



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



Escola Tècnica  
Superior d'Enginyeria  
Informàtica

Escola Tècnica Superior d'Enginyeria Informàtica  
Universitat Politècnica de València

# Identificación de pólenes utilizando técnicas de visión por computador

Trabajo Fin de Grado

**Grado en Ingeniería Informática**

**Autor:** Felipe Peinado Sanchis

**Tutor:** Dr. Fernando López García

2014-2015

Trabajo de Fin de Grado  
Grado en Ingeniería Informática

# Identificación de pólenes utilizando técnicas de visión por computador

Felipe Peinado Sanchis  
Director: Dr. Fernando López García

02/09/2015



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

Universidad Politécnica de Valencia  
Escuela Técnica Superior de Informática



# Resumen

En este trabajo se evalúa la viabilidad de un sistema automático de clasificación de pólenes en imágenes de miel obtenidas con microscopio basado en un modelo que usa como discriminantes atributos derivados del estudio del color y la textura.

En primer lugar se establece cómo podría favorecer la Visión por Computador al campo de la Melisopalinología para solucionar las necesidades de clasificación de pólenes con unos niveles de garantía similares a los del criterio experto.

Realizamos una serie de experimentos a partir de un conjunto de 12 imágenes con 138 pólenes identificados y clasificados de manera manual. En dichos experimentos se contempla la viabilidad de los descriptores suaves derivados del color y la textura como factores discriminantes de clasificación entre especies. Estos atributos se obtienen en tres espacios de color distintos: escala de grises, RGB y CIE LAB, siendo este último perceptualmente uniforme. Los resultados se evalúan de acuerdo a la tasa de aciertos del modelo, y se estudia detalladamente el comportamiento del modelo con respecto a las categorías más conflictivas.

A continuación realizamos más experimentos utilizando atributos derivados de la matriz de co-ocurrencia de la textura de las imágenes, y se presentan las configuraciones que mejores resultados han dado para el conjunto de pólenes de los que disponemos. También aplicamos de forma análoga la extracción de atributos en el dominio de las frecuencias usando wavelets y analizamos sus patrones binarios locales, definiendo qué combinación de descriptores resulta más efectiva.

Observando la evolución en la tasa de acierto de nuestro modelo, finalmente proponemos varios trabajos futuros para la continuidad del estudio.



# Agradecimientos

A todos aquellos, amigos y familiares, que me han ayudado a seguir planteándome nuevos retos y a aquellos que me han facilitado las herramientas para superarlos.

A los doctores Fernando López y Gabriela Andreu, sin duda este trabajo no podría haberlo realizado sin su ayuda y tutela.

Agradecer a Dña Isabel Escriche Roberto (catedrática) y a Dña María del Sol Juan Borrás (investigadora) del instituto Universitario de Ingeniería Agronómica y del Medio Natural que nos hayan cedido generosamente imágenes de polenes etiquetadas. Sin su gran aportación este trabajo no tendría sentido.



# Índice general

<b>Abstract</b>	<b>I</b>
<b>Agradecimientos</b>	<b>III</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Motivación . . . . .	1
1.2. Organización de este trabajo . . . . .	2
1.3. Melisopalinología . . . . .	2
1.4. Reconocimiento automático de pólenes . . . . .	4
<b>2. Estudio de atributos de imágenes</b>	<b>5</b>
2.1. Color . . . . .	5
2.1.1. RGB . . . . .	6
2.1.2. CIE LAB . . . . .	6
2.2. Textura . . . . .	8
2.2.1. Matrices de co-ocurrencia . . . . .	9
2.3. Descriptores . . . . .	11
2.3.1. Descriptores suaves de color y textura . . . . .	11
2.3.2. Descriptores de Haralick . . . . .	12
2.3.3. Descriptores basados en wavelets . . . . .	15
2.3.4. Patrones Binarios Locales (LBP) . . . . .	18
<b>3. Experimentación</b>	<b>21</b>
3.1. Datos de los que disponemos . . . . .	21
3.2. Segmentación y clasificación automáticas . . . . .	23
3.2.1. Experimentación sobre la imagen . . . . .	24
3.2.1.1. Obtención de las imágenes y del polen muestra . . . . .	24
3.2.1.2. Obtención de los atributos . . . . .	26
3.2.1.3. Barrido de la imagen . . . . .	27
3.2.1.4. Evaluación de resultados . . . . .	28
3.3. Segmentación manual y clasificación automática . . . . .	30
3.3.1. Usando descriptores suaves de color y textura . . . . .	30
3.3.1.1. Preprocesado de las imágenes y elaboración de la BBDD . . . . .	31
3.3.1.2. Obtención de los atributos . . . . .	31
3.3.1.3. Generando el modelo . . . . .	32
3.3.1.4. Evaluación de resultados . . . . .	33

3.3.2.	Usando matrices de co-ocurrencia . . . . .	36
3.3.2.1.	Preprocesado de las imágenes y elaboración de la BBDD . . . . .	37
3.3.2.2.	Obtención de los atributos . . . . .	37
3.3.2.3.	Generando el modelo . . . . .	38
3.3.2.4.	Evaluación de resultados . . . . .	38
3.3.3.	Usando descriptores suaves y matrices de co-ocurrencia . . . . .	41
3.3.3.1.	Evaluación de resultados . . . . .	41
3.3.4.	Experimentación usando wavelets . . . . .	43
3.3.4.1.	Preprocesado de las imágenes y elaboración del banco de filtros log-Gabor . . . . .	44
3.3.4.2.	Obtención de los atributos . . . . .	44
3.3.4.3.	Generando el modelo . . . . .	44
3.3.4.4.	Evaluación de resultados . . . . .	45
3.3.5.	Usando descriptores suaves, matrices de co-ocurrencia y wavelets . . . . .	46
3.3.5.1.	Evaluación de resultados . . . . .	46
3.3.6.	Experimentación usando Patrones Binarios Locales (LBP) . . . . .	47
3.3.6.1.	Preprocesado de las imágenes . . . . .	47
3.3.6.2.	Obtención de los atributos y generación del modelo . . . . .	47
3.3.6.3.	Evaluación de resultados . . . . .	48
3.3.7.	Usando descriptores suaves, matrices de co-ocurrencia, wavelets y patrones binarios locales . . . . .	50
3.3.7.1.	Evaluación de resultados . . . . .	50
<b>4.</b>	<b>Conclusiones</b>	<b>53</b>
4.1.	Trabajo futuro . . . . .	55
<b>A.</b>	<b>Anexo A: Imágenes de pólenes proporcionadas</b>	<b>57</b>
<b>B.</b>	<b>Anexo B: Muestras segmentadas manualmente.</b>	<b>65</b>
B.1.	Helianthus . . . . .	65
B.2.	Prunus . . . . .	66
B.3.	Rosmarinus . . . . .	66
B.4.	Polygonaceas . . . . .	67
B.5.	Diploaxis . . . . .	67
B.6.	Eryngium . . . . .	68
B.7.	Erica . . . . .	68
B.8.	Quercus sp. . . . .	69
B.9.	Trifolium . . . . .	69
B.10.	Tipo Genista . . . . .	69
B.11.	Lavandula . . . . .	69
B.12.	Taraxacum sp. . . . .	70
B.13.	Sedum . . . . .	70
B.14.	Thymus . . . . .	70

B.15.L.Stoechas . . . . .	70
<b>Bibliografía</b>	<b>71</b>



# 1. Introducción

## 1.1. Motivación

El avance imparable del desarrollo tecnológico es uno de los principales motores de mejora continua en la industria. La revolución informática ha permitido mejorar mediante la automatización la eficacia y la eficiencia de una gran cantidad de procesos que previamente se realizaban de manera manual y cuyo rendimiento queda supeditado a las limitaciones del potencial humano.

Entre estas aplicaciones de la informática industrial se encuentra la Visión por Computador, que es una rama derivada de la Inteligencia Artificial encargada de procesar y analizar imágenes del mundo real para obtener información relevante a partir de ellas. Estas imágenes pueden proceder de fuentes diversas, por ejemplo imágenes bidimensionales, secuencias de vídeo, estructuras multidimensionales procedentes de escáneres médicos, etc. En su faceta de aplicación técnica, la visión por computador tiene como objetivo hacer uso de modelos y teorías para construir sistemas de visión para fines diversos como el reconocimiento y clasificación de objetos, reconstrucción de escenas, detección de eventos, aprendizaje, inspección de defectos, restauración de imágenes, etc...

No obstante no todas las tareas son susceptibles de ser automatizadas fácilmente. El factor humano de la experiencia profesional y su adaptabilidad ante variaciones en los problemas son elementos claves en la toma de muchas decisiones, y por ello surgen muchas dificultades a la hora de generar sistemas expertos capaces de alcanzar los niveles de eficacia requeridos. Estas barreras, no triviales, fuerzan a la industria a invertir en procesos de ingeniería para analizar la viabilidad de desarrollar sistemas tecnológicos que solventen estas necesidades.

Una de estas tareas se da en la industria de la miel: El recuento y clasificación de los pólenes que forman parte de la miel es una tarea clave en el análisis de la calidad y la procedencia de la mezcla, y es una tarea que se realiza de manera manual, a través de imágenes de microscopio y que requiere del criterio de palinólogos y melisopalinólogos experimentados para ser funcional. Entre los atributos que fijan el criterio del analista experto se encuentran el color, la textura y propiedades morfológicas.

El objetivo del presente documento es, dado un conjunto de imágenes de microscopio procedentes de miel, establecer la viabilidad de un modelo, en un sistema software basado en MATLAB, capaz de solventar con una fiabilidad similar a la de un palinólogo experto la detección y clasificación de los pólenes que aparecen en dicha imagen.

Para ello nos basaremos inicialmente en un estudio de un conjunto de descriptores obtenidos a través del color y la textura y posteriormente se seguirá estudiando con atributos derivados del estudio más exhaustivo de las texturas.

Como en cualquier investigación, en este proyecto se aspira a obtener unos resultados cuya validez sea medible. La evaluación de nuestro sistema se realizará primero mediante análisis del éxito del enfoque dada una serie de mapas de cercanía creados sobre las imágenes ejemplo, y posteriormente mediante la tasa de aciertos en diferentes condiciones de un conjunto de modelos de clasificación propuestos.

## 1.2. Organización de este trabajo

Este documento se organiza de la siguiente manera: En este capítulo se plantean los objetivos de este trabajo y su estructuración. Además, se introduce la melisopalínología como ciencia orientada al estudio de los pólenes en la miel y la necesidad de automatizar ciertos procesos mediante la Visión por Computador.

En el Capítulo 2 describiremos los fundamentos teóricos sobre los que se sustentan los elementos que hemos escogido como discriminantes para nuestros sistemas de clasificación.

El Capítulo 3 versa sobre la aplicación técnica de dichos fundamentos y el proceso de experimentación realizado, estando dividido en bloques según el sistema y los criterios utilizados. En este apartado serán presentadas las diferentes técnicas utilizadas para realizar el experimento, el preprocesamiento de los datos y los resultados finales.

Finalmente, el Capítulo 4 mostrará las conclusiones a las que se ha llegado y el trabajo futuro propuesto por el autor.

En el Anexo A se recogen las 12 imágenes de miel obtenidas con microscopio en las que se basa este trabajo y en el Anexo B la segmentación manual de las imágenes en distintas clases de pólenes.

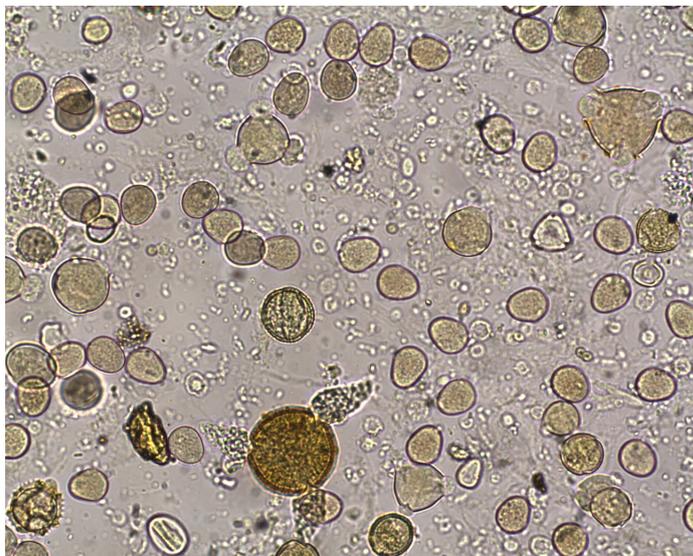
## 1.3. Melisopalínología

La miel es un producto natural, una disolución acuosa concentrada de hidratos de carbono, diversas enzimas, aminoácidos, ácidos orgánicos, minerales, sustancias aromáticas, pigmentos ceras, granos de polen, etc [1].

Derivada de la palinología, la melisopalínología (de la voz griega “melissa”, miel y de “palinología”, ciencia de los pólenes) es una rama de la botánica dedicada al estudio de la presencia de pólenes en la miel con el objetivo de determinar y autenticar la procedencia geográfica de dichas mieles y la identificación de la calidad del producto, basándose en la presencia significativa de determinados tipos de pólenes [2].

El primer trabajo relacionado con la miel a nivel microscópico data del final del siglo XIX, [3] y el trabajo de autores como Griebel [4], Erdtman [5] o Zander [6] entre los años 30 y 50 asientan las bases sobre las que se erigirán las técnicas analíticas posteriormente [7]. Como referencia, la Comisión Internacional para Botánica Apícola (ICBB, del inglés International Commission for Bee Botany) impulsó el desarrollo de un método unificado, desarrollado por Louveaux y otros en el año 1978 y que aun es referencia en múltiples laboratorios para analizar la calidad de la miel y garantizar su origen geográfico mediante el análisis espectral [7].

A día de hoy la tarea de identificación de estos pólenes se realiza mayoritariamente de manera visual. De manera inicial, aunque no garantiza una total precisión, los expertos pueden identificar los tipos predominantes de polen dentro de una muestra de miel atendiendo a atributos como el color del conjunto [8]. No obstante, muchos expertos recomiendan complementar esto con un análisis microscópico posterior para identificar correctamente su composición y origen [9] [10] Para ello es necesaria una inspección de diferentes imágenes obtenidas por microscopio, como la que se puede observar en la siguiente figura.



**Figura 1.1.:** Muestra de miel tomada con microscopio

Como se puede intuir a partir de la imagen, es una metodología que consume mucho tiempo y que además requiere de la dedicación de un profesional experimentado durante todo el proceso de clasificación. Dichos resultados, además, son la interpretación subjetiva del experto basada en su propio criterio. Es por esto que los métodos automáticos de identificación y clasificación de pólenes se vuelven necesarios para superar estas limitaciones en el procedimiento manual [11].

## 1.4. Reconocimiento automático de pólenes

Una de las primeras referencias que se encuentran sobre la necesidad de automatizar el recuento de pólenes procede de J.R. Flenley, en el año 1968 [12]. En esta obra, el autor destaca dos problemas que pueden dificultar esta tarea: que el polen puede tener infinitas orientaciones dentro de la imagen, y que algunos granos queden parcialmente enfocados en la imagen.

Siguiendo la estela de Flenley, H. J. L. White también ahonda en el problema de la obtención de imágenes adecuadas [13], ya que desarrollar un conjunto de criterios de clasificación que sea realmente eficaz antes de diseñar el sistema en sí no resulta tan trivial como podría parecerlo. Estos problemas son reseñados a día de hoy como barreras a superar incluso en publicaciones recientes como las de Rodríguez Damián y otros, que proponen modelos basados en diferentes tipos de análisis[14].

La mayoría de propuestas incluyen estudiar la imagen ateniéndose a sus atributos geométricos o de forma, de color, del dominio de frecuencia, y del estudio de texturas, siendo para ello utilizados métodos de análisis específicos y exhaustivos siendo después clasificados por sistemas de redes neuronales o máquinas de vector soporte. [15] [16]

Hay quórum entre todas las partes interesadas para continuar desarrollando un sistema que permita solventar este problema. En adición a estos problemas se añaden aspectos como la dificultad de recolectar datos suficientes y el hecho de que hasta hace relativamente poco tiempo las prestaciones de los equipos informáticos eran insuficientes para el reconocimiento de imágenes digitales. Incluso ahora el tiempo de cómputo necesario, aunque para nada prohibitivo, es elevado según qué circunstancias.

En definitiva, un conjunto adecuado de características, y una mejora en los procesos de obtención de información de las imágenes son la clave para obtener un sistema completo que finalmente automatice nuestro problema.

## 2. Estudio de atributos de imágenes

Este capítulo conforma la base teórica sobre la cual se realizará una batería de experimentos a la colección de imágenes microscópicas de miel que se nos ha proporcionado. Es aquí, pues, donde introduciremos los diferentes atributos que nos servirán para extraer información del procesamiento de las imágenes, es decir, el color y la textura. Ambos apartados contendrán información de cómo obtener información relevante de ellas, y plantearemos definiciones sobre cómo pueden ayudarnos en nuestro objetivo. En el último apartado de este capítulo describiremos cómo extraer los descriptores obtenidos de dichos atributos.

### 2.1. Color

El uso del color en el procesamiento de imágenes viene respaldado por dos factores importantes [17]. El primero es que el color es un descriptor muy potente que la mayoría de veces nos ayuda a identificar objetos dentro de un escenario, y el segundo es que el ser humano es capaz de distinguir entre miles de colores, teniendo en cuenta las intensidades y tonos, mientras que solo somos capaces de distinguir entre unas pocas decenas de niveles de gris. Este último punto obviamente atañe a los análisis supervisados por personas.

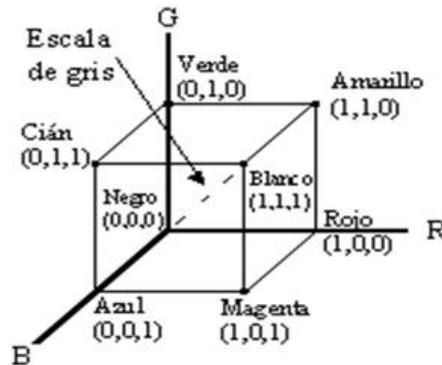
La percepción e interpretación del color por parte del ser humano es un proceso que todavía despierta debate y discusión, aunque sin embargo la naturaleza física del color sí que puede ser explicada formalmente en base a experimentos y modelos teóricos. Históricamente fue Isaac Newton el descubridor de que cuando un rayo de sol atraviesa un prisma de cristal surge por el otro extremo como un conjunto de colores en un espectro uniforme que va del rojo al violeta.

Este espectro de colores derivados de la luz visible se corresponde con una pequeña banda de frecuencias en el espectro electromagnético, y es el comportamiento de cada objeto a la hora de reflejar la luz en dichas frecuencias la que otorga la propiedad del color. Por poner un ejemplo, un objeto verde absorbe de la luz toda la energía del rango excluyendo principalmente aquellas entre 500 y 570 nm, que la refleja, siendo percibida por nuestros ojos con el color verde.

Para facilitar la catalogación de los colores se propone el concepto de espacios de color. Estos modelos son en esencia una especificación formada por un sistema de coordenadas y un subespacio dentro de ese sistema donde cada color se representa como un único punto. En este capítulo detallaremos los espacios RGB y CIE LAB.

### 2.1.1. RGB

En el modelo o espacio RGB se plantea un espacio de coordenadas cartesianas donde cada color aparece descompuesto en tres componentes primarios, consistentes en los colores rojo, verde y azul.



**Figura 2.1.:** Cubo RGB normalizado. (Fuente: Universidade da Coruña)

En este modelo, como se puede apreciar en la anterior figura, el color se define como puntos dentro del cubo definidos como vectores que surgen del origen de coordenadas. En este modelo la escala de grises se corresponde con un vector equidistante a los tres componentes.

Las imágenes representadas según este modelo están compuestas por una imagen o canal, para cada uno de los tres colores, que se combinan para ofrecer la información necesaria para cada píxel. Esta información puede ser más o menos completa dependiendo del rango de valores que toma cada color. Por ejemplo, si cada una de estas tres imágenes tiene un valor por píxel de 8 bits, el conjunto será una imagen con 24 bits por píxel, lo que nos da un rango de  $(2^8)^3 = 16,777,216$  colores.

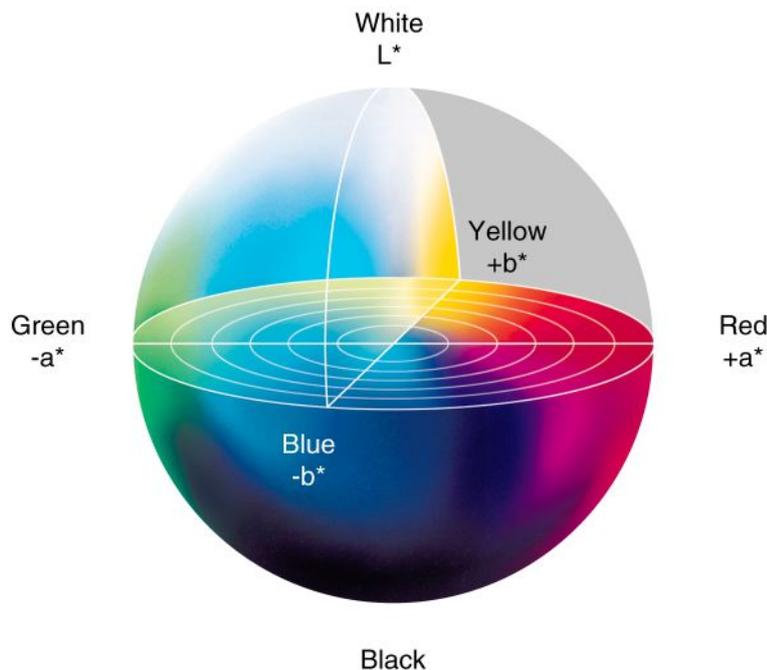
### 2.1.2. CIE LAB

La Comisión Internacional de Iluminación, CIE (del francés Commission Internationale de L'Eclairage) nos plantea un espacio de color perceptualmente uniforme que nos permite medir diferencias entre distintos puntos de una manera más aproximada a como lo hace el cerebro humano.

