

Detección de cáncer de pulmón en imágenes de tomografía computarizada

utilizando

Red neuronal artificial basada en retropropagación

Característica de matrices de coocurrencia de nivel de gris

Lilik Anifah

Departamento de Informática, Facultad de Ingeniería
Universitas Negeri Surabaya
Indonesia lilikanifah@unesa.ac.id

Rina Harimurti

Departamento de Informática, Facultad de Ingeniería
Universitas Negeri Surabaya
Indonesia rinaharimurti@unesa.ac.id

Puput Wanarti Rusimamto

Departamento de Ingeniería Eléctrica, Facultad de Ingeniería
Universitas Negeri Surabaya Indonesia
puputwanarti@unesa.ac.id

Muerte de causa maligna y podría crecer a partir de cada tipo de célula en el cuerpo humano [1].

El cáncer de pulmón es la causa más común de muerte por cáncer, no solo la principal causa de muerte masculina, sino también la población femenina, 13,6% y las muertes por cáncer de pulmón 11,1%. El porcentaje de muertes puede disminuir si el cáncer puede ser detectado y detectado, la posibilidad de que se produzca un error aumentará.

Haryanto

Departamento de Ingeniería Eléctrica, Facultad de Ingeniería
Universitas Trunojoyo Madura
Indonesia

Haryanto_utm@yahoo.com

Zaimah Permatasari

Departamento de Ingeniería Eléctrica, Facultad de Ingeniería
Universitas Negeri Surabaya Indonesia
zaimah.permatasari@gmail.com

Adam Ridianto Muhamad

Departamento de Ingeniería Eléctrica, Facultad de Ingeniería
Universitas Negeri Surabaya Indonesia
adamridianto@unesa.ac.id

Resumen: El cáncer de pulmón es la causa más común de muerte por cáncer en el mundo. La detección temprana del cáncer de pulmón ayudará enormemente a salvar al paciente. Esta investigación se centra en la detección del cáncer de pulmón utilizando la función de matrices de co-ocurrencia de nivel de grises (GLCM) basadas en la retropropagación de la red neuronal Artificial. Los datos pulmonares utilizados proceden de la base de datos de archivos de imágenes de cáncer, los datos utilizados consistieron en 50 imágenes de TC. La imagen de TC se agrupa en 2 grupos, normal y cáncer de pulmón. Los pasos de esta investigación son: preprocesamiento de imágenes, segmentación de la región de interés, extracción de características y detección de cáncer de pulmón mediante la retropropagación de redes neuronales. Los resultados muestran que el sistema puede detectar imágenes de TC * de pulmón normal y cáncer de pulmón con una precisión del 80%. Con suerte, se utiliza para ayudar al personal médico y la investigación a detectar el estado del cáncer de pulmón.

Palabras clave: componente; cáncer de pulmón; backpropagación; red neuronal; glcm; Imagen de CT

1. INTRODUCCIÓN

El cáncer es el crecimiento de células anormales y descontroladas. Puede dañar el tejido circundante que se extiende lejos de su origen.

más alto. Hay dos tipos principales de cáncer de pulmón, el cáncer de pulmón de células pequeñas (CPCP) y el cáncer de pulmón de células no pequeñas (CPCNP) [2].

La detección temprana del cáncer de pulmón ayudará a recuperar al paciente. El instrumento que se utiliza para detectar el cáncer de pulmón es a través de una tomografía computarizada (CT Scan). La imagen de la tomografía computarizada dará resultados diferentes entre pulmones normales y anormales, y también el estadio del cáncer de pulmón. El proceso de inspección de esta manera, por supuesto, requiere un equipo que sea muy complejo, cxcpsivc y rcquirc cxcprts en el proceso. Incluso los cxcprts también pueden cometer errores al distinguir btwnccn pulmones normales y anormales. Por lo tanto, muchos investigadores han proporcionado varias soluciones alternativas que pueden ayudar a los médicos utilizando diversas técnicas de procesamiento de imagc [3].

La red neuronal artificial es una parte del sistema de procesamiento de la información diseñada para imitar como el funcionamiento del cerebro humano. Resuelve un problema realizando el proceso de aprendizaje a través de la sinapsis del peso de sus sinapsis [4].

