



# **ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO**

## **PROCESAMIENTO DIGITAL DE IMÁGENES EN DETECCIÓN DE MATERIAL CONTAMINANTE INORGÁNICO EN ORILLAS DE LAGUNAS**

**MARÍA VICTORIA URGILEZ JARAMILLO**

**Trabajo de Titulación modalidad: Proyecto de Investigación y Desarrollo,  
presentado ante el Instituto de Posgrado y Educación Continua de la ESPOCH,  
como requisito parcial para la obtención del grado de:**

**MAGÍSTER EN SISTEMAS DE TELECOMUNICACIONES**

**Riobamba – Ecuador**

**Enero 2022**

©2022, María Victoria Urgilez Jaramillo

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo cita bibliográfica del documento, siempre y cuando se reconozca el Derechos de Autor.

## ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO

### CERTIFICACIÓN

EL TRIBUNAL DEL TRABAJO DE TITULACIÓN CERTIFICA QUE:

El Trabajo de Titulación modalidad Proyectos de Investigación y Desarrollo, denominado: **PROCESAMIENTO DIGITAL DE IMÁGENES EN DETECCIÓN DE MATERIAL CONTAMINANTE INORGÁNICO EN ORILLAS DE LAGUNA**, de responsabilidad de la señorita María Victoria Urgilez Jaramillo, ha sido minuciosamente revisado y se autoriza su presentación.

Tribunal:

Ing. Oswaldo Geovanny Martínez Guashima M.Sc

PRESIDENTE

\_\_\_\_\_

FIRMA

Ing. José Luis Jinez Tapia Mag.

DIRECTOR

\_\_\_\_\_

FIRMA

Ing. Vanessa Alexandra Vasconez Nuñez MEng.

MIEMBRO DEL TRIBUNAL

\_\_\_\_\_

FIRMA

Ing. José Luis Morales Gordon Mag.

MIEMBRO DEL TRIBUNAL

\_\_\_\_\_

FIRMA

**Riobamba, enero 2022**

## DERECHOS INTELECTUALES

Yo, María Victoria Urgilez Jaramillo, declaro que soy responsable de las ideas, doctrinas y resultados expuestos en el **Trabajo de Titulación modalidad Proyectos de Investigación y Desarrollo**, y que el patrimonio intelectual generado por la misma pertenece exclusivamente a la Escuela Superior Politécnica de Chimborazo.

-----  
María Victoria Urgilez Jaramillo  
C.I. 2100253844

Yo, María Victoria Urgilez Jaramillo, declaro que el presente trabajo de titulación es de mi autoría y los resultados del mismo son auténticos. Los textos en el documento que provienen de otras fuentes están debidamente citados y referenciados.

Como autor, asumo la responsabilidad legal y académica de los contenidos de este Trabajo de Titulación; el patrimonio intelectual pertenece a la Escuela Superior Politécnica de Chimborazo.

-----  
María Victoria Urgilez Jaramillo

C.I. 2100253844

## **DEDICATORIA**

El presente trabajo de titulación está dedicado a mis hijas y esposo que con su paciencia, amor y apoyo me han ayudado a terminar esta meta en mi vida y a todas aquellas personas que con sus palabras de aliento me ayudaron a seguir adelante.

## **AGRADECIMIENTO**

Me gustaría agradecer de todo corazón a mi Tutor y miembros del Tribunal, así como a mis profesores de la Maestría por su apoyo y enseñanzas que me permitieron plasmar en este trabajo de investigación mis ideas y estudios adquiridos.

A mi familia por su apoyo y comprensión.

María Victoria

## TABLA DE CONTENIDO

RESUMEN.....	xiv
ABSTRACT .....	xv
INTRODUCCIÓN .....	1

### CAPÍTULO I

1	PLANTEAMIENTO DEL PROBLEMA.....	4
1.1	Situación problemática .....	4
1.2	Formulación del problema .....	5
1.3	Preguntas directrices o específicas de la investigación .....	5
1.4	Justificación .....	6
1.5	Objetivos .....	6
1.5.1	<i>Objetivo General</i> .....	6
1.5.2	<i>Objetivos Específicos</i> .....	7
1.6	Hipótesis.....	7

### CAPÍTULO II

2	MARCO TEÓRICO .....	8
2.1	Antecedentes .....	8
2.2	Contaminación ambiental por basura.....	10
2.2.1	<i>Contaminación por plástico</i> .....	11
2.2.2	<i>Contaminación por Hojalata</i> .....	12
2.2.3	<i>Contaminación por vidrio</i> .....	13
2.3	Trabajos desarrollados con respecto al tema de estudio .....	13
2.4	Elementos de percepción visual .....	15
2.4.1	<i>Estructura del ojo humano</i> .....	15
2.4.2	<i>Percepción de color</i> .....	15
2.4.3	<i>Captación de la imagen por cámara</i> .....	15
2.4.4	<i>Iluminación</i> .....	16
2.5	<i>Imagen Digital</i> .....	16
2.6	Visión por Computador .....	17
2.7	Procesamiento digital de Imágenes.....	18



2.7.1	<i>Adquisición</i> .....	19
2.7.2	<i>Pre-procesamiento</i> .....	21
2.7.3	<i>Segmentación</i> .....	22
2.7.4	<i>Representación y descripción</i> .....	22
2.7.5	<i>Reconocimiento</i> .....	23
2.7.6	<i>Interpretación</i> .....	23
2.8	<b>Detección de Objetos</b> .....	24
2.8.1	<i>Témlate Maching</i> .....	24
2.8.1	<i>Machine Learning</i> .....	26
2.8.2	<i>Detección de objetos por Cascada</i> .....	26
2.8.3	<i>HOG</i> .....	27
2.8.4	<i>LBP</i> .....	30
2.8.5	<i>Aggregate Channel Features (ACF)</i> .....	32
2.8.6	<i>Deep Learning</i> .....	33
2.8.7	<i>Otros métodos de reconocimiento de objetos</i> .....	33

### CAPÍTULO III

3	<b>DISEÑO DEL <i>HARDWARE</i> Y <i>SOFTWARE</i> DEL SISTEMA</b> .....	34
3.1	<b>Requerimientos del hardware</b> .....	34
3.2	<b>Requerimientos de <i>Software</i></b> .....	35
3.3	<b>Arquitectura del sistema</b> .....	35
3.3.1	<i>Módulo de Adquisición de Imágenes</i> .....	36
3.3.2	<i>Módulo Procesamiento de Imágenes</i> .....	37
3.4	<b>Selección de los elementos <i>hardware</i> del sistema</b> .....	38
3.4.1	<i>Phantom 3 Professional</i> .....	38
3.4.2	<i>Ordenador portátil</i> .....	39
3.5	<b>Selección de <i>software</i> para el sistema</b> .....	40
3.5.1	<i>PIX4D</i> .....	40
3.5.2	<i>Matlab</i> .....	41
3.6	<b>Diagrama de flujo del programa para el Módulo de Adquisición de Imágenes</b> .....	42
3.7	<b>Diagrama de flujo para el Módulo de Procesamiento de Imágenes</b> .....	42

## CAPÍTULO IV

<b>4</b>	<b>PRUEBAS Y RESULTADOS .....</b>	<b>45</b>
<b>4.1</b>	<b>Preprocesamiento .....</b>	<b>45</b>
<b>4.2</b>	<b>Entrenamiento de los detectores .....</b>	<b>46</b>
<b>4.2.1</b>	<b><i>Detectores de Botellas .....</i></b>	<b>48</b>
4.2.1.1	<i>Detector de objetos en cascada.....</i>	48
4.2.1.2	<i>Detector ACF .....</i>	49
<b>4.2.2</b>	<b><i>Evaluación de los detectores .....</i></b>	<b>50</b>
4.2.2.1	<i>Matriz de confusión.....</i>	50
4.2.2.2	<i>Cálculo de la sensibilidad y precisión .....</i>	51
<b>4.3</b>	<b>Resultados.....</b>	<b>51</b>
<b>4.3.1</b>	<b><i>Cálculo del área contaminada por botellas plásticas.....</i></b>	<b>56</b>
4.3.1.1.	<i>Filtrado de Imagen.....</i>	56
4.3.1.2.	<i>GSD (GROUND SAMPLING DISTANCE).....</i>	57
	<b>CONCLUSIONES.....</b>	<b>59</b>
	<b>RECOMENDACIONES.....</b>	<b>61</b>
	<b>BIBLIOGRAFÍA</b>	
	<b>ANEXOS</b>	

## ÍNDICE DE TABLAS

**Tabla 1-4:** Datos de Precisión y Sensibilidad del detector en cascada con descriptores HOG..... 52

**Tabla 2-4:** Datos de Precisión y Sensibilidad del detector en cascada con descriptores LBP ..... 54

## ÍNDICE DE FIGURAS

<b>Figura 1-2.</b>	Tiempo estimado para biodegradarse.....	11
<b>Figura 2-2.</b>	Código de Identificación de Resina.....	12
<b>Figura 3-2.</b>	Etapas procesamiento digital de Imágenes.....	19
<b>Figura 4-2.</b>	Técnica de coincidencia de plantillas.....	25
<b>Figura 5-2.</b>	Representación de Boosting.....	27
<b>Figura 6-2.</b>	Clasificador Débil.....	27
<b>Figura 7-2.</b>	Representación del Gradiente.....	28
<b>Figura 8-2.</b>	Histogramas de Orientación de Gradientes.....	29
<b>Figura 9-2.</b>	Normalización de Bloques.....	30
<b>Figura 10-2.</b>	Invarianza en Local Binary Pattern.....	30
<b>Figura 11-2.</b>	Local Binary Pattern básico.....	31
<b>Figura 12-2.</b>	Histogramas de LBB y LBP Uniforme.....	31
<b>Figura 13-2.</b>	<i>Aggregate channel features</i> para la detección de rostro.....	32
<b>Figura 1-3.</b>	Arquitectura Sistema.....	36
<b>Figura 2-3.</b>	Diagrama de bloques del módulo de adquisición de imágenes.....	37
<b>Figura 3-3.</b>	Diagrama de bloques del módulo de procesamiento de imágenes.....	37
<b>Figura 4-3.</b>	<i>Drone Phantom 3 Professional</i> .....	38
<b>Figura 5-3.</b>	Ordenador Sony VAIO core i7.....	39
<b>Figura 6-3.</b>	Herramientas de <i>software</i> PIX 4d en los planes de vuelo.....	41
<b>Figura 7-3.</b>	Herramientas de procesamiento de imagen y visión por computadora de Matlab...	41
<b>Figura 8-3.</b>	Diagrama de flujo del módulo de adquisición de imágenes.....	42
<b>Figura 9-3.</b>	Diagrama de bloques del módulo de procesamiento de imágenes.....	44
<b>Figura 1-4.</b>	Preprocesamiento.....	45
<b>Figura 2-4.</b>	Tabla de etiquetado en Matlab.....	46
<b>Figura 3-4.</b>	Detección HOG.....	49
<b>Figura 4-4.</b>	Detección ACF.....	49
<b>Figura 5-4.</b>	Matriz de Confusión.....	50
<b>Figura 6-4.</b>	Grafica lineal de la precisión de acuerdo al número de etapas.....	53
<b>Figura 7-4.</b>	Grafica de Precisión y Sensibilidad HOG.....	53
<b>Figura 8-4.</b>	Grafica lineal de la precisión de acuerdo al número de etapas.....	54
<b>Figura 9-4.</b>	Grafica de Precisión y Sensibilidad LBP.....	55
<b>Figura 10-4.</b>	Grafica lineal de la precisión y tasa de error promedio logarítmica.....	56
<b>Figura 11-4.</b>	Aplicación de filtros y binarización para el cálculo del área.....	57
<b>Figura 12-4.</b>	Cálculo de GSD.....	58

## **ÍNDICE DE ANEXOS**

**ANEXO A** HOJAS DE ESPECIFICACIONES DRONE PHANTOM 3

**ANEXO B** TABLA COMPARATIVA SOFTWARE DE FOTOGRAMETRÍA

**ANEXO C** CÓDIGO DE PROGRAMACIÓN MATLAB

## RESUMEN

El objetivo fue aplicar procesamiento digital de imágenes en la detección de material contaminante inorgánico en orillas de lagunas, para ello se realizó un sistema de detección de objetos capaz de identificar botellas plásticas presentes en el ecosistema mediante el entrenamiento de un detector. En el desarrollo de la investigación se trabajó con métodos teórico-empírico con enfoque cuantitativo y cualitativo; Para el entrenamiento del detector se trabajó con diferentes algoritmos de detección basados en los siguientes descriptores: Patrones Binarios Locales (LBP), Histograma de gradientes orientados (HOG) y Características de Canales Agregados (ACF) se evaluó en base a su sensibilidad y precisión. Como Resultado de la evaluación se determinó que el sistema de detección por cascadas con descriptor de Histograma de gradientes orientados tiene una mayor precisión en el reconocimiento de la botella plástica con porcentaje del 70% en su detección. Se diseñó el sistema de detección con dos módulos, el primero, módulo de adquisición de imágenes, es el encargado de la captura, selección y almacenamiento de los datos para ello se hace uso de un Vehículo Aéreo no Tripulado (VANT) conjuntamente con el *software* de mapeo *PIX4D Capture*, mientras que en el segundo módulo, el módulo de procesamiento de imágenes, se estandarizan los datos, se mejora las características del objeto en las imágenes con el uso de los filtros de ruido, contraste y ajuste de histogramas, se ejecuta el algoritmo de detección y se calcula el área que ocupa la botella en pixeles para finalmente mostrar como resultado el área de las botellas detectadas en la zona de análisis.

Palabras Clave: <PROCESAMIENTO DIGITAL DE IMAGENES>, <VEHICULO AEREO NO TRIPULADO>, <PATRONES BINARIOS LOCALES (LBP)>, <HISTOGRAMA DE GRADIENTES ORIENTADOS (HOG)>, < DETECTOR DE CARACTERISTICAS DE CANAL AGREGADO (ACF) >.



0003-DBRA-UPT-IPEC-2022

## **ABSTRACT**

The objective of this work was to apply digital image processing in the detection of inorganic contaminant material on the shores of lagoons, for which an object detection system capable of identifying plastic bottles present in the ecosystem was carried out by calibrating a detector. In the development of the research, we worked with theoretical-empirical methods of quantitative and qualitative approaches. For the calibrating of the detector, different detection algorithms were used based on the following descriptors: Local Binary Patterns (LBP), Histogram of oriented gradients (HOG) and Aggregated Channel Characteristics (ACF) which were evaluated based on their sensitivity and precision. As a result of the evaluation, it was determined that the cascade detection system with a Histogram descriptor of oriented gradients has a greater precision in the recognition of plastic bottles with a detection percentage of 70%. The detection system was designed with two modules, the first, the image acquisition module, in charge of capturing, selecting, and storing the data for which an Unmanned Aerial Vehicle (UAV) is used together with the software of mapping PIX4D Capture. In the second module, the image processing module, the data is standardized, the characteristics of the object in the images are improved with the use of noise, contrast, and histogram adjustment filters, the detection algorithm is executed and the area occupied by the bottle in pixels is calculated so as to finally present the area of the bottles detected in the analysis zone.

Keywords: <DIGITAL IMAGE PROCESSING>, <UNMANNED AIRCRAFT VEHICLE>, <LOCAL BINARY PATTERNS (LBP)>, <ORIENTED GRADIENT HISTOGRAM (HOG)>, <AGGREGATED CHANNEL FEATURE DETECTOR (ACF)>.

## INTRODUCCIÓN

El ser humano se acostumbrado a utilizar el medio ambiente de diferente manera con el objetivo de satisfacer sus necesidades sin la conciencia necesaria para disminuir el impacto ambiental generado por miles de años de destrucción de la naturaleza como la tala de árboles, la caza indiscriminada, desperdicios orgánicos e inorgánicos entre muchas otras. Sin embargo, en la actualidad uno de los grandes problemas que afectan de manera indiscriminada a todo el planeta es el uso excesivo de plástico desechado en el ambiente generando que los animales lleguen a morir por el consumo de este material contaminante y con un tiempo de degradación de al menos 450 años.

El facilismo al que está acostumbrado el ser humano ha permitido el consumo colosal del plástico en el diario vivir, ya sea a modo de botellas, fundas, ropa, juguetes, dispositivos médicos, etc.

El plástico o también llamado polímero es un derivado del petróleo, que a su vez es un recurso no renovable, uno de los productos de polímero con más consumo en el mercado son las botellas plásticas para su utilización en la comercialización de líquidos especialmente, también conocido como tereftalato de polietileno (PET), fue inscrito con la denominación de un polímero para fibra. (Quishpe & Ortiz, 2018, p. 3).

En este proyecto nos centramos en la problemática producida por las botellas plásticas que debido a la falta de práctica de normas medioambientales en su mayoría son desechadas en ambientes como lago o lagunas especialmente aquellas que son de uso turístico, a pesar de ello algunas de estas botellas pueden ser reutilizadas o recicladas.

De acuerdo a la información presentada por la revista *The Guardian* se publicó que en el año 2016 se vendieron 480.000 millones de botellas plásticas en todo el mundo, la mayoría de estas botellas plásticas fueron utilizadas para bebidas embotelladas y solo el 50% de esa cantidad se acopió para su adecuado reciclaje y el 7% de aquellas acopiadas se convirtió en nuevas botellas, en cambio, la mayor parte de las botellas plásticas fabricadas acaba en vertederos o en los océanos. De acuerdo con el reporte global de preferencias de envases de *Euromonitor International* para el 2021 se predice 583.300 millones de ventas de estos productos plásticos. (Quishpe & Ortiz, 2018, p. 3)



Ecuador actualmente no cuenta con un adecuado manejo de los desechos ya que seis de cada diez hogares reciclan según datos del INEC. (INEC, 2020, p. 8) Es necesaria la concientización de la sociedad acerca del impacto que causa el desecho de materiales contaminantes en el medio ambiente dañando la fauna flora y generando contaminación en el agua. Es por ello que muchos de los avances tecnológicos que se han ido desarrollando a lo largo de estos años han ido dirigidos a esta problemática en el país.

Uno de los grandes avances tecnológicos explotados actualmente a nivel mundial es la inteligencia artificial, que tiene como subramas a la visión por computadora, *matching learning* y *deep learning*, que no es más que un conjunto de algoritmos capaces de enseñar a una computadora a realizar tareas de aprendizaje automático. En este proyecto se verá reflejado el uso de estas tecnologías para el desarrollo del proyecto.

Dentro de las diferente tecnologías antes expuestas se analizó el uso del procesamiento digital de imágenes para la detección de objetos en fotografías o videos tomadas por *drones* o con el uso de cámara convencional, en este proyecto se determina la cantidad de material plástico (Botellas plásticas) presente a orillas de lagunas, este estudio se puede usar como precedente para la detección de material plástico en cualquier tipo de ambiente generando una mayor porcentaje de reciclaje al implementarlas en los botaderos o industrias que deseen hacer uso de este material. Por lo antes expuesto se ha desarrollado la investigación cuyo objetivo es aplicar procesamiento digital de imágenes en la detección de material contaminante inorgánico en orillas de lagunas, y para el cumplimiento de la misma se convino la utilización de los siguientes métodos y técnicas:

Métodos Teóricos:

- Revisión documental, para conocer información relacionada a materiales de contaminación inorgánica a la orilla de lagunas, procesamiento digital de imágenes, reconocimiento de objetos, y trabajos relacionados al objeto de estudio, así como avances tecnológicos.
- Histórico-lógico, empleado para estudiar la evolución de los sistemas para el procesamiento digital de imágenes en el reconocimiento de objetos.
- Sistematización, para aplicar la información bibliográfica y experimental al diseño del sistema.
- Análisis y síntesis, para evaluar los resultados, llegar a conclusiones y recomendaciones.

Métodos empíricos:

- Experimentación, para realizar la simulación y comprobación de los módulos que se describen en la investigación.
- Medición, para la evaluación del sistema durante la implementación práctica de la investigación.
- Observación, para la validación del sistema diseñado.

Para cumplir con el objetivo trazado se llevarán a cabo las siguientes tareas de Investigación:

- Estudio y selección de técnicas de procesamiento de imágenes y reconocimiento de objetos.
- Revisión bibliográfica sobre material contaminación inorgánico.
- Diseño e implementación del *hardware* y *software*.
- Pruebas y puesta a punto de los algoritmos.
- Valoración de resultados y análisis de errores.

Resultados esperados: Disponer de un prototipo de sistema de procesamiento digital de imágenes en la detección de material contaminante inorgánico en orillas de lagunas que sea modular, reconozca material inorgánico (botellas de plástico), y presente el porcentaje del área de contaminación.

La presente investigación consta de un Resumen e Introducción, cuatro capítulos, conclusiones y recomendaciones. El **primer capítulo**, se revisa la problemática y se definen objetivos generales y específicos de esta investigación. El **segundo capítulo** define el marco teórico enfocado al tema de estudio, como material contaminante inorgánico, procesamiento digital de imágenes, técnicas de reconocimiento de objetos. El **tercer capítulo** contempla la metodología de la investigación, el *hardware* y *software* del sistema, donde se define la arquitectura, y características del sistema, en el **Capítulo cuatro** se integra los componentes del sistema, y se evalúa; finalmente en las **Conclusiones**, se resumen los resultados fundamentales alcanzados en la investigación y se resaltan aspectos importantes. Y en las **Recomendaciones** se plantea algunos objetivos que se pudieran tener en cuenta para el desarrollo de investigaciones posteriores y trabajos futuros.

## CAPÍTULO I

### 1 PLANTEAMIENTO DEL PROBLEMA

En este contempla la situación de la problemática y los objetivos generales y específicos como requisitos necesarios para el desarrollo de la presente investigación.

#### 1.1 Situación problemática

La contaminación por basura orgánica e inorgánica es uno de los grandes problemas que se presenta en el país, Ecuador produce aproximadamente 4 millones de toneladas de desechos al año, pero no toda esta cantidad debería terminar en los rellenos. Del millón de toneladas de residuos sólidos que podrían reciclarse en el país, únicamente entre un 15 y un 25% logra este fin. (El Comercio, 2017) Muchos de los residuos de comida chatarra como latas, botellas de vidrio, plástico o papel entre otros son botados al entorno provocando contaminación ambiental en ríos, lagos, lagunas, barrancos, quebradas, entre otros.

