

**UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ**

**Colegio de Ciencias e Ingeniería**

**Detección de Leones Marinos en las Islas Galápagos: una  
Comparación entre los Modelos de Aprendizaje Profundo  
YOLO-NAS y YOLOv8**

**Kevin Angel Gil Bazán**

**Angelo Abel Gil Bazán**

**Ingeniería Electrónica y Automatización**

Trabajo de fin de carrera presentado como requisito  
para la obtención del título de  
Ingeniero Eléctronico

Quito, 25 de Julio de 2023

# **UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ**

**Colegio de Ciencias e Ingeniería**

## **HOJA DE CALIFICACIÓN DE TRABAJO DE FIN DE CARRERA**

**Detección de Leones Marinos en las Islas Galápagos: una  
Comparación entre los Modelos de Aprendizaje Profundo  
YOLO-NAS y YOLOv8**

**Kevin Angel Gil Bazán**

**Angelo Abel Gil Bazán**

**Nombre del profesor, Título académico**

**Diego Benitez, Ph.D.,**

Quito, 25 de Julio de 2023

**© DERECHOS DE AUTOR**

Por medio del presente documento certifico que he leído todas las Políticas y Manuales de la Universidad San Francisco de Quito USFQ, incluyendo la Política de Propiedad Intelectual USFQ, y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo quedan sujetos a lo dispuesto en esas Políticas.

Asimismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo en el repositorio virtual, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior del Ecuador.

Nombres y apellidos: Kevin Angel Gil Bazán  
Código: 00213404  
Cédula de identidad: 2000082772  
Lugar y fecha: Quito, 25 de Julio de 2023

Nombres y apellidos: Angelo Abel Gil Bazán  
Código: 00213402  
Cédula de identidad: 2000082764  
Lugar y fecha: Quito, 25 de Julio de 2023

## **ACLARACIÓN PARA PUBLICACIÓN**

**Nota:** El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en <http://bit.ly/COPETHeses>.

## **UNPUBLISHED DOCUMENT**

**Note:** The following capstone project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on <http://bit.ly/COPETHeses>.

## RESUMEN

Los leones o lobos marinos (*Zalophus Wollebaeki*) son una especie protegida, y el monitoreo efectivo es crucial para la preservación del hábitat y los estudios de comportamiento. Sin embargo, el conteo manual de leones marinos es laborioso y propenso a errores. En este artículo, exploramos el uso de dos modelos de redes neuronales convolucionales estándar (YOLO-NAS y YOLOv8) para la detección de leones marinos como un paso preliminar hacia la automatización del proceso de conteo. Para ello se recolectó un dataset de imágenes y videos de lobos marinos en su ambiente natural en las Islas Galápagos. Los resultados demuestran que ambos modelos exhiben capacidades de detección prometedoras, identificando con éxito a casi todos los leones marinos en las imágenes. En particular, YOLOv8 demuestra ser más confiable para detectar leones marinos en condiciones desafiantes y complejas, mientras que YOLO-NAS se destaca en la identificación de una mayor cantidad de individuos, incluidos los de menor tamaño. Estos hallazgos allanan el camino para futuros desarrollos en herramientas automatizadas de conteo de leones marinos, simplificando los esfuerzos de conservación y avanzando en nuestra comprensión de esta especie protegida.

Palabras clave: Deep learning, detección lobos marinos, YOLO-NAS, YOLOv8

## ABSTRACT

Sea lions (*Zalophus Wollebaeki*) are a protected species, and effective monitoring is crucial for habitat preservation and behavioral studies. However, manual sea lion counting is laborious and error prone. In this paper, we explore the use of two standard convolutional neural network models (YOLO-NAS and YOLOv8) for sea lion detection as a preliminary step towards automating the counting process. For this purpose, a dataset of images and videos of sea lions was collected in their natural environment in the Galapagos Islands. The results demonstrate that both models exhibit promising detection capabilities, successfully identifying almost all sea lions in the images. YOLOv8 proves to be more reliable in detecting sea lions under challenging and complex conditions, while YOLO-NAS excels in identifying a larger number of individuals, including those of smaller size. These findings pave the way for future developments in automated sea lion counting tools, streamlining conservation efforts, and advancing our understanding of this protected species.

Keywords: Deep learning, sea lion detection, YOLO-NAS, YOLOv8.

## **AGRADECIMIENTOS**

Los autores agradecen al Tnlgo. Carlos Chauca del GAD Municipal de San Cristóbal su ayuda al facilitar los vídeos de leones marinos obtenidos con drones y que fueron utilizados para extraer imágenes para este estudio.

## TABLA DE CONTENIDO

Introducción .....	11
Materiales y métodos .....	14
Arquitectura YOLO-NAS.....	14
YOLO-V8 Architecture .....	15
Base de datos.....	16
Configuración experimental.....	17
Preprocesamiento de vídeo y creación de conjunto de datos.....	17
Entrenamiento, validación y particiones de prueba .....	18
Anotación de datos.....	19
Configuración de modelos .....	19
Transfer learning.....	19
Métricas de evaluación .....	21
Conteo vía detección.....	22
Resultados y discusión.....	22
Evaluación y rendimiento .....	23
Conclusiones y trabajo futuro .....	28
Referencias bibliográficas.....	29
Anexo A: Resultados detección del lobo marino.....	33
Anexo B: Resultados detección del lobo marino en entornos desafiantes.....	37



**ÍNDICE DE TABLAS**

Tabla 1. Comparison of evaluation metrics YOLO NAS vs YOLOv8. ....22

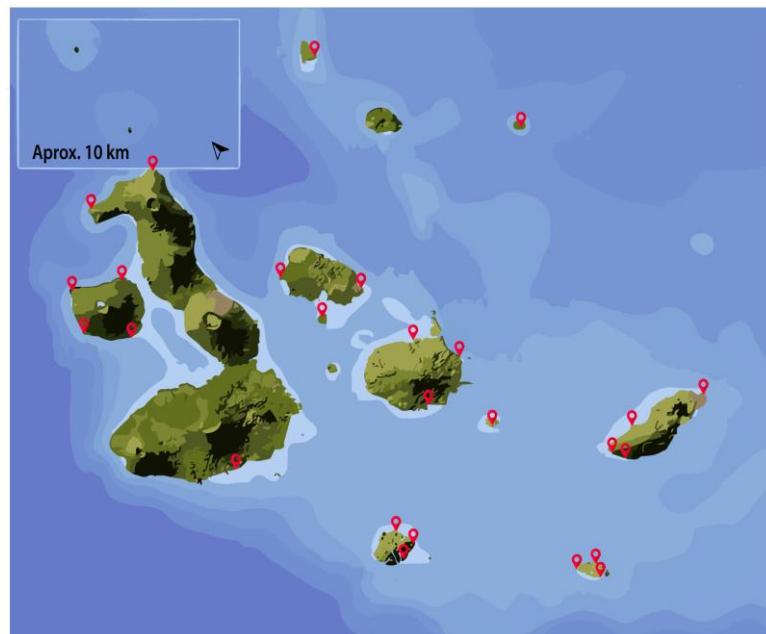
## ÍNDICE DE FIGURAS

Figura 1. Principales colonias de cría de lobos marinos de Galápagos. ....	11
Figura 2. Arquitectura para ambos modelos YOLO (NAS y v8) .....	16
Figura 3. mAP para ambos modelos YOLO (NAS y v8) .....	23
Figura 4. Rendimiento de recuperación para ambos modelos YOLO (NAS y v8) .....	24
Figura 5. Rendimiento de CLS Loss para ambos modelos YOLO (NAS y v8). ....	25
Figura 6. Rendimiento DFL Loss para ambos modelos YOLO (NAS y v8).....	26

## INTRODUCCIÓN

El océano es un ecosistema fascinante y vital para nuestro planeta, vasto y hogar de una gran diversidad de formas de vida marina, muchas de las cuales enfrentan desafíos constantes para su supervivencia. Entre los tesoros de este mundo acuático se encuentran las Islas Galápagos, un archipiélago único en el Pacífico Oriental y el mundo. Estas islas volcánicas son el hogar del lobo marino de Galápagos, una especie emblemática que ha adaptado su vida en el océano y en las playas rocosas de las islas.