Se han realizado estudios específicos en la detección del cáncer de pulmón mediante varios métodos. Singh et al (2010) han realizado una investigación sobre el desarrollo del algoritmo Celular Neural Network (CNN) para detectar los límites y áreas del cáncer de pulmón a partir de imágenes de rayos X [5]. Detección pulmonar pulmonar utilizando redes neuronales artificiales y métodos de agrupamiento difuso investigados por [6].

Resca.tc:h por B: a.it'u Sa.ptalk.ar cáncer
Neunl i.rnuges II '). Otro se ha hecho
por Zagreb, sistema de Croastim, clasificación de enfoque
de.sth.ma aud c: hreaic obstructive pu.b: noua.ry disease
(CORD) 'tuzzy rajcø a.tid tl: g entrenado neuronal
network [8].

Es research foq.wes Detección de cáncer LMT * en Cr
Scan I.zczge Usm.g Artificial Neva.J Network Based Citay

según GLCM maüi% para extraer sus características:
homogeneidad, e.eagy, varia.ac; g,

GLCM luego entrenado El método Lazo de redes
artificial es el proceso de retropropagaciónFormación ABM tbe
realizado una prueba de CT de El los resultados obtienen
I, hal remolacha)

El proceso de researeb es la figura 2. La coalición de la
retropropagación de redes neuronales utilizada en esta
investigación es ecu: sist de 3 lay «, I; aF I. is in: put • what from 5
oodcg, layet 2 arc g I.ycr 3 (ou: put 1

Matrices de nivel (GLCM) Con suerte usc para ayudar a médicos y
para d <ect Cuneet

II. METHODOLOGY

A. Mai € ri, gi original— de d: ue used consitió o SO
The lung data used ori CT-inm.ug • es- CT-
Cancerimagingarchive Database, data irr • nc está garantizado
en 2 clustcrs.,
y cáncer de pulmón. Imagen de CT N «-mal consistente
de 20 inwges y 30 imágenes de tomografía. Figure I (a.) Muestra
empate
example De li.ii'ig CT-itna: ge (b) * haws lungcancer
imager

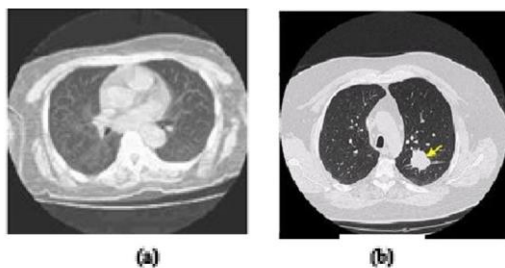


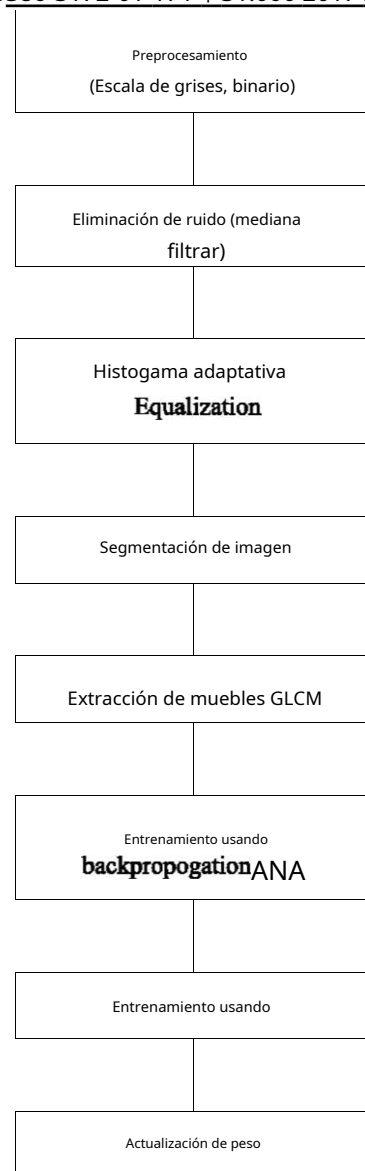
Fig 1. Lung CT image (a) Normal (b) Lung cancer

Métodos

Antes de detectar la imagen de tomografía SRATUS @
eaãaae • gmetit es el primer paso prgpcwi: ttg by seale
eliminar el ruido utilizando para mejorar el control
de la ecuación del histograma ulaptive.
y después paocesscd segmect *tion aplicado para encontrar
segtraetit región. de interés, las características son
objeto oc chaneteri \$ iics * comió dividido nznzely notu: rul
y características nHiEciol. T'be aniuml featue es una
combinación de iroa.ge, como brighness utd bordes o
el objeto. 'Mientras que d: tc C'Gfic: id es mediante la operación
lhe tal histograma de nivel gruy, etc.

