

Intra-perinodular Textural Transition (Ipris): A 3D Descriptor for Nodule Diagnosis on Lung CT

Mehdi Alilou^(✉), Mahdi Orooji, and Anant Madabhushi

Center for Computational Imaging and Personalized Diagnostics,
Case Western Reserve University, 2071 Martin Luther King Drive,
Cleveland, OH 44106, USA
me.alilou@gmail.com

Abstract. This paper presents Ipris (Intra-perinodular textural transition), a new radiomic method, to automatically distinguish between benign and malignant nodules on routine lung CT scans. Ipris represents a minimal set of quantitative measurements which attempt to capture the transition in textural appearance going from the inside to the outside of the nodule. Briefly the approach involves partitioning the 3D volume and interface of the nodule into K nested shells. Then, a set of 48 Ipris features from 2D slices of the shells are extracted. The features pertain to the spiculations, intensity and gradient sharpness obtained from intensity differences between inner and outer voxels of an interface voxel. The Ipris features were used to train a support vector machine classifier in order to distinguish between benign (granulomas) from malignant (adenocarcinomas) nodules on non-contrast CT scans. We used CT scans of 290 patients from multiple institutions, one cohort for training ($N = 145$) and the other ($N = 145$) for independent validation. Independent validation of the Ipris approach yielded an AUC of 0.83 whereas, the established textural and shape radiomic features yielded a corresponding AUC of 0.75, while the AUCs for two human experts (1 pulmonologist, 1 radiologist) yielded corresponding AUCs of 0.69 and 0.73.

1 Introduction

Lung nodules characterization on CT images is a challenging clinical task. Granulomas which are benign presentations but appear as malignant nodules on chest CT scans are considered among the most difficult tumor confounders to discern. Granulomas and adenocarcinomas (malignant lesions) are often indistinguishable on both CT and PET scans, both nodules appearing “hot” on PET. Hence, many people with benign nodules are subjected to unnecessary surgical procedures due to the inability in making confident diagnostic predictions with respect to the nodule on CT. Consequently there is a need for discriminating radiomic features for improved characterization of lung nodules on CT scans.

Recently, there has been substantial interest in the use of textural, intensity or shape analysis [1] for radiomic [5] characterization of lung nodules. For example, in [2] the authors analyzed shape (e.g. surface area, volume and surface to

volume ratio) features together with textural and intensity features extracted from CT data of lung and oropharyngeal cancers. Applying unsupervised clustering to the extracted image features, they found an association with underlying gene-expression profiles of lung cancer patients. Way et al. [7] developed an automated 3D active contour segmentation method and extracted morphological and textural features from the segmented nodules. A leave-one out cross validation of their method yielded an AUC of 0.83 in a data set of 44 malignant and 52 benign nodules. In a related study Shah et al. [6] achieved AUC values between 0.68 and 0.92 with 48 malignant and 33 benign nodules. One common attribute associated with the majority of these previous radiomic related approaches for lung nodule characterization is the fact that they involve features pertaining to the nodule alone. Additionally recent studies [4] have suggested that texture based features tend to be affected by the choice of scanner, reconstruction kernel, and slice thickness.

Recently there has been a growing appreciation of the role and prevalence of lymphocytic infiltration associated with malignant lung nodules [3]. The infiltration appears within the perinodular space of malignant nodules, which in turn appears to cause differential textural patterns adjacent to the nodule. Lymphocytic infiltration does not typically co-occur with granulomas and benign nodules however. Hence by quantitatively capturing the textural attributes from the interior through to the exterior of the nodule, one might possibly distinguish malignant from benign lung nodules. Some attempts have been proposed towards this end through the concept of margin sharpness. One version of the margin sharpness descriptor was presented in [8] where the sharpness of the intensity transition across the lesion boundary was calculated. The feature was employed for content based image retrieval applications for liver and lung scans.

However, the previous margin sharpness approaches tend to limit the interrogation of intensity changes to the nodule interface rather than interrogating the whole core of the tumor. The ability to interrogate not just the margin transition, but the textural patterns as they change from within the core of the nodule all the way to outside to nodule could provide for a more comprehensive portrait of the heterogeneity within the nodule.

