Data Analytics and Artificial Intelligence
Vol: 2(5), 2022
REST Publisher; ISBN: 978-81-948459-4-2
Website: http://restpublisher.com/book-series/data-analytics-and-artificial-intelligence/

A Study of Pulmonary Carcinoma Using Image Processing

Moulieswaran E,Kalpana G
SRM Institute of Science and Technology, Kattankulathur, Chennai, TamilNadu, India.
*Corresponding author Email:emoulieswaran@gmail.com

Abstract. In the previous few decades, one of the best reliable imaging techniques has been Computed Tomography (CT). To diagnose pulmonary carcinoma, early or later through CT scan images, many image processing has been used. Pre-processing, Segmentation, Extraction, filter, Classification are the few techniques that are usually used in detecting the initial stage of pulmonary carcinoma using image processing. To get CT images from the LIDC (Lung image database consortium) dataset such as NIH/NCI. DICOM (Digital imaging in communication in medicine) is a protocol on behalf of transmission among image devices. CT scan images may lead to error and delay the finding of pulmonary carcinoma. The purpose of sharing digital medical images is to easily access data. Many viewers maintain the image format (DICOM) and read large numbers of data.Keywords: CT (Computed Tomography), Pre-Processing, Enhancement, Segmentation, Extraction.

1. Introduction

The image processing technique is majorly used to enhance unprocessed image capture from various cameras or scans from various regions. It is shown that the image processing is very useful for the finding of tumor cells [1]. The data can be retrieved easily using image processing. A technique based on regular parameters using image processing for the investigation of lung cancer is reported [1]. These techniques are used usually in the army, clinical, investigational areas etc[2]. Image processing is divided into several stages, which include image pre-processing, Denoising, optimization, and edge detection, among others [3].Cancer is currently the most dangerous disease that has confronted human life and resulted in death [4]. Pulmonary carcinoma is defined by independent cell grows in the lung [5]. Pulmonary carcinoma is another name for Lung cancer. Common indications are cough (which includes blood whoop); lose weight, mild and severe chest pain. Pulmonary carcinoma has two types they are NSCLC and SCLC [6]. In men, pulmonary carcinoma, the most common identified cancer, and the maximum pulmonary carcinoma is in Eastern Europe, Central and Eastern Asia [7]. Globally, around 19.3 million cases and in 2020 nearly 10 million expire. Female breast cancer has surpassed pulmonary carcinoma in terms of prevalence, with 2.3 million cases (11.7 percent), followed by lung, prostate, colorectal, and stomach cancers (11.4 percent, 7.3 percent, 10.0 percent, and 5.6 percent). Pulmonary Carcinoma is the cause of cancer deaths, accounting for 1.8 million deaths (18% of all cancer deaths), followed by colorectal, stomach, liver, and female breast cancers (9.4 percent, 7.7 percent, 8.3 percent, 6.9 percent) [8].

2. Methodology

Here are some image processing methodologies: image collection, pre-processing, segmentation, extraction, classification, and the use of various filtering techniques and results. Collecting Image: Collect CT scan images for research from the LIDC dataset such as NIH/NCI. The advantage of the CT image is improved clearness, less noise, and distortion. A standard protocol, DICOM is used for transmission among imaging devices. These images were from CT scans where CT images have lower noise. 512 * 512 pixels are the image size. Pulmonary carcinoma Data set:

1. 1. UCI Machine Learning Repository https://archive.ics.uci.edu/ml/datasets/lung+cancer
2. Public Access to the Cancer Imaging Archive (TCIA) https: //wiki. cancerimagingarchive.net/display/Public/LIDC-IDRI-Kaggle
   https: //www. kaggle. com/datasets/mohamedhanyyy/chest-ctscan-images
   https://www.kaggle.com/datasets/fanbyprinciple/luna-lung-cancer-dataset.

