Universidad Politécnica de Madrid
Escuela Técnica Superior de Ingenieros
Informáticos

POLITÉCNICA

Algoritmos de segmentación semántica
para anotación de imágenes

TESIS DOCTORAL

César Antonio Ortiz Toro
Ingeniero en Informática

Madrid 2019

Algoritmos de segmentación semántica para anotación de imágenes

*Memoria que presenta para optar al título de Doctor en Software,
Sistemas y Computación*

César Antonio Ortiz Toro
Ingeniero en Informática

Dirigida por las Doctoras

Consuelo Gonzalo Martín, Ernestina Menasalvas Ruiz

Departamento de Arquitectura y Tecnología de Sistemas Informáticos
Escuela Técnica Superior de Ingenieros Informáticos

Madrid 2019

Tribunal nombrado por el Mgfo. y Excmo. Sr. Rector de la Universidad Politécnica de Madrid, el día de de 2019

PRESIDENTE:

VOCAL:

VOCAL:

VOCAL:

SECRETARIO:

Realizado el acto de defensa y lectura de Tesis el día

Calificación:

EL PRESIDENTE

EL SECRETARIO

LOS VOCALES

Agradecimientos

Quiero expresar mi mas sincero agradecimiento a mis directoras de tesis, Dra. Consuelo Gonzalo Martín y Dra. Ernestina Menasalvas Ruiz, por su apoyo, dedicación e infinita paciencia. Sin ellas este trabajo no hubiera sido posible.

Me gustaría extender este agradecimiento a todos los miembros del MIDAS, pasados y presentes, que han sido una valiosa ayuda y con los que he disfrutado enormemente trabajando.

Finalmente, recordar al resto de las personas que han estado a mi lado durante todo este tiempo; no hace falta mencionarlos, ellos saben quienes son.

Resumen

Desde el punto de vista de la cognición humana quizás una de las habilidades más significativas sea la capacidad de identificar y agrupar patrones visuales en elementos o conceptos abstractos significativos. Este proceso de agrupación perceptiva es la base de la capacidad para comprender y catalogar la información presente en las imágenes. Con el desarrollo de la informática, se plantea la posibilidad de extender los conceptos que rigen la percepción humana a sistemas computacionales, en lo que se conoce generalmente como anotación semántica de imágenes. Los métodos de anotación semántica de imágenes parten de características visuales de bajo nivel con las que se entrena un modelo con el que anotar nuevas imágenes.

La capacidad de inferencia de este modelo de clasificación depende de la representatividad de las características visuales con las que se ha entrenado, en relación al problema que se trata. Características locales asociadas a objetos en la imagen son mucho más representativas que características globales extraídas de la imagen completa. Sin embargo, el proceso de segmentación de una imagen representa un problema en sí mismo. Sencillamente no es posible asegurar que los resultados de esta sean adecuados para su uso como base para extraer características.

En este contexto, se propone como objetivo general de la Tesis el estudio, desarrollo y aplicación de diferentes algoritmos para abordar problemas relacionados con la anotación semántica de imágenes, trabajando a un nivel de particionado de la imagen inferior al de una segmentación, una sobre-segmentación. Una sobre-segmentación se define como un particionado bajo un criterio de uniformidad perceptual, sin buscar un significado semántico específico. Conocidos comúnmente como superpíxeles, el resultado de una sobre-segmentación respeta las fronteras presentes en las imágenes, produce elementos perceptualmente significativos y es más sencilla de generar que una segmentación. En consecuencia, un superpíxel es un candidato adecuado como elemento base para la extracción de características. En particular, nos centraremos en la caracterización de imágenes a nivel de sobre-segmentación como base para la anotación semántica, planteando como casos de estudio problemas dentro de los dominios de imagen médica (bidimensional y volumétrica) e imagen óptica en teledetección (multiespectral e hiperespectral).

Abstract

From the perspective of human cognition, perhaps one of the most significant abilities is the capacity of identifying and grouping visual patterns into significant abstract elements or concepts. This process of perceptual grouping is the basis of the capacity to understand and catalog the information present in the images. With the development of informatics, the possibility of extending the concepts that govern human perception to computational systems is proposed, in what is generally known as semantic image annotation. The semantic image annotation methods train a model to annotate new images from low-level visual features.

The inference capacity of this classification model depends on the representativeness of the visual features used for training. Local features associated with objects in the image are much more representative than global features extracted from the whole image. However, the process of segmentation of an image represents a problem itself. It is just not possible to guarantee that the results of a segmentation are suitable for use as a basis for extracting features.

In this context, the main objective of the Thesis is the study, development, and application of different algorithms to address problems related to semantic image annotation, working at a lower image partitioning level than a segmentation, an over-segmentation. An over-segmentation is defined as a partitioning under a criterion of perceptual uniformity, without looking for a specific semantic meaning. Commonly known as superpixels, the result of an over-segmentation respects the boundaries present in the images, generates perceptually significant elements and is easier to produce than a segmentation. Consequently, a superpixel is a convenient candidate as the base element for the extraction of features. In particular, we will focus on the characterization of images at the level of over-segmentation as the basis for semantic annotation, presenting as case studies problems within the domains of medical image (two-dimensional and volumetric) and remote sensing optical image (multispectral and hyperspectral).

Índice

Agradecimientos	VII
Resumen	IX
Abstract	XI
1. Introducción	1
1.1. Motivación	4
1.2. Descripción del problema	5
1.3. Objetivos	7
1.4. Estructura de la Tesis	8
1.5. Contribuciones	9
2. Estado del Arte	11
2.1. Preprocesado para anotación semántica	12
2.2. Extracción de características para anotación semántica	15
2.2.1. Características espectrales	17
2.2.2. Características texturales	18
2.2.3. Características basadas en forma	20
2.2.4. Características basadas en el entorno	21
2.3. Algoritmos de clasificación para anotación semántica	22
2.3.1. Algoritmos de clasificación generativos	22
2.3.2. Algoritmos de clasificación discriminativos	24
2.3.3. Clasificación perezosa	27
2.4. Segmentación de imágenes	28
2.4.1. Métodos ascendentes	29
2.4.2. Métodos descendentes	31
2.4.3. Métricas para la evaluación de la segmentación de imágenes	32
3. Extracción de características basadas en sobre-segmentación de imágenes	35

3.1. Superpíxeles	36
3.1.1. Algoritmos para la generación de superpíxeles	37
3.2. Caracterización de superpíxeles para anotación semántica . . .	42
3.2.1. Caracterización de superpíxeles mediante sus momentos estadísticos	43
3.2.2. Características de forma aplicadas a superpíxeles . . .	45
3.2.3. Matriz de co-ocurrencia de grises y medidas derivadas . . .	54
4. Segmentación de imágenes a color utilizando un índice de rugosidad basado en superpíxeles	61
4.1. Materiales y métodos	62
4.1.1. Conjuntos aproximados	63
4.1.2. Histón	64
4.1.3. Índice de rugosidad	64
4.1.4. Segmentación mediante un índice de rugosidad basado en superpíxeles	65
4.2. Resultados	67
5. Anotación semántica de imágenes médicas basadas en sobre-segmentación	73
5.1. Detección de mitosis en imágenes histopatológicas de cáncer de mama.	74
5.1.1. Materiales y métodos	75
5.1.2. Resultados	78
5.2. Detección de espinas dendríticas en microscopía confocal. . .	80
5.2.1. Materiales y métodos	82
5.2.2. Resultados	86
5.3. Clasificación de volúmenes MRI para detección de Alzheimer . . .	87
5.3.1. Materiales y métodos	90
5.3.2. Resultados	96
5.3.3. Información demográfica de las diferentes divisiones usados en el proceso de validación	102
5.3.4. Resultados de las evaluaciones individuales para cada división de datos	105
6. Anotación semántica de imagen de teledetección basada en sobre-segmentación	107
6.1. Segmentación de imagen multiespectral para teledetección. . .	108
6.1.1. Materiales y Métodos	109
6.1.2. Resultados	113
6.2. Clasificación en imagen hiperespectral.	126
6.2.1. Materiales y métodos	127

6.2.2. Resultados	131
7. Discusión y Conclusiones	139
7.1. Conclusiones	140
7.2. Desarrollo futuro	143
Bibliografía	145

Índice de figuras

1.1. Problemas asociados a la anotación semántica de imágenes: Detección de objetos (personas) (imagen superior izquierda (Fu et al., 2018)). Reconocimiento de objetos en una escena urbana (imagen superior derecha, (Garcia Freitas et al., 2018)). Segmentación de una tomografía computarizada de tórax (imagen inferior izquierda, imagen original parte de la "RIDER Lung CT collection" (Clark et al., 2013)). Anotación semántica de una escena urbana (imagen inferior derecha (Yasrab, 2018)).	3
1.2. Representación del intervalo semántico en el contexto del análisis de imágenes.	4
1.3. Diferentes segmentaciones manuales para una misma image, hechas por diferentes operadores (Imágenes procedentes del "The Berkeley Segmentation Dataset and Benchmark" (Martin et al., 2001)).	6
2.1. Algunos ejemplos de problemas asociados a la captura de imágenes. En la fila superior se puede ver un ejemplo de ruido gaussiano, artefacto por metal en un CT parte de la "RIDER Lung CT collection" (Clark et al., 2013), y un pliegue en la preparación de un estudio histológico (Rastogi et al., 2013). En la fila inferior se ve un ejemplo de un artefacto por movimiento en un MRI (Bier et al., 2018), sobre-exposición en una fotografía (Yoo et al., 2015) y efecto de "bandeado" en una imagen de satélite (Song et al., 2018).	14
2.2. Skewness; diferentes tipos de distribución.	18
2.3. Ejemplo de distribuciones asociadas a diferentes tipos de curtosis: A) leptocúrtica B) mesocúrtica C) platicúrtica D) uniforme.	19
2.4. Segmentación de imágenes; esquema de organización de los diferentes métodos.	29

3.1. Conjunto de imágenes de prueba seleccionado de la "VisTex database". En cada imagen se muestra el nombre que se utilizará para denominarla y sus valores globales de correlación (corr.) y contraste (contr.).	43
3.2. Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la media calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación la intensidad media de cada imagen (M.).	46
3.3. Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la desviación estándar calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación la desviación estándar de cada imagen (D.E.).	47
3.4. Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la asimetría calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación la asimetría de cada imagen (A.).	48
3.5. Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la curtosis calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación la curtosis media de cada imagen (C.).	49
3.6. Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la solidez calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto.	51

- 3.7. Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando el número de Euler calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. 52
- 3.8. Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la excentricidad calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. 53
- 3.9. Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la entropía calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación la entropía media de cada imagen (E.). 56
- 3.10. Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la homogeneidad calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación el contraste medio para cada imagen de las diferentes pruebas (Hom.). 57
- 3.11. Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando el contraste calculado a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación el contraste medio para cada imagen de las diferentes pruebas (Ctr.). 58

3.12.	Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la correlación calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación la correlación media para cada imagen de las diferentes pruebas (Cor.).	59
4.1.	Diagrama de flujo del método desarrollado para segmentación de imágenes mediante un índice de rugosidad basado en superpíxeles.	65
4.2.	Conjunto de imágenes utilizados como prueba.	68
4.3.	Segmentación de imágenes a color. Ejemplo de resultados para 10 segmentos, para las imágenes Im1, Im5, Im11, Im14 e Im17; la primera columna muestra una segmentación mediante histogramas, la segunda el resultado del método original mediante índice de rugosidad y la tercera el producto de una segmentación mediante índice de rugosidad basado en superpíxeles.	72
5.1.	Esquema general del proceso de detección de mitosis propuesto.	75
5.2.	Proceso de selección de superpíxeles candidatos. Detalle de la imagen normalizada original (izquierda) con la segmentación mediante superpíxeles obtenida a partir de su representación HE correspondiente superpuesta, imagen de representación/-relación azul asociada (centro) y conjunto de candidatos a mitosis (derecha) destacados en color.	76
5.3.	Ejemplos de ocurrencias mitóticas. El superpíxel correspondiente al núcleo mitótico aparece marcado en tonos rojizos, los segmentos circundantes asociados están resaltados en tonos verdosos.	77
5.4.	Algunos ejemplos de clasificación de mitosis. Verdaderos positivos contorneados en verde, falsos positivos contorneados en rojo y falsos negativos sin contorno.	78
5.5.	"Slice"(izquierda) y reconstrucción volumétrica (derecha) de una neurona captada mediante microscopia confocal.	81
5.6.	Segmentación supervóxel de una pila de imágenes dendrita (slice, detalle). Los colores carecen de significado especial.	83
5.7.	Vector de suma en el método de regularidad del vecindario local (proyección máxima). Supervóxeles azulados indican un vector más corto, rojizo significa vectores más largos.	83

5.8. Árbol de mínima expansión con el gráfico final que representa el gráfico de vástago superpuesto. La escala está en vóxeles. . .	84
5.9. Supervóxeles en dendritas etiquetados con el número de trayectos más largos en un árbol de extensión mínima del que forma parte. Un tono azulado indica un supervóxel con menor número de caminos, un tono rojizo uno con más caminos. . .	85
5.10. Distancias de supervóxel al tronco dendrítico en el método del gráfico de distancia del vástago dendrítico (proyección máxima). Un supervóxel azulado indica un elemento más cercano, rojizo significa más lejano.	85
5.11. Imagen dendrítica con el lomo dendrítico agrupado resultados del método de regularidad del vecindario local superpuesto. Proyección máxima (izquierda) y reconstrucción de volumen (derecha). Los colores no tienen un significado especial. . . .	86
5.12. Proyección máxima de los resultados de la detección de la columna dendrítica agrupada (arriba) y de la detección mediante la distancia (abajo) del gráfico del vástago dendrítico. Las diferencias entre los métodos se pueden ver en los resultados de las áreas mal definidas (1 y 2) y la columna dendrítica cerca del tallo dendrítico (3, 4, 5 y 6) aparecen resaltados. . .	87
5.13. Muestra de diferentes resultados de detección de espinas dendrítica agrupada (Detalles). Las detecciones de espinas dendríticas parciales por los resultados del método de regularidad del vecindario local (fila superior) se emparejan con los resultados de la detección por el método de distancia del gráfico del tallo dendrítico (fila inferior).	88
5.14. Esquema del método propuesto para la clasificación de volúmenes MRI.	91
5.15. Cortes de volumen, seleccionados del conjunto de datos OASIS-1 preprocesado, mostrando ejemplos de segmentación de tejidos. Cortes originales del MRI (izquierda), probabilidades de materia gris (centro izquierda), probabilidades de materia blanca (centro derecha), líquido cefalorraquídeo (derecha). . .	93
5.16. Corte de volumen del conjunto de datos OASIS-1 preprocesado, mostrando las probabilidades de segmentación del tejido y su corte agregado correspondiente como una imagen RGB. Probabilidades de materia gris (izquierda), probabilidades de materia blanca (centro izquierda), líquido cefalorraquídeo (centro derecha) y agregado (derecha) (probabilidades de materia gris como azul, probabilidades de materia blanca como verde, líquido cefalorraquídeo como rojo.	94
5.17. Gráfica de codo representando los valores propios de los 30 primeros componentes.	99

- 6.1. Resultados de la segmentación para imagen CHILE1 (fila superior) y para CANADA1 (fila inferior). Primera columna: Representación en falso color (infrarrojo cercano/verde/rojo). Los resultados del método original y del método propuesto para 15 conglomerados finales se muestran en la segunda y tercera columna, respectivamente. Los colores no tienen un significado especial. 117
- 6.2. Resultados de la segmentación para la imagen CANADA2 (fila superior) y para MARYLAND (fila inferior). Primera columna: Representación en falso color (infrarrojo cercano/verde/rojo). Los resultados del método original y del método propuesto para 15 conglomerados finales se muestran en la segunda y tercera columna, respectivamente. Los colores no tienen un significado especial. 118
- 6.3. Fila superior: Detalle de la esquina superior derecha de la imagen CANADA2. Posición del extracto (imagen izquierda); resultados del método original (imagen central) y del método propuesto (imagen derecha). Fila inferior: detalles de la zona del campo de golf de la imagen MARYLAND. Posición del extracto (imagen izquierda); resultados del método original (imagen central) y del método propuesto (imagen derecha). Los colores no tienen un significado especial. 120
- 6.4. Comparación de métodos de fusión de regiones. Imagen CHILE2 en la fila superior (15 agrupaciones espectrales finales); imagen ARGENTINA (10 agrupaciones espectrales finales) en la fila inferior. En la primera columna se presenta una representación en falso color de las imágenes de prueba (infrarrojo cercano/verde/rojo). Los resultados para el método de combinación de regiones de distancia de color y co-ocurrencia en el método de combinación de región de superpíxeles se muestran en la segunda y tercera columna, respectivamente. Los colores no tienen un significado especial. 121
- 6.5. Comparación de métodos de fusión de regiones; 5 clusters espectrales finales. Fila superior: imagen CHILE2. Distancia de color (imagen izquierda) y co-ocurrencia en superpíxeles (imagen derecha). Fila central: imagen ARGENTINA. Distancia de color (imagen izquierda) y co-ocurrencia en superpíxeles (imagen derecha). Los colores no tienen un significado especial. Fila inferior: imagen MARYLAND. Distancia de color (imagen izquierda) y co-ocurrencia en superpíxeles (imagen derecha). Los colores no tienen un significado especial. 123

6.6. Comparación de métodos de combinación de regiones. La co- ocurrencia en superpíxeles da como resultado la caracteriza- ción de la columna izquierda y la caracterización de superpíxe- les en la columna derecha. Fila superior: imagen CHILE2 (10 clusters espectrales finales). Fila central: imagen CANADA2 (10 clusters espectrales finales). Fila inferior: imagen CHILE1 (10 clusters espectrales finales). Los colores no tienen un sig- nificado especial.	124
6.7. Comparación de métodos de combinación de regiones. La co- ocurrencia en superpíxeles da como resultado la caracteriza- ción de la columna izquierda y la caracterización de superpíxe- les en la columna derecha. Fila superior: imagen ARGENTI- NA (15 clusters espectrales finales). Fila central: imagen CA- NADA1 (10 clusters espectrales finales). Fila inferior: imagen MARYLAND (10 clusters espectrales finales). Los colores no tienen un significado especial.	125
6.8. Esquema del método desarrollado para la clasificación de imá- genes hiperespectrales.	128
6.9. Escena Indian Pines. Falso color (banda 125 a rojo, 70 a ver- de y 30 a azul) (imagen superior izquierda). Verdad terreno (superior derecha) y resultados del método propuesto (inferior derecha).	132
6.10. Escena PaviaU. Falso color (banda 80 a rojo, 28 a verde y 9 a azul) (imagen izquierda). Verdad terreno (centro) y resultados del método propuesto (derecha).	133
6.11. Escena Salinas. Falso color (banda 125 a rojo, 70 a verde y 30 a azul) (izquierda). Verdad terreno (centro) y resultados del método propuesto (derecha).	134

Índice de Tablas

3.1. Coeficiente de correlación de Pearson entre los diferentes momentos estadísticos extraídas y los parámetros de la sobresegmentación. Se presenta la media para el grupo de imágenes de prueba y mayores y menores valores obtenidos.	44
3.2. Coeficiente de correlación de Pearson entre las diferentes características de forma extraídas y los parámetros de la sobresegmentación. Se presenta la media para el grupo de imágenes de prueba y mayores y menores valores obtenidos.	50
3.3. Coeficiente de correlación de Pearson entre las diferentes características de Haralick extraídas y los parámetros de la sobresegmentación. Se presenta la media para el grupo de imágenes de prueba y mayores y menores valores obtenidos.	55
4.1. Comparación del número de segmentos obtenidos mediante el método desarrollado con los resultados de un método de umbralización basado en histogramas y los resultados del método basado en el índice de rugosidad, tanto por canal de color como para el resultado final combinado.	69
4.2. Resultados del criterio de evaluación de Levine y Nazif para el algoritmo original, el método propuesto y una segmentación utilizando histogramas. Los mejores valores para cada criterio (Uniformidad intra-región, disparidad inter-regional, y valor agregado) y para cada segmentación aparecen resaltados en gris.	70
5.1. Resultados medios para la evaluación del proceso de detección de mitosis. Se muestra la sensibilidad, la especificidad, la precisión, el "F-score" y el coeficiente de correlación de Matthews (CCM).	79

5.2. Características demográficas de la población estudiada (base de datos OASIS-1). Los valores de la escala de clasificación de demencia clínica (CDR) se indican como el número de pacientes diagnosticados como demencia muy leve/demencia leve/demencia moderada. Los valores del Mini-Mental State Examination (MMSE) se indican como media $\pm$ desviación estándar.	96
5.3. Características demográficas del subconjunto de datos de mayores de 60 años (tabla superior) y del subconjunto de datos reducido (tabla inferior). El género se indica como masculino/-femenino. Los valores de la escala de clasificación de demencia clínica (CDR) se indican como el número de pacientes diagnosticados como demencia muy leve/demencia leve/demencia moderada. Los valores del Mini-Mental State Examination (MMSE) se indican como media $\pm$ desviación estándar.	97
5.4. Resultados de la evaluación que muestran una comparación entre las métricas de evaluación medias propuestas a partir del uso de diferentes "kernels" (lineal, polinomial y radial) para el clasificador SVM para el conjunto de datos de 60 años y más, para las 10 repeticiones. Los mejores resultados aparecen destacados en negrita.	100
5.5. Resultados de la evaluación que muestran una comparación entre la media de las métricas de evaluación de la línea de base de Clínica y el método propuesto (SVM entrenado usando un "kernel" lineal) para el conjunto de datos de los sujetos de 60 años o más (tabla superior) y para el conjunto de datos reducido (tabla inferior).	100
5.6. Matriz de confusión media para las pruebas de validación, que muestra la tasa de clasificación en porcentaje para el conjunto de datos de los sujetos de 60 años o más (tabla superior) y para el conjunto de datos reducido (tabla inferior). Las tablas de la izquierda corresponden a la línea de base de Clínica, las de la derecha corresponden al método propuesto ("kernel" lineal).	101
5.7. Información demográfica para cada división del conjunto de datos de 60 años o más. El sexo se indica como masculino/femenino, los valores de la Escala de Calificación de Demencia Clínica (CDR, por sus siglas en inglés) se indican como el número de pacientes diagnosticados como demencia muy leve/demencia leve/ demencia moderada. Los valores de Edad y Mini-Mental State Examination (MMSE) se indican como la media $\pm$ desviación estándar.	103

5.8. Información demográfica para cada división de prueba del conjunto de datos reducido. El sexo se indica como masculino/femenino, los valores de la Escala de Calificación de Demencia Clínica (CDR) se indican como el número de pacientes diagnosticados como demencia muy leve/demencia leve/demencia moderada. Los valores del Mini-Mental State Examination (MMSE) se indican como la media de la desviación estándar de $\pm$	104
5.9. Resultados de la evaluación que muestran una comparación entre las métricas de evaluación para cada división de prueba utilizando la línea de base de Clínica (tabla superior) y el método propuesto (SVM entrenado usando un "kernel" lineal) para el conjunto de datos de sujetos de 60 años o más (tabla inferior).	105
5.10. Resultados de la evaluación que muestran una comparación entre las métricas de evaluación para cada división de la prueba, utilizando la línea de base de Clínica (tabla superior) y el método propuesto (SVM entrenado usando un "kernel" lineal) para el conjunto de datos reducido (tabla inferior).	106
6.1. Comparación del número de conglomerados espectrales producidos por el método propuesto con los resultados de un método de umbralización basado en histogramas y los resultados del método basado en el índice de rugosidad, tanto por banda espectral como por el resultado final combinado.	114
6.2. Los resultados de los criterios de evaluación de Levine y Nazif para el algoritmo original y el método propuesto, ordenados por métodos de segmentación y fusión de regiones (donde C. Dist es la distancia de color, Cooc. Sup. representa la co-ocurrencia en el método de superpíxeles y Sup. Ch. es el método de caracterización de superpíxeles). Los mejores resultados para cada criterio (Uniformidad intra-región, disparidad interregional, y los valores agregados) para cada segmentación aparecen resaltados en gris oscuro. Los mejores resultados teniendo en cuenta sólo el método de segmentación también se muestran en gris claro.	115
6.3. Métricas de precisión por clase para la escena PaviaU Ratio de verdaderos positivos (TP), ratio de falsos positivos (FP), precisión, exhaustividad y valor-F. Los resultados para la caracterización de la escena original se presentan en "C. Intens.", los resultados del volumen representando las tasas de cambio en "C. Tasas", y el vector agregado de las dos caracterizaciones en "V. agregado".	131

6.4.	Métricas de precisión por clase para la escena PaviaU Ratio de verdaderos positivos (TP), ratio de falsos positiva (TF), precisión, exhaustividad y valor-F.	135
6.5.	Matriz de confusión para la escena PaviaU.	135
6.6.	Métricas de precisión por clase para la escena Salinas. Ratio de verdaderos positivos (TP), ratio de falsos positiva (TF), precisión, exhaustividad y valor-F.	136
6.7.	Matriz de confusión para la escena Salinas.	136
6.8.	Métricas de precisión por clase para la escena Indian pines, Ratio de verdaderos positivos (TP), ratio de falsos positiva (TF), precisión, exhaustividad y valor-F.	137
6.9.	Matriz de confusión para la escena Indian Pines.	137

Capítulo 1

Introducción

Desde el punto de vista de la cognición humana, quizás una de las habilidades más significativas sea la capacidad de localizar rápidamente patrones visuales y agruparlos automáticamente en elementos o conceptos abstractos significativos. Así, partiendo del trabajo presentado en (Palmer, 1999), podemos definir el concepto de agrupamiento o asociación perceptual como la capacidad visual del ser humano para abstraer información de alto nivel a partir de una estructura de bajo nivel, una imagen, sin ningún conocimiento específico del contenido de ésta. Este proceso de agrupación perceptiva es la base de la capacidad para comprender y catalogar la información presente en estas imágenes (Elmqvist y Fekete, 2010).

Los avances tecnológicos de los últimos años han permitido plantear la posibilidad de extender los conceptos que rigen la percepción humana a sistemas computacionales, dando como resultado la oportunidad de realizar automáticamente tareas asociadas al procesamiento de la información visual de las imágenes tales como la indexación, la identificación o el reconocimiento de los elementos que representa. Aunque desde una perspectiva humana el proceso de agrupación jerárquica pueda parecer trivial, su automatización, su implementación en un programa informático, no es un proceso trivial.

Las diferentes tareas asociadas a la interpretación de información visual en una imagen pueden englobarse bajo el concepto general de anotación automática de imágenes. La anotación automática de imágenes (AIA, Automatic Image Annotation) es el proceso por el cual, a partir de un conjunto de características extraídas sobre una imagen se asocian imágenes no etiquetadas con un conjunto de palabras/conceptos que describen su contenido usando técnicas de aprendizaje automático (Smeulders et al., 2000). Si bien no existe una división formal de los diferentes problemas asociados al proceso de anotación automática de imágenes, atendiendo a la nomenclatura usada comúnmente en la literatura (p. ej. en (Sermanet et al., 2013)) es posible categorizarlos en 5 tipos (se muestran algunos ejemplos de estos problemas en Fig. 1.1).

- Segmentación de imágenes: La segmentación se define como el proceso de dividir una imagen en N entidades significativas no solapadas ((Shapiro y Stockman, 2001)), sin depender de información externa a la propia imagen, y por tanto sin que, *a priori*, podamos asignar una clase específica a ninguna de las regiones.
- Detección de objetos: Similar al problema de reconocimiento de objetos, este caso plantea únicamente dos clases: objeto/no objeto. Por tanto, el resultado se compondrá de regiones etiquetadas como "objeto" y el resto de la imagen descartada o clasificada como "fondo". En principio, la presencia del objeto en la imagen está garantizada.
- Reconocimiento de objetos: El reconocimiento de objetos plantea la búsqueda de uno o varios objetos pertenecientes a un conjunto de posibles clases dentro de una imagen. No se garantiza la presencia de todas las clases dentro de una imagen. El resultado será un conjunto de áreas asociadas a la clase a la que pertenece el objeto detectado.
- Clasificación de imágenes: En la clasificación de imágenes, se asocia la imagen completa con una o varias clases, sin tener en cuenta si éstas describen la imagen o solamente una parte de esta.
- Anotación semántica: En un proceso de anotación semántica se asocia a cada píxel de la imagen una de las etiquetas de un conjunto de clases definidas por la semántica del problema tratado, dividiendo la imagen completamente en regiones diferentes.

La AIA puede considerarse un tipo de clasificación multicategoría. Como la mayoría de los procesos de clasificación, la anotación automática de imágenes se basa en la creación de un modelo supervisado, una conceptualización general del problema capaz de predecir el comportamiento de futuras instancias a partir de un conjunto de ejemplos suministrados previamente (Kotsiantis et al., 2007). Este modelo depende de un proceso de extracción de características, casi siempre asociado a una segmentación previa. Esta capacidad de inferencia está definida por la "calidad", la representatividad en relación al problema que se plantea resolver, de las características que sea posible extraer de forma automática de la imagen y que se usarán para entrenar dicho modelo.

Sin embargo, ¿cómo podemos definir qué características pueden considerarse representativas? ¿qué convierte a parte de una imagen en algo representativo? ¿qué significa ser "representativo"? Estas preguntas reflejan el problema subyacente en los procesos de anotación automática de imágenes y, en general, todo problema de modelización, el problema conocido como "intervalo semántico". Partiendo del trabajo presentado en (Hare et al., 2006), es posible definir el concepto de intervalo semántico como la diferencia entre

Figura 1.1: Problemas asociados a la anotación semántica de imágenes: Detección de objetos (personas) (imagen superior izquierda (Fu et al., 2018)). Reconocimiento de objetos en una escena urbana (imagen superior derecha, (Garcia Freitas et al., 2018)). Segmentación de una tomografía computarizada de tórax (imagen inferior izquierda, imagen original parte de la "RIDER Lung CT collection" (Clark et al., 2013)). Anotación semántica de una escena urbana (imagen inferior derecha (Yasrab, 2018)).

la representación de conocimiento contextual en su forma original o natural y su formalización dentro de un modelo. Si bien no es un fenómeno exclusivo de la anotación automática de imágenes, la extracción de conocimiento a partir de una imagen representa uno de los ejemplos más claros del problema. Se parte de un conjunto de etiquetas de alto nivel que describen una imagen o partes de ella, "bosque", "lesión", "tumor", ..., sin embargo sólo es posible hacer asociaciones utilizando la información disponible en una estructura de bajo nivel, una matriz de píxeles, esto es, una imagen (una representación de este concepto puede verse en 1.2). Debe entenderse la magnitud del problema teniendo en cuenta que, para una matriz de píxeles, incluso conceptos

Figura 1.2: Representación del intervalo semántico en el contexto del análisis de imágenes.

básicos como forma o color implican métodos de formalización matemática específicos, que no tienen por qué ser ni intuitivos, ni únicos, ni robustos (Upadhyaya y Dixit, 2016).

1.1. Motivación

Reconocimiento de objetos, detección de objetos, anotación semántica... son procesos clave en infinidad de tareas de investigación, de recuperación de imágenes y de gestión de contenidos. Realizados manualmente por un operador, estos procesos implican interpretar y anotar una imagen a partir de sus conocimientos sobre la temática de las imágenes y su criterio personal. A pesar de la existencia de herramientas para facilitar su realización, como por ejemplo LabelMe¹ (Rapson et al.) o RectLabel², el proceso de etiquetado manual es posible únicamente cuando el problema implica un volumen reducido de imágenes, no solamente por ser un proceso laborioso, sino por la naturaleza subjetiva de este. Incluso en problemas con una semántica asociada limitada, como por ejemplo en la detección de tumores en imagen radiológica, diferentes expertos pueden etiquetar un mismo elemento de forma distinta. Por lo tanto, definir que es una anotación "correcta" es una tarea compleja, lo que hace que la anotación de imágenes no sea un problema bien definido.

Estos problemas han impulsado la investigación de diferentes métodos relacionadas con las diferentes tareas englobadas dentro del concepto de AIA. En estos métodos, se extraen características visuales de bajo nivel de las imágenes, con las que se entrena un modelo para anotar nuevas imágenes. Esto permite limitar los problemas asociados al etiquetado manual al subconjunto utilizado para la creación de un modelo de clasificación, en general mucho más pequeño que el que corresponde al problema total.

¹<https://wikivisually.com/wiki/LabelMe>

²<https://rectlabel.com/>

Como ya se ha mencionado, la representatividad del modelo de clasificación asociado a un problema de AIA depende de las características visuales con las que se ha entrenado y, como se señala en (Hentschel et al., 2007), se da por hecho que características locales asociadas a objetos en la imagen son mucho más representativas que características globales extraídas de la imagen completa, observación que parece confirmarse en estudios como (Shyu et al., 1998), (Lisin et al., 2005), (Rabinovich et al., 2007) o (Liu y Xia, 2010). Se puede ver en (Guo et al., 2007) como las características extraídas a nivel de objeto reducen la variabilidad espectral de las clases; por otro lado, puesto que un objeto representa un área definida, es posible extender el tipo de característica extraíbles a características basadas en propiedades espaciales, texturales y contextuales, hecho que permite potencialmente mejorar la representatividad del modelo de clasificación. Por tanto, el proceso de obtener los objetos de una imagen, la segmentación, se presenta como una fase decisiva dentro de los diferentes procesos englobados bajo el concepto de AIA.

1.2. Descripción del problema

La extracción de características partiendo de objetos se encuentra con el obstáculo de depender de una segmentación adecuada de la imagen, lo que representa un problema en sí mismo. El problema del intervalo semántico también está presente los procesos de segmentación. De nuevo, qué se considera relevante, representativo o sencillamente un objeto, depende del contexto y el observador. Esto queda patente en Fig. 1.3, donde se muestran varias segmentaciones manuales de una imagen de la "Berkeley Segmentation Dataset and Benchmark" (Martin et al., 2001), realizadas por diferentes operadores humanos. Puede verse como, por ejemplo, las formas de dividir la montaña y el cielo varían de forma significativa, lo que implica que la percepción de la escena difiere entre diferentes sujetos, sin que se pueda realmente decir que ninguna de ellas es incorrecta. Por lo tanto, definir que es una segmentación "correcta" es una tarea compleja, lo que hace que la anotación de imágenes no sea un problema bien definido. Citando a (Fu y Mui, 1981): "el problema de segmentación de la imagen es básicamente un problema de percepción psicofísica y, por lo tanto, no es susceptible de una solución puramente analítica". Se plantea por tanto la siguiente pregunta: ¿Es posible extraer características de una imagen a nivel local sin depender de un particionado en objetos?

En este contexto, se propone como objetivo general de la Tesis el estudio, desarrollo y aplicación de diferentes algoritmos para abordar problemas relacionados con la anotación automática de imágenes, trabajando a un nivel de particionado de la imagen inferior al de una segmentación, una sobre-segmentación.

Figura 1.3: Diferentes segmentaciones manuales para una misma image, hechas por diferentes operadores (Imágenes procedentes del "The Berkeley Segmentation Dataset and Benchmark" (Martin et al., 2001).

Se puede definir una sobre-segmentación como el proceso por el cual los objetos presentes en una imagen son a su vez divididos o fragmentados en sub-objetos bajo un criterio de uniformidad perceptual (Eggleson, 1998), en principio sin significado semántico específico. Una sobre-segmentación respeta las fronteras presentes en las imágenes, produce elementos perceptualmente significativos (Achanta et al., 2012a) y es más sencilla de generar que una segmentación (Wang et al., 2014). Lo que es más, conocidos comúnmente como superpíxeles, el resultado de una sobre-segmentación codifica directamente algunos de los principios básicos en los que se basa la percepción humana (Humphrey, 1924), como son los de semejanza y proximidad (Michaelson y Arens, 2017) y representa un objeto útil de cara a implementar el resto de estos principios, como puede verse en (Ren y Malik, 2003a). Un superpíxel es, en consecuencia, un candidato adecuado como elemento base para la extracción de características.

En esta Tesis, se aborda por tanto el estudio de la sobre-segmentación de imágenes como herramienta para la extracción de características para anotación semántica. De esta manera, se estudiará el comportamiento de la sobre-segmentación de imágenes como herramienta para la extracción de

características, y a partir de esta, se desarrollarán nuevos métodos de caracterización para diversos procesos asociados a la anotación automática de imágenes. En particular, se plantearán como casos de estudio problemas dentro de los dominios de la imagen médica (bidimensional y volumétrica) y la imagen óptica de teledetección (pancromática, multiespectral e hiperespectral).

Si bien puede parecer que estos dos dominios, imagen médica e imagen de teledetección, tienen muy pocos puntos en común, ambos presentan propiedades que los hacen muy interesantes en el estudio de la segmentación para anotación automática de imágenes. La semántica asociada a estos dominios está centrada en sus respectivos ámbitos, siendo en general limitada y, hasta cierto punto, no-ambigua, lo que nos permite centrarnos en la extracción de características, no en el proceso de clasificación. Además, puesto que ambos dominios incluyen imagen multidimensional, ya sea como imagen volumétrica (multidimensional en el plano espacial) en el contexto médico, ya sea como imagen multiespectral (imagen multidimensional en el plano espectral) en el contexto de teledetección, podemos abordar la extensión de este trabajo más allá de la imagen bidimensional. Por último, ambos dominios abarcan un amplio rango de problemas diferentes, en muchos casos aún pendientes de resolución; los casos de uso con los que se realizará la evaluación representan problemas reales.

Definido como el proceso de asignar a cada píxel una clase especificada por la semántica del problema a tratar, la anotación semántica representa dentro del conjunto de tareas asociadas a la anotación automática de imágenes el caso más general. Esta Tesis se centrará por tanto en el caso de la anotación semántica de imágenes. Sin embargo, puesto que trabajamos con problemas reales con necesidades específicas también se presentarán, en los capítulos 5 y 6, ejemplos de otras tareas asociadas a la AIA.

1.3. Objetivos

La consecución del objetivo global de esta Tesis conlleva superar los siguientes objetivos parciales:

- Revisar y clasificar las actuales técnicas de anotación automática de imágenes con especial énfasis en los algoritmos de segmentación y sobre-segmentación en los que se basan.
- Determinar y evaluar un conjunto de descriptores de bajo nivel a partir de una segmentación basada en superpíxeles/supervóxeles que permitan la adecuada caracterización de los objetos contenidos en una imagen. En el caso de ser estos descriptores cualitativos, desarrollo de un vocabulario adecuado a una segmentación basada en superpíxeles/supervóxeles.

- Proponer y desarrollar nuevas características/descriptores en imagen bidimensional y volumétrica basadas en sobre-segmentación mediante superpíxeles/supervóxeles.
- Validar los descriptores propuestos en escenarios reales, a partir de un conjunto seleccionado de casos de uso en imagen médicas (bidimensional y volumétrica) e imagen de teledetección (pancromáticas, multiespectrales e hiperespectrales).

1.4. Estructura de la Tesis

Esta Tesis está organizada de la siguiente manera:

- En el capítulo 2 se presenta un estudio exhaustivo sobre las diferentes técnicas asociadas al dominio de la anotación semántica de imágenes, siguiendo la metodología común a los diferentes trabajos publicados en este tema. Por lo tanto se presentará una visión general de las diferentes técnicas de pre-procesado de imágenes, las características de bajo nivel más habituales que pueden extraerse de una imagen, y los modelos de clasificación que han sido utilizados en problemas de anotación automática de imágenes, así como de los métodos más comunes usados para segmentación de imágenes y las bases sobre cómo evaluar los resultados de ésta.
- El capítulo 3 presenta el concepto de sobre-segmentación mediante superpíxeles/supervóxeles y desarrolla su uso como primitiva base para la extracción de características de bajo nivel en imágenes.
- El capítulo 4, como extensión del uso de superpíxeles como elemento mínimo en una imagen, se propone una extensión al concepto de índice de rugosidad basada en superpíxeles para la segmentación de imágenes.
- En el capítulo 5 se presentan los diferentes algoritmos desarrollados para la anotación semántica de imágenes médicas a partir de sobre-segmentación basada en superpíxeles, comenzando con una introducción a las características específicas de estos tipos de imagen. Como casos de estudio se plantea la detección de espinas dendríticas en imágenes de microscopia confocal y la detección de mitosis en imágenes histopatológicas de cáncer de mama, ambos casos representativos de un problema de detección de objetos. Además también se abordará la clasificación de volúmenes MRI para la detección de la enfermedad de Alzheimer, siendo este caso un problema de clasificación de imagen (volumétrica).
- En el capítulo 6 se presentan los diferentes algoritmos desarrollados para la anotación semántica de imagen de teledetección a partir de

sobre-segmentación basada en superpíxeles, tras un resumen de las características que definen este tipo de imágenes. Como casos de estudio se utilizará la segmentación de imágenes multiespectrales de satélite a partir de índice de rugosidad basado en superpíxeles (un problema de segmentación) y la generación de un mapa de coberturas a partir de imagen hiperespectral, un problema típico de anotación semántica.

- El capítulo 7 discute y concluye los resultados generales y las contribuciones del trabajo presentado en los capítulos anteriores. Finalmente se exponen algunas de las posibles líneas a seguir para dar continuidad a este trabajo.

1.5. Contribuciones

Esta Tesis aporta una serie de contribuciones al campo de la segmentación semántica de imágenes. Se detallan brevemente a continuación.

- Se ha realizado un estudio en profundidad del comportamiento de diferentes características de color, forma y textura extraídas partiendo de una sobre-segmentación mediante superpíxeles.
- Se ha desarrollado un método para la segmentación de imagen a color, una extensión de la propuesta presentada en (Mushrif y Ray, 2008), basada en índice de rugosidad. En este método se presentan los conceptos de histón e índice de rugosidad basados en superpíxeles, donde se utilizan como forma de adaptar lo que se considera el vecindario de un píxel a las características específicas de la imagen,
- Se ha implementado un método automático para la detección de mitosis en imagen histopatológica de cáncer de mama basado en la combinación de diferentes tipos de características sencillas extraídas a partir de una sobre-segmentación mediante superpíxeles. Una extensión de este método se ha presentado en (Ortiz et al., 2017b).
- Se han implementado dos métodos para identificar y reconstruir espinas dendríticas en imágenes de microscopía confocal. Estos métodos aprovechando las posibilidades que ofrecen las caracterizaciones realizadas a partir del entorno y la posición de un superpíxel. Este trabajo ha sido presentado en (Ortiz et al., 2014).
- Se ha desarrollado un método para identificar pacientes con la enfermedad de Alzheimer a partir de volúmenes MRI. Este método parte del concepto de histón basado en superpíxeles. Así, se implementa la extensión de dicho concepto a volúmenes a color y se plantea su utilización como elemento de caracterización textural. Este trabajo se ha publicado en (Ortiz et al., 2018).

- Se ha desarrollado un método para la segmentación de imágenes multiespectrales, una extensión del método de segmentación de imágenes a color visto anteriormente. Esta extensión implica la adaptación del cálculo del índice de rugosidad basado en superpíxeles al dominio multiespectral. Así mismo, como parte del desarrollo, se presentan dos métodos diferentes para cuantificar relaciones entre regiones en una imagen, basados en también en una sobre-segmentación mediante superpíxeles. Este trabajo ha sido publicado en (Ortiz et al., 2015)
- Se ha implementado un método de caracterización de imágenes para la generación de mapas temáticos a partir de volúmenes hiperespectrales, utilizando el concepto de índice de rugosidad basado en superpíxeles, con la idea de explotar la relación espacial/espectral presente en este tipo de imágenes. Este trabajo se ha presentado en (Ortiz et al., 2017a).

Capítulo 2

Estado del Arte

En su definición más básica, una imagen bidimensional es una matriz numérica. Cada posición de esta matriz se conoce como un píxel, generalmente considerado la unidad de análisis más pequeña dentro de una imagen. La equivalencia en el mundo real de un píxel vendrá determinado por los modelos de geometría del sensor y de la escena. Un píxel no tiene significado por sí mismo, pero sus agrupaciones perceptuales forman una de las bases del sistema visual humano, los objetos. En (Bharathi y Subashini, 2011) se utiliza la siguiente definición de objeto en el contexto de una imagen: "Un objeto [...] es una porción identificable de una imagen que puede ser interpretada como una sola unidad". Esta definición abarca desde un conjunto de píxeles representando un objeto del mundo real hasta un segmento resultado de una sobre-segmentación (como conjunto de píxeles visualmente homogéneos) y, a su vez, nos da una idea de lo complejo que puede ser definir conceptos que, en principio, parecen intuitivos, y por tanto lo crucial que puede llegar a ser un proceso de segmentación.

La segmentación de imágenes, en general realizada antes del paso de extracción de características y después de otros procesos previos de procesamiento, es un paso casi imprescindible en una anotación semántica. Como se ha mencionado previamente, una representación adecuada de las características presentes en una imagen mejorará significativamente el rendimiento de las técnicas de anotación semántica. Específicamente, la ventaja de un enfoque basado en segmentación, frente a uno basado en píxeles, permite la posibilidad no sólo de obtener múltiples vectores de características de cada imagen, sino también el incorporar información contextual, como medidas de complejidad espacial de áreas o relaciones entre diferentes objetos (Blaschke et al., 2004). Además, la extracción de características a nivel de objeto reducirá la variabilidad interna intrínseca a cualquier imagen, ofreciendo una descripción mucho más precisa del contenido de la imagen.

En cualquier caso, el proceso para asociar un significado semántico a este conjunto de objetos depende tanto de las características específicas de la

imagen como del conjunto de términos semánticos que incluye el vocabulario que se quiera asociar a la imagen. En su forma más general, este proceso implica tres fases, preprocesado de la imagen (donde se puede incluir el proceso de segmentación), extracción de características y clasificación. El primer paso en un proceso de anotación semántica es el preprocesamiento. Los detalles específicos de este proceso dependen en gran medida de los tipos de imágenes utilizadas en cada caso de estudio. Sobre la imagen/conjunto de imágenes preprocesadas se calcula un conjunto de características. En la fase de clasificación el/los conjuntos de características extraídos se utilizarán para crear un modelo de clasificación, que se usará para etiquetar nuevas muestras.

Cada una de las fases en las que se puede dividir la anotación semántica representa un problema en sí mismo, cada uno con su propio conjunto de técnicas y algoritmos asociado. Dependiendo de la naturaleza del problema que se trata de resolver, un proceso de anotación semántica utilizará una combinación u otra de estos. En este capítulo repasaremos las principales técnicas asociadas a cada una de las etapas del proceso, señalando algunos ejemplos de su uso en tareas de anotación semántica, con especial énfasis en el proceso de segmentación.

