The Use of Low-Dose CT Intra- and Extra-Nodular Image Texture Features to Improve Small Lung Nodule Diagnosis in Lung Cancer Screening

Rongkai Yan[1], *IEEE Member,* Saeed Ashrafinia[2,3], *IEEE Member,* Seyoun Park[4], Junghoon Lee[4], Linda C. Chu[5], Cheng Ting Lin[5], Amira Hussien[5], Nagina Malguria[5], Jon Steingrimsson[6], Arman Rahmim[2,3], *IEEE Senior Member,* Peng Huang[1, 6*]

***Abstract*–Standard computed tomography (CT) scan is performed on lung cancer patients for progression and lesion classification. However, low-dose CT (LDCT) is commonly used in lung cancer screening for high-risk people. Extensive studies have shown that computer-aided diagnosis (CAD) using standard CT could greatly improve the diagnostic accuracy of early lung cancer. Unlike standard CT imaging, the application of radiological texture features extracted by radiologists on LDCT imaging is not well established due to lower resolution and higher variations. The purpose of this study is to investigate possible diagnosis value of texture features by comparing the classification performance of radiologic reading with radiologic reading combined with computer-aided texture features. A total of 186 biopsy-confirmed control and lung cancer cases were obtained from the National Lung Screening Trial (NLST). Cases were matched by various clinical parameters including age, gender, smoking status, chronic obstructive pulmonary disease (COPD) status, body mass index (BMI) and image appearances. We compared the subjective diagnosis of benign/malignant with the consensus readings of three radiologists. We then developed a CAD framework that imports radiologic reading features and extracts CAD features for heterogeneity quantification and data analysis. A total of 1342 CAD features were extracted. After eliminating highly correlated and redundant features, the remaining 458 features were given to a random forest algorithm, and a predicted probability of malignancy score (Pm) was calculated. Patients were grouped into 140 training (70 biopsy-positive for cancer and 70 negatives) and 46 testing (20 positives and 26 negatives) sets, and a threshold value over Pm (0.5) was then used to classify the test set into cancer and non-cancer. Clinical accuracy [sensitivity, specificity, positive predictive value (PPV), and negative predictive value (PV)] were [0.95, 0.88, 0.86, 0.96] and [0.70, 0.69, 0.64, 0.75] for the CAD and radiologic reading, respectively. The CAD framework incorporating the clinical reading with the texture features extracted from LDCT increased the PPV and reduced the false positive (FP) rate in the early diagnosis of lung cancer. This approach shows promise for improving the accuracy of lung cancer diagnosis in the clinical environment, especially in patients with well-established risk factors.**

[1]Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, MD

[2]Department of Electrical and Computer Engineering, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD

[3]Division of Nuclear Medicine and Molecular Imaging, Department of Radiology and Radiological Sciences, Johns Hopkins University School of Medicine, Baltimore, MD

[4]Radiation Oncology and Molecular Radiation Sciences, Johns Hopkins University School of Medicine, Baltimore, MD

[5]Department of Radiology and Radiological Sciences, Johns Hopkins University School of Medicine, Baltimore, MD

[6]Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD

[*]Corresponding author's email: phuang12@jhmu.edu

I. Introduction

Within the past decade, the incidence of lung cancer had grown due to environmental factors and increased detection partially due to the use of low-dose computed tomography (LDCT) screening. The National Lung Screening Trial (NLST) demonstrated a 20% reduction in lung cancer mortality with LDCT screening [1], but the radiologic diagnosis of lung cancer remains challenging when nodules are small (<20mm). Benign lung nodules are difficult to distinguish from early lung cancers, and false positive cases frequently lead to invasive procedures putting patients at risk of untoward complications [2][3]. Radiologists have put significant efforts towards quantitative methods to improve the accuracy of lung cancer diagnosis. Extensive studies have shown that computer-aided diagnosis (CAD) could greatly improve the diagnostic accuracy [4]. However, there are limited studies – to our knowledge - that diagnose lung cancer by analyzing the texture features extracted from LDCT images. A possible explanation is that LDCT images have generally lower resolution and calculated features exhibit higher variation [5].

II. Methods

We randomly selected the matched-control sample of 186 participants with non-calcified lung nodules that were 4mm-20mm in size and underwent a biopsy in the NLST. Variables used for matching were age, gender, smoking status, COPD status, BMI, study year of the positive screen, and screening results. LDCT scans prior to lung biopsy were randomly split into training (70 cancers+70 benign controls) and validation (20 cancers + 26 benign controls) sets. Three radiologists read the images for consensus regarding the three dominant nodule locations using the same clinical radiologic software. The nodule segmentation was semi-automatically drawn with intra-

nodular, peri-nodular, and extra-nodular area using our in-house software [6]. Our radiologists extracted 18 radiologic features from which they have made the diagnosis for each patient. Then these features were combined with computer-generated 1342 features in the machine learning analysis. We also performed reliability analysis of the radiologist's features using 20 randomly selected images. Kappa statistic was used to evaluate interrater reliability.

