

**UNIVERSIDAD NACIONAL DE FRONTERA**

**Facultad de Ingeniería de Industrias Alimentarias y  
Biotecnología**

**Escuela Profesional de Ingeniería de Industrias Alimentarias**



**TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE  
INGENIERO DE INDUSTRIAS ALIMENTARIAS**

**TÍTULO: “Machine learning y discriminación de la  
carne de pulpo (*Octopus mimus*) y calamar gigante  
(*Dosidicus gigas*) usando imágenes hiperespectrales”**

Autor: Br. William Daniel Vera Jimenez

Asesor: Dr. Wilson Manuel Castro Silupu

**Registro: TI – EPIIA – 026**

**Sullana – Perú**

**2022**

## **Dedicatoria**

Este trabajo va dedicado a:

- Mi madre Rosa Jiménez, como una pequeña retribución a tanto esfuerzo para que pudiese formarme como profesional.
- La memoria de mi padre Héctor Vera y mi tío Carlos Jiménez.

## Agradecimiento

Mi estima y gratitud a todos los profesionales que contribuyeron a la realización de la presente tesis;

Al Dr. Wilson Manuel Castro Silupu, quien, con su experiencia y compromiso, encaminó y fortaleció notablemente esta investigación.

A las Ingenieras; Erika Quinde y Susana Trelles, profesionales del Laboratorio de Inocuidad de Alimentos de la UNF, por el apoyo brindado en el proceso de adquisición de las imágenes hiperespectrales.

Asimismo, agradezco al programa “*Financiamiento de investigaciones conducentes al grado académico de bachiller o título profesional de la Universidad Nacional de Frontera - 2021*”, por la subvención para la ejecución de la tesis.

	<b>UNIVERSIDAD NACIONAL DE FRONTERA</b> Escuela Profesional de .....			
	<b>ANEXO 1 - B</b> <b>Carta de Conformidad del Asesor</b>			
Factor 7: Código Estándar 22-23-24 Acreditación-SINEACE <sup>11</sup>	Criterios 1,3,4,10 de Acreditación-ICACIT <sup>12</sup>	Versión: 001	Aprobado: Resolución N° 168/2021-UNF/CO	Fecha: 25/05/2021

Señor

**Luis Alfredo Espinoza Espinoza**

Coordinador (e) de la Facultad de Ingeniería de Industrias Alimentarias y Biotecnología  
 Universidad Nacional de Frontera

El suscrito, en mi condición de asesor del señor **William Daniel Vera Jimenez** de la Escuela Profesional de **Ingeniería de Industrias Alimentarias** informo que la Tesis titulada **"Machine learning y discriminación de la carne de pulpo (*Octopus mimus*) y calamar gigante (*Dosidicus gigas*) usando imágenes hiperespectrales"** se encuentra conforme a lo indicado en el Reglamento de Grados y Títulos.

Sullana, 30 de diciembre de 2021

  
 .....  
 Dr. Wilson Manuel Castro Silupu  
 DNI: 40322327

<sup>11</sup> SINEACE (2016) Modelo de Acreditación para Programas de Estudios de Educación Superior Universitaria. Disponible: <https://www.sineace.gob.pe/wp-content/uploads/2014/08/Anexo-1-nuevo-modelo-programas-Resolucion-175.pdf>

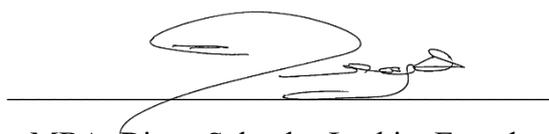
<sup>12</sup> ICACIT (2021) Criterios Generales, Informe Autoestudio ICACIT. Disponible: <https://icacitperu.org/web/es/acreditacion/criterios-de-acreditacion/acreditacion-en-ingenieria>

## Jurado Evaluador

A handwritten signature in black ink, reading "Vallejos M", written over a horizontal line.

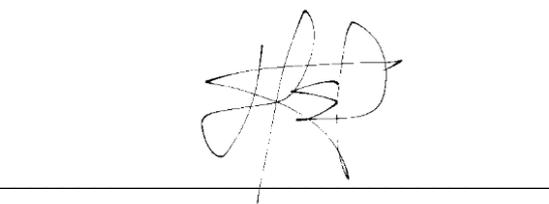
MBA. Leandro Alonso Vallejos More

Presidente

A handwritten signature in black ink, written over a horizontal line.

MBA. Diego Salvador Lachira Estrada

Secretario

A handwritten signature in black ink, written over a horizontal line.

Dr. Wilson Manuel Castro Silupu

Vocal



**UNIVERSIDAD  
NACIONAL DE  
FRONTERA**

### Acta de Evaluación de Sustentación del Informe de Tesis

Siendo las 14:00 horas del día 11 del mes de enero del año 2022 se reunieron en la sala virtual <https://meet.google.com/vvu-xiyj-ggg?authuser=0> de la Universidad Nacional de Frontera, los miembros del Jurado de Tesis para evaluar el Informe de Tesis, denominado: "Machine learning y discriminación de la carne de pulpo (*Octopus mimus*) y calamar gigante (*Dosidicus gigas*) usando imágenes hiperespectrales".

Siendo sustentado en sesión pública por el autor: Br. William Daniel Vera Jimenez, como requisito para obtener el Título Profesional de Ingeniero en Industrias Alimentarias

Terminada la sustentación, se procedió a la defensa del Informe de Tesis, etapa en que los miembros del Jurado de Tesis formularon sus inquietudes y preguntas de manera individual, las que fueron respondidas por el(los) sustentante(s).

Seguidamente, el Jurado solicitó el retiro de todos los asistentes y del (los) sustentante(s) de la sala virtual o física según sea el caso; el Jurado de Tesis determinó la calificación concedida a la sustentación del Informe de Tesis para la Obtención de Título Profesional, en términos de:

**Aprobado (a)** con el calificativo de Muy Bueno - dieciocho (18), levantándose la sesión a las 14:55 horas del mismo día. Se concluye el acto de sustentación, suscribiendo el acta.

**Presidente**

MBA. Leandro Alonso Vallejos More

**Secretario**

MBA. Diego Salvador Lachira Estrada

**Vocal**

Dr. Wilson Manuel Castro Silupu

No.	DESCRIPCIÓN	RECOMENDAR	
		SI	NO
1.	Recomendar para presentar en eventos	X	
2.	Recomendar para publicación	X	
3.	Recomendar para patente	X	
4.	Recomendar para Meritorio		
5.	Recomendar para Laureado		

Código: TI – EPIIA – 026

## Índice

Resumen .....	x
Abstract.....	xi
I. INTRODUCCIÓN .....	12
1.1. Pulpo ( <i>Octopus mimus</i> ).....	14
1.2. Calamar gigante ( <i>Dosidicus gigas</i> ).....	15
1.3. Clasificación taxonómica de las especies en estudio .....	15
1.4. Composición química de las especies en estudio .....	16
1.5. Demanda de pulpo y calamar gigante.....	17
1.6. Impacto económico .....	18
1.7. El fraude en productos marinos .....	19
1.8. Metodologías empleadas en la autenticación de productos marinos .....	21
1.9. Fundamento de la técnica de imágenes hiperespectrales .....	22
1.10. Machine learning.....	25
1.11. Antecedentes .....	26
2. MATERIALES Y MÉTODOS .....	27
2.1 Material biológico.....	27
2.2 Metodología.....	30
2.3 Obtención de las muestras .....	31
2.4 Sistema de imágenes hiperespectrales. ....	31
2.5 Preprocesamiento de las imágenes .....	32
2.6 Aplicación de modelos de Machine Learning. ....	33
2.7 Análisis de los datos. ....	34
3. RESULTADOS.....	37
3.1 Perfiles espectrales de la carne de <i>O. mimus</i> y <i>D. gigas</i> .....	37
3.2 Rendimiento de los modelos de clasificación de machine learning. ....	39
4. DISCUSIÓN.....	41
5. CONCLUSIONES .....	42
6. RECOMENDACIONES .....	42
7. REFERENCIAS BIBLIOGRÁFICAS .....	43
ANEXO 1. Glosario .....	53
ANEXO 2. Evidencias fotográficas.....	54

**Índice de tablas**

<b>Tabla 1</b> Clasificación taxonómica de <i>O. mimus</i> y <i>D. gigas</i> .....	15
<b>Tabla 2</b> Composición química de <i>O. mimus</i> y <i>D. gigas</i> .....	17
<b>Tabla 3</b> Atributos para la evaluación sensorial de cefalópodos frescos o refrigerados .	29
<b>Tabla 4</b> Estructura general de la matriz de confusión y sus medidas de rendimiento ...	35
<b>Tabla 5</b> Estadísticos de precisión (accuracy) y medida f2 (f2-measure).....	41

## Índice de figuras

<b>Figura 1</b> Principales destinos de las exportaciones peruanas de pulpo ( <i>O. mimus</i> )....	12
<b>Figura 2</b> Desembarques anuales de pulpo ( <i>O. mimus</i> ) en Perú.....	13
<b>Figura 3</b> Variación porcentual de los desembarques anuales de pulpo ( <i>O. mimus</i> ) en Perú.....	14
<b>Figura 4</b> Esquema de las principales partes externas de a) <i>O. mimus</i> y b) <i>D. gigas</i> .....	16
<b>Figura 5</b> Precios FOB por tonelada de las exportaciones peruanas de <i>O. mimus</i> y <i>D. gigas</i> congelado, durante el periodo 2014 - 2020.....	19
<b>Figura 6</b> Modos de adquisición de las imágenes hiperespectrales .....	23
<b>Figura 7</b> Hipercubo tridimensional de datos espectrales.....	25
<b>Figura 8</b> Localización geográfica del Puerto de Paita .....	28
<b>Figura 9</b> Etapas del estudio .....	30
<b>Figura 10</b> Sistema de adquisición de imágenes hiperespectrales .....	32
<b>Figura 11</b> Espectros completos de las muestra de <i>O. mimus</i> y <i>D. gigas</i> en estado fresco, congelado y cocido .....	38
<b>Figura 12</b> Espectros medios de las muestras de <i>O. mimus</i> y <i>D. gigas</i> .....	39
<b>Figura 13</b> Matrices de confusión. (a-b) Discriminante; (b-c) SVM; y (d-e) KNN .....	40
<b>Figura 15</b> Lavado de muestras con agua potable.....	54
<b>Figura 17</b> .....	54
<b>Figura 16</b> Pelado y cortado de muestras.....	54
<b>Figura 19</b> Escaneo comandado desde el software Lumi .....	55
<b>Figura 18</b> Proceso de adquisición de imágenes hiperespectrales .....	55
<b>Figura 21</b> Matriz de confusión del modelo LDA, con mejor rendimiento.....	55
<b>Figura 20</b> Obtención de perfiles espectrales usando el software Matlab .....	55

## Resumen

El pulpo es altamente susceptible al fraude por sustitución con otros cefalópodos de menor valor comercial, principalmente con calamar gigante, lo cual atenta contra el bienestar y la satisfacción de los consumidores, de modo que, es necesario contar con métodos rápidos, efectivos y de preferencia no invasivos, como la técnica de las imágenes hiperespectrales acopladas a herramientas de machine learning, para detectar esta práctica ilegal. Por tanto; el objetivo de esta investigación fue determinar si existe efecto en la aplicación de la técnica de machine learning sobre la discriminación de la carne de pulpo (*Octopus mimus*) y calamar gigante (*Dosidicus gigas*) usando perfiles espectrales. Se recolectaron muestras frescas de ambas especies en el puerto de Paita; se seleccionaron los brazos, por ser la parte de mayor similitud, a los cuales se les despojó la piel y se cortaron en trozos de 1.0 cm de longitud aproximadamente, obteniéndose 300 trozos por especie, divididos equitativamente para obtener muestras frescas, congeladas y cocidas. Las imágenes hiperespectrales se obtuvieron por reflectancia, en el rango espectral de 400-1000 nm., luego se segmentaron las imágenes para obtener los perfiles medios y se probaron modelos de machine learning para evaluar su rendimiento en base a los estadísticos de precisión y medida f-2. Los cálculos se realizaron en el software Matlab 2019a, aplicando 30 repeticiones y validación cruzada (K-fold = 5). Los resultados evidenciaron alto nivel de precisión para los modelos de Análisis Discriminante Lineal (Accuracy 99.9 %, Medida F-2 99.94%) y Maquinas de Vectores de Soporte (Accuracy 99.1 %, Medida F-2 99.06 %). Se concluye que existe efecto en la técnica de machine learning sobre la discriminación de la carne de pulpo y calamar gigante usando perfiles espectrales, con tasas de éxito de hasta 99.9%.

