



Universidad de las Ciencias Informáticas

DIRECCIÓN DE FORMACIÓN POSTGRADUADA

Tesis presentada en opción al título de
Máster en Informática Aplicada

“Algoritmo basado en técnicas de segmentación de imágenes de tomografía de tórax para aumentar el acierto en la identificación de nódulos pulmonares solitarios”

Autora: Ing. Arelys Rivero Castro

Tutor: MSc. Hector Raúl González Díez

La Habana, X de diciembre de 2014

“Año 56 de la Revolución”

DECLARACIÓN DE AUTORÍA

Yo Arelys Rivero Castro, con carné de identidad 86121511374, declaro que soy la autora principal del resultado que expongo en la presente memoria titulada “Algoritmo basado en técnicas de segmentación de imágenes de tomografía de tórax para aumentar el acierto en la identificación de nódulos pulmonares solitarios”, para optar por el título de Máster en Informática Aplicada. Autorizo a la Universidad de las Ciencias Informáticas a hacer uso de la misma en su beneficio, así como los derechos patrimoniales con carácter exclusivo.

Para que así conste, firmo la presente declaración jurada de autoría en Ciudad de La Habana a los X días del mes de diciembre del año 2014.

Firma del Maestrante

AGRADECIMIENTOS

Especialmente a Yoel por pelearme, por ser compañero y amigo, por su apoyo y su comprensión.

A Michi y Renier por ser hermanos y amigos en todo momento, por su cariño y su incondicionalidad.

A mi familia, a Jean Carlos por regalar una sonrisa cuando más la necesito y por sus ocurrencias. A mi hermana, a mami y a papi por sus innumerables consejos.

A mis abuelos por su ejemplo, los que están y los que ya se marcharon.

A la gran familia del 41 103 por el apoyo y el carisma (los Elos y los P...).

Agradezco a todos los que dieron su aporte en el desarrollo de la investigación, Yosvani, Yoanny, Freddy.

RESUMEN

Identificar el cáncer de pulmón en fases tempranas, ha sido en los últimos años una tarea priorizada por la comunidad científica. Como alternativa para minimizar el impacto de esta enfermedad se han desarrollado algoritmos para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax; que proporcionan a los especialistas, una segunda opinión en la interpretación de los resultados, con altos valores de sensibilidad pero con un alto número de falsos positivos por imágenes. En esta investigación se exponen los resultados obtenidos al desarrollar un algoritmo, utilizando métodos de segmentación de imágenes, para la identificación de nódulos pulmonares solitarios, que aumenta el acierto en la identificación de dichas estructuras. Se utilizó durante la fase de segmentación la combinación de los filtros Máxima Densidad Local e Iris, que se enfocan en las características morfológico-radiográficas de los nódulos pulmonares solitarios. Los resultados arrojados por el algoritmo desarrollado, fueron contrastados con las anotaciones realizadas en imágenes publicadas en The Lung Image Database Consortium Image Collection (LIDC-IDRI) y se obtuvo un 79% de acierto en la detección de nódulos pulmonares solitarios, con un promedio de 5,02 falsos positivos por imagen, para estructuras de 3-30 mm de diámetro.

PALABRAS CLAVE: algoritmos de identificación, nódulos pulmonares solitarios, procesamiento de imágenes médicas, segmentación de imágenes, tomografía de tórax

ÍNDICE

INTRODUCCIÓN	6
CAPÍTULO 1. FUNDAMENTOS TEÓRICOS METODOLÓGICOS DE LAS TÉCNICAS UTILIZADAS PARA LA IDENTIFICACIÓN DE NÓDULOS PULMONARES SOLITARIOS EN IMÁGENES DE TOMOGRAFÍA DE TÓRAX	13
1.1. CARACTERÍSTICAS MORFOLÓGICO-RADIOGRÁFICAS DE LOS NÓDULOS PULMONARES SOLITARIOS ----	13
1.2. SISTEMAS DE DIAGNÓSTICO ASISTIDOS POR ORDENADOR (CAD)	15
1.3. TENDENCIAS EN EL DESARROLLO DE SISTEMAS CAD PARA LA DETECCIÓN DE NÓDULOS PULMONARES EN IMÁGENES DE TOMOGRAFÍA	17
1.4. NÓDULOS PERIFÉRICOS	21
1.5. ALGORITMOS DE PROCESAMIENTO PARA LA SEGMENTACIÓN DE IMÁGENES	22
1.5.1. <i>Segmentación basada en umbralización</i>	23
1.5.2. <i>Segmentación orientada a regiones</i>	23
1.5.3. <i>Segmentación basada en transformaciones morfológicas</i>	24
1.5.4. <i>Segmentación basada en bordes</i>	24
1.6. CONCLUSIONES DEL ESTADO DE ARTE	25
1.7. CONCLUSIONES DEL CAPÍTULO	26
CAPÍTULO 2. ALGORITMO PARA LA CLASIFICACIÓN DE NÓDULOS PULMONARES SOLITARIOS EN IMÁGENES DE TOMOGRAFÍA DE TÓRAX	28
2.1. PROPUESTA DE SOLUCIÓN	28
2.2. FUNCIONAMIENTO DEL FILTRO LOCAL DENSITY MAXIMUM (LDM)	32
2.3. DESCRIPCIÓN DEL FILTRO DE IRIS	35
2.4. DESCRIPCIÓN DE LA SECUENCIA DE PASOS DESARROLLADOS POR EL ALGORITMO PARA IDENTIFICAR LOS NÓDULOS PULMONARES SOLITARIOS	37
2.5. CONCLUSIONES DEL CAPÍTULO	39
CAPÍTULO 3. ANÁLISIS Y VALORACIÓN DE LOS RESULTADOS OBTENIDOS POR EL ALGORITMO PROPUESTO PARA LA IDENTIFICACIÓN DE NÓDULOS PULMONARES SOLITARIOS EN IMÁGENES DE TOMOGRAFÍA DE TÓRAX	41

3.1. INDICADORES ESTADÍSTICOS BÁSICOS PARA EVALUAR EL DESEMPEÑO DE UN PROCEDIMIENTO DIAGNÓSTICO -----	41
3.2. AUTOMATIZACIÓN DEL CÁLCULO DE LOS INDICADORES DE VALIDEZ DEL ALGORITMO PARA LA IDENTIFICACIÓN DE NÓDULOS PULMONARES SOLITARIOS -----	44
3.3. ANÁLISIS DE LOS RESULTADOS OBTENIDOS-----	46
3.4. CONCLUSIONES DEL CAPÍTULO -----	48
CONCLUSIONES -----	49
RECOMENDACIONES-----	50
REFERENCIAS BIBLIOGRÁFICAS-----	51
ANEXOS-----	59

INTRODUCCIÓN

La utilización de las tecnologías de la información y las comunicaciones en el sector de la salud ha producido transformaciones en sus procesos; propiciando entornos de trabajo más organizados, con un rápido acceso a la información, facilidades para el intercambio de conocimiento, la formación y la colaboración médica. La creación de técnicas imagenológicas como la ecografía (US, por sus siglas en inglés), la tomografía computarizada (CT, por sus siglas en inglés) (1, 2), la resonancia magnética (MR, por sus siglas en inglés) o la medicina nuclear (MN, por sus siglas en inglés), situaron la Radiología, a la cabecera de las especialidades médicas. La adopción de estas técnicas imagenológicas, ha marcado un punto crucial en el desarrollo de las ciencias médicas.

Actualmente se utilizan novedosos equipos con altas prestaciones de hardware y software que unidos a un conjunto de algoritmos y técnicas de procesamiento de imágenes digitales (3), ayudan al personal médico a brindar mejores y más rápidos servicios. Para la adquisición de las imágenes digitales se utilizan variedades de técnicas basadas en los rayos X, la resonancia magnética nuclear, emisiones radioisotópicas, entre otras. Las tecnologías existentes posibilitan la recogida de gran cantidad de datos referentes a las diferentes estructuras anatómicas, permitiendo su reproducción en medios visuales.

La imagen digital con respecto a las técnicas de radiografía convencional, presupone dos características relevantes que magnifican su utilidad; la resolución espacial (píxel por pulgada o centímetro) y la densidad o profundidad (niveles de grises que se pueden representar). Los departamentos de diagnóstico por imágenes se han beneficiado con la adopción de la imagen digital como técnica principal para el diagnóstico, algunos de los aportes que se evidencian son:

- ✓ Reducción considerable del consumo de recursos materiales por concepto de impresión de placas radiográficas.
- ✓ Reducción del tiempo de realización de los exámenes médicos y por tanto la atención de un mayor número de pacientes en los servicios de diagnóstico por imágenes.
- ✓ Visualización y análisis detallado de estructuras y lesiones al utilizar algoritmos computacionales para el procesamiento de las imágenes y la realización de mediciones de estructuras de interés.
- ✓ Facilidad de almacenamiento de los estudios a largo plazo, lo que posibilita la realización de estudios de comparación.
- ✓ Rapidez en la ejecución de consultas de segunda opinión entre especialistas geográficamente distantes mediante el uso de la telemedicina.

Por sus principios creacionales algunas modalidades como la CT y la MR, son especialmente útiles para la diferenciación entre un tipo de tejido y otro, por las diversas formas en que el tejido libera la energía absorbida durante el proceso de realización del estudio médico. Este elemento posibilita la identificación de tejido patológico (2) y por tanto de tejido cancerígeno en las diferentes estructuras anatómicas. El diagnóstico por imágenes médicas es una de las formas más rápidas y efectivas para la detección del cáncer de pulmón. La CT posibilita la detección de un alto porcentaje de nódulos pulmonares con mayor precisión que otras técnicas imagenológicas (4, 5), debido a su calidad, determinada por su alta resolución y las especificidades que presupone la utilización de los rayos X.

La importancia de identificar correctamente este tipo de tejido cancerígeno parte de la necesidad de realizar un diagnóstico temprano, con la intención de reducir el alto índice de mortalidad que provocan los tumores malignos. La mayoría de los pacientes a los que se les diagnostica cáncer de pulmón en su fase inicial se les detecta casuísticamente, acuden a la consulta médica a causa de otra afección. La dificultad en el diagnóstico de esta enfermedad radica en que los síntomas suelen ser tardíos, y por ello hasta el 80% de los pacientes se presentan en estadios no resecables¹ (6).

El cáncer de pulmón es uno de los que tiene mayor incidencia en el ser humano, responsable de los mayores índices de mortalidad oncológica a nivel mundial. Es la primera causa de mortalidad por cáncer en el varón y la tercera después del de colon y mama en la mujer. (7) Se calcula que alrededor del 90% de los pacientes con cáncer de pulmón mueren a causa de esta enfermedad (8).

Según el Anuario Estadístico de Salud emitido por el Ministerio de Salud Pública de Cuba (MINSAP) (9) en el año 2014, los tumores malignos representan la primera causa de muerte en la isla. En el año 2013 el número de defunciones por tumores malignos fue de 22 868, de estas, 5 212 están asociados a tráquea, bronquios y pulmón, lo que representa una tasa de 46,7 por cada 100 000 habitantes; comportándose como el grupo de mayor prevalencia tanto en hombres como mujeres. En la Tabla 1 se muestra el número de defunciones por causa de cáncer de tráquea, bronquios y pulmón desglosadas en hombre y mujeres.

Tabla 1: Tasa de defunciones en Cuba debido a tumores malignos en tráquea, bronquios y pulmón (9)

Grupos	Defunciones por cáncer de TBP	Tasa por cada 100 000 habitantes
Hombres	3 296	59.2
Mujeres	1 916	34.3

¹ Que se puede extraer por medio de cirugía.

Por el impacto social que representa el cáncer de pulmón en el mundo, la comunidad científica internacional ha mostrado interés en la detección de dicha enfermedad desde sus fases más tempranas, a partir de la utilización de las técnicas de diagnóstico por imágenes médicas. Los más novedosos equipos de adquisición de imágenes médicas, generan un elevado volumen de imágenes y realizan cada vez, cortes de menor grosor, lo que facilita que el especialista cuente con valiosa información para emitir su diagnóstico.

A pesar de las ventajas claramente evidenciadas por el uso de esta tecnología, el proceso de diagnóstico utilizando las imágenes médicas es engorroso. La revisión de mucha información produce cansancio y falta de concentración en los especialistas. La interpretación de las imágenes es muy subjetiva y está determinada en gran medida por el nivel de experticia que posee el imagenólogo. Esta experiencia se adquiere con los años de trabajo y es relevante para alcanzar el mayor número de aciertos posibles, durante el proceso de detección de los nódulos pulmonares y evitar la realización de biopsias y procedimientos molestos o invasivos al paciente. En ocasiones se requiere de consultas de segunda opinión para reforzar el criterio del especialista que realiza el diagnóstico, ralentizando el proceso de atención al paciente.

Para apoyar la toma de decisiones durante el proceso de diagnóstico por imágenes médicas, se han desarrollado a nivel internacional sistemas de Diagnóstico Asistido por Ordenador (CAD, por sus siglas en inglés) (10–12), los cuales brindan al imagenólogo una segunda opinión sobre el caso analizado. Diversos autores han publicado trabajos donde se evidencian las ventajas al integrar sistemas CAD a sistemas de adquisición, almacenamiento, transmisión y visualización de imágenes médicas (PACS, por sus siglas en inglés) (13, 14); entre las que sobresale una mejoría en la capacidad diagnóstica de los especialistas y la reducción del tiempo de revisión de los estudios médicos. Teniendo en cuenta estos criterios se han desarrollado un conjunto de investigaciones que se enfocan en la detección de nódulos pulmonares solitarios, utilizando imágenes de tomografía de tórax (15–18). Dichas soluciones se enmarcan fundamentalmente en tres fases: segmentación, reconocimiento de formas y clasificación (19). Puntualmente en la fase de segmentación de las estructuras nodulares existen inconvenientes que dificultan el proceso de identificación:

- ✓ Algunos nódulos pulmonares son de muy pequeño tamaño, lo que propicia que sean confundidos con estructuras ruidosas; esto propicia que puedan ser eliminados al utilizar alguna técnica de preprocesamiento en la imagen o al realizar reducción de falsos positivos.

- ✓ Las características de algunos nódulos no son completamente representativas, en ocasiones presentan patrones de calcificación, textura o forma que complejizan su identificación.
- ✓ Poseen características morfológico-radiográficas similares a otras estructuras como los bronquios y los vasos sanguíneos que también se encuentran dentro de la región pulmonar.

En algunos trabajos publicados por diversos autores se evidencian estas deficiencias, tal es el caso de (20, 21), donde no es posible diferenciar los nódulos pulmonares solitarios de otras estructuras contenidas en la región pulmonar, solo identificando estructuras de alto valor de intensidad. En (22) a pesar de obtener altos valores de sensibilidad es engorroso crear estructuras que puedan simular de forma precisa los patrones de forma, calcificación y textura de los nódulos pulmonares solitarios, lo que trae consigo la detección de un elevado número de falsos positivos.

Por la complejidad que presupone la identificación de las estructuras nodulares, la mayoría de las investigaciones se enfocan en alcanzar los mayores índices de sensibilidad en la fase de segmentación. Generalmente la reducción de los falsos positivos es una tarea delegada a la fase de clasificación; por lo que al concluir la etapa de identificación de los nódulos pulmonares solitarios, no se obtienen favorables valores de acierto, al alcanzar elevados valores de sensibilidad, pero también un alto número de falsos positivos por imagen médica analizada.

Después de analizar la situación antes expuesta, se identifica como **problema científico**: ¿Cómo aumentar el acierto en la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax?

Se define como **objeto de estudio**: proceso de identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax.

Para dar solución a la problemática antes mencionada se ha definido como **objetivo general**: desarrollar un algoritmo basado en técnicas de segmentación de imágenes de tomografía de tórax, enfocadas en sus características morfológico-radiográficas, para aumentar el acierto en la identificación de nódulos pulmonares solitarios.

El **campo de acción** se enfoca en las técnicas de segmentación como parte del proceso de identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax.

Para guiar la investigación y comprobar los resultados se propone la siguiente **hipótesis**: si se utiliza un algoritmo basado en técnicas de segmentación de imágenes de tomografía de tórax enfocadas en sus

características morfológico-radiográficas, aumentará el acierto en la identificación de nódulos pulmonares solitarios.

Objetivos específicos:

1. Establecer los fundamentos teórico-metodológicos para el desarrollo de un algoritmo de identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax.
2. Desarrollar métodos de segmentación y reconocimiento de formas para seleccionar las regiones de interés en las imágenes de tomografía de tórax.
3. Validar los resultados alcanzados por el algoritmo de identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax, contrastándolos con los obtenidos por algoritmos similares descritos durante el estudio del estado del arte.

Novedad de la investigación

Como resultado de la investigación se obtendrá un algoritmo capaz de identificar y señalar al radiólogo las estructuras nodulares a partir del análisis del contenido de las imágenes de tomografía de tórax. Al utilizar la combinación de los filtros Máxima Densidad Local e Iris durante el proceso de segmentación e identificación de las estructuras nodulares, se pueden obtener altos valores de acierto, a partir de una elevada sensibilidad y un bajo número de falsos positivos por imagen médica analizada.

Aporte práctico de la investigación

El algoritmo propuesto, al integrarse a la solución PACS desarrollada en el Centro de Informática Médica (CESIM) de la Universidad de las Ciencias Informáticas (UCI), pudiera servir de apoyo a los imagenólogos durante el proceso de toma de decisiones para la emisión de un diagnóstico certero, al identificar nódulos pulmonares solitarios en imágenes de tomografía de tórax.