Image pre-processing: The process involves improving the data by removing unwanted distortions or improving specific images for subsequent processing. It is used to enhance the images, like lines, boundaries, and textures. To remove noise from the image, different filtering techniques have been used based on the noise in the image. Pre-processing techniques for noise reduction are Gaussian, Salt and Paper, Poisson, impulse, and Speckle. These are the filtering techniques such as Denoisingare Spatial Domain (Mean or Wiener, Non-Linear or Median) and Transform Domain. Image Segmentation : The strategy for labelling every pixel of an image through segmentation is comparable, and certain visual features can be shared. There are two approaches: the Marker-Controlled Watershed Segmentation Approach and the Threshold Approach.It is divided into different subcategories called image segments. It aids in reducing the difficulty of an image by making image processing or analysis much simpler. Feature Extraction: It is an important step in the dimensionality reduction procedure; an initial set of image data is fragmented and reduced into manageable sets. This method can be broken down into two general steps: feature selection and classification. Shape measurements are physical dimensional measurements that characterize an image's appearance. Only the following characteristics can be extracted: area, perimeter, eccentricity, contrast, correlation, energy, and homogeneity. Classification: This categorizes detects the nodule as benign or malignant. Here are the few classifiers that have been used for pulmonary carcinoma are SVM, CNN, ANN, KNN. Use Various Filtering Techniques: Filter techniques are used to improve and adjust digital images. It is used for blurring, sharpening, edge detection and noise reduction. Filter techniques are mostly used to eliminate high frequencies (flattening techniques) and low frequencies (image enhancement, edge detection).

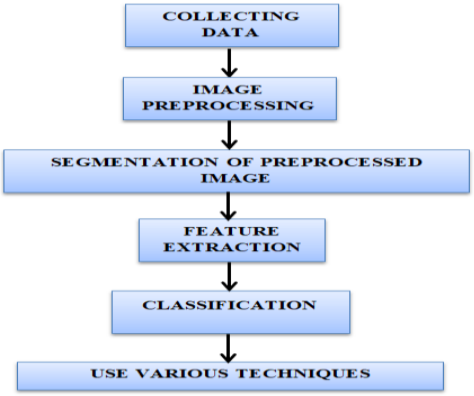

Figure 1 Lung Cancer Detection Techniques

3. Applications

Some image processing applications which are applicable these days: Agricultural- Image processing techniques are better suited to the agricultural domain. This produces more seeds in agricultural areas for harvest management. This approach cannot be manually implemented to perform various tasks. Image processing techniques are used to sort and grade fresh fruits. Harvest Classification is a fuzzy logic analysis technique [9]. Communication- It has several operations such as telecommunication, video conferencing used with the support of image processing techniques. Image processing techniques are used to identify the features during video conferencing. For analysis of vocal interaction image processing, coding is performed. Recognition of Character- Manually written, and printed files can be easily identified using image processing applications. Medical-Image processing techniques are used in clinical areas to improve performance and recognize body parts. Image processing techniques are used in the skull and chest. This is also used for the study, diagnosis, identification, and treatment of disease or illness. Medical Imaging focuses mainly on the identification and unveiling of internal structures that are covered by the skin and bones. This method is mainly useful for specialists conducting laparoscopic surgery to show the inside sections without actually opening the body. It uses magnetic resonance imaging (MRIs) and CT scans.

TABLE 1. Comparisons of medical imaging [10]

Types	X-Ray	Ultrasound	MRI	CT-Scan
Resolution	2000 * 2500	The typical required resolu-tion for medical ultra-sound is 1 mm	typical resolutions of around 1.5×1.5×4 mm.	0.5 mm along the x to y axis, 0.5–0.625 mm along the z axis
Continuous Speed	Slow	No	Fast	Average
Charge	Less	Moderate	Very High	Less
Statistics Acquisition	Low	Low	High	High