2.1. Preprocesado para anotación semántica

Preprocesado (o preprocesamiento) es el nombre común para el conjunto de operaciones realizadas a las imágenes obtenidas directamente de un sensor con el objetivo de prepararlas para su análisis e interpretación. El preprocesado de una imagen implica generalmente la eliminación o minimización de elementos no deseados dentro la imagen, por ejemplo, ruido, fallos del sensor, partes desenfocadas, etc. Si bien la naturaleza destructiva de muchas de las técnicas de preprocesado pueden plantear dudas sobre su aplicación, el uso adecuado de estas puede dar como resultado una mejor extracción de características tanto a nivel local como global, como se plantea en (Krig, 2016). Podemos ver en (Tushabe y Wilkinson, 2007) como, dependiendo del problema y el ámbito al que esté asociada la imagen, será necesario elegir qué aspectos y partes de la imagen pueden ser ignorados, mejorados o filtrados.

Atendiendo a su propósito, podemos dividir las operaciones de preprocesamiento en aquellas realizadas para corregir defectos de la imagen que puedan afectar a procesos posteriores y aquellas aplicadas con el objetivo de realzar características que nos interesa aprovechar en pasos posteriores. En la Fig. 2.1 podemos ver algunos ejemplos típicos de estos defectos.

- Correcciones sobre el sensor: Como cualquier instrumento científico, los sensores utilizados para adquirir las imágenes están sujetos a limitaciones físicas, limitaciones relacionadas con su diseño y a fallos derivados

de su uso. Estos errores son conocidos habitualmente como artefactos y se definen como cualquier característica que aparece en una imagen que no está presente en el objeto original (Hsieh, 2009). Es posible que estos errores sean irrelevantes o fácilmente descartables en una clasificación manual, pero pueden cambiar completamente las características que se extraen de la imagen y por tanto los modelos generados a partir de ellas. La naturaleza de estos artefactos irá asociada al tipo de sensor que se utiliza e implicará un conjunto de técnicas específicas para su corrección. Particularizando, por ejemplo, al ámbito médico de la imagen de resonancia magnética, podemos ver algunos ejemplos de artefactos y técnicas de preprocesado asociadas en (Smith, 2010).

- Correcciones de iluminación: Una distribución irregular en la iluminación, entendida ésta como el conjunto de intensidades que puede adoptar un elemento en una escena, puede introducir áreas sombreadas en la imagen que dificultan distinguir las texturas y estructuras locales; lo que es más, una iluminación desigual dentro de una conjunto de imágenes podría introducir diferencias artificiales en su caracterización, añadiendo un sesgo a los resultados. De entre los métodos de corrección de iluminación se pueden destacar aquellos basados en filtrado de rango, ecualización de histograma y remapeado de tablas de color (Liang et al., 2010).
- Filtrado de Ruido: En su definición más amplia, podemos considerar el ruido como la información innecesaria en una imagen. Más formalmente, particularizando al dominio de las imágenes, el ruido se define como variaciones aleatorias en el brillo o la información espectral de una imagen (Motwani et al., 2004). Cómo aparece representado, cuál es su origen y hasta qué punto puede ser eliminado depende de nuevo del tipo de sensor que se ha utilizado para adquirir la imagen. La preeminencia de la presencia de ruido en las imágenes ha hecho que existan multitud de filtros (paso bajo, paso alto, medio, mediana, etc...) para tratar de eliminar o minimizar la presencia de este fenómeno en éstas (Farooque y Rohankar, 2013).
- Correcciones geométricas: Si una escena aparece desplazada, girada o se ha tomado desde una perspectiva inadecuada, puede ser útil corregir su geometría antes de realizar la extracción de características, pues algunos descriptores pueden ser sensibles a cambios de posición y orientación. En cualquier caso, si el proceso de clasificación implica múltiples imágenes, las condiciones en las que estas han sido adquiridas pueden variar de forma significativa, incluyendo una nueva variable en el problema. El preprocesamiento para la corrección geométrica es un paso común en problemas de teledetección, como puede verse en (Jensen y Lulla, 1987).

- Correcciones de color: Puede ser necesario redistribuir la saturación del color o corregir los artefactos de iluminación en el canal de intensidad, de cara a minimizar las variaciones de color en un conjunto de imágenes. Típicamente, el tono de color es uno de los atributos más difíciles de corregir, y puede ser necesario el uso de técnicas interactivas o cambios a espacios de color diferentes al RGB (Sharma y Bala, 2002).

En los procesos de mejora de imágenes, el objetivo es acentuar ciertas características de la imagen para su posterior análisis o para mejorar la visualización de la imagen. La mejora de la imagen es útil para la extracción de características, el análisis general de la imagen y la visualización. El proceso de mejora en sí mismo no aumenta el contenido de información inherente en los datos, simplemente enfatiza ciertas características especificadas de la imagen. Algunas de las técnicas de mejora son:

- Iluminación: La iluminación en una imagen no solo puede corregirse, es también posible mejorarla. La iluminación global puede mejorar-

Figura 2.1: Algunos ejemplos de problemas asociados a la captura de imágenes. En la fila superior se puede ver un ejemplo de ruido gaussiano, artefacto por metal en un CT parte de la "RIDER Lung CT collection" (Clark et al., 2013), y un pliegue en la preparación de un estudio histológico (Rastogi et al., 2013). En la fila inferior se ve un ejemplo de un artefacto por movimiento en un MRI (Bier et al., 2018), sobre-exposición en una fotografía (Yoo et al., 2015) y efecto de "bandeado" en una imagen de satélite (Song et al., 2018).

se mediante operaciones sencillas de remapeado de tablas de color o ecualizaciones de histograma. La iluminación local se puede mejorar utilizando filtros de gradiente, ecualización de histograma local y filtros de rango.

- Mejoras de enfoque y desenfoque: En la fase de preprocesamiento pueden emplearse muchos métodos de filtrado conocidos para el enfoque y desenfoque. Por ejemplo, para compensar los artefactos de alias de píxeles introducidos por la rotación que pueden manifestarse como píxeles borrosos que oscurecen los detalles finos, se pueden utilizar filtros de nitidez para mejorar las características de los bordes antes de los cálculos de gradientes.
- Realce de bordes: El realce de bordes mejora la nitidez de la imagen compensando posibles desenfoques en la imagen al destacar las transiciones entre áreas presentes en la imagen (Papari y Petkov, 2011).
- Falso color: Si bien en principio no sería relevante de cara a la extracción de características, puesto que no se añade información, es posible enriquecer la interpretación de una imagen a color re-ordenando o descartando bandas espectrales en una imagen para facilitar la visualización de características que no son fácilmente discernibles de otra manera.

En el contexto de la anotación semántica de imágenes, se puede incluir la segmentación como una operación de preprocesamiento orientada a la mejora de la imagen, si bien hablaremos de este proceso en la sección 2.4.

2.2. Extracción de características para anotación semántica

El ser humano es perfectamente capaz de extraer toda la información visual presente en una imagen; sin embargo, no es el caso de los algoritmos matemáticos. El análisis automático de la información visual contenida en una imagen pasa por la obtención de una representación adecuada para la realización de una clasificación del contenido presente en estas. Este proceso se conoce como extracción de características y se realiza mediante descriptores visuales. Citando a Kumar y Bathia en (Kumar y Bhatia, 2014): "Los descriptores visuales deben contener la información necesaria para distinguir entre clases, ser insensibles a variabilidades mínimas en la entrada, y ser limitadas en número, para permitir un cómputo eficiente de las funciones discriminantes y limitar la cantidad de datos de formación requeridos".

Un descriptor visual se define como una función o algoritmo que extrae una característica o conjuntos de características específicas de una imagen,

expresándola como un vector numérico (Stoklasa, 2016). Formalmente, y particularizando a imágenes bidimensionales, siendo una imagen una función $I(x, y)$ definida sobre el dominio de imagen Ω , podemos definir un descriptor de imagen como una función vectorial $\psi(I) = \vec{c} = (c_1, \dots, c_k)$. El vector de características resultado, $\vec{c}$, representará, por ejemplo, información estadística general, propiedades texturales o información sobre la forma y el aspecto de las diferentes regiones que componen una imagen.

Un descriptor visual puede aplicarse a una imagen completa, a regiones de la imagen resultado de una segmentación, o a pequeñas regiones que rodean a puntos considerados de interés en la imagen. Tendremos así descriptores que producirán características globales, regionales o locales. Caracterizando a nivel global, la imagen estará representada por un vector de características multidimensional que describe la información de toda la imagen. Las ventajas de las características globales se basan en su sencillez de cálculo, que generalmente se traduce en menores necesidades de "hardware" para el procesamiento de las imágenes y mayor velocidad en la aplicación del descriptor visual. Sin embargo las características globales se ven afectadas por la heterogeneidad inherente a los conjuntos de imágenes, siendo sensibles a transformaciones geométricas en la imagen y cambios de posición u oclusiones entre los elementos que la componen.

Como ya se ha mencionado, en una representación de características a nivel de regiones, tendremos para cada imagen múltiples vectores, cada uno asociado a una región diferente. El uso de regiones añade la posibilidad de utilizar características específicas de forma, al poder ser una región un objeto irregular, o particularizar características texturales al conjunto de posibles objetos detectados en la imagen, lo que implica una mejora en la capacidad de discriminación de estos descriptores, a costa de un aumento en su complejidad y tiempo de cálculo (Salahat y Qasaimeh, 2017).

Una representación de características a nivel local describe pequeñas porciones de la imagen asociada a un punto específico, el punto clave o punto de interés. El cálculo de las características locales consta de dos etapas básicas: la detección del punto de interés y la descripción de la región. La decisión sobre lo que se considera un punto de interés y su región asociada depende de muchos factores, existiendo multitud de detectores de punto de interés diferentes, como puede verse en (Li et al., 2015b). Las estructuras locales son más distintivas y estables que regiones más grandes o una imagen completa. Los descriptores de imágenes locales, por ejemplo, han mostrado en general mejor rendimiento que los descriptores globales (Jégou et al., 2010) en tareas de búsqueda de imágenes, a costa, eso sí, de un incremento en las necesidades de procesamiento necesarias para la tarea.

Entrando con más detalle en los diferentes tipos de posibles caracterizaciones de imágenes y dependiendo del tipo de información extraída, es posible dividir estos en cuatro grupos: características espectrales, características de

forma, características basadas en el entorno y características texturales.

2.2.1. Características espectrales

Las propiedades espectrales de una imagen, el color en el caso de las fotografías, son una característica visual básica en la percepción humana (Caelli, 1981). A partir de estas se han desarrollado infinidad de posibles caracterizaciones (Garcia-Lamont et al., 2018), basadas en medidas de similitud, cuantificaciones de intensidad de color, descriptores de color dominante, etc... . Se puede, por ejemplo, señalar las incluidas dentro del estándar MPEG-7 (ISO/IEC 15938 (Martínez et al., 2002) formalmente conocido como *Multimedia Content Description Interface*) como son el descriptor de espectro de color o la descripción compacta de color dominante. Teniendo en cuenta su uso casi universal, merece la pena examinar con más detenimiento las más sencillas de estas características, los momentos estadísticos.

Matemáticamente hablando un momento estadístico (Spanos, 1999) se define como una medida cuantitativa específica de la distribución de un conjunto de puntos. El concepto de momento procede del campo de la física, donde su uso principal es la caracterización de sistemas de masas puntuales. Si bien en estadística los valores de los puntos ya no son masas, los momentos utilizados siguen estando relacionados con sus homónimos físicos. Si los puntos representan masa, entonces el momento cero es la masa total, el primer momento dividido por la masa total es el centro de la masa, y el segundo momento es la inercia rotacional. Si los puntos representan la densidad de probabilidad, entonces el momento cero será la probabilidad total (es decir, uno), el primer momento es la media del conjunto y el segundo momento central es la varianza. También significativos para la caracterización de imágenes son el tercer momento central, la asimetría ("skewness"), y el cuarto momento central, la curtosis.

Cuando un conjunto de valores tiene una tendencia central suficientemente fuerte, es decir, una tendencia a agruparse alrededor de algún valor en particular, entonces puede ser útil caracterizar el conjunto por unos pocos parámetros que están relacionados con sus momentos. La más conocida es la media de los valores $x_1, \dots, x_N$, que estima el valor en torno al cual se produce la agrupación central.

La varianza, el segundo momento central de una distribución, se define como la expectativa de la desviación al cuadrado de una variable aleatoria de su media. Informalmente, mide hasta qué punto un conjunto de números están separados de su valor promedio.

El "skewness" es una medida de la asimetría de la distribución de probabilidad de una variable aleatoria sobre su media. El valor de asimetría puede ser positivo, negativo, o indefinido. La interpretación cualitativa de la asimetría es complicada y poco intuitiva. Por ejemplo, para una distri-

Figura 2.2: Skewness; diferentes tipos de distribución.

bución unimodal, el sesgo negativo indicaría que la cola del lado izquierdo de la función de densidad de probabilidad es más larga que la derecha sin distinguir entre estos dos tipos de forma. En las distribuciones multimodales y distribuciones discretas, los valores de asimetría también son complejos de interpretar, como se muestra en (Mardia, 1970). Todo esto lo hace en principio difícil de utilizar como un descriptor visual si bien en situaciones específicas, como en (Samonds et al., 2014), ha demostrado su utilidad. Una interpretación gráfica del concepto de "skewness" puede verse en la figura 2.2.

El cuarto momento central es una medida de la "pesadez" de la cola de la distribución, llamada curtosis, comparado con una distribución normal de varianza similar. Si una distribución tiene colas importantes, la curtosis será alta (distribución leptocúrtica); por el contrario, las distribuciones de cola ligera tienen una curtosis baja (distribución platicúrtica). De nuevo, la interpretación de un valor específico de curtosis puede ser compleja, lo cual hace complicado su uso como descriptor visual. Aun así existen trabajos, como (Liu et al., 2004), donde la curtosis forma parte de un vector de características para clasificación. Una interpretación gráfica del concepto de curtosis puede verse en la figura 2.3

2.2.2. Características texturales

La descripción de la textura de una imagen ha constituido uno de los grandes temas en el campo del procesamiento de imágenes. Aunque no existe una definición formal de textura, intuitivamente se puede asociar a diversas propiedades de las imágenes, tales como suavidad, rugosidad, regularidad, etc. A veces se describe la textura de una imagen como la variación entre los niveles de intensidad en el entorno inmediato de cada pixel (Haralick y Shapiro, 1992). También se puede describir como un atributo que representa la

Figura 2.3: Ejemplo de distribuciones asociadas a diferentes tipos de curtosis: A) leptocúrtica B) mesocúrtica C) platocúrtica D) uniforme.

distribución espacial de los niveles de intensidad en una región dada de una imagen digital (Jain et al., 1995). Hay que tener en cuenta que los descriptores texturales dependen directamente del contexto y son de poca utilidad a menos que se entienda la forma en que se pueden usar y las características del conjunto de datos del cual se derivan.

La forma más sencilla de representación textural quizás sea a través del histograma de una imagen. En un sentido matemático general, un histograma es una función m_i que suma el número de observaciones dentro de cada elemento de un conjunto de categorías no solapadas (conocidas como intervalos o "bins"). En el contexto de la caracterización de imágenes, un histograma se define como una categorización de la distribución de elementos en una matriz de intensidades (Pearson, 1895). El conocido gráfico de barras asociado a un histograma es simplemente una forma de representar este.

A pesar de su sencillez, el uso de histogramas como descriptor de imágenes ha mostrado buenos resultados en tareas de clasificación de imágenes para indexado y recuperación, siendo este uso propuesto por primera vez en (Swain y Ballard, 1991). Un histograma es un descriptor robusto ante traslaciones y rotaciones de elementos en una escena, aunque, dado que en un histograma se pierdan las características espaciales no es posible utilizarlo para codificar relaciones espaciales entre píxeles. Dado su interés en el desarrollo de esta Tesis, el uso de histogramas como descriptor de imágenes se verá con más profundidad en el capítulo 4.

En 1973, en (Haralick et al., 1973) se propuso un conjunto de características texturales; con el tiempo estas características, conocidas como ca-

racterísticas de Haralick, han llegado a ser unos de los descriptores visuales más conocidos y utilizados. Este trabajo presentaba 14 características texturales derivadas de la llamada matriz de co-ocurrencia, a veces también llamada matriz de dependencia espacial o matriz de co-ocurrencia de nivel gris (GLCM), donde se describe la distribución de los valores de intensidad contiguos en una imagen. Las características de Haralick han sido utilizadas por ejemplo en la clasificación de resonancias magnéticas para el seguimiento de cancer de próstata (Gnep et al., 2017), seguridad mediante biometría del iris (Subban et al., 2017), clasificación de imágenes de microscopía de fluorescencia (Boland et al., 1998) o detección de fallos en procesos industriales (Mery y Berti, 2003).

Si bien las características de Haralick son quizás las características texturales más utilizadas, se han desarrollado gran cantidad de descriptores de texturas. Presentadas en (Tamura et al., 1978), las características de Tamura implican formulas "ad hoc" para el cálculo de atributos texturales como la regularidad, el contraste o la autosemejanza. Derivados del concepto de polinomios de Zernike (von F, 1934), las características de Zernike (Teague, 1980) aprovechan las propiedades de estos para codificar propiedades texturales de una imagen. También MPEG-7 especifica un conjunto de descriptores basados en contenido para permitir a los usuarios caracterizar imágenes basándose en criterios visuales (Martínez et al., 2002).

2.2.3. Características basadas en forma

Tal y como se plantea en (Lim y Galoogahi, 2010), la forma es un parámetro esencial en la identificación y el reconocimiento de objetos reales. En (Yang et al., 2008) se clasifican, en términos generales, las técnicas de extracción de formas en dos grupos principales: métodos basados en contornos y métodos basados en regiones (es común que en la literatura relacionada, las regiones aparezcan denominadas como "blobs"). Los métodos basados en contornos extraen características utilizando únicamente las fronteras de las regiones, los métodos basados en regiones también tienen en cuenta el interior de dichas regiones, incluyendo posibles huecos en su interior y/o áreas inconexas pero también parte de dicha región. Debido a que las técnicas basadas en el contorno utilizan sólo una parte de la región, son más sensibles al ruido que las técnicas basadas en la región, ya que los pequeños cambios en la forma afectan significativamente al contorno de la forma y tienden a ser métodos complejos, por lo que en general no son muy utilizados. Aun así, destacan los resultados obtenidos utilizando ensamblaje de fragmentos de contorno caracterizados mediante el contexto local en (Nelson y Selinger, 1998) o el uso de cadenas de segmentos de contorno aproximadamente rectos (kAS) conectados (Ferrari et al., 2008) en problemas generales de reconocimiento de objetos.

Sin embargo, en muchos casos se han utilizado con éxito para caracterizar

imágenes descriptores de forma regionales muy sencillos, que incluyen, área, circularidad, excentricidad... Los descriptores individuales de forma más simples no son robustos (Gong et al., 1994); por lo que es común combinarlos para crear descriptores más efectivos. Así, en (Daramola y Ademola, 2016) y en (Ding y Gunasekaran, 1994), el tamaño de las regiones aparece como una de las principales características, en (Gudi et al., 1994) el vector de características utilizado para reconocer hojas de árboles incluye la circularidad y la excentricidad entre otras.

Aplicaciones específicas pueden requerir descriptores visuales basados en forma más complejos, utilizándose sobre imágenes en dominios concretos, como el reconocimiento de texto a partir de diccionarios de formas en (Sharma et al., 2009), el reconocimiento de posturas corporales usando el descriptor de forma de contorno de MPEG-7 en (Moghaddam y Piccardi, 2010) o la identificación de pájaros utilizando un descriptor de forma de Fourier en (Liu et al., 2007).

2.2.4. Características basadas en el entorno

La importancia de la relación espacial como elemento de caracterización de una imagen ha sido destacada en muchos trabajos, partiendo de (Freeman, 1975) e incluyendo lingüística, filosofía, psicología experimental e informática. A pesar de la aceptación de su importancia como característica definitoria, la mayoría de las discusiones sobre características relacionadas con el entorno sólo se centran en una pequeña fracción de las características reales, en general con estudios individuales.

Partiendo de los estudios presentados en (Saad et al., 2015) podemos dividir las características basadas en el entorno en 3 categorías:

- Topológicas: codifican relaciones espaciales entre objetos invariantes a transformaciones de rotación translación o escalado.
- Orientación: describen la ubicación de un objeto como una relación de orientación total o parcial con un objeto de referencia.
- Métricas: contienen relaciones entre objetos que pueden ser físicamente expresadas en términos de escala. Según (Hernández-Gracidas y Sucar, 2007), la relación espacial entre un objeto y su referencia de objeto cambiará en la escala, pero no se verá afectada por la rotación o la traslación.

2.3. Algoritmos de clasificación para anotación semántica

Partiendo del trabajo presentado en (Kotsiantis et al., 2007), podemos definir un proceso de clasificación como una forma de aprendizaje automático en la que dos o más grupos o clases de objetos distintas se conocen de antemano y sus características se utilizan para agrupar objetos adicionales. Dependiendo de si tenemos únicamente una descripción de las clases o un conjunto de muestras de cada una de ellas, estaremos hablando de una clasificación no supervisada o supervisada.

En una clasificación no supervisada, las identidades de las diferentes clases deben especificarse, ya que generalmente no son conocidas *a priori* porque no existe un conjunto de entrenamiento o porque las características dentro de la escena no están bien definidas. La imagen se dividirá en diferentes clases de acuerdo con algunos criterios predeterminados como puedan ser forma, tamaño, color o textura, que actúan como características (Sathya y Abraham, 2013). Si la clasificación no supervisada depende únicamente de características intrínsecas a la imagen, esto es, no hay una semántica previa que nos condicione el proceso de clasificación (no estamos buscando objetos verdes en una imagen porque sabemos que son árboles, aunque una vez terminada la clasificación no supervisada veamos que, efectivamente en la imagen los objetos verdes corresponden a árboles), se puede interpretar como un proceso de segmentación, concepto que será examinado en profundidad en la siguiente sección.

En una clasificación supervisada, la identidad y ubicación en la imagen de las clases de interés son conocidas al haber sido etiquetadas previamente, generalmente por un experto. El analista intenta localizar áreas específicas en las imágenes que representen ejemplos homogéneos y representativos de las clases que se desean etiquetar en las imágenes. Estas áreas se denominan comúnmente conjunto de entrenamiento, porque las características de estas áreas conocidas se utilizan para entrenar el modelo de cara a realizar la clasificación. A partir de este modelo se evalúa el resto de la imagen/imágenes y se asigna cada píxel una clase, en general la que en el modelo de como resultado la mayor probabilidad de ser miembro.

En este trabajo nos centraremos en los métodos de clasificación supervisada. Dentro de los métodos supervisados, nos encontramos con dos "familias" principales de clasificadores: generativos y discriminativos.

2.3.1. Algoritmos de clasificación generativos

Un clasificador generativo, también conocido como probabilístico, (Taskar et al., 2001) trata de inferir el modelo a partir del cual se ha generado un conjunto de datos, esto es, para un conjunto de elementos $x_1, x_2, \dots, x_n$ y un

conjunto de clases $c_1, c_2, \dots, c_n$ se busca encontrar la probabilidad *a posteriori* $p(x|c)$ de que un elemento pertenezca a una clase determinada, usando modelos bayesianos. Si bien las probabilidades previas del modelo $p(c_i)$ se pueden modelar encontrando la frecuencia de muestras pertenecientes al concepto c_i , la parte crítica en una clasificación probabilista es modelar las probabilidades condicionales $p(x|c_i)$, siendo esta la principal diferencia entre distintos métodos probabilísticos.

2.3.1.1. Métodos no paramétricos

Los métodos no paramétricos calculan las probabilidades condicionales sin ninguna suposición previa sobre la distribución de las características de la imagen. La distribución real de características se aprende de las características de las muestras de entrenamiento. Existen dos tipos de aproximaciones a los enfoques no paramétricos:

- Los métodos basados en modelos de relevancia definen las distribuciones conjuntas elemento/clase y calculan las probabilidades *a posteriori* de los elementos de prueba para cada clase, asignando al elemento la clase más probable. Los métodos basados en modelos formulan un problema de optimización para aprender la probabilidad de un elemento c dado un vector de características. Destacar dentro de los métodos basados en relevancia los trabajos de anotación de imágenes y vídeo utilizando modelos de relevancia espacial continua (Continuous-space Relevance Model) y modelos de relevancia múltiple de Bernoulli (Multiple Bernoulli Relevance Models, MBRM) en (Lavrenko et al., 2004)) y (Feng et al., 2004)) respectivamente.
- Los clasificadores basados en modelos temáticos consideran el conjunto de entrenamiento como muestras de una mezcla específica de temas, donde cada tema es una distribución de probabilidad sobre las características de la imagen y las clases presentes en el problema. Algunos ejemplos típicos de un clasificador basado en modelos temáticos son el análisis semántico latente (Latent Semantic Analysis, LSA) (Dumais, 2004) o la asignación latente de Dirichlet (Latent Dirichlet allocation, LDA) (Blei et al., 2003), este último usado con éxito en la clasificación general de imágenes (Zhang et al., 2011a).

2.3.1.2. Métodos paramétricos

En un enfoque paramétrico de clasificación se supone que el espacio de característica sigue un cierto tipo de distribución continua y conocida. Por lo tanto, la probabilidad condicional $p(x|c)$ se modela utilizando esta distribución de características. Las características se agrupan y cuantifican primero;

luego se construye el modelo de probabilidad condicional para cada conglomerado. La probabilidad condicional $p(x|c)$ se suele modelar como una distribución gaussiana multivariada, basada en la matriz de covarianza calculadas a partir de los vectores de la característica.

El uso de enfoques paramétricos para clasificación de imágenes ha sido limitado, pues existen varios problemas asociados a estos. En primer lugar, estos enfoques suelen asumir una distribución de características particular que puede no ser cierta. En segundo lugar, los parámetros del modelo se aprenden normalmente mediante métodos de optimización, que en la mayoría de los casos pueden no converger, siendo el cálculo del proceso de optimización muy costoso. En tercer lugar, para los modelos de anotación concepto por concepto, el modelo para cada concepto se aprende independientemente y no se captan las posibles correlaciones entre conceptos. Aun así, en (Blei y Jordan, 2003) se plantea con éxito el uso de esta clase de modelos para la clasificación de imágenes.

2.3.2. Algoritmos de clasificación discriminativos

Los métodos de clasificación discriminativos, también conocidos como métodos no probabilísticos o condicionales, consideran cada etiqueta por separado y en base a cada una de ellas entrenan un modelo específico. Como se muestra en (Taskar et al., 2002), aunque son muy eficaces, una desventaja importante es que requieren definir de antemano el vocabulario de las etiquetas. Por lo tanto, estos enfoques no manejan bien los escenarios a gran escala, en los que puede tener miles de etiquetas y el vocabulario puede cambiar con el tiempo.

2.3.2.1. Clasificación lineal

El principio subyacente a los clasificadores lineales pasa por tomar una decisión de clasificación basada en el valor de una combinación lineal de características. Estos clasificadores funcionan bien para problemas prácticos como la clasificación de documentos y, en general, para problemas con muchas variables (características), alcanzando niveles de precisión comparables a los de los clasificadores no lineales más complejos, a la vez que se tarda menos tiempo en entrenarlos y utilizarlos (Yuan et al., 2012). Para un problema de clasificación de dos clases, se puede visualizar el funcionamiento de un clasificador lineal dividiendo un espacio de entrada de alta dimensión con un hiperplano: todos los puntos de un lado del hiperplano se clasifican como la clase "a", mientras que los pertenecientes al otro lado clasifican parte de la clase "b".

Los modelos de clasificación lineales están representados en algunos de los primeros trabajos de anotación semántica para imágenes desarrollados, en (Smith et al., 1985) o (Sabot) donde se utilizaban a menudo modelos

mixtos para definir una distribución conjunta sobre las características de la imagen y las etiquetas.

2.3.2.2. Clasificación mediante máquina de vectores de soporte

Encuadradas originalmente dentro de los clasificadores lineales, una máquina de vectores de soporte (support vectors machine, SVM, (Suykens y Vandewalle, 1999)) es un método general de aprendizaje supervisado capaz de realizar una separación binaria entre clases. En un SVM la clasificación se realiza buscando el plano o, dependiendo de la dimensionalidad del problema, el hiperplano que mejor diferencie las dos clases. La idea es obtener lo que se denomina un margen máximo a cada lado del hiperplano, seleccionando uno de separación equidistante a las muestras más cercanas entre sí de cada clase. Sólo se consideran los datos que definen los bordes (los vectores de soporte) de esos márgenes. La búsqueda del hiperplano de separación en estos espacios, normalmente de muy alta dimensión, se hará implícitamente utilizando las llamadas funciones "kernel". La posibilidad de elegir y diseñar estos "kernels" dotan al método de una gran flexibilidad. Desde un punto de vista algorítmico, la optimización del margen geométrico representa un problema de optimización cuadrática con restricciones lineales, que pueden resolverse mediante técnicas de programación cuadrática estándar.

En general, el uso de SVM produce buenos resultados en comparación con otros clasificadores, especialmente para grandes conjuntos de datos, y es notablemente robusto frente a variaciones en el sesgo o varianza del modelo, así como a conjuntos de datos de entrenamiento pequeños (Meyer et al., 2003). El SVM puede clasificar tanto datos lineales como no lineales debido al uso de mapeo de kernel. La ventaja de SVM sobre otros clasificadores es que logra límites de clase óptimos al encontrar la distancia máxima entre clases.

Un SVM es un clasificador binario. Sin embargo, es muy habitual que la semántica que queremos asociar a una imagen implique la existencia de más de dos conceptos semánticos. El enfoque más común a la hora de tratar un SVM como un clasificador multiclase implica entrenar un SVM separado para cada concepto con cada SVM generando un valor de probabilidad (Hsu y Lin, 2002). Durante la fase de prueba, las decisiones de todos los clasificadores se fusionan para obtener la etiqueta de clase final de una imagen de prueba. El nivel base consiste en múltiples clasificadores binarios y el segundo nivel fusiona las decisiones de los clasificadores del nivel base.

2.3.2.3. Clasificación mediante árboles de decisión

Un árbol de decisión (Breiman, 2017) es una herramienta de clasificación en múltiples etapas. Dependiendo del número de decisiones tomadas en cada nodo interno del árbol, un árbol de decisión puede llamarse árbol bina-

rio o n-ario. A diferencia de otros modelos de clasificación cuyas relaciones entrada-salida son difíciles de describir, la relación entrada-salida en un árbol de decisión puede expresarse utilizando reglas comprensibles para el ser humano, por ejemplo, reglas "if/then". Este método, a veces conocido como inducción de regla, presenta varias ventajas. Primero, el proceso de razonamiento detrás del modelo resulta claramente evidente cuando se examina el árbol y resulta útil cuando se desea ver el modo en que los atributos de los datos pueden dividir o particionar la población en subconjuntos relevantes para el problema. Esto contrasta con otras técnicas de modelado de caja negra en las que la lógica interna puede resultar difícil de averiguar. Para etiquetar una nueva muestra, se atraviesa el árbol desde el nodo raíz hasta un nodo de hoja utilizando el valor del atributo de la nueva muestra. La decisión de la muestra es el resultado de la hoja final a la se llega (Rokach y Maimon, 2008).

En comparación con otros métodos de aprendizaje, un árbol de decisión es fácil de interpretar y entender y puede aprender con un pequeño número de muestras. También es robusto ante datos incompletos y ruidosos (Guo et al., 2004). Estas ventajas han dado lugar a una multitud de algoritmos para la creación de árboles de decisión.

El primer algoritmo basado en árboles de decisión publicado fue THAID (Doyle, 1973). En este trabajo se plantea la que será la estrategia general más común en la generación de árboles de decisión; esto es una búsqueda exhaustiva en el espacio definido por los vectores de características, donde cada nodo se crea minimizando la probabilidad de que en ese nodo un elemento sea etiquetado incorrectamente (lo que se conoce generalmente como un índice de "impureza"). THAID específicamente emplea una medida de impureza del nodo basada en la distribución más probable del vector de características para cada nodo.

Esta estrategia ha demostrado ser muy popular, dando lugar a una serie de algoritmos muy utilizados. Uno de los árboles de decisión más simples, ID3, propuesto en (Quinlan, 1986), está basado en el concepto de entropía, entendido como ganancia de información. C4.5 (Quinlan, 2014) y CART (Breheny, 1984) son dos algoritmos de árbol de clasificación posteriores que siguen este enfoque. C4.5, una extensión del algoritmo ID3, utiliza el concepto de entropía como función de impureza, mientras que CART (Classification And Regression Tree), algoritmo que en su momento fue tan popular que se ha convertido en un término genérico para designar las diferentes técnicas asociadas a los árboles de decisión), utiliza una generalización de la varianza binomial llamada índice de Gini. Estos algoritmos, a diferencia del THAID, primero crean un árbol demasiado grande y luego lo podan a un tamaño más pequeño para minimizar un estimado del error de clasificación.

2.3.2.4. Clasificación basada en redes de neuronas

La habilidad de las redes neuronales para aprender patrones también se ha utilizado para la anotación semántica de imágenes como puede verse, por ejemplo en (Sonka et al., 2014). La red neuronal tiene la ventaja de que las salidas de las neuronas de la capa de salida están determinadas por las capas anteriores y los bordes de conexión. No necesita ningún otro ajuste de parámetros ni ninguna suposición sobre la distribución de la característica. Sin embargo, las redes neuronales presentan varios problemas. Así la exactitud de la clasificación depende del número de capas y neuronas ocultas, número que de hecho resulta muy difícil de justificar, al igual que la elección de las funciones de activación apropiadas tal y como se muestra en (Del Frate et al., 2007). Una red neuronal funciona como una caja negra, lo que significa que la relación exacta entre la entrada y la salida no es transparente pudiendo ser difícil de explicar. Finalmente, suelen ser costosos en tiempo de cómputo.

Hoy en día son especialmente populares las redes neuronales convolucionales (convolutional neural network, CNN) derivadas de los trabajos presentados en (Fukushima y Miyake, 1982). Las CNNs, aplicadas a imágenes, representan básicamente varias capas de filtros de convolución donde, en lugar de conectar cada neurona de entrada a cada neurona de salida como en una red neuronal tradicional, cada región de la entrada está conectada a una neurona en la salida. El propósito principal de los filtros de convolución es la extracción de características de la imagen de entrada. Estos filtros preservan la relación espacial entre los píxeles al operar en pequeños cuadrados sobre los datos de entrada, no a nivel de píxel. Cada capa aplica diferentes filtros de convolución, típicamente cientos o miles de veces, dependiendo del tamaño de la imagen y combina sus resultados. Durante la fase de entrenamiento, una CNN aprende automáticamente los valores de sus filtros en función de la tarea para la que está siendo entrenado. Al clasificar una nueva muestra, cada neurona de salida genera una medida de confianza y la clase correspondiente a la medida máxima indica la decisión sobre la muestra.

2.3.3. Clasificación perezosa

Finalmente, y fuera de las categorías típicas en las que se encuadran los algoritmos de clasificación, se encuentra la conocida como clasificación perezosa. Los clasificadores "perezosos" difieren de los enfoques de aprendizaje inductivos descritos anteriormente al no generar directamente un modelo de clasificación. Si bien dependen también de un conjunto de entrenamiento, el proceso de entrenamiento como tal no existe. En el momento de etiquetar una nueva instancia, un clasificador perezoso realizará una búsqueda entre todos los ejemplos del conjunto de entrenamiento utilizando un criterio de similitud determinado para asignarle a esta su clase correspondiente (Aha, 1997). Algunos ejemplos típicos de clasificadores perezosos son el aprendizaje

basado en instancias (Aha et al., 1991) y los sistemas bayesianos perezosos (Zheng et al., 1999).

Dentro del conjunto de métodos de clasificación perezosa encontramos los clasificadores basados en distancia. Entendiendo un vector de características normalizado como una secuencia de valores, es posible interpretarlo como un conjunto de coordenadas que definen la posición de cada sujeto en un espacio n -dimensional, donde n corresponde al número de características extraídas de dichos sujetos. Esto permite el uso de algoritmos de agrupamiento por proximidad espacial como métodos de clasificación. Quizás el mejor ejemplo de clasificador mediante proximidad espacial sea el método del vecino más próximo (k-Nearest Neighbor, k-NN). K-NN es un clasificador no paramétrico desarrollado en (Fix y Hodges, 1951), donde para clasificar una imagen desconocida, representada por un vector de característica, se calcula las distancias entre el punto y los puntos en el conjunto de datos de entrenamiento, siendo asignado el punto a la clase más común entre sus vecinos más cercanos.

2.4. Segmentación de imágenes

La segmentación es el proceso por el cual se realiza una división de una imagen en varios elementos coherentes, sin conocimientos previos sobre la imagen y sin buscar que los elementos resultado tengan un significado (Shapiro y Stockman, 2001). Formalmente, partiendo de (Gonzalez et al., 2002) se puede definir una segmentación como el proceso de particionado de una imagen en regiones conectadas no solapadas, homogéneas en relación a alguna característica o características específicas, como puedan ser intensidad, color o textura. El objetivo de una segmentación es encontrar las regiones S_i más grandes posibles para las que se cumple un criterio de homogeneidad. Sea $P(-)$ una función que evalúa una condición de homogeneidad. Se espera que una segmentación cumpla para todas las regiones $P(S_k) = \text{verdadero}$ y $P(S_i \cup S_j) = \text{falso}$ para dos regiones adyacentes S_i y S_j cualquiera.

El principal problema dentro de un proceso de segmentación es definir un criterio de homogeneidad P para que la segmentación producida sea la más adecuada para una aplicación particular. ¿Cómo definir el criterio P para que los segmentos se correspondan únicamente con los objetos y regiones del mundo real representados en la imagen? ¿Qué son exactamente los objetos y áreas que nos interesan? Afortunadamente, se pueden encontrar algunas vías para desarrollar algoritmos de segmentación a partir de la investigación en percepción humana. Así, los elementos similares en color, forma o posición espacial tienden a agruparse y es en general en esta clase de características de bajo nivel, color, textura, fronteras en las que se suelen basar los algoritmos de segmentación

La ambigüedad inherente a un proceso de segmentación ha hecho que

Figura 2.4: Segmentación de imágenes; esquema de organización de los diferentes métodos.

este sea uno de los problemas más estudiados en el campo del tratamiento de imágenes. Existen cientos de técnicas de segmentación diferentes porque no existe un método único que sea bueno para todas las imágenes. Diferentes métodos son adecuados para diferentes problemas de segmentación y diferentes tipos de imágenes. Atendiendo a la "dirección" en la que se realiza la segmentación, es posible agrupar los diferentes algoritmos en dos enfoques diferentes: métodos libres (métodos ascendentes) y métodos basados en conocimiento (métodos descendentes). Se puede ver un esquema de los diferentes métodos de segmentación en relación con estos enfoques en Fig. 2.4.

2.4.1. Métodos ascendentes

Los métodos ascendentes giran alrededor del concepto de agrupación, con el objetivo de unir píxeles con propiedades visuales consistentes de acuerdo con un cierto criterio de similitud. Debido a la falta de suposiciones sobre la forma geométrica del objeto de interés, estos métodos no tienen restricciones en relación a que puede considerarse una solución admisible, que es una propiedad deseada en algunos casos, pero también inconveniente en otros, ya no existe una base sobre la que rechazar resultados erróneos debido a la variabilidad de la intensidad, las oclusiones, la presencia de ruido, etc.

El método de segmentación más sencillo es la umbralización. Convierte la imagen de entrada en una máscara binaria de manera que todos los píxeles con una intensidad inferior a un umbral especificado se etiqueten como "fondo" y todos los demás como "objeto". La búsqueda del umbral óptimo para cada imagen ha dado lugar a multitud de algoritmos (Yeung, 2018), admitiendo muchos de ellos la posibilidad de calcular múltiples umbrales, extendiendo el problema a un caso de multiumbralización.

Los métodos basados en crecimiento de regiones representan otra aproximación a la segmentación de tipo ascendente (Sonka et al., 2014). En general se parte de un conjunto de pequeñas regiones iniciales (algunos métodos comienzan con regiones que consisten inicialmente en un simple píxel) y un criterio de homogeneidad $P(-)$ para las regiones vecinas. Si el criterio indica que la homogeneidad no se rompe incluso después de una fusión entre regiones vecinas, estas regiones se fusionan. Es posible refinar este algoritmo añadiendo un proceso opuesto de separación. Cuando una región no cumple el criterio de homogeneidad, se divide en regiones más pequeñas. Estas regiones más pequeñas pueden volver a fusionarse con regiones vecinas más adelante.

La segmentación utilizando cortes en grafos se plantea como un problema de optimización del etiquetado en un grafo. El enfoque utiliza la representación de grafo de la imagen del problema y un método de minimización de energía/errores para encontrar la solución (Boykov et al., 1998). El algoritmo precisa de un conjunto de marcadores que definan los diferentes objetos a segmentar, que pueden ser creados manual o automáticamente (como resultado de un paso de preprocesamiento anterior). Formalmente, la función de energía junto con los marcadores pueden ser entendidos como un modelo. El método de corte gráfico trata de ajustar el modelo a la representación en grafo de los datos de la imagen. Los métodos de corte gráfico son muy populares como herramienta de segmentación manual, ya que el usuario es capaz de corregir errores de segmentación muy fácilmente, necesitando únicamente añadir uno o más marcadores nuevos en las regiones problemáticas.

Los métodos de segmentación basados en agrupamientos, "clustering", son una extensión de las técnicas de análisis en minería de datos a imágenes. Generalmente, codifican los píxeles en un espacio vectorial de características, y después realizan agrupamientos en el espacio definido por dichos vectores. Estos métodos tienden a dividir la imagen en regiones pequeñas porque la mayoría de las características de bajo nivel son en realidad estadísticas locales (Sathya y Manavalan, 2011). Algoritmos de segmentación basados en agrupamiento muy conocidos son la transformación de divisoria (Beucher y Meyer, 1992) o los ya vistos K-means y mezcla de gaussianas.

Finalmente destacar que, si bien la detección de contornos no es realmente un método de segmentación (en los términos formulados al principio de esta sección), si ha dado lugar a diferentes procesos utilizados para seg-

mentar imágenes. La detección de contornos está íntimamente relacionada con el concepto de segmentación. Si tenemos una segmentación, siempre podemos obtener contornos cerrados a partir de los límites de los segmentos. Desafortunadamente, no tiene por qué ser obligatorio que existan contornos cerrados, por lo que el proceso opuesto para obtener regiones de los contornos es más complicado, como se muestra en (Jacobs, 1996). La detección de contornos es una tarea común dentro de la visión por ordenador, similar a la detección de bordes. La diferencia entre contornos y bordes es que los bordes corresponden a la variación de los valores de intensidad, mientras que los contornos deben corresponder a objetos salientes. Existen multitud de métodos para detección de contornos por cómo puede verse por ejemplo en (Joshi y Sivaswamy, 2006) ó (Mairal et al., 2008), que han dado lugar a una variedad de métodos para la segmentación de imágenes (Arbelaez et al., 2011).

2.4.2. Métodos descendentes

Los enfoques descendentes, también denominados técnicas basadas en modelos o métodos basados en conocimiento, parten de una plantilla de segmentación genérica para la imagen. La plantilla, o modelo de segmentación, planteada se ajusta dentro de unos márgenes para adaptarse a la imagen, o al menos para minimizar el error entre modelo e imagen tanto como sea posible (Sonka et al., 2014). La diferencia entre el modelo, el error, y la imagen se utiliza para confirmar o rechazar la segmentación y, a veces, también para implementar modificaciones del modelo que conduzcan a un mejor ajuste. Los métodos basados en el conocimiento, asumen que el espacio de soluciones admisibles es limitado, y buscan una solución que sea un compromiso entre la producida a partir de las observaciones y la expresada en el espacio modelo.

Presentado en (Brunelli, 2009), la correspondencia de plantillas es un representante típico de los enfoques de los métodos descendentes donde los objetos se representan con modelos, a menudo denominados plantillas, y se busca la mejor coincidencia posible en la imagen. Este planteamiento tan sencillo sin embargo, esconde un conjunto de problemas que deben de ser abordados y resueltos, principalmente como abordar transformaciones como la escala y la rotación. El enfoque más básico consiste en probar la plantilla en todas las transformaciones posibles (posiciones, rotaciones, escalas), pero este enfoque es muy intensivo desde el punto de vista computacional. Otro problema surge en el momento de evaluar la similitud entre la plantilla y las partes de la imagen. Existe todo un conjunto de medidas típicas de similitud que puede aplicarse no solo a la plantilla completa, sino de cara a dotarlo al método de flexibilidad ante distorsiones y transformaciones geométricas, la plantilla se puede dividir en partes pequeñas (parches) que se colocan relativamente al punto de referencia de toda la plantilla, como puede verse

en (Boiman y Irani, 2007). Esta aproximación es muy utilizada en el ámbito de la medicina, donde recibe el nombre de segmentación basada en atlas. El atlas es una imagen de referencia asociada a una segmentación de un conjunto de imágenes que representa la verdad terreno. En este contexto, es muy importante elegir el atlas que se utilizará para la segmentación. Por lo tanto, se han desarrollado enfoques de segmentación multi-atlas (por ejemplo, en (Aljabar et al., 2009)) con técnicas de construcción de atlas muy sofisticadas para capturar las variaciones de forma del objeto de interés.

Los modelos deformables se basan en curvas o superficies que se mueven bajo la influencia de un conjunto de fuerzas definidas, tanto internas como externas. Las fuerzas externas deforman el modelo hacia la forma y posición deseadas, mientras que las fuerzas internas mantienen el modelo coherente (por ejemplo, determinan su suavidad). Existen dos tipos básicos de modelos deformables: modelos deformables paramétricos y modelos deformables geométricos. Los modelos deformables paramétricos se conocen como serpientes, modelos de contorno activos o curvas de minimización de energía. Las serpientes fueron propuestas en (Kass et al., 1988). Las serpientes fueron diseñadas como modelos interactivos por lo tanto, deben ser inicializados cerca de las regiones deseadas para garantizar un buen rendimiento en aplicaciones no interactivas. En (Cohen, 1991) se introduce el concepto de fuerza de inflado interno, llamada fuerza de globo, que propaga las serpientes sobre bordes espurios hacia las regiones deseadas y las hace menos sensibles a las condiciones iniciales. Los llamados modelos geométricos deformables (Caselles et al., 1993), implican la representación implícita del contorno como el punto cero de una función a resolver. De esta forma se obtienen algunas ventajas sobre los modelos deformables paramétricos, por ejemplo, no sufre de problemas de topología (los contornos pueden dividirse y fusionarse naturalmente).

2.4.3. Métricas para la evaluación de la segmentación de imágenes

La evaluación de la calidad de la segmentación de una imagen es un problema complejo, especialmente cuando no hay un conjunto de evaluación con el que comparar. Qué puede considerarse una buena segmentación está sujeto a criterios subjetivos y diferentes aplicaciones pueden funcionar mejor usando diferentes segmentaciones. Si bien los criterios de una buena segmentación son a menudo dependientes de la aplicación y difíciles de definir explícitamente, para muchas aplicaciones, incluso en una evaluación visual, la diferencia entre una segmentación adecuada y una inferior es notable.

