

La investigación reportada en esta tesis es parte de los programas de investigación del CICESE (Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California).

La investigación fue financiada por el SECIHTI (Secretaría de Ciencia, Humanidades, Tecnología e Innovación).

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México). El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo o titular de los Derechos de Autor.

**Centro de Investigación Científica y de Educación
Superior de Ensenada, Baja California**



**Maestría en Ciencias
en Ecología Marina**

**Automatización en la identificación y seguimiento de
tiburones ballena (*Rhincodon typus*) a partir del uso de drones
y la aplicación de modelos de inteligencia artificial**

Tesis
para cubrir parcialmente los requisitos necesarios para obtener el grado de
Maestro en Ciencias

Presenta:

Paola Judith Delgado García

Ensenada, Baja California, México
2025

Tesis defendida por
Paola Judith Delgado García

y aprobada por el siguiente Comité

Dr. Oscar Sosa Nishizaki
Director de tesis

Dr. Emiliano García Rodríguez
Codirector de tesis

Dra. Luz Erandi Saldaña Ruiz

Dr. Irving Hussein López Nava



Dr. Rafael Andrés Cabral Tena
Coordinador del Posgrado en Ecología Marina

Dra. Ana Denise Re Araujo
Directora de Estudios de Posgrado

Resumen de la tesis que presenta **Paola Judith Delgado García** como requisito parcial para la obtención del grado de Maestro en Ciencias en Ecología Marina.

Automatización en la identificación y seguimiento de tiburones ballena (*Rhincodon typus*) a partir del uso de drones y la aplicación de modelos de inteligencia artificial

Resumen aprobado por:

Dr. Oscar Sosa Nishizaki
Director de tesis

Dr. Emiliano García Rodríguez
Codirector de tesis

El tiburón ballena (*Rhincodon typus*) es una especie migratoria conocida por formar agregaciones temporales en zonas costeras con florecimientos de organismos plantónicos, como es el caso de Bahía de los Ángeles, Baja California, México. Actualmente esta especie está clasificada como vulnerable tanto a nivel nacional como internacional, debido a diversas amenazas, lo que hace crucial su monitoreo. En este contexto, los videos tomados desde drones resultan una aproximación tecnológica prometedora para estudiar a los tiburones en su hábitat natural sin causar mayores perturbaciones. Y para sustraer la información de los videos, los modelos de inteligencia artificial ofrecen una alternativa eficaz para su análisis. Para desarrollar un protocolo de monitoreo con drones de tiburón ballena en Bahía de los Ángeles, durante el mes de octubre de 2023, se realizaron 59 vuelos de dron en Bahía de los Ángeles, incluyendo 42 vuelos desde embarcaciones y 17 desde tierra, logrando documentar 23 avistamientos de tiburones ballena. Esta información fue utilizada para entrenar y probar modelos de inteligencia artificial mediante dos distintos enfoques: DeepLabCut (DLC) y Multi Scale Patch (MSP) que permitiera identificar la presencia de tiburones de una forma videográfica automatizada para ser contabilizados. Ambos enfoques se basan en arquitecturas de redes neuronales convolucionales, incorporando técnicas de aumento de datos y aprendizaje por transferencia para mejorar su desempeño. El modelo más eficaz desarrollado en este estudio fue el MSP-1920, con un desempeño de Puntaje F1 macro de 0.91. A través del análisis estadístico de diferentes variables ambientales, logísticas y el comportamiento del tiburón durante la recolección de datos, se determinó que estas variables no afectan significativamente la identificación videográfica de tiburones ballena por el modelo MSP-1920. Con este modelo se espera proporcionar una herramienta que permita ampliar el conocimiento ecológico de la especie en el área, y mejorar las medidas manejo y conservación en la bahía y en hábitats similares.

Palabras clave: tiburón ballena, identificación videográfica automatizada, seguimiento, inteligencia artificial, drones

Abstract of the thesis presented by **Paola Judith Delgado García** as a partial requirement to obtain the Master of Science degree in Marine Ecology

Automated Identification and Monitoring of Whale Sharks (*Rhincodon typus*) Using Drones and Artificial Intelligence Models

Abstract approved by:

Dr. Oscar Sosa Nishizaki
Thesis Director

Dr. Emiliano García Rodríguez
Thesis Director

The whale shark (*Rhincodon typus*) is a migratory species known to form temporary aggregations in coastal areas with blooms of planktonic organisms, as in Bahía de los Ángeles, Baja California, Mexico. This species is classified as vulnerable nationally and internationally due to various threats, making monitoring crucial. In this context, videos taken from drones are a promising technological approach to studying sharks in their natural habitat without causing significant disturbances. Moreover, to extract information from videos, artificial intelligence models offer an effective alternative for analysis. To develop a drone monitoring protocol for whale sharks in Bahía de los Ángeles, during October 2023, 59 drone flights were carried out in Bahía de los Ángeles, including 42 flights from boats and 17 from land, managing to document 23 whale shark sightings. This information was used to train and test artificial intelligence models using two different approaches, DeepLabCut (DLC) and Multi-Scale Patch (MSP), which allowed the presence of sharks to be identified in an automated videographic way to be counted. Both approaches are based on convolutional neural network architectures, incorporating data augmentation and transfer learning techniques to improve performance. The most effective model developed in this study was the MSP-1920, with a Macro F1 Score performance of 0.91. The statistical analysis of different environmental and logistic variables and the shark's behaviour during data collection determined that these variables do not significantly affect the video identification of whale sharks by the MSP-1920 model. This model is expected to provide a tool that expands the ecological knowledge of the species in the area and improves management and conservation measures in the bay and similar habitats.

Keywords: whale shark, identification, tracking, artificial intelligence, drones

Dedicatoria

A mis padres, Eva y Enrique, por impulsarme a siempre superarme y brindarme todo el apoyo para lograrlo.

A mi familia, su amor incondicional es mi motor para continuar y un lugar seguro para resguardarme ante la adversidad.

A Ricardo y Neji, quienes me salvan la vida todos los días.

Agradecimientos

Al Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California (CICESE) y al posgrado en Ecología Marina por permitirme cursar este posgrado en su institución.

Al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT) ahora Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) por el apoyo económico durante mis estudios de maestría y para la realización de la tesis.

Al Proyecto FORDECYT-PRONACES/139022/2020 “Uso sustentable de la megafauna marina en Áreas Naturales Protegidas: aspectos socioeconómicos y la aplicación de tecnologías emergentes para su monitoreo” por el apoyo en el financiamiento de esta tesis.

A mi grandioso comité de tesis, fue una maravillosa y enriquecedora experiencia. Emiliano, gracias por aportar tu gran experiencia en trabajo de campo e investigación, así como por tu carácter tan sereno y bromista, sin duda me ayudó a tomar las cosas con calma, a disfrutar del proceso. Óscar, gracias por recibirme en tu laboratorio, por ayudarme a ser más organizada y plantear mis ideas en escrito, por tu gran sentido del humor, y por siempre estar al tanto de mi bienestar. Luz, mil gracias por todo el apoyo y contribuciones, fuiste mi primer contacto con el laboratorio y siempre me has inspirado confianza, sobre todo al inicio cuando no sabía cómo iniciar esta investigación. Hussein, gracias por ampliar mi panorama en este trabajo interdisciplinario, por el gran interés que has mostrado en el proyecto desde el inicio y el gran apoyo brindado para que pudiera desarrollar este trabajo desde el ámbito computacional.

Al grupo de monitoreo PEJESAPO, en especial a Vanesa Vázquez, los capis Octavio Díaz ‘El Pelón’ y Octavio Arce, que me apoyaron en la recolección de los datos y me ayudaron a afinar mis ojos en busca de tiburones ballena.

A mis compañeros y amigos del Departamento de Ciencias de la Comunicación. Joan, millones de gracias por tu apoyo en mi incursión a la programación y los modelos de inteligencia artificial, gracias por responder todas mis dudas y ayudarme a desarrollar la parte más compleja de este trabajo. Eliaf, gracias por enseñarme el uso del software DLC y resolver las dudas que surgieran en el camino, siempre admiré tu capacidad por explicar temas tan complejos de forma tan sencilla.

Al laboratorio de ecología pesquera: Carmencita, Zuri, Masao, Rafa, Nadia y mi maravilloso comité Oscar, Emi y Luz. Gracias por la alegría con el que todos van a trabajar, su trato tan cálido y su disponibilidad para apoyar a quien lo necesite.

A mi familia: Eva, Enrique, Frida, Iván, Shakti, Abi, Fernando y Fermín. Gracias por siempre impulsarme a superarme y por apoyarme en todo lo que necesité. Todo lo que soy y he logrado se los debo a ustedes.

A mis amigos, la familia que uno escoge. A Ale, Caro, César, Eli y Jesús, gracias por sus anécdotas, sus palabras de aliento y compañía, por disfrutar de este camino conmigo. A el Chichí, Lalo, Marce y Alejandra, por ser mi familia en bahía, nos divertimos mucho y siempre estuvieron al pie del cañón para apoyarme cuando lo necesitaba. A Jean y Karla, agradezco poder contar con ustedes, cada plática me permitió aligerar la carga debido a su comprensión y las carcajadas que soltamos juntas. A Erika, gracias por tus charlas, tus bromas y por siempre escucharme sin juzgar.

A mi bebecito perruno Neji, llegaste a mí desde que inicié la maestría y me has acompañado en momentos tristes donde no permito que nadie me observe o se acerque. Nuestras caminatas y juegos me permiten desconectar de todo, son el verdadero escape de mi día a día.

A Ricardo, gracias por el apoyo que me has brindado en cada uno de mis proyectos desde hace algunos años. Gracias por ayudarme en el camino de convertirme la mejor versión de mí misma, a ser perseverante y no decaer ante el primer fallo, así como a controlar mejor mi carácter (jaja).

Tabla de contenido

	Página
Resumen en español.....	ii
Resumen en inglés.....	iii
Dedicatoria	iv
Agradecimientos.....	v
Lista de figuras.....	ix
Lista de tablas	xi
Capítulo 1. Introducción	1
1.1 Antecedentes	4
1.2 Justificación	7
1.3 Hipótesis.....	8
1.4 Objetivos	9
1.4.1 Objetivo general	9
1.4.2 Objetivos específicos.....	9
Capítulo 2. Metodología	10
2.1 Área de estudio	10
2.2 Características oceanográficas.....	11
2.3 Recolección de datos.....	12
2.3.1 Descripción de vuelos de dron pilotados desde tierra.....	13
2.3.2 Descripción de vuelos de dron pilotados desde embarcación.....	14
2.3.3 Datos recolectados.....	15
2.4 Entrenamiento de modelos.....	16
2.4.1 Enfoque de entrenamiento mediante software DeepLabCut (DLC)	18
2.4.1.1 Preprocesamiento de datos	18
2.4.1.2 Fase de entrenamiento y prueba	19

2.4.2	Enfoque de entrenamiento a partir de Parches Multi-Escala (MSP)	19
2.4.2.1	Preprocesamiento de datos	19
2.4.2.2	Fase de entrenamiento y prueba	21
2.4.2.3	Procesamiento posterior.....	22
2.4.2.4	Mejora del seguimiento en modelos bajo el enfoque MSP.....	23
2.5	Comparación del rendimiento de los diferentes enfoques	24
2.6	Variables que afectan la efectividad de los modelos.....	28
Capítulo 3.	Resultados.....	31
3.1	Datos recolectados en campo	31
3.2	Análisis y comparación del desempeño de los modelos de identificación y seguimiento del tiburón ballena	32
3.3	Análisis de las variables que afectan la efectividad de los modelos	37
Capítulo 4.	Discusión	40
4.1	Virtudes y oportunidades de mejora del modelo	41
4.2	Comparación del mejor modelo respecto a otros estudios.....	45
4.3	Ventajas y oportunidades de los enfoques implementados: DLC vs MSP	48
4.4	Uso futuro del mejor modelo obtenido en el monitoreo de tiburón ballena.....	49
Capítulo 5.	Conclusiones.....	51
	Literatura citada	52
	Anexo	60

Lista de figuras

Página

Figura 1. Área de estudio: El Rincón - Subzona de Uso Público Tiburón Ballena en la Reserva de la Biosfera Zona Marina Bahía de los Ángeles, Canal de Ballenas y de Salsipuedes.	10
Figura 2. Ubicación de las zonas sobrevoladas con dron. Los polígonos rojos corresponden a los vuelos piloteados desde tierra en distintas playas de la bahía. Los puntos rojos corresponden a los vuelos piloteados desde embarcaciones.....	13
Figura 3. Esquema de vuelos pilotados desde tierra (A) y desde embarcación (B).	14
Figura 4. Bitácora de campo para vuelos piloteados desde tierra (#1) y desde embarcación (#2).	16
Figura 5. Proceso empleado según el enfoque implementado: DLC y MSP.	17
Figura 6. Esqueleto de tiburón ballena compuesto de 11 marcadores para el entrenamiento del enfoque DeepLabCut.	18
Figura 7. Arquitectura ResNet50.....	19
Figura 8. Técnicas de aumento de datos.....	20
Figura 9. Arquitectura DenseNet121.	21
Figura 10. Proceso de prueba del modelo MSP.	22
Figura 11. Distribución de pesos para promedio ponderado según la cercanía a la imagen actual.....	22
Figura 12. Generación de predicción en modelos MSP-192, MSP-192O y MSP-96O.	23
Figura 13. Proceso de delimitación de parches.	24
Figura 14. Generación de coordenadas para cuadro delimitador con base en el enfoque utilizado. ..	25
Figura 15. Ejemplos de la comparación del área predicha contra la verdadera.....	26
Figura 16. Matriz de confusión de los valores reales contra los predichos.	26
Figura 17. Resultados de la recolección de videos. El color <i>naranja</i> representa a los videos grabados mediante el pilotaje de dron desde tierra y el <i>azul</i> desde embarcación. a) Cantidad de tiburones avistados según el tipo de vuelo. b) Cantidad de tiburones avistados con respecto a las semanas de trabajo de campo. c) Tiburones avistados con respecto a la zona y tipo de vuelo. d) Tiburones avistados con respecto al horario y tipo de vuelo.	31
Figura 18. Comparación de modelos mediante la métrica Exactitud.	32
Figura 19. Comparación de modelos mediante la métrica Puntaje F1-macro.	33
Figura 20. Comparación de modelos mediante la métrica Cobertura de Área.	34

Figura 21. Comparación visual de la identificación de tiburones con base en distintos modelos.	35
Figura 22. Gráficos de distribución de los residuales. a) Modelo RL #1, b) Modelo GAM #1.....	39
Figura 23. Comparación de la predicción de los modelos del video 22.2_PlayaBlanca-PM-1 ante un caso fácil y uno difícil.....	60
Figura 24. Comparación de la predicción de los modelos del video 25.1_Tb2 ante un caso fácil y uno difícil.	60
Figura 25. Comparación de la predicción de los modelos del video 25.1_Tb7 ante un caso fácil y uno difícil.	61
Figura 26. Comparación de la predicción de los modelos del video 25.1_Tb9 ante un caso fácil y uno difícil.	61
Figura 27. Comparación de la predicción de los modelos del video 25.1_Tb12 ante un caso fácil y uno difícil.	62
Figura 28. Comparación de la predicción de los modelos del video 25.1_Tb13 ante un caso fácil y uno difícil.	62
Figura 29. Comparación de la predicción de los modelos del video 27.1_Angelitos-AM-2 ante un caso fácil y uno difícil.....	63
Figura 30. Comparación de la predicción de los modelos del video 27.3_Estero-AM-1 ante un caso fácil y uno difícil.	63
Figura 31. Comparación de la predicción de los modelos del video 28.4_Rampa-PM-1 ante un caso fácil y uno difícil.	64

Lista de tablas

	Página
Tabla I. Distribución de imágenes por clase y fase para el modelo MSP.....	20
Tabla II. Nombres de los diferentes modelos predictivos desarrollados.....	23
Tabla III. Clasificación de variables con efecto en el desempeño de los modelos.....	28
Tabla IV. Desempeño de los modelos con base a cada video del conjunto de prueba.	36
Tabla V. Características de los videos analizados.....	37
Tabla VI. Mejores modelos de regresión lineal.....	38
Tabla VII. Mejores Modelos Aditivos Generalizados (GAM).	38

Capítulo 1. Introducción

En las últimas dos décadas, el uso de vehículos no tripulados y sensores remotos, como los *drones*, han sido recurrentemente empleados en estudios poblacionales de especies marinas con distribución en aguas superficiales o en zonas someras. En estos estudios se ha podido determinar la variación espacial de la estructura y diversidad de ensamblajes de megafauna marina como delfines, ballenas, tiburones, rayas y tortugas (Kelaher et al., 2023); estimar densidades de elasmobranquios en zonas costeras someras (Kiszka et al., 2016); cuantificar el impacto de las condiciones ambientales en la densidad de tiburones (Digiacoimo et al., 2020); así como realizar conteos de pinnípedos con imágenes aéreas térmicas en playas con vegetación abundante (Gooday et al., 2018), por mencionar algunos ejemplos. Los *drones* permiten coleccionar datos con una resolución espacial y temporal mayor al muestreo realizado a bordo de embarcaciones, así como por ser sencillos de manipular y presentan una alta rentabilidad para su uso en campo (Hodgson et al., 2013; Linchant et al., 2015). Sin embargo, existen diversos retos a resolver relacionados con la fiabilidad en la interpretación de las imágenes obtenidas por los drones para poder producir resultados robustos (Colefax et al., 2023). Uno de los principales desafíos es el procesamiento manual de la información videográfica recopilada, lo que vuelve al uso de drones un método complicado de implementar (Gray et al., 2019). En busca de resolver este desafío, se ha apostado por la automatización del procesamiento de las imágenes mediante el uso de modelos de inteligencia artificial (IA) (Goodwin et al., 2022).

La principal rama de IA involucrada en el análisis de datos audiovisuales es el aprendizaje profundo (*Deep Learning*). Este enfoque ofrece una ventaja significativa en el desempeño de modelos predictivos en comparación con los métodos estadísticos tradicionales y de aprendizaje por máquina (*machine learning*) debido a su capacidad de manejar grandes volúmenes de datos y extraer características complejas de alta dimensión, facilitado por el incremento de la capacidad de cómputo actual (Guo et al., 2020). El aprendizaje profundo, desarrollado para simular el proceso de análisis y aprendizaje del cerebro humano, se ejecuta mediante algoritmos conocidos como redes neuronales artificiales (ANN). Estas redes se componen principalmente de una capa de entrada, una de salida y varias capas ocultas intermedias. Cada capa oculta contiene un número diferente de neuronas totalmente conectadas entre sí, que extraen características de las capas anteriores y las transmiten a la siguiente mediante operaciones de activación, lo que le introduce no linealidad a la red y le permite aprender y representar patrones complejos (Wu & Feng, 2018).

A partir de este concepto, surgieron las redes neuronales convolucionales (CNN), una variante que permite el análisis de imágenes. Las CNN se basan en la adición de capas convolucionales para extraer características importantes de las imágenes. Estas capas aplican filtros (*kernels*), representados por matrices pequeñas, que se desplazan por la imagen de entrada y, mediante operaciones matemáticas, extraen bordes, texturas y patrones, generando con ellas un mapa de características. Este mapa pasa por capas de agrupamiento (*pooling*) que reducen su dimensionalidad al extraer los valores más importantes de una región del mapa de características, disminuyendo así la cantidad de parámetros y el coste computacional. Posteriormente, la información pasa a la estructura base de una ANN, que incluye capas de activación y capas totalmente conectadas (*Fully Connected*) que toman las características extraídas y realizan las predicciones, obteniendo los resultados finales en la capa de salida (Li et al., 2022).

Aplicando estos modelos en investigación ecológica, la capa de entrada correspondería a los datos del objeto de estudio; las capas ocultas, a la arquitectura neuronal elegida junto con sus operaciones; y la capa de salida, a la predicción según la tarea seleccionada. Las tareas pueden ser de tipo 1) Clasificación y reconocimiento, donde se genera una etiqueta para cada categoría estudiada, como la presencia o ausencia de una especie en la imagen; 2) Detección y localización, donde se determina la posición del objeto en los datos, como la detección de un individuo en una imagen mediante un cuadro delimitador; 3) Segmentación semántica, donde se clasifica cada píxel de una imagen en una categoría específica, como distinguir entre diferentes especies presentes en una imagen; 4) Segmentación de instancias, donde se diferencia entre instancias individuales dentro de la misma categoría, como identificar y diferenciar cada individuo de cada especie en una imagen (Guo et al., 2020); y 5) Seguimiento visual de objetos, donde se implementa un detector de objetos que permite encontrar la correspondencia entre las detecciones y los objetivos ya rastreados como el monitoreo continuo de la trayectoria de un objeto dentro de fotogramas continuos en un video (Krebs et al., 2017).

La capacidad que tiene la IA para automatizar el procesamiento de la información y, en conjunto con el gran alcance espaciotemporal y fácil operación de los drones, vuelven a estas tecnologías altamente viables para su uso en estudios ecológicos como lo son el monitoreo de especies. Siendo vitales para la conservación de especies en bajo alguna categoría de riesgo como el tiburón ballena.

El tiburón ballena es una especie altamente migratoria que presenta agregaciones estacionales en distintas playas alrededor del mundo, siendo México uno de los países que visita anualmente. El Golfo de California representa una de las agregaciones estacionales más importantes del país y, sobre todo, Bahía de los Ángeles (Peregrín-Tovar, 2014; Ramírez-Macías et al., 2012). Su agregación se deriva de la presencia

de la población de *Acartia clausi*, principal fuente de alimento de la especie, el cual es un copépodo que presenta las mayores abundancias durante primavera en la bahía (Lavaniegos et al., 2012). Este fenómeno sucede anualmente durante los meses de mayo a diciembre (Peregrín-Tovar, 2014), propiciando el desarrollo de ecoturismo mediante la observación y nado con la especie, siendo un gran atractivo para visitantes nacionales e internacionales debido al comportamiento dócil de la especie con el humano, lo que permite conocerlo y acercarse a él, desarrollándose así una actividad económica de gran importancia para la comunidad (SEMARNAT, 2014).

La agregación de la especie en la bahía fue parte de los fundamentos para establecer la Reserva de la Biosfera de la Zona Marina Bahía de los Ángeles, Canal de Ballenas y Salsipuedes en 2007 (Danemann & Ezcurra, 2008). Actualmente, como parte de la estrategia de manejo y conservación del tiburón ballena, en la Reserva de Bahía de los Ángeles se lleva a cabo un monitoreo directo de la agregación anual a través de natación y buceo libre, con el objetivo de describir los patrones de abundancia, uso del área y comportamiento de la especie (SEMARNAT, 2014). Durante la temporada de agregación se realizan viajes cada dos o tres días en embarcaciones, durante los cuales se recolecta información incluyendo el conteo de organismos para el cálculo de la abundancia (Peregrín-Tovar, 2014), su foto-identificación a través fotografías de la parte lateral de la cabeza, ya que ahí se encuentra un patrón de coloración individual característico (Jaramillo-Gil, 2020; Speed et al., 2007); así como la estimación de tallas mediante la comparación del organismo con la embarcación (Jaramillo-Gil, 2020). Esta información también permite evaluar el periodo en que los individuos permanecen dentro de las agregaciones (Arzoumanian et al., 2005), la composición de sus tallas, sexos y edades (Rowat et al., 2011), e inclusive en algunos casos su tasa de crecimiento y la ecología de alimentación (Clark & Nelson, 1997).

El monitoreo directo presenta algunas limitaciones en su realización. La principal es que el tiburón ballena puede pasar desapercibido cuando se encuentra fuera del campo de visión del observador, como cuando se encuentra sumergido (Peregrín-Tovar, 2014). En cuanto a la toma de imágenes para la foto-identificación, si esta se realiza desde diferentes ángulos puede afectar en la correcta distinción de los individuos (Arzoumanian et al., 2005; Speed et al., 2007), así la presencia de un comportamiento evasivo por parte de algunos tiburones ballena ante el desarrollo de la actividad (Jaramillo-Gil, 2020). Así mismo, la falta de un método robusto para la estimación de las tallas de una manera más precisa (Jeffreys et al., 2013).

El presente trabajo busca determinar la efectividad del uso de drones y la aplicación de modelos de inteligencia artificial en la automatización de la identificación y seguimiento de tiburones ballena

(*Rhincodon typus*), como base para una mejora alternativa en los métodos de monitoreo de la especie en Bahía de los Ángeles.

