IOP Publishing

Lung Cancer and Imaging

Ayman El-Baz and Jasjit S Suri

Chapter 3

Quantitative malignancy recognition of lung cancer using non-invasive image modalities

Chung-Ming Lo

3.1 Introduction

Currently, lung cancer is the most common cancer worldwide. Although various medications and therapies have been developed, the death rate is still high [1, 2]. Like in other cancers, a possible reason for this phenomenon is the high prevalence of late-stage cancers. That is, the cancers are not detected earlier. One way to reduce the late detection is lung cancer screening using low-dose computed tomography to find early-stage peripheral lung cancers [3]. However, computed tomography cannot detect endobronchial lesions. Therefore, bronchoscopic examination plays a necessary role as an imaging modality in diagnosing endobronchial/tracheal lesions [4, 5].

About 85% of lung cancers are non-small-cell lung cancer (NSCLC) [4, 5]. Among these, 80% are adenocarcinomas (ACs) and squamous cell carcinomas (SCCs). The differentiation by type in the clinic leads to the following different treatment strategies. Patients with advanced lung ACs tend to have mutations of the epidermal growth factor receptor (EGFR) which respond well to EGFR-tyrosine kinase inhibitors (EGFR-TKIs) but SCCs do not [6–8]. Cytotoxic chemotherapy or radiotherapy have been the standard treatments to deal with SCCs in the past and recently immune checkpoint blockade therapy has gradually become the new standard to treat advanced SCC with a high expression of programmed death ligand 1 (PD-L1). Consequently, patients with different types of lung cancer should undergo different treatment. The determination of the correct histological type is the key starting point.

An immunohistochemical (IHC) panel is commonly used to determine the histological type, however, this takes a couple of days. For endobronchial lesions, bronchoscopic biopsy can be used as a convenient way to acquire tumor tissue and diagnose the pathological type. The limitation is that only a small amount of tissue is extracted. Due to the heterogeneity of malignant tissues, uncertainty is associated

doi:10.1088/978-0-7503-2540-0ch3

with this invasive procedure. During cancer detection, bronchoscopic imaging techniques are convenient to detect abnormalities. An accompany diagnosis of histological type using bronchoscopy would be helpful in determining staging and subsequent treatment. Both traditional white-light bronchoscopy (WLB) and auto-fluorescent imaging (AFI) may have the potential ability to detect and diagnose endobronchial premalignant lesions in the bronchial mucosa [6–8]. The extent of the mass and the resection margins can be also estimated simultaneously [9]. In the literature, the comparison of AFI and WLB showed higher detection rates and sensitivities but variable specificity [10]. Inter-observer variability should be reduced to make these approaches more practical [9].

Machine-learning based automatic classification has been widely used for medical images. After image preprocessing, such as contrast enhancement or color analysis, meaningful diagnostic features from bronchoscopy can be extracted and combined to generate a classification model. Each case is given a probability indicating the likelihood of it being a certain type. Analyzing the whole tumor area with so-called computer-aided diagnosis (CAD) would be more objective and efficient.

A previous CAD system was developed to classify normal mucosa and lung tumors using WLB [11] and achieved an accuracy of 80%. In this study, we further extend the application to differentiate cancer types using WLB and AFI. In the experiment, via bronchoscopy, textures in multiple color channels of RGB are separately quantified to classify ACs and SCCs. The hue–saturation–value (HSV) color space translates the color component, making it more intuitive to provide a better color interpretation, such as the results mentioned in medical applications. The proposed method may facilitate the physicians' decision making in a clinical environment.

3.2 Materials and methods

3.2.1 Patient information

This study analyzing bronchoscopy images was approved by the institutional review board of Shuang Ho Hospital (New Taipei City, Taiwan). Due to its retrospective nature, informed consent from patients was waived. From September 2015 to April 2017, 70 patients at Shuang Ho Hospital were examined under BF-F260 (Olympus Optical, Tokyo, Japan). Among the 70 cases, 36 had normal mucosa findings while 34 had neoplastic findings using WLB and AFI. The pathological diagnosis of the 34 abnormal neoplastic cases was determined by bronchoscopic biopsies and a pathologist. In the image observations, only 23 endobronchial tumors were clearly recognized with WLB and AFI without confounding by bleeding. Other cases were difficult to interpret using human observation. Additionally, two SCCs, two unknown carcinomas, and one tracheal tumor were excluded due to small size. Finally, the enrolled tumor database included 12 ACs (patients aged 42–83 years) and 11 SCCs (for WLB, only 10; patients aged 50–90 years). The AC and SCC samples are shown in figures 3.1 and 3.2. The tumor areas that appeared in the WLB or AFI were determined and delineated by a physician as regions of interest (ROIs).

