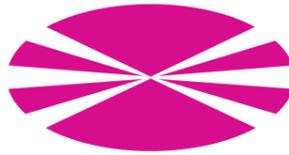


**UNIVERSIDADE DA CORUÑA**

FACULTADE DE INFORMÁTICA

Departamento de Tecnoloxías da Información e as Comunicaci3ns



OPTIMIZACIÓN PARAMÉTRICA DE TÉCNICAS AVANZADAS  
DE SEGMENTACIÓN DE IMÁGENES BIOMÉDICAS: UNA  
CONTEXTUALIZACIÓN A LA PATOLOGÍA.

**Tesis doctoral**

DOCTORANDO:

M<sup>a</sup> Carmen Bóveda Álvarez

DIRECTORES:

BERNARDINO ARCAY VARELA

ALFONSO CASTRO MARTÍNEZ

A Coruña, Julio de 2014



UNIVERSIDADE DA CORUÑA

FACULTADE DE INFORMÁTICA

DEPARTAMENTO DE TECNOLOXÍAS DA INFORMACIÓN E AS  
COMUNICACIÓNS

TESIS DOCTORAL

OPTIMIZACIÓN PARAMÉTRICA DE TÉCNICAS AVANZADAS  
DE SEGMENTACIÓN DE IMÁGENES BIOMÉDICAS: UNA  
CONTEXTUALIZACIÓN A LA PATOLOGÍA.

**DOCTORANDO:**

**M<sup>a</sup> Carmen Bóveda Álvarez**

**DIRECTORES:**

**BERNARDINO ARCAY VARELA  
ALFONSO CASTRO MARTÍNEZ**

**A Coruña, Julio de 2014**



D. **BERNARDINO ARCA Y VARELA**, Catedrático de Universidad en el Área de Ciencias de la Computación e Inteligencia Artificial de la *Universidade da Coruña* y D. **ALFONSO CASTRO MARTÍNEZ**, Profesor Contratado Doctor en el Área de Ciencias de la Computación e Inteligencia Artificial de la *Universidade da Coruña* HACEN CONSTAR QUE:

La memoria “**Optimización paramétrica de técnicas avanzadas de segmentación de imágenes biomédicas: una contextualización a la patología.**” ha sido realizada por D<sup>a</sup>. M<sup>a</sup> **CARMEN BÓVEDA ÁLVAREZ**, bajo nuestra dirección, en el *Departamento de Tecnoloxías da Información e as Comunicacións* de la *Universidade da Coruña*, y constituye la Tesis que presenta para optar al Grado de Doctor en Informática de la *Universidade da Coruña*.

A Coruña, Xuño de 2014

D. Bernardino Arcay Varela

D. Alfonso Castro Martínez



*A mi familia*



## **ABSTRACT**

Lung cancer produces high mortality susceptible of being reduced by 20% detecting carcinogenic nodules in the early stages using chest CT. If processing algorithms are unadjusted scans for detecting nodules may lose 17% of them.

We are looking for optimal parameterization of segmentation algorithms based on fuzzy clustering is, to delimit the constituent pixel image of a possible pulmonary nodule.

On the same set of prototype images of LIDC we have realised segmentations with various combinations of preconditions of algorithms, and defined and performance measures used methods based on empirical validation discrepancy subjected to descriptive statistics allow you to compare and get the suitable parameter values input to the problem. Showing that essentially the kind of solitary pulmonary nodule density written in the picture along with the parameter algorithms, number of clusters, are the most influential in the performance of all methods factors. Being optimal segment between eight and ten classes those images with node density tarnished and all other three classes to take to a single cluster node carcinogen.



## RESUMO

O cancro de pulmón produce unha alta mortalidade susceptible de diminuírse nun 20% ao detectar nos primeiros estadios posibles nodosidades canceríxenas empregando CT de peito. Se os algoritmos de procesado das tomografías non están axustados para a detección de nodosidades pódense perder un 17% delas.

Búscase unha parametrización óptima dos algoritmos de segmentación, baseados en clusterización difusa, para delimitar píxels da imaxe constituíntes dunha posible nodosidade pulmonar.

Sobre un mesmo conxunto prototipo de imaxes do LIDC realizáronse segmentacións con diversas combinacións de precondicións nos algoritmos, e definíronse e empregáronse índices de rendemento baseados en métodos de validación empírica por diverxencia que sometidos a unha estatística descritiva permiten comparar e obter os valores de parámetros de entrada axeitados ao problema.

Demostrándose que fundamentalmente o tipo de densidade da Nodosidade Pulmonar Senlleira contida nas imaxes xunto co parámetro dos algoritmos, número de clústers, son os factores máis influentes no rendemento de todos os métodos empregados. Sendo óptimo segmentar entre oito e dez clases aquelas imaxes con nodosidade de densidade deslustrada e en tres clases todas as demais para ter nun único clúster a nodosidade canceríxena.



## RESUMEN

El cáncer de pulmón produce una alta mortandad susceptible de reducirse en un 20% al detectar en los primeros estadios posibles nódulos cancerígenos utilizando CT de tórax. Si los algoritmos de procesamiento de las tomografías no están ajustados para la detección de nódulos se pueden perder un 17% de ellos.

Se busca una parametrización óptima de los algoritmos de segmentación, basados en clusterización difusa, para delimitar píxeles de la imagen constituyentes de un posible nódulo pulmonar.

Sobre un conjunto prototipo de imágenes del LIDC se han realizado segmentaciones con diversas combinaciones de precondiciones en los algoritmos, así como definido y utilizado índices de rendimiento basados en métodos de validación empírica por discrepancia que sometidos a una estadística descriptiva permiten comparar y obtener los valores de parámetros de entrada adecuados al problema.

Demostrándose que fundamentalmente el tipo de densidad del Nódulo Pulmonar Solitario contenido en las imágenes junto con el parámetro de los algoritmos, número de clústeres, son los factores más influyentes en el rendimiento de todos los métodos empleados. Siendo óptimo segmentar entre ocho y diez clases aquellas imágenes con nódulo de densidad deslustrada y en tres clases todas las demás, para tener en un único clúster al nódulo cancerígeno.



# ÍNDICE DE CONTENIDOS

<b>1</b>	<b><i>INTRODUCCIÓN</i></b> .....	<b>1</b>
1.1	Justificación del trabajo .....	1
1.2	Hipótesis y Objetivos.....	4
1.3	Organización de la memoria.....	5
<b>2</b>	<b><i>CONTEXTUALIZACIÓN</i></b> .....	<b>7</b>
2.1	Los pulmones y el cáncer.....	7
2.2	Imágenes.....	16
2.3	Algoritmos de procesado analizados. ....	31
2.4	Sistemas CAD para imágenes de Pulmón .....	46
<b>3</b>	<b><i>METODOLOGÍA</i></b> .....	<b>54</b>
3.1	Contexto de los estudios: imágenes de la Base de Datos LIDC .....	54
3.2	Conjuntos de prueba y validación.....	55
3.3	Pre-procesado pulmonar.....	58
3.4	Método empleado y combinaciones de valores paramétricos analizadas .....	64
3.5	Criterios utilizados para la clasificación de las pruebas .....	67
3.6	Índices de rendimiento de las pruebas .....	76
<b>4</b>	<b><i>RESULTADOS Y DISCUSIÓN</i></b> .....	<b>82</b>
4.1	Análisis de valores paramétricos .....	83
4.2	Validación de valores paramétricos .....	141
4.3	Validación cruzada.....	152
4.4	Discusión .....	156
<b>5</b>	<b><i>CONCLUSIONES Y FUTUROS DESARROLLOS</i></b> .....	<b>165</b>
5.1	Conclusiones .....	165
5.2	Propuestas .....	167
<b>6</b>	<b><i>BIBLIOGRAFÍA</i></b> .....	<b>168</b>



# 1 INTRODUCCIÓN

En este capítulo se indica cual es el propósito del trabajo y la justificación del mismo, destacando su magnitud e importancia; se plantean las hipótesis y especifican los objetivos de la tesis realizada. Finalizándolo con la descripción de la organización de esta memoria.

## *1.1 Justificación del trabajo*

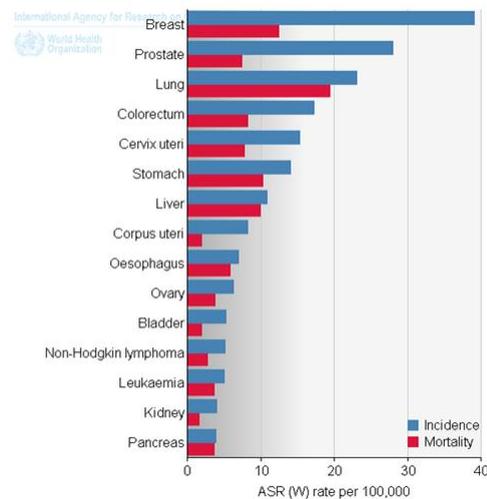
El Instituto Nacional del Cáncer (NCI) de EEUU difundió on line el 5 de noviembre de 2010 los resultados iniciales de un estudio -iniciado en agosto de 2002- sobre métodos de detección para reducir el número de muertes por cáncer de pulmón [NLST, 2010]. Detectaron cánceres en estadios relativamente iniciales utilizando la tomografía computarizada espiral de baja dosis en un grupo de enfermos y la radiografía convencional de pecho en otro grupo, informando que hubo un 20% menos de muertes por cáncer de pulmón entre los participantes en el estudio que habían sido evaluados con las tomografías axiales computarizadas de baja dosis.

El cáncer de pulmón es la causa principal de mortalidad por cáncer en todo el mundo. Ello justifica los esfuerzos para poder reducir la mortalidad como mínimo en ese 20% y así potencialmente salvar un número bastante considerable de personas de los estragos de esta enfermedad.

### 1.1.1 Importancia de la detección temprana del cáncer. Cifras

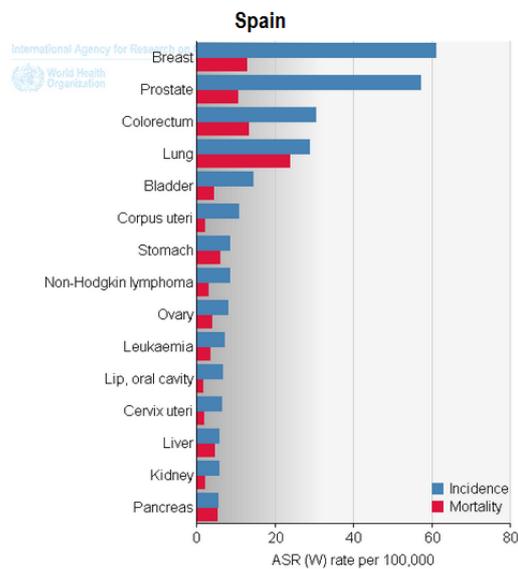
Analizando las gráficas del proyecto GLOBOCAN2008 de la International Agency for Research on Cancer (IARC) que forma parte de la OMS [GLOBOCAN2008, 2010], se aprecia que desde hace varias décadas, el cáncer de pulmón ha causado más muertes que cualquier otro tipo de cáncer, este hecho se refleja en las barras rojas de la Figura 1.

En el análisis de la incidencia y mortalidad, tanto en hombres como en mujeres de distintas edades de diversas regiones del mundo, destaca el mayor número de casos en hombres que en mujeres, siendo significativo que cada año, mueran más mujeres por cáncer de pulmón que por cáncer de mama, a pesar de ser este último de mucha mayor incidencia. Habiendo sido la mortalidad por cáncer de pulmón de 1.38 millones de personas, un 18.2% del total de los canceres en 2008, según las estadísticas oficiales conocidas en el momento de este acceso [GLOBOCAN2008, 2010].



**Figura 1. Cánceres más frecuentes en ambos sexos a nivel mundial.**

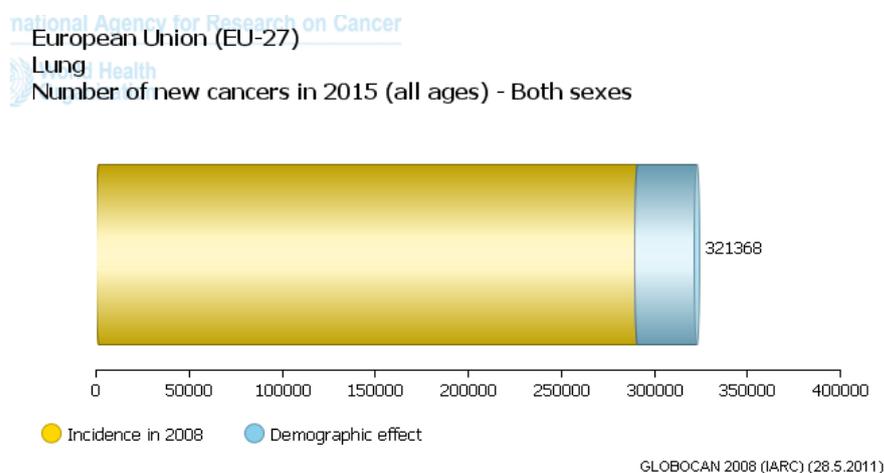
En España este tipo de cáncer es el que presenta mayores cifras de mortandad, aún siendo el cuarto en incidencia (véase Figura 2).



**Figura 2. Incidencia de cánceres en España.**

Según los estudios realizados, para España, por el instituto Carlos III de Madrid, el rasgo más patente es el ascenso de la mortalidad debida al cáncer de pulmón (1.75% cada año) que corresponde a los nuevos patrones de consumo de tabaco detectados en mujeres, y que afecta a las generaciones nacidas después de 1940, a pesar de que dicho cáncer en los varones mostró signos de estabilización durante los últimos 10 años [ISCIII, 2009].

En la Unión Europea, las predicciones de incidencias de nuevos cánceres de pulmón, para el año 2015, se indican en la Figura 3, para todas las edades y sexo.



**Figura 3. Predicción de nuevos cánceres de pulmón para 2015.**

El cáncer de pulmón es de los de peor pronóstico, ya que se suele diagnosticar cuando está en un estadio muy avanzado. Por ello para revertir su alta mortalidad y aumentar y mejorar la supervivencia de los pacientes es vital detectarlo precozmente.

Utilizando la Tomografía Axial Computarizada (TAC) se pueden detectar nódulos pulmonares incipientes en un tiempo reducido de exploración, pues esta técnica permite obtener imágenes de alta resolución espacial y con gran contraste de todo el tórax.

### 1.1.2 Trascendencia contextualizada del procesado de imagen

Los radiólogos y técnicos interpretan secuencias de imágenes del tórax obtenidas mediante la tomografía axial computarizada de baja dosis. El número de imágenes por cada estudio realizado es muy elevado, más de 300 imágenes en cada uno, y en cada imagen es difícil distinguir un nódulo en su fase inicial de otros objetos pulmonares; por ello es imprescindible disponer de herramientas que ayuden en el diagnóstico, lo cual justifica la necesidad de desarrollar y optimizar el uso de algoritmos para detectar nódulos pulmonares de forma automática, lo que implica procesar las imágenes de forma conveniente y adecuada.

Las imágenes médicas de estos estudios de TAC no sólo contienen la información de interés diagnóstico, sino que también contienen otros objetos y datos erróneos que las

distorsionan como es el ruido de carácter aleatorio, y artefactos debidos a imperfecciones en el dispositivo de adquisición. Por ello una vez adquirida la imagen, es necesario procesarla, bien aplicándole filtros, modificando el contraste, y/o realzando determinadas zonas de dicha imagen; todo ello con el objetivo de aumentar la visibilidad de estructuras u objetos que permitan mejorar la interpretación diagnóstica, y así contribuir a mejorar la detección de nódulos, y por lo tanto a disminuir el número de muertes por cáncer de pulmón.

## 1.2 Hipótesis y Objetivos

Este trabajo pretende contribuir a incrementar el grado de automatización en el proceso de diagnóstico precoz del cáncer pulmonar, maximizando la detección de los píxeles que representan al Nódulo Pulmonar Solitario en las imágenes de las Tomografías Axiales Computarizadas de tórax, utilizando para ello la técnica de segmentación de imágenes mediante algoritmos de clustering difuso. En concreto los algoritmos: Fuzzy C-Means, Adaptive Fuzzy C-Means, Spatially Kernelized Fuzzy C-Means, Fuzzy K-Nearest Neighbor, y como referencia el algoritmo CMeans.

Utilizando imágenes de la base de datos [LIDC, 2005] que tienen asociada una máscara informada por radiólogos en la que se indican los píxeles correspondientes a los puntos pertenecientes al nódulo y su borde, se han planteado las siguientes hipótesis de trabajo:

- *Hipótesis 1:* Un algoritmo es óptimo cuando maximiza el número de píxeles detectados dentro del tumor y minimiza el número de píxeles fuera del contorno del tumor marcado por los radiólogos en la máscara.
- *Hipótesis 2:* Minimizar un entorno de la máscara es equivalente a minimizar cualquier entorno mayor que ella.
- *Hipótesis 3:* El entorno a minimizar será aquel determinado por los expertos como dudoso y fuera del límite marcado - por ellos –como nódulo.
- *Hipótesis 4:* La perturbación de los valores de los parámetros de entrada a los algoritmos permite buscar los máximos puntos del tumor, y posteriormente elegir el que dé lugar a un valor mínimo de puntos fuera de él.

- *Hipótesis 5:* El mínimo de píxeles del clúster nódulo coincide con un número mínimo de píxeles para la zona marcada en la máscara.

Así, los objetivos establecidos han sido:

- Buscar cual de los algoritmos de segmentación, objeto del estudio empírico, es el más adecuado para la detección de los píxeles constituyentes del Nódulo Pulmonar Solitario en las imágenes de Tomografía Axial Computarizada.
- Encontrar para el tipo de imágenes “corpus del trabajo”, el número de clústeres adecuado tal que en uno de ellos esté perfectamente representado y aislado el tumor, independientemente del significado de los demás clústeres de la imagen.
- Analizar la influencia de los demás parámetros de entrada a los algoritmos para determinar en qué forma influyen en la partición y poder realizar una buena elección en los valores que se les asignarán.
- Elegir para cada algoritmo aquella combinación de valores de parámetros que permitan maximizar en la segmentación resultante los puntos de la imagen pertenecientes al nódulo pulmonar solitario (NPS) y minimizar los errores, incrementando así la precisión en la clasificación de los píxeles pertenecientes al clúster nódulo.

### *1.3 Organización de la memoria*

En este primer capítulo se ha justificado la importancia del trabajo realizado y se han planteado las hipótesis y objetivos, indicando que este estudio tiene como misión evaluar la precisión de las pruebas de segmentación realizadas con las imágenes de CT de tórax, y considerar adecuadas dichas pruebas si clasifican los píxeles correspondientes a los nódulos correctamente, de acuerdo a la máscara asociada elaborada y validada por expertos radiólogos.

El capítulo 2 presenta el contexto del trabajo, exponiendo en primer lugar una somera descripción de los pulmones, el cáncer de pulmón y concretando en lo que es el Nódulo

Pulmonar Solitario, continua describiendo la imagen medica, la tomografía computarizada, el estándar DICOM y la imagen digital y su procesamiento, para finalmente citar el estado del arte de parte de los trabajos científicos publicados para detectar, clasificar y diagnosticar nódulos pulmonares utilizando imágenes TAC de tórax.

El capítulo 3 presenta el método de trabajo que se ha aportado y utilizado en esta tesis. Comenzando por la descripción del corpus de las imágenes utilizadas en la parte de pruebas y en la parte de validación de valores de parámetros óptimos, sigue con la descripción del algoritmo de preprocesado empleado, explica la base de la que se parte para analizar y comparar los resultados de las segmentaciones, definiendo y describiendo los índices empleados para medir su rendimiento y determinar cuáles conjuntos de valores de parámetros son óptimos.

En el capítulo 4 se presentan los resultados obtenidos. Comenzando con un estudio del rendimiento de las pruebas efectuadas con los algoritmos ejecutados utilizando diversos valores de parámetros de entrada para la valoración de los resultados empíricos, y comprobación de las hipótesis planteadas. Se analiza el comportamiento paramétrico en cada algoritmo. Se presenta una estadística descriptiva de los resultados por cada uno de los algoritmos empleados, comparando los rendimientos de las diversas pruebas según la tipología del Nódulo Pulmonar Solitario a segmentar, como forma de obtención de los valores óptimos específicos; también se indaga en la sensibilidad y la especificidad de las pruebas de validación, observando las divergencias entre los resultados de test y de validación. Se completa este capítulo describiendo la validación cruzada realizada con los valores paramétricos de algoritmos e imágenes TAC objeto de estudio de otro investigador, con la finalidad de ver la capacidad de generalización de los resultados obtenidos en este trabajo. Y se completa el capítulo de resultados con la discusión final.

Por último, en el capítulo 5 se exponen las conclusiones principales obtenidas a lo largo de la realización de esta tesis doctoral y las propuestas para futuros desarrollos.

## 2 CONTEXTUALIZACIÓN

En este capítulo se describe el contexto del trabajo desarrollado, desde tres perspectivas: una divulgativa donde se indica brevemente que es el cáncer y su ubicación en los pulmones, centrándose en el Nódulo Pulmonar Solitario, otra técnica en la que se definen los conceptos de imagen digital, TAC, DICOM, y análisis de imágenes para desarrollar la descripción de los algoritmos empleados en esta tesis, y finalmente se relata que son los sistemas CAD y se citan trabajos científicos publicados de sistemas CAD para la detección de nódulos pulmonares.

### *2.1 Los pulmones y el cáncer*

Los pulmones son dos órganos esponjosos y ligeros, los más grandes del organismo, que se encuentran situados en la cavidad torácica. El pulmón derecho es mayor que el izquierdo porque el corazón se encuentra situado en ese lugar. El pulmón derecho se divide en tres porciones, llamadas lóbulos, y el izquierdo en dos. Los pulmones están recubiertos por una membrana, llamada pleura, que los protege y les da elasticidad para facilitar sus movimientos. Hay diferentes estructuras que entran y salen de los pulmones y que realizan distintas funciones. Mediante el movimiento de fuelle del tórax, los pulmones aspiran el aire obteniendo el oxígeno y expulsando de su interior el dióxido de carbono. El aire inspirado pasa por la laringe y la tráquea desde donde se distribuye a través de los bronquios a los pulmones. La tráquea, que lleva el oxígeno procedente de la boca y las fosas nasales, a la altura de los pulmones se divide en los bronquios; estos se ramifican en múltiples conductos, llamados bronquiolos. Los bronquiolos poseen unas terminaciones minúsculas denominadas alvéolos. Estos son unas pequeñas cavidades que están en contacto con los vasos sanguíneos que irrigan los pulmones. Las venas y las arterias se encargan de transportar la sangre desde los pulmones al organismo y desde éste a los pulmones. La sangre llega hasta los alvéolos y allí es donde se intercambian los gases, dejando el dióxido de carbono en los pulmones para ser expulsado y el oxígeno en la sangre para que sea transportado al resto del organismo [Elmundosalud, 2009].

En la Figura 4 se muestra la ubicación y estructura de los pulmones, indicándose también la disposición y la separación de los distintos lóbulos en cada pulmón.

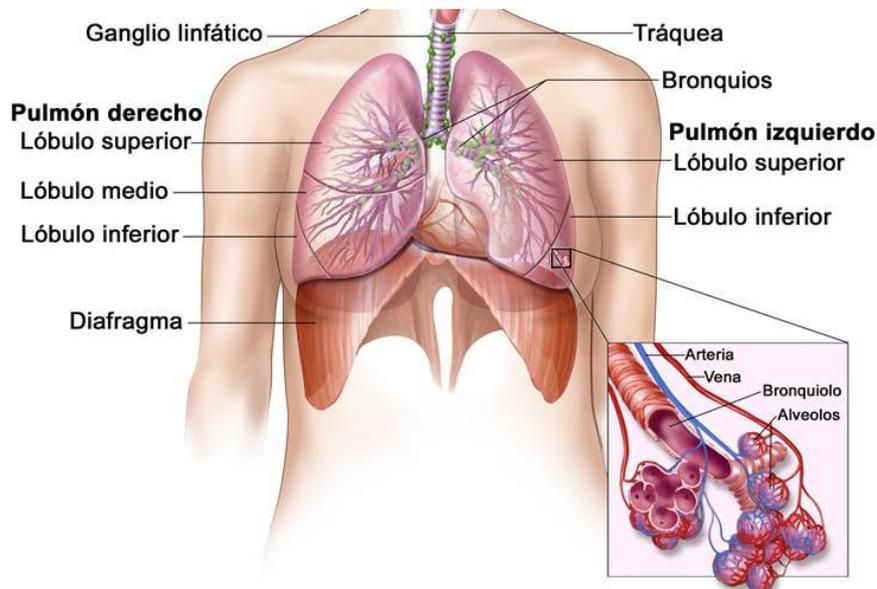


Figura 4. Ubicación y estructura de los pulmones.

El carcinoma o cáncer de pulmón se forma en los tejidos del pulmón, por lo general en las células que recubren los conductos de aire, habitualmente en las paredes internas de los bronquios, y al crecer puede obstruir el paso del aire y alterar la respiración. Por este motivo, produce generalmente falta de aire, ahogo y fatiga.

La palabra carcinoma es el término empleado para nombrar enfermedades que tienen en común la transformación de una célula normal en otra que se comporta de forma muy peligrosa para el cuerpo humano.

El cuerpo humano está constituido por células que se dividen periódicamente y de forma regular para reemplazar a células ya envejecidas o muertas, y mantener así la integridad y el funcionamiento correcto de los distintos órganos que lo forman.

El proceso de división de las células está regulado por unos mecanismos de control que indican a la célula cuándo comenzar a dividirse y cuándo permanecer estática. En una célula, con dichos mecanismos de control alterados, se puede producir una división incontrolada de la misma que con el tiempo degenera en un tumor o nódulo [AECC, 2011].

Si las células que forman dicho tumor no poseen la capacidad de invadir y destruir otros órganos se trata de un tumor benigno. Pero si estas células además de crecer sin control sufren nuevas alteraciones y adquieren la facultad de infiltrarse invadiendo tejidos y órganos a su alrededor, y se trasladan a otras partes del organismo proliferando en ellas entonces se produce un tumor maligno ó cáncer.

Dependiendo del tipo de tejido en el que surge se distinguen los términos:

- **Carcinoma:** originado en tejidos que cubren la superficie interna o externa del cuerpo como el seno, pulmón, colon, vejiga, próstata.
- **Sarcoma:** tiene lugar en tejidos que sostienen el cuerpo como los huesos, los músculos o el tejido conectivo que rodea órganos.
- **Linfoma:** cuando ocurre en los ganglios linfáticos, donde maduran las células del sistema inmunológico.
- **Leucemia:** originado en células inmaduras de la sangre, que se generan en la médula ósea, y que se acumulan en el torrente sanguíneo.

Cuando, en concreto, las células afectadas son las de los pulmones se le denomina carcinoma de pulmón.

### 2.1.1 Carcinoma de pulmón: Nódulo Pulmonar Solitario

Se denomina así al crecimiento anormal de las células del pulmón. El 90% de los carcinomas de pulmón son de dos tipos: broncogénico o de células no pequeñas y microcítico o de células pequeñas. El 10% restante está compuesto por carcinoides o tumores neuroendocrinos [SEOM, 2010].

El carcinoma de pulmón de células pequeñas o microcítico recibe este nombre por el tamaño de las células vistas con el microscopio. Se estima que alrededor del 20% de todos los carcinomas son microcíticos. Estas células pequeñas se multiplican rápidamente y se pueden formar grandes tumores; teniendo además mayor capacidad para extenderse a otros órganos. El tumor primario se produce, generalmente, cerca de los bronquios expandiéndose hacia la zona central de los pulmones. Mientras que el de células no pequeñas representa prácticamente el 80% restante de los carcinomas de pulmón y se extiende más lentamente que el de células pequeñas [Libby, 2004].

Las variedades de carcinoma no-microcítico de pulmón son el carcinoma escamoso o epidermoide y el adenocarcinoma. El carcinoma escamoso es el más frecuente, representando el 40% de los carcinomas no microcíticos. Suele localizarse en la parte central de los pulmones, y con frecuencia se necrosa en su interior. Tiene un crecimiento relativamente lento. El adenocarcinoma representa el 30% de los carcinomas no microcíticos. Suele localizarse en zonas más periféricas de los pulmones, por lo que frecuentemente afecta a la pleura y pared torácica.

Al clasificar un tumor hay que determinar en qué fase se encuentra, para conocer su extensión y gravedad. Para ello se utiliza, mayoritariamente, el sistema TNM. Estas siglas hacen referencia a tres aspectos del cáncer: la T se refiere al tamaño y localización del

tumor, la N a la afectación de los ganglios linfáticos y la M a la afectación o no de otros órganos.

A continuación se muestra las referencias utilizadas para indicar el tamaño y la localización:

- T0: no hay signos de un tumor primario.
- Tis: Carcinoma in situ que no se ha diseminado a tejidos vecinos
- T1: tumor menor o igual a 3 cm de diámetro máximo, rodeado por tejido pulmonar o pleura visceral.
- T2: tumor mayor de 3 centímetros o tumor de cualquier tamaño que invade la pleura visceral.
- T3: tumor de cualquier tamaño con extensión directa a la pared costal (incluidos los tumores de la cisura superior), diafragma, pleura mediastínica o pericardio; sin afectación del corazón, grandes vasos, tráquea, esófago, cuerpos vertebrales; o un tumor del bronquio principal.
- T4: tumor de cualquier tamaño con infiltración del mediastino o del corazón, grandes vasos, tráquea, esófago, cuerpos vertebrales o con derrame pleural maligno.

En la parte izquierda de la Figura 5 se muestran estos estadios del carcinoma de pulmón. El estadio Tis indica que se encuentra todavía concentrado en su lugar original y aún no se ha esparcido a otros órganos. Los Estadio I, Estadio II y Estadio III indican que cuanto más alto es su número más se ha extendido, ya sea porque el tumor ha crecido o porque ha invadido órganos vecinos. Es en el Estadio IV cuando se ha esparcido haciendo metástasis en otros órganos, es decir, ha formado nuevos tumores en otras partes del cuerpo.

Esta información es sumamente relevante para el oncólogo, pues le permitirá pronosticar cómo puede seguir avanzando el carcinoma y así definir qué tipo de tratamiento debe aplicarle al paciente, con qué urgencia lo necesita, y cuáles son las posibilidades de recuperación.

Para diagnosticar el carcinoma de pulmón se utilizan diversas técnicas: pruebas analíticas (analizan los componentes de sangre, orina...), pruebas de imagen (en las que se obtienen imágenes del interior del cuerpo) y estudio de tejidos (mediante biopsia o citología).

Las pruebas de imagen son importantes a la hora de determinar la localización, tamaño y extensión de la enfermedad, y se emplean para ello la radiografía y la Tomografía Computarizada (TC).

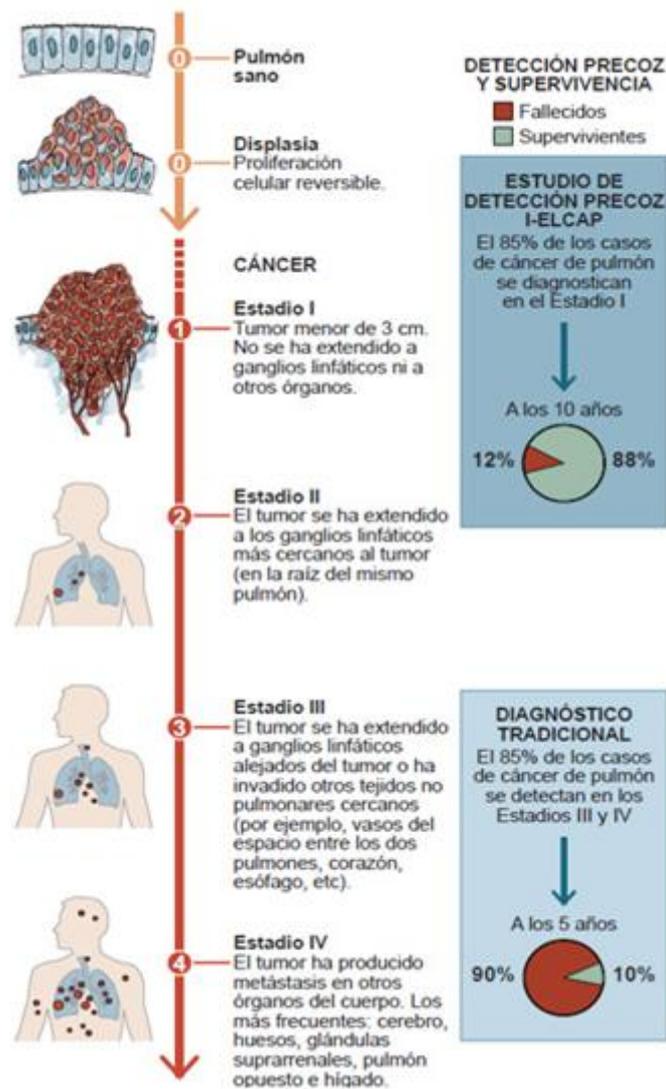


Figura 5. Fases en el carcinoma de pulmón [ELCAP, 2010].

En esta tesis se ha trabajado con imágenes de Tomografías Computarizadas, en las que el tumor se corresponde con el Estadio I; dicho tumor es conocido como Nódulo Pulmonar Solitario (NPS). Tal como se indica gráficamente en la parte derecha de la Figura 5, es en esa fase (Estadio I) en la que tiene especial importancia detectar de manera precoz el cáncer de pulmón, pues de ello depende que la supervivencia del paciente pueda llegar a ser de un 88%.

### 2.1.1.1 Definición y clasificación

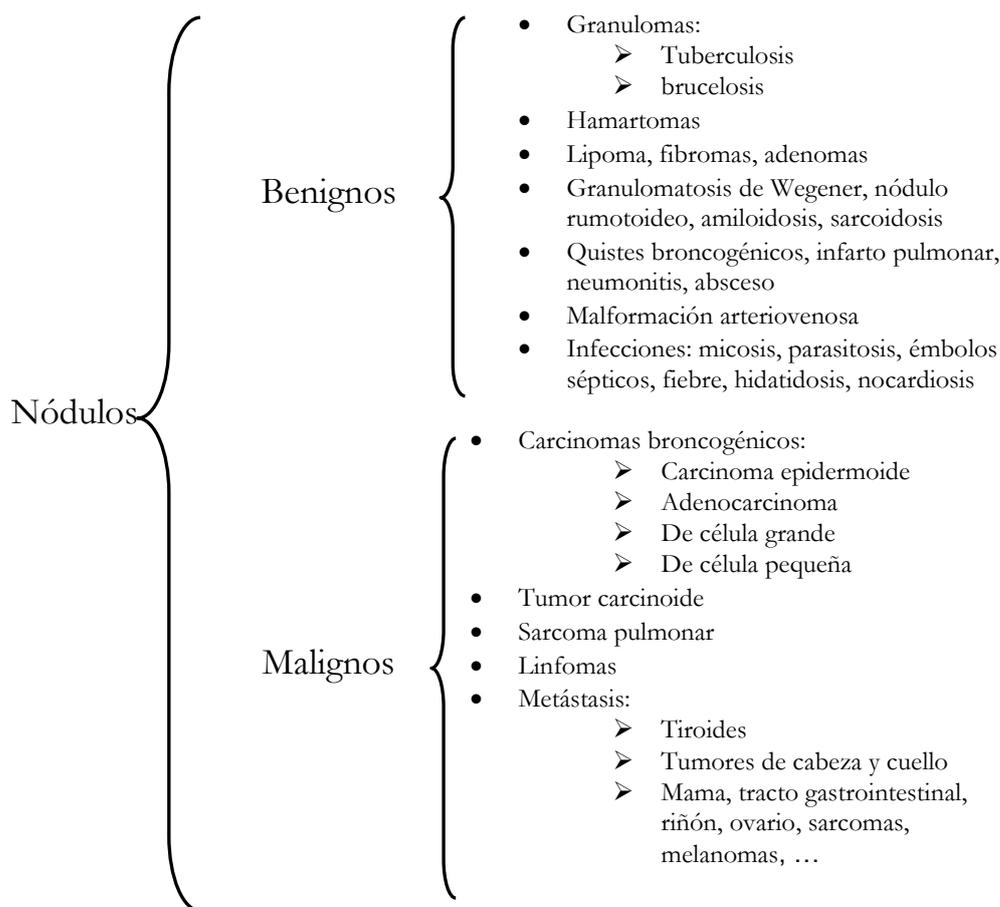
Se denomina Nódulo Pulmonar Solitario (NPS) a la lesión de forma ovalada o redondeada cuyo diámetro es menor o igual a 3 cm (si supera los 3 cm ó los 6 cm según

diversos autores, se designa como masa), de localización intrapulmonar y que se encuentra rodeada por tejido pulmonar normal.

El 60 % de todos los NPS son benignos, aunque en España algunos autores consideran que el porcentaje de NPS malignos es mayor [Estévez, 2001]. Y todos los tipos histológicos de carcinoma de pulmón pueden presentarse como NPS [Gabrielli M., 2007].

### 2.1.1.2 Etiología

Las causas más frecuentes de aparición del Nódulo Pulmonar Solitario se describen en el Esquema 1. Los granulomas son la causa más frecuente de los NPS benignos, y en el caso de los nódulos malignos es el carcinoma broncogénico el más frecuente. La probabilidad de que un NPS sea benigno o maligno depende de muchos factores clínicos y radiográficos.



Esquema 1. Causas más frecuentes de NPS.

Entre un 40-50% resultan ser malignos, y de ellos, el 75% corresponden a carcinomas broncogénicos. El 50-60% restante de los NPS suelen ser lesiones benignas, de las cuales el 80% son procesos inflamatorios, granulomas asociados a tuberculosis o enfermedades micóticas [Gabrielli, 2007].

Al evaluar a un paciente que tiene un NPS, antes de caracterizar dicho nódulo se comprueba si existen factores clínicos que indiquen la posibilidad de un carcinoma pulmonar subyacente. Un NPS es poco probable que sea de origen metastásico, en ausencia de un cáncer primario conocido.

### 2.1.1.3 Factores de riesgo

En la Tabla 1 se resume la valoración del riesgo de carcinoma en pacientes con Nódulo Pulmonar Solitario. La incidencia se incrementa progresivamente a partir de los 40 y hasta los 80 años de edad, aumenta en pacientes fumadores siendo directamente proporcional a la cantidad de cigarrillos que consuman diariamente. También aumenta el riesgo si el paciente tiene historia de cáncer extrapulmonar, o bien fibrosis pulmonar. Además se determinan otros factores clínicos tales como la exposición laboral a agentes químicos por ejemplo el asbesto, e historial de viajes por zonas donde la tuberculosis e infecciones micóticas son endémicas.

Alrededor del año 1920 ya se observó en la población, un gran incremento de fallecimientos por carcinoma de pulmón. Existiendo trabajos publicados por los autores Lombard y Doering en 1928 y Müller en 1939, sobre la posible relación entre el hábito de fumar y dicho carcinoma, pero esta cuestión suscitó verdadero interés a partir de 1950 cuando comenzó la abrumadora evidencia de que el hábito de fumar es una causa muy importante en el aumento de la incidencia del carcinoma pulmonar [ASEH, 2010].

<b>VARIABLE</b>	<b>RIESGO DE CARCINOMA</b>		
	<i>Bajo</i>	<i>Medio</i>	<i>Elevado</i>
<b><i>Diámetro del nódulo</i></b>	Menos de 1,5 cm	De 1,5 a 2,2 cm	Mayor de 2,2 cm
<b><i>Edad del paciente (años)</i></b>	Menor de 45	Entre 45-60	Mayor de 60
<b><i>Categoría de fumador</i></b>	Nunca ha fumado.	Menos de 20/día	Más de 20/día
<b><i>Tiempo de cese de fumar</i></b>	Más de 7 años	Menos de 7 años	Nunca lo ha dejado
<b><i>Márgenes NPS</i></b>	Lisos	Festonados	Corona radiada o espiculados

Tabla 1. Valoración del riesgo de carcinoma en pacientes con NPS.

### 2.1.1.4 Caracterización

El carcinoma pulmonar es 1,5 veces más probable que ocurra en el pulmón derecho que en el izquierdo, cerca del 70% se localizan en los lóbulos superiores. Sin embargo, los nódulos benignos se distribuyen igualmente a través de los lóbulos superiores e inferiores [Gabielli, 2007].

Un NPS raramente se hace evidente en una radiografía de tórax si mide menos de 9 mm de diámetro, sin embargo en una Tomografía Computarizada (CT) se ven nódulos de tamaños menores; lo cual permite que para la detección precoz del cáncer sea éste último el método empleado.

El tamaño del NPS no es un predictor fidedigno de benignidad, pero los nódulos de mayor tamaño (cercano a los 3 cm de diámetro) tienen una mayor probabilidad de ser malignos. Más del 90% de los nódulos menores de 2 cm son benignos. El estudio de screening realizado por el Early Lung Cancer Action Project mostró que sólo el 8% de las lesiones menores de 1cm de diámetro fueron malignas [Winer\_Muram, 2006].

Existen dos patrones, véase Figura 6, que son relativamente específicos para la detección de carcinoma: uno de ellos es el “*Signo de Corona Radiada*” que consiste en una extensión lineal muy fina de 4-5 mm hacia fuera del nódulo descrito en imágenes de TC, que se homologa con la “*aparición espiculada*” descrito en la radiografía de tórax. La presencia de bordes festoneados se asocia con una probabilidad intermedia de presentar carcinoma, mientras que los bordes lisos y bien definidos son sugerentes de enfermedad benigna.

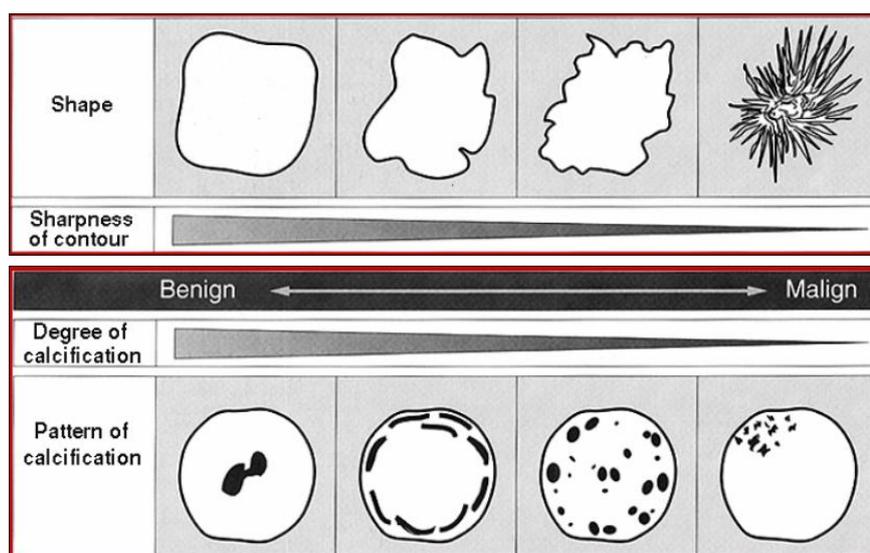


Figura 6. Patrones de calcificación y contorno del NPS [FMV, 2011].

Los patrones de calcificación de un NPS, mostrados en la Figura 6 , que sugieren benignidad son: calcificación central (típico de un granuloma), concéntrica (anillo de calcificación), laminar, y calcificación en “popcorn” (frecuentemente visto en los hamartomas). Los patrones de calcificación punteados o excéntricos, distróficos, irregulares o asimétricos se asocian más frecuentemente con la presencia de carcinoma [Yeong, 2007].

El patrón de crecimiento puede estimarse cuando se tienen imágenes previas. El tiempo en que los tumores broncogénicos malignos doblan su volumen rara vez es menor a un mes o más de un año, sin embargo, algunos tumores pulmonares como el adenocarcinoma y el tumor carcinoide pueden tener tiempos de duplicación mayores de dos años [Winer\_Muram, 2006] [Yeong, 2007].

Tradicionalmente en las radiografías de tórax, las lesiones se han considerado benignas cuando permanecen estables durante dos años. Sin embargo, los carcinomas broncoalveolares y carcinoides pueden permanecer ocasionalmente estables durante dos o más años.

El uso de la estabilidad como indicador de lesiones benignas es útil como método de valoración cuando se cuenta con imágenes de alta resolución que permiten mediciones exactas del crecimiento del nódulo. La Tomografía Computarizada de alta resolución permite mejorar la estimación del tamaño nodular y de sus características de crecimiento, pues permite distinguir cambios de tamaño de 0,3 mm [Yeong, 2007].

### **2.1.1.5 Protocolos de diagnóstico y tratamiento**

A lo largo del tiempo se han propuesto diversas estrategias diagnóstico-terapéuticas para abordar el NPS. Algunos autores proponen esperar y observar el crecimiento del nódulo con controles radiológicos en pacientes con bajo riesgo de malignidad. Otros señalan que se debe realizar una biopsia para determinar si el NPS es candidato a cirugía. Por último, varios grupos aconsejan la cirugía como primera opción, a menos que haya una fuerte sospecha de benignidad.

Para poder establecer la mejor estrategia diagnóstico-terapéutica, costo-efectiva para un NPS, primero debe conocerse la probabilidad de malignidad de la lesión. Los datos concernientes con los factores de riesgo que permiten definir dicha probabilidad de malignidad han sido revisados y corroborados en varios estudios [Gabrielli, 2007] y, resumidos en la Tabla 1 página 13.

Para finalizar este apartado se muestran, a continuación en la Figura 7, los protocolos de diagnóstico utilizados actualmente a ambos lados del Atlántico, en los cuales queda claramente especificada la utilización de la tomografía computacional como técnica diagnóstica.

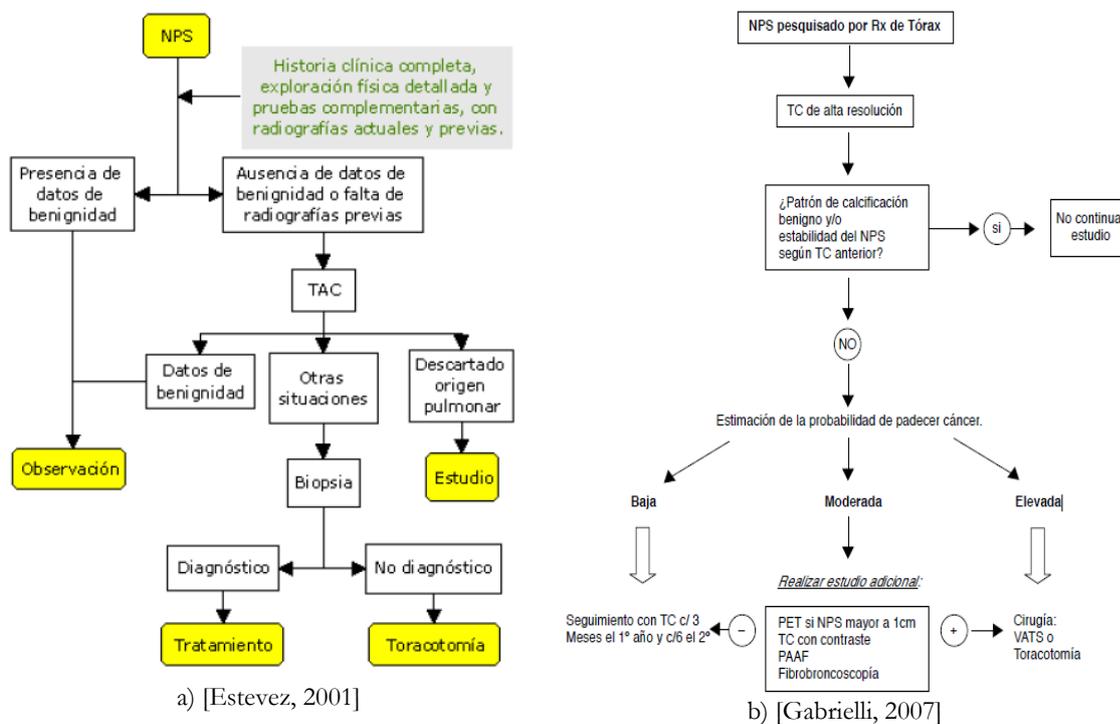


Figura 7. Protocolos de diagnóstico.

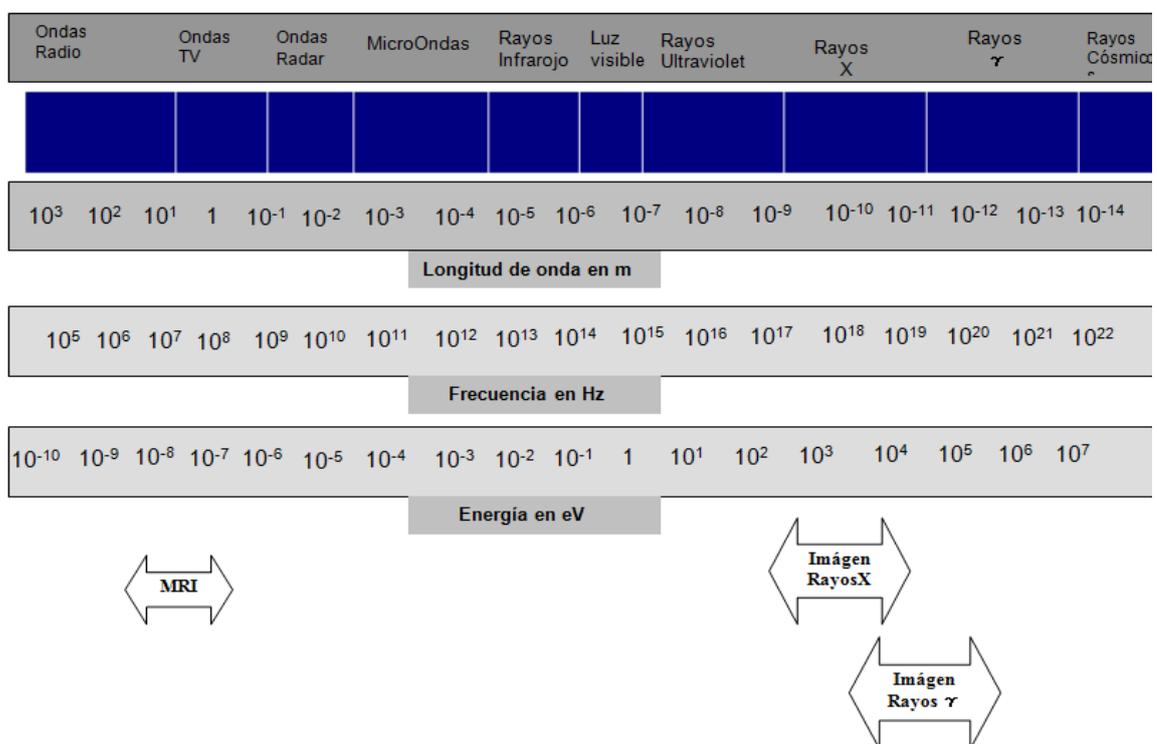
## 2.2 Imágenes

En este apartado se describe de forma somera la imagen médica, concretando que es la Tomografía Axial Computarizada y se detalla el estándar DICOM utilizado; además se especifica la segmentación como operación de procesamiento de la imagen digital y el empleo de algoritmos de clústering y lógica difusa en ella.

### 2.2.1 La imagen médica

Se define imagen médica como un conjunto de técnicas y procesos utilizados para crear imágenes del cuerpo humano, o partes de él, sin tener que abrirlo. Se utiliza con propósitos clínicos para diagnosticar o examinar enfermedades y para el estudio de la anatomía de un cuerpo.

Actualmente la imagen médica, en su modalidad digital, se ha convertido en un instrumento fundamental en la práctica clínica porque permite detectar y tratar patologías. El objetivo general de la imagen médica es adquirir información útil sobre los procesos fisiológicos y órganos del cuerpo humano utilizando fuentes de energías internas o externas. En la Figura 8 se realiza una clasificación de modalidades de imágenes médicas según la fuente de energía empleada para obtenerla. La naturaleza de la energía utilizada para irradiar a un paciente es la que determina el nombre de la modalidad de la imagen médica obtenida.



**Figura 8. Clasificación modalidades de Imagen Médica según Fuente de Energía.**

Las radiografías y la tomografía computarizada utilizan rayos X, la medicina nuclear (PET, SPECT) utiliza rayos  $\gamma$ , la resonancia magnética utiliza ondas de radio (radiación electromagnética), y la ecografía utiliza energía ultrasónica. Según la interacción de esta energía con los tejidos del cuerpo se distinguen radiaciones ionizantes, las cuales inducen radiaciones químicas, y radiaciones no ionizantes en las que únicamente se produce calentamiento.

Cuando se clasifican las imágenes médicas utilizando como criterio la capacidad de separar objetos a diferentes profundidades, se diferencia entre imágenes proyectivas en las que se produce una superposición sobre un plano, e imágenes tomográficas que proporcionan varios cortes. En la Tabla 2 se muestran las distintas modalidades de imagen médica junto a los valores que caracterizan a cada una de ellas, y en la Figura 9 se ponen ejemplos de las imágenes de éstas modalidades.

MODALIDAD	TÉCNICA IMAGEN	TOMOGRÁFICA	ENERGÍA	IONIZANTE	FUNCIONAL
Radiología	Convencional	No	Rayos X	Si	No
	Digital	No	Rayos X	Si	No
	<b>TAC</b>	<b>SI</b>	<b>RAYOS X</b>	<b>SI</b>	<b>NO</b>
Resonancia Magnética	fMRI	Si	Ondas Radio	No	Si
	MRI	Si	Ondas Radio	No	No
Medicina Nuclear	PET	Si	Rayos $\gamma$	Si	Si
Nuclear	SPECT	Si	Rayos $\gamma$	Si	Si
Ecografía	Ecografía	Si	Ultrasonido	No	No
Endoscopia	Endoscopia	No	Luz	No	No

Tabla 2. Características de las modalidades de imagen médica.

En las imágenes médicas es difícil generar información válida para ser procesada, pues existe gran cantidad de ruido y una enorme variabilidad. Hay dos parámetros importantes a considerar en ellas: el contraste, lo que se ve en la imagen, y la resolución, es decir, el grado de detalle espacial y temporal. Por ello se hace necesario utilizar programas de diagnóstico asistido por computador, conocidos como sistemas CAD; descritos en el apartado 2.4 como sistemas que pretenden ayudar al radiólogo a la hora de localizar zonas sospechosas de la imagen.

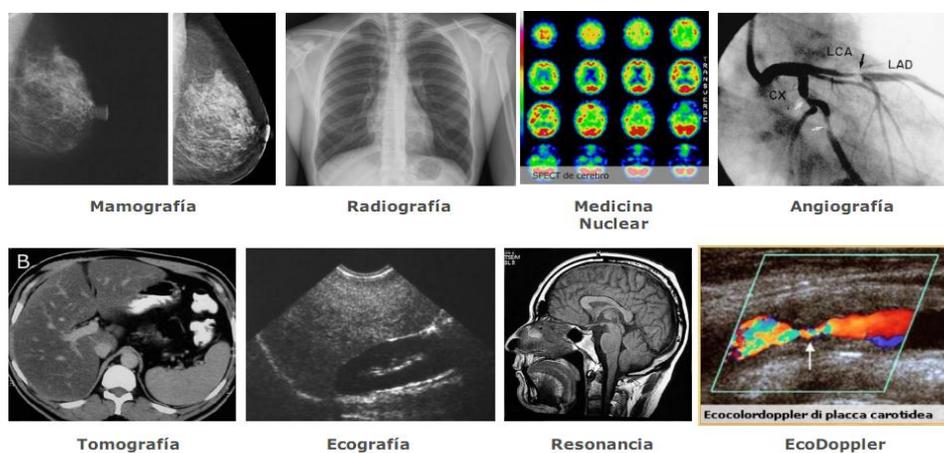


Figura 9. Diversas modalidades de imágenes médicas [SAIS, 2011].

Las imágenes objeto de esta tesis son las Tomografías Axiales Computarizadas, denotadas con las siglas TAC o TC. La técnica utilizada en ellas es una opción no invasiva de diagnóstico por imagen, de gran utilidad para la exploración prácticamente de todo el cuerpo, aunque principalmente se utiliza en las áreas de Tórax, Abdomen, Sistema cardiovascular y Sistema musculoesquelético. En concreto, en este trabajo se utilizan las TAC de Tórax.

### 2.2.2 Tomografía Computarizada

La palabra “tomografía” proviene del griego: *tomos* significa corte o sección, y *grafía* deriva de *grafein* cuyo significado es grabar o escribir.

La Tomografía Axial Computarizada, conocida por la sigla TAC, es una modalidad de imagen que obtiene cortes axiales del cuerpo; para ello se mide la atenuación de un haz de rayos X muy estrecho y que atraviesa el cuerpo desde múltiples proyecciones [Cho93].

En el año 1917, el matemático Austriaco J. Radon, estableció los fundamentos matemáticos de la TAC demostrando que era posible reconstruir un objeto bidimensional o tridimensional, a partir de un conjunto de infinitas proyecciones [Corbo, 2004]. El físico A.M. Cormack utilizó en 1963 dichos resultados para aplicaciones en medicina, surgiendo así la llamada tomografía computarizada. Posteriormente en 1972, cuando Godfrey Hounsfield trabajaba de ingeniero en la compañía de discos EMI, demostró públicamente el funcionamiento de este sistema. Y en 1982, este ingeniero británico recibió el premio Nobel de Física, compartido con el físico Alan Cormack, de la Tufts University. Habían demostrado que el haz de rayos X que atravesaba un objeto, contenía información de todos los constituyentes del material que había atravesado.

Durante mucho tiempo se han utilizado radiografías de tórax para detectar señales de carcinoma de pulmón en los pacientes con síntomas, pero esas pruebas no permitían reducir las muertes originadas por dicho carcinoma, debido al estadio avanzado en el que se encontraban. Sin embargo, con la tomografía computarizada se pueden detectar tumores más pequeños que con la radiografía, y por lo tanto más incipientes, por ello se utiliza como nueva tecnología para “salvar vidas”.

La tomografía computarizada más sencilla utiliza un escáner que usa un haz de rayos X, y un único detector. La fuente de rayos X y el detector están conectados, de tal modo que se mueven de forma sincronizada, radiando el objeto desde diferentes direcciones [TSID,

2011]. Cuando el conjunto fuente-detector efectúa un barrido para obtener la imagen, las estructuras internas del cuerpo atenúan el haz de rayos X según sus respectivos valores de número atómico y densidad de masa. La intensidad de la radiación detectada se recoge en el detector en forma de mapa de atenuaciones. Este mapa se utiliza para resolver un complejo sistema de ecuaciones que permitirá conocer los valores de los diferentes coeficientes de atenuación, los cuales se corresponden con los valores de cada punto de la imagen final. La reconstrucción de una imagen cualquiera con la Tomografía Computarizada requiere un computador digital que resuelva dicho sistema de ecuaciones (aproximadamente 30.000 ecuaciones), y por eso a la tomografía se la califica como computarizada.

