



## **Identificación de tiburones con métodos de visión artificial**

Mariana Paniagua Lopera

Trabajo de grado presentado como requisito para optar al título de Ingeniera Electrónica  
otorgado por UdeA

Asesor:

David Fernández Mc Cann, PhD.

Coasesores:

Juan Pablo Serna López, PhD.

Karla Cirila Garcés García, Ph.D

Universidad de Antioquia

Facultad de Ingeniería

Ingeniería Electrónica

Medellín

2024

<b>Cita</b>	(Paniagua Lopera, M. 2023)
<b>Referencia</b>	Paniagua Lopera, M (2024). <i>Identificación de tiburones con métodos de visión artificial</i> , Trabajo de grado, Ingeniería Electrónica, Universidad de Antioquia, Medellín, 2024.
<b>Estilo APA 7 (2020)</b>	



Grupo de Electrónica de Potencia, Automatización y Robótica (GEPAR).  
 Grupo de Geografía, Limnología, Modelación Ambiental (GeoLimna).

Coasesora de trabajo de grado Universidad Veracruzana.: Karla Cirila Garcés García.



Centro de Documentación de Ingeniería (CENDOI)

**Repositorio Institucional:** <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - [www.udea.edu.co](http://www.udea.edu.co)

**Rector:** John Jairo Arboleda Céspedes.

**Decano/Director:** Julio César Saldarriaga Molina.

**Jefe departamento:** Eduard Emiro Rodríguez Ramírez.

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

## **Resumen**

El presente trabajo se enfocó en el desarrollo de un programa que mediante técnicas de visión artificial y procesamiento digital de imágenes permita la identificación automática de distintas especies de tiburones presentes en las costas colombianas y mexicanas, mediante la extracción y el análisis de los coeficientes de Fourier de la aleta dorsal.

Los datos que se utilizaron para realizar la identificación fueron imágenes de aletas tomadas en Tamiahua, Veracruz, México por estudiantes de biología marina de la Universidad Veracruzana e imágenes sacadas de bases de datos de la plataforma de investigación colaborativa iNaturalist. Estas fotos fueron minuciosamente escogidas para obtener mejores resultados al momento de la comparación, fueron segmentadas mediante una inteligencia artificial llamada SAM perteneciente a meta y procesadas en Python.

## Contenido

Resumen .....	3
Contenido .....	4
Lista de Figuras .....	5
Lista de Tablas .....	6
1. Introducción .....	7
2. Objetivos .....	9
3. Marco Teórico .....	10
3.1. Estado del arte .....	10
3.2. Visión artificial .....	11
3.3. Erosión.....	12
3.4. Dilatación.....	12
3.5. Segment Anything Model (SAM) .....	12
3.6. Descriptores Genéricos de Fourier .....	13
3.7. Análisis de Correlación .....	15
3.8. Modelo de Color HSV .....	15
4. Metodología .....	16
4.1. Recopilación de datos .....	16
4.2. Procesamiento de los datos.....	17
5. Resultados y análisis .....	20
5.1. Resultados con SAM .....	22
5.2. Resultados con segmentado clásico.....	24
6. Conclusiones .....	29
7. Póster.....	29
8. Referencias Bibliográficas .....	31

## Lista de Figuras

Figura 1. Ejemplo de segmentación de instancias con ovejas [13].....	13
Figura 2. (a) Imagen en el espacio polar, (b) Imagen después de la conversión al espacio cartesiano [18].....	14
Figura 3. Imagen de un <i>Carcharhinus falciformis</i> sacada de iNaturalist [20].....	16
Figura 4. Fotografía de una <i>Sphyrna lewini</i> . .....	17
Figura 5. Metodología propuesta. ....	18
Figura 6. Selección del área de interés. ....	18
Figura 7. Metodología para segmentación con SAM.....	19
Figura 8. Metodología para segmentación con color. ....	19
Figura 9. Aleta segmentada. ....	20
Figura 10.. <i>Carcharhinus perezii</i> [21].....	22
Figura 11. Segmentación de la imagen. ....	22
Figura 12. Histograma de los descriptores de la aleta dorsal de un tiburón arenoso.....	23
Figura 13. Salida del programa segmentación con SAM.....	23
Figura 14. Aleta de <i>Galeocerdo cuvier</i> .....	24
Figura 15. Separación entre la aleta y el fondo. ....	25
Figura 16. Máscara <i>Galeocerdo cuvier</i> después de la erosión. ....	25
Figura 17. Histograma de los descriptores de la aleta dorsal de un tiburón tigre. ....	26
Figura 18. Salida del programa segmentación clásica. ....	26
Figura 19. Segmentación SAM vs forma clásica por umbrales. ....	27
Figura 20. Resultados <i>Galeocerdo cuvier</i> con SAM. ....	28

**Lista de Tablas**

Tabla 1. Nombres de las especies de tiburón. .... 10

Tabla 2. Descriptores genéricos de Fourier promedio de cada especie. .... 21

## 1. Introducción

En la actualidad, el identificar las especies de tiburones capturadas en los puntos de desembarque es un trabajo tedioso debido a la velocidad de las dinámicas pesqueras. Es por esta razón, que el objetivo de este trabajo fue desarrollar un programa que sea capaz de identificar 6 especies de tiburón pertenecientes al golfo de México, 3 de las cuales habitan en el pacífico colombiano. Las técnicas de procesamiento digital de imágenes e inteligencia artificial permitieron el análisis y la clasificación de las imágenes de dichos tiburones; esta clasificación se realizó mediante los descriptores genéricos de Fourier de la aleta dorsal, los cuales varían dependiendo de la especie.

Los tiburones son especies acuáticas pertenecientes al grupo de peces condriictios, caracterizadas por poseer un esqueleto de cartílago, escamas placoideas y órganos copuladores externos en los machos (llamados cláspers). Estos animales ocupan diversos ambientes tanto marinos, como estuarinos y dulceacuícolas, y diversas zonas, desde la zona tropical a las zonas polares [1]. Entre estos ecosistemas se encuentran el caribe y el pacífico tanto mexicano como colombiano. En la actualidad, existen 111 especies de tiburones en México y 76 especies de tiburones en Colombia, lo que sitúa a México y a Colombia como dos de los tres países con más fauna de peces cartilaginosos de América Latina junto con Brasil [2].

En México los tiburones son un recurso pesquero muy importante a causa de su demanda económica, debido a que se comercializa la mayoría de sus partes. Por otra parte, en Colombia la caza de tiburones está prohibida según el decreto 281 de 2021 [3], por lo que para el país la principal función de estos animales es el control poblacional. Por estas razones, tanto económicas como ambientales es que surge la necesidad de tener un registro del número de individuos de cada especie que son capturados en puerto y para ello es necesaria una identificación rápida y eficiente de las especies, específicamente de las siguientes 6 especies las cuales son las más comunes en los mares colombianos y mexicanos: *Carcharhinus falciformis*, *Galeocerdo cuvier*, *Carcharhinus limbatus*, *Carcharhinus obscurus*, *Carcharhinus perezii* y *Sphyrna lewini*. Actualmente, la identificación la realiza un personal especializado como lo son los biólogos marinos mediante guías y catálogos como la “*Guía para la identificación de especies de tiburones, rayas y quimeras de Colombia*” [1] y el “*Catálogo de aletas, tronchos y cabezas de tiburones en el pacífico mexicano*” [4]. Ambos libros utilizados y referenciados a lo largo de la realización del proyecto.

Es por esta necesidad que surge la idea de crear un programa que sea capaz de identificar estas especies de tiburones por medio de procesamiento digital de imágenes, utilizando una foto del tiburón o de su aleta dorsal. A esta foto, se le realizó un proceso de segmentación, el cuál se encarga de separar la aleta de los demás elementos que hay a su alrededor. A esta imagen ya segmentada y binarizada, se le aplicó una limpieza de impurezas mediante erosiones y dilataciones. Una vez que se tiene la imagen en blanco y negro limpia, se encuentra su centro y con este se procede a encontrar sus primeros 50 descriptores genéricos de Fourier, que son los que poseen la mayor cantidad de información y finalmente estos descriptores se comparan con los promedios de los descriptores correspondientes a cada tiburón. De esta manera, seleccionando el que posea la mayor correlación, se elige el tiburón que más se asemeje.

## **2. Objetivos**

### **A. Objetivo general**

Desarrollar un algoritmo utilizando visión artificial utilizando técnicas como filtros y operaciones de Morfología, segmentación y algoritmos de Post-Procesamiento, que permitan la identificación de 6 especies de tiburón a partir del perfil de las aletas y así poder inventariar su presencia en los puntos de desembarque pesquero.

### **B. Objetivos específicos**

- Determinar un método basado en técnicas de visión artificial e inteligencia artificial aplicada, que permita extraer el perfil de la aleta de un tiburón a partir de una imagen.
- Crear una base de datos de aletas de algunos tiburones que habitan en el golfo mexicano en colaboración con la Universidad Veracruzana.
- Desarrollar un algoritmo para la identificación de especies de tiburón a partir del perfil de sus aletas.
- Verificar en campo el algoritmo propuesto y su eficiencia al momento de la identificación.