1.1 Antecedentes

El tiburón ballena es el pez más grande del mundo, alcanzando longitudes estimadas de 14 a 18 metros, el cual es fácilmente distinguible por su patrón de coloración particular de puntos y líneas claras sobre un fondo oscuro y una superficie clara en el vientre, así como su cabeza grande y su boca ensanchada horizontalmente (Compagno, 2001). Es una especie altamente migratoria, que se distribuye en aguas tropicales y subtropicales alrededor del mundo con temperaturas entre los 20°C - 35°C, formando agregaciones costeras temporales debido a la presencia de florecimientos de organismos plantónicos (Rowat & Brooks, 2012). Se alimenta de copépodos, eufáusidos, segéstidos y peces pelágicos pequeños ingeridos por el tiburón mediante filtración pasiva y/o succión (Compagno, 2001; Hacohe-Domené et al., 2006; Hsu et al., 2007; Nelson & Eckert, 2007; Taylor, 2007). Su estrategia de vida es de tipo 'K', lo que implica un crecimiento lento, maduración tardía y gran longevidad (Ramírez-Macías et al., 2012), lo cual, junto con su alta movilidad y relativa baja abundancia, desencadena que sea una especie vulnerable a amenazas como la captura incidental, la colisión con embarcaciones, el turismo no regulado y el cambio climático (Compagno, 2001; Pierce & Norman, 2016).

Como consecuencia del efecto de estas amenazas, la especie se encuentra clasificada bajo distintas categorías de riesgo a nivel nacional e internacional. En la lista roja de la Unión Internacional para la Conservación de la Naturaleza (UICN) se encuentra evaluada como 'En Peligro' y en la NOM-059-SEMARNAT-2010 como 'Amenazado'. Debido a este estatus, se han desarrollado distintas estrategias para su conservación como la regulación de su pesca, consumo y comercio al estar enlistado en el Apéndice II de CITES a nivel internacional y la prohibición de su pesca en México (CONAPESCA, 2007). Así mismo, la protección de los ecosistemas donde se distribuye mediante el establecimiento de áreas naturales marinas protegidas. Por ello, es relevante el mantenimiento de un programa de monitoreo de la especie, así como la aplicación de tecnología y métodos que vuelvan más robustas las observaciones.

Dentro de los aspectos clave a monitorear en una población se encuentra el conteo de organismos, la foto-identificación y la estimación de tallas de los individuos (Jaramillo-Gil, 2020), ya que permiten conocer parámetros demográficos como abundancia, densidad y grado de residencia poblacional, así como patrones de conectividad entre regiones (Ramírez-Macías et al., 2012), con el fin de supervisar el estado

de las especies en el tiempo. Aplicando la tecnología de drones en el monitoreo, para la estimación de abundancia, comúnmente se emplean métodos de transectos lineales por su capacidad de optimizar el conteo de organismos y el uso de la batería del dron (Rex et al., 2023), pero conlleva desafíos como el doble conteo de los individuos debido al movimiento de los animales, problemas de detección cuando los individuos se ocultan, se sumergen o por la fatiga del observador que puede omitir organismos tras largas jornadas de trabajo, así como errores de identificación al confundir especies o elementos del paisaje con los objetivos de estudio (Brack et al., 2018). Para mitigar estos inconvenientes, se ha propuesto el traslape de un porcentaje de las imágenes tomadas en áreas contiguas, reduciendo así el riesgo de doble conteo (Brown et al., 2023; Hodgson et al., 2013); y el uso de filtros especiales en las lentes, como los de infrarrojo o visión térmica que mejoran la detección de individuos (Gooday et al., 2018; Hodgson et al., 2013; Ma et al., 2024).

En relación con la abundancia y aspectos de migración poblacional, tales como el grado de residencia y conectividad, resulta recurrente la utilización de técnicas de foto-identificación para el reconocimiento de individuos que presentan características morfológicas distintivas, como es el caso de las ballenas y delfines que exhiben estas características en su zona dorsal. Estas actividades se han visto facilitadas por el uso de imágenes aéreas obtenidas mediante drones e incluso satélites, debido a la mejora en la resolución de las cámaras incorporadas en estos equipos, así como a su operación remota que minimiza la posibilidad de causar disturbios a las especies estudiadas (Maglietta et al., 2022; Patton et al., 2023).

Otro aspecto clave en el monitoreo de fauna marina es la estimación de tallas de los individuos, para lo cual el uso de fotografías aéreas con drones ha resultado ser una técnica eficaz. En especies como las mantarrayas y los tiburones ballena se recurre a contrastar al organismo junto a un objeto de referencia de longitud conocida, ya sea transportado por un nadador (Whitehead et al., 2022) o dejándolo a la deriva cerca del animal (Setyawan et al., 2022), lo que permite generar modelos aproximados con mayor precisión para estimar la longitud del individuo.

A pesar de los avances en el proceso de monitoreo, los errores derivados de la manipulación humana persisten, especialmente durante el procesamiento de la gran cantidad de datos recopilados en campo, lo que puede resultar en identificaciones erróneas debido al cansancio y la saturación del operador (Goodwin et al., 2022). Así mismo, las condiciones ambientales continúan representando un reto al afectar la visibilidad de las cámaras debido a la nubosidad, viento, turbidez del agua y efectos de la radiación solar en el mar, como la absorción y reflexión (Jones et al., 2022). Para mitigar estos problemas, algunos autores

han optado por ajustar el ángulo de la cámara en que graba el dron (Brown et al., 2023), programar vuelos en horarios con menor reflejo solar, solapar imágenes y utilizar filtros polarizados (Hodgson et al., 2013).

Aunado a lo anterior, la IA se presenta como una solución viable para mejorar el proceso de monitoreo de especies. En la última década, su uso en el procesamiento de datos ecológicos ha crecido notablemente, sobre todo en tareas de clasificación, identificación y segmentación de objetos (Pichler & Hartig, 2023). Con estos métodos, se han procesado imágenes obtenidas por drones para identificar organismos marinos, tanto en la superficie como bajo el agua, y su clasificación entre distintas especies; además de obtener medidas morfométricas, realizar foto identificación, estimar el tamaño de agregaciones y abundancias poblacionales, así como en estudios de comportamiento en cetáceos, dugongos, tortugas marinas y tiburones (Bogucki et al., 2018; Dimauro et al., 2022; Dujon et al., 2021; Gray et al., 2019; Guirado et al., 2019; Maire et al., 2015). Igualmente, en tiburones se ha enfocado su uso para seguridad en playas, como forma de alerta temprana para usuarios y salvavidas en playas donde se comparte afluencia con tiburones para reducir los riesgos de incidentes (Kaur et al., 2024).

La evaluación del desempeño de los modelos de IA basado en análisis de imágenes se realiza mediante el cálculo de métricas según la tarea empleada: clasificación, detección de objetos o segmentación. Para la clasificación de imágenes se utilizan las métricas de exactitud (*accuracy*), la precisión (*precision*), la sensibilidad (*recall*) y puntaje F1 (*F1-score*), calculadas a partir de la matriz de confusión del modelo (verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos). La exactitud evalúa el desempeño general del modelo, pero no es adecuada para datos desbalanceados. La precisión mide qué tan acertadas son las predicciones, mientras que el recuerdo indica la capacidad del modelo para detectar los casos reales. El puntaje F1, útil en datos desbalanceados, combina precisión y recuerdo en un promedio armónico (Rainio et al., 2024). En cuanto a tareas de detección de objetos y segmentación, es ampliamente utilizado el valor de precisión media (mAP) y la Intersección sobre Unión (IoU), donde se comparan las áreas predichas por el modelo con respecto a las áreas verdaderas (Ditria et al., 2020).

Entre los desafíos de las CNN están el alto costo computacional para el entrenamiento de los modelos (Nath et al., 2019) y la necesidad de una gran cantidad de datos para entrenar el modelo (Fu et al., 2024). Para solventar la escasez de datos, se implementan técnicas como el 'Aprendizaje por Transferencia' (*Transfer Learning*), que permite adaptar modelos pre-entrenados a otras tareas o a imágenes de especies distintas al objeto de estudio del proyecto; así como técnicas de 'Aumento de Datos', que consiste en aplicar transformaciones geométricas (giro, recorte, zoom, brillo, etc.) y de color (contraste, brillo,

saturación, etc.) a las imágenes y con ello, incrementar la cantidad de datos de entrenamiento disponibles (Goodwin et al., 2022).

Así mismo, se han desarrollado programas accesibles para no programadores y que requieren una baja cantidad de datos de entrenamiento, como DeepLabCut (Mathis et al., 2018). Este programa de código libre es ampliamente utilizado en estudios de comportamiento animal en entornos controlados, utiliza redes pre-entrenadas y requiere una cantidad mínima de datos etiquetados para realizar tareas de seguimiento de individuos mediante un esqueleto definido por el usuario, el cual corresponde a una estructura de puntos clave que indican partes clave del cuerpo del organismo estudiado (como las articulaciones de orejas, patas, entre otros) conectadas entre sí mediante líneas o aristas, lo que permite dar seguimiento al movimiento (Mathis et al., 2018).

El estudio del tiburón ballena no ha quedado exento de la utilización de estas herramientas. Los drones se han implementado, principalmente, para realizar fotogrametría aérea de individuo (Whitehead et al., 2022). Si bien, se han llevado a cabo más investigaciones de monitoreo aéreo para la estimación de abundancias (Hodgson & Marsh, 2007; Rowat et al., 2009), el estudio de la ecología de forrajeo de la especie (Ketchum et al., 2013) y el análisis del comportamiento frente al turismo (Raudino et al., 2016), han sido realizadas a bordo de aviones tripulados. En cuanto a la aplicación de técnicas de IA, se ha utilizado el aprendizaje profundo mediante modelos de detección de objetos y segmentación para optimizar el proceso de foto-identificación en imágenes subacuáticas (Immonen et al., 2023; Kholiavchenko, 2022; Nepovinskykh et al., 2024). Además, se ha empleado el aprendizaje automático con modelos de bosques y ensambles no lineales para identificar comportamientos de forrajeo (Whitehead et al., 2021) y desarrollar modelos de idoneidad de hábitat (Daye, 2023). Finalmente, el uso combinado de drones e IA ha sido explorado en estudios recientes, como en la evaluación del impacto del ecoturismo en el comportamiento del tiburón ballena (Gayford et al., 2023).

1.2 Justificación

Bahía de los Ángeles presenta las características biológicas, ecológicas y oceanográficas que permiten la agregación estacional del tiburón ballena (Rowat & Brooks, 2012). Se ha hipotetizado que la región asociada a esta bahía puede llegar ser un área de refugio para la especie ante variaciones climáticas, ya que mantiene una temperatura menor y producción primaria mayor ante fenómenos como El Niño-Oscilación del Sur (ENOS) a diferencia del resto del Golfo de California (Jaramillo-Gil et al., 2023). Esta

agregación permite el desarrollo de actividades turísticas de observación y nado, una importante fuente de ingreso para la comunidad (SEMARNAT, 2014), lo cual requiere un manejo adecuado.

Las actividades de monitoreo de la especie son importantes para determinar parámetros poblacionales como la abundancia, tiempos de residencia, madurez sexual, así como estimar patrones de conectividad regional y estado de la especie (Brooks et al., 2010; Jaramillo-Gil et al., 2023). A su vez, permiten establecer medidas adecuadas en su conservación y manejo, como determinar la capacidad de carga para el desarrollo de actividades turísticas sustentables (Peregrín-Tovar, 2014) y limitar el tráfico marino en busca de reducir la probabilidad de choques con embarcaciones (Pierce & Norman, 2016).

Uno de los grandes desafíos que conlleva el monitoreo tradicional de la especie es el gran esfuerzo requerido (Gray et al., 2019). Ante ello, los drones y la IA presentan una alternativa para el monitoreo de fauna ya que maximiza la probabilidad de detección de los organismos estudiados durante un tiempo de muestreo menor (Hodgson et al., 2013), permitiendo un procesamiento de datos más eficiente (Christie et al., 2016) y una reducción del sesgo por operación humana (Goodwin et al., 2022). Además, reduce la perturbación en zonas inaccesibles o protegidas (Christie et al., 2016) como lo es el área de estudio de esta tesis.

Por ello, el presente trabajo recurre a la utilización de estas nuevas tecnologías en busca de automatizar las observaciones de la especie y disminuir los costos operativos con el fin de aumentar la información disponible y con ello, una mejora en las medidas correspondientes al manejo del tiburón ballena en la región.

1.3 Hipótesis

Considerando que el uso de drones e IA han permitido: 1) mejorar los métodos de monitoreo para la estimación de las abundancias, la foto-identificación y la medición de individuos; 2) que su uso ha permitido la reducción de esfuerzos en el trabajo de campo, un mayor acceso a zonas remotas, disminuir la perturbación de los organismos y su hábitat, 3) que su uso ha reducido el sesgo asociado a las tareas humanas y eficientizado el procesamiento de las imágenes e información para estudiar distintas especies marinas alrededor del mundo y 4) considerando las condiciones ambientales de Bahía de los Ángeles; se espera que la aplicación de estas herramientas produzcan resultados de una manera más eficiente y mejoren el estudio del tiburón ballena en esta bahía.

1.4 Objetivos

1.4.1 Objetivo general

Determinar la efectividad del uso de drones e inteligencia artificial para la identificación y seguimiento de tiburón ballena en Bahía de los Ángeles mediante un protocolo de monitoreo con drones y uso de modelos de aprendizaje profundo para mejorar el estudio del tiburón ballena en los sitios de agregaciones.

1.4.2 Objetivos específicos

- Recolectar videos de tiburón ballena con drones pilotados desde tierra y desde una embarcación para formar una base de datos que permita entrenar y probar los modelos de inteligencia artificial.
- Entrenar modelos de inteligencia artificial mediante dos distintos enfoques para resolver tareas de identificación y seguimiento del tiburón ballena en videos recolectados con dron.
- Evaluar y comparar el rendimiento de los modelos generados para determinar su efectividad en el monitoreo de tiburón ballena.
- Identificar, clasificar y determinar el efecto de las variables ambientales y operativas que afectan la identificación del tiburón ballena sobre el desempeño de los modelos de inteligencia artificial mediante el empleo de modelos estadísticos.

Capítulo 2. Metodología

2.1 Área de estudio

Bahía de los Ángeles se ubica al sureste del estado de Baja California, en la Región de las Grandes Islas (RGI) dentro del Golfo de California (GC) y pertenece al municipio de San Quintín. Es una zona de relevancia ecológica para el estado y el país ya que se encuentra dentro de dos áreas naturales protegidas, la parte terrestre pertenece al Área de Protección de Flora y Fauna Valle de los Cirios y la parte marina a la Reserva de la Biosfera Zona Marina Bahía de los Ángeles, Canal de Ballenas y de Salsipuedes. Es hábitat de una extensa variedad de especies tanto endémicas como migratorias siendo uno de ellos el tiburón ballena, el cual visita a esta bahía anualmente durante los meses de mayo a noviembre, lo que permite el desarrollo de actividades turísticas como la observación y nado de la especie, siendo una fuente importante de ingresos para la comunidad (SEMARNAT, 2024). Dentro de la reserva marina se designó la Subzona de Uso Público Tiburón Ballena (SUPTB) a la parte sur de la bahía (Figura 1), denominada localmente como El Rincón, donde se presenta la mayor cantidad y frecuencia de avistamientos de la especie dentro de la bahía (SEMARNAT, 2014; Peregrín-Tovar, 2014; Jaramillo-Gil, 2020).

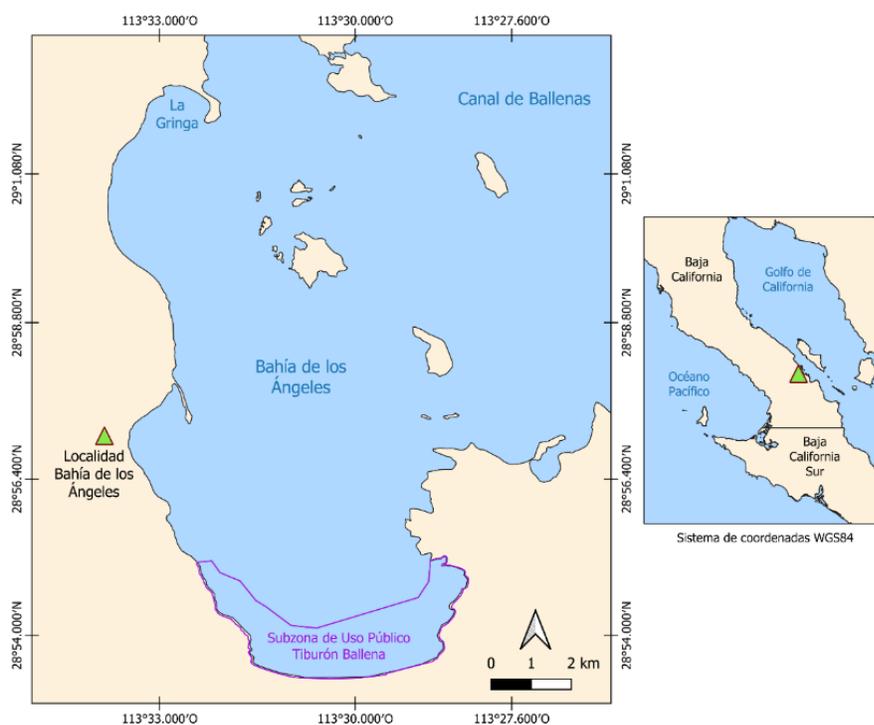


Figura 1. Área de estudio: El Rincón - Subzona de Uso Público Tiburón Ballena en la Reserva de la Biosfera Zona Marina Bahía de los Ángeles, Canal de Ballenas y de Salsipuedes.

En la SUPTB solo se permiten actividades de bajo impacto como monitoreo ambiental, nado y observación de tiburón ballena, así como pesca artesanal, de consumo y de fomento, siempre y cuando sea fuera de la temporada de agregación. Estas actividades deben apegarse a las reglas de operación y códigos de conducta estipulados en el plan de manejo de la reserva que se enfocan en la protección del tiburón ballena. En ellos se encuentran la delimitación del área con boyas, desarrollo de actividades turísticas con guías locales y autorizado, navegación a una velocidad menor de cuatro nudos y en embarcaciones de hasta 10 metros de eslora, permiso de nado con tiburón a un máximo de cuatro nadadores por inmersión en una distancia no menor de un metro del individuo y límite de 17 embarcaciones simultáneamente en la zona practicando la actividad turística, por mencionar algunos. Así mismo, el plan de manejo establece el desarrollo de actividades de monitoreo de la especie que pueden o no desarrollarse a la par de la actividad turística, en busca de conocer el estado de la especie y que permitan el manejo adecuado de la especie (SEMARNAT, 2018).

Gracias a la información recopilada en los monitoreos desde hace más de 20 años se ha podido caracterizar la agregación de tiburones ballenas presente en la bahía. La agregación se compone, en su mayoría, por machos juveniles con tallas menores a los nueve metros de longitud total, de los cuales entre el 30% y 50% presentan fidelidad al sitio y entre el 50% y 70% son nuevos individuos (Nates-López, 2015; Ramírez-Macías et al., 2012). Diversos autores han hipotetizado que la abundancia de alimento y temperaturas relativamente más bajas durante el verano que en el resto del GC, son los principales factores que explican la presencia y recurrencia de la especie en la zona (Hernández-Nava & Álvarez-Borrego, 2013; Jaramillo-Gil, 2023; Lavaniegos et al., 2012; Nates-López, 2015; Nelson & Eckert, 2007), manteniéndolas incluso durante la fase cálida del ENOS, siendo así un refugio para la especie (Jaramillo-Gil et al., 2023).

2.2 Características oceanográficas

La RGI debe su nombre a la presencia de las dos islas de mayor tamaño del GC: Ángel de la Guarda y Tiburón. Esta región cuenta con grandes variaciones en su batimetría que promueven la interacción de dos masas de agua: el Agua del GC con salinidad mayor a 35 UPS y temperatura menor a 12°C en los primeros 100 m de profundidad y por debajo, el Agua Subsuperficial Subtropical con salinidad entre los 34.5-35 UPS y temperaturas entre los 9.0-18 °C (Torres-Orozco, 1993). Estas masas de agua sufren una mezcla vertical intensa debido a las corrientes de marea que provoca una alta productividad y las temperaturas más bajas del golfo, por lo cual es considerada un área de refugio para diversos organismos marinos ante fenómenos climatológicos como ENOS (Argote et al., 1995; Tershly et al., 1991). Caso similar

sucede en el Canal de Ballenas, una zona que se encuentra entre la Isla Ángel de la Guarda y el archipiélago frente a Bahía de los Ángeles, donde las corrientes de marea permiten el ingreso de la Agua Subsuperficial Subtropical al canal generando surgencia constante (Argote et al., 1995; López et al., 2006; Paden et al., 1993; Soto-Mardones et al., 1999). Adentrándose hacia la costa se encuentra Bahía de los Ángeles, donde la batimetría, el viento, la temperatura y las masas de agua adyacentes interactúan generando las mismas condiciones de surgimiento de aguas con altos contenidos de nutrientes y mezcla vertical, siendo idóneas para el sustento una gran cantidad de especies marinas, como tiburones (*Squatina californica* y *Mustelus californicus*); rayas (*Rhinobatos productus*, *Myliobatis californica* y *Grymnura marmorata*); tortugas marinas (*Chelonia mydas*); y mamíferos marinos tales como ballenas (*Balaenoptera physalus*), delfines (*Delphinus capensis*) y lobos marinos (*Zalophus californianus californianus*) (SEMARNAT, 2014).

Esta bahía de profundidad somera (<50m) mantiene un alto intercambio de agua con el Canal de Ballenas, a través de sus cuatro canales (Prof. ~ 50 m) y el patrón de vientos característicos de la zona (SEMARNAT, 2014). Durante el invierno, el viento proviene del norte-noroeste provocando el ingreso del agua del Canal de Ballenas desde el noroeste con flujo en dirección sur, promoviendo temperaturas superficiales del agua (SST) entre 14°-22°C y salinidad entre 34.6-35.3 UPS.

Caso contrario sucede en el verano, donde los vientos que provienen del este-sureste permiten el ingreso de agua del Canal de Ballenas a la parte sur, donde fluye hacia el oeste y con un giro hacia el sur, manteniendo SST entre 24°-29°C y salinidad entre 35.3-35.6 UPS. En el caso de primavera y otoño, se presentan variaciones en la dirección e intensidad del viento, principalmente provienen del oeste que junto con largos periodos de calmas dominados por brisas que promueven un intercambio entre el interior y exterior de la bahía generando una alta productividad en la zona (Amador-Buenrostro et al., 1991; Nava-Ruiz, 2018).

2.3 Recolección de datos

Se realizaron 59 vuelos de dron (nueve horas en total) en el área de estudio que corresponde a la SUPTB dentro de la Reserva Marina, así como la zona popularmente conocida como La Rampa, que es el muelle principal de la localidad (Figura 2). Los vuelos fueron realizados durante el mes de octubre del 2023, siendo 42 pilotados desde tierra y 17 pilotados desde una embarcación (Figura 2). Para los vuelos, se utilizó un dron comercial DJI Phantom 4 con una cámara con resolución 4k (Ultra HD 3840x2160 pixeles).

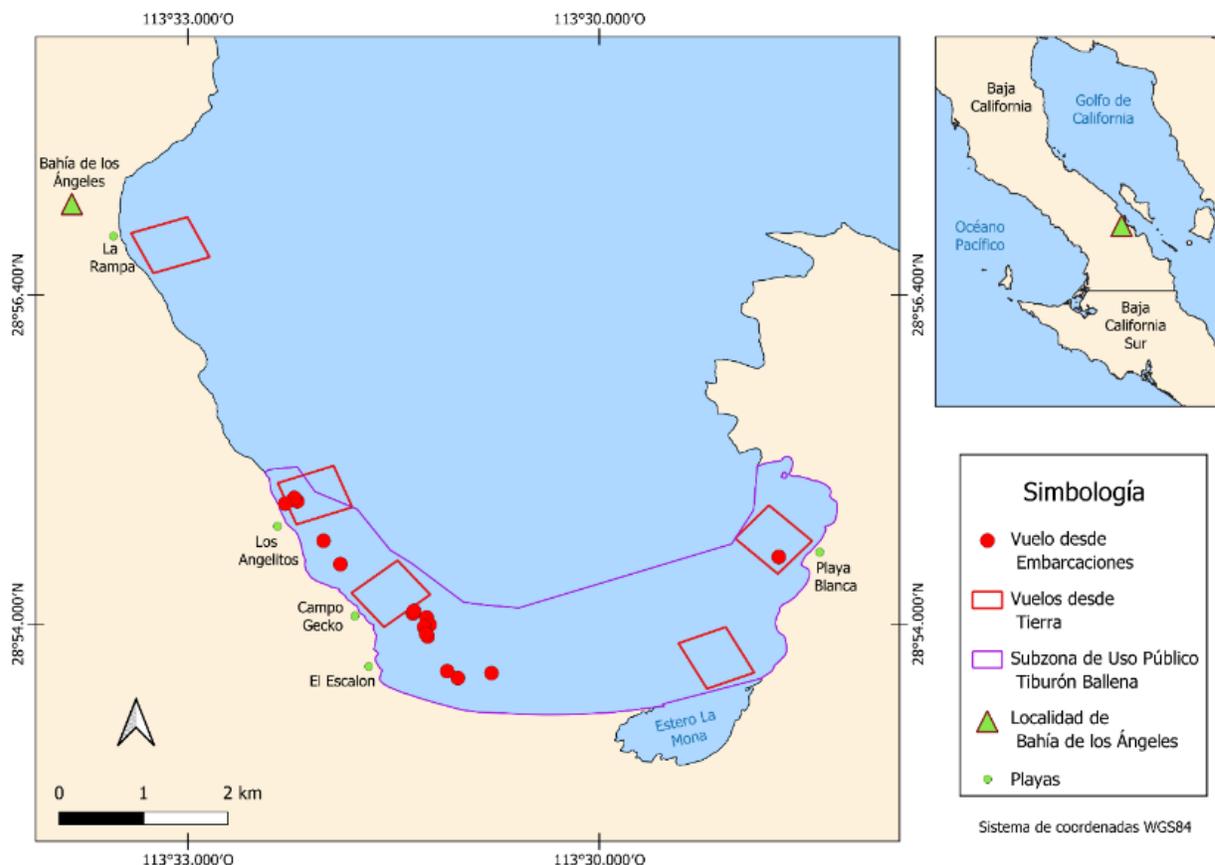


Figura 2. Ubicación de las zonas sobrevoladas con dron. Los polígonos rojos corresponden a los vuelos pilotados desde tierra en distintas playas de la bahía. Los puntos rojos corresponden a los vuelos pilotados desde embarcaciones.