El resultado final de la reconstrucción de cada corte axial es una matriz de números, y cada elemento de la matriz corresponde a un píxel de la imagen. La visualización de la imagen se realiza utilizando una escala de grises. Los valores almacenados en esta matriz son proporcionales al coeficiente de atenuación del tejido situado espacialmente en la misma posición que el píxel correspondiente. A estos valores se los denomina números TC y son calculados de acuerdo a la siguiente expresión:

$$TC = (\mu_{material} \cdot E - \mu_{agua} \cdot E) / K$$

Donde E representa la energía efectiva del haz de rayos X,  $\mu_{material}$  y  $\mu_{agua}$  son los coeficientes lineales de atenuación del agua y del material en estudio respectivamente y K es una constante que depende del diseño del equipo. Universalmente se ha adoptado la escala Hounsfield, la cual asigna el valor cero 0 al agua, y el -1000 al aire. De esta forma, los materiales más densos como el tejido óseo o los metales quedan en un rango que abarca desde cero hasta 1000 ó 3000 unidades Hounsfield (UH), dependiendo de la escala utilizada por el tomógrafo. Las Unidades Hounsfield y la densidad presentan una correspondencia lineal, conocida como función o curva de calibración.

En la Figura 10 se muestra el rango dinámico, desde -1000 hasta 1000 unidades Hounsfield, para diferentes tejidos del cuerpo humano.

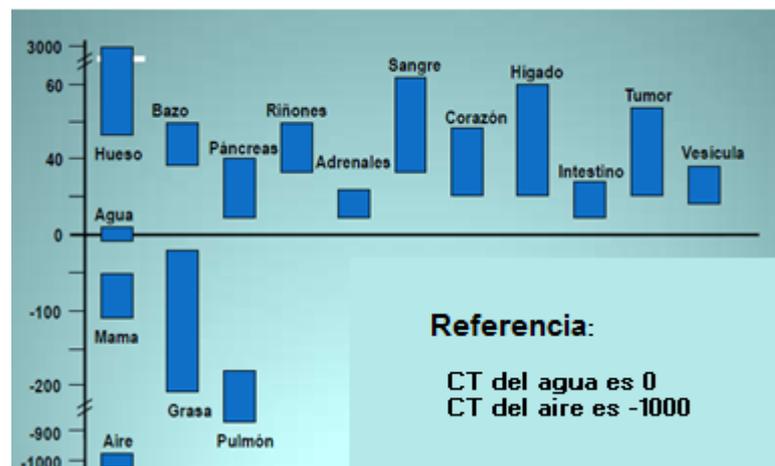


Figura 10. Rango dinámico de las unidades Hounsfield para tejidos del cuerpo humano.

La información que proporciona cualquier imagen radiológica está comprendida en los 16 niveles de grises que el ojo humano puede distinguir. Para establecer una compatibilidad entre el rango dinámico de las unidades Hounsfield y esta escala de grises se asignan 1000 UH al blanco y  $-1000$  UH al negro.

Los escáneres actuales almacenan los datos de las imágenes en discos duros del ordenador; y aunque para el almacenamiento de estas imágenes existen numerosos formatos útiles en el campo de la radiología, para el manejo, almacenamiento, impresión y transmisión de las imágenes que se utilizan en la mayoría de los centros médicos, destaca y se utiliza el estándar conocido por el término DICOM.

### 2.2.3 DICOM

El estándar Digital Imaging and Communication in Medicine, conocido por las siglas DICOM, se utiliza para facilitar la integración de los diferentes dispositivos y sistemas que tratan las imágenes médicas. [DICOM, 2011\_2009]

El comité formado por el American College of Radiology (ACR) y la National Electrical Manufacturers Association (NEMA) desarrolló, en 1983, un método estándar para transmitir información e imágenes médicas. Fue en 1985 cuando publicaron la primera versión “el estándar ACR-NEMA”, donde definieron los protocolos de comunicación y el tipo de información a intercambiar entre los diferentes dispositivos, además de definir el formato digital para almacenar las imágenes como objetos.

En sus inicios fue pensado para almacenar e intercambiar únicamente imágenes e información relacionada con las áreas de radiología y cardiología. Posteriormente en 1993, la versión 3.0 que incluía numerosas mejoras para las comunicaciones estandarizadas, se amplió a otros campos de la medicina, y le cambiaron el nombre por Digital Imaging and Communications in Medicine (DICOM). Actualmente, DICOM es sostenido por los miembros del DICOM Standards Committee que está integrado por organizaciones, vendedores de hardware y software para PACS y otros grupos de interés general.

DICOM es un estándar orientado a objetos, en el que una entidad del mundo real como un paciente, una visita, una imagen, etc. es modelada como un objeto. La información almacenada por cada objeto es la de los atributos que lo caracterizan.

Utiliza el modelo conceptual Entidad-Relación para representar la imagen médica y todo su entorno de tal forma, que la información está organizada tal como indica el diagrama Entidad\_Relación (E/R) de la Figura 11.

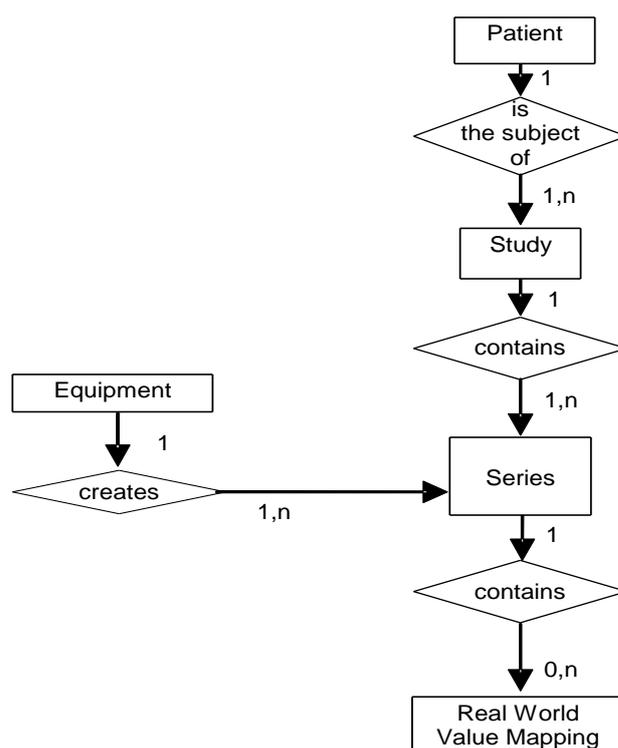


Figura 11. Diseño conceptual del estándar DICOM [DICOM, 2009].

La unidad elemental de DICOM, denominada Service-Object Pair (SOP), está formada por dos elementos: el Information Object Definition (IOD) y la “Clase de Servicio”. Los IOD definen el contenido de las imágenes médicas del centro, y las Clases de Servicio definen lo que hay que hacer con ese contenido.

Las SOP son la unidad elemental de DICOM y todo lo que DICOM hace se basa en la utilización de estas clases.

DICOM puede definir en la clase SOP, por ejemplo Imágenes de TC. El objeto de información TC y la clase de servicio de almacenamiento, se combinan para formar la “CT image storage SOP class”. Una vez definidos los objetos de interés y todas sus características, DICOM define las operaciones que pueden ser ejecutadas y sobre qué objetos; a estas operaciones se las llama “DICOM Message Service” (DIMSE). Así, para que dos objetos se comuniquen utilizan mensajes entre ellos.

Desde la perspectiva de la implementación, cada uno de los ficheros DICOM tiene *una cabecera y un cuerpo*. La cabecera contiene multitud de campos estandarizados correspondientes a los atributos de los objetos; en ella se especifican tanto datos administrativos (datos del paciente, hospital donde se realizó, entre otros) como datos sobre la imagen. El cuerpo contiene la imagen, que puede estar comprimida en distintos estándares.

El formato de fichero es muy complejo, debido a la gran cantidad de campos que se especifican en la cabecera, así como a los varios tipos de cabecera que permite y la multitud de formatos en los que puede estar grabada la imagen.

Específicamente, un fichero DICOM consta de cuatro partes: Preámbulo y prefijo identificativo del fichero, Meta-cabecera, Cabecera y finalmente la Imagen (aunque desde el punto de vista del formato, la imagen es un elemento más de la cabecera). El fichero en formato DICOM comienza con un preámbulo de 128 bytes, cuyo uso está definido por la implementación. El prefijo identificativo permite a las implementaciones diferenciar si un fichero es DICOM o no. La cabecera y la meta-cabecera están formadas por campos con la información específica de la imagen, incluyendo a la propia imagen.

De todos estos campos, los que interesan para esta tesis son los que contienen la información necesaria para procesar y visualizar la imagen correctamente. En los campos de la cabecera y meta-cabecera está toda la información necesaria para que una implementación del estándar sea capaz de procesar y visualizar correctamente la imagen o imágenes almacenadas en el fichero, además de todos los datos asociados a ellas que el personal médico necesita para interpretar correctamente las imágenes [DICOM, 2009].

DICOM facilita la conexión entre los equipos de imágenes médicas especificando el conjunto de protocolos que deben ser seguidos por los dispositivos conformes al estándar para las comunicaciones en red (véase en Figura 12. Modelo de Comunicación DICOM.).

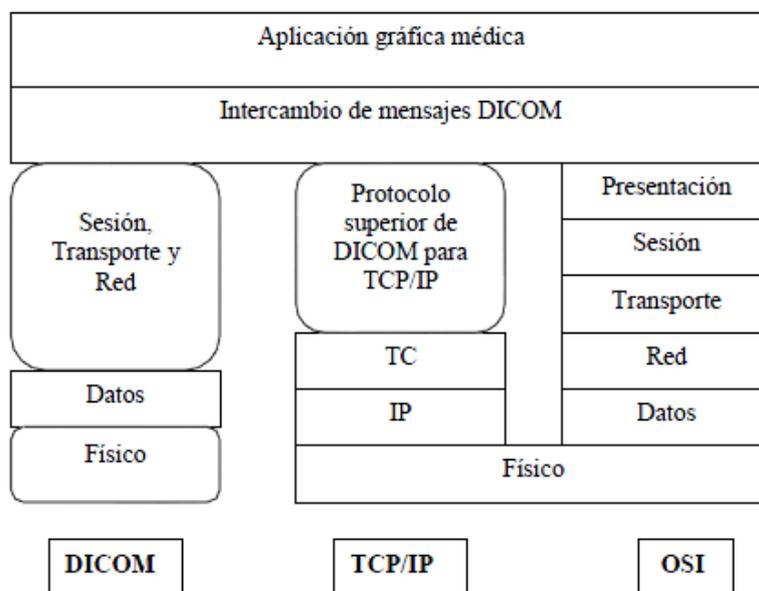


Figura 12. Modelo de Comunicación DICOM.

Este estándar indica la sintaxis y semántica de los comandos junto con la información asociada que puede ser intercambiada usando estos protocolos. También detalla la información que debe suministrarse y el conjunto de servicios, para que se puedan almacenar datos, que tienen que seguir los dispositivos conformes a dicho estándar.

#### 2.2.4 La imagen digital y su procesamiento

Formalmente una imagen es una función bidimensional de la intensidad de la luz  $f(x, y)$ , donde  $x$  e  $y$  son las coordenadas espaciales en un punto y  $f(x, y)$  es el valor proporcional al brillo (o nivel de gris) de la imagen en ese punto. Una imagen digital  $f(x, y)$  se ha discretizado tanto en las coordenadas espaciales como en el brillo.

Dicha función  $f$  debe ser estrictamente positiva y finita por ser una función de energía tal como se indica en la siguiente expresión:

$$f(x, y) = i(x, y) \cdot r(x, y)$$

Siendo  $0 < i(x, y) < \infty$  y donde  $0 < r(x, y) < 1$ . Con  $i(x, y)$  se representa la intensidad de luz que incide sobre un objeto y con  $r(x, y)$  la luz reflejada por dicho objeto la cual toma valores entre 0 (absorción total) y 1 (reflexión total).

Se expresa la imagen digital como una matriz cuyos índices identifican un punto de la imagen y el valor del correspondiente elemento de la matriz indica el nivel de gris o

intensidad digital en ese punto. A cada uno de los elementos de una distribución digital de este tipo se le denomina pixel, abreviatura de elemento de la imagen en su denominación inglesa. Si la imagen tiene tres dimensiones,  $f(x, y, z)$ , entonces el elemento se nombra con el término “voxel”. Siendo el píxel la unidad mínima de información en una imagen de dos dimensiones y el voxel en una imagen tridimensional.

Las imágenes objeto de este trabajo se corresponden con un par ordenado  $(M, N)$ , donde  $M$  es el número de filas y  $N$  es el número de columnas de la matriz, cuyo tamaño de imagen -en bits- viene dado por el producto  $M \times N \times k$ , donde  $k$  es el nivel de cuantificación o profundidad de color; además, el número de niveles de gris es el valor  $2^k$ , es decir la cantidad de niveles de cuantificación que pueden mostrarse en una imagen y la resolución espacial es la cantidad de píxeles por unidad de longitud en cada eje de coordenadas de la imagen.

Para trabajar con estos valores en la computadora, el nivel de brillo o valor de cada píxel (nivel de gris) es cuantificado a códigos binarios enteros positivos. El valor 0 corresponde a la tonalidad del negro y el valor más alto ( $2^k$ ) al blanco.

El tamaño idóneo de una imagen digital varía dependiendo de su aplicación. En la Tabla 3 se muestra, a nivel orientativo, una lista de tamaños de algunas imágenes médicas convencionales. Cuanto más se aproxime la imagen digitalizada a la original mayor capacidad de almacenamiento y procesamiento necesita.

Tipo de imágenes		Matriz píxeles	Bits por pixel	Tamaño (MB)	Imágenes porEstudio
<b>Radiología</b>	Tórax	3500x4300	16	28,7	2
	Columna	2400x3000	16	13,7	1
	Mano	1800x2400	16	8,2	2
<b>Mamografía</b>		3600x4800	16	33	4
<b>Tomografía</b>		512x512	16	0,5	400
<b>Resonancia</b>		512x512	16	0,5	500
<b>Ecografía</b>		512x512	8	0,3	20
<b>Eco-Doppler</b>		512x512	24	0,8	12
<b>Medicina Nuclear</b>		128x128	24	0,1	30

Tabla 3. Tamaño según tipo de imágenes médicas.

El número de píxeles existentes por cada valor de nivel de gris de una imagen se representa mediante un histograma como el de la Figura 13.

Tanto la resolución espacial como el nivel de gris o cuantificación definen en gran medida la calidad de la imagen. Si dicha imagen se va a utilizar con fines de diagnóstico entonces

será necesario visualizar las diferentes estructuras anatómicas existentes en ella y permitir detectar cualquier signo de patología existente. Por ello la calidad es fundamental y cualquier defecto en la citada imagen disminuye la capacidad de utilizarla como método de diagnóstico médico. El ruido, los artefactos, la borrosidad, la distorsión, el insuficiente contraste son factores que restan calidad a la imagen. Para mitigar estos efectos indeseables se utilizan diferentes técnicas de filtrado de imágenes en la etapa de preprocesado.

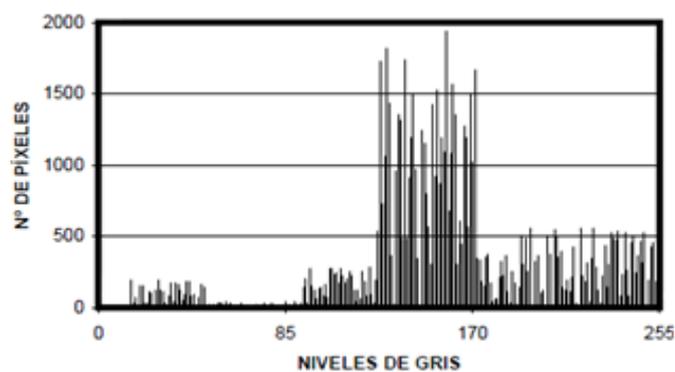


Figura 13. Histograma de una imagen de 256 niveles de gris.

En una imagen radiológica el valor del píxel,  $f(x, y)$ , representa la estructura anatómica en el espacio, definiéndose la imagen en el dominio espacial. Una imagen también puede ser representada por sus componentes de frecuencia espacial  $(u, v)$ , utilizándose para ello herramientas matemáticas. La imagen  $f(x, y)$  transformada al dominio de la frecuencia se representa por  $F(u, v)$ , la cual es nuevamente una imagen digital pero no tiene el parecido visual de  $f(x, y)$ . La información que se visualiza en el dominio de la frecuencia no es fácilmente visible en el dominio espacial y la calidad de una imagen es más fácil medirla en el dominio de la frecuencia.

En este trabajo se han utilizado imágenes de tomografías computarizadas en el dominio espacial, siendo por tanto más difícil de cuantificar la calidad de las mismas.

Para mejorar el aspecto de las imágenes y hacer evidentes detalles que interese resaltar se realiza un procesado de imágenes. Este procesamiento digital se inicia cuando la imagen es considerada una matriz en la que cada uno de sus elementos tiene un valor numérico correspondiente a la luminosidad promedio de cada píxel en la citada imagen. Con la finalidad de resaltar los detalles convenientes se modifican los valores de la luminosidad de los píxeles mediante operaciones o transformaciones matemáticas aplicadas a la imagen [González, 1996].

Así pues, tratar imágenes digitales implica procesarlas y analizarlas. Con el procesado de imágenes se pretende mejorarlas y con el análisis extraer de ellas características para identificarlas y clasificarlas.

El procesado de imágenes se caracteriza por proporcionar soluciones específicas para cada problema, teniendo en cuenta que, lo que funciona bien en un caso puede ser totalmente inadecuado en otro.

### **2.2.4.1 Segmentación de imágenes**

La Segmentación es una de las operaciones más utilizadas en el procesado digital de imágenes. Segmentar es un proceso que consiste en dividir una imagen en objetos o partes para clasificarlos. A la salida de este proceso se obtienen, generalmente, los datos de los píxeles que constituyen el contorno de una región determinada, o bien toda la región. En este proceso se realiza una asignación de píxeles a uno de varios conjuntos disjuntos que comparten una propiedad común.

Segmentar es de las tareas más complicadas debido a la necesidad de representar, mediante una estructura entendible por el ordenador, los diferentes elementos segmentados; habiendo pérdida de información debido al paso de una estructura 3D a 2D que se basa solamente en la intensidad de los píxeles. Asimismo la estructura que se pretende identificar depende en gran medida del objetivo a conseguir. Este problema se ve dificultado por la falta de estándares y metodologías, además de por las características que presentan las imágenes digitales: artefactos debidos al muestreo, aliasing espacial, baja relación señal-ruido, aspecto similar para diferentes elementos, formas complejas o dificultad de incorporar conocimiento a la segmentación.

Los métodos de segmentación de imágenes monocromáticas se suelen basar en una de las propiedades básicas de los valores del nivel de gris: la discontinuidad y la similitud [González, 2004]. Los métodos basados en la detección de discontinuidades consisten en dividir la imagen partiendo de las variaciones de luminosidad presentes en ella, suponiendo que los cambios bruscos de intensidad se corresponden con los bordes de los objetos que se quiere delimitar, mientras que los cambios suaves se corresponden con las zonas homogéneas de los objetos. Siendo los algoritmos más representativos de este método: los detectores de líneas y los detectores de bordes presentes en la imagen. Y los principales métodos basados en la similitud son la umbralización, y las técnicas de segmentación por regiones. Estas últimas buscan dividir la imagen en regiones o clústeres,

maximizando la homogeneidad de los píxeles en una región y la heterogeneidad entre regiones.

La formulación básica de la segmentación orientada a regiones se expresa formalmente a continuación.

Dada una imagen  $R$  y  $R_1, \dots, R_n$  el conjunto de regiones resultantes de su segmentación, debe cumplirse que:

- a)  $\bigcup_{i=1}^n R_i = R$
- b)  $R_i$  es una región conexa, con  $i=1, \dots, n$
- c)  $R_i \cap R_j = \emptyset$  para todo  $i$  y  $j$ , con  $i \neq j$
- d)  $P(R_i) = \text{VERDADERO}$  para  $i=1, \dots, n$
- e)  $P(R_i \cup R_j) = \text{FALSO}$  para  $i \neq j$

Donde  $P(R_i)$  es un predicado lógico sobre los puntos del conjunto  $R_i$ , y  $\emptyset$  es el conjunto vacío.

Todavía los algoritmos más utilizados, dentro de esa técnica de segmentación, son los algoritmos de clústering difuso ya que se caracterizan por tener elevada inmunidad al ruido y mayor capacidad para manejar información ambigua; razones por las que se consideran adecuados para trabajar con las imágenes digitales médicas diagnósticas de CT [Castro, 2004].

Existen las técnicas de segmentación por modelos que buscan combinar conocimiento sobre el elemento que se quiere reconocer en la imagen y la información disponible en la propia imagen. Dentro de éstas se encuentran técnicas muy diferentes como las Redes de Neuronas Artificiales (RNA), modelos de campos aleatorios de Markov (modelo estadístico que puede ser usado dentro de los métodos de segmentación) o los modelos deformables.

La mayoría de los métodos de segmentación pueden tratarse como problemas de optimización, donde la segmentación deseada es la que minimiza alguna función de energía, definida para una aplicación en particular. La ventaja, de considerar la segmentación como un problema de optimización, es que se definen de manera precisa los aspectos deseables de la segmentación. Es normal que, para diferentes aplicaciones, se necesiten diferentes funciones de energía.

### 2.2.4.2 Algoritmos de Clustering y Lógica Difusa

A principios del siglo pasado R. Tryon introdujo el término de “análisis clúster”; dicho término engloba a diferentes algoritmos de clasificación cuyo objetivo es la organización sucesiva de la información hasta convertirla en conjuntos comprensibles y manejables a los que se denomina clústeres [Tryon, 1939]. Por lo tanto, esta técnica clasifica objetos similares en clases o categorías diferentes, agrupándolos en base a algún tipo de criterio como la división del conjunto inicial de datos en subconjuntos (clústeres); de forma que los elementos asignados a cada uno de ellos compartan algún atributo común, que generalmente es la proximidad de acuerdo a alguna medida predefinida de la distancia (city-block, distancia euclídea, distancia Chebychev, etc.) [Hartigan 2001].

El objetivo básico de los algoritmos de segmentación es definir una partición del espacio. Los algoritmos de agrupamiento o clustering de datos, tanto por sus características como por su frecuente aplicación en problemas de reconocimiento de patrones, son una técnica que puede adaptarse al problema de la clasificación no supervisada de los píxeles constituyentes de los nódulos pulmonares en las imágenes de CT de tórax.

Los algoritmos de clustering forman grupos, clases o clústeres a partir de una imagen. Para ello reciben como datos de entrada, además de la citada imagen a segmentar, el vector de características y parámetros específicos como, por ejemplo, el número de clústeres deseados; dichos algoritmos, de forma autónoma y sin supervisión de ningún tipo, agrupan los píxeles de la imagen recibida en clústeres a partir de la información disponible. De tal forma que los grupos o clústeres resultantes de la segmentación comparten características análogas y simbolizan la forma en la que se analiza y describe el mundo por parte de los humanos. Así cuando se dividen objetos en grupos se realiza una tarea de clustering y cuando se asignan objetos particulares a dichos grupos se está haciendo una clasificación.

Los métodos de clustering existentes difieren, a grandes rasgos, en la forma de estructurar los clústeres. Aquellos que encuentran clústeres que corresponden a una partición del conjunto de objetos son señalados como métodos de Hard-clustering o clustering particional; el más conocido de ellos es el algoritmo CMeans, también llamado K-Means. Y a los métodos que asignan a cada objeto un valor de pertenencia con respecto a cada clúster se les conoce como métodos de Soft-clustering, siendo el más representativo de este tipo de clustering el algoritmo Fuzzy C-Means.

Exceptuando el algoritmo de clustering particional CMeans, los demás algoritmos empleados en esta tesis son Soft\_clustering y por tanto utilizan la lógica borrosa para clasificar los píxeles en clústeres.

La Lógica Difusa o borrosa, también conocida como lógica fuzzy, es una generalización de la teoría general de conjuntos que permite imitar el comportamiento de la lógica humana. Surge como una forma de procesar información con un alto grado de imprecisión o ambigüedad, permitiendo que un elemento pertenezca de forma parcial a uno o a varios conjuntos. Se utiliza para realizar operaciones en los conceptos que están fuera de las definiciones de la lógica convencional de los conjuntos clásicos, en los cuales un elemento sólo tiene dos opciones: pertenecer o no pertenecer a un conjunto. Así, en la lógica difusa un elemento puede tener infinitos grados intermedios entre los estados: *verdadero* y *falso*.

Heráclito (s IV a.C.) propuso que las cosas podían ser verdad o no verdad simultáneamente. En la década de 1900 Lukasicwicz empezó proponiendo una lógica de tres valores junto con las matemáticas que la describían -el tercer valor asignado entre verdadero y falso lo definió con el significado de “posible”- para posteriormente trabajar en lógica plurivalente, la cual puede admitir distintas cantidades de valores de verdad: desde tres, hasta infinito [Pérez, 2007].

Lofti Asker Zadeh, profesor de la Universidad de California en Berkeley, publicó el trabajo “Fuzzy Sets”, en el cual describió las matemáticas de la teoría de conjuntos difuso y con ello la lógica difusa [Zadeh, 1965].

Los conjuntos difusos son utilizados para representar la indecisión, pueden contener elementos de forma parcial, es decir, la propiedad  $x \in A$  puede ser cierta con un grado de verdad. Se mide dicha “pertenencia” con un número,  $\mu_A(x)$  entre 0 y 1, llamado grado de pertenencia de  $x$  a  $A$ . Si su valor es 0,  $x$  no pertenece a  $A$ ; si  $0 < \mu_A(x) < 1$ ,  $x$  pertenece a  $A$  de una manera parcial, y si es 1, entonces  $x \in A$  totalmente. Siendo conocido  $\mu_A(x)$  como el valor de la función de pertenencia. De esta forma en un conjunto difuso, la función de pertenencia es la que asocia a cada objeto del conjunto un valor, comprendido entre el intervalo  $[0,1]$  según el grado de pertenencia del objeto al conjunto difuso. Se suele normalizar el grado de pertenencia máximo a 1.

Con los conjuntos difusos se pueden definir sub-conjuntos, tal que cualquier elemento pueda pertenecer a ellos en diferentes grados. Para cada conjunto difuso existe asociada una función de pertenencia que le caracteriza; dicha función indica en qué medida un

elemento forma parte de ese conjunto difuso. Las formas de las funciones de pertenencia más típicas son trapezoidales, lineales y curvas.

El centro de gravedad de un conjunto difuso es único, y se corresponde con lo que se denomina centro o centroide, es decir, el vector considerado como representante de un clúster particular.

### *2.3 Algoritmos de procesamiento analizados.*

Para abordar la clasificación automática de los píxeles del NPS de las imágenes se han seleccionado algoritmos basados en la partición de datos.

Los algoritmos convencionales de segmentación (CMeans) consideran que un píxel sólo puede pertenecer a una clase. Al analizar las imágenes de TAC, se observa que la frontera entre objetos es difusa, por eso lo adecuado es trabajar con conjuntos difusos; y por tanto, utilizar algoritmos de segmentación difusa, que presentan claras ventajas sobre los métodos de clustering tradicionales, dado que permiten asignar a cada píxel una pertenencia parcial o distribuida entre los clústeres en los que se divide la imagen.

Se ha utilizado el clásico algoritmo CMeans, como referencia, para poder establecer una comparación en los resultados con los demás algoritmos empleados; como método más utilizado de clustering difuso, se ha incluido la variante difusa del algoritmo CMedias, el algoritmo Fuzzy CMeans, que permite especificar el grado de pertenencia de cada elemento de entrada a los diferentes clústeres obtenidos.

Se utiliza el algoritmo Fuzzy K-Nearest Neighbor, como formalización difusa de los métodos de análisis clúster denominados “del vecino más cercano” (KNN); en este método se determina la función de pertenencia de cada vector de entrada, a los distintos grupos, mediante el estudio de las distancias de cada elemento a un conjunto de muestras cercanas a cada clase. Es necesario proporcionar externamente y con anterioridad, a la ejecución del algoritmo, los representantes de cada clúster para que se realice correctamente el agrupamiento.

Las imágenes biomédicas de TAC presentan inhomogeneidades entre píxeles que deben pertenecer a una misma clase. Ello ha exhortado a escoger algoritmos diseñados para compensar dicha falta de uniformidad; se ha elegido pues, el Adaptive Fuzzy CMeans por estar diseñado para segmentar imágenes con inhomogeneidades de tipo multiplicativo. Asimismo, también se ha empleado el algoritmo Spatially Kernelized Fuzzy CMeans porque compensa las inhomogeneidades al añadir un término de penalización espacial,

además de emplear una métrica de distancia basada en un kernel, y permitir que el etiquetado de los píxeles tenga en cuenta a los píxeles adyacentes, denominados vecinos. En los apartados siguientes se expone la descripción teórica y la implementación particular de cada algoritmo de clustering seleccionado. Indicando en la Tabla 4 los acrónimos que se utilizan para referenciar a estos algoritmos a lo largo de la memoria. Se verá que, la mayoría de los algoritmos de clústering difuso elegidos, intentan minimizar una función objetivo cuyos parámetros principales son los grados de pertenencia, y los parámetros que determinan la forma y localización de los clústeres.

<b>Método</b>	<b>Identificador</b>	<b>Origen</b>
<b>C-Means</b>	CMeans	[MacQueen, 1967]
<b>Fuzzy C-Means</b>	FCM	[Bezdek, 1981]
<b>Fuzzy K-Nearest Neighbor</b>	FKNN	[Keller, 1985]
<b>Adaptive Fuzzy C-Means</b>	AFCM	[DL Pham, 1999]
<b>Spatially Kernelized Fuzzy C-Means</b>	SKFCM	[Zhang, 2004a]

**Tabla 4. Algoritmos de segmentación Clustering empleados.**

La implementación y evaluación de dichos algoritmos de clustering se ha llevado a cabo mediante el software Matlab. Éste es un paquete interactivo que proporciona un entorno de computación y desarrollo de aplicaciones integrado, orientado a la ejecución de proyectos en los que se encuentren implicados tanto cálculos matemáticos como la visualización gráfica de los mismos [Matlab 2010]. En los últimos años, Matlab se ha convertido en un estándar, pues integra de forma coherente los requisitos esenciales de un sistema de computación técnico (análisis numérico, cálculo matricial, procesamiento de señales, visualización gráfica, herramientas para aplicaciones específicas, capacidad de ejecución en múltiples plataformas, etc.), en un entorno donde los problemas y sus soluciones se expresan en lenguaje de alto nivel basado en vectores y matrices, sin necesidad de hacer uso de la programación tradicional [Hahn 2007].

### 2.3.1 C-Means.

El algoritmo CMeans, conocido también como K-means o K-medias es un método de agrupamiento por partición, desarrollado por MacQueen [MacQueen, 1967] Se basa en elegir inicialmente los centroides, datos representativos de los C grupos considerados clases o clústeres, y en asignar cada dato o pixel de entrada al clúster más semejante o cercano. Posteriormente se computa de nuevo el centro y se repite el proceso de clasificación hasta que se alcance la condición de convergencia. Esta técnica se ha aplicado

con éxito a grandes volúmenes de datos de entrada, ya que su complejidad es lineal, una vez que se han determinado tanto el número de clases ( $C$ ) como el número de ciclos del algoritmo. Es una técnica de clústering capaz de encontrar rápidamente una solución pero no garantiza una configuración óptima de los clústeres, pues tiene una alta dependencia de la elección de los centroides iniciales.

En la Figura 14 se puede ver el flujo de control de este algoritmo, descrito exhaustivamente a continuación.

Debido a su sensibilidad a la inicialización, es preciso e imprescindible además de conocer de antemano el número de grupos en los que se desea clasificar los vectores de entrada, elegir convenientemente los centroides de cada una de estas  $k$  clases de forma supervisada.

Una vez determinados los centroides, pixeles iniciales para cada clúster, se examinan todos los puntos a clasificar para determinar a cuál de las clases se parecen más. El grado de este parecido se determina mediante un criterio de medida, comúnmente una distancia Euclídea:

$$x_j \in C_j \Leftrightarrow \|x_i - v_j\| < \|x_i - v_k\|, \forall k = 1, \dots, C, k \neq j, \forall i = 1, \dots, N$$

Dónde  $C_j$  especifica la clase a la que se asignará el elemento de entrada  $x_j$ . Siendo  $v_j$  el centroide de dicha clase.

Cuando se han terminado de examinar todos los puntos (a excepción de los puntos tomados inicialmente como muestras), y se han asociado con una clase determinada, se realiza una nueva clasificación; dicha clasificación considera a todos los puntos de la imagen, incluidos los correspondientes a las muestras iniciales, para calcular de nuevo los centroides o centros de gravedad de las clases, como se indica en la expresión siguiente:

$$v_k = \frac{1}{n_k} \sum_{i=1}^{n_k} x_i, \forall k = 1, \dots, C$$

Donde  $n_k$  es el número de elementos de entrada asignados a la clase  $C_k$  durante la iteración.

Esta fase de asignación y cálculo de los centroides continúa iterativamente hasta que se alcanza la condición de finalización o convergencia del algoritmo; dicha convergencia se produce cuando la variación en los centros de cada clase no supere un cierto umbral ajustable, y con lo cual no se produce ninguna redistribución de los vectores de entrada entre una iteración y la siguiente.

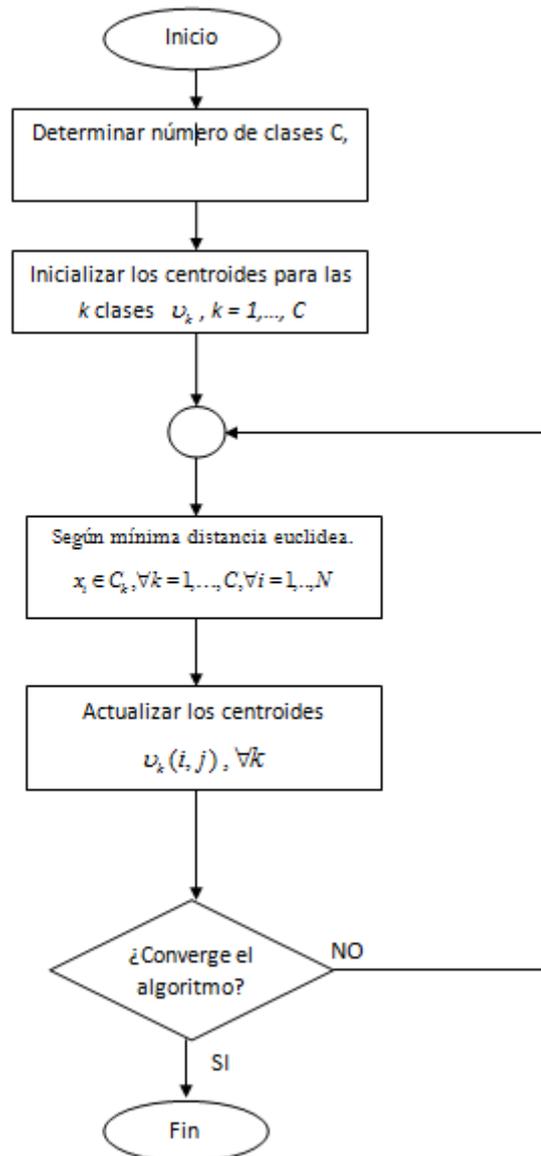


Figura 14. Ordinograma del algoritmo CMEANS.

### 2.3.2 Fuzzy C-Means

Ruspini fue el primero en incorporar la teoría de lógica difusa en los algoritmos de agrupamiento [Ruspini, 1970], después Dunn propone y desarrolla el algoritmo [Dunn, 1973] y lo mejora y generaliza Bezdek [Bezdek, 1981]. Es utilizado con frecuencia en el reconocimiento de patrones y es la base de muchos de los algoritmos de clústering difusos disponibles actualmente.

Este algoritmo, Fuzzy C-Means -conocido por las siglas FCM-, es un método de agrupamiento que generaliza al algoritmo C-Means para clustering difuso; permite que un

pixel pertenezca a dos o más clústeres a la vez, mediante la especificación de un grado de pertenencia a cada grupo o clúster.

Se basa en el método de división de conjuntos iterativo de Picard; definidos éstos como un esquema de aproximación sucesiva. Se utiliza normalmente para resolver ecuaciones diferenciales; empieza con unos valores iniciales y mediante sustituciones sucesivas converge en una solución, aproximando así, el cálculo del error mínimo cuadrado de la siguiente función objetivo:

$$J_{FCM} = \sum_{i=1}^N \sum_{k=1}^C u_{ik}^m \|x_i - v_k\|^2$$

La imagen  $X = \{x_1, x_2, x_3, \dots, x_N\}$  es el conjunto de elementos que es preciso clasificar,  $N$  es el número de píxeles;  $C$  es el número de clústeres en el que segmentar la imagen,  $u_{ik}^m$  representa al factor de pertenencia del pixel  $i$ -es decir el elemento  $x_i$  - a la clase o clúster referenciado por  $k$ , cuyo centroide es  $v_k$ . Finalmente,  $m$  es un factor de ponderación que debe tomar valores de tipo real en el intervalo abierto  $(1, \infty)$ , indicando el grado de factor fuzzy de los centros de los clústeres resultantes.

El operador  $\|\dots\|$  especifica cualquier condición de la similitud entre un dato de entrada y el centroide o elemento representativo del centro de la clase, como puede ser la distancia euclídea. De este modo la partición de la imagen de entrada se efectúa mediante la optimización iterativa de la función  $J_{FCM}$ , con la actualización sucesiva tanto de las funciones de pertenencia  $u_{ik}^m$ , como de los centroides  $v_k$ , como se muestra en los pasos del algoritmo representados en la Figura 15.

Como también se indica en el ordinograma, debe conocerse a priori el número de clústeres en los que se va a dividir la imagen, y pasarle una muestra de cada uno de los clústeres para que sea utilizada como el centroide inicial en la clase correspondiente.

La matriz de pertenencias -indicada a continuación- tiene  $N$  columnas, tantas como elementos haya que clasificar; y tiene  $C$  filas que se corresponden con los clústeres en los que segmentar la imagen  $X$

$$U = \begin{pmatrix} u_{11} & u_{12} & \dots & u_{1N} \\ u_{21} & u_{22} & \dots & u_{2N} \\ \vdots & \vdots & & \vdots \\ u_{C1} & u_{C2} & \dots & u_{CN} \end{pmatrix}$$

Debe cumplirse que en cada columna de  $U$  la suma de los elementos es igual a 1.

$$\sum_{i=1}^C u_{ik} = 1, \forall k$$

$$u_{ik} \in [0,1] \quad \text{siendo} \quad 1 \leq i \leq C, 1 \leq k \leq N.$$

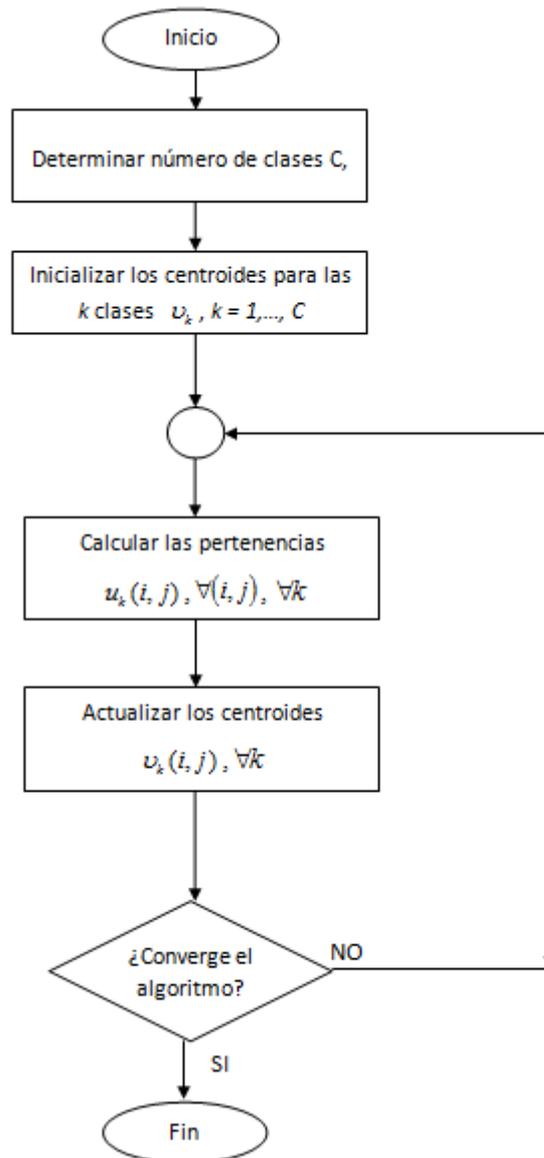


Figura 15. Ordinograma del algoritmo FCM.

El cálculo de las nuevas pertenencias se realiza según la siguiente expresión:

$$u_{ik} = \left( \sum_{j=1}^C \left( \frac{\|x_k - v_i\|}{\|x_k - v_j\|} \right)^{\frac{2}{m-1}} \right)^{-1}$$

Y para la actualización de los nuevos prototipos de centroides, se aplica la formula expresada a continuación:

$$v_i = \frac{\sum_{k=1}^N u_{ik}^m \cdot x_k}{\sum_{k=1}^N u_{ik}^m}$$

Se considera que el algoritmo converge cuando la diferencia entre los centroides de cada clúster entre dos iteraciones consecutivas cumplen  $\|v_{actual} - v_{anterior}\| < \varepsilon$  donde  $\varepsilon$  es el valor de error permitido, que funciona como criterio de finalización o de parada.

### 2.3.3 Fuzzy K-Nearest Neighbor

El algoritmo FKNN (Fuzzy K-Nearest Neighbor) es una versión difusa de los métodos de análisis clúster [Keller, 1985] denominados “del vecino más cercano” (KNN); cuyo paradigma es que un nuevo pixel se va a clasificar en la clase o clúster más frecuente a la que pertenecen sus “K” vecinos más cercanos.

Ha sido utilizado en aplicaciones biomédicas, tanto para segmentar imágenes como para clasificarlas en bases de datos biomédicas.

Los métodos basados en el paradigma KNN determinan la pertenencia de cada vector de entrada de los diferentes clústeres, calculando la distancia a un conjunto de muestras representativas de los mismos, de tal modo que cada elemento de entrada se asocia a la clase a la que pertenezca la muestra más similar a él de entre las K más próximas. Estas “K muestras más próximas” son los denominados “vecinos”.

En la versión difusa, FKNN, para determinar la pertenencia de los pixeles de la imagen de entrada se considera la pertenencia de cada pixel a todos los clústeres posibles, en lugar de asociar cada pixel estrictamente a una clase y se calcula la función de pertenencia de dicho elemento a los distintos clústeres existentes. Es decir, al ser un algoritmo difuso, un píxel puede tener un 90% de pertenencia a la clase A, el 8% a la clase B y el 2% a la C. Con esto se entiende que ese píxel esta dentro de la clase A, pero que además está más cercano a B que a C.

En la Figura 16 se muestra el funcionamiento general del algoritmo FKNN; indicando que inicialmente necesita conocer el número de clases en las que se va a dividir la imagen a clasificar. Además éste algoritmo también necesita como parámetro de entrada el

conjunto de muestras prototípicas de cada clase o clúster, el cual debe incluir al menos tantos elementos como vecinos se consideren. Dicho conjunto de muestras, necesarias para poner en funcionamiento este algoritmo, ha de estar clasificado de forma difusa, es decir, debe asociarse a cada muestra un grado de pertenencia a cada uno de los clústeres considerados.

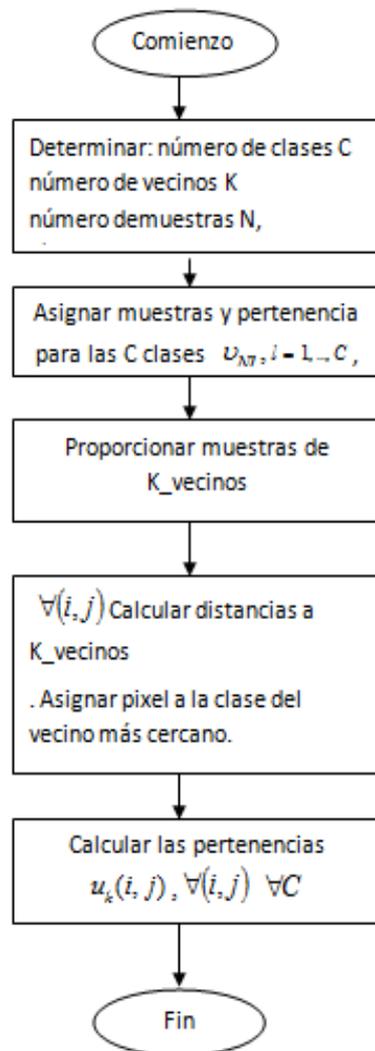


Figura 16. Ordinograma del algoritmo Fuzzy K-Nearest Neighbor.

Buscando los k vecinos más próximos de cada punto de la imagen a segmentar se obtienen las distancias que los separan del conjunto de muestras representativas de cada clúster. Y así, conociendo las distancias, se calculan los grados de pertenencia de los puntos o píxeles a todos los clústeres considerados; teniendo en cuenta tanto la distancia de cada punto de la imagen a todos sus vecinos, como el grado de pertenencia de cada

uno de estos vecinos a los clústeres objetivo. Para ello se implementa la siguiente expresión:

$$u_i(x) = \frac{\sum_{j=1}^K u_{ij} \left( \frac{1}{\|x - x_j\|^{\frac{2}{m-1}}} \right)}{\sum_{j=1}^K \left( \frac{1}{\|x - x_j\|^{\frac{2}{m-1}}} \right)}$$

Donde  $u_{ij}$  representa el factor de pertenencia de la  $j$ -ésima muestra al  $i$ -ésimo clúster. El valor  $x_j$  representa una de las  $K$  muestras más cercanas al píxel que se desea clasificar;  $x$  representa el píxel actual;  $m$  es el factor de ponderación de la distancia del elemento  $x$  a las muestras y  $u_i(x)$  representa el grado de pertenencia a la clase  $i$  del píxel  $x$ . Se puede observar que esta expresión valora tanto la distancia del píxel estudiado a todos sus vecinos como el factor de pertenencia de cada uno de esos vecinos a los clústeres objetivo.

#### 2.3.4 Adaptive Fuzzy C-Means

En las imágenes médicas, generalmente, la intensidad no es homogénea a lo largo de la imagen produciendo un efecto sombra, denominado “artefacto”, que hace que la intensidad de los píxeles que pertenecen a la misma clase, varíe a lo largo del dominio de la imagen. Este efecto puede generar errores graves cuando se segmenta la imagen utilizando métodos de clasificación basados en la intensidad de los píxeles. Las variaciones de intensidad en una imagen se pueden modelar al realizar el producto de la imagen original con un campo multiplicador, que varíe suave y lentamente a lo largo y ancho de dicha imagen.

Hay dos tipos de estrategias para segmentar imágenes con intensidades no homogéneas. La primera corrige la imagen antes de segmentarla, aplicando por separado primero un algoritmo de corrección y después otro de segmentación. En este primer método, una vez corregida la imagen, el algoritmo de segmentación asume que la intensidad es homogénea. Se han implementado diferentes versiones, mediante uso de polinomios, tales como la construcción de una superficie de corrección spline, usando como guía puntos de referencia seleccionados manualmente y mediante la aplicación de filtros homomórficos para eliminar los efectos de las inhomogeneidades.

La segunda estrategia consiste en compensar el efecto sombra a la vez que se segmenta la imagen. Con esta aproximación se puede utilizar información intermedia de la segmentación mientras se realiza la corrección.

El algoritmo Adaptive Fuzzy C-Means (AFCM), descrito en [DL Pham, 1999a] y en [DL Pham, 1999b], utiliza esta segunda estrategia. Es un algoritmo de clustering no supervisado, diseñado para obtener la segmentación de imágenes que están sujetas a intensidades no homogéneas, mediante la clasificación difusa de píxeles. El algoritmo se definió modificando la función objetivo del algoritmo Fuzzy CMeans (FCM). Dicha modificación es para incluir un campo multiplicador,  $m(i, j)$ , que permita que los centroides de cada clase varíen a lo largo de la imagen; este campo multiplicador se obtiene para cada iteración mediante una técnica multinivel [Briggs, 1987a] y [Briggs, 1987b] que mejora la eficiencia del algoritmo.

La función objetivo del algoritmo Adaptive Fuzzy C-Means, mostrada a continuación, toma como base la del Fuzzy C-Means y añade un término multiplicador  $m(i, j)$  que modela las variaciones en el brillo de la imagen, manteniendo así las ventajas del algoritmo Fuzzy C-Means, a la vez que es aplicable a imágenes con intensidades no homogéneas.

$$\begin{aligned}
J_{AFCM} = & \sum_{ij} \sum_{k=1}^C u_k(i, j)^2 \| y(i, j) - m(i, j)v_k \|^2 \\
& + \lambda_1 \sum_{ij} \left( (D_i * m(i, j))^2 + (D_j * m(i, j))^2 \right) \\
& + \lambda_2 \sum_{ij} \left( (D_{ii} * m(i, j))^2 + 2(D_{ij} ** m(i, j))^2 \right) + (D_{jj} * m(i, j))^2
\end{aligned}$$

Donde  $D_i$  y  $D_j$  son los operadores de diferencias finitas a lo largo de filas y columnas y  $D_{ii} = D_i * D_i$ ,  $D_{jj} = D_j ** D_j$  y  $D_{ij} = D_j * D_j$  son las diferencias finitas de segundo orden. Los símbolos  $*$  y  $**$  denotan los operadores de convolución en el dominio discreto, en una y dos dimensiones, respectivamente. Los dos últimos operandos, controlados por los parámetros  $\lambda_1$  y  $\lambda_2$  son las expresiones de regularización de primer y segundo orden operados con el campo multiplicador. La expresión reguladora de primer orden, penaliza los campos multiplicadores que presentan grandes variaciones. La de segundo orden también penaliza esta alta variación, pero especialmente penaliza los campos multiplicadores que presentan discontinuidades.

Si se asume que las funciones de pertenencia  $u_k(i, j)$  y los centroides  $v_k$  representativos de cada clase son conocidos, el campo multiplicador que minimiza la función objetivo será aquel que acerque los centroides a los datos, y a la vez varíe lenta (lo que controlan ambas expresiones de regularización) y suavemente (restricción mantenida por la expresión reguladora de segundo orden). Sin ambas expresiones reguladoras, siempre se podría encontrar un campo multiplicador que hiciese cero la función objetivo. Los

parámetros  $\lambda_1$  y  $\lambda_2$  se deben ajustar de acuerdo a la magnitud y suavidad de las variaciones de intensidad a lo largo de la imagen. Para imágenes homogéneas se emplean grandes valores para  $\lambda_1$  y  $\lambda_2$ .

La función objetivo del Adaptive Fuzzy C-Means puede minimizarse de una manera similar a la del algoritmo Fuzzy C-Means. Se realiza la primera derivada con respecto a  $u_k(i, j)$ ,  $v_k$  y  $m(i, j)$  e igualando a cero dichas ecuaciones, se obtienen las tres condiciones necesarias para que la función  $J_{AFCM}$  sea mínima.

En la Figura 17 aparece el ordinograma del algoritmo AFCM completo; considerando que el algoritmo converge si el máximo cambio en las funciones de pertenencia para todos los píxeles entre dos iteraciones consecutivas es menor que el umbral fijado.

La función de pertenencia que ha sido implementada para este algoritmo es:

$$u_k(i, j) = \frac{\|y(i, j) - m(i, j)v_k\|^{-2}}{\sum_{l=1}^C \|y(i, j) - m(i, j)v_l\|^{-2}}, \forall(i, j), k = 1, \dots, C$$

El cálculo de los nuevos centroides en cada iteración se realiza utilizando la expresión

$$v_k = \frac{\sum_{ij} u_k(i, j)^2 m(i, j) y(i, j)}{\sum_{ij} u_k(i, j)^2 m(i, j)^2}, k = 1, \dots, C$$

Y para obtener el nuevo campo multiplicativo se resuelve para  $m(i, j)$  la siguiente ecuación:

$$\begin{aligned} & y(i, j) \sum_{k=1}^C u_k^2(i, j) v_k \\ &= m(i, j) \sum_{k=1}^C u_k^2(i, j) v_k^2 \\ &+ \lambda_1 (m(i, j) ** H_1(i, j)) \\ &+ \lambda_2 (m(i, j) ** H_2(i, j)) \end{aligned}$$

Donde:  $H_1(i, j) = D_i * \tilde{D}_i + D_j * \tilde{D}_j$ , y  $H_2(i, j) = D_{ii} * + 2(D_{ij} ** D_{ij}) + D_{jj} * \tilde{D}_{jj}$ .

Para resolver la ecuación iterativamente, empleando los esquemas de Jacobi o Gauss-Seidel, se requiere un elevado número de iteraciones para alcanzar la convergencia; por ello se emplea un algoritmo multigrad. Este último resuelve la ecuación de una forma más eficiente, trabajando con varias resoluciones. Y así, la complejidad computacional se reduce drásticamente, reemplazando las iteraciones que normalmente tendrían lugar en un

grid fino, con iteraciones en un grid de menor resolución. Los detalles del método se pueden consultar en [DL Pham, 1999a], cuyo ordinograma está en la Figura 17.

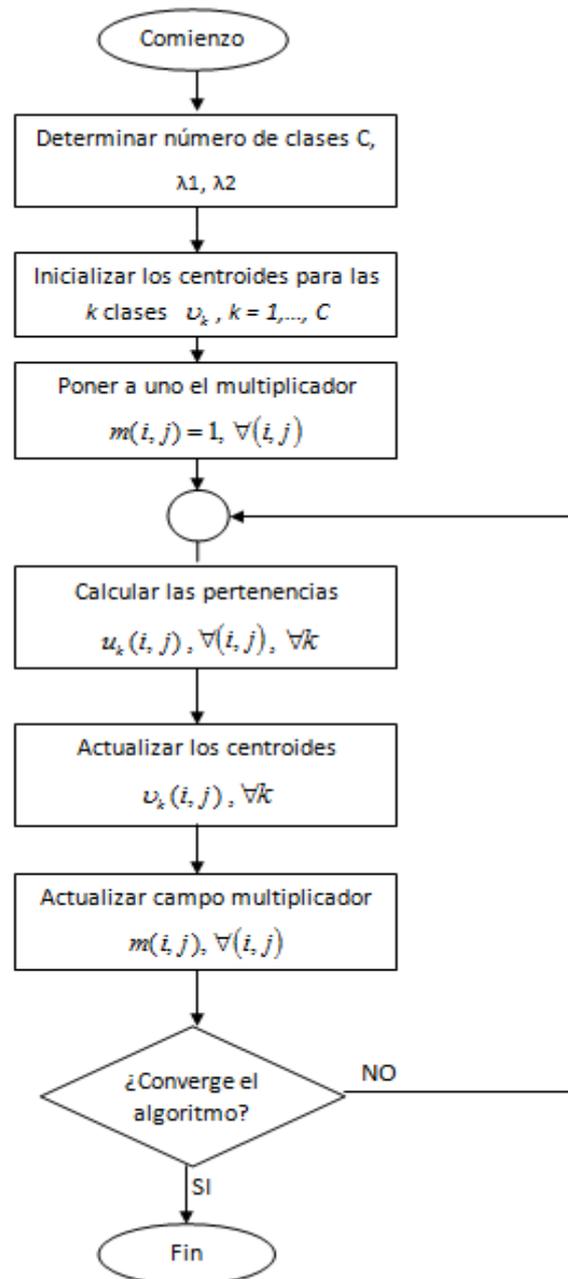


Figura 17. Ordinograma del algoritmo AFKM.

### 2.3.5 Spatially Kernelized Fuzzy C-Means

Es más sencillo determinar la forma de los clústeres en un espacio de más dimensiones, mapear los datos a ese espacio, hacer los cálculos y volver a la dimensión actual, pero esto es un proceso muy costoso. Sin embargo, existen las denominadas funciones kernel que

mediante un “truco” matemático devuelven el valor de la distancia entre dos puntos en un espacio de mayor dimensión.

El algoritmo Spatially Kernelized Fuzzy C-Means (SKFCM) [Zhang, 2004a] está basado en el Fuzzy C-Means (FCM). Pero en lugar de emplear la distancia euclídea emplea una métrica de distancia basada en un kernel. Es decir, el algoritmo calcula la distancia entre los datos y los centroides mediante una función kernel, que permite aumentar la insensibilidad al ruido que puedan tener las imágenes.

En estos métodos se realiza de forma implícita una translación no lineal del espacio de datos al espacio de características,  $\Phi: X \rightarrow F(x \rightarrow \Phi(x))$ , donde un conjunto de datos  $x_1, \dots, x_n \subseteq X$  de un espacio de baja dimensión se mapea a un espacio de características de una dimensionalidad mucho mayor; ello tiene por objetivo convertir un problema no lineal en el espacio de entrada (espacio original de los datos) en otro potencialmente lineal en el espacio de mayor dimensionalidad para facilitar su resolución. Computacionalmente es un procedimiento bastante costoso, pues en primer lugar se debe calcular  $\Phi(x)$ , transformación no lineal y de altas dimensiones, y después calcular el producto interno. Sin embargo, gracias a las funciones kernel, se puede calcular el producto interno, en el espacio de características directamente desde el espacio de entrada, sin necesidad de calcular el mapeo  $\Phi$ . Puesto que un kernel, en el espacio de características, se puede representar como una función  $K$  de la siguiente forma  $K(x, y) = \langle \Phi(x), \Phi(y) \rangle$  donde  $\langle \Phi(x), \Phi(y) \rangle$  representa el producto interno.

Así pues, una característica importante es que ese producto interno entre  $\Phi(x)$  y  $\Phi(y)$  se puede calcular implícitamente en el espacio de características  $F$ , directamente desde el espacio de entrada sin necesidad de calcular explícitamente ese mapeo  $\Phi$ . Ahora bien, el espacio de entrada debe ser un conjunto no vacío.

Las funciones kernel más utilizadas son las funciones gaussianas, polinómicas, sigmoidales, de base radial, y tangente hiperbólica.

