

# **Detección y clasificación automática de embarcaciones por parámetros acústicos.**

*Tesis final presentada para obtener el título de  
Ingeniero de Sonido de la Universidad Nacional de Tres  
de Febrero (UNTREF)*

**TESISTA: Tomás Martín (38.416.291)**

**TUTOR/A: Ing. Alexander Marino**

**COTUTOR/A: Dra. Vanesa Reyes**



## **AGRADECIMIENTOS**

En primer lugar, debo agradecer a la Universidad Nacional de Tres de Febrero y en particular a los coordinadores de la carrera de Ingeniería de Sonido por generar un espacio donde adquirir conocimientos variados y profundos del fenómeno sonoro. Gracias a todas las personas que fueron partícipes de mi formación académica directa o indirectamente.

Un sentido agradecimiento al Grupo de Investigación de Acústica Subacuática (GIAS) por introducir y transmitir su vocación por esta ciencia. Al personal de la Fundación Cethus por acercar las inquietudes y necesidades que dieron lugar a la presente investigación. A mi tutor, Ing. Alexander Marino y cotutora, Dra. Vanesa Reyes por guiarme durante todo este proceso.

Gracias a mis amigos y compañeros por los buenos momentos compartidos. Ustedes hicieron posible el hecho de poder recorrer toda una carrera universitaria estando lejos de familiares y lugar de origen. En especial agradecer a Ciro, con quien pude compartir todos estos años de estudio.

Por último, pero no menos importante, quiero agradecer a mi familia que, a pesar de la distancia, me brindó su incondicional apoyo durante todo este trayecto. En particular a mi madre Rosana por haberme dado el cariño y motivación para alcanzar mis metas.

## **DEDICATORIA**

*A mi vieja, por su apoyo incondicional.*

*Este logro también es suyo.*

# ÍNDICE DE CONTENIDOS

RESUMEN .....	VI
ABSTRACT .....	VII
1. ESTRUCTURA DE LA INVESTIGACIÓN .....	1
2. INTRODUCCIÓN .....	2
2.1. FUNDAMENTACIÓN .....	2
2.2. OBJETIVOS .....	3
2.2.1. OBJETIVO GENERAL .....	3
2.2.2. OBJETIVOS ESPECÍFICOS .....	3
3. ESTADO DEL ARTE Y MARCO TEÓRICO .....	5
3.1. ACÚSTICA SUBACUÁTICA .....	5
3.2. IMPACTO ANTROPOGÉNICO EN FAUNA MARINA .....	6
3.3. RUIDO MARINO .....	7
3.4. RUIDO DE EMBARCACIONES .....	7
3.4.1. RUIDO DE HÉLICE .....	8
3.4.2. RUIDO DE MAQUINARIA .....	10
3.4.3. RUIDO POR FLUJO DE AGUA .....	10
3.4.4. COMBINACIÓN DE FUENTES .....	11
3.5. TRANSPORTE MARÍTIMO .....	12
3.6. DESCRIPTORES ACÚSTICOS .....	14
3.6.1. ENERGÍA POR TERCIOS DE OCTAVA .....	15
3.6.2. FIRMA ACÚSTICA .....	15
3.7. REPRESENTACIÓN DIGITAL DE SEÑALES. ....	16
3.7.1. FRECUENCIA DE MUESTREO .....	16
3.7.2. DECIMACIÓN .....	16
3.7.3. NORMALIZACIÓN .....	17
3.8. ALGORITMOS DE DETECCIÓN Y CLASIFICACIÓN EN GRANDES BASES DE DATOS .....	17
3.8.1. K-NEAREST NEIGHBORS (KNN) .....	19
3.8.2. SUPPORT VECTOR MACHINE (SVM) .....	19
3.8.3. ÁRBOL DE DECISIÓN (DT) .....	20
3.8.4. RANDOM FOREST .....	21
3.9. MÉTRICAS DE DESEMPEÑO .....	22
4. DISEÑO DE LA INVESTIGACIÓN, VARIABLE Y MUESTRA .....	24

4.1. MANEJO DE DATOS .....	25
4.2. COMPOSICIÓN DE LA BASE DE DATOS.....	25
4.3. DESCRIPTORES ACÚSTICOS .....	29
4.3.1. ALGORITMO DE CÁLCULO DE DESCRIPTORES. ....	31
4.4. PRE-PROCESAMIENTO DEL DATASET.....	36
4.5. DETECTOR AUTOMÁTICO DE EMBARCACIONES.....	38
4.5.1. EXPERIMENTACIÓN CON ALGORITMOS PARA DETECTORES.....	39
4.5.2. EVALUACIÓN DE CAPACIDAD DE GENERALIZACIÓN Y SESGO. ....	40
4.6. CLASIFICADOR AUTOMÁTICO DE EMBARCACIONES.....	41
5. RESULTADOS.....	42
5.1. PARÁMETROS ACÚSTICOS Y VARIABLES DE ENTRADA. ....	42
5.1.1. MEL-FREQUENCY CEPSTRAL COEFFICIENT (MFCC) .....	44
5.1.2. MEL SPECTROGRAM .....	45
5.1.3. SPECTRAL CONTRAST .....	46
5.1.4. FIRMA ACÚSTICA .....	46
5.1.5. VECTOR DE ÍNDICES DE TONALIDAD .....	47
5.2. DETECTOR AUTOMÁTICO DE EMBARCACIONES.....	49
5.2.1. EVALUACIÓN DE LOS MODELOS.....	51
5.3. CLASIFICADOR AUTOMÁTICO DE EMBARCACIONES.....	53
6. CONCLUSIONES.....	59
7. TRABAJOS A FUTURO.....	61
BIBLIOGRAFÍA .....	62
ANEXO I.    Evaluación de capacidad de generalización de detectores.....	66
ANEXO II.    CÓDIGOS DESARROLLADOS .....	72

## ÍNDICE DE FIGURAS

Figura 1.	Representación gráfica de las fuentes de ruido en un barco.....	8
Figura 2.	Representación gráfica de los tipos de cavitación en la hélice.....	9
Figura 3.	Cavitación de vórtice en la punta.....	10
Figura 4.	Espectrogramas de distintos tipos de barcos.....	13
Figura 5.	Representación gráfica del proceso de decimación.....	16
Figura 6.	Representación gráfica del funcionamiento de kNN. ....	19
Figura 7.	Representación gráfica del funcionamiento de SVM.....	20
Figura 8.	Representación gráfica del funcionamiento de un árbol de decisión.....	21
Figura 9.	Representación gráfica del funcionamiento de Random Forest.....	22
Figura 10.	Diagrama en bloques del diseño de la investigación. ....	24
Figura 11.	Ejemplo de detección visual de embarcaciones en Raven Pro 1.6. ....	30
Figura 12.	Representación gráfica de firma acústica .....	32
Figura 13.	Representación de la detección de picos.....	33
Figura 14.	Representación de índices de tonalidad en espectrograma.....	35
Figura 15.	Diagrama en bloques de la etapa de pre-procesamiento.....	37
Figura 16.	Diagrama en bloques de la etapa de detección.....	39
Figura 17.	Representación de resultados del índice de tonalidad.....	43
Figura 18.	Representación de los MFCC.....	44
Figura 19.	Representación de los espectrogramas de Mel.....	45
Figura 20.	Representación del Contraste Espectral.....	46
Figura 21.	Representación de las firmas acústicas.....	47
Figura 22.	Representación los índices de tonalidad.....	48
Figura 23.	Detecciones sobre espectrograma de larga duración obtenidas con DT.....	52

## ÍNDICE DE TABLAS

Tabla 1.	Descriptores acústicos subacuáticos .....	14
Tabla 2.	Distribución de muestras de distintos grupos. ....	28
Tabla 3.	Distribución frecuencial de los subrangos.....	34
Tabla 4.	Mediana (Q1-Q3) de lo descriptores únicos.....	42
Tabla 5.	Hiperparámetros que maximizan la performance para cada modelo .....	50
Tabla 6.	Resultados de las métricas de desempeño de los distintos modelos. ....	50
Tabla 7.	Hiperparámetros que maximizan la performance para cada modelo .....	54
Tabla 8.	Métricas de desempeño por categoría - Modelo KNN.....	55
Tabla 9.	Métricas de desempeño por categoría - Modelo SVM.....	56
Tabla 10.	Métricas de desempeño por categoría - Modelo RF.....	56
Tabla 11.	Métricas de desempeño por categoría - Modelo DT. ....	57



## ÍNDICE DE ANEXO

ANEXO I.	EVALUACIÓN DE CAPACIDAD DE GENERALIZACIÓN DE DETECTORES. ....	66
ANEXO II.	CÓDIGOS DESARROLLADOS .....	72

## ÍNDICE DE FIGURAS ANEXO

FIGURA A1.	DETECCIONES SOBRE ESPECTROGRAMA DE LARGA DURACIÓN OBTENIDAS CON KNN. ....	66
FIGURA A2.	DETECCIONES SOBRE ESPECTROGRAMA DE LARGA DURACIÓN OBTENIDAS CON SVM. ....	66
FIGURA A3.	DETECCIONES SOBRE ESPECTROGRAMA DE LARGA DURACIÓN OBTENIDAS CON DT. ....	67
FIGURA A4.	DETECCIONES SOBRE ESPECTROGRAMA DE LARGA DURACIÓN OBTENIDAS CON RF. ....	67
FIGURA A5.	DETECCIONES SOBRE ESPECTROGRAMA DE LARGA DURACIÓN OBTENIDAS CON SVM. ....	68
FIGURA A6.	DETECCIONES SOBRE ESPECTROGRAMA DE LARGA DURACIÓN OBTENIDAS CON KNN. ....	68
FIGURA A7.	DETECCIONES SOBRE ESPECTROGRAMA DE LARGA DURACIÓN OBTENIDAS CON RF. ....	69
FIGURA A8.	DETECCIONES SOBRE ESPECTROGRAMA DE LARGA DURACIÓN OBTENIDAS CON DT. ....	69
FIGURA A9.	DETECCIONES SOBRE ESPECTROGRAMA DE LARGA DURACIÓN OBTENIDAS CON SVM. ....	70
FIGURA A10.	DETECCIONES SOBRE ESPECTROGRAMA DE LARGA DURACIÓN OBTENIDAS CON KNN. ....	70
FIGURA A11.	DETECCIONES SOBRE ESPECTROGRAMA DE LARGA DURACIÓN OBTENIDAS CON RF. ....	71
FIGURA A12.	DETECCIONES SOBRE ESPECTROGRAMA DE LARGA DURACIÓN OBTENIDAS CON DT. ....	71

## RESUMEN

En el presente trabajo de tesis se desarrollan algoritmos de detección y clasificación de embarcaciones basados en características acústicas. Se busca proporcionar herramientas para automatizar el análisis en el monitoreo acústico pasivo (PAM) y sentar las bases para futuras mejoras en los sistemas propuestos.

La metodología empleada consta de tres etapas principales. En primer lugar, se realiza un estudio exhaustivo del comportamiento acústico de las embarcaciones y se crea una base de datos con diversas muestras. En segundo lugar, se obtienen y evalúan diferentes descriptores acústicos, tanto estandarizados como propuestos. Se destaca el parámetro "índice de tonalidad" por su capacidad para distinguir entre embarcaciones y no embarcaciones.

En la tercera etapa, se desarrollan y evalúan cuatro modelos de aprendizaje automático para la detección y clasificación de embarcaciones: Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest (RF) y Decision Tree (DT). Se calculan métricas de desempeño para los algoritmos de detección, obteniendo resultados satisfactorios en todos los modelos, con valores de F1 superiores a 0,95. Además, se evalúa la capacidad de generalización de los modelos en archivos de larga duración independientes de los conjuntos de datos utilizados para el entrenamiento, demostrando una buena capacidad para detectar embarcaciones en condiciones realistas.

En cuanto a la clasificación, se encontró que el modelo SVM obtiene los mejores resultados en todas las categorías, con valores de F1 superiores a 0,8, seguido por el modelo KNN con un rendimiento aceptable, valores de F1 superiores a 0,75.

**Palabras clave:** detección de embarcaciones, clasificación acústica, características acústicas, aprendizaje automático, monitoreo acústico pasivo.

## **ABSTRACT**

In this thesis, algorithms for vessel detection and classification based on acoustic characteristics are developed. The aim is to provide tools for automating analysis in passive acoustic monitoring (PAM) and lay the foundation for future improvements in the proposed systems.

The methodology employed consists of three main stages. Firstly, an exhaustive study of vessel acoustic behavior is conducted, and a diverse sample database is created. Secondly, different acoustic descriptors, both standardized and proposed, are obtained and evaluated. The 'tonality index' parameter is emphasized for its ability to distinguish between vessels and non-vessels.

In the third stage, four machine learning models, namely Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest (RF), and Decision Tree (DT), are developed and evaluated for vessel detection and classification. Performance metrics for the detection algorithms are calculated, yielding satisfactory results in all models, with F1 scores above 0.95. Additionally, the model's generalization ability is evaluated on independent long-duration files, separate from the training datasets, demonstrating good capacity for detecting vessels under realistic conditions.

Regarding classification, it was found that the SVM model achieves the best results in all categories, with F1 scores above 0.8, followed by the KNN model with acceptable performance, with F1 scores above 0.75.

**Keywords:** vessel detection, acoustic classification, acoustic characteristics, machine learning, passive acoustic monitoring

# 1. ESTRUCTURA DE LA INVESTIGACIÓN

En el primer capítulo se exponen los fundamentos principales que movilizan esta tesis de grado. Se introduce a la problemática, se explican las razones por las cuales resulta importante el abordaje del tema, se especifica el objetivo general y sus correspondientes objetivos específicos.

En el segundo capítulo se expone tanto una revisión del estado del arte de las principales investigaciones llevadas a cabo en el estudio del ruido de embarcaciones como el marco teórico necesario para la elaboración y correcta comprensión de la presente investigación. El marco teórico está dividido en tres secciones. En primer lugar, se describen los distintos eventos acústicos generados por las embarcaciones. Luego se presentan descriptores acústicos. Por último, se muestran principios y consideraciones a tener en cuenta en algunos modelos de aprendizaje automático.

En el tercer capítulo se desarrolla el procedimiento llevado a cabo en la investigación. En primer lugar, se realiza un estudio exhaustivo de los descriptores acústicos de las señales de interés. En segundo lugar, se desarrollan y evalúan algoritmos para detectar y clasificar los registros en cuestión.

En el cuarto capítulo se presentan los resultados obtenidos. Se examinan los resultados de los descriptores acústicos utilizados. Se analizan y discuten los resultados de desempeño de los modelos de aprendizaje automático, así como su capacidad de generalización en diferentes escenarios.

En el quinto capítulo se exponen las conclusiones derivadas de la presente investigación. Se resumen los hallazgos más relevantes y se discuten las contribuciones y limitaciones de la tesis.

En el sexto capítulo se presentan las recomendaciones para trabajos futuros basadas en los resultados y conclusiones obtenidas. Se identifican áreas de mejora y posibles líneas de investigación que podrían ampliar y mejorar el desempeño de los algoritmos de detección y clasificación propuestos.

## 2. INTRODUCCIÓN

### 2.1. FUNDAMENTACIÓN

El ruido producto de la actividad humana, conocido como antropogénico, es cada vez mayor en los mares y océanos del mundo. Los eventos sonoros de cortos períodos de tiempo y alta intensidad han sido objeto de estudio debido a su potencial para causar lesiones físicas en mamíferos marinos [1-3]. Sin embargo, el incremento del transporte marítimo supone una exposición a ruidos de baja intensidad en largos períodos de tiempo lo que resulta en consecuencias potencialmente graves en una variedad de organismos marinos, incluidos mamíferos, peces y cefalópodos [4-6]. Algunos de los efectos que producen son: enmascaramiento de sonidos biológicamente significativos, estrés crónico, modificación en la conducta y cambios en la atención [7-10]. Esta es la razón por la cual se debe prestar especial atención en la actividad antropogénica en el mar.

Dentro del territorio argentino, la Fundación Cethus<sup>1</sup> es una de las principales organizaciones sin fines de lucro encargadas de investigar, monitorear, estudiar y divulgar conocimiento sobre los distintos tipos de cetáceos del Mar Argentino para su conservación. Desde el año 2010, la fundación implementa métodos de monitoreo que buscan caracterizar señales acústicas subacuáticas [11]. En la actualidad, en conjunto con el Grupo de Investigación de Acústica Submarina (GIAS) de la Universidad Nacional de Tres de Febrero (UNTREF), se encuentra en etapa de desarrollo de una sonoboya autónoma de Monitoreo Acústico Pasivo (PAM) que registre eventos sonoros (tanto de cetáceos como antropogénicos) para enviarlos automáticamente a un servidor en la nube con el cual se permite el procesamiento remoto de la información.

En la presente investigación se analizan las características acústicas de distintos sonidos de embarcaciones. Dichos eventos se obtienen a partir de una inspección visual de espectrogramas de registros sonoros provenientes de diversas fuentes. En primer lugar, se analizan las grabaciones recolectadas por Fundación Cethus durante la campaña antártica (2014-2020).

---

<sup>1</sup> La Fundación Cethus es una organización no gubernamental sin fines de lucro fundada en Buenos Aires en 1992 que realiza trabajos de investigación y conservación sobre cetáceos. Sitio web: [www.cethus.org](http://www.cethus.org).

En segundo lugar, a partir de una intensiva búsqueda, se intenta encontrar posibles bases de datos de libre acceso en internet. Además, se establece comunicación con los autores de trabajos similares solicitando el acceso a sus bases de datos de este tipo de registros de forma colaborativa.

Se pretende encontrar particularidades que permitan detectar y clasificar entre distintos tipos de embarcaciones en base a descriptores acústicos estandarizados y personalizados. El lenguaje utilizado para realizar los algoritmos es Python basándose en diversas herramientas de Inteligencia Artificial. A partir de este desarrollo, se buscan abordar posibles soluciones computacionales que permitan discriminar los mencionados eventos, lo que permitirá reducir significativamente tanto la memoria de registro de datos como la energía necesaria para el envío de la información al servidor en el marco del proyecto que está desarrollando la fundación.

En una primera instancia se genera una base de datos (BD) conformada tanto por señales de audio registradas en la campaña antes mencionada como por grabaciones provenientes de otras fuentes. Luego, se lleva a cabo una investigación exhaustiva para determinar los descriptores más efectivos en la diferenciación de los registros acústicos de embarcaciones y por último se desarrollan y evalúan algoritmos que permitan detectar y clasificar automáticamente las mencionadas señales.

## **2.2. OBJETIVOS**

### **2.2.1. OBJETIVO GENERAL**

El objetivo principal de la presente investigación es el estudio de señales acústicas producidas por embarcaciones para el desarrollo de sistemas de detección y clasificación basados en descriptores acústicos y aprendizaje automático.

### **2.2.2. OBJETIVOS ESPECÍFICOS**

- Crear una amplia base de datos con registros sonoros diversos clasificados en distintas categorías.
- Analizar características acústicas del ruido antropogénico submarino.
- Establecer los descriptores que caractericen a los eventos sonoros.

- Desarrollar y evaluar algoritmos en Python capaces de detectar y clasificar el ruido inducido por el hombre mediante inteligencia artificial.
- Realizar propuestas de mejoras de desarrollo e investigación.

### 3. ESTADO DEL ARTE Y MARCO TEÓRICO

#### 3.1. ACÚSTICA SUBACUÁTICA

La Acústica Subacuática desempeña un papel fundamental en el estudio y comprensión del sonido en el medio acuático. Esta disciplina se centra en el análisis de la propagación, interacción y detección de ondas sonoras en el agua, abarcando diversas áreas que van desde la monitorización de la vida marina hasta el control de impacto ambiental de las actividades antropogénicas. En esta sección, se abordan algunos aspectos clave de la acústica subacuática que son relevantes para la presente investigación.

A diferencia de la propagación en el aire, el sonido en el agua se propaga a velocidades superiores y presenta una menor atenuación. Estas características permiten que las ondas sonoras viajen distancias más largas con una pérdida de energía reducida. Por otro lado, la propagación del sonido en el agua se ve influenciada por diversos factores ambientales. La profundidad del agua, la temperatura, la salinidad y la presencia de corrientes oceánicas son los principales elementos que afectan la velocidad de propagación y la atenuación del sonido subacuático. Aunque es deseable contar con instrumentación precisa para obtener información específica sobre estos parámetros, en ausencia de los mismos, se suelen utilizar valores aproximados. Por ejemplo, para agua salada, se considera una velocidad de propagación de alrededor de 1500 m/s, mientras que en agua dulce esta velocidad se estima aproximadamente en 1435 m/s.

Es importante tener en cuenta que el nivel de presión sonora en el aire se calcula utilizando un valor de referencia asociado al umbral de audición humana a 1 kHz (20  $\mu$ Pa), mientras que en el agua se utiliza un valor de referencia de 1  $\mu$ Pa. Esta diferencia en los niveles de referencia debe ser considerada al realizar mediciones y análisis acústicos en el entorno subacuático.

Asimismo, la ubicación de una fuente de sonido desempeña un papel relevante en la propagación del sonido en el agua. Una fuente se considera puntual cuando la energía decae 6 dB por cada duplicación de la distancia. En general, debido a las reflexiones causadas por los cambios de medios de propagación (superficie del agua y fondo marino), el océano se comporta como una guía de onda acústica, con una atenuación de 3 dB por



cada duplicación de la distancia. Este fenómeno es especialmente notable para bajas frecuencias, como las generadas por los motores de las embarcaciones.

En el marco de la Acústica Subacuática, se han desarrollado una amplia gama de técnicas y herramientas para la detección, clasificación y localización de fuentes sonoras en el medio acuático. Estas técnicas incluyen el monitoreo acústico pasivo también conocido como PAM (por sus siglas en inglés, Passive Acoustic Monitoring), que se basa en el uso de dispositivos para registrar y analizar de manera continua los sonidos presentes sin necesidad de una intervención activa en el ambiente marino.

### **3.2. IMPACTO ANTROPOGÉNICO EN FAUNA MARINA**

Como se informa en la introducción de la presente investigación, el incremento de la actividad humana en los océanos supone un impacto negativo sobre la fauna marina. Por lo tanto, resulta imprescindible recopilar información acústica que permita estimar el nivel de daño producido por la degradación en el paisaje sonoro en determinadas regiones. La recopilación de estos datos debe ser a largo plazo y en grandes extensiones del océano para permitir inferir si el incremento de nivel de ruido afecta la distribución y/o el comportamiento de las distintas especies marinas.

Los avances en la detección automatizada de eventos sonoros permiten el rápido procesamiento de grandes volúmenes de datos, incluidas vocalizaciones de animales y el sonido antropogénico que pueden ser relevantes para políticas de conservación medioambiental. Esta información es especialmente valiosa para la investigación y la gestión de los ecosistemas marinos, en los que la detección visual de especies puede ser considerablemente más difícil que en los ecosistemas terrestres [12].

En la actualidad, se han desarrollado diversas herramientas que permiten reconocer sonidos de casos particulares de animales. No obstante, es importante destacar que PAM no solo se utiliza como una herramienta para el seguimiento de la fauna marina, sino también para el estudio y análisis de actividades antropogénicas y paisajes sonoros marinos en general.

