Improving the Detection Rate of CAD-Methods for Lung Nodule Detection in CT Scans Through Features of Local Curvature.

Introduction

Cancer is the leading cause of death in Canada, ahead of cardiovascular disease, cerebrovascular diseases and chronic lower respiratory diseases [1]. Almost half of all Canadians are projected to develop cancer within their lifetime and a quarter of all Canadians expected to die of the disease [1]. In 2014, amongst nonsmokers and smokers, lung cancer had become the third most common cancer in new cancer cases for males at 14%, and the second most common new cancer for females at 13% [1,2]. However, the risk of lung cancer for smokers, relative to the risk amongst those who have never smoked before, is increased by an order of tenfold or more [3]. Unfortunately, lung cancer ranks as the leading cause of cancer deaths by a significant margin for both sexes, having caused 26.5% of all Canadian cancer-related deaths in 2014[1]. While the five-year survival rate of lung cancer is generally poor, averaging 5-10%[3], the survival rate is strongly correlated with the stage of the disease at diagnosis. Those patients diagnosed at stage I have the highest rate of five-year survival at 35%, whereas patients diagnosed at stage III have a survival rate of 6% and patients diagnosed at stage IV do not tend to survive longer than 2 years [4]. Hence, it is vitally important to diagnose lung cancer as early as possible to improve patient outcomes.

Traditionally, mass scanning by radiography has been unpromising, but low-dose computed topography (CT) has shown great promise in mass screening programs for lung cancers [5]. Sone et al [5] found that over the course of a three-year mass screening trial of repeat CT scans, they were able to allow for a detection rate nearly 11 times greater than the expected annual number of lung cancers. More impressively, 88% of lung cancers identified in the trial were in stage 1, leading to medical intervention with the highest chance of a positive patient outcome. Therefore, great strides could be taken to reduce the mortality rate of lung cancers by implementing a more widespread mass-screening program capable of catching new lung cancer cases in their early stages.

However, repeat mass screenings of the general population require many trained experts who can examine the scans for possible suspicious masses. Unfortunately as the scale of the screenings increases, it can be difficult or uneconomical to find a suitable number of these trained experts. In a study on mammography screenings, Morton et al [6], found that up to 77% of newly detected breast cancers were present in prior mammograms. However, in the prior mammograms, the breast cancer was missed due to perceptual errors by the lone radiologist performing the screening. Furthermore, they found that when two or more radiologists examined each mammography screening, 5-15% more breast cancer cases were detected. However, this system of double blind screening is not practiced in North America due to the prohibitive resource cost associated with implementing it [6]. Similarly, perceptual errors in the analysis of lung cancer screenings are likely to occur. These errors are often due to complex and potentially misleading anatomical structures within the lung, such as soft tissue, blood vessels, or airways[7]. Thus, a computer-aided diagnostic (CAD) method that is capable of detecting suspicious masses, or a method of reducing the irrelevant anatomical structures within the lungs would greatly benefit any mass-screening program of lung cancers.

Thus I propose a method that seeks to improve both the accuracy of existing CAD methods and provide a method for identifying irrelevant anatomical structures based on local second order shape analysis. The key insight of the proposed method is to exploit the observation that nodules often appear more spherical and bloblike compared to irrelevant anatomical structure such as blood vessels, which tend to appear elongated and cylindrical. Further, these differences in shape can be automatically detected through second derivative analysis, but not through first derivative analysis. In doing so, the proposed method could provide a means by which existing methods could avoid false negatives generated from interfering irrelevant anatomy in the vicinity of a cancerous mass. Furthermore, the method could also be used to help radiologists by providing a CT scan with irrelevant anatomical structures diminished or removed, leading to more accurate single blind screenings.

Related works:

Typically, CAD methods attempt to distinguish the more bloblike lung nodules from the more elongated and cylindrical shaped blood vessels or more planar organ walls. Often this distinction is based on methods involving template-matching based on gradient intensities and orientation [8,9,10], identifying properties of an ellipsoid fit to the region of interest [7] or identifying properties of eccentricity and circularity of candidate regions [11].

However, these methods rely on approximations of shape based on information taken from the first derivative, which does not fully capture the local shape information. Furthermore, they do not necessarily distinguish if irrelevant anatomy is present in the volume of interest alongside ~~of~~ the potential lung nodule. As such, irrelevant anatomy may mislead many of these methods, leading to false negatives. Hence, a more accurate picture of the shape characteristics could be obtained with eigenanalysis of the second-order Hessian, which can be used to determine local shape.