El plástico es uno de los materiales más utilizados en el diario vivir, ya sea a modo de botellas, fundas, ropa, juguetes, dispositivos médicos, etc. Los plásticos son definidos como: *“Materiales compuestos principalmente por moléculas grandes (polímeros) que se fabrican sintéticamente o, si se producen naturalmente, están muy modificados.”* (Strong, 2005, p. 1) Su nombre se debe a la propiedad que tiene de moldearse a cualquier forma, sin embargo, lo que lo convierte en favorito de las industrias es el bajo costo de su producción. Debido a todas sus propiedades es usado en diferentes campos lo que conlleva a que el aumento de material desechado haya ido creciendo a pasos agigantados provocando que lagunas, ríos, barrancos, calles, mares y océanos estén llenos de este material.

De acuerdo al artículo publicado por el semanario *“The Economist”* desde 1950 se han producido 6300 millones de toneladas de desecho plásticos, donde 80% de ellos han sido depositados en vertederos y el medio ambiente, se ha botado en aguas alrededor de 10.000 de toneladas existiendo 50.000.000 millones de micropartículas de plástico en los océanos, que son irrecuperables. (The Economist, 2018),

La contaminación es uno de los grandes problemas ambientales a lo que nos enfrentamos hoy en día, debido a que falta de interés en concientizar a las personas ha provocado que muchos ríos o zonas de recreación como orillas de ríos, lagos o lagunas queden llenos de desperdicios. Entre los principales desperdicios encontrados se pueden destacar las botellas de plástico las mismas que actualmente pueden ser recicladas, así también como latas de aluminio, botellas de vidrio entre otros.

Según el diario el Comercio los datos de la Tercera Comunicación Nacional del Ecuador revelan que el sector de residuos es el responsable del 4,19% del total de emisiones de gases contaminantes a la atmósfera, recalando que actualmente alrededor de la mitad de los municipios aún no cuenta con un manejo adecuado de estos desechos sólidos. El ministro de ambiente del Ecuador indica que: A través de la separación de los residuos en los hogares, se puedan reutilizar los materiales y crear negocios para cada material como para el plástico y el papel. Granizo recaló que esto es labor de los municipios y el trabajo del MAE será apoyarlos y acompañarlos en el proceso. (El comercio, 2018)

Bajo este contenido se busca desarrollar la investigación que trata sobre la detección de material contaminante inorgánico (botellas de plástico) en orillas de lagunas mediante el uso del procesamiento digital de imágenes, determinando un estimado del área.

## **1.2 Formulación del problema**

¿Cómo procesar digitalmente las imágenes para la detección de botellas plásticas en orillas de lagunas y estimar el área contaminada?

## **1.3 Preguntas directrices o específicas de la investigación**

¿Qué material contaminante se puede encontrar en las orillas de lagunas y sus efectos al ecosistema?

¿Qué características mínimas se necesitan para la adquisición de datos?

¿Qué técnicas permite mejorar la calidad de la imagen para continuar con su análisis?

¿Qué técnicas de PDI se usan para la detección material contaminante a orillas de lagunas?

¿Es posible la detección de botellas plásticas en las orillas de lagunas mediante el uso de técnicas de procesamiento de imágenes?

## **1.4 Justificación**

En el Ecuador no existen datos reales de la cantidad de material plástico como botellas y fundas, así como latas de aluminio y demás basura inorgánica que se vota en ríos, lagos, lagunas o el mar ni un control de la cantidad de material no degradable que se depositan cerca de ríos calles o barrancas, por lo que se podría decir que se desconoce el impacto ambiental que está causando el desperdicio de estos materiales en el ambiente.

El país no cuenta con un adecuado manejo de los desechos ya que seis de cada diez hogares reciclan según datos del INEC. Es necesaria la concientización de la sociedad acerca del impacto que causa el desecho de materiales contaminantes en el medio ambiente dañando la fauna, flora y generando contaminación en el agua. (INEC, 2020, p. 2)

El uso del procesamiento digital de imágenes para la detección de objetos en fotografías o videos tomadas por *drones* o con el uso de cámara convencional se encuentra en auge por lo que es una tecnología nueva que permitirá en este proyecto, determinar la cantidad de material plástico, aluminio y otros a orillas de lagunas, este estudio se puede usar como precedente para la detección de material plástico en cualquier tipo de ambiente generando una mayor porcentaje de reciclaje al implementarlas en los botaderos o industrias que deseen hacer uso de este material.

## **1.5 Objetivos**

### ***1.5.1 Objetivo General***

Aplicar procesamiento digital de imágenes en la detección de material contaminante inorgánico en orillas de lagunas.

### ***1.5.2 Objetivos Específicos***

- Analizar las características de los contaminantes inorgánicos y los efectos que casusa en el medioambiente.
- Investigar, seleccionar y aplicar el proceso de adquisición de datos que se ajusten a los requerimientos del sistema.
- Establecer los procesos precisos para el reconocimiento de material contaminante en las imágenes previamente recolectadas y estimar el área de contaminación por desechos en la zona analizada.
- Realizar la validación del sistema implementado.

### **1.6 Hipótesis**

Mediante el uso de técnicas de procesamiento digital de imágenes se permitirá detectar botellas plásticas en imágenes obtenidas a orillas de lagunas.

## CAPÍTULO II

### 2 MARCO TEÓRICO

En este capítulo se presenta los antecedentes de la problemática, se hace una introducción al concepto de visión por computadora, así como a las técnicas que se han estado empleando para el reconocimiento de objetos; se describe el procesamiento digital de imágenes, y algunos trabajos desarrollados.

#### 2.1 Antecedentes

La contaminación es un tema preocupante en la actualidad, ya que la vida marina está sufriendo daños irreparables por la cantidad de desechos que terminan en ríos, lagos, lagunas e incluso llegando al océano. Entre los materiales más usados y que en mayor porcentaje se encuentran en el medio ambiente podemos recalcar al plástico, el mismo que hasta el 2015 ha generado 6300 millones de toneladas de los cuales el 9% es reciclado, el 12 % incinerado y el 79% acumulado en entornos naturales. (BBC, 2017)

La gran cantidad de plástico fue generado por la idea de artículos desechables, como envases, cucharas, vasos, botellas, y una gran cantidad de objetos de uso diario. Las botellas de plástico son unos de los artículos más usados en 2016, se han generado 480 000 millones de las cuales 110 000 millones son de Coca Cola, generando alrededor de 20 000 botellas por segundo aumentando drásticamente la contaminación ambiental. A pesar de los esfuerzos que se han hecho los últimos años para fomentar el reciclaje, menos del 50 % de botellas son recicladas, de acuerdo a la BBC solo el 7% de estas se convierten en nuevas. (BBC, 2017)

Los envases de PET (tereftalato de polietileno) son los productos de plásticos más usados en la actualidad por las industrias alimenticias por ser irrompible, económico, liviano, impermeable y reciclable; además PET es la resina con mejores características para el reciclado, según Greenpeace. Estos envases demoran en degradarse aproximadamente 700 años a pesar de ello su uso ha ido en aumento, para su elaboración se usan grandes cantidades de petróleo, con 24 millones de galones se producen 1.000 millones de botellas. (Semana S.A, 2016) Para reducir el impacto ambiental creado por este material, se pretende incentivar a empresas recicladoras y productoras de envases, a reutilizar los

envases PET, esta decisión disminuiría en gran parte el problema de botellas que contaminan ríos, lagos o playas además de otros entornos naturales. La educación ambiental es otra de las maneras que se puede trabajar para reciclar este y otros tipos de materiales ya sean plásticos u otros.

En los últimos años se han realizado grandes descubrimientos para reducir el impacto ambiental, entre ellos se puede destacar a la empresa Unilever junto con *Ioniqa & Indorama Ventures*, quienes han desarrollado un mecanismo para convertir los residuos de PET en plástico virgen. El reciclaje de botellas PET se lo realiza de manera manual con operadores en los botaderos sin embargo existen maquinarias como son *Reserve Vending Machine (RVM)* donde se almacena botellas PET y entregan un recibo que será intercambiado con incentivos, seis de estas máquinas según MAE serán colocadas en diferentes ciudades del país para concientizar a la ciudadanía. (MAE, 2014)

El procesamiento de imágenes, usado para el reconocimiento de objetos con el paso de los años ha ido mejorando de forma impresionante, se ha aplicado principalmente para la localización y clasificación de objetos en imágenes obtenidas de redes sociales, *drones* o cámaras digitales ya sea en ambiente controlados como laboratorios o en el entorno natural. El objetivo de este proceso es el de reducir su tasa de error tanto que sea igual que el ojo humano. (Bejerano, 2014)

El reconocimiento de plástico a través de imágenes se ha estudiado en diferentes artículos cabe recalcar los siguientes proyectos:

- Kumar y Prasad (2012) en el reconocimiento de botellas y su clasificación de acuerdo al tipo de material del que se encuentran fabricadas como vidrio, metal polietileno, y las botellas de polietileno de baja densidad (LDPE).
- También se han desarrollado algoritmos para la identificación y el reconocimiento de las botellas de plástico (botellas de leche) en movimiento en una cinta transportadora sin embargo la deformación de las botellas de leche durante el proceso de reciclado no garantiza una clasificación al 100%, este estudio se realizó bajo condiciones de iluminación y adquisición controladas (Scavino y otros, 2009), así como el uso de técnicas de visión artificial para identificar HDPE de calidad alimentaria, ejemplificada por las botellas de leche de plástico (Ioannis & Bharath, 2011).
- Reconocimiento de materiales plásticos a través de técnicas de visión artificial en color al objeto de su Implementación en una línea de triaje de una E.R.S.U. este artículo presenta estudios



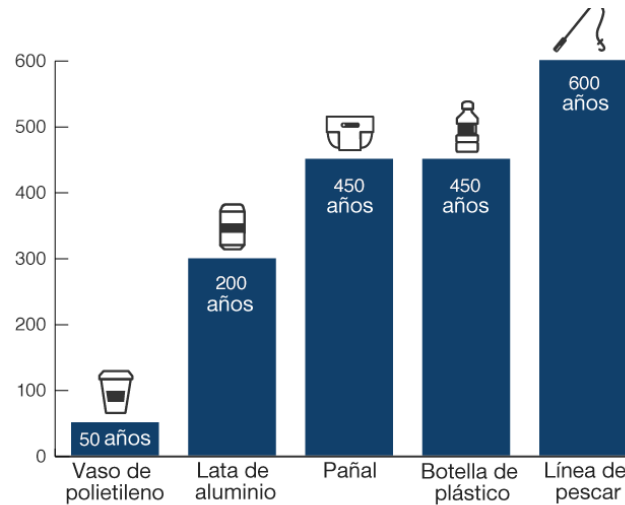
preliminares que permitan evaluar la viabilidad técnica del reconocimiento de materiales plásticos situados sobre la cinta de triaje manual (Martínez y otros, 2001).

En vista de lo expuesto en el presente proyecto se realizará el desarrollo del proceso de reconocimiento de imagen para la detección de botellas plásticas en un ambiente no controlado como son las orillas de lagunas, con el uso de un *drone* para la recolección de las imágenes hacer analizadas, complementando de esta manera los estudios antes mencionados y ayudando a la detección de material contaminante en el medio ambiente y estimar el área de contaminación de las zonas analizadas.

## **2.2 Contaminación ambiental por basura**

La contaminación ambiental es un problema preocupante en todo el mundo, sin embargo, existen algunos tipos de contaminación que no podemos controlar por ejemplo la basura que se genera por el uso de comida chatarra como fundas, botellas y latas de aluminio son materiales que se pueden reciclar y reutilizar. En el país no se ha incentivado a la gente a la separación de basura ni al reciclaje incluso muchas personas todavía continúan botando basura en áreas naturales como ríos, lagos lagunas sin darse cuenta generando contaminación para sí mismos, este tipo de comportamiento se lo puede encontrar especialmente en lugares turísticos parques, ríos, lagunas.

Entre los materiales que mayor uso se da y se desecha se puede recalcar a las botellas plásticas (PET), entre otras siendo esta, la que mayor tiempo tiene en degradarse como se observa en la siguiente figura.



El tiempo exacto varía según el tipo de producto y condiciones ambientales

**Figura 1-2.** Tiempo estimado para biodegradarse

**Fuente:** (NOAA/Woods Hole Sea Grant, 2018)

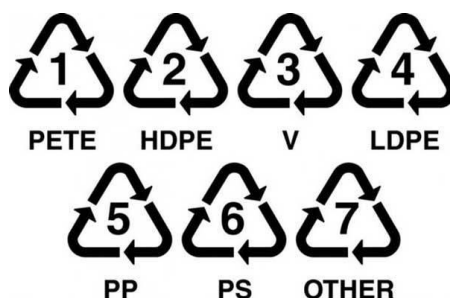
### 2.2.1 Contaminación por plástico

El plástico es el material usado actualmente en la mayoría de productos de tipo textil, médico, de empaquetamiento, recreación, alimenticio, entretenimiento, transporte, industrias y productos del hogar, debido a sus características como versatilidad, fácil de moldearse, baja densidad, ligeros, algunos poseen aislamiento térmico y eléctrico, sin embargo, otros pueden ser usados como conductores, resistentes a la corrosión, duraderos, etc. Este material está compuesto por una parte de polímeros orgánicos y sintéticos, por lo que, si las propiedades físicas de un plástico determinado no cumplen plenamente los requisitos, se puede modificar su equilibrio de propiedades con relleno, colores, agentes espumantes, retardantes de llama, plastificantes, entre otros, para satisfacer la demanda de una aplicación específica. (PlasticsEurope, 2018)

El problema de los productos plásticos radica en la idea que se tiene de ellos como de un solo uso como botellas de bebidas, pañales, cubiertos entre otros. Las botellas plásticas son uno de los tipos más comunes de desechos. En algunos países se han implementado medidas para la reducción de su consumo como las maquinarias para depositar las botellas generándoles a los consumidores una pequeña ganancia por cada botella que se recicle, de esta manera se pretende aumentar el porcentaje de reciclaje de botellas. (BBC, 2017)

Con el fin de reducir la contaminación por botellas de plástico se ha emprendido un programa de recolección y clasificación del material utilizado, de tal manera que, para distinguir el tipo de plástico usado en cada envase se ha implementado una codificación impresa en la misma. En la siguiente figura se puede apreciar la codificación de identificación de la resina usado en las diferentes botellas.

Este proceso se lo ha realizado con el fin de que los consumidores conozcan los efectos nocivos que puede tener el reutilizar una botella plástica, entre los estudios que se han hecho se ha detectado la presencia de actividad estrogénica, lo que podría dañar las hormonas y afectar la salud del consumidor. (Mercola, 2015)



**Figura 2-2.** Código de Identificación de Resina

Fuente: (BM IMPEX ENGINEERING, 2017)

Las botellas que se encuentran en el mercado son en su mayoría del tipo PETE y HDPE. El Tereftalato de Polietileno (PET o PETE) es usado generalmente en botellas de bebidas como botellas de jugos y gaseosas. A este tipo de resina aplicándole la tecnología necesaria se puede convertir en nuevas botellas u otro tipo de artículos, por ello es la que se recicla más. Mientras que el Polietileno de alta densidad (HDPE) es ligeramente peligroso y usado en envases de leche, aguas, jugos, cereales, bolsas, recubrimiento de cajas, champús, líquidos de limpieza del hogar y químicos industriales. (Toledo, 2013, pp. 46-48)

### 2.2.2 Contaminación por Hojalata

Se considera hojalata a los desechos de aluminio debido a que la mayoría de utensilios de cocina, latas para los refrescos, papel de aluminio, revestimientos de las viviendas, fabricación de electrodomésticos, recipientes, entre otros usan el aluminio para la producción de productos de uso cotidiano. El aluminio es un material que puede tardar entre 10 y 100 años en degradarse, dependiendo

de su espesor, y genera óxido de hierro, perjudicial para la tierra y el ser humano. Además, especies como las tortugas marinas o numerosas aves pueden ver amenazada su supervivencia cuando quedan enganchadas a las anillas de estos envases. (Ambientum, 2019)

### **2.2.3 Contaminación por vidrio**

4.000 años es lo que tarda una botella de vidrio en descomponerse, poniendo en riesgo el entorno y su ecosistema. Debido a que los animales pueden cortarse o ingerirlo haciendo imposible su digestión. Además, es uno de los factores de riesgo de incendio más peligrosos. (Ambientum, 2019)

Una vez descritos los principales contaminantes inorgánicos se elige el reconocimiento de botellas plásticas por ser estas las que mayor contaminación producen.

## **2.3 Trabajos desarrollados con respecto al tema de estudio**

En la Universidad Politécnica de Valencia se desarrolló un proyecto de carrera titulado Reconocimiento de objetos mediante *WebCam* en tiempo real cuya finalidad era la implementación de un sistema capaz de capturar, detectar y dar seguimiento a objetos mediante una *WebCam*. Llegaron a la conclusión que el procesamiento y la interpretación correcta de los objetos contenidos en una imagen es una tarea compleja y computacionalmente cara. Cualquier reconocimiento de objetos requiere una capacidad de cómputo muy elevada debido a que las imágenes a procesar son matrices tridimensionales grandes. Si bien el programa desarrollado no presenta mayores dificultades en interpretar objetos contenidos en imágenes estáticas cargadas desde un archivo, la implementación en tiempo real con imágenes obtenidas desde una *WebCam* presenta numerosas dificultades. (Rossius, 2013).

En la Universidad de Barcelona el proyecto de Reconocimiento Automático de Formas permite reconocer una imagen, o una parte de la misma recortada por el usuario, seleccionada desde la cámara o desde la galería del dispositivo y muestra como resultado final, la imagen más parecida a la original encontrada en una base de datos de imágenes guardada en la memoria del dispositivo. El problema de este proyecto es que mientras más grande es la base de datos el sistema de reconocimiento tarda mucho más y necesita capacidad de procesamiento alto. (Sanz, 2013)

Google cuenta con una aplicación de reconocimiento de imágenes con inteligencia artificial denominada Google Lens. La API es rápida y puede clasificar imágenes en categorías muy concretas, detecta caras y las emociones que reflejan, lee textos en varios idiomas directamente desde una fotografía. Sin embargo, no siempre acierta y sólo dice el nombre del objeto que piensa estar observando, si son varios objetos en una imagen se enfoca en uno solo. (Krasis Consulting S.L.U, 2015)

En las Memorias de la Décima Sexta Conferencia Iberoamericana en Sistemas, Cibernética e Informática (CISCI 2017) se presentó un trabajo sobre la revisión del procesamiento digital de imágenes, segmentación para la detección de áreas contaminadas por petróleo. El mismo permite identificar regiones de imágenes digitales contaminadas por petróleo en ambientes controlados, donde se analiza las imágenes con un clasificador controlado y adiestrado para que seleccione áreas contaminadas diferenciando la contaminación por petróleo y aceite quemado. (Bohorquez et al., 2017)

El proyecto *LitterDrone (Development and Exploitation of Innovative Tools for Remote Marine Litter Control and Management through UAVs)* desarrolla un sistema piloto para la caracterización de basura marina en las costas de Europa, con el uso de nuevas tecnologías como vehículos aéreos no tripulados equipados con cámaras de alta resolución que llevarán a cabo muestreos en playas pertenecientes al programa oficial de seguimiento de basuras marinas en playas de cara a probar su potencial técnico y comercial. Con la información recogida a través de *LitterDrone* se podrán detectar con mayor facilidad los puntos de acumulación de basura en la costa, así como la composición de la basura presente en ellos, pudiendo así diseñar actuaciones de prevención y mitigación más precisas y eficientes. (Litterdrone, 2019)

La mayoría de los trabajos desarrollados si bien reconocen objetos o formas, no se les han dado un uso más práctico y que contribuya al bien común. Sin embargo, existen dos trabajos relevantes como el de detección de áreas contaminadas por petróleo o aceite quemado a través de procesamiento digital de imágenes, y el proyecto *LitterDrone* para detectar acumulación de basura en el Parque Nacional Marítimo-Terrestre de las Islas Atlánticas de Galicia.

En el Ecuador no se han desarrollado trabajos con respecto al procesamiento digital de imágenes en detección de material contaminante inorgánico en orillas de lagunas, por lo que este proyecto sería el primero en ser desarrollado.

Para saber cómo funciona el reconocimiento de imágenes es necesario estudiar algunos términos como los que se describen a continuación.

## **2.4 Elementos de percepción visual**

### **2.4.1 Estructura del ojo humano**

El ojo humano tiene la capacidad de captar imágenes con gran nitidez e identificar movimiento, profundidad, color, y enfocar cualquier objeto en diferentes distancias; es capaz de adaptarse a diferentes niveles de iluminación; tiene un diámetro de 20 mm y es casi esférico. El tejido resistente y transparente que cubre la superficie anterior del ojo es llamado córnea, que junto con el iris son los encargados de permitir el paso de la luz, atravesando la lente del cristalino hasta llegar a la parte trasera de la esfera ocular llamada retina que está recubierta por una capa de células fotosensibles la cual recibe una pequeña imagen invertida del mundo exterior. La lente del cristalino altera su forma para enfocar la imagen. (García, 2016, pp. 30-31)

### **2.4.2 Percepción de color**

Se percibe el color mediante los conos, que son los sensores que traducen la energía lumínica en señales nerviosas en el ojo humano de acuerdo a la banda de longitudes de onda son más sensibles, los mismos se pueden dividir en tres clases: Los sensores tipo alfa ( $\alpha$ ) tienen mayor sensibilidad a 480 nanómetros (azul), el tipo beta ( $\beta$ ) a 540 nm (verde) y el tipo gama ( $\gamma$ ) a 570 nm (rojo). Con la combinación de estas tres señales se realiza la identificación del color en la imagen y se percibe una gran variedad de colores. (García, 2016, p. 32)

### **2.4.3 Captación de la imagen por cámara**

Una cámara cuenta con sensores de captación de la imagen los mismos son componentes sensibles a la luz que modifican su señal eléctrica en función de la intensidad luminosa que perciben. La señal eléctrica que transmiten los fotodiodos es en función de la intensidad luminosa que reciben, su espectro, y el tiempo de integración. Cuando accionamos el disparador de la cámara, la luz reflejada por el objeto atraviesa la lente, esta luz sensibiliza al sensor de captación de imagen, el cual descompone estos haces en colores rojo, verde y azul; el sensor convierte estas señales en cargas

eléctricas analógicas para ser transferidas a la electrónica del dispositivo. Mediante la electrónica incorporada en la cámara se transforma los datos análogos a un formato digital para su posterior almacenamiento en la memoria de la misma. (Tecnología & Informática, 2019)

Los elementos a tener en cuenta al momento de elegir una cámara para obtener mejores fotografías son los megapíxeles, una lente de calidad, un procesador rápido, estabilización de imagen, detección de rostros y lentes intercambiables. (Tecnología & Informática, 2019)

#### **2.4.4 Iluminación**

El aspecto más importante tanto para observar como para obtener una imagen con cámaras es la iluminación. Para efectos del desarrollo de este trabajo la iluminación juega un papel fundamental ya que la mejor imagen es aquella que tiene mayor contraste, pues se distinguen las áreas de interés, y se destacan del fondo. (Garcia, 2016, p. 42) La luz es reflejada de dos maneras:

- Por reflexión especular, cada rayo incidente se refleja en una única dirección como en el caso de algunos metales o el espejo, que pueden saturar el sensor de la cámara, en este caso lo mejor es utilizar alguna técnica de iluminación especial.
- Las reflexiones difusas son tenues pero estables, la intensidad de reflexión puede ser de 10 a 1000 veces menor que la fuente de luz.