Textuc tiene una caracterización muy hermosa; Mjor. amplia
gama de textura irn, agcv utilizada por hurna: u sistema visual: us
fir interpretación subgrwnnty iznprovd izn.Ø, ge repaired

978-1-5386-3172-0 / 17 / \$ 31.000 2017 IEEE



Cerdo L. Process of thewe.arã

111. Rm.n.tsæmDtscusm

Procesamiento de una imagen

Los métodos de procesamiento de imágenes utilizados fijan filetes medje.n binarios y el histograma adaptativo tienen s; larld.art di: tagtiSOit píxeles. En la figura se muestra el resultado de la irritación en escala de grises. En la figura se muestra la imagen binaria (c) Eliminación de ruido. resultado usando el filtro medio (d)

a una imagen en escala de grises que es el paso para convertir a una imagen binuy. El método de filtro de la mediana es muy útil para eliminar

Figura 3Cc) el ruido de la imagen en el cáncer de pulmón « remolacha reducida, adi1Ft: ivg para iff • se calidad, así que piénselo es más claro e inelegible. Esto puede ser se <l. de la imagen de resultado o histograma adaptativo ecuación, la imagen de t'C \$ ull.s tv; ad '

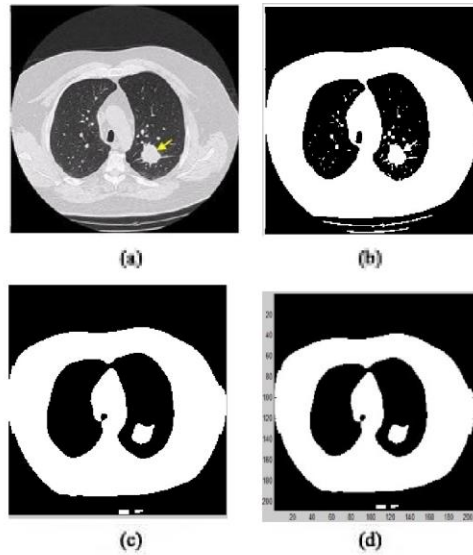


Fig 3. (a) Grayscale image (b) Blurry image (c) Noise removal result using median filter (d) Results of Histogram adaptive equalization process

B. Image Segmentation

Segmentation method used to get the area of lung cancer is regionprops used to find the value of this centroid is represented in the form of coordinates X and Y which states the central axis coordinates. Once obtained the value of centroid then done the process of labeling. Labeling is used to recognize all objects that has the potential as a nodule.

Objects that are candidates of the affected area will be normalized and then searched its feature information. Figure 4 is example result of image segmentation.

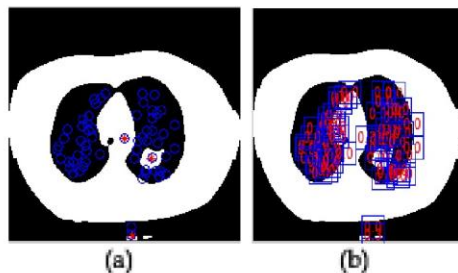


Fig 4. (a) Centroid location (b) Region of object in CT image

C. Feature Extraction

In texture feature extraction, the difference is the texture that is the determinant characteristic of the image. The statistical technique for texture feature extensions is GLCM. The technique is performed by performing a displacement to find for the gray traces of each pixel separated by a fixed distance d and angle θ . Usually the angles used are $0^\circ, 45^\circ, 90^\circ$, and 135° [9]. GLCM features extraction used are homogeneity, energy, contrast, correlation, and variance.