Sato et al [12] showed that shape analysis obtained from eigenanalysis of the local image Hessian could be used to distinguish different types of tissue when rendering visualizations of scanned data. Mendonça et al. [13] developed a model-based curvature tensor analysis of local shape for lesion detection in CT scans. Mendonça et al chose to use the curvature tensor rather than using analysis of the image Hessian to avoid mischaracterizations that can occur at inflection points of an intensity profile. Some of these issues are apparent in methods that use averaged characteristics of the Hessian to form feature sets [14]. Similarly, Epstein et al [15] were able to detect colonic polyps based on local curvature features. Additionally Daniels et al [16], showed that eigenanalysis of the local image Hessian can be used to distinguish tube, blob and plane like structures within an organ. In particular, both of these methods were able to show a high degree of success in distinguishing between nodules and blood vessels or organ walls in the medical imaging of a colon. Hence, applying a similar method of eigenanalysis of the second-order image Hessian to lung CT scans could be used to identify what anatomical structure each voxel in the scan belongs to.

Proposed Research:

A. Description of method:

The proposed method will be implemented through the following stages: lung isolation, first and second order partial derivative estimation, shape classification, and finally nodule detection.

First, the lungs themselves must be isolated in the 3D CT scan. This can be accomplished using standard methods based on the distinguishable Hounsfield unit (a measure of pressure) value of voxels within the lung as compared to voxels belonging to anatomical structures such as fat or bones [17]. Thus, the 3D internal volume of the lungs consists of all voxels within this target Hounsfield value range, with small gaps in the volume filled in using a standard 3d morphological closure operator.

Next, the first and second order partial derivatives at each voxel in the CT scan must be computed. Typically, methods that only use the first-order partial derivatives, such as those detailed in section 2, use finite difference methods for their efficiency. However these methods can be unstable in medical image scans and are generally not suitable for second-order partial derivatives [18]. Instead, the proposed method will use deriche filters, which have shown promise in producing stable approximation of second-order partial derivatives in medical imaging [19].

Once the partial derivative information is generated, the method can begin to perform local shape analysis at each voxel. For each voxel in the internal volume of the lung, the method will generate a second-order Hessian matrix that can be analyzed using Eigen decomposition. From this decomposition, the three eigenvectors (principal directions) provide information on the direction of curvature of the local surface of the voxel, while the eigenvalues (principal values) describe the rate of curvature in each of these directions. Hence, if the local surface is cylindrical, then one eigenvalue should be large while the other two are relatively small, showing one preferred direction of curvature. If the local surface is planar, then there should be two preferred directions where one eigenvalues should be relatively small, while the other two are large. If the shape is bloblike than all three eigenvalues should be relatively large, showing a high degree of curvature in each of the principal directions.

Now that this shape information has been determined at each voxel, it can be used to improve existing methods or remove irrelevant anatomical structures from a CT scan. For methods that focus on identifying nodules by metrics computed on a volume of interest, the influence of voxels belonging to irrelevant anatomical structure could be reduced based on the previous local shape classification. Each align to the ideal blob shape. Similarly irrelevant anatomy can be filtered if the principal values of the voxels in a volume of interest show that local shape is cylindrical or planar. In doing so existing, algorithms will be able to ignore false negatives generated from irrelevant anatomy coincidental to a lung nodule in a volume of interest. Similarly, voxels belonging to an irrelevant anatomical structure can be diminished or removed from a scan presented to a radiologist for screening.

B. Evaluation of the method

To test the effectiveness of this method, existing lung nodule detection methods will be used as a baseline of performance. This baseline will then be compared to the detection results on data where the influence of voxels considered by each method has been determined in the previously described manner, based upon local shape characteristics. Each method will be trained upon the Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI) provided by the National Cancer Institute [20]. The LIDC/IDRI Database contains 1018 clinical cases, each composed of images from a clinical thoracic CT scan and the results of a two-phase image annotation process performed by four experienced thoracic radiologists. Each radiologist is first shown the CT scan, to annotate any suspicious masses. Then, each radiologist was asked to independently review the anonymized marks generated from each of the radiologists before rendering a final opinion, seeking to identify all possible lung nodules in each CT scan without requiring forced consensus. From this database, a set of positive samples, i.e volumes of interests containing a long nodule, and a set of negative samples, i.e a set of volumes of interest containing no lung nodules, can be generated. These two sets can then be divided into training, validation, and testing sets for which the aforementioned methods can be trained and evaluated upon. For the purposes of evaluation, a true positive will occur if an algorithm successfully detects a volume of interest containing a lung nodule as indicated by the radiologists and a false positive will occur if the algorithm believes there is a lung nodule within a volume of interest that the radiologists did not annotate. Similarly, a true negative will occur if the algorithm does not identify a lung nodule within a volume of interest that contains no annotations by the radiologists. However, if the algorithm identifies a lung nodule within a volume of interest that the radiologists did annotate as containing a suspicious mass, it will be considered a false negative. If the proposed method is successful, then we should expect to see a decrease in the number of false negatives and an increase in true positives, comparative to the rates seen in the baseline methods. Should the proposed method fail to demonstrate an improvement, or should it incidentally generate a statistically significant increase in false positives or false negatives, then it will be deemed unsuccessful.