## Palabras clave

Fraude alimentario, cefalópodos, imágenes hiperespectrales, aprendizaje automático.

### **Abstract.**

Octopus is highly susceptible to fraud by substitution with other cephalopods of lower commercial value, mainly giant squid, which threatens the welfare and satisfaction of consumers, so it is necessary to have fast, effective and preferably non-invasive methods, such as hyperspectral imaging coupled with machine learning tools, to detect this illegal practice. Therefore, the objective of this research was to determine if there is an effect on the application of the machine learning technique on the discrimination of octopus (*Octopus mimus*) and giant squid (*Dosidicus gigas*) meat using spectral profiles. Fresh samples of both species were collected in the Paita Port; the arms were selected because they were the most similar part, the skin was removed and they were cut into pieces of approximately 1.0 cm in length, obtaining 300 pieces per species, divided equally to obtain fresh, frozen and cooked samples. Hyperspectral images were taken by reflectance, in the 400-1000 nm. spectral range, then images were segmented to obtain mean profiles and machine learning models were tested to evaluate their performance based on Accuracy and F-2 measure. The calculations were performed in Matlab 2019a software, applying 30 repetitions and cross-validation (K-fold = 5). The results evidenced high level of accuracy for the Linear Discriminant Analysis models (Accuracy 99.9 %, F-2 Measure 99.94 %) and Support Vector Machines (Accuracy 99.1 %, F-2 Measure 99.06 %). It is concluded that there is an effect of the machine learning technique on the discrimination of octopus and giant squid meat using spectral profiles, with success rates up to 99.9%.

### **Keywords**

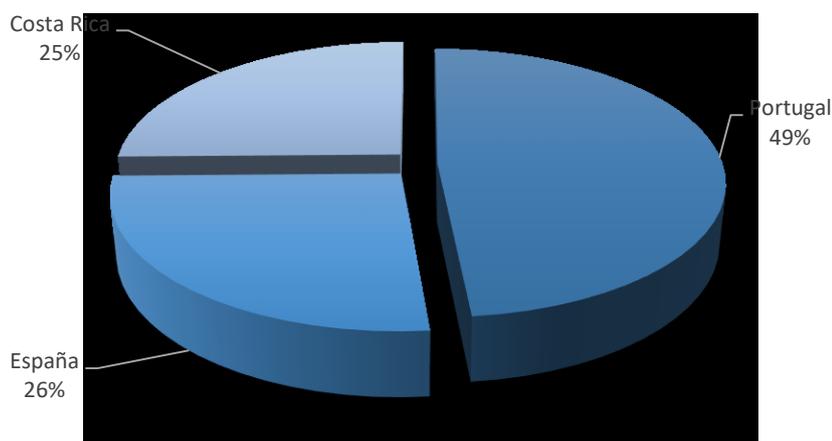
Food fraud, cephalopods, hyperspectral imaging, machine learning

## I. INTRODUCCIÓN

El pulpo es un recurso marino de alto valor nutritivo y comercial, cuya demanda ha reflejado un crecimiento sostenido en los últimos años (Mouritsen & Styrbæk, 2018). El pulpo es considerado un alimento tradicional muy apreciado en los mercados internacionales (Boyle & Rodhouse, 2007). En este rubro, las exportaciones peruanas tienen una participación significativa y sus principales mercados de destino se muestran en la Figura 1.

**Figura 1**

*Principales destinos de las exportaciones peruanas de pulpo (*O. mimus*)*



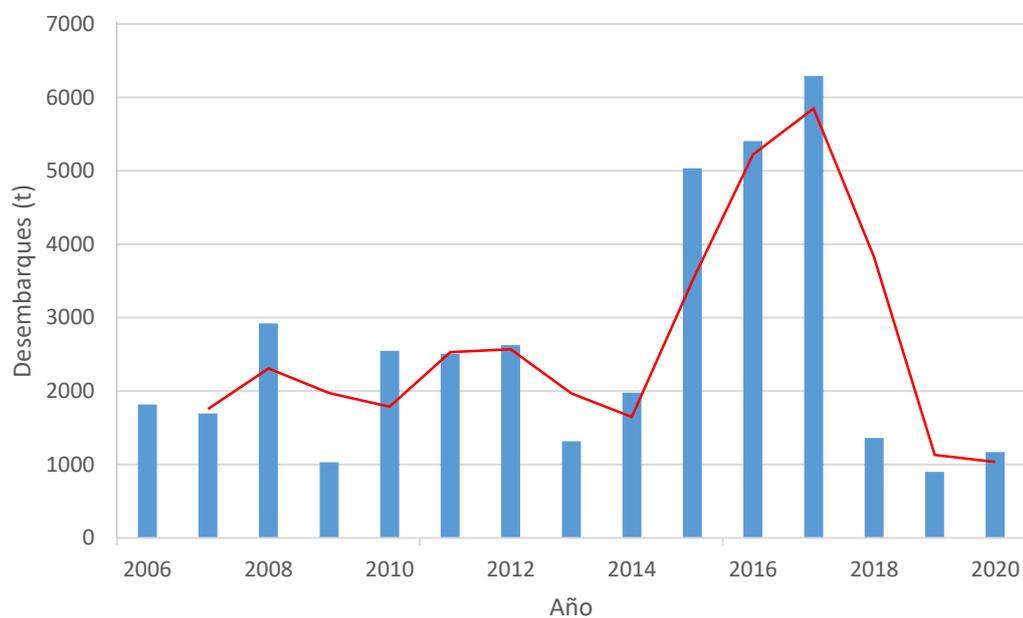
**Fuente:** Adaptado del Sistema Integrado de Información de Comercio Exterior del Perú - (SIICEX, 2021a).

En cuanto al volumen de capturas de pulpo en el Perú, Figura 2, existe gran variabilidad; con periodos de escasez y otros de relativa disponibilidad. Considerando el último quindenio; la mayoría de desembarques anuales no superó las 2 mil toneladas. Sin embargo; del 2015 al 2017 se produjo un crecimiento significativo,

alcanzando un pico máximo de 6,289 toneladas durante el 2017. La caída abrupta a partir del año 2018, coincide con la etapa post Fenómeno del Niño, lo cual podría haber afectado significativamente la reproducción de esta especie en las aguas del mar peruano.

**Figura 2**

*Desembarques anuales de pulpo (*O. mimus*) en Perú*

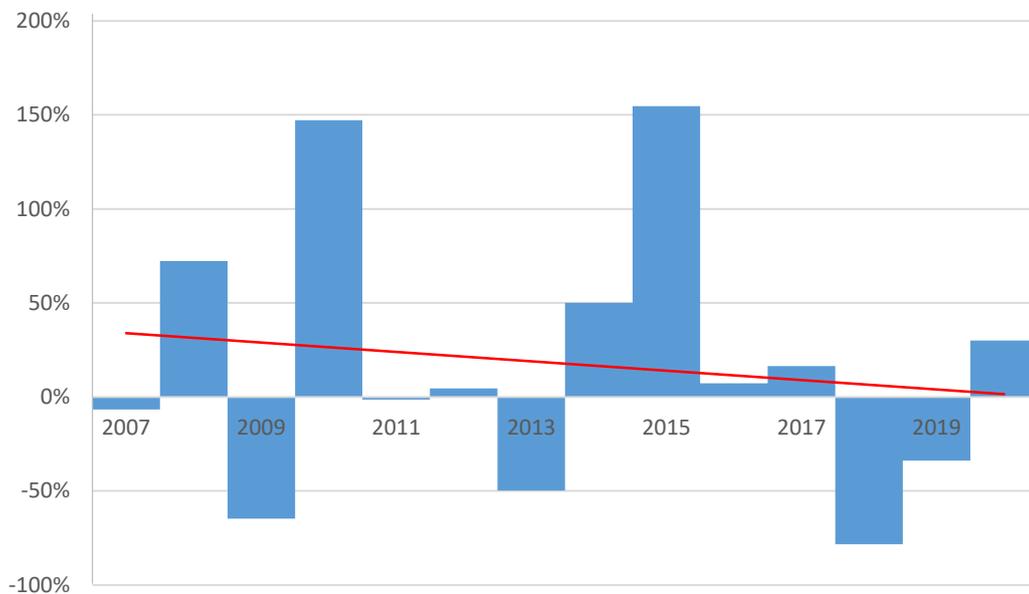


*Fuente:* Ministerio de la Producción del Perú, Produce (2021).

Ilustrando la variación porcentual de las capturas de *O. mimus*, Figura 3, se observa una tendencia decreciente en la variación de capturas anuales, registrando la cifra más baja en el 2018, periodo en el cual descendió en 78%; esta reducción en las capturas genera incrementos en el precio, lo cual alienta el fraude bajo la modalidad de sustitución con especies similares de menor valor económico, principalmente calamar gigante (Españeira & Vieites, 2012).

**Figura 3**

*Variación porcentual de los desembarques anuales de pulpo (*O. mimus*) en Perú*



*Fuente.* Ministerio de la Producción del Perú, Produce (2021).

### **1.1. Pulpo (*Octopus mimus*)**

*O. mimus*, comúnmente conocido como pulpo mimus o pulpo de Gould (Sauer et al., 2019), es una especie bentónica que se distribuye en la costa del Pacífico, entre el norte Peruano y el centro de Chile, su habitat comprende las praderas de algas con fondos rocosos (Zúñiga et al., 2014). La pesquería de *O. mimus* se inició a finales de los 80's, siendo comercializado bajo la denominación de *O. vulgaris*, posteriormente, se determinó mediante estudios genéticos que los taxones correspondían a *O. mimus* (Warnke et al., 2000), la única especie explotada en el Pacífico sudeste (Markaida & Gilly, 2016a).

## 1.2. Calamar gigante (*Dosidicus gigas*)

*Dosidicus gigas*, comúnmente conocido como calamar gigante, es un cefalópodo de la familia Ommastrephidae, presente en las aguas del pacífico oriental desde México hasta Chile, su hábitad son las aguas profundas asociadas a la corriente de Humbolt y la corriente de California (Jian et al., 2020). La pesquería de *D. gigas*, a gran escala, data desde el año 1991, con la participación de grandes embarcaciones rusas y japonesas. Actualmente, predomina la pesca artesanal, ocupando el segundo lugar en desembarques totales por año, solamente superado por la anchoveta peruana (Rovegno, 2021).

## 1.3. Clasificación taxonómica de las especies en estudio

La clasificación taxonómica de las especies en estudio, se muestran en la tabla 1; *O. mimus* y *D. gigas*, pertenecen al mismo Reino, Filo y Clase. A partir de la categoría Orden se establecen las diferencias taxonómicas entre ambas especies en base a su morfología.

**Tabla 1**

*Clasificación taxonómica de O. mimus y D. gigas*

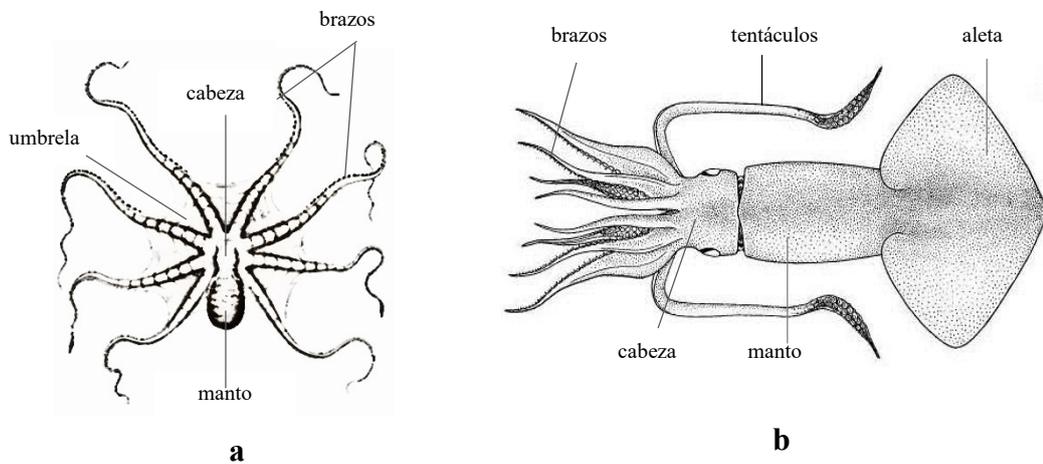
Clasificación	Taxón	
	<i>O. mimus</i>	<i>D. gigas</i>
Reino	Animalia	Animalia
Filo	Mollusca	Mollusca
Clase	Cephalopoda	Cephalopoda
Orden	Octopoda	Oegopsida
Familia	Octopodidae	Ommastrephinae
Género	<i>Octopus</i>	<i>Dosidicus</i>
Especie	<i>Octopus mimus</i>	<i>Dosidicus gigas</i>

Fuente: Registro Mundial de Molluscos (*Mollusca Base*, s. f.)

