SciencePG
Science Publishing Group

Detection of Lung Cancer Using Digital Image Processing Techniques: A Comparative Study

Munimanda Prem Chander[1], M. Venkateshwara Rao[2], T. V. Rajinikanth[3]

[1]Department of Computer Science and Engineering GIT, GITAM UNIVERSITY, Visakhapatnam, India

[2]Department of Information Technology, GIT, GITAM UNIVERSITY, Visakhapatnam, India

[3]Department of Computer Science and Engineering, Srinidhi Institute of Science and Technology, Hyderabad, India

Email address:

premchandermunimanda@gmail.com (M. P. Chander), mandapati_venkat@yahoo.com (M. V. Rao), rajinitv@gmail.com (T. V. Rajinikanth)

Abstract: This paper focuses on early stage lung cancer detection. Lung cancer is prominent cancer as it states large number of deaths of more than a million every year. It creates need of detecting the lung nodule at early stage in Computer Tomography medical images. So to detect the occurrence of cancer nodule at early stage, the requirement of methods and techniques is increasing. There are different methods and techniques existing but none of them provide a better accuracy of detection. One of the techniques is content based image retrieval Computer Aided Diagnosis System (CAD) for early detection of lung nodules from the Chest Computer Tomography (CT) images. This optimization algorithm allows physicians to identify the nodules present in the CT lung images in the early stage hence the lung cancer. The MATLAB image processing toolbox based implementation is done on the CT lung images and the classifications of these images are carried out. The performance measures like the classification rate and the false positive rates are analyzed.

Keywords: Classification, Lung Cancer Detection, Accuracy, Image Processing Techniques

1. Introduction

In nature, lung disease plays a major role in health issue. In any form of lung disease mainly the breathing gets affected, here are some common forms of lung diseases. Acute bronchitis, asthma, Chronic Obstructive Pulmonary Disease (COPD), Acute Respiratory Distress Syndrome (ARDS) and Lung cancer. This annual report provides the estimated numbers of new cancer cases and deaths in 2015, as well as current cancer incidence, mortality, and survival statistics and information on cancer symptoms, risk factors, early detection, and treatment. In 2015, there will be an estimated 1,658,370 new cancer cases diagnosed and 589,430 cancer deaths in the US.

As per World cancer report 2016 lung cancer is the most common cause of cancer-related death in men and women, The major causes of the lung diseases are smoking, inhaling the drugs, smoke and allergic materials. The computed tomography (CT) images assists in detecting the extreme of the lung diseases. For the analysis of the proposed method CT image is sufficient also the visibility of soft tissue is better. There are several types of lung cancer, and these are divided into two main groups: Small cell lung cancer and non-small cell lung cancer which has three subtypes: Carcinoma, Adeno carcinoma and Squamous cell carcinomas [1].

Proper lung cancer staging is essential for optimizing therapy and assessing prognosis. A cancer's stage is based on the size or extent of the primary tumor and whether it has spread to nearby lymph nodes or other areas of the body. A number of different staging systems are used to classify cancer. Lung cancer is the second most commonly diagnosed cancer in both men and women. An estimated 224,390 new cases of lung cancer are expected in 2016, accounting for about 14% of all cancer diagnoses. The mentioned CAD schemes were not applied directly to the domain of lung cancer screening, yet they provide theoretical justification and potential to be effective tools to improve CT lung cancer screening. One has to examine carefully the reported results

and their applicability to lung cancer screening.

