



**UAEM** | Universidad Autónoma  
del Estado de México

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE  
MÉXICO

INGENIERÍA EN COMPUTACIÓN

---

**Desarrollo de algoritmos evolutivos para la segmentación  
automática de imágenes**

---

Tesis que presenta

**Jared CERVANTES CANALES**

Para obtener el Grado de

**Maestro en Ciencias**

Asesor de Tesis:

**Dr. Farid GARCÍA-LAMONT**

Texcoco, Estado de México.

Octubre del 2021



# Abstract

A vision system is made up of several steps, each of which influences the final result. One of those steps is segmentation, it allows to obtain only the region of interest from the object and eliminate the background. Segmentation is very important because it allows to improve the quality of the segmented region, improve the quality of the extracted features and at the same time reduce the noise introduced by a poor quality of features. There are several segmentation techniques in the literature, each technique with one or more variables to tune. Choosing the appropriate technique to segment a certain image set is not an easy task, some times various factors can influence the segmentation's quality such as pre-processing and data set. In addition, small variations in lighting and changes in intensity can have influence on the quality of the segmentation. Sometimes an expert is necessary to select the segmentation technique and fine-tune the associated variables. In this Thesis, automatic algorithms are developed for the selection of segmentation techniques and associated variables. The techniques developed are implemented and compared using several sets of images. The experimental results obtained are discussed and compared in this Thesis. These algorithms will facilitate and simplify the selection of the segmentation techniques to be used, as well as the necessary variables and implemented pre-processing.



# Resumen

Un sistema de visión está compuesto por varios pasos, cada uno de ellos influye en el resultado final. Uno de esos pasos es la segmentación, esta permite obtener solo la región que interesa del objeto y eliminar el fondo. La segmentación es muy importante pues permite mejorar la calidad de la región segmentada, mejorar la calidad de las características extraídas y al mismo tiempo reducir el ruido introducido por una mala calidad de características. Existen diversas técnicas de segmentación en la literatura, cada una de ellas con una o varias variables para afinar. Elegir la técnica adecuada para segmentar un determinado conjunto de datos no es una tarea fácil, ya que en algunas ocasiones diversos factores pueden influir en la calidad de la segmentación como preprocesamiento y conjunto de imágenes. Aunado a lo anterior, pequeñas variaciones en la iluminación y cambios de intensidad pueden influir demasiado en la calidad de la segmentación. En algunas ocasiones es necesario un experto para seleccionar la técnica de segmentación y poner a punto las variables asociadas. En esta Tesis, se desarrollan algoritmos automáticos para selección de técnicas de segmentación y variables asociadas. Las técnicas desarrolladas son implementadas y comparadas utilizando varios conjuntos de datos. Los resultados experimentales obtenidos son discutidos y comparados en esta Tesis. Estos algoritmos permitirán facilitar y simplificar la selección de la técnica de segmentación a utilizar, así como las variables necesarias y preprocesamiento implementado.



# Índice general

<b>Abstract</b>	<b>III</b>
<b>Resumen</b>	<b>V</b>
<b>Índice general</b>	<b>VII</b>
<b>Índice de figuras</b>	<b>XI</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Problemática . . . . .	3
1.2. Justificación . . . . .	4
1.3. Objetivos . . . . .	4
1.4. Estado del Arte . . . . .	5
1.4.1. Técnicas basadas en umbral . . . . .	5
1.4.2. Técnicas basadas en <i>Clustering</i> . . . . .	6
1.4.3. Técnicas basadas en detección de bordes . . . . .	7
1.4.4. Técnicas basadas en regiones . . . . .	8
1.4.5. Técnicas basadas en Redes neuronales . . . . .	9
1.5. Organización de la Tesis . . . . .	11
<b>2. Preliminares</b>	<b>13</b>
2.1. Técnicas de segmentación . . . . .	13
2.1.1. Técnicas de segmentación basadas en umbral . . . . .	16
<b>Segmentación por frontera global</b> . . . . .	16
<b>Frontera adaptativa</b> . . . . .	17
<b>Método de Otsu</b> . . . . .	18
2.1.2. Métodos basados en bordes . . . . .	20
<b>Detección de bordes</b> . . . . .	20
<b>Método del gradiente</b> . . . . .	21

	<b>Laplaciano</b> . . . . .	21
2.1.3.	Métodos Basados en <i>Clustering</i> . . . . .	22
	<b>K-means</b> . . . . .	23
2.1.4.	Métodos Basados en regiones . . . . .	23
	<b>Dividir y Fusionar</b> . . . . .	24
2.1.5.	Métodos Basados en IA . . . . .	24
	<b>Redes neuronales</b> . . . . .	24
	<b>Máquinas de soporte vectorial</b> . . . . .	24
2.2.	Espacios de color . . . . .	25
2.2.1.	Espacio RGB . . . . .	25
2.2.2.	Espacio HSV . . . . .	26
2.2.3.	Espacio HSI . . . . .	27
2.3.	Algoritmos Genéticos . . . . .	28
	<b>Elementos de un algoritmo genético.</b> . . . . .	28
2.3.1.	Algoritmo genético básico . . . . .	29
2.3.2.	Población inicial . . . . .	30
2.3.3.	Selección de individuos . . . . .	30
2.3.4.	Cruza . . . . .	31
2.3.5.	Mutación . . . . .	31
2.3.6.	Condición de paro . . . . .	32
<b>3.</b>	<b>Metodología</b> . . . . .	<b>33</b>
3.1.	Conjunto de datos . . . . .	36
3.2.	Selección de imágenes . . . . .	38
3.2.1.	Sefexa Image Segmentation Tool . . . . .	38
3.3.	Selección de técnicas y parámetros . . . . .	40
3.3.1.	Población inicial y representación del cromosoma . . . . .	41
	Caso 1 (9 variables) . . . . .	42
	Caso 2 (11 variables) . . . . .	43
3.3.2.	Evaluación de función de aptitud . . . . .	45
	Probabilistic Rand Index (PRI) . . . . .	45
	Global Consistency Error (GCE) . . . . .	46
	Variation of Information (VOI) . . . . .	47
3.3.3.	Selección . . . . .	47

3.3.4.	Cruza . . . . .	48
3.3.5.	Mutación . . . . .	49
3.3.6.	Condición de paro . . . . .	49
3.3.7.	Elitismo . . . . .	50
<b>4.</b>	<b>Resultados experimentales</b>	<b>51</b>
4.1.	Conjuntos de datos . . . . .	51
4.1.1.	Conjunto de datos ICL . . . . .	51
4.1.2.	Conjunto de datos enfermedades del café . . . . .	52
4.1.3.	Conjunto de datos PlantVillage . . . . .	53
4.1.4.	Conjunto de datos LSM . . . . .	54
4.2.	Segmentación de los conjuntos de datos . . . . .	55
4.2.1.	Conjunto de datos ICL . . . . .	55
4.2.2.	Conjunto de datos enfermedades del café . . . . .	56
4.2.3.	Conjunto de datos PlantVillage . . . . .	57
4.2.4.	Conjunto de datos LSM . . . . .	58
4.3.	Resultados y análisis comparativo . . . . .	59
4.3.1.	Convergencia del algoritmo genético . . . . .	59
4.4.	Discusiones . . . . .	60
4.4.1.	Resultados obtenidos con el conjunto de datos Enfermedades del café . . . . .	61
4.4.2.	Resultados obtenidos con el conjunto de datos PlantVillage . . . . .	64
4.4.3.	Resultados obtenidos con el conjunto de datos LSM . . . . .	67
<b>5.</b>	<b>Conclusiones</b>	<b>71</b>
	<b>Bibliografía</b>	<b>73</b>



# Índice de figuras

1.1. Artículos clasificados del Estado del Arte . . . . .	10
2.1. Segmentación semántica <sup>1</sup> . . . . .	14
2.2. Segmentación de instancias <sup>2</sup> . . . . .	14
2.3. Métodos de segmentación más elementales . . . . .	15
2.4. Filtro utilizado para el Laplaciano . . . . .	22
2.5. Modelo RGB . . . . .	26
2.6. Modelo HSV <sup>3</sup> . . . . .	27
2.7. Representación gráfica de las expresiones que se usa en la genética . . . . .	29
2.8. Operador de Cruza . . . . .	31
2.9. Operador de Mutación . . . . .	32
3.1. Componentes de un sistema de visión . . . . .	33
3.2. Técnicas y parámetros necesarios para segmentar . . . . .	34
3.3. Metodología . . . . .	36
3.4. Conjuntos de datos utilizados en los experimentos . . . . .	37
3.5. Selección del conjunto de imágenes y obtención de parámetros de segmentación . . . . .	39
3.6. Interfaz del software Sefexa . . . . .	39
3.7. Segmentación por Sefexa . . . . .	40
3.8. Diagrama de flujo de un algoritmo genético . . . . .	41
3.9. Cadena cromosómica de nueve variables . . . . .	42
3.10. Cadena cromosómica de once variables . . . . .	44
3.11. Representación del porcentaje de aptitud . . . . .	48
3.12. Cruza de dos puntos . . . . .	49
4.1. Conjunto de datos ICL . . . . .	52
4.2. Conjunto de datos enfermedades de café . . . . .	53
4.3. Conjunto de datos PlantVillage . . . . .	54
4.4. Conjunto de datos LSM . . . . .	55

4.5. Segmentación del conjunto de datos ICL . . . . .	56
4.6. Segmentación del conjunto de datos enfermedades del Café . . . . .	57
4.7. Segmentación del conjunto de datos PlantVillage . . . . .	58
4.8. Segmentación del conjunto de datos LSM . . . . .	59
4.9. Distancia promedio entre individuos . . . . .	60
4.10. Mejor individuo obtenido y aptitud promedio de individuos . . . . .	60
4.11. The Probabilistic Rand Index (PRI) para cada imagen en el conjunto de imágenes de cafe . . . . .	61
4.12. El Global Consistency Error (GCE) para cada imagen en el conjunto de imágenes de cafe . . . . .	62
4.13. La Variance of Information (VoI) para cada imagen en el conjunto de imágenes de cafe . . . . .	63
4.14. The Probabilistic Rand Index (PRI) for each image in PlantVillage dataset	65
4.15. The Global Consistency Error (GCE) for each image in PlantVillage dataset	66
4.16. The Variance of Information (VoI) for each image in PlantVillage dataset	67
4.17. The Probabilistic Rand Index (PRI) for each image in LSM dataset . . . .	68
4.18. The Global Consistency Error (GCE) for each image in LSM dataset . . .	69
4.19. The Variance of Information (VoI) for each image in LSM dataset . . . .	70

# Capítulo 1

## Introducción

El desarrollo de sistemas automáticos de visión y procesamiento de imágenes ha sido de gran interés para resolver problemas de la vida real. Muchas áreas como medicina [26] [15] [68], ingeniería civil [61], Astronomía [57], seguridad[62], robótica [38] y agricultura [9] se apoyan y hacen uso de las técnicas de procesamiento de imágenes y visión para procesar y analizar la información que pueda ser utilizada para tomar decisiones.

En la actualidad, aunque los sistemas de procesamiento de imágenes digitales y los sistemas de visión han avanzado mucho, presentan varios problemas cuando se implementan. Los sistemas de visión dependen de varios pasos, y cada paso se realiza mediante alguna técnica que depende de sus variables. En general, un sistema de visión consiste en preprocesamiento, segmentación, extracción de características, selección de características y clasificación, entre otros. Uno de los pasos más importantes en un sistema de visión es la segmentación.

La segmentación es un paso muy importante y sensible en un sistema de visión. Pequeños cambios en el preprocesamiento, variables de cada técnica, iluminación o intensidad de las imágenes, o incluso operaciones morfológicas como dilatación o erosión para pulir la imagen final pueden provocar una mala segmentación. El principal objetivo de este trabajo es el desarrollo de algoritmos automáticos para la segmentación de conjuntos de imágenes.

La segmentación se define como el proceso de dividir la imagen en múltiples regiones. La segmentación más básica es la segmentación binaria, que divide la imagen en solo dos regiones, la región de interés y el fondo de la imagen. Este proceso facilita el análisis de la imagen, seccionando y obteniendo su parte más representativa y eliminando la que no representa interés.

Formalmente, la segmentación de imágenes de genes se define como [22]:

$$I_{segmentada} = \cup_{i=1}^n Ri \quad (1.1)$$

Tal que:

$$\cap_{i=1}^n Ri \neq \emptyset \quad (1.2)$$

Donde  $n$  representa el número de segmentos

$$R_k = \{(i, j) \in N^2 \mid I(i, j) = \delta_k\} \quad (1.3)$$

Siendo  $I(i, j)$  el valor del píxel local en  $(i, j)$  en la imagen de entrada  $I$  y  $\delta_k$  es el valor del umbral en el  $k$  lugar.

En otras palabras, la segmentación denota un proceso mediante el que la imagen de entrada se divide en regiones no superpuestas. Cada región es homogénea y conectada. Cada región en una imagen segmentada debe satisfacer las propiedades de homogeneidad y conectividad. Se considera que la región es homogénea si todos los píxeles de la región satisfacen las condiciones de homogeneidad definidas por uno o más atributos de píxel, como intensidad, color, textura, etc.

Existen diversas técnicas de segmentación en la literatura, cada una de ellas con una o varias variables para afinar. Elegir la técnica adecuada para segmentar un determinado conjunto de datos no es una tarea fácil. Múltiples factores pueden influir en la calidad de la segmentación como técnicas de preprocesamiento, variaciones en la iluminación, cambios de intensidad y demás.

En algunas ocasiones es necesario un experto que evalúe todos los aspectos del conjunto de imágenes, realice experimentos para seleccionar la técnica de segmentación adecuada y ponga a punto las variables asociadas. Sin embargo, el cálculo de tales variables para una correcta segmentación de la imagen tiene una complejidad combinatoria muy alta, además, no existe una solución general para todos los casos de segmentación. Debido a que el espacio de solución es muy grande es necesario y justificado el uso de algoritmos evolutivos.

En los últimos años, los algoritmos evolutivos han llegado a ser una herramienta importante para resolver problemas de optimización. La optimización se ve como un problema de decisión que implica encontrar la mejor solución dentro de un conjunto finito de posibles soluciones [30] [33].

Las técnicas y conceptos de optimización no se limitan a ninguna disciplina en particular y desempeñan un papel cada vez más importante en la solución y el modelado en

diversos campos de la ciencia que van desde la bioinformática, la economía, el diseño de estructuras, hasta el modelado y automatización.

Es por ello, que el desarrollo de sistemas de optimización que mejoren los resultados es un reto importante en procesamiento digital de imágenes. Con particular relevancia, también han sido utilizados en problemas de procesamiento de imágenes, incluyendo: detección de bordes, reconocimiento facial, segmentación de imagen, compresión de imágenes y extracción de características en imágenes.

En esta Tesis, se desarrollan algoritmos automáticos para selección de técnicas de segmentación y variables asociadas. Las técnicas desarrolladas son implementadas y comparadas utilizando varios conjuntos de datos. En este Capítulo introductorio se explica la justificación de este trabajo, se traza el planteamiento del problema, la descripción del objetivo general y los específicos, un breve resumen de las aportaciones anteriores, y finalmente se describe la organización del resto del documento.

## 1.1. Problemática

Un mundo en donde dispositivos tienen la habilidad de detectar objetos, anomalías en un conjunto de imágenes, personas o animales, gestos o señas en imágenes es una realidad que se vive en el presente. Detectar objetos ha sido una tarea que incluye varios procesos y que se ha implementado en diversos campos de aplicación. El proceso básico que conlleva esta labor está centrado en la segmentación de imágenes para una posterior extracción de características.

Sin embargo, identificar los procesos adecuados para una óptima segmentación no es un proceso fácil. Esto requiere en algunos casos un profundo conocimiento de cada proceso involucrado. Una correcta obtención de variables involucradas y técnicas necesarias para una correcta segmentación solo es obtenida a través de años de experiencia y aunado a ello, múltiples experimentos previos para llegar a una solución adecuada.

Es por lo que es indispensable el desarrollo de técnicas automáticas de segmentación. Aunado a lo anterior, obtener el conjunto de procesos necesarios para una segmentación óptima no es sencillo, pues la combinación de procesos es muy grande. Además, no existe una solución general para todos los casos de segmentación. Es por lo que, debido a un espacio de solución muy grande es necesario y justificado el uso de algoritmos evolutivos.

En este trabajo, se desarrolla un algoritmo genético para una correcta segmentación en un determinado conjunto de imágenes. Cada individuo de nuestra población es una

posible solución al problema de la mala segmentación, asimismo el individuo está conformado por distintas variables donde pueden contener técnicas de preprocesamiento, técnicas morfológicas, espacios de color, canales y técnicas de segmentación. El trabajo del algoritmo genético es evaluar a cada individuo, seleccionar a los mejores y hacer una cruce y mutación de cada uno de ellos para así mejorar los individuos en cada generación.

Probar cada una de estas combinaciones de procesos, y encontrar la más adecuada es vital para obtener la extracción de características adecuada y conseguir sistemas automáticos con un mejor desempeño y capaces de implementarse en situaciones de la vida real.

## 1.2. Justificación

La segmentación de imágenes es un proceso muy importante en procesamiento digital de imágenes y sistemas de visión cuyo estudio en los últimos años se ha acentuado para resolver múltiples problemas en la literatura. La correcta segmentación de imágenes influye directamente en el desempeño de un sistema de visión. Es por lo que, el estudio, desarrollo de nuevos algoritmos y aplicación de estos es de suma importancia.

El desarrollo de algoritmos de segmentación automáticos podría permitir mejorar la segmentación de conjuntos de imágenes bajo condiciones de iluminación y fondo similares. Estos algoritmos permitirán facilitar y simplificar la selección de técnicas de segmentación a utilizar, así como las variables necesarias y preprocesamiento implementado.

Para lograr obtener un modelo adecuado para la segmentación de un conjunto de imágenes, es necesario primero comprender la influencia de cada uno de los pasos previos y posteriores a la segmentación, entender como influyen las variables en particular y en combinación al obtener una segmentación.

## 1.3. Objetivos

El objetivo general de esta tesis es desarrollar algoritmos para la segmentación automática de conjuntos de imágenes basados en algoritmos evolutivos.

Se pretende alcanzar los siguientes objetivos específicos:

1. Analizar e identificar procesos que influyen en la segmentación de imágenes.

2. Determinar técnicas de segmentación a utilizar, así como parámetros a utilizar para cada técnica de segmentación
3. Implementar los algoritmos de segmentación automática basados en algoritmos genéticos.

## 1.4. Estado del Arte

En los últimos años ha existido un magno avance en el desarrollo de algoritmos de segmentación de imágenes utilizando algoritmos evolutivos (AE). Las numerosas técnicas desarrolladas pueden clasificarse de acuerdo con las siguientes categorías:

1. Segmentación con AE usando técnicas basadas en umbral
2. Segmentación con AE usando técnicas basadas en *Clustering*
3. Segmentación con AE usando técnicas basadas en detección de bordes
4. Segmentación con AE usando técnicas basadas en regiones
5. Segmentación con AE usando técnicas basadas en redes neuronales

### 1.4.1. Técnicas basadas en umbral

Las técnicas basadas en umbral son métodos ampliamente utilizados en numerosas aplicaciones por su facilidad de implementación. En estos algoritmos, el problema principal es encontrar el umbral óptimo mediante el análisis del histograma, ya que esta tarea suele ser computacionalmente muy costosa, se han utilizado diferentes técnicas desde muy básicas hasta técnicas más complejas ayudadas de algoritmos evolutivos [5] [48] [60] [40]. Los algoritmos propuestos pueden obtener una frontera (solo región deseada y fondo) o fronteras con múltiples niveles (diferentes objetos deseados y fondo). Es claro que el tiempo de cómputo y complejidad se incrementa exponencialmente al aumentar el número de umbrales.

Banimelhem y Yahya [5] utilizan un algoritmo genético para encontrar los umbrales óptimos de segmentación. Los autores utilizan una función de aptitud que avalúa la varianza entre objetos en la imagen y la búsqueda es realizada sobre un cromosoma

de tamaño  $L \times n$ , donde  $L$  es el número de intensidades y  $n$  el número de umbrales utilizado.

Oliveira y Yamanaka [48] Utilizan un algoritmo genético para obtener la frontera óptima utilizando el algoritmo de Otsu [49]. Los autores utilizan la varianza intraclase como función de aptitud y un tamaño de cromosoma definido por el número de intensidades. En sus resultados, los autores comparan sus resultados con el método de Otsu con un umbral y dos umbrales.

De manera similar [40] emplea un algoritmo genético para la selección del umbral óptimo, los picos y valles del histograma. Los autores argumentan que este nuevo enfoque mejora la segmentación a un menor coste computacional que otros métodos como Otsu, teniendo como criterio de paro una función de histograma normalizada. Esta idea la comparte [6] con la diferencia que en [40] agregan una transformada para reducir el tamaño de histogramas y así acelerar la convergencia del algoritmo genético.

En [53] los autores proponen un algoritmo de segmentación de imágenes binarias que contienen objetos circulares y elípticos que han sido afectados artificialmente con diferentes tipos de ruido para probar la calidad de segmentación. El algoritmo propuesto permite seleccionar operadores morfológicos.

Los autores en [51] utilizan un algoritmo de evolución diferencial para encontrar la máxima varianza entre clases, esto es, umbral óptimo del algoritmo de Otsu [49]. Los autores realizan un paso de preprocesamiento para reducir el ruido en la imagen.

### 1.4.2. Técnicas basadas en *Clustering*

Las técnicas de *Clustering* son métodos muy importantes en Aprendizaje máquina. Estas técnicas se denominan técnicas de aprendizaje no supervisado y permiten obtener subgrupos de objetos o datos que son obtenidos basándose en un grado de similitud o disimilitud. Objetos muy similares entre si son agrupados o etiquetados dentro de un grupo, mientras que objetos muy disimilares entre si son agrupados o etiquetados en diferentes grupos. Esto es, los algoritmos utilizados para agrupar maximizan la similitud entre datos que pertenecen a la misma clase y minimizan la similitud que debe haber entre datos que pertenecen a distintas clases.

La implementación de algoritmos genéticos, como apoyo en *Clustering*, ha mejorado drásticamente los avances en una correcta segmentación [1] [58] [16] [2].

En [58] los autores presentan una técnica de *Clustering* difuso basada en evolución diferencial multiobjetivo. Los autores argumentan que las técnicas que optimizan un

solo objetivo no obtienen buenos resultados de segmentación. En sus resultados comparan la técnica propuesta con técnicas que optimizan un solo objetivo. La técnica propuesta codifica los centros de grupos y optimiza múltiples medidas de validez simultáneamente.

En [1] los autores utilizan tres métodos de segmentación de *k-means*, el trabajo del algoritmo genético es ubicar las regiones de color en el espacio HSI para identificar todos los píxeles deseados y no deseados. A esta solución se le asigna el nombre de GAHSI que es la opción de utilizar un algoritmo genético en un solo espacio de color que es el HSI (iniciales de *hue* -tono-, *saturation* -saturación- e *intensity* -intensidad-).

Algunas propuestas como [16] describen un método en donde los cromosomas se formulan en una estructura jerárquica para encontrar el número exacto de segmentos en una imagen de prueba automáticamente. El algoritmo genético se utiliza en conjunto con el algoritmo *Fuzzy c-means* (FCM) para encontrar el número correcto de segmentos. En [2] los autores proponen un nuevo método que combina el algoritmo genético y el algoritmo *K-Means* para segmentar imágenes médicas. En esta técnica combinada, se utiliza un algoritmo genético de longitud de cadena variable para la determinación de los centros óptimos de las regiones segmentadas.

Más recientemente, [45] y [12] han propuesto algoritmos de segmentación basados en técnicas evolutivas. En [45] los autores proponen un algoritmo de búsqueda local de *Clustering* automático, el cual no requiere conocimiento a priori sobre el número de grupos. El algoritmo es hábil para encontrar automáticamente el número de grupos y su configuración.

Chakraborty y Mali [12] proponen un método de segmentación basado en *Fuzzy c-means* para mejorar la eficiencia de segmentación en imágenes biomédicas. El algoritmo utiliza funciones de membresía difusas para localizar las posiciones óptimas de los centros de grupos. El enfoque no depende de las posiciones iniciales de los centros de grupos elegidos.

### 1.4.3. Técnicas basadas en detección de bordes

Los bordes en una imagen son obtenidos por los cambios locales de intensidad. La detección de bordes es definida como el proceso de encontrar los cambios de intensidad en una imagen. La detección de bordes puede ser utilizada como una técnica de segmentación al utilizarse con otros métodos.