2 Brief Overview and Novel Contributions

This paper presents Ipris, a radiomic approach, to automatically distinguish between benign and malignant nodules on routine lung CT scans. Ipris represents a minimal set of quantitative measurements which attempt to capture the transitional heterogeneity from the intra- to the peri-nodular space. Briefly the steps of Ipris computation involve segmentation of the nodule volume across contiguous slices. Then, the 3D interface of the nodule was partitioned automatically into K nested shells using morphological operations. Next, a set of 48 Ipris features from 2D slices of the shells are extracted. The most predictive Ipris features are then identified using feature selection and then used to train a support vector machine classifier in order to distinguish between benign and malignant nodules.

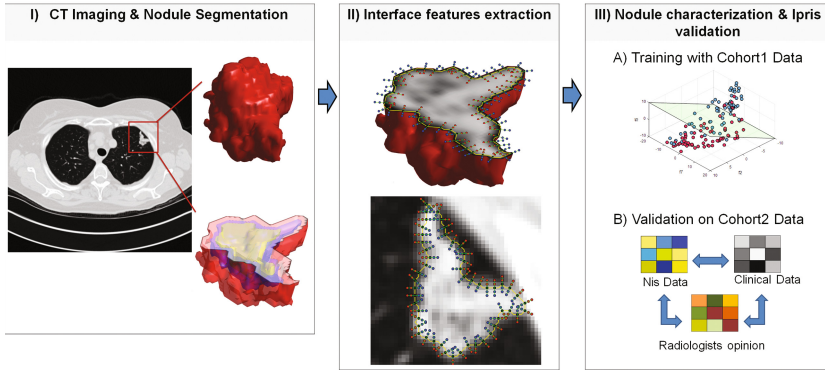


Fig. 1. The main work flow of Ipris includes, (I) manual segmentation of the nodules by radiologists, partitioning the nodules into k shells, (II) extraction of 3D nodules interface features from shells and (III.a) Classification of the nodules into benign and malignant categories by a classifier trained with Ipris features. (III.b) Evaluation of the Ipris by comparing its performance against texture based radiomic features and expert human readers on an independent test set, distinct from the training set used to constructing the Ipris classifier.

The novel contributions of our work are as follows. To the best of our knowledge, Ipris is the first attempt to quantitatively capture and evaluate the role of transitional heterogeneity from the intra- to the peri-nodular space to discriminate between granulomas (benign) and malignant nodules on lung CT images. Unlike previous related attempts at capturing descriptors of margin sharpness, is not limited to evaluating heterogeneity patterns solely at the nodule surface. In fact Ipris is a 3D feature that attempts to capture the intensity and gradient transitions as well as a minimal set of textural speculation related statistics, from the inner core of the lesion all the way to the peri-nodular space, immediately outside the nodule boundary. Ipris involves partitioning the internal volume of the nodule into annular, nested shells in order to systematically capture heterogeneity, irregularities and differences in growth patterns between malignant and benign nodules.

The robustness of Ipris was examined via the independent validation and also ensuring that the training and testing sets comprised of cases from multiple different sites. Additionally, our approach was evaluated on CT scans with slice thicknesses that varied from 1 to 5 mm. We perform an exhaustive and rigorous evaluation of our approach on a cohort of 290 patients divided randomly into training and testing sets, with extensive human-machine comparison studies involving two different human readers. Additionally, Ipris was also quantitatively compared on the test set against the performance of the state of the art texture features for lung nodule characterization. Figure 1 summarizes the main work flow of our method.

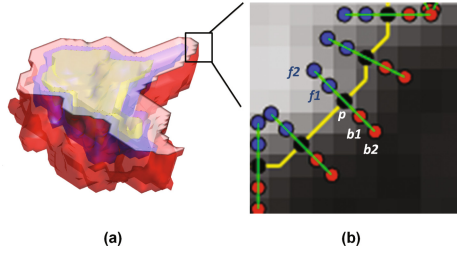


Fig. 2. (a) A nodule partitioned into three shells. (b) 2D representation of the outer shell’s border pixels and their corresponding normal lines including inner (f_i) and outer (b_i) pixels.

3 Methodology

3.1 Shell Definition

Let $\Gamma = \{1, \dots, H\} \times \{1, \dots, W\} \times \{1, \dots, D\}$ be a three-dimensional image lattice and v be the binary volume of a nodule defined as $v : \Gamma \rightarrow \{0, 1\}$. The nodules volume is partitioned into k shells such that $v = \{s_1, \dots, s_k\}$, $s_{i-1} \subset s_i$ and $\bigcup_{i=1}^k s_i = v$. Figure 2.a illustrates shells of a nodule.