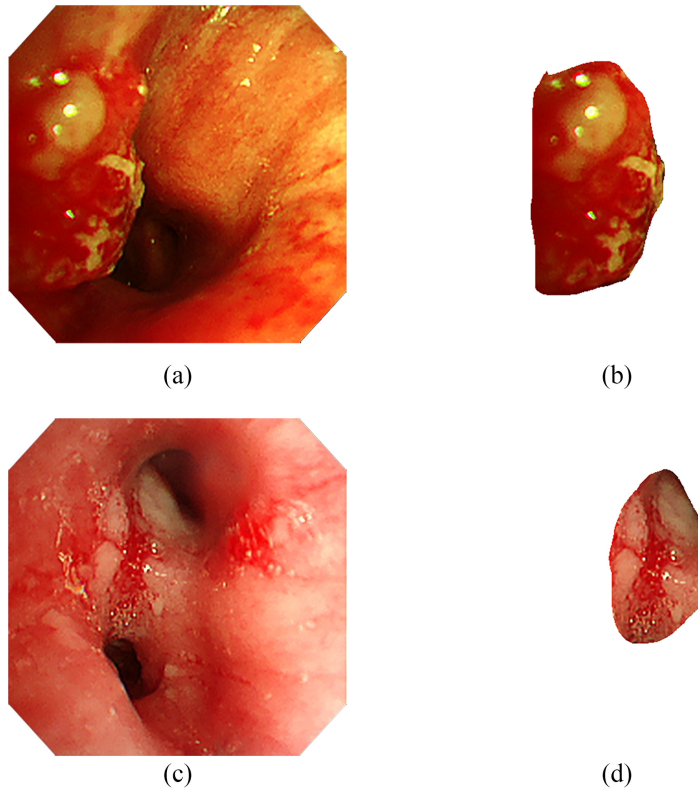

Figure 3.1. Samples of (a) adenocarcinomas in white-light bronchoscopy, (b) the corresponding tumor area and (c) squamous cell carcinomas in white-light bronchoscopy, (d) the corresponding tumor area.

3.2.2 Multichannel features

Medical images include various image appearances according to the tissue properties and how they are interpreted [12–14]. Tissues such as skin, fat, and bone can be differentiated by the rate of absorption of radiation in x-ray imaging or computed tomography. Different tissues have different reflecting echoes, which is the basic theory of ultrasound. These imaging modalities generate a gray-scale representation of organs inside the human body because they are determined by signal intensity. Compared to gray-scale images, the bronchoscopy images used in detecting lung cancer in this study have visible light as the illuminance source. These color images are composed of multi-channels, i.e. the standard red (R), green (G), and blue (B).

Day-to-day imaging examinations in hospitals mostly generate gray-scale images, which belong to the remit of the radiology department. Thus, previous CAD systems have quantified the image features of ultrasound, magnetic resonance imaging, and so on [15, 16]. Gray-scale images have only one channel, i.e. the brightness. In most situations, the diagnostic information extracted from the gray-sale is already sufficient for medical decision making, such as the differentiation of benign and malignant [15–19]. In this study, the proposed CAD systems use image features extracted from multiple channels to classify different lung cancer types. The