Impacto de la propuesta

La propuesta realizada en la presente investigación puede contribuir significativamente en:

- ✓ La reducción del tiempo de lectura de los estudios médicos influyendo directamente en el proceso de atención al paciente.
- ✓ La formación de nuevos especialistas en el área de la Imagenología.
- ✓ La comunidad científica contará con un nuevo algoritmo que podrá ser contrastado con los desarrollados por otros investigadores.

Entre los métodos utilizados en esta investigación se destacan los siguientes:

Teóricos:

Histórico Lógico: al realizar un análisis crítico valorativo de la información contenida en las fuentes bibliográficas consultadas, con el objetivo de conocer cuáles son las tendencias y el comportamiento del desarrollo de software, en torno a la utilización de las técnicas de segmentación de imágenes para identificar nódulos pulmonares solitarios en las imágenes de tomografía de tórax. Como parte inicial de la investigación se realiza un estudio del estado del arte de la problemática planteada. Se analizan diversas soluciones existentes, se identifican fuentes de datos confiables así como los algoritmos de procesamiento y segmentación de imágenes más utilizados.

Analítico Sintético: para descomponer el problema de la investigación en elementos concretos de la solución y así profundizar en las particularidades del proceso de segmentación, llevado a cabo en las imágenes de tomografía de tórax. Luego se realiza una síntesis en la solución propuesta.

Inductivo Deductivo: teniendo en cuenta el funcionamiento del proceso de diagnóstico por imágenes médicas se evalúa la problemática existente, para determinar aspectos particulares y desarrollar una propuesta de solución.

Hipotético Deductivo: para el análisis de la información derivada del uso de otros métodos y de los componentes que conforman el objeto de investigación. Se utiliza en la definición de la hipótesis y para arribar a conclusiones mediante el uso de procedimientos deductivos. Se realiza el planteamiento de objetivos específicos, los cuales son resueltos durante el desarrollo de la investigación.

Modelación: con el fin de crear abstracciones e interpretar la realidad. Se utiliza al crear modelos de funcionamiento de los flujos de trabajos enmarcados en el área de radiología y de las técnicas utilizadas en el desarrollo de la investigación.

Empíricos:

Experimento: con el objetivo de identificar cuáles son las combinaciones más efectivas de los parámetros utilizados por el filtro LDM para alcanzan el mejor índice de sensibilidad en la identificación de los nódulos pulmonares solitarios.

Observación: al analizar los valores resultantes de la ejecución del algoritmo propuesto y corroborar que estos se corresponden con los resultados esperados; los cuáles se encuentran en el fichero XML de cada serie analizada, pública en The lung Image Database Consortium.

El documento está estructurado en tres capítulos, cuyo contenido se describe a continuación:

Capítulo 1. Fundamentos teórico metodológicos de las técnicas utilizadas para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax: agrupa los principales conceptos vinculados a la investigación. Se evidencia el estudio realizado para analizar las tendencias existentes en torno al desarrollo de los Sistemas de Diagnóstico Asistido por Computador destinados a la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax. Se describen los principales métodos de segmentación de imágenes médicas, a partir de la definición de cuáles de ellos pueden ser relevantes para obtener resultados satisfactorios en la solución del problema planteado en la investigación.

Capítulo 2. Algoritmo para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax: se realiza una descripción detallada de cuáles fueron las técnicas de segmentación de imágenes médicas utilizadas para la identificación de las estructuras candidatas a ser nódulos pulmonares solitarios. Se evidencia el comportamiento que tiene el algoritmo en cada paso de procesamiento. Se describen las técnicas utilizadas y se presentan las características de los principales métodos desarrollados en la fase de identificación de las estructuras nodulares.

Capítulo 3. Análisis y valoración de los resultados obtenidos por el algoritmo propuesto para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax: se describen los indicadores a partir de los cuales se realizará la valoración del algoritmo desarrollado para la identificación de nódulos pulmonares solitarios y las características del estándar de oro utilizado para comprobar los resultados obtenidos. Queda evidenciado el procedimiento llevado a cabo para el cálculo de los indicadores y se especifican los valores de sensibilidad y el promedio de falsos positivos por imágenes obtenidos por el algoritmo propuesto, para dar solución a la problemática. Se muestran los resultados alcanzados por dos de los algoritmos analizados durante el estudio del estado del arte y se comparan con los alcanzados por la propuesta realizada.

CAPÍTULO 1. Fundamentos teórico metodológicos de las técnicas utilizadas para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

En este capítulo se expone el marco conceptual y las principales características radiográficas de los nódulos pulmonares solitarios (NPS); así como la estructura fundamental de los CAD. Se sintetizan los procedimientos llevados a cabo por algunos autores que desarrollan el tema en cuestión, con el objetivo de identificar tendencias, patrones, principales algoritmos y técnicas en la segmentación de imágenes, utilizadas para la identificación de las estructuras nodulares a través de imágenes de tomografía de tórax.

1.1. Características morfológico-radiográficas de los nódulos pulmonares solitarios

El cáncer de pulmón se forma en los tejidos de esta zona, por lo general, en las células que recubren las vías respiratorias. Los dos tipos más importantes de cáncer de pulmón son el de células pequeñas y el de células no pequeñas. (23) El cáncer de pulmón es la causa más frecuente de mortalidad por cáncer en el mundo. (24, 25) La detección temprana de la enfermedad, cuando aún está en la fase de nódulo pulmonar, se ha convertido en un elemento importante, debido a los altos índices de prevalencia y mortalidad que esta representa.

En (26–28) se denomina nódulo pulmonar a cualquier opacidad no lineal única, rodeada por parénquima pulmonar normal, menor de tres centímetros de diámetro, no asociada a adenopatías², atelectasia³ o neumonía. La lesión pulmonar que supera los 3 cm de diámetro se conoce como masa pulmonar.

El diagnóstico por imágenes médicas se ha convertido en los últimos años en una de las técnicas más utilizadas debido a la capacidad de detectar lesiones, incluso en su fase asintomática. Las imágenes de CT son las más utilizadas para la detección de nódulos pulmonares solitarios (NPS) (26, 29–31). La utilización de equipos médicos de alta tecnología facilita la detección y descripción de las estructuras nodulares cuando estas se encuentran en fases tempranas de la enfermedad.

² Se define por la presencia de ganglios linfáticos anormales en número, tamaño o consistencia.

³ Disminución del volumen pulmonar.

CAPÍTULO 1. Fundamentos Teóricos Metodológicos de las técnicas para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

Los NPS poseen un grupo de características desde el punto de vista morfológico-radiológico que posibilitan una aproximación diagnóstica a partir del análisis de las imágenes médicas. Se manifiestan en la imagen de CT como estructuras redondeadas de alta densidad al igual que el cuerpo, la camilla del paciente y algunas estructuras internas del pulmón como los vasos, las venas y los bronquios. Dichos valores de densidad disminuyen desde el centro hacia los bordes de la estructura nodular, por lo que si se realiza un mapa de gradiente del NPS la dirección del vector gradiente es hacia el centro del mismo.

Algunas de las características más significativas que permiten la identificación de los NPS a partir de imágenes de CT son: contorno, localización, calcificación, densidad, tamaño y velocidad de crecimiento. (32–35)

Atendiendo al contorno se pueden clasificar en lisos, espiculados y lobulados. El primero de ellos tiene criterio de benignidad, no siendo este el caso de los dos últimos que tienen criterio de malignidad. Los NPS presentan diversos patrones de calcificación, pueden ser laminares, centrales, sólidos o globales, palomitas de maíz, dispersos y excéntricos (Ver figura 1). Desafortunadamente del 28%-63% de los nódulos benignos son no calcificados lo que complejiza el proceso de identificación de estas estructuras.



Figura 1: Patrones de clasificación de los nódulos pulmonares solitarios en imágenes de CT (5)

Según su localización se clasifican en centrales, periféricos y subpleurales y según la densidad o textura se clasifican en sólidos (sólidos no calcificados), calcificados (sólidos calcificados), vidrios deslustrados (esmerilados). Los NPS más frecuentes son los sólidos. Cerca del 34% de los no sólidos y entre el 40%-50% de los parcialmente sólidos menores de 1,5 mm son malignos, riesgo que aumenta con el tamaño de los nódulos; o en caso de los semisólidos si la parte sólida se localiza en el centro. De los sólidos son malignos solo el 15% de los menores de 1 mm. (28)

De la misma forma en que existe diversidad en las características ya mencionadas, la estructura interna de los NPS puede ser tejido blando, líquido, grasa y aire. Estas especificaciones al igual que la textura puede ser determina por los radiólogo a partir de los valores de Hounsfield (HU, por sus siglas en inglés) (36, 37) correspondientes a cada uno de los tejidos que componen los NPS.

1.2. Sistemas de Diagnóstico Asistidos por Ordenador

Un CAD (10, 38) integra el procesamiento de imágenes, la visión artificial, las matemáticas, la física y las estadísticas, con las técnicas computarizadas que apoyan al radiólogo en el proceso de toma de decisiones. El radiólogo utiliza la salida de este tipo de sistema como una “segunda opinión” y toma la decisión final del caso. Esta área estudia el diseño de sistemas que reconocen patrones en los datos y los usan para clasificar objetos de forma automática y tomar decisiones. Aplicar estas técnicas permite analizar la imagen digitalizada de un modo diferente a como lo haría un observador humano.

Los CAD detectan regiones de interés y las caracterizan de forma cuantitativa, para a través de patrones y algoritmos matemáticos, detectar anomalías y zonas sospechosas. Su objetivo final no es el de realizar un diagnóstico completamente automático, sino ayudar a que el profesional mejore su rendimiento diagnóstico, dirigiendo su atención hacia aquellos puntos con mayor riesgo de existencia de una afección. Con la utilización de sistemas de este tipo el rendimiento del ordenador no tiene que ser comparable o mejor que el de los médicos, sino complementario. La ganancia potencial se debe al efecto sinérgico obtenido mediante la combinación de la competencia del radiólogo y la capacidad del ordenador.

Los sistemas CAD son muy versátiles y se han utilizado para dar solución a múltiples problemas en diversas áreas de la sociedad, entre ellas la medicina. Se han desarrollado numerosas investigaciones y sistemas CAD especializados en patologías muy concretas, de estructuras específicas del cuerpo.

Algunos ejemplos de ellos son los desarrollados para la detección de cáncer de mama (39, 40), cáncer de próstata (41, 42), lesiones de la piel (43, 44) y cáncer de pulmón (5, 30, 45, 46).

Existen sistemas CAD que emplean métodos estadísticos y asumen que las imágenes pueden contener uno o más objetos y que cada objeto pertenece a uno de varios tipos, categorías o clases de patrones. Para poder realizar su trabajo cuentan con una estructura dividida en fases, donde cada fase tiene un objetivo bien definido en el proceso de transformar la imagen médica original en una imagen médica modificada. La metodología de los sistemas CAD es muy similar al de los sistemas de reconocimiento de patrones, en la Figura 2 se muestra la estructura de este tipo de sistemas.

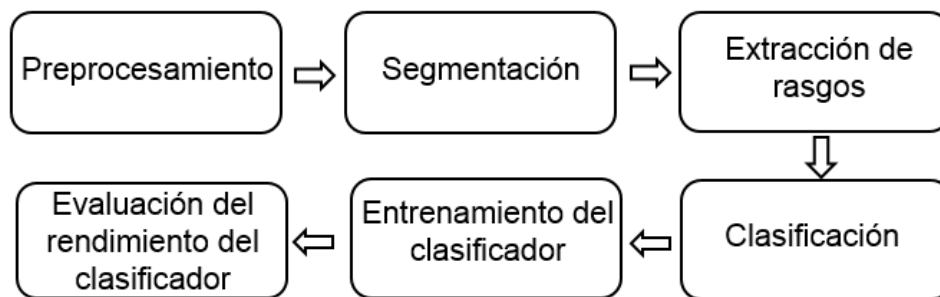


Figura 2: Estructura de un CAD (fuente: elaboración propia)

En la fase de *Preprocesamiento* se le realizan transformaciones a la imagen con el objetivo de eliminar ruido y armonizarla en caso de sufrir diferencias en niveles de exposición en distintos puntos. Se resaltan bordes para que en fases posteriores se puedan delimitar con mayor facilidad las diferentes estructuras representadas en la imagen.

El objetivo de la etapa de *Segmentación*, es dividir la imagen en las partes que la constituyen, hasta un nivel en el cual se aíslen las regiones de interés, para detectar estructuras importantes y definir las como regiones para analizar en fases posteriores.

Una vez que se han vencido las dos primeras fases, se procede con el reconocimiento de patrones, que se basa en la *Extracción de Rasgos* de las regiones previamente segmentadas. Los rasgos descriptivos están determinados por las estructuras que se quieren procesar y están asociados a la forma, tamaño, compacidad, textura, relación con estructuras próximas, entre otras. Al finalizar este proceso se tienen identificadas y caracterizadas las regiones de interés, lo cual posibilita su *Clasificación*.

A partir de las regiones previamente identificadas, se determina cuáles de ellas son las estructuras candidatas y se destacan las mismas para que el radiólogo le preste atención especializada. Es el radiólogo en esta fase, quien decide finalmente, qué es o no relevante para el diagnóstico médico y su clasificación.

Con el objetivo de realizar la identificación de los nódulos pulmonares solitarios en imágenes de tomografía durante el proceso de diagnóstico médico, son de interés para la presente investigación las primeras fases del desarrollo de los sistemas CAD. Se pretende abarcar hasta la fase de Extracción de rasgos en la cual se seleccionan algunas estructuras candidatas a ser nódulos pulmonares teniendo en cuenta sus características morfológico-radiográficas.

1.3. Tendencias en el desarrollo de sistemas CAD para la detección de nódulos pulmonares en imágenes de tomografía

Con el objetivo de identificar estructuras nodulares a partir de la utilización de imágenes de tomografía de tórax se han desarrollado CAD. A continuación se resumen las técnicas de segmentación utilizadas por algunos autores.

En (47) segmentan los pulmones sección por sección en un proceso bidimensional, aplicando un umbral basado en niveles de gris para cada sección, con el objetivo de identificar la región pulmonar dentro del cuerpo del paciente y diferenciarla de las demás regiones. Se selecciona un punto entre dos picos que aparecen en el histograma de la imagen, que diferencian el tejido blando del tejido denso. De este modo obtienen una primera aproximación a la máscara pulmonar, la que mejoran con una serie de pasos adicionales entre ellos el "rolling-ball". Se prosigue empleando umbralización en base a niveles de gris para generar el volumen pulmonar segmentado. Para obtener los candidatos a nódulos pulmonares, son utilizados múltiples umbrales y la combinación de una serie de técnicas en 2D y 3D.

En (21) realizan la segmentación de los pulmones a partir del análisis de la densidad volumétrica en el histograma de la imagen. Luego del análisis de los picos resultante se fija un umbral de -375 HU, con el objetivo de separar inicialmente el parénquima pulmonar del resto de las estructuras anatómicas. A todos aquellos vóxeles con intensidad menor que el umbral definido se les da valor 1, quedando de color blanco en la imagen resultante; los que están por encima del umbral reciben valor 0 y quedan de color negro. Continúan aplicando la técnica de componentes conexas en 3D para identificar la región pulmonar; siendo

CAPÍTULO 1. Fundamentos Teóricos Metodológicos de las técnicas para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

esta la que mayor área representa comparada con el resto de las estructuras que pueden estar representadas en la imagen médica. Aplicando cierre morfológico utilizando un objeto en forma de disco se obtienen como resultado, los pulmones completamente blancos. De esta forma, se logra un resultado corregido que incluye los posibles nódulos periféricos que pudieron ser excluidos de la región pulmonar durante el proceso inicial de umbralización. Finalmente se obtiene la máscara pulmonar que es el resultado de la extracción de la zona pulmonar de la imagen original. La imagen es llevada a escala de grises ($\text{Gray Level} = \text{HU} + 1024$). Posteriormente se le aplica a la imagen el filtro LDM y al resultado obtenido se le realiza una reducción de falsos positivos a partir de la comprobación de la forma y el tamaño de las estructuras. Para este último paso se utilizan estructuras predefinidas y se especifica que su valor fue determinado de forma experimental.

En (31) se segmentan los pulmones utilizando umbralización basada en valores de niveles de gris, para ello se utilizan el histograma de la imagen. Se obtiene una primera aproximación a la máscara pulmonar en la cual quedan excluidos los nódulos periféricos. Para corregir este resultado se realiza una dilatación utilizando una estructura en forma de disco, seguida de una erosión para que la imagen retorne a su tamaño original. El diámetro del disco es variable y se realizan varias iteraciones. Al obtener la máscara pulmonar se realiza una umbralización. Con el objetivo de hacer un análisis volumétrico de las estructuras contenidas en la máscara pulmonar se realizan grupos que contienen 18 imágenes, se analizan atendiendo al diámetro (no mayor de 3 cm), volumen (no mayor de $14,14 \text{ mm}^3$) y el valor de elongación o capacidad en 2D y 3D. El factor de elongación en 2D se corresponde con la relación de la distancia entre el eje mayor y el menor, de un rectángulo o una elipse que encierra el posible nódulo. El factor de elongación 3D se calculó por la relación de los ejes máximo y mínimo, a partir de las coordenadas de la estructura del vóxel candidato. Adicionalmente se realiza una correspondencia morfológica multinivel (3D), utilizando formas de 3, 6, 9 y 12 mm de diámetro.

