

Classification of Malignancy in Suspicious Lesions Using Autofluorescence Bronchoscopy

Tomaž Finkšt¹ – Jurij Franc Tasič² – Marjeta Zorman Terčelj³ – Marko Meža^{2,*}

¹ University of Ljubljana, Faculty of Mechanical Engineering, Slovenia

² University of Ljubljana, Faculty of Electrical Engineering, Slovenia

³ University Medical Centre, Clinical Department of Pulmonary Diseases and Allergy, Slovenia

This paper presents a novel approach to the classification of bronchial tissue as either malignant or precancerous based on autofluorescence bronchoscopy (AFB) images. The study consisted of 44 images, of which 22 were confirmed as malignant and 22 as nonmalignant precancerous cases. Our approach starts with the detection of a region of interest (ROI). This is followed by an analysis of semi-normal intensity distributions in gray-scale images of red and green components of the previously identified ROI. Based on the results of this analysis, features are computed, which are then used to build an image-classification model. This model classifies the tissue images into malignant/nonmalignant classes. We utilized several classification algorithms, i.e., naive Bayes, K-nearest-neighbor (K-NN), and support vector machine (SVM) with dot kernel. The criteria used when testing their performance were accuracy, sensitivity, specificity, and the area under the curve. Wilcoxon's signed-rank test was used to confirm the accuracy of the classification method. The proposed method was compared to a similar approach reported by Buoutris et al., who analyzed the texture features in a gray-level co-occurrence matrix (GLCM). Using the best-performing classification algorithm (SVM with dot kernel), the accuracy of the proposed approach (95.8 %) was better than that reported by Bountris et al. (92.1 %).

Keywords: image segmentation, edge detection, autofluorescence bronchoscopy, machine learning

Highlights

- A new procedure for the classification of bronchoscopic images into malignant and precancerous cases is presented.
- Wilcoxon's signed-rank test confirmed the efficiency of the new classification procedure.
- The small number of image features (< 10) used in the classification means that the computing-power requirements are small enough for the implementation of the algorithm in an embedded system.
- As machine learning is an essential component of the new procedure, the method can be easily adapted to image characterization in other technical fields.

0 INTRODUCTION

Lung cancer is a malignant disease of the bronchial tubes. The progress of the disease is usually rapid in cases of symptomatic small-cell and non-small-cell lung carcinomas. Early detection of the malignant and precancerous lesions is mandatory for a successful treatment. Autofluorescence bronchoscopy (AFB) is used for the detection of degenerated tissue on the inner surfaces of the bronchial tubes [1]. In cases of a positive AFB finding, a laboratory analysis of a tissue sample is used to confirm or reject the result.

AFB was developed from white-light bronchoscopy (WLB), which was used as a diagnostic tool. In contrast to WLB, AFB utilizes blue light (wavelengths of 380 nm to 460 nm) [2] to [4]. Precancerous (early-stage) lesions are differentiated from healthy bronchial tissue based on the differences between the fluorescent light that is reflected. In the subsequent pathophysiological examination, dysplastic lesions, metaplastic lesions, and carcinomas in situ are characterized by more layers of cells, with diameters ranging from micrometers

to millimeters. They are differentiated from healthy tissue in a pathological examination by the formation of veins under the basal membrane [1] and different concentrations of fluorophores. AFB utilizes the difference in the fluoroscopic properties of malignant and nonmalignant tissues to detect malignancy, making it a superior diagnostic procedure to WLB. In AFB, a healthy mucus membrane emits more green than red light, and unhealthy tissue emits more red than green. Only a little blue light is reflected, either from the healthy or unhealthy tissue. Different implementations of stimuli and sensors can add their particularities to the general picture of fluorescent bronchoscopy.

Many clinical studies have confirmed the utility of AFB in the detection of cellular changes in early-stage lung cancer, with noninvasive neoplastic lesions, small lesions, preneoplastic lesions, and neoplastic changes in tissue detected using state-of-the-art AFB diagnostic systems. Although such detection is possible using WLB, its performance is not as good as that of AFB [2], [5] to [7]. In addition to its utilization of changes in fluorescence in the detection of

*Corr. Author's Address: University of Ljubljana, Faculty of Electrical Engineering, Tržaška cesta 25, 1000 Ljubljana, Slovenia, marko.meza@fe.uni-lj.si

precancerous cells, the superior performance of AFB is attributed to its utilization of algorithms in image characterization. A number of studies showed that AFB significantly outperformed WLB in the detection of preneoplastic and early neoplastic lesions [2], [3] and [8].

However, AFB suffers from a lack of specificity when compared to WLB, with larger numbers of false positives (FPs). According to some studies, the incidence of FPs was as high as 30 % [2], [6], [8] to [10]. Inflammation and other non-life-threatening conditions, which result in changes in cell fluorescence, contribute to the high FP rate in AFB. Thus, to confirm the AFB findings, a biopsy is needed. Biopsies are time consuming and costly and can be distressing for the patient [2] and [10]. Thus far, none of the commercial AFB diagnostic systems (e.g., AFB system LIFE[®], Karl Storz[®], D-light[®], Pentax SAFE 3000[®], and DAFE[®]) have managed to resolve the FP problem.

AFB is designed to produce images in the Red Green Blue (RGB) color space. However, the fluorescent emissions contain only the R and G component. Thus, the image has to be transformed into a grey-scale image, where one of the two colors is represented by the minimum intensity and the other color is represented by the maximum intensity [11].