La función objetivo del algoritmo Spatially Kernelized Fuzzy C-Means toma como base la del Fuzzy C-Means; dicha función es modificada sustituyendo la distancia Euclídea, entre el valor de intensidad de cada punto y el valor de los centroides, por una distancia en un espacio de más alta dimensionalidad [Zhang, 2004].

$$J_{SKFCM} = \sum_{ij} \sum_{k=1}^C u_k(i, j)^m \|\Phi(x(i, j)) - \Phi(v_k)\|^2$$

Donde  $\Phi$  es un mapeado no lineal implícito. Se aplica la sustitución kernel y se obtiene:

$$\begin{aligned} \|\Phi(x(i, j)) - \Phi(v_k)\| &= (\Phi(x(i, j)) - \Phi(v_k))^T (\Phi(x(i, j)) - \Phi(v_k)) \\ &= \Phi(x(i, j))^T \Phi(x(i, j)) - \Phi(v_k)^T \Phi(x(i, j)) - \Phi(x(i, j))^T \Phi(v_k) + \Phi(v_k)^T \Phi(v_k) \\ &= \mathbf{K}(x(i, j), x(i, j)) + \mathbf{K}(v_k, v_k) - 2\mathbf{K}(x(i, j), v_k) \end{aligned}$$

Utilizando una función kernel de base radial Gaussiana, donde  $\mathbf{K}(x, x) = 1$ , la función objetivo se puede expresar como:

$$J_{KFCM} = \sum_{ij} \sum_{k=1}^C u_k(i, j)^m (1 - \mathbf{K}(x(i, j), v_k))$$

Además, a la función objetivo  $J_{KFCM}$  se le añade un operando de penalización, que incluye la información espacial del vecindario, para que la clasificación de cada punto no solo dependa de su valor de intensidad sino también de la de sus píxeles vecinos en la imagen. Considerando como píxeles vecinos los que forman parte de la ventana que rodea al punto en la imagen.

Una vez añadido el operando de penalización, cuya finalidad es que actúe como un regulador para ayudar a compensar el ruido en la imagen, la función objetivo del algoritmo Spatially Kernelized Fuzzy C-Means queda como se indica a continuación:

$$J_{SKFCM} = J_{KFCM} + \frac{\alpha}{N_R} \sum_{ij} \sum_{k=1}^C u_k(i, j)^m \sum_{r \in N_{ij}} (1 - u_k(r))^m$$

Sustituyendo  $J_{KFCM}$  por su expresión, se obtiene finalmente:

$$\begin{aligned} J_{SKFCM} &= \sum_{ij} \sum_{k=1}^C u_k(i, j)^m (1 - \mathbf{K}(x(i, j), v_k)) \\ &+ \frac{\alpha}{N_R} \sum_{ij} \sum_{k=1}^C u_k(i, j)^m \sum_{r \in N_{ij}} (1 - u_k(r))^m \end{aligned}$$

Donde  $N_{ij}$  representa la vecindad de  $x(i, j)$ , es decir, el conjunto de píxeles que existen en una ventana en torno a  $x(i, j)$  sin incluirlo.  $N_R$  es el número de elementos de la ventana  $N_{ij}$ . El parámetro  $\alpha$  controla el efecto de la penalización y su valor debe ser mayor que cero y menor que uno.

La función objetivo del Spatially Kernelized Fuzzy C-Means se minimiza cuando se asignan altos valores de pertenencia a píxeles cuyo valor de intensidad está próximo al del centroide para el clúster dado, y bajos valores a intensidades de píxeles muy alejados.

Derivando la función objetivo con respecto a  $u_k(i, j)$  y a  $v_k$ , e igualándolas a cero, al igual que se ha hecho en el caso del AFCM se obtienen las dos condiciones necesarias para que  $J_{SKFCM}$  sea mínima.

En la Figura 18 se muestra el ordinograma del algoritmo Spatially Kernelized Fuzzy C-Means.

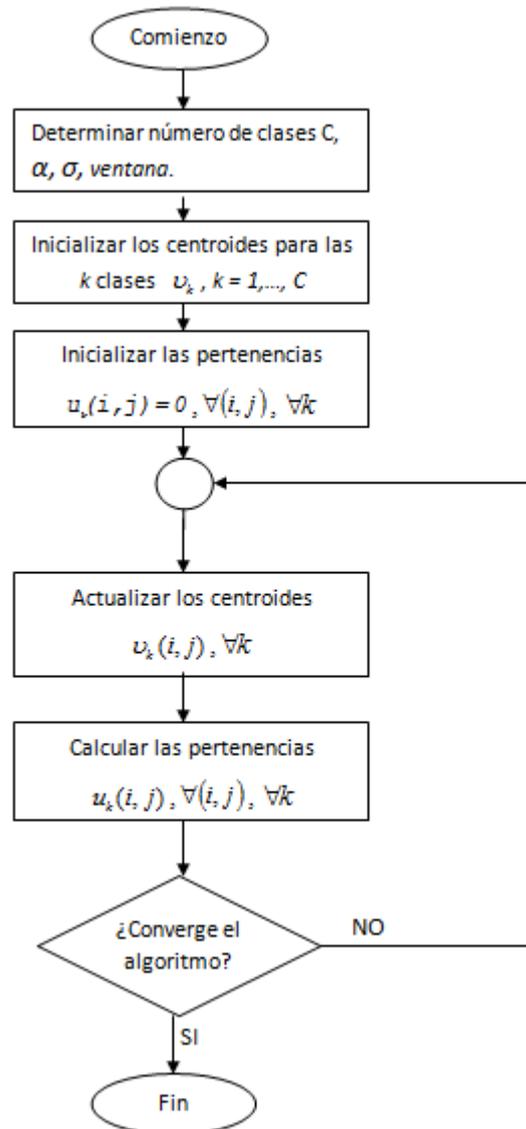


Figura 18. Ordinograma del algoritmo SKFCM.

Se considera que el algoritmo converge si el máximo cambio en las funciones de pertenencia para todos los píxeles entre dos iteraciones consecutivas es menor que el umbral fijado.

La función de pertenencia implementada para este algoritmo está dada por la expresión:

$$U_k(i, j) = \frac{\left( (1 - K(x(i, j), v_k)) + \left( \alpha \sum_{r \in N_k} (1 - u_k(r))^m / N_R \right) \right)^{-1/(m-1)}}{\sum_{l=1}^c \left( (1 - K(x(i, j), v_l)) + \left( \alpha \sum_{r \in N_k} (1 - u_l(r))^m / N_R \right) \right)^{-1/(m-1)}}$$

La expresión utilizada para el cálculo del nuevo centro de cada clúster en cada iteración es:

$$v_k = \frac{\sum_{ij} u_k^m(i, j) K(x(i, j), v_k) x(i, j)}{\sum_{ij} u_k^m(i, j) K(x(i, j), v_k)}$$

Y al igual que en el caso de los algoritmos Fuzzy C-Means y Adaptive Fuzzy C-Means se considera que el algoritmo converge cuando se cumple que la diferencia entre los centroides de cada clúster entre dos iteraciones consecutivas es inferior a épsilon ( $\|v_{actual} - v_{anterior}\| < \varepsilon$ ).

## 2.4 Sistemas CAD para imágenes de Pulmón

Un CAD (Computer-Aided Diagnosis) es un sistema de ayuda al diagnóstico asistido por ordenador, su diseño es un campo de aplicación del Análisis de Imágenes. La Radiología, Ecografía, Tomografía Computarizada, Resonancia Magnética, Mamografía, Medicina Nuclear, Densitometría, Angiografía diagnóstica y terapéutica son tecnologías, dentro del área de Diagnóstico Médico, en las que se utiliza el ordenador como herramienta de apoyo al diagnóstico de un gran número de patologías y posteriores intervenciones.

Lo habitual es que para diagnosticar un cáncer se realicen entre otras pruebas “test de Imagen” como radiografías, ecografías, resonancias magnéticas, etc. Sin embargo, las imágenes realizadas no son fáciles de interpretar; lo cual origina que en ellas no se detecten un porcentaje importante de tumores. Resulta complicado detectar lesiones debido a múltiples factores como por ejemplo la poca calidad de la imagen, la ocultación de objetos en la imagen debido al paso de 3D a 2D, las lesiones de reducido tamaño difíciles de ver por el ojo humano, la subjetividad en la interpretación, el cansancio y la falta de concentración, la inexistencia de una búsqueda sistemática y exhaustiva en la imagen, etc.

Se ha demostrado que la habilidad, para localizar patrones que tienen cierta similitud con el fondo de la imagen, requiere una búsqueda sistemática por toda la imagen [Treisman, 1985]. Sin embargo, estudios que analizaron el movimiento de ojo de radiólogos que exploraron una imagen, durante el análisis de una mamografía, demostraron que dicho

movimiento no es sistemático [Krupinski, 1996]. A consecuencia de todo esto, los “test de Imagen” utilizados para el diagnóstico precoz del cáncer pueden tener limitaciones graves. De hecho, se calcula que entre un 10% y un 25% de los tumores pasan inadvertidos a los ojos del radiólogo [Azpizarte, 2006], [Cheng, 2005] por ello, los sistemas de diagnóstico asistido por computador CAD, pueden ser una gran ayuda en la interpretación de este tipo de imágenes. Además, permiten a los radiólogos, reducir el tiempo empleado en el análisis de las secuencias de tomografías computarizadas [Beigelman-Aubry, 2007].

Estos sistemas de ayuda al diagnóstico asistido por ordenador utilizan técnicas de procesamiento de imágenes. Emplean algoritmos para analizar las imágenes originales detectando regiones de interés (ROI), y mejoran la presentación e interpretación de las mismas. Los resultados que producen se utilizan como una segunda opinión, previa a la decisión final del médico, dirigiendo la atención hacia zonas de la imagen sospechosas de contener alguna anomalía.

En la Figura 19 se muestra que el objetivo final de un sistema CAD es mejorar el rendimiento diagnóstico del experto humano encargado de interpretar la imagen médica, dicho rendimiento puede medirse mediante los valores de *sensibilidad* ( $S$ ) y *especificidad* ( $E$ ) conseguidos

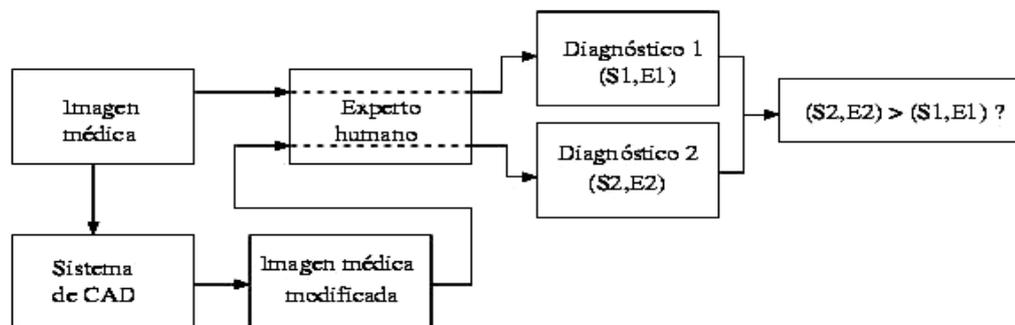
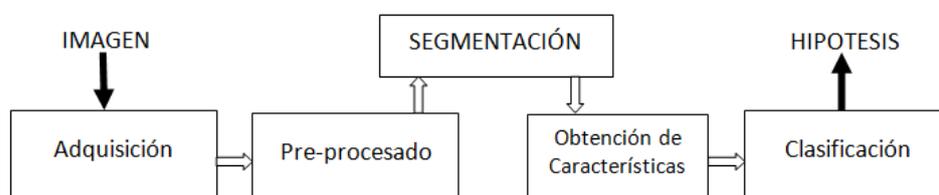


Figura 19. Diagnóstico con y sin la ayuda de un sistema de CAD.([LLobet Azpizarte, 2006]).

La investigación en estos sistemas para el análisis de la imagen pulmonar tuvo su comienzo en la década de 1970, con el desarrollo de técnicas específicas para el análisis de radiografías pulmonares [vanGinneken, 2001]. Hoy en día, con el carcinoma de pulmón, se utilizan sistemas de Diagnóstico Asistido por Computador para ayudar a establecer las características de las lesiones sospechosas de malignidad, y determinar su estadificación; todo ello permite aplicar un tratamiento, control y seguimiento al paciente.

En la Figura 20 se indica la estructura utilizada por la mayoría de los sistemas CAD, para detección de nódulos pulmonares. Tras la adquisición de una imagen se realiza un pre-procesado, mediante el cual se realiza y acondiciona la imagen; en el caso de esta tesis, en concreto, se identifica la región que ocupan los pulmones izquierdo y derecho y se eliminan el resto de objetos de la imagen. Esta etapa tiene un papel importante en el proceso de detección de nódulos pulmonares ya que permite aumentar la precisión y reducir el coste computacional en las siguientes etapas del proceso. Tras la segmentación se obtiene la matriz de características representativa de la imagen. Características que permiten distribuir los puntos de la imagen, entre las distintas clases o categorías que ocupen regiones en el espacio n-dimensional, lo más compactas y disjuntas posible.



**Figura 20. Modelo de sistema CAD.**

Los primeros trabajos sobre CAD han sido publicados en la década de 1980, en particular en el campo de la mamografía con la detección de microcalcificaciones [Chan, 1987], en angiografía con la determinación del tamaño de los vasos [Fujita, 1987], y en radiografía de tórax con la detección de nódulos pulmonares [Giger, 1988].

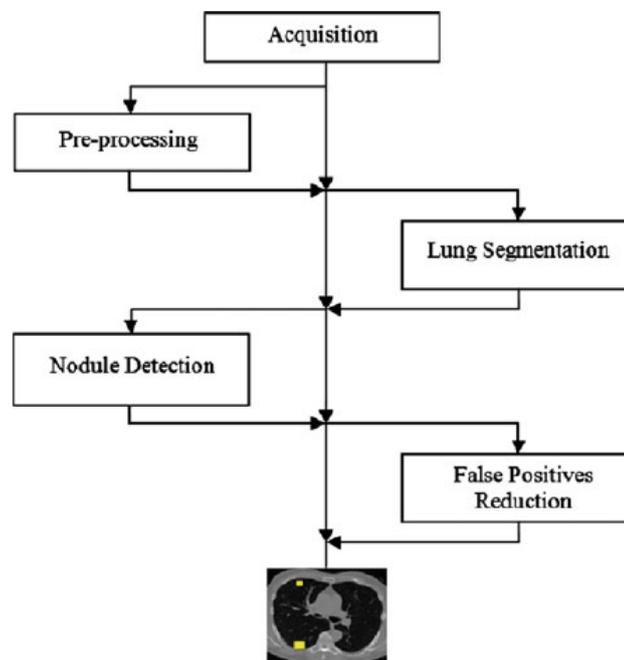
Las mejoras tecnológicas introducidas en los tomógrafos han incrementado la producción de imágenes durante un escaneo, dando lugar a “una explosión de datos” [Rubin, 2000] cuya interpretación es facilitada utilizando técnicas de procesamiento de imágenes. Causa y consecuencia de esto es la intensificación producida en los últimos años, del número de ponencias y artículos de investigación sobre temas relacionados con el sistema CAD de tórax.

A continuación se cita una pequeña revisión bibliográfica sobre la detección y diagnóstico de nódulos pulmonares en imágenes de CT de tórax.

Las imágenes empleadas en los trabajos publicados provienen de bases de datos de imágenes públicas como LIDC (Lung Image Database Consortium), ELCAP (Early Lung Cancer Action Program Public Lung Image Database), MID (Medical Image Database) y

privadas (obtenidas en los hospitales con los que trabajan o colaboran los distintos autores).

En [Li, 2007] se expone una revisión de los trabajos publicados hasta 2004 sobre la detección de nódulos pulmonares. Hay muchos artículos donde se describen métodos para detectar, de forma automática o semiautomática, nódulos pulmonares. De forma genérica la Figura 20, vista anteriormente, ilustra las fases que estos métodos utilizan. Hay investigadores que incluyen todas las fases y otros que sólo emplean un subconjunto de ellas. La Figura 21 muestra la estructura más utilizada por los sistemas de detección de nódulos pulmonares.



**Figura 21. Estructura genérica de los métodos de detección de nódulos pulmonares [Lee, 2010].**

La exactitud es transcendental en la etapa de pre-procesado de un sistema CAD pues, un proceso no riguroso representa entre el 5% y el 17% de errores en la clasificación final [Armato, 2004]. Por tanto para mejorar la calidad, reduciendo ruido y artefactos, y con ello mejorar la interpretación de las imágenes adquiridas, se utilizan distintos tipos de filtros: filtros paso-bajo con parámetros gaussianos [Garnavi, 2005], filtros de mediana [Kim, 2007], suavizados gaussianos [Pu, 2008], [Gori, 2006] y [Retico, 2008]. Con un “ring average filter” se suprimen nódulos y se enfatizan con un “matched filter” [Arimura,

2004]. Se enfatizan las imágenes con un filtro “Laplacian of Gaussian” [Diciotti, 2008], [Sluimer, 2003] y las regiones de la imagen con filtros morfológicos [Kostis, 2003], [Bae, 2005]. Se combinan filtros cilíndricos con filtros esféricos [Chang, 2004]. Para dar “textura” a los distintos tipos de nódulos se utilizan filtros anisotrópicos [Kostis, 2003]. Los filtros mejorados de escalas múltiples son empleados por [Li, 2004] y los filtros basados en la enfatización “dot-like” por [Fetita, 2003], [Sun, 2007].

Para aumentar la fiabilidad, exactitud y precisión, y a la vez disminuir el coste computacional, en la detección de los nódulos pulmonares, se identifican los pulmones y se eliminan los demás elementos de la imagen. Las técnicas de segmentación para separar el pulmón de los tejidos envolventes tanto en [Leader, 2003] como en [Armato, 2004] están basadas en reglas; las cuales consideran el pulmón como una región de tonalidad oscura en la imagen, al ser una bolsa de aire, y así distinguirlo de las regiones a su alrededor, cuya tonalidad es diferente [Leader, 2003]. Sitúan el tórax, mediante la umbralización de niveles de gris y el análisis de componentes, y encuentran la tráquea buscando zonas circulares llenas de aire. Cuando ya están ubicados los pulmones y las vías respiratorias, se separan los pulmones y se eliminan la tráquea y los bronquios; realizando operaciones morfológicas de dilatación y erosión para obtener los pulmones sin huecos y con bordes suaves [Sluimer, 2006].

Para aislar las regiones de nódulos en la imagen [Aoyama, 2003] utiliza programación dinámica; para segmentar los nódulos [Kawata, 2000] emplea modelos deformables; también [Kim, 2003] emplea un modelo deformable, y además combinaciones de umbral y relleno de regiones.

Las técnicas de segmentación más utilizadas para encontrar candidatos a nódulos son: algoritmos “snake” [Itai, 2007], modelos multi-escala gaussianos ajustados para la segmentación de nódulos elipsoidales [Okada, 2005], aproximaciones basadas en morfología matemática para eliminar de la imagen aquellas partes que son irrelevantes ([Kostis, 2003]; [Kuhnigk, 2006]), umbralización [Wei, 2002],[Paik, 2004]; umbralización y operaciones morfológicas para reducir el volumen parcial y “beam hardening effects” [Kanawaza, 1998], detección de bordes Canny y “morphological closing”[Lee,2001], regiones de crecimiento para segmentar tejido pulmonar [Diciotti, 2008]; métodos de conectividad fuzzy centrados en la fuerte afinidad entre la semilla y cada punto específico [Udupa, 2002]. También se combinan regiones de crecimiento con métodos de conectividad fuzzy [Dehmeshki, 2008]; se utiliza umbralización simple para separar los nódulos candidatos de la imagen de fondo [Giger, 1994], [El\_Baz, 2007]; se hace uso de

umbrales basados en el histograma para separar el pulmón de los demás componentes del tórax [Wang, 2007] y [Retico, 2008]. Se emplea umbralización de niveles de gris, seguida de operaciones morfológicas, y crecimiento de regiones 3D, para separar tres clases de nódulos pulmonares [Fetita, 2003]; se realiza la segmentación automática de los nódulos pulmonares utilizando crecimiento de regiones [Sun, 2007].

Uno de los métodos utilizados para encontrar regiones de interés (ROI) es disponer de los valores de densidad de los píxeles para cada corte de un estudio, y con la unión de todos los cortes conseguir una ROI de 3D que se compara con una plantilla o modelo para identificar los nódulos [Osman, 2007]. También se utilizan filtros de realce, para objetos esféricos, y clasificación neuronal, en regiones de vóxeles seleccionadas, para reducir falsos positivos [Retico, 2008].

En la detección de nódulos pulmonares se comprueba si el patrón está presente en la imagen, y se identifica su ubicación. Aunque hay investigadores que se basan únicamente en los resultados de la segmentación del pulmón; otros utilizan ambas cosas para detectar los nódulos con mayor precisión [Lee, 2010].

Una vez que en las imágenes son detectados los nódulos candidatos, se calculan sus características; dichas características se utilizan de entrada a los clasificadores que determinan si el objeto candidato pertenece al grupo de nódulos, o al grupo de no nódulos. Las características más utilizadas para la clasificación son la densidad de los píxeles, la descripción de las formas, el tamaño, y la relación espacial.

Como clasificadores se utilizan diversas técnicas: basadas en reglas, análisis lineal discriminante(LDA), utilización de plantillas (Template Matching), técnicas de clusterización, máquinas de soporte vectorial (SVM), redes neuronales, clasificadores bayesianos, y algoritmos genéticos. En los artículos publicados se proponen clasificadores automáticos basados en reglas para clasificar nódulos y no nódulos [Li, 2008]. También, con un enfoque basado en reglas se caracteriza el estado de malignidad de los nódulos [Kostis, 2003]. Se utiliza un clasificador basado en el análisis discriminante lineal [Kawata, 1998], [Kawata, 2000]. Para clasificar los nódulos candidatos, producidos por umbralización múltiple, se emplean clasificadores LDA [Armato, 1999]; los cuales se basan en características de niveles de gris, características de bordes, características morfológicas, y características clínicas para discriminar entre nódulos benignos y malignos [Aoyama, 2003]. También se utiliza el LDA basado en distribución de la distancia Mahalanobis para clasificar el tipo de nódulo [Kim, 2007]. Se desarrolla un clasificador basado en texturas que clasifican nódulos, basándose en características tales como el

tamaño y la forma [Kim, 2003]. Se utilizan clasificadores LDA y SVM [Sluimer, 2003], y redes de neuronas artificiales para detectar nódulos [Arimura, 2004]. Para detectar patrones en nódulos benignos y malignos se desarrollan múltiples ensayos con redes de neuronas artificiales [Suzuki, 2005]. Y para clasificar los patrones de nódulos y no nódulos se entrena un clasificador bayesiano [McCulloch, 2004]. En [Lee, 2001] se combinaron algoritmos genéticos, y Template matching. En [Kawata, 2000] se propone un clasificador discriminante lineal, que utiliza el resultado del algoritmo CMeans sobre diferentes características obtenidas del histograma topológico sobre un grupo de datos de nódulos benignos y malignos. [Lee, 2010] desarrolló una clasificación de conjunto utilizando clustering para mejorar el rendimiento de la clasificación de los nódulos pulmonares.

Una vez finalizada la etapa de clasificación, y para ayudar a delimitar y reducir el trabajo de los radiólogos, muchos de los trabajos publicados sobre detección de nódulos tienen por objetivo la reducción de falsos positivos. Entre dichos trabajos destacan un clasificador de reglas para la reducción de falsos positivos [Armato, 2004], y también clasificadores de reglas junto con redes neuronales [Arimura, 2004]. Los falsos positivos se eliminan donde hay valores de intensidad mayores que las de los nódulos vecinos [Osman, 2007]; se implementan test de esfericidad para reducirlos [Chang, 2004]. Y se utilizan algoritmos genéticos [Boroczky, 2006], junto con plantillas [Lee, 2001]. Además de analizar vóxeles, del vecindario en 3D de la región de interés, mediante una aproximación neuronal para determinar la probabilidad de falsos positivos [Gori, 2006] y [Retico, 2008]).

En general, todos los métodos mencionados anteriormente pueden mejorarse utilizando una reducción de falsos positivos adecuada; aquellos sistemas que utilizan algoritmos de detección multi-etapa son los que obtienen mejor rendimiento [Lee, 2010].

En los últimos años, para ayudar al diagnóstico del cáncer de pulmón, se han desarrollado múltiples métodos automáticos. A nivel comercial los primeros sistemas CAD fueron desarrollados por la compañía R2 Technology, aparecieron en EEUU en 1988. Estos primeros sistemas se utilizaban para la detección de lesiones en mamografía. Posteriormente, a partir del año 2001, las empresas Deus Technologies y Mitsubishi Space Software han desarrollado distintos sistemas CAD, con el objetivo de facilitar la detección de nódulos pulmonares en radiografías e imágenes de tórax. Actualmente, hay varios miles de sistemas CAD utilizados para ayudar a los radiólogos en la detección precoz del cáncer, en muchos hospitales, clínicas y centros de detección de todo el mundo. Destacando casas comerciales como [Mediantechologies, 2011] que están investigando y desarrollando

diferentes tipos de sistemas de CAD para la detección, evaluación y seguimiento de las lesiones identificadas en imágenes de Tomografías Computarizadas de tórax.

La investigación y desarrollo de sistemas CAD sobre TC, además de realizarse con el objetivo de detectar y realizar un seguimiento de los nódulos pulmonares, también se utiliza para otras finalidades sirvan de ejemplo los ensayos terapéuticos de enfisema pulmonar [Brown, 2007].

### 3 METODOLOGÍA

En este capítulo se presenta el método de trabajo empleado. Comenzando por la descripción del corpus de las imágenes utilizadas para la realización de la parte de pruebas y de la parte de validación de valores de parámetros óptimos, describiendo a continuación las fases del algoritmo de preprocesado empleado; sigue con la descripción de la metodología empleada para la obtención de las mejores combinaciones de valores paramétricos. Se explica la base teórica utilizada que permite definir los índices para cuantificar el rendimiento de las segmentaciones, y así poder analizar y comparar los resultados de las pruebas, y con ello determinar cuáles combinaciones de valores de parámetros son óptimas.

#### *3.1 Contexto de los estudios: imágenes de la Base de Datos LIDC*

Existen diversas bases de datos de imágenes CT de pulmón de carácter público y privadas que permiten evaluar el rendimiento de las pruebas de segmentación mediante un análisis retrospectivo.

Actualmente entre las bases de datos públicas de CT de pulmón digitalizadas destacan las siguientes:

- Early Lung Cancer Action Program (ELCAP) Public Lung Image Database.
- ELCAP Public Lung Database to Address Drug Response.
- Lung Image Database Consortium (LIDC) in National Imaging Archive.
- Imaging Database Resources Initiative (IDRI)
- Medical Image Database.

Todas ellas están supervisadas y en ellas se especifican, de algún modo, las regiones o píxeles correspondientes a nódulos pulmonares.

En la realización de este trabajo se ha utilizado un conjunto de Tomografías Axiales Computarizadas (TAC) pulmonares de alta resolución provenientes de la base de datos del Lung Image Database Consortium, conocido este consorcio por las siglas LIDC, la cual es accesible desde el National Biomedical Imaging Archive (NBIA). El objetivo del LIDC es el desarrollo de un repositorio de TAC pulmonares de referencia que sirvan para el desarrollo y evaluación de sistemas CAD que permitan la detección de nódulos

pulmonares. En su elaboración han colaborado cinco instituciones académicas de Estados Unidos: Cornell University; the University of California, Los Angeles; the University of Chicago; the University of Iowa and the University of Michigan [Armato, 2004].

Actualmente esta base de datos está formada por imágenes de tomografía computarizada, almacenadas en formato DICOM e informadas en archivos con formato XML.

Cada una de las imágenes fue analizada y anotada por cuatro expertos radiólogos en dos fases. En una primera fase realizaron una revisión ciega, en la que cada radiólogo analizó las imágenes independientemente de los demás y las registró anotando los nódulos encontrados. En una segunda fase las anotaciones realizadas en la primera fase fueron distribuidas entre todos los expertos participantes, los cuales analizaron nuevamente las imágenes. En esta segunda fase cada radiólogo decidía si incluía o ignoraba los nódulos indicados por los demás expertos, pudiendo incluso excluir nódulos que él mismo anotara en la primera fase. En los nódulos menores de tres milímetros, los radiólogos sólo marcaban un punto sobre el nódulo (el centroide); en los mayores de tres milímetros perfilaron su contorno [Armato, 2004].

En este trabajo únicamente han sido utilizadas imágenes con un nódulo certificado (NPS) habiendo descartado aquellas imágenes que no tenían un Nódulo Pulmonar Solitario.

### *3.2 Conjuntos de prueba y validación*

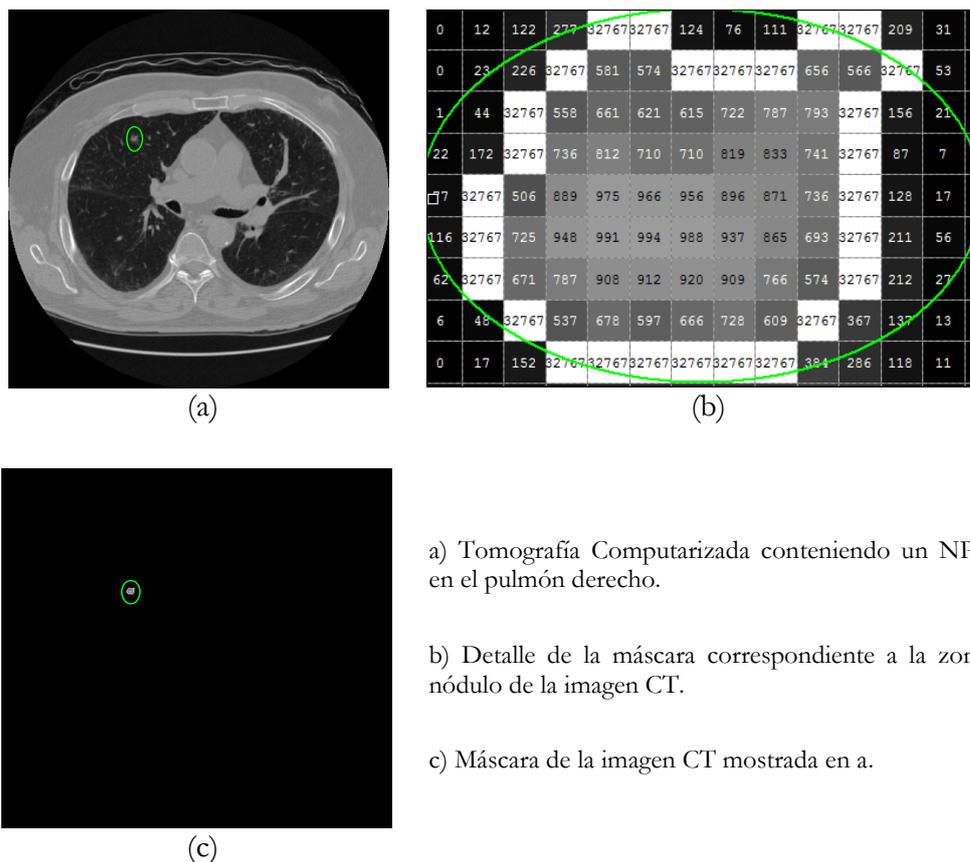
La Base Datos de Imágenes utilizada disponía inicialmente de 23 estudios, con los que se ha trabajado en esta tesis aplicando los algoritmos de segmentación objeto de estudio a las imágenes que tenían Nódulos Pulmonares Solitarios.

Las imágenes estaban almacenadas en dos carpetas, una denominada CT y otra identificada como PMAP con las máscaras de los NPS correspondientes a las tomografías computarizadas del primer directorio mencionado.

Se han analizado todas las máscaras correspondientes a las imágenes de CT disponibles originariamente y en ellas se encontraban los 23 estudios seleccionados de Nódulos Pulmonares Solitarios. De cada estudio han sido extraídas dos imágenes con NPS. En total fueron elegidas 46 imágenes, cada una de ellas conteniendo un NPS, divididas en dos grupos: uno para realizar las pruebas test de búsqueda y ajuste de valores paramétricos óptimos en los algoritmos de segmentación, y otro grupo distinto para la evaluación de los resultados. En dicha evaluación, con el segundo grupo de imágenes, se validan o valoran los valores de parámetros con los que se habían obtenido los mejores resultados sobre el

primer grupo de imágenes. Todas las imágenes de cada grupo (test y validación) pertenecen a estudios diferentes.

En la Figura 22a se muestra un ejemplo de imagen CT en la que hay un NPS circunscrito en un círculo verde; en el apartado *b* de dicha figura se exhibe, a nivel de píxel, un recorte de la máscara correspondiente a la región de interés marcada, presentando en el apartado *c* la imagen completa de la máscara.



a) Tomografía Computarizada conteniendo un NPS en el pulmón derecho.

b) Detalle de la máscara correspondiente a la zona nódulo de la imagen CT.

c) Máscara de la imagen CT mostrada en a.

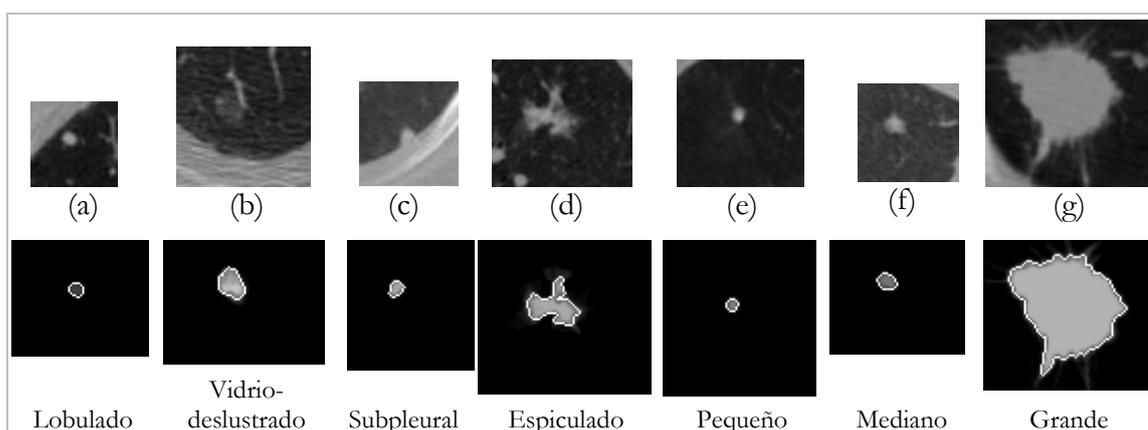
**Figura 22. Tipo de imagen y su correspondiente máscara de la BD inicial del LIDC.**

Los valores que aparecen en la máscara (Figura 22b) son números enteros positivos, resultantes de la valoración obtenida después de que los radiólogos (LIDC) hubiesen señalado cada píxel de la CT como perteneciente o no al nódulo. Si todos los expertos habían coincidido en anotar dicho píxel como perteneciente al nódulo entonces el valor dado al citado píxel es 1000, si ningún radiólogo del grupo lo había anotado como perteneciente al nódulo entonces el valor asignado es 0. Así los demás valores son el resultado de las coincidencias o no, entre los radiólogos, para los píxeles pertenecientes a la región de interés. En aquellos puntos donde el valor obtenido por los píxeles fue 500, se estableció el punto de corte en el que un píxel pertenece al nódulo, reasignando en la

máscara el valor máximo 32767 a dichos píxeles. La correspondencia entre los puntos de la CT y su máscara se realizó mediante las posiciones de las coordenadas de cada píxel.

La elección de las tomografías computarizadas se ha realizado con el objetivo de que dichas imágenes tuviesen nódulos que funcionasen de prototipos representando a las posibles tipologías de NPS, y así poder ver el efecto de los distintos algoritmos de segmentación sobre las imágenes diferentes. En este pequeño subconjunto hay imágenes con nódulos de diferente morfología (lobulados, espiculados), de distintas densidades (calcificados, en vidrio deslustrado), con diversa localización (sub-pleurales, centrales), y de varios tamaños (subcentimétricos y con diámetro mayor del centímetro). En la Figura 23 se muestran algunos ejemplos de nódulos y las correspondientes máscaras empleadas en este trabajo.

Las cabeceras DICOM de las imágenes indican que cada tomografía computarizada tiene una resolución espacial de 0,7031mm, entre centro de píxeles, y un tamaño de 512x512 píxeles, con una profundidad de 16 bits por píxel. Teniendo en cuenta dicha resolución y con la información de las máscaras correspondientes a cada CT se ha calculado el diámetro de cada uno de los NPS contenido en las imágenes. Así del total de tomografías utilizadas para las pruebas, el 78% corresponden a imágenes con nódulos de diámetro inferior a 10mm (40% menores de 6 mm, y 38% entre 6 y 10 mm), y el 22% son mayores de un centímetro y menores de 3 cm.



**Figura 23. Nódulos y sus máscaras con diferentes formas, localización, densidad y tamaño de las CT utilizadas.**

Se emplearon mayoritariamente imágenes con nódulos subcentimétricos, con el objetivo de focalizar la atención en detectar con más precisión, los píxeles correspondientes a los nódulos pulmonares en el primer estadio de cáncer. Ya que el primer estadio es una etapa

en la que es muy importante conocer el incremento de tamaño del NPS, en un periodo de tiempo de dos a seis meses desde la detección del susodicho nódulo pulmonar, y con ello poder determinar finalmente la malignidad del mismo [ELCAP, 2010].

El cáncer de pulmón es 1,5 veces más frecuente en el pulmón derecho que en el izquierdo [Conde, 2009], [Garland, 1961]. Para reflejar este hecho, las imágenes que han sido utilizadas en este trabajo están ubicadas el 65% en el pulmón derecho, y el 35% en el izquierdo.

Cuando los expertos buscan nódulos en una tomografía, comienzan valorando la existencia de nódulos en contacto con la superficie de la pleura visceral, costillas, o cisuras; los nódulos localizados en estas zonas son los denominados sub-pleurales. Se dispone de un 13% de nódulos sub-pleurales, que precisamente debido a su localización han sido los más difíciles de pre-procesar, por ello se han perdido en alguna imagen pixeles pertenecientes a la zona nódulo.

Las características de los bordes de los nódulos que indican malignidad incluyen: irregularidad, espiculación, y lobulación. Se han utilizado un 30% de imágenes con nódulos espiculados, y un 26% de CT con NPS de bordes lobulados. Los bordes suaves no siempre indican benignidad ya que, hasta una tercera parte de las lesiones malignas tienen un margen así, y muchas de éstas pueden ser metástasis [Gurney, 1993].

Los NPS más frecuentes son los sólidos, pero la probabilidad de malignidad más alta corresponde a los no sólidos [Conde, 2009]; en esta tesis, los nódulos pobremente definidos en vidrio deslustrado, corresponden a un 13% de las imágenes de CT utilizadas como material de trabajo.

Los grupos de imágenes seleccionados representan a los posibles nódulos que pueden encontrarse en las diversas bases de datos de tomografías helicoidales de alta resolución existentes a día de hoy, tanto en tamaño, localización, forma como en densidad de intensidad. Se han elegido así, para que los resultados que en principio se pudiesen obtener fuesen aproximadamente iguales a los que se hubiesen obtenido considerando un conjunto de CT más amplio.

### *3.3 Pre-procesado pulmonar*

El objetivo perseguido al realizar el pre-procesamiento es delimitar los pulmones, restringiendo el espacio de búsqueda. Aislado los pulmones se pretende eliminar la mayoría de elementos superfluos, que podrían afectar al análisis de los nódulos.

En la Figura 24 se ilustran las estructuras que componen una imagen de una tomografía computarizada de tórax de un corte concreto. En ella se puede observar que además del pulmón izquierdo y derecho se aprecian otros elementos como la columna, la caja torácica, corazón, esófago, venas y arterias; además de poder incluir elementos externos como la camilla y sábanas (véase Figura 25).

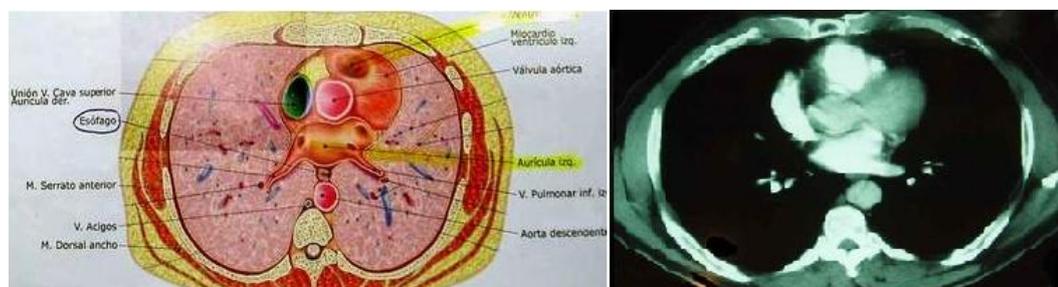


Figura 24. Dibujo aproximado a la tomografía mostrada [Medspain, 2011].

Cuando se analiza un histograma correspondiente a una tomografía computarizada, por ejemplo la Figura 25, se puede observar como los píxeles o puntos correspondientes a distintos tejidos tienen intensidades o brillos similares. En dicha figura se indican, con letras rojas numeradas, distintas zonas de la anatomía en la imagen CT de tórax, que se corresponden con el mismo intervalo de intensidades en el histograma (denotado por A y señalado con una flecha roja). En concreto, el NPS de la tomografía – denotado por a1- coincide en intensidad de pixel con el de tejido muscular y válvulas (indicados por a2, a3, a4, y a5), estos últimos tienen los mismos niveles de luminosidad y contraste que el NPS de la tomografía mostrada.

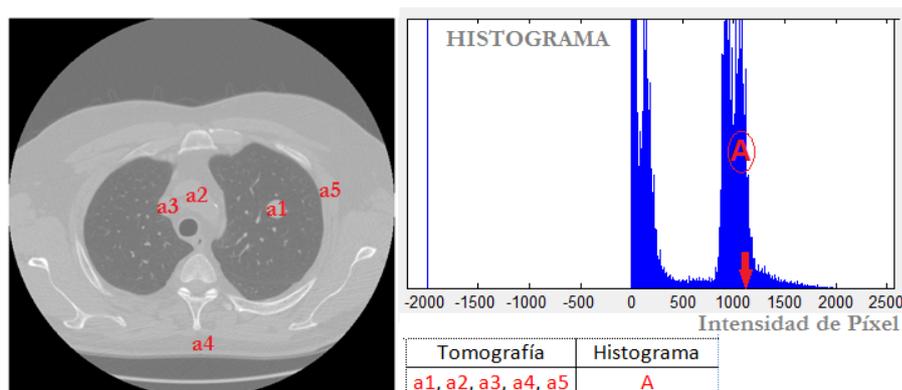
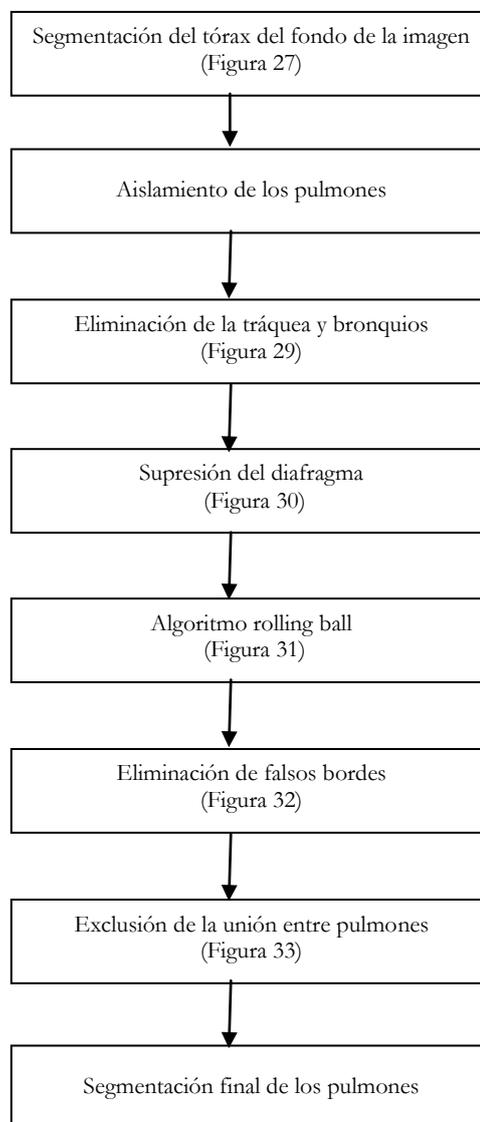


Figura 25. Asociación entre intensidades del histograma y objetos del CT.

Así pues, gran parte de la dificultad para detectar píxeles correspondientes a la región de interés, es debido a que el nódulo tiene unas unidades Hounsfield iguales y próximas a las

de otros elementos de la imagen. Se trata, pues, de eliminar aquellos elementos de las tomografías que distorsionen el proceso de identificación de los píxeles del nódulo. Por lo cual, es muy importante disminuir el espacio de búsqueda del clúster nódulo, y de esta forma reducir probables falsos positivos fuera de los tejidos pulmonares; y así también se contribuye a reducir el tiempo de ejecución de los algoritmos en la posterior segmentación de la imagen.

Se describe a continuación como se ha obtenido la región pulmonar, liberada de elementos que no son objeto de interés en este trabajo. En la Figura 26 se indican las diferentes etapas seguidas en la pre-segmentación de los pulmones, con cada tomografía.



**Figura 26. Fases del pre-procesado.**

Para aislar los pulmones se emplea como estrategia una segmentación incremental. La finalidad es separar y eliminar en cada fase un elemento no deseado de la imagen TC.

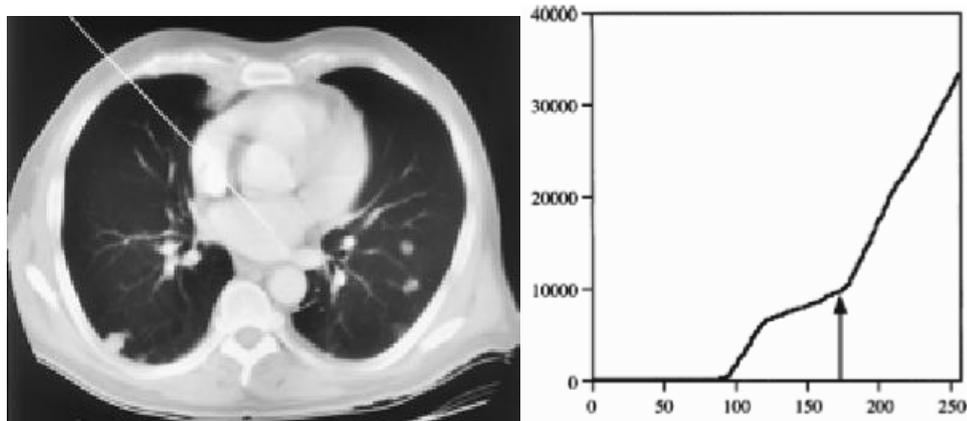


Figura 27. Separación del fondo y del tórax.

En una segunda fase se calcula el histograma de la zona resultante. Para la delimitación inicial de los pulmones, se utiliza como umbral el valor definido en el valle entre la moda blanca y oscura del histograma del tórax (Figura 28). Después de aislar los pulmones se elimina la tráquea y los bronquios. En la Figura 29 se muestra un ejemplo del resultado de este proceso.

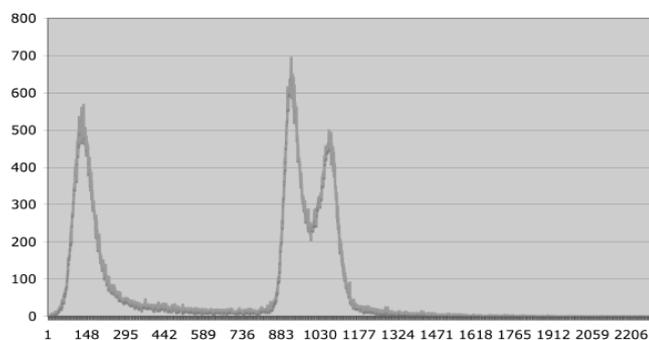


Figura 28. Histograma del tórax.

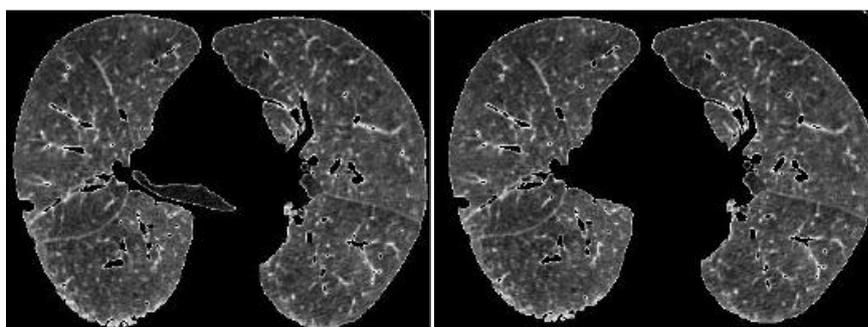


Figura 29. Eliminación de tráquea y bronquios.

Si la tomografía es de la parte inferior de los pulmones, en contacto con el diafragma, se procede a eliminarlo ya que su nivel de gris puede producir confusiones en el análisis del corte. Para ello, se detectan en las imágenes binarias obtenidas de los pulmones aquellas zonas no presentes con un diámetro mayor de 25 mm, normalmente queda eliminado en la segunda fase al tener un nivel de gris superior al de los pulmones, marcándose para que no aparezca en la imagen final; véase el resultado en la Figura 30.



Figura 30. Eliminación del diafragma.

En este punto del proceso, los pulmones se encuentran aislados; sin embargo durante alguna de las operaciones, debido a las diferencias en el nivel de gris, o por la localización de los cortes, se pueden haber eliminado zonas de interés como los nódulos situados en el borde de alguno de los pulmones, al presentar un nivel de gris muy similar. Para recuperarlos se aplica un algoritmo rolling ball, consistente en un elemento estructural circular para detectar cavidades en el borde de los pulmones, y en ese caso recuperarlas mediante un algoritmo de crecimiento de llenado; este último proceso se muestra en la Figura 31.

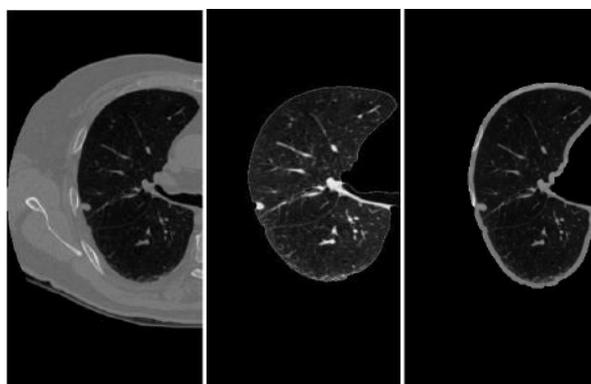
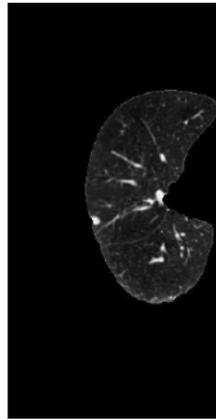


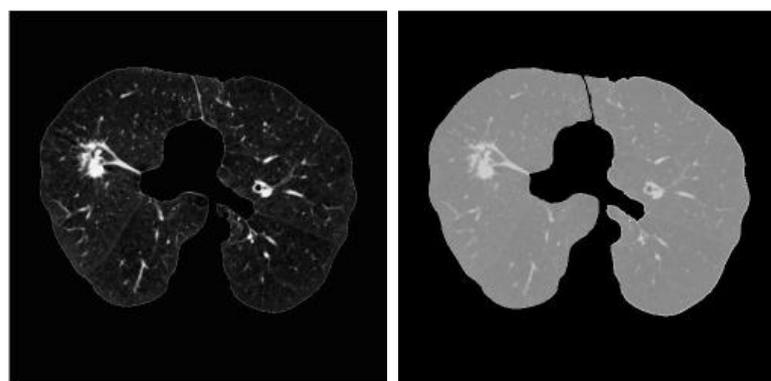
Figura 31. Resultado de aplicar el algoritmo de Rolling ball.

Al aplicar el algoritmo de Rolling ball, debido a la operación de dilatación, queda un falso borde de anchura mayor a un pixel. Se emplea, en la siguiente fase, una operación de erosión para eliminar el falso borde. Obteniéndose, tras este proceso, la imagen pre procesada como se indica en la Figura 32.



**Figura 32. Eliminación del falso borde.**

Después de la operación de rolling ball, en aquellas tomografías donde los pulmones derecho e izquierdo se encuentran muy próximos, aparece otro efecto no deseado, es la unión de ambos pulmones. Para separarlos, se traza una línea en la mitad de la imagen, y si a ambos lados de ella, se presentan valores superiores al nivel medio se utiliza un algoritmo de crecimiento de regiones para su marcación y eliminación. En la Figura 33 se muestra este último resultado.



**Figura 33. Separación de la unión de los pulmones.**

### 3.4 Método empleado y combinaciones de valores paramétricos analizadas

Tras el pre-procesado de las imágenes se aplican los algoritmos de segmentación a los cortes pulmonares preprocesados, en los cuales se han minimizado los elementos que distorsionan la delimitación del nódulo pulmonar solitario.

La selección adecuada de los valores de los parámetros de entrada a los algoritmos, descritos en el apartado 2.3 es muy importante para realizar una buena segmentación, y ésta no es una tarea fácil, pues la relación entre estos valores de entrada y el resultado final de la segmentación no es evidente, ni inmediata. Por ello, además de la valoración del resultado visual de la segmentación, también se calculan y utilizan los valores tipificados de los índices empleados, descritos más adelante en el apartado 3.6 pág.76. Estos valores permiten decidir de forma objetiva, dentro de un mismo algoritmo, las pruebas que presentan los mejores resultados, y así de forma implícita, los valores óptimos para ser utilizados en los parámetros de entrada. Una vez seleccionadas las mejores pruebas por algoritmo, se utiliza el coeficiente de variación para comparar los mejores resultados de segmentación entre los diferentes algoritmos, aplicados a las mismas imágenes. Y así de este modo, determinar el algoritmo más adecuado tras analizar cual tiene un comportamiento con menos dispersión y mayor media entre las pruebas realizadas al grupo de imágenes.

En las tablas mostradas a continuación en este apartado, se indican por cada algoritmo estudiado, los parámetros cuyos valores son objetivo de optimización, con la finalidad de obtener la segmentación más precisa del NPS en las imágenes de CT de tórax. Las pruebas de segmentación realizadas, con un algoritmo concreto sobre una imagen, se han ejecutado utilizando ocurrencias tomadas de las combinaciones de valores de los parámetros de entrada a cada algoritmo; dichos valores de parámetros están indicados en la columna *Rango de valores* de las citadas tablas. Cuando los datos de dichas columnas, están entre corchetes, indican todos los valores que han sido empleados en las pruebas, desde la cuantía inferior especificada hasta la superior incluida; y cuando los valores extremos están expresados entre llaves, entonces han sido asignados, como valores de parámetros, un subconjunto del formado por los extremos especificados.

Con el algoritmo de referencia CMeans los parámetros de entrada, que se han empleado en la optimización, son el tipo de muestra utilizada como centroide inicial, denotada por

tm, y el número de clases en las que segmentar la imagen, denotado por c; los valores considerados para ambos parámetros se especifican en la Tabla 5.

<b>CMeans</b>	<b>Parámetros</b>	<b>Descripción</b>	<b>Rango de valores</b>
	<b>tm</b>	Tipo de muestra utilizada como centroide	Automática, Manual
	<b>c</b>	Número de clústeres	[3, 15]

**Tabla 5. Parámetros analizados del algoritmo CMeans.**

Teniendo en cuenta la función objetivo del algoritmo Fuzzy C-Means, los parámetros de entrada variables que han sido utilizados son: los centroides iniciales, el factor fuzzy m, y el número de clústeres. El parámetro c, número de clústeres, es un entero positivo, que indica al algoritmo en cuantas clases deberá quedar segmentada la imagen resultante del procesado, con este método FCM. Los centroides iniciales, para cada clúster a determinar en cada una de las imágenes, son calculados de forma automática, o escogidos manualmente. Y el factor fuzzy, m, es el exponente de la función de pertenencia. Los valores utilizados para estos parámetros están especificados en la Tabla 6.

<b>FCM</b>	<b>Parámetros</b>	<b>Descripción</b>	<b>Rango de valores</b>
	<b>tm</b>	Tipo de muestra de los centroides	Automática, Manual
	<b>m</b>	Factor fuzzy	{1,...,10 <sup>10</sup> }
	<b>c</b>	Número de clústeres	{3..., 30}

**Tabla 6. Parámetros analizados del algoritmo Fuzzy C-Means.**

En la versión difusa del algoritmo “del vecino más cercano”, también se utilizan como parámetros de entrada variables: el tipo de muestra utilizada como centroide(tm), el parámetro fuzzy (m), y el número de clases (c) en las que dividir la imagen, para que en una de ellas quede diferenciado el nódulo a precisar; teniendo en este método como parámetros específicos: el número de vecinos, y el número de muestras para caracterizar a los vecinos próximos. Se denota por v el número de vecinos, considerados relevantes, por ser más próximos a los píxeles a clasificar en cada clúster, y n al número de muestras por cada vecino. Los valores utilizados en las diversas segmentaciones, están especificados en la Tabla 7.

<b>FKNN</b>	<b>Parámetros</b>	<b>Descripción</b>	<b>Rango de valores</b>
	<b>tm</b>	Tipo de muestra utilizada como centroide	Automática, Manual
	<b>n</b>	Número de muestras	[3, 13]
	<b>v</b>	Vecinos	[3, 6]
	<b>m</b>	Factor fuzzy	{1,..., 400}
	<b>c</b>	Número de clústeres	[3, 10]

**Tabla 7. Parámetros analizados del algoritmo Fuzzy K-Nearest Neighbor.**

En la Tabla 8 se indican los parámetros utilizados como valores de entrada al método Adaptive Fuzzy C-Means: tipo de muestra (tm) y número de clústeres (c), y los parámetros de penalización de grandes variaciones de intensidad (L1) y de discontinuidad (L2) en las imágenes.

<b>AFCM</b>	<b>Parámetros</b>	<b>Descripción</b>	<b>Rango de valores</b>
	<b>tm</b>	Tipo de muestra utilizada como centroide	Automática, Manual
	<b>L1</b>	Lambda1	{0,..., 10 <sup>10</sup> }
	<b>L2</b>	Lambda2	{0,..., 10 <sup>10</sup> }
	<b>c</b>	Número de clústeres	[3, 10]

**Tabla 8. Parámetros analizados del algoritmo Adaptive Fuzzy C-Means.**

Para el método Spatially Kernelized Fuzzy C-Means, se especifican en la Tabla 9, junto con los parámetros tipo de muestra (tm), número de clases (c) y factor fuzzy (m), los parámetros específicos que incluyen información espacial de la ventana de vecinos (v), alfa (a) y sigma (s) para compensar las inhomogeneidades.