En la figura 4, se pueden apreciar las principales características morfológicas externas de *O. mimus* y *D. gigas*. Los ejemplares enteros se pueden diferenciar a simple vista; *O. mimus* tiene el cuerpo en forma de saco y carece de tentáculos, mientras que, *D. gigas* tiene cuerpo alargado con terminación en forma de aleta, además posee dos tentáculos. Se observa que los brazos son la parte de mayor similitud entre ambas especies.

**Figura 4**

*Esquema de las principales partes externas de a) O. mimus y b) D. gigas*



*Fuente.* Adaptado de Aguirre-Urreta & Cichowolski,(2007)

#### 1.4. Composición química de las especies en estudio

La composición química de las especies en estudio, según el Ministerio de Salud de Perú (2017), se muestra en la tabla 2; *O. mimus* destaca por su mayor contenido de calcio, hierro, vitamina A y vitamina C. Por su parte, *D. gigas* contiene mayores niveles de proteína, carbohidratos totales, cenizas y fósforo. En cuanto al resto de elementos, tales como; grasa total, zinc, tiamina, riboflavina y niacina, los niveles son muy cercanos entre ambas especies.

**Tabla 2***Composición química de O. mimus y D. gigas*

Elemento	Unidad	<i>O. mimus</i>	<i>D. gigas</i>
Agua	g	81.6	78.1
Proteína	g	13.6	16.0
Grasa total	g	1.4	1.1
Carbohidratos totales	g	2.2	3.1
Cenizas	g	1.1	1.7
Calcio	mg	53	9
Fósforo	mg	126	180
Zinc	mg	1.68	1.53
Hierro	mg	3.0	0.08
Vitamina A	ug	45	10
Tiamina	mg	0.02	0.02
Riboflavina	mg	0.05	0.06
Niacina	mg	1.49	1.80
Vitamina C	mg	3.2	2.3

*Fuente.* Ministerio de Salud de Perú (2017).

### 1.5. Demanda de pulpo y calamar gigante

Actualmente, existe una demanda progresiva de fuentes proteicas marinas y, ante la sobreexplotación de diversos peces comerciales, el consumo de cefalópodos se ha incrementado (Gonzalez-Pestana et al., 2021; Pierce et al., 2014). Los cefalópodos; calamares, pulpos y sepias, son muy apreciados en los principales mercados internacionales, tales como; Asia, Europa y Norteamérica (Boyle & Rodhouse, 2007),

estas especies se caracterizan por su rápido desarrollo, breve esperanza de vida y una gran capacidad de adaptación a la variabilidad climática, precisamente este último rasgo les permite incrementar su población rápidamente, lo que les confiere una especial importancia a nivel económico y ecológico (Doubleday et al., 2016). Los calamares predominan en la pesquería mundial de cefalópodos con aproximadamente el 80% de las capturas, mientras que los pulpos y las sepias comprenden cerca del 10% cada uno (Sauer et al., 2019).

En Latinoamérica, se ha evidenciado un crecimiento significativo en este tipo de pesquería; en 1990 los cefalópodos representaban el 0.5% de los desembarques totales de la región, 25 años después se incrementó a 7%, lo cual constituye un tercio de la pesquería mundial. México, Perú y Chile, abarcan casi el total de los desembarques, donde los calamares resaltan por su abundancia, mientras que el pulpo genera preocupación en cuanto a su sostenibilidad (Markaida & Gilly, 2016b), esta situación, obliga a los estados a establecer límites de pesca por temporada, lo cual repercute en el alza del precio de pulpo, condición que es aprovechada por algunos comerciantes para lucrar de forma fraudulenta.

### **1.6. Impacto económico**

De acuerdo a las estadísticas del Sistema Integrado de Información de Comercio Exterior del Perú - (SIICEX, 2021b), *O. mimus* tiene una mejor cotización que *D. gigas* en el mercado internacional, llegando en promedio a triplicar el precio FOB por tonelada, tal como se muestran en la figura 5.

**Figura 5**

*Precios FOB por tonelada de las exportaciones peruanas de *O. mimus* y *D. gigas* congelado, durante el periodo 2014 - 2020.*



*Fuente.* Adaptado del Sistema Integrado de Información de Comercio Exterior del Perú – SIICEX (2021).

Como se puede apreciar en la figura 5, durante el último septenio los precios FOB por tonelada de las especies en estudio han sufrido variaciones; en el caso de *O. mimus* se evidencia un incremento significativo durante los años 2018 y 2019 en respuesta a la escasez producida a causa de los fenómenos climáticos, efecto similar se produjo con *Dosidicus gigas* (PROMPERÚ, 2020), en todos los casos, el precio de *Octopus mimus* fue ampliamente superior.

### **1.7. El fraude en productos marinos**

El fraude alimentario consiste en la adulteración o falsificación intencional de alimentos con fines económicos (Comisión Europea, 2018), esta problemática persiste

desde los orígenes del comercio y puede ocurrir en cualquier nivel de la cadena productiva, atentando contra la sostenibilidad y la seguridad del consumo (Visciano & Schirone, 2021). Los productos marinos son altamente susceptibles al fraude, debido a la complejidad de su naturaleza y la forma en que se organizan las cadenas de suministro a nivel global (Ghidini et al., 2019), se estima que el 34% de las operaciones comerciales de estos productos se ven afectadas por prácticas fraudulentas (Saubi, 2018).

Los productos marinos, por lo general, son comercializados para consumo humano directo y se pueden encontrar en diferentes presentaciones; frescos, cocidos, congelados y en conservas (Mouritsen & Styrbæk, 2021), en su estado natural son evidentes las diferencias entre cada especie, sin embargo, durante el procesamiento pasan por una serie de operaciones que anulan las peculiaridades taxonómicas resultando muy difícil su identificación (Shao et al., 2021), dando lugar al fraude por sustitución, donde especies de gran valor comercial son reemplazadas por otras similares de menor valor, constituyendo la modalidad de fraude más frecuente en productos del mar (Pardo & Jiménez, 2020), esta práctica ilegal se impulsa para obtener ventajas económicas, en desmedro del bienestar y satisfacción del consumidor (Munekata et al., 2021; Spink et al., 2019).

Las incidencias de fraudes en productos de mar son frecuentes, así lo evidencian estudios realizados en diversas partes del mundo; en Bélgica, (Deconinck et al., 2020) reportaron casos de fraude en bacalao del atlántico procesado y lenguado común, en distintos niveles de la cadena, principalmente en el comercio mayorista y servicios de alimentos. En Italia, (Acutis et al., 2019) implementaron un plan de monitoreo para detectar sustitución de pescados, determinando que el mero era objeto de fraude con mayor frecuencia, y los supermercados eran los puntos de venta donde prevalecían la

mayoría de casos. En Madrid, (Horreo et al., 2019), determinaron que cerca del 30% de los pescados que se ofrecen en los restaurantes resultaron etiquetados de manera incorrecta. Por su parte, (Guardone et al., 2017) realizaron una encuesta sobre el etiquetado erróneo de diversos productos de mar desembarcados en Pisa, Italia, procedentes de países no Europeos, encontrando el mayor porcentaje en aquellos a base de cefalópodos (43.8%) de origen asiático.

### **1.8. Metodologías empleadas en la autenticación de productos marinos**

Para la autenticación de estos productos, se han desarrollado diversos métodos, dentro de ellos, la inspección visual es frecuentemente utilizada, este método se basa en la identificación de características morfológicas de cada espécimen, para lo cual se toman como referencia las Guías de Identificación publicadas por la FAO, este método sólo es aplicable a ejemplares enteros, en un óptimo estado de desarrollo (Espiñeira & Lago, 2016). Otros métodos analíticos que se viene empleando en los últimos años son aquellos basados en técnicas moleculares (Wang et al., 2021; Blanco-Fernandez et al., 2021; Kotsanopoulos et al., 2021) y cromatográficas (Nehal et al., 2021), estos métodos se pueden aplicar en cualquier estado de desarrollo y procesamiento, se caracterizan por su alta eficiencia y especificidad, sin embargo, presentan ciertas desventajas al tratarse de métodos destructivos, altamente exigentes en personal y equipamiento especializado (Velasco et al., 2021).

Ante esta situación, se ha incrementado la demanda de metodologías de análisis no destructivas, rápidas y fiables para la autenticación de pescados y mariscos (Suman et al., 2021; Brooks et al., 2021), encontrando en las imágenes hiperespectrales (HSI) una gran alternativa para imponerse a las limitaciones presentadas por los métodos analíticos tradicionales (Kamruzzaman et al., 2015). La tecnología HSI cumple con las

condiciones rapidez, fiabilidad y no destructiva, además, resulta apropiada para el monitoreo en línea (Suthar et al., 2017).

### **1.9. Fundamento de la técnica de imágenes hiperespectrales**

Las imágenes hiperespectrales integran la espectroscopía y la visión por computadora en un solo sistema, para obtener información espacial y espectral a través de la interacción entre la luz y las moléculas de la materia (Feng et al., 2019), la energía resultante se puede cuantificar por reflectancia, transmitancia e interactancia, la elección de la metodología depende fundamentalmente del tipo de muestra y el objetivo del análisis (Hu, Dong, & Liu, 2016).

En el método de reflectancia, la fuente de luz va dispuesta en el mismo sentido de la cámara, con un ángulo de inclinación adecuado, lo cual permite que la energía luminosa incida sobre la muestra biológica y una pequeña cantidad se refleja en la capa superficial generando reflejos especulares, el resto de luz penetra la muestra, se dispersa, se absorbe y nuevamente surge a la superficie provocando reflejos difusos. En consecuencia, la reflectancia capta señales especulares y difusas, en función a la variación de la intensidad producida por los efectos de reflexión, absorción y difusión de la luz en distintas longitudes de onda. Esta técnica es muy útil para evaluar atributos externos de la muestra (Qin et al., 2017).

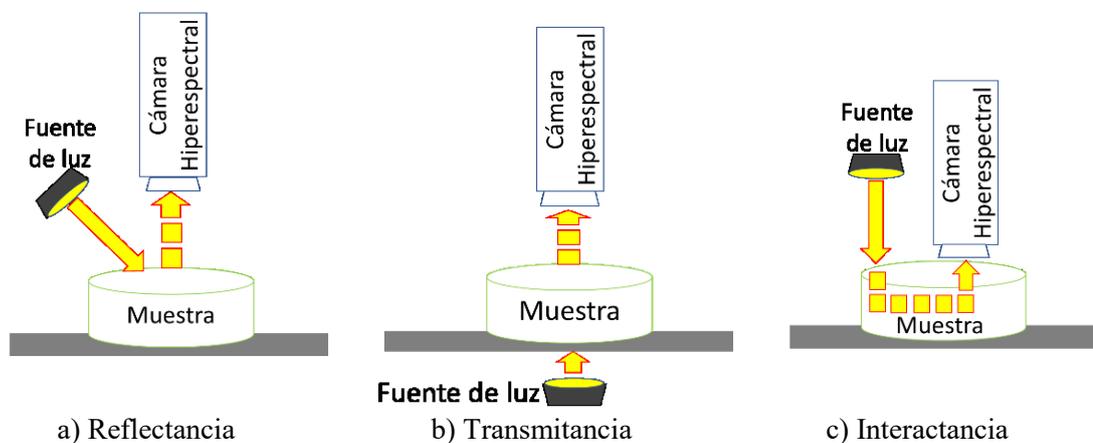
En el método de transmitancia, la fuente de luz y la cámara se disponen en lados opuestos de la muestra formando un ángulo lineal para la detección, la luz se difunde entre la muestra y traspasa para generar señales difusas detectables a diferentes longitudes de onda. El estado físico, tamaño y forma de la muestra influyen sobre la trayectoria de la luz, atenuando las señales, por lo que deberá prestarse especial

atención a estos parámetros. Esta técnica es útil para evaluar componentes y defectos internos de la muestra (Lu et al., 2020).