Early Detection of Lung Cancer

Screening with low-dose spiral computed tomography (LDCT) has been shown to reduce lung cancer morality by 20% compared to standard chest x-ray among adults with at least a 30 pack-year smoking history who were current smokers or had quit within 15 years. The American Cancer Society guidelines for the early detection of lung cancer endorse a process of shared decision making between clinicians who have access to high-volume, high-quality lung cancer screening programs and current or former smokers (quit within 15 years) who are 55 to 74 years of age, in good health, and with at least a 30 pack-year history of smoking. Shared decision making should include a discussion of the benefits, uncertainties, and harms associated with lung cancer screening. Appropriate treatment for lung cancer is based on whether the tumor is small cell (13%) or non-small cell (83%), as well as other tumor characteristics. Based on type and stage of cancer, as well as specific molecular characteristics of cancer cells, treatments can include surgery, radiation therapy, chemo -therapy, and/or targeted therapies. For early stage non-small cell lung cancers, surgery is usually the treatment of choice; chemo-therapy (sometimes in combination with radiation therapy) may be given as well. Advanced-stage non-small cell lung cancer patients are usually treated with chemotherapy, targeted drugs (or a combination of the two), or immunotherapy. Chemo therapy, alone or combined with radiation, is the usual treatment for small cell lung cancer; on this regimen, a large percentage of patients experience remission, though the cancer often returns.

2. Related Work

Aparna kanatte e*t al* [1] describes Lung carcinoma is one of the most lethal of cancers worldwide. Positron emission tomography (PET) data has greater sensitivity and specificity in the staging of lung cancer than computed tomography (CT) or magnetic resonance imaging (MRI). By using k-nearest neighbor and support vector machines (SVM) classifiers. Wavelet features with SVM classifier gave a consistent accuracy of 97% with an average sensitivity and specificity of 0.81 and 0.99 respectively.

Yongbum Lee *et. al* [2] proposed novel template matching technique based on genetic algorithm(GA) template matching (GATM)for detecting nodules existing in lung area, for Computer- Aided Diagnosis (CAD) systems to detect lung nodule in helical X-ray pulmonary computed tomography (CT) images. The GA was used to determine the target position in the observed image efficiently and to select an adequate template image from database. By using this method detection of nodule rate is about 72%.

Samuel H Hawkins *et. al* [3] presented on the focusing on cases of the adenocarcinoma nonsmall cell lung cancer tumor subtype from a larger data set. Comparison of classifiers for future selection approaches. Classifiers can be used to build to predict survival time.

Xing CHEN *et. al* [4] proposed a noninvasive detection method of lung cancer combined with a sort of virtual l SAW gas sensors and imaging recognition method. Patients breathe goes through an electronic nose with solid phase micro extraction (SMPE) and capillary column for pre-concentration and separation of volatile organic compounds (VOCs) respectively.

Noha Lee *et. al* [5] provides insight into effectiveness of lung cancer screening and assesses the potential of computer-aided design developments. CAD systems for lung tissue discrimination, nodule discrimination, and nodule characterization. CAD Technology may help radiologists alter the benefit- cost calculus of CT sensitivity and specificity in lung cancer screening.

David S *et. al* [6] developed CAD algorithm to detect lung nodule and polyp detection using CAD normal surface overlap method in helical Computed Tomography (CT) images.

Jyh-Shyan *et. al.* [7] developed a neural- digital computer-aided diagnosis system (CAD) system based on a parameterized two level convolution neural network (CNN) Architecture and on a special multipliable output encoding procedure. The developed architecture was trained, tested and evaluated specially on the problem of diagnosis of lung cancer nodule found on digitized chest radio- graphs. The system performs automatic suspect localization feature extraction and diagnosis of a particular pattern class-aimed at a high degree of "true-positive fraction" detection and "low positive fraction" detection.

Freedman *et al.* [8] assessed the detection of signs consistent with lung cancer on chest radiographs and breast cancer on mammograms. There are systems for other diseases and other types of images under development; however, this process depends on the availability of an accurate database. The performance of a Siemens Lung Care CT CAD system against a radiologist-generated database of CT imaged nodules. CAD and individual radiologists participating in the evaluation performed similarly.

Mac Redmond R *et al.* [9] presented the prevalence of respectable lung cancer detected by LDCCT at baseline screening was low at 0.23%, but there was a high rate of significant incidental pathology. Low dose chest computed tomographic scanning (LDCCT) can detect early stage asymptomatic lung cancer in a high risk urban population. Four hundred forty patients underwent surgery for primary lung cancer, and 45 normal cases were selected. Eight radiologists participated in observer tests.