Los lobos marinos de Galápagos pertenecen al género *Zalophus* y la especie *Wollebaeki* (Villegas-Amtmann et al., 2008). Estos carismáticos mamíferos han sido estudiados durante décadas por su comportamiento social, además de ser cruciales en los ecosistemas marinos de Galápagos. Estos son depredadores de peces, calamares y otros animales que habitan el mar, convirtiéndose en controladores de las poblaciones de estas especies. Se encuentran distribuidos por todo el archipiélago de las Islas Galápagos como se observa en la *Figura 1*.



*Figura 1. Principales colonias de cría de lobos marinos de Galápagos. Adaptado de (Paez-Guevara, 2017).*

La población del lobo marino ha disminuido drásticamente en los últimos 30 años (Riofrío-Lazo et al., 2017), por lo que, actualmente está considerada en peligro de extinción por la

Unión Internacional para la Conservación de la Naturaleza (UICN) (Páez-Rosas et al., 2020). Son diversos los factores que han llevado a esta especie a estar amenazada, eventos oceanográficos fuertes como el fenómeno de "El Niño" pueden llegar a reducir cerca del 50% de su población, y la recuperación de esta puede tardar hasta 7 años (Denkinger et al., 2015). La mortalidad de las crías asciende hasta el 100% durante estos fenómenos ambientales, debido a la reducción del suministro de alimento en las islas (Denkinger et al., 2017).

La contaminación, la pérdida de hábitat y sobre-explotación afectan las zonas costeras del mundo (Páez-Rosas et al., 2020), y actualmente el lobo marino enfrenta otro grave problema, enfermedades que atraviesan las barreras de especies de mamíferos, al transmitirse de especies domésticas a especies silvestres. El moquillo canino (CDV), es transmitido por un virus altamente contagioso y representa una serie amenaza para el lobo marino de Galápagos (Patricio Vega-Mariño et al., 2023). Además, existen otras enfermedades como la *Dirofilaria immitis*, una enfermedad infecciosa provocada por un parásito relacionado con animales domésticos (Gregory et al., 2023).

Por lo tanto, el seguimiento poblacional de esta especie es de suma importancia para su conservación. Desde 2015 se realizan viajes de monitoreo para evaluar el estado de las condiciones ambientales, así como las poblaciones y su relación con los sitios que habitan (Riofrío-Lazo et al., 2017).

Los biólogos y científicos marinos han estado monitoreando las poblaciones de lobos marinos utilizando métodos tradicionales, manuales y costosos, ya que requieren mucho tiempo y recursos (Chirag Padubidri et al., 2021). Además, cabe señalar que muchas veces el hábitat de estos animales se encuentra en lugares de difícil acceso (Paez-Guevara, 2017). Sin embargo, los avances en Deep Learning e inteligencia artificial permiten utilizar métodos y tecnologías para automatizar y agilizar el proceso de reconocimiento y conteo de lobos marinos, reduciendo así el margen de error presente en los procesos manuales (Christin et al., 2019). Por ejemplo, en (Chirag

Padubidri et al., 2021), se usaron dos modelos de aprendizaje profundo, un UNet básico sin ningún extractor de características y otro con el extractor de características EfficientNet-B5, junto con fotografías aéreas y mapas de densidad, para contar la población de leones marinos de Steller en las islas Aleutianas occidentales en Alaska, logrando un error cuadrático medio (RMSE) de 1,88.

A pesar de eso, el desarrollo de herramientas para el conteo automático de lobos marinos en Galápagos aún no se ha llevado a cabo y no ha recibido la atención que merece. Además, se deben desarrollar y adaptar algoritmos para el reconocimiento de las especies locales de lobos marinos.

En las últimas décadas, el rápido desarrollo de tecnologías basadas en hardware ha abierto nuevas posibilidades para que los científicos recopilen datos multimodales en varios dominios de aplicación, como ómica, bioimagen, imágenes médicas e interfaz (cerebro/cuerpo)-máquina (Mahmud et al., 2018). La biología y la medicina son campos con grandes cantidades de datos, pero estos son complejos y, a menudo, malinterpretados o mal entendidos, por lo que el aprendizaje profundo es una técnica apropiada para resolver problemas en estos campos (Ching et al., 2018).

En este proyecto nos enfocamos en el desarrollo de un sistema de reconocimiento y conteo de lobos marinos basado en técnicas de Deep Learning, específicamente usando redes neuronales convolucionales (CNN), debido a su gran potencial para la detección y clasificación de objetos, convirtiéndose en una herramienta prometedora para esta tarea (Li et al., 2023).

Para cumplir con los objetivos propuestos se han implementado dos métodos estándar YOLO-NAS y YOLO-V8, que son nuevas arquitecturas basadas en YOLO lanzadas en 2023. YOLO-NAS ofrece capacidades superiores en términos de mAP (alta precisión promedio) y latencia de inferencia, en relación con los modelos YOLOv7 y YOLOv8, además del

recientemente lanzado YOLOv6-v3.0 (Deci, 2023). Para observar diferencias, se ha propuesto comparar ambas arquitecturas en igualdad de parámetros y condiciones.

La aplicación de estas tecnologías y técnicas tienen un potencial revolucionario para la conservación del medio ambiente y de especies amenazadas (Lamba et al., 2019). Con el Deep Learning es posible monitorear las poblaciones de lobos marinos, brindando a investigadores, científicos y guardaparques herramientas eficientes, precisas y escalables para obtener información sobre la conservación y salud del medio marino de esta especie. Este trabajo puede presentar la base para futuras aplicaciones de Deep Learning en la conservación de otras especies silvestres en hábitad marinos y también terrestres, contribuyendo a la protección y preservación de especies únicas en el mundo como las existentes en las Islas Galápagos.

## **MATERIALES Y MÉTODOS**

### ***Arquitectura YOLO-NAS.***

YOLO-NAS, lanzado en mayo de 2023, un modelo de detección de objetos de una etapa que utiliza el nuevo algoritmo de búsqueda neuronal (NAS), implementa la arquitectura YOLO en su núcleo y se basa en la versión más reciente YOLOV8, donde la arquitectura se ha mejorado significativamente y ofrece mayor precisión, rendimiento y velocidad de detección.

Estas son algunas de las características clave de YOLO-NAS:

- Alta precisión: se ha demostrado que YOLO-NAS logra resultados de vanguardia en una variedad de puntos de referencia de detección de objetos.
- Inferencia en tiempo real: YOLO-NAS puede detectar objetos en tiempo real, lo que lo hace adecuado para una variedad de aplicaciones.
- Etapa única: YOLO-NAS es un modelo de etapa única, lo que significa que puede detectar objetos en un solo paso a través de la imagen. Esto lo hace más rápido y eficiente que los modelos de dos etapas.

- Basado en NAS: YOLO-NAS está basado en NAS, lo que significa que se ha optimizado para la detección de objetos, lo que da como resultado un modelo que es más preciso y eficiente que los modelos tradicionales.

YOLO-NAS, es un modelo de detección de objetos potente y versátil que se puede usar para varias aplicaciones, es rápido, preciso y fácil de usar (Deci, 2023).

### ***YOLO-V8 Architecture***

YOLO-V8, lanzado en enero del 2023, es el octavo modelo en la serie YOLO orientado a la detección de objetos, clasificación de imágenes y segmentación de instancias de vanguardia en imágenes y vídeos en tiempo real (Benali Amjoud & Mustapha Amrouch, 2023). Este modelo es una versión que presenta una serie de mejoras comparados con modelos anteriores como mayor rendimiento en términos de velocidad y precisión (Jocher-Chaurasia & Qui, 2023).

YOLO-V8 propone un enfoque que se basa en la detección de una sola pasada, lo que quiere decir que la detección y segmentación de objetos lo realiza en una sola operación, lo que lo convierte en un modelo más rápido convirtiéndose en una de sus principales ventajas, lo que lo hace ideal para aplicaciones en tiempo real. Otra ventaja es su alta precisión en la detección de objetos (Hussain, 2023).

YOLO-V8 presenta una CNN preentrenada en un masivo conjunto de datos que incluyen imágenes y etiquetas de personas, animales, vehículos y otros objetos. Para la detección, divide la imagen en celdas para predecir la probabilidad de que exista un objeto en cada celda, cuando esta es alta YOLO-V8 es capaz de predecir la posición, el tamaño y el tipo de objeto.

Las principales características son:

- Arquitectura avanzada de backbone y neck: YOLOv8 adopta una arquitectura de última generación de backbone y neck, obteniendo un mejor rendimiento de extracción de características y detección de objetos.