En particular, este espacio permite que cuando nos movemos de un color a otro, la diferencia perceptual va a estar relacionada con la distancia entre dichos puntos (particularmente la euclídea) a lo largo de todo el espacio, por lo que se postula como una alternativa mucho más uniforme que RGB para tareas donde el color tiene gran peso en el análisis del conjunto.

En este modelo se nos proponen tres componentes: L, A y B. L es un componente que representa la iluminación dentro de un rango de 0 a 100, que comprenden el

negro y el blanco como extremos. Los ejes A y B son opuestos, y vienen a representar los gradientes rojo-verde y amarillo-azul. Los límites de estos ejes dependerán de la implementación específica del espacio de color debido a procesos como la conversión de RGB a LAB (que requiere conversiones intermedias), pero los rangos más habituales son  $\pm 100$  o de  $-128$  a  $+127$ . Podemos visualizar el modelo como una esfera del que obtenemos un plano circular de colores con coordenadas A y B perpendicular a su eje, con coordenada L, de manera similar a la siguiente figura:



**Figura 2.2.:** Interpretación visual del espacio CIE LAB

Para convertir correctamente una imagen RGB a CIE LAB se necesita primero convertir la imagen al espacio CIE XYZ, del que se origina CIE LAB. El proceso es el siguiente [18]:

Conversión lineal de RGB a XYZ:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0,412453 & 0,357580 & 0,180423 \\ 0,212671 & 0,715160 & 0,072169 \\ 0,019334 & 0,119193 & 0,950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.1)$$

Conversión no lineal de XYZ a CIE LAB:

$$L = \begin{cases} 116(Y/Y_n)^{1/3} - 16 & \text{si } Y/Y_n > 0,008856 \\ 903,3(Y/Y_n) & \text{else} \end{cases} \quad (2.2)$$

$$a = 500(f(X/X_n) - f(Y/Y_n)) \quad (2.3)$$

$$b = 200(f(Y/Y_n) - f(Z/Z_n)) \quad (2.4)$$

donde

$$f(t) = \begin{cases} t^{1/3} & \text{si } Y/Y_n > 0,008856 \\ 7,787t + 16/116 & \text{else} \end{cases} \quad (2.5)$$

$X_n, Y_n, Z_n$  son los valores de X, Y y Z para el “iluminante”, punto blanco de referencia. Si se sigue la recomendación ITU-R BT.709 (International Telecommunication Union - Recommendation), este iluminante debe ser el D65, con los valores  $[X_n Y_n Z_n] = [0.95045 \ 1 \ 1.088754]$ [19].

## 2.2. Textura

La textura es una propiedad inherente a la superficie de los objetos. Puesto que los usos de las texturas como discriminantes surgen de la habilidad innata de la percepción humana es un área muy interesante para la Visión por Computador. Podemos definir el término, cuando hace referencia a una imagen, como la cuantificación de las variaciones de tono dentro de una región de interés continua en la imagen. En el procesamiento de imágenes se considera que las texturas contienen dos elementos principales: El primero, que las texturas están formadas por componentes simples llamados primitivas de texturas; El segundo, que la textura está definida por la relación espacial entre dichas primitivas.

Históricamente se han propuesto diversos atributos de textura basados entre otros en estadística de primer orden o globales (como la media, desviación estándar, histograma), o estadísticas de segundo orden, como las obtenidas a partir de las matrices de co-ocurrencia [20]. La hipótesis subyacente en estas matrices es que en una imagen la relación espacial que los tonos tienen entre ellos almacena cierta información de la textura [21] y son los métodos más populares de obtener información a partir de ellas.

Otras propuestas incluyen el uso de histogramas de textura para encontrar patrones, como el LBP (Local Binary Pattern) [22]; modelos estocásticos, como los campos aleatorios de Markov o de Gibbs, [21] o medir su dimensión fractal [23], entre otras muchas.



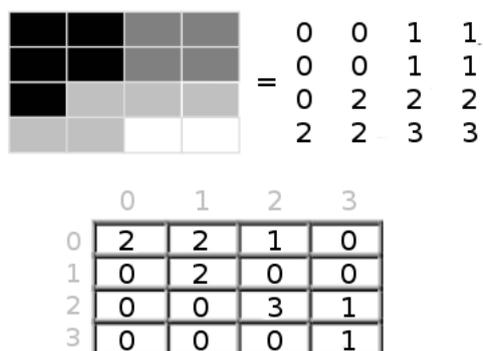
**Figura 2.3.:** Diversas texturas. (Fuente: Subdimension studios)

En los próximos apartados presentaremos algunos de estos atributos, puesto que partiremos de dichas propuestas para realizar nuestro trabajo.

### 2.2.1. Matrices de co-ocurrencia

En los años 70 Robert Haralick propuso un conjunto de 14 medidas de textura basadas en la dependencia espacial de los tonos de grises [24]. Para ello hace uso matemáticamente de las matrices de co-ocurrencia dentro de los niveles de gris.

La matriz de co-ocurrencia (GLCM a partir de ahora) describe la frecuencia con la que un nivel de gris aparece, según una relación espacial determinada, junto a otro valor de gris, dentro de una región de interés llamada ventana. La GLCM es, pues, un resumen de la forma en que los píxeles de determinado valor están situados con respecto a otros píxeles de otros valores.



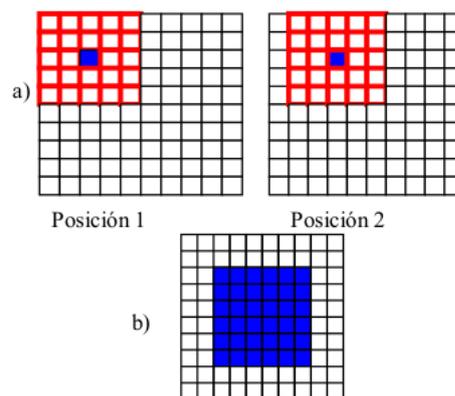
**Figura 2.4.:** GLCM derivada de una imagen de 4x4 píxeles con 4 niveles de gris. (Fuente: [25])

Como podemos ver en la figura 2.4 (donde el valor 0 se corresponde con el negro y el 3 con el blanco), el valor 3 que se encuentra en la posición (2, 2) de la matriz inferior de la figura implica que en la imagen se encuentran 3 ocurrencias de píxeles con valor 2 situados a la distancia predeterminada (una posición a la derecha en nuestro caso) con respecto a un píxel con valor 2.

Hay que tener en cuenta que este procedimiento está parametrizado según unas cuantas variables distintas: La simetría, el tamaño de la ventana, la banda espectral de entrada y el número de bits de salida y la componente espacial del análisis, es decir, el ángulo y la distancia entre píxeles de interés.

La simetría en la matriz de co-ocurrencia está contemplada en la publicación de Haralick, por lo que es necesaria para el cálculo de esos atributos. Para obtener la simetría basta con computar la relación entre un píxel y su vecino al igual que la recíproca entre el vecino y el píxel original.

Respecto al tamaño de la ventana, esta debe ser cuadrada. El tamaño relativo de la ventana y de los objetos en la imagen determinan la utilidad de esta medida para la clasificación. Es recomendable que la ventana sea menor que el objeto y lo suficientemente grande como para capturar la variabilidad del mismo [25].

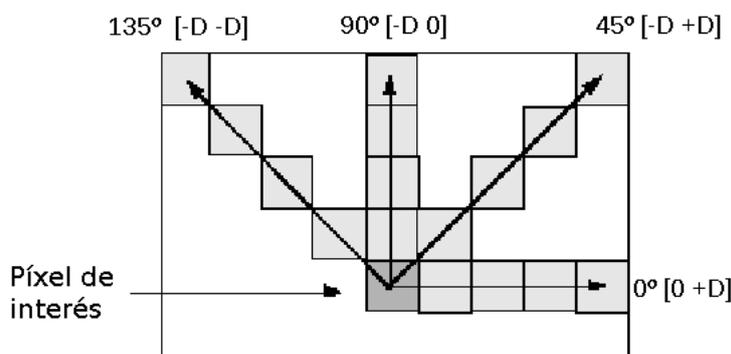


**Figura 2.5.:** En a) se muestra sobre una imagen de dimensiones 10x10 píxeles, una ventana móvil de 5x5 en dos posiciones y el píxel central que recibe el resultado. En b) los píxeles que reciben el resultado y las 2 filas y las 2 columnas del borde de la imagen que reciben el valor de textura calculados en los píxeles cercanos. (Fuente: [25])

Para dotar de un tamaño adecuado a la GLCM de salida, la banda espectral de entrada puede verse simplificada a menos valores para reducir la dispersión en sus datos. De esta manera sería como si la escala de grises tuviese menos valores posibles. El valor adecuado de la banda no es fácilmente discernible ya que implica pérdida de precisión, y requiere de ajustes según el tipo de trabajo al que está destinado. Por ejemplo, no es muy práctico trabajar con 256 niveles de gris cuando nuestra fuente

son imágenes de 100x100 píxeles, ya que un porcentaje mayoritario de los valores de dicha matriz será 0.

La relación espacial entre el píxel de referencia y su vecino puede ser en cualquiera de las 8 direcciones, pero puesto que trabajaremos con matrices simétricas solo usaremos cuatro. Estas direcciones son desplazamientos angulares de 0°, 45°, 90° y 135°, adicionalmente la distancia también puede variar para que no sea el objetivo el píxel inmediatamente adjunto.



**Figura 2.6.:** Posibles combinaciones de relación espacial en la GLCM. (Fuente: Mathworks)

A partir de los resultados de esta matriz y de la combinación de múltiples parámetros podemos extraer información interesante para nuestro objetivo, como veremos a continuación.

## 2.3. Descriptores

En este apartado introducimos cuáles son aquellos descriptores que utilizaremos en nuestro análisis y profundizaremos en su definición y características. La combinación de los resultados a la hora aplicar la extracción de estos descriptores será presentada en un capítulo aparte.

### 2.3.1. Descriptores suaves de color y textura

Estos descriptores se obtienen del estudio de los histogramas de valores de cada canal de color. Se describen como “suaves” porque recogen propiedades de color y textura sin recurrir a métodos más sofisticados, exhaustivos o específicos. [18]

En nuestro caso, puesto que de manera visible podemos discernir entre diferentes clases de pólenes ateniéndonos a diferencias evidentes de textura y color, vamos a aplicar estas técnicas con el objetivo de extraer información de cada canal de color

mediante la media y lo complementaremos con la desviación típica para medir como varía cada componente en toda la imagen.

$$\sigma(z) = \sqrt{\frac{\sum_{i=1}^L (z_i - m)^2}{L - 1}} \quad (2.6)$$

Tal que  $z$  sea la variable aleatoria,  $L$  el tamaño del conjunto de datos y  $m$  el valor medio de los valores de  $z$ .

Para obtener información de las texturas, realizaremos cálculos simples estadísticos. A partir del histograma de cada componente del color podemos calcular los momentos centrales para cada histograma (del 2º al 5º, por razones de eficiencia).

Podemos definir el momento  $n$  de  $z$  sobre su media como:

$$\mu_n(z) = \sum_{i=1}^L (z_i - m)^n p(z_i) \quad (2.7)$$

Tal que  $z$  es la variable aleatoria,  $p(z)$  el histograma,  $L$  el número de distintos valores de las variables y  $m$  el valor medio de  $z$ .

$$m = \sum_{i=1}^L z_i p(z_i) \quad (2.8)$$

Así pues, con este método para obtener descriptores, podemos generar un vector de 18 componentes como máximo: 6 por cada canal de color, la media, la desviación típica, y los momentos centrales del histograma del 2 al 5. La combinación de estos descriptores nos proporcionarán diversos resultados, como veremos en el capítulo de experimentación.

### 2.3.2. Descriptores de Haralick

En este apartado introducimos el conjunto de descriptores definidos por Haralick [24] y que son obtenidos a partir de la matriz de co-ocurrencia definida en el apartado anterior una vez ha sido normalizada para obtener la probabilidad de aparición en la imagen, es decir, que la suma de sus elementos es igual a 1.

Del total de los posibles descriptores que se pueden extraer de la manipulación estadística y de los momentos de la GLCM, Haralick propone 14 descriptores :

- Segundo Momento Angular:

$$f_1 = \sum_i \sum_j p(i, j)^2 \quad (2.9)$$

- Contraste:

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \right\}, |i - j| = n \quad (2.10)$$

- Correlación:

$$f_3 = \frac{\sum_i^{N_g} \sum_j^{N_g} (ij)p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (2.11)$$

donde las medias y desviaciones típicas se hacen sobre  $p_x$  y  $p_y$ , las funciones de densidad de la probabilidad parcial de x e y.

- Suma de cuadrados (Varianza):

$$f_4 = \sum_i \sum_j (i - \mu)^2 p(i, j) \quad (2.12)$$

- Momento de diferencia inversa (Homogeneidad):

$$f_5 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j) \quad (2.13)$$

- Media de la suma:

$$f_6 = \sum_{k=2}^{2N_g} k p_{x+y}(k) \quad (2.14)$$

donde x e y son las coordenadas de la entrada en la GLCM y  $p_{x+y}(i)$  la probabilidad de las coordenadas de la GLCM sumando x+y.

- Varianza de la suma:

$$f_7 = \sum_{k=2}^{2N_g} (k - f_6)^2 p_{x+y}(k) \quad (2.15)$$

- Entropía de la suma:

$$f_8 = - \sum_{k=2}^{2N_g} p_{x+y}(k) \log \{p_{x+y}(i)\} \quad (2.16)$$

- Entropía:

$$f_9 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log \{p(i, j)\} \quad (2.17)$$

- Diferencia de varianza:

$$f_{10} = \sum_{k=0}^{N_g-1} \left[ k - \sum_{l=0}^{N_g-1} l p_{x-y}(l) \right]^2 p_{x-y}(k) \quad (2.18)$$

- Diferencia de entropía:

$$f_{11} = - \sum_{k=0}^{N_g-1} p_{x-y}(k) \log \{p_{x-y}(k)\} \quad (2.19)$$

- Info. medida de correlación 1:

$$f_{12} = \frac{f_9 - HXY1}{\text{máx}(HX, HY)} \quad (2.20)$$

- Info. medida de correlación 2:

$$f_{13} = (1 - \exp[-2(HXY2 - f_9)])^{1/2} \quad (2.21)$$

- Coeficiente de correlación máxima:

$$f_{14} = (\text{segundo mayor eigenvalue de } Q)^{1/2} \quad (2.22)$$

No obstante, como sugieren algunos autores, es recomendable evaluar posibles subconjuntos e incluso otros como Soh y Tsatsoulis [26] proponen este conjunto de 10 descriptores:

- Segundo Momento Angular: Ecuación 2.9
- Contraste: Ecuación 2.10
- Correlación: Ecuación 2.11
- Homogeneidad: Ecuación 2.13
- Entropía: Ecuación 2.17
- Auto-correlación:

$$f_{15} = \sum_i \sum_j (ij) p(i, j) \quad (2.23)$$

- Disimilaridad:

$$f_{16} = \sum_i \sum_j |i - j| * p(i, j) \quad (2.24)$$

- Sombreado de agrupamientos:

$$f_{17} = \sum_i \sum_j (i + j - \mu_x - \mu_y)^3 p(i, j) \quad (2.25)$$

- Prominencia de agrupamientos:

$$f_{18} = \sum_i \sum_j (i + j - \mu_x - \mu_y)^4 p(i, j) \quad (2.26)$$

- Probabilidad Máxima:

$$f_{19} = \max_{i,j} p(i, j) \quad (2.27)$$

Como se puede apreciar, además de usar ciertos atributos de Haralick también se añaden dos parámetros de agrupamientos con el objetivo de emular el comportamiento de nuestra percepción [27].

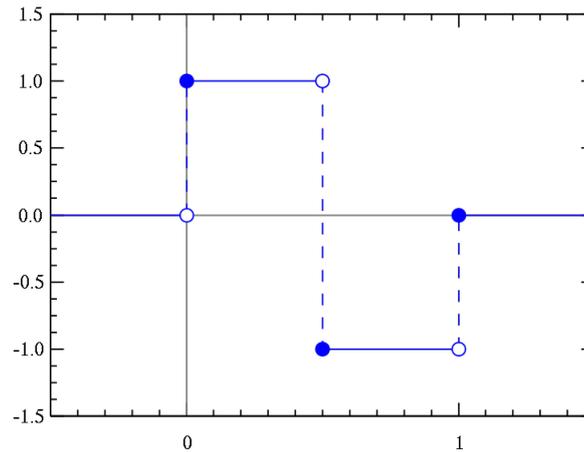
Estas variantes en el uso de descriptores nos ofrecen la oportunidad de poder configurar un conjunto a medida de nuestras necesidades y es que no solo podemos alternar el uso de éstos, si no que podemos combinar los provenientes de diversas GLCMs con configuraciones distintas.

Como veremos en el siguiente capítulo, donde se experimenta con estos atributos y con los obtenidos en el apartado anterior, la combinación de ambos es necesaria para evitar que se produzcan errores por sobreajuste.

### 2.3.3. Descriptores basados en wavelets

A diferencia de los anteriores descriptores, las wavelets existen en el dominio de las frecuencias, o lo que es lo mismo, reflejan su movimiento periódico respecto a su frecuencia, en contraposición al tiempo. Las wavelets nos permiten estudiar como cambia la señal o imagen con respecto a determinadas frecuencias, y al espacio (en el caso de las imágenes). Resumiendo: son funciones de tipo onda definidas tanto en el tiempo como en la frecuencia para describir una señal o imagen.

Su desarrollo se produce de manera temprana con el trabajo de Alfréd Haar a principios del siglo XX [28], pese a que el término “wavelet” no surgió hasta bastantes décadas después. Haar definió una serie de funciones con tal de conseguir un sistema ortonormal para el espacio de las funciones de cuadrado integrable en el intervalo  $[0, 1]$ , generando la ondícula más simple posible, conocida como wavelet de Haar.



**Figura 2.7.:** Wavelet de Haar

Ya en la década de los 70, George Zweig postula la Transformada Wavelet Continua estudiando las reacciones del oído al sonido.

Al igual que Zweig, Jean Morlet, Alex Grossmann y Pierre Goupillard se encuentran entre los miembros destacados que desarrollaron la teoría de las wavelets en esta época. En la posterior década encontramos a Yves Meyer, Stéphane Mallat, Ingrid Daubechies, Ronald Coifman, Ali Akansu o Victor Wickerhauser como investigadores destacados en este campo. Nos centraremos a continuación en el trabajo de Dennis Gabor dentro del campo de los wavelets por su directa aplicación en el campo del procesamiento de imágenes para el análisis de texturas.

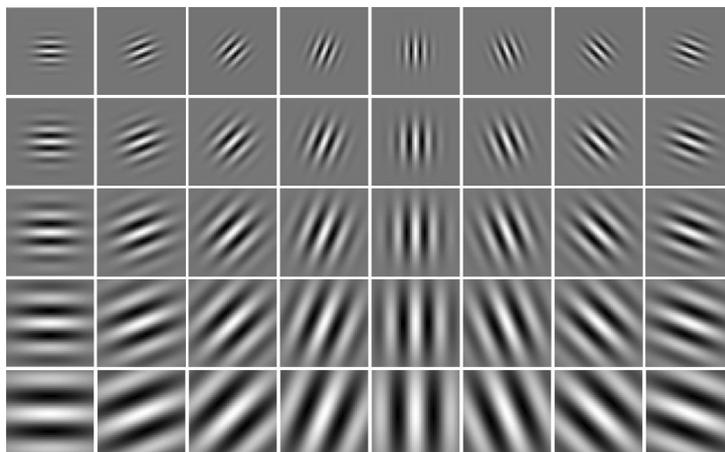
La transformada de Fourier, que nos proporciona una excelente resolución en el dominio de las frecuencias tiene la contrapartida de que nos hace perder toda información temporal (en el caso de señales) o espacial (en el caso de las imágenes). La transformada de Gabor contrarresta este problema estableciendo una ventana sobre las funciones base de la transformada de Fourier a la que se aplica un impulso gaussiano, de forma que la función resultante al ser transformada nos proporciona la relación tiempo-frecuencia. La ecuación de la transformada de Gabor es la siguiente:

$$G_x(t, f) = \int_{-\infty}^{\infty} e^{-2\Pi(\tau-t)^2} e^{-j2\Pi f\tau} x(\tau) d\tau \quad (2.28)$$

Las wavelets de Gabor, pues, tienen la particularidad de que minimizan la incertidumbre sobre la información que lleva la ondícula, ya que reducen la desviación típica en los dominios de frecuencia y tiempo.

Los filtros de Gabor bidimensionales, que están directamente relacionados con éstas últimas wavelets, son filtros lineales parametrizables según su orientación y escala.

Son utilizados de manera efectiva para la representación y discriminación de texturas, ya que el comportamiento de la percepción en el sistema visual humano se considera equiparable al análisis de imágenes a partir de los estos filtros [29, 30].



**Figura 2.8.:** Banco de filtros de Gabor de 4 escalas y 8 orientaciones

Una alternativa a los filtros de Gabor son los filtros log-Gabor propuestos por Field [31]. A diferencia de los filtros de Gabor normales, estos no tienen la limitación del máximo ancho de banda limitado a una octava y componente DC distinto de cero. [32]. Field, además, demuestra que estos filtros de base logarítmica son más similares al comportamiento de las células en la visión humana [31, 33].

Estos filtros están definidos en el dominio de las frecuencias como funciones gaussianas cambiadas con respecto al origen por la singularidad de la función logarítmica [34].

