Cancer Lungs Detection on CT Scan Image Using Artificial Neural Network Backpropagation Based Gray Level Coocurrence Matrices Feature

by Puput Wanarti

Submission date: 01-Jul-2019 11:42AM (UTC+0700)
Submission ID: 1148323275
File name: Puput_ICACSIS_IEEE_2017.pdf (798.89K)
Word count: 2415
Character count: 12585

Cancer Lungs Detection on CT Scan Image Using Artificial Neural Network Backpropagation Based Gray Level Coocurrence Matrices Feature

Lilik Anifah
Informatics Department, Faculty of Engineering
Universitas Negeri Surabaya
Indonesia
lilikanifah@unesa.ac.id

Haryanto
Electrical Engineering Department, Faculty of Engineering
Universitas Trunojoyo Madura
Indonesia
Haryanto_utm@yahoo.com

Rina Harimurti
Informatics Department, Faculty of Engineering
Universitas Negeri Surabaya
Indonesia
rinaharimurti@unesa.ac.id

Zaimah Permatasari
Electrical Engineering Department, Faculty of Engineering
Universitas Negeri Surabaya
Indonesia
zaimah.permatasari@gmail.com

Puput Wanarti Rusimamto
Electrical Engineering Department, Faculty of Engineering
Universitas Negeri Surabaya
Indonesia
puputwanarti@unesa.ac.id

Adam Ridiantho Muhamad
Electrical Engineering Department, Faculty of Engineering
Universitas Negeri Surabaya
Indonesia
adamridianto@unesa.ac.id

***Abstract*— Lung cancer is the most common cause of cancer death in the world. Early detection of lung cancer will greatly help to save the patient. This research focuses on detection of lung cancer using Artificial Neural Network Back-propagation based Gray Level Co-occurrence Matrices (GLCM) feature. The lung data used originates from the Cancerimagingarchive Database, data used consisted of 50 CT-images. CT-image is grouped into 2 clusters, normal and lung cancer. The steps of this research are: image preprocessing, region of interest segmentation, feature extraction, and detection of lung cancer using Neural Network Back-propagation. The results shows system can detect CT-image of normal lung and lung cancer with accuracy of 80%. Hopefully use to help medical personnel and research to detect lung cancer status.**

Keywords—component; lung cancer; backpropagation; neural network; glcm; CT-image

I. Introduction

Cancer is the growth of abnormally and uncontrolled cells. It can damage the surrounding tissue spread far from its origin. Malignant cause death and it could grow from every cell type in the human body [1].

Lung cancer is the most common cause of cancer deaths, not only the highest cause of male deaths, but also female population, 13.6% and cancer deaths lungs 11.1%. The percentage of deaths may decrease if the cancer can be detected and treated early, the possibility of recovery will be higher. There are two main types of lung cancer, Small Cell Lung Cancer (SCLC) and Non-Small Cell Lung Cancer (NSCLC) [2].

Early detection of lung cancer will help to recover the patient. Instrument used to detect lung cancer is through CT Scan (Computed Topography). The image of CT Scan will give different results between normal and abnormal lungs, and also the stage of lung cancer. Inspection process in this way of course requires equipment that is quite complex, expensive and require experts in the process. Even experts can also make mistakes in distinguishing between normal and abnormal lungs. Therefore, many researchers have provided various alternative solutions that can help physicians by utilizing various image processing techniques [3].

Artificial neural network is one of the information processing system designed imitate like the workings of the human brain. It solving a problem by doing the learning process through the synapse of the weight of its synapses [4]. Several studies have been done in determining lung cancer by several methods. Research has been conducted by Singh et al (2010) on the development of the Celular Neural Network (CNN) algorithm to detect the limits and areas of lung cancer from X-ray images [5]. Lung Cancer Detection Using Artificial Neural Network and Fuzzy Clustering Methods investigated by [6].

978-1-5386-3172-0/17/$31.00

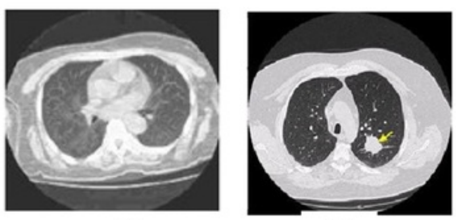
Research conducted by Almas Pathan, Bairu Saptalkar detects lung cancer using Neural Network on X-ray images [7]. Other research has been done by Zagreb, Croastia, system for classification of asthma and chronic obstructive pulmonary disease (COPD) based on fuzzy rules and the trained neural network [8].