- Head Ultralytics dividido sin anclaje: YOLOv8 cuenta con un head Ultralytics dividido sin anclaje para mejorar la precisión, además, de un proceso de detección más eficiente en comparación con los métodos que se basan en anclaje.
- Varios modelos preentrenados: YOLOv8 proporciona una serie de modelos preentrenados para cumplir con diversas tareas y requisitos de rendimiento, lo que le facilita encontrar el modelo adecuado para su caso de uso específico (Jocher-Chaurasia & Qui, 2023).

YOLO-V8 es un potente modelo para la detección de objetos por su gran velocidad, precisión y flexibilidad, lo que le permite ser utilizado en distintas aplicaciones.

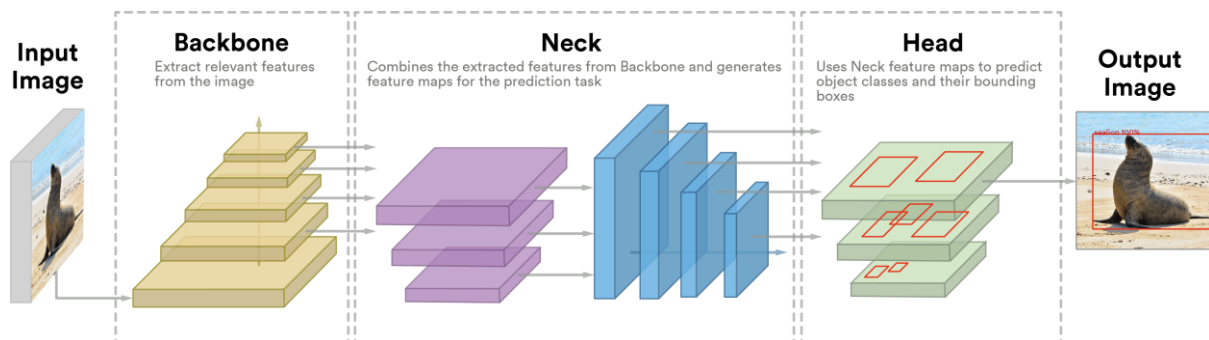


Figura 2. La arquitectura de las dos versiones de Yolo se compone de tres partes principales que incluyen la columna vertebral,

### **Base de datos**

La base de datos utilizada en este proyecto es un conjunto de imágenes que contienen lobos marinos del género *Zalophus* y especie *Wollebaeki* con sus características distintivas, estas fueron obtenidas a través de fotografías y videos tomados en playas de San Cristóbal como Playa Mann, Lobería, Punta Carola, Puerto Chino, Ochoa, Isla Lobos, entre otros. Dado que el lobo marino se encuentra distribuido por todo el archipiélago de Galápagos, incluso en lugares de difícil acceso, se ha contado con el apoyo de guías naturalistas y personas dedicadas al turismo para ampliar la base de datos con los lobos marinos que habitan islas como Floreana, Santa Cruz, Isabela e islas deshabitadas como Española. Así mismo, se contó con el apoyo de las autoridades del GAD Municipal de San Cristóbal para obtener videos a través de drones de diferentes playas



como las que se encuentran en Bahía Naufragio, que es donde se concentra la mayor colonia de lobos marinos de esta isla. Como resultado se obtuvieron aproximadamente 3.000 imágenes de lobos marinos capturadas en diferentes ubicaciones y momentos del día, con distintos niveles de iluminación y desde variados ángulos de visión con el propósito de asegurar la diversidad del conjunto de datos.

### ***Configuración experimental***

#### ***Preprocesamiento de vídeo y creación de conjunto de datos***

En este paso, se recopilaron imágenes de lobos marinos a partir de fuentes de vídeo. Para extraer los fotogramas de los vídeos, se utilizó la biblioteca OpenCV para decodificar y obtener todas las imágenes. Cada vídeo tiene una duración de 30 segundos aproximadamente y 60 fps. En el proceso de extracción de fotogramas de un vídeo, se aplicó una estrategia de muestreo para seleccionar una imagen cada 10 fotogramas. Esta decisión se basó en la observación de que existían fotogramas borrosos y fotogramas que se repetían consecutivamente sin cambios significativos en la escena.

La presencia de fotogramas borrosos puede deberse a varios factores, como movimiento rápido en la escena, baja calidad de la fuente de vídeo o problemas en la captura. Estos fotogramas borrosos pueden afectar negativamente la calidad y la utilidad de los datos.

De igual manera, la presencia de fotogramas repetidos consecutivamente sin cambios significativos no aporta información adicional y puede redundar en una sobre representación de ciertos momentos de la secuencia de vídeo.

Esta estrategia de muestreo garantizó que los fotogramas seleccionados fueran más representativos y capturaran cambios significativos en la secuencia de vídeo, a la vez que se mejoraba la calidad general del conjunto de datos de los fotogramas extraídos.

Después de este proceso, se recopilaron 1000 imágenes válidas en formato JPG. Dado que este número no era suficiente para entrenar los modelos de aprendizaje profundo sin incurrir en sobre ajuste, se aplicó una técnica de aumento de datos.

### ***Entrenamiento, validación y particiones de prueba***

Se utilizó la técnica de partición estratégica de conjuntos de datos para el entrenamiento y evaluación de los modelos YoloNas y YoloV8. Las imágenes se distribuyeron aleatoriamente utilizando python. Para evitar la contaminación entre los conjuntos de entrenamiento, validación y prueba, se llevó a cabo una partición adecuada. En total, se trabajó con un conjunto de imágenes completo.

El conjunto de entrenamiento se compuso del 70% de las imágenes disponibles, lo que proporcionó una base sólida para el desarrollo y ajuste del modelo. Este conjunto se utilizó para entrenar el modelo y ajustar sus parámetros internos.

Con la finalidad de supervisar y optimizar los hiper-parámetros del modelo, se reservó el 20% de las imágenes para el conjunto de validación. Aquí, se evaluaron diferentes configuraciones de hiper-parámetros y se seleccionó la combinación óptima que maximizó el rendimiento del modelo.

El 10% restante de las imágenes se destinó al conjunto de prueba. Estas imágenes no se utilizaron durante el entrenamiento ni la selección de hiper-parámetros, y se mantuvieron completamente separadas hasta la etapa final. Este conjunto se utilizó para evaluar el rendimiento imparcial y objetivo del modelo en datos no vistos previamente.

La partición cuidadosa de los conjuntos de datos de entrenamiento, validación y prueba permitió una evaluación imparcial y rigurosa de los modelos. Además, garantizó que los modelos no se hayan ajustado excesivamente a los datos de prueba y demostró la capacidad de generalización del modelo en escenarios del mundo real (Kahlout & Ekler, 2021).

Esta técnica de partición es ampliamente aceptado y utilizado en el campo del aprendizaje automático para garantizar resultados confiables y una evaluación precisa del rendimiento del modelo (Quang Vinh Nguyen et al., 2021).

### ***Anotación de datos***

Se utilizó la herramienta de Labeling para realizar las anotaciones de los cuadros delimitadores de los lobos marinos en todas las imágenes. Las etiquetas de la clase "sealion" se aplicaron de manera precisa, sin espacios entre los objetos y sus cajas delimitadoras. Estas anotaciones generaron un archivo de texto que contenía información sobre los límites del cuadro y la etiqueta de clase correspondiente a cada región marcada. Tanto el conjunto de datos experimental de cuadros como el archivo de anotaciones fueron utilizados para entrenar las dos versiones de Yolo.

### ***Configuración de modelos***

Para los modelos YOLO NAS y YOLO V8, se configuraron varios hiper-parámetros similares, lo que nos permitió comparar que modelo tiene mejor desempeño. En ambas arquitecturas se aplicó fine tuning. Se escogió el optimizador Adam por que presenta mayor eficiencia en comparación con otros optimizadores (Quang Vinh Nguyen et al., 2021). El algoritmo Adam actualiza los pesos y sesgos en el entrenamiento del modelo. La tasa de aprendizaje asociada al optimizador se estableció en 0.001, este parámetro afecta la velocidad de convergencia durante el entrenamiento. El tamaño del lote fue de 5, 4 workers y el total de épocas para la etapa de entrenamiento se estableció en 50.

### ***Transfer learning***

El Transfer Learning es una estrategia ampliamente empleada en los campos del aprendizaje automático y el aprendizaje profundo, especialmente en problemas de detección de objetos en imágenes. Esta técnica consiste en aprovechar el conocimiento y características

adquiridos por un modelo previamente entrenado en una tarea específica y aplicarlo a otra tarea relacionada.