2.3.1 Descripción de vuelos de dron pilotados desde tierra

Los vuelos fueron realizados en cinco playas: La Rampa, Los Angelitos, Gecko, Desembocadura del Estero y Playa Blanca (Figura 2). Estas playas fueron elegidas por encontrarse dentro de la SUPTB (a excepción de La Rampa), contar con altos registros de individuos durante temporadas de agregación anteriores (Peregrín-Tovar, 2014; SEMARNAT, 2024), así como por la accesibilidad a la zona.

No fue posible encontrar un acceso terrestre a la zona central de la SUPTB debido a condiciones del terreno, así como la privatización de tierras. Los polígonos que se sobrevolaron contaron con dimensiones de 700 m de largo y 600 m de ancho (Figura 3), los cuales, fueron divididos en cuatro transectos lineales separados 500 m entre sí, con un solapamiento de 10%. Las actividades de vuelo, despeje y aterrizaje fueron realizadas de manera autónoma con la asistencia del software DJI pilot (versión 2.5).

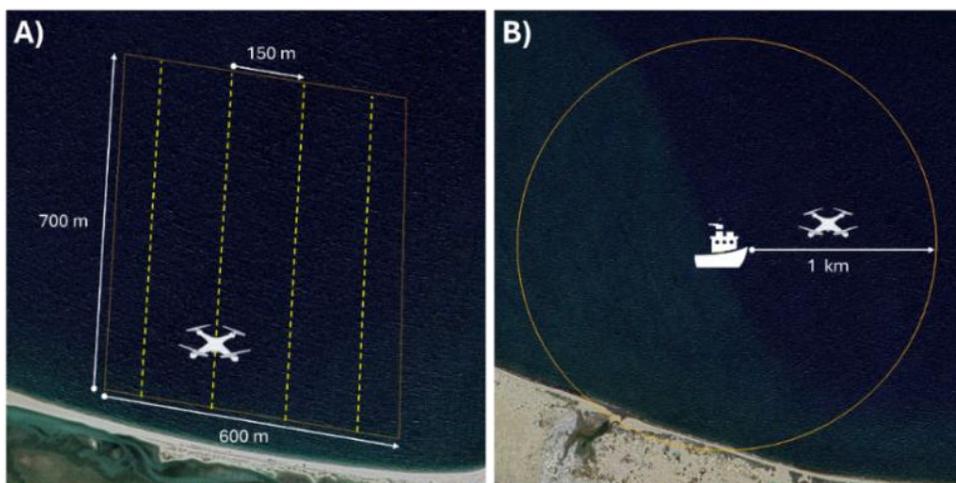


Figura 3. Esquema de vuelos pilotados desde tierra (A) y desde embarcación (B).

Las dimensiones de los polígonos fueron determinadas con base en la autonomía del dron y la logística. Respecto a la autonomía del dron, se consideró no sobrepasar la distancia máxima de conexión entre la nave con el control remoto permitido por las aplicaciones de DJI (~1 km), así como no exceder una duración de 20 minutos por batería, la cual puede disminuir conforme la temperatura ambiental aumenta.

En cuanto a la logística, el tamaño del polígono permitió muestrear entre 2 y 3 playas por día con dos repeticiones (mañana y tarde), considerando las horas de luz diaria, el ángulo de incidencia del sol (reflejo solar en el mar), así como el tiempo de traslado, que iba desde 5 min a la playa más cercana hasta aproximadamente 45 minutos a la más lejana.

Con base en lo anterior, se dividió al área de estudio en parte norte (La Rampa, Angelitos y Gecko) y parte sureste (Estero y Playa Blanca), donde se realizaron 2 vuelos por día en cada sitio. El vuelo matutino fue entre 9am-12pm y el vespertino entre 2pm-4pm a una altitud de vuelo de 100 m. Cada polígono requería 15 min de vuelo (tomando en cuenta el despegue y el aterrizaje), por lo que era necesario solo 1 batería por sitio.

2.3.2 Descripción de vuelos de dron pilotados desde embarcación

Este tipo de vuelos fueron realizados a la par de viajes de monitoreo del grupo comunitario PEJESAPO donde se recorría la SUPTB. La búsqueda de individuos fue realizada de manera visual por parte de los tripulantes de la embarcación. Una vez que eran avistados los individuos, la embarcación se posicionaba

a una distancia no menor de 5 m del tiburón (SEMARNAT, 2024), por lo que, mientras se realizaba el monitoreo tradicional del organismo, el piloto de dron aprovechaba para volar sobre el mismo individuo sin sobrepasar la distancia de segura de conexión (1 km, Figura 3) y el tiempo permitido por la autonomía de la batería (~20 min). La altitud de vuelo era de 50 metros para el seguimiento del organismo y de 100 metros para identificar la embarcación o la presencia de algún otro individuo cerca; si este último escenario se cumplía, se procedía a realizar el proceso descrito anteriormente.

El vuelo, el despegue y el aterrizaje fueron realizados de manera manual, y asistidos con la aplicación de DJI pilot. Para el despegue y el aterrizaje se contó con la asistencia de un tripulante de la embarcación, quien sostenía y recibía al dron con sus manos al aire en la zona de la proa. Debido a lo anterior y la movilidad de la embarcación por el oleaje, se desactivaron las medidas de seguridad del dron sobre el regreso al origen (zona de despegue) al finalizar el vuelo o en caso de emergencia, así como la prohibición del descenso debido a la detección de obstáculos (personas u objetos cercanos).

La velocidad del viento fue la principal limitante de los vuelos, ya que no debía sobrepasar los 10 m/s. Por lo que, se revisaba la predicción del estado del tiempo un día antes y al inicio del día en la aplicación Windy (Windytv, República Checa), así como antes de iniciar el vuelo en cada sitio con el anemómetro.

2.3.3 Datos recolectados.

Durante cada vuelo, se registraron en una bitácora de campo datos de localización (sitio, coordenadas, fecha y hora), condiciones ambientales (turbidez, velocidad del viento, porcentaje de nubosidad y porcentaje de brillo solar) así como observaciones de la especie (en caso de encontrar) (Figura 4).

La escala de turbidez fue registrada con base en el tono del mar y la visibilidad del fondo, basado en Hodgson y colaboradores (2013), con las categorías siguientes: 1) Profundo con el fondo no visible y color mayormente azul 2) Profundo con el fondo no visible y agua turbia con color 50% verde y 50 azul, y 3) Somero con el fondo no visible y agua turbia, color mayormente verde.

La velocidad del viento fue registrada en unidades 'm/s' con un anemómetro marca Brunton ADC-Summit. La nubosidad y el brillo solar fueron determinados con respecto al porcentaje de cobertura de nubes en el cielo y el reflejo del sol en el mar, respectivamente.

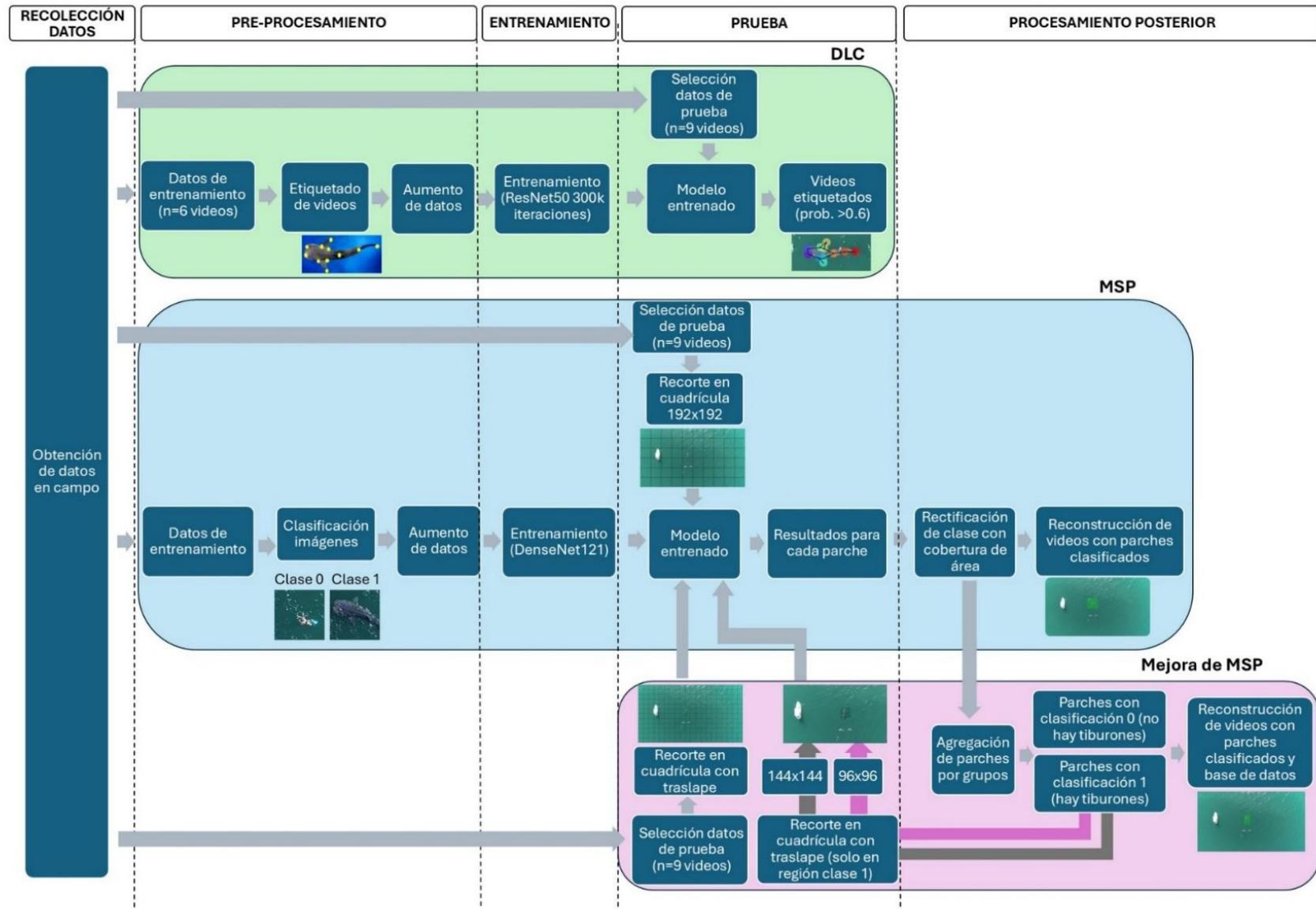


Figura 5. Proceso empleado según el enfoque implementado: DLC y MSP.

2.4.1 Enfoque de entrenamiento mediante software DeepLabCut (DLC)

2.4.1.1 Preprocesamiento de datos

Se seleccionaron seis videos para conformar el conjunto de datos de entrenamiento, siendo tres pilotados desde tierra y otros tres desde embarcación, los cuales fueron procesados a resolución FullHD (1920x1080). De estos videos, mediante la utilización del programa DLC (Mathis et al, 2018; Nath et al., 2019), se etiquetaron 20 imágenes representativas de cada uno con un esqueleto de 11 marcadores que indican características morfológicas relevantes del tiburón: dos en la cabeza, tres para el inicio y final de cada aleta pectoral, uno para la primera aleta dorsal, uno para la quilla caudal y uno para la punta de la aleta caudal (Figura 6). Este esqueleto fue propuesto para esta investigación.

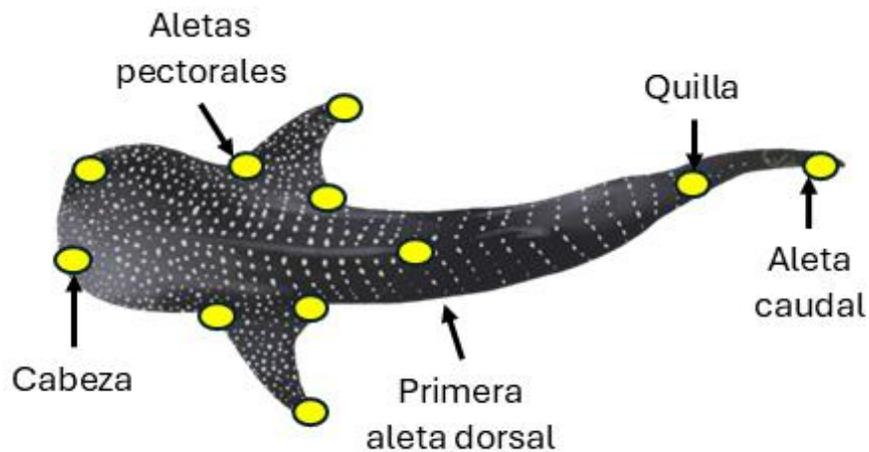


Figura 6. Esqueleto de tiburón ballena compuesto de 11 marcadores para el entrenamiento del enfoque DeepLabCut.

Con las 180 imágenes etiquetadas, el programa realizó un aumento de datos predeterminado mediante la utilización de la biblioteca "Imgaug". Esta función aplica transformaciones como rotación, escalado, desplazamiento, cambios de brillo y contraste, entre otras, de manera aleatoria en cada iteración del entrenamiento. La magnitud del aumento no es una cantidad fija, ya que el aumento de imágenes es distinto para cada época de entrenamiento, lo que resulta en múltiples versiones aumentadas de las imágenes originales (Mathis et al., 2018).

2.4.1.2 Fase de entrenamiento y prueba

Utilizando el conjunto de datos etiquetado, se entrenó la arquitectura ResNet50 (Insafutdinov et al 2016 & He et al 2016), con 300k iteraciones (Figura 7). Esta arquitectura fue seleccionada debido a recomendaciones por parte de los desarrolladores del software ya que, al ser una red troncal (*backbone*), está pre-entrenada con una gran base de datos lo que permite ser utilizada como punto de partida para cualquier tarea, así como por exhibir un mejor rendimiento durante la fase de entrenamiento y prueba entre las 200k y 300k iteraciones (Mathis et al., 2018).

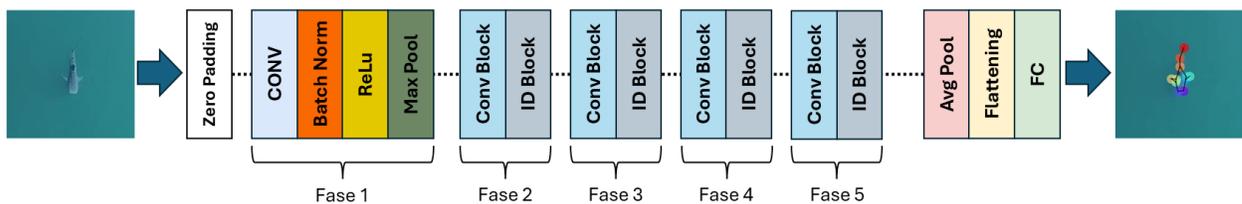


Figura 7. Arquitectura ResNet50.

Por cada video del conjunto de prueba se obtuvo como resultado las coordenadas (X, Y) y la verosimilitud (*likelihood*) para cada marcador del esqueleto propuesto en cada una de las imágenes de las que se compone cada video. El software utiliza un umbral de verosimilitud predeterminado de 0.6 para definir que el marcador identificado corresponde al tiburón, el cual también fue utilizado en el análisis.

2.4.2 Enfoque de entrenamiento a partir de Parches Multi-Escala (MSP)

2.4.2.1 Preprocesamiento de datos

El enfoque MSP se centra en la clasificación de regiones de las imágenes, por lo que, para conformar los conjuntos de entrenamiento y validación fue necesario seleccionar manualmente imágenes (*frames*) de cada video grabado. Este proceso busca incluir una variedad de elementos que puedan representar los ambientes reales del sitio donde se observa al tiburón ballena. La inclusión de estos elementos variados permite mejorar la precisión de las predicciones del modelo. Entre los elementos seleccionados se encuentran la presencia de tiburones, otras especies animales, embarcaciones, diferentes colores del agua de mar y distintas intensidades de brillo solar, entre otros aspectos.

Por ser una tarea de clasificación de objetos y con el fin de identificar a los individuos de tiburón ballena presentes en la imagen, el problema fue abordado desde un enfoque binario. Para ello, se recortaron manualmente regiones de cada imagen extraída para conformar dos conjuntos de datos: tiburón (1) y no tiburón (0). Los recortes del primer conjunto de datos se realizaron colocando al tiburón en la parte central de la zona de recorte manteniendo un borde equilibrado tanto en la vertical como en la horizontal; para los casos en que el tiburón se encontraba en un extremo de la imagen y/o solo podía observarse una parte del cuerpo, se mantuvo el borde en las zonas que lo permitían. Este borde fue de suma importancia ya que, posteriormente, se aplicaron técnicas de aumento de datos necesarias debido a la baja cantidad de observaciones de tiburón ballena con los que contábamos.

Para conformar el segundo conjunto de datos (no tiburón), simplemente se buscó obtener una cantidad similar de recortes que se tuvieron para la otra clase y así evitar el sesgo en el modelo. Los recortes fueron representativos de la variedad de objetos y condiciones ambientales observables en cada video que no correspondían a un tiburón ballena desde la presencia de aves, tortugas, embarcaciones, hasta diferentes tonalidades del agua y reflejo del brillo solar.

De esta manera, el conjunto de datos original fue de 1,413 imágenes (Tabla I). Como se mencionó previamente, los modelos de aprendizaje profundo necesitan una gran cantidad de datos de entrenamiento para realizar las tareas de una manera más precisa, se optó por utilizar técnicas de aumento de datos mediante transformaciones geométricas (rotación, corte y *zoom*) y de color (brillo, contraste y saturación), como se muestra en la Figura 8. Logrando así que la cantidad de imágenes creciera a 41,500, siendo 24 310 para el conjunto de datos de entrenamiento y 17 190 para validación.

Tabla I. Distribución de imágenes por clase y fase para el modelo MSP.

Fase	Cantidad de imágenes por clase		
	Tiburón (1)	No tiburón (0)	Total
Original	603	810	1,413
Aumento de datos	17,190	24,310	41,500
Conjunto de datos de entrenamiento	13,752	19,448	33,200
Conjunto de datos de validación	3,438	4,862	8,300



Figura 8. Técnicas de aumento de datos.

2.4.2.2 Fase de entrenamiento y prueba

Con las imágenes anteriores se probaron distintas arquitecturas como modelos candidatos, optando por la DenseNet 120 (Figura 9) ya que fue la que mostró un mejor desempeño. Además, esta arquitectura presenta múltiples ventajas en cuestión de entrenamiento: 1) Facilita la propagación de información y mejora el flujo de gradientes a través de la red. 2) Es altamente eficiente en el uso de parámetros, ya que reutiliza los mapas de características aprendidos en capas anteriores, reduciendo significativamente la cantidad de parámetros requeridos sin comprometer el rendimiento. 3) Esta eficiencia también se traduce en una menor tendencia al sobreajuste, lo que es especialmente útil en conjuntos de datos limitados. Finalmente, 4) permite entrenar redes más profundas y precisas con menor demanda de memoria y computación en comparación con otras arquitecturas como ResNet (Huang et al., 2016).

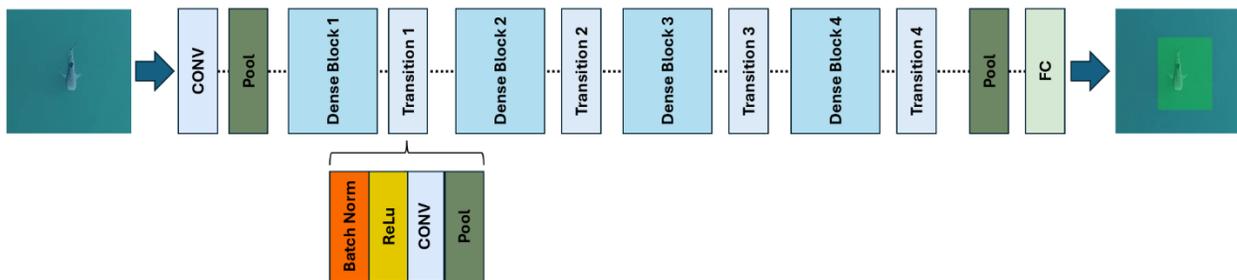


Figura 9. Arquitectura DenseNet121.

Las imágenes de entrenamiento y validación fueron re-escaladas a un tamaño de 224x224 píxeles para contar con uniformidad. De esta manera, fueron ingresadas al modelo y utilizadas para su entrenamiento y posteriormente su validación.

El conjunto de datos de prueba fue conformado por los mismos nueve videos de prueba que para el enfoque DLC. Al ser un enfoque que funciona con imágenes individuales, se extrajeron imágenes cada 0.1 segundo de cada video. A partir de ello, se recortó cada imagen con base a una cuadrícula con parches de 192x192 píxeles, los cuales fueron re-escalados a 224x224 píxeles por ser el tamaño de aprendizaje del modelo. Las imágenes fueron ingresadas al modelo para su prueba en lotes (*batch size*) de 2,048, obteniendo como resultado un valor de probabilidad para cada parche de cada video ingresado al modelo. Tomando como base el otro enfoque (DLC), determinamos el umbral de 0.6 como la probabilidad mínima para determinar que en ese parche se encuentra un tiburón (Figura 10Figura 10).

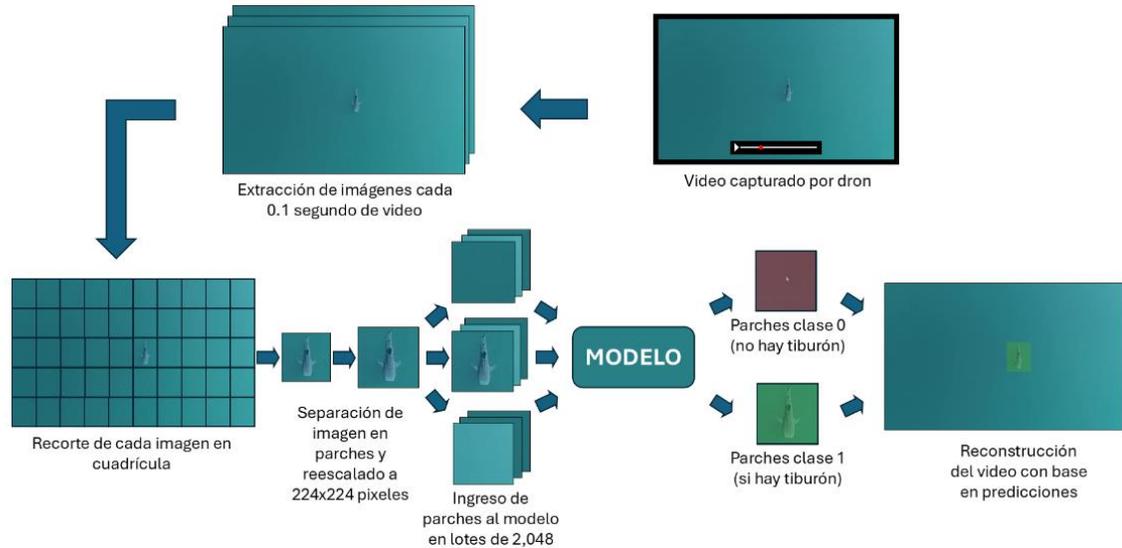


Figura 10. Proceso de prueba del modelo MSP.

2.4.2.3 Procesamiento posterior

Con el objetivo de reducir los falsos positivos, se realizó un ajuste temporal a los resultados para cada fotograma por segundo de video mediante un promedio ponderado de la probabilidad de las 10 imágenes anteriores, buscando que las imágenes inmediatamente anteriores tuvieran más peso que aquella más alejada. De manera que, en las cuatro inmediatamente anteriores se encontraba el 50% del peso, las siguientes tres de 30%, las siguientes dos de 15% y la última de 5% (Figura 11).



Figura 11. Distribución de pesos para promedio ponderado según la cercanía a la imagen actual.

