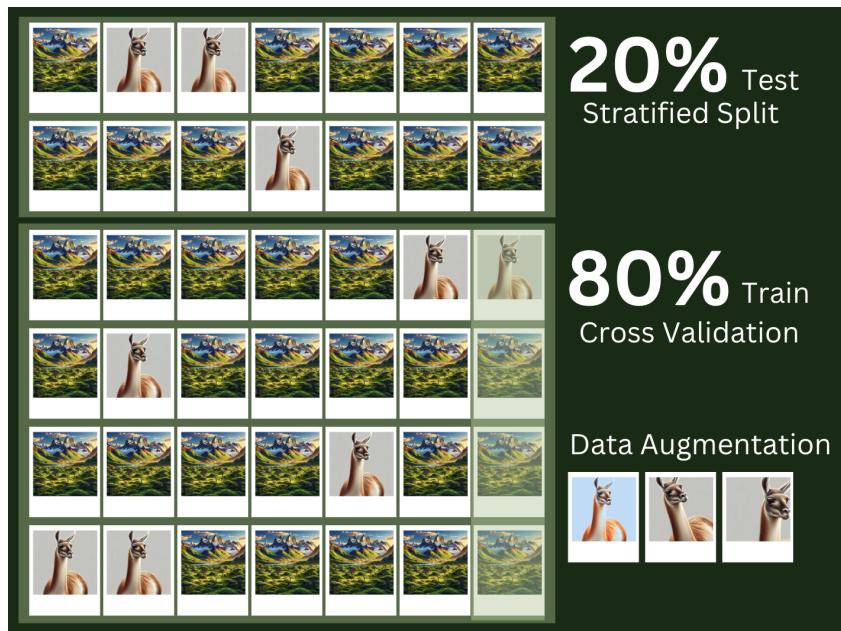


Figura 34: Seteo de la Experimentación

Las métricas utilizadas para la evaluación de los modelos de clasificación mencionados son:

- **Accuracy:** Proporción de predicciones correctas sobre el total de predicciones.

$$Accuracy = \frac{TP + TN}{Total}$$

Es una métrica fundamental para evaluar el rendimiento general del modelo. La clasificación se considerará satisfactoria de lograrse una exactitud mayor a la precisión en la tarea presentada por la situación actual, la cual es la exactitud esperada para un humano.

- **Recall:** Con el objetivo de medir la proporción de ejemplos positivos que el modelo ha identificado correctamente.

$$Recall = \frac{TP}{TP + FN}$$

En proyectos de conservación es crucial asegurar que no se omitan avistamientos de especies, lo cual es fundamental para la identificación precisa de falsos negativos. Dada la naturaleza imperativa de la preservación de la biodiversidad, se da prioridad al recall para maximizar la detección de individuos y garantizar que ningún espécimen relevante sea inadvertido en el análisis.

- **Precisión:** Esta métrica se centra en los Falsos Positivos. La precisión mide la proporción de identificaciones correctas frente al total de identificaciones positivas realizadas por el modelo. Una precisión alta significa que hay una menor proporción de Falsos Positivos.

$$Precisión = \frac{TP}{TP + FP}$$

Al priorizar la precisión, se minimiza la incidencia de falsos positivos, lo cual es esencial para evitar el desvío de recursos de conservación hacia especies no objetivo.

- **Average Precision (AUC-PR):** Para obtener una visión holística del comportamiento del modelo a lo largo de diferentes umbrales de decisión, se estudiará la curva Precisión-Recall y, en particular, el área bajo esta curva (AUC-PR). Un valor de AUC-PR más alto indica un rendimiento general superior en la clasificación, teniendo en cuenta tanto la Precisión como el Recall en cada umbral.

$$AUC - PR = \int_0^1 Precision(r) dr \text{ siendo } r = Recall$$

En resumen, los modelos se evalúan maximizando el Average Precision dado un accuracy mayor o igual a la precisión humana identificada por la tarea. Así, los modelos buscarán maximizar su poder de clasificación siendo capaces de mejorar la

exactitud de un humano. Finalmente, el modelo elegido será aquel que logre el máximo Average Precision.

$$\text{Función Objetivo} = \max(\text{AUC-PR}) \text{ sujeto a Accuracy } \geq P_t$$

Siendo P_t la precisión humana de la tarea t . Los humanos son capaces de identificar una presencia animal en su 96.6%. La tarea de detección de especies tiene un accuracy humano del 96.6% y de cantidades es el 90%.

Además, las métricas de Recall y Precisión serán utilizadas para definir los thresholds de validación. Es decir, la confianza en la predicción necesaria para llegar a valores máximos en Recall o Specificity, dependiendo el caso. Esta minimización se ve limitada por el costo beneficio que implica validar (**Anexo 4**).

Desarrollo de la Solución

Tarea 1: Detección de Imágenes con Animales.

En la ejecución de esta tarea los modelos procesan cada imagen transformada en tensores, como fue previamente descrito, y llevan a cabo una clasificación binaria, proporcionando la probabilidad que determinan con respecto a la presencia de un animal.

Para cada arquitectura pre-entrenada con los pesos de ImageNet se llevó a cabo un proceso de optimización mediante la afinación de hiperparámetros. Se varían tanto el número de capas como el valor de la tasa de aprendizaje con el objetivo de maximizar la precisión media (Auc-Pr).

De acuerdo con las evaluaciones, todos los modelos superan el nivel de Accuracy requerido para trascender la tasa de acierto humano inherente al proceso, establecida en el 88.43%. Esta referencia se ilustra en la **Figura 35** mediante una línea gris intermitente. Dadas estas condiciones, se selecciona el modelo VGG16, por su superioridad en términos de Average Precision, registrando un destacado 98.3%. Este modelo llegó a su máxima performance sin capas agregadas y con un learning rate del 0.001. Si bien ResNet50 muestra resultados comparables, se prioriza VGG16 debido a su mayor eficiencia computacional.

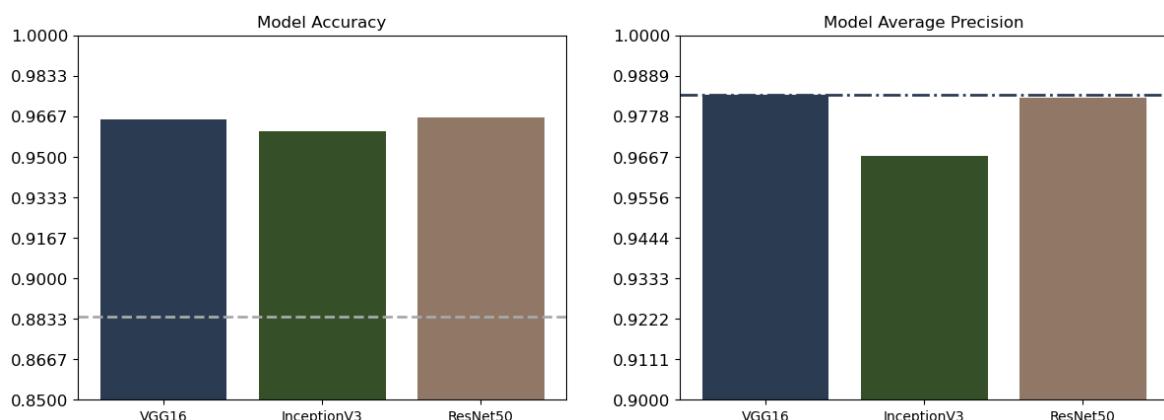


Figura 35: Accuracy de los Modelos de Presencia de Animal.

Figura 36: Average Precision de los Modelos de Presencia de Animal.

El modelo óptimo se empleó para establecer los márgenes de confianza para los cuales se definirá el rango de imágenes para las que se solicitará validación por parte de expertos. Se sometieron los resultados de prueba al algoritmo de optimización de costo beneficio entre el costo de error y de validación. Los umbrales óptimos del área de validación se situaron entre 0.01 y 0.96. Es decir, las imágenes cuya confianza de clasificación se sitúe en este intervalo, deberán ser validadas.