<b>SKFCM</b>	<b>Parámetros</b>	<b>Descripción</b>	<b>Rango de valores</b>
	<b>tm</b>	Tipo de muestra utilizada como centroide	Automática, Manual
	<b>v</b>	Ventana	{3x3,...,10x10}
	<b>s</b>	Sigma	{50,..., 300}
	<b>a</b>	Alfa	{0.1,...,0.8}
	<b>m</b>	Factor fuzzy	{1,...,128}
	<b>c</b>	Número de clústeres	[3, 10]

**Tabla 9. Parámetros analizados del algoritmo Spatially Kernelized Fuzzy C-Means.**

Existen dos parámetros comunes a todos los algoritmos empleados: el tipo de muestra y el número de clases o clústeres. El tipo de muestra, utilizada como semilla inicial de cada clase, para cada segmentación tiene como valores en todos los métodos: manual y automática; y el valor para el parámetro número de clústeres en que segmentar la imagen, con el objetivo de aislar adecuadamente el NPS, ha sido, al menos, un número entero del intervalo cerrado entre 3 y 10, en todos los métodos.

Las muestras de tipo “manual”, para el clúster nódulo, se han tomado de la imagen entre los píxeles situados en la zona correspondiente al tumor. Y las muestras automáticas se han calculado a través de un algoritmo que detecta máximos y mínimos locales en el histograma de la imagen.

Como se ha mencionado anteriormente, el número de clústeres en que segmentar la imagen ha variado de forma sistemática, para todos los algoritmos, entre 3 y 10. Se le ha asignado un número mínimo de clústeres igual a tres, porque es necesario distinguir en la imagen, al menos, la región fondo de imagen, la región parénquima pulmonar, y la región de interés nódulo.

Se han realizado pruebas aisladas, no sistematizadas, con los diversos algoritmos considerando para éste parámetro valores superiores a 100 clústeres, pero se ha desistido de sistematizarlos, pues, los aciertos conseguidos eran ridículos y para ello eran necesarios tiempos de ejecución excesivamente grandes. También se han realizado pruebas con valores entre 50 y 100 clases por segmentación con resultados negativos. Finalmente se ha optado por sistematizar algunas pruebas con ocurrencias entre 10 y 30 clústeres para el algoritmo FCM, siendo la amplia mayoría de las segmentaciones realizadas entre 3 y 10 clústeres.

### *3.5 Criterios utilizados para la clasificación de las pruebas*

El análisis ROC (Receiver Operating Characteristic) es la técnica de preferencia para evaluar nuevos métodos de diagnóstico por imagen, en Radiología y en especial en las tomografías computarizadas de tórax.

Los resultados originados sobre las imágenes procesadas, tras la ejecución de los algoritmos de segmentación, es necesario analizarlos, evaluarlos y compararlos de forma precisa y objetiva; pues de alguna manera, el clasificar a un paciente con la presencia o ausencia de nódulos cancerígenos puede depender de ello. Los radiólogos realizan estas conjeturas a través de “pruebas diagnósticas”, de las cuales es trascendente conocer su exactitud o precisión. La curva ROC es una de las herramientas más usadas para evaluar la exactitud de estos “diagnósticos”.

En el informe especial [Dodd, 2004] se justifica la utilización de las curvas ROC, como metodología a seguir para el estudio, análisis, y presentación de resultados de las imágenes del LIDC.

La curva ROC se utilizó durante la Segunda Guerra Mundial en el análisis de imágenes de radar, “para evaluar la capacidad de los operadores de radares para distinguir entre señales con significado (avión enemigo al acecho y el ruido de fondo)” [Marcum, 1947]. A partir de las primeras presentaciones de la Teoría de Detección de Señales, en la década de 1950, esta teoría asume un papel central en Psicología [Green y Swets, 1966] surgiendo del

análisis estadístico de decisión, y a finales de la década de 1960 tienen lugar las primeras aplicaciones de esta teoría al área médica, siendo en la actualidad el análisis ROC un método generalizado en esta disciplina.

Hoy en día está muy extendida la utilización de las curvas ROC, jugando un papel fundamental en evaluación de diagnóstico y clasificación, en detección de señales, psiquiatría, medicina y psicología, y recientemente también en meteorología, inteligencia artificial y economía. Desde la perspectiva de las Ciencias de la Computación, el análisis ROC se utiliza, cada vez más, como una herramienta para evaluar los efectos discriminantes entre los diferentes métodos, siendo Spackman uno de los primeros en utilizarlas para la evaluación y comparación de algoritmos [Spackman, 1989]. Es evidente que el análisis ROC, proporciona un método para seleccionar los modelos posiblemente óptimos y descartar aquellos cuyo coste/beneficio no sea el adecuado [Egan, 1975].

A continuación se describen los fundamentos formales de esta herramienta, en la cual se ha basado este estudio de resultados de los distintos experimentos realizados con los algoritmos de segmentación sobre las imágenes utilizadas como patrón.

### 3.5.1 Clasificador de rendimiento

Un clasificador es una función que permite decidir, a partir de la evaluación de una “condición”, cuáles de un conjunto de instancias están en un grupo o en otro. Pudiendo tener un resultado binario, ordinal o continuo. Es un resultado binario cuando sólo existen dos posibles resultados, Positivo y Negativo, denotados por P y N. Ocurre un resultado ordinal si este puede tomar distintos valores, ordenados siguiendo una escala establecida. Y el resultado es continuo cuando el resultado puede ser cualquier valor, dentro de un intervalo real especificado.

Se considera un problema de clasificación binaria T, es decir test o prueba, cuyos resultados se etiquetan como las clases Positiva o Negativa, y siendo S el estado real de la situación a clasificar con los valores Falso o Verdadero.

Para analizar una prueba o test, se establece una comparación entre el resultado T y el estatus S. Esta comparación se realiza utilizando una matriz o tabla, en la cual se relacionan las filas (Negativa, Positiva) con las columnas (Falso, Verdadero), tal como se muestra en la Tabla 10.

En este trabajo se trata de clasificar los píxeles de la imagen segmentada, con el objetivo de determinar si un píxel clasificado como perteneciente al clúster nódulo, es punto

cancerígeno según su correspondencia con la máscara asociada informada por los radiólogos. Así un Falso Positivo ocurrirá cuando la prueba predice que el resultado es Positivo, y es falso que el píxel se corresponda con un punto cancerígeno del nódulo marcado como tal en la máscara. Un Falso Negativo, por el contrario, ocurre cuando el resultado de la prueba es Negativo, sugiriendo que no es punto correspondiente al nódulo, cuando el estado real S es verdadero.

		Estado real S	
		Falso	Verdadero
Resultado de la prueba T	Negativa	<b>Verdadero Negativo (TN)</b>	<b>Falso Negativo (FN)</b>
	Positiva	<b>Falso Positivo (FP)</b>	<b>Verdadero Positivo (TP)</b>

Tabla 10. Resultados de comparar la prueba T con el estado real S.

En la comparación de la prueba T y el estado real S son posibles dos errores (FP, FN) y dos aciertos (TN, TP). Para evaluar la validez de la prueba es necesario calcular cuánto se aproxima la medida al valor real que se pretende medir. Si la prueba fuese perfecta los FP y los FN serían 0.

La Tabla 10 en el campo de la Inteligencia Artificial se conoce como matriz de confusión, y estadísticamente es una tabla de contingencia; en dicha matriz o tabla se registran y analizan la relación entre las variables que permiten calcular los índices de validez o probabilidades indicados en la Tabla 11, también conocidos como índices de la eficacia de una prueba [Metz, 1986b].

	Significado	Fracción
<b>ESPECIFICIDAD</b>	Probabilidad de obtener un resultado negativo cuando el estatus real es falso.	$\frac{TN}{(TN + FP)}$
<b>SENSIBILIDAD</b>	Probabilidad de obtener un resultado positivo cuando el estatus real es cierto.	$\frac{TP}{(FN + TP)}$
<b>valor PREDICTIVO NEGATIVO</b>	Fracción entre resultados válidos y resultados negativos de la prueba	$\frac{TN}{TN + FN}$
<b>valor PREDICTIVO POSITIVO</b>	Fracción entre resultados válidos y resultados positivos de la prueba	$\frac{TP}{TP + FP}$

Tabla 11. Índices de eficacia de una prueba.

En general se utilizan la sensibilidad y la especificidad junto con los valores predictivo positivo y predictivo negativo como índices de validez.

La Tabla 12 muestra los resultados expresados en fracciones para la prueba T y estado S si se representan los resultados de las variables test T, y estado S en binario, con el significado indicado a continuación:

$$T = \begin{cases} 1 & \text{positivo para píxel del Clúster\_nódulo en la posición } (x,y) \\ 0 & \text{negativo para píxel de Clúster\_NO nódulo en la posición } (x,y) \end{cases}$$

$$S = \begin{cases} 1 & \text{Verdadero si existe punto\_nódulo en posición } (x,y) \\ 0 & \text{Falso si no existe punto\_nódulo en posición } (x,y) \end{cases}$$

	S = 0	S = 1
T = 0	<b>Especificidad</b> TNf = P[T=0   S=0]	Fracción falso negativo FNf = P[T=0   S=1]
T = 1	Fracción falso positivo FPf = P[T=1   S=0]	<b>Sensibilidad</b> TPf = P[T=1   S=1]

Tabla 12. Resultado de un test binario en fracciones.

Este modelo se aplica a la situación que se necesita modelizar o representar. Siendo necesario, por lo tanto, un punto de corte  $c$  para establecer el umbral a partir del cual se considerarán los resultados de una prueba positivos o negativos para un estado real  $S$  verdadero. Estableciendo este punto de corte se pueden transformar los resultados en binarios. Y así obtener un resultado  $T$  dicotómico:

$$\text{Si } T \geq c \Rightarrow \text{positivo}$$

$$\text{Si } T < c \Rightarrow \text{negativo}$$

Un píxel de la imagen segmentada, clasificado en el clúster nódulo, será considerado TP cuando su correspondiente posición en la máscara tenga un valor mayor o igual a 500. Las máscaras asignan el valor 32767 a los puntos que bordean la lesión nódulo; los demás puntos de la máscara tienen asociado un valor en el intervalo de los números naturales de  $[0, 1000]$ . Cuando el valor de un punto en la máscara, asociado con el píxel de la imagen segmentada, es mayor de 500 (punto de corte establecido por los expertos que realizaron las máscaras) entonces es un punto correspondiente al nódulo maligno. En la imagen segmentada se clasifican los píxeles pertenecientes al clúster nódulo en verdaderos positivos (TP), falsos positivos (FP), verdaderos negativos (TN) y falsos negativos (FN).

De esta forma, los resultados para un valor de corte  $c$ , quedan resumidos en fracciones como muestra la Tabla 13.

	S = 0	S = 1
T < c	<p><b>Especificidad</b></p> $TNf(c) = P[T < c   S=0]$	<p>Fracción falso negativo</p> $FNf(c) = P[T < c   S=1]$
T ≥ c	<p>Fracción falso positivo</p> $FPf(c) = P[T ≥ c   S=0]$	<p><b>Sensibilidad</b></p> $TPf(c) = P[T ≥ c   S=1]$

**Tabla 13. Resultados de un test en fracciones.**

En [Pepe, 2003] se propone la utilización de  $TPf$  y  $FPf$  para cuantificar la precisión de una prueba, sugiriendo así la utilización de un gráfico bidimensional que permita ver rápidamente lo precisa que es una prueba. Situando la  $FPf$  en el eje de las abscisas y la  $TPf$  en el eje de las ordenadas. Este gráfico situado en un cuadrado unitario tendrá un solo punto ( $FPf$ ,  $TPf$ ) en el primer cuadrante del sistema de coordenadas para el caso del test binario. Siendo un test con precisión perfecta cuando se cumpla que  $FPf = 0$  y  $TPf = 1$ .

Y en el caso de una prueba con resultado continuo se muestran, en el cuadrado unitario, todos los pares  $(FPf(c), TPf(c)), \forall c$ . Este gráfico que muestra la exactitud diagnóstica, es la Curva ROC.

### 3.5.2 Curva ROC

La curva ROC, resume toda la información contenida en las tablas de contingencia  $2 \times 2$  que originan todos los puntos de corte, y que, por esta razón, es mucho más informativa que sólo las medidas de efectividad como la sensibilidad o la especificidad que se refiere sólo a un punto de corte, supuestamente óptimo. En esta curva se muestran todos los pares de sensibilidad y complemento de la especificidad, resultantes de la variación continua de todos los puntos de corte en todo el rango de resultados observados. A continuación se describe formalmente dicha curva.

Según [Pepe, 2003] las características operacionales ( $FPf$ ,  $TPf$ ), donde  $c$  es el punto de corte o umbral de decisión, son:

$$FPf(c) = P[T \geq c | S = 0],$$

$$TPf(c) = P[T \geq c | S = 1],$$

Y permiten definir la curva ROC de la siguiente forma:

$$ROC(.) = \{FPf(c), TPf(c), c \in (-\infty, \infty)\}$$

Cuando aumenta el umbral de decisión  $c$  se cumple:

$$\lim_{c \rightarrow \infty} TPf(c) = 0 \text{ y } \lim_{c \rightarrow \infty} FPf(c) = 0$$

Y cuando el umbral  $c$  disminuye

$$\lim_{c \rightarrow -\infty} TPf(c) = 1 \text{ y } \lim_{c \rightarrow -\infty} FPf(c) = 1$$

A partir de lo cual se deduce que la curva ROC es una función monótonamente creciente en el cuadrante positivo, pudiendo expresarla de la siguiente forma:

$$ROC(.) = \{p, ROC(p), p \in (0,1)\}$$

Es decir, si  $p$  es un punto en la curva, la función ROC a cada valor de  $p$  le hace corresponder  $TPf(c)$ , donde  $c$  es un punto de corte particular, para el cual  $FPf(c) = p$ .

En la Figura 34 se muestra un ejemplo de curva ROC, y sobre ella el valor que toma la función para el punto  $p$ . Se puede operar sobre cualquier punto de la curva ROC utilizando el valor de corte apropiado.

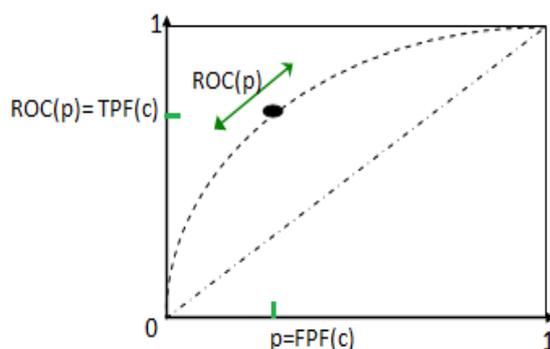


Figura 34. Curva ROC.

### 3.5.3 Área bajo la curva ROC

El índice más utilizado como resumen de la exactitud de una prueba diagnóstica es el área bajo la curva ROC, se denota por AUC (Área Under Curve); véanse las curvas A y B en la Figura 35 que toman valores entre 0.5 y 1, en el caso de una prueba perfecta la curva

tendrá el valor 1.0, y en el caso de una prueba de poca información como la curva C su valor será como máximo 0.5.

Para evaluar de manera global la capacidad de discriminación entre dos pruebas se calcula esa área que queda debajo de la curva ROC, sirviendo dicha área como índice de comparación entre pruebas.

El Área Bajo la Curva ROC se define como la siguiente integral:

$$AUC = \int_0^1 ROC(p)dp$$

Si dos pruebas están ordenadas tal que la prueba A, véase Figura 35 [Bradley, 1997], es uniformemente mejor que la B entonces se cumple que:

$$ROC_A(p) \geq ROC_B(p), \forall p \in (0,1) \Rightarrow AUC_A \geq AUC_B$$

Es decir, estadísticamente las áreas bajo la curva también están ordenadas [Pepe, 2003].

Entonces si una curva ROC es uniformemente mejor que otra, tendrá mayor área bajo ella, y la prueba correspondiente será la elegida debido a su mayor capacidad de discriminación.

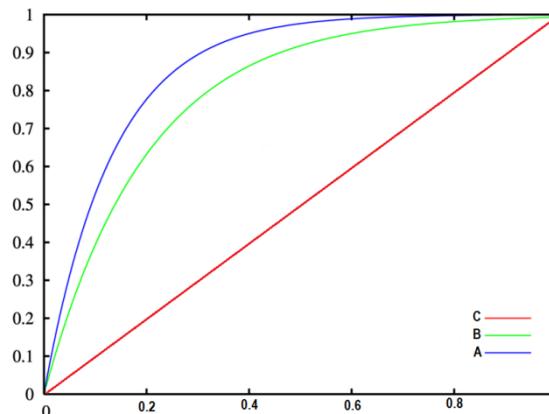


Figura 35. Curvas ROC [Bradley, 1997].

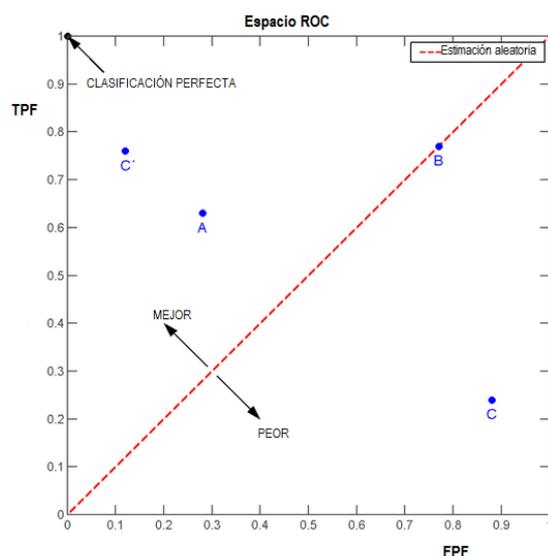
### 3.5.4 Espacio ROC

Un espacio ROC se define por FPf y TPf como ejes de abscisas y ordenadas respectivamente. Cada resultado de la predicción de una instancia de la matriz de confusión representa un punto en el espacio ROC.

En este espacio se relacionan los niveles de sensibilidad (fracción de verdaderos positivos TPf) y el complementario de la especificidad ( $1 - TNf$ ), de tal forma que el extremo

superior izquierdo, o coordenada (0,1) del espacio ROC, es el punto de clasificación perfecta, representando un 100% de sensibilidad (ningún falso negativo) y un 100% de especificidad (ningún falso positivo).

La zona de clasificación óptima corresponde a la esquina superior izquierda, ver en la Figura 36 donde se maximizan los aciertos y se minimizan los fallos. En dicha figura la línea diagonal, correspondiente a una prueba aleatoria, divide el espacio en zonas de buena o mala clasificación. Los puntos por encima de la diagonal indican unos resultados con buena clasificación, mientras que los puntos por debajo de la diagonal indican una peor clasificación [Fawett, 2006]. En dicha figura, hay cuatro clasificadores discretos, el punto A es el mejor del grupo A, B, C. El punto C es el peor, sin embargo el simétrico C' es el mejor. Una clasificación totalmente aleatoria daría un punto a lo largo de la línea diagonal, denominada también línea de no-discriminación, desde el extremo izquierdo hasta la esquina superior derecha. El punto B representa una prueba no informativa pues existe la misma probabilidad de realizar una clasificación buena que otra peor; los puntos sobre la diagonal del gráfico expresaran lo que sería una prueba inútil.



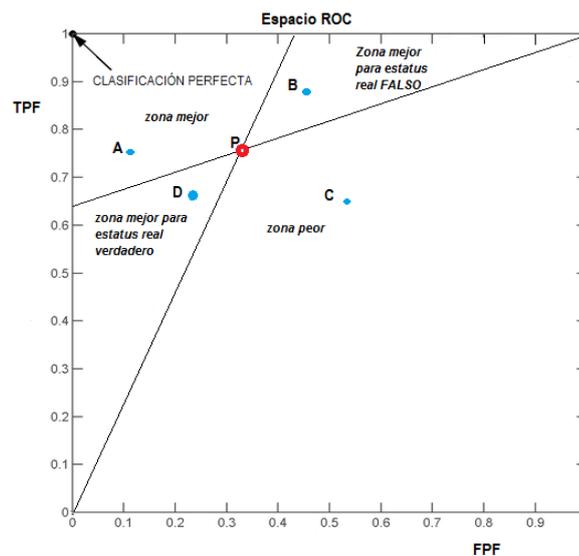
**Figura 36. Espacio ROC y línea diagonal, de estimación aleatoria, que lo divide en áreas de buena o mala clasificación.**

Para comparar una prueba P con otra cualquiera, en el espacio ROC, conviene trazar una recta que pase por el origen de coordenadas (0,0)- cero fallos y cero aciertos- y por el punto P, trazando una segunda recta que pase por el punto de coordenadas (1,1)- 100% fallos y 100% aciertos- y por ese punto P, tal que el espacio ROC quede dividido en cuatro zonas como se indican en la Figura 37, de forma que cualquier otra prueba cuyo

resultado se represente en este espacio estará ubicada en alguna de las cuatro zonas definidas como zona mejor, zona mejor para estatus real falso, zona peor y zona mejor para estatus real verdadero.

Dentro del espacio unitario ROC (Figura 37), según donde se encuentre la prueba a comparar se tendrá una de las siguientes situaciones:

- El punto A está en la zona mejor, en esta región la prueba A es siempre superior a la prueba P, tanto para confirmar un estatus real verdadero como un estatus real falso.
- El punto B está en la zona mejor para un estatus real falso.
- El punto C está en la zona peor, en esta zona la prueba C siempre es inferior a la prueba P.
- El punto D está en la zona mejor para un estatus real verdadero.



**Figura 37. Regiones de clasificación al comparar pruebas en el espacio ROC.**

Todas estas consideraciones se refieren únicamente a la capacidad discriminadora de un clasificador. La decisión sobre qué prueba es más adecuada, es siempre algo más complicado, puesto que intervienen otros aspectos, principalmente: el coste de la prueba y la valoración de las consecuencias que conllevan un falso positivo o un falso negativo.

Para decidir de forma no arbitraria, cuales pruebas de las realizadas presentan los mejores resultados, se definen y utilizan los índices descritos en el siguiente apartado.

### 3.6 Índices de rendimiento de las pruebas

Se definen unos indicadores de referencia, como protocolo de trabajo, para gestionar el conjunto de pruebas realizadas al segmentar las imágenes, y para obtener información objetiva, clara y precisa de los test ejecutados. Para dicha definición e implementación de los indicadores de referencia se utilizan el área del polígono formado por los segmentos de las rectas que se cortan en el punto por el cual pasa la hipotética curva ROC y los ejes de abscisas y ordenadas de dicho espacio, y la maximización de la diferencia de la distancia entre los puntos antagónicos de Acierto total y Fallo total.

Estos dos aspectos permiten obtener una visión rápida y global de la valoración de las pruebas realizadas, facilitando la elección de la mejor combinación de valores de parámetros para cada algoritmo, y asimismo la comparación entre algoritmos.

La Figura 38 muestra una sub-matriz de la imagen que contiene la máscara del nódulo correspondiente a una de las imágenes tomográficas utilizadas. En color rojo están los píxeles confirmados por los expertos como correspondientes a puntos cancerígenos en la CT asociada. El color morado es para los píxeles en los que los expertos no coincidían en la valoración de los mismos (explicación dada en el apartado 3.2). Las celdas con valor cero y en fondo negro simbolizan el resto de la imagen-máscara hasta completar la matriz de 512x512 píxeles

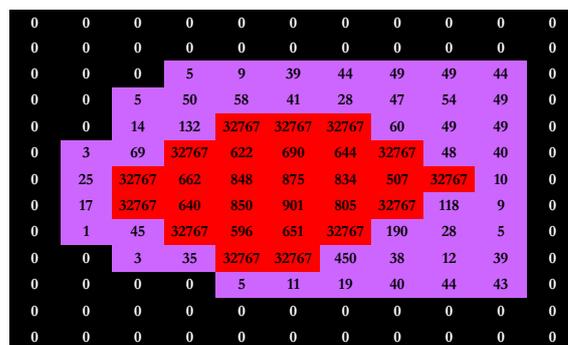


Figura 38 Submatriz con parte coloreada correspondiente a la máscara del nódulo de una CT.

En cada prueba realizada se han clasificado los píxeles de la imagen resultado de la segmentación con las etiquetas TP, FPe, FPr, TN y FN.

Para describir de forma contextual el significado de cada una de estas etiquetas, se supone que en el resultado de la imagen segmentada, por uno de los algoritmos con una combinación determinada de valores de parámetros, todos los píxeles pertenecientes al clúster nódulo están representados en la matriz de la Figura 39, la cual tiene una

correspondencia unívoca con los representados en la Figura 38 acordes a la imagen-máscara.

Se denota con TP (verdaderos positivos) a los píxeles que perteneciendo al clúster nódulo, se corresponden con los indicados por los expertos como puntos cancerígenos. Se denomina FPe (Falsos Positivos en el entorno) a aquellos píxeles que perteneciendo al clúster nódulo, no hay acuerdo mayoritario entre los expertos acerca de si son puntos cancerígenos. Por FPr (Falsos Positivos en el resto) se contabiliza cualquier píxel que perteneciendo al clúster nódulo no tiene ninguna relación con él, pero está ubicado en la imagen segmentada en el citado clúster. Y los píxeles clasificados como FN (falsos negativos) son todos aquellos que no perteneciendo al clúster nódulo tienen la calificación de elemento cancerígeno dada por los expertos. Finalmente, se denomina con TN (verdaderos negativos) a aquellos píxeles que no pertenecen a la clase nódulo y no son, considerados por la mayoría de los expertos, puntos pertenecientes al nódulo.

FPr											
FPr											
FPr											
FPr	FPr	FPe	FPr								
FPr	FPr	FPe	FPe	TP	TP	TP	FPe	FPe	FPe	FPr	FPr
FPr	FPe	FPe	TP	TP	TP	TP	TP	FPe	FPe	FPr	FPr
FPr	FPe	TP	TP	TP	FN	FN	TP	TP	FPe	FPr	FPr
FPr	FPe	TP	TP	TP	TP	TP	TP	FPe	FPe	FPr	FPr
FPr	FPe	FPe	TP	TP	TP	TP	FPe	FPe	FPe	FPr	FPr
FPr	FPr	FPr	FPr	FPe	FPe	FPe	FPe	FPe	TN	FPr	FPr
FPr											
FPr											

Figura 39. Asignación de TP, FN, FPe, FPr y TN para el clúster nódulo en correspondencia con la máscara de la Figura 38.

Se ha diferenciado entre FPe y FPr, porque una de las hipótesis de trabajo ha sido:” si se minimizan los píxeles que son falsos positivos en el entorno del nódulo entonces también se minimizarán los falsos positivos en el resto de la imagen”.

### 3.6.1 Índice Difko\_ok

Con el objetivo de maximizar los TP y disminuir los FP, se buscaron valores para los parámetros de los algoritmos, tal que produjesen como resultado en la prueba de segmentación, un punto en la zona ROC lo más próximo posible al punto ideal de clasificación perfecta (véase la representación de dicho punto en la Figura 37). Este punto ideal de acierto total, se denotará por punto OK y es el correspondiente a una prueba con valor de fracciones  $TP_f = 1$ , y de  $FP_f = 0$ .

Así, tal como se indica en la Figura 40, representante del espacio ROC, se renombra el punto óptimo, OK, donde la fracción de falsos positivos es cero y la fracción de verdaderos positivos es uno; y se denomina punto KO, al punto pésimo, en el cual la fracción de falsos positivos es uno y la de verdaderos positivos es cero.

En esta tesis interesa buscar aquellas pruebas que, al representarlas en este espacio, estén próximas al punto OK y alejadas del punto KO. Ambos puntos se encuentran en los extremos de la diagonal negativa del espacio unitario ROC (véase Figura 40), siendo por tanto la distancia entre ellos igual a  $\sqrt{2}$ , pues las coordenadas del gráfico acordes al espacio ROC representan medidas de probabilidad, y por tanto varían entre 0 y 1.

Identificando por  $d_2$  a la distancia que separa una prueba  $p$  del punto KO, y denotando por  $d_1$  la distancia a la que se encuentra la prueba  $p$  del punto OK, entonces la prueba  $p$  será mejor cuanto mayor sea el valor dado por la diferencia entre  $d_2$  y  $d_1$ . Ello indica, que serán óptimas aquellas pruebas cuya diferencia de distancia,  $d_2-d_1$ , sea aproximadamente 1.4, pues coincidirían los puntos que representan los resultados de las pruebas con el punto OK del espacio unitario ROC.

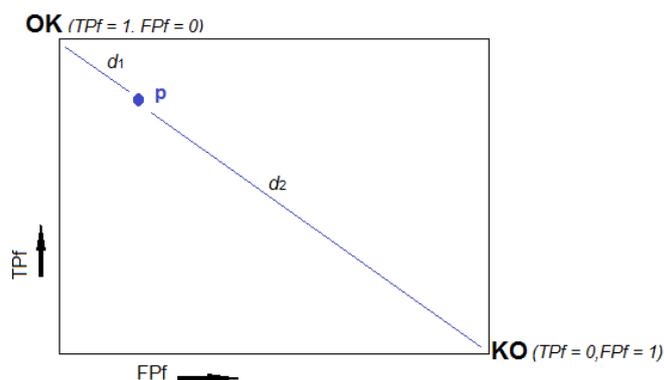


Figura 40 Espacio ROC con los puntos KO y OK.

Siendo el objetivo encontrar aquellas pruebas  $p$  que maximicen la diferencia de distancia entre los susodichos puntos KO y OK, se justifica la maximización de esta distancia como índice de búsqueda de las pruebas óptimas. A esta distancia se la ha denominado con el identificador  $Difko\_ok$ .

Una prueba se considera útil si este índice  $Difko\_ok$  es mayor que cero ( $d_2-d_1 > 0$ ), es decir si la prueba se ubica en el espacio ROC por encima de la diagonal positiva, y por tanto la prueba aporta valor; dicha prueba será perfecta si tiene un 100% de aciertos y un 0% de fallos, siendo en este último caso el valor del índice  $Difko\_ok$  igual a  $\sqrt{2}$  (valor del segmento de la diagonal negativa del espacio unitario ROC).

Para interpretar cualitativamente el valor del índice Difko\_ok se indica en Tabla 14 la escala lineal propuesta con seis etiquetas.

Valoración	Mala	Insignificante	Baja	Moderada	Buena	Muy Buena
DifKo_ok	< 0	0,0 – 0,28	0,29 - 0,57	0,58 – 0,84	0,85 – 1,13	1,14 – 1,4

Tabla 14. Valoración del índice DifKo\_ok.

### 3.6.2 Índice Azope

Cuanto mayor es el área de la zona peor, mejor es la prueba comparada (ver en la Figura 37. Regiones de clasificación al comparar pruebas en el espacio ROC.).

En el apartado 3.5.3 se ha descrito AUC (Área Under Curve), como el índice más utilizado para resumir la exactitud de un test, y en el apartado 3.5.4 se detalla la división en cuatro regiones a partir de un punto (TPf, FPf) representado en el espacio ROC. Se utilizan ambos conceptos para definir, de forma específica, el índice Azope que permita calcular la precisión de cada prueba realizada en este trabajo, a partir del área de la zona peor en el espacio unitario, donde se ubica dicha prueba.

Se definen en el espacio ROC dos puntos, denominados punto NULO y punto TOTAL. El punto NULO es el punto origen de coordenadas de dicho espacio unitario, donde el valor de aciertos y fallos coincide en 0; y el punto TOTAL es el punto en el que aciertos y fallos toman el valor máximo 1.

Este nuevo índice, Azope, se obtiene calculando la superficie de un polígono en el espacio ROC; dicho polígono está formado por cuatro segmentos  $\bar{s}_1$ ,  $\bar{s}_2$ ,  $\bar{s}_3$  y  $\bar{s}_4$ , tal como se muestran en la Figura 41. El segmento  $\bar{s}_1$  es el fragmento de la recta que pasa por el punto NULO y por el punto P (prueba representada en el espacio ROC), los puntos P y TOTAL son los extremos del segmento  $\bar{s}_2$ ; por último,  $\bar{s}_3$  y  $\bar{s}_4$  son los segmentos correspondientes a los ejes de ordenadas y de abscisas del espacio unitario ROC respectivamente.

La utilización del índice Azope ha permitido descartar todas aquellas pruebas para las que el cálculo del índice fue inferior a 0.5, en todo el grupo de imágenes segmentadas. Las pruebas descartadas se consideran no válidas, por estar situadas en el espacio ROC por debajo de la diagonal principal. Además, esto ha facilitado analizar los valores que tenían los parámetros de entrada para los algoritmos procesados, permitiendo acotar los posibles valores óptimos para dichos parámetros.

Para obtener el índice Azope se calcula la recta que pasa por el origen del espacio ROC, punto NULO de coordenadas (0,0), y por el punto representante de la prueba P, en análisis con sus correspondientes porcentajes de aciertos y fallos; por tanto, las coordenadas en el espacio ROC de este punto P tienen los valores (TPf, FPf).

A continuación se indica la ecuación de la recta que pasa por ambos puntos:

$$Y = m_1 \cdot X$$

Siendo  $m_1$  la pendiente de dicha recta, calculada según la siguiente expresión:

$$m_1 = TPf / (\varepsilon + FPf)$$

Se ha utilizado épsilon,  $\varepsilon$ , como valor despreciable para evitar una división por cero cuando en alguna prueba la fracción de falsos positivos,  $FPf$ , sea 0.

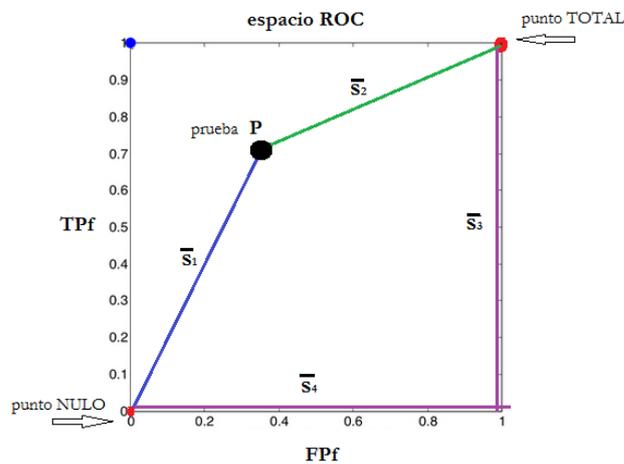


Figura 41. Segmentos delimitando el Área de Zona Peor en el espacio ROC

Además se calcula la recta, que contiene al segmento  $\overline{s_2}$ , de pendiente  $m_2$ , que pasa por el punto P, en análisis, y por el punto TOTAL -situado en el vértice superior derecho del espacio ROC-, punto (1,1):

$$m_2 = ((1 - TPf) / (1 + \varepsilon - FPf)); \quad Y = m_2 \cdot (X - FPf) + TPf$$

Así, el índice Azope se obtiene calculando el valor de la integral bajo las rectas para el intervalo [0,1] del espacio unitario ROC:

$$Azope = \int_0^{FPf} (TPf / FPf) \cdot X \, dX + \int_{FPf}^1 (((1 - TPf) / (1 - FPf)) \cdot (X - FPf) + TPf) \, dX$$

Este índice permite eliminar aquellas pruebas no significativas. Dichas pruebas críticas tienen un valor de índice Azope inferior a 0.5, en el espacio unitario ROC. Además, el dominio de este índice pertenece a los números reales positivos, y permite calcular los valores tipificados de las medias de los test realizados, con el objetivo de estimar pruebas

diferentes realizadas con un mismo algoritmo. Y así, determinar de forma cuantitativa y objetiva, cual podría ser la mejor combinación de valores de parámetros para cada uno de los algoritmos de segmentación, empleados con las imágenes de tipo TAC de tórax, para la detección de los diferentes tipos de Nódulos Pulmonares Solitarios.

Para interpretar cualitativamente el valor del índice Azope, en las diferentes pruebas de segmentación, se indica en Tabla 15 la escala propuesta para este índice.

<b>Valoración</b>	Mala	Insignificante	Baja	Moderada	Buena	Muy Buena
<b>Azope</b>	<0,5	0,51 - 0,60	0,61 - 0,70	0,71 - 0,80	0,81 – 0,90	0,91 – 1,0

**Tabla 15. Valoración del índice Azope.**

## 4 RESULTADOS Y DISCUSIÓN

Se presentan en este capítulo en el apartado 4.1 los resultados obtenidos, mediante métodos empíricos:- observación y cálculo de rendimiento-, sobre el grupo de imágenes test o imágenes de prueba para la búsqueda de los valores de parámetros adecuados que permiten segmentar el NPS existente en las imágenes con la mayor precisión posible con cada algoritmo de segmentación. Posteriormente en el apartado 4.2 se muestran los resultados obtenidos, mediante técnicas de discrepancia, al segmentar el segundo grupo de imágenes para la validación con los parámetros seleccionados por haber alcanzado un rendimiento mayor con las imágenes del conjunto test.

Sobre ambos grupos de imágenes se ha aplicado, previamente al método de segmentación, el algoritmo de pre-procesado (apartado 3.3).

Se denomina prueba a la obtención de los resultados de la ejecución de un algoritmo de segmentación específico, realizada con unos valores concretos en los parámetros de entrada a dicho algoritmo, sobre el conjunto de imágenes.

En todas las pruebas realizadas han sido calculados los aciertos y fallos de la clasificación de los píxeles, tanto en el nódulo como en su entorno y resto de la imagen; basándolos en el número de TP, FP, TN y FN obtenidos, para tras el análisis espacial de los píxeles correspondientes al nódulo obtener la fracción de verdaderos positivos (TPf) y la fracción de falsos positivos en el entorno y en el resto de la imagen (FPef y FPrf).

En este contexto del análisis de los resultados obtenidos se considera rendimiento, de una prueba realizada con un algoritmo, a la cuantía que permite estimar la precisión en la segmentación, es decir, cuanto de buena es una prueba, con las precondiciones utilizadas para la ejecución del algoritmo, según el clúster nódulo resultante maximice los TP en la comparación con la zona ROI (de la máscara correspondiente a la imagen segmentada), y en el entorno y resto de la imagen segmentada minimice los FP.

La estimación del rendimiento de cada prueba ejecutada con los algoritmos, y sobre cada imagen se ha efectuado calculando los índices Azope y Difko\_ok

Por cada algoritmo, para la búsqueda de las pruebas óptimas sobre el conjunto de imágenes de test se ha empleado como herramienta la estadística descriptiva aplicada a los valores de los índices de rendimiento.

#### *4.1 Análisis de valores paramétricos*

La combinación óptima de parámetros es aquella que proporciona como resultado de la segmentación un número máximo de verdaderos positivos en el clúster nódulo, y mínimo de falsos positivos en su entorno y en el resto de la imagen. Si la imagen no tiene un buen contraste los nódulos son difíciles de detectar por lo que hay que realizar un ajuste más fino en la combinación de valores de los parámetros de entrada de los algoritmos de procesado para que permitan mejorar dicho contraste.

Al tener por objetivo calibrar el ajuste de parámetros que son precondition de los algoritmos de segmentación, se ha ejecutado múltiples veces el mismo algoritmo con cada una de las imágenes pertenecientes al grupo conjunto de pruebas. Se han efectuado tantas pruebas de segmentación con cada imagen como combinaciones de valores han sido asignadas a los parámetros de entrada.

Se justifica la utilización conjunta de los índices utilizados en el análisis de rendimiento de las pruebas realizadas y la confirmación y/o refutación de las hipótesis planteadas en el subapartado 4.1.1.

El comportamiento e interacción de los parámetros cuyos mejores valores se pretenden buscar para el tipo de imágenes investigadas, se comenta en el subapartado 4.1.2 para los algoritmos utilizados, y en dicho apartado se muestra la relación existente entre el valor de los parámetros de entrada de los algoritmos de segmentación y el tiempo de ejecución. En el índice Difko\_ok está basada la estadística descriptiva que muestra el comportamiento de los algoritmos al segmentar este tipo de CT (subapartado 4.1.3). En el subapartado 4.1.4 se indica el resumen de comparar los mejores resultados obtenidos de las diversas pruebas realizadas con cada algoritmo según las características del NPS existente en las imágenes, haciendo uso para ello de los valores tipificados de la media del índice Azope. Finalmente en la sección 4.1.5 se compara el rendimiento de los métodos de segmentación empleados mediante el factor de covarianza, pues desde un punto de vista práctico y genérico se considera una medida de la inversa de la relación señal ruido de las imágenes segmentadas que produce cada algoritmo.

A lo largo de este apartado se muestran recortes de las imágenes segmentadas donde se ha añadido falso color en la zona ROI de las imágenes resultantes del procesado para apreciar la clasificación en clústeres a nivel de píxel, asignando el color rojo para representar los correspondientes píxeles de la zona resultante segmentada como nódulo. Ajuntando además, el recorte de la máscara de la imagen segmentada correspondiente

para que se pueda establecer fácilmente una comparación visual entre el resultado y lo que se pretende conseguir. En el recorte de la máscara del NPS también se ha utilizado el color rojo como indicativo de la zona nódulo marcada por los expertos y en color morado se muestra la zona correspondiente al entorno del nódulo en la que la mayoría de los expertos no habían coincidido en clasificarla como zona de píxeles correspondientes a puntos cancerígenos.

#### 4.1.1 Pruebas realizadas e hipótesis confirmadas y refutadas

No todas las pruebas realizadas han proporcionado resultados adecuados en el sentido de proporcionar resultados predictivos aprovechables. Se han realizado más de veinte mil ejecuciones de los algoritmos en estudio utilizando las diversas combinaciones de parámetros, expuestas en el apartado 3.4, sobre el conjunto de las 23 imágenes pertenecientes al grupo de imágenes test. De las pruebas realizadas, casi la mitad del total ha permitido obtener resultados aceptables en al menos algunas de las imágenes segmentadas. En la Tabla 16 se desglosa por algoritmo el porcentaje de pruebas válidas con algún resultado positivo en lo referente al rendimiento de los métodos de segmentación empleados.

<b>Algoritmo</b>	<b>Pruebas válidas</b>
CMeans	46%
FCM	50%
FKNN	55%
AFCM	43%
SKFCM	53%

**Tabla 16. Porcentaje de pruebas válidas por algoritmo con rendimiento positivo en alguna imagen.**

Se han utilizado los índices Azope y Difko\_ok para comparar las técnicas de segmentación en estudio, de modo que no sea arbitrario escoger aquella que ofrezca mejores tasas de clasificación de los píxeles en el clúster nódulo. En la Tabla 17, en cuya cabecera se indican los valores extremos del dominio de los índices, están tabuladas las cuantías mínimas y máximas alcanzadas por ellos, en las diversas pruebas realizadas con cada uno de los algoritmos al segmentar las imágenes con sus diversas precondiciones.

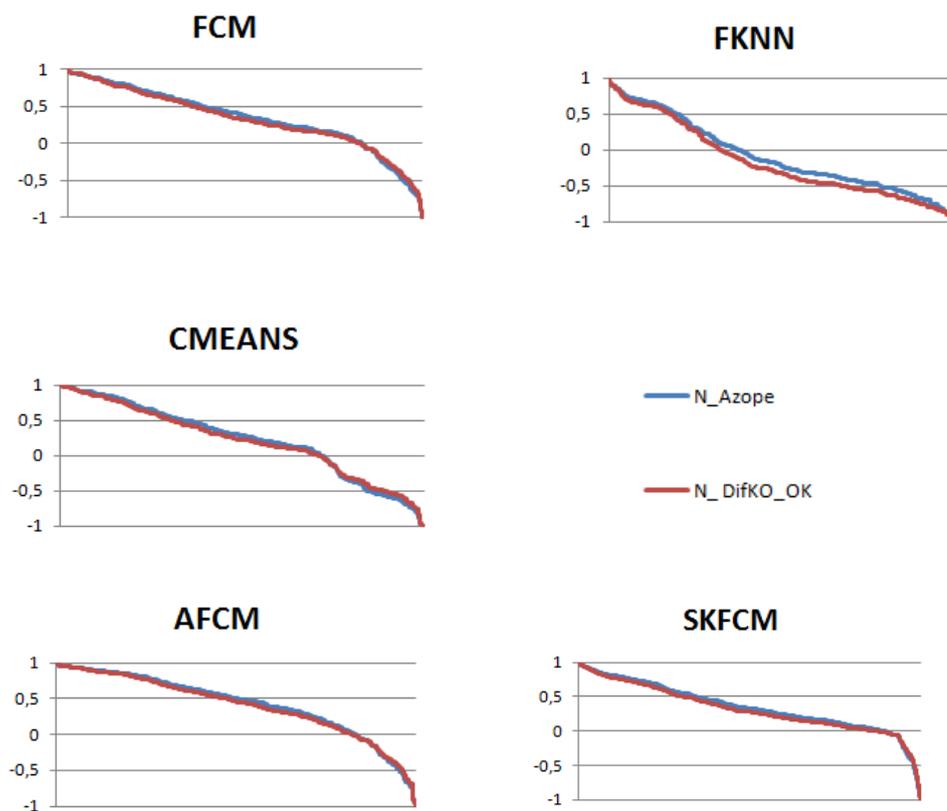
<b>Algoritmo</b> <i>[mín, Máx]</i>	<b>IndiceAzope</b> <i>[0; 1]</i>	<b>IndiceDifkook</b> <i><math>[-\sqrt{2}; \sqrt{2}]</math></i>
<b>CMEANS</b>	[0.13 , 0.97 ]	[ -1.05 , 1.32 ]
<b>FCM</b>	[ 0.07, 0.99 ]	[ -1.19, 1.36 ]
<b>FKNN</b>	[ 0.47 , 0.99 ]	[ -0.08 , 1.36 ]
<b>AFCM</b>	[ 0.07 , 0.99 ]	[ -1.18 , 1.36 ]
<b>SKFCM</b>	[ 0.07 , 0.99 ]	[ -1.19 , 1.36 ]

**Tabla 17. Valores extremos alcanzados por los índices de rendimiento en las pruebas de segmentación.**

Para comprobar la coherencia entre los índices empleados con respecto a las pruebas realizadas se normalizaron los valores de los índices obtenidos por las pruebas, escalonándolos en el intervalo  $[-1, +1]$ . En la Figura 42 se muestra para cada algoritmo, el resultado normalizado de los dos índices (como indica la leyenda de las gráficas N\_Azope y N\_Difko\_ok) por cada una de las pruebas ejecutadas con las imágenes del grupo test.

Los índices normalizados se solapan para los resultados de casi todos los algoritmos. El solapamiento es prácticamente total, existiendo un pequeño desfase entre los valores de ambos índices con el algoritmo FKNN para las pruebas con un rendimiento normalizado por debajo de cero. Ello puede deberse a que en las pruebas realizadas con el FKNN existen más falsos positivos en el resto de la imagen y un menor número de TP que en las pruebas realizadas con los otros algoritmos. Así al haber una coincidencia casi total, pero no absoluta, en la valoración del rendimiento de los algoritmos para la detección de los píxeles correspondientes al NPS de cada imagen, se ha decidido utilizar los dos de forma que se complementen en la tarea de análisis y búsqueda de los valores de parámetros que generan los mejores resultados con cada algoritmo.

Se ha focalizado el análisis en aquellas pruebas cuyos dos índices alcanzan los valores más altos, por ser en las que se obtiene un mayor rendimiento al segmentar las imágenes aislando en un único clúster, el Nódulo Pulmonar Solitario existente. Procediendo de la siguiente manera: sobre el conjunto de resultados con todas las imágenes segmentadas en las diferentes pruebas, primero se realizó la búsqueda de los mejores resultados de segmentación, a nivel de cada imagen, calculando la sensibilidad y especificidad de cada resultado segmentado y observando la representación de cada imagen segmentada, y considerando los valores obtenidos por los índices de rendimiento junto con el tiempo empleado en dicha segmentación; posteriormente se realizó el cálculo del rendimiento de cada prueba sobre todo el conjunto de imágenes.



**Figura 42 Resultados normalizados de los índices Azope y Difko\_ok para pruebas realizadas con los algoritmos.**

La combinación de valores, utilizada como precondition en la ejecución de cada algoritmo, que produce buenos resultados sobre una imagen en particular, no implica que produzca buenos resultados sobre la totalidad del conjunto de imágenes; esto indujo a buscar los mejores resultados obtenidos para el conjunto de todas las imágenes de pruebas con una misma combinación de valores para los parámetros de entrada del algoritmo en observación, y realizar el mismo proceso para cada subgrupo de imágenes caracterizado por su tipología de nódulo.

En la Tabla 18 se muestran resultados de dos pruebas diferentes ( $p_1$  y  $p_{74}$ ) realizadas sobre la misma imagen. La valoración con los dos índices de rendimiento (Azope y Difko\_ok) es muy buena. Y los correspondientes recortes de la región de interés de dicha imagen segmentada se muestran en la Figura 43, en la cual se visualiza, al comparar con la máscara la  $p_{74}$ , como al aumentar los verdaderos positivos se incrementan en el entorno del nódulo y el resto de la imagen los falsos positivos. Aunque en el ejemplo concreto se muestra el resultado de la segmentación con el algoritmo Fuzzy CMeans, ocurre lo mismo

con los demás algoritmos empleados sobre las distintas imágenes en las diferentes pruebas.

Imagen	Prueba	Tiempo	TPf	Fpef	FPrf	Azope	Difko_ok
Ima14_5	p_01	43,73	0,88	0,03	0,02	0,93	1,18
Ima14_5	p_74	527,68	1,00	0,14	0,03	0,93	1,17

Tabla 18. Resultados de diferentes pruebas de FCM para una misma imagen con valoración de rendimiento muy buena.

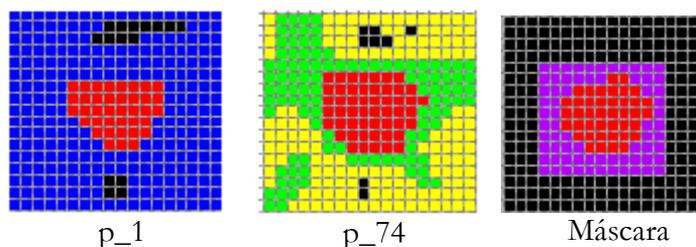


Figura 43. Nódulo de la imagen 14\_5 segmentado con el algoritmo FCM en las pruebas indicadas en la Tabla 18.

El objetivo de “Elegir para cada algoritmo aquella combinación de valores de parámetros que permiten maximizar en la segmentación resultante los puntos de la imagen pertenecientes al nódulo pulmonar solitario (NPS) y minimizar los errores, incrementando la precisión en la clasificación de los píxeles pertenecientes al clúster nódulo”, se cumple con todos los algoritmos utilizados. Satisfaciendo dicho objetivo, se confirma la hipótesis1. Como ejemplo de ello véase en la Tabla 19 y su correspondiente Figura 44 donde aumentando la fracción de verdaderos positivos disminuye la fracción de falsos positivos en el entorno del nódulo (p\_1012 respecto a la prueba p\_1000); en este caso los índices de rendimiento de las pruebas evolucionan desde una valoración baja hacia una valoración muy buena con la combinación de valores de parámetros de la prueba p\_1012. Estos resultados mostrados corresponden a pruebas diferentes ejecutadas con precondiciones en las que se han perturbado los valores de los parámetros de entrada. Lo cual confirma también la hipótesis4.

IdImagen	Id_pru	Tiempo	TPf	Fpef	FPrf	Azope	DifKo_ok
Imagen23_16	p_1000	989,44	0,48	0,11	0,01	0,68	0,48
Imagen23_16	p_1012	73,07	0,97	0,04	0,00	0,97	1,31

Tabla 19. Resultados de segmentación con el AFCM para la imagen 23\_16 con nódulo pulmonar mayor de 3 cm.

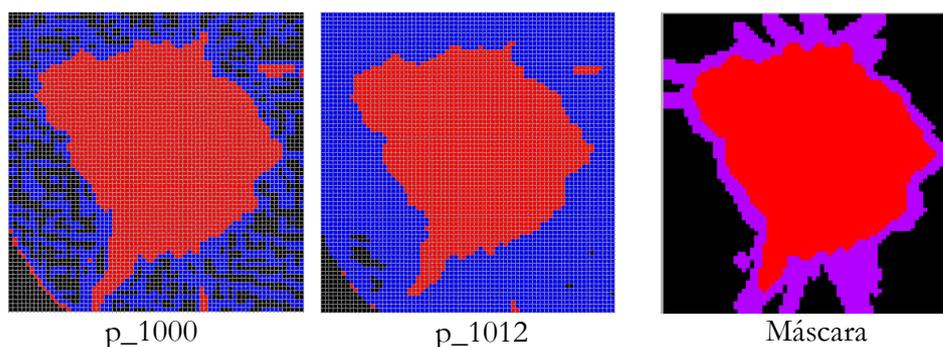


Figura 44. Resultados de segmentación del algoritmo AFKM con la Imagen 23\_19 correspondientes a la Tabla 19.

Sin embargo, al modificar en las pruebas las combinaciones de valores de parámetros buscando maximizar los aciertos y minimizar los fallos, se alcanza un punto donde el incremento de aciertos implica un aumento en los falsos positivos. Como muestra de ello, véase la Tabla 20 y su correspondiente Figura 45, donde los índices Azope y Difko\_ok presentan una valoración muy buena en la p\_08, pero al intentar aumentar la fracción de TP (en la p\_77) aumenta también la fracción de FP a 0.02 en el entorno del nódulo.

IdImagen	Id_pru	Tiempo	TPf	Fpef	FPrf	Azope	DifKo_ok
Imagen05_9	p_77	36,64	0,85	0,02	0,01	0,91	1,15
Imagen05_9	p_08	2409,87	0,81	0,00	0,01	0,91	1,10

Tabla 20. Resultados con el SKFCM con valoración muy buena de los índices

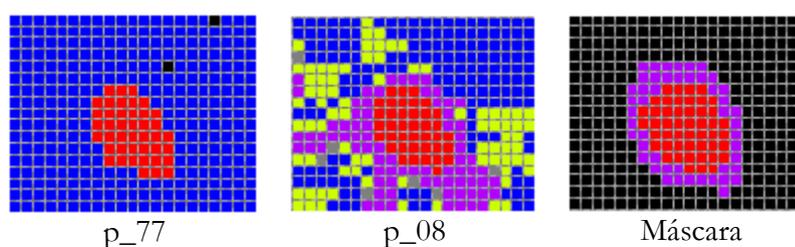


Figura 45. Nódulo de la imagen 5\_9 segmentado con el algoritmo SKFCM en las pruebas indicadas en la Tabla 20.

Y ello ocurre con todos los algoritmos analizados y sobre todas las imágenes utilizadas. En la Tabla 21 y Figura 46 se muestra como la prueba p18 proporciona mejor resultado que las pruebas p\_2003 y p\_2033 a pesar de aumentar los falsos positivos en el entorno, pues, en proporción mejora más los verdaderos positivos.

IdImagen	Id_pru	Tiempo	TPf	Fpef	FPrf	Azope	DifKo_ok
Imagen17_5	p_18	94,81	0,97	0,12	0,01	0,93	1,19
Imagen17_5	p_2003	238,32	0,88	0,05	0,01	0,91	1,17
Imagen17_5	p_2033	23,16	0,71	0,00	0,01	0,86	0,93

Tabla 21. Resultados de segmentar la imagen 17\_5 con el método SKFCM.

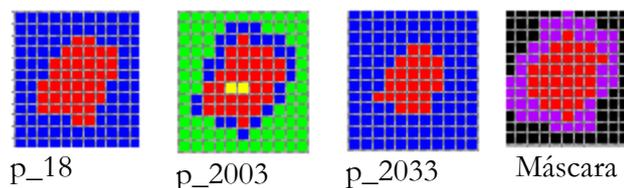


Figura 46. Nódulo de la imagen 17\_5 segmentado con el algoritmo SKFCM en las pruebas indicadas en la Tabla 21.

En todos los algoritmos analizados, la hipótesis2 planteada en esta tesis no se cumple, tal como se corrobora en los resultados concretos de segmentación del algoritmo SKFCM, mostrados en la Tabla 22, y en su correspondiente Figura 47. Se muestran los recortes de la ROI de la imagen 6\_5 segmentada según las pruebas p\_73, p\_85 y p\_86, y se han aumentado los aciertos (0.97) en la prueba p\_73 respecto a las pruebas p\_85 y p\_86 disminuyendo los falsos positivos en el entorno (0.0 frente a 0.02), sin embargo se ha mantenido sin reducir la fracción de falsos positivos en el resto de la imagen.

IdImagen	Id_pru	Tiempo	TPf	Fpef	FPrf	Azope	DifKo_ok
Imagen06_5	p_73	35,37	0,97	0,00	0,01	0,99	1,36
Imagen06_5	p_85	17,99	0,85	0,02	0,01	0,92	1,14
Imagen06_5	p_86	22,23	0,27	0,02	0,00	0,62	0,29

Tabla 22. Resultados de segmentación para la Imagen 6\_5 con el algoritmo SKFCM.

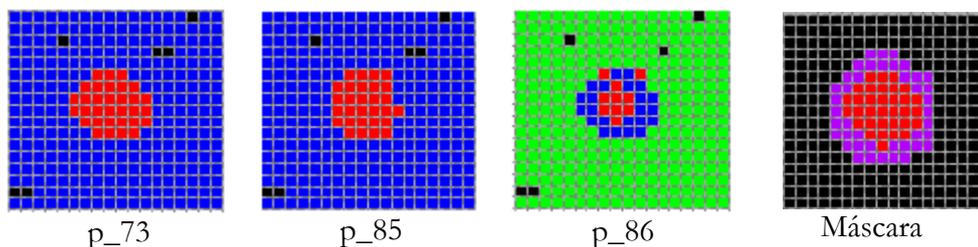


Figura 47. Nódulo de la imagen 6\_5 segmentado con el algoritmo SKFCM en las pruebas indicadas en la Tabla 22.

La hipótesis<sup>5</sup> tampoco se cumple, puesto que existen pruebas en las que las imágenes segmentadas obtenidas no tienen ningún verdadero positivo en la zona nódulo marcada por los expertos.

#### 4.1.2 Comportamiento paramétrico por algoritmo

Los índices utilizados facilitan el trabajo de comparación de pruebas, para elegir las que presentan los mejores resultados, y con ello obtener los valores óptimos para los parámetros de entrada; sin embargo, no explican las interacciones entre los distintos parámetros de un algoritmo, únicamente indican si la combinación de valores para dichos parámetros es mejor que otra de las utilizadas.

En las pruebas realizadas se han utilizado como semillas, para cada clúster de la segmentación, dos tipos de muestras: unas obtenidas automáticamente, y otras manuales. Para conseguir los ejemplares automáticamente se ha empleado un algoritmo que, conocido el número de clústeres en los que segmentar la imagen, obtiene los centroides iniciales a partir de los máximos o picos en el histograma de la imagen. Y los ejemplares manuales se han tomado directamente en las regiones donde se encuentra el NPS en la imagen tomográfica.

