

ARTÍCULO ORIGINAL

Algoritmo para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax**Algorithm for identification of solitary pulmonary nodules in chest tomography scan images**

**Arelys Rivero Castro,^I Yoel Rivera Suárez,^{II} Yosvani Borges González,^{III}
Yoanny Naranjo Gorrín^{IV}**

^IIngeniera. Universidad de las Ciencias Informáticas, Cuba. E-mail: arcastro@uci.cu

^{II}Ingeniero. Universidad de las Ciencias Informáticas, Cuba. E-mail:

yrsuarez@uci.cu

^{III}Ingeniero. Ministerio de Comunicaciones, Cuba. E-mail: yosvani.borges@mic.cu

^{IV}Ingeniero. Hospital Pediátrico William Soler, Cuba. E-mail:

yoanny@cardiows.sld.cu

RESUMEN

La identificación del cáncer de pulmón en fases iniciales ha sido en los últimos años una tarea priorizada de la comunidad científica. Esta enfermedad representa la primera causa de muerte en el varón y la tercera después del cáncer de colon y mama en la mujer. La realización de estudios imagenológicos contribuye a la detección temprana de esta enfermedad. El elevado volumen de imágenes generado por los equipos médicos provoca la revisión de mucha información para emitir un diagnóstico médico. Con frecuencia se requiere la valoración de varios especialistas para llegar a un diagnóstico acertado, retardando el proceso de atención al paciente.

En la presente investigación se exponen los resultados obtenidos al desarrollar un algoritmo utilizando métodos de procesamiento de imágenes, para la identificación de nódulos pulmonares solitarios. La utilización de sistemas que dirigen la atención de los especialistas a regiones candidatas en la imagen, proporcionando una segunda opinión en la interpretación de los resultados, pudiera mejorar la consistencia y agilizar el proceso de diagnóstico. 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 77.78 % de acierto en la detección de nódulos pulmonares solitarios.

Palabras clave: algoritmos de identificación, cáncer de pulmón, nódulos pulmonares solitarios, procesamiento de imágenes médicas, tomografía de tórax.

ABSTRACT

The identification of lung cancer at early stages has been in recent years a prioritized task for the scientific community. This disease is the leading cause of death in men and the third after the colon and breast cancer in women. Performing imaging studies contributes to the early detection of this disease. The high volume of images generated by medical equipment leads to reviewing much information to issue a medical diagnosis. Often are required the assessment of several specialists to reach an accurate diagnosis, slowing the process of patient care.

In the present investigation are exposed the results obtained to develop an algorithm using image processing methods for the identification of solitary pulmonary nodules. The use of systems that direct the attention of specialists to candidate regions in the image, providing a second opinion in the interpretation of results could improve consistency and agility in the diagnostic process. The results obtained by the developed algorithm were compared with annotations in images published in The Lung Image Database Consortium Image Collection (LIDC-IDRI) and was obtained 77.78 % accuracy in the detection of solitary pulmonary nodules.

Key words: identifying algorithm, lung cancer, solitary pulmonary nodules, medical image processing, chest tomography.

INTRODUCCIÓN

La medicina es uno de los sectores que más se ha beneficiado por la influencia de las nuevas tecnologías en la sociedad; principalmente aquellas que están destinadas al diagnóstico por imágenes, incrementando la capacidad de análisis y el estudio de diversas enfermedades. La imagenología médica, comprende la realización de exámenes diagnósticos y terapéuticos, en los cuales se utilizan equipos que generan imágenes del cuerpo humano, permitiendo la visualización de estructuras internas y tejido patológico, posibilitando avanzar en el tratamiento de alguna enfermedad sin necesidad de intervenir quirúrgicamente al paciente.¹

La creación de técnicas imagenológicas como la ecografía, la tomografía computarizada (TC),^{2,3} la resonancia magnética o la medicina nuclear, situaron a la Radiología, a la cabecera de las especialidades médicas. Desde su adopción, ha marcado un punto crucial en el desarrollo de las ciencias médicas.

Las imágenes generadas por los equipos médicos de estas modalidades poseen dos características fundamentales que la magnifican con respecto a las imágenes de radiografía analógica, siendo estas la resolución espacial¹ (cantidad de píxeles por pulgada o cm)⁴ y la densidad o profundidad (niveles de grises que se pueden representar),⁵ que posee la imagen.⁶

Entre las múltiples ventajas que ofrecen las imágenes digitales sobre las analógicas se destacan el bajo uso de dosis radiante, la eliminación del uso de líquidos

reveladores y fijadores, el fácil almacenamiento y acceso inmediato; así como la posibilidad de modificar las características de las imágenes,⁷⁻⁹ facilitando la detección de hemorragias, lesiones de órganos internos, e incluso para la detección de enfermedades de carácter patológico como es el caso del cáncer de pulmón.¹⁰⁻¹²

Teniendo en cuenta el impacto del cáncer de pulmón en la población cubana y la inexistencia de un sistema informático que facilite a los radiólogos la identificación de estructuras nodulares; el objetivo de esta investigación es desarrollar un algoritmo que señale a los especialistas aquellas estructuras que representen una lesión de este tipo. De esta forma se pueden aumentar o contribuir con los altos valores de aciertos durante el proceso de diagnóstico médico. La valoración conjunta de los resultados alcanzados por el algoritmo y el criterio médico de los especialistas en imagenología, puede repercutir positivamente en el diagnóstico temprano y certero del cáncer de pulmón en la Isla.

