



Universidad
de La Laguna

INTEMARES



Mapeo de alta resolución asistido por aprendizaje profundo de densidades de *Dendrophyllia cornigera* (Lamarck, 1816) y *Phakellia ventilabrum* (Linnaeus, 1767) en el Sistema de Cañones de Avilés y Capbreton, Golfo de Vizcaya

High-resolution mapping assisted by deep learning of *Dendrophyllia cornigera* (Lamarck, 1816) and *Phakellia ventilabrum* (Linnaeus, 1767) densities in the Avilés and Capbreton Canyon System, Bay of Biscay

Alberto Gayá Vilar

Máster en Biología Marina: Biodiversidad y Conservación

Julio de 2023

AUTORIZACIÓN DEL TUTOR/ES

El **Dr. Sabrina Clemente Martín**, Profesor Contratada Doctora de la Universidad de La Laguna y el **Dr. Elena Prado Ortega** Investigadora en el **Instituto Español de Oceanografía, Centro Oceanográfico de Santander (COST-IEO, CSIC)**, como Tutor Académico y Tutor Externo, respectivamente,

DECLARAN:

Que la memoria presentada por **D. Alberto Gayá Vilar** titulada “**Mapeo de alta resolución asistido por aprendizaje profundo de densidades de *Dendrophyllia cornigera* (Lamarck, 1816) y *Phakellia ventilabrum* (Linnaeus, 1767) en el Sistema de Cañones de Avilés y Capbreton, Golfo de Vizcaya**”, ha sido realizada bajo su dirección y consideran que reúne todas las condiciones de calidad y rigor científico requeridas para optar a su presentación como Trabajo de Fin de Máster, en el Máster Oficial de Postgrado de Biología Marina: Biodiversidad y Conservación de la Universidad de La Laguna, curso académico 2022-2023.

Y para que así conste y surta los efectos oportunos, firman el presente informe favorable en San Cristóbal de La Laguna a **07 de julio** de 2023.

PRADO
ORTEGA, MARIA
ELENA
(AUTENTICACIÓN)
N)

Firmado digitalmente
por PRADO ORTEGA,
MARIA ELENA
(AUTENTICACIÓN)
Fecha: 2023.07.07
17:20:37 +02'00'

Fdo. **Dr. Elena Prado Ortega**

CLEMENTE
MARTIN MARIA
SABRINA -
78700248K

Firmado digitalmente
por CLEMENTE
MARTIN MARIA
SABRINA - 78700248K
Fecha: 2023.07.07
15:51:40 +01'00'

Fdo. **Dr. Sabrina Clemente Martín**

ÍNDICE

Resumen.....	1
Abstract.....	1
Introducción.....	2
Materiales y métodos.....	5
Área de estudio.....	5
Imágenes submarinas ROTV Politolana.....	7
Preprocesamiento de imágenes y entrenamiento de algoritmos.....	9
Etiquetado automático de especies en imágenes submarinas utilizando aprendizaje profundo.....	10
Selección del modelo de detección de objetos.....	12
Flujo de trabajo continuo para la generación mapas de densidad de especie.....	13
Resultados.....	16
Comparativa de modelos de detección de objetos.....	16
Estudio de densidad de especies.....	20
Discusión.....	21
Conclusión.....	25
Agradecimientos.....	26
Bibliografía.....	27

Resumen

En España, el hábitat 1170 Arrecifes está ampliamente distribuido, desde la línea de costa hasta los fondos profundos. Entre su variedad de tipologías, este trabajo se ha centrado en los afloramientos rocosos de la plataforma circalitoral del mar Cantábrico, donde una de las principales comunidades está estructurada por el coral amarillo *Dendrophyllia cornigera* y la esponja de copa *Phakellia ventilabrum*. El objetivo principal de este estudio es establecer un flujo de trabajo continuo basado en modelos de aprendizaje profundo para extraer datos de densidad de especies a partir de imágenes en bruto obtenidas por un ROV. Para lograrlo, se evaluaron y compararon diferentes modelos de detección de objetos, como YOLOv7, YOLOv8 y Faster R-CNN, en plataforma continental aledaña a dos Sistemas de Cañones Submarinos (SCS). Los modelos fueron entrenados y validados en imágenes del SCS de Avilés. Posteriormente, se validaron los modelos previamente entrenados en imágenes del SCS de Capbreton, con el objetivo de determinar si los modelos eran capaces de detectar las especies en un ambiente diferente al que fueron entrenados. Los resultados destacan que YOLOv8 sobresalió en precisión, recuperación y F1, obteniendo una precisión del 92.4% en el SCS Avilés. En el SCS Capbreton, el modelo también supo generalizar de manera efectiva ambas especies. Con el modelo YOLOv8 previamente entrenado, se generaron 27.668 anotaciones automáticas. Las anotaciones permitieron calcular una densidad media de 1,07 y 3,82 individuos/m² de *P. ventilabrum* y *D. cornigera*, respectivamente, en las dos zonas de estudio con un área barrida de 5647,48 m². Estos resultados resaltan el potencial del aprendizaje profundo para mejorar la eficiencia y precisión en el seguimiento de ecosistemas marinos vulnerables, permitiendo tomar decisiones informadas que puedan tener un impacto positivo en la conservación del medio marino.

Palabras clave: Inteligencia artificial, Ecosistema marino vulnerable, Flujo de trabajo continuo, Modelo de detección de objetos, Red Natura 2000

Abstract

In Spain, the 1170 Reef habitat is widely distributed, from the coastline to deep waters. Among its variety of typologies, this work has focused on the rocky outcrops of the circalittoral platform of the Cantabrian Sea, where communities are structured by the yellow coral *Dendrophyllia cornigera* and the cup sponge *Phakellia ventilabrum*. The main objective of this study was to establish a pipeline based on deep learning models to extract species density data from raw images obtained by ROV. To achieve this, different object detection models, such as YOLOv7, YOLOv8 and Faster R-CNN, were evaluated and compared in various Submarine Canyon Systems (SCS). The models were trained and validated on images from the Avilés SCS. Subsequently, the previously trained models were validated on images from the Capbreton SCS, with the aim of determining whether the models were capable of detecting species in an environment different from the one they were trained on. The results highlight that YOLOv8 excelled in accuracy, recall and F1, achieving an accuracy of 92.4% in the Avilés SCS. In the Capbreton SCS, the model also effectively generalized both species. With the previously trained YOLOv8 model, 27,668 automatic annotations were generated. The annotations allowed to calculate an average density of 1.07 and 3.82 individuals/m² for *P.*

ventilabrum and *D. cornigera*, respectively, in the two study areas with a swept area of 5647.48 m². These results highlight the potential of deep learning to improve efficiency and accuracy in monitoring vulnerable marine ecosystems, allowing informed decisions to be made that can have a positive impact on marine conservation.

Keywords: Artificial intelligence, Vulnerable marine ecosystem, Pipeline, Object detection model, Natura 2000 Network

Introducción

La Directiva Hábitat (Directiva 92/43/CEE) crea la red “Natura 2000”, red de espacios europeos que trata de mantener o, si es posible, restablecer, un estado de conservación favorable de determinados tipos de hábitats naturales y de ciertas especies animales y vegetales. La red Natura 2000 de ámbito marino es parte integrante de la red ecológica europea Natura 2000, y constituye la aplicación de la Directiva Hábitat y la Directiva Aves (Directiva 2009/147/CE) en el medio marino, consideradas las dos herramientas legislativas más importantes para la conservación de la biodiversidad en Europa. La red Natura 2000 está compuesta por los Lugares de Importancia Comunitaria (LIC), que con el tiempo pasan a designarse Zonas Especiales de Conservación (ZEC), y las Zonas de Especial Protección para Aves (ZEPA).

Entre los tipos de hábitats marinos que figuran en el Anexo I de la Directiva Hábitats (92/43/CEE) como tipos de hábitats naturales de interés comunitario cuya conservación requiere la designación de Zonas Especiales de Conservación (ZEC) se encuentra el Hábitat 1170 Arrecifes. Se consideran arrecifes en sentido de la Directiva a todos aquellos sustratos duros compactos que afloran sobre fondos marinos en la zona sublitoral (sumergida) o litoral (intermareal), ya sean de origen biogénico o geológico.

En España el hábitat 1170 Arrecifes está distribuido por todas las costas y aguas marinas, y se puede encontrar desde la línea de costa hasta los fondos profundos, pudiendo ocupar amplias superficies. Dentro de esta gran variedad de tipologías del hábitat 1170, este trabajo se ha centrado en dos afloramientos rocosos dentro de la plataforma circalitoral del mar Cantábrico. Estos afloramientos son considerados ecosistemas marinos vulnerables (EMV) debido a su importancia como puntos críticos de biodiversidad y funcionamiento de los ecosistemas en las

profundidades marinas (FAO, 2009). Los sustratos rocosos circalitorales, situados en el sistema fital por debajo del nivel máximo de distribución de las fanerógamas marinas y algas fotófilas, y extendiéndose hasta la máxima profundidad de las algas esciáfilas, se caracterizan por una luz muy difusa y condiciones hidrodinámicas más constantes que en los lechos superiores. La profundidad a la que comienza la zona circalitoral depende directamente de la intensidad de la luz que llega al fondo marino. La mayoría de los fondos rocosos circalitorales están dominados por especies animales, debido a la debilidad de la luz. El número de especies que caracterizan estos fondos marinos puede ser muy elevado, dependiendo de las diferentes zonas geográficas, la geomorfología del fondo y los diferentes factores que les afectan (Dominguez-Carrió et al., 2022).

En la plataforma rocosa circalitoral cantábrica, las comunidades están compuestas por un gran número de especies de esponjas y coralígeno, que proveen de estructura tridimensional a estos hábitats, pertenecientes al hábitat 1170 Arrecifes. Sin embargo, a pesar de ser especies estructurantes, su pequeño tamaño y su distribución altamente fraccionada sobre el fondo dificulta enormemente su cartografiado. A la vez, el seguimiento de estas especies y la distribución de las comunidades en el tiempo y espacio resulta esencial para asegurar la protección de estos hábitats. La utilización de imágenes adquiridas a través de vehículos operados remotamente (ROV) se posiciona como una herramienta valiosa para abordar este desafío.