Por otro lado, hasta la fecha no existe a nivel nacional e internacional regulación de mediciones y análisis de información bioacústica para realizar una EIA en un contexto subacuático. Ante esta problemática, PAM se presenta como una herramienta viable para

el monitoreo a largo plazo en ubicaciones remotas. De esta forma se obtiene información confiable, comparable y objetiva con la cual sea posible desarrollar una metodología estandarizada de evaluación [13].

### **3.3. RUIDO MARINO**

Cuando se habla de ruido, es importante tener en cuenta la diferencia entre el ambiental y el de fondo. Estos términos se usan indistintamente, pero en realidad esto no debería ocurrir. El ambiental es aquel evento sonoro que se percibe constantemente, excluido el sonido propio generado por la recepción de hidrófonos y cualquier fuente que pueda identificarse individualmente. El ruido de fondo está formado tanto por el ambiental como todos los otros sonidos identificables individuales.

Hay casos en los que un evento identificable puede considerarse ambiental porque normalmente se encuentra presente; por ejemplo, el tráfico marítimo en un puerto concurrido o un gran grupo de mamíferos marinos haciendo vocalizaciones continuas. Si bien se puede identificar la fuente general de estos, no se pueden distinguir barcos o animales individuales.

La exploración geofísica (petróleo y gas), la construcción de parques eólicos, ejercicios militares y transporte comercial, han dado como resultado niveles de presión sonora que se estiman en diez veces más altos en la actualidad que hace algunas décadas [14]. Al igual que los efectos sobre la fauna marina, las características acústicas son específicas de cada actividad. Para la presente investigación resulta de especial interés comprender particularmente el ruido antropogénico producido por embarcaciones. En la siguiente subsección se presentan detalles de este evento acústico en cuestión.

### **3.4. RUIDO DE EMBARCACIONES**

En general, la hélice y propulsor son las fuentes de ruido preponderante para la mayoría de las embarcaciones. En particular, las que operan a alta potencia o las que operan con hélices de velocidad de rotación constante a baja potencia a menudo tienen un ruido dominante de hélice/propulsor [15]. Los buques de baja potencia y, en particular, los que cuentan con motores diésel, a menudo tienen un ruido dominante procedente de la radiación estructural excitada por la maquinaria a bordo. Algunas embarcaciones se

caracterizan por una fuerte radiación de ruido estructural debido a sistemas hidráulicos, engranajes, compresores u otra maquinaria potencialmente ruidosa. Por otro lado, las embarcaciones de alta velocidad pueden tener una fuerte radiación debido a la generación de vórtices alrededor de aberturas y discontinuidades en el casco. Los buques diseñados para producir un bajo nivel de ruido pueden tener una radiación procedente de fuentes auxiliares, como bombas, ventiladores u otras fuentes de baja potencia mecánica. Las fuentes típicas de ruido de los barcos se muestran en la Figura 1.

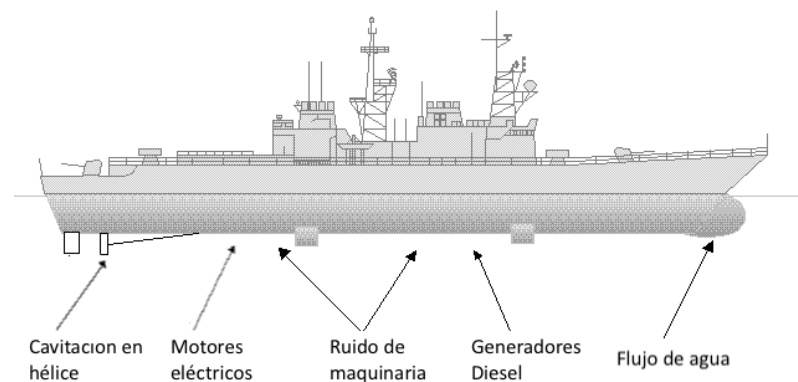


Figura 1. Representación gráfica de las fuentes de ruido en un barco.

### 3.4.1. RUIDO DE HÉLICE

El ruido de la hélice se produce por una serie de mecanismos generadores de ruido, cavitación de vórtice en la punta, diferentes tipos de cavitación de pala, cavitación de vórtice en el eje, pulsos de presión generados por las palas giratorias de la hélice y ruidos tonales debido a la resonancia entre las frecuencias naturales de las palas y los vórtices del borde de salida. En la Figura 2 se muestran diferentes tipos de ruido de cavitación.

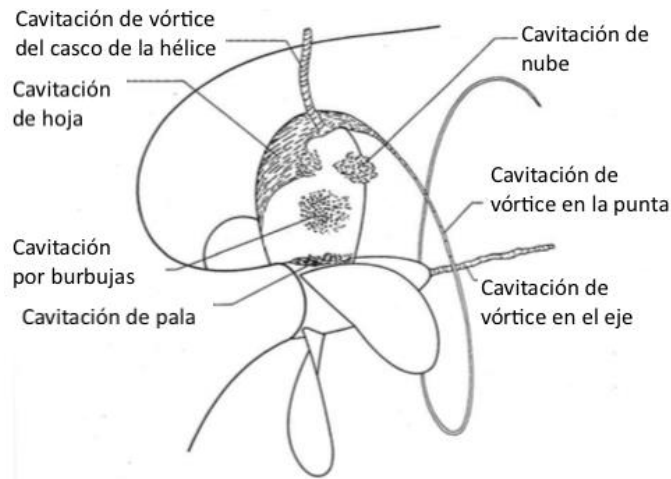


Figura 2. Representación gráfica de los tipos de cavitación en la hélice.

El ruido de la hélice depende en gran medida de la carga. A medida que aumenta la carga en las palas de la hélice, la presión en el lado de succión se vuelve lo suficientemente baja como para que ocurra la cavitación. La cavitación de las palas es muy fluctuante y normalmente será más fuerte cerca de la parte superior del diámetro de la hélice. Además, la diferencia de presión entre el lado de succión y el lado de presión de la pala de la hélice provocará vórtices.

Las hélices pueden tener diámetros grandes de hasta casi 10 m, más comúnmente entre 2 a 6 m. Por otro lado, es común encontrarse con barcos que tienen dos hélices. En tales casos, se producirá interferencia entre los dos campos de presión y la fuente de ruido será muy direccional. En todos los casos, el campo de presión se reflejará desde el casco y estos reflejos también interferirán con el campo de presión original, lo que hará que el patrón de origen sea bastante complejo. Los fenómenos de interferencia generalmente solo tienen importancia en el dominio de baja frecuencia que involucra la frecuencia de paso de la pala y los dos o tres armónicos más altos de la frecuencia de paso de la pala.

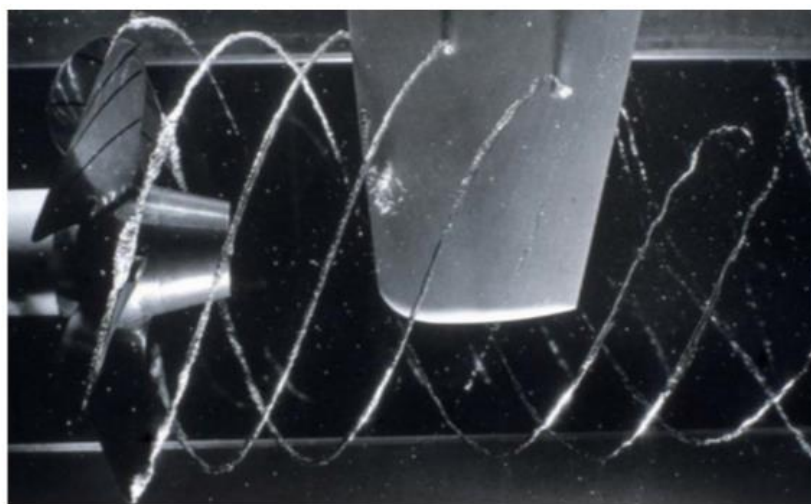


Figura 3. Cavitación de vórtice en la punta.

### 3.4.2. RUIDO DE MAQUINARIA

Una embarcación contiene una gran cantidad de maquinaria generadora de ruido. Parte de la misma está acoplada a tuberías, conductos o pozos que transmiten la energía del ruido por estructuras o fluidos sobre grandes áreas. Todas estas fuentes de ruido están contenidas en un casco parcialmente sumergido en agua. La transmisión de vibraciones estará fuertemente influenciada por la interacción entre la fuente y las placas sumergidas de la carcasa, así como las frecuencias de resonancia de las mismas. Finalmente, la vibración de las placas provoca fluctuaciones de presión en el agua, algunas de las cuales se irradian como sonido al campo lejano. El ruido de baja frecuencia puede estar asociado con la vibración global de grandes áreas del casco que se mueven en fase a una misma frecuencia, mientras que los modos de vibración locales de áreas más pequeñas del casco serán más típicos para la vibración a frecuencias más altas. Para embarcaciones de baja potencia o para embarcaciones con maquinaria firmemente sujeta, la radiación de ruido estructural puede ser dominante en general o en partes del rango de frecuencia. Las fuentes de maquinaria significativas pueden estar ubicadas a lo largo del casco.

### 3.4.3. RUIDO POR FLUJO DE AGUA

El casco en sí, así como todos los apéndices sujetos al casco, generarán turbulencias y vórtices cuando se muevan por el agua y, en ocasiones, pueden generar un

ruido significativo, especialmente a alta velocidad. Asimismo, cualquier abertura en el casco (por ejemplo, túneles de propulsión) puede generar variaciones de presión a medida que el agua pasa a través de ellas. Si la frecuencia del vórtice o la frecuencia de variación de la presión se acoplan a cualquier frecuencia natural estructural, puede surgir un fuerte ruido tonal. Dichas fuentes pueden ubicarse a lo largo de la extensión del casco y tanto el nivel como su contenido frecuencial que dependerá de la velocidad del flujo.

#### **3.4.4. COMBINACIÓN DE FUENTES**

Las diversas fuentes de ruido están ubicadas en diferentes lugares del casco, actúan a distintas profundidades, pueden variar su nivel, ubicación y el contenido de frecuencia dependiendo de las condiciones de operación. Además, un barco puede tener una eslora superior a 400 m, aunque la mayoría de los barcos suelen estar en el rango de 50 a 350 m. Por todas estas razones, es muy complejo describir un barco como una fuente de ruido ideal. Por otro lado, si se registra el ruido producido desde una distancia razonablemente larga, puede parecer una fuente puntual monopolar. Sin embargo, a larga distancia, las propiedades de transmisión del mar y la influencia del ruido de fondo pueden afectar el evento acústico registrado dificultando la capacidad de detección o clasificación del mismo.

La capacidad para determinar las características acústicas de las distintas fuentes depende también del ruido de fondo en la zona de grabación. Teniendo en cuenta que la pérdida por dispersión esférica teórica es de 40 dB a una distancia de 100 m, un barco relativamente silencioso puede quedar enmascarado con el ruido de fondo a una distancia de esta magnitud, mientras que un barco ruidoso puede reconocerse fácilmente a distancias mucho más largas.

La mayoría de los barcos se moverán a cierta velocidad durante la observación, mientras que la posición de grabación se realizará desde una ubicación fija. Por lo tanto, el aspecto y las distancias relativas entre las diferentes fuentes de ruido y el observador cambiarán continuamente. Además, la directividad de las diferentes fuentes puede causar una variación significativa en el nivel aparente de la fuente a medida que cambia el aspecto. Esta es la razón por la cual describir con precisión una embarcación como fuente de ruido a partir de las observaciones en el campo lejano se vuelve una tarea sumamente compleja.

### 3.5. TRANSPORTE MARÍTIMO

En la actualidad, los barcos son una de las principales fuentes de ruido antropogénico en el océano. Debido a que no solo aumenta la cantidad de embarcaciones sino también su tamaño y potencia de propulsión, el nivel producido también lo hace proporcionalmente. El transporte marítimo no se distribuye uniformemente por todo el océano. Este se concentra en áreas donde se encuentran rutas marítimas y puertos, principalmente en el hemisferio norte.

Si bien el sonido producido por barcos es predominantemente de baja frecuencia, también es de banda ancha debido a la cavitación producida por la rotación de hélices. Dada la naturaleza del mismo, el sonido de una sola embarcación puede viajar grandes distancias y posiblemente combinarse con otras fuentes sonoras aumentando los niveles de ruido ambiental.

En términos de grandes embarcaciones comerciales, se encuentran diferentes tipos como transporte de pasajeros y carga, buques cisterna, pesca, dragado, suministro en alta mar, rompehielos, investigación y militares, cada uno con su propia firma acústica. McKenna et al. [16] midieron niveles de presión sonora producidos por siete tipos de barcos comerciales en el Canal de Santa Bárbara. Registraron niveles de presión y combinándolo con sistemas de información de tráfico marítimo, pudieron estimar los niveles de presión que generan referenciado a 1 metro de distancia. Los niveles estimados oscilan entre 182,1 y 188,1 dB (1  $\mu$ Pa RMS - 1 m) para frecuencias entre 20 y 1000 Hz. Por otro lado, afirman que el ruido es generado principalmente por cavitación, presentando un pico máximo entre 50 y 150 Hz pero que puede extenderse hasta 10 kHz.

En su investigación, los autores encontraron que la radiación sonora en los barcos evaluados es asimétrica, presentando diferencias de entre 5 a 10 dB entre popa y proa. Con el fin de enriquecer la caracterización del evento sonoro, en la Figura 4 se muestran espectrogramas de 3 distintos tipos de embarcaciones.

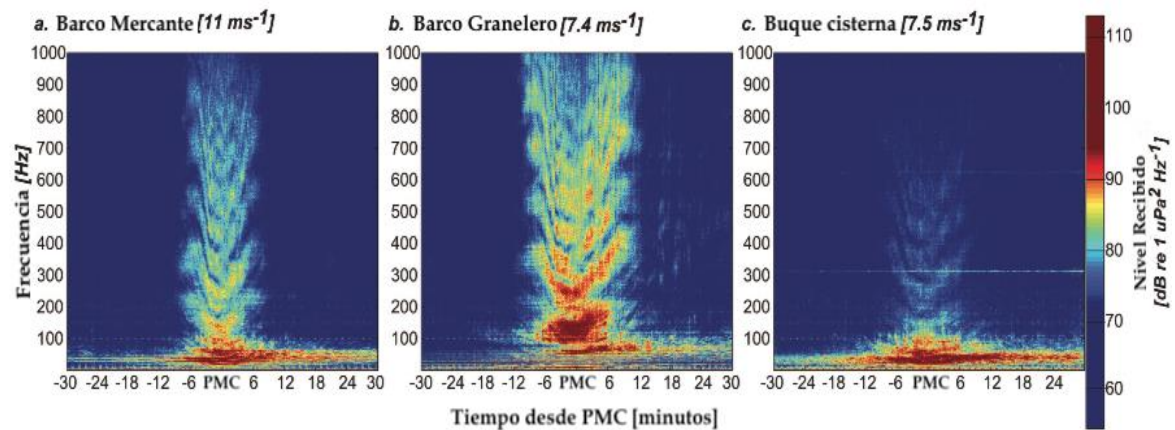


Figura 4. Espectrogramas de distintos tipos de barcos [28].

A partir de la Figura 4 puede afirmarse que el ruido generado depende considerablemente del tipo de navío. Se observan dos características principales, componentes tonales por debajo de 120 Hz y una distribución energética con forma de U por encima de esta.

Las líneas tonales son el resultado de la cavitación de las palas de la hélice y sus armónicos. Por otro lado, los patrones de interferencia en forma de U son explicados principalmente por la interferencia constructiva y destructiva producto de interacciones entre distintos caminos de propagación (por ejemplo, sonido directo y reflejado en la superficie marítima) [17]. Este fenómeno es conocido como efecto espejo de Lloyd y depende de la profundidad de la fuente y el receptor, la distancia de la fuente al receptor, las propiedades de la columna de agua y la reflexión del fondo [18].

Jensen et al. [19] midieron niveles de ruido de dos embarcaciones pequeñas utilizadas habitualmente con fines recreativos en zonas costeras. Como resultado obtuvieron niveles de presión entre 138 y 152 dB (1  $\mu$ Pa RMS - 1 m) para frecuencias entre 0,2 a 40 kHz. Por otro lado, al analizar la distribución energética de las señales encontraron que si bien sigue siendo predominantemente en bajas frecuencias, cuenta con mayor proporción de energía por encima de 1 kHz en comparación con barcos de mayor tamaño.

El potencial de propagación a larga distancia del ruido de baja frecuencia generado por embarcaciones comerciales supone un gran riesgo para mamíferos marinos. En particular, cetáceos que se comunican en dichas bandas de frecuencia corren riesgo de enmascaramiento, lo que puede tener múltiples impactos negativos. Por ejemplo, se



incrementa el riesgo de colisión de los cetáceos y puede alterar sus patrones migratorios y comportamentales. Estos problemas son evitables si se diseñan rutas de navegación que no sean coincidentes con áreas claves para el normal desarrollo de la vida marina.

### 3.6. DESCRIPTORES ACÚSTICOS

En la Tabla 1 se presentan los descriptores estandarizados que se utilizan en la presente investigación.

Tabla 1. Descriptores acústicos subacuáticos

Parámetro			Definición
<b>Centroide Espectral</b>	[Hz]	<b>Spectral Centroid</b>	Indica la posición del "centro de gravedad" del espectro de acuerdo a las amplitudes frecuenciales
<b>Ancho de Banda Espectral</b>	[Hz]	<b>Spectral Bandwidth</b>	Rango de frecuencias en el espectro donde se encuentra concentrada la mayor parte de la energía.
<b>Planitud Espectral</b>		<b>Spectral Flatness</b>	Medida de la variabilidad espectral que indica qué tan plana es la distribución de amplitudes.
<b>Tasa de Cruce por Cero</b>		<b>Zero-Crossing rate</b>	Cantidad de veces que la señal cruza el eje horizontal en relación a su tiempo.
<b>Estimación de Fundamental</b>	[Hz]	<b>Fundamental Estimation</b>	Estimación de la frecuencia fundamental de una señal de audio.
<b>Coeficiente Cepstral de Frecuencia Mel</b>		<b>Mel-frequency cepstral coefficient</b>	Coeficientes que representan la forma espectral de una señal de audio en una escala mel.
<b>Espectrograma de Mel</b>		<b>Mel Spectrogram</b>	Representación espectral de una señal de audio en la escala mel.
<b>Contraste Espectral</b>		<b>Spectral Contrast</b>	Medida que cuantifica las diferencias en amplitud entre regiones del espectro de audio.

Si bien existen otros parámetros acústicos comúnmente utilizados en sistemas de detección y clasificación automática, se han seleccionado los descriptores presentados anteriormente debido a su origen en la reconocida biblioteca librosa [20]. Esta elección se basa en que dicha librería es ampliamente utilizada en la comunidad científica para problemáticas similares a las que intenta abordar la presente tesis [21,22]. Al utilizar estos descriptores, se obtiene información clave que facilita la identificación y distinción de diversos eventos acústicos, lo que contribuye a la solidez y relevancia de la investigación.

A continuación, se presentan algunas ecuaciones necesarias para calcular los parámetros acústicos utilizados en la presente investigación.

### **3.6.1. ENERGÍA POR TERCIOS DE OCTAVA**

En la norma IEC 61260-3 [23] se definen los intervalos de frecuencias estandarizados donde se obtiene un único valor representativo de la energía que contenga cada banda. Como se muestra en la ecuación 1, se calcula como la sumatoria de la energía sobre el resultado de la transformada de Fourier discreta (DFT por sus siglas en inglés) filtrada.

$$E(i) = \sum_{t=1}^k h(X(i)) \quad (1)$$

Donde  $X(i)$  representa el valor de magnitud de cada banda de frecuencia,  $i$  representa cada una de las bandas y  $h$  representa el filtrado de la señal, en este caso en particular, por tercios de octava.

### **3.6.2. FIRMA ACÚSTICA**

Una firma acústica ( $F_a$ ) es la caracterización frecuencial de toda fuente sonora. Toda fuente acústica posee una distribución en frecuencia que la hace diferenciable del resto. La misma depende de cualidades físicas de generación de sonido. Por otro lado, el medio de propagación influye en la manera en que la señal acústica es captada por un receptor.

El análisis de firma acústica es un complemento importante usualmente utilizado en PAM para rastrear buques de guerra. Por lo tanto, resulta en un componente de especial interés en tareas de clasificación de distintas embarcaciones.

### 3.7. REPRESENTACIÓN DIGITAL DE SEÑALES.

Para el correcto procesamiento de señales acústicas resulta considerablemente importante tener en cuenta ciertos aspectos de la representación digital de las mismas.

#### 3.7.1. FRECUENCIA DE MUESTREO

La frecuencia de muestreo se define como la tasa de muestras por segundo que se toman sobre una señal analógica. Es un factor importante a tener en cuenta ya que está directamente relacionada con la utilización de recursos computacionales y con la frecuencia máxima que puede representar la señal digital. A una mayor frecuencia de muestreo, mayor cantidad de datos para representar la señal. Por lo tanto, si bien es posible representar frecuencias superiores esto supone un incremento en el costo computacional.

#### 3.7.2. DECIMACIÓN

Decimar es un proceso donde se reduce la cantidad de muestras por segundo de una representación digital buscando optimizar el costo computacional de un algoritmo. En la Figura 5 se muestra ilustrativamente como se realiza el proceso de decimación.

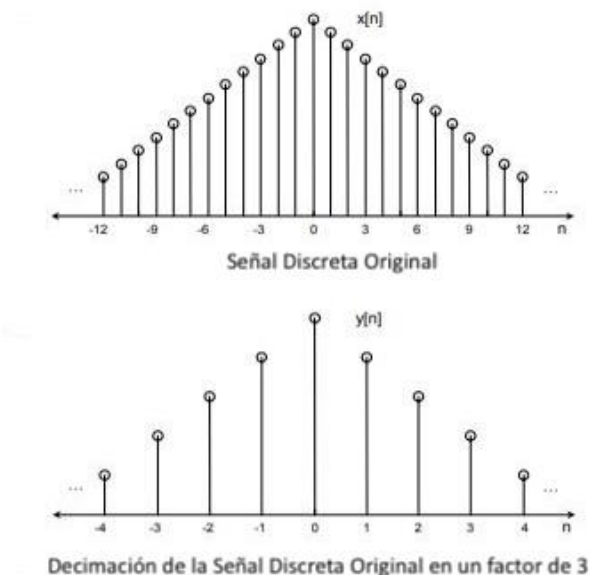


Figura 5. Representación gráfica del proceso de decimación.

### 3.7.3. NORMALIZACIÓN

La normalización surge de analizar y procesar la información digital para lograr estandarización con el objetivo de trabajar sobre un rango bien definido de amplitudes. En otras palabras, mediante la normalización se busca ajustar valores medidos en diferentes escalas a una escala común con el objetivo de poder hacer comparaciones coherentes entre conjuntos de muestras. Este proceso es considerablemente necesario en sistemas de reconocimiento de patrones y particularmente en algoritmos basados en aprendizaje automático. En la presente investigación se utilizan las normalizaciones definidas en las ecuaciones 2 y 3.

$$X_n = \frac{X_n}{X_{max}} \quad (2)$$

$$X_n = \frac{X_n - \bar{X}}{s} \quad (3)$$

Donde  $X_n$  representa el valor de magnitud de cada muestra,  $X_{max}$  el valor de magnitud máxima del conjunto de muestras,  $\bar{X}$  su valor promedio y  $s$  su desvío estándar.