Obteniendo de la siguiente manera el promedio ponderado de la probabilidad que sería la probabilidad final para cada imagen evaluada (1):

$$Probabilidad = \frac{\sum \text{Imagen} * \text{Peso}}{\text{Cantidad total de imágenes}} \quad (1)$$

2.4.2.4 Mejora del seguimiento en modelos bajo el enfoque MSP.

Se aplicó una mejora en el seguimiento del individuo en las predicciones mediante ajustes en el tamaño de los parches, el solapamiento entre ellos, la determinación de regiones positivas de la imagen y la delimitación del área predicha. En busca de simplificar los nombres de los modelos para una mejor comprensión del texto y los gráficos, se asignaron los siguientes nombres (ver Tabla II).

Tabla II. Nombres de los diferentes modelos predictivos desarrollados.

Nombre modelo	Significado
DLC	DeepLabCut
MSP-192	Parches de 192x192 pixeles
MSP-192O	Parches de 192x192 pixeles con solapamiento
MSP-96O	Parches de 96x96 pixeles con solapamiento
MSP-96OB	Parches de 96x96 pixeles con solapamiento y cuadro delimitador
MSP-96OB.75	Parches de 96x96 pixeles con solapamiento y cuadro delimitador recortado a $\frac{3}{4}$ cada lado desde centroide
MSP-96OB.5	Parches de 96x96 pixeles con solapamiento y cuadro delimitador recortado a $\frac{1}{2}$ cada lado desde centroide

El modelo MSP-192 se tomó como referencia y se modificó la forma de recortar las imágenes de entrada según el tamaño de los parches y la presentación del resultado final utilizando cuadros delimitadores. Para ello, se realizaron tres iteraciones del modelo, empleando diferentes tamaños de parches (192x192 pixeles, 144x144 pixeles y 96x96 pixeles) y un solapamiento entre ellos del 50% tanto en la vertical como en la horizontal. En la primera iteración (MSP-192O), el proceso fue aplicado a toda la imagen. Para la segunda y tercera iteración el análisis se centró en las regiones clasificadas como positivas (clasificación 1, presencia de tiburón) en la iteración anterior, en lugar de abarcar toda la imagen, limitándose solo a las áreas donde el modelo identificaba tiburones (Figura 12).

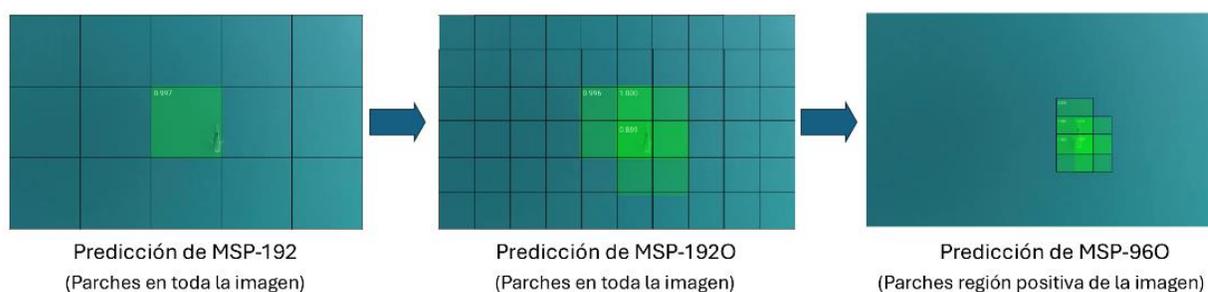


Figura 12. Generación de predicción en modelos MSP-192, MSP-192O y MSP-96O.

Para determinar dichas regiones, se agruparon los parches positivos mediante intersecciones y adyacencias, añadiendo un píxel al ancho y al largo de cada parche clasificado como 1, y mediante una matriz de adyacencias se formaron grupos en diferentes regiones de la imagen. De esta manera, el resultado final de la tercera iteración (MSP-96O) fue más preciso en cuanto a la ubicación y el tamaño del individuo, a la par que se mitigaron los falsos positivos obtenidos en las predicciones anteriores. De estas iteraciones se obtuvieron los resultados en crudo de los parches traslapados (MSP-192O y MSP-96O) y, en la última iteración, se delimitaron las zonas con intersecciones en más de dos parches (MSP-96OB). Con el fin de que los cuadrantes se ajustaran mejor al tamaño del tiburón, se realizó una reducción de los parches basada en un centroide. A partir de los resultados de la tercera iteración y delimitados (MSP-96OB), los polígonos se recortaron en altura y anchura un 75% (MSP-96OB.75) y un 50% (MSP-96OB.5) con respecto al centroide (Figura 13). En todos los casos en los que se utilizaron cuadros delimitadores, la probabilidad final fue calculada como el promedio de los cuadrantes agrupados.

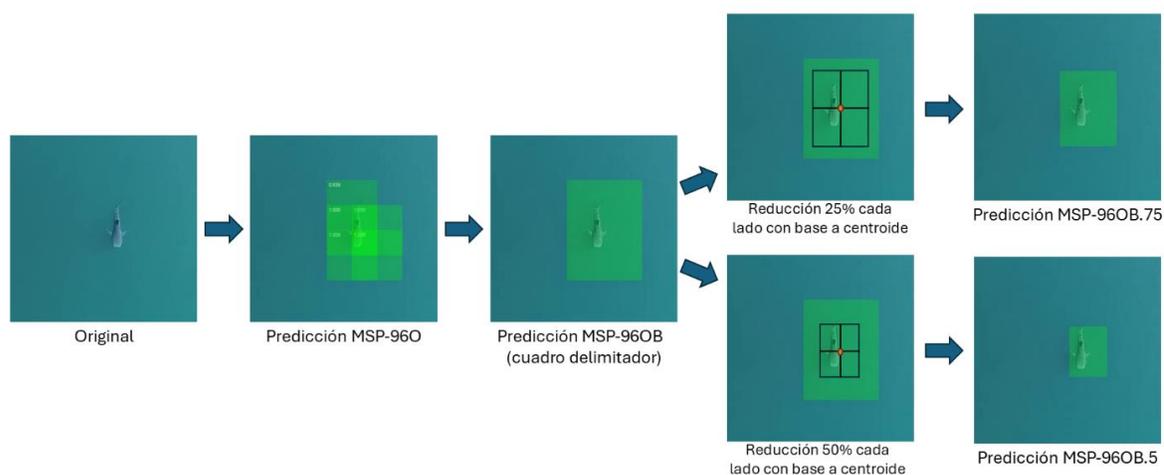


Figura 13. Proceso de delimitación de parches.

2.5 Comparación del rendimiento de los diferentes enfoques

Se realizó una comparación de los resultados de cada enfoque contra el área verdadera (*ground truth*), con base en la comparación de las áreas cubiertas. Para ello, fue necesario revisar manualmente cada segundo de video para localizar al tiburón en la imagen, anotando las coordenadas (X, Y) así como el ancho y alto del rectángulo formado al rodearlo, obteniendo así la ubicación real del individuo. En el caso del enfoque de DLC, debido a que se cuenta con una posición (X, Y) y una verosimilitud (*likelihood*) para cada marcador del cuerpo, fue necesario obtener los valores mínimos y máximos de 'X' y 'Y', para así obtener la figura para la comparación. En el caso del enfoque de MSP, se utilizaron los datos tal cual se obtuvieron

del modelo, ya que la salida fue un par de coordenadas iniciales (X, Y) así como el ancho y alto que equivalen al tamaño del parche (Figura 14).

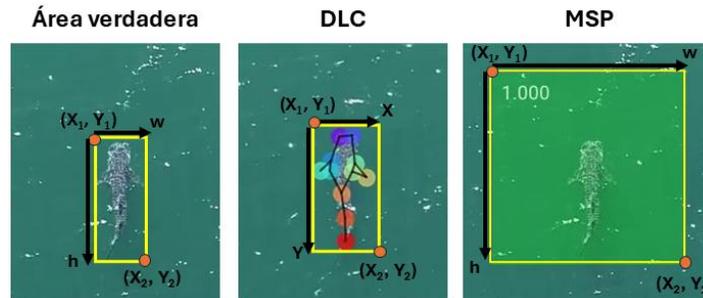


Figura 14. Generación de coordenadas para cuadro delimitador con base en el enfoque utilizado.

Las ecuaciones utilizadas para determinar los valores de (X_1, Y_1) y (X_2, Y_2) para poder delimitar el contorno que rodea al tiburón se muestra a continuación (2):

$$\begin{aligned} X_1 &= X_1 \\ Y_1 &= Y_1 \\ X_2 &= X_1 + w \\ Y_2 &= Y_1 + h \end{aligned} \quad (2)$$

Para el enfoque DLC, primero fue necesario determinar los valores mínimos y máximos para las coordenadas X, Y mediante los marcadores *-key points-* (3):

$$\begin{aligned} X_1 &= \operatorname{argmin} \forall \text{ key point } (X) \\ Y_1 &= \operatorname{argmin} \forall \text{ key point } (Y) \\ w &= \operatorname{argmax} \forall \text{ key point } (X) - X_1 \\ h &= \operatorname{argmax} \forall \text{ key point } (Y) - Y_1 \\ X_2 &= X_1 + w \\ Y_2 &= Y_1 + h \end{aligned} \quad (3)$$

Una vez que contamos con estos datos, se formaron los rectángulos y se procedió a comparar la cobertura de área para determinar el porcentaje de cobertura del área predicción con respecto a la verdadera (Figura 15). Para ello, fue necesario determinar la intersección entre los dos polígonos a comparar, obtener su área y proceder a calcular la cobertura de área (CA) de la siguiente manera (4):

$$\text{Cobertura de área (CA)} = \frac{\text{Área de intersección}}{\text{Área verdadera}} \quad (4)$$

De esta manera, se calculó la CA de cada fotograma por segundo de video, así como la media y la desviación estándar de esta métrica para determinar y comparar la efectividad del seguimiento de cada enfoque.

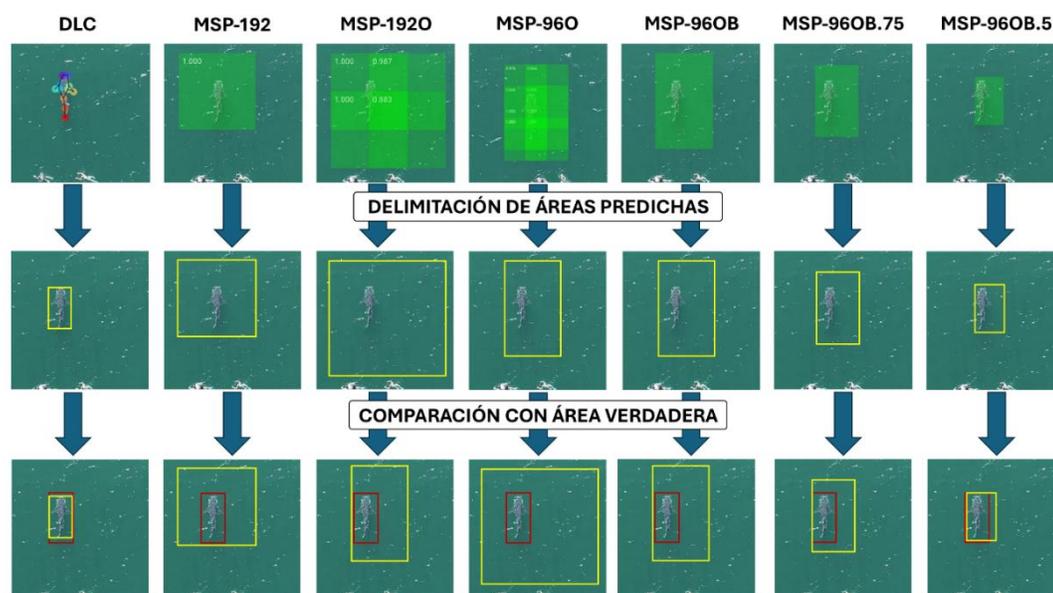


Figura 15. Ejemplos de la comparación del área predicha contra la verdadera.

Además, se calcularon métricas que nos permiten evaluar el rendimiento en cuanto a tareas de clasificación de los modelos de clasificación. Para ello, se hace un recuento de cuatro variables relacionadas con los valores reales y los predichos correspondientes a la identificación del tiburón (Figura 16): verdaderos positivos (TP), verdaderos negativos (TN), falsos positivos (FP) y falsos negativos (FN). Los dos primeros representan una correcta identificación por parte del modelo y los dos últimos, una incorrecta por diferir con el área verdadera.

		REALIDAD	
		Positivo	Negativo
PREDICCIÓN	Positivo	<p>TP (Verdadero Positivo)</p>	<p>FP (Falso Positivo)</p>
	Negativo	<p>FN (Falso Negativo)</p>	<p>TN (Verdadero Negativo)</p>

Figura 16. Matriz de confusión de los valores reales contra los predichos.

A partir de ellos, se calculan las métricas de Exactitud (5), Precisión (6), Recuerdo o Sensibilidad (7), Puntaje F1 (8) y Puntaje F1 macro (9) de la siguiente manera:

$$Exactitud = \frac{TP + TN}{TP + FP + FN + TN} \quad (5)$$

$$Precisión = \frac{TP}{TP + FP} \quad (6)$$

$$Sensibilidad = \frac{TP}{TP + FN} \quad (7)$$

$$Puntaje F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (8)$$

$$F1 \text{ macro} = \frac{1}{n} \sum_{i=1}^n F1_i \quad (9)$$

donde n es el número de clases y $F1_i$ es el valor de Puntaje F1 para la clase i .

La métrica de ‘Exactitud’ mide la proporción de predicciones correctas (tiburón y no tiburón) respecto al total de imágenes evaluadas. Es decir, es la proporción de imágenes clasificadas correctamente, tanto tiburón ballena como no tiburón, sobre el total de imágenes.

La ‘Precisión’ mide la proporción de imágenes correctamente identificadas como tiburón ballena (clase positiva) entre todas las predicciones realizadas por el modelo que fueron clasificadas como tiburón ballena, ya sean correctas o incorrectas. Es útil para evaluar cuántas de las predicciones positivas del modelo son realmente correctas. La ‘Sensibilidad’, también conocida como Tasa de Verdaderos Positivos, mide la proporción de tiburones ballena correctamente identificados entre todas las instancias realmente positivas (es decir, imágenes que contienen tiburón ballena). En otras palabras, indica la capacidad del modelo para detectar correctamente todas las instancias de tiburón ballena, sin importar las predicciones incorrectas.

El ‘Puntaje F1’ es la media armónica entre ‘Precisión’ y ‘Sensibilidad’, proporcionando un balance entre ambas métricas. Es especialmente útil en escenarios con clases desbalanceadas, donde la precisión por sí solo podría ser engañosa. El Puntaje F1 logra equilibrar la necesidad de precisión (pocas predicciones incorrectas como tiburón ballena) y la necesidad de sensibilidad (detectando todos los tiburones ballena posibles).

2.6 Variables que afectan la efectividad de los modelos

A partir de la experiencia en campo y las predicciones generadas por los modelos, se identificaron y cuantificaron las variables con potencial impacto en su desempeño, y están descritas en la Tabla III. Estas ocho variables fueron clasificadas a su vez por tipo, debido a la dependencia con algún aspecto (variables ambientales, del movimiento del tiburón y de la logística del trabajo de campo).

Tabla III. Clasificación de variables con efecto en el desempeño de los modelos.

Variable	Tipo	Niveles	Descripción
Horario de vuelo (Tiempo transcurrido desde el amanecer)	Ambiental / logística	1	0-2 hrs
		2	2-4 hrs
		3	4-6 hrs
		4	6-8 hrs
		5	8-10 hrs
Brillo solar	Ambiental	0	0%
		1	1%-15%
		2	15-30%
		3	>30%
Fuente solar	Ambiental / logística	1	Superior
		2	Superior Derecha
		3	Superior Izquierda
		4	Inferior
		5	Inferior Derecha
		6	Inferior Izquierda
		7	Derecha
		8	Izquierda
Turbidez	Ambiental	1	Profundo con el fondo no visible, color mayormente azul
		2	Profundo con el fondo no visible y agua turbia
		3	Somero con el fondo no visible y color mayormente verde
Altitud de vuelo	Logística	1	Baja altitud (50 m)
		2	Gran altitud (100 m)
Rotación del dron	Tiburón	0	No cambia la dirección de grabación
		1	Cambia la dirección de grabación
Sumergimiento del tiburón	Tiburón	0	El tiburón se encuentra en superficie
		1	El tiburón se sumerge parcial o completamente
Presencia de objetos externos	Logística / ambiental	0	Sin presencia de objetos cerca al tiburón
		1	Presencia de fauna cerca al tiburón
		2	Presencia de nadadores cerca al tiburón
		3	Presencia de pangas cerca al tiburón
		4	Fauna oculta a tiburón
		5	Nadadores ocultan al tiburón
		6	Pangas ocultan al tiburón

A partir de esta clasificación, se realizaron modelos lineales y Modelos Aditivos Generalizados (GAM's, por sus siglas en inglés) para determinar si existía una relación entre las variables independientes con la predicción del mejor modelo, mediante el programa de R (R Core Team, 2024). Se probaron varios modelos donde se eligió a la clase predicha, la probabilidad de identificación y la tasa de falsos negativos como variable de respuesta. Para el caso de las variables independientes, se tomaron las descritas en la Tabla III, en solitario y en grupos.

El modelo de regresión lineal fue utilizado para conocer si existe una relación lineal entre la probabilidad o clasificación predicha con respecto a las variables descritas anteriormente, pero representando la dependencia de las observaciones dado que en este caso contamos con varias observaciones para un mismo individuo. Para ello, se utilizó la función *'lm'* de la paquetería *'lmtest'* (Zeileis & Hothorn, 2002). A continuación, se describe la fórmula matemática:

$$Y' = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p + \epsilon \quad (10)$$

Donde Y' la variable de respuesta siendo en este caso la probabilidad de identificación del tiburón o la clase (tiburón = 1/no tiburón = 0), β_0 es el intercepto, β_p representan los coeficientes asociados a las variables independientes X_p , y ϵ el error o residuales.

Los GAM's fueron usados para encontrar posibles relaciones no lineales entre la variable de respuesta con respecto a las variables explicativas. Se utilizó el paquete *'mgcv'* en R que es un 'vehículo de cálculo GAM mixto con estimación automática de suavidad' (Wood, 2017), siendo la fórmula del modelo base descrita a continuación:

$$Y' = \beta_0 + f_1(X_1) + f_2(X_2) + f_3(X_3) + \dots + f_p(X_p) + \epsilon \quad (11)$$

Donde Y' es la variable de respuesta siendo en este caso la probabilidad de identificación del tiburón, β_0 es el intercepto (una constante), f_p son las funciones no paramétricas suavizantes del modelo, X_p son las variables independientes siendo en este caso los factores ya categorizados y mencionados anteriormente (Tabla III), y ϵ los residuales. Se utilizó la distribución beta para analizar la variable respuesta de la probabilidad de la identificación, ya que esta distribución es útil al manejar datos proporcionales como estos. En el caso de analizar la clase predicha, se utilizó una distribución binomial.

Se analizaron los residuales de cada modelo desarrollado. Para ello, se ejecutaron diversas pruebas en busca de comprobar si los residuales cumplían con los siguientes supuestos:

- Aleatoriedad. Para ello se contrastaron los valores ajustados contra los residuales en busca de encontrar una dispersión en el gráfico lo que indicaría que cumple con este supuesto. En cambio, si se observa una agrupación o patrón en el gráfico se infiere que no existe aleatoriedad. Esta prueba se realiza de manera cualitativa a falta de una prueba estadística para ello.
- Normalidad. Para ello, se aplicó la prueba de Shapiro-Wilk donde se busca probar si se rechaza o no se rechaza la hipótesis nula, la cual estipula que los residuales tienen una distribución normal. Si el valor p de la prueba es > 0.05 no se rechaza la hipótesis nula, si es < 0.05 se rechaza.
- Homocedasticidad. Se refiere a que los residuales tengan varianzas constantes. Para ello, se aplicó la prueba de Breusch-Pagan donde la hipótesis nula estipula que los residuales tienen varianza constante y, por ende, son homogéneos, esta se cumple si el valor de la prueba es $p > 0.05$. En caso contrario, se rechaza la hipótesis nula y tomamos la hipótesis alterna que indica que no se cumple este supuesto.
- Independencia. Indica que no existe correlación entre los errores. Para ello, se utilizó la prueba de Durbin-Watson para probar la hipótesis nula, la cual estipula que los residuales son independientes. Esta no se rechaza si el valor p de la prueba es > 0.05 . En caso contrario, se rechaza la hipótesis nula y tomamos la hipótesis alterna que indica que no se cumple este supuesto.
- Media cero. Para ello, se calcula la media de los residuales buscando que se aproximen a cero.

Capítulo 3. Resultados

En este capítulo, se presentan los hallazgos derivados de los análisis realizados sobre los datos recolectados en campo y el desempeño de los modelos de inteligencia artificial empleados para identificar y seguir tiburones ballena. Inicialmente, se describen las características y distribución de los avistamientos de tiburones ballena, destacando las condiciones y patrones observados durante los vuelos de dron. Posteriormente, se evalúa y compara el rendimiento de los modelos mediante métricas de desempeño como exactitud, puntaje F1 macro y cobertura de área. Finalmente, se analiza el impacto de las variables ambientales y operativas en la efectividad de los modelos, explorando la capacidad predictiva de los modelos estadísticos utilizados.

3.1 Datos recolectados en campo

Se registraron 23 avistamientos de tiburón ballena mediante vuelos de dron, siendo 17 grabados durante vuelos pilotados desde embarcaciones y 6 pilotados desde tierra (Figura 17, sección a). La semana con mayores registros fue la tercera, del 23 al 28 de octubre (Figura 17, sección b), donde se avistaron 15 tiburones desde embarcación y 10 desde tierra.

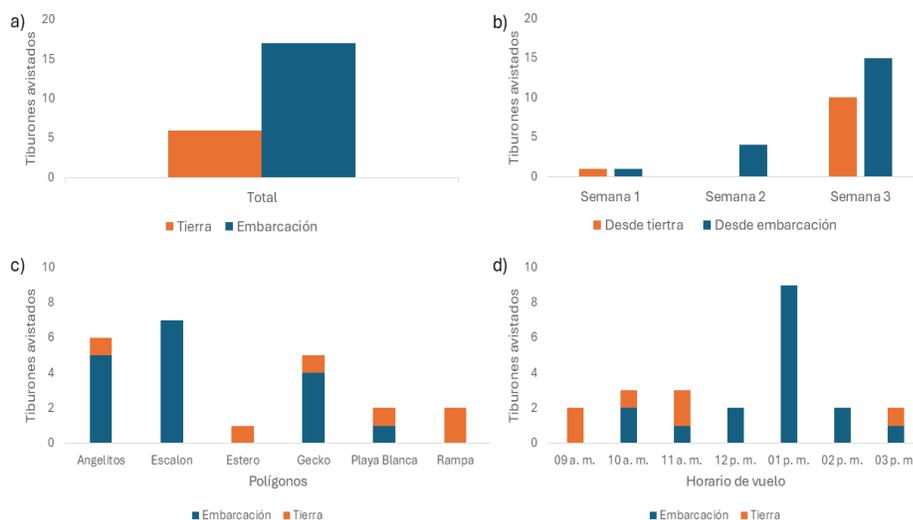


Figura 17. Resultados de la recolección de videos. El color *naranja* representa a los videos grabados mediante el pilotaje de dron desde tierra y el *azul* desde embarcación. a) Cantidad de tiburones avistados según el tipo de vuelo. b) Cantidad de tiburones avistados con respecto a las semanas de trabajo de campo. c) Tiburones avistados con respecto a la zona y tipo de vuelo. d) Tiburones avistados con respecto al horario y tipo de vuelo.

En cuanto a las zonas de observación (Figura 17, sección c), El Escalón ($n = 7$), Angelitos ($n = 6$) y Campo Gecko ($n = 5$) presentaron el mayor número de avistamientos. En cuanto a los horarios de observación (Figura 17, sección d), la mayor presencia de tiburones ballena fue al mediodía (1pm, $n = 9$) para los vuelos pilotados desde embarcaciones, y por la mañana ($n > 2$), entre las 9am y 11 am, para en los vuelos pilotados desde tierra.

3.2 Análisis y comparación del desempeño de los modelos de identificación y seguimiento del tiburón ballena

Con base en la métrica Exactitud (Figura 18) se observa el mejor modelo corresponde al MSP-192O (caja verde) por contar con una media de 0.97 y el peor modelo corresponde a DLC (caja azul) con una media de 0.70. Las primeras cajas, correspondientes a los modelos DLC (color azul) y MSP-192 (color naranja), presentan una mayor variación con un rango de 0.55 (0.40-0.95) y de 0.30 (0.66-0.96), respectivamente. En contraste, para los modelos que utilizan solapamiento de los parches analizados, el rango de valores se reduce a 0.20 (0.80-1), siendo el modelo MSP-192O (caja verde) el que presenta mejores resultados de dispersión con un rango de 0.12 (0.88-1). Esto indica que la aplicación de solapamiento permite una mejora en el rendimiento de los modelos con respecto a la clasificación de las imágenes, sobre todo, al reducir el rango de variación de los resultados.