En lugar de entrenar un modelo desde cero, se utiliza un modelo que ya ha sido entrenado en un conjunto de datos masivo y se le hacen ajustes en sus pesos y parámetros para adaptarlo a nuestro beneficio.

Los modelos pre-entrenados han aprendido representaciones de alto nivel y características generales en los datos, lo que los hace útiles para tareas relacionadas (Huang et al., 2017).

Investigaciones anteriores han demostrado que, utilizando un modelo pre-entrenado como punto de partida, se puede acelerar significativamente el proceso de entrenamiento y lograr resultados satisfactorios con conjuntos de datos limitados. Se ha demostrado que el uso del aprendizaje por transferencia, incluso en situaciones donde los conjuntos de datos de origen y destino son diferentes, supera a la inicialización aleatoria de los pesos de la CNN logrando una convergencia más eficiente y rápida (Situ et al., 2023).

Las versiones de YOLO NAS y YOLO V8 utilizados en este estudio son CNN recientes reconocidas por su versatilidad y capacidad de generalizar el conocimiento aprendido en conjuntos de datos distintos a aquellos en los que se entrenó originalmente. Esto se traduce en una mejora significativa en el rendimiento de los modelos en la detección de objetos en nuevos conjuntos de datos.

Para el desarrollo de este estudio se utiliza como punto de partida para el aprendizaje por transferencia los modelos pre-entrenados YOLO NAS y YOLO V8 con el conjunto de datos de COCO 2017, que contiene 330k imágenes a disposición con 80 categorías de objetos (Lin et al., 2014).

### *Métricas de evaluación*

Para evaluar el rendimiento de los dos modelos Yolo, se utilizan varias métricas de evaluación que incluyen el tamaño del parámetro, tiempo de inferencia por imagen, tiempo de entrenamiento, precisión, recuperación, puntuación F1 y mAP (precisión media promedio).

- El tamaño de los parámetros se relaciona con el número de parámetros que tiene el modelo que se entrena.
- El tiempo de Inferencia representa el promedio de la velocidad de detección del lobo marino en 30 imágenes de pruebas.
- La precisión se define como la cantidad de objetos que se detectaron en comparación con la totalidad de objetos presentes en una imagen.
- La recuperación (recall) es la cantidad de objetos encontrados correctamente en relación con los realmente presentes. La precisión y recuperación tienen intenciones contrarias: mientras menos cantidad de falsos negativos se generen más alta es la precisión, pero más baja la recuperación, y mientras menos falsos positivos se produzcan existirá menos precisión, pero mayor recuperación.
- El área que se encuentra debajo de la curva PR se llama mAP (precisión media promedio), y representa la precisión promedio de todas las clases que están especificadas en el modelo de prueba.
- La puntuación de F1 tiene en cuenta la recuperación y la precisión, brindando un equilibrio entre las dos. Una mayor puntuación en el ámbito de la F1 implica que el modelo está más entrenado. La pendiente PR (precisión-recuperación) expresa la relación entre la precisión y la recuperación.

Se estableció un umbral de IoU (intersección sobre unión) en 0.5.

Para calcular las métricas, se utilizan los siguientes términos: TP (verdaderos positivos) es el número de casos correctamente etiquetados como positivos, TN (verdaderos negativos) es el

número de casos correctamente etiquetados como negativos, FP (falsos positivos) es el número de casos incorrectamente etiquetados como positivos y FN (falsos negativos) es el número de casos que son positivos, pero están etiquetados como negativos. Además, N representa el número total de clases. Estos términos son fundamentales para calcular las métricas de evaluación en la detección de objetos.

$$Precision = \frac{TP}{TP + FP}$$

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

$$F_1 = 2 \times \frac{Precision * Recall}{Precision + Recall}$$

$$AP = \int_0^1 p(r) dr$$

$$mAP = \frac{1}{N} \sum AP_t$$

### Conteo vía detección

Este método utiliza un algoritmo para identificar en la imagen a los individuos mediante un detector visual de objetos. Estos son anotados mediante los cuadros delimitadores o etiquetas. Una vez identificados los lobos marinos el algoritmo realiza el conteo de estos.

## RESULTADOS Y DISCUSIÓN

Algorithm	Parameter Size (MB)	Training Epochs	Training Time (h)	Time Inference (ms)	Precision	Recall	mAP @0.5	F1-score	mAP @0.5:0.95	Cls loss
YOLO-NAS	177.83	50	16.032	143.39	0.782	0.9803	0.8543	0.869	0.7834	0.930
YOLOv8	136.7	50	15.048	150.9	0.835	0.825	0.884	0.83	0.653	0.528

Tabla 1. Comparación detallada de las métricas de evaluación entre YOLO NAS y YOLOv8.

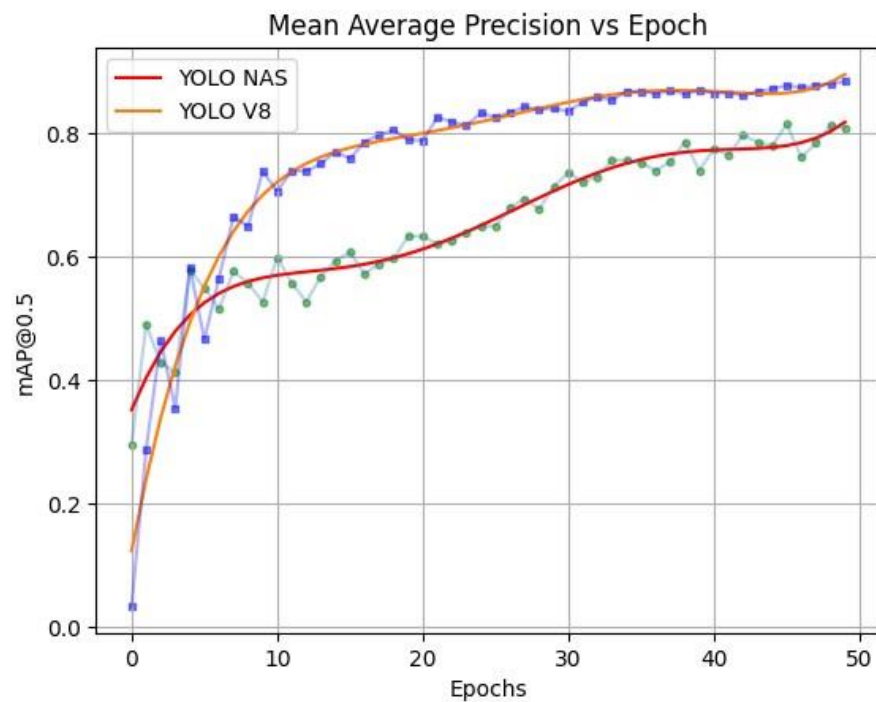
Se compararon dos versiones de YOLO para la detección de lobos marinos en un conjunto de datos de 3000 imágenes tomadas en las Islas Galápagos. Los dos modelos fueron entrenados utilizando las mismas imágenes de entrenamiento y se evaluaron utilizando las mismas imágenes de validación y prueba. Los resultados experimentales se muestran en la *Tabla 1*. Estos resultados

del experimento sugieren que Yolo Nas y Yolo V8 son prometedores para la detección de lobos marinos en la naturaleza.

La configuración experimental utilizada para el entrenamiento de ambos CNN se describe en la sección *Configuración experimental* del documento.

### ***Evaluación y rendimiento***

La *Tabla 1* muestra una evaluación detallada del desempeño de cada CNN en las imágenes de validación. Yolo V8, con 68,124,531 de parámetros, superó a Yolo Nas, que tiene 66,905,371 de parámetros. Yolo V8 tiene un mejor rendimiento que Yolo Nas en términos de mAP sobre un umbral de IoU de 0,5. Esto sugiere que YOLO V8 es más preciso y tiene un mejor rendimiento general que YOLO NAS.



*Figura 3. Precisión media promedio frente a rendimiento de recuperación para ambos modelos YOLO (NAS y v8)*

En la *Figura 3* muestra que el valor de mAP para YOLOv8 aumenta significativamente en las primeras 10 épocas y luego se estabiliza. El incremento acelerado se detiene y se observa un

efecto de convergencia general. Esto sugiere que YOLOv8 se estabiliza después de aproximadamente 40 épocas y no mejora significativamente después de este punto.