This research focuses Detection of Lung Cancer in CT Scan Image Using Artificial Neural Network Based Gray Level Coocurrence Matrices (GLCM) feature. Hopefully use to help medical personnel and research to detect lung cancer status.

II. Methodology

A. Material

The lung data used originates from the cancerimagingarchive Database, data used consisted of 50 CT-images. CT-image is grouped into 2 clusters, normal and lung cancer. Normal CT-image consisting of 20 images and 30 lung cancer tomography images. Figure 1(a) shows the example of normal lung CT-image while (b) shows lung cancer image.



(a) (b)

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

B. Methods

Before detecting the status, tomography image enhancement is done. The first step is prepocessing by doing grayscale and binary, then remove noise using median filter, to improve the contrast adaptive histogram equalization is used, and afterwards the segmentation processed applied to find segment the region of interest. Features are unique characteristics of an object, features are divided into two, namely natural features and artificial features. The natural feature is an ever-present part of the image, such as the brightness and edges of the object. While the artificial feature is a feature obtained by a particular operation on the image, such as gray level histogram, etc [9].

Texture is very helpful characterization for a wide range of image. Texture commonly used by human visual systems for recognition and interpretation [10]. The intensity value which subsequently improved image repaired according to GLCM matrix to extract its features: homogeneity, energy, contrast, correlation, and variance.

The GLCM feature extraction results then trained using Artificial Neural Networks. The method used is backpropagation network. After the training process then performed a test of CT image of the lungs, the results will get the level of accuracy of the system that has been designed. Process in this research is figure in figure 2. Configuration of Neural Network Backpropagation used in this research is consist of 3 layer, layer 1 is input layer where there are from 5 nodes, layer 2 there are 8 nodes and layer 3 (output layer) 1 node.

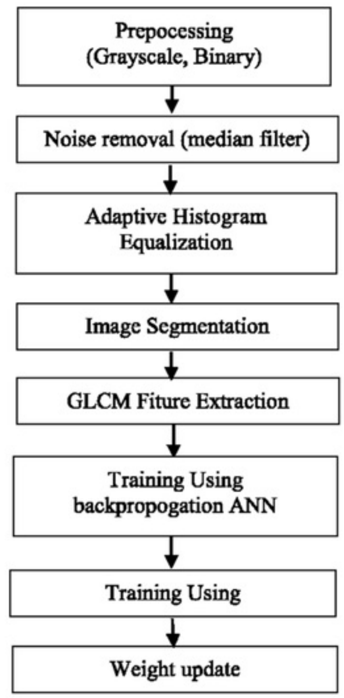


Fig 2. Process of the research

III. Results and Discussion

A. Image Processing

Image processing methods used for intensity improvement are grayscale, binary, median filters and adaptive histogram equations. The CT-images used have standart dimension 207x208 pixels. The results of preprocessing is shown is Figure 3. The result of Grayscale image is shown in Fig 3 (b), binary image is shown (c) Noise removal result using median filter (d) Results of Histogram adaptive equalization process.

Grayscale process for converting images from rgb image to grayscale image which is a step to convert to binary image. The median filter method is very useful for eliminating ouliners/noise. Fig 3(c) shows the image noise in lung cancer tomography has been reduced. The adaptive histogram equations aims to improve the image quality so that it is clearer and more intelligible. This can be seen from the image of the result of the adaptive histogram equation, where the image results are ready for image segmentation process.

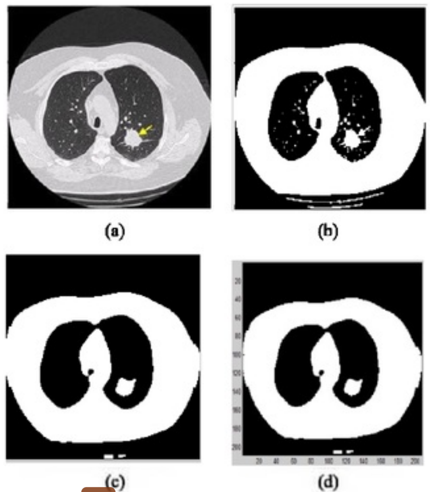


Fig 3. (a) *Grayscale* image (b) *Binary image* (c) Noise removal result using median filter (d) Results of Histogram adaptive equalization process

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.