Although a number of studies have described the use of advanced technologies in diagnosis and treatment [12] to [14], only a few studies have focused on improving the specificity of AFB [15] to [17]. In one study of the spectra of autofluorescent light, the authors added endoscopic optical spectroscopy to AFB to improve its specificity [17]. Bard et al. [15] reported a real-time analysis of the autofluorescent spectrum while manipulating the tissue. The procedure was prone to tissue bleeding. Zeng et al. [16] and Terčelj et al. [17] used experimental apparatuses consisting of an AFB machine and a laboratory spectrometer in noncontact tissue examinations. However, modifications to AFB machines add to their complexity, and as a result, operators need to develop additional skills. Furthermore, the examination times are longer because the patient has to remain intubated, and this should be restricted to the shortest possible duration. Goujon et al. [18] and Qu et al. [19] described the use of follow-up, off-line analyses of spectral images in AFB to reduce the FP results. Other authors reported numerical manipulations of the red and green emissions, combined with different methods for thresholding the images into true positive and FP groups [18] and [20]. They reported a reliability of 79 % using these methods. They also reported that

the addition of the results of WLB slightly improved the findings. In other AFB studies, Bountris et al. [21] and [22] transformed the original color space (RGB) to a more suitable color model (hue saturation value (HSV)) for the purpose of tissue examinations. They reported that the resulting image textures, aided by neural networks, minimized the FP results. Using this approach, it was possible to distinguish between the malignant and the inflamed tissue in the processed images. Haritou et al. [23] developed a computer tool for the off-line classification of AFB images. They used the HSV color space and performed a texture analysis in which sets of characteristic features were defined, depending on the classification type of the image. The texture features they used were Laws' texture energy (TE) and first- and second-order statistics of the data in gray-level co-occurrence matrices (GLCMs). They combined these classification methods with feature-selection methods to develop a reliable and efficient classification method. The final result was a color texture vector consisting of 10 features, which was evaluated for irregularities (malignancy/disease) using a custom neural network. The off-line tool was developed for a procedure evaluation and suggestions for refinements by physicians.

In the present paper we describe the development of a method for the detection of suspicious lesions (regions of interest (ROI)) on AFB images and the classification of these areas into malignant/nonmalignant tissue. First, we created and evaluated the method for the detection of the ROI in comparison with a human expert (see Section 3.1). Next, we extracted a set of image features based on a classifier, which we used to classify the ROI into malignant/nonmalignant tissue. Subsequently, we evaluated the performance of the classifier by comparing its findings with ground-truth data, which were based on the findings (histological analysis) of a tissue biopsy. Finally, we compared the performance of our approach with that of the method developed by Bountris et al. [21] and [22].

1 METHODS

1.1 Data Acquisition

AFB images obtained from patients who attended the Department of Lung Disease and Allergy at the University Clinical Center in Ljubljana, Slovenia were included in the study. All the patients underwent WLB. In cases where malignancy was suspected, AFB (LIFE[®]) images were obtained. The LIFE[®] system consists of a light source, a light fiber, and a camera

with an optical coating. The camera was calibrated before each AFB examination. Using this system, a healthy mucous membrane appears green on the image, and malignant and premalignant lesions appear red or red-brown due to the different fluorescence properties of non-healthy tissue. In all suspicious cases, a biopsy was performed, and the samples were histologically analyzed. The biopsy site was based on the findings of the AFB examination (i.e., abnormal fluorescence properties of the tissue).

The following data were obtained: images and videos of AFB and WLB examinations of 50 patients, results of histological tissue analyses, and other non-relevant study data. No personal data were collected from the database. The study was approved by the Medical Ethics Committee in Slovenia.

The exclusion criteria were the absence of histological data and blurred images due to various causes (e.g., the presence of blood and mucus resulting from intubation, pinching the tissue while obtaining the biopsy sample, and excessive tissue pathology). The final study consisted of 44 images. A distinct area of suspicion and abnormal fluorescence was present in at least one area in each image. The results of the histological analysis of the biopsy tissue samples revealed 22 cases of malignant changes and 22 cases of inflammation (Table 1).

On the WLB images, medical experts marked the ROI (i.e., suspected area of malignancy).

Table 1. Image database

	Malignant tissue changes	Other irregularities
Number of images	22	22

1.2 Procedure for Machine-Supported AFB Reading

Developments in software design, embedded systems, and workstations have had dramatic influences on most aspects of modern life. AFB needs to exploit these modern technologies in diagnostic processes.

A machine-supported AFB system autonomously annotates suspicious areas on an image. An expert then interprets the image reading and makes a diagnosis. Machine support makes the work less stressful, decreases the likelihood of errors, and increases the productivity. Machine-supported AFB reading can be designed to work in real-time or off-line. The latter is easier to implement. A major benefit of machine-supported reading is that it improves the specificity of the AFB readings.

Fig. 1 depicts the steps involved in machine-supported AFB reading. The steps are briefly described in the following subsections.

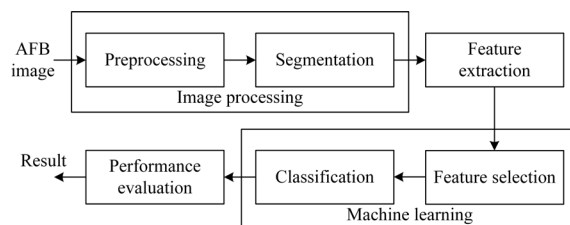


Fig. 1. Flowchart of AFB machine reading

1.2 Preprocessing of AFB Images

Preprocessing is carried out to standardize the image quality (parameters) and produce uniform images for further processing. During preprocessing, the color intensity, contrast, and image sharpness are modified. To improve the homogeneity of the pixel intensity, border-intensity enhancement, morphological closing, and Gaussian low-pass filtering (BCG) are employed [24].

1.2.1 BCG Method

Gaussian BCG low-pass filtering involves border intensity enhancement, morphological closing, and Gaussian low-pass filtering [24]. The border intensity enhances the contrast where the borders between different textures are detected. Morphological closing involves dilation, followed by erosion of the surfaces, based on a predefined set of rules with the purpose of image structuring on defined surfaces. As a result, the border lines between the surfaces are smoothed, small spots are integrated into the surfaces, and straits are withdrawn. Gaussian low-pass filtering reduces the contrast among the different surfaces of the image.

The background and defined objects (ROI) are processed separately using different settings. The background is processed for uniformity, and the ROI is processed for detail enhancement.