Matemáticamente se pueden representar en dos filtros: radial y angular:

$$\hat{G}(\rho, \theta) = \hat{G}_\rho \hat{G}_\theta = e^{-\frac{1}{2} \left[ \frac{\log\left(\frac{\rho}{u_0}\right)}{\log\left(\frac{\alpha_\rho}{u_0}\right)} \right]^2} e^{-\frac{1}{2} \left[ \frac{(\theta - \theta_0)}{\alpha_\theta} \right]^2} \quad (2.29)$$

donde  $\rho$  y  $\theta$  representan las coordenadas polares,  $u_0$  es la frecuencia central,  $\theta_0$  es el ángulo de orientación y los parámetros  $\alpha_\rho$  y  $\alpha_\theta$  determinan la escala y ancho de banda angular respectivamente [34].

El procedimiento suele consistir en generar un banco de filtros de Gabor o log-Gabor formado por filtros individuales de diversas escalas y orientaciones, que se convolucionan con la imagen. Las imágenes resultantes se consideran sitas en el “espacio de Gabor” y a partir de ellas se extraen los descriptores deseados para el análisis.

A partir de cálculos estadísticos simples sobre el histograma de las imágenes convolucionadas podemos obtener sus momentos estándar: media, desviación típica,

asimetría y curtosis.

De esta forma, un banco de 24 filtros (4 escalas y 6 orientaciones) como el propuesto por Nava [35] producirá un vector de 96 elementos para describir la textura de cada polen.

En nuestros experimentos vamos a plantear obtener esos mismos descriptores, y posteriormente veremos como afectan al estudio las diferentes combinaciones de los mismos.

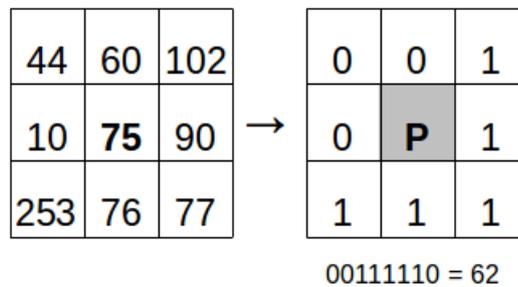
### 2.3.4. Patrones Binarios Locales (LBP)

Descritos inicialmente en 1994 [36, 37], los Patrones Binarios Locales, LBP a partir de ahora, son una herramienta habitual en la clasificación de texturas [34, 22].

Los LBP se pueden obtener de la siguiente forma:

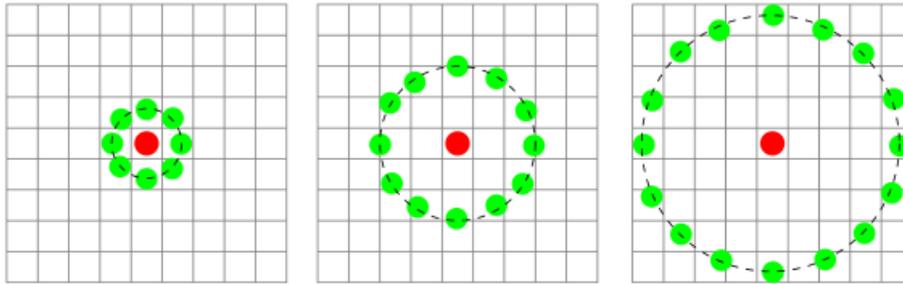
Para cada píxel  $P$  en la imagen, examinamos los píxeles de su entorno (8-vecindad) y para cada uno de estos 8 píxeles comparamos su valor con  $P$ , siguiendo un orden circular en sentido (favorable o en contra) de las agujas del reloj. Si el valor del píxel vecino es mayor o igual que  $P$  se le asigna un 1 y en el caso contrario un 0.

Una vez tenemos asignados los 8 píxeles vecinos, a partir de sus valores codificamos una cadena binaria siguiendo el mismo orden que anteriormente, lo que nos proporciona un número binario, que en decimal está comprendido en el rango  $[0, 255]$ . Si computamos todos los píxeles de la imagen, el resultado es una matriz de las mismas dimensiones de la imagen.



**Figura 2.9.:** LBP - Calculando el valor del píxel central  $P$

Nótese que en vez de usar los píxeles inmediatamente vecinos podría asignarse una distancia arbitraria. Esto implicaría una cantidad más elevada de píxeles vecinos, que generarían por lo tanto cadenas de valor para  $P$  en un rango  $[0, 2^n - 1]$ . Si computamos de esta forma todos los píxeles de la imagen, el resultado ya no será una matriz de valores representable como una imagen en escala de grises.



**Figura 2.10.:** LBP sobre diversas distancias (Fuente Wikimedia)

A continuación, de estas matrices obtenidas del proceso de analizar sus LBP se puede computar su histograma sobre la frecuencia de cada número. De este histograma podemos obtener descriptores como hemos hecho anteriormente: Momentos estándar obtenidos de cálculos estadísticos simples.



## 3. Experimentación

En este capítulo se explicarán los experimentos que se han llevado a cabo. Podemos distinguir dos bloques diferentes:

- Un primer bloque donde se realiza el primer intento de clasificar y segmentar la imagen de manera automática según un mapa de distancias ateniéndonos solo a la extracción de atributos mediante los descriptores suaves de color y textura definidas en el apartado 2.3.1. Aquí se analizarán los diferentes resultados de trabajar con CIE LAB o con RGB como espacios de color.
- Un segundo bloque donde, segmentaremos de manera manual y clasificaremos de manera automática. Para ello realizamos una base de datos documental de los diferentes tipos de pólenes de los que se dispone y se genera un modelo de clasificación del tipo k-NN, experimentando con las diferentes maneras de generarlo. Aquí trabajaremos con tres maneras distintas de clasificar la imagen: con los descriptores suaves de color y textura; con los descriptores de Haralick y los propuestos por Soh y Tsatsoulis para las matrices de co-ocurrencia; con la obtención de atributos mediante el filtrado log-Gabor, y finalmente con los patrones locales binarios.

A lo largo del capítulo explicaremos como se han construido los diferentes modelos y que resultados hemos obtenido de cada uno de ellos, así como el preprocesamiento que hemos tenido que realizar a los datos que se nos han proporcionado para poder trabajar con ellos.

Remarcamos que para la elaboración de estos experimentos se ha utilizado MATLAB.

### 3.1. Datos de los que disponemos

Para la realización de estos experimentos se nos proporciona un conjunto de 12 imágenes tomadas con microscopio de mieles con varios tipos de pólenes en ellas. Junto a estas se nos proporciona una clasificación manual de cada tipo de objeto en las imágenes realizada por personal experimentado en la materia. En total se nos presentan 18 categorías que pueden verse en la siguiente figura.

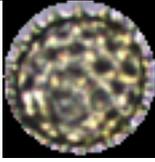
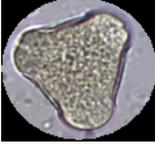
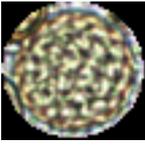
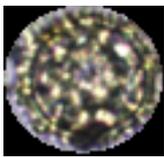
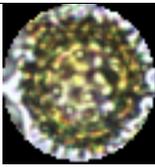
Tipo	Muestra	Tipo	Muestra
Impureza		Helianthus	
Prunus		Rosmarinus	
Polygonaceas		Diploaxis	
Eryngium		Erica	
Quercus sp.		Trifolium	
Tipo Genista		Lavandula	
Gota Glicerina		Taraxacum sp.	
Sedum		Thymus	
L. Stoechas		Compuestas	

Figura 3.1.: Tabla con las categorías de pólenes y objetos

Como se puede observar, no todas las categorías se corresponden con tipos de pólenes, ya que pueden aparecer distintos cuerpos en la imagen.

Las imágenes completas se encuentran en el Anexo A de este documento y están formateadas como mapa de bits, por lo tanto sin compresión y tienen una resolución media de 560x450 píxeles.

Una posterior catalogación de cada polen en las imágenes nos desvela la siguiente tabla de contenidos.

Tipo	Número de muestras	Ratio
2 - Helianthus	50	36.23 %
3 - Prunus	6	4.35 %
4 - Rosmarinus	11	7.97 %
5 - Polygonaceas	8	5.80 %
6 - Diplotaxis	7	5.07 %
7 - Eryngium	4	2.90 %
8 - Erica	13	9.42 %
9 - Quercus sp.	14	10.14 %
10 - Trifolium	1	0.72 %
11 - Tipo Genista	4	2.90 %
12 - Lavandula	2	1.45 %
14 - Taraxacum sp.	3	2.17 %
15 - Sedum	6	4.35 %
16 - Thymus	3	2.17 %
17 - L. Stoechas	1	0.72 %
18 - Compuestas	5	3.62 %

**Cuadro 3.1.:** Cantidad de muestras por especie y ratio sobre el total

Teniendo en cuenta esta tabla a simple vista se aprecia lo que puede resultar un problema de cara a los experimentos: hay categorías que están infrarrepresentadas donde apenas tenemos un único ejemplar.

Para la realización de estos experimentos se ha tenido en cuenta el conjunto completo de los pólenes, por lo que los resultados pueden verse afectados por el hecho de que algunas categorías no tengan los suficientes ejemplares.

## 3.2. Segmentación y clasificación automáticas

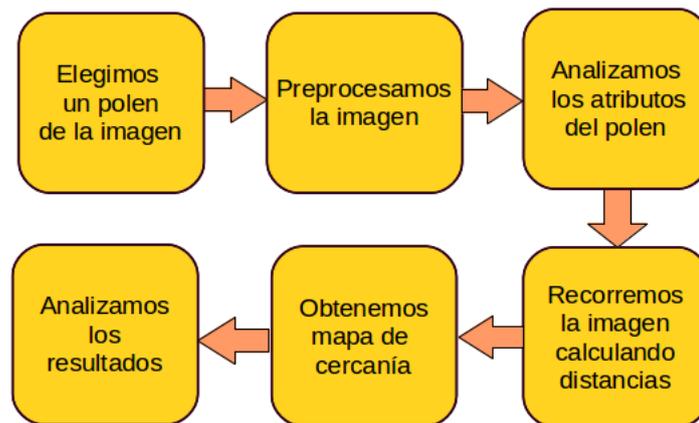
En este apartado se detallan los experimentos realizados mediante segmentación y clasificación automáticas con los atributos extraídos de los descriptores suaves de color y textura.

### 3.2.1. Experimentación sobre la imagen

El objetivo de este primer bloque de experimentos es obtener un mapa de cercanía realizado por un sistema supervisado de clasificación de pólenes que se base en sus descriptores suaves de color y textura como los ya descritos previamente, dada una imagen de miel (sobre los espacios RGB o CIE LAB comentados en el Capítulo 2) donde seleccionaremos manualmente un polen.

Para ello se diseñó un sistema que tuviese en cuenta estos requerimientos:

1. El usuario debe seleccionar un polen dentro de la imagen y a partir de ahí analizamos sus atributos y buscamos otros pólenes del mismo tipo dentro de la imagen.
2. El proceso de búsqueda de pólenes similares debe recorrer todos los píxeles de la imagen en cuestión, y esta tarea debe ser lo más eficiente posible
3. El mapa de distancias euclídeas de salida debe mostrar visiblemente las regiones candidatas a ser el centroide de un polen de la misma especie que el supervisado.



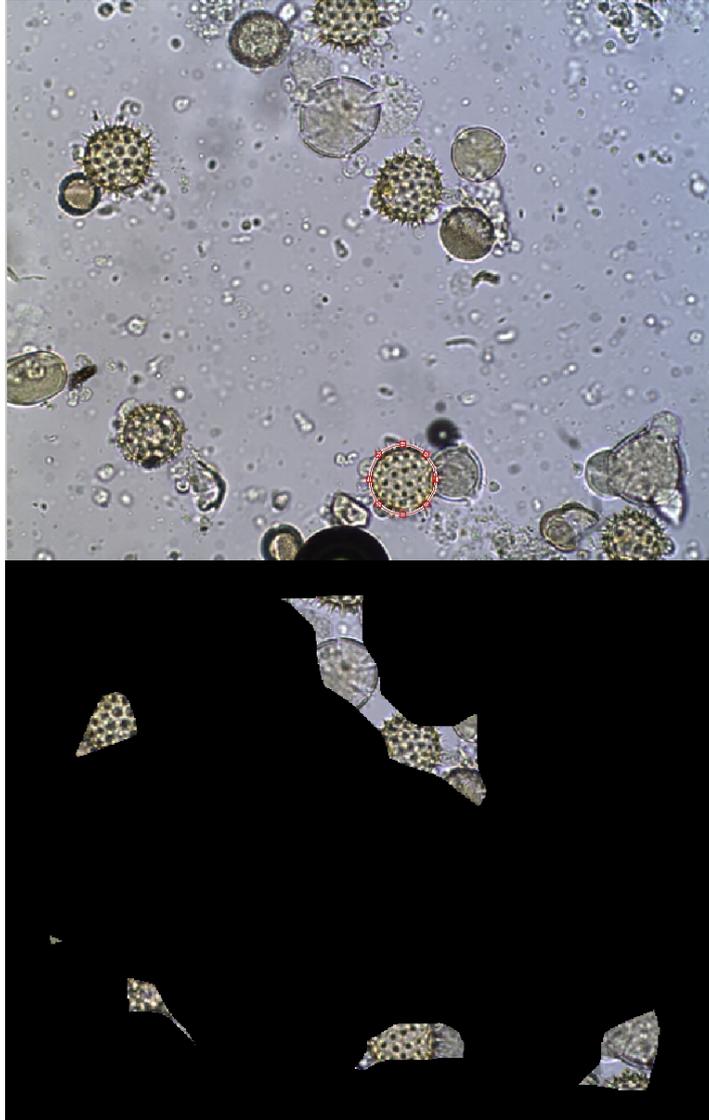
**Figura 3.2.:** Secuencia de pasos del experimento de segmentación y clasificación automática

De esta forma con identificar un polen de cada especie, el sistema identificaría los demás. A lo largo de esta sección veremos como se han ido cubriendo estas restricciones.

#### 3.2.1.1. Obtención de las imágenes y del polen muestra

Para este experimento vamos a trabajar con las imágenes de miel completas.

Ya que es el usuario quien debe seleccionar un polen como muestra, primero contemplaremos la necesidad de crear un procedimiento para seleccionar una región de interés (ROI, del inglés Region of Interest a partir de ahora). Para ello utilizaremos las herramientas que nos proporciona el toolkit de imagen digital de MATLAB y permitiremos al usuario crear una elipse sobre el polen que desee.



**Figura 3.3.:** Selección de polen (rodeado en rojo) y posterior delimitación de las zonas factibles.

Sobre esta selección, que internamente MATLAB renderiza como un cuadrilátero, analizaremos solo la región elíptica utilizando máscaras binarias para obtener su vector de descriptores suaves de color y textura.

En el momento que la selección está realizada, procedemos a post procesar la ima-

gen, de manera que descartemos todas las regiones que no puedan ser candidatas a centroide de un polen de la misma especie que el seleccionado. Para ello realizamos este procedimiento:

1. Aplicamos un algoritmo (desarrollado por Xose R. Fdez-Vidal del Depto. Física Aplicada de la Universidad de Santiago)[38] para la obtención de un mapa de saliencia a partir de la máxima alineación de fase utilizando filtros monogénicos. Este algoritmo aplica modelos de visualización similares a los de la percepción humana para destacar regiones de la imagen que captan nuestra atención por contener elementos sobresalientes en ellos. En las pruebas realizadas sobre las 12 imágenes de las que disponemos, el 100 % de los pólenes aptos para el análisis son preservados. De esta manera reducimos las regiones plausibles de contener un centroide de un polen en un rango entre el 30-80 % dependiendo de qué imagen de miel es la que se procesa.
2. A la imagen resultante de aplicarle el algoritmo de saliencia la sometemos a un proceso de erosión de las regiones donde no cabría un polen del tamaño seleccionado, acogiéndonos a su radio, lo cual reduce también significativamente el área factible, dependiendo, de nuevo, de la imagen tratada y del polen seleccionado.
3. Posteriormente, para eliminar las regiones próximas al borde, descartamos toda área que esté a menos distancia de los bordes que el radio del polen seleccionado. Esto se realiza para evitar errores, ya que no podemos garantizar la clasificación de un polen incompleto.

Nótese en la figura 3.3 que dada la selección del polen marcado en la imagen superior, las regiones factibles resultantes en la imagen inferior incluirán como poco todos los centroides de los pólenes que cumplen las condiciones previas: estar en áreas donde quepa el polen y no estar a una distancia menor del borde de la imagen al radio de nuestra selección.

### 3.2.1.2. Obtención de los atributos

Teniendo en cuenta que las imágenes proporcionadas se nos otorgan en el espacio de colores RGB, es necesario un paso intermedio en el caso de trabajar con CIE LAB para convertirlo a este espacio. En nuestro sistema utilizamos un algoritmo propuesto por Mark Ruzon en Matlab Exchange <sup>1</sup>que nos garantiza la aplicación de la recomendación BT.709 de ITU-R usando el punto blanco de referencia D65 y que nos proporciona un error entre conversiones RGB-LAB-RGB de menos del  $10^{-5}$ . Estas imágenes convertidas tendrán sus componentes A y B en el rango  $[-110, 110]$ <sup>2</sup>.

---

<sup>1</sup>Disponible en el momento de la redacción en MATLAB Central, apartado Exchange bajo el nombre “RGB2Lab.m”.

<sup>2</sup>Desde la versión r2014b MATLAB nos garantiza la misma funcionalidad mediante el método “rgb2lab”.

En cualquiera de los dos casos, se obtiene un vector de 6 elementos para cada canal de color que incluye:

- Media.
- Desviación típica.
- Segundo momento central.
- Tercer momento central.
- Cuarto momento central.
- Quinto momento central.

Para la obtención de estos elementos realizamos un algoritmo propio que tiene en cuenta que la ROI es elíptica, por lo que creamos nuestras propias herramientas de generar histogramas, acordes a la máscara binaria de la selección y para calcular la media y desviación típica de esta subregión.

A lo largo del experimento, se hacen combinaciones manuales de los distintos descriptores para probar su idoneidad, en el apartado de resultaremos comentaremos aquellos que han dado mejores resultados.

### 3.2.1.3. Barrido de la imagen

Gracias al preprocesamiento de la imagen ahora disponemos de una reducida área donde es factible que se encuentre el centroide del polen. En esta fase generaremos un mapa de cercanía píxel a píxel calculando la distancia entre la ROI seleccionada y una región similar llevada a ese punto del mapa.

A todos los efectos nuestro método recorre todos estos píxeles factibles y los considera como centroides de una región similar a la seleccionada por el usuario y de manera análoga a dicha región, se calcula su vector de descriptores suaves, con el añadido de que inmediatamente se calcula la distancia euclídea entre este vector y el calculado previamente.

$$d(v, w) = \sqrt{\sum_{i=1}^n (v_i - w_i)^2} \quad (3.1)$$

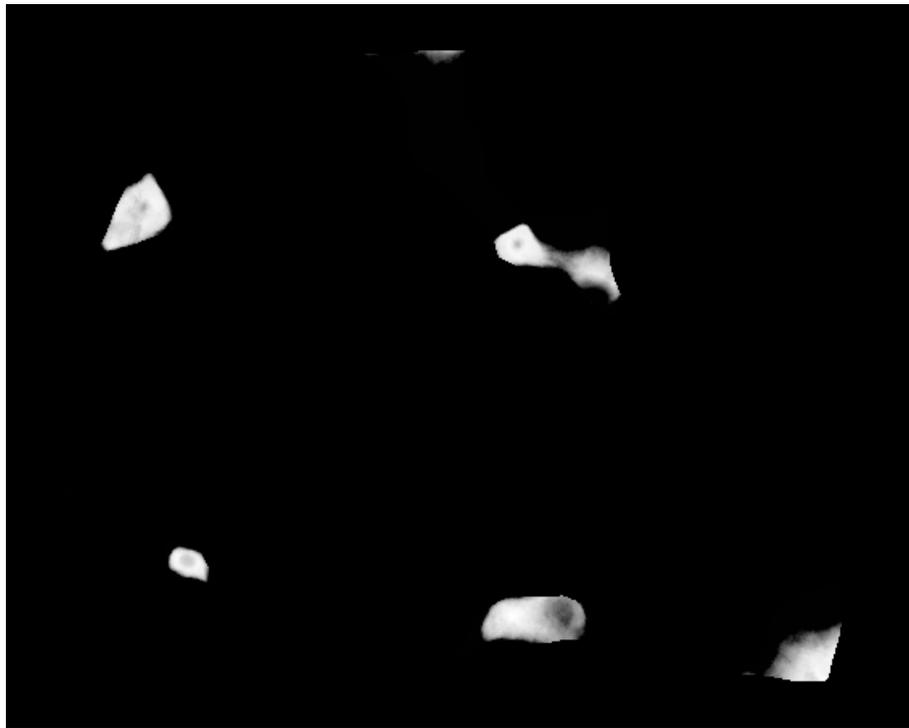
Tal que  $n$  es el número de parámetros seleccionados, típicamente 6 (para cada canal) por el número de descriptores suaves.

Una vez calculada la distancia, esta se almacena dentro de una matriz de las mismas dimensiones que la imagen original, en la posición correspondiente a su píxel centroide. De esta manera una vez se recorre toda la matriz tenemos un mapa con la cercanía a nuestra ROI.

Para presentar esta información de manera visual vamos a normalizar esta matriz en el rango 0-255 de forma que tratando a la matriz como una imagen de escala de grises, visualmente se aprecien las regiones candidatas a ser pólenes de la misma especie.

#### 3.2.1.4. Evaluación de resultados

Haciendo pruebas seleccionando distintos pólenes sobre las 12 imágenes y a partir de los mapas de cercanía, observamos a simple vista que, con independencia del tipo de polen escogido, el sistema es capaz de reconocer áreas donde efectivamente se encuentra un polen y en la abrumadora mayoría de casos se corresponden con regiones próximas al centroide, pero estos no son necesariamente de la misma especie.



