A Comprehensive Review on Various State-Of-The-Art Techniques for Early Lung Cancer Detection

Abhir Bhandary¹ & Dr. Ananth Prabhu G²
¹Department of Information Science & Engineering
NMAM Institute of Technology, Nitte, Karkala, Karnataka, India
²Department of Computer Science & Engineering
Sahyadri College of Engineering & Management, Mangaluru, Karnataka, India

Abstract: Lung cancer is the first cause of deaths among different types of cancer. Early, chest X-ray films were used for lung cancer diagnosis. Nonetheless, X-ray films are 2D projection images that usually overlap different tissues hiding small lung cancer nodules. Computed Tomography (CT) is one of the best imaging modalities for detection of small lung nodules due to the fast acquisition times of successive images with high resolution and the minimization of artifacts caused by abdominal movement. However, there are many difficulties in CT interpretation such as the limitation of the human visual system, deficient training and experience, and factors like fatigue or distraction that leads to high rates of missed cancers. Here, a second opinion is usually required to reduce erroneous rates. In this research article we have analysed various existing techniques for early lung cancer detection.

INTRODUCTION

Nowadays the data collected in the medical knowledge of many different diseases is of great importance. Research on the data and obtain useful results and patterns of disease, one of the goals of this data is used. The high volume of data and the resulting confusion is a problem that prevents you achieves remarkable results. Thus, the image analysis and data mining to overcome this problem and to obtain useful relationships between risk factors used in disease. Image mining is the extraction of knowledge hidden in the pictures, the relationship between data, images or patterns that are stored in the image are not clear and transparent discussed. Image mining is more than just the development of data mining in the area of the image. According to the up-to-date statistics from American Cancer Society [1], lung cancer is the leading cause of cancer related deaths with over 159,000 deaths estimated for the United States alone in 2013, and the overall 5-year survival rate for lung cancer is merely 16%. The survival rate increases to 52% if it is localized, and decreases to 4% if it has metastasized. Therefore, to detect lung cancer at earlier stages is of great importance [2], and computer-aided detection (CADe) in supplement to radiologists’ diagnosis has become a promising tool to serve such purposes [3]. While detection of pulmonary nodules has a crucial effect on the diagnosis of lung cancer, but the detection is a nontrivial task, not only because the appearance of pulmonary nodules varies in a wide range, but also because the nodule densities have low contrast against adjacent vessel segments and other lung tissues.

Computed tomography (CT) has been shown as the most popular imaging modality for nodule detection [2], [4], because it has the ability to provide reliable image textures for the detection of small nodules. The development of lung nodule CADe systems using CT imaging modality has made good progress over the past decade [5], [6]. Generally, such CADe systems consist of three stages: (1) image preprocessing, (2) initial nodule candidates (INCs) identification, and (3) false positive (FP) reduction of the INCs with preservation of the true positives (TPs). Image mining is an interdisciplinary field of expertise on machine vision, image processing, image retrieval, data mining, machine learning, databases and artificial intelligence based. In addition, in each of these areas there has been a lot of research; research in image mining is a new problem. A high volume of images, such as satellite images, medical images and digital photographs are produced on a daily basis that if these images are analysed, very useful information in our possession are. There are two basic techniques of image mining, first technique in the exploration of a large number of independent pictures to do. The second technique is explored in a series of integrated and linked images do. The main objective of image analysis of all significant patterns from images without knowing the details of the content of the images. This means
that without a basic knowledge of image content is as clever as input a set of images, to extract important patterns. Patterns can be extracted from various types of classified patterns, symbologies, dependent patterns, temporal patterns and spatial patterns.

**ANALYZING VARIOUS LATEST STATE-OF-THE-ART TECHNIQUES FOR EARLY LUNG CANCER DETECTION**

In the article, “Lung Nodules Classification Based on Growth Changes and Registration Technology” by Tong Jia, Yukun Bai, Hao Zhang, DongYue Chen, XiaoSheng Yu, ChengDong Wu doubling time to complete the diagnosis of small benign and malignant nodules is computed, solving the difficult problem of their diagnosis. Before calculating the growth rate of lung nodules, a new method is proposed for registration on follow-up CT image registration. First, lung CT images global rigid registration. In the rigid registration framework, we research the registration accuracy for different method. The global rigid image registration results will be estimated in the initial parameters as the next non-rigid registration later. Then registration method based on B-spline partial registration. In the process of local non-rigid registration for lung nodules region, adding constraints to prevent overdriving effect caused by similar measure. In the end, method is feasible and it can get a better result.