MATERIALES Y MÉTODOS

Durante el desarrollo de la investigación se utilizaron los siguientes métodos científicos:

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 cómo se ha comportado el desarrollo de software en torno a la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax.

Analítico Sintético: para descomponer el problema de la investigación en elementos concretos de la solución.

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

Experimento: para realizar la validación de los resultados obtenidos por el algoritmo de identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax a partir de las imágenes diagnosticadas que se encuentran disponibles en The Lung Image Database Consortium Image Collection (LIDC-IDRI).

Herramienta y tecnologías informáticas empleadas

C# 4.0: es un lenguaje de programación orientado a objetos y permite a los desarrolladores crear una amplia gama de aplicaciones que se ejecuten en la plataforma .NET Framework, su sintaxis es sencilla, fácil de utilizar y permite los conceptos de encapsulación, herencia y polimorfismo.¹³

Visual Studio 2012: es un Entorno de Desarrollo Integrado (IDE, por sus siglas en inglés), que soporta varios lenguajes de programación incluido Visual C#. Incluye herramientas que simplifican todo el proceso de desarrollo de aplicaciones, de principio a fin. Permite realizar una administración del ciclo de vida de las aplicaciones e incorpora otras pruebas que ayudan a garantizar la calidad del código en todo momento.¹⁴

Librería EmguCV 2.4.2: EmguCV es un wrapper de la librería de procesamiento de imágenes OpenCV que permite realizar invocaciones directas a funciones de OpenCV.^{15,16} Es una librería de código abierto y multiplataforma diseñada con el lenguaje de programación C#. 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.^{17,18}

Librería CALIB: es una solución de software para llevar a cabo los procesos de lectura, procesamiento y visualización de imágenes médicas, provenientes de múltiples equipos y de diversas modalidades.¹⁹ Le permite a los desarrolladores visualizar las imágenes de tomografía en formato DICOM. Fue desarrollada en Centro de Informática Médica perteneciente a la Universidad de las Ciencias Informáticas.

Sistemas para Diagnóstico Asistido por Ordenador

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 patológico² 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 TC posibilita la detección de un alto porcentaje de nódulos pulmonares con mayor precisión que otras técnicas imagenológicas,^{20,21} debido a su calidad, determinada por su alta resolución y las especificidades que presupone la utilización de los rayos X.

La detección temprana del cáncer de pulmón es relevante, en la actualidad más del 90 % de la identificación de nódulos pulmonares se realiza de forma casual.²² Estos nódulos pueden aparecer en un 0.2 % en radiografías de baja dosis, pero su incidencia ha aumentado considerablemente con el uso de la tomografía computarizada,²³⁻²⁵ debido a que la evaluación de la presencia de nódulos en una radiografía solo es eficiente si el nódulo mide al menos 9 mm;²¹ sin embargo, en la TC se pueden detectar estructuras nodulares menores a 1 mm de diámetro.^{21,25} Con la utilización de los sistemas para el Diagnóstico Asistido por Computadora (CAD, por sus siglas en inglés) se puede agilizar el proceso de detección de estructuras nodulares. Los CAD son una tecnología que combina la radiología con el procesamiento de imágenes y la inteligencia artificial. Son capaces de dirigir la atención del radiólogo a regiones candidatas dentro de la imagen a ser nódulos pulmonares,²⁶⁻²⁸ proporcionando una segunda opinión a la hora de interpretar los resultados, mejorando así el rendimiento y la consistencia del diagnóstico.

Proceso de identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax

El proceso de identificación de nódulos pulmonares solitarios da inicio una vez que el paciente se realiza un estudio de tomografía de tórax. Las imágenes generadas son revisadas por el especialista en las estaciones de visualización de la institución médica, quien las interpreta y emite un diagnóstico a partir de los hallazgos detectados durante el proceso de observación. El proceso de diagnóstico a través de imágenes médicas es engorroso. Los equipos de adquisición de imágenes más modernos, generan un elevado volumen de información; propiciando que el especialista tenga que revisar muchas imágenes de un mismo estudio, lo que pudiera provocar cansancio y agotamiento en el radiólogo. En varias ocasiones aparecen nódulos muy pequeños o estructuras indefinidas, que tienden a confundir al especialista. Cómo las estructuras nodulares son de alta densidad el especialista realiza variaciones de este parámetro, en busca de que los nódulos queden resaltados en la imagen que se analiza. La habilidad para identificar estas estructuras se adquieren con los años. Generalmente el proceso de diagnóstico requiere de amplio nivel de experticia por parte del radiólogo y en ocasiones se requiere una consulta de segunda para la emisión de un diagnóstico eficiente,

retardando el proceso de atención al paciente. Este inconveniente se puede mitigar con la utilización de sistemas capaces de dirigir la atención de los radiólogos a regiones candidatas a ser nódulos pulmonares dentro de la imagen.

Situación social sobre el cáncer de pulmón a nivel internacional y en Cuba

Los nódulos pulmonares solitarios (NPS) se pueden identificar atendiendo a la forma de su estructura interna, calcificación, esfericidad, lobulación, espiculación, textura y localización. Los primeros pueden ser de tejido blando, líquido, grasa y aire. Atendiendo a su calcificación son los llamados rosita de maíz, laminado, sólido, no central, central y ausente. Los nódulos con las características de esfericidad son los lineales, ovoides y redondos. Los lobulados pueden ser de marcada lobulación o no lobulados. Los espiculados se identifican por ser no espiculados o con marcada espiculación. Atendiendo a su textura los nódulos se caracterizan en sólidos no calcificados, calcificados y vidrio deslustrado. Según su localización pueden ser centrales o periféricos^{2, 23,25,29}.