Por otro lado, el valor de mAP para Yolo Nas aumenta más lentamente en cada época y experimenta variaciones de intensidad que no afectan la tendencia creciente. Esto sugiere que Yolo Nas no se estabiliza tan rápidamente como YOLO V8 y continúa mejorando incluso después de 50 épocas. Aunque el valor de mAP de Yolo Nas en el umbral de IoU de 0,5 es inferior al de YOLO V8. Yolo NAS presenta un mayor valor de mAP con un umbral de IoU de 0,5 a 0,95. Esto indica que Yolo Nas es más preciso para la detección de objetos pequeños, mientras que YOLO V8 es más eficaz para la detección de objetos grandes.

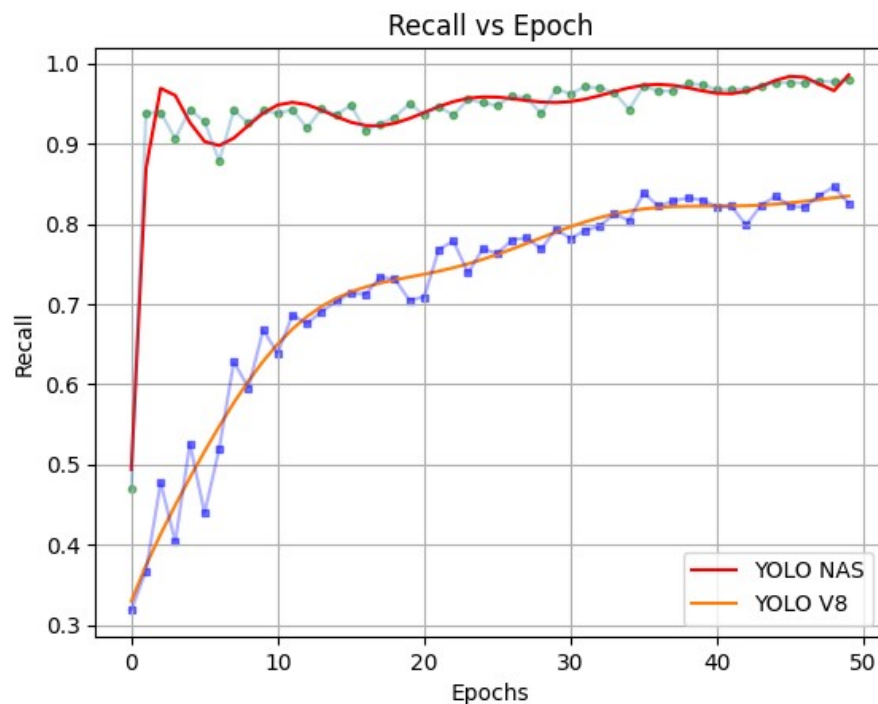


Figura 4. Rendimiento de recuperación frente a época para ambos modelos YOLO (NAS y v8)

En la *Figura 4* se muestra que el modelo Yolo Nas presenta un mayor valor de recuperación, 98.03%. La curva en Yolo Nas alcanza rápidamente más del 90% de recuperación antes de la época 5 lo que indica que el modelo logra identificar correctamente a la mayoría de los lobos marinos en las imágenes desde las primeras etapas del entrenamiento. Esto demuestra la eficiencia del modelo para aprender de forma correcta las características y patrones relevantes que



distinguen los lobos marinos. Al contrario, Yolo V8 se incrementa lentamente y parece que converge desde la época 40 obteniendo un valor de 82.5%. Esto sugiere que el modelo YoloV8 tiene mayores dificultades para identificar y recuperar algunos lobos marinos en las imágenes. Esto demuestra que el modelo requiere más entrenamiento para mejorar su capacidad de reconocer correctamente los lobos marinos.

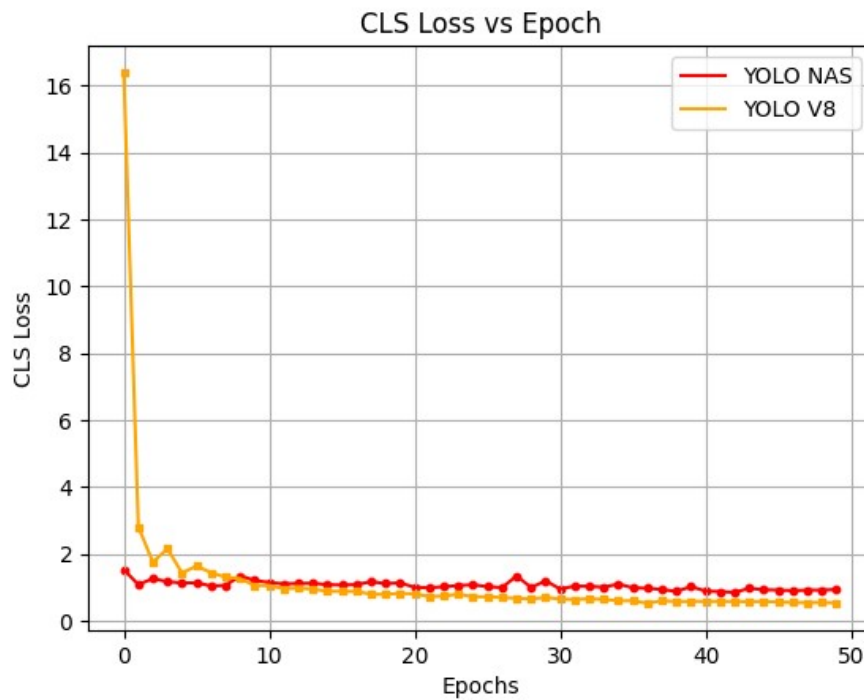


Figura 5. Rendimiento de CLS Loss vs Epoch para ambos modelos YOLO (NAS y v8).

Las curvas obtenidas de la pérdida de clasificación en la *Figura 5* muestra que ambos modelos convergen rápidamente. Esto demuestra la excelente capacidad de Yolo Nas y Yolo V8 de clasificar a los lobos marinos desde las primeras épocas. Sin embargo, Yolo V8 obtiene un menor error de clasificación obteniendo un valor de 0.52 en comparación con Yolo Nas que tiene 0.93. Lo que sugiere un mejor desempeño de Yolo v8 para la clasificación.

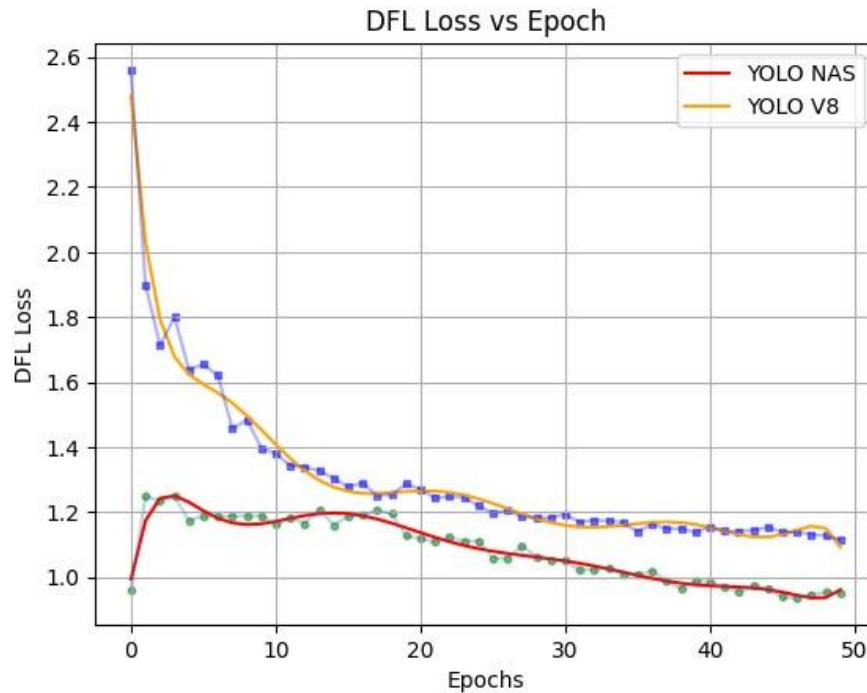


Figura 6. Rendimiento DFL Loss vs Epoch para ambos modelos YOLO (NAS y v8)

En la *Figura 6* de la pérdida de focalización distributiva Yolo Nas logra un mínimo mayor que Yolo V8. Esta métrica indica que Yolo Nas logra un menor error de localización y está generando predicciones precisas de las ubicaciones de los lobos marinos en el conjunto de imágenes.