El uso de vehículos submarinos genera una gran cantidad de muestras *in situ*, no destructivas, representativas y potencialmente repetibles, y ello permite no sólo una caracterización completa de la diversidad bentónica sino que también podría sentar las bases para un programa de seguimiento a largo plazo (Dominguez-Carrió et al., 2022). Sin embargo, procesar toda esta información ha creado un cuello de botella, dado que la anotación de datos visuales, es decir, la identificación y etiquetado de especies en las imágenes, exige mucho tiempo, trabajo y dinero (Weinstein, 2017). Además, los ecosistemas profundos son entornos complicados con condiciones de luz desequilibradas, bajo contraste y la presencia de oclusión y camuflaje de organismos. Bajo estas circunstancias, los objetos capturados por la cámara del ROV se vuelven difíciles de identificar (Song et al., 2022). Para abordar estos problemas y obtener información cuantitativa de las imágenes submarinas, se han desarrollado nuevas herramientas para el análisis automatizado de las características de las imágenes. Uno de los enfoques más prometedores es el uso de técnicas de aprendizaje profundo basadas en redes

neuronales, una combinación de inteligencia artificial y visión artificial. Esto implica la aplicación de múltiples capas de algoritmos de aprendizaje automático altamente interconectadas para lograr resultados mejorados a partir de las imágenes sin procesar (Olden et al., 2008; Le Cun et al., 2015). Estas técnicas ya han logrado resultados formidables realizando diferentes tareas de ecología marina como clasificación de corales (Bhandarkar et al., 2022; Mahmood et al., 2017; Raphael et al., 2020), detección y clasificación de peces (Zhong et al., 2022; Siddiqui et al., 2018; Knausgård et al., 2021), o identificación de diferentes tipos de fauna bentónica (Abad-Uribarren et al., 2022; Song et al., 2022; Liu & Wang 2022).

Dentro del campo del aprendizaje profundo, los detectores de objetos se pueden dividir en dos grupos: los detectores de dos etapas y los detectores de una sola etapa. Los detectores de dos etapas incluyen Faster R-CNN (Ren et al., 2017), que genera primero una serie de propuestas de regiones (RPN) y luego determinan la categoría y la ubicación del objeto. Por otro lado, los detectores de una sola etapa, como YOLO (Redmon et al., 2016), identifican y ubican objetos en una imagen en un solo paso. Estos modelos de detección de objetos pueden ser utilizados como herramientas para automatizar los procesos de identificación de especies y generar mapas de densidad precisos para hacer seguimientos del ecosistema. Así, los estudios que emplean estos modelos permiten tener un registro detallado de los ecosistemas y ayudan a identificar patrones en la salud del hábitat y en la biodiversidad que alberga.

Los ecosistemas marinos están sometidos a numerosas amenazas e impactos, como el cambio climático, la contaminación por desechos, la pesca comercial y la minería en aguas profundas. Por esta razón, es importante generar un mapeo detallado de sus ecosistemas más vulnerables para poder tomar decisiones informadas sobre su gestión y conservación. Un mapeo preciso puede ayudar a identificar áreas críticas que requieren protección y a desarrollar estrategias efectivas para mitigar los impactos negativos en el ecosistema (Rodríguez-Basalo et al., 2022).

En el presente estudio, se llevó a cabo una evaluación de la densidad de especies estructurantes del hábitat 1170 en dos zonas de plataforma rocosa circalitoral del Cantábrico. Para ello se utilizaron modelos de detección de objetos que permiten identificar y etiquetar las especies en las imágenes submarinas de forma automática. En primer lugar se realizó una comparación de modelos de detección de objetos con diferentes arquitecturas neuronales,

incluyendo YOLOv7, YOLOv8 y Faster R-CNN, para determinar cuál funciona mejor en el contexto de la generación de mapas de densidad de especies a partir de imágenes geolocalizadas obtenidas de un transecto. El modelo que obtuvo las mejores métricas se utilizó para generar un flujo de trabajo continuo que permitió obtener mapas de densidad de especies detallados. El objetivo final fue elaborar un mapa base inicial en el seguimiento del estado de salud del ecosistema, mediante la descripción geográfica detallada de especies estructurantes del hábitat 1170, o como soporte en la ayuda a la toma de decisiones de las medidas de gestión de las zonas Red Natura 2000. Los resultados esperados de la aplicación de este modelo incluyen la generación automática de datos de densidad y presencia geográfica para las especies bentónicas *D. cornigera* y *P. ventilabrum*, y la presentación de los resultados a través de mapas de densidades de las especies.

Materiales y métodos

Área de estudio

La zona de estudio se centra en dos afloramientos rocosos de la plataforma circalitoral, situados dentro del Sistema de Cañones Submarinos de Avilés (SCSA), declarado Lugar de Interés Comunitario (LIC), y en la zona del Sistema de Cañones Submarinos de Capbreton (SCSC), ubicados ambos en el Mar Cantábrico al sur del Golfo de Vizcaya (Figura 1). Actualmente el LIC de SCSA está siendo objeto de diferentes estudios que favorezcan su declaración desde su actual figura de LIC a Zona de Especial Conservación (ZEC). Así mismo, la zona del SCSC está siendo objeto de estudio para su declaración como nuevo LIC para pasar a formar parte de Red Natura 2000 marina. Estos estudios se enmarcan dentro de las acciones llevadas a cabo en el proyecto LIFE IP INTEMARES.

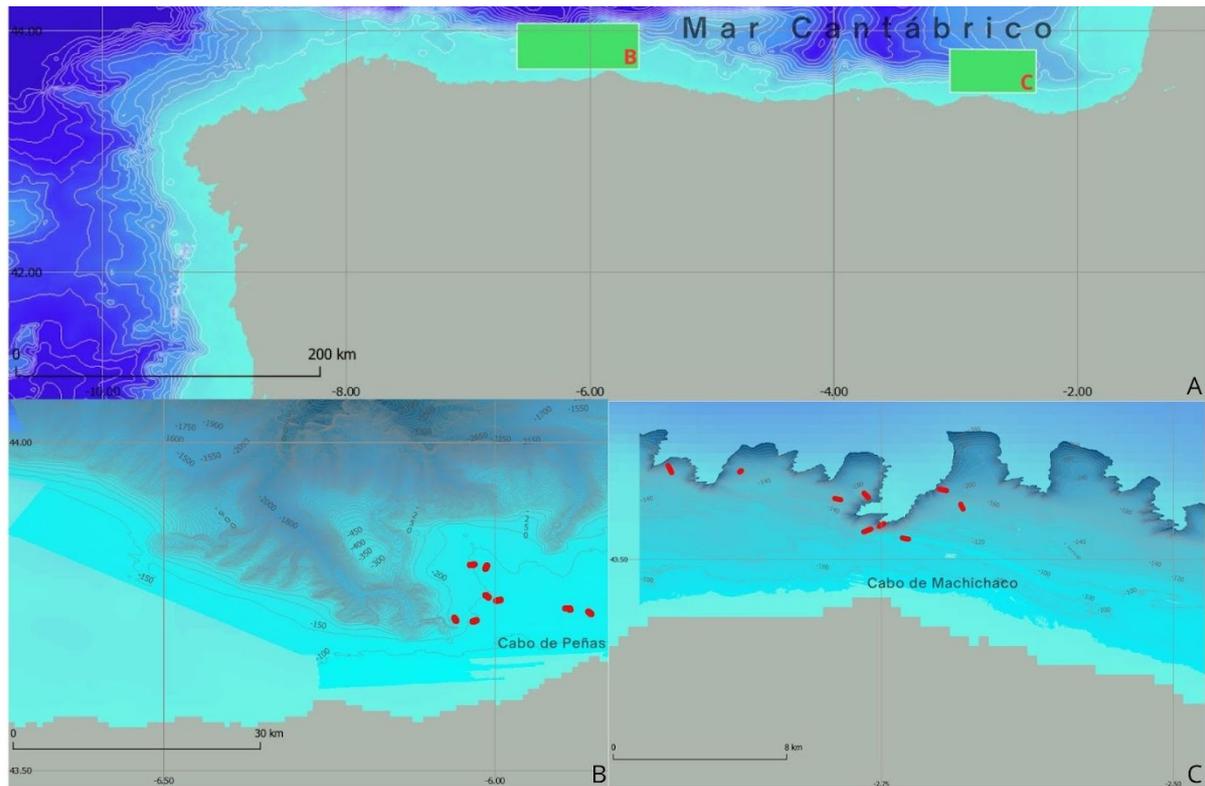


Figura 1. Mapa de ubicación (A) que muestra las áreas de estudio destacadas mediante un rectángulo verde, situadas en el Mar Cantábrico. Además, se presentan las representaciones detalladas de la batimetría del sistema de cañones submarinos de Avilés (B) y del sistema de cañones submarinos de Capbreton (C), donde se destacan los transectos ROTV identificados mediante puntos rojos

A lo largo de todo el Golfo de Vizcaya, la plataforma continental es generalmente estrecha, un rasgo característico típico de los márgenes continentales compresivos (Ercilla et al., 2008), y además se caracteriza por la existencia de afloramientos rocosos debido principalmente a los mecanismos de transporte sedimentario asociados a la dinámica oceanográfica. Estos afloramientos rocosos son zonas que albergan una variedad de comunidades bentónicas, muchas de las cuales se encuentran dentro de lo que se conoce como hábitat 1170, y conforman un hábitat heterogéneo y complejo capaz de albergar una gran riqueza de especies. Dentro de estas comunidades se han seleccionado dos especies como especies objetivo dentro de este estudio, debido a su capacidad estructurante del hábitat 1170 en toda la plataforma circalitoral rocosa del Cantábrico (Rodríguez-Basalo et al., 2022). Estas especies son el coral amarillo *Dendrophyllia cornigera* (Lamarck, 1816) y la esponja copa *Phakellia ventilabrum* (Linnaeus, 1767) (Figura 2).

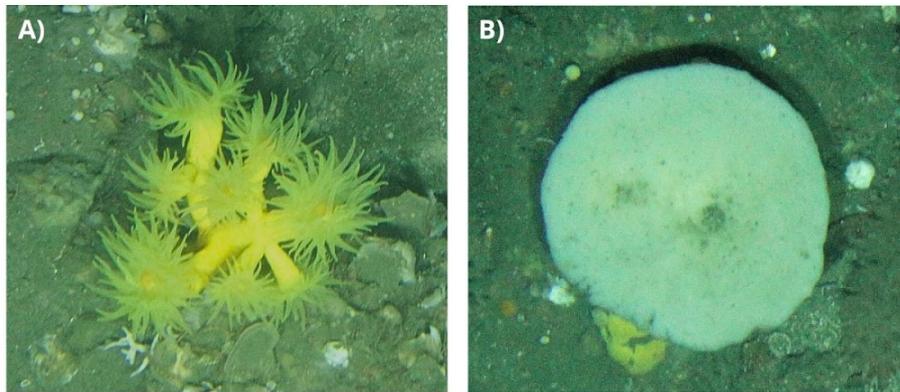


Figura 2. Fotografía detallada de las especies *Dendrophyllia cornigera* (A) y *Phakellia ventilabrum* (B).