Special effects	high emis-sion	Non-Ionizing	No Hazards reported	Small Risk of irradia-tion
Accessib-ility	Maxi-mum	Maxi-mum	Less than CT	Without much trouble
Exposure to Radiation	Low	High	None	Moderate
Time Taken	15 Mins	30 Mins to 1 Hour	30 to 45 Minutes	5 Minutes

Specialists can look at the mysterious or unknown third dimension of the body in this way. When using a CT scanner, fragments inside can be seen, and diseased parts can be easily detected and found. When anticipating the MRI, it gets a signal from the body's magnetic elements converts its magnetic field and covers scanned data into internal organ images with the aid of software [10]. Commercial -Image processing techniques are used in a variety of operations such as bar code, signature, and stocking. With the help of the image processing approach, signatures and barcodes can be recognized confidentially.

4. Related Work

Shanthi et al [11]proposed an automated pulmonary carcinoma prediction and detection system based on multiple levels of data collection, extraction, selection, classification, and identification. The Cancer Genome Atlas database was used to collect CT images. The primary features are offset specification and statistical measure extraction via the Gray - level Co-occurrence matrix (GLCM). The texture features of the lung regions of interest (ROI) will be used to identify the object. A Gabor filter is a linear filter with a Gaussian function increased by it. The two-dimensional Gabor filters were provided by

$$G_c[i,j] = Be\ i^2 + j^2/2a^2\ \ cos(\ 2\pi\ f(\ i\ cos\theta + j\ sin\ \theta\ ))$$

$$G_s[i,j] = Ce\ i^2 + j^2/2a^2\ \sin(2\pi\ f(\ i\ cos\theta + j\ sin\ \theta\ ))$$

The SDS method is used to evaluate subsets in a much more effective way. In the initial phases, each machine will be tasked with combining a portion. This procedure will be repeated several times to determine the current status. SDS techniques can handle large processing, Naive Bayes classifier, which implies the parameters that contribute to classification are mutually connected. NN was much more efficient for finding pulmonary carcinoma for helping oncologists. WasudeoRahane et al [12] proposed a technique such as Image acquisition, Noise reduction, GrayScale Conversion, Binarization, Segmentation and Support Vector machine that has been used in the proposed work. This research paper is to identify pulmonary carcinoma using median filter and segmentation to get accurate results.Ashwini et al [13]proposed an image processing technique was used to identify and estimate pulmonary carcinoma cells. The multi SVM machine classifier has been used to determine whether the nodule was malignant or benign. Image processing has been used to improve, remove noise, segment, and extract images. MATLAB is used for all of these techniques. With an accuracy of 97% the multi SVM classifier identified whether the cell was normal or abnormal. Anjali Sevani et al [14] proposed an image processing technique used to enhance the image using a median filtering technique, Dilation, Erosion, opening and closing, feature extraction. This research paper is to use different methodologies to find the different stages of pulmonary carcinoma easily.

Results:

A. Median Filter

The median filter makes use of the entire 3x3-neighborhood: 18, 65, 80, 75, 88, 55, 110, 158, 230

B. Opening

The scientific description of the opening is as follows:

I o B = (I-B) + B

C. Closing

The mathematical description of the closing is as follows:

I•B = (I+B) – B

5. Summary of Related Work with Accuracy

TABLE 2. Image Processing Mechanism

References	Image Processing Mechanism	Classifier mechanism	Accuracy
[15]	Watershed Segmentation	SVM	92%
[19]	Grey Level Co-Occurrence Matrix, clustering, and K-means	Multi SVM	97%
[18]	WMHE	Deep-LearningWith Instant-aneouslyTr	98.42%

		ained Neural-Networks 'DITNN'	
[12]	Grey Level Co-Occurrence Matrix GaborFilter	Decision Tree 'DT', Naïve Bayes 'NB', Neural Network 'NN'	87.41%, 88.52%, 89.63 %.