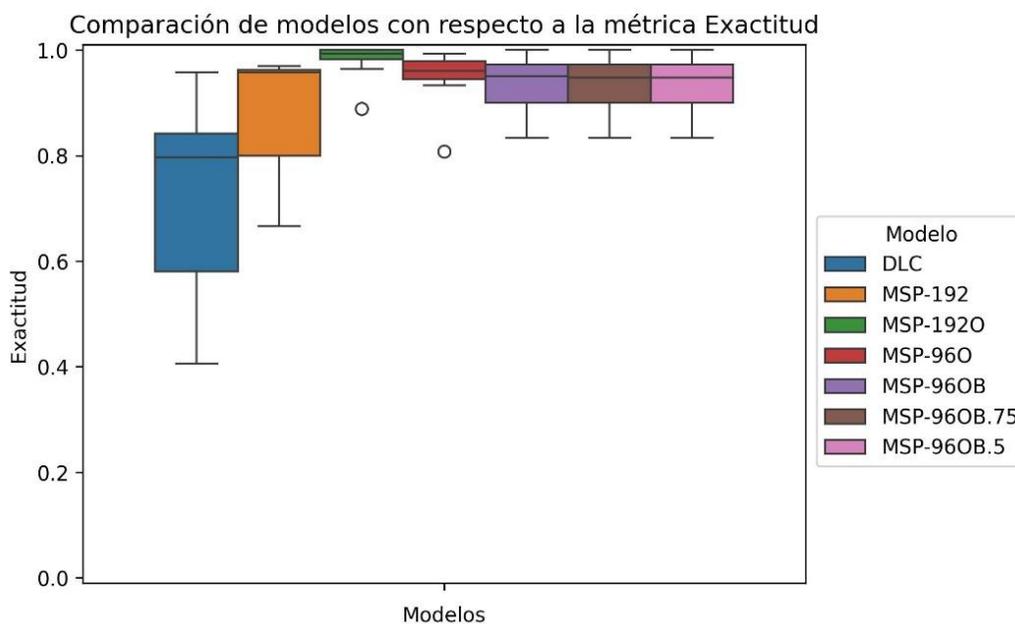


Figura 18. Comparación de modelos mediante la métrica Exactitud.

Debido a que la métrica de Exactitud puede estar sesgada cuando las clases son balanceadas, se recurrió a presentar y sustentar el desempeño de estos modelos con base en la métrica Puntaje F1 macro (Figura 19). Si bien, el desempeño de esta métrica sigue un patrón similar a la evaluada anteriormente, es evidente que la variación en todos los modelos incrementa. Aquellos con un desempeño menor y con alta variación fueron el DLC y el MSP-96O, con un rango de 0.57 (0.38-0.95) y de 0.51 (0.48-0.99), respectivamente. Con respecto a los modelos MSP-192, MSP-96OB.75 y MSP-96OB.5, su desempeño aumenta, pero la variación se mantiene alta, con un rango de 0.36 (0.64-1.00). Los modelos que cuentan con menor variación en el desempeño son MSP-192O, con un rango de 0.25 (0.74-1.00), y MSP-96OB, con un rango de 0.23 (0.77-1.00), pero este último presenta un punto extremo de 0.63. Por lo que, con base en el análisis de esta métrica, el mejor modelo es el MSP-192O, corroborando lo expuesto en la métrica anterior.

Es consistente que la aplicación de solapamiento (enfoque MSP) permite una mejora en el rendimiento de los modelos con respecto a la clasificación de las imágenes, sobre todo, reduciendo el rango de variación de los resultados.

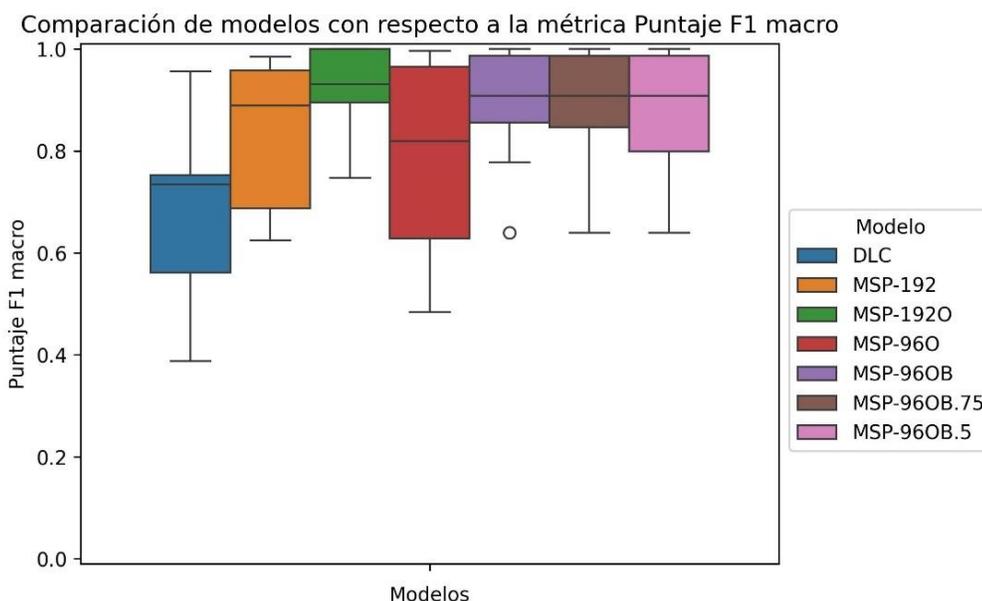


Figura 19. Comparación de modelos mediante la métrica Puntaje F1-macro.

Por ello, en la Figura 20 podemos observar que en los modelos MSP-96OB, MSP-96OB.75 sus predicciones se basan en el reconocimiento de más del 80% del área verdadera, a excepción por dos videos, en los cuales puede reducirse hasta el 77% para el MSP-96OB.75. El modelo MSP-96OB.5 mantiene una cobertura entre el 65% y el 75%, a excepción de dos casos atípicos (60% y 83%). El modelo MSP-192O presenta una

cobertura entre 57% y 83%, a excepción de dos casos (93% y 50%). En el modelo MSP-192 su cobertura ronda 55%-70% pero puede variar hasta el 77%. Por último, los modelos con las coberturas más bajas y con las mayores variaciones fueron el DLC, su principal distribución es entre el 45%-59% con variación hasta 78% y 1 caso atípico (18%), y el MSP-96O, con su distribución principal entre 38%-59% pero con variaciones de 27% hasta 68%. Con respecto a esta métrica, el modelo que cuenta con una mayor cobertura de área durante la predicción es el MSP-96OB.

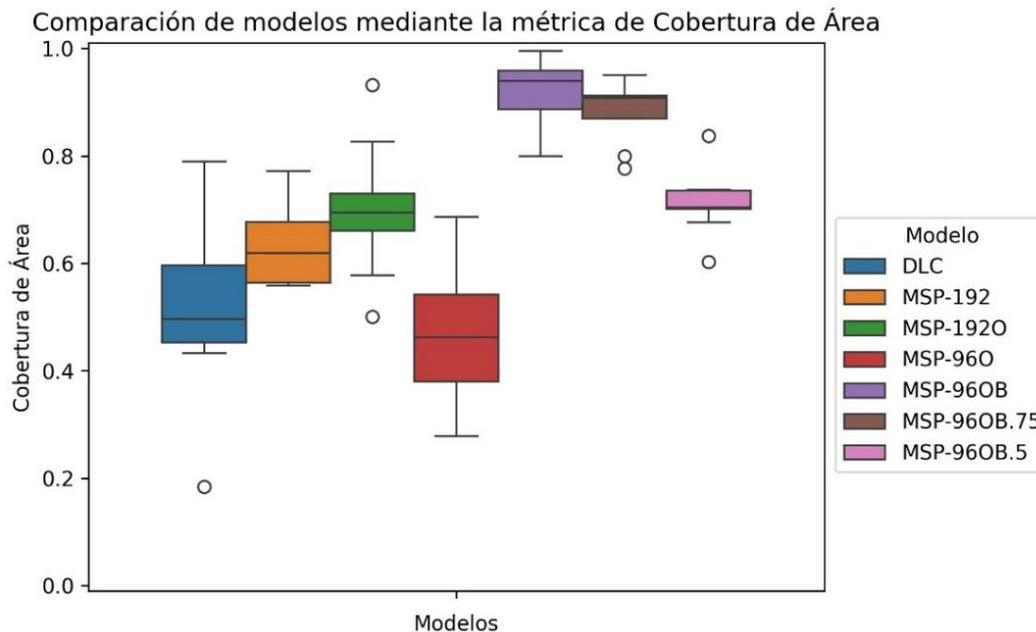


Figura 20. Comparación de modelos mediante la métrica Cobertura de Área.

El desempeño de los modelos puede verse afectado por variables como la altitud de vuelo, el reflejo solar en el mar, la presencia de objetos externos al tiburón, el comportamiento de sumergimiento del tiburón y movimientos del dron, además del enfoque empleado, el tamaño y el solapamiento de los parches. En la Figura 21 se puede notar de manera visual, la afectación de estas variables en el desempeño de los modelos. El escenario que se muestra en la primera fila se destaca como “el ideal” por tener un fondo claro y uniforme, sin afectación de objetos externos ni reflejos del sol en el mar, deriva en una correcta identificación en todos los modelos. En la segunda fila se muestra un escenario más retador, donde el tiburón ballena se encuentra sumergido, sorprendentemente solo el modelo DLC no lo identifica mientras que los demás lo identifican en su totalidad o al menos una parte de su cuerpo. La tercera fila presenta una de las escenas más retadoras para todos los modelos ya que intervienen variables como la presencia de nadadores y de la embarcación, el movimiento del agua y el sumergimiento del tiburón ballena, lo que generó que solo en los modelos con los parches más grandes (MSP-192 y MSP-192O) lograron identificarlo.

MSP-96OB, MSP-96OB.75 y MSP-96OB.5). Por ende, se determina que el mejor modelo es el MSP-192O, siendo consistente con lo expuesto anteriormente.

Tabla IV. Desempeño de los modelos con base a cada video del conjunto de prueba.

Video	Modelo	Exactitud	Puntaje F1 Macro	Cobertura de Área	Video	Modelo	Exactitud	Puntaje F1 Macro	Cobertura de Área
22.2_PlayaBlanca-PM-1	DLC	0.825	0.562	0.790	25.1_Tb13	DLC	0.861	0.925	0.496
	MSP-192	0.920	0.660	0.706		MSP-192	0.957	0.978	0.619
	MSP-192O	0.988	0.747	0.717		MSP-192O	1.000	1.000	0.660
	MSP-96O	0.980	0.628	0.463		MSP-96O	0.993	0.997	0.380
	MSP-96OB	0.900	0.640	0.887		MSP-96OB	0.973	0.986	0.959
	MSP-96OB.5	0.900	0.640	0.735		MSP-96OB.5	0.973	0.986	0.676
	MSP-96OB.75	0.900	0.640	0.869		MSP-96OB.75	0.973	0.986	0.922
25.1_Tb2	DLC	0.406	0.387	0.532	27.1_Angelitos-AM-2	DLC	0.958	0.956	0.653
	MSP-192	0.962	0.890	0.564		MSP-192	0.960	0.958	0.772
	MSP-192O	0.993	0.910	0.501		MSP-192O	1.000	1.000	0.827
	MSP-96O	0.808	0.484	0.278		MSP-96O	0.970	0.892	0.542
	MSP-96OB	0.957	0.909	0.939		MSP-96OB	1.000	1.000	0.996
	MSP-96OB.5	0.957	0.909	0.714		MSP-96OB.5	1.000	1.000	0.702
	MSP-96OB.75	0.957	0.909	0.908		MSP-96OB.75	1.000	1.000	0.912
25.1_Tb7	DLC	0.797	0.643	0.597	27.3_Estero-AM-1	DLC	0.842	0.737	0.453
	MSP-192	0.968	0.887	0.675		MSP-192	0.800	0.688	0.677
	MSP-192O	0.995	0.931	0.679		MSP-192O	0.982	0.895	0.932
	MSP-96O	0.945	0.620	0.402		MSP-96O	0.961	0.657	0.648
	MSP-96OB	0.950	0.856	0.911		MSP-96OB	0.947	0.885	0.941
	MSP-96OB.5	0.921	0.799	0.704		MSP-96OB.5	0.947	0.885	0.837
	MSP-96OB.75	0.946	0.847	0.870		MSP-96OB.75	0.947	0.885	0.909
25.1_Tb9	DLC	0.604	0.753	0.466	28.4_Rampa-PM-1	DLC	0.500	0.486	0.185
	MSP-192	0.969	0.984	0.562		MSP-192	0.667	0.625	0.600
	MSP-192O	1.000	1.000	0.578		MSP-192O	0.889	0.800	0.729
	MSP-96O	0.933	0.965	0.321		MSP-96O	0.947	0.819	0.687
	MSP-96OB	0.994	0.997	0.987		MSP-96OB	0.833	0.778	0.800
	MSP-96OB.5	0.994	0.997	0.703		MSP-96OB.5	0.833	0.778	0.737
	MSP-96OB.75	0.994	0.997	0.951		MSP-96OB.75	0.833	0.778	0.800
25.1_Tb12	DLC	0.581	0.735	0.432					
	MSP-192	0.800	0.889	0.558					
	MSP-192O	0.964	0.982	0.695					
	MSP-96O	0.979	0.989	0.462					
	MSP-96OB	0.871	0.931	0.842					
	MSP-96OB.5	0.871	0.931	0.603					
	MSP-96OB.75	0.871	0.931	0.777					

3.3 Análisis de las variables que afectan la efectividad de los modelos

Las variables presentes en cada uno de los videos se describen cualitativamente en la Tabla V. Los primeros seis videos enlistados fueron pilotados desde embarcaciones, los cuales tienen una duración entre 30 segundos hasta 4 minutos, se suscitaron principalmente por la tarde, con un brillo solar de medio a alto, turbidez de media a alta, así como con una afectación visible mayor del brillo solar, rotación del dron, sumergimiento del tiburón y presencia de objetos externos. Los últimos tres videos de la tabla corresponden a los piloteados desde tierra, los cuales son de duración corta (5 s a 24 s), fueron desempeñados tanto en la mañana como en la tarde, con un brillo solar nulo o bajo, turbidez variante (baja-media-alta), con altitud de vuelo alto, con presencia de objetos externos, pero sin representación de sumergimiento del tiburón.

Tabla V. Características de los videos analizados.

Nombre del video	Duración de video (min:s)	Variables							
		Horario de vuelo	Brillo Solar	Fuente Solar	Turbidez	Altitud de Vuelo	Rotación del dron	Sumergimiento	Presencia de objetos externos
22.2_PlayaBlanca-PM-1	01:19	Tarde	Bajo-Medio	Cambiante	Media	50 m	Si	Si	Si
25.1_Tb2	01:08	Tarde	Medio-Alto	Sin cambios	Media	50 m	No	Si	Si
25.1_Tb7	04:01	Tarde	Medio-Alto	Cambiante	Baja	50 m	Si	Si	Si
25.1_Tb9	02:33	Tarde	Medio-Alto	Cambiante	Alta	50 m	Si	Si	Si
25.1_Tb12	00:30	Tarde	Bajo-Medio	Cambiante	Alta	50 m	Si	Si	Si
25.1_Tb13	00:35	Tarde	Medio	Sin cambios	Alta	50 m	No	Si	Si
27.1_Angelitos-AM-2	00:24	Mañana	Nulo	Sin cambios	Baja	100 m	No	No	Si
27.3_Estero-AM-1	00:18	Mañana	Bajo	Cambiante	Media	100 m	No	No	No
28.4_Rampa-PM-1	00:05	Tarde	Bajo	Sin cambios	Alta	100 m	No	No	Si

A partir de la clasificación de estas variables, según la Tabla III (véase en la sección 2.6), y la aplicación de modelos de regresión lineal (RL) y modelos aditivos generalizados (GAM) se observó lo siguiente. En los RL

(Tabla VI), las variables de brillo solar y horario de vuelo fueron las más significativas tanto para la probabilidad de identificación como la clase. Sin embargo, solo explican entre el 14% y el 20% de la variabilidad del modelo.

Tabla VI. Mejores modelos de regresión lineal.

Modelo	Valor p		R ²	Supuesto cumplidos
	θ_0	θ_x		
1) Probabilidad = 0.57 + 0.47 brillo solar ₁ + 0.41 brillo solar ₂ + 0.39 brillo solar ₃ - 0.28 video estero - 0.13 video playa blanca - 0.57 video rampa - 0.25 video Tb12 + ϵ	<0.001	<0.001	0.2	Homocedasticidad (p=0.5) y media cero, pero gráficamente se percibe alta correlación
2) Clase = 0.58 + 0.48 brillo solar ₁ + 0.40 brillo solar ₂ + 0.37 brillo solar ₃ - 0.38 video estero - 0.15 video playa blanca - 0.56 video rampa - 0.29 video Tb12 + ϵ	<0.001	<0.001	0.14	Homocedasticidad (p=0.5) y media cero, pero gráficamente se percibe alta correlación

En los GAM (Tabla VII), el horario de vuelo, la altitud, la turbidez y el sumergimiento fueron las variables más significativas a la hora de determinar la probabilidad de la identificación. En cuanto a la determinación de clase, la variable de agentes externos fue la de mayor influencia. Sin embargo, la devianza para ambos modelos no fu significativa (41.1 y 21.4, respectivamente).

Tabla VII. Mejores Modelos Aditivos Generalizados (GAM).

Modelo	Distribución	Valor p		R ²	Devianza	AIC	Supuestos cumplidos
		B_0	B_x				
1) Probabilidad = 3.09 - 0.38 horario de vuelo ₃ - 0.70 horario de vuelo ₄ - 1.56 horario de vuelo ₅ - 2.27 altitud ₂ - 0.33 sumergimiento ₁ - 0.41 video Tb7	Beta	<0.05	<0.05	0.149	41.1	-21088.99	Ninguno
2) Clase = 1.22 agentes externos ₁ + 1.56 video playa blanca + 2.23 video Tb13 + 1.87 video Tb7 + 3.86 video Tb9	Binomial	<0.05	<0.05	0.157	21.4	296.6508	Ninguno

A pesar de que en la Tabla VI y la Tabla VII se presentan los mejores modelos estadísticos, basados en la significancia de sus variables, no se identificó un modelo estadístico que cumpliera plenamente con los supuestos, en específico, de normalidad e independencia. Esta violación de supuestos se evidenció al analizar gráficamente los residuos (Figura 22), donde la gráfica Q-Q mostró que los residuos no siguen la línea esperada de normalidad, y en la comparación de Residuales vs Ajustados se observaron patrones en lugar de una distribución aleatoria. Por lo tanto, se infiere que estos modelos no capturan bien la relación entre las variables con la predicción de los modelos de IA.

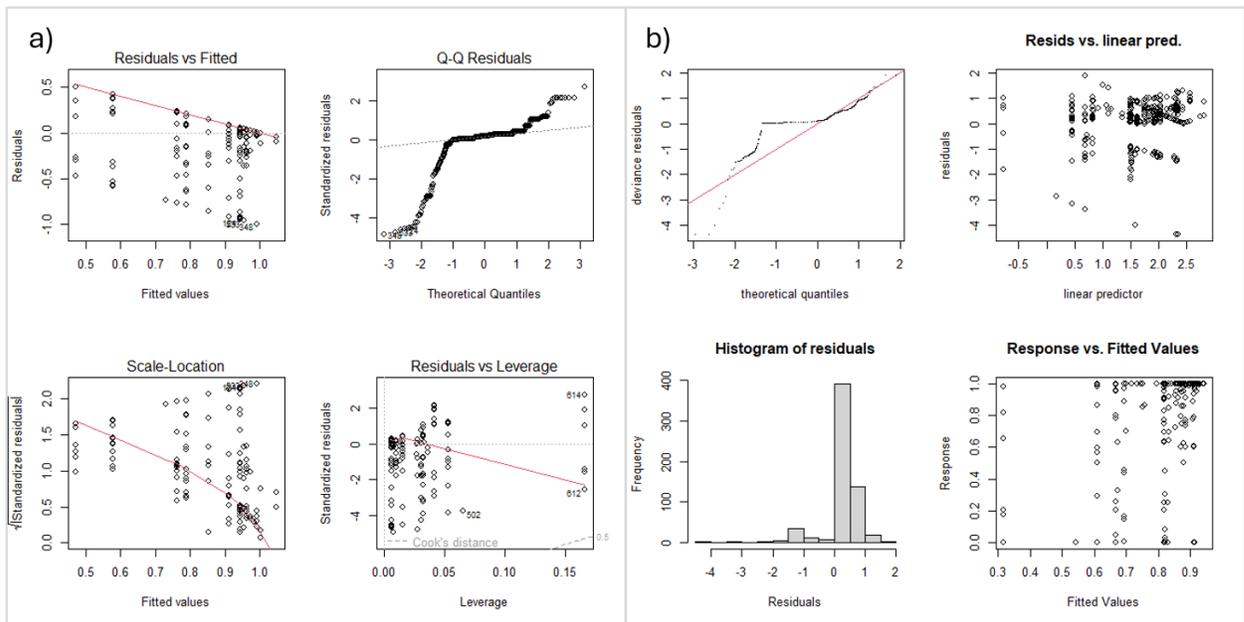


Figura 22. Gráficos de distribución de los residuos. **a)** Modelo RL #1, **b)** Modelo GAM #1.

Los resultados obtenidos en este capítulo destacan el impacto positivo de la aplicación de solapamiento en los modelos de predicción, mejorando su precisión y reduciendo la variabilidad de los resultados. Entre los modelos evaluados, MSP-1920 se posiciona como el más robusto y consistente bajo diversas métricas de desempeño. Si bien las variables ambientales y operativas como la altitud de vuelo, el brillo solar y la turbidez, presentan indicios de afectación en la efectividad de los modelos, los análisis estadísticos realizados evidenciaron un incumplimiento de los supuestos de normalidad e independencia. Esto sugiere que la relación entre dichas variables y las predicciones de los modelos de IA podría no ser lineal ni adecuadamente representada. Estos hallazgos destacan la necesidad de explorar enfoques más complejos y adaptativos para modelar de manera precisa la influencia de las condiciones ambientales en el desempeño de los modelos, sentando una base sólida para futuros estudios y aplicaciones en el monitoreo de tiburones ballena en Bahía de los Ángeles y otras áreas de agregación de la especie.

Capítulo 4. Discusión

En esta tesis se evaluó el uso de imágenes obtenidas por dron y su análisis a través de modelos predictivos de inteligencia artificial (IA) como una forma de poder contar con una metodología más eficiente de monitoreo de tiburón ballena en Bahía de los Ángeles. Así mismo, se comparó la efectividad de identificación de los organismos utilizando dos métodos que permitieran la automatización de las actividades de monitoreo.

Con base en 23 avistamientos aéreos de tiburón ballena y aplicando técnicas de IA se logró clasificar e identificar la presencia de tiburones ballena. Posteriormente se comparó la efectividad entre los modelos y se seleccionó al mejor modelo con base en las métricas comúnmente utilizadas en la literatura como Puntaje F1-score, Exactitud, así como una de desarrollo propio, llamada Cobertura de Área (CA). De esta manera, se demostró que los modelos predictivos bajo el enfoque MSP muestran un mejor desempeño que DLC, incluso al aplicar solapamiento de imágenes es posible reducir la variación y aumentar el rendimiento, siendo el modelo MSP-1920 el mejor modelo con un Puntaje F1 macro promedio de 0.91 y CA promedio de 0.65. Valores altos en la métrica CA implican una mejor predicción, ya que esta métrica compara el cuadrante generado a partir de la predicción del modelo (área predicha) con el generado a partir de la posición real del tiburón ballena en la imagen (área verdadera). Sin embargo, cuando los valores son medios, como en el caso de MSP-1920, no siempre representa una mala predicción, sino que el modelo es capaz de reconocer al tiburón incluso con solo observar una parte de su cuerpo.

Al aplicar diferentes modelos de análisis estadístico se pudo discernir que las variables físicas presentes durante la recolección de los datos (ambientales como la condición de la luz, logísticas como la altitud de vuelo y relativas al comportamiento del tiburón como el sumergimiento, por mencionar algunas) no manifiestan una afectación directa en la identificación de tiburones por parte de los modelos predictivos. Cabe mencionar que, durante la etapa de recolección de datos, se encontró una alta cantidad de avistamiento de tiburones ballena alrededor de la 1pm (mediodía), horario que concuerda con lo expuesto por Nelson & Eckert (2007) aunque Peregrín-Tovar (2014) encontró un mayor avistamiento en el horario de 3pm-6pm. Cabe mencionar que, en ambos estudios los avistamientos fueron realizados desde embarcaciones. Sin embargo, en el caso de este trabajo, esta alta recurrencia de avistamientos durante ese horario se encuentra mayormente relacionado con la logística de los recorridos de monitoreo a bordo de una embarcación, pero con la confirmación a partir del vuelo de dron. Esto no pudo ser comparado con el método de pilotaje desde tierra ya que el avistamiento de organismos fue escaso y no fue realizado

durante dichos horarios ya que se evitaban las condiciones altas de reflejo del sol ya que reducen la visualización de los objetos presentes en la imagen, incluidos los tiburones ballena. Así mismo, la zona con mayor avistamiento de tiburones ballena fue El Escalón, seguido por Los Angelitos y Gecko, lo que corresponde a la zona oeste de la SUPTB y concuerda con lo reportado por Peregrín-Tovar (2014) y el grupo comunitario a cargo del monitoreo de la especie.