Yolo Nas sobresale en las detecciones de lobos marinos que están solos en la imagen, como se muestra en *Anexo A*. Sin embargo, cuando la imagen contiene múltiples individuos, a menudo superpuestos o con oclusión causado por elementos del ambiente como rocas o árboles, u otros lobos marinos y cuando presentan un gran cambio en su tamaño y posición o cuando están muy cerca unos de otros, ambos modelos tienden a detectar los ejemplares que están menos cubiertos por otros, como se evidencia en *Anexo B*. Esto sugiere que los modelos presentan poco desempeño en discriminar los límites de las instancias cuando las características se superponen.

Así mismo, Yolo Nas, realiza etiquetas incorrectas en imágenes que contienen como fondo rocas del mismo color y forma que un lobo marino, como se observa en *Anexo B*.

Para solucionar este problema, se podría aumentar la cantidad de épocas de entrenamiento y agregar más imágenes que presenten fondos rocosos al conjunto de datos, lo que permitiría que el modelo aprenda a diferenciar mejor el fondo de los objetos. Además, específicamente en Yolo Nas se podría ajustar las múltiples transformaciones que aplica a las imágenes en el preprocesamiento de los datos de entrenamiento. Esto ayudaría al modelo a ser más sensible y diferenciar las características que distinguen a las rocas de los lobos marinos.

Ambos modelos mantienen un buen rendimiento cuando predicen colonias de lobos marinos que aparecen en grandes cantidades y dispersos. En este contexto, el modelo Yolo Nas muestra un mejor desempeño que Yolo V8 al detectar la mayoría de los ejemplares.

Estos hallazgos resaltan las diferencias en el comportamiento y el rendimiento de ambos modelos. Yolo Nas muestra una mayor tasa de recuperación y pérdida de focalización distributiva. Por otro lado, YOLO V8 presenta un mejor mAP (precisión media promedio), como también pérdida de clasificación superior y una mayor sensibilidad para diferenciar entre rocas y lobos marinos, lo que sugiere una mayor estabilidad y precisión en las predicciones.

## CONCLUSIONES Y TRABAJO FUTURO

Los resultados sugieren que Yolo Nas y Yolo V8 son modelos prometedores ya que lograron detectar casi todos los lobos marinos en las imágenes. Sin embargo, Yolo V8 parece tener una mayor robustez en la detección en condiciones difíciles y complejas, pero Yolo Nas suele detectar un mayor número de individuos en las imágenes incluso cuando son de tamaño pequeño.

Es relevante mencionar que, aunque no se llevó a cabo un ajuste exhaustivo de los hiper-parámetros (que no estaba dentro del alcance de este trabajo), los modelos analizados en este trabajo obtuvieron excelentes detecciones, logrando prometedores resultados.

Se recomienda en un futuro estudio incrementar el conjunto de imágenes y afinar aún más los hiper-parámetros para alcanzar una mayor eficacia. Además, se sugiere ampliar el entrenamiento a otras clases de especies dentro de las Islas Galápagos como iguanas marinas o tortugas gigantes, con el objetivo de utilizar estos modelos en diversas aplicaciones para la conservación de especies nativas y endémicas.

## REFERENCIAS BIBLIOGRÁFICAS

- Villegas-Amtmann, S., Costa, D. P., Tremblay, Y., Salazar, S., & Auriol-Gamboa, D. (2008). Multiple foraging strategies in a marine apex predator, the Galapagos sea lion *Zalophus wollebaeki*. *Marine Ecology Progress Series*, 363, 299–309.
- Riofrío-Lazo, M., Arreguín-Sánchez, F., & Páez-Rosas, D. (2017). Population abundance of the endangered Galapagos sea lion *Zalophus wollebaeki* in the southeastern Galapagos archipelago. *PLoS One*, 12(1), e0168829.
- Páez-Rosas, D., & Guevara, N. (2017). Management strategies and conservation status of Galapagos sea lion populations at San Cristóbal Island, Galapagos, Ecuador. *Tropical Pinnipeds: Bio-Ecology, Threats and Conservation*, 159–175.
- Kahlout, K. M., & Ekler, P. (2021). Algorithmic Splitting: A Method for Dataset Preparation. *IEEE Access*, 9, 125229–125237. doi:10.1109/ACCESS.2021.3110745
- Nguyen, Q. H., Ly, H.-B., Ho, L. S., Al-Ansari, N., Van Le, H., Tran, V. Q., ... Pham, B. T. (2021). Influence of Data Splitting on Performance of Machine Learning Models in Prediction of Shear Strength of Soil. *Mathematical Problems in Engineering*.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2022). *An introduction to statistical learning : with applications in R* (Second edition). Boston: Springer.
- Christin, S., Hervet, É., & Lecomte, N. (2019). Applications for deep learning in ecology. *Methods in Ecology and Evolution*, 10(10), 1632–1644. doi:10.1111/2041-210X.13256
- Mahmud, M., Kaiser, M. S., Hussain, A., & Vassanelli, S. (2018). Applications of Deep Learning and Reinforcement Learning to Biological Data. *IEEE Transactions on Neural Networks and Learning Systems*, 29(6), 2063–2079. doi:10.1109/TNNLS.2018.2790388
- Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., ... Others. (2018). Opportunities and obstacles for deep learning in biology and medicine. *Journal of The Royal Society Interface*, 15(141), 20170387.

- Li, J., Xu, W., Deng, L., Xiao, Y., Han, Z., & Zheng, H. (2023). Deep learning for visual recognition and detection of aquatic animals: A review. *Reviews in Aquaculture*, 15(2), 409–433. doi:10.1111/raq.12726
- Deci. (2023). YOLO-NAS by Deci Achieves State-of-the-Art Performance on Object Detection Using Neural Architecture Search. Retrieved from <https://deci.ai/blog/yolo-nas-object-detection-foundation-model/>
- Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLOv8 (Version 8.0.0). Retrieved from <https://github.com/ultralytics/ultralytics>
- Situ, Z., Teng, S., Feng, W., Zhong, Q., Chen, G., Su, J., & Zhou, Q. (2023). A transfer learning-based YOLO network for sewer defect detection in comparison to classic object detection methods. *Developments in the Built Environment*, 15, 100191. doi:10.1016/j.dibe.2023.100191
- Huang, Z., Pan, Z., & Lei, B. (2017). Transfer Learning with Deep Convolutional Neural Network for SAR Target Classification with Limited Labeled Data. *Remote Sensing*, 9(9). doi:10.3390/rs9090907
- Zaheer, R., & Shaziya, H. (2019). A Study of the Optimization Algorithms in Deep Learning. *2019 Third International Conference on Inventive Systems and Control (ICISC)*, 536–539. doi:10.1109/ICISC44355.2019.9036442
- Hussain, M. (2023). YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection. *Machines*, 11(7). doi:10.3390/machines11070677
- Amjoud, A. B., & Amrouch, M. (2023). Object Detection Using Deep Learning, CNNs and Vision Transformers: A Review. *IEEE Access*, 11, 35479–35516. doi:10.1109/ACCESS.2023.3266093

- Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., ... Dollár, P. (2015). Microsoft COCO: Common Objects in Context. *arXiv [Cs.CV]*. Retrieved from <http://arxiv.org/abs/1405.0312>
- Padubidri, C., Kamilaris, A., Karatsiolis, S., & Kamminga, J. (2021). Counting sea lions and elephants from aerial photography using deep learning with density maps. *Animal Biotelemetry*, 9(1), 1–10.
- Páez-Rosas, D., Moreno-Sánchez, X., Tripp-Valdez, A., Elorriaga-Verplancken, F. R., & Carranco-Narváez, S. (2020). Changes in the Galapagos sea lion diet as a response to El Niño-Southern Oscillation. *Regional Studies in Marine Science*, 40, 101485.
- Denkinger, J., Gordillo, L., Montero-Serra, I., Murillo, J. C., Guevara, N., Hirschfeld, M., ... Dan, M. (2015). Urban life of Galapagos sea lions (*Zalophus wollebaeki*) on San Cristobal Island, Ecuador: colony trends and threats. *Journal of Sea Research*, 105, 10–14.
- Vega-Mariño, P., Olson, J., Howitt, B., Criollo, R., Figueroa, L., Orlando, S. A., ... Garcia-Bereguian, M. A. (2023). A recent distemper virus outbreak in the growing canine populations of Galapagos Islands: a persistent threat for the endangered Galapagos Sea Lion. *Frontiers in Veterinary Science*, 10, 1154625.
- Lamba, A., Cassey, P., Segaran, R. R., & Koh, L. P. (2019). Deep learning for environmental conservation. *Current Biology*, 29(19), R977–R982.
- Gregory, T. M., Livingston, I., Hawkins, E. C., Loyola, A., Cave, A., Vaden, S. L., ... Others. (2023). DIROFILARIA IMMITIS IDENTIFIED IN GALAPAGOS SEA LIONS (*ZALOPHUS WOLLEBAEKI*): A WILDLIFE HEALTH AND CONSERVATION CONCERN. *Journal of Wildlife Diseases*.
- Denkinger, J., Guevara, N., Ayala, S., Murillo, J. C., Hirschfeld, M., Montero-Serra, I., ... Others. (2017). Pup mortality and evidence for pathogen exposure in Galapagos sea lions (*Zalophus*

wollebaeki) on San Cristobal Island, Galapagos, Ecuador. *Journal of Wildlife Diseases*, 53(3), 491–498.