Basavanna *et al.* [11] provided guidelines for clinically relevant FPR and TPR measures [10] as well as special ROC methods for cancer screening, [15-18] and also noted that since the underlying prevalence of cancer in average risk populations is very low, the FPR should be very small for acceptable cancer screening of asymptomatic people.

Bagyasri *et al.* [15], the authors described a theoretical analysis on how to combine classifiers with an optimal decision rule and optimal ROC curve. To provide clinically relevant definitions for *sensitivity* and *specificity* [12-13].

Anam Tariq *et al.* [17] investigated the effect of a CAD scheme on radiologist performance in the detection of lung cancers on chest radiographs.

Ruchka *et al.* [14] reported a CAD system for nodule detection using a difference-image technique. They compared several rule-based schemes for identifying nodules. A massive-training ANN (MTANN) [34] reduced the false positives.

Danshensong *et al.* [18] reported on a CAD scheme to help radiologists improve the detection of pulmonary nodules in chest radiographs by focusing on false positive reduction. They could reduce the number of false positives to 44.3% with a small increase in the number of true positives of 2.3%.

Kesav kancherla *et al* [21-23], obtained an accuracy of 81% using 71 features related to shape, intensity and color in our previous work. By adding the nucleus segmented features we improved the accuracy to 87%. Nucleus segmentation is performed by using Seeded region growing segmentation method. Our results demonstrate the potential of nucleus segmented features for detecting lung cancer in early stage.

R. Sah *et al.* [24] Five year survival in stage I carcinoma was not influenced by histological type, while there was statistically significant difference in survival between adenocarcinoma (0%) and squamous cell carcinoma (46%) in stage II disease, with early stage lung cancer reported patient-specific models for detecting lung nodules for use in screening and follow-up surveillance. In the last decades, a large body of research has been reported in the field of lung nodule detection and classification [2], [16-19].

S. Sone *et al* [25-27] most lung cancers are not localized when first detected, but early detection is mandatory to improve prognosis. Since curable early cases are hard to visualize with conventional chest radiography, a new diagnostic means must be found. We assessed whether population-based mass screening with a spiral computed tomography scanner could contribute substantially to detection of smaller cancers, and decrease mortality.

Kramer *et al.* [28] described an automated method to distinguish benign and malignant solitary nodules. Fifty-five chest radiographs were discriminated using LDA and ANN for feature combination and classification. Comparisons with manual grading showed that LDA had an AUC value of 88.6%, whereas manual identification resulted in an AUC value of 85.4%.

J. C. Nesbitt *et al* [29-30] Reducing morphologic differences by placing patients in groups based on the TNM subset and refinement in categorization by matching TNM subsets based on histology and other factors can improve considerably homogeneity and enhance prognostic predictability. The development of more accurate measures for predicting prognosis may serve to clarify the roles of primary and adjuvant treatment, particularly in those patients with early-stage disease associated with poor prognostic factors in whom the potential for long-term survival is reduced.

B. J. Flehinger *et al.* [35] 70 percent of the stage I patients in each program who were treated surgically survived more than five years, but there were only two five-year survivors among those who did not have surgery. We conclude that patients with lung cancers detected in stage I by chest x-ray film and treated surgically have a good chance of remaining free of disease for many years. Those stage I lung cancers which are not respected progress and lead to death within five years. Therefore, every effort should be made to detect and treat lung cancer early in high-risk populations.