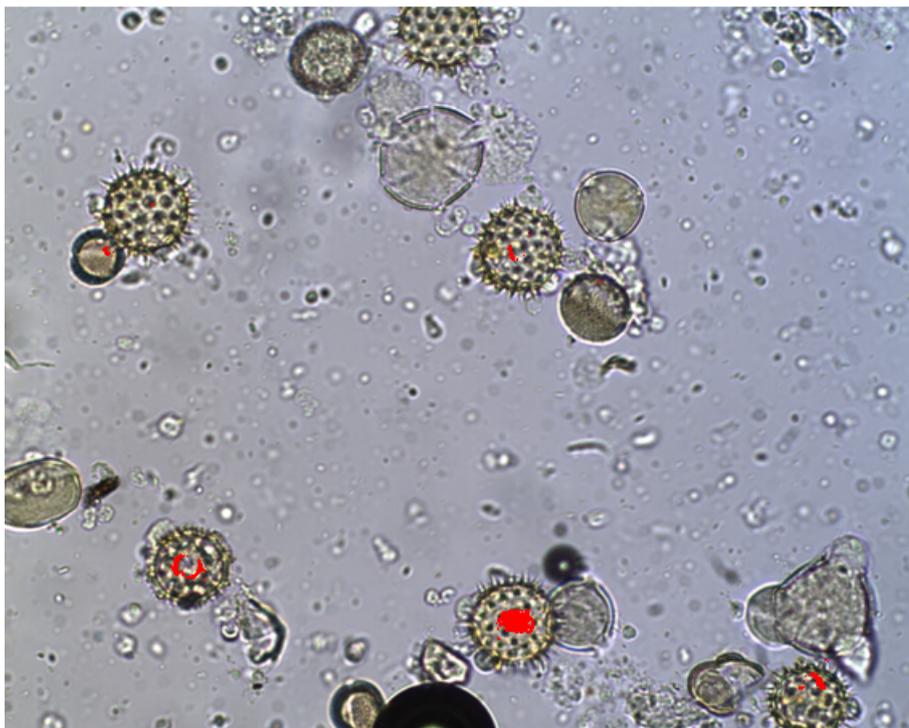
**Figura 3.4.:** Mapa de grises resultante de la selección de la figura anterior

No observamos diferencias significativas entre trabajar con RGB o CIE LAB, de la misma forma que observamos un comportamiento más coherente al trabajar solo con los atributos de media, desviación típica y segundo momento central.

En la figura 3.4 se puede observar esto. Las agrupaciones de puntos cercanos (blancos) se corresponden con áreas muy cercanas a los centroides de los pólenes.

Realizamos un método simple para ver donde se sitúan los X píxeles más cercanos a nuestra región y observamos que estos se agrupan de manera correcta en los centros

de algunos pólenes, con independencia de que sean aciertos o no. También observamos en algunos casos relativamente sencillos para la vista humana que algunos pólenes de la misma especie no están siquiera entre los 1000 puntos más cercanos.



**Figura 3.5.:** Situación de los 500 puntos más similares a nuestro polen (marcados en rojo).

Como observamos en la anterior figura, realizada a partir de los datos previamente presentados, la mayor agrupación de puntos cercanos se corresponde obviamente con el polen de tipo *Helianthus* seleccionado, pero la siguiente mayor agrupación está en torno al centroide de un polen de la especie *Taraxacum* sp. y hay un *Helianthus* que prácticamente no tiene ningún punto entre estos 500 primeros.

Esta situación se repite con mucha frecuencia en todas las imágenes: Los puntos más cercanos al 0 en nuestra matriz de distancias van a pertenecer a áreas céntricas pertenecientes a varios pólenes, pero sin diferenciar entre especies correctamente. Esto es especialmente evidente en imágenes con mayor acumulación de pólenes y seleccionando un tipo *Helianthus*, donde aparecen muchas regiones pertenecientes a otras especies y se omiten varios pólenes que sí que serían objetivos.

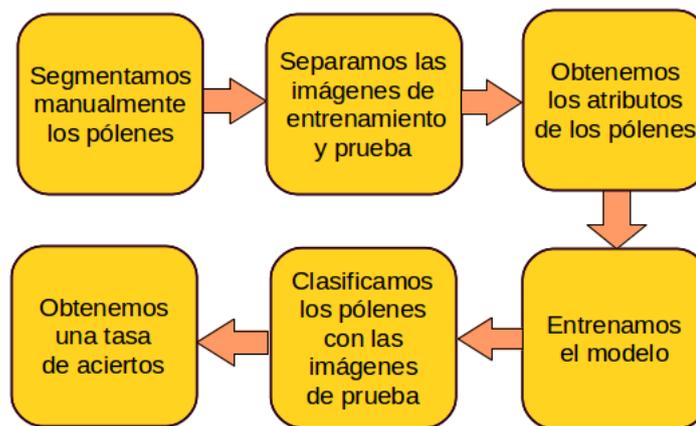
En conclusión, ni ajustando el espacio de colores ni la cantidad de atributos extraídos, el sistema es incapaz de clasificar correctamente los pólenes, teniendo además un comportamiento muy similar ambos espacios de colores en igualdad de atributos. Esto es de esperar, ya que, como se ha comentado en el capítulo 1.4 tanto la orientación de los pólenes, que siendo aun de la misma especie pueden tener multitud de

variaciones, como la captura de la imagen, que puede conllevar desenfoces, cambios de color, iluminación, etc. afectan directamente a estos descriptores suaves.

### 3.3. Segmentación manual y clasificación automática

Replanteamos el problema al no obtener buenos resultados con la aproximación anterior. En esta ocasión vamos a partir de una segmentación manual previa de los pólenes y realizaremos un modelo de clasificación automática que, mediante aproximación a los  $k$ -vecinos más cercanos nos verifique que los atributos obtenidos por los descriptores suaves de color y textura, así como los obtenidos del estudio de las matrices de co-ocurrencia y los wavelets son suficientes para realizar esta labor de clasificación.

La estructura del experimento va a ser constante a lo largo de este apartado, y se resume en el siguiente diagrama:



**Figura 3.6.:** Secuencia de pasos de los experimentos de segmentación manual y clasificación automática

#### 3.3.1. Usando descriptores suaves de color y textura

Este experimento debe tener en cuenta las siguientes circunstancias:

1. Debe realizarse una segmentación manual, por lo que todas las imágenes de los pólenes deben ser exploradas para obtener muestras lo más ajustada y uniformemente posible, que después serán clasificadas según su especie.
2. Se entrenará un modelo de clasificación de vecinos cercanos a tenor de estos parámetros: el número de vecinos cercanos ( $k$ ) y el porcentaje de imágenes que entrenarán a nuestro modelo y el porcentaje que se destinará a testearlo.

3. Tanto en RGB como en CIE LAB, se analizarán todas las imágenes de los pólenes de cada categoría destinados a entrenamiento y se obtendrá una colección de atributos global para esa especie, que quedará etiquetada.

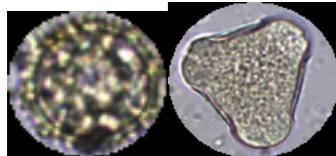
En esta sección se detallará como se han ido solventando estas necesidades.

#### 3.3.1.1. Preprocesado de las imágenes y elaboración de la BBDD

Vamos a divergir en este experimento del procedimiento utilizado en el experimento anterior. En este caso, como queremos verificar la técnica de clasificación, vamos a ignorar las imágenes proporcionadas como conjunto y vamos a hacer uso solamente de los pólenes, que extraeremos previamente.

Mediante un programa simple desarrollado en MATLAB extraemos, mediante elipses relativamente ajustadas a la superficie del polen, nuevas imágenes con tan solo el polen. Por el funcionamiento del formato de mapa de bits, estas imágenes son cuadriláteros con un borde negro alrededor de dicha elipse.

A la hora de analizar los pólenes se tendrá en cuenta esta peculiaridad. de la misma forma se aprecia que la irregularidad de algunas especies de pólenes puede suponer que la muestra extraída contenga más información visual del fondo de la que sería deseable. No obstante, debido a que solo ocurre en dos categorías (Prunus y Eryngium en algún ejemplo específico) se decide seguir con este procedimiento. Esto se aprecia en la figura 3.7, donde vemos que una parte significativa de la muestra derecha (de la especie Prunus) no pertenece al polen en sí.



**Figura 3.7.:** Ejemplos de Taraxacum sp. y Prunus extraídos de las imágenes.

La base de datos de los pólenes queda compuesta tal cual se define en este capítulo, en el apartado 3.1, teniendo dos colecciones, una en color (RGB) y otra en escala de grises. Tras un proceso de catalogación manual y posterior revisión, verificamos que ambas colecciones servirán también para el posterior experimento.

#### 3.3.1.2. Obtención de los atributos

Al igual que ocurría en el experimento anterior, se realiza una conversión de RGB a CIE LAB cuando se trabaja en este espacio de color. Dicha conversión se realiza de manera análoga, por lo que no se discutirá de nuevo en este apartado.

Para los tres casos (escala de grises, RGB y CIE LAB), se obtiene un vector de 6 elementos por cada canal de color.

Para la obtención de estos elementos utilizaremos los algoritmos del anterior experimento. Ya que estos tienen en cuenta que la ROI es elíptica, para que funcionen correctamente vamos a implementar un pequeño método que umbralice a partir de la imagen del polen y genere una máscara binaria para delimitar el área que nos interesa.

A lo largo del experimento se hacen combinaciones de los distintos estadísticos para probar su idoneidad. En el apartado de resultados comentaremos aquellos que han dado mejores resultados.

### 3.3.1.3. Generando el modelo

Los modelos basados en k-NN o k-vecinos más cercanos son unos de los sistemas de aprendizaje automático más simples. El modelo recibe como entrada un conjunto de muestras de entrenamiento, en nuestro caso vectores de atributos con una etiqueta de clasificación, que son tratados como puntos dentro de un espacio multidimensional. Este espacio queda poblado por instancias de diferentes clases, y cuando se le presenta un objeto a clasificar, el modelo le otorga una clase con respecto a la clase mayoritaria de entre sus k-vecinos más cercanos. Es decir, cuando  $k=3$ , se buscan las 3 instancias de pólenes más cercanas a nuestro polen a clasificar y se clasifica de acorde con la clase mayoritaria.

Como se puede comprobar el número k de vecinos es un parámetro clave en el proceso y dependiendo de su valor puede variar mucho la tasa de aciertos, sobre todo, al tener muy pocos ejemplos de algunas categorías de pólenes. Por ejemplo, si en la categoría *Taraxacum sp.* solo tenemos 3 imágenes, de las cuales solo 2 se usan para entrenamiento, una  $k=4$  no permitirá que se lleguen a clasificar pólenes con esa especie porque saldrán vecinos de otras categorías.

De acuerdo con la tabla resumen de nuestra BBDD de pólenes, las siguientes categorías quedan excluidas directamente por la razón que las acompaña:

- (clase 10) - *Trifolium*: por tener tan solo una imagen.
- (clase 17) - *L.Stocheas*: por tener tan solo una imagen.
- (clase 18) - Compuestas: al no ser una especie de pólenes como tal no procede modelizar a partir de ellas.

Posteriormente veremos que otras categorías con pocas imágenes deben ser descartadas para el correcto funcionamiento del modelo, ya que k será mayor que el número de muestras disponibles en el conjunto de entrenamiento para la clase de polen.

Para evitar este problema vamos a generar varios modelos, de acorde a la proporción de imágenes de cada categoría que se destinarán a entrenamiento y a probar el modelo. Los procedimientos se explican a continuación.

Un inconveniente importante que vamos a encontrar a lo largo del estudio es que al trabajar con diferentes descriptores obtenidos de diversos atributos de la imagen, los

valores de dichos descriptores pueden variar muchísimo entre sí: Los momentos centrales de mayor grado son de varios órdenes de magnitud superiores a, por ejemplo, la media o la desviación típica. Esto implica que al trabajar con un clasificador como el nuestro, estos descriptores de magnitud superior cobren mucha más relevancia que los más pequeños. El estudio se realiza con conocimiento de esta problemática al no tener peso los discriminantes.

#### **Particiones variables**

En el modelo de particiones variables vamos a testear 3 porcentajes distintos de muestras de entrenamiento sobre el total que nos servirá para crear el modelo. Dichos porcentajes son 50 %, 70 %, y 80 %. Para todos los porcentajes se realizarán modelos k-NN con k=1, 2, 3, 4 y 5.

Las diferencias en el número de muestras de las categorías conllevan que variar el porcentaje, por ejemplo del 70 % al 80 %, afecte principalmente a las categorías con más muestras. Nótese que si disponemos sólo de 3 muestras, con ambos porcentajes se destinarán 2 imágenes a entrenamiento.

Como hemos comentado previamente, las clases con menos muestras no podrán ser clase mayoritaria para los valores de k más altos, así que se experimentará también excluyendo aquellas categorías que no reúnan un número mínimo de muestras. De esta forma tendríamos que:

- Exigiendo un mínimo de 4 imágenes se excluyen las siguientes categorías: Lavandula (clase nº12), Taraxacum sp. (clase nº14), Thymus (clase nº16).
- Exigiendo un mínimo de 6 imágenes se excluyen las siguientes categorías: Eryngium (clase nº7), Tipo Genista (clase nº11), Lavandula (clase nº12), Taraxacum sp. (clase nº14), Thymus (clase nº16).

#### **Leave-One-Out**

Leave-One-Out es un caso especial de partición de elementos. En este caso se genera el modelo a partir de todos los ejemplos de los que disponemos excepto del que vamos a clasificar. Es decir, que si queremos clasificar una de las 136 muestras de polen de las que disponemos de categorías con más de 1 muestra, nuestro modelo se entrenaría con 135 imágenes y se utilizaría para clasificar la restante.

De esta manera clasificamos todos los pólenes de nuestro subconjunto con modelos realizados a partir de todos los demás. De nuevo la k será un factor determinante del rendimiento del modelo, ya que algunas categorías estarán muy poco representadas.

Al igual que en particionado variable, también se experimentará con retirar las categorías menos pobladas para ver el comportamiento con las k de valor más alto.

#### **3.3.1.4. Evaluación de resultados**

Para evaluar el resultado de aplicar el modelo se presenta una tasa de aciertos global.

La combinación de estadísticos que se ha utilizado para los presentes resultados es la que mejores resultados ha ofrecido, que coincide con la que comentábamos en el apartado anterior: Media, Desviación típica, Segundo Momento Central.

En los posteriores subapartados se ofrecen algunos resultados resumidos a partir de las combinaciones más relevantes de particiones y umbral de imágenes mínimas por categoría, con respecto al espacio de color que estemos utilizando. En el Anexo B se encuentran todos los resultados de la batería de pruebas realizadas al completo.

En conclusión podemos destacar que a la vista de los resultados obtenidos no podemos contar con que nuestro modelo sea lo suficientemente adecuado para clasificar los pólenes que se nos han proporcionado. Como alternativa a este modelo, plantearemos más adelante la necesidad de contar con algún atributo extraído de forma más específica, como los descriptores de textura.

### Escala de grises

Para categorías con 2 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Partición 50 %	49.51 %	52.38 %	50.79 %	49.21 %	52.38 %
Partición 70 %	52.78 %	50.00 %	52.78 %	<b>58.33 %</b>	<b>58.33 %</b>
Partición 80 %	48.15 %	51.85 %	44.44 %	51.85 %	51.85 %
Leave-One-Out	56.49 %	49.62 %	<b>59.54 %</b>	58.02 %	61.07 %

Para categorías con 4 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Partición 50 %	53.33 %	56.67 %	58.33 %	53.33 %	56.67 %
Partición 70 %	60.61 %	60.61 %	<b>63.64 %</b>	<b>63.64 %</b>	60.61 %
Partición 80 %	58.33 %	58.33 %	54.17 %	58.33 %	58.33 %
Leave-One-Out	61.79 %	60.16 %	<b>65.04 %</b>	<b>64.23 %</b>	<b>65.85 %</b>

Para categorías con 6 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Partición 50 %	58.93 %	62.50 %	58.93 %	58.93 %	60.71 %
Partición 70 %	67.74 %	<b>70.97 %</b>	<b>70.97 %</b>	67.74 %	67.74 %
Partición 80 %	65.00 %	65.00 %	65.00 %	70.00 %	70.00 %
Leave-One-Out	65.22 %	68.70 %	69.57 %	<b>71.30 %</b>	<b>70.43 %</b>

Como podemos comprobar, la tasa de aciertos es significativamente baja como para que el sistema sea lo suficientemente satisfactorio. La hipótesis planteada en el

apartado anterior respecto a si afectan las categorías con poca representación a este modelo (y por lo tanto a la tasa de aciertos) es correcta, ya que podemos observar como se incrementan los aciertos conforme dichas categorías desaparecen del modelo.

También podemos observar como el modelo basado en Leave-One-Out obtiene mejores resultados que las otras versiones para prácticamente toda k, excepto en la última tabla.

#### RGB

Para categorías con 2 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Partición 50 %	<b>65.08 %</b>	<b>65.08 %</b>	60.32 %	61.90 %	63.49 %
Partición 70 %	55.56 %	55.56 %	61.11 %	58.33 %	58.33 %
Partición 80 %	59.26 %	<b>66.67 %</b>	51.85 %	48.15 %	48.15 %
Leave-One-Out	64.12 %	54.20 %	<b>64.89 %</b>	62.60 %	60.31 %

Para categorías con 4 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Partición 50 %	<b>68.33 %</b>	65.00 %	63.33 %	65.00 %	65.00 %
Partición 70 %	60.61 %	57.58 %	63.64 %	60.61 %	60.61 %
Partición 80 %	<b>66.67 %</b>	62.50 %	54.17 %	50.00 %	54.17 %
Leave-One-Out	<b>68.29 %</b>	60.98 %	<b>67.48 %</b>	65.04 %	65.04 %

Para categorías con 6 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Partición 50 %	<b>71.43 %</b>	<b>71.43 %</b>	<b>71.43 %</b>	69.64 %	69.64 %
Partición 70 %	61.29 %	61.29 %	64.52 %	64.52 %	64.52 %
Partición 80 %	70.00 %	70.00 %	70.00 %	70.00 %	65.00 %
Leave-One-Out	<b>71.30 %</b>	70.43 %	70.43 %	69.57 %	70.43 %

Podemos observar que con el añadido del color el sistema mejora su precisión en la mayoría de situaciones, aunque no lo suficiente.

A diferencia de lo ocurrido en los experimentos sobre la escala de grises, aquí vemos que los mejores resultados se obtienen con las k más pequeñas. En los restantes casos podemos observar como al aumentar el tamaño de los porcentajes destinados a entrenamiento mejoran los resultados.

**CIE LAB**

Para categorías con 2 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Partición 50 %	52.38 %	<b>58.73 %</b>	57.14 %	53.97 %	52.38 %
Partición 70 %	47.22 %	47.22 %	47.22 %	47.22 %	47.22 %
Partición 80 %	44.44 %	44.44 %	37.04 %	40.74 %	33.33 %
Leave-One-Out	53.44 %	47.33 %	<b>58.02 %</b>	54.96 %	<b>59.54 %</b>

Para categorías con 4 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Partición 50 %	55.00 %	<b>60.00 %</b>	<b>60.00 %</b>	56.67 %	55.00 %
Partición 70 %	51.52 %	48.48 %	48.48 %	48.48 %	54.55 %
Partición 80 %	50.00 %	45.83 %	37.50 %	41.67 %	37.50 %
Leave-One-Out	59.35 %	56.91 %	<b>63.41 %</b>	<b>61.79 %</b>	<b>64.23 %</b>

Para categorías con 6 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Partición 50 %	58.93 %	64.29 %	62.50 %	60.71 %	58.93 %
Partición 70 %	54.84 %	54.84 %	51.61 %	58.06 %	54.84 %
Partición 80 %	60.00 %	60.00 %	45.00 %	55.00 %	45.00 %
Leave-One-Out	<b>64.35 %</b>	<b>67.83 %</b>	<b>69.57 %</b>	<b>66.96 %</b>	<b>66.96 %</b>

Lo primero que llama la atención de estos resultados es el descenso en la tasa de aciertos con respecto a los resultados obtenidos en RGB, siendo más similares a los de la escala de grises.

Podemos concluir de la misma manera que en el anterior caso, que los mejores resultados se obtienen cuando retiramos los conjuntos con menos muestras, puesto que perjudican a las k superiores.

### 3.3.2. Usando matrices de co-ocurrencia

En este apartado vamos a trabajar con los descriptores de textura obtenidos de la matriz de co-ocurrencia.

Este experimento tendrá características muy similares al anterior descrito, incluyendo su evaluación mediante tasa de aciertos, aunque hay que destacar las siguientes diferencias:

1. Solo trabajaremos en escala de grises.
2. Los atributos se extraerán de una matriz de co-ocurrencia derivada de la imagen.

En esta sección se detallará como se ha formulado este experimento.

#### 3.3.2.1. Preprocesado de las imágenes y elaboración de la BBDD

Al igual que en el experimento anterior, se utiliza la misma BBDD de imágenes en escala de grises para garantizar que los resultados sean coherentes. Por lo tanto, esta BBDD de los pólenes queda compuesta tal cual se define en este capítulo, en el apartado 3.1.

Partiendo de los resultados obtenidos anteriormente podemos extrapolar que los mejores resultados se darán con los conjuntos que más imágenes tienen (concretamente los conjuntos que superen las 6 muestras en total), por lo que también estudiaremos como afecta esta preselección de categorías según su número de imágenes a nuestro modelo.

#### 3.3.2.2. Obtención de los atributos

De la misma forma que anteriormente trabajaremos para extraer los descriptores a partir de la información de los niveles de gris del mapa de bits y también utilizaremos la misma técnica para trabajar solo con la región de nuestro interés.

Como se ha establecido en el capítulo 2.2.1 vamos a obtener una matriz de co-ocurrencia de los píxeles de la imagen para extraer de ella un conjunto de descriptores, un total de 22. De estos descriptores seleccionaremos los subconjuntos propuestos por Haralick y por Soh y Tsatsoulis, que se han descrito en el capítulo 2.3.2.