### 3. Marco Teórico

#### 3.1. Estado del arte

En la revisión de trabajos anteriores para la identificación de tiburones se tomaron como referencia los libros utilizados por los biólogos para identificar las especies de tiburón basándose en la forma de su aleta dorsal estos libros fueron la “*Guía para la identificación de especies de tiburones, rayas y quimeras de Colombia*” [1] y el “*Catálogo de aletas, tronchos y cabezas de tiburones en el pacífico mexicano*” [4]. Por otra parte, para parte de identificación con procesamiento de imagen se utilizaron como referencia artículos como “*Técnicas de procesamiento digital de imágenes para detección de plagas y enfermedades en cultivos: una revisión*” [5].

Para la identificación correcta de la especie del tiburón se buscaron los nombres de los tiburones científicos y los nombre con los que son conocidos comúnmente los cuales fueron adquiridos directamente de un especialista, con la Doctora Karla Cirila Garcés García. Además, el nombre coloquial dado por los pescadores.

<b>Nombre Científico</b>	<b>Nombre común</b>	<b>Nombre Coloquial</b>
<i>Carcharhinus falciformis</i>	Tiburón sedoso	Prieto Grande
<i>Galeocerdo cuvier</i>	Tiburón tigre	Tintorera
<i>Carcharhinus limbatus</i>	Tiburón de puntas negras	Puntas negras
<i>Carcharhinus obscurus</i>	Tiburón arenero	Gambuso
<i>Carcharhinus perezii</i>	Tiburón coralino	Punta negra <sup>1</sup>
<i>Sphyrna lewini</i>	Tiburón martillo común	Cornuda

Tabla 1. Nombres de las especies de tiburón.

---

<sup>1</sup> Se le dice tiburón de punta negra por la coloración negra de su nariz. No confundir con la especie *Carcharhinus limbatus*.

### 3.2. Visión artificial

El IBM (International Business Machines) define la visión artificial como un campo de la inteligencia artificial que permite a los ordenadores y sistemas extraer información a través de imágenes digitales, videos y otras entradas visuales [6], mientras que Intel la define en pocas palabras, como la tecnología que brinda a los equipos industriales la capacidad de “ver” lo que están haciendo y tomar decisiones rápidas basadas en lo que observan [7]. En resumen, se podría decir que la visión artificial es un espacio dentro del mundo de la inteligencia artificial que dota a las maquinas la capacidad de ver, y de acuerdo a las imágenes o frames que capturan, obtener información y tomar decisiones en base a esta.

Todo el proceso de identificación se hace mediante el procesamiento digital de las imágenes de las aletas. El procesamiento digital de imágenes se puede definir como la serie de procesos o técnicas que se aplican digitalmente sobre una imagen. Muchos lo definen como el proceso en el cual entra una imagen y genera como salida otra imagen, definición que se queda un tanto corta para este proyecto. Sin embargo, este trabajo encuentra justo en medio de lo definido anteriormente como visión artificial y lo definido como procesamiento de imágenes. Es por ello por lo que en este trabajo se tomará el paradigma de que definen González y Woods en su libro Digital Image Processing [8] y consiste en considerar tres tipos de procesos computarizados: procesos de nivel bajo, medio y alto. Los procesos de bajo nivel implican operaciones primitivas como reducir el ruido, mejorar el contraste y mejorar la nitidez de la imagen y se caracteriza por el hecho de que tanto sus entradas como sus salidas son imágenes. El procesamiento de imágenes de nivel medio implica tareas como la segmentación (particionamiento una imagen en regiones u objetos), la descripción y la clasificación (reconocimiento) de individuos objetos. Un proceso de nivel medio se caracteriza por el hecho de que sus insumos generalmente son imágenes, pero sus resultados son atributos extraídos de esas imágenes (por ejemplo, bordes, contornos y la identidad de los objetos individuales). Finalmente, procesamiento de nivel superior, el cual se acerca más al objetivo de dotar a una máquina de funciones humanas, en este caso “ver”. Este proyecto se ubica en el nivel medio, cerca de la visión artificial puesto que no solo se basa en realizar operaciones primitivas, si no, que es capaz de realizar una clasificación a partir de una serie de datos que surgen a partir de una imagen.

Los métodos de visión artificial utilizados en este proyecto fueron: la morfología matemática (erosiones y dilataciones) y la segmentación de instancias. La morfología matemática es una

herramienta que se utiliza para extraer las componentes numéricas que se utilizan para representar y describir una región [8]. Gracias a esto se puede obtener otro punto de vista de una imagen resaltando detalles y características medibles que sin estas transformaciones no serían tan fáciles de notar. En otras palabras, una transformación morfológica,  $T$ , dará lugar a otro conjunto  $T(x)$ , alterando la forma del conjunto original [9]. Estas operaciones se realizan gracias a un elemento estructurante llamado *kernel* mucho más pequeño que la imagen original, el cual, es utilizado como referencia al momento de realizar operaciones como la erosión y la dilatación, las cuales se aplican al desplazarlo por la imagen. El kernel puede ser de cualquier tamaño y forma dependiendo de las necesidades de la situación.

### 3.3. Erosión

Una erosión es una operación que consiste en eliminar los pixeles existentes en el borde de un objeto con el fin de hacerlo más pequeño, o en caso de que la figura sea una línea hacerla más delgada, para ello se define el pixel de salida en el mínimo, es decir, en 0 y tomará este valor si cualquiera de los pixeles que lo rodean tienen un valor de 0 [10].

Matemáticamente se define como:

$$A \ominus B = \{z | (B)_z \subseteq A\} [8]$$

### 3.4. Dilatación

Por otro lado, la dilatación es una operación que tiene como objetivo hacer más grande el objeto o la línea más gruesa definiendo los pixeles de salida en su valor máximo '1', es decir, tomará el valor de 1 si cualquiera de los pixeles que pertenece al fondo conectan con los pixeles del objeto [11].

$$A \oplus B = \{z | (B)_z \cap A \neq \emptyset\} [8]$$

### 3.5. Segment Anything Model (SAM)

Ahora, la segmentación de instancias es un método de detección de objetos que consiste en separar una imagen en instancias individuales, en otras palabras, dibuja el límite de cada instancia de objeto independiente de su tamaño y forma [12]. Lo que caracteriza la segmentación de instancias es que no solo delimita el objeto, sino que también detecta los pixeles que pertenecen a cada instancia. Por ejemplo, en una habitación en donde se encuentran varios balones otro tipo de segmentación como la segmentación semántica no es capaz de

diferenciar un balón de otro. En cambio, la segmentación de instancias hace la diferencia entre un balón y otro, reconociendo que son dos objetos diferentes a pesar de que posean la misma forma.

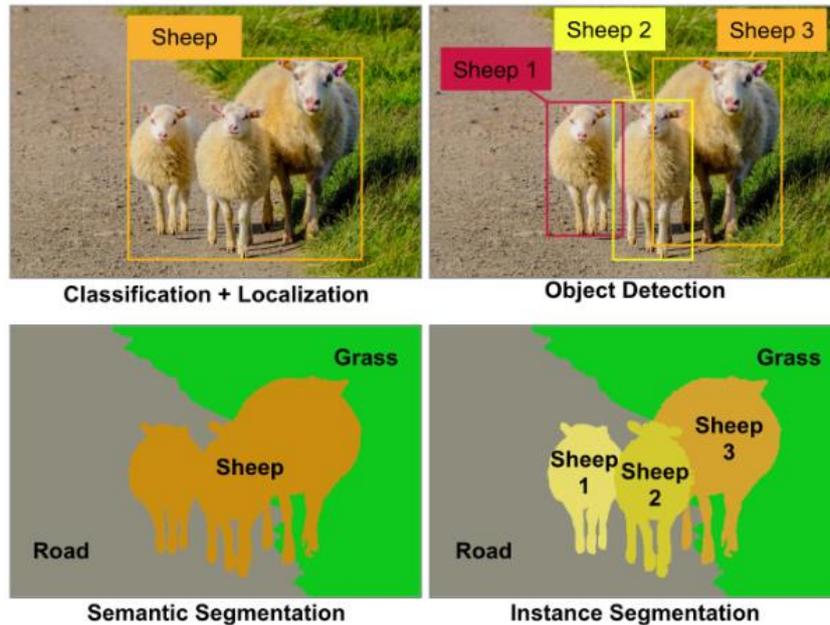


Figura 1. Ejemplo de segmentación de instancias con ovejas [13]

Esta segmentación de instancias se llevó a cabo mediante una inteligencia artificial llamada Segment Anything Model (SAM) perteneciente a Meta AI. SAM es un sistema de segmentación rápido enfocado en la segmentación de instancias con generalización sin lección o zero-shot generalization en inglés. Este modelo posee la capacidad de trabajar con datos nunca antes vistos [14]. SAM fue entrenado mediante una técnica auto supervisada llamada Mean Absolute Error (MAE) y de un entrenamiento supervisado a gran escala, ambos mediante el conjunto de datos SA-1B que es un dataset de imágenes y máscaras de segmentación de instancias de gran tamaño, es por esto por lo que SAM es capaz de segmentar objetos de cualquier tipo hasta objetos que nunca ha visto, porque ha aprendido a identificar las características generales de un objeto [15].