Si el objeto tiene una buena iluminación se pueden lograr muy buenos resultados en el sistema de reconocimiento de objetos. La visión artificial depende de una iluminación apropiada.

#### **2.5 Imagen Digital**

Se puede definir una imagen como una función bidimensional de luz e intensidad, representado matemáticamente como la función  $f(x,y)$  para cada punto de la imagen  $(x,y)$ , Una imagen consta de dos partes la iluminación  $i(x,y)$  y la reflectividad  $r(x,y)$

$$f(x,y) = i(x,y) * r(x,y) \quad (1)$$

Dónde  $i(x,y)$  representa la luz que incide sobre un objeto y  $r(x,y)$  la cantidad de luz que refleja. (Alegre y otros, 2003)

Las imágenes o gráficos digitales son representaciones bidimensionales constituidas por bits (1 y 0). La adquisición de imágenes y la evolución de los programas de computación hicieron posible el uso de imágenes digitales en diferentes áreas, su captura se puede realizar por cámaras digitales capturando la imagen y representadas por pequeñas cuadrículas denominados pixeles definida como *Picture Element, es el valor de color o intensidad asociado a cada elemento de la matriz; tiene coordenadas espaciales. Puede tomar valores dentro del rango [0, 255]*. (Alegre y otros, 2003)

Una imagen puede ser continua respecto a  $(x,y)$  al igual que en amplitud, por lo que es necesaria su digitalización y muestreo para convertir la imagen a formato digital. Para análisis se utilizan imágenes en gris o a color, la imagen en nivel de gris consiste en el nivel de intensidad de las imágenes monocromáticas. Usualmente se trabaja con imágenes a nivel de gris extrayendo sus características en matrices y las imágenes a color en el modelo RGB (rojo, verde y azul), sin embargo, se puede trabajar en otros modelos como CMYK, HSL,  $L^*a^*b$  o YUV.

Se representa a una imagen como una matriz de pixeles  $M \times N$ , siendo el pixel la representación de la menor unidad homogénea dentro de una imagen. Las imágenes en color se forman por una combinación de imágenes individuales en 2-D, es decir en el sistema de color RGB, una imagen en color se compone de tres matrices de intensidad para los colores rojo, verde y azul. De esta manera se puede usar muchas técnicas de procesamiento de imágenes a nivel de gris en cada uno de los componentes de la imagen a color de forma individual (López, 2016). Existen diferentes programas que permiten el procesamiento de imágenes digitales, en este proyecto se tratará con imágenes obtenidas por cámaras digitales estándares en color (RGB) en el lenguaje de programación Matlab.

## **2.6 Visión por Computador**

La Visión por Computador trata de imitar la funcionalidad del sistema de visión; acciones como diferenciar un carro o clasificar objetos dentro de un espacio visible son el diario vivir de las personas. El desarrollo de esta cualidad en un ordenador es una tarea compleja que lleva años de investigación y muchos recursos. Sin embargo, se han producido importantes avances en cuanto a estos tipos de proyectos. Las redes neuronales profundas, específicamente las redes neuronales convolucionales han



dado paso a este salto de nivel. Algunas aplicaciones destacadas que se han desarrollado son el reconocimiento facial, de señales de tráfico o los automóviles autónomos.

Cuando se habla de visión por computador se encuentra inmersa la visión artificial que engloba al proceso óptico mediante el cual un sistema inteligente es capaz de extraer información de un entorno para su interpretación mediante el uso de la computadora. Debido a la complejidad de la visión humana y cómo esta identifica formas y colores al mismo tiempo, los sistemas de visión artificial resultaron básicamente imprácticos. Sin embargo, se los ha adaptado para simular procesos no tan complejos pero prácticos, capaces de procesar cantidades de datos mucho mayores y en menos tiempo si se trata de tareas repetitivas. La precisión matemática, permite un análisis mucho más detallado y extenso por parte de un sistema basado en la visión artificial, permitiendo la creación sistemas con características diferentes en cuanto a componentes y capacidades, según la tarea a implementar.

Para iniciar un proceso de reconocimiento es necesaria la cuantificación digital del entorno, logrando ser capaz de procesar la información obtenida y aplicar los algoritmos adecuados para su procesamiento y correcta interpretación. Para evitar bucles infinitos debido a la variedad de patrones por reconocer es necesario realizar un preprocesado de la imagen, esto contribuye a que el sistema sea capaz de identificar el objeto independientemente de su iluminación, color o posición dentro de la imagen, y que por lo tanto se adelante a conocer todas las infinitas posibles variaciones que el objeto puede presentar. (Rossius, 2013)

## **2.7 Procesamiento digital de Imágenes**

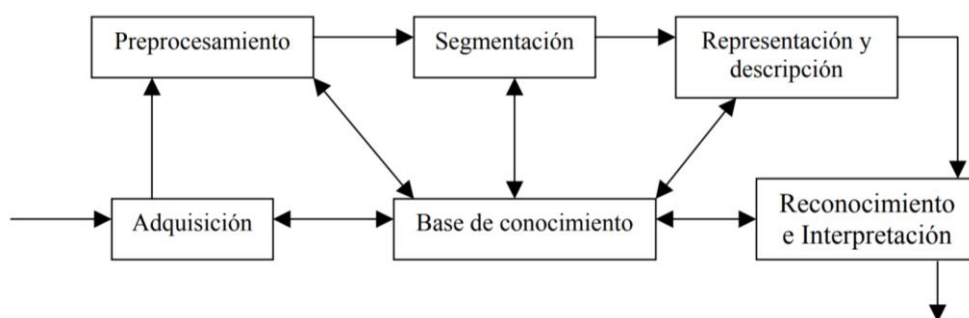
El procesamiento de imágenes, consiste en el uso de técnicas computacionales para llegar a encontrar y comparar características de objetos representados en imágenes, con el fin de tomar una decisión o no. se lo define como:

El proceso digital de imagen aporta tecnología informática para el estudio integral de IMAGENES. Se entiende bajo tal denominación al conjunto de opciones que permiten receptionar y optimizar la información contenida en imágenes de todo tipo (micro, macro, aéreas, remotas, color, monocromo, infrarrojas, etc.). Esta información puede referirse tanto a aspectos directamente visibles en las imágenes como a otros recursos ocultos a la percepción visual y que pueden ser detectados por estimulación de los materiales mediante diversas radiaciones selectivas (infrarroja, térmica, radar, rayos X, etc.) (Barriuso, 1991, p. 1)

Hasta el momento no se ha podido definir donde empieza o termina el área de aplicación del procesamiento digital de imágenes, sin embargo, se ha ido avanzando en el uso de este mecanismo para diferentes áreas de las que cabe destacar su uso en la identificación de algunas especies de fitoplancton, zooplancton, bacterias, virus en imágenes de tejido de camarón, cromosomas del abulón rojo, azul y amarillo, parásitos de peces. Así como también en el estudio de superficies marinas y estudios de oleajes. (Álvarez, 2012)

La arqueología utiliza este proceso en el reconocimiento de materiales con componentes de cerámica o metálicos, así como en la visión remota de la superficie de la tierra constituyendo el fundamento del proceso denominado *Remote Sensing* o teledetección. Con el uso de la tecnología se puede analizar imágenes remotas (por ejemplo, procedentes de satélites), generando la detección de cualificaciones de cultivos, materiales geológicos, composición textural del suelo, etc. (Barriuso, 1991)

Las etapas del procesamiento digital de imágenes son:



**Figura 3-2.** Etapas procesamiento digital de Imágenes

Fuente: (Martínez y otros, 2001)

### 2.7.1 Adquisición

Se denomina adquisición al proceso de obtener la imagen por captura a través de cámaras digitales o digitalización por medio de escáneres. El proceso de adquisición se da mediante cuatro tareas:

a) El proceso de formación de imágenes mediante el cual una información luminosa en 3D (imagen real) es proyectada en un plano 2D (imagen digital). Para ellos existen algunos dispositivos como las cámaras digitales que imitan el proceso del ojo humano. (Gómez & Guerrero, 2016, p. 13)

b) Representación de imágenes digitales a través de una matriz o un vector bidimensional de números binarios (0, 1), se representan como pequeños cuadros, en forma de mosaico individual denominados píxeles. En la imagen binaria cada celda de la matriz está representada por 1 píxel que equivale 1 bit. Si el bit es 0, el color es negro; si el bit es 1, el color es blanco. En una imagen en escala de grises cada celda de la matriz está representada por 1 píxel que equivale 1 byte, que permite 256 niveles de gris; desde 0 que es color negro, hasta 255 que es color blanco. Y por último una imagen a color cada celda de la matriz está representada por 1 píxel que equivale a 3 bytes. Cada píxel consta de 3 valores (Rojo, Verde, Azul), un byte por color y tiene 16,7 millones de colores posibles. (Gómez & Guerrero, 2016, p. 13)

c) Existen muchos dispositivos de captura y digitalización de imágenes; como cámaras, videocámaras, y cada una se distingue de la otra según la resolución y calidad de la fotografía, así como existen muchos tipos de escáner como planos, de microfilm; utilizados para diferentes actividades según la necesidad. Pero la definición básica de los escáneres es que son dispositivos de entrada en el ordenador que permite la captura de una imagen, documento de texto o fotografía, y lo transfiere en bits de información, los cuales puede entender y manejar un ordenador. En cambio, una cámara fotográfica digital en vez de capturar y almacenar fotografías en películas fotográficas como las cámaras fotográficas convencionales, lo hace digitalmente mediante un dispositivo electrónico, o en cinta magnética usando un formato analógico como muchas cámaras de video.

d) Los formatos de fichero de imagen son una forma estándar de organizar y almacenar los datos que representan la imagen. Existen diferentes tipos de formatos según la necesidad de almacenamiento, que se diferencian según los niveles de profundidad admitidos, tipo de compresión y otras características como la posibilidad de definir de transparencias o diferentes imágenes en un mismo archivo. Algunos de los formatos más utilizados y estándares son BMP, TIFF, GIF y JPG.

*BMP (Windows Bitmap)*: este formato permite muchos niveles de profundidad 1 bit por píxel (2 colores), 4 bits (16 colores), 8 bits (escala de grises o paleta), 16 bits (Hi-color) y 24 bits (3 bytes - True-color). A esta clase de archivos puede seleccionarse una compresión RLE (*Run Length Encoding*) sin pérdida de calidad. No sufre pérdidas de calidad y por tanto resulta adecuado para guardar imágenes que se desean manipular posteriormente. Guarda gran cantidad de información de la imagen. No es adecuado para transmisión por red. (Gómez & Guerrero, 2016, p. 15)

El formato TIFF (*Tagged Image File Format*) propuesto para trabajos de impresión de alta resolución y calidad. Es un formato muy flexible, basado en tags (etiquetas). Es ideal para archivar archivos originales. Utiliza cualquier profundidad de color de 1 a 32 bits. Produce archivos muy grandes.

Formato GIF (*Graphics Interchange Format*) está basado en el uso de paletas de hasta 256 colores. Este formato usa el algoritmo de compresión LZW (compresión sin pérdida) sin embargo el uso de paletas implica una pérdida de información de color. Permite definir transparencias. Los ficheros pueden contener múltiples imágenes, esto permite crear animaciones sencillas y está diseñado específicamente para comprimir imágenes digitales. Admite gamas de menor número de colores y esto permite optimizar el tamaño del archivo que contiene la imagen. Adecuado para transmisión en redes.

Formato JPEG (*Joint Photographic Experts Group*) orientado al almacenamiento de imágenes fotográficas. Admite imágenes en escala de grises (1 byte por píxel) y RGB (3 bytes por píxel). Incluye un mecanismo avanzado de compresión, que puede ajustarse a distintas *ratios* de compresión. En la mayoría de los casos, consigue una ratio compresión/calidad mucho mejor que los otros formatos. Nivel de compresión ajustable. Es ideal para publicar fotografías en la web siempre y cuando se configuren adecuadamente dimensiones y compresión. Si se trabaja con un JPEG guardando en disco tras cada operación, la imagen se va degradando. No es recomendable para archivar originales. (Gómez & Guerrero, 2016, p. 15)

### **2.7.2 Pre-procesamiento**

Radica en el uso de técnicas para mejorar la imagen, en caso de ser adquiridas por video es necesario aplicar un conjunto de métodos para estabilizarlas, las mismas que después serán inducidas al proceso de mejora, restauración, compresión o descompresión de acuerdo a las necesidades del sistema. Este proceso se logra mediante el análisis de histogramas, uso de filtros, estiramientos de contrastes entre otros. (Mejía, 2005, pág. 2) Algunas técnicas utilizadas comúnmente son:

Reducción del ruido, las imágenes digitales son propensas a contener diferentes tipos de ruido entre los que se destacan el ruido Gaussiano, el ruido sal y pimienta entre otros, para mejorar la imagen se pueden aplicar filtros lineales, de promediación, mediana e incluso redes neuronales previamente entrenadas.

Técnicas de realce en el Dominio espacial son aquellas técnicas cuyas operaciones se efectúan sobre el valor del píxel, sin tener en cuenta los píxeles vecinos. Estas transformaciones se agrupan en dos principales técnicas, transformaciones de intensidad simples y procesado del histograma.

Técnicas de realce en dominio de frecuencia consideran a la imagen digital como una matriz cuyos píxeles representan un valor proporcional al nivel de brillo de la imagen. Sin embargo, la misma imagen puede ser considerada como una función no periódica, y definirse en otro espacio bidimensional cuyos ejes vengán determinados por la amplitud y frecuencia para cada dirección de la imagen. Esto es lo que se conoce como el dominio de la frecuencia. El fundamento de los métodos del dominio de frecuencias es el uso del teorema de convolución. La hipótesis de partida es que la imagen que queremos mejorar fruto del mecanismo de captación de la imagen ha sufrido una convolución con una determinada función definida por las lentes del aparato de medida, el medio en que se tomara la imagen o cualquier otro efecto y que además está contaminada por ruido aleatorio. (Aldalu & Santamaría, 2012, pp. 31-32)

### **2.7.3 Segmentación**

Consiste en dividir la imagen en regiones previo a su análisis, cada región será representada con una etiqueta. Las regiones estarán formadas por píxeles que poseen características afines como niveles de gris similares, variaciones extremas en niveles de grises encontrados, contornos de objetos entre otros. Se pueden distinguir dos técnicas de segmentación, las que se basan en las diferencias entre niveles de grises como detección de puntos, líneas, bordes, entre otras y las que se basan en las similitudes como la umbralización, la división y fusión de regiones, el crecimiento de regiones por agregación de píxeles. (Alegre et al., 2003, pp. 108-119)

### **2.7.4 Representación y descripción**

Después de segmentar en regiones la imagen y etiquetarlas se procede a representar y describir el conjunto de píxeles segmentados de una manera adecuada, la representación de una región se la puede realizar de dos maneras.

- a. En términos de sus características externas (contorno), se utiliza cuando se tiene interés en características de la forma se pueden aplicar las técnicas de códigos de cadena, lados del contorno, representaciones poligonales, firmas, esqueleto, entre otros.

- b. En términos de sus características internas (píxeles de la región), se da cuando se necesita analizar las propiedades internas como color textura, etc. Se puede trabajar con técnicas como la topológica, ejes y áreas.

La elección de un esquema de representación consiste en hacer que los datos sean útiles en una computadora. Es decir, una región puede estar representada por su límite, y el límite puede describirse por características tales como su longitud y el número de concavidades que contiene. Se puede trabajar con los dos tipos de representaciones de acuerdo a las necesidades del problema a resolver. A partir de ello se obtiene los descriptores del objeto. Las características seleccionadas como descriptores deben ser tan insensibles como sea posible a las variaciones en el tamaño de la región, la traducción y la rotación. Ciertas técnicas usadas son la de longitud de contorno, número forma y descriptores de Fourier entre otros. (Alegre et al., 2003, p. 120)

### **2.7.5 Reconocimiento**

En esta etapa se reconoce el objeto u patrón de acuerdo a las necesidades del sistema, los enfoques para el reconocimiento de patrones computarizado se pueden dividir en dos áreas principales:

- a) Teoría de la decisión, trata de patrones descritos usando descriptores cuantitativos (longitud área textura, etc.)
- b) Estructural, trata sobre patrones mejor representados por información simbólica, como cadenas, y descritos por las propiedades y relaciones entre esos símbolos.

El aprendizaje a partir de patrones es la base del reconocimiento de objetos, en esta etapa se utilizan técnicas de aprendizaje tanto para la teoría de la decisión como para los enfoques estructurales tales como distancias, correspondencias de cadena, estadísticas, redes neuronales, correlación.

### **2.7.6 Interpretación**

Consiste en interpretar o asignar un significado de acuerdo a los objetos reconocidos, se identifica la escena mediante diferentes técnicas sistema de producción, redes semánticas y sistemas expertos, entre otros.

## 2.8 Detección de Objetos

La detección de objetos es un procedimiento de visión por computadora que permite la localización de diferentes tipos de objetos ya sea dentro de imágenes o videos mediante varias técnicas de procesamiento. Actualmente existen métodos de visión por computadora usadas para la detección de objetos de acuerdo a las aplicaciones y los recursos disponibles. A continuación, se explican algunas técnicas de detección y descriptores de objetos.

### 2.8.1 *Témlate Maching*

Es una técnica de detección de objetos, que consiste en la detección de pequeñas partes de una imagen que concuerden con una imagen establecida denominada plantilla o patrón, el reconocimiento se realiza por emparejamiento: las imágenes, o regiones de imagen, se establecen en comparación con las plantillas comparándolas de tal forma que su apariencia (imagen) juega un papel evidente y primordial. (Brunelli, 2009, p. 43)

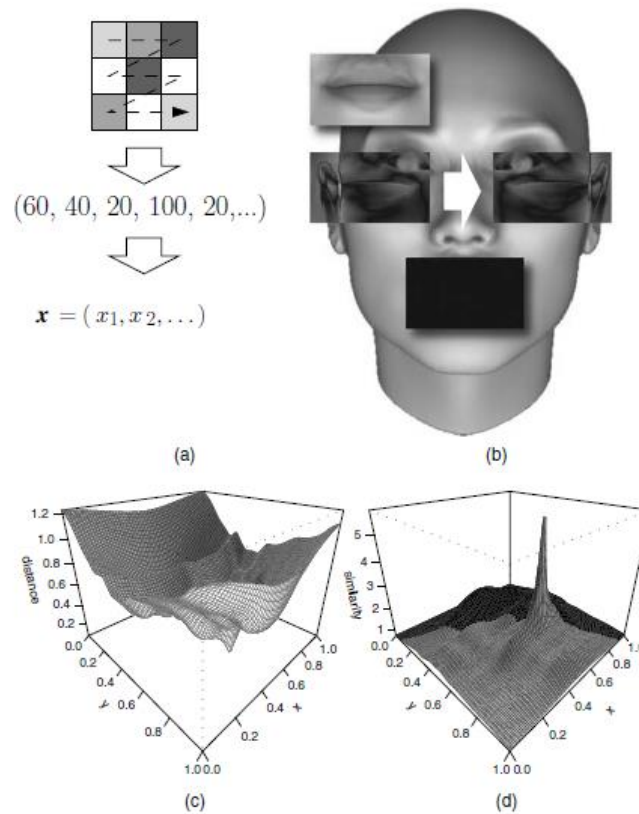
En una plantilla no todos sus valores de intensidad (imagen) son exactamente iguales al objeto a identificar, algunas presentan diferentes perturbaciones que pueden ser provocadas por ruido aditivo, blanco o de sal y pimienta comunes en la adquisición de datos, igualmente los puntos de vista de un objeto pueden generar variación, así como la iluminación y el sensor de imágenes usado para la captura de la misma.

La técnica de coincidencia de plantillas más simple, consiste en la representación de una plantilla (valores de intensidad de luz) en forma de vector ( $x$ ) que se puede comparar calculando su distancia en el espacio vectorial asociado con una imagen congruente representada como el vector ( $y$ ). (Brunelli, 2009, pp. 1-3) en las siguientes ecuaciones se describen el cálculo de menor distancia y similitud para la toma de decisión. Un valor pequeño de  $d(x, y)$  o un valor alto de  $s(x, y)$  es indicativo de similitud de patrones. Si  $x$  es representativo de la plantilla, se busca otras instancias de la misma superponiéndola a otras imágenes, o partes de las mismas, buscando las ubicaciones de menor distancia  $d(x, y)$  o mayor similitud  $s(x, y)$ .

$$d(x, y) = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2 = \frac{1}{N} \|x - y\|^2 \quad (2)$$

$$s(x, y) = \frac{1}{1 + d(x, y)} \quad (3)$$

En la Figura 4-2. se puede visualizar como las plantillas son representadas a manera de vectores (a) y se comparan calculando su distancia en el espacio vectorial asociado. La plantilla se mueve sobre la imagen, como una ventana deslizante (b), y la diferencia entre la plantilla y la imagen se cuantifica utilizando la Ecuación (2), buscando el valor mínimo (c), o la Ecuación (3) , buscando el valor máximo (d).



