Barrett's Esophagus Identification Using Color Co-occurrence Matrices

Luis A. de Souza Jr.
Department of Computing
Federal University of São Carlos
São Carlos, Brazil
E-mail:luis.souza@dc.ufscar.br

Alanna Ebigbo, Andreas Probst, Helmut Messmann
Klinikum Augsburg III
Medizinische Klinik
Augsburg, Germany
E-mail: {alana.ebigbo, andreas.probst, helmut.messmann}
@klinikum-augsburg.de

João P. Papa
Department of Computing
São Paulo State University
Bauru, Brazil
E-mail:papa@fc.unesp.br

Robert Mendel, Christoph Palm
Regensburg Medical Image Computing
Ostbayerische Technische Hochschule Regensburg
Regensburg, Germany
E-mail: christoph.palm@oth-regensburg.de

***Abstract*—In this work, we propose the use of single channel Color Co-occurrence Matrices for texture description of Barrett's Esophagus (BE) and adenocarcinoma images. Further classification using supervised learning techniques, such as Optimum-Path Forest (OPF), Support Vector Machines with Radial Basis Function (SVM-RBF) and Bayesian classifier supports the context of automatic BE and adenocarcinoma diagnosis. We validated three approaches of classification based on patches, patients and images in two datasets (MICCAI 2015 and Augsburg) using the color-and-texture descriptors and the machine learning techniques. Concerning MICCAI 2015 dataset, the best results were obtained using the blue channel for the descriptors and the supervised OPF for classification purposes in the patch-based approach, with sensitivity nearly to 73% for positive adenocarcinoma identification and specificity close to 77% for BE (non-cancerous) patch classification. Regarding the Augsburg dataset, the most accurate results were also obtained using both OPF classifier and blue channel descriptor for the feature extraction, with sensitivity close to 67% and specificity around to 76%. Our work highlights new advances in the related research area and provides a promising technique that combines color and texture information, allied to three different approaches of dataset pre-processing aiming to configure robust scenarios for the classification step.**

***Keywords*-Barrett's Esophagus; Co-occurrence Matrices; Machine Learning; Texture Analysis**

I. Introduction

The incidence of Barrett's esophagus (BE) and Barrett's adenocarcinoma in the west of the globe have risen significantly in the past decade. Because of their close association with the metabolic syndrome, this trend is expected to continue rising in the next years [1], [2], [3]. The early diagnosis of Esophageal adenocarcinoma (EA) is critical for the diseases' remission and justifies the necessity of efficient surveillance, detection and characterization. However, the detection of dysplastic regions and their characterization of abnormalities within BE-diagnosed patients can be challenging, especially for endoscopists presenting lack of experience for the evaluation. Even considering the dangerousness of the disease, when detected at the early stages, the injured tissue can be treated with very high rates of remission (93% after 10 years of treatment) [2], [4], [5].

The computer-aided analysis of BE may be one powerful instrument and has been subject of intensive research in the past years [6]. Up to now, mainly handcrafted features of endoscopic images based on texture and color were extracted and classified subsequently. For instance, Van der Sommen [7] designed a system for the automatic extraction of features for detecting and delineating early neoplastic tissue regions in patients diagnosed with BE, followed by some other relevant studies in the same field [8], [9], [10] that aimed to assess the feasibility of adenocarcinoma classification in endoscopic images of BE diagnosed patients. Souza et al. [11] also conducted a study introducing two approaches to distinguish between BE and adenocarcinoma: (i) the Optimum-Path Forest (OPF) [12], [13] classifier; and (ii) the use of Bag-of-Visual-Words [14], [15] using points-of-interest extracted from endoscopic images using Speed-Up Robust Features [16] and Scale Invariant Feature Transform [17] techniques for the feature vector calculation [6].

There are, in addition, some image processing techniques that can describe the image in different ways, providing feature vectors based on color or texture of the injured region. One of these techniques is the Co-occurrence Matrix (CM), which usually employs gray-scale images to encode texture information. However, there are some new approaches considering the influence of both color and texture for the CM calculation that can provide different descriptions for the BE and adenocarcinoma context [18].

