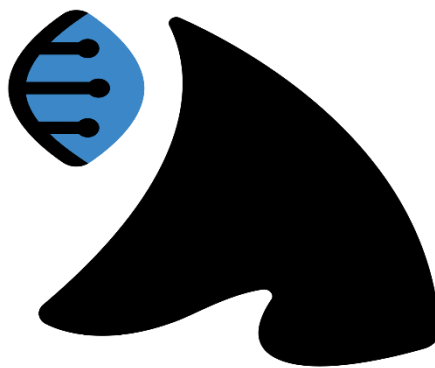


Panorama Actual de Datasets de Elasmobranquios: Revisión Comparativa



e-Lasmobranc

Revisión de Datasets de Elasmobranquios

Este documento presenta una revisión de los datasets existentes sobre especies de elasmobranquios, abarcando tanto aquellos de acceso público y libre disponibilidad como los privados o accesibles bajo solicitud. Se ofrecen breves descripciones de sus características más relevantes, destacando fortalezas y particularidades. Además, se incluye una tabla resumen final que sintetiza los datos clave de cada uno. Esta revisión permite tener una visión general del estado actual de los recursos disponibles, identificar vacíos o limitaciones existentes y orientar futuras mejoras en nuestros propios desarrollos e investigaciones.

Índice de Fuentes de Datos

Datasets Específicos de Elasmobranquios	2
Shark Species Dataset	2
Mozambique Sharks & Rays	2
Sharks and Rays Aquarium Tracking.	3
Aquarium Multi-species Dataset	3
Shark species	4
Shark Labeled Image Dataset	5
Shark Computer Vision Project by Ticon Dataset	5
SharkTrack: an accurate, generalisable software for streamlining shark and ray underwater video análisis	6
MEDLEM database, a data collection on large Elasmobranchs in the Mediterranean and Black seas	6
SharkPulse/Shark Detector	7
Tiburones y Rayas CRAM	8
Sharks and Rays (Andy Murch)	8
Datasets Generales con Elasmobranquios	10
FishNet	10
FathomNet (Plataforma Pública)	10
iNaturalist	11
Tabla Resumen Datasets	12

Datasets Específicos de Elasmobranquios

Shark Species Dataset

Acceso: Dataset abierto disponible en [Roboflow Universe](#).

Tamaño: 1,546 imágenes etiquetadas de tiburones.

Categorías: 16 especies de tiburones, incluyendo ejemplos como Blacktip Shark, Hammerhead Shark, Bull Shark, Whale Shark, Whitetip Shark, Lemon Shark, Great White Shark, Thresher Shark, entre otras.

Anotaciones: Cada imagen cuenta con bounding boxes que identifican la especie de tiburón como clase.

Contexto: Compilado como parte de un proyecto de detección y seguimiento de peces (MSc Fish Detection Tracking). Las imágenes provienen de diversas fuentes de dominio público, en su mayoría tomadas en entornos subacuáticos.

Uso sugerido: Tareas de clasificación y detección automática de tiburones.



Mozambique Sharks & Rays

Acceso: Dataset abierto disponible en [Roboflow Universe](#).

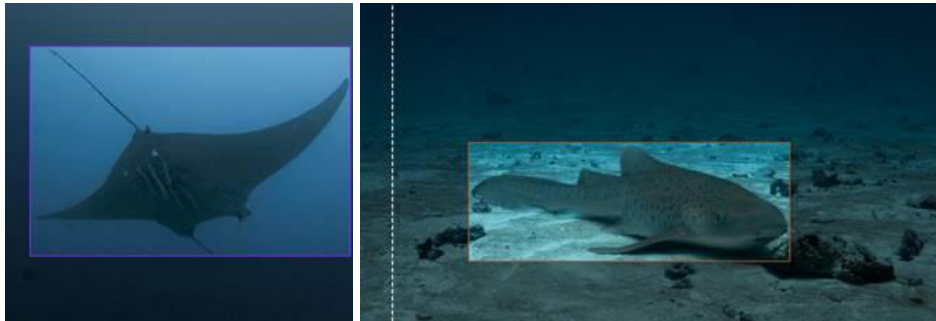
Tamaño: 383 imágenes con anotaciones.