Es posible proponer un conjunto de características deseables y medibles en una segmentación (como puede verse en el trabajo seminal de Haralick et al. en (Haralick et al., 1973)) que pueden ser calculadas y utilizadas como criterio para una evaluación cuantitativa no supervisada. Este tipo de eva-

luación presenta importantes ventajas; es completamente objetiva y, quizás la más crítica, no requiere referencias anotadas con las que comparar. Pero estos criterios cuantitativos no supervisados no pueden proporcionar una medida absoluta de la calidad de una segmentación, ya que sólo un conjunto específico de características en una imagen son realmente medibles.

En (Zhang y Gerbrands, 1994) se propone una clasificación de los métodos de evaluación como "analíticos", "bondad empírica" o "discrepancia empírica", correspondiendo la discrepancia empírica al uso de una verdad terreno. Los métodos analíticos intentan caracterizar un algoritmo en sí mismo en términos de principios, requisitos, complejidad, etc. sin referencia a una implementación concreta del algoritmo o un resultado con datos prueba. Por ejemplo, es posible evaluar las características generales de la información de la imagen que se utiliza para realiza la segmentación (Liedtke et al., 1987), (Zhang y Gerbrands, 1991). También se puede usar como métrica de evaluación medidas de complejidad y eficiencia del algoritmo (Garbay, 1986), o su respuesta a modelos de datos teóricos o a perturbaciones en la imagen, como ruido añadido (Gerbrands, 1988). Mientras que en dominios limitados, como la detección de fronteras esto puede ser útil, en general la falta de una contexto general para la segmentación de imágenes limita los resultados proporcionados por estos métodos.

Los métodos basados en una bondad empírica evalúan los algoritmos calculando a partir de propiedades específicas sobre la imagen segmentada, sin conocer *a priori* el resultado deseado en esa segmentación. La ventaja de esta clase de métodos es que sólo requieren que el usuario defina una métrica de bondad, sin necesidad de contar con una verdad terreno, y pueden ser usadas incluso como una manera de ajustar la efectividad de un algoritmo de segmentación durante su ejecución. Por otro lado, las métricas de bondad representan en el mejor de los casos una aproximación heurística al problema, y, dependiendo de las suposiciones que implique dicha heurística, pueden mostrar un fuerte sesgo hacia algunas propiedades de la imagen en particular. Dentro del conjunto de métodos basados en la bondad empírica de una segmentación (Zhang et al., 2008a) podemos destacar el uso de la uniformidad de niveles de grises en una región (Levine y Nazif, 1985b), medidas de entropía inter e intra-regional (Zhang et al., 2003) o una medición de las diferencias de color en las fronteras de las regiones (Erdem et al., 2004).

Capítulo 3

Extracción de características basadas en sobre-segmentación de imágenes

El funcionamiento de la capacidad de agrupación perceptiva del ser humano es un campo de investigación cuyo origen se encuentra en el siglo pasado, en los estudios sobre cognición perceptiva de (Humphrey, 1924), y es conocido como psicología de la Gestalt. Uno de los principios fundamentales postulado por esta es la llamada ley de la "Prägnanz" (traducido generalmente del alemán como "pregnancia" o, mejor aún, "principio de la buena forma"), planteado en (Pratt, 1936) que afirma la tendencia de la experiencia perceptiva a organizarse de la forma más simple posible. Que las partes de una figura tengan "buena forma" implica que existe una propiedad o caracterización común a estas con las que es posible definir claramente unidades autónomas dentro del conjunto de partes que componen una imagen. Este concepto hace que en una imagen, aunque haya figuras que se superponen e interfieran entre si sea posible verlas como partes diferentes de un conjunto, al prevalecer sus propiedades de "buena forma".

El principio de la "buena forma" se articula alrededor de un conjunto de principio o leyes que, según la psicología de la Gestalt, sistematizan como los seres humanos perciben naturalmente los objetos como patrones y objetos organizados. Dentro de estos, destacamos cuatro: el principio de semejanza, según el cual la mente agrupa los elementos similares en una misma entidad; el principio de proximidad, que predice que objetos próximos tenderán a verse como una unidad; el principio de simplicidad, que implica que el individuo tenderá a agrupar sus campos perceptivos de la forma más sencilla posible y el principio de continuidad, por el cual los elementos que mantienen un patrón o dirección se percibirán como parte de un mismo grupo.

Si bien una anotación semántica perceptualmente correcta de una imagen debería de cumplir con todos o la mayoría los principios de la Gestalt, las

aproximaciones al proceso de segmentación mediante la aplicación de estos no han sido especialmente efectiva, puesto que muchas de las heurísticas necesarias para forzar su cumplimiento resultaron inadecuadas para muchos escenarios del mundo real como puede verse en (Richtsfeld et al., 2012).

En este contexto, parece adecuado partir de un conjunto de primitivas más simples que un objeto, pero y a diferencia de un píxel, con significado perceptual propio. La idea detrás del concepto de superpíxel surge del hecho de que la división de una imagen en píxeles no es realmente una división natural, sino simplemente un artefacto del dispositivo que captura las imágenes. Sería más natural, y presumiblemente más eficiente, trabajar con entidades perceptiblemente significativas obtenidas a partir de un proceso de agrupación de bajo nivel.

3.1. Superpíxeles

Un superpíxel, definido en (Ren y Malik, 2003b), se describe como una región perceptualmente uniforme en la imagen. El uso de superpíxeles no proporciona una segmentación de los objetos reales presentes en una imagen, sino una sobre-segmentación de la imagen, compuesta por áreas pequeñas y homogéneas en color y/o textura. Como se plantea en (Ren y Malik, 2003a) los superpíxeles representan un sub-objeto, una primitiva conveniente desde la cual calcular las características locales de una imagen, capturando la redundancia en la imagen y reduciendo en gran medida la complejidad de las tareas posteriores de procesamiento de imágenes.

Dado que los superpíxeles son agrupaciones perceptuales de píxeles, la mayoría de los algoritmos de segmentación no supervisados pueden utilizarse para la generación de superpíxeles. De hecho, la mayoría de los autores no hacen una distinción explícita entre algoritmos de superpíxel y algoritmos generales para sobre-segmentación; es común, como puede verse (Schick et al., 2012) o (Neubert y Protzel, 2012), el uso de estos los términos de forma intercambiable. Existen algoritmos de sobre-segmentación no diseñados explícitamente para generar superpíxeles y que, sin embargo, comparten muchas características con los algoritmos para generación de superpíxeles. Así, en (Veksler et al., 2010) se distinguen los algoritmos de superpíxeles de los algoritmos de segmentación que se ejecutan en "modo de sobre-segmentación". Más formalmente, es posible usar la diferenciación propuesta en (Neubert y Protzel, 2013); los algoritmos para generación de superpíxeles deben ofrecer alguna forma de control sobre el número de superpíxeles generados y sus propiedades, mientras que un algoritmo sobre-segmentación es esencialmente un algoritmo para segmentación general configurado en modo sobre-segmentación.

Puesto que los superpíxeles son creados con el propósito de servir como elementos básicos a operaciones posteriores (a este respecto, una segmenta-

ción mediante superpíxeles puede considerarse un paso de pre-procesamiento), cabe esperar que los algoritmos para la generación de superpíxeles produzcan un conjunto de primitivas con una serie de propiedades deseables de cara a su uso en una anotación semántica (Lv, 2013):

- Computacionalmente eficiente: El uso de superpíxeles reduce la complejidad de las imágenes de cientos de miles de píxeles a sólo unos pocos miles de superpíxeles.
- Eficiente desde el punto de vista de la representación: El uso de superpíxeles permite modelar interacciones por pares de largo alcance entre elementos, mientras que en el caso de una cuadrícula de píxeles estas están restringidas a píxeles adyacentes.
- Perceptualmente significativos: cada superpíxel es una unidad perceptualmente consistente, es decir, todos los píxeles de un superpíxel presentarán una asociación por uniformidad, en general en color o textura.
- Casi completo: Como resultado de una sobre-segmentación, los superpíxeles conservan la mayoría de las estructuras presentes en la imagen. La pérdida de detalles y fronteras en la imagen es mínima en el paso de una cuadrícula de píxeles a un mapa de superpíxeles.

Como se muestra en (Michaelsen y Arens, 2017), estas propiedades codifican en un superpíxel los principios de la Gestalt de proximidad y similitud a nivel de píxeles. De esta manera, para una segmentación no supervisada, el uso de superpíxeles proporciona una forma para extraer características complejas y discriminatorias; pudiendo además reducir el tiempo de formación e inferencia.

3.1.1. Algoritmos para la generación de superpíxeles

Existen multitud de algoritmos para la generación de superpíxeles que implementan de una manera u otra las características deseables planteadas en el punto anterior. Estas características se traducen en una serie de propiedades comunes a las sobre-segmentaciones resultado que estos algoritmos producen, como puede verse por ejemplo en (Levinshtein et al., 2009) ó (Schick et al., 2012):

- Partición: Los superpíxeles deben definir una partición de la imagen, es decir, los superpíxeles deben estar disociados y asignar una etiqueta a cada píxel.
- Adherencia a los límites: Los superpíxeles deben preservar las fronteras presentes en la imagen. Como están definidas estas fronteras dependerá del tipo de imagen y del escenario de aplicación de la sobresegmentación.

- **Conectividad:** Se espera que los superpíxeles representen conjuntos de píxeles conectados o como mínimo cercanos.
- **Regularidad y suavidad:** En ausencia de fronteras dentro de la imagen, los superpíxeles deben ser compactos, estar repartidos regularmente y mostrar límites suaves.
- **Eficiencia:** Los superpíxeles deben generarse eficientemente. En general se espera que el uso de superpíxeles no afecte, o incluso acelere, el desempeño de las tareas para las que son usados.
- **Número controlable de superpíxeles:** Debe de existir un mecanismo de control que permita influir sobre el número de superpíxeles a utilizar.

Algunos de estos requisitos pueden formularse explícitamente, por ejemplo, (Liu et al., 2011) exige que los superpíxeles no disminuyan el rendimiento alcanzable de los pasos de procesamiento subsiguientes, y que las propiedades antes descritas se cumplan con el menor número de superpíxeles. En (Achant et al., 2012b) se propone incluso como propiedad de los superpíxeles que estos aumenten significativamente el rendimiento del resto de los pasos de procesamiento subsiguientes.

A partir de la clasificación presentada en (Stutz et al., 2018) podemos señalar los siguientes algoritmos para la generación de superpíxeles:

- **Superpíxeles mediante cortes normalizados** (Ren y Malik, 2003a). Derivado del algoritmo original de cortes normalizados, (Boykov et al., 1998), la generación de superpíxeles mediante cortes normalizados adopta una función de minimización de energía mucho más sensible que utilizada en el método original para obtener una sobre-segmentación de la imagen.
- **Felzenswalb & Huttenlocher** (Felzenszwalb y Huttenlocher, 2004). Propuesto en 2004, este es un enfoque también basado en grafos, originalmente orientado hacia la segmentación de imágenes. El algoritmo se basa en una medida que describe si existe evidencia de un límite entre segmentos. Usando una segmentación inicial donde cada píxel es su propio segmento, el algoritmo itera fusionando segmentos basados en dicha medida.
- **QuickShift** (Vedaldi y Soatto, 2008). Originalmente concebido como método general de segmentación, QuickShift se plantea como un algoritmo de búsqueda de la moda que, en lugar de desplazar cada punto de forma iterativa hacia una media local, forma un árbol de enlaces entre los vecinos cercanos donde se maximiza una función de densidad derivada de la presencia de grupos de píxeles similares. Puede ser categorizado como un método de ascenso de gradiente.

- Turbopíxeles (Levinshtein et al., 2009). Uno de los primeros algoritmos explícitamente diseñado para la generación de superpíxeles, Turbopíxels se basa en el uso de contornos activos. A partir de un conjunto de centros o núcleos iniciales, cada superpíxel crece por medio de la evolución de estos contornos, dirigido por una medida de afinidad entre píxeles.
- Superpíxeles Compactos, superpíxeles de Intensidad Constante (Veksler et al., 2010). Métodos relacionados entre sí, que plantean un refinamiento iterativo de un sobre-particionado semilla. Inicialmente, la imagen está cubierta por áreas cuadradas superpuestas de tal manera que cada píxel aparece en el interior de varios de estos cuadros. Cada cuadro representa un superpíxel y cada píxel puede asignarse a uno de los cuadros superpuestos. La asignación para cada píxel se calcula utilizando el algoritmo de α -expansión propuesto en (Boykov et al., 2001).
- SEEDS, (Supixels extracted via energy-driven sampling, (Van den Bergh et al., 2012)). También clasificable como método de ascenso de gradiente, SEEDS implementa un proceso de refinamiento a partir de una división en rejilla de la imagen, donde los superpíxeles finales son el resultado de intercambiando iterativo de píxeles y/o bloques de píxeles entre vecinos, utilizando histogramas de color como medida de semejanza.
- Superpíxeles mediante tasa de entropía (Liu et al., 2011). Este método propone la maximización de una función objetivo consistente en dos componentes: la tasa de entropía de un recorrido aleatorio en el grafo que representa la imagen y una función de equilibrio entre segmentos. La tasa de entropía favorece la formación de conglomerados compactos y homogéneos, mientras que la función de equilibrio fomenta conglomerados de tamaños similares.
- Superpíxeles vía Optimización Pseudo-Booleana (Zhang et al., 2011b). Similar al método de superpíxeles compactos/superpíxeles de intensidad constante, este algoritmo divide la imagen en franjas verticales y horizontales superpuestas de tal manera que cada píxel está cubierto exactamente por dos franjas verticales y dos franjas horizontales. Los píxeles son asignados a una banda horizontal y otra vertical, utilizando una definición propia del concepto de flujo máximo, etiquetando de forma independiente bandas verticales y bandas horizontales. Juntas, las etiquetas correspondientes a las franjas horizontales y las etiquetas de las tiras verticales forman una sobre-segmentación.
- Superpíxeles mediante preservación de topología, TPS (Tang et al., 2012). El objetivo de TPS es generar una segmentación de superpíxe-

les que represente una topología regular de cuadrícula, es decir, que los superpíxeles se puedan organizar en un array en el que cada superpíxel tenga una posición consistente y ordenada. TPS actúa "dibujando" los límites entre superpíxeles a partir de un mapa de las fronteras presentes en la imagen. El método parte de un conjunto de semillas colocadas regularmente en la imagen y traza los caminos óptimos entre estas semillas siguiendo el mapa de fronteras.

- Superpíxeles por relajación de contorno (Conrad et al., 2013). Este enfoque representa una aproximación estadística de la tarea de segmentación por superpíxeles en la que se supone que una imagen es el resultado de múltiples procesos estocásticos, cada uno asociado a un superpíxel, que dan lugar al valor del píxel x_n en el canal de la imagen c . Este conjunto de procesos estocásticos resulta en un conjunto de funciones de energía cuya optimización produce como resultado una sobre-segmentación mediante superpíxeles.
- Superpíxeles basados en inundación (FFC) (Pan et al., 2014). En FFC se plantean una serie de restricciones de compacidad, suavidad y adherencia a los bordes en patrón hexagonal definido por el número de superpíxeles especificado por el usuario. Los superpíxeles se generarán mediante un proceso de crecimiento a partir de los mínimos locales y un refinamiento posterior.
- Superpíxeles mediante caminos aleatorios perezosos (Shen et al., 2014). Este método utiliza el concepto de "lazy random walks" (Aldous y Fill, 2002) para asignar a cada píxel la probabilidad de pertenecer a un superpíxel semilla. Los superpíxeles finales son el resultado de la optimización iterativa de una función de energía definida a partir de las características texturales de la imagen y el tiempo de desplazamiento en los caminos definidos.
- SLIC(Simple Linear Iterative Clustering) (Achanta et al., 2012a). SLIC implementa una agrupación local basada en el método de las K-medias para generar una segmentación de superpíxeles. SLIC puede por tanto clasificarse como método de ascenso de gradiente. Este algoritmo se utiliza a menudo como base para la comparación de otros algoritmos y es particularmente interesante debido a su simplicidad.

Todos los enfoques para generación de superpíxeles plantean sus propias ventajas e inconveniente. Si bien existen estudios que muestran que los algoritmos basados en agrupamiento pueden ser más eficaces que los basados en grafos (ver (Stutz et al., 2018)), la elección final de un método específico dependerá de la aplicación en particular donde estos se vayan a utilizar. En esta Tesis nos hemos decantado por el uso de SLIC.

Simple Linear Iterative Clustering (SLIC) (Achanta et al., 2012a), es una versión espacialmente localizada del método de segmentación de las k-medias. SLIC produce superpíxeles de tamaño regular, distribuidos uniformemente sobre la imagen. Este algoritmo ofrece también un notable control sobre el aspecto de la segmentación final mediante un parámetro de compacidad y es notablemente rápido, cumpliendo con todas las propiedades deseables en un superpíxel (Schick et al., 2012). Las sobre-segmentaciones producidas mediante SLIC se adhieren a las fronteras presentes en una imagen tan bien o mejor que otros métodos, tanto desde el punto de vista de métricas como el error de infra-segmentación, (Neubert y Protzel, 2012), como de la varianza explicada (Moore et al., 2008), tal y como se muestra en (Stutz et al., 2018) o (Csillik, 2016). SLIC ha sido utilizado en el desarrollo de multitud de aplicaciones, como estudios histopatológicos en (Bejnordi et al., 2015) o (Zormpas-Petridis et al., 2018), clasificación de fotografías (Mostajabi et al., 2015) y generación de mapas de cobertura para imagen de satélite (Csillik, 2017). Finalmente, la simplicidad del algoritmo de SLIC nos permite realizar de forma sencilla la extensión a imagen volumétrica o pilas multiespectrales e hiperespectrales; todo esto lo hace especialmente adecuado para el desarrollo de esta Tesis.

3.1.1.1. Generación de superpíxeles mediante SLIC

La generación de superpíxeles mediante SLIC comienza con el muestreo de una serie de semillas espaciadas regularmente, seguido de un proceso de agrupamiento mediante k-medias, donde las semillas se utilizan como centroides para generar segmentos en la primera iteración. SLIC introduce una nueva medida de distancia D_w , donde la distancia entre dos píxeles diferentes aparece ponderada entre la distancia de color, d_c , y la distancia espacial, d_s , definidas como:

$$D_w = \sqrt{\left(\frac{d_c}{m}\right)^2 + \left(\frac{d_s}{In}\right)^2} \quad (3.1)$$

$$d_c = \sqrt{\sum_{s_i \in S} (I(x_1, y_1, s_i) - I(x_2, y_2, s_i))^2} \quad d_s = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (3.2)$$

d_c y d_s representan el color y la distancia espacial entre los píxeles $I(x_1, y_1, s_i)$ y $I(x_2, y_2, s_i)$ en la banda espectral s_i , In es el intervalo de muestreo de los centroides de los grupos de píxeles, y m controla la compacidad de los superpíxeles. La distancia de color asegura la homogeneidad del superpíxel, y la restricción espacial fuerza la compacidad del superpíxel.

3.2. Caracterización de superpíxeles para anotación semántica

Como se muestra en (Galavis et al., 2010), (Balagurunathan et al., 2014) o (Yan et al., 2015), existen múltiples factores que influyen en las características que se extraen de las imágenes, ya sea el tipo de sensor, los ajustes en el momento de adquirir la imagen, el pre-procesado al que han sido sometidas etc... Esto incluye, en el caso de una extracción de características a nivel de objeto, el resultado de la segmentación que se usa como elemento de partida y por tanto, en el caso de partir de una sobre-segmentación, el resultado de dicha sobre-segmentación. Así pues, un paso esencial a la hora de proponer el uso de superpíxeles como elemento base para la extracción de características es saber cómo se van a comportar estas como descriptores visuales, si los resultados que obtengamos pueden ser generalizados y/o hasta qué punto dependen de los parámetros específicos de la sobre-segmentación que estamos usando como base para la extracción de características.

Con objeto de contestar a estas preguntas, se analizará el comportamiento de un conjunto de características extraídas a partir de la sobre-segmentación de un grupo de imágenes, de 512×512 píxeles, parte del conjunto de texturas de referencia de "The VisTex database" (Picard et al., 1995). Con la idea de conseguir un conjunto heterogéneo de imágenes, y por lo tanto significativo, se han escogido imágenes con diferentes combinaciones de contraste, como medida de la uniformidad a nivel local de la imagen, y correlación, como una medida del "orden" de una imagen expresada como la dependencia lineal entre píxeles adyacentes, calculadas a partir de su matriz de co-ocurrencia de grises, (Baraldi y Parmiggiani, 1995). El conjunto de imágenes de prueba se presenta en la figura 3.1.

De cara a analizar el comportamiento de la extracción de características a nivel de superpíxel se considerará cada imagen como una única unidad textural. Para cada una de las imágenes se generará un conjunto de sobre-segmentaciones utilizando SLIC, variando el tamaño del superpíxel y su factor de compacidad. Específicamente se utilizarán tamaños comprendidos entre los 3 y los 65 píxeles, con un intervalo de 2 píxel entre cada prueba, y factores de compacidad situados entre las 50 y las 2425 unidades, con un intervalo de 75 entre cada una de las pruebas. Esto da como resultado un total de 32×32 posibles combinaciones de parámetros para la segmentación, 1024 sobre-segmentaciones por imagen. Sobre estas sobre-segmentaciones se aplicará un descriptor visual específico y, puesto que se consideran las imágenes como unidades texturales únicas, para cada sobre-segmentación se utilizará la media del vector de características extraído. Se presenta la media para el grupo de imágenes de prueba y mayores y menores valores obtenidos. De cara a facilitar la visualización y el análisis de los resultados, estos se presentarán como gráficos de contornos con 15 isolíneas marcadas.

Figura 3.1: Conjunto de imágenes de prueba seleccionado de la "VisTex database". En cada imagen se muestra el nombre que se utilizará para denominarla y sus valores globales de correlación (corrl.) y contraste (contr.).

3.2.1. Caracterización de superpíxeles mediante sus momentos estadísticos

Con el objetivo de analizar las diferencias en el comportamiento de los momentos estadísticos cuando son extraídos a partir de una imagen sobre-

Característica	Tamaño del superpíxel			Valor de compacidad		
	Mínimo	Máximo	Media	Mínimo	Máximo	Media
Media	-0.5556	0.3275	-0.2418	-0.4422	0.1569	-0.1800
Desviación est.	0.6856	0.8523	0.7559	0.4238	0.6077	0.5101
Asimetría	-0.8609	0.8953	0.4662	-0.2141	0.4698	0.1674
Curtosis	-0.4023	0.8204	0.1406	-0.7477	0.0888	-0.5562

Tabla 3.1: Coeficiente de correlación de Pearson entre los diferentes momentos estadísticos extraídas y los parámetros de la sobre-segmentación. Se presenta la media para el grupo de imágenes de prueba y mayores y menores valores obtenidos.

segmentada se ha realizado el cálculo de la media, la desviación estándar, la asimetría y la curtosis. En la Tabla 3.1 se muestra la magnitud de la asociación (lineal) entre los valores obtenidos para las caracterizaciones mediante momentos estadísticos y los dos parámetros de los que depende la sobre-segmentación, medida mediante el coeficiente de correlación de Pearson. Puede verse tanto la media como los mayores y menores valores en el grupo de imágenes de prueba. La Fig. 3.2 representa el gráfico de contorno para las pruebas a partir de la media calculado a nivel de sobre-segmentación. Como podemos observar en dicha figura, si bien aparecen variaciones significativas para diferencias pequeñas en los parámetros de la sobresegmentación, la media de los superpíxeles se mantiene en general próxima a la media de la imagen, sin que exista una asociación entre los cambios en los parámetros de la segmentación y la media de los superpíxeles de la unidad textural. La combinación de la media con una ventana deslizante, presentada en (Huang et al., 1979), es un filtro para el suavizado de imágenes bien conocido, sin embargo la media calculada a nivel de superpíxel resulta en una cuantización o "resumen" de la imagen, conservando las fronteras presentes en la imagen.

Existe una relación directa entre la desviación estándar media de una imagen calculada a nivel de superpíxel y los parámetros de dicha segmentación, tal y como puede apreciarse tanto por sus valores de correlación como en Fig. 3.3, donde se representa la gráfica de contorno para los resultados de la desviación estándar calculada a partir de una sobre-segmentación para el conjunto de pruebas. Esta relación está especialmente marcada para el tamaño del superpíxel (Coeficiente de correlación de Pearson medio 0.8523, valor- $P < 0,05$ en todas las pruebas) que para su compacidad. Estas relaciones son de esperar, no solo áreas más pequeñas suelen implicar mayor homogeneidad local sino que el parámetro de compacidad usado en SLIC está específicamente diseñado para reducir la varianza en la generación de superpíxeles al lastrar la heterogeneidad espectral local.

Al ser relaciones homogéneas, y generalmente centradas, la asimetría y

la curtosis tienen la ventaja de ser invariables a los cambios de luminancia en las imágenes, lo cual en principio los hacen muy interesantes para su uso como características de clasificación. Sin embargo, estos estimadores no muestran un comportamiento predecible cuando son calculados a partir de distribuciones multimodales (Von Hippel, 2005), (Westfall, 2014). Una de las ventajas de su uso con superpíxeles es que, dada su naturaleza de elemento pequeño y visualmente homogéneo, las posibilidades de que se den superpíxeles con distribución multimodal son reducidas. Sin embargo, ni a partir de los resultados de correlación ni en Fig. 3.4 y Fig. 3.5 es posible mantener que el comportamiento de la caracterización partiendo de superpíxeles tenga relación con los parámetros de la segmentación, dependiendo más de las características específicas de la textura.

Los experimentos practicados en el conjunto de pruebas utilizando la media, la desviación estándar, la asimetría y la curtosis calculadas a partir de una sobre-segmentación muestran un comportamiento estable. Esto decir, pequeños cambios en los parámetros de tamaño del superpíxel y valor de compacidad se asocian a pequeños cambios en los resultados de la extracción de características a partir de una segmentación mediante superpíxeles utilizando SLIC. Así mismo, tanto media como desviación estándar muestran un comportamiento predecible en relación a los parámetros de la sobre-segmentación, si bien esto no pudo ser afirmado categóricamente en relación a la asimetría y la curtosis.

3.2.2. Características de forma aplicadas a superpíxeles

Una de las ventajas que comparte la extracción de características utilizando una sobre-segmentación con el uso de caracterizaciones a partir de objetos es la existencia de áreas definidas en la imagen, lo que implica la presencia de formas, con propiedades específicas y medibles. Dada la sencillez estructural de un superpíxel, no todas las características de forma pueden ser aplicadas a este caso. Se proponen por tanto algunas de las más conocidas y utilizadas (Loncaric, 1998).

- Solidez: La solidez es la relación entre los píxeles que compone un superpíxel con los píxeles que se encuentran dentro de su caja delimitadora asociada.
- Número de Euler: El número de Euler es la relación entre el número de píxeles dentro del polígono delimitador convexo y aquellos que forman parte del área a la que se asocia dicho polígono.
- Excentricidad: La excentricidad se define como el ángulo entre el eje x y el eje mayor de una elipse con los mismos segundos momentos que el superpíxel. El valor está en grados, oscilando entre -90 grados y 90 grados, siendo normalizado a valores entre 0 y 1.

Figura 3.2: Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la media calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación la intensidad media de cada imagen (M.).

Figura 3.3: Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la desviación estándar calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación la desviación estándar de cada imagen (D.E.).

Figura 3.4: Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la asimetría calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación la asimetría de cada imagen (A.).

Figura 3.5: Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la curtosis calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación la curtosis media de cada imagen (C.).

Característica	Tamaño del superpíxel			Valor de compacidad		
	Mínimo	Máximo	Media	Mínimo	Máximo	Media
Solidez	0.1311	0.5130	0.3240	0.7592	0.9380	0.8476
Excentricidad	-0.7707	-0.4993	-0.6913	-0.8083	0.0574	-0.3660
N. de Euler	-0.1365	0.2457	0.1180	-0.6554	-0.3942	-0.5462

Tabla 3.2: Coeficiente de correlación de Pearson entre las diferentes características de forma extraídas y los parámetros de la sobre-segmentación. Se presenta la media para el grupo de imágenes de prueba y mayores y menores valores obtenidos.

En la Tabla 3.2 se muestra la magnitud de la asociación (lineal) entre los valores obtenidos para las caracterizaciones mediante las características de forma propuestas y los dos parámetros de los que depende la sobre-segmentación, medida mediante el coeficiente de correlación de Pearson. Puede verse tanto la media como los mayores y menores valores en el grupo de imágenes de prueba.

Mayores valores de solidez aparecen asociados a valores de compacidad más altos. Existe por tanto una relación entre la solidez media de una textura, y la compacidad del superpíxel (Coeficiente de correlación de Pearson medio 0.8476, valor-P $< 0,05$ en todas las pruebas), calculada a partir de una sobre-segmentación, lo cual resulta también evidente en el gráfico de contorno para la solidez, Fig. 3.6. En el caso del número de Euler, si bien se aprecia en Fig. 3.7 como los mayores valores se concentran cuando tenemos superpíxeles grandes y muy compactos, esta relación no parece ser lineal para todo el rango de valores en los parámetros, y no se refleja en las medidas de correlación entre el número de Euler y los parámetros de la sobre-segmentación.

No es posible extraer una tendencia general de Fig. 3.8, más allá de un incremento en la excentricidad para valores reducidos de compacidad. Aunque esta relación se ve reflejada hasta cierto punto en la magnitud de la asociación entre compacidad de superpíxel y excentricidad, teniendo en cuenta su definición, el valor de la excentricidad estará más relacionado con la geometría específica de la imagen que con los parámetros de una sobre-segmentación.

Puede concluirse, para el caso de características de forma aplicadas a superpíxeles, que si bien como era de esperar son sensible a los parámetros utilizados en la sobre-segmentación, en el caso de la solidez y el número de Euler se muestra un comportamiento estable y predecible. La excentricidad aparece demasiado marcada por el resultado de la sobre-segmentación como para poder afirmar categóricamente algo sobre su comportamiento, si bien existe una tendencia a incrementarse con el uso de superpíxeles cada vez más

Figura 3.6: Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la solidez calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto.

Figura 3.7: Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando el número de Euler calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto.

Figura 3.8: Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la excentricidad calculada a partir de una sobresegmentación (valor de intensidad), para el conjunto de pruebas propuesto.

compactos.

3.2.3. Matriz de co-ocurrencia de grises y medidas derivadas

Recordemos las definiciones dadas en (Haralick y Shapiro, 1992) y (Jain et al., 1995) de textura, esto es, la variación de intensidad en el entorno inmediato de cada píxel y la distribución espacial de los niveles de intensidad en una región dada de una imagen digital. Podemos ver un factor común entre ambas; la presencia de una vecindad o una región dada en la imagen, sin que exista una definición específica de estas. Tradicionalmente esto implica métodos basados en ventanas deslizantes, una división en pequeñas áreas cuadradas, ventanas, asociadas a cada píxel. Estas áreas no guardan una relación real con el vecindario perceptual de un píxel, no respetando fronteras o estructuras presentes en la imagen.

Dada la dependencia de las características de Haralick de un área específica para el cálculo de una matriz de co-ocurrencia, la posibilidad de que esta sea un área homogénea, con significado perceptual propio justifica su cálculo a partir de una segmentación de superpíxeles. Dentro del conjunto de características propuestas en (Haralick y Shapiro, 1992) se seleccionará el contraste, la homogeneidad, la entropía y la correlación para la realización de las pruebas.

- La entropía representa una medida de la aleatoriedad o el grado de desorden presente en la imagen. El valor de entropía es mayor cuando todos los elementos de la matriz de co-ocurrencia son iguales y disminuye según aumentan la disparidad entre estos.
- El contraste es una medida de las variaciones de intensidad o niveles de grises entre el píxel de referencia y su vecino. Mayores contrastes reflejan grandes diferencias de intensidad en la matriz de co-ocurrencia de grises.
- La homogeneidad mide la proximidad de la distribución de los elementos de la matriz de co-ocurrencia a su diagonal. A medida que aumenta la homogeneidad, el contraste, por lo general, disminuye.
- La correlación se deriva del segundo momento angular de la matriz de co-ocurrencia y representa una medida de la similitud de un píxel con sus vecinos en el área usada para crear dicha matriz.

En la Tabla 3.3 se muestra la magnitud de la asociación (lineal) entre los valores obtenidos para las caracterizaciones mediante las características de Haralick propuestas y los dos parámetros de los que depende la sobre-segmentación, medida mediante el coeficiente de correlación de Pearson. Puede verse tanto la media como los mayores y menores valores en el grupo de imágenes de prueba.

Característica	Tamaño del superpíxel			Valor de compacidad		
	Mínimo	Máximo	Media	Mínimo	Máximo	Media
Correlación	0.3899	0.6085	0.5108	-0.0354	0.0798	0.0121
Contraste	0.5498	0.8834	0.7400	0.0713	0.1256	0.0986
Entropía	-0.4258	-0.3739	-0.4015	-0.0645	-0.0210	-0.0341
Homogeneidad	-0.6157	-0.5050	-0.5779	-0.0946	-0.0325	-0.0603

Tabla 3.3: Coeficiente de correlación de Pearson entre las diferentes características de Haralick extraídas y los parámetros de la sobre-segmentación. Se presenta la media para el grupo de imágenes de prueba y mayores y menores valores obtenidos.

Puede verse en los gráficos de contorno calculados a partir de sobre-segmentaciones mostrados en Fig. 3.9, Fig. 3.10, Fig. 3.11 y Fig. 3.12, correspondientes a entropía, homogeneidad, contraste y correlación, respectivamente, como para estas características existe una relación, directa en el caso de la correlación y el contraste, inversa para entropía y la homogeneidad, entre el tamaño del superpíxel y su valor, por encima de los cambios que puedan darse debido al uso de diferentes factores de compacidad. Atendiendo a los valores de correlación entre estas caracterizaciones y los parámetros de la sobre-segmentación, esta relación no puede considerarse lineal; mientras que para tamaños de superpíxel altos los valores de la caracterización tienden a permanecer estable, tamaños de superpíxel pequeños (a partir de tamaños menores de 10 píxeles, aproximadamente) hacen que esta descienda rápidamente.

Además, en Fig. 3.11 se muestra una relación directa entre el contraste de una imagen y el contraste de dicha imagen calculado utilizando una sobre-segmentación, siendo esta inferior en aquellas texturas con un contraste global menor (Piedra 4, Arena 2, Terreno 8, Tela 16, Ladrillo 2). En menor medida, esto también ocurre para la homogeneidad calculada utilizando una sobre-segmentación, como puede verse en Fig. 3.10). Por último, a partir de Fig. 3.12 se puede constatar como los valores de correlación dependerán de la presencia de áreas visualmente homogéneas dentro de la imagen, como se dan, por ejemplo en Arena 2 o Terreno 2, o en menor medida Tela 16 y Tela 2. Los experimentos con las cuatro características de Haralick propuestas muestran como no solamente se comportan de forma estable y hasta cierto punto, predecible, sino que su valor dependerá de características globales de la imagen.

Figura 3.9: Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la entropía calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación la entropía media de cada imagen (E.).

Figura 3.10: Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la homogeneidad calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación el contraste medio para cada imagen de las diferentes pruebas (Hom.).

Figura 3.11: Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando el contraste calculado a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación el contraste medio para cada imagen de las diferentes pruebas (Ctr.).

Figura 3.12: Gráfica de contornos para los conjuntos de prueba, mostrando la relación entre el tamaño del superpíxel (representado en las abscisas), el parámetro de compacidad (representado en las ordenadas) y la caracterización de la imagen utilizando la correlación calculada a partir de una sobre-segmentación (valor de intensidad), para el conjunto de pruebas propuesto. Se incluye a efectos de comparación la correlación media para cada imagen de las diferentes pruebas (Cor.).

Capítulo 4

Segmentación de imágenes a color utilizando un índice de rugosidad basado en superpíxeles

Uno de los métodos de segmentación de imágenes más sencillos es la umbralización mediante histogramas. Estos métodos operan en el dominio espectral de la imagen; no se consideran las relaciones espaciales entre píxeles adyacentes, únicamente los niveles de intensidad en cada canal de color. Cada uno de los canales de color de la imagen es segmentada de forma independiente sin tener en cuenta posibles relaciones espaciales de intensidades intra e inter-canal. Una mejora evidente a estos métodos es su ampliación al dominio espacial de una imagen. Esto abre la posibilidad de extender métodos de segmentación basados en histogramas adaptándolos a su uso con superpíxeles.

Los métodos de segmentación basados en histogramas asumen una correspondencia entre regiones de imagen homogéneas y picos en los histogramas. En su forma más sencilla, se supone la existencia en la imagen de dos regiones diferenciadas: un objeto y un fondo. El objeto estará definido por un conjunto de niveles de intensidad representado por uno de los picos más prominentes en el histograma. El fondo de la imagen estará definido por otro conjunto de niveles de intensidad, representado en el histograma por otro de los picos más destacados. Los valles, el nivel de gris con menos representantes en el histograma entre los dos picos, se utilizarán para trazar los rangos de los grupos de niveles intensidad de objeto y fondo. A pesar de las evidentes limitaciones del método, la segmentación mediante umbralización aparece como base en diversos trabajos de anotación semántica tales como (Kartikeyan et al., 1998), (Antonie, 2008) o (Saad et al., 2011). Para imáge-

nes más complejas es posible aplicar este método de forma recursiva en los rangos calculados, dividiéndolos en rangos más pequeños, como se puede ver en (Zhang et al., 2008b), (Al-amri et al., 2010) o (He et al., 2009).

El concepto de histón, propuesto por Mohabey y Ray en (Mohabey y Ray, 2000), permite extender el concepto de histograma añadiendo relaciones entre diferentes canales de color. Si bien los resultados son superiores a la umbralización mediante histogramas, las segmentaciones basadas sólo en histones tienden a sub-representar pequeñas regiones homogéneas en la imagen, por lo que se desarrolló una extensión a este trabajo, utilizando histogramas e histones desde la perspectiva de la teoría de conjuntos aproximados en (Mushrif y Ray, 2008). Esta extensión a la segmentación mediante histones (cuyos fundamentos se describirán en la siguiente sección) produce resultados notablemente superiores comparados tanto con su método original como con otros métodos basados en histogramas. La simplicidad de esta propuesta y sus buenos resultados lo convierten en una buena opción para su uso como método de segmentación no supervisado de imágenes.

El índice de rugosidad desarrollado por Mushrif y Ray se basa en una ventana de 3×3 píxeles alrededor de cada píxel, un concepto de vecindario muy limitado, como puede verse en (Yue et al., 2012). Esto hace que regiones pequeñas aparezcan ahora sobre-representadas, hasta tal punto que el ruido presente en estas puede tener el mismo peso que regiones homogéneas más grandes y significativas. Una mejora evidente es el uso de un vecindario más flexible, directamente relacionado con las características del píxel analizado.

4.1. Materiales y métodos

El método original de Mohabey y Ray se divide en 3 fases: cálculo del índice de rugosidad para cada canal de color; selección de picos y sus correspondientes valles, utilizando el índice de rugosidad y la fusión de los segmentos generados en la fase anterior, utilizando el concepto de distancia de color. En este capítulo se propone un método de segmentación basado en el de Mohabey y Ray que utiliza como vecindario para calcular el índice de rugosidad una sobre-segmentación previa de la imagen mediante superpíxeles, en lugar de una ventana pre-definida para cada píxel. Teniendo en cuenta la definición de un superpíxel como una región perceptualmente uniforme, esta sobre-segmentación representa una forma natural de proporcionar un vecindario adaptado a las características de la imagen y por tanto una mejora del método original. Un esquema del método propuesto puede verse en Fig. 4.1.

Previamente a la descripción detallada del método propuesto, se van a desarrollar las bases teóricas del mismo en las siguientes subsecciones.

4.1.1. Conjuntos aproximados

Un conjunto aproximado ("rough set"), descrito por primera vez por el matemático Zdzislaw I. Pawlak en (Pawlak, 1982) y más tarde formalizado en (Pawlak, 1992), es una aproximación al concepto de vaguedad o ambigüedad. Consideremos un sistema de información $S = (U, A)$ donde U , el universo, es un conjunto finito, no vacío de objetos $U = u_1, u_2, \dots, u_n$, y A es un conjunto finito, no vacío de atributos $A = a_1, a_2, \dots, a_n$. Cada atributo $a \in A$ define una función de información $f_a : U \rightarrow V_a$ donde V_a es el dominio del atributo a , definido como el conjunto de valores de a . Por cada $B \subseteq A$ hay una relación de equivalencia asociada $IND(B)$ definida como:

$$IND(B) = \{(u, u') \in U \times U \mid \forall a \in B \ a(u) = a(u')\} \quad (4.1)$$

donde $IND(B)$ se denomina relación de indiscernibilidad. Si $S = (u, u') \in IND(B)$, los objetos x y x' son indiscernibles entre sí teniendo en cuenta los atributos de B , ya que tienen los mismos valores para este conjunto de atributos B . $U/IND(B)$ indica el conjunto de clases de equivalencia de $IND(B)$. En general, un conjunto $X \subseteq U$ no podrá representarse exactamente utilizando las clases de equivalencia inducidas por un subconjunto de atributos P , ya que es posible que el conjunto contenga objetos indistinguibles en función de estos atributos. Sin embargo, dado cualquier subconjunto de atributos P , $X \subseteq U$ puede definirse de forma "aproximada" empleando dos conjuntos exactos llamados aproximaciones inferiores ($\underline{P(X)}$) y superiores ($\overline{P(X)}$) donde

$$\underline{P(X)} = \bigcup_{u \in U} \{P(u) \mid P(u) \subseteq X\} \quad (4.2)$$

$$\overline{P(X)} = \bigcup_{u \in U} \{P(u) \mid P(u) \cap X \neq \emptyset\} \quad (4.3)$$

La aproximación inferior consiste en todas los elementos que con seguridad pertenecen al conjunto U , la aproximación superior consiste en todos los elementos que posiblemente pertenezcan al conjunto U . La diferencia entre la aproximación superior e inferior se denomina región límite y define el conjunto de objetos que no se pueden incluir realmente en X , pero que tampoco pueden ser descartados. A diferencia de otros métodos de análisis de incertidumbre, el análisis aproximado de conjuntos no requiere información adicional o interpretaciones subjetivas para determinar la composición del conjunto, sólo utiliza la información presentada dentro de los datos dados (Duntsch y Gediga, 1997). La teoría de conjuntos aproximados proporciona una forma objetiva de análisis (Pawlak et al., 1995).

4.1.2. Histón

Una histón (Mohabey y Ray, 2000) es un contorno añadido en la parte superior de cualquiera de los histogramas existentes de los componentes espectrales de la imagen. Explota la correlación entre los píxeles vecinos tanto en el mismo como en otros planos espectrales como método para la segmentación de la imagen. En un histón, la colección de todos los puntos que caen bajo la esfera de color similar¹ de un radio predefinido, llamado umbral de similitud o "expanse", (E), pertenece a una única categoría de dicho histograma.

Para generar un histón se evalúa el número de puntos encapsulados en la esfera de color similar, para cada categoría de intensidad g en el histograma base, y se añaden en el histograma a dicha categoría. La definición de un histón, en una imagen $I(x, y, s, s)$ de tamaño $M \times N$, donde s representan los canales de color de la imagen, se presenta como:

$$H_{s_i} = \sum_{x=1}^N \sum_{y=1}^M (1 + S(x, y)) \delta(I(x, y, s_i) - g) \quad \text{for } 0 \leq g \leq L-1 \text{ and } s_i \in S_p \quad (4.4)$$

donde $\delta(\cdot)$ es la delta de Kronecker, L es el número total de niveles de intensidad en cada una de las componentes espectrales (por lo tanto, $\delta(I(x, y, s_i) - g)$ es una definición de histograma) y $S(x, y)$ es una función de similitud definida para comprobar si un componente de este vecindario pertenece a la esfera de color similar. Esta función viene definida por la ecuación 4.5 y se basa en una medida de distancia, $d_t(x, y)$, correspondiendo esta a la suma de las distancias de color de una de los canales $s_1, \dots, s_i, \in S_p$ que componen la imagen, en cualquier píxel (x, y) de un vecindario de tamaño $P \times Q$.

$$S(x, y) = \begin{cases} 1 & \text{if } d_t(x, y) < E \\ 0 & \text{otherwise} \end{cases} \quad (4.5)$$

$$d_t(x, y) = \sum_{p \in P} \sum_{q \in Q} \sqrt{\sum_{s_i \in S_p} (I(x, y, s_i) - I(p, q, s_i))^2}$$

4.1.3. Índice de rugosidad

Histogramas e histones pueden ser relacionados desde el punto de vista de la teoría de conjuntos aproximados. El valor del histograma tradicional ($h_{s_i}(g)$) se define como una aproximación inferior, una colección de píxeles con un valor espectral específico; y su histón derivado ($H_{s_i}(g)$) se define como una aproximación superior, una colección de píxeles que podrían pertenecer

¹La esfera de color similar es la región en el espacio de intensidad espectral donde todos los valores de intensidad que caen en esa región pueden clasificarse como el mismo valor de color

Figura 4.1: Diagrama de flujo del método desarrollado para segmentación de imágenes mediante un índice de rugosidad basado en superpíxeles.

a un valor espectral específico. Las diferencias entre la aproximación inferior y superior en una categoría de intensidad muestran la falta o presencia de homogeneidad en color o similitud espacial. Para medir estas diferencias, en (Mushrif y Ray, 2008) se propone un índice de rugosidad ("roughness index"), cuya expresión correspondiente se define como:

$$\rho = 1 - \frac{|h_{s_i}(g)|}{|H_{s_i}(g)|}, \quad 0 \leq g \leq L - 1, \quad s_i \in S_p \quad (4.6)$$

En las regiones de color homogéneas, el índice de rugosidad tenderá a acercarse a 1, mientras que en las regiones límite, será más próximo a 0.

4.1.4. Segmentación mediante un índice de rugosidad basado en superpíxeles

De la misma forma que el método original, el método desarrollado se divide en 3 fases: cálculo del índice de rugosidad; selección de picos y valles y la fusión de segmentos generados en la fase anterior, utilizando el concepto de distancia de color. A diferencia del método original, el cálculo del índice de rugosidad ya no depende de un vecindario predefinida para cada píxel, sino que se utiliza una sobre-segmentación previa de la imagen mediante superpíxeles, para definir dicho vecindario. No solo eso, el uso de una sobre-segmentación mediante superpíxeles también permite definir el "expanse" en la generación de los histones de forma dinámica. Un esquema del método desarrollado puede verse en Fig. 4.1.