En (5) se aplica un umbral adaptativo como primer paso en el proceso de segmentación, el cual permitirá separar dentro de la imagen original las regiones de alta y baja densidad. Las regiones de alta densidad quedaran de color blanco, las cuales serán el cuerpo del paciente, la camilla, el tejido denso dentro del pulmón, los nódulos y la grasa. Zonas como el fondo y el aire del pulmón pertenecen a las regiones de baja intensidad y quedarán de color negro. Sobre el resultado se aplica una técnica de etiquetado de regiones (técnica de supresión) para suprimir el fondo y todas las regiones blancas, manteniendo las

CAPÍTULO 1. Fundamentos Teóricos Metodológicos de las técnicas para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

regiones restantes. A continuación se elimina el fondo y las regiones blancas, quedando el contenido correspondiente al tejido denso dentro de los pulmones. Las dos últimas imágenes obtenidas, se combinan por medio de una operación lógica OR, proporcionando como resultado las regiones pulmonares que incluyen el tejido denso en su interior.

Se realiza una operación morfológica de dilatación, que combinada una vez más por medio de una operación OR, permitiendo recuperar parte del hilio pulmonar que no había sido incluida. Como los nódulos periféricos pueden quedar excluidos de la máscara pulmonar al utilizar la umbralización adaptativa, se utiliza una estructura en forma de disco para corregirla. Como los pulmones habían sido previamente dilatados, se realiza una erosión para devolverlos a su tamaño original. El autor prosigue con la utilización del filtro Iris como umbral de decisión para generar los puntos semillas. A los puntos generados se le aplica la segmentación mediante crecimiento de regiones, lo que brindará como resultado las estructuras pulmonares. Como parte de un proceso experimental, en una de las variantes del algoritmo que implementa el autor, realiza la reducción de falsos positivos a partir de la discriminación de aquellas estructuras nodulares resultantes del filtro de Iris que poseen valor de intensidad menor de 150 HU.

En (20) utilizan para la identificación de estructuras nodulares un umbral fijo a partir del análisis del histograma de la imagen, de igual forma analizando los picos resultantes. Para realizar este proceso se lleva la imagen a 16 bit ($\text{GrayLevel} = 0.3 * \text{RedPixel} + 0.59 * \text{GreenPixel} + 0.11 * \text{BluePixel}$) con el objetivo de minimizar la búsqueda sobre la imagen. De forma experimental es seleccionado el umbral de 150 y los píxeles con mayor valor de intensidad se pintan en negro y los menores en blanco. Posteriormente se utiliza el método de componentes conexas para quedarse solamente con el área que corresponde a la región pulmonar. Se lleva a cabo un cierre morfológico utilizando una estructura en forma de disco para corregir la máscara pulmonar defectuosa, luego del proceso de umbralización. Para retornar la región pulmonar a sus dimensiones originales se aplica una erosión. Para obtener la máscara pulmonar original se hacen corresponder los píxeles blancos en la imagen original. Para finalizar el proceso de segmentación aplica el algoritmo LDM, identificando de esta manera las estructuras que representan máximos locales.

En la Tabla 2 se resumen las técnicas utilizadas por un grupo de autores con el objetivo de identificar nódulos pulmonares solitarios en imágenes de tomografía de tórax; así como los resultados alcanzados en cuanto a sensibilidad y especificidad. Algunos de los resultados mostrados, no son precisamente del

CAPÍTULO 1. Fundamentos Teóricos Metodológicos de las técnicas para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

proceso de segmentación sino el resultado final del algoritmo luego de la clasificación de las estructuras nodulares en benignas y malignas, lo que da la medida de que los resultados en la fase previa deben haber sido satisfactorios.

Tabla 2: Resumen de datos y resultados obtenidos en investigaciones para identificar nódulos pulmonares solitarios relacionados anteriormente (fuente: elaboración propia)

Artículo	Datos	Resultados	Etapas de resultados	Técnicas utilizadas
Automated detection of lung nodules in CT scans: Preliminary results (47)	43 casos Subconjunto de 20 casos (1 o 2 nódulos por imágenes)	Sensibilidad: 70% FP: 1,5 por imagen Sensibilidad: 89% FP: 1,3 por imagen	clasificación	2D y 3D
Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm (21)	8 series (60-80 imágenes)	Sensibilidad: 84,2% FP: 5 por imagen	segmentación	2D y 3D
Pulmonary Nodules: Automated Detection on CT Images with Morphologic Matching Algorithm—Preliminary Results (22)	20 casos (164 nódulos de 3-5 mm de diámetro)	Sensibilidad: 95,1% (156 de 164 nódulos) FP: 6,9 para estructuras de 3 mm FP: 4,0 para estructuras de 5 mm	segmentación	2D y 3D
Desarrollo de un sistema de diagnóstico asistido por computador para detección de nódulos pulmonares en tomografía computarizada multicorte (5)	41 pacientes, 4 150 imágenes (64 nódulos)	Sensibilidad: 80,0% FP: 2 por imagen Sensibilidad: 100% FP: 46.5 por caso	clasificación	2D y 3D
Implementation of Computer Aided Diagnosis System for Lung Cancer Detection (20)	No se especifica	Sensibilidad 96%	segmentación	2D

1.4. Nódulos periféricos

Los nódulos al igual que el cuerpo del paciente y la camilla son estructuras que se representan en la imagen como zonas de alta densidad. Cuando se utilizan técnicas de umbralización para la segmentación del área pulmonar, si existen nódulos periféricos, se producen errores en la máscara resultante. El nódulo se confunde con la estructura que se encuentra fuera del área pulmonar por lo que en la máscara resultante no se encuentra incluida la estructura nodular que hace contacto con la pleura⁴. Ver Figura 3.

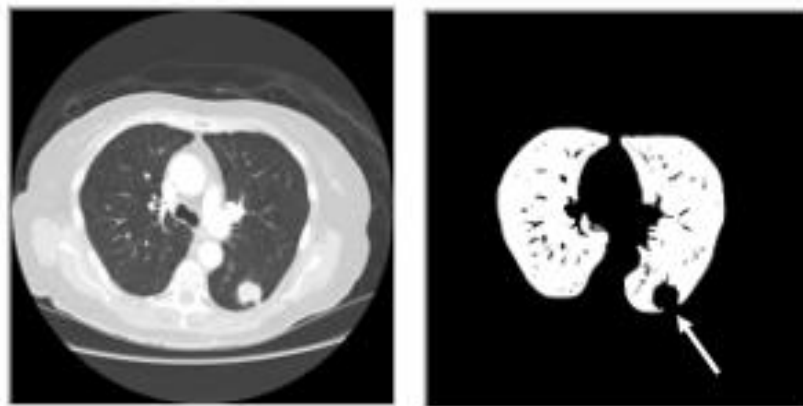


Figura 3: Nódulo periférico o subpleural excluido de la máscara pulmonar al aplicar filtro de umbralización (fuente: elaboración propia)

Para dar solución a este inconveniente se han utilizado numerosas técnicas tales como: cierre morfológico con disco (31), rolling-ball (47, 48), determinación de puntos frontera o técnicas de relleno (49, 50) y contornos activos (51–54).

En (5) se realiza un análisis del funcionamiento de algunas de estas técnicas y sus desventajas al utilizarlas durante la corrección de la máscara pulmonar. La primera técnica mencionada consiste en dilatar y luego erosionar la estructura (el pulmón en este caso). Si se aplican estas transformaciones morfológicas con parámetros pequeños puede que no se consiga el efecto deseado y si se realiza con parámetros muy grandes se puede deformar la máscara del pulmón. La utilización de la determinación de los puntos frontera se lleva a cabo mediante la comparación de las curvaturas en los puntos de la frontera de pulmón. Esta técnica deja trazos, ángulos y bordes rectos, provocando que muchas veces la zona

⁴ Membrana que recubre el exterior de los pulmones y el interior de la cavidad del pecho. (72)

donde se aloje el nódulo periférico no sea rellenada completamente, aunque es suficiente para su detección.

Los contornos activos han demostrado ser una herramienta útil para diferentes aplicaciones en el desarrollo de sistemas CAD; para la segmentación del área pulmonar y la segmentación de los nódulos contenidos dentro de dicha estructura (55). Esta técnica puede ser implementada de forma automática y semiautomática. La variante semiautomática ofrece mejores resultados pero requiere la intervención del radiólogo para establecer los puntos de control y construir la estructura.

Se ha decidido seleccionar la técnica de cierre con disco, utilizada en (20, 22, 56), debido a que ofrece resultados favorables de forma automática y es de bajo costo computacional. Cierre con disco es una técnica que consiste en aplicar las operaciones de dilación y erosión morfológica con un elemento estructural en forma de disco, con el objetivo de rellenar orificios en pequeños fragmentos.

1.5. Algoritmos de procesamiento para la segmentación de imágenes

El objetivo de la fase de segmentación es dividir la imagen para poder diferenciar los objetos o estructuras que están contenidas dentro de ella. En (57) se define la segmentación como el particionamiento de una imagen en un conjunto de regiones homogéneas y significativas, de tal manera que los píxeles en cada región particionada, poseen un conjunto idéntico de propiedades o atributos. Estos conjuntos de propiedades de la imagen pueden incluir los niveles de grises, el contraste, los valores espectrales, o propiedades texturales.

Los métodos de segmentación a emplear están determinados por el tipo de imagen, la estructura anatómica u objeto a identificar y el objetivo específico que se quiere lograr con el proceso. Elementos como el ruido, volúmenes parciales y movimientos en la imagen, también pueden tener consecuencias significativas en el desempeño de los algoritmos de segmentación, cada imagen y tipo de tejido tiene un conjunto de características propias. (58) Existen diversos de estos métodos de segmentación (59, 60) , no existe una solución genérica que se aplique a todas las situaciones, habitualmente se utiliza una combinación de ellos para obtener los resultados deseados sobre la imagen.

En (18) se manifiesta que existen dos enfoques fundamentales para llevar a cabo el proceso de segmentación de la región pulmonar: basada en umbralización y segmentación por modelos deformables. En esta investigación específicamente es factible la utilización del enfoque basado en umbralización

debido a que en las imágenes de tomografía, los tejidos pulmonares se representan con diferentes valores de HU con respecto a otras estructuras anatómicas como es el caso del corazón, el hígado y el tejido óseo (11).

Las técnicas basadas en umbralización son muy utilizadas; estas son especialmente útiles cuando se trata de desarrollar aplicaciones donde la velocidad de procesamiento es un factor importante (59). Para la segmentación del área pulmonar diversos autores hacen uso de la umbralización conjuntamente con otras técnicas como Otsu (61), operaciones morfológicas (62), detección de bordes (31, 63), componentes conexas (64, 65), cierre con disco o “rolling ball” (5, 66).

1.5.1. Segmentación basada en umbralización

La umbralización es un método que busca segmentar imágenes escalares, creando una partición binaria de las intensidades de las imágenes. Una umbralización trata de determinar un valor de intensidad, llamado umbral, que separa las clases deseadas. La umbralización se utiliza para la segmentación de imágenes, donde estructuras diferentes, tienen intensidades contrastantes u otras características diferenciables. Esta técnica posee un alto grado de aceptación. (59)

Es factible el uso de técnicas de umbralización para la segmentación de los pulmones en las imágenes de CT. En dichas imágenes el cuerpo del paciente, la camilla, el tejido denso, los nódulos, los huesos y los órganos son regiones de alta densidad; mientras que el fondo de la imagen, el aire y el gas están representados como regiones de baja densidad. (5)

1.5.2. Segmentación orientada a regiones

El objetivo de la segmentación orientada a regiones (60) es la búsqueda directa de regiones de interés en la imagen. Tiene su base en las reglas de similitud y en la conectividad de los píxeles. Las regiones se forman mediante píxeles que tengan conectividad y presenten alguna propiedad de similitud y discrepancia respecto al resto de los píxeles que no pertenecen a la región.

La segmentación orientada a regiones es una técnica que verifica las siguientes condiciones:

- ✓ La unión de todas las regiones generadas cubren la imagen original a segmentar por completo.
- ✓ Las regiones obtenidas son conexas según la distancia seleccionada.
- ✓ Las regiones obtenidas son disjuntas.

- ✓ Los píxeles pertenecientes a una determinada región cumplen algún criterio de similitud.
- ✓ La condición de similitud que se cumple en los píxeles de una región, no se cumple para los píxeles de la unión de dos o más regiones.

Los métodos de segmentación orientados a regiones dependen fundamentalmente del criterio considerado para determinar la presencia de una región y de la condición que decide las adiciones de píxeles a la frontera de una región. Algunos de estos métodos son: crecimiento de regiones por agregación de píxeles, crecimiento con mezclado, división y fusión de regiones (59, 60).

1.5.3. Segmentación basada en transformaciones morfológicas

Las transformaciones morfológicas se encargan de cambiar la forma y estructura de los objetos. Estas herramientas matemáticas permiten obtener componentes que dan una idea de la forma y estructura de los objetos que forman la imagen. Permiten modificar estas formas para separar los objetos unos de otros, obtener contornos primarios, descomponer formas complejas en formas más simples, obtener contornos dentro de entornos ruidosos y reconstruir elementos distorsionados. (60) Existen un conjunto de operaciones morfológicas tales como: dilatación, erosión, apertura y cierre.

1.5.4. Segmentación basada en bordes

Los límites de regiones y los bordes están estrechamente relacionados, a menudo hay un fuerte ajuste en la intensidad en los límites de dichas regiones. En (67) se plantea que las fronteras son bordes unidos que caracterizan la forma de un objeto, estos por tanto, son útiles para calcular rasgos geométricos como tamaño y orientación. Los bordes corresponden a cambios locales significativos en una imagen y se encuentran en la frontera entre dos regiones. Se podría decir que los bordes son una discontinuidad de algún tipo de función de intensidad de la imagen. Existen muchos tipos de bordes diferentes, algunos son debidos a las sombras que se producen sobre los objeto, otros a la variación de su reflectancia o incluso por la textura de los mismos. Finalmente se define un borde como el cambio local significativo en la intensidad de la imagen que usualmente está asociado con una discontinuidad, bien de la intensidad o su derivada. Las discontinuidades pueden ser de dos tipos:

- ✓ Discontinuidad de salto: donde la intensidad cambia abruptamente, de un valor en un lado de la discontinuidad a otro con valor diferente en el lado opuesto.

- ✓ Discontinuidad de línea o paso: donde las intensidades cambian abruptamente volviendo en poco espacio a su valor inicial.

Los saltos y pasos son raros en las imágenes y lo que se observan son rampas y tejados como se muestra en la Figura 4.

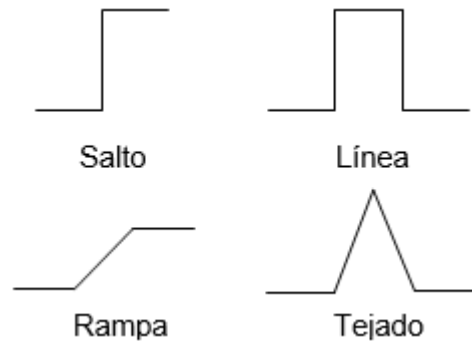


Figura 4: Perfiles de bordes. Modelos idealizados, fila superior y modelos observados, fila inferior (67)

1.6. Conclusiones del estudio del estado del arte

Al realizar el análisis bibliográfico relacionado con las investigaciones que abordan el tema de la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax, se evidencia que la comunidad científica internacional posee intensiones marcadas en la resolución de esta problemática, debido a su impacto en la mortalidad y repercusión en la calidad de vida de las personas. Son diversas y numerosas las soluciones existentes para la identificación de estructuras nodulares en este tipo de imágenes, lo que demuestra que no existe una única combinación de técnicas de segmentación de imágenes que facilite la obtención de buenos resultados.

A pesar de existir varios enfoques para llevar a cabo el proceso de segmentación, se manifiestan tendencias en el uso de la umbralización; determinadas por sus buenos resultados (en imágenes de tomografía, por sus características), su facilidad de implementación y su bajo costo computacional. Durante la fase de obtención de la máscara pulmonar se utiliza frecuentemente, al emplear el enfoque basado en umbralización, estrategias como: umbralización basada en niveles de gris, umbralización adaptativa y análisis del histograma de la imagen. Durante la fase de identificación de los nódulos pulmonares solitarios, se evidencia el uso de técnicas tales como umbralización adaptativa, utilización de

CAPÍTULO 1. Fundamentos Teóricos Metodológicos de las técnicas para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

varios umbrales y el filtro LDM. En esta última fase obtienen mejores resultados aquellas técnicas que utilizan más de un umbral, logrando incluir como candidatas, un mayor número de estructuras.

Se evidencia que aquellas investigaciones que utilizan principalmente técnicas de segmentación puntualmente enfocadas en las características morfológica-radiológicas de las estructuras nodulares, tienden a obtener resultados más concretos y enfocados a la problemática. Es frecuente ver que como se combinan técnicas de procesamiento de imágenes en 2D y 3D, teniendo en cuenta que en cada una de las imágenes se representa un corte o fragmento de una estructura.