### 3.8. ALGORITMOS DE DETECCIÓN Y CLASIFICACIÓN EN GRANDES BASES DE DATOS

Durante la última década, el costo de recopilar y almacenar datos acústicos se ha reducido drásticamente, por lo que suele ser normal recopilar terabytes de información en un solo proyecto [24,25]. A medida que aumenta el volumen de datos acústicos, el proceso de extracción de información relevante se vuelve más costoso y largo.

El aprendizaje automático tiene el potencial de identificar señales de interés dentro de grandes conjuntos de datos de forma relativamente barata y con mayor consistencia que los analistas humanos [26]. Hasta el momento, se han utilizado diversos algoritmos con el objetivo de lograr una clasificación precisa y eficiente. Entre los algoritmos empleados se encuentran Support Vector Machine [27], k-Nearest Neighbors [28], Multilayer Perceptron [29], Random Forest [30] y Decision Tree [31]. Estos algoritmos han demostrado su eficacia en trabajos previos relacionados con la clasificación de sonidos en diferentes dominios, incluyendo la identificación de embarcaciones [32–35].

En el ámbito de la detección y clasificación de grandes volúmenes de información, como los generados por grabaciones en PAM, se han observado avances significativos

mediante el uso de redes neuronales profundas. Estas arquitecturas han demostrado tener tasas bajas de falsos positivos en comparación con otros algoritmos alternativos empleados en este campo. Sin embargo, es esencial tener en cuenta que la capacidad de generalización de los resultados depende en gran medida de la diversidad y representatividad del conjunto de datos de entrenamiento utilizado. Si el conjunto de datos carece de variabilidad en términos de las situaciones de grabación y los fenómenos acústicos capturados, la capacidad del algoritmo para obtener buenos resultados se verá limitada a las condiciones específicas en las que se recopilaron los registros acústicos utilizados para entrenar el modelo. Por lo tanto, es crucial asegurarse de contar con un conjunto de datos extenso y diverso, que abarque una amplia gama de situaciones y escenarios acústicos, a fin de lograr resultados más generales y aplicables en diferentes contextos.

Cabe destacar que los algoritmos mencionados se encuentran disponibles para su implementación en la biblioteca scikit-learn (sklearn) [36]. Esta biblioteca se destaca por su alta calidad de documentación, su diseño claro y su facilidad de uso, lo que la convierte en una herramienta accesible y ampliamente adoptada para el aprendizaje automático. Estas características hacen que sklearn sea especialmente útil para la presente tesis, ya que permite comparar resultados obtenidos con diferentes tipologías de algoritmos. Por otro lado, es importante tener en cuenta que la selección adecuada de los hiperparámetros puede tener un impacto significativo en el rendimiento y la generalización de estos algoritmos. Los hiperparámetros pueden definirse como parámetros externos al modelo que influyen en el proceso de entrenamiento y afectan su capacidad para generalizar y obtener resultados satisfactorios en datos nuevos.

Resulta imprescindible mencionar algunas de las principales fuentes de error en el entrenamiento de modelos de aprendizaje automático. Por un lado, el sesgo se refiere a las simplificaciones excesivas que subestiman las relaciones complejas en los datos, lo que puede introducir errores en los resultados del modelo. Por otro lado, la varianza se relaciona con la sensibilidad del modelo a las fluctuaciones en los datos de entrenamiento y puede conducir al sobreajuste, donde el modelo se adapta demasiado a los detalles específicos de los datos de entrenamiento y tiene dificultades para generalizar en datos nuevos. Además, es importante tener en cuenta la existencia de un tipo de sesgo debido a

la manera en que los datos son recopilados, conocido como sesgo en los datos. Este sesgo puede ocurrir por una recopilación, grabación o etiquetado inadecuado de datos, lo que afecta el rendimiento y la capacidad de generalización del modelo.

A continuación, con el objetivo de introducir características principales e hiperparámetros, se presenta una breve explicación de los algoritmos que se utilizan en la presente investigación.

### 3.8.1. K-NEAREST NEIGHBORS (KNN)

El algoritmo kNN es un método simple y eficiente en el que un objeto se clasifica según la mayoría de sus vecinos más cercanos en el espacio de características. En otras palabras, el algoritmo busca los  $k$  vecinos más cercanos a ese punto en función de una medida de distancia y le asigna una etiqueta de clase basándose en la categoría que aparece con mayor frecuencia entre sus vecinos más cercanos. En la Figura 6, se representa gráficamente la lógica de funcionamiento del algoritmo en cuestión.

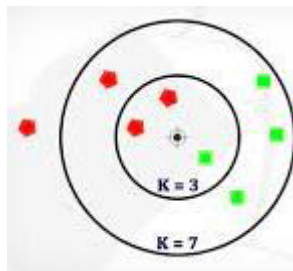


Figura 6. Representación gráfica del funcionamiento de kNN.

Al implementar kNN con sklearn, uno de los hiperparámetros clave es el valor de " $k$ ", que representa el número de vecinos más cercanos considerados para la clasificación. Otros hiperparámetros incluyen la métrica de distancia utilizada para calcular la cercanía entre los puntos y las estrategias de ponderación utilizadas para asignar diferentes pesos a los vecinos.

### 3.8.2. SUPPORT VECTOR MACHINE (SVM)

El algoritmo SVM es ampliamente utilizado en la clasificación de datos debido a su capacidad para construir límites de decisión óptimos entre diferentes clases. SVM busca encontrar un hiperplano en que mejor separe las clases. El objetivo es encontrar el

hiperplano que tenga la mayor distancia posible a los puntos más cercanos de cada categoría, llamados vectores de soporte. En la Figura 7, se representa gráficamente la lógica de funcionamiento del algoritmo en cuestión.

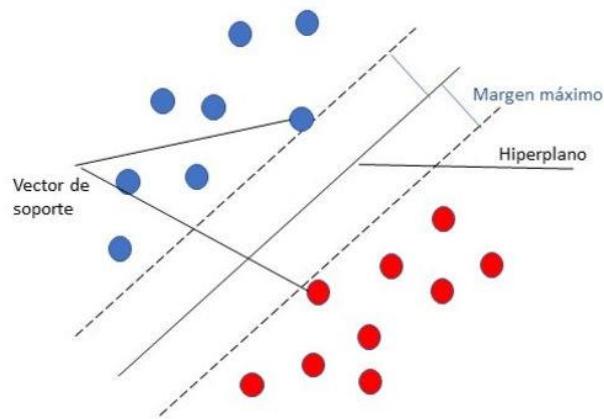


Figura 7. Representación gráfica del funcionamiento de SVM.

Al implementar SVM con sklearn, se pueden ajustar varios hiperparámetros. El más importante es el tipo de kernel utilizado, que determina la forma del hiperplano y la capacidad del algoritmo para manejar datos no lineales. Algunos ejemplos comunes de kernels son el lineal, el polinómico y el radial (RBF). Además, se incluye el parámetro de regularización  $C$ , que controla la penalización por clasificar incorrectamente puntos de datos. Un valor más alto de  $C$  indica un ajuste más estricto a los datos de entrenamiento, lo que puede conducir a una mayor precisión, pero también a un mayor riesgo de sobreajuste. Por otro lado, el hiperparámetro  $\gamma$  afecta el alcance de influencia de un solo ejemplo de entrenamiento. Un valor más alto de  $\gamma$  significa que solo los puntos de datos más cercanos tendrán un impacto significativo en la clasificación, mientras que un valor más bajo implica un alcance más amplio de influencia.

### 3.8.3. ÁRBOL DE DECISIÓN (DT)

El árbol de decisión es un algoritmo de aprendizaje automático utilizado para clasificar datos. Utiliza una estructura de árbol donde cada nodo interno representa una característica y cada rama representa una posible respuesta a esa característica. Las hojas

del árbol representan las clases o categorías finales a las que se asignan los datos. El árbol de decisión busca establecer reglas claras de decisión y distinguir diferentes clases basándose en las características de entrada.

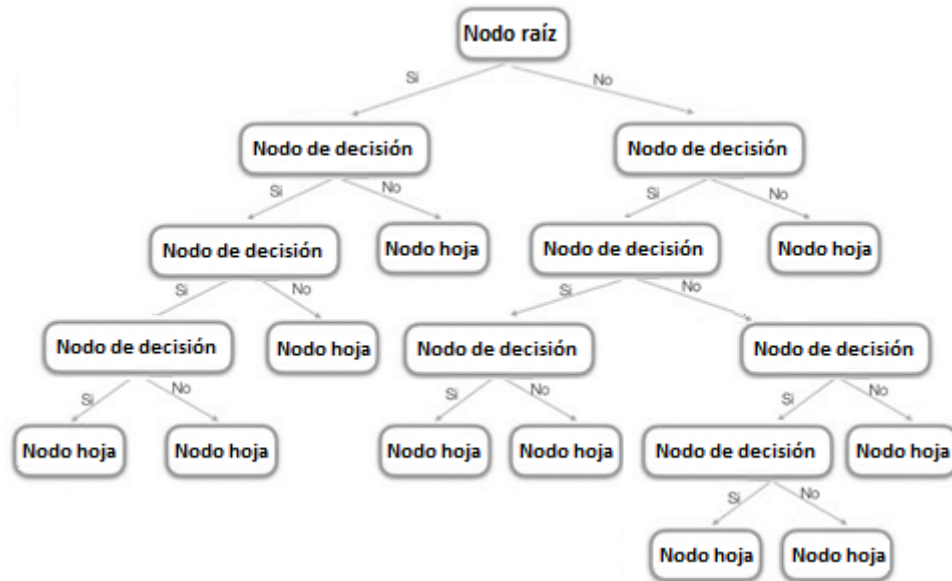


Figura 8. Representación gráfica del funcionamiento de un árbol de decisión.

Al implementar el árbol de decisión con sklearn, se pueden ajustar hiperparámetros como el criterio de división y la profundidad máxima del árbol. El criterio de división, determina la medida utilizada para evaluar la calidad de una división en el árbol, y puede ser "gini" o "entropía". El criterio "gini" se basa en el índice de impureza de Gini, mientras que "entropía" utiliza la entropía de información para medir la homogeneidad de las divisiones. Por otro lado, la profundidad máxima del árbol, limita el número de niveles que puede tener el árbol, lo que puede afectar la capacidad para capturar relaciones complejas en los datos. De establecerse un valor elevando de este hiperparámetro, se incrementa la tendencia al sobreajuste del modelo.

### 3.8.4. RANDOM FOREST

El algoritmo Random Forest es un método de aprendizaje de conjunto que utiliza múltiples árboles de decisión para realizar la clasificación. Cada árbol de decisión en el conjunto se entrena en una muestra aleatoria de los datos de entrenamiento y genera una predicción. Luego, la predicción final se determina mediante la combinación de las predicciones individuales de cada árbol.



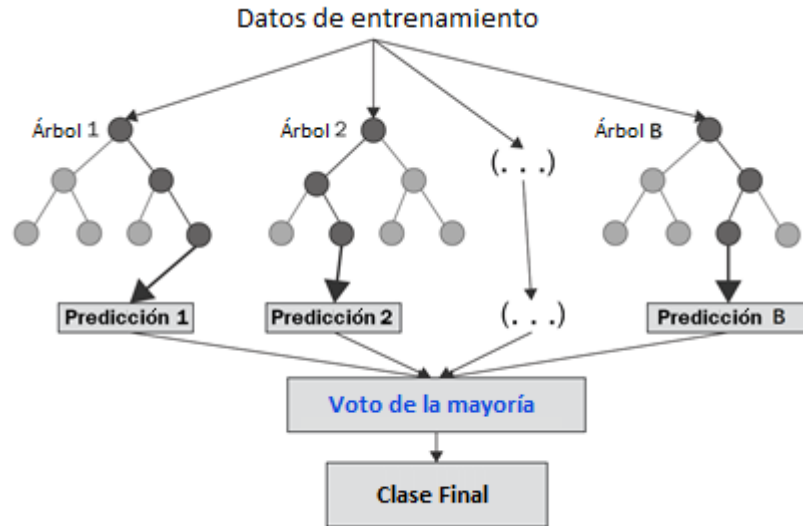


Figura 9. Representación gráfica del funcionamiento de Random Forest.

Al implementar Random Forest con sklearn, se pueden ajustar varios hiperparámetros como el número de árboles en el conjunto (`n_estimators`), la profundidad máxima de los árboles (`max_depth`), el número mínimo de muestras requeridas para dividir un nodo interno (`min_samples_split`), el criterio de división utilizado para evaluar la calidad de una división (`criterion`), el número mínimo de muestras requeridas en las hojas finales (`min_samples_leaf`) y la cantidad máxima de características consideradas en cada división (`max_features`).

### 3.9. MÉTRICAS DE DESEMPEÑO

Para caracterizar el funcionamiento de sistemas de detección y clasificación basados en aprendizaje automático se suelen utilizar los descriptores que se presentan a continuación [37,38].

- **Exactitud:** Es la relación entre los resultados correctos sobre el total de los resultados obtenidos. Este parámetro en particular suele dar resultados imprecisos cuando las bases de datos están desbalanceadas.

$$Exactitud = \frac{PV + NV}{PV + NV + PF + NF} \quad (4)$$

Donde PV (positivo-verdadero) y NV (negativo-verdadero) son los resultados positivos y los negativos clasificados correctamente. PF (positivo-falso) y NF (negativo-falso) son los resultados clasificados incorrectamente.

- Precisión: Definida como la relación entre los resultados positivos correctos sobre el total de los resultados positivos obtenidos.

$$Precisión = \frac{PV}{PV + PF} \quad (5)$$

- Exhaustividad: Es la relación entre los resultados positivos correctos sobre el total de los casos positivos.

$$Exhaustividad = \frac{PV}{PV + NF} \quad (6)$$

- Valor F1: Se utiliza para combinar las medidas de precisión y exhaustividad en un solo valor.

$$F1 = 2 \cdot \frac{Precisión \cdot Exhaustividad}{Precisión + Exhaustividad} \quad (7)$$

## 4. DISEÑO DE LA INVESTIGACIÓN, VARIABLE Y MUESTRA

Para el desarrollo de la presente investigación, resultó imprescindible la obtención de una gran cantidad de datos para el entrenamiento de los algoritmos. En la Figura 10 se muestra un diagrama en bloque de los pasos necesarios para el desarrollo de la presente investigación.

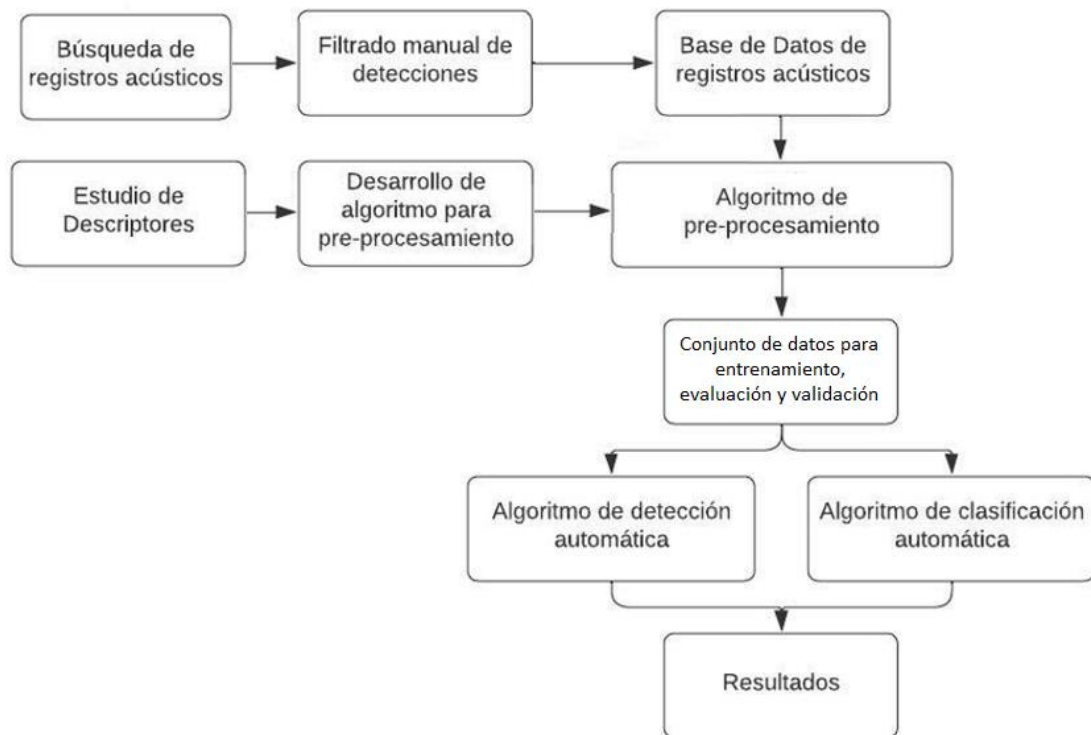


Figura 10. Diagrama en bloques del diseño de la investigación.

En primer lugar, se realizó una búsqueda exhaustiva de registros acústicos donde pudiera haber señales de embarcaciones. Seguidamente, a partir la inspección visual de sus espectrogramas se generaron los conjuntos de datos.

En segundo lugar, se llevó a cabo un estudio de las distintas señales acústicas de interés con el fin de seleccionar y diseñar los descriptores a utilizar como entrada para los posteriores algoritmos. En este sentido, se desarrolló un algoritmo con el objetivo de pre procesar las muestras de audio y conseguir los parámetros que conforman los conjuntos de datos de entrada para la etapa de detección y clasificación.

En tercer lugar, se desarrolló y evaluó el desempeño de los distintos modelos. El entrenamiento y validación se llevó a cabo a partir de los registros antropogénicos representativos del fenómeno acústico en cuestión.

#### **4.1. MANEJO DE DATOS**

En el presente estudio, los eventos de interés se consiguieron a partir de la detección de sonidos mediante la inspección visual de sus espectrogramas y la posterior confirmación auditiva correspondiente. Para llevar a cabo esta tarea se utilizó un programa especializado para inspección de espectros y señales temporales bioacústicas. En este caso en particular se utilizó el software Raven Pro [39].

Dependiendo de la cantidad de eventos disponibles, la base de datos se complementó con otras fuentes de datos alternativas para conseguir un mayor volumen de muestras. Vale la pena mencionar que los datos de BD son utilizados tanto en etapa de entrenamiento como para la validación.

#### **4.2. COMPOSICIÓN DE LA BASE DE DATOS.**

Para la conformación de la base de datos a utilizar para entrenar, probar y validar los algoritmos se utilizan diversas fuentes de registros sonoros. Para el desarrollo del sistema de detección automática resultó necesario conformar una base de datos que cuente con 2 conjuntos de datos: embarcaciones (E) y no embarcaciones (NE). En ambos casos, para determinar que se encuentra, o no, un evento acústico generado por embarcaciones se realizó una inspección visual de sus espectrogramas. Para el desarrollo del sistema de clasificación automática, se construyó una base de datos que consta de tres conjuntos de datos: embarcaciones grandes (GE) con eslora mayor a 20 m, embarcaciones medianas (ME) con eslora entre 12 y 20 m, y embarcaciones pequeñas (PE) con eslora menor a 12 m. Estos conjuntos de datos surgieron al subdividir en tres categorías distintas el grupo de embarcaciones (E) utilizado durante la etapa de detección.

En primer lugar, se analizaron las bases de datos provistas por la Fundación Cethus. A partir de esta experiencia se obtienen 27 archivos de 1 minuto de duración. Se considera que esta cantidad de eventos acústicos es insuficiente para poder entrenar, probar y validar los algoritmos a desarrollar. Además, al desconocerse el tipo de embarcación que generó

cada señal, estas muestras no son de utilidad para algoritmos basados en aprendizaje supervisado. Por lo tanto, se procedió a buscar posibles bases de datos de público acceso en internet.

A partir de un extenso estudio de los distintos trabajos de investigación desarrollados, se encuentra que en la mayoría de los casos utilizan sus propias bases de datos privadas. Por esta razón, se intenta establecer contacto con sus autores para solicitar su colaboración. A consecuencia de esta operación se consiguieron dos bases de datos externas distintas.

La primera base de datos, provista por Alfio Yori Fernández<sup>2</sup>, consta de 69 archivos de audio con duración variable, siendo el archivo más corto de 15 segundos. Las grabaciones fueron realizadas cerca del Golfo Corcovado en la Patagonia chilena. Cada muestra se encuentra etiquetada y se conoce el largo de eslora de cada embarcación. Cabe mencionar que la base de datos no cuenta con grabaciones de ruido de fondo. Además, se desconoce la metodología utilizada para la grabación de los audios.

La segunda base de datos [40], proporcionada por David Santos-Domínguez<sup>3</sup>, consta de 94 archivos de audio. Estas grabaciones fueron realizadas en las proximidades del puerto de Vigo, situado en el noroeste de España. Al igual que en la base de datos anterior, la duración de las grabaciones varía, siendo el registro de menor duración de 15 segundos. Las muestras también se encuentran etiquetadas, aunque no se dispone de información sobre las esloras de las embarcaciones. No obstante, se cuenta con el nombre de cada embarcación y fotografías lo que permitió obtener la longitud de cada una de ellas a través del proyecto MarineTraffic<sup>4</sup> para realizar la clasificación.

Considerando la variabilidad de los registros acústicos conseguidos, se estableció un formato para las muestras de audio con el objetivo de conseguir una estandarización para su posterior procesamiento.

---

<sup>2</sup> Alfio Yori Fernandez, Doctor en Ciencias de la Ingeniería Fernández de la Universidad Austral de Chile

<sup>3</sup> David Santos-Domínguez, Ingeniero en Acoustics Lab, AtlanTTic Research Center, Universidad de Vigo

<sup>4</sup> MarineTraffic es una plataforma en línea que proporciona información sobre el tráfico marítimo a nivel mundial. Link de acceso: [www.marinetraffic.com](http://www.marinetraffic.com)

Cada evento que conforma la base de datos es de un único canal, con una duración de 15 segundos, 16 bits de resolución y una frecuencia de muestreo de 22050 Hz. Se eligió esta frecuencia de muestreo porque permite la correcta representación de frecuencias inferiores a 10 kHz, suficientes para caracterizar el fenómeno acústico de interés. Por otro lado, de los archivos de mayor duración se extrajeron como máximo 5 submuestras consiguiendo una cantidad superior de muestras de duración estandarizada.

La justificación para extraer un máximo de 5 submuestras se basa en la intención de evitar la creación de un desequilibrio entre las muestras extraídas de diferentes archivos de audio. Al tener una cantidad excesiva de muestras de un solo archivo y pocas de otros en comparación, se podría generar un sesgo en el análisis posterior que se realice. De esta manera, se asegura que cada archivo tenga entre 1 y 5 muestras, evitando así una sobrerrepresentación o subrepresentación de algún archivo en particular en el conjunto de datos analizados.

A partir de la extracción de submuestras explicada anteriormente, se obtuvieron 131 muestras del primer conjunto de datos, distribuidas en 39 GE, 54 ME y 38 PE. En cuanto al segundo conjunto de datos, se consiguieron 448 muestras, distribuidas en 137 GE, 98 ME, 75 PE y 138 NE. De esta manera, se logra una distribución equitativa y representativa de los diferentes archivos de audio en ambos conjuntos de datos, lo que permitió realizar análisis comparativos y evaluaciones objetivas en el marco de la investigación.