A. Radiologic Features

LDCT scans in the validation set were reviewed by three radiologists who were blinded to biopsy outcomes. These human readings were used to compare with computer's reading to evaluate the added value of CAD. Radiologic features include size, location, density, shape, margin, calcification, and attenuation. The features below are commonly used in radiologist field. The primary radiologic criteria for nodule benignity was size with a threshold of 6mm because the majority of benign nodules are less than 6mm according to 2016 NCCN guideline version [7]. However, a smaller size nodule does not exclude malignancy [8]. The margin of the nodule was described as smooth, spiculated, or lobulated. Lobulation and spiculation were common and occurred at a significantly higher frequency among malignant nodules [9]. The pleural tag is a linear area of high attenuation surrounded by aerated lung, originating from the edge of the mass and extending peripherally to contact the pleural surface. This feature indicates that the tumor invades the fissure or pleural [10]. The attachment to the vessel is a sign that tumor nutrition comes from the vessel. Nodule homogeneity is a feature that is found more often in benign lesions than lung cancer. Calcifications within nodules typically represent dystrophic calcium from necrotic tissue and favor a benign diagnosis, although eccentric and stippled calcification patterns are indeterminate for malignancy [11]. A cancer-like feature from radiologist reading was defined into three scores: a score of 1 is negative for suspicious features and considered benign without any clinical significance; a score of 2 is a suspicious nodule with indeterminate diagnosis, and a score of 3 is a positive nodule suspicious for lung cancer.

B. Computational Features

Our in-house texture analysis software was used to extract a total of 1342 texture features. Among them, 1108 features were generated from Radiomics features by using various parameter combinations in the gray-level co-occurrence matrix, gray-level run-length matrix, and wavelet filters. The number of features was then reduced by feature selection. First, features with high correlation (R>0.95) were identified, one of which was selected as a representative and others were eliminated. Then, features with zero variance were subsequently eliminated. The remaining 458 features were feed into a random forest algorithm comprising 5000 trees using the training data only. This computation identified 38 top CAD features and derived a predicted probability of malignancy score (Pm). A cutoff value of 0.5 was specified to analyze the training set. Texture features were then extracted from the test set images to determine their malignancy (Pm) scores and subsequently their positive or negative image-based predictions. One of the 38 features identified is the Cp feature calculated through formula ($A^3/36\pi V^2$) where A is the sum of surface areas of all small clusters with high CT intensity inside the nodule and V is the sum of volumes of these clusters. Another CAD feature d2m2 was calculated as the mean of the second derivatives in the nodule margin area ROC analysis was used to compare the accuracy of CAD prediction, NLST investigator's diagnosis, and readings from three experienced board-certified thoracic radiologists.

III. RESULTS

Malignant and benign cases were well balanced in demographics, diameter, margin and attenuation of dominant nodules, total number and mean diameter of spiculated solid non-calcified nodules. The valuable from radiologist reading score was 0.70. Malignant cases had significantly higher Cp than benign nodules (Wilcoxon rank test p=0.005, figure 1). The Cp was also found to be positively correlated with the percent of tumor cells (Spearman correlation R=0.2446, pp=0.0169). The malignant tumors tend to have the lower d2m2 feature than benign features although the difference is not statistically significant (p=0.27, Wilcoxon rank test). However, it was found to be negatively correlated with the percent of tumor cells (Spearman R=-0.2079, p=0.0423) and positively correlated the inflammatory cells (R=0.2673, p=0.0096).

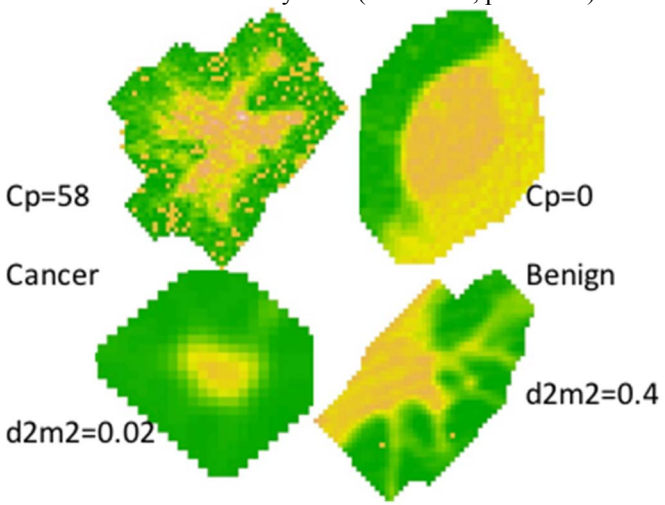

Figure 1. Cp feature: An example showing a malignant nodule on the left and a benign nodule on the right with their Cp value.