Las técnicas de segmentación basadas en bordes trabajan bien sobre imágenes que tienen variación de intensidad pequeñas, sin embargo, en imágenes que tienen una alta

variación la segmentación puede no ser buena. Algunos métodos han sido propuestos con el fin de enfrentar este problema. En [14] los autores proponen un algoritmo basado en *Particle Swarm Optimization* (PSO) que permite mejorar la detección de bordes en comparación con técnicas clásicas como Canny y Prewitt.

Kumar y Raheja [37] proponen un método de optimización basado en Colonia de Hormigas (ACO por sus siglas en inglés *Ant Colony Optimization*). En el método propuesto, los valores de intensidad de la imagen de entrada son transformados en valores de feromonas que son utilizados para detectar los bordes en la imagen.

Diaz et al [17] proponen un algoritmo de segmentación basado en un algoritmo genético. El algoritmo genético es utilizado para reducir falsos bordes a partir de los bordes delgados y cerrados. De acuerdo con los autores, el algoritmo es sensitivo a ruido.

#### 1.4.4. Técnicas basadas en regiones

Estas técnicas funcionan de una forma parecida a las técnicas de *clustering*, sin embargo, no son iguales. A diferencia de las técnicas de *Clustering*, las técnicas basadas en regiones utilizan dos pasos generalmente, en un primer paso se divide recursivamente la imagen utilizando un criterio de homogeneidad y en un segundo paso se agrupan las regiones con intensidades similares una vez más utilizando un criterio de homogeneidad. Es claro que estos criterios influyen en la división de regiones y posterior crecimiento de regiones [73].

En [19] los autores utilizan un grafo con pesos no direccionado para describir una imagen sobreesegmentada. En el grafo cada nodo representa una región y los pesos de los bordes miden la disimilitud entre regiones de acuerdo con sus intensidades. Los autores utilizan una heurística para obtener el grafo.

Freedman y Daniel [21] mejoran el método de caminata aleatoria para segmentar. En su nuevo algoritmo proponen una nueva forma de calcular los pesos de los bordes y utilizan una distribución de probabilidad de colores. Además, substituyen el tradicional grafo por un nuevo grafo basado en sobreesegmentación de cambio medio.

Shi y Malik [59] proponen un novedoso método de segmentación basado en grafos. El método en lugar de enfocarse en características locales y sus consistencias de los datos de la imagen, se enfoca en la impresión global de una imagen. Los autores utilizan el criterio *Normalized cuts* para medir la disimilitud total entre diferentes grupos y la similitud entre individuos de un grupo.

Faris et al. [3] propone un sistema de segmentación de regiones utilizando PSO. El algoritmo SRG inicia con un píxel semilla y compara este con sus píxeles vecinos de acuerdo a algunos atributos como intensidad o textura. En cada iteración los va uniendo cuando son lo suficientemente parecidos, este proceso de prueba y ordenamiento se repite para los siguientes vecinos. Los autores utilizan *Particle Swarm Optimization* (PSO) para automatizar la selección del píxel semilla inicial y el valor de umbral que mide la diferencia entre el píxel y sus vecinos.

En [7] los autores utilizan un algoritmo genético para encontrar la combinación de parámetros óptimos que maximizan un criterio de calidad de segmentación. Los autores utilizan el algoritmo de segmentación *Phoenix* [39], que es una técnica de segmentación de división recursiva. Los autores argumentan que el algoritmo puede trabajar con variaciones en la imagen que afectan la calidad de segmentación debido a que la función objetivo que utilizan se adapta a esas variaciones.

#### 1.4.5. Técnicas basadas en Redes neuronales

Por otro lado, los algoritmos genéticos no solo se han ocupado de apoyo, si no se han combinado con métodos basados en redes neuronales. En la mayoría de los casos, los algoritmos genéticos son utilizados para optimizar los parámetros de la red neuronal [24].

Gómez y Flores [24] utilizan una Red neuronal acoplada por pulsos (PCNN) para segmentación de imágenes. Sin embargo, el desempeño de PCNN se ve afectado por los hiperparámetros, para enfrentar este problema, los autores utilizan un algoritmo de evolución diferencial. En sus experimentos los autores utilizan el índice de validez de *Clusters* como función de aptitud.

Otros casos como [18] combina una red neuronal con un algoritmo genético para una segmentación de tejidos en imágenes de ultrasonido, donde la topología y el número de los nodos de la red neuronal, es determinada automáticamente por los algoritmos genéticos durante el entrenamiento. También es el caso de [69] artículo en el que se propone un nuevo algoritmo de segmentación de imágenes utilizando una combinación de una red neuronal auto generada y el algoritmo genético. Usando algoritmo genético se seleccionan como semillas las muestras óptimas y estas se toman como árboles neuronales iniciales, y así se genera un bosque neuronal auto generado entrenando el resto de las muestras.

Existen procesos semi automáticos de segmentación como se describe en [32] donde se necesita un experto en el área que pueda delimitar la zona para identificar el tejido

óseo esponjoso que se quiera segmentar en una tomografía computarizada, y posteriormente el algoritmo genético busca simultáneamente los filtros gráficos correctos, parámetros precisos para estos filtros, así como el orden correcto de aplicación de los filtros.

A continuación, se presenta un cuadro conceptual (1.1), donde se puede apreciar una manera de clasificar los artículos previos.

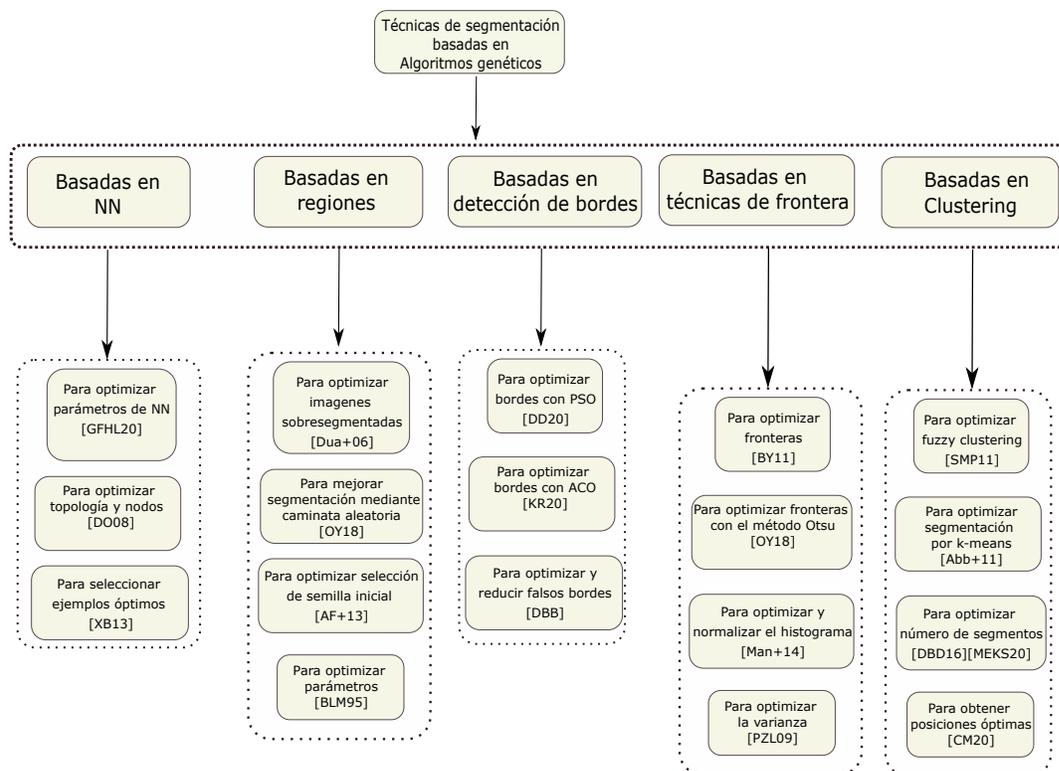


FIGURA 1.1: Artículos clasificados del Estado del Arte

Algunos artículos discuten algoritmos genéticos para mejorarlos o combinarlos con una única técnica de segmentación. Otros utilizan algunos operadores morfológicos, pero de la misma forma, se aplican al resultado de una sola técnica o combinación. Algunas propuestas no son del todo automáticas y requieren la intervención de un experto en el área, y otros enfoques corren el riesgo de caer en un óptimo local y resultar en una mala segmentación.

## 1.5. Organización de la Tesis

El resto de la Tesis se organiza de la siguiente manera:

El Capítulo 2 contiene una descripción general de las técnicas de segmentación y su clasificación, las métricas de desempeño que evaluarán nuestro trabajo y el marco teórico de un algoritmo genético.

El Capítulo 3 muestra la metodología propuesta de este trabajo, los procesos y las herramientas utilizadas. Se detalla un algoritmo genético para ser aplicado en la solución de los problemas de segmentación.

El Capítulo 4 contiene los resultados experimentales de cada conjunto de datos probados, utilizando métricas de desempeño y comparando la efectividad con técnicas de segmentación en el estado del arte.

Finalmente, el Capítulo 5 está compuesto por el resumen del trabajo completo, donde también se comentan algunas conclusiones e ideas para trabajos posteriores.



# Capítulo 2

## Preliminares

### 2.1. Técnicas de segmentación

El procesamiento digital de imágenes comprende varios métodos y técnicas que son utilizadas para extraer información de una imagen. Uno de los pasos más importantes lo constituye la segmentación de imágenes que permite descomponer una imagen en sus partes constituyentes, es decir, los objetos que son de interés y el fondo. Para ello se utilizan ciertas características locales que permiten distinguir un objeto de otro, simplificando la representación de una imagen.

Un píxel representa el elemento de una imagen (*picture element* por sus siglas en inglés), cada píxel nos da la información necesaria en términos de su nivel en gris, color, textura, forma, etc. Cada intensidad de un píxel es representada por un simple número entre 0 y 255. La segmentación es el procedimiento de asignación de un valor a cada píxel que conforma una imagen, formando grupos de píxeles con el mismo valor. En otras palabras, la segmentación es una clasificación con múltiples clases donde cada clase tiene características o propiedades específicas de la imagen a examinar [71] [35].

La segmentación se basa en los siguientes principios:

1. Similitud: Cada uno de los píxeles de un elemento tiene valores parecidos con respecto a alguna propiedad específica. En este principio, se estudia la regularidad en los valores del nivel de gris.
2. Discontinuidad: Los objetos y más concretamente las regiones que los distinguen destacan del entorno y tienen por tanto unos bordes definidos.
3. Conectividad: los píxeles pertenecientes al mismo objeto o región tienen que ser contiguos, es decir, tienden a agruparse constituyendo regiones homogéneas.

La segmentación se puede representar de dos maneras: segmentación semántica y segmentación de instancias.



FIGURA 2.1: Segmentación semántica <sup>1</sup>

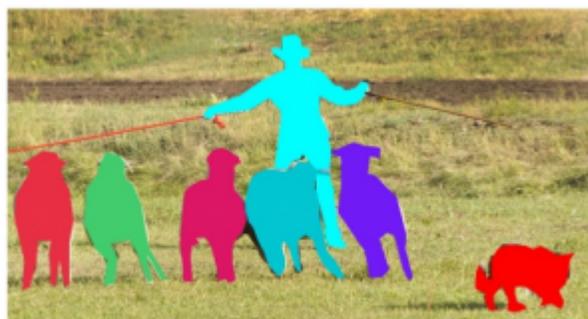


FIGURA 2.2: Segmentación de instancias <sup>2</sup>

Donde la segmentación semántica se refiere al proceso de clasificar varios objetos en un mismo grupo (ver Figura 2.2 <sup>1</sup>). Mientras que la segmentación de instancias permite clasificar diferentes objetos a varios grupos con las mismas características (ver Figura 2.2 <sup>2</sup>).

Por otro lado, la segmentación panóptica es la unión de la segmentación semántica y la segmentación de instancias. El resultado de una segmentación panóptica es un nuevo modelo que tendrá las peculiaridades de los dos modelos anteriores, la etiqueta del píxel (segmentación semántica) y la predicción para cada instancia de píxel (segmentación de instancias).

Las técnicas de segmentación de imágenes son utilizadas para dividir una imagen en varios segmentos, estas partes son utilizadas para localizar los objetos y sus fronteras

<sup>1</sup>Figura tomada de [<https://informatica.blogs.uoc.edu/la-segmentacion-semantica-y-sus-benchmarks/>]

<sup>2</sup>[Figura tomada de <https://informatica.blogs.uoc.edu/la-segmentacion-semantica-y-sus-benchmarks/>]

en la imagen. Una efectiva segmentación es realizada cuando se separan los objetos en la imagen en un tiempo computacional aceptable y describiendo las fronteras del objeto perfectamente.

De forma general, los métodos de segmentación pueden ser divididos en dos grandes partes: métodos de segmentación basados en píxeles y métodos de segmentación basados en *texels*, los más estudiados son los primeros. A lo largo de las últimas décadas se han desarrollado diversas técnicas de segmentación, en la Figura 2.3 se describen las técnicas de segmentación más elementales.

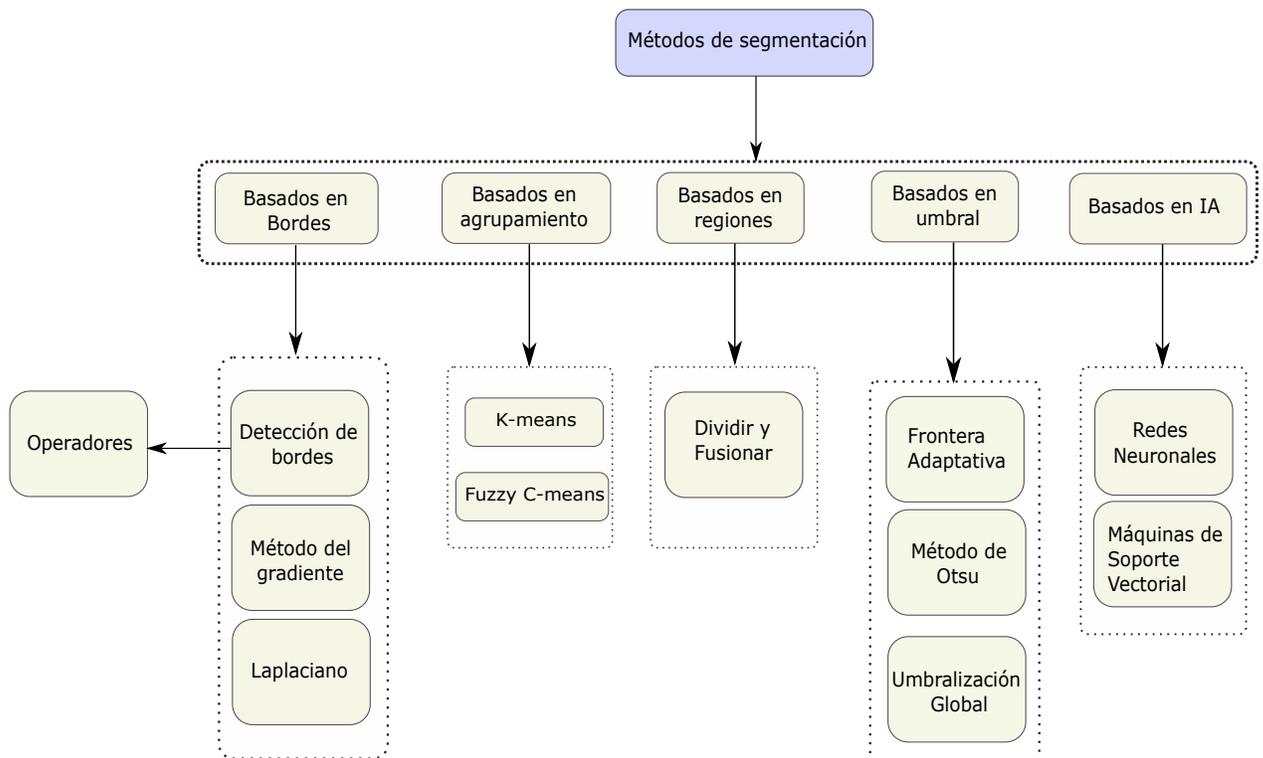


FIGURA 2.3: Métodos de segmentación más elementales

En los últimos años ha existido un magno avance en el desarrollo de algoritmos de segmentación de imágenes. Las numerosas técnicas desarrolladas pueden clasificarse de acuerdo con las siguientes categorías:

1. Técnicas de segmentación basadas en umbral
2. Técnicas basadas en *Clustering*
3. Técnicas basadas en detección de bordes
4. Técnicas basadas en regiones

## 5. Técnicas basadas en redes neuronales

### 2.1.1. Técnicas de segmentación basadas en umbral

#### Segmentación por frontera global

Esta es una técnica de segmentación rápida de coste computacional bajo. La obtención del umbral está considerada bajo diversos criterios, ya que puede ser manejado por el mismo usuario, donde se establece un punto de umbral o se obtiene de forma automática [71].

Para obtener el umbral de manera automática es posible hacerlo de dos formas, encontrando el umbral con la suma del píxel con la intensidad más alta con el píxel de intensidad más baja y dividirlo entre dos, como se muestra en la Ecuación (2.1). El otro modo es mediante la normalización del histograma de una imagen  $I$  de tamaño  $M \times N$ , esta técnica es sencilla porque al obtener el histograma normalizado, que como bien se sabe es la sumatoria de las probabilidades de existencia de cada píxel, no considera información espacial, sino solamente la distribución de grises en la imagen, la Ecuación (2.2) muestra cómo obtener el histograma normalizado.

$$th = \frac{\text{máx}(I) - \text{mín}(I)}{2} \quad (2.1)$$

$$th = \frac{\sum_{i=1}^m \sum_{j=1}^n f(i, j)}{M \times N} \quad (2.2)$$

En pocas palabras, la segmentación por frontera global permite convertir una imagen de niveles de gris en una imagen binaria, esto por medio de la Ecuación (2.3), donde si cumple la condición de umbral, la salida será 1, en caso contrario 0.

$$G(x, y) = \begin{cases} 1, & f(x, y) > th \\ 0, & f(x, y) < th \end{cases} \quad (2.3)$$

donde  $th$  = Representa el valor del umbral asignado. La razón del uso de un histograma es la distribución binomial que se maneja, para así poder implementar un umbral que permita visualizar diversas características de la misma imagen evitando de esta forma la perdida de información que podría llegar a ser importante.

### Frontera adaptativa

La frontera adaptativa al igual que la frontera global, están basadas en un umbral, ambos devuelven una imagen binaria, pero la diferencia que existe entre ambos es el algoritmo para la obtención del umbral. El algoritmo aquí presente muestra cómo obtener el umbral por este método.

1. Se comienza seleccionando una frontera inicial  $th(i)$ , donde  $i = 1$ . La frontera inicial puede ser cualquiera, en la mayoría de los casos  $th(1) = 127$ . Este umbral puede ser de igual forma obtenido con la Ecuación (2.1).
2. Se agrupan los píxeles usando  $th(1)$  creando así dos grupos. Donde el primer grupo es para todos aquellos píxeles con intensidades mayores a  $th$  y el segundo para los píxeles con intensidades menores a  $th$ . La Ecuación (2.4) y la Ecuación (2.5) muestran el agrupamiento de píxeles.

$$G1 \in [0, \dots, th] \quad (2.4)$$

$$G2 \in [th + 1, \dots, L] \quad (2.5)$$

donde  $L =$  Intensidad máxima presente en la imagen

3. Se obtienen las medias de cada grupo. Las Ecuaciones (2.6) y (2.7) muestran las ecuaciones utilizadas.

$$\mu_1 = \sum_{i=0}^m \frac{G1(i)}{m} \quad (2.6)$$

$$\mu_2 = \sum_{i=m+1}^n \frac{G2(i)}{n} \quad (2.7)$$

4. Una vez obtenida la media de ambos grupos, se actualiza la frontera  $th(i)$  utilizando la Ecuación (2.8).

$$th(i) = \frac{\mu_1 + \mu_2}{2} \quad (2.8)$$

5. Finalmente, este proceso se realiza hasta que  $th(i)$  no se mueva con respecto al resultado de dos iteraciones anteriores. En caso contrario se siguen ejecutando todos los pasos. Una vez obtenido el umbral se aplica la misma condición mostrada en la Ecuación (2.3).

## Método de Otsu

El método de Otsu, llamado así en honor a Nobuyuki Otsu que lo desarrolló en 1979 [49], utiliza técnicas estadísticas, para resolver el problema. En concreto, se utiliza la variancia, que es una medida de la dispersión de valores, en este caso se trata de la dispersión de los niveles de gris.

El método de Otsu calcula el valor umbral de forma que la dispersión dentro de cada segmento sea lo más pequeña posible, pero al mismo tiempo sea la más alta posible entre segmentos diferentes. Para ello se calcula el cociente entre ambas variancias y se busca un valor umbral donde este cociente sea máximo.

El método Otsu elige el umbral óptimo maximizando la varianza entre clases (*intra-class variance*) mediante una búsqueda exhaustiva. Si bien hay diferentes métodos para hallar un umbral, la mayoría de ellos no dan buenos resultados cuando se trabaja con imágenes del mundo real debido a la presencia de ruido, histogramas planos o una iluminación inadecuada. Por el contrario, el método de Otsu fue uno de los mejores métodos de selección de umbral para imágenes.

A medida que el número de clases de una imagen aumenta, el método de Otsu necesita mucho más tiempo para seleccionar un umbral adecuado. Para determinar el umbral de una imagen eficientemente, se obtiene una varianza entre dos diferentes clases de intensidades agrupadas que se modifican por cada iteración.

La importancia del método de Otsu radica en que es automático, es decir, no necesita supervisión humana ni información previa de la imagen antes de su procesamiento. Para realizar segmentación mediante el procedimiento del método Otsu existe un algoritmo que permite llevar a cabo este tipo de segmentación. Dicho algoritmo se explica paso a paso a continuación.

1. Para llevar a cabo la segmentación por el método Otsu es necesario normalizar el histograma de la imagen.

2. Una vez normalizada la imagen, se asume que los píxeles son divididos en dos clases:  $C_0$  con niveles de gris hasta un umbral  $K$ , y  $C_1$  con niveles de gris empezando por  $K + 1$  hasta la intensidad más alta  $L$ . Las expresiones (2.9) y (2.10) muestran el agrupamiento para cada clase.

$$C_0 \subset \{1, 2, \dots, k\} \quad (2.9)$$

$$C_1 \subset \{k + 1, \dots, L\} \quad (2.10)$$

3. Se obtiene la distribución de las probabilidades de ocurrencia de las intensidades que existen en cada clase, para poder obtenerlas se aplica la ecuación (2.11) para la clase  $C_0$  y la Ecuación (2.12) para la clase  $C_1$ .

$$w_0 = P_r(C_0) = \sum_{i=1}^k P_i = w(k) \quad (2.11)$$

$$w_1 = P_r(C_1) = \sum_{i=k+1}^L P_i = 1 - w(k) \quad (2.12)$$

4. En seguida se obtienen las medias respectivas por cada clase. Para la clase  $C_0$  se ocupa la Ecuación (2.13) y para la clase  $C_1$  se emplea la Ecuación (2.14).

$$\mu_0 = \sum_{i=1}^k i P_r(i | C_0) = \sum_{i=1}^k i \frac{P_i}{w_0} = \frac{\mu_k}{w(k)} \quad (2.13)$$

$$\mu_1 = \sum_{i=k+1}^L i P_r(i | C_1) = \sum_{i=k+1}^L i \frac{P_i}{w_1} = \frac{\mu_L}{1 - w(k)} \quad (2.14)$$

5. Se obtienen las varianzas, para la clase  $C_0$  ocupamos la Ecuación (2.15) y para la clase  $C_1$  ocupamos la Ecuación (2.16).

$$\sigma_0^2 = \sum_{i=1}^k (i - \mu_0)^2 P_i(i | C_0) = \sum_{i=1}^k \frac{(i - \mu_0)^2 P_i}{w_0} \quad (2.15)$$

$$\sigma_1^2 = \sum_{i=k+1}^L (i - \mu_1)^2 P_i(i | C_1) = \sum_{i=k+1}^L \frac{(i - \mu_1)^2 P_i}{w_1} \quad (2.16)$$

- Finalmente se calcula la Ecuación (2.17) para cada iteración. Una vez se tenga el resultado final por cada nivel de intensidad, se toma el valor resultante más pequeño como umbral.

$$\sigma_w^2 = w_0\sigma_0^2 + w_1\sigma_1^2 \quad (2.17)$$

### 2.1.2. Métodos basados en bordes

La segmentación basada en bordes permite segmentar una imagen identificando los bordes o fronteras de los objetos. Las fronteras son localizadas donde la intensidad en la imagen cambia y también es utilizada para identificar discontinuidades en la imagen. El propósito de la detección de bordes es reducir significativamente la cantidad de datos a ser procesados en la imagen mientras que preserva las propiedades estructurales de la imagen [66] [65].