Categorías: 11 especies de tiburones y rayas, además de otros animales marinos en menor proporción. Ejemplos: Manta, Shark, Guitar Fish, Smalleye Stingray, Eagle Ray, Turtle, Whale, Devil Ray, entre otros.

Anotaciones: Las imágenes incluyen *bounding boxes* que identifican tiburones y rayas, clasificadas por especie o al menos por tipo general.

Contexto: Desarrollado en el marco de iniciativas de conservación marina en Mozambique, centrado en un conjunto reducido de especies locales. Las imágenes son principalmente subacuáticas.

Uso sugerido: Detección y clasificación automática de tiburones y rayas en fotografías.



Sharks and Rays Aquarium Tracking.

Acceso: Abierto para descarga en [Dryad](#).

Tamaño: Secuencias de vídeo e imágenes del Oceanário de Lisboa, con decenas de miles de cuadros. Ejemplo: 66 secuencias de entrenamiento de tiburones (15,812 cuadros anotados) y 133 secuencias de entrenamiento de rayas (28,168 cuadros anotados).
Tamaño total aproximado: 53.7 GB.

Categorías: 2 clases principales: tiburón y raya.

Anotaciones: Anotaciones completas de *tracking* y detección, incluyendo *bounding boxes* cuadro a cuadro y IDs de trayectoria para cada individuo.

Contexto: Datos capturados en el tanque principal del Oceanário de Lisboa, diseñados para el desarrollo de métodos automatizados de seguimiento de peces en entornos de acuario.

Uso sugerido: Algoritmos de clasificación, detección y seguimiento de tiburones y rayas.

Aquarium Multi-species Dataset

Acceso: Abierto en [Kaggle](#) (*Aquarium Dataset*) y en [Roboflow Universe](#) (*Aquarium Combined*).

Categorías: 7 clases de animales marinos en acuarios: pez, tiburón, mantarraya (*stingray*), medusa, pingüino, frailecillo y estrella de mar.

Tamaño: 638 imágenes en total, generalmente con múltiples animales por imagen.

Anotaciones: Incluye *bounding boxes* para cada una de las 7 categorías.

Contexto: Dataset compilado a partir de fotografías tomadas en acuarios públicos, principalmente imágenes subacuáticas de fauna marina en entornos controlados.

Uso sugerido: Entrenamiento de modelos para clasificación, conteo y detección multiclase en imágenes subacuáticas.



Shark species

Acceso: Abierto en [Kaggle](https://www.kaggle.com/datasets/ashleythornley/shark-species).

Categorías: 14 especies de tiburones, incluyendo ejemplos como tiburón peregrino, de punta negra, azul, toro, martillo, limón, mako y endermero.

Tamaño: 1,549 imágenes en total.

Anotaciones: Etiquetas simples con el nombre de la especie para cada imagen.

Contexto: Predominantemente imágenes subacuáticas, aunque muchas parecen generadas automáticamente.

Uso sugerido: Entrenamiento de modelos de clasificación para la diferenciación automática de especies de tiburones.



Shark Labeled Image Dataset

Acceso: Disponible en images.cv.

Categorías: Una única clase: tiburón.

Tamaño: 537 imágenes de tiburones, aparentemente recopiladas de diversas fuentes.

Anotaciones: Etiquetas simples por clase (no incluye *bounding boxes* ni segmentación).

Contexto: No especificado; se presume que las imágenes provienen de colecciones mixtas sin un proyecto de fondo concreto.

Uso sugerido: Tareas de clasificación de tiburones en imágenes.



Shark Computer Vision Project by Ticon Dataset

Acceso: Disponible en [Roboflow Universe](https://roboflow.com/dataset/shark-computer-vision-project-by-ticon-dataset).

Categorías: Una única clase que indica la presencia de tiburón (las imágenes sin ejemplar aparecen etiquetadas como *null*).

Tamaño: 1,999 imágenes, de las cuales 1,996 contienen tiburones.

Anotaciones: *Bounding boxes* sobre los individuos de la categoría tiburón; en su mayoría, un cuadro por imagen, con un total de 2,286 anotaciones.

Contexto: No especificado; se presume orientado a tareas de detección de tiburones en entornos subacuáticos.