En tanto, el método de interactancia, comprende aspectos de las dos metodologías anteriores, la fuente de luz se ubica de forma paralela a la cámara para que los reflejos especulares no sean captados, a su vez, permite que la luz penetre una cierta cantidad de tejido debajo de la superficie, de modo tal que el grosor de la muestra no afecte el rendimiento y calidad de la imagen. Este método destaca por su capacidad para obtener información más profunda sobre la composición y estructura de la materia en estudio, debido al menor efecto de superficie, sin embargo, puede resultar inapropiada para muestras muy pequeñas (Hu, Dong, Liu, et al., 2016).

### Figura 6

#### *Modos de adquisición de las imágenes hiperespectrales*



*Fuente.* Adaptado de Castro, (2016).

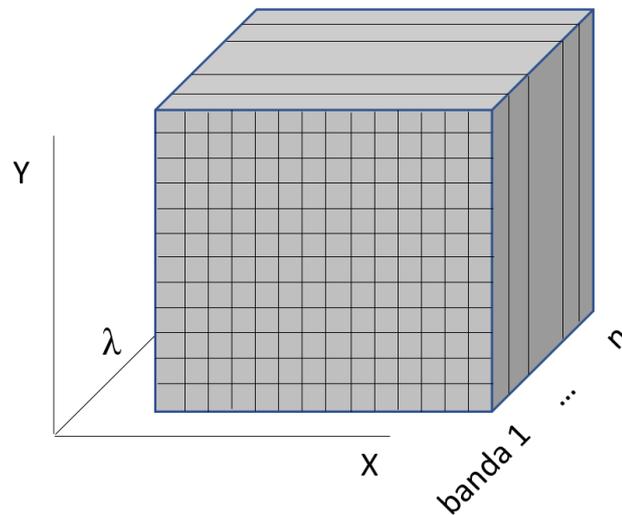
Para capturar una imagen, existen tres métodos principales de escaneo: por punto o píxel, por líneas y por área. La elección del sistema de captura de imágenes obedece al uso práctico y al costo (Khaled et al., 2021). En el escaneo por punto el espectro se adquiere en una sola ubicación espacial, desplazando el detector o la muestra para cada

ubicación en la que se requiera obtener espectros, este método demanda mayor tiempo en la obtención de imágenes hiperespectrales y por lo general se usa en HSI Raman. El método de escaneo por líneas, como su nombre lo sugiere, se realiza línea por línea, siendo necesario desplazar la muestra a una velocidad constante y regulada, este método puede ser hasta 100 veces más rápido que el escaneo por puntos, resultando adecuado para control de calidad de procesos en línea. Por último, en el escaneo de área, se captura la imagen sobre una muestra fija, para este método no se requiere contar con un espectrógrafo, puesto que las imágenes en diferentes longitudes de onda se obtienen mediante; Filtro Sintonizable de Cristal Líquido (LCTF) o un Filtro Acústico-Óptico Sintonizable (AOTF), este método solo es aplicable para imágenes multiespectrales o sistemas análogos en los que se requiere un número limitado de longitudes de onda (Kamruzzaman & Sun, 2016).

La información obtenida se ordena en hipercubos tridimensionales  $(x;y;\lambda)$ , donde “x” y “y” forman las dimensiones espaciales y “ $\lambda$ ” constituye la dimensión espectral, figura 7. Por lo general, el hipercubo contiene una mezcla de varios componentes, incluyendo objetos empleados como portamuestra, ruido espectral, interferencias espaciales e información redundante, siendo necesario separar la información de interés y eliminar todo tipo de ruidos e interferencias, para lo cual existe una vasta cantidad de algoritmos (Amigo, 2020).

**Figura 7**

*Hipercubo tridimensional de datos espectrales*



*Fuente.* Adaptado de (Maraphum et al., 2020)

### **1.10. Machine learning**

Para la clasificación de los datos espectrales de alimentos, generalmente se emplean algoritmos de aprendizaje automático supervisado (Ru et al., 2019), esta técnica requiere datos de entrada etiquetados a partir de los cuales se entrena un modelo de clasificación o predicción para el reconocimiento de patrones que permitan obtener las salidas deseadas. Entre las principales herramientas de aprendizaje supervisado tenemos; Análisis Discriminante Lineal, Bayes Ingenuo, Regresión Logística, K-Vecino más Cercano, Máquina de Vectores de Soporte, Bosque aleatorio, Árboles de Decisión y Red Neuronal Artificial.

### 1.11. Antecedentes

Entre los trabajos publicados acerca de uso de HSI para discriminación de especies marinas, destaca el realizado por Grassi et al., (2018), quienes diseñaron un método de visión artificial para diferenciar filetes de *Gadus morhua* y *Melanogrammus aeglefinus*; evaluaron un total de 181 muestras en estado fresco, de las cuales 90 correspondían a la primera especie. Las imágenes fueron obtenidas por HSI de reflectancia en el rango de 400 a 1000 nm de longitud de onda. Desarrollaron modelos de análisis discriminante lineal, basados en la variable de la textura superficial, logrando una tasa de clasificación del 86.05 al 92.31%.

Por su parte, Sun et al., (2019), evaluaron la capacidad de HSI para discriminar el músculo de *Litopenaeus vannamei* cultivados en agua dulce y agua salada; analizaron un total de 728 muestras conformadas por músculos de 1 cm de longitud. Las imágenes fueron obtenidas por HSI de reflectancia en el rango de 900 – 1700 nm de longitud de onda. La mayor tasa de precisión alcanzada fue de 92.2 %, mediante el modelo de selección directa secuencial aplicada en conjunto con el análisis discriminante de cuadrados parciales (SFS-PLS-DA).

En otro estudio, Xu et al., (2017), evaluó la capacidad del sistema de visión por computadora (CVS) y dos sistemas de HSI por reflectancia, en los rangos de 400 a 1000 nm y 897 – 1753 nm de longitud de onda ( $\lambda$ ), para discriminar entre filetes de salmón procedentes de sistemas de crianza orgánico y convencional. En total se evaluaron 80 muestras de filetes de cada sistema de crianza. Los modelos se desarrollaron a partir de Clasificadores de bosques aleatorios (RF), análisis discriminante de mínimos cuadrados parciales (PLS-DA) y máquina de vectores de soporte (SVM). La mejor tasa de clasificación (98.2 %) se obtuvo con HSI en el rango de  $\lambda=400-1000$  nm, aplicando el modelo SVM.

En consecuencia, los estudios mencionados evidencian que el análisis de imágenes hiperespectrales poseen gran potencial para la discriminación fiable entre carne de especies marinas similares. No obstante, se requiere estudiar en diferentes condiciones de proceso (fresco, congelado y cocido); de manera que, la presente investigación permitirá cubrir un vacío del conocimiento actual con significancia para la industria alimentaria. Por tanto, la presente investigación se realizó con el objetivo de determinar si existe efecto en la aplicación de la técnica de Machine learning sobre la discriminación de la carne de pulpo y calamar gigante usando imágenes hiperespectrales.

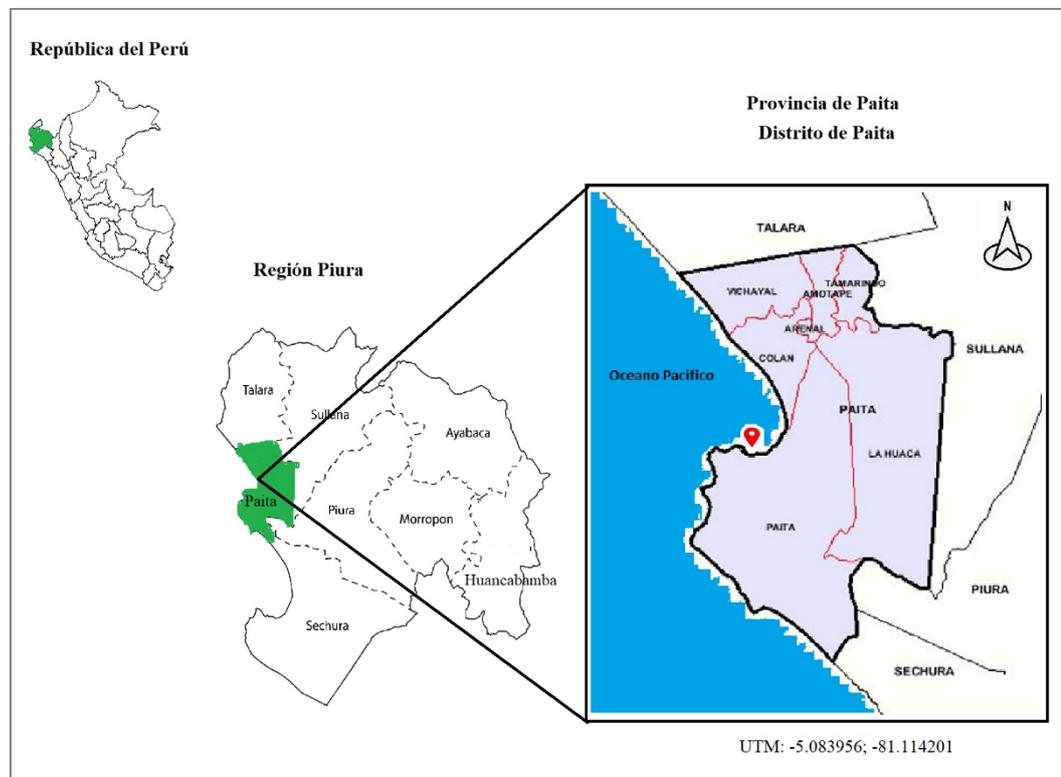
## **2. MATERIALES Y MÉTODOS**

### **2.1 Material biológico**

Los ejemplares de *Octopus mimus* y *Dosidicus gigas* fueron adquiridos en el muelle de Paita, Piura, Perú (UTM: -5.083956, -81.114201), figura 8, zona donde se realizan los desembarques, para asegurar que el producto se encuentre en estado fresco.

## Figura 8

### Localización geográfica del Puerto de Paita



*Fuente.* Adaptado de Google Maps.

Para la selección de los ejemplares, se aplicó muestreo por conveniencia; verificando el estado de frescura mediante inspección organoléptica, según los atributos para evaluación sensorial de cefalópodos frescos o refrigerados publicados por la Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO, 1999). Ver tabla 3.