**Figura 4-2.** Técnica de coincidencia de plantillas.

Fuente: (Brunelli, 2009)

Este método puede funcionar muy bien en entornos controlados donde la sensibilidad, iluminación y posición del objeto se encuentran controlados ya que esta técnica toma en cuenta la forma del objeto en forma total, es decir todos los pixeles de la plantilla como propiedad de la plantilla.



### 2.8.1 *Machine Learning*

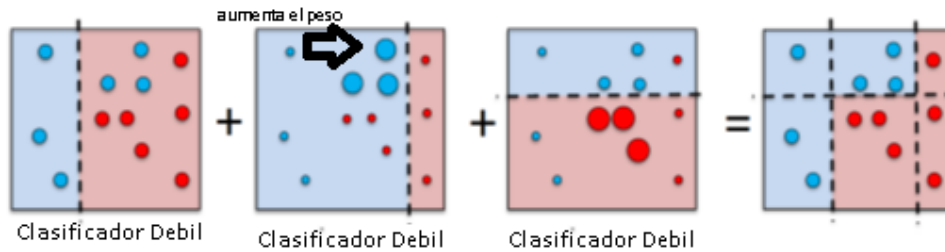
El enfoque estándar de *machine learning* inicia con la recopilación de imágenes y la selección de las características más relevantes de cada imagen. Las mismas se añaden a un modelo de *machine learning*, el cual las divide según la categoría a la que pertenece, utiliza esa información durante el análisis y la clasificación de nuevos objetos. La combinación de diversos algoritmos y métodos de extracción de características pueden crear un modelo de reconocimiento de objetos preciso. (The MathWorks, Inc., 2019) Las técnicas más utilizadas de *machine learning* son:

- Extracción de características HOG (*Histogram of oriented gradients*) con un modelo SVM (Máquinas de vectores de soporte) de *machine learning*.
- Modelos *bag-of-words* con características tales como SURF (*Speeded up Robust Features*) y MSER (*Maximally Stable Extremal Regions*).
- El algoritmo *Viola-Jones*, que se puede utilizar para reconocer una serie de objetos, incluidas caras y torsos.

Los recursos determinantes para elegir entre *Machine Learning* y *Deep Learning* es una GPU potente y gran cantidad de imágenes de entrenamiento etiquetadas. Para trabajar con las técnicas *Deep Learning* son necesarios los dos recursos; *machine learning* puede ser la mejor opción si no se cuenta con alguno de los dos recursos.

### 2.8.2 *Detección de objetos por Cascada*

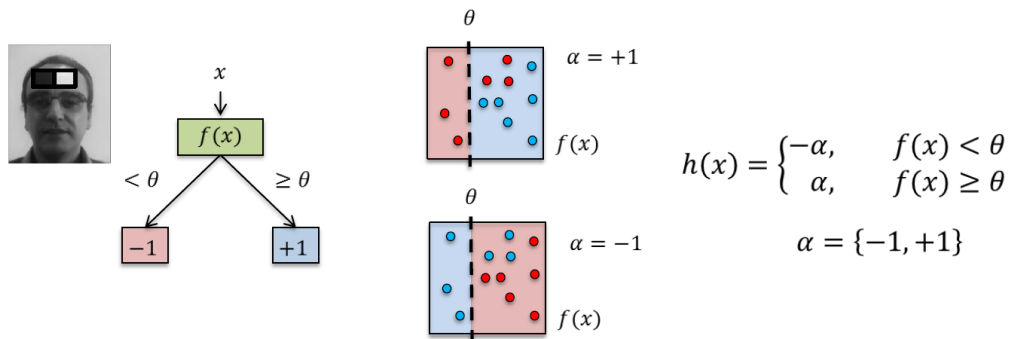
Es una técnica de detección que usa un clasificador en cascada que consta de etapas, donde cada etapa es un conjunto de aprendizaje de clasificadores débiles. Los clasificadores débiles son clasificadores simples construidos a partir de alguna característica del objeto, se denominan *Decision stump*. Cada etapa se entrena mediante una técnica llamada Boosting, el mismo que es un algoritmo de aprendizaje automático que consiste en la combinación de clasificadores simples o débiles para obtener un clasificador robusto, cada clasificador simple aprende dando un peso diferente a cada ejemplo, dando un peso mayor a los ejemplos mal clasificados por los clasificadores previos y un peso menor a los ejemplos bien clasificados, en la figura 5-2 se muestra un ejemplo simple de este método. (López, 2016)



**Figura 5-2.** Representación de *Boosting*

Fuente: (Valveny, 2016)

*Decision stump* es un árbol de decisión binario de profundidad 1, su objetivo es el cálculo de una función de característica para cada imagen de entrada y fijar un umbral para decidir los ejemplos positivos y negativos, desentendiéndose de una sola característica de un descriptor o clasificador los más comunes son LBP (*Local Binary Pattern*), HOG y HAAR usados en detección de rostros y peatones. En la siguiente figura se puede apreciar un ejemplo de un clasificador débil que depende de una sola característica de HAAR.



**Figura 6-2.** Clasificador Débil.

Fuente: (Valveny, 2016)

### 2.8.3 HOG

Es un descriptor de características basado en la información de la orientación del gradiente de áreas locales de una imagen (Contorno de la imagen). El gradiente es un cambio de la intensidad de la imagen en una cierta dirección, en el que el cambio de dirección es máximo, está compuesto por la dirección y la magnitud del cambio de dirección máximo.

El gradiente de un pixel, se calcula obteniendo la diferencia de intensidad de los pixeles vecinos horizontales y verticales, como se muestra en la ecuación (4) donde I es la intensidad dx es la diferencia de intensidad horizontal dy vertical.

$$\begin{aligned} dx &= I(x + 1, y) - I(x - 1, y) \\ dy &= I(x, y + 1) - I(x, y - 1) \end{aligned} \quad (4)$$

A partir de estas definiciones se puede obtener la dirección  $\theta$  y magnitud  $g$  de la orientación del gradiente como se ilustra en la siguiente ecuación.

$$\begin{aligned} \theta(x, y) &= \arctan \frac{dx}{dy} \\ g(x, y) &= \sqrt{dx^2 + dy^2} \end{aligned} \quad (5)$$

El cálculo del gradiente se puede realizar tanto en imágenes a color o en escala de grises, en las imágenes a color se realiza el mismo proceso en cada una de las los canales R G B y luego se selecciona el gradiente del canal con magnitud mayor.

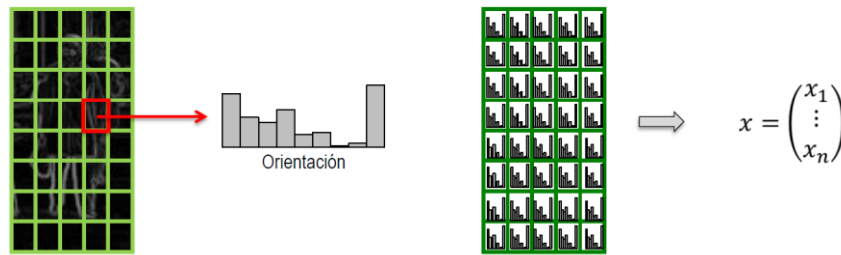


**Figura 7-2.** Representación del Gradiente

**Fuente:** (Valveny, 2016)

El gradiente nos ayuda obtener información de la forma del objeto, pero debido a que esta representación es a nivel de pixel es difícil utilizarlo en un clasificador por ello es necesario convertir el gradiente de los pixeles en una representación global de toda la imagen, en forma de vector de características que capture la forma global del objeto, es así que HOG logra esta representación mediante los siguientes pasos: Primero divide la imagen en celdas y de cada una de ellas se obtiene un histograma de las orientaciones de gradientes. A continuación, se combinan todos los histogramas

de las celdas para obtener la representación global de toda la imagen en forma de vector de características, este proceso se puede visualizar en la figura 8-2. así como en la ecuación (7).



**Figura 8-2.** Histogramas de Orientación de Gradientes

Fuente: (Valveny, 2016)

$$\omega_k(x, y) = \max \left( 0, 1 - \frac{\theta(x, y) - \theta_k}{\delta\theta} \right) \quad (6)$$

$$h_{ij}(k) = \sum_{(x,y) \in C_{ij}} \omega_k(x, y) g(x, y) \quad (7)$$

Dónde:

$h(k)$ : el histograma en el intervalo  $k$

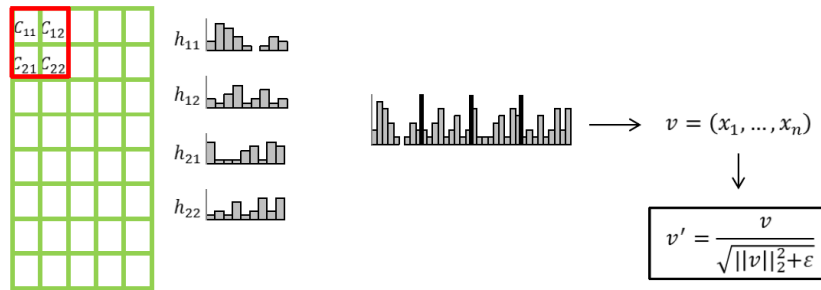
$C$ : celda

$\omega_k(x, y)$ : factor de asignación

$\delta\theta$  rango de los intervalos de los valores de los histogramas.

$g(x, y)$ : magnitud del gradiente

Debido a que es necesario conseguir la máxima invarianza posible en cualquier descriptor se hace necesario la introducción de bloques en el cálculo de histogramas, pues los cambios de iluminación hacen que la intensidad de los pixeles cambia el valor de los histogramas generando varianza en la magnitud de los gradientes de la celda para reducir este problema es necesario una normalización local a cada una de las zonas de la imagen denominas bloques que no es más que una agrupación de celdas. La normalización de los histogramas se realiza a nivel de bloque utilizando la norma L2. (López, 2016)

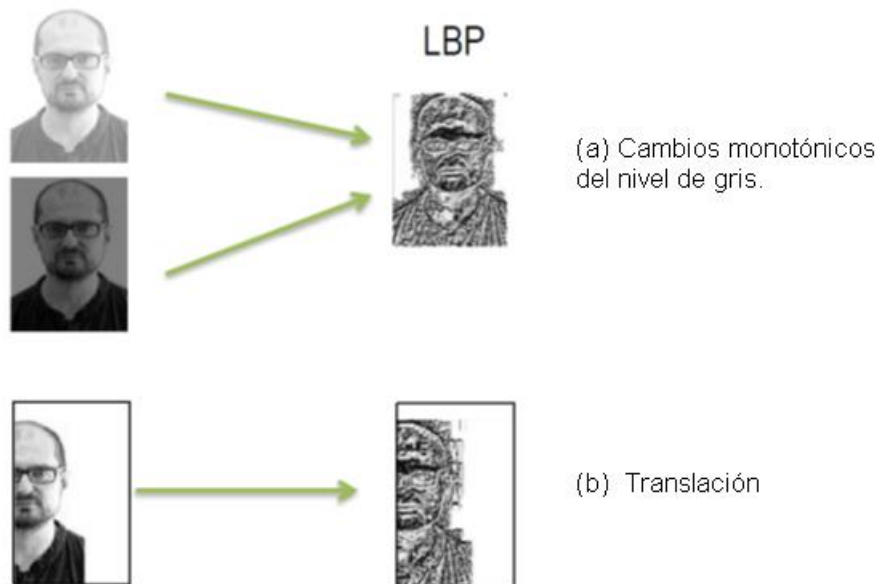


**Figura 9-2.** Normalización de Bloques

Fuente: (Valveny, 2016)

### 2.8.4 LBP

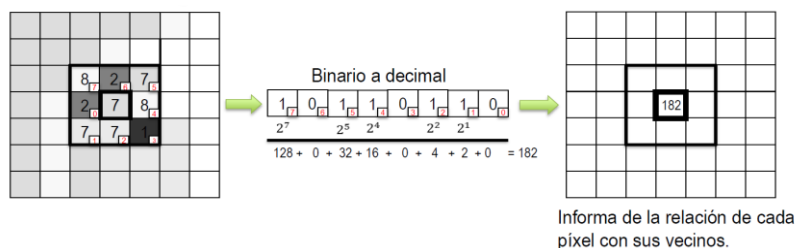
Local Binary Pattern es un descriptor que toma la relación de cada pixel con sus vecinos usando códigos binarios, se usa para identificar propiedades locales y características de partes individuales de una imagen. Es un método poderoso y sencillo con excelentes resultados, utilizado en varios estudios, especialmente en el reconocimiento facial y segmentación de texturas. Una de sus ventajas es la invarianza a la traslación (b) y a los cambios monótonos del nivel de gris (a), ya que cada pixel depende solamente de los vecinos, se puede observar estas propiedades en la figura 12-2.



**Figura 10-2.** Invarianza en Local Binary Pattern.

Fuente: (Valveny, 2016)

LBP básico consisten en tomar un pixel que será el pixel central y definir una vecindad (formada por todos los pixeles vecinos), donde a cada pixel vecino se le aplicar la siguiente regla: Si el valor de intensidad del pixel vecino es mayor o igual al valor del pixel central se dará el valor de 1, caso contrario será 0, la intensidad del pixel varía entre 0 y 255. Esta regla crea un código binario el mismo que es convertido a decimal tomando el pixel central este valor como el nuevo valor de intensidad. Como se muestra en la figura 11-2.

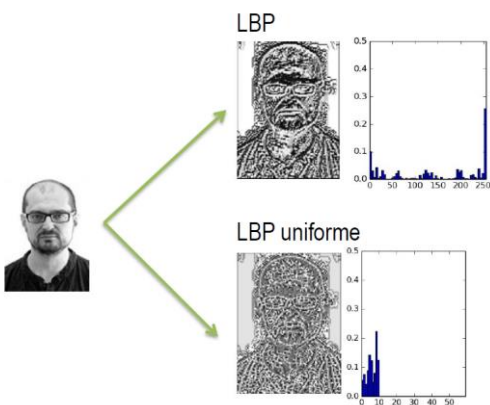


**Figura 11-2.** Local Binary Pattern básico.

Fuente: (López, 2016)

A partir de la idea básica de LBP existen diferentes variantes al definir distintos tipos de vecindades, así como también al determinar un nivel umbral para comparar entre los vecinos logrando ser más robustos a pequeñas variaciones en el nivel de gris debidos a ruidos que pueden aparecer en la adquisición de los datos.

LBP Uniforme es una de las variantes de LBP básico que consiste en disminuir el número de patrones debido a que hay muchos patrones que no son utilizados y se reducen a los que más se utilizan, para ello se define una medida de uniformidad U que se define como el número de transiciones 0-1 en la representación binaria.



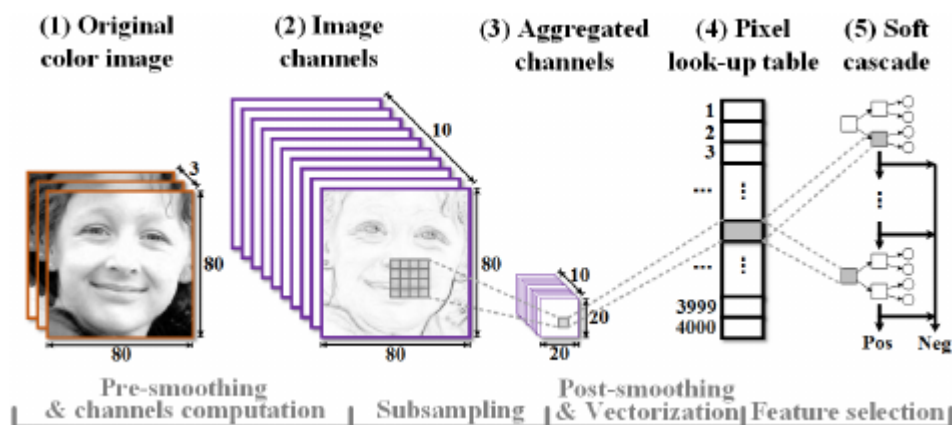
**Figura 12-2.** Histogramas de LBB y LBP Uniforme.

Fuente: (López, 2016)

### 2.8.5 Aggregate Channel Features (ACF)

El detector ACF es una variación de *Integral Channel Features* (ICF) que consiste en extraer características directamente como valores de píxeles en canales extendidos generando varios canales de imagen a partir de una imagen de entrada. Los canales comúnmente utilizados son RGB en las imágenes de color, pero se pueden agregar muchos otros canales. Se puede definir un canal como un mapa registrado de la imagen original donde los píxeles de salida se asignan a los píxeles de entrada mediante alguna transformación lineal o no. (Dollár et al., 2009, p. 4)

Se pueden extraer varias características de estos canales extendidos, se han definido dos tipos de canales: De primer orden, que consisten en extraer características de un solo canal sumando píxeles en regiones rectangulares fijas, denotadas como  $f(C)$ . (Dollár et al., 2009, p. 4) De orden superior, que son funciones obtenidas mediante la combinación de dos o más funciones de cada canal de primer orden.



**Figura 13-2.** *Aggregate channel features* para la detección de rostro.

Fuente: (Bin et al., 2014)

Los canales se vuelven a calcular a múltiples escalas para extraer un conjunto de características de canal que pueden representar todo el espacio de escala. ACF tiene ventajas, como una representación más rica, una velocidad de detección acelerada y una localización más precisa de los objetos en las imágenes cuando se usa junto con un método de refuerzo. (Monedero, 2017, p. 15)

### **2.8.6 *Deep Learning***

Un modelo claro de lo que es *Deep Learning* son las redes neuronales convolucionales CNN, que aprenden automáticamente las características inherentes de un objeto a fin de identificarlo. Existen dos enfoques en cuanto al reconocimiento de objetos mediante *Deep Learning*:

- El primer enfoque es el entrenamiento de un modelo desde cero que recopila una gran cantidad de datos etiquetados, luego diseña una arquitectura de red que permite el aprendizaje de las características principales y se crea el modelo. La desventaja de este enfoque es el requerimiento de gran cantidad de datos para el entrenamiento, y con ello la configuración de las capas y las ponderaciones en la CNN. (The MathWorks, Inc., 2019)
- El segundo enfoque es un modelo de *deep learning* previamente entrenado, los mismos emplean transferencia del aprendizaje que implica un proceso de ajuste de un modelo previamente entrenado. Inicia con una red existente, como por ejemplo AlexNet o GoogLeNet, proporcionando información nueva que contiene clases previamente desconocidas. Este método disminuye el tiempo y proporciona resultados más rápidos. (The MathWorks, Inc., 2019)

### **2.8.7 *Otros métodos de reconocimiento de objetos***

En función de la aplicación resulta mejor utilizar enfoques más básicos de reconocimiento de objetos como:

- Coincidencia de plantillas en donde se emplea una imagen pequeña o plantilla, para localizar regiones coincidentes en una imagen más amplia.
- Segmentación de imágenes y análisis de blobs que utilizan propiedades simples de los objetos, tales como tamaño, color o forma.



## CAPÍTULO III

### 3 DISEÑO DEL *HARDWARE* Y *SOFTWARE* DEL SISTEMA

En este capítulo se presenta el diseño de *hardware* y *software* del sistema para el procesamiento digital de imágenes en detección de material contaminante inorgánico en orillas de lagunas. De acuerdo a los requerimientos que debe cumplir esta investigación se realiza el diseño en bloques en cada uno de sus módulos, se justifica la selección de los componentes que lo integran, así como el *software* a implementar.

#### 3.1 Requerimientos del hardware

De acuerdo al estudio teórico realizado en el capítulo anterior se puede definir los requerimientos del diseño. Estos son:

- Cámara digital de alta resolución con capacidad 4K y HD (*High Definition*). El contaminante inorgánico a reconocer son las botellas de plástico para ello las imágenes deben ser capturadas con una alta resolución para obtener las mayores características de la forma de la botella.
- Vehículos aéreos no tripulados (VANT). Para obtener una mayor visibilidad e iluminación del área contaminada y capturar imágenes HD es necesario la tomarlas desde el aire, durante el día especialmente en momento donde el sol no ilumine completamente a la botella, además debe proporcionarse el área a fotografiar, siendo el *drone*, la mejor opción para la captura de las imágenes.
- Teléfono Inteligente. Debido al uso de un *drone* se hace necesario el control de este equipo a través de un teléfono Inteligente o tableta.
- Unidad de almacenamiento de gran capacidad, las imágenes previamente capturadas por su alta resolución poseen un tamaño especialmente grande, además deben ser procesadas y almacenadas en un disco duro de al menos 500GB.

- Ordenador de alto rendimiento. El sistema necesita entrenamiento previo con gran cantidad de imágenes y la resolución de la imagen debe ser robusta, por lo que el ordenador debe tener las siguientes características:
  - o procesador Intel I7.
  - o 32 GB de memoria RAM.
  - o Disco duro SSD de al menos 128 GB
  - o Tarjeta gráfica 2080 Ti.

### **3.2 Requerimientos de *Software***

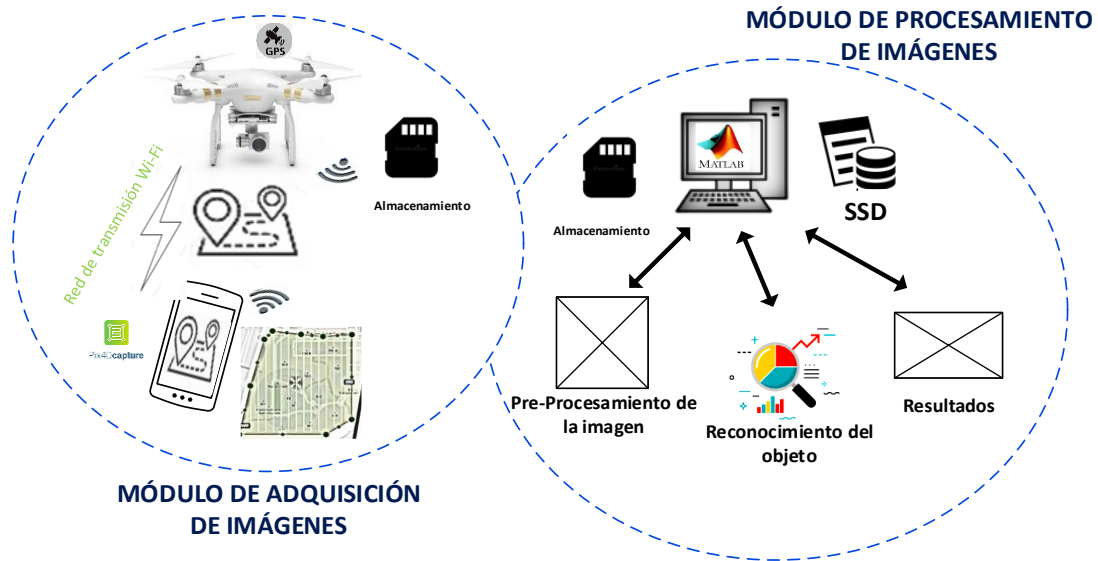
En los proyectos de investigación la etapa más importante es la elección del *software* a utilizar ya que el mismo determina el funcionamiento correcto del sistema a desarrollar. A continuación, se especifican los requerimientos de *software*.

