A Novel CADe System for Detection of Pulmonary Nodules in Thoracic CT Imaging

Sameera R
Applied Electronics and Instrumentation
Younus College of engineering and technology
Kollam, India

Safuvan T
Assistant Professor
Dept. of Electronics and Communication
Younus College of engineering and technology
Kollam, India

***Abstract*— Computer-aided detection (CADe) of pulmonary nodules is critical to assisting radiologists in early identification of lung cancer from computed tomography (CT) scans. This paper proposes a novel CADe system based on a hierarchical vector quantization (VQ) scheme. Compared with the commonly-used simple thresholding approach, the high-level VQ yields a more accurate segmentation of the lungs from the chest volume. In identifying initial nodule candidates (INCs) within the lungs, the low-level VQ proves to be effective for INCs detection and segmentation, as well as computationally efficient compared to existing approaches. False-positive (FP) reduction is conducted via rule-based filtering operations in combination with a feature-based support vector machine classifier. The proposed system was validated on 205 patient cases from the publically available online Lung Image Database Consortium database, with each case having at least one juxtapleural nodule annotation.**

Keywords*— Computer-aided detection (CADe), computed tomography (CT) imaging, vector quantization (VQ).

I.INTRODUCTION

According to the up-to-date statistics from the American cancer society, 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 five-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, and computer-aided detection (cade) in supplement to radiologists' diagnosis has become a promising tool to serve such purposes. 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 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, 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. 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). In the preprocessing stage, the system aims to largely reduce the search space to the lungs, where a segmentation of the lungs from the entire chest volume is usually required. Because of the high image contrast between lung fields and the surrounding body tissue, image intensity-based simple thresholding is effective, and is currently the most commonly used technique for lung segmentation. However, the determination of an accurate threshold is greatly affected by image acquisition protocols, scanner types, as well as the inhomogeneity of intensities in the lung region, especially toward the segmentation of pathological lungs with severe pathologies. This paper proposes an adaptive solution to mitigate the difficulty of thresholding-based method in lung segmentation.

After defining the search space (i.e., the lung volume), incs detection is the next step to build a cade system. Various incs detection techniques have been extensively studied in recent years, such as multiple thresholding.

The most commonly used multiple thresholding approach aims to find connected components of similar image gray-values. Though intensity-based thresholding methods are computationally cheaper than other pattern-recognition techniques for the detection of INCS, they also suffer considerable drawbacks. For example, it is difficult to adaptively determine multiple thresholds, because pulmonary nodules with a wide intensity range are embedded in an inhomogeneous parenchyma background.

On the other hand, pattern-recognition techniques. Are complicated and usually computational intensive. This paper proposes an efficient means which shares the adaptive natures of pattern-recognition techniques and the simplicity of intensitybased thresholding methods. Sufficient detection power for nodule candidates is inevitably accompanied by many (obvious) fps. A rule-based filtering operation is often employed to cheaply and drastically reduce the number of obvious fps, so that their influence on the computationally more expensive learning process can be eliminated. In general, fp reduction using machine learning has been extensively studied in the literature.

II. LITERATURE SURVEY

[1] Maintaining a statewide cancer registry that meets both National Program of Cancer Registries and Centers for Disease Control and Prevention (CDC) high quality data standards and North American Association of Central Cancer

Registries (NAACCR) gold certification is accomplished through collaborative funding efforts.

[2] Describes the choice of treatment for patients with cancers diagnosed as a result of screening are selected by the treating physician in conjunction with the participant. However, each participating institution must be committed to document, for each diagnosed case of lung cancer, the timing and nature of the intervention(s) (if any) and also the prospective course in respect to manifestations of metastases. The development and refinement of the screening protocol has been a concern of the ELCAP (Early Lung Cancer Action Program) Group for more than two decades, and it has been updated in the framework of the International Conferences organized by this Group and in the resultant international consortium on screening for lung cancer, I-ELCAP.

[3] Describes an image-based CAD system for early detection of prostate cancer using DCE-MRI is introduced. Prostate cancer is the most frequently diagnosed malignancy among men and remains the second leading cause of cancer-related death in the USA with more than 238,000 new cases and a mortality rate of about 30,000 in 2013. Therefore, early diagnosis of prostate cancer can improve the effectiveness of treatment and increase the patient's chance of survival. Currently, needle biopsy is the gold standard for the diagnosis of prostate cancer. However, it is an invasive procedure with high costs and potential morbidity rates. Additionally, it has a higher possibility of producing false positive diagnosis due to relatively small needle biopsy samples.

[4] Describes there is no clear consensus regarding the definition of a pulmonary nodule. Yet, "nodule" is one of the most common words found in chest CT reports. A committee of the Fleischer Society on CT nomenclature defined a pulmonary nodule as "a round opacity, at least moderately well marinated and no greater than 3 cm in maximum diameter.