Considering the growth of studies in which BE and adenocarcinoma evaluation is performed by means of machine learning and image processing techniques, the main contribution of this paper is the evaluation of Color Co-occurrence Matrices for the description of the dysplastic tissue in BE

diagnosed patients. Such assessment provides a novel BE and adenocarcinoma identification approach in which color and texture information can be combined to improve the classification results. Some previous works already evaluated the impact of color and texture information independently for the BE and adenocarcinoma description [7], [19], [6], showing promising results. However, the use of both phenomena in a single descriptor has been poorly studied in this context.

The remainder of this paper is organized as follows: Section II presents a brief background about color and texture combination using Co-occurrence Matrices. Section III discusses the methodology employed in this work, and Section IV presents the experimental results. Finally, Section V states discussions and future works.

II. Theoretical Background

A. Color and texture combination

The "parallel concept" [18] for color and texture analysis considers both phenomena for data description separated. While color is measured globally by means of histogram calculation, the texture is characterized by the relationship of the intensities of neighboring pixels ignoring their color. The processing of both information, i.e., color and texture, is performed independently, being combined subsequently to compose a final feature vector (Figure 1, (a)). The parallel approach can present advantages; however, the view on texture as a structure purely based on intensity is simplified.

The "sequential concept" [18] uses color analysis as a first means, in which the pattern is composed of segmented color primitives obtained by clustering the color histogram. Some previous works showed how useful the sequential approach could be for some tasks, such as industrial quality control and defect detection in granite images [20], [21]. However, the concept of colored texture primitives may not provide generalization support (Figure 1, (b)).

In the "integrative concept" [18], the information dependency between both color and texture is taken into account for feature extraction purposes. There are two strategies for the integrative color-texture combination: single- and multi-channel (Figure 1, (c)). The single-channel method analyses the gray-scale texture on each color channel separately, providing a subtle use for color information restricting the intensity pattern to the wavelength interval associated with that color channel [18]. The single channel approach is suitable for methods based on the gray-scale domain. These concepts have been proposed for well-known textural feature description, such as Wavelet-based [22], Gabor filters [23], [24], [25] and Markov Random Fields [26], showing very promising results through the years.

B. Gray-scale and Single-Channel Co-occurrence Matrices

Co-occurrence Matrices are defined as the relationship between the values of a central pixel p and its neighboring $\eta(p)$ [27]. Given a gray-scale image I, a pixel p contains two information: its value $I(p) \in [0, 255]$ and its position $p = (m, n)$, such that $m, n \in \mathbb{N}$.