Teniendo en cuenta que existen autores que en sus investigaciones obtienen elevados valores de sensibilidad y otros son capaces de reducir un número importante de falsos positivos; se considera necesario el desarrollo de un algoritmo de identificación de nódulos pulmonares solitarios que combine aquellas técnicas que se enmarcan en la caracterización de las estructuras mencionadas. Es preciso obtener un alto acierto, a partir de un elevado valor de sensibilidad y con la menor cantidad posible de falsos positivos por imagen médica analizada. Para alcanzar estos resultados se utilizarán técnicas que realicen el procesamiento solo en 2D por considerarse más sencillas y menos costosas en tiempo de ejecución, ver Tabla 3.

Tabla 3: Técnicas de segmentación a emplear en el desarrollo del algoritmo de identificación de nódulos pulmonares solitarios

Extracción de la región pulmonar	Detección de candidatos a nódulos pulmonares	Reducción de falsos positivos
<ul style="list-style-type: none">- Umbralización adaptativa (utilizando método de Otsu)- Inversión- Componentes conexas- Cierre con disco	<ul style="list-style-type: none">- LDM- Filtro Iris	Atendiendo a: <ul style="list-style-type: none">- tamaño- esfericidad

1.7. Conclusiones del capítulo

- ✓ Los pasos básicos seguidos por los desarrolladores de algoritmos para la identificación de nódulos pulmonares solitarios son: extracción de la región pulmonar, detección de los nódulos pulmonares, reducción de falsos positivos.

CAPÍTULO 1. Fundamentos Teóricos Metodológicos de las técnicas para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

- ✓ Existe una tendencia a utilizar técnicas de umbralización durante el proceso de segmentación, para la obtención de la máscara pulmonar. Se considera que ofrecen resultados satisfactorios en el proceso de segmentación de imágenes de tomografía. Son ampliamente utilizadas debido a su facilidad de implementación y bajo costo computacional al analizar la imagen médica.
- ✓ Al aplicar la técnica de umbralización para la extracción de la máscara pulmonar, esta última debe ser corregida, a partir de la realización de otras operaciones sobre la imagen, para incluir los nódulos periféricos que fueron eliminados en dicha fase.
- ✓ Al analizar las características de los nódulos pulmonares solitarios, su forma de representación en las imágenes de tomografía y los resultados obtenidos por otros autores, se considera una variante factible la utilización de algoritmos como LDM e Iris; teniendo en cuenta que su combinación pudiera propiciar alcanzar altos valores de sensibilidad y la reducción de un considerable número de falsos positivos, durante el proceso de identificación de estructuras nodulares.

CAPÍTULO 2. Algoritmo para la clasificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

En este capítulo se explican las transformaciones realizadas a las imágenes médicas con el objetivo de reducir el área de análisis e identificar las estructuras candidatas a ser nódulos pulmonares solitarios. Los procedimientos realizados se enmarcan en cada una de las etapas de segmentación por las que transcurre la imagen: extracción de la región pulmonar, detección de los nódulos pulmonares y reducción de falsos positivos. Se muestra sobre la imagen original los resultados obtenidos al aplicar las técnicas seleccionadas. Se describen los principales métodos implementados para el desarrollo del algoritmo.

2.1. Propuesta de solución

A partir de los estudios realizados a las técnicas de segmentación utilizadas por varios autores para la identificación de estructuras nodulares, se propone la realización de un algoritmo capaz de detectar nódulos pulmonares solitarios en imágenes de tomografía de tórax. Teniendo en cuenta las características morfológicas-radiológicas con que se presentan los nódulos pulmonares solitarios en las imágenes de tomografía, el algoritmo propuesto analiza la imagen seleccionada por el radiólogo, en busca de estructuras de alta densidad dentro de la región pulmonar y luego ejecuta un proceso discriminatorio al seleccionar solo aquellas estructuras que cumplan ciertos criterios de tamaño y circularidad. Las estructuras resultantes de este proceso son analizadas con el objetivo de seleccionar finalmente solo aquellas en las que los valores de intensidad crecen hacia el centro de la misma. Los resultados obtenidos se muestran de forma automática en la imagen original, a partir del señalamiento de las estructuras nodulares que pudieran ser de interés para el especialista durante el proceso de diagnóstico médico.

La reducción de falsos positivos se puede realizar en diversos momentos del proceso de segmentación. Puede resultar factible hacerlo intercalado con las técnicas de identificación y de esta forma se reduce el tiempo de análisis al contar con menos estructuras a procesar. En este caso en particular es factible la reducción antes de aplicar el filtro de Iris teniendo en cuenta que este último es complejo y requiere de la realización de un conjunto de operaciones matemáticas; siendo significativo evitar el procesamiento innecesario de un conjunto de estructuras.

Extracción de la región pulmonar

Para obtener la solución deseada, la imagen médica debe ser segmentada con el objetivo de ir reduciendo zonas de análisis y ser más eficiente al identificar las estructuras candidatas. La imagen de tomografía transita por una serie de pasos, comenzando con una umbralización binaria. Al analizar las características de este tipo de imágenes es determinante resaltar que los valores de densidad no son uniformes, varían según los parámetros de calibración del escáner y el espesor a lo largo del cuerpo del paciente. A partir de este elemento se decide la implementación de un umbral adaptativo y se escoge el método automático Otsu (68, 69); el cual elige un umbral óptimo maximizando la varianza entre clases, mediante una búsqueda exhaustiva. Con este procedimiento se separan las regiones de alta y baja densidad representadas en la imagen, Figura 5(b).

Con la imagen umbralizada se procede a la identificación de la región pulmonar, para lo cual se aplica la técnica de detección de componentes conexas. Esta técnica se basa en la identificación de las estructuras de color blanco que se encuentran dentro de la imagen, por lo que es necesario invertir la misma, Figura 5(c). Al concluir la inversión de la imagen se procede a aplicar correctamente la técnica de detección de componentes conexas, Figura 5(d). Dicha técnica permite identificar las estructuras aisladas y eliminar aquellas que no cumplan los requerimientos de tamaño para clasificar como regiones pulmonares, Figura 5(e).

Luego de haber identificado las regiones correspondientes a los pulmones, se aplica el método "Flood-Fill", que identifica todos los píxeles que sean de igual color al punto inicial y a su vez contiguos entre sí y los sustituye por el color de relleno, para obtener una imagen en la que se aprecia el tejido denso dentro de los pulmones. Posteriormente se realiza una operación lógica XOR entre esta última imagen y la imagen de la Figura 5(e). Dicha operación compara los píxeles entre ambas imágenes, si estos son iguales da como resultado 0 y 1 en caso contrario. El resultado de esta operación se invierte para obtener una primera aproximación a la máscara pulmonar, Figura 5(f).

Existen estructuras nodulares que se localizan en la periferia de los pulmones. Estas estructuras de densidad similar al cuerpo y la camilla del paciente, pueden ser eliminadas tras la umbralización. Para dar solución a este problema se aplicó el filtro cierre con disco para suavizar los bordes de los pulmones y poder incluir las estructuras que pudieran ser de interés para el especialista, Figura 5(g). En las investigaciones analizadas no se especifica el tamaño y las iteraciones utilizadas para la ejecución de esta

técnica y es por ello que dichos valores se seleccionan de forma experimental. Se definió una estructura de tamaño 3x3 y se realizaron 6 iteraciones durante su aplicación. Después de haber obtenido la máscara pulmonar corregida, se procede a la identificación de la región pulmonar, Figura 5(h), llevando a cabo una correlación de este último resultado con respecto a la imagen original.

Detección de los nódulos pulmonares

Como los nódulos pulmonares solitarios, aparecen en una imagen de tomografía como una estructura de forma circular con altos valores de intensidad, se aplicó el algoritmo Local Density Maximum (LDM). Se comienza aplicando un umbral inicial que pudiera ser el máximo valor de densidad de la imagen, que luego irá disminuyendo según el valor de paso previamente definido, hasta alcanzar el valor del umbral de parada. LDM compara las estructuras identificadas en el nivel de umbral actual, con las estructuras identificadas en el nivel de umbral previo y analiza cuáles de estas estructuras pueden ser incluidas como posibles estructuras nodulares, Figura 5(i).

Reducción de falsos positivos

Una vez que se tienen aquellas estructuras que están representadas en la imagen médica con altos valores de intensidad, se lleva a cabo un proceso de reducción de falsos positivos, a partir de las siguientes características de los nódulos pulmonares solitarios.

✓ Tamaño

Se eliminan aquellas estructuras que poseen un diámetro que no se encuentra en el intervalo de 3 a 30 mm. Las que se encuentran por encima del límite superior no son consideradas nódulos pulmonares solitarios y por las características de los datos y las imágenes con las que se cuenta para la validación del algoritmo desarrollado se decide restringir el límite inferior.

✓ Circularidad

Para analizar qué tan circular es una estructura, se determina la diferencia entre el radio máximo y el mínimo. La proporcionalidad entre la diferencia y el radio máximo no puede ser mayor del 25%.

$$dif(r1, r2) / \max(r1, r2) \leq 0,25$$

Detección de los nódulos pulmonares

Al depurar las estructuras resultantes del análisis de los altos valores de intensidad, las candidatas a ser nódulos pulmonares presentan un patrón en el cual, dichos valores disminuyen desde el centro hacia los bordes, es decir, crecen hacia el centro de la estructura. En un mapa de gradiente, los nódulos aparecen como zonas donde los gradientes convergen; es por ello que el análisis del campo de vectores de gradiente se escoge como un método indicado para la identificación de dichas estructuras. Teniendo en cuenta estas características se aplica a la imagen resultante una variante de estos filtros de gradiente, denominada filtro de Iris, se muestran los resultados sobre la imagen original, Figura 5(j).

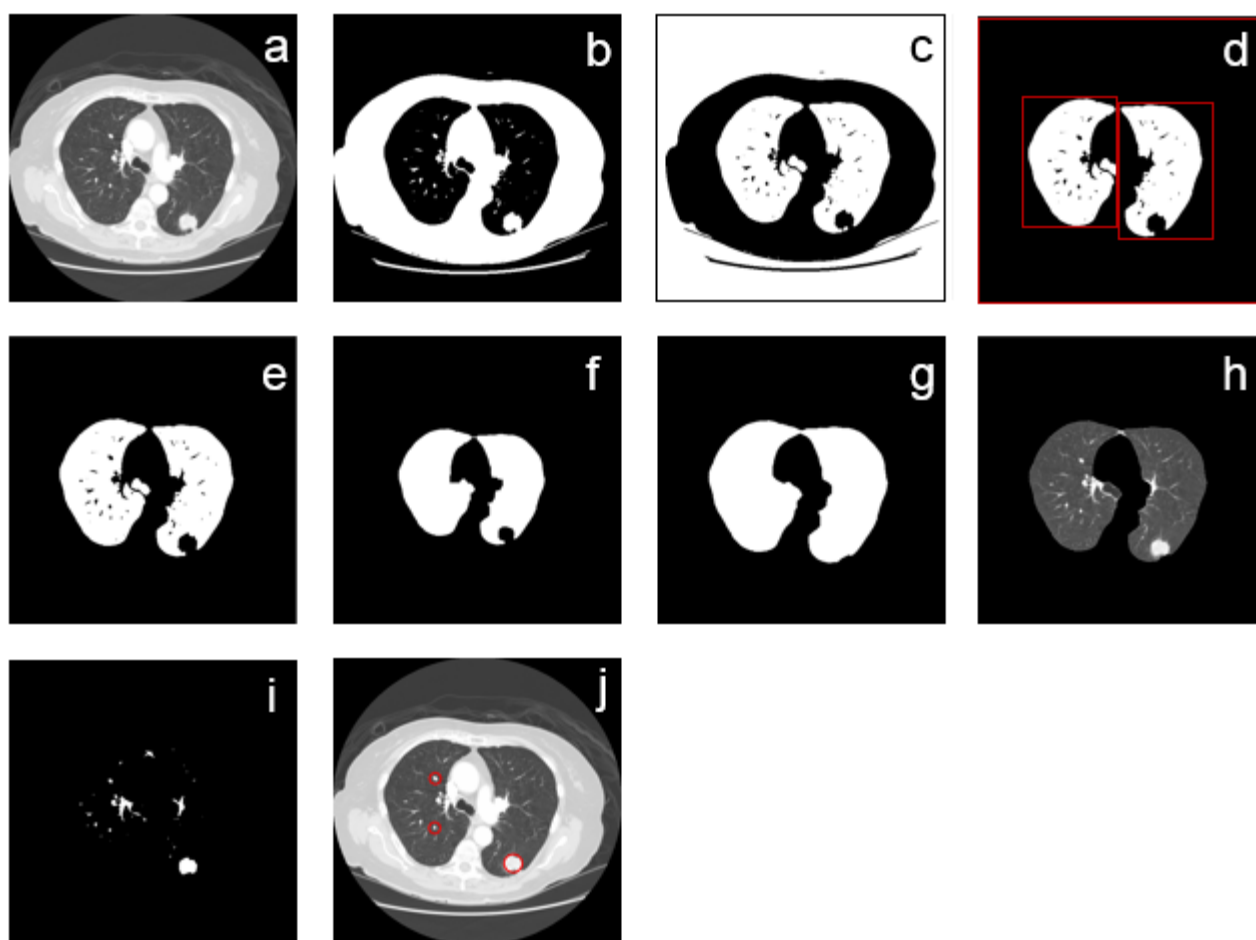


Figura 5: Proceso de identificación de nódulos pulmonares solitarios: (a) imagen original de imagen original de tomografía, (b) umbralización adaptativa con el método de Otsu, (c) inversión de la imagen

umbralizada, (d) detección de componentes conexas en (c), (e) identificación de las regiones correspondiente a los pulmones, (f) primera aproximación de la máscara pulmonar, (g) máscara pulmonar corregida luego de haber aplicado cierre con disco, (h) delimitación de la región pulmonar mediante una comparación píxel a píxel entre las figuras (a) y (g), (i) aplicación del filtro LDM sobre la región pulmonar (h), (j) resultado del filtro Iris mostrado en la imagen original.

2.2. Funcionamiento del filtro Local Density Maximun (LDM)

LDM (20, 21) es una técnica que se enfoca en la determinación de estructuras como los vasos sanguíneos, las paredes bronquiales y los nódulos pulmonares solitarios que manifiestan en la imagen de tomografía, valores de densidad superiores a los del parénquima pulmonar. Por lo tanto, esta técnica se centra en encontrar objetos que tienen alta densidad en su vecindario.

Para el cálculo de las estructuras candidatas a ser resultantes de la aplicación del algoritmo, se utilizan los siguientes valores:

1. Valor de umbral inicial: especifica el valor de la intensidad donde empieza el umbral.
2. Valor de paso del umbral: especifica el valor de paso en que el umbral se reduce en cada momento.
3. Umbral de parada: es el valor de la intensidad en que se detiene el algoritmo.
4. Densidad mínima de máximo local.
5. Relación de valor local: especificando el cambio de volumen del objeto con respecto al volumen de su caja circundante.

Todos los valores se determinan experimentalmente. Inicialmente, se selecciona el valor de la densidad máxima de la imagen como el valor de umbral inicial, todos los objetos se encuentran por encima de este valor de umbral. Dicho valor se reduce sucesivamente atendiendo al valor de paso y hasta el umbral de parada. La imagen de CT se convierte en blanco y negro a partir del valor de umbral que se analiza en cada momento. En cada paso del proceso se determinan las componentes conexas.

Los objetos detectados en el nivel de umbral actual, pueden contener algunos objetos detectados en el nivel de umbral anterior. Para un objeto i dado, en el nivel de umbral actual, existe m objetos detectados en el nivel anterior, de tal manera, que el objeto i contiene todos estos m objetos dentro de ella.

El valor cero para m significa que el objeto recién detectado no contiene ningún objeto del nivel anterior dentro de ella. Esto significa, que este objeto es nuevo y se añade al conjunto de nódulos detectados actualmente. El valor 1 de m , significa que el objeto actual contiene sólo un objeto del nivel anterior. Para determinar las especificaciones anteriormente descritas, se define un cuadro delimitador de un objeto detectado, como el rectángulo más pequeño que contiene completamente la estructura identificada. Para los cálculos correspondientes se definen los siguientes parámetros:

1. dx : es la diferencia entre la coordenada x en el objeto.
 2. dy : es la diferencia entre la coordenada y en el objeto.
 3. d : se define como el cuadrado del máximo valor de dx y dy .
 4. $R1$ es el número de píxeles del objeto actual dividido por d .
 5. $R2$ es el número de píxeles del objeto anterior dividido por d .
 6. R es la relación de $R1$ a $R2$.
- ✓ Si el valor de R es menor que la relación que se define en el primer paso, entonces esto significa que R cae drásticamente. En este escenario, el objeto actual se interpretará como una meseta y el objeto anterior sustituye al nuevo objeto. Entiéndase meseta, como una región, en la cual dos máximos locales están lado a lado en el mismo vecindario o contorno.
 - ✓ Si cualquier R es mayor o el objeto anterior no cumple los criterios de ser máximo local, el objeto actual sustituye al objeto anterior.
 - ✓ Si el valor de m es mayor que 1, significa que el objeto actual contiene más de un objeto anterior. Todos esos objetos anteriores, que satisfacen los criterios de ser máximos locales como se indica en la primera etapa, son reconocidos como máximos locales, mientras que el objeto actual se reconoce como meseta.
 - ✓ Si solo uno de los objeto anteriores cumple con los criterios del primer paso, entonces se trata como en el segundo caso (es decir, $m = 1$). Si ninguno de los objetos anteriores cumple los criterios, se trata como el primer caso ($m = 0$). El proceso de búsqueda de máximos locales termina cuando el umbral alcanza el valor mínimo de la densidad especificado en la primera etapa.