##### 4.1.2.1 Parámetros de entrada al algoritmo CMeans

En la Figura 48 se indica el comportamiento del algoritmo CMeans mostrando el valor medio del índice Difko\_ok frente al número de clústeres en dos series de pruebas: una que inicializa las semillas de los centroides con muestras automáticas, y otra con muestras manuales. El rendimiento es el mismo para los dos tipos de muestras utilizadas, estando ambas gráficas solapadas en la figura.

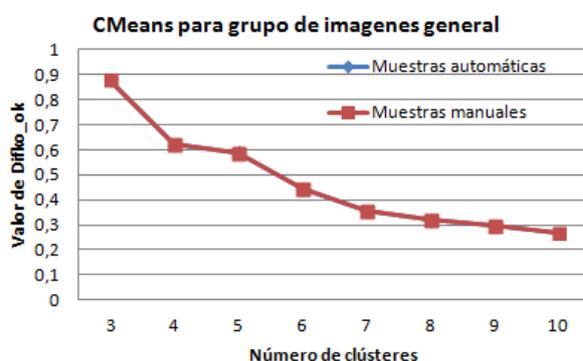


Figura 48. Rendimiento de las pruebas con el algoritmo CMEANS en función del número de clústeres para tipo de muestras automáticas y manuales.

Existe un subgrupo de imágenes que contienen un NPS denominado de tipo vidrio deslustrado donde la valoración del rendimiento es mala, pues, la media de Difko\_ok toma valores negativos para las segmentaciones de cualquier número de clústeres, y ello ocurre con los distintos tipos de muestras consideradas. Lógicamente, al excluir del grupo de imágenes test estas tomografías preprocesadas que contienen los nódulos de tipo vidrio deslustrado, aumenta el rendimiento de las pruebas realizadas sobre el grupo de las imágenes restantes, siendo igualmente similar trabajar con muestras manuales que con las obtenidas automáticamente; y disminuye también el rendimiento cuando aumenta el número de regiones de la segmentación.

Al analizar el comportamiento según el tamaño de los clústeres, independientemente del tipo de muestra, el rendimiento disminuye al aumentar el número de particiones en la imagen. Igual respuesta presentan los resultados según forma y localización de los nódulos en las imágenes segmentadas, siguiendo la misma tendencia que lo mostrado en la gráfica de la Figura 48.

Con este algoritmo de referencia se concluye que independientemente de la tipología del nódulo contenido, en cada una de las imágenes del grupo de prueba, disminuye fuertemente el rendimiento a medida que aumenta el número de clases de la imagen segmentada. Además se repite el mismo comportamiento, tanto sean las muestras manuales como automáticas, del mismo modo que tomando como centroide inicial la moda de intensidad de los píxeles de la ROI.

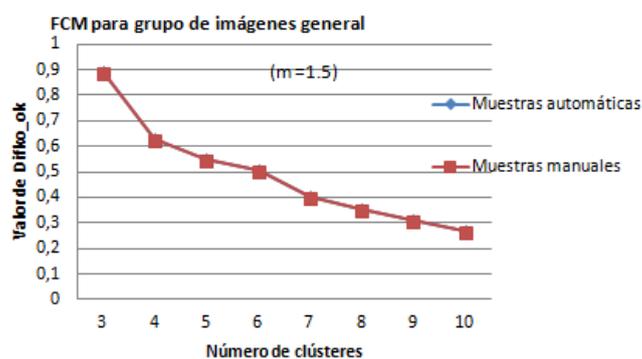
#### **4.1.2.2 Parámetros de entrada al algoritmo FCM**

Se analiza la influencia de los parámetros, en el método Fuzzy-CMeans, comparando las pruebas realizadas, sobre las imágenes a segmentar, al variar el número de clústeres y modificar el factor fuzzy ( $m$ ), además de considerar distintos tipos de semillas para los centroides.

En la Figura 49 se aprecia como disminuye el rendimiento (valor medio del índice Difko\_ok obtenido en la segmentación del grupo completo de imágenes) cuando aumenta el número de clústeres en la segmentación. Se solapa la tendencia del rendimiento, para las muestras utilizadas como centroides del tipo automáticas y del tipo manuales, teniendo en ambos casos el mismo comportamiento en relación al número de clases elegidas para la segmentación. La tendencia en el rendimiento es la misma, tanto para un factor fuzzy bajo como para un valor más alto en dicho parámetro, ello indica que el parámetro número de clústeres es el de mayor influencia en el rendimiento de la segmentación. Véase dicha

tendencia de rendimiento en la Figura 49, cuya gráfica corresponde a un valor fuzzy 1.5. La cual indica que para el tipo de imágenes tratadas, el valor óptimo para el parámetro número de clústeres se obtiene al particionar en tres clústeres, pues con este valor la segmentación realizada alcanza los valores más altos de rendimiento.

El parámetro tipo de muestras utilizado en la inicialización de los centroides, influye en la velocidad de convergencia del algoritmo, pero no se ha apreciado su influencia en el rendimiento como se comentó anteriormente.



**Figura 49. Rendimiento de las pruebas con el algoritmo FCM según el número de clústeres para tipo de muestras automáticas y manuales.**

Al realizar el estudio de campo sobre los resultados obtenidos, se ha observado que un subgrupo de imágenes (el 13% del grupo general) tiene un comportamiento diferente. Obteniéndose con ellas rendimientos de valoraciones malas, que afectaban la media de rendimiento del grupo, empeorándola considerablemente. Dicho subgrupo corresponde a las imágenes que contienen un tipo de NPS denominado en “en vidrio deslustrado”. En la Figura 50 se muestra este hecho para unos valores concretos de parámetros de entrada, y en dicha figura también se muestra la influencia del parámetro factor fuzzy en el rendimiento de las pruebas de segmentación; se muestran tres series diferentes: el grupo deslustrados se refiere al formado por las imágenes atípicas comentadas, el grupo general corresponde al grupo completo de las imágenes de prueba, y el tercer grupo es el conjunto de las imágenes de prueba sin segmentaciones de imágenes con NPS deslustrados. Cuando aumenta el factor fuzzy, a valores no convencionales, se producen pequeñas modificaciones en el grupo general, pero en el de imágenes con tipología deslustrada incrementa, considerablemente, el rendimiento de la segmentación para valores mayores de 2. Ello es debido a que esta tipología de nódulos deslustrados corresponde a píxeles con valores de intensidad más próximos al parénquima pulmonar; y así se ajusta mejor

con más grado de borrosidad para la determinación de pertenencia de los píxeles al clúster nódulo. Además de necesitar un mayor número de clases para poder aislar este tipo de nódulo en un único clúster.

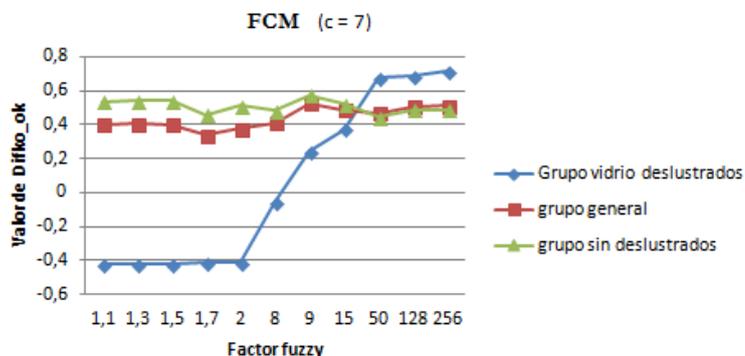


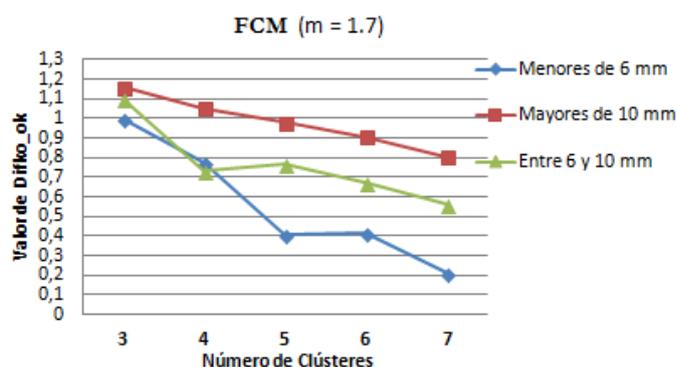
Figura 50. Rendimiento de las pruebas con el algoritmo FCM según el factor fuzzy empleado en las imágenes segmentadas con 7 clases.

Estas observaciones motivaron el realizar pruebas aisladas, con algunas imágenes, considerando para el parámetro número de clases valores entre 50 y 100 clústeres, pero se ha desistido sistematizarlos pues los aciertos conseguidos eran ridículos, y para ello eran necesarios tiempos de ejecución excesivamente grandes siendo laborioso el proceso de inicialización de los centroides. Finalmente, se ha optado por sistematizar pruebas con ocurrencias entre 10 y 30 clústeres obteniéndose cualitativamente buenos resultados, con un máximo de aciertos segmentando estas imágenes en 15 clústeres, y a partir de este número de clases disminuyen los verdaderos positivos, tal como se resume en la Tabla 23. En dicha tabla se muestra que mejora el valor medio del índice Difko\_ok para los valores de factor fuzzy altos (50, 128 y 256), y un número de clústeres: 15, 20 y 30; alcanzándose el mejor resultado ( $\text{mediaDifko\_ok} = 0,92$ ) con 15 clases y factor fuzzy 256.

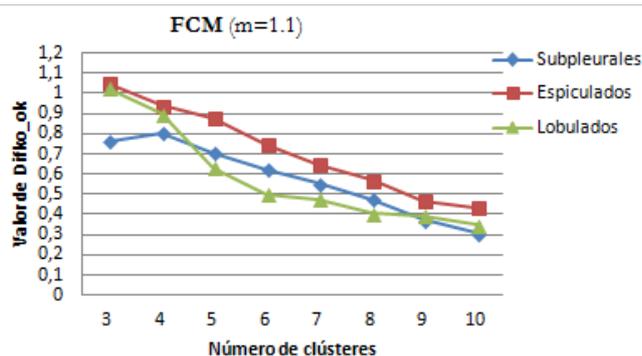
Media de valoresDifko_ok		Número de clústeres		
		15	20	30
Factor fuzzy	1,1	-0,33	-0,27	-0,16
	1,5	-0,33	-0,27	-0,18
	2	-0,28	-0,25	-0,22
	50	0,91	0,85	0,53
	128	0,91	0,85	0,56
	256	0,92	0,86	0,56

Tabla 23. Rendimiento de pruebas con el algoritmo FCM en las imágenes con NPS en vidrio deslustrado.

El rendimiento de las pruebas con el algoritmo, para todos los subgrupos de imágenes, disminuye conforme se segmenta en un mayor número de clases. Permitiendo la variación en el factor fuzzy, aumentar la precisión de la segmentación en el clúster nódulo. Esto se observa cuando se analizan los resultados por tamaño de diámetro del NPS contenido en las imágenes (véase Figura 51a), donde el número de clases con mejor rendimiento es tres para todos los segmentos de tamaños considerados y con diversos valores de factor fuzzy. Así, a medida que aumenta el número de clústeres disminuye el rendimiento, siendo éste mejor para las imágenes con diámetros de nódulos mayores. Las fluctuaciones en la tendencia a empeorar el rendimiento en los subgrupos de imágenes con nódulos subcentimétricos, pueden ser causadas por contener algunas de estas imágenes NPS en vidrio deslustrado; y, como se ha mostrado anteriormente, este tipo de nódulo se precisa mejor segmentando en un mayor número de clústeres y con valores atípicos de factor fuzzy.



(a)



(b)

**Figura 51a-b. Rendimiento de las pruebas con el algoritmo FCM según el número de clústeres en las imágenes segmentadas con NPS clasificados por tamaño, forma y localización.**

En la Figura 51b aparece resumido de forma gráfica, el comportamiento del algoritmo para los subgrupos de imágenes con nódulo cuya morfología de bordes es espiculada (en

rojo), y lobulada (en verde). Además en la misma gráfica se dibuja el comportamiento del rendimiento de las pruebas con el algoritmo para aquellas imágenes segmentadas conteniendo nódulos subpleurales (en azul). El menor rendimiento se obtiene en la delimitación de los nódulos subpleurales. Para los tres tipos de imágenes se alcanzan valores de Difko\_ok con una valoración superior o como mínimo buena, tomando como factor fuzzy valores en los intervalos [1.7, 2].

### 4.1.2.3 Parámetros de entrada al algoritmo FKNN

Los parámetros de entrada para la ejecución de este algoritmo son los indicados en la Tabla 7, el más influyente es el número de clústeres. Se ha utilizado mayoritariamente el tipo de muestra manual, pues los resultados con respecto a las pruebas en las que se utilizan semillas generadas de forma automática eran ligeramente mejores, siendo también manuales las muestras consideradas por cada vecino.

En la Figura 52 se muestran los resultados del rendimiento de las pruebas con el algoritmo según el número de clases utilizadas, mostrando en las gráficas el comportamiento del mismo según la densidad del tipo de NPS contenido en las imágenes.

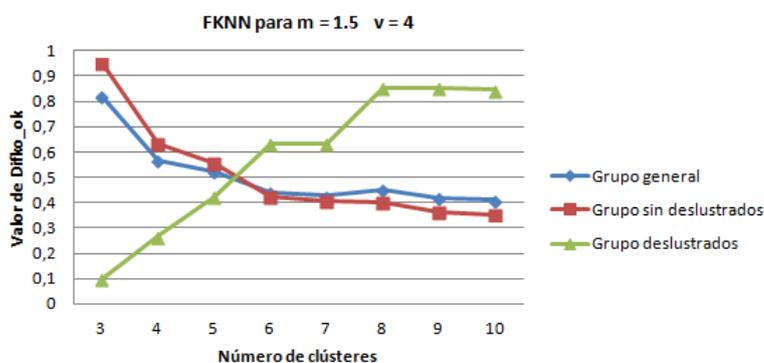


Figura 52. Rendimiento de las pruebas con el algoritmo FKNN en función del número de clústeres.

En color verde se muestran los resultados del sub-grupo de imágenes en vidrio deslustrado, y en rojo el comportamiento del algoritmo con el subgrupo de imágenes que tienen nódulo definido sólido o semi-sólido; siendo estos últimos los correspondientes al 87% de las imágenes del grupo de prueba para la búsqueda de valores de parámetros. Se observa distinto comportamiento en la tendencia del rendimiento. En el caso de las imágenes con nódulos con componente sólido o semisólido se obtiene una valoración

muy buena, al segmentar en tres clases dichas imágenes; sin embargo, para las CT que contienen nódulos en vidrio deslustrado es necesario segmentar como mínimo en seis clústeres para alcanzar, al menos, una valoración moderada. Las gráficas de la figura corresponden a segmentaciones realizadas con valor fuzzy 1.5 y considerando 4 vecinos por muestra de cada clúster a segmentar.

Se indica en la leyenda de la Figura 53, el número de clústeres empleados en segmentar las imágenes del subgrupo con nódulos en vidrio deslustrado. En las distintas pruebas, con diversos valores para los parámetros de entrada, se obtiene una valoración de Difko\_ok muy buena empleando 8, 9 y 10 clústeres. A medida que disminuye el valor en el número de clústeres a segmentar, disminuye el rendimiento de las pruebas con el algoritmo para este subgrupo de imágenes; este hecho es apreciable en el gráfico, sin conocer los valores del resto de parámetros que influyen en el resultado de la segmentación de las múltiples pruebas.

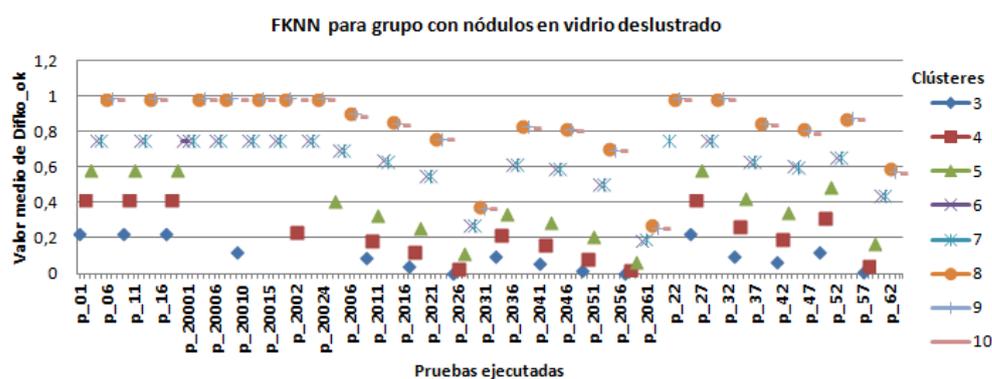


Figura 53. Rendimiento de las pruebas con diversos valores para los parámetros de entrada al segmentar el sub-grupo de imágenes con NPS en vidrio deslustrado con el algoritmo FKNN.

Al analizar la influencia del parámetro número de vecinos, en la segmentación realizada con 8, 9 ó 10 clases del sub-grupo de imágenes con nódulos en vidrio deslustrado, se observó que los resultados disminuyen su rendimiento cuando se incrementa el número de vecinos. Lo cual parece originado por la dificultad y variabilidad en la asignación de muestras adecuadas para los vecinos del clúster nódulo, y si se consideran más vecinos dicha dificultad se ve incrementada, pues en la clasificación de los píxeles para el clúster nódulo influyen las pertenencias de todos y cada uno de los vecinos asignados.

La influencia del factor fuzzy, para imágenes con este tipo de nódulo en vidrio deslustrado, es tal que para valores altos (mayores de 15) tiene el mismo rendimiento para 3 que para 4 vecinos, pero para valores fuzzy bajos ( $m < 4$ ) produce mejor rendimiento una vecindad de valor no mayor de 3, aunque se requiere un número de al menos 8 muestras por cada vecino (un total de 24 muestras como mínimo). Ello es debido a la naturaleza de esta tipología de nódulos, pues en la misma ROI están entremezclados píxeles de intensidades similares al parénquima pulmonar.

Para las demás imágenes, las de tipología de nódulo sólido o semisólido (el 87% de las imágenes grupo de prueba), el número óptimo de clases en que particionarlas es tres, y con un número de vecinos mayor o igual a cuatro; sin embargo el rendimiento decrece cuando aumentamos el número de muestras por vecino, debiendo ser este parámetro ajustado a un valor coincidente con el número de vecinos considerado.

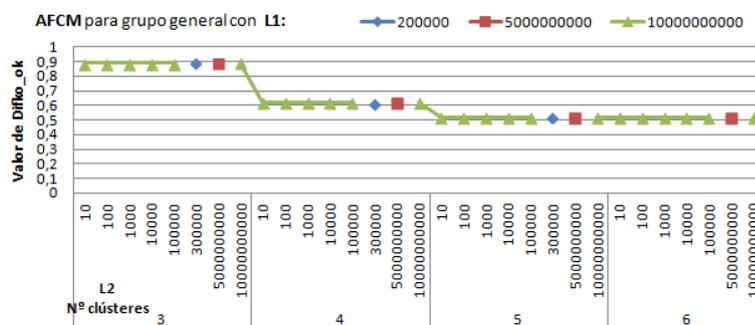
Para cualquier tamaño del tipo de nódulo contenido en las imágenes, el comportamiento del parámetro número de clústeres es tal que disminuye el rendimiento de las pruebas a medida que aumenta el número de clases en la segmentación. Siendo el número de clústeres, también con el algoritmo FKNN, el parámetro más influyente; los demás parámetros permiten realizar pequeños ajustes, para mejorar el rendimiento de la segmentación a realizar sobre las imágenes; existiendo una interdependencia entre ellos que permite mejorar la precisión en la clasificación de los píxeles pertenecientes a la ROI. En esta interdependencia se ha observado que en el ajuste a realizar si disminuye el valor del factor fuzzy se debe aumentar el número de vecinos, y si aumenta el valor fuzzy lo adecuado es disminuir el número de vecinos.

#### **4.1.2.4 Parámetros de entrada al algoritmo AFCM**

Este algoritmo se ha codificado en Matlab para un factor fuzzy igual a 2, así pues, los parámetros variables a analizar, tal como se ha indicado en la Tabla 8, son el número de clústeres, tipo de muestra, y los parámetros de regularización de primer y segundo orden: Lambda1 y Lambda2.

La relación entre el rendimiento y el número de clústeres para este algoritmo es casi la misma utilizando muestras automáticas que manuales, y al igual que en algoritmos anteriores el rendimiento de las pruebas disminuye a medida que el valor del número de clases se incrementa. Se muestra este hecho, para diferentes valores de L1 y L2, en la Figura 54, en la que para valores muy grandes de  $L1 \geq 200000$  independientemente de los

valores de L2, el promedio de la Difko\_ok obtenida vale lo mismo para igual número de clústeres, disminuyendo su valor conforme aumenta el número de regiones segmentadas en la imagen.



**Figura 54. Rendimiento de las pruebas con el algoritmo AFCM en función del número de clústeres y de L2 para distintos valores de L1.**

Para valores bajos de L1 (10; 100; 1000; 10000), el rendimiento de las pruebas con este algoritmo también presenta los mejores resultados en las segmentaciones en tres clústeres. Asimismo, las imágenes que contienen nódulos cuya densidad es en vidrio deslustrado, también producen resultados atípicos como los mostrados anteriormente con los algoritmos FCM y FKNN; dichos resultados, considerando un número de clústeres en el intervalo [3, 7] y valores del parámetro L2 entre 0 y 1000000000 son bastante malos, pues se obtienen valores negativos y en los mejores casos con un promedio de Difko\_ok inferior a 0.5, con una valoración moderada como máximo.

La Figura 55 muestra los promedios de rendimiento de las pruebas, al procesar las imágenes que contienen NPS en vidrio deslustrado, donde se realizaron segmentaciones con valores de 8, 9 y 10 clústeres para valores bajos (entre 0 y 10) tanto del parámetro de primer orden L1 como del de segundo orden L2. Bajo dichas condiciones se obtienen buenas valoraciones en los resultados de este subgrupo de imágenes específicas. Para las imágenes con esta tipología de nódulo, los parámetros L1 y L2 deben tener valores tan pequeños como se ha indicado porque interesa la imagen sin apenas suavizar para que no se “diluyan” las intensidades de los píxeles constituyentes de este tipo de NPS.

Analizados los rendimientos de las pruebas con las restantes imágenes del grupo test (con nódulos sólidos o semisólidos) se ha observado que el comportamiento e interacción entre los parámetros de entrada es tal que se deben utilizar 3 ó 4 clases como número de clústeres, y valores tanto de L1 como de L2 al menos iguales a 10000. No se deben utilizar

más de cuatro clases porque la función de pertenencia con un número de clases mayor no es capaz de producir una división del espacio del problema en la que los píxeles de los nódulos queden claramente separados del resto. Y los valores para las expresiones de regularización L1 y L2 no deben ser inferiores a 10000 porque es necesario el suavizado en la imagen para no incrementar los falsos positivos en el entorno del nódulo contenido en las imágenes.

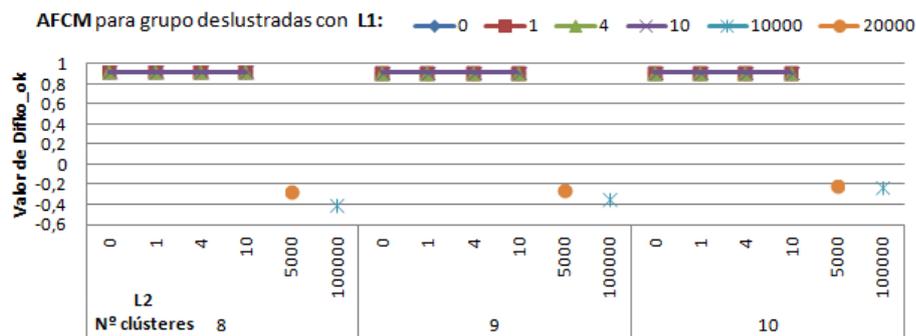


Figura 55. Rendimiento de las pruebas con el algoritmo AFCM para imágenes con nódulos en vidrio deslustrado.

Los datos obtenidos, agrupadas las imágenes según el diámetro del nódulo contenido en las tomografías, indican que en aquellas con nódulo de mayor tamaño, es adecuado utilizar cuatro clases en el proceso de segmentación, mientras que en aquellas otras con nódulos de diámetro menor se obtienen mejores resultados utilizando sólo tres clases.

En los subgrupos de imágenes con nódulos de diámetro subcentimétrico, se puede incrementar el rendimiento de la segmentación en tres clases, ajustando los valores de L1 y L2. Se observó que con valores de L2 muy bajos ( $L2 < 20$ ), disminuye el rendimiento de la segmentación en los nódulos de mayor tamaño, si se incrementa el número de clústeres. Esto último también lo justifica el hecho de que, al aumentar el valor de los parámetros lambda (L1, L2) se incrementa el rendimiento del resultado para aquellas imágenes con nódulos mayores del centímetro segmentadas en tres clases; a consta de disminuirlo en las imágenes con nódulos sub-centimétricos.

Para los subgrupos de imágenes con nódulos de formas lobuladas y espiculadas, y para los de localización subpleural se indica el comportamiento del rendimiento de la segmentación en función del parámetro L2 en la Figura 56.

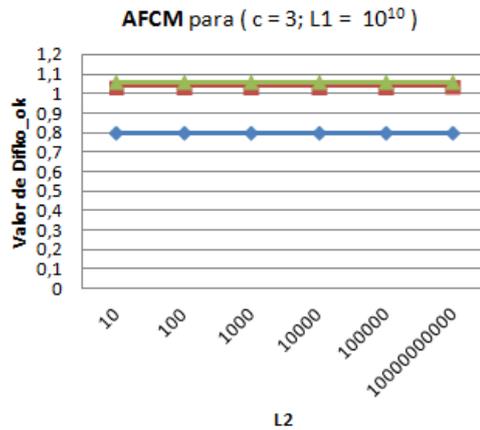


Figura 56. Rendimiento de las pruebas con el algoritmo AFCM para imágenes segmentadas con nódulos lobulados, espiculados y subpleurales.

En casi todas las pruebas el rendimiento más elevado se obtiene segmentando en tres clústeres (independientemente de los demás parámetros), por ello se analizó el comportamiento de los parámetros L1 y L2 para un valor constante de  $c = 3$ . Se ha observado que para valores de L1 altos ( $L1 > 10000$ ) el resultado es el mismo para cualquier valor de L2, al igual que sucede con las imágenes con nódulos de diferentes tamaños.

Finalmente se muestra mediante la Figura 57 como el comportamiento del algoritmo AFCM, para ambas expresiones de regularización L1 y L2 con valores muy altos, es el mismo que el del método de segmentación FCM; presentándose en concreto los valores de rendimiento para la segmentación de imágenes con nódulos de diferentes tamaños.

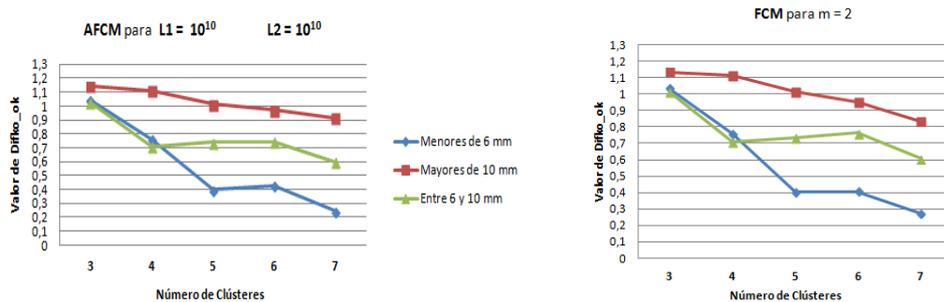


Figura 57. Rendimientos de los algoritmos AFCM con valores muy elevados de L1 y L2 y FCM con factor fuzzy igual a 2.

#### 4.1.2.5 Parámetros de entrada al algoritmo SKFCM

Con este algoritmo son muchos los parámetros que se han ajustado, para obtener un buen rendimiento. En la Tabla 9, se ha indicado que los parámetros de entrada para la ejecución del algoritmo han sido: tipo de muestra, número de clústeres, factor fuzzy, alfa, sigma y ventana de píxeles vecinos.

En el análisis del comportamiento del rendimiento de las pruebas frente al parámetro número de clústeres se ha observado la misma tendencia que con los anteriores algoritmos comentados, dicho rendimiento disminuye a medida que se incrementan las clases al segmentar las imágenes del grupo general. Siendo el citado parámetro número de clústeres más influyente en los resultados que el factor fuzzy. Esto se muestra en las gráficas de la Figura 58, en las que el parámetro número de clústeres corresponde al eje de abscisas, y las distintas series muestran rendimientos para diversos valores fuzzy indicados en la leyenda del área gráfica.

La influencia del tipo de muestra indica que se percibe una pequeña mejoría en el rendimiento al utilizar como semilla, para una segmentación en tres clústeres, el tipo de muestra manual.

Los parámetros alfa( $a$ ), sigma( $s$ ), ventana ( $v$ ), y factor fuzzy ( $m$ ) se utilizan para calibrar la maximización de los aciertos en el nódulo, y minimizar los fallos en su entorno y por consiguiente en el resto de la imagen, permitiendo realizar un ajuste fino.

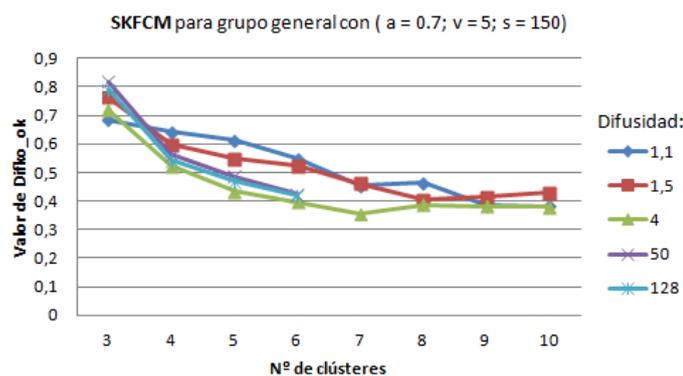


Figura 58. Rendimiento de las pruebas con el algoritmo SKFCM con respecto al número de clústeres.

También este algoritmo muestra diferente comportamiento en la segmentación, según el tipo de densidad del nódulo presente en las imágenes. Así, se hace necesario utilizar diferentes valores de parámetros para segmentar las imágenes con nódulos en vidrio

deslustrado. Esta diferencia en el comportamiento de las pruebas con el algoritmo, según el tipo de nódulo de las imágenes a segmentar, se muestra en la Figura 59a-b.

El rendimiento de las pruebas para las imágenes con NPS en vidrio deslustrado aumenta cuando en la segmentación se incrementa el número de clases. Y con las demás imágenes se obtienen mejores resultados al decrementar el número de clústeres.

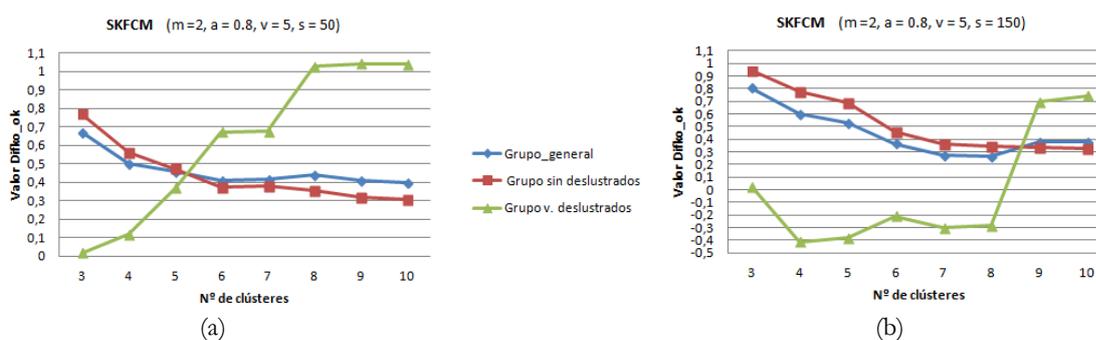


Figura 59a-b. Rendimiento de las pruebas con el algoritmo SKFCM según densidad de nódulo contenido en las imágenes.

Manteniendo constante el valor del factor fuzzy, las imágenes con NPS deslustrados obtienen mejor segmentación para el clúster nódulo si se disminuye el valor del parámetro sigma. Para las demás imágenes es necesario aumentar el valor en dicho parámetro. Ello indica que el valor de entrada sigma, después del número de clústeres y del factor fuzzy, es crítico, dando lugar a importantes variaciones en la clasificación de la ROI.

En las imágenes con nódulos sólidos ó semisólidos (grupo sin deslustrados), para valores de  $m \leq 4$ , el rendimiento de las pruebas empeora cuanto mayor es el tamaño del parámetro ventana. Sin embargo para valores altos del factor fuzzy ( $m \geq 50$ ) el rendimiento, cuando aumenta el parámetro ventana, mejora ligeramente (Figura 60 para el grupo general) o se mantiene según la tipología de nódulo contenido en las imágenes.

Al incrementar el valor del parámetro alfa, en la segmentación de imágenes con nódulos sólidos y/o semisólidos, mejora ligeramente el rendimiento de las pruebas con el algoritmo. Ello puede ser debido a que los nódulos no están perfectamente definidos y se necesita un valor de penalización alfa alto. Aumentando el valor de sigma en la gaussiana del kernel, mejora el rendimiento en la segmentación de los nódulos de diámetro menor al centímetro; mientras que aislar en un clúster, un nódulo mayor de 10mm, empeora el rendimiento si aumenta sigma.

Y el parámetro ventana de pixeles tiende a mejorar el rendimiento en torno a un tamaño 5x5, y en aquellas imágenes con nódulos subpleurales se necesita una ventana de más

vecinos para una mejor definición del clúster. Siendo este último subgrupo de imágenes, con el que más bajo rendimiento se obtiene. Para conseguir mejores resultados con dicho subgrupo es conveniente aumentar el factor fuzzy a un valor 50, y también aumentar el tamaño de la ventana de pixeles vecinos ( $v = 10$ ) con sigma 150 para incrementar algo más su rendimiento.

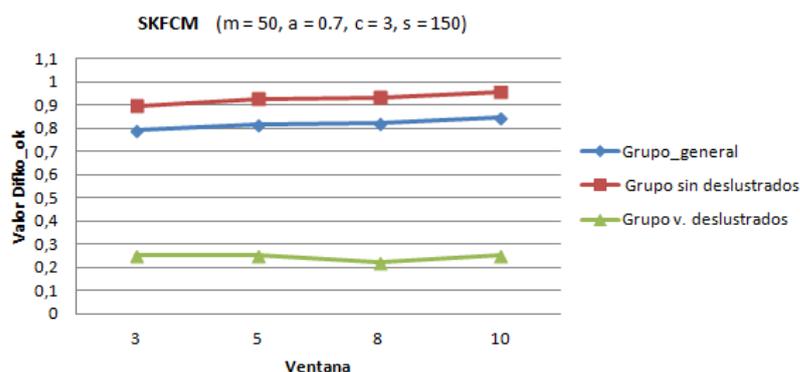


Figura 60. Rendimiento de las pruebas con el algoritmo SKFCM en función del tamaño de ventana.

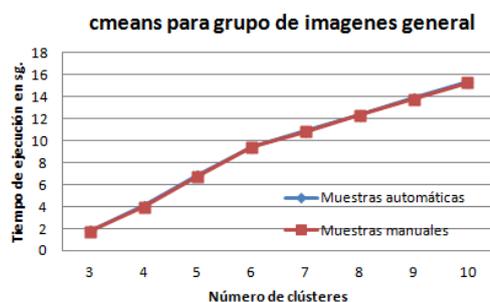
El orden de influencia en la variación de los valores de los parámetros, que influyen en el rendimiento de las pruebas con el algoritmo al procesar las imágenes ha sido en primer lugar el número de clases, a continuación el factor fuzzy seguido de sigma y alfa, y finalmente la ventana de pixeles vecinos.

#### 4.1.2.6 Relación entre el tiempo de ejecución y los parámetros

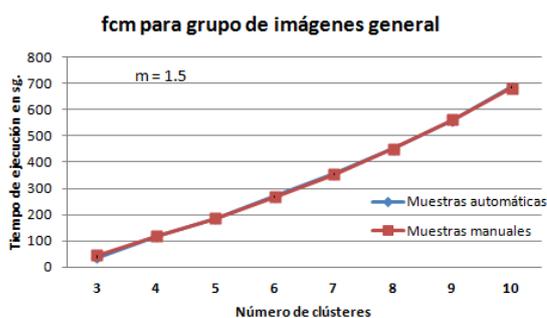
Durante la realización de este estudio, uno de los mayores inconvenientes ha sido la cantidad de tiempo necesario para el procesamiento de los algoritmos de segmentación. El tiempo empleado para segmentar cada imagen depende del algoritmo a procesar, teniendo especial relevancia los valores de entrada variables de cada algoritmo; dicho tiempo de segmentación cambia según la combinación de parámetros empleada, véase en la Figura 61a-b y la Figura 62-c la influencia del número de clústeres. Estas diferencias en tiempo de ejecución son debidas a que las precondiciones de los algoritmos influyen en el número de iteraciones que deben realizar para alcanzar la convergencia.

Los datos de las gráficas mostrados en este apartado corresponden a la media aritmética obtenida de los tiempos de ejecución de las pruebas realizadas sobre los diversos grupos de imágenes y con las ocurrencias de parámetros indicadas en ellas.

Utilizando iguales valores en los parámetros de segmentación sobre los mismos grupos de imágenes, en cada algoritmo, con los dos tipos de muestras no se han visto diferencias significativas en cuanto al tiempo de ejecución de las pruebas realizadas, esto también se muestra gráficamente en la Figura 61a-b y Figura 62a-c, donde además, se puede apreciar la dependencia polinómica existente con el parámetro Número de Clústeres.



(a)



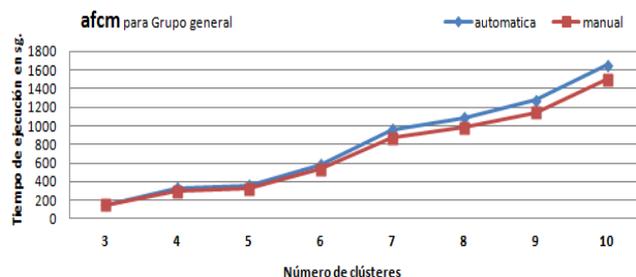
(b)

Figura 61a-b. Resultados de la media del tiempo de ejecución con los algoritmos CMeans y FCM frente al número de clústeres, utilizando como tipo de semillas de centroides muestras automáticas y manuales.

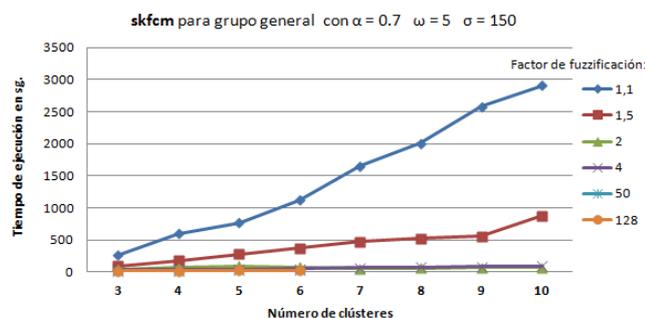
Los valores de los demás parámetros de entrada al algoritmo Adaptive Fuzzy C-Means también tienen un peso muy importante, pues, dependiendo de los valores en los términos de regularización de primer y segundo orden ( $L1$  y  $L2$ ), así será la cantidad de tiempo necesaria para la ejecución del algoritmo. En la Figura 62a se aprecia como para las mismas ocurrencias de parámetros, a mayor número de clústeres se sigue verificando que mayor es el tiempo de ejecución empleado en la segmentación.

En el tiempo de ejecución del algoritmo Spatially Kernelized Fuzzy C-Means, influye más el factor fuzzy ( $m$ ) que el número de clústeres; en la Figura 62b se muestra como para valores del factor fuzzy bajos,  $m \in \{1.1, 1.5\}$ , el incremento del tiempo en función del número de clústeres es más acentuado, y a medida que  $m$  aumenta disminuye el tiempo de

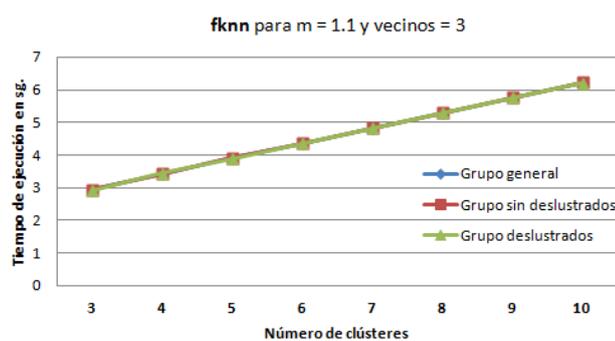
ejecución, aunque sigue siendo mayor a medida que hay más clústeres por segmentación. A partir de un valor fuzzy mayor o igual a 2 el tiempo de ejecución no es tan dispar. Los tiempos de ejecución en el algoritmo Fuzzy K-Nearest Neighbor son menores que en los demás algoritmos, aunque también se incrementan al aumentar el número de clústeres en la segmentación, tal como se ilustra en la Figura 62c.



(a)



(b)



(c)

Figura 62a-c. Resultados de la media del tiempo de ejecución de pruebas realizadas con los algoritmos AFCM, SKFCM y FKNN.

Los valores de tiempo consumido se indican más a nivel orientativo que de precisión, pues estos valores permiten ver las diferencias de tiempos medios de ejecución, entre los algoritmos empleados para la segmentación de las imágenes utilizadas. En la Tabla 24 se indican los valores medio de tiempo de ejecución. La estimación de tiempo realizada se

corresponde únicamente con el tiempo de ejecución del algoritmo de segmentación. Para segmentar una imagen en tres clústeres el tiempo medio empleado por el algoritmo CMeans es 1.75 sg, por el algoritmo Fuzzy K-Nearest Neighbor es 2.94 sg, con el Fuzzy C-Means se emplea una media de tiempo de 45.30 sg, con el Spatially Kernelized Fuzzy C-Means ya es necesario utilizar una media de tiempo de 97.35 sg, y con el algoritmo Adaptive Fuzzy C-Means el tiempo medio en segundos se eleva a 1050.54 sg.

Se ha buscado un ajuste para las curvas del tiempo de ejecución de los algoritmos frente al número de clústeres, por ser este parámetro el más influyente y determinante como se ha visto anteriormente. El mayor límite de confianza en el ajuste se consigue con curvas polinómicas. En la Figura 63 se muestra el resultado de realizar un ajuste lineal para cada uno de los algoritmos empleados.

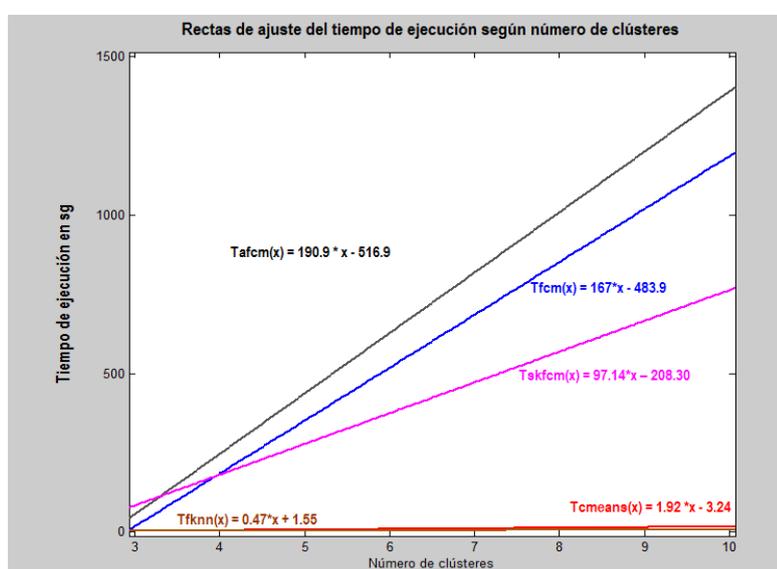


Figura 63. Representación del ajuste lineal del tiempo de ejecución de los algoritmos en función del número de clústeres empleado en la segmentación.

algoritmo	C = 3	C = 4	C = 5	C = 6	C = 7	C = 8	C = 9	C = 10
CMeans	1.75	3.97	6.76	9.43	10.81	12.29	13.74	15.27
FCM	45.30	117.05	185.93	268.35	354.08	451.56	561.13	628.11
FKNN	2.94	3.43	3.90	4.36	4.82	5.29	5.76	6.22
AFCM	1050.54	1488.21	2191.88	2219.13	2278.82	2306.32	3379.91	3606.50
SKFCM	97.35	183.60	284.94	376.68	479.30	527.50	559.09	876.25

Tabla 24 Consumo del recurso tiempo promedio de ejecución, en sg, de los algoritmos de segmentación en las pruebas más significativas.

### 4.1.3 Descriptivos del rendimiento de las pruebas seleccionadas

Para reforzar de forma objetiva el análisis de los resultados, y considerar a la vez los aciertos y fallos, se han empleado los valores de los índices Difko\_ok y Azope obtenidos a partir de las pruebas de segmentación, realizadas con las imágenes del grupo test o prueba, para calcular sus parámetros estadísticos. Así pues, para tener un enfoque descriptivo se han calculado los valores máximo, mínimo, media, mediana, desviación estándar y desviación de la mediana, además de las medidas de forma simetría y curtosis de los índices de rendimiento de las pruebas realizadas.

Se muestran resumidas por cada algoritmo, en forma tabular, aquellas pruebas donde mejores resultados se han obtenido para cada una de las imágenes en particular, y para el grupo de imágenes en general. En las correspondientes tablas sólo se muestran aquellos datos cuyos valores tanto de media como de mediana alcanzaron, como mínimo, una valoración moderada ( $\text{Difko\_ok} \geq 0.58$ ).

En concreto para el algoritmo Fuzzy CMeans se presentan en primer lugar los resultados a nivel de cada imagen del grupo; dichos resultados corresponden a una prueba ejecutada con una combinación de valores de parámetros de entrada específicos, describiendo por cada imagen segmentada los valores para los índices de rendimiento Difko\_ok y Azope, la valoración cualitativa asignada a los mismos, ya indicadas en la Tabla 14 y en la Tabla 15 respectivamente, y además las cuantías de aciertos y fallos.

Para todos los algoritmos se especifican los mejores resultados obtenidos, en cada prueba con todo el grupo de imágenes test, mostrando la media aritmética y la desviación típica estimada, dado que ambos parámetros se utilizan de manera extensa como medidas de tendencia central y dispersión, por ser dos de las medidas más estables; pero al presentar los datos marcadas asimetrías, a pesar de su uso generalizado, se considera conveniente añadir otros estadísticos que las complementen para evitar sesgos en la interpretación de dichos resultados. El hecho de que los resultados presenten asimetrías con todos los métodos empleados es debido a que las imágenes segmentadas provienen de diferentes estudios, de diferentes pacientes y con tipologías de nódulos muy variables.

#### 4.1.3.1 Resultados con el algoritmo CMeans

Los índices considerados y calculados de cada segmentación concreta, realizada con este algoritmo, han sido utilizados para el cálculo de los parámetros estadísticos, con la finalidad de medir el rendimiento de las pruebas sobre el conjunto total de imágenes test.

El rango de resultados para estas pruebas, por ejemplo la prueba con identificador 2 de la Tabla 25 y de la Tabla 26, es muy amplio (0.90) oscilando entre valoraciones malas (-1.05) y muy buenas (1.28). Se alcanzan medias buenas (0.88) en algunas pruebas, pero con un grado de dispersión alto (0.48). Los valores del parámetro mediana son superiores cuantitativamente (1.09) y cualitativamente siguen teniendo una valoración buena.

<b>ID_CMeans</b>	<b>Mínimo</b>	<b>Máximo</b>	<b>Media</b>	<b>Mediana</b>
<b>1</b>	-0,35	1,30	0,88	1,09
<b>2</b>	-1,05	1,28	0,62	0,90
<b>5</b>	-0,75	1,14	0,35	0,43
<b>9</b>	-0,35	1,30	0,88	1,09
<b>10</b>	-1,05	1,32	0,62	0,90
<b>17</b>	-0,86	1,30	0,80	1,09
<b>26</b>	-0,86	1,30	0,80	1,09

**Tabla 25 Medidas de centralización del índice Difko\_ok de los resultados de segmentación seleccionados para el algoritmo CMEANS sobre el grupo completo de imágenes de prueba.**

Los valores de la desviación estándar típica indican gran dispersión o heterogeneidad en los resultados, lo cual indica la existencia de subgrupos con valores demasiado bajos.

La distribución de los resultados es asimétrica a la izquierda para todas las pruebas consideradas al ser negativa su simetría, situándose la mediana de todas las pruebas a la derecha de la media. Analizando las concentraciones de datos en torno a la media (parámetro curtosis) se obtienen distribuciones leptocúrticas (mayores de 3) en todas las pruebas. Es decir, los resultados están concentrados alrededor de los valores centrales (entre la media y la mediana), habiendo más resultados con valores superiores a la media, por estar situados a su derecha.

<b>ID_CMeans</b>	<b>Amplitud</b>	<b>Desviación mediana</b>	<b>Desviación. estándar</b>	<b>Simetría</b>	<b>Curtosis</b>
<b>1</b>	1,66	0,20	0,48	-1,37	3,86
<b>2</b>	2,33	0,31	0,70	-1,27	3,58
<b>5</b>	1,89	0,22	0,47	-0,63	3,03
<b>9</b>	1,66	0,20	0,48	-1,37	3,86
<b>10</b>	2,37	0,31	0,70	-1,26	3,56
<b>17</b>	2,17	0,20	0,60	-1,44	4,08
<b>26</b>	2,17	0,20	0,60	-1,44	4,08

**Tabla 26 Medidas de dispersión y forma del índice Difko\_ok de los resultados de segmentación seleccionados para el algoritmo CMEANS sobre el grupo completo de imágenes de prueba.**

En general con el algoritmo CMeans, las pruebas seleccionadas presentan distribuciones asimétricas, y las medias estadísticas están afectadas por puntuaciones muy extremas, porque los valores de las desviaciones típicas indican que las puntuaciones individuales se apartan mucho de la media.

#### 4.1.3.2 Resultados con el algoritmo Fuzzy CMeans

Se señalan en la Tabla 27, los valores de los índices de rendimiento, y de las fracciones de aciertos y fallos de la Prueba9. Los parámetros estadísticos de tendencia central se muestran en la Tabla 28, y los parámetros de dispersión y forma en la Tabla 29, todos ellos, correspondientes a dicha prueba (ID\_FCM = 9), y a las demás seleccionadas por alcanzar una valoración moderada ( $Difko\_ok \geq 0.58$ ) tanto en la media como en la mediana.

FCM Prueba9	IdImagen	Azope	Difko_ok	Valoración	TPf	Fpef	FPrf
	Imagen23_16	0,97	1,31	Muy buena	0,98	0,05	0,00
	Imagen08_8	0,96	1,29	Muy buena	1,00	0,07	0,01
	Imagen17_5	0,96	1,29	Muy buena	0,94	0,02	0,01
	Imagen15_18	0,96	1,29	Muy buena	0,98	0,07	0,01
	Imagen12_7	0,97	1,30	Muy buena	0,99	0,06	0,02
	Imagen05_9	0,95	1,26	Muy buena	0,96	0,06	0,01
	Imagen06_5	0,95	1,25	Muy buena	1,00	0,10	0,02
	Imagen19_23	0,96	1,28	Muy buena	0,99	0,07	0,01
	Imagen14_5	0,93	1,18	Muy buena	0,88	0,03	0,02
	Imagen20_8	0,91	1,12	Buena	0,99	0,17	0,02
	Imagen04_13	0,89	1,07	Buena	0,85	0,08	0,02
	Imagen13_4	0,84	0,91	Buena	0,70	0,02	0,01
	Imagen16_11	0,91	1,12	Buena	1,00	0,17	0,01
	Imagen09_17	0,83	0,85	Buena	1,00	0,34	0,00
	Imagen11_8	0,84	0,89	Buena	1,00	0,32	0,02
	Imagen22_10	0,84	0,89	Buena	1,00	0,32	0,01
	Imagen03_10	0,78	0,78	Moderada	0,86	0,30	0,02
	Imagen01_10	0,80	0,75	Moderada	0,99	0,40	0,01
	Imagen10_30	0,76	0,66	Moderada	1,00	0,47	0,01
	Imagen02_8	0,70	0,50	Baja	0,88	0,50	0,03
	Imagen07_11	0,49	-0,01	Mala	0,00	0,01	0,83
	Imagen18_22	0,45	-0,12	Mala	0,73	0,84	0,19
	Imagen21_15	0,35	-0,35	Mala	0,66	0,96	0,16

Tabla 27 Valores de los índices, aciertos y fallos obtenidos como resultado de la segmentación con el algoritmo FCM para 3 clases y factor fuzzy 1.5 utilizando las imágenes del conjunto de pruebas.

Los valores específicos de la Prueba9 se han calculado a partir de los resultados de la segmentación en tres clústeres, y empleando un factor fuzzy,  $m = 1.5$ , sobre las imágenes del conjunto test. De las veintitrés imágenes segmentadas por la Prueba9, hay nueve resultados con valoración muy buena, siete con valoración buena, tres con valoración moderada, y como valores muy extremos que alteran la media del resultado una baja y dos malas.

Con este algoritmo, los descriptivos calculados (resumidos en la Tabla 28 y en la Tabla 29), tienen un recorrido amplio pues van desde valores mínimos malos ( $<0$ ) hasta valoraciones buenas tanto en medias (0.92) como en medianas (1.09). Los valores de las medianas son superiores, en todas las ocurrencias, a los valores de las medias porque los resultados se encuentran, en todas las pruebas seleccionadas, distribuidos de forma asimétrica con cola por la izquierda (simetría  $< 0$ ). Excepto la prueba 70, todas las demás muestran una gran concentración de datos en la región central de la distribución (curtosis  $> 3$ ). La prueba con identificador 70 es platicúrtica, es decir, en la región central no tiene una alta concentración de datos. Esta prueba corresponde a segmentaciones realizadas empleando un número de clústeres igual a 6, y con un factor fuzzy de valor 9; estas precondiciones en el algoritmo FCM implican una distribución de los datos menos concentrada. Todas las demás pruebas corresponden a segmentaciones en 3 ó 4 clústeres y con unos valores de factor fuzzy entre 1.1 y 2.

ID_FCM	Mínimo	Máximo	Media	Mediana
9	-0,35	1,31	0,89	1,07
17	-0,40	1,31	0,88	1,05
18	-1,14	1,30	0,62	0,91
25	-0,35	1,30	0,88	1,09
26	-1,07	1,32	0,62	0,90
41	-0,32	1,33	0,89	1,07
49	-0,12	1,36	0,92	1,09
70	-0,50	1,32	0,59	0,62

**Tabla 28 Medidas de centralización del índice Difko\_ok de los resultados de segmentación seleccionados del algoritmo FCM sobre el grupo completo de imágenes de prueba.**

Las desviaciones típicas indican que los resultados individuales se apartan mucho de la media. Para ilustrar estas heterogeneidades véase en concreto la Figura 64, en la que se muestran las ROI de imágenes segmentadas, y sus respectivas máscaras, correspondientes con los valores estadísticos mínimo, mediana y máximo de la Prueba con ID\_FCM = 9.

ID_FCM	Amplitud	Desviación mediana	Desviación estándar	Simetría	Curtosis
9	1,66	0,22	0,48	-1,35	3,87
17	1,71	0,22	0,48	-1,42	4,07
18	2,44	0,30	0,71	-1,12	3,17
25	1,66	0,20	0,48	-1,37	3,87
26	2,40	0,31	0,70	-1,24	3,52
41	1,65	0,22	0,48	-1,32	3,75
49	1,49	0,19	0,45	-1,21	3,31
70	1,82	0,40	0,51	-0,30	2,14

Tabla 29 Medidas de dispersión y forma del índice Difko\_ok de los resultados de segmentación seleccionados del algoritmo FCM sobre el grupo completo de imágenes de prueba.

Las ocurrencias de estadísticos correspondientes a la Prueba9 aparecen en la Tabla 28 y la Tabla 29 y sus valores de aciertos y fallos junto con sus valores de índices de rendimiento están en la Tabla 27.

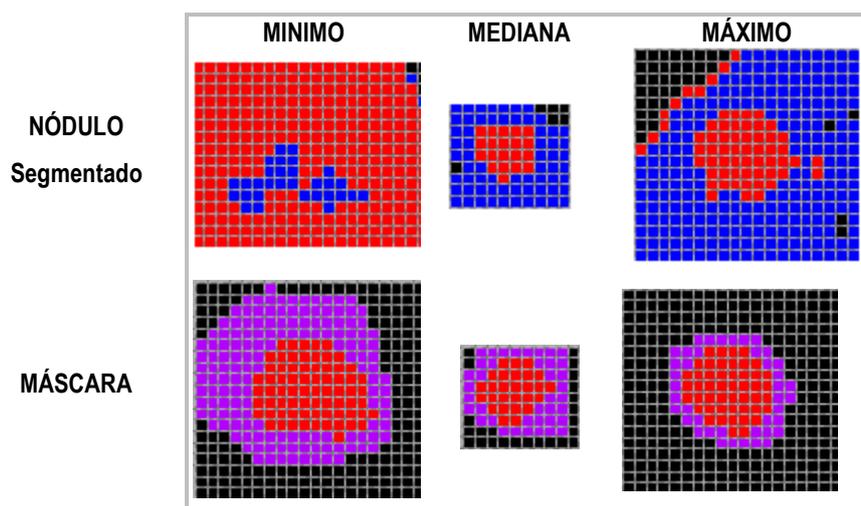


Figura 64. Visualización de segmentaciones de NPS de la Prueba9 de la Tabla 28.

#### 4.1.3.3 Resultados con el algoritmo Fuzzy K-Nearest Neighbor

En la Tabla 30 y en la Tabla 31 se muestran los estadísticos referidos al índice Difko\_ok, de aquellas pruebas que han proporcionado los mejores resultados en los parámetros estadísticos de tendencia central media y mediana ( $\text{Difko\_ok} \geq 0.58$ ), correspondientes al algoritmo Fuzzy K-Nearest Neighbor.