### 3.6. Descriptores Genéricos de Fourier

Los descriptores genéricos de Fourier o GFD con una familia de descriptores geométricos que se basan en la descomposición de una imagen en su serie de Fourier. Estos descriptores se caracterizan por ser invariantes al tamaño, la rotación o la traslación de la forma del objeto y además por tener un bajo costo computacional [16]. Los GFD se aplican a la forma de una región y utilizan la transformación de puntos cartesianos a coordenadas polares, esta

transformación se aplica a todos los pixeles de la región y se colocan los nuevos valores en una imagen cartesiana rectangular normal [17].

Los pasos para realizar esta conversión consisten en que primero se debe normalizar la imagen dividiendo cada pixel por su valor máximo. Después, se debe rotar la imagen normalizada en el sentido contrario a las agujas del reloj en un delta de ángulo muy pequeño. Luego, los valores de los pixeles a lo largo del eje x positivo teniendo en cuenta el centro como punto de referencia, se copian y se pegan en una matriz como elementos de una fila. Estos pasos se repiten hasta que la imagen gire 360°. El resultado de aplicar estos pasos es pasar una imagen en el espacio polar al espacio cartesiano, convirtiéndose en una imagen rectangular normal [16] [18].

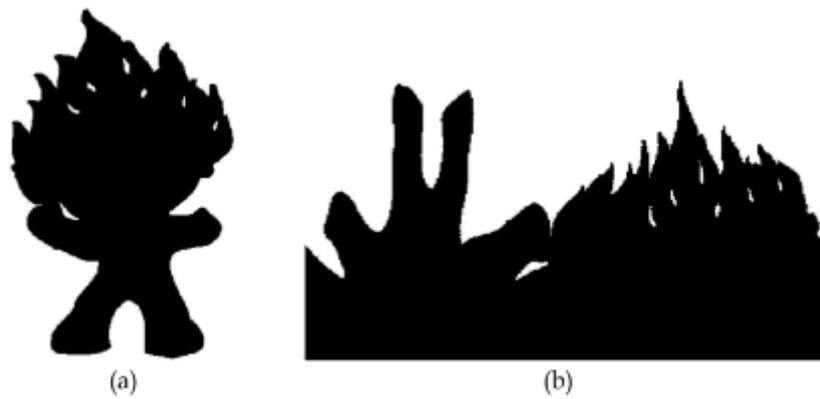


Figura 2. (a) Imagen en el espacio polar, (b) Imagen después de la conversión al espacio cartesiano [18]

Para adquirir la transformada de Fourier se aplica con la siguiente fórmula:

$$pf(\rho, \phi) = \sum_r \sum_i f(r_i \theta_i) \exp \left[ j2\pi \left( \frac{r}{R} \rho + \frac{2\pi i}{T} \phi \right) \right] \quad (1)$$

Dónde:

$0 \leq r = \sqrt{[(x - g_x)^2 + (y - g_y)^2]} < R$ , con  $(g_x, g_y)$  siendo el centro de masa,  $R$  y  $T$  siendo el radio y las resoluciones angulares respectivamente y con  $\theta_i = i \left( \frac{2\pi}{T} \right)$ ,  $0 \leq \rho \leq R_i$ ,  $0 \leq \phi < T$ .

Finalmente, los coeficientes genéricos de Fourier se hallan de la siguiente manera:

$$GFD = \left\{ \frac{|pf(0,0)|}{area}, \frac{|pf(0,1)|}{|pf(0,0)|}, \dots, \frac{|pf(0,n)|}{|pf(0,0)|}, \dots, \frac{|pf(m,0)|}{|pf(0,0)|}, \dots, \frac{|pf(m,n)|}{|pf(0,0)|} \right\} \quad (2)$$

Dónde área es el área del círculo delimitador en el que reside la imagen polar,  $m$  es el número máximo de frecuencias radiales seleccionadas y  $n$  es el número de frecuencias angulares seleccionadas [18].

### **3.7. Análisis de Correlación**

El coeficiente de correlación es una medida que indica el grado de relación que poseen dos variables, este no distingue entre variable dependiente y la independiente. La correlación se encarga de medir la asociación, no la causalidad; en esto consiste la diferencia del análisis de correlación respecto de la regresión y de la convolución, donde si una variable depende de otra en una muestra de sujetos promedio, puede generalizarse este resultado [19].

En la correlación se tiene en cuenta tanto el signo, como el valor. Si el signo de la correlación es positivo, significa que las variables tienen una relación directa. En cambio, si el signo de la correlación es negativo, significa que poseen una relación inversa. Por otra parte, si es 0, significa que las variables no tienen ninguna relación entre sí. Ahora, en términos de magnitud. Entre más cercana esté la correlación a los valores extremos (1 y -1) poseerá una tendencia más fuerte, en otras palabras, más estarán relacionadas las variables.

### **3.8. Modelo de Color HSV**

El modelo HSV por sus siglas en ingles Hue, Saturation y Valor, es un modelo de color muy utilizado en el mundo del procesamiento de imágenes y visión artificial y consiste en representar un color por medio de en 3 componentes el tono, la saturación y el valor.

El Tono (H), es el encargado de definir el color real en sí. La saturación (S), representa la intensidad o pureza del color. Un color completamente saturado es vibrante y rico, mientras que un valor de saturación bajo hace que el color parezca opaco o grisáceo y por último el valor (V), el cual significa el brillo o luminosidad del color. Un valor alto corresponde a blanco o un tono muy claro, mientras que un valor bajo representa negro o un tono muy oscuro.

## 4. Metodología

La metodología utilizada en este proyecto consta de 4 etapas, la recopilación de datos, el procesamiento de los datos, la aplicación del algoritmo seleccionado y la evaluación del desempeño del algoritmo.

### 4.1. Recopilación de datos

La recopilación de datos en este caso de imágenes de aletas de tiburón y de tiburones en sí, se llevó a cabo de dos formas, la primera fuente de imágenes fue la plataforma de investigación colaborativa iNaturalist, la cual está llena de fotografías tomadas por personas de todo el mundo. Estas fotografías son catalogadas en sus respectivas especies mediante un sistema de reconocimiento de imágenes y son corregidas por los mismos usuarios, muchos de ellos biólogos o entusiastas de la biología. De este catálogo se obtuvieron aproximadamente 140 imágenes tanto de aletas como de cuerpo completo de los 6 tiburones. Se seleccionaron las imágenes manualmente, eligiendo las fotografías con mejor calidad y en las cuales se pudiera apreciar correctamente la aleta de tiburón sin obstáculos visuales de por medio y sin que la aleta dorsal de confunda con el fondo. Similares a la siguiente imagen:

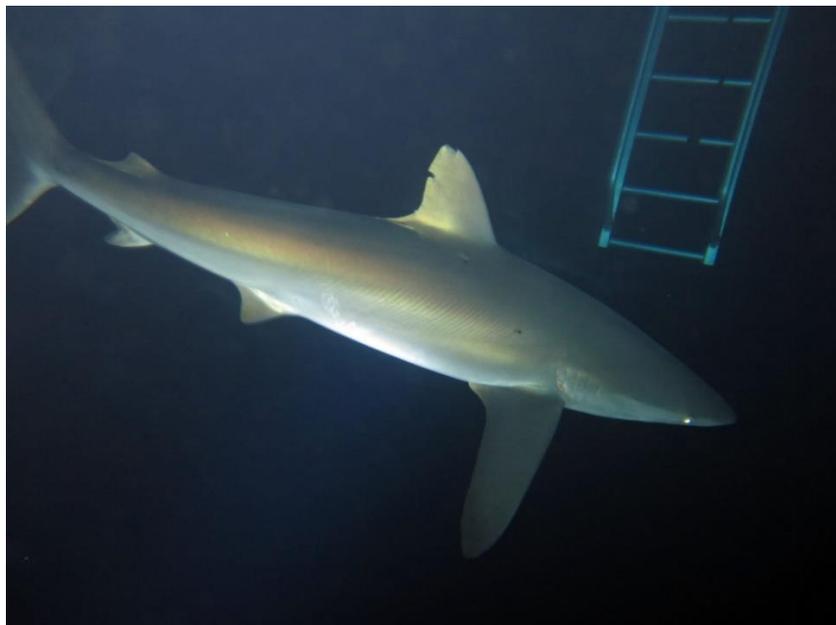


Figura 3. Imagen de un *Carcharhinus falciformis* tomada de iNaturalist [20]

Otra fuente de datos fueron las fotografías tomadas directamente en el golfo de México, en una localidad pesquera llamada Tamiahua ubicada en la región Huasteca Alta del estado de Veracruz. Estas fotografías fueron tomadas por profesores y estudiantes del programa de biología marina de la Universidad Veracruzana.



*Figura 4. Fotografía de una Sphyrna lewini.*

#### **4.2. Procesamiento de los datos**

Una vez que se cargó la imagen al programa se recorta el objeto de interés (la aleta dorsal) esto para asegurarse que el elemento más grande dentro del conjunto de máscaras producidas por SAM corresponda a la aleta dorsal. El recorte se realizó gracias a la herramienta de detectar un evento que posee la librería CV2, la cual detecta la acción ‘click’ del mouse. Se tomaron las coordenadas de la esquina superior izquierda y de la esquina inferior derecha para formar un rectángulo. Una vez obtenidas las coordenadas del objeto de interés la imagen se transforma del espacio de color BGR utilizado por defecto en CV2 a RGB para que no surjan problemas con las demás librerías.

Para el desarrollo del programa se siguió la metodología representada en el diagrama de la *Figura 5*. Este diagrama muestra los pasos que se utilizan, desde el procesamiento de la imagen entrante hasta la salida (Las dos especies de tiburón que más se asemejen).