III. PROPOSED METHOD

VQ was originally used for data compression in signal processing, and became popular in a variety of research fields such as speech recognition, face detection, image compression and classification, and image segmentation. It allows for the modeling of probability density functions by the distribution of prototype vectors. The general VQ framework evolves two processes: 1) the training process which determines the set of codebook vector according to the probability of the input data; and 2) the encoding process which assigns input vectors to the codebook vectors. The well-known Linde– Buzo–Gray (LBG) algorithm has been widely used for the design of vector quantizer. The algorithm aims to minimize the mean squared error and guarantees to converge to the local optimality. A very important but difficult task in the CADe of lung nodulesis the detection of INCs, which aims to search for suspicious 3-D objects as nodule candidates using specific strategies. This step is required to be characterized by a sensitivity that is as close to 100% as possible, in order to avoid setting a priori upper bound on the CADe system performance. Meanwhile, the INCs should minimize the number of FPs to ease the following FP reduction step. This section presents our hierarchical VQ scheme for automatic detection and segmentation of INCs.

IV. METHODOLOGIES

Detection of pulmonary nodules has a crucial effect on the diagnosis of lung cancer, but the detection is a nontrivial Computed tomography (CT) has been shown as the most popular imaging modality for nodule detection, 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 task, not only because the appearance of pulmonary nodules varies n a wide range, but also because nodule densities have low contrast against adjacent vessel segments and other lung tissues

Mainly three modules

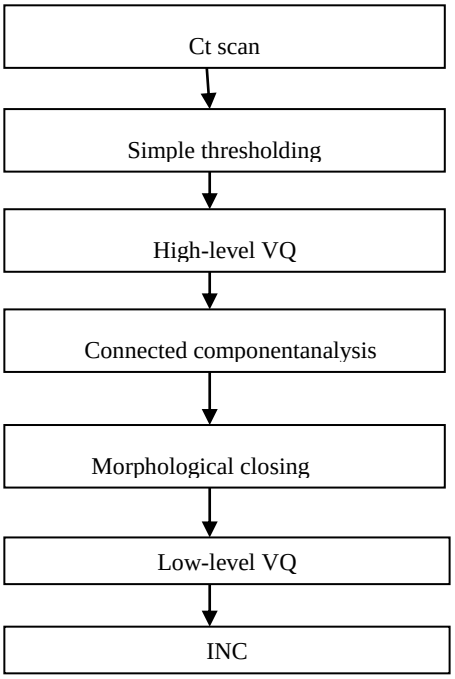
A. Self-adaptive VQ algorithm

VQ was originally used for data compression in signal processing, and became popular in a variety of research fields such as speech recognition, face detection, image compression and classification, and image segmentation. It allows for the modelling of probability density functions by the distribution of prototype vectors. The general VQ framework evolves two processes: 1) the training process which determines the set of codebook vector according to the probability of the input data; and 2) the encoding process which assigns input vectors to the codebook vectors. The well-known Linde– Buzo–Gray (LBG) algorithm has been widely used for the design of vector quantizer. The algorithm aims to minimize the mean squared error and guarantees to converge to the local optimality.

B. Incs detection via a hierarchical VQ scheme

A very important but difficult task in the CADe of lung nodules is the detection of INCs, which aims to search for suspicious 3-D objects as nodule candidates using specific strategies. This step is required to be characterized by a sensitivity that is as close to 100% as possible, in order to avoid setting a priori upper bound on the CADe system performance. Meanwhile, the INCs should minimize the number of FPs to ease the following FP reduction step. This section presents our hierarchical VQ scheme for automatic detection and segmentation of INCs.

Block diagram of hierarchical VQ scheme



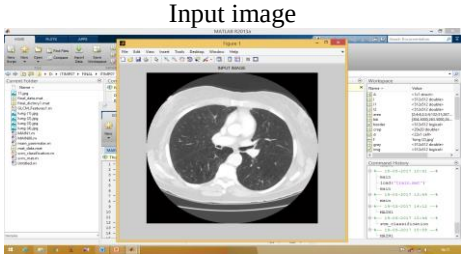
C.False positive reduction from incs

1. Rule-Based Filtering Operations:
It is challenging to thoroughly separate nodules from attached structures due to their similar intensities, especially for the juxta-vascular nodules (the nodules attached to blood vessels). Since the thickness of blood vessels varies considerably (e.g., from small veins to large arteries), a 2-D morphological opening disk with radius of 1 up to 5 pixels was adopted to detach vessels at different degrees.
2. Feature-Based SVM Classification:
A supervised learning strategy is carried out using the SVM classifier to further reduce FPs. Our feature-based SVM classifier relies on a series of features extracted from each of the remaining INC after rule-based filtering operations.

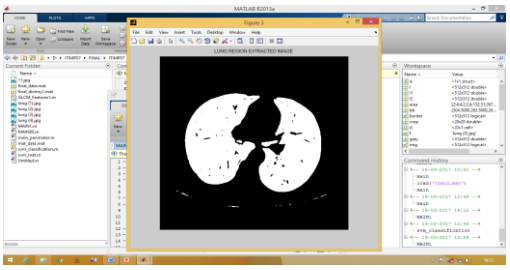
V.EXPERIMENTAL RESULTS

A.INC Detection and Rule-Based Filtering Performance

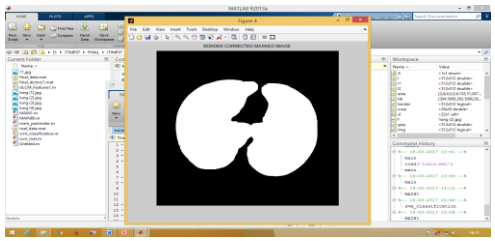
Input image



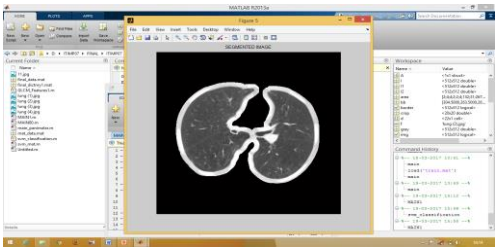
Lung region extraction



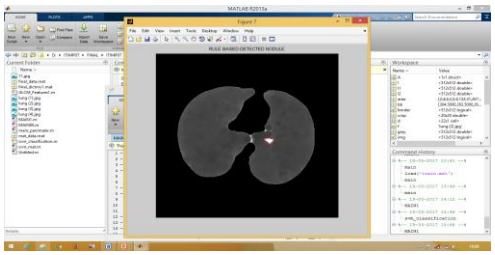
Border corrected lung mask



Segmented lung mask



Rule based detected nodule



VI.DISCUSSION AND CONCLUSION

In this paper, a novel CADe system was proposed for fast and adaptive detection of pulmonary nodules in chest CT scans. Based on our previous work of self-adaptive online VQ for image segmentation, we developed a hierarchical VQ scheme for INCs detection. The high-level VQ proves to be feasible to replace the commonly-used simple thresholding scheme for extraction of the lungs with higher accuracy, as well as comparable processing time and automation level. The following low-level VQ illustrates adequate detection power for nonGGO nodules, and is computationally more efficient than the state of-the-art approaches. In this study, simple expert rules were firstly employed to exclude obvious FPs from being considered by the sophisticated feature-based SVM classifier. The SVM classification results indicated that gradient features contributed the most against any of the other

three groups of features (geometric, geometric, intensity, and Hessian features). The forward feature selection strategy showed that the SVM classifier performed the best in the "gradient + intensity" feature space rather than for any other feature combinations. 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. In particular, we reported the performance of our system for the detection of juxta-pleural nodules. The proposed hierarchical INCs detection approach is fast, adaptive, and fully automatic. 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.

ACKNOWLEDGMENT

The author thankfully acknowledgment Mr. Safuvan T, Assistant Professor, YCET, without whose guidance and supervision, this would not been possible. I would like to extent my gratitude to the Head of the Department of Electronics and communication Mr. Rajeev S K who were always ready to help me with ideas and suggestions for rectifying the mistakes that crept up time to time during the completion of this venture. I would also like to thank my friends and last but not the least the staff of ECE department for their support and encouragement. Above all, I am thankful to the god almighty.

REFERENCES

[1] R. Siegel, D. Naishadham, and A. Jemal, "Cancer statistics," CA Cancer J. Clin., vol. 63, pp. 11–30, 2013.

[2] C. I. Henschke, D. I. McCauley, D. F. Yankelevitz, D. P. Naidich, G. McGuinness, O. S. Miettinen, D. M. Libby, M. W. Pasmantier, J. Koizumi, N. K. Altorki, and J. P. Smith, "Early lung cancer action roject: Overall design and findings from baseline screening," Lancet,vol. 354, no. 9173, pp. 99–105, 1999.

[3] A. El-Baz and J. Suri, Lung Imaging and Computer Aided Diagnosis. Boca Raton, FL, USA: CRC Press, 2011.

[4] H. MacMahon, J. H. M. Austin, G. Gamsu, C. J. Herold, J. R. Jett, D. P. Naidich, E. F. Patz, and S. J. Swensen, "Guidelines for management of small pulmonary nodules detected on CT scans: A statement from the Fleischner society," Radiology, vol. 237, no. 2, pp. 395–400, 2005.

[5] B. van Ginneken, S. G. Armato 3rd, B. de Hoop, S. van Amelsvoort-van de Vorst, T. Duindam, M. Niemeijer, K. Murphy, A. Schilham, A. Retico, M. E. Fantacci, N. Camarlinghi, F. Bagagli, I. Gori, T. Hara, H. Fujita, G. Gargano, R. Bellotti, S. Tangaro, L.Bola˜nos, F. De Carlo, P. Cerello, S. Cristian Cheran, E. Lopez Torres, M. Prokop, "Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: The ANODE09 study," Med. Image Anal., vol. 14, no. 6, pp. 707–722, 2010.

[6] A. El-Baz, G. M. Beache, G. Gimel'farb, K. Suzuki, K. Okada, A. Elnakib, A. Soliman, and B. Abdollahi, "Computer-aided diagnosis systems for lung cancer: Challenges and methodologies," Int. J. Biomed. Image., vol. 2013, pp. 1– 46, 2013.