Según la Organización Mundial de Salud (OMS) el cáncer es una de las principales causas de muerte a nivel mundial. Se le atribuye 8,2 millones de defunciones ocurridas en todo el mundo en el 2012. Entre los diferentes tipos de cáncer, el que mayor incidencia tuvo en ese mismo año, es el de pulmón con 1,59 millones de muertes.³⁰

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.³¹ Se calcula que alrededor de 90 % de los pacientes con cáncer de pulmón mueren a causa de esta enfermedad.²²

En el Anuario Estadístico de Salud del año 2013 emitido por el Ministerio de Salud Pública de Cuba (MINSAP)³² en abril del 2014, se especifica que los tumores malignos representan la primera causa de muerte en la Isla. El número de defunciones asociadas a estas enfermedades tumorales se elevó a 22 868, de estas, 5 212 se localizaron en tráquea, bronquios y pulmón (TBP), representando una tasa de 46,7 % por cada 100 000 habitantes; comportándose como el grupo de mayor prevalencia tanto en hombres como mujeres. El número de defunciones asociados a estas localizaciones en el sexo masculino asciende a 3 296, representando una tasa del 59.2 % y en el sexo femenino a 1 916, con una tasa de 34.3 %.

Análisis de tendencia en el desarrollo de sistemas CAD para la detección de nódulos pulmonares solitarios

En el transcurso de los años se han publicado diferentes artículos donde se especifican las técnicas desarrolladas para el diseño de algoritmos útiles para la detección de estructuras candidatas a ser nódulos pulmonares solitarios, como son:

Armato et al.¹² segmentan los pulmones sección por sección en un proceso bidimensional, aplicando un umbral basado en niveles de gris para cada sección, esto lo realizan con el objetivo de identificar la región pulmonar dentro del cuerpo del paciente y diferenciarla de las demás regiones. 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". Prosiguen empleando umbralización en base a niveles de gris para generar el

volumen pulmonar segmentado. Para obtener los candidatos a nódulos pulmonares, utilizan múltiples umbrales y una combinación de una serie de técnicas de 2D y 3D.

Gurcan et al.³³ utilizan un primer umbral en -800 Unidades Hounsfield (HU, por sus siglas en inglés) para excluir la región externa del tórax. Como la densidad del pulmón en imagen de Tomografía Computarizada (TC) varía según la profundidad de la inspiración, el ancho del haz y la calibración del escáner. Este umbral no es suficiente por lo que a continuación utilizan el histograma de la imagen dentro de la región torácica y mediante una técnica de "clustering" conocida como "k-means" segmentan cada región pulmonar dentro del tórax utilizando únicamente los niveles de gris de la imagen.

Zhao et al.³⁴ con el objetivo de identificar estructuras nodulares, comienzan aplicando un umbral en -375 HU, para diferenciar regiones de alta y baja densidad. Continúan aplicando la técnica de componentes conexas en 3D con el fin de identificar la región pulmonar. Aplicando cierre morfológico obtienen como resultado los pulmones completamente blancos. Al obtener la máscara pulmonar utilizan el algoritmo Local Density Maximum (LDM), con el fin de identificar las estructuras candidatas a ser nódulos pulmonares solitarios.

Bae et al.³⁵ segmentan los pulmones utilizando umbralización basada en valores de niveles de gris, para ello 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 un filtro de cierre con disco a través de 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 realizan un 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 mm³) 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 voxel candidato. Adicionalmente se realiza una correspondencia morfológica multinivel (3D) utilizando formas de 3, 6, 9 y 12 mm de diámetro.

Souto et al.¹⁷ dividen el proceso de segmentación en tres etapas: aplicación de un umbral adaptativo, reducción de regiones aisladas y segmentación de las regiones pulmonares. Al concluir el segmentado ya cuentan solamente con la región de los pulmones, a los que le aplican un filtro denominado filtro de iris; obteniendo como resultado una imagen en la cual la estructura del fondo y las restantes regiones torácicas han sido eliminadas. A continuación aplican valores umbrales, obteniendo una imagen binaria que contiene las regiones sospechosas de ser nódulos pulmonares.

Suárez Cuenca²¹ se aplica un umbral adaptativo como primer paso en el proceso de segmentación, el cual le permitirá separar dentro de la imagen original las regiones de alta y baja densidad. Las primeras quedarán de color blanco, las cuales serán el cuerpo, la camilla, el tejido denso del pulmón, los nódulos y la grasa. Zonas como el fondo y el aire del pulmón pertenecen a las regiones de baja intensidad, quedarán de color negro. Sobre el resultado se aplica una técnica de etiquetado de regiones para poder suprimir el fondo y todas las regiones blancas (técnica de supresión), manteniendo presente las regiones restantes. A continuación vuelve a eliminar el fondo y las regiones blancas, quedando el contenido correspondiente al

tejido denso dentro de los pulmones. Estas últimas dos imágenes obtenidas las combina por medio de una operación lógica OR, proporcionado como resultado las regiones pulmonares que incluyen el tejido denso en su interior. A continuación realiza una operación morfológica de dilatación, que combinada una vez más por medio de un OR, le permite recuperar parte del hilio pulmonar que no había sido incluida. Como los pulmones habían sido previamente dilatados, realiza una erosión para devolverlos a su tamaño original. Al concluir el proceso de segmentación, el autor prosigue con la utilización de una versión del filtro iris, como umbral de decisión para generar los puntos semillas. A los puntos generados le aplica la segmentación mediante crecimiento de regiones, lo que brindara como resultados las estructuras pulmonares.