- *Software* de Planimetría. Para obtener las imágenes es necesario definir el área a sobrevolar, trazar la ruta con un plan de vuelo, que el *drone* tomará. Una vez trazada la ruta se debe enviar la altura, posición de la cámara, velocidad, y solapamiento de imágenes para la captura de las mismas. De cada zona fotografiada se debe guardar las coordenadas, el tamaño de *pixel* en el terreno (GSD) que usará posteriormente para el cálculo del área contaminada. Por ello es necesario el uso de un *software* de planificación de vuelo gratuito.
- Lenguaje de programación capaz de aplicar filtros para retirar el ruido, estandarizar el tamaño de todas las imágenes a un valor determinado y ajustar el contraste, diseñar y ejecutar algoritmos de detección de objetos y capaz de resolver cálculos numéricos básicos. En este apartado se utiliza la plataforma de programación y cálculo numérico MATLAB.

### **3.3 Arquitectura del sistema**

La concepción general propuesta para el sistema se presenta en la figura 1-3, que muestra la arquitectura construida para el procesamiento digital de imágenes en la detección de material contaminante inorgánico en orillas de lagunas. Se presenta una arquitectura de dos módulos el primero de Adquisición de Imágenes y el segundo de Procesamiento de Imágenes.

En la siguiente figura se puede observar la arquitectura del sistema de reconocimiento de contaminación inorgánica (botellas plásticas).



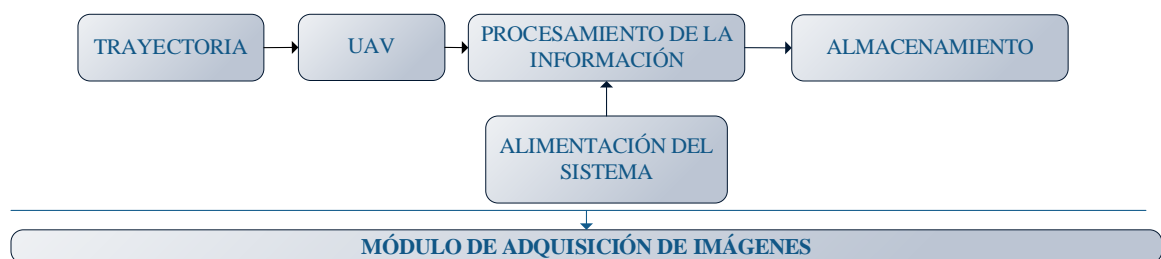
**Figura 1-3.** Arquitectura Sistema

Realizado por: Urgilez M. 2019

### 3.3.1 Módulo de Adquisición de Imágenes

El Módulo de Adquisición de Imágenes es el encargado de la captura y almacenamiento de las imágenes, las mismas que pueden ser guardadas en la nube, tarjeta de memoria u otro sistema de almacenamiento. Está integrado por un UAV (Vehículo Aéreo no Tripulado) para acceder a las zonas de contaminación, el mismo cuenta con una cámara profesional que toma las imágenes pertinentes, a su vez este se encuentra monitorizado con conexión Wi-Fi a un teléfono inteligente que cuenta con la app adecuada para proveerle la información necesaria para la captura de imágenes. De esta manera se recoge la información de manera eficaz, eficiente y con el menor tiempo posible. Una vez obtenidas las imágenes éstas son enviadas al módulo de procesamiento de imágenes mediante conexión. A continuación se presenta en la siguiente figura el diagrama de bloques del módulo de adquisición de imágenes integrado por 5 bloques, el de trayectoria donde consta el área a sobrevolar, la altura y el tipo de trayectoria este proceso se lo puede realizar de una manera sencilla a través de la aplicación *PIX4D Capture*, encargado de enviarle toda la información de vuelo mediante Wi-Fi al bloque UAV, mismo que a través de la cámara profesional que tiene incorporada captura las imágenes, dichas

imágenes son almacenadas en una tarjeta de memoria o también se puede usar la nube si fuese necesario, de esta manera se transmiten al siguiente módulo.

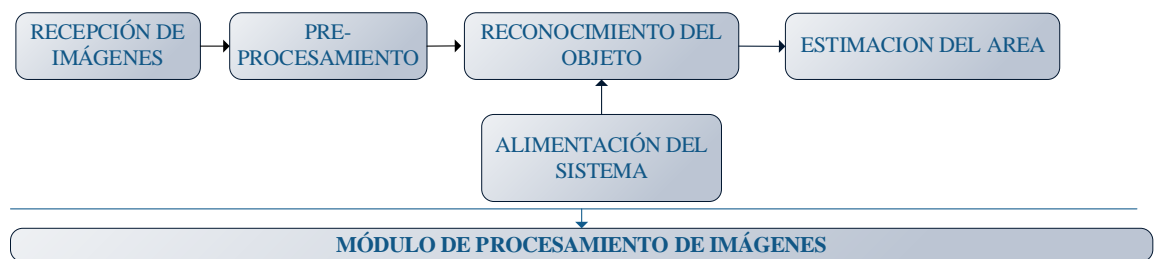


**Figura 2-3.** Diagrama de bloques del módulo de adquisición de imágenes

Realizado por: Urgilez M. 2019

### 3.3.2 Módulo Procesamiento de Imágenes

El módulo de procesamiento de imágenes es el encargado de recibir, preparar, hacer el reconocimiento de los objetos de interés de las imágenes obtenidas en el módulo anterior, para trabajar en este módulo es necesario el uso de un ordenador de alta gama por la gran cantidad de información que se necesita procesar y del *software* que se utiliza. Este módulo está compuesto por 5 bloques, el que recibe las imágenes tomadas y almacenadas, para pasar al pre-procesamiento de las imágenes en donde se retira el ruido, se convierten a un mismo tamaño de píxeles y se mejora el contraste; una vez pre-procesadas pasan al siguiente bloque en donde el sistema experto hace un reconocimiento del objeto con las características relevantes del mismo en este caso botellas de plástico. Identificadas las zonas de contaminación por botellas de plástico se pasa al bloque final que es la generación de un reporte de datos que muestra la imagen etiquetada de todos los objetos contaminantes (botellas de plástico) para proceder a calcular el área que ocupan las botellas en las imágenes y proceder a estimar una aproximación del área contaminada.



**Figura 3-3.** Diagrama de bloques del módulo de procesamiento de imágenes.

Realizado por: Urgilez M. 2019

### 3.4 Selección de los elementos *hardware* del sistema

A continuación, se detalla las características principales de los elementos *hardware* que conforman el sistema de procesamiento digital de imágenes en la detección de material contaminante inorgánico en orillas de lagunas, las especificaciones técnicas de los mismos se pueden ver al final del documento en el (Anexo A).

#### 3.4.1 *Phantom 3 Professional*

Phantom 3 Professional es un cuadricóptero de marca DJI de interfaz amigable posee una cámara digital 4K y HD. Tiene un estabilizador integrado que permite minimizar su peso y tamaño. Trabaja con los sistemas de posicionamiento satelital GPS/GLONASS (*Global Positioning System / Global Navigation Satellite System*) que permiten que se muevan con precisión en caso de tener desactivado uno de ellos., su velocidad máxima de ascenso es 5 m /s, la velocidad máxima de descenso 3 m / s, y presenta una máxima velocidad 16 m / s (modo ATTI), su ángulo de inclinación máxima es 35 °, posee un máximo tiempo de vuelo de 23 minutos. La cámara tiene un sensor 1 / 2.3" CMOS con píxeles efectivos 12.4 M (píxeles totales: 12.76 M) con lente FOV 94 ° 20 mm (formato equivalente a 35 mm) f / 2.8 enfoque a rango ISO 100-3200 (video) 100-1600 (foto), el tamaño de la imagen s 4000 × 3000 con un solo tiro, se dispara en ráfaga 3/5/7 cuadros, Max Video Bitrate 60 Mbps, foto en JPEG, DNG (RAW) y vídeo MP4, MOV (MPEG-4 AVC / H.264), Tarjetas SD compatible Micro SD Capacidad máxima: se requieren 64 GB Clase 10 o UHS-1. (SZ DJI Technology Co. Ltd, 2017)



**Figura 4-3.** *Drone Phantom 3 Professional*

**Fuente:** (SZ DJI Technology Co. Ltd, 2017)

### 3.4.2 Ordenador portátil

En el actual proyecto se hizo necesario el uso de un ordenador de alta gama debido a que en el aprendizaje automático usado como parte del desarrollo del sistema es preciso el uso de un GPU potente, para poder disminuir el periodo de aprendizaje del algoritmo de detección. La duración de este aprendizaje depende de la cantidad de información a procesar, que puede durar días e incluso meses, así mismo el uso de un GPU de gran potencia disminuirá este tipo de proceso a horas o días, mejorando la eficiencia del módulo de reconocimiento y haciendo posible trabajar con un mayor número de imágenes para el aprendizaje del algoritmo, refinando la detección de los objetos a analizar.

Para trabajar con pequeños proyectos (por debajo de 100 imágenes a 14 MP) es necesario 8 GB RAM; 15 GB SSD de espacio libre; para proyectos medianos (entre 100 y 500 imágenes a 14 MP) debe poseer 32 GB RAM, 60 GB SSD de espacio libre; y para proyectos grandes (sobre las 2000 imágenes a 14 MP) su capacidad debe ser de 64 GB RAM, 120 GB SSD de espacio libre. (Martorell, 2019) El ordenador a usar debe tener al menos las siguientes características:

- Windows 8, 10 de 64 bits,
- CPU *quad-core hexa-core* Intel i7/Xeon,
- GeForceGPU compatible con OpenGL 3.2
- 8 GB RAM
- *Hard Disk SSD*.

Es por lo antes expuesto que se usó una computadora portátil SONY VAIO que cuenta con las características necesarias para trabajar con el sistema implementado.



**Figura 5-3.** Ordenador Sony VAIO *core i7*.

Fuente: (Martorell, 2019)

### 3.5 Selección de *software* para el sistema

El sistema de procesamiento de digital de imágenes requiere de *software* especializado, para que en combinación con el *hardware* se correlacionen y satisfagan los requerimientos del sistema. El *software* para manejar los módulos del sistema de procesamiento digital de imágenes en la detección de material contaminante inorgánico en orillas de lagunas, se realiza con programación estructurada y orientada a objetos.

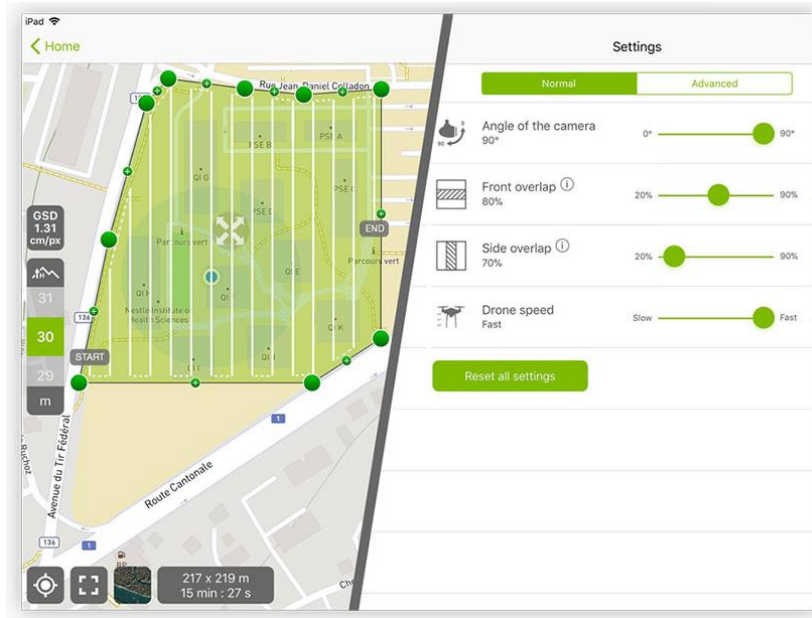
Las herramientas de programación utilizadas para el desarrollo del *software*, se detallan a continuación:

#### 3.5.1 *PIX4D*

Es un *software* de mapeo y fotogrametría que permite la digitalización de escenarios en 2D y 3D con *drones* a través de aplicaciones de vuelo y plataformas en la nube. Cuenta con varias herramientas de *software* para fotometría y mapeo con *drones*, en la figura 6-3 se observan una planificación de vuelo con *PIX 4D Capture*.

Para efectos de desarrollo de este sistema se utiliza la herramienta *PIX 4D Capture* para la planificación de vuelo del *drone*, mapeo y modelado, esta herramienta además cuenta con la compatibilidad para el uso con los diferentes tipos *DRONE* de marca DJI, Parrot y Yuneec. Sus principales características son flexibilidad para el vuelo multicopter, precisión en el momento de definir la altura estableciendo el ángulo de la cámara, almacenamiento directo desde el dispositivo al *software*. (PIX 4D SA., 2019)

En el anexo B se puede visualizar una tabla comparativa de este *software* y otros de similar aplicación. Se identifico a PIX4D como el mejor *software* de fotogrametría en este proyecto debido a su compatibilidad con *Android, Iphone, Ipad*, su fácil manejo y especialmente porque permite trabajar con *drones* de diferentes marcas presentes en el mercado. (Pandrone, 2018)

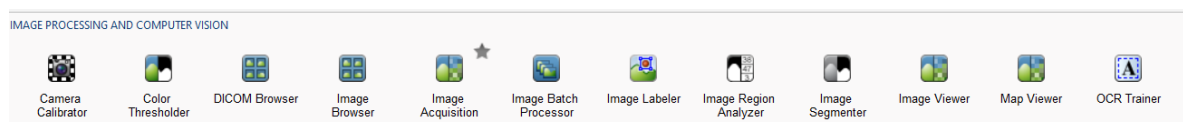


**Figura 6-3.** Herramientas de *software* PIX 4d en los planes de vuelo.

Fuente: (PIX 4D SA., 2019)

### 3.5.2 *Matlab*

Es un lenguaje matemático de alto desempeño diseñado por *The MathWorks, Inc.* que trabaja con una programación basada en matrices y que cuenta con varias Apps dependiendo del tipo de programación que se desee realizar. Entre las principales áreas de estudio con los que cuenta Matlab, se pueden destacar los que se presentan en la figura 7-3. En este proyecto se trabaja con las herramientas disponibles en *Image Processing Toolbox*, capaces de proporcionar algoritmos complejos y apps de flujo de trabajo para el correcto procesamiento, análisis y visualización de imágenes y video, permitiendo el uso de diferentes filtros ya definidos en este lenguaje. Puede llevar a cabo segmentación de imágenes, mejora de imágenes, reducción de ruido, transformaciones geométricas, registro de imágenes y procesamiento de imágenes 3D. Muchas funciones del toolbox soportan la generación de código C/C++ para la generación y desarrollo de sistemas embebidos. (The MathWorks, Inc., 2019)



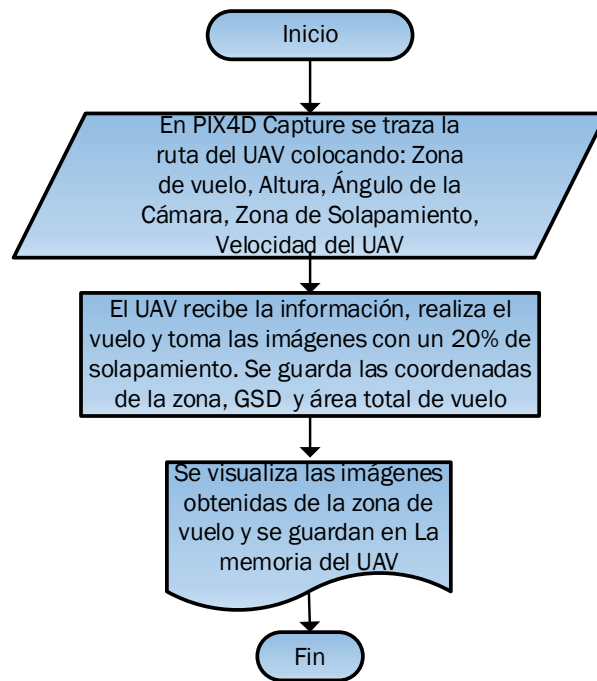
**Figura 7-3.** Herramientas de procesamiento de imagen y visión por computadora de Matlab

Fuente: (The MathWorks, Inc., 2019)



### 3.6 Diagrama de flujo del programa para el Módulo de Adquisición de Imágenes

El Módulo de Adquisición de Imágenes es el encargado de capturar las imágenes de la contaminación presente en orillas de lagunas. En la app *Pix4Dcapture* se coloca lo parámetros de vuelo del UAV: zona de vuelo, altura a la que se va a encontrar el UAV, ángulo de la cámara, velocidad, y solapamiento de imágenes; esta aplicación guarda las coordenadas de la zona y el área de la misma. Una vez colocados los parámetros se inicia el vuelo con el UAV éste captura las imágenes con un 20% de solapamiento; para finalizar las imágenes obtenidos se guardan en la memoria del UAV.



**Figura 8-3.** Diagrama de flujo del módulo de adquisición de imágenes

Realizado por: Urgilez M. 2019

### 3.7 Diagrama de flujo para el Módulo de Procesamiento de Imágenes

El módulo de procesamiento de imágenes contiene el programa principal del sistema de procesamiento digital de imágenes en la detección de material contaminante inorgánico en orillas de lagunas. El mismo se muestra en la siguiente figura y se detalla a continuación:

Encabezado.

- Se declara las librerías necesarias para la utilización de las diferentes sentencias.
- Se declara las funciones pertinentes.

Preprocesamiento: se llama a tres procedimientos

- 1) Modificar el tamaño de la imagen de acuerdo al método de detección a usar.
- 2) Disminución del ruido de la imagen utilizando una red neuronal profunda,
- 3) Mejora del contraste con Distribución uniforme de histogramas sin ruido.

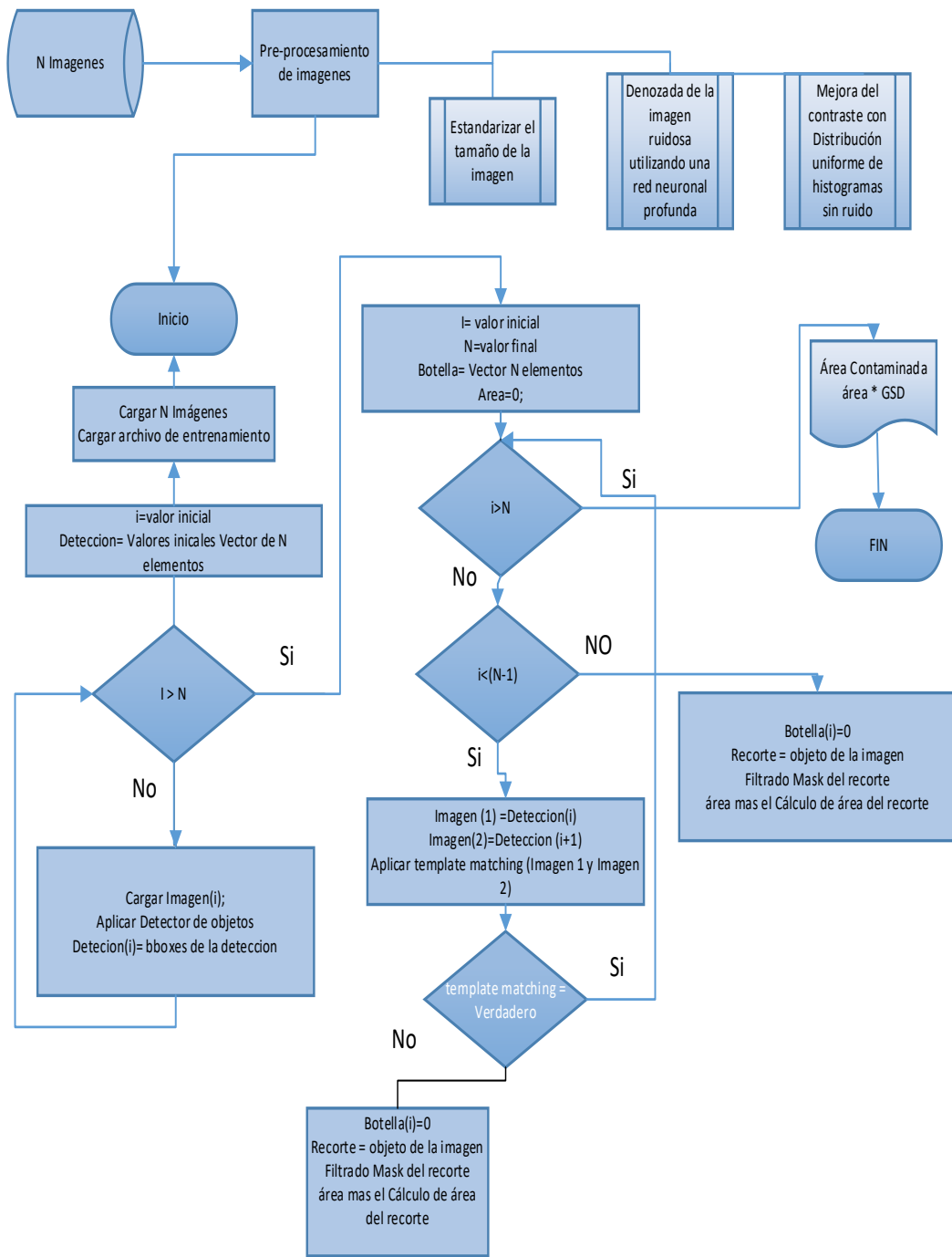
Luego del pre-procesamiento de la imagen se realiza el procesamiento de detección de imágenes el mismo que está comprendido por tres procedimientos de clasificación de la imagen, esto con el objetivo de reconocer el mejor detector para la clase. Se probó 3 tipos diferentes de detectores y se realizaron pruebas y su evaluación e identificó el mejor. En este apartado se describe el sistema implementado usando un detector básico.