Debido a que la cantidad de muestras en los distintos grupos de las bases de datos es significativamente diferente, se consideró necesario asegurar una cantidad similar de muestras de ambas bases de datos durante el entrenamiento del modelo. De esta forma se evita que el modelo tenga un sesgo hacia una base de datos en particular y se consiga un rendimiento desequilibrado en la clasificación de nuevas muestras. Para lograr esto, se optó por descartar muestras que hayan sido extraídas de un mismo archivo, hasta lograr homogeneizar la cantidad de muestras en los distintos grupos de ambas bases de datos. Este proceso consistió en eliminar progresivamente las muestras provenientes de archivos de los que se les extrajeron una mayor cantidad, con el objetivo de lograr conjuntos de datos balanceados en términos de la cantidad de muestras por archivo.

Por último, dado que no se dispone de grabaciones de ruido de fondo en el primer conjunto de datos, se decidió extraer muestras de ruido de fondo de grabaciones de la

Fundación Cethus para completar dicha categoría en el conjunto de entrenamiento del modelo. Aunque estas grabaciones no provienen directamente de las zonas de estudio, su inclusión permite ampliar la variedad de registros de ruido de fondo en el conjunto de entrenamiento y mejorar la capacidad de generalización del modelo.

En la Tabla 2 se presenta la distribución final de las muestras en cada uno de los grupos de embarcaciones y de ruido de fondo en ambas bases de datos.

Tabla 2. Distribución de muestras de distintos grupos.

Origen de DB	Tipo de registro	Categoría de embarcación	Cantidad
Alfio Yori Fernández	E	GE	39
		ME	40
		PE	38
David Santos- Domínguez	E	GE	40
		ME	40
		PE	40
	NE	-	120
Fundación Cethus	NE	-	120

La combinación de las bases de datos se considera una estrategia apropiada para el desarrollo de algoritmos de detección y clasificación basados en inteligencia artificial. La utilización de múltiples bases de datos puede aportar diversos beneficios en este ámbito, como una mayor cantidad de datos disponibles para entrenar los modelos y una mayor

diversidad en los registros acústicos, lo que puede mejorar la capacidad de generalización de los modelos. Además, cada conjunto de datos puede presentar diferentes características y limitaciones que pueden ser compensadas por la inclusión de otros conjuntos. Por esta razón, para entrenar los algoritmos se utilizaron ambas bases de datos agrupadas como un único conjunto de datos.

En consecuencia, para la etapa de detección del modelo se dispone de 237 muestras de embarcaciones (E) y 240 muestras de no embarcaciones (NE). Por otro lado, para la etapa de clasificación se utilizan un total de 79, 80 y 78 muestras para GE, ME y PE respectivamente.

### **4.3. DESCRIPTORES ACÚSTICOS**

A partir de la revisión del estado del arte de sistemas de detección y clasificación de ruido de embarcaciones es que se decide utilizar la firma acústica de los registros sonoros como uno de los descriptores del evento acústico. La misma se obtiene a partir de una reducción de la DFT en bandas de tercio de octava estandarizadas por la norma IEC 61260-3. Para los fines de esta investigación se utilizan las bandas definidas entre 5 y 4000 Hz.

En la Figura 11 se muestra, a modo de ejemplo, la representación de un espectrograma de embarcaciones. En el gráfico superior puede observarse la representación de 0 a 2,6 kHz mostrando la característica forma en U de su espectrograma mientras que en el gráfico inferior se observa en detalle el comportamiento tonal del evento acústico en cuestión.



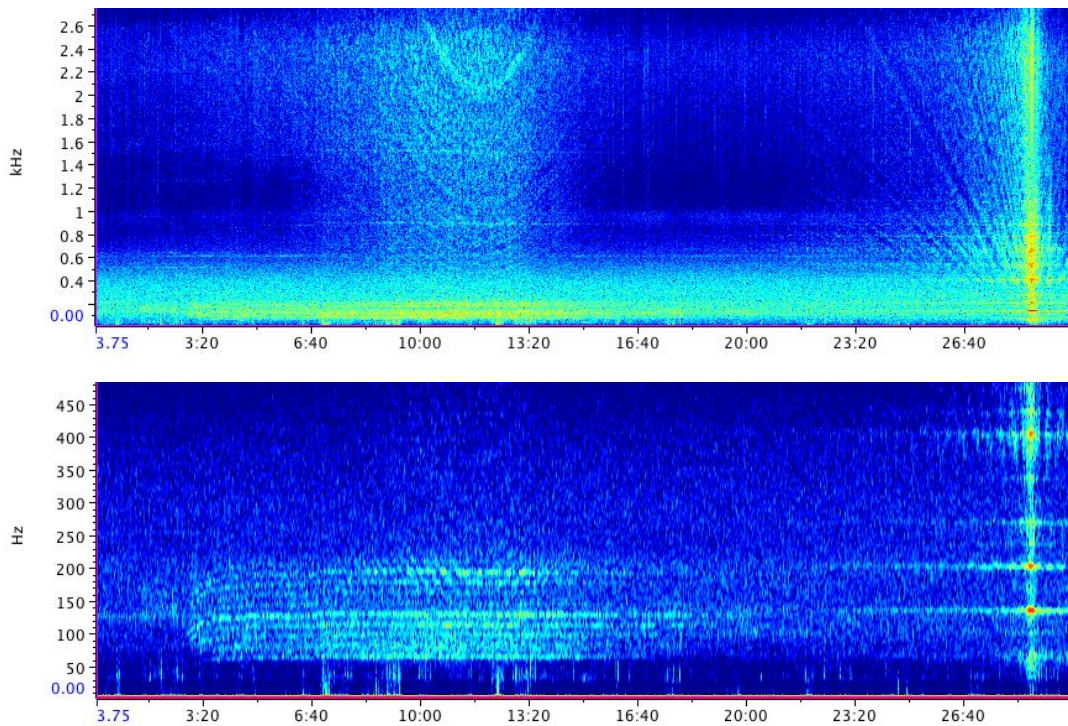


Figura 11. Ejemplo de detección visual de embarcaciones en software Raven Pro 1.6.  
Representación completa (arriba) y representación de su contenido tonal (abajo).

Por otra parte, debido a la naturaleza del fenómeno acústico, el cual se distingue por tener componentes tonales, se desarrollaron 2 descriptores personalizados. El primero de ellos se utiliza para calcular un índice de tonalidad. Este parámetro busca representar la cantidad de componentes tonales en los registros acústicos mediante un único valor. El segundo descriptor representa coeficientes de tonalidad por bandas de frecuencias. El mismo está basado en el análisis espectral de la señal y considera tanto la estabilidad temporal de la energía como la relación con bandas de frecuencia adyacentes.

Por otro lado, se calcularon algunos parámetros disponibles en la biblioteca librosa, con el objetivo de buscar diferencias entre resultados de las distintas categorías. A partir de la comparación de los resultados obtenidos de cada parámetro, se podría inferir cuáles parámetros serían de mayor utilidad como entrada en los posteriores algoritmos de detección y clasificación. Es decir, aquellos parámetros que presenten una marcada diferencia entre las distintas categorías son considerados como los más eficientes para discriminar las señales. Es importante destacar que se utilizó la biblioteca librosa debido a que esta es ampliamente aceptada en sistemas de procesamiento de señales acústicas. Además, es una herramienta de código abierto y cuenta con una gran cantidad de usuarios

y desarrolladores en la comunidad científica. Por lo tanto, su uso garantiza la compatibilidad y facilidad de integración con otras herramientas de procesamiento de señales acústicas, lo que facilita el análisis y la comparación de resultados obtenidos por otros investigadores.

#### **4.3.1. ALGORITMO DE CÁLCULO DE DESCRIPTORES.**

A partir de las muestras de audio se necesita calcular y obtener los distintos parámetros acústicos y variables para conformar los conjuntos de datos de entrenamiento y evaluación a utilizar por los posteriores algoritmos de detección y clasificación automática. Se desarrolló un algoritmo con cinco etapas: carga de archivos de audio; obtención de la firma acústica; cálculo del índice de tonalidad; obtención de las frecuencias picos y sus amplitudes relativas; cálculo de parámetros acústicos estandarizados y exportación de los resultados.

En primer lugar, se leen los nombres de todos los archivos que se encuentren dentro de un directorio establecido y se procede a cargar cada uno de ellos de manera individual. A continuación, se realiza la normalización con el objetivo de estandarizar las amplitudes de las señales y, si es necesario, se lleva a cabo una decimación para asegurar que todas las muestras tengan una frecuencia de muestreo ( $f_s$ ) de 11025 Hz. Este proceso de estandarización es de vital importancia para independizarse del nivel que tenga cada señal y garantiza que los resultados obtenidos del cálculo de los distintos parámetros acústicos sean comparables entre sí. La elección de dicha frecuencia se basó en que 11025 Hz es la  $f_s$  mínima de los archivos en los cuales se extraen las muestras. Además, esta  $f_s$  permite representar correctamente hasta aproximadamente 5500 Hz, lo cual se considera suficiente para los fines de la presente investigación.

Mediante el empleo de un algoritmo auxiliar se obtiene la respuesta en frecuencia en tercios de octava. Si por el contrario se decidiera utilizar filtros de banda de octava, las características tonales se verían comprometidas por el proceso de integración energética que las conforma. La firma acústica (FA) surge del procesamiento de la muestra mediante filtros de sexto orden, dando como resultado 30 valores correspondientes a bandas de frecuencia entre 5 y 4000 Hz. Se decide utilizar este rango de frecuencias ya que es donde se supone que se almacena la mayor cantidad de información del evento acústico en

cuestión. Finalmente, buscando optimizar el proceso de aprendizaje de los algoritmos, cada firma acústica es normalizada. Se dividen los valores de las bandas por el máximo de la FA y de esta manera se logra una representación energética con un valor máximo de 1. Esta normalización puede hacerse ya que la información relevante en la FA es la relación de amplitudes entre bandas de frecuencia en lugar de los valores absolutos de las mismas.

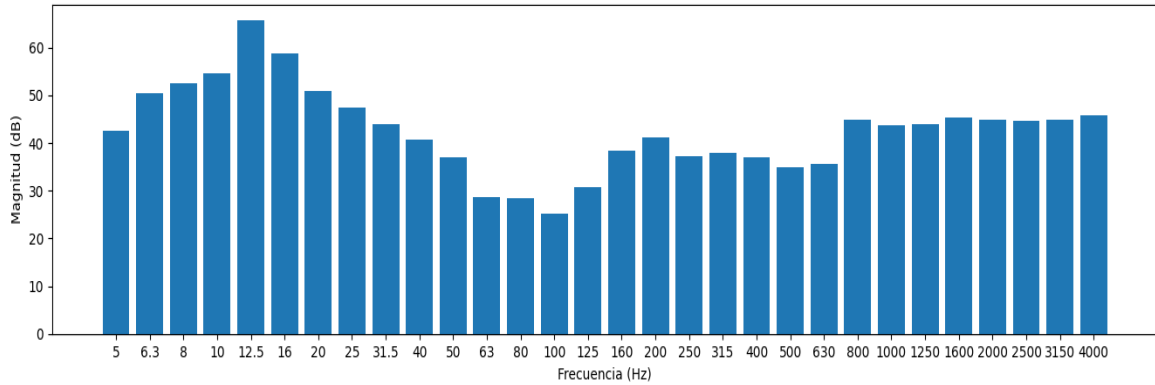


Figura 12. Representación gráfica de firma acústica

Para obtener el índice de tonalidad, se calcula la respuesta en frecuencia del registro acústico, la cual luego es suavizada mediante un filtro de media móvil para obtener su envolvente frecuencial. A continuación, se identifican los picos de frecuencia que superan en 15 dB a un umbral definido por la envolvente frecuencial. Para calcular el descriptor se utiliza una función sigmoide modificada, la cual se ajusta para que los resultados estén comprendidos entre 0 y 1, siendo valores más cercanos a 1 indicativos de una mayor cantidad de picos tonales presentes.

$$IT = 2 \cdot \left( \frac{1}{1 + e^{-\frac{\text{num\_picos}}{\text{num\_frec}} * 12000}} - 0,5 \right) \quad (8)$$

Donde "IT" es el índice de tonalidad, "num<sub>picos</sub>" es el número de picos de frecuencia identificados, "num<sub>frec</sub>" es la cantidad de elementos frecuenciales que se obtienen a partir de la FFT. Para definir la función sigmoide se resta 0,5 al valor obtenido, de manera que esta quede definida entre -0,5 y 0,5, y posteriormente se multiplica por 2 para que el límite superior sea 1. El factor de escala de 12000 que multiplica en el exponente presente en el denominador se ha obtenido empíricamente y permite tener valores bajos para menos de 3 picos detectados y valores altos para más de 6 picos detectados.

En la Figura 13 se muestra, a modo de ejemplo, la representación de la respuesta en frecuencia en rojo, el umbral de detección de picos en azul y los picos detectados como cruces negras.

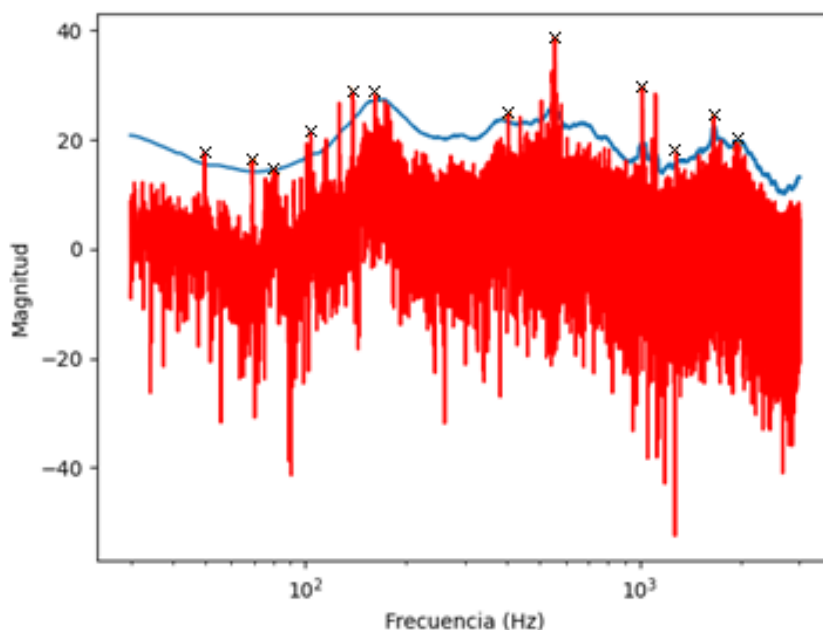


Figura 13. Representación de la detección de picos.

El segundo descriptor personalizado es un vector de índices de tonalidad que se obtiene mediante el procesamiento del espectrograma de cada archivo de audio. En primer lugar, se calcula el espectrograma y se restringen las frecuencias entre 20 y 200 Hz. A continuación, se realiza un recorrido por cada bin de frecuencia en el espectrograma para calcular su respectivo índice de tonalidad. Este cálculo considera tanto la estabilidad temporal de la energía en el bin como la diferencia de energía con las frecuencias adyacentes.

Posteriormente, se divide el espectro en subrangos de frecuencia previamente establecidos, los cuales están diseñados para tener una mayor resolución frecuencial en las bajas y una menor en las altas frecuencias. En cada subrango, se identifica la frecuencia que exhibe el índice de tonalidad máximo, y se registran tanto los valores máximos de los índices de tonalidad como las correspondientes frecuencias en listas independientes. En la Tabla 3 se describen dichos subrangos en detalle.

Tabla 3. Distribución frecuencial de los subrangos.

Subrangos de Frecuencias	Ancho de Banda
20 - 50 [Hz]	5 [Hz]
50 - 80 [Hz]	10 [Hz]
80 - 120 [Hz]	20 [Hz]
120 – 200 [Hz]	40 [Hz]

Por último, se aplica una transformación basada en una ecuación sigmoide modificada al vector de índices de tonalidad. Esta transformación tiene como objetivo conseguir valores tendientes a 0 para índices de tonalidad inferiores a 0,3, y valores cercanos a 1 para índices superiores a 0,7. De esta manera, se logra una mejor representación y discriminación de las características tonales presentes en la señal de audio.

$$tonVect = \left( \frac{1}{1 + e^{-(tonMax*11-5)}} \right) \quad (9)$$

Donde "tonVect" representa los índices de tonalidad modificados, mientras que "tonMax" corresponde a los índices de tonalidad originales. Para lograr esto, se resta un valor constante de 5 a cada elemento de "tonMax", lo que tiene como efecto desplazar la curva sigmoide hacia la parte positiva del eje x. Además, se utiliza un factor de escala de 11 en el exponente de la función sigmoide para ajustar el ancho de la curva en el eje x. De esta manera, los resultados obtenidos mediante la función sigmoide están limitados entre 0 y 1 para valores típicos de índices de tonalidad en el rango de 0,1 a 0,8.

En la Figura 14 se presenta, a modo de ejemplo, los índices de tonalidad en el espectrograma. El nivel de intensidad de las líneas punteadas refleja el valor del índice de

tonalidad, lo que permite identificar claramente las frecuencias donde el índice alcanza valores considerables y desaparece cuando tiende a cero.

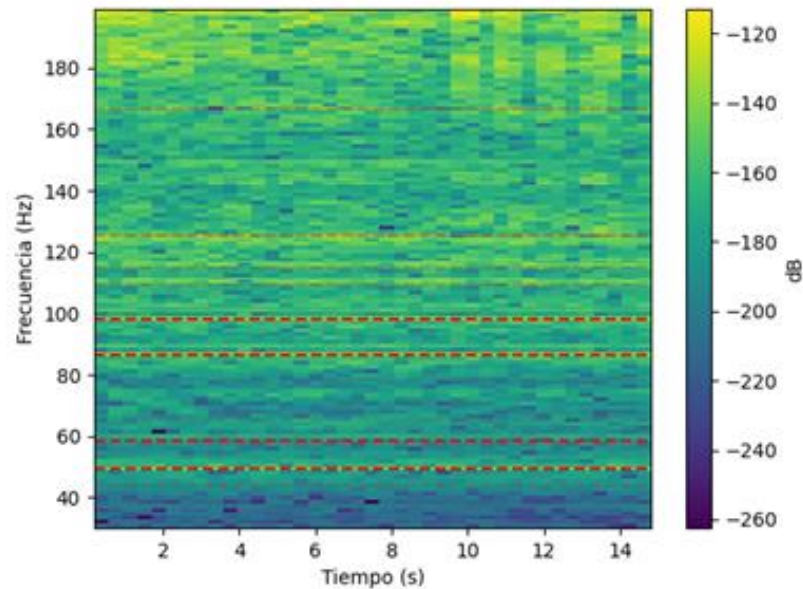


Figura 14. Representación de índices de tonalidad en espectrograma.

Por otro lado, se procede a calcular los parámetros acústicos estándar. Se calculan los siguientes descriptores disponibles en la biblioteca librosa:

- Spectral Centroid
- Spectral Bandwidth
- Spectral Flatness
- Zero-Crossing rate
- Fundamental Estimation (método de Yin)
- Mel-frequency cepstral coefficient (MFCC)
- Mel Spectrogram
- Spectral Contrast

Una vez obtenidos los resultados de todos los parámetros acústicos utilizando la biblioteca librosa, se procede a su procesamiento para reducir la cantidad de datos y facilitar su análisis.

Algunos de los resultados obtenidos son vectores (por ejemplo, Spectral Centroid), los cuales son producto del cálculo del parámetro en ventanas temporales. Para reducir la

cantidad de datos de estos vectores, se calcula su mediana. Se calcula este estadístico en lugar de su media debido a que no presentan una distribución suficientemente normal.

Otros descriptores resultados son matrices (por ejemplo, MFCC) que proporcionan una representación en dos dimensiones (temporal y frecuencial). Para simplificar estas matrices, se calcula la mediana en la dimensión temporal, manteniendo la dimensión frecuencial separada en las distintas bandas de frecuencia. De esta manera, se obtiene un vector por cada matriz de resultados, lo que permite una representación más compacta y manejable de los datos.

Por último, el algoritmo se encarga de concatenar parámetros estandarizados y personalizados en vectores para almacenarlos y exportarlos en un archivo con formato “.csv”.

#### **4.4. PRE-PROCESAMIENTO DEL DATASET.**

En esta etapa, se lleva a cabo el preprocesamiento de las muestras de datos acústicos con el objetivo de prepararlos adecuadamente antes de utilizarlos como entrada en los algoritmos de detección y clasificación de embarcaciones. Se divide en tres etapas principales: el acondicionamiento de la información, la gestión de outliers y la normalización de los datos. En la Figura 15 se presenta un diagrama en bloques que muestra las tres etapas antes mencionadas.

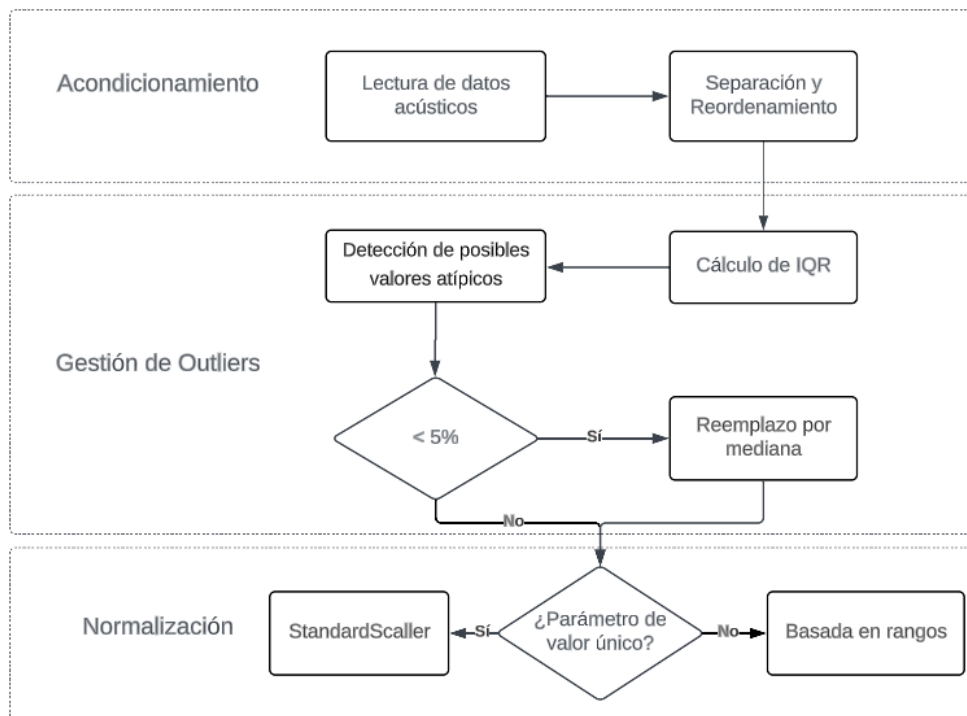


Figura 15. Diagrama en bloques de la etapa de pre-procesamiento.

El proceso de acondicionamiento de la información comienza con la lectura de los datos acústicos almacenados en archivos Excel. Estos archivos contienen los parámetros acústicos asociados a las muestras de embarcaciones. Una vez cargados los datos, se procede a realizar una serie de manipulaciones con el fin de garantizar su calidad y coherencia.