In the present study, the BCG parameters were: the projected size of the surface in morphological closing, projected surface size deviation σ , window size Ω , and targeted smoothness in the Gaussian core. Small changes in intensity were considered to denote the inhomogeneity. Consequently, the surfaces should not be small compared to the whole image.

Fig. 2 presents a flowchart of the AFB image-preprocessing steps.

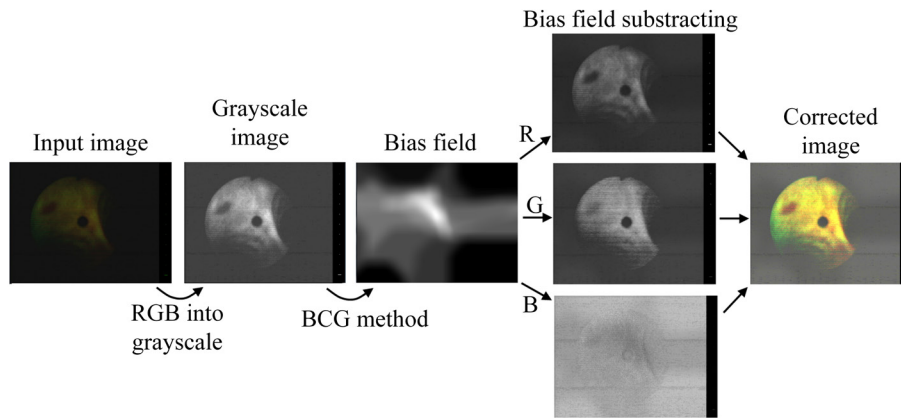


Fig. 2. Flowchart of preprocessing for an AFB image

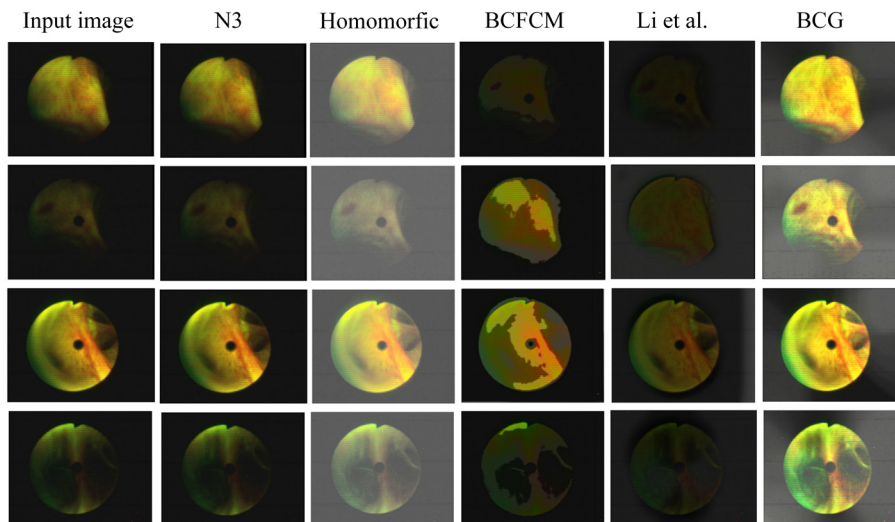


Fig. 3. Assessment of the different preprocessing steps for an AFB image

Following the preprocessing of the AFB image, suspicious areas were detected using the BCG method. Alternatives to the BCG method are the N3 method [25], homomorphic filtration [26], border closing fuzzy C-means [27] and the method of Li et al. [28]. The practical advantages of the BCG method include its computational efficiency, thereby increasing the speed of the operation, as well as its robustness to noise and a dynamic range of possible inputs. The method is particularly useful for processing database images produced in similar environments.

1.3 Image Segmentation and Definition of the ROI

The intensity of the fluorescence light in AFB is influenced by the excitation wavelength λ , the absorption, the scattering, and the distance [29] as well as the different fluorescence properties of malignant

and non-malignant tissue. The latter is used to discriminate malignant and non-malignant tissue.

During the imaging, the intensity of each image was calibrated for different distances. The definition of the ROI was based on the preprocessed AFB image. All non-healthy tissue was included in the ROI. Malignant changes, inflammation, bruises, and the excessive age degeneration of the tissue appeared similar on the image and needed to be investigated in detail. Once defined, the ROI was compared with the WLB image and the results of the physician's observation.

The R/G intensity ratio of the surface of the image was calculated, and the results were presented in a gray image, followed by filtering. The signal-to-noise (S/N) ratio was increased by lowering the noise while taking care to preserve the edges. Different filters were used: simple averaging, Gaussian filtering, and

anisotropic diffusion filtering, and the last of these yielded the best results.

1.3.1 Anisotropic Filtering

Anisotropic filtering for image processing was performed in accordance with the anisotropic diffusion method of Perona and Malik [30]. Using this method, a line (or curve) of filtering is first defined, which is then performed along this line. The filter is used to preserve the edges and smooth out the surfaces [31]. The process is iterative and semi-reversible. The differential equations used in anisotropic diffusion filtering have been described earlier [31].

The requirements are:

- causality: Iterations must not produce new (false) features;
- at a minimum, edge sharpness and position must be preserved;
- intra-region filtering is prohibited;
- edges should ideally be enhanced.

Filtering is described by Eq. (1):

$$\partial_t u = \text{div}(g(x, y, t)\nabla u) = g(x, y, t)\Delta u + \nabla g \nabla u, \quad (1)$$

where u is the intensity of the image element, and $g(x, y, t)$ is the diffusion coefficient (function of position and time). As filtering intensifies with the number of iterations, the diffusion coefficient g was small (0) on the edges and large (1) within the surfaces. The brightness gradient $|\nabla u(x, y, t)|$ was used for g modulation. The diffusion coefficient g was a non-negative, monotonously decreasing function of the intensity gradient ∇u , with the boundary condition $g(0)=1$. Shaping g changed the diffusion rate. When $g=1$, the diffusion became linear, and the properties of the image had no effect on the filtering.