4.1.4.1. Superpíxeles como primitiva para el cálculo de índice de rugosidad

El uso de superpíxeles como vecindario para el cálculo de histones garantiza una relación espacial directa entre un píxel y su vecindario, respetando las características y fronteras entre las áreas presentes en la imagen. De esta manera, los superpíxeles permiten definir lo que consideramos una región homogénea, no basada en un valor fijo, sino adaptada a las características específicas de la imagen. Puesto que la segmentación utilizando superpíxeles ya produce áreas localmente homogéneas, es posible cuantificar la homogeneidad local utilizando la desviación estándar media de la intensidad del conjunto de superpíxeles. Así, a diferencia del "expanse" utilizado por Mohabey y Ray, un umbral de valor "100", es posible definir este de forma dinámica para cada imagen. Esto es, un píxel estará dentro de la esfera de color si la distancia entre la intensidad media del superpíxel del que forma parte y la intensidad de dicho píxel es inferior a la desviación estándar local media del plano de color en el espacio definido por los superpíxeles, en cada una de sus componentes de color.

Denotemos el conjunto de superpíxeles resultantes de una segmentación de imagen como $Cs = \{Cs_1, Cs_2, \dots, Cs_{Ns}\}$ donde Ns es el número total de superpíxeles, $I(x_{Cs_j}, y_{Cs_j}, s_i)$ representa el centroide del superpíxel Cs_j en canal de color $s_i \in S_p$ y Np_j es el número de píxeles en dicho superpíxel; la función de similaridad $S_2(x, y)$ se define como:

$$S_2(x, y) = \begin{cases} 1 & \text{if } \forall s_i \in S_p \quad d_{ts_i}(x, y) < E_{s_i} \\ 0 & \text{otherwise} \end{cases} \quad (4.7)$$

$$d_{ts_i}(x, y) = |I(x, y, s_i) - I(x_{Cs_j}, y_{Cs_j}, s_i)| \quad I(x, y, s_i) \in Cs_j \quad (4.8)$$

$$E_{s_i} = \frac{1}{Ns} \sum_{j=1}^{Ns} \frac{1}{Np_j} \sum_{I \in Cs_j} I(x, y, s_i) \quad (4.9)$$

A partir de esta formulación para S_2 , el cálculo de un histón ($H_{s_{2i}}(g)$) derivado del histograma ($h_{s_i}(g)$) se definirá como:

$$H_{s_{2i}} = \sum_{x=1}^N \sum_{y=1}^M (1 + S(x, y)) \delta(I(x, y, s_i) - g) \quad \text{for } 0 \leq g \leq L-1 \text{ and } s_i \in S_{2p} \quad (4.10)$$

quedando la expresión correspondiente al índice de rugosidad basado en superpíxeles como:

$$\rho = 1 - \frac{|h_{s_i}(g)|}{|H_{s_{2i}}(g)|}, \quad 0 \leq g \leq L-1, \quad s_i \in S_{2p} \quad (4.11)$$

4.1.4.2. Selección de picos y valles y fusión de regiones

Después de seleccionar los picos significativos de forma similar al método original (un pico se considera significativo si su altura es superior al 20 % del valor medio del índice de rugosidad y se encuentra a más de 10 unidades de intensidad de otro pico), los valles se obtienen encontrando los valores mínimos entre picos adyacentes. De acuerdo con la ubicación de los picos y valles, los segmentos para cada canal de color se calculan y combinan en una segmentación base.

La segmentación base obtenida a partir de picos y valles genera muchas regiones pequeñas, algunas de ellas compuestas por muy pocos píxeles, por lo que resulta necesario realizar la fusión de regiones. Al igual que en el método original, se ha utilizado el proceso propuesto por (Cheng et al., 2002a) para la fusión de regiones, un método iterativo donde se combinan regiones de acuerdo a una distancia de color definida como la diferencia mínima entre valores espectrales medios correspondientes.

4.2. Resultados

El método desarrollado se evaluará utilizando un conjunto de 20 imágenes provenientes de la "Berkeley Segmentation Dataset and Benchmark" (Martin et al., 2001). Como métrica de evaluación se ha elegido el criterio agregado de Levine y Nazif (Levine y Nazif, 1985a). Este criterio valora la distribución de intensidades en la imagen en relación a la segmentación, lo que le hace especialmente adecuado para la evaluación de métodos de segmentación dependientes del particionado de las intensidades de los canales de color, como el presentado. A pesar de su antigüedad, este criterio ha demostrado un rendimiento similar o mejor que criterios de evaluación no supervisados más modernos (Chabrier et al., 2006).

El criterio agregado de Levine y Nazif evalúa el equilibrio en una segmentación entre la uniformidad intra-regional (definida como uno menos la varianza normalizada de intensidades dentro de cada región) y la disparidad inter-regional (como la suma ponderada de los contrastes de las regiones, normalizada utilizando el porcentaje de bordes compartidos entre clases). La uniformidad intra-regional y la disparidad inter-regional son medidas complementarias que evalúan la homogeneidad y el carácter distintivo de las regiones resultado de una segmentación. Uniformidad intra-regional, disparidad inter-regional y criterio agregado de Levine y Nazif aumentan su valor a medida que aumenta la calidad percibida de la segmentación.

En la Tabla 4.1 se compara el número de segmentos obtenidos (tanto para cada canal de color como el total tras el proceso de fusión) con el método desarrollado con los que resultan utilizando únicamente una segmentación mediante histogramas y con una implementación del método original

Figura 4.2: Conjunto de imágenes utilizados como prueba.

Imagen	Método	Número de agrupaciones				Imagen	Método	Número de agrupaciones			
		Azul	Verde	Rojo	Total			Azul	Verde	Rojo	Total
im1	Hist	5	9	8	27	im11	Hist	10	7	10	79
	I.R.	21	9	20	87		I.R.	18	10	17	120
	I.R.Sp	23	11	21	94		I.R.Sp	23	11	22	120
im2	Hist	11	6	15	128	im12	Hist	11	11	7	188
	I.R.	18	12	21	230		I.R.	20	22	12	357
	I.R.Sp	24	13	21	224		I.R.Sp	23	33	20	372
im3	Hist	9	4	10	66	im13	Hist	5	3	4	19
	I.R.	22	11	19	201		I.R.	6	6	10	29
	I.R.Sp	23	12	21	203		I.R.Sp	22	12	21	58
im4	Hist	9	4	4	43	im14	Hist	10	5	9	90
	I.R.	12	9	21	142		I.R.	18	10	14	172
	I.R.Sp	22	10	22	157		I.R.Sp	21	13	19	188
im5	Hist	7	5	9	64	im15	Hist	14	7	10	95
	I.R.	18	11	18	135		I.R.	14	7	11	103
	I.R.Sp	19	11	22	145		I.R.Sp	19	10	21	133
im6	Hist	10	6	11	134	im16	Hist	8	3	8	69
	I.R.	14	10	15	176		I.R.	21	14	21	173
	I.R.Sp	25	13	18	212		I.R.Sp	23	13	25	178
im7	Hist	15	8	4	138	im17	Hist	13	5	4	72
	I.R.	23	11	16	260		I.R.	18	10	15	146
	I.R.Sp	25	13	21	282		I.R.Sp	24	13	24	153
im8	Hist	7	3	6	37	im18	Hist	9	7	8	19
	I.R.	21	10	13	140		I.R.	17	11	13	46
	I.R.Sp	23	13	22	145		I.R.Sp	21	11	19	56
im9	Hist	8	5	5	92	im19	Hist	13	7	12	137
	I.R.	12	7	13	171		I.R.	19	9	13	168
	I.R.Sp	21	13	22	248		I.R.Sp	21	11	18	181
im10	Hist	12	6	10	105	im20	Hist	8	6	7	52
	I.R.	21	11	17	148		I.R.	18	25	18	162
	I.R.Sp	23	12	22	164		I.R.Sp	24	34	20	156

Tabla 4.1: Comparación del número de segmentos obtenidos mediante el método desarrollado con los resultados de un método de umbralización basado en histogramas y los resultados del método basado en el índice de rugosidad, tanto por canal de color como para el resultado final combinado.

mediante índice de rugosidad de Mushrif y Ray (2008). Podemos observar como tener en cuenta la distribución espacial da lugar a un número considerablemente mayor de segmentos, tanto para cada canal de color como para la segmentación final, especialmente en el método desarrollado. Si bien esto no garantiza una mejor segmentación, especialmente cuando va a haber una fase de fusión para reducir el número final de segmentos espectrales, abre la posibilidad de una representación más precisa de los matices y detalles presentes en la imagen.

En la Tabla 4.2 se presentan los resultados de la evaluación cuantitativa no supervisada utilizando el criterio agregado de Levine y Nazif. En esta Tabla se muestran los valores obtenidos la uniformidad intra-regional, la disparidad inter-regional y el total agregado, para segmentaciones resultado de fijar en 5, 10 y 15 el número de segmentos. Si bien no ocurre en todas las

		Uniformidad Intra-regional			disparidad inter-regional			Levine& Nazif agregado		
		Histogr.	Histón	Superpx.	Histogr.	Histón	Supepx.	Histogr.	Histón	Superpx.
im1	5	0.84926	0.83327	0.84346	0.15369	0.1639	0.18112	0.50147	0.49859	0.51229
	10	0.88262	0.86722	0.88569	0.12959	0.12768	0.13338	0.50611	0.49745	0.50954
	15	0.89075	0.88211	0.89429	0.1174	0.11755	0.11686	0.50407	0.49983	0.50557
im2	5	0.55773	0.41819	0.56356	0.28382	0.26101	0.32267	0.42077	0.3396	0.44311
	10	0.53589	0.55151	0.5904	0.24949	0.23026	0.25826	0.39269	0.39088	0.42433
	15	0.4938	0.56725	0.60313	0.22383	0.22739	0.22998	0.35882	0.39732	0.41655
im3	5	0.54379	0.59319	0.63313	0.31582	0.30811	0.32038	0.4298	0.45065	0.47676
	10	0.42623	0.52946	0.5496	0.24008	0.21433	0.22061	0.33316	0.3719	0.38511
	15	0.31117	0.53155	0.52226	0.20513	0.20777	0.21574	0.25815	0.359	0.369
im4	5	0.70967	0.65572	0.66235	0.24404	0.23634	0.23462	0.47685	0.44603	0.44848
	10	0.66558	0.66806	0.71998	0.19253	0.1971	0.1862	0.42906	0.43258	0.45309
	15	0.60566	0.67967	0.70185	0.18081	0.1854	0.17532	0.39323	0.43254	0.43859
im5	5	0.72404	0.73349	0.73166	0.26862	0.27375	0.27761	0.49633	0.50362	0.50463
	10	0.7439	0.73914	0.75282	0.25075	0.24069	0.2532	0.49732	0.48991	0.50301
	15	0.7568	0.71968	0.76818	0.23442	0.23171	0.23377	0.49561	0.47569	0.50098
im6	5	0.7198	0.69079	0.66364	0.23849	0.23646	0.22222	0.47915	0.46362	0.44293
	10	0.67692	0.60282	0.59583	0.20204	0.20596	0.22374	0.43948	0.40439	0.40978
	15	0.64917	0.5731	0.59894	0.19512	0.19472	0.22045	0.42214	0.38391	0.40969
im7	5	0.49778	0.50304	0.55558	0.37771	0.39091	0.37996	0.43774	0.44698	0.46777
	10	0.40227	0.52329	0.47295	0.36964	0.35057	0.36452	0.38595	0.43693	0.41874
	15	0.23055	0.45262	0.44492	0.3394	0.34397	0.33495	0.28497	0.39829	0.38993
im8	5	0.74974	0.76523	0.7547	0.20931	0.22245	0.25269	0.47952	0.49384	0.5037
	10	0.68295	0.72745	0.70772	0.15597	0.17162	0.18853	0.41946	0.44954	0.44813
	15	0.60296	0.72038	0.73091	0.14502	0.17184	0.16731	0.37399	0.44611	0.44911
im9	5	0.38465	0.49631	0.56046	0.28703	0.29974	0.29163	0.33584	0.39802	0.42605
	10	0.25668	0.43138	0.531	0.21967	0.21021	0.21884	0.23818	0.32079	0.37492
	15	0.18785	0.4103	0.49087	0.20052	0.18866	0.21167	0.19419	0.29948	0.35127
im10	5	0.62236	0.69909	0.61413	0.21303	0.24726	0.21225	0.41769	0.47318	0.41319
	10	0.72081	0.7342	0.71013	0.18884	0.17595	0.16775	0.45483	0.45508	0.43894
	15	0.73673	0.7612	0.74957	0.17876	0.15817	0.17039	0.45774	0.45968	0.45998
im11	5	0.81299	0.80957	0.81947	0.26224	0.2417	0.28404	0.53761	0.52564	0.55176
	10	0.7639	0.71675	0.81997	0.21604	0.21489	0.19766	0.48997	0.46582	0.50881
	15	0.71661	0.70129	0.76688	0.18926	0.18858	0.19546	0.45294	0.44494	0.48117
im12	5	0.62327	0.60869	0.65953	0.30003	0.28248	0.3099	0.46165	0.44559	0.48472
	10	0.59818	0.61019	0.60936	0.29438	0.30219	0.29597	0.44628	0.45619	0.45266
	15	0.58069	0.5989	0.61901	0.28279	0.28926	0.29045	0.43174	0.44408	0.45473
im13	5	0.67685	0.48191	0.85556	0.2072	0.2721	0.28957	0.44202	0.37701	0.57256
	10	0.60229	0.45983	0.77105	0.18798	0.15846	0.21112	0.39514	0.30914	0.49109
	15	0.54753	0.39722	0.79778	0.18382	0.1472	0.16436	0.36568	0.27221	0.48107
im14	5	0.59126	0.60378	0.64232	0.27018	0.30155	0.35738	0.43072	0.45266	0.49985
	10	0.55452	0.59605	0.60992	0.26896	0.27527	0.28562	0.41174	0.43566	0.44777
	15	0.5295	0.54915	0.58903	0.25417	0.25863	0.28034	0.39184	0.40389	0.43468
im15	5	0.76662	0.78053	0.75471	0.19086	0.18978	0.23266	0.47874	0.48515	0.49368
	10	0.75626	0.74651	0.7824	0.15509	0.14158	0.14055	0.45568	0.44404	0.46148
	15	0.73075	0.73491	0.75604	0.13822	0.14172	0.14462	0.43448	0.43832	0.45033
im16	5	0.48883	0.74219	0.7324	0.27338	0.24894	0.25455	0.3811	0.49556	0.49348
	10	0.47058	0.70547	0.74625	0.24508	0.20552	0.20128	0.35783	0.45549	0.47377
	15	0.45273	0.70193	0.71787	0.23677	0.18838	0.18556	0.34475	0.44516	0.45171
im17	5	0.32573	0.26102	0.16885	0.55373	0.53427	0.46205	0.43973	0.39765	0.31545
	10	0.22117	0.28278	0.3055	0.52928	0.50341	0.50761	0.37523	0.39309	0.40655
	15	0.18239	0.20255	0.24135	0.47721	0.47282	0.50688	0.3298	0.33769	0.37411
im18	5	0.90336	0.89965	0.8954	0.27762	0.28309	0.29751	0.59049	0.59137	0.59646
	10	0.9299	0.93207	0.93179	0.24158	0.20153	0.20302	0.58574	0.5668	0.5674
	15	0.93215	0.93177	0.93988	0.21592	0.17817	0.17432	0.57404	0.55497	0.5571
im19	5	0.75248	0.77539	0.76444	0.21114	0.19247	0.20806	0.48181	0.48393	0.48625
	10	0.74078	0.7465	0.74635	0.17747	0.17588	0.17606	0.45913	0.4611	0.4612
	15	0.71225	0.71883	0.72034	0.17651	0.17276	0.17235	0.44438	0.4458	0.44635
im20	5	0.66045	0.82057	0.81633	0.13642	0.17697	0.22065	0.39844	0.49877	0.51849
	10	0.65531	0.81998	0.83873	0.12963	0.1253	0.1247	0.39247	0.47264	0.48172
	15	0.64226	0.81324	0.84145	0.11657	0.1093	0.096772	0.37941	0.46127	0.46911

Tabla 4.2: Resultados del criterio de evaluación de Levine y Nazif para el algoritmo original, el método propuesto y una segmentación utilizando histogramas. Los mejores valores para cada criterio (Uniformidad intra-región, disparidad inter-regional, y valor agregado) y para cada segmentación aparecen resaltados en gris.

imágenes, cosa por otra parte comprensible dada la variabilidad inherente del conjunto de prueba, todos los criterios usados en la evaluación se muestran en general favorables al método desarrollado. Esto es especialmente evidente en el caso tanto del criterio agregado como de la uniformidad inter-regional. Por lo tanto, podemos afirmar que el uso de un índice de rugosidad basado en superpíxeles no solo produce segmentos visualmente más homogéneos que el método original o que una segmentación basada en histogramas, sino que además mantiene mejor el equilibrio entre regiones visualmente uniformes y regiones visualmente dispares entre sí.

Si bien ni una mayor uniformidad intra-regional, ni una mayor disparidad inter-regional pueden asociarse directamente a una mejor anotación semántica, intuitivamente podemos ver como objetos más uniformes y distinguibles entre sí resultarán en vectores de características con una menor variabilidad intra-clase y más diferenciables entre sí. Una comparación con los resultados de la segmentación (10 segmentos) para algunas de las imágenes puede verse en Fig. 4.3. El color para los resultados proporcionados ha sido elegido para maximizar la legibilidad de las imágenes, usando (Harrower y Brewer, 2003), y no tiene ningún significado adicional.

Figura 4.3: Segmentación de imágenes a color. Ejemplo de resultados para 10 segmentos, para las imágenes Im1, Im5, Im11, Im14 e Im17; la primera columna muestra una segmentación mediante histogramas, la segunda el resultado del método original mediante índice de rugosidad y la tercera el producto de una segmentación mediante índice de rugosidad basado en superpíxeles.

Capítulo 5

Anotación semántica de imágenes médicas basadas en sobre-segmentación

El término "imagen médica" engloba el conjunto de técnicas y procesos utilizados para obtener imágenes del cuerpo humano o de partes del mismo, de cara a la realización de diferentes procedimientos médicos, incluyendo aquellos relacionados con la investigación médica (Flower, 2012). A día de hoy, las diferentes modalidades de imagen médica son cruciales en todos los niveles de atención facultativa, siendo parte integral de multitud de procedimientos diagnósticos. Aunque el juicio médico/clínico puede ser suficiente antes del tratamiento de muchas afecciones, el uso de servicios de diagnóstico por imágenes es fundamental para confirmar, evaluar y documentar correctamente el curso de de estas, así como para valorar la respuesta a los tratamiento.

Como campo de investigación científica, dependiendo del contexto, la imagen médica puede considerarse una sub-disciplina de la medicina en general, la ingeniería biomédica o la física médica. En un sentido general, se puede describir el funcionamiento conceptual de las técnicas de imagen médica como la medida de la interacción con el cuerpo de algún tipo de "radiación" con el que este ha sido escaneado (Louis, 1992). En qué consiste este escaneo, cuál es el resultado de esta interacción y cómo interpretarla, dependerá del tipo de imagen médica, donde se incluyen modalidades tan diferentes entre sí como son radiología, endoscopia, termografía, fotografía médica y microscopia...

La popularización e importancia del diagnóstico basado en imagen médica ha incrementado enormemente la investigación realizada en este campo y con ello el número de colecciones de imágenes médicas disponibles. Dado el volumen de datos disponible, el proceso de anotación manual de imágenes médicas es cada vez más costoso en términos de trabajo humano, estando ex-

puesto a errores y sesgos asociados a las limitaciones de la anotación manual. En consecuencia, las diferentes técnicas asociadas a la anotación semántica han ido adquiriendo una gran relevancia en el contexto de la imagen médica. En general, el rendimiento de los métodos automáticos para análisis de imagen médica no tiene por qué ser mejor que el de los médicos, es suficiente con que sea complementario a estos, como una forma de "segunda opinión" (Doi, 2007).

5.1. Detección de mitosis en imágenes histopatológicas de cáncer de mama.

Uno de los principales biomarcadores para el pronóstico de pacientes con cáncer de mama es la evaluación visual del tejido tumoral, el cual, en un examen microscópico, presenta diferencias sutiles comparado con el tejido normal. En estas evaluaciones, el recuento de mitosis¹ se muestra como un importante predictor de la agresividad y severidad tumoral (Elston y Ellis, 1991). Sin embargo, el gran tamaño de las imágenes histopatológicas y la variabilidad de aspecto de las ocurrencias mitóticas hacen del conteo manual de mitosis un proceso lento, tedioso y sujeto a fallos. La detección automática de núcleos mitóticos podría reducir el trabajo del patólogo y proporcionar resultados más consistentes, mejorando la calidad y rapidez del diagnóstico.

Tradicionalmente los métodos de detección automática de mitosis se basan en características definidas específicamente, con la idea de codificar propiedades relevantes de las apariciones mitóticas². Estos métodos incluyen y combinan varias características, calculadas para cada píxel o utilizando áreas cuadradas de tamaño fijo alrededor de estos. Se pueden ver ejemplos del uso de características morfológicas en (Sommer et al., 2012) y (Irshad, 2013), estadísticas en (Khan et al., 2012) y texturales en (Tek, 2013). Los métodos basados en características específicas, de la misma manera que la etiquetación manual de mitosis, sufren problemas relacionados con la variabilidad en textura y morfología asociada a las apariciones mitóticas y su similitud con otras estructuras presentes en la imagen.

Una segunda aproximación ha surgido con la popularización de las redes neuronales convolucionales (CNN). Existen múltiples enfoques para la detección de mitosis utilizando una CNN como clasificador de píxeles (Chen et al., 2016) o usando ventanas de tamaño fijo (Cruz-Roa et al., 2014). Si bien el uso de CNNs es capaz de encontrar patrones de características en las imágenes que son difíciles de describir usando clasificación mediante características, estos métodos dependen de la existencia de un conjunto de entrenamiento relevante y de gran tamaño. Por otro lado, el entrenamiento

¹El número de células en proceso de división

²Las células en proceso de división capturadas en un corte histológico

Figura 5.1: Esquema general del proceso de detección de mitosis propuesto.

y las pruebas usando CNNs son computacionalmente exigentes, requiriendo mucho tiempo.

En esta Tesis se aborda el problema de la detección de mitosis como un problema general de clasificación de segmentos en imágenes, donde se genera un modelo para etiquetar cada objeto resultante de una segmentación, en este caso mediante superpíxeles, como mitosis/no mitosis, a partir de un conjunto de características sencillas (momentos estadísticos y características de forma y textura)

5.1.1. Materiales y métodos

La solución propuesta al problema planteado sigue la estructura típica de los métodos de detección de mitosis, utilizando características específicas, la cual incluye las etapas de pre-procesamiento de imágenes, extracción de características y entrenamiento del clasificador. Se verá en las siguientes subsecciones como el método propuesto desarrolla estas etapas. Un esquema general de este método se muestra en Fig. 5.1.

5.1.1.1. Pre-procesamiento

El proceso de obtención de una imagen histopatológica suele requerir el tinto de la muestra de tejido, de cara a resaltar las estructuras presentes en este. En el caso de las imágenes de histopatología mamaria se utiliza comúnmente una combinación de dos tintes, hematoxilina (H) y eosina (E) (Veta et al., 2014). Esta tinción hace que las regiones nucleares y citoplasmáticas se muestren en tonos azules y púrpuras, mientras que el material extracelular tiende hacia gradaciones de rosa, dando a la imagen su aspecto característico. Sin embargo, las inconsistencias en la preparación de las tinciones histológicas (problemas de sobre-tintado, fallos en la fijación de las tintas y diferencias de color entre las imágenes, entre otros) pueden dificultar

Figura 5.2: Proceso de selección de superpíxeles candidatos. Detalle de la imagen normalizada original (izquierda) con la segmentación mediante superpíxeles obtenida a partir de su representación HE correspondiente superpuesta, imagen de representación/relación azul asociada (centro) y conjunto de candidatos a mitosis (derecha) destacados en color.

el análisis de estos tejidos.

Para normalizar el aspecto de las imágenes histopatológicas, estas son procesadas usando el método presentado en (Macenko et al., 2009), basado en el análisis de componentes principales. Este método da como resultado 3 imágenes diferentes, la imagen histopatológica normalizada (una imagen RGB) y dos imágenes representando los tintados individuales mediante hematoxilina y eosina. La combinación de estas dos bandas de color, generalmente conocida como imagen HE, es segmentada con una versión de SLIC específicamente modificada para trabajar con imágenes de 2 canales. Esta segmentación definirá el conjunto de candidatos a mitosis con el que se va a trabajar.

Es posible reducir el número de potenciales candidatos a mitosis aprovechando la apariencia de las figuras mitóticas como objetos hipercromáticos. Para acentuar las diferencias entre los posibles candidatos a la clasificación como mitosis y el resto, el fondo, se calcula la llamada representación/relación azul (Chang et al., 2012) a partir de las imágenes histopatológicas RGB normalizadas. En esta representación se codifica la relación entre el canal de color azul y los rojos y verde, dándose un valor más alto a los píxel con una intensidad azul elevada en relación a sus componentes rojo y verde. Para cada superpíxel generado en la imagen HE, se calcula su valor medio correspondiente en la imagen representación/relación azul, y aquellos en el percentil superior al 95 % son seleccionados como posibles candidatos mitóticos. En Fig. 5.2 se ilustran las etapas del proceso de selección de candidatos.

5.1.1.2. Extracción de características

A partir de la imagen HE, se plantea para cada superpíxel candidato un conjunto de características que serán utilizadas como un vector agregado con

Figura 5.3: Ejemplos de ocurrencias mitóticas. El superpíxel correspondiente al núcleo mitótico aparece marcado en tonos rojizos, los segmentos circundantes asociados están resaltados en tonos verdosos.

el que se caracterizará dicho superpíxel. Se utilizan tres tipos diferentes de características, cuya descripción general puede encontrarse en el capítulo 3:

- De la información de intensidad de los píxeles de cada uno de los canales de color, en este caso hematoxilina y eosina, para cada superpíxel candidato se extrae la media, la desviación estándar, la asimetría y la curtosis.
- A partir de la representación binaria del superpíxel, se calculan tres rasgos de forma: número de Euler, solidez y excentricidad.
- Utilizando la matriz de co-ocurrencia de grises asociada a cada superpíxel, para cada uno de los canales de color de la imagen (hematoxilina y eosina), se obtendrán cuatro rasgos texturales: el contraste, la homogeneidad, la entropía y la correlación.

Durante el proceso de división, en el entorno de una célula también se dan cambios específicos que pueden ser utilizados para la detección de ocurrencias mitóticas (Lipponen et al., 1994). Así, una célula mitótica además de los cambios en su núcleo, suele presentar rasgos como una membrana nuclear mal definida, protuberancias alrededor de los bordes y diferentes manchas en el citoplasma. Considerando el entorno de una célula el conjunto de todos los superpíxeles que la rodean, como se muestra en Fig. 5.3, el vector de características de cada candidato contendrá también la media de los rasgos de estos superpíxeles para cada candidato. Junto a las características del superpíxel candidato, esto produce un vector agregado final de 38 características diferentes.

5.1.1.3. Clasificación

Como método para generar el modelo de clasificación se ha seleccionado "Random Forests". Teniendo en cuenta que una mitosis no es un proceso común (se espera que una imagen histopatológica sana contenga una o ninguna mitosis), incluso habiendo realizado un proceso de selección previo existirá un notable desequilibrio entre las posibles clases (mitosis/no mitosis) en el

Figura 5.4: Algunos ejemplos de clasificación de mitosis. Verdaderos positivos contorneados en verde, falsos positivos contorneados en rojo y falsos negativos sin contorno.

conjunto de candidatos. Con el fin de reducir el sesgo de clasificación resultado de este desequilibrio en el momento de crear el modelo, en el conjunto de entrenamiento primero se reducirá el número de segmentos no mitóticos a la décima parte y luego se realizará una operación de sobre-muestreo sobre los candidatos mitóticos mediante SMOTE (Synthetic Minority Oversampling Technique) (Chawla et al., 2002). Debe tenerse en cuenta que el uso de SMOTE como técnica de sobre-muestreo representa un equilibrio entre la precisión y la sensibilidad, un modelo creado a partir de un conjunto de entrenamiento equilibrado mediante esta técnica da lugar a mas verdaderos positivos, pero también implica un aumento de los falsos positivos.

5.1.1.4. Conjunto de datos: TUPAC16

El método desarrollado ha sido entrenado y evaluado usando el conjunto de datos TUPAC16, publicado para el "MICCAI'16 Grand Challenge on Mitosis Assessment" (Veta et al., 2018). El conjunto de datos utilizado consiste en imágenes de 73 casos de cáncer de mama, de diferentes centros de médicos, 23 tomadas con un microscopio Aperio ScanScope XT, y 50 con un Leica SCN400 (aumento de $\times 40$, resolución espacial de $0,25 \mu m/\text{píxel}$ en ambos casos), correspondientes a un área real de $2 mm^2$ (5657×5657 píxeles). Las ubicaciones de las posibles mitosis han sido anotadas por dos patólogos de forma independiente. Las coincidencias entre los patólogos se consideradas una mitosis. Las discrepancias se presentan a un panel de dos expertos adicionales, quienes toman la decisión final.

5.1.2. Resultados

Como estrategia de evaluación se utilizará validación cruzada de 10 iteraciones sobre el conjunto de candidatos a mitosis. Qué se considera un verdadero positivo sigue las directrices utilizadas en el "MICCAI'16 Grand

Característica	Sensitividad	Especificidad	Precisión	F-score	CCM
Media	0,088	0,005	0,106	0,096	0,092
Asimetría	0,006	0,000	0,091	0,011	0,022
Curtosis	0,000	0,000	0,000	0,000	-0,002
Desviación est.	0,006	0,002	0,015	0,008	0,006
Color Agregado	0,235	0,003	0,317	0,270	0,269
Solidez	0,006	0,002	0,016	0,009	0,006
Excentricidad	0,024	0,014	0,010	0,014	0,006
N. de Euler	0,029	0,017	0,011	0,015	0,007
Forma Agregado	0,076	0,021	0,022	0,034	0,030
Correlación	0,000	0,000	0,000	0,000	0,000
Contraste	0,012	0,003	0,022	0,015	0,012
Entropía	0,029	0,002	0,104	0,046	0,052
Homogeneidad	0,029	0,004	0,039	0,034	0,029
Textura Agregado	0,071	0,004	0,091	0,079	0,075
Total Agregado	0,441	0,004	0,436	0,439	0,435

Tabla 5.1: Resultados medios para la evaluación del proceso de detección de mitosis. Se muestra la sensibilidad, la especificidad, la precisión, el "F-score" y el coeficiente de correlación de Matthews (CCM).

Challenge on Mitosis Assessment". Una detección se considerará un verdadero positivo sólo si su distancia a la ubicación de una mitosis anotada es inferior a $7,5 \mu m$ (30 píxeles). Si hay múltiples detecciones dentro de un radio de 30 píxeles de una mitosis anotada, se cuentan únicamente como un solo verdadero positivo. La segmentación se realiza con nuestra propia implementación de SLIC en MATLAB, la reducción del sesgo entre clases y la clasificación se realiza utilizando WEKA. Una descripción de las métricas de evaluación de rendimiento utilizadas puede verse en el (Powers, 2011).

Dependiendo de las características presentes en la región y de su complejidad, una segmentación mediante superpíxeles (para un intervalo de muestreo de 25 píxeles, elegido para aislar la mayor parte de las mitosis en un único superpíxel) podría producir aproximadamente 80.000 segmentos. Tras el proceso de selección de los posibles candidatos, una histopatología típica resultará en un número de entre 300 y 400 candidatos a mitosis. Esto implica, dependiendo de la imagen, en un sesgo entre clases de entre 100 y 200 no mitosis por cada mitosis. Cabe señalar que un pequeño porcentaje de mitosis (globalmente, entre el 5 % y el 8 % en nuestras pruebas), generalmente mal definidas o en sus últimas etapas, podría perderse en el proceso de selección de candidatos.

La Tabla 5.1 muestra el resultado medio de las 10 iteraciones del proceso de validación, tanto para el vector agregado como para las características in-

dividuales y las características agrupadas por tipo. Se presentan únicamente las métricas de rendimiento asociadas a las ocurrencias mitóticas positivas y el coeficiente de correlación de Matthews; debido al desequilibrio entre clases de los conjuntos de pruebas, incluir los valores asociados a los candidatos no mitóticos, (correctamente clasificados en un 99.9 %) daría una imagen sesgada del rendimiento del método. Los resultados incluyen el efecto de la pérdida de mitosis en el proceso de selección de candidatos.

Es posible ver cómo, individualmente, ninguna de las características utilizadas da lugar a un modelo de clasificación razonable, mejorando los resultados con el uso de diferentes vectores agregados. Si bien no es posible hacer una comparación directa con otros métodos dada las diferencias en el conjunto de pruebas y en las estrategias de validación, se puede ver como la caracterización de superpíxeles, a pesar de la sencillez de su planteamiento, produce resultados ("F-score" 0,44) en consonancia con otros métodos de detección de mitosis (MICCAI Grand Challenge).

5.2. Detección de espinas dendríticas en microscopía confocal.

La microscopia confocal es una forma especializada de microscopia de fluorescencia de campo amplio (Pawley, 2010), donde se genera un volumen tridimensional seleccionando diferentes planos del corte histológico mediante la supresión activa de la señal proveniente de los planos desenfocados. Uno de los múltiples usos de esta técnica de microscopia es la obtención de imágenes volumétricas de neuronas sin necesidad de destruir el tejido donde se encuentran, con suficiente resolución como para distinguir pequeñas estructuras, como las espinas dendríticas, como puede verse en Fig. 5.5.

Las espinas dendríticas son pequeñas protuberancias que emergen de las dendritas. Estas estructuras definen la interconectividad de la red neuronal. En neurobiología, la identificación y reconstrucción de espinas dendríticas a partir de grandes conjuntos de imágenes microscópicas es una herramienta importante para el estudio de las funciones neuronales y sus propiedades biofísicas. Muchas funciones neuronales, como el aprendizaje y la memoria, están estrechamente relacionadas con la aparición o desaparición de esas conexiones neuronales (Alvarez y Sabatini, 2007). Por lo tanto, el estudio de las espinas dendríticas es un importante área de investigación en neurobiología.

Por lo general las espinas dendríticas son etiquetadas manualmente, lo que consume mucho tiempo y es susceptible al sesgo inherente del operador humano, incluso contando con la ayuda de computadoras. Para hacer frente a estos problemas, varios métodos se han propuesto para detectar y medir espinas dendríticas automáticamente, sin, o con interacción humana limitada.

Figura 5.5: "Slice"(izquierda) y reconstrucción volumétrica (derecha) de una neurona captada mediante microscopía confocal.

Las técnicas de tintado fluorescente utilizadas comúnmente en microscopía confocal no permiten visualizar las estructuras internas de la neurona (Belichenko y Dahlström, 1995). De esta manera, las espinas dendríticas vienen definidas por su posición y su forma respecto al tallo neuronal, no siendo evidente ninguna característica visual que nos permita su clasificación utilizando diferencias de intensidad o análisis textural. Se plantea por tanto la detección de espinas dendríticas como un problema de clasificación mediante forma y/o entorno. Los trabajos existentes sobre la detección de espinas dendríticas suelen dividirse en dos grupos: métodos basados en clasificación directa y métodos basados en la extracción del tallo o línea central.

Los métodos orientados a la clasificación separan los elementos de una imagen en diferentes categorías utilizando modelos basados en el la forma de la dendrita y sus características de vecindad. Por ejemplo, un clasificador muy sencillo, mediante la agrupación de vóxeles a partir de la distancia a la superficie dendrítica es propuesto en (Rodríguez et al., 2008). Un algoritmo más complejo es utilizado en (Li et al., 2009), basado en un clasificador usando la curvatura de la superficie, la distancia al área central de la dendrita y la relación entre las normales de la superficie y el área central en puntos específicos de la neurona.

Los métodos basados en la extracción de líneas centrales reducen los objetos presentes en el volumen confocal a una representación basada en grafos y detectan en estos el/los posibles tallos de la dendrita, las líneas centrales. Las espinas dendríticas aparecerán como pequeñas protuberancias asociadas a estas líneas centrales. Existen diferentes enfoques para la extracción del tallo dendrítico, generalmente basados en algoritmos genéricos de esqueletonización. En (Zhou et al., 2009) se propone un modelo de ajuste binario local para generar un conjunto de niveles de intensidad volumétricos desde donde

generar la línea central. El trabajo presentado en (Xu et al., 2006) adapta el conocido método de esqueletización de "fuego de pasto" (grassfire) (Leymarie y Levine, 1992) para calcular la región central del tallo dendrítico. En (Janoos et al., 2009) se propone la creación de una estructura curvilínea utilizando una aproximación de la función media geodésica, definida por la superficie de la dendrita. En (Zhang et al., 2007), la esqueletonización se realiza también utilizando un detector de estructuras curvilíneas, y las espinas dendríticas se detectan a través del algoritmo de "marcha rápida" (Sethian, 1996), tomando las regiones centrales de las espinas como puntos iniciales. También relacionados con los enfoques basados en la extracción de línea central de dendrita, algunos métodos proponen el uso de modelos para representar el volumen de la neurona. Por ejemplo, en ((Jaeger et al., 2010) y (He et al., 2003)) se emplean modelos paramétricos basado en estructuras cilíndricas como una aproximación de la topología neuronal.

Este caso de uso desarrolla dos métodos automatizados para detectar y extraer espinas dendríticas independientemente de su tamaño o morfología (Jones y Powell, 1969), utilizando una segmentación mediante supervóxeles, basados en la estructura física de las dendritas neuronales.

5.2.1. Materiales y métodos

Es posible dividir los métodos desarrollados en tres fases: segmentación del volumen confocal, clasificación de los segmentos, y agrupación de los segmentos clasificados en espinas dendríticas individuales. Como primer paso, común a ambos métodos propuestos, se separarán las dendritas del fondo en el volumen confocal, utilizando un valor umbral predefinido para seleccionar los supervóxeles correspondientes a la dendrita. La segmentación del volumen confocal se realiza utilizando una versión de SLIC modificada para trabajar específicamente con volúmenes de intensidad.

La principal diferencia entre los dos métodos desarrollados se encuentra en el proceso de clasificación de los segmentos: el primero de ellos utiliza la regularidad local en el espacio definido por la segmentación. El segundo se basa en la construcción de un grafo representativo de la dendrita.

5.2.1.1. Detección por regularidad en el vecindario local

Una dendrita puede ser vista como una estructura cilíndrica que presenta pequeñas protuberancias, las espinas dendríticas. Como se mencionó en el capítulo 3, el tamaño de los superpíxeles/supervóxeles generados por el SLIC es uniforme, lo que resulta en que, excluyendo el fondo del volumen confocal, la mayoría de supervóxeles se generen dentro del tallo dendrítico, distribuidos más o menos regularmente. Sin embargo, en las espinas dendríticas se generará un menor número de supervóxeles, distribuidos entre la cabeza y el cuello de la espina dendrítica, quedando la distribución de estos marcada por

la geometría general del tallo dendrítico. Un ejemplo de esta segmentación sobre un corte se puede ver en Fig. 5.6

Inspirados en los trabajos para segmentación de mitocondrias en (Lucchi et al., 2012), aprovechamos la distribución de supervóxeles generados por SLIC. La vecindad de los supervóxeles situados dentro de un tallo dendrítico será, hasta cierto punto, regular. Por otra parte, la vecindad de un supervóxel situado dentro de una espina dendrítica vendrá definida por la presencia del fondo circundante y el cuello de dicha espina dendrítica. De esta manera, podemos afirmar que:

$$\left\| \sum C_s - n(C_s) \right\| < \left\| \sum C_d - n(C_d) \right\| \quad (5.1)$$

donde $n(x) = \{C_1, C_2, C_3, \dots, C_n\}$ es el vecindario del punto x , C_s es un supervóxel parte de la espina dendrítica y C_d es un supervóxel del tallo dendrítico. Esta discrepancia puede utilizarse para detectar las espinas dendríticas, como se puede ver en Fig. 5.7.

Este método utiliza únicamente información local del vecindario para generar un modelo con el que detectar espinas dendríticas, por lo que puede definirse como un método orientado a la clasificación.

5.2.1.2. Detección por distancia al tallo dendrítico

El uso de una segmentación mediante supervóxeles permite plantear un segundo método, uno basado en la extracción de la línea central a partir de un grafo describiendo la estructura de la dendrita. Es posible simplificar este

Figura 5.6: Segmentación supervóxel de una pila de imágenes dendrítica (slice, detalle). Los colores carecen de significado especial.

Figura 5.7: Vector de suma en el método de regularidad del vecindario local (proyección máxima). Supervóxeles azulados indican un vector más corto, rojizo significa vectores más largos.

grafo para dejar sólo una representación del tallo dendrítico, y detectar las posibles dendritas a partir de la distancia esta representación.

Para generar el grafo base que representa la dendrita, se unen los puntos centrales de supervóxeles contiguos, y se etiqueta cada vértice con el número de vóxeles que compone la frontera entre estos. A partir de este grafo se calcula el árbol de recubrimiento mínimo (un ejemplo del árbol resultante puede verse en Fig. 5.8). Puesto que el tronco dendrítico representa una estructura lineal mucho más grande que las espinas dendríticas, el "tronco" del árbol de recubrimiento mínimo estará formado principalmente por los nodos definidos por los supervóxeles presentes en el tallo dendrítico.

Para aislar este "tronco", se utiliza un algoritmo basado en un modelo de optimización mediante colonia de hormigas (Dorigo et al., 1996). Para cada nodo del grafo, se calcula la ruta más larga posible dentro del árbol de recubrimiento mínimo. La estructura de árbol del grafo asegura que estas rutas atravesarán el "tronco" del árbol. Los nodos situados dentro del "tronco" del árbol estarán presentes en muchas más rutas que los que están dentro de las ramas, como se puede ver en la Fig. 5.9. Este algoritmo puede encontrar la representación del tallo dendrítico en el grafo a pesar de la posible existencia de bucles visuales, bifurcaciones dentro de la dendrita o la presencia de múltiples dendritas aisladas en el volumen confocal. La distancia del centro de masa de los supervóxeles a los vértices del grafo que representa el tallo se utilizará para clasificar supervóxeles, dado que esta distancia será mayor en los supervóxeles generados en las espinas dendríticas que en la de aquellos que forman parte del tallo dendrítico. El resultado de este método se presenta en Fig. 5.10.

Figura 5.8: Árbol de mínima expansión con el gráfico final que representa el gráfico de vástago superpuesto. La escala está en vóxeles.

Figura 5.9: Supervóxels en dendritas etiquetados con el número de trayectos más largos en un árbol de extensión mínima del que forma parte. Un tono azulado indica un supervóxel con menor número de caminos, un tono rojizo uno con más caminos.

Figura 5.10: Distancias de supervóxel al tronco dendrítico en el método del gráfico de distancia del vástago dendrítico (proyección máxima). Un supervóxel azulado indica un elemento más cercano, rojizo significa más lejano.

5.2.1.3. Agrupación de supervóxels

Puesto que el uso de supervóxels da como resultado una sobre-segmentación de la imagen y por tanto de la dendrita, el último paso en nuestros clasificadores es agrupar los supervóxels clasificados como espinas dendríticas. La agrupación se realiza utilizando el concepto de máximas esferas inscritas, de forma similar a (Silin y Patzek, 2006). Desde el centro de cada supervóxel candidato, se calcula el volumen esférico más grande que no incluya el centro de un supervóxel no candidato. Los supervóxels se agregarán teniendo en cuenta los solapamientos entre las esferas calculadas. Cada uno de estos agregados se etiqueta como una espina dendrítica diferente, como se muestra en la Fig. 5.11.

5.2.1.4. Conjunto de prueba

Como conjunto de datos de prueba, se utilizan cuatro volúmenes diferentes de microscopía confocal (Leica SP 5 con un objetivo de inmersión en glicerol de 63x) codificados como pilas de imágenes de 16 bits, cada imagen representando una porción del volumen, con una resolución de $0,13mm \times 0,13mm \times 0,4mm$. Aunque puede haber artefactos resultado de la presencia de ruido o efectos de fotoblanqueo, no se ha aplicado ningún tipo de pre-procesamiento en el volumen.

Figura 5.11: Imagen dendrítica con el lomo dendrítico agrupado resultados del método de regularidad del vecindario local superpuesto. Proyección máxima (izquierda) y reconstrucción de volumen (derecha). Los colores no tienen un significado especial.

5.2.2. Resultados

Si bien no se dispone de una clasificación base con la que comparar los resultados, la evaluación de estos por parte de expertos ha sido notablemente positiva para ambos métodos. La detección mediante regularidad del vecindario local es efectiva detectando pequeñas espinas dendríticas, incluidas espinas alineadas con el eje z . Sin embargo, las espinas dendríticas más grandes podrían generar una distribución de supervóxeles similar a la distribución en el tallo dendrítico. En este caso algunos supervóxeles podrían estar mal clasificados, como se ve en la Fig. 5.12. También puede haber falsos positivos en áreas cuya topología no es la habitual en una espina dendrítica, como los extremos, o en menor medida, en bifurcaciones.

Por otra parte, el uso de la detección mediante distancia al grafo del tallo dendrítico permite etiquetar fácilmente espinas dendríticas grandes y medianas. Sin embargo, como se menciona en (Janoos et al., 2009) los métodos basados en la extracción de la línea central pueden ser muy sensibles a pequeños cambios en la forma del objeto. El uso de supervóxel minimiza este problema, ya que la esqueletonización no depende de los vóxeles sino de un objeto de orden superior, como es un supervóxel, pero aun así es posible perder características cercanas al tallo dendrítico, especialmente en áreas estrechas o pobremente definidas dentro del volumen dendrítico. Algunos ejemplos de estos problemas se muestran en la Fig. 5.13.

5.3. Clasificación de volúmenes MRI para detección de Alzheimer

La enfermedad de Alzheimer (Alzheimer's disease, AD) es la forma predominante de demencia en ancianos, y se espera que el número de pacientes se multiplique en los próximos años (Ferri et al., 2005). La enfermedad de Alzheimer, en sus primeras etapas, implica alteraciones a pequeña escala en el cerebro definidas por la presencia de ovillos neurofibrilares y depósitos de placa beta amiloide ($A\beta$) (Braak y Braak, 1997). Estas alteraciones resultan principalmente en daños a las sinapsis, seguido de degeneración en los axones neuronales y, en última instancia, degeneración del árbol dendrítico y el perikaryon, conduciendo a atrofia en regiones específicas del cerebro (Serrano-Pozo et al., 2011), especialmente relevantes en áreas como los hipocampos derecho e izquierdo, los giros temporales, el giro cingulado y los precúneos (Ishii et al., 2005).