En primer lugar, se lleva a cabo la gestión de outliers. Los cuales pueden entenderse como valores atípicos que se encuentran fuera del rango esperado en un conjunto de datos. En este caso, se decidió reemplazarlos por la mediana correspondiente a cada parámetro acústico, a menos que la cantidad de potenciales outliers representen más del 5% del conjunto de datos. Esta decisión se basa en la consideración de que, de obtenerse una gran cantidad de resultados por fuera del rango, esto puede deberse a una característica propia de la dispersión de parámetro y no necesariamente a un valor atípico. Por otro lado, es importante señalar que debido al conjunto de datos no era considerablemente extenso, en lugar de descartar las muestras consideradas como valores



atípicos, se optó por reemplazarlas por la mediana. Esto permitió evitar la reducción en la cantidad de datos del conjunto.

Para identificar los outliers, se utiliza el criterio del rango intercuartil (IQR, por sus siglas en inglés), calculando los límites inferior y superior como 1,5 veces el IQR por debajo y por encima del primer y tercer cuartil respectivamente.

Además, se aplica una normalización de los datos para asegurar que todas las características se ubiquen dentro de un rango comparable y mejorar así la performance de las etapas subsiguiente. Para los parámetros de valor único, se utiliza el escalador "StandardScaler" disponible en la biblioteca sklearn, el cual transforma los datos para que tengan una media de cero y una desviación estándar de uno. Por otro lado, para los parámetros formados por vectores con múltiples elementos, se realiza una normalización basada en su rango de valores con el objetivo de preservar las relaciones de valores entre los elementos del vector. Esta decisión se fundamenta en la necesidad de mantener la información sobre las relaciones de valores existentes entre los elementos que componen cada vector.

#### **4.5. DETECTOR AUTOMÁTICO DE EMBARCACIONES.**

En esta sección, se aborda la experimentación y desarrollo de detectores automáticos de embarcaciones basado en parámetros acústicos y técnicas de aprendizaje automático. El proceso de construcción del detector se divide en dos etapas fundamentales: pruebas experimentales con distintos modelos y la evaluación de la capacidad de generalización de los mismos. En la Figura 16 se presenta un diagrama en bloques que muestra las dos etapas antes mencionadas.

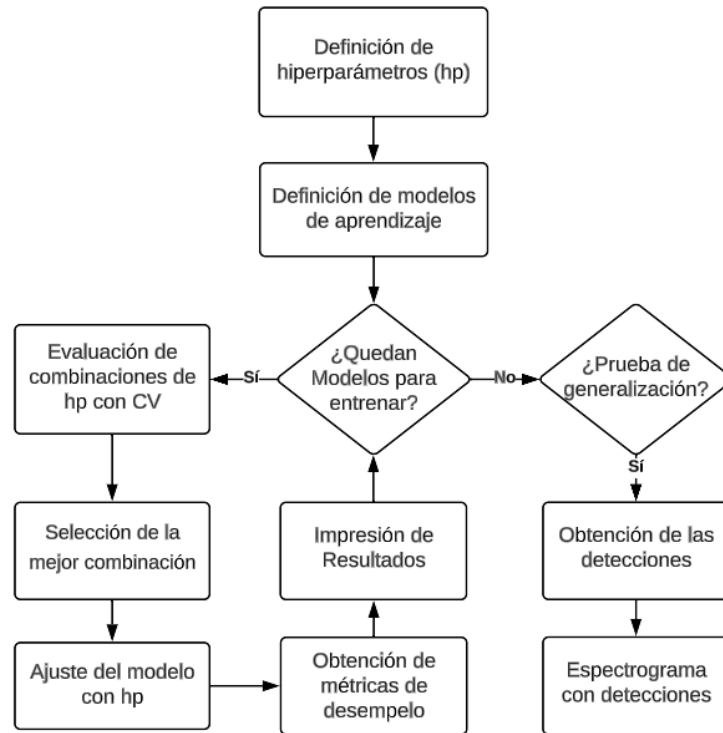


Figura 16. Diagrama en bloques de la etapa de detección.

#### 4.5.1. EXPERIMENTACIÓN CON ALGORITMOS PARA DETECTORES.

En esta etapa, se procede a evaluar el desempeño de distintos algoritmos de aprendizaje automático para determinar cuál de ellos es el más idóneo para los fines de la presente tesis.

Considerando las características específicas de los parámetros acústicos y la disponibilidad limitada de muestras para desarrollar sistemas basados en redes neuronales profundas se desarrollan y evalúan los siguientes algoritmos: KNN, SVM, RF y DT. Estos algoritmos son ampliamente utilizados en el campo del aprendizaje automático y se consideran pertinentes para la tarea de detección y clasificación de embarcaciones.

Para cada algoritmo, se lleva a cabo una búsqueda exhaustiva de los hiperparámetros óptimos mediante el empleo de la herramienta GridSearchCV de la biblioteca sklearn. Esta técnica de optimización permite explorar sistemáticamente diversas combinaciones de hiperparámetros y determinar aquellas que maximizan el rendimiento del modelo.

GridSearchCV utiliza validación cruzada para evaluar el rendimiento del modelo en diferentes particiones del conjunto de entrenamiento. En lugar de usar un único conjunto

de entrenamiento, este se divide en subconjuntos llamados "folds" y se realiza el entrenamiento múltiples veces. Esto proporciona estimaciones más robustas del rendimiento del modelo. Es menester mencionar que los subconjuntos de entrenamiento y evaluación se obtuvieron a partir de dividir en un 80% y 20% el conjunto de datos total respectivamente. Además, en el presente desarrollo, se establece en 5 la cantidad de folds para realizar la validación cruzada.

Una vez ajustados los hiperparámetros, se obtiene el rendimiento de cada modelo mediante el cálculo de las métricas de desempeño explicadas en el marco teórico sobre el conjunto de datos de evaluación. A partir de los resultados obtenidos, se efectúa una comparación exhaustiva entre los algoritmos evaluados, teniendo en cuenta tanto su rendimiento como sus características intrínsecas. Dicha evaluación contribuye a la selección del algoritmo más adecuado para el conjunto de datos y los parámetros acústicos utilizados.

#### **4.5.2. EVALUACIÓN DE CAPACIDAD DE GENERALIZACIÓN Y SESGO.**

Con el propósito de evaluar la capacidad de generalización de los modelos, se procesan grabaciones proporcionadas por la fundación Cethus utilizando el mismo flujo de pre-procesamiento mencionado anteriormente. Estas grabaciones contienen eventos acústicos generados por embarcaciones, aunque no se conoce su categoría específica. Por lo tanto, se decide utilizar estas grabaciones para evaluar los diferentes modelos de detección en ejemplos reales externos al conjunto de datos de entrenamiento. Para este fin, se utilizan tres archivos de audio de 30 minutos de duración, donde se observa la presencia de embarcaciones mediante la inspección visual de los espectrogramas.

En primer lugar, se segmentan los archivos en fragmentos de 15 segundos y se calculan los parámetros acústicos correspondientes. En segundo lugar, estos fragmentos se someten posteriormente a los diferentes modelos de detección para su evaluación. Es importante destacar que, aunque se emplearon muestras de las grabaciones proporcionadas por la fundación Cethus para ampliar el conjunto de datos de no embarcaciones, las muestras correspondientes a las embarcaciones extraídas de estas grabaciones son independientes de las muestras utilizadas en el proceso de entrenamiento. Por lo tanto, esta experimentación permite evaluar la capacidad de

generalización de los diferentes modelos en situaciones reales y desvinculadas del conjunto de entrenamiento.

Esta etapa de evaluación tiene como objetivo determinar la capacidad de los modelos para identificar las embarcaciones en registros que presentan el fenómeno completo de un barco en circulación. Es decir, se analizan archivos que contienen instantes tanto con ruido de fondo sin embarcaciones como con la presencia de embarcaciones.

Los resultados que se obtienen de esta experimentación proporcionan información adicional sobre el comportamiento de los modelos y sirven como una validación adicional en un contexto de detección de embarcaciones más representativo de la realidad. Además, permiten evaluar el sesgo inherente a los sistemas de grabación utilizados y las regiones en donde se registraron los eventos acústicos entre otros factores.

#### **4.6. CLASIFICADOR AUTOMÁTICO DE EMBARCACIONES.**

En esta sección, se desarrollan y evalúan modelos de clasificadores automáticos de embarcaciones basados en parámetros acústicos y técnicas de aprendizaje automático. Dado que la disponibilidad de muestras es limitada, en esta etapa se centra en la realización de pruebas experimentales con diferentes modelos de clasificación sin la posibilidad de evaluar la capacidad de generalización y el posible sesgo del conjunto de datos.

En este contexto, se realizó un proceso de preprocesamiento de los datos acústicos, siguiendo los mismos pasos descritos en la sección 3.4. Posteriormente, se empleó un enfoque similar al explicado en la sección 3.5 para la experimentación con los distintos modelos. La principal diferencia radica en que en este caso se utilizaron muestras divididas en cuatro categorías: GE, ME, PE y NE.

Con el objetivo de evaluar el rendimiento de los modelos de clasificación, se obtienen las métricas de desempeño descritas en el marco teórico para cada una de las categorías mencionadas. Estas métricas proporcionan una medida cuantitativa del desempeño de los modelos en la clasificación de las embarcaciones en función de sus características acústicas. Mediante la comparación de las métricas de desempeño, se busca identificar el algoritmo más idóneo para la clasificación de las embarcaciones en cada una de las categorías.

## 5. RESULTADOS

### 5.1. PARÁMETROS ACÚSTICOS Y VARIABLES DE ENTRADA.

Con el objetivo de evaluar la capacidad de los descriptores como datos de entrada para los algoritmos de detección y clasificación, se llevan a cabo dos enfoques diferentes para representar los resultados. Por un lado, se presentan los descriptores de valor único en una tabla comparativa. Por otro lado, para los resultados que son vectores con un conjunto de valores, se emplea un código en Python para generar distintos gráficos. Todos los resultados se presentan diferenciados entre las categorías previamente mencionadas (NE, GE, ME y PE).

La Tabla 4 muestra los valores de mediana y rango intercuartil de los parámetros acústicos. Estas métricas se seleccionaron debido a que los parámetros no mostraron una distribución normal según el test de Kolmogorov-Smirnov [41].

Tabla 4. Mediana (Q1-Q3) de los descriptores únicos.

Parámetro	GE	ME	PE	NE
Spectral Flatness	0,066 (0,037-0,104)	0,057 (0,039-0,091)	0,087 (0,049-0,123)	0,018 (0,002-0,131)
Spectral Centroid [Hz]	1737,2 (1526,5-1913,2)	1634,3 (1452,8-1909,9)	1950,9 (1698,7-2346,0)	1410,6 (321,5-2036,2)
Spectral Bandwidth [Hz]	1381,6 (1294,4-1495,2)	1381,1 (1282,7-1509,2)	1467,6 (1420,9-1532,6)	1367,8 (747,9-1490,4)
Zero-crossing rate	0,20 (0,16-0,25)	0,18 (0,14-0,25)	0,26 (0,16-0,32)	0,05 (0,00-0,24)
Frec. Fundamental [Hz]	19,4 (14,5-22,4)	17,7 (14,5-21,2)	20,2 (16,2-30,0)	18,3 (16,7-20,4)
Índice de Tonalidad	0,83 (0,6-0,94)	0,83 (0,58-0,92)	0,87 (0,79-0,94)	0,01 (0,00-0,26)

Como puede observarse en la tabla 4, no existe un parámetro individual que permita clasificar de manera definitiva entre las diferentes categorías de embarcaciones. Los valores de los distintos parámetros se encuentran en un rango similar entre las categorías. Sin embargo, se observa una tendencia general en algunos parámetros acústicos. Por ejemplo, las PE muestran valores relativamente más altos en el parámetro Spectral Centroid en comparación con las ME y GE. Esto podría sugerir que las embarcaciones pequeñas tienden a tener una distribución de energía más centrada hacia frecuencias superiores.

Por otro lado, se observa que el índice de tonalidad muestra resultados insatisfactorios para clasificar entre los diferentes tipos de embarcaciones. No obstante, este parámetro es el que presenta una mayor capacidad para distinguir entre embarcaciones y no embarcaciones. Se evidencia esta diferenciación en el rango de valores de Q1 a Q3, que varía entre 0,00 y 0,26 para las no embarcaciones, mientras que las embarcaciones presentan valores de Q1 superiores a 0,58. En otras palabras, se observa que el 75% de las muestras correspondientes a E presentan un índice de tonalidad superior a 0,58 mientras que el 75% de las muestras correspondientes a NE exhiben un índice de tonalidad inferior a 0,26. Este resultado es consistente con el comportamiento tonal producto de la cavitación generada por la rotación de hélices que se menciona en el marco teórico. La Figura 17 muestra un gráfico de cajas donde se evidencia dicha separación.

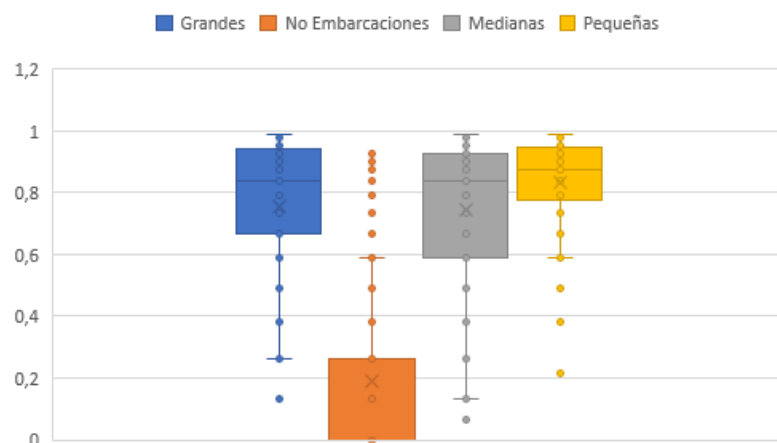


Figura 17. Representación de resultados del índice de tonalidad.

Al analizar los gráficos de los parámetros acústicos representados como vectores, en un primer análisis no se logra apreciar una diferenciación marcada entre las distintas categorías (NE, GE, ME y PE). Sin embargo, es importante destacar que estos gráficos pueden revelar detalles o patrones complejos que podrían resultar relevantes para sistemas de detección y clasificación basados en inteligencia artificial. Con el objetivo de explorar en mayor profundidad en busca de características relevantes, a continuación, se presentan los gráficos correspondientes a cada parámetro, seguidos de su análisis pertinente.

### 5.1.1. MEL-FREQUENCY CEPSTRAL COEFFICIENT (MFCC)

En la Figura 18 se presentan los resultados obtenidos del cálculo de los MFCC.

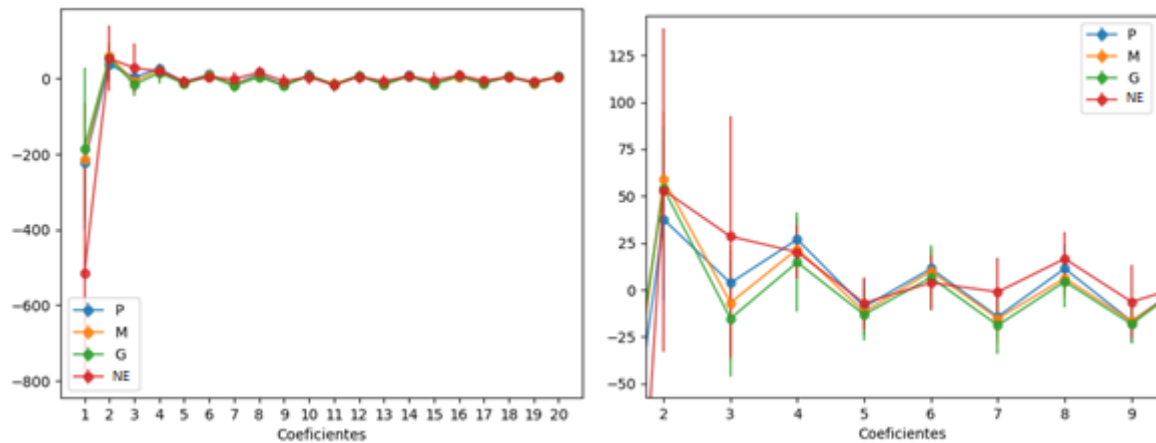


Figura 18. Representación de los MFCC

Como se puede apreciar en la representación completa de los MFCC, no se observa una diferenciación clara entre las distintas categorías. Sin embargo, al graficar el rango de coeficientes del 2 al 9, se presenta una diferencia en la distribución de resultados entre E (P, M y G) y NE. Esto es especialmente notable en el tercer coeficiente, donde NE muestra una mediana superior a 25, mientras que las embarcaciones presentan valores cercanos a 0. Además, se observa que a medida que aumenta el tamaño de la embarcación, dicho coeficiente parece disminuir. Sin embargo, es importante mencionar que los valores de los coeficientes están dentro del rango de Q1-Q3, lo que indica que este análisis no es determinante por sí solo para la clasificación de las embarcaciones.

### 5.1.2. MEL SPECTROGRAM

En la Figura 19 se presentan los resultados obtenidos del cálculo de los espectrogramas de Mel.

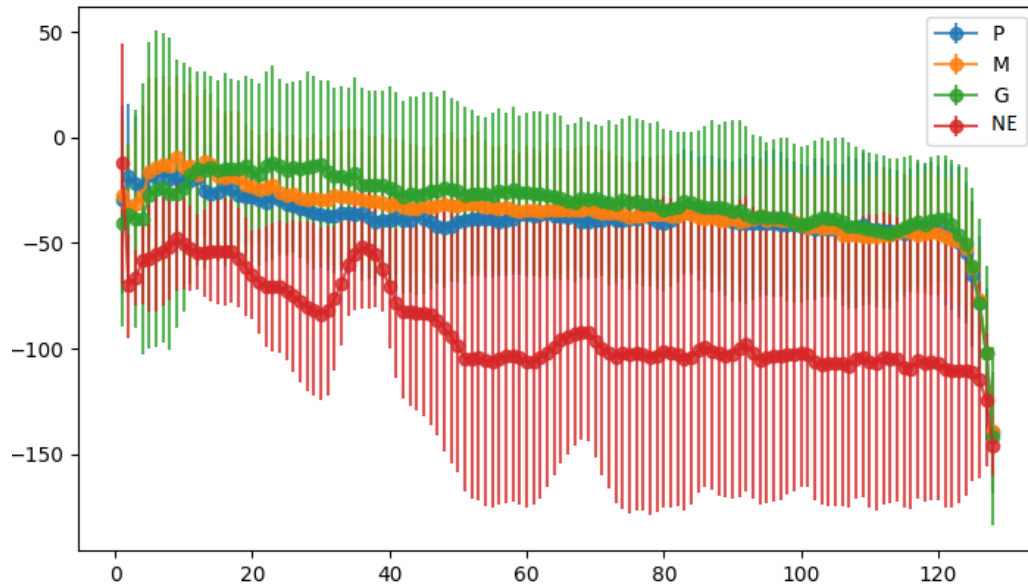


Figura 19. Representación de los espectrogramas de Mel.

En la representación de los espectrogramas de Mel, se puede observar una leve diferenciación entre las categorías de embarcaciones, especialmente en los coeficientes inferiores a 60. Sin embargo, es importante destacar que la diferenciación no es tan clara debido a que todos los resultados se encuentran dentro de los rangos intercuartiles (Q1-Q3) para las distintas categorías. A pesar de esto, al utilizar algoritmos de inteligencia artificial, no se descarta la posibilidad de que este parámetro sea útil, ya que esta tecnología tiene la capacidad de identificar y aprovechar características complejas para la clasificación.

Por otro lado, la categoría NE presenta resultados considerablemente inferiores en comparación con todas las categorías de embarcaciones. Esto sugiere que los espectrogramas de Mel pueden ser útiles para distinguir entre E y NE. Sin embargo, es importante destacar que se observa una alta dispersión de los coeficientes en la categoría NE, donde los valores se superponen con los de las embarcaciones (E). Esta variabilidad puede afectar el potencial de separación entre categorías.



### 5.1.3. SPECTRAL CONTRAST

En la Figura 20 se presentan los resultados obtenidos del cálculo del Contraste Espectral.

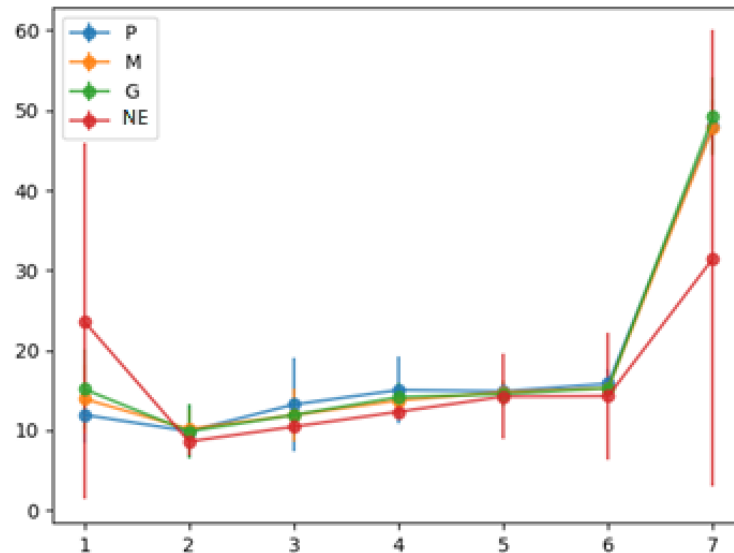


Figura 20. Representación del Contraste Espectral.

En la representación del Contraste Espectral, no se aprecia una diferenciación clara entre las distintas categorías. Si bien se observa una diferencia en los resultados de la categoría de no embarcaciones para los coeficientes 1 y 7 en comparación con las demás categorías, los resultados de las E están dentro del rango intercuartil de NE. Esto indica que, aunque existen algunas variaciones en los valores del Contraste Espectral, no proporcionan una diferenciación inequívoca entre las categorías.

### 5.1.4. FIRMA ACÚSTICA

Como se menciona en la sección de descriptores acústicos, la firma acústica (FA) se obtiene mediante el filtrado por tercios de octava en el rango de frecuencias de 5 y 4000 Hz. En la figura 21 se presentan los resultados obtenidos del cálculo de FA.

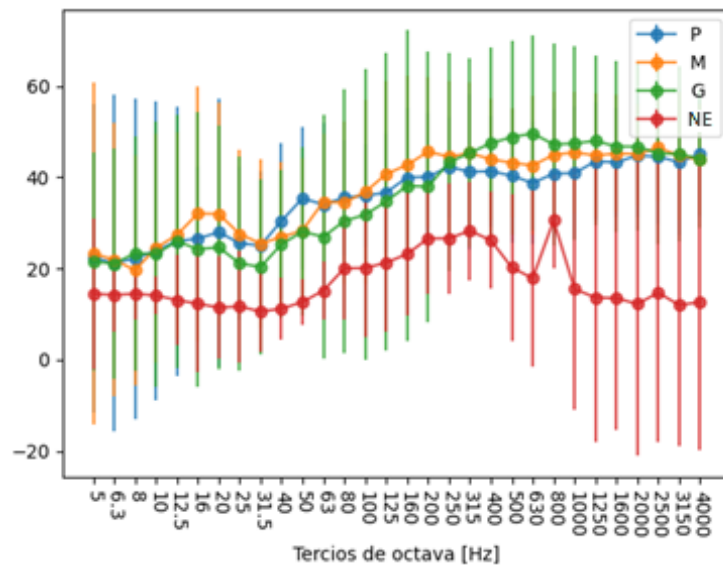


Figura 21. Representación de las firmas acústicas.