Figure 2a shows a right lower lobe 8-mm nodule with a smooth margin, without suspicious features (e.g. spiculation, lobulation, vessel involvement, bubble-like lucency). However, the Cp score was 21—a value that is significantly predictive of malignancy, and was confirmed through biopsy. Figure 2b shows a right lower lobe 18-mm nodule with halo sign, spiculation, lobulation, and connecting vessel that was visually diagnosed to be a malignancy. Our CAD found the Cp score to be 1.273 which was more consistent with the benign nodule.

In the validation set, Pm scores from malignant cases were significantly higher than benign nodules (p<0.0001) as can be seen in figure 3.

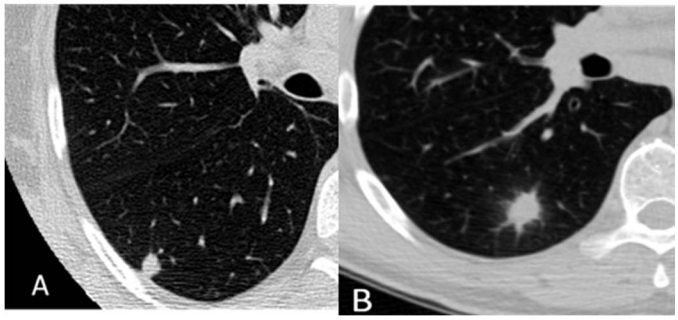

Figure 2. CT image presentation: A) A 8-mm biopsy-proven lung cancer on computed tomography scan. The nodule is oval and without spiculation, lobulation, the vessel involved, or bubble-like lucency. B) An 18-mm benign nodule with halo sign, spiculation, lobulation, and big vessel connected.

The calculated values of [sensitivity, specificity, PPV, NPV] from CAD and the radiologists' consensus readings were [0.95, 0.88, 0.86, 0.96] and [0.70, 0.69, 0.64, 0.75], respectively (figure 4). In the reliability study of 18 radiologist's rating scales using 20 randomly selected images, features with good reproducibility are: the presence of fat (kappa=0.96), calcification (kappa=0.89), GGO indicator (kappa=0.85). Features with poor reproducibility are: ragged/microlobulation (kappa=0.20), vessel abnormality (kappa=0.21), pleural tag (kappa=0.24), and tumor disappearance rate (kappa=0.24). The cancer-like scale from radiologist reading was not included in the reliability study because all radiologists agreed that it is difficult to reach a consensus among radiologists on this reading.

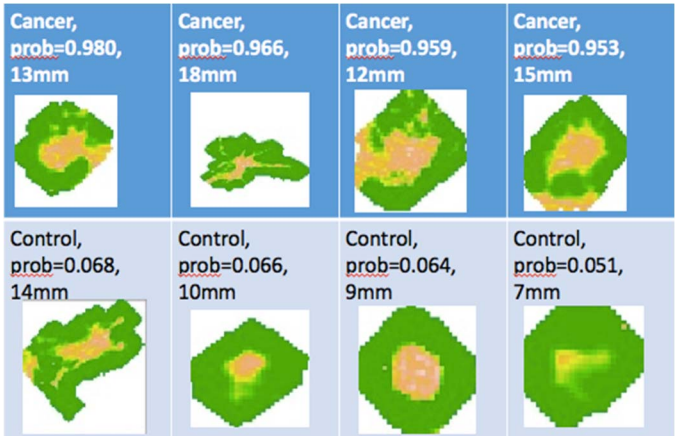

Figure 3. Nodules with highest and lowest predicted probabilities of malignancy.

IV. Discussion

Commonly used radiologic criteria for lung cancer diagnosis may lead to a high false-positive rate when nodules are small. While NLST demonstrated a lung cancer mortality decrease of 20%, false-positive results accounted for 96.4% of positive cases over 3 rounds of LDCT screening[12]. The downside is that most of these positive nodules turn out to be indolent and clinically insignificant. We improved upon the radiologists' diagnostic accuracy by incorporating a CAD framework that extracts consider radiologic features as well as radiomic features that cannot be detected visually.