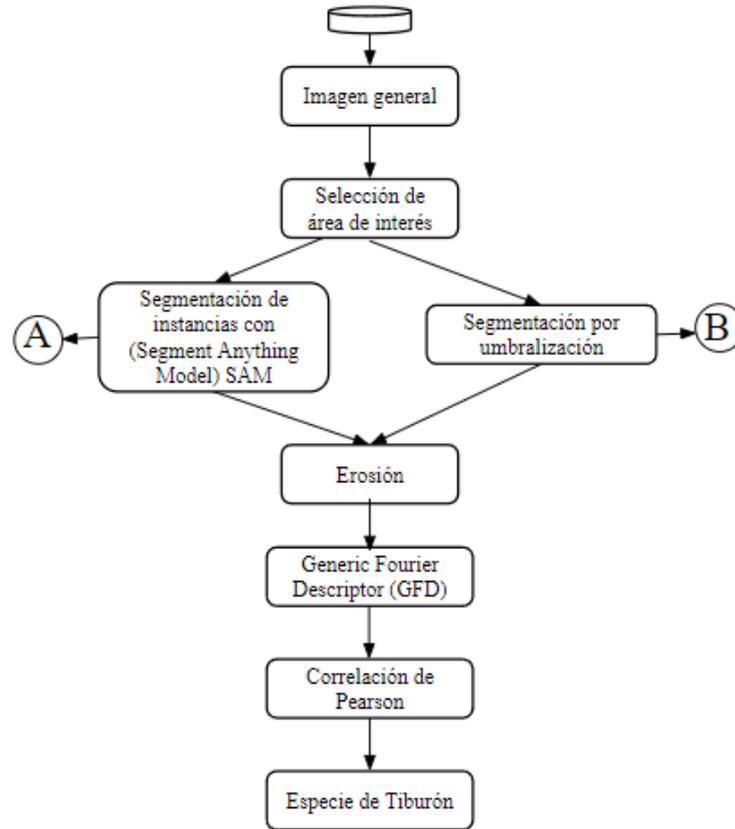


Figura 5. Metodología propuesta.

El primer paso antes de comenzar con el procesamiento de la imagen es realizar una selección del área de la aleta para facilitar la segmentación de la misma, pues esto hace que la aleta sea el principal objeto de interés en el nuevo espacio.



Figura 6. Selección del área de interés.

La siguiente parte consistió en la segmentación de la aleta y se dividió en dos partes, una segmentación mediante la AI de Meta llamada Segment Anything Model (SAM) y una segmentación mediante operaciones clásicas de procesamiento digital de imágenes.

La línea A se realizó con la siguiente metodología:

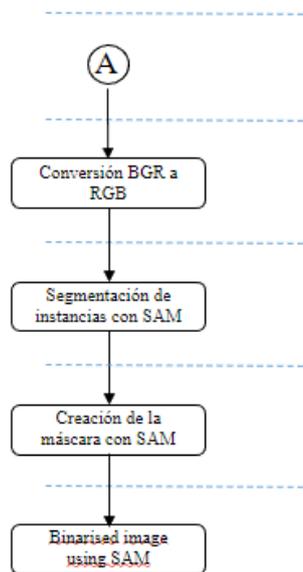


Figura 7. Metodología para segmentación con SAM.

Inicialmente se realizó una conversión de BGR a RGB para que el segmento de imagen seleccionado en la *figura 6* sea compatible con otras librerías. Una vez realizado este paso, se ingresó la imagen a los métodos de segmentación, máscaras y binarización proporcionados por la Inteligencia artificial ‘SAM’.

La línea B se realizó con la siguiente metodología:

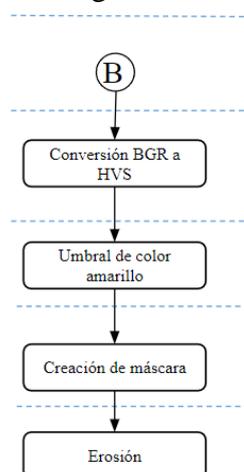
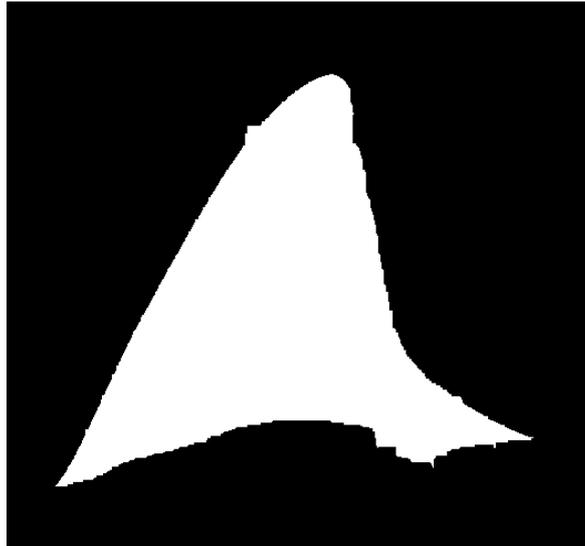


Figura 8. Metodología para segmentación con color.

Como la imagen que ingresó en esta línea tiene un fondo amarillo, lo que se hizo fue pasar el espacio de color de BGR a HVS para poder detectar el color. Una vez en este espacio, se procedió a hallar los umbrales (alto y bajo) para este color. Finalmente se hallaron las máscaras y se aplicó una erosión para eliminar el posible ruido.

Sea cual sea el método que se eligió, a la salida de las líneas se retorna una imagen en blanco y negro correspondiente a la aleta, tal y como se muestra en la *figura 9*.



*Figura 9. Aleta segmentada.*

En el proceso de segmentado se asegura que el objeto aleta estuviera centrado y no tocara los bordes de la imagen, esto con el propósito de calcular los descriptores genéricos de Fourier.

## **5. Resultados y análisis**

El procesamiento de las imágenes se realizó para los datasets de cada una de las 6 especies de tiburón, estos datasets contaban con entre 15 y 30 imágenes. A cada una de las imágenes se le calcularon los primeros 30 descriptores genéricos de Fourier (GFD), ya que, mediante un análisis visual de estos descriptores se llegó a la conclusión de que son estos primeros 30 los que capturan la mayor cantidad significativa de información con respecto a la forma de la aleta dorsal. Una vez obtenidos los descriptores de cada una de las aletas del dataset, se procedió a calcular la media para obtener como resultado final una columna con los Descriptores promedio de cada especie, como se muestran en la *tabla 1*.

<i>G. cuvier</i>	<i>C. falciformis</i>	<i>C. limbatus</i>	<i>C. obscurus</i>	<i>C. perezi</i>	<i>S. lewini</i>
0.611595	0.908016	0.744197	0.815407	0.819054	0.801442
0.061082	0.061287	0.046726	0.061488	0.058153	0.052019
0.118628	0.146924	0.119994	0.084714	0.142796	0.197667
0.183962	0.13261	0.189584	0.163241	0.161514	0.146564
0.064572	0.089647	0.095951	0.081169	0.069568	0.073949
0.083351	0.050153	0.055686	0.030044	0.069033	0.091467
0.262122	0.167828	0.188898	0.208286	0.177207	0.15445
0.049148	0.052291	0.036312	0.034039	0.047242	0.055465
0.055911	0.09372	0.08091	0.075989	0.100222	0.097897
0.110283	0.062881	0.098736	0.092848	0.080298	0.069549
0.046942	0.065035	0.056732	0.061058	0.053929	0.059121
0.056995	0.0547	0.063177	0.038767	0.065942	0.064944
0.033737	0.01547	0.025122	0.022553	0.018177	0.022462
0.040222	0.025234	0.016763	0.028013	0.022901	0.021579
0.030462	0.02719	0.031737	0.029472	0.036396	0.041817
0.036626	0.019696	0.019722	0.026157	0.017838	0.020667
0.031914	0.029506	0.019162	0.029664	0.019374	0.017298
0.025809	0.028291	0.031859	0.032399	0.031563	0.024452
0.022707	0.014146	0.014988	0.016743	0.017321	0.021317
0.020922	0.012144	0.011625	0.017021	0.010412	0.015334
0.016554	0.014712	0.016171	0.016615	0.023672	0.010568
0.020694	0.013757	0.014363	0.012528	0.012555	0.011715
0.022084	0.012521	0.017981	0.018526	0.014631	0.018914
0.019529	0.011047	0.013686	0.015591	0.014221	0.011179
0.015997	0.008876	0.007679	0.011891	0.00931	0.006882
0.010654	0.008364	0.011579	0.00984	0.011476	0.01005
0.012756	0.008626	0.011364	0.011287	0.007733	0.008749
0.01319	0.010478	0.007129	0.007767	0.008675	0.00944
0.01275	0.007677	0.010248	0.010593	0.012862	0.009712
0.013734	0.0073	0.009453	0.008955	0.008217	0.00867

Tabla 2. Descriptores genéricos de Fourier promedio de cada especie.

Una vez obtenida esta tabla se puede empezar a analizar el rendimiento del algoritmo, esto se hace realizando un análisis de similaridad mediante la correlación de Pearson entre los descriptores que corresponden a la imagen que se desea identificar con cada una de las columnas de la *tabla 1*. Al final, se retornan las 2 especies que tuvieron el índice de correlación más cercano a 1, es decir, los descriptores más similares a los descriptores de la imagen de entrada.