**Tabla 3***Atributos para la evaluación sensorial de cefalópodos frescos o refrigerados*

<b>Atributo sensorial</b>		<b>Descripción</b>
Color	Piel	<ul style="list-style-type: none"> <li>• Brillante</li> <li>• Apagado</li> <li>• Blanquecino</li> </ul>
	Carne	<ul style="list-style-type: none"> <li>• Blanco nacarado</li> <li>• Blanco calizo</li> <li>• Rosáceo</li> <li>• Amarillo claro</li> </ul>
Adherencia	Piel	<ul style="list-style-type: none"> <li>• Adherente a la carne</li> <li>• Fácilmente separable de la carne</li> </ul>
	Carne	<ul style="list-style-type: none"> <li>• Muy firme</li> <li>• Firme</li> <li>• Ligeramente blanda</li> </ul>
Textura	Brazos/Tentáculos	<ul style="list-style-type: none"> <li>• Resistentes al desgarramiento</li> <li>• Fáciles de arrancar</li> </ul>
		<ul style="list-style-type: none"> <li>• Fresco</li> <li>• A algas</li> <li>• Ligero o ninguno</li> <li>• A tinta</li> </ul>

*Fuente.* Directrices para la evaluación sensorial del pescado y los mariscos en laboratorio (FAO, 1999)

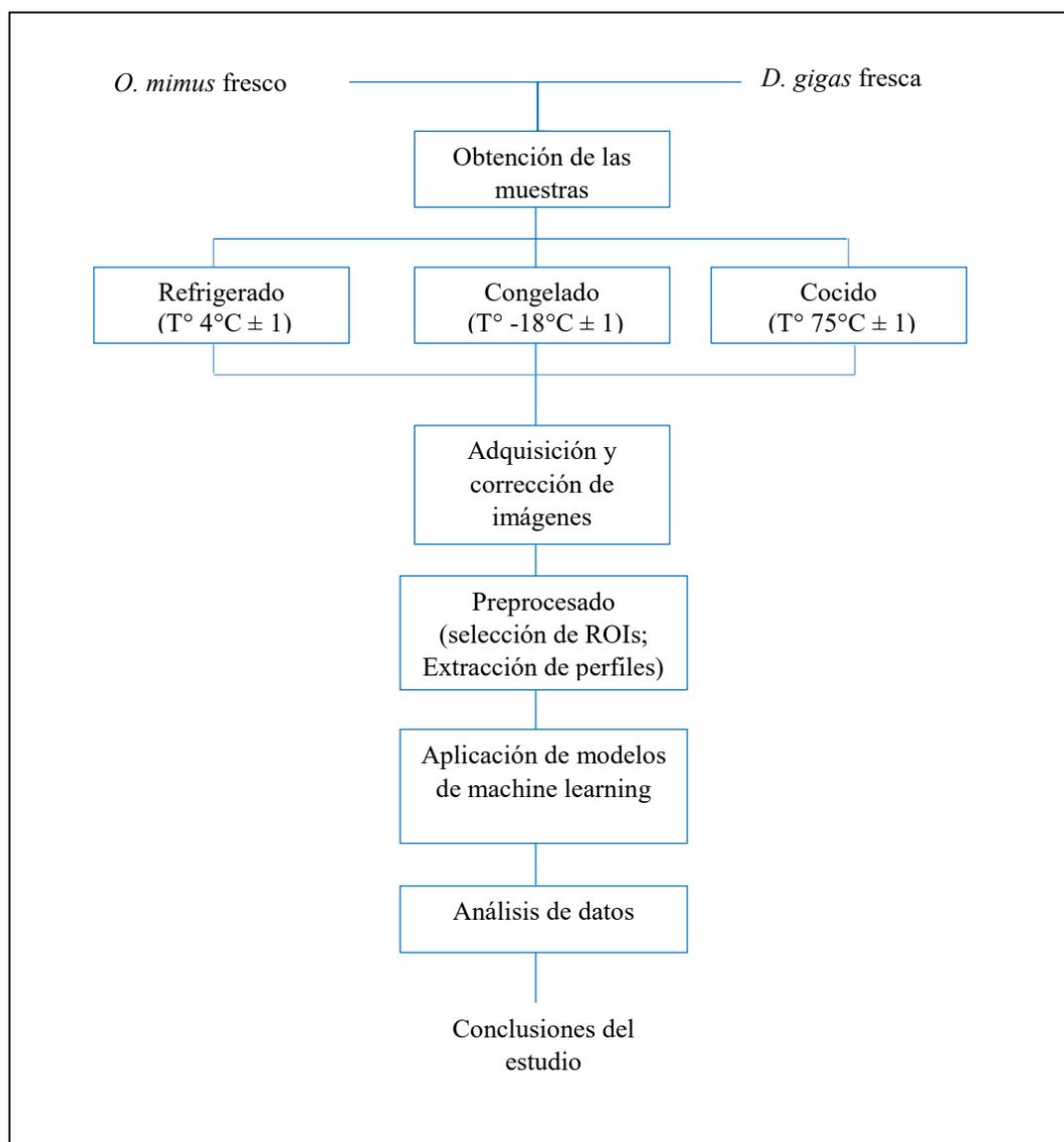
Asimismo, se verificó la talla mínima de captura, siendo 1 kg para *O. mimus* (Produce, 2019), mientras que para *D. gigas* se registró la longitud promedio al no existir regulación acerca de la talla mínima de captura. Las muestras se transportaron al laboratorio en cajas térmicas acondicionadas a una temperatura aproximada de 4°C, siendo procesadas dentro de las 24 horas siguientes.

## 2.2 Metodología

La metodología empleada en el presente trabajo, Figura 9, se basó en la secuencia para el análisis de hiperespectros, elaborados tomando como referencia la metodología aplicada por (Castro et al., 2019).

**Figura 9**

*Etapas del estudio*



*Fuente.* Adaptado de Castro et al., (2019).

### 2.3 Obtención de las muestras

Los ejemplares fueron transportados en caja térmica acondicionada con hielo a una temperatura de  $4^{\circ} \pm 1^{\circ}\text{C}$ ., una vez recepcionados en laboratorio se lavaron con agua potable, seleccionando los tentáculos, por ser la parte de mayor similitud entre ambas especies, luego se procedió a retirar la piel y se cortaron en trozos de 1 cm de longitud aproximadamente, lo que constituyó la unidad muestral, en total se emplearon 600 muestras; 300 por cada especie, a los cuales se les aplicó los siguientes tratamientos:

- Muestras frescas ( $n = 100$ ); se almacenaron a  $4 \pm 1^{\circ}\text{C}$  en refrigerador de convección (Daewoo FR-146RN, China).
- Muestras congeladas ( $n = 100$ ); se sometieron a temperatura de  $-25 \pm 1^{\circ}\text{C}$  por 18 horas, hasta que el centro del producto alcanzó los  $-18^{\circ}\text{C} \pm 1^{\circ}\text{C}$ , en congelador de laboratorio (Pol-Eko-Aparatura ZLN 85, Polonia).
- Muestras cocidas ( $n = 100$ ); se aplicó tratamiento térmico a una temperatura de  $100^{\circ}\text{C}$  por 8 minutos, hasta que el centro del producto alcanzó los  $75 \pm 1^{\circ}\text{C}$ , empleando cocina de inducción.

La temperatura interna de las muestras se midió con termómetro digital (Hanna Checktemp, USA), con rango de  $-50$  a  $150^{\circ}\text{C}$  y precisión de  $\pm 0.2^{\circ}\text{C}$ .

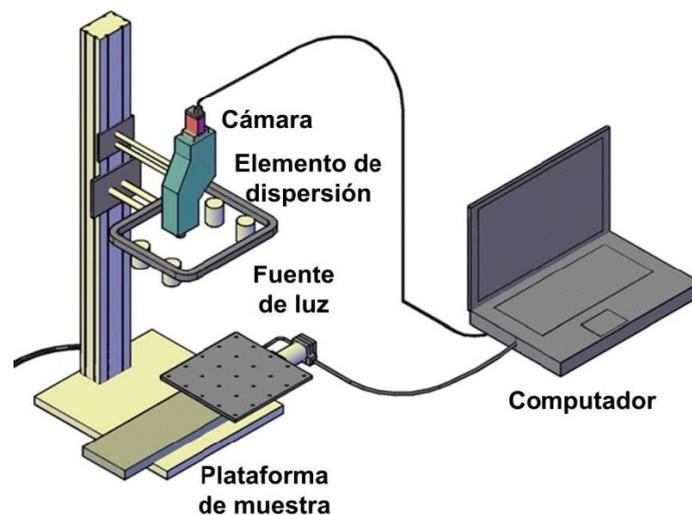
### 2.4 Sistema de imágenes hiperespectrales.

Se empleó un sistema HSI-VNIR por reflectancia, compuesto por un una cámara Hiperespectral con rango de  $400 - 1000$  nm. (Specim FX10e, Finlandia), seis lámparas de luz halógena dispuestas en un ángulo de  $45^{\circ}$ , una plataforma

portamuestra móvil de 40 x 21.4 cm y un ordenador con sistema operativo Windows. Los elementos del sistema HSI-VNIR se presentan en la figura 10.

**Figura 10**

*Sistema de adquisición de imágenes hiperespectrales*



*Fuente.* Tomado de (Castro, 2016)

La distancia entre la cámara y la muestra fue de 30 cm, se empleó el software Lumi Scanner 2019 para la adquisición de imágenes, con una velocidad de la plataforma 9 mm/s. Con resolución espacial de 2 mm/píxel y resolución espectral de 4 mm/píxel.

## 2.5 Preprocesamiento de las imágenes

Para obtener los datos de la reflectancia, fue preciso realizar la corrección de las imágenes hiperespectrales en bruto, aplicando la ecuación (1), de esta forma se descartaron los ruidos originados por la cámara y la intensidad de luz

innecesaria, generando imágenes con un nivel de reflectancia real. Las imágenes de referencia, blanco y oscuro, se obtuvieron usando como patrones; una cinta teflón estándar (reflectancia de  $\sim 99\%$ ) y cubierta del lente (reflectancia de  $\sim 0\%$ ), respectivamente, según lo descrito por (Dixit et al., 2021).

$$I_c = \frac{I_o - D}{W - D} \quad \text{Ecuación (1)}$$

Donde:

- $I_c$  : Imagen corregida
- $I_o$  : Imagen en bruto
- $D$  : Imagen de referencia oscura
- $W$  : Imagen de referencia blanca

A la imagen corregida se extrajo la región de interés (ROI) para la obtención de los espectros medios, mediante un Guide implementado en script de Matlab.

## 2.6 Aplicación de modelos de Machine Learning.

Se aplicaron modelos de aprendizaje supervisado, siendo necesario etiquetar los datos de entrada etiquetados a partir de los cuales se entrena un modelo de clasificación o predicción para el reconocimiento de patrones que permitan obtener las salidas deseadas. Para evaluar la idoneidad en la clasificación de los datos espectrales, se aplicaron tres algoritmos de aprendizaje automático supervisado, basados en; Análisis Discriminante, Máquinas de Vectores de Soporte y K – Vecinos Más Cercanos.

El análisis discriminante clasifica los datos en base a la distribución normal y la covarianza, partiendo del supuesto que las matrices de covarianzas son iguales

(Restaino et al., 2011). Este modelo genera combinaciones lineales, cuadráticas o cúbicas, dependiendo la distribución de los datos, para cada una de las variables dando lugar a las funciones discriminantes, buscando maximizar el cociente de la matriz interclases y la matriz entreclases. La Función Discriminante busca que las clases se separen y se reduzca la dimensionalidad, sirviendo de eje para la diferenciación de clases (Jesus et al., 2003).

Las Máquinas de Soporte Vectorial (SVM), son un grupo de métodos de aprendizaje automático supervisado, útiles para clasificación y regresión. Estos algoritmos generan un hiperplano o línea de decisión a partir de la cual se separan las clases, a su vez, manejan un margen entre los puntos más cercanos a las clases, delimitado por dos líneas denominadas vectores de soporte, conformados por los puntos de datos más próximos al hiperplano, mientras mayor sea el margen, mejor será la clasificación (Cano et al., 2017).

Los modelos basados en K- Vecinos Más Cercanos, son de tipo no paramétricos, basado en instancias, por tanto, requiere almacenar una gran cantidad de datos y un mayor tiempo de procesamiento. Para encontrar los puntos similares (vecinos) más cercanos, se calcula la distancia entre el punto a clasificar y el resto de puntos similares que se encuentra en el subconjunto de entrenamiento (Forina et al., 2009)

## **2.7 Análisis de los datos.**

Con cada modelo de machine learning aplicado se realizaron 30 repeticiones con validación cruzada ( $K\text{-fold} = 5$ ). Los indicadores de desempeño

(Accuracy y F2-measure) de cada modelo se evaluaron a partir de la matriz de confusión.

Una matriz de confusión se compone de dos dimensiones; clases reales y de predicción. Las filas contienen los datos de las clases reales y las columnas contienen los datos de las clases de predicción. La tabla 4 representa la estructura de una matriz de confusión para clasificación multiclase a partir de la matriz de la cual se calculan los indicadores de rendimiento del modelo, tales como; accuracy, precisión, recall, y f2-measure. Estas medidas se establecen por el número de aciertos y errores del modelo de clasificación (Castro et al., 2019).

**Tabla 4**

*Estructura general de la matriz de confusión y sus medidas de rendimiento*

		Clase de predicción $C_j$				
		$C_1$	...	$C_i$	...	$C_n$
Clase real $C_i$	$C_1$	TN <sub>i</sub>	TN <sub>i</sub>	FP <sub>i</sub>	TN <sub>i</sub>	TN <sub>i</sub>
	...	TN <sub>i</sub>	TN <sub>i</sub>	FP <sub>i</sub>	TN <sub>i</sub>	TN <sub>i</sub>
	$C_i$	FN <sub>i</sub>	FN <sub>i</sub>	TP <sub>i</sub>	FN <sub>i</sub>	FN <sub>i</sub>
	...	TN <sub>i</sub>	TN <sub>i</sub>	FP <sub>i</sub>	TN <sub>i</sub>	TN <sub>i</sub>
	$C_n$	TN <sub>i</sub>	TN <sub>i</sub>	FP <sub>i</sub>	TN <sub>i</sub>	TN <sub>i</sub>

*Fuente.* Tomado de Castro et al., (2019).