En el caso del modelo predictivo de presencia de animales, se determinó que validando entre los umbrales del 0.01 al 0.96, que representan el 3.88% de las imágenes, se logra un coste óptimo.

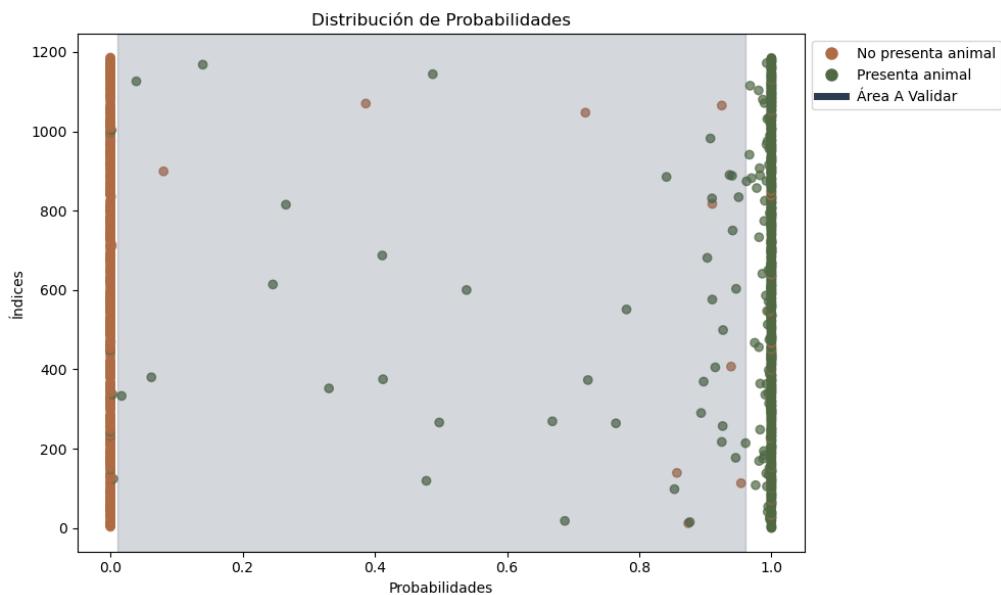


Figura 37: Distribución de Probabilidades y Área a Validar en Presencia de Animal.

Se ha establecido que el 95.94% de las imágenes que no presentan animales pueden ser excluidas de la validación manual. Esto indica una alta confiabilidad del sistema en identificar correctamente ausencias de fauna en las capturas. Por otro lado, un 92.74% de las imágenes que sí contienen animales también pueden ser descartadas de la validación manual.

Tarea 2: Detección de Especies.

En el proceso de modelización para la detección de especies, se seleccionaron exclusivamente aquellas imágenes que evidencian la presencia de animales. Es relevante destacar que el 4.2% de estas capturas muestra múltiples especies simultáneamente. Dada la potencial ambigüedad que estas imágenes podrían introducir en el modelo, debido a su patrón confuso para el modelo, se tomó la decisión de excluir tales registros del conjunto de datos. Así, el corpus final de entrenamiento con el que se trabajó consta de 2,162 imágenes.

En primera instancia, y a modo de primer hito, se centró la estrategia de detección de especies en la identificación específica de guanacos, dada su amplia participación en el dataset. Para estructurar los datos, se implementó la técnica de Stratified Split, asignando el 80% de los datos al conjunto de entrenamiento. Dado que las imágenes de guanacos son preponderantes en este conjunto, se recurrió a técnicas de Data Augmentation para ampliar el repertorio de imágenes de otras especies. De esta manera, se consiguió un conjunto de datos equilibrado entre imágenes de guanacos y de otras especies.

Durante la etapa de entrenamiento, se analizaron manualmente los casos extremos de falsos positivos y negativos en la predicción del modelo, para investigar la existencia de errores sistemáticos en la inteligencia artificial. Este proceso reveló ciertas discrepancias en el etiquetado original, particularmente situaciones en las que se asignó una especie incorrectamente a una imagen, es decir, dado el error humano, existían fallas en la supervisión. Tales anomalías se ajustaron adecuadamente para prevenir distorsiones en el aprendizaje de los modelos y, consecuentemente, se consiguió una mejora en su desempeño.

A pesar de los esfuerzos de modelización, ninguno de los modelos propuestos supera la eficacia discernible del análisis humano. Sin embargo, el modelo VGG16 se destaca entre sus pares, exhibiendo una Average Precision del 96.66%, como se muestra en la **Figura 39**.

Debido a que dos de los modelos alcanzaron un índice de accuracy superior al 0.8 propuesto, como se observa en la **Figura 38**, se consideró oportuno evaluar modelos que predicen la categoría del animal y las cinco principales especies potenciales. Ambos enfoques fueron examinados en el mismo paso, debido a una estimación de complejidad comparable. Sin embargo, en el caso de la predicción de categoría animal, los resultados no excedieron un índice de precisión de 0.45 en datos de prueba, lo que indica que este modelo se desempeña por debajo de la simple aleatoriedad y, por lo tanto, debe ser descartado. Similarmente, el modelo que propone un 'top 5' de especies también mostró limitaciones, sin alcanzar un índice de precisión superior al 0.27. Conociendo esto, es en vano evaluar un modelo de predicción de especies puntuales.

Se utilizará el modelo de VGG16 que clasifica la presencia de guanacos para extraer de la validación manual aquellos que tienen una confianza suficiente de serlo. Nuevamente esa confianza será definida por el algoritmo de optimización de costo beneficio entre el costo de error y de validación.

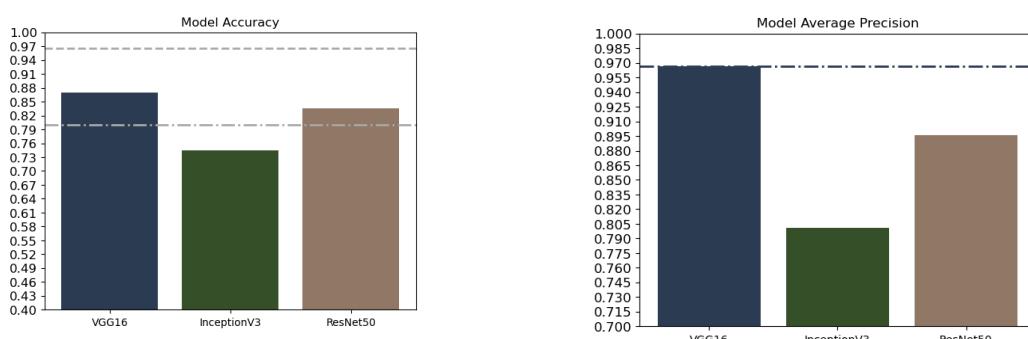


Figura 38: Accuracy de los Modelos de Presencia de Guanacos.

Figura 39: Average Precision de los Modelos de Presencia de Guanacos.

La estrategia propuesta reduce eficazmente el número de imágenes que contienen animales y que requieren validación, excluyendo específicamente aquellas con alta probabilidad de contener un guanaco. Solo las imágenes que presentan una probabilidad menor al 71% de representar un guanaco serán sometidas a validación, como se observa en la **Figura 40**, este enfoque promete mejorar la precisión humana y ahorrar el esfuerzo de validación de una parte significativa de las imágenes. Se calcula que la proporción de imágenes sujetas a revisión será del 50.81% del total de imágenes que exhiben animales.