Uso sugerido: Clasificación y detección de tiburones en imágenes subacuáticas.

SharkTrack: an accurate, generalisable software for streamlining shark and ray underwater video análisis

Acceso: [Disponible bajo demanda](#) (actualmente no responde).

Categorías: Una clase general de elasmobranquios (sin clasificación automática por especie); en las anotaciones manuales se incluye identificación específica de tiburones y rayas.

Tamaño: 6,800 imágenes con *bounding boxes*, provenientes de 77 videos BRUVS obtenidos en 25 ubicaciones globales. El conjunto incluye 1,740 imágenes de rayas y 1,076 de tiburones.

Anotaciones: *Bounding boxes* para cada individuo, con IDs de trayectoria para *tracking*.

Contexto: Dataset generado como parte del desarrollo del software SharkTrack, orientado a la detección, seguimiento y análisis de elasmobranquios principalmente en videos subacuáticos.

Uso sugerido: Detección, clasificación y seguimiento de tiburones y rayas en videos subacuáticos.



MEDLEM database, a data collection on large Elasmobranchs in the Mediterranean and Black seas

Acceso: [Disponible bajo demanda](#).

Categorías: 34 especies y 2 categorías agrupadas de elasmobranquios grandes del Mediterráneo y Mar Negro. Ejemplos destacados: 1,868 individuos de devil ray (*Mobula mobular*), 935 individuos de basking shark (*Cetorhinus maximus*), 622 individuos de blue shark (*Prionace glauca*) y 342 individuos de great white shark (*Carcharodon carcharias*) entre otras.

Tamaño: 3,197 registros documentando 4,963 individuos.

Anotaciones: Datos espaciales y temporales, avistamientos y varamientos, parámetros biológicos (longitud, peso, sexo), artes de pesca, y fuentes diversas.

Contexto: Como parte de una herramienta de monitoreo y conservación, apoyando el seguimiento de especies grandes en el Mediterráneo y Mar Negro.

Uso sugerido: Investigación sobre distribución, biología y estado de conservación de elasmobranquios grandes.

SharkPulse/Shark Detector

Acceso: Visualización disponible a través de su [plataforma web](#). No se ofrece descarga directa de los datos.

Categorías: 611 especies de tiburones y rayas entre las que destacan *Triaenodon obesus*, *Myliobatis tenuicaudatus*, *Trygonoptera imitata*, *Heterodontus portusjacksoni* y *Rhincodon typus*.

Tamaño: Base de datos global con aproximadamente 100,000 avistamientos recopilados de múltiples fuentes dispersas (redes sociales, contribuciones ciudadanas, literatura científica, medios de comunicación, entre otras).

Anotaciones: Imágenes acompañadas de datos temporales, ubicación geográfica e identificación de la fuente original.

Contexto: SharkPulse surge como una plataforma de ciencia ciudadana combinada con inteligencia artificial para llenar vacíos de información sobre la distribución y el estado de conservación de los elasmobranquios, apoyando iniciativas de investigación, manejo y conservación marina a nivel global.

Uso sugerido: Exploración y análisis de datos globales sobre avistamientos de tiburones y rayas, así como generación de conocimiento para apoyar esfuerzos de conservación y monitoreo.



Tiburones y Rayas CRAM

Acceso: Disponible para visualización a través de la página web de la [Fundación CRAM](#).

Categorías: Incluye 13 especies de elasmobranquios documentadas en el contexto del Mediterráneo.

Tamaño: Base de datos con 45 imágenes de diferentes especies de tiburones y rayas.

Anotaciones: Imágenes acompañadas del nombre de la especie identificada.

Contexto: Material publicado por la Fundación CRAM como parte de un proyecto piloto realizado entre 2008 y 2009, enfocado en la conservación y gestión pesquera de elasmobranquios, basado en el análisis de capturas desembarcadas en lonjas catalanas, en colaboración con Obra Social Caja Madrid.

Uso sugerido: Consulta y referencia para iniciativas de conservación, identificación de especies mediterráneas y estudios sobre el impacto de las pesquerías locales.



