International Journal of Advances in Electrical Engineering

E-ISSN: 2708-4582
P-ISSN: 2708-4574
IJAEE 2024; 5(2): 67-73

www.electricaltechjournal.com
Received: 10-07-2024
Accepted: 16-08-2024

Muhammad D Hassan
Collage of Health and Medical Techniques, Northern Technical University, Kirkuk, Iraq

Suad Kakil Ahmed
Collage of Health and Medical Techniques, Northern Technical University, Kirkuk, Iraq

Correspondence
Muhammad D Hassan
Collage of Health and Medical Techniques, Northern Technical University, Kirkuk, Iraq

Statistic and morphological image processing for automated lung cancer identification

Muhammad D Hassan and Suad Kakil Ahmed

DOI: https://doi.org/10.22271/27084574.2024.v5.i2a.74

Abstract
Among the population of Jordan, lung cancer is the second most usually diagnosed form of the disease. Over the last several years, there has been an increasing amount of data that demonstrates that early identification of lung cancer may allow for more rapid treatment action. These pieces of evidence have been accumulated over the course of the last few years. Screening programs for lung cancer have been established all over the globe as a result of this, which has supplied the motivation for their implementation. For the goal of finding regions of lung cancer, a computer-aided detection (CAD) system is presented within the scope of this investigation. The objective of this system is to make use of computed tomography (CT) pictures. It is now being deployed as a "second reader" in order to assist radiologists in concentrating their attention on areas that may be overlooked during visual interpretation. This is being done in order to help radiologists concentrate their attention on relevant regions. First, the CT images are thresholder, then the sections that are created are labeled, and finally, certain diagnostic characteristics of each location are obtained for further analysis and interpretation. It is via this process that segmentation is achieved. The proposed computer-aided design (CAD) system is comprised of these three major processes, which are listed and described below. The study is taught, evaluated, and verified with the use of photographs gathered from forty-five different patients. The findings that were acquired are an exact match for the diagnosis that was made by the radiologist throughout the process of identifying the defective regions and quantitatively measuring their size, position, and boundaries, in addition to exhibiting their other diagnostic features. Additionally, the suggested approach is able to identify locations that have been incorrectly categorized.

Keywords: Labeling, lung cancer, thresholding, computer tomography

Introduction

Cancer, which is the second most frequent cause of death in Jordan, is a substantial contribution to morbidity and mortality among the Jordanian population. Heart problems are the primary cause of death, leaving cancer as the second most common cause of death [1].

There are around 12.7% of all cancer cases that have been diagnosed in Iraq that are lung cancer. Lung cancer accounts for 86% of all cases of lung cancer in men, whereas only 14% of cases are found in females. When compared to other kinds of cancer, lung cancer is the most prevalent among men [2, 3].

One of the tumors that is believed to be one of the most challenging to treat is lung cancer. When lung cancer is detected in its early stages, it might provide an opportunity for medical therapy that can help reduce the risk of the disease. The computed tomography (CT) scan is a typical method that gives information about three-dimensional (3D) cross sectional images of the lung. This approach is helpful in identifying the condition that is being experienced. Radiologists are need to devote a significant amount of time to the interpretation of CT scans when dealing with a big collection of patient pictures. It may be possible to employ computer-aided detection (CAD) technologies in order to lessen the amount of labor that radiologists have to do. Due to the fact that computer-aided diagnosis (CAD) systems may be deployed as a "second reader," they can also assist in enhancing the performance of radiologists in the identification of primary lung cancer. It is anticipated that this will assist radiologists in concentrating their attention on areas that may contain lung cancer [4, 5].

The automated detection of lung nodules in CT scans has been developed by a number of different researchers who have built a variety of distinct approaches. These researchers have contributed to the development of this technique. Numerous research groups were inspired to develop identification strategies in order to enable the rapid detection of illnesses based on image analysis techniques as a result of improvements in image reconstruction techniques and the availability of digital modalities. This was done in order to facilitate the rapid detection of diseases. In order to make the process of identifying diseases more expedient, this project was set up. Standard image analysis procedures, on the other hand, either make use of a manual tracing technique or a semi-automatic approach. Both of these approaches are distinct from one another in terms of the manner in which the radiologist would intervene. "Tracing" techniques are the name given to each of these methods of investigation. As a result of the need for automated image analysis methods and the advancements that have been made in the computer industry, the researchers were inspired to suggest more advanced image analysis approaches in order to address the issue of automation [6, 7].

A wide range of labs and organizations have collaborated to develop methods for the computer-assisted segmentation of CT images of the lungs. It has been shown that these methodological approaches are successful [8, 9]. This was done in order to get a better understanding of the lungs. This was done in order to improve accuracy. However, manual procedures are not only arduous but also vulnerable to fluctuations that may occur both between observers and within observers. During the process of image segmentation, it is possible to be guided. This contrast may be seen in the picture. Identifying the borders that divide the left and right lungs was the objective of the two-dimensional edge tracking technique that was used in order to accomplish this aim. There have been other people who have used three-dimensional area development by making use of seed sites that were well delineated [10]. A significant amount of human engagement is necessary in many semi-automatic systems in order to pick threshold values or change the segmentation that is consequently produced. When the volume averaging has a tendency to diminish the contrast at the border of the lung, the anterior and posterior junction lines are manually provided in the research that Kalender *et al*. conducted in order to discriminate between the right and left lungs. This is done in instances when the contrast is present. Brown *et al*. also proposed a knowledge-based and automated approach for the segmentation of chest CT images. This technique was published here. Inside the framework of their approach, low-level image processing methods are directed by anatomical information that is stored inside a semantic network. On the other hand, Brown *et al*. used dynamic programming in order to search for the junction lines in an automated manner, in contrast to Kalender *et al*., who required the input of humans in order to define the anterior junction lines. For the purpose of locating the junction lines, this was carried out.

For the approaches that have been provided, there are several restrictions that prevent them from being directly applicable for detection. However, the fundamental purpose of these articles is to provide a technique for the segmentation of the airway tree as well as the segmentation of the lung tissue and the separation of it from other regions in the surrounding area. Although there are some restrictions to the scope of work that these articles cover, their primary objective is to give a method for the segmentation both of these things. In the process of diagnosing or identifying problems in the lung tissues, they do not take into consideration any application for the segmentation of the lungs and airways. These documents have the potential to serve as a foundation for the work that is being suggested.

The identification of lung nodules in helical CT images was made possible by Armato and colleagues via the development of an automated approach. In order to make use of the volumetric image data obtained from a CT scan, this method integrates both two-dimensional and three-dimensional analysis. They sought to use the lung segmentation procedure in order to identify lung nodules; however, the identification of even the smallest of lung nodules is not adequate to result in a diagnosis that is significant. In addition to this, the radiologist is required to be familiar with the features of each nodule that is detected, which are not provided in this research.

Satoh and his colleagues developed a computer-aided detection (CAD) system with the intention of identifying possible nodule candidates at an early stage, commencing with the present, and making use of early helical CT scanning of the thorax. This procedure was carried out in order to achieve the aforementioned objective. However, their work did not take into account the category of discovered nodules or the display of any features that would facilitate their identification. That being said, the primary emphasis of their research was on the development of a system that is capable of identifying very small nodules [11].

Conducted a performance evaluation of an automated classifier. This evaluation was carried out with the intention of achieving the aforementioned objective. In this study, Armato and colleagues introduced a completely automated computational technique for the identification of lung nodules in CT images. The purpose of this method was to discover lung malignancies that could be overlooked during visual interpretation. The detection of lung nodules is accomplished using a three-step procedure, beginning with two-dimensional processing, then moving on to three-dimensional analysis, and finally utilizing an automatic classifier. A technique of detecting nodules in CT scans was suggested and put into practice by Awi and colleagues.

A computer-aided design (CAD) program that is capable of interpreting a CT data set and automatically recognizing pulmonary nodules on the photographs is the subject of further study that is now being conducted. In the case of very small nodules, it is anticipated that the approach will function more effectively. Small nodules, which may be overlooked during visual interpretation, are often the primary focus of the computer-aided design (CAD) systems that have been created. Those works provide a computer-aided design (CAD) system that does not take into account the many aspects that will assist the radiologist in reducing the amount of work he or she has to do and speeding up the diagnosing process. Not only do radiologists demand a clear list of data that describes the internal morphology, size, location, and border shape of the nodule, but they also want two-dimensional and three-dimensional measurements of the nodule that has been recognized when it is found.

A computerized system that has the benefit of recognizing lung nodules and showing its statistical characteristics, in addition to detecting abnormal areas in the lung tissue and finding its diagnostic measures, will be the emphasis of the current suggested CAD system in this study. This system will also be able to locate diagnostic measurements. It is noteworthy that this system will be developed with the intention of assisting the radiologist in arriving at an accurate diagnosis. A higher level of diagnostic confidence in lung cancer is anticipated to result from the use of the suggested method.

The Methodology

The creation of databases and the validation of diagnoses: This research project's validation database included sixty individuals who were diagnosed with thoracic conditions. A computed tomography (CT) scan is performed on these individuals at King Hussein Cancer Center between the dates of April 19, 2018, and May 20, 2024. All of these patients were male, with forty of them being male and five of them being female. It was determined that a cancer was present after a bronchoscopy and histological examination were carried out. At the time of the exams, the patients' ages ranged from 49 to 69 years old, with 59 being the average age. The range of ages that were present was around 49 to 69 years. Every single patient was given a unique number in order to ensure that their privacy was maintained at all times. A training data set is the initial set of data, while the sample data set is the second set of data. The data is segregated into two sets. It includes forty patients, which accounts for 79% of the total data. It is used in the process of training the classifier to ensure accurate modeling. It is the validation data set that makes up the second set. Two percent, or nine patients, make up this group. The suggested classifier's performance is shown with the help of the validation data.

To carry out the CT tests, a CT scanner (model number CT-S5VA47A; manufactured by Siemens Medical Systems in Germany) was used. The size of the image matrix for each reconstructed CT segment was 1024 × 1024 pixels. There are a total of 3,150 photos included inside the database, which is comprised of 45 scan sets. The number of images contained within each set ranges from 52 to 94, with 72 being the average number of images contained within each set.

On two separate occasions, the visual diagnosis of the CT pictures was carried out. The initial diagnostic was performed by the team from the King Hussein Cancer Center while the scanning procedure was being carried out. This was done in an attempt to obtain the best possible degree of efficiency and accuracy in the system that was created. In a later period of time, Jordan Hospital was the location where the second diagnosis was confirmed.

The development and implementation of an automated system for the diagnosis and classification of primary lung cancer

One of the aspects of the computer-aided design (CAD) system that was built is that it can perform the role of a second reader for the CT scan images. The CT DICOM pictures are converted from 18 bits to 10 bits stored Bitmap images using this system, which is responsible for the conversion. In order to distinguish the lung into its component parts, a threshold labeling method is carried out. A method called the Otsu thresholding algorithm serves as the foundation for the thresholding technique. This algorithm is intended to lessen the amount of variance that is present within the group. This step involves applying Run-Length encoding labeling to the binary picture to identify all discernible locations inside the lung tissue. In summary, a set of the region's border. All these features are ascertained by applying the histogram characteristics to the detected locations.

The radiologist is able to recognize, detect, and categorize the nature of the detected area by comparing and recognizing the features of the region. This allows the radiologist to determine whether or not the detected region is a normal region. An illustration of the block diagram of the algorithm for automated detection and classification may be seen in Figure 1.

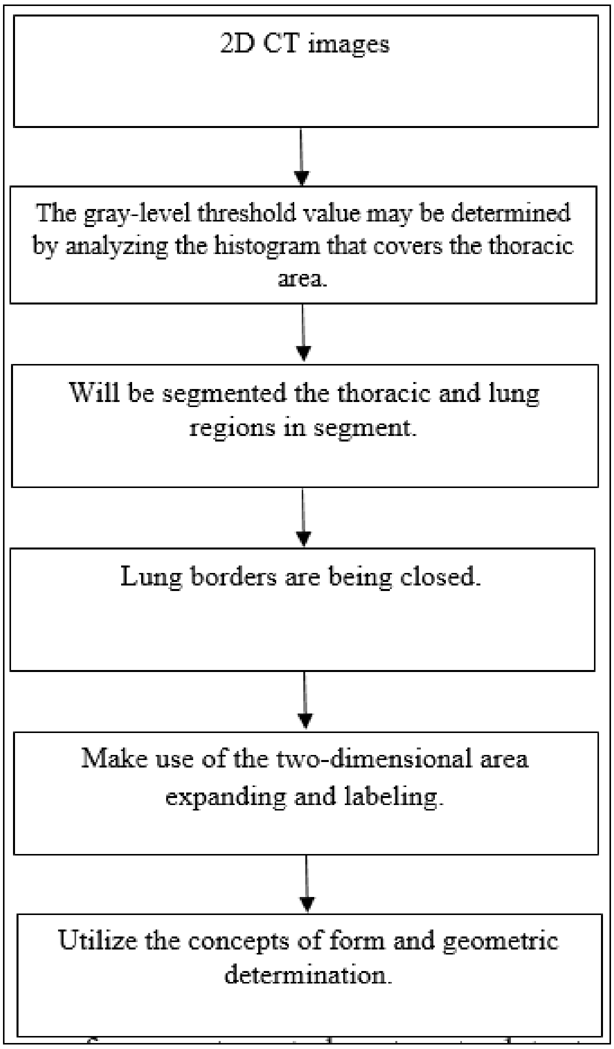

Fig 1: The overarching scheme for an automated system to detect and classify primary lung cancer inside CT images is shown in the flow diagram.

Labelling at the threshold

A direct approach of delineating the object's regions involves executing a threshold labeling operation, whereby a binary value of 1 is allocated to each pixel exhibiting a sufficiently elevated value. When a pixel is designated the binary value 1, it signifies the possibility that it is part of an object of interest. Any pixel lacking a sufficiently high value is given a binary value of 0. Having this designation indicates that there is a potential that it is a component. The variance for the subset of values less than or equal to t is defined by the threshold t, whereas the variance for the subset of values beyond t is also defined by the same threshold. Otsu proposes that the best threshold is the threshold that minimizes the weighted sum of group variances. This is the definition by which the best threshold is defined. The weights represent the probability that are associated with each of the groupings. Following this, the optimal threshold t may be identified by doing a straightforward sequential search over all of the possible values of t in order to identify the threshold (t) that minimizes the value of $\sigma_{\omega}^{2}(t)$. It is possible to limit this to a search between the two modes in many different scenarios. Nevertheless, the identification of the modes is, in all actuality, the same thing as the identification of the values that separate the modes respectively.

Following the creation of the binary image by the Otsu threshold, the lung tissue is manually selected in order to exclude and ignore the regions that are located around the lung tissue and are outside of the region of interest. This is done in order to get the desired result. This is done in order to get the desired results. A binary-1 pixel in the lung area is chosen for the selection, and the selection is based on a 4-adjacency relation. The system takes into account any binary-1 pixels that are related to the pixel that is being chosen, but it ignores any other pixels. A representation of the binary picture that includes the parts that need processing is shown in Figure 2.

Marking of components that are connected

The labeling procedure for linked components is responsible for carrying out the unit transition from pixel to region levels. Each pixel along a path with a binary value of 1 is given the same identifying label with the same value. This label is assigned to every pixel with a binary-1 value and is connected to all other pixels along the route. An identifying label is assigned to a potential item area. This label serves to identify that specific area. An example of a grouping process is the labeling of connected components. It is the Run-length encoding of the binary picture that serves as the foundation for the labeling operation that is used in this approach. This approach employs a Run-length encoding of the input picture prior to resolving the equivalence classes and thereafter re-labeling the runs according to the identified equivalence classes. This occurs prior to the resolution of the equivalence classes. Subsequently, they proceed to analyze these runs, allocate initial labels, and document label correspondences in a local equivalence database. At a certain juncture, they reclassify the runs based on the determined equivalence classes.

Extrapolation of features

The computation of global attributes for each area that is created by the linked components labeling algorithm is what the analysis phase entails. During the process of interpreting the CT scan, the radiologist takes into account a wide range of various characteristics in order to determine whether or not the patient has primary lung cancer. These parameters contain the dimensions of each region, which include its width and length, radius, the morphological feature of its border, and, last but not least, its spatial histogram. There are many more factors that are included in these parameters. The following is a list of characteristics that the radiologist makes an effort to take into account. The radiologist will have the ability to simply make a judgment about the lesion that is being investigated since these measures are made accessible for each location.

The system that is being presented is capable of automatically computing all of the attributes that are required by the radiologist in order to categorize the areas that have been classified from a CT picture. When the CT film is used, it is difficult for the radiologist to view the shape of the area border clearly. This approach, however, demonstrates the shape of the area border. Taking advantage of this benefit is something that this system does. Additionally, this approach provides the radiologist with the ability to explore the homogeneity of the region in issue, in addition to presenting numerical figures for the length, breadth, radius, and area of the region. The radiologist will be able to arrive at a more accurate opinion and diagnosis of the areas that have been identified with the use of this information. In order to achieve this goal, it is necessary to reduce the inaccuracies that are brought about by manually measuring the attributes of the area, as well as the errors that are brought about by fluctuations that occur both between and within the observer.

For the purpose of evaluating the effectiveness of the proposed computer-aided design (CAD) system, 750 photographs that were taken from the validation set are used. In all, there are 3,001 nodules in the photos. In order to establish the specificity, accuracy, precision, and recall of the system, a computation is performed using the formulas that are shown below:

Particularity = TN/(FP+TN) (1)

Accuracy = (TN + TP)/(FP+TN+FN+TP) (2)

Exactness = TP/(FP + TP) (3)

Recall = TP/(FN+TN) (4)

In this context, TP signifies true positives, representing areas identified by the proposed system and classified as positive by both hospitals; FP denotes false positives, indicating areas detected by the proposed system but classified as negative by both hospitals; FN refers to false negatives, representing areas overlooked by the proposed system yet classified as positive by both hospitals; TN indicates true negatives, referring to areas missed by the proposed system and classified as negative by both hospitals. The specialists from both hospitals are tasked with calculating the proportions of true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) in this research. The outcome is deemed relevant to performance metrics when both hospitals achieve agreement.

The outcomes

The graphic clearly demonstrates that the run-length algorithm successfully identified and localized all viable locations without any losses. While the technique that has been presented has the potential to be useful in localizing sick locations, it also has the capability of enabling the feature extraction algorithm to emphasize the impact that these areas have on the patient's medical record.

There is a major indication that permits the radiologist to execute his or her diagnostic, and that indication is the internal morphology of the site that was found. It is conceivable that the existence of necrosis, cavitations, or calcification is indicated by regions that include heterogeneous areas. It is necessary to utilize the spatial histogram of the CAD system in order to conduct an analysis of the degree of homogeneity that is present within each particular detected area in order to establish that the suggested system is capable of exhibiting the behavior of the specified tissue. An illustration of the various characteristics that are associated with the region that is shown by a red star in figure 3 is provided. The purpose of this demonstration is to show that the system is capable of recognizing both tiny and big nodules, and to demonstrate that it can detect both types of nodules.

Table 1: Specified values for the performance characteristics of the system that is being proposed

Quantified and forecasted	Positive	Negative
Positive	382(TP)	5(FN)
Negative	17(FP)	378(TN)

Table 2: Measures of performance indicators

The Indicator of Performance	Percentage Value (%)
Exactness	96.77
Recall	98.89
Accuracy	98.00
Particularity	97.06

When it comes to the 53 patients, radiologists from both King Hussein Center and Jordan Hospital are now analyzing the lesions that are creating worry. This is done in order to confirm the findings acquired from the automated system as well as the results the radiologists got. The findings that were acquired for 73% of the patients are utilized as the training set, while the remaining 27%, which consists of nine patients, are used as the validation set. Table 1 presents the performance parameters that are associated with the system that is being suggested. The suggested systems' performance indicators are shown in Table 2, which may be found here.

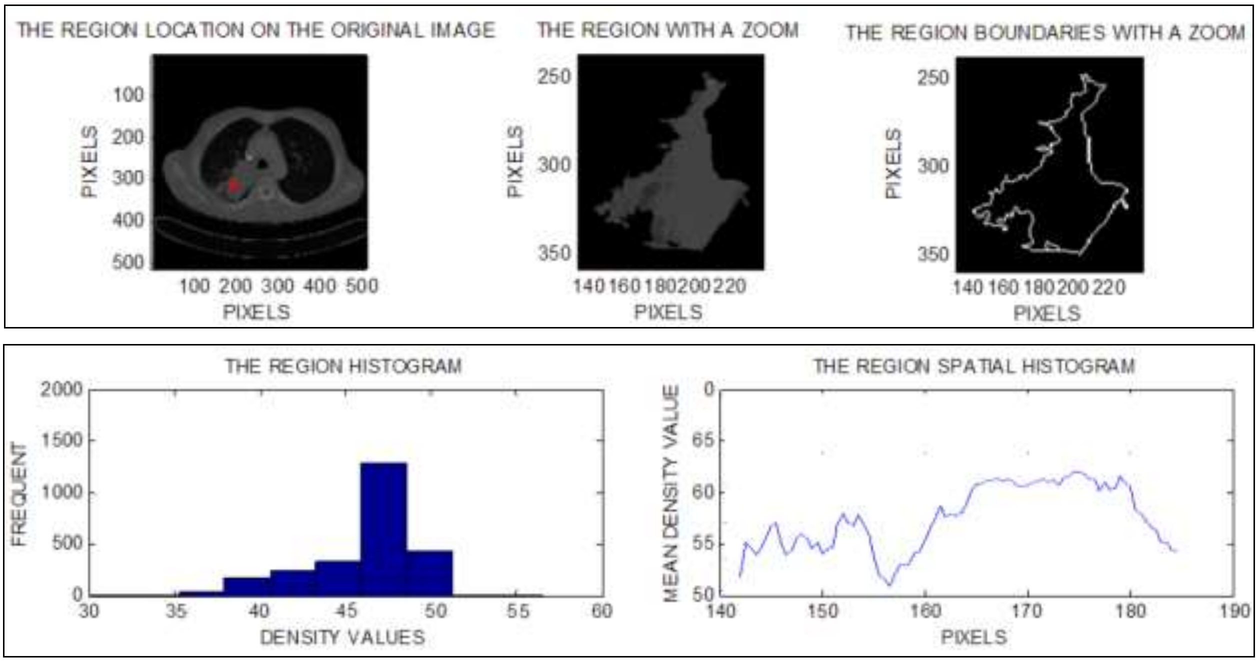

Fig 3: The top second half of the picture is represented by the characteristics of the red star

Discussion

Performing a comparison study with other works that are comparable to this one that are accessible in the literature allows for the discussion and highlighting of the strengths and shortcomings of this work. This research presents a system that autonomously establishes the threshold value for each selected CT image using the Otsu algorithm, effectively identifying the pixels corresponding to the lungs, in comparison to the work proposed by Shiying *et al*. This is conducted to facilitate a comparison between the two systems. A manually closure of the lung boundary is performed before the application of the 3D linked components labeling. In this method, the areas that are placed at the border of the lung and extend into the region of the mediastinum are closed off in order to achieve the desired effect. This is one of the limits that there is a possibility of encountering when it comes to the functioning of the system that was advised. In the event that the area is manually closed, it is possible that pixels from the mediastinum region will be included. This might result in an inconsistent decline in the accuracy of the system's estimated measurements occurring. However, when manually closing is done to the lung borders in a cautious manner, it will reduce the number of mistakes that may be created. Furthermore, the suggested approach takes into account the whole of the picture, eliminating the need to divide the lung into left and right sections, as Shiying *et al*. have pointed out. A few human selection and closure procedures are included in the segmentation operation that is being presented in this work. It is important to highlight that this segmentation operation is not a completely automated segmentation. When it comes to distinguishing lung pixels during the thresholding operation, the segmentation operation that is being discussed in this study is not only adequate but also exhibits a high degree of accuracy. This is in contrast to the findings that Shiying and his colleagues acquired. Through meticulous closure of the border areas, it is possible to achieve a labeling efficiency of one hundred percent for the regions that are contained inside the lung.

Armato and his colleagues developed a system capable of automatically identifying lung nodules. Based on the multiple-gray-level thresholding method, ten percent of the fifty nodules in the database were excluded from the first set of nodule candidates. Ten percent of the nodules were included into the database. The inefficiency of the extraction method was the cause. Conversely, as seen in Table 2, the proposed methodologies consider all nodules, achieving excellent levels of accuracy, precision, specificity, and recall percentages.

Additionally, the classification findings that were shown in other study that was examined are restricted to the data that was supplied to support their detection. There are many different forms, locations, morphologies, pathological representations, and boundaries that make cancer regions difficult to construct. Because of this, it is not possible to ensure that the categorization systems that have been provided are accurate and robust. Nodule margins and internal morphology are two essential characteristics that the radiologist must understand to accurately finalize their diagnosis. These parameters supplement the location and size of the nodule, which are the two primary factors concerning the nodule. Furthermore, cancer nodules that have been found are not restricted by size and/or location. In the study presented here, the computer-aided design (CAD) system that has been provided is able to supply the radiologist with a multitude of useful diagnostic characteristics that enable him or her to draw conclusions about the discovered region. In addition, the system indicates the location of the area, as well as its size, edges, and interior shape. The technique determines not only the presence of tiny nodules but also the presence of masses and lesions in addition to detecting them.

In order to assist the radiologist in locating extremely tiny defective areas, the computer-aided design (CAD) system that is being developed is anticipated to give a tool that will provide assistance. When the radiologist has this information, it will be much simpler for them to determine whether or not the region that has been found is impacted by cancer. Furthermore, it will be used to assess the rate of change in tumor development. The technology will be capable of detecting even subtle changes in size and will provide a more accurate measurement of the diagnosed lesion, regardless of the direction of the change. This is primarily because identifying the direction of change may be very challenging by manual selection or visual inspection. Furthermore, there exists the possibility of an unmonitored rate of change that cannot be detected by human monitoring. This is a possibility. The CAD system is able to determine the size of the discovered nodules in each CT scan that is performed on a patient over the course of a period of time. The radiologist is able to determine the effectiveness of any changes made to the scans in the future by comparing the values that were measured with the values that were measured in the past. It is possible for the system to measure variations in values that are smaller than 0.85 millimeters. During the early phases of detection, the system will be an invaluable instrument for determining the pace of progress in the treatment process. Additionally, it will be of great use to the medical practitioner in selecting the appropriate drug.

Due to the fact that the system was able to identify extremely tiny nodules that the radiologist would not be able to readily perceive, it may be employed as a tool to investigate the emergence of a nodule as well as the progression of the illness over the course of time. The most important aspect of this study is this contribution, which indicates that the suggested system is capable of carrying out the duties that have been outlined.

Conclusions

It is suggested in this research that an intelligent and automated system for the identification and categorization of nodules be developed. The computer-aided design (CAD) system can interpret DICOM CT images and use modern image processing techniques to enhance the segmentation and detection of large lesions and tiny cancerous lesions. This leads to a more precise cancer diagnosis. This is likely to occur while the lesions are in the early physiological stages and may potentially be addressed by surgical intervention. The technology may also disclose the dimensions of the discovered malignancy, serving as a substitute for the manual measuring procedure conducted by the radiologist to assess the breadth (transverse) and length (anterior-posterior) of the cancerous region. The radiologist has a significant obstacle when it comes to the detection of cancer area margins during the diagnostic process. The margin is one of the elements that is used during the diagnostic process to indicate the development of the tumor. During the course of this investigation, the system that is being studied shows each identified area border in order to facilitate the process of detecting region boundaries that are sensitive to fluctuations in observer. In late stages of cancer, there is a chance that low-contrast lesions may show calcification, necrosis, and cavitation. Radiologists encounter challenges and potential sources of inaccuracy when interpreting areas with less contrast. The proposed approach may provide a histogram, enabling the radiologist to assess the homogeneity of the cancer-affected region.

The computer-aided design (CAD) system that is being constructed will play a significant part in identifying minute variations in the size of the nodules that have been identified. This will provide the radiologist with the capacity to analyze and quantify the growth factor of the tumor over a period of time. If the treatment is administered at an earlier stage in the progression of the illness, it will improve the effectiveness of the treatment and make it easier for the physician to choose the appropriate drug.

References

1. Manikandan T, Bharathi N. Lung cancer detection using fuzzy auto-seed cluster means morphological segmentation and SVM classifier. J Med Syst. 2016;40(1):1-9. DOI: 10.1007/s10916-016-0624-1.
2. Wang S, Yang DM, Rong R, Zhan X, Fujimoto J, Liu H, *et al*. Artificial intelligence in lung cancer pathology image analysis. Cancers (Basel). 2019;11(11):1673. DOI: 10.3390/cancers11111673.
3. Yu KH, Zhang C, Berry GJ, Altman RB, Ré C, Rubin DL, *et al*. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat Commun. 2016;7:12474. DOI: 10.1038/ncomms12474.
4. Chaudhary A, Singh SS. Lung cancer detection on CT images by using image processing. In: 2012 International Conference on Computing Sciences; 2012 Sep 18-19; Roorkee, India. IEEE; c2012. p. 142-146. DOI: 10.1109/ICCS.2012.61.
5. Krishnamurthy S, Narasimhan G, Rengasamy U. Three-dimensional lung nodule segmentation and shape variance analysis to detect lung cancer with reduced false positives. Proc Inst Mech Eng H. 2016;230(1):58-70. DOI: 10.1177/0954411914568506.
6. Zhang G, Jiang S, Yang Z, Gong L, Ma X, Zhou Z, *et al*. Automatic nodule detection for lung cancer in CT images: A review. Comput Biol Med. 2018;103:287-300. DOI: 10.1016/j.compbiomed.2018.11.001.
7. Singh GAP, Gupta PK. Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans. Neural Comput

Appl. 2019;31(10):6863-6877.
DOI: 10.1007/s00542-018-4051-x.

8. Hunter LA, Krafft S, Stingo F, Choi H, Martel MK, Kry SF, *et al*. High quality machine-robust image features: Identification in non-small cell lung cancer computed tomography images. Med Phys. 2013;40(12):121916. DOI: 10.1118/1.4839675.
9. Song J, Liu Z, Zhong W, Huang Y, Ma Z, Dong D, *et al*. Non-small cell lung cancer: Quantitative phenotypic analysis of CT images as a potential marker of prognosis. Sci. Rep. 2016;6:38282.
DOI: 10.1038/srep38282.
10. Lavanya M, Kannan PM, Arivalagan M. Lung cancer diagnosis and staging using firefly algorithm fuzzy C-means segmentation and support vector machine classification of lung nodules. Int J Biomed Eng. Technol. 2021;37(2):185-200.
DOI: 10.1504/IJBET.2021.115144.
11. Binczyk F, Prazuch W, Bozek P, Polanska J. Radiomics and artificial intelligence in lung cancer screening. Transl Lung Cancer Res. 2021;10(2):1186-1198.
DOI: 10.21037/tlcr-21-280.