Each s is consisted of n 2D slices (layers) $s_i = \{l_1, \dots, l_n\}$ and each 2D slice consisted of j boundary pixels $l_i = \{p_1, \dots, p_j\}$. The slope of normal at a boundary pixel $p_i = (x, y)$, is computed using the co-ordinates of two adjacent pixels $p_{i-1} = (x_{i-1}, y_{i-1}), p_{i+1} = (x_{i+1}, y_{i+1})$ of it and m_{p_i} is defined as:

$$m_{p_i} = \frac{1}{\tan^{-1}\left(\frac{y_{i-1}-y_{i+1}}{x_{i-1}-x_{i+1}}\right)}. \tag{1}$$

The normal line at a boundary pixel p_i is then divided into foreground (f) and background (b) pixels.

3.2 Ipris Features

The average gradient difference of every p_i is then computed based on gradient values over f and b via:

$$dG_{p_i} = \frac{1}{Q} \sum_{q=1}^Q \frac{\nabla f_q - \nabla b_q}{2q}, \tag{2}$$

where, $Q = \frac{R}{2} - 1$ when R is the number of pixels sampled over the normal line of pixel p_i and $\nabla f_r, \nabla b_r$ are the gradient magnitude values of foreground and background pixels along the normal line. Accordingly, the intensity difference profile dI_{p_i} at pixel p_i is calculated based on Eq. 4 by plugging the intensity- instead of gradient values. In addition to dG_{p_i} and dI_{p_i} , the average gradient sharpness at pixel p_i is defined as:

$$aG_{p_i} = \frac{1}{R} \sum_{r=1}^R M_r, \quad (3)$$

where, M_r is the gradient magnitude value of the r th sample over the normal line. Similarly, the entropy of the gradient magnitudes over the p_i s normal line is calculated via:

$$\varepsilon_{p_i} = \sum_{r=1}^R M_r \log_2 M_r, \quad (4)$$

Finally, for each shell $s \in v$ we calculated the mean, standard deviation, minimum and maximum of the dG_{p_i} , dI_{p_i} , aG_{p_i} , ε_{p_i} according to the constituent border pixels $p \in l$ of the 2D slices $l \in s$ of a shell.

4 Experimental Results and Discussion

4.1 Data Description and Pre-processing

Our study comprised of CT scans of 290 patients from multiple institutions. The data set of 290 patients divided randomly into training and testing sets. One cohort for training (N = 145) and the other (N = 145) for independent validation. Both of the cohorts consisted of 73 malignant and 72 benign nodules. All patients had previously undergone surgical wedge resection for a suspicious nodule and have a histopathologically confirmed diagnosis for having a malignant or benign nodule. The number of slices per scan ranged from 126 to 385 and slice thickness of the CT scans ranged from 1–6 mm. Each slice had a XY planar resolution of 512×512 pixels with a 16 bit gray scale resolution in Hounsfield Units (HU). The volume of interest containing the nodules was manually segmented across contiguous slices by an expert cardiothoracic radiologist with 20 years of experience in interpreting chest CT scans, via a hand-annotation tool in 3D-Slicer[®] software.

4.2 Experiment 1: Evaluating Ipris via Support Vector Machine Classifiers

The most informative Ipris features was selected and ranked using Minimum redundancy, Maximum Relevance (mRMR) feature selection algorithm. The 3 most informative features were found to be (1) gray profile of the second shell (dI_{s_2}), (2) entropy of the gradient magnitudes of the outer shell (ε_{s_1}) and (3) the mean gradient sharpness of the outer shell (aG_{s_1}). Note that, the most informative feature was found to correspond to the second outer shell. This appears to strongly suggest that interface features should not be limited to the margin itself and it is critical to interrogate the textural transition from within the core of the nodule to the peri-nodular space and not just focus on the nodule surface. The top ranked Ipris features were used to train a Support Vector

Machine (SVM) classifier in order to distinguish benign from malignant nodules on non-contrast CT scans. To ensure a completely independent validation, we trained the SVM classifier using data from D_{train} and validated the classifier on a different D_{test} . The linear kernel was empirically determined for training the SVM using D_{train} . The SVM classifier yielded an AUC=0.83 in conjunction with the top ranked Ipris features on the independent validation set ($n = 145$ scans).