Ejaz et al.¹⁸ utilizan para la identificación de estructuras nodulares un umbral fijo a través del histograma de la imagen, con el fin de separar las regiones de alta y baja densidad. Obtienen la región pulmonar a través de la técnica de componentes conexas. Aplican cierre morfológico para rellenar la máscara de color blanco. Corrigen la máscara pulmonar haciendo uso del método de cierre con disco. Para terminar el proceso identifican las estructuras candidatas a ser nódulos pulmonares con el algoritmo Local Density Maximum (LDM).

Una de las clasificaciones de los nódulos pulmonares según su localización son los llamados, periféricos o subpleural, que reciben su nombre debido a que aparecen cerca de la pleura³ y aunque no se originen en ella, sí aparecen en la región periférica y como consecuencia a esto, en un futuro pueden ejercer contacto con ella, debido a su crecimiento. Los nódulos periféricos tienen altos valores de densidad al igual que la camilla y el cuerpo del paciente. Esta característica provoca que al utilizar el método de umbralización para la obtención de la máscara pulmonar, queden excluidos de dicha región, llevándose a cabo un proceso de segmentación incorrecto e imposibilitando la posterior detección del nódulo pulmonar solitario.^{18,35}

Para dar solución a este inconveniente de la exclusión de los nódulos periféricos de la máscara pulmonar existen diversas técnicas como rolling-ball^{21,37} y contornos activos.³³ Los autores de la presente investigación decidieron utilizar cierre con disco³⁵⁻⁴¹ debido a que ofrece resultados favorables de forma automática y cuenta con 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.

Al analizar los resultados obtenidos por los algoritmos estudiados durante la revisión del estado del conocimiento respecto al tema de la identificación de nódulos pulmonares solitarios, se evidencia que estos obtuvieron resultados significativos, en la tabla 1 se relacionan algunos de ellos.

FP: Falsos positivos (Estructuras identificadas como nódulos pulmonares cuando realmente no lo son).

Notas: En algunos algoritmos los resultados son luego de la fase de segmentación y en otros luego de haber clasificado las estructuras en malignas o benignas, se utilizan solo como guía en la investigación.

Tabla 1. Resumen de resultados de las investigaciones analizadas

Investigaciones	Resultados
Armato et al. (12)	Sensibilidad: 89 % con 1,3 FP por imagen
Zhao et al. (27)	Sensibilidad: 84,2 % con 5 FP por imagen
Bae et al. (28)	Sensibilidad: 95,1 % con 6,9 FP (para estructuras de 3 mm) y 4,0 FP (para estructuras de 5 mm)
Suárez Cuenca (14)	Sensibilidad: 80,0 % con 3,9 FP por caso
Ejaz et al. (30)	Sensibilidad 96 % no se especifican valores de FP

(Fuente: creada por los autores)

A partir del estudio realizado a los procedimientos llevados a cabo por varios autores para la realización de un CAD útiles para la detección de estructuras nodulares, se propone la realización de un algoritmo capaz de detectar nódulos pulmonares solitarios en imágenes de tomografía de tórax. Los autores del trabajo coinciden en que las tendencias estudiadas no deben ser implementadas íntegramente de esa manera, por lo que se decidió la utilización de varias técnicas empleadas en una u otra de las investigaciones analizadas.

Las técnicas seleccionadas para el desarrollo del algoritmo se detallan a continuación:

Para diferenciar las regiones de alta y baja densidad se usará una umbralización adaptativa, donde el umbral se adaptará a la densidad de la imagen. Con el fin de identificar la región pulmonar se utilizará la técnica de componentes conexas, permitiendo identificar las regiones conexas dentro de la imagen y con solo eliminar las más grandes y más pequeñas se obtiene la máscara de forma rápida y eficiente. Para corregir la máscara pulmonar se empleará la técnica de cierre con disco porque es automática, eficiente y de bajo costo computacional. Con el objetivo de identificar las regiones candidatas se hará uso del algoritmo Local LDM,^{34,41} por los buenos resultados que arroja el mismo en la identificación de estructuras con altos valores de densidad dentro de la imagen de TC.

RESULTADOS

El algoritmo propuesto analiza la imagen seleccionada por el radiólogo, en busca de estructuras de alta densidad dentro de la región pulmonar, mostrando un resultado de forma automatizada, señalando las estructuras nodulares que pudieran ser de interés. Para obtener la solución deseada como se muestra en la figura 1, la imagen de tomografía (fig. 1a) transita por una serie de pasos, comenzando con una umbralización binaria basada en un umbral adaptativo. Para realizar este procedimiento se utiliza el método automático de Otsu, el cual elige umbral óptimo maximizando la varianza entre clases mediante una búsqueda exhaustiva.³⁵ Esta selección está determinada porque los valores de densidad de la imagen de tomografía de tórax no son uniformes, varía según los parámetros de calibración del escáner y el espesor a lo largo del cuerpo del paciente. Con este procedimiento se separan las regiones de alta y baja densidad, (fig. 1b).