Sharks and Rays (Andy Murch)

Acceso: Disponible a través de su sitio web oficial [SharksAndRays](#). La plataforma ofrece acceso visual a las imágenes y opciones de adquisición mediante licencias comerciales, educativas o científicas previa solicitud.

Tamaño: La base de datos de imágenes contiene más de 2,000 fotografías en alta resolución de elasmobranquios. Además, su proyecto anterior, [Elasmodiver](#), alberga una colección aún más extensa, con más de 10,000 imágenes que incluyen tiburones, rayas, peces guitarra y quimeras capturados en hábitats marinos de todo el mundo.

Categorías: Cubre 12 órdenes taxonómicos utilizados en sus modelos actuales, abarcando una amplia variedad de especies marinas. Incluye numerosas especies de tiburones y rayas, desde géneros comunes hasta especies de aguas profundas raras.

Anotaciones: Todas las imágenes incluyen identificación de la especie correspondiente, organizada de forma taxonómica y acompañada de una guía de identificación.

Contexto: Proyecto fotográfico y científico orientado a la documentación, identificación y conservación de especies de elasmobranquios, proporcionando recursos visuales de alta calidad para investigación, educación y difusión.

Uso sugerido: Identificación de especies, materiales educativos, apoyo a investigaciones científicas y uso comercial bajo licencia.



Datasets Generales con Elasmobranquios

FishNet

Acceso: Disponible en [GitHub](#).

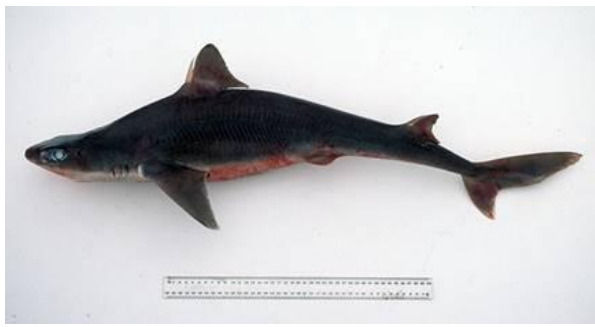
Categorías: 8 clases taxonómicas, 83 órdenes, 463 familias, 3,826 géneros y 17,357 especies de organismos acuáticos; entre ellas, 7 corresponden a rayas y 12 a tiburones (por ejemplo: Dasyatidae, Gymnuridae, Rajidae, Alopiidae, Lamnidae, Carcharhinidae).

Tamaño: 94,532 imágenes organizadas taxonómicamente, incluyendo 1,740 imágenes de rayas y 1,076 de tiburones.

Anotaciones: *Bounding boxes* y 22 rasgos funcionales agrupados en tres categorías: hábitat, rol ecológico y valor nutricional.

Contexto: Diseñado para apoyar la investigación en ecología acuática y contribuir a la conservación de ecosistemas marinos. Las imágenes provienen de fuentes diversas, incluyendo fotos subacuáticas, ilustraciones y registros de pesca.

Uso sugerido: Clasificación, detección de peces y predicción de rasgos funcionales en diferentes especies.



FathomNet (Plataforma Pública)

Acceso: Disponible públicamente a través de [FathomNet.org](#) (requiere registro para descarga masiva de datos).

Categorías: Más de 2,200 conceptos taxonómicos marinos, incluyendo peces, invertebrados, medusas, corales, esponjas, cefalópodos, así como objetos no biológicos (equipos submarinos, desechos marinos). Entre estos se incluyen tiburones y rayas aunque en una primera búsqueda, no se han identificado especies concretas de elasmobranquios.

Tamaño: Más de 175,000 imágenes subacuáticas.

Anotaciones: En su gran mayoría *Bounding boxes* e identificaciones taxonómicas expertas.

Contexto: FathomNet surge para resolver las limitaciones del aprendizaje automático en oceanografía, como la falta de estandarización de datasets, escasez de conjuntos de imágenes abiertas y falta de agregación de datos curados.

Uso sugerido: Entrenamiento y aplicación de modelos de aprendizaje automático para clasificación y detección de organismos marinos, monitoreo automatizado de biodiversidad, navegación y análisis de datos multifuente, así como soporte para la investigación científica y conservación oceánica.

iNaturalist

Acceso: Visualización y descarga parcial públicamente accesible en su página web [inaturalist](https://www.inaturalist.org). No se permite descarga total del conjunto o subcategorías, pero existen datasets preparados para competiciones como [iNaturalist Dataset 2021](https://www.inaturalist.org/competitions) (una cantidad mucho más reducida de especies e imágenes).