4.3 Experiment 2: Evaluating Ipris Against Established Textural and Shape Features

We compared the Ipris based classifier C_{Ipris} with a classifier trained with the established textural and shape radiomic features C_{rad} . In this regard, a total of 669 radiomic features including 645 2D texture and intensity along with 24 3D shape features were extracted from the volume of interest (i.e., nodule area). The texture features consisted of local binary pattern, gradient, Gabor filter, Laws-Laplacian pyramids, Laws and Haralick features. The shape features include some of the geometrical properties of the nodules such as size, compactness, eccentricity, elongation, convexity and sphericity.

While C_{Ipris} yielded an AUC of 0.83 on the validation set, C_{rad} yielded a corresponding AUC of 0.75. Moreover the mean Ipris feature extraction run time per one scan was 0.7s, while it was 1.2s for the radiomic features.

Ipris outperformed well established shape and textural features from both the perspective of AUC and computational efficiency on the validation set. Additionally, we found that in the several cases, shape and texture radiomics fail to correctly discriminate the nodules owing to overlapping shape and textural appearance. Interestingly, Ipris worked well in these cases. As an example, Fig. 3 illustrates a 4 different CT scans including two malignant and two benign nodules (a, b, c, d). Two instances of the corresponding textural feature maps including Haralick correlation and entropy are shown in (e, f, g, h) and (i, j, k, l) panels. The corresponding Ipris feature vectors are shown in the form of a bar plot in (m, n, o, p) panels. Note that, the height of each column of the bar plots is a reflection of the Ipris feature value. As may be appreciable from Fig. 3(m, n, o, p), unlike the texture maps which appear almost identical for both nodules, Ipris appears to significantly over-express for the benign nodule and appears to be largely under-expressing for the malignant nodule.

4.4 Experiment 3: Human-Machine Comparison

In order to perform human machine comparison, the classification performance of C_{Ipris} was compared against the diagnosis of two human experts. A board certified attending radiologist with seven years of experience in thoracic radiology and a pulmonologist with 7 years of experience in reading chest CT scans served as Readers 1 and 2 respectively. Both readers were blinded to the true histopathologic diagnosis of the validation set. Each reader was asked to assign a score between 1 to 5 to each nodule with 1 referring to a diagnosis of to “benign”

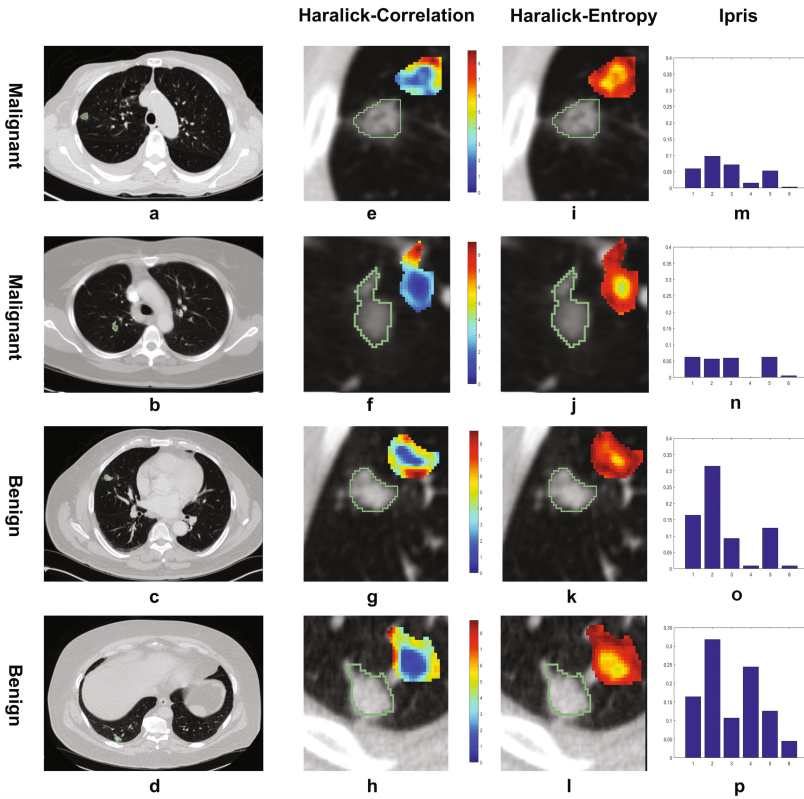


Fig. 3. Four CT scans including two malignant and two benign nodules (a, b, c, d). Two instances of the corresponding textural feature maps including Haralick correlation and entropy (e, f, g, h) and (i, j, k, l). The corresponding Ipris feature vectors (m, n, o, p). Note that, the height of each column of the bar plots is a resection of a Ipris feature value. Unlike the texture maps which appear almost identical for both nodules, Ipris appears to significantly over-express for the benign nodule and appears to be largely under-expressing for the malignant nodule