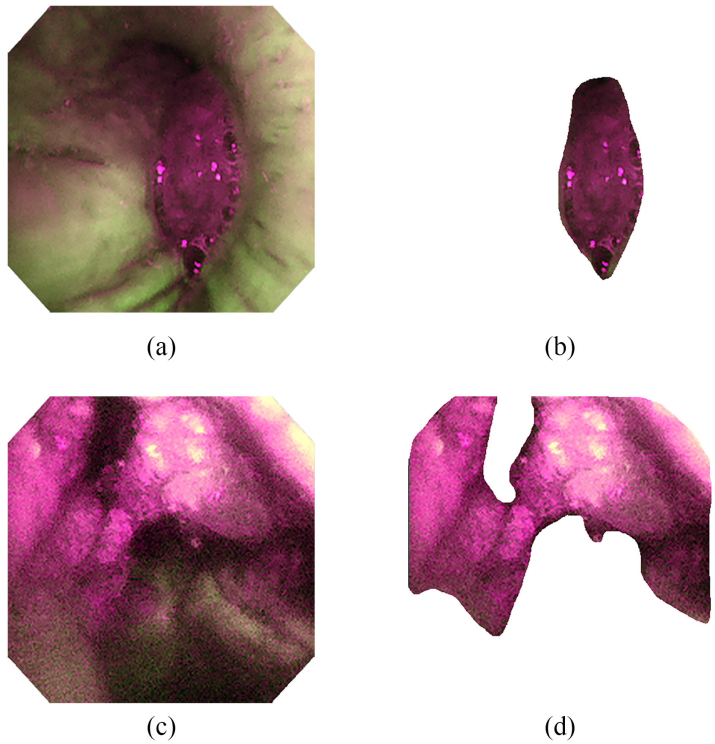

Figure 3.2. Samples of (a) adenocarcinomas in autofluorescent imaging, (b) the corresponding tumor area and (c) squamous cell carcinomas in autofluorescent imaging, (d) the corresponding tumor area.

bronchoscopy images generated from WLB and AFI are in color. Therefore, calculating quantitative image features from multiple channels rather than the conventional gray-scale is the main purpose of the experiment. In addition, the original standard RGB is not suitable for describing color. In order to extract meaningful color features for tissue characterization, the color space transformation is performed first to convert RGB to hue (H), saturation (S), and value (V). After that, textural features are individually extracted from the H, S, and V channels for the pattern recognition of lung cancers.

3.2.2.1 HSV transformation

As mentioned previously, unlike other medical images, bronchoscopy generates color images when using either WLB or AFI. Also, in clinical examinations, physicians have to detect abnormalities according to the color characteristics of tissues. Since color information plays such an important role, the use of color space substantially affects the classification result. The RGB color space is the original color space used by image capture modules and the JPEG format. However, it is not intuitive to describe color using the RGB composition. A better color space used to describe color is HSV [20]. HSV was proposed to imitate human perception. When the physician detects abnormal tissues from background tissues, they must depend greatly on hue, saturation, and brightness. Hue means the main color such as red or

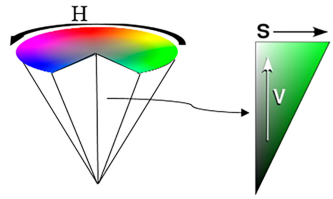

Figure 3.3. The HSV color space presents hue (H), saturation (S), and value (V).

yellow. Saturation indicates how rich the hue is, and brightness can tell us the level of luminance. Figure 3.3 shows the shape of the HSV color space.

3.2.2.2 Textural features

Pattern recognition in medical images always depends on the correlations between pixels, that is texture analyses [15, 16]. Texture is a second-order statistic, which means feature generation involving two or more pixels together rather than a lone pixel. Previous studies in the literature have used textural features to classify tissues in ultrasound or other imaging modalities [15, 16]. Here, the texture features were extracted from three channels (H, S, V) and combined to present the differences between malignant types.

The target tissues were just the tumor area, which was delineated by the physicians who carried out the bronchoscopy. Next, the gray-level co-occurrence matrix (GLCM) [21] was used to describe the joint frequencies of pair-wise pixel combinations. Because color images have three channels, the GLCM was individually applied to three channels. The whole delineated area was scanned to generate co-occurrence matrices $P = [p(i,j|d,\theta)]$ showing the frequencies of two adjacent pixels at a distance d and a direction θ. As per the equation listed below, the two gray-scale pixel values were i and j. Four offset directions, $\theta = 0°$, 45°, 90°, and 135°, were used to find the correlations between the two adjacent pixels, as shown in figure 3.4. To achieve rotation invariance, these four directions were combined in the matrix. The statistics of the matrix listed below are the 14 GLCM textural features:

$$\text{Autocorrelation} = \sum_i \sum_j (p_x - \mu_x)(p_y - \mu_y)/\sigma_x\sigma_y; \tag{3.1}$$

$$\text{Contrast} = \sum_n n^2 \left\{ \sum_i \sum_j p(i,j) \right\}, |i-j| = n; \tag{3.2}$$

$$\text{Correlation} = \frac{\sum_i \sum_j (i-\mu_x)(j-\mu_y)p(i,j)}{\sigma_x\sigma_y}; \tag{3.3}$$

$$\text{Cluster prominence} = \sum_i \sum_j (i+j-\mu_x-\mu_y)^4 p(i,j); \tag{3.4}$$

$$\text{Cluster shading} = \sum_i \sum_j (i+j-\mu_x-\mu_y)^3 p(i,j); \tag{3.5}$$

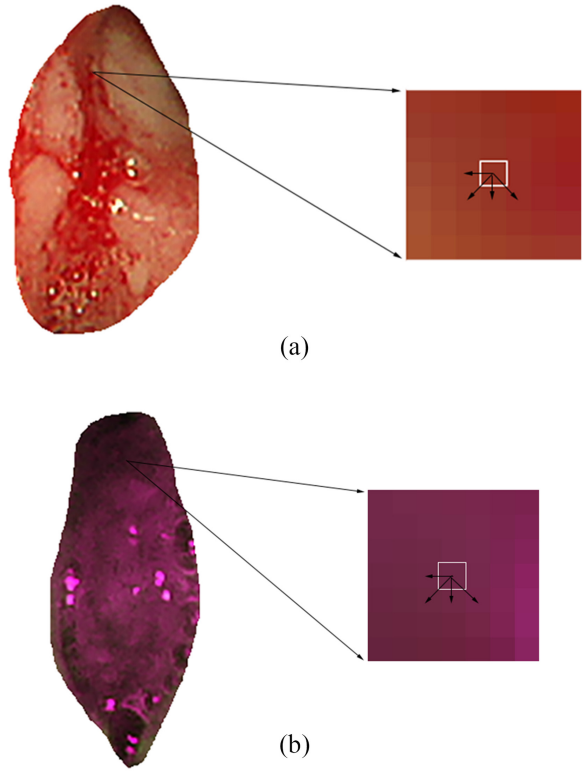

Figure 3.4. Texture features are correlations between adjacent pixels in (a) white-light bronchoscopy and (b) autofluorescent imaging.

$$\text{Dissimilarity} = \sum_i \sum_j p(i, j)|i - j|; \tag{3.6}$$

$$\text{Energy} = \sum_i \sum_j p(i, j)^2; \tag{3.7}$$

$$\text{Entropy} = -\sum_i \sum_j p(i, j)\log(p(i, j)); \tag{3.8}$$

$$\text{Homogeneity} = -\sum_i \sum_j \frac{1}{1 + i - j} p(i, j); \tag{3.9}$$

$$\text{Difference variance} = \sum_i i^2 p_{x-y}(i); \tag{3.10}$$

$$\begin{aligned} \text{Difference entropy} = &- \sum_i p_{x+y}(i)\log(p_{x+y}(i)); \\ &\times \frac{HXY - HXY1}{\max\{HX, HY\}} \\ HXY = \text{Entropy} = &- \sum_i \sum_j p(i, j)\log(p(i, j)), \end{aligned} \tag{3.11}$$

$$\begin{aligned}\text{Information measure of correlation} &= HXY1\\ &= -\sum_i\sum_j p(i,j)\log(p_x(i)p_y(j))\\ HX &= \text{entropy of } p_x,\\ HY &= \text{entropy of } p_y;\end{aligned} \tag{3.12}$$

$$\text{Inverse difference normalized} = \sum_i\sum_j \frac{1}{1+|i-j|}p(i,j); \quad \text{and} \tag{3.13}$$

$$\text{Inverse difference moment} = \sum_i\sum_j \frac{1}{1+(i-j)^2}p(i,j); \tag{3.14}$$

where μ_x, μ_y, σ_x, and σ_y are the means and standard deviations of the distributions of $p(i,j|d,\theta)$:

$$\mu_x = \sum_i i\sum_j p(i,j),\ \ \mu_y = \sum_j j\sum_i p(i,j) \tag{3.15}$$

and

$$\sigma_x^2 = \sum_i (i-u_x)^2\sum_j p(i,j),\quad \sigma_y^2 = \sum_j (j-u_y)^2\sum_i p(i,j). \tag{3.16}$$