Determinación de los valores de los parámetros necesarios para la ejecución del filtro LDM

En las investigaciones donde se aborda el tema referente a la implementación del filtro LDM, se evidencia que el método presenta altos valores de sensibilidad al ser utilizado en las imágenes de tomografía para la identificación de nódulos pulmonares solitarios. Estas plantean que la definición de los parámetros necesarios para su ejecución, es determinada de forma experimental y no se especifican los parámetros utilizados o estos no son aplicables. Teniendo en cuenta estos elementos, en la presente investigación se realiza el procedimiento que se describe a continuación, para la determinación de los valores de los parámetros del filtro LDM.

Se escoge un conjunto de imágenes que se encuentran públicas en The Lung Image Database Consortium Image Collection (70–72), de una muestra de 14 series se seleccionan 5 de ellas (LIDC-IDRI-0003, LIDC-IDRI-0005, LIDC-IDRI-0006, LIDC-IDRI-0008, LIDC-IDRI-00013). El criterio de selección de las series está determinado por la diversidad en cuanto a tamaño y patrones de calcificación de las estructuras contenidas en las imágenes de dichas series. El objetivo es ganar la mayor representatividad posible, con respecto a las principales características de los nódulos; con la intención de determinar la combinación indicada entre los valores de: paso, relación entre las estructuras de nivel actual-previo y la densidad mínima del máximo local para identificar el mayor número posible de estructuras.

Durante las pruebas se someten las imágenes a varios procesos de umbralización en el cual se modifican los valores de los parámetros utilizados por el algoritmo LDM. El paso se varía de 10 a 70 avanzando 10 cada vez, la proporción desde 0.1 a 1 con paso 0.1 y el tamaño mínimo de 10 a 70 con paso 10. Los valores de inicio y fin del rango de valores de intensidad se mantuvieron fijos en 215 y 10 respectivamente.

Al término de cada proceso se almacenan en un fichero los resultados de sensibilidad alcanzados por el algoritmo con la combinación de valores correspondiente. Todos los ficheros resultantes son analizados en busca de la combinación de valores más coincidente. El resultado final de este proceso es la combinación de umbrales que se muestra a continuación:

- ✓ 50, 0.8, 20
- ✓ 20, 0.1, 20
- ✓ 30, 0.9, 20
- ✓ 40, 0.9, 20

✓ 10, 1.0, 20

Aquellos nódulos pulmonares solitarios que son semisólidos o no sólidos, pueden no ser identificados al utilizar la misma combinación de valores que se utiliza para una estructura sólida y es por ello que para garantizar mayor sensibilidad se utiliza para dar solución a esta problemática varias combinaciones de los parámetros del algoritmo LDM. Si bien es cierto que la utilización de varias combinaciones aumenta el valor de sensibilidad en la identificación de nódulos pulmonares, también aumenta el número de falsos positivos; de igual forma al restringir por el tamaño y la forma de las estructuras y aplicar luego el filtro de Iris, se reduce un número importante de falsos positivos. Como resultado de la umbralización no todas las estructuras presentan las características de circularidad y tamaño indicadas, para luego ser identificadas como nódulos pulmonares solitarios.

2.3. Descripción del filtro de Iris

En (5) se plantea que el filtro de Iris fue diseñado para realzar regiones redondeadas no homogéneas y un nódulo, idealmente, se considera una estructura esférica con densidad variable. Este método se encarga de determinar cuáles de las estructuras que fueron previamente identificadas como estructuras de alto valor de intensidad, presenta patrón de gradiente dirigido hacia el centro de la misma.

Para aplicar el filtro de Iris en el algoritmo desarrollado, se analizan las estructuras contenidas dentro de la región pulmonar, que fueron resultantes del proceso de aplicación del filtro LDM. Para realizar este proceso se utiliza la imagen original y la imagen binarizada resultante del procesamiento anterior. A estas estructuras previamente seleccionadas, se les aplica el filtro de gradiente sobel, con el objetivo de calcular el vector gradiente en cada píxel. Con el resultado de este proceso es posible obtener la orientación del gradiente con respecto al eje x, a partir de la siguiente expresión:

$$\alpha = \tan^{-1} \frac{G_C(x',y')}{G_R(x',y')}$$

Para estimar la media de convergencia del campo de vectores gradientes hacia un punto, se analizan los gradientes de los píxeles del entorno del punto sobre el cual se aplica. Será mayor la salida del filtro, mientras más vectores de gradientes estén dirigidos hacia el punto de análisis. Se analizan entonces los gradientes a lo largo de las N líneas, a modo de radios de circunferencia radiadas desde P (punto de

análisis), ver Figura 6. Para delimitar el entorno del punto, se utiliza el parámetro Rmax (radio máximo) y se analizan para cada línea todos los píxeles desde el punto P hasta la distancia máxima (Rmax).

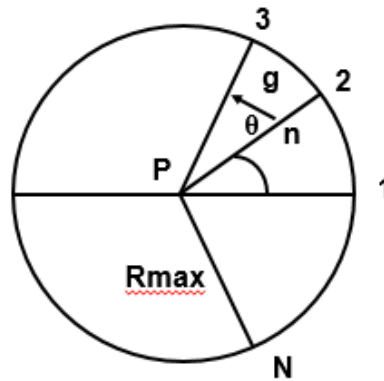


Figura 6: Punto de aplicación del filtro (P) y N líneas de análisis (5)

En (5) se define el índice de convergencia (IC) del vector gradiente (g) en un punto Q_i , de coordenadas (x', y') , con respecto a un punto P, como el coseno del ángulo que forma el vector con la línea que une ambos puntos. El punto Q_i , de coordenadas (x', y') se encuentra en la intersección del vector g con la línea 2. Para cada punto Q_i , el filtro calcula la media de todos los índices de convergencia a lo largo de la línea $\overline{PQ_i}$. A esto se le llama índice de convergencia modificado (ICM) y se determina con la siguiente expresión:

$$ICM = \frac{1}{\overline{PQ_i}} \sum_P^{Q_i} \cos \theta (x', y') dl$$

La salida del filtro Iris en el punto P se define como:

$$OP = \frac{1}{N} \sum_{i=0}^{N-1} \max_{Rmin \leq n \leq Rmax} \frac{1}{N} \sum_{m=1}^n \cos \theta im$$

Donde el final de la expresión representa la media de los índices de convergencia desde el punto P hasta Q_i , a lo largo de la línea i-ésima y el índice n representa la distancia en píxeles, que separa los puntos P y Q_i .

```
1 FiltroIris(imagenOriginal, imagenBinarizada, iteraciones)
2 Inicio
3 Sobel ← AplicarSobel(imagenOriginal)
4 Para k ← 0 Hasta anchoImagen Con Paso 1 Hacer
5     Para l ← 0 Hasta altoImagen Con Paso 1 Hacer
6         Si mascara(k,l) != 255 Entonces
7             iris(k,l) = 0
8         SiNo Entonces
9             sumal ← 0
10            Para i ← 0 Hasta iteraciones Con Paso 1 Hacer
11                maxco ← -2
12                Para n ← radioMinimo Hasta radioMaximo Con Paso 1 Hacer
13                    suma ← 0
14                    Para m ← 1 Hasta n Con Paso 1 Hacer
15                        Calcular x
16                        Calcular y
17                        Si x < ancho and x >= 0 and y < alto and y >= 0 Entonces
18                            calcularAngulo
19                            suma ← suma + coseno(angulo)
20                        FinSi
21                    FinPara
22                    Si suma/n > maxco Entonces
23                        maxco ← suma/n
24                    FinSi
25                FinPara
26                sumal ← sumal + maxco
27            FinPara
28            iris(k,l) ← sumal/iteraciones
29        FinPara
30    FinPara
31    return Image(iris)
32 FinProcedimiento
```

Figura 7: Pseudocódigo utilizado del filtro de Iris (fuente: elaboración propia)

2.4. Descripción de la secuencia de pasos desarrollados por el algoritmo para identificar los nódulos pulmonares solitarios

Durante el desarrollo del algoritmo de identificación de nódulos pulmonares solitarios propuesto en la presente investigación, se utilizó la librería EmguCV 2.4.2, que es un un wrapper de la librería de procesamiento de imágenes OpenCV, permite realizar invocaciones directas a funciones de OpenCV (73, 74). Es una librería de código abierto y multiplataforma diseñada con el lenguaje de programación C# (75). Esta librería ha sido seleccionada ya que incluye funcionalidades esenciales para el procesamiento de imágenes médicas y por su infraestructura fácil de usar. Otras de las ventajas que se tuvo en cuenta fue la

capacidad de serializar imágenes en ficheros XML. Ha sido utilizada por varios investigadores en el proceso de identificación de nódulos pulmonares solitarios (76, 77).

Para realizar el proceso de lectura y visualización de las imágenes médicas en formato DICOM se utilizó la librería CALIB (78). Dicho componente fue desarrollado en Centro de Informática Médica perteneciente a la Universidad de las Ciencias Informáticas.

Para hacer uso del algoritmo para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax, el especialista debe utilizar un visor de imágenes médicas. Dicho visor permitirá cargar una imagen de tomografía de tórax en formato DICOM. Para que el sistema visualice las estructuras candidatas, la imagen en análisis transcurre por una serie de transformaciones que se mencionan a continuación.

1. Se ejecuta el método *PulmonaryNoduleDetection* que posee en su interior todas las técnicas a utilizar. Este método se encuentra en la clase *ImageAnalysis*.
2. La primera técnica que se ejecuta es *Threshold*, que se implementa en la misma clase *ImageAnalysis*; pero requiere de la funcionalidad *cvThreshold* que se encuentra en la librería *Emgu.CV.CvInvoke* para aplicar el método de umbralización basado en Otsu y poder obtener el umbral óptimo a utilizar.
3. Se invierte la imagen con la funcionalidad *Invert* que se encuentra en la clase *ImageAnalysis*, esta funcionalidad necesita para su correcta realización del método *_MorphologyEx* que se encuentra en la librería *Emgu.CV.Image*.
4. Al concluir la inversión de la imagen el algoritmo prosigue en la identificación de la región pulmonar, la cual realiza con la utilización de la funcionalidad *PulmonaryRegion* que se ha implementado en la clase controladora *ImageAnalysis*.
5. A la región pulmonar resultante en el paso anterior se le aplica la técnica *Mask* que se encuentra dentro del método *PulmonaryNoduleDetection*. Esta técnica ejecuta la funcionalidad *Fill4Corners* que se encuentra en la clase *Util*, permitiendo rellenar la región pulmonar, dando como resultado la máscara pulmonar.
6. Para evitar que los pulmones se peguen al borde de la imagen cuando se está aplicando el cierre con disco, es usado el método *ExpandBackground* que se encuentra en la clase *Util*.

7. Con el fondo expandido se procede a aplicar *ImproveMask*, la misma utiliza iterativamente la funcionalidad *CloseDisk* de la clase *Util*. La técnica *ImproveMask* se utiliza con el fin de incluir los nódulos periféricos que fueron excluidos de la máscara pulmonar en el proceso de umbralización.
8. Para que la máscara pulmonar retome su posición y tamaño original se utilizan los métodos *AdjustPosition* y *ReduceBackground*, que se encuentran en la clase *Util*.
9. Con la máscara pulmonar corregida, se procede a la identificación de la región pulmonar, lo cual se realiza con el método *OriginalPulmonaryRegion* desarrollado en la clase *ImageAnalysis*.
10. Cuando se obtiene la región de interés se le aplica el algoritmo LDM, el cual se realiza con el método *LdmVariableThresHold* con el fin de identificar las estructuras con altos valores de densidad que se encuentran en la imagen que se está analizando. Se utiliza un método auxiliar *OverlapComponent* que se encuentra en la clase *Util*, que es quien determina la estructura candidata entre las estructuras actuales y las previas.
11. Con el resultado arrojado por el algoritmo LDM, se identifican varias estructuras que por sus características no pueden ser nódulos pulmonares solitarios, por lo que se precede a la eliminación de las mismas. Esta acción es ejecutada a través del método *ReductionFalsePositive*, el cual se implementado en su totalidad en la clase controladora *ImageAnalysis*.
12. Las estructuras de alto valor de intensidad son analizadas para identificar en cuáles de ellas dicho valor crece hacia el centro, para ello se aplica al resultado anterior el filtro de Iris a través del método *ApplyIrisForRegions*. Se hace uso del método *IrisFilter* que es donde se analiza la tendencia del gradiente en las regiones conexas previamente identificadas.
13. Dicho método devuelve las regiones que cumplen con las características señaladas.
14. Se señalan en la imagen original las estructuras candidatas a ser nódulos pulmonares solitarios haciendo uso del método *ShowCandidateNodules*.

2.5. Conclusiones del capítulo

- ✓ En este capítulo se presentó la propuesta para la implementación del algoritmo de identificación de nódulos pulmonares solitarios, basado en técnicas de procesamiento de imágenes que se enfocan en sus características morfológicas-radiográficas.
- ✓ Se describen las especificaciones llevadas a cabo en cada fase del proceso y se mostraron las imágenes con los resultados obtenidos al aplicar cada una de las técnicas principales.

- ✓ Se evidencia que es factible utilizar la combinación del filtro LDM y el filtro Iris para la identificación de los nódulos pulmonares solitarios. Con la utilización de LDM se obtiene alta sensibilidad, mientras que con Iris se reduce un número importante de falsos positivos por imagen médica analizada.
- ✓ El filtro LDM es utilizado con más de una combinación de valores de sus parámetros de ejecución, para obtener una mayor sensibilidad en la identificación de los nódulos pulmonares solitarios.

CAPÍTULO 3. Análisis y valoración de los resultados obtenidos por el algoritmo propuesto para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

En este capítulo se evidencia el valor de sensibilidad obtenido por el algoritmo desarrollado para la identificación de nódulos pulmonares solitarios. Se realiza la validación de los resultados del algoritmo propuesto, utilizando un conjunto de imágenes previamente diagnosticadas y publicadas en The Lung Image Database Consortium Image Collection (LIDC-IDRI). Se describen los resultados alcanzados, a partir de una comparación de los valores de sensibilidad y el número de falsos positivos por imagen, con respecto a los obtenidos por dos de los algoritmos relacionados en el estado del arte.

3.1. Indicadores estadísticos básicos para evaluar el desempeño de un procedimiento diagnóstico

Se considera que una prueba diagnóstica es buena cuando ofrece resultados positivos en enfermos y negativos en pacientes sanos, con el menor rango de error posible. Por lo tanto, las condiciones que deben ser exigidas en un test diagnóstico son principalmente tres (79):

- ✓ **Validez:** grado en que un test mide lo que se supone que debe medir, es decir, la frecuencia con la que los resultados obtenidos con este test pueden ser confirmados por otros más complejos y rigurosos.
- ✓ **Reproductividad:** capacidad de un test de ofrecer los mismos resultados cuando se repite su aplicación en circunstancias similares. La reproductividad viene determinada por la variabilidad biológica del hecho observado, la introducida por el propio observador y la derivada del propio test.
- ✓ **Seguridad:** certeza de que un test predecirá la presencia o ausencia de enfermedad en un paciente. La seguridad viene determinada por el valor predictivo de un resultado positivo o negativo, es decir, la probabilidad de que siendo un test positivo el paciente esté realmente enfermo.

CAPÍTULO 3: Análisis y valoración de los resultados obtenidos por el algoritmo propuesto para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

Los parámetros que miden la validez de una prueba diagnóstica son la sensibilidad y la especificidad. Miden la discriminación diagnóstica de una prueba en relación a un criterio de referencia, que se considera verdad; es decir, se comparan los resultados de la aplicación de la prueba con lo que muestra una situación existente. A la situación existente válida para comparar los resultados se le llama “prueba de referencia” o “patrón de oro” (80). El resultado de una prueba diagnóstica puede ser positivo o negativo, pero estos pueden ser correctos o incorrectos, dando lugar a cuatro tipos de resultados: verdaderos positivos (VP), verdaderos negativos (VN), falsos positivos (FP) y falsos negativos (FN). Ver Tabla 4.

Tabla 4: Relación entre los resultados de una prueba diagnóstica y la presencia o ausencia de una enfermedad (79)

Resultado de la prueba	Verdadero diagnóstico	
	Enfermo	Sano
Positivo	Verdaderos Positivos (VP)	Positivo
Negativo	Falsos Negativos (FN)	Negativo

La sensibilidad (S) indica cuan “sensible” es la prueba a la presencia de la enfermedad. Se refiere a la proporción de sujetos que tienen la enfermedad y el resultado de la prueba realizada da positivo (81). Puede estimarse como la proporción de pacientes enfermos que obtuvieron un resultado positivo en la prueba diagnóstica, es decir, la proporción de verdaderos positivos, o de enfermos diagnosticados, respecto del total de enfermos en la población de estudio (79):

$$S = \frac{VP}{VP + FN}$$

Como en esta investigación se creó un entorno para la identificación de nódulos pulmonares solitarios, se ajusta la definición de sensibilidad y especificidad a la temática abordada, quedando de la siguiente forma:

Sensibilidad: representa la proporción de estructuras que son identificadas como nódulos pulmonares solitarios cuando estas realmente lo son.

Especificidad: Proporción de estructuras que no se identificaron como nódulos pulmonares cuando estas realmente no lo son (pudiendo ser: bronquios, vasos sanguíneos, etc.).