Cada vez resulta más evidente que, cuando a un paciente se le diagnostica la enfermedad de Alzheimer, la atrofia ya está bien establecida en el cerebro. La presencia clínica más temprana de los síntomas que pueden eventualmente progresar a un diagnóstico clínico de la enfermedad de Alzheimer se clasifican generalmente como deterioro cognitivo amnésico leve (MCI, Mild amnesic Cognitive Impairment) (Dubois y Albert, 2004). El MCI se considera un estado de deterioro clínico, generalmente pérdida de memoria, más

Figura 5.12: Proyección máxima de los resultados de la detección de la columna dendrítica agrupada (arriba) y de la detección mediante la distancia (abajo) del gráfico del vástago dendrítico. Las diferencias entre los métodos se pueden ver en los resultados de las áreas mal definidas (1 y 2) y la columna dendrítica cerca del tallo dendrítico (3, 4, 5 y 6) aparecen resaltados.

Figura 5.13: Muestra de diferentes resultados de detección de espinas dendrítica agrupada (Detalles). Las detecciones de espinas dendríticas parciales por los resultados del método de regularidad del vecindario local (fila superior) se emparejan con los resultados de la detección por el método de distancia del gráfico del tallo dendrítico (fila inferior).

allá de lo que puede esperarse para la edad y educación del sujeto sin llegar a cumplir con los criterios de clasificación de la demencia. Aunque no todos los pacientes con MCI desarrollarán AD, ya es posible detectar atrofia cerebral en esa etapa (Tabatabaei-Jafari et al., 2015). En los individuos en las primeras fases de la enfermedad de Alzheimer, los volúmenes entorrinales ya se han reducido en un 20 %-30 % y los volúmenes del hipocampo en un 15 %-25 %. Las estimación de la progresión de la atrofia en los casos diagnosticados de enfermedad de Alzheimer, entre el 0,8 % y el 2 % anual, sugieren que el proceso de atrofia asociado a la enfermedad en áreas como el lóbulo temporal medio debe haber estado activo durante varios años antes del diagnóstico o incluso antes de la presencia de síntomas.

El proceso de degeneración puede ser visualizado a través de varias modalidades de imagen médica, las cuales han demostrado ser un valioso biomarcador de la etapa y la agresividad potencial del aspecto neurodegenerativo de la patología (Jack Jr et al., 2010). Como resultado, en las últimas dos décadas el uso de diferentes modalidades de imagen médica ha pasado de ser un mero instrumento secundario en el diagnóstico de la enfermedad de Alzheimer a ser una de las principales herramientas. No solo eso, los avances en neuroimagen han proporcionado poderosas herramientas para el estudio de enfermedades relacionadas con el cerebro. Así, la resonancia magnética funcional en estado de reposo (fMRI) (Vemuri et al., 2012), la visualización de la conectividad cerebral funcional intrínseca, o el uso de radiotrazadores TEP desarrollados para permitir la visualización *in vivo* de agregaciones tau y placas $A\beta$ (Wong et al., 2010), (Clark et al., 2012), se han convertido en importantes biomarcadores en la investigación de la enfermedad de Alzhei-

mer. Sin embargo, a pesar de estos avances, la tomografía por emisión de positrones (PET) y la resonancia magnética (MRI) siguen siendo las principales herramientas en los estudios relacionados con la AD, especialmente en esta última, dada su disponibilidad, naturaleza no invasiva y la relativa ausencia de molestias para el paciente. Sin embargo, los cambios tempranos en la enfermedad de Alzheimer son sutiles y es difícil distinguir los patrones mediante la evaluación radiológica convencional. Por lo tanto, sigue siendo difícil establecer biomarcadores fiables para el diagnóstico y el seguimiento de la progresión de la enfermedad, especialmente en las primeras fases. Esto ha llevado al desarrollo de numerosos métodos automáticos para la evaluación de la atrofia cerebral a partir de imágenes MRI.

Existe una gran cantidad de investigación relacionada con la detección de la AD en MRIs asistida por ordenador. Teniendo en cuenta el trabajo presentado en (Rathore et al., 2017), estos métodos pueden agruparse en tres categorías, dependiendo del tipo de características que se utilicen para evaluar las variaciones estructurales y cómo son extraídas. Así, existen métodos basados en mapas de densidad, ya sea utilizando todo el cerebro como unidad o basándose en una parcelación; métodos derivados del estudio de la superficie cortical, también de forma global o local; y métodos basados únicamente en el examen de regiones específicas del cerebro.

Los métodos de mapeado de densidad buscan patrones de atrofia a partir de los volúmenes de materia blanca (WM), materia gris (GM) y líquido cefalorraquídeo (CSF) generados por los métodos de morfometría basada en vóxeles (VBM) (Ashburner y Friston, 2000). Si bien es posible crear modelos directamente utilizando estos volúmenes como un vector de características, como puede verse en (Magnin et al., 2009) y (Klöppel et al., 2008) es más común, teniendo en cuenta la alta dimensionalidad del espacio de características que estos volúmenes representan, recurrir a diferentes métodos de reducción de características, supervisados o no supervisados. En (Salvatore et al., 2015) las dimensiones de los mapas de densidad WM y GM se reducen mediante el análisis de componentes principales (PCA) y el resultado se utiliza para entrenar a un clasificador basado en SVM. En (Beheshti et al., 2015) la reducción de las dimensiones de los mapas de GM se realiza utilizando como características únicamente la distribución de la intensidad de los vóxeles. Otra forma de selección de características se basa en el uso de un atlas de partición cerebral para obtener mediciones regionales de las características anatómicas para cuestionar la presencia de áreas de tejido anormal como en (Tzourio-Mazoyer et al., 2002) o (Shattuck et al., 2008).

Los métodos derivados de la superficie cortical utilizan cambios sutiles extraídos a nivel de vértice sobre la superficie del cerebro, representados principalmente como medidas de espesor cortical. Al igual que en los métodos de mapeado de densidad, estas mediciones pueden utilizarse directamente (Li et al., 2014) o procesarse de cara a su reducción de dimensionalidad. Algunos

ejemplos de esta última categoría se encuentran en el trabajo presentado en (Park et al., 2012) donde la superficie cortical se modela utilizando mallas tridimensionales y el espesor cortical se extrae mediante la parametrización de estas mallas o el método presentado en (Cho et al., 2012) donde los datos de espesor de la superficie cortical se convierten al dominio de frecuencia y la dimensionalidad es reducida mediante el filtrado de componentes de alta frecuencia.

Los métodos basados en el análisis de regiones específicas del cerebro se centran únicamente en las áreas donde se concentra la degeneración neuronal producto de la AD. Esto implica habitualmente el uso de biomarcadores extraídos del hipocampo (Sørensen et al., 2016), (Jaba et al., 2017)) como volumen, forma o rasgos texturales. Otras áreas del cerebro también se han utilizado en (Tang et al., 2015), donde se realizan estudios morfológicos utilizando varias regiones, incluyendo el hipocampo derecho e izquierdo, el tálamo y los ventrículos laterales, como biomarcadores para la predicción de AD.

5.3.1. Materiales y métodos

Relacionado con los métodos derivados de los mapas de densidad, en este trabajo se realiza la adaptación de la versión del concepto de histón presentada en el capítulo 4 de esta Tesis a imágenes volumétricas a color. Esto permite su uso como elemento textural en la clasificación de MRIs, con el fin de diferenciar a pacientes con enfermedad de Alzheimer de pacientes no afectados. El uso de histones, como parte del método de segmentación de Mushrif y Ray (Mushrif y Ray, 2008) ha sido utilizado en algunas modalidades de imagen médica (Senthilkumaran y Rajesh, 2009), incluyendo MRI (Xie et al., 2015). Sin embargo, el potencial de un histón, en cierto modo un histograma extendido con información contextual de cada píxel/vóxel, como vector de características ha permanecido esencialmente inexplorado hasta ahora.

En comparación con otros métodos basados en mapas de densidad, los métodos texturales pueden identificar patrones de intensidad y relaciones ocultas al ojo humano sin ayuda (Kassner y Thornhill, 2010). El trabajo realizado presenta un biomarcador sencillo a partir del cerebro completo, basado en la relación entre las probabilidades del tejido cerebral de corresponder a materia gris, materia blanca o líquido cefalorraquídeo, en lugar de una medición única de la intensidad del vóxel como en (Maani et al., 2015) o (Martínez-Murcia et al., 2017). La partición del volumen MRI en estos tres volúmenes de probabilidad permite que cada uno de ellos se equipare a un canal de color y por tanto realizar la caracterización del volumen de MRI dentro del espacio de color RGB.

El proceso de clasificación presentado se divide en tres etapas, como se puede ver en Fig. 5.14. En la primera etapa, pre-procesamiento, se lleva a

Figura 5.14: Esquema del método propuesto para la clasificación de volúmenes MRI.

cabo la corrección de posibles diferencias en intensidad, una normalización espacial sobre el conjunto de datos completo, y se segmentan los volúmenes en materia blanca, materia gris y líquido cefalorraquídeo. En la segunda etapa, se calcula un conjunto de histones basados en una segmentación mediante supervóxeles a partir del volumen agregado de GM, WM y CSF, que se utilizarán como vector de características. Sobre este vector de características se realiza una reducción de dimensionalidad mediante la selección de componentes principales. El vector de características resultante se utiliza para entrenar un clasificador basado en SVM en la tercera etapa.

5.3.1.1. Software Clínica

Para trabajar con un flujo de trabajo de pre-procesamiento estandarizado, compatible con múltiples bases de datos de neuroimagen, el pre-procesamiento del volumen y la gestión general del conjunto de datos se lleva a cabo utilizando el software Clínica. Clínica es una plataforma de software para estudios de investigación en neurociencia, desarrollada por el Laborato-

rio ARAMIS³ en el ICM⁴ de París. El proceso general de pre-procesamiento de un MRI-T1 implica segmentación de tejido, corrección de sesgo y normalización espacial. Clinica proporciona una interfaz para usar de manera sencilla el procedimiento planteado en SPM (Statistical Parametric Mapping) (Henson y Friston, 2006), que realiza todos estos procesos simultáneamente, en un procedimiento conocido como "segmentación unificada" (Ashburner y Friston, 2005).

SPM modela el cerebro como una capa de líquido cefalorraquídeo que rodea las materias blanca y gris. La probabilidad *a priori* de que un vóxel contenga materia gris o blanca puede determinarse utilizando un atlas probabilístico de tipos de tejidos. Este método plantea las intensidades del volumen como una mezcla de k funciones gaussianas, cada una de ellas modelada por su media, varianza y una proporción conocida de mezcla de tejidos, esta última generada a partir de múltiples atlas. La regla de Bayes se utiliza para producir la probabilidad *a posteriori* de cada clase de tejido. Usando este enfoque, dos vóxeles con intensidades idénticas pueden ser identificados como tejido diferente. Un ejemplo de los resultados obtenidos se puede ver en la Fig. ??.

Como paso final, el proceso calcula una plantilla de deformación para el conjunto de imágenes pre-procesado, aplicando el algoritmo de difeomorfismo de DARTEL a la imagen MRI⁵. DARTEL (Diffeomorphic Anatomical Registration Through Exponentiated Lie. Algebra(Goto et al., 2013)) plantea la normalización espacial de los volúmenes MRI a partir de un "campo de flujos" bidireccional.

DARTEL comienza tomando el resultado producido por una segmentación GM/WM/CSF y alineándolo lo más precisamente posible a un conjunto de mapas de probabilidad tisular, utilizando únicamente mediante transformaciones rígidas. A partir de la media de todas las imágenes, se crea una plantilla inicial que se utiliza para el registro simultáneo de tejidos entre imágenes. El modelo creado permite calcular las deformaciones específicas de cada una de las imágenes individuales. Finalmente, se vuelve a generar una plantilla, promediando la inversa de las deformaciones aplicadas. Este proceso se repite varias veces, hasta que la nueva plantilla no produce cambios relevantes al ser aplicada a las imágenes.

Por último, Clinica ofrece un módulo de aprendizaje automático que permite realizar clasificaciones mediante la combinación de diferentes entradas, algoritmos y estrategias de validación. Este módulo depende de la librería scikit-learn (Pedregosa et al., 2011) de python.

³www.aramislab.fr

⁴www.icm-institute.org

⁵En anatomía computacional, un sistema difeomorfo es un sistema diseñado para asignar distancias métricas en el espacio de imágenes anatómicas, con el fin de permitir la cuantificación y comparación de cambios morfométricos en las estructuras presentes

Figura 5.15: Cortes de volumen, seleccionados del conjunto de datos OASIS-1 preprocesado, mostrando ejemplos de segmentación de tejidos. Cortes originales del MRI (izquierda), probabilidades de materia gris (centro izquierda), probabilidades de materia blanca (centro derecha), líquido cefalorraquídeo (derecha).

5.3.1.2. Extracción de características: histones basados en super-vóxeles para la detección de Alzheimer en imágenes estructurales de MRI

Como puede verse en (Callaert et al., 2014), en el proceso de segmentación unificada existe una relación entre la degeneración neuronal y la presencia de vóxeles con una probabilidad relativamente baja de que forme parte de un tipo específico de tejido. Dado que la enfermedad de Alzheimer tiende a manifestarse como atrofia en zonas específicas del cerebro, en estos casos el proceso de segmentación tisular tenderá a mostrar zonas "ambiguas", difíciles de clasificar como uno u otro tipo de tejido. En este trabajo, el uso de histones

Figura 5.16: Corte de volumen del conjunto de datos OASIS-1 pre-procesado, mostrando las probabilidades de segmentación del tejido y su corte agregado correspondiente como una imagen RGB. Probabilidades de materia gris (izquierda), probabilidades de materia blanca (centro izquierda), líquido cefalorraquídeo (centro derecha) y agregado (derecha) (probabilidades de materia gris como azul, probabilidades de materia blanca como verde, líquido cefalorraquídeo como rojo).

como característica textural explota tanto la disminución del volumen total de GM y WM como los cambios en la relación entre GM, WM y CSF.

Es posible equiparar los diferentes volúmenes de probabilidad (representados como mapas de intensidad de 8 bits) a los canales de un volumen agregado a color (véase Fig. 5.16). Utilizando una versión de SLIC modificada para trabajar con volúmenes de color, se llevará a cabo la sobre-segmentación mediante supervóxeles de dicho volumen agregado. De forma similar al desarrollo mostrado en el capítulo 4 de esta Tesis, el uso de supervóxeles como vecindario para el cálculo de histones garantiza una relación espacial directa entre un vóxel y su vecindario, respetando las áreas y fronteras entre las áreas presentes en la imagen. Puesto que la segmentación utilizando supervóxeles ya produce áreas localmente homogéneas, es posible cuantificar la homogeneidad local utilizando la desviación estándar media de la intensidad del conjunto de supervóxeles, permitiendo definir el "expanse" para la generación de histones de forma dinámica.

Para calcular un histón, H_{SV_i} , se considera que un vóxel está dentro de la esfera de color si la distancia entre la intensidad media de su correspondiente supervóxel y la intensidad de ese vóxel es menor que la desviación local media del volumen de probabilidad en el espacio definido por los supervóxeles, en cada uno de los volúmenes de probabilidad. Sea $Cs = \{Cs_1, Cs_2, \dots, Cs_{Ns}\}$ el conjunto de supervóxeles resultantes en una segmentación de imagen donde Ns es el número total de supervóxeles; $I(x_{Cs_j}, y_{Cs_j}, z_{Cs_j}, v_i)$ representa el centroide del supervóxel Cs_j en el volumen de probabilidad v_i en V_p . V_p es el conjunto de volúmenes de probabilidad $V_p = \{WM, GM, CLF\}$; y Np_j es el número de vóxeles en ese supervóxel. La función de similitud $SV(x, y, z)$

se define como:

$$SV(x, y, z) = \begin{cases} 1 & \text{if } \forall s_i \in V_p \quad d_{ts_i}(x, y, z) < E_{v_i} \\ 0 & \text{otherwise} \end{cases} \quad (5.2)$$

$$d_{ts_i}(x, y, z) = |I(x, y, z, v_i) - I(x_{Cs_j}, y_{Cs_j}, v_i)| \quad I(x, y, z, v_i) \in Cs_j \quad (5.3)$$

$$E_{v_i} = \frac{1}{N_s} \sum_{j=1}^{N_s} \frac{1}{Np_j} \sum_{I \in Cs_j} I(x, y, z, v_i) \quad (5.4)$$

El vector de características obtenido está formado por 768 componentes (un histón diferente para cada volumen de probabilidad, cada uno con 256 niveles de intensidad).

Trabajar con vectores de características muy grandes hace que un clasificador sea propenso tener problemas de sobreajuste. Para disminuir el tamaño del vector de características representado por los histones, el método desarrollado recurre al análisis de componentes principales (PCA). PCA (Wold et al., 1987) transforma un conjunto de variables en una combinación lineal de estas, llamadas componentes principales (PCA), sin pérdida de información. Para un conjunto de datos estandarizado, los componentes principales pueden calcularse como los vectores propios normalizados de la matriz de covarianza de las variables originales y pueden ordenarse por la cantidad de varianza que cada componente principal explica. Desde un punto de vista geométrico, cada componente puede verse como la dirección de maximización de la varianza de las muestras, sin relación con componentes anteriores, cuando se proyectan sobre el propio componente. El número de componentes extraídos es igual al número de variables que se analizan, de cara a crear un modelo de clasificación se selecciona un subconjunto. Por lo general, sólo los primeros componentes representan cantidades significativas de varianza, correspondiendo el resto a cantidades triviales de esta.

5.3.1.3. Conjunto de datos: OASIS

Los sujetos incluidos en este estudio han sido obtenidos de la "Open Access Series of Imaging Studies" (OASIS) (Marcus et al., 2007). OASIS es un proyecto orientado a proporcionar libre acceso a conjuntos de volúmenes MRI cerebrales a la comunidad científica. OASIS está disponible a través del Centro de Investigación de la Enfermedad de Alzheimer de la Universidad de Washington, el Dr. Randy Buckner del Instituto Médico Howard Hughes (HHMI) de la Universidad de Harvard, el Grupo de Investigación en Neuroinformática (NRG) de la Facultad de Medicina de la Universidad de Washington y la Red de Investigación en Informática Biomédica (BIRN).

		No-demencia					Demencia				
Grupo de edad	N.	n.	media	M/F	CDR	MMSE	n	media	M/F	CDR	MMSE
<20	19	19	18.53	10/9			0	0	0	0/0/0	
20s	119	119	22.82	51/68			0	0	0	0/0/0	
30s	16	16	33.38	11/5			0	0	0	0/0/0	
40s	31	31	45.58	10/21		29.5±1.5	0	0	0	0/0/0	
50s	33	33	54.36	11/22		29.4±0.8	0	0	0	0/0/0	
60s	40	25	64.88	7/18		29.1±1.3	15	66.13	6/9	12/3/0	22.9±5.5
70s	83	35	73.37	10/25		29.1±1	48	74.42	20/19	32/15/1	24.6±4
80s	62	30	84.07	8/22		28.7±1.2	32	82.88	13/19	22/9/1	24.5±3.9
90s	13	8	91.00	1/7		28.3±1.6	5	92.00	2/3	4/1/0	23.8±1.9
	416	316		119/197		29.1±1.1	100		41/59	70/28/2	24.3±4.16

Tabla 5.2: Características demográficas de la población estudiada (base de datos OASIS-1). Los valores de la escala de clasificación de demencia clínica (CDR) se indican como el número de pacientes diagnosticados como demencia muy leve/demencia leve/demencia moderada. Los valores del Mini-Mental State Examination (MMSE) se indican como media± desviación estándar.

Para este estudio ha utilizado específicamente la "Cross-sectional MRI Data in Young, Middle-Aged, Non-demented and Demented Older Adults" de OASIS (OASIS-1). Esta colección es un conjunto de estudios transversales compuesta por 416 individuos de ambos sexos, todos diestros, con edades comprendidas entre los 18 y los 96 años. El conjunto incluye 100 pacientes (mayores de 60 años) con un diagnóstico clínico de enfermedad de Alzheimer que va de muy leve a moderado. Un resumen de las características demográficas de la OASIS-1 se muestra en la Tabla 5.2.

5.3.2. Resultados

Se han utilizado dos subconjuntos de datos, generados a partir de OASIS-1. Las diferencias en las características demográficas y clínicas entre los gru-

pos de pacientes (AD y CN) se evalúan utilizando el Test-T para la edad y el MMSE, y la prueba χ^2 de Pearson para el género. El nivel de significación se ha fijado en 0,05. El primer subconjunto incluye a todos los sujetos de 60 años o más. En este subconjunto se encontraron diferencias significativas entre los pacientes de control y los pacientes con enfermedad de Alzheimer en tanto en MMSE como en género. Las diferencias de género entre grupos, así como la gran variabilidad general en la edad del conjunto de datos, son factores que pueden dar lugar a un sesgo en los resultados de la clasificación. Teniendo esto en cuenta, se ha utilizado un segundo conjunto de datos reducido, sus sujetos seleccionados aleatoriamente bajo el criterio de minimizar las diferencias de género entre grupos y, en la medida de lo posible, descartar los valores atípicos para disminuir la desviación estándar de la edad. En la Tabla 5.3 se detalla un resumen de las características demográficas para la población de ambos conjuntos. La segmentación y extracción de características se lleva a cabo con nuestra propia implementación de SLIC en MATLAB, la reducción de dimensionalidad y la clasificación de sujetos se lleva a cabo utilizando R. Una descripción de las métricas de evaluación del rendimiento utilizadas puede verse en (Powers, 2011).

Para llevar a cabo la creación de un modelo a partir de los volúmenes MRI, con el fin de separar a los pacientes con enfermedad de Alzheimer de los pacientes no afectados, se ha seleccionado un SVM para generar el modelo, dado que generalmente produce buenos resultados a pesar de posibles problemas de sesgo o varianza en los vectores de características, como puede verse en (Meyer et al., 2003).

Diagnóstico	Número	Edad	Género	Media MMSE	CDR
Enfermedad de Alzheimer	100	76.6 $\pm$ 7.77	41/59	24.3 $\pm$ 4.16	70/28/2
Cognitivo normal	98	75.9 $\pm$ 8.07	26/72	29.93 $\pm$ 3.8	

Diagnóstico	Número	Edad	Género	Media MMSE	CDR
Enfermedad de Alzheimer	70	75 $\pm$ 5.5	23/47	24.8 $\pm$ 4.1	49/20/1
Cognitivo normal	69	75 $\pm$ 4.47	25/44	29.04 $\pm$ 3.77	

Tabla 5.3: Características demográficas del subconjunto de datos de mayores de 60 años (tabla superior) y del subconjunto de datos reducido (tabla inferior). El género se indica como masculino/femenino. Los valores de la escala de clasificación de demencia clínica (CDR) se indican como el número de pacientes diagnosticados como demencia muy leve/demencia leve/demencia moderada. Los valores del Mini-Mental State Examination (MMSE) se indican como media $\pm$ desviación estándar.

Con el fin de obtener estimaciones imparciales para los resultados, se utiliza una estrategia de "holdout", siguiendo las recomendaciones presentadas en (Varoquaux et al., 2017). Cada conjunto de datos se divide aleatoriamente diez veces en dos grupos: un conjunto de entrenamiento y un conjunto de pruebas (80 %-20 %). El proceso de división utilizado preserva la distribución de edad y género entre conjunto de pruebas y entrenamiento. Cada conjunto de entrenamiento se utiliza para entrenar un clasificador. Sus correspondientes conjuntos de pruebas se utilizan con fines de evaluación. Los conjuntos de entrenamiento del primer subconjunto (todos los mayores de 60) también se utilizan para determinar el "kernel" adecuado para el clasificador SVM. La información demográfica individual para cada división se presenta en 5.3.3. A efectos de comparación, los resultados proporcionados por el módulo de clasificación de Clínica se usan como referencia (Samper-Gonzalez et al., 2017). Los mapas de probabilidad de materia gris modulados por DARTEL del conjunto de entrenamiento son utilizados para calcular un "kernel" lineal a partir de la matriz de Gram de sus vectores de característica (todos los vóxeles en el volumen). Este "kernel" es la entrada para un SVM genérico cuyo parámetro de coste se optimiza para conseguir el modelo con la mejor precisión equilibrada posible, mediante una búsqueda exhaustiva en cuadrícula. Este proceso se repite tanto en los conjuntos de datos de mayores de 60 años como en el reducido, para cada división.

5.3.2.1. Parámetros para SVM y selección de "kernel"

Para decidir el tamaño del subconjunto de componentes principales a usar, se utiliza el método conocido como "método del codo". El "método del codo" (Cattell, 1966) representa la covarianza que explica cada valor propio frente al ordinal del componente principal. Este método implica un análisis visual de dicha representación, donde se busca una "flexión" o "codo" entre los componentes con valores propios relativamente grandes y aquellos con valores propios pequeños. Se asume que más allá de este "codo" la varianza explicada no aumentará significativamente con la adición de más vectores propios, por lo que utilizar estos en una clasificación aportará poco o nada al modelo. La Fig. 5.17 muestra el gráfico obtenido, trazando la varianza explicada en términos de los primeros 30 componentes principales del conjunto de datos analizado. Como se puede ver, el "codo" se da entre los componentes principales 5º y 6º, por lo tanto los primeros 5 componentes parecen ser suficientes para describir la varianza presente en los datos.

Dado que no se conocen las características específicas del conjunto de datos procesado y, por lo tanto, no sabemos cuál es el "kernel" más adecuado, se realizarán pruebas con "kernels" genéricos lineales, polinómicos y radiales. A cada uno de los subconjuntos de entrenamiento del conjunto de datos de sujetos de 60 años o más le se aplicará una metodología de validación cruzada con 10 iteraciones, repitiendo el experimento de plegado 10 veces para cada

Figura 5.17: Gráfica de codo representando los valores propios de los 30 primeros componentes.

"kernel". Para evitar posibles problemas de sobreajuste de los "kernels" no se lleva a cabo ningún paso de optimización en los modelos presentados. Se utilizan los parámetros por defecto para el SVM en R (1 para el parámetro de coste, $1/(\text{datos dimensión})$ para el parámetro gamma y 3 para el parámetro grado). De los resultados presentados en la Tabla 5.4, seleccionaremos un "kernel" lineal para realizar el resto de los experimentos, ya que representa una mejora en todas las medidas de rendimiento evaluadas.

5.3.2.2. Clasificación

La tabla 5.5 muestra la media de los resultados obtenidos con el conjunto de datos de sujetos de 60 años o más (tabla superior) así como con el subconjunto reducido (tabla inferior) utilizando tanto la referencia obtenida con Clínica como el método desarrollado basados en histones. Los resultados de la evaluación individual para cada iteración del "holdout" pueden encontrarse en 5.3.4.

El uso de histones como características representa una mejora en todas las métricas de evaluación propuestas (prueba de McNemar $p < 0,05$ para todas las pruebas, excepto la 5ª iteración). Como puede apreciarse tanto en la matriz de confusión (ver 5.5, fila superior) como en los valores de predicción negativos y positivos (NPV, PPV), los errores de clasificación en el modelo generado con Clínica tienen un ligero sesgo hacia los falsos

	Preci.	Preci. B.	NPV	PPV	Sensit.	Especif.	F-score
histón: SVM kernel lineal	0.6918	0.6913	0.7473	0.6562	0.8012	0.5813	0.7214
histón: SVM kernel polinómico	0.5995	0.5989	0.6181	0.5897	0.6760	0.5217	0.6266
histón: SVM kernel radial	0.6503	0.65004	0.7034	0.6187	0.7738	0.5262	0.6874

Tabla 5.4: Resultados de la evaluación que muestran una comparación entre las métricas de evaluación medias propuestas a partir del uso de diferentes "kernels" (lineal, polinomial y radial) para el clasificador SVM para el conjunto de datos de 60 años y más, para las 10 repeticiones. Los mejores resultados aparecen destacados en negrita.

	Preci.	Preci. E.	NPV	PPV	Sensit.	Especif.	F-score
Clinica:Vóxeles como características	0.6643	0.6730	0.6415	0.7045	0.615	0.7241	0.6481
histón: SVM kernel lineal	0.8071	0.8001	0.8226	0.7838	0.8541	0.7620	0.8158

	Preci.	Preci. B.	NPV	PPV	Sensit.	Especif.	F-score
Clinica:Vóxeles como características	0.7089	0.7234	0.7088	0.7168	0.6845	0.7355	0.6925
histón: SVM kernel lineal	0.8366	0.8334	0.8637	0.8242	0.8783	0.7885	0.8475

Tabla 5.5: Resultados de la evaluación que muestran una comparación entre la media de las métricas de evaluación de la línea de base de Clínica y el método propuesto (SVM entrenado usando un "kernel" lineal) para el conjunto de datos de los sujetos de 60 años o más (tabla superior) y para el conjunto de datos reducido (tabla inferior).

negativos, mientras que, para el método propuesto, los errores se asocian principalmente con los falsos positivos. Cabe señalar las notables diferencias entre los resultados presentados en la Tabla 5.4 y la Tabla 5.5 para el conjunto de datos de sujetos de 60 años o más. Es posible achacar estas diferencias a las diferentes estrategias de evaluación entre los dos casos. La validación cruzada no sólo tiene un sesgo pesimista en sus resultados (Varoquaux et al., 2017), sino que las poblaciones en las iteraciones de una validación cruzada son completamente aleatorios, sin respetar la distribución de edad o género del conjunto de datos original.

Como era de esperar, utilizando el conjunto de datos reducido sin el sesgo impuesto por diferencias significativas de género entre los grupos de AD y CN y una menor variabilidad en la edad, los resultados mejoran tanto para el modelo base generado con Clínica como para el método propuesto, como puede verse en la Tabla 5.5(Tabla inferior). En este caso, el método propuesto proporciona también mejores resultados para todas las métricas de

		Valor predicho				Valor predicho	
		AD	CN			AD	CN
Valor actual	AD	62.7	30.5	Valor actual	AD	83.7	23.9
	CN	37.2	69.4		CN	16.2	76.1

		Valor predicho				Valor predicho	
		AD	CN			AD	CN
Valor actual	AD	66.2	25.8	Valor actual	AD	87.9	21.6
	CN	33.3	71.1		CN	12.1	78.3

Tabla 5.6: Matriz de confusión media para las pruebas de validación, que muestra la tasa de clasificación en porcentaje para el conjunto de datos de los sujetos de 60 años o más (tabla superior) y para el conjunto de datos reducido (tabla inferior). Las tablas de la izquierda corresponden a la línea de base de Clínica, las de la derecha corresponden al método propuesto ("kernel" lineal).

evaluación en comparación con los resultados con el conjunto de datos de 60 años o más. La distribución en los errores no difiere del escenario anterior (ver Tabla 5.6, fila inferior). En este caso, no podemos afirmar que los resultados de la prueba de McNemar sean significativos, ya que esta prueba se considera inadecuada para tamaños de muestrales pequeños (Jonsson, 1993).

Podemos afirmar por tanto que los resultados experimentales, tanto para mayores de 60 años como en el subconjunto reducido, han demostrado una mejora significativa en el desempeño para la clasificación de sujetos con AD en relación a los CN, en comparación con el uso de vóxeles como vector de características (modelo generado por Clínica). Aunque dadas las diferencias de edad, género, progreso de la enfermedad y/o calidad de imagen entre poblaciones de estudio es imposible hacer una comparación directa, en general podemos afirmar que los resultados obtenidos son comparables o mejores que los de métodos texturales similares (Maani et al., 2015), (Termenon y Graña, 2012) o (Martínez-Murcia et al., 2017). Por otro lado, la actual implementación del método, en el que los histones se calculan en todo el cerebro, no se centra en los patrones espaciales específicos de degeneración cerebral asociados con la enfermedad de Alzheimer. Esto puede limitar la capacidad del método para discriminar en presencia de atrofia cerebral asociada a otras patologías, o incluso en pacientes muy ancianos.

Finalmente mencionar como el uso de un flujo de trabajo estandarizado, proporcionado por Clínica, representa un paso importante hacia la reproducibilidad de este trabajo y su comparabilidad con desarrollos futuros. Sin embargo, debe tenerse en cuenta que este paso de pre-procesamiento puede ser costoso desde el punto de vista computacional, especialmente si se consideran conjuntos de datos de gran tamaño.

5.3.3. Información demográfica de las diferentes divisiones usados en el proceso de validación

En esta subsección se presenta la información demográfica para cada división utilizada en la validación, tanto para el conjunto de datos de mayores de 60 años, como para el subconjunto de datos reducido.

		Conjunto de Entrenamiento						Conjunto de prueba					
Diag.		N.	Edad	Género	MMSE	CDR	N.	Edad	Género	MMSE	CDR		
División 1	AD	76	76.4±6.89	29/47	24.43±4.45	56/18/2	24	77.7±7.84	12/12	23.3±3.14	14/10/0		
	CN	80	75.2±8.62	23/57	29.02±1.17		18	79.11±10.09	3/15	28.66±1.32			
División 2	AD	74	76.6±6.85	31/43	24.5±3.98	55/18/1	26	77.1±7.96	10/16	23.76±4.67	15/10/1		
	CN	82	76.59±8.48	21/61	28.92±1.17		16	72.43±10.87	5/11	29.06±1.38			
División 3	AD	80	76.42±7.52	33/47	24.4±4.19	58/21/1	20	78.1±5.14	8/12	23.85±4.14	12/7/1		
	CN	70	75.46±8.91	20/56	28.97±1.25		22	77.5±9.25	6/16	28.9±1.06			
División 4	AD	79	76.3±6.59	36/43	24.64±4.07	56/21/2	21	77.7±8.94	5/16	23.09±4.35	14/7/0		
	CN	77	75.88±9.36	19/58	23.09±1.11		21	76.047±7.04	7/14	28.55±1.46			
División 5	AD	78	77.28±7.46	30/48	23.93±4.28	52/24/2	22	74.9±5.5	11/11	25.68±3.46	18/4/0		
	CN	78	76.14±8.82	21/57	28.91±1.23		20	75.05±9.76	5/15	29.15±1.8			
División 6	AD	84	76.73±6.83	33/51	24.57±4.02	62/21/1	16	76.87±8.73	8/8	23±4.77	8/7/1		
	CN	72	75.77±9.42	21/51	29.02±1.11		26	74.3±7.8	5/21	28.76±1.45			
División 7	AD	80	77.11±7.25	35/45	24.52±4.1	57/23/0	20	75.35±6.53	6/14	23.5±4.39	13/5/2		
	CN	76	76.22±9.03	19/57	28.93±1.21		22	74.86±8.92	7/15	29.04±1.21			
División 8	AD	77	76.55±7.31	31/46	23.92±4.39	51/24/2	23	77.43±6.51	10/13	25.65±2.99	19/4/0		
	CN	79	75.6±9.28	21/58	29±1.23		19	77.21±7.72	5/14	28.78±1.13			
División 9	AD	76	76.94±7.52	34/42	24.94±3.77	57/18/1	24	76.1±5.75	7/17	22.33±4.77	13/10/1		
	CN	80	75.9±8.72	21/59	28.96±1.26		18	79.94±10.33	5/13	28.94±0.93			
División 10	AD	81	76.7±7.34	31/50	24.25±4.17	56/25/0	19	77±6.23	10/9	24.57±3.14	14/3/2		
	CN	75	76.12±9.17	19/56	29±1.2		23	75.26±8.49	7/16	28.82±1.23			

Tabla 5.7: Información demográfica para cada división del conjunto de datos de 60 años o más. El sexo se indica como masculino/femenino, los valores de la Escala de Calificación de Demencia Clínica (CDR, por sus siglas en inglés) se indican como el número de pacientes diagnosticados como demencia muy leve/ demencia leve/ demencia moderada. Los valores de Edad y Mini-Mental State Examination (MMSE) se indican como la media±desviación estándar.

		Conjunto de Entrenamiento					Conjunto de pruebas				
	Diag.	N.	Edad	Género	MMSE	CDR	N.	Edad	Género	MMSE	CDR
División 1	AD	54	75.88±4.87	20/24	24.37±3.79	38/16/0	16	75.87±4.55	8/8	25.5±5.03	11/4/1
	CN	55	75.1±6.80	22/33	28.9±1.33		14	73.92±5.26	3/11	29.13±0.66	
División 2	AD	56	75.85±4.73	21/35	24.92±3.83	40/16/0	14	76±5.11	7/7	23.42±5	9/4/1
	CN	53	74.73±6.15	21/32	28.96±1.25		16	75.31±7.76	4/12	28.93±1.18	
División 3	AD	53	75.77±5.04	21/32	24.11±4.13	35/17/1	17	75.16±6.36	7/10	26.23±3.61	14/3/0
	CN	56	75.16±6.36	22/34	28.89±1.26		13	73.61±7.19	3/10	29.23±1.09	
División 4	AD	56	75.82±5	22/34	24.5±4.24	38/17/1	14	76.14±3.86	6/8	25.14±3.54	11/3/0
	CN	53	75.5±6.33	20/33	29.05±1.19		16	72.75±6.8	5/11	28.62±1.31	
División 5	AD	53	75.9±4.68	20/33	24.33±3.68	36/16/1	17	75.58±5.17	8/9	25.52±5.22	13/4/0
	CN	56	75.39±5.17	21/35	28.92±1.26		13	72.61±5.82	4/9	29.07±1.11	
División 6	AD	52	75.67±5.03	20/32	24.11±4.07	35/16/1	18	76.5±3.98	8/10	26.11±3.9	14/4/0
	CN	57	74.94±6.43	20/37	28.87±1.29		12	74.5±7.11	5/7	29.33±0.77	
División 7	AD	54	76.07±5	21/33	25.31±4.02	41/12/1	16	75.25±3.99	7/9	22.31±3.57	8/8/0
	CN	55	74.76±6.5	19/36	28.98±1.19		14	75.28±6.74	6/8	28.85±1.4	
División 8	AD	55	75.87±4.58	24/31	24.96±3.76	40/15/1	15	76.93±5.59	4/11	23.4±5.12	9/5/0
	CN	54	74.64±5.59	19/35	28.88±1.25		15	74.64±6.54	6/9	29.2±1.14	
División 9	AD	54	76.18±4.98	24/30	24.7±3.97	37/16/1	16	74.87±3.96	4/12	24.37±4.61	12/4/0
	CN	55	74.3±6.14	19/36	28.98±1.2		14	73.14±7.76	6/8	28.65±1.35	
División 10	AD	60	75.76±4.5	27/33	24.81±4.05	42/18/0	10	76.6±6.34	1/9	23.5±4.4	6/3/1
	CN	49	74.26±6.35	15/34	29.04±1.11		20	76.35±6.8	10/10	28.75±1.48	

Tabla 5.8: Información demográfica para cada división de prueba del conjunto de datos reducido. El sexo se indica como masculino/femenino, los valores de la Escala de Calificación de Demencia Clínica (CDR) se indican como el número de pacientes diagnosticados como demencia muy leve/demencia leve/demencia moderada. Los valores del Mini-Mental State Examination (MMSE) se indican como la media de la desviación estándar de $\pm$.

5.3.4. Resultados de las evaluaciones individuales para cada división de datos

En esta subsección se presenta los resultados individuales para cada una de las diferentes divisiones usadas en el proceso de validación, tanto para el conjunto de datos de mayores de 60 años, como para el subconjunto de datos reducido.

	Prec.	Prec. B.	NPV	PPV	Sensit.	Espec.	F- score
División 1	0.7	0.7098	0.6315	0.8181	0.5625	0.8571	0.6667
División 2	0.63333	0.6205	0.6190	0.6666	0.4285	0.8125	0.5217
División 3	0.7	0.8597	0.625	0.7857	0.6470	0.7692	0.7097
División 4	0.7333	0.74107	0.8333	0.6666	0.8571	0.625	0.75
División 5	0.8	0.7964	0.8235	0.7692	0.7692	0.8235	0.7692
División 6	0.7667	0.7579	0.8235	0.6923	0.75	0.7778	0.72
División 7	0.7	0.6964	0.6923	0.70588	0.75	0.6428	0.7273
División 8	0.6897	0.6827	0.6154	0.75	0.7059	0.6667	0.7273
División 9	0.7	0.7	0.8	0.6	0.75	0.6667	0.6667
División 10	0.6666	0.6696	0.625	0.7142	0.625	0.7142	0.6667

	Prec.	Prec. B.	NPV	PPV	Sensit.	Espec.	F- score
División 1	0.7857	0.7777	0.7647	0.8	0.8333	0.7222	0.8163
División 2	0.8095	0.81	0.7222	0.875	0.8076	0.8125	0.84
División 3	0.8333	0.8409	1	0.7407	1	0.6818	0.8510
División 4	0.7857	0.7857	0.8	0.7727	0.8095	0.7619	0.7906
División 5	0.7857	0.7886	0.7391	0.8421	0.7272	0.85	0.7804
División 6	0.8095	0.8341	0.95	0.6818	0.9375	0.7307	0.7894
División 7	0.8095	0.8113	0.85	0.7727	0.85	0.7727	0.8095
División 8	0.7857	0.7814	0.7778	0.7916	0.826	0.7368	0.8085
División 9	0.7857	0.7708	0.8	0.7778	0.875	0.6667	0.8235
División 10	0.7857	0.7708	0.8	0.7778	0.875	0.6667	0.8235

Tabla 5.9: Resultados de la evaluación que muestran una comparación entre las métricas de evaluación para cada división de prueba utilizando la línea de base de Clínica (tabla superior) y el método propuesto (SVM entrenado usando un "kernel" lineal) para el conjunto de datos de sujetos de 60 años o más (tabla inferior).

	Prec.	Prec. B.	NPV	PPV	Sensit.	Espec.	F- score
División 1	0.619	0.6082	0.5625	0.6538	0.7083	0.5	0.68
División 2	0.6667	0.7321	0.5357	0.9286	0.5	0.9375	0.65
División 3	0.7143	0.7159	0.75	0.6818	0.75	0.6818	0.7143
División 4	0.6905	0.7019	0.6538	0.75	0.5714	0.8095	0.6486
División 5	0.7381	0.7426	0.6957	0.7895	0.6818	0.8	0.7317
División 6	0.5476	0.5366	0.6522	0.4211	0.5	0.5769	0.4571
División 7	0.6905	0.6905	0.7143	0.6667	0.7	0.6818	0.6829
División 8	0.6429	0.6528	0.5833	0.7222	0.5652	0.7368	0.6341
División 9	0.6429	0.6624	0.56	0.7647	0.5417	0.7778	0.6341
División 10	0.6905	0.6875	0.7083	0.6667	0.6316	0.7391	0.6486

	Prec.	Prec. B.	NPV	PPV	Sensit.	Espec.	F- score
División 1	0.8666	0.8660	0.8571	0.875	0.875	0.8571	0.875
División 2	0.8333	0.8303	0.8235	0.8461	0.7857	0.875	0.8148
División 3	0.8333	0.8348	0.7857	0.875	0.8235	0.8461	0.8484
División 4	0.8666	0.875	1	0.7777	1	0.75	0.875
División 5	0.8	0.7873	0.8181	0.7894	0.8823	0.6923	0.8333
División 6	0.8333	0.8333	0.7692	0.8823	0.8333	0.8333	0.8571
División 7	0.8333	0.8258	0.9090	0.7894	0.9375	0.7142	0.8571
División 8	0.8667	0.8611	0.8889	0.8333	0.8333	0.8889	0.8333
División 9	0.8	0.7991	0.7857	0.8125	0.8125	0.7857	0.8125
División 10	0.8812	0.8718	0.921	0.8238	0.875	0.7727	0.8485

Tabla 5.10: Resultados de la evaluación que muestran una comparación entre las métricas de evaluación para cada división de la prueba, utilizando la línea de base de Clinica (tabla superior) y el método propuesto (SVM entrenado usando un "kernel" lineal) para el conjunto de datos reducido (tabla inferior).

Capítulo 6

Anotación semántica de imagen de teledetección basada en sobre-segmentación

El uso de imagen de teledetección (en el contexto de esta Tesis) tiene como objetivo la captura, tratamiento y análisis de imágenes de la superficie terrestre (Schowengerdt, 2006). Una imagen de teledetección es una representación de la energía reflejada o irradiada por el suelo, medida ya sea a partir de sensores situados en satélites aviones o drones. Los datos obtenidos por teledetección se utilizan para clasificar y cartografiar grandes áreas, en sustitución de métodos tradicionales de cartografía, que requieren estudios de campo complejos y costosos. Así, podemos encontrar este tipo de imágenes involucradas en una gran variedad de aplicaciones; el estudio, desarrollo y planificación urbana en general (Weng et al., 2018) y de infraestructuras en particular (Ge et al., 2018); el estudio del uso de los suelos (Novara et al., 2018); estudios ambientales (Wu et al., 2014), aprovechamiento de recursos naturales (Ibarrola-Ulzurrun et al., 2018); gestión y monitorización de recursos agrícolas y minerales (Moran et al., 1997); gestión y monitorización de desastres naturales (Klomp, 2016) etc... Al igual que sucede con la imagen médica, el volumen de datos adquiridos mediante la teledetección ha ido creciendo gradualmente con el avance del tiempo y la mejora de las técnicas de teledetección, haciendo que su análisis manual se haya convertido en un procedimiento complejo y engorroso.

Es posible clasificar las diferentes aplicaciones de las imágenes de teledetección en dos tipos diferentes (Mather y Tso, 2016). Por un lado, es posible encontrar aplicaciones donde se busca la medición de una propiedad física específica de la superficie, como una manera de obtener información de sus características intrínsecas o el estado de esta, como por ejemplo al realizar mediciones de la temperatura del mar. Otro de los posibles objetivos para el uso de imagen de teledetección es la identificación de la naturaleza de las

áreas superficiales; es decir, el paso de los posibles valores simbólicos presentes en una imagen (textura moteada, área verde) a conceptos semánticos (campo de cultivo, área urbana). El proceso de asignar a cada posible región en la imagen una etiqueta de un conjunto de clases predefinidas, expresadas típicamente en términos de cobertura del suelo en lo que se conoce como mapa temático, siendo la generación de un mapa temático uno de los casos mas típicos de anotación semántica (Bratanu et al., 2011).

6.1. Segmentación de imagen multispectral para teledetección.

La segmentación de imágenes es un paso esencial en la comprensión de imágenes de alto nivel en el ámbito de la teledetección, formando parte integral de un gran número de aplicaciones como el reconocimiento y seguimiento de objetos (Friedman y Russell, 1997), la clasificación de cobertura y uso de la tierra (Fröhlich et al., 2013). La segmentación de imágenes, y específicamente la segmentación de imágenes multispectrales, es un problema complejo donde, a pesar de la gran cantidad de trabajos desarrollados sobre el tema, no puede considerarse que exista realmente una solución válida para todas las clases de imágenes (Pal y Pal, 1993).

Los algoritmos para segmentación de imágenes multispectrales suelen estar basados en métodos de crecimiento de regiones (Beucher y Meyer, 1992) y de transformación de divisoria (Li y Xiao, 2007) (Karantzas y Argialas, 2006) o (Chen et al., 2006). Técnicas basadas en agrupamiento se han aplicado con éxito a la segmentación de imágenes multispectrales por satélite, como k-means (Koonsanit et al., 2012), "fuzzy" c-means (Yugander et al., 2012) o agrupamiento basado en densidad (Tran et al., 2003).