Makaju et al [15]The watershed segmentation technique for recognizing and SVM for categorizing small firm lumps as malignant or benign were used in this research paper to differentiate the malignant nodules from the CT scan. The proposed approach identifies cancer with a consistency of 92 percent, which is higher than traditional model's consistency of 86.6 percent. When compared to traditional work, most development is based on proposed work. Meanwhile, this model did not categorize the tumor into different stage of the disease.Shakeel et al [16] proposed to identify and predict the pulmonary carcinoma via and DITNN. Firstly, CT images from a public dataset, namely the 'CIA' Cancer Imaging Archive. Secondly, the WMHE improves the image pre-processing technique for improving the quality of the CT image. Thirdly, improved CT images were segmented and sent to IPCT.Fourthly, extract a pixel similar value to extract from the segmented region. As a result, with 98.42 percent accuracy and a false positive rate of 0.038, classifier techniques are much more efficient and accurate in predicting benign tumors in CT images.PrathameshGawade and R.P. Chauhan [17]the methodology used in this research paper is to diagnose at initial and critical phases using intelligence, computational techniques, and various distortion removals via edge detection techniques and algorithms. Different image processing techniques have been used, such as the input image, GrayScale image, few filters (like high pass, median), threshold segmentation, watershed algorithm, morphological operations. The purpose is to analyze images at the lowest level. To improve the pixel intensity and then, convert discrete to digital image, segments into pixels, then reconstruct the image with better quality.Amir Roointan et al [18] proposed to identify the possibility factors of pulmonary carcinoma in early part they widely used biosensor technology. For early diagnosis and healing of lung nodules, this technology has a major role. This approach is expected to be diagnosed using bio sensing to identify risks that have remedial and medical implications in the future. Various bio sensing features have been reviewed briefly. These features were used to identify pulmonary carcinoma disease biomarkers. Moritz Schwyzer et al [19]. Machine learning was used in this study to detect lung nodules in a specified situation of ultraviolet PET scans using fluorodeoxyglucose positron emission tomography (FDG-PET) imaging. Lung nodules are detected by the artificial nervous system. The levels of responsiveness were 95.9 % and 91.1 %, respectively. The artificial neural network achieved accuracy of 98.1 percent with a regular dose of 98.1 percent and 94.2 percent with an ultra-flow dose of 3.3 percent. As an added benefit, the system increases the precision of testing lung tumors.H. Guo et al [20]the Knowledge-based Analysis of Mortality Prediction Network (KAMP-Net) is a prediction method that has been proposed. This method improves prediction by extracting clinical knowledge and features exposed by a dual stream network (DSN). When compared to other methods, the KAMP - Net can perform and achieve more. Low-dose CT (LDCT) is more accurate than X-rays in detecting pulmonary carcinoma and reduces death rates. Another approach, such as a novel approach using CNN and SVM to find the results, has been used.M. B. Khumancha et al [21]The two datasets used in this research paper are data1 and data2. The LIDC/IDRI for detecting nodules in Data1 is accessed. Kaggle is used to access Data2 for cancer detection. The CNN module is used to detect nodules in the datasets, and the second procedure detects malignant symptoms in the dataset. The accuracy is 90.78 percent based on the true count, while the precision is calculated to be 89.24 percent.Guobin Zhang et al [22]this study used a computer-aided design (CAD) approach to detect lung nodules automatically. The CT description is being used in the current development. It was studied and discovered to produce better results in terms of high accuracy and low positive predictive value. CAD systems used a variety of techniques.