Formula of feature extraction (homogeneity, energy, contrast, correlation, and variance are (Lee & Choi,2010):

$$\text{Contrast} = \sum_i \sum_j (i - j)^2 C(i, j) \quad (1)$$

$$\text{Energy} = \sum_i \sum_j C^2(i, j) \quad (2)$$

$$\text{Entropy} = \sum_i \sum_j C(i, j) \log C(i, j) \quad (3)$$

$$\text{Homogeneity} = \sum_i \sum_j \frac{C(i, j)}{1 + |i - j|} \quad (4)$$

$$\text{Correlation} = u \sum_i^i = 1 \sum_j^j = 1 \quad (5)$$

Homogeneity Table 1 shows that the image of lung cancer has higher homogeneity value compared with normal image. Range homogeneity of abnormal lung image ranged from 0,9796-0,9816, whereas normal lung image between 0,9793-0,9805, the higher homogeneity value, hence the higher uniformity of intensity of an image.

TABLE I. HOMOGENEITY FEATURE

Image	Homogeneity	
	Lungs Cancer	Normal
1	0.9796	0.9793
2	0.9816	0.9805
3	0.9781	0.9790
4	0.9835	0.9838
5	0.9809	0.9745
6	0.9803	0.9873
7	0.9791	0.9794
8	0.9680	0.9798
9	0.9809	0.9843
10	0.9834	0.9838

TABLE II. ENERGY FEATURE

Image	Energy	
	Lungs Cancer	Normal
1	0.5148	0.4697
2	0.4719	0.4630
3	0.4909	0.5060
4	0.5216	0.5258
5	0.4645	0.5493
6	0.5271	0.6094
7	0.4847	0.4605
8	0.5704	0.5499
9	0.4814	0.0315
10	0.5146	0.4880

Table 2 shows below the image of lung cancer having a higher energy value than the normal image. Range of image energy of lung cancer ranged from 0,5148-0,4719, whereas normal image between 0,4697-0,4630.

TABLE III. CONTRAST FEATURE

Image	Contrast	
	Lungs Cancer	Normal
1	0.0408	0.0414
2	0.0368	0.0390
3	0.0437	0.0420

Imagen	Contraste	
	Cáncer de lun	Normal
4	0.0331	0.0325
	0.0382	0.0510
	0.0395	0.0255
7	0.0418	0.0412
8	0.0641	0.0403
	0.0383	0.0315
10	0.0331	0.0327

La Tabla 3 muestra el rango de contraste de la imagen pulmonar normal en el rango de 0.414-0.0390, mientras que la imagen pulmonar no es normal entre 0.0408-0.0368.

Según la Tabla 4, el rango de correlación de imágenes de cáncer de pulmón varió de 2,6779-2,5487, mientras que la imagen normal entre 2,5640-2,5097.

TABLA IV. FUNCIÓN DE CORRELACIÓN

Imagen	Correlación	
	Cáncer de lun	Normal
	2.6779	2.5640
	2.5487	2.5097
3	2.6447	2.6622
	2.6746	2.6823
5	2.5157	2.7544
6	2.7000	2.8420
7	2.6114	2.4980
	2.7864	2.7444
	2.4263	2.7723
10	2.6586	2.5872

TABLA V. CARACTERÍSTICAS DE LOS VARIANOS

Imagen	Varians	
	Cáncer de lun	Normal
	2.6779	2.5640
	2.5487	2.5097
3	2.6447	2.6622
4	2.6746	2.6823
	2.5157	2.7544
	2.7000	2.8420
7	2.6114	2.4980
8	2.7864	2.7444
	2.4263	2.7723
	2.6586	2.5872

La tabla 5 muestra que la imagen de pulmón normal tiene un valor más alto que la imagen de cáncer de pulmón. El rango de energía de imagen normal osciló entre 0,2886-0,2948, mientras que la imagen de cáncer de pulmón entre 0,2573-0,2857.

Los resultados de homogenea, contraste, contraste, correlación y variables variables en las Tablas 1-5 anteriores a algunos valores se superponen entre imágenes normales y de cáncer, por lo que no es posible realizar una programación lineal. Es por eso que la retropropagación ANN solía detectar entre imágenes cancerosas y normales.