A partir de los resultados obtenidos, se observan algunas diferencias en las medianas de las distintas categorías. Sin embargo, es importante tener en cuenta que los valores se encuentran dentro de los rangos intercuartiles, lo que dificulta determinar de manera visual su potencialidad. No obstante, como se mencionó previamente en el análisis del espectrograma de Mel, los algoritmos de inteligencia artificial tienen la capacidad de descubrir características complejas, lo que sugiere que este parámetro podría ser útil a pesar de no se apreciarse una diferenciación clara a simple vista.

Además, al comparar las categorías de NE con la de E, en estas últimas se observan diferencias en las bandas de bajas frecuencias. Esto puede indicar que la firma acústica podría ser de utilidad para distinguir entre dichas categorías.

### 5.1.5. VECTOR DE ÍNDICES DE TONALIDAD

Como se menciona en la sección de descriptores acústicos, el vector de índices de tonalidad (tonVect) se obtiene mediante la estimación de un índice de tonalidad por subrangos de frecuencia, con mayor resolución frecuencial para bajas frecuencias que para altas frecuencias. En la Figura 22 se muestran los resultados obtenidos del cálculo de tonVect.

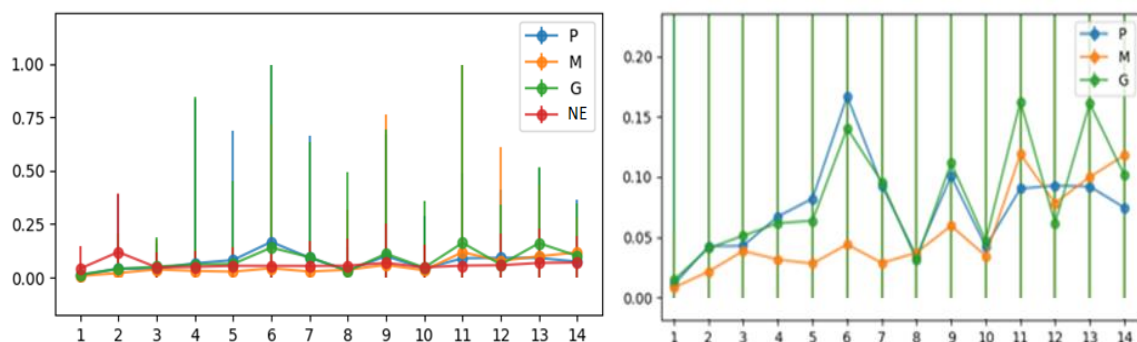


Figura 22. Representación los índices de tonalidad.

En los gráficos presentados, puede apreciarse que los valores de índices de tonalidad para las diferentes categorías de embarcaciones son superiores en comparación con la categoría de no embarcaciones. Esto sugiere que el vector *tonVect* puede ser un parámetro útil para distinguir entre E y NE. Al analizar la dispersión de los posibles valores, los índices de tonalidad de NE en todos los subrangos se concentran por debajo de 0,5. Además, si se considera desde el tercero en adelante, los resultados son inferiores a 0,25.

Adicionalmente, se puede observar una tendencia similar en las embarcaciones de distinto tamaño, con picos en los subrangos 6, 9, 11 y 13 de índices de tonalidad. El hecho de que presenten valores en las mismas bandas de frecuencia sugiere que las diferentes embarcaciones producen señales acústicas tonales superpuestas en esas frecuencias. Este hallazgo es relevante, ya que demuestra que la información de las frecuencias tonales generadas no es la principal opción como parámetro de clasificación para distinguir entre diferentes tipos de embarcaciones, al menos con las bases de datos disponibles.

Por otro lado, es relevante destacar la amplia variabilidad en la dispersión de los datos y la proximidad de los valores de las medianas a cero. Esto indica que muchas de las muestras exhiben resultados cercanos a cero. No obstante, se observan subrangos donde los índices de tonalidad aumentan, esto sugiere una tendencia hacia valores altos en dichos subrangos.

Dado este comportamiento, es recomendable revisar el algoritmo para obtener una mayor cantidad de muestras con índices de tonalidad altos. No obstante, es importante tener precaución para evitar que los resultados se concentren en índices de tonalidad máximos. Esto implica un compromiso entre la sensibilidad y la especificidad del algoritmo.

En el contexto de la presente tesis, se decidió avanzar con el desarrollo de los algoritmos de detección y clasificación, con la recomendación a futuro de realizar ajustes y mejoras en algoritmo que calcula los índices de tonalidad. Es fundamental tener en cuenta que cada ajuste de valor debe ser evaluado en un conjunto amplio de muestras, analizando detalladamente la distribución de los resultados. El objetivo es lograr un algoritmo más efectivo y preciso en la clasificación de las embarcaciones.

En base a lo expuesto, se puede concluir que el vector de índices de tonalidad tiene potencial como datos de entrada en un algoritmo de clasificación basado en inteligencia artificial. No obstante, es importante tener en cuenta que existe una superposición de valores de índices de tonalidad entre las diferentes categorías de embarcaciones. Esto implica que los resultados obtenidos mediante el cálculo de este parámetro no logran una separación clara de las categorías de embarcaciones.

A pesar de las limitaciones mencionadas, es posible que este aporte información relevante y contribuya en combinación con otros descriptores acústicos para mejorar la precisión del sistema de clasificación.

## **5.2. DETECTOR AUTOMÁTICO DE EMBARCACIONES**

En esta sección, se realiza una evaluación exhaustiva de diferentes hiperparámetros de los modelos utilizando la técnica de validación cruzada mediante GridSearchCV. Se evalúan los algoritmos de KNN, SVM, Random Forest y Decision Tree para la detección de embarcaciones. El objetivo es determinar el modelo y los hiperparámetros que maximizan el Valor F1 en el conjunto de evaluación. Además, se presentan los resultados de los diferentes modelos sobre archivos independientes de su conjunto de entrenamiento, lo que permite evaluar la capacidad de generalización y el rendimiento de los algoritmos en situaciones reales. La Tabla 5 y 6 muestran los resultados obtenidos para los distintos hiperparámetros evaluados en esta investigación.

Tabla 5. Hiperparámetros que maximizan la performance para cada modelo

Modelo	Hiperparámetros
KNN	'algorithm': 'auto', 'n_neighbors': 9, 'weights': 'distance'
SVM	'C': 0.1, 'degree': 2, 'gamma': 1, 'kernel': 'poly'
RF	'criterion': 'entropy', 'max_depth': 15, 'max_features': 'log2', 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 200
DT	'criterion': 'entropy', 'max_depth': 7, 'max_features': 'sqrt', 'min_samples_leaf': 1, 'min_samples_split': 2, 'splitter': 'best'

Tabla 6. Resultados de las métricas de desempeño de los distintos modelos.

Modelo	Valor F1	Exhaustividad	Precisión	Exactitud
KNN	0,97	1,0	0,94	0,97
SVM	0,99	1,0	0,98	0,99
RF	0,97	1,0	0,94	0,97
DT	0,96	0,98	0,94	0,96

En general, todos los modelos evaluados han obtenido resultados destacables en la detección de embarcaciones y no embarcaciones. El modelo SVM alcanzó el puntaje más alto de 0,99 en Valor F1 en el conjunto de datos de evaluación, seguido por KNN y RF con puntajes de 0,97. El modelo DT obtuvo un puntaje ligeramente inferior pero aun considerablemente alto, con un puntaje de 0,96

Es interesante observar que todos los modelos han logrado un puntaje perfecto en el conjunto de entrenamiento, lo cual sugiere un posible sobreajuste al conjunto de

entrenamiento utilizado debido a la optimización de los hiperparámetros. Sin embargo, los puntajes en el conjunto de evaluación son también altos, lo que indica capacidad de generalización de los modelos.

Vale la pena mencionar que se esperaban resultados considerablemente altos ya que, como se explica en la sección anterior, existe una clara separación en los resultados obtenidos para estas categorías en algunos parámetros, como, por ejemplo, el índice de tonalidad. Esta observación es consistente con la capacidad de los modelos para discernir entre embarcaciones y no embarcaciones.

Por otro lado, al analizar los resultados de la tabla, se puede observar que Precisión, Exhaustividad y Exactitud para la detección de embarcaciones son altos en todos los modelos evaluados, lo que indica que los modelos son capaces de identificar la mayoría de las embarcaciones presentes en el conjunto de prueba mientras minimizan los falsos positivos.

Es importante tener en cuenta que estos resultados deben interpretarse considerando ciertos factores. Aunque la validación cruzada se ha utilizado para mitigar el sobreajuste y se ha obtenido una buena precisión en el conjunto de prueba, el tamaño del conjunto de datos es relativamente pequeño, lo que podría limitar la capacidad de generalización de los modelos a nuevos datos. Por lo tanto, se recomienda realizar evaluaciones en conjuntos de datos más grandes y diversos.

### **5.2.1. EVALUACIÓN DE LOS MODELOS**

Como se describe en la sección de diseño de la investigación, se realiza una evaluación del sistema desarrollado utilizando grabaciones proporcionadas por la fundación Cethus. Estas grabaciones contienen eventos acústicos producidos por embarcaciones, los cuales fueron identificados mediante la inspección visual de sus espectrogramas. El objetivo principal de esta experiencia es la de analizar la capacidad de generalización de los modelos desarrollados. A modo ilustrativo, se presenta en la Figura 23 un gráfico de esta evaluación, el cual aporta a la comprensión del análisis presente en esta sección. Cabe mencionar que la totalidad de los gráficos de cada modelo se encuentran disponibles en el Anexo I.

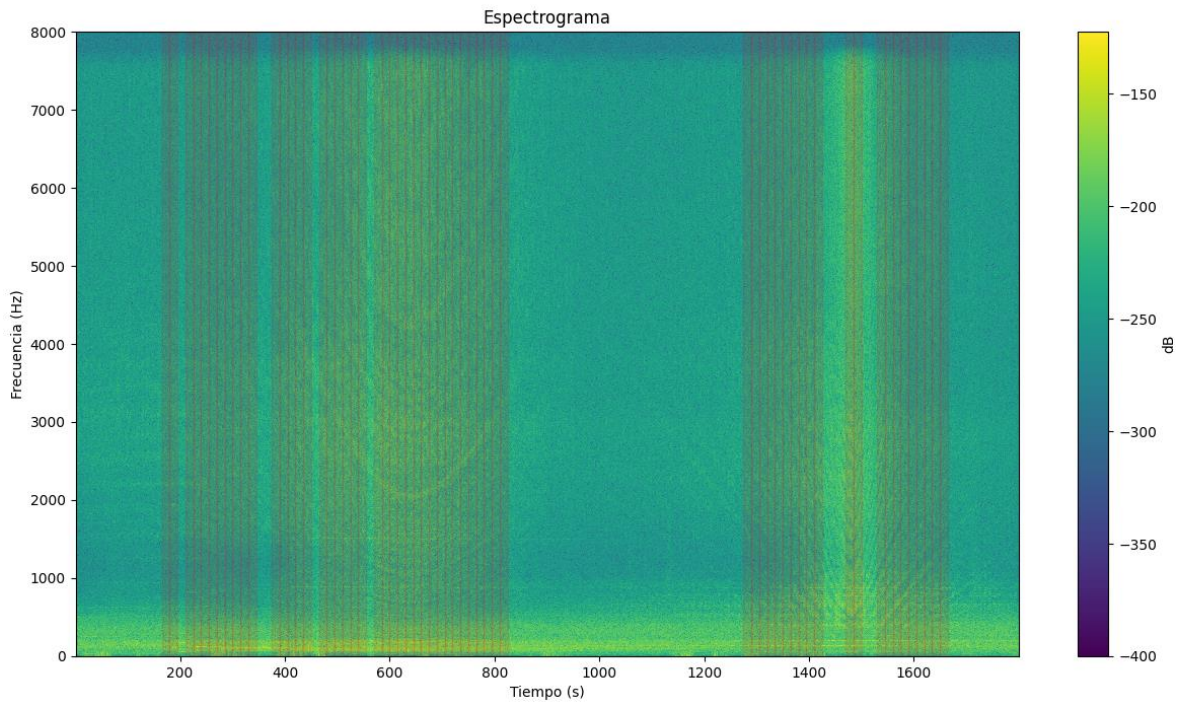


Figura 23. Detecciones sobre espectrograma de larga duración obtenidas con DT.

Al analizar los gráficos del Anexo I, se observa que, en términos generales, todos los algoritmos demuestran una buena capacidad para detectar embarcaciones, descartando efectivamente las áreas donde no se observan señales correspondientes. No se aprecian diferencias significativas en el rendimiento de los diferentes modelos, ya que prácticamente detectan embarcaciones en los mismos fragmentos.

Sin embargo, se identifican algunos casos en los que los modelos presentan dificultades en la detección de embarcaciones. Como se puede observarse en la figura 23, se encuentran fragmentos que deberían ser considerados como embarcaciones, pero no son detectados. Esta limitación puede atribuirse al enfoque principal del desarrollo, que se centra en las características tonales de las embarcaciones. En momentos en los que la energía captada carece de suficiente contenido tonal, es posible que no se identifiquen correctamente como eventos de embarcaciones. Para superar esta limitación, se propone como trabajo futuro implementar una lógica de detección que considere la densidad de fragmentos clasificados como embarcaciones, de modo que se reconozca todo el bloque de tiempo como un evento acústico producido por embarcaciones. Además, esto podría fortalecer el sistema al evitar detecciones erróneas de embarcaciones en momentos en los que no están presentes.

A pesar de los resultados obtenidos en este estudio, se reconoce que el rendimiento de los modelos podría mejorar significativamente si se dispusiera de un conjunto de datos más amplio y diverso. Sería beneficioso contar con una mayor cantidad de muestras de embarcaciones que reflejen la variabilidad en los sistemas de grabación, los paisajes acústicos y las diferencias inherentes a las embarcaciones. Esto permitiría una mejor adaptación del modelo a diferentes escenarios y aumentaría su capacidad de generalización.

A pesar de las limitaciones antes mencionadas, los resultados obtenidos proporcionan una base sólida para el desarrollo de sistemas de detección de embarcaciones, y se espera que sirva como punto de partida para futuras investigaciones y desarrollos en el campo. Por otro lado, los cuatro modelos evaluados han mostrado un rendimiento destacado en la detección de embarcaciones. No obstante, considerando la simplicidad y eficacia del algoritmo de Decision Tree, se recomienda su implementación como modelo de detección de embarcaciones.

### **5.3. CLASIFICADOR AUTOMÁTICO DE EMBARCACIONES**

En esta sección, se realiza una evaluación exhaustiva de diferentes hiperparámetros de los modelos utilizando la técnica de validación cruzada mediante GridSearchCV. Se evalúan los algoritmos de KNN, SVM, Random Forest y Decision Tree para la clasificación de embarcaciones. El objetivo es determinar el modelo y los hiperparámetros que maximizan el Valor F1 en el conjunto de evaluación. En la Tabla 7 se presentan los resultados hiperparámetros seleccionados en esta investigación. Con el objetivo de introducir los hiperparámetros representados en la Tabla 7, se proporciona una descripción de los mismos en la Tabla 8.



Tabla 7. Hiperparámetros que maximizan la performance para cada modelo

Modelo	Hiperparámetros
KNN	'algorithm': 'auto', 'n_neighbors': 5, 'weights': 'distance'
SVM	'C': 10, 'degree': 2, 'gamma': 0.1, 'kernel': 'rbf'
RF	'criterion': 'entropy', 'max_depth': 10, 'max_features': 'sqrt', 'min_samples_leaf': 2, 'min_samples_split': 2, 'n_estimators': 100
DT	'criterion': 'entropy', 'max_depth': 10, 'max_features': 'sqrt', 'min_samples_leaf': 2, 'min_samples_split': 5, 'splitter': best

Tabla 8. Hiperparámetros que maximizan la performance para cada modelo

Modelo	Hiperparámetro	Descripción
<b>KNN</b>	algorithm	Algoritmo utilizado para calcular los vecinos más cercanos.
	n_neighbors	Número de vecinos más cercanos considerados.
	weights	Peso aplicado a los vecinos en función de su distancia.
<b>SVM</b>	C	Parámetro de regularización.
	degree	Grado del kernel polinomial.
	gamma	Coeficiente del kernel.
	kernel	Tipo de kernel utilizado.
<b>DT - RF</b>	criterion	Criterio utilizado para medir la calidad de una división.
	max_depth	Profundidad máxima del árbol.
	max_features	Número de características a considerar en cada división.
	min_samples_leaf	Número mínimo de muestras requeridas en un nodo hoja.
	min_samples_split	Número mínimo de muestras requeridas para dividir un nodo.
<b>RF</b>	n_estimators	Número de árboles en el bosque.
<b>DT</b>	splitter	Estrategia utilizada para seleccionar la mejor división.

En las Tablas 9-12 se presentan las métricas de desempeño correspondientes a las distintas categorías, divididas por modelo de aprendizaje automático. En términos generales, mediante el uso del modelo SVM se obtienen los mejores resultados en todas las categorías. Además, se observa consistentemente en todos los modelos una clara habilidad para clasificar correctamente las muestras correspondientes a la categoría NE.

Es importante señalar que se decidió incluir la categoría NE en esta experimentación por dos motivos fundamentales. En primer lugar, en caso de que la etapa de detección presente falsos positivos, la etapa de clasificación posterior podría descartarlos al asignarlos a la categoría NE. En segundo lugar, para futuros trabajos y considerando posibles mejoras, podría ser recomendable utilizar una única etapa de clasificación que englobe todas las muestras correspondientes a la categoría NE, en lugar de utilizar dos etapas separadas para detección y clasificación. Es fundamental considerar que los conjuntos de datos utilizados en la experimentación son limitados, lo que afecta directamente sobre la capacidad de clasificación de los distintos modelos.

Tabla 9. Métricas de desempeño por categoría - Modelo KNN.

<b>Categoría</b>	<b>Valor F1</b>	<b>Exhaustividad</b>	<b>Precisión</b>
Grandes	0,80	0,78	0,82
Medianas	0,79	0,85	0,73
Pequeñas	0,76	0,76	0,76
No Embarcaciones	0,97	0,94	1,00

El modelo KNN muestra un rendimiento aceptable en todas las categorías, con valores F1 competitivos y un equilibrio razonable entre precisión y exhaustividad. Si bien los resultados obtenidos son significativos, es importante tener en cuenta que no alcanzan niveles óptimos para una clasificación precisa en todas las situaciones. Además, el

algoritmo KNN puede presentar limitaciones en escenarios con cierto grado de complejidad y variabilidad.

Tabla 10. Métricas de desempeño por categoría - Modelo SVM.

<b>Categoría</b>	<b>Valor F1</b>	<b>Exhaustividad</b>	<b>Precisión</b>
Grandes	0,88	0,83	0,94
Medianas	0,93	1,00	0,87
Pequeñas	0,94	0,94	0,94
No Embarcaciones	1,00	1,00	1,00

Los resultados del modelo SVM son notables, demostrando su capacidad para detectar correctamente todas las categorías. Esto indica que el modelo SVM es efectivo en la clasificación de embarcaciones

Resulta interesante destacar que se observa una disminución en el rendimiento del algoritmo en la categoría de GE en términos de exhaustividad. Esto sugiere que el modelo puede estar perdiendo la detección de algunas muestras de GE. No obstante, todas las métricas de evaluación superan el umbral de 0,80, lo cual se considera satisfactorio, especialmente teniendo en cuenta las limitaciones propias de utilizar un conjunto de datos de pocas muestras.

Tabla 11. Métricas de desempeño por categoría - Modelo RF.

<b>Categoría</b>	<b>Valor F1</b>	<b>Exhaustividad</b>	<b>Precisión</b>
Grandes	0,77	0,67	0,92
Medianas	0,61	0,85	0,61

Pequeñas	0,88	0,88	0,88
No Embarcaciones	1,00	1,00	1,00

El modelo RF presenta un buen rendimiento en la detección de NE y PE. Sin embargo, se observa una disminución en el rendimiento en las categorías de ME y GE lo que sugiere que tiene dificultades para distinguir correctamente las embarcaciones de mayor tamaño.

Tabla 12. Métricas de desempeño por categoría - Modelo DT.

<b>Categoría</b>	<b>Valor F1</b>	<b>Exhaustividad</b>	<b>Precisión</b>
Grandes	0,61	0,56	0,67
Medianas	0,62	0,69	0,56
Pequeñas	0,58	0,65	0,52
No Embarcaciones	0,86	0,75	1,00

El modelo DT muestra resultados inferiores en todas las categorías en comparación con los otros modelos. Esto sugiere que el modelo puede tener dificultades para lograr clasificaciones precisas de manera general. Es importante mencionar que este comportamiento era esperado, ya que los modelos de árbol de decisión suelen tener limitaciones cuando se enfrentan a problemas de clasificación complejos y cuando dispone de un número reducido de muestras para el entrenamiento.

Además, al comparar los valores de F para el conjunto de datos de entrenamiento (F1-E) y de prueba (F1-P) en los distintos modelos, se observa únicamente un notable deterioro en el caso del modelo DT (0,93 de F1-E y 0,62 F1-P). Esto es un indicador de sobreajuste del modelo al conjunto de entrenamiento, lo cual resulta en una pérdida de

capacidad de generalización. Dicho resultado es coherente con los bajos valores de las métricas de desempeño presentadas en la Tabla 12, reforzando la idea de que el modelo DT no es adecuado para esta clasificación en particular.

Por último, es importante tener en cuenta que los resultados obtenidos en esta etapa son válidos únicamente dentro de este contexto específico y no se pueden extrapolar directamente a otras situaciones o conjuntos de datos diferentes. Es necesario destacar que la cantidad limitada de muestras utilizadas en esta experiencia puede introducir un sesgo en los algoritmos debido a la dependencia en el conjunto de datos específico. En otras palabras, es posible que las muestras disponibles tengan características o limitaciones que no representen completamente la variabilidad y diversidad de las embarcaciones presentes en distintos sistemas de registro.

Además, es menester mencionar que no se dispone de registros acústicos independientes de las bases de datos utilizadas para el entrenamiento y evaluación. Por lo tanto, queda pendiente para trabajos futuros la evaluación de la capacidad de generalización de los modelos y la consideración del posible sesgo del conjunto de datos. Para llevar a cabo estas validaciones, sería necesario contar con un conjunto de datos más representativo y diverso, lo que permitiría realizar un análisis más exhaustivo, similar al presentado en la etapa de detección.

## 6. CONCLUSIONES

En el presente trabajo se describe la metodología para el desarrollo de algoritmos de detección y clasificación de embarcaciones basándose en sus características acústicas. Esta experiencia tiene como objetivo proporcionar información para desarrollar herramientas que permitan ahorrar tiempo y trabajo a los usuarios dedicados a la monitorización ambiental, gestión del tráfico marítimo y seguridad marítima. Además, sienta las bases para futuras investigaciones en la mejora de la de los sistemas de detección y clasificación acústica de embarcaciones aquí propuestos.