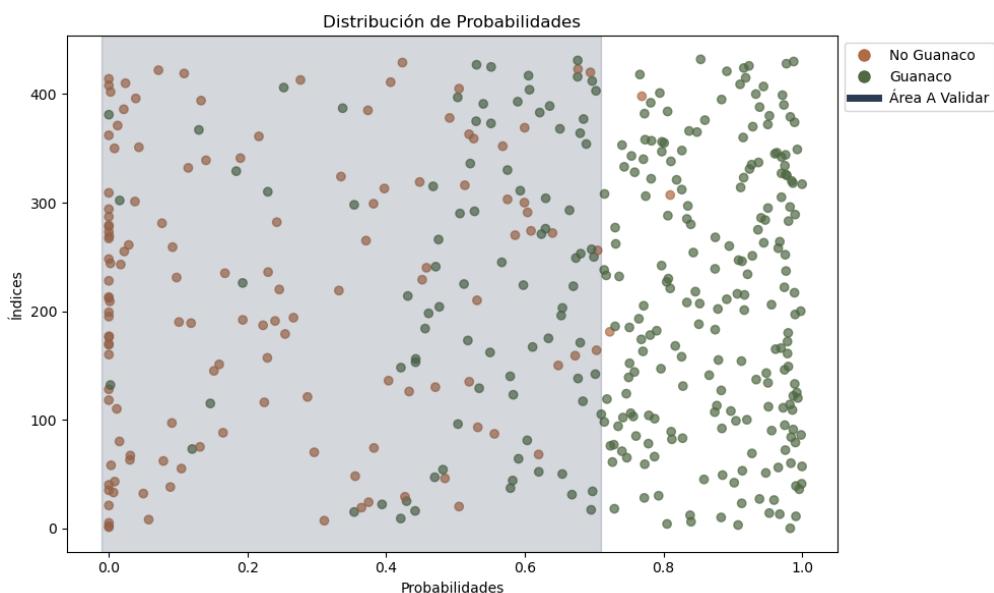


Figura 40: Distribución de Probabilidades y Área a Validar en Presencia de Guanacos.

Tarea 3: Recuento de Cantidades.

De acuerdo con estudios similares (Norouzzadeh et al., 2018), esta tarea ha demostrado ser la que obtiene menor grado de confiabilidad. El estándar de referencia para esta tarea se ha establecido en un 63.1% de precisión. A partir de este antecedente, se anticipa que podría ser el aspecto menos susceptible a la automatización. Adicionalmente, se carece de un conjunto de entrenamiento sólido y coherente, ya que las etiquetas no se asignaron siguiendo un criterio uniforme en relación con la imagen correspondiente, tal como se identificó en etapas anteriores.

Partiendo de esa base, se decide evaluar el primer objetivo estratégico que implica la detección de la presencia de un único animal. Se decidió utilizar, para esta tarea, el modelo de *Megadetector*, dado que su diseño consiste en detecciones individuales dentro de la imagen provista.

El proceso implementado con *MegaDetector* involucra una rigurosa evaluación de todas las imágenes que han sido previamente clasificadas como conteniendo presencia animal. En esta fase, el modelo se encarga de identificar y reconocer

individuos específicos dentro de estas imágenes. Posteriormente, se aplica un criterio de filtrado, descartando aquellas detecciones cuya confianza sea inferior al 20%. Este umbral se ha establecido en base a las recomendaciones óptimas de sensibilidad del modelo. Finalmente, se procede a un conteo meticuloso de las detecciones realizadas en cada imagen, presentando el número total de detecciones identificadas. Este enfoque asegura una evaluación precisa y eficiente de la presencia animal en el conjunto de datos estudiado.



Figura 41: Representación de Funcionalidad de Megadetector

La adopción de esta metodología permitirá obtener resultados no supervisados, lo cual es crucial debido a la incongruencia observada entre las imágenes y sus correspondientes etiquetas en la base de datos de *Rewilding*. Esta metodología es esencial para realizar un análisis detallado de los casos con alta probabilidad de falsos positivos y negativos. Un examen minucioso de las imágenes que presentan una alta probabilidad de errores en el etiquetado **reveló que un 83,68% de estas contenía inexactitudes en sus etiquetas**. Este hallazgo subraya la importancia de nuestro enfoque para asegurar la precisión en la clasificación y el etiquetado de datos. A pesar de que se han identificado numerosas inconsistencias en el etiquetado de la cantidad de elementos en las imágenes, estas no se han considerado al evaluar los errores humanos en el proceso de etiquetado. Cabe destacar que el criterio empleado para la anotación de cantidades no siempre refleja de manera precisa el contenido de la imagen individual. En ciertas ocasiones, el etiquetado parece basarse en la variación observada entre una imagen y la siguiente, en lugar de una descripción detallada y aislada de cada imagen en sí.

Con el fin de perseguir el primer objetivo, de identificación de animales aislados, se evalúa el modelo con respecto a la base de datos corregida. *Megadetector* logra un nivel de accuracy de 90.74%, superando la capacidad humana (90%), exhibiendo una Average Precision del 94.37%. Se llevaron a cabo evaluaciones adicionales empleando modelos fundamentados en aprendizaje supervisado utilizando el

conjunto de datos corregido. Sin embargo, estos modelos no replicaron el nivel de eficacia observado en *Megadetector*, como se documenta en la **Figura 42 y 43**.

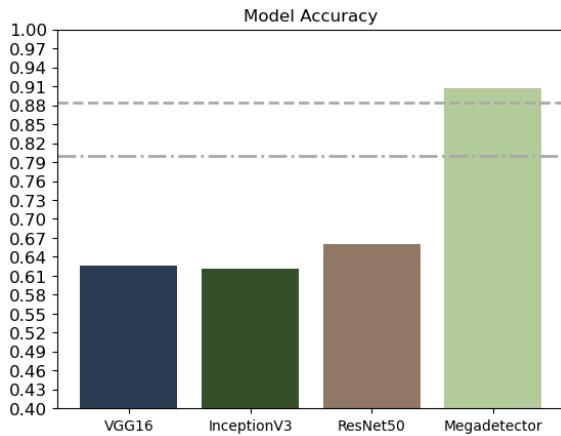


Figura 42: Accuracy de los Modelos de Presencia de Animales Aislados.

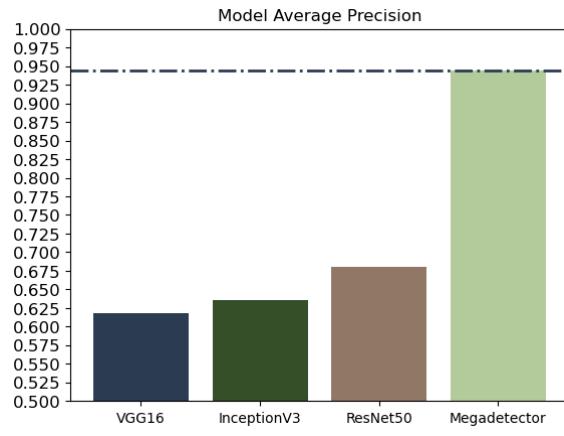


Figura 43: Average Precision de los Modelos de Presencia de Animales Aislados.

Cómo el modelo de *Megadetector* fue lo suficientemente preciso, superando el 80%, se evalúa el siguiente nivel de dificultad dado por la clasificación exacta de cantidades. Se lleva a cabo el mismo procedimiento. Los resultados dieron un accuracy menor al implícito en la tarea manual, siendo este del 68.21%. Por lo que no se adopta el modelo.

Continuando con la identificación de presencia de animales aislados, se aplica la heurística que define el umbral a validar. Se determinó que validando entre los umbrales del 0.01 al 0.69, que representan el 58.21% del total de imágenes que exhiben animales, se logra un coste óptimo.