Conclusion:

In this proposal, I have described an experimental method that can potentially address some of the major issues facing the implementation of a mass-screening program for lung cancer. While other methods of lung nodule detection exist, they are prone to generating false negative when dealing with a volume of interest that contains both a lung nodule and irrelevant anatomy. However, there exist methods in other medical imaging domains that have shown promise in separating different types of tissue depending on the local shape characteristics. The proposed method described previously will take advantage of these local shape characteristics to improve the detection rate of existing methods. Furthermore, the method is capable of providing a more easily interpreted scan to a radiologist, where factors that could typically lead to a false negative are mitigated. As the 5-year survival rate of patients whose diagnosis occurred in the first stage of the disease is far higher than those diagnosed in a later stage, it is crucial that detection rates and regular screenings of lung cancers be improved. Thus, if the proposed method can be successfully developed, it could provide vital benefits to a screening program that would address one of the most widespread and single deadliest forms of cancer that currently afflict the Canadian population.

References:

[1] Canadian Cancer Society's Advisory Committee on Cancer Statistics. *Canadian Cancer Statistics 2014*. Toronto, ON: Canadian Cancer Society; 2014. pp 28-39

[2] What are the key statistics about lung cancer? *The American Cancer Society*, 2015, http://www.cancer.org/cancer/lungcancer-non-smallcell/detailedguide/non-small-cell-lung-cancer-key-statistics

[3] WORLD CANCER REPORT 2008, *International Agency for Research on Cancer*, Lyon, France, WHO Press, World Health Organization 2008 pp 390-396

[4] Lung cancer survival statistics, *Cancer Research UK,* 2014, http://www.cancerresearchuk.org/cancer-info/cancerstats/types/lung/survival/lung-cancer-survival-statistics

[5] S., Sone, et al. "Results of Three-Year Mass Screening Programme for Lung Cancer Using Mobile Low-Dose Spiral Computed Tomography Scanner." *British Journal of Cancer* 84.1 (2001): 25.

[6] M. Morton, et al. "Screening Mammograms: Interpretation with Computer-aided Detection—Prospective Evaluation 1." *Radiology 239.2* (2006): 375-383.

[7] G. Zhanyu, et al. "Computer-Aided Detection Of Lung Nodules: False Positive Reduction Using a 3D Gradient Field Method and 3D Ellipsoid Fitting." *Medical physics 32.8* (2005): 2443-2454.

[8] M. Yang, et al. "False Positive Reduction in Lung GGO Nodule Detection with 3d Volume Shape Descriptor." *Acoustics, Speech and Signal Processing*, 2007. ICASSP 2007. *IEEE International Conference on. Vol. 1. IEEE*, 2007.

[9] T. Maxine, et al. "A Novel Computer-Aided Lung Nodule Detection System for CT Images." *Medical physics* 38.10 (2011): 5630-5645.

[10] B. Zhao, et al. "Automatic Detection Of Small Lung Nodules On CT Utilizing A Local Density Maximum Algorithm". *Journal of Applied Clinical Medical Physics*, [S.l.], p. 248 - 260, jun. 2003. ISSN 15269914.

[11] M. Alilou, et al. "A Comprehensive Framework for Automatic Detection of Pulmonary Nodules in Lung Ct Images." *Image Analysis & Stereology 33.1* (2014): 13-27.

[12] S. Yoshinobu, et al. "Tissue Classification Based On 3D Local Intensity Structures for Volume Rendering." *Visualization and Computer Graphics, IEEE Transactions on 6.2* (2000): 160-180.

[13] P. Mendonça, et al. "Model-Based Analysis of Local Shape for Lesion Detection in CT Scans." *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005. Springer Berlin Heidelberg*, 2005. 688-695.

[14] B. Sahiner, et al. "False-positive reduction using Hessian features in computer-aided detection of pulmonary nodules on thoracic CT images."*Medical Imaging. International Society for Optics and Photonics*, 2005.

[15] M. Epstein, et al. "Hessian Matrix-Based Shape Extraction and Volume Growing for 3D Polyp Segmentation in CT Colonography", *CT Colonography*. *INTECH Open Access Publisher,* 2010.

[16] D Florie, et al. "Quantification of Collagen Orientation in 3D Engineered Tissue." *3rd Kuala Lumpur International Conference on Biomedical Engineering 2006. Springer Berlin Heidelberg*, 2007.

[17] Walter O'Dell. "Automatic Segmentation of Tumor-Laden Lung Volumes from the LIDC Database." *SPIE Medical Imaging. International Society for Optics and Photonics*, 2012.

[18] R. Deriche, et al. "3D Edge Detection Using Recursive Filtering: Application to Scanner Images." *CVGIP: Image Understanding 53.1* (1991): 76-87.

[19] O. Monga and S. Benayoun. "Using Partial Derivatives of 3D Images to Extract Typical Surface Features." *Computer vision and image understanding61.2* (1995): 171-189.

[20] S. Armato III, et al. "The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans." Medical physics 38.2 (2011): 915-931.