Las medidas de rendimiento se calcularon a partir de los siguientes elementos de la matriz de confusión:

- TP<sub>i</sub>** **Verdaderos positivos.** Número de muestras cuyo valor real es positivo con respecto a la clase “*i*” y el modelo las predice como positivas dentro de esta clase.
- TN<sub>i</sub>** **Verdaderos negativos.** Número de muestras cuyo valor real es negativo con respecto a la clase “*i*” y el modelo las predice fuera de esta clase.
- FP<sub>i</sub>** **Falsos positivos.** Número de muestras cuyo valor real es negativo con respecto a la clase “*i*” y el modelo las predice como positivas dentro de esta clase.

Los cuatro elementos de la matriz de confusión descritos anteriormente permiten calcular los indicadores de rendimiento de los modelos de clasificación; exactitud (accuracy), precisión (precision), sensibilidad (recall), especificidad (f2-measure).

**Exactitud o accuracy;** es la proporción del total de predicciones realizadas correctamente por el modelo. Resulta del cociente de las predicciones verdaderas totales entre el total de observaciones, ver ecuación 2.

$$Exactitud = \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \quad \text{Ecuación (2)}$$

**Precisión o precisión;** es la proporción del total de predicciones positivas realizadas correctamente por el modelo. Resulta del cociente de las predicciones verdaderas positivas entre todas las predicciones positivas, ver ecuación 3.

$$\textit{Precisión} = \frac{TP_i}{TP_i + TN_i} \quad \text{Ecuación (3)}$$

**Sensibilidad o recall;** es la proporción de valores reales positivos que el modelo predijo correctamente. Resulta del cociente de las predicciones verdaderas positivas entre el total de los valores reales positivos, ver ecuación 4.

$$\textit{Sensibilidad} = \frac{TP_i}{TP_i + FN_i} \quad \text{Ecuación (4)}$$

**Especificidad o F2-Measure;** es la proporción de valores reales negativos que el modelo predijo correctamente. Resulta del cociente de las predicciones verdaderas negativas entre el total de los valores reales negativos, ver ecuación 5.

$$\textit{Especificidad} = \frac{TN_i}{TN_i + FP_i} \quad \text{Ecuación (5)}$$

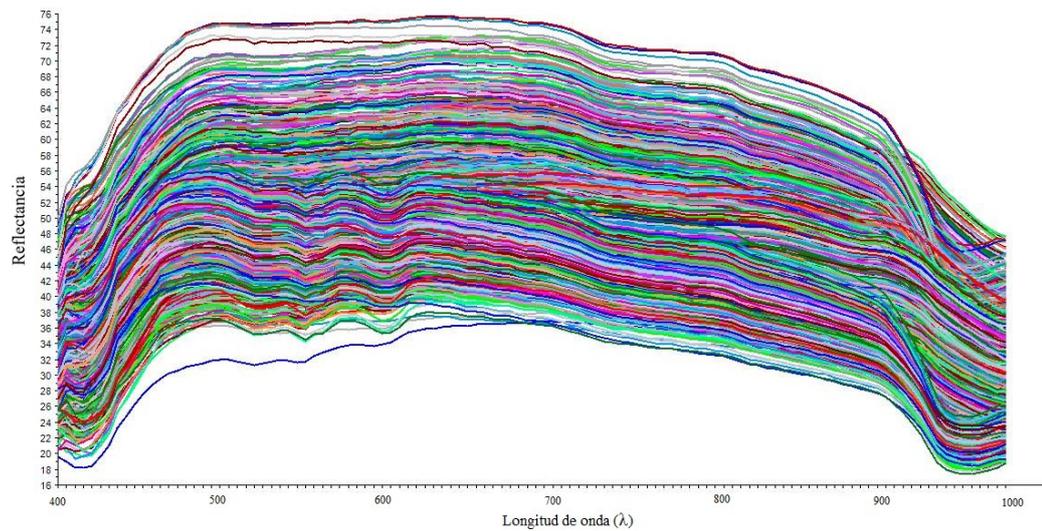
### 3. RESULTADOS

#### 3.1 Perfiles espectrales de la carne de *O. mimus* y *D. gigas*.

En la figura 12, se observan los espectros completos de las muestras de *O. mimus* y *D. gigas*, en estado; fresco, congelado y cocido, se aprecia una gran cantidad de bandas superpuestas a las que posteriormente se extrajeron los espectros medios.

**Figura 11**

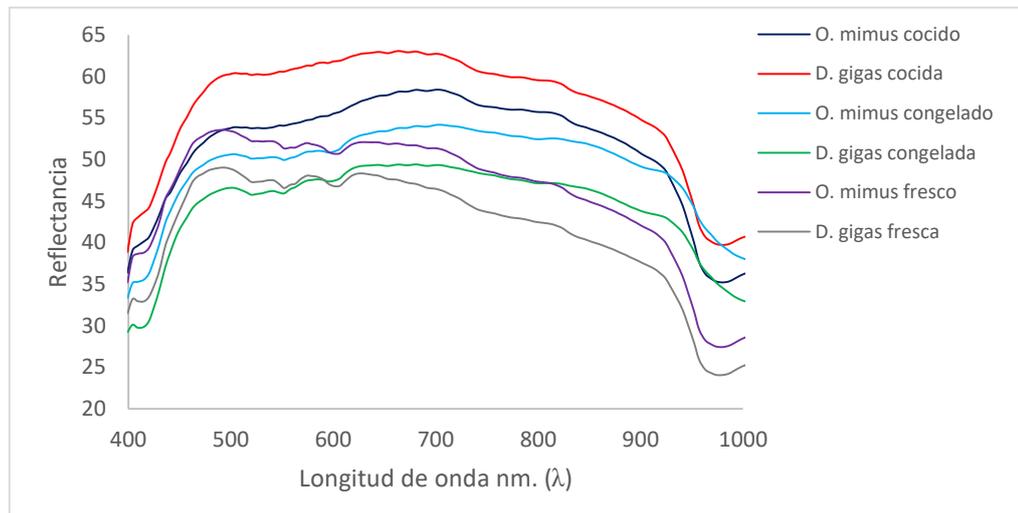
*Espectros completos de las muestra de O. mimus y D. gigas en estado fresco, congelado y cocido*



En la Figura 12, se presentan los espectros medios. Las muestras de *O. mimus* en estado fresco y congelado tuvieron valores de reflectancia mayores a los de *D. gigas*. Sin embargo, en las muestras sometidas a cocción, se aprecia que *D. gigas* obtuvo niveles de reflectancia más altos que *O. mimus*. En cada uno de los tratamientos existe una diferencia notoria entre los valores de absorbancia, no obstante, se evidencia un patrón de reflectancia similar.

**Figura 12**

*Espectros medios de las muestras de O. mimus y D. gigas*



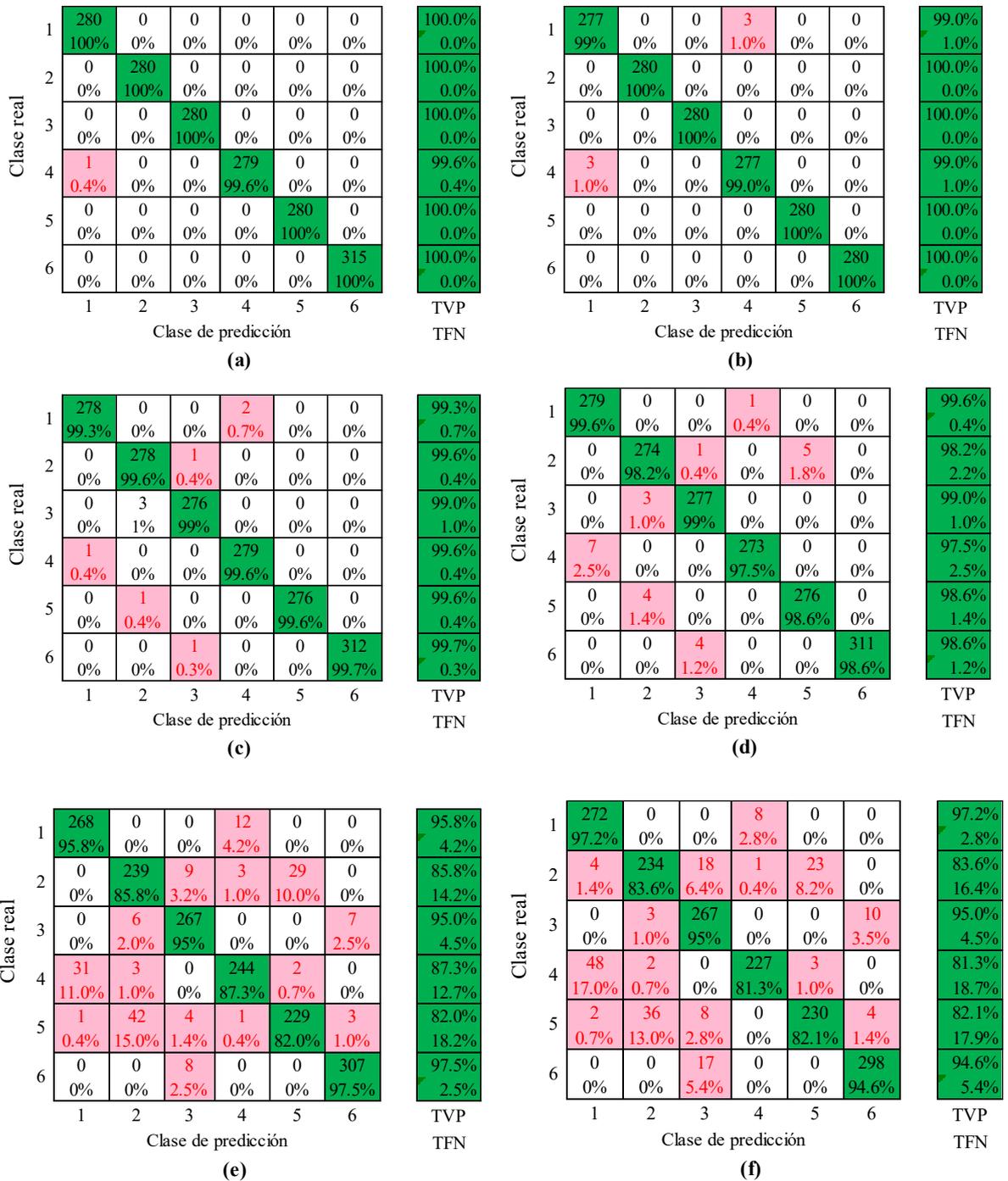
### 3.2 Rendimiento de los modelos de clasificación de machine learning.

En la figura 14, se presentan las matrices de confusión de los modelos aplicados sobre los cuales se calcularon los indicadores estadísticos de Accuracy y F2-Measure.

En la tabla 5, podemos observar los indicadores estadísticos de los modelos de machine learning usados; discriminante lineal, discriminante cuadrático, Máquina de Soporte Vectorial Cuadrático, Máquinas de Soporte Vectorial Cúbico, Vecinos Más Cercano Fino, Vecino Más Cercano Ponderado. El modelo Discriminante Lineal mostró el mejor desempeño con respecto a los modelos basados en SVM y KNN, obteniendo una precisión general de 99.9% y un F2-Measure igual a 99.94%. Todos los modelos probados, excepto KNN Ponderado, permitieron obtener un Accuracy (precisión) y F2-Measure mayor al 90%.

Figura 13

Matrices de confusión. (a-b) Discriminante; (b-c) SVM; y (d-e) KNN



**Tabla 5***Estadísticos de precisión (accuracy) y medida f2 (f2-measure)*

Modelo	Accuracy		F2-Measure	
	%	$\sigma$	%	$\Sigma$
Discriminante lineal	99.90	0.00	99.94	0.08
Discriminante cuadrático	99.70	0.00	99.64	0.55
SVM Cuadrático	99.10	0.07	99.06	0.52
SVM Cubico	98.50	0.11	98.53	0.54
KNN fino	90.60	0.00	90.44	5.23
KNN Ponderado	89.10	0.00	88.91	3.81

#### 4. DISCUSIÓN

En el presente trabajo de investigación, se obtuvo una alta tasa de discriminación usando los datos espectrales completos obtenidos por un sistema HSI de reflectancia en el rango de 400-1000 nm, el mayor rendimiento fue de 99.9% para el modelo Discriminante Lineal, parecido a lo reportado por Sun et al., (2019), quienes utilizaron el sistema HSI-NIR en el rango de 874-1734 nm y tras una selección de variables óptimas, obtuvieron 99.2% de precisión aplicando selección directa secuencial combinada con análisis discriminante de mínimos cuadrados parciales (SFS-PLS-DA) para clasificar músculos de camarón según su sistema de crianza.