3.2.2.3 Statistical analysis

Quantitative texture features were extracted from three channels. To determine whether they had diagnostic ability, these features were evaluated to determine if they were statistically significant in the differentiation of cancer type. The Kolmogorov–Smirnov test [22] was first used to evaluate if the texture features were normally distributed. Features having normal or non-normal distributions were then tested using the corresponding Student's *t*-test [22] and Mann–Whitney *U*-test [22]. The features with a *p*-value of <0.05 show a statistically significant difference. Texture features were also combined in the binary logistic regression model [23]. Using backward elimination, the most relevant feature combination can be selected during the elimination process.

Due to the limited size of the collected database, the leave-one-out cross-validation method [24, 25] was used to validate the generalization ability of the prediction model. With k cases, in each iteration $k-1$ are used for training and the remaining case was used to test the trained model. Using a biopsy-proven result as the ground truth of the cancer type, each case was given a malignancy probability. With different thresholds, five performance indices, including the accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), were obtained. The probability formula combined selected features to indicate the likelihood of being a certain type of cancer:

$$P = \frac{1}{1+e^{-(c_1\times x_1+c_2\times x_2+\cdots+c_n\times x_n+c_0)}}. \tag{3.17}$$

Performance index comparisons were evaluated using the chi-squared test in SPSS software (version 16 for Windows; SPSS, Chicago, IL, USA). The trade-offs between sensitivity and specificity under different thresholds were illustrated using the receiver operating characteristic (ROC) curve. A_z, the area under an ROC curve, was also evaluated using a bivariate chi-squared test in the ROCKIT software (C Metz, University of Chicago, Chicago, IL, USA).

3.2.3 Result

3.2.3.1 White-light bronchoscopy

From HSV color space, every channel has 14 GLCM textural features. In the test results, as shown in table 3.1, the significant features ($p < 0.05$) are correlation, cluster prominence, cluster shading, and difference variance in the S channel, and cluster prominence and cluster shading in the V channel. The logistic regression classifier selected correlation (S), difference variance (S), and cluster shading (V) to generate the prediction model and achieve an accuracy of 86% (19/22), a sensitivity of 90% (9/10), a specificity of 83% (10/12), a PPV of 82% (9/11), and an NPV of 91% (10/11). The AUC was 0.82. Only one SCC was misclassified.

3.2.3.2 Autofluorescent imaging

Two kinds of bronchoscopy images were used in the experiment to automatically classify malignant types. Same as WLB texture features, 42 AFI texture features were tested first. The information measure of the correlation in S and correlation in V had significant p-values <0.05 as shown in table 3.2. By combining these two features in the logistic regression classifier, the model achieved an accuracy of 83% (19/23), a sensitivity of 73% (8/11), a specificity of 92% (11/12), and $A_z = 0.82$. A further experiment was carried out using RGB features, and the information measure of the correlation in R and G were significant. Combining the features achieved an accuracy of 57% (13/23), a sensitivity of 73% (8/11), a specificity of 42% (5/12), and $A_z = 0.67$. Obviously, HSV performed better than RGB, in particular for specificity (p-value = 0.009 4) and only one AC was misclassified.

Table 3.1. Significant textural features in the hue–saturation–value (HSV) color space tested by the Student's t-test.

	AC	SCC	
Texture feature	Mean ± SD	Mean ± SD	p
Correlation (S)	0.987 ± 0.004	0.979 ± 0.004	<0.001[a]
Cluster prominence (S)	1589.937 ± 818.998	738.032 ± 433.412	<0.01[a]
Cluster shading (S)	130.865 ± 66.141	64.554 ± 36.635	<0.01[a]
Difference variance (S)	20.257 ± 9.692	12.559 ± 6.827	<0.05[a]
Cluster prominence (V)	298.707 ± 206.022	143.151 ± 80.896	<0.05[a]
Cluster shade (V)	−32.370 ± 20.906	−16.407 ± 8.691	<0.05[a]

[a] A p-value <0.05 indicates a statistically significant difference.