Las matrices de co-ocurrencia, recordemos, tienen diversos parámetros que las hacen únicas, entre ellos se encuentra el número de niveles de gris que definen la salida. Puesto que las imágenes de pólenes son considerablemente pequeñas hemos de reducir las dimensiones de la salida para evitar que la matriz resultante sea muy dispersa. Siguiendo las recomendaciones de Soh y Tsatsoulis utilizaremos 8 niveles de gris en nuestros experimentos.

Adicionalmente para garantizar que las diferencias de iluminación o enfoque entre muestras no perjudiquen al modelo consideramos combinar varias veces los subconjuntos de atributos, pero obteniéndolos de matrices de co-ocurrencia realizadas con diferentes parámetros. Así pues consideramos un total de:

- 4 matrices con distancia  $d=1$ : relaciones en  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  y  $135^\circ$
- 4 matrices con distancia  $d=2$ : relaciones en  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  y  $135^\circ$

- 4 matrices con distancia  $d=3$ : relaciones en  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  y  $135^\circ$

Esto nos da un total de 12 matrices que nos proveerán de 156 descriptores en el caso del subconjunto de Haralick o de 120 descriptores en el caso de Soh y Tsatsoulis.

Tendremos que estudiar si el uso de tantos descriptores perjudica a nuestro modelo (la maldición de la dimensionalidad) así que se realizarán pruebas al respecto reduciendo el número de distancias usadas.

A pesar del aumento significativo de atributos el tiempo de cálculo es similar al de apartados anteriores.

### 3.3.2.3. Generando el modelo

Este experimento hereda las características definidas en el apartado anterior, ya que a efectos prácticos es el mismo procedimiento pero con distintos atributos. Así pues, el modelo cambiará según la cantidad de imágenes destinadas a entrenamiento y a test, según la  $k$  y según la combinación de atributos elegidos.

De nuevo se diferenciará entre los experimentos con particiones variables (50 %, 70 % y 80 %) y los Leave-One-Out.

### 3.3.2.4. Evaluación de resultados

En los posteriores subapartados separaremos los resultados según el subconjunto elegido, basados en los estudios de Haralick y de Soh y Tsatsoulis, y ofreceremos conclusiones teniendo en cuenta las combinaciones más relevantes de particiones y número de imágenes mínimas por categoría.

En conclusión observamos que la tasa de aciertos se reduce drásticamente cuando analizamos por sí solos estos descriptores en comparación con los descriptores suaves de color y textura.

**Descriptores de Haralick** Los resultados que se encuentran en este apartado se obtienen de la aplicación de un conjunto de descriptores similar al propuesto por Haralick, y que se ha discutido en el apartado 2.3.2.

Para categorías con 2 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part. 50 %	30.00 %	20.00 %	17.50 %	27.50 %	25.00 %
Part. 70 %	26.09 %	21.74 %	26.09 %	21.74 %	21.74 %
Part. 80 %	21.05 %	15.79 %	21.05 %	21.05 %	<b>31.78 %</b>
Leave-One-Out	30.12 %	27.71 %	<b>32.53 %</b>	30.12 %	<b>33.73 %</b>

Para categorías con 4 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	32.43 %	29.73 %	24.32 %	<b>37.84 %</b>	29.73 %
Part.70 %	30.00 %	30.00 %	35.00 %	25.00 %	25.00 %
Part.80 %	25.00 %	25.00 %	31.25 %	25.00 %	<b>37.50 %</b>
Leave-One-Out	34.67 %	30.67 %	<b>36.00 %</b>	34.67 %	<b>37.33 %</b>

Para categorías con 6 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	<b>42.42 %</b>	33.33 %	30.30 %	39.39 %	30.30 %
Part.70 %	38.89 %	22.22 %	38.89 %	16.67 %	33.33 %
Part.80 %	33.33 %	<b>41.67 %</b>	<b>41.67 %</b>	33.33 %	<b>41.67 %</b>
Leave-One-Out	40.30 %	29.85 %	40.30 %	38.81 %	35.82 %

En las anteriores tablas se observa que la tasa global de aciertos más elevada del conjunto de pruebas se obtiene con un umbral de 6 imágenes mínimo para las categorías, y el porcentaje ideal de imágenes para entrenamiento y para el test está entre el 80 % y el Leave-One-Out.

Lo primero que llama la atención es que es evidente que por sí sola esta configuración da resultados muy por debajo de los obtenidos con los descriptores suaves de color y textura.

Como se ha comentado previamente, variamos el número de descriptores de textura usados (variando los conjuntos de descriptores obtenidos con diferentes distancias) en el análisis según los parámetros anteriores y obtenemos tasas de acierto relativamente similares. En la tabla siguiente vemos como varía la tasa de aciertos para la partición al 50 % del conjunto de categorías con 2 imágenes mínimo:

Distancias	k=1	k=2	k=3	k=4	k=5
d=1	25.00 %	<b>27.50 %</b>	17.50 %	20.00 %	25.00 %
d=1+d=2	25.00 %	22.50 %	15.00 %	22.50 %	<b>27.50 %</b>
d=1+d=2+d=3	<b>30.00 %</b>	20.00 %	17.50 %	<b>27.50 %</b>	25.00 %

Por lo tanto, y siguiendo el camino que otros estudios han mostrado, asumimos que dadas nuestras circunstancias en la mayoría de casos las tasas de acierto serán similares cuando trabajamos con todos los descriptores juntos [26] [39].

**Descriptores de Soh y Tsatsoulis** Los resultados que se encuentran en este apartado se obtienen de la aplicación de un conjunto de descriptores similar al propuesto por Soh y Tsatsoulis, y que se ha discutido en el apartado 2.3.2.

Para categorías con 2 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	32.50 %	25.00 %	25.00 %	22.50 %	30.00 %
Part.70 %	<b>39.13 %</b>	30.43 %	26.09 %	21.74 %	26.09 %
Part.80 %	<b>36.84 %</b>	<b>36.84 %</b>	31.58 %	26.32 %	31.58 %
Leave-One-Out	27.71 %	<b>33.73 %</b>	26.51 %	25.30 %	27.71 %

Para categorías con 4 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	35.14 %	27.03 %	21.62 %	24.32 %	35.14 %
Part.70 %	<b>45.00 %</b>	35.00 %	30.00 %	25.00 %	35.00 %
Part.80 %	<b>43.75 %</b>	<b>43.75 %</b>	<b>43.75 %</b>	31.25 %	37.50 %
Leave-One-Out	32.00 %	37.33 %	29.33 %	30.67 %	29.33 %

Para categorías con 6 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	36.36 %	33.33 %	33.33 %	42.42 %	<b>51.52 %</b>
Part. 70 %	<b>50.00 %</b>	33.33 %	27.78 %	27.78 %	38.89 %
Part.80 %	<b>50.00 %</b>	<b>50.00 %</b>	<b>50.00 %</b>	41.67 %	41.67 %
Leave-One-Out	32.84 %	25.37 %	32.84 %	29.85 %	26.87 %

Podemos observar un comportamiento similar al expuesto en el apartado anterior: El porcentaje de partición idóneo está en torno al 80 % y el número de muestras por categoría influye en el modelo.

Al igual que previamente procedemos a variar el número de atributos para justificación el uso de todos los atributos definidos. En la siguiente tabla se exponen estos resultados, en los que obtenemos de nuevo números similares, siendo los mejores los derivados de los conjuntos más completos en la mayoría de casos.

Distancias	k=1	k=2	k=3	k=4	k=5
d=1	25.00 %	22.50 %	22.50 %	27.50 %	22.50 %
d=1+d=2	<b>32.50 %</b>	25.00 %	22.50 %	20.00 %	27.50 %
d=1+d=2+d=3	<b>32.50 %</b>	25.00 %	25.00 %	22.50 %	<b>30.00 %</b>

De nuevo observamos que los conjuntos más completos son los que mejor resultados dan.

Podemos concluir que los resultados obtenidos en este apartado son ligeramente superiores a los realizados en el anterior, con descriptores de Haralick. No obstante los resultados individuales de trabajar con descriptores obtenidos de la matriz de co-ocurrencia son significativamente inferiores a los realizados en los primeros experimentos con los descriptores suaves de color y textura.

#### 3.3.3. Usando descriptores suaves y matrices de co-ocurrencia

En este apartado vamos a combinar ambos conjuntos de descriptores para observar si la unión de dichos descriptores mejora la tasa de aciertos de nuestro modelo. El procedimiento es similar al del apartado anterior, por lo que experimentamos en escala de grises añadiendo al modelo los discriminantes más significativos del análisis de descriptores suaves de color y textura.

##### 3.3.3.1. Evaluación de resultados

En los posteriores subapartados separaremos los resultados según el subconjunto elegido, basados en los estudios de Haralick y de Soh y Tsatsoulis combinados con los descriptores suaves de color y textura, y ofreceremos conclusiones teniendo en cuenta las combinaciones más relevantes de particiones de muestras para entrenamiento y categorías con un número mínimo de muestras.

**Descriptores suaves de color y textura más descriptores de Haralick** Los resultados que se encuentran en este apartado se obtienen de los descriptores suaves de color y textura destacados en el apartado anterior y de un conjunto de descriptores similar al propuesto por Haralick, y que se ha discutido en el apartado 2.3.2.

Para categorías con 2 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part. 50 %	47.62 %	52.38 %	57.14 %	49.21 %	50.79 %
Part. 70 %	58.33 %	55.56 %	<b>63.89 %</b>	<b>66.67 %</b>	<b>66.67 %</b>
Part. 80 %	59.26 %	59.26 %	55.56 %	59.26 %	55.56 %
Leave-One-Out	<b>59.54 %</b>	48.85 %	53.44 %	54.96 %	57.25 %

Para categorías con 4 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	50.00 %	53.33 %	58.33 %	51.67 %	53.33 %
Part.70 %	63.64 %	63.34 %	<b>66.67 %</b>	<b>69.70 %</b>	<b>72.73 %</b>
Part.80 %	<b>66.67 %</b>	62.50 %	58.33 %	58.33 %	62.50 %
Leave-One-Out	64.23 %	60.16 %	57.72 %	60.16 %	59.35 %

Para categorías con 6 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	58.93 %	60.71 %	57.14 %	51.79 %	53.57 %
Part.70 %	70.97 %	74.19 %	70.97 %	<b>77.42 %</b>	<b>77.42 %</b>
Part.80 %	<b>75.00 %</b>	<b>75.00 %</b>	70.00 %	<b>75.00 %</b>	<b>75.00 %</b>
Leave-One-Out	70.43 %	69.57 %	64.35 %	66.09 %	66.09 %

En las anteriores tablas se observa que la tasa global de aciertos más elevada del conjunto de pruebas se obtiene con un umbral de 6 imágenes mínimo para las categorías, y el porcentaje ideal de imágenes para entrenamiento y para el test está entre el 70 % y el 80 %.

Lo primero que llama la atención es que es evidente que por sí sola esta configuración da resultados significativamente superiores a los obtenidos tan solo con los descriptores suaves de color y textura en su modalidad más eficaz y a los obtenidos usando sólo descriptores de Haralick de las matrices de co-ocurrencia.

### Descriptores suaves de color y textura más descriptores de Soh y Tsatsoulis

Los resultados que se encuentran en este apartado se obtienen de la aplicación de los descriptores suaves de color y textura y de un conjunto de descriptores similar al propuesto por Soh y Tsatsoulis, y que se ha discutido en el apartado 2.3.2.

Para categorías con 2 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	61.90 %	57.14 %	58.73 %	<b>63.49 %</b>	<b>63.49 %</b>
Part.70 %	<b>69.44 %</b>	61.11 %	<b>63.89 %</b>	61.11 %	<b>66.67 %</b>
Part.80 %	59.26 %	55.56 %	51.85 %	55.56 %	55.56 %
Leave-One-Out	61.83 %	56.49 %	58.78 %	58.78 %	58.02 %

Para categorías con 4 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	66.67 %	60.00 %	63.33 %	66.67 %	66.67 %
Part.70 %	<b>78.79 %</b>	63.64 %	<b>72.73 %</b>	<b>69.70 %</b>	<b>69.70 %</b>
Part.80 %	<b>70.83 %</b>	62.50 %	66.67 %	62.50 %	62.50 %
Leave-One-Out	68.29 %	61.79 %	64.23 %	64.23 %	60.98 %

Para categorías con 6 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	71.43 %	67.86 %	67.86 %	69.64 %	71.43 %
Part. 70 %	<b>83.87 %</b>	77.42 %	77.42 %	74.19 %	70.97 %
Part.80 %	<b>80.00 %</b>	<b>85.00 %</b>	<b>75.00 %</b>	<b>80.00 %</b>	75.00 %
Leave-One-Out	72.17 %	74.78 %	71.30 %	71.30 %	66.96 %

Podemos observar un comportamiento similar al expuesto en el apartado anterior: El porcentaje de partición idóneo está en torno al 70-80 % y el número de muestras por categoría influye en el modelo.

Podemos concluir que los resultados obtenidos en este apartado son ligeramente superiores a los realizados en el anterior, con descriptores de Haralick. Es de destacar que los resultados de trabajar con descriptores obtenidos de la matriz de co-ocurrencia combinados con los descriptores suaves de color y textura dan índices de aciertos muy superior y en la línea de lo que se espera del modelo.

#### 3.3.4. Experimentación usando wavelets

En este apartado vamos a trabajar con wavelets, obteniendo descriptores a partir de la aplicación de filtros de log-Gabor.

Este experimento seguirá en la línea establecida en los anteriores apartados, incluyendo su evaluación mediante tasa de aciertos. De la misma forma, hemos de tener en cuenta que:

1. Solo trabajaremos en escala de grises.
2. Puesto que cada muestra es irregular en tamaño y forma, para un análisis más equitativo, las muestras se escalarán mediante interpolación bicúbica a un cuadrado de dimensiones NxN donde N es la dimensión más grande de todo el conjunto de muestras.

En esta sección se detallará como se ha formulado este experimento.

### 3.3.4.1. Preprocesado de las imágenes y elaboración del banco de filtros log-Gabor

Para realizar el experimento y siguiendo las recomendaciones generales [40] ajustamos todas las imágenes a la misma proporción de aspecto, eligiendo para ello el lado mayor de todo nuestro conjunto de muestras como referencia para elaborar un cuadrado de dicha dimensión. De esta forma nos aseguramos no perder información espacial. La interpolación de las muestras a este tamaño se realiza mediante interpolación bicúbica.

Para la elaboración de los filtros utilizaremos las recomendaciones de diseño de los parámetros de Fischer, Cristóbal y Redondo [41, 42] en las que destacan:

- Frecuencia de la primera escala = 0.5.
- Ratio interesalar = 2
- Mejora para respuestas planas
- Normalización para cada sub-banda.
- Sin reducción de tamaño o subsampleo
- Continuidad periódica
- Filtro lowpass.
- Filtro de fase compleja

Por lo tanto se crea un banco de filtros de 4 escalas y 6 orientaciones para cada muestra, que se convolucionarán para crear un total de 24 imágenes por cada muestra.

### 3.3.4.2. Obtención de los atributos

De la misma forma que anteriormente trabajaremos para extraer los descriptores estadísticos simples a partir del histograma de las imágenes convolucionadas.

En nuestro caso disponemos de la información relativa a los 4 primeros momentos: la media, la desviación típica, la asimetría y la curtosis.

Por lo tanto para cada muestra obtendremos un máximo total de 96 descriptores.

### 3.3.4.3. Generando el modelo

Nuevamente, el experimento hereda las características definidas en los apartados anteriores. Así pues, el modelo cambiará según la cantidad de imágenes destinadas a entrenamiento y a test, según la  $k$  y según la combinación de atributos elegidos.

De nuevo se diferenciará entre los experimentos con particiones variables (50 %, 70 % y 80 %) y los Leave-One-Out.

### 3.3.4.4. Evaluación de resultados

Los resultados que se encuentran en este apartado se obtienen de la aplicación exclusiva de los descriptores de log-Gabor explicados previamente y cuya base teórica se ha discutido en el apartado 2.3.3.

Puesto que anteriormente al trabajar con los momentos centrales hemos visto que los resultados óptimos se obtenían con tan solo los primeros 2 o 3 descriptores, vamos a realizar una prueba previa sobre los mismos. Estos resultados están obtenidos del análisis con una partición del 70 % de las muestras para test, con categorías con 6 muestras mínimo:

Momentos	k=1	k=2	k=3	k=4	k=5
1º	43.33 %	53.33 %	50.00 %	36.67 %	43.33 %
1º+2º	50.00 %	50.00 %	46.67 %	43.33 %	46.67 %
1º+2º+3º	<b>53.33 %</b>	<b>56.67 %</b>	<b>53.33 %</b>	53.33 %	50.00 %
1º+2º+3º+4º	40.00 %	53.33 %	50.00 %	<b>56.67 %</b>	43.33 %

Por lo tanto las pruebas se realizarán teniendo en cuenta tan solo los primeros 3 descriptores para un mayor índice de aciertos.

Para categorías con 2 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part. 50 %	33.33 %	<b>42.86 %</b>	28.57 %	20.63 %	19.05 %
Part. 70 %	<b>45.71 %</b>	37.14 %	<b>45.71 %</b>	42.86 %	40.00 %
Part. 80 %	40.74 %	40.74 %	<b>44.44 %</b>	<b>44.44 %</b>	<b>48.15 %</b>
Leave-One-Out	40.77 %	40.77 %	43.85 %	43.08 %	41.54 %

Para categorías con 4 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	35.00 %	41.67 %	26.67 %	20.00 %	20.00 %
Part.70 %	<b>50.00 %</b>	<b>50.00 %</b>	<b>53.12 %</b>	46.88 %	46.88 %
Part.80 %	45.83 %	45.83 %	<b>54.17 %</b>	<b>50.00 %</b>	<b>58.33 %</b>
Leave-One-Out	45.08 %	49.18 %	48.36 %	45.90 %	45.08 %

Para categorías con 6 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	39.29 %	46.43 %	25.00 %	23.21 %	19.64 %
Part.70 %	53.33 %	<b>56.67 %</b>	53.33 %	53.33 %	50.00 %
Part.80 %	<b>55.00 %</b>	55.00 %	<b>60.00 %</b>	<b>60.00 %</b>	<b>65.00 %</b>
Leave-One-Out	48.25 %	52.63 %	51.75 %	49.12 %	49.12 %

Al igual que en los experimentos previos, se observa que en general la tasa global de aciertos más elevada del conjunto de pruebas se obtiene con un umbral de 6 imágenes mínimo para las categorías, y el porcentaje ideal de imágenes para entrenamiento y para el test está entre el 70 % y el 80 %.

A la vista de estos resultados, se evidencia que los descriptores obtenidos mediante filtros de log-Gabor no son lo suficientemente aptos por sí solos para poder clasificar adecuadamente los pólenes.

### 3.3.5. Usando descriptores suaves, matrices de co-ocurrencia y wavelets

En este apartado vamos a combinar todos los conjuntos de descriptores vistos hasta ahora para observar si la unión de los mismos mejora la tasa de aciertos de nuestro modelo. El procedimiento es similar al del apartado anterior, por lo que experimentaremos en escala de grises añadiendo al modelo los discriminantes más significativos del análisis de descriptores suaves de color y textura además de los derivados de la matriz de co-ocurrencia.

#### 3.3.5.1. Evaluación de resultados

Utilizaremos los descriptores que mejores resultados nos han dado en los experimentos realizados en el apartado 3.3.3, combinados con los descriptores obtenidos mediante el filtrado de log-Gabor y ofreceremos conclusiones teniendo en cuenta las combinaciones más relevantes de particiones de muestras para entrenamiento y categorías con un número mínimo de muestras.

Para categorías con 2 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part. 50 %	61.90 %	58.73 %	58.73 %	61.90 %	58.73 %
Part. 70 %	<b>71.43 %</b>	<b>62.86 %</b>	<b>65.71 %</b>	<b>65.71 %</b>	<b>62.86 %</b>
Part. 80 %	66.67 %	59.26 %	59.26 %	55.56 %	55.56 %
Leave-One-Out	62.31 %	59.23 %	58.46 %	59.23 %	57.69 %

Para categorías con 4 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	65.00 %	61.67 %	61.67 %	65.00 %	61.67 %
Part.70 %	<b>78.12 %</b>	<b>68.75 %</b>	<b>71.88 %</b>	<b>71.88 %</b>	<b>68.75 %</b>
Part.80 %	75.00 %	66.67 %	66.67 %	66.67 %	62.50 %
Leave-One-Out	69.67 %	66.39 %	63.93 %	64.75 %	63.11 %

Para categorías con 6 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	67.86 %	67.86 %	66.07 %	71.43 %	66.07 %
Part.70 %	83.33 %	76.67 %	76.67 %	76.67 %	70.00 %
Part.80 %	<b>85.00 %</b>	<b>85.00 %</b>	<b>80.00 %</b>	<b>80.00 %</b>	<b>75.00 %</b>
Leave-One-Out	74.56 %	73.68 %	69.30 %	71.05 %	67.54 %

Una vez más en las anteriores tablas se observa que el análisis combinado de diversos descriptores da mejores resultados que cuando se trabaja con ellos de manera individual. Además, esta combinación nos proporciona la tasa global de aciertos más elevada de todos los conjuntos propuestos hasta ahora.

También se repite en las pruebas que los resultados óptimos se obtienen con un umbral de 6 imágenes mínimo para las categorías, y con un porcentaje ideal de imágenes para entrenamiento y para el test entre el 70 % y el 80 %.

Podemos llegar a la conclusión de que esta configuración es la que mejores resultados da con respecto a las anteriores propuestas, no obstante la mejora en la tasa de aciertos no es muy acentuada. Como observación independiente a la tasa de aciertos, hay que mencionar que el coste computacional de estas pruebas es considerablemente más elevado que en los anteriores casos.

#### 3.3.6. Experimentación usando Patrones Binarios Locales (LBP)

En este apartado vamos a trabajar con los Patrones Binarios Locales. Este experimento seguirá en la línea establecida en los anteriores apartados, incluyendo su evaluación mediante tasa de aciertos. Trabajaremos en escala de grises.