Como primer paso del sistema de detección y decisión se carga un número  $N$  de imágenes y el archivo de entrenamiento previamente probado y evaluado, se genera la variable  $i$  inicializada en cero, que permite recorrer todas las imágenes cargadas y la tabla detección donde se registrarán las detecciones (*bbox*) de cada una de las imágenes.

Al terminar la clasificación se procede al análisis de la tabla, para ello se utiliza un nuevo bucle que recorre las detecciones y toma la imagen en esa posición y la imagen adyacente esto se hace para evitar la redundancia de objetos en las imágenes.

Se verifica que los objetos detectados no se encuentren en ambas imágenes, para ello se aplica el detector de plantillas *Template matching* quien me devuelve un true o false, si el objeto está en las dos imágenes, si ese es el caso se sigue a la siguiente imagen. En caso de que no se cumpla la condición se recorta el área definida por los *bbox* de la botella detectado y se aplica una función de filtros y binarización para calcular el área de la botella detectada.

El área se ira sumando según avance el índice  $i$  por cada una de las celdas de la tabla detección. Cuando haya llegado al final de la tabla detección se visualizará el área total de pixeles ocupados en la imagen. Para definir el área aproximada en unidades de medida de longitud es necesario su multiplicación por el GSD definido en la adquisición de datos por *PIX 4D Capture*, de esta manera se puede obtener una estimación del espacio ocupada por las botellas.



**Figura 9-3.** Diagrama de bloques del módulo de procesamiento de imágenes

Realizado por: Urgilez M. 2019

## CAPÍTULO IV

### 4 PRUEBAS Y RESULTADOS

En este capítulo se explicarán los métodos y, pruebas experimentales realizadas para evaluar el detector de objetos, para ello se comparará las medidas de precisión y sensibilidad obtenidas en los diferentes detectores. Se analiza el detector en Cascada con los descriptores HOG y LBP, así como el Detector ACF.

#### 4.1 Preprocesamiento

Se utiliza un preprocesamiento de las imágenes previo su uso, cambiando el tamaño de las imágenes a una dimensión estándar (1080x1080 o 416x416) de acuerdo a las necesidades del detector, se reduce el Ruido de la imagen con la herramienta de Matlab *denoisingNetwork* una red neuronal convolucional pre-entrenada de eliminación de ruido.'DnCNN' y se ajusta el contraste de la imagen aplicada en las imágenes a color.



**Figura 1-4.** Preprocesamiento

**Realizado por:** Urgilez M. 2019

Previo al entrenamiento del clasificador objetos en Cascada en Matlab se necesitó trabajar de la siguiente manera:

- Las imágenes se clasifican en dos carpetas una con las imágenes del objeto a buscar y una carpeta con las imágenes de fondo.
- La carpeta que contiene las imágenes positivas se divide en dos partes, una parte para el entrenamiento y otra parte para la evaluación del detector. Para la separación de las imágenes se creó una función llamada carpeta que consiste en tomar 80% de las imágenes positivas al azar y colocarlas en una carpeta denominada *train*, y el 20% restante en una carpeta denominada *test*. Los que se utilizaron para el entrenamiento y validación respectivamente.
- Mediante el uso de la app de Matlab *Image Labeler* se etiqueta las regiones de interés rectangulares (ROI o BBox) de las imágenes de entrenamiento y evaluación, las mismas que se guardarán en un archivo tipo mat que contiene la información de las etiquetas para cada imagen.

	1 imageFilename	2 Botellas
1	'G:\reconocimiento de objetos matlab\detección botellas\train\ima38.bmp'	[1137,106,95,122]
2	'G:\reconocimiento de objetos matlab\detección botellas\train\ima56.bmp'	[980,91,85,97]
3	'G:\reconocimiento de objetos matlab\detección botellas\train\ima130.bmp'	[279,7,72,102;670,40,127,64]
4	'G:\reconocimiento de objetos matlab\detección botellas\train\ima167.bmp'	[279,2,79,82]
5	'G:\reconocimiento de objetos matlab\detección botellas\train\ima185.bmp'	[281,31,100,103]
6	'G:\reconocimiento de objetos matlab\detección botellas\train\ima222.bmp'	[501,86,106,102;985,5,156,40]
7	'G:\reconocimiento de objetos matlab\detección botellas\train\ima259.bmp'	3x4 double
8	'G:\reconocimiento de objetos matlab\detección botellas\train\ima295.bmp'	3x4 double
9	'G:\reconocimiento de objetos matlab\detección botellas\train\ima314.bmp'	3x4 double
10	'G:\reconocimiento de objetos matlab\detección botellas\train\ima332.bmp'	3x4 double
11	'G:\reconocimiento de objetos matlab\detección botellas\train\ima351.bmp'	3x4 double
12	'G:\reconocimiento de objetos matlab\detección botellas\train\ima387.bmp'	3x4 double
13	'G:\reconocimiento de objetos matlab\detección botellas\train\ima406.bmp'	3x4 double
14	'G:\reconocimiento de objetos matlab\detección botellas\train\ima424.bmp'	3x4 double
15	'G:\reconocimiento de objetos matlab\detección botellas\train\ima479.bmp'	[609,457,89,161;1003,20,166,138]
16	'G:\reconocimiento de objetos matlab\detección botellas\train\ima498.bmp'	[662,579,78,142;1112,122,169,156]
17	'G:\reconocimiento de objetos matlab\detección botellas\train\ima516.bmp'	[1112,272,169,154]
18	'G:\reconocimiento de objetos matlab\detección botellas\train\ima535.bmp'	[1117,416,164,177;634,1,91,60]

**Figura 2-4.** Tabla de etiquetado en Matlab

Realizado por: Urgilez M. 2019

- La aplicación *Image Labeler* admite todos los formatos de archivo de imagen admitidos por *imread*.

## 4.2 Entrenamiento de los detectores

Se crean algoritmos de entrenamiento con 2079 imágenes positivas y 2254 imágenes negativas entrenados con Clasificador de objetos en Cascada y descriptores HOG, LBP y ACF. Es importante

revisar que las imágenes del fondo o negativas no contengan ninguna de las imágenes del objeto para que el entrenamiento sea el adecuado, así mismo el etiquetado de las botellas debe ser el correcto para obtener un buen aprendizaje.

Para el entrenamiento del detector en Cascada con descriptores HOG y LBP, se siguen los siguientes pasos:

- Se cargan las imágenes positivas con su tabla de etiquetas (ROI) ya definidas previamente y se especifica el nombre de la carpeta que contiene las imágenes negativas o de fondo.
- Se establece el número de etapas y tasa de falsos positivos con el que se va a entrenar al detector.
- Se procede a crear el algoritmo de detección mediante el entrenamiento del detector con las imágenes ya cargadas en el lenguaje de programación.
- Finalmente se obtiene un archivo de entrenamiento en formato xml el mismo que se usara posteriormente para la detección de objetos.

Para el entrenamiento del detector ACF, se siguen los siguientes pasos:

- Se cargan las imágenes positivas con su tabla de etiquetas (*BBox*) ya definidas previamente y se especifica el nombre de la carpeta que contiene las imágenes positivas.
- Se establece el número de estados y Factor de muestras negativas con el que se va a entrenar al detector.
- Se procede a crear el algoritmo de detección mediante el entrenamiento del detector con las imágenes ya cargadas en el lenguaje de programación. Para el entrenamiento se utiliza instancias positivas de los objetos en las imágenes proporcionadas en la tabla de etiquetas y recopila automáticamente instancias negativas de las imágenes durante el entrenamiento
- Finalmente se obtiene un archivo de entrenamiento en formato xml el mismo que se usara posteriormente para la detección de objetos.

Para la detección de objetos en imágenes se sigue los siguientes pasos:

- Leer una imagen de entrada: esta es la imagen que se desea analizar y debe de estar en formato aceptados por la librería OpenCV en este proyecto se usaron imágenes .jpeg.
- Cargar el archivo con la cascada de clasificadores entrenada *BOOTS*, el que debe estar definido por el número de etapas y tasa de falsos positivos usados en su entrenamiento, para llevar así, un control de datos para su evaluación posterior.

- Para la detección se trabaja con el método de ventana deslizante, el mismo que se pasara por toda la imagen analizando las características obtenidas en el entrenamiento y la de la región de interés, en caso de que la imagen no contenga las características definidas por el entrenamiento se procede a descartarla, en caso de que la región de interés tenga similitud con las características del entrenamiento se procede a continuar en la detección , la misma que volverá realizar un análisis con clasificadores débiles o canales de características, la imagen que llegue hasta el final de las etapas se la define como Imagen positiva y se precisa la región ROI.
- El número de etapas o estados nos indica las veces que se detecta una zona de la imagen como positiva para finalmente establecer que existe el objeto en dicha zona.

#### **4.2.1 Detectores de Botellas**

##### *4.2.1.1 Detector de objetos en cascada*

En cada etapa del clasificador se etiqueta la región definida por la ubicación actual de una ventana deslizante como positiva o negativa. Positiva si se encontró la botella y negativa en caso contrario. Si la etiqueta es negativa, la clasificación de esta región está completa y el detector desliza la ventana a la siguiente ubicación. Si la etiqueta es positiva, el clasificador pasa la región a la siguiente etapa. El detector determina que se ha encontrado un objeto cuando la etapa final clasifica la región como positiva. Con esta definición se pueden obtener estos resultados:

- Reales positivos ocurre cuando una botella se clasifica correctamente.
- Falso positivo cuando se clasifica erróneamente una región.
- Falso negativo cuando una botella se clasifica erróneamente como negativo.

Para que funcione bien, cada etapa de la cascada debe tener una tasa baja de falsos negativos. Si una etapa etiqueta incorrectamente un objeto como negativo, la clasificación se detiene y no se puede corregir el error. Sin embargo, cada etapa puede tener una alta tasa de falsos positivos. Incluso si el detector etiqueta incorrectamente un no objeto como positivo, puede corregir el error en etapas posteriores.



**Figura 3-4.** Detección HOG

**Realizado por:** Urgilez M. 2019

#### 4.2.1.2 Detector ACF

El detector de objetos ACF utiliza características de canal agregado para la detección de botellas, este se basa en el previo entrenamiento establecido utilizando una estructura donde se encuentran la ubicación de las imágenes de entrenamiento. El detector devuelve un *BBox* con la ubicación de la etiqueta y confianza



**Figura 4-4.** Detección ACF

**Realizado por:** Urgilez M. 2019

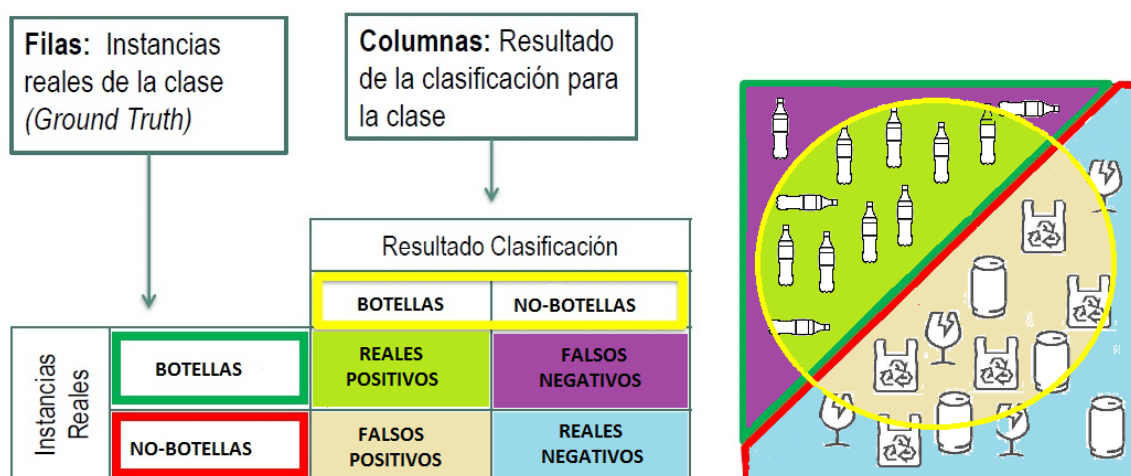


## 4.2.2 Evaluación de los detectores

Para la evaluación del detector se va integrar los conceptos de matriz de confusión, precisión y sensibilidad.

### 4.2.2.1 Matriz de confusión

Consiste en una matriz de datos que permite determinar variables de medida y evaluación del detector. La matriz está compuesta por las filas (Instancias reales de la clase) y columnas (Resultados de la clasificación). En la figura 5-4. se observa un ejemplo de la matriz de confusión compuesta por los reales positivos, que son todas las botellas detectadas correctamente, los falsos positivos conformados por todos los objetos que no son botellas y el sistema lo detectan como una botella, los falsos negativos que representan todas las botellas que no fueron detectadas y los reales negativos conformados por los objetos que no son botellas y no las detectaron como botellas. En este caso no se puede definir los valores de los Reales negativos porque solo se tiene la detección de una clase en este caso las botellas.



**Figura 5-4.** Matriz de Confusión

Realizado por: Urgilez M. 2019

#### 4.2.2.2 Cálculo de la sensibilidad y precisión

La sensibilidad se define como la eficiencia en la clasificación de todos los elementos que son de la clase. Se calcula con la Ecuación(8) (López, 2016)

$$\text{Sensibilidad} = \frac{\text{Reales positivos}}{\text{Reales positivos} + \text{Falsos Negativos}} \quad (8)$$

La precisión indica la calidad de respuesta del clasificador, definido con la ecuación (9).

$$\text{Precisión} = \frac{\text{Reales positivos}}{\text{Reales positivos} + \text{Falsos Positivos}} \quad (9)$$

### 4.3 Resultados

A continuación, se presentan los resultados de la evaluación de los algoritmos de detección: Detector en cascada con descriptores HOG, detector en cascado con descriptores LBP y finalmente el detector ACF. Para ello se realizaron los entrenamientos de los algoritmos con diferentes parámetros de entrada, y se determinó los valores del número de etapas y la tasa de falso positivos que dan mejor resultado con cada detector.

Se realizó la Evaluación de detector en Cascada con descriptores HOG con la ayuda de una función llamada evaluando detector, programada en el entorno de Matlab. Los resultados se presentan en la tabla 1-4. se puede distinguir que el mejor detector se da con un número de etapas de 30 y una tasa de falsos positivos de 0.25 dando como resultado una precisión del 70%. y una sensibilidad del 23%

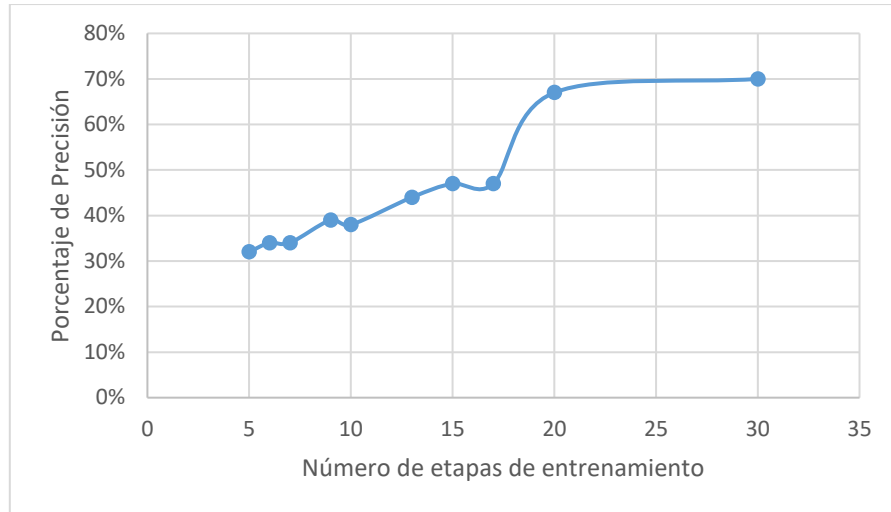
**Tabla 1-4:** Datos de Precisión y Sensibilidad del detector en cascada con descriptores HOG

Numero De Etapas	Tasa De Falsos Positivos	Tasa De Verdaderos Positivos	Precisión	Sensibilidad
5	0.10	0.995	0.32	0.36
5	0.15	0.995	0.31	0.32
5	0.20	0.995	0.30	0.29
5	0.25	0.995	0.30	0.23
5	0.30	0.995	0.30	0.21
6	0.10	0.995	0.34	0.41
7	0.10	0.995	0.34	0.41
9	0.10	0.995	0.39	0.42
10	0.10	0.995	0.38	0.23
13	0.10	0.995	0.44	0.40
15	0.10	0.995	0.47	0.39
17	0.10	0.995	0.47	0.38
20	0.15	0.995	0.67	0.38
30	0.20	0.995	0.65	0.26
30	0.25	0.995	0.70	0.23
30	0.30	0.995	0.66	0.27

**Fuente:** Pruebas realizadas durante la implementación del sistema.

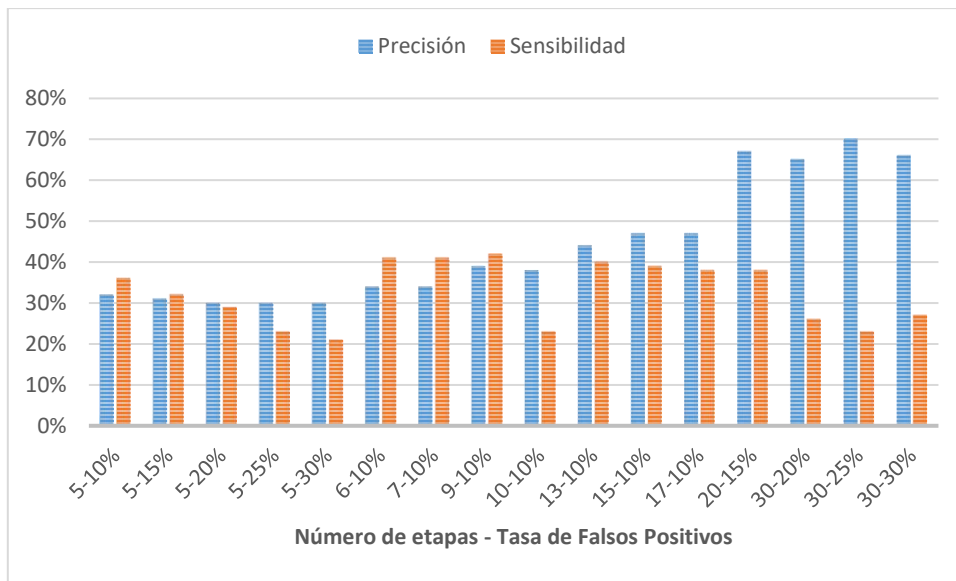
**Realizado por:** Urgilez M. 2019

Para una mejor visualización de los resultados se creó la figura 6-4 que describe la máxima precisión que se obtuvo de acuerdo al número de etapas utilizadas para el entrenamiento.



**Figura 6-4.** Grafica lineal de la precisión de acuerdo al número de etapas.

Realizado por: Urgilez M. 2019



**Figura 7-4.** Grafica de Precisión y Sensibilidad HOG.

Realizado por: Urgilez M. 2019

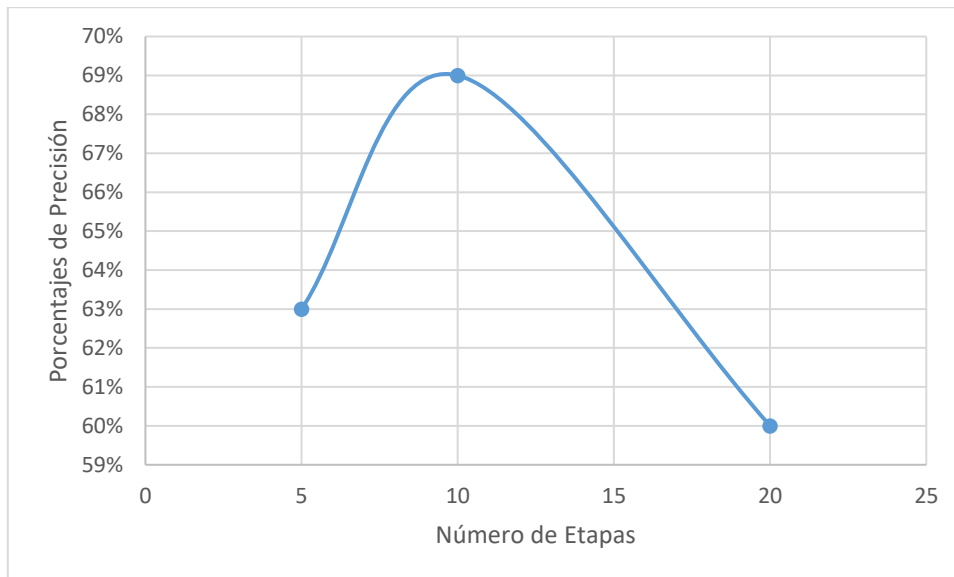
Al igual que el detector anterior, se realizó la Evaluación de detector en Cascada con descriptores LBP con la ayuda de una función denominada evaluando detector programada en el entorno de programación Matlab. En la tabla se puede distinguir que el mejor detector se da con un numero de etapas de 10 y una tasa de falsos positivos de 0.05 dando como resultado una precisión del 69%. y una sensibilidad del 45% como se muestra en la figura 9-4.