AC, adenocarcinoma; SCC, squamous cell carcinoma; S, saturation; V, value; SD, standard deviation.

Table 3.2. Significant HSV textural features and corresponding *p*-values evaluated using Student's *t*-test.

	AC	SCC	
Feature	Mean ± SD	Mean ± SD	*p*-value
Information measure of correlation (S)	−0.854 ± 0.015	−0.821 ± 0.039	<0.05[a]
Correlation (V)	0.984 ± 0.003	0.976 ± 0.013	<0.05[a]

[a] A *p*-value <0.05 indicates a statistically significant difference.

AC, adenocarcinoma; SCC, squamous cell carcinoma; S, saturation; V, value; SD, standard deviation.

3.2.4 Discussion

Currently, bronchoscopy is widely used in detecting and taking biopsies of endobronchial lesions of lung cancer. If physicians could obtain an initial classification of carcinomas using only bronchoscopy images, their practical use in clinical contexts would be much greater. Therefore, this study proposed using quantitative image features to generate a prediction model for a rapid and accurate assessment of malignancies. Based on the color textures in the HSV color space, the proposed CAD system can generate more objective diagnoses with less inter-observer variability than was reported in the previous literature as the diagnosis limitation. Using WLB, the CAD system achieved an accuracy of 86% (19/22), a sensitivity of 90% (9/10), a specificity of 83% (10/12), a PPV of 82% (9/11), and an NPV of 91% (10/11). AFI had better sensitivity and specificity to clinically detect early endobronchial lesions. In our result, using GLCM textural features from HSV images achieved an accuracy of 83% (19/23), a sensitivity of 73% (8/11), and a specificity of 92% (11/12), which was better than the results achieved using RGB features.

The HSV model separates the brightness component (V) of a color from its chrominance components (H and S). Previously, Sural *et al* used HSV to develop a workflow in image segmentation and color histogram generation. The HSV features achieved substantial improvement in content based image retrieval [26]. For medical images, HSV has also shown its comparably better ability in the definition of existing trauma colors than RGB [27]. HSV transformation is not a time-consuming or computation intensive process. Using the HSV color texture, an initial assessment of tissue malignancy using bronchoscopy images can be an alternative approach to the histological examination of biopsy specimens. Clinical physicians could obtain rapid results before a final pathological diagnosis is available. This alternative procedure could facilitate physician availability for more severe and urgent lung cancer patients.

Although this study did not collect many cases, the experiments using WLB and AFI to classify malignant lung tumors is the first study exploring the ability of the image processing and logistic regression classifier. We showed that both WLB and AFI can achieve accuracies higher than 80%. A previous CAD system achieved a similar accuracy (83% versus 80%) in classifying normal mucosa and tumors using WLB [18]. It also demonstrates the strong potential for the use of image features to classify different types of tumors. In future studies, automatic lesion detection should be added to improve the practical use of the proposed CAD system. While

carrying out bronchoscopy, physicians could detect more possible abnormalities and perform a rapid assessment of malignancy and subsequent treatment. Also, combining WLB and AFI may achieve higher accuracy, although the clinical use is not so practical.

3.3 Conclusion

This study proposed using WLB and AFI image features to classify different malignant types. HSV transformation and texture features were used to strengthen the classification ability. Both WLB and AFI achieved a diagnostic accuracy higher than 80%, which can provide an assessment suggestion to physicians in clinical use.

References

[1] Ferlay J *et al* 2015 Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012 *Int. J. Cancer* **136** E359–86

[2] Fitzmaurice C *et al* 2017 Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study *JAMA Oncol.* **3** 524–48

[3] National Lung Screening Trial Research Team 2011 Reduced lung-cancer mortality with low-dose computed tomographic screening *N. Engl. J. Med.* **365** 395–409

[4] Zaric B *et al* 2011 Autofluorescence imaging videobronchoscopy in the detection of lung cancer: from research tool to everyday procedure *Expert Rev. Med. Dev.* **8** 167–72

[5] Zaric B *et al* 2013 Diagnostic value of autofluorescence bronchoscopy in lung cancer *Thorac. Cancer* **4** 1–8