CAPÍTULO 3: Análisis y valoración de los resultados obtenidos por el algoritmo propuesto para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

En investigaciones similares (5, 47), en la que se aborda el tema de la identificación de nódulos pulmonares solitarios, se plantea que no es provechoso hacer un análisis del total de falsos positivos detectados por el algoritmo, luego de la ejecución de una serie de imágenes. El valor total de este indicador, visto de esta forma, resulta muy elevado y no da la medida real del comportamiento del algoritmo. Teniendo en cuenta estos elementos es que se contrasta el valor de sensibilidad alcanzado con respecto al promedio de falsos positivos por imagen médica. De esta forma se aprecia qué tan acertado es el resultado del algoritmo, al identificar las estructuras nodulares en cada una de las imágenes analizadas.

Se escoge como “patrón de oro” para validar los resultados arrojados por el algoritmo implementado The Lung Image Database Consortium Image Collection (70–72), base de datos pública que se utiliza en la comunidad científica para el desarrollo y la evolución de los CAD. Las imágenes contenidas en dicha base de datos están en formato DICOM (Digital imaging and communications in medicine) y se encuentran anonimizadas para garantizar la privacidad de los pacientes. La colimación de las series de imágenes publicadas no es superior a los 3 mm. Para señalar las lesiones presentes en las imágenes se utilizaron los criterios de 12 imagenólogos, que colaboran con el consorcio. Dichos especialistas realizaron una revisión de las imágenes, utilizando una interfaz de ordenador, en la cual se les permitió realizar mediciones y variar los parámetros de ancho y centro de ventana, para finalmente marcar las lesiones identificadas.

Por cada una de las series de imágenes que contiene la base de datos del consorcio, se dispone de un fichero que describe en formato XML, los hallazgos de los especialistas. Se especifican las estructuras que son consideradas nódulos pulmonares solitarios, a través de la descripción de un conjunto de coordenadas que representan su borde. Se puntualiza la imagen médica en la que está contenida; así como características de especulación, lobulación, textura, densidad, entre otras características distintivas relativas a dichas estructuras.

Para medir los resultados obtenidos, se verifica la correlación entre los resultados representados en el XML y los arrojados por el algoritmo propuesto. Realizando una analogía con lo descrito en la Tabla 4, para calcular los valores de sensibilidad se utiliza la relación de la Tabla 5.

CAPÍTULO 3: Análisis y valoración de los resultados obtenidos por el algoritmo propuesto para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

Tabla 5: Relación entre los resultados de los ficheros XML y la arrojadas por el algoritmo desarrollado (fuente: elaboración propia)

Resultados del algoritmo	Resultados del XML	
	Presente	Ausente
Presente	Verdaderos Positivos (VP)	Presente
Ausente	Falsos Negativos (FN)	Ausente

3.2. Automatización del cálculo de los indicadores de validez del algoritmo para la identificación de nódulos pulmonares solitarios

Para validar los resultados obtenidos en el desarrollo de los algoritmos para la identificación de nódulos pulmonares solitarios, se decide analizar y comparar los indicadores, determinando la correspondencia con respecto a los resultados descritos en el fichero XML de cada una de las series de LIDC-IDRI. Teniendo en cuenta que no es necesario garantizar la representatividad de un tipo u otro de estructura nodular, se escogieron las series que no fueron utilizadas durante el proceso de análisis de los valores utilizados por el filtro LDM. La muestra inicial fue seleccionada utilizando una técnica de “Muestreo Accidental”. Las utilizadas para la validación de los resultados fueron: LIDC-IDRI-0001, LIDC-IDRI-0011, LIDC-IDRI-0012, LIDC-IDRI-0014, LIDC-IDRI-0016, LIDC-IDRI-0018, LIDC-IDRI-0021, LIDC-IDRI-0024, LIDC-IDRI-0027.

Para la realización de las pruebas al igual que en estudios similares (33), fueron eliminadas de las series, aquellas imágenes que no contenían la región pulmonar, con el objetivo de minimizar el tiempo de análisis y de no generar falsos positivos innecesarios. Como el volumen de imágenes a analizar es elevado y puede ser engorroso el proceso de contabilizar los indicadores necesarios para el cálculo de los valores de sensibilidad y los falsos positivos por imagen, se decide automatizar las pruebas. Para llevar a cabo este proceso se realizan los siguientes pasos:

Paso 1: Leer el fichero XML correspondiente a la serie de la imagen que se analiza, para identificar las estructuras que son nódulos pulmonares solitarios.

El fichero XML tiene una estructura definida para facilitar la lectura de los hallazgos detectados por los especialistas, durante la fase de diagnóstico médico, ver Anexo 1. Del conjunto de estructuras contenidas

CAPÍTULO 3: Análisis y valoración de los resultados obtenidos por el algoritmo propuesto para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

en el fichero se desecharon aquellas que miden menos de 3 mm o más de 30 mm, conjuntamente con aquellas, que a pesar de estar identificadas en el fichero, no fueron consideradas estructuras nodulares por los radiólogos que participaron en el proceso de diagnóstico. Las coordenadas de estas últimas estructuras se encuentran enmarcadas en las etiquetas: <nonNodule> </nonNodule>

Paso 2: Ejecutar el algoritmo.

Paso 3: Determinar correspondencia entre las estructuras señaladas en el XML y las estructuras resultantes del análisis realizado por el algoritmo en cuestión.

Para poder calcular los valores de sensibilidad y el número de falsos positivos por imagen al ejecutar el algoritmo desarrollado, es preciso determinar la correspondencia existente entre las estructuras descritas en el fichero XML y las identificadas por dichos algoritmos, es decir ver en qué medida son coincidentes.

A partir de la correlación existente entre las estructuras analizadas se contabiliza los valores de los indicadores: VP, FP y FN. El pseudocódigo utilizado para dicha operación se relaciona a continuación:

```
1  AreaDeIterseccion(centroCirculo1, centroCirculo2, circulo1Radio, circulo2Radio)
2  Inicio
3      distanciaCirculo1Circulo2 ← RaízCuadrada((centroCirculo1.X- centroCirculo2.X)^2
4          +(centroCirculo1.Y- centroCirculo2.Y)^2)
5      Si distanciaCirculo1Circulo2 > circulo1Radio + circulo2Radio Entonces
6          return Área de Intersección ← 0
7      FinSi
8      SiNo Si distanciaCirculo1Circulo2 ≤ |circulo1Radio - circulo2Radio|
9          and circulo1Radio ≥ circulo2Radio Entonces
10         return Área de Intersección ← Área del Circulo2
11     FinSi
12     SiNo Si distanciaCirculo1Circulo2 ≤ |circulo1Radio - circulo2Radio|
13         and circulo1Radio < circulo2Radio Entonces
14         return area de Intersección ← Área del Circulo1
15     FinSi
16     SiNo
17         phi ← arco coseno((circulo1Radio^2 + distanciaCirculo1Circulo2^2 - circulo2Radio^2)
18             / (2 * circulo1Radio^2 * distanciaCirculo1Circulo2)) * 2)
19         theta ← (arco coseno (circulo2Radio^2 + distanciaCirculo1Circulo2^2 - circulo1Radio^2)
20             / (2 * circulo2Radio^2 * distanciaCirculo1Circulo2)) * 2)
21         area1 ← 0.5 * theta * circulo2Radio^2 - 0.5 * circulo2Radio^2 * seno(theta)
22         area2 ← 0.5 * phi * circulo1Radio^2 - 0.5 * circulo1Radio^2 * seno(phi)
23         return areaIterseccion ← area1 + area2
24     FinSi
25 FinProcedimiento
```

Figura 8: Pseudocódigo del método para determinar la intersección entre las estructuras señalizadas en el XML y las devueltas por el algoritmo desarrollado

CAPÍTULO 3: Análisis y valoración de los resultados obtenidos por el algoritmo propuesto para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

Paso 4: Calcular el valor de sensibilidad alcanzado por el algoritmo; así como la cantidad promedio de falsos positivos por imagen médica analizada, a partir de los indicadores obtenidos en el paso 3.

3.3. Análisis de los resultados obtenidos

Con el objetivo de valorar los resultados obtenidos por el algoritmo desarrollado, a partir de la propuesta realizada en el capítulo anterior, se decide implementar dos de los algoritmos relacionados en el estado del arte, para así poder comparar los indicadores de sensibilidad y el promedio de falsos positivos por imagen médica analizada. Como se ha mencionado anteriormente no existe una técnica de segmentación por excelencia para la identificación de nódulos pulmonares solitarios. Los algoritmos escogidos para realizar la comparación fueron los relacionados en (5), teniendo en cuenta el análisis que se hace en esa investigación y los resultados alcanzados en la misma. De igual forma, es escogida la estrategia llevada a cabo en (65), a partir de los resultados de sensibilidad que se alcanza al utilizar el filtro LDM. Durante el estudio realizado en el estado del arte, se identificó que los algoritmos seleccionados y el algoritmo desarrollado están en igualdad de condiciones; teniendo en cuenta que la base de la segmentación realizada, es el uso de técnicas de umbralización. El algoritmo referido en (5) utiliza algunas técnicas en 3D pero estas fueron implementadas en 2D para poder establecer una comparación en igualdad de condiciones.

Para la validación de los resultados alcanzados se utilizaron 9 series, para un total de 766 imágenes, en las cuales se encontraban representadas 172 estructuras nodulares, analizadas desde un plano en 2D. Las estructuras analizadas varían en tamaño de 3-30 mm. A continuación se muestran los detalles.

Tabla 6: Resultados alcanzados por el algoritmo propuesto (AP) para la identificación de nódulos pulmonares solitarios conjuntamente con dos de los analizados en el estado del arte.

Indicadores	Serie LIDC-IDRI-0001			Serie LIDC-IDRI-0011			Serie LIDC-IDRI-0012		
	AP	(20)	(5)	AP	(20)	(5)	AP	(20)	(5)
VP	8	8	6	18	23	8	29	17	21
FP	474	1745	225	480	930	342	507	1036	372
FN	1	1	3	10	5	20	13	15	21
S:	0,89	0,89	0,67	0,64	0,82	0,29	0,69	0,53	0,50
FPI	5	17,45	2,25	4,85	9,39	3,45	4,88	9,96	3,58

CAPÍTULO 3: Análisis y valoración de los resultados obtenidos por el algoritmo propuesto para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

TI	100	99	104
-----------	------------	-----------	------------

	Serie LIDC-IDRI-0014			Serie LIDC-IDRI-0016			Serie LIDC-IDRI-0018		
Indicadores	AP	(20)	(5)	AP	(20)	(5)	AP	(20)	(5)
VP	6	4	4	15	13	12	15	14	13
FP	476	1012	364	498	980	185	433	853	289
FN	0	2	2	8	10	11	3	4	5
S:	1,00	0,67	0,67	0,65	0,57	0,52	0,83	0,78	0,72
FPI	5,17	11,00	3,96	5,03	9,90	1,87	4,61	9,07	3,07
TI	92			99			94		

	Serie LIDC-IDRI-0021			Serie LIDC-IDRI-0024			Serie LIDC-IDRI-0027		
Indicadores	AP	(20)	(5)	AP	(20)	(5)	AP	(20)	(5)
VP	10	8	9	10	7	8	16	13	12
FP	442	779	264	533	633	360	468	861	258
FN	2	4	3	3	6	5	5	8	9
S:	0,83	0,67	0,75	0,77	0,54	0,62	0,76	0,62	0,57
FPI	4,80	8,47	2,87	6,20	7,36	4,19	4,98	9,16	2,74
TI	92			86			94		

FPI: promedio de falsos positivos por imagen

TI: total de imágenes analizadas

Al ejecutar el algoritmo desarrollado se obtuvo un valor de sensibilidad de un 79% con un promedio de 5,02 falsos positivos por imagen. Para lograr estos resultados se utilizaron varias combinaciones de parámetros del filtro LDM conjuntamente con el filtro de Iris.

Luego de la ejecución del algoritmo propuesto en (20) se obtuvo un valor de sensibilidad de 68% con un promedio de 10,20 falsos positivos por imagen. El valor de sensibilidad alcanzado por este algoritmo con los datos utilizados, es inferior al relacionado en el estado del arte. Este resultado estuvo determinado porque el algoritmo LDM utilizando un solo umbral, no fue sensible a la detección de nódulos periféricos y algunos que no presentan marcados patrones de calcificación; mostrándose estos últimos como estructuras de densidad no tan representativa y representándose mucho más disgregados en las imágenes.

CAPÍTULO 3: Análisis y valoración de los resultados obtenidos por el algoritmo propuesto para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

El algoritmo propuesto en (5) obtuvo un valor de sensibilidad de 59% con un promedio de 3,11 falsos positivos por imagen. Los resultados alcanzados por este algoritmo son inferiores a los relacionados en el estado del arte porque independientemente que los datos utilizados no son los mismos, al utilizar como paso previo a la aplicación del filtro de Iris una umbralización adaptativa, aquellas estructuras de muy bajo valor de intensidad pueden no quedar representada en la imagen y por tanto no serán candidatas a nódulos pulmonares solitarios. Existen además otro conjunto de estructuras que no demuestran patrones definidos de circularidad (como es el caso de muchos nódulos periféricos) y por ello quedan eliminados durante la fase de reducción de falsos positivos. Sin embargo, es importante señalar que es inferior el número de falsos positivos con respecto al algoritmo propuesto en (20), teniendo en cuenta que el filtro de Iris es mucho más excluyente que el filtro LDM, al escoger como candidatas, solo las estructuras en las que el vector gradiente fluye hacia el centro de la misma.

3.4. Conclusiones del capítulo

- ✓ Se validaron los resultados alcanzados por el algoritmo propuesto en esta investigación para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax utilizando 9 series contenidas en The Lung Image Database Consortium Image Collection. Se obtuvo una sensibilidad de un 79% y un promedio de 5,02 falsos positivos por imagen médica, para estructuras de 3-30 mm de diámetro.
- ✓ Se contrastaron los resultados alcanzados por el algoritmo propuesto con respecto a dos de los algoritmos identificados durante el estudio del estado del arte y quedó evidenciada la factibilidad al utilizar la combinación de filtros LDM e Iris. Estas técnicas se enfocan en las características morfológico-radiográficas de dicha estructuras durante la fase de identificación y arrojan buenos resultados, LDM por su alta sensibilidad e Iris por su enfoque discriminatorio.

CONCLUSIONES

Durante el desarrollo de esta investigación se arribaron a las siguientes conclusiones, que evidencian el cumplimiento de los objetivos propuestos:

- ✓ Para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax se pueden utilizar diversas técnicas de segmentación; pero existe una tendencia al uso de la umbralización por sus buenos resultados y su bajo costo computacional.
- ✓ Utilizar la combinación del filtro LDM e Iris para la identificación de los nódulos pulmonares, posibilita alcanzar notables resultados de sensibilidad en la identificación de dichas estructuras, con un menor número de falsos positivos por imagen.
- ✓ Se desarrolló un algoritmo para la identificación de nódulos pulmonares solitarios que alcanza resultados favorables para ser utilizado por los imagenólogos como apoyo durante el proceso de diagnóstico médico.
- ✓ Se validaron los resultados obtenidos utilizando un conjunto de imágenes públicas en The Lung Image Database Consortium Image Collection (LIDC-IDRI) y se obtuvo una sensibilidad de un 79% para un promedio de 5,02 falsos positivos por imagen médica, para estructuras de 3-30 mm de diámetro; demostrando mayor aciertos que otras combinaciones de técnicas utilizadas por otros autores de trabajos similares.

RECOMENDACIONES

- ✓ Seleccionar un conjunto de datos mayor al escogido en la investigación, para optimizar y ajustar la combinación de parámetros utilizada para la ejecución del filtro LDM.
- ✓ Implementar el filtro de Iris en 3D con el objetivo de realizar un análisis volumétrico de las estructuras candidatas a ser nódulos pulmonares solitarios y así evitar el descarte de estructuras que pueden manifestar el patrón de aumento de la intensidad en profundidad.
- ✓ Desarrollar un algoritmo para la clasificación de nódulos pulmonares en benignos y malignos, con el objetivo de culminar el desarrollo del sistema de diagnóstico asistido por ordenador.