Con este algoritmo ninguna de las pruebas seleccionadas tiene valoración mala, pues el valor mínimo para estos resultados no es negativo, a diferencia de lo que sucedía con el algoritmo CMeans, y los valores máximos de los resultados oscilan dentro de una

valoración muy buena aunque las medias calculadas en ningún caso llegan a tener dicho acuerdo cualitativo. El valor alcanzado en las medias es bueno ( $\text{media} \geq 0.85$ ) en las pruebas con identificador 41 y 2017, y moderado en las demás. En todas las pruebas la media es inferior a la mediana y los valores de simetría son negativos, esto corrobora que los valores de los resultados están más concentrados a la derecha de la media que a su izquierda. Como el valor de curtosis es negativo en todas los casos, y siendo la distribución con asimetría a la izquierda, los datos no están concentrados entre la media y la mediana; de esto se deduce que existen bastantes valores con valoración buena y muy buena.

ID_FKNN	Mínimo	Máximo	Media	Mediana
33	0,00	1,31	0,82	0,97
41	0,00	1,36	0,87	1,08
57	0,00	1,30	0,78	0,97
58	0,00	1,30	0,77	0,88
2017	0,00	1,34	0,90	1,13
2027	0,00	1,24	0,74	0,79
2033	0,00	1,31	0,80	0,93
2049	0,00	1,30	0,88	1,15
2059	0,00	1,27	0,68	0,76

Tabla 30. Medidas de centralización del índice Difko\_ok de los resultados de segmentación seleccionados del algoritmo FKNN sobre el grupo completo de imágenes de prueba.

ID_FKNN	Amplitud	Desviación. mediana	Desviación. estándar	Simetría	Curtosis
33	1,31	0,29	0,43	-0,61	1,99
41	1,36	0,20	0,44	-0,68	2,09
57	1,30	0,29	0,45	-0,56	1,84
58	1,30	0,26	0,37	-0,70	2,67
2017	1,34	0,17	0,45	-0,86	2,31
2027	1,24	0,20	0,36	-0,67	2,61
2033	1,31	0,30	0,44	-0,56	1,86
2049	1,30	0,14	0,46	-0,82	2,20
2059	1,27	0,24	0,35	-0,50	2,41

Tabla 31 Medidas de dispersión y forma del índice DifKo\_ok de los resultados de segmentación seleccionados del algoritmo FKNN sobre el grupo completo de imágenes de prueba.

Observando las desviaciones estándar se aprecia claramente que los datos se encuentran alejados de las respectivas medias mayoritariamente. Como la forma es platicúrtica con cola por la izquierda los resultados tienen mayoritariamente valores bastante superiores a la media, pero también existen valores extremos con resultados que disminuyen dicha media.

#### 4.1.3.4 Resultados con el algoritmo Adaptive Fuzzy CMeans

El algoritmo Adaptive Fuzzy C-Means ha sido codificado con un factor constante 2 de valor fuzzy, y por tanto este factor no se utilizó como parámetro de entrada variable. Así, en las pruebas de segmentación con este algoritmo, no se pudieron probar segmentaciones con diferentes valores de factor fuzzy, por ello, es probable que se tengan valores mínimos negativos debido a que se han restado posibilidades de precisar un ajuste más fino en los resultados de las segmentaciones para el clúster nódulo, algo que si se ha realizado con los demás algoritmos, al variar el valor de dicho parámetro.

Al igual que sucedía con el algoritmo CMeans, las pruebas proporcionan resultados con un rango de valores muy amplio: en la Prueba55 se alcanzó una diferencia entre el valor mínimo y máximo del rendimiento de 2.43, véanse Tabla 32 y Tabla 33. En todas las pruebas el valor de la mediana es superior a la media. Y dichas medias alcanzan valoraciones mayoritariamente buenas.

ID_AFCM	Mínimo	Máximo	Media	Mediana
2	-0,07	1,29	0,86	0,97
9	-0,40	1,33	0,88	1,09
17	-1,07	1,30	0,77	0,95
33	-0,40	1,33	0,88	1,09
48	-0,38	1,31	0,89	1,07
51	-0,40	1,31	0,89	1,09
55	-1,12	1,29	0,62	0,93
85	-0,38	1,31	0,89	1,08
107	-1,01	1,29	0,78	0,99
108	-0,94	1,29	0,60	1,03
111	-0,72	1,31	0,84	1,07
112	-0,75	1,36	0,63	1,08
2001	0,00	1,32	0,93	1,13
2011	0,00	1,34	0,89	1,12

Tabla 32. Medidas de centralización del índice Difko\_ok de los resultados de segmentación seleccionados del algoritmo AFCM sobre el grupo completo de imágenes de prueba.

Los resultados, en todos los casos, se encuentran distribuidos de forma asimétrica con cola a la izquierda; estando los resultados de las pruebas mayoritariamente concentrados entre las medidas de tendencia central.

Analizando la desviación estándar se puede afirmar que entre las mejores pruebas seleccionadas con este algoritmo, hay resultados de segmentaciones de las imágenes que se encuentran alejados de sus respectivas medias. Estando los valores concentrados

mayoritariamente por encima de la media, teniendo valores mínimos tan extremos y habiendo pocos valores por debajo de la media, se concluye que en general con este algoritmo se obtienen buenos resultados.

ID_AFCM	Amplitud	Desviación mediana	Desviación estándar	Simetría	Curtosis
2	1,37	0,24	0,43	-1,06	2,90
9	1,73	0,20	0,49	-1,54	4,36
17	2,37	0,18	0,61	-1,92	5,87
33	1,73	0,20	0,49	-1,54	4,34
48	1,69	0,22	0,48	-1,41	4,07
51	1,71	0,20	0,48	-1,45	4,17
55	2,41	0,30	0,71	-1,06	2,97
85	1,70	0,22	0,48	-1,39	3,99
107	2,30	0,18	0,61	-1,91	5,76
108	2,23	0,25	0,79	-0,80	1,98
111	2,03	0,18	0,57	-1,82	5,30
112	2,12	0,22	0,76	-0,76	1,88
2001	1,32	0,15	0,44	-1,14	2,89
2011	1,34	0,17	0,44	-0,98	2,56

Tabla 33. Medidas de dispersión y forma del índice Difko\_ok de los resultados de segmentación seleccionados del algoritmo AFCM sobre el grupo completo de imágenes de prueba.

#### 4.1.3.5 Resultados con el algoritmo Spatially Kernelized Fuzzy CMeans

En la Tabla 34 y en la Tabla 35 se muestran las medidas de centralización y dispersión de los valores Difko\_ok resultado de las pruebas seleccionadas que fueron ejecutadas con el algoritmo Spatially Kernelized Fuzzy C-Means. El rango de datos entre el mínimo valor alcanzado por la imagen con peores resultados y aquella con la que mejores resultados se han conseguido para la misma prueba es amplio, debido a que en la mayoría de las pruebas siempre hay alguna imagen con valoración mala en la clasificación de los píxeles del clúster nódulo (véase como ejemplo la Prueba18).

Las pruebas (con identificadores 9, 69, 77) en las que se consiguen valoraciones buenas tanto para las medias como para las medianas, presentan los datos poco concentrados entre las medidas centrales.

Las pruebas leptocúrticas que presentan una mayor concentración de datos entre la media y la mediana (curtosis  $> 3$ ) tienen desviaciones estándar altas.

ID_SKFCM	Mínimo	Máximo	Media	Mediana
9	-0,02	1,36	0,92	1,05
17	0,00	1,34	0,76	0,86
18	-1,15	1,27	0,60	0,80
57	0,00	1,29	0,67	0,70
69	0,00	1,31	0,82	0,94
77	0,00	1,36	0,85	0,94
2010	-1,12	1,31	0,64	0,88
2034	-0,79	1,30	0,60	0,84
2057	-0,01	1,35	0,68	0,80

Tabla 34. Medidas de centralización del índice Difko\_ok de los resultados de segmentación seleccionados del algoritmo SKFCM sobre el grupo completo de imágenes de prueba.

ID_SKFCM	Amplitud	Desviación mediana	Desviación estándar	Simetría	Curtosis
9	1,38	0,26	0,45	-0,92	2,59
17	1,34	0,25	0,40	-0,56	2,44
18	2,42	0,30	0,68	-1,48	4,28
57	1,29	0,27	0,39	-0,36	2,06
69	1,31	0,24	0,39	-0,55	2,05
77	1,36	0,33	0,42	-0,45	1,85
2010	2,42	0,28	0,71	-1,41	3,91
2034	2,10	0,35	0,66	-0,74	2,17
2057	1,36	0,24	0,40	-0,31	2,13

Tabla 35. Medidas de dispersión y forma del índice Difko\_ok de los resultados de segmentación seleccionados del algoritmo SKFCM sobre el grupo completo de imágenes de prueba.

En general con este algoritmo, las pruebas seleccionadas presentan distribuciones asimétricas, y las medias estadísticas están afectadas por puntuaciones muy extremas, pues, los valores de las desviaciones típicas indican que hay puntuaciones individuales que se apartan mucho de las medias. Esta dispersión tan elevada en las pruebas puede ser debida a que el algoritmo discrimina bastante, detectando las diferencias existentes, téngase en cuenta que las imágenes contienen NPS de diversas características y provienen de estudios diferentes.

#### 4.1.4 Resultados por tipología de nódulo

En el apartado anterior se ha mostrado como las pruebas seleccionadas, en cada uno de los algoritmos, presentan distribuciones asimétricas, y las medias estadísticas estaban afectadas por puntuaciones extremas, dando lugar a valores en las desviaciones típicas que

indican puntuaciones individuales que se apartan mucho de la media. Como consecuencia de estos resultados, se clasificó el grupo de imágenes de análisis en varios subgrupos, atendiendo al tamaño, forma, localización y densidad del NPS que dichas imágenes contienen, con la finalidad de analizar los resultados de la segmentación de las imágenes agrupadas según la tipología del nódulo, y ver si se obtendrían resultados más homogéneos, con menor grado de dispersión.

Por lo tanto, se han analizado los rendimientos de la segmentación de los cortes tomográficos que compartían la misma tipología de nódulo; obteniéndose para estos valores una distribución menos asimétrica que con todo el conjunto de imágenes de pruebas.

Pero cuando se ejecuta un determinado algoritmo, utilizando una combinación de valores de parámetros concreta, los resultados obtenidos dan lugar a una distribución distinta a la generada por los resultados conseguidos con diferentes valores de parámetros. Por lo cual, se comparan los resultados de pruebas diferentes, realizadas empleando el mismo algoritmo sobre igual subgrupo de imágenes, para ver cual prueba destaca más y obtener así, la combinación óptima de los valores de entrada a cada uno de esos algoritmos. Por ello se tipifican las variables utilizadas, para medir comparativamente el rendimiento de aquellas pruebas de entre las que mejores resultados produjeron, y así obtener cuál de ellas produce un resultado bueno y con menos heterogeneidad para las imágenes con la tipología de nódulo considerada.

La fórmula empleada para tipificar las variables de medida es:

$$Z = \frac{x - \mu}{\sigma}$$

Donde  $x$  es la media del índice considerado y calculado, para una prueba,  $\mu$  es la media de los  $x$  para el grupo de tipología de nódulo estudiado, y  $\sigma$  es la desviación estándar de  $x$ . Así de esta forma,  $Z$  tendrá una media cuyo valor es cero y una desviación típica de valor uno, con la que resulta más fácil comparar los resultados de rendimiento en cada algoritmo. Considerando que las pruebas con mejores resultados se obtienen a partir de aquellas que obtengan una mayor puntuación normalizada.

El índice de rendimiento utilizado para calcular los valores tipificados ha sido el índice Azope, dado que es una medida con sólo valores positivos, y en el dominio del otro índice, Difko\_ok, existen valores negativos como se ha señalado en la Tabla 17.

En la Tabla 36, Tabla 37 y Tabla 38 se resumen, por cada algoritmo según la tipología de nódulo contenida en las tomografías, las tres combinaciones de valores paramétricos que

produjeron los resultados más elevados. Además por cada uno de los algoritmos, desde la Figura 65 hasta la Figura 71, se muestran recortes de la ROI de imágenes segmentadas (conteniendo NPS de las características exigidas) correspondientes a los resultados de las pruebas con el valor tipificado más alto, y para facilitar la comparación visual de dichos resultados en las regiones de interés pixel a pixel se adjuntan las pertinentes máscaras.

#### **4.1.4.1 Según tamaño de NPS**

En la clasificación por tamaños se utilizan diferentes grupos correspondiéndose a los intervalos de diámetros de nódulos (ver Tabla 36) menores de 6mm, entre 6 y 10 mm y mayores de 10mm [Castañer, 2011], [Conde, 2009].

	PARÁMETROS (diámetro>1cm)	ValorTip	PARÁMETROS ([6, 10] mm)	ValorTip	PARÁMETROS (Diámetro< 6mm)	ValorTip
CMEANS	(c = 3; tm = manual)	0,60	(c = 3; tm = manual)	0,77	(c = 3; tm = manual)	0,69
	(c = 3; tm = automática)	0,60	(c = 3; tm = automática)	0,77	(c = 3; tm = automática)	0,69
	(c = 3; tm = manual1)	0,60	(c = 3; tm = manual1)	0,77	(c = 3; tm = manual1)	0,68
FCM	(c = 3; tm = manual; m =256)	1,10	(c = 3; tm = manual; m =256)	1,31	(c = 3; tm = manual; m = 1.5)	0,89
	(c = 3; tm = manual; m = 50)	1,00	(c = 3; tm = manual; m = 50)	1,16	(c = 3; tm = manual; m = 2)	0,86
	(c = 3; tm = manual; m = 1.7)	0,85	(c = 3; tm = manual; m = 1.7)	0,79	(c = 3; tm = manual; m = 1.1)	0,81
FKNN	(c = 3; v= 6 ; n = 6 ; m = 1.1 )	1,03	(c = 3; v= 4; n = 4; m = 2)	0,71	(c = 3; v= 6 ; n = 6 ; m = 2)	1,02
	(c = 3 ; v= 4; n = 4 ; m = 1.5 )	0,82	(c = 3 ; v= 5; n = 5 ; m = 2)	0,65	(c = 3 ; v= 5; n = 5 ; m = 2)	0,95
	(c = 3; v= 4; n = 4; m = 2)	0,52	(c = 3 ; v= 4; n = 4 ; m = 1.5 )	0,60	(c = 3 ; v=4 ; n = 4 ; m = 4)	0,84
AFCM	(c = 4; L1 = 10000; L2 = 10000)	0,70	(c = 3; L1 = 10000 ; L2 = 10)	0,82	(c = 3; L1 = 50000; L2 =500000)	0,63
	(c = 3; L1 = 50000; L2 = 500000 )	0,63	(c = 3; L1 = 10000 ; L2 = 1000)	0,71	(c = 3; L1 = 50000; L2 = 300000)	0,62
	(c = 3 ; L1 = 1E+10 ; L2 = 1E+10)	0,63	(c = 4; L1 = 10000; L2 = 100000)	0,62	(c = 3; L1 = 1E+10; L2 = 1E+10)	0,60
SKFCM	(c =3 ; a = 0.7; v = 5 ; s = 150; m = 50)	0,93	(c = 3; a= 0.7; v = 5; s = 300; m = 1.5)	1,11	(c = 3; a = 0.7; v = 5; s = 300; m = 1.5)	1,31
	(c = 3; a = 0.7; v = 10; s = 150 ; m = 50)	0,91	(c = 3; a = 0.7; v = 5 ; s = 150; m = 50)	1,00	(c = 3; a = 0.7; v = 10; s = 150 ; m = 50)	1,02
	(c = 3; a = 0.7; v = 5 ; s = 150 ; m = 1.5)	0,90	(c = 3; a = 0.7; v = 10; s = 150 ; m = 50)	0,94	(c = 3 ; a = 0.7; v = 5 ; s = 150; m = 50)	0,74

Tabla 36. Resumen de pruebas óptimas por algoritmo para imágenes agrupadas por tamaños.

#### ***4.1.4.1.1 Nódulos mayores de 10 mm***

En los resultados conseguidos con el algoritmo CMeans tiene más influencia el número de clústeres que el tipo de muestra (teniendo en cuenta que el centroide inicial siempre ha sido seleccionado dentro de la zona de interés para el clúster nódulo). De hecho para igual número de clústeres los resultados son prácticamente iguales independientemente del tipo de muestras que se han elegido, esto último es debido a que las semillas iniciales son claramente pertenecientes a la región de interés.

El mejor resultado se produce con una segmentación en tres clústeres, pues aquellas donde se segmenta en más clústeres son las de valores tipificados más bajos. El valor tipificado más alto corresponde a tres clases, y al observar pixel a pixel dicha segmentación es la más fiel a la máscara. Pues los nódulos segmentados en cuatro clústeres tienen asignados los píxeles más externos de los mismos a una nueva clase, lo que implica la disminución de verdaderos positivos en la zona nódulo.

Con el algoritmo Fuzzy C-Means, para las imágenes con nódulos de diámetro mayor que 1cm, los valores tipificados más altos del rendimiento de las pruebas, se consiguen con una segmentación en tres clases y valores fuzzy mayores de 1.5. De hecho, para la división de la imagen en tres clústeres, el rendimiento es más elevado cuanto mayor es el factor fuzzy. Así pues, se ha obtenido que la prueba de mayor rendimiento para esta tipología de nódulo haya sido la que segmenta la CT en 3 clústeres y con factor fuzzy 256. Ello puede ser porque cuanto mayor es el parámetro fuzzy más pertenecen todos los píxeles a todos los clústeres.

Para el algoritmo Fuzzy K-Nearest Neighbor la mayoría de las pruebas que presentan el mejor rendimiento segmentan las imágenes en tres clústeres; la prueba con valor tipificado más alto es la correspondiente a la combinación de valores: 3 clústeres, factor fuzzy 1.1, y 6 vecinos con 6 muestras. El hecho de ser los nódulos de tamaño mayor a 1cm, indica mayor probabilidad de unas intensidades de píxeles claramente diferenciables del tejido circundante; así, para obtener los mejores resultados con este algoritmo es recomendable considerar un mayor número de vecinos y muestras, para la delimitación del clúster que constituye la región de interés. Lo más adecuado es segmentar en 3 clústeres, a pesar de incrementar los FP en el resto de la imagen puesto que al hacerlo en más clases disminuyen los TP dado que un porcentaje de los píxeles correspondientes al nódulo se asocian a otra clase, pues con este método, para esos píxeles, se encuentran más similitudes con los vecinos del borde del nódulo.

La segmentación en tres ó cuatro clústeres es la que produce mejores resultados en el algoritmo Adaptive Fuzzy C-Means para las imágenes de nódulos mayores de los 10mm. En las pruebas seleccionadas, los parámetros del algoritmo que permiten compensar las inhomogeneidades tienen valores altos para este tipo de nódulos (el campo multiplicador de primer orden -L1- no está por debajo de 10000), pues así permiten realizar un suavizado en los bordes de dichos nódulos.

Con el algoritmo Spatially Kernelized Fuzzy C-Means el mayor rendimiento para este subgrupo de imágenes se obtiene claramente segmentándolas en 3 clústeres. Para los mejores resultados es suficiente con una ventana de 5x5 para incluir la información espacial del vecindario de cada uno de los píxeles del nódulo a clasificar.

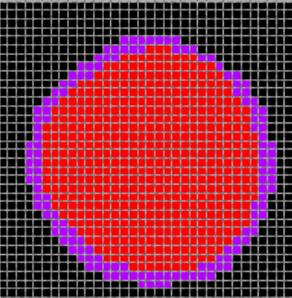
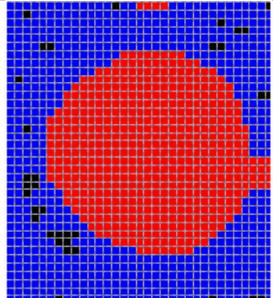
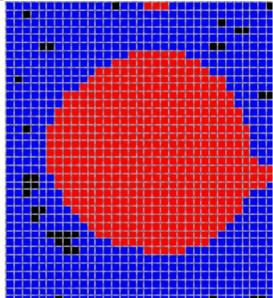
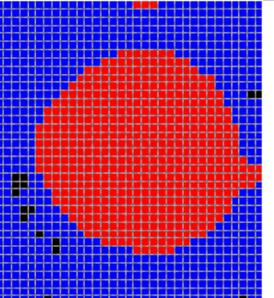
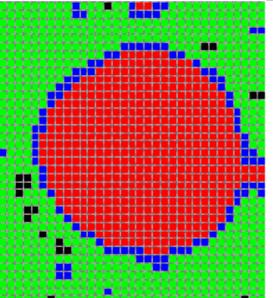
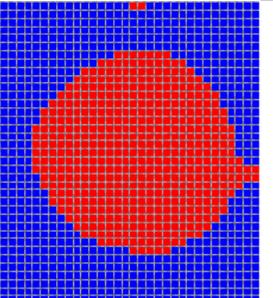
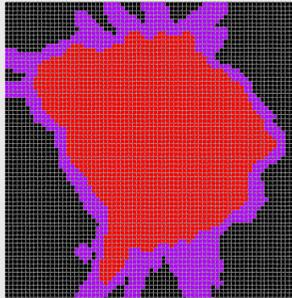
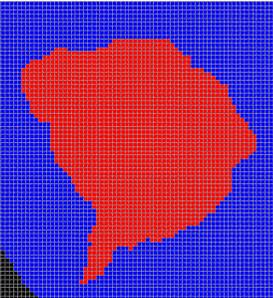
MÁSCARAS	CMEANS (c = 3; tm = manual)	FCM (c = 3; tm = manual; m = 256)	FKNN (c = 3; v = 6; n = 6; m = 1.1)	AFCM (c = 4; L1 = 10000; L2 = 10000)	SKFCM (c = 3; a = 0.7; v = 5; s = 150; m = 50)
					
					

Figura 65. ROI segmentadas por algoritmo para algunas imágenes con NPS de tamaño centimétrico.

#### ***4.1.4.1.2 Nódulos de diámetro entre 6 y 10 mm***

Utilizando el método CMeans, los píxeles correspondientes a los nódulos cuyo diámetro es inferior al centímetro pero mayor de 5 mm son clasificados de forma más precisa cuando el número de clases utilizada para segmentar la imagen es tres, no habiendo diferencias en los valores tipificados entre el tipo de muestra manual, y la calculada automáticamente buscando los máximos en el histograma.

El rendimiento de las pruebas con el algoritmo CMeans disminuye al clasificar los píxeles del corte CT en cuatro clústeres, pues al incrementar el número de clústeres los píxeles correspondientes a la zona nódulo quedan repartidos entre las clases según su nivel de intensidad, al estar este método basado en la distancia euclídea existente entre los píxeles y los centroides de cada clase. Se apreció en la segmentación de estos nódulos, empleando 4 clústeres, que la zona nódulo aparece claramente diferenciada en dos clases: una más interna y otra bordeándola no perteneciente al clúster objetivo. Sin embargo en los resultados en 3 clústeres no existe tal división, con lo que la sensibilidad de la prueba es bastante más elevada, corroborando esto el valor tipificado más alto en la mejor combinación de valores paramétricos.

En la segmentación, de las imágenes que contienen la tipología de nódulo con diámetro de tamaño intermedio, con el algoritmo Fuzzy CMeans obtienen mayor valor tipificado las pruebas en las que el número de clases empleadas ha sido tres y el tipo de muestra para los centroides iniciales es manual. El valor del factor fuzzy que mejor correspondencia establece en la pertenencia parcial de los píxeles a los clústeres es alto, siendo la prueba con mejor media tipificada (1.31) ajustada con un valor de  $m = 256$ ; de forma que este valor en el exponente de la matriz de pertenencia de la función objetivo del algoritmo, fuerza a que los píxeles más próximos al centroide pertenezcan al clúster nódulo, y los resultados tienen un mayor número de verdaderos positivos a pesar de que existe alguna imagen en la que se incrementan los falsos positivos en el entorno del nódulo.

Con el método Fuzzy K-Nearest Neighbor se debe proporcionar como parámetro de entrada, además del número de clases en que segmentar la imagen y el factor fuzzy, un número ( $n$ ) mínimo de muestras prototipo por clúster cómo de vecinos ( $v$ ) sean considerados. Los mejores resultados se han obtenido con un número de clústeres igual a tres, siendo los valores tipificados más altos los correspondientes a un valor de parámetro número de muestras por clúster igual al número de vecinos utilizado como precondition,

en todos los test cuya segmentación fue implementada con tres clases. La prueba con valor tipificado más alto corresponde con un número de vecinos igual a 4, utilizando también  $n=4$  muestras por cada uno de los tres clústeres en que se segmentaron las imágenes. Un número mayor de muestras por clúster produce peores resultados, al igual que un número mayor de vecinos debido a la similitud en los valores de intensidad entre píxeles de puntos cancerígenos y píxeles pertenecientes a la zona correspondiente a tejido sano.

El haber codificado el algoritmo Adaptive Fuzzy C-Means con un factor  $m = 2$ , condicionó a buscar la mejor combinación de parámetros de entrada variables ajustando de forma más precisa los términos de penalización L1 y L2, permitiendo este algoritmo encontrar buenos rendimientos al segmentar en cuatro clases con valores de penalización de discontinuidades muy elevados, aunque los mejores resultados siguen obteniéndose con tres clústeres como en los demás algoritmos.

Dentro del conjunto formado por las combinaciones de valores de parámetros, con las que se han obtenido valores tipificados más altos, utilizando el algoritmo Spatially Kernelized Fuzzy C-Means, están aquellas pruebas en las que se ha segmentado en tres clústeres el conjunto de imágenes CT. Para esas pruebas (segmentadas en tres clústeres) el valor de control de penalización alfa, que da lugar a las pruebas con mejor resultado, es de una cuantía 0.7, y el parámetro que permite influir en la clasificación de cada píxel según el valor de su vecindad es  $v = 5$ . Ocurre que para un valor fuzzy bajo,  $m = 1.5$ , el valor del parámetro de la función kernel más adecuado es el más alto para las pruebas seleccionadas ( $s = 300$ ), sin embargo si se emplea un valor de  $m$  alto (50), el valor ajustado para sigma debe reducirse a la mitad (150). Con valores de sigma más altos se incrementan los falsos positivos en el entorno del nódulo.

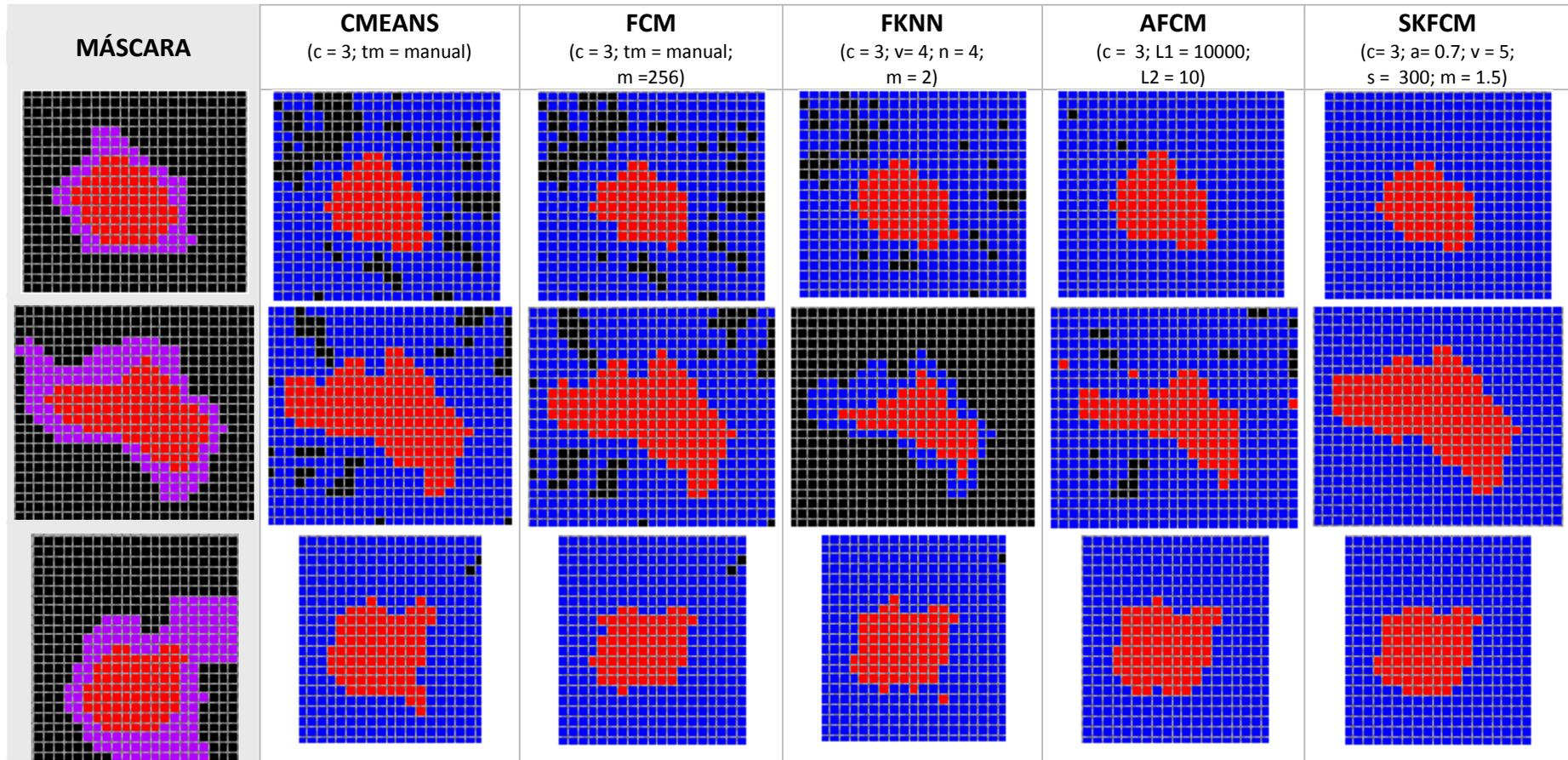


Figura 66. ROI segmentadas por algoritmo para algunas imágenes con NPS de tamaño subcentimétrico.

#### **4.1.4.1.3 Nódulos de diámetro menor a 6 mm**

Con los nódulos de tamaño inferior a 6 milímetros, los resultados obtenidos utilizando como método el algoritmo CMeans, y que permiten mayor número de píxeles clasificados como verdaderos positivos y menor número de falsos positivos, se consiguen al segmentar en tres clústeres. No observando diferencias entre utilizar centroides obtenidos manualmente y al calcularlos (tm = automática) a través del algoritmo que busca máximos locales en la ROI. Al segmentar en cuatro clases los píxeles más externos de la lesión quedan incluidos en un nuevo clúster, lo cual reduce significativamente los aciertos en la zona a aislar en una única clase. También para el algoritmo FCM, los mejores resultados se obtienen segmentando en 3 clústeres. El parámetro fuzzy que permite obtener una mayor precisión en la segmentación del clúster nódulo es con un valor 1.5, con el cual se obtienen valores tipificados de 0.89.

El algoritmo Fuzzy K-Nearest Neighbor tiene mayor rendimiento al segmentar las imágenes en tres clases, y con valor de factor fuzzy 2; el comportamiento es mejor para un número de vecinos 5 ó 6, y además en las pruebas de mejor rendimiento el número de muestras nunca supera al número de vecinos, esto último es natural al considerar que son pocos píxeles los constituyentes de estos nódulos pequeños.

En el Adaptive Fuzzy C-Means, los valores tipificados más elevados se obtienen para las pruebas en las que se ha segmentado en tres clústeres. Siendo mejores los valores tipificados cuando ambos parámetros de penalización son grandes ( $L1 = 50000$  y  $L2 = 500000$ ), lo cual puede ser indicativo de que las CT con nódulo solitario parecen bastante homogéneas, a diferencia del subconjunto formado por las imágenes conteniendo nódulo de tamaño intermedio (entre 6 y 10 mm) en las que había imágenes con NPS en vidrio deslustrado, pudiendo ser considerado esto último como heterogeneidad o ruido en la imagen.

El resumen comparativo de las pruebas que mejores resultados han permitido obtener para este subconjunto de imágenes cuando el método empleado ha sido el Spatially Kernelized Fuzzy C-Means también indica que lo adecuado es segmentar en tres clústeres. A pesar de disponer de mayor número de parámetros para ajustar la precisión en la clasificación de los píxeles, predominan los rendimientos tipificados más altos cuando en la segmentación han sido empleadas sólo tres clases, tanto con valores fuzzy convencionales ( $m=1.5$ ) como con valores altos ( $m= 50$ ); y lo mismo ha ocurrido con los valores del parámetro sigma.

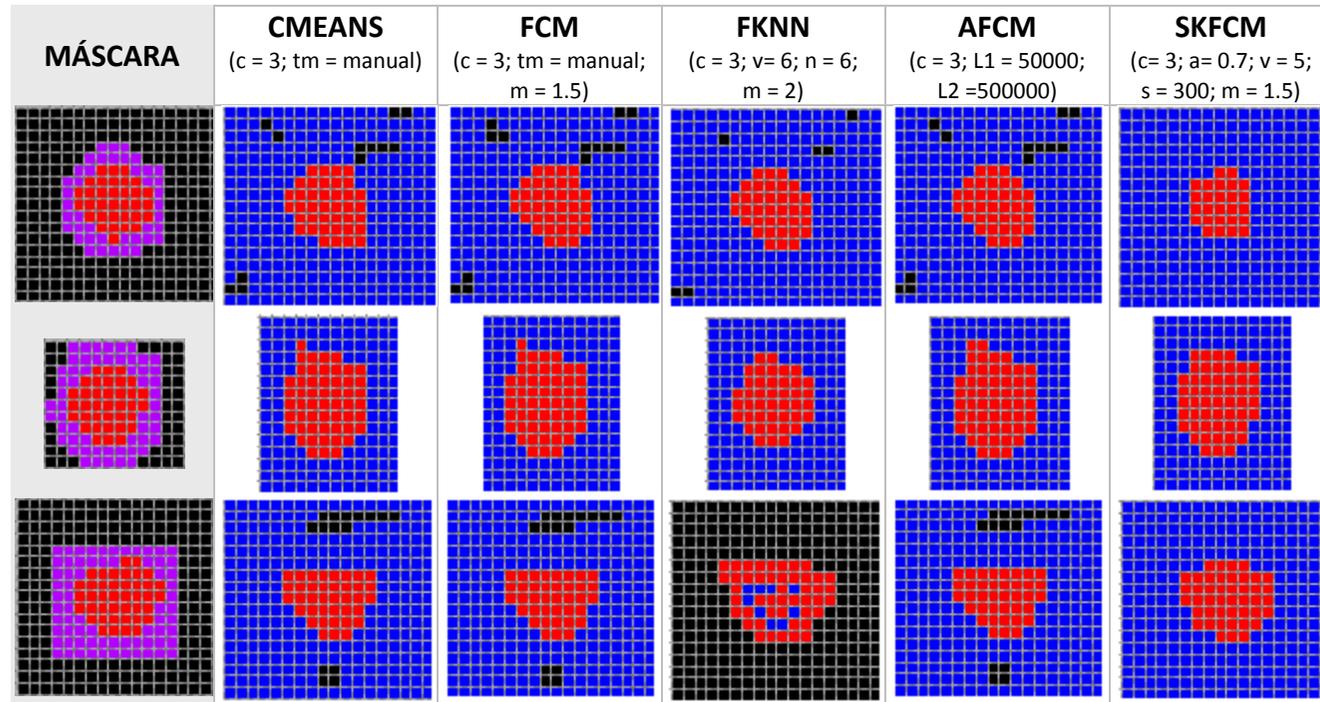


Figura 67. ROI segmentadas por algoritmo para algunas imágenes con NPS de tamaño inferior a 6 mm.

#### 4.1.4.2 Según forma de NPS

En este apartado se resumen los resultados de agrupar las imágenes diferenciando sólo entre aquellas con nódulo cuyo contorno tiene formas lobulares, y aquellas otras con NPS con espículas prolongándose o invadiendo el parénquima pulmonar.

	PARÁMETROS (Lobulados)	ValorTip	PARÁMETROS (Espiculados)	ValorTip
CMEANS	(c = 3; tm = manual)	0,72	(c = 3; tm = manual)	0,95
	(c = 3; tm = automática)	0,72	(c = 3; tm = automática)	0,95
	(c = 3; tm = manual1)	0,7	(c = 4; tm = manual)	0,15
FCM	(c = 3; tm = manual; m = 2)	0,73	(c = 3; tm = manual; m = 1.7)	0,89
	(c = 3; tm = manual; m = 1.5)	0,71	(c = 3 ; tm = manual; m = 1.3 )	0,78
	(c = 3 ; tm = manual; m = 1.3 )	0,7	(c = 3; tm = manual; m = 1.5)	0,74
FKNN	(c = 3; v= 6 ; n = 6 ; m = 2)	1,22	(c = 3 ; v= 5; n = 5 ; m = 2)	0,86
	(c = 3 ; v= 5; n = 5 ; m = 2)	1,13	(c = 3; v= 4; n = 4; m = 2)	0,76
	(c = 3 ; v=4 ; n = 4 ; m = 4 )	0,91	(c = 3; v= 6 ; n = 6 ; m = 2)	0,64
AFCM	(c = 3; L1 = 10000; L2 = 1000)	0,81	(c = 4; L1 = 10000; L2 = 10000)	0,90
	(c = 3; L1 = 10000; L2 = 10)	0,74	(c = 3; L1 = 50000; L2 = 300000)	0,66
	(c = 3; L1 = 100000; L2 = 10000)	0,54	(c = 3; L1 = 50000; L2 =500000 )	0,65
SKFCM	(c= 3; a= 0.7; v = 5; s = 300; m = 1.5)	1,41	(c= 3; a= 0.7; v = 5; s = 300; m = 1.5)	1,18
	(c=4 ; a = 0.7 ; v = 5; s = 150; m = 1.1)	1,14	(c = 3; a= 0.7; v = 5 ; s = 150 ; m = 1.5)	0,84
	(c = 3; a= 0.7; v = 10; s = 150 ; m = 50)	0,84	(c=4 ; a = 0.7 ; v = 5 ; s = 150; m = 1.1)	0,77

Tabla 37. Resumen de pruebas óptimas por algoritmo para imágenes agrupadas por forma de NPS.

##### 4.1.4.2.1 Lobulados

Es notorio que cuando el parámetro número de clases tiene valor 3 se alcanzan los mayores resultados tipificados en la media (superiores o iguales a 0.70), con todos los algoritmos empleados. Se obtiene mejor rendimiento con las pruebas que separan los píxeles de las imágenes en tres clústeres, pues la fragmentación de las mismas en cuatro clústeres lleva a realizar división de zonas dentro del propio nódulo y con ello disminuir la precisión en la clasificación de los píxeles de la zona de interés.

Con el Fuzzy C-Means los valores tipificados más altos son para las segmentaciones en 3 clústeres y un factor fuzzy 2; para esta tipología de nódulo el parámetro m con valores superiores a 2 ha generado peores resultados.

Con el Fuzzy K-Nearest Neighbor para las imágenes con NPS de márgenes lobulados tienen mayor valor tipificado las pruebas que emplean un factor fuzzy 2, y con seis vecinos mejor que con un número menor.

Los rendimientos que destacan, con el algoritmo Adaptive Fuzzy C-Means, obtenidos por las pruebas que han utilizado 3 clases, para segmentar esta tipología de imágenes emplean unos valores de L1 y L2 grandes, siendo la penalización de discontinuidades del campo multiplicador ( $L2 = 1000$ ) de un orden de magnitud inferior a la penalización de las grandes variaciones ( $L1 = 10000$ ).

Y con el algoritmo Spatially Kernelized Fuzzy C-Means, aunque se obtiene resultados buenos en la partición en cuatro clases, los tipificados más altos se corresponden con las pruebas en las que se han empleado 3 clases; destacando las pruebas en las que se ha empleado un termino de penalización  $\alpha = 0.7$  para permitir que el clúster nódulo sea más homogéneo, y utilizando una ventana de vecindario  $v = 5 \times 5$ , con el valor sigma de la función kernel gaussiana  $s = 300$  y factor fuzzy  $m = 1.5$ .

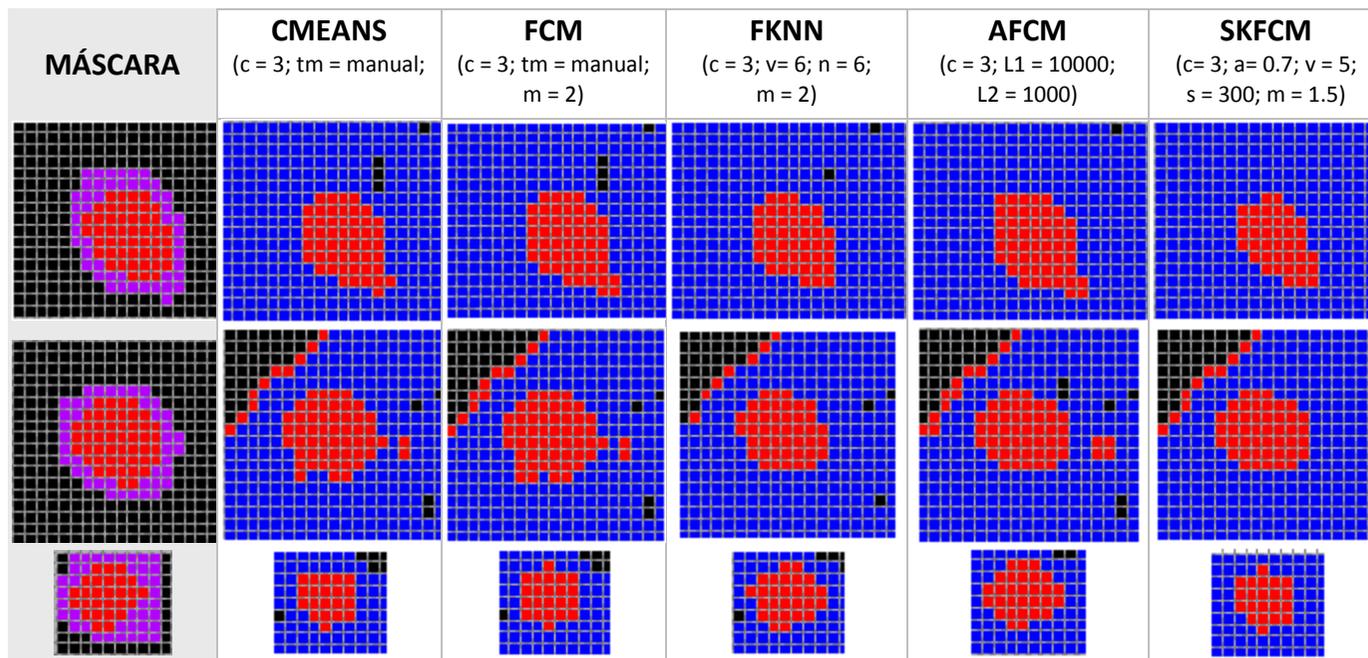


Figura 68. ROI segmentadas por algoritmo para algunas imágenes con NPS de bordes lobulados.

#### ***4.1.4.2.2 Espiculados***

Bajo este epígrafe se ha congregado el resumen de las pruebas seleccionadas dentro de las realizadas, y tipificado su rendimiento, con el subgrupo de imágenes cuyo nódulo presenta irradiaciones lineales desde su borde entremezclándose con el parénquima pulmonar.

En el método CMeans, los datos destacados corresponden a la utilización de 3 clústeres; para poder tener agrupados los píxeles correspondientes al NPS de cada imagen en uno sólo de ellos. Al comparar las CT que han sido segmentadas en 4 clases, se aprecia la pérdida de píxeles clasificados como verdaderos positivos.

Con el algoritmo Fuzzy C-Means lo mejor es segmentar en tres clústeres, y con el valor fuzzy 1.7, observándose menor número de falsos positivos que los encontrados en los resultados de emplear otros valores fuzzy.

También con el método Fuzzy K-Nearest Neighbor, conviene segmentar en 3 clases. Siendo con un valor de  $m = 2$  con los que se obtiene mejores prestaciones. Destaca la prueba en que el número de vecinos y muestras por vecino es 5.

Para el algoritmo Adaptive Fuzzy C-Means el valor tipificado más alto de media (0.90) se obtiene para el subgrupo de nódulos segmentando en 4 clústeres y penalizando por igual las grandes variaciones y las discontinuidades del campo multiplicador con valores de L1 y L2 de 10000.

Para este subgrupo de imágenes con nódulo de contorno espiculado, el Spatially Kernelized Fuzzy C-Means presenta los mejores resultados: segmentando en 3 clases con un factor fuzzy de 1.5, ventana de vecindario 5x5, alfa 0.7 y sigma 300, empeorando la segmentación cuando la anchura del kernel disminuye a valores por debajo de 300.

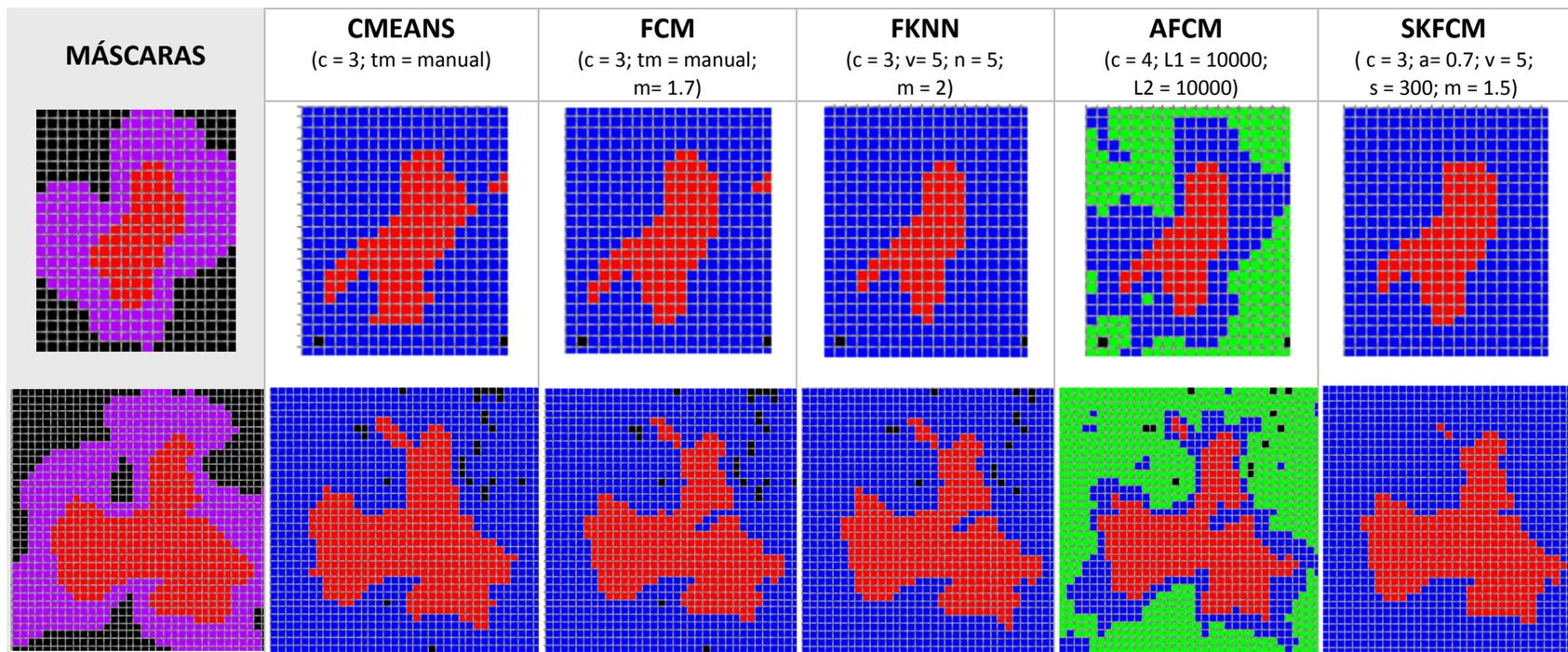


Figura 69. ROI segmentadas por algoritmo para algunas imágenes con NPS de forma espiculada.

#### 4.1.4.3 Pruebas óptimas según densidad de NPS

Según la atenuación de los rayos X, respecto al parénquima pulmonar, se tienen nódulos de alta densidad, igual densidad o baja densidad. La mayoría de los nódulos son de mayor densidad que el parénquima pulmonar. Se analizan, en esta clasificación, los nódulos más difíciles de segmentar por ser difusos e imprecisos: los nódulos cuya densidad es baja y denominados “en vidrio deslustrado”.

Con el algoritmo CMeans no se han logrado obtener resultados aceptables para esta tipología de NPS. Con los demás algoritmos empleados, si se han encontrado valores de parámetros que permitieron una clasificación buena en un único clúster.

Las imágenes que contienen un nódulo en vidrio deslustrado (menos denso e indefinido y con una distribución difusa) se podrían considerar “con ruido”. El resultado de éste hecho es que los algoritmos fuzzy permiten segmentar y aislar aceptablemente en un clúster, los nódulos con densidad en vidrio deslustrado porque la ventaja de estos algoritmos es que si un pixel está afectado por ruido, la segmentación se ve afectada en una cantidad fraccional, mientras que en el algoritmo CMeans la clasificación entera puede verse afectada.

La prueba del algoritmo Fuzzy C-Means que ha dado los valores tipificados más altos es, con diferencia, la que emplea un factor fuzzy  $m = 256$  y segmenta las imágenes en 15 clústeres. Observándose en estos resultados que aunque no muy buenos, sí son aceptables, pues existe un número elevado de verdaderos positivos, sin haber excesivos falsos positivos.

Con el algoritmo Fuzzy K-Nearest Neighbor se ha obtenido un buen rendimiento al precisar los píxeles pertenecientes al clúster nódulo, quizás debido al hecho de tener que utilizar como entrada al algoritmo al menos tantas muestras como vecinos a considerar en cada clase a segmentar. Siendo especialmente necesario utilizar un número de muestras superior al número de vecinos empleado, y a la vez menos vecinos por clúster dada la naturaleza difusa y poco definida de esta tipología de nódulo. La prueba con valores de parámetros ( $c = 8$ ;  $v = 3$ ;  $n = 8$ ;  $m = 2$ ), cuyos resultados destacan, utiliza 3 vecinos con 8 muestras por cada uno de los ocho clústeres segmentados.

Con el algoritmo Adaptive Fuzzy C-Means para una segmentación en 8 ó 10 clases es necesario que los términos penalizadores tomen valores muy bajos:  $L1=10$  y  $L2= 10$ . Así, con estos valores, dicho algoritmo procesa mejor las inhomogeneidades de las imágenes originadas por los píxeles de este tipo de NPS de aspecto poco definido y difuso, a la vez

que se realiza la segmentación de las CT preprocesadas. Por ésta razón las pruebas con estos valores de parámetros de entrada presentan los valores tipificados más elevados, dentro del conjunto test seleccionados.

Las pruebas de mayor rendimiento obtenidas empleando el método Spatially Kernelized Fuzzy C-Means que tienen los valores tipificados más altos corresponden a los test que han empleado un parámetro sigma de la función kernel gaussiana con un valor relativamente pequeño (50), y un valor alto de alfa (0.8) para penalizar el “ruido” en las imágenes, debido dicho ruido a la poca definición de los píxeles de este tipo de nódulo.

	PARÁMETROS (Deslustrados)	ValorTip	PARÁMETROS (subpleurales)	ValorTip
CMEANS	(c = 3; tm = manual)	0,73	(c = 4; tm = manual)	0,47
	(c = 3; tm = automática)	0,73	(c = 4; tm = automática)	0,47
	(c = 3; tm = manual1)	0,73	(c = 3; tm = manual1)	0,44
FCM	(c = 15; tm = manual; m = 256)	1,77	(c = 3; tm = manual; m =256)	0,74
	(c = 20; tm = manual; m = 512)	1,24	(c = 3; tm = manual; m = 1.7)	0,72
	(c = 30; tm = manual; m =258)	1,23	(c = 3; tm = manual; m = 50)	0,61
FKNN	(c = 8 ; v= 3 ; n = 8; m = 2 )	1,94	(c = 3; v= 4; n = 4; m = 2)	1,25
	(c = 7 ; v= 5 ; n = 9 ; m = 1.1 )	1,25	(c = 3 ; v= 4; n = 4 ; m = 1.5 )	1,14
	(c = 5; v= 4 ; n = 6; m = 1.1)	0,72	(c = 3 ; v= 5; n = 5 ; m = 2)	0,80
AFCM	(c = 8 ; L1 = 10 ; L2 = 10)	1,95	(c = 3 ; L1 = 10 ; L2 = 10)	0,85
	(c = 10 ; L1 = 10 ; L2 = 10)	1,93	(c = 4; L1 = 100000; L2 = 10000)	0,76
	(c = 3; L1 = 10000 ; L2 = 10)	0,32	(c = 4; L1 = 10000 ; L2 = 10000)	0,64
SKFCM	(c = 9; a = 0.8; v = 5 ; s = 50; m = 2)	1,69	(c = 3; a = 0.7; v = 10; s = 150; m = 50)	1,20
	(c = 6; a = 0.8; v = 5; s = 50 ; m = 2)	1,09	(c = 3; a = 0.8 ; v = 5; s = 50 ; m = 2)	0,99
	(c = 5; a = 0.1; v = 10 ; s = 150 ; m = 4)	0,73	(c = 3 ; a = 0.7; v = 5 ; s = 150; m = 50)	0,94

Tabla 38. Resumen de pruebas óptimas por algoritmo para imágenes agrupadas por densidad y localización.

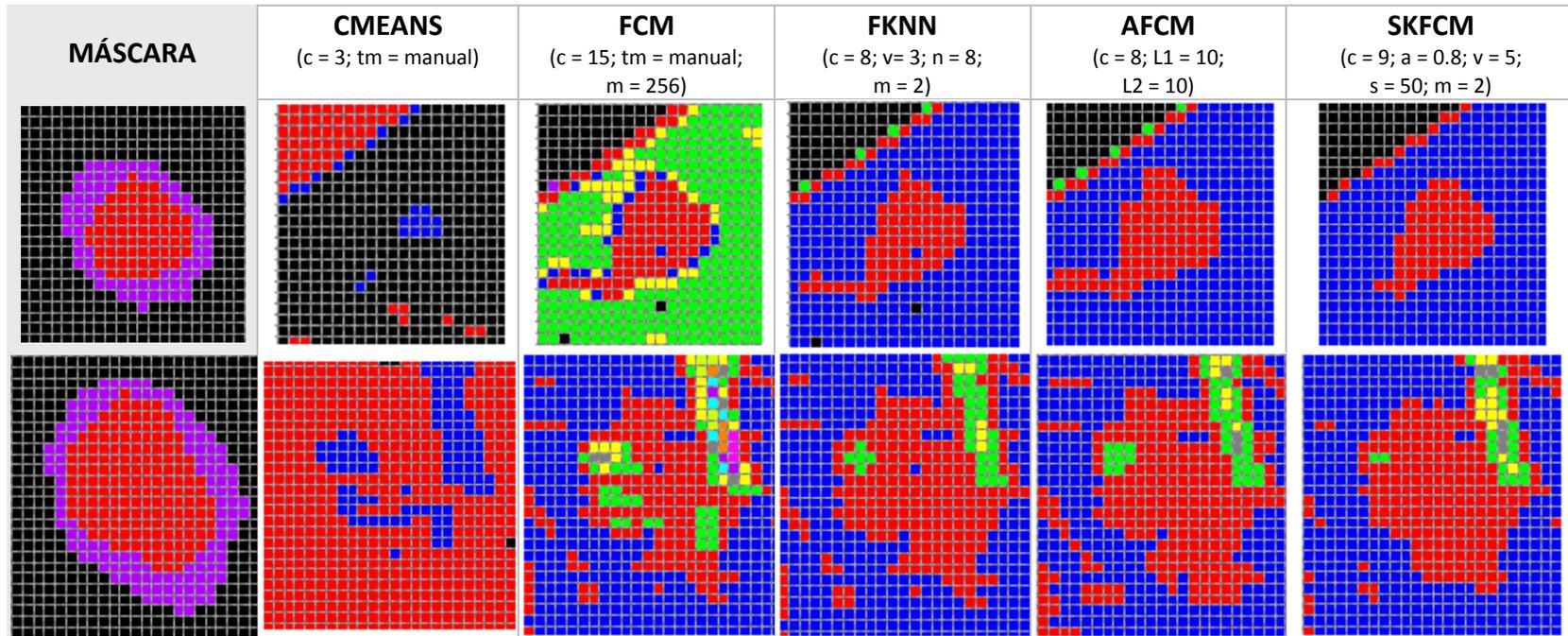


Figura 70. ROI segmentadas por algoritmo para algunas imágenes con NPS de densidad en vidrio deslustrado

#### 4.1.4.4 Pruebas óptimas según localización

Los nódulos pulmonares están ubicados con mayor frecuencia en los lóbulos superiores y concretamente en el lóbulo superior derecho; aquí sólo se hace referencia a los nódulos subpleurales, por la dificultad de su procesado al estar solapados en parte de la pleura.

Los valores tipificados más altos con el método CMeans corresponden a la segmentación que ha utilizado 4 clases. En las pruebas los resultados son aceptables, aunque tienen un elevado número de píxeles clasificados como clúster-nódulo y que pertenecen al entorno del mismo; pero al considerar la hipótesis de que dicho entorno no es tejido cancerígeno se consideran falsos positivos. Este número de falsos positivos en el entorno es más elevado cuando se utiliza como parámetro inicial 3 clústeres.

En el caso del algoritmo Fuzzy CMeans, la mejor segmentación para los nódulos subpleurales se consigue segmentando en tres clases y utilizando un factor fuzzy 256. En las ROI segmentadas con dicho valor, se ha logrado reducir más los falsos positivos en el entorno que en las demás pruebas cuyos valores tipificados también son altos.

Los resultados del algoritmo Fuzzy K-Nearest Neighbor indican que con las pruebas mejores se han empleado 3 clústeres. La combinación de un factor fuzzy 2, con la consideración de 4 vecinos por clúster y 4 muestras por cada uno de los clústeres a segmentar es la que permite que aumenten los píxeles clasificados como TP, a pesar de que en alguna imagen se incrementa ligeramente el número de falsos positivos en el entorno del nódulo. Además para el mismo factor fuzzy, y mayor número de vecinos y muestras los resultados empeoran.

Es el algoritmo AFCM con coeficientes de penalización  $L1 = 100000$  y  $L2 = 10000$ , junto con el CMeans, con los que segmentando en cuatro clases se obtiene una clasificación aceptable. En dichas segmentaciones los píxeles correspondientes al clúster nódulo quedan agrupados y delimitados por píxeles pertenecientes a otro clúster que los separa del parénquima pulmonar. Este cuarto clúster no resta verdaderos positivos al clúster nódulo

Finalmente, en los mejores resultados obtenidos con el algoritmo Spatially Kernelized Fuzzy C-Means, se aprecia el predominio de 3 clases como la clusterización con mejor rendimiento; consiguiendo la mejor media con una ventana de vecindario  $10 \times 10$ , un valor de parámetro Sigma de la función de kernel gaussiano 150, y controlando la penalización un valor de Alfa 0.7 que producen unos valores tipificados de media 1.20 si el factor fuzzy es 50.