También se ha abordado la idea de combinar información espectral y espacial. Este enfoque generalmente funciona mejor que los métodos limitados a un dominio de datos único. Por ejemplo, en (Pachk et al., 2003) se desarrollan dos clasificadores estadísticos, uno basado en información espectral y otro en información espacial y se combinan sus resultados. En (Haralick y Dinstein, 1975) se propuso un método bautizado como "agrupación espacial", una combinación de crecimiento regional y análisis de histogramas. En (Matas y Kittler, 1995) una combinación de crecimiento de regiones y análisis local de características mediante histogramas se usa para extraer posibles regiones de una imagen multispectral. El concepto de relajación probabilística (Rosenfeld et al., 1976) como método de segmentación multispectral, se ha propuesto tanto en (Eklundh et al., 1980) como en (Hsiao y Sawchuk, 1989). En (Banerjee et al., 2014) una combinación de agrupación basada en árboles de extensión mínima y análisis de densidad muestra resultados prometedores.

En el capítulo 4 se mostraron las posibilidades del uso de un índice de

rugosidad basado en superpíxeles para la segmentación de imagen a color. En esta sección se presentan los resultados de la extensión de este método a imágenes multiespectrales de teledetección como forma de combinar información espectral y espacial. Aunque la implementación actual utiliza cuatro bandas espectrales, puede adaptarse fácilmente al número de bandas espectrales presentes en la imagen.

6.1.1. Materiales y Métodos

Al igual que el método original de segmentación (Mushrif y Ray, 2008), el proceso desarrollado se divide en tres etapas: el cálculo del índice de rugosidad a partir de una sobre-segmentación utilizando SLIC; el cálculo de los umbrales para cada componente espectral utilizando los picos y valles significativos del índice de rugosidad, de donde se obtendrá una segmentación inicial de la imagen, y la fusión de regiones. La adaptación del proceso a imágenes multiespectrales implica cambios en la forma de calcular el índice de rugosidad y en como se realiza el proceso de elección de picos y valles, usando dicho índice de rugosidad. En este trabajo, cada uno de los componentes de la segmentación se denomina agrupación espectral y los resultados antes y después de la fusión de regiones se conocen como agrupaciones espectrales base y agrupaciones espectrales finales, respectivamente.

Uno de los problemas que se presenta a la hora de realizar esta extensión es la fusión de regiones utilizando el concepto de distancia de color. La distancia de color, como se plantea en (Vergés-Llahí y Sanfeliu, 2005), no es un concepto trivial y resulta difícil de definir adecuadamente en el ámbito multiespectral. Por tanto, y aprovechando las características intrínsecas de una segmentación de superpíxeles, también se han desarrollado dos métodos de fusión de regiones alternativos diferentes (uno realizando la fusión a nivel de píxeles y otro a nivel de superpíxeles).

6.1.1.1. Índice de rugosidad basado en superpíxeles en imágenes multiespectrales de satélite

Se ha realizado la adaptación del método planteado en el capítulo 4 al espacio definido por el número de bandas espectrales presentes en la imagen. Se denominará al conjunto de los superpíxeles resultado de la sobre-segmentación de la imagen como $Cs = \{Cs_1, Cs_2, \dots, Cs_{Ns}\}$ donde Ns es el número total de superpíxeles, $I(x_{Cs_j}, y_{Cs_j}, s_i)$ representa el centroide de un superpíxel, Cs_j es la banda espectral $s_i \in S_p$ y Np_j es el número de píxeles en dicho superpíxel; la función de similitud $SM(x, y)$ queda definida como:

$$SM(x, y) = \begin{cases} 1 & \text{if } \forall s_i \in S_p \quad d_{ts_i}(x, y) < E_{s_i} \\ 0 & \text{en otro caso} \end{cases} \quad (6.1)$$

$$d_{ts_i}(x, y) = |I(x, y, s_i) - I(x_{Cs_j}, y_{Cs_j}, s_i)| \quad I(x, y, s_i) \in Cs_j \quad (6.2)$$

$$E_{s_i} = \frac{1}{Ns} \sum_{j=1}^{Ns} \frac{1}{Np_j} \sum_{I \in Cs_j} I(x, y, s_i) \quad (6.3)$$

La relación entre el histograma y el histon calculado a partir la función de similaridad anterior dará el índice de rugosidad asociado a la imagen multiespectral.

6.1.1.2. Detección de picos y valles en índices de rugosidad de imagen multiespectral

En el momento de seleccionar picos significativos, la imagen multiespectral presenta un desafío específico; es normal que, para una banda espectral, la mayoría de los valores de intensidad aparezcan concentrados en un rango muy estrecho, haciendo que esta banda aparezca sub-representada en la segmentación final en el caso de utiliza valores fijos para definir que es un pico significativo, como en el método original de Mushrif y Ray. Este caso puede detectarse comprobando la distribución de los valores en el histograma, utilizando los cuantiles correspondientes. De esta forma, se sustituye la definición de pico original por una dinámica, donde cada histograma se comprueba y se considera "apretado" si al menos el 90 % de sus valores de intensidad se agrupan en un rango inferior al 50 % de los valores de intensidad posibles. En este caso, la distancia mínima posible entre picos se reduce linealmente desde el valor original de $10 * (max(I(s_i) - min(I(s_i)))$ a un mínimo de $2 * (max(I(s_i) - min(I(s_i)))$, cuando el 90 % de los valores se agrupan en menos del 10 % del intervalo total.

6.1.1.3. Fundamentos de fusión de regiones

Los métodos basados en índice de rugosidad tratan cada banda espectral individualmente y tienden a resaltar pequeñas diferencias regionales de homogeneidad, generando más agrupaciones espectrales base que los métodos basados en histogramas. El método original utiliza el algoritmo propuesto en (Cheng et al., 2002b) para la fusión de regiones. Este algoritmo es en esencia un algoritmo de fusión basado en la distancia de color, diseñado para fotografía y su extensión a un dominio multiespectral no garantiza resultados adecuados.

Para evitar depender de una distancia de color, se proponen dos métodos de fusión de agrupaciones espectrales diferentes, uno basado en píxeles, que se denominará "co-ocurrencia en superpíxeles", y otro basado en segmentos (en este caso superpíxel), denominado "caracterización de superpíxeles".

Ambos comparten la misma estructura subyacente; se define una distancia de similitud entre las agrupaciones espectrales obtenidos utilizando su relación con la sobre-segmentación de la imagen. Esta distancia se utiliza para construir una matriz de distancia entre pares de elementos (las agrupaciones espectrales de base o los superpíxeles). A partir de esta matriz se puede crear un árbol binario de agrupamientos jerárquicos que muestre las relaciones entre elementos, utilizando el algoritmo UPGMA (Unweighed Pair Group Method with Arithmetic Mean) (Legendre y Legendre, 1998). El árbol resultante es una jerarquía de varios niveles, en la que los agrupamientos de un nivel se unen en un agrupamiento del siguiente nivel. Esta estructura permite decidir la escala de agrupación (en nuestro caso, el número final de agrupamiento) más apropiada para una aplicación.

6.1.1.4. Fusión de regiones basada en co-ocurrencia entre superpíxeles

Como se mencionó anteriormente, cada superpíxel representa un área estadísticamente homogénea. Esto significa no solamente que un número limitado de agrupaciones espectrales base aparecerán en un superpíxel concreto, sino que también que estas agrupaciones van a estar estrechamente relacionados en un sentido perceptivo. Por lo tanto, es posible abordar el problema de la fusión de regiones como un problema de agrupación no supervisada de agrupaciones espectrales base, teniendo en cuenta su co-ocurrencia dentro de cada superpíxel.

Sea $Sb = \{Sb_1, Sb_2, \dots, Sb_{Nsb}\}$ un conjunto de agrupaciones espectrales base donde Nsb representa el número total de agrupaciones espectrales base. Se define la relación entre la agrupación Sb_m y la agrupación Sb_n como el porcentaje medio de cobertura de área por Sb_n en cada uno de los superpíxeles en los que están presente también la agrupación espectral base Sb_m , dividido por el porcentaje total cubierto por Sb_n en la imagen. Esta relación se calcula para cada par de agrupaciones para crear una matriz de co-ocurrencia $X(Nsb, Nsb)$ que puede ser utilizada directamente por UPGMA, quedando definida como se muestra en:

$$X(m, n) = \begin{cases} 0 & \text{if } m = n \\ \frac{1}{Nsb_n / (N * M)} \sum_{i=1}^{Ns} \frac{Cs_i^{Sb_n}}{Np_i} & \text{en otro caso} \end{cases} \quad (6.4)$$

Nsb_n representa el número de píxeles que pertenece al grupo espectral n y $Cs_i^{Sb_n}$ representan los píxeles en el superpíxel Cs_i que pertenece al grupo espectral Sb_n .

6.1.1.5. Fusión de regiones basada en caracterización de superpíxels

El uso de superpíxeles como base para el análisis de imágenes basado en objetos (Object-based Image Analysis, OBIA) es un desarrollo relativamente reciente en el análisis de teledetección y representa una forma práctica de mejorar el rendimiento de clasificación sin introducir costes computacionales adicionales, como se muestra en (Blaschke, 2010) o (Garcia-Pedrero et al., 2015). Dado que ya partimos de segmentación mediante superpíxeles, es posible utilizar esta como punto de partida para el análisis de la imagen con el doble propósito de fusionar regiones y proporcionar una aproximación de bajo nivel para un posible análisis OBIA.

El propósito original de la UPGMA es la construcción de fenogramas taxonómicos, árboles que reflejaran las similitudes fenotípicas entre especies o grupos de especies. Podemos considerar el conjunto de superpíxeles como una población compuesta de diferentes subespecies, su fenotipo caracterizado por las agrupaciones espectrales base presentes en cada superpíxel. Áreas similares en la imagen producirán superpíxeles con una distribución similar de agrupaciones espectrales base. Por lo tanto, podemos considerar el proceso de fusión de regiones como un problema de agrupamiento no supervisado, utilizando la distribución de los grupos espectrales dentro de los superpíxeles como base para una medida de distancia. Cada superpíxel en el conjunto de superpíxeles resultantes en segmentación, $Cs = \{Cs_1, Cs_2, \dots, Cs_{Ns}\}$, puede describirse en relación con el conjunto base de agrupaciones espectrales, $Sb = \{Sb_1, Sb_2, \dots, Sb_{Nsb}\}$, entendido como una secuencia de valores $Vs = \{Vs_1, Vs_2, \dots, Vs_{Nsb}\}$ indicando el porcentaje del área en cada superpíxel Cs_i cubierta por cada uno de las agrupaciones espectrales de base Sb_j (6.5), para crear una matriz de similaridad $X(Ns, Nsb)$ entre superpíxeles. A partir de esta matriz de matriz de similaridad es posible crear un árbol binario de agrupamientos jerárquicos utilizando UPGMA.

$$Vs_i = \sum_{j=1}^{Nsb} \frac{Cs_i^{Sb_j}}{Np_i} \quad i \in 1 \dots Ns \quad (6.5)$$

6.1.1.6. Conjunto de pruebas

Se han seleccionado seis imágenes de satélite diferentes como conjunto de datos de prueba para este método. La prueba se realiza utilizando las bandas espectrales azul, verde, rojo e infrarrojo cercano.

Las imágenes CANADA1 y CANADA2 han sido tomadas por el satélite Landsat 7, con una resolución espacial de 30 metros. CANADA1 es un extracto de 995×1100 píxeles de una imagen tomada el 23 de Agosto de 1999, y muestra el área suroeste del lago Old Wives, en Canadá. CANADA2 forma parte de una imagen adquirida sobre la isla de Baffin en Julio de 2001, tiene

una dimensión de 1052×1224 píxeles y muestra el área que rodea la bahía de Bowman.

Las imágenes MARYLAND y ARGENTINA, con una resolución espacial de 4 metros, proceden del satélite OrbView-3. MARYLAND fue tomada en agosto de 2006 y representa una segmento de 988×1218 píxeles del área sur de Annapolis, en el estado de Maryland, EEUU. ARGENTINA es parte (1024×1148 píxeles) de una imagen tomada en febrero de 2006, y muestra la región suroeste del pueblo de Gobernador Costa.

La imagen CHILE1 es una imagen registrada por el satélite Worldview-2, cuyo sensor multiespectral tiene una resolución espacial de 1,8 metros. CHILE1 es una imagen de 2048×2048 píxeles tomada el 11 de Septiembre del 2011. El área corresponde a una zona rural ubicada en la región de Valparaíso en la Comuna de Los Andes, Chile.

La imagen CHILE2 es un corte de 1388×1350 píxeles de una imagen tomada por el satélite QuickBird-2, con una resolución de 2,4 metros por píxel. Adquirida en Diciembre de 2011, correspondiente a un área de la Comuna de Calle Larga, también en la región de Valparaíso, Chile.

6.1.2. Resultados

Se realizará una evaluación no supervisada así como una evaluación en términos cualitativos tanto del método completo como específicamente de las diferencias entre las estrategias de fusión de regiones. Al igual que en el capítulo 4, se utilizará como criterio de evaluación el criterio agregado de Levine y Nazif (Levine y Nazif, 1985a).

6.1.2.1. Evaluación de resultados para la segmentación basada en índice de rugosidad

Compararemos el método completo de segmentación de rugosidad basado en superpíxeles con una implementación del método de segmentación original (Mushrif y Ray, 2008) adaptado a imágenes multiespectrales, y los dos algoritmos de fusión propuestos con el algoritmo de fusión de distancia de color original presentado en (Cheng et al., 2002b).

La Tabla 6.1 muestra el número de conglomerados espectrales obtenidos mediante el método de rugosidad basado en superpíxeles y los resultados del método basado en el índice de rugosidad original, comparados con los resultados de un método de umbral basado en histograma, antes de la fase de fusión de conglomerados espectrales base, tanto para cada banda espectral como para el resultado final combinado. Podemos observar cómo, al tener en cuenta la distribución espacial de la imagen, tanto el método del índice de rugosidad original de Mushrif y Ray como el método desarrollado suelen proporcionar un número considerablemente mayor de agrupaciones espectrales para cada componente espectral, tal y como sucedía en imágenes a

Imagen	Método	Número de agrupaciones				
		Azul	Verde	Rojo	Infra.	Final
CHILE2	Histograma	3	4	5	9	63
	Índice de rugosidad	25	20	26	16	141
	Índice de rugosidad. Superpíxeles	17	21	22	16	171
MARYLAND	Histograma	8	10	11	11	178
	Roughness index	13	13	14	9	88
	Índice de rugosidad. Superpíxeles	18	19	23	14	177
CANADA2	Histograma	2	1	2	2	9
	Índice de rugosidad	4	5	4	4	38
	Índice de rugosidad. Superpíxeles	12	13	14	18	188
CHILE1	Histograma	4	6	7	6	80
	Índice de rugosidad	10	9	11	13	121
	Índice de rugosidad. Superpíxeles	10	11	12	13	125
ARGENTINA	Histograma	1	5	5	9	76
	Índice de rugosidad	10	12	13	16	174
	Índice de rugosidad. Superpíxeles	11	12	16	11	176
CANADA1	Histograma	2	1	2	2	7
	Índice de rugosidad	11	11	9	6	121
	Índice de rugosidad. Superpíxeles	10	12	9	8	110

Tabla 6.1: Comparación del número de conglomerados espectrales producidos por el método propuesto con los resultados de un método de umbralización basado en histogramas y los resultados del método basado en el índice de rugosidad, tanto por banda espectral como por el resultado final combinado.

color. Aunque, de nuevo, esto no garantiza una mejor segmentación, abre la posibilidad de una representación más precisa de los matices y detalles de la imagen.

En la Tabla 6.2 se presentan los resultados de la evaluación cuantitativa no supervisada a partir del criterio agregado de Levine y Nazif, donde aparecen los resultados de uniformidad intra-regional, disparidad inter-regional y el agregado total. De forma similar al caso de imágenes a color, la uniformidad intra-regional y el criterio agregado son claramente favorables al método de rugosidad basada en superpíxeles, si bien el uso de este método no parece destacar especialmente desde el punto de vista de la disparidad inter-regional. Por lo tanto, podemos afirmar que el uso de la rugosidad basada en superpíxeles para segmentación de imagen multiespectral produce segmentos significativamente más homogéneos que el método original de Mushrif y Ray, pero no más dispares. Cabe señalar que en general, las diferencias numéricas en esta evaluación cuantitativa entre los dos métodos no son muy pronunciadas.

Desde un punto de vista cualitativo, se compararán los resultados del método original modificado al dominio multiespectrales (implicando dicha adaptación la selección de picos adaptativa y la modificación del algoritmo

Imagen	N. Clusters	Uniformidad intra-regional						Variabilidad inter-regional						Valor agregado					
		Índice de rugosidad			Índice de rugosidad. Superpíxeles			Índice de rugosidad			Índice de rugosidad. Superpíxeles			Índice de rugosidad			Índice de rugosidad. Superpíxeles		
		C. Dist	Coo. Sup.	Sup. Ch.	C. Dist	Coo. Sup.	Sup. Ch.	C. Dist	Coo. Sup.	Sup. Ch.	C. Dist	Coo. Sup.	Sup. Ch.	C. Dist	Coo. Sup.	Sup. Ch.	C. Dist	Coo. Sup.	Sup. Ch.
CANADA1	5	0.7343	0.6650	0.7011	0.7809	0.7332	0.7300	0.2137	0.1307	0.2020	0.2490	0.2080	0.2151	0.4740	0.3978	0.4515	0.5150	0.4706	0.4725
	10	0.6025	0.6648	0.6144	0.6852	0.7013	0.6484	0.1787	0.1715	0.1752	0.1726	0.1755	0.1785	0.3906	0.4182	0.3948	0.4289	0.4384	0.4134
	15	0.6274	0.6432	0.5228	0.6495	0.6357	0.5457	0.1712	0.1624	0.1753	0.1716	0.1654	0.1779	0.3993	0.4028	0.3491	0.4105	0.4005	0.3618
CANADA2	5	0.8764	0.7670	0.8594	0.8306	0.8564	0.8310	0.1640	0.1711	0.2844	0.1761	0.1636	0.1897	0.5202	0.4690	0.5719	0.5034	0.5100	0.5103
	10	0.7678	0.6676	0.7081	0.8307	0.8158	0.7611	0.1877	0.1780	0.1745	0.1753	0.1703	0.1739	0.4777	0.4228	0.4413	0.5030	0.4931	0.4675
	15	0.6960	0.6725	0.6562	0.8132	0.8001	0.7003	0.1741	0.1739	0.1749	0.1749	0.1765	0.1764	0.4350	0.4232	0.4156	0.4940	0.4882	0.4383
ARGENTINA	5	0.7871	0.7939	0.7045	0.7795	0.7291	0.6945	0.1770	0.1791	0.1819	0.2142	0.1685	0.1668	0.4821	0.4865	0.4432	0.4969	0.4488	0.4306
	10	0.7288	0.6973	0.6152	0.6975	0.6278	0.5389	0.1656	0.1770	0.1722	0.1742	0.1724	0.1633	0.4472	0.4372	0.3937	0.4359	0.4001	0.3511
	15	0.6563	0.6649	0.4498	0.6423	0.6105	0.4154	0.1698	0.1687	0.1705	0.1696	0.1693	0.1618	0.4130	0.4168	0.3102	0.4060	0.3899	0.2886
MARYLAND	5	0.8541	0.8943	0.8567	0.8632	0.9003	0.8495	0.4556	0.4391	0.4158	0.3856	0.4739	0.4801	0.6548	0.6667	0.6362	0.6244	0.6871	0.6648
	10	0.8015	0.8094	0.6350	0.8323	0.8881	0.7592	0.4428	0.4377	0.4123	0.4193	0.4632	0.4688	0.6221	0.6235	0.5237	0.6258	0.6757	0.6140
	15	0.7504	0.7314	0.5337	0.8213	0.8896	0.7189	0.4337	0.4388	0.4120	0.4171	0.4571	0.4582	0.5921	0.5851	0.4729	0.6192	0.6734	0.5886
CHILE1	5	0.8233	0.8817	0.8471	0.8187	0.8851	0.8312	0.3124	0.2959	0.2943	0.2986	0.2825	0.2973	0.5679	0.5888	0.5707	0.5586	0.5838	0.5642
	10	0.8088	0.8408	0.8328	0.8013	0.8744	0.8074	0.2897	0.2895	0.2909	0.3031	0.2894	0.2923	0.5492	0.5651	0.5618	0.5522	0.5819	0.5498
	15	0.7984	0.8276	0.7331	0.8020	0.8378	0.7808	0.2869	0.2856	0.2895	0.2886	0.2851	0.2885	0.5427	0.5566	0.5113	0.5453	0.5614	0.5347
CHILE2	5	0.7568	0.8594	0.8967	0.7515	0.8239	0.9053	0.2310	0.2002	0.2302	0.2363	0.2048	0.1969	0.4939	0.5298	0.5635	0.4939	0.5143	0.5511
	10	0.7463	0.8547	0.8009	0.7439	0.8125	0.8989	0.2044	0.1931	0.2031	0.1977	0.1972	0.2004	0.4753	0.5239	0.5020	0.4708	0.5048	0.5496
	15	0.7449	0.7664	0.7859	0.7310	0.8041	0.8068	0.2017	0.1897	0.2033	0.1972	0.1929	0.1958	0.4733	0.4781	0.4946	0.4641	0.4985	0.5013

Tabla 6.2: Los resultados de los criterios de evaluación de Levine y Nazif para el algoritmo original y el método propuesto, ordenados por métodos de segmentación y fusión de regiones (donde C. Dist es la distancia de color, Coo. Sup. representa la co-ocurrencia en el método de superpíxeles y Sup. Ch. es el método de caracterización de superpíxeles). Los mejores resultados para cada criterio (Uniformidad intra-región, disparidad interregional, y los valores agregados) para cada segmentación aparecen resaltados en gris oscuro. Los mejores resultados teniendo en cuenta sólo el método de segmentación también se muestran en gris claro.

de fusión de segmentos propuesto en (Cheng et al., 2002b) para su uso en imágenes multiespectrales) y el método desarrollado usando el método de fusión de co-ocurrencia en superpíxeles ya que ambos métodos de segmentación producen segmentos agrupando píxeles. Dado que, independientemente de las combinaciones entre los métodos de segmentación y los métodos de fusión de regiones, las conclusiones generales que se obtienen de esta evaluación no cambian significativamente, los resultados y el comportamiento de los métodos de fusión de regiones se discutirán de forma independiente en la siguiente subsección. El color para los resultados proporcionados ha sido elegido para maximizar la legibilidad de las imágenes, usando (Harrower y Brewer, 2003), y no tiene ningún significado adicional.

La principal diferencia entre los dos métodos de segmentación, antes de la fase de fusión de áreas, es lo que se entiende por vecindad de un píxel y, por tanto, cómo y dónde se aplica la función de similitud que define la esfera de color para el cálculo del histon. Como hemos mencionado anteriormente, la segmentación basada en la rugosidad original define la vecindad de un píxel como un área pequeña y fija a su alrededor. Aunque es capaz de detectar la homogeneidad asociada a pequeñas variaciones en áreas bien definidas, las regiones heterogéneas o las diferencias pequeñas pero significativas entre áreas tienden a aparecer infra-representadas. Por otro lado, el uso de superpíxeles aísla eficazmente las regiones presentes en la imagen en sus propios vecindarios.

En la Fig. 6.1 se presentan los resultados para CHILE1 (fila superior) y CANADA1 (fila inferior) para 15 agrupaciones espectrales finales. Como se puede observar, en la imagen CHILE1, el uso de superpíxeles como vecindario (imagen derecha) permite discernir mejor las diferencias dentro y entre las áreas de cultivo. Los resultados de la imagen CANADA1 muestran que, a pesar del método original presentado en (Mushrif y Ray, 2008) que muestra más variedad en áreas compactas como las pequeñas lagunas, las diferencias sutiles asociadas al terreno central y las características largas y estrechas relacionadas con las corrientes de agua están adecuadamente representadas. En consecuencia, tanto para CANADA1 como para CHILE1, el método propuesto refleja mejor la distribución de las regiones en las imágenes.

En Fig. 6.2 se muestran los resultados de las imágenes CANADA2 y MARYLAND (columna central con el método original, columna derecha con el método basado en superpíxeles). CANADA2 (fila superior) es un área agrícola de baja resolución espacial, que muestra una notable variedad de tipos de terreno y cultivos; MARYLAND (fila inferior) es una imagen urbana de alta resolución espacial que presenta un conjunto heterogéneo de edificaciones, vegetación, suelo desnudo y áreas de agua. Las regiones relevantes en ambas imágenes aparecen divididas y diseminadas; como un conjunto de pequeños campos de cultivo en CANADA2 y como una red de construcciones y carreteras en MARYLAND. Al Tratar este tipo de imágenes como un conjunto

Figura 6.1: Resultados de la segmentación para imagen CHILE1 (fila superior) y para CANADA1 (fila inferior). Primera columna: Representación en falso color (infrarrojo cercano/verde/rojo). Los resultados del método original y del método propuesto para 15 conglomerados finales se muestran en la segunda y tercera columna, respectivamente. Los colores no tienen un significado especial.

Figura 6.2: Resultados de la segmentación para la imagen CANADA2 (fila superior) y para MARYLAND (fila inferior). Primera columna: Representación en falso color (infrarrojo cercano/verde/rojo). Los resultados del método original y del método propuesto para 15 conglomerados finales se muestran en la segunda y tercera columna, respectivamente. Los colores no tienen un significado especial.

de vecindarios aislados podemos obtener una mejor representación de las características presentes en el terreno, para el mismo número de agrupaciones espectrales finales, es posible capturar mejor las diferencias entre regiones, independientemente de su distribución en la imagen. Una vez más, el método original tiende a sobre-representar áreas compactas en los resultados, como el lago en la parte superior derecha de CANADA2 o el campo de golf en la parte media derecha de la imagen de MARYLAND. Por otro lado, el uso de superpíxeles como vecindario mantiene la estructura general de la imagen, representada en CANADA2 por las diferencias entre campos de cultivo y en MARYLAND por las diferencias en áreas forestales y zonas urbanas. En estas imágenes, esto también se refleja en la cantidad de agrupaciones espectrales base que se obtienen antes del paso de fusión, significativamente mayor para el método propuesto, como se presenta en la Tabla 6.1.

En Fig. 6.3 también se observa un segundo problema derivado de la definición limitada del vecindario en el método original. Los detalles de la imagen espectralmente cercanos a áreas homogéneas pero no relacionadas pueden ser segmentados como parte de estas áreas, como puede verse en MARYLAND con las zonas correspondientes al agua y las sombras presentes en la imagen y, en CANADA2, con las aguas poco profundas del lago y partes del bosque cercano. Estos artefactos no están presentes en los resultados del método propuesto. Como se verá en la siguiente subsección, este problema también puede producirse, o exacerbarse, utilizando la distancia de color como método de fusión de regiones.

6.1.2.2. Evaluación de técnicas de fusión de áreas

A partir de los resultados de la evaluación cuantitativa en la Tabla 6.2, no podemos decir realmente cual de los métodos de fusión de regiones desarrollado es preferible, si bien ambos muestran una ventaja evidente sobre el uso de distancia de color. Tanto el criterio agregado de Levine y Nazif como la uniformidad intra-región muestran mejores resultados cuando se aplican a los resultados obtenidos utilizando co-ocurrencia en superpíxeles. Desde el punto de vista de las disparidades inter-regionales, ninguno de los métodos es preferible. Teniendo en cuenta sólo la evaluación cualitativa, podemos concluir que si bien los métodos desarrollados muestran mejores resultados que la distancia de color, cual de ellos es preferible depende más de la imagen específica que del método de fusión de segmentos.

El uso de co-ocurrencia en superpíxeles reemplaza un método basado en la proximidad espectral por otro basado también en la proximidad espacial, con la idea de que esta ya ha sido tenida en cuenta durante la sobre-segmentación mediante superpíxeles. En la Fig. 6.4, en la imagen CHILE2 (fila superior) aunque ambos métodos mantienen la estructura general de la imagen, la co-ocurrencia en superpíxeles (columna derecha) representa mejor las diferencias relevantes entre y dentro de los diferentes campos de cultivo. En la

Figura 6.3: Fila superior: Detalle de la esquina superior derecha de la imagen CANADA2. Posición del extracto (imagen izquierda); resultados del método original (imagen central) y del método propuesto (imagen derecha). Fila inferior: detalles de la zona del campo de golf de la imagen MARYLAND. Posición del extracto (imagen izquierda); resultados del método original (imagen central) y del método propuesto (imagen derecha). Los colores no tienen un significado especial.

Figura 6.4: Comparación de métodos de fusión de regiones. Imagen CHILE2 en la fila superior (15 agrupaciones espectrales finales); imagen ARGENTINA (10 agrupaciones espectrales finales) en la fila inferior. En la primera columna se presenta una representación en falso color de las imágenes de prueba (infrarrojo cercano/verde/rojo). Los resultados para el método de combinación de regiones de distancia y co-ocurrencia en el método de combinación de región de superpíxeles se muestran en la segunda y tercera columna, respectivamente. Los colores no tienen un significado especial.

imagen ARGENTINA (fila inferior), las diferencias espectrales dentro de los estanques de agua asociadas a la presencia de vegetación son mayores que entre los distintos tipos de terreno, y existe una evidente relación espacial entre estas. Como resultado, el método original de distancia de color (columna central) preserva las sutiles diferencias en los estanques a expensas de las áreas presentes en el resto del terreno, mientras que el uso de co-ocurrencia en superpíxeles (columna derecha), aunque fusiona los estanques en una sola área, mantiene las diferencias generales en el terreno.

Las diferencias entre los métodos de fusión de regiones son especialmente significativas cuando buscamos una segmentación compuesta por un pequeño número de áreas. En la Fig.6.5 podemos observar los resultados de una segmentación con 5 clusters espectrales finales para las imágenes CHILE2 (fila superior) y ARGENTINA (fila inferior), tanto para el método de fusión mediante distancia de color (columna izquierda) como el de co-ocurrencia en superpíxeles (columna derecha). En la imagen ARGENTINA, la distancia de color como método de fusión de regiones se centra en los detalles de la imagen, dejando esencialmente sólo las partes prominentes de los estanques de agua y las colinas rocosas. Por otro lado, al utilizar la co-ocurrencia en superpíxeles, las principales áreas presentes en la imagen permanecen perfectamente identificables. En la imagen CHILE2 vemos, por ejemplo, que a pesar de perder detalles como las estructuras individuales en las áreas urbanas (en este caso formadas por los conglomerados espectrales asociados principalmente al asfalto para techos y a la tierra seca), las áreas urbanas permanecen como un área distinta.

Otra consecuencia del uso de un método de fusión de regiones basado en la proximidad espacial es que las agrupaciones espectrales base sin relación espacial, pero por lo demás similares, se mantendrán separados. Podemos ver también en Fig. 6.5 (fila inferior) cómo el uso de la distancia de color fusiona el agua y los parches de sombra en la imagen MARYLAND, ya que tienen una distribución espectral similar. Sin embargo, al utilizar co-ocurrencia en superpíxeles, estas dos áreas características distintas permanecen separadas, ya que no existe una clara relación de co-ocurrencia entre el agua y las sombras en la imagen.

El segundo método de fusión de regiones, la caracterización de superpíxeles, es un método basado en segmentos que aprovecha la distribución de las agrupaciones espectrales base para describir cada superpíxel como secuencias de valores. En las Fig. 6.6 y Fig. 6.7 pueden observarse algunos ejemplos de los resultados de la combinación de regiones utilizando la caracterización de superpíxeles en comparación con la combinación de áreas utilizando la co-ocurrencia en superpíxeles (la columna izquierda muestra la co-ocurrencia en los resultados de superpíxeles; la columna derecha muestra los resultados de la caracterización de superpíxeles). Como método basado en segmentos, la caracterización superpíxel evita el efecto "sal y pimienta" asociado a los

Figura 6.5: Comparación de métodos de fusión de regiones; 5 clusters espectrales finales. Fila superior: imagen CHILE2. Distancia de color (imagen izquierda) y co-ocurrencia en superpíxeles (imagen derecha). Fila central: imagen ARGENTINA. Distancia de color (imagen izquierda) y co-ocurrencia en superpíxeles (imagen derecha). Los colores no tienen un significado especial. Fila inferior: imagen MARYLAND. Distancia de color (imagen izquierda) y co-ocurrencia en superpíxeles (imagen derecha). Los colores no tienen un significado especial.

Figura 6.6: Comparación de métodos de combinación de regiones. La co-ocurrencia en superpíxeles da como resultado la caracterización de la columna izquierda y la caracterización de superpíxeles en la columna derecha. Fila superior: imagen CHILE2 (10 clusters espectrales finales). Fila central: imagen CANADA2 (10 clusters espectrales finales). Fila inferior: imagen CHILE1 (10 clusters espectrales finales). Los colores no tienen un significado especial.

Figura 6.7: Comparación de métodos de combinación de regiones. La co-ocurrencia en superpíxeles da como resultado la caracterización de la columna izquierda y la caracterización de superpíxeles en la columna derecha. Fila superior: imagen ARGENTINA (15 clusters espectrales finales). Fila central: imagen CANADA1 (10 clusters espectrales finales). Fila inferior: imagen MARYLAND (10 clusters espectrales finales). Los colores no tienen un significado especial.

métodos basados en píxeles que pueden ocurrir en imágenes complejas o de alta resolución con una alta varianza local como puede verse en la imagen CHILE2 y CANADA2 (Fig. 6.6) o ARGENTINA y MARYLAND (Fig. 6.7). Los campos de cultivo en las imágenes CHILE1 y CHILE2 (Fig. 6.6) muestran cómo el uso de la caracterización de superpíxeles también prioriza las diferencias generales entre áreas bien definidas sobre la segmentación de pequeños detalles.

Por otro lado, a pesar de un fuerte componente espacial dado por la sobre-segmentación, el método de caracterización de superpíxeles se basa en la proximidad espectral, por lo que hay una tendencia a enfatizar la homogeneidad en áreas grandes o mal definidas como se puede ver en las imágenes ARGENTINA o CANADA1 (Fig. 6.7). También en la imagen CANADA1, las corrientes de agua aparecen fragmentadas o completamente ignoradas, ya que las características representadas por áreas pequeñas o irregulares, especialmente si éstas no se han reflejado en la propia segmentación mediante superpíxeles, pueden perderse. En cualquier caso, tomando como referencia la fusión de áreas a partir de distancia de color, la estructura general de la imagen aparece mejor definida y las regiones espectralmente similares pero no relacionadas permanecen también diferenciadas.

Comparando los resultados entre el método original y el desarrollado, podemos concluir que el índice de rugosidad basado en superpíxeles permite capturar mejor las diferencias asociadas a áreas fragmentadas y de estructura compleja en la imagen, o sutiles diferencias espectrales asociadas a áreas más grandes pero no relacionadas entre sí. Los dos métodos de fusión de regiones propuestos, comparados con una fusión de regiones utilizando distancia de color, mantienen también mejor la estructura general de la imagen, sin importar el número de agrupaciones espectrales finales, y permiten que en la segmentación final regiones de la imagen espectralmente cercanas, pero de otras maneras no relacionadas se mantengan diferenciadas. El uso del algoritmo de segmentación presentado no está ligado a ningún tipo de sensor multiespectral, sino que representa un enfoque general para la combinación de información espacial y espectral en la segmentación de imágenes. Por lo tanto, el algoritmo puede adaptarse fácilmente a diferentes tipos de imágenes multiespectrales o multibanda.

6.2. Clasificación en imagen hiperespectral.

Las imágenes de teledetección hiperespectral se caracterizan por la adquisición de datos tridimensionales (3D), compuestas por dos dimensiones espaciales con información sobre la posición de los píxeles y una dimensión espectral con vectores de reflectancia de alta dimensión (Zhong et al., 2015). La ventaja de usar datos hiperespectrales sobre datos multiespectrales, imagen a color o pancromática, es la capacidad de definir características de

superficie con una resolución espectral muy alta. Sin embargo, es precisamente esta elevada resolución espectral la que hace que la clasificación de este tipo de imágenes sea, desde hace varias décadas, un importante reto en teledetección.

En general, la clasificación de imágenes hiperespectrales implica el uso efectivo de toda la información espectral asociada a cada píxel. El elevado número de bandas espectrales hace que los modelos creados para clasificación en imagen hiperespectral se vean limitados por el llamado fenómeno Hughes, es decir, se necesita una gran cantidad de muestras etiquetadas para que de los datos de alta dimensión se obtengan resultados fiables (Alonso et al., 2011). Es común plantear diversas formas de reducción de características para tratar de minimizar este hecho; podemos ver trabajos basados en el uso de PCA (Rodarmel y Shan, 2002), en análisis independiente de componentes (ICA) (Shah et al., 2004), el aprendizaje múltiple (Li et al., 2015a) o la selección directa de las bandas más representativas a partir de los datos hiperespectrales (denominado selección de bandas) (Chang et al., 1999) jerarquización de bandas (Kumar et al., 2002) o agrupación de las mismas (Martínez-Usó et al., 2007). Sin embargo, todas estas técnicas pueden implicar la pérdida de información relevante durante el proceso de reducción de la dimensionalidad.

Afortunadamente, al igual que en otros tipos de imágenes para teledetección, es muy posible que píxeles adyacentes pertenezcan a la misma tipo de cobertura, lo cual posibilita utilizar información espacial como complemento a la caracterización espectral, existiendo numerosos trabajos que aprovechan este hecho en la creación de modelos de clasificación para imagen hiperespectral (Ghamisi et al., 2017). Así, podemos ver métodos donde se utilizan diversos filtros espaciales, como filtros de propagación basados en suavizado (Chen et al., 2018), filtros de suavizado con preservación de fronteras (EPF, edge-preserving filtering) (Kang et al., 2014) o jerarquías de filtros de convolución (Zhou y Wei, 2016) para obtener características que contienen información sobre la estructuración espacial de la imagen. Alternativamente, es posible extraer información espacial/espectral fusionada a partir de filtros Gabor (Shen y Jia, 2011).

6.2.1. Materiales y métodos

En esta Tesis se han visto las posibilidades de un índice de rugosidad basado en superpíxeles para la segmentación tanto de imagen a color (en el capítulo 4) como en imagen multiespectral (en la sección anterior). En esta sección se muestra la adaptación de este proceso de segmentación para su uso en la clasificación de imágenes hiperespectrales como una forma de obtener un vector de características que explote la relación espacial/espectral en escenas hiperespectrales.

La adaptación del proceso de segmentación a imágenes hiperespectrales

implica adaptar la forma de calcular el índice de rugosidad en la clasificación de imágenes multiespectrales al volumen hiperespectral. Con la idea de obtener información no solo de las intensidades correspondientes a cada banda espectral del volumen, sino también de la evolución de ésta entre las diferentes bandas espectrales, se segmentará también un volumen calculado a partir de las tasas de cambio entre bandas.

En Fig. 6.8 se describe el método de clasificación propuesto. Este comprende las siguientes fases:

1. Se parte de dos versiones del volumen hiperespectral; la escena original y otro volumen definido, para cada píxel, por la tasa de cambio entre bandas espectrales.
2. De cara a maximizar las diferencias locales de intensidad, se ha aplicado a cada banda de cada una de los dos cubos una función de contraste local basada en histograma. (CLAHE Contrast-limited adaptive histogram equalization, (Zuiderveld, 1994))
3. Se aplica entonces, el algoritmo de segmentación por índice de aproximación basado en supervóxeles a ambos cubos.
4. Para cada píxel se genera un vector de características concatenado los valores correspondientes de los dos volúmenes segmentados.
5. Se genera un modelo de clasificación a nivel de píxeles utilizando los vectores de características calculados.

Figura 6.8: Esquema del método desarrollado para la clasificación de imágenes hiperespectrales.

6.2.1.1. Extracción de características: segmentación de imagen hiperespectral mediante un índice de rugosidad basados en supervóxeles

En la adaptación del método de segmentación de imágenes a color a segmentación de imágenes multiespectrales en teledetección presentada en la anterior sección, se planteó el uso de una segmentación mediante superpíxeles, específicamente utilizando SLIC, tanto para la creación de un conjunto de vecindarios como para el cálculo de los histones y sus correspondientes índices de rugosidad. Como ya vimos en los casos de uso de detección de espinas dendríticas en imagen de microscopía confocal y en el clasificación de volúmenes MRI para detección de la enfermedad de Alzheimer, el proceso para la generación de superpíxeles usando SLIC se puede adaptar fácilmente a un volumen tridimensional. Específicamente, se recurrirá a la versión de SLIC desarrollada para el caso de uso de detección de espinas dendríticas en microscopía confocal. Si bien desde un punto de vista conceptual un volumen de microscopía confocal y un imagen hiperespectral son completamente diferentes, uno representa principalmente información espacial, el otro se compone mayoritariamente de información espectral, desde un punto de vista metodológico no presentan diferencias.

El uso de supervóxeles para segmentar un cubo hiperespectral reduce éste a un conjunto de volúmenes localmente homogéneos. Esto nos permite explotar adecuadamente la relación espacial/espectral presente en la imagen hiperespectral. Sea el conjunto de supervóxeles resultado de una segmentación $Vs = \{Vs_1, Vs_2, \dots, Vs_{Ns}\}$, donde Ns es el número total de supervóxeles, $I(x_{Cs_j}, y_{Cs_j}, s_i)$ representa un vóxel dentro del cubo hiperespectral, Vs_j el centroide del supervóxeles y Np_j el número de vóxeles en dicho supervóxel. La función de similitud $SH(x, y, z)$ se define como:

$$SH(x, y, z) = \begin{cases} 1 & \text{if } \forall s_i \in V_p \quad d_{ts_i}(x, y, z) < E \\ 0 & \text{en otro caso} \end{cases} \quad (6.6)$$

$$d_{ts}(x, y, z) = |I(x, y, z) - I(x_{Vs_j}, y_{Vs_j})| \quad I(x, y, z) \in Cs_j \quad (6.7)$$

$$E_{v_i} = \frac{1}{Ns} \sum_{j=1}^{Ns} \frac{1}{Np_j} \sum_{I \in Vs_j} I(x, y, z) \quad (6.8)$$

siendo E la media de la desviación local dentro de cada supervóxel. La relación entre el histograma y el histon calculado a partir la función de similitud anterior dará el índice de rugosidad asociado al cubo hiperespectral.

Cada uno de los máximos presentes en el índice de rugosidad calculado representa uno o varios volúmenes de intensidad uniforme dentro del cubo

hiperespectral. La selección de los máximos considerado relevantes mantiene el criterio mostrado en la sección anterior. Los segmentos finales se determinan a partir de los valores mínimos entre máximos adyacentes. El proceso planteado no depende de que el proceso de segmentación obtenga un número específico de segmentos, por lo que se omitirá la fase de fusión de regiones.

6.2.1.2. Clasificación

Para cada píxel parte del volumen hiperespectral, su vector de características corresponderá a su equivalencia en la dimensión espectral del volumen/-segmentación resultado de aplicar el proceso de extracción de características descrito anteriormente, tanto al volumen hiperespectral original como al correspondiente a las tasas de cambio.

Dependiendo del tipo de sensor hiperespectral utilizado, es fácil que los vectores de características creados den como resultado grandes espacios dimensionales (específicamente, el doble del número de bandas espectrales originales menos una). En este caso se pretende evitar el realizar un proceso de reducción de características. Teniendo esto en cuenta, se ha elegido "random forest" como método de clasificación, dado que no solamente es considerado un método de clasificación notablemente preciso, sino que, al funcionar como un ensamblado promedio de múltiples modelos, ruidosos pero aproximadamente imparciales, puede manejar grandes vectores de características sin necesidad de hacer un filtrado previo.

6.2.1.3. Conjunto de pruebas

Para realizar la validación de la metodología propuesta se han utilizado 3 escenas hiperspectrales. Las dos primera, Indian Pines y Salinas, han sido registradas usando el "Airborne Visible/Infrared Imaging Spectrometer" (AVIRIS) con una resolución espacial de 20 metros. Indian Pines tiene un tamaño de 145×145 píxeles y 220 bandas espectrales. Presenta 16 clases etiquetadas, la mayoría de ellas diferentes tipos de cultivos. La escena Salinas, parte del valle Salinas en California, tiene un tamaño de 512×217 píxeles y 224 bandas espectrales. En esta imagen hay también información etiquetada de 16 clases diferentes. La tercera escena corresponde a una zona de la universidad de Pavía (imagen PaviaU), adquirida usando el "Reflective Optics System Imaging Spectrometer" (ROSIS). La imagen tiene un tamaño de 610×340 píxeles, está formada por 103 bandas espectrales con una resolución espacial de 1.3 metros y presenta información para 9 clases diferentes. Todas las escenas son de libre uso.

		Referencia	C. Intens.	C. Tasas	V. agregado
Indian Pines	Kappa	0,68	0,914	0,918	0,927
	valor-F ponderado	0,711	0,925	0,928	0,937
	valor-F medio	0,694	0,909	0,873	0,933
Pavia U	Kappa	0,867	0,920	0,953	0,968
	valor-F ponderado	0,898	0,940	0,65	0,976
	valor-F medio	0,88	0,933	0,947	0,969
Salinas	Kappa	0,914	0,988	0,981	0,990
	valor-F ponderado	0,922	0,989	0,984	0,992
	valor-F medio	0,851	0,987	0,987	0,989

Tabla 6.3: Métricas de precisión por clase para la escena PaviaU Ratio de verdaderos positivos (TP), ratio de falsos positiva (TF), precisión, exhaustividad y valor-F. Los resultados para la caracterización de la escena original se presentan en "C. Intens.", los resultados del volumen representando las tasas de cambio en "C. Tasas", y el vector agregado de las dos caracterizaciones en "V. agregado".

6.2.2. Resultados

Como conjunto de entrenamiento, para cada escena, se ha seleccionado aleatoriamente un 5 % de la verdad terreno presente (con un mínimo de 5 muestras), dejando el resto como conjunto de validación. Una descripción de las métricas de evaluación del rendimiento utilizadas puede verse en (Powers, 2011).

En la tabla 6.3 se muestran el valor kappa, el valor-F medio y el valor-F ponderado con el número de muestras por clase para las diferentes escenas usadas durante la validación del método propuesto. Se presentan así mismo estos valores para el método aplicado individualmente al cubo hiperespectral original y la pila representando la tasa de cambio entre bandas espectrales. Se incluye como referencia los resultados de clasificación obtenidos cuando el vector de características de cada píxel está definido por los valores espectrales de dicho píxel, también utilizando "random forest".

Podemos ver como el método propuesto es significativamente mejor que la clasificación directa, comparable con métodos recientes también basados en explotar la relación espacial presente en una imagen hiperespectral, (p.ej. (Zhang et al., 2017); (Ghamisi et al., 2014)). Destacamos también los buenos resultados obtenidos utilizando el método de clasificación de forma individual en la escena original y el volumen representando las tasas de cambio.