A central concern in nodule detection is the high rate of false positives when sensitivity is increased to detect subtle nodules. A nodule is deemed a false positive result if it led to a completely negative workup or more than 12 months of follow-up with no cancer diagnosis. Reducing false positive rates while maintaining high sensitivity is still a difficult problem. Data mining techniques include LDA [34], rule-based approaches (a set of "if-then" statements), combinations of these two [22], artificial neural networks (ANNs), and maximum-margin based discriminators such as the SVM. Novel methodologies for searching have been introduced. Matching for detection [35], unsupervised clustering techniques [33], and a local density maximum algorithm [32]. Methods to improve discrimination of nodules from lung tissue include subtraction of vessels by region-growing.

3. Summary of Lung Nodule Detection and Classification Work

Image processing and data mining classification techniques are useful together have shown significant improvement in medical industry in terms of prediction, detection and decision making of lung cancer in early stage. Table 1 shows the summary of image processing and classification work, accuracy and sensitivity of various techniques.

Table 1. *Summary of Lung Nodule Detection and Classification.*

Author	Images	Technique		Year	Accuracy
		Segmentation algorithm	Classifier		
Aparna kanatte [1]	PET	Standard Uptake Values	KNN, SVM	2008	97%
Danshen song [18]	CT	Entropy Threshold	SVM	2000	85
Yongbum Lee [2]	CT	GATM	Classification	2001	72%
Atiyeh Hashemi [20]	CT	Region Growing	ANN	2013	95%
Basavanna [11]	CT	Max- Min cluster Algorithm	KNN, DT	2016	85%
Fatma Taher [12]	Sputum	Hopfield Neural Network(HNN)	Bayesian	2012	88.62%

Author	Images	Technique		Year	Accuracy
		Segmentation algorithm	**Classifier**		
Prashanth [13]	PET	Threshold Segmentation	KNN, SVM	2014	97%
Ruchika [14]	CT	Mean Shift Algorithm	Mathematical Morphology	2015	86%
Bhagyasri [15]		Marker controlled water shed Algorithm	Classification	2014	85.25%
Kesav kancherla [21]	Sputum	Speed region growing	Random forest	2013	87%
S K Vijai Anand [19]	CT	Optimal Therosholding	Back Propagation Network	2010	86.30%
S. Sivakumar [16]	CT	Weighted fuzzy probabilistic based clustering	SVM	2013	80.36%
Anam Tariq [17]	CT	Trshold segmentation	Neuro- fuzzy	2013	95%

4. Conclusion

In this paper, summarized the literature on different types of technologies of CAD approaches to improve lung cancer diagnosis through lung tissue discrimination, nodule detection/classification, and nodule characterization. CAD applications have shown high sensitivity but must also demonstrate high specificity to avoid cost-intensive, inconvenient, or even harmful follow-up procedures to rule out misclassified lung lesions. However, current protocols for nodule detection report many false positives, requiring substantial improvements in the technology and its applications prior to clinical application of CAD in the practice of lung cancer screening.

References

[1] Aparna Kanakatte, Nallasamy Mani, Bala Srinivasan, Jayavardhana Gubbi, "Pulmonary Tumor Volume Detection from Positron Emission Tomography Images", International Conference on Biomedical Engineering and Informatics, pp: 213-217, 2008.

[2] Yongbum Lee, Takeshi Hara, Hiroshi, shigeki and Takeo: IEEE Transactions on Medical Iamaging vol 20 No 7, July 2001.

[3] Samuel H Hawkins, Yoganand Balagurunathan, Virendra kumar, Lawrence, and Robert IEEE Access, vol 2, 2014.

[4] Xing CHEN, Mingfu CAO, Yan Hao Proceedings of the 2005 IEEE Engineering in medical and Biology 27th Annual Conf. China Sep 1-4, 2005.

[5] Noha Lee, Andrew F. L Aine, Guillermo, Jeffrey and John K. Gohagan IEEE Review in Bio Medical Engineering, vol 2, 2009.

[6] David S paik, Christipher, Geoffrey, rubeen, Rubak Acar, Joyoni Dey and Sandy Nepel: IEEE Transactions on Medical Iamaging vol 23 No 26, June 2004.