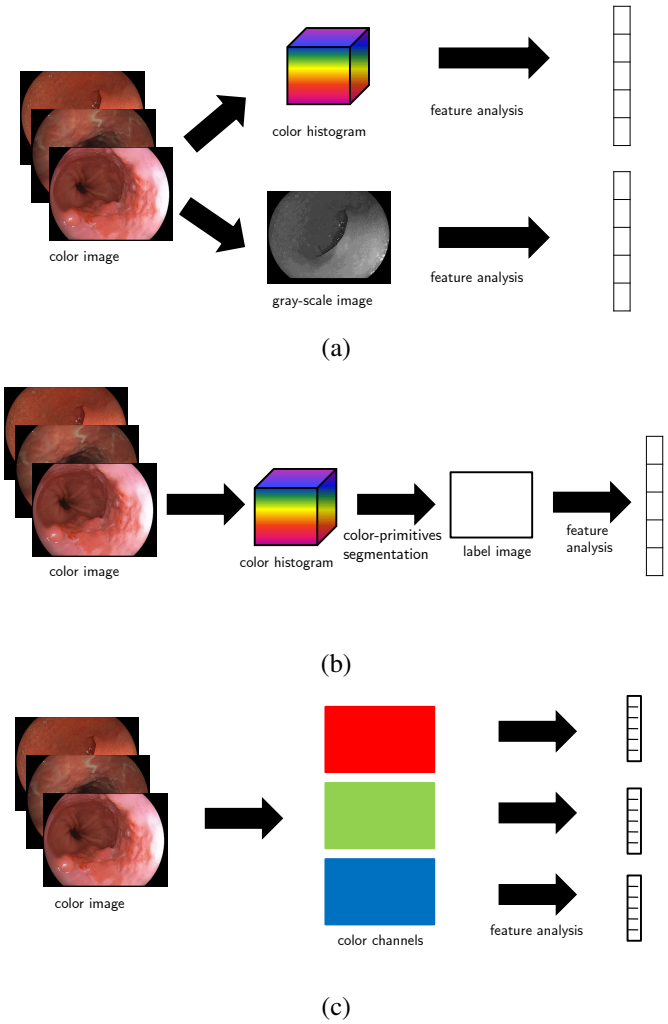

Fig. 1. Color and texture concepts: (a) parallel concept for color texture analysis; (b) sequential concept for color texture analysis and (c) integrative single-channel color texture analysis (adapted from [18]).

Let η_p be the neighborhood of p such that $p^* \in \eta_p$ when $d(p, p^*) \leq \mathbf{D}$, in which $d(p, p^*)$ stands for the polar distance between p and $p^\star$. Let C^d be a co-occurrence matrix defined over distance d such that each element is computed as follows:

$$C^d_{i,j} = P(I(p) = i \wedge I(p^*) = j), \tag{1}$$

such that $p^* \in \eta_p$. In other words, $C^d_{i,j}$ encodes the probability P of transition between brightness values from i to j. Additionally, it is well known that one must compute one co-occurrence matrix for each orientation angle.

Since C^d is symmetric for each orientation angle according to d, Palm [18] proposed to combine the different co-occurrence matrices into a single one that encodes distinct orientation angles. Also, eight Haralick features [27] were extracted (homogeneity, contrast, correlation, variance, inverse difference moment, entropy, correlation I, correlation II) and distributed over the four feature groups proposed by Gotlieb and Kreyszig [28].

Such approach allows the use of large values for **D**, which is basically the radius of a discrete circle. Computing CMs for four different angles and constant radius **D**, one can obtain 8×4 rotationally-dependent features. In order to be rotationally independent, we compute the mean and variance of each Haralick feature, thus ending up with an 8×2 dimensional gray-scale CM feature (GCF) space.

The Single-Channel Co-occurrence Matrices (SCMs) [18] stand for the successively use of gray-scale CMs in each k color channel separately (for RGB system, k = 1,2,3). Such matrices are computed using the very same Equation 1, but applied to each color channel. Thus, the corresponding rotational invariant single-channel Co-occurrence features (SCFs) consist of K feature vectors SCF^k presenting analogous behavior of GCF according to k. Therefore, the evaluation becomes a k-dimensional problem, once each k color-channel will provide a different descriptor to be analyzed. The advantage comes with the possibility of evaluation of each color channel independently, analyzing its impact on the texture information composition in a combined color and intensity texture information. The information profit by analyzing intensity independent color textures is quite high, being a brand new way of evaluation for color- and texture-based problems.

III. Methodology

In this section, we describe the datasets, pre-processing and feature extraction procedures, classification techniques and approaches employed in this work.

A. Datasets

1) MICCAI Dataset: The experiments were conducted using the dataset of BE-and adenocarcinoma-diagnosed patients provided at the "MICCAI 2015 EndoVis Challenge"[1]. Such dataset is composed of 100 endoscopic images of the lower esophagus from 39 individuals, in which 22 present BE and 17 present early-stage adenocarcinoma. For each patient, a different number of samples was available (ranging from one to eight), with a total of 50 samples showing BE and cancerous tissue areas and 50 images showing only BE without cancer. The injured tissue in cancerous images has been delineated by five different endoscopy experts. Figure 2 shows some samples and their respective delineation performed by the experts.

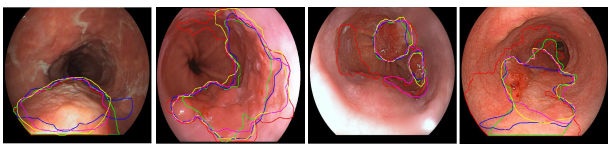

Fig. 2. Four MICCAI database samples with their respective delineations provided by five different experts.