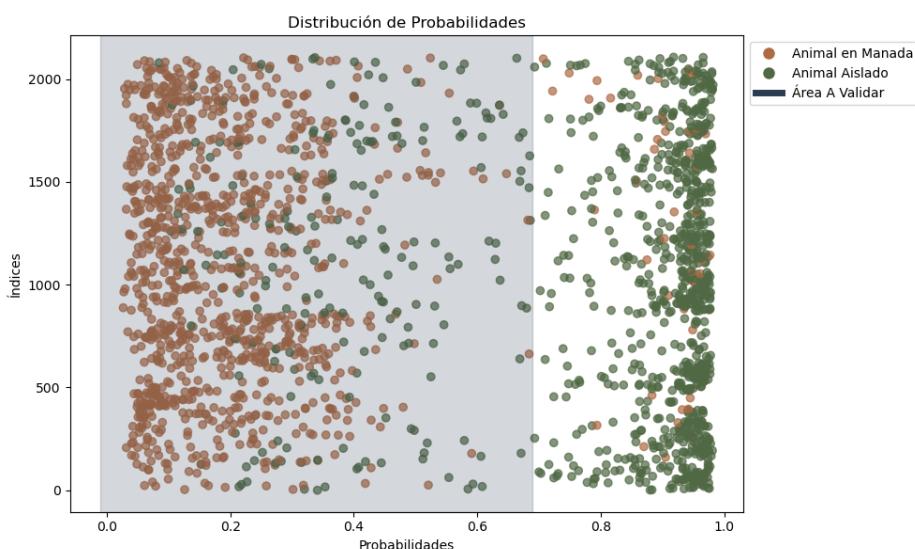


Figura 44: Distribución de Probabilidades y Área a Validar en Presencia de Animales Aislados.

Proceso Semi-Automatizado

Tras un minucioso análisis de la confiabilidad de los modelos seleccionados, se determinó la fracción de imágenes cuyas predicciones poseen un grado de certeza tal que omite de la validación humana. La precisión en la clasificación para la detección de ausencia de animales alcanza un nivel que permite excluir del proceso de validación manual al 95.94% de las imágenes que no evidencian presencia animal.

En lo concerniente a las imágenes que se clasifican con una confianza elevada de contener animales, que constituyen el 92.74% del total, se excluye la necesidad de validación para aquellas imágenes identificadas confiadamente como guanacos solitarios. Los modelos para la detección de guanacos y de animales aislados permiten omitir la validación en el 49.19% y el 41.79% de las imágenes con presencia animal respectivamente. De esta forma, se estima que el 19.06% de las imágenes con animales pueden ser clasificadas sin necesidad de validación posterior.

Los modelos de detección de guanacos y animales aislados permiten la no validación de las imágenes que presentan animales en 49.19% y 41.79% respectivamente. Por lo que, el porcentaje de imágenes que presentan animales que no se deberán validar representa el 14.14%, se calcula:

$$\text{Porcentaje de Animales No Validadas} = 92.74\% * 49.19\% * 41.79\% = 19.06\%$$

Por consiguiente, la proporción total de imágenes etiquetadas automáticamente asciende al 71.43%.

$$\text{Porcentaje de Etiquetadas Automáticamente} = 95.94\% \times 68.12\% + 19.06\% \times 31.88\% = 71.43\%$$

$$\text{Porcentaje de Validadas} = 1 = 71.43\% = 28.57\%$$



Figura 45: Representación de Imágenes a Validar

El experto deberá **validar únicamente el 28.57% de las imágenes**. Bajo este esquema, y referenciando el volumen de imágenes procesadas en el año 2023, el análisis humano se limitaría a tan solo 33,900 imágenes de un total de 118,769 revisadas en el escenario actual. Esta eficiencia se ve potenciada por la posibilidad de examinar las imágenes en secuencias rápidas, lo cual agiliza aún más el proceso. Una representación visual del conjunto de imágenes que requieren validación puede consultarse en la **Figura 45**.

Resultados y Conclusiones del Desarrollo

Tras un periodo intensivo de investigación y desarrollo en la aplicación de técnicas de Inteligencia Artificial para la conservación de la biodiversidad, se han alcanzado hallazgos significativos. La integración de algoritmos avanzados y análisis de imágenes ha permitido crear soluciones robustas para la identificación eficiente de especies. Este avance representa un paso monumental en la mejora de procesos relacionados con la biodiversidad y su preservación. Los cálculos precedentes a los resultados plasmados son detallados en el **Anexo 5**. A continuación se presentan los resultados a partir de los KPIs planteados.

Tiempo Total del Proceso de Etiquetado

Los tiempos de procesamiento para la solución propuesta de las imágenes, evaluados en una computadora personal convencional, se desagregan de la siguiente forma:

- **Tiempo de Conversión a Tensor:** Este proceso debe ser aplicado a todas las imágenes e implica 0.1 segundos por imagen.
- **Tiempo de Predicción Modelo VGG16:** Modelos de este carácter tiene dos funcionalidades: predecir la presencia de animal y predecir la presencia de guanaco. En primera instancia se aplicará a todas las imágenes tensorificadas, pero para la segunda funcionalidad se utilizará únicamente en las que hayan sido clasificadas como animal previamente. Implica 0.07 segundos por imagen.
- **Tiempo de Predicción de Megadetector:** Este modelo es utilizado para predecir los guanacos que se presentan solos en la imagen. Por lo tanto únicamente se aplicará para aquellos casos en los que se predijo que la imagen contenía un animal y que el mismo era un guanaco. Este proceso implica un tiempo de 8 segundos por imagen.

Tiempo de Conversion de Imagen a Tensor	Tiempo de Predicción del Modelo con Arquitectura VGG16	Tiempo de Predicción del Modelo Megadetector
0.1 seg / imagen	0.07 seg / imagen	8 seg / imagen

Tabla 10: Tiempos de Procesamiento Automático

Es importante señalar que este tiempo no interfiere con las labores humanas, permitiendo su ejecución en segundo plano, sin perturbar las actividades principales

de los biólogos. La información estará preparada para iniciar el proceso de validación dos días posteriores al inicio de la ejecución. Este intervalo de procesamiento equivale a 1.99 jornadas laborales continuas durante las cuales la interfaz ejecuta los modelos predictivos de manera autónoma, sin requerir intervención supervisora.

Además, se agrega el tiempo de validación manual sobre las imágenes que no obtuvieron la confianza suficiente por los modelos.

	Tiempo Total	Jornadas Laborales Ocupadas
Clasificación Manual	152.75 horas	19.1 Jornadas
Solución Semi-Automática	Automático: 7.89 hs Validación Manual: 12.87 hs Total: 20.71hs	1.6 Jornadas Ocupadas ▼ -86.44% 2.6 Jornadas Totales ▼ -86.45%

Tabla 11: Resultados de Tiempos para Situación Actual y Solución.

Como se observa en la **Tabla 11**, al humano le llevará aproximadamente 1.6 jornadas laborales validar las imágenes, esto resulta en una mejora porcentual del 91.62% con respecto a la situación anterior. A diferencia de la situación inicial, este trabajo puede ser compartido entre el personal ya que la interfaz gráfica promueve la estandarización del proceso.

El procedimiento semi-automatizado ahora se completa en tan solo 2.6 jornadas laborales, lo que significa que los resultados estarán listos en ese plazo tras iniciar la ejecución. Esta optimización ha llevado a una **mejora del 86.45% en la eficiencia del proceso global**. Si se considera únicamente el componente manual del proceso y se excluye el tiempo de ejecución automática, **el ahorro de tiempo se eleva a un 91.62%**. Estos números reflejan no sólo una operación más ágil, sino también una transformación en la eficiencia operativa.

Cantidad Promedio de Imágenes Analizadas por Hora

En el caso de la solución propuesta la cantidad de imágenes procesadas por hora se formulan de la siguiente forma:

	Cantidad de Imágenes por Hora
Clasificación Manual	777 imágenes/hr
Solución Semi-Automática	5735 imágenes/hr ▲ 638.02%

Tabla 12: Resultados de Cantidad de Imágenes por Hora para Situación Actual y Solución.