Imágenes submarinas ROTV Politolana

Se emplearon imágenes submarinas de alta resolución obtenidas utilizando el vehículo remolcado operado a distancia (ROTV) Politolana (Sánchez & Rodríguez, 2013). El ROTV Politolana, diseñado por el Centro Oceanográfico de Santander del Instituto Español de Oceanografía (IEO-CSIC), es capaz de sumergirse hasta una profundidad máxima de 2000 m. Para el estudio del fondo marino, el ROTV Politolana utiliza métodos fotogramétricos y está equipado con una cámara de alta resolución, telemetría bidireccional y un sistema de posicionamiento acústico. Este vehículo está equipado con cuatro punteros láser acoplados a una distancia precisa de 25 cm entre ellos. Esta configuración permite realizar mediciones con precisión y obtener datos detallados durante cada despliegue. Además, el vehículo adquiere imágenes y videos en alta definición que se sincronizan con los datos ambientales, lo que garantiza la obtención de conjuntos de datos completos en cada inmersión.

En total, en este trabajo se han utilizado 20 transectos, con un área promedio de 285 m² y una longitud media de 410 m por transecto. En el SCSA, las imágenes fueron adquiridas durante la campaña oceanográfica INTEMARES A4 Avilés (2017). Por otra parte, en el SCSC, se llevaron a cabo transectos fotográficos durante las campañas INTEMARES-Capbreton 0619 y 0620 (2019 y 2020). Estos transectos, tanto en el SCSA como en SCSC, se llevaron a cabo en un rango de profundidad de entre 90 y 300 m.

El ROTV Politolana captura fotografías utilizando un intervalo de tiempo de entre 0,5 s y 20 s, según la configuración escogida para el muestreo, lo que proporciona datos

representativos del hábitat y las comunidades bentónicas a caracterizar. Estas imágenes permiten obtener una visión completa del fondo marino en alta resolución (figura 3). En total, se aportaron 5012 imágenes entre el SCSA y SCSC para este estudio.

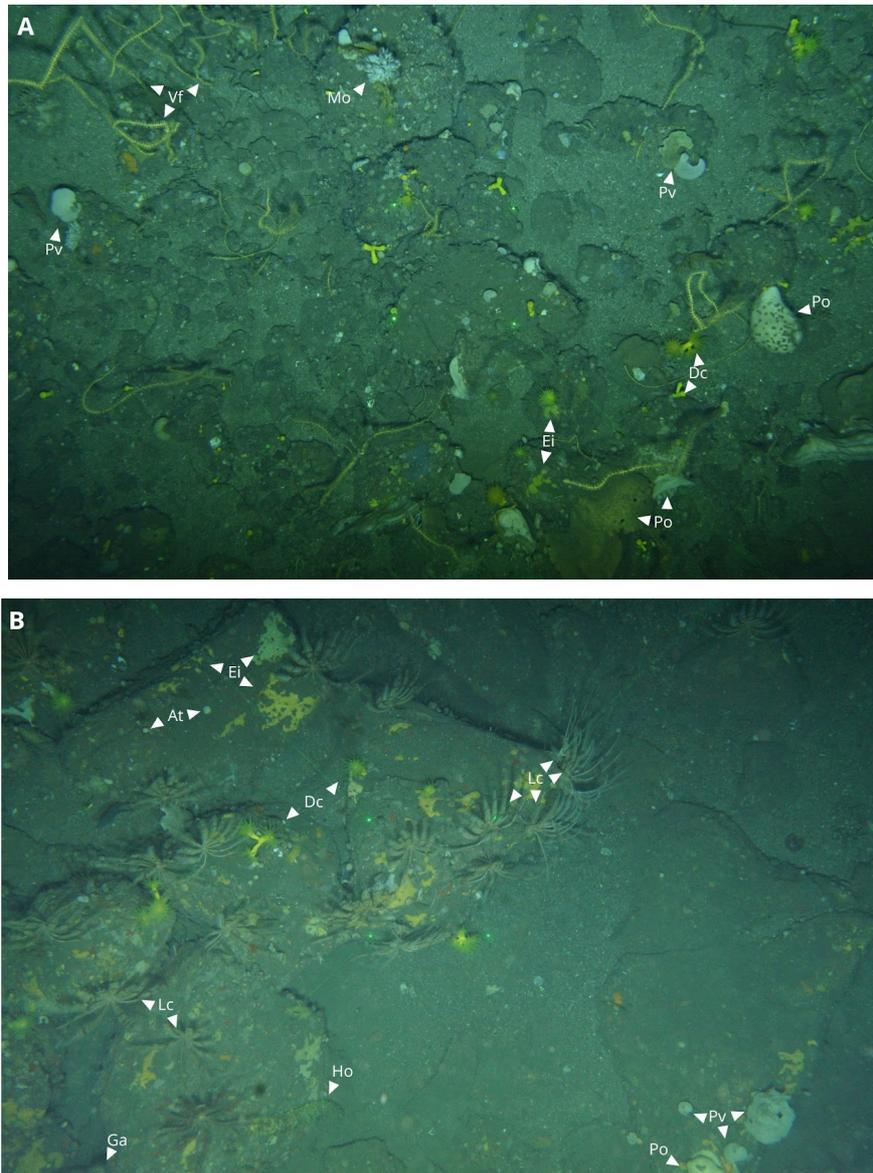


Figura 3: Imágenes obtenidas mediante el uso del ROTV, en las cuales se pueden observar ejemplares dispersos de las especies *Phakellia ventilabrum* (Pv) y *Dendrophyllia cornigera* (Dc) sobre sustrato rocoso. En el Sistema submarino de Capbreton (A), se identificaron esponjas incrustantes (Ei) y otros organismos pertenecientes al filo *Porifera* (Po), así como los corales *Viminella flagellum* (Vf) y *Madrepora oculata* (Mo). En el Sistema submarino de Avilés (B), se observaron esponjas incrustantes (Ei), la esponja *Artemisina transiens* (At) y otros organismos pertenecientes al filo *Porifera* (Po), además de *Leptometra celtica* (Lc), un ejemplar de la familia *Holothuroidea* (Ho) y el erizo *Gracilechinus acutus* (Ga). Los láseres son visibles en el centro de las imágenes.

Preprocesamiento de imágenes y entrenamiento de algoritmos

Para evaluar la generalización del modelo de detección, se seleccionaron aleatoriamente 60 imágenes de un conjunto de 300 imágenes del SCSA como conjunto de validación. Las 240 imágenes restantes se utilizaron como imágenes de entrenamiento en una proporción de 8 a 2. Se aseguró que no hubiera imágenes ni individuos repetidos entre los conjuntos de entrenamiento y validación para evitar el sobreajuste del modelo.

Se utilizó el software de anotación de datos de imagen Supervisely (<https://supervise.ly/>) para dibujar cuadros delimitadores en las especies objetivo *D. cornigera* y *P. ventilabrum*. Las imágenes se etiquetaron asegurando que todo el individuo estuviera representado con la menor área de fondo posible (figura 4).

En cuanto al conjunto de datos de entrenamiento, se realizaron 527 y 1045 anotaciones para las clases *D. cornigera* y *P. ventilabrum*. La anotación se llevó a cabo por científicos expertos entrenados. Para equilibrar el balance de clases, se utilizó la función de objetos voladores de Supervisely (<https://github.com/supervisely-ecosystem/flying-objects>). Esta función genera datos sintéticos para tareas de detección de objetos. Se anotaron ejemplares de ambas clases como máscaras y la aplicación efectuó aumentos en los objetos y los distribuyó en distintos fondos seleccionados. Con la creación de 20 imágenes sintéticas, se lograron equilibrar las clases con 2330 y 2605 anotaciones para las clases *D. cornigera* y *P. ventilabrum*, respectivamente (figura 4).

Además, para evaluar la capacidad de generalización del modelo en un entorno diferente al utilizado para el entrenamiento, se seleccionaron aleatoriamente 60 imágenes de la zona de SCSC, zona donde podemos encontrar nuestras dos especies de interés sometidas a diferentes presiones ambientales. Estas imágenes se utilizaron como imágenes de validación para determinar si el modelo era capaz de detectar correctamente las especies objetivo.

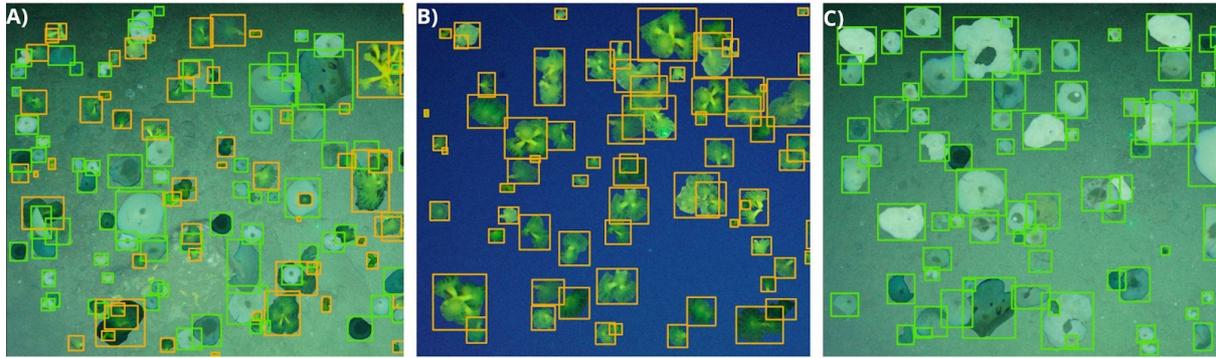


Figura 4. Imágenes sintéticas generadas utilizando el software de anotación de datos de imagen Supervisely y la función de objetos voladores para aumentar el tamaño de la muestra y aplicar aumentos en las especies objetivo *D. cornigera* y *P. ventilabrum*. (A) Ejemplares de ambas especies. (B) Sólo ejemplares de *D. cornigera*. (C) Sólo ejemplares de *P. ventilabrum*. Cada imagen contiene una media de 120 anotaciones.