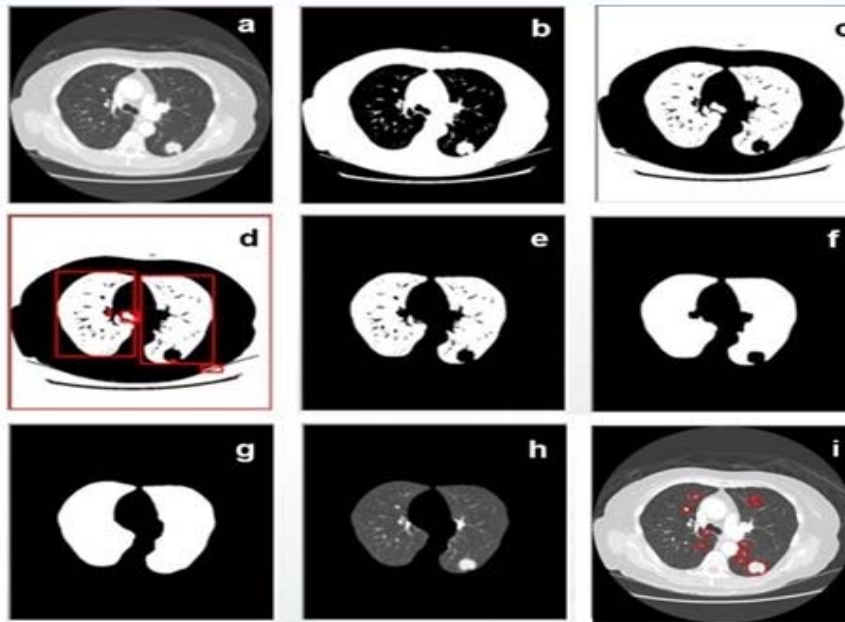


Fig. 1. Proceso de identificación de nódulos pulmonares solitarios: (a) imagen original de imagen original de tomografía, (b) umbralización adaptativa con el método de Otsu, (c) inversión de la imagen umbralizada, (d) detección de componentes conexos en (c), (e) identificación de las regiones correspondiente 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) región de interés sobre la que se aplicara el algoritmo LDM para identificar las estructuras candidatas a nódulos pulmonares

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, (fig. 1c). Al concluir la inversión de la imagen se procede a aplicar correctamente la técnica de detección de componentes conexas, (fig. 1d). Dicha técnica permite identificar las estructuras aisladas y eliminar las más pequeñas y las más grandes. Al finalizar este paso se aplica una depuración sobre las estructuras identificadas previamente, eliminando las regiones más pequeña y las más grandes obteniendo como resultado las regiones correspondientes a los pulmones, como se observa en la (fig. 1e).

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 imagen y la imagen de la (fig. 1e), esta 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, (fig. 1f).

Existen estructuras nodulares que se localizan cerca del borde 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 un cierre con disco para suavizar los bordes de los pulmones y poder incluir las estructuras que pudieran ser de interés para el especialista, (fig. 1g). Después de

haber obtenido la máscara pulmonar corregida, se procede entonces a la identificación de la región pulmonar (fig. 1h), sobre la que se aplicará una umbralización para identificar todo el tejido denso dentro de esta zona. La imagen resultante se utilizará como máscara para identificar la región de interés que será analizada, (fig. 1i).

Los nódulos pulmonares solitarios aparecen en una imagen de tomografía como una estructura de forma circular con alto valores de densidad, disminuyendo hacia los bordes. Para la detección de estas estructuras se aplicó el algoritmo Local Density Maximum (LDM), el cual comienza aplicando un umbral inicial que pudiera ser el máximo valor de densidad de la imagen y luego irá disminuyendo según el valor de paso con que cambiara al próximo nivel de umbral, hasta alcanzar el valor del umbral de parada. LDM compara las estructuras identificados en el nivel de umbral actual, con las estructuras identificadas en el nivel de umbral previo y analiza cuales de estas estructuras pueden ser incluidas como posibles estructuras nodulares.

Para la validación del algoritmo de identificación de nódulos pulmonares se realizó un experimento. La población seleccionada son las imágenes del The Lung Image Database Consortium Image Collection (LIDC-IDRI). Se utilizó la técnica de Muestreo Accidental. El muestreo es una herramienta de la investigación científica, cuya función básica es determinar que parte de una población debe examinarse, con la finalidad de hacer inferencias sobre dicha población. El Muestreo Accidental es un tipo de muestreo no probabilístico, en él se incluyen todos los elementos disponibles, seleccionándolos arbitrariamente sin tener en cuenta ninguna técnica especial, hasta llegar a la cantidad deseada. Este muestreo no garantiza la representatividad de la muestra, pero puede usarse en casos donde eso no constituya un requisito.⁴⁰

El objetivo del experimento es validar que las estructuras nodulares contenidas en las imágenes de tomografía de tórax analizadas, no han sido eliminadas durante el proceso de segmentación. Se escogieron aleatoriamente 16 series, que contienen un total de 2110 imágenes de tomografía de tórax. Por cada una de las series existe un fichero XML que describe los rasgos que caracterizan los nódulos pulmonares. De las 2110 imágenes contenidas en las series se escogieron 377, teniendo en cuenta que solo estas contienen estructuras nodulares. Los resultados del experimento se muestran en la tabla 2.