En esta sección se detallará como se ha formulado este experimento.

##### 3.3.6.1. Preprocesado de las imágenes

Para realizar estos experimentos, utilizaremos las mismas muestras segmentadas de las imágenes en escala de grises para garantizar que los resultados sean coherentes con los anteriores experimentos. Por lo tanto, esta base de muestras de pólenes queda compuesta tal cual se define en este capítulo, en el apartado 3.1.

##### 3.3.6.2. Obtención de los atributos y generación del modelo

De la misma forma que anteriormente trabajaremos para extraer los descriptores estadísticos simples a partir del histograma de las matrices resultantes tras analizar sus Patrones Binarios Locales.

Para nuestro caso tendremos que estudiar el comportamiento del modelo al aplicar la media, la desviación típica, y los 4 momentos centrales. Estos descriptores son análogos a los utilizados para el análisis utilizando descriptores suaves de color y textura del capítulo 3.3.1. y tendremos que estudiar si, al igual que ha ocurrido en dicho experimento, los mejores resultados se obtienen solo con un subconjunto de dichos descriptores.

De manera análoga a los experimentos previos, el modelo cambiará según la cantidad de imágenes destinadas a entrenamiento y a test, según la  $k$  y según la combinación de atributos elegidos, también diferenciaremos entre los experimentos con particiones variables (50 %, 70 % y 80 %) y los Leave-One-Out.

### 3.3.6.3. Evaluación de resultados

Los resultados que se encuentran en este apartado se obtienen de la aplicación exclusiva de los descriptores obtenidos a partir de los LBP explicados previamente y cuya base teórica se ha discutido en el apartado 2.3.4.

En primer lugar vamos a analizar la influencia del parámetro de las distancias de los vecinos en la tasa de aciertos. Para ello vamos a probar diversas combinaciones de distancias aplicadas sobre el análisis de las muestras con una partición del 70 % de las mismas para test, y con categorías con 6 muestras mínimo. Hemos utilizado como referencia los 6 descriptores posibles en estas pruebas:

Distancias	k=1	k=2	k=3	k=4	k=5
1	43.33 %	60.00 %	60.00 %	60.00 %	60.00 %
2	63.33 %	63.33 %	56.67 %	60.00 %	63.33 %
3	56.67 %	<b>66.67 %</b>	60.00 %	56.67 %	<b>66.67 %</b>
4	53.33 %	<b>66.67 %</b>	60.00 %	60.00 %	63.33 %
1+2	63.33 %	<b>66.67 %</b>	63.33 %	60.00 %	63.33 %
<b>3+4</b>	<b>70.00 %</b>	63.33 %	<b>66.67 %</b>	<b>66.67 %</b>	66.67 %
1+2+3	66.67 %	63.33 %	63.33 %	<b>66.67 %</b>	<b>70.00 %</b>
2+3+4	66.67 %	<b>66.67 %</b>	60.00 %	63.33 %	<b>70.00 %</b>
1+2+3+4	66.67 %	<b>66.67 %</b>	60.00 %	63.33 %	<b>70.00 %</b>

Considerando estos resultados, parece que lo más adecuado es utilizar una combinación de las distancias con radio 3 y 4.

Puesto que anteriormente al trabajar con los momentos centrales hemos visto que los resultados óptimos se obtenían con tan solo los primeros 2 o 3 descriptores (media, desviación típica y momentos centrales), vamos a realizar una prueba previa sobre los mismos. Estos resultados están obtenidos del análisis con una partición del 70 % de las muestras para test, con categorías con 6 muestras mínimo aplicando las distancias de radio 3 y 4 combinadas:

### 3.3 Segmentación manual y clasificación automática

Momentos	k=1	k=2	k=3	k=4	k=5
1 <sup>o</sup>	33.33 %	40.00 %	40.00 %	30.00 %	36.67 %
1 <sup>o</sup> +2 <sup>o</sup>	43.33 %	43.33 %	60.00 %	60.00 %	56.67 %
1 <sup>o</sup> +2 <sup>o</sup> +3 <sup>o</sup>	63.33 %	<b>63.33 %</b>	63.33 %	<b>70.00 %</b>	<b>66.67 %</b>
1 <sup>o</sup> +2 <sup>o</sup> +3 <sup>o</sup> +4 <sup>o</sup>	66.67 %	<b>63.33 %</b>	<b>66.67 %</b>	<b>70.00 %</b>	63.33 %
1 <sup>o</sup> +2 <sup>o</sup> +3 <sup>o</sup> +4 <sup>o</sup> +5 <sup>o</sup>	66.67 %	<b>63.33 %</b>	<b>66.67 %</b>	<b>70.00 %</b>	<b>66.67 %</b>
<b>1<sup>o</sup>+2<sup>o</sup>+3<sup>o</sup>+4<sup>o</sup>+5<sup>o</sup>+6<sup>o</sup></b>	<b>70.00 %</b>	<b>63.33 %</b>	<b>66.67 %</b>	66.67 %	<b>66.67 %</b>
<b>6<sup>o</sup></b>	<b>70.00 %</b>	<b>63.33 %</b>	<b>66.67 %</b>	66.67 %	<b>66.67 %</b>

Analizando las dos últimas filas de la tabla podemos observar el hecho de que los resultados no se alteran cuando utilizamos exclusivamente el último momento o lo combinamos con cualquiera de los demás descriptores. Para asegurarnos de que no es solo una coincidencia en el porcentaje de aciertos, analizamos la matriz de confusión de las diferentes muestras y confirmamos que la clasificación ha sido exactamente la misma.

Observando los valores de este momento sobre las muestras observamos que su magnitud es de varias órdenes superior a los demás descriptores, hecho que nos lleva a pensar que las desviaciones en el valor de este momento a lo largo de las distintas categorías cobran más relevancia para el modelo de clasificación que la combinación de las desviaciones de los valores de los demás descriptores.

No obstante, este descriptor es el que mejores resultados otorga a priori, por lo que realizaremos nuestra batería de pruebas con respecto a los parámetros obtenidos: distancias combinadas de radio 3 y 4, y el quinto momento central. Estos son los resultados:

Para categorías con 2 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part. 50 %	49.21 %	52.38 %	52.38 %	55.56 %	<b>57.14 %</b>
Part. 70 %	<b>60.00 %</b>	51.43 %	51.43 %	<b>60.00 %</b>	<b>57.14 %</b>
Part. 80 %	<b>59.26 %</b>	<b>59.26 %</b>	<b>55.56 %</b>	51.85 %	48.15 %
Leave-One-Out	50.00 %	43.85 %	47.69 %	50.00 %	53.08 %

Para categorías con 4 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	55.00 %	53.33 %	53.33 %	58.33 %	60.00 %
Part.70 %	<b>62.50 %</b>	<b>56.25 %</b>	<b>56.25 %</b>	<b>62.50 %</b>	<b>62.50 %</b>
Part.80 %	<b>62.50 %</b>	50.00 %	54.17 %	50.00 %	54.17 %
Leave-One-Out	54.92 %	50.82 %	52.46 %	53.28 %	56.56 %

Para categorías con 6 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	58.93 %	55.36 %	60.71 %	64.29 %	62.50 %
Part.70 %	<b>70.00 %</b>	63.33 %	<b>66.67 %</b>	<b>66.67 %</b>	<b>66.67 %</b>
Part.80 %	<b>70.00 %</b>	<b>70.00 %</b>	65.00 %	65.00 %	65.00 %
Leave-One-Out	61.40 %	55.26 %	57.02 %	59.65 %	64.04 %

Al igual que en los experimentos previos, se observa que en general la tasa global de aciertos más elevada del conjunto de pruebas se obtiene con un umbral de 6 imágenes mínimo para las categorías, y el porcentaje ideal de imágenes para entrenamiento y para el test está entre el 70 % y el 80 %.

Los resultados obtenidos son relativamente similares en su tasa de acierto a los obtenidos con los descriptores suaves de color y textura, y es evidente que por sí solos no nos sirven para clasificar nuestros pólenes.

### 3.3.7. Usando descriptores suaves, matrices de co-ocurrencia, wavelets y patrones binarios locales

En este apartado vamos a estudiar los resultados de aplicar todos los conjuntos de descriptores vistos hasta ahora para observar si la combinación de los mismos mejora la tasa de aciertos de nuestro modelo. De nuevo, el procedimiento es similar al del apartado anterior: Experimentaremos en escala de grises añadiendo al modelo los discriminantes más significativos de los análisis de descriptores previos.

#### 3.3.7.1. Evaluación de resultados

Utilizaremos los descriptores que mejores resultados nos han dado en los experimentos realizados en el apartado 3.3.5, sumado al último momento central sobre el histograma, que nos otorgaba los mejores resultados. Los resultados se presentarán separados según las diferentes combinaciones de particiones de muestras para entrenamiento y categorías con un número mínimo de muestras.

Para categorías con 2 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part. 50 %	49.21 %	52.38 %	52.38 %	55.56 %	57.14 %
Part. 70 %	<b>60.00 %</b>	51.43 %	51.43 %	<b>60.00 %</b>	<b>57.14 %</b>
Part. 80 %	59.26 %	<b>59.26 %</b>	<b>55.56 %</b>	51.85 %	48.15 %
Leave-One-Out	50.00 %	43.58 %	47.69 %	50.00 %	53.08 %

Para categorías con 4 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	55.00 %	53.33 %	53.33 %	58.33 %	60.00 %
Part.70 %	<b>62.50 %</b>	<b>56.25 %</b>	<b>56.25 %</b>	<b>62.50 %</b>	<b>62.50 %</b>
Part.80 %	<b>62.50 %</b>	50.00 %	54.17 %	50.00 %	54.17 %
Leave-One-Out	54.92 %	50.82 %	52.46 %	53.28 %	56.56 %

Para categorías con 6 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	58.93 %	55.36 %	60.71 %	64.29 %	62.50 %
Part.70 %	<b>70.00 %</b>	63.33 %	<b>66.67 %</b>	<b>66.67 %</b>	<b>66.67 %</b>
Part.80 %	<b>70.00 %</b>	<b>70.00 %</b>	65.00 %	65.00 %	65.00 %
Leave-One-Out	61.40 %	55.26 %	57.02 %	59.65 %	64.04 %

Podemos observar que en esta ocasión la tasa de aciertos desciende, el análisis combinado de todos estos descriptores da peores resultados que cuando se trabajaba sin LBP. Puesto que el descriptor elegido en este caso es el último momento y nuestros descriptores no están normalizados para la clasificación, vamos a repetir el experimento utilizando los descriptores estadísticos que han dado mejores resultados en anteriores pruebas: la media y la desviación típica.

Para categorías con 2 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part. 50 %	61.90 %	58.73 %	58.73 %	61.90 %	58.73 %
Part. 70 %	<b>71.43 %</b>	<b>62.86 %</b>	<b>65.71 %</b>	<b>65.71 %</b>	<b>62.86 %</b>
Part. 80 %	66.67 %	59.26 %	59.26 %	55.56 %	55.56 %
Leave-One-Out	62.31 %	59.23 %	58.46 %	59.23 %	57.69 %

Para categorías con 4 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	65.00 %	61.67 %	61.67 %	65.00 %	61.67 %
Part.70 %	<b>78.12 %</b>	<b>68.75 %</b>	<b>71.88 %</b>	<b>71.88 %</b>	<b>68.75 %</b>
Part.80 %	75.00 %	66.67 %	66.67 %	66.67 %	62.50 %
Leave-One-Out	69.67 %	66.39 %	63.93 %	64.75 %	63.11 %

Para categorías con 6 imágenes mínimo y particiones variables y Leave-One-Out:

Porcentaje entrenamiento	k=1	k=2	k=3	k=4	k=5
Part.50 %	67.86 %	67.86 %	64.29 %	71.43 %	66.07 %
Part.70 %	83.33 %	76.67 %	76.67 %	76.67 %	70.00 %
Part.80 %	<b>85.00 %</b>	<b>85.00 %</b>	<b>80.00 %</b>	<b>80.00 %</b>	<b>75.00 %</b>
Leave-One-Out	74.56 %	73.68 %	69.30 %	71.05 %	67.54 %

En esta ocasión observamos como efectivamente, los resultados son mejores que los obtenidos previamente trabajando con el último momento central.

Por el contrario se observa que la tasa de acierto ya no mejora con respecto a los resultados obtenidos en el capítulo 3.3.5, concretamente obteniendo los mismos números en ambos experimentos en prácticamente todos los casos. Un estudio más detallado de la clasificación nos desvela que hay variaciones en los resultados de categorías cuyas muestras eran anteriormente clasificadas de forma errónea, pero que no alteran la tasa de aciertos ya que el modelo las etiqueta equivocadamente de nuevo.

## 4. Conclusiones

En este proyecto se han estudiado los resultados de la aplicación de un conjunto de descriptores para realizar un sistema de clasificación automático de pólenes a partir de imágenes de microscopio de miel. Estos descriptores están derivados del color y la textura de la imagen.

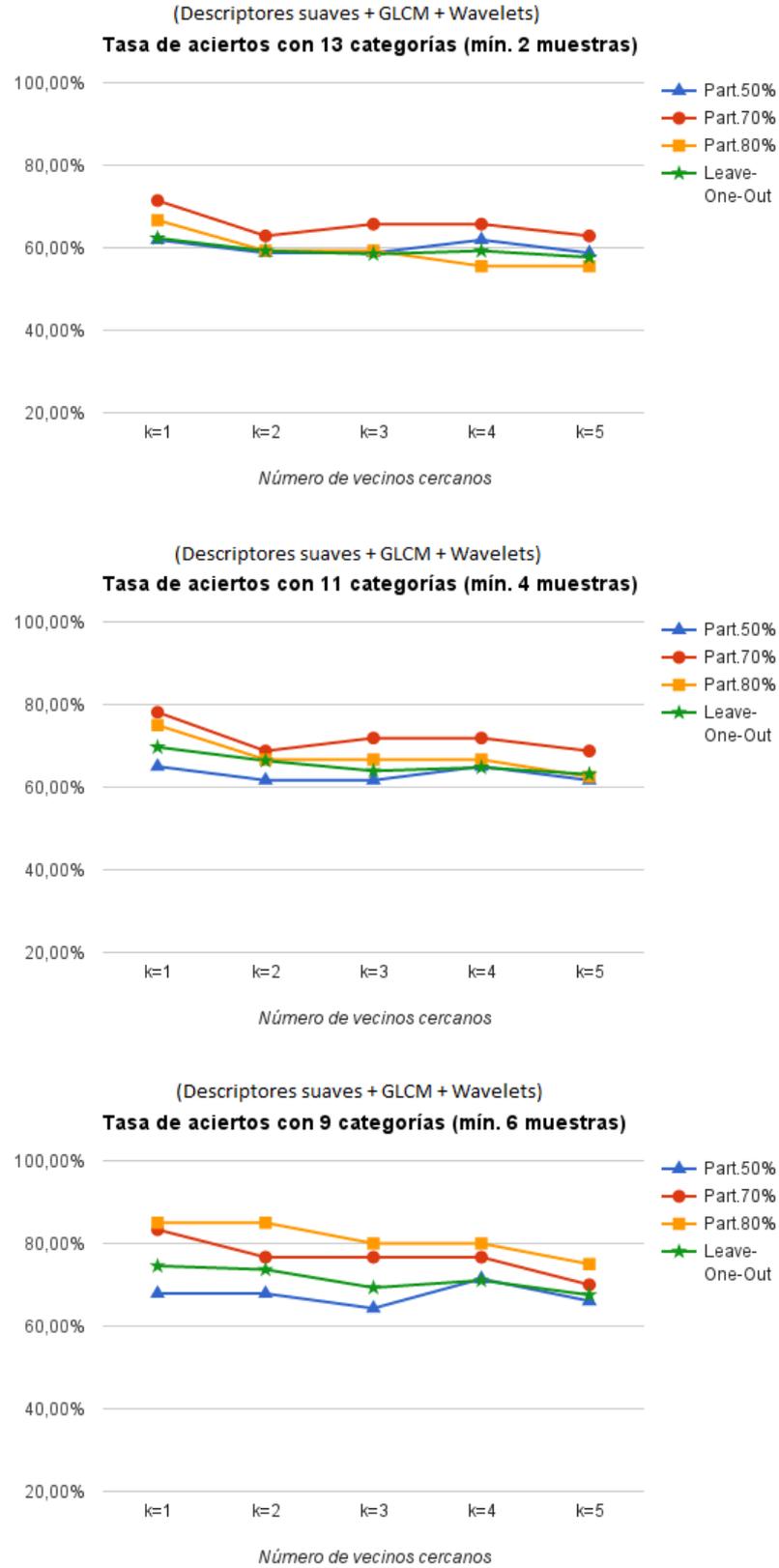
En el contexto industrial, estas tareas se realizan principalmente de manera manual según el criterio experto.

Para realizar nuestro estudio hemos dispuesto de una colección de 12 imágenes en color de mieles y de la que hemos podido extraer 138 muestras de pólenes de 16 clases distintas y hemos trabajado con las herramientas que el entorno MATLAB nos ofrece en su versión r2014a.

Sobre estos recursos se ha llevado a cabo una serie de experimentos consistentes en observar la tasa de aciertos de un modelo propuesto con diferentes parámetros. Los primeros experimentos, con segmentación y clasificación automáticas, se han realizado sobre la misma imagen para obtener regiones euclídeamente cercanas con respecto a los descriptores suaves de color y textura de una muestra seleccionada por el usuario, utilizando para ello tres esquemas de color: escala de grises, RGB y CIE LAB.

Considerando que los primeros resultados han sido negativos en cuanto a la clasificación, a partir de este punto se ha optado por realizar una segmentación manual para evaluar diferentes métodos de clasificación automática, creando una base de datos de muestras para observar si los descriptores por sí solos serían suficientes para un análisis adecuado de nuestros datos y, posteriormente, se ha repetido el estudio trabajando con los descriptores de textura derivados de las matrices de co-ocurrencia, descriptores obtenidos de wavelets y con patrones binarios locales. El método elegido para esta clasificación ha sido k-NN sin ningún tipo de normalización de descriptores.

Teniendo en cuenta los resultados de los diversos modelos propuestos y considerando algunos aspectos de su comportamiento, nos planteamos que algunas especies de pólenes no tienen la cantidad de muestras necesarias para poder clasificarlos de manera precisa.



**Cuadro 4.1.:** Resultados de aplicar descriptores suaves, GLCM y wavelets, variando el número de categorías según su número de muestras. Se muestran las cuatro diferentes particiones de entrenamiento y test.

Como se puede observar en el cuadro 4.1 los mejores resultados globales se han obtenido al combinar descriptores suaves, matrices de co-ocurrencia y wavelets, alcanzando tasas de acierto máximas de un 85 % cuando clasificamos únicamente pólenes de cuyas categorías tenemos más de 6 muestras, con un 80 % de esas muestras destinadas a entrenamiento y un número de vecinos cercanos entre 1 y 3. Los porcentajes de acierto no han variado significativamente al añadir los patrones binarios locales a esta combinación, por lo que se han omitido.

Así pues, a la vista de los resultados obtenidos y considerando los buenos resultados en la segmentación automática, podemos establecer que la mejora en la tasa de aciertos al combinar todos los descriptores es lo suficientemente significativa como para considerar seguir trabajando en la misma línea, estudiando qué combinaciones de descriptores nos proporciona las mayores garantías a la hora de discriminar entre especies, con perspectivas de obtener un modelo de clasificación con características adecuadas para responder a los estándares de calidad que requiere la industria.

### 4.1. Trabajo futuro

En vista de los resultados, se abren varias posibles líneas de estudio de cara al futuro.

Por un lado, para garantizar con seguridad la adecuación del modelo a escenarios de uso habituales en la industria, habría que obtener nuevas muestras para nuestro estudio, especialmente de aquellas categorías de las que disponemos menos muestras y también de aquellas cuya morfología sea parcialmente irregular, ya que obtener diferentes ángulos o perspectivas focales podría enriquecer la base de datos cuyos parámetros nos sirven para discriminar entre especies.

Por otra parte podríamos estudiar con más detalle los errores a la hora de clasificar e intentar encontrar algún patrón de fallos (como por ejemplo, confusiones entre dos especies) que sea solucionable mediante algún atributo específico, bien de color y textura, o bien morfológico, que nos permita reducir este tipo de errores específicos.

Por último se puede explorar el comportamiento de otros modelos de clasificación ante este problema: Como se ha mencionado en el capítulo de experimentación, una clasificación realizada mediante discriminantes normalizados o dotando de pesos distintos a cada atributo, muy probablemente nos de más garantías de éxito. Y también sería interesante analizar cuáles de nuestros atributos no aportan información valiosa para la clasificación, por ejemplo mediante un análisis de discriminación lineal.

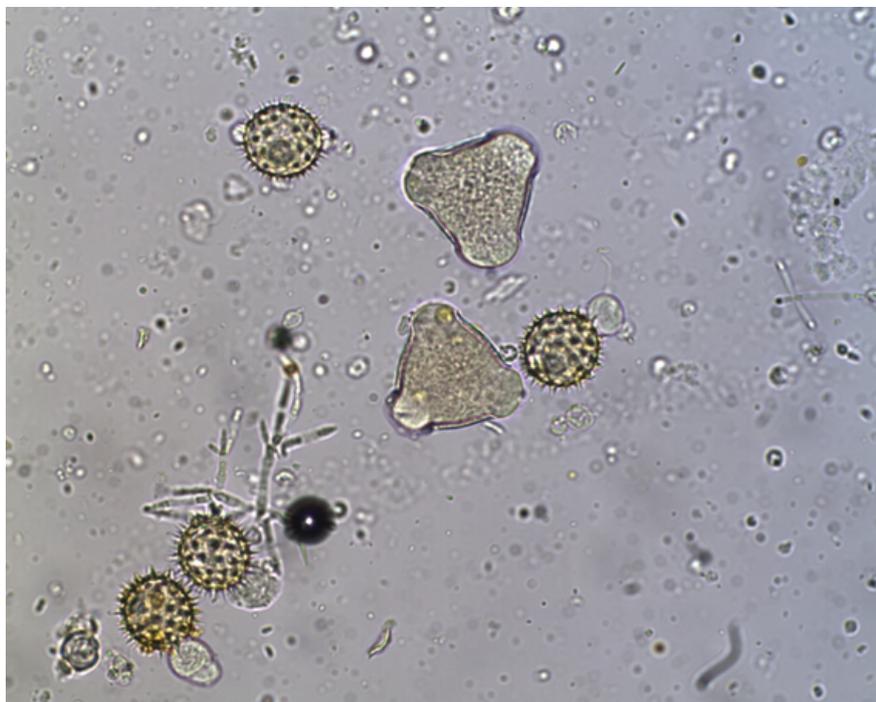
De la misma forma, el modelo de clasificación podría ser plasmado como un árbol de clasificación donde se examinen en diferentes niveles los atributos del polen.