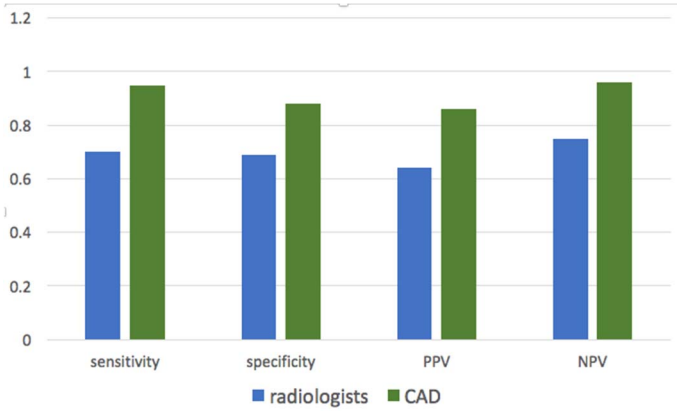

Figure 4. Comparison of diagnostic accuracy between radiologist and CAD in the validation set

The Cp score is the primary feature of the CT image that was found significantly correlated with the probability of malignancy, independent of nodule size, spiculation, lobulation, or pleural tag. A higher Cp score for a nodule means a low attenuation diffusion function. Lower Cp means that the attenuation inside the nodule is smooth and homogeneous. The improved diagnostic performance due to independent information from new image texture features and higher prediction accuracy on the independent validation data from Pm score compared to three radiologists' reading increased the PPV and reduced the false positive (FP) rate in the early diagnosis of lung cancer. Adoption of this approach into routine clinical practice will require confirmation of these results with additional prospective randomized studies using participants from LDCT screening programs.

V. Conclusion

This study demonstrated that the LDCT image texture analysis combined with radiologist's interpretations could increase the PPV rate and reduce the FP rate in the early diagnosis of lung cancer. This approach shows promise for improving the accuracy of lung cancer diagnosis in the clinical environment, especially in patients with well-established risk factors. Adoption of this approach into routine clinical practice will require confirmation of these results with additional prospective randomized studies using participants from LDCT screening programs.

VI. Acknowledgment

The authors thank the NCI for access to the NCI data collected by the NLST. This work was sponsored by Johns Hopkins-Allegheny Health Network Cancer Research funding and Johns Hopkins University (2015 Discovery Award).

References

[1] D. R. Aberle *et al.*, "Reduced Lung Cancer Mortality with Low-Dose Computed Tomographic Screening.," *N. Engl. J. Med.*, vol. 365, no. 5, pp. 395–409, 2011.

[2] P. J. Mahadevia, L. A. Fleisher, K. D. Frick, J. Eng, S. N. Goodman, and N. R. Powe, "Lung cancer screening with helical computed tomography in older adult smokers: a decision and cost-effectiveness analysis," *JAMA*, vol. 289, no. 3, pp. 313–322, 2003.

[3] P. Nanavaty, M. S. Alvarez, and W. Michael Alberts, "Lung Cancer

Screening: Advantages, controversies, and applications," *Cancer Control*, vol. 21, no. 1, pp. 9–14, 2014.

[4] D. Chen, P. Huang, and X. Cheng, "A concrete statistical realization of Kleinberg's stochastic discrimination for pattern recognition. Part I. Two-class classification," *Ann. Stat.*, vol. 31, no. 5, pp. 1393–1412, 2003.

[5] P. B. Bach *et al.*, "Benefits and harms of CT screening for lung cancer: a systematic review.," *JAMA*, vol. 307, no. 22, pp. 2418–29, 2012.

[6] Y. Gu *et al.*, "Automated delineation of lung tumors from CT images: Method and evaluation," *Mol. Imaging Biol.*, vol. 14, p. S349, 2012.

[7] D. E. Wood, "National Comprehensive Cancer Network (NCCN) Clinical Practice Guidelines for Lung Cancer Screening," *Thoracic Surgery Clinics*, vol. 25, no. 2. pp. 185–197, 2015.

[8] D. R. Aberle, F. Abtin, and K. Brown, "Computed tomography screening for lung cancer: has it finally arrived? Implications of the national lung screening trial.," *J. Clin. Oncol.*, vol. 31, no. 8, pp. 1002–8, 2013.

[9] A. O. Soubani, "The evaluation and management of the solitary pulmonary nodule," *Postgrad. Med. J.*, vol. 84, no. 995, pp. 459–466, 2008.

[10] C. V Zwirewich, S. Vedal, R. R. Miller, and N. L. Muller, "Solitary pulmonary nodule: high-resolution CT and radiologic-pathologic correlation.," *Radiology*, vol. 179, no. 2, pp. 469–476, 1991.

[11] A. Khan, H. Al-Jahdali, C. Allen, K. Irion, S. Al Ghanem, and S. Koteyar, "The calcified lung nodule: What does it mean?," *Ann. Thorac. Med.*, vol. 5, no. 2, p. 67, 2010.

[12] H. J. W. L. Aerts *et al.*, "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach," *Nat. Commun.*, vol. 5, 2014.