Tabla 2. Experimento para validar los resultados del proceso de segmentación

No. de la serie	CIS	CIN	Cantidad de imágenes en las que se incluyen los nódulos durante el proceso de segmentación utilizando el algoritmo propuesto			CN	CNDAP	CNNDAP
			Todos	Parcial	No incluye			
LIDC-IDRI-0001	133	14	8	4	2	18	12	6
LIDC-IDRI-0003	140	35	29	5	1	45	38	7
LIDC-IDRI-0005	133	13	8	5	0	14	9	5
LIDC-IDRI-0006	133	20	14	3	3	25	18	7
LIDC-IDRI-0008	133	12	8	3	1	14	10	4
LIDC-IDRI-0011	128	38	31	4	3	51	42	9
LIDC-IDRI-0012	128	38	29	6	3	51	41	10
LIDC-IDRI-0013	125	45	38	4	3	84	76	8
LIDC-IDRI-0014	114	9	6	0	3	9	6	3
LIDC-IDRI-0016	141	32	22	5	5	38	28	10
LIDC-IDRI-0018	127	20	15	4	1	21	14	7
LIDC-IDRI-0021	149	23	16	5	2	26	19	7
LIDC-IDRI-0024	133	15	10	3	2	18	13	5
LIDC-IDRI-0027	133	25	19	3	3	27	20	7
LIDC-IDRI-0028	141	10	6	3	1	10	6	4
LIDC-IDRI-0030	119	28	20	4	4	35	26	9
Total	2110	377	279	61	37	486	378	108
			74.00 %	16.18 %	9.82 %		77,78 %	22.22 %

Fuente: creada por los autores

CIS: cantidad de imágenes de la serie.

CIN: cantidad de imágenes con nódulos.

Todos: el algoritmo incluye todos los nódulos existentes en la imagen.

Parcial: el algoritmo incluye una parte del total de nódulos existentes en la

imagen.

No incluye: el algoritmo no logra incluir durante el proceso de segmentación ninguno de los nódulos existentes.

CN: cantidad de nódulos en la serie.

CNDAP: Cantidad de nódulos detectados por el algoritmo propuesto.

CNNDAP: Cantidad de nódulos no detectados por el algoritmo propuesto.

Del total de imágenes con nódulos identificadas en la muestra, en 279 se incluyen todos los nódulos durante el proceso de segmentación, independientemente a que existan otras estructuras, representando un 74 % del total. En 61 imágenes no se identificaron todos los nódulos existentes y en 37 no se identificaron ninguno de los nódulos, representando un 16.18 % y 9.82 % del total respectivamente.

Los resultados arrojados por el algoritmo desarrollado para la identificación de nódulos pulmonares son similares a los alcanzados por los algoritmos analizados. A pesar de que no fueron probados con la misma data, da una idea de cuán buenos o malos fueron dichos resultados. Se obtuvo una sensibilidad cercana al 80 % con un promedio de 4 FP por imagen. En el caso particular de esta investigación no es recomendable decidir si es mejor mostrar muchos falsos positivos o minimizar el porcentaje de sensibilidad porque en ambos casos existe repercusión negativa para el paciente. Cuando se obtienen altos valores de sensibilidad se manifiestan de igual forma elevados valores de especificidad determinada por el número de falsos positivos.

CONCLUSIONES

Utilizando el algoritmo desarrollado para la identificación de nódulos pulmonares solitarios en imágenes de tomografía de tórax se obtuvo un 77.78 % de acierto, al analizar 377 imágenes que contienen dichas estructuras. Con ello se puede contribuir a la detección temprana de la enfermedad en la población cubana y por tanto mejorar la calidad de vida del paciente y los indicadores de salud.

REFERENCIAS BIBLIOGRÁFICAS

1. Montañó LM. Imagenología y detectores en medicina. Cinvestav [Internet] 2007 [citado 2014 Feb 4]; 26: 16-23. Disponible en: http://www.cinvestav.mx/Portals/0/SiteDocs/Sec_Difusion/RevistaCinvestav/enero-marzo2007/imagenologia.pdf
2. Romero A. Principios de la Tomografía Computarizada. Fed Enseñ CCOO Andal. 2009 [citado 2014 Feb 4]; 4. Disponible en: <http://www.feandalucia.ccoo.es/indcontei.aspx?d=3398&s=5&ind=177>
3. Corbo DN. Tomografía Axial Computada. Urug Fac Med E Ing Univ Repúb Orient Urug. 2004.
4. Passariello G. Imágenes Médicas. Adquisición, Análisis, Procesamiento e Interpretación [Internet]. 1a ed. Equinoccio; 1995 [citado 2014 Feb 4]. Disponible en: http://www.researchgate.net/publication/233728000_Imagenes_Mdicas_Adquisicin_analisis_procesamiento_e_interpretacin