Las técnicas de Sobel y el operador de Prewitt son clasificadas como técnicas de primer orden. Estas técnicas detectan los bordes mediante la búsqueda de los máximos y mínimos de la primera derivada de la imagen. La disimilitud en los píxeles es el principio básico de las técnicas basadas en detectores de bordes.

#### Detección de bordes

El borde en una imagen se obtiene al detectar una considerable diferencia de intensidad entre dos píxeles. Esto es, las técnicas de detección de bordes encuentran discontinuidades de color o cambios abruptos de brillo en ciertas regiones de la imagen.

Una segmentación exitosa basada en bordes se debe a tres pasos clave: Detectar bordes, eliminar bordes irrelevantes y conectar o agrupar [22] [13].

Las desventajas de estas técnicas de segmentación son:

- La técnica sensible al ruido.
- La técnica no es apropiada si el borde de la imagen está mal definido. La presencia de ello provocará que el rendimiento no sea óptimo.
- Existen errores considerables en la segmentación, si existen muchos bordes en la imagen.

**Operadores de detección de bordes** El operador de Sobel es un operador clásico de detección de borde de primer orden utilizado para aproximar la magnitud absoluta del gradiente de cada píxel mediante matrices de convolución o un par de filtros de  $3 \times 3$ . Estos filtros estiman los gradientes en las direcciones horizontal y vertical y las magnitudes de la pendiente.

Los operadores de Sobel son menos sensibles al ruido (ya que utiliza *kernels* de convolución más grandes) y puede suavizar la imagen de entrada en mayor medida. Pero produce bordes más gruesos, para evitar este problema se necesita adicionalmente un proceso de adelgazamiento. El detector de bordes de Prewitt funciona de una manera similar al operador de Sobel, pero utiliza un *kernel* diferente. El operador de Prewitt es un detector de bordes, basado en filtros pequeños.

El detector de bordes de Canny es un operador de detección de bordes óptimo utilizado para detectar una amplia gama de bordes en imágenes. La entrada de un operador de Canny es una imagen de escala de grises [8].

### Método del gradiente

La idea principal es que la magnitud del gradiente es mucho mayor en el borde que en puntos distanciados al borde.

Desafortunadamente, en la práctica, este método obtiene elementos que no pertenecen a la estructura y regularmente faltan elementos que sí pertenecen. Para superar este problema, existen diversas estrategias de optimización. Una posible solución es asignar un coste a cada elemento del borde y encontrar un borde cuyo coste total es el menor entre todos los posibles. Este método exige un elevado poder computacional.

### Laplaciano

Para el método del gradiente se ocupa una sola derivada para diferenciar la imagen, pero en el método de Laplaciano se ocupa una segunda derivada para estimar mejor la posición de la orilla, donde la segunda derivada cruza cero.

El Laplaciano de una imagen  $I(x, y)$  es una derivada de orden dos definida por:

$$\nabla^2 I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \quad (2.18)$$

En general se suele tomar el valor negativo del Laplaciano. Al igual que en el caso del gradiente se puede implementar en forma digital de varias formas.

Puesto que el Laplaciano es un operador de derivación, la suma de los coeficientes debe ser cero. además, el coeficiente asociado con el píxel central debe ser positivo y los demás coeficientes negativos o ceros. En la Figura (2.4) podemos ver un filtro para el cálculo del Laplaciano.

0	-1	0
-1	4	-1
0	-1	0

FIGURA 2.4: Filtro utilizado para el Laplaciano

En este caso la expresión para determinar el Laplaciano está dada por la Ecuación (2.19).

$$4z_5 - (z_2 + z_4 + z_6 + z_8) \quad (2.19)$$

Aunque el Laplaciano responde a transiciones en la intensidad de la imagen, se emplea en pocas ocasiones en la práctica. Debido a que es un operador de segunda derivada es sensible en exceso a la presencia de ruido.

### 2.1.3. Métodos Basados en *Clustering*

Basados en encontrar regiones o subregiones donde existe una constante en esa parte de la imagen. La similitud de los píxeles es el principio básico de segmentación basada en agrupamiento.

En estos métodos, los píxeles se agrupan en regiones más grandes según su similitud de acuerdo con criterios de similitud predefinidos y considerando las relaciones espaciales de adyacencia entre píxeles. Ejemplos simples de criterios de similitud podrían ser [22]:

- La diferencia de intensidad absoluta entre un píxel candidato y el píxel inicial debe estar dentro de un rango definido
- La diferencia de intensidad absoluta entre un píxel candidato y la intensidad media de funcionamiento de la región de cultivo debe estar dentro de un rango definido

→ La diferencia entre la desviación estándar en intensidad sobre una vecindad local específica del píxel candidato y la de una vecindad local del píxel candidato debe (o no debe) exceder un cierto umbral, este es un criterio básico de rugosidad / suavidad.

### **K-means**

*K-means* es un método de agrupamiento, que tiene como objetivo la partición de un conjunto de  $n$  elementos en  $k$  grupos en el que cada elemento pertenece al grupo más cercano. Es generalmente un método utilizado en minería de datos para calcular la distribución de un conjunto de elementos con respecto a un número de centroides definido.

El algoritmo que se utilizará está basado en los pasos aquí mostrados.

1. Se eligen  $k$ -puntos como centroides.
2. Se calcula la distancia de todos los elementos, con respecto a los centroides. Generalmente esta distancia es calculada haciendo uso de distancias euclidianas y agregando el punto al centroide con la distancia más corta.
3. Una vez obtenido el grupo de puntos por cada centroide, se calcula la media por cada punto y se actualiza la posición del nuevo centroide.
4. Este proceso se realiza hasta que no existan cambios en los centros.

#### **2.1.4. Métodos Basados en regiones**

Los métodos basados en regiones son métodos que inician por lo general con un píxel y van agregando píxeles de acuerdo con un parámetro de similitud.

Punam Thakare y Parandwadi [54] argumentan que la técnica de región creciente es uno de los métodos más populares de segmentación. Este método inicia con un píxel y va agregando píxeles en base a su similitud formando una región. El proceso total es repetido hasta que todos los píxeles conectados pertenecientes a la misma región son fácilmente encontrados. Sin embargo, una desventaja de este método es que la presencia de ruido provoca un pobre desempeño y es computacionalmente muy costosa.

## Dividir y Fusionar

Como su nombre lo indica, el primer paso de este método es dividir nuestra imagen principal, con base en una estructura de datos de un árbol cuádruple. La imagen se divide en cuatro cuadrantes, y cada cuadrante se puede dividir en cuatro más, este proceso se puede repetir una y otra vez hasta alcanzar el objetivo de tener píxeles de un mismo color en un solo cuadrante [71] [64] [52].

El segundo paso fusiona las regiones adyacentes y similares. Este último proceso se repite hasta que no sea posible realizar más fusiones [71] [47].

### 2.1.5. Métodos Basados en IA

Los métodos basados en inteligencia artificial (IA) son una serie de técnicas que emulan características o capacidades exclusivas del intelecto humano, maximizando el rendimiento por medio del aprendizaje de la experiencia y ajustándose a nuevas aportaciones.

#### Redes neuronales

Las redes neuronales artificiales (RNA) son un modelo donde simulan el comportamiento de las redes neuronales biológicas, su objetivo es resolver problemas mediante el flujo de información ajustando valores en los enlaces de estas neuronas. En general, una red neuronal artificial está constituida de tres capas: capa de entrada, capa(s) oculta(s) y capa de salida.

Una de las ventajas de este modelo, es el alto grado de paralelismo que permite reducir el tiempo computacional de solución de un problema. Este modelo aprende y corrige su error mediante una técnica llamada propagación hacia atrás.

#### Máquinas de soporte vectorial

Las máquinas de soporte vectorial (SVM por sus siglas en inglés *Support Vector Machines*), es un método de clasificación binaria. Las SVM permiten obtener un hiperplano óptimo, de modo que el margen de separación entre las dos clases se amplía al máximo.

En el contexto de segmentación, se puede utilizar para clasificar píxeles, aquellos píxeles que tienen ciertas características en común pueden ser tratados como objetos de interés, y aquellos píxeles que tengan un valor diferente, son tratados como fondo.

## 2.2. Espacios de color

El espacio de color es una organización o gama de colores específica en una imagen, video, ordenador o dispositivo. En otras palabras, es una combinación de pigmentos de colores primarios utilizando una función de mapeo. El objetivo de un modelo de color es facilitar la especificación de los colores de una forma normalizada. En otras palabras, un modelo de color podría ser definido como un sistema de coordenadas tridimensional en el que cada color queda representado por un único punto.

La importancia del color radica en que es un poderoso descriptor para identificar y extraer un objeto de una escena. También se debe de mencionar que aun cuando los humanos pueden distinguir miles de colores, a su vez solo distingue dos docenas de grises. Los modelos de color más habituales son RGB (rojo, verde, azul) para monitores y videocámaras; CMY (cian, magenta, amarillo) para impresoras y YIQ (donde Y es brillo e I y Q son componentes cromáticos) que es el estándar para televisión [25].

En la literatura, los espacios de color para el procesamiento de imágenes en color son los siguientes: RGB, HSV (tono, saturación, valor), HSI (tono, saturación, intensidad), CIE XYZ, NTSC, YUV y YCbCr [22]. A continuación, se presenta las características de los espacios de color previos.

### 2.2.1. Espacio RGB

En el modelo RGB cada color aparece en sus componentes espectrales: rojo, verde y azul. Este modelo está basado en el sistema de coordenadas cartesianas. El subespacio de color de interés es el tetraedro mostrado en la Figura (2.5), en el cual los valores RGB están en tres vértices.

El color negro corresponde al punto de origen y el blanco se sitúa en el vértice opuesto, el más alejado del origen. En este modelo, la escala de grises se extiende a lo largo de una diagonal desde el punto negro al punto blanco, y los colores son puntos dentro del tetraedro definidos por los vectores desde el origen.

Por conveniencia, todos los vectores son normalizados, de modo que el tetraedro de la figura es el tetraedro unitario, es decir, todos los valores de R, G y B están en el rango [0,1]. Las imágenes en este modelo se forman por la combinación en diferentes proporciones de cada uno de los colores primarios RGB.

El espacio RGB se podría decir que es un modelo simple, en varios estudios se emplea para el procesamiento de color o cuando es necesario transformar colores a un espacio de color diferente.

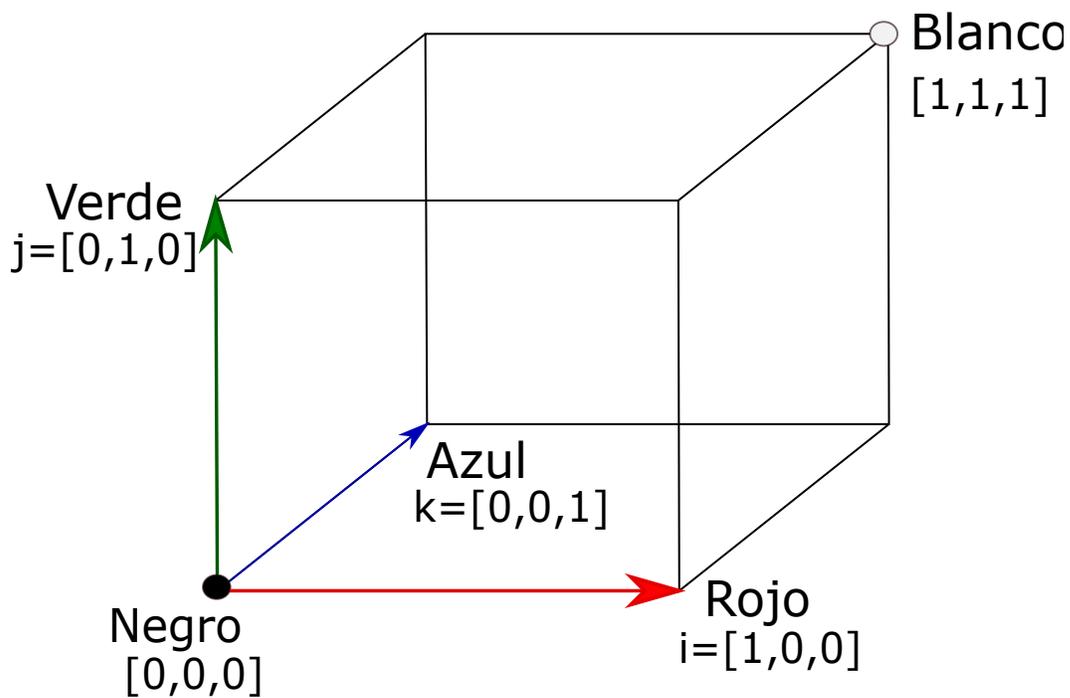
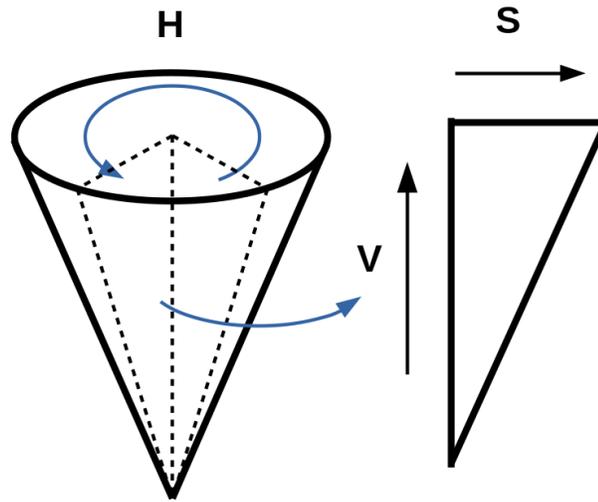


FIGURA 2.5: Modelo RGB

### 2.2.2. Espacio HSV

En este espacio, el color se representa por los componentes tono (H), saturación (S) y valor (V). El tono es la característica cromática que describe un color puro; por ejemplo, amarillo, naranja, rojo, etc. La saturación es una medida de cómo se diluye el tono en la luz blanca; El valor es la intensidad o el brillo del color.

El espacio HSV se convierte en una herramienta útil para desarrollar algoritmos de procesamiento de imágenes basados en algunas propiedades de la percepción humana del color [25] [31].

FIGURA 2.6: Modelo HSV <sup>3</sup>

El tono está en el rango  $[0, 2\pi] \subset \mathbb{R}$ ; la saturación está en el rango real  $[0, 1]$ , mientras que el valor suele estar en el rango  $[0, 255] \subset \mathbb{R}$ . El espacio HSV tiene forma donde la punta es el color negro y se va aclarando hasta llegar a blanco que es la base, el radio y la altura del cono representan los componentes de saturación y valor, respectivamente. En la Figura (2.6 <sup>3</sup>) se puede observar su representación.

Teniendo en cuenta que, para los colores negro, blanco y gris, el parámetro de tono no está definido, debido a que estos colores se consideran singularidades dentro de este espacio de color; porque no tienen una cromaticidad específica [22].

### 2.2.3. Espacio HSI

Sus siglas corresponden a H (Tono), S (Saturación), I (Intensidad). El modelo HSI tiene una utilidad por dos características:

- La componente de intensidad, I, está desacoplada de la información cromática contenida en la imagen.
- Las componentes de tono y saturación están estrechamente relacionadas con la forma en que el ser humano percibe el color.

<sup>3</sup>[Tomado de [https://www.researchgate.net/figure/Figura-23-Espacio-de-colores-HSV\\_fig2\\_326551993](https://www.researchgate.net/figure/Figura-23-Espacio-de-colores-HSV_fig2_326551993)]

Los rangos de los componentes son los mismos que el espacio HSV. Es decir, el tono está en el rango  $[0, 2\pi] \subset R$ ; la saturación está en el rango real  $[0, 1]$ , mientras que la intensidad está en el rango  $[0, 255] \subset R$ . La diferencia es que el espacio HSI tiene forma de doble cono, por lo tanto, es diferente calcular sus componentes.

## 2.3. Algoritmos Genéticos

Los algoritmos genéticos (AGs), parten de la premisa de la teoría de la evolución para optimizar una solución a partir de mejorar en cada generación una población. Para ello, emplean la selección natural de los individuos más aptos como solucionadores de problemas de optimización [23] [50].

La optimización se realiza mediante el intercambio natural de material genético entre los padres. Los descendientes se forman a partir de genes individuales de los padres. Se evalúa la aptitud de la descendencia y los individuos con mejor aptitud son los que tienen más probabilidades de reproducirse.

El algoritmo genético (GA) [23] se puede usar para encontrar el subconjunto de características [27] [46] [4] [34] [70] [55] en donde los bits del cromosoma representan si la característica está incluida o no. Se puede encontrar el máximo global para la función objetivo que proporciona el mejor subconjunto subóptimo. Aquí nuevamente la función objetivo es el desempeño del predictor.

### Elementos de un algoritmo genético.

Para ejecutar un AG, se requiere de una población de individuos. Cada individuo, es un candidato a ser la solución del problema tratado, o permite llegar a la solución a partir de este.

Cada individuo de la población se representa con una cadena binaria y se denomina genotipo del individuo que es análoga al cromosoma en el sistema biológico.

Cada genotipo representa a puntos  $x$  del espacio de búsqueda del problema. A cada punto  $x$  se le denomina fenotipo. Se usa el termino gen para referirse a la codificación de una determinada característica del individuo. Cada gen puede tomar distintos valores que son llamados alelos.

En la Figura (2.7) se puede observar gráficamente la representación de algunas expresiones.

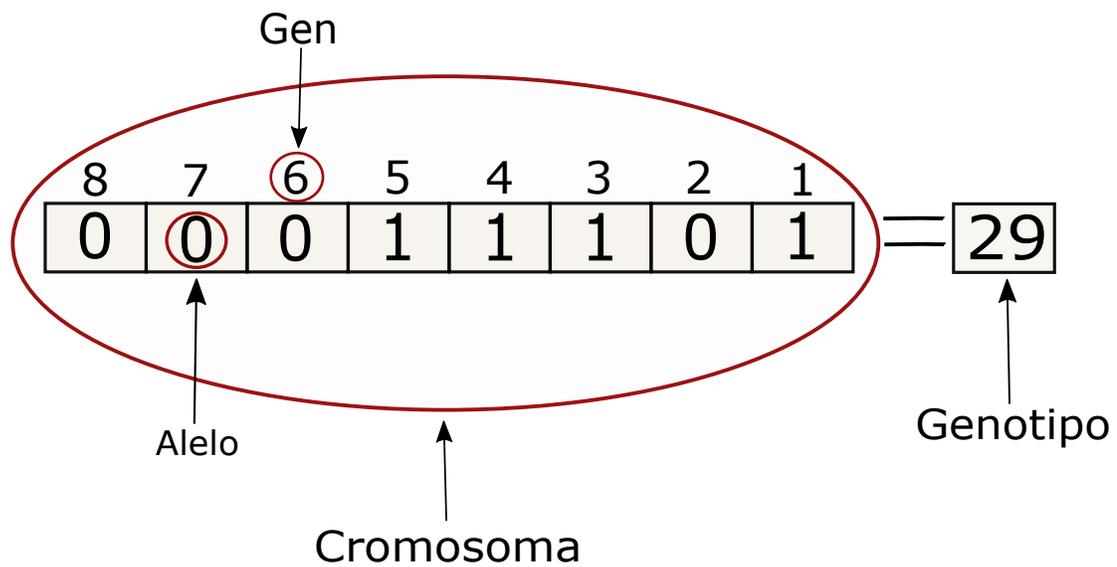


FIGURA 2.7: Representación gráfica de las expresiones que se usa en la genética

### 2.3.1. Algoritmo genético básico

Se expone un algoritmo genético básico, con el objetivo de explicar con claridad el funcionamiento de un AG. El término básico o simple, es debido a que en cada una de sus etapas se aplican las elecciones más sencillas posibles. El algoritmo inicia con una población generada aleatoriamente. La función de adaptación es una función matemática para la que se busca el valor óptimo en un determinado intervalo. El algoritmo entra a un ciclo donde el primer paso es una selección de individuos. Esta selección se realiza de tal manera que solo permanezcan los individuos mejor adaptados. Los individuos por cruzar se eligen de forma consecutiva, ya que se supone que el proceso de selección ha reubicado a los individuos de forma eficiente.

Se aplica una mutación aleatoria y se determina el nivel de adaptación de la nueva generación de individuos. El criterio de paro es un número máximo de generaciones, aunque en otros casos se fijan 203 generaciones en las que no hubo mejora de aptitudes.

El esquema general de un algoritmo genético básico se puede observar en el Algoritmo (1)

Las partes de este algoritmo se describe con más detalle a continuación:

---

<b>Entrada:</b>	VARIABLES, tamaño de cromosoma, espacio de búsqueda, etc.
<b>Salida:</b>	Mejor solución

---

- 1: *Generar una población inicial*
- 2: *Evaluar aptitud de población inicial*
- 3: **WHILE** *condición de paro no se cumple* **Do**
- 4: *Selección de individuos para la reproducción*
- 5: *Cruza de individuos*
- 6: *Mutación de individuos*
- 7: *Obtener la nueva generación* **END**

---

**Algoritmo 1:** Algoritmo genético básico

### 2.3.2. Población inicial

Los individuos de la población inicial suelen ser cadenas de ceros y unos generados de forma completamente aleatoria. Es decir, se va generando cada gen, con una función que devuelve un cero o un uno con igual probabilidad. Es importante dotar al algoritmo genético de población con suficientemente variedad, para poder explorar todas las zonas del espacio de búsqueda.

### 2.3.3. Selección de individuos

La idea básica de selección es utilizar una distribución de probabilidad de selección de una cadena, donde la probabilidad es directamente proporcional a la función de aptitud. Es decir, el proceso de selección debe favorecer la cantidad de copias de los individuos más aptos. Las técnicas de selección usadas pueden clasificarse en tres grupos: selección proporcional, selección mediante torneo y selección de estado uniforme. Sin embargo, en este trabajo solo se analizarán algunas técnicas del grupo selección proporcional, para un estudio más a fondo sobre las demás técnicas puede consultar la referencia.

**Ruleta** Este método ha sido el más comúnmente utilizado desde los inicios de los AGs. El algoritmo presenta el problema de que el individuo menos apto puede ser seleccionado más de una vez. Sin embargo, su popularidad se debe a su simplicidad. El algoritmo de la Ruleta es el siguiente:

- Calcular la suma de valores esperados  $T$ .
- Repetir  $N$  veces ( $N$  es el tamaño de la población).

- Generar un número aleatorio  $r$  entre 0.0 y  $T$ .
- Ciclar a través de los individuos de la población sumando los valores esperados hasta que la suma sea mayor o igual a  $r$ .
- El individuo que exceda esta suma es el seleccionado.

### 2.3.4. Cruza

Este es un método de fusión sobre la información genética de dos individuos. Este proceso provee un mecanismo para heredar características a su descendencia donde intervienen ambos padres.

La forma más simple del operador de cruce es el cruce de un solo punto, que consiste en seleccionar una posición en la cadena de ambos padres e intercambiar las partes divididas por dicha posición. La Figura (2.8) muestra un ejemplo de cruce.

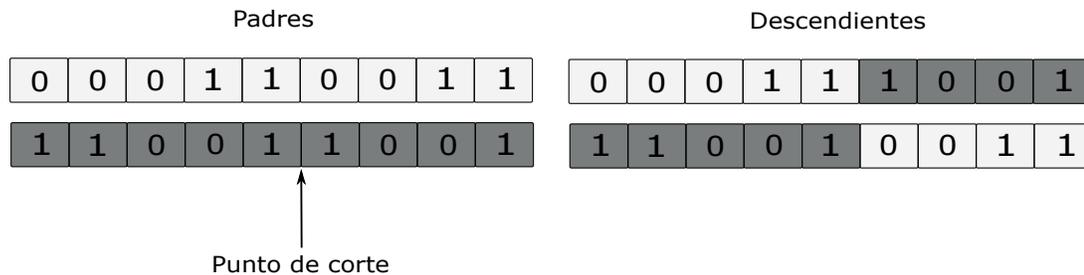


FIGURA 2.8: Operador de Cruza

### 2.3.5. Mutación

La mutación es un proceso donde el material genético puede ser alterado en forma aleatoria, debidamente a un error en la reproducción o la deformación de genes. A diferencia de la genética humana, la probabilidad en un algoritmo genético es mayor y puede ser seleccionado. De hecho, en un algoritmo genético, la mutación es un mecanismo que evita caer en mínimos locales.