El modelo SVM también mostró un rendimiento adecuado en la discriminación de la carne de *O. mimus* y *D. gigas*, con una precisión total de 99.1 %, similar a

lo obtenido por Xu et al., (2017), quienes reportaron una tasa de clasificación correcta de 98.2% con el modelo SVM, empleando el mismo sistema de adquisición de imágenes hiperespectrales, dicho modelo mostró un alto rendimiento con los datos espectrales completos y con longitudes de onda óptimas.

## 5. CONCLUSIONES

Se evidenció que las imágenes hiperespectrales combinadas con la técnica de machine learning tienen gran capacidad para discriminar la carne de *Octopus mimus* y *Dosidicus gigas*. Los resultados obtenidos en la presente investigación muestran que el Análisis Discriminante Lineal (LDA) logró clasificar la carne de ambas especies, en estado fresco, congelado y cocido, con una precisión mayor al 99%. Por tanto, existe evidencia suficiente para concluir que existe efecto de la técnica de machine learning en la discriminación de la carne de pulpo (*O. mimus*) y calamar gigante (*D. gigas*) usando perfiles de imágenes hiperespectrales.

## 6. RECOMENDACIONES

- Identificar otras especies de pescados y mariscos, susceptibles de fraude para evaluar el potencial de discriminación de las imágenes hiperespectrales combinadas con la técnica de Machine Learning.
- Investigar la factibilidad de implementar un sistema de monitoreo in situ para detección de fraudes en pescados y mariscos, a través de la selección de características relevantes, así como el diseño y desarrollo de dispositivos portátiles.

## 7. REFERENCIAS BIBLIOGRÁFICAS

- Acutis, P. L., Cambiotti, V., Riina, M. V., Meistro, S., Maurella, C., Massaro, M., Stacchini, P., Gili, S., Malandra, R., Pezzolato, M., Caramelli, M., & Bozzetta, E. (2019). Detection of fish species substitution frauds in Italy: A targeted National Monitoring Plan. *Food Control*, *101*, 151-155. <https://doi.org/10.1016/j.foodcont.2019.02.020>
- Aguirre-Urreta, M., & Cichowolski, M. (2007). *Cephalopoda* (pp. 441-476).
- Amigo, J. M. (2020). Hyperspectral and multispectral imaging: Setting the scene. *Data Handling in Science and Technology*, *32*, 3-16. <https://doi.org/10.1016/B978-0-444-63977-6.00001-8>
- Blanco-Fernandez, C., Ardura, A., Masiá, P., Rodriguez, N., Voces, L., Fernandez-Raigoso, M., Roca, A., Machado-Schiaffino, G., Dopico, E., & Garcia-Vazquez, E. (2021). Fraud in highly appreciated fish detected from DNA in Europe may undermine the Development Goal of sustainable fishing in Africa. *Scientific Reports 2021 11:1*, *11*(1), 1-10. <https://doi.org/10.1038/s41598-021-91020-w>
- Boyle, P., & Rodhouse, P. (2007). Cephalopods: Ecology and Fisheries. En *Cephalopods: Ecology and Fisheries*. Blackwell Publishing Ltd. <https://doi.org/10.1002/9780470995310>
- Brooks, C., Parr, L., Smith, J. M., Buchanan, D., Snioch, D., & Hebishy, E. (2021). A Review of Food Fraud and Food Authenticity across the Food Supply Chain, with an Examination of the Impact of the COVID-19 Pandemic and Brexit on Food Industry. *Food Control*, 108171. <https://doi.org/10.1016/j.foodcont.2021.108171>
- Cano, G., García-Rodríguez, J., Orts, S., García-García, A., Peña-García, J., Pérez-Garrido, A., & Pérez-Sánchez, H. (2017). Predicción de solubilidad de fármacos usando máquinas de soporte vectorial sobre unidades de procesamiento gráfico.

- Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería*, 33(1), 97-102. <https://doi.org/10.1016/j.rimni.2015.12.001>
- Castro, W. (2016). *Aplicación de la tecnología de imágenes hiperespectrales al control de calidad de productos agroalimentarios de la región de Amazonas (Perú)* [Tesis doctoral, Universitat Politècnica de València]. <https://riunet.upv.es/handle/10251/63250>
- Castro, W., Oblitas, J., De-La-Torre, M., Cotrina, C., Bazán, K., & Avila-George, H. (2019). Classification of Cape Gooseberry Fruit According to its Level of Ripeness Using Machine Learning Techniques and Different Color Spaces. *IEEE Access*, 7, 27389-27400. <https://doi.org/10.1109/ACCESS.2019.2898223>
- Deconinck, D., Volckaert, F. A. M., Hostens, K., Panicz, R., Eljasik, P., Faria, M., Monteiro, C. S., Robbens, J., & Derycke, S. (2020). A high-quality genetic reference database for European commercial fishes reveals substitution fraud of processed Atlantic cod (*Gadus morhua*) and common sole (*Solea solea*) at different steps in the Belgian supply chain. *Food and Chemical Toxicology*, 141, 111417. <https://doi.org/10.1016/J.FCT.2020.111417>
- Dixit, Y., Al-Sarayreh, M., Craigie, C. R., & Reis, M. M. (2021). A global calibration model for prediction of intramuscular fat and pH in red meat using hyperspectral imaging. *Meat Science*, 108405. <https://doi.org/10.1016/j.meatsci.2020.108405>
- Doubleday, Z. A., Prowse, T. A. A., Arkhipkin, A., Pierce, G. J., Semmens, J., Steer, M., Leporati, S. C., Lourenço, S., Quetglas, A., Sauer, W., & Gillanders, B. M. (2016). Global proliferation of cephalopods. *Current Biology*, 26(10), R406-R407. <https://doi.org/10.1016/J.CUB.2016.04.002>

- Espiñeira, M., & Lago, F. (2016). Advances in Authenticity Testing for Fish Speciation. *Advances in Food Authenticity Testing*, 415-440. <https://doi.org/10.1016/B978-0-08-100220-9.00015-1>
- Espiñeira, M., & Vieites, J. M. (2012). Rapid method for controlling the correct labeling of products containing common octopus (*Octopus vulgaris*) and main substitute species (*Eledone cirrhosa* and *Dosidicus gigas*) by fast real-time PCR. *Food Chemistry*, 135(4), 2439-2444. <https://doi.org/10.1016/j.foodchem.2012.07.056>
- Comisión Europea. (2018). *Fraude agroalimentario*. Fraude Agroalimentario. [https://ec.europa.eu/food/safety/agri-food-fraud\\_en](https://ec.europa.eu/food/safety/agri-food-fraud_en)
- FAO. (1999). 8. *EVALUACION DE LA CALIDAD DEL PESCADO*. <http://www.fao.org/3/v7180s/v7180s09.htm>
- Feng, L., Zhu, S., Liu, F., He, Y., Bao, Y., & Zhang, C. (2019). Hyperspectral imaging for seed quality and safety inspection: A review. *Plant Methods* 2019 15:1, 15(1), 1-25. <https://doi.org/10.1186/S13007-019-0476-Y>
- Forina, M., Casale, M., & Oliveri, P. (2009). 4.04—Application of Chemometrics to Food Chemistry. En S. D. Brown, R. Tauler, & B. Walczak (Eds.), *Comprehensive Chemometrics* (pp. 75-128). Elsevier. <https://doi.org/10.1016/B978-044452701-1.00124-1>
- Gonzalez-Pestana, A., Mangel, J. C., Alfaro-Córdova, E., Acuña-Perales, N., Córdova-Zavaleta, F., Segura-Cobeña, E., Benites, D., Espinoza, M., Coasaca-Céspedes, J., Jiménez, A., Pingo, S., Moscoso, V., Alfaro-Shigueto, J., & Espinoza, P. (2021). Diet, trophic interactions and possible ecological role of commercial sharks and batoids in northern Peruvian waters. *Journal of Fish Biology*, 98(3), 768-783. <https://doi.org/10.1111/jfb.14624>

- Grassi, S., Casiraghi, E., & Alamprese, C. (2018). Fish fillet authentication by image analysis. *Journal of Food Engineering*, 234, 16-23. <https://doi.org/10.1016/j.jfoodeng.2018.04.012>
- Guardone, L., Tinacci, L., Costanzo, F., Azzarelli, D., D'Amico, P., Tasselli, G., Magni, A., Guidi, A., Nucera, D., & Armani, A. (2017). DNA barcoding as a tool for detecting mislabeling of fishery products imported from third countries: An official survey conducted at the Border Inspection Post of Livorno-Pisa (Italy). *Food Control*, 80, 204-216. <https://doi.org/10.1016/J.FOODCONT.2017.03.056>
- Horreo, J. L., Fitze, P. S., Jiménez-Valverde, A., Noriega, J. A., & Pelaez, M. L. (2019). Amplification of 16S rDNA reveals important fish mislabeling in Madrid restaurants. *Food Control*, 96, 146-150. <https://doi.org/10.1016/J.FOODCONT.2018.09.020>
- Hu, M. H., Dong, Q. L., & Liu, B. L. (2016). Classification and characterization of blueberry mechanical damage with time evolution using reflectance, transmittance and interactance imaging spectroscopy. *Computers and Electronics in Agriculture*, 122, 19-28. <https://doi.org/10.1016/J.COMPAG.2016.01.015>
- Hu, M. H., Dong, Q. L., Liu, B. L., & Opara, U. L. (2016). Prediction of mechanical properties of blueberry using hyperspectral interactance imaging. *Postharvest Biology and Technology*, 115, 122-131. <https://doi.org/10.1016/J.POSTHARVBIO.2015.11.021>
- Jesús, M., Alcázar, A., Pablos, F., & Neves, G. (2003). Diferenciación del origen geográfico de cervezas mediante técnicas de reconocimiento de patrones según su contenido mineral. *Cerveza y malta*, ISSN 0300-4481, N<sup>o</sup>. 157, 2003, pags. 21-25, 40.

- Jian, W., Mengting, Q., Wei, Y., Xinjun, C., & Bilin, L. (2020). Habitat Variations of the Jumbo Flying Squid *Dosidicus Gigas* Off Peru Under El Niño And La Niña Events with Different Intensities. *Journal of Fishery Sciences of China*, 27(10), 1095-1103. <https://doi.org/10.3724/SP.J.1118.2020.20054>
- Kamruzzaman, M., Makino, Y., & Oshita, S. (2015). Non-invasive analytical technology for the detection of contamination, adulteration, and authenticity of meat, poultry, and fish: A review. *Analytica Chimica Acta*, 853(1), 19-29. <https://doi.org/10.1016/J.ACA.2014.08.043>
- Kamruzzaman, M., & Sun, D. W. (2016). Introduction to Hyperspectral Imaging Technology. *Computer Vision Technology for Food Quality Evaluation: Second Edition*, 111-139. <https://doi.org/10.1016/B978-0-12-802232-0.00005-0>
- Khaled, A. Y., Parrish, C. A., & Adedeji, A. (2021). Emerging nondestructive approaches for meat quality and safety evaluation—A review. *Comprehensive Reviews in Food Science and Food Safety*, 20(4), 3438-3463. <https://doi.org/10.1111/1541-4337.12781>
- Kotsanopoulos, K. V., Exadactylos, A., Gkafas, G. A., Martsikalis, P. V., Parlapani, F. F., Boziaris, I. S., & Arvanitoyannis, I. S. (2021). The use of molecular markers in the verification of fish and seafood authenticity and the detection of adulteration. *Comprehensive Reviews in Food Science and Food Safety*, 20(2), 1584-1654. <https://doi.org/10.1111/1541-4337.12719>
- Lu, Y., Saeys, W., Kim, M., Peng, Y., & Lu, R. (2020). Hyperspectral imaging technology for quality and safety evaluation of horticultural products: A review and celebration of the past 20-year progress. *Postharvest Biology and Technology*, 170, 111318. <https://doi.org/10.1016/J.POSTHARVBIO.2020.111318>