[7] Jyh-Shyan, Shih-Chung, Akira, Mattew and Seong: IEEE Transactions on Medical Iamaging vol 15, No 2, April 1996.

[8] M. Freedman, "Improved small volume lung cancer detection with computer-aided detection: Database characteristics and imaging of response to breast cancer risk reduction strategies," *Ann. NY Acad. Sci.*, vol. 1020, pp. 175–89, 2004.

[9] R. MacRedmond *et al.*, "Screening for lung cancer using low dose CT scanning," *Thorax*, vol. 59, pp. 237–41, 2004.

[10] Z. Liu and M. Tan, "ROC-based utility function maximization for feature selection and classification with applications to high-dimensional protease data, biometrics," *J. Int. Biometr. Soc.*, vol. 64, no. 4, pp. 1155–1161, 2008.

[11] Basavanna et. al, "Tumor Cell Identification in Medical Images using Image Processing Techniques" October, 2016.

[12] Fatma Taher, Naoufel Werghi and Hussain Al-Ahmed, "Bayesian Classification Artificial Neural Network Methods for Lung Cancer Early Diagnosis", IEEE, pp: 773-776, 2012.

[13] Prasshanth Naresh, Rajashree Shettar, "Image Processing and classification Techniques for early Detection of Lung Cancer for Preventive Healthcare: October, 2016.

[14] Ruchika, Ashima "Detection of Lung cancer in CT Images using Mean Shift Algorithm", volume 5, issue 5, May 2015.

[15] Bhagyasri G. Patil, Prof Sanjeev "Cancer cells Detection using Digital Image Processing Methods", Vol 3 Issue 4 March 2014.

[16] S. Sivakumar, Dr. C. Chandrasekar, "Lung Nodule Detection Using Fuzzy Clustering and Support Vector Machines", International Journal of Engineering and Technology (IJET), Vol 5 No 1, pp: 179-185, Feb-Mar 2013.

[17] Anam Tariq, M. Usman Akram and M. Younus Javed, "Lung Nodule Detection in CT Images using Neuro Fuzzy Classifier", Fourth International Workshop on Computational Intelligence in Medical Imaging (CIMI), pp: 49-53, 2013.

[18] Dansheng Song, Tatyana A. Zhukov, Olga Markov, Wei Qian, Melvyn S. Tockman, "Prognosis of stage i lung cancer patients through quantitative analysis of centrosomal features", ie, pp: 1607-1610, 3012.

[19] S. K. Vijai Anand, "Segmentation coupled Textural Feature Classification for Lung Tumor Prediction", ICCCCT, pp: 518-524, 2010.

[20] Atiyeh Hashemi, Abdol Hamid Pilevar, Reza Rafeh, "Mass Detection in Lung CT Images Using Region Growing Segmentation and Decision Making Based on Fuzzy Inference System and Artificial Neural Network", I. J. Image, Graphics and Signal Processing, 6, pp: 16-24, 2013.

[21] Kesav Kancherla, Srinivas Mukkamala, "Early Lung Cancer Detection using Nucleus Segmentation based Features", IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), pp: 91-95, 2013.

[22] M. Aoyama, Q. Li, S. Katsuragawa, H. MacMahon, and K. Doi, "Automated computerized scheme for distinction between benign and malignant solitary pulmonary nodules on chest images," *Med. Phys.*, vol. 29, pp. 701–8, 2002.

[23] J. A. Swets, "ROC analysis applied to the evaluation of medical imaging techniques," *Invest. Radiol.*, vol. 14, pp. 109–21, 1979.

[24] R. Sah *et al.*, "Results of surgical treatment of stage I and II lung cancer," *J. Cardiovasc. Surg.*, vol. 37, pp. 169–172, 1996.

[25] S. Sone *et al.*, "Mass screening for lung cancer with mobile spiral computed tomography scanner," *Lancet*, vol. 351, pp. 1242–5, 1998.

[26] C. I. Henschke *et al.*, "Early lung cancer action project: Overall design and findings from baseline screening," *Lancet*, vol. 354, pp. 99–105, 1999.