Todos los modelos de detección de objetos se basaron en los mismos pesos de preentrenamiento del conjunto de datos COCO (Common Objects in Context), que es un conjunto de datos ampliamente utilizado en la investigación de visión por computadora. Los modelos se entrenaron con las imágenes seleccionadas del SCSA, tanto con y sin aumento de datos, durante 200 épocas para mejorar su precisión y rendimiento. Durante cada época, los modelos procesaron todo el conjunto de datos de entrenamiento una vez y ajustaron sus parámetros para minimizar el error de predicción. Esto permitió que los modelos mejoraran continuamente su capacidad para detectar correctamente las especies objetivo.

Etiquetado automático de especies en imágenes submarinas utilizando aprendizaje profundo

Para la anotación automática de nuestro conjunto de imágenes, se desarrolló un marco basado en el aprendizaje profundo en el que se consideraron tres arquitecturas diferentes de redes neuronales profundas. Una red neuronal consta de capas de entrada que reciben los datos, un núcleo de procesamiento con capas ocultas y capas de salida que proporcionan el resultado del modelo. El término "profundo" hace referencia al número de capas ocultas en la estructura de la red neuronal. Una red neuronal se entrena para producir el resultado deseado ajustando sus parámetros internos, llamados pesos, en función del error entre la salida del modelo y la respuesta correcta. Este ajuste se realiza mediante un proceso llamado descenso del gradiente (Schmidhuber, 2014).

En este estudio, se compararon tres modelos de detección de objetos con diferentes arquitecturas neuronales: YOLOv7 y YOLOv8 (Bochkovskiy et al., 2022; Karmaker et al., 2023), modelos de una sola etapa, y Faster R-CNN (Ren et al., 2017), un modelo de dos etapas. Los modelos de una etapa YOLO han demostrado tener ventajas en comparación con Faster R-CNN (Abdulghani & Menekşe Dalveren 2022; Maity et al., 2021; Zheng et al., 2022), por lo que se incluyeron las dos últimas versiones del modelo en la comparación.

YOLOv7 y YOLOv8 son modelos de detección de objetos en tiempo real que emplean redes neuronales convolucionales (CNN) para identificar y localizar objetos en imágenes de manera eficiente. YOLOv7 utiliza la arquitectura ELAN, que mejora la capacidad de aprendizaje y convergencia de las redes profundas. Por otro lado, YOLOv8 integra avances en aprendizaje profundo y visión por computadora, como la adopción de estructuras de atención y el uso de bloques de convolución de dilatación, lo que resulta en una mayor velocidad y precisión en la detección de objetos. En este estudio, se utilizó el modelo YOLOv8x para el entrenamiento de YOLOv8, mientras que se utilizó el modelo YOLOv7-E6E para el entrenamiento de YOLOv7.

Faster R-CNN es un modelo de detección de objetos de dos etapas que combina una CNN para extraer características y una Red de Propuestas de Región (RPN) para generar propuestas de alta calidad. La RPN predice los límites y puntuaciones de objetividad en cada posición de la imagen y se entrena de extremo a extremo. Estas propuestas son luego utilizadas por Faster R-CNN en la etapa de detección y clasificación de objetos (Ren et al., 2017). En este estudio, se utilizó el modelo X-101-32x8d para el entrenamiento de Faster R-CNN.

Las versiones YOLOv8x, YOLOv7-E6E y Faster R-CNN X-101-32x8d fueron elegidas para este estudio debido a su rendimiento en la detección de objetos. Según nuestros datos y la información disponible en los repositorios de GitHub, estas versiones presentan un alto rendimiento en términos de precisión y velocidad en la detección de objetos.

Además, para el entrenamiento de los modelos de detección de objetos, se utilizó Google Colab Pro, una plataforma que nos proporcionó la GPU NVIDIA A100-SXM, lo que permitió procesar grandes cantidades de datos de manera eficiente y acelerar el entrenamiento de los modelos. Google Colab también permitió compartir fácilmente el código y los resultados del estudio con los miembros del equipo.

Selección del modelo de detección de objetos

Para evaluar el desempeño de los diferentes modelos en la tarea de detectar *D. cornigera* y *P. Ventilabrum* en la zona de plataforma del SCSA y SCSC, se seleccionaron tres métricas ampliamente utilizadas para tareas de detección de objetos. En concreto, se compararon y analizaron las métricas de precisión, recuperación y F1 Score para los modelos de detección de objetos YOLOv8, YOLOv7 y Faster R-CNN. La precisión mide qué tan precisas son las predicciones del modelo, es decir, el porcentaje de predicciones que son correctas. La recuperación (o la sensibilidad de un clasificador) mide qué tan bien el modelo encuentra todos los aspectos positivos, es decir, cuántos de los positivos reales captura nuestro modelo etiquetándolos como positivos. Por último, la puntuación F1 es un índice que evalúa el equilibrio entre precisión y recuperación, muy utilizado en deep learning para comparar el rendimiento de dos modelos en la misma tarea. La precisión, recuperación y F1 se calculan con las siguientes ecuaciones (1), (2) y (3) (Van Rijsbergen, 1974):

$$precisión = \frac{Verdaderos\ Positivos}{Verdaderos\ Positivos + Falsos\ Positivos} \quad (1)$$

$$recuperación = \frac{Verdaderos\ Positivos}{Verdaderos\ Positivos + Falsos\ Negativos} \quad (2)$$

$$F1 = 2 \cdot \frac{precisión \cdot recuperación}{precisión + recuperación} \quad (3)$$

Se utilizó una intersección de unión (IoU) de 0,5 para la evaluación de los modelos de detección de objetos. IoU es una métrica que mide la superposición entre la detección y el objeto real, calculada como la intersección entre ambos dividida por su unión (figura 5). Un IoU de 0,5 significa que el 50% del área del objeto real está cubierta por la detección. Además, se midieron las métricas para dos umbrales de confianza distintos (0,5 y 0,6), lo cual es importante para evaluar la detección del modelo en diferentes niveles de certeza.

Los conjuntos de datos de entrenamiento de Avilés, con y sin aumentos, se utilizaron para los conjuntos de datos de validación de Avilés y Capbreton. El conjunto de datos de Capbreton se utilizó para evaluar la capacidad de los modelos para generalizar en otros sistemas de cañones submarinos sometidos a diferentes presiones y condiciones ambientales.

Los modelos se compararon con los distintos entrenamientos para dos conjuntos de datos de validación: uno del SCSA y otro del conjunto de datos del SCSC. La evaluación exhaustiva de los modelos en diferentes escenarios y condiciones es fundamental para obtener resultados precisos y confiables en la detección de objetos en sistemas submarinos.

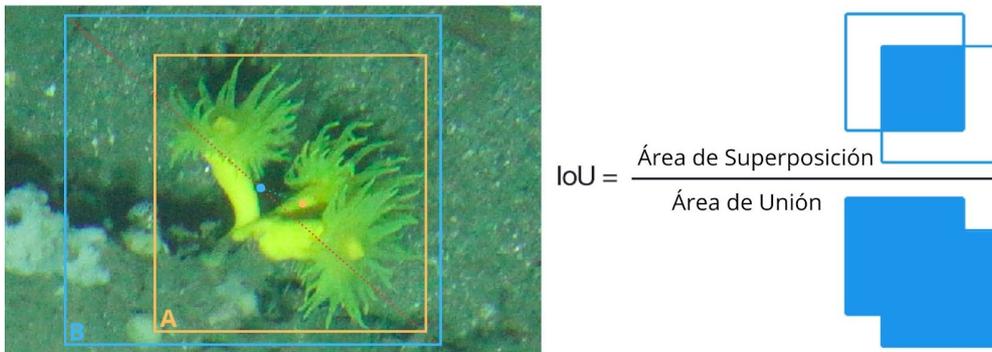


Figura 5. La imagen muestra el cálculo de la intersección de unión (IoU) para la detección de *D. cornigera*. En naranja (A) se muestra el cuadro delimitador de la anotación y en azul (B) el de la inferencia, así como el centro de cada caja y la distancia entre ambos puntos. El IoU se calcula como la intersección entre ambas cajas dividida por su unión

Flujo de trabajo continuo para la generación mapas de densidad de especie

Se implementó un flujo de trabajo continuo en Google Colab Pro para automatizar el proceso de generación de mapas de densidad de especie a partir de imágenes de transectos sin procesar (figura 6). Las imágenes se sincronizaron con los datos de telemetría del ROTV, lo que permitió obtener información de la profundidad, las coordenadas y la altura del ROTV respecto al lecho marino para cada imagen de transecto. Además, con el programa ImageJ (versión 1.53o) se midió manualmente la distancia entre las marcas de los punteros láser del ROTV en 50 imágenes, para obtener el área que abarcaba cada imagen en función de su resolución. Con estos datos, se entrenó un modelo de regresión sencillo utilizando técnicas de aprendizaje automático (machine learning) para relacionar el área con la altura del ROTV respecto al lecho marino. Las predicciones del modelo se utilizaron para calcular el área del resto de las imágenes.

Paralelamente, se llevaron a cabo predicciones con el modelo YOLOv8, previamente entrenado con nuestros datos (YOLOv8-SCS), para analizar las imágenes capturadas en los transectos. Cabe destacar que, en nuestra comparación de modelos de detección de objetos, YOLOv8 obtuvo las mejores métricas con un IoU de 0,5 y un umbral de confianza del 0,6 (figura 7). Por consiguiente, establecimos estos parámetros para realizar las predicciones y generar las inferencias en nuestro flujo de trabajo. Posteriormente, los resultados obtenidos se combinaron con el área y las coordenadas de cada imagen, con el fin de calcular el número de individuos por metro cuadrado en cada una.

Finalmente, se utilizó el software QGIS (versión 3.22) para generar los mapas a partir de los datos de densidad de especies obtenidos en cada punto de imagen de los transectos. Para este propósito, se aplicó una simbología basada en una graduación del tamaño del símbolo en función de la categoría de valores de densidad (Schmidt et al., 2022). La serie de datos de densidad se clasificó en intervalos según su valor y a cada intervalo se le asignó un tamaño de símbolo, siendo mayor el tamaño para mayores densidades. Este enfoque permitió visualizar claramente las zonas con mayor densidad de especies en los mapas generados.