### 5.1. Resultados con SAM

A continuación, se ingresó al software una imagen con la que nunca se había trabajado. Esta imagen corresponde a un tiburón de la especie *Carcharhinus perezii* o tiburón coralino.



Figura 10.. *Carcharhinus perezii* [21]

Al ser la imagen de entrada una fotografía de cuerpo entero, se utilizó el método de segmentación de instancias con SAM. En la *figura 11* se puede apreciar la segmentación y la creación de la máscara sobre la aleta dorsal del espécimen.



Figura 11. Segmentación de la imagen.

Como se puede observar SAM realiza un proceso de segmentado bastante preciso, a pesar de que hallan otros objetos alrededor de la aleta como el coral o el buzo, la inteligencia artificial no presentó ningún inconveniente.

Una vez obtenida la mascarará de la aleta, el programa procedió a calcular los Descriptores de Fourier respectivos. Estos coeficientes se encargan de describir la forma de la aleta.

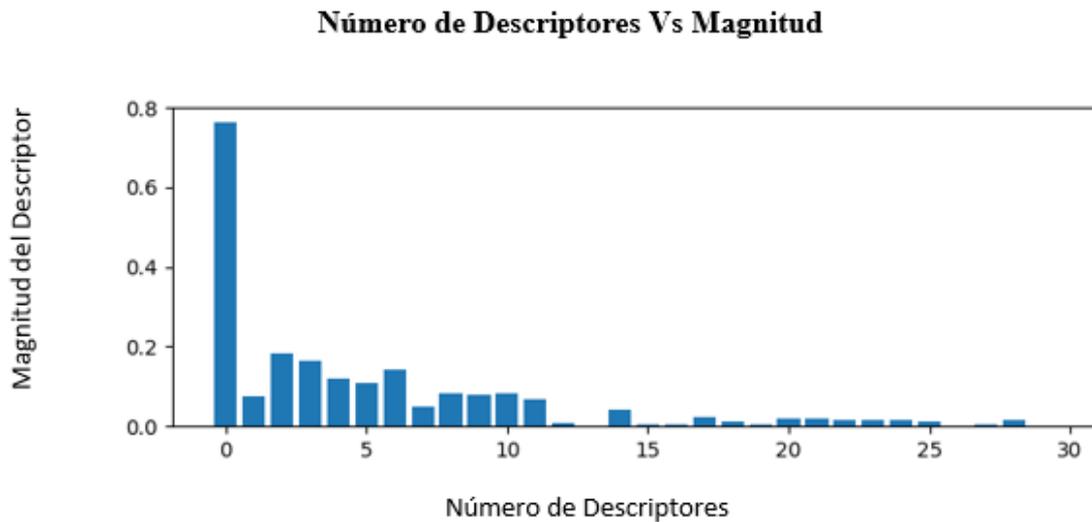


Figura 12. Histograma de los descriptores de la aleta dorsal de un tiburón arenero.

Una vez calculados los descriptores se procede con el análisis de similitud utilizando la correlación de Pearson, gracias al cual, el programa toma una decisión y entrega como resultado las dos especies más similares.

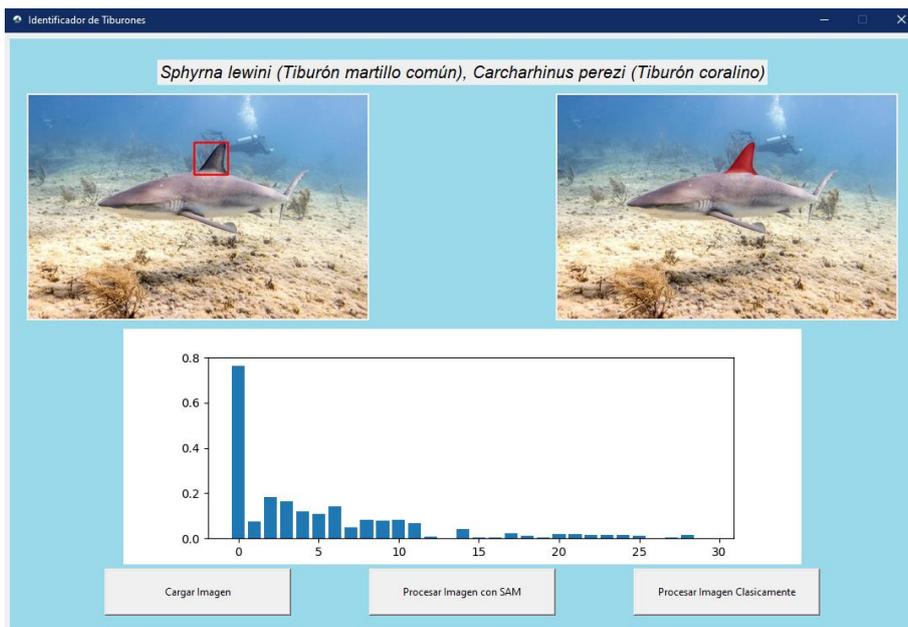


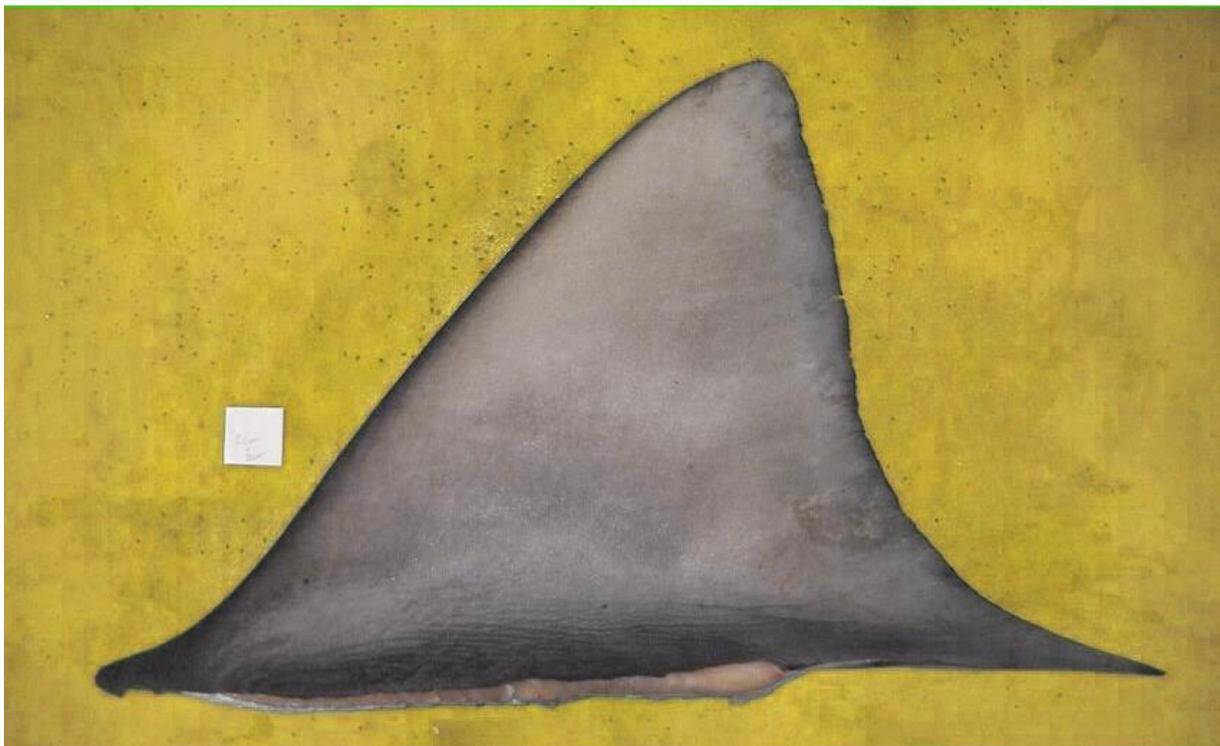
Figura 13. Salida del programa segmentación con SAM.

En este caso se puede ver como el software regresó las dos opciones más probables, que fueron la *Sphyrna lewini* y el *Carcharhinus perezi*. Este último, es efectivamente la especie correspondiente a la imagen ingresada.

Como se puede observar, la salida del programa se compone de 4 elementos. Las dos posibles opciones de la especie de tiburón, una imagen de la foto original y la foto segmentada, los descriptores y los botones para continuar con otra identificación.