[27] D. S. Gierada, T. K. Pilgrim, M. Ford, R. M. Fagerstrom, T. R. Church, H. Nath, K. Garg, and D. C. Strollo, "Lung cancer: inter observer agreement on interpretation of pulmonary findings at low-dose CT screening," *Radiol.*, vol. 246, no. 1, pp. 265–272, 2008.

[28] B. S. Kramer and O. W. Brawley, "Cancer screening," *Hematol. Oncol. Clin. North Amer.*, vol. 14, pp. 831–48, 2000.

[29] L. Berlin, "Liability of performing CT screening for coronary artery disease and lung cancer," *Amer. J. Roentgenol.*, vol. 179, pp. 837–42, 2002.

[30] S. J. Swensen *et al.*, "Screening for lung cancer with low-dose spiral computed tomography," *Amer. J. Respir. Crit. Care Med.*, vol. 165, pp. 508–13, 2002.

[31] J. C. Nesbitt *et al.*, "Survival in early-stage lung cancer," *Ann. Thorac. Surg.*, vol. 60, pp. 466–472, 1995.

[32] C. I. Henschke *et al.*, "CT screening for lung cancer: Suspiciousness of nodules according to size on baseline scans," *Radiology*, vol. 231, pp. 164–8, 2004.

[33] L. L. Humphrey, S. Teutsch, and M. Johnson, "Lung cancer screening with sputum cytologic examination, chest radiography, and computed tomography: An update for the U.S. Preventive Services Task Force," *Ann. Intern. Med.*, vol. 140, pp. 740–53, 2004.

[34] J. Gohagan *et al.*, "Baseline findings of a randomized feasibility trial of lung cancer screening with spiral CT scan vs chest radiograph: The lung screening study of the National Cancer Institute," *Chest*, vol. 126, pp. 114–21, 2004.

[35] B. J. Flehinger, M. Kimmel, and M. R. Melamed, "Survival from early lung cancer: Implications for screening," *Chest*, vol. 101, pp. 1013–1018, 1992.

Biography

M. Prem Chander has been pursuing Ph. D degree in Computer Science and Engineering from GITAM UNIVERSITY, Visakhapatnam, A. P India, as well as working as Asst. Professor of CSE Department in MNR Engineering College, Hyderabad. As well as obtained M. Tech. in Computer Science and Engineering from Jawaharlal Nehru Technological University(JNTU), Anantapur, AP, INDIA in 2011. Prior to his professional career, he obtained Bachelor of Technology (B. Tech) in Computer Science and Information Technology from Sri Datta Institute of Engineering and Science, Hyderabad. His total teaching experience is 12 years.

M. Venkateshwara Rao has obtained his Ph. D degree in C. S. E branch. He is working as Associate Professor in the Department of Information Technology, in Gitam Institute of Technology, GITAM UNIVERSITY, Visakhapatnam, AP, India. He obtained his M. Sc (Electronics), M. Phil (Opto Electronics), and M. Tech (CST). His teaching and Industrial experience is 23 years. His writings have appeared in numerous Professional conferences and Journals (International journals-5 National conferences-2.

T. V. Rajini Kanth has obtained his Ph.D. degree in C. S. E. branch from Osmania university, Hyderabad in July, 2008 and M. Tech. (C. S. E.) degree from Osmania University, Hyderabad in January, 2001. His specialization area in research is "Spatial Data mining". He obtained his PGDCS degree from HCU, Hyderabad in 1996. He received his M. Sc. (Applied mathematics) degree in the year 1989 from S. V. University, Tirupati as University Ranker. He is currently working as professor in CSE, at SNIST, Hyderabad. His writings have appeared in numerous Professional conferences and Journals (International journals-29, national Journals-3, International conferences- 23, national conferences-1, Total =56). Under his guide ship 5 Ph.D. Research Scholars got awarded (3-JNTUH & 3-ANU).