Un análisis más detallado de los resultados obtenidos se incluye en las tablas 6.4, 6.6 y 6.8 correspondientes a las matrices de confusión del conjunto de pruebas (PaviaU, Salinas e Indian Pines, respectivamente) y en las tablas 6.5, 6.7 y 6.9 donde se ven las métricas de precisión por clase para cada imagen

de prueba (PaviaU, Salinas e Indian Pines, respectivamente). En general, se puede ver como el método presenta un comportamiento similar en casi todas sus clases, concentrándose los fallos en las clases con muy poca representación en la imagen, lo que implica pocas muestras para el entrenamiento. De esta manera, por ejemplo, en la escena "Indian Pines" los fallos se concentran en las áreas de grava (93 muestras en total) y las zonas de hierba cortada (74 muestras en total).

Una evaluación visual de los resultados, en Fig. 6.9, Fig. 6.10 y Fig. 6.11, muestra como el método desarrollado produce áreas homogéneas, bien definidas. Puede verse como los fallos están asociados a las clases con muy pocas muestras en la imagen, tal y como se deducía de las métricas de precisión para cada imagen y en las fronteras entre regiones, donde puede esperarse cierta ambigüedad. Ejemplos de este comportamiento se pueden observar en la escena "Salinas", en las fronteras entre las zonas de lechugas (tipo 4) y maíz, o los límites entre las zonas de maíz (tipo 1), soja (tipo 1) y soja (tipo 2) en la escena "Indian Pines".

Figura 6.9: Escena Indian Pines. Falso color (banda 125 a rojo, 70 a verde y 30 a azul) (imagen superior izquierda). Verdad terreno (superior derecha) y resultados del método propuesto (inferior derecha).

Figura 6.10: Escena PaviaU. Falso color (banda 80 a rojo, 28 a verde y 9 a azul) (imagen izquierda). Verdad terreno (centro) y resultados del método propuesto (derecha).

Figura 6.11: Escena Salinas. Falso color (banda 125 a rojo, 70 a verde y 30 a azul) (izquierda). Verdad terreno (centro) y resultados del método propuesto (derecha).

	TP Ratio	FP Ratio	Prec.	Exhaust.	Valor-F
Asfalto	0,960	0,008	0,958	0,960	0,959
Hierba	0,998	0,010	0,987	0,998	0,993
Grava	0,960	0,003	0,946	0,960	0,953
Tierra	0,958	0,001	0,988	0,958	0,973
Pavimento	0,984	0,000	1,000	0,984	0,992
Ladrillo	0,982	0,001	0,994	0,982	0,988
Árboles	0,912	0,000	0,998	0,912	0,953
Metal	0,922	0,008	0,916	0,922	0,919
Sombras	1,000	0,000	0,999	1,000	0,999
Media ponderada	0,977	0,007	0,975	0,976	0,976

Tabla 6.4: Métricas de precisión por clase para la escena PaviaU Ratio de verdaderos positivos (TP), ratio de falsos positiva (TF), precisión, exhaustividad y valor-F.

	Asfalto	Hierba	Grava	Tierra	Pavimento	Ladrillo	Árboles	Metal	Sombras
Asfalto	6050	9	15	0	0	6	2	217	0
Hierba	0	17689	0	25	0	2	0	0	0
Grava	7	4	1914	0	0	8	0	61	0
Tierra	0	110	0	2789	0	10	0	0	1
Pavimento	11	7	0	0	1257	0	0	2	0
Ladrillo	0	74	0	10	0	4690	0	3	0
Árboles	97	0	0	0	0	0	1152	14	0
Metal	152	26	95	0	0	1	0	3223	0
Sombras	0	0	0	0	0	0	0	0	899

Tabla 6.5: Matriz de confusión para la escena PaviaU.

	TP Ratio	FP Ratio	Prec.	Exhaust.	Valor-F
Brócoli-1	1,000	0,000	1,000	1,000	1,000
Brócoli-2	0,995	0,000	1,000	0,995	0,997
Barbecho-1	0,998	0,001	0,964	0,998	0,981
Barbecho-2	0,982	0,001	0,980	0,982	0,981
Barbecho-3	0,984	0,001	0,987	0,984	0,985
Rastrojos	0,998	0,000	0,996	0,998	0,997
Viñedos-1	0,991	0,000	1,000	0,991	0,995
Viñedos-2	0,990	0,004	0,984	0,990	0,987
Maíz	1,000	0,001	0,994	1,000	0,997
Apio	0,972	0,001	0,991	0,972	0,981
Lechugas-1	0,989	0,000	0,990	0,989	0,990
Lechugas-2	1,000	0,000	0,998	1,000	0,999
Lechugas-3	0,980	0,001	0,949	0,980	0,964
Lechugas-4	0,943	0,000	0,981	0,943	0,961
Viñedos-3	0,982	0,002	0,988	0,982	0,985
Viñedos-4	0,989	0,000	0,998	0,989	0,993
Media ponderada	0,989	0,001	0,989	0,989	0,989

Tabla 6.6: Métricas de precisión por clase para la escena Salinas. Ratio de verdaderos positivos (TP), ratio de falsos positiva (TF), precisión, exhaustividad y valor-F.

	Br-1	Br-2	Bar-1	Bar-2	Bar-3	Rast	Viñ-1	Viñ-2	Maíz	Apio	Le-1	Le-2	Le-3	Le-4	Viñ-3	Viñ-4
Brócoli-1	1908	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Brócoli-2	0	3520	0	0	0	0	0	19	0	0	0	0	0	0	0	0
Barbecho-1	0	0	1873	2	2	0	0	0	0	0	0	0	0	0	0	0
Barbecho-2	0	0	0	1300	24	0	0	0	0	0	0	0	0	0	0	0
Barbecho-3	0	0	15	21	2503	2	0	0	0	0	3	0	0	0	0	0
Rastrojos	0	0	0	0	1	3753	0	7	0	0	0	0	0	0	0	0
Viñedos-1	0	0	1	0	0	0	3368	27	0	0	0	0	0	0	0	4
Viñedos-2	0	0	9	0	6	0	0	10605	0	2	6	0	0	0	79	0
Maíz	0	0	0	0	0	0	0	0	5891	1	0	0	0	0	0	0
Apio	0	0	40	1	0	0	1	2	37	3028	0	0	0	5	0	0
Lechugas-1	0	0	5	2	0	0	0	0	0	0	1003	4	0	0	0	0
Lechugas-2	0	0	0	0	0	0	0	0	0	0	0	1830	0	0	0	0
Lechugas-3	0	0	0	0	0	0	0	0	0	3	0	0	853	14	0	0
Lechugas-4	0	0	0	0	0	0	0	0	0	23	0	0	35	958	0	0
Viñedos-3	0	0	0	0	0	13	0	109	0	0	1	0	0	0	6781	0
Viñedos-4	0	0	0	0	0	0	0	8	0	0	0	0	11	0	0	1697

Tabla 6.7: Matriz de confusión para la escena Salinas.

	TP Ratio	FP Ratio	Prec.	Exhaust.	Valor-F
Alfalfa	0,951	0,001	0,867	0,951	0,907
Maíz-t1	0,938	0,036	0,808	0,938	0,868
Maíz-t2	0,906	0,005	0,941	0,906	0,923
Maíz-t3	0,867	0,000	0,980	0,867	0,920
pastos	0,843	0,000	0,997	0,843	0,914
Hierba/árboles	0,997	0,004	0,954	0,997	0,975
Hierba cortada	1,000	0,005	0,333	1,000	0,500
Paja volteada	0,982	0,000	1,000	0,982	0,991
Avena	1,000	0,000	1,000	1,000	1,000
Soja- t1	0,862	0,004	0,961	0,862	0,909
Soja- t2	0,964	0,014	0,957	0,964	0,961
Soja- t3	0,760	0,013	0,784	0,760	0,772
Trigo	0,995	0,000	1,000	0,995	0,997
Bosque	1,000	0,001	0,996	1,000	0,998
Edificios	0,978	0,000	0,994	0,978	0,986
Grava	0,216	0,000	1,000	0,216	0,355
media ponderada	0,928	0,010	0,935	0,928	0,928

Tabla 6.8: Métricas de precisión por clase para la escena Indian pines, Ratio de verdaderos positivos (TP), ratio de falsos positiva (TF), precisión, exhaustividad y valor-F.

	Alf	Mz1	Mz2	Mz3	Pas	Ha	Hc	Paj	Avena	Sj1	Sj2	Sj3	Tr	Bs	Ed	Gr
Alfalfa	39	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Maíz-t1	0	1272	7	0	0	0	0	0	0	8	68	1	0	0	0	0
Maíz-t2	0	66	714	1	0	0	0	0	0	0	7	0	0	0	0	0
Maíz-t3	0	13	0	195	0	14	0	0	0	0	3	0	0	0	0	0
Pastos	0	1	8	0	386	13	45	0	0	0	4	0	0	0	1	0
Hierba/árboles	0	0	0	0	0	691	0	0	0	0	0	0	0	2	0	0
Hierba cortada	0	0	0	0	0	0	23	0	0	0	0	0	0	0	0	0
Paja volteada	6	0	0	0	0	2	0	446	0	0	0	0	0	0	0	0
Avena	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0
Soja- t1	0	45	20	0	0	0	0	0	0	796	19	41	0	1	1	0
Soja- t2	0	49	4	0	0	4	1	0	0	24	2249	0	0	1	0	0
Soja- t3	0	126	5	3	0	0	0	0	0	0	0	428	0	1	0	0
Trigo	0	0	0	0	1	0	0	0	0	0	0	0	193	0	0	0
Bosque	0	0	0	0	0	0	0	0	0	0	0	0	0	1201	0	0
Edificios	0	0	1	0	0	0	0	0	0	0	0	7	0	0	358	0
Grava	0	0	0	0	0	0	0	0	0	0	0	69	0	0	0	19

Tabla 6.9: Matriz de confusión para la escena Indian Pines.

Capítulo 7

Discusión y Conclusiones

El camino desde una imagen a la interpretación de información visual de esta, a su anotación semántica automática, es largo y complejo. En el centro de este camino se encuentra la creación de un modelo de clasificación a partir de un conjunto de características. La capacidad de inferencia del modelo de clasificación, esto es, la capacidad del modelo de extraer información, en esencia la capacidad de etiquetar correctamente nuevas instancias de un sistema de anotación automática de imágenes vendrá dada por la representatividad de las características con las que ha sido entrenado. Características locales asociadas a objetos presentes en la imagen son mucho más representativas que características globales extraídas de la imagen completa. Por tanto, obtener los objetos presentes en una imagen, una segmentación, es una fase decisiva en los problemas asociados a la anotación semántica. Sin embargo, el propio proceso de segmentación representa un problema en sí mismo; sencillamente no es posible asegurar que exista una correspondencia entre los objetos presentes en una imagen y el resultado de una segmentación. Se abre entonces la posibilidad de utilizar un conjunto de primitivas más sencillas que los objetos para extraer información de una imagen, pero, y a diferencia de un píxel, con significado perceptual propio, como es el caso de los superpíxeles.

En este contexto, esta Tesis se ha centrado en el estudio de la sobre-segmentación de imágenes como herramienta para la extracción de características para anotación semántica. De esta manera, en esta Tesis se afronta el estudio, desarrollo y aplicación de los superpíxeles como elemento para realizar una caracterización. A lo largo de esta Tesis, se ha considerado cuidadosamente la naturaleza y comportamiento de una sobre-segmentación de imágenes y se han desarrollado métodos de extracción de características aprovechando esa naturaleza. El resultado es un estudio de la capacidad de un proceso de sobre-segmentación, como método para proporcionar primitivas útiles, para la caracterización de imágenes en problemas asociados a la anotación semántica.

7.1. Conclusiones

En el capítulo 3, a partir de un conjunto de imágenes/textura se evalúa la evolución de los valores medios de un conjunto de posibles características (específicamente, la media, la desviación estándar, la asimetría la curtosis, la solidez, el número de Euler, la elongación, el contraste, la excentricidad, la homogeneidad, la entropía y la correlación). Esta evaluación es un paso esencial a la hora de entender la idoneidad del uso de superpíxeles como elemento base para la extracción de características, pues representa entender cómo se van a comportar éstas como descriptores visuales, si los resultados que obtengamos pueden ser generalizados y/o hasta qué punto dependen de los parámetros específicos de la sobre-segmentación que estamos usando como base para la extracción de características. Se ha comprobado cómo estas características se comportan de una forma estable en relación a los parámetros de una sobre-segmentación utilizando SLIC, es decir pequeños cambios en los parámetros de la sobre-segmentación se asocian a pequeños cambios en los resultados de la extracción de características. Así mismo, la mayoría de las características utilizadas varían de forma predecible en relación a los parámetros con los que se ha generado la sobre-segmentación. En general, el tamaño del superpíxel tiene una influencia mucho más marcada que el factor de compacidad en el valor de las características extraídas.

En el capítulo 4 se ha desarrollado un método para la segmentación de imagen a color, una extensión de la propuesta presentada en (Mushrif y Ray, 2008). Una sobre-segmentación no solo es un conjunto de primitivas útiles para la extracción de características, sino que permite definir de forma natural cual es el entorno cercano de un píxel. De esta forma, en este método se explora la idea de explotar la homogeneidad espectral local de una imagen utilizando una segmentación mediante superpíxeles, como una forma de crear un vecindario natural para cada píxel. El algoritmo se divide en tres etapas. En la primera etapa, se calcula el índice de rugosidad basado en superpíxeles de cada uno de los componentes de color de la imagen. En la segunda etapa, se crea una segmentación inicial de la imagen utilizando los picos y valles significativos del índice de rugosidad. En la tercera se realiza la fusión de los elementos de la segmentación inicial para obtener el número de segmentos deseados, utilizando el concepto de distancia de color.

Una evaluación cuantitativa de los resultados, a partir del criterio agregado de Levine y Nazif, señala como el método desarrollado se muestra superior en un 61 % de las pruebas realizadas frente al método original y a una segmentación mediante histogramas (individualmente, en un 66 % de las pruebas frente al método original y en un 83 % comparado con una segmentación mediante histograma). Respecto a la disparidad inter-regional, el método desarrollado produce mejores resultados en un 46 % de las pruebas (individualmente se muestra superior en un 66 % de las pruebas frente al mé-

todo original y en un 53 % frente a una segmentación mediante histograma). Podemos afirmar por tanto como el método desarrollado produce segmentos notablemente más homogéneos y, en general, mejor diferenciados entre sí que el método original o que una segmentación que no tiene en cuenta relaciones espaciales en la imagen. Cuando estos dos criterios se utilizan como medida de calidad agregada, los resultados son claramente favorables al método desarrollado, siendo superior a los métodos de referencia en el 76 % de las pruebas llevadas a cabo.

En el capítulo 5, en el contexto de la validación de las capacidades de la extracción de características a nivel de sobre-segmentación, se presentan una serie de casos de uso en relación con diferentes modalidades de imagen médica.

Como primer caso de uso se ha implementado un método automático para la detección de mitosis en imagen histopatológica de cáncer de mama basado en la combinación de diferentes tipos de características sencillas extraídas de superpíxeles. El uso de estas características de forma individual no permite generar un modelo capaz de detectar ocurrencias mitóticas. Sin embargo, su uso como un vector agregado, a pesar de la sencillez del planteamiento general del método propuesto, muestra resultados (valor-F de 0,439 en las pruebas realizadas) en la media de otros métodos para detección de mitosis en imagen histopatológica.

El segundo caso de uso presenta dos métodos para identificar y reconstruir espinas dendríticas en imágenes de microscopía confocal, aprovechando las posibilidades que ofrecen las caracterizaciones realizadas a partir del entorno y la posición de un superpíxel (en este caso concreto, supervóxeles). El primer método, se desarrolla a partir del concepto de la regularidad local en el espacio definido por la segmentación mediante supervóxeles. El segundo método utiliza supervóxeles como unidad mínima para construir un árbol representativo de la dendrita. La evaluación de ambos métodos por expertos en el campo muestra resultados muy positivos. Lo que es más, el primer método, si bien puede pasar por alto algunas partes en las espinas dendríticas más grandes, es capaz de detectar espinas dendríticas pequeñas o mal definidas. El segundo método está diseñado para detectar de forma fiable las espinas dendríticas grandes y medianas. De esta forma, pueden considerarse que son métodos complementarios.

En el tercer caso de uso se ha desarrollado un método para identificar pacientes con la enfermedad de Alzheimer a partir de volúmenes MRI. Este método parte de uno de los conceptos presentados en el capítulo 4, el histón basado en superpíxeles. Este caso de uso implementa la extensión de dicho concepto a volúmenes a color y plantea su utilización como elemento de caracterización textural. Para ello, se genera un modelo utilizando SVM, usando como vector de características histones calculados a partir del agregado de una categorización en materia gris, materia blanca y líquido

cefalorraquídeo del volumen MRI, tratando dicho agregado como si fuera un volumen a color. Como conjunto de datos para la validación se ha utilizado la base de datos OASIS-1, realizando su normalización espacial y la correspondiente segmentación en tejidos mediante el proceso proporcionado por el software "Clinica". Los resultados obtenidos al validar el método desarrollado muestran una significativa ventaja en comparación con una clasificación base proporcionada por Clinica. De esta manera, se pasa de una precisión equilibrada del 67 % a una del 80 %, para un conjunto de pruebas donde se utilizan todos los pacientes de más de 60 años y de una precisión equilibrada del 72 % al 83 % en el caso de un subconjunto de pruebas donde la edad y el género del conjunto de los sujetos con la enfermedad de Alzheimer y de control son estadísticamente similares.

El capítulo 6 muestra dos casos de uso relacionados con la imagen de teledetección: la extensión del método presentado en el capítulo 4 a imagen multiespectral, y su adaptación como forma de generar una caracterización para la clasificación en imágenes hiperespectrales.

El primer caso de uso muestra la extensión del método de segmentación de imágenes a color desarrollado en el capítulo 4 a imágenes multiespectrales. Esta extensión implica adaptar el cálculo del índice de rugosidad basado en superpíxeles al dominio multiespectral y ajustar el algoritmo de selección de picos y valles a las peculiaridades de este tipo de imágenes. Así mismo, se abandona el concepto de distancia de color para realizar la fusión de regiones en favor de dos métodos basados en la sobre-segmentación mediante superpíxeles.

El resultado de la evaluación del método de segmentación de imagen multiespectral es consistente con lo obtenido para la segmentación de imágenes a color. Utilizando el criterio agregado de Levine y Nazif, la extensión del método desarrollado a imagen multiespectral obtiene mejores resultados que el método original en 77 % de las pruebas en el caso la uniformidad intra-regional y en el 72 % de estas en el caso del criterio agregado. Ni el método original ni el desarrollado destacan desde el punto de vista de la disparidad inter-regional. Las pruebas realizadas sobre los dos métodos de fusión de regiones muestran su ventaja sobre la fusión mediante distancia de color. De esta forma se obtiene con estos métodos de fusión, en conjunto, mejores resultados en el 61 % y el 66 % de las pruebas evaluando la uniformidad intra-regional y el criterio agregado, respectivamente. Desde el punto de vista de una evaluación cualitativa, el resultado final del método de segmentación para imagen multiespectral propuesto mantiene mejor la estructura general de la imagen, sin importar el número de segmentos finales, conservando pequeños detalles y teniendo en cuenta sutiles diferencias espectrales en área de mayor tamaño.

En el segundo caso de uso, se muestra la adaptación del método de segmentación de imágenes a color desarrollado en el capítulo 4 para la extrac-

ción de características como base para la generación de mapas temáticos a partir de imagen hiperespectral, con la idea de explotar la relación espacial/espectral presente en este tipo de imágenes. Cada píxel en un volumen hiperespectral se caracteriza a partir de los segmentos calculados utilizando el índice de rugosidad a nivel de supervóxeles, tanto para el volumen original como para uno representando las tasas de cambio entre bandas espectrales del volumen. Para cada caso de pruebas, los modelos se han creado partiendo con un conjunto de entrenamiento formado por un 5 % de las instancias disponibles. El modelo resultante utilizando la caracterización propuesta es significativamente superior a un modelo de referencia creado utilizando el volumen hiperespectral original. De esta forma, se pasa de un valor-F medio de 0,69, 0,88 y 0,85 en la clasificación de referencia, para las escenas de prueba "Indian pines", "PaviaU" y "Salinas" respectivamente, a valores-F medios de 0,93, 0,96 y 0,98 en el método desarrollado. Visualmente, los mapas temáticos generados por el método desarrollado muestran áreas homogéneas y bien definidas, concentrándose los fallos en las clases con muy poca información para entrenamiento y en las fronteras entre regiones.

A la luz del trabajo realizado, podemos concluir que esta Tesis demuestra las posibilidades de una sobre-segmentación mediante superpíxeles como conjunto de primitivas para la extracción de características para aplicaciones de anotación semántica. Una segmentación mediante superpíxeles, ya sea en imagen bidimensional, ya sea en sus extensiones a diferentes tipos de imágenes volumétricas, se presenta como un conjunto de primitivas flexible y fácil de caracterizar, que, a pesar de su sencillez, ha permitido el desarrollo de múltiples algoritmos para la resolución de problemas asociados a la anotación semántica en contextos muy diferentes entre sí.

7.2. Desarrollo futuro

A pesar de haber conseguido los objetivos planteados en esta Tesis, la amplitud del tema tratado permite proponer un conjunto de posibles desarrollos futuros con los cuales dar continuación a los trabajos realizados. Los resultados obtenidos muestran en que en principio cualquier problema de clasificación de imágenes puede beneficiarse del uso de superpíxeles como elemento de extracción de características. No solo eso, dado que las características extraídas a partir de una segmentación se comportan de forma estable, es posible utilizar estas como forma de reducir el número de parámetros utilizado por SLIC, ajustando el comportamiento de la sobre-segmentación a las peculiaridades de la imagen. Así, de manera similar a lo planteado en (Achanta et al., 2012b) o (Boemer et al., 2018), es posible utilizar sus características texturales para ajustar de forma individual en cada superpixel el parámetro de compacidad. Por otro lado sería posible plantear el problema de sobre-segmentación como un problema de minimización de la desviación

estándar local, donde, en lugar de utilizar un número fijo de superpíxeles, se subdividen estos hasta que desviación local media de la segmentación alcanza un umbral o se llega a un número determinado de superpíxeles.

Por supuesto, cada uno de los casos de uso presentados representa un problema en sí mismo, siendo posible extender su desarrollo.

El trabajo presentado para la detección de mitosis en imagen histopatológica muestra un notable paralelismo con el concepto de radiomics (Kumar et al., 2012). El análisis de este trabajo desde esa perspectiva no solamente sugiere nuevos conjuntos de características que aplicar al problema, sino que muestra como el uso de radiomics partiendo de una sobre-segmentación, a pesar de la existencia de algunos trabajos relacionados como (Zhang et al., 2015) o (Bharath et al., 2017), no parece haber sido convenientemente explotado.

Los notables resultados obtenidos con el método propuesto para identificar pacientes con la enfermedad de Alzheimer a partir de volúmenes MRI sugieren la extensión del trabajo a otros problemas asociados al caso de estudio. La discriminación entre el trastorno cognitivo normal y el leve o la predicción de la posible evolución de este a enfermedad de Alzheimer a partir de volúmenes MRI representan casos donde el método desarrollado puede ser aplicado sin apenas cambios. Por otro lado, teniendo en cuenta el planeamiento general de los métodos texturales basados en mapas de densidad, entre los que se encuentra el desarrollado, debería ser posible mejorar la capacidad de discriminación de la clasificación, modificando el proceso para centrar el análisis en las regiones específicas del cerebro donde se concentra la degeneración neuronal producto de la AD.

Los desarrollos de segmentación de imagen a color y multiespectral y la caracterización de imágenes hiperespectrales para la generación de mapas temáticos se basan en el concepto de índice de rugosidad, directamente derivado de la idea de conjunto aproximado. Sin embargo, la formulación original no es la única aproximación a la idea del análisis de incertidumbre. No solamente se ha desarrollado diferentes aproximaciones a la idea de conjuntos aproximados, como los presentados en (Greco et al., 2000) o en (Yao, 2007)), sino que partiendo de la misma idea que estos y considerados como un concepto complementario al de conjuntos aproximados (Pawlak, 1995), es posible formular el concepto de conjunto difuso. Existen precedentes de la extensión del método original el método original de Mushrif y Ray para segmentación de imágenes a color utilizando una de las múltiples aproximaciones al concepto de conjunto difuso (Mushrif y Ray, 2009). Es por tanto una posible línea de investigación para los métodos de segmentación desarrollados en este trabajo la extensión del concepto de histón basado en superpíxeles utilizando aproximaciones alternativas al concepto de conjuntos aproximados y/o conjuntos difusos.

Bibliografía

*Y así, del mucho leer y del poco dormir,
se le secó el cerebro de manera que vino
a perder el juicio.*

Miguel de Cervantes Saavedra

- ACHANTA, R., SHAJI, A., SMITH, K., LUCCHI, A., FUA, P. y SÄSTRUNK, S. SLIC Superpixels Compared to State-of-the-art Superpixel Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34(11), páginas 2274 – 2282, 2012a. ISSN 0162-8828. A previous version of this article was published as a EPFL Technical Report in 2010: <http://infoscience.epfl.ch/record/149300>. Supplementary material can be found at: <http://ivrg.epfl.ch/research/superpixels>.
- ACHANTA, R., SHAJI, A., SMITH, K., LUCCHI, A., FUA, P., SÜSTRUNK, S. ET AL. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*, vol. 34(11), páginas 2274–2282, 2012b.
- AHA, D. W. Editorial. En *Lazy learning*, páginas 7–10. Springer, 1997.
- AHA, D. W., KIBLER, D. y ALBERT, M. K. Instance-based learning algorithms. *Machine learning*, vol. 6(1), páginas 37–66, 1991.
- AL-AMRI, S. S., KALYANKAR, N. V. y D., K. S. Image segmentation by using threshold techniques. *CoRR*, vol. abs/1005.4020, 2010.
- ALDOUS, D. y FILL, J. Reversible markov chains and random walks on graphs. 2002.
- ALJABAR, P., HECKEMANN, R. A., HAMMERS, A., HAJNAL, J. V. y RUECKERT, D. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *Neuroimage*, vol. 46(3), páginas 726–738, 2009.
- ALONSO, M. C., MALPICA, J. A. y DE AGIRRE, A. M. Consequences of the hughes phenomenon on some classification techniques. En *Proceedings of the ASPRS 2001 annual conference*, páginas 1–5. 2011.

- ALVAREZ, V. A. y SABATINI, B. L. Anatomical and physiological plasticity of dendritic spines. *Annu. Rev. Neurosci.*, vol. 30, páginas 79–97, 2007.
- ANTONIE, L. Automated segmentation and classification of brain magnetic resonance imaging. *C615 Project*, páginas 1–15, 2008.
- ARBELAEZ, P., MAIRE, M., FOWLKES, C. y MALIK, J. Contour detection and hierarchical image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, vol. 33(5), páginas 898–916, 2011.
- ASHBURNER, J. y FRISTON, K. J. Voxel-based morphometry-the methods. *Neuroimage*, vol. 11(6), páginas 805–821, 2000.
- ASHBURNER, J. y FRISTON, K. J. Unified segmentation. *Neuroimage*, vol. 26(3), páginas 839–851, 2005.
- BALAGURUNATHAN, Y., GU, Y., WANG, H., KUMAR, V., GROVE, O., HAWKINS, S., KIM, J., GOLDFOF, D. B., HALL, L. O., GATENBY, R. A. ET AL. Reproducibility and prognosis of quantitative features extracted from ct images. *Translational oncology*, vol. 7(1), páginas 72–87, 2014.
- BANERJEE, B., VARMA, S., BUDDHIRAJU, K. y EETI, L. Unsupervised multi-spectral satellite image segmentation combining modified mean-shift and a new minimum spanning tree based clustering technique. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 7(3), páginas 888–894, 2014. ISSN 1939-1404.
- BARALDI, A. y PARMIGGIANI, F. An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33(2), páginas 293–304, 1995.
- BEHESHTI, I., DEMIREL, H., INITIATIVE, A. D. N. ET AL. Probability distribution function-based classification of structural mri for the detection of alzheimer's disease. *Computers in biology and medicine*, vol. 64, páginas 208–216, 2015.
- BEJNORDI, B. E., LITJENS, G., HERMSEN, M., KARSSEMEIJER, N. y VAN DER LAAK, J. A. A multi-scale superpixel classification approach to the detection of regions of interest in whole slide histopathology images. En *Medical Imaging 2015: Digital Pathology*, vol. 9420, página 94200H. International Society for Optics and Photonics, 2015.
- BELICHENKO, P. V. y DAHLSTRÖM, A. Studies on the 3-dimensional architecture of dendritic spines and varicosities in human cortex by confocal laser scanning microscopy and lucifer yellow microinjections. *Journal of neuroscience methods*, vol. 57(1), páginas 55–61, 1995.

- VAN DEN BERGH, M., BOIX, X., ROIG, G., DE CAPITANI, B. y VAN GOOL, L. Seeds: Superpixels extracted via energy-driven sampling. En *European conference on computer vision*, páginas 13–26. Springer, 2012.
- BEUCHER, S. y MEYER, F. The morphological approach to segmentation: the watershed transformation. *Optical Engineering-New York-Marcel Dekker Incorporated-*, vol. 34, páginas 433–433, 1992.
- BHARATH, H., COLLEMAN, S., SIMA, D. M. y VAN HUFFEL, S. Tumor segmentation from multimodal mri using random forest with superpixel and tensor based feature extraction. En *International MICCAI Brainlesion Workshop*, páginas 463–473. Springer, 2017.
- BHARATHI, P. y SUBASHINI, P. Optimization of image processing techniques using neural networks: a review. *WSEAS Transactions on Information Science and Applications*, vol. 8(8), páginas 300–328, 2011.
- BIER, B., RAVIKUMAR, N., UNBERATH, M., LEVENSTON, M., GOLD, G., FAHRIG, R. y MAIER, A. Range imaging for motion compensation in c-arm cone-beam ct of knees under weight-bearing conditions. *Journal of Imaging*, vol. 4(1), página 13, 2018.
- BLASCHKE, T. Object based image analysis for remote sensing. *{ISPRS} Journal of Photogrammetry and Remote Sensing*, vol. 65(1), páginas 2 – 16, 2010. ISSN 0924-2716.
- BLASCHKE, T., BURNETT, C. y PEKKARINEN, A. Image segmentation methods for object-based analysis and classification. En *Remote sensing image analysis: Including the spatial domain*, páginas 211–236. Springer, 2004.
- BLEI, D. M. y JORDAN, M. I. Modeling annotated data. En *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*, páginas 127–134. ACM, 2003.
- BLEI, D. M., NG, A. Y. y JORDAN, M. I. Latent dirichlet allocation. *Journal of machine Learning research*, vol. 3(Jan), páginas 993–1022, 2003.
- BOEMER, F., RATNER, E. y LENDASSE, A. Parameter-free image segmentation with slic. *Neurocomputing*, vol. 277, páginas 228–236, 2018.
- BOIMAN, O. y IRANI, M. Detecting irregularities in images and in video. *International journal of computer vision*, vol. 74(1), páginas 17–31, 2007.
- BOLAND, M. V., MARKEY, M. K. y MURPHY, R. F. Automated recognition of patterns characteristic of subcellular structures in fluorescence microscopy images. *Cytometry: The Journal of the International Society for Analytical Cytology*, vol. 33(3), páginas 366–375, 1998.

- BOYKOV, Y., VEKSLER, O. y ZABIH, R. Markov random fields with efficient approximations. En *Computer vision and pattern recognition, 1998. Proceedings. 1998 IEEE computer society conference on*, páginas 648–655. IEEE, 1998.
- BOYKOV, Y., VEKSLER, O. y ZABIH, R. Fast approximate energy minimization via graph cuts. *IEEE Transactions on pattern analysis and machine intelligence*, vol. 23(11), páginas 1222–1239, 2001.
- BRAAK, H. y BRAAK, E. Frequency of stages of alzheimer-related lesions in different age categories. *Neurobiology of aging*, vol. 18(4), páginas 351–357, 1997.
- BRATASANU, D., NEDELICU, I. y DATCU, M. Bridging the semantic gap for satellite image annotation and automatic mapping applications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4(1), página 193, 2011.
- BREHENY, P. Classification and regression trees. 1984.
- BREIMAN, L. *Classification and regression trees*. Routledge, 2017.
- BRUNELLI, R. *Template matching techniques in computer vision: theory and practice*. John Wiley & Sons, 2009.
- CAELLI, T. *Visual perception*. Elsevier, 1981.
- CALLAERT, D. V., RIBBENS, A., MAES, F., SWINNEN, S. P. y WENDE-ROTH, N. Assessing age-related gray matter decline with voxel-based morphometry depends significantly on segmentation and normalization procedures. *Frontiers in Aging Neuroscience*, vol. 6, página 124, 2014. ISSN 1663-4365.
- CASELLES, V., CATTÉ, F., COLL, T. y DIBOS, F. A geometric model for active contours in image processing. *Numerische mathematik*, vol. 66(1), páginas 1–31, 1993.
- CATTELL, R. B. The scree test for the number of factors. *Multivariate Behavioral Research*, vol. 1(2), páginas 245–276, 1966. PMID: 26828106.
- CHABRIER, S., EMILE, B., ROSENBERGER, C. y LAURENT, H. Unsupervised performance evaluation of image segmentation. *EURASIP Journal on Applied Signal Processing*, vol. 2006, páginas 217–217, 2006.
- CHANG, C.-I., DU, Q., SUN, T.-L. y ALTHOUSE, M. L. A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification. *IEEE transactions on geoscience and remote sensing*, vol. 37(6), páginas 2631–2641, 1999.

- CHANG, H., LOSS, L. A. y PARVIN, B. Nuclear segmentation in h&e sections via multi-reference graph cut. En *International symposium biomedical imaging*. 2012.
- CHAWLA, N. V., BOWYER, K. W., HALL, L. O. y KEGELMEYER, W. P. Smote: Synthetic minority over-sampling technique. *J. Artif. Int. Res.*, vol. 16(1), páginas 321–357, 2002. ISSN 1076-9757.
- CHEN, H., DOU, Q., WANG, X., QIN, J. y HENG, P.-A. Mitosis detection in breast cancer histology images via deep cascaded networks. En *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, AAAI'16, páginas 1160–1166. AAAI Press, 2016.
- CHEN, Z., JIANG, J., JIANG, X., FANG, X. y CAI, Z. Spectral-spatial feature extraction of hyperspectral images based on propagation filter. *Sensors*, vol. 18(6), página 1978, 2018.
- CHEN, Z., ZHAO, Z., GONG, P. y ZENG, B. A new process for the segmentation of high resolution remote sensing imagery. *International Journal of Remote Sensing*, vol. 27(22), páginas 4991–5001, 2006.
- CHENG, H.-D., JIANG, X. y WANG, J. Color image segmentation based on homogram thresholding and region merging. *Pattern recognition*, vol. 35(2), páginas 373–393, 2002a.
- CHENG, H. D., JIANG, X. H. y WANG, J. Color image segmentation based on homogram thresholding and region merging. *Pattern Recognition*, vol. 35, páginas 373–393, 2002b.
- CHO, Y., SEONG, J.-K., JEONG, Y., SHIN, S. Y., INITIATIVE, A. D. N. ET AL. Individual subject classification for alzheimer's disease based on incremental learning using a spatial frequency representation of cortical thickness data. *Neuroimage*, vol. 59(3), páginas 2217–2230, 2012.
- CLARK, C. M., PONTECORVO, M. J., BEACH, T. G., BEDELL, B. J., COLEMAN, R. E., DORAISWAMY, P. M., FLEISHER, A. S., REIMAN, E. M., SABBAGH, M. N., SADOWSKY, C. H. ET AL. Cerebral pet with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid- β plaques: a prospective cohort study. *The Lancet Neurology*, vol. 11(8), páginas 669–678, 2012.
- CLARK, K., VENDT, B., SMITH, K., FREYMAN, J., KIRBY, J., KOPPEL, P., MOORE, S., PHILLIPS, S., MAFFITT, D., PRINGLE, M. ET AL. The cancer imaging archive (tcia): maintaining and operating a public information repository. *Journal of digital imaging*, vol. 26(6), páginas 1045–1057, 2013.

- COHEN, L. D. On active contour models and balloons. *CVGIP: Image understanding*, vol. 53(2), páginas 211–218, 1991.
- CONRAD, C., MERTZ, M. y MESTER, R. Contour-relaxed superpixels. En *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, páginas 280–293. Springer, 2013.
- CRUZ-ROA, A., BASAVANHALLY, A., GONZÁLEZ, F., GILMORE, H., FELDMAN, M., GANESAN, S., SHIH, N., TOMASZEWSKI, J. y MADABHUSHI, A. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. En *Medical Imaging 2014: Digital Pathology*, vol. 9041, página 904103. International Society for Optics and Photonics, 2014.
- CSILLIK, O. Superpixels: The end of pixels in obia. a comparison of state-of-the-art superpixel methods for remote sensing data. 2016.
- CSILLIK, O. Fast segmentation and classification of very high resolution remote sensing data using slic superpixels. *Remote Sensing*, vol. 9(3), página 243, 2017.
- DARAMOLA, S. A. y ADEMOLA, A. Vehicle classification algorithm using size and shape. *International Journal of Soft Computing and Engineering*, vol. 6(1), páginas 78–81, 2016.
- DEL FRATE, F., PACIFICI, F., SCHIAVON, G. y SOLIMINI, C. Use of neural networks for automatic classification from high-resolution images. *IEEE transactions on geoscience and remote sensing*, vol. 45(4), páginas 800–809, 2007.
- DING, K. y GUNASEKARAN, S. Shape feature extraction and classification of food material using computer vision. *Transactions of the ASAE*, vol. 37(5), páginas 1537–1545, 1994.
- DOI, K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized medical imaging and graphics*, vol. 31(4-5), páginas 198–211, 2007.
- DORIGO, M., MANIEZZO, V. y COLORNI, A. The ant system: Optimization by a colony of cooperating agents. *IEEE Trans. on Systems, Man and Cybernetics-Part B*, vol. 26(1), páginas 29–41, 1996.
- DOYLE, P. The use of automatic interaction detector and similar search procedures. *Journal of the Operational Research Society*, vol. 24(3), páginas 465–467, 1973.
- DUBOIS, B. y ALBERT, M. L. Amnestic mci or prodromal alzheimer's disease? *The Lancet Neurology*, vol. 3(4), páginas 246–248, 2004.

- DUMAIS, S. T. Latent semantic analysis. *Annual review of information science and technology*, vol. 38(1), páginas 188–230, 2004.
- DUNTSCH, I. y GEDIGA, G. Statistical evaluation of rough set dependency analysis. *International Journal of Human-Computer Studies*, vol. 46(5), páginas 589–604, 1997.
- EGGLESTON, P. Understanding oversegmentation and region merging. *Vision Systems Design*, 1998.
- EKLUNDH, J. O., YAMAMOTO, H. y ROSENFELD, A. A relaxation method for multispectral pixel classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2(1), páginas 72–75, 1980. ISSN 0162-8828.
- ELMQVIST, N. y FEKETE, J.-D. Hierarchical aggregation for information visualization: Overview, techniques, and design guidelines. *IEEE Transactions on Visualization and Computer Graphics*, vol. 16(3), páginas 439–454, 2010.
- ELSTON, C. y ELLIS, I. pathological prognostic factors in breast cancer. i. the value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology*, vol. 19(5), páginas 403–410, 1991. ISSN 1365-2559.
- ERDEM, C. E., SANKUR, B. y TEKALP, A. M. Performance measures for video object segmentation and tracking. *IEEE Transactions on Image Processing*, vol. 13(7), páginas 937–951, 2004.
- VON F, Z. Beugungstheorie des schneidenverfahrens und seiner verbesserten form, der phasenkontrastmethode. *Physica*, vol. 1(7-12), páginas 689–704, 1934.
- FAROOQUE, M. A. y ROHANKAR, J. S. Survey on various noises and techniques for denoising the color image. *International Journal of Application or Innovation in Engineering & Management (IJAIEEM)*, vol. 2(11), páginas 217–221, 2013.
- FELZENSZWALB, P. F. y HUTTENLOCHER, D. P. Efficient graph-based image segmentation. *International journal of computer vision*, vol. 59(2), páginas 167–181, 2004.
- FENG, S., MANMATHA, R. y LAVRENKO, V. Multiple bernoulli relevance models for image and video annotation. En *null*, páginas 1002–1009. IEEE, 2004.
- FERRARI, V., FEVRIER, L., JURIE, F. y SCHMID, C. Groups of adjacent contour segments for object detection. *IEEE transactions on pattern analysis and machine intelligence*, vol. 30(1), páginas 36–51, 2008.

- FERRI, C., PRINCE, M., BRAYNE, C., BRODATY, H., FRATIGLIONI, L., GANGULI, M., HALL, K., HASEGAWA, K., HENDRIE, H., HUANG, Y., JORM, A., MATHERS, C., MENEZES, P., RIMMER, E. y SCAZUFCA, M. Global prevalence of dementia: A delphi consensus study. *Lancet*, vol. 366(9503), páginas 2112–2117, 2005. ISSN 0140-6736.
- FIX, E. y HODGES, J. An important contribution to nonparametric discriminant analysis and density estimation. *International Statistical Review*, vol. 3(57), páginas 233–238, 1951.
- FLOWER, M. A. *Webb's physics of medical imaging*. CRC Press, 2012.
- FREEMAN, J. The modelling of spatial relations. *Computer graphics and image processing*, vol. 4(2), páginas 156–171, 1975.
- FRIEDMAN, N. y RUSSELL, S. Image segmentation in video sequences: A probabilistic approach. En *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*, UAI'97, páginas 175–181. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1997. ISBN 1-55860-485-5.
- FRÖHLICH, B., BACH, E., WALDE, I., HESE, S., SCHMULLIUS, C. y DENZLER, J. Land cover classification of satellite images using contextual information. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. II-3/W1, páginas 1–6, 2013.
- FU, K.-S. y MUI, J. A survey on image segmentation. *Pattern recognition*, vol. 13(1), páginas 3–16, 1981.
- FU, X., YU, R., ZHANG, W., WU, J. y SHAO, S. Delving deep into multiscale pedestrian detection via single scale feature maps. *Sensors*, vol. 18(4), página 1063, 2018.
- FUKUSHIMA, K. y MIYAKE, S. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. En *Competition and cooperation in neural nets*, páginas 267–285. Springer, 1982.
- GALAVIS, P. E., HOLLENSSEN, C., JALLOW, N., PALIWAL, B. y JERAJ, R. Variability of textural features in fdg pet images due to different acquisition modes and reconstruction parameters. *Acta Oncologica*, vol. 49(7), páginas 1012–1016, 2010.
- GARBAY, C. Image structure representation and processing: a discussion of some segmentation methods in cytology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (2), páginas 140–146, 1986.
- GARCIA FREITAS, P., DA EIRA, L., SANTOS, S. y FARIAS, M. On the application lbp texture descriptors and its variants for no-reference image quality assessment. *Journal of Imaging*, vol. 4(10), página 114, 2018.

- GARCIA-LAMONT, F., CERVANTES, J., LÓPEZ, A. y RODRIGUEZ, L. Segmentation of images by color features: a survey. *Neurocomputing*, vol. 292, páginas 1–27, 2018.
- GARCIA-PEDRERO, A., GONZALO-MARTÍN, C., FONSECA-LUENGO, D. y LILLO-SAAVEDRA, M. A geobia methodology for fragmented agricultural landscapes. *Remote Sensing*, vol. 7(1), página 767, 2015. ISSN 2072-4292.
- GE, Y., CHEN, Y., JIA, Y., GUO, X. y HU, S. Dynamic monitoring the infrastructure of major ports in sri lanka by using multi-temporal high spatial resolution remote sensing images. *Journal of Geographical Sciences*, vol. 28(7), páginas 973–984, 2018.
- GERBRANDS, J. J. Segmentation of noisy images. 1988.
- GHAMISI, P., BENEDIKTSSON, J. A. y ULFARSSON, M. O. Spectral-spatial classification of hyperspectral images based on hidden markov random fields. *IEEE Trans. Geoscience and Remote Sensing*, vol. 52(5), páginas 2565–2574, 2014.
- GHAMISI, P., PLAZA, J., CHEN, Y., LI, J. y PLAZA, A. J. Advanced spectral classifiers for hyperspectral images: A review. *IEEE Geoscience and Remote Sensing Magazine*, vol. 5(1), páginas 8–32, 2017.
- GNEP, K., FARGEAS, A., GUTIÉRREZ-CARVAJAL, R. E., COMMANDEUR, F., MATHIEU, R., OSPINA, J. D., ROLLAND, Y., ROHOU, T., VINCENDEAU, S., HATT, M. ET AL. Haralick textural features on t2-weighted mri are associated with biochemical recurrence following radiotherapy for peripheral zone prostate cancer. *Journal of Magnetic Resonance Imaging*, vol. 45(1), páginas 103–117, 2017.
- GONG, Y., ZHANG, H. ET AL. An image database system with content capturing and fast image indexing abilities. En *Multimedia Computing and Systems, 1994., Proceedings of the International Conference on*, páginas 121–130. IEEE, 1994.
- GONZALEZ, R. C., WOODS, R. E. ET AL. Digital image processing. 2002.
- GOTO, M., ABE, O., AOKI, S., HAYASHI, N., MIYATI, T., TAKAO, H., IWATSUBO, T., YAMASHITA, F., MATSUDA, H., MORI, H. ET AL. Diffeomorphic anatomical registration through exponentiated lie algebra provides reduced effect of scanner for cortex volumetry with atlas-based method in healthy subjects. *Neuroradiology*, vol. 55(7), páginas 869–875, 2013.
- GRECO, S., MATARAZZO, B. y SLOWINSKI, R. *Multicriteria classification by dominance-based rough set approach*. Università degli studi, 2000.