5. Martínez Á. Bases Teóricas, Digitalización y Análisis de Imágenes. Manual de Telepatología [Internet] [citado 2014 Feb 6] p.11-6. Disponible en: https://www.seap.es/c/document_library/get_file?uuid=531ccea-9c9f-4f58-8855-3229e8307924&groupId=10157
6. Ordoñez CA. FORMATOS DE IMAGEN DIGITAL. Rev Digit Univ [Internet] 2005 [citado 2014 May 13];5(7). Disponible en: http://www.revista.unam.mx/vol.6/num5/art50/may_art50.pdf
7. Mugarra CF, Chavarría M. La Radiología Digital: Adquisición de imágenes [Internet]. Universidad de Valencia, España; 2003 [citado 2014 May 8]. Disponible en: http://www.conganat.org/SEIS/is/is45/IS45_33.pdf
8. Radiodiagnóstico digital oral. Relación entre La Radiología Convencional Y La Radiología Digital [Internet]. [citado 2014 Feb 5]. Disponible en: http://www.radiodiagnosticodigitaloral.com/joomla/index.php?option=com_content&view=article&id=90:relacion-entre-la-radiologia-convencional-y-la-radiologia-digital&catid=47:modulo-1&Itemid=76
9. Barbieri PG, Flores GJ, Escribano BM. Actualización en radiología dental. Radiología convencional Vs digital. Radiol Conv Vs Digit 2006 [citado 2014 May 25];22(2). Disponible en: http://scielo.isciii.es/scielo.php?pid=S0213-12852006000200005&script=sci_arttext
10. Ramírez JC, Arboleda C, McCollough CH. Tomografía computarizada por rayos X: fundamentos y actualidad. Rev Ing Bioméd 2008 [citado 2014 May 25];2(4):16-7. Disponible en: http://scholar.google.com/citations?view_op=view_citation&hl=es&user=Y_8d7PsAAA&citation_for_view=Y_8d7PsAAA:Zph67rFs4hoC
11. Llobet R, Pérez JC, Paredes R. Técnicas Reconocimiento de Formas Aplicadas al Diagnóstico de Cáncer Asistido por Ordenador. Rev E Salud [Internet] 2006 [citado 2014 Feb 10];2(6). Disponible en: <http://www.revistaesalud.com/index.php/revistaesalud/article/view/110>
12. Armato III SG, Giger ML, MacMahon H. Automated detection of lung nodules in CT scans: Preliminary results. Am Assoc Phys Med 2001 [citado 2014 Feb 10];28:1552. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/11548926>
13. Microsoft. Visual C# resources [Internet] 2014 [citado 2014 May 25]. Disponible en: <http://msdn.microsoft.com/en-us/vstudio/hh341490.aspx>
14. Microsoft. Visual Studio [Internet] 2014 [citado 2014 May 25]. Disponible en: <http://www.visualstudio.com/>
15. GNU Operating System. EmguCV [Internet] 2014 [citado 2014 May 25]. Disponible en: http://www.emgu.com/wiki/index.php/Main_Page
16. Bradski G, Kaehler A. Computer Vision with the OpenCV Library [Internet]. United States of America; 2008 [citado 2014 May 26]. Disponible en: <http://www.cs.haifa.ac.il/dkeren/ip/OReilly-LearningOpenCV.pdf>
17. Souto M, Tahoces PG, Suárez JJ, Lado MJ, Remy M, Remy J, et al. Detección automática de nódulos pulmonares en tomografía computarizada. Un estudio preliminar*. Radiología 2008 [citado 2014 May 25];50(5):387-92. Disponible en:

<http://www.elsevier.es/ct-revista-radiologia-119-articulo-deteccion-automatica-nodulos-pulmonares-tomografia-13127633>

18. Ejaz N, Javed S, Sajid Z. Implementation of Computer Aided Diagnosis System for Lung Cancer Detection. Lect Notes Softw Eng 2013 [citado 2014 May 25]; 1(4): 364-9. Disponible en: www.inse.org/papers/78-VC020.pdf
19. Sánchez M, Vallés AE. Framework CALIB. Procesos de lectura, procesamiento y visualización de imágenes médicas. Ciudad de La Habana: Universidad de las Ciencias Informáticas, Facultad 7; 2010.
20. Mainieri JA, Brenes A. Cáncer de Pulmón. Asoc Costarric Neumol Cir Tórax 2003 [citado 2014 May 25]; 45(1). Disponible en: http://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S0001-60022003000500006&lng=en&nrm=iso&tlng=es
21. Suárez JJ. Desarrollo de un sistema de diagnóstico asistido por computador para detección de nódulos pulmonares en tomografía computarizada multicorte. [PhD]. [Santiago de Compostela]; 2009.
22. Costa DM, Prado Y, Lozano JL, Plasencia CC, Riesgo Cosme Y de la C. Principales aspectos clinicoepidemiológicos del cáncer de pulmón. MEDISAN 2011 [citado 2014 May 25]; 15(8). Disponible en: http://scielo.sld.cu/scielo.php?pid=S1029-30192011000800008&script=sci_arttext
23. Young Kim H, Mog Shim Y, Soo Lee K, Han J, A Yi C, Kyung Kim Y. Persistent Pulmonary Nodular Ground-Glass Opacity at Thin-Section CT: Histopathologic Comparisons. Radiology 2007 [citado 2014 May 25]; 245(1). Disponible en: <http://pubs.rsna.org/doi/full/10.1148/radiol.2451061682>
24. Erasmus JJ, Connolly JE, MacAdams HP, Roggli VL. Solitary Pulmonary Nodules: Part I. Morphologic Evaluation for Differentiation of Benign and Malignant Lesions. RadioGraphics 2000 [citado 2014 May 26]; 20(1): 43-58. Disponible en: http://pubs.rsna.org/doi/10.1148/radiographics.20.1.g00ja0343?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%3dpubmed
25. Winer H. The Solitary Pulmonary Nodule. Radiology. 2006 [citado 2014 May 26]; 239(1): 34-49. Disponible en: http://pubs.rsna.org/doi/10.1148/radiol.2391050343?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%3dpubmed
26. Oporto S. Detection of Microcalcification in Digital Mammograms by Improved-MMGW Segmentation Algorithm. Am Coll Radiol. 2013 [citado 2014 May 26]; 10. Disponible en: <http://www.sciencedirect.com/science/journal/15461440/10/10>
27. Carrillo de Gea JM, García G. DETECCIÓN DE NORMALIDAD Y ANORMALIDAD EN RADIOGRAFÍAS DE TÓRAX USANDO LBP [Internet]. Universidad de Murcia; 2008 [citado 2014 Jun 16]. Disponible en: <http://dis.um.es/ginesgm/files/inv/pfc-jmcg/pfc-jmcg.pdf>
28. Sluimer I, Schilham A, Prokop M, Ginneken B van. Computer Analysis of Computed Tomography Scans of the Lung: A Survey. IEEE Trans Med IMAGING. 2006 [citado 2014 Jun 16]; 25(4). Disponible en: http://www.researchgate.net/publication/7172159_Computer_analysis_of_computed_tomography_scans_of_the_lung_a_survey