To enhance the edges while smoothing out the surfaces, Eqs. (2) or (3) was used:

$$g(\nabla u) = e^{-\left(\frac{|\nabla u|}{\lambda}\right)^2}, \quad (2)$$

$$g(\nabla u) = \frac{1}{1 + \left(\frac{|\nabla u|}{\lambda}\right)^2}. \quad (3)$$

Eq. (2) was used to enhance the contrast intensity, and Eq. (3) was used to enhance the filtering on wide surfaces and modulate g [32].

The parameters were the number of iterations n and the spatial scaling factor λ of the gradient u . Noisy images require a larger λ , and iterations n wipe out the image details while preserving well-defined edges.

A numerically similar application, solving the Fourier law of heat flow, was speed optimized for implementation in an embedded system [33]. The author reported that this method was 110 times faster than the existing commercial solutions. The mathematics employed in the present study are similar to those used in the method described in [33]. Thus, verified guidelines for the speed optimization of filtering for real-time, machine-supported AFB diagnostics exist.

1.3.2 Canny Algorithm for Edge Detection

The Canny algorithm proceeds as follows. First, a Gaussian filter is used for moderate filtering of an AFB image [34]. Parameter σ (standard deviation of the Gaussian distribution) modulates the filtering intensity. The diffusion coefficient g is calculated by Eq. (4), and the angle of searching for an edge is calculated by Eq. (5).

$$g(x, y) = \sqrt{G_x^2 + G_y^2}, \quad (4)$$

$$\alpha(x, y) = \arctan \frac{G_y}{G_x}, \quad (5)$$

G_x and G_y are calculated using the Sobel, Prewitt or Roberts's operator. The edge points determined in Eq. (4) give rise to edges with the magnitude of the image gradient. The algorithm then tracks the top of these edges, and any pixels that are not on the top of the edge are set to zero to produce a thin line in the output. This is known as non-maximal suppression. The edge pixels are then thresholded by hysteresis thresholding, which is based on the use of two thresholds: $T1$ and $T2$, and $T1 > T2$. Edge pixels with values greater than $T2$ are considered strong edge pixels, and edge pixels with values between $T1$ and $T2$ are considered weak edge pixels. Values below $T1$ form surfaces.

A flowchart of the Canny algorithm is presented in Fig. 4. As shown in the figure, as the Gaussian filter lowered the noise level, the S/N ratio increased. The horizontal (x axes) and vertical (y axes) gradients of the image element intensities were calculated. The gradient magnitude was calculated using Eq. (4). Local maxima were found. All the other gradients were set to 0.

The threshold filter separates the edge pixels from the background. The inbuilt hysteresis in the filter allows for a distinction between the edges and the background. Pixels with intensities below the

lower threshold form the background, and those with intensities above the upper threshold form edges. Pixels in-between these thresholds represent edges if they have edge-neighbors; otherwise, they form the background.

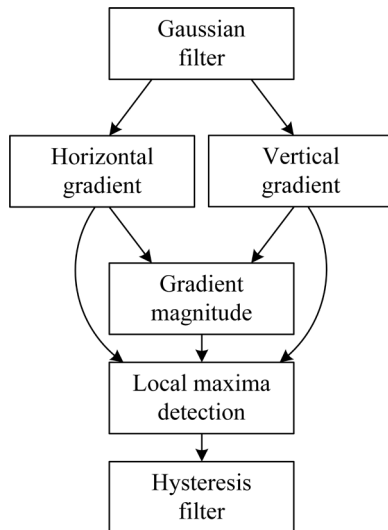


Fig. 4. Flowchart of AFB image preprocessing

1.4 Feature Extraction

The surface-classification method of Bountris et al. [21] and [22] was used. This method is based on a feature study. Ideally, only a few features with lots of content are identified. The presence of multiple features increases the noise level, which interferes with the efficiency of the surface classification. Excessive image preprocessing prior to the feature search can potentially result in the depletion of TE. To preserve TE, we repeated the feature search using a combination of 25 Laws' image filters [21] and first- and second-order statistics. Each ROI was mapped from the RGB to the HSV color space. Component H was used in the analysis. The combination of 25 Laws' image filters and the results of the ROI mapping yielded 25 TE images. Seven first-order statistics were calculated for each image, resulting in 175 TE features. The calculation of GLCM ($d=1$, $\theta=45^\circ$), and 22 first-order statistics produced an additional 550 TE features. Each ROI was characterized by 725 TE features, which were ordered into feature vectors. The dimension of each feature vector was impractically high. Information-rich features are mandatory for differentiating between the malignant and FP areas. Therefore, we employed principal component analysis (PCA) to generate a reduced number of information-rich features.

1.5 Feature Reduction

PCA was performed to reduce the data into smaller numbers of information-enriched features [35] and [36]. These features (inputs to the method) were then mapped into a space with fewer dimensions. The data variance in lower-dimensional space needs to be high to guarantee a high correlation between the image properties and the corresponding image feature. Thus, PCA was performed to transform the correlated (linearly dependent) features into uncorrelated (linearly independent) features. The resulting features were orthogonal and formed the base of the feature space.

The features were ranked according to their information relevance, with features having higher variance considered more important in the image. The most important features were used in the image analysis, and the least important features were considered noise [35]. PCA was also applied to the feature vectors. A subset of the resulting features was then used to build a machine-learning model for the classification of the samples into malignant and nonmalignant cases. In the iterative PCA, which was run using different settings, we selected the setting that resulted in the best classification models.

1.6 Image Classification

The bronchial images were classified using a combination of machine learning and calculation and the classification of information-rich features, all of which were implemented using computationally efficient algorithms. The following supervised-learning classification algorithms were tested in the creation of the classification models: naive Bayes classifier, K-nearest-neighbor (K-NN), and support vector machines with dot-kernel-type classifier (SVM). The dot kernel used was defined by $k(x, y) = x \times y$ [37].