Otros estudios, como comentábamos en el Capítulo 1, utilizan redes neuronales, o máquinas de vector soporte para modelizar la clasificación. Estos modelos, que podemos afirmar que sí que son más complejos de computar, no garantizan una mayor eficacia a la hora de obtener resultados, por lo que habría que estudiar su adecuación.



## A. Anexo A: Imágenes de pólenes proporcionadas

A continuación se adjuntan las 12 imágenes a partir de las cuales se ha realizado todo el trabajo presente en este documento. Se ofrece un resumen de la cantidad de pólenes de cada especie que aparecen en las mismas en el cuadro 3.1 del Capítulo 3, así como una tabla con las categorías y un ejemplo de cada en la Figura 3.1 del mismo Capítulo.



**Figura A.1.:** Imagen 1

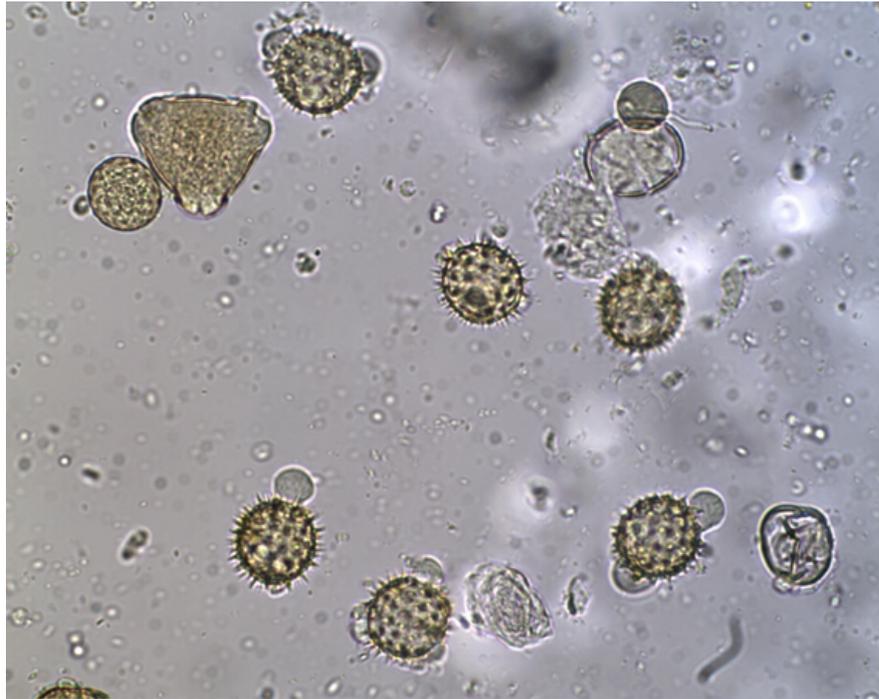
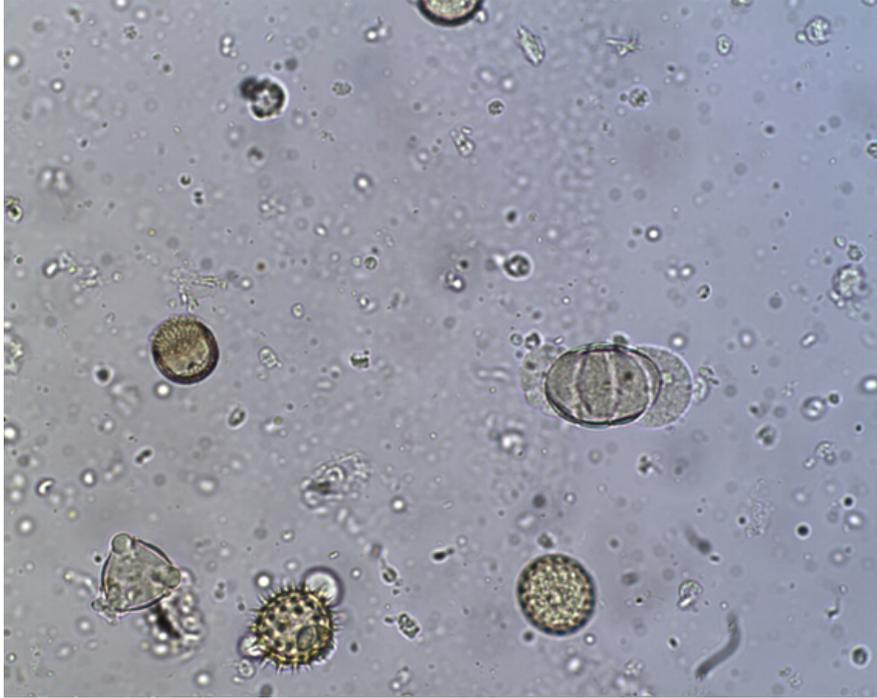


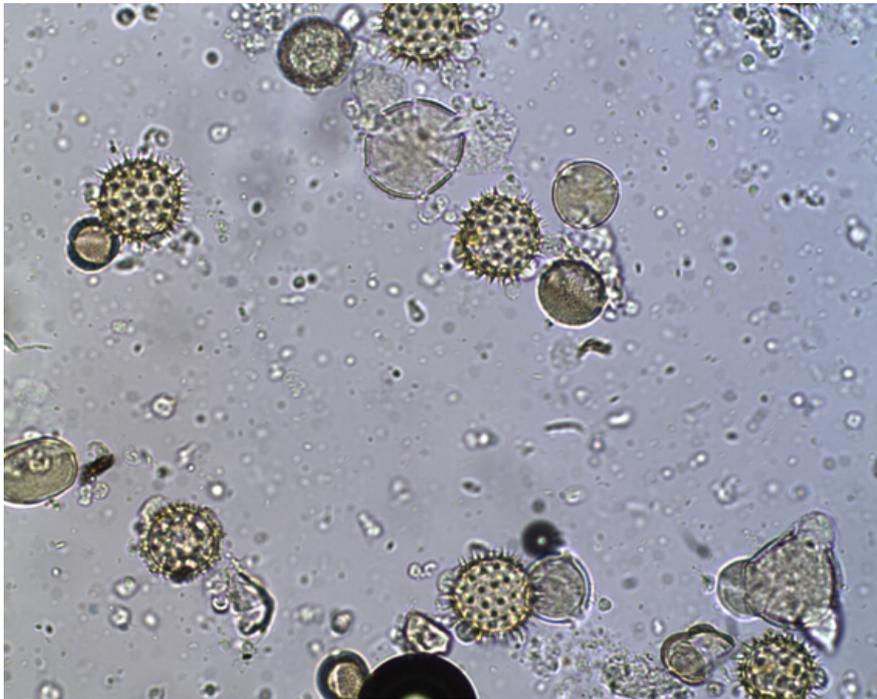
Figura A.2.: Imagen 2



Figura A.3.: Imagen 3



**Figura A.4.:** Imagen 4



**Figura A.5.:** Imagen 5



Figura A.6.: Imagen 6

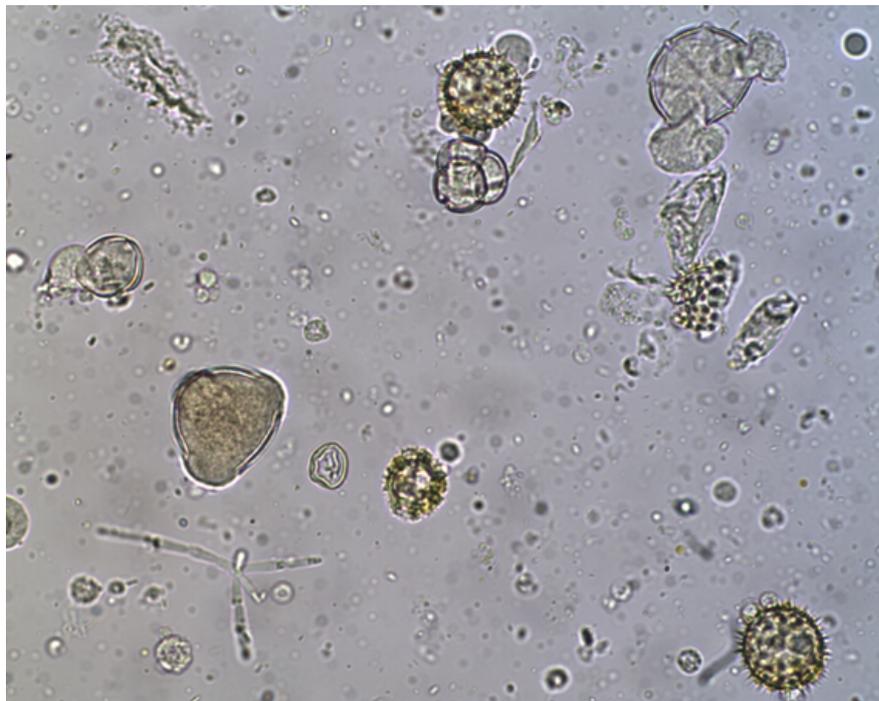
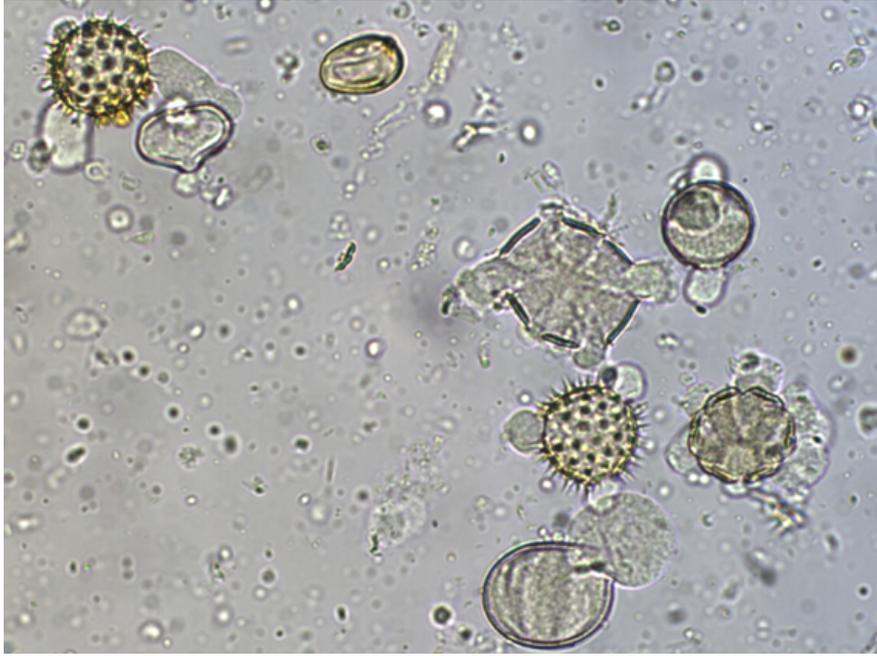
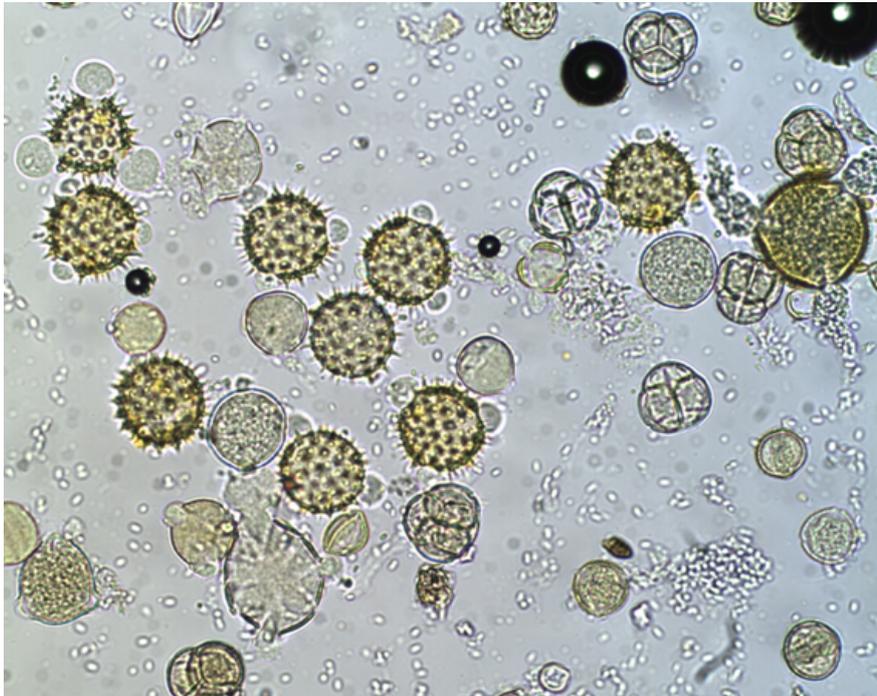


Figura A.7.: Imagen 7



**Figura A.8.:** Imagen 8



**Figura A.9.:** Imagen 9

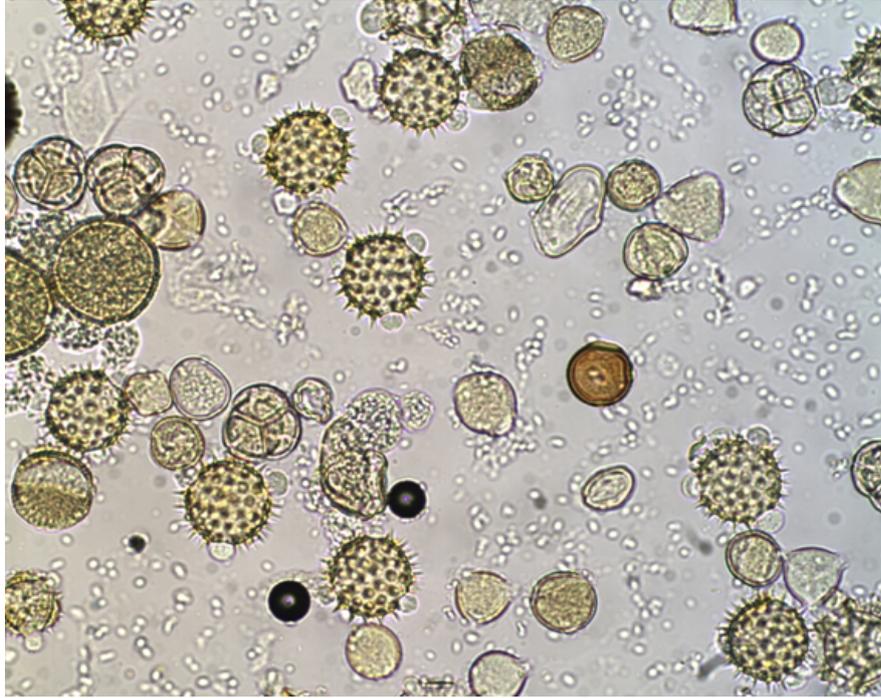


Figura A.10.: Imagen 10

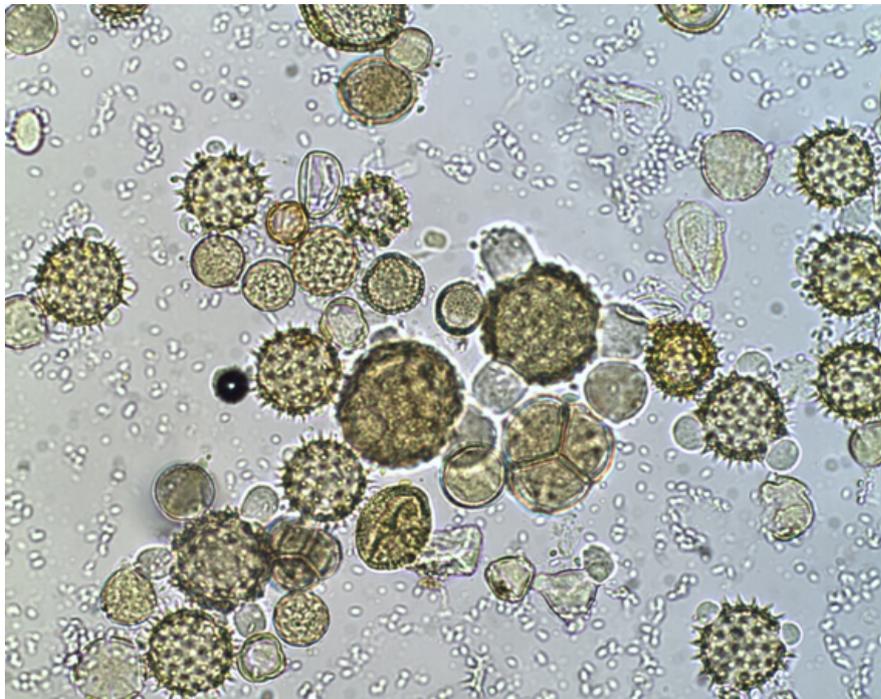
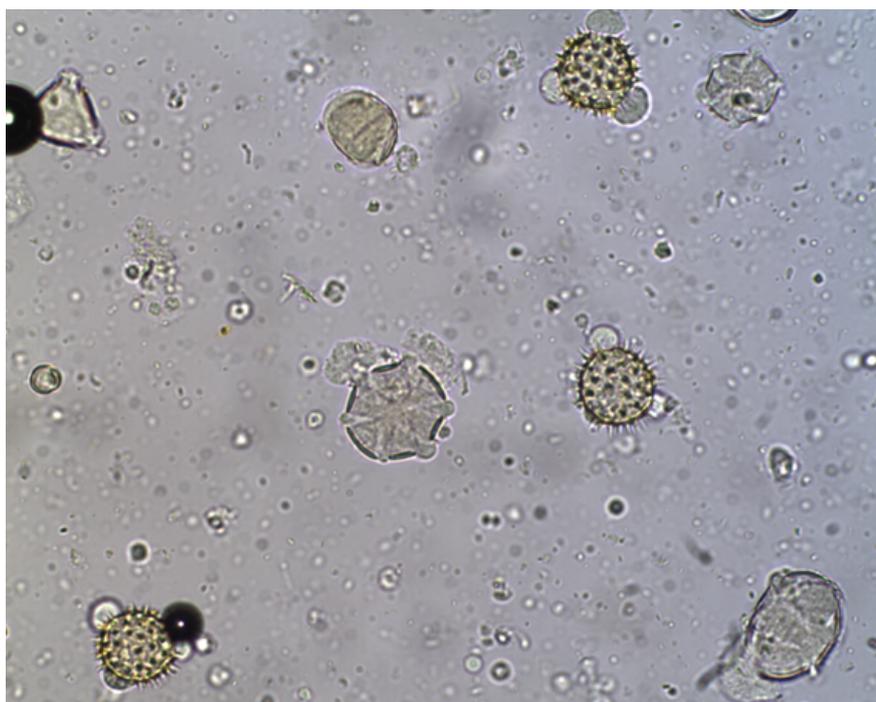