[6] Lam S, MacAulay C, leRiche J C and Palcic B 2000 Detection and localization of early lung cancer by fluorescence bronchoscopy *Cancer* **89** 2468–73

[7] Hirsch F R *et al* 2001 Fluorescence versus white-light bronchoscopy for detection of preneoplastic lesions: a randomized study *J. Natl. Cancer Inst.* **93** 1385–91

[8] Sharma D, Newman T G and Aronow W S 2015 Lung cancer screening: history, current perspectives, and future directions *Arch. Med. Sci.* **11** 1033–43

[9] Kennedy T C, Lam S and Hirsch F R 2001 Review of recent advances in fluorescence bronchoscopy in early localization of central airway lung cancer *Oncologist* **6** 257–62

[10] Sun J *et al* 2011 The value of autofluorescence bronchoscopy combined with white light bronchoscopy compared with white light alone in the diagnosis of intraepithelial neoplasia and invasive lung cancer: a meta-analysis *J. Thorac. Oncol.* **6** 1336–44

[11] Benz M, Rojas-Solano J R, Kage A, Wittenberg T, Munzenmayer C and Becker H D 2010 Computer-assisted diagnosis for white light bronchoscopy: first results *CHEST* **138** 433A

[12] Chang R-F, Lee C-C and Lo C-M 2019 Quantitative diagnosis of rotator cuff tears based on sonographic pattern recognition *PLoS One* **14** e0212741

[13] Lo C-M, Hung P-H and Hsieh K L-C 2019 Computer-aided detection of hyperacute stroke based on relative radiomic patterns in computed tomography *Appl. Sci.* **9** 1668

[14] Moon W K *et al* 2018 Quantitative breast density analysis using tomosynthesis and comparison with MRI and digital mammography *Comput. Methods Programs Biomed.* **154** 99–107

[15] Hsieh K L-C, Chen C-Y and Lo C-M 2017 Radiomic model for predicting mutations in the isocitrate dehydrogenase gene in glioblastomas *Oncotarget* **8** 45888
[16] Hsieh K L-C, Chen C-Y and Lo C-M 2017 Quantitative glioma grading using transformed gray-scale invariant textures of MRI *Comput. Biol. Med.* **83** 102–8
[17] Moon W K, Chen I-L, Chang J M, Shin S U, Lo C-M and Chang R-F 2017 The adaptive computer-aided diagnosis system based on tumor sizes for the classification of breast tumors detected at screening ultrasound *Ultrasonics* **76** 70–7
[18] Hsieh K L-C, Lo C-M and Hsiao C-J 2017 Computer-aided grading of gliomas based on local and global MRI features *Comput. Methods Programs Biomed.* **139** 31–8
[19] Hsieh K L-C, Tsai R-J, Teng Y-C and Lo C-M 2017 Effect of a computer-aided diagnosis system on radiologists' performance in grading gliomas with MRI *PLoS One* **12** e0171342
[20] Oliveira R B, Mercedes Filho E, Ma Z, Papa J P, Pereira A S and Tavares J M R 2016 Computational methods for the image segmentation of pigmented skin lesions: a review *Comput. Methods Programs Biomed.* **131** 127–41
[21] Haralick R M, Shanmuga K and Dinstein I 1973 Textural features for image classification *IEEE Trans. Syst. Man Cybern.* **Smc3** 610–21
[22] Field A P 2009 *Discovering Statistics Using SPSS* 3rd edn (Los Angeles, CA: SAGE)
[23] Hosmer D W 2000 *Applied Logistic Regression* 2nd edn (New York: Wiley)
[24] Alpaydin E 2004 *Introduction to Machine Learning* (Cambridge, MA: MIT Press)
[25] Refaeilzadeh P, Tang L and Liu H 2009 Cross-validation *Encyclopedia of Database Systems* (Berlin: Springer), pp 532–8
[26] Sural S, Qian G and Pramanik S 2002 Segmentation and histogram generation using the HSV color space for image retrieval *Proc. of Int. Conf. on Image Processing, 2002* **vol 2** (Piscataway, NJ: IEEE), p I–II
[27] Georgieva L, Dimitrova T and Angelov N 2005 RGB and HSV colour models in colour identification of digital trauma images *Proc. of the Int. Conf. CompSysTech*