29. Meza P, Gutreiman K, Núñez S. NÓDULO PULMONAR SOLITARIO. Rev MEDICA COSTA RICA Centroam. 2007 [citado 2014 Jun 16];155-9. Disponible en: <http://www.binasss.sa.cr/revistas/rmcc/CONTENIDOS.htm>
30. Organización Mundial de la Salud (OMS). Cáncer [En línea]. 2013 [citado 2014 Feb 22]. Disponible en: <http://www.who.int/mediacentre/factsheets/fs297/es/>
31. Blázquez B, Varela R. DIAGNOSIS DECISION SUPPORT SYSTEMS. Leganés, España?: Departamento de Teoría de la Señal y Comunicaciones, Universidad Carlos III de Madrid; 2010.
32. DIRECCIÓN DE REGISTROS MÉDICOS Y ESTADÍSTICAS DE SALUD. ANUARIO ESTADÍSTICO DE SALUD 2013 [Internet]. La Habana: MINISTERIO DE SALUD PÚBLICA; 2014 [citado 2014 Feb 22]. Disponible en: <http://www.sld.cu/servicios/estadisticas/>
33. Gurcan MN, Sahiner B, Petrick N, Chan HP, Kazerooni EA, Cascade PN, et al. Lung nodule detection on thoracic computed tomography images: Preliminary evaluation of a computer-aided diagnosis system. Med Phys 2002 [citado 2014 Feb 22];29. Disponible en: <http://scitation.aip.org/content/aapm/journal/medphys/29/1>
34. Zhao B, Gamsu G, Ginsberg MS, Jiang L, Schwartz LH. Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm. J Appl Clin Med Phys 2003 [citado 2014 Feb 22];4(3):248-60. Disponible en: http://www.unboundmedicine.com/medline/journal/Journal_of_applied_clinical_medical_physics?start=990&next=true
35. Bae KT, Kim JS, Na YH, Gi Kim K, Kim JH. Pulmonary Nodules: Automated Detection on CT Images with Morphologic Matching Algorithm- Preliminary Results1. Radiology 2005 [citado 2014 Feb 22];236:286-94. Disponible en: <http://pubs.rsna.org/doi/abs/10.1148/radiol.2361041286>
36. Retico A, Delogu P, Fantacci M, Gori I, Preite A. Lung nodule detection in low-dose and thin-slice computed tomography. 2008 [citado 2014 Feb 22];38:525-34. Disponible en: <http://www.computersinbiologyandmedicine.com/article/S0010-4825%2808%2900020-6/abstract>
37. Bellotti R, DE Carlo F, Gargano G, Tangano S, Cascio D, Catanzaritti E, et al. A CAD system for nodule detection in low-dose lung CTs based on region growing and new active contour model. Med Phys. 2007 [citado 2014 Feb 22];34(12):4901-10. Disponible en: http://www.researchgate.net/publication/5654351_A_CAD_system_for_nodule_detection_in_low-dose_lung_CT_s_based_on_region_growing_and_a_new_active_contour_model
38. Ejaz N, Javed S, Sajid Z. Implementation of Computer Aided Diagnosis System for Lung Cancer Detection. Lect Notes Softw Eng 2013 [citado 2014 Feb 22];1(4):364-9. Disponible en: www.inse.org/papers/78-VC020.pdf
39. Hernández RA, Coello Z. El Proceso de Investigación Científica. Ciudad de La Habana: Editorial Universitaria; 2011.

¹Se refiere a la finura de detalles visibles en una imagen, cuanto menor es el área terrestre representada por cada píxel en una imagen digital, mayores son los detalles que pueden ser captados y mayor es la resolución espacial.

²Son aquellas estructuras nodulares que se encuentran alojadas en la periferia de la región pulmonar, cercanas a la pleural; lo que trae consigo que se le conozca también como nódulo subpleural.

³Es una membrana serosa que recubre a los pulmones.

Recibido: 19 de mayo de 2014.

Aprobado: 5 de marzo de 2015.