La forma más sencilla de mutación consiste en cambiar el valor de una de las posiciones de la cadena. Si el valor es cero pasa a uno, y si es uno pasa a cero. La Figura (2.9) muestra un ejemplo:

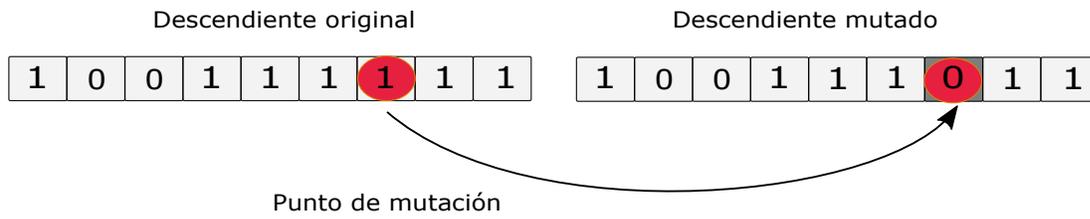


FIGURA 2.9: Operador de Mutación

### 2.3.6. Condición de paro

Es necesario especificar las condiciones en las que el algoritmo deja de evolucionar y se presenta la mejor solución encontrada. La condición de paro más sencilla se presenta al detectar que la mayor parte de la población ha convergido a una forma similar, careciendo de la suficiente diversidad para que tenga sentido continuar con la evolución.

# Capítulo 3

## Metodología

En este Capítulo, se describe paso a paso la metodología utilizada para optimizar la segmentación. La detección de objetos en un sistema de visión ha sido de gran importancia para los avances computacionales y aplicaciones en nuestra vida cotidiana. Las técnicas de visión han sido utilizadas en numerosas aplicaciones como detectar enfermedades en hojas de plantas, detectar diferentes tipos de cáncer, reconocimiento facial para cámaras de seguridad, y detección de construcciones son algunas de las muchas aplicaciones de estos sistemas de visión, donde algunos de estos ejemplos de sistemas de visión, necesitan pasar previamente por un proceso de segmentación. La Figura 3.1 muestra el proceso completo de un sistema de visión y muestra como la segmentación es un solo paso dentro de ese proceso.



FIGURA 3.1: Componentes de un sistema de visión

Para segmentar una imagen existen muchas técnicas de segmentación y preprocesamiento que se pueden utilizar, cada uno de ellas o la combinación de estas, nos darán una determinada calidad en la segmentación. Por otro lado, Cada imagen tiene ciertas características y cada una de las imágenes puede verse afectada por cambios en la iluminación o ruido, por lo tanto, la calidad de segmentación será diferente y dependerá del conjunto de técnicas utilizadas.

Probar cada una de las técnicas de segmentación, cada una de las técnicas de preprocesamientos y escoger la combinación óptima es una tarea muy difícil y costosa. Este trabajo a menudo es realizado por un experto, sin embargo, pequeños cambios en la iluminación o en el ruido pueden afectar significativamente la calidad de la segmentación.

Es por eso la importancia de incluir un algoritmo genético que realice la búsqueda de la combinación adecuada de parámetros y técnicas para obtener la segmentación óptima.

Este trabajo solo incluye algunas técnicas de segmentación en combinación con algunas técnicas de preprocesamiento y operadores morfológicos utilizando algunos espacios de color. La Figura 3.2 muestra a grandes rasgos lo dicho anteriormente. En este Capítulo se describen paso a paso el proceso llevado a cabo.

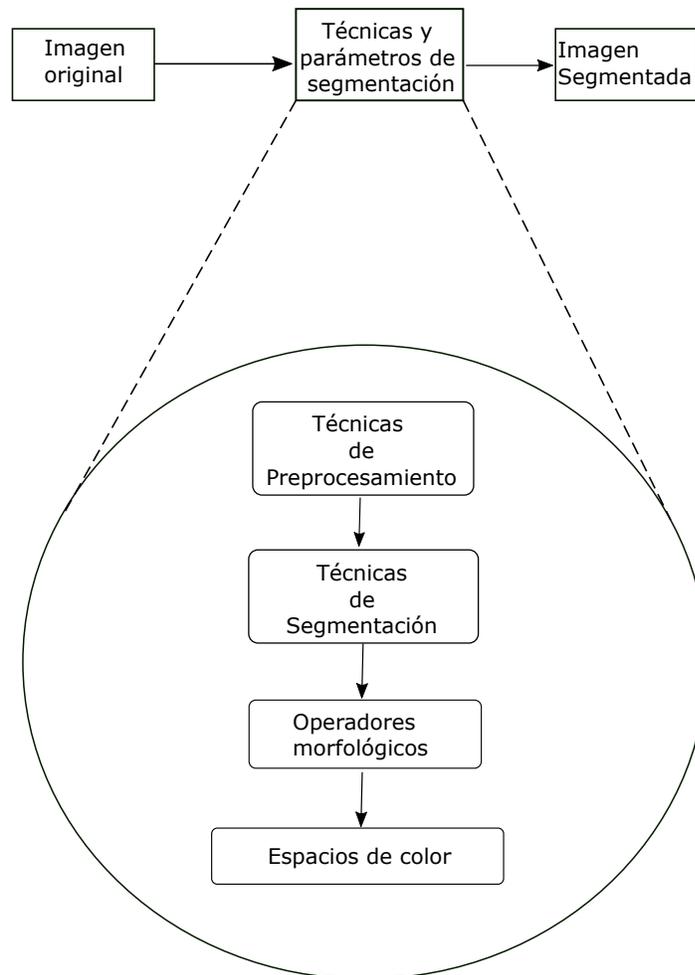


FIGURA 3.2: Técnicas y parámetros necesarios para segmentar

La solución propuesta dentro de la metodología de la Figura 3.2 comienza tomando una imagen original, a la que llamaremos “Imagen E,” con esta imagen se utilizan un conjunto de técnicas con parámetros asociados que fueron previamente optimizadas utilizando un algoritmo genético y se obtiene una “Imagen S”.

El algoritmo genético busca la técnica que mejor resultados obtiene a partir de varias técnicas de preprocesamiento, segmentación, operadores morfológicos y espacios de color. Cada combinación de técnicas de segmentación y parámetros se codifican y conforman un cromosoma que representa un individuo. Un conjunto inicial de individuos conforma la población inicial (ver Figura 3.3). Cada individuo es evaluado o cuantificada su aptitud mediante alguna técnica para evaluar la calidad de la segmentación.

Mientras mayor aptitud tenga el individuo mayor probabilidad de sobrevivir tendrá y mayor probabilidad de que alguno de sus genes (combinación de técnicas) pasen a la siguiente generación. Este proceso se realiza iterativamente por varias generaciones hasta que se cumple una condición de paro o se obtiene la segmentación óptima.

En el momento que el genético concluye se ha encontrado la combinación de técnicas y parámetros correctos para lograr una buena segmentación en el conjunto de imágenes. En las siguientes secciones se describen en detalle cada uno de los pasos de la metodología propuesta.

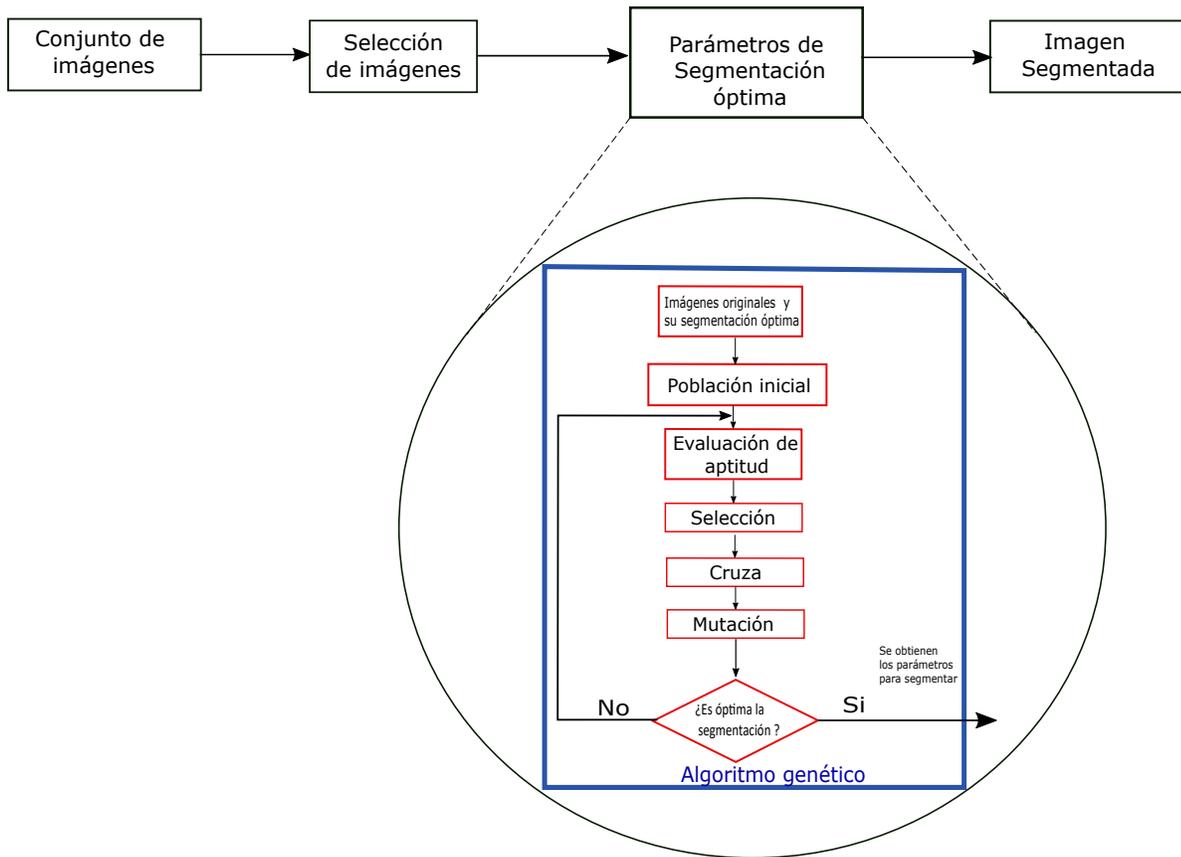


FIGURA 3.3: Metodología

### 3.1. Conjunto de datos

El primer paso es la selección del conjunto de datos. El método propuesto está diseñado para trabajar en conjuntos de datos con fondo constante. Es claro que pueden existir algunas pequeñas variaciones, sobre todo en iluminación y ruido, sin embargo, de forma general los conjuntos de imágenes tienen pocas variaciones. La figura muestra varias muestras de conjuntos de imágenes, cada fila representa un conjunto.

La primera fila pertenece a un conjunto de imágenes de hojas de plantas ICL, que es una colección de hojas de la Universidad de Hefei [28]. El conjunto de datos (ver Figura 3.4) es muy fácil de segmentar pues no tiene cambios significativos en la iluminación y el fondo es totalmente blanco. Solo algunos pequeños cambios en la superficie de la hoja podrían afectar la segmentación de la hoja. Sin embargo, en general parece ser

muy fácil encontrar las técnicas de segmentación y parámetros para una segmentación óptima.



FIGURA 3.4: Conjuntos de datos utilizados en los experimentos

En la segunda fila el conjunto de datos pertenece a una colección de imágenes y videos de palabras en el lenguaje de señas Mexicano [20]. En este caso, aun cuando el fondo es negro y las personas están vestidas de negro, Es posible ver en las imágenes diferentes cambios en la iluminación. Estos cambios de iluminación afectan considerablemente la segmentación de la imagen. Encontrar una combinación de técnicas y parámetros que segmenten perfectamente el conjunto de imágenes en este caso es una tarea más compleja.

Finalmente, la tercera fila pertenece al conjunto de datos Plantvillage [44], este contiene imágenes de plantas con diferentes afectaciones o enfermedades. El conjunto de datos, aunque parece que tiene el mismo fondo este cambia de intensidad y tonalidad. Además, las imágenes presentan varias sombras que afectan la calidad de la segmentación. Aunado a lo anterior, el brillo en las hojas también puede provocar una mala segmentación.

## 3.2. Selección de imágenes

El algoritmo genético debe calcular la idoneidad de una combinación de técnicas y parámetros para obtener la mejor combinación. Para ello, podemos calcular la aptitud evaluando cómo se comporta una combinación de parámetros con una imagen.

Podría dar resultados erróneos, ya que es mejor evaluar la aptitud en varias imágenes en lugar de en una sola imagen. En esta tesis se seleccionan varias imágenes que contienen la variedad de características de todo el conjunto; se segmentan a mano y se etiquetan como segmentación óptima.

Las segmentaciones obtenidas se comparan con las óptimas, y cuanto más similares son a la segmentación óptima, mayor es la aptitud de la combinación de técnicas y parámetros. El subconjunto de imágenes que segmentaremos a mano y usaremos con el algoritmo genético para obtener los parámetros óptimos se selecciona manualmente. Sin embargo, es posible utilizar algoritmos de agrupamiento para obtener las imágenes más representativas de las imágenes de forma automática.

### 3.2.1. Sefexa Image Segmentation Tool

Esta herramienta es la que utilizamos para segmentar el subconjunto de imágenes con características variadas. Nos referimos a características variadas a imágenes con mucha iluminación o imágenes con poca iluminación, imágenes que presentan sombras, etc. Esto nos permite que el algoritmo genético aprenda a segmentar sobre un amplio espectro del conjunto de imágenes.

El uso de varias imágenes permite disminuir considerablemente el error al segmentar sobre el total del conjunto de datos, ya que nos permite obtener los parámetros necesarios sobre imágenes con diversas características. Como se puede apreciar en la Figura 3.5.

La tercera fila muestra imágenes con distintos niveles de iluminación, incluso imágenes con sombras, así como imágenes con imagen objetivo con distinta tonalidad. Estas imágenes con características diferentes entre sí deben seleccionarse para mejorar la calidad de segmentación global.



caso de la segmentación manual es muy sencillo segmentar una imagen, sin embargo, es casi imposible segmentar cada una de las imágenes de forma manual cuando trabajamos en conjuntos de datos muy grandes. En estos casos, es indispensable un sistema de segmentación automático o semi automático, como es el caso propuesto.

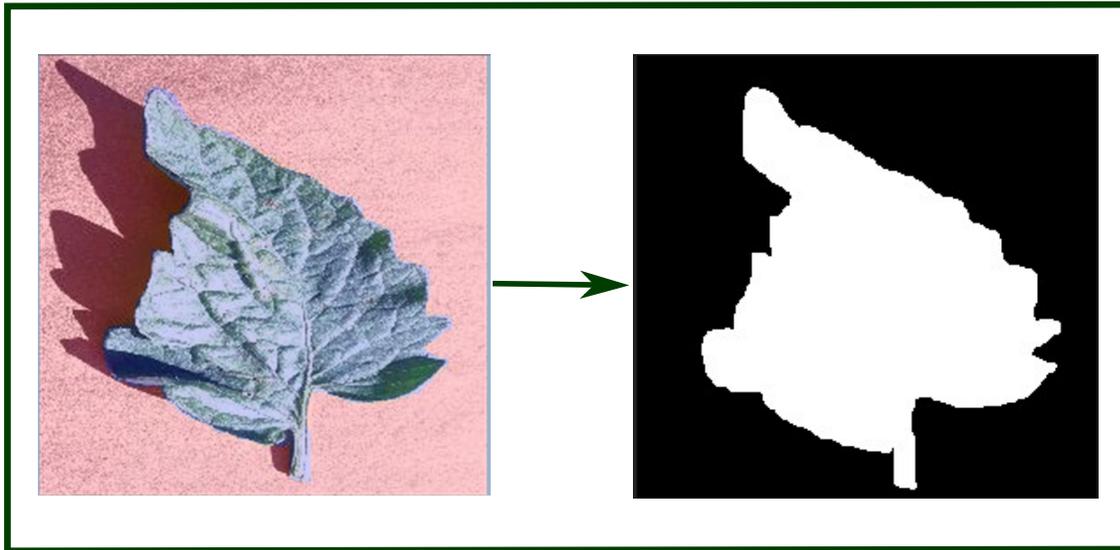


FIGURA 3.7: Segmentación por Sefexa

### 3.3. Selección de técnicas y parámetros

Esta es la parte medular de la metodología propuesta. El problema de segmentación puede ser visto como un problema de optimización, donde  $\Omega = \{S^1, \dots, S^n\}$  es el conjunto de parámetros de segmentación candidatos, esto es  $S = \{p_1, p_2, \dots, p_t\}$ , por lo tanto, la segmentación óptima  $S^* \in \Omega$  deberá satisfacer lo siguiente:

$$\forall S \in \Omega : f(S^*) < f(S) \quad (3.1)$$

donde  $f(\cdot)$  representa la función de aptitud, que mide la relación entre la segmentación obtenida y la segmentación óptima. A mayor semejanza o a menor disimilitud la aptitud es mayor.

Debido a que los algoritmos genéticos nos sirven para optimizar funciones en base a una función de aptitud, en esta tesis se propone un algoritmo genético que obtenga las técnicas adecuadas de segmentación y sus parámetros. Los algoritmos genéticos son una de las heurísticas más utilizadas para resolver problemas reales de optimización.

Este algoritmo está basado en la evolución natural de las especies. Los algoritmos genéticos buscan dentro de un conjunto de soluciones la mejor solución utilizando un conjunto de individuos inicial que representa a un conjunto de soluciones, estas soluciones las recombina y muta en base a un criterio de selección generalmente basado en la aptitud del individuo, produciendo nuevos individuos con mejor aptitud en cada generación. Este proceso se repite hasta alcanzar una solución o cumplir con un criterio de paro. El algoritmo genético involucra los siguientes pasos que son mostrados en la Figura 3.8.

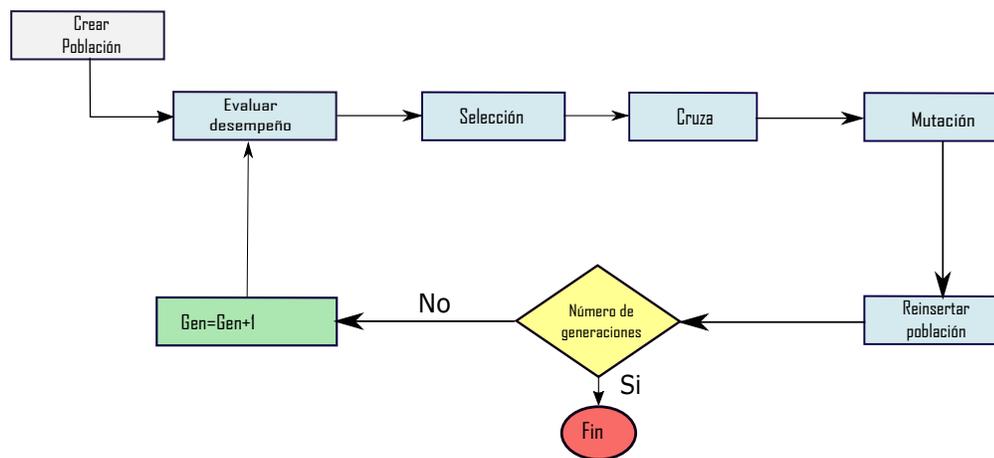


FIGURA 3.8: Diagrama de flujo de un algoritmo genético

### 3.3.1. Población inicial y representación del cromosoma

Así como en la biología evolucionan las especies a través del tiempo y sobreviven los individuos más fuertes o los más aptos. En un conjunto de posibles soluciones a un problema, cada una de esas soluciones funcionan como si fueran un individuo, por lo tanto, tenemos una población de posibles soluciones donde se busca encontrar los individuos con más aptitud a través de ciertas generaciones.

La creación de una población se hace de una manera aleatoria, donde el número de alelos del cromosoma depende del número de variables con que se trabaja. Una población con  $n$  individuos es definida mediante  $\mathbf{S}^g = \{\mathbf{s}_1^g, \mathbf{s}_2^g, \dots, \mathbf{s}_n^g\}$ , donde  $g$  representa la generación actual. Cada individuo es representado por un vector de dimensión  $d$  y es definido como  $\mathbf{s}_i^g = [s_{i,1}^g, s_{i,2}^g, \dots, s_{i,d}^g]$

Las variables en el espacio de búsqueda que constituyen los fenotipos son codificadas en genotipos. El genotipo de un individuo es definido por su cromosoma, mientras que la posible solución del problema de ese cromosoma es el fenotipo. Los operadores genéticos son aplicados al genotipo en cada iteración, esto permite mejorar a los individuos.

En nuestro caso se realizaron dos experimentos. Un primer experimento con solo 9 variables y un segundo caso con 11 variables que se realizó para mejorar aún más los resultados obtenidos con 9 variables, debido a que los resultados no eran muy buenos. El uso de espacios de color con el algoritmo propuesto con 11 variables ayudo a mejorar la calidad de la segmentación

### Caso 1 (9 variables)

En el primero, se creó una población aleatoriamente de 120 individuos con 9 variables que componen a cada uno. Esto es,  $\mathbf{S}^g = \{\mathbf{s}_1^g, \mathbf{s}_2^g, \dots, \mathbf{s}_{120}^g\}$  y  $\mathbf{s}_i^g = [s_{i,1}^g, s_{i,2}^g, \dots, s_{i,9}^g]$

Los 120 individuos son aleatoriamente inicializados en el espacio de búsqueda  $[s_k^L, s_k^U]$  donde  $k$  representa la  $k$ -ésima variable del cromosoma, L y U representan las cotas inferior y superior respectivamente para cada variable. Para cada variable se describen las cotas utilizadas. La Imagen (3.9) se puede observar nuestras variables representadas en una cadena cromosómica, donde cada alelo representa cada variable de nuestro algoritmo.



FIGURA 3.9: Cadena cromosómica de nueve variables

Donde:

- Filtro promedio: Es una variable discreta que puede tomar el valor de 1 si se utiliza un filtro promedio o 0 si no se utiliza este filtro promedio.
- Tamaño: Es una variable discreta y se refiere al tamaño del filtro promedio, puede tomar los valores de 3, 5, 7 y 9.
- Filtro Gaussiano: Es una variable discreta que indica que se está utilizando un filtro gaussiano con un numero 1, y un 0 si no se está aplicando un filtro gaussiano.

- Desviación Estándar: La desviación estándar del filtro gaussiano es una variable continua que puede tomar valores entre 0 a 1.
- Segmentación: Esta variable representa el tipo de segmentación que se va a emplear, puede tomar 4 valores que son los siguientes.
  - Clustering
  - PCA
  - Frontera adaptativa
  - Método de Otsu
- Dilatación: La dilatación puede tomar valores discretos de 1 a 4 donde representan los tamaños del círculo a dilatar. Mientras más grande sea el círculo, la dilatación es mayor.
- Erosión: La erosión, así como la dilatación, cuenta con 4 valores de tamaño para esta variable discreta. Mientras más grande sea el círculo de erosión mayor será la erosión.

Las últimas dos variables se repiten para mejorar la calidad de la segmentación. El tipo de dilatación y erosión utilizadas es diferente y la combinación de estas genera una segmentación final diferente dependiendo de la combinación utilizada.

### Caso 2 (11 variables)

Después de observar algunas pruebas con el algoritmo de nueve variables, se decidió agregar dos variables más para mejorar la calidad de la segmentación. Para el segundo caso, se creó una población aleatoriamente de 120 individuos con 11 variables que componen a cada uno. Esto es,  $\mathbf{S}^g = \{\mathbf{s}_1^g, \mathbf{s}_2^g, \dots, \mathbf{s}_{120}^g\}$  y  $\mathbf{s}_i^g = [s_{i,1}^g, s_{i,2}^g, \dots, s_{i,11}^g]$

Los 120 individuos son aleatoriamente inicializados en el espacio de búsqueda  $[s_k^L, s_k^U]$  donde  $k$  representa la  $k$ -ésima variable del cromosoma,  $L$  y  $U$  representan las cotas inferior y superior respectivamente para cada variable. Para cada variable se describen las cotas utilizadas.

En la Imagen (3.10) se observa el cromosoma completo.

Donde:

- Espacio de color: Comprende el espacio de color que puede tomar 5 variables.



FIGURA 3.10: Cadena cromosómica de once variables

- RGB
  - NTSC
  - HSV
  - XYZ
  - YCBCR
- Canal: El canal tiene 4 configuraciones que se describen a continuación.
    1. Todos los canales
    2. Primer canal
    3. Segundo canal
    4. Tercer canal

A partir de la población creada, se procede a evaluar cada uno de los individuos para saber su aptitud o, dicho de otra manera, la calidad de la segmentación obtenida a partir de las variables utilizadas.

Normalmente cada variable es decodificada introduciéndolo en la Fórmula (3.2) para obtener su valor. En nuestro caso, como se puede apreciar en las Figuras 3.9 y 3.10 casi todas las variables son discretas, solo una variable es continua, esta es la desviación estándar utilizada para el filtro Gaussiano.