- Maraphum, K., Saengprachatanarug, K., Aparatana, K., Izumikawa, Y., & Taira, E. (2020). Mapeo espacial de Brix y contenido de humedad en el tallo de la caña de azúcar utilizando imágenes hiperespectrales. *Journal of Near Infrared Spectroscopy*, 28(4), 167-174. <https://doi.org/10.1177/0967033520905370>
- Markaida, U., & Gilly, W. F. (2016a). Cephalopods of Pacific Latin America. *Fisheries Research*, 173, 113-121. <https://doi.org/10.1016/j.fishres.2015.09.014>
- Markaida, U., & Gilly, W. F. (2016b). Cephalopods of Pacific Latin America. *Fisheries Research*, 173, 113-121. <https://doi.org/10.1016/J.FISHRES.2015.09.014>
- Mouritsen, O. G., & Styrbæk, K. (2018). Cephalopod Gastronomy—A Promise for the Future. *Frontiers in Communication*, 3, 38. <https://doi.org/10.3389/fcomm.2018.00038>
- Mouritsen, O. G., & Styrbæk, K. (2021). Buying, Preparing, and Storing Cephalopods. *Octopuses, Squid & Cuttlefish*, 93-114. [https://doi.org/10.1007/978-3-030-58027-8\\_7](https://doi.org/10.1007/978-3-030-58027-8_7)
- Munekata, P. E. S., Domínguez, R., Pateiro, M., & Lorenzo, J. M. (2021). Introduction to food fraud. *Food Toxicology and Forensics*, 1-30. <https://doi.org/10.1016/B978-0-12-822360-4.00002-9>
- Nehal, N., Choudhary, B., Nagpure, A., & Gupta, R. K. (2021). DNA barcoding: A modern age tool for detection of adulteration in food. <https://doi.org/10.1080/07388551.2021.1874279>, 41(5), 767-791. <https://doi.org/10.1080/07388551.2021.1874279>
- Pardo, M. Á., & Jiménez, E. (2020). DNA barcoding revealing seafood mislabeling in food services from Spain. *Journal of Food Composition and Analysis*, 91, 103521. <https://doi.org/10.1016/J.JFCA.2020.103521>

- Pierce, G. J., Portela, J., Pierce, G. J., & Portela, J. (2014). Fisheries Production and Market Demand. *Cephalopod Culture*, 41-58. [https://doi.org/10.1007/978-94-017-8648-5\\_3](https://doi.org/10.1007/978-94-017-8648-5_3)
- Produce. (2019). *Pesca y consumo responsable*. Tabla de medidas mínimas de pescados, crustáceos y moluscos. <https://pescayconsumoresponsable.produce.gob.pe/tallas-minimas.html>
- PROMPERÚ. (2020). *Desarrollo del comercio exterior pesquero y acuícola en el Perú. Informe anual 2019* (pp. 1-83).
- Qin, J., Kim, M. S., Chao, K., Chan, D. E., Delwiche, S. R., & Cho, B.-K. (2017). Line-Scan Hyperspectral Imaging Techniques for Food Safety and Quality Applications. *Applied Sciences* 2017, Vol. 7, Page 125, 7(2), 125. <https://doi.org/10.3390/APP7020125>
- Restaino, E., Fassio, A., & Cozzolino, D. (2011). Discrimination of meat patés according to the animal species by means of near infrared spectroscopy and chemometrics. Discriminación de muestras de paté de carne según tipo de especie mediante el uso de la espectroscopia en el infrarrojo cercano y la quimiometria. *CyTA - Journal of Food*, 9(3), 210-213. <https://doi.org/10.1080/19476337.2010.512396>
- Rovegno, N. (2021). *Pota*. En: *Sociedad Peruana de Derecho Ambiental, Wikipesca Perú. Plataforma colaborativa sobre la pesca en el Perú*. <https://www.mardelperu.pe/pesca/13/pesqueria-pota>
- Ru, C., Li, Z., & Tang, R. (2019). A Hyperspectral Imaging Approach for Classifying Geographical Origins of *Rhizoma Atractylodis Macrocephalae* Using the Fusion of Spectrum-Image in VNIR and SWIR Ranges (VNIR-SWIR-FuSI). *Sensors* 2019, Vol. 19, Page 2045, 19(9), 2045. <https://doi.org/10.3390/S19092045>

- Sauer, W. H. H., Gleadall, I. G., Downey-Breedt, N., Doubleday, Z., Gillespie, G., Haimovici, M., Ibáñez, C. M., Katugin, O. N., Leporati, S., Lipinski, M. R., Markaida, U., Ramos, J. E., Rosa, R., Villanueva, R., Arguelles, J., Briceño, F. A., Carrasco, S. A., Che, L. J., Chen, C.-S., ... Pecl, G. (2019). World Octopus Fisheries. *https://doi.org/10.1080/23308249.2019.1680603*, 29(3), 279-429. <https://doi.org/10.1080/23308249.2019.1680603>
- Shao, S., Xiaojie, Z., Lingfeng, K., & Qi, L. (2021). Molecular Identification of Dried Shellfish Products Sold on the Market Using DNA Barcoding. *J. Ocean Univ. China (Oceanic and Coastal Sea Research)*, 20(4), 931-938. <https://doi.org/10.1007/s11802-021-4682-7>
- SIICEX. (2021a). *Ficha Comercial*. [https://www.siicex.gob.pe/siicex/portal5ES.asp?\\_page\\_=172.17100&\\_portletid\\_=sfichaproductoinit&scriptdo=cc\\_fp\\_init&pproducto=%20177%20&pnomproducto=%20Pulpo%20Congelado](https://www.siicex.gob.pe/siicex/portal5ES.asp?_page_=172.17100&_portletid_=sfichaproductoinit&scriptdo=cc_fp_init&pproducto=%20177%20&pnomproducto=%20Pulpo%20Congelado)
- SIICEX. (2021b). *Productos Pesqueros*. [https://www.siicex.gob.pe/siicex/portal5ES.asp?\\_page\\_=402.39500](https://www.siicex.gob.pe/siicex/portal5ES.asp?_page_=402.39500)
- Spink, J., Chen, W., Zhang, G., & Speier-Pero, C. (2019). Introducing the Food Fraud Prevention Cycle (FFPC): A dynamic information management and strategic roadmap. *Food Control*, 105, 233-241. <https://doi.org/10.1016/J.FOODCONT.2019.06.002>
- Suman, M., Cavanna, D., Sammarco, G., Lambertini, F., & Loffi, C. (2021). Fighting food frauds exploiting chromatography-mass spectrometry technologies: Scenario comparison between solutions in scientific literature and real approaches in place in industrial facilities. *TrAC Trends in Analytical Chemistry*, 142, 116305. <https://doi.org/10.1016/J.TRAC.2021.116305>

- Sun, D., Weng, H., He, X., Li, L., He, Y., & Cen, H. (2019). Combining near-infrared hyperspectral imaging with elemental and isotopic analysis to discriminate farm-raised pacific white shrimp from high-salinity and low-salinity environments. *Food Chemistry*, 299, 125121. <https://doi.org/10.1016/j.foodchem.2019.125121>
- Suthar, G., Huang, J. Y., & Chidangil, S. (2017). Optimisation and evaluation of hyperspectral imaging system using machine learning algorithm. <https://doi.org/10.1117/12.2296863>, 10438, 121-127. <https://doi.org/10.1117/12.2296863>
- Velasco, A., Ramilo-Fernández, G., Denis, F., Oliveira, L., Shum, P., Silva, H., & Sotelo, C. G. (2021). A New Rapid Method for the Authentication of Common Octopus (*Octopus vulgaris*) in Seafood Products Using Recombinase Polymerase Amplification (RPA) and Lateral Flow Assay (LFA). *Foods* 2021, Vol. 10, Page 1825, 10(8), 1825. <https://doi.org/10.3390/FOODS10081825>
- Visciano, P., & Schirone, M. (2021). Food frauds: Global incidents and misleading situations. *Trends in Food Science & Technology*, 114, 424-442. <https://doi.org/10.1016/J.TIFS.2021.06.010>
- Wang, N., Xing, R.-R., Zhou, M.-Y., Sun, R.-X., Han, J.-X., Zhang, J.-K., Zheng, W.-J., & Chen, Y. (2021). Application of DNA barcoding and metabarcoding for species identification in salmon products. <https://doi.org/10.1080/19440049.2020.1869324>, 38(5), 754-768. <https://doi.org/10.1080/19440049.2020.1869324>
- Warnke, K., Söller, R., Blohm, D., & Saint-Paul, U. (2000). Rapid differentiation between *Octopus vulgaris* Cuvier (1797) and *Octopus mimus* Gould (1852), using randomly amplified polymorphic DNA. *Journal of Zoological Systematics and*

*Evolutionary Research*, 38(2), 119-122. <https://doi.org/10.1046/j.1439-0469.2000.382134.x>

Xu, J. L., Riccioli, C., & Sun, D. W. (2017). Comparison of hyperspectral imaging and computer vision for automatic differentiation of organically and conventionally farmed salmon. *Journal of Food Engineering*, 196, 170-182. <https://doi.org/10.1016/j.jfoodeng.2016.10.021>

Zúñiga, Ó., Olivares, A., & Rosas, C. (2014). *Octopus mimus*. *Cephalopod Culture*, 397-413. [https://doi.org/10.1007/978-94-017-8648-5\\_21](https://doi.org/10.1007/978-94-017-8648-5_21)

**ANEXO 1. Glosario**

ADN : Ácido Desoxirribonucleico

FAO : Organización de las Naciones Unidas para la Agricultura y la Alimentación

FOB : “Free on board”, término de comercio internacional que implica entrega de la mercancía sobre el buque, en el puerto de origen

HSI : Imágenes hiperespectrales

LDA : Análisis discriminante lineal

KNN : Vecinos más cercanos

SVM : Máquina de soporte vectorial

VNIR : Área visible e infrarroja cercana del espectro electromagnético

## ANEXO 2. Evidencias fotográficas

**Figura 14**

*Caja térmica acondicionada con hielo para el transporte de muestras frescas*



**Figura 145**

*Lavado de muestras con agua potable*



**Figura 166**

*Pelado y cortado de muestras*



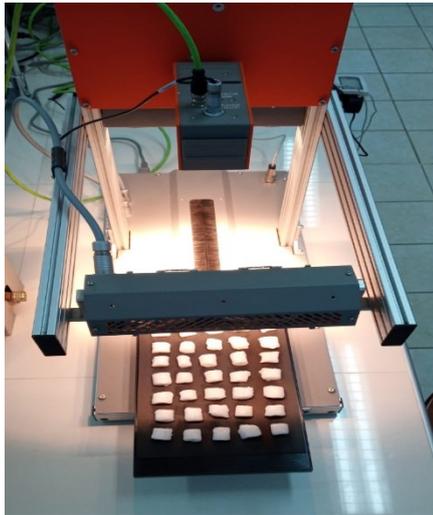
**Figura 157**

*Muestras listas para adquisición de imágenes*



**Figura 188**

*Proceso de adquisición de imágenes hiperespectrales*



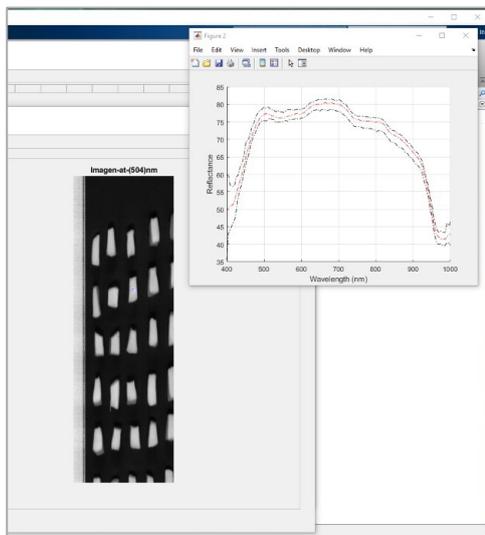
**Figura 179**

*Escaneo comandado desde el software Lumi*



**Figura 20**

*Obtención de perfiles espectrales usando el software Matlab*



**Figura 19**

*Matriz de confusión del modelo LDA, con mejor rendimiento*