## 5.2. Resultados con segmentado clásico

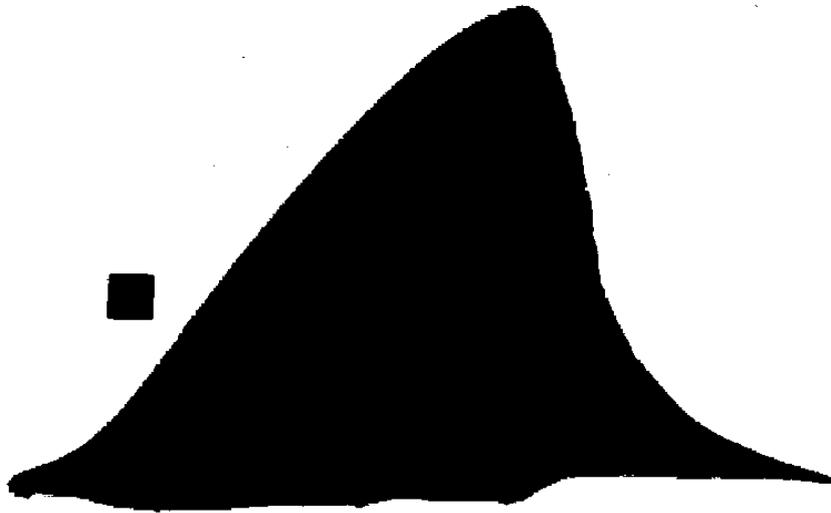
Para el segmentado clásico se requiere de una imagen de la aleta preferiblemente ya separada del trozo, sin ningún objeto adicional alrededor como la imagen de la *figura 14*. Esta imagen fue una de las fotografías proporcionadas por la Universidad Veracruzana. Esta imagen corresponde a la aleta dorsal de un *Galeocerdo cuvier*, también conocido como tiburón tigre.



*Figura 14. Aleta de Galeocerdo cuvier.*

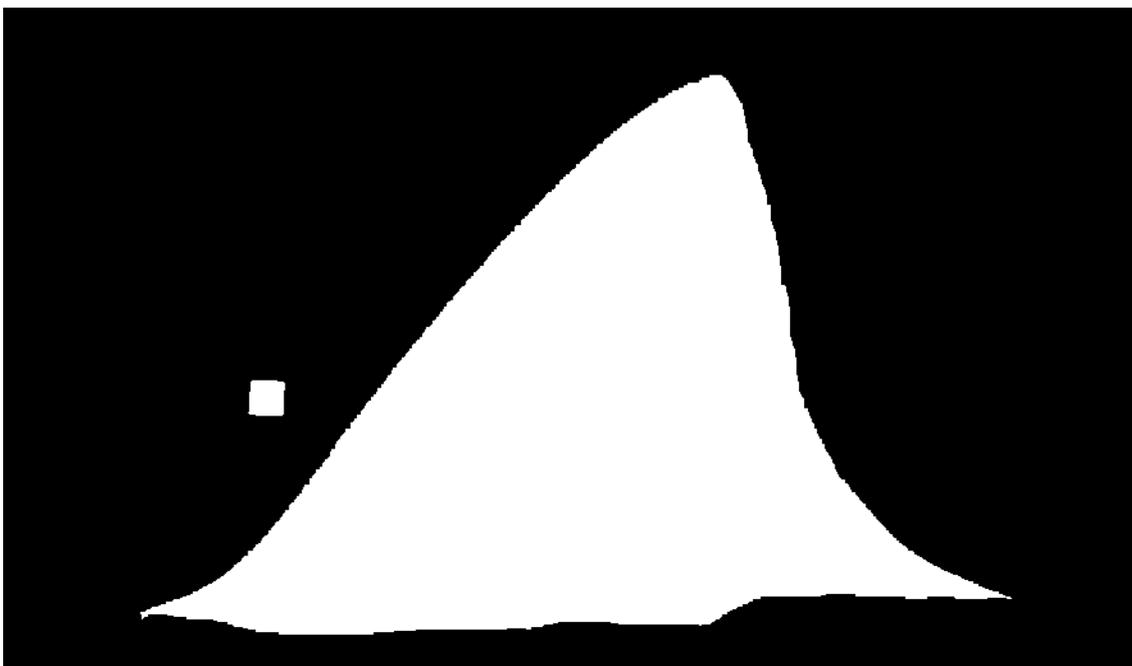
El segmentado de esta imagen se realiza por medio de la línea B definida en la metodología, que consistía en identificar el fondo amarillo definiendo los umbrales de color en el espacio HSV entre [20, 100, 100] y [30, 255, 255], donde el primer valor del arreglo corresponde al tono, el

segundo a la saturación máxima y el último al valor máximo de brillo. Para lograr separar la imagen de la siguiente manera.



*Figura 15. Separación entre la aleta y el fondo.*

En primera instancia, la umbralización no realiza una segmentación exacta de la aleta, pues alrededor de la imagen se pueden observar manchas negras al rededor, que corresponden a elementos que difieren del color del fondo, como lo son en este caso las manchas de sangre. Es por ello por lo que se realizan operaciones de erosión sobre la imagen sementada, para eliminar este ruido.



*Figura 16. Máscara Galeocerdo cuvier después de la erosión.*

A pesar de realizar las iteraciones de erosión para eliminar la mayor cantidad de ruido posible, la máscara se creó con el cuadrado blanco correspondiente a un trozo de papel que se encontraba a su lado. Esto no afecta en gran medida el resultado ya que el programa está diseñado para tomar la máscara más grande y es a esta, a la cual se le calculan los descriptores de Fourier.

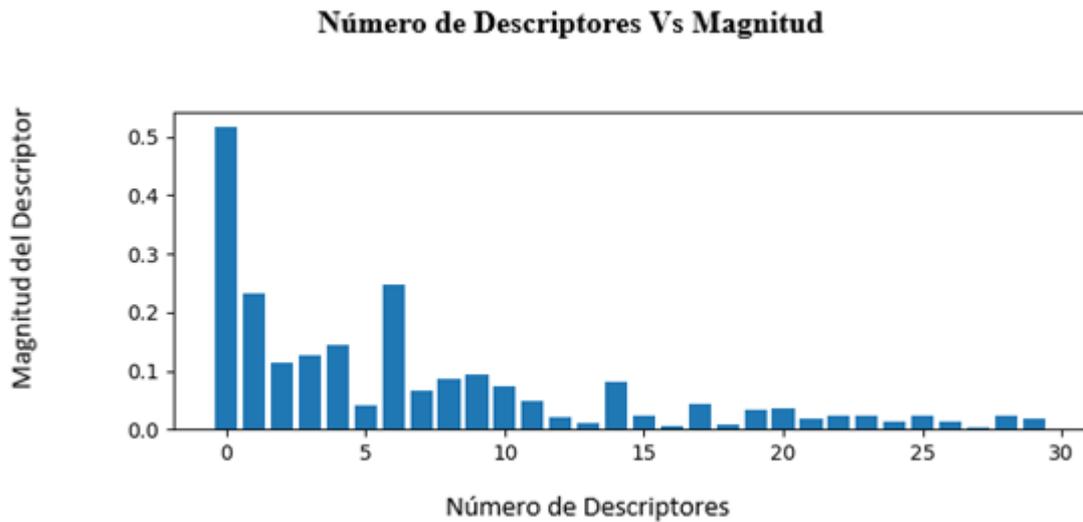


Figura 17. Histograma de los descriptores de la aleta dorsal de un tiburón tigre.

Finalmente, el programa realizó el análisis de correlación y retorno las dos especies de tiburón con más similitud.

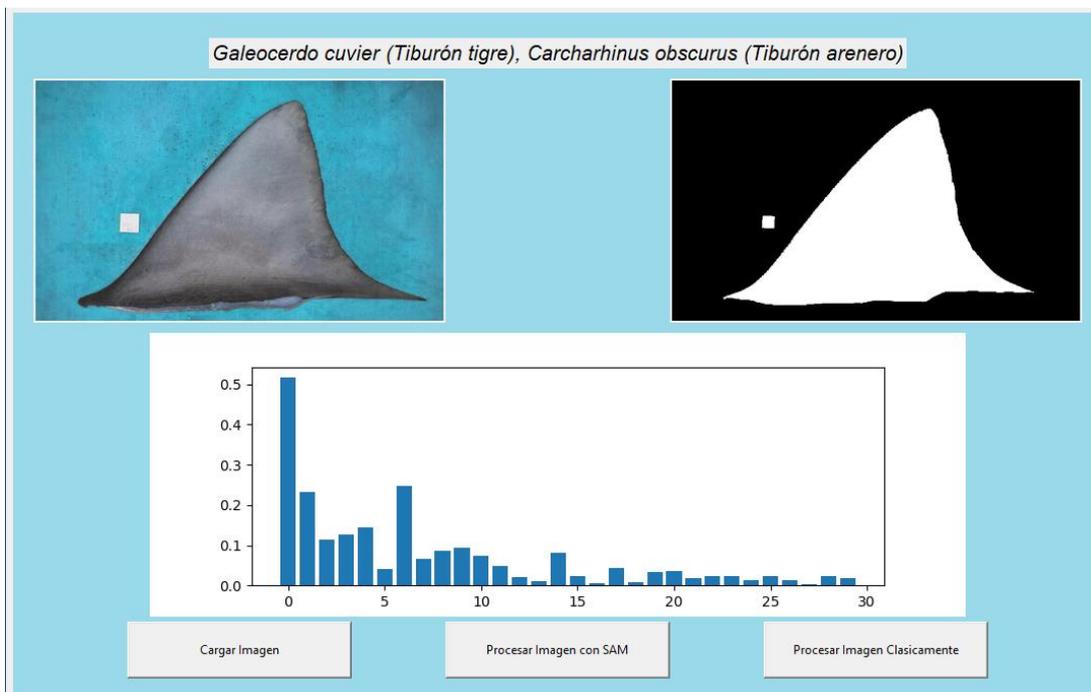
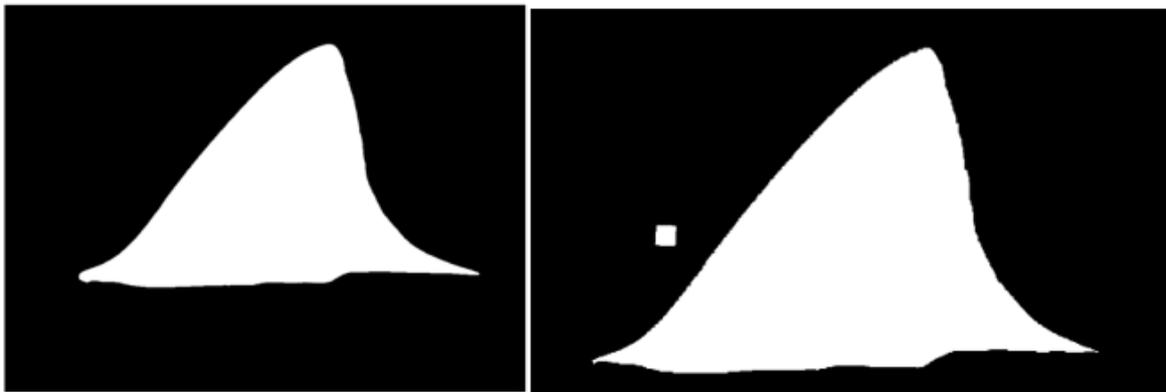


Figura 18. Salida del programa segmentación clásica.