2 referring “mostly benign”, 3 being “not sure”, 4 being “mostly malignant” and 5 being “malignant”. To help human readers, we have developed a scoring tool which enables readers to preview the CT scans, identify the target nodules and enter their scores. AUC values were computed based on a hard decision using a threshold of >3 and >0.5 respectively for the human readers and the machine classifier. CT based diagnosis of the two human readers including on the validation set were found to be 0.69 and 0.73 respectively. Interestingly C_{Ipris} outperformed both human readers.

5 Conclusion

In this work, we presented Ipris, a novel radiomic method, to automatically distinguish between benign and malignant nodules on routine lung CT scans. Ipris attempts to capture the transitional heterogeneity from the intra- to the perinodular space and exploits the fact that the transitional patterns may be substantially different between benign and malignant nodules on CT scans. On an independent validation set, Ipris was compared against well established radiomic features and interpretations of two human readers. Ipris yielded a better performance compared to established radiomic features in terms of both classification AUC computational efficiency. Significantly, Ipris also was found to perform substantially better compared to two human expert readers, a pulmonologist and a thoracic radiologist with 7 years of experience reading chest CT scans. Additionally, even though this was not explicitly evaluated, Ipris appears to be robust to the slice thickness of the CT scans, since the datasets considered in this work involved 1–5 mm thickness. Future work will involve integrating Ipris with established radiomic features to further improve classification performance.

Acknowledgments. Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under award numbers 1U24CA199374-01, R01CA202752-01A1 R01CA208236-01A1 R21CA179327-01; R21CA195152-01 the National Institute of Diabetes and Digestive and Kidney Diseases under award number R01DK098503-02, National Center for Research Resources under award number 1 C06 RR12463-01 the DOD Prostate Cancer Synergistic Idea Development Award (PC120857); the DOD Lung Cancer Idea Development New Investigator Award (LC130463), the DOD Prostate Cancer Idea Development Award; the DOD Peer Reviewed Cancer Research Program W81XWH-16-1-0329, W81XWH-15-1-0613, the Case Comprehensive Cancer Center Pilot Grant VelaSano Grant from the Cleveland Clinic the Wallace H. Coulter Foundation Program in the Department of Biomedical Engineering at Case Western Reserve University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Alilou, M., et al.: An integrated segmentation and shape based classification scheme for distinguishing adenocarcinomas from granulomas on lung CT. *Med. Phys.* (2017). doi:[10.1002/mp.12208](https://doi.org/10.1002/mp.12208)
2. Aerts, H., et al.: Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature Commun.* **5**, 4006 (2014)
3. Brambilla, E., et al.: Prognostic effect of tumor lymphocytic infiltration in resectable non-small-cell lung cancer. *J. Clin. Oncol.* **34**(11), 1223–30 (2016)
4. He, L., Huang, Y., Ma, Z., Liang, C., Liang, C., Liu, Z.: Effects of contrast-enhancement, reconstruction slice thickness and convolution kernel on the diagnostic performance of radiomics signature in solitary pulmonary nodule. *Sci. Rep.* **6**, 34921 (2016)
5. Lambin, P., et al.: Radiomics: extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer* **48**, 441–446 (2012)

6. Shah, S.K., et al.: Computer-aided diagnosis of the solitary pulmonary nodule. *Acad. Radiol.* **12**, 570–575 (2005)
7. Way, T.W., et al.: Computer-aided diagnosis of pulmonary nodules on CT scans: segmentation and classification using 3D active contours. *Med. Phys.* **33**, 2323–2337 (2006)
8. Xu, J., Napel, S., Greenspan, H., Beaulieu, C.F., Agrawal, N., Rubin, D.: Quantifying the margin sharpness of lesions on radiological images for content-based image retrieval. *Med. Phys.* **39**(9), 5405–5418 (2012)