- GUDI, C., NANDYAL, S., DATTATREYA, S. y PATIL, P. Leaf classification based on shape features using backpropagation neural networks. 1994.
- GUO, L., MA, Y., CUKIC, B. y SINGH, H. Robust prediction of fault-proneness by random forests. En *Software Reliability Engineering, 2004. ISSRE 2004. 15th International Symposium on*, páginas 417–428. IEEE, 2004.
- GUO, Q., KELLY, M., GONG, P. y LIU, D. An object-based classification approach in mapping tree mortality using high spatial resolution imagery. *GIScience & Remote Sensing*, vol. 44(1), páginas 24–47, 2007.
- HARALICK, R. y DINSTEIN, I. A spatial clustering procedure for multi-image data. *Circuits and Systems, IEEE Transactions on*, vol. 22(5), páginas 440–450, 1975. ISSN 0098-4094.
- HARALICK, R., SHANMUGAM, K. y DINSTEIN, I. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, vol. SMC-3(6), páginas 610–621, 1973. ISSN 0018-9472.
- HARALICK, R. M. y SHAPIRO, L. G. *Computer and robot vision*. Addison-wesley, 1992.
- HARE, J. S., LEWIS, P. H., ENSER, P. G. y SANDOM, C. J. Mind the gap: another look at the problem of the semantic gap in image retrieval. En *Multimedia Content Analysis, Management, and Retrieval 2006*, vol. 6073, página 607309. International Society for Optics and Photonics, 2006.
- HARROWER, M. y BREWER, C. A. ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. *The Cartographic Journal*, vol. 40(1), páginas 27–37, 2003. ISSN 0008-7041.
- HE, S., NI, J., WU, L., WEI, H. y ZHAO, S. Image threshold segmentation method with 2-d histogram based on multi-resolution analysis. En *Computer Science Education, 2009. ICCSE '09. 4th International Conference on*, páginas 753–757. 2009.
- HE, W., HAMILTON, T. A., COHEN, A. R., HOLMES, T. J., PACE, C., SZAROWSKI, D. H., TURNER, J. N. y ROYSAM, B. Automated three-dimensional tracing of neurons in confocal and brightfield images. *Microsc. Microanal*, vol. 9, páginas 296–310, 2003.
- HENSON, R. y FRISTON, K. Statistical parametric mapping: the analysis of functional brain images. 2006.
- HENTSCHEL, C., STOBBER, S., NÜRNBERGER, A. y DETYNIECKI, M. Automatic image annotation using a visual dictionary based on reliable image segmentation. En *International Workshop on Adaptive Multimedia Retrieval*, páginas 45–56. Springer, 2007.

- HERNÁNDEZ-GRACIDAS, C. y SUCAR, L. E. Markov random fields and spatial information to improve automatic image annotation. En *Pacific-Rim Symposium on Image and Video Technology*, páginas 879–892. Springer, 2007.
- HSIAO, J. y SAWCHUK, A. Supervised textured image segmentation using feature smoothing and probabilistic relaxation techniques. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11(12), páginas 1279–1292, 1989. ISSN 0162-8828.
- HSIEH, J. Image artifacts: appearances, causes, and corrections. 2009.
- HSU, C.-W. y LIN, C.-J. A comparison of methods for multiclass support vector machines. *IEEE transactions on Neural Networks*, vol. 13(2), páginas 415–425, 2002.
- HUANG, T., YANG, G. y TANG, G. A fast two-dimensional median filtering algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 27(1), páginas 13–18, 1979.
- HUMPHREY, G. The psychology of the gestalt. *Journal of Educational Psychology*, vol. 15(7), página 401, 1924.
- IBARROLA-ULZURRUN, E., MARCELLO, J. y GONZALO-MARTIN, C. Advanced classification of remote sensing high resolution imagery. an application for the management of natural resources. En *Developments and Advances in Intelligent Systems and Applications*, páginas 1–13. Springer, 2018.
- IRSHAD, H. Automated mitosis detection in histopathology using morphological and multi-channel statistics features. *Journal of pathology informatics*, vol. 4, 2013.
- ISHII, K., KAWACHI, T., SASAKI, H., KONO, A. K., FUKUDA, T., KOJIMA, Y. y MORI, E. Voxel-based morphometric comparison between early- and late-onset mild alzheimer’s disease and assessment of diagnostic performance of z score images. *American Journal of Neuroradiology*, vol. 26(2), páginas 333–340, 2005.
- JABA, L., SHANTHI, V. y SINGH, D. Estimation of hippocampus volume from mri using imagej for alzheimer’s diagnosis. *Atlas Journal of Medical and Biological Sciences*, vol. 1(1), páginas 15–20, 2017.
- JACK JR, C. R., WISTE, H. J., VEMURI, P., WEIGAND, S. D., SENJEM, M. L., ZENG, G., BERNSTEIN, M. A., GUNTER, J. L., PANKRATZ, V. S., AISEN, P. S. ET AL. Brain beta-amyloid measures and magnetic resonance imaging atrophy both predict time-to-progression from mild cognitive impairment to alzheimer’s disease. *Brain*, vol. 133(11), páginas 3336–3348, 2010.

- JACOBS, D. W. Robust and efficient detection of salient convex groups. *IEEE transactions on pattern analysis and machine intelligence*, vol. 18(1), páginas 23–37, 1996.
- JAEGER, S., PALANIAPPAN, K., CASAS-DELUCCHI, C. S. y CARDOSO, M. C. Classification of cell cycle phases in 3d confocal microscopy using pcna and chromocenter features. En *Proceedings of the Seventh Indian Conference on Computer Vision, Graphics and Image Processing, ICVGIP '10*, páginas 412–418. ACM, New York, NY, USA, 2010. ISBN 978-1-4503-0060-5.
- JAIN, R., KASTURI, R. y SCHUNCK, B. G. *Machine vision*, vol. 5. McGraw-Hill New York, 1995.
- JANOOS, F., MOSALIGANTI, K., XU, X., MACHIRAJU, R., HUANG, K. y WONG, S. T. Robust 3d reconstruction and identification of dendritic spines from optical microscopy imaging. *Medical image analysis*, vol. 13(1), páginas 167–179, 2009.
- JÉGOU, H., DOUZE, M., SCHMID, C. y PÉREZ, P. Aggregating local descriptors into a compact image representation. En *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, páginas 3304–3311. IEEE, 2010.
- JENSEN, J. R. y LULLA, K. Introductory digital image processing: a remote sensing perspective. 1987.
- JONES, E. y POWELL, T. Morphological variations in the dendritic spines of the neocortex. *Journal of cell science*, vol. 5(2), páginas 509–529, 1969.
- JONSSON, R. Exact properties of mcnemar’s test in small samples. 1993.
- JOSHI, G. D. y SIVASWAMY, J. A simple scheme for contour detection. En *VISAPP (1)*, páginas 236–242. 2006.
- KANG, X., LI, S. y BENEDIKTSSON, J. A. Spectral–spatial hyperspectral image classification with edge-preserving filtering. *IEEE transactions on geoscience and remote sensing*, vol. 52(5), páginas 2666–2677, 2014.
- KARANTZALOS, K. y ARGIALAS, D. Improving edge detection and watershed segmentation with anisotropic diffusion and morphological levellings. *International Journal of Remote Sensing*, vol. 27(24), páginas 5427–5434, 2006.
- KARTIKEYAN, B., SARKAR, A. y MAJUMDER, K. A segmentation approach to classification of remote sensing imagery. *International Journal of Remote Sensing*, vol. 19(9), páginas 1695–1709, 1998.

- KASS, M., WITKIN, A. y TERZOPOULOS, D. Snakes: Active contour models. *International journal of computer vision*, vol. 1(4), páginas 321–331, 1988.
- KASSNER, A. y THORNHILL, R. Texture analysis: a review of neurologic mr imaging applications. *American Journal of Neuroradiology*, vol. 31(5), páginas 809–816, 2010.
- KHAN, A. M., EL-DALY, H. y RAJPOOT, N. M. A gamma-gaussian mixture model for detection of mitotic cells in breast cancer histopathology images. En *Pattern Recognition (ICPR), 2012 21st International Conference on*, páginas 149–152. IEEE, 2012.
- KLOMP, J. Economic development and natural disasters: A satellite data analysis. *Global Environmental Change*, vol. 36, páginas 67–88, 2016.
- KLÖPPEL, S., STONNINGTON, C. M., CHU, C., DRAGANSKI, B., SCAHILL, R. I., ROHRER, J. D., FOX, N. C., JACK JR, C. R., ASHBURNER, J. y FRACKOWIAK, R. S. Automatic classification of mr scans in alzheimer's disease. *Brain*, vol. 131(3), páginas 681–689, 2008.
- KOONSANIT, K., JARUSKULCHAI, C. y EIUMNOH, A. Parameter-free k-means clustering algorithm for satellite imagery application. En *Information Science and Applications (ICISA), 2012 International Conference on*, páginas 1–6. 2012.
- KOTSIANTIS, S. B., ZAHARAKIS, I. y PINTELAS, P. Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering*, vol. 160, páginas 3–24, 2007.
- KRIG, S. Image pre-processing. En *Computer Vision Metrics*, páginas 35–74. Springer, 2016.
- KUMAR, G. y BHATIA, P. K. A detailed review of feature extraction in image processing systems. En *Advanced Computing & Communication Technologies (ACCT), 2014 Fourth International Conference on*, páginas 5–12. IEEE, 2014.
- KUMAR, S., GHOSH, J. y CRAWFORD, M. M. Hierarchical fusion of multiple classifiers for hyperspectral data analysis. *Pattern Analysis & Applications*, vol. 5(2), páginas 210–220, 2002.
- KUMAR, V., GU, Y., BASU, S., BERGLUND, A., ESCHRICH, S. A., SCHA-BATH, M. B., FORSTER, K., AERTS, H. J., DEKKER, A., FENSTERMA-CHER, D. ET AL. Radiomics: the process and the challenges. *Magnetic resonance imaging*, vol. 30(9), páginas 1234–1248, 2012.
- LAVRENKO, V., MANMATHA, R. y JEON, J. A model for learning the semantics of pictures. En *Advances in neural information processing systems*, páginas 553–560. 2004.

- LEGENDRE, P. y LEGENDRE, L. *Numerical Ecology*. Developments in Environmental Modelling. Elsevier Science, 1998. ISBN 9780080537870.
- LEVINE, M. y NAZIF, A. M. Dynamic measurement of computer generated image segmentations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PAMI-7(2), páginas 155–164, 1985a. ISSN 0162-8828.
- LEVINE, M. D. y NAZIF, A. M. Dynamic measurement of computer generated image segmentations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (2), páginas 155–164, 1985b.
- LEVINSHTEIN, A., STERE, A., KUTULAKOS, K. N., FLEET, D. J., DICKINSON, S. J. y SIDDIQI, K. Turbopixels: Fast superpixels using geometric flows. *IEEE transactions on pattern analysis and machine intelligence*, vol. 31(12), páginas 2290–2297, 2009.
- LEYMARIE, F. y LEVINE, M. D. Simulating the grassfire transform using an active contour model. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (1), páginas 56–75, 1992.
- LI, J., HUANG, X., GAMBA, P., BIOUCAS-DIAS, J. M., ZHANG, L., BENEDIKTSSON, J. A. y PLAZA, A. Multiple feature learning for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53(3), páginas 1592–1606, 2015a.
- LI, P. y XIAO, X. Multispectral image segmentation by a multichannel watershed-based approach. *International Journal of Remote Sensing*, vol. 28(19), páginas 4429–4452, 2007.
- LI, Q., ZHOU, X., DENG, Z., BARON, M., TEYLAN, M. A., KIM, Y. y WONG, S. T. A novel surface-based geometric approach for 3d dendritic spine detection from multi-photon excitation microscopy images. En *Bio-medical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*, páginas 1255–1258. IEEE, 2009.
- LI, S., YUAN, X., PU, F., LI, D., FAN, Y., WU, L., CHAO, W., CHEN, N., HE, Y. y HAN, Y. Abnormal changes of multidimensional surface features using multivariate pattern classification in amnesic mild cognitive impairment patients. *Journal of Neuroscience*, vol. 34(32), páginas 10541–10553, 2014.
- LI, Y., WANG, S., TIAN, Q. y DING, X. A survey of recent advances in visual feature detection. *Neurocomputing*, vol. 149, páginas 736–751, 2015b.

- LIANG, L., HE, W.-P., LEI, L., ZHANG, W. y WANG, H.-X. Survey on enhancement methods for non-uniform illumination image [j]. *Application Research of Computers*, vol. 5, página 008, 2010.
- LIEDTKE, C., GAHM, T., KAPPEL, F. y AEIKENS, B. Segmentation of microscopic cell scenes. *Analytical and quantitative cytology and histology*, vol. 9(3), páginas 197–211, 1987.
- LIM, K.-L. y GALOOGAHI, H. K. Shape classification using local and global features. En *Image and Video Technology (PSIVT), 2010 Fourth Pacific-Rim Symposium on*, páginas 115–120. IEEE, 2010.
- LIPPONEN, P., AALTOOMAA, S., KOSMA, V.-M. y SYRJÄNEN, K. Apoptosis in breast cancer as related to histopathological characteristics and prognosis. *European Journal of Cancer*, vol. 30(14), páginas 2068–2073, 1994.
- LISIN, D. A., MATTAR, M. A., BLASCHKO, M. B., LEARNED-MILLER, E. G. y BENFIELD, M. C. Combining local and global image features for object class recognition. En *Computer vision and pattern recognition-workshops, 2005. CVPR workshops. IEEE Computer society conference on*, páginas 47–47. IEEE, 2005.
- LIU, D. y XIA, F. Assessing object-based classification: advantages and limitations. *Remote Sensing Letters*, vol. 1(4), páginas 187–194, 2010.
- LIU, M.-Y., TUZEL, O., RAMALINGAM, S. y CHELLAPPA, R. Entropy rate superpixel segmentation. En *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, páginas 2097–2104. IEEE, 2011.
- LIU, Y., TEVEROVSKIY, L., CARMICHAEL, O., KIKINIS, R., SHENTON, M., CARTER, C. S., STENGER, V. A., DAVIS, S., AIZENSTEIN, H., BECKER, J. T. ET AL. Discriminative mr image feature analysis for automatic schizophrenia and alzheimer's disease classification. En *International conference on medical image computing and computer-assisted intervention*, páginas 393–401. Springer, 2004.
- LIU, Y., ZHANG, J., TJONDRONEGORO, D. y GEVE, S. A shape ontology framework for bird classification. En *Digital Image Computing Techniques and Applications, 9th Biennial Conference of the Australian Pattern Recognition Society on*, páginas 478–484. IEEE, 2007.
- LONCARIC, S. A survey of shape analysis techniques. *Pattern recognition*, vol. 31(8), páginas 983–1001, 1998.
- LOUIS, A. Medical imaging: state of the art and future development. *Inverse Problems*, vol. 8(5), página 709, 1992.

- LUCCHI, A., SMITH, K., ACHANTA, R., KNOTT, G. y FUA, P. Supervoxel-based segmentation of mitochondria in em image stacks with learned shape features. *IEEE Trans. Med. Imaging*, vol. 31(2), páginas 474–486, 2012.
- LV, J. An approach for superpixels using uniform segmentation and reciprocal nearest neighbors clustering. *Journal of Theoretical & Applied Information Technology*, vol. 47(3), 2013.
- MAANI, R., YANG, Y. H. y KALRA, S. Voxel-based texture analysis of the brain. *PloS one*, vol. 10(3), página e0117759, 2015.
- MACENKO, M., NIETHAMMER, M., MARRON, J. S., BORLAND, D., WOOSLEY, J. T., GUAN, X., SCHMITT, C. y THOMAS, N. E. A method for normalizing histology slides for quantitative analysis. En *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*, páginas 1107–1110. IEEE, 2009.
- MAGNIN, B., MESROB, L., KINKINGNÉHUN, S., PÉLÉGRINI-ISSAC, M., COLLIOT, O., SARAZIN, M., DUBOIS, B., LEHÉRICY, S. y BENALI, H. Support vector machine-based classification of alzheimer's disease from whole-brain anatomical mri. *Neuroradiology*, vol. 51(2), páginas 73–83, 2009.
- MAIRAL, J., LEORDEANU, M., BACH, F., HEBERT, M. y PONCE, J. Discriminative sparse image models for class-specific edge detection and image interpretation. En *European Conference on Computer Vision*, páginas 43–56. Springer, 2008.
- MARCUS, D. S., WANG, T. H., PARKER, J., CSERNANSKY, J. G., MORRIS, J. C. y BUCKNER, R. L. Open access series of imaging studies (oasis): Cross-sectional mri data in young, middle aged, nondemented, and demented older adults. *J. Cognitive Neuroscience*, vol. 19(9), páginas 1498–1507, 2007. ISSN 0898-929X.
- MARDIA, K. V. Measures of multivariate skewness and kurtosis with applications. *Biometrika*, vol. 57(3), páginas 519–530, 1970.
- MARTIN, D., FOWLKES, C., TAL, D. y MALIK, J. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. En *Proc. 8th Int'l Conf. Computer Vision*, vol. 2, páginas 416–423. 2001.
- MARTÍNEZ, J. M., KOENEN, R. y PEREIRA, F. Mpeg-7: the generic multimedia content description standard, part 1. *IEEE multimedia*, vol. 9(2), páginas 78–87, 2002.

- MARTINEZ-MURCIA, F. J., GÓRRIZ, J. M., RAMÍREZ, J., SEGOVIA, F., SALAS-GONZALEZ, D., CASTILLO-BARNES, D., ILLÁN, I. A., ORTIZ, A., INITIATIVE, A. D. N. ET AL. Evaluating alzheimer's disease diagnosis using texture analysis. En *Annual Conference on Medical Image Understanding and Analysis*, páginas 470–481. Springer, 2017.
- MARTÍNEZ-USÓ, A., PLA, F., SOTOCA, J. M. y GARCÍA-SEVILLA, P. Clustering-based hyperspectral band selection using information measures. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45(12), páginas 4158–4171, 2007.
- MATAS, J. y KITTLER, J. Spatial and feature space clustering: Applications in image analysis. En *Proceedings CAIP'95, Computer Analysis of Images and Patterns*, páginas 162–173. Springer, 1995.
- MATHER, P. y TSO, B. *Classification methods for remotely sensed data*. CRC press, 2016.
- MERY, D. y BERTI, M. A. Automatic detection of welding defects using texture features. *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 45(10), páginas 676–681, 2003.
- MEYER, D., LEISCH, F. y HORNIK, K. The support vector machine under test. *Neurocomputing*, vol. 55(1-2), páginas 169–186, 2003.
- MICCAI GRAND CHALLENGE. Tumor proliferation assessment challenge (tupac16). <http://tupac.tue-image.nl>, ????
- MICHAELSEN, E. y ARENS, M. Hierarchical grouping using gestalt assessments. En *Computer Vision Workshop (ICCVW), 2017 IEEE International Conference on*, páginas 1702–1709. IEEE, 2017.
- MOGHADDAM, Z. y PICCARDI, M. Human action recognition with mpeg-7 descriptors and architectures. En *Proceedings of the first ACM international workshop on Analysis and retrieval of tracked events and motion in imagery streams*, páginas 63–68. ACM, 2010.
- MOHABEY, A. y RAY, A. Rough set theory based segmentation of color images. En *Fuzzy Information Processing Society, 2000. NAFIPS. 19th International Conference of the North American*, páginas 338–342. 2000.
- MOORE, A. P., PRINCE, S. J., WARRELL, J., MOHAMMED, U. y JONES, G. Superpixel lattices. En *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, páginas 1–8. IEEE, 2008.
- MORAN, M. S., INOUE, Y. y BARNES, E. Opportunities and limitations for image-based remote sensing in precision crop management. *Remote sensing of Environment*, vol. 61(3), páginas 319–346, 1997.

- MOSTAJABI, M., YADOLLAHPOUR, P. y SHAKHNAROVICH, G. Feedforward semantic segmentation with zoom-out features. En *Proceedings of the IEEE conference on computer vision and pattern recognition*, páginas 3376–3385. 2015.
- MOTWANI, M. C., GADIYA, M. C., MOTWANI, R. C. y HARRIS, F. C. Survey of image denoising techniques. En *Proceedings of GSPX*, páginas 27–30. 2004.
- MUSHRIF, M. M. y RAY, A. K. Color image segmentation: Rough-set theoretic approach. *Pattern Recognition Letters*, vol. 29(4), páginas 483 – 493, 2008. ISSN 0167-8655.
- MUSHRIF, M. M. y RAY, A. K. A-ifs histon based multithresholding algorithm for color image segmentation. *IEEE signal processing letters*, vol. 16(3), páginas 168–171, 2009.
- NELSON, R. C. y SELINGER, A. A cubist approach to object recognition. En *Computer Vision, 1998. Sixth International Conference on*, páginas 614–621. IEEE, 1998.
- NEUBERT, P. y PROTZEL, P. Superpixel benchmark and comparison. En *Proc. Forum Bildverarbeitung*, vol. 6. 2012.
- NEUBERT, P. y PROTZEL, P. Evaluating superpixels in video: Metrics beyond figure-ground segmentation. En *BMVC*. 2013.
- NOVARA, A., PISCIOTTA, A., MINACAPILLI, M., MALTESE, A., CAPODICI, F., CERDÀ, A. y GRISTINA, L. The impact of soil erosion on soil fertility and vine vigor. a multidisciplinary approach based on field, laboratory and remote sensing approaches. *Science of the Total Environment*, vol. 622, páginas 474–480, 2018.
- ORTIZ, C., GONZALO-MARTÍN, C., GARCÍA-PEDRERO, A. y MENASALVAS RUIZ, E. Supervoxels-based histon as a new alzheimer’s disease imaging biomarker. *Sensors*, vol. 18(6), página 1752, 2018.
- ORTIZ, C. A., GONZALO-MARTÍ, C., PENA, J. M. y MENASALVAS, E. 3d dendrite spine detection-a supervoxel based approach. En *International Conference on Rough Sets and Intelligent Systems Paradigms*, páginas 359–366. Springer, 2014.
- ORTIZ, C. A., GONZALO MARTÍN, C., GARCÍA PEDRERO, Á. y MENASALVAS RUIZ, E. Superpixel-based roughness measure for multispectral satellite image segmentation. *Remote Sensing*, vol. 7(11), páginas 14620–14645, 2015.

- ORTIZ, C. A., GONZALO MARTÍN, C., GARCÍA PEDRERO, Á., RODRÍGUEZ-ESPARRAGÓN, D. y MENASALVAS RUIZ, E. Clasificación de imágenes hiperespectrales mediante un índice de aproximación basado en supervóxeles. En *Teledetección. Nuevas plataformas y sensores aplicados a la gestión del agua, la agricultura y el medio ambiente*, páginas 451–454. Editorial Universitat Politècnica de València, 2017a.
- ORTIZ, C. A., MARTÍN, C. G., PEDRERO, A. G., GONZALEZ, A. R. y MENASALVAS, E. Mitosis detection in breast cancer using superpixels and ensemble classifiers. En *International Conference on Practical Applications of Computational Biology & Bioinformatics*, páginas 137–145. Springer, 2017b.
- PACLIK, P., DUIN, R. P., VAN KEMPEN, G. M. y KOHLUS, R. Segmentation of multi-spectral images using the combined classifier approach. *Image and Vision Computing*, vol. 21(6), páginas 473–482, 2003.
- PAL, N. R. y PAL, S. K. A review on image segmentation techniques. *Pattern Recognition*, vol. 26(9), páginas 1277 – 1294, 1993. ISSN 0031-3203.
- PALMER, S. E. *Vision science: Photons to phenomenology*. MIT press, 1999.
- PAN, X., ZHOU, Y., ZHANG, C. y LIU, Q. Flooding based superpixels generation with color, compactness and smoothness constraints. En *Image Processing (ICIP), 2014 IEEE International Conference on*, páginas 4432–4436. IEEE, 2014.
- PAPARI, G. y PETKOV, N. Edge and line oriented contour detection: State of the art. *Image and Vision Computing*, vol. 29(2-3), páginas 79–103, 2011.
- PARK, H., YANG, J.-J., SEO, J. y LEE, J.-M. Dimensionality reduced cortical features and their use in the classification of alzheimer’s disease and mild cognitive impairment. *Neuroscience letters*, vol. 529(2), páginas 123–127, 2012.
- PAWLAK, Z. Rough sets. *International Journal of Computer & Information Sciences*, vol. 11(5), páginas 341–356, 1982. ISSN 0091-7036.
- PAWLAK, Z. *Rough Sets: Theoretical Aspects of Reasoning About Data*. Kluwer Academic Publishers, Norwell, MA, USA, 1992. ISBN 0792314727.
- PAWLAK, Z. Vagueness and uncertainty: a rough set perspective. *Computational intelligence*, vol. 11(2), páginas 227–232, 1995.
- PAWLAK, Z., GRZYMALA-BUSSE, J., SLOWINSKI, R. y ZIARKO, W. Rough sets. *Commun. ACM*, vol. 38(11), páginas 88–95, 1995. ISSN 0001-0782.

- PAWLEY, J. *Handbook of biological confocal microscopy*. Springer Science & Business Media, 2010.
- PEARSON, K. Contributions to the mathematical theory of evolution. ii. skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London*, vol. 186(Part I), páginas 343–424, 1895.
- PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O., BLONDEL, M., PRETTENHOFER, P., WEISS, R., DUBOURG, V., VANDERPLAS, J., PASSOS, A., COUNAPEAU, D., BRUCHER, M., PERROT, M. y DUCHESNAY, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, vol. 12, páginas 2825–2830, 2011.
- PICARD, R., GRACZYK, C., MANN, S., WACHMAN, J., PICARD, L., CAMPBELL, L. y NEGROPONTE, N. Vision texture database. *the Media Laboratory, MIT, Cambridge, Massachusetts*, 1995.
- POWERS, D. M. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. 2011.
- PRATT, C. Principles of gestalt psychology. 1936.
- QUINLAN, J. R. Induction of decision trees. *Machine learning*, vol. 1(1), páginas 81–106, 1986.
- QUINLAN, J. R. *C4. 5: programs for machine learning*. Elsevier, 2014.
- RABINOVICH, A., VEDALDI, A. y BELONGIE, S. J. *Does image segmentation improve object categorization?*. Department of Computer Science and Engineering, University of California, San Diego, 2007.
- RAPSON, C. J., SEET, B.-C., NAEEM, M. A., LEE, J. E., AL-SARAYREH, M. y KLETTE, R. Reducing the pain: A novel tool for efficient ground-truth labelling in images. ????
- RASTOGI, V., PURI, N., ARORA, S., KAUR, G., YADAV, L. y SHARMA, R. Artefacts: a diagnostic dilemma—a review. *Journal of clinical and diagnostic research: JCDR*, vol. 7(10), página 2408, 2013.
- RATHORE, S., HABES, M., IFTIKHAR, M. A., SHACKLETT, A. y DAVATZIKOS, C. A review on neuroimaging-based classification studies and associated feature extraction methods for alzheimer’s disease and its prodromal stages. *NeuroImage*, vol. 155, páginas 530–548, 2017.
- REN, X. y MALIK, J. Learning a classification model for segmentation. En *null*, página 10. IEEE, 2003a.

- REN, X. y MALIK, J. Learning a classification model for segmentation. En *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, páginas 10–17 vol.1. 2003b.
- RICHTSFELD, A., ZILLICH, M. y VINCZE, M. Implementation of gestalt principles for object segmentation. En *Pattern Recognition (ICPR), 2012 21st International Conference on*, páginas 1330–1333. IEEE, 2012.
- RODARMEL, C. y SHAN, J. Principal component analysis for hyperspectral image classification. *Surveying and Land Information Science*, vol. 62(2), páginas 115–122, 2002.
- RODRIGUEZ, A., EHLENBERGER, D. B., DICKSTEIN, D. L., HOF, P. R. y WEARNE, S. L. Automated three-dimensional detection and shape classification of dendritic spines from fluorescence microscopy images. *PloS one*, vol. 3(4), página e1997, 2008.
- ROKACH, L. y MAIMON, O. Z. *Data mining with decision trees: theory and applications*, vol. 69. World scientific, 2008.
- ROSENFELD, A., HUMMEL, R. A. y ZUCKER, S. W. Scene labeling by relaxation operations. *IEEE Transactions on Systems, Man, and Cybernetics*, (6), páginas 420–433, 1976.
- SAAD, M. N., MUDA, Z., ASHAARI, N. S., HAMID, H. A. ET AL. The spatial relation features for describing objects relationships within image. En *Electrical Engineering and Informatics (ICEEI), 2015 International Conference on*, páginas 126–131. IEEE, 2015.
- SAAD, N. M., BAKAR, S. A., ABDULLAH, A. R., SALAHUDDIN, L., MUDA, S. y MOKJI, M. Brain lesion segmentation from diffusion-weighted mri based on adaptive thresholding and gray level co-occurrence matrix. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, vol. 3(2), páginas 1–13, 2011.
- SABOL, D. Interpretation of residual images: Spectral mixture analysis of aviris images, owens valley, california. En *Proc. 2nd AVIRIS Workshop, R. ????*
- SALAHAT, E. y QASAIMAH, M. Recent advances in features extraction and description algorithms: A comprehensive survey. En *Industrial Technology (ICIT), 2017 IEEE International Conference on*, páginas 1059–1063. IEEE, 2017.
- SALVATORE, C., CERASA, A., BATTISTA, P., GILARDI, M. C., QUATTRO-NE, A. y CASTIGLIONI, I. Magnetic resonance imaging biomarkers for the early diagnosis of alzheimer’s disease: a machine learning approach. *Frontiers in neuroscience*, vol. 9, página 307, 2015.

- SAMONDS, J. M., POTETZ, B. R. y LEE, T. S. Sample skewness as a statistical measurement of neuronal tuning sharpness. *Neural computation*, vol. 26(5), páginas 860–906, 2014.
- SAMPER-GONZALEZ, J., BURGOS, N., FONTANELLA, S., BERTIN, H., HABERT, M.-O., DURRLEMAN, S., EVGENIOU, T., COLLIOT, O., INITIATIVE, A. D. N. ET AL. Yet another admi machine learning paper? paving the way towards fully-reproducible research on classification of alzheimer's disease. En *International Workshop on Machine Learning in Medical Imaging*, páginas 53–60. Springer, 2017.
- SATHYA, B. y MANAVALAN, R. Image segmentation by clustering methods: performance analysis. *International Journal of Computer Applications*, vol. 29(11), 2011.
- SATHYA, R. y ABRAHAM, A. Comparison of supervised and unsupervised learning algorithms for pattern classification. *International Journal of Advanced Research in Artificial Intelligence*, vol. 2(2), páginas 34–38, 2013.
- SCHICK, A., FISCHER, M. y STIEFELHAGEN, R. Measuring and evaluating the compactness of superpixels. En *Pattern Recognition (ICPR), 2012 21st International Conference on*, páginas 930–934. IEEE, 2012.
- SCHOWENGERDT, R. A. *Remote sensing: models and methods for image processing*. Elsevier, 2006.
- SENTHILKUMARAN, N. y RAJESH, R. A study on rough set theory for medical image segmentation. *International journal of recent trends in Engineering*, vol. 2(2), páginas 236–238, 2009.
- SERMANET, P., EIGEN, D., ZHANG, X., MATHIEU, M., FERGUS, R. y LECUN, Y. Overfeat: Integrated recognition, localization and detection using convolutional networks. *arXiv preprint arXiv:1312.6229*, 2013.
- SERRANO-POZO, A., FROSCH, M. P., MASLIAH, E. y HYMAN, B. T. Neuropathological alterations in alzheimer disease. *Cold Spring Harbor perspectives in medicine*, vol. 1(1), página a006189, 2011.
- SETHIAN, J. A. A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences*, vol. 93(4), páginas 1591–1595, 1996.
- SHAH, C. A., ARORA, M. K. y VARSHNEY, P. K. Unsupervised classification of hyperspectral data: an ica mixture model based approach. *International Journal of Remote Sensing*, vol. 25(2), páginas 481–487, 2004.
- SHAPIRO, L. y STOCKMAN, G. C. Computer vision. 2001. ed: Prentice Hall, 2001.

- SHARMA, D. V., LEHAL, G. S. y MEHTA, S. Shape encoded post processing of gurmukhi ocr. En *2009 10th International Conference on Document Analysis and Recognition*, páginas 788–792. IEEE, 2009.
- SHARMA, G. y BALA, R. *Digital color imaging handbook*. CRC press, 2002.
- SHATTUCK, D. W., MIRZA, M., ADISETIYO, V., HOJATKASHANI, C., SALAMON, G., NARR, K. L., POLDRACK, R. A., BILDER, R. M. y TOGA, A. W. Construction of a 3d probabilistic atlas of human cortical structures. *Neuroimage*, vol. 39(3), páginas 1064–1080, 2008.
- SHEN, J., DU, Y., WANG, W. y LI, X. Lazy random walks for superpixel segmentation. *IEEE Transactions on Image Processing*, vol. 23(4), páginas 1451–1462, 2014.
- SHEN, L. y JIA, S. Three-dimensional gabor wavelets for pixel-based hyperspectral imagery classification. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49(12), páginas 5039–5046, 2011.
- SHYU, C.-R., BRODLEY, C., KAK, A., KOSAKA, A., AISEN, A. y BRODERICK, L. Local versus global features for content-based image retrieval. En *Content-Based Access of Image and Video Libraries, 1998. Proceedings. IEEE Workshop on*, páginas 30–34. IEEE, 1998.
- SILIN, D. y PATZEK, T. Pore space morphology analysis using maximal inscribed spheres. *Physica A: Statistical and Theoretical Physics*, vol. 371(2), páginas 336–360, 2006.
- SMEULDERS, A. W., WORRING, M., SANTINI, S., GUPTA, A. y JAIN, R. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (12), páginas 1349–1380, 2000.
- SMITH, M. O., JOHNSON, P. E. y ADAMS, J. B. Quantitative determination of mineral types and abundances from reflectance spectra using principal components analysis. *Journal of Geophysical Research: Solid Earth*, vol. 90(S02), 1985.
- SMITH, T. B. Mri artifacts and correction strategies. *Imaging in Medicine*, vol. 2(4), página 445, 2010.
- SOMMER, C., FIASCHI, L., HAMPRECHT, F. A. y GERLICH, D. W. Learning-based mitotic cell detection in histopathological images. En *Pattern Recognition (ICPR), 2012 21st International Conference on*, páginas 2306–2309. IEEE, 2012.
- SONG, Q., WANG, Y., YAN, X. y GU, H. Remote sensing images stripe noise removal by double sparse regulation and region separation. *Remote Sensing*, vol. 10(7), página 998, 2018.

- SONKA, M., HLAVAC, V. y BOYLE, R. *Image processing, analysis, and machine vision*. Cengage Learning, 2014.
- SØRENSEN, L., IGEL, C., LIV HANSEN, N., OSLER, M., LAURITZEN, M., ROSTRUP, E. y NIELSEN, M. Early detection of alzheimer's disease using mri hippocampal texture. *Human brain mapping*, vol. 37(3), páginas 1148–1161, 2016.
- SPANOS, A. *Probability theory and statistical inference: econometric modeling with observational data*. Cambridge University Press, 1999.
- STOKLASA, R. *Semantic-Segmentation of Biomedical Images based on Visual Similarity*. Tesis Doctoral, 2016.
- STUTZ, D., HERMANS, A. y LEIBE, B. Superpixels: an evaluation of the state-of-the-art. *Computer Vision and Image Understanding*, vol. 166, páginas 1–27, 2018.
- SUBBAN, R., SUSITHA, N. y MANKAME, D. P. Efficient iris recognition using haralick features based extraction and fuzzy particle swarm optimization. *Cluster Computing*, páginas 1–12, 2017.
- SUYKENS, J. A. y VANDEWALLE, J. Least squares support vector machine classifiers. *Neural processing letters*, vol. 9(3), páginas 293–300, 1999.
- SWAIN, M. J. y BALLARD, D. H. Color indexing. *International journal of computer vision*, vol. 7(1), páginas 11–32, 1991.
- TABATABAEI-JAFARI, H., SHAW, M. E. y CHERBUIN, N. Cerebral atrophy in mild cognitive impairment: A systematic review with meta-analysis. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, vol. 1(4), páginas 487–504, 2015.
- TAMURA, H., MORI, S. y YAMAWAKI, T. Textural features corresponding to visual perception. *IEEE Transactions on Systems, man, and cybernetics*, vol. 8(6), páginas 460–473, 1978.
- TANG, D., FU, H. y CAO, X. Topology preserved regular superpixel. En *Multimedia and Expo (ICME), 2012 IEEE International Conference on*, páginas 765–768. IEEE, 2012.
- TANG, X., HOLLAND, D., DALE, A. M., YOUNES, L. y MILLER, M. I. Baseline shape diffeomorphometry patterns of subcortical and ventricular structures in predicting conversion of mild cognitive impairment to alzheimer's disease. *Journal of Alzheimer's Disease*, vol. 44(2), páginas 599–611, 2015.

- TASKAR, B., ABBEEL, P. y KOLLER, D. Discriminative probabilistic models for relational data. En *Proceedings of the Eighteenth conference on Uncertainty in artificial intelligence*, páginas 485–492. Morgan Kaufmann Publishers Inc., 2002.
- TASKAR, B., SEGAL, E. y KOLLER, D. Probabilistic classification and clustering in relational data. En *International joint conference on artificial intelligence*, vol. 17, páginas 870–878. Lawrence Erlbaum Associates LTD, 2001.
- TEAGUE, M. R. Image analysis via the general theory of moments. *JOSA*, vol. 70(8), páginas 920–930, 1980.
- TEK, F. B. Mitosis detection using generic features and an ensemble of cascade adaboosts. *Journal of pathology informatics*, vol. 4, 2013.
- TERMENON, M. y GRAÑA, M. A two stage sequential ensemble applied to the classification of alzheimer’s disease based on mri features. *Neural Processing Letters*, vol. 35(1), páginas 1–12, 2012.
- TRAN, T., WEHRENS, R. y BUYDENS, L. Knn density-based clustering for high dimensional multispectral images. En *Remote Sensing and Data Fusion over Urban Areas, 2003. 2nd GRSS/ISPRS Joint Workshop on*, páginas 147–151. 2003.
- TUSHABE, F. y WILKINSON, M. Image preprocessing for compression: Attribute filtering. En *World Congress on Engineering and Computer Science*. Citeseer, 2007.
- TZOURIO-MAZOYER, N., LANDEAU, B., PAPATHANASSIOU, D., CRIVELLO, F., ETARD, O., DELCROIX, N., MAZOYER, B. y JOLIOT, M. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage*, vol. 15(1), páginas 273–289, 2002.
- UPADHYAYA, N. y DIXIT, M. A review: Relating low level features to high level semantics in cbir. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 9(3), páginas 433–444, 2016.
- VAROQUAUX, G., RAAMANA, P. R., ENGEMANN, D. A., HOYOS-IDROBO, A., SCHWARTZ, Y. y THIRION, B. Assessing and tuning brain decoders: cross-validation, caveats, and guidelines. *NeuroImage*, vol. 145, páginas 166–179, 2017.
- VEDALDI, A. y SOATTO, S. Quick shift and kernel methods for mode seeking. En *European Conference on Computer Vision*, páginas 705–718. Springer, 2008.

- VEKSLER, O., BOYKOV, Y. y MEHRANI, P. Superpixels and supervoxels in an energy optimization framework. En *European conference on Computer vision*, páginas 211–224. Springer, 2010.
- VEMURI, P., JONES, D. T. y JACK, C. R. Resting state functional mri in alzheimer's disease. *Alzheimer's research & therapy*, vol. 4(1), página 2, 2012.
- VERGÉS-LLAHÍ, J. y SANFELIU, A. Evaluation of distances between color image segmentations. En *Iberian Conference on Pattern Recognition and Image Analysis*, páginas 263–270. Springer, 2005.
- VETA, M., HENG, Y. J., STATHONIKOS, N., BEJNORDI, B. E., BECA, F., WOLLMANN, T., ROHR, K., SHAH, M. A., WANG, D., ROUSSON, M. ET AL. Predicting breast tumor proliferation from whole-slide images: the tupac16 challenge. *arXiv preprint arXiv:1807.08284*, 2018.
- VETA, M., PLUIM, J. P., VAN DIEST, P. J. y VIERGEVER, M. A. Breast cancer histopathology image analysis: A review. *IEEE Transactions on Biomedical Engineering*, vol. 61(5), páginas 1400–1411, 2014.
- VON HIPPEL, P. T. Mean, median, and skew: Correcting a textbook rule. *Journal of Statistics Education*, vol. 13(2), 2005.
- WANG, C., CHEN, J. y LI, W. Review on superpixel segmentation algorithms. *Application research of Computers*, vol. 31(1), páginas 6–12, 2014.
- WENG, Q., QUATTROCHI, D. y GAMBA, P. E. *Urban remote sensing*. CRC press, 2018.
- WESTFALL, P. H. Kurtosis as peakedness, 1905–2014. rip. *The American Statistician*, vol. 68(3), páginas 191–195, 2014.
- WOLD, S., ESBENSEN, K. y GELADI, P. Principal component analysis. *Chemometrics and intelligent laboratory systems*, vol. 2(1-3), páginas 37–52, 1987.
- WONG, D. F., ROSENBERG, P. B., ZHOU, Y., KUMAR, A., RAYMONT, V., RAVERT, H. T., DANNALS, R. F., NANDI, A., BRAŠIĆ, J. R., YE, W. ET AL. In vivo imaging of amyloid deposition in alzheimer's disease using the radioligand 18f-av-45 (flobetapir f 18). *Journal of Nuclear Medicine*, vol. 51(6), páginas 913–920, 2010.
- WU, C., DU, B. y ZHANG, L. Slow feature analysis for change detection in multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52(5), páginas 2858–2874, 2014.

- XIE, C.-H., LIU, Y.-J. y CHANG, J.-Y. Medical image segmentation using rough set and local polynomial regression. *Multimedia Tools and Applications*, vol. 74(6), páginas 1885–1914, 2015.
- XU, X., CHENG, J., WITT, R. M., SABATINI, B. L. y WONG, S. T. A shape analysis method to detect dendritic spine in 3d optical microscopy image. En *Biomedical Imaging: Nano to Macro, 2006. 3rd IEEE International Symposium on*, páginas 554–557. IEEE, 2006.
- YAN, J., CHU-SHERN, J. L., LOI, H. Y., KHOR, L. K., SINHA, A. K., QUEK, S. T., THAM, I. W. y TOWNSEND, D. Impact of image reconstruction settings on texture features in 18f-fdg pet. *Journal of nuclear medicine*, vol. 56(11), páginas 1667–1673, 2015.
- YANG, M., KPALMA, K. y RONSIN, J. A survey of shape feature extraction techniques. 2008.
- YAO, Y. Decision-theoretic rough set models. En *International Conference on Rough Sets and Knowledge Technology*, páginas 1–12. Springer, 2007.
- YASRAB, R. Ecru: An encoder-decoder based convolution neural network (cnn) for road-scene understanding. *Journal of Imaging*, vol. 4(10), página 116, 2018.
- YEUNG, A. W. An updated survey on statistical thresholding and sample size of fmri studies. *Frontiers in human neuroscience*, vol. 12, página 16, 2018.
- YOO, Y., IM, J. y PAIK, J. Low-light image enhancement using adaptive digital pixel binning. *Sensors*, vol. 15(7), páginas 14917–14931, 2015.
- YUAN, G.-X., HO, C.-H. y LIN, C.-J. Recent advances of large-scale linear classification. *Proceedings of the IEEE*, vol. 100(9), páginas 2584–2603, 2012.
- YUE, X., MIAO, D., ZHANG, N., CAO, L. y WU, Q. Multiscale roughness measure for color image segmentation. *Information Sciences*, vol. 216, páginas 93–112, 2012.
- YUGANDER, P., SHESHAGIRI, B., SUNANDA, K. y SUSMITHA, E. Multiple kernel fuzzy c-means algorithm with als method for satellite and medical image segmentation. En *Devices, Circuits and Systems (ICDCS), 2012 International Conference on*, páginas 244–248. 2012.
- ZHANG, H., FRITTS, J. E. y GOLDMAN, S. A. An entropy-based objective evaluation method for image segmentation. En *Storage and Retrieval Methods and Applications for Multimedia 2004*, vol. 5307, páginas 38–50. International Society for Optics and Photonics, 2003.

- ZHANG, H., FRITTS, J. E. y GOLDMAN, S. A. Image segmentation evaluation: A survey of unsupervised methods. *computer vision and image understanding*, vol. 110(2), páginas 260–280, 2008a.
- ZHANG, H., FRITTS, J. E. y GOLDMAN, S. A. Image segmentation evaluation: A survey of unsupervised methods. *Comput. Vis. Image Underst.*, vol. 110(2), páginas 260–280, 2008b. ISSN 1077-3142.
- ZHANG, J., KHALVATI, F., WONG, A. y HAIDER, M. A. Superpixel-based prostate cancer detection from diffusion magnetic resonance imaging. *Journal of Computational Vision and Imaging Systems*, vol. 1(1), 2015.
- ZHANG, R., ZHANG, L., WANG, X.-J. y GUAN, L. Multi-feature pls for combining visual features in image annotation. En *Proceedings of the 19th ACM international conference on Multimedia*, páginas 1513–1516. ACM, 2011a.
- ZHANG, S., LI, S., FU, W. y FANG, L. Multiscale superpixel-based sparse representation for hyperspectral image classification. *Remote Sensing*, vol. 9(2), página 139, 2017.
- ZHANG, Y. y GERBRANDS, J. J. Transition region determination based thresholding. *Pattern Recognition Letters*, vol. 12(1), páginas 13–23, 1991.
- ZHANG, Y. y GERBRANDS, J. J. Objective and quantitative segmentation evaluation and comparison. *Signal processing*, vol. 39(1-2), páginas 43–54, 1994.
- ZHANG, Y., HARTLEY, R., MASHFORD, J. y BURN, S. Superpixels via pseudo-boolean optimization. 2011b.
- ZHANG, Y., ZHOU, X., WITT, R. M., SABATINI, B. L., ADJEROH, D. y WONG, S. T. Automated spine detection using curvilinear structure detector and lda classifier. En *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on*, páginas 528–531. IEEE, 2007.
- ZHENG, Z., WEBB, G. I. y TING, K. M. Lazy bayesian rules: A lazy semi-naive bayesian learning technique competitive to boosting decision trees. En *In Proc. 16th International Conf. on Machine Learning*. Citeseer, 1999.
- ZHONG, Z., FAN, B., DUAN, J., WANG, L., DING, K., XIANG, S. y PAN, C. Discriminant tensor spectral-spatial feature extraction for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, vol. 12(5), páginas 1028–1032, 2015.

- ZHOU, W., LI, H. y ZHOU, X. 3d neuron dendritic spine detection and dendrite reconstruction. *International Journal of Computer Aided Engineering and Technology*, vol. 1(4), páginas 516–531, 2009.
- ZHOU, Y. y WEI, Y. Learning hierarchical spectral–spatial features for hyperspectral image classification. *IEEE transactions on cybernetics*, vol. 46(7), páginas 1667–1678, 2016.
- ZORMPAS-PETRIDIS, K., FAILMEZGER, H., ROXANIS, I., BLACKLEDGE, M., JAMIN, Y. y YUAN, Y. Capturing global spatial context for accurate cell classification in skin cancer histology. En *Computational Pathology and Ophthalmic Medical Image Analysis*, páginas 52–60. Springer, 2018.
- ZUIDERVELD, K. Contrast limited adaptive histogram equalization. *Graphics gems*, páginas 474–485, 1994.