1.7 Assessment

To assess the performance of the classification model, we computed an accuracy sensitivity and specificity measure and then compared the results with those of other studies [21] and [22]. As the number of images was small, we applied the procedure of uniform sampling, with exchange (bootstrap) [38] to assess the efficiency of the machine learning.

2 PROPOSED NOVEL METHOD FOR FEATURE EXTRACTION

The method for extracting the texture features draws on the work of Bountris et al. [21] and [22]. These texture features and an image histogram were used to produce the FOS. The histogram (HSV space) graphically showed the distribution of the color intensity and provided direction on additional information that was needed to improve the image reading.

In our approach, we created features by plotting the red (R) and green (G) channel histogram of a ROI in one graph. A fixed area within a ROI was chosen and analyzed in all cases. Both the R and G histograms of the ROI showed bell-type distributions. Fig. 5 presents an example of a histogram where the x axis shows the channel intensity (0 to 255, 8-bit resolution) and the y axis shows the number of corresponding intensity pixels within the ROI. The histograms of both channels were then approximated using the Gaussian-approximation function. The function (see Eq. 6) for each channel was defined by the values of the mean μ , the standard deviation of σ , and the peak value M . The computed Gaussian approximation of the red and green channels of the selected image is also shown in Fig. 5. The histograms of all the images were similar. In all cases, although μ , σ , and M differed, a semi-normal distribution was always present.

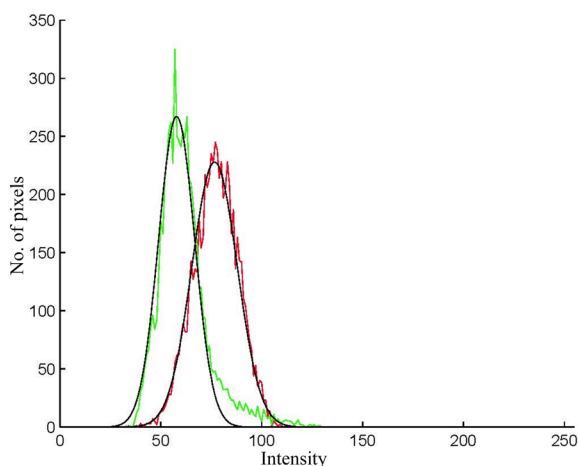


Fig. 5. Histogram of the R and G plane in an AFB image

Eq. (6) is the scaled Gaussian function that was used to model the R and G histograms in Fig. 5.

$$y(x, M, \mu, \sigma^2) = \frac{M}{\sqrt{2\sigma^2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (6)$$

The parameters of the Gaussian functions (μ , σ , and M) and the cross-section of the two curves then served as features for the machine-learning model. Fig. 6 shows a flowchart of the feature extraction.

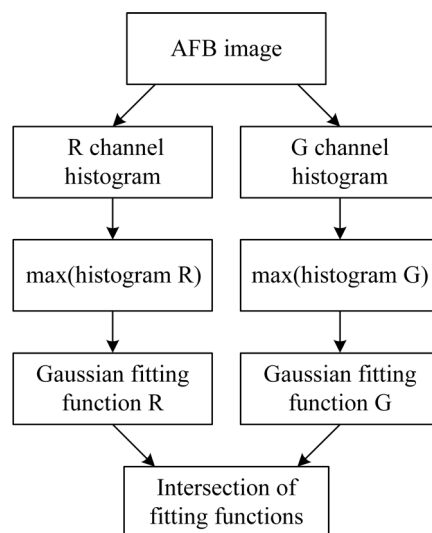


Fig. 6. Feature extraction

In total, seven features were extracted from each sample. The extracted features were then used to build classification models for the discrimination of the malignant and nonmalignant cases. Next, we tested the performance of the classification models containing the extracted features using naive Bayes, K-NN, and SVM with dot kernel machine-learning algorithms. As a performance parameter, we observed the accuracy measure.

3 RESULTS

3.1 Detection of a Suspicious Area

In each WLB image, a medical expert first selected the ROI, marking what he/she considered a suspicious area (malignant tissue). The expert-marked ROI was then compared with the machine-defined ROI. When comparing the machine-defined ROI with the expert-defined ROI, the former was considered as ground truth. In the comparison, each pixel of the machine- and expert-defined ROI was compared.

False-positive errors (FPEs) and false-negative errors (FNEs) were recorded. The mean values and standard deviations of the FPEs and FNE during the image processing were calculated in the RGB and HSV color spaces (Table 2). It is clear that the image processing in the HSV color space produced slightly

better results than the image processing in the RGB space.

Table 2. Comparison of ROI detection methods

	FPE [%]		FNE [%]	
	Mean	Std.	Mean	Std.
RGB	11.1	2.4	11.0	3.7
HSV	8.7	2.5	8.4	2.7

3.2 Classification

We compared the classification method of Bountris et al. [21] and [22] with our proposed method. As we were not able to obtain data from the authors of the reported method, we evaluated both approaches based on our own data.

The dataset consisted of 44 images, in which a ROI was identified. Malignancy was confirmed in 22 of the images. Irregularities present in the other 22 images had raised a suspicion of malignancy. These were biopsied but found to be nonmalignant. We performed randomized, stratified subsampling of 15 samples from the malignant dataset and 15 samples from the nonmalignant dataset. These sample cases were then used for learning. The other 14 cases (seven from the malignant set and seven from the nonmalignant set) were used for testing (Table 3).

Table 3. Learning and evaluation of the image sets

	Malignant tissue	Other tissue irregularities
Learning set	15	15
Testing set	7	7
Sum of images	22	22

3.2.1 Generation of Texture Features Using the Method of Bountris et al. [21] and [22]

We repeated the feature-extraction method described by Bountris et al. [21] and [22]. Fig. 7 outlines the

process used to produce 25 TE images of the ROI from the initial image [23]. The ROI was first defined in the AFB image. The area of the ROI, from which a biopsy was taken, was the area for machine processing. The area was then transformed into HSV color space. The H component was used for the convolution with 25 Laws' masks. This resulted in 25 TE images. Twenty-nine texture statistics were then calculated from each image, resulting in 725 features per ROI.