6. Summary of Related Work

TABLE 3.Techniques

References	Techniques	Future Enhancement
[23]	1. Image Segmentation 2. Feature Extraction	Try the different filtering techniques like Weiner filter which removes the noise that has a corrupted signal.
[13]	1. Image acquisition 2. Noise reduction 3. GrayScale Conversion 4. Binarization 5. Segmentation	Identify the disease which helps in reducing the abnormal cells.
[24]	1. Fusion Method 2. Image Segmentation	Evaluate large and different datasets to achieve greater

		robustness.
[22]	1. Input Image 2. GrayScale image 3. High pass filter 4. Median filter 5. Threshold segmentation 6. Watershed algorithm 7. Morphological operations	Other types of cancer, such as breast cancer and skin cancer, can also be treated with it. It's also used in medical studies.
[17]	1. Pre-processing 2. Enhancement 3. Median filter 4. Binary image processing 5. Dilation 6. Erosion 7. Opening 8. Closing 9. Feature extraction	Other types of cancer, such as brain tumors, skin cancer, stomach cancer, and breast cancer, can also be treated using the proposed methods.

NidhiS.Nadkarni and Prof.SangamBorkar[23] proposed a technique to detect pulmonary carcinoma using CT image and image segmenting and feature extraction to find cancer. Different techniques and methodology are used in Data Collection,Pre-processing, Segmentation, Extraction, and classification. In this paper, median filter, mathematical morphological operation, and geometrical features are used. Zheng.S et al [24]CNN and MIP (Maximum intensity projection) images were used to investigate a pulmonary nodule discovery. Improved pulmonary nodules in MRI assessment with CT scans to increase the likelihood of detecting nodules. MIP image thickness was high. Due to that, accuracy level has been increased in detection of lung nodules.

7. Conclusion

This indicates that the method of detecting this disease is very important and necessary in avoiding serious phases and reducing its global measurement. There are only a few algorithms that have been improved (for example, Gabor filter, Watershed Segmentation, Grey Level Co- Occurrence Matrix, Weighted Mean Histogram Equalization). This paper covers a system utilized for pulmonary carcinoma deductions using CT images, PET and MRI. In most of the image processing techniques, image segmentation and feature extraction have been used to get better segmentation.

References

[1]. A. K. Tiwari, "PREDICTION OF LUNG CANCER USING IMAGE PROCESSING," *Advanced Computational Intelligence: An International Journal,* vol. 3, no. 1, pp. 1-9, 2016.

[2]. Shaik Naseera, G.K. Rajini, B. Venkateswarlu, Jasmin Pemeena Priyadarisini M., " A Review on Image Processing Applications in Medical Field," *Research J. Pharm. and Tech,* vol. 10, pp. 3556-3560, 2017.

[3]. A. Asuntha, A. Brindha, S. Indirani, and A. Srinivasan, "Lung cancer detection using SVM algorithm and optimization techniques," *Journal of Chemical and Pharmaceutical,* vol. 9, pp. 3198-3203, 2016.

[4]. S. S. Priya and B. Ramamurthy, "Lung cancer detection using image processing," *Research Journal of Pharmacy Technology,* vol. 11, pp. 2045-2049, 2018.

[5]. Robert L. Keith , MD, Lung Cancer (Lung Carcinoma), Division of Pulmonary Sciences and Critical Care Medicine, University of Colorado School of Medicine, July 2020.

[6]. "Non-Small Cell Lung Cancer Treatment (PDQ®)–Patient Version," 27 August 2021.

[7]. "World Cancer Report: Cancer Research for Cancer Prevention," ISBN 978-92-832-0448-0, 2020.

[8]. Hyuna Sung, Jacques Ferlay, Rebecca L.Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, Freddie Bray, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *A Cancer Journal for Clinicians Global Cancer Statistics,* vol. 71, no. 3, pp. 209-249, 04 February 2021.

[9]. Ranu Gorai, "A Survey of Digital Image Processing," *International Journal of Research in Engineering Technology and Science,* vol. 6, 2016.

[10]. N. Goel, A. Yadav, and B. M. Singh, "Medical image processing: A review," in *2016 Second International Innovative Applications of Computational Intelligence on Power, Energy and Controls with their Impact on Humanity (CIPECH)*, Ghaziabad, 2016.