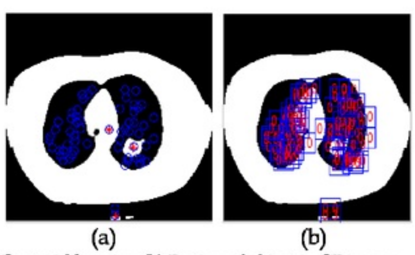


Fig 4. (a) Centroid location (b) Region of object in CT image

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°,45 °, 90 °, 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):

$$Contrast = \sum i \sum j (i-j)^2 C(i,j) \quad (1)$$

$$Energy = \sum i \sum j C^2(i,j) \quad (2)$$

$$Entropy = \sum i \sum j C(i,j) \log C(i,j) \quad (3)$$

$$Homogeneity = \sum i \sum j \frac{C(i,j)}{1+|i+j|} \quad (4)$$

$$Correlation = u\sum_i^i = 1 \sum_j^i = 1 \quad (5)$$

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	0.4959	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

Image	Contrast	
	Lungs Cancer	*Normal*
4	0.0331	0.0325
5	0.0382	0.0510
6	0.0395	0.0255
7	0.0418	0.0412
8	0.0641	0.0403
9	0.0383	0.0315
10	0.0331	0.0327

The Table 3 shows the normal lung image contrast range in the range of 0.414-0.0390, whereas the lung image is not normal between 0.0408-0.0368.

Based on Table 4 range of image correlation of lung cancer ranged from 2,6779-2,5487, while the normal image between 2,5640-2,5097.

TABLE IV. CORRELATION FEATURE

Image	Correlation	
	Lungs Cancer	*Normal*
1	2.6779	2.5640
2	2.5487	2.5097
3	2.6447	2.6622
4	2.6746	2.6823
5	2.5157	2.7544
6	2.7000	2.8420
7	2.6114	2.4980
8	2.7864	2.7444
9	2.4263	2.7723
10	2.6586	2.5872

TABLE V. VARIANS FEATURE

Image	Varians	
	Lungs Cancer	*Normal*
1	2.6779	2.5640
2	2.5487	2.5097
3	2.6447	2.6622
4	2.6746	2.6823
5	2.5157	2.7544
6	2.7000	2.8420
7	2.6114	2.4980
8	2.7864	2.7444
9	2.4263	2.7723
10	2.6586	2.5872

Table 5 shows that the normal lung image has a higher value than the image of lung cancer. Range of normal image energy ranged between 0.2886-0.2948, while the image of lung cancer between 0.2573-0.2857.

The results of the homogeneity, energy, contrast, correlation, and variance features in Table 1-5 above some values are overlap between normal and cancer images, therefore it is not possible to do linear programming. This is why ANN backpropagation used to detect between cancerous and normal images.

D. Cancer Lung Detection

Training is done using backpropagation, each category of image has a different target value. In the normal image of the target used 1, the image of target lung cancer used is 0. The testing is done using the input parameters and update weight obtained in training process (learning rate 0.3, *hidden layer value* 20, and epoch 1000). We decide applied learning rate 0.3, *hidden layer value* 20, and epoch 1000 based on the experiment that already done, using that parameter produce the highest accuracy than another parameter. Table 6 is the results of the experiments.

TABLE VI. LEARNING RATE VARIATION TOWARDS MSE VALUE

Learning rate	MSE
0.4	0.08
0.6	0.06
0.3	0.09

Confusion matrix (Table 7) shows that lung cancer CT-image data classified as lung cancer 80%, and normal data classified as normal is 80%.

TABLE VII. CONFUSION MATRIX

Prediction	Lung cancer	Normal	Total
Lung cancer	80%	20%	100%
Normal	20%	80%	100%

IV. CONCLUTION

The preprocessing stage, image segmentation, feature extraction, and learning process have been done. The weight generated from the learning process is used to test 50 CT-image data. The results show that the system works with accuracy 80%. In order to obtain higher accuracy, further research is needed by improving the preprocessing process, image segmentation, feature extraction, and learning process.

ACKNOWLEDGMENT

This research was supported by Universitas Negeri Surabaya. We would also like to show our gratitude to Cancerimagingarchive Database for sharing their lung CT-image and biomarker this research.

REFERENCES

[1] American Cancer Society. (2013). What are the risk factors for non-small cell lung cancer?. 19 Desember 2013. < http : // www . cancer . org>.

[2] Ancuceanu, R. V., and Victoria, I, 2004, Pharmacologically Active Natural Compounds for Lung Cancer, Altern. Med. Rev., 9, 4, 402-419

[3] M.S.Tarawneh, "Lung Cancer Detection Using Image Processing Techniques", Leonardo Electronic Journal of Practices and Technologies, Issue 20, pp. 147-158, January-June 2012.

[4] Kusumadewi, S.. *Membangun Jaringan Syaraf Tiruan dengan Matlab dan Exel Link*. Yogyakarta: Graha Ilmu.