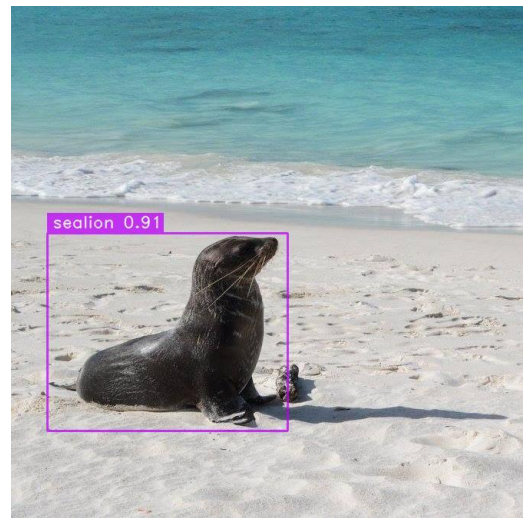
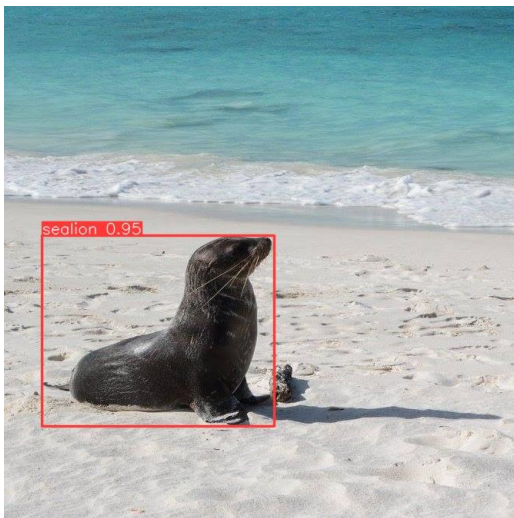
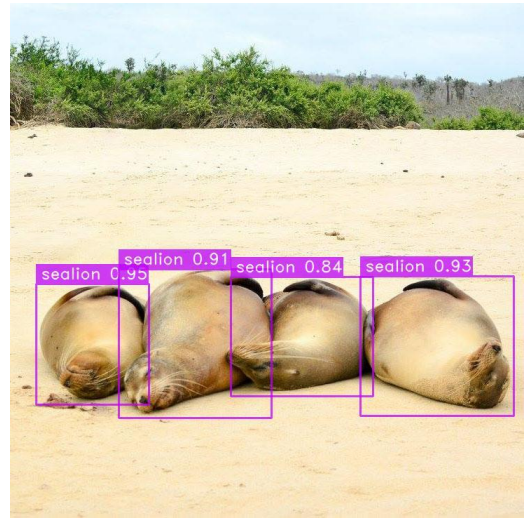
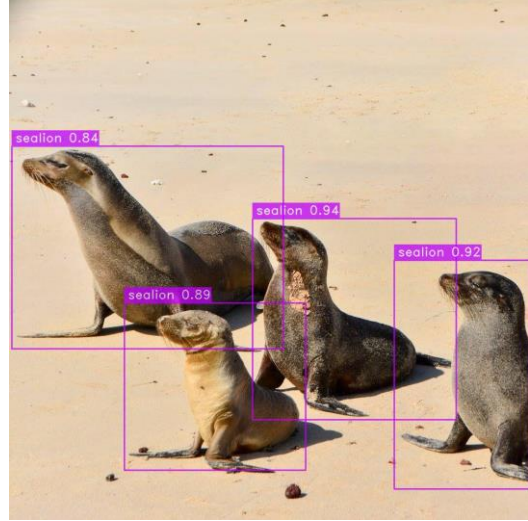
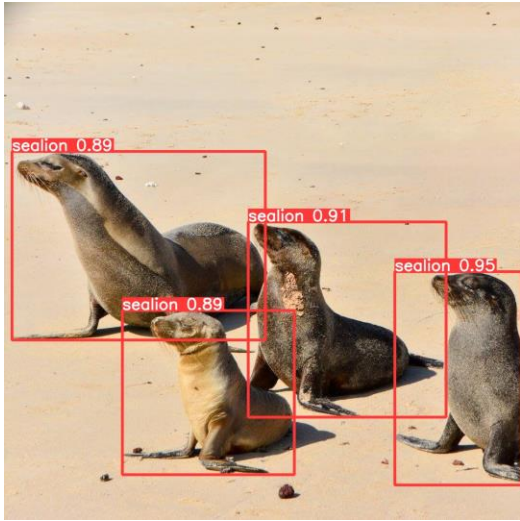
Aharon, S., Louis-Dupont, Ofri Masad, Yurkova, K., Lotem Fridman, Lkdci, ... Eran-Deci. (2021). Super-Gradients. *GitHub Repository*. doi:10.5281/ZENODO.7789328



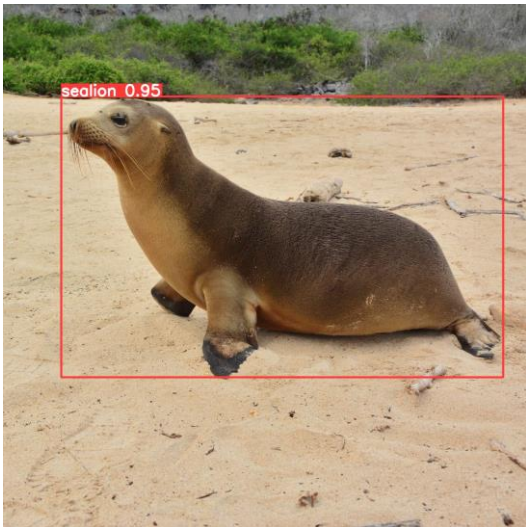
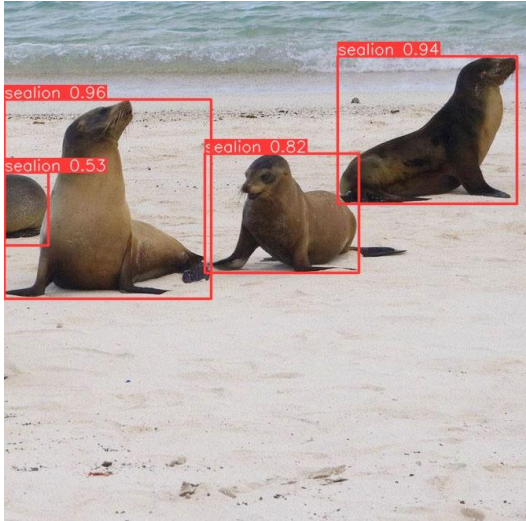
# ANEXO A: RESULTADOS DETECCIÓN DEL LOBO MARINO