Categorías: Miles de especies de fauna y flora global, incluyendo un total de 518,576 especies, de entre las cuales 731 corresponden con elasmobranquios.

Tamaño: Más de 200 millones de observaciones totales y específicamente, 142,265 registros de elasmobranquios, provenientes de contribuciones ciudadanas en todo el mundo, acompañados de fotos y localización georreferenciada.

Anotaciones: Imágenes acompañadas de datos temporales, ubicación geográfica, identificador de especie y categoría taxonómica superior, con posibles detalles adicionales como etapa de vida, sexo y estado del organismo.

Contexto: iNaturalist es una de las plataformas de ciencia ciudadana más grandes del mundo, orientada a documentar biodiversidad global a través de contribuciones fotográficas georreferenciada.

Uso sugerido: Mapeo de biodiversidad, análisis geográfico, clasificación de especies y apoyo a conservación y monitoreo.



Tabla Resumen Datasets

Dataset	Tamaño	Categorías	Anotaciones	Acceso público/directo
Shark Species Dataset	1,546 imágenes	16 especies tiburones	Bounding boxes	Sí (Roboflow)
Mozambique Sharks & Rays	383 imágenes	11 especies tiburones/rayas	Bounding boxes	Sí (Roboflow)
Aquarium Tracking (Lisboa)	~53.7 GB video (imagenes)	2 clases (tiburón, raya)	Bounding boxes + tracking	Sí (Dryad)
Aquarium Multi-species Dataset	638 imágenes	7 clases marinas	Bounding boxes	Sí (Kaggle, Roboflow)
Shark Species (Kaggle)	1,549 imágenes	14 especies tiburones	Etiqueta simple	Sí (Kaggle)
Shark Labeled Image Dataset	537 imágenes	1 clase (tiburón)	Etiqueta simple	Sí (images.cv)
Shark Computer Vision (Ticon)	1,999 imágenes	1 clase (tiburón)	Bounding boxes	Sí (Roboflow)
SharkTrack	6,800 imágenes	2 clases (tiburón, raya)	Bounding boxes + tracking	No (bajo demanda)
MEDLEM	3,197 registros (4,963 individuos)	34 especies + 2 grupos	Datos espaciales, biológicos, varamientos	No (bajo demanda)
SharkPulse / Shark Detector	~100,000 imagenes	611 especies tiburones/rayas	Datos temporales / geográficos	No (visualización web)
Tiburones y Rayas CRAM	45 imágenes	13 especies	Etiqueta simple	Parcial (visualización web CRAM)
Sharks and Rays (Andy Murch)	>2,000 imágenes (+10,000 Elasmobranch)	12 órdenes taxonómicos	Etiqueta + guía taxonómica	No (previa solicitud)
FishNet	94,532 imágenes (1,816)	17,357 especies (12 tiburones, 7 rayas)	Bounding boxes + rasgos funcionales	Sí (GitHub)
FathomNet	>175,000 imágenes (-)	>2,200 conceptos marinos	Bounding boxes	No (requiere registro)
iNaturalist	>200 mill. imagenes (142,265)	518,576 especies (731 elasmobranchios)	Datos temporales, geográficos, taxonómicos	Parcial (visual, competiciones limitadas)

Este proyecto se desarrolla con la colaboración de la Fundación Biodiversidad del Ministerio para la Transición Ecológica y el Reto Demográfico, a través del Programa Pleamar, y se cofinancia por la Unión Europea por el FEMPA (Fondo Europeo Marítimo, de Pesca y de Acuicultura).

El informe ha sido elaborado por:

Departamento de Tecnología Informática y Computación (dTIC)	Centro de Investigación Marina (CIMAR)
Andrés Fuster Guilló	Francisca Giménez Casalduero
Marcelo Saval Calvo	Irene Antón Linares
Ismael Beviá Ballesteros	Isabel Abel Abellán
Jorge Azorín López	Nieves Aranda Garrido