**Tabla 2-4:** Datos de Precisión y Sensibilidad del detector en cascada con descriptores LBP

Numero De Etapas	Tasa De Falsos Positivos	Tasa De Verdaderos Positivos	Precisión	Sensibilidad
5	0.05	0.995	0.63	0.47
5	0.10	0.995	0.55	0.48
5	0.30	0.995	0.32	0.46
10	0.05	0.995	0.69	0.45
10	0.10	0.995	0.66	0.45
10	0.15	0.995	0.67	0.44
10	0.20	0.995	0.67	0.45
10	0.25	0.995	0.65	0.46
10	0.30	0.995	0.60	0.46
20	0.20	0.995	0.60	0.46

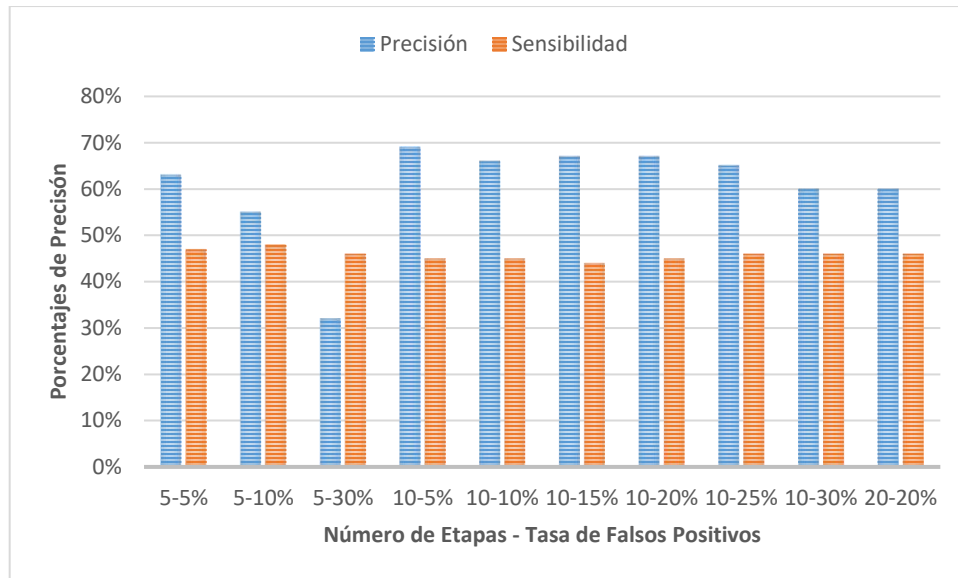
**Fuente:** Pruebas realizadas durante la implementación del sistema.

**Realizado por:** Urgilez M. 2019



**Figura 8-4.** Grafica lineal de la precisión de acuerdo al número de etapas

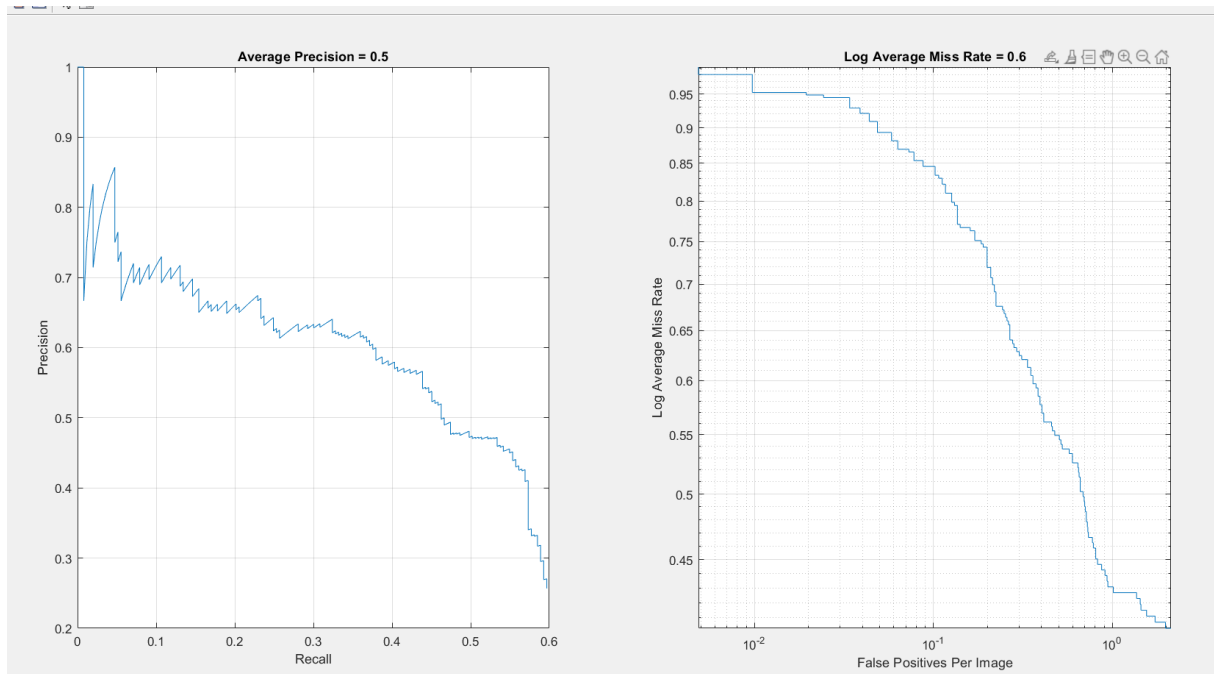
**Realizado por:** Urgilez M. 2019



**Figura 9-4.** Grafica de Precisión y Sensibilidad LBP.

**Realizado por:** Urgilez M. 2019

Finalmente se realizó la Evaluación de detector ACF, con la ayuda de una función denominada evaluación programada en el entorno de Matlab. Esta función devuelve el promedio de la precisión tomada para cada imagen y la tasa de error promedio logarítmica de los resultados de detección. En la imagen se puede observar que a medida que la sensibilidad aumenta la precisión disminuye, de igual manera si la tasa de error promedio logarítmica disminuye los Falsos positivos por imagen disminuyen. La precisión de este detector es del 50% y la sensibilidad (*recall*) es del 45%. El detector utiliza 10 estados o número de etapas de entrenamiento y un Factor de ejemplo negativos de 10.



**Figura 10-4.** Grafica lineal de la precisión y tasa de error promedio logarítmica

**Realizado por:** Urgilez M. 2019

En el proceso de reconocimiento de imágenes se pudo determinar que el detector de objetos en cascada HOG tiene una mayor precisión en la detección de botellas plásticas, y una menor sensibilidad con 70% y 23 % respectivamente. Además, se pudo comprobar que el procesamiento digital de imágenes permite el reconocimiento de botellas plásticas en el entorno, en ambientes no controlados, en este caso a orillas de lagunas con una precisión de 70 % en sus detecciones.

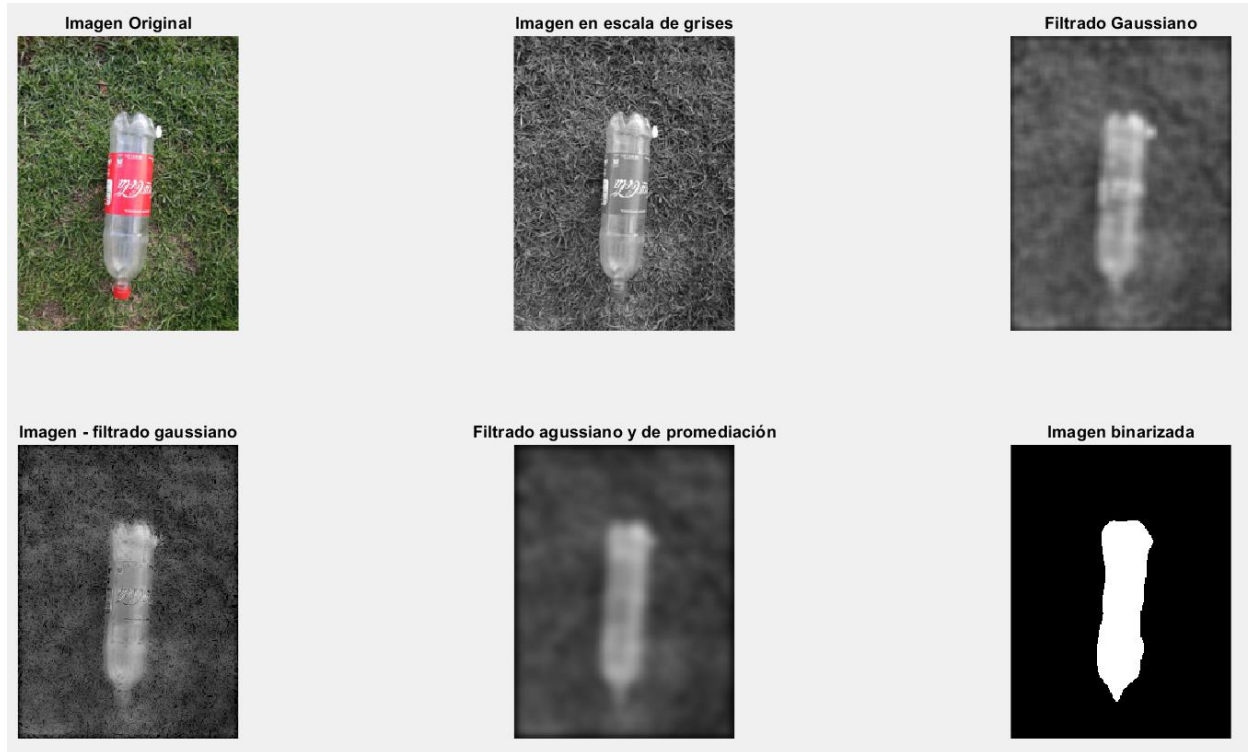
#### **4.3.1 Cálculo del área contaminada por botellas plásticas**

##### **4.3.1.1. Filtrado de Imagen**

Para el cálculo del área de una botella se crea una función que tiene como entrada la imagen del objeto a color y devuelve la estimación del área de los objetos en una imagen binaria, cuyo valor corresponde aproximadamente al número total de píxeles que ocupa el objeto en la imagen.

Se trabajó con un filtro de promediación de un tamaño de [9,9] y filtros gaussianos de paso bajo rotacionalmente de tamaño [9,9] con desviación estándar de 6 y 3 respectivamente. Se utilizaron sumas y restas de matrices y finalmente se binariza la imagen filtrada.

En la figura 11-4 se puede observar las diferentes etapas de filtrado de la imagen y finalmente el resultado de la aplicación de estos filtros en la imagen de botella.



**Figura 11-4.** Aplicación de filtros y binarización para el cálculo del área

**Realizado por:** Urgilez María, (2019)

#### 4.3.1.2. GSD (GROUND SAMPLING DISTANCE)

GSD es una medida de la cantidad de centímetros por cada píxel en una imagen, tiene como limitación el muestreo limitado por las características de la cámara empleada. Se puede definir también como una medida que permite relacionar la resolución de la imagen obtenida respecto a la superficie fotografiada, esta medición generalmente está en  $\text{píxel}/\text{cm}^2$  (Cristancho, 2020)

Para el cálculo del área es necesario conocer el GSD (*GROUND SAMPLING DISTANCE*), al usar *Pix4D Capture* para la adquisición de datos, este *software* nos entrega el GSD de las imágenes. Si no se utiliza el *DRONE* para la adquisición de las imágenes se puede calcular el GSD conociendo ciertos datos de la cámara, en este caso se usó *GROUND SAMPLING DISTANCE CALCULATOR* en Excel como se visualiza en la figura 8-4.





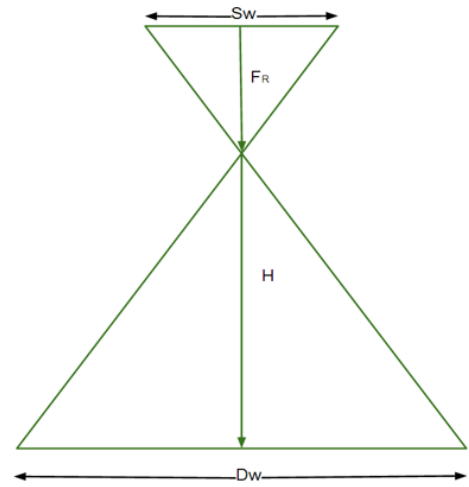
## GROUND SAMPLING DISTANCE CALCULATOR

### Instructions

1. Enter the Sensor Width (millimeters) in cell **C14**
2. Enter the Focal Length (millimeters) in cell **C15** (real focal length, not 35 mm equivalent)
3. Enter the Flight Height (meters) in cell **C16**
4. Enter the Image Width (pixels) in cell **C17**
5. [Optional] Enter the Image Height (pixels) in cell **C18**
6. Hit Enter

### Calculator

Sw	12,7	= the sensor width of the camera (millimeters)
Fr	1,79	= the focal length of the camera (millimeters)
H	10	= the flight height (meters)
imW	1080	= the image width (pixels)
imH	1080	= the image height (pixels)
GSD	6,57	= Ground Sampling Distance (centimeters/pixel)
Dw	71	= width of single image footprint on the ground (meters)
DH	71	= height of single image footprint on the ground (meters)



Dw = The footprint width / distance covered on the ground by one image in width direction

[www.pix4d.com](http://www.pix4d.com)

**Figura 12-4.** Cálculo de GSD

Fuente: (PIX 4D SA., 2019)

GSD se calcula con la siguiente ecuación:

$$GSD = \frac{Sw * H * 100}{Fr * imW} \quad (10)$$

El GSD calculado se usa para multiplicar por la suma de las áreas detectadas dándonos la estimación del área contaminada por botellas en  $\text{cm}^2$ .

## CONCLUSIONES

- Los envases de plástico PET son los de mayor uso en la industria alimenticia, a pesar de que se degradan aproximadamente a los 700 años, es por ello que en muchos países del mundo se ha prohibido el uso de plástico en todas sus formas, y así disminuir el impacto ambiental que ha generado a nivel mundial, incluso se ha desarrollado un mecanismo para convertir los residuos de PET en plástico virgen generando mayor reciclaje de este material.
- Tras el análisis de las zonas de las lagunas, debido especialmente a su flora y fauna se hace necesario trabajar con un sistema de recolección de imágenes capaz de sobrevolar la superficie de la laguna, un *drone*, evitando obstáculos como las totoras que sin un cuidado especial pueden llegar a crecer y formarse en gran cantidad en sus alrededores impidiendo de esta manera el uso de otro tipo de adquisición de imágenes. Para el manejo del *drone* se usó *PIX4D Capture*, que es una de las mejores aplicaciones gratuitas de planificación de vuelo de *drones* para mapeo y modelado 3D proporcionando telemetría en vivo e incluye además información como altitud de vuelo, GSD y velocidad de vuelo necesarias para la implementación del sistema.
- En el proceso de reconocimiento de imágenes se pudo determinar que el detector de objetos en cascada HOG tiene una mayor precisión en la detección de una botella, y una menor sensibilidad con valores de 70% y 23% respectivamente. HOG es mucho más robusto que LBP, sin embargo, el proceso conlleva una gran cantidad de recursos informáticos para la extracción de características de cada imagen.
- De acuerdo a las pruebas realizadas se pudo observar que mientras mayor es la altura de toma de imagen, el detector obtendrá un mayor número de Falsos Positivos, esto se debe principalmente a la poca información que se puede recolectar de las imágenes donde no se puede apreciar con precisión las características del objeto a detectar, de igual manera al tener una mayor tasa de sensibilidad en los detectores, la cantidad de Falsos positivos aumenta y disminuye la precisión en la detección, así como se detectan varias veces partes del objeto a identificar en este caso las botellas, este problema se da debido a que las intensidades representadas en los píxeles de la imagen de una botella tienen una gran varianza a la presencia de diferentes tipos de iluminación.

- Mientras mayor sea la cantidad imágenes utilizada para el entrenamiento del algoritmo de detección se tendrá un detector más robusto que posea una buena precisión en sus decisiones. Sin embargo, se debe tomar en cuenta que el uso de una base de datos excesiva conllevaría a un proceso de entrenamiento que puede llegar a generarse en semanas.

## RECOMENDACIONES

- Para el buen funcionamiento del sistema se debe trabajar con imágenes de alta resolución y captadas a una altura de 4m, tomadas en horas donde el sol no incida directamente sobre los objetos a fotografiar, de preferencia en las mañanas antes de las 10 am y en las tardes después de las 4 pm, para garantizar un mínimo porcentaje de error en sus detecciones.
- Dependiendo del tipo de objeto que se desee reconocer es importante obtener imágenes de entrenamiento similares a las que se van a utilizar en la realidad.
- Se recomienda entrenar al sistema con una base de datos de al menos 5000 imágenes positivas 10000 imágenes negativas. En el futuro para obtener una mejor precisión.
- La elección de los filtros y funciones utilizadas en el pre procesamiento debe ser determinada de acuerdo al tipo de objeto que se desea reconocer y las características presentes en el mismo, se recomienda hacer un estudio y análisis de diferentes tipos de filtros morfológicos de iluminación, contraste o ruido.
- Para futuros trabajos se puede utilizar algoritmos más complejos de *deep learning* o Redes Neuronales Convolucionada CNN para mejorar la precisión del detector.

## BIBLIOGRAFÍA

- Aldalu, B., & Santamaría, M. (2012). *Realce de Imágenes*. Argentina: Departamento de Ingeniería, Departamento de Matemática Universidad Nacional del Sur Bahía Blanca.
- Alegre, E., Fernandez, R., Sanchez, L., & Mostaza, J. (2003). *Procesamiento digital de image: fundamentos y prácticas con Matlab*. Universidad de León. Secretariado de Publicaciones y Medios Audiovisuales.
- Álvarez, J. (2012). *CICESE*. <http://usuario.cicese.mx/~josue/aplicaciones.php>
- Ambientum. (2019). *El portar profesional del medio ambiente*. <https://www.ambientum.com/ambientum/residuos/tiempo-basura.asp>
- Barriuso, J. (1991). Aplicaciones del proceso digital e imagenes en la arqueologia. *Complutum*, 319-323. Complutum.
- BBC, M. (11 de 12 de 2017). *Science Magazine*. <http://www.bbc.com/mundo/noticias-42304901>
- Bejerano, P. G. (21 de 08 de 2014). *Blogthinkbig.com*. <https://blogthinkbig.com/reconocimiento-de-objetos>
- Bin, Y., Junjie, Y., Zhen, L., & Stan, .. L. (29 de Septiembre de 2014). Aggregate channel features for multi-view face detection. *IEEE*.
- BM IMPEX ENGINEERING. (10 de enero de 2017). *usounicoprofesional*. Uso único: tipos de plásticos: <https://usounicoprofesional.com/uso-unico-tipos-de-plasticos/>
- Bohorquez, F., Vargas, Á., & Guzmán, I. (2017). REVISIÓN DEL PROCESAMIENTO DIGITAL DE IMÁGENES, SEGMENTACIÓN PARA LA DETECCIÓN DE ÁREAS CONTAMINADAS POR PETRÓLEO.
- Brunelli, R. (2009). *TEMPLATE MATCHING TECHNIQUES IN COMPUTER VISION*. Reino Unido: John Wiley & Sons Ltd,.
- Constituyente, A. N. (1998). *Constitución de la República del Ecuador*. Montecristi, Ecuador: Asamblea Nacional.
- Cristancho, C. (2020). Análisis de imágenes aéreas como apoyo para el seguimiento de cultivos de pasto para consumo bovino. Bogota. [https://ciencia.lasalle.edu.co/ing\\_automatizacion/783](https://ciencia.lasalle.edu.co/ing_automatizacion/783)
- Dollár, P., Tu, Z., Perona, P., & Belongie, S. (2009). *Integral Channel Features*. California: BMVC.
- El Comercio. (28 de 04 de 2017). Ecuador tiene un déficit en reciclar basura. Ecuador. <https://www.elcomercio.com/tendencias/sociedad/ecuador-deficit-reciclar-basura-contaminacion.html>
- El comercio. (7 de Enero de 2018). El aumento de basura, sin freno en la región. *El comercio*.