Con base en los datos proporcionados (**Tabla 12**), la tasa de procesamiento de imágenes ha experimentado una notable mejora. Originalmente, el sistema procesaba imágenes a un ritmo de 777 imágenes por hora. Sin embargo, con las optimizaciones implementadas, este número ha aumentado a 5,735 imágenes por hora. Esta elevación representa una **mejora de aproximadamente el 638.02% en la eficiencia del procesamiento.**

Precisión de Etiquetado

Previo a la implementación de la solución, la precisión de la clasificación se estimaba en un 90%, teniendo en cuenta el error humano intrínseco. Si se confiara en la solución parcialmente automatizada se lograría una precisión en el proceso del 94.6%, dado la tasa de error expuesta en la **Tabla 13**.

	Tasa de Error	Cantidad de Errores
Clasificación Manual	10%	11,877 errores
Solución Semi-Automática	5.6%  -43.97%	3,243 errores  -72.7%

Tabla 13: Resultados de Errores para la Situación Actual y la Solución.

Este incremento en la precisión, que representa casi un 44%, no solo refleja la eficacia de las mejoras aplicadas, sino que también **garantiza una confiabilidad significativamente mayor** en las predicciones y resultados generados. En términos prácticos, esto se traduce en una **reducción de errores del 72.7%**, lo que a su vez conlleva una gestión más eficiente y certera de los recursos y decisiones basadas en dichos resultados. Por ejemplo, si se implementaran ambos escenarios en el conjunto de datos del año 2023, el enfoque vigente resultaría en 11,877 incidencias erróneas, mientras que la metodología propuesta reduciría significativamente la cifra a solo 3,243 equivocaciones. Este tipo de avance reafirma el compromiso con la excelencia y la continua búsqueda de perfección en las operaciones y análisis realizados.

Costo del Proceso

El nuevo escenario presenta costos en recursos humanos, errores implícitos y, además, se le agrega un paquete anual que involucra el mantenimiento, los recursos de profesionales, comunicación periódica y capacitación del personal por \$1,000,000 (Noviembre, 2023). Además, se asigna un 10% adicional del presupuesto total como reserva para contingencias, por lo que el valor final del paquete será de \$1,100,000.

	Costo Total
Clasificación Manual	\$56.5
Solución Semi-Automática	\$33.24

Tabla 14: Resultados de Costos de la Situación Actual y Solución.

La adopción del nuevo proceso semi-automatizado conlleva una significativa optimización financiera para la organización. Se ha evidenciado una contundente reducción en los costos, pasando de un desembolso de \$56.5 a tan solo \$33.24. Esto representa un **ahorro del 41.18%**, lo que es una clara manifestación de una gestión más eficiente y una implementación tecnológica acertada. Este drástico descenso no solo refleja una notable economía en términos monetarios, sino que también subraya el potencial de innovar y adaptarse a soluciones tecnológicas de vanguardia.

Ratio de Imágenes Útiles

El porcentaje de imágenes con relevancia directa para las estimaciones de población en Rewilding es meramente el 0.79% del total de la base (**Anexo 2**). Esto implica que un vasto 99% de las imágenes almacenadas no aporta valor sustancial para los objetivos de la Fundación.

	Ratio de Imágenes Útiles
Clasificación Manual	0.79%
Solución Semi-Automática	100%

Tabla 15: Resultados de Ratios de Imágenes Útiles de Situación Actual y Solución.

Sin embargo, con la implementación del sistema de procesamiento semi-automatizado propuesto, **se reconoce el valor inherente de cada imagen** al ser almacenadas en su totalidad. Esta revalorización deriva del hecho de que las imágenes fortalecerán y refinarán los modelos predictivos, incrementando su precisión. Adicionalmente, este conjunto de datos no solo será crucial para las áreas de estudio actuales, sino que también servirá como una herramienta esencial si se opta por expandir las iniciativas de *Rewilding* a otras regiones.

Conclusiones

Este proyecto se ha enfocado en la mejora y automatización del proceso de etiquetado de imágenes de vida silvestre, utilizando tecnologías avanzadas de Inteligencia Artificial para analizar imágenes de cámaras trampa. Con un objetivo claro de incrementar la eficiencia y precisión en la estimación de la abundancia de especies, el proyecto ha abordado con éxito los retos asociados al manejo de grandes volúmenes de datos. El resultado ha sido una solución semi-automatizada que no solo mejora significativamente el proceso de clasificación de imágenes sino que también demuestra cómo la sinergia entre la inteligencia artificial y la intervención humana puede ser un catalizador potente para proyectos de conservación de la biodiversidad.

La adopción del proyecto proporcionará al cliente una interfaz visual intuitiva para la ejecución y validación de modelos de clasificación de imágenes, lo que facilita significativamente el proceso de monitoreo y análisis de datos. Esta interfaz permitiría a los usuarios interactuar de manera eficiente con el modelo, ajustar parámetros y validar los resultados de forma inmediata, lo que agiliza la toma de decisiones basada en los datos obtenidos. Además, el proyecto incluye una base de datos robusta y organizada que asegura la consistencia en la clasificación de imágenes. Esto no solo optimiza la recuperación y el manejo de la información sino que también garantiza la integridad y la trazabilidad de los datos a lo largo del tiempo, lo que resulta crucial para estudios de estimación de poblaciones y para la verificación de la calidad de los datos en investigaciones de conservación y ecología.

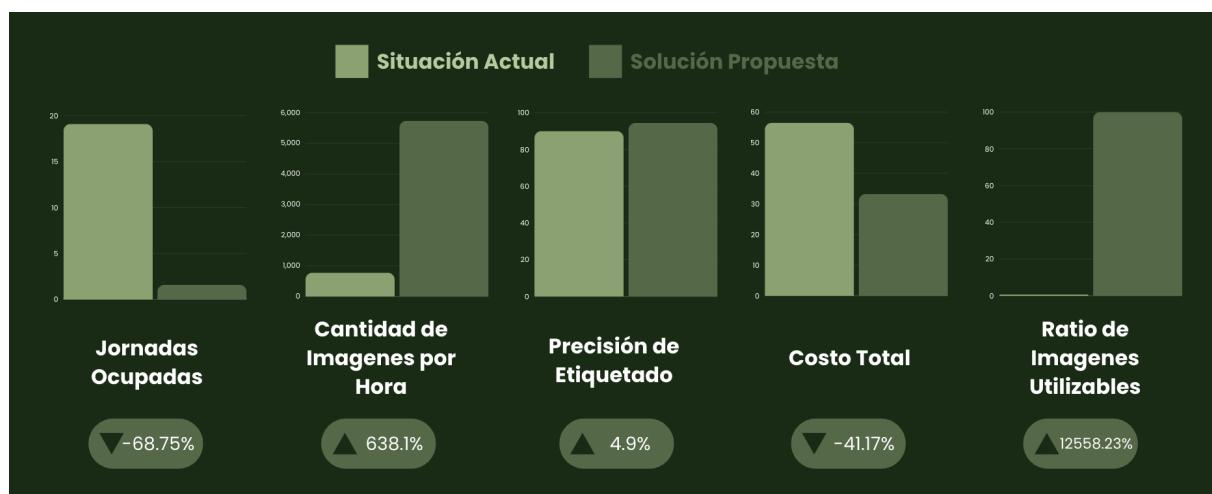


Figura 46: Variación Porcentual de KPIs

La implementación de la solución propuesta ha generado mejoras significativas en los indicadores relevantes para el proceso de clasificación automática de imágenes de animales capturadas por cámaras trampa. Se ha logrado una reducción del tiempo necesario para completar las jornadas totales en un 68.75%. En cuanto a la cantidad de imágenes procesadas por hora, se ha incrementado en un notable 638.1%.

La precisión del etiquetado, un aspecto crítico para la calidad del análisis de datos, se ha incrementado en un 4.9%. Este avance es fundamental para la precisión de estudios de biodiversidad y para el seguimiento de la fauna.

Desde una perspectiva económica, el costo total del proceso se ha reducido en un 41.17%. Finalmente, el ratio de imágenes utilizables se ha incrementado a un 12558.23%, lo que refleja un aprovechamiento completo del conjunto de datos y una minimización del descarte de imágenes.