[1]https://endovissub-barrett.grand-challenge.org/

2) Augsburg Dataset: A dataset provided by the Augsburg Klinikum, Medizinische Klinik III, was also used for the experiments. Such dataset is composed of 76 endoscopic images (esophagus) obtained from different patients with adenocarcinoma (34 samples) and BE (42 samples). The images were annotated (manual segmentation of the adenocarcinoma's and Barrett's areas) by an expert from the Augsburg Klinikum. The ground-truth diagnosis was validated by biopsy process. Some Augsburg dataset samples can be observed in Figure 3.

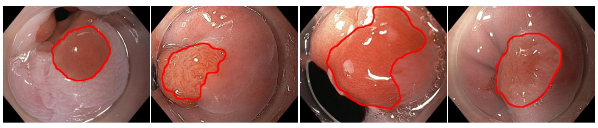

Fig. 3. Four Augsburg database samples with their respective delineations provided by the expert.

B. Pre-processing

Concerning the pre-processing step, the images were split into patches to be used in different approaches for classification purposes. Considering that databases present different image resolutions, the patch size was chosen in order to cover the entire image without overlapping. Regarding the MICCAI database, each image was split into 48 patches of 200×200 pixels, resulting in $4,800$ non-overlapped patches for the whole database. The Augsburg database images were split into 20 patches per image, with sizes of 270×270 and resulting in $1,520$ non-overlapped patches.

Additionally, it is important to notice that for the patch labeling process, the experts' annotations were considered for both datasets. Concerning the MICCAI database, the intersection area of the five experts' delineations was considered the correct adenocarcinoma region including the fuzzy delineation area (area of confusion among the delineations). Concerning patches that cross this region, the numbers of cancerous and non-cancerous pixels were compared, being the final label defined by the majority of pixels inside the patch. With respect to the Augsburg data, we used the only delineated area available. Notice that an analogous procedure for labeling patches was employed as well.

C. Feature Extraction

In order to consider a color-and-texture evaluation protocol for the automatic identification of BE and adenocarcinoma, the integrative single-channel co-occurrence matrix was applied. For each color channel, SCMs and Haralick features were computed and further used for learning purposes. Notice that the same set of features were extracted from the gray-scale images, which were obtained using the mean pixel values of each channel.

The experimental protocol was composed of three distinct evaluations: (i) first, we considered color and texture information from each channel separately, (ii) then the same set of features (i.e., from each channel) were concatenated to produce a single feature vector, and (iii) the color and texture extraction

techniques used previously were also considered for the gray-scale images.

Since each color channel and gray-scale feature vector consist of 16 elements, the composite descriptors (RGB) comprise 48 features (i.e., 3×16). Further, the Principal Component Analysis (PCA) was applied to reduce the number of features to a 16-dimensional space using the Single Value Decomposition and the covariance matrix approach for the largest eigenvalues definition, thus ending up with the same single-channel descriptor dimension. For all approaches, the SCMs were calculated with three different radii (i.e., 1, 5 and 10) to assess the impact on their representation concerning the classification results.

D. Classification

After the feature extraction using the integrative single-channel co-occurrence matrices, the descriptors from the databases were used as input to the following supervised learning techniques:

- OPF: supervised classifier with complete graph proposed by Papa et al. [12], [13];
- SVM-RBF: Support Vector Machines with Radial Basis Function kernel and parameters optimized by cross-validation [29];
- Bayes: standard Bayesian classifier.

Regarding the OPF and SVM-RBF classifiers, we used the open-source libraries LibOPF [30] and LibSVM [29], respectively. With respect to the Bayesian classifier, we employed our own implementation.

E. Approaches

This work employs three different approaches for the database pre-processing and classification: patch-based, patient patch-based, and image-based approach. Regarding the patch-based approach, 80% of all patches were randomly selected for training, while 20% of the remaining ones were used for testing purposes, being such partitioning process employed for 20 runs for both databases. Therefore, the patch-based classification step was conducted to discriminate between patches from BE and adenocarcinoma classes without taking into account information about the patients. Concerning MICCