Industrial Engineering Journal
ISSN: 0970-2555
Volume : 50, Issue 7, No. 1, July 2021

THE INVOLVEMENT OF MORPHOLOGICAL PROCESSING IN IMAGE PROCESSING FOR THE EXACT AND EARLY STAGE IDENTIFICATION OF LUNG CANCER USING MEDIAN FILTERS

P Anandhi, Dept of Electronics and Communication Engineering, Sree Venkateswara College Of Engineering, Nellore (Dt), Andhra Pradesh, India.
P.Ravi Kumar, Dept of Electronics and Communication Engineering, Sree Venkateswara College Of Engineering, Nellore (Dt), Andhra Pradesh, India.
N.Kesav Kumar,Dept of Electronics and Communication Engineering, Sree Venkateswara College Of Engineering, Nellore (Dt), Andhra Pradesh, India.
R. Prapulla Kumar, Dept of Computer Science and Engineering, Sree Venkateswara College Of Engineering, Nellore (Dt), Andhra Pradesh, India.

ABSTRACT
Many people lose their lives each year to cancer, one of the most deadly and most prevalent diseases. Lung cancer is the most common and has the greatest fatality rate of among all major forms of cancer. Computed tomography scans are employed to detect lung cancer since they give a clear view of the tumour within the body and enable the tracking of the tumor's development. Despite the fact that CT is favoured over other imaging modalities, visual interpretation of these CT scan pictures may be laborious, risky, and delay the discovery of lung cancer. As a result, various medical professions use image processing methods to find lung cancers early. A computerized system was used in this investigation to identify lung cancer in CT scan picture is conferred. The method for diagnosing lung cancer is proposed to partition the lung region that is relevant using mathematical morphological procedures. For picture pre-processing, this method makes use of techniques like median filtering. Support vector machines are employed to classify CT scan images into normal and abnormal groups by computing geometric attributes from the retrieved region of interest.

1. INTRODUCTION
The mortality rate from cancer is increasing daily, making it a significant worldwide public health problem. The most common and fatal kind of illness that can affect both sexes is cancer of the lungs. the development of cancerous nodules (malignant lung tumours) as a result of unregulated cell proliferation in the lung tissues is referred to as lung cancer, also known as carcinoma. Smoking and tobacco use are the two primary causes of malignant lung nodules. Only 14% of lung cancer patients overall survive the condition after all stages, over a time span of 5–6 years. categories of lung cancer according to cell features are small cell lung cancer and non-small cell lung cancer. Between 15 and 20% of cancer cases are small cell lung cancer, whereas between 80 and 85% of cases are non-small cell lung cancer. By presenting patients access to appropriate quick therapy, early detection of lung cancer can consequently raise the survival prospects up to 60–70% and lower the fatality rate.Generally speaking, lung cancer is divided into four phases according to the severity: Stage I lung cancer is localized, Stage II and III chest cancers are localised, and Stage IV lung cancer is dispersed by chest to another body areas. Lung cancer may be identified using a variety of imaging methods, including Positron Emission Tomography (PET), Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and Chest X-rays. CT scan images are typically chosen over images from other types of imaging since they're more trustworthy, crisper, and less distorted. Visual interpretation of datasets is a difficult, drawn-out, and very individual process. This elevates the possibility of human error and boosts the possibility that cancer could have been classified incorrectly. Because of this, having a computerised approach is essential for helping the

radiologist diagnose lung cancer accurately. The approach developed for the system includes dataset collection, pre-processing, lung segmentation, feature extraction, and classification.

2. LITERATURE SURVEY

One of the most important illnesses to affect humans in recorded history is lung cancer. By 2030, 10,000,000 people will have died from lung cancer globally, according to the World Health Organisation. The advanced Non-Small Cell Lung Cancer (NSCLC) 5-year endurance rate [1] is still dismally poor. The improvement in diagnostic/prognostic or predictive accuracy that may be achieved by quantitative image feature analysis is hypothesized to have an effect on a large number of patients [2]. Clinical computed tomography (CT) scans presented in conformity by industry standards were employed in the current investigation to extract image features. The first and most important stage is to precisely outline the lung tumours as to minimise variability for feature extraction. For the best radiation oncology, lung tumours must be accurately delineated. Radiation oncologists or radiologists often manually outline the tumor's perimeter while analysing CT images for tumours. Manual segmentation typically underestimates the lesion's volume [3] and is highly variable [4,5] in order to ensure that the entire lesion is detected. A consistent, precise segmentation is crucial because image characteristics (including texture and shape-related variables) are sensitive to minuscule tumour border alterations. Therefore, a highly automatic, accurate, and repeatable lung tumour separation method would constitute an important step forward. The accurate identification of soft tissue lesions from a specific modality like CT, PET, or MRI is of utmost importance for computer-aided diagnosis (CAD), computer-aided surgery, radiation therapy planning, and medical research. The properties of tumours are often extremely comparable to those of the adjacent normal tissue, noise from the image capture procedure, and the great heterogeneity of cancer tumours. The intensity-based or morphological methodologies used in conventional medical picture segmentation techniques [6–9] can occasionally provide inaccurate tumour segmentation.Definiens AG [12] and Merck & Co., Inc. collaborated to create the lung tumour analysis (LuTA) tool [10] that is a part of the Definiens Cognition Network Technology [11]. It is a prototype application that illustrates how organs and tumours may be automatically and partially identified in CT scans. We suggest a novel delineation technique based on employing several seed locations with region expanding with the goal of solving the aforementioned shortcomings of the "Click & Grow" algorithm [13]. By employing an initial seed point to create a region where numerous seed points are automatically generated, the new method makes advantage of the properties of the original approach. Multi modal segmentation can be created using the several areas that were grown. Recently, many medical image applications have relied heavily on ensemble segmentation, meaning the term for an assortment of several input segmentation (multiple executions employing the same segmentation method but with various initialization that are combined to generate consensus segmentation) [14–15]. In this study, we demonstrate how such a strategy minimizes variability between observers with far fewer operators interactions than the initial approach.

3. EXISTING SYSTEM

Currently, doctors use their visual interpretation skills to identify lung cancer in CT images. Graphical interpretation of datasets is a difficult, drawn-out, and very individual process. Due to this, there is a greater chance that someone will make an oversight and that cancer will be categorized incorrectly.
The earlier methods included studying pictures from magnetic resonance imaging, computerized tomography scans, and mammography. Professionalism helps doctors diagnose the condition and establish the cancer's phases. In addition to some surgical techniques, chemotherapy, radiation, and targeted therapy are all used in the treatment of cancer. This analysis takes a very long time, costs a lot

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 50, Issue 7, No. 1, July 2021

of money, and hurts a section of the body.

4. PROPOSED SYSTEM

Median filters are employed in the suggested technique to detect lung cancer in CT images for improved enhancement. Lung cancer is accurately classified using Support Vector Machine (SVM). Figure illustrates the suggested approach for CT image-based lung cancer diagnosis.

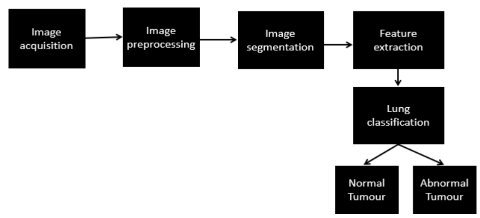

Fig1: schematic of the projected system

SOFTWARE

MATLAB was the primary piece of software we utilised to construct this project. MATRIX LABORATORY is the abbreviation for the software. A high performance language for technical computing is called MATLAB. A sophisticated matrix and matrix operation support has been built into the data analysis and visualisation software MATLAB. Additionally, MATLAB has a strong built-in programming language and excellent graphics capabilities. The utilisation of collections of MATLAB programmes created to serve certain tasks has contributed to MATLAB's rise to prominence as a tool. These collections of software are referred to as toolboxes, and the image processing toolbox in particular is the one that interests us.

5. METHODOLOGY

Data Collection

The initial step is to obtain lung CT images of cancer patients. The images were collected from the Cancer Imaging Archive database for research reasons. The images are stored in DICOM format. The image database contains computerised tomography images of patients both with and without lung cancer.

Image Pre-Processing

The goal of the preprocessing step of an image is to reduce any undesired distortions already existing in the picture and to highlight any features that will be helpful during the processing stage. picture enhancement and picture smoothing are two of the primary phases. To get rid of any unwelcome noise that is present in the image, smoothing is used. Since salt and pepper noise frequently appears in CT scan images, median filtering has been established to be a very effectual method for removing this impulse noise while maintaining the edges. The greatest consequence for image smoothing come from median filtering since it reduces noise without distorting the picture.

Digital photos are enhanced using image processing techniques to generate better results for processing later. Since artefacts brought on by contrast variations in the image have an impact on image quality, contrast adjustment is done to improve the image. Brightness modification increases the contrast between pixels in an image by altering input pixel values so that 1% of the data is saturated at both high and low input image levels of brightness by default.

Image Segmentation

Segmentation is the process of removing the necessary region of interest from the picture. The

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 50, Issue 7, No. 1, July 2021

acquisition of the lung area from binary pictures may be done effectively using mathematical morphological techniques. In our procedure, binary pictures were initially created from the preprocessed grayscale photos. To remove unnecessary components from the binary picture, a morphological opening operation was carried out using a disc structuring element. After that, the opened image was enhanced and given a distinct border operation. The holes and crevices in the lungs were filled to create the lung masks. In order to offer us the segmented tumour region, a final exclusive OR operation is done to the lung mask output and clear border output.

Feature Extraction

The most crucial stage in turning input data into necessary features is feature extraction. In this step, significant features from the segmented region of interest are extracted, and these features are used as input for categorising CT scan images. By removing three geometrical elements, it is possible to determine the size and form of the lung tumour. The characteristics of the malignant lung nodule include its size, shape, and eccentricity.

Area: This scalar value represents the total number of pixels that the malignant lung nodule has gained. The total of the pixel areas in the binary picture that are qualified by value 1 is used to calculate the area from the binary image.

Perimeter: This scalar value indicates the total number of pixels that are present at the lung tumor's boundary. By adding the pixels that were detected with the value 1 at the outline of the lung nodule, the perimeter of the binary picture is calculated..

Eccentricity: Other names for this statistic are irregularity index (I), circularity, and roundness.

Eccentricity is equivalent to 1 for a circular form and less than 1 for every other shape..

$$\text{Eccentricity} = \text{length of major axis}/\text{lengthofminoraxis}$$

Classification

The CT scan pictures are classified as normal or abnormal during the classification step. In our approach, CT scans will be utilised to diagnose lung cancer using the SVM algorithm. SVM categorize are supervised learning models that examine input data and categorise it based on pattern. The SVM classifier divides the training dataset into two groups and creates a model utilising it. The SVM algorithm then divides the testing dataset into two classes, assigning fresh instances to each class. Thus, the SVM classifier sorts the lung CT images into two groups by locating the ideal hyper plane.

ADVANTAGES

1. Better cancer nodules identification accuracy than the best available model.
2. determines if the lung cancer found is aggressive or benign.
3. Removes salt and pepper noises from images which causes false detection of cancer.

APPLICATIONS

The medical area is where image processing is mostly used to identify lung cancer in CT scans. Considering the fact that CT is preferred above various types of imaging, it might be difficult to interpret these CT scan images visually. As a result, the medical industry commonly uses methods of image processing to spot lung cancers in its earliest stages.

6. EXPERIMENTAL RESULTS

MATLAB software was used in the suggested system's implementation. The Cancer Imaging Archive (TCIA) provided the database for this investigation. The lung cancer patient's CT scan picture and associated histogram are shown in Fig. 2. An image's histogram is a visual representation of the picture that shows the distribution of pixels among the various grey levels.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 50, Issue 7, No. 1, July 2021

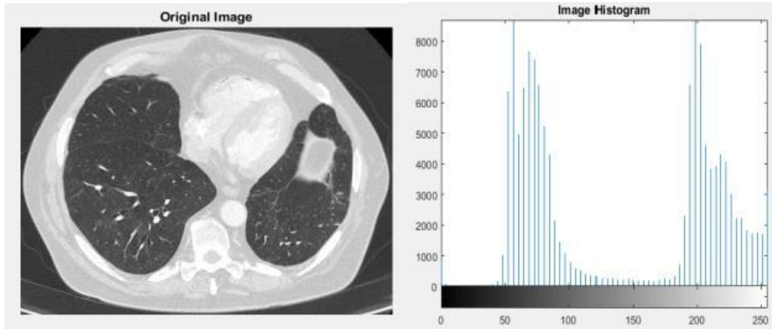

Fig. 2. CT scan of a malignant lungs and its histogram

The lung CT scans were then subjected to a 3*3 median filtering technique to remove salt and pepper noise while maintaining the boundaries of the image.

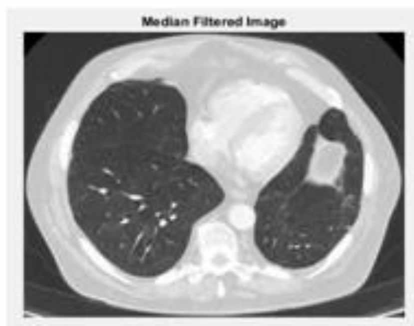

Fig.3.Median filtering method for image smoothing

Employing brightness modification, picture augmentation was carried out later in the preparation stage.

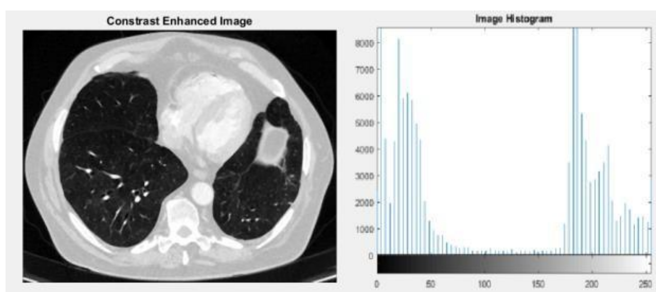

Fig. 4. Opposition enhanced image and histogram of image

As illustrated in Fig., the lung masks and tumour area have been created based on the CT picture. 5.

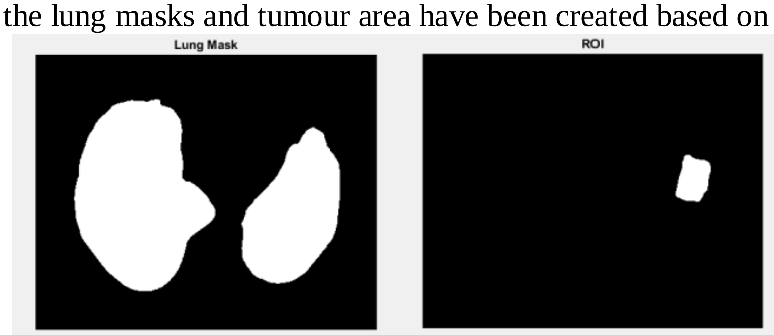

Fig.5.mask of the lungs and the removed tumour area

From the fragmented tumour region, three geometrical features—area, the perimeter, and weirdness—

were then retrieved.

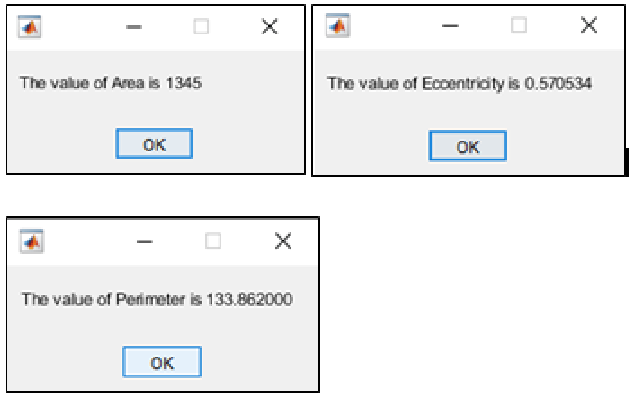

Fig. 6. Message boxes with features

The SVM classification algorithm will be used in the final step to assess when the CT scan pictures are cancerous or non cancerous.

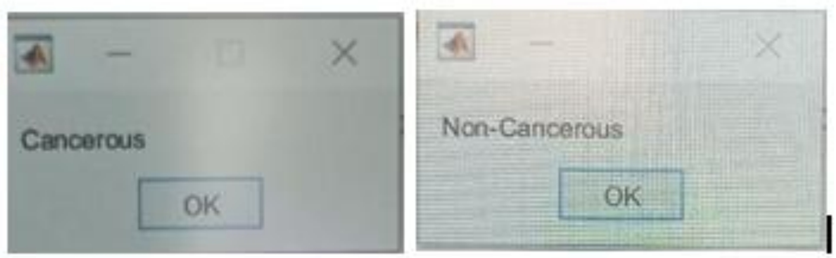

Fig.7.Message boxes with final result

CONCLUSION

With the help of an image processing technique, this work effectively created a system for the automatic detection of lung cancer in CT images. The selected method works well for improving, segmenting, and collecting data from CT images. Utilizing the median filtering method, the erratic noise in the shots was effectively removed without harming the images. The lung and tumour regions can be properly segmented utilizing mathematical morphological techniques. Area, perimeter, and eccentricity—three geometrical properties—were retrieved from the segmented tumour region and used as inputs by the classifier to separate healthy from pathological lung CT scans. Thus, the proposed method enables the accurate and early stage detection of lung cancer.