El análisis en este proyecto destacó la importancia de estandarizar las columnas categóricas para eliminar duplicados causados por pequeñas variaciones textuales, crucial para garantizar la coherencia y precisión de los datos, lo que mejoró los análisis subsecuentes. Se enfrentó al desafío de precisar las rutas de imágenes en la base de datos. Inicialmente, solo un 14.7% de las imágenes coincidían con las rutas generadas, debido a la variabilidad en la nomenclatura y almacenamiento de archivos. Tras un minucioso proceso, se logró una corrección del 99.4% de los datos. La estandarización y corrección de las rutas de los archivos en la base de datos se traduce en una mejora significativa para el proyecto, asegurando la consistencia y precisión en el manejo de la información y minimizando errores potenciales, vital en proyectos de gran escala y complejidad.

Estas mejoras no solo reflejan la eficacia de la solución propuesta, sino que también ponen de manifiesto la importancia de adoptar tecnologías avanzadas y metodologías optimizadas en la investigación y conservación de la vida silvestre. La capacidad para procesar y etiquetar eficientemente grandes volúmenes de imágenes abre nuevas posibilidades para el monitoreo ambiental y la toma de decisiones basada en datos en tiempo real, contribuyendo significativamente al campo de la ciencia de datos aplicada a la ecología y conservación.

En el caso de que la solución fuera adoptada para el procesamiento de la información del año corriente las mejoras en términos absolutos serían las descritas en la **Tabla 16**.

	Cantidad de Jornadas Laborales	Cantidad de Errores	Costo Total
Clasificación Manual	19.1 Jornadas	11,877 errores	\$6,722,325.4
Solución Semi-Automática	3.59 Jornadas	3,113 errores	\$2,762,766

Tabla 16: Resultados Obtenidos en 2023.

Futuras Direcciones

El proyecto actual establece una plataforma sólida para múltiples direcciones de desarrollo futuro significativas:

- Escalabilidad y Expansión del Proyecto:** El diseño y metodología implementados en este proyecto facilitan su expansión a nuevas áreas geográficas y la integración de cámaras adicionales, sin que ello implique un

incremento proporcional en los costos operativos. Esta característica potencia la escalabilidad del proyecto y fomenta una mejora continua.

- **Implementación de Edge Computing:** La integración de tecnologías de Edge Computing ofrece la posibilidad de procesar imágenes en el lugar de captura, permitiendo la eliminación automática de capturas sin presencia animal. Esta estrategia es viable tras alcanzar una confianza robusta en el modelo de detección de animales, minimizando la necesidad de re-entrenamiento continuo. Es crucial evaluar el balance entre los beneficios del re-entrenamiento del modelo y la eficiencia operativa ganada mediante el uso de Edge Computing para la eliminación de imágenes.
- **Automatización de Reportes de Insights:** Basándose en los datos generados por el proceso de clasificación, se propone el desarrollo de un sistema automatizado para la generación de reportes analíticos. Este sistema podría integrarse a la interfaz de usuario ya desarrollada, incluyendo indicadores personalizados según los requisitos de *Rewilding*.
- **Evaluación Continua y Mejora de Implementaciones:** La monitorización y análisis continuo de los resultados y la experiencia de usuario proporcionados por la organización permitirá una comprensión más profunda de las futuras implementaciones. Esta evaluación constante es clave para la mejora de los supuestos iniciales del proyecto y optimizará el proceso de incorporación en nuevas regiones u organizaciones.

Estas direcciones no solo buscan mejorar la eficiencia y efectividad del proyecto actual, sino también explorar nuevas vías para potenciar su impacto en el campo de la conservación de la vida silvestre mediante la tecnología de clasificación de imágenes.

Anexo

Anexo 1: Interfaz de procesamiento y validación de datos.

Se realizaron dos interfaces de usuario. La primera permite al usuario procesar las imágenes de nuevas muestras, sobre las cuales se generan los tensores correspondientes y se corren los modelos predictivos. La segunda permite al usuario visualizar las imágenes y validar manualmente el output del modelo. Esta validación aplicará los cambios sobre la base de datos final y permitirá datos consistentes para realizar un nuevo entrenamiento a futuro.

A continuación se ofrece una descripción de cómo utilizar la interfaz de usuario para utilizar los modelos de clasificación y validar los etiquetados que sean necesarios. Los mismos, serán disponibilizados en un archivo .exe, el cual se podrá ejecutar desde la computadora personal del usuario.

Primeramente, se ofrece un botón que permite seleccionar la carpeta de imágenes correspondiente, una vez seleccionada, al presionar el botón de procesamiento de imágenes, se correrán todos los modelos necesarios. De no haber problemas, se muestra un mensaje que refiere a un proceso exitoso.

Una vez ocurrido esto, las imágenes sin animales y que no precisan validación son movidas a una carpeta para tal fin. Además, se genera un archivo de valores separados por comas con toda la metadata asociada a las imágenes, que el usuario puede nombrar a gusto.



Figura 47: Etiquetado de Especie

Al ejecutar la segunda interfaz de usuario, se disponibiliza un botón que permite cargar el .csv con las rutas y predicciones a validar, cuando se realiza esto, se muestran las imágenes correspondientes y se pide insertar la cantidad de animales, sugiriendo siempre 1 por default, y seleccionar la especie que se observa de entre las opciones que se mostrarán ordenadas por la proporción de datos que representa la especie en los análisis previos. A la demostración sobre los datos de muestreo 2023, se incorpora la función de visualización de ráfagas (visualizar simultáneamente hasta 8 imágenes tomadas en un intervalo de N segundos), con opción de selección múltiple y clasificación simultánea.

Anexo 2: Estimación de Proporción de Imágenes sin Animales en 2023.

Se estima el tiempo que llevaría etiquetar la base del año corriente basado en la cantidad promedio de imágenes con animales etiquetadas de los años anteriores para ambas zonas. Sauce presentó un total de 859, 753 y 685 animales etiquetados en los últimos tres años respectivamente. Ibérica presenta la información de los últimos dos años, siendo sus totales 223 y 125.

$$\text{Promedio de Imágenes Etiquetadas con Animales por Año} = \frac{223+125}{2} + \frac{859+753+685}{3} = 940$$

$$\text{Porcentaje de Base con Animales Etiquetados} = \frac{940}{118,769} \times 100 = 0.79\%$$

Por lo que el porcentaje de la base que no se debería etiquetar es de:

$$\text{Porcentaje de Base no Etiquetable} = 100\% - 0.79\% = 99.21\%$$

Además, como en el 2022 se almacenaron todas las imágenes que presentan animales (sin importar si están etiquetadas) se utiliza esta información para calcular el promedio de imágenes que presentan animales del total de la base. La base de imágenes del 2022 tiene un tamaño de 35,756 en Sauce y 2112 en Ibérica.

$$\text{Porcentaje de Base con Animales} = \frac{35,756 + 2,112}{118,769} \times 100 = 31.88\%$$

$$\text{Porcentaje de Base sin Animales} = 100\% - 31.88\% = 68.12\%$$

Anexo 3: Cálculo de Tasa de Error, Tiempos y Costos para escenarios Automatizados y Semiautomatizados.