[5] S. Singh, R. Vijay, Y. Singh, "Artificial Neural Network and Cancer Detection", IOSR Journal of Computer Engineering, pp. 20-24.

[6] F. Taher, N. Werhi, H. Al-Ahmad, R. Sammouda, "Lung Cancer Detection by Using Artificial Neural Network and Fuzzy Clustering Methods", American Journal of Biomedical Engineering 2012, 2(3), pp. 136-142.

[7] Almas Pathan, Bairu.K.saptalkar, "Detection and Classification of Lung Cancer Using Artificial Neural Network", International Journal on Advanced Computer Engineering and Communication Technology Vol-1 Issue:1.

[8] A. Badnjevic, M. Cifrek, D. Koruga, D. Osmankovic, "Neuro-fuzzy classification of asthma and chronic obstructive pulmonary disease", BMC Med Inform Decis Mak. 2015; 15(Suppl 3): S1. Published online 2015 Sep 11.

[9] R.M.Haralick, K. Shanmugam, R.M.Haralick, K. Shanmugam,, "Textural features for image classification," *IEEE Trans. System Man. Cybernetics*, vol. SMC-3, pp. 610–621, 1973.

[10] D. Tian, "A Review on Image Feature Extraction and Representation Techniques", International Journal of Multimedia and Ubiquitous Engineering Vol. 8, No. 4, July, 2013, pp 385-396.

Cancer Lungs Detection on CT Scan Image Using Artificial Neural Network Backpropagation Based Gray Level Coocurrence Matrices Feature

ORIGINALITY REPORT

23%	14%	16%	16%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	docplayer.net Internet Source	4%
2	Submitted to Vels University Student Paper	3%
3	www.ijcsmc.com Internet Source	2%
4	ijcsit.com Internet Source	1%
5	"Table of contents", 2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 2017 Publication	1%
6	Lilik Anifah, Tati Latifah R. Mengko, Mauridhi Hery Purnomo, I Ketut Eddy Purnama. "Osteoarthritis Classification Using Self Organizing Map Based Gray Level Run Length Matrices", 2017 5th International Conference	1%

on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME), 2017
Publication

7	es.scribd.com Internet Source	1%
8	Submitted to Universiti Teknologi Malaysia Student Paper	1%
9	Submitted to College of Engineering Trivandrum Student Paper	1%
10	link.springer.com Internet Source	1%
11	f1000.com Internet Source	1%
12	Submitted to University of Moratuwa Student Paper	1%
13	Akkoyun, Ozgur, Nicola Careddu, and Graziella Marras. "An Investigation of Colour Properties of Orosei Limestones (Italy) by Using Computer Program", Key Engineering Materials, 2013. Publication	1%
14	M. Octaviano Pratama, Aniati Murni Arymurthy. "Automatic land cover classification of geotagged images using ID3, Naïve Bayes and	<1%

Random Forest", 2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 2017
Publication

15	M. Bigioni. "Antitumour effect of combination treatment with Sabarubicin (MEN 10755) and cis-platin (DDP) in human lung tumour xenograft", Cancer Chemotherapy and Pharmacology, 09/2008 Publication	<1%
16	M. Anwar Ma'sum, Ito Wasito, Adi Nurhadiyatna. "Intelligent K-Means clustering for expressed genes identification linked to malignancy of human colorectal carcinoma", 2013 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 2013 Publication	<1%
17	www.ijert.org Internet Source	<1%
18	www.ijeit.com Internet Source	<1%
19	Submitted to University of Sheffield Student Paper	<1%
20	www.thieme-connect.com Internet Source	<1%

21 ccrc.farmasi.ugm.ac.id
Internet Source
<1%

22 "Emerging Intelligent Computing Technology and Applications", Springer Science and Business Media LLC, 2009
Publication
<1%

23 Submitted to Universiti Teknologi Petronas
Student Paper
<1%

24 "31st International EuroPACS Meeting", International Journal of Computer Assisted Radiology and Surgery, 2013
Publication
<1%

25 N Kavya, N Usha, N. Sriraam, D Sharath, Prabha Ravi. "Breast Cancer Detection using Non Invasive Imaging and Cyber Physical System", 2018 3rd International Conference on Circuits, Control, Communication and Computing (I4C), 2018
Publication
<1%

26 "Soft Computing for Problem Solving", Springer Science and Business Media LLC, 2019
Publication
<1%

27 L. Shajy, P. Smitha, E. Boney Shanker, Varghese Paul. "Segmentation and feature extraction of sputum cell for early detection of lung cancer", 2014 First International
<1%

Conference on Computational Systems and Communications (ICCSC), 2014
Publication

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		