FUTURESCOPE

Future improvements will concentrate on kinds of part-solid tumours, which produce unstable tumour borders when given various start seed sites. Another issue that has to be solved is the issue of inadequate tumour segmentation with large tumours.

REFERENCES

1. D.H. Johnson, W.J. Blot, D.P. Carbone ,“ Cancer of the lung: Non-small cell lung cancer and small celllungcancer”,in:
2. M.D.Abeloff,J.O.Armitage,J.E.Niederbuber,M.B.Kastan,W.G.McKenna(Eds.),Abeloff’sclinicalon cology, Churchill Livingstone/Elsevier,Philadelphia,2008.

3. R.J.Gillies,A.R.Anderson,R.A.Gatenby,D.L.Morse,"Thebiologyunderlyingmolecularimaginginoncology: from genome to anatome andbackagain,"ClinicalRadiology65(2010)517–521.
4. J. Rexilius, H.K. Hahn, M. Schluter, H.Bourquain, H.O. Peitgen, " Evaluation of accuracy in MSlesionvolumetry usingrealisticlesionphantoms,"AcademicRadiology12 (2005)17–24.
5. P. Tai, J. Van Dyk, E. Yu, J. Battista, L.Stitt,T.Coad,"Variabilityoftargetvolume delineation in cervical esophagealcancer,"InternationalJournalofRadiationOncology,Biology,Physics42(1998)277–288.
6. J.S.Cooper,S.K.Mukherji,A.Y.Toledano, C. Beldon, I.M. Schmalfuss, R.Amdur,S.Sailer,L.A.Loevner,P.Kousouboris,K.K.Ang,J.Cormack,J.
7. Sicks, "An evaluation of the variability oftumor-shape definition derived by experi-encedobserversfromCTimagesofsupraglottic carcinomas (ACRIN protocol6658)," International Journal of RadiationOncology,Biology,Physics67(2007)
8. 972–975.
9. S.Hojjatoleslami,J.Kittler,"Regiongrowing:anewapproach,"IEEETransac-tions on Image Processing 7 (1998) 1079–1084.
10. J.Dehmeshki,H.Amin,M.Valdivieso,X.Ye, " Segmentation of pulmonary nodulesin thoracic CT scans: a region growingapproach," IEEE Transactions on MedicalImaging27(2008) 467–480.
11. J.Dijkers,C.VanWijk,F.Vos,J.Florie,
12. Y. Nio, H. Venema, R. Truyen, L. vanVliet,"SegmentationandsizemeasurementofpolypsinCTcolonography, Medical Image ComputingandComputer-AssistedIntervention"–MICCAI2005 (2005) 712–719.
13. A.B.LeLu,M.Wolf,J.Liang,M.Salganicoff,D.Comaniciu,"AccuratepolypsegmentationforDCTcolongraphyusingmulti-stagedprobabilisticbinarylearningandcompositionalmodel,"in:IEEEConferenceonComputerVisionandPatternRecognition, 2008, pp. 1–8.
14. C. Bendtsen, M. Kietzmann, R. Korn, P.Mozley, G. Schmidt, G. Binnig, "X-raycomputedtomography:semiautomatedvolumetricanalysisoflate-stagelungtumorsasabasisforresponseassessments,"InternationalJournalofBiomedicalImaging(2011).(2011)Article ID 361589.
15. M.Athelogou,G.Schmidt,A.Sch¨ape,M.Baatz,G.Binnig,"Cognitionnetworktechnology–anovelmultimodalimageanalysistechniqueforautomaticidentificationandquantificationof
16. biologicalimagecontents,"ImagingCellularandMolecularBiologicalFunctions(2007)407–422. DefiniensA.G.,/http://www.definiens.comS.
17. Y.Gu, V.Kumar, L.O.Hall, D.B.Goldgof,R.Korn,C.Bendtsen,R.A.Gatenby,R.J.Gillies, "Automated Delineation of LungTumorsfromCTImages:MethodandEvaluation,"in:WorldMolecularImagingCongress, San Diego,CA,USA,2011, pp.373.
18. J. Huo, K. Okada, W. Pope, M. Brown,"Sampling-based ensemble segmentationagainstinter-operatorvariability,"in:Proc.SPIE, 2011, pp. 796315.
19. J.Huo,E.M.vanRikxoort,K.Okada,H.J.Kim,W.Pope,J.Goldin,M.Brown,"Confidence-basedensembleforGBMbraintumorsegmentation,"in:Proc.SPIE,2011, pp. 79622P.