Las conclusiones de la presente tesis se pueden dividir en tres partes. En primer lugar, las que surgen del estudio del comportamiento acústico de las embarcaciones y la creación de la base de datos. En segundo lugar, las conclusiones de la obtención y análisis de parámetros. Por último, las conclusiones derivadas de la experimentación con diversos modelos de aprendizaje automático para la detección y clasificación de embarcaciones.

La primera experiencia revela la complejidad del análisis de eventos acústicos, con dificultades en la identificación de características que permitan separar las señales. A partir de la inspección visual de los espectrogramas se observa que los diferentes tipos de embarcaciones generan eventos superpuestos en frecuencia. Se destaca la información que provee el comportamiento tonal del fenómeno en cuestión. En cuanto a la conformación de la base de datos, se destaca la escasez de muestras de acceso público, lo que dificulta la construcción de conjuntos de datos amplios y variados para poder representar correctamente el fenómeno acústico en cuestión.

De la segunda experiencia, se concluye que los parámetros acústicos evaluados presentan resultados similares entre categorías de embarcaciones. En particular, el índice de tonalidad muestra resultados satisfactorios en cuanto a la separación entre E y NE (con valores entre 0,5 a 1 para E y entre 0 a 0,3 para NE), pero no es de utilidad para clasificar los diferentes tipos de embarcaciones. Del estudio del vector de índices de tonalidad se revela que las señales acústicas tonales producidas por las diferentes categorías de E se superponen en frecuencias, lo cual es consistente con lo que inferido en la primera etapa.

En la tercera experiencia, en términos de algoritmos de detección, se observa un desempeño destacable en los diferentes modelos, con valores F1 superiores a 0,95. Si bien

el modelo SVM es el que presenta el mejor desempeño, se sugiere la elección del algoritmo de “Decision Tree” como modelo de detección debido a su simplicidad. En cuanto a la clasificación, se encuentra que mediante SVM se obtienen los mejores resultados en todas las categorías, con valores F1 superiores a 0,8, seguido por el modelo KNN con un rendimiento aceptable, con valores F1 superiores a 0,75. Por otro lado, DT tiene resultados inferiores y muestra indicios de sobreajuste al conjunto de datos de entrenamiento. Es importante destacar que los conjuntos de datos utilizados fueron limitados, lo que afecta la capacidad de generalización de los modelos.

En términos generales se considera que, si bien los sistemas desarrollados pueden ser ampliamente mejorados a futuro, esta experiencia proporciona información valiosa y una metodología sólida para el desarrollo de detectores y clasificadores automáticos de embarcaciones. Además, se considera que se obtuvieron resultados satisfactorios y que el objetivo de la presente tesis fue alcanzado con éxito. En la siguiente sección se comentan posibles trabajos a futuro.

## 7. TRABAJOS A FUTURO

- Evaluar las posibles diferencias en parámetros acústicos al dividir las muestras por bases de datos. Este análisis permitiría cuantificar el grado de influencia en los resultados debido al entorno acústico y a las características del sistema de grabación, entre otros aspectos.

- Mejorar el vector de índices de tonalidad con el objetivo de reducir la dispersión de resultados y maximizar la discriminación entre E y NE.

- Investigar y desarrollar otros descriptores personalizados que ayuden en la clasificación de las embarcaciones. La exploración de nuevos descriptores podría capturar características que mejoren la capacidad de clasificación.

- Evaluar la influencia de cada parámetro acústico en el rendimiento de los modelos. Este análisis permitiría determinar los parámetros acústicos más relevantes para la detección y clasificación, optimizando así la eficiencia del sistema.

- Desarrollar un algoritmo que calcule la densidad de detecciones en el tiempo para detectar correctamente las regiones temporales en las que se encuentran las embarcaciones. Esto ayudaría a evitar falsos negativos (detecciones de NE en regiones de E) y falsos positivos (detecciones de E espurias).

- Ampliar el conjunto de datos para robustecer los algoritmos, especialmente en el caso del clasificador. Obtener un dataset amplio y diverso permitiría implementar modelos basados en redes neuronales, lo cual se recomienda ampliamente para problemáticas complejas como la clasificación de las embarcaciones.

- Con el fin de evaluar la capacidad de generalización del algoritmo de clasificación, sería conveniente obtener grabaciones etiquetadas de larga duración e independientes del conjunto de datos de entrenamiento.

- Implementar un sistema similar al desarrollado en esta tesis, pero con la capacidad de detectar y clasificar embarcaciones directamente en cualquier archivo de audio utilizado como entrada. Esto proporcionará una herramienta práctica y funcional para la detección y clasificación automatizada de embarcaciones.



## BIBLIOGRAFÍA

- [1] Evans, D.L., England, G.R.E. "Joint interim report Bahamas Marine Mammal Stranding Event of 15-16 March 2000". Tech. Rep., US Department of Commerce and Secretary of Navy. 59 pp. 747-756. December 2001.
- [2] Lucke, K., Siebert, U., Lepper, P.A., Blanchet, M.A. "Temporary shift in masked hearing thresholds in a harbor porpoise (*Phocoena phocoena*) after exposure to seismic airgun stimuli". *J. Acoust. Soc. Am.* 125, pp. 4060-4070. Jun 2009.
- [3] Bailey, H., Senior, B., Simmons, D., Rusin, J., Picken, G., Thompson, P.M. "Assessing underwater noise levels during pile-driving at an offshore windfarm and its potential effects on marine mammals". *Marine Pollution Bulletin* 60, pp. 888-897. 2010.
- [4] Richardson, W.J., Würsig, B. "Influences of man-made noise and other human actions on cetacean behaviour". *Marine and Freshwater Behaviour and Physiology* 29, pp. 183-209. 1997.
- [5] Popper, A.N., Hastings, M.C. "The effects of anthropogenic sources of sound on fishes". *Journal of Fish Biology* 75, pp. 455-489. 2009.
- [6] André, M., Solé, M., Lenoir, M., Durfort, M., Quero, C., Mas, A., Lombarte, A., van der Schaar, M., López-Bejar, M., Morell, M., Zaugg, S., Houégnigan, L. "Low-frequency sounds induce acoustic trauma in cephalopods". *Frontiers in Ecology and the Environment* 9, pp. 489-493. 2011.
- [7] Clark, C.W., Ellison, W.T., Southall, B.L. Hatch, L., Van Parijs, S.M. Frankel, A., Ponirakis, D. "Acoustic masking in marine ecosystem: intuitions analysis, and implication". *Marine Ecology Progress Series* 395. pp. 201-222. 2009.
- [8] Wright, A.J., Soto, N.A., Baldwin, A.L., Bateson, M., Beale, C.M., Clark, C., Deak, T., Edwards, E.F., Fernandez, A., Godinho, A., Hatch, L.T., Kakuschke, A., Lusseau, D., Martineu, D., Romero, M.L., Weilgart, L.S., Wintle. B.A., Notarbartolo-di Sciara, G., Martin, V. "Do marine mammals experience stress related to anthropogenic noise?". *International Journal of Comparative Psychology* 20, pp. 274-316. 2007.
- [9] Picciulin, M., Sebastianutto, L., Codarin, A., Farina, A., Ferrero, E.A. "In situ behavioural responses to boat noise exposure of *Gobius cruentatus* and *Chromis chromis* living in a

Marine Protected Area". *Journal of Experimental Marine Biology and Ecology* 386. pp. 125-132. 2008.

[10] Chan, A., Giraldo-Perez, P., Smith, S., Blumstein, D.T. "Anthropogenic noise affects risk assessment and attention: the distracted prey hypothesis". *Biology Letters* 6, pp. 458-461. 2010.

[11] Fundación Cethus, Lineas de investigación, Extraído el 19 de Agosto de 2021, [www.cethus.org](http://www.cethus.org)

[12] Shiu, Y., Palmer, K. J., Roch, M. A., Fleishman, E., Liu, X., Nosal, E. M., Helbe T., Cholewiak, D., Gillispie, D., Klinck, H. "Deep neural networks for automated detection of marine mammal species". *Nature Research*. 2020.

[13] Erbe, C. "Underwater passive acoustic monitoring & noise impacts on marine fauna". *Australian Acoustical Society*. Enero 2013.

[14] Gardiner, J., Morrin, E., Todd, I., Todd, E., *Marine Mammal Observer & Passive Acoustic Monitoring Handbook*, 1er edición, Pelagic Publishin, Reino Unido. 2015.

[15] Abrahamsen, K., "The ship as an underwater noise source". *Acoustical Society of America*. 2012.

[16] McKenna, M.F., Ross, D., Wiggins, S.M., Hildebrand, J.A. "Underwater radiated noise from modern commercial ships". *Journal of the Acoustical Society of America* 131, pp. 92–103. 2012.

[17] Wales, S., Heitmeyer, R. "An ensemble source spectra model for merchant ship-radiated noise," *Journal of the Acoustical Society of América* 111, pp. 1211–1231. 2002.

[18] Wilmut, M. J., Chapman, N. R., Heard, G. J., Ebbeson, G. R. "Inversion of Lloyd Mirror Field for determining a source's track," *IEEE J. Ocean. Eng.* 32, pp. 940–947. 2007.

[19] Jensen, F.H., Bejder, L., Wahlberg, M., Soto, N.A., Johnson, M., Madsen, P.T. "Vessel noise effects on delphinid communication". *Marine Ecology Progress Series* 395, pp. 161–175. 2009-b.

[20] McFee, Brian, Colin Raffel, Dawen Liang, Daniel PW Ellis, Matt McVicar, Eric Battenberg, Oriol Nieto. "librosa: Audio and music signal analysis in python." In *Proceedings of the 14th python in science conference*, pp. 18-25. 2015..

- [21] Yildirim, M. E. "Ship Type Recognition using Deep Learning with FFT Spectrums of Audio Signals." *El-Cezeri* 10, no. 1, pp. 57-65. 2023
- [22] Gimse, Håkon. "Classification of Marine Vessels using Sonar Data and a Neural Network". Master's thesis, Norwegian University of Science and Technology. 2019.
- [23] IEC-International Electrotechnical Commision, Electroacoustics - Octave-band and fractional-octave-band filters. Part 3: Periodic tests, 2016.
- [24] Blackwell, S. B., Natopms, C. S., McDonald T. L., Thode, A. M., Mathias, D., Kim, K. H., Greene C. R. Jr, Macrander M. A., "Effects of airgun sounds on bowhead whale calling rates: evidence for two behavioral thresholds". *Journal Plos One* 10. 2015.
- [25] Hildebrand, J. A., Baumann-Pickering, S., Frasier, K. E., Trickey, J. S., Merkens, K. P., Wiggins, S. M., McDonald M. A., Garrison, L. P., Harris, D., Marques, T. A., Thomas, L. "Passive acoustic monitoring of beaked whale densities in the Gulf of Mexico". *Sci. Rep* 5. 2015.
- [26] Ciresan, D. C., Meier, U., Masci, J., Schmidhuber, J. "A committee of neural networks for traffic sign classification". *Proc. Int. Jt. Conf. Neural Netw.USA*. 2011.
- [27] Wang, J. C., Wang, J. F., He, K. W., & Hsu, C. S. "Environmental sound classification using hybrid SVM/KNN classifier and MPEG-7 audio low-level descriptor." In *The 2006 IEEE international joint conference on neural network proceedings*, pp. 1731-1735, 2006.
- [28] Chen, C. H., Huang, W. T., Tan, T. H., Chang, C. C., & Chang, Y. J. "Using k-nearest neighbor classification to diagnose abnormal lung sounds." *Sensors*, vol. 15, no. 6, pp. 13132-13158, 2015.
- [29] Gupta, C. N., Palaniappan, R., Swaminathan, S., & Krishnan, S. M. "Neural network classification of homomorphic segmented heart sounds." *Applied soft computing*, vol. 7, no. 1, pp. 286-297, 2007.
- [30] Xia, X., Togneri, R., Sohel, F., & Huang, D. "Random forest classification based acoustic event detection." In *2017 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 163-168, 2017.
- [31] Romero, J., Luque, A., & Carrasco, A. "Animal Sound Classification using Sequential Classifiers." In *BIOSIGNALS*, pp. 242-247, 2017.

- [32] Pollara AS, Lignan G, Boulange L, Sutin A, Salloum H. "Specifics of demon acoustic signatures for large and small boats." *The Journal of the Acoustical Society of America* 141, pp. 3991-3991, 2017.
- [33] Slamnoiu G, Radu O, Rosca V, Pascu C, Damian R, Surdu G, Curca E, Radulescu A. "DEMON-type algorithms for determination of hydro-acoustic signatures of surface ships and of divers." *IOP Conference Series: Materials Science and Engineering* 145, 2016.
- [34] Chung KW, Sutin A, Sedunov A, Bruno M. "DEMON acoustic ship signature measurements in an urban harbor." *Advances in Acoustics and Vibration*, 2011.
- [35] Hanson D, Antoni J, Brown G, Emslie R. "Cyclostationarity for Passive Underwater Detection of Propeller Craft: A Development of DEMON Processing." *Proceedings of Acoustics*, 2008.
- [36] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, J.-M., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- [37] Ng., A., "Machine learning." Curso Online de la Universidad de Stanford, Julio 2016.
- [38] Marafioti, A., Identificación automática de instrumentos musicales acústicos a partir de señales monofónicas digitales, Tesis de Grado, Universidad Nacional de Tres de Febrero, 2016.
- [39] K. Lisa Yang Center for Conservation Bioacoustics. (2019). Raven Pro: Interactive Sound Analysis Software (Version 1.6.1) [Computer software]. Ithaca, NY: The Cornell Lab of Ornithology. Disponible en: <http://ravensoundsoftware.com/>.
- [40] Santos-Domínguez, D., Torres-Guijarro, S., Cardenal-López, A., Pena-Gimenez, A. "ShipsEar: An underwater vessel noise database". *Applied Acoustics*, 113, pp. 64-69, 2016.
- [41] Marino, A., "Detección de clicks para implementación de monitoreo acústico pasivo de delfines", Tesis de grado, Universidad Nacional de Tres de Febrero, 2017.

## ANEXO I. Evaluación de capacidad de generalización de detectores.

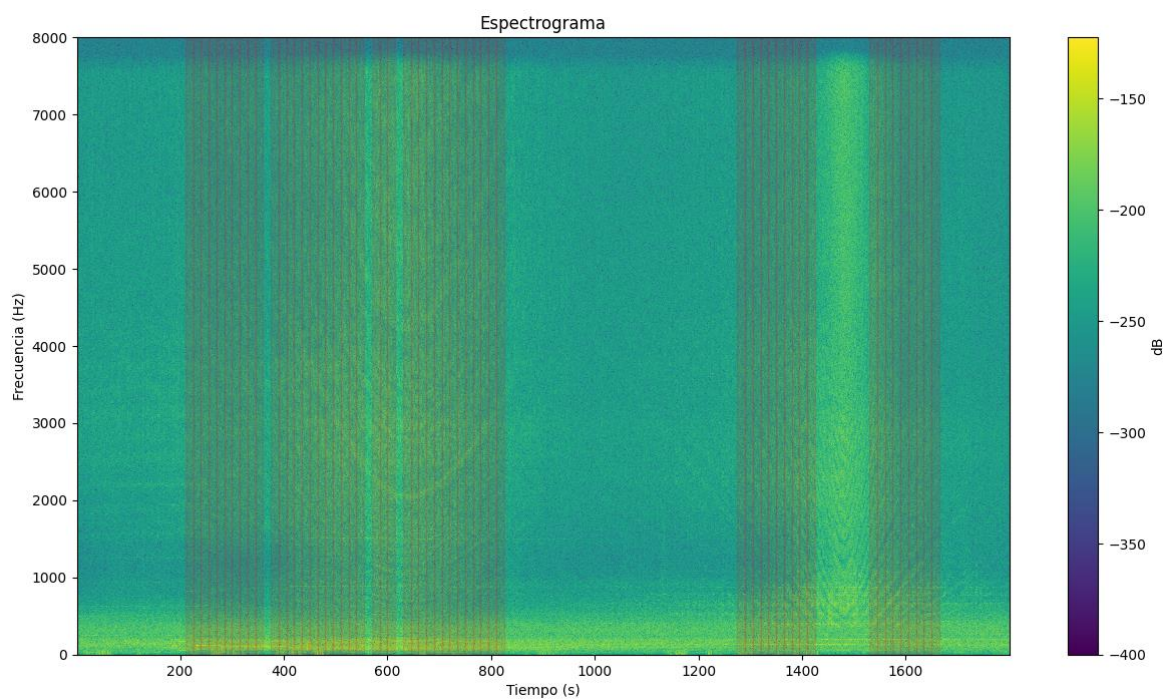


Figura A1. Detecciones sobre espectrograma de larga duración obtenidas con KNN.

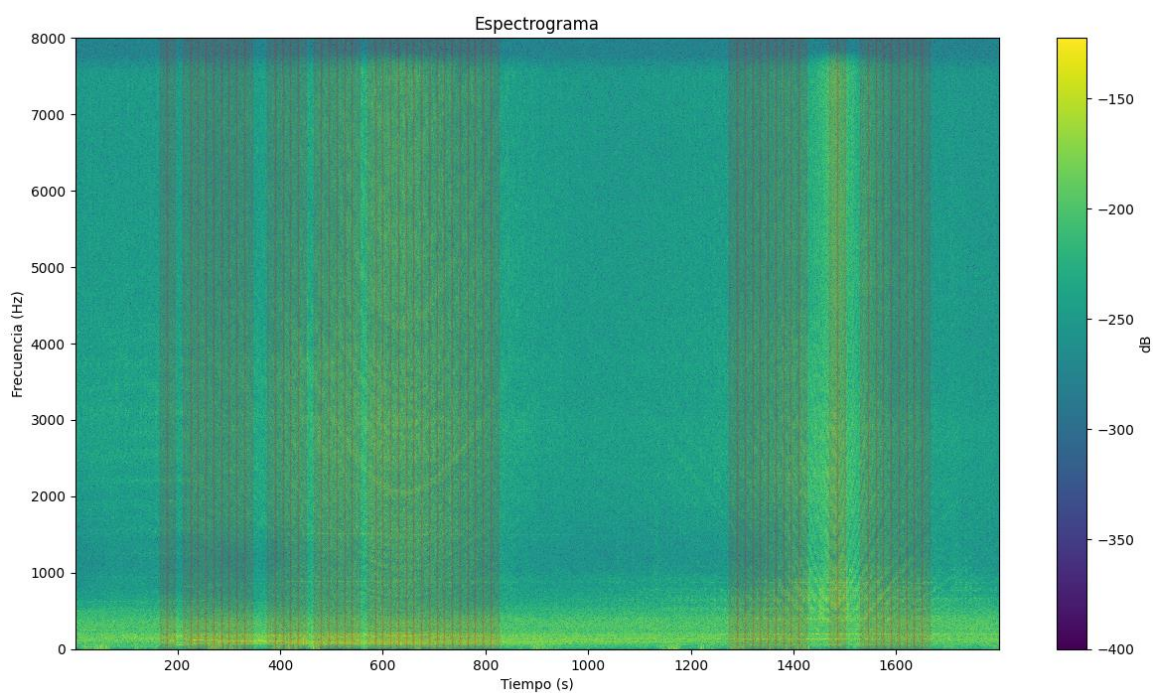


Figura A2. Detecciones sobre espectrograma de larga duración obtenidas con SVM.



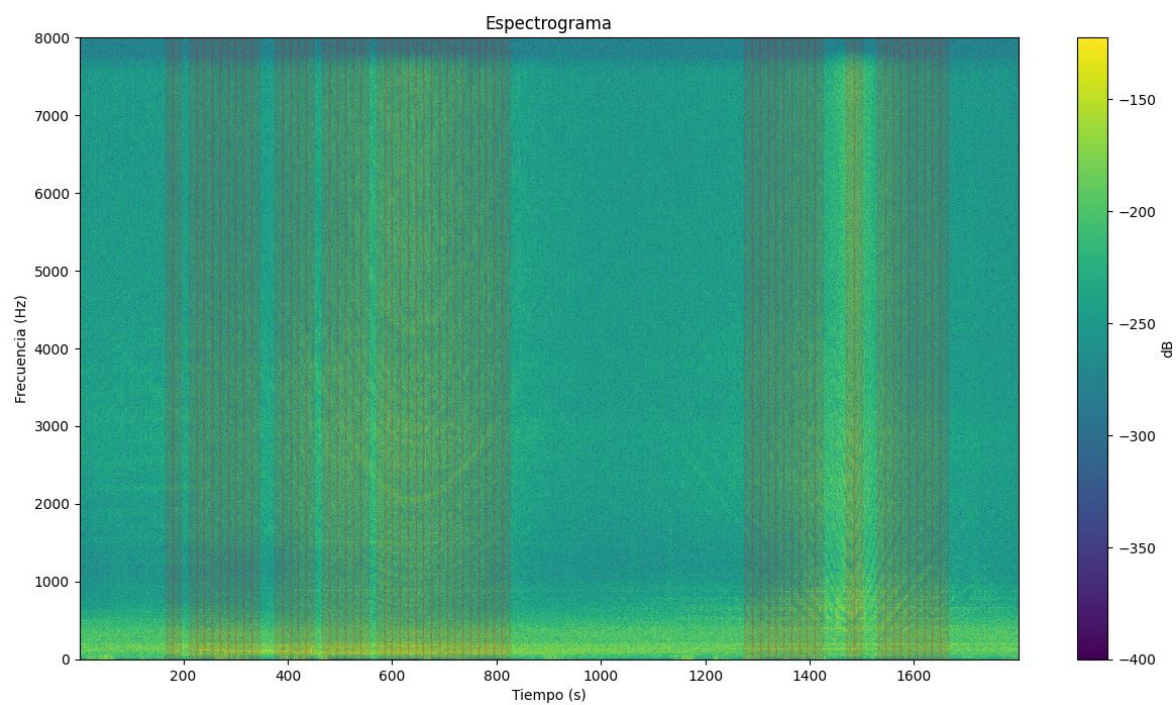


Figura A3. Detecciones sobre espectrograma de larga duración obtenidas con DT.

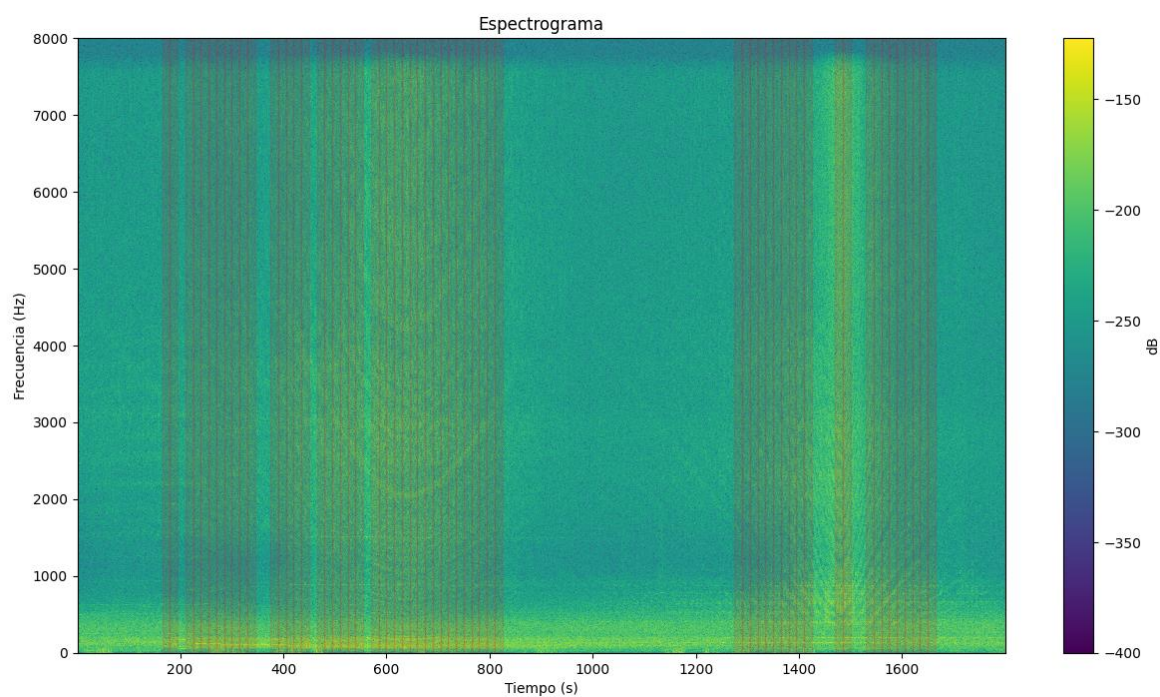


Figura A4. Detecciones sobre espectrograma de larga duración obtenidas con RF.