Selection of texture features using the method of Bountris et al. [21] and [22]

PCA was applied to reduce the feature space. In our case, we reduced 725 features to 10 independent features. These features were then used for machine learning. Several machine-learning algorithms (naive Bayes, K-NN, and SVM with dot kernel) were tested. The SVM with the dot kernel produced the best classification results. The results obtained were similar to those reported by Bountris et al. [21] and [22], who reported an accuracy of 91.2 % in their dataset. Therefore, we believe that we have managed to repeat their method accurately.

Proposed Method for Feature-Set Generation

We used the same learning/test dataset to evaluate the proposed method of feature extraction. We extracted a total of seven features from each ROI. The extracted features were Gaussian fitted parameters of both the red and green channels and the intersection points of both fitting curves. For a description of the feature-extraction method, please see Section 2. Naive Bayes, K-NN, and SVM with dot-kernel machine-learning algorithms were tested and used in the creation of the model.

3.3 Classification Results

The classification results obtained using the proposed method were evaluated by comparing them with the results obtained using the methods of Bountris

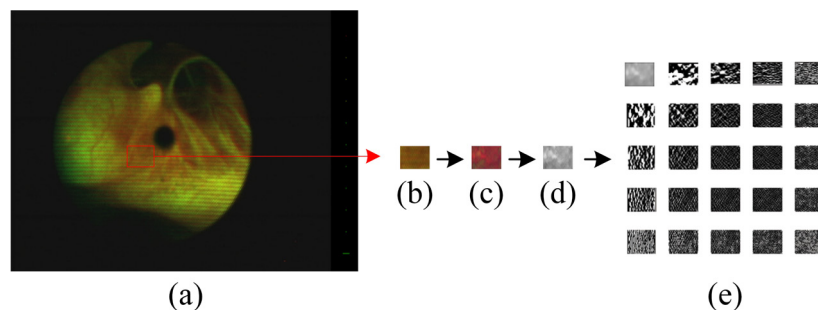


Fig. 7. Process of texture-feature selection

et al. [21] and [22]. The results of both methods are presented in Table 4.

Table 4. Comparison of both methods

	Method of Bountris et al. [21] and [22]	Proposed method
Classifier	Accuracy [%]	Accuracy [%]
K-NN	90.7	90.9
Naive Bayes	90.4	91.5
SVM with dot kernel	92.1	95.8
Average accuracy	91.1	92.7

Accuracy was considered as the most important parameter in the assessment of both methods. Other parameters that were assessed are presented in Table 5. The naive Bayes classifier was used to obtain the data in Table 5. The proposed approach yielded better results, with substantially less computing time than the method proposed by Bountris et al. [21] and [22].

Table 5. Comparison of other measures in the assessment of the methods (according to the naive Bayes classifier)

NB	Bountris et al. method [21] and [22]	Proposed method
Sensitivity [%]	81.82	95.45
Specificity [%]	90.91	86.36
Accuracy [%]	93.33	88.33
Recall [%]	83.33	95.00
AUC	0.883	0.958

Wilcoxon's signed-rank test was used to compare the performance of both methods. To validate the null hypothesis that both methods would produce similar results, using a statistical confidence level of $\alpha = 0.05$. The results of Wilcoxon's signed-rank test are shown in Table 6.

Table 6. Results of Wilcoxon's signed-rank test of both methods

	Accuracy	Recall	AUC
p	0.0112	0.0052	0.00063323

4 DISCUSSION AND CONCLUSION

We implemented and successfully tested a novel approach to the detection and classification of malignant tissue in AFB images. Using the proposed approach, we were able to successfully discriminate malignant and nonmalignant tissue in AFB images in borderline cases.

In the proposed approach, a suspicious area on an image was identified using the BCG method. Suspicious areas detected by an expert were then compared with those identified using the proposed algorithm. The algorithm produced usable results. It detected the ROI, performing slightly better when used in the HSV color space.

From each suspicious area (ROI), a tissue biopsy sample was obtained. This tissue sample was sent for histological analysis, and the results were used as ground truth when determining whether a tissue sample was malignant or nonmalignant. From the determined ROI and ground-truth data, a test/learn dataset was created for the classification of each ROI as malignant/nonmalignant.

We tested two approaches to tissue classification. We repeated the approach described by Bountris et al. [21] and [22] and compared it with our approach. The main difference between the two approaches was the different set of features that were extracted from the AFB images. The method of Bountris et al. [21] and [22] is based on texture features, whereas our approach is based on the difference in the fluoroscopic properties of malignant and nonmalignant tissue. We discriminated between the two types of tissue based on the red and green distributions in the histogram of the AFB image. We modeled the red and green channel histograms using a Gaussian fitting function and used the fitted parameters as features to discriminate between the two types of tissue.

Using the proposed approach, we obtained similar results to those reported by Bountris et al. [21] and [22]. Therefore, we believe that our implementation of their approach was accurate. The proposed approach yielded better accuracy than that reported by Bountris et al. [21] and [22]. They achieved an accuracy of 92.1 %. With our approach, using the SVM algorithm with the dot kernel, we achieved an accuracy of 95.8 %. In addition, our approach is computationally more efficient, as the number of features used for the modeling is small ($n = 7$). In contrast, a large number of features ($n = 725$) is used in the approach of Bountris et al. [21] and [22]. The relatively small number of features enabled us to create a classification model based on a relatively small sample set. Furthermore, over-fitting did not occur due to the small number of features compared to the number of samples.

We should stress that we tested other schemes in the validation of the models' performance (cross-fold validation). We omitted one scheme (data not reported). We achieved similar and consistent performance results with these tests.