REFERENCIAS BIBLIOGRÁFICAS

1. NEUMOSUR, Suplemento. TOMOGRAFIA AXIAL COMPUTERIZADA DE TORAX: PRINCIPALES PATRONES RADIOLÓGICOS. PAPEL DE LA RMN EN EL ESTUDIO DE LA PATOLOGÍATORACICA. 2006.
2. CALZADO, A and GELEIJNS, J. Tomografía computarizada. Evolución, principios técnicos y aplicaciones. *Revista Física Médica*. 2011. Vol. 11, no. 3, p. 163–180.
3. MEDINA, Rubén and BELLERA, Jesús. Bases del Procesamiento de Imágenes Médicas. *Universidad de Los Andes, Facultad de Ingeniería, Grupo de Ingeniería Biomédica de la ULA (GIBULA)* [En línea]. Disponible en: http://www.saber.ula.ve/redtelemedicina/TallerTelemedicina/j_bellera-01.pdf
4. MAINIERI-HIDALGO, José A. and BRENES-DITTEL, Alejandro. Cáncer de Pulmón. *Asociación Costarricense de Neumología y Cirugía de Tórax*. 2003. Vol. 45, no. 1.
5. SUÁREZ CUENCA, Jorge Juan. *Desarrollo de un sistema de diagnóstico asistido por computador para detección de nódulos pulmonares en tomografía computarizada multicorte*. Tesis PhD. Santiago de Compostela, 2009.
6. CABANES VARONA, Lixaida, RODRÍGUEZ VÁZQUEZ, uan Carlos, GASSIOT NUÑO, Carlos de Jesús, RAMOS QUEVEDO, Adela, MORALES SÁNCHEZ, Lisvenia and LIMA GUERRA, Ailema. Carcinoma broncogénico: In fluencia de las demoras diagnósticas y terapéuticas en la supervivencia. Hospital «Hermanos Ameijeiras». 2003-2009. *Neumol Cir Torax*. 2011. Vol. 70, no. 2, p. 84–93.
7. BLÁZQUEZ VALLES, Beatriz and VARELA IZQUIERDO, Raúl. *DIAGNOSIS DECISION SUPPORT SYSTEMS*. Leganés, España : Departamento de Teoría de la Señal y Comunicaciones, Universidad Carlos III de Madrid, 2010.
8. COSTA MONTANÉ, Daniel Marino, PRADO LAGE, Yulien, LOZANO SALAZAR, Jorge Luis, PLASENCIA ASOREY, C. Carolina and RIESGO COSME, Yalili de la Caridad. Principales aspectos clinicoepidemiológicos del cáncer de pulmón. *MEDISAN*. 2011. Vol. 15.
9. DIRECCIÓN DE REGISTROS MÉDICOS Y ESTADÍSTICAS DE SALUD. *ANUARIO ESTADÍSTICO DE SALUD 2013* [En línea]. La Habana: MINISTERIO DE SALUD PÚBLICA, 2014. Disponible en: <http://files.sld.cu/dne/files/2014/05/anuario-2013-esp-e.pdf>
10. DOI, Kunio. *Computer-Aided Diagnosis in Medical Imaging: Historical Review, Current Status and Future Potential*. *Computerized Medical Imaging and Graphics*. March 2007. Vol. vol 31, no. 4-5, p. p. 198–211. DOI 10.1016/j.compmedimag.2007.02.002. /
11. EL-BAZ, Ayman, BEACHE, Garth M., GIMEL'FARB, Georgy, SUZUKI, Kenji, OKADA, Kazunori, ELNAKIB, Ahmed, SOLIMAN, Ahmed and ABDOLLAHI, Behnoush. Computer-Aided

- Diagnosis Systems for Lung Cancer: Challenges and Methodologies. *International Journal of Biomedical Imaging*. 2013. Vol. 2013, p. 1–46. DOI 10.1155/2013/942353.
12. BELLOTTI, Roberto, DE CARLO, F., GARGANO, Gianfranco, TANGARO, Sabina, CASCIO, Donato, CATANZARITI, Ezio, CERELLO, P., CHERAN, S.C., DELOGU, P., DE MITRI, Ivan, FULCHERI, Cchristian, GROSSO, Daniele, RETICO, Alessandra, SQUARCIA, S., TOMMASI, Eleonora and GOLOSIO, Bruno. A CAD system for modele detecction in low-dose lung CTs based on region growing and new active contour model. *Medical Physics*. 2007. DOI 10.1118/1.2804720.
 13. LE, Anh H. T., LIU, Brent and HUANG, H. K. Integration of computer-aided diagnosis/detection (CAD) results in a PACS environment using CAD–PACS toolkit and DICOM SR. *International Journal of Computer Assisted Radiology and Surgery*. 2009. Vol. 4, no. 4, p. 317–329.
 14. BOGONI, Luca, KO, Jane P., ALPERT, Jeffrey, ANAND, Vikram, FANTAUZZI, John, FLORIN, Charles H., WAN KOO, Chi, MASON, Derek, ROM, William, SHIAU, Maria, SALGANICOFF, Marcos and NAIDICH, David P. Impact of a Computer-Aided Detection (CAD) System Integrated into a Picture Archiving and Communication System (PACS) on Reader Sensitivity and Efficiency for the Detection of Lung Nodules in Thoracic CT Exams. *J Digit Imaging*. 2012. Vol. 25, no. 6. DOI 10.1007/s10278-012-9496-0.
 15. GOO, Jin Mo. A computer-aided diagnosis for evaluating lung nodules on chest CT: the current status and perspective. *Korean Jornal of Radiology*. 2011. Vol. 12, no. 2, p. 145–155. DOI 10.3348/kjr.2011.12.2.145.
 16. SOUTO BAYARRI, Miguel, SUÁREZ-CUENCA, Jorge Juan, TAHOCES, Pablo G, REVEL, Marie-Pierre, DELHAYE, Damien, CARREIRA, José M, REMY-JARDIN, Martine and REMY, Jacques. Automatic detection of pulmonary nodules: Evaluation of performance using two different MDCT scanners. *ournal of Biomedical Graphics and Computing*. 2012. Vol. 2, no. 2. DOI 10.5430/jbgc.v2n2p55.
 17. DOLEJŠÍ, Martin. Automatic Two-Step Detection of Pulmonary Nodules. DOI 10.1117/12.709161.
 18. FIRMINO, Macedo, MORAIS, Antônio H, MENDOÇA, Roberto M, DANTAS, Marcel R, HEKIS, Helio R and VALENTIM, Ricardo. Computer-aided detection system for lung cancer in computed tomography scans: Review and future prospects. *BioMedical Engineering OnLine*. 2014. Vol. 13, no. 41.
 19. FERNÁNDEZ PRADA, María, LLANOS MÉNDEZ, Aurora and GÓMEZ, Rebeca Isabel. Eficacia del diagnóstico asistido por ordenador aplicado al cribado de cáncer de mama. *Agencia de Evaluación de Tecnologías Sanitarias de Andalucía*. 2013.
 20. EJAZ, Naveed, JAVED, Shazia and SAJID, Zeeshan. Implementation of Computer Aided Diagnosis System for Lung Cancer Detection. *Lecture Notes on Software Engineering*. 2013. Vol. 1, no. 4, p. 364–369. DOI 10.7763/LNSE.2013.V1.78.

21. ZHAO, Binsheng, GAMSU, Gordon, GINSBERG, Michelle S., JIANG, Li and SCHWARTZ, Lawrence H. Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm. *JOURNAL OF APPLIED CLINICAL MEDICAL PHYSICS*. 2003. Vol. 4, no. 3, p. 248–260.
22. BAE, Kyongtae T., KIM, Jin-Sung, NA, Yong-Hum, GI KIM, Kwang and KIM, Jin-Hwan. Pulmonary Nodules: Automated Detection on CT Images with Morphologic Matching Algorithm—Preliminary Results1. *Radiology*. 2005. Vol. 236, p. 286–294. DOI 10.1148/radiol.2361041286.
23. *Cáncer de pulmón* [En línea]. Instituto Nacional del Cáncer de los Institutos Nacionales de la Salud de Estado Unidos, 2014. Disponible en: <http://www.cancer.gov/espanol/tipos/pulmon>
24. *Cáncer* [En línea]. Organización Mundial de la Salud (OMS), 2013. Disponible en: <http://www.who.int/mediacentre/factsheets/fs297/es/>
25. *aecce contra el cáncer* [En línea]. Asociación Española Contra el Cáncer- mejorando la vida de las personas, 2013. Disponible en: <https://www.aecc.es/Paginas/PaginaPrincipal.aspx>
26. BASTARRIKA, G, CANO, D, HERNÁNDEZ, C, ALONSO-BURGOS, A, GONZÁLEZ, I, VILLANUEVA, A, VIVASA, I and ZULUETAB, J. Detección y caracterización del nódulo pulmonar por tomografía computarizada multicorte. *Radiología*. 2007. Vol. 49, no. 4, p. 237–249.
27. GÓMEZ DE TERREROS CARO, F.J., GÓMEZ-ESTERN AGUILAR, C. and DISDIER VICENTE, C. ACTUALIZACIÓN EN EL NÓDULO PULMONAR SOLITARIO. *Neumosur*. 2008. Vol. 19, no. 4, p. 207–217.
28. WINER-MURAM, Helen T. The Solitary Pulmonary Nodule. *Radiology*. abril 2006. Vol. 239, no. 1, p. 34–49. DOI <http://dx.doi.org/10.1148/radiol.2391050343>.
29. KERIK, Nora E., ALEXANDERSON-ROSAS, Erick, FERMON-SCHWAYCER, Salomón, RUÍZ-RAMÍREZ, O. Leonel and ALONSO-LÓPEZ, Fred. Utilidad de la tomografía por emisión de positrones (PET) en el cáncer pulmonar. *Gaceta médica de México*. 2004. Vol. 140, no. 5.
30. SOUTOA, M., TAHOCES, P.G., SUÁREZ CUENCA, J. J., LADO, M.J., REMY-JARDIN, M., REMY, J. and VIDAL, J.J. Detección automática de nódulos pulmonares en tomografía computarizada. Un estudio preliminar*. *Radiología*. 2008. Vol. 50, p. 387–392.
31. BAE, Kyongtae T., KIM, Jin-Sung, NA, Yong-Hum, GI KIM, Kwang and KIM, Jin-Hwan. Pulmonary Nodules: Automated Detection on CT Images with Morphologic Matching Algorithm—Preliminary Results1. *Radiology*. 2005. Vol. 236, p. 286–294. DOI 10.1148/radiol.2361041286.
32. ERASMUS, Jeremmy F., CONNOLLY, Jhon E., MCADAMS, H. Page and ROGGLI, Victor L. Solitary Pulmonary Nodules: Part I: Morphologic Evaluation for Differentiation of Benign and Malignant Lesions. *RadioGraphics*. 2000. Vol. 20, no. 1, p. 59–66.

33. SOUTOA, M., TAHOSES, P.G., SUÁREZ CUENCA, J. J., LADO, M.J., REMY-JARDIN, M., REMY, J. and VIDAL, J.J. Detección automática de nódulos pulmonares en tomografía computarizada. Un estudio preliminar*. *Radiología*. 2008. Vol. 50, p. 387–392.
34. JEONG, Y J, YI, CA and LEE, KS. Nódulos pulmonares solitarios: detección, caracterización y guías para su diagnóstico y tratamiento*. 2008. Vol. 50, no. 3. DOI 10.1016/S0033-8338(08)71964-7.
35. COSTA MONTANÉ, Daniel Marino, PRADO LAGE, Yulien, LOZANO SALAZAR, Jorge Luis, PLASENCIA ASOREY, C. Carolina and RIESGO COSME, Yalili de la Caridad. Principales aspectos clinicoepidemiológicos del cáncer de pulmón. *MEDISAN*. 2011. Vol. 15.
36. RAMÍREZ GIRALDO, Juan Carlos, ARBOLEDA CLAVIJO, Carolina and MCCOLLOUGH, Cynthia H. Tomografía computarizada por rayos X: fundamentos y actualidad. *Revista Ingeniería Biomédica*. 2008. Vol. 2, no. 4, p. 16–17.
37. RAMÍREZ PATIÑO, Juan Fernando, ANDREA ISAZA, Jérica, MARIAKA, Isabela and VÉLEZ ZEA, Jaime Andrés. Unidades Hounsfield como instrumento para la evaluación de la desmineralización ósea producida por el uso de exoprótesis. *Engineering Faculty, University of Antioquia*. 2013. Vol. 66, p. 159–167.
38. DOI, Kunio. Current status and future potential of computer-aided diagnosis in medical imaging. *The British Journal of Radiology* [online]. 2005. Vol. 78. DOI 10.1259/bjr/82933343. Available from: http://tzklabo.met.nagoya-u.ac.jp/CHICAGO/KRL_Papers/KunioDoi/24_CurrentStatusCAD.pdf
39. CHENG, H.D., CAI, Xiaopeng, CHEN, Xiaowei, HU, Liming and LOU, Xueling. Computer-aided detection and classification of microcalcifications in mammograms: a survey. *Pattern Recognition*. 2003. Vol. 36, no. 12, p. 2967–2991. DOI DOI: 10.1016/S0031-3203(03)00192-4.
40. LARA TAMBURRINO, María Del Carmen and ZORRILLAB, Álvaro Olmedo. Detección temprana y diagnóstico del cáncer mamario. *Revista de la Facultad de Medicina de la UNAM*. 2011. Vol. 54, no. 1.
41. LLOBET, Rafael, PÉREZ-CORTÉS, Juan C., TOSELLI, Alejandro H. and JUAN, Alfons. Computer-aided detection of prostate cancer. *International Journal of Medical Informatics*. 2006. Vol. 76, p. 557–556. DOI doi:10.1016/j.ijmedinf.2006.03.001.
42. MORADI, Mehdi, MOUSAVI, Parvin and ABOLMAESUMI, Purang. Computer-Aided Diagnosis of Prostate Cancer with Emphasis on Ultrasound-Based Approaches: A Review. *Ultrasound in Medicine & Biology*. 2007. Vol. 33, no. 7, p. 1010–1028. DOI 10.1016/j.ultrasmedbio.2007.01.008.
43. CHMID-SAUGEON, Philippe, GUILLOD, Joel and THIRAN, Jean-Philippe. Towards a computer-aided diagnosis system for pigmented skin lesions. *Computerized Medical Imaging and Graphics*. 2003. Vol. 27, p. 65–78.

44. BARZEGARI, Masoomeh, GHANINEZHAD, Haiedeh, MANSOORI, Parisa, TAHERI, Arash, NARAGHI, Zahra S and ASGARI, Masood. Computer-aided dermoscopy for diagnosis of melanoma. *BMC Dermatology*. 2005. Vol. 5, no. 8. DOI 10.1186/1471-5945-5-8.
45. FERREIRA DA SILVA SOUSA, João Rodrigo, CORREA SILVA, Aristófanés, CARDOSO DE PAIVA, Anselmo and ACATAUASSÚ NUNES, Rodolfo. Methodology for automatic detection of lung nodules in computerized tomography images. *computer methods and programs in biomedicine*. 2010. Vol. 98, no. 1, p. 1–14. DOI 10.1016/j.cmpb.2009.07.006.
46. BRANDMAN, Scott and KO, Jane P. Pulmonary Nodule Detection, Characterization, and Management With Multidetector Computed Tomography. *J Thorac Imaging*. 2011. Vol. 26, no. 2, p. 90–105.
47. ARMATO III, Samuel G., GIGER, Maryellen L. and MACMAHON, Heber. Automated detection of lung nodules in CT scans: Preliminary results. *American Association of Physicists in Medicine*. 2001. Vol. 28, p. 1552. DOI <http://dx.doi.org/10.1118/1.1387272>.
48. RETICO, Alessandra, DELOGU, Pasquale, FANTACCI, MariaEvelina, GORI, I and PREITE MARTINEZ, A. Lung nodule detection in low-dose and thin-slice computed tomography. 2008. Vol. 38, p. 525–534. DOI 10.1016/j.compbiomed.2008.02.001.
49. GURCAN, Metin N., SAHINER, Berkman, PETRICK, Nicholas, CHAN, Heang-Ping, KAZEROONI, Ella A., CASCADE, Philip N. and HADJIISKI, Lubomir. Lung nodule detection on thoracic computed tomography images: Preliminary evaluation of a computer-aided diagnosis system. *Medical Physics*. 2002. Vol. 29. DOI <http://dx.doi.org/10.1118/1.1515762>.
50. KO, Jane P and BETKE, Margrit. Chest CT: Automated Nodule Detection and Assessment of Change over Time—Preliminary Experience. *Radiology*. 2001. Vol. 218, no. 1, p. 267–273.
51. Lankton and Tannenbaum. Localizing Region-Based Active Contours. *IEEE TRANSACTIONS ON MEDICAL IMAGING*. 2008. Vol. 17, no. 11, p. 2029 – 2039. DOI 10.1109/TIP.2008.2004611.
52. LI, Hua, THORSTAD, Wade L., BIEHL, Kenneth J., LAFOREST, Richard, SU, Yi, SHOGHI, Koresh I., DONNELLY, Eric D., LOW, Daniel A. and LU, Wei. A novel PET tumor delineation method based on adaptive region-growing and dual-front active contours. *Medical Physics*. 2008. Vol. 35, no. 8, p. 3711–3721. DOI 10.1118/1.2956713.
53. BELLOTTI, R., DE CARLO, F., GARGANO, G., TANGARO, S., CASCIO, D., CATANZARITI, E., CERELLO, P., CHERAN, S. C., DELOGU, P., DE MITRI, I., FULCHERI, C., GROSSO, D., RETICO, A., SQUARCIA, S., TOMMASI, E. and GOLOSIO, Bruno. A CAD system for nodule detection in low-dose lung CTs based on region growing and a new active contour model. *Medical Physics*. 2007. Vol. 34, no. 12. DOI 10.1118/1.2804720_.

54. SHAHVARAN, Zahra, KAZEMI, Kamran, SADEGH HELFROUSH, Mohammad and JAFARIAN, Nassim. Region-based Active Contour Model based on Markov Random Field to Segment Images with Intensity Non-Uniformity and Noise. *Journal of Medical Signals and Sensors*. 2012. Vol. 2, no. 1, p. 17–24.
55. SUÁREZ CUENCA, Jorge Juan, TAHOSES, Pablo G, LADO, M.J., SOUTO, M. and VIDAL, J.J. Detección automática del nódulo pulmonar subpleural basada en el empleo de contornos activos. *CISTI*. 2008. Vol. 1, p. 585–591.
56. GURCAN, Metin N., SAHINER, Berkman, PETRICK, Nicholas, CHAN, Heang-Ping, KAZEROONI, Ella A., CASCADE, Philip N. and HADJIISKI, Lubomir. Lung nodule detection on thoracic computed tomography images: Preliminary evaluation of a computer-aided diagnosis system. *Medical Physics*. 2002. Vol. 29. DOI <http://dx.doi.org/10.1118/1.1515762>.
57. ACHARYA, Tinku and RAY, Ajoy K. *Image Processing. Principles and Applications*. A JOHN WILEY & SONS, 2005. ISBN 13 978-0-471-71998-4.
58. COTO, Ernesto. Métodos de Segmentación de Imágenes Médicas. *Universidad Central de Venezuela. Facultad de Ciencias* [En línea]. 2003. Disponible en: <http://lcg.ciens.ucv.ve/~ernesto/nds/CotoND200305.pdf>
59. GONZÁLEZ,, Rafael C and WOODS, Richard E. *Digital Image Processing*. 2a ed. Prentice Hall, 2002. ISBN 0-201-18075-8.
60. GONZÁLEZ RAMOS, Ana, MARTÍNEZ, Francisco Javier, PERNÍA, Alpha Verónica, ALBA ELÍAS, Fernando, CASTEJÓN LIMAS, Manuel, ORDIERES MERÉ, Joaquín and VERGARA GONZÁLEZ, Eliseo. *Técnicas y Algoritmos Básicos de Visión Artificial*. Universidad de la Rioja. Servicio de Publicaciones. 2006. ISBN 84-689-9345-X.
61. JC, Ross, RS, Estepar, A, Diaz, C-F, Westin, R, Kikinis, E, Silverman and G, Washko. Lung extraction, lobe segmentation and hierarchical region assessment for quantitative analysis on high resolution computed tomography images. *Springer*. 2009. Vol. 5762, p. 690–698.
62. VAN RIKXOORT, EM, DE HOOP, B, VAN DE VORST, S, PROKOP, M and VAN GINNEKEN, B. Automatic segmentation of pulmonary segments from volumetric chest ct scans. 2009. Vol. 28, no. 4, p. 621–630.
63. PU, J, ROOS, J, YI, CA, NAPEL, S, RUBIN, GD and PAIK, DS. Adaptive border marching algorithm: Automatic lung segmentation on chest CT images. 2008. Vol. 32, p. 452–462.
64. WEI, Q, HU, Y, GELFAND, G and MACGREGOR, JH. Segmentation of lung lobes in high-resolution isotropic ct images. 2009. Vol. 56, no. 5, p. 1383–1393.

65. EJAZ, Naveed, JAVED, Shazia and SAJID, Zeeshan. Implementation of Computer Aided Diagnosis System for Lung Cancer Detection. *Lecture Notes on Software Engineering*. 2013. Vol. 1, no. 4, p. 364–369. DOI 10.7763/LNSE.2013.V1.78.
66. ARMATO, SG and SENSAKOVIC, WF. Automated lung segmentation for thoracic CT: Impact on computer-aided diagnosis. 2004. Vol. 11, p. 1011–1021.
67. MOLINA, R. *Introducción al Procesamiento y Análisis de Imágenes Digitales*. Universidad de Granada, Departamento de Ciencias de la Computación e IA, 1998.
68. JIANZHUANG, Liu, XIAN, Xidian Univ, WENQING, Li and TIAN YUPENG. Automatic thresholding of gray-level pictures using two-dimension Otsu method. *IEEE*. 1991. Vol. 1, p. 325 – 327. DOI 10.1109/CICCAS.1991.184351.
69. An Improved Image Segmentation Algorithm Based on Two- Dimensional Otsu Method. *Information Science Letters*. 2012. Vol. 1, no. 2, p. 77–83.
70. G. ARMATO III, Samuel. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans. *Medical Physics*. 2011. Vol. 38, no. 2, p. 915–931. DOI 10.1118/1.3528204.
71. MCNITT-GRAY, Michael F, ARMATO, Samuel G, MEYER, Charles R, REEVES, Anthony P, MCLENNAN, Geoffrey, PAIS, Richie C, FREYMANN, John, BROWN, Matthew S, ENGELMANN, Roger M, BLAND, Peyton H, LADERACH, Gary E, PIKER, Chris, GUO, Junfeng, TOWFIC, Zaid, QING, David P-Y, YANKELEVITZ, David F, ABERLE, Denise R, VAN BEEK, Edwin J R, MACMAHON, Heber, KAZEROONI, Ella A, CROFT, Barbara Y and CLARKE, Laurence P. The lung Image Database Consortium (LIDC) data collection process for nodule detection and annotation. *Academia de Radiología*. 2007. Vol. 14, no. 12.
72. MCNITT-GRAY, Michael F, ARMATO, Samuel G, MEYER, Charles R, REEVES, Anthony P, MCLENNAN, Geoffrey, PAIS, Richie C, FREYMANN, John, BROWN, Matthew S, ENGELMANN, Roger M, BLAND, Peyton H, LADERACH, Gary E, PIKER, Chris, GUO, Junfeng, TOWFIC, Zaid, QING, David P-Y, YANKELEVITZ, David F, ABERLE, Denise R, VAN BEEK, Edwin J R, MACMAHON, Heber, KAZEROONI, Ella A, CROFT, Barbara Y and CLARKE, Laurence P. The Lung Image Database Consortium (LIDC). An Evaluation of Radiologist Variability in the Identification of Lung Nodules on CT Scans. *Academia de Radiología*. 2007. Vol. 14, no. 11, p. 1409–1421.
73. *EmguCV* [En línea]. 2014. [Accedido 4 Abril 2014]. Disponible en: http://www.emgu.com/wiki/index.php/Main_Page
74. BRADSKI, Gary and KAEHLER, Adrian. *Computer Vision with the OpenCV Library* [En línea]. United States of America, 2008. [Accedido 4 Abril 2014]. ISBN 978-0-596-51613-0. Disponible en: <http://www.cs.haifa.ac.il/~dkeren/ip/OReilly-LearningOpenCV.pdf>

75. *Visual C# resources* [En línea]. 2014. [Accedido 27 Mayo 2014]. Disponible en: <http://msdn.microsoft.com/en-us/vstudio/hh341490.aspx>
76. SOUTO, M., TAHOCES, P.G., SUÁREZ CUENCA, J. J., LADO, M.J., REMY-JARDIN, M., REMY, J. and VIDAL, J.J. Detección automática de nódulos pulmonares en tomografía computarizada. Un estudio preliminar*. *Radiología*. 2008. Vol. 50, p. 387–392.
77. EJAZ, Naveed, JAVED, Shazia and SAJID, Zeeshan. Implementation of Computer Aided Diagnosis System for Lung Cancer Detection. *Lecture Notes on Software Engineering*. 2013. Vol. 1, no. 4, p. 364–369. DOI 10.7763/LNSE.2013.V1.78.
78. SÁNCHEZ DIEGUEZ, Maikel and VALLÉS GÁMEZ, Antonio Enrique. *Framework CALIB. Procesos de lectura, procesamiento y visualización de imágenes médicas*. Tesis de Grado. Ciudad de La Habana: Universidad de las Ciencias Informáticas, Facultad 7, 2010.
79. PITA FERNÁNDEZ, S and PÉRTEGAS DÍAZ, S. Pruebas diagnósticas: Sensibilidad y especificidad. *Cuadernos de Atención Primaria*. 2003. Vol. 10, p. 120–124.
80. CUEVAS RENAUD, Corina and ALEJO MARTÍNEZ, Amalia. *Sensibilidad y Especificidad de una prueba* [En línea]. 2010. Universidad Nacional Autónoma de México, Facultad de Psicología. Disponible en: <http://www.psicol.unam.mx/Investigacion2/pdf/SENSIBILIDAD%20Y%20ESPECIFICIDAD.pdf>
81. FLORKOWSKI, Christopher M. Sensitivity, Specificity, Receiver-Operating Characteristic (ROC) Curves and Likelihood Ratios: Communicating the Performance of Diagnostic Tests. *The Clinical Biochemist Reviews*. 2008. Vol. 29, no. 1, p. 83–87.

ANEXOS

Anexo 1: Estructura del XML correspondiente a cada una de las series de imágenes médicas que se encuentran publicadas en “The Lung Image Database Consortium Image Collection”

```
<?xml version="1.0" encoding="iso-8859-1"?>
  Version header
  <LidcReadMessage xmlns="http://www.nih.gov"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.nih.gov
  http://troll.rad.med.umich.edu/lidc/LidcReadMessage.xsd">
    links to schema and xsd file
    <ResponseHeader>
      Beginning of the response header
      <Version>1.7</Version>
      Version Number
      <DateRequest>2000-01-01</DateRequest>
      Date of the original request to read the series
      <TimeRequest>00:00:00</TimeRequest>
      Time of that request

      <ServicingSite>anonymous</ServicingSite>
      Code for the site performing the read - set to anonymous
      <TaskDescription>Second unblinded read</TaskDescription>
      Code for the reading task being requested (in this case, it is the
      second phase, unblinded read)

      <DateService>2000-01-01</DateService>
      Date that the request was serviced (completed)
      <TimeService>00:00:00</TimeService>
      Time that the request was serviced
      <ResponseDescription>1 - Reading complete</ResponseDescription>
      Used for internal communication; only those files marked as "1-Reading
      complete" will be part of public release.
      <SeriesInstanceUID>1.3.6.1.4.1.9328.50.3.1888</SeriesInstanceUID>
      The DICOM field Series Instance UID (0020,000E)
    </ResponseHeader>
      End of response header
    <readingSession>
      A reading session consists of a set of markings done by a single reader at
      a single phase (either the blinded or unblinded phase - in these files only
      the unblinded read results are reported)
      <annotationVersion>3.12</annotationVersion>
      annotation version
      <servicingRadiologistID></servicingRadiologistID>
      Radiologist id - set to anonymous
      <unblindedReadNodule>
        Beginning of nodule marking information
        <noduleID>4983</noduleID>
        Nodule identifier - unique id for nodule
        <characteristics>
```

For a nodule > 3 mm, each reader is asked to subjectively assess the nodule's characteristics as described on LIDC CDE page <http://cdebrowser.nci.nih.gov/CDEBrowser/> and as follows:

<subtlety>5</subtlety>
Radiologist assessment of nodule subtlety on 1-5 scale
 (1 extremely subtle
 2
 3
 4
 5 obvious)
 <internalStructure>1</internalStructure>
Radiologist assessment of nodule internal structure
 (1 soft tissue,
 2 fluid,
 3 fat,
 4 air)
 <calcification>6</calcification>
Radiologist assessment of internal calcification of nodule
 (1 Popcorn Popcorn Appearance
 2 Laminated Laminated Appearance
 3 Solid Solid Appearance
 4 Non-Central Non-Central Appearance
 5 Central Central Calcification
 6 Absent)
 <sphericity>4</sphericity>
Radiologist assessment of shape of nodule in terms of its roundness/sphericity with only 3 terms defined:
 (1 Linear Linear Appearance
 2
 3 Ovoid Ovoid Appearance
 4
 5 Round Round Appearance)
 <margin>4</margin>
Radiologist assessment of nodule margin on a 1-5 scale with only the extreme values explicitly defined:
 (1 Poorly Defined Poorly Defined
 5 Sharp Sharp Margin)
 <lobulation>1</lobulation>
Radiologist assessment of nodule lobulation on a 1-5 scale with only the extreme values explicitly defined:
 (1 No Lobulation No Lobulation
 5 Marked Marked Lobulation)
 <spiculation>1</spiculation>
Radiologist assessment of nodule spiculation on a 1-5 scale with only the extreme values explicitly defined:
 (1 No Spiculation No Spiculation
 5 Marked Marked Spiculation)
 <texture>5</texture>
Radiologist assessment of nodule internal texture with only 3 terms defined:
 (1 Non-Solid/Ground Glass Opacity
 2
 3 Part Solid/Mixed
 4
 5 Solid Texture)
 <malignancy>5</malignancy>

Radiologist subjective assessment of likelihood of malignancy of this nodule (ASSUMING 60-year-old male smoker)

- 1 Highly Unlikely for Cancer
- 2 Moderately Unlikely for Cancer
- 3 Indeterminate Likelihood
- 4 Moderately Suspicious for Cancer
- 5 Highly Suspicious for Cancer)

</characteristics>

End of characteristics section

<roi>

<roi> signifies beginning of region of interest description

roi is a boundary describing either:

(a) a single voxel for nodules < 3 mm

(b) a complete contour of the nodule that describes the first voxel just outside the nodule. The format is one where the z position is first specified (using imageZposition AND imageSOP_UID), then the coordinates of connected boundary points within each x-y plane.

NOTE: x-y plane convention is that (1,1) is upper left

<imageZposition>-410.350006</imageZposition>

The third part of the tuplet in DICOM field (0020,0032) to describe z location of this slice on which the nodule is visualized and contoured

<imageSOP_UID>1.3.6.1.4.1.9328.50.3.1892</imageSOP_UID>

DICOM field (0008,0018) of that location; Please note that this also specifies the name of corresponding image file (1.3.6.1.4.1.9328.50.3.1892.dcm). The user can open that image file to see the image on which the nodule is visualized.

<inclusion>TRUE</inclusion>

Element to describe whether the following voxels are to be included (value = TRUE) or excluded (value=FALSE)

<edgeMap>

Element to describe that this is beginning of the edge map - the collection of these points make up the boundary of the nodule

<xCoord>412</xCoord>

x coordinate of edge map point

<yCoord>312</yCoord>

y coordinate of edge map point

</edgeMap>

end of edge map point

<edgeMap>

next edge map point - NOTE that this nodule has more than one point, indicating that this is a nodule > 3 mm

<xCoord>412</xCoord>

x coordinate of point

<yCoord>311</yCoord>

y coordinate of point

</edgeMap>

End of edge map - NOTE that this nodule has more than one point, indicating that this is a nodule > 3 mm

<edgeMap>

<xCoord>412</xCoord>

<yCoord>310</yCoord>

</edgeMap>

<edgeMap>

```
<xCoord>412</xCoord>
<yCoord>309</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>412</xCoord>
  <yCoord>308</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>411</xCoord>
  <yCoord>307</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>411</xCoord>
  <yCoord>306</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>411</xCoord>
  <yCoord>305</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>410</xCoord>
  <yCoord>305</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>409</xCoord>
  <yCoord>306</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>409</xCoord>
  <yCoord>307</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>409</xCoord>
  <yCoord>308</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>409</xCoord>
  <yCoord>309</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>410</xCoord>
  <yCoord>310</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>410</xCoord>
  <yCoord>311</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>411</xCoord>
  <yCoord>312</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>412</xCoord>
  <yCoord>312</yCoord>
</edgeMap>
</roi>
```

end of the roi definition for that z position (Slice)

<roi>

the roi continues on to another image

<imageZposition>-411.100006</imageZposition>

The third part of the tuple in DICOM field (0020,0032) to describe z location of this slice on which the nodule is visualized and contoured

<imageSOP_UID>1.3.6.1.4.1.9328.50.3.1893</imageSOP_UID>

DICOM field (0008,0018) of that location

<inclusion>TRUE</inclusion>

<edgeMap>

<xCoord>412</xCoord>

<yCoord>313</yCoord>

</edgeMap>

<edgeMap>

<xCoord>412</xCoord>

<yCoord>312</yCoord>

</edgeMap>

<edgeMap>

<xCoord>413</xCoord>

<yCoord>311</yCoord>

</edgeMap>

<edgeMap>

<xCoord>413</xCoord>

<yCoord>310</yCoord>

</edgeMap>

<edgeMap>

<xCoord>413</xCoord>

<yCoord>309</yCoord>

</edgeMap>

<edgeMap>

<xCoord>413</xCoord>

<yCoord>308</yCoord>

</edgeMap>

<edgeMap>

<xCoord>413</xCoord>

<yCoord>307</yCoord>

</edgeMap>

<edgeMap>

<xCoord>412</xCoord>

<yCoord>306</yCoord>

</edgeMap>

<edgeMap>

<xCoord>411</xCoord>

<yCoord>305</yCoord>

</edgeMap>

<edgeMap>

<xCoord>410</xCoord>

<yCoord>305</yCoord>

</edgeMap>

<edgeMap>

<xCoord>409</xCoord>

<yCoord>306</yCoord>

</edgeMap>

<edgeMap>

<xCoord>409</xCoord>


```
<yCoord>307</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>409</xCoord>
  <yCoord>308</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>409</xCoord>
  <yCoord>309</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>409</xCoord>
  <yCoord>310</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>409</xCoord>
  <yCoord>311</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>410</xCoord>
  <yCoord>312</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>411</xCoord>
  <yCoord>313</yCoord>
</edgeMap>
<edgeMap>
  <xCoord>412</xCoord>
  <yCoord>313</yCoord>
</edgeMap>
</roi>
<roi>
  <imageZposition>-411.850006</imageZposition>
  <imageSOP_UID>1.3.6.1.4.1.9328.50.3.1894</imageSOP_UID>
  <inclusion>TRUE</inclusion>
  <edgeMap>
    <xCoord>413</xCoord>
    <yCoord>315</yCoord>
  </edgeMap>
  <edgeMap>
    <xCoord>414</xCoord>
    <yCoord>314</yCoord>
  </edgeMap>
  <edgeMap>
    <xCoord>415</xCoord>
    <yCoord>314</yCoord>
  </edgeMap>
</roi>
```