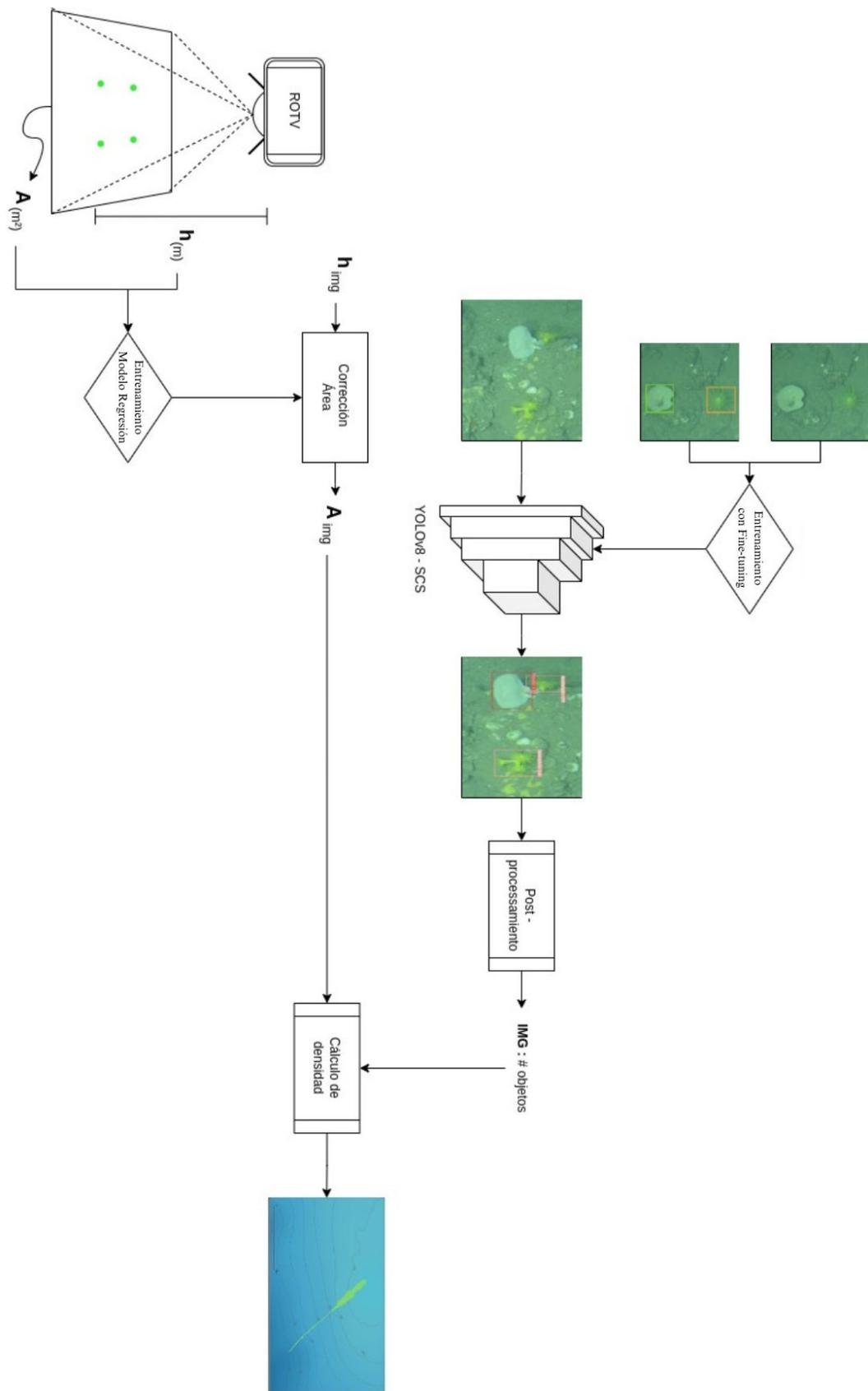


Figura 6. Representación gráfica del flujo de trabajo continuo implementado en Google Colab para la generación automatizada de mapas de densidad de especie a partir de imágenes de transectos sin procesar y datos de telemetría del ROTV.

Resultados

Comparativa de modelos de detección de objetos

En la comparativa entre los modelos de una etapa YOLOv7 y YOLOv8 y el modelo de dos etapas Faster R-CNN, se evaluó la precisión y eficiencia de los modelos al ser entrenados con y sin aumento de datos. Los resultados, presentados en la tabla 1, indican una mejora significativa en el rendimiento de los tres modelos al ser entrenados con el aumento de datos. En particular para la especie *D. cornigera*, los modelos de una etapa mostraron una mayor mejora en su rendimiento. YOLOv8 alcanzó una tasa de recuperación del 18% y un F1 del 11,4%, mientras que YOLOv7 mostró mejoras por encima del 7,7% en todas sus métricas. Para la especie *P. ventilabrum*, el modelo Faster R-CNN obtuvo la mayor diferencia de mejora en términos de recuperación y F1 con un 8,5% y un 5,2% respectivamente. Por otro lado, el modelo YOLOv8 obtuvo diferencias de un 5,7% en términos de precisión.

Tabla 1. Diferencia en el rendimiento de los modelos YOLOv8, YOLOv7 y Faster R-CNN en términos de precisión, recuperación y F1 para las especies *D. cornigera* y *P. ventilabrum* con y sin aumento de datos. Los valores representan la diferencia entre el rendimiento del modelo con aumento de datos menos el rendimiento del modelo sin aumento de datos.

Modelo	Especie	Precisión (%)	recuperación (%)	F1 (%)
YOLOv8	<i>D. cornigera</i>	1,4	18,5	11,4
YOLOv7	<i>D. cornigera</i>	7,7	10,4	9,6
Faster R-CNN	<i>D. cornigera</i>	0,4	1,6	1,5
YOLOv8	<i>P. ventilabrum</i>	5,7	3,5	3,3
YOLOv7	<i>P. ventilabrum</i>	0,5	3,1	2,3
Faster R-CNN	<i>P. ventilabrum</i>	3,5	8,5	5,2

Según los resultados presentados en la figura 7, el modelo YOLOv8 tiene el mejor rendimiento en términos de precisión, recuperación y F1 para ambos umbrales de confianza 0,5 y 0,6. En el SCSA, para *P. ventilabrum*, el modelo YOLOv8 obtiene valores superiores al 92,3% para un umbral de 0,6. Para *D. cornigera* en el mismo sistema, el modelo YOLOv8 obtuvo unos resultados muy superiores al resto de modelos con una precisión del 92,4%, un

recuperación del 91,0% y un F1 del 91.7% para un umbral de 0,6. Además, YOLOv8 también ha sido el modelo que mejor ha generalizado al obtener las mejores métricas para ambas especies en el SCSC.

En general, el rendimiento de los tres modelos mejora al aumentar el umbral de 0,5 a 0,6 en términos de precisión y F1. Estos resultados indican que el modelo YOLOv8 es el más adecuado para la tarea en cuestión y que un umbral de 0,6 puede proporcionar un mejor rendimiento en términos de precisión y F1.

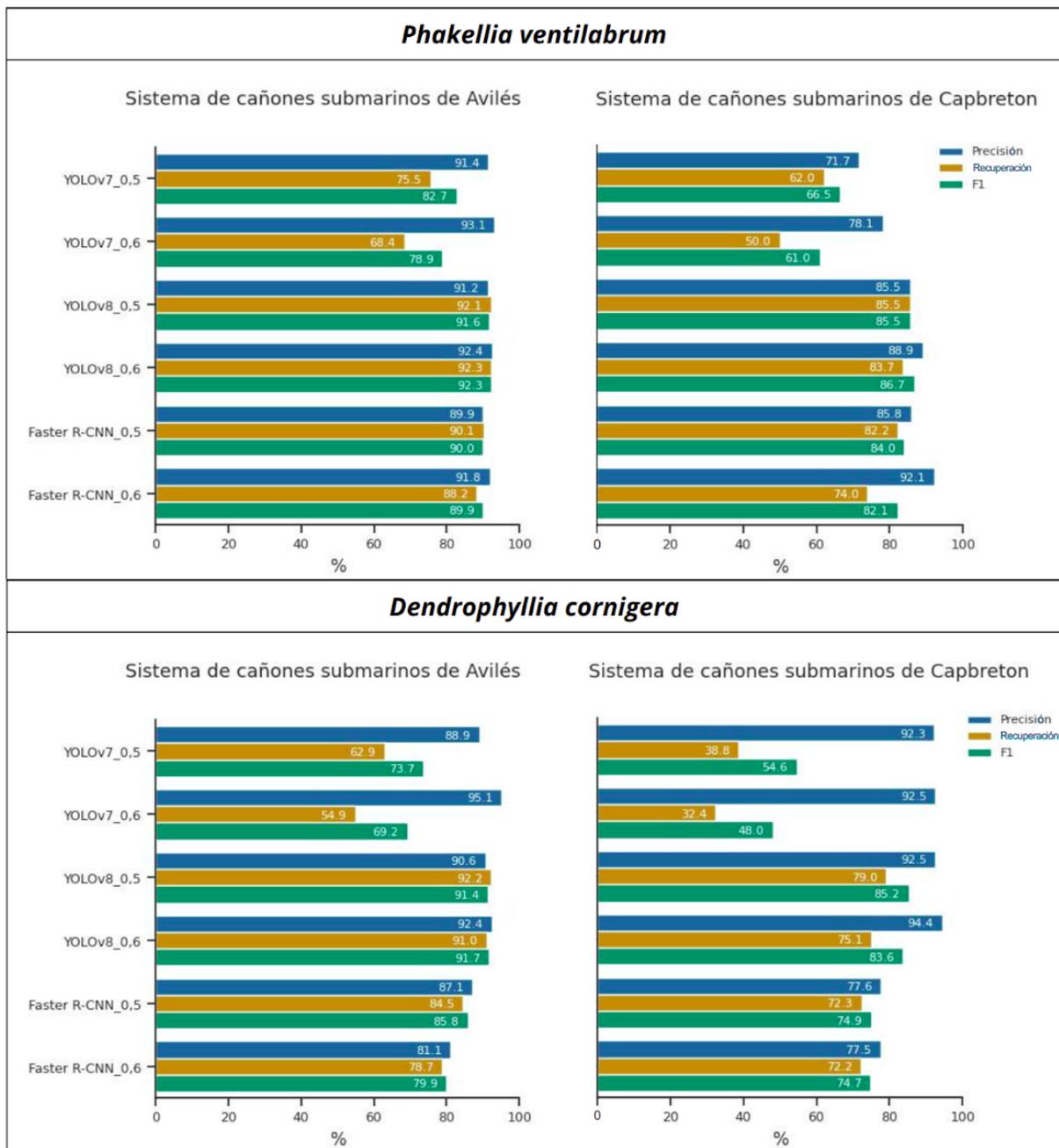


Figura 7. Comparación de métricas para la detección de las especies *P. ventilabrum* y *D. cornigera* utilizando los modelos YOLOv8, YOLOv7 y Faster R-CNN en los SCSCA y SCSC. Se analizaron las métricas de precisión, recuperación y F1 utilizando datos aumentados y unos umbrales de confianza del 0,5 y 0,6.