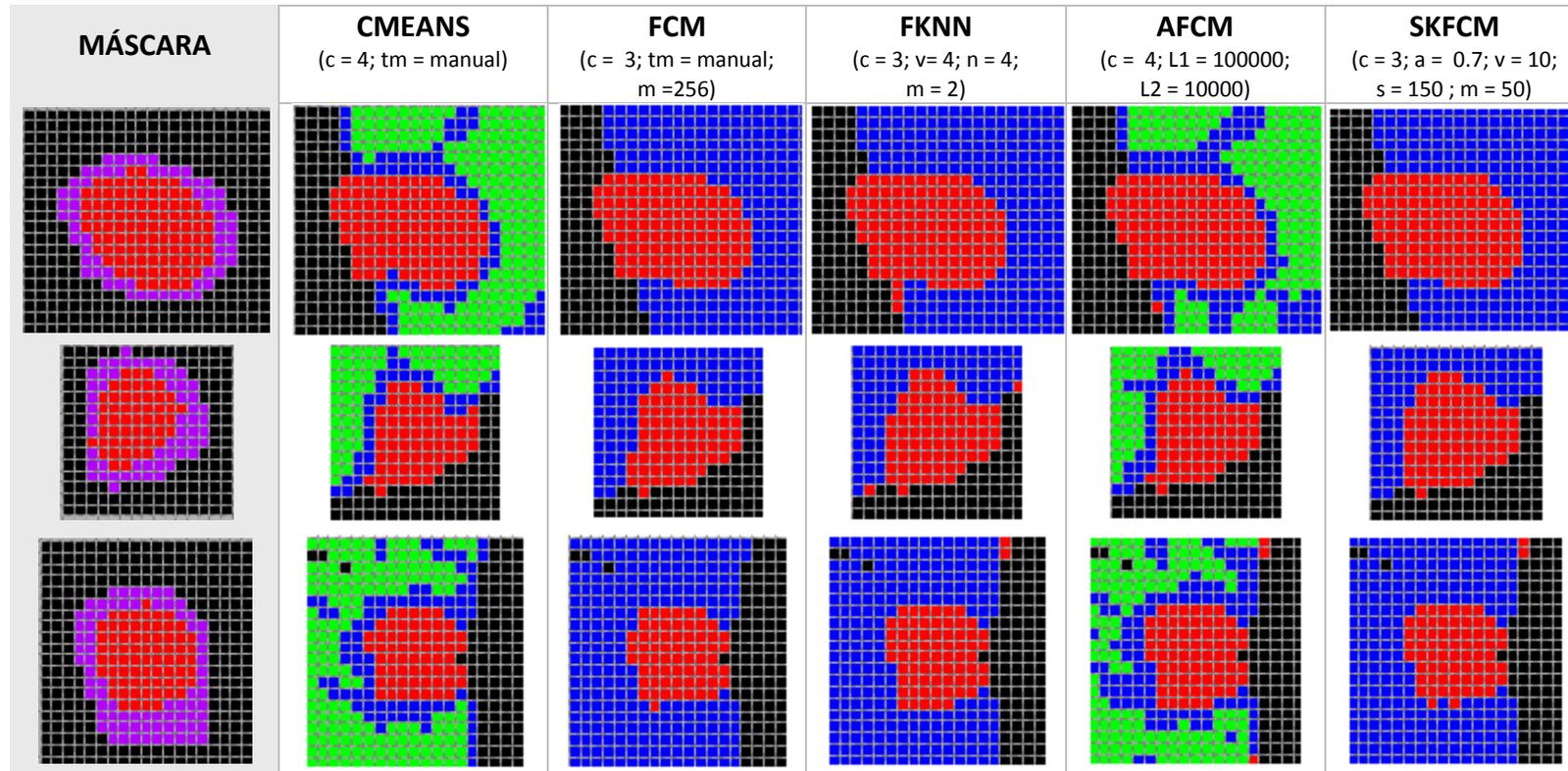


Figura 71. ROI segmentadas por algoritmo para algunas imágenes con NPS de localización subpleural.

#### 4.1.5 Comparación del rendimiento de pruebas óptimas entre algoritmos

A partir de la prueba que produce los mejores resultados en cada algoritmo, con cada agrupación de imágenes según su tipología, se busca cual es más adecuado para segmentar dicha tipología de nódulos. Por ello se calcula el coeficiente de variación; y se realiza el análisis comparativo del rendimiento obtenido por las pruebas de test óptimas entre los diferentes algoritmos de segmentación, y así, determinar en qué casos es conveniente utilizar cada una de las técnicas consideradas, por ser con la que mayor media y menos dispersión se origina en los resultados. Se elige aquella prueba que produzca menos variabilidad respecto a su media, es decir, la que produzca unos resultados más homogéneos en todo el subgrupo de imágenes.

Así pues, se determina objetivamente la mayor o menor homogeneidad de los resultados producidos por los distintos algoritmos (sobre las imágenes agrupadas por tipología de nódulo), calculando el cociente entre la desviación típica y la media del índice de rendimiento de dichas pruebas óptimas. Es conveniente que para el cálculo de coeficiente de variación se utilice un índice con valores no negativos. Por ello, se emplea el índice Azope, y así asegurar que la media siempre es positiva, evitando problemas con dicho coeficiente.

Se propondrá, como algoritmo más conveniente, aquel cuya prueba óptima produzca el coeficiente de variación más bajo, es decir, que la relación entre su desviación típica y su media sea la menor; obteniendo así el de menor dispersión y mayor media en todo el subgrupo de imágenes considerado.

##### **4.1.5.1 Por tamaño de nódulo**

A nivel visual son más fáciles de identificar los nódulos de mayor tamaño. Con todos los algoritmos utilizados, a excepción del Adaptive Fuzzy CMeans, los mejores resultados para los nódulos mayores de un centímetro se han conseguido segmentando las imágenes en tres clústeres. Los valores más bajos de coeficiente de variación se obtienen con el método Fuzzy C-Means (0.15), en el cual se ha utilizado un valor fuzzy 256; un coeficiente de variación próximo a éste corresponde al algoritmo Spatially Kernelized Fuzzy C-Means para también un valor de factor fuzzy alto ( $m = 50$ ). El algoritmo CMeans es el que presenta mayor dispersión (0.28) para éste sub-grupo de imágenes, como se indica en la Tabla 39.

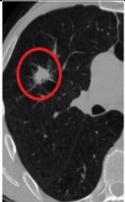
Diámetro $\geq$ 10 mm	Algoritmo	Parámetros	CV
	<b>CMeans</b>	(c = 3; tm = certeza)	0,28
	<b>FCM</b>	(c = 3; tm = certeza; m =256)	0,15
	<b>FKNN</b>	(c = 3; v = 6 ; n = 6 ; m = 1.1 )	0,20
	<b>AFCM</b>	(c = 4; L1 = 10000; L2 = 10000)	0,25
	<b>SKFCM</b>	(c = 3 ; a = 0.7; v = 5 ; s = 150; m = 50)	0,18

Tabla 39. Comparativa final de algoritmos para imágenes con nódulos mayores de 1cm.

Desde el punto de vista del diagnóstico médico existe un gran interés en encontrar y delimitar aquellos nódulos de pequeño tamaño que puedan sugerir, con cierto grado de certeza, la existencia de carcinoma. Para los nódulos menores de un centímetro y mayores a medio centímetro, los mejores valores de coeficiente de variación se obtienen utilizando el Fuzzy C-Means con un valor fuzzy alto ( $m = 256$ ); teniendo éste método casi el triple de mejor relación entre la dispersión y la media que si se emplea en la segmentación el método CMeans, y teniendo casi el doble de mejor la relación entre la dispersión y la media que los restantes algoritmos (véase Tabla 40).

$6\text{mm} \leq$ Diámet $<$ 10 mm	Algoritmo	Parámetros	CV
	<b>CMeans</b>	(c = 3; tm = certeza)	0,29
	<b>FCM</b>	(c = 3; tm = certeza; m =256)	0,11
	<b>FKNN</b>	(c = 3; v = 4; n = 4; m = 2)	0,24
	<b>AFCM</b>	(c = 3; L1 = 10000 ; L2 = 10)	0,24
	<b>SKFCM</b>	(c = 3; a = 0.7; v = 5; s = 300; m = 1.5)	0,25

Tabla 40 Comparativa final de algoritmos para imágenes con nódulos cuyo diámetro es entre 6 y 10 mm.

El término de “cáncer precoz” se refiere a aquellos cánceres incipientes cuyo tamaño es muy pequeño, menor o igual a medio centímetro, y que generalmente tiene buen pronóstico. Aunque la detección precoz no siempre garantice una evolución favorable, porque hay algunos tumores de muy pequeño tamaño que conducen a una rápida metástasis. Se utiliza como indicador de riesgo de malignidad el tiempo de doblaje del diámetro del nódulo, y los TAC multicorte permiten evaluar el tamaño de los nódulos y precisar mejor el crecimiento del mismo [Castañer, 2011] [Yeong, 2007]. Por lo cual, es importante precisar en el segmentado los máximos TP y mínimos FP correspondientes a los puntos de la imagen en la ROI.

En este caso la relación entre dispersión y media, sobre las segmentaciones resultantes son similares con todos los métodos empleados, como se resumen en la Tabla 41; a nivel numérico el menor valor se consigue con el algoritmo de referencia CMeans. Siendo los algoritmos que consideran en sus parámetros al vecindario, los que introducen algo más de variabilidad.

Diámetro < 6 mm	Algoritmo	Parámetros	CV
	<b>CMeans</b>	(c = 3; tm = certeza)	0,10
	<b>FCM</b>	(c = 3; tm = certeza; m = 1.5)	0,11
	<b>FKNN</b>	(c = 3; v= 6 ; n = 6 ; m = 2)	0,13
	<b>AFCM</b>	(c = 3; L1 = 50000; L2 = 500000 )	0,11
	<b>SKFCM</b>	(c= 3; a= 0.7; v = 5; s = 300; m = 1.5)	0,13

Tabla 41. Comparativa final de algoritmos para imágenes con nódulos de diámetro inferior a 6 mm

#### 4.1.5.2 Por localización del nódulo

Los nódulos subpleurales se encuentran situados en contacto con la pleura en los bordes del pulmón, son difíciles de precisar porque en el pre-procesado al intentar eliminar la pleura -ya que ésta no pertenece al parénquima pulmonar- se pierden algunos píxeles correspondientes a éste nódulo. En la detección de los píxeles de estos nódulos el algoritmo que presentó los mejores resultados con una menor variabilidad ha sido el Spatially Kernelized Fuzzy C-Means. Dicha comparativa, con las correspondientes combinaciones de valores de parámetros óptimos, se muestra en la Tabla 42.

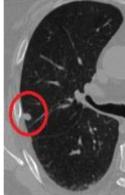
SUBPLEURALES:	Algoritmo	Parámetros	CV
	<b>CMeans</b>	(c = 4; tm = certeza)	0,19
	<b>FCM</b>	(c = 3; tm = certeza; m =258)	0,18
	<b>FKNN</b>	(c = 3; v= 4; n = 4; m = 2)	0,16
	<b>AFCM</b>	(c = 4; L1 = 100000; L2 = 10000)	0,19
	<b>SKFCM</b>	(c = 3; a= 0.7; v = 10; s = 150 ; m = 50)	0,15

Tabla 42 Comparativa final de algoritmos para imágenes con nódulos de localización subpleural.

### 4.1.5.3 Por densidad del nódulo

Cuando se clasifica un NPS según la densidad se está haciendo referencia a la atenuación de los rayos X respecto al parénquima pulmonar. Atendiendo a ésta propiedad se distinguen nódulos sólidos o semisólidos, y en vidrio deslustrado. Siendo los nódulos en vidrio deslustrado con contenido no sólido, y por tanto menos densos e indefinidos, con distribución difusa. Por ello es más difícil segmentarlos dado que los píxeles que lo representan están entremezclados con los píxeles de otras zonas.

Se ha conseguido segmentarlos, con un rendimiento bastante bueno; pero utilizando unos valores de parámetros que difieren de forma notable respecto a los empleados con los demás NPS sólidos o semisólidos. Con el algoritmo de referencia CMeans no se ha logrado ningún resultado aceptable para este tipo de nódulos. Destacan los resultados conseguidos con los algoritmos Spatially Kernelized Fuzzy C-Means y Fuzzy K-Nearest Neighbor, en los que se han utilizado 9 y 8 clústeres respectivamente, y un valor de factor fuzzy 2; ambos métodos permiten considerar al vecindario como parámetro de entrada. En la Tabla 43, se aprecia como el Spatially Kernelized Fuzzy C-Means produce el mejor rendimiento con menor dispersión y mayor media en los resultados.

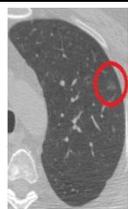
VIDRIODESLUSTRADO	Algoritmo	Parámetros	CV
	CMeans	(c = 3; tm = certeza)	0,16
	FCM	(c = 15; tm = certeza; m = 256)	0,09
	FKNN	(c = 8 ; v= 3 ; n = 8; m = 2 )	0,05
	AFCM	(c = 8 ; L1 = 10 ; L2 = 10)	0,10
	SKFCM	(c = 9; a = 0.8; v = 5 ; s = 50; m = 2)	0,04

Tabla 43. Comparativa final de algoritmos para imágenes con nódulos cuya densidad es en vidrio deslustrado.

### 4.1.5.4 Por forma del nódulo

Según la clasificación por forma, los nódulos pueden ser redondos, lobulares e irregulares. Estas formas se ven modificadas por los márgenes (véase en la Figura 6 los patrones de calcificación y contorno), que si están bien definidos permiten ver una brusca transición en la intensidad con respecto al tejido que le rodea. Los nódulos lobulares tienen pequeñas ondulaciones en el contorno del nódulo. Los resultados finales de segmentación para este subgrupo de imágenes se presentan en la Tabla 44. Siendo el Fuzzy C-Means, Adaptive Fuzzy C-Means y Spatially Kernelized Fuzzy C-Means los algoritmos de mejor

comportamiento, referido éste al menor valor de coeficiente de variación (0.06) producido en la segmentación del subgrupo de lobulados, y habiendo segmentado para los métodos Fuzzy C-Means y Adaptive Fuzzy C-Means con un valor fuzzy 2.

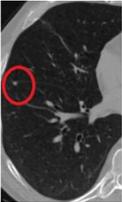
LOBULADOS	Algoritmo	Parámetros	CV
	<b>CMeans</b>	(c = 3; tm = certeza)	0,08
	<b>FCM</b>	(c = 3; tm = certeza; m = 2)	0,06
	<b>FKNN</b>	(c = 3; v = 6 ; n = 6 ; m = 2)	0,09
	<b>AFCM</b>	(c = 3; L1 = 10000 ; L2 = 1000)	0,06
	<b>SKFCM</b>	(c = 3; a = 0.7; v = 5; s = 300; m = 1.5)	0,06

Tabla 44 Comparativa final de algoritmos para imágenes con nódulos con márgenes lobulares.

Aquellas imágenes cuyo NPS presenta líneas rectas que se irradian desde el nódulo, es decir márgenes espiculados que a veces son difíciles de visualizar cuando el tamaño de las líneas es muy pequeño, confundándose con el tejido circundante, en general para su detección es mejor utilizar el algoritmo Adaptive Fuzzy C-Means. Tal como se indica en la Tabla 45, dicho método presenta, para éste sub-grupo de imágenes, un coeficiente de variación de 0.06 en contraste con los algoritmos Fuzzy K-Nearest Neighbor y Spatially Kernelized Fuzzy C-Means cuyos resultados tienen más del doble de la relación entre dispersión y media con éstas imágenes.

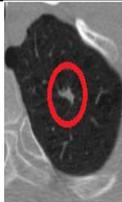
ESPICULADOS	Algoritmo	Parámetros	CV
	<b>CMeans</b>	(c = 3; tm = certeza)	0,09
	<b>FCM</b>	(c = 3; tm = certeza; m = 1.7)	0,07
	<b>FKNN</b>	(c = 3 ; v = 5; n = 5 ; m = 2)	0,16
	<b>AFCM</b>	(c = 4; L1 = 10000; L2 = 10000)	0,06
	<b>SKFCM</b>	(c = 3; a = 0.7; v = 5; s = 300; m = 1.5)	0,13

Tabla 45. Comparativa final de algoritmos para imágenes con nódulos con márgenes espiculados.

#### 4.2 Validación de valores paramétricos

De cada uno de los 23 estudios de cortes tomográficos, descritos en el apartado 3.2, se ha tomado una imagen distinta, y próxima al corte tomográfico al utilizado en las pruebas de análisis de valores paramétricos. Disponiendo así, para la valoración de los valores de parámetros óptimos, de 23 Nódulos Pulmonares Solitarios contenidos en los cortes de

CT de Tórax, de los cuales el 74% corresponden a nódulos con diámetro inferior a 10mm (39% menores de 6mm y 35% entre 6 y 10mm) y el 26% son mayores de un centímetro, un 13% de nódulos sub-pleurales, un 30% de imágenes con nódulos espiculados y un 26% de CT con NPS lobulados y un 13% con nódulos de densidad en vidrio deslustrado. Se han aplicado específicamente los algoritmos objeto de estudio sobre las imágenes que forman el grupo de validación, segmentando con los valores de parámetros de entrada indicados en la Tabla 46. Estos valores de parámetros se consideran los más adecuados, por ser con ellos con los que se han logrado mejores rendimientos en los resultados sobre el primer grupo de imágenes para las pruebas test.

	<b>CMEANS</b>	<b>FCM</b>	<b>FKNN</b>	<b>AFCM</b>	<b>SKFCM</b>
<b>General</b>	(c = 3; tm = manual)	(c = 3; tm = manual; m = 1.7)	(c = 3; v= 4; n = 4; m = 2)	(c = 3; L1 = 50000; L2 = 500000)	(c= 3; a= 0.7; v = 5; s = 300; m = 1.5)
<b>GRANDES</b>	(c = 3; tm = manual)	(c = 3; tm = manual; m =256)	(c = 3; v= 6; n = 6; m =1.1)	(c = 4; L1 = 10000; L2 = 10000)	(c =3 ; a = 0.7; v = 5; s = 150; m = 50)
<b>PEQUEÑOS</b>	(c = 3; tm = manual)	(c = 3; tm = manual; m = 1.5)	(c = 3; v= 6; n = 6 ; m = 2)	(c = 3; L1 = 50000; L2 =500000 )	(c= 3; a= 0.7; v = 5; s = 300; m = 1.5)
<b>MEDIANOS</b>	(c = 3; tm = manual)	(c = 3; tm = manual; m =256)	(c = 3; v= 4; n = 4; m = 2)	(c = 3; L1 = 10000 ; L2 = 10)	(c= 3; a= 0.7; v = 5; s = 300; m = 1.5)
<b>DESLUSTRADOS</b>	(c = 3; tm = manual)	(c =15; tm = manual; m = 256)	(c = 8 ; v= 3; n = 8; m = 2)	(c = 8; L1 = 10; L2 = 10)	(c = 9; a = 0.8; v = 5; s = 50; m = 2)
<b>SUBPLEURALES</b>	(c = 4; tm = manual)	(c = 3; tm = manual; m =256)	(c = 3; v= 4; n = 4; m = 2)	(c = 4; L1 = 100000; L2 =10000)	(c = 3; a= 0.7; v = 10; s = 150; m = 50)
<b>LOBULARES</b>	(c = 3; tm = manual)	(c = 3; tm = manual; m = 2)	(c = 3; v= 6; n = 6 ; m = 2)	(c = 3; L1 = 10000; L2 = 1000)	(c= 3; a= 0.7; v = 5; s = 300; m = 1.5)
<b>ESPICULADOS</b>	(c = 3; tm = manual)	(c = 3; tm = manual; m = 1.7)	(c = 3 ; v= 5; n = 5 ; m = 2)	(c = 4; L1 = 10000; L2 = 10000)	(c= 3; a= 0.7; v = 5; s = 300; m = 1.5)

**Tabla 46. Valores paramétricos de entrada óptimos para segmentar con cada algoritmo el grupo de imágenes general y los subgrupos según la tipología de nódulos contenida.**

Con este conjunto de valores de parámetros seleccionados en la primera parte del experimento, se han ejecutado los algoritmos de segmentación estudiados utilizando el grupo de imágenes para validación que previamente se han pre-segmentado; y haciendo uso de técnicas empíricas de discrepancia se han calculado los aciertos y fallos, tomando como referencia las máscaras realizadas por los expertos de este segundo grupo de imágenes.

Se muestra a continuación, en primer lugar una comparativa de forma gráfica y tabular de los resultados obtenidos en la validación de los valores paramétricos empleados por los algoritmos, según el arquetipo de nódulo contenido en las imágenes; y posteriormente se finaliza mostrando, en el espacio ROC, la sensibilidad y el complementario de la especificidad, conseguida al segmentar con cada uno de los algoritmos el grupo de

imágenes CT de validación , empleando para ello los valores de parámetros seleccionados como óptimos.

#### 4.2.1 Aciertos y fallos por algoritmo y tipología de NPS

Para cada algoritmo se muestran comparativamente la media de la fracción de verdaderos positivos (TP), media de la fracción de falsos positivos en el entorno del nódulo (FPe), y la media de la fracción de falsos positivos en el resto de la imagen (FPr).

La Figura 72 muestra gráficamente la comparación de los resultados entre los algoritmos empleados al segmentar las imágenes con nódulos de diámetro superior a los 10 mm.

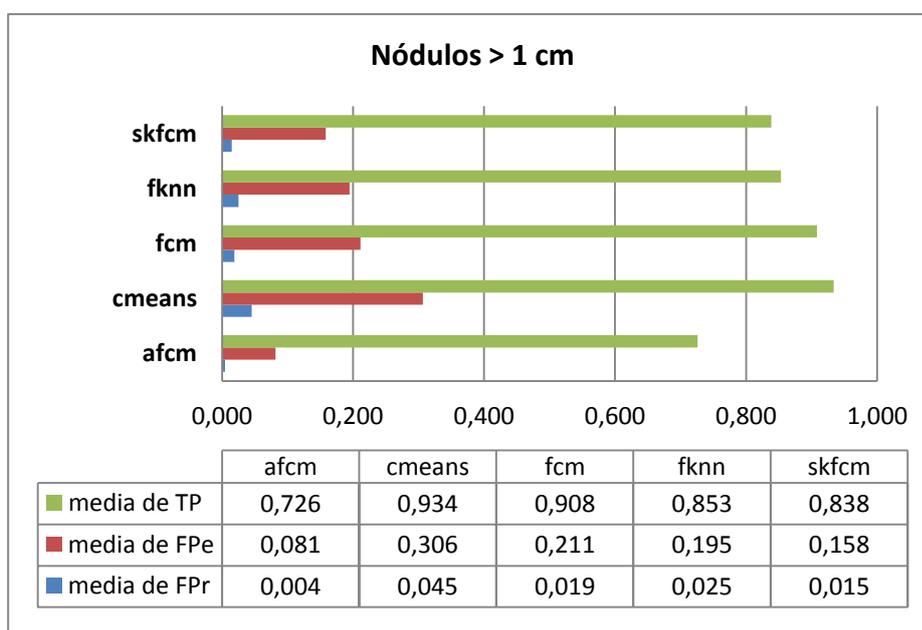
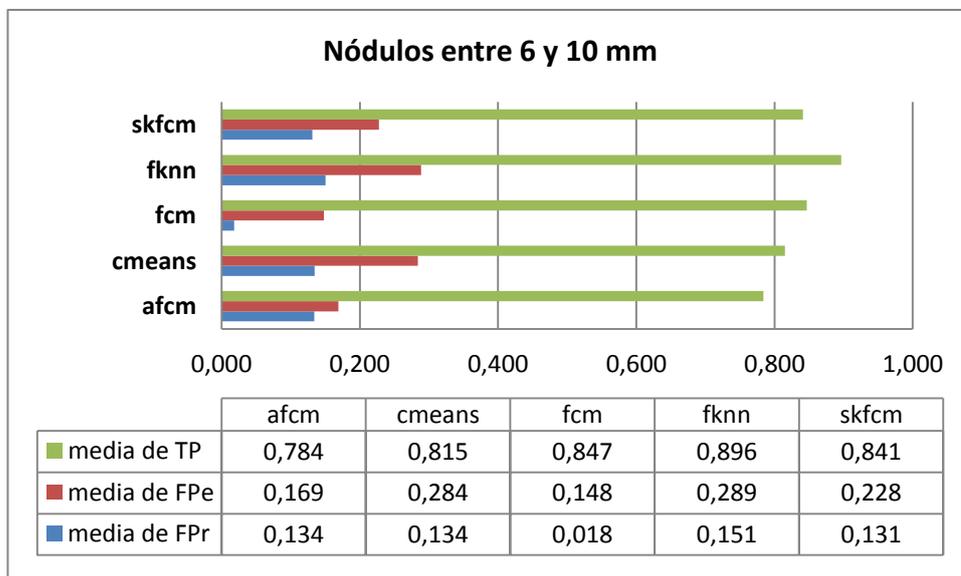


Figura 72. Comparación de aciertos y fallos para subgrupo de imágenes con NPS mayores de 1 cm.

El algoritmo de referencia CMeans es el que presenta un criterio menos estricto al clasificar los píxeles de este tipo de nódulo, pues siendo elevado el porcentaje de aciertos (93,4%) presenta un elevadísimo porcentaje de fallos en el entorno de estos nódulos (30,6%); y además en el resto de la imagen es el que más píxeles clasifica, un 4.51%, como falsos positivos. En el otro extremo está el algoritmo Adaptive Fuzzy C-Means con el menor porcentaje de falsos positivos, produce un 8.15% en el entorno del nódulo y tan sólo un 0.44% en el resto de la imagen pero es el que alcanza un valor de aciertos inferior (72.60%). Se perfila como más adecuado, el algoritmo Fuzzy CMeans pues alcanza aciertos del 90.84%, un 21.13% y 1.86% de falsos positivos en el entorno y resto de la

imagen respectivamente. Los algoritmos Fuzzy K-Nearest Neighbor y Spatially Constrained Kernelized (con un porcentaje de aciertos del 85.31% y Fuzzy C-Means (83.83% de media) presentan una valoración intermedia.

Para las imágenes con nódulos de tamaño intermedio, cuyo diámetro oscila entre 6 y 10 mm, se han obtenido los resultados de aciertos y fallos medios especificados en la Figura 73, en la cual se pueden comparar gráficamente dichas fracciones de aciertos y fallos. Los algoritmos que producen mayor porcentaje de falsos positivos en el entorno del nódulo son el CMeans (28.41%) y el Fuzzy K-Nearest Neighbor (28.85%); siendo también los métodos con mayor número en porcentaje de falsos positivos en el resto de la imagen: el CMeans con un 13.44% y el Fuzzy K-Nearest Neighbor con un 15,06%. Con el algoritmo Adaptive Fuzzy C-Means se tiene el menor porcentaje de aciertos (78.40%), a la vez que es de los de menor porcentaje de fallos en el entorno del nódulo (13.91%). Al igual que para las imágenes con nódulos mayores de 1 centímetro el método Fuzzy CMeans, es el que teniendo un buen porcentaje de aciertos (84.69%), permite obtener un número más bajo en el porcentaje de falsos positivos tanto en el entorno (14.82%) como en el resto (1.82%). Es pues, el algoritmo Fuzzy CMeans el que tiene mejor prestación al permitir obtener más aciertos con menos fallos.



**Figura 73. Comparación de aciertos y fallos para subgrupo de imágenes con NPS cuyo diámetro está entre 6 y 10 mm.**

Los resultados obtenidos con los algoritmos de segmentación para las imágenes con nódulos de diámetro inferior o igual a medio centímetro están resumidos en la Figura 74. El porcentaje más alto de verdaderos positivos se obtiene con el algoritmo Adaptive

Fuzzy C-Means (media 92.06%), el porcentaje menor de aciertos (85.42) se obtiene con el Fuzzy K-Nearest Neighbor; el cual tiene el menor porcentaje de falsos positivos en el entorno (10.75%). Los métodos CMeans y Fuzzy CMeans ofrecen resultados similares en el porcentaje de verdaderos positivos (media 91.43); exhibiendo un porcentaje ligeramente mayor de falsos positivos el CMeans, tanto en el entorno del nódulo como en el resto. El algoritmo Spatially Kernelized Fuzzy C-Means, no siendo el de menor porcentaje de aciertos, muestra un menor resultado que los demás en la media (87.26%), un 13.38% de falsos positivos en el entorno y un 1.42% de fallos en el resto.

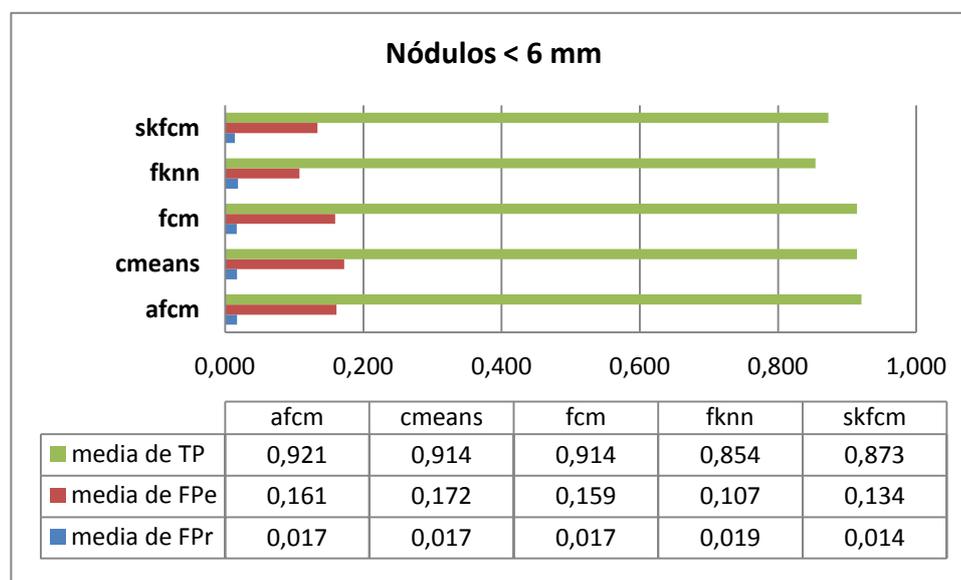


Figura 74. Comparación de aciertos y fallos para subgrupo de imágenes con NPS de diámetro inferior a 6 mm.

Se han agrupado las imágenes de validación que contenían un nódulo solitario localizado en la pleura periférica de los pulmones y se las ha denominado subgrupo de imágenes con NPS subpleural. En la Figura 75 se resumen los resultados derivados de la aplicación de los algoritmos objeto de esta tesis sobre el subgrupo de imágenes con nódulo subpleural. Los mejores porcentajes de aciertos corresponden a los métodos Spatially Kernelized Fuzzy C-Means y Fuzzy K-Nearest Neighbor con una media de 95.65% y 95.29% respectivamente. El algoritmo con menor valor en el porcentaje de verdaderos positivos es el CMeans (media 83.09%). Los resultados de aciertos con los algoritmos Fuzzy C-Means y Adaptive Fuzzy C-Means están en una zona intermedia con unos valores de media de 89.45% y de 90.37% respectivamente. Los falsos positivos en el entorno son bastante elevados, oscilando su porcentaje entre los 24.62 % para el CMeans y los 27.00% para el Spatially Kernelized Fuzzy C-Means.

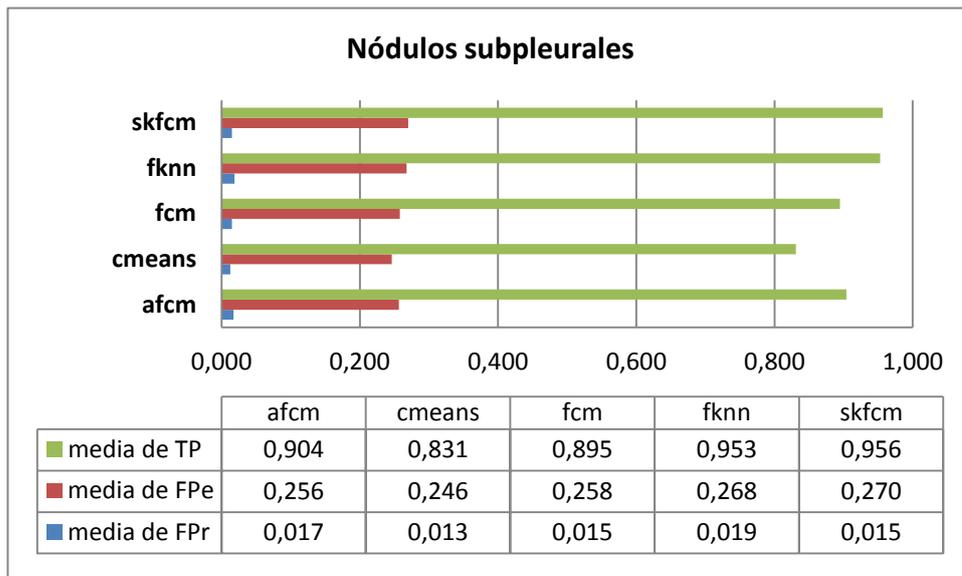


Figura 75. Comparación de aciertos y fallos para subgrupo de imágenes con NPS cuya localización es subpleural.

La segmentación del subconjunto de imágenes de validación que tienen un NPS en vidrio deslustrado permitió obtener los resultados mostrados en la Figura 76.

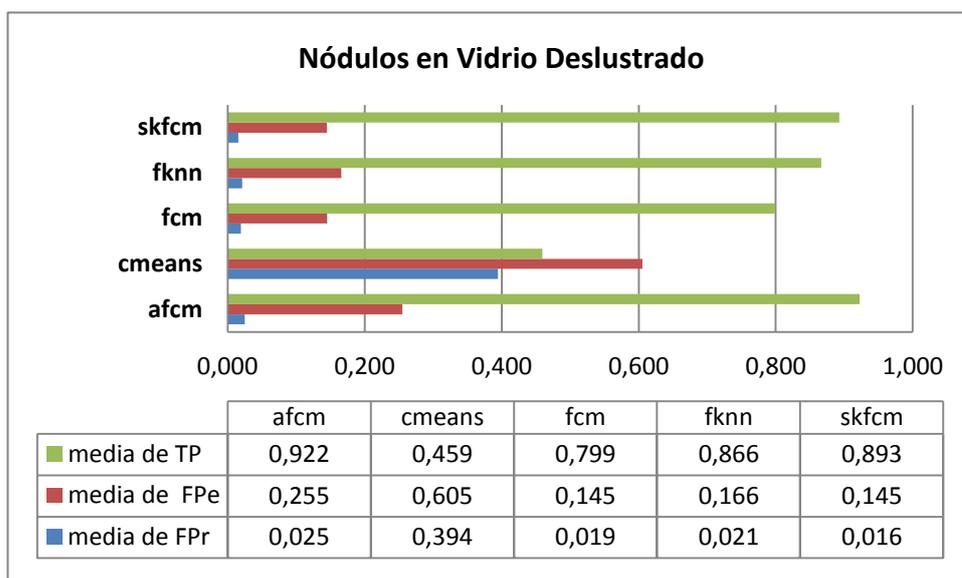


Figura 76. Comparación de aciertos y fallos para subgrupo de imágenes con NPS en vidrio deslustrado.

Claramente se ve que el algoritmo CMeans no presenta unos resultados aceptables para las imágenes con esta tipología de nódulo, pues la media de verdaderos positivos (45.93%) es muy baja, siendo la media de los falsos positivos en el entorno del nódulo (60.53%), y en el resto de la imagen (39.43%) muy malos. El mayor porcentaje de aciertos se consigue

con el Adaptive Fuzzy C-Means, media de 92.25%, siendo la media de porcentaje de fallos poco más que aceptable (25.50%) en el entorno y en el resto (2.51%). Con el algoritmo Spatially Kernelized Fuzzy C-Means se logra la mejor relación entre aciertos (89.26%) y fallos en el entorno (14.46%), y en el resto (media 1.59%). El Fuzzy K-Nearest Neighbor y el Fuzzy C-Means produce buenos resultados: un 86.63% de media de verdaderos positivos para el primero y un 79.93% para el segundo; siendo los fallos de estos un 16.59% y un 14.54% en el entorno, respectivamente.

A la vista de los resultados, al igual que sucedía con el grupo de imágenes test, los algoritmos con parámetros que consideran al vecindario de los píxeles a clasificar son los que realizan las mejores segmentaciones de las imágenes para este clúster nódulo. Además, también resulta evidente que el algoritmo de referencia CMeans no es apto para la segmentación de imágenes en las que sea necesario precisar los píxeles pertenecientes a un NPS con densidad en vidrio deslustrado.

La Figura 77 muestra los resultados de las medias de verdaderos positivos y falsos positivos, recogidos de los resultados de la segmentación con las imágenes que contenían un nódulo con bordes lobulados. Los algoritmos Adaptive Fuzzy C-Means (92.62%) y Fuzzy C-Means (93.40%) tienen más aciertos que el algoritmo de referencia CMeans (92.33%), y los fallos en el entorno del Adaptive Fuzzy C-Means (9.22%) son menores que los del CMeans (16.37%) y del Fuzzy C-Means (16.11%). El algoritmo Spatially Kernelized Fuzzy C-Means consigue una mejor relación entre aciertos (90.02%) y fallos (11.79%) que los anteriores. El Fuzzy K-Nearest Neighbor tiene menos porcentaje de aciertos (87.48%) y también menor porcentaje de fallos (8.93).

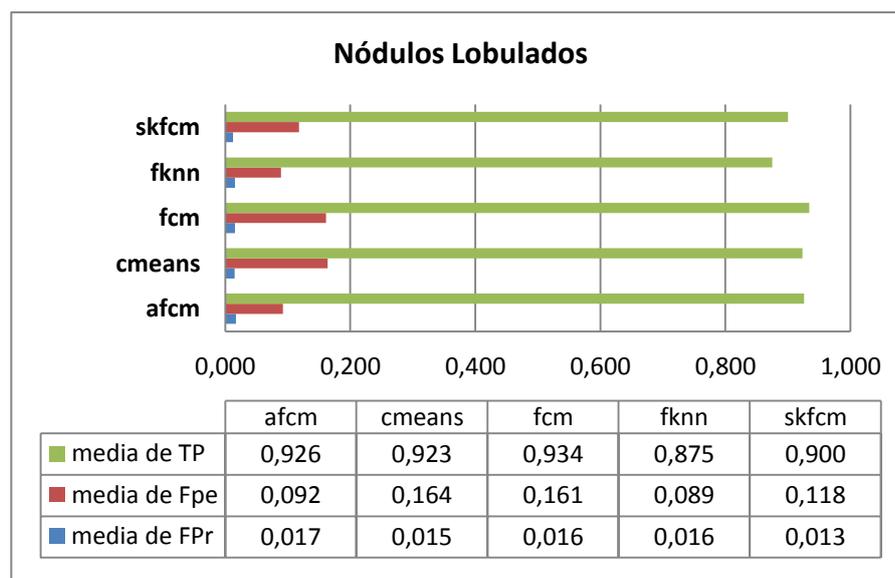
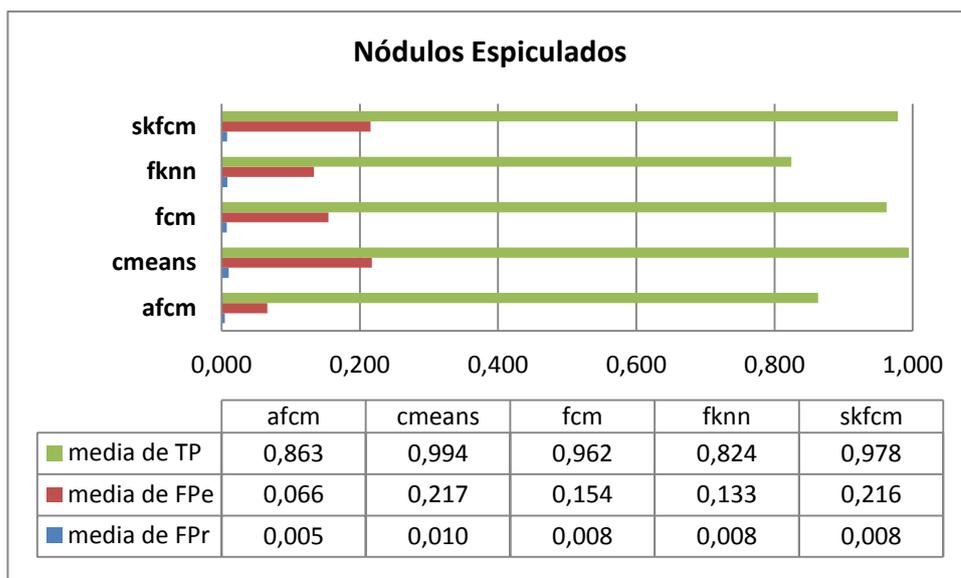


Figura 77. Comparación de aciertos y fallos para subgrupo de imágenes con NPS con márgenes lobulados.

Destaca ligeramente el algoritmo Adaptive Fuzzy C-Means por presentar en los resultados más aciertos y menos fallos que los demás.

La Figura 78 indica los resultados para el subgrupo de imágenes de validación que contienen un NPS con espículas en los bordes. En el porcentaje de aciertos destaca el algoritmo CMeans, con un 99.42%, pero este algoritmo es el que presenta mayor cantidad de falsos positivos (21.73%). El algoritmo Spatially Kernelized Fuzzy C-Means con una media de verdaderos positivos de 97.85%, y de falsos positivos 21.58% tiene resultados similares al de referencia CMeans. Con los resultados menos destacados figura el Fuzzy K-Nearest Neighbor (aciertos 82.44%, y fallos de 13.35%). Es el algoritmo Fuzzy C-Means el que destaca con un 96.21% en los verdaderos positivos y con un 15.44% de falsos positivos en el entorno y un 0.78% en el resto. El algoritmo con menor porcentaje de fallos es el Adaptive Fuzzy C-Means (6.61%) siendo sus aciertos de un 86.31%.



**Figura 78. Comparación de aciertos y fallos para subgrupo de imágenes con NPS cuyos márgenes tienen espículas.**

#### 4.2.2 Sensibilidad y Especificidad de las pruebas de validación.

Todos los resultados obtenidos con las pruebas generales por algoritmo han sido significativos. En general su sensibilidad y especificidad han sido altas. Al representar en el espacio ROC -para cada algoritmo- el resultado de la prueba general sobre las imágenes de validación, todos los pares (TPf, FPe) se localizan, claramente, por encima de la diagonal principal.

Los puntos que tienen la mejor relación entre verdaderos positivos y falsos positivos se encuentran en el intervalo  $[0, 0.2]$  para abscisas y  $[0.8, 1]$  para ordenadas; considerada esta esquina superior izquierda del espacio ROC la zona de clasificación óptima.

En el apartado anterior se ha corroborado que el algoritmo CMeans no es adecuado para la segmentación de imágenes que contienen un nódulo con densidad en vidrio deslustrado. Entonces, separadas dichas imágenes del conjunto total de validación, los resultados de segmentación con este método son bastante parecidos a los obtenidos con los demás algoritmos.

En la Tabla 47 se muestra la estadística de aciertos y fallos en la región de interés de las imágenes de validación segmentadas, y se adjunta su representación en el espacio ROC. Todos los pares  $(FPeF, TPf)$  que representan el resultado de la segmentación del clúster nódulo por cada imagen del conjunto utilizado para la validación, dentro de este espacio de representación, se encuentran localizados por encima de la diagonal principal, siendo por lo tanto todos los valores significativos. La sensibilidad de esta prueba es alta dado que la fracción de verdaderos positivos (TPf) tiene un valor medio de 0.95 con una desviación estándar de 0.08, y el complementario de la especificidad, dependiente de la imagen concreta segmentada, con una media de 0.19 y dispersión de 0.16.

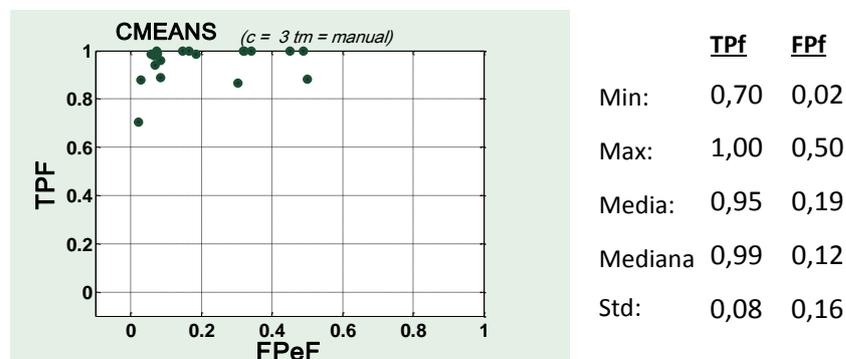


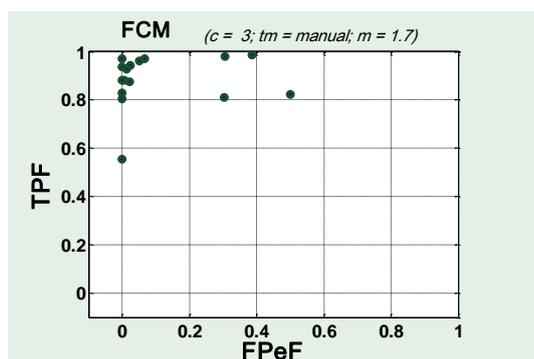
Tabla 47. Resultados en el espacio ROC y estadísticos de la segmentación del algoritmo CMEANS para el grupo de imágenes de validación con NPS sólido o semisólido.

La Tabla 48 resume los resultados obtenidos con el algoritmo Fuzzy C-Means al segmentar las imágenes en tres clústeres, eligiendo las muestras de los centroides iniciales manualmente sobre la zona nódulo, y empleando un factor fuzzy de valor 1.7.

Todos los pares  $(FPeF, TPf)$  representados en el espacio ROC están alejados, por encima, de la diagonal principal, ello evidencia la validez de la prueba al ser todos sus puntos significativos. La mayoría de los puntos están en la zona óptima del espacio ROC, con una

sensibilidad próxima al 100%, indicativo ello de la práctica inexistencia de Falsos Negativos, y con una especificidad también alta por ser muy pocos los Falsos Positivos existentes en los resultados correspondientes a dichos puntos.

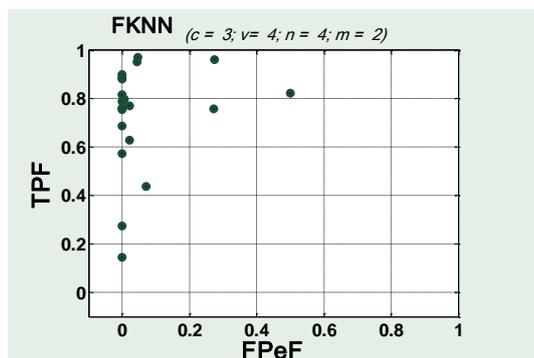
La media de aciertos es 0.91 con desviación 0.11, y la de fallos 0.13 con desviación 0.16 en la región de interés de las imágenes segmentadas. El rango de valores para los fallos es desde un mínimo de 0.00 hasta un máximo de 0.50, y para los aciertos su rango va desde un valor mínimo 0.56 hasta el máximo 1.00.



	<u>TPf</u>	<u>FPf</u>
Min:	0,56	0,00
Max:	1,00	0,50
Media:	0,91	0,13
Mediana	0,94	0,04
Std:	0,11	0,16

Tabla 48. Resultados en el espacio ROC y estadísticos de la segmentación con el algoritmo FCM para grupo de imágenes de validación con NPS sólido o semisólido.

Se observan en la Tabla 49 los estadísticos de los resultados de la prueba realizada con el algoritmo Fuzzy K-Nearest Neighbor, en la que se han empleado como parámetros óptimos a validar un número de 3 clases para segmentar las imágenes con un factor fuzzy 2, y considerando 4 vecinos y 4 muestras por clúster.



	<u>TPf</u>	<u>FPf</u>
Min:	0,15	0,00
Max:	0,97	0,50
Media:	0,73	0,06
Mediana	0,78	0,00
Std:	0,22	0,13

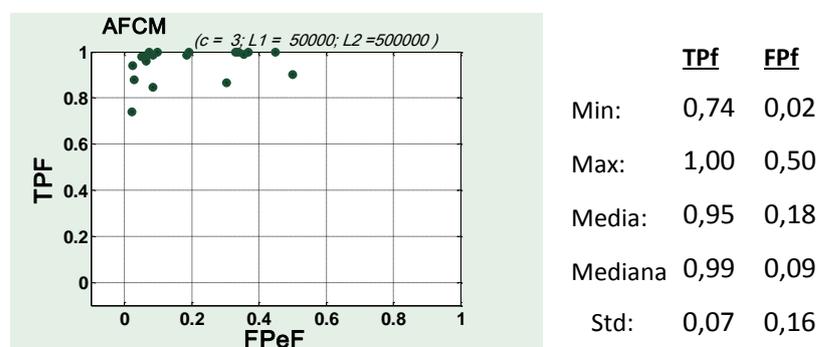
Tabla 49. Resultados en el espacio ROC y estadísticos de la segmentación con el algoritmo FKNN para el grupo de imágenes de validación con NPS sólido o semisólido.

Con el algoritmo Fuzzy K-Nearest Neighbor no están la mayoría de los pares (TPf, FPf) en la zona óptima, aunque todos los puntos están ubicados por encima de la diagonal

principal y por tanto son valores significativos. La sensibilidad para este algoritmo indica la existencia de falsos negativos y la especificidad apunta a la existencia de falsos positivos en la región de interés. Se ha obtenido una media de 0.73 (desviación estándar de 0.22) en los aciertos y media de 0.06 en los fallos. Oscilando el rango de la fracción de verdaderos positivos entre 0.15 y 0.97, y el de la fracción de falsos positivos entre 0.00 y 0.50.

Se indica en la Tabla 50 el resumen de la prueba de validación realizada con el algoritmo Adaptive Fuzzy C-Means para la que se han empleado 3 clases, un valor 50000 de penalización de grandes variaciones (L1) y una penalización de discontinuidades L2 con valor 500000. La representación gráfica en el espacio ROC muestra una concentración de puntos en la zona óptima del espacio unitario. Siendo la media de aciertos 0.95 con una desviación estándar de 0.07 y la media de fallos 0.18 con desviación 0.16. El punto con el valor más bajo de verdaderos positivos ha sido 0.74 y el mínimo para los falsos positivos fue 0.02.

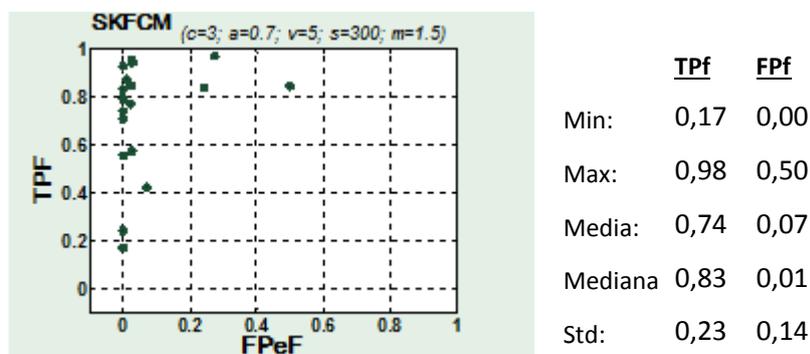
La sensibilidad alcanzó máximos de 100%, no habiendo en dicha imagen segmentada ningún falso negativo y la especificidad máxima ha sido del 98% (pues el mínimo de FPf es 0.02). Todos los pares (TPf, FPf) son suficientemente significativos al estar ubicados muy por encima de la diagonal en el espacio ROC.



**Tabla 50. Resultados en el espacio ROC y estadísticos de la segmentación con el algoritmo AFCM para grupo de imágenes de validación con NPS sólido o semisólido.**

Finalmente, la Tabla 51 corresponde al resumen de las segmentaciones realizadas con el algoritmo Spatially Kernelized Fuzzy C-Means, validado con los valores de parámetros óptimos: Número de clústeres = 3, alfa = 0.7, ventana de vecindad = 5x5, sigma = 300, valor fuzzy = 1.5. En ella se aprecia que todos los pares de valores (TPf, FPf) son significativos, porque están ubicados por encima de la diagonal del espacio de representación. Se observan varios puntos ubicados en la zona óptima del espacio ROC,

aunque no son mayoría, permitiendo obtener en algunas imágenes segmentadas un 98% de sensibilidad y un 100% de especificidad. Los resultados presentan una media en la fracción de verdaderos positivos de 0.74, siendo la desviación típica de 0.23 y la fracción de falsos positivos con media de 0.07.



**Tabla 51. Resultados en el espacio ROC y estadísticos de la segmentación con el algoritmo SKFCM para grupo de imágenes de validación con NPS sólido o semisólido.**

Finalmente destacar que los algoritmos CMeans y Adaptive Fuzzy C-Means presentan una concentración de datos en toda la zona superior del espacio ROC, pues sus valores corresponden con mayores fracciones de falsos positivos y de verdaderos positivos, siendo por tanto adecuado emplear estos algoritmos cuando el estatus real es falso. El algoritmo Fuzzy C-Means presenta mayor concentración de datos en la zona óptima, siendo por tanto el algoritmo que presenta mayor poder discriminante en la clasificación de los píxeles correspondientes a puntos pertenecientes a los NPS para todas las imágenes en general. Y los algoritmos que entre sus parámetros tienen en cuenta al vecindario (Fuzzy K-Nearest Neighbor y Spatially Kernelized Fuzzy C-Means) muestran una concentración de datos próxima al eje TPF, es decir, con una fracción de falsos positivos próxima a cero, siendo adecuado utilizar estos últimos algoritmos cuando el estatus real es verdadero.

### 4.3 Validación cruzada

Para verificar los valores óptimos de los parámetros y su validez con este tipo de imágenes, se considera adecuado disponer de CT de tórax distintas a las utilizadas hasta el momento. Para ello se ha pedido a un investigador del grupo – denotado por G1 de aquí en adelante-, las imágenes de sus estudios CT de tórax con las que había obtenido el peor y el mejor resultado al segmentarlas con sus propios valores de parámetros. A la vez se

han utilizado las imágenes del grupo de pruebas de este trabajo – denotadas por G2- con las que se había obtenido el peor y el mejor resultado; para, estas últimas, ser segmentadas con los valores de parámetros utilizados por el otro investigador (G1).

Este proceso se ha realizado para los algoritmos Fuzzy CMeans y Spatially Kernelized Fuzzy C-Means.

Las máscaras de las imágenes de G1 difieren de las máscaras utilizadas en este trabajo, pues sólo indican con 1 ó 0 si los píxeles de la imagen respectiva corresponden a un punto del nódulo o no. No distinguiendo la zona difusa del entorno del nódulo, como ocurre con las máscaras utilizadas para las imágenes de esta tesis.

En la segmentación comparativa, sobre cada imagen, se han utilizado los mismos centroides tanto en las pruebas con valores de parámetros del investigador G2 como los del investigador G1. En la combinación de valores de parámetros a utilizar hay coincidencia en la segmentación en tres clases para la máxima detección de píxeles pertenecientes al nódulo. Se muestran en la Tabla 52 los valores utilizados como parámetros de entrada por los investigadores G1 y G2 para la segmentación del NPS de cualquier imagen de CT con el algoritmo Fuzzy CMeans.

<b>FCM</b>	<b>c</b>	<b>m</b>
<b>G1</b>	3	2
<b>G2</b>	3	1,7

**Tabla 52. Valores de parámetros utilizados en la validación cruzada para el algoritmo Fuzzy CMeans.**

La función kernel, utilizada por el investigador G1 y la empleada por G2 con el algoritmo Spatially Kernelized Fuzzy C-Means, hace corresponder diferente valor al parámetro Sigma por ser distinta la expresión asignada al denominador, tal como se especifica a continuación:

<b>G1</b>	<b>G2</b>
$K(x, y) = e^{\frac{-(x-y)^2}{2\sigma^2}}$	$K(x, y) = e^{\frac{-(x-y)^2}{\sigma^2}}$

La Tabla 53 indica los valores de entrada al algoritmo Spatially Kernelized Fuzzy C-Means para las imágenes a segmentar utilizados por los investigadores G1 y G2.

Las imágenes utilizadas por los investigadores, con las que ambos han obtenido los peores resultados al segmentar el clúster nódulo, se han vuelto a segmentar empleando los

valores óptimos de parámetros obtenidos en esta tesis para la tipología de nódulo con densidad en vidrio deslustrado, y cuyos valores se indican en la Tabla 54.

SKFCM	c	m	vecindad	Alfa	Sigma
G1	3	2	3	0,5	$\sqrt{2 * 550^2}$
G2	3	1,5	5	0,7	300

Tabla 53. Valores de parámetros utilizados en la validación cruzada para el algoritmo Spatially Kernelized Fuzzy C-Means.

Valores parámetros entrada	
FCM	(c=15; m = 256)
SKFCM	(c = 9; a = 0.8; v = 5 ; s = 50; m = 2)

Tabla 54. Valores de parámetros para segmentar imágenes con NPS en vidrio deslustrado.

#### 4.3.1 Resultados empleando las imágenes del investigador G2

La Tabla 55 muestra los resultados obtenidos, con el algoritmo Spatially Kernelized Fuzzy C-Means, para la segmentación de las imágenes clasificadas como peor y mejor del grupo de pruebas utilizado en esta tesis; las columnas Difkook\_G1 y Difkook\_G2 muestran los rendimientos de dichas segmentaciones al aplicar los valores de parámetros de los investigadores G1 y G2 respectivamente. Con los valores de parámetros generales los resultados son malos en la imagen calificada como peor, consiguiendo buenos resultados al segmentar la imagen “peor” con la combinación óptima para la tipología de nódulo deslustrado conseguida en esta tesis.

SKFCM	Cali_Im	Difkook_G1	Difkook_G2
	Peor	0,00	0,00
	Mejor	0,53	1,36
	Deslustrada	0,00	1,08

Tabla 55. Resultados con el Spatially Kernelized Fuzzy C-Means para las imágenes de G2

Las imágenes utilizadas en este apartado disponen de máscaras con zona nódulo y zona entorno del mismo, por lo cual se indican en la Tabla 56 la fracción de aciertos y de fallos en la clasificación de los píxeles del nódulo, del entorno y del resto de la imagen,

obtenidas en las segmentaciones tanto con los valores de parámetros del investigador G1 como del investigador G2.