Figura A.11.: Imagen 11



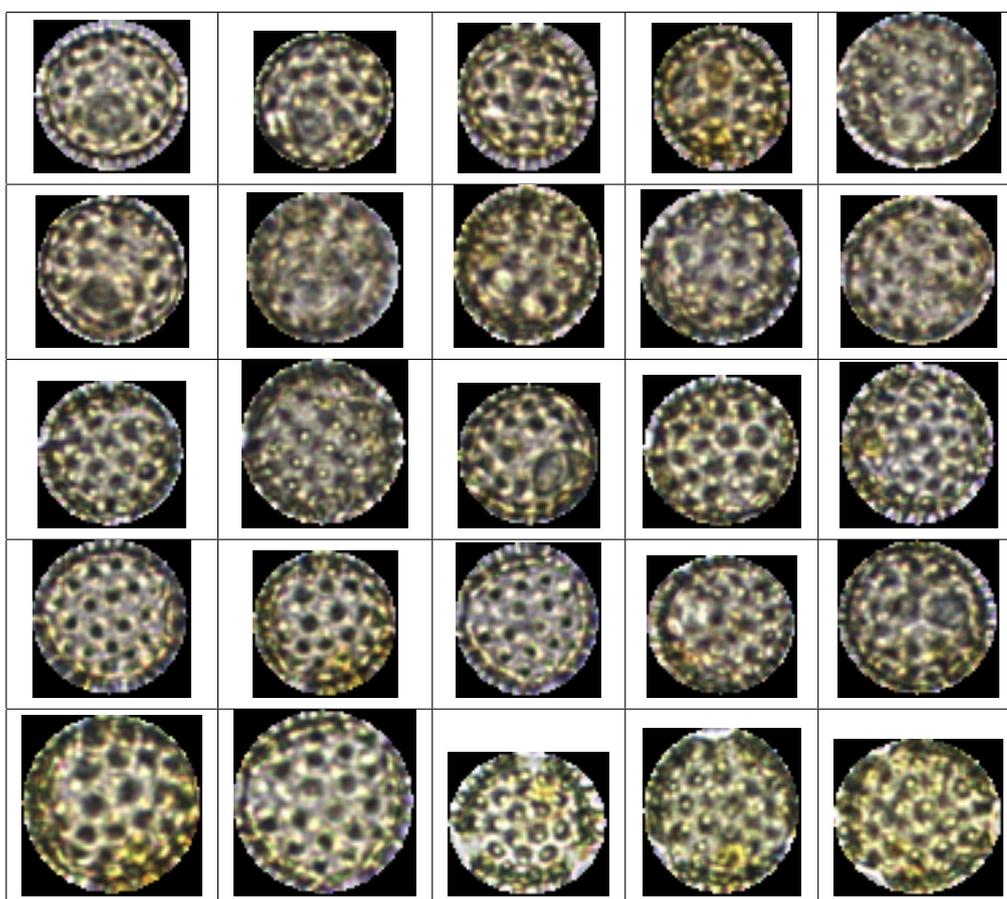
**Figura A.12.:** Imagen 12

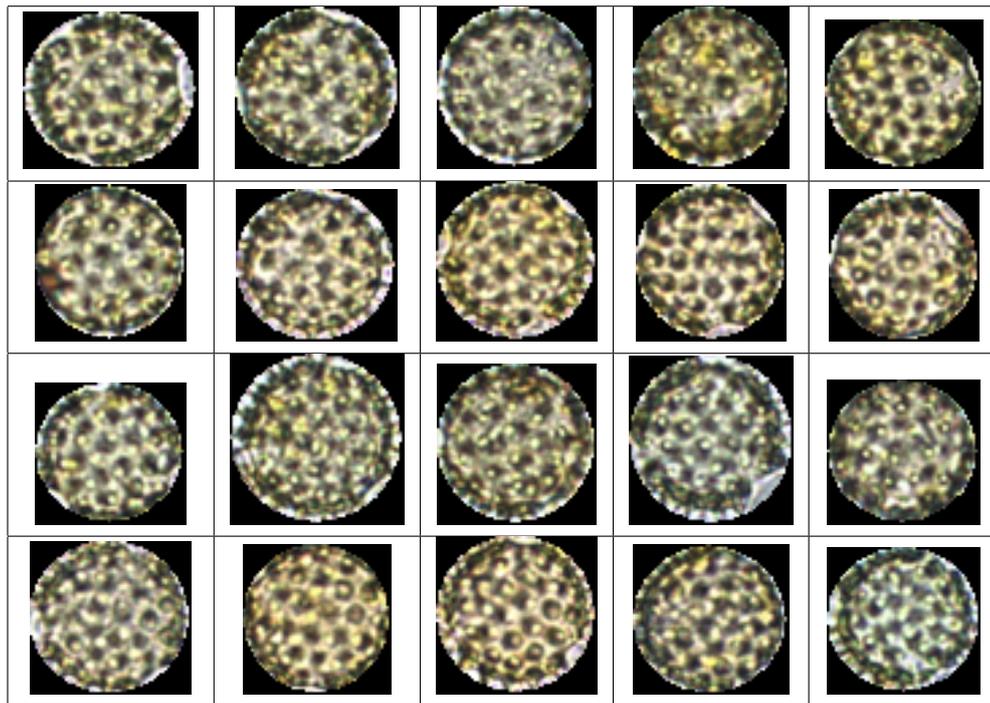


## B. Anexo B: Muestras segmentadas manualmente.

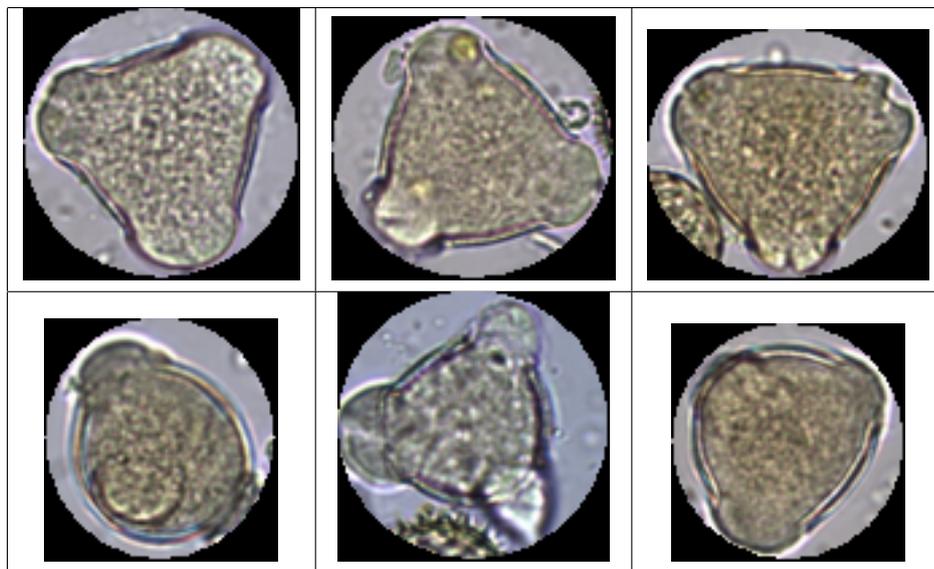
A continuación se adjuntan las 138 muestras obtenidas a partir de las imágenes del Anexo A, separadas por especie de polen.

### B.1. Helianthus

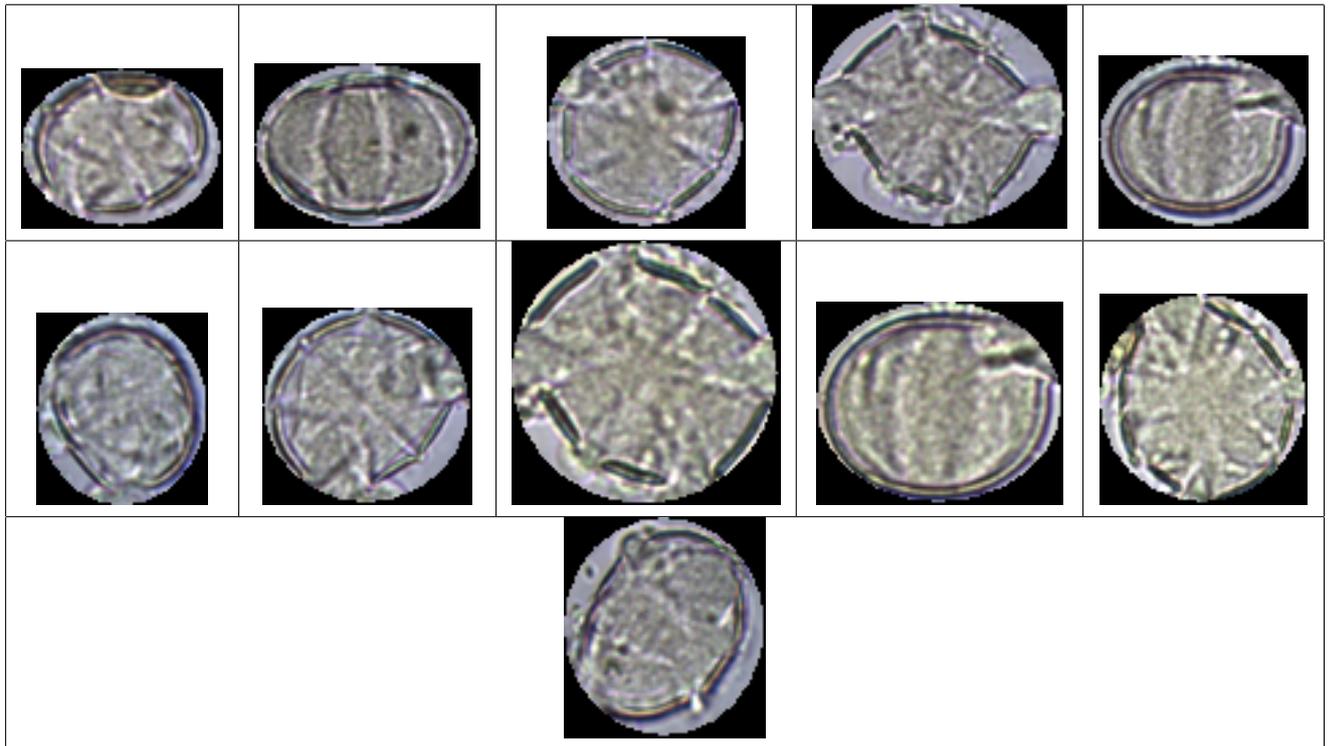




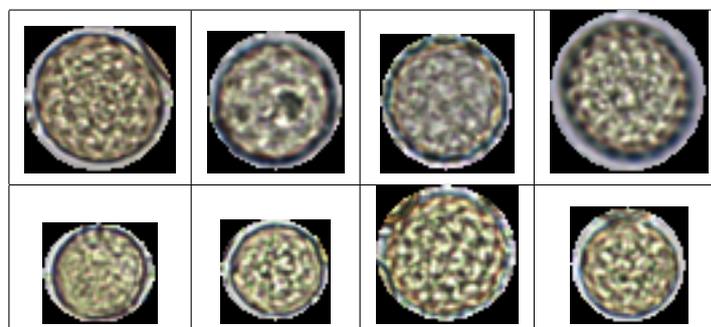
## B.2. Prunus



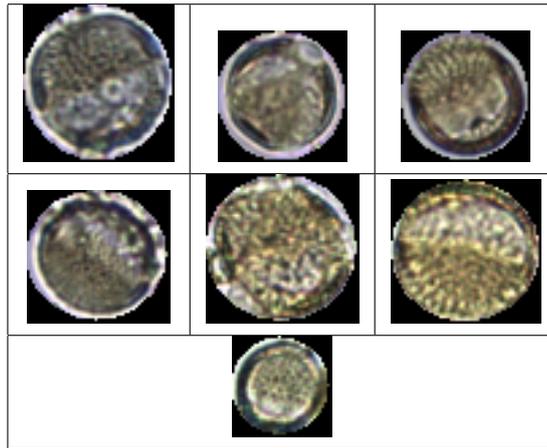
## B.3. Rosmarinus



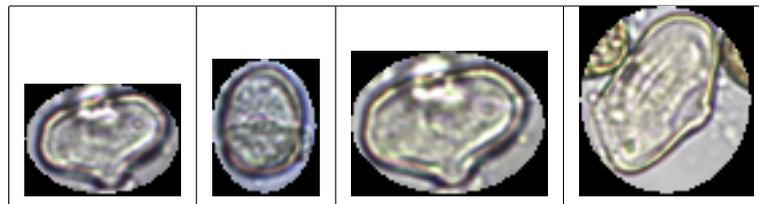
#### B.4. Polygonaceas



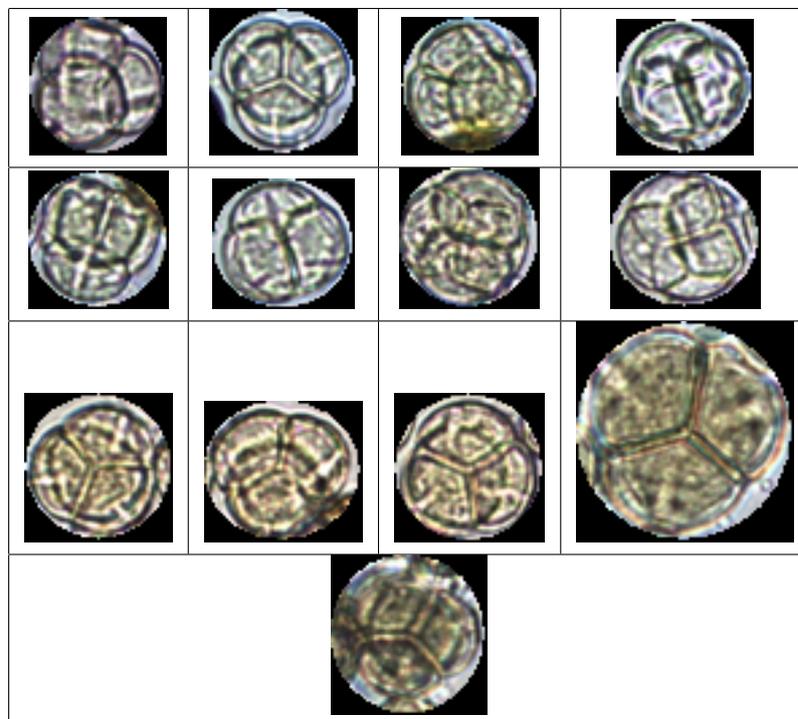
#### B.5. Diplotaxis



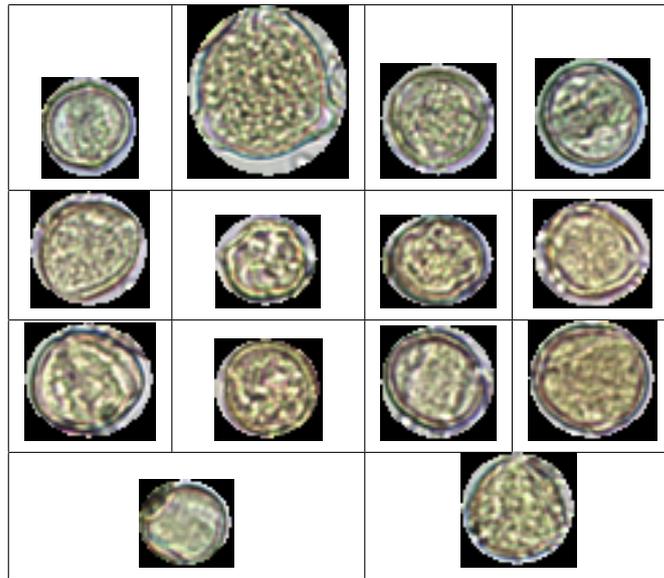
### B.6. Eryngium



### B.7. Erica



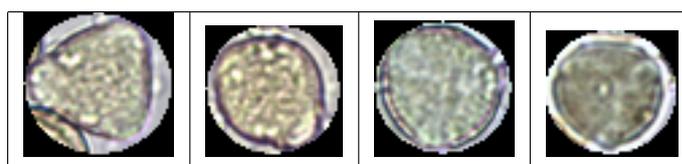
**B.8. Quercus sp.**



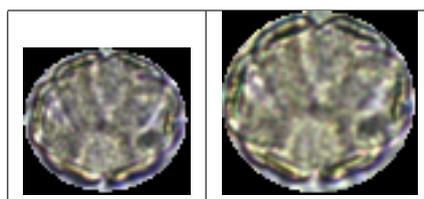
**B.9. Trifolium**



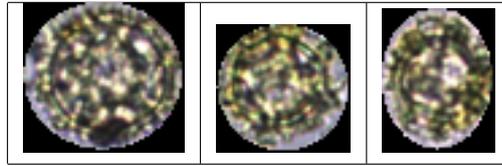
**B.10. Tipo Genista**



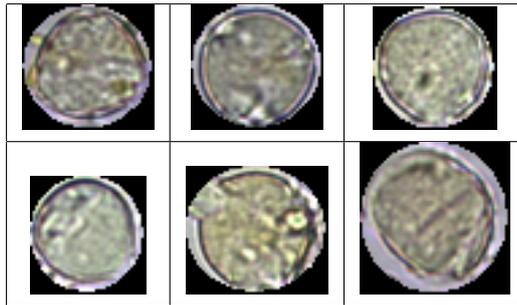
**B.11. Lavandula**



**B.12. Taraxacum sp.**



**B.13. Sedum**



**B.14. Thymus**



**B.15. L.Stoechas**



# Bibliografía

- [1] H.D. Belitz and W. Grosch. *Food Chemistry*. Ed. Acribia S.A., 1988.
- [2] Y. Kaya, M.E. Erez, O. Karabacak, L. Kayci, and M. Fidan. *An automatic identification method for the comparison of plant and honey pollen based on GLCM texture features and artificial neural network*. Grana, vol.52-1, 2013.
- [3] R. Pfister. *Versuch einer Mikroskopie des Honigs*. Foscungsber Lebensmitt. Bez. Hyg. Chem. Pharm. (Munich), 1895.
- [4] C. Griebel. *Zur Polleanalyse des Honigs*. Z. F. Unters. Lebensmittel, 1930.
- [5] G. Erdtman. *An introduction to pollen analysis*. Chronica Bot., Waltham MA, 1943.
- [6] E. Zander. *La palinología al servicio del estudio de la miel*. An. Edaf. Fisiol. Veget. 9, 1950.
- [7] W. Von Der Ohe, L. Oddo, M. Piana, M. Morlot, and P. Martin. *Harmonized methods of melissopalynology*. Apidologie, Springer Verlag, 35, 2004.
- [8] W. Kirk. *A Colour Guide to Pollen Loads of the Honey Bee*. International Bee Research Association, New York, USA., 1994.
- [9] M. Campos, K. Markham, K. Mitchell, and A. da Cunha. *An approach to the characterization of bee pollens via their flavonoid/phenolic profiles*. Phytochemical Analysis 8, 1997.
- [10] M.P. de Sá-Otero and E. Díaz-Losada. *Método de determinación del origen geográfico del polen apícola comercial*. Lazaroa 23, 2002.
- [11] E. Stillman and J.R. Flenley. *The needs and prospects for automation in palynology*. Quat. Sci. Rev. 15, 1996.
- [12] J.R. Flenley. *The Problem of Pollen Recognition*. Canberra, Australia, 1968.
- [13] H.J.L. White. *Preliminary Research into the Possibilities of Automated Pollen Counting*. Pollen Et Spores, vol. XXX, 1988.
- [14] M. Rodríguez-Damián, E. Cernadas, A. Formella, M. Fernández-Delgado, and P. de Sá-Otero. *Automatic detection and classification of grains of pollen based on shape and texture*. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 36, 2006.
- [15] G. Allen. *An Automated Pollen Recognition System*. Massey University, 2006.

- 
- [16] S. Shafiee, S. Minaei, N. Moghaddam-Charkari, and M. Barzegar. *Honey characterization using computer vision system and artificial neural networks*. Elsevier, 2014.
- [17] R.C. Gonzalez and R.E. Woods. *Digital Image Processing (3rd Ed.)*. Addison-Wesley, 2007.
- [18] F. López García. *Real-Time Surface Grading Ceramic Tiles*. Universidad Politécnica de Valencia, 2005.
- [19] Y. Ohno. *CIE Fundamentals for color measurements*. International Conference on Digital Printing Technologies, Vancouver, 2000.
- [20] J.R. Jensen. *Introductory Digital Image Processing (2nd Ed.)*. Ed. Prentice Hall, 1996.
- [21] M. Hassner and J. Sklansky. *The use of Markov random fields as models of Texture*. CG and Image Processing, 1980.
- [22] T. Ojala, M. Pietikainen, and T. Maenpaa. *Multiresolution gray-scale and 718 rotation invariant texture classification with local binary patterns*. IEEE 719 Trans. Pattern anal. Mach. Intell 24, 2002.
- [23] B.B. Mandelbrot. *The fractal geometry of nature*. Freeman, NY, 1983.
- [24] R. Haralick and S. Shamnugam. *Textural Features for Image Classification*. IEEE Transactions on Cybernetics, Vol SMC-3, No. 2, 1973.
- [25] M. Presutti. *La matriz de co-ocurrencia en la clasificación multiespectral*. Universidad Nacional de La Plata, 2004.
- [26] L.K. Soh and C. Tsatsoulis. *Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices*. IEEE Trans. Geosci. Remote Sensing, 1999.
- [27] R.W. Connors, M.M. Trivedi, and C.A. Harlow. *Segmentation of a high-resolution urban scene using texture operators*. Comput. Vision, Graph., Image Processing, vol. 25, 1984.
- [28] A. Haar. *Zur Theorie der orthogonalen Funktionensysteme*. Mathematische Annalen 69, 1910.
- [29] J.G. Daugman. *Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters*. Journal of the Optical Society of America A 2, 1985.
- [30] S. Marcelja. *Mathematical description of the responses of simple cortical cells*. Journal of the Optical Society of America 70, 1980.
- [31] D.J. Field. *Relation between the statistics of natural images and the response properties of cortical cells*. Journal of the Optical Society of America A 4, 1987.
- [32] X. Gao, F. Sattar, and R. Venkateswarlu. *Multiscale corner detection of gray level images based on log-Gabor wavelet transform*. IEEE Trans. Circuits Syst. Video Technol. 17, 2007.

- [33] P. Kovési. *Image features from phase congruency*. VIDERE- Journal of computer vision research., 1999.
- [34] J.V. Marcos, R. Nava, G. Cristóbal, R. Redondo, , and B. Escalante-Ramírez. *Automated pollen identification using microscopic imaging and texture analysis*. Micron Volume 68, 2014.
- [35] R. Nava, B. Escalante-Ramírez, and G. Cristóbal. *Texture image retrieval based on log-Gabor features*. Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. Springer Berlin., 2012.
- [36] T. Ojala, M. Pietikainen, and D. Harwood. *Performance evaluation of texture measures with classification based on Kullback discrimination of distributions*. Proceedings of the 12th IAPR International Conference on Pattern Recognition 1, 1994.
- [37] T. Ojala, M. Pietikainen, and D. Harwood. *A Comparative Study of Texture Measures with Classification Based on Feature Distributions*. Pattern Recognition vol. 29, 1996.
- [38] F. López García, X. Fdez-Vidal, X. Pardo, and R. Dosil. *A Biologically Inspired Vision System for Enclosed Environments*. Pattern Recognition. Chapter 13,pp.249-266.IN-TECH, 2009.
- [39] P. C. Chen and T. Pavlidis. *Segmentation by texture using a co- occurrence matrix and a split-and-merge algorithm*. Comput. Graph. Image Processing, vol. 10, no. 2, 1979.
- [40] B.S. Manjunath and W.Y. Ma. *Texture features for browsing and retrieval of image data*. IEEE Transactions on Pattern Analysis and Machine Intelligence Volume 18, 1996.
- [41] S. Fischer, G. Cristóbal, and R. Redondo. *Sparse overcomplete Gabor Wavelet representation based on local competitions*. IEEE Trans. on Image Processing 15, 2006.
- [42] S. Fischer, F. Sroubek, L. Perrinet, R. Redondo, and G. Cristóbal. *Self-invertible log-Gabor wavelets*. Int. Journal of Computer Vision 75, 2007.