4.1 Virtudes y oportunidades de mejora del modelo

Las arquitecturas de redes neuronales convolucionales han demostrado ser una herramienta útil para la identificación de organismos marinos en imágenes, siendo aplicadas en el conteo de tiburones en imágenes submarinas (Villon et al., 2024), la detección de tiburones en imágenes de dron para la protección de bañistas (Sharma et al., 2022), y la identificación en tiempo real de especies de tiburón con base en imágenes de drones (Purcell et al., 2022). En este caso, estas arquitecturas fueron utilizadas para la identificación y seguimiento de tiburones ballena mediante videos grabados por dron. El mejor modelo, MSP-1920, exhibió un alto grado de precisión en cuanto a la identificación de tiburones ballena (Puntaje F1-score macro = 0.91), a pesar de las condiciones adversas a las cuales se enfrentó como la alta turbidez de la bahía asociada a la alta productividad primaria (Lavaniegos et al., 2012), el alto brillo solar presente en la mayoría de los videos relacionado con la logística del trabajo de campo y la presencia tanto de fauna asociada al tiburón como de nadadores y embarcaciones cercanas al organismo durante la grabación de los videos.

Dentro de las fortalezas del modelo MSP-1920 se encuentra su alta efectividad, ya que logró identificar el 85% de las imágenes correspondientes a cada clase (tiburón y no tiburón) del total de imágenes inspeccionadas (sensibilidad = 0.85) y el 96% de las veces que se identificó una clase por el modelo, esta clasificación fue correcta (precisión = 0.96). Ante ello, la métrica F1-score nos proporciona una descripción más precisa del desempeño del modelo, ya que su cálculo proviene del equilibrio de la precisión y la sensibilidad. El valor del Puntaje F1 macro para el mejor modelo fue de 0.91, lo que indica que es eficaz al identificar correctamente a los tiburones ballena y minimiza la clasificación incorrecta con respecto a otras especies u objetos presentes en las imágenes analizadas. Es importante recalcar que, como cualquier otro modelo que utilice CNN, a medida que se amplíe el conjunto de datos de entrenamiento y se representen la mayor diversidad de escenarios reales, el modelo continuará mejorando su capacidad de generalización y aumentará su poder predictivo ante nuevos datos (Colomer, 2024). Es importante recalcar que la predicción provista por el modelo fue de cada 0.1 segundo de video analizado y solo se generó el área

verdadera para cada segundo del video (cada 30 fotogramas). Por lo que, al interpolar la predicción de los datos intermedios con el ajuste temporal y compararlos con el área verdadera del fotograma más cercano del cual se dispone de información, cambios rápidos en las condiciones de operación, ambiental o de comportamiento del tiburón podrían no verse reflejadas, afectando en la clasificación del modelo.

A diferencia del mejor modelo (MSP-192O), aquellos con tamaño de parche 96x96 píxeles (MSP-96OB, MSP-96OB.75 y MSP-96OB.5) sobresalen con respecto a los otros modelos por que puede identificar al tiburón en un área más precisa, lo que reduce las clasificaciones incorrectas debidas a la injerencia objetos distintos al tiburón que podrían compartir espacio cuando existen cuadrantes de mayor tamaño. Esto se denota con respecto a la métrica denominada como “Cobertura de área” ($CA > 0.80$), un símil a la métrica “Intersection over Union” (IoU) utilizada en tareas de detección de objetos, la cual permite rectificar la clasificación predicha por el modelo mediante la identificación y la comparación del porcentaje cubierto por el parche (área predicha) con respecto al ocupado por el tiburón (área verdadera). Si el parche sobrepasa el umbral de 10% de traslapo ($CA > 0.1$), se considera que corresponde al tiburón, lo que reduce los falsos positivos. En cambio, el modelo MSP-192O y los escenarios (videos) donde esta métrica cuenta con puntuaciones bajas (CA promedio= 0.65), el presente modelo sigue siendo efectivo, ya que indica su capacidad para identificar al tiburón con al menos la mitad de su cuerpo de su cuerpo presente en el parche, aunque también puede ser debido al solapamiento de sus parches que genere que el tiburón se encuentre entre dos parches contiguos.

El buen desempeño de los modelos está relacionado con características particulares de los tiburones ballena. Entre ellas, se puede destacar su comportamiento más común y activo durante el día: el forrajeo, comportamiento asociado a la alimentación por medio de la filtración de cantidades grandes de agua mientras el organismo nada lentamente en la superficie o justo debajo de ella (Montero-Quintana et al., 2020) en zonas en donde se concentra una alta abundancia de sus presas (plancton y pequeño necton) durante la temporada de agregación en la bahía (Hernández-Nava & Álvarez-Borrego, 2013; Lavaniegos et al., 2012). Este comportamiento genera una alta accesibilidad para localizar y seguir al tiburón ya que sucede en la superficie, la mayor parte del día y puede continuar hasta que se presente una perturbación, mayormente relacionada de su interacción con humanos, lo que provoca su escape mediante el incremento en su nado o sumergimiento (Montero-Quintana et al., 2020).

Aunado al comportamiento, el tiburón fue avistado nadando solo o en cercanías de aves y tortugas, siendo esta baja presencia de fauna asociada lo que facilitó que el modelo predijera correctamente la presencia del tiburón, ya que no existieron objetos cercanos que pudieran generar confusión en la predicción del

modelo. Se ha reportado la presencia de megafauna marina (delfines, atunes, lobos marinos, ballenas y tiburones) junto con tiburón ballena impulsado por la presencia de bancos de peces de carnada tanto en la bahía (Lenihan, 2023), sin embargo, esto podría no ser un factor de afectación o confusión para el modelo ya que no presentan el mismo tamaño, morfología y color que el tiburón ballena. Esto representa una gran ventaja de nuestro trabajo ya que, en animales con comportamiento gregario, con similitud con el entorno y fauna asociada (Dujon et al., 2021), la probabilidad de detección disminuye, como es el caso del monitoreo de focas, tortugas marinas y aves.

Entre las limitaciones del presente trabajo se encuentra la autonomía del Dron. Los drones comerciales presentan limitaciones en la extensión y duración del vuelo debido a la limitación de la carga de la batería y la distancia de operación (Hodgson et al., 2017), siendo un desafío en estudios que requieren monitoreo de grandes áreas o zonas alejadas de la costa (Hensel et al., 2018). Si bien, cubrir una gran cantidad de área durante una menor cantidad de vuelos puede ser contrarrestado aumentando la altitud de vuelo (Nolan et al., 2021), esto reduce la capacidad de monitorear áreas más profundas y de identificar correctamente a las especies o simplemente no identificarlas (Hensel et al., 2018). Sumado a ello, el cambio en la altitud de vuelo y el tamaño de los individuos puede provocar problemas en la identificación mediante la aplicación de arquitecturas de CNN debido al tamaño angular de los objetos, entre más pequeños sean con referencia a la imagen, su detección y clasificación se vuelve más complicada (Purcell et al., 2022). Para ello, se recomienda que los objetos a identificar representen un tamaño mayor a 80 píxeles de la totalidad de la imagen y, para aquellos que sean más pequeños se recomienda monitorearse a una altitud no mayor a 50m para lograr la detección y, posteriormente, se puede reducir la altitud a 25 m para mejorar la clasificación (Purcell et al., 2022). En el presente trabajo, los casos de prueba donde la altitud de vuelo fue de 100m (pilotados desde tierra) contaron con un desempeño mejor en la identificación por parte de todos los modelos bajo el enfoque MSP, esto puede deberse a que fueron videos con condiciones ambientales y de vuelo ideales (bajo brillo solar, sin rotación del dron y sin perturbación del animal), de corta duración y con baja disponibilidad de datos (menor cantidad de videos colectados). Sin embargo, para el enfoque DLC, el modelo fue capaz de localizar a los tiburones 100m, pero confundió los puntos clave correspondientes a la cabeza con los de la cola y viceversa.

Al examinar la relación entre las variables identificadas y la identificación de tiburones ballena mediante la aplicación de modelos lineales (RL y GAM; Figura 22), se muestra una falta de ajuste atribuible a la baja variabilidad de los datos. Aunque se cuenta con un elevado número de observaciones, estas solo se distribuyen entre nueve individuos. Además, algunos de estos individuos tienen más observaciones que

otros, dependiendo de la duración de las grabaciones. Por lo que, es posible que los modelos estadísticos aquí desarrollados no representen adecuadamente la relación entre las variables analizadas.

Si bien, al desarrollar las pruebas estadísticas (RL y GAM) no encontramos una relación lineal entre las variables ambientales con la predicción del modelo, empíricamente se pudieron reconocer ciertos aspectos que pueden afectar a los modelos de clasificación de imágenes mediante el uso de técnicas de IA. Uno de ellos es la presencia de reflejo solar en el mar que afecta tanto a la predicción del modelo como el pilotaje del dron. Se ha documentado que las horas de mayor actividad registrada del tiburón ballena en Bahía de los Ángeles se da durante el día (Nates-López, 2015) por lo cual, es un factor importante por considerar en la logística de los estudios de tiburón ballena mediante uso de drones. Como alternativa a este problema se encuentra el uso de cámaras multiespectrales y filtros polarizados en el lente del dron para mejorar la visualización de objetos en el ambiente marino (Colefax et al., 2023). Una alternativa a los filtros polarizados puede ser el ajuste de los ángulos de la cámara, ya que permite disminuir o anular la incidencia del reflejo solar en la cámara. Así mismo, se puede complementar con la aplicación de técnicas de procesamiento de imágenes aéreas previo a la incorporación en modelos de IA, como la normalización de los colores de fondo (Mitchell et al., 2022) o el uso de filtros espectrales en ciertas longitudes de onda que permiten diferenciar el cuerpo de los animales del resto del paisaje (Purcell et al., 2022). Sin embargo, el uso de filtros debe ser en cámaras diseñadas para ello, ya que Colefax y colaboradores (2023) encontraron que aplicar un tratamiento a las imágenes RGB de cámaras comerciales y accesibles (GoPro) mediante filtros de banda ancha de color verde (525 y 550 nm) no representó una mejora en la detección de animales marinos durante vuelos de dron.

Otra variable importante es la alta turbidez del mar en la bahía, derivada de la alta productividad de la zona (Lavaniegos et al., 2012), lo cual puede afectar la detección del tiburón ya que aguas turbias o con alta sedimentación pueden reducir la visibilidad (Hodgson et al., 2017). Esto además se relaciona con el sumergimiento subsuperficial de los individuos pueden ser poco o nada visibles (Hensel et al., 2018). Esto impacta en estudios de abundancias y densidades poblacionales debido a una subrepresentación, omisión o confusión de las especies (Gray et al., 2018); en estudios de fotogrametría, ya que se disminuye la nitidez del contorno del cuerpo y dificulta que el modelo distinga los píxeles correspondientes al individuo de los del entorno circundante, fenómeno conocido como “floración” (Dujon et al., 2021; Whitehead et al., 2022; Rex et al., 2024).

No obstante, a pesar de que estas dificultades se encuentran presentes en los datos utilizados para el entrenamiento y prueba de los modelos, aquellos bajo el enfoque MSP obtuvieron rendimientos buenos

y aceptables (Puntaje F1 macro > 0.78), lo que demuestra el valor de esta propuesta al probar los modelos en condiciones adversas, que probablemente reflejan la realidad que se enfrentaría al implementar un modelo de monitoreo de bajo costo.

4.2 Comparación del mejor modelo respecto a otros estudios

El mejor modelo desarrollado en esta tesis, denominado 'MSP-1920', demostró desempeños comparables e incluso superiores a los reportados para modelos aplicados al monitoreo de diversas especies en ambientes marinos silvestres como por ejemplo tortugas marinas, distintas especies de tiburones y otros peces. Este modelo usa la arquitectura de una DenseNet121, una red compuesta por capas completamente conectadas que permiten un mejor flujo de información y disminuye el problema de desvanecimiento de gradientes (Huang et al., 2016). Su desempeño fue óptimo alcanzando un puntaje F1 macro de 0.91, una cobertura de área de 0.70, una precisión macro de 0.96 y una sensibilidad macro de 0.85 (Tabla V, Sección 4.1). Además, mantuvo estos altos niveles de precisión y estabilidad frente a variables ambientales adversas, como condiciones de alto brillo solar, turbidez, y la presencia de agentes externos, incluyendo personas, fauna y objetos, sin requerir procesamiento previo de las imágenes. Estos resultados subrayan la robustez del modelo y su capacidad para operar eficientemente en condiciones reales, caracterizadas por incertidumbre y variabilidad.

No obstante, existen modelos que utilizan arquitecturas más simplificadas en busca de generar una menor demanda de recursos computacionales que les permita realizar actividades de monitoreo en tiempo real o como una alternativa para proyectos con recursos limitados. Ante esto, se han propuesto modelos con aplicación a seguridad de playas que realizan identificación de tiburones en tiempo real mediante arquitecturas existentes y simples como la Small U-Net (Barbelian et al., 2021), la cual está compuesta por 23 capas convolucionales, 5 capas maxpooling y 5 capas de concatenación con 2 millones de parámetros que puede procesar imágenes de Ultra Alta Definición (UHD – 4k) con un desempeño de puntaje f1 de 0.862 para la identificación de distintas especies de tiburones. Así mismo, se han implementado modelos simplificados para la identificación de tortugas marinas que puedan aplicados en computadoras austeras mediante el desarrollo de arquitecturas propias. Dujon y colaboradores (2021) desarrollaron una arquitectura con una estructura de convolución-convolución-pooling que presenta una sensibilidad de 0.75 y una precisión de 0.20. Por su parte, la arquitectura de Gray y colaboradores (2018) cuenta con dos capas convolucionadas intercaladas con capas max pooling, seguidas de dos capas completamente

conectadas y una capa de salida softmax, la cual mostró un desempeño de precisión del 0.16 y una sensibilidad del 0.76.

Si bien, la utilización de modelos más simples, en relación con la DenseNet121 utilizada en este trabajo, puede no ser adecuada para tareas más complejas aun cuando muestran desempeños buenos para su uso en tiempo real y aceptables para su uso en monitoreos con recursos limitados. La red Small U-Net es más compacta y menos compleja que la DenseNet121 al contar con un menor número de parámetros y el empleo de capas de concatenación en lugar de capas completamente conectadas (Barbelian et al., 2021) lo que reduce significativamente el poder computacional necesario y constituye una ventaja para resolver tareas en tiempo real. No obstante, los modelos basados en la arquitectura DenseNet121 presentan ventajas significativas para abordar tareas más complejas (Huang et al., 2016), como podría ser el análisis morfométrico. Estas ventajas se derivan de su elevada profundidad, caracterizada por la presencia de múltiples capas y completamente conectadas, permitiendo un mayor y mejor flujo de la información y, con ello, la captura de características más detalladas y complejas en comparación con la Small U-Net (Huang et al., 2016). Este nivel de detalle es particularmente relevante en el monitoreo del tiburón ballena, especialmente bajo condiciones adversas, como niveles elevados de brillo solar o cuando los individuos se encuentran en la zona subsuperficial de la columna de agua, donde los desafíos asociados a la detección y el seguimiento se vuelven más pronunciados.

El modelo desarrollado por Dujon y colaboradores (2021), presenta un desempeño deficiente ya que el modelo identifica muchas tortugas marinas (sensibilidad de 0.75) pero muchas de ellas son falsas detecciones (precisión baja de 0.2). La afectación del desempeño de este modelo se debe a que la simplicidad de la red no permite hacer una correcta identificación en escenas grabadas a altitudes de vuelo fuera de los 30-50 m, con un posicionamiento a mayor profundidad en la columna de agua o por la presencia de un alto grado de turbidez y reflejo del sol en el mar, incluso al aplicar técnicas de aumento de datos (rotando y volteando las imágenes). Factores que no fueron limitadores para el correcto desempeño del modelo desarrollado en esta tesis, debido a la robustez que confiere el uso de una red más compleja y pre-entrenada como la DenseNet121, así como por la utilización de más técnicas de aumento de datos (cambio de color: saturación, brillo y contraste), lo que permitió representar una mayor variedad de ambientes donde se podría presentar el tiburón ballena.

Caso similar sucede con el modelo desarrollado por Gray et al. (2018) para la identificación de tortugas marinas. Este modelo exhibió resultados similares con sensibilidad de 0.765 y una precisión de 0.16, permitiendo mejorar la identificación entre un 8% y un 9% con respecto a la revisión manual. Este bajo

rendimiento se deriva tanto de la simplicidad de la arquitectura seleccionada como de la baja cantidad de datos entrenamiento (~10k) y con alto desbalance de las clases (10:1, negativos-positivos). Si bien, nuestro modelo exhibe mejores resultados debido al uso de una red más compleja, también intervienen factores como el uso de aprendizaje por transferencia, el uso de técnicas de aumento de datos, y que los datos están balanceados que permitieron una correcta la identificación de tiburones ballena.

Existen modelos que exhiben mejores resultados con arquitecturas más complejas como el desarrollado por Kaur y colaboradores (2024). Este modelo utiliza una arquitectura híbrida entre una red neuronal convolucional (CNN) y una máquina de vectores de soporte (SVM) para analizar imágenes de dron e identificar vida silvestre correspondiente a cinco grandes categorías: mamíferos, aves, reptiles, anfibios y peces. La red fue entrenada con una base de datos compuesta por 3,397 imágenes y mostró un desempeño alto con puntaje F1 mayor al 0.94 para todas las clases. Si bien, su desempeño fue mayor al exhibido por el mejor modelo de este trabajo (puntaje F1 macro de 0.91), presenta dificultades por las condiciones de iluminación y ángulo de la cámara durante la captura, lo cual no fue exhibido por el modelo presente en esta tesis. La base de datos utilizada en esta tesis muestra una variedad de condiciones ambientales (turbidez y brillo solar alto), presencia de agentes externos (fauna, nadadores y embarcaciones) así como relacionados con la logística de los vuelos (brillo solar, rotación dron y altitud de vuelo) las cuales no exhiben una afectación directa en el desempeño del modelo. Si bien, la adición de SVM aumenta la capacidad discriminativa del modelo, representa un mayor coste computacional para el presente proyecto que podría ser solventado, en su caso, con un incremento de los datos de entrenamiento. Esto puede resolverse con el establecimiento de un monitoreo constante del tiburón ballena mediante usos de dron, así como por la colaboración entre los distintos usuarios de la bahía como habitantes, prestadores de servicio, turistas y científicos que realicen legalmente grabaciones de dron en la zona.

Dentro de las ventajas de nuestro modelo se encuentra el uso de aprendizaje por transferencia mediante la utilización de una red ampliamente reconocida en tareas de clasificación de imágenes (Elharrouss et al., 2024) como lo es la DenseNet121, lo que permite contar con una gran variedad de objetos representados en imágenes recopilados de distintas fuentes reduciendo así el impacto de sesgos inherentes a los datos (Munappy et al., 2022). Así mismo, la incorporación de técnicas de incremento de datos (geométricas y de color) que permitieron resolver la cuestión de la escasez de datos disponibles y proporcionar un modelo con un alto rendimiento sin necesidad de una arquitectura extremadamente compleja (Mumuni & Mumuni, 2022).

4.3 Ventajas y oportunidades de los enfoques implementados: DLC vs MSP

En esta investigación se utilizaron dos enfoques para el entrenamiento de los modelos: implementando un software el uso libre DeepLabCut (DLC) y mediante la programación del proceso en su totalidad para la implementación de modelos de CNN (MSP), las cuales conllevan ciertas ventajas y desventajas en su uso.

En el caso de los DLC, sus ventajas están relacionadas con la facilidad de realizar todas las fases relacionadas con el desarrollo de un modelo de aprendizaje profundo debido a que utiliza una interfaz que permite su disponibilidad para cualquier usuario, no solo programadores. Este requiere un bajo número de datos para el entrenamiento y la forma de etiquetarlos es bastante sencilla, utilizando esqueletos predeterminados con posibilidad de adecuarlos al objeto de estudio. También, existe la posibilidad de obtener predicciones de manera rápida, dependiendo del poder de cómputo con el que se cuente, y a un bajo esfuerzo empleado por el operador en comparación con el otro enfoque. Por último, tiene la posibilidad de refinar el modelo mediante la corrección específica de las predicciones con bajos resultados (Nath et al., 2019). Sin embargo, entre sus desventajas encontramos que necesita un hardware moderno para un funcionamiento eficiente y rápido, con mínimo 32 GB de RAM y 8GB de GPU. Sobre el desempeño del modelo, tiene la incapacidad de dar seguimiento a puntos (partes del cuerpo) que se encuentren ocluidos (Nath et al., 2019), así como ha sido mayormente utilizado en experimentos de laboratorio y de comportamiento animal, contando con poca representación de casos relacionados con la presente investigación en ambientes marinos.

Para el caso del enfoque completamente programable (MSP), su principal ventaja es que el modelo puede mejorar su eficacia con el aumento de su complejidad, ya que es posible adicionar más capas y conexiones entre ellas según la tarea a realizar y el poder de cómputo que se tenga. Entre sus desventajas encontramos la alta necesidad de datos de entrenamiento, así como el gran esfuerzo (~90%) que requiere la preparación de los datos para un correcto desempeño del modelo (Purcell et al., 2022). Esto puede ser solventado mediante el incremento de los datos de entrenamiento conforme estos se van generando o recopilando en campo (Gray et al., 2018), la aplicación de técnicas de aumento de datos para resolver la escasez de datos (Mumuni & Mumuni, 2022) así como la implementación de algoritmos de detección de anomalías y herramientas de limpieza de datos para aumentar su calidad (Whang et al., 2021).

No obstante, independientemente del enfoque, ambos cuentan con una red troncal basada en una red neuronal convolucional, por lo que, existen factores que afectan al desempeño por igual sea cual sea el enfoque empleado. Uno de ellos es la alta probabilidad al sobreajuste del modelo reflejado en un

rendimiento bajo al analizar nueva información (Purcell et al., 2022). Esto puede ser mitigado brindando una mejor representación de todos los escenarios posibles a presentarse en los datos y que así permita una mejor generalización por parte del modelo para mejorar su predicción (Colomer, 2024). Para ello se recurre a aplicar técnicas de aumento de datos y la retroalimentación del conjunto de entrenamiento mediante la incorporación de nuevos datos conforme estos son recopilados (Whang et al., 2021). Otro factor es el tamaño angular de los objetos, entre más pequeños se vean en la imagen, más se complica su detección y clasificación (Purcell et al., 2022), lo cual se relaciona con la altitud de vuelo y el tamaño de los individuos. Para ello, se recomienda que los objetos sean más grandes a 80 píxeles en la totalidad de la imagen y aquellos que sean más pequeños, sean monitoreados a una altitud no mayor a 50 metros para que se logre la detección, para posteriormente se reduzca la altitud a 25 m mejorar la clasificación (Purcell et al., 2022). Estas recomendaciones de altitudes de vuelo pueden variar dependiendo de la tarea a realizar, así como del tamaño de las especies a monitorear.

4.4 Uso futuro del mejor modelo obtenido en el monitoreo de tiburón ballena

El modelo MSP-1920, basado en una red DenseNet121, muestra un alto desempeño ($f1\text{-score} = 0.91$), siendo bastante efectivo para identificar a los tiburones ballena avistados durante vuelos de dron. La identificación precisa de individuos en imágenes permite un conteo directo, etapa básica para estimar la abundancia y la densidad poblacional. Una vez que se logra realizar esta tarea de manera automatizada, permite la reducción del sesgo humano (Gray et al., 2019) y aumenta la reproducibilidad de los resultados (Whang et al., 2021). El uso de este modelo asegura que las detecciones sean consistentes, incluso cuando los individuos se encuentran en contextos visualmente complejos, como una exposición parcial de su cuerpo por encontrarse en los límites de la imagen, por la oclusión de alguna embarcación o por comportamiento de buceo. Esto mejora la precisión de los cálculos de densidad que dependen de identificar correctamente los tiburones dentro del área evaluada y puede ser complementado con técnicas de detección de objetos donde se automatiza la generación de los cuadros delimitadores (Purcell et al., 2022). Así mismo, se pueden generar modelos especializados para cada tipo de escenario que permitan mejorar la identificación y seguimiento.

A su vez, si el modelo se complementa con técnicas de análisis geométrico o fotogrametría, como la calibración de las imágenes con la altura de vuelo y la resolución de la cámara (Rex et al., 2024; Stone & Davis, 2024), es posible medir el tamaño de los tiburones ballena detectados. También se pueden recurrir a técnicas de segmentación de instancias (Gray et al., 2019; Maire et al., 2015) para la extracción del

cuerpo de los tiburones de una manera automática. Esto es fácilmente aplicable ya que las capacidades del modelo desarrollado en esta tesis disminuyen la confusión que pueda tener al identificar al tiburón.