La Figura 8 presenta los resultados de detección parcial obtenidos mediante el uso de los modelos YOLOv8, Faster R-CNN y YOLOv7. Se estableció un umbral de confianza del detector en 0,6 y un umbral de IoU en 0,5. El primer conjunto (1) de imágenes (figura 8A) comparativas indica que los tres modelos clasifican erróneamente algunas esponjas que no son *P. ventilabrum* como si lo fueran. Faster R-CNN presenta una mayor cantidad de falsas

detecciones para *P. ventilabrum*, confundiendo con otras esponjas y restos de conchíferos e incluso con *D. cornigera* al anotar esponjas amarillas incrustantes como si fueran esta especie. Por otro lado, los tres modelos no detectan la totalidad de las *D. cornigera* presentes en las imágenes. Se observa una menor detección en ejemplares simples que en aquellos que forman una comunidad compleja. YOLOv7 presenta una mayor cantidad de ejemplares de *D. cornigera* no detectados, dejando sin detectar más del 60% de los ejemplares presentes en las imágenes. Sin embargo, cabe destacar la capacidad de detección de YOLOv8 con altas densidades de ejemplares de menor tamaño de *D. cornigera*.

El segundo conjunto de imágenes (figura 8B) comparativas indica que Faster R-CNN clasifica como *D. cornigera* y *P. ventilabrum* una esponja compleja de gran tamaño y también comete errores al identificar dos ejemplares de esponjas incrustantes como *D. cornigera*. Por otro lado, como se puede observar, los modelos basados en YOLO no han sido capaces de detectar los ejemplares de *P. ventilabrum* en su totalidad.

En general, los resultados indican que YOLOv8 presenta una mayor eficacia en la detección y una mayor precisión en el conteo en comparación con el resto de modelos evaluados. El algoritmo demuestra un mejor desempeño para la detección de estas especies estructurantes del bentos marino.

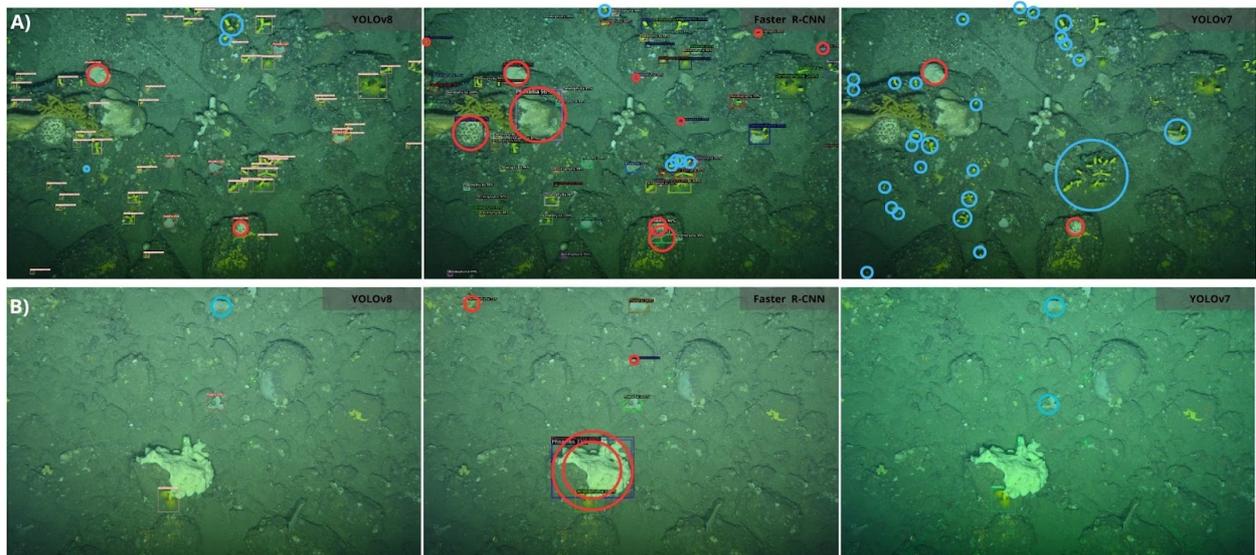


Figura 8. Comparación de los resultados de detección de los diferentes algoritmos para imágenes obtenidas del SCSC (A y B). Los círculos rojos representan una detección falsa y los círculos azules representan una detección omitida. En el marco derecho de la imagen aparece el modelo de detección utilizado. Esta figura muestra cómo los diferentes modelos se desempeñan en la detección de las especies *D. cornigera* y *P. ventilabrum*.

Estudio de densidad de especies

Se procesaron un total de 5021 imágenes de transectos a través de flujo de trabajo continuo creado para la detección automática de las especies objetivo. Los transectos cubrieron un área de 5647,48 m², con una superficie media cubierta de 282,37 m². Como resultado, se obtuvieron 27668 anotaciones automáticas generadas por el modelo YOLOv8-SCS, de las cuales 6087 corresponden a la especie *P. ventilabrum* y 21581 a la especie *D. cornigera*. Esto resultó en una densidad media de 1,01 individuos/m² de *P. ventilabrum* y 3,07 individuos/m² de *D. cornigera* en el SCSA y de 1,18 individuos/m² de *P. ventilabrum* y 4.98 individuos/m² de *D. cornigera* en el SCSC. La información generada permitió cartografiar las densidades de estas especies y comparar las dos zonas de estudio (Figura 9).

La plataforma rocosa del SCSC presentó las mayores densidades para ambas especies. En particular, se observó una densidad máxima de 60,56 individuos/m² para *D. cornigera* y 12,96 individuos/m² para *P. ventilabrum*. Estos datos de densidades máximas se encontraron en el mismo transecto a una profundidad media de 160,08 m y donde se registró una temperatura del agua de 11,72 °C. Con estos datos se generó un mapa de densidad de especie que muestra los puntos con mayores densidades de estas especies (Figura 9).

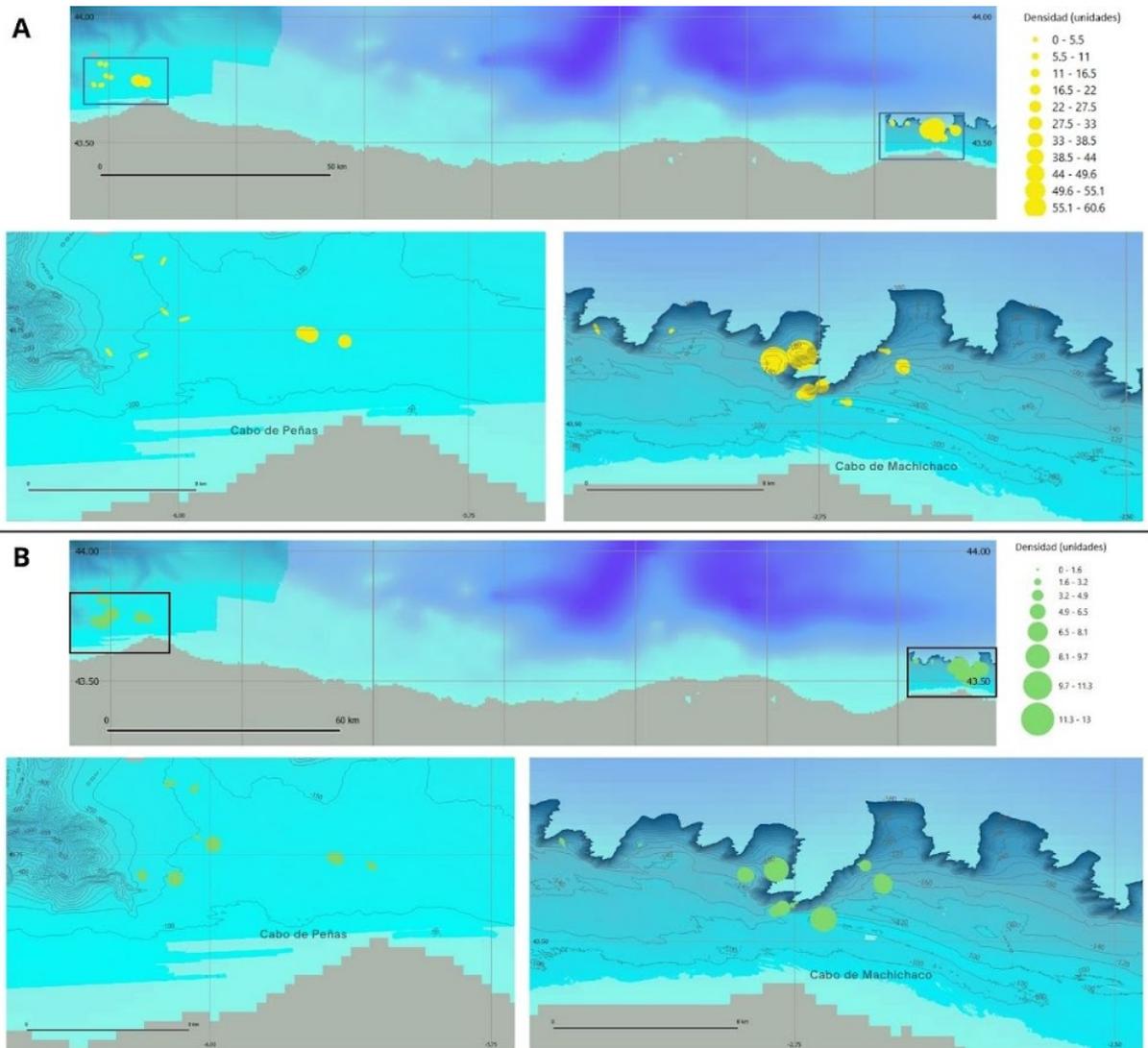


Figura 9. Mapas de densidad de especies para las especies *Dendrophyllia cornigera* (A) y *Phakellia ventilabrum* en el Sistema de Cañones Submarinos de Avilés (abajo a la izquierda) y en el Sistema de Cañones Submarinos de Capbreton (abajo a la derecha). En la parte superior se muestra un mapa de situación a gran escala que representa la ubicación de ambos sistemas de cañones. La tabla de atributos indica la simbología utilizada para representar los diferentes intervalos de densidad de las especies.