- García, E. (27 de noviembre de 2016). Detección y clasificación de objetos dentro de un salón de clases empleando técnicas de procesamiento digital de imágenes. México: Universidad Autónoma Metropolitana.
- García, G. (2013). *Procesamiento Audiovisual. Adquisición y representación de imágenes*. Murcia: Dept. de Informática y Sistemas, Universidad de Murcia. <http://dis.um.es/profesores/ginesgm>
- Gómez, D., & Guerrero, A. (mayo de 2016). Estudio y análisis de técnicas para procesamiento digital de imágenes. *Tesis de Ingeniería*. universidad tecnológica de Pereira.
- Gonzales, R., Woods, R., & Steven, E. (2004). *digital Image Processing Using Matlab*. New Jersey: PEARSON Prentice Hall.
- INEC. (Noviembre de 2020). *Instituto Nacional de estadísticas y Censos*. [https://www.ecuadorencifras.gob.ec/documentos/web-inec/Encuestas\\_Ambientales/Hogares/Hogares%202019/MOD\\_AMB\\_HOGAR\\_ESPND\\_2019\\_11.pdf](https://www.ecuadorencifras.gob.ec/documentos/web-inec/Encuestas_Ambientales/Hogares/Hogares%202019/MOD_AMB_HOGAR_ESPND_2019_11.pdf)
- Ioannis, A., & Bharath, A. A. (2011). *Using Machine Vision to Sort Plastics*. Londres: Department of Bioengineering, Imperial College London.
- Krasis Consulting S.L.U. (2015). *Campus MVP*. <https://www.campusmvp.es/recursos/post/Google-presenta-su-API-de-reconocimiento-de-imagenes-con-inteligencia-artificial.aspx>
- Kumar, D., & Prasad, G. V. (2012). *Image Processing Techniques to Recognize and Classify Bottle Articles*. International Journal of Computer Applications.
- Litterdrone. (2019). *Development and Exploitation of Innovative Tools for Remote Marine Litter Control and Management through UAVs*. <http://litterdrone.eu/index.php/proyecto/>
- López, A. (2016). "DETECCIÓN DE OBJETOS". Barcelo España.
- MAE. (2014). *Ministerio del Ambiente Ecuador*. <http://www.ambiente.gob.ec/maquinaria-para-recolectar-botellas-pet-contribuira-al-reciclaje-en-municipios/>
- Martínez, E., Jaramillo, M., Fernández, J., García, M., Marcos, A., Moral, F., & Lopez, F. (2001). *RECONOCIMIENTO DE MATERIALES PLÁSTICOS A TRAVÉS DE*. Dpto. de Expresión Gráfica. UNEX.
- Martorell, A. (2019). *Asociación Geoinnova*. <https://geoinnova.org/blog-territorio/mejor-pc-la-fotogrametria/>
- MathWorks. (2018). *MATLAB*. <https://la.mathworks.com/products/matlab.html>
- Mejía, J. (2005). *Apuntes de Procesamiento digital de Imágenes*. Mexico: Facultad de Ingeniería UASPL.
- Mejía, O. (19 de Mayo de 2011). Computación en la nube. 80, 45-52. ContactoS.
- Mercola. (08 de 04 de 2015). Mercola: <https://www.google.com.ec/amp/s/articulos.mercola.com/sitios/articulos/archivo/2015/04/08/amp/uso-de-plasticos.aspx>

- Monedero, A. (2017). *Adaptación automática de la detección de personas a la escena*. Madrid: Universidad Autónoma de Madrid.
- NOAA/Woods Hole Sea Grant. (2018). “*The known unknowns of plastic pollution*”. [http://c.files.bbc.co.uk/10BA3/production/\\_99151586\\_grafico\\_plastico\\_04\\_nc.png](http://c.files.bbc.co.uk/10BA3/production/_99151586_grafico_plastico_04_nc.png)
- Pandrone. (Febrero de 2018). 6 planificador de misiones gratuito para dron. Francia.
- PIX 4D SA. (15 de 10 de 2019). *PIX 4D*. <https://www.pix4d.com/>
- PlasticsEurope. (2018). *¿QUÉ ES EL PLÁSTICO?* PlasticsEurope: <https://www.plasticseurope.org/es/about-plastics/what-are-plastics>
- Quishpe, J., & Ortiz, J. (Septiembre de 2018). Ética en la producción de botellas plásticas y su contaminación al medio ambiente. *Revista Caribeña de Ciencias Sociales*. <https://www.eumed.net/rev/caribe/2018/09/botellas-plasticas-ambiente.html>
- Rossius, S. (2013). Reconocimiento de objetos mediante WebCam en tiempo real . *Tesis Doctoral*. . Universidad politecnica de Valencia.
- Sanz, J. A. (septiembre de 2013). Reconocimiento automático de formas. Barcelona: Universidad de Barcelona.
- Scavino, E., Wahab, D., Basri, H., & Mustafa, M. M. (2009). Application of automated image analysis to the identification and extraction of recyclable plastic bottles. *Journal of Zhejiang University SCIENCE A*.
- Semana S.A. (13 de 10 de 2016). *Semana Sostenible*. <http://sostenibilidad.semana.com/negocios-verdes/articulo/plastico-pet-un-amigable-pero-no-inofensivo/36282>
- Strong, B. A. (2005). *Plastics: materials and processing*. Devorah Yarnell.
- SZ DJI Technology Co. Ltd. (06 de 07 de 2017). *PHANTOM 3 PROFESSIONAL*. Phantom\_3\_Professional\_User\_Manual\_v1.8\_en: <https://www.dji.com/phantom-3-pro>
- Tecnología & Informática. (2019). *La cámara digital de fotos*. <https://tecnologia-informatica.com/camara-digital-fotos-lentes-cuidados/>
- The Economist. (3 de Marzo de 2018). <https://www.economist.com/news/international/21737498-so-far-it-seems-less-bad-other-kinds-pollution-about-which-less-fuss-made>
- The MathWorks, Inc. (2019). *MathWorks*. <https://la.mathworks.com/solutions/image-video-processing/object-recognition.html>
- Toledo, I. (abril de 2013). Reciclamiento del PET.
- Valveny, E. (2016). Detector basado en HAAR/Adaboost. *Detector de Objetos*. Barcelona.

# ANEXOS

## ANEXO A HOJAS DE ESPECIFICACIONES *DRONE PHANTOM 3*



Hojas Técnicas de los elementos que conforman el sistema de procesamiento digital de imágenes en la detección de material contaminante inorgánico en orillas de lagunas.

# PHANTOM 3

PROFESSIONAL


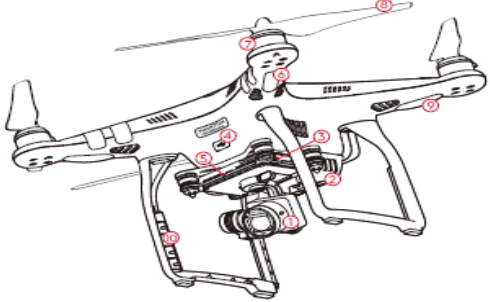
## Quick Start Guide

V1.2



### Phantom 3 Professional

It's time to meet your Phantom 3 Professional.  
The Phantom 3 Professional's camera records video at up to 4K and captures 12 megapixel photos.  
Review the diagram below for a full list of your Phantom 3 Professional's parts:



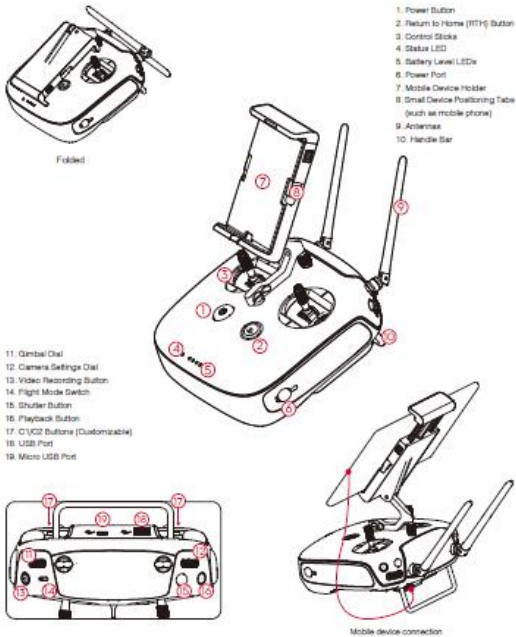
1. Gimbal and Camera	9. Aircraft Status Indicators
2. Vision Positioning System	10. Antennae
3. Camera Micro SD Card Slot	11. Intelligent Flight Battery
4. Aircraft Micro USB Port	12. Power Button
5. Camera Status Indicator	13. Battery Level Indicators
6. Front LEDs	14. Link Button
7. Motors	15. Camera Micro USB Port
8. Propellers	



## Remote Controller

This powerful Remote Controller allows you to pilot and maneuver your Phantom 3 Professional at distances over 16,400 feet (5 km)\* away, while putting selected camera controls at your fingertips.

Built into your Remote Controller is a rechargeable LiPo battery and DJI Lightbridge, which when paired with a compatible mobile device gives you a live HD view from the Phantom's camera.



1. Power Button
2. Return to Home (RTH) Button
3. Control Sticks
4. Status LED
5. Battery Level LED
6. Power Port
7. Mobile Device Holder
8. Small Device Positioning Tab (such as mobile phone)
9. Antenna
10. Handle Bar

11. Gimbal Dial
12. Camera Settings Dial
13. Video Recording Button
14. Flight Mode Switch
15. Shutter Button
16. Playback Button
17. C/VC2 Buttons (Customizable)
18. USB Port
19. Micro USB Port

\* This maximum transmission distance was tested in a lab environment and is for reference only. The maximum operating distance may vary depending on conditions in your immediate surroundings.

## Fly Safe

DJI encourages you to enjoy flying your Phantom 3 Professional in a safe, responsible, and smart way. To do this, it is important to understand some basic flight guidelines, both for your protection and for the safety of those around you.

1. **Fly in Open Areas:** Always fly in locations that are free and clear of buildings, trees, power lines, and other obstacles. Do not fly above or near people or animals.
2. **Maintain Control at All Times:** Even when using DJI autopilot functions such as Auto-Takeoff, Auto-Landing, and Auto-Return to Home, always keep your hands on the Remote Controller and maintain control of your aircraft when it is in flight.
3. **Maintain Line of Sight:** Keep your aircraft in sight at all times, and avoid flying behind buildings or other obstacles that may block your view.
4. **Monitor Your Altitude:** For the safety of all-sized aircraft and other air traffic, always fly at altitudes less than 400 feet (120 meters) above ground level, or in line with your local laws and regulations.

Visit <http://flysafe.dji.com/ho-fly> for more information on critical safety features such as No-Fly Zones.



### Calibrating the Compass:

Only calibrate the compass when the DJI GO app or the status indicator prompt you to do so. Observe the following rules when calibrating your compass:

1. DO NOT calibrate your compass where there is a chance of strong magnetic interference, such as magnets, parking structures, and steel environments underground.
2. DO NOT carry ferromagnetic materials with you during calibration such as cellular phones.
3. The DJI GO app will prompt you to resolve the compass issue if the compass is affected by strong interference after calibration is complete. Follow the prompted instructions to resolve the compass issue.

### Environmental Considerations:

1. Do not fly in severe weather conditions. This includes high winds (speeds of 20 mph or more), snow, rain, and fog.
2. Only fly in open areas. Tall buildings and steel structures may affect the accuracy of the on-board compass and GPS signal.
3. Avoid obstacles, crowds, high-voltage power lines, trees, and bodies of water.
4. Minimize electromagnetic interference by avoiding areas with high levels of electromagnetic, including mobile phone base stations, radio transmission towers, or Wi-Fi hotspots.
5. Aircraft and battery performance are subject to environmental factors such as air density and temperature. Be very careful when flying 10,000 feet (3 km) or more above sea level, as battery and aircraft performance may not be at peak efficiency.
6. The Phantom 3 Professional cannot operate in P Mode or use GPS at polar latitudes. Only ATTI Mode and the Vision Positioning system will be operational.

### P Mode:

In this mode, the Phantom 3 Professional can use GPS and the Vision Positioning system, allowing it to hover accurately in position indoors and out. When GPS is available, a Home Point will be locked so that the Phantom 3 Professional can Return to Home if the Remote Controller signal is lost.



To enable P Mode, toggle the Flight Mode Switch to the P position.

### There are three states in P mode:

**P-GPS:** GPS works best when outdoors and in a wide open area. Your Phantom 3 Professional uses GPS to hover in place when there is a strong GPS signal.

**P-VISIO:** If GPS is not available, the aircraft can use the Vision Positioning system to hover accurately.

**P-ATTI:** Neither GPS nor the Vision Positioning system are available. The aircraft will only use its barometer for altitude and other on-board sensors for attitude stabilization.

Note that the Vision Positioning system may not work properly when the Phantom 3 Professional is flying over water, over surfaces without a clear pattern, or in a low-light environment.



### Return to Home:

When there is a strong GPS signal, the aircraft will be able to record a Home Point and return to that Home Point when required. The Home Point location is recorded when the GPS signal icon in the DJI GO app is either yellow or green.

The aircraft will return to the Home Point automatically in the following scenarios (all require a strong GPS signal):

**Smart RTH:** When the RTH button on the Remote Controller or in the DJI GO app is pressed.

**Low-Battery RTH:** A notification will appear in the DJI GO app requesting the pilot to take action when the battery level falls under a certain level.

**Failsafe RTH:** When the Remote Controller's signal is lost.



While returning to the Home Point, the aircraft's altitude can be adjusted to avoid obstacles. Tall buildings may affect the Remote Controller's signal. The Failsafe Return to Home procedure will be triggered if the signal is lost. Be sure to fly higher than any nearby buildings to avoid crashing.

## ANEXO B TABLA COMPARATIVA SOFTWARE DE FOTOGRAMETRÍA

Tabla comparativa de aplicaciones gratuitas destinadas a planificar misiones para los *drones* DJI y Parrot. (*Pix4D Capture*, DJI GS PRO, *Drone Deploy*, *Drone Harmony*, Hangar360, Droneblock).

designed by 	Free app ? 	iPhone 	iPad 	Android 	Plan mapping 	Circular 	Peripheral mapping 	Waypoint route 	360° pan 	Ergonomy 	Mission saving 	Capture preview 	DJI 	Parrot 
 Pix4D Capture	✓	✓	✓	✓	✓	✓	✗	✗	✗	☆☆☆☆	✗	✗	✓	✓
 DJI GS PRO	✓	✗	✓	✗	✓	✓\$	✗	✓	✗	☆☆☆	✓	✗	✓	✗
 DroneDeploy	✓	✓	✓	✓	✓	✗	✗	✗	✗	☆☆☆	✓	✗	✓	✗
 Drone Harmony	✓	✗	✓	✓	✓	✓	✓	✓	✗	☆☆☆	✓	✓	✓	✗
 Hangar360	✓	✓	✗	✗	✗	✗	✗	✗	✓	☆☆☆	✗	✗	✓	✗
 Drone Blocks	✓	✓	✗	✗	✗	✗	✗	✓	✓	☆☆	✓	✗	✓	✗

Fuente: (Pandrone, 2018)

## ANEXO C CÓDIGO DE PROGRAMACIÓN MATLAB

Código del programa principal de procesamiento digital de imágenes.

### CÓDIGO PREPROCESAMIENTO.

```
function [Ic]=preprocesamiento (Image)% % Cargar la imagen
Icolor=imresize(Image,[256 256]);
% elimine ruido
[IcolorR,IcolorG,IcolorB] = imsplit(Icolor);
net = denoisingNetwork('dncnn');
IcolorR = denoiseImage(IcolorR,net);
IcolorG = denoiseImage(IcolorG,net);
IcolorB = denoiseImage(IcolorB,net);
IRGB = cat(3,IcolorR,IcolorG,IcolorB);
% ajustar el contraste
Ic= imadjust(IRGB,stretchlim(IRGB));
```

### CÓDIGO MATLAB DETECTOR OBJECT CASCADE HOG

```
% Entrenamiento
%% Load positive images and bounding boxes of bikes
load('tabletrain1080.mat');
%Select the bounding boxes for stop signs from the table.
positiveInstances = tabletrain1080(:,1:2);
%% Step 2: Specify folder with negative images
negativeFolder = [pwd '\NO BOTELLA'];
%% Step 3: Train the detector
NumStages = 30;
FAR = 0.25;
%trainCascadeObjectDetector('objectdetector30_10.xml', tablebotella, negativeFolder,...
% 'NumCascadeStages', NumStages, 'FalseAlarmRate', FAR);
%trainCascadeObjectDetector('objectdetector13_45.xml',positiveInstances, ...
% negativeFolder,'FalseAlarmRate',FAR,'NumCascadeStages',NumStages);
trainCascadeObjectDetector('objectdetector30_25.xml', tabletrain1080, negativeFolder,...
'ObjectTrainingSize', 'Auto', 'NegativeSamplesFactor', 2,...
```

```
'NumCascadeStages', NumStages, 'FalseAlarmRate', FAR,...  
'TruePositiveRate', 0.995, 'FeatureType', 'HOG');
```

### **CODIGO DE MATLAB DEL PROGRAMA CARPETA.**

```
%% dividir aleatoriamente los datos trainig y test  
valFolderName='D:\DETECTOR EN CASCADA 1080\BOTELLA';  
files = dir(fullfile(valFolderName));  
N = length(files);  
tf=randperm(N)>(0.80*N);  
if not(isfolder(fullfile('train')))  
    mkdir train  
end  
if not(isfolder(fullfile('test')))  
    mkdir test  
end  
%%  
for i=3:length(tf)  
    files_re = files(i).name;  
    if tf(i) == 1  
        copyfile (fullfile(valFolderName,files_re),'test')  
    else  
        copyfile (fullfile(valFolderName,files_re),'train')  
    end  
end  
end
```

### **CÓDIGO DE LA FUNCIÓN ÁREA**

```
% Calculo del area  
function [a]=area(Ic)  
I=rgb2gray(Ic);  
filtro=fspecial('gaussian', [9 9], 6);  
If=imfilter(I,filtro);  
background=(I-If);
```

```

I= I-background;
I= I-background;
filtro=fspecial('gaussian', [9 9], 3);
If=imfilter(I,filtro);
filtro=fspecial('average', [9 9]);
If=imfilter(If,filtro);
BW1=imbinarize(If);
a=bwarea(BW1);

```

## CÓDIGO EVALUANDO PRECISIÓN OBJECT CASCADE

```

clc
clear all
close all
load('gTruthtest1080.mat')
testGTruth = selectLabels(gTruthtest1080,'botella');
%%
evaluationData = objectDetectorTrainingData(gTruthtest1080,...
    'SamplingFactor',1,'WriteLocation','D:\DETECTOR EN CASCADA 1080\test');
%%
imds = imageDatastore(fullfile('D:\DETECTOR EN CASCADA 1080\test'));
numImages = height(evaluationData);
result(numImages,:) = struct('imageFilename',[],'botella',[]);
%%
addpath('D:\DETECTOR EN CASCADA 1080\XML')
detector = vision.CascadeObjectDetector('objectdetector20_20.xml');
%%
for i = 1:numImages
    % Read Image
    I = readimage(imds,i);
    % Detect the object of interest
    bboxes = detector(I);
    % Store result
    result(i).botella = bboxes;

```

```

    result(i).imageFilename=imds.Files(i);
end
% Convert structure to table
results = struct2table(result);
%% PRECISION
%ordenar las tablas
evaluationData = sortrows(evaluationData,'imageFilename','ascend');
results = sortrows(results,'imageFilename','ascend');
%% Evaluate Metrics
[tasa, precision, precision_traslape, tasasensibilidad,
sensibilidad]=evaluandoDetector(evaluationData,results);

```

## **CÓDIGO MATLAB DETECTOR ACF**

```

%% detector ACF
clc
clear all;
close all;
load('gTruthtrain1080.mat');
botellaGTruth = selectLabels(gTruthtrain1080,'botella');
%%
train = objectDetectorTrainingData(botellaGTruth,'SamplingFactor',2,...
'WriteLocation','D:\detector ACF 2\BOTELLA');
%train = objectDetectorTrainingData(botellaGTruth);
%% entrenar al detector
detector = trainACFObjectDetector(train,'NumStages',10,'NegativeSamplesFactor',10 );
%% Guardar detector
save('DetectorACF20-10.mat','detector');

```

## CODIGO EVALUACION ACF

```
clc
clear all
close all
load('gTruthtest1080.mat')
testGTruth = selectLabels(gTruthtest1080,'botella');
load('DetectorACF20-10.mat');
%%
evaluationData = objectDetectorTrainingData(gTruthtest1080,...
    'SamplingFactor',1,'WriteLocation','test');
%%
imds = imageDatastore(fullfile('test'));
numImages = height(evaluationData);
result(numImages,:) = struct('Boxes',[],'Scores',[]);
%%
for i = 1:numImages
% Read Image
    I = readimage(imds,i);
% Detect the object of interest
    [bboxes,scores] = detect(detector,I,'SelectStrongest',true);
    [selectedBbox,selectedScore] = selectStrongestBbox(bboxes,scores,'OverlapThreshold',0.1);
    if length(selectedScore)>3
        [~,idx] = max(selectedScore);
% idx=find(selectedScore>13);
        selectedBbox1(1,:)=selectedBbox(idx,:);
        selectedScore1(1,:)=selectedScore(idx,:);
        selectedScore(idx,:) = [];
        [~,idx] = max(selectedScore);
        selectedBbox1(2,:)=selectedBbox(idx,:);
        selectedScore1(2,:)=selectedScore(idx,:);
        selectedScore(idx,:) = [];
        [~,idx] = max(selectedScore);
        selectedBbox1(3,:)=selectedBbox(idx,:);
        selectedScore1(3,:)=selectedScore(idx,:);
```

```

        selectedScore(idx,:) = [];
    else
        selectedBbox1=selectedBbox;
        selectedScore1=selectedScore;
    end
    % Store result
    result(i).Boxes = selectedBbox1;
    result(i).Scores = selectedScore1;
end
% Convert structure to table
results = struct2table(result);
%%
evaluationData = sortrows(evaluationData,'imageFilename','ascend');
% results = sortrows(results,'imageFilename','ascend');
overlap = 0.1;
% Evaluate Metrics
[ap,recall,precision] = evaluateDetectionPrecision(results...
    ,evaluationData(:,2),overlap);
[am,fppi,missRate] = evaluateDetectionMissRate(results,evaluationData(:,2),overlap);
%% Plot Metrics
subplot(1,2,1);
plot(recall,precision);
xlabel('Recall');
ylabel('Precision');
title(sprintf('Average Precision = %.1f', mean( precision)))
grid on
subplot(1,2,2);
loglog(fppi, missRate);
xlabel('False Positives Per Image');
ylabel('Log Average Miss Rate');
title(sprintf('Log Average Miss Rate = %.1f', mean(missRate)))
grid on
%%
rmpath('EvaluationData');

```



## CÓDIGO EVALUANDO DETECTOR

```
function [F1,P1, P, R, S] = evaluandoDetector(G,D)
sumMejorTraslapeD = 0;%%
Ndet = 0;%predicciones totales
sumMejorTraslapeG = 0;
Precision=0;
Sensibilidad=0;
sumPrecision=0;
Nobj = 0;% Numero de objetos
RPos=0;% Reales Positivos
FPos=0;% Falso Positivos
FNeg=0; % Falsos Negativos
j=1;
k=1;
for im=1:height(D)
    % Calculando matriz de traslape
    IndiceTraslape = bboxOverlapRatio(G.botella{im,1},D.botella{im,1});
    ndet = size(D.botella{im,1},1);
    Ndet = Ndet + ndet;
    nobj = size(G.botella{im,1},1);
    Nobj = Nobj + nobj;
    if ndet>0 && nobj>0
        rpos=numel(IndiceTraslape(IndiceTraslape>0));
        RPos=RPos+rpos;
        fpos=ndet-rpos;
        FPos=FPos+fpos;
        % Calculando Presicion1
        Precision=rpos/(rpos+fpos);
        j=j+1;
        sumPrecision=sumPrecision+Precision;
        % Calculando Precisión traslape
        sumMejorTraslapeD = sumMejorTraslapeD + sum(max(IndiceTraslape,[],1));
        % Calculando Sensibilidad
        if nobj>rpos
```

```
fneg=nobj-rpos;
FNeg=FNeg+fneg;
s=rpos/(rpos+fneg);
Sensibilidad=Sensibilidad+s;
k=k+1;
end
% Calculando Sensibilidad traslape
sumMejorTraslapeG = sumMejorTraslapeG + sum(max(IndiceTraslape,[],2));
end
end
P1 = sumPrecision/j;
P = sumMejorTraslapeD/Ndet;
R = sumMejorTraslapeG/Nobj;
S = Sensibilidad/k;
F1 = 2*P*R/(P+R);
```