Sin embargo, independientemente del paso siguiente a tomar, es imperativo aumentar la cantidad de videos sobre tiburones distintos mediante un protocolo de recolección de datos en campo, permitiendo así incrementar las representaciones de los diferentes escenarios en los que se puede encontrar a un tiburón ballena en la bahía. Además, se deben probar o seleccionar estrategias para resolver problemas como la afectación del reflejo solar, la turbidez y el fenómeno denominado como “floración” (Rex et al., 2024), que puede afectar tanto en su detección como en tareas más específicas como la estimación del tamaño de los individuos.

Otro tema crucial es el poder de cómputo y la inversión necesaria. Es vital considerar estos aspectos en el proyecto, ya que la principal motivación de realizar este tipo de investigaciones es generar alternativas que apoyen a la comunidad de Bahía de los Ángeles y los voluntarios del grupo comunitario para que continúen realizando los monitoreos anuales de la especie de manera más autónoma y económicamente sostenible e incluso extenderlo a otras investigaciones relacionadas con la ecología de la especie. Ante ello, uno de los principales retos es la extensa área por cubrir para el monitoreo de la especie, un polígono de 1,069 hectáreas considerando solo la SUPTB (SEMARNAT, 2014), para lo que se debe considerar el tiempo de traslado, la accesibilidad a las playas y la autonomía del dron en cuanto a la duración de la batería y el rango de conexión con el control remoto.

Capítulo 5. Conclusiones

Este trabajo demostró la viabilidad de implementar técnicas de IA para el monitoreo automatizado de tiburones ballena en Bahía de los Ángeles, utilizando imágenes capturadas por drones. Entre los modelos evaluados, el modelo MSP-1920 mostró un desempeño sobresaliente, con buenos valores en las métricas de Puntaje F1 macro (0.91) y Cobertura de Área (0.65), esta última debido al solapamiento de las imágenes donde el tiburón por lo general se encontraba entre dos parches distintos. Estas métricas destacan la capacidad del modelo para identificar correctamente a los tiburones ballena en escenarios visualmente complejos.

La eficacia del modelo MSP-1920 se mantuvo incluso bajo condiciones adversas, como alta turbidez del agua, brillo solar intenso y presencia de agentes externos (nadadores, embarcaciones y fauna). Los modelos de RL y GAM no identificaron una relación significativa entre variables ambientales y el desempeño del modelo. Sin embargo, estos modelos pueden no estar representando adecuadamente estas relaciones. A pesar de esto, se identificaron limitaciones inherentes a la autonomía de los drones y la necesidad de ampliar la base de datos de entrenamiento, lo cual es crucial para mejorar la generalización del modelo en futuras aplicaciones.

El protocolo que incluye la combinación de imágenes captadas desde un dron y su interpretación a través de IA no solo establece un estándar para el monitoreo de tiburones ballena en zonas marinas protegidas, sino que también constituye una herramienta valiosa para estudios de conservación ya que es una base para la automatización de estimaciones de abundancia y densidad poblacional con una mayor precisión al enfoque tradicional. Además, su diseño es escalable para tareas avanzadas, como el análisis morfométrico, mediante la integración de técnicas de fotogrametría o de segmentación.

Finalmente, este proyecto enfatiza la importancia de generar soluciones tecnológicas accesibles y sostenibles para apoyar a las comunidades locales y grupos de monitoreo. Así como la relevancia de la investigación multidisciplinaria para abordar la complejidad del problema estudiado, al obtener una visión más holística del fenómeno lo que enriquece el análisis. Con una inversión continua en datos y recursos computacionales, el enfoque aquí desarrollado tiene el potencial de revolucionar el monitoreo de tiburones ballena, contribuyendo en su conocimiento y conservación en Bahía de los Ángeles.

Literatura citada

- Amador-Buenrostro, A., Serrano-Guzmán, S., & Argote-Espinoza, M. (1991). Numerical model of the circulation induced by the wind at Bahía de Los Angeles, B.C., Mexico. *Ciencias Marinas*, 17(3), 39–57. <https://doi.org/10.7773/cm.v17i3.822>
- Argote, M. L., Amador, A., Lavin, M. F., & Hunter, J. R. (1995). Tidal dissipation and stratification in the Gulf of California. *Journal of Geophysical Research*, 100(C8). <https://doi.org/10.1029/95jc01500>
- Arzoumanian, Z., Holmberg, J., & Norman, B. (2005). An astronomical pattern-matching algorithm for computer-aided identification of whale sharks *Rhincodon typus*. *Journal of Applied Ecology*, 42(6), 999–1011. <https://doi.org/10.1111/j.1365-2664.2005.01117.x>
- Barbelian, M. A., Dinu, C., & Pietreanu, C. V. (2021). Deep learning approach on shark attack risk assesment using real-time autonomous surveillance systems. *U.P.B. Sci. Bull., Series D*, 83. https://www.scientificbulletin.upb.ro/rev_docs_arhiva/fulldb9_373484.pdf
- Bogucki, R., Cygan, M., Khan, C. B., Klimek, M., Milczek, J. K., & Mucha, M. (2018). Applying deep learning to right whale photo identification. *Conservation Biology*, 33(3), 676–684. <https://doi.org/10.1111/cobi.13226>
- Brack, I. V., Kindel, A., & Oliveira, L. F. B. (2018). Detection errors in wildlife abundance estimates from Unmanned Aerial Systems (UAS) surveys: Synthesis, solutions, and challenges. *Methods in Ecology and Evolution*, 9(8), 1864–1873. <https://doi.org/10.1111/2041-210X.13026>
- Brooks, K., Rowat, D., Pierce, S. J., Jouannet, D., & Vely, M. (2010). Seeing spots: Photo-identification as a regional tool for whale shark identification. *Western Indian Ocean J. Mar. Sci*, 9(2), 185–194. <https://www.ajol.info/index.php/wiojms/article/view/73980>
- Brown, A. M., Allen, S. J., Kelly, N., & Hodgson, A. J. (2023). Using unoccupied aerial vehicles to estimate availability and group size error for aerial surveys of coastal dolphins. *Remote Sensing in Ecology and Conservation*, 9(3), 340–353. <https://doi.org/10.1002/rse2.313>
- Christie, K. S., Gilbert, S. L., Brown, C. L., Hatfield, M., & Hanson, L. (2016). Unmanned aircraft systems in wildlife research: Current and future applications of a transformative technology. *Frontiers in Ecology and the Environment*, 14(5), 241–251. <https://doi.org/10.1002/fee.1281>
- Clark, E., & Nelson, D. R. (1997). Young whale sharks, *Rhincodon typus*, feeding on a copepod bloom near La Paz, Mexico. *Environmental Biology of Fishes*, 50(1), 63–73. <https://doi.org/10.1023/A:1007312310127>
- Colefax, A. P., Walsh, A. J., Purcell, C. R., & Butcher, P. (2023). Utility of spectral filtering to improve the reliability of marine fauna detections from drone-based monitoring. *Sensors*, 23(22). <https://doi.org/10.3390/s23229193>
- Colomer, D. (2024). *Deep learning for leopard individual identification: An adaptive angular margin approach*. <https://doi.org/10.48550/arXiv.2411.01962>
- Compagno, L. J. V. (2001). *Sharks of the world: an annotated and illustrated catalogue of shark species known to date* (Vol. 2). Food and Agriculture Organization of the United Nations. <https://openknowledge.fao.org/handle/20.500.14283/x9293e>

- Danemann, G. D., & Ezcurra, E. (Eds.). (2008). *Bahía de los Ángeles recursos naturales y comunidad. Línea Base 2007*. Instituto Nacional de Ecología. <https://biblioteca.semarnat.gob.mx/janium/Documentos/Ciga/Libros2013/CD002095.pdf>
- Daye, D. (2023). *Predicting habitat suitability of migratory sharks using machine learning methods* [Tesis de Maestría, University of Rhode Island]. Repositorio institucional. <https://digitalcommons.uri.edu/theses/2312>
- Digiacommo, A. E., Harrison, W. E., Johnston, D. W., & Ridge, J. T. (2020). Elasmobranch use of nearshore estuarine habitats responds to fine-scale, intra-seasonal environmental variation: Observing coastal shark density in a temperate estuary utilizing Unoccupied Aircraft Systems (UAS). *Drones*, 4(4), 1–18. <https://doi.org/10.3390/drones4040074>
- Dimauro, G., Simone, L., Carlucci, R., Fanizza, C., Lomonte, N., & Maglietta, R. (2022). Automated and non-invasive UAV-based system for the monitoring and the group size estimation of dolphins. *2022 7th International Conference on Smart and Sustainable Technologies, SpliTech 2022*. <https://doi.org/10.23919/SpliTech55088.2022.9854266>
- Ditria, E. M., Lopez-Marcano, S., Sievers, M., Jinks, E. L., Brown, C. J., & Connolly, R. M. (2020). Automating the analysis of fish abundance using object detection: Optimizing animal ecology with deep learning. *Frontiers in Marine Science*, 7. <https://doi.org/10.3389/fmars.2020.00429>
- Dujon, A. M., Ierodiaconou, D., Geeson, J. J., Arnould, J. P. Y., Allan, B. M., Katselidis, K. A., & Schofield, G. (2021). Machine learning to detect marine animals in UAV imagery: Effect of morphology, spacing, behaviour and habitat. *Remote Sensing in Ecology and Conservation*, 7(3), 341–354. <https://doi.org/10.1002/rse2.205>
- Elharrouss, O., Akbari, Y., Almadeed, N., & Al-Maadeed, S. (2024). Backbones-review: Feature extractor networks for deep learning and deep reinforcement learning approaches in computer vision. *Computer Science Review*, 53, 100645. <https://doi.org/10.1016/J.COSREV.2024.100645>
- Fu, X., Jiang, J., Wu, X., Huang, L., Han, R., Li, K., Liu, C., Roy, K., Chen, J., Mahmoud, N. T. A., & Wang, Z. (2024). Deep learning in water protection of resources, environment, and ecology: Achievement and challenges. *Environmental Science and Pollution Research*, 31(10), 14503–14536. <https://doi.org/10.1007/s11356-024-31963-5>
- Gayford, J. H., Pearse, W. D., De La Parra Venegas, R., & Whitehead, D. A. (2023). Quantifying the behavioural consequences of shark ecotourism. *Scientific Reports*, 13(1). <https://doi.org/10.1038/s41598-023-39560-1>
- Gooday, O. J., Key, N., Goldstien, S., & Zawar-Reza, P. (2018). An assessment of thermal-image acquisition with an unmanned aerial vehicle (UAV) for direct counts of coastal marine mammals ashore. *Journal of Unmanned Vehicle Systems*, 6(2), 100–108. <https://doi.org/10.1139/juvs-2016-0029>
- Goodwin, M., Halvorsen, K. T., Jiao, L., Knausgård, K. M., Martin, A. H., Moyano, M., Oomen, R. A., Rasmussen, J. H., Sørvalen, T. K., & Thorbjørnsen, S. H. (2022). Unlocking the potential of deep learning for marine ecology: overview, applications, and outlook. *ICES Journal of Marine Science*, 79(2), 319–336. <https://doi.org/10.1093/icesjms/fsab255>

- Gray, P. C., Bierlich, K. C., Mantell, S. A., Friedlaender, A. S., Goldbogen, J. A., & Johnston, D. W. (2019). Drones and convolutional neural networks facilitate automated and accurate cetacean species identification and photogrammetry. *Methods in Ecology and Evolution*, *10*(9), 1490–1500. <https://doi.org/10.1111/2041-210X.13246>
- Gray, P. C., Fleishman, A. B., Klein, D. J., McKown, M. W., Bézy, V. S., Lohmann, K. J., & Johnston, D. W. (2018). A convolutional neural network for detecting sea turtles in drone imagery. *Methods in Ecology and Evolution*, *10*(3), 345–355. <https://doi.org/10.1111/2041-210X.13132>
- Guirado, E., Tabik, S., Rivas, M. L., Alcaraz-Segura, D., & Herrera, F. (2019). Whale counting in satellite and aerial images with deep learning. *Scientific Reports*, *9*(1). <https://doi.org/10.1038/s41598-019-50795-9>
- Guo, Q., Jin, S., Li, M., Yang, Q., Xu, K., Ju, Y., Zhang, J., Xuan, J., Liu, J., Su, Y., Xu, Q., & Liu, Y. (2020). Application of deep learning in ecological resource research: Theories, methods, and challenges. *Science China Earth Sciences*, *63*(10), 1457–1474. <https://doi.org/10.1007/s11430-019-9584-9>
- Hacohen-Domené, A., Galván-Magaña, F., & Ketchum-Mejia, J. (2006). Abundance of whale shark (*Rhincodon typus*) preferred prey species in the southern Gulf of California, Mexico. *Cybium*, *30*(4), 99–102. <https://doi.org/https://doi.org/10.26028/cybium/2006-304supp-013>
- Hensel, E., Wenclawski, S., & Layman, C. A. (2018). Using a small, consumer-grade drone to identify and count marine megafauna in shallow habitats. *Latin American Journal of Aquatic Research*, *46*(5), 1025–1033. <https://doi.org/10.3856/vol46-issue5-fulltext-15>
- Hernández-Nava, M. F., & Álvarez-Borrego, S. (2013). Zooplankton en una área de alimentación del tiburón ballena (*Rhincodon typus*) en la Bahía de los Ángeles (Golfo de California). *Hidrobiológica*, *23*(2), 198–208. https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0188-88972013000200008
- Hodgson, A., Kelly, N., & Peel, D. (2013). Unmanned aerial vehicles (UAVs) for surveying marine fauna: A dugong case study. *PLoS ONE*, *8*(11). <https://doi.org/10.1371/journal.pone.0079556>
- Hodgson, A., & Marsh, H. (2007). *An evaluation of aerial surveys for monitoring whale shark abundance and distribution: Future options*. <https://doi.org/10.13140/RG.2.1.1127.2167>
- Hodgson, A., Peel, D., & Kelly, N. (2017). Unmanned aerial vehicles for surveying marine fauna: Assessing detection probability. *Ecological Applications*, *27*(4), 1253–1267. <https://doi.org/10.1002/eap.1519>
- Hsu, H. H., Joung, S. J., Liao, Y. Y., & Liu, K. M. (2007). Satellite tracking of juvenile whale sharks, *Rhincodon typus*, in the Northwestern Pacific. *Fisheries Research*, *84*(1), 25–31. <https://doi.org/10.1016/j.fishres.2006.11.030>
- Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2016). *Densely connected convolutional networks*. <https://doi.org/10.48550/arXiv.1608.06993>
- Immonen, V., Nepovinskykh, E., Eerola, T., Stewart, C. V., & Kälviäinen, H. (2023). Combining feature aggregation and geometric similarity for re-identification of patterned animals. *Camera Traps, AI and Ecology*. <https://doi.org/10.48550/arXiv.2308.06335>

- Jaramillo-Gil, S. (2020). *Estimación del tamaño de la agregación de tiburón ballena (Rhincodon typus) en Bahía de los Ángeles, Baja California, con base en foto-identificaciones* [Tesis de Maestría, Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California]. Repositorio Institucional. <http://cicese.repositorioinstitucional.mx/jspui/handle/1007/3309>
- Jaramillo-Gil, S., Pardo, M. A., Vázquez-Haikin, A., Bolaños-Jiménez, J., & Sosa-Nishizaki, O. (2023). Whale shark abundance forecast: The interannual hotspot effect. *Journal of Applied Ecology*, 60(6), 954–966. <https://doi.org/10.1111/1365-2664.14406>
- Jeffreys, G. L., Rowat, D., Marshall, H., & Brooks, K. (2013). The development of robust morphometric indices from accurate and precise measurements of free-swimming whale sharks using laser photogrammetry. *Journal of the Marine Biological Association of the United Kingdom*, 93(2), 309–320. <https://doi.org/10.1017/S0025315412001312>
- Jones, A., Bruce, E., Davies, K. P., & Cato, D. H. (2022). Enhancing UAV images to improve the observation of submerged whales using a water column correction method. *Marine Mammal Science*. <https://doi.org/10.1111/mms.12994>
- Kaur, A., Kukreja, V., Chattopadhyay, S., Verma, A., & Sharma, R. (2024). Empowering wildlife conservation with a fused CNN-SVM deep learning model for multi-classification using drone-based imagery. *2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation, IATMSI 2024*. <https://doi.org/10.1109/IATMSI60426.2024.10502645>
- Kelaher, B. P., Monteforte, K. I., Morris, S. G., Schlacher, T. A., March, D. T., Tucker, J. P., & Butcher, P. A. (2023). Drone-based assessment of marine megafauna off wave-exposed sandy beaches. *Remote Sensing*, 15(16). <https://doi.org/10.3390/rs15164018>
- Ketchum, J. T., Galván-Magaña, F., & Klimley, A. P. (2013). Segregation and foraging ecology of whale sharks, *Rhincodon typus*, in the southwestern Gulf of California. *Environmental Biology of Fishes*, 96(6), 779–795. <https://doi.org/10.1007/s10641-012-0071-9>
- Kholiavchenko, M. (2022). *Comprehensive deep learning pipeline for whale shark recognition* [Tesis de Maestría, Rensselaer Polytechnic Institute]. Repositorio institucional. <https://dspace.rpi.edu/items/972a3860-883b-4a44-95bb-aa3e632857fd>
- Kiszka, J. J., Mourier, J., Gastrich, K., & Heithaus, M. R. (2016). Using unmanned aerial vehicles (UAVs) to investigate shark and ray densities in a shallow coral lagoon. *Marine Ecology Progress Series*, 560, 237–242. <https://doi.org/10.3354/meps11945>
- Krebs, S., Duraisamy, B., & Flohr, F. (2017). A survey on leveraging deep neural networks for object tracking. *IEEE 20th International Conference on Intelligent Transportation Systems (ITSC): Workshop*, 411–418. <https://doi.org/10.1109/ITSC.2017.831790>
- Lavaniegos, B. E., Heckel, G., & Ladrón De Guevara, P. (2012). Variabilidad estacional de copépodos y cladóceros de Bahía de los Ángeles (Golfo de California) e importancia de *Acartia clausi* como alimento del tiburón ballena. *Ciencias Marinas*, 38(1A), 11–30. <https://doi.org/10.7773/cm.v38i1A.2017>
- Lenihan, P. (2023). *Aerial observations of the whale sharks of Bahía de Los Ángeles and their associations with baitfish there and throughout the world* [Tesis de Maestría, University of Miami].

Repositorio institucional: <https://scholarship.miami.edu/esploro/outputs/report/Aerial-observations-of-the-whale-sharks/991031970419802976>

- Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2022). A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE Transactions on Neural Networks and Learning Systems*, 33(12), 6999–7019. <https://doi.org/10.1109/TNNLS.2021.3084827>
- Linchant, J., Lisein, J., Semeki, J., Lejeune, P., & Vermeulen, C. (2015). Are Unmanned Aircraft Systems (UAS) the future of wildlife monitoring? A review of accomplishments and challenges. *Mammal Review*, 45(4), 239–252. <https://doi.org/10.1111/mam.12046>
- López, M., Candela, J., & Argote, M. L. (2006). Why does the Ballenas Channel have the coldest SST in the Gulf of California? *Geophysical Research Letters*, 33(11). <https://doi.org/10.1029/2006GL025908>
- Ma, G., Li, W., Bao, H., Roberts, N. J., Li, Y., Zhang, W., Yang, K., & Jiang, G. (2024). UAV equipped with infrared imaging for cervidae monitoring: Improving detection accuracy by eliminating background information interference. *Ecological Informatics*, 81. <https://doi.org/10.1016/j.ecoinf.2024.102651>
- Maglietta, R., Carlucci, R., Fanizza, C., & Dimauro, G. (2022). Machine learning and image processing methods for cetacean photo identification: A systematic review. *IEEE Access*, 10, 80195–80207. <https://doi.org/10.1109/ACCESS.2022.3195218>
- Maire, F., Mejias Alvarez, L., & Hodgson, A. (2015). *Automating marine mammal detection in aerial images captured during wildlife surveys: A deep learning approach*. <https://doi.org/10.1007/978-3-319-26350-2>
- Mitchell, J., Scott-Holland, T., Argent, J., Emmert, K., Butcher, P., Meager, J., & Mikitis, M. (2022). *Queensland SharkSmart drone trial final report*. <https://doi.org/https://doi.org/10.13140/RG.2.2.16581.12002>
- Montero-Quintana, A. N., Vázquez-Haikin, J. A., Merklng, T., Blanchard, P., & Osorio-Beristain, M. (2020). Ecotourism impacts on the behaviour of whale sharks: an experimental approach. *ORYX*, 54(2), 270–275. <https://doi.org/10.1017/S0030605318000017>
- Mumuni, A., & Mumuni, F. (2022). Data augmentation: a comprehensive survey of modern approaches. *Array*, 16. <https://doi.org/10.1016/j.array.2022.100258>
- Munappy, A. R., Bosch, J., Olsson, H. H., Arpteg, A., & Brinne, B. (2022). Data management for production quality deep learning models: challenges and solutions. *Journal of Systems and Software*, 191. <https://doi.org/10.1016/j.jss.2022.111359>
- Nates-López, A. (2015). *Caracterización individual del uso del hábitat del tiburón ballena (Rhincodon typus) en Bahía de los Ángeles mediante el uso de foto-identificación y telemetría acústica* [Tesis de Maestría, Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California]. Repositorio institucional. <http://cicese.repositorioinstitucional.mx/jspui/handle/1007/1500>

- Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M., & Mathis, M. W. (2019). Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature Protocols*, *14*(7), 2152–2176. <https://doi.org/10.1038/s41596-019-0176-0>
- Nava-Ruiz, V. M. (2018). *Estacionalidad, morfología y potencial tóxico de Pseudo-nitzschia spp. de la Zona de las Grandes Islas y Alto golfo de California* [Tesis de Maestría, Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California]. Repositorio institucional. <http://cicese.repositorioinstitucional.mx/jspui/handle/1007/2247>
- Nelson, J. D., & Eckert, S. A. (2007). Foraging ecology of whale sharks (*Rhincodon typus*) within Bahía de Los Angeles, Baja California Norte, México. *Fisheries Research*, *84*(1), 47–64. <https://doi.org/10.1016/j.fishres.2006.11.013>
- Nepovinnikh, E., Chelak, I., Eerola, T., Immonen, V., Kälviäinen, H., Kholiavchenko, M., & Stewart, C. V. (2024). Species-agnostic patterned animal re-identification by aggregating deep local features. *International Journal of Computer Vision*, *132*(9), 4003–4018. <https://doi.org/10.1007/s11263-024-02071-1>
- Nolan, G., Kurfess, F. J., Gounder, K., Tan, D., Daly, C., & Skae, C. (2021). Deep learning at a distance: Remotely working to surveil sharks. *2021 ASEE Annual Conference*. <https://doi.org/10.18260/1-2--36896>
- Paden, C. A., Winant, C. D., & Abbott, M. R. (1993). Tidal and atmospheric forcing of the upper ocean in the Gulf of California. 2. Surface heat flux. *Journal of Geophysical Research*, *98*(C11). <https://doi.org/10.1029/93jc02190>
- Patton, P. T., Cheeseman, T., Abe, K., Yamaguchi, T., Reade, W., Southerland, K., Howard, A., Oleson, E. M., Allen, J. B., Ashe, E., Athayde, A., Baird, R. W., Basran, C., Cabrera, E., Calambokidis, J., Cardoso, J., Carroll, E. L., Cesario, A., Cheney, B. J., ... Bejder, L. (2023). A deep learning approach to photo-identification demonstrates high performance on two dozen cetacean species. *Methods in Ecology and Evolution*, 1–15. <https://doi.org/10.1111/2041-210X.14167>
- Peregrín-Tovar, A. A. (2014). *Bases para la estimación de la capacidad de carga turística y propuestas para su implementación para la actividad de avistamiento de tiburón ballena (Rhincodon typus) en Bahía de los Ángeles, B.C.* [Tesis de Maestría, Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California]. Repositorio institucional. <http://cicese.repositorioinstitucional.mx/jspui/handle/1007/1446>
- Pichler, M., & Hartig, F. (2023). Machine learning and deep learning— a review for ecologists. *Methods in Ecology and Evolution*, *14*(4), 994–1016. <https://doi.org/10.1111/2041-210X.14061>
- Pierce, S. J., & Norman, B. (2016). *Rhincodon typus*. En *IUCN Red List of Threatened Species*. <https://doi.org/10.2305/IUCN.UK.2016-1.RLTS.T19488A2365291.en>
- Purcell, C. R., Walsh, A. J., Colefax, A. P., & Butcher, P. (2022). Assessing the ability of deep learning techniques to perform real-time identification of shark species in live streaming video from drones. *Frontiers in Marine Science*, *9*. <https://doi.org/10.3389/fmars.2022.981897>
- Rainio, O., Teuho, J., & Klén, R. (2024). Evaluation metrics and statistical tests for machine learning. *Scientific Reports*, *14*(1). <https://doi.org/10.1038/s41598-024-56706-x>