Como se muestra en la *Figura 18*. La segmentación mediante la definición de umbrales para separar la imagen del fondo también fue capaz de identificar exitosamente la especie de tiburón correspondiente a la aleta ingresada.

Ahora, se realizó la segmentación de la misma aleta del *Galeocerdo cuvier* mediante SAM. Ya que esta opción también se puede utilizar para las imágenes con el fondo amarillo. Para contrastar las diferencias entre ambos métodos de segmentación. A la izquierda la máscara realizada por SAM y a la derecha la máscara creada por el método de la umbralización.



*Figura 19. Segmentación SAM vs forma clásica por umbrales.*

En la imagen segmentada con SAM se puede observar más claramente lo mencionado anteriormente. Los bordes son mucho más suaves y definidos, sin contar que SAM elimina automáticamente todo el ruido que rodea la imagen. A diferencia de la máscara creada por el método de umbrales que posee bordes menos definidos y las partículas de ruido más grandes no son eliminadas. A pesar de estas diferencias y como se puede observar en la *figura 20*. Ambas opciones fueron capaces de identificar correctamente la especie. La diferencia radica en la velocidad de procesamiento de ambos métodos. SAM al ser un modelo de segmentación de instancias que funciona con inteligencia artificial es mucho más demandante en recursos computacionales y en tiempo, es por ello que garantiza mejores resultados. En cambio, la segmentación por umbrales requiere menos capacidad de cómputo y actúa a una velocidad mucho mayor a la de SAM.

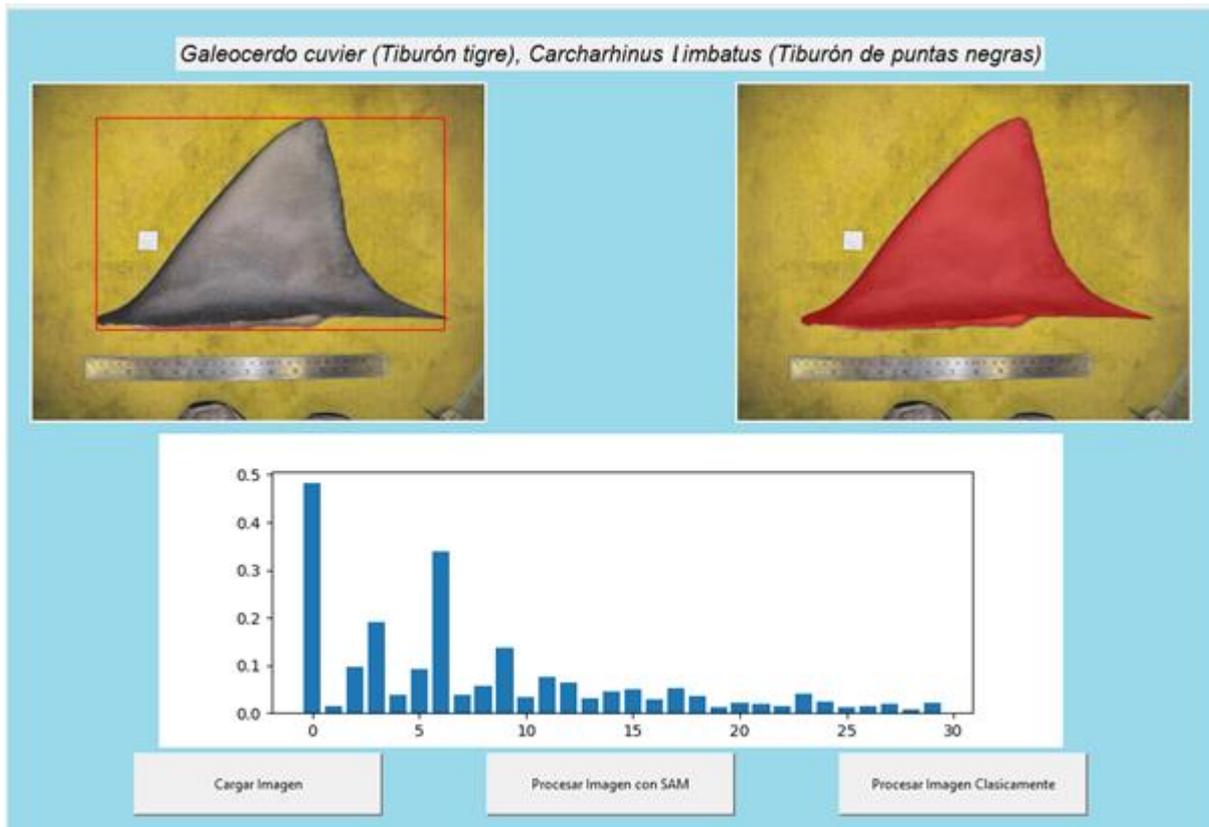


Figura 20. Resultados *Galeocerdo cuvier* con SAM.

Para la evaluación del programa no es útil calcular las métricas para cada una de las especies, puesto que se cuentan con muy pocos datos para testing, 40 imágenes en total (aproximadamente 6 datos por especie) lo que darían unos resultados que no representan correctamente el algoritmo. Es por ello, que la evaluación se hará sobre 100% del conjunto de test realizando una clasificación binaria. Clasificó bien o clasificó mal.

De las 40 imágenes nuevas ingresadas al programa para su respectiva prueba se identificaron correctamente 28 de los 40 tiburones. Lo que significa que el algoritmo posee una precisión del 70%.

## 6. Conclusiones

Segment Anything Model al ser una inteligencia artificial especializada en segmentación de instancias realiza una mejor separación de la aleta dorsal del tiburón, lo que mejora el cálculo de los descriptores de Fourier puesto que entrega un borde más suave y una forma mejor definida que la entrega la segmentación clásica por umbralización. Es por ello por lo que la identificación de especies es mucho más efectiva por este método.

Ambos métodos de segmentación son capaces de identificar correctamente una especie de tiburón, la utilización de un método u otro dependerá de la necesidad planteada. Si se requiere mayor precisión a la hora de segmentar y por lo tanto un resultado más preciso se utiliza la segmentación con SAM. En cambio, si se requiere velocidad y no se posee una gran cantidad de recursos de cómputo entonces la segmentación clásica es la mejor opción.

El conjunto de datos es notablemente limitado, la mayoría de las imágenes utilizadas provienen iNaturalist y fueron seleccionadas manualmente puesto que las imágenes son tomadas por gente común y no todas contaban con buena calidad. La otra parte de las imágenes obtenidas para el dataset provienen de las fotografías tomadas por la Universidad Veracruzana en el momento en que los pescadores arribaban al puerto, esto también presentaba un dificultad puesto que la pesca del día e incluso de la semana dependía de factores externos como las suradas, la entrada de Norte (Masa de aire frío que se extiende por todo el noroeste de México) y el clima, esto sin contar que cada tiburón tiene una época diferente del año y algunos se encuentran en peligro de extinción.

A pesar de que el conjunto de datos utilizado para el programa es relativamente pequeño, la precisión del algoritmo es bastante aceptable (70%).

Analizando el algoritmo se reveló que las imágenes en las cuales se puede apreciar en su totalidad la aleta dorsal de tiburón (Si rotación en el eje vertical) arroja mejores resultados puesto que se capta en su totalidad la forma de la aleta y produce una mejora en el cálculo de los descriptores de Fourier.,

# Identificación de tiburones con métodos de visión artificial



**UNIVERSIDAD DE ANTIOQUIA**

Facultad de Ingeniería

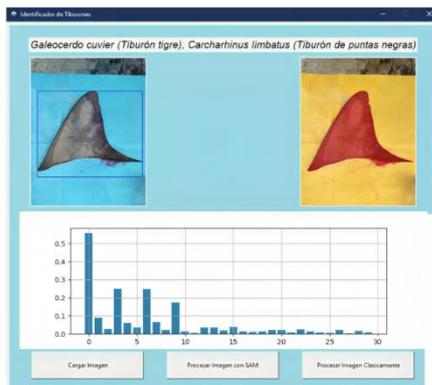
**PRACTICANTE:** Mariana Paniagua Lopera

**PROGRAMA:** Ingeniería Electrónica

**ASESORES:** David Fernández Mc Cann  
Juan Pablo Serna López  
Karla Cirila Garcés García

Semestre de la práctica: 2023-2

Este trabajo se enfocó en el desarrollo de un programa que mediante técnicas de visión artificial y procesamiento digital de imágenes permita la identificación automática de 6 especies de tiburones presentes en las costas colombianas y mexicanas, mediante la extracción y el análisis de los descriptores de Fourier de la aleta dorsal.