The results of this study show that it is possible to partially automate the detection and classification of malignancy in suspicious areas during AFB examinations. However, the final interpretation of suspicious areas depends on the decision of an expert in the field.

6 REFERENCES

- [1] Wang, K.P., Mehta, A.C., Turner Jr, J.F. (eds.) (2011). *Flexible Bronchoscopy*. John Wiley & Sons, Chichester.
- [2] Wang, Y., Wang, Q., Feng, J., Wu, Q. (2013). Comparison of auto fluorescence imaging bronchoscopy and white light bronchoscopy for detection of lung cancers and precancerous lesions. *Patient Preference and Adherence*, vol. 7, p. 621-631, DOI:10.2147/PPA.S46749.
- [3] Liu, Z., Zhang, Y., Li, Y.P., Ma, J., Shi, F., Zhao, D.F., Li, J.M., Zhang, Y.Z. (2016). Clinical relevance of using autofluorescence bronchoscopy and white light bronchoscopy in different types of airway lesions. *Journal of Cancer Research and Therapeutics*, vol. 12, no. 1, p. 69-72, DOI:10.4103/0973-1482.147731.
- [4] van den Bergh, H. (2003). Early detection of lung cancer and the role of endoscopic fluorescence imaging. *Medical Laser Application*, vol. 18, no. 1, p. 20-26, DOI:10.1078/1615-1615-00083.
- [5] Bezel, P., Tischler, V., Robinson, C., Baumueller, S., Bode-Lesniewska, B., Kohler, M., Freitag, L., Franzen, D. (2016). Diagnostic value of bronchoalveolar lavage for diagnosis of suspected peripheral lung cancer. *Clinical Lung Cancer*, vol. 17, no. 5, p. e151-e156, DOI:10.1016/j.clcc.2015.12.012.
- [6] Edell, E., Lam, S., Pass, H., Miller, Y.E., Sutedja, T., Kennedy, T., Loewen, G. Keith, R.L. (2009). Detection and localization of intraepithelial neoplasia and invasive carcinoma using fluorescence-reflectance bronchoscopy: an international, multicenter clinical trial. *Journal of Thoracic Oncology*, vol. 4, no. 1, p. 49-54, DOI:10.1097/JTO.0b013e3181914506.
- [7] Häussinger, K., Becker, H., Stanzel, F., Kreuzer, A., Schmidt, B., Strausz, J., Cavaliere, S., Herth, F., Kohlhäufel, M., Müller, K.-M., Huber, R.-M., Pichlmeier, U., Bolliger, Ch.T. (2005). Autofluorescence bronchoscopy with white light bronchoscopy compared with white light bronchoscopy alone for the detection of precancerous lesions: a European randomised controlled multicentre trial. *Thorax*, vol. 60, no. 6, p. 496-503, DOI:10.1136/thx.2005.041475.
- [8] Moghissi, K., Dixon, K., Stringer, M.R. (2008). Current indications and future perspective of fluorescence bronchoscopy: a review study. *Photodiagnosis and Photodynamic Therapy*, vol. 5, no. 4, p. 238-246, DOI:10.1016/j.pdpdt.2009.01.008.
- [9] McGregor, H.C., Short, M.A., McWilliams, A., Shaipanich, T., Ionescu, D.N., Zhao, J., Wang, W., Chen, G., Lam, S. Zeng, H. (2016). Real-time endoscopic Raman spectroscopy for in vivo early lung cancer detection. *Journal of Biophotonics*, vol. 10, no. 1, p. 98-110, DOI:10.1002/jbio.201500204.
- [10] Leong, S., Shaipanich, T., Lam, S., Yasufuku, K. (2013). Diagnostic bronchoscopy—current and future perspectives. *Journal of Thoracic Disease*, vol. 5, no. 5, p. S498-S510, DOI:10.3978/j.issn.2072-1439.2013.09.08.
- [11] Lee, P., van den Berg, R.M., Lam, S., Gazdar, A.F., Grunberg, K., McWilliams, A., LeRiche, J., Postmus, P.E., Sutedja, T.G. (2009). Color fluorescence ratio for detection of bronchial dysplasia and carcinoma in situ. *Clinical Cancer Research*, vol. 15, no. 14, p. 4700-4705, DOI:10.1158/1078-0432.CCR-08-1644.
- [12] Cajal, C., Santolaria, J., Samper, D., Garrido, A. (2015). Simulation of laser triangulation sensors scanning for design and evaluation purposes. *International Journal of Simulation Modelling*, vol. 14, no. 2, p. 250-264, DOI:10.2507/IJSIMM14(2)6.296.
- [13] Jerbic, B., Nikolic, G., Chudy, D., Svaco, M., Sekoranja, B. (2015). Robotic application in neurosurgery using intelligent visual and haptic interaction. *International Journal of Simulation Modelling*, vol. 14, no. 1, p. 71-84, DOI:10.2507/IJSIMM14(1)7.290.
- [14] Martinović, M., Stoić, A., Kiš, D. (2008). Segmentation of the CT image using self-organizing neural networks. *Tehnički vjesnik – Technical Gazette*, vol. 15, no. 4, p. 23-28.
- [15] Bard, M.P., Amelink, A., Skurichina, M., Hegt, V.N., Duin, R.P., Sterenborg, H.J., Hoogsteden, H., Aerts, J.G. (2006). Optical spectroscopy for the classification of malignant lesions of the bronchial tree. *Chest*, vol. 129, no. 4, p. 995-1001, DOI:10.1378/chest.129.4.995.
- [16] Zeng, H., Petek, M., Tercelj Zorman, M., McWilliams, A., Palcic, B., Lam, S. (2004). Integrated endoscopy system for simultaneous imaging and spectroscopy for early lung cancer detection. *Optics Letters*, vol. 29, no. 6, p. 587-589, DOI:10.1364/OL.29.000587.
- [17] Tercelj, M., Zeng, H., Petek, M., Rott, T., Palcic, B. (2005). Acquisition of fluorescence and reflectance spectra during routine bronchoscopy examinations using the ClearVu Elite™ device: Pilot study. *Lung Cancer*, vol. 50, no. 1, p. 35-42, DOI:10.1016/j.lungcan.2005.05.028.
- [18] Goujon, D., Zellweger, M., Radu, A., Grosjean, P., Weber, B.-C., van den Bergh, H., Monnier, P., Wagnières, G. (2003). In vivo autofluorescence imaging of early cancers in the human tracheobronchial tree with a spectrally optimized system. *Journal of Biomedical Optics*, vol. 8, no. 1, p. 17-25, DOI:10.1117/1.1528594.
- [19] Qu, J.Y., Chang, H., Xiong, S. (2002). Fluorescence spectral imaging for characterization of tissue based on multivariate statistical analysis. *JOSA A*, vol. 19, no. 9, p. 1823-1831, DOI:10.1364/JOSAA.19.001823.
- [20] Gabrecht, T., Radu, A., Zellweger, M., Lovisa, B., Goujon, D., Grosjean, P., van der Bergh, H., Monnier, P., Wagnières, G. (2007). Autofluorescence bronchoscopy: Clinical experience with an optimized system in head and neck cancer patients. *Medical Laser Application*, vol. 22, no. 3, p. 185-192, DOI:10.1016/j.mla.2007.09.003.
- [21] Bountris, P., Apostolou, A., Haritou, M., Passalidou, E., Koutsouris, D. (2009). Combined texture features for improved classification of suspicious areas in autofluorescence bronchoscopy. *9th International Conference on Information Technology and Applications in Biomedicine*, p. 1-4, DOI:10.1109/ITAB.2009.5394448.