In the article, “Automatic Lung Nodule Segmentation and Classification in CT Images Based on SVM” by Elmar Rendon-Gonzalez and Volodymyr Ponomaryov an automatic CADe system is presented for segmentation and classification of lung nodules classification in CT images with a sensibility of 84.93%, specificity of 80.92% and an accuracy performance of 78.08%. The contributions of the proposed method in CAD system design are as follows: a) It can select lung nodule parenchymal tissue accurately; b) It has the ability to select the suspicious zones effectively and can reduce in a proper way FPs; c) Quality criterions of designed system shows a good enough performance and can distinguish between benign and malignant nodules.

In the article, “A Computer Aided Diagnosis for detection and classification of lung nodules” by Lakshmi Narayanan A and Prof. Jeeva J. B the CT images of lungs are examined for the possible detection of lung nodules. Initially the images are pre-processed where the contrast levels are adjusted and user is allowed to crop the image to select the ROI. Then segmentation is done and nodules are identified from the ROI. The features are extracted from the nodules and it is given as an input to artificial neural networks for classification. A feed forward network is used for this purpose. The CAD system is able to identify the nodule that lies over the blood vessel. An overall accuracy of 92.2% and a very low false positive rate of 0.9% were achieved by the proposed method. Texture features can be used to decrease the false positive rate and thereby increasing the detection accuracy.

In the article, “Computer-Aided Detection of Pulmonary Nodules based on SVM in Thoracic CT Images” by Parinaz Eskandarian and Jamshid Bagherzadeh a novel CADE system was proposed for detection of pulmonary nodules in chest CT scans. Based on our previous work of self-adaptive online SVM for image segmentation, we developed a SVM scheme for INCs detection. The SVM classification results indicated that gradient features contributed the most against any of the other two groups of features (geometric, intensity). The forward feature selection strategy showed that the SVM classifier performed the best in the gradient with intensity feature space rather than for any other feature combinations. In terms of free response ROC analysis, the proposed CADEs system achieved an overall sensitivity of 90% at 4 FPs per scan. Compared with existing CADE systems evaluated on the same lung image LIDC database, our approach showed a comparable detection capability but a lower computational cost. The presented CADE system yields comparable detection accuracy and more computational efficiency than existing systems, which demonstrates the feasibility of our CADE system for clinical utility.

In the article, “Improved Fuzzy C-Means Clustering Algorithm for Automatic Detection of Lung Nodules” by Fan Liao and Chunxia Zhao method inherits the advantages of the traditional FCM algorithm, and proposes the pre-treatment using of human visual attention mode, complete fuzzy clustering based on visual focus of attention, which provides an important foundation for the next doctor's diagnosis, and improves the efficiency and accuracy of diagnosis.

In the article, “Lung Nodule Diagnosis from CT Images Based on Ensemble Learning” by Farzad Vasheghani Farahani, Abbas. Ahmadi and M.H. Fazel Zarandi the effectiveness of the proposed ensemble intelligent system is evaluated for diagnosis of lung nodules, experiments are conducted on the LIDC database mentioned above. In the first phase, image processing techniques including preprocessing, segmentation, post processing and feature extraction were used. Then, specific features (Roundness, Circularity,
In the article, “Automatic Immuno scoring of Immuno histochemistry Images of Human Lung Cancer Tissue Samples: A Screening Tool,” by Shu Zhang, Jianfei Huang, Jianguo Zhang, Mingyan Qiu, Huiqun Wu, Yifei Liu, Yan Wang, Qin Jin, Wei Wang, Huijun Zhu and Hua Huang the image based lung cancer stage prediction is simple, sensitive, and easy to perform. Application to large sample series is imperative to evaluate correlation with other generally accepted techniques.

In the article, “Lung Nodule Volume Measurement using Digital Chest Tomosynthesis” by D. Hadházi, B. Czétényi, Á. Horváth, G. Orbán, Á. Horváth and G. Horváth A simple way to measure nodule volume is to determine the diameter of the nodules in the coronal slices, and assuming that all of the nodules form ideal spheres—the volume of the nodules is estimated with the volume of these spheres. However, this approach is rather inaccurate in real cases according to our practice and to too. A better way is if the nodules are modelled with ellipsoids. This also needs to segment nodules in coronal slices and to estimate the nodule size in the perpendicular direction. For segmentation we apply gradient vector flow-based ACM Snake while for the estimation of the size of projection of nodules in the in-depth direction feature-based machine learning methods are applied.

In the article, “Lung Nodule Classification Using Deep Features in CT Images” by Devinder Kumar, Alexander Wong and David A. Clausi a CAD classifier system for classifying lung nodules as either malignant or benign is presented. The proposed system uses deep features extracted from an autoencoder for annotations provided by up-to four radiologists for 157 patients to precisely create a strong representation of nodules. Using the LIDC dataset, we showed that the proposed system convincingly outperforms the state-of-the-art method on overall accuracy metric even after experimenting with almost five times the data size (4323 vs. 914) used in the state-of-the-art method and considering the biopsy level clinical decision as ground truth. This is because the deep features not only take the different conventional semantic features like lobulation, speculation etc. into account but they also take into account the association between them.

In the article, “Computer Aided Diagnosis System for Early Lung Cancer Detection” by Fatma Taher, Nauufel Werghi and Hussain Al-Ahmad a novel CAD system for lung cancer early diagnosis. A set of different features such as a nucleus to cytoplasm ratio, curvature, circularity, Eigenvectors ratio and density of the nucleus region were extracted from the nucleus region to be used in the diagnostic process. Two classification techniques were applied to the selected features: ANN and SVM. The comparison demonstrates a clear superiority of the SVM classifier, where it gives the highest accuracy. Therefore, experimentation suggests that the SVM classifier is much better than the ANN. The SVM exhibits an elegant and methodological choice to classify the sputum images to benign and malignant. At the current stage, the SVM obtained has a high accuracy of 97%, which is superior to other methods. The preliminary results of the CAD system have yielded promising results that would supplement the use of current technologies for diagnosing lung cancer. In future studies, we plan to increase the elements of the feature extraction for better classification and use various combinations of pre-existing features, and to consider other nonlinear modalities in the SVM.

**CONCLUSION**

The detection of lung cancer can be done in several ways, such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and X-ray. All these methods consume a lot of resources in terms of both time and money, in addition to their invasiveness. Recently, scientists have proven that the non-invasive technique of sputum cell analysis can assist in the successful diagnosis of lung cancer. All of this motivates us to come up with an automatic diagnostic system for early detection of lung cancer based on the analysis of the sputum colour images. Especially, when dealing with large amounts of data, in addition to relieving doctors from tedious and routinely task. The design and development of sputum colour image segmentation is an extremely challenging task. Hence in this research article we have analysed various existing techniques for early lung cancer detection.

**REFERENCES**

2. R.N. Strickland, Tumor detection in non stationary backgrounds, IEEE Transactions on Medical Imaging 13(1 une) (1994) 491--499

Compactness, Ellipticity, and Eccentricity) of interested objects were fed to classification phase. In the classification phase, these five features were used as inputs for single classifiers (MLP, KNN, and SVM) and ensemble system. Then, majority voting method was applied to combine results of base classifiers in committee.


5. Cherry BuAungan, Xiuying Dagan Feng, Stefan Eberl, and Michael Fulham, Lung segmentation and tumor detection from CT thorax volumes of FDG PET-CT scans by template registration and incorporation of functional information, Nuclear Science Symposium Conference Record, pp. 5349-5352, 2008

6. Takeshi Hara, Hiroshi Fujita, Yongbum Lee, Hitoshi Yoshimura and Shoji Kido, Automated lesion detection methods for 2D and 3D chest x-ray images


23. Wenyuan Qi. medical image non-rigid registration fast algorithm research based on a


67. Syedia Tahseen Fatima Bokhari, Muhammad Sharif, Mussarat Yasmin, Steven Lawrence Fernandes, Fundus Image Segmentation and Feature Extraction for the Detection of
Glaucmae: A New Approach, Current Medical Imaging Reviews, Special issue on Frontiers in Intelligent Systems Development for Medical Image Processing & Computer Vision, USA


74. He, Dong-Jian, “Improved saliency toolbox/Itti model for region of interest extraction Optical Engineering,” vol.50, pp: 100-101, September 2011.


90. N. Esfandiari, M. R. Babavalian, A.-M. E. Moghadam, and V. K. Tabar, “Knowledge discovery in medicine: Current issue and future
91. S. M. Naqi, Muhammad Sharif, Mussarat Yasmin, Steven Lawrence Fernandes, Lung Nodule Detection Using Polygon Approximation and Hybrid Features from Lung CT Images, Current Medical Imaging Reviews, Special issue on Frontiers in Intelligent Systems Development for Medical Image Processing & Computer Vision, USA
92. Steven Lawrence Fernandes, G. Josemin Bala, “Robust Human Skin Detection using OpenCV on ARM based ODROID XU4 Heterogeneous Multi-Processing Device” ACM – ICPS, ISBN No 978-1-4503-3962-9