El tamaño de la sección del cromosoma depende del espacio de búsqueda necesario. En nuestro caso, para el cálculo de la desviación estándar no es necesaria una búsqueda minuciosa, así que se utilizaron saltos de 0.1 y dado que la desviación estándar la acotamos con valores entre 0 y 1, solo se necesitaron 4 alelos para codificarla.

$$x_i = x_i^l + \left( \frac{x_i^u - x_i^l}{2^\beta - 1} \right) (\gamma) \quad (3.2)$$

Donde:

$\beta$  = Número de alelos del cromosoma

$x_i^l$  = Límite inferior

$x_i^u$  = Límite superior

$\gamma$  = El número binario del cromosoma en decimal

La longitud de cada subcadena binaria depende del tamaño del espacio de búsqueda y el número de decimales requeridos para la variable decodificada. Si cada variable de decisión está dada por una cadena de longitud  $L$  y existen  $n$  variables, entonces la longitud final del cromosoma será de tamaño  $L \times n$ .

### 3.3.2. Evaluación de función de aptitud

Una vez decodificado el cromosoma de cada individuo, este tendrá cada una de las variables necesarias para segmentar una imagen. Las variables definen el tipo de pre-procesamiento, la técnica de segmentación, operación morfológica y espacio de color de la segmentación. Cada combinación de estas variables nos da como resultado una imagen segmentada con una calidad de segmentación particular.

La evaluación o calidad de esta segmentación con las variables que lo componen se obtiene realizando una comparación con la segmentación manual o segmentación óptima (*Ground truth*). *Ground truth* es un término comúnmente utilizado para referirse a la información proporcionada por la observación directa en contraposición a la información proporcionada por inferencia.

Para realizar esta comparación entre la información directa y la información obtenida por inferencia se utilizan varias métricas de desempeño VOI, PRI y GCE que se describen a continuación.

#### Probabilistic Rand Index (PRI)

El índice probabilístico define la exactitud de una segmentación bajo un punto de vista estadístico. Este índice fue introducido en [56] para medir la similitud de métodos de agrupamiento y modificado por [63] para evaluar similitud en segmentación.

Una vez segmentada una imagen, esta imagen segmentada (S) tiene etiquetas de píxeles de clase 0 o de clase 1. La imagen segmentada manualmente (G) también tiene etiquetas de clase 0 y 1, sin embargo, sus etiquetas fueron obtenidas manualmente de acuerdo con un conocimiento previo.

El índice de probabilidad estadístico (PRI, por sus siglas en inglés) contabiliza el número de pares de etiquetas de píxeles que son iguales en la imagen segmentada S y la imagen segmentada manualmente G. PRI obtiene el promedio total y lo acota

en el rango  $[0,1]$ . Un valor de 0 significa que no existen consistencia entre la imagen segmentada manualmente y la obtenida mediante un algoritmo. Mientras que un valor de 1 significa que la segmentación obtenida con el algoritmo es idéntica a la obtenida manualmente.

La segmentación de una imagen puede ser descrita en la forma de números binarios  $I(l_i^{s_k} = l_j^{s_k})$  sobre cada par de píxeles  $(x_i, x_j)$ . La distribución de estos números sigue una distribución de Bernoulli y nos da una variable aleatoria con un valor esperado denotado por  $p_{ij}$ . El índice probabilístico de dos segmentaciones puede ser definido como:

$$PR(S_{test}, S_k) = \frac{1}{N} \sum [I(l_i^{S_{test}} = l_i^{S_k})p_{ij} + I(l_i^{S_{test}} \neq l_i^{S_k})(1 - p_{ij})] \quad (3.3)$$

donde  $N$  es el número de píxeles,  $S_k$  es el conjunto de segmentaciones reales,  $p_{ij}$  es la probabilidad de que las etiquetas en la segmentación realizada sean iguales a la segmentación real. En la práctica, la media es utilizada para calcular  $p_{ij}$ . El índice probabilístico toma valores en el rango  $[0,1]$ , donde una puntuación de cero indica que el etiquetado de la imagen de prueba es totalmente opuesto a la segmentación real, y una puntuación de 1 indica que la imagen segmentada real y la imagen segmentada de prueba son iguales para cada par de píxeles.

Una posible desventaja del índice PRI radica en que esta tiene un rango dinámico muy pequeño y debido a que comúnmente las imágenes son muy grandes, esto provoca que los resultados de PRI sean muy similares.

### Global Consistency Error (GCE)

GCE permite calcular el grado de traslape de regiones en la imagen, este fue propuesto en [41] y [42]. GCE fue ideada para cuantificar la calidad de la segmentación en diferentes niveles de detalle. Sea  $R(S, p_i)$  el conjunto de píxeles en la segmentación  $S$  que contienen al píxel  $p_i$ , el error de refinamiento local se define como:

$$E(S_1, S_2, p_i) = \frac{|R(S_1, p_i)/R(S_2, p_i)|}{|R(S_1, p_i)|} \quad (3.4)$$

Este error no es simétrico con respecto a las segmentaciones comparadas y toma el valor de cero cuando  $S_1$  es un refinamiento de  $S_2$  en el píxel  $p_i$ . GCE es definido como

$$GCE(S_1, S_2) = \frac{1}{n} \min \left\{ \sum_i E(S_1, S_2, p_i), \sum_i x_i E(S_2, S_1, p_i) \right\} \quad (3.5)$$

### Variation of Information (VOI)

La variación de la Información (VoI) [43] define la distancia entre dos segmentaciones como la entropía condicional promedio de una segmentación dada la otra, y por lo tanto mide aproximadamente la cantidad de aleatoriedad en una segmentación que no puede ser explicado por el otro.

VOI mide la suma de la pérdida y ganancia de información entre dos imágenes segmentadas (una real y otra de prueba), es decir mide el grado en que una clase puede explicar a la otra. La métrica VOI es no negativa, con valores más bajos indican una mayor similitud. Es basado en la relación entre un punto y su clase.

La variación de la información es una medida de la distancia entre dos particiones (de elementos) de clase. Una clase con píxeles  $X_1, X_2, \dots, X_k$  está representado por una variable aleatoria  $X$  con  $X = \{1, \dots, K\}$ . de tal manera que  $p_i = |X_i|/n_i \in X$  y  $n = \sum_i X_i$  la variación de información entre dos clases  $X$  e  $Y$  por lo representado se define como

$$VI(X, Y) = H(X) + H(Y) - 2I(X, Y) \quad (3.6)$$

donde  $H(X)$  es la entropía de  $X$  y  $I(X, Y)$  es la información similar entre  $X$  e  $Y$ .  $VI(x, y)$  mide la asignación de píxeles para una clase de artículos  $X$  reduce la incertidumbre acerca de píxeles del tema en clase  $Y$ .

En resumen, lo que harán las métricas de desempeño será sacar las coordenadas en nuestras imágenes segmentadas (binarias), asignando el valor de 1 a los píxeles de color blanco y un 0 a los píxeles de color negro, dando por acierto a los que coincidan y una penalización a los que no. La función de aptitud será esta comparación, a menor diferencia mayor aptitud y mayor porcentaje de supervivencia de ese individuo.

### 3.3.3. Selección

A partir de la población inicial y calculadas sus respectivas aptitudes, el siguiente paso es la selección. En la selección, la probabilidad de cada individuo a ser escogido y que su simiente pase a la siguiente generación es directamente proporcional a su aptitud.

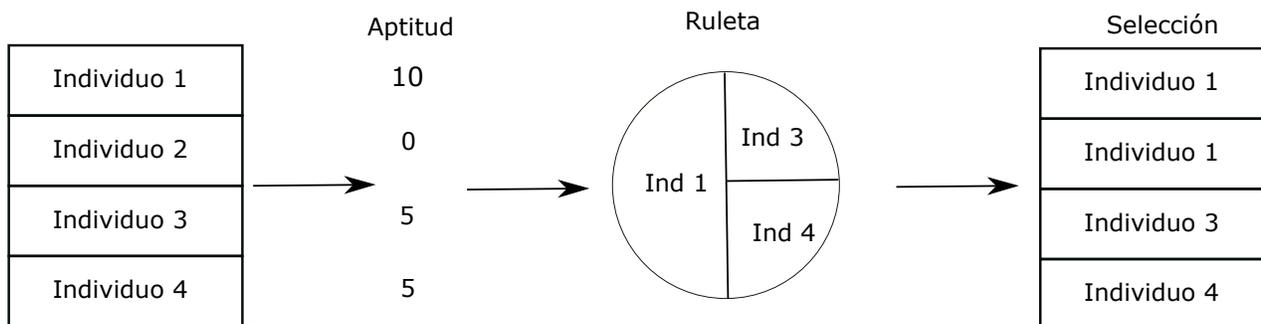


FIGURA 3.11: Representación del porcentaje de aptitud

La técnica de selección que se utilizó es selección por el método de ruleta, Esta técnica, asigna a cada individuo un segmento de la una rueda. La magnitud del segmento de la rueda de cada individuo depende de la aptitud obtenida. Mientras mayor sea la aptitud obtenida, mayor es el segmento de rueda que obtiene. Para seleccionar un individuo se hace girar la rueda de la ruleta y en el punto donde se detiene la rueda indicará al individuo que se seleccionará para que sobreviva en la próxima generación. La Figura 3.11 muestra el proceso de selección por el método de la ruleta. El algoritmo de selección por el método de la ruleta es descrito en el algoritmo 2

---

**Entrada:** Aptitud de cada cromosoma  $f_i$ .

**Salida:** Cromosomas seleccionados

---

- 1: *Calcular la suma total de aptitudes de los  $n$  individuos  $\sum_{i=1}^n f_i$*
  - 2: *Generar un número aleatorio  $r \in [0, 1]$*
  - 3: **Obtener el valor de  $s = r \times \sum_{i=1}^n f_i$**
  - 4: *Si  $s \leq \sum_{i=1}^k f_i$  entonces se selecciona el  $k$ th cromosoma para la siguiente generación*
  - 5: *Repetir pasos 2 a 4 hasta que el número de cromosomas seleccionados sea igual a  $n$ .*
- 

**Algoritmo 2:** Algoritmo de Selección ruleta

### 3.3.4. Cruza

Biológicamente el cruce es cuando dos cromosomas se enrollan y de ahí comienza el proceso de mitosis donde la célula que tiene a los dos cromosomas se divide obteniendo información de ambos cromosomas. Este mecanismo de herencia es muy importante en algoritmos genéticos y permite definir la forma como pasará la información genética de padres a hijos.

La cruce permite fusionar la información genética de dos individuos y pasarla a los hijos. Aunque en la literatura existen diferentes métodos de cruce, en nuestros experimentos se realizó cruce de dos puntos. La Figura 3.12 muestra gráficamente el procedimiento para la cruce de dos puntos.

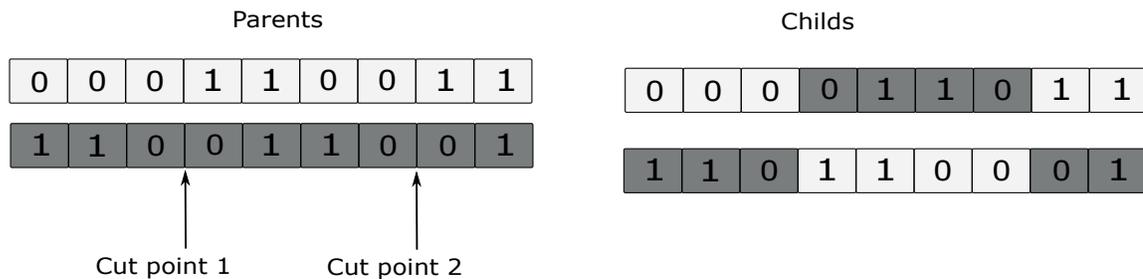


FIGURA 3.12: Cruce de dos puntos

### 3.3.5. Mutación

La mutación permite modificar el cromosoma de forma aleatoria, esto se replica de la evolución de las especies cuando existe un error en la reproducción, conocido como error de copiado o se genera una deformación de los genes. En los algoritmos genéticos la probabilidad de mutación suele ser mucho más grande que en la de cualquier organismo biológico.

La mutación permite que eventualmente un cromosoma de un salto a lo desconocido al modificarse aleatoriamente su cromosoma. Este operador es muy importante para el algoritmo genético ya que impide que el algoritmo quede atrapado en mínimos locales.

La mutación se aplica a un solo gen, generalmente es implementada con una probabilidad baja, como 0.01 o 0.001. La mutación simplemente implica invertir el valor de un bit en un cromosoma. Por ejemplo, con una tasa de mutación de 0.01, podría esperarse que un gen en un cromosoma de 100 genes pudiera invertirse.

### 3.3.6. Condición de paro

La condición de paro es muy importante para un algoritmo genético. Existen dos condiciones de paro regularmente utilizadas 1) cuando el algoritmo genético ha llegado al número de generaciones estipuladas de antemano y 2) cuando se detecta que la

mayor parte de la población ha convergido, provocando que sus cromosomas carezcan de suficiente diversidad.

La primera puede provocar que el paro se realice prematuramente y la solución obtenida no sea adecuada, mientras que la segunda solo detiene el proceso cuando los cromosomas son muy similares entre sí y seguir con la evolución carece de sentido.

En el método propuesto, el algoritmo finaliza después de 10 iteraciones en las que no existe mejora en la evolución de los individuos. La cadena con mayor precisión se almacena con la precisión obtenida.

### 3.3.7. Elitismo

El elitismo generalmente se utiliza en AG para no perder o eliminar al individuo con la mejor aptitud entre las generaciones. Esta técnica selecciona al individuo más apto o los individuos más aptos y los pasa intactos a la siguiente generación. En cada generación, se obtiene al individuo con la mejor aptitud y se compara con la generación anterior. Si no se obtiene un cromosoma que mejore la mejor aptitud hasta el momento, la cadena anterior aún permanece intacta en la nueva generación.

En el método propuesto no se utiliza elitismo puro, utilizamos una forma modificada en la que un solo cromosoma (el de mejor aptitud) de cada generación se copian en la siguiente generación.

# Capítulo 4

## Resultados experimentales

En este Capítulo mostramos los resultados obtenidos con el algoritmo y metodología propuesta. En cada Sección, describimos los conjuntos de datos y los parámetros utilizados para llegar a los resultados obtenidos. Además, se muestran las comparaciones obtenidas con otros métodos de segmentación en el estado del arte.

### 4.1. Conjuntos de datos

#### 4.1.1. Conjunto de datos ICL

El conjunto de datos de hojas ICL es un conjunto de imágenes de hojas de plantas que ha sido muy utilizado los últimos años en muchos trabajos de investigación [10] [11] [67] [72], es llamado así pues fue recolectado por miembros del Laboratorio de Cómputo Inteligente en el Instituto de Máquinas Inteligentes de la academia China de Ciencias. Probablemente es el conjunto de datos más grande disponible para investigación.

El conjunto de datos contiene 16851 imágenes de 220 especies. Cada especie tiene entre 26 y 1078 ejemplos. En todas las imágenes los recolectores cuidaron que el fondo sea blanco lo que permite contrastar los colores de la hoja y ayuda a mejorar la calidad de la segmentación. Además, también cuidaron que la iluminación fuera adecuada, esto permite reducir o eliminar partes con demasiada iluminación, mejorando la calidad de la segmentación. La Fig. 4.1 muestra diferentes imágenes de plantas seleccionadas al azar.



FIGURA 4.1: Conjunto de datos ICL

#### 4.1.2. Conjunto de datos enfermedades del café

El conjunto de datos de hojas Roya del café fue obtenido del sitio web [data.mendeley.com](http://data.mendeley.com), y recolectado por los autores de [36]. Contiene 739 imágenes de hojas sanas y cuatro enfermedades comunes de la especie *coffea arabica* que fueron tomadas en cultivos de café en Santa María de Marechal, Floreano, Brasil. Todas las fotos se tomaron en un ambiente controlado ya que fueron puestas en un fondo blanco. Las imágenes originales del conjunto tienen una resolución de 2048 píxeles de ancho por 1024 de píxeles de alto.

Aunque todas las imágenes tienen fondo blanco, los recolectores no cuidaron que la iluminación sea adecuada, esto provoca problemas al segmentar. La Fig. 4.2 muestra diferentes imágenes seleccionadas al azar.



FIGURA 4.2: Conjunto de datos enfermedades de café

### 4.1.3. Conjunto de datos PlantVillage

El conjunto de datos de PlantVillage fue originalmente utilizado en [29] consta de 54303 imágenes de hojas sanas y no saludables divididas en 38 categorías por especies y enfermedades.

El conjunto de datos PlantVillage contiene 38 clases y 54.305 imágenes de 14 especies de plantas diferentes en total, 12 de las cuales están sanas, 26 de las cuales están enfermas. Las imágenes del conjunto de datos son imágenes en color de distintos tamaños. En este conjunto de datos el fondo en las imágenes varía muy poco, sin embargo, los recolectores no tuvieron cuidado en la iluminación lo que aumenta la dificultad para segmentar adecuadamente las imágenes. La Fig. 4.3 muestra diferentes plantas y enfermedades seleccionados al azar.



FIGURA 4.3: Conjunto de datos PlantVillage

#### 4.1.4. Conjunto de datos LSM

El conjunto de datos del lenguaje de señas Mexicano (LSM) es un conjunto de datos obtenido por los autores para identificar y traducir automáticamente el LSM [20]. El conjunto de datos contiene videos e imágenes de 249 palabras del LSM los autores utilizan fondo y ropa específicos para mejorar la calidad de segmentación de manos y cara. Esto permite a los autores mejorar el desempeño de la traducción.

Los autores recopilaron 2480 videos de palabras de signos. Las señales realizadas en estos videos provienen de once personas diferentes en promedio para cada una de las 249 palabras recopiladas. Cada video contiene un solo signo de la lengua mexicana, en estos videos el individuo está frente a la cámara, con la mano izquierda y derecha en reposo, luego el individuo comienza el movimiento del gesto de la mano, luego de eso, el individuo termina con las manos en reposo.

De los videos, el número de imágenes obtenidas es de 37200 y en promedio 15 imágenes representan una palabra. Es decir, el conjunto de posiciones de manos y cara y formas descritas por las manos nos describen una palabra del LSM.

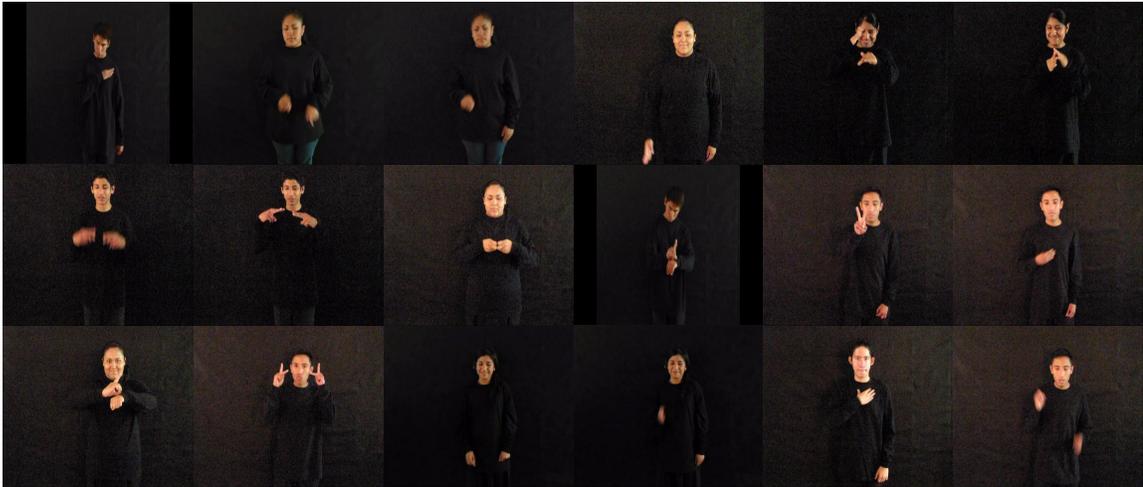


FIGURA 4.4: Conjunto de datos LSM

## 4.2. Segmentación de los conjuntos de datos

En esta sección se muestran los resultados de segmentación obtenidas con las técnicas clásicas en los diferentes conjuntos de datos utilizados.

### 4.2.1. Conjunto de datos ICL

El conjunto de datos ICL es uno de los conjuntos de imágenes de plantas mejor elaborado. La mayoría de las imágenes tiene fondo blanco y algunas pocas tienen fondo azul cielo, sin embargo, en las pruebas realizadas se puede apreciar que esto no afecta la calidad de la segmentación. La Figura 4.5 muestra los resultados obtenidos con Otsu y PCA.

En los experimentos realizados se excluyeron los dos métodos propuestos en esta tesis, pues las imágenes no necesitan de algoritmos elaborados para segmentar las imágenes. El uso de algoritmos de segmentación básicos como Otsu y PCA permiten segmentar adecuadamente las imágenes con una alta calidad. En la Figura 4.5 la primera fila muestra algunas imágenes seleccionadas al azar, la segunda fila muestra la segmentación obtenida con la técnica de segmentación de Otsu y la tercera fila muestra los resultados obtenidos al utilizar el método PCA al segmentar.

Como se puede apreciar, en este caso debido a que los recolectores de las imágenes se preocuparon por tener un fondo blanco o azul y una iluminación adecuada, el uso de algoritmos básicos permite obtener una buena calidad de segmentación y obtener fácilmente la región que necesitamos.



FIGURA 4.5: Segmentación del conjunto de datos ICL

#### 4.2.2. Conjunto de datos enfermedades del café

El segundo conjunto de datos empleado es el conjunto de datos de enfermedades del café [36]. Aunque todas las imágenes del conjunto de datos también tienen fondo blanco, en este caso los autores no tuvieron cuidado con la iluminación y las imágenes presentan algunas sombras. Estas sombras presentan algunos problemas para algunas técnicas de segmentación básicas. La Figura 4.6 muestra en la primera fila 4 imágenes al azar del conjunto de datos. Las filas restantes muestran los resultados de segmentación obtenidos con las diferentes técnicas utilizadas.

La segunda fila muestra los resultados obtenidos con el método de Otsu. El método de Otsu es uno de los métodos de segmentación más utilizado debido a su desempeño. Sin embargo, este tiene problemas para segmentar algunas imágenes con mala iluminación, sobre todo en las sombras de la imagen o incluso confunde algunas áreas dentro de la hoja que se encuentran con una tonalidad más clara del resto de la hoja.

La tercera fila muestra la segmentación obtenida por el método PCA. Con el método PCA al igual que Otsu, se sigue teniendo problemas al segmentar debido a la mala iluminación. La cuarta y quinta columna se muestran los resultados de segmentación obtenidos con los métodos propuestos. En estos se puede observar una segmentación de mejor calidad.

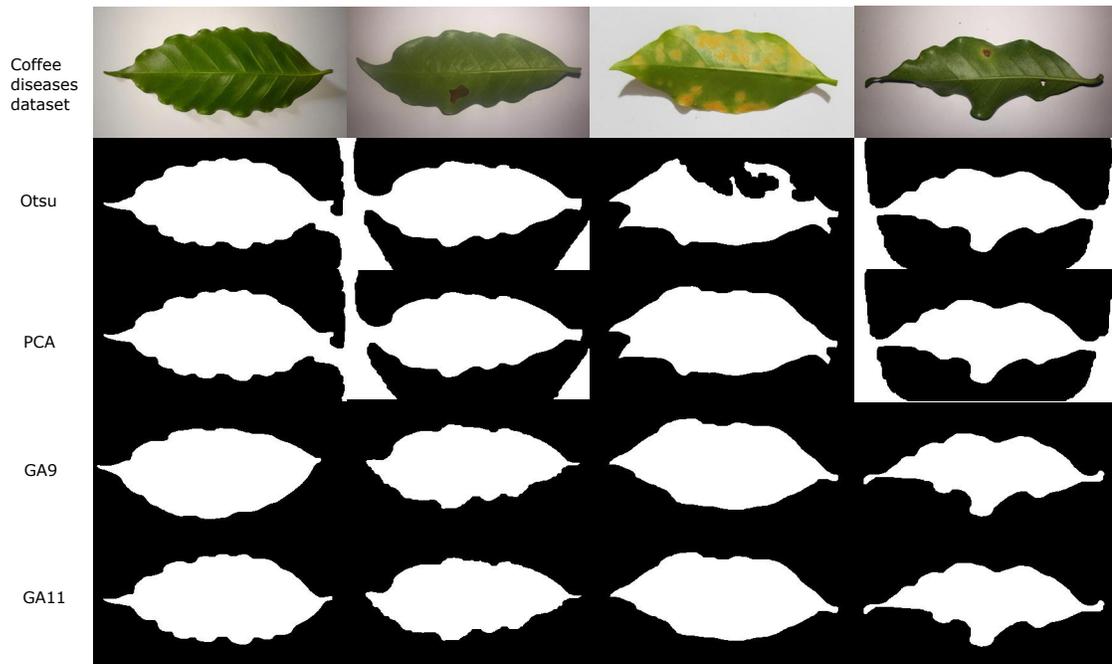


FIGURA 4.6: Segmentación del conjunto de datos enfermedades del Café

### 4.2.3. Conjunto de datos PlantVillage

Este conjunto de datos tiene más problemas al segmentar debido a que los recolectores no tuvieron cuidado en la iluminación y en algunos casos el fondo de la hoja. La Figura 4.7 muestra la primera fila con las imágenes seleccionadas al azar. Las filas 2,3,4 y 5 muestran las imágenes de la fila 1 segmentadas utilizando los métodos de Otsu, PCA, GA9 y GA11 respectivamente.