Siguiendo la lógica de cálculos de los indicadores propuestos en el Business Case, se evalúan los resultados del escenario automatizado tomando los estudios como benchmark.

$$\text{Error Automatico Estimado} = (FN + E_E + E_c) \frac{n}{N} + FP \times \frac{1-n}{N}$$

$$\text{Error Automatico Estimado} = (3.8\% + 5.1\% + 36.9\%) \times 31.88\% + 2.49\% \times 68.12\% = 16.3\%$$

De manera similar se realiza el cálculo del tiempo promedio que genera cada imagen al clasificarla:

$$\text{Tiempo por Imagen} = \frac{3600}{6480} = 0.56 \text{ segundos}$$

Adicionalmente, es posible calcular el costo implícito asociado con la clasificación de cada imagen en el proceso vigente. Este cálculo permitirá, posteriormente, proyectar el gasto total correspondiente a cada campaña de captura de imágenes.

$$\text{Costo Error por Imagen} = \frac{1-n}{N} \times FP \times C_{FP} + \frac{n}{N} \times (FN \times C_{FN} + E_E \times C_E + E_c \times C_c)$$

$$\begin{aligned} \text{Costo Error por Imagen} &= 68.12\% \times 2.49\% \times \$1000 \\ &+ 31.88\% \times (3.8\% \times \$300 + 5.1\% \times \$1300 + 36.9\% \times \$650) = \$118.2 \end{aligned}$$

Al proyectar estos costos al volumen total de imágenes presentes en la base de datos, se obtiene una cifra representativa del costo total acumulado. Se le agrega el costo del paquete de servicio.

$$\text{Costo Total Automático} = 118,769 \times \$118.2 + \$1100000 = \$13,727,936.25$$

$$\text{Costo Total por Imagen} = \frac{\$13,727,936.25}{118,769} = \$115.59$$

Para el caso Semi-Automatizado, se toman nuevamente los benchmarks planteados por estudios similares, agregando el paso extra de validación manual.

Al no poder aplicar la heurística de optimización de umbrales (**Anexo 5**) sobre los modelos benchmark, se supone que se validará el 90% de las imágenes mal etiquetadas. Se procede al cálculo:

$$\text{Error Semi-Automático Estimado} = 16.3\% \times 0.1 + 5\% \times 0.9 = 3.43\%$$

$$\text{Cantidad de Segundos por Imagen} = 1.84 \text{ segundos}$$

$$\text{Costo Total Semi-Automático} = 118,769 \times (\$9.6 + \$3.2 + \$0.04) + \$1100000 = \$2,625,498.29$$

$$\text{Costo Total por Imagen} = \frac{\$2,625,498.29}{118,769} = \$22.69$$

Anexo 4: Heurística de Optimización de Umbrales

La heurística implementada para definir el área de validación opera bajo un esquema optimizado que equilibra precisión y coste. La función optimizar_umbral (**Figura 48**) ajusta los límites de confianza para determinar que imágenes deben ser validadas manualmente y cuáles pueden ser clasificadas automáticamente. El objetivo es minimizar la suma del coste de validación manual y el coste asociado a errores de clasificación.

En la práctica, se establecen umbrales inferior y superior dentro de los cuales las predicciones del modelo se consideran inciertas y por ende, se someten a revisión humana. Las predicciones que caen fuera de estos umbrales se aceptan como están, asumiendo que el modelo es suficientemente confiable en esos rangos.

El algoritmo calcula los costes de falsos positivos, falsos negativos y la validación manual, encontrando el rango de umbrales que resulta en el menor coste total posible.

```

1 # Funcion que busca optimizar el costo para identificar los umbrales de validacion.
2 def optimizar_umbral(y_test, y_proba, tarea, CFP = 1000, CFN = 300, CV = 20, TV = 0.579):
3     # Inicializacion de variables
4     limInf = 0
5     limSup = 1
6     CostoMin = float('inf')
7     Vop = 0
8     # Determinacion de costos.
9     if tarea == 'especie':
10         CE = CFP + CFN
11     else:
12         CE = (CFP + CFN) / 2
13     # Ciclo que minimizando el costo.
14     # Balancea el costo generado por errores de clasificacion con el costo de validacion.
15     for umbral_inf in np.arange(0.01, 0.5, 0.01):
16         for umbral_sup in np.arange(0.5, 0.99, 0.01):
17             # Determina la cantidad de imagenes que caen entre los umbrales.
18             V = np.sum((y_proba >= umbral_inf) & (y_proba <= umbral_sup))
19             # Costo generado por la validacion.
20             C = V * CV
21             # Costo de Error de Validacion
22             C += V * TV * CE
23             # Determina el error que conlleva no validar fuera de los umbrales.
24             if tarea == 'presencia':
25                 # Errores por Falsos Positivos fuera de los umbrales.
26                 FP = np.sum((y_proba < umbral_inf) | (y_proba > umbral_sup) & (y_test == 0))
27                 # Errores por Falsos Negativos fuera de los umbrales.
28                 FN = np.sum((y_proba < umbral_inf) | (y_proba > umbral_sup) & (y_test == 1))
29                 # Costos generados por errores de clasificacion.
30                 C += FP * CFP + FN * CFN
31             else:
32                 # Errores generados.
33                 E = np.sum((y_proba < umbral_inf) | (y_proba > umbral_sup))
34                 # Costo generado por errores de clasificacion.
35                 C += E * CE
36             # En el caso de que el costo sea menor al minimo actual, se actualiza el minimo.
37             if C < CostoMin:
38                 Vop = V
39                 CostoMin = C
40                 limInf = umbral_inf
41                 limSup = umbral_sup
42
43 print(f"Mejor rango de umbral: {round(limInf,2)} - {round(limSup,2)} con un costo de: ${CostoMin}")
44 print(f"Validara {Vop} imagenes, correspondientes al {round(Vop * 100 / len(y_test),2)}% del total de las imagenes.")
45
46 return limInf, limSup

```

Figura 48: Código de Heurística de Optimización de Umbrales

Anexo 5: Cálculo de KPIs para Solución Propuesta Semi-Automática

Por último para la implementación de nuestra solución, se sigue la lógica Semi-Automatizada utilizando los tiempos y capacidades de predicción de nuestros modelos.

Tiempo Total del Proceso:

Si la base de datos correspondiente al año 2023 se procesa de acuerdo con el protocolo vigente en la situación actual, los tiempos de procesamiento se establecerán de la siguiente manera:

$$Tiempo Ejecucion Manual_{2023} = \frac{118,769 \times 4.63 \text{ seg}}{60 \times 60} = 152,75 \text{ hs}$$

$$Jornadas Ejecucion Manual_{2023} = \frac{152,75 \text{ hs}}{8 \text{ hs}} = 19.1 \text{ Jornadas}$$

Se presentan las variables pertinentes y el cálculo de los tiempos implicados en la **Tabla 17**.

	Cantidad de Imágenes	Proporción de Presencia Animal	Porcentaje de Confianza en la Predicción del Modelo
Clasificación de Presencia Animal			Animales: 92.74% No Animales: 95.94%
Clasificación de Especies	118,769	31.88%	Guanacos: 49.19%
Clasificación de Cantidad			Aislados: 41.79%
Tiempo de Conversión de Imagen a Tensor		Tiempo de Predicción del Modelo con Arquitectura VGG16	Tiempo de Predicción del Modelo Megadetector
0.1 seg / imagen		0.07 seg / imagen	8 seg / imagen

Tabla 17: Variables para Cálculo de Tiempo en Escenario Automático

El tiempo que conlleva procesar las imágenes por los modelos de inteligencia artificial se calcula de la siguiente forma:

$$\text{Tiempo Automática por Imagen} = 0.1 s + 0.07 s + 31.88\% \times (0.07 s + 49.19\% \times 8 s) = 1.45 \text{ seg}$$

$$\text{Tiempo Ejecución Automática}_{2023} = \frac{118,769 \times 1.45 \text{ seg}}{60 \times 60} = 47.84 \text{ hs}$$

$$\text{Días Ejecución Automática}_{2023} = \frac{47.84 \text{ hs}}{24 \text{ hs}} = 2 \text{ días}$$

El tiempo requerido por la validación se calcula, sabiendo que se deben validar el 28.57% de las imágenes con un tiempo por imagen de 1.36 segundos (**Anexo 3**):

$$\text{Tiempo Validación por Imagen} = 28.57\% \times 1.36 \text{ seg} = 0.39 \text{ seg}$$

$$\text{Tiempo de Validación}_{2023} = \frac{118,769 \times 0.39 \text{ seg}}{60 \times 60} = 12.87 \text{ hs}$$

$$\text{Jornadas de Validación}_{2023} = \frac{12.87 \text{ hs}}{8 \text{ hs}} = 1.6 \text{ Jornadas}$$