## Resultados

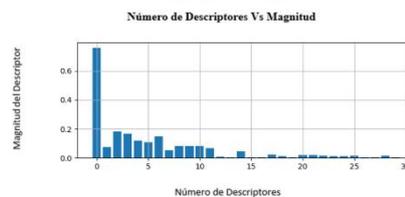
El procesamiento de las imágenes se realizó para de cada una de las 6 especies de tiburón, estos datasets contaban con entre 15 y 30 imágenes. A cada una de las imágenes se le calcularon los primeros 30 descriptores genéricos de Fourier (GFD). Así, a la imagen que ingresa al sistema se le calculan los GFD y se le realiza un análisis de similitud con respecto a los GFD obtenidos de cada especie

N°	G. cuvier	C. falciformis	C. limbatus	C. obscurus	C. perez	S. lewini
1	0.611595	0.908016	0.744197	0.815407	0.819054	0.801442
2	0.061082	0.061287	0.046726	0.061488	0.058153	0.052019
...	...	...	...	...	...	...

Imagen segmentada por SAM.



Descriptores de la aleta dorsal.



## SAM vs Forma Clásica

**Galeocerdo cuvier (Tiburón tigre)**



## Introducción

En la actualidad, el identificar las especies de tiburones capturadas en los puntos de desembarque es un trabajo tedioso debido a la velocidad de las dinámicas pesqueras. Esta identificación es necesaria debido a la importancia comercial y ambiental que poseen estos peces. Es por esto que surge la necesidad de tener un software que sea capaz de realizar una identificación rápida y eficiente de las especies, específicamente las especies Carcharhinus falciformis, Galeocerdo Cuvier, Carcharhinus limbatus, Carcharhinus obscurus, Carcharhinus perez y Sphyrna lewini.



## Metodología

La metodología utilizada en este proyecto consta de 4 etapas, la recopilación de datos, el procesamiento de los datos (procesado por SAM o procesado clásico), la aplicación del algoritmo seleccionado y la evaluación del desempeño del algoritmo.



## Objetivos

- ✓ Desarrollar un algoritmo utilizando técnicas de visión artificial que permitan la identificación de 6 especies de tiburón a partir del perfil de las aletas
- ✓ Determinar un método basado en técnicas de visión artificial e inteligencia artificial aplicada, que permita extraer el perfil de la aleta de un tiburón a partir de una imagen.
- ✓ Crear una base de datos de aletas de algunos tiburones que habitan en el golfo mexicano en colaboración con la Universidad Veracruzana.
- ✓ Desarrollar un algoritmo para la identificación de especies de tiburón a partir del perfil de sus aletas.

## Conclusiones

- ✓ SAM al ser una inteligencia artificial especializada en segmentación de instancias separa mejor la aleta, lo que mejora el cálculo de los descriptores de Fourier.
- ✓ Si se requiere mayor precisión a la hora de segmentar y un mejor resultado se utiliza SAM. En cambio, si se requiere velocidad y se tienen pocos recursos de cómputo entonces será con la segmentación clásica
- ✓ A pesar de que el conjunto de datos utilizado para el programa es relativamente pequeño, la precisión del algoritmo es bastante aceptable (70%).
- ✓ Analizando el algoritmo se reveló que las imágenes en las cuales se puede apreciar en su totalidad la aleta dorsal del tiburón, arrojan mejores resultados

### DATOS DE CONTACTO DEL AUTOR:

3116010179 | +57 3116010179 | mariana.paniagual@udea.edu.co | marianap15 | <https://www.linkedin.com/in/mariana-paniagual/>

## 7. Referencias Bibliográficas

- [1] P. A. Mejía-Falla, A. F. Navia, and V. Puentes, “Guía para la identificación de especies de tiburones, rayas y quimeras de Colombia,” 2011.
- [2] El Ministerio de Ambiente y Desarrollo Sostenible, “Tiburones, Rayas y Quimeras en Colombia,” <https://tiburones.invemar.org.co/#s1>.
- [3] Ministerio de Ambiente y Desarrollo Sostenible, “Decreto 281 de 2021,” pp. 1–4, May 281AD.
- [4] J. C. Castellanos Betancourt, C. E. Ramírez Santiago, and J. L. Castillo Géniz, *Catálogo de aletas, tronchos y cabezas de tiburones en el Pacífico mexicano*, vol. 1. 2013.
- [5] J. A. Gómez-Camperos, H. Y. Jaramillo, and G. Guerrero-Gómez, “Técnicas de procesamiento digital de imágenes para detección de plagas y enfermedades en cultivos: una revisión,” *Ingeniería y Competividad*, vol. 24, May 2021.
- [6] International Business Machines Corporation, “¿Qué es la visión artificial?” <https://www.ibm.com/es-es/topics/computer-vision>.
- [7] intel, “What Is Machine Vision?” <https://www.intel.la/content/www/xl/es/manufacturing/what-is-machine-vision.html>.
- [8] R. Gonzalez and Z. Faisal, *Digital Image Processing Fourth Edition*. 2019.
- [9] T. Olivares Montes and P. A. Cuenca Castillo, “La Morfología Matemática en el Tratamiento Digital de Imágenes,” *Revista de la Facultad de Educación de Albacete*, 1996.
- [10] MathWorks, “Tipos de operaciones morfológicas,” <https://es.mathworks.com/help/images/morphological-dilation-and-erosion.html>.
- [11] A. Fernández-Caballero, M. G. Manzano Arjona, E. A. González, and S. M. Tomé, “Una Perspectiva de la Inteligencia Artificial en su 50 Aniversario,” *Actas*, vol. 2, jul. 2006.
- [12] ArcGIS, “Introducción al aprendizaje profundo,” <https://pro.arcgis.com/es/pro-app/latest/help/analysis/deep-learning/what-is-deep-learning-.htm#:~:text=La%20segmentaci%C3%B3n%20de%20instancias%20es,de%20la%20forma%20d el%20tejado>.
- [13] O’REILLY, “Introducing capsule networks,” [https://www.oreilly.com/content/introducing-capsule-networks/?cmp=tw-data-na-article-ainy18\\_thea](https://www.oreilly.com/content/introducing-capsule-networks/?cmp=tw-data-na-article-ainy18_thea).
- [14] Meta AI, “Segment Anything Model (SAM),” <https://segment-anything.com/>.
- [15] A. Kirillov *et al.*, “Segment Anything.” [Online]. Available: <https://segment-anything.com>.
- [16] J. L. Mundy and A. Zisserman, “Fourier-based shape descriptors,” *IEEE Trans Pattern Anal Mach Intell*, 1991.
- [17] A. K. Jain, “Fundamentals of Digital Image Processing,” *Control of Color Imaging Systems*, 2018, [Online]. Available: <https://api.semanticscholar.org/CorpusID:53624641>
- [18] Peng-Yeng Yin, *Pattern Recognition: Techniques, Technology and Applications*. Springer, 2006.

- [19] H. Elorza Pérez Tejada, "Estadística para las ciencias sociales, del comportamiento y de la salud."
- [20] S. Le-Bris, "Tiburón Sedoso (*Carcharhinus falciformis*)," <https://www.inaturalist.org/observations/150330769>.
- [21] B. Illis, "Tiburón Coralino," <https://www.inaturalist.org/observations/205673980>.

## **Agradecimientos**

Quiero expresar mi más sincero agradecimiento a todas las personas que contribuyeron de alguna manera en la realización de este trabajo de grado:

A mi familia, especialmente a mis padres, Fabian y Claudia, a mi hermano Miguel, a mi abuela Stella, a mis tías Dolly, Mary y Paola, mi bisabuela Ofelia y a mi tío Jorge que en paz descansen. Por su amor, apoyo incondicional y por creer en mí en todo momento.

A mis amigos, en especial a Alex, Santiago, Paula, Erika, Sergio, Carolina, Pablo, Esteban y Salomé por su aliento, comprensión y por estar siempre ahí, brindándome su amistad y su ánimo.

A mis profesores, especialmente al Dr. David Fernández, al Dr. Juan Pablo Serna, al Dr. Andrés López por su orientación, enseñanzas y por compartir sus conocimientos conmigo durante estos años de formación académica.

Un agradecimiento muy especial a la Dra. Karla Garcés, quien no solo fue mi profesora, sino que también me abrió las puertas de su hogar en México, brindándome su apoyo, conocimiento y amistad en un momento crucial de mi investigación.

A mis mascotas, Candy y Vicente, por su compañía incondicional, por alegrar mis días y por estar siempre a mi lado, incluso cuando escribía este informe.

Y, por último, pero no menos importante, a Britney Spears y a Karina de Aespa, cuya música me acompañó en muchos momentos de este proceso y me dio la motivación necesaria para seguir adelante.

¡Gracias a todos!