FEATURE EXTRACTION OF LUNG CANCER USING IMAGE ANALYSIS TECHNIQUES

L.T. ALAYUE, B.S. GOSHU[#], ENDRIS TAJU

Department of Physics, Dire Dawa University, [#]e-mail: belaysitotaw@gmail.com

Abstract. Lung cancer is one of the most life-threatening diseases. It is a medical problem that needs accurate diagnosis and timely treatment by healthcare professionals. Although CT is preferred over other imaging modalities, visual interpretation of CT scan images may be subject to error and can cause a delay in lung cancer detection. Therefore, image processing techniques are widely used for early-stage detection of lung tumors. This study was conducted to perform pre-processing, segmentation, and feature extraction of lung CT images using image processing techniques. We used the MATLAB programming language to devise a stepwise approach that included image acquisition, pre-processing, segmentation, and features extraction. A total of 14 lung CT scan images in the age group of 55–75 years were downloaded from an open access repository. The analyzed images were grayscale, 8 bits, with a resolution ranging from 151×213 to 721×900, and Digital Imaging and Communications in Medicine (DICOM) format. In the pre-processing stage median filter was used to remove noise from the original image since it preserved the edges of the image, whereas segmentation was done through edge detection and threshold analysis. The results show that solid tumors were detected in three CT images corresponding to patients aged between 71 and 75 years old. Our study indicates that image processing plays a significant role in lung cancer recognition and early-stage treatment. Health professionals need to work closely with medical physicists to improve the accuracy of diagnosis.

Key words: CT image, lung cancer detection, MATLAB, image processing.

INTRODUCTION

Cancer is the second-leading cause of death and an important barrier to increasing the life expectancy of human beings [10]. Among all different types of cancers, lung cancer is the most prevalent, having the highest mortality rate [3].

Computed tomography (CT) scans are used to identify the lung cancer as it provides a detailed picture of the tumor and tracks its growth [2]. Although CT is preferred over other imaging modalities, visual interpretation of CT scan images may be subject to error and can cause a delay in lung cancer detection [2, 4, 25].

Therefore, image processing techniques are widely used in medical fields for early-stage detection of lung tumors [33].

The rate of lung cancer exposure in five years period is 16 %; however, if the nodules are detected at an early stage, the survival rate can be increased to 81–85 % [18]. Lung nodule detection is primarily done manually by trained pulmonary radiologists with the help of computer-aided diagnosis systems [10, 12]. Such systems are designed to be highly sensitive for nodule examinations and classifications [2, 10, 12, 14, 15].

The main issues within the high-level processing of biomedical images result from their complex nature. Image analysis helps to interpret an image using a computer and specifically designed software [16, 17, 19, 31]. The purpose of this work is to assess feature extraction of lung cancer using image analysis techniques.

LITERATURE REVIEW

IMAGE PROCESSING

Image processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it. It is a type of signal processing in which the input is an image, and the output may be image or characteristics/features associated with that image [19]. Nowadays, image processing is a rapidly growing research field. It forms core research area within engineering and computer science disciplines too. Image processing basically includes the following three steps: importing the image via image acquisition tools, analysing and manipulating the image, generating an output consisting of an altered image or a report based on image analysis.

FEATURE EXTRACTION

Feature extraction is a part of a process, in which an initial set of the row data is divided and reduced to manageable groups. The physical dimensional measurement from an image describes the physical appearance of the image. The most common features extracted are area, centroid, bounding box, and mean [23, 34]. The physical measurement, texture feature extraction uses the gray level co-occurrence matrix (GLCM) methods. The GLCM based extraction characterizes the texture by calculating the proportion of pairs of pixels with specific values and then extracting statistical measures from the matrix. Statistical parameters generated through GLCM values included contrast, correlation, energy, and homogeneity [13, 21, 35].

Image segmentation

Image segmentation is the process by which a digital image is partitioned into various subgroups (of pixels) called image objects, which can reduce the complexity of the image, and thus the image becomes simpler to analyze [19, 26, 27, 33]. This process assists in detecting critical parts of the image that are not displayed in the original image using thresholding and edge detection [19]. Edge detection is used to identify the border of the image. The gradient operators of the Canny methods are better than others [28, 33]. The image border is well demarcated using the Canny, Sobel and Prewitt method showed by [23, 28].

Thresholding is one of the most powerful tools for image segmentation. The implementation of threshold helps in two main issues: removal of unnecessary information from the image (background activity) and the appearance of details not easily detected in the original image. The thresholding method has the advantages of smaller storage space, fast processing speed, and ease in manipulation. Thresholding of the given images in this study was done at the value of threshold (*Thr)* is 0.99.

MATERIALS AND METHODS

Image processing and analysis were done using a high-level programming language MATLAB [24]. It is used for technical computing, as well as for biomedical data analysis [1, 7, 31]. It integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation.

In this work, the images were taken from the lung cancer CT scan image database, lung cancer.org [37]. The acquired images were all grayscale, 8-bit, DICOM format with a resolution ranging from $151 \times 213 \times 721 \times 900$ [23, 37].

SAMPLE SIZE

A total of 14 lung CT scan images for men aged between 55 and 75 years were collected from the online database [23, 37].

PROCEDURES

Detection of lung cancer using image analysis techniques was conducted in four steps or procedures shown in Figure 1.
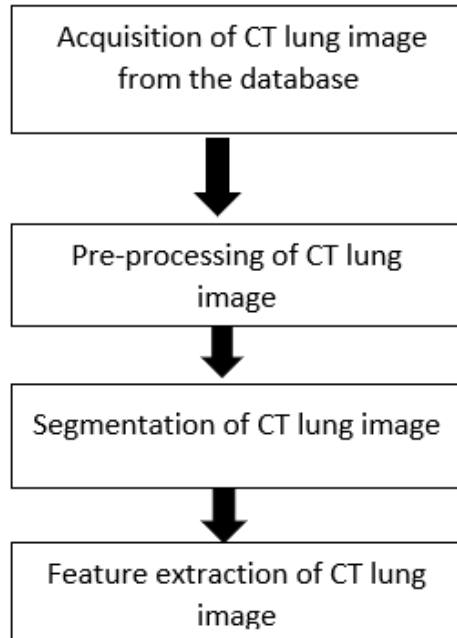
Fig. 1. Block diagram of the CT image feature extraction process [19].

ACQUISITION OF LUNG CT SCAN IMAGE

The first step in image processing is the collection of images from the available database. The images are in DICOM format. CT scan images contain prognostic information, provide pictures of tissues, organs, and skeletal structures to obtain a suitable representation of the image for the segmentation process, low noise, and distortion. The gray level mean pixels value, centroid, entropy, and bounding box were computed.

Rescaling

The images used in the analysis are resized in different scales to determine how various sizes affect the analysis process. Different image sizes carry different information that's why the best image size needs to be examined in detail. The purpose of image resizing is to produce a lower data size, which hastens the processing time [25]. The size of the input image was rescaled by [512, 512] since the input image should be in square matrix form for segmentation and feature extraction.

Image enhancement

Image enhancement is the process of adjusting digital images so that the results are more suitable for display or further image analysis. In some cases, the image has unrelated or insufficient information for extracting due to different defects [7, 8].

Background removal

The background portion is the unwanted region that cannot be considered for analysis and was removed by choosing the region of interest in the given image.

Median filter

The median filter is one of the well-known order-statistic filters due to its good performance for some specific noise types such as "Gaussian", "random", and "salt and pepper" noises [7, 8]. Accordingly, the median filter, the center pixel of the $M \times N$ neighbor pixel is replaced by the median value of the corresponding window. We use this filter to remove the noise pixels.

Image segmentation

Image segmentation is the process in which a digital image is partitioned into multiple segments that correspond to different classes of pixels. It represents the image in clear view and is done by the following techniques:

Edge detection

Edge is a set of connected pixels that lie on the boundary between two regions. Local transitions among different image intensities constitute an edge [24]. So, edge detection is performed by evaluating the intensity gradient. The edge detection process involves small kernels that convolve with an image to estimate the first-order directional derivatives of the image brightness distribution [3, 8, 9, 32]. Kernels are pre-defined groups of edge models that match each image segment of a fixed size. The edge value is calculated by forming a matrix centered on each pixel. If the value is not smaller than a given threshold, then the pixel is categorized as an edge. Edge detection algorithms are grouped into two categories, namely, gradient and Laplacian operator respectively. In this study, edge detection was done using the gradient operator [24, 32].

Feature extraction

A feature is a significant piece of information that provides a more entailed understanding of its nature. In this study, textural feature extraction was made based

on the Gray Level Co-occurrence Matrix (GLCM). GLCM is a matrix where the rows and columns are equivalent to the number of gray levels, G in the image [36]. In this research, texture-based features were analyzed, including area, centroid, bounding box, energy, entropy, contrast, correlation, homogeneity, and statistical features such as mean [14, 19].

Area

The area of the image portion is obtained by the summation of the total number of constituent pixels.

Energy

The energy corresponding to a certain image describes the uniformity of the constituent pixels. It lies between 0 and 1. If the image is perfectly uniform, then the corresponding energy level has a value of 1.

Correlation

The correlation measures the extent to which a pixel value is correlated with the values of neighboring pixels. Its value lies between –1 and +1. The neighboring pixels is –1, for perfectly negatively correlation, and +1 for the positively correlation.

Homogeneity

It gives the distribution value of the closeness of the GLCM to the diagonal of GLCM. It takes values between 0 and 1.

Contrast

The contrast is expressed in terms of local variations in the GLCM that determines the intensity difference between a pixel and its neighborhood. The contrast is 0 for a uniform image. Contrast adjustment remaps image intensity values to the full display range of the data type. An image with high contrast has sharp borders between black and white portions.

Mean

The mean is obtained by calculating the arithmetic mean of all the values of gray pixels that constitute the region of interest (ROI) of the image.

Centroid

Centroid is the geometric center of an image. It is a fixed point located at the intersection of all hyperplanes of symmetry within that image. The centroid point is mathematically expressed as the coordinates "[x, y] = centroid (polyin, image)" [20, 25].

Bounding box

Bounding box is an imaginary rectangle that serves as a point of reference for object detection and creates a collision box for the image. Data annotators draw these rectangles over images, outlining the object of interest within each image by defining its x and y coordinates.

RESULTS AND DISCUSSIONS

RESULTS

Characteristics of input images

In this work, 14 CT scan images were considered of male adults in the age range of 55–75 years [23, 37]. The mean and standard deviation of the age are 66.07±6.73. The characteristics of the analyzed images are presented in Figures 2–5.

a)   b)   c)   d)
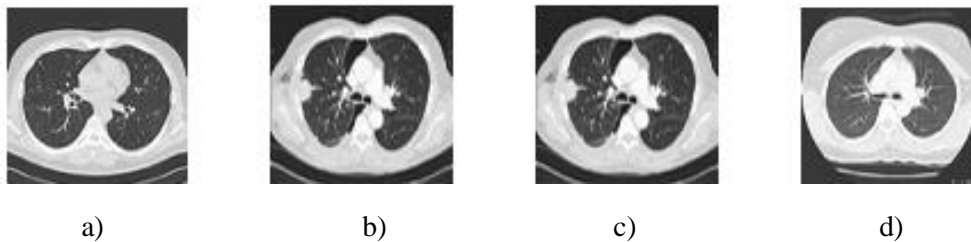
Fig. 2. Input images of CT scan [23, 37] a) Im1, 55 years old, b) Im2, 56 years old, c) Im3, 58 years old, and d) Im4, 61 years old.

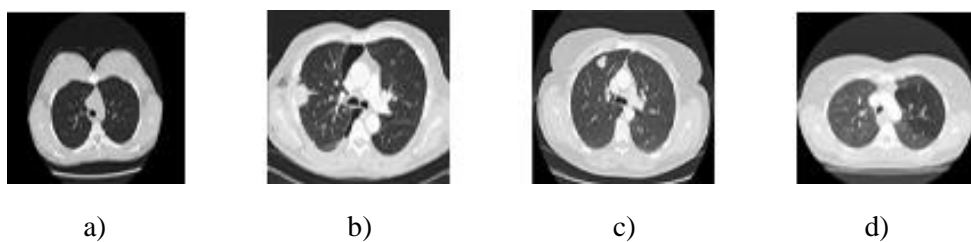a)   b)   c)   d)

Fig. 3. Input images of CT scan [23, 37] a) Im5, 63 years old, b) Im6, 64 years old, c) Im7, 67 years old, and d) Im9, 67 years old.

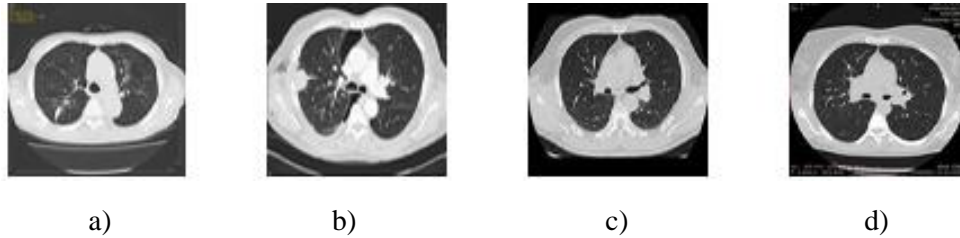a)   b)   c)   d)

Fig. 4. Input images of CT scan [23, 37] a) Im9, 69 years old, b) Im10, 71 years old, c) Im11, 72 years old, and d) Im12, 73 years old.
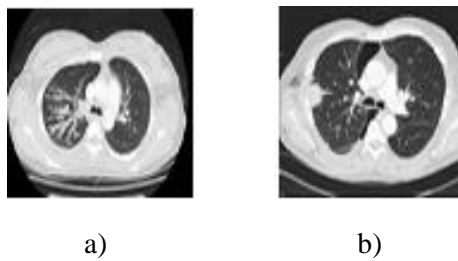

a)   b)

Fig. 5. Input images of CT scan [23, 37] a) Im13, 74 years old, and b) Im14, 75 years old.

Background removal and selecting the ROIs of an image are significant steps in the analysis. It is especially relevant when the geometry of the sample does not cover the whole area measured. The surrounding image is usually composed of highly noisy fields, and thus, it might hinder the performance of another process. Moreover, removing it implies a substantial saving of computing time. We did the background removal, contrast adjustment, and brightness adjustment of our sample images shown in Figures 6–19. Figures 6–19(a) shows the rescaled images of the lung, Figures 6–18(b) shows the background of images that were removed, Figures 6–19(c) shows the contrast adjusted image using morphological processes, and Figures 6–19(d) shows the brightness adjusted images of all samples.

a)   b)   c)   d)
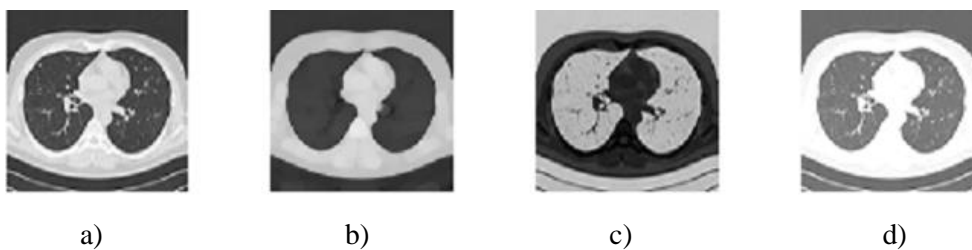
Fig. 6. Image characteristics of Im1 a) rescaled, b) background removed, c) contrast adjusted, and d) brightness adjusted images.
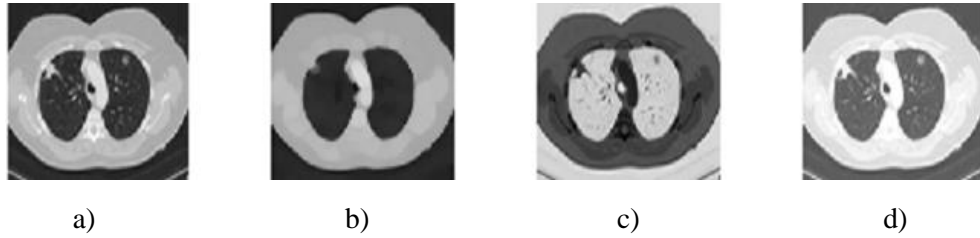
Fig. 7. Image characteristics of Im2, a) rescaled, b) background removed, c) contrast adjusted, and d) brightness adjusted image.
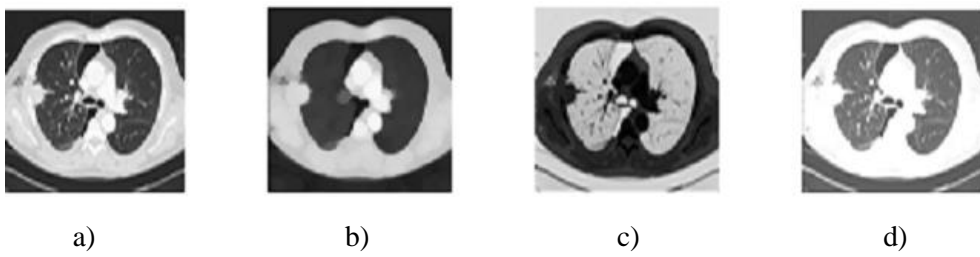

Fig. 8. Image characteristics of Im3, a) rescaled, b) background removed, c) contrast adjusted, and d) brightness adjusted image.

Fig. 9. Image characteristics of Im4, a) rescaled, b) background removed, c) contrast adjusted, and d) brightness adjusted image.
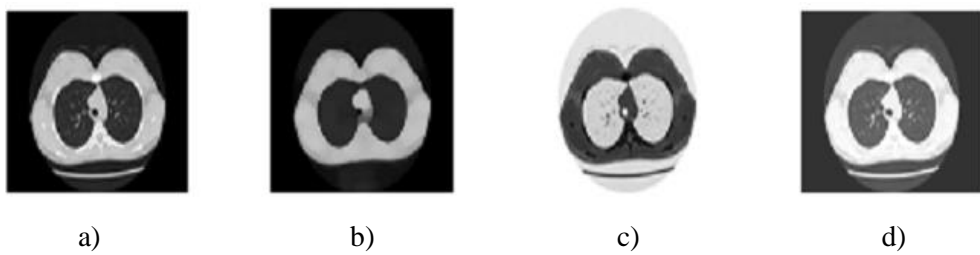

Fig. 10. Image characteristics of Im5, a) rescaled, b) background removed, c) contrast adjusted, and d) brightness adjusted image.

Fig. 11. Image characteristics of Im6, a) rescaled, b) background removed, c) contrast adjusted, and d) brightness adjusted image.

Fig. 12. Image characteristics of Im7, a) rescaled, b) background removed, c) contrast adjusted, and d) brightness adjusted image.
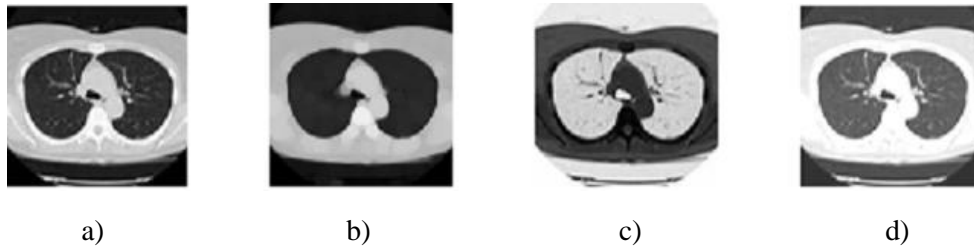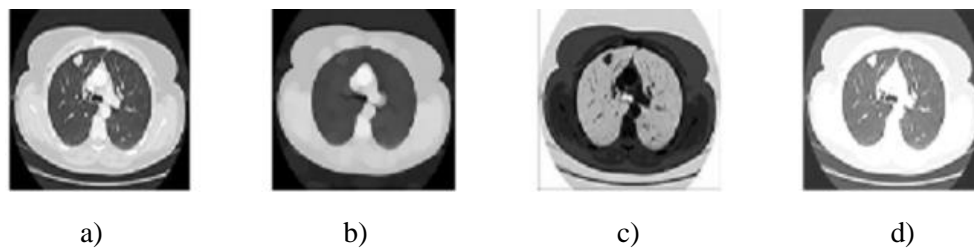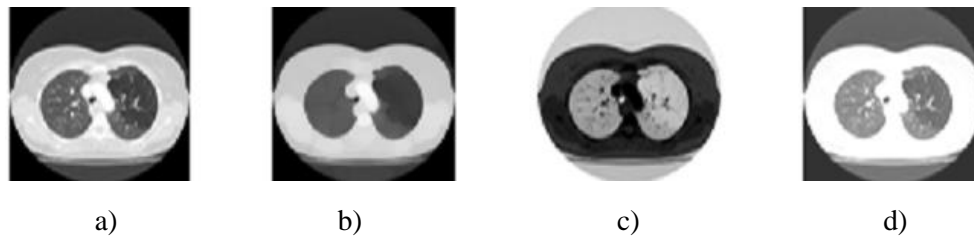

Fig. 13. Image characteristics of Im8, a) rescaled, b) background removed, c) contrast adjusted, and d) brightness adjusted image.
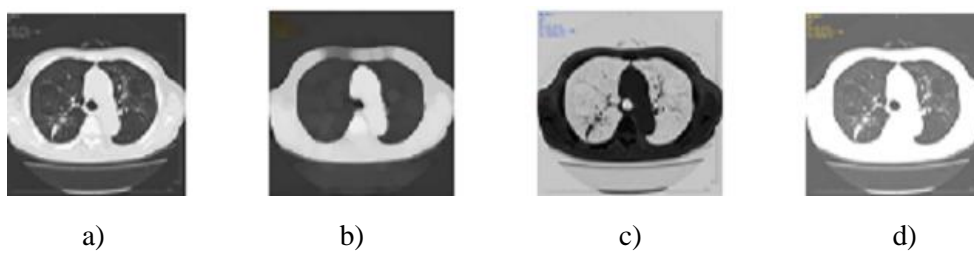

Fig. 14. Image characteristics of Im9, a) rescaled, b) background removed, c) contrast adjusted, and d) brightness adjusted image.

a)   b)   c)   d)

Fig. 15. Image characteristics of Im10, a) rescaled, b) background removed, c) contrast adjusted, and d) brightness adjusted image.

a)   b)   c)   d)

Fig. 16. Image characteristics of Im11, a) rescaled, b) background removed, c) contrast adjusted, and d) brightness adjusted image.

a)   b)   c)   d)

Fig. 17. Image characteristics of Im12, a) rescaled, b) background removed, c) contrast adjusted, and d) brightness adjusted image.

a)   b)   c)   d)

Fig. 18. Image characteristics of Im13, a) rescaled, b) background removed, c) contrast adjusted, and d) brightness adjusted image.

a)   b)   c)   d)
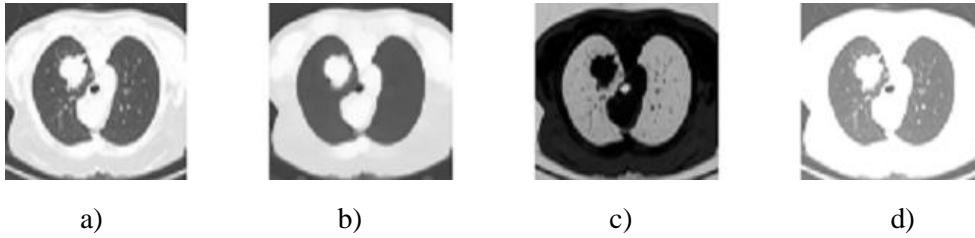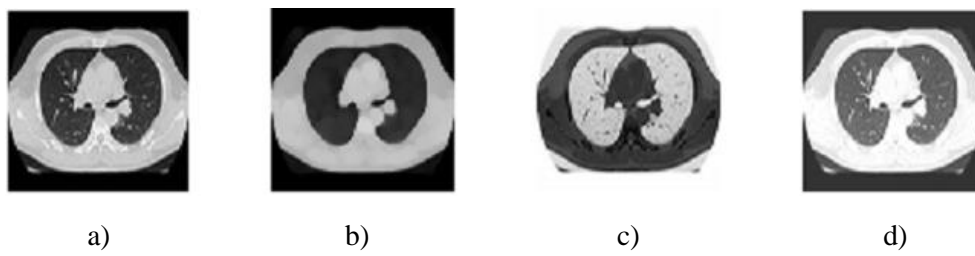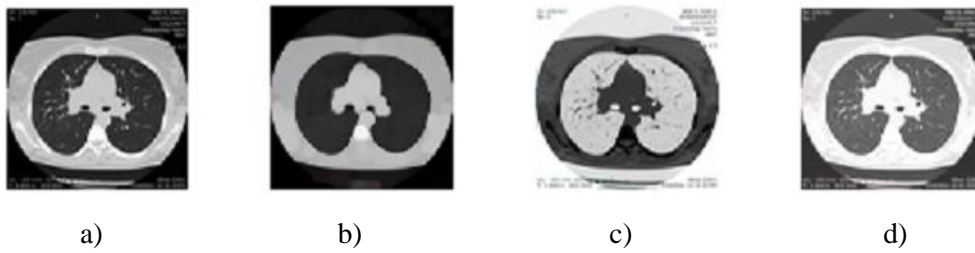
Fig. 19. Image characteristics of Im14, a) rescaled, b) background removed, c) contrast adjusted, and
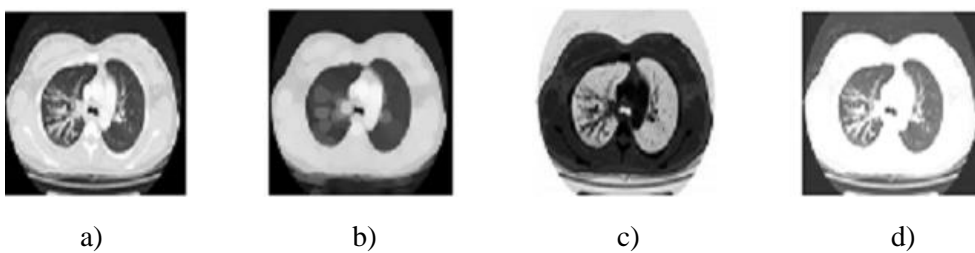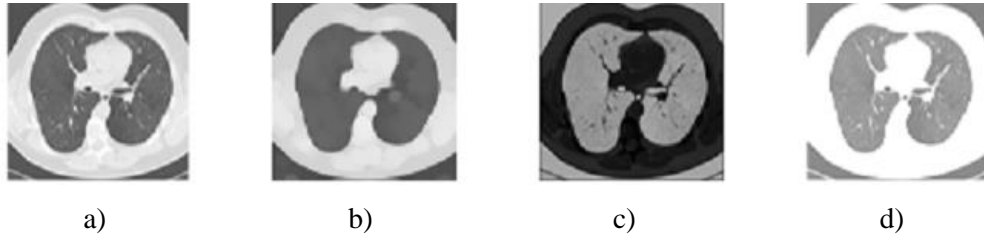d) brightness adjusted image.

Segmentation is the process of partitioning a digital image into multiple portions. The goal of segmentation is to simplify and change the representation of an image into something more meaningful and easier to analyze. In this study, image segmentation was done through edge detection and threshold analysis of the given image as follows.

Edge is used for finding the border of objects in the image. We did this through the Canny method since the image border is well demarcated [5, 6]. The image border is well demarcated in the Canny method as shown in Figs. 20–22 below.

Thresholding is one of the most powerful tools for image segmentation. The thresholding method has the advantages of smaller storage space, fast processing speed, and ease in manipulation. Thresholding of the given image was done at the value of $Thr = 0.99$, and the results are shown in Figure 20–22. Out of fourteen input images [3], lung cancers were found only in three of the input images, shown in Figures 20–22.

a)   b)   c)

Fig. 20. Results of the segmented image (Im10) through Canny edge detection and thresholding method, a) background removed, b) edge detection with Canny methods, and c) segmented image at a value of $Thr = 0.99$.

a)   b)   c)

Fig. 21. Results of the segmented image (Im11) through Canny edge detection and thresholding method, a) background removed, b) Edge detection with Canny methods, and c) Segmented image at a value of $Thr = 0.99$.
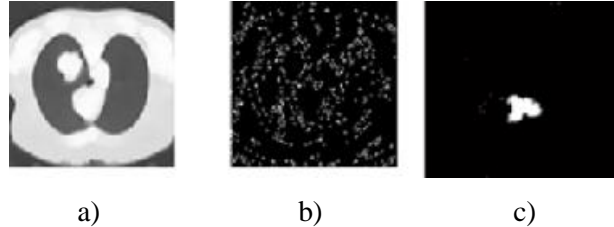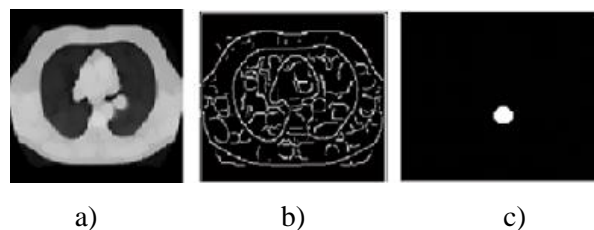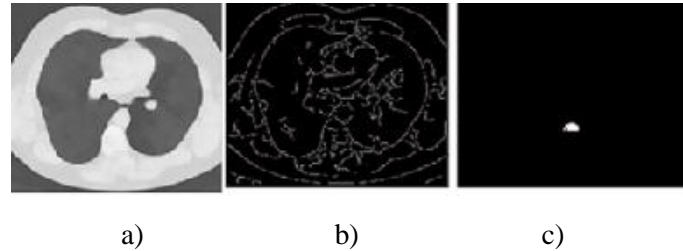
a)   b)   c)

Fig. 22. Results of the segmented image (Im14) through Canny edge detection and thresholding method, a) background removed, b) edge detection with Canny methods, and c) segmented image at a value of *Thr* = 0.99.

Once the segmentation is done, information is extracted, physical dimensional measured, and texture features extracted. The physical dimensional measurement describes and characterizes the physical appearance of the image. Extracted area, centroid, bounding box, and mean are shown in Table 1.

Table 1

Extracted features of lung CT scan images after appropriate image processing techniques

Input images	Physical dimensional measurement features			
	Area	Centroid	Bounding box	Mean
Im1	14502	[501.6103 221.2377 2]	[88.5000 31.5000 0.5000 773 514 3]	2.24e+03
Im2	12765	[487.2127 175.5391 2]	[88.5000 31.5000 0.5000 699 413 3]	1.92e+03
Im3	7323	[405.4674 146.0274 2]	[88.5000 31.5000 0.5000 513 377 3]	1.59e+03
Im4	16758	[478.9348 239.5843 2]	[88.5000 31.5000 0.5000 697 442 3]	1.96e+03
Im5	14580	[446.5934 209.5924 2]	[88.5000 31.5000 0.5000 641 410 3]	1.86e+03
Im6	7032	[408.9006 149.1822 2]	[88.5000 31.5000 0.5000 513 377 3]	1.59e+03
Im7	12816	[413.9913 183.6053 2]	[88.5000 31.5000 0.5000 698 406 3]	1.87e+03
Im8	14376	[462.8066 247.9821 2]	[88.5000 31.5000 0.5000 657 473 3]	2.01e+03
Im9	13824	[410.4625 237.6454 2]	[86.5000 31.5000 0.5000 698 469 3]	1.99e+03
Im10	9363	[370.4777 205.4771 2]	[86.5000 31.5000 0.5000 554 365 3]	1.64e+03
Im11	13755	[413.9716 222.0757 2]	[86.5000 31.5000 0.5000 702 437 3]	1.95e+03
Im12	15951	[482.9707 241.2770 2]	[86.5000 31.5000 0.5000 701 469 3]	2.03e+03
Im13	14112	[482.3310 192.2798 2]	[88.5000 31.5000 0.5000 697 413 3]	1.90e+03
Im14	7095	[402.7336 143.9742 2]	[88.5000 31.5000 0.5000 513 377 3]	1.59e+03

Table 2

Texture analysis of lung CT scan images

Input images	Text features				Lung cancer status
	Contrast	Correlation	Energy	Homogeneity	
Im1	0.43	0.98	0.57	0.99	Absent
Im2	0.50	0.98	0.54	0.99	Absent
Im3	0.54	0.98	0.51	0.99	Absent
Im4	0.43	0.98	0.55	0.99	Absent

Input images	Text features				Lung cancer status
	Contrast	Correlation	Energy	Homogeneity	
Im5	0.57	0.97	0.53	0.99	Absent
Im6	0.47	0.98	0.52	0.99	Absent
Im7	0.40	0.98	0.55	0.99	Absent
Im8	0.54	0.97	0.55	0.99	Absent
Im9	0.42	0.98	0.56	0.99	Absent
Im10	0.56	0.98	0.52	0.99	Present
Im11	0.49	0.98	0.55	0.99	Present
Im12	0.51	0.98	0.55	0.99	Absent
Im13	0.40	0.98	0.55	0.99	Absent
Im14	0.52	0.98	0.52	0.99	Present

The second order statistics are based on a co-occurrence matrix, which includes contrast, correlation coefficients, energy, homogeneity, and status of lung cancer, is shown in Table 2. The contrast values of present and absent of lung cancers were between 0.40 and 0.57. The correlation coefficients of all images were 0.98, except Im8 in which it was 0.97. The energy parameter for both presence and absence of lung cancers were ranged from 0.51 to 0.57. The homogeneity for all images was the same.

The results illustrated in Table 2 show that the contrast for the three cancerous images Im10, Im11, and Im12 were 0.56, 0.49, and 0.52 respectively. Different works show that thresholding techniques is an efficient way for image segmentation and would be a good candidate to significantly differentiate between the presence and absence of cancer images [3, 14]. The segmentation is based directly on pixel intensities [14].

DISCUSSION

In this work, three of the lung CT images indicated the presence of cancer as shown by the segmentation outputs. From the total acquired images, Im10, Im11, and Im14 were found to indicate lung cancer. The detected nodules had different sizes, as shown by the feature analysis result [26, 29].

In the pre-processing stage, image enhancement techniques were employed to improve the quality of the input image for further image processing. Image enhancement was done through image rescaling, filtering with morphological factors, brightness and contrast adjustment, dilation, erosion, histogram equalization, and median filtering.

The image pre-processing stage starts with image enhancement to improve the interpretability or perception of information included in the image for human viewers or to provide better input for other automated image processing techniques. In this work, image enhancement was done through contrast adjustment. The enhancement was comparable with other works [11, 16, 30].

Image enhancement was also done through contrast adjustment, which was also consistent with other studies [19, 22–24, 28]. The acquired image does not usually present the desired contrast. Thus, improvement of contrast is needed as the organ shape, boundaries, and internal functionality can be better detected. In addition, organ demarcation can be achieved in many cases without removing the background activity. The MATLAB command that implements contrast processing is the *imadjust* [20, 24]. In this technique, the contrast of an image can be enhanced or reduced based on the need.

Edge detection was applied for finding the border of objects in the image [27]. We did this through one of the gradient operators of the Canny methods, edge detections are consistent with other works [28, 33]. The image border is well demarcated using the Canny method compared with Sobel and Prewitt edge detection method showed by [22, 28].

In the feature extraction stage, physical dimensional measurement and texture-based features were extracted. In this work, we extracted area, centroid, bounding box, and mean. This work is consistent with the method and results of other studies [13, 21].

Following the physical measurement, texture feature extraction was done using the gray level co-occurrence matrix (GLCM) methods shown in Table 1 and 2. Statistical parameters were generated through GLCM values included showing contrast, correlation, energy, and homogeneity. This work is consistent with other image processing studies [17, 21, 35].

In this work, the segmentation outputs of three lung CT images indicate the presence of cancer. This result is in line with the American Cancer Society [30] evidence as most people diagnosed with lung cancer are 65 years old and above, with the average at diagnosis being 70. According to the American Cancer Society, the largest group with a new lung cancer diagnosis was the age 70 to 74 years, followed by the 65–69 years old group [33].

CONCLUSION

Currently, lung cancer is the second contributor to morbidity and mortality in the human population. In this research work, pre-processing, segmentation, and feature extraction of the CT images of the lung was performed using an original software written in MATLAB programming language. First, a stepwise enhancement was applied to the original image to improve the quality. Then, the median filter method was applied to reduce the noise of the image. Image segmentation was done through edge detection and threshold analysis. In the feature extraction stage, some physical and statistical features were extracted to support the diagnosis of abnormality on the lung tissue. The image analysis procedure developed here will improve the accuracy of the supports of the medical healthcare providers involved in the diagnosis of lung cancer. Overall, image processing will reduce morbidity and mortality of lung cancer

patients through timely and accurate diagnosis. Moreover, our results suggest that quantitative CT texture analysis has the potential to differentiate primary lung cancer from granulomatous lesions using thresholding techniques for segmentation. Nevertheless, texture analysis is not expected to replace tissue diagnosis but may be added to the radiologist's arsenal used to characterize spiculated or lobulated solid lung nodules especially in areas where granulomatous infections are endemic.

The validity and specificity of the image resulted from CT can be increased using image processing for each suspected patient. Image processing combined with medical know-how will improve the accuracy of diagnosis. Therefore, image processing, and analysis techniques should be done to make the clinician's decisions more reliable. Image processing needs to be done to reduce false positive and negative results, which plays a significant role in lung cancer recognition and early-stage treatment.

R E F E R E N C E S

1. AL-TARAWNEH, M.S., Lung cancer detection using image processing techniques, *Leonardo Electronic Journal of Practices and Technologies,* 2012*,* **11**(21), 147–158.
2. ARMATO, S.G., F. LI, M.L. GIGER, H. MACMAHON, S. SONE, K. DOI, Lung cancer: Performance of automated lung nodule detection applied to cancers missed in a CT screening program, *Radiology*, 2002*,* **225**(3), 685–692.
3. BIANCONI, F., M.L. FRAVOLIN, R. BELLO-CEREZO, M. MINESTRINI, M. SCIALPI, B. PALUMBO, Evaluation of shape and textural features from CT as prognostic biomarkers in non-small cell lung cancer, *Anticancer research*, 2018, **38**, 2155–2160.
4. BRAY, F., J. FERLAY, I. SOERJOMATARAM, R.L. SIEGEL, L.A. TORRE, A. JEMAL, Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA: A Cancer Journal for Clinicians,* 2018*,* **68**(6), 394–424.
5. CHAUDHARY, A., S. S. SINGH, Lung cancer detection on CT images by using image processing, Paper presented at the *International Conference on Computing Sciences*, 2012, 142–146.
6. DOLEJŠÍ, M., *Detection of Pulmonary Nodules from CT Scans*, Master's Thesis, Czech Technical University in Prague, 2007.
7. EL-BAZ, A., G.M. BEACHE, G. GIMEL'FARB, K. SUZUKI, K. OKADA, A. ELNAKIB, B. ABDOLLAHI, Computer-aided diagnosis systems for lung cancer: Challenges and methodologies, *International Journal of Biomedical Imaging,* 2013**, 2013,** 1–45.
8. FITZMAURICE, C., D. ABATE, N. ABBASI *et al*., Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2017: A systematic analysis for the Global Burden of Disease Study, *JAMA Oncology*, 2019, **5**(12), 1749–1768.
9. FUJITA, H., X. ZHANG, S. KIDO, T. HARA, X. ZHOU, Y. HATANAKA, R. XU, Computer-aided diagnosis: The emerging of three CAD systems induced by Japanese health care needs, *Comput. Methods Prog. Biomed.*, 2008, **92**, 238–248.
10. HARALICK, R.M., K. SHANMUGAM, I.H. DINSTEIN, Textural features for image classification, *IEEE Transactions on Systems, Man, and Cybernetics*, 1973, **6**, 610–621
11. JACOBS, C., E.M. VAN RIKXOORT, T. TWELLMANN, E.T. SCHOLTEN, P.A. DE JONG, J.M. KUHNIGK, C. SCHAEFER-PROKOP, Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images, *Medical Image Analysis,* 2014*,* **18**(2), 374–384.

12. JENA, S.R., T. GEORGE, N. PONRAJ, Featur*e* extraction and classification techniques for the detection of lung cancer*:* A detailed survey, Paper presented at the *International Conference on Computer Communication and Informatics (ICCCI)*, Jan. 23–25, 2019, Coimbatore, India, http://dx.doi.org/10.1109/ICCCI.2019.8822164.

13. JOON, P., A. JATAIN, S.B. BHASKAR, Lung cancer detection using image processing techniques, *International Journal of Engineering Science,* 2017, **7**(4), 6497–6450.

14. KAHAN, S.A., S. HUSSAIN, S. YANG, K. LQBA, Efective and reliable framework for lung nodules detection from CT scan images, *Scientific Reports*, 2019, 9(4989), 1–14, https://doi.org/10.1038/s41598-019-41510-9.

15. KATRE, P.R., A. THAKARE, Detection of lung cancer stages using image processing and data classification techniques, Paper presented at the *2nd International Conference for Convergence in Technology (I2CT)*, Mumbai, 2017, 402–404.

16. KAUR, A.R., Feature extraction and principal component analysis for lung cancer detection in CT scan images, *International Journal of Advanced Research in Computer Science and Software Engineering*, 2013*,* **3**(3), 187–190.

17. KIRK, S., Y. LEE, P. KUMAR, J. FILIPPINI, B. ALBERTINA, M. WATSON, J. LEMMERMAN, Radiology data from the cancer genome atlas lung squamous cell carcinoma [tcga-lusc] collection, *The Cancer Imaging Archive*, 2016a. doi: 10.7937/K9/TCIA.2016. TYGKKFMQ.

18. KNIGHT, S.B., P.A. CROSBIE, H. BALATA, J. CHUDZIAK, T. HUSSELL, C. DIVE, Progress and prospects of early detection in lung cancer, *Open Biology*, 2017*, 7*(9), 170070. doi:10.1098/rsob.170070

19. KUMAR, G., P.K. BHATIA, A detailed review of feature extraction in image processing system*s*, Paper presented at the *Fourth International Conference on Advanced Computing & Communication Technologies*, 2014, 5–12, DOI 10.1109/ACCT.2014.74.

20. LYRA, M., A. PLOUSSI, A. GEORGANTZOGLOU, P. IONESCU, MATLAB as a tool in nuclear medicine image processing, *MATLAB - a Ubiquitous Tool for the Practical Engineer,* 2011*,* 477–500. doi: 10.5772/19999.

21. MENDOZA, F., R. LU, Basics of image analysis, in: *Hyperspectral Imaging Technology in Food and Agriculture*, B. Park, R. Lu, eds., Springer, 2015, pp. 9–56, DOI 10.1007/978-1-4939-2836-1_2.

22. MIAH, M., M.A. YOUSUF, Detection of lung cancer from CT image using images processing and neural network*,* Paper presented at the *International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, Savar, Bangladesh, 2015, 1–6.

23. NADERNEJAD, E., S. SHARIFZADEH, H. HASSANPOUR, Edge detection techniques: Evaluations and comparisons, *Applied Mathematical Sciences, 2008, **2**(31), 1507–1520.

24. NADKARNI, N.S., S. BORKAR, Detection of lung cancer in CT images using image processing, Paper presented at the *3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, 2019, place?, https://doi.org/10.1109/ICOEI.2019.8862577.

25. NAGAR, S., *Introduction to MATLAB for Engineers and Scientists: Solutions for Numerical Computation and Modeling,* Apress, New York, pp. 1–217.

26. NAIDICH, D.P., A.A. BANKIER, H. MACMAHON, C.M. SCHAEFER-PROKOP, M. PISTOLESI, J.M. GOO, J.H. AUSTIN, Recommendations for the management of subsolid pulmonary nodules detected at ct: A statement from the Fleischner society, *Radiology, 2013,* **266**(1), 304–317

27. PARVEEN, S.S., C. KAVITHA, A review on computer-aided detection and diagnosis of lung cancer nodules, *International Journal of Computers & Technology,* 2012**, 3**(3a), 393–400.

28. PRIYA, S.S., B. RAMAMURTHY, Lung cancer detection using image processing techniques, *Research Journal of Pharmacy and Technology,* 2018*, **11**(5), 2045–2049.

29. SHRIVAKSHAN, G., C. CHANDRASEKAR, A comparison of various edge detection techniques used in image processing, *International Journal of Computer Science Issues (IJCSI),* 2012*, **9**(5), 269–278.

30. TARVER, T., Cancer facts & figures, *Journal of Consumer Health on the Internet,* American Cancer Society, 2012**, 16**(3), 366-367. doi:10.1080/15398285.2012.701177.
31. TERAMOTO, A., H. FUJITA, O. YAMAM URO, T. TAMAKI, Automated detection of pulmonary nodules in PET/CT images: Ensemble false-positive reduction using a convolutional neural network technique, *Medical Physics,* 2016*, **43**(6), 2821–2827.
32. THAKUR, S., M. SHARMA, A Review: An effective approach for lung cancer detection using various algorithm, *International Journal of Computer Science Trends and Technology* (IJCST), 2016, **4**(5), 140–142.
33. THIRUMARAN, J., S. SHYLAJA, Medical image processing – an introduction, *International Journal of Science and Research (IJSR)*, 2015, **4**(11), 1197–1199.
34. TIWARI, A.K., Prediction of lung cancer using image processing techniques: A review, *Advanced Computational Intelligence: An International Journal (ACII)*, 2016, **3**(1), 1–9.
35. TOPRAK, A., I. GÜLER, Suppression of impulse noise in medical images with the use of the fuzzy adaptive median filter, *Journal of Medical Systems,* 2006*, **30**(6), 465–471.
36. TUN, K.M.M., A.S. KHAING, Feature extraction and classification of lung cancer nodule using image processing techniques, *Int. J. Eng. Res. Technol.(IJERT)*, 2014*, **3**(3), 2278–0181.
37. ***BING.COM/images, Lung cancer CT scan, https://www.bing.com/images/search?q=Lung+ Cancer+CT+Scan&form=IRBPRS&first=1&tsc Image Basic Hover, Retrieved March 23, 2021.