[11]. S. Shanthi and N. Rajkumar, "Lung Cancer Prediction Using Stochastic Diffusion Search (SDS) Based Feature Selection and Machine Learning Methods," *Neural Processing Letters,* pp. 1-14, 2020.

[12]. Wasudeo Rahane, Himali Dalvi, Yamini Magar, Anjali Kalane, "Lung Cancer detection using image processing and machine learning Health care," in *Proceeding of 2018 IEEE International Conference on Current Trends toward Converging Technologies*, Coimbatore, India., 2018.

[13]. P. Ashwini, S. Antony, and V. Kanchana, "Categorizing the stages of lung cancer using Multi SVM Classifier," *International Journal of Research in Pharmaceutical Sciences,* vol. 10, no. 3, pp. 2323-2328, 2019.

[14]. Anjali Sevani, Hardik Modi Sagar Patel, Himanshu Patel, "Implementation of Image Processing Techniques for Identifying Different Stages of Lung Cancer," *International Journal of Applied Engineering Research,* vol. 13, no. 0973-4562, pp. 6493-6499, 2018.

[15]. S. Makaju, P. Prasad, A. Alsadoon, A. Singh, and A. Elchouemi, "Lung cancer detection using ct scan images," *Procedia Computer Science,* vol. 125, pp. 107-114, 2018.

[16]. P. M. Shakeel, M. Burhanuddin, and M. I. Desa, "Lung cancer detection from CT image using improved profuse clustering and deep learning instantaneously trained neural," *Measurement,* vol. 145, pp. 702-712, 2019.

[17]. Prathamesh Gawade and R.P. Chauhan, "Detection of lung cancer using image processing techniques," *International Journal of Advanced Technology and Engineering Exploration,* vol. 3(25), no. 2394-5443, NOVEMBER-2016.

[18]. Amir Roointan, Tanveer Ahmad Mir, Shadil Ibrahim Wani, Mati-ur-Rehman, Khalil Khadim Hussain, Bilal Ahmed, Shugufta Abrahim, Amir Savardashtaki, Ghazaal Gandomani, Molood Gandomani, Raja Chinnappan, Mahmood H Akhtar, "Early detection of lung cancer biomarkers through biosensor technology: A review," *Journal of Pharmaceutical and Biomedical Analysis,* vol. 164, pp. 93-103, 2019.

[19]. Schwyzer M, Ferraro DA, Muehlematter UJ, Curioni-Fontecedro A, Huellner MW, von Schulthess GK, Kaufmann PA, Burger IA, Messerli M., "Automated detection of lung cancer at ultralow dose PET/CT by deep neural networks," *Lung Cancer,* Dec 2018.

[20]. H. Guo, U. Kruger, G. Wang, M. K. Kalra and P. Yan, "Knowledge-Based Analysis for Mortality Prediction From CT Images," *IEEE Journal of Biomedical and Health Informatics,* vol. 24, no. 2, pp. 457-464, Feb 2020.

[21]. Khumancha, M.B., Barai, A. & Rao, "Lung Cancer Detection from Computed Tomography (CT) Scans using Convolutional Neural Network," in *IEEE10th International Conference on Computing, Communication and Networking Technologies* , 2019.

[22]. Guobin Zhang, Shan Jiang, Zhiyong Yang, Li Gong, Xiaodong Ma, Zeyang Zhou, Chao Bao, Qi Liu, "Automatic nodule detection for lung cancer in CT images: A review," *Computers in Biology and Medicine,* vol. 103, pp. 287-300, 2018.

[23]. Nidhi S.Nadkarni, Prof. Sangam Borkar, "Detection of lung cancer in CT images using image processing," in *Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2019)*, 2019.

[24]. Zheng S, Guo J, Cui X, Veldhuis RNJ, Oudkerk M, van Ooijen PMA, "Automatic Pulmonary Nodule Detection in CT Scans Using Convolutional Neural Networks Based on Maximum Intensity Projection," *IEEE Trans Med Imaging,* pp. 707-805, 2020.