Discusión

El presente estudio aborda la necesidad de aportar soluciones eficientes para el seguimiento, protección y conservación de los ecosistemas marinos vulnerables (EMV) de la plataforma circalitoral, mediante la implementación de algoritmos de identificación automática y el cálculo de densidades sin intervención humana. Estos ecosistemas son cruciales para mantener la biodiversidad marina y desempeñan un papel vital en la provisión de servicios ecosistémicos esenciales (Ríos et al., 2022). Para ello, hemos utilizado herramientas de aprendizaje profundo para ayudar en la identificación y mapeo de los EMV de dos sistemas de

cañones submarinos, los SCSA y SCSC. En particular, hemos elegido YOLOv8 como el algoritmo más adecuado para esta tarea debido a su capacidad para procesar información de manera eficaz y su arquitectura sofisticada que incluye funciones de pérdida avanzadas (Lin, 2023). Los resultados obtenidos proporcionan información útil para la gestión de áreas Natura 2000, aumentando el conocimiento sobre los hábitats de plataforma circalitoral rocosa y la distribución de especies vulnerables como *D. cornigera* y *P. ventilabrum* en el área de estudio.

Según nuestros resultados, YOLOv8 ha demostrado ser el modelo más eficaz en la detección de nuestras dos especies de estudio: *D. cornigera* y *P. ventilabrum*. En particular, ha logrado alcanzar valores de F1 superiores al 91,7% para ambas especies en el SCSA, lo que indica un alto nivel de precisión y recuperación en la detección. A pesar de las notables diferencias morfológicas entre ejemplares de la misma especie, como la diversidad en formas, tamaños y complejidad de las colonias de *D. cornigera*, y la variabilidad en tamaños y las formas de *P. ventilabrum*, YOLOv8 ha logrado detectar ambas especies con precisiones superiores al 92,4% en el SCSA. Estos resultados respaldan el estudio realizado por Li et al. (2023), que concluye que YOLOv8 es un modelo adecuado para condiciones complejas, mostrando una notable universalidad y robustez en la detección de objetos en imágenes con variabilidad y ruido.

El éxito en la detección de nuestras especies también puede atribuirse al entrenamiento previo de los modelos en conjuntos de datos a gran escala, como COCO (Lin et al., 2014), y al aumento de datos mediante imágenes sintéticas. Este enfoque ha sido beneficioso para abordar el desbalance de clases y ha mejorado nuestras métricas. En particular, observamos una mejora del 11,4% en el F1, al incluir las imágenes sintéticas, para la clase *D. cornigera* con YOLOv8. Además, YOLOv8 ha demostrado una excelente capacidad de generalización al obtener buenos resultados en las predicciones en el SCSC. Sin embargo, se observó que YOLOv7 no logró generalizar adecuadamente, obteniendo métricas inferiores al cambiar al SCSC. Por lo tanto, la validación con imágenes de Capbreton (SCSC) fue crucial para evaluar el sobreajuste de los modelos y su impacto en las métricas.

La comparación entre los SCSA y SCSC ha proporcionado una oportunidad única para evaluar la capacidad de generalización de los modelos y detectar posibles problemas de sobreajuste. A nivel ecológico y de gestión es muy importante contar con este tipo de evaluación sobre las diferencias en las densidades de especies estructurantes en estas zonas.

Este trabajo ha permitido obtener una cartografía de detalle de las variaciones en las densidades de las especies seleccionadas. Y puede ser la base para una posterior determinación de las causas de dichas variaciones de densidad, que podrían ser atribuidas a diferencias en las condiciones ambientales de las zonas de estudio, a diferentes niveles de presión antropogénica o probablemente a una combinación de ambas.

Según los resultados obtenidos, encontramos densidades superiores en el SCSC para *D. cornigera* y *P. ventilabrum*, en comparación con el SCSA. Además, encontramos esponjas blancas incrustantes y otros tipos de esponjas que se podían confundir fácilmente con *P. ventilabrum*. En el SCSA, por otro lado, se observan comunidades coralinas complejas pero menos numerosas, así como un mayor número de esponjas incrustantes y menor de esponjas tridimensionales. Estas diferencias han tenido un impacto en el rendimiento de los modelos en ambos sistemas de estudio, sin embargo, y a pesar de las diferentes características de las zonas, los modelos demostraron una alta precisión y capacidad para detectar objetos en ambas. Cabe destacar que este trabajo se centra en el desarrollo de una herramienta de anotación automática y no en proporcionar una interpretación ecológica detallada de los datos obtenidos por dicha herramienta.

Una vez que se identificó YOLOv8 como el modelo más eficaz para la detección de nuestras especies objetivo, se procedió a automatizar el proceso de obtención de datos ecológicos a partir de imágenes en bruto. La utilización de algoritmos de aprendizaje profundo permite generar mapas de densidad de especies precisos y eficientes (figura 9). El tiempo requerido para realizar todo el etiquetado de imágenes mediante la utilización de estos modelos automáticos se reduce drásticamente en comparación con el tiempo empleado para esta misma tarea por expertos, que manualmente deben buscar y etiquetar cada una de las especies presentes en miles de imágenes. Por lo tanto, estos modelos aportan una gran ventaja, dado que consiguen en un tiempo mínimo información valiosa para la realización de posteriores estudios poblacionales y ecológicos en profundidad, que redunden en la mejora de las medidas de gestión aplicadas en los ecosistemas estudiados.

La presencia de altas densidades de especies pertenecientes al hábitat 1170 Arrecifes, respalda la necesidad de establecer regulaciones y medidas de gestión sostenibles para preservar la biodiversidad en el SCSA y SCSC. Estos ecosistemas están siendo considerados para su protección y conservación a través de su designación como ZEC y LIC, gracias al proyecto

LIFE IP INTERMARES, dicho proyecto se encuentra alineado con el objetivo global de alcanzar el 30 % de áreas marinas protegidas en 2030 bajo el paraguas de la Red Natura 2000. Este estudio demuestra la importancia de utilizar tecnologías avanzadas para estudiar de forma exhaustiva estos ecosistemas complejos y profundos sirviendo como base para el establecimiento de medidas legislativas adecuadas, apoyadas en información científica rigurosa con un alto grado de detalle, para así mantener el equilibrio de la estructura de estos ecosistemas bentónicos, de las relaciones tróficas que soportan, y asegurar la sostenibilidad de las pesquerías que mantienen.

La implementación de un flujo de trabajo continuo como herramienta de gestión y conservación de los SCSCA y SCSC ofrece la oportunidad de monitorear las densidades de las especies de estudio de la zona de plataforma. Además, el flujo de trabajo continuo tiene el potencial de evolucionar y adaptarse a las necesidades de futuros estudios, para obtener datos en tiempo real, abarcar un mayor número de especies y proporcionar información más detallada sobre los tamaños de los organismos detectados. Además, su capacidad de replicación en el tiempo favorecerá el seguimiento de la evolución ambiental de estas zonas o su posible respuesta a las medidas de gestión aplicadas. Varios estudios, como el de Zhong et al. (2022), han demostrado con éxito el uso de modelos de detección de objetos basados en YOLO para la identificación de animales marinos en tiempo real, lo que respalda la validez de nuestro enfoque. Para la extracción de mediciones de las especies hay modelos de segmentación de última generación prometedores como el modelo SAM (Segment Anything Model) que nos permitiría obtener el área que ocupa la especie en la imagen a tiempo real (Kirillov et al., 2023). Para garantizar la precisión de estas medidas, es fundamental establecer un protocolo de actuación para la captura de imágenes con ROV que minimice la posibilidad de errores, sincronizando adecuadamente los datos del ROV con la imagen capturada. Asimismo, también resultaría interesante ampliar el rango de especies detectadas (Li et al., 2022), lo que permitiría ampliar el alcance del estudio y obtener información más detallada sobre la estructura de las comunidades en la zona. Esta herramienta aportaría una perspectiva holística, pudiendo detectar cambios en los patrones de vida marina en términos de densidad, tallas medias y biomásas, para evaluar adecuadamente el impacto de las actividades humanas en estos hábitats. Se facilitaría así la adopción de medidas para proteger y conservar estos ecosistemas marinos únicos.

En relación a posibles mejoras del presente trabajo, se sugiere utilizar un conjunto más amplio y diverso de imágenes de alta calidad para entrenar el modelo, además de implementar técnicas de aumento de datos adicionales. Está comprobado que al aumentar el número de imágenes durante el entrenamiento, se mejora significativamente la capacidad de generalización y precisión del modelo de detección de objetos (Eversberg & Lambrecht, 2021; Zoph et al., 2020).

Conclusión

1. El modelo YOLOv8 resultó eficaz para la detección del coral amarillo *Dendrophyllia cornigera* y la esponja copa *Phakellia ventilabrum*, dos especies clave de la plataforma rocosa del circalitoral del Mar Cantábrico.
2. Se desarrolló una herramienta poderosa y precisa, dentro de un flujo de trabajo continuo, que permite la detección automática de especies objetivo a partir de imágenes de transecto sin procesar de la plataforma circalitoral, a través de vehículos operados remotamente (ROV).
3. Los resultados muestran que los tres modelos (YOLOv7, YOLOv8 y Faster R-CNN) mejoran su rendimiento al ser entrenados con aumento de datos y que YOLOv8 es el modelo que presenta el mejor rendimiento en términos de precisión, recuperación y F1 para ambos umbrales de confianza 0,5 y 0,6.
4. La implementación de esta herramienta en la zona de plataforma del Sistemas de Cañones Submarinos de Avilés (SCSA) y Capbreton (SCSC) permitió realizar un seguimiento de estos ecosistemas marinos vulnerables, con mapas detallados de densidad de las especies objetivo que indicaron que la plataforma rocosa del SCSC presentó las mayores densidades *D. cornigera* y *P. ventilabrum*.
5. La implementación de tecnologías basadas en el aprendizaje profundo constituyen una metodología eficiente y precisa para el muestreo y seguimiento de las poblaciones bentónicas sésiles. Esto resulta esencial para respaldar la protección y conservación de la biodiversidad en estos ecosistemas.

Agradecimientos

Quiero expresar mi más sincero agradecimiento a todas las personas y proyectos que han contribuido a este trabajo. En primer lugar, a Elena Prado, doctora del IEO-CSIC Santander, por ser la persona más importante en este trabajo. Desde el primer momento en que contacté con ella, siempre ha sido una persona muy atenta que me ha apoyado y dirigido en todo momento. Su amabilidad, profesionalidad y dedicación han sido fundamentales para el éxito de este proyecto. También agradezco a Adolfo Cobo, doctor en la Universidad de Cantabria, por su ayuda con los problemas de enfoque del proyecto y con los algoritmos. Agradezco a Sergio Sierra, doctorando de la Universidad de Cantabria, por su ayuda con los modelos. Y a Sabrina Clemente, mi tutora en la ULL, por su valiosa ayuda con la escritura del artículo.