- Ramírez-Macías, D., Vázquez-Haikin, A., & Vázquez-Juárez, R. (2012). Whale shark *Rhincodon typus* populations along the west coast of the Gulf of California and implications for management. *Endangered Species Research*, 18(2), 115–128. <https://doi.org/10.3354/esr00437>
- Raudino, H., Rob, D., Barnes, P., Mau, R., Wilson, E., Gardner, S., & Waples, K. (2016). Whale shark behavioural responses to tourism interactions in Ningaloo Marine Park and implications for future management. *Conservation Science W. Aust*, 10(2). www.dpaw.wa.gov.au/CSWAjournal
- Rex, P. T., Abbott, K. J., Prezgay, R. E., & Lowe, C. G. (2024). The effects of depth and altitude on image-based shark size measurements using UAV surveillance. *Drones*, 8(10), 547. <https://doi.org/10.3390/drones8100547>
- Rex, P. T., May, J. H., Pierce, E. K., & Lowe, C. G. (2023). Patterns of overlapping habitat use of juvenile white shark and human recreational water users along southern California beaches. *PLoS one*, 18(6), e0286575. <https://doi.org/10.1371/journal.pone.0286575>
- Rowat, D., Brooks, K., March, A., McCarten, C., Jouannet, D., Riley, L., Jeffreys, G., Perri, M., Vely, M., & Pardigon, B. (2011). Long-term membership of whale sharks (*Rhincodon typus*) in coastal aggregations in Seychelles and Djibouti. *Marine and Freshwater Research*, 62(6), 621–627. <https://doi.org/10.1071/MF10135>
- Rowat, D., & Brooks, K. S. (2012). A review of the biology, fisheries and conservation of the whale shark *Rhincodon typus*. *Journal of Fish Biology*, 80(5), 1019–1056. <https://doi.org/10.1111/j.1095-8649.2012.03252.x>
- Rowat, D., Gore, M., Meekan, M. G., Lawler, I. R., & Bradshaw, C. J. A. (2009). Aerial survey as a tool to estimate whale shark abundance trends. *Journal of Experimental Marine Biology and Ecology*, 368(1), 1–8. <https://doi.org/10.1016/j.jembe.2008.09.001>
- Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT). (2014). *Programa de manejo reserva de la biosfera zona marina Bahía de los Ángeles, Canales de Ballenas y de Salsipuedes* (SEMARNAT & CONANP, Eds.). https://simec.conanp.gob.mx/pdf_libro_pm/1_libro_pm.pdf
- Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT). (2018). *Programa de Acción para la Conservación de la Especie Tiburón Ballena (*Rhincodon typus*)* (SEMARNAT & CONANP, Eds.). <https://www.gob.mx/conanp/documentos/programa-de-accion-para-la-conservacion-de-especies-pace-tiburon-ballena-rhincodon-typus?state=published>
- Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT). (2024). *Plan de manejo tipo para la conservación y aprovechamiento no extractivo de *Rhincodon typus* (tiburón ballena) a través de la observación y nado en Bahía de los Ángeles, Baja California. Temporada 2024*. https://www.gob.mx/cms/uploads/attachment/file/919159/PMT_Tiburon_Ballena_Bahia_de_Los_Angeles_2024.pdf
- Setyawan, E., Stevenson, B. C., Izuan, M., Constantine, R., & Erdmann, M. V. (2022). How big is that manta ray? A novel and non-invasive method for measuring reef manta rays using small drones. *Drones*, 6(3). <https://doi.org/10.3390/drones6030063>
- Soto-Mardones, L., Marioni, S., & Pares-Sierra, A. (1999). Variabilidad espaciotemporal de la temperatura superficial del mar en el Golfo de California. *Ciencias Marinas*, 25(1), 1–30. <http://www.redalyc.org/articulo.oa?id=48025101>

- Speed, C. W., Meekan, M. G., & Bradshaw, C. J. A. (2007). Spot the match - wildlife photo-identification using information theory. *Frontiers in Zoology*, 4. <https://doi.org/10.1186/1742-9994-4-2>
- Stone, T. C., & Davis, K. J. (2024). Using unmanned aerial vehicles to estimate body volume at scale for ecological monitoring. *Methods in Ecology and Evolution*. <https://doi.org/10.1111/2041-210X.14457>
- Taylor, J. G. (2007). Ram filter-feeding and nocturnal feeding of whale sharks (*Rhincodon typus*) at Ningaloo Reef, Western Australia. *Fisheries Research*, 84(1), 65–70. <https://doi.org/10.1016/j.fishres.2006.11.014>
- Tershy, B. R., Breese, D., & Álvarez-Borrego, S. (1991). Increase in cetacean and seabird numbers in the Canal de Ballenas during an El Niño-Southern Oscillation event. *Marine Ecology Progress Series*, 69, 299–302. <https://www.int-res.com/articles/meps/69/m069p299.pdf>
- Torres-Orozco, E. (1993). *Análisis volumétrico de las masas de agua del Golfo de California* [Tesis de Maestría, Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California]. Repositorio institucional. <http://cicese.repositorioinstitucional.mx/jspui/handle/1007/2881>
- Whang, S. E., Roh, Y., Song, H., & Lee, J.-G. (2021). *Data collection and quality challenges in deep learning: a data-centric AI perspective*. <http://arxiv.org/abs/2112.06409>
- Whitehead, D. A., Ayres, K. A., Gayford, J. H., Ketchum, J. T., Galván-Magana, F., & Christiansen, F. (2022). Aerial photogrammetry of whale sharks (*Rhincodon typus*) in the Bay of La Paz, using an unoccupied aerial vehicle. *Marine Biology*, 169(7). <https://doi.org/10.1007/s00227-022-04085-0>
- Whitehead, D. A., Magaña, F. G., Ketchum, J. T., Hoyos, E. M., Armas, R. G., Pancaldi, F., & Olivier, D. (2021). The use of machine learning to detect foraging behaviour in whale sharks: a new tool in conservation. *Journal of Fish Biology*, 98(3), 865–869. <https://doi.org/10.1111/jfb.14589>
- Wu, Y. chen, & Feng, J. wen. (2018). Development and Application of Artificial Neural Network. *Wireless Personal Communications*, 102(2), 1645–1656. <https://doi.org/10.1007/s11277-017-5224-x>

Anexo A

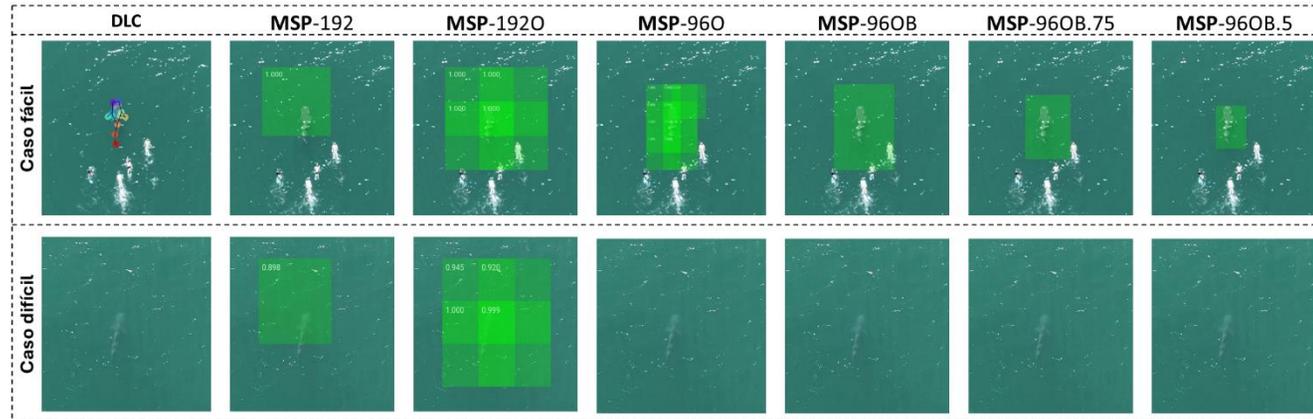


Figura 23. Comparación de la predicción de los modelos del video 22.2_PlayaBlanca-PM-1 ante un caso fácil y uno difícil.

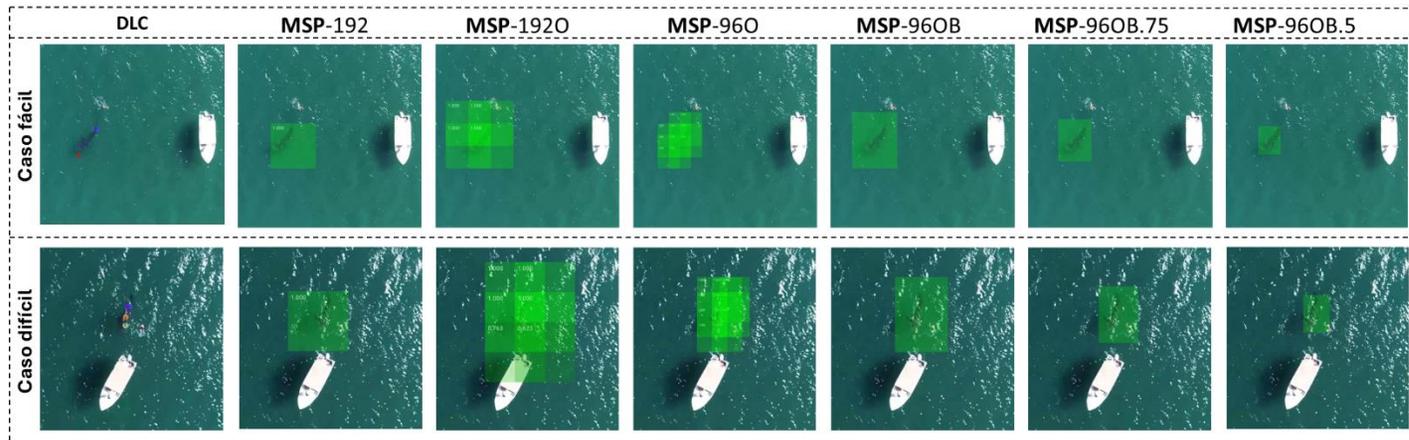


Figura 24. Comparación de la predicción de los modelos del video 25.1_Tb2 ante un caso fácil y uno difícil.

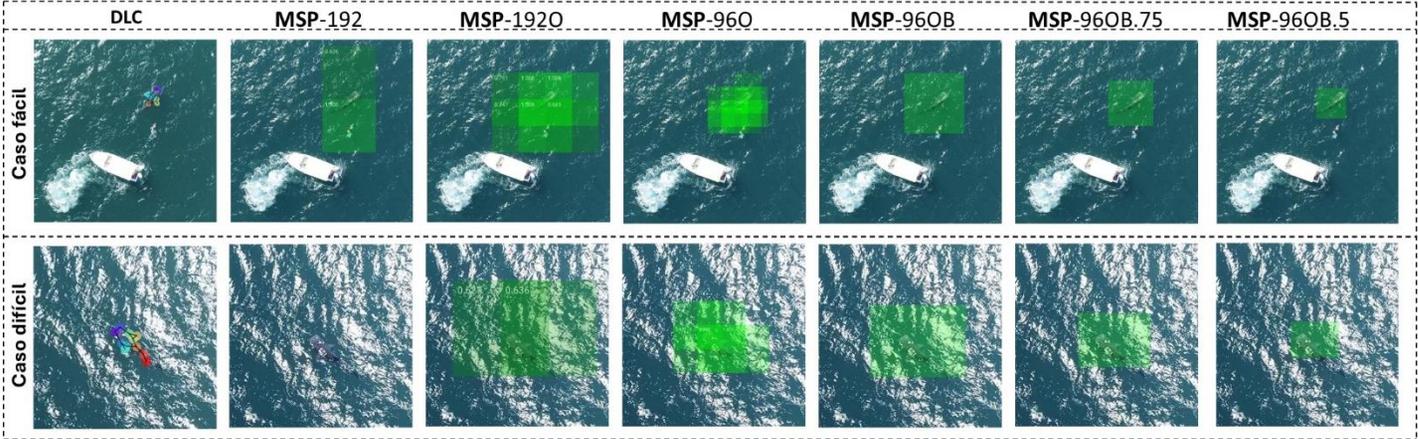


Figura 25. Comparación de la predicción de los modelos del video 25.1_Tb7 ante un caso fácil y uno difícil.

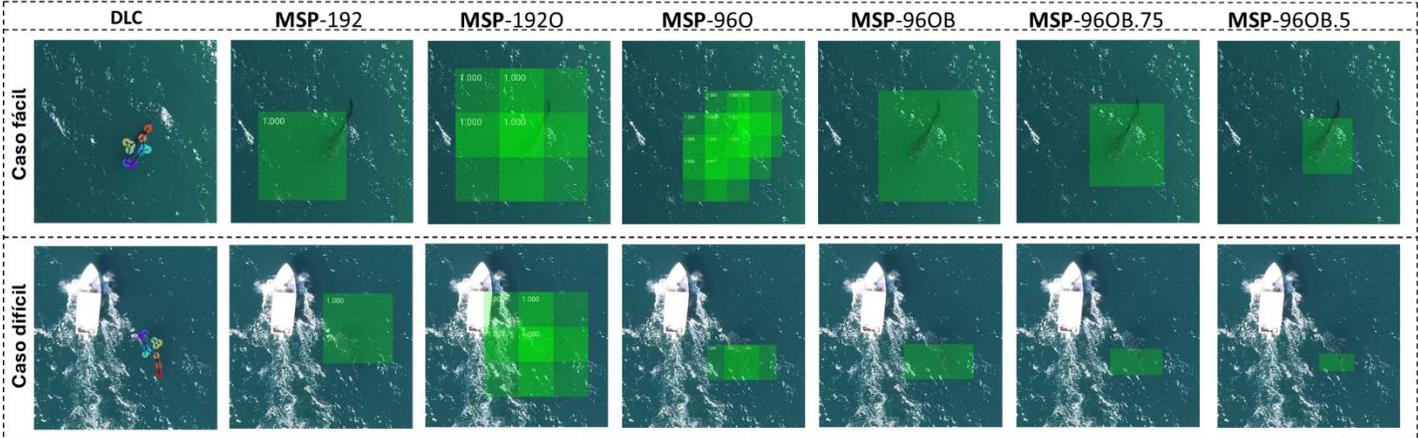


Figura 26. Comparación de la predicción de los modelos del video 25.1_Tb9 ante un caso fácil y uno difícil.

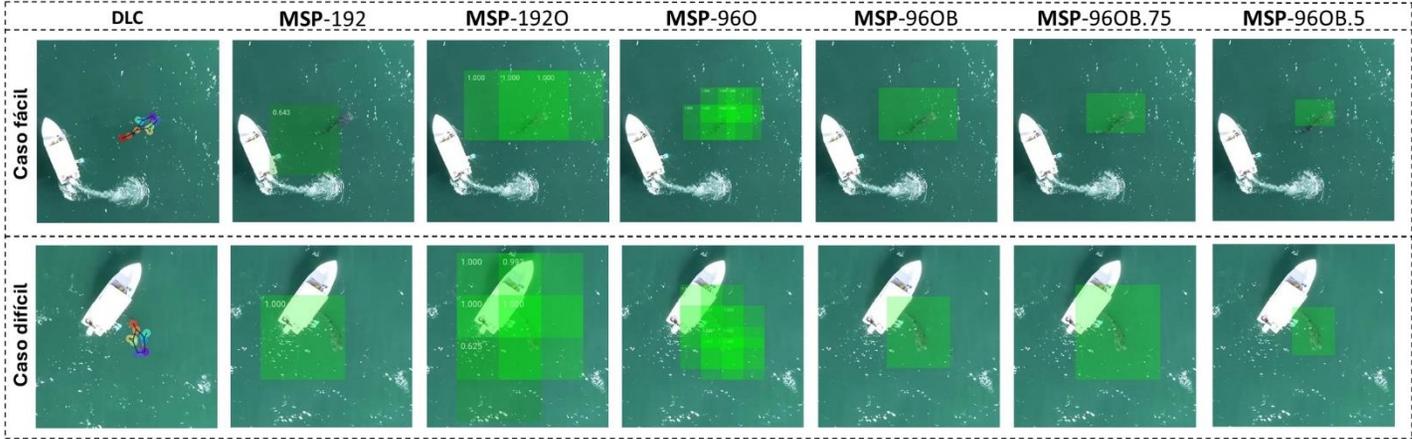


Figura 27. Comparación de la predicción de los modelos del video 25.1_Tb12 ante un caso fácil y uno difícil.

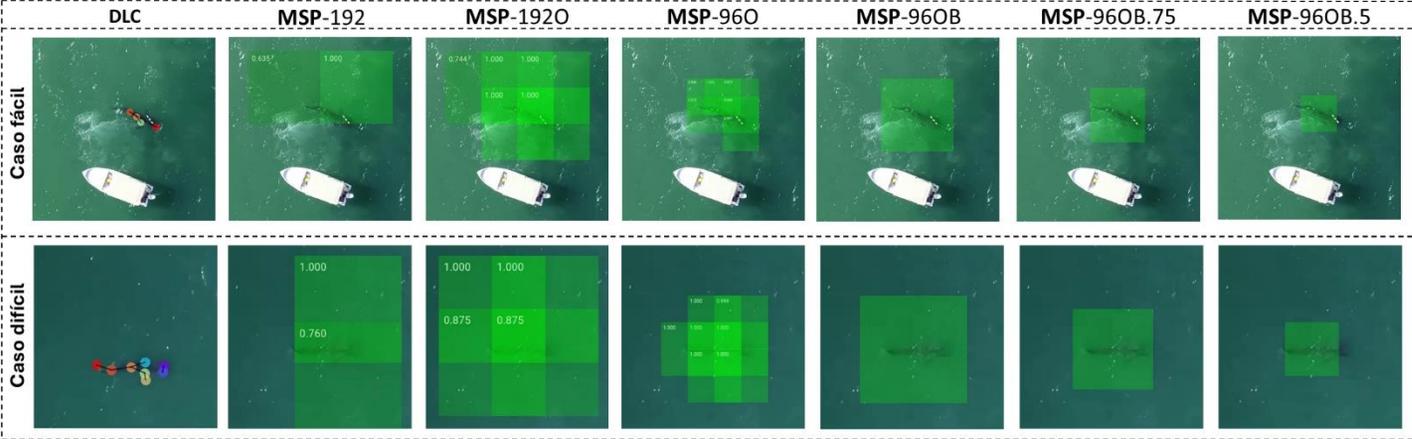


Figura 28. Comparación de la predicción de los modelos del video 25.1_Tb13 ante un caso fácil y uno difícil.

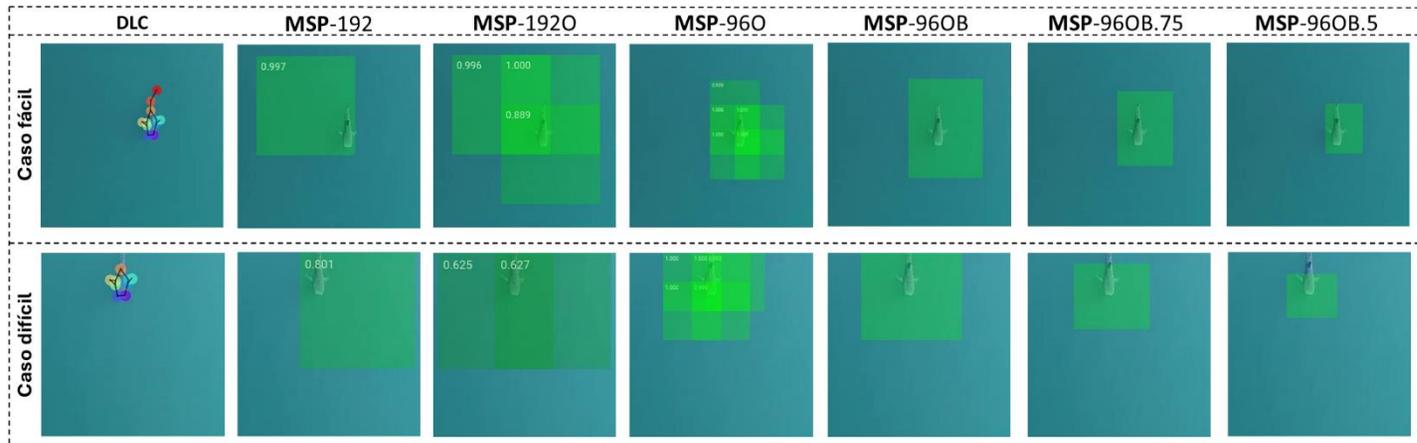


Figura 29. Comparación de la predicción de los modelos del video 27.1_Angelitos-AM-2 ante un caso fácil y uno difícil.

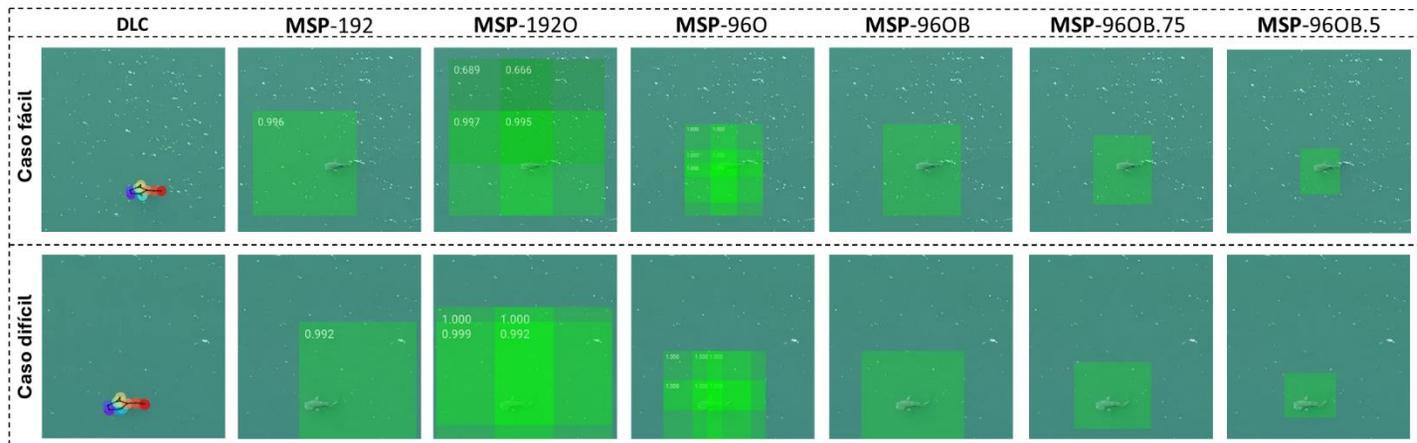


Figura 30. Comparación de la predicción de los modelos del video 27.3_Estero-AM-1 ante un caso fácil y uno difícil.

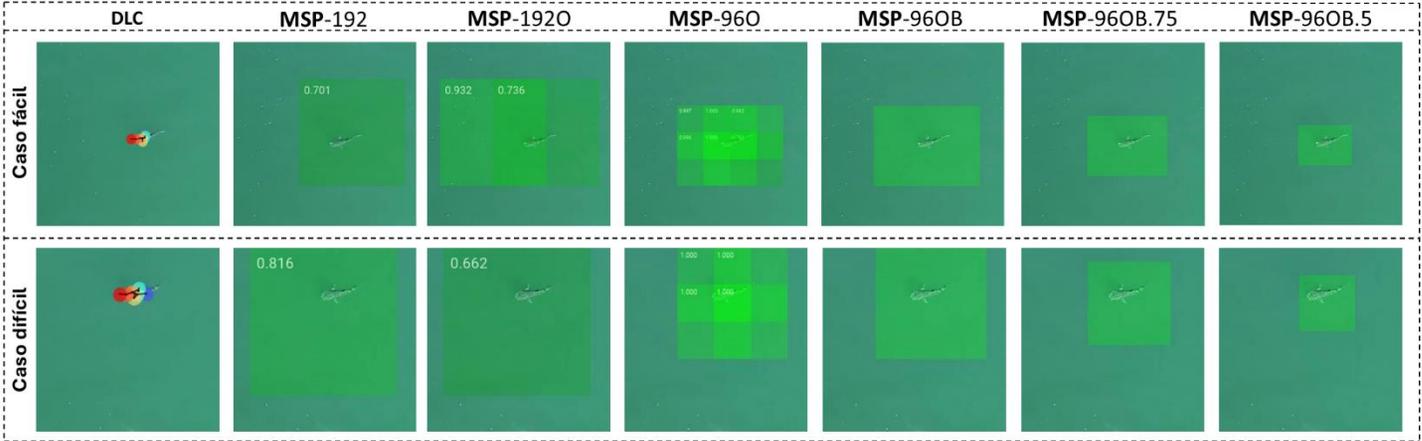


Figura 31. Comparación de la predicción de los modelos del video 28.4_Rampa-PM-1 ante un caso fácil y uno difícil.