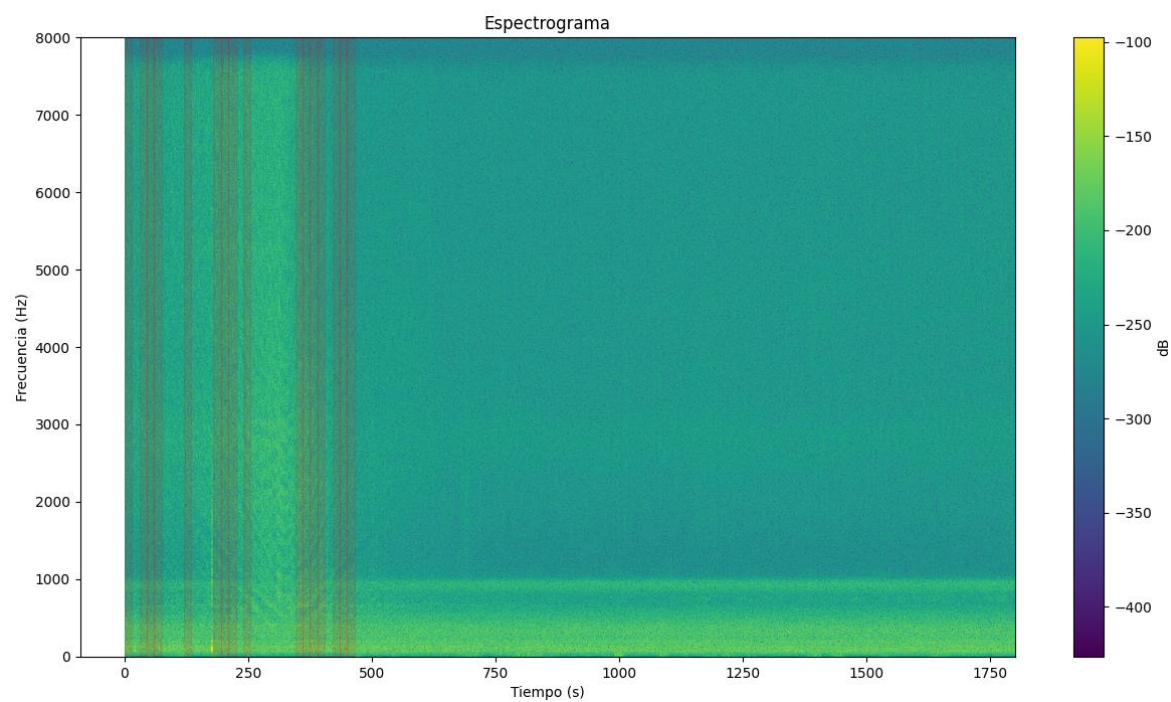


Figura A5. Detecciones sobre espectrograma de larga duración obtenidas con SVM.

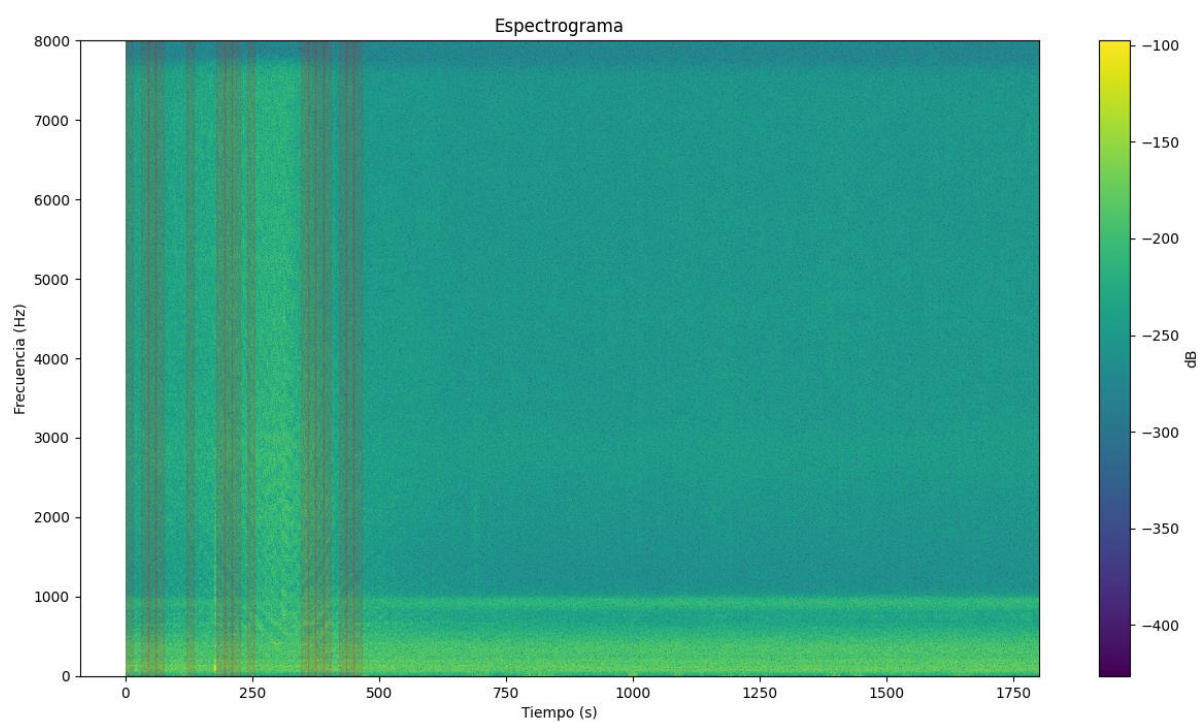


Figura A6. Detecciones sobre espectrograma de larga duración obtenidas con KNN.



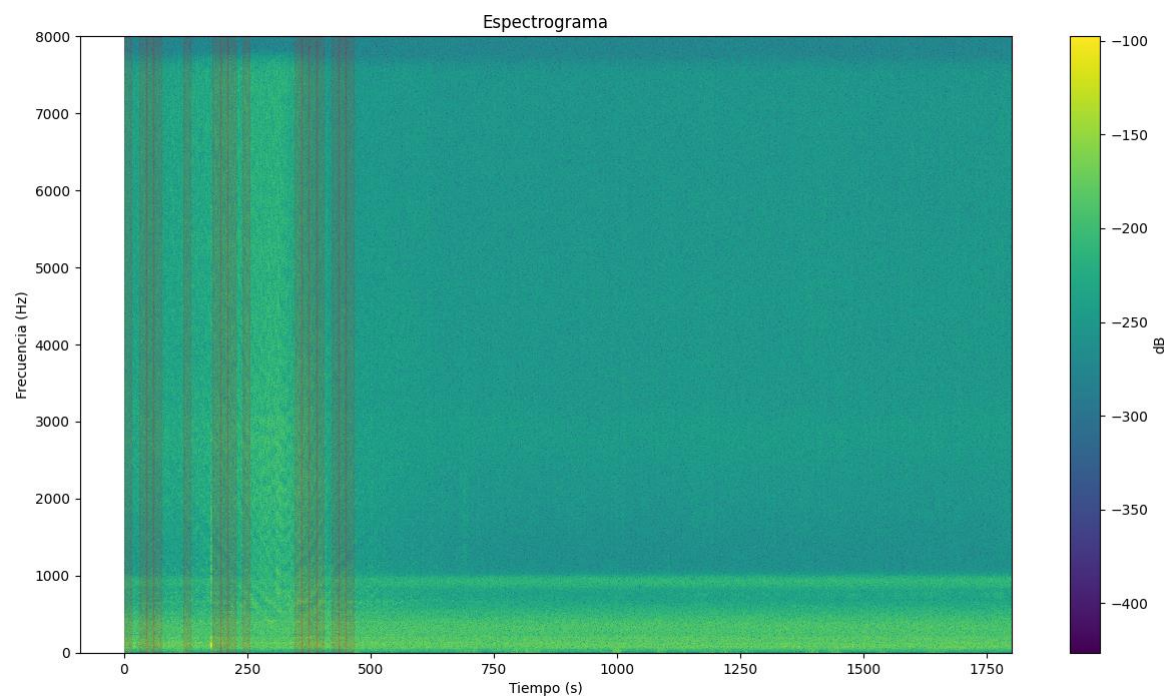


Figura A7. Detecciones sobre espectrograma de larga duración obtenidas con RF.

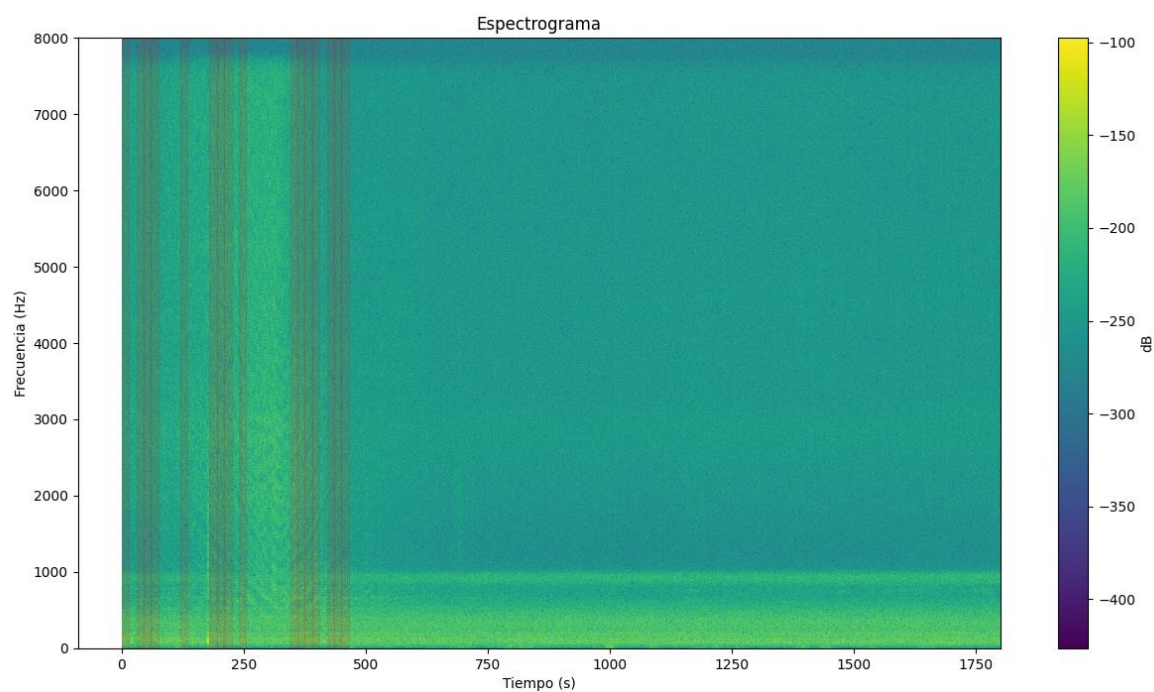


Figura A8. Detecciones sobre espectrograma de larga duración obtenidas con DT.



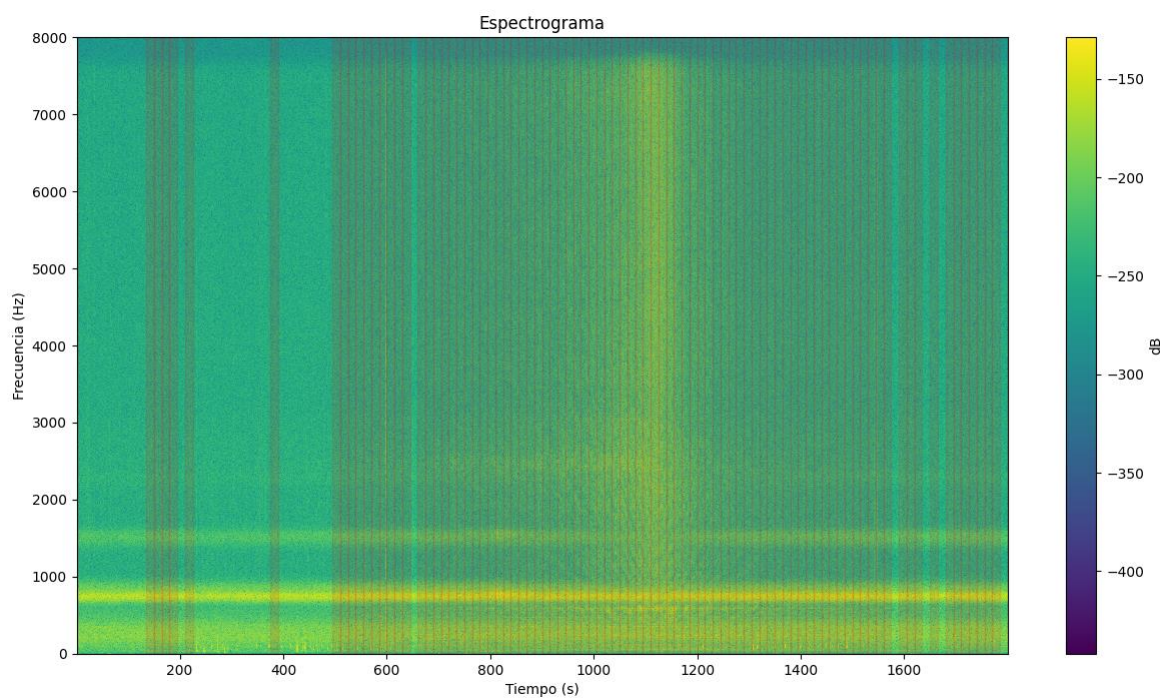


Figura A9. Detecciones sobre espectrograma de larga duración obtenidas con SVM.

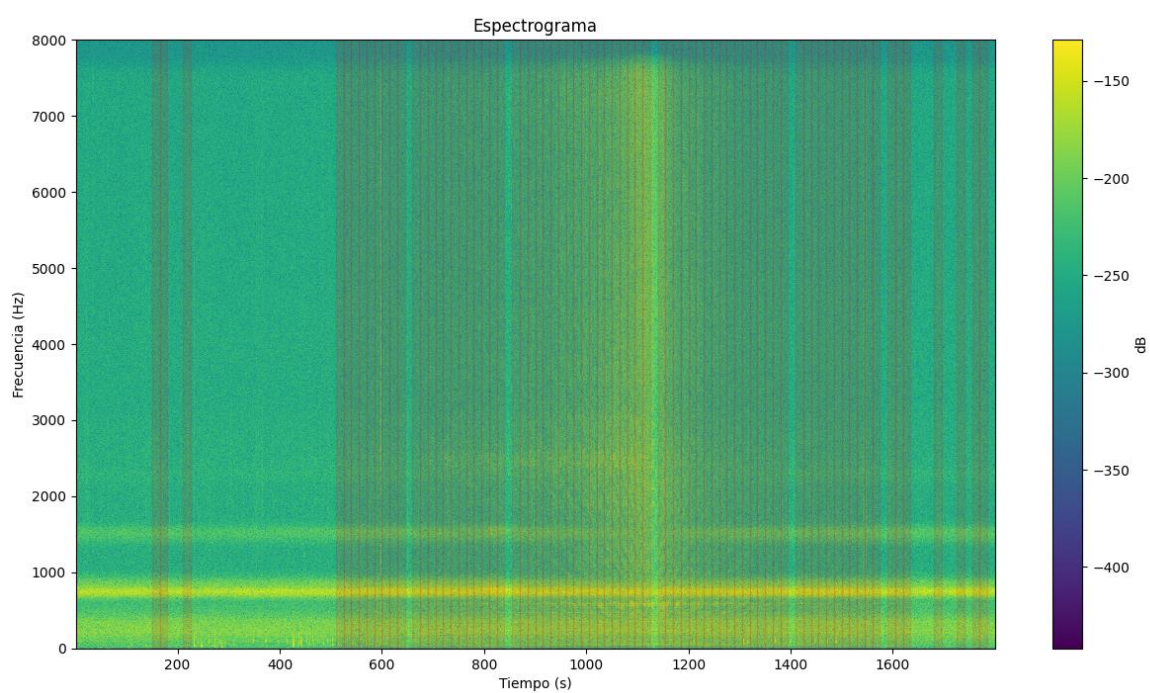


Figura A10. Detecciones sobre espectrograma de larga duración obtenidas con KNN.

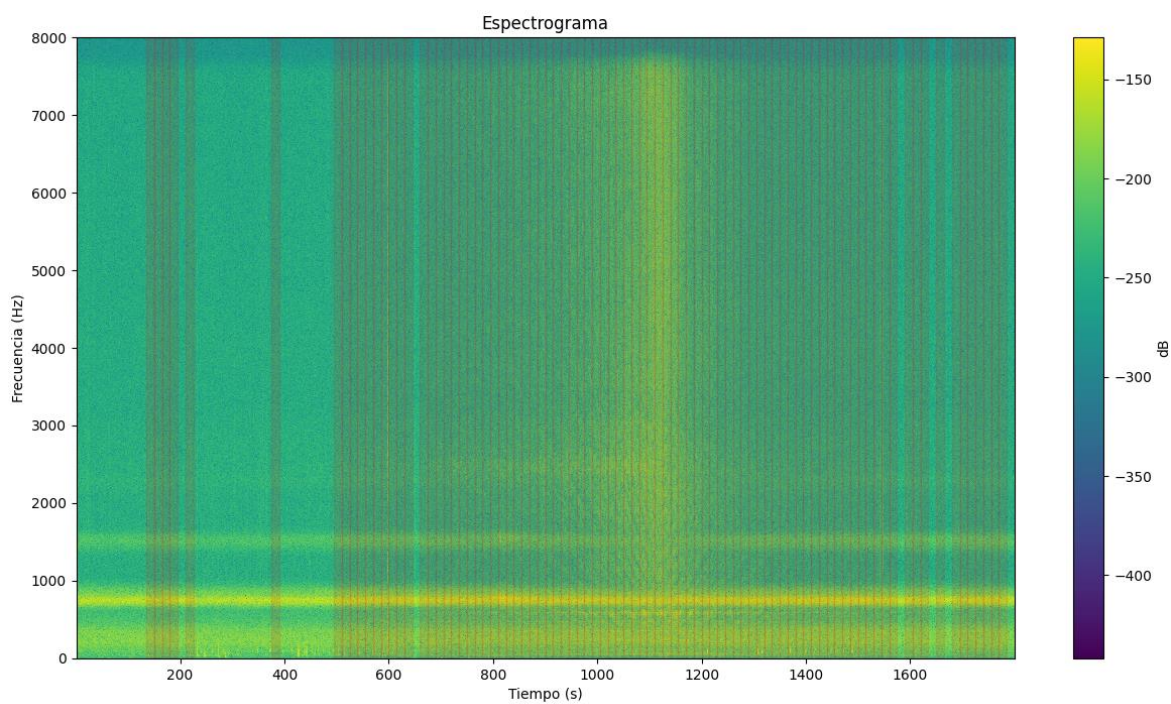


Figura A11. Detecciones sobre espectrograma de larga duración obtenidas con RF.

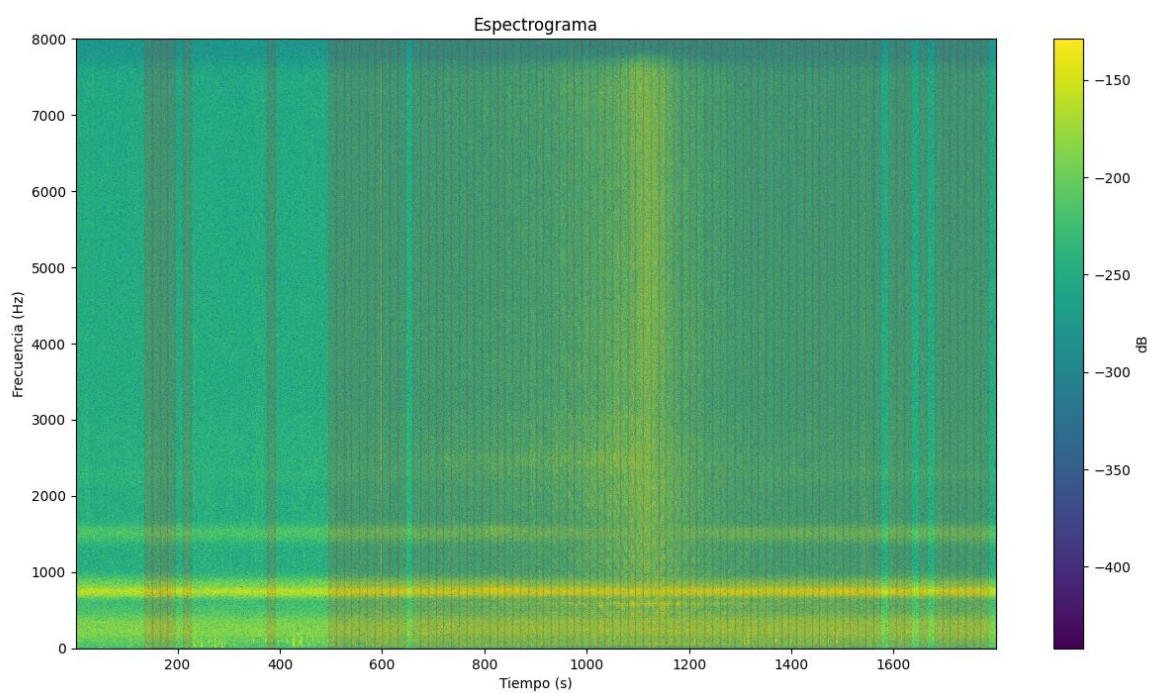


Figura A12. Detecciones sobre espectrograma de larga duración obtenidas con DT.

## **ANEXO II. CÓDIGOS DESARROLLADOS**

El código del algoritmo desarrollado en este trabajo de tesis continuará siendo actualizado y mejorado para optimizar su rendimiento. Este es el motivo por el cual no se proporciona explícitamente el código, pero el mismo está disponible en el siguiente repositorio de GitHub:

<https://github.com/TomyMartin/deteccion-clasificacion-embarcaciones.git>

En caso de que el repositorio no se encuentre disponible por cualquier motivo, los interesados pueden solicitar acceso enviando un correo electrónico a la dirección:

[tomasheto@hotmail.es](mailto:tomasheto@hotmail.es)