El tiempo total que conlleva el proceso de clasificación de imágenes en horas es de:

$$\text{Tiempo Total Solución Semi-Automática}_{2023} = 47.84 \text{ hs} + 12.87 \text{ hs} = 60.71 \text{ hs}$$

$$\text{Jornadas Total Solución Semi-Automática}_{2023} = \frac{60.71 \text{ hs}}{8 \text{ hs}} = 7.59 \text{ Jornadas}$$

Cantidad de Imágenes por Hora:

Para calcular la cantidad de imágenes que se procesaran por hora se tiene en cuenta tanto el tiempo que tarda en procesarlas los modelos predictivos como el tiempo que se tarda en validar aquellas que necesiten un refuerzo en su clasificación.

$$Cantidad\ Imagenes\ por\ Hora\ Manual_{2023} = \frac{60 \times 60}{4.63\ seg} = 777\ imagenes/hr$$

$$Cantidad\ Imagenes\ por\ Hora\ Solución\ Semi-Automática_{2023} = \frac{60 \times 60}{1.45+0.39} = 1957\ imagenes/hr$$

Precisión de Etiquetado:

La cantidad de imágenes que se clasificaran mal en caso de aplicar la metodología actual está dada por:

$$Cantidad\ de\ Errores\ Manual_{2023} = 118,769 \times 10\% = 11,877\ errores$$

Para el caso de aplicar la solución semi automatizada, la tasa de error implícita en el proceso se explica según los datos en la **Tabla 18**.

	Porcentaje de Confianza en la Predicción del Modelo	Porcentaje de Error Asumido Bajo esa Confianza
Clasificación de Presencia Animal	Animales: 92.74% No Animales: 95.94%	Animales: 2.38% No Animales: 1.45%
Clasificación de Especies	Guanacos: 49.19%	Guanacos: 1.41%
Clasificación de Cantidad	Aislados: 41.79%	Aislados: 5.22%

Tabla 18: Variables para Cálculo de Tasa de Error Asumido en Escenario Automático.

La precisión en la predicción de ausencia de individuos se determina calculando el producto de la proporción de imágenes que no contienen animales en la base de datos, el porcentaje de estas predicciones realizadas con alta confianza y la tasa de error asociada a dichas predicciones.

$$Error\ Pred\ No\ Animal = 68.12\% \times 95.94\% \times 1.45\% = 0.95\%$$

La exactitud en la detección de la presencia de especímenes se establece mediante una fórmula compuesta que incorpora diversos factores: inicialmente se toma la fracción de imágenes que efectivamente exhiben animales dentro del conjunto total de datos. A este valor se le aplica el porcentaje de dichas imágenes que son clasificadas con un alto grado de certeza como guanacos. Este producto se ajusta posteriormente por la tasa de error vinculada a la identificación de guanacos, así como por el resultado de la interacción entre el porcentaje de aciertos con elevada confianza en la identificación de animales en solitario y la tasa de error correspondiente a tal categorización.

$$Error\ Pred\ Animal = 31.88\% \times 49.19\% \times (1.41\% + 41.79\% \times 5.22\%) = 0.56\%$$

Es importante aclarar que se asume que todas las predicciones erráticas del modelo de presencia animal no forman parte de este error ya que el modelo de predicción de cantidades no encontrará animales o el modelo de predicción de guanacos no encontrará un guanaco, por lo que estas imágenes serán validadas por el humano.

$$Error Automatico Asumido = 0.95\% + 0.56\% = 1.51\%$$

$$Cantidad de Errores Automatico_{2023} = 118,769 \times 1.51\% = 1793 errores$$

Ahora, las imágenes que se les asigna validación tienen una tasa de error dada por este proceso. Se asume que las tasas de error en el proceso de validación son 50% menores que en la situación actual.

$$Proporción de Imágenes a Validar (AV) = 28.57\%$$

$$Tasa Error Validación = 31.88\% \times \left(\frac{16.7\%+3.4\%+10\%}{2} \right) \times 80.94\% + 68.12\% \times \frac{0.6\%}{2} \times 4.06\% = 3.89\%$$

$$Cantidad de Errores Validación_{2023} = 118,769 \times 28.57\% \times 3.89\% = 1320 errores$$

$$Error de Solución Semi-Automatica_{2023} = 1.51\% + 3.89\% = 5.4\%$$

$$Cantidad de Errores Solución Semi-Automatica_{2023} = 1793 + 1320 = 3113$$

Costos del Proceso:

Para calcular los costos del proceso se utilizará los errores detallados en el paso anterior y los costos definidos previamente.

$$Costo Error Pred No Animal = 0.95\% \times \$300 = \$2.85$$

$$Costo Error Pred Animal = 31.88\% \times 49.19\% \times (1.41\% \times \$1300 + 41.79\% \times 5.22\% \times \$650) = \$5.10$$

$$Costo Error Automatico Asumido = \$2.85 + \$5.10 = \$7.95$$

$$Costo Error Validación = 31.88\% \times 80.94\% \times \frac{16.7\% \times \$300 + 3.4\% \times \$1300 + 10\% \times \$650}{2}$$

$$+ 68.12\% \times 4.06\% \times \frac{0.6\%}{2} \times \$1000 = \$24.27$$

$$Costo Operacional por Validación = 28.57\% \times 1.36 seg \times \frac{\$10,730}{8 \times 60 \times 60} = \$0.14$$

$$Costo de Servicio = 1,100,000$$

$$Costo Total = 118,769 \times (\$7.95 + \$24.27 + \$0.14) + 1,100,000 = \$4,944,552.53$$

Anexo 6: Repositorio de GitHub

Para consultar el código fuente, recursos y documentación asociada al proyecto, se puede acceder al repositorio oficial en GitHub a través del siguiente enlace:

[RewildingPA](#)

Es recomendable seguir las instrucciones detalladas en el archivo README.md dentro del repositorio para obtener una comprensión adecuada sobre la estructura del proyecto y los pasos para su implementación.

Anexo 7: Presentación

Se adjunta el link a la presentación final del proyecto realizada para la defensa de este. En ella se hace un resumen visual del proyecto.

[Presentación](#)

Bibliografía

Fundación Rewilding Argentina. (2023, marzo 28). Fundación Rewilding Argentina. <https://www.rewindingargentina.org/>

Morris, D. (n.d.). *MegaDetector: MegaDetector is an AI model that helps conservation folks spend less time doing boring things with camera trap images.* <https://github.com/agentmorris/MegaDetector>

IWildCam 2020 - FGVC7. (n.d.). Kaggle.com. Retrieved September 11, 2023, from <https://www.kaggle.com/competitions/iwildcam-2020-fgvc7>

Technology: Wildlife insights - Google. (n.d.). Sustainability. Retrieved September 20, 2023, from <https://sustainability.google/operating-sustainably/stories/wildlife-insights/>

Fennell, M., Beirne, C., & Burton, A. C. (2022). Use of object detection in camera trap image identification: Assessing a method to rapidly and accurately classify human and animal detections for research and application in recreation ecology. *Global Ecology and Conservation*, 35(e02104), e02104. <https://doi.org/10.1016/j.gecco.2022.e02104>

Norouzzadeh, M. S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M. S., Packer, C., & Clune, J. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences of the United States of America*, 115(25). <https://doi.org/10.1073/pnas.1719367115>

Johansson, Ö., Samelius, G., Wikberg, E., Chapron, G., Mishra, C., & Low, M. (2020). Identification errors in camera-trap studies result in systematic population overestimation. *Scientific Reports*, 10(1), 6393.

<https://doi.org/10.1038/s41598-020-63367-z>

Chesson, M. (2012). *An investigation of human-error rates in wildlife photographic identification; Implications for the use of citizen scientists.* University of Massachusetts Amherst.

<https://doi.org/10.7275/2386185>