SKFCM	Cali_Im	fTP_G1	fFPe_G1	fFPr_G1	fTP_G2	fFPe_G2	fFPr_G2
Peor		1,00	1,00	1,00	0,00	0,00	0,82
Mejor		1,00	0,56	0,02	0,97	0,00	0,01
<i>Deslustrada</i>		<i>1,00</i>	<i>1,00</i>	<i>1,00</i>	<i>0,83</i>	<i>0,05</i>	<i>0,02</i>

Tabla 56. Fracción de aciertos y fallos con el algoritmo Spatially Kernelized Fuzzy C-Means para las imágenes de G2.

Con las citadas imágenes de G2, los resultados del algoritmo Fuzzy CMeans se indican en la Tabla 57 y en la Tabla 58.

FCM	Cali_Im	Difkook_G1	Difkook_G2
Peor		-0,01	-0,01
Mejor		1,29	1,36
<i>Deslustrada</i>		<i>-0,01</i>	<i>1,09</i>

Tabla 57. Resultados con el Fuzzy C-Means para las imágenes de G2

El rendimiento más alto en la mejor imagen con los valores paramétricos de G2 (1.36) es debido a la minimización de falsos positivos en el entorno del nódulo; la segmentación con los valores paramétricos de G1 produce más verdaderos positivos (1.00) a consta de aumentar los falsos positivos en el entorno (0.07) y en el resto de la imagen. Con los valores de parámetros generales los resultados son malos en la imagen calificada como peor, consiguiendo buenos resultados (1.09) al segmentar la imagen “peor” con la combinación óptima de G2 para la tipología de nódulo deslustrado.

FCM	Cali_Im	fTP_G1	fFPe_G1	fFPr_G1	fTP_G2	fFPe_G2	fFPr_G2
Peor		0,00	0,01	0,83	0,00	0,01	0,83
Mejor		1,00	0,07	0,02	0,97	0,00	0,01
<i>Deslustrada</i>		<i>0,00</i>	<i>0,01</i>	<i>0,83</i>	<i>0,84</i>	<i>0,07</i>	<i>0,02</i>

Tabla 58. Fracción de aciertos y fallos con el método Fuzzy C-Means para las imágenes de G2

### 4.3.2 Resultados empleando las imágenes del investigador G1

En la Tabla 59 y en la Tabla 60 se muestran los resultados obtenidos en la segmentación de las imágenes etiquetadas como peor y mejor por el investigador G1, en las cuales se aplicaron los valores específicos de parámetros utilizados por G2 y por G1.

SKFCM	Cali_Im	Difkook_G1	Difkook_G2	FTP_G1	fFP_G1	FTP_G2	fFP_G2
	Mejor	1,24	1,32	0,90	0,01	0,95	0,01
	Peor	-0,01	1,36	0,00	0,01	1,00	0,03
	<i>Deslustrada</i>	<i>-0,01</i>	<i>1,39</i>	<i>0,00</i>	<i>0,01</i>	<i>1,00</i>	<i>0,01</i>

Tabla 59. Resultados con el Spatially Kernelized Fuzzy C-Means para las imágenes de G1.

En el caso de la imagen calificada “peor”, se ha vuelto a segmentar utilizando los valores de parámetros óptimos conseguidos para la tipología de nódulos de densidad en vidrio deslustrada, consiguiendo buenos resultados para los dos algoritmos analizados.

FCM	Cali_Im	Difkook_G1	Difkook_G2	FTP_G1	fFP_G1	FTP_G2	fFP_G2
	Mejor	1,31	1,15	0,95	0,02	0,85	0,01
	Peor	-0,01	-0,01	0,00	0,01	0,00	0,01
	<i>Deslustrada</i>	<i>-0,01</i>	<i>1,39</i>	<i>0,00</i>	<i>0,01</i>	<i>1,00</i>	<i>0,01</i>

Tabla 60. Resultados con el método Fuzzy C-Means para las imágenes de G1.

## 4.4 Discusión

Se decidió utilizar algoritmos de clustering difuso para la detección de las ROI, debido a su facilidad para manejar información multidimensional, lo cual permite adaptarlos fácilmente a la clasificación de imágenes; puesto que su baja sensibilidad al ruido debería permitir discernir mejor entre los nódulos y el resto de elementos en la imagen. Además estos algoritmos son adecuados por su capacidad para manejar información ambigua, que suele haber en imágenes médicas debido a la baja relación señal/ruido. En este trabajo se han analizado los resultados proporcionados por algunos de estos algoritmos de clustering difuso, que utilizan diferentes estrategias en la clasificación de los píxeles que componen la imagen.

Se ha preferido analizar un número pequeño de imágenes para determinar mejor las diferentes características que influyen en el resultado, y detectar los posibles problemas

que se puedan dar, por ello en las pruebas para la búsqueda de valores óptimos para los parámetros de cada algoritmo se han utilizado sólo 23 imágenes; todas ellas con NPS, provenientes de diferentes estudios y con características distintas.

Las diferentes segmentaciones fueron obtenidas variando los valores de las combinaciones de parámetros, para cada algoritmo, sobre cada una de las imágenes. Así mismo, para la validación de los mejores valores de parámetros obtenidos se han utilizado otras 23 imágenes distintas provenientes de los mismos estudios, tomadas del Lung Image Database Consortium en enero del año 2005. Dichas imágenes seleccionadas contienen una muestra significativa de la variedad de NPS que se puede encontrar en las imágenes de CT de Tórax.

Existen múltiples trabajos dedicados a determinar la existencia de NPS en las CT, tales como los referenciados en el apartado 2.4. Este trabajo se centra en la detección de la ROI con la finalidad de reducir el número de falsos positivos y aumentar los TP en dicha zona.

Se considera que para contribuir a los CAD, hacen falta métodos de segmentación que preserven la morfología de los distintos tipos de nódulos y sitúen el lugar exacto y extensión de la zona con tumor, por ello se focaliza el trabajo en precisar al máximo los píxeles que forman el NPS, puesto que la región de interés de las imágenes se puede contrastar con las correspondientes mascarar puntuadas por los expertos en la zona concreta de existencia del nódulo. Siendo el objetivo encontrar, para cada uno de los algoritmos utilizados, la combinación de valores paramétricos que permitan agrupar en un único clúster a todos los píxeles constituyentes del nódulo, y que a dicho clúster no pertenezcan píxeles del entorno de dicho nódulo. Además de comparar los mejores resultados obtenidos con los distintos algoritmos, se debe determinar cuál de ellos es el más adecuado, cuantitativamente con mayor capacidad de discriminación según el tipo de nódulo contenido en las CT.

En un principio se pretendía que todos los puntos del nódulo estuviesen agrupados en un solo clúster, y que en dicho clúster estuviesen minimizados los puntos correspondientes a otras zonas de la imagen. Motivo éste para, mediante un preprocesado, intentar eliminar de la imagen todos aquellos elementos que pudiesen originar falsos positivos porque sus intensidades de píxel estuviesen en la misma zona del histograma que los píxeles constituyentes del nódulo. Para ello, el proceso previo a la aplicación de los distintos algoritmos sobre cada una de las imágenes ha sido realizar la etapa de preprocesado ya descrita, algo realizado mayoritariamente en los trabajos de clasificación sobre imágenes

CT referenciados en el capítulo 2. El hecho de realizar la etapa de preprocesado ha facilitado la tarea de segmentación, permitiendo que la ejecución de los algoritmos se realizase empleando menor tiempo computacional porque se fragmenta la imagen en menos particiones.

Se planteó sobre las imágenes CT, como segunda hipótesis de trabajo que: "minimizar el entorno del nódulo, referenciado en la máscara como zona de incertidumbre, sea equivalente a minimizar cualquier entorno mayor que el del NPS". El análisis de los resultados obtenidos llevó a rechazar dicha hipótesis. Pues en varias de las segmentaciones al maximizar los verdaderos positivos en el nódulo y minimizar los falsos positivos en su entorno, origina que en el resto de la imagen según la CT concreta segmentada lo mismo se incrementen los falsos positivos como ocurra todo lo contrario.

Al observar los resultados de las imágenes segmentadas se aprecia que píxeles correspondientes a vasos sanguíneos son clasificados como pertenecientes al clúster nódulo, pues los valores de intensidad de dichos píxeles son similares a los del tumor en un número elevado de imágenes, dando lugar a un mayor número de falsos positivos en los resultados.

Mayoritariamente, en la comunidad investigadora, para la valoración del rendimiento de pruebas realizadas con algoritmos para la detección de nódulos en las imágenes de CT se está utilizando la representación con curvas ROC [Dodd, 2004]; pero se muestran los resultados de una prueba sobre un gran número de imágenes, y en este trabajo se utiliza un grupo pequeño de imágenes y sobre ellas se realiza un número elevado de pruebas con valores paramétricos diferentes. Por ello, para cuantificar la valoración del rendimiento de dichas pruebas, realizadas sobre las mismas imágenes con cada uno de los algoritmos, se han diseñado, han implementado y han utilizado - tomando como base el espacio ROC- los índices de rendimiento de prueba nombrados por los acrónimos Difko\_ok y Azope; con los cuales se pretende determinar de manera objetiva los valores de parámetros que permiten obtener las mejores segmentaciones con cada una de las imágenes, y poder calcular también los estadísticos sobre el conjunto de imágenes, para de esta manera determinar los valores óptimos de los parámetros.

Los nódulos contenidos en las CT tienen distintas formas, tamaños, localizaciones y densidades con intensidades del grupo de píxeles que representan al NPS diferentes. Aunque la valoración de los márgenes y las formas de dichos nódulos es dependiente del observador, y no todos los radiólogos coincidirían en la clasificación, las observaciones de los resultados de cada prueba de segmentación sobre cada una de las imágenes utilizadas,

indican que en el proceso de segmentación tienen una gran influencia las características del NPS; ya que según su tipología así será más o menos precisa la clasificación de los píxeles del clúster nódulo, y es más adecuado para cada tipo utilizar una combinación concreta de valores paramétricos.

En todos los algoritmos utilizados se deben proporcionar unos centroides iniciales, lo ideal es que cada centroide describa perfectamente a cada clúster, para de esta forma poder aumentar la homogeneidad intraregión y la distancia entre regiones. En el análisis del comportamiento paramétrico de los distintos algoritmos empleados se considera más adecuado el centroide obtenido del tipo de muestra manual, pues aunque en los resultados no se han apreciado diferencias significativas mejoran ligeramente los resultados. La elección de dichas muestras manuales de las imágenes se realizó al escoger los píxeles de mayor intensidad sobre la ROI, por ser visualmente los más fáciles de detectar y que además coinciden con aquellos para los que los radiólogos en la máscara asignaron un valor superior a 900.

El parámetro factor fuzzy ( $m$ ), se ha utilizado como variable de entrada en casi todos los algoritmos difusos empleados. Los valores grandes de  $m$  agrandan la diferencia de la ponderación entre términos, es decir,  $m$  actúa de factor diferenciador entre dos situaciones: una en que el píxel a clasificar se encuentra cerca del centroide, y otra en la que el píxel se encuentre cerca de la frontera de la  $K$ -ésima clase difusa. Cuando  $m$  es mayor que los valores convencionales los píxeles clasificados toman sus valores de pertenencia en el intervalo abierto  $(0, 1)$  de tipo real, y cuando  $m$  toma valores bajos los píxeles de la imagen segmentada toman sus valores de pertenencia más próximos a los valores del conjunto binario  $\{0, 1\}$ .

Todas las pruebas realizadas, con los algoritmos empleados, presentan la misma tendencia en el rendimiento frente al número de clústeres. En general con los métodos difusos, los mejores resultados, y como se obtiene un rendimiento más alto en dichas pruebas de segmentación es dividiendo las imágenes en tres clústeres. Ello es debido a que la función de pertenencia con un número de clases mayor no es capaz de producir una división del espacio del problema en la que los píxeles del nódulo queden claramente separados del resto, dependiendo que los píxeles se clasifiquen pertenecientes al nódulo de la imagen del número de clústeres. De la partición en tres clústeres: uno corresponde al fondo de imagen, otro al parénquima pulmonar y el tercero al clúster nódulo. El comportamiento de dicho parámetro ha sido bastante uniforme en los métodos empleados, disminuyendo el rendimiento de las pruebas realizadas con los algoritmos al aumentar dicho número de

particiones. Claramente el 87% de las imágenes, correspondiente al grupo total desprovisto de aquellas que contenían NPS con densidad de tipo vidrio deslustrado, obtiene un rendimiento óptimo segmentando en tres clústeres. Y para el 13% de imágenes restantes, conteniendo tipología de nódulo en vidrio deslustrado, se debe segmentar en un número mayor de clases, para lograr agrupar en un único clúster los píxeles pertenecientes al nódulo; en concreto con el algoritmo Fuzzy CMeans se ha necesitado utilizar 15 clases y un grado de factor fuzzy elevado ( $m=256$ ), en los demás algoritmos fuzzy se han obtenido buenos resultados segmentando en un número menor de clases (8 ó 9 clústeres) al disponer de más parámetros que permiten calibrar el ajuste.

Para dichas imágenes con NPS deslustrado es adecuado utilizar un algoritmo que tenga en cuenta al vecindario de forma explícita, en nuestro caso se utilizan el Fuzzy K-Nearest Neighbor y el Spatially Kernelized Fuzzy C-Means pues los píxeles de este tipo de nódulo están más dispersos y difuminados sobre el parénquima pulmonar. Destacando que con el algoritmo Fuzzy K-Nearest Neighbor se deben utilizar sólo tres o cuatro vecinos, porque si aumenta el número de vecinos disminuye el rendimiento de las pruebas, puesto que éste es sensible a la calidad de las muestras del vecindario, y si no son las adecuadas se degrada la calidad de la segmentación. Y la ventana de vecindario más adecuada a considerar con el algoritmo Spatially Kernelized Fuzzy C-Means es de  $5 \times 5$ , pues al aumentarla disminuyen los aciertos de la zona nódulo aumentando los falsos positivos en el entorno; excepto en las imágenes con nódulos subpleurales, donde es conveniente utilizar una ventana  $10 \times 10$  para poder discriminar mejor la pertenencia de los píxeles correspondientes al tejido pleural.

En general el factor fuzzy más adecuado se encuentra en el intervalo  $[1.5, 2]$ , pero concretamente para esas imágenes con nódulos en vidrio deslustrado es necesario con el Fuzzy C-Means utilizar valores altos, 256, y con el Spatially Kernelized Fuzzy C-Means un valor 50. También hay que matizar que con el algoritmo Adaptive Fuzzy C-Means para segmentar adecuadamente estos nódulos de densidad en vidrio deslustrado hay que utilizar unos campos multiplicadores con valor inferior o igual a 10, para obtener una imagen no muy suavizada para que no se "diluyan" las intensidades de los píxeles de la ROI.

El algoritmo CMeans no es la solución más adecuada para utilizar sobre todo el conjunto de imágenes, debido a las limitaciones de la función de pertenencia que utiliza, pues con él no se consiguen resultados de calidad sobre todo el grupo de imágenes. De hecho en las imágenes con la tipología en vidrio deslustrado sólo se han conseguido resultados

aceptables y buenos empleando algoritmos de clústering difuso, pues poseen mayor capacidad para manejar la información ambigua de los píxeles constituyentes de dicha tipología de NPS.

La cuantificación del rendimiento de las pruebas ha permitido encontrar aquellas con las que se han obtenido buenos resultados, es decir, maximizar los verdaderos positivos en la ROI, y minimizar los falsos positivos en el entorno del nódulo. Esto se ha podido realizar así, por disponer de máscaras informadas que permitieron compararlas con los resultados segmentados de la clasificación a nivel de píxel individual, y no interesar el resto de los puntos de la imagen segmentada por desconocerse la clase a la que pertenecen sus píxeles. En el resumen del análisis estadístico, realizado para cada algoritmo sobre las segmentaciones que habían proporcionado los mejores resultados, vistos en el apartado 4.1.2, se ha constatado que la distribución de los resultados por prueba no es normal, en la gran mayoría de las pruebas es asimétrica con cola por la izquierda, por ello la mediana está en la zona de distribución donde hay un mayor valor de rendimiento. Al estar las medias sesgadas hacia el lado negativo de la distribución, por ser inferiores a las medianas, para la mayoría de los nódulos segmentados del conjunto, se obtienen valoraciones buenas, aunque siempre existen algunas imágenes en las que los resultados segmentados son atípicos. Así, en la mayoría de las mejores pruebas seleccionadas, el valor de la media es inferior a la medida representativa de dónde está el grupo, pues dicho valor medio está aminorado por los pocos resultados atípicos. Dichos resultados atípicos dan lugar a la existencia de una gran desviación estándar, la cual indica que las imágenes son heterogéneas, con diferencias entre ellas. Este resultado es coherente con el hecho de que cada una de las imágenes empleadas tanto en el grupo test como en el grupo de validación proviene de estudios diferentes y de pacientes distintos que tienen NPS de características diversas. Además, la alta cuantificación de la dispersión de los valores en cada prueba, puede ser indicativo del poder discriminante de los algoritmos empleados, y por ello se detecten las diferencias existentes en las diversas imágenes. También hay que tener en cuenta que no existe un valor ideal en la desviación estándar, y el valor máximo posible de dicha desviación es  $(\text{máximo}-\text{mínimo})/2$ , siendo el valor mínimo el obtenido por las imágenes con resultados atípicos.

En los resultados visuales de las segmentaciones obtenidas se aprecia que es más precisa la segmentación con los nódulos circulares que con los espiculares, para todos los métodos empleados; ello es debido a la propia naturaleza del nódulo que es más uniforme, y en el que es más fácil de detectar con precisión el cambio de intensidad en los píxeles que lo

integran. Las imágenes con nódulos de tamaño superior al centímetro producen resultados con gran dispersión en todos los algoritmos utilizados, ello puede ser debido a dos motivos: por una parte estos nódulos tienen un número elevado de espículas que a veces se confunden con tractos fibrosos en el parénquima pulmonar, y por otra parte estos nódulos de mayor tamaño tienen un mayor número de píxeles, con más variedad en los niveles de intensidad, en el entorno que los delimita. Y además dicho entorno está fusionado con el parénquima pulmonar dando lugar a una indefinición entre lo que es nódulo y lo que no. Esto aparece reflejado en las máscaras de los nódulos de las imágenes que son objeto del estudio, y que como objetivo se había planteado reducir la pertenencia al clúster nódulo para estos puntos de la imagen; dichos puntos del entorno del tumor son dudosos para los propios radiólogos ya que su valoración en la máscara se sitúa por debajo de 500.

Se ha observado que en los resultados concernientes a las imágenes con nódulos de tamaño entre 6 y 10 mm existe más dispersión que en el resto de las imágenes; lo cual es debido a que entre las CT con estos tamaños de nódulo, se encuentran aquellos cuya densidad es en vidrio deslustrado, y estos últimos es necesario detectarlos segmentando las imágenes que los contienen en un mayor número de clases, algo no conveniente para la gran mayoría del grupo de imágenes.

El algoritmo Fuzzy CMeans es el que mejor rendimiento proporciona al segmentar las imágenes con nódulos de este tamaño intermedio; siendo dicho método en el que menos influencia tiene el escaso contraste existente en los nódulos de este tamaño cuya densidad es en vidrio deslustrado.

En el subgrupo de imágenes con nódulos subpleurales existe dispersión en los resultados de las segmentaciones. Esto es debido a que al preprocesar la imagen estando estos nódulos en contacto con la pleura, en unas imágenes hay menos puntos de la zona nódulo que en otras tras la eliminación de parte de la pleura, y existen más puntos clasificados falsos positivos en el entorno del nódulo.

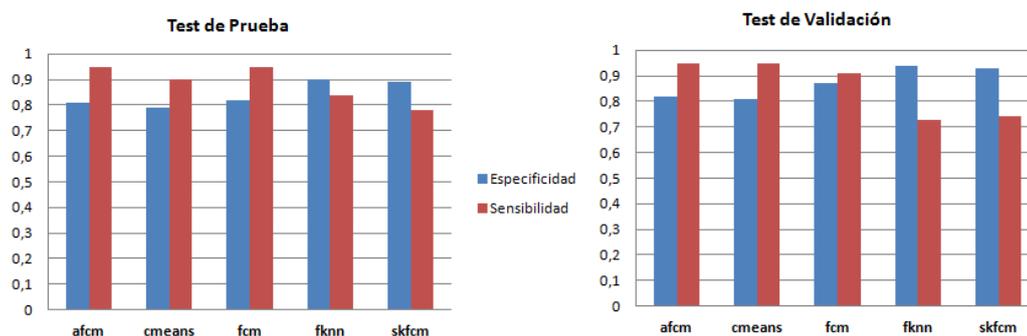
En los demás subgrupos de imágenes se ha logrado reducir la dispersión respecto a la media sobre todo el conjunto de imágenes de prueba. Destacando que en los nódulos con forma espiculada, más difíciles de procesar debido a su naturaleza, es el algoritmo Adaptive Fuzzy CMeans el que obtiene mejores resultados segmentados, porque este método a medida que realiza la segmentación compensa las inhomogeneidades de la zona nódulo en la imagen, pues con los campos multiplicadores de primer orden L1 y de

segundo orden L2 realiza un suavizado en la ROI de la imagen, lo que da lugar a una disminución de falsos positivos en el entorno del nódulo.

Los algoritmos que tienen en cuenta las características de los píxeles del vecindario, Fuzzy K-Nearest Neighbor y Spatially Kernelized Fuzzy CMeans, son los que proporcionan mayor rendimiento con menor dispersión en los nódulos cuya densidad es en vidrio deslustrado, porque priorizan la información espacial que contienen los píxeles que pertenecen al clúster que se etiqueta como nódulo. Teniendo además ambos algoritmos, tanto en los test de pruebas como en el test de de validación, resultados con valores más altos de especificidad que de sensibilidad. Así pues, estos algoritmos que tienen en cuenta los píxeles vecinos de forma específica en sus parámetros de entrada, son más útiles si el criterio a emplear para la clasificación es más estricto, es decir, si es necesario que tenga más ponderación la existencia de fracciones pequeñas de falsos positivos. Pero en general el algoritmo Fuzzy K-Nearest Neighbor presenta resultados muy dependientes de la calidad de las muestras iniciales de los centroides, y resulta bastante laboriosa la asignación del número de muestras del vecindario.

Las pruebas con el algoritmo Adaptive Fuzzy CMeans dan lugar a resultados con mayor sensibilidad que especificidad, siendo por tanto más útil emplear este algoritmo cuando interesa más el hecho de detectar verdaderos positivos que falsos positivos, es decir emplear un criterio más laxo.

Estos índices de eficacia, sensibilidad y especificidad, se muestran en la Figura 79 como resultado de la comparativa, al segmentar, utilizando en los parámetros de entrada los valores óptimos con los algoritmos empleados por este trabajo sobre los dos grupos de imágenes: el de prueba y el de validación.



**Figura 79. Resultado de la segmentación por algoritmo, en términos de sensibilidad y especificidad, sobre el conjunto de imágenes de Prueba y de Validación.**

Se observa que existen diferencias entre los resultados obtenidos para el test de pruebas y para el test de validación, pues aunque las imágenes pertenecen a los mismos estudios son de cortes diferentes.

## 5 CONCLUSIONES Y FUTUROS DESARROLLOS

En este capítulo se enumeran las principales contribuciones de esta tesis, mencionando en un segundo apartado las líneas de trabajo futuro.

### 5.1 Conclusiones

Tras sistematizar, ordenar y presentar el resumen de los datos resultantes de las segmentaciones de imágenes que permitieron deducir los valores de parámetros óptimos a emplear con cada algoritmo, y orientar en la demostración de las hipótesis planteadas y en la consecución de los objetivos propuestos se concluye que:

- Se han definido índices de rendimiento de prueba, denominados con los acrónimos Azope y Difko\_ok sobre el espacio unitario ROC, para la valoración objetiva de las segmentaciones.
- Se ha propuesto un método de trabajo para sistematizar la valoración de los resultados de las segmentaciones, y establecido una metodología para encontrar la combinación de valores de parámetros óptimos en la segmentación de regiones de interés (NPS) valorando la calidad de dichas imágenes segmentadas.
- Ninguno de los algoritmos de segmentación empleados proporciona una clasificación perfecta de los píxeles de los NPS; aunque los resultados obtenidos, a pesar de ser modestos, son significativos. Y para aquellos píxeles de los nódulos para los que la máscara demuestra que había un alto acuerdo entre los especialistas médicos (puntos con una puntuación igual o mayor a 900), estos algoritmos son capaces de detectarlos casi al 100%.
- Se ha refutado la hipótesis2 ante la evidencia de los resultados. No se puede establecer una relación entre falsos positivos en el entorno del nódulo y falsos positivos en el resto de la imagen segmentada, dado que los falsos positivos en el resto son muy dependientes de la imagen concreta a particionar, y el hecho de minimizar los falsos positivos en dicho entorno no implica su minimización en el resto de la imagen. También se ha refutado la hipótesis5 por encontrar segmentaciones donde son nulos los verdaderos positivos del clúster nódulo correspondientes a la ROI.
- En el 87% de las imágenes de esta tesis, lo adecuado es segmentar en tres clústeres; por ser éste el número de clases óptimo que permite aislar, con mayor

precisión, en un único clúster al nódulo de la imagen preprocesada. Y en el resto de las imágenes segmentar en un mayor número de clases, siendo su número concreto dependiente del algoritmo a emplear y de los valores de los demás parámetros de entrada.

- El parámetro de entrada a los algoritmos de mayor influencia y más peso en el rendimiento de las pruebas es el número de clústeres en los que segmentar la imagen, y después el factor fuzzy. El ajuste fino de los demás parámetros de entrada incide de forma significativa en el aumento o disminución de los aciertos o fallos. Como aportación, la Tabla 46 muestra los valores óptimos de los parámetros de segmentación por algoritmo después de verificarse la consecución de las hipótesis1, hipótesis3 e hipótesis4 planteadas inicialmente.
- El tamaño, la forma y fundamentalmente la densidad del Nódulo Pulmonar Solitario contenido en las imágenes influyen en el rendimiento de todas las pruebas realizadas con los algoritmos empleados. Por ello, sin conocer la tipología de NPS existente, se debe proceder a segmentar, tras el preprocesado de las imágenes TAC, utilizando los valores óptimos generales del algoritmo elegido; y con aquellas imágenes en las que se obtengan puntuaciones de rendimiento muy extremas y atípicas por la izquierda de la distribución de datos, se ejecutará una nueva segmentación aplicando en dicho algoritmo las precondiciones de valores específicos correspondientes a la tipología NPS en vidrio deslustrado.

## 5.2 Propuestas

Se considera conveniente, como mejora a este trabajo:

- Incluir en la etapa de preprocesado la identificación de vasos sanguíneos en la presegmentación para facilitar la separación de nódulos y vasos, y reducir así los falsos positivos.
- Realizar una validación de los valores óptimos para los parámetros de entrada utilizando una base de imágenes mayor.
- Utilizar el método de trabajo desarrollado en esta tesis para sistematizar la búsqueda de valores paramétricos con otros algoritmos de segmentación diferentes.
- Emplear algoritmos genéticos para determinar y ajustar los valores paramétricos de segmentación.
- Comparar los valores paramétricos óptimos obtenidos vía algoritmos genéticos y vía el método de trabajo desarrollado en esta tesis.
- Si utilizando una gran base de datos para validar los valores paramétricos se obtuviesen distribuciones de los datos normales, entonces utilizando una estadística inferencial realizar un cálculo de probabilidades para conocer qué resultados de segmentación muy buenos o muy malos serían improbables por azar según la relación existente entre la desviación típica y la distribución normal.
- Trabajar con las imágenes en el dominio de la frecuencia para la detección y determinación de NPS.

## 6 BIBLIOGRAFÍA

[**AAAI,1994**] Proc. 1994 AAI Spring Symposium. *Medical Applications of Computer Vision*. W. Wells and E. Grimson, eds., pp. 1-200. Mar. 1994.

[**Abdel-Mottaleb,1998**] M Abdel-Mottaleb, S Krishnamachari and NJ Mankovik. *Performance Evaluation of Clustering Algorithms for Scalable Image Retrieval*. In KW Bowyer and PJ Phillips. *Empirical Evaluation Techniques in Computer Vision*. IEEE Computer Press, pp45-56. 1998.

[**Abraira, 1997**] Abraira Santos V. *Precisión de las clasificaciones clínicas*. Tesis Doctoral, UCM. 1997.

[**AECC, 2011**] *Asociación Española Contra el Cancer*. Accesible online via <https://www.aecc.es/SobreElCancer/elcancer/Paginas/Comosediagnostica.aspx>, 2011.

[**Alonso, 2000**] Alonso A, Arcay B, Castro A. *Analysis and Evaluation of Hard and Fuzzy Clustering Segmentation Techniques in Burned Patients Images*, *Image and Vision Computing*, 18:1045-54, 2000.

[**Anaesthetist, 2011**] *The magnificent ROC*. accesible online vía <http://www.anaesthetist.com/mnm/stats/roc/>. 2011.

[**Aoyama, 2003**] Aoyama,M., Li,Q., Katsuragawa, S., et al.: Computerized scheme for determination of the likelihood measure of malignancy for pulmonary nodules on low-dose CT images. *Med. Phys.* 30, 387– 394, 2003.

[**Arimura, 2004**] Arimura, H., Katsuragawa, S., Suzuki, K., et al. *Computerized scheme for automated detection of lung nodules in low-dose computed tomography images for lung cancer screening*. *Acad. Radiol.* 11, 617–629, 2004.

[**Armato, 1999**] Armato S.G., GigerM.L., Moran C.J., et al. *Computerized detection of pulmonary nodules on CT scans*. *Radiographics* 19: 1303–1311, 1999.

[**Armato, 2004a**] Armato III S.G., McLennan G., M.F. McNitt-Gray for the Lung Image Database Consortium Research Group. *Lung Image Database Consortium: Developing a Resource for the Medical Imaging Research Community*. *Radiology*, 232, pp. 739-748, 2004.

[**Armato, 2004b**] S. G. Armato and W. F. Sensakovic, *Automated lung segmentation for thoracic CT*, *Acad. Radiol.*, vol. 11, n. 9, pp. 1011–1021, 2004.

[**ASEH, 2010**] Asociación de la Sociedad Española de Hipertensión. *Historia del razonamiento estadístico*. Accesible online vía <http://www.seh-lelha.org/historiastat.htm>. 2010.

[**Azpitarte, 2006**] Azpitarte R. LL., Pérez J.C., Paredes R. *Técnicas Reconocimiento De Formas Aplicadas Al Diagnóstico De Cáncer Asistido Por Ordenador*. RevistaeSalud.com. Vol. 2, n. 7 III Trim. 2006.

[**Bae, 2005**] Bae, K.T., Kim, J.S., Na, Y.H., et al. *Pulmonary nodules: automated detection on CT images with morphologic matching algorithm— preliminary results*. Radiology 236, 286–294, 2005.

[**Beigelman-Aubry, 2007**] Beigelman-Aubry C, and others. *Computer-aided detection of solid lung nodules on follow-up MDCT screening: evaluation of detection, tracking, and reading time*. AJR Am J Roentgenol. ;189(4):948-55. 2007.

[**Bezdek, 1981**] Bezdek, J.C. *Pattern recognition with fuzzy objective function algorithms*. New York: Plenum. 1981.

[**Bezdek 1984**] Bezdek, J. C., Ehrlich, R., Full, W. *FCM: the Fuzzy C-Means Clustering Algorithm* Computer and Geosciences, vol. 10, n. 2-3, pp. 191-203, 1984.

[**Bezdek, 1987**] J.C. Bezdek, R.J. Hathaway, M.J. Sabin and W.T. Tucker. *Convergence Theory for Fuzzy C-Means Counterexamples and Repairs*. IEEE Transactions on Systems, Man and Cybernetics, 17:5, pp. 873-877. 1987.

[**Bezdek 2005**] Bezdek, J. C., Keller, J., Krisnapuram, R., Pal, N. R. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*. Springer, 2005.

[**Bradley, 1997**] Bradley A. P. *The use of the area under the roc curve in the evaluation of machine learning algorithms*. Pattern Recognition, Vol. 30, n. 7, pp. 1145-1159, 1997

[**Briggs, 1987a**] Briggs, William L. *A Multigrid Tutorial. Presented by Van Emden Henson. Center for Applied Scientific Computing*. Lawrence Livermore National Laboratory. Accesible online via <http://www.math.ust.hk/~mawang/teaching/math532/mgtut.pdf>.

[**Briggs, 1987b**] Briggs, William L. *A Multigrid Tutorial part two. Presented by Van Emden Henson. Center for Applied Scientific Computing*. Lawrence Livermore National Laboratory. Accesible online en [https://computation.llnl.gov/casc/people/henson/mgtut/ps/mgtut\\_part2.pdf](https://computation.llnl.gov/casc/people/henson/mgtut/ps/mgtut_part2.pdf)

[**Brown, 2007**] M.S. Brown and others, *CAD in clinical trials: Current role and architectural requirements*, Computerized Medical Imaging and Graphics , n.31, pp. 332–337, 2007.

[**Burckhardt, 2000**] K. Burckhardt, Ch. Gerber, J. Hodler, H. Nötzli and G.Székely. *Precision of distance determination using 3D to 2D projections: The error of migration measurement using X-ray images*. Medical Image Analysis, pp 375-388, 2000.

[Burgueño, 1995] Burgueño M.J. and others. *Las curvas ROC en la evaluación de las pruebas diagnósticas*. Med Clin. n. 104, pp. 661-670. 1995

[Castañer, 2011] Castañer González E. *Nódulo y masa pulmonar*. Udiat / Parc Taulí, accesible online vía [http://www.radiologsdecatalunya.cat/formacio/resums/GB07BT07\\_R.pdf](http://www.radiologsdecatalunya.cat/formacio/resums/GB07BT07_R.pdf).

[Castro, 2004] Castro A. *Estudio y validación de algoritmos avanzados de segmentación en imágenes médicas. Integración en un sistema de información*. Tesis Doctoral, UDC. 2004.

[Castro, 2006] Castro A, Boveda C and Arcay B. *Analysis of fuzzy clustering algorithms for the segmentation of burn wounds photographs*. Lecture Notes in Computer Science, vol. 4142, pp. 491–501, Povia de Varzim 2006.

[Castro, 2007] Castro A, Boveda C and Arcay B. *A new method for initialization of clustering algorithms based on histogram analysis*. In The Seventh IASTED International Conference on Visualization, Imaging and Image Processing, pp. 176-181, Mallorca 2007.

[Castro, 2010a] Castro A, Boveda C, Rey A and Arcay B. *An analysis of different clustering algorithms for ROI detection in high resolutions CT lung images*. Lecture Notes in Computer Science. Vol 6374, pp 241-248, Varsovia 2010.

[Chang, 2004] Chang, S., Emoto, H. Metaxas, D.N., et al.: *Pulmonary micronodule detection from 3-d chest CT*. Medical image computing and computer-assisted intervention, vol. 3217, pp. 821–828. Springer, Berlin, 2004.

[Cheng, 2005] H.D. Cheng, X.J. Shi, R. Min, et al. *Approaches for automated detection and classification of masses in mammograms*. Pattern Recognition, page in press, 2005.

[Cohen, 1960] Cohen J. *A coefficient of agreement for nominal scales*. Educ Psychol Mea. 20: 37-46. 1960.

[Conde, 2009] Conde Valero A., E. Navascués Martínez. *Estudio del nódulo pulmonar solitario. Manual de Neumología Neumosur*. 3ª ed. Editorial Madrid SCM. pp. 233-42. 2009.

[Corbo, 2004] Corbo Pereira D.N. *Tomografía Axial Computada*. XIII Seminario de Ingeniería biomédica. Universidad de la República Oriental del Uruguay .2004

[Courtney, 2001] P. Courtney, NA Thacker. *Performance Characterisation in Computer Vision: The Role of Statistics in Testing and Design*. Imaging and Vision Systems: Theory, Assessment and Applications. Jacques Blanc-Talon and Dan Popescu (Eds.). NOVA Science Books, 2001.

[Dehmeshki, 2008] Dehmeshki, J., Amin, H., Valdivieso, M., et al. *Segmentation of pulmonary nodules in thoracic CT scans: a region growing approach*. IEEE Trans. Medi. Imaging 27, 467–480, 2008.

- [**DelRio, 2008**] D. del Río Medina, C. Bocanegra Sánchez et al. *La cabecera del estándar DICOM*. Revista eSalud.com, vol. 4, n. 16, pp. 1-23, 2008.
- [**DelRio, 2009**] D. del Río Medina, C. Bocanegra Sánchez et al. *La imagen 2D en el estándar DICOM*. Revista eSalud.com, vol.5, n 17, pp. 1-20, 2009.
- [**Diciotti, 2008**] Diciotti, S., Picozzi, G., Falchini, M., et al. *3D segmentation algorithm of small lung nodules in spiral CT images*. IEEE Trans. Inf. Technol. Biomedical 12, 7–19, 2008
- [**DICOM, 2009**] *Digital Imaging and Communications in Medicine*. Accesible on line vía <ftp://medical.nema.org/medical/dicom/> , 2009.
- [**DL Pham, 1999a**] DL Pham, JL Prince. *An adaptive fuzzy C-means algorithm for image segmentation in the presence of intensity inhomogeneities*. Pattern Recognition Letters 20 (1999) 57-68
- [**DL Pham, 1999b**] DL Pham, JL Prince. *Adaptive Fuzzy Segmentation of Magnetic Resonance Images*. IEEE Transactions on Medical Imaging, Vol.18, Nº 9, Sept. 1999.
- [**Dodd, 2004**] Dodd Lori E. and others. *Assessment Methodologies and Statistical Issues for Computer-Aided Diagnosis of Lung Nodules in Computed Tomography: Contemporary Research Topics Relevant to the Lung Image Database Consortium*. Academic Radiology. Vo.11, n 4, pp 462-475. 2004.
- [**Dorfman, 1992**] D.D. Dorfman, K.S. Berbaum, and C.E. Metz. *Receiver operating characteristic rating analysis: generalization to the population of readers and patients with the jackknife method*. Investigative Radiology, vol. 27, pp. 723-731. 1992.
- [**Dunn, 1973**] Dunn J.C. *A Fuzzy Relative of the ISODATA Process and its Use in Detecting Compact Well-Separated Clusters*, Journal of Cybernetics 3: 32-57. 1973.
- [**Egan, 1975**] Egan, JP. *Signal Detection Theory and ROC Analysis*. New York. Academic Press. 1975.
- [**El-Baz, 2007**] El-Baz, A., Gimel'farb, G., Falk, R., et al. *A new CAD system for early diagnosis of detected lung nodules*. IEEE Int. Conf. Image Process. 2:II-461–II-464, 2007.
- [**ELCAP, 2010**] *Early Lung Cancer Action Project*. Accesible online vía <http://www.ielcap.org/>. 2010.
- [**Elmundosalud, 2009**] *Cáncer de pulmón*. Especiales elmundosalud. Accesible online vía <http://www.elmundo.es/elmundosalud/especiales/cancer>. 2009.
- [**Faardit, 2011**] *Diagnóstico por Imágenes y Cáncer de Pulmon*. Accesible on line vía <http://edu.faardit.org.ar/>. 2011.

[Fawcett, 2006] T. Fawcett. *An introduction to Roc Analysis*. Pattern Recognition Letters 27, 861-874, 2006.

[Fetita, 2003] Fetita, C.I., Prêteux, F.J., Beigelman-Aubry, C., et al. *3-D automated lung nodule segmentation inHRCT*. LectureNotes in Computer Science, vol. 2878/2003, pp. 626–634. Springer, Berlin, 2003.

[FMV, 2011] *Nódulo Pulmonar Solitario*. Servicio de diagnóstico por imágenes. Facultad de Medicina Virtual. Universidad de Buenos Aires, accesible online vía <http://www.fmv-uba.org.ar/>. 2011.

[Forsyth, 2002] D.A. Forsyth, J. Ponce. *Computer Vision A Modern Approach*. Prentice Hall. 2002.

[Freeman, 1972] H. Freeman, T.S. Huang, A. Rosenfeld, and A. van Dam, eds. *Computer Graphics and Image Processing*. Academic Press, New York. 1972.

[Gabielli, 2007] Gabielli M, Muñoz S., Passalacqua S. et all. *Nódulo Pulmonar Solitario: Desafío diagnóstico y terapéutico*. Cuad. Cir., vol. 21, pp. 65-74, 2007.

[Gamavi, 2005] Garnavi, R., Baraani-Dastjerdi, A., Abrishami Moghaddam, H., et al.: *A new segmentation method for lung HRCT images*. In: Lovell, B.C., Maeder, A.J., Caelli, T., Ourselin, S., (eds.): *Proceedings of the Digital Imaging Computing: Techniques and Applications*, p. 8. IEEE CS Press, Cairns Convention Centre, Brisbane, Australia, 2005.

[Garland, 1961] Garland LH. *Bronchial carcinomas: lobar distribution of lesions in 250 case*. Calif Me; 94: 7-8. 1961.

[Giger, 1994] Giger, M.L., Bae, K.T., MacMahon, H. *Computerized detection of pulmonary nodules in computed tomography images*. Invest. Radiol. 29, 459–465, 1994.

[Givens, 1985] J.A. Givens Jr, M.R. Gray and K.J. Meller. *A Fuzzy K-Nearest Neighbour Algorithm*. IEEE Trans. Syst., Man, Cybern. July/August, 1985.

[GLOBOCAN2008, 2010] Ferlay J, Shin HR, Bray F, Forman D, Mathers C and Parkin DM. GLOBOCAN 2008, *Cancer Incidence and Mortality Worldwide*. IARC CancerBase No. 10. Lyon, France: International Agency for Research on Cancer. Accesible online via: <http://globocan.iarc.fr>. 2010.

[González, 1996] R.C. González, R.E. Woods. *Tratamiento digital de las imágenes*. Addison-Wesley.1996.

[González, 2004] R.C. González, R.E. Woods, S.L: Eddins. *Digital Image Processing*. Prentice Hall.2004.

[Gordon, 1997] I. Gordon. *Theories of Visual Perception*. John Wiley and Son, 1997.

[**Gori, 2006**] Gori, I., Bellotti, R., Cerello, P., et al. *Lung nodule detection in screening computed tomography*. IEEE Nuclear Science Symposium Conference Record. IEEE, 2006.

[**Gurney, 1993**] Gurney JW. *Determining the likelihood of malignancy in solitary pulmonary nodules with Bayesian análisis*. I. Theory. Radiology 186, 405-13. 1993.

[**Hahn 2007**] Hahn, B., Valentine, D. *Essential MATLAB for Scientists and Engineers*, 3rd Edition. Newnes, 2007.

[**Hartigan 2001**] Hartigan, J. A. *Statistical Clustering. International Encyclopaedia of the Social and Behavioral Sciences*, pp. 15014-15019, Elsevier Ltd., 2001.

[**Hnasell,2008**] Hnasell D.M, Bankier A. A, MacMahon H, *Fleischner Society: Glossary of Terms for Thoracic Imaging*. Radiology: Volume 246: Number 3—March 2008.

[**Hu, 2001**] S. Hu, E. A. Hoffman, and J. M. Reinhardt, *Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images*, IEEE Trans. Med. Imag., vol. 20, no. 6, pp. 490–498, Jun. 2001.

[**ISCIII, 2009**] *Monitorización de la situación del cáncer en España*. Instituto Carlos III de Madrid, accesible online vía <http://www.isciii.es/>. 2010.

[**Itai, 2007**] Itai, Y., Hyoungeop, K., Ishida, T., et al. *A segmentation method of lung areas by using snakes and automatic detection of abnormal shadow on the areas*. Int. J. Innov. Comput. Inf. Control 3, 277– 284, 2007.

[**Kanazawa, 1998**] Kanazawa, K., Kawata, Y., Niki, Y., et al. *Computer-aided diagnosis for pulmonary nodules based on helical CT images*. Comput. Med. Imaging Graph. 22, 157–167, 1998.

[**Kanungo, 1995**] T. Kanungo, M.Y. Jaisimha, J. Palmer and R.M. Haralick. *A Methodology for Quantitative Performance Evaluation of Detection Algorithms*. IEEE Transactions on Image Processing, vol. 4, no .12, pp. 1667-1674. 1995.

[**Kaufman 1990**] Kaufman, L., Rousseeuw, P. J. *Finding Groups in data: An introduction to Cluster Analysis*. John Wiley & Sons, 1990.

[**Kawata, 1998**] Kawata, Y., Niki, N., Ohmatsu, H., et al. *Quantitative surface characterization of pulmonary nodules based on thin-section CT images*. IEEE Trans. Nuclear Sci. 45, 2132–2138, 1998.

[**Kawata, 2000**] Kawata, Y., Niki, N., Ohmatsu, H., et al. *Hybrid classification approach of malignant and benign pulmonary nodules based on topological and histogram features*. Medical Image Computing and Computer-assisted Intervention—MICCAI 2000, vol. 1935/2000, pp. 297–306. LNCS Springer Berlin/Heidelberg, Berlin, 2000.

[**Kim, 2003**] Kim, D.Y., Kim, J.H., Noh, S.M., et al. *Pulmonary nodule detection using chest CT images*. Acta. Radiol. 44, 252–257, 2003.

[**Kim, 2007**] Kim, H., Nakashima, T., Itai, Y., et al. *Automatic detection of ground glass opacity from the thoracic MDCT images by using density features*. In: International Conference on Control, Automation and Systems, pp. 1274–1277. IEEE Xplore, COEX, Seoul, Korea, 2007.

[**Kostis, 2003**] Kostis, W.J., Reeves, A.P., Yankelevitz, D.F., et al. *Three dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images*. IEEE Trans. Medical Imaging 22, 1259–1274, 2003.

[**Krupinski, 1996**] E.A. Krupinski. *Visual scanning patterns of radiologists searching mammograms*. Academic Radiology, 3(2):137.144, 1996.

[**Landis, 1977**] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. N° 33: 159-174. 1977.

[**Leader, 2003**] J. K. Leader, B. Zheng, R. M. Rogers, F. C. Scieurba, A. Perez, B. E Chapman, S. Patel, C. R. Fuhrman, and D. Gur, *Automated lung segmentation in X-ray computed tomography*, Acad. Radiol., vol. 10, n. 11, pp. 1224–1236, 2003.

[**Lee,1990a**] S.U. Lee and Y.M. Lim. *On the Color Image Segmentation Algorithm Based on the Treshholding and the Fuzzy c-Means Techniques*. Pattern Recognition. 1990.

[**Lee,1990b**] S. U. Lee, S.Y. Chung and R.H. Park. *A comparative performance study of several global thresholding techniques for segmentation*. Computer Vision, Graphics and Image Processing, 52, pp 171-190, 1990.

[**Lee, 2001**] Lee, Y., Hara, T., Fujita, H., et al. *Automated detection of pulmonary nodules in helical CT images based on an improved template- matching technique*. IEEE Trans. Med. Imaging 20, 595– 604, 2001.

[**Li, 2000**] P. Li, BD. Corner, and S. Paquette. *Segmenting 3D surface scan data of the human body by 2D projection*. Proc. SPIE, vol. 3985, pp 172-177, 2000.

[**Li B., 2001**] B. Li and J. M. Reinhardt, *Automatic generation of 3-D shape models and their application to tomographic image segmentation*, Proc. SPIE (Medical Imaging), vol. 4322, pp. 311–322, 2001.

[**Li, 2004**] Li, Q., Doi, K. *New selective enhancement filter and its application for significant improvement of nodule detection on computed tomography*. In: Proc. of SPIE, vol. 5370, pp. 1–9, 2004.

[**Li, 2007**] Li Q . *Recent progress in computer-aided diagnosis of lung nodules on thin-section CT*. Computerized Medical Imaging and Graphics, vol 31, pp.248-257. 2007.

[**Li, 2008**] Li, Q., Li, F., Doi, K. *Computerized detection of lung nodules in thin-section CT images by use of selective enhancement filters and an automated rule-based classifier*. Acad. Radiol. 15, 165–175, 2008.

[**Libby, 2004**] Libby DM, Smith JP, Altorki NK, Pasmantier MW, Henschke CL. *Managing the small pulmonary nodule discovered by CT*. CHEST. 125(4), pp1522-1529. 2004.

[**Marcum, 1947**] *A Statistical Theory of Target Detection by Pulsed Radar*. Disponible on line en [http://www.rand.org/pubs/research\\_memoranda/RM754.html](http://www.rand.org/pubs/research_memoranda/RM754.html). 2010.

[**Matlab, 2010**] *The MathWorks: Matlab and Simulink for Technical Computing*. Accesible online vía <http://www.mathworks.com/>, 2011.

[**McQueen, 1967**] J.B. McQueen, *Some methods for classification and analysis of multivariate observation*". 5th Berkeley Symposium on Mathematical Statistics and Probability, pp. 281-297, Berkeley, 1967.

[**Mediantechologies, 2011**] *Median Technologies*, accesible on line vía <http://www.mediantechologies.com>, 2011.

[**Medspain, 2011**] *Imágenes*, accesible on line vía <http://www.medspain.com>, 2011.

[**Metz, 1985**] Metz C.E.. *Evaluation of image quality by ROC analysis: concepts, techniques, and future possibilities*. Japanese Journal of Radiological Technology, vol. 4, pp. 990-1002. 1985.

[**Metz, 1986a**] Metz C.E. *ROC methodology in radiologic imaging*. Investigative Radiology, vol. 21, pp. 720-733. 1986.

[**Metz, 1986b**] Metz C.E. *Statistical Analysis of ROC Data in Evaluating Diagnostic Performance*. Multiple Regression Analysis: Applications in the Health Sciences, number 13, edited by Donald E. Herbert and Raymond H. Myers. 365–384. American Institute of Physics, 1986.

[**Metz, 1989**] Metz C.E. *Some practical issues of experimental design and data analysis in radiological ROC studies*. Investigative Radiology, vol. 24, pp. 234-245, 1989.

[**Metz, 1992**] Metz C.E. *Evaluation of medical images*. In: *The Formation, Handling, and Evaluation of Medical Images* (AE Todd Pokropek and MA Viergever, eds.). Berlin, Springer-Verlag, pps. 277-302, 1992.

[**Metz, 1996**] Metz C.E. *Evaluation of digital mammography by ROC analysis*. Digital Mammography '96., vol. 1119 of International Congress Series., pp. 61-68. June 9-12. 1996.

[**Meyers, 1964**] P. Meyers, C. Nice, Jr., H. Becker, W. Nettleton, J. Sweeney, and G. Meckstroth. *Automated computer analysis of radiographic images*. Radiology, vol. 83, pp. 1029-1034. 1964.

[**NLSTR, 2011**]. *The National Lung Screening Trial: Overview and Study Design*. Radiology: Vol. 258: n 1 pp. 243-253. 2011.

[**Okada, 2005**] Okada, K., Comaniciu, D., Krishnan, A. *Robust anisotropic Gaussian fitting for volumetric characterization of pulmonary nodules in multislice ct*. IEEE Trans. Med. Imaging 24, 409– 423, 2005.

[**Pagano, 2001**] Pagano M. and Gauvreau K. *Fundamentos de Bioestadística*. Thomson Learnig 2001.

[**Paik, 2004**] Paik, D.S., Beaulieu, C.F., Rubin, G.D., et al. *Surface normal overlap: a computer-aided detection algorithm with application to colonic polyps and lung nodules in helical CT*. IEEE Trans. Med. Imaging 23, 661–675, 2004.

[**Pepe, 2003**] Pepe M.S. *The statistical evaluation of medical tests for classification and prediction*. New York: Oxford. 2003.

[**Pérez, 2007**] Pérez I., León B. *Lógica difusa para principiantes. Teoría y práctica*. Publicaciones UCAB. 2007.

[**Pham, 1999**] Pham D.L., Prince J. *An adaptive algorithm for image segmentation in the presence of intensity inhomogeneities*. Pattern Recognition Letters, 20, pp. 57-68, 1999.

[**Pu, 2008**] Pu, J., Roos, J., Yi, C.A., et al. *Adaptive border marching algorithm: automatic lung segmentation on chest CT images*. Comput. Med. Imaging Graph. 32, 452–462, 2008.

[**Retico, 2008**] Retico, A., Delogu, P., Fantacci, M.E., et al. *Lung nodule detection in low-dose and thin-slice computed tomography*. Comput. Biol. Med. 38, 525–534, 2008.

[**ROCCH, 2009**] *ROC analysis and the ROC Convex Hull* accessible online via [http://home.comcast.net/~tom.fawcett/public\\_html/ROCCH](http://home.comcast.net/~tom.fawcett/public_html/ROCCH), 2009.

[**Rubin, 2000**] G. D. Rubin, *Data explosion: the challenge of multidetector-row CT*, Eur. J. Radiol., vol. 36, pp. 74–80, 2000.

[**Ruspini, 1970**] Ruspini Enrique H. . *Numerical methods for fuzzy clustering*. Information Sciences. Vol. 2, Issue 3 Pages 319-350. 1970,

[**SAIS, 2011**] *Sociedad Argentina de Informática y Salud*. Accesible on line <http://www.sais.org.ar>. 2011.

[**SEOM, 2010**] *Sociedad Española de Oncología Médica*, accesible on line vía <http://www.seom.org/en/infopublico>. 2010.

- [**Silva, 2001**] A. Silva, J. S. Silva, B. S. Santos, and C. Ferreira, *Fast pulmonary contour extraction in X-ray CT images: a methodology and quality assessment*, Proc. SPIE, vol. 4321, pp. 216–224, 2001.
- [**Sluimer, 2003**] Sluimer, I.C., van Waes, P.F., Vierever, M.A., et al. *Computer- aided diagnosis in high resolution CT of the lungs*. Med. Phys. 30, 3081–3090, 2003.
- [**Sluimer, 2006**] Sluimer, I.C., et al. *Computer Analysis of Computed Tomography Scans of the Lung: A Survey*. IEEE Transactions on Medical Imaging. Vol. 25.n.4. 2006.
- [**Sonka, 2000**] M. Sonka, J.M. Fitzpatrick. *Handbook of Medical Imaging*. Volume 2. Medical Image Processing and Analysis. SPIE. 2000.
- [**Spackman, 1989**] Spackman, K. A, *Signal detection theory: Valuable tools for evaluating inductive learning*. In: Proceedings of the Sixth International Workshop on Machine Learning. San Mateo, CA, pp. 160-163, Morgan Kaufman. 1989.
- [**Suarez, 2009a**] Suárez-Cuenca J.:*Desarrollo de un sistema de diagnóstico asistido por computador para detección de nódulos pulmonares en tomografía computarizada multicorte*. Tesis Doctoral, USC. 2009
- [**Suarez, 2009b**] Suárez-Cuenca J, Tahoces PG, Souto M, Lado MJ, Remy-Jardin M, Remy J, Vidal JJ: *Application of the iris filter for automatic detection of pulmonary nodules on computed tomography images*. Computers in Biology and Medicine 39: pp.921-933. 2009
- [**Sun, 2007**] Sun, S.-S., Li, H., Hou, X.-R., et al. *Automatic segmentation of pulmonary nodules in CT images*. In: 1st International Conference on Bioinformatics and Biomedical Engineering, ICBBE., pp. 790–793. IEEE, 2007.
- [**Swets, 1979**] J.A. Swets. *ROC analysis applied to the evaluation of medical imaging techniques*. Investigative Radiology, vol. 14, pp. 109-121, 1979.
- [**Torres, 2010**] Torres A. *Curvas ROC para datos de supervivencia. Aplicación a Datos Biomédicos*. Proyecto Master, USC. 2010.
- [**Treisman, 1985**] Treisman A. *Pre-attentive processing in vision*. Computer Vision, Graphics and Image Processing, 31:156.177, 1985.
- [**Tryon, 1939**] Tryon, R. C. *Cluster analysis*. McGraw-Hill, 1939.
- [**Tsid, 2011**] *TSID*. Accesible on line via: [www.tsid.net](http://www.tsid.net), 2011.
- [**Udupa, 2002**] Udupa, J.K., Saha, P.K., Lotufo, R.A. *Relative fuzzy connectedness and object definition: theory, algorithms, and applications in image segmentation*. IEEE Trans. Pattern Anal. Mach. Intell. 24, 1485–1500, 2002.

[**van Ginneken, 2010**] van Ginneken et al. *Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: the ANODE09 study*. Medical Image Analysis vol. 14 pp. 707- 722, 2010.

[**van Ginneken, 2001**] van Ginneken B., Haar B.M., and M.A. Viergever, *Computer-Aided Diagnosis in Chest Radiography: A Survey*. IEEE Transactions on Medical Imaging, IEEE Press, 20 (12), pp. 1228-12241, 2001.

[**Wang, 2007**] Wang, P., DeNuzio, A., Okunieff, P., et al. *Lung metastases detection in CT images using 3D template matching*. Med. Phys. 34, 915–922, 2007.

[**Winer\_Muram, 2006**] Winer\_Muram HT, *The Solitary Pulmonary Nodule*. Radiology, 239(1), pp34-49. 2006.

[**Winsberg, 1967**] F. Winsberg, M. Elkin, J. Macy, V. Bordaz, and W. Weymouth. *Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis*. Radiology, vol. 89, pp. 211-215. 1967.

[**Yeong, 2007**] Jeong YJ, Yi CA, Lee KS, *Solitary Pulmonary Nodules: Detection Characterization, and Guidance for Further Diagnostic Workup and Treatment*. AJR 2, 188, pp 57-68. 2007.

[**Zadeh, 1965**] Zadeh L.A. *Fuzzy Sets*. Information and Control, n 8 pp. 338-353. 1965.

[**Zhang, 2003**] D.-Q Zhang., S.-C Chen.. A novel kernelized fuzzy C-means algorithm with application in medical image segmentation. Artificial Intelligence in Medicine 32, 37 – 50 pp. 1224–1236, 2003.

[**Zhang, 1992**] Zhang Y.J. and Gerbrands JJ. *Segmentation evaluation using ultimate measurement accuracy*, SPIE 1657, 449-460, 1992.

[**Zhang, 1994**] Zhang Y.J. and Gerbrands JJ. *Objective and quantitative segmentation evaluation and comparison*, Signal Processing, 39 :43-54, 1994.

[**Zhang, 1996**] Zhang Y.J. *A survey on evaluation methods for image segmentation*. Pattern Recognition, vol. 29, no. 8, pp. 1335-1346. 1996.

[**Zhang, 1997**] Zhang Y.J. *Evaluation and comparison of different segmentation algorithms*. Pattern Recognition Letters, 18, pp. 963-974. 1997.

[**Zhang, 2004a**] Zhang D.Q, Leung Y.W. *Improved possibilistic C-means clustering algorithms* IEEE Transactions on Fuzzy Systems. Vol.12, n2 pp 209-217.2004

[**Zhang, 2004b**] Zhang D.Q., Chen S.C. *A novel kernelized fuzzy C-means algorithm with application in medical image segmentation*. Artificial Intelligence in Medicine 32, 37 – 50, 2004.