Es claro ver en la segunda y tercera fila que los algoritmos de Otsu y PCA no presentan buenos resultados cuando se presentan problemas de iluminación en la imagen. Esto afecta demasiado al proceso debido a que por lo general la segmentación es un paso en el reconocimiento de patrones, obtener una imagen mal segmentada provocará que las características extraídas sean muy malas y por consiguiente un sistema de reconocimiento deficiente.

Además de ello, es claro ver en las imágenes que incluso algunas sombras las segmenta como si fueran parte del objeto a segmentar (Fila 2, Columna 5 y Fila 2, Columna 2). Por otro lado, los algoritmos propuestos GA9 y GA11 muestran resultados muy buenos al segmentar las imágenes.

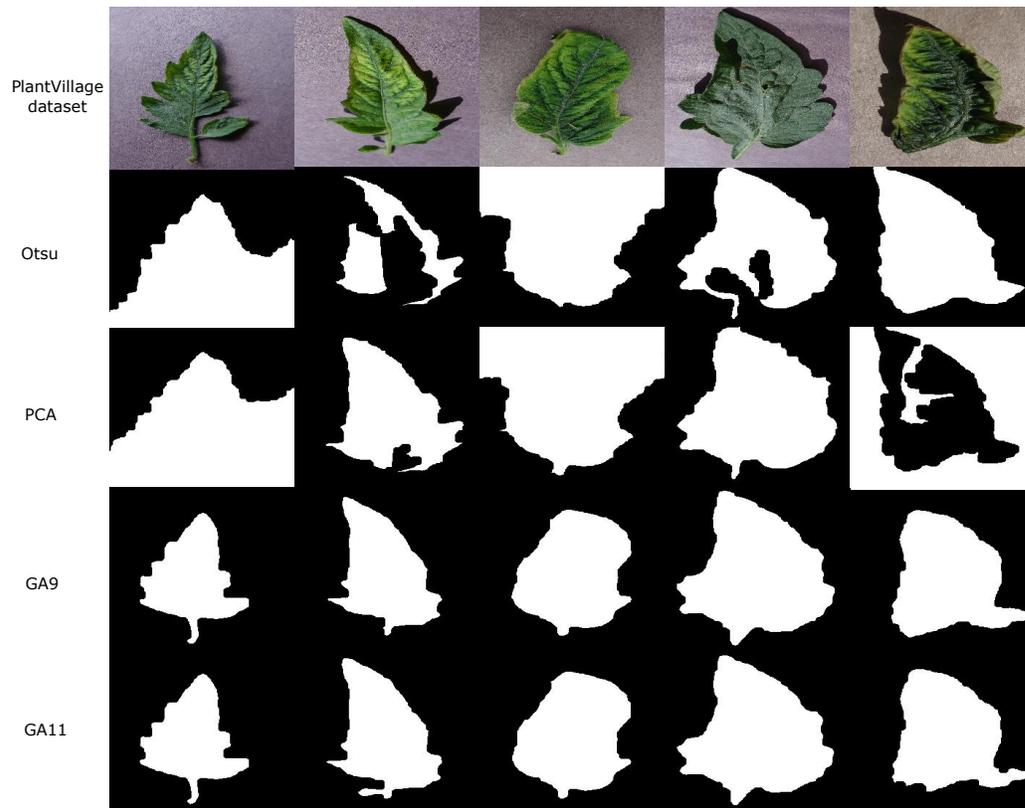


FIGURA 4.7: Segmentación del conjunto de datos PlantVillage

#### 4.2.4. Conjunto de datos LSM

Finalmente, el conjunto de datos LSM contiene imágenes del lenguaje de señas mexicano. Todas las imágenes fueron tomadas con fondo y vestimenta negros. Parece que el conjunto de imágenes es muy fácil de segmentar, sin embargo, algunas imágenes presentan problemas de iluminación.

Esto se puede apreciar en las dos primeras imágenes de la fila 1. Los métodos de Otsu y PCA tienen problemas para segmentar la imagen, aun el algoritmo GA (presenta problemas pues solo segmenta la parte de la cara y no las manos). Esto es de vital importancia ya que la dirección de una o ambas manos, así como la posición de alguna de las manos o contacto de alguna de las manos con el cuerpo son patrones que se buscan para identificar la seña realizada.

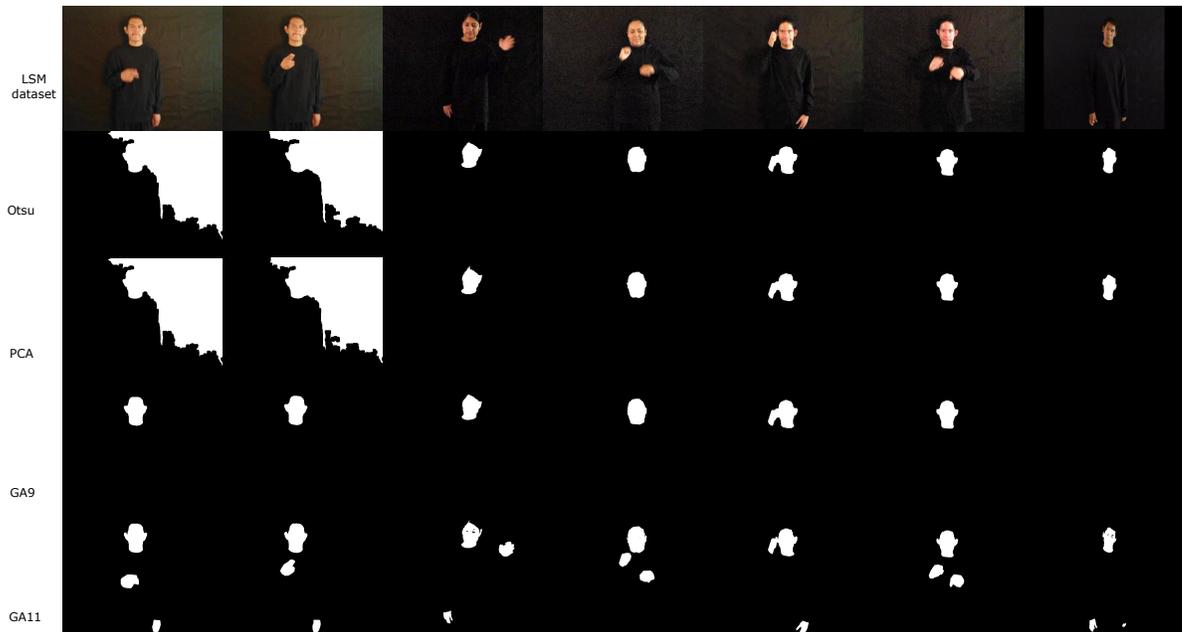


FIGURA 4.8: Segmentación del conjunto de datos LSM

La Figura 4.8 muestra que las segmentaciones obtenidas con Otsu y PCA no logran obtener las regiones de las manos que son muy importantes para el reconocimiento de las señas. En los experimentos realizados se puede apreciar que aun el algoritmo propuesto GA9 tiene problemas para encontrar el conjunto de variables adecuadas para segmentar la imagen. Por otro lado, el algoritmo AG11 obtiene una mejor calidad de segmentación en todas las imágenes.

### 4.3. Resultados y análisis comparativo

En esta Sección mostramos los resultados obtenidos utilizando el algoritmo genético y las distintas métricas de comparación.

#### 4.3.1. Convergencia del algoritmo genético

En los experimentos llevados a cabo se utilizaron poblaciones iniciales con 120 individuos. En la mayoría de los experimentos el algoritmo genético convergió a un mínimo antes de las 80 generaciones. La gráfica 4.9 muestra la distancia promedio entre individuos por cada generación. También muestra como durante las primeras generaciones la distancia entre individuos es muy grande y conforme van pasando las generaciones se va acotando y reduciéndose esta distancia entre individuos.

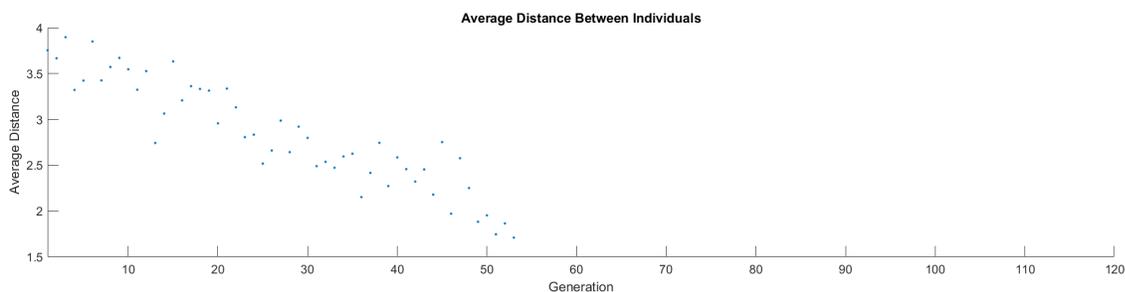


FIGURA 4.9: Distancia promedio entre individuos

En todos los experimentos se utilizó minimización. Debido a que Probabilistic Rand Index (PRI) esta acotado entre 0 y 1 con valores de 1 como los deseados. La aptitud se escaló utilizando  $Fitness = 10 - pri$ , esa es la razón por la que la aptitud del mejor individuo es de 9,02846.

En la gráfica los puntos en azul representan la aptitud media de los individuos. Se debe mencionar que cuando un individuo tiene una aptitud muy mala, se implementó en el algoritmo una penalización muy fuerte para evitar que el genético seleccionara esos individuos. Los puntos en negro en la gráfica representan el mejor individuo en la generación. En la gráfica se puede apreciar cómo se van acortando las distancias y disminuyendo la penalización en cada generación.

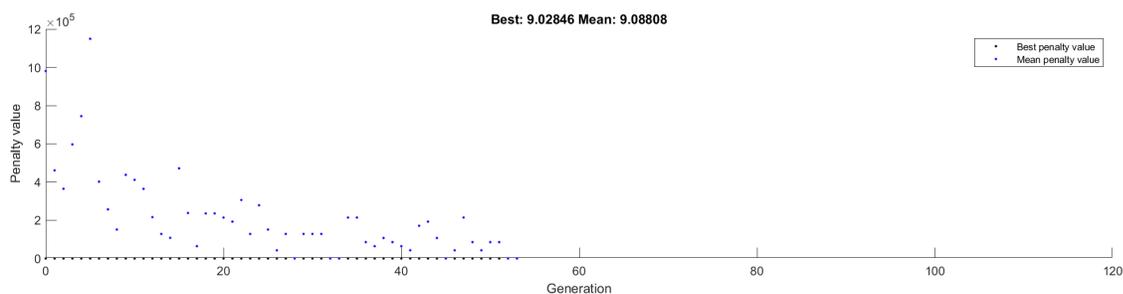


FIGURA 4.10: Mejor individuo obtenido y aptitud promedio de individuos

## 4.4. Discusiones

En esta Sección se discuten los resultados obtenidos en los diferentes conjuntos de datos y utilizando las métricas Probabilistic Rand Index (PRI), Global Consistency Error (GCE) y The Variance of Information (VoI). En la tesis se utilizaron 4 conjuntos de datos, sin embargo, en esta Sección solo se discuten los resultados obtenidos en 3 conjuntos de datos (Coffee diseases, PlantVillage y LSM datasets), debido a que el primer conjunto

de datos (ICL) es muy fácil de segmentar y los resultados eran muy similares entre sí, se excluyeron de esta discusión.

#### 4.4.1. Resultados obtenidos con el conjunto de datos Enfermedades del café

Para el conjunto de datos Coffee diseases, como se pudo apreciar en la Sección anterior los algoritmos Otsu y PCA segmentan muy bien en algunos casos, sin embargo, cuando existen problemas de iluminación, las partes mal iluminadas las segmenta como si fueran parte de la región deseada. Este patrón se repite en muchas de las imágenes. Para estas comparaciones se utilizaron 16 imágenes al azar. Estas fueron segmentadas manualmente y comparadas con los resultados obtenidos con las técnicas Otsu, PCA y los algoritmos propuestos GA9 y GA11.

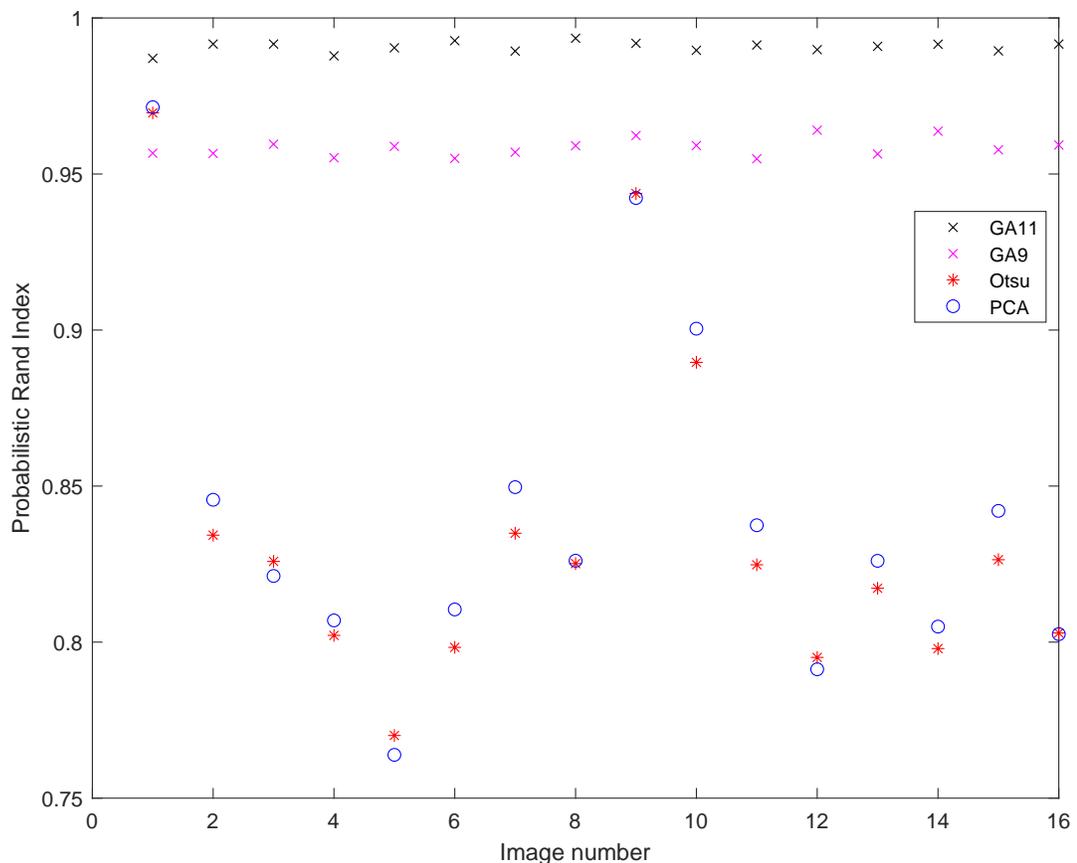


FIGURA 4.11: The Probabilistic Rand Index (PRI) para cada imagen en el conjunto de imágenes de café

Las métricas PRI, GCE y VoI calculan la similitud entre dos imágenes segmentadas. En el caso de la métrica PRI su rango está entre  $[0, 1]$ , mientras más cercano a 1 es más parecida la segmentación obtenida a la segmentación manual.

La Figura 4.11 muestra los resultados obtenidos utilizando la métrica de comparación de PRI con los diferentes métodos de segmentación. En la Figura se muestra el método propuesto GA9 ('x' magenta), GA11 ('x' negra) y el algoritmo de segmentación de Otsu ('\*' en rojo) y la segmentación basada en PCA (círculos azules) y se compararon todas las imágenes con el conjunto de datos *ground truth*. El promedio de la métrica PRI en todas las imágenes fue GA11: 0,9907, GA9: 0,9585, PCA: 0,8401, Otsu: 0,8349.

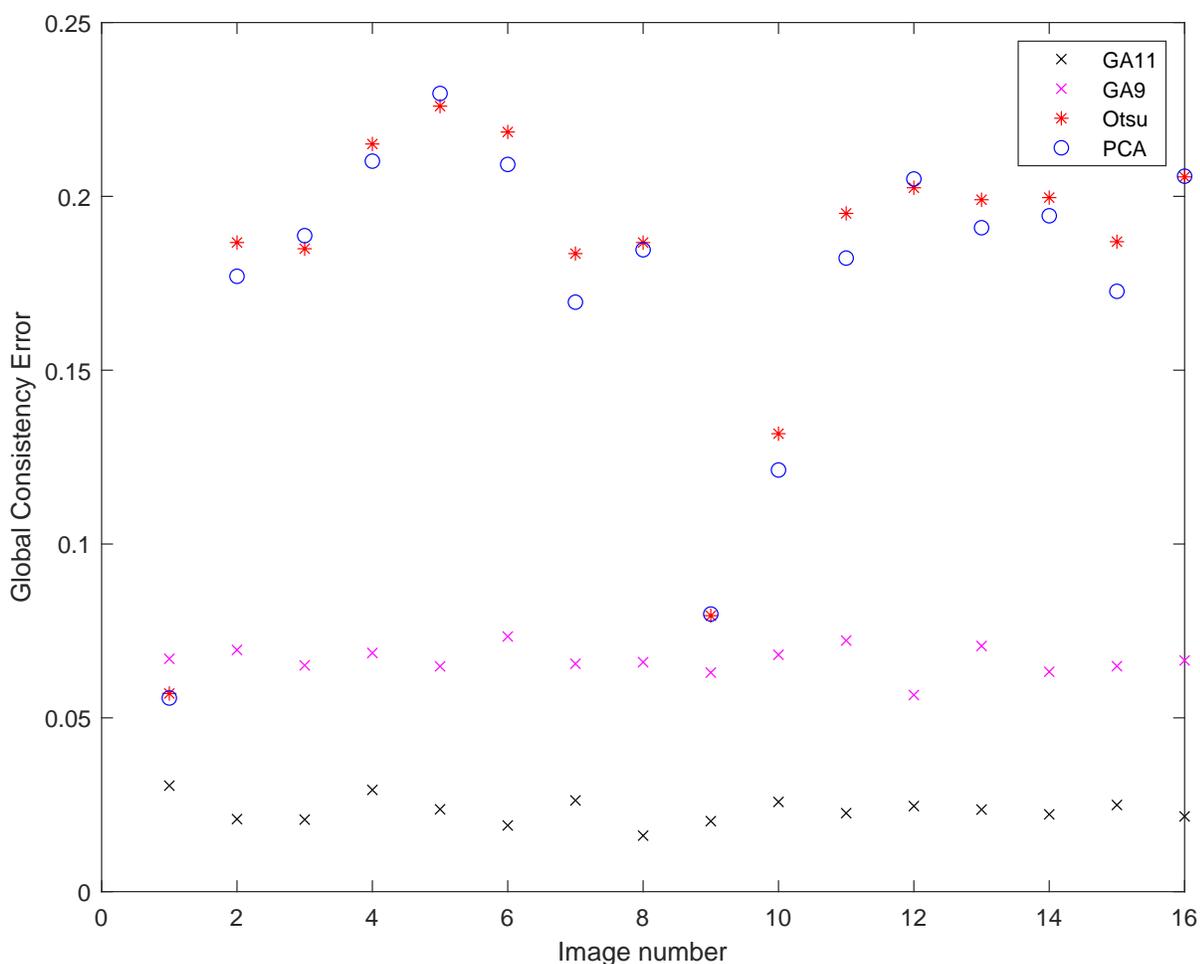


FIGURA 4.12: El Global Consistency Error (GCE) para cada imagen en el conjunto de imágenes de café

En el caso de GCE, el rango de valores de la métrica de comparación está entre  $[0, 1]$  mientras más cercano a 0 indica que la segmentación obtenida es más similar a la

segmentación manual o *Ground truth*.

La Figura 4.12 muestra el GCE para cada imagen con los diferentes métodos de segmentación. Se midieron el GCE para los métodos propuestos GA9 ('x' magenta) y GA11 ('x' negro) y el algoritmo de segmentación de Otsu ('\*' rojo) y la segmentación basada en PCA (círculos azules) y se compararon todas las imágenes con el conjunto de datos *Ground truth*. El GCE promedio en todas las imágenes fue GA11: 0.0233, GA9: 0.0667, PCA: 0.1736, Otsu: 0.1787.

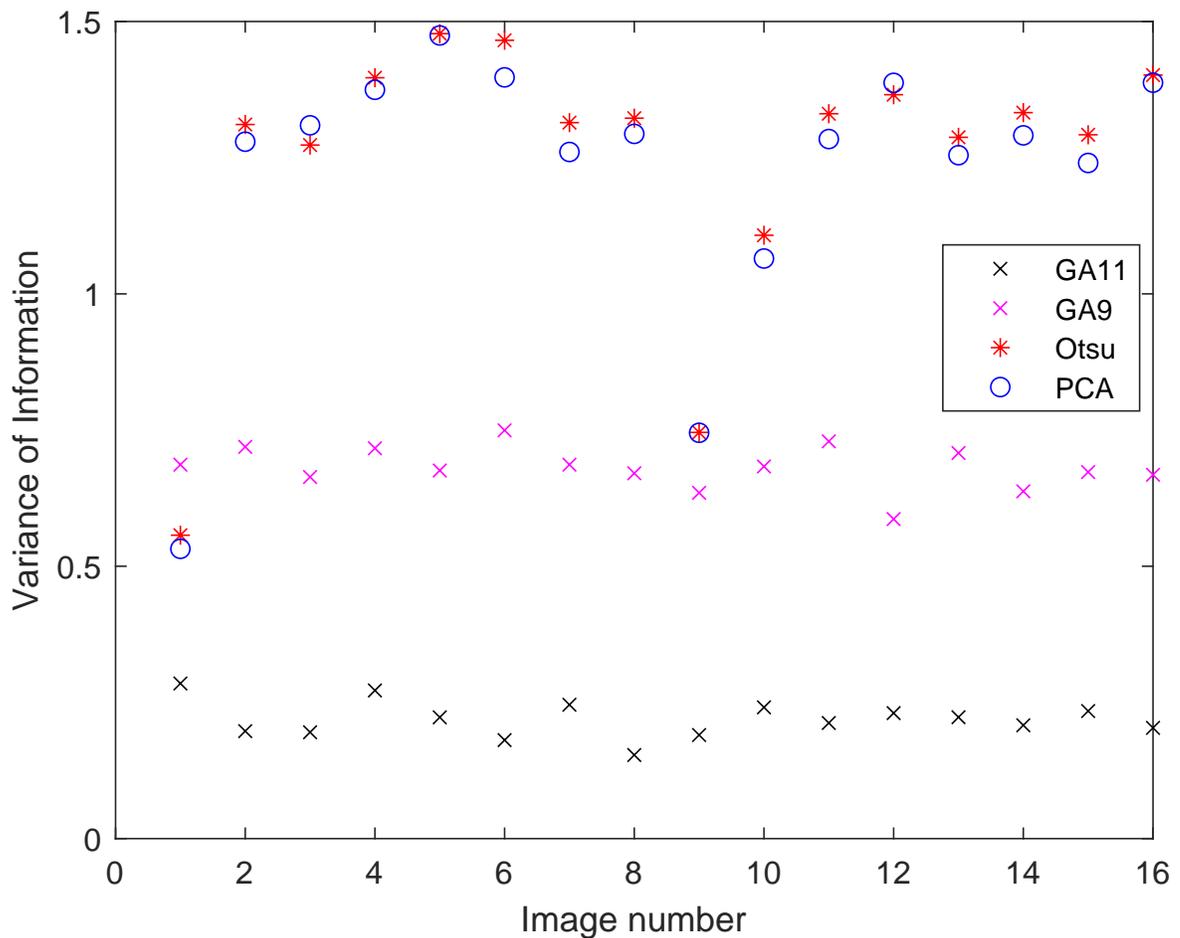


FIGURA 4.13: La Variance of Information (VoI) para cada imagen en el conjunto de imágenes de café

En el caso de la métrica Varianza de la Información, el rango de valores esta en  $[0, +\infty]$ . De forma similar al caso de GCE, cuanto más cerca de 0 indica que el resultado de la segmentación obtenida es similar a la segmentación manual o *Ground truth*.