## YOLOv8

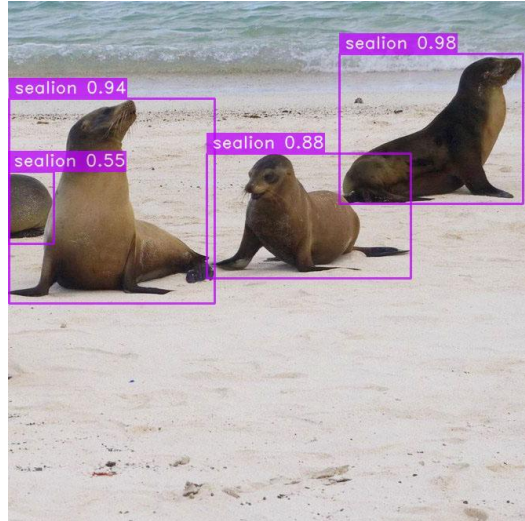
## YOLO-NAS



### YOLOv8



### YOLO-NAS

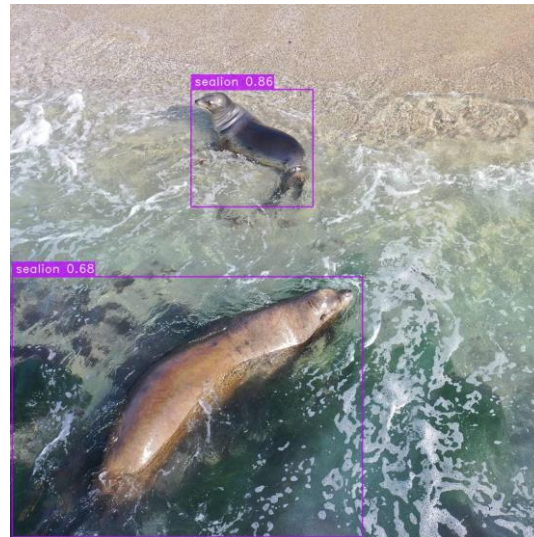




### YOLOv8



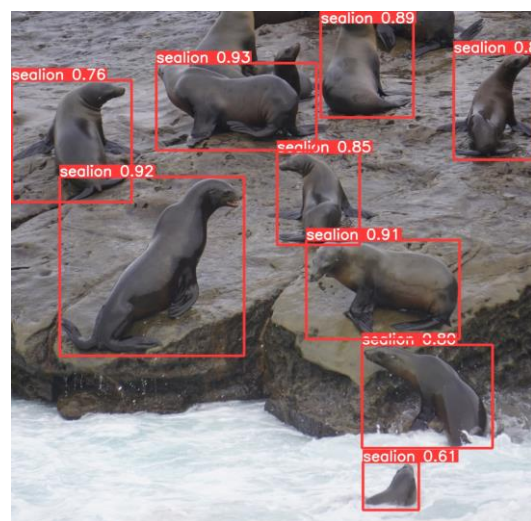
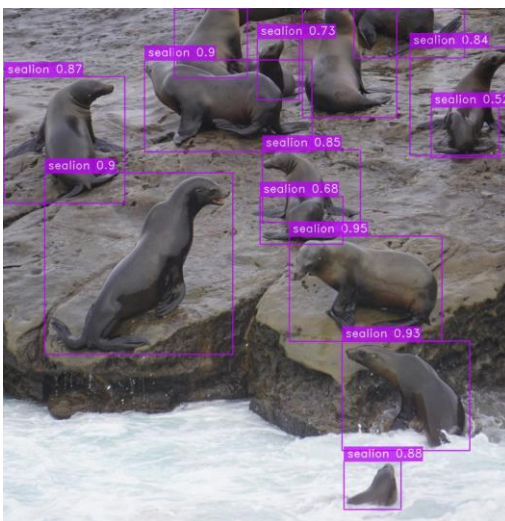
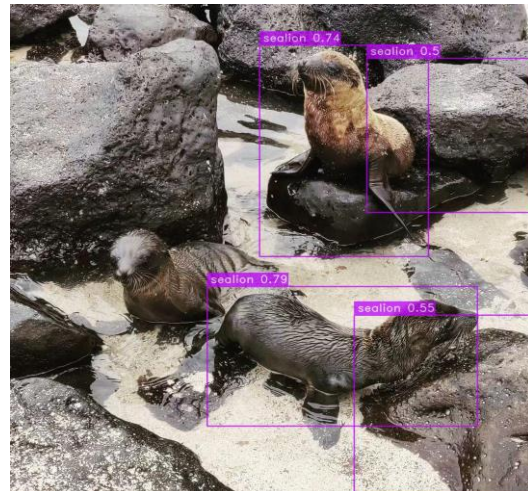
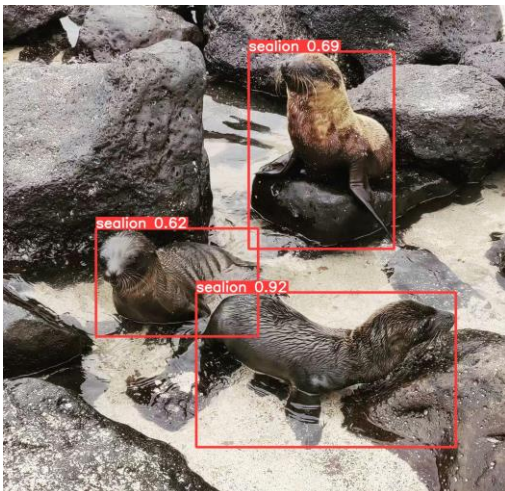
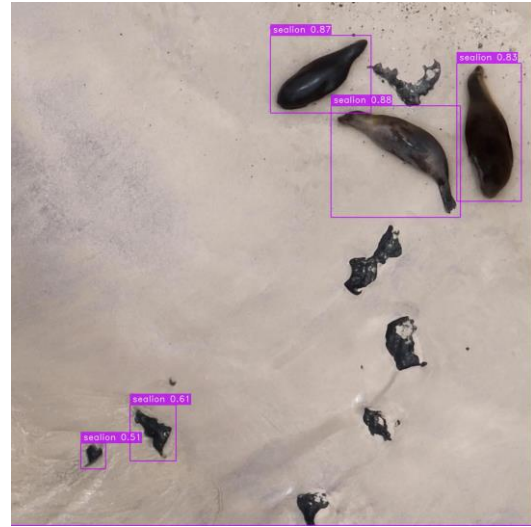
### YOLO-NAS



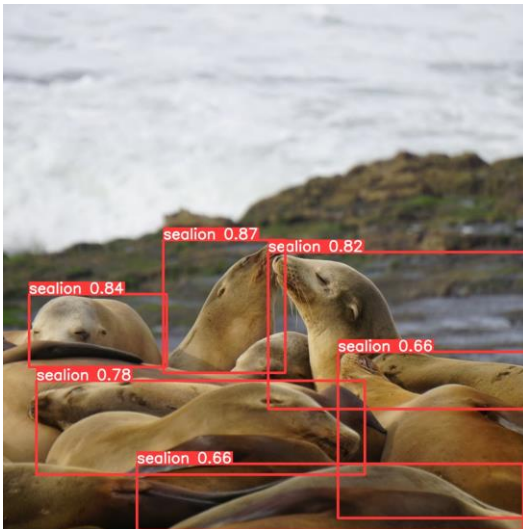
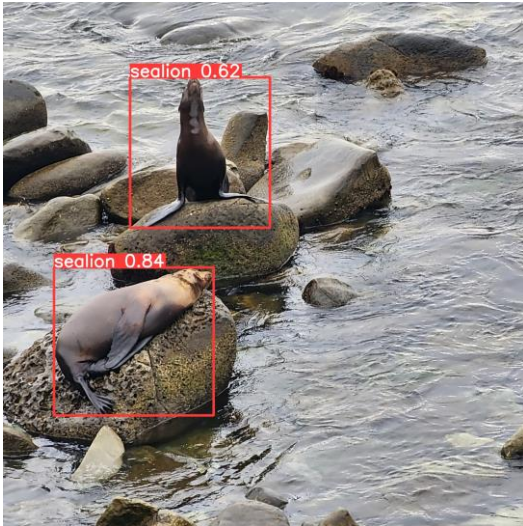
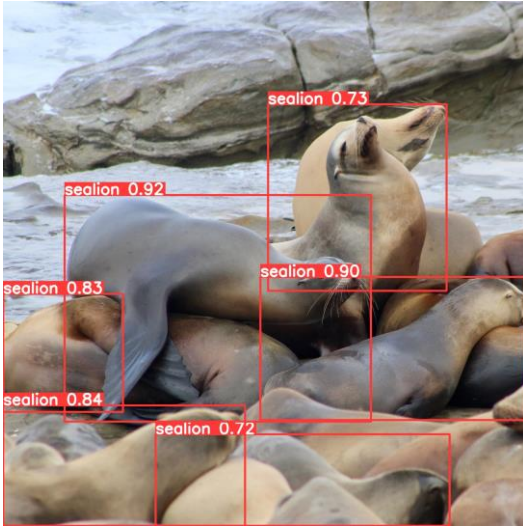
# ANEXO B: RESULTADOS DETECCIÓN DEL LOBO MARINO EN ENTORNOS DESAFIANTES

YOLOv8

YOLO-NAS



### YOLOv8



### YOLO-NAS