D. Detección de cáncer de pulmón

El entrenamiento se realiza mediante retropropagación, cada categoría de imagen tiene un valor objetivo diferente. En la imagen normal del objetivo utilizado I, la imagen del cáncer de pulmón objetivo utilizado es 0. La prueba se realiza utilizando los parámetros de entrada y el peso de actualización obtenido en el proceso de entrenamiento (tasa de Icarning 0,3, valor de capa oculta 20 y época 1000). Decidimos la tasa de aprendizaje aplicada 0.3, el valor de la capa oculta 20 y la época 1000 en función del experimento que ya se realizó, utilizando ese parámetro para producir la mayor precisión que otro parámetro. La Tabla 6 son los resultados de los experimentos.

TABLA VI. VARIACIÓN DE LA TASA DE APRENDIZAJE HACIA EL VALOR DE MSE

Tasa de aprendizaje	MSE
	0,08
0,6	0,06
0,3	0,09

La matriz de confusión (Tabla 7) muestra que los datos de CTimgc de cáncer de pulmón clasificados como cáncer de pulmón son 80% y los datos normales clasificados como normales son 80%.

TABLA VII. MATRIZ DE CONFUSIÓN

PredicGon	Cáncer de lun	Normal	Total
Cáncer de lun	80%	20%	100%
Normal	20%		100%

IV. CONCLUSION

Se han realizado la etapa de preprocesamiento, la segmentación de imágenes, la extracción de características y el proceso de aprendizaje. El peso generado a partir del proceso de aprendizaje se utiliza para probar 50 datos de imagen CT. Los resultados muestran que el sistema funciona con una precisión del 80%. Para obtener una mayor precisión, se necesita más investigación mejorando el proceso de preprocesamiento, la segmentación de imágenes, la extracción de características y el proceso de aprendizaje.

ACQqOWLEDGNffNT

Esta investigación fue apoyada por Universitas Negeri Surabaya. También nos gustaría mostrar nuestra gratitud a Cancerimagingarchive Database por compartir su imagen de TC pulmonar y su biomarcador en esta investigación.

REFERENCIAS

- [1] Sociedad Estadounidense del Cáncer. (2013). ¿Cuáles son los factores de riesgo del cáncer de pulmón de células no pequeñas? 19 de diciembre de 2013. <http://www.cancer.org>.
- [2] Ancuccanu, RV y Victoria, I, 2004, farmacológicamente activo Compuestos naturales para el cáncer de pulmón, Altrn. Medicina. Rcv., 9, 4, 402-419

- [3] MStarawnch, "Detección del cáncer de pulmón mediante técnicas de procesamiento de imágenes", Revista electrónica Leonardo o Prácticas y tecnologías, número 20, págs. 147-158, enero-junio de 2012.
- [4] Kusumadwi, S. . Membangun Jaringan Syaraf Tiruan dengan Matlab y Exel Link. Yogyakarta: Graha I] mu.
- [5] S. Singh, R. Vijay, Y. Singh, "Red neuronal artificial y detección del cáncer", IOSR Journal of Computcr Engincrring, págs. 20-24.
- [6] F. Tahcr, N. Wcrhi, H. Al-Ahmad, R. Sammouda, "Dctcctcion del cáncer de pulmón mediante el uso de redes naturales artificiales y métodos de agrupamiento difuso", American Journal of Biomedical Engineering 2012, 2 (3), págs. 136 -142.
- [7] Almas Pathan, Bairu.K.saptalkar, "Detección y clasificación del cáncer de pulmón utilizando una red neuronal oficial", Revista internacional en Tecnología de comunicación y tecnología informática avanzada Vol- I
- Problema: I-
- [8] A. Badnjevic, M. Cifrck, D. Koruga, D. Osmankovic, "Clasificación neuro-difusa del asma y la enfermedad pulmonar obstructiva crónica", BMC Med Inform Decis Mak. 2015; 15 (Supl. 3): SI. Publicado en línea el 11 de septiembre de 2015.
- [9] RMHaralick, K. Shanmugam, RMHaralick, K Shanmugam „" Características de textura para la clasificación de imágenes ", IEEE Trans. System Man. Cibernética, vol. SMC-3, págs. 610-621, 1973.
- [10] D. Tian, "A Rcvicw on Image Feature Extraction and Rcprcscntation Techniques ", Revista Internacional de Multimedia e Ingeniería Ubicua vol. 8, No. 4, julio de 2013, págs. 385-396.