La Figura 4.13 muestra el VoI para cada imagen con los diferentes métodos de segmentación. Se midió el VoI para los métodos propuestos GA9 ('x' magenta) y GA11

('x' negro) y el algoritmo de segmentación de Otsu ('\*' rojo) y la segmentación basada en PCA (círculos azules) y se compararon todas las imágenes con el conjunto de datos *Ground truth*. Los resultados experimentales con VoI promedian en todas las imágenes los siguientes resultados GA11: 0,2183, GA9: 0,6805, PCA: 1,2235, Otsu: 1,2488.

En este caso, los algoritmos Otsu y PCA segmentan bien la hoja, pero agregan una parte que no pertenece a la hoja debido a la mala iluminación. El algoritmo de segmentación debe segmentar solo la parte de la imagen para extraer mejores características y mejorar el rendimiento al clasificar.

#### 4.4.2. Resultados obtenidos con el conjunto de datos PlantVillage

Para el conjunto de datos de PlantVillage, la segmentación es más difícil porque la iluminación no se ocupó por completo y las técnicas de segmentación confunden la región requerida con las sombras. Debido a que muchas imágenes tienen este problema, la segmentación es generalmente muy pobre. Se utilizaron ciento once imágenes aleatorias para estas comparaciones. Estos fueron segmentados manualmente y comparados con los resultados obtenidos con las técnicas Otsu, PCA y los algoritmos GA9 y GA11 propuestos.

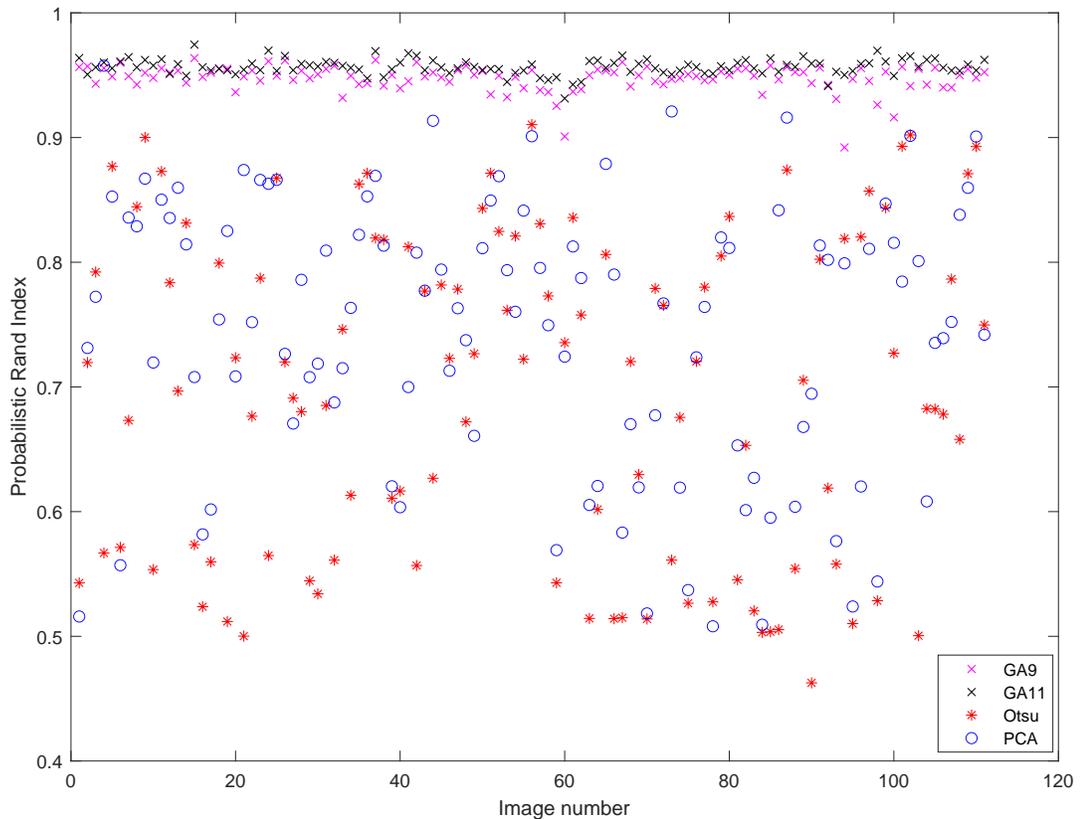


FIGURA 4.14: The Probabilistic Rand Index (PRI) for each image in Plant-Village dataset

La Figura 4.14 muestra los resultados obtenidos utilizando la métrica PRI para cada imagen con los diferentes métodos de segmentación. La Figura muestra los resultados obtenidos con el método propuesto GA9 ('x' en magenta) y GA11 ('x' en negro) y el algoritmo de segmentación de Otsu ('\*' en rojo) y la segmentación basada en PCA (círculos azules) y se compararon todas las imágenes con el conjunto de datos *Ground truth*. El promedio de la métrica PRI en todas las imágenes fue de GA11: 0,9566, GA9: 0,9475, PCA: 0,7446, Otsu: 0,6971.

La Figura 4.15 muestra la métrica GCE para cada imagen con los diferentes métodos de segmentación. La Figura muestra los resultados obtenidos con el método propuesto GA9 ('x' en magenta), GA11 ('x' en negro) y el algoritmo de segmentación de Otsu ('\*' en rojo) y la segmentación basada en PCA (círculos azules) y se compararon todas las imágenes con el conjunto de datos *Ground truth*. El promedio de la métrica GCE en todas las imágenes fue GA11: 0.0742, GA9: 0.0808, PCA: 0.2466, Otsu: 0.2845.

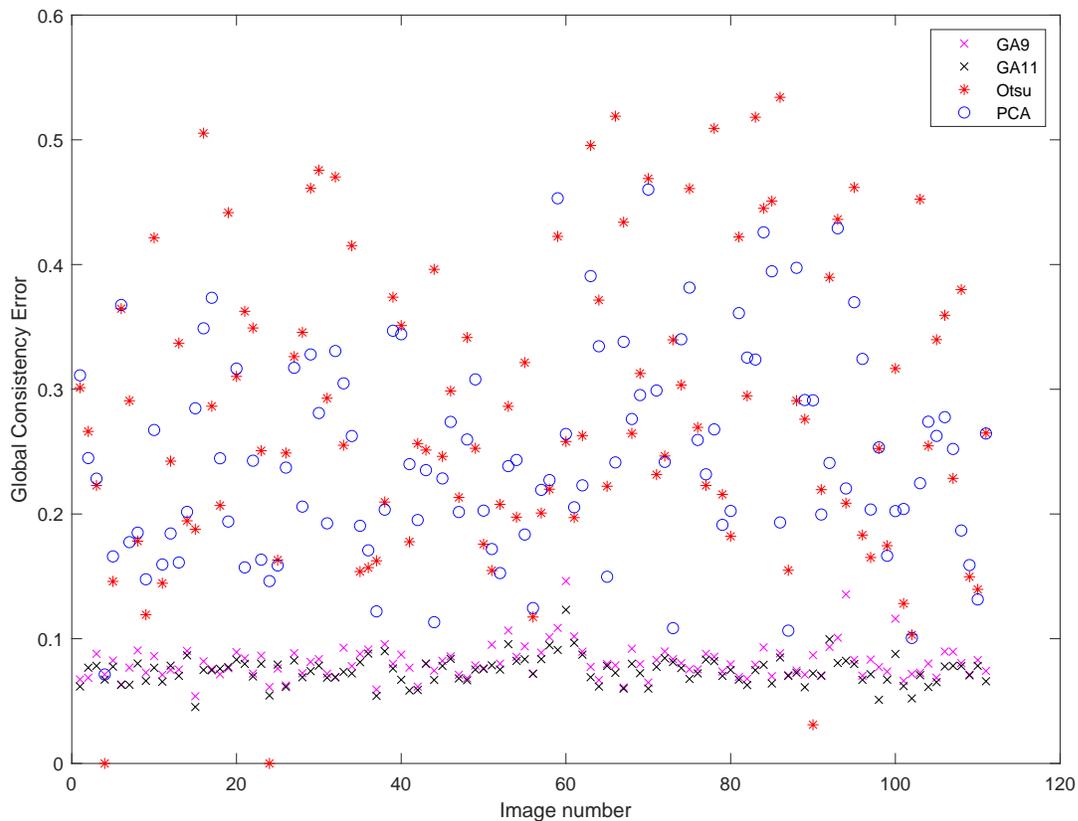


FIGURA 4.15: The Global Consistency Error (GCE) for each image in Plant-Village dataset

La Figura 4.16 muestra el VoI para cada imagen con los diferentes métodos de segmentación. La Figura muestra los resultados obtenidos con el método propuesto GA9 ('x' en magenta), GA11 ('x' en negro) y el algoritmo de segmentación de Otsu ('\*' en rojo) y la segmentación basada en PCA (círculos azules) y se compararon todas las imágenes con el conjunto de datos *Ground truth*. El promedio de la métrica GCE en todas las imágenes fue GA11: 0.7431, GA9: 0.8091, PCA: 1.7293, Otsu: 1.9585.

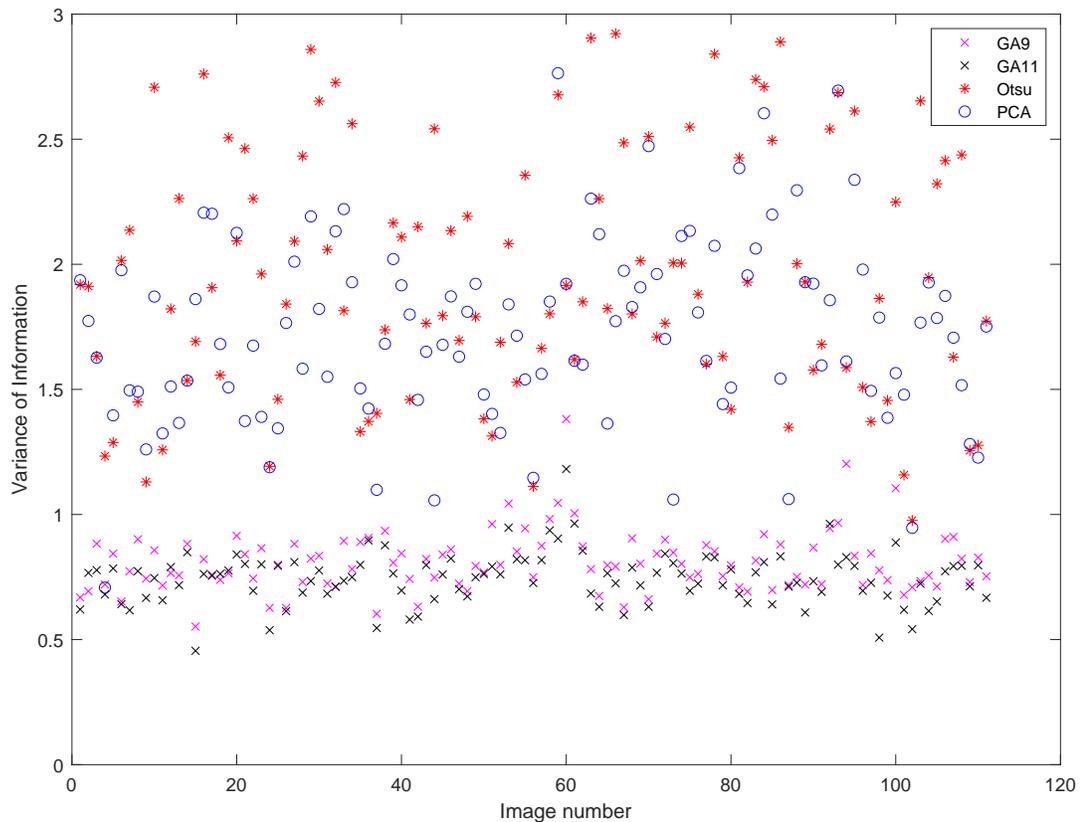


FIGURA 4.16: The Variance of Information (VoI) for each image in PlantVillage dataset

En las tres gráficas 4.14 4.15 4.16 es fácil ver como los resultados son mejores con las técnicas propuestas. En las gráficas comparativas con las tres métricas siempre los algoritmos propuestos mejoran los resultados en comparación con las técnicas clásicas.

#### 4.4.3. Resultados obtenidos con el conjunto de datos LSM

Para el conjunto de datos LSM, la segmentación es aún más difícil porque no solo hay una región para segmentar, sino varias regiones. Lo anterior parece no ser tomado en cuenta por las métricas de comparación, ya que en los resultados experimentales llevados a cabo no se ve reflejado.

El conjunto de datos utilizado presenta varios problemas de iluminación que complican la segmentación. Se utilizaron veintiséis imágenes aleatorias para estas comparaciones. Estos fueron segmentados manualmente y comparados con los resultados obtenidos con las técnicas Otsu, PCA y los algoritmos GA9 y GA11 propuestos.

La Figura 4.17 muestra los resultados obtenidos utilizando la métrica PRI para cada imagen con los diferentes métodos de segmentación. La Figura muestra los resultados obtenidos con el método propuesto GA9 ('x' en magenta) y GA11 ('x' en negro) y el algoritmo de segmentación de Otsu ('\*' en rojo) y la segmentación basada en PCA (círculos azules) y se compararon todas las imágenes con el conjunto de datos *Ground truth*. El promedio de la métrica PRI en todas las imágenes fue GA11: 0,9925, GA9: 0,9668, PCA: 0,9701, Otsu: 0,9378.

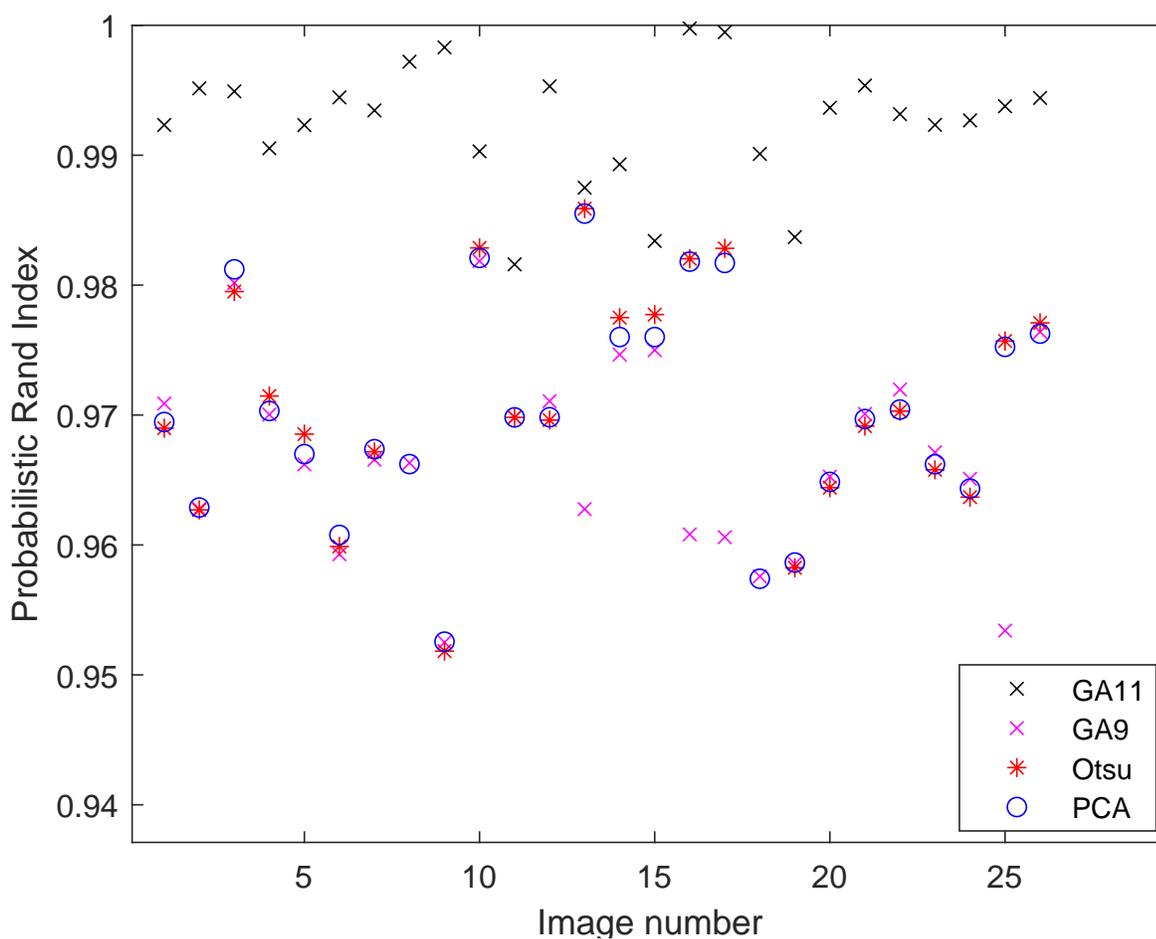


FIGURA 4.17: The Probabilistic Rand Index (PRI) for each image in LSM dataset

La Figura 4.18 muestra la métrica GCE para cada imagen con los diferentes métodos de segmentación. La Figura muestra los resultados obtenidos con el método propuesto GA9 ('x' en magenta), GA11 ('x' en negro) y el algoritmo de segmentación de Otsu ('\*' en rojo) y la segmentación basada en PCA (círculos azules) y se compararon todas las

imágenes con el conjunto de datos *Ground truth*. El promedio de la métrica GCE en todas las imágenes fue GA11: 0.0240, GA9: 0.0316, PCA: 0.0185, Otsu: 0.0394.

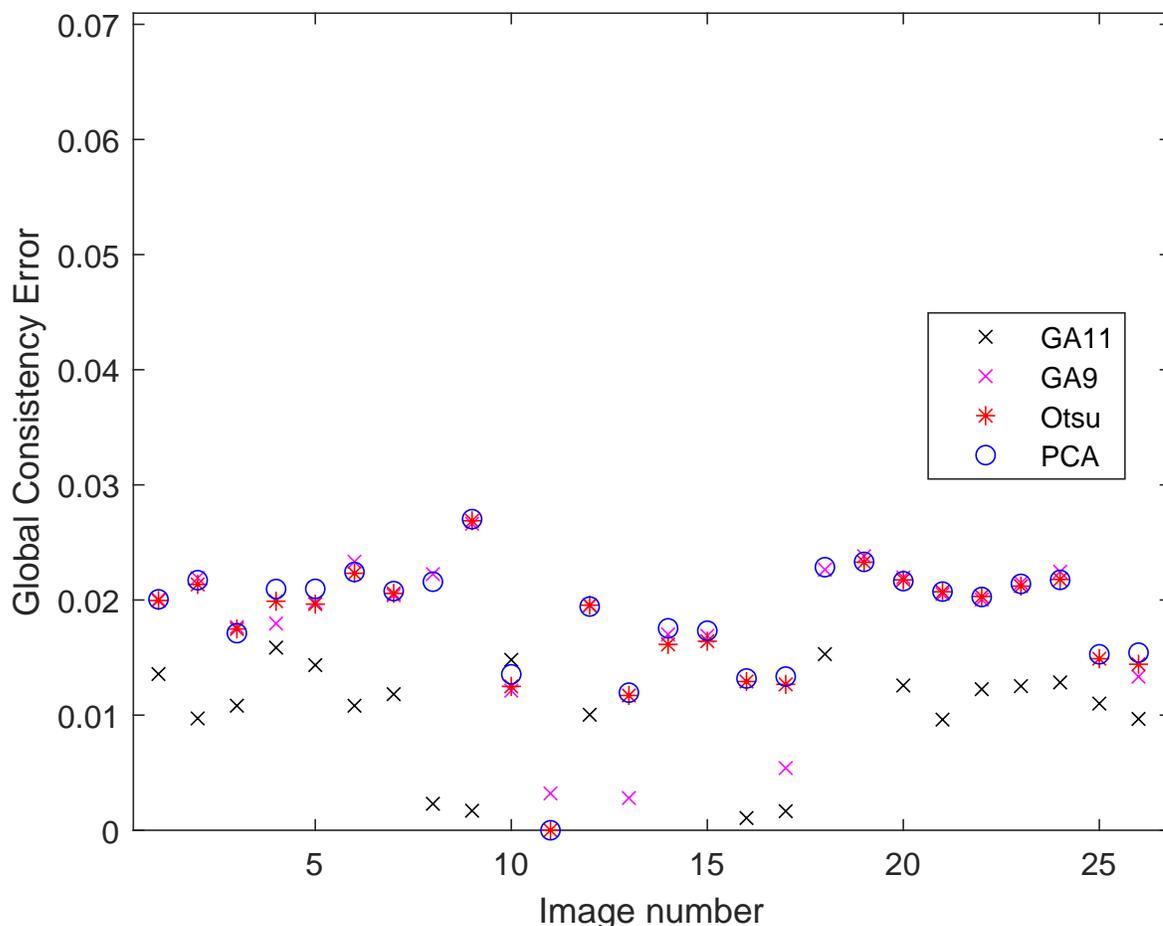


FIGURA 4.18: The Global Consistency Error (GCE) for each image in LSM dataset

La Figura 4.19 muestra el VoI para cada imagen con los diferentes métodos de segmentación. La Figura 4.19 muestra los resultados obtenidos con el método propuesto GA9 ('x' en magenta), GA11 ('x' en negro) y el algoritmo de segmentación de Otsu ('\*' en rojo) y la segmentación basada en PCA (círculos azules) y se compararon todas las imágenes con el conjunto de datos *Ground truth*. El promedio de la métrica GCE en todas las imágenes fue GA11: 0.1198, GA9: 0.2341, PCA: 0.2310, Otsu: 0.3150.

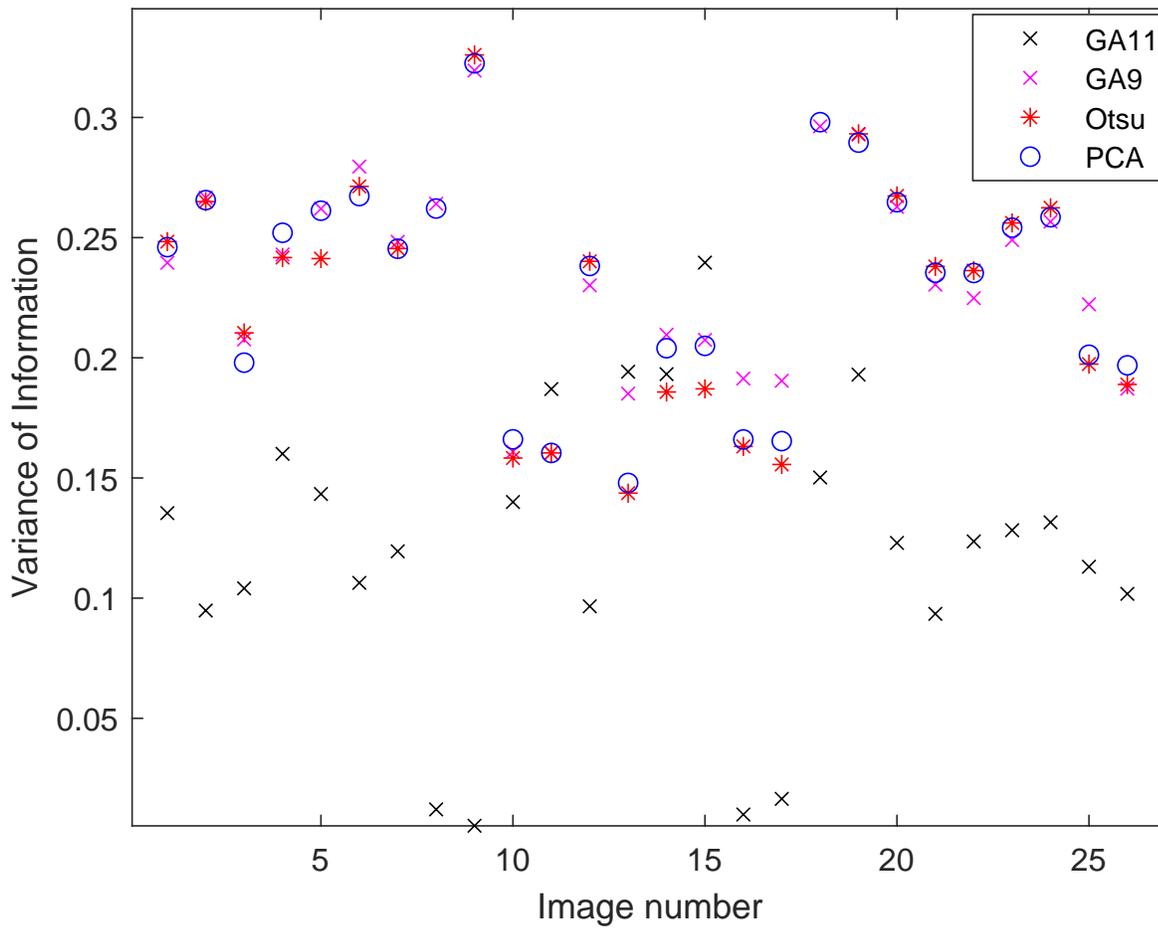


FIGURA 4.19: The Variance of Information (VoI) for each image in LSM dataset

En este conjunto de datos, los resultados obtenidos con PCA son incluso mejores que los obtenidos con el algoritmo GA9 propuesto; Se realizaron varias pruebas para verificar los resultados y los resultados son muy similares. Los aquí reportados son una muestra de los obtenidos.