- [22] Bountris, P., Haritou, M., Passalidou, E., Apostolou, N., Koutsouris, D. (2009). Detection and classification of suspicious areas in autofluorescence bronchoscopy. *World Congress on Medical Physics and Biomedical Engineering*, p. 1838-1841, DOI:10.1007/978-3-642-03882-2_488.
- [23] Haritou, M., Bountris, P., Passalidou, E., Koklonis, K., Koutsouris, D. (2014). An image analysis tool for the classification of lesions suspicious for malignancy in autofluorescence bronchoscopy. *Biomedical Spectroscopy and Imaging*, vol. 3, no. 2, p. 167-183, DOI:10.3233/BSI-140079.
- [24] Chambers, O., Milenkovic, J., Tasic, J.F. (2014). A pre-processing scheme for real-time registration of dynamic contrast-enhanced magnetic resonance images. *Journal of Real-Time Image Processing*, p. 1-10, DOI:10.1007/s11554-014-0468-0.
- [25] Sled, J.G., Zijdenbos, A.P., Evans, A.C. (1997). A comparison of retrospective intensity non-uniformity correction methods for MRI. *Biennial International Conference on Information Processing in Medical Imaging*, p. 459-464, DOI:10.1007/3-540-63046-5_43.
- [26] Brinkmann, B.H., Manduca, A., Robb, R.A. (1998). Optimized homomorphic unsharp masking for MR grayscale inhomogeneity correction. *IEEE Transactions on Medical Imaging*, vol. 17, no. 2, p. 161-171, DOI:10.1109/42.700729.
- [27] Ahmed, M.N., Yamany, S.M., Mohamed, N., Farag, A.A., Moriarty, T. (2002). A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Transactions on Medical Imaging*, vol. 21, no. 3, p. 193-199, DOI:10.1109/42.996338.
- [28] Li, C., Huang, R., Ding, Z., Gatenby, J.C., Metaxas, D.N., Gore, J.C. (2011). A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI. *IEEE Transactions on Image Processing*, vol. 20, no. 7, p. 2007-2016, DOI:10.1109/TIP.2011.2146190.
- [29] Fawzy, Y.S., Petek, M., Tercej, M., Zeng, H. (2006). In vivo assessment and evaluation of lung tissue morphologic and physiological changes from non-contact endoscopic reflectance spectroscopy for improving lung cancer detection. *Journal of Biomedical Optics*, vol. 11, no. 4, p. 044003, DOI:10.1117/1.2337529.
- [30] Perona, P., Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, p. 629-639, DOI:10.1109/34.56205.
- [31] Weickert, J. (1998). *Anisotropic Diffusion in Image Processing*, p. 59-60. Teubner, Stuttgart.
- [32] Kovesi, P.D. (2000). MATLAB and Octave functions for computer vision and image processing. from <http://www.csse.uwa.edu.au/~pk/Research/MatlabFns/#match>, accessed on 2016-05-25.
- [33] Jenko, M. (2015). Numerical cooking for pasteurized soft boiled eggs. *Strojniški vestnik - Journal of Mechanical Engineering*, vol. 61, no. 5, p. 319-329, DOI:10.5545/sv-jme.2014.2187.
- [34] Maini, R., & Aggarwal, H. (2009). Study and comparison of various image edge detection techniques. *International Journal of Image Processing*, vol. 3, no. 1, p. 1-11.
- [35] Theodoridis, S., Pikrakis, A., Koutroumbas, K., Cavouras, D. (2010). *Introduction to Pattern Recognition: A Matlab Approach*. Academic Press, Oxford.
- [36] Duda, R.O., Hart, P.E., Stork, D.G. (2012). *Pattern Classification*. John Wiley & Sons, New York.
- [37] Hofmann, M., Klinkenberg, R. (eds.). (2013). *RapidMiner: Data Mining Use Cases and Business Analytics Applications*. Chapman & Hall/CRC Data Mining and Knowledge Discovery Series. CRC Press, Boca Raton.
- [38] Molinaro, A.M., Simon, R., Pfeiffer, R.M. (2005). Prediction error estimation: a comparison of resampling methods. *Bioinformatics*, vol. 21, no. 15, p. 3301-3307, DOI:10.1093/bioinformatics/bti499.