Además, quiero expresar mi agradecimiento a Juan Huguet, mi cuñado y doctor especializado en deep learning. Aunque no ha colaborado directamente en el proyecto, su ayuda y conocimientos en programación han sido fundamentales para mi desarrollo. Juan me inició en el mundo de la inteligencia artificial y la programación, y ha sido un pilar básico en mi vida. Quiero reconocer y agradecer su dedicación y esfuerzo.

Esta investigación se ha realizado en el ámbito de los proyectos Nature + LIFE INDEMARES (07/NAT/E/000732) e INTEMARES, parcialmente financiado por la convocatoria LIFE+ “Naturaleza y Biodiversidad” de la Comisión Europea (LIFE15 IPE ES 012) y coordinado por la Fundación Biodiversidad del Ministerio para la Transición Ecológica de España, con la participación de diferentes instituciones científicas y ONG's.

Gracias a todos por vuestra valiosa contribución.

Bibliografía

Abad-Uribarren, A., Prado, E., Sierra, S., Cobo, A., Rodríguez-Basalo, A., Gómez-Ballesteros, M., & Sánchez, F. (2022). Deep learning-assisted high resolution mapping of vulnerable habitats within the Capbreton Canyon System, Bay of Biscay. *Estuarine, Coastal and Shelf Science*, 275(107957), 107957. <https://doi.org/10.1016/j.ecss.2022.107957>

Abdulghani, A. M. A. G., & Menekşe Dalveren, G. G. (2022). Moving object detection in video with algorithms YOLO and faster R-CNN in different conditions. *European Journal of Science and Technology*. <https://doi.org/10.31590/ejosat.1013049>

Bhandarkar, S. M., Kathirvelu, S., & Hopkinson, B. M. (2022). Object detection in 3D coral ecosystem maps from multiple image sequences. *2022 26th International Conference on Pattern Recognition (ICPR)*.

[BOE.es](https://www.boe.es) (1992). *DOUE-L-1992-81200 Directiva 92/43/CEE del Consejo, de 21 de mayo de 1992, relativa a la conservación de los hábitats naturales y de la fauna y flora silvestres.* (s/f). [Boe.es](https://www.boe.es). Recuperado el 16 de mayo de 2023, de <https://www.boe.es/buscar/doc.php?id=DOUE-L-1992-81200>

Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *ArXiv*, abs/2207.02696.

Dominguez-Carrió, C., Riera, J. L., Robert, K., Zabala, M., Requena, S., Gori, A., Orejas, C., Lo Iacono, C., Estournel, C., Corbera, G., Ambroso, S., Uriz, M. J., López-González, P. J., Sardá, R., & Gili, J.-M. (2022). Diversity, structure and spatial distribution of megabenthic communities in Cap de Creus continental shelf and submarine canyon (NW Mediterranean). *Progress in Oceanography*, 208(102877), 102877. <https://doi.org/10.1016/j.pocean.2022.102877>

Ercilla, G., Casas, D., Estrada, F., Vázquez, J. T., Iglesias, J., García, M., Gómez, M., Acosta, J., Gallart, J., & Maestro-González, A. (2008). Morphosedimentary features and recent depositional architectural model of the Cantabrian continental margin. *Marine Geology*, 247(1–2), 61–83. <https://doi.org/10.1016/j.margeo.2007.08.007>

Eversberg, L., & Lambrecht, J. (2021). Generating images with physics-based rendering for an industrial object detection task: Realism versus domain randomization. *Sensors (Basel, Switzerland)*, 21(23), 7901. <https://doi.org/10.3390/s21237901>

FAO, 2009. International Guidelines for the Management of Deep-Sea Fisheries in the High-Seas. Food and Agriculture Organization of the United Nations, Rome.

flying-objects: Research app to generate synthetic data for detection / segmentation / instance segmentation tasks. (s/f). Github.com. Recuperado el 4 de Agosto de 2023, de <https://github.com/supervisely-ecosystem/flying-objects>

Karmaker, A., & Ray, P. (2023). A Comprehensive Review of YOLO: From YOLOv1 to YOLOv8 and Beyond. ArXiv, abs/2304.00501.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., & Girshick, R. (2023). Segment Anything. <https://doi.org/10.48550/ARXIV.2304.02643>

Knausgård, K. M., Wiklund, A., Sjørdalen, T. K., Halvorsen, K. T., Kleiven, A. R., Jiao, L., & Goodwin, M. (2022). Temperate fish detection and classification: a deep learning based approach. *Applied Intelligence*, 52(6), 6988–7001. <https://doi.org/10.1007/s10489-020-02154-9>

Lamarck, J.-B. M. (1816). Histoire naturelle des animaux sans vertèbres. Tome second. Verdière. 568 pp

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>

Li, J., Xu, W., Deng, L., Xiao, Y., Han, Z., & Zheng, H. (2022). Deep learning for visual recognition and detection of aquatic animals: A review. *Reviews in Aquaculture*. <https://doi.org/10.1111/raq.12726>

Linnaeus, C. (1766). Systema naturae per regna tria naturae: secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. Ed, 533–1327.

Liu, Y., & Wang, S. (2022). Corrigendum to “A quantitative detection algorithm based on improved faster R-CNN for marine benthos”. *Ecological Informatics*, 70(101655), 101655. <https://doi.org/10.1016/j.ecoinf.2022.101655>

Mahmood, A., Bennamoun, M., An, S., Sohel, F., Boussaid, F., Hovey, R., Kendrick, G., & Fisher, R. B. (2017). Deep learning for coral classification. En *Handbook of Neural Computation* (pp. 383–401). Elsevier.

Maity, M., Banerjee, S., & Sinha Chaudhuri, S. (2021). Faster R-CNN and YOLO based Vehicle detection: A Survey. 2021 5th International Conference on Computing Methodologies and Communication (ICCMC).

Olden, J. D., Lawler, J. J., & Poff, N. L. (2008). Machine learning methods without tears: a primer for ecologists. *The Quarterly Review of Biology*, 83(2), 171–193. <https://doi.org/10.1086/587826>

Prado, E., Rodríguez-Basalo, A., Cobo, A., Ríos, P., & Sánchez, F. (2020). 3D fine-scale terrain variables from underwater photogrammetry: A new approach to benthic microhabitat modeling in a circalittoral rocky shelf. *Remote Sensing*, 12(15), 2466. <https://doi.org/10.3390/rs12152466>

Raphael, A., Dubinsky, Z., Iluz, D., & Netanyahu, N. S. (2020). Neural Network recognition of marine benthos and corals. *Diversity*, 12(1), 29. <https://doi.org/10.3390/d12010029>

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>

Ríos, P., Altuna, Á., Frutos, I., Manjón-Cabeza, E., García-Guillén, L., Macías-Ramírez, A., Ibarrola, T. P., Gofas, S., Taboada, S., Souto, J., Álvarez, F., Saiz-Salinas, J. I., Cárdenas, P., Rodríguez-Cabello, C., Lourido, A., Boza, C., Rodríguez-Basalo, A., Prado, E., Abad-Uribarren, A., ... Cristobo, J. (2022). Avilés Canyon System: Increasing the benthic

biodiversity knowledge. *Estuarine, Coastal and Shelf Science*, 274(107924), 107924. <https://doi.org/10.1016/j.ecss.2022.107924>

Rodríguez-Basalo, A., Ríos, P., Arrese, B., Abad-Uribarren, A., Cristobo, J., Ibarrola, T. P., Gómez-Ballesteros, M., Prado, E., & Sánchez, F. (2022). Mapping the habitats of a complex circalittoral rocky shelf in the Cantabrian Sea (south Bay of Biscay). *Estuarine, Coastal and Shelf Science*, 273(107912), 107912. <https://doi.org/10.1016/j.ecss.2022.107912>

Sánchez, F.; Rodríguez, J.M. POLITOLANA, a new low cost towed vehicle designed for the characterization of the deep-sea floor. In Proceedings of the Martech 2013 5th International Workshop on Marine Technology, Girona, Spain, 9–11 October 2013.

Schmidhuber, J. (2014). Deep learning in neural networks: An overview. En arXiv [cs.NE]. <http://arxiv.org/abs/1404.7828>

Schmidt, C., Muñoz, G., Lancaster, L. T., Lessard, J.-P., Marske, K. A., Marshall, K. E., & Garroway, C. J. (2022). Population demography maintains biogeographic boundaries. *Ecology Letters*, 25(8), 1905–1913. <https://doi.org/10.1111/ele.14058>

Siddiqui, S. A., Salman, A., Malik, M. I., Shafait, F., Mian, A., Shortis, M. R., & Harvey, E. S. (2018). Automatic fish species classification in underwater videos: exploiting pre-trained deep neural network models to compensate for limited labelled data. *ICES Journal of Marine Science: Journal Du Conseil*, 75(1), 374–389. <https://doi.org/10.1093/icesjms/fsx109>

Song, P., Li, P., Dai, L., Wang, T., & Chen, Z. (2022). *Boosting R-CNN: Reweighting R-CNN samples by RPN's error for underwater object detection*. <https://doi.org/10.48550/ARXIV.2206.13728>

Supervisely: unified OS for computer vision. (s/f). Supervisely.com. Recuperado el 26 de Abril de 2023, de <https://supervisely.com/>

Van Rijsbergen, C. J. (1974). Foundation of evaluation. *The Journal of Documentation; Devoted to the Recording, Organization and Dissemination of Specialized Knowledge*, 30(4), 365–373. <https://doi.org/10.1108/eb026584>

Weinstein, B. G. (2018). A computer vision for animal ecology. *The Journal of Animal Ecology*, 87(3), 533–545. <https://doi.org/10.1111/1365-2656.12780>

Zhong, J., Li, M., Qin, J., Cui, Y., Yang, K., & Zhang, H. (2022). Real-time marine animal detection using Yolo-based deep learning networks in the coral reef ecosystem. *ISPRS - International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences*, XLVI-3/W1-2022, 301–306. <https://doi.org/10.5194/isprs-archives-xlvi-3-w1-2022-301-2022>

Zoph, B., Cubuk, E. D., Ghiasi, G., Lin, T.-Y., Shlens, J., & Le, Q. V. (2020). Learning data augmentation strategies for object detection. En *Computer Vision – ECCV 2020* (pp. 566–583). Springer International Publishing.