# Capítulo 5

## Conclusiones

En esta tesis, se proponen dos nuevos algoritmos de segmentación de imágenes para conjuntos de datos. Los algoritmos propuestos permiten obtener la mejor combinación de variables y algoritmos de segmentación para segmentar una imagen. Los algoritmos fueron probados y comparados con las técnicas Otsu, y segmentación basada en PCA. Se probaron cuatro conjuntos de imágenes en los experimentos llevados a cabo. Los experimentos realizados muestran cómo la mala iluminación afecta a los algoritmos básicos, mientras que los algoritmos propuestos encuentran una buena combinación de variables para segmentar.

El cálculo de la identificación de regiones de una imagen que son homogéneas y conectadas se considera un problema de búsqueda combinatoria que consume mucho tiempo y, en ocasiones, las definiciones de estas regiones son inadecuadas. Los AG son eficaces en el caso de problemas combinatorios, ya que permiten la exploración paralela del espacio de búsqueda.

En consecuencia, esta Tesis propone un nuevo enfoque de segmentación basado en AG que puede mejorar el tiempo computacional y darnos una mejor calidad de segmentación. El método consiste en probar aleatoriamente toda una generación de posibles soluciones, seleccionar las mejores opciones y luego mejorarlas a través de la evolución; cuando se encuentran los parámetros necesarios, se aplican al conjunto restante de imágenes. Se observó que podían producirse errores al clasificar la luz de fondo como píxeles de interés aplicando individualmente varios métodos de segmentación a una imagen. La combinación de operadores morfológicos y de preprocesamiento reduce estos problemas.

Con el análisis de las principales técnicas de segmentación y su respectiva clasificación, entendemos que no existe una técnica perfecta para la segmentación de imágenes.

Cada imagen es individual y tiene sus características. Por esta razón, no se puede generalizar un solo método. En la mayoría de los casos, el resultado de una buena segmentación combinará varias técnicas, operadores y técnicas de preprocesamiento. Teniendo en cuenta lo anterior, el uso de algoritmos genéticos para segmentación automática pueden dar mejores resultados, reduciendo costos y acelerando el proceso.

Como trabajo futuro se debe considerar el uso de algoritmos de segmentación con más variables para mejorar la calidad de la segmentación en conjuntos de imágenes más complejos. También es fundamental estudiar las métricas de comparación de las imágenes segmentadas y proponer nuevas métricas de rendimiento. Especialmente en conjuntos de imágenes con varias regiones requeridas. Incluso agregando factores de penalización para regiones pequeñas pero significativas.

# Bibliografía

- [1] M. Abbasgholipour y col. «Color image segmentation with genetic algorithm in a raisin sorting system based on machine vision in variable conditions». En: *Expert Systems with Applications* 38.4 (2011), págs. 3671-3678. DOI: [10.1016/j.eswa.2010.09.023](https://doi.org/10.1016/j.eswa.2010.09.023).
- [2] AidinSakhavati Akbar ShahrzadKhashandarag MirkamalMirnia. «A New Method for Medical Image Clustering Using Genetic Algorithm». En: *IJCSI International Journal of Computer Science Issues* 10.1 (ene. de 2013).
- [3] Ali Qusay Al-Faris y col. «Breast MRI Tumour Segmentation Using Modified Automatic Seeded Region Growing Based on Particle Swarm Optimization Image Clustering». En: *Advances in Intelligent Systems and Computing*. Springer International Publishing, 2013, págs. 49-60. DOI: [10.1007/978-3-319-00930-8\\_5](https://doi.org/10.1007/978-3-319-00930-8_5).
- [4] Alex Alexandridis y col. «A two-stage evolutionary algorithm for variable selection in the development of RBF neural network models». En: *Chemometrics and Intelligent Laboratory Systems* 75.2 (2005), págs. 149-162. DOI: [10.1016/j.chemolab.2004.06.004](https://doi.org/10.1016/j.chemolab.2004.06.004).
- [5] O. Banimelhem y Ya Yahya. «Multi-Thresholding Image Segmentation Using Genetic Algorithm». En: 2011.
- [6] Vitoantonio Bevilacqua y Giuseppe Mastronardi. «Image Segmentation Using a Genetic Algorithm». En: *Soft Computing Applications*. Physica-Verlag HD, 2003, págs. 115-126. DOI: [10.1007/978-3-7908-1768-3\\_11](https://doi.org/10.1007/978-3-7908-1768-3_11).
- [7] Bir Bhanu, Sungkee Lee y John Ming. «Adaptive image segmentation using a genetic algorithm». En: *IEEE Transactions on Systems, Man, and Cybernetics* 25.12 (1995), págs. 1543-1567. DOI: [10.1109/21.478444](https://doi.org/10.1109/21.478444).
- [8] John Canny. «A Computational Approach to Edge Detection». En: *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI*-8.6 (1986), págs. 679-698. DOI: [10.1109/tpami.1986.4767851](https://doi.org/10.1109/tpami.1986.4767851).

- [9] Miguel Á. Castillo-Martínez y col. «Color index based thresholding method for background and foreground segmentation of plant images». En: *Computers and Electronics in Agriculture* 178 (2020), pág. 105783. DOI: [10.1016/j.compag.2020.105783](https://doi.org/10.1016/j.compag.2020.105783).
- [10] Jair Cervantes y col. «Análisis Comparativo de las técnicas utilizadas en un Sistema de Reconocimiento de Hojas de Planta». En: *Revista Iberoamericana de Automática e Informática Industrial RIAI* 14.1 (2017), págs. 104-114. DOI: [10.1016/j.riai.2016.09.005](https://doi.org/10.1016/j.riai.2016.09.005).
- [11] Jair Cervantes y col. «Complex Identification of Plants from Leaves». En: *Intelligent Computing Methodologies*. Springer International Publishing, 2018, págs. 376-387. DOI: [10.1007/978-3-319-95957-3\\_41](https://doi.org/10.1007/978-3-319-95957-3_41).
- [12] Shouvik Chakraborty y Kalyani Mali. «Fuzzy Electromagnetism Optimization (FEMO) and its application in biomedical image segmentation». En: *Applied Soft Computing* 97 (2020), pág. 106800. DOI: [10.1016/j.asoc.2020.106800](https://doi.org/10.1016/j.asoc.2020.106800).
- [13] Toby Breckon Chris Solomon. *Fundamentals of Digital Image Processing*. John Wiley Sons Inc, 20 de dic. de 2010. 344 págs. ISBN: 0470844736. URL: [https://www.ebook.de/de/product/3996688/chris\\_solomon\\_toby\\_breckon\\_fundamentals\\_of\\_digital\\_image\\_processing.html](https://www.ebook.de/de/product/3996688/chris_solomon_toby_breckon_fundamentals_of_digital_image_processing.html).
- [14] Naveen Singh Dagar y Pawan Kumar Dahiya. «Edge Detection Technique using Binary Particle Swarm Optimization». En: *Procedia Computer Science* 167 (2020), págs. 1421-1436. DOI: [10.1016/j.procs.2020.03.353](https://doi.org/10.1016/j.procs.2020.03.353).
- [15] Antonios Danelakis, Theoharis Theoharis y Dimitrios A. Verganelakis. «Survey of automated multiple sclerosis lesion segmentation techniques on magnetic resonance imaging». En: *Computerized Medical Imaging and Graphics* 70 (2018), págs. 83-100. DOI: [10.1016/j.compmedimag.2018.10.002](https://doi.org/10.1016/j.compmedimag.2018.10.002).
- [16] Sourav De, Siddhartha Bhattacharyya y Paramartha Dutta. «Automatic magnetic resonance image segmentation by fuzzy intercluster hostility index based genetic algorithm: An application». En: *Applied Soft Computing* 47 (2016), págs. 669-683. DOI: [10.1016/j.asoc.2016.05.042](https://doi.org/10.1016/j.asoc.2016.05.042).
- [17] I. Diaz, J. Branch y P. Boulanger. «A Genetic Algorithm to Segment Range Image by Edge Detection». En: *2005 International Conference on Industrial Electronics and Control Applications*. IEEE. DOI: [10.1109/icieca.2005.1644380](https://doi.org/10.1109/icieca.2005.1644380).

- [18] Z DOKUR y T OLMEZ. «Tissue segmentation in ultrasound images by using genetic algorithms». En: *Expert Systems with Applications* 34.4 (2008), págs. 2739-2746. DOI: [10.1016/j.eswa.2007.05.002](https://doi.org/10.1016/j.eswa.2007.05.002).
- [19] Abraham Duarte y col. «Improving image segmentation quality through effective region merging using a hierarchical social metaheuristic». En: *Pattern Recognition Letters* 27.11 (2006), págs. 1239-1251. DOI: [10.1016/j.patrec.2005.07.022](https://doi.org/10.1016/j.patrec.2005.07.022).
- [20] Josué Espejel-Cabrera y col. «Mexican sign language segmentation using color based neuronal networks to detect the individual skin color». En: *Expert Systems with Applications* 183 (2021), pág. 115295. DOI: [10.1016/j.eswa.2021.115295](https://doi.org/10.1016/j.eswa.2021.115295).
- [21] Daniel Freedman. «An improved image graph for semi-automatic segmentation». En: *Signal, Image and Video Processing* 6.4 (2010), págs. 533-545. DOI: [10.1007/s11760-010-0181-9](https://doi.org/10.1007/s11760-010-0181-9).
- [22] Farid Garcia-Lamont y col. «Segmentation of images by color features: A survey». En: *Neurocomputing* 292 (2018), págs. 1-27. DOI: [10.1016/j.neucom.2018.01.091](https://doi.org/10.1016/j.neucom.2018.01.091).
- [23] David E. Goldberg y John H. Holland. En: *Machine Learning* 3.2/3 (1988), págs. 95-99. DOI: [10.1023/a:1022602019183](https://doi.org/10.1023/a:1022602019183).
- [24] Wilfrido Gómez-Flores y Juanita Hernández-López. «Automatic adjustment of the pulse-coupled neural network hyperparameters based on differential evolution and cluster validity index for image segmentation». En: *Applied Soft Computing* 97 (2020), pág. 105547. DOI: [10.1016/j.asoc.2019.105547](https://doi.org/10.1016/j.asoc.2019.105547).
- [25] Rafael Gonzalez. *Digital image processing*. Upper Saddle River, N.J: Prentice Hall, 2002. ISBN: 0201180758.
- [26] Jatin Gupta, Sumindar Kaur Saini y Mamta Juneja. «Survey of denoising and segmentation techniques for MRI images of prostate for improving diagnostic tools in medical applications». En: *Materials Today: Proceedings* 28 (2020), págs. 1667-1672. DOI: [10.1016/j.matpr.2020.05.023](https://doi.org/10.1016/j.matpr.2020.05.023).
- [27] Elisseff A. Kaelbling Guyon I. «An introduction to variable and feature selection». En: *Journal of Machine Learning Research* (2003), 1157-1182.
- [28] Rongxiang Hu y col. «Multiscale Distance Matrix for Fast Plant Leaf Recognition». En: *IEEE Transactions on Image Processing* 21.11 (2012), págs. 4667-4672. DOI: [10.1109/tip.2012.2207391](https://doi.org/10.1109/tip.2012.2207391).

- [29] David. P. Hughes y Marcel Salathe. «An open access repository of images on plant health to enable the development of mobile disease diagnostics». En: (25 de nov. de 2015). arXiv: [1511.08060](https://arxiv.org/abs/1511.08060) [cs.CY].
- [30] Bram Huygens y col. «Application of evolutionary algorithms to optimise one- and two-dimensional gradient chromatographic separations». En: *Journal of Chromatography A* 1628 (2020), pág. 461435. DOI: [10.1016/j.chroma.2020.461435](https://doi.org/10.1016/j.chroma.2020.461435).
- [31] Seiji Ito y col. «An image segmentation method using histograms and the human characteristics of HSI color space for a scene image». En: *Artificial Life and Robotics* 10.1 (2006), págs. 6-10. DOI: [10.1007/s10015-005-0352-x](https://doi.org/10.1007/s10015-005-0352-x).
- [32] K. Janc y col. «Genetic algorithms as a useful tool for trabecular and cortical bone segmentation». En: *Computer Methods and Programs in Biomedicine* 111.1 (2013), págs. 72-83. DOI: [10.1016/j.cmpb.2013.03.012](https://doi.org/10.1016/j.cmpb.2013.03.012).
- [33] S. Jose y C. Vijayalakshmi. «Design and Analysis of Multi-Objective Optimization Problem Using Evolutionary Algorithms.» En: *Procedia Computer Science* 172 (2020), págs. 896-899. DOI: [10.1016/j.procs.2020.05.129](https://doi.org/10.1016/j.procs.2020.05.129).
- [34] Delphine. Jouan-Rimbaud y col. «Genetic Algorithms as a Tool for Wavelength Selection in Multivariate Calibration». En: *Analytical Chemistry* 67.23 (1995), págs. 4295-4301. DOI: [10.1021/ac00119a015](https://doi.org/10.1021/ac00119a015).
- [35] Muhammad Waseem Khan. «A Survey: Image Segmentation Techniques». En: *International Journal of Future Computer and Communication* (2014), págs. 89-93. DOI: [10.7763/ijfcc.2014.v3.274](https://doi.org/10.7763/ijfcc.2014.v3.274).
- [36] Renato A. Krohling. *BRACOL - A Brazilian Arabica Coffee Leaf images dataset to identification and quantification of coffee diseases and pests*. 2019. DOI: [10.17632/YY2K5Y8MXG.1](https://doi.org/10.17632/YY2K5Y8MXG.1).
- [37] Akshi Kumar y Sahil Raheja. «Edge Detection using Guided Image Filtering and Enhanced Ant Colony Optimization». En: *Procedia Computer Science* 173 (2020), págs. 8-17. DOI: [10.1016/j.procs.2020.06.003](https://doi.org/10.1016/j.procs.2020.06.003).
- [38] Paulo Roberto Gardel Kurka y Aldo André Díaz Salazar. «Applications of image processing in robotics and instrumentation». En: *Mechanical Systems and Signal Processing* 124 (2019), págs. 142-169. DOI: [10.1016/j.ymsp.2019.01.015](https://doi.org/10.1016/j.ymsp.2019.01.015).
- [39] K. Laws y col. «The Phoenix Image Segmentation System: Description and Evaluation». En: 1982.

- [40] S. Manikandan y col. «Multilevel thresholding for segmentation of medical brain images using real coded genetic algorithm». En: *Measurement* 47 (2014), págs. 558-568. DOI: [10.1016/j.measurement.2013.09.031](https://doi.org/10.1016/j.measurement.2013.09.031).
- [41] D. Martin y col. «A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics». En: *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*. IEEE Comput. Soc. DOI: [10.1109/iccv.2001.937655](https://doi.org/10.1109/iccv.2001.937655).
- [42] David Royal Martin. «An Empirical Approach to Grouping and Segmentation». Tesis doct. EECS Department, University of California, Berkeley, 2003. URL: <http://www2.eecs.berkeley.edu/Pubs/TechRpts/2003/5252.html>.
- [43] Marina Meilă. «Comparing clusterings». En: *Proceedings of the 22nd international conference on Machine learning - ICML '05*. ACM Press, 2005. DOI: [10.1145/1102351.1102424](https://doi.org/10.1145/1102351.1102424).
- [44] Sharada P. Mohanty, David P. Hughes y Marcel Salathé. «Using Deep Learning for Image-Based Plant Disease Detection». En: *Frontiers in Plant Science* 7 (2016). DOI: [10.3389/fpls.2016.01419](https://doi.org/10.3389/fpls.2016.01419).
- [45] Seyed Jaleleddin Mousavirad, Hossein Ebrahimpour-Komleh y Gerald Schaefer. «Automatic clustering using a local search-based human mental search algorithm for image segmentation». En: *Applied Soft Computing* 96 (2020), pág. 106604. DOI: [10.1016/j.asoc.2020.106604](https://doi.org/10.1016/j.asoc.2020.106604).
- [46] Songyot Nakariyakul y David P. Casasent. «An improvement on floating search algorithms for feature subset selection». En: *Pattern Recognition* 42.9 (2009), págs. 1932-1940. DOI: [10.1016/j.patcog.2008.11.018](https://doi.org/10.1016/j.patcog.2008.11.018).
- [47] Jifeng Ning y col. «Interactive image segmentation by maximal similarity based region merging». En: *Pattern Recognition* 43.2 (2010), págs. 445-456. DOI: [10.1016/j.patcog.2009.03.004](https://doi.org/10.1016/j.patcog.2009.03.004).
- [48] Pedro Ventura de Oliveira y Keiji Yamanaka. «Image Segmentation Using Multi-level Thresholding and Genetic Algorithm: An Approach». En: *2018 2nd International Conference on Data Science and Business Analytics (ICDSBA)*. IEEE, 2018. DOI: [10.1109/icdsba.2018.00078](https://doi.org/10.1109/icdsba.2018.00078).
- [49] Nobuyuki Otsu. «A Threshold Selection Method from Gray-Level Histograms». En: *IEEE Transactions on Systems, Man, and Cybernetics* 9.1 (1979), págs. 62-66. DOI: [10.1109/tsmc.1979.4310076](https://doi.org/10.1109/tsmc.1979.4310076).

- [50] Stelios E. Papadakis y col. «A Genetic Based Approach to the Type I Structure Identification Problem». En: *Informatica* 16.3 (2005), págs. 365-382. DOI: [10.15388/informatica.2005.104](https://doi.org/10.15388/informatica.2005.104).
- [51] Zhenkui Pei, Yanli Zhao y Zhen Liu. «Image segmentation based on Differential Evolution algorithm». En: *2009 International Conference on Image Analysis and Signal Processing*. IEEE, 2009. DOI: [10.1109/iasp.2009.5054643](https://doi.org/10.1109/iasp.2009.5054643).
- [52] Bo Peng, Lei Zhang y Jian Yang. «Iterated Graph Cuts for Image Segmentation». En: *Computer Vision – ACCV 2009*. Springer Berlin Heidelberg, 2010, págs. 677-686. DOI: [10.1007/978-3-642-12304-7\\_64](https://doi.org/10.1007/978-3-642-12304-7_64).
- [53] B.D. Phulpagar y R.S. Bichkar. «Segmentation of Images Containing Multiple Intensity Levels using Genetic Algorithms». En: *International Journal of Computer Applications* 126.9 (2015), págs. 29-39. DOI: [10.5120/ijca2015906186](https://doi.org/10.5120/ijca2015906186).
- [54] Thakare Punam; «A Study of Image Segmentation and EdgeDetection Techniques». En: *International Journal on Computer Science and Engineering (IJCSE)* vol 3 (2011), págs. 899-905.
- [55] W. Punch y col. «Further Research on Feature Selection and Classification Using Genetic Algorithms». En: *ICGA*. 1993.
- [56] William M. Rand. «Objective Criteria for the Evaluation of Clustering Methods». En: *Journal of the American Statistical Association* 66.336 (1971), págs. 846-850. DOI: [10.1080/01621459.1971.10482356](https://doi.org/10.1080/01621459.1971.10482356).
- [57] J.E. Ruiz y col. «Optimization of digital image processing to determine quantum dots' height and density from atomic force microscopy». En: *Ultramicroscopy* 184 (2018), págs. 234-241. DOI: [10.1016/j.ultramic.2017.09.004](https://doi.org/10.1016/j.ultramic.2017.09.004).
- [58] Indrajit Saha, Ujjwal Maulik y Dariusz Plewczynski. «A new multi-objective technique for differential fuzzy clustering». En: *Applied Soft Computing* 11.2 (2011), págs. 2765-2776. DOI: [10.1016/j.asoc.2010.11.007](https://doi.org/10.1016/j.asoc.2010.11.007).
- [59] Jianbo Shi y J. Malik. «Normalized cuts and image segmentation». En: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22.8 (2000), págs. 888-905. DOI: [10.1109/34.868688](https://doi.org/10.1109/34.868688).
- [60] R. Srikanth y K. Bikshalu. «Multilevel thresholding image segmentation based on energy curve with harmony Search Algorithm». En: *Ain Shams Engineering Journal* 12.1 (2021), págs. 1-20. DOI: [10.1016/j.asej.2020.09.003](https://doi.org/10.1016/j.asej.2020.09.003).

- [61] Ran Tao y col. «Analysis of Mafic rocks Microstructure damage and failure Process under Compression Test Using Quantitative Scanning Electron Microscopy and Digital Images Processing». En: *Engineering Fracture Mechanics* 231 (2020), pág. 107019. DOI: [10.1016/j.engfracmech.2020.107019](https://doi.org/10.1016/j.engfracmech.2020.107019).
- [62] Leslie Ching Ow Tiong, Seong Tae Kim y Yong Man Ro. «Multimodal facial biometrics recognition: Dual-stream convolutional neural networks with multi-feature fusion layers». En: *Image and Vision Computing* 102 (2020), pág. 103977. DOI: [10.1016/j.imavis.2020.103977](https://doi.org/10.1016/j.imavis.2020.103977).
- [63] R. Unnikrishnan y M. Hebert. «Measures of Similarity». En: *2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05) - Volume 1*. IEEE, 2005. DOI: [10.1109/acvmot.2005.71](https://doi.org/10.1109/acvmot.2005.71).
- [64] Sara Vicente, Vladimir Kolmogorov y Carsten Rother. «Graph cut based image segmentation with connectivity priors». En: *2008 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2008. DOI: [10.1109/cvpr.2008.4587440](https://doi.org/10.1109/cvpr.2008.4587440).
- [65] Hongzhi Wang y John Oliensis. «Generalizing edge detection to contour detection for image segmentation». En: *Computer Vision and Image Understanding* 114.7 (2010), págs. 731-744. DOI: [10.1016/j.cviu.2010.02.001](https://doi.org/10.1016/j.cviu.2010.02.001).
- [66] Jie Wang, Lili Ju y Xiaoqiang Wang. «An Edge-Weighted Centroidal Voronoi Tessellation Model for Image Segmentation». En: *IEEE Transactions on Image Processing* 18.8 (2009), págs. 1844-1858. DOI: [10.1109/tip.2009.2021087](https://doi.org/10.1109/tip.2009.2021087).
- [67] Xuan Wang y col. «Leaf Recognition Based on Elliptical Half Gabor and Maximum Gap Local Line Direction Pattern». En: *IEEE Access* 8 (2020), págs. 39175-39183. DOI: [10.1109/access.2020.2976117](https://doi.org/10.1109/access.2020.2976117).
- [68] Min Xian y col. «Automatic breast ultrasound image segmentation: A survey». En: *Pattern Recognition* 79 (2018), págs. 340-355. DOI: [10.1016/j.patcog.2018.02.012](https://doi.org/10.1016/j.patcog.2018.02.012).
- [69] Fengying Xie y Alan C. Bovik. «Automatic segmentation of dermoscopy images using self-generating neural networks seeded by genetic algorithm». En: *Pattern Recognition* 46.3 (2013), págs. 1012-1019. DOI: [10.1016/j.patcog.2012.08.012](https://doi.org/10.1016/j.patcog.2012.08.012).
- [70] Jihoon Yang y Vasant Honavar. «Feature Subset Selection Using a Genetic Algorithm». En: *Feature Extraction, Construction and Selection*. Springer US, 1998, págs. 117-136. DOI: [10.1007/978-1-4615-5725-8\\_8](https://doi.org/10.1007/978-1-4615-5725-8_8).

- 
- [71] Nida M. Zaitoun y Musbah J. Aqel. «Survey on Image Segmentation Techniques». En: *Procedia Computer Science* 65 (2015), págs. 797-806. DOI: [10.1016/j.procs.2015.09.027](https://doi.org/10.1016/j.procs.2015.09.027).
- [72] Shanwen Zhang, Xiaowei Wu y Zhuhong You. «Jaccard distance based weighted sparse representation for coarse-to-fine plant species recognition». En: *PLOS ONE* 12.6 (2017). Ed. por Zhao Zhang, e0178317. DOI: [10.1371/journal.pone.0178317](https://doi.org/10.1371/journal.pone.0178317).
- [73] Steven W. Zucker. «Region growing: Childhood and adolescence». En: *Computer Graphics and Image Processing* 5.3 (1976), págs. 382-399. DOI: [10.1016/s0146-664x\(76\)80014-7](https://doi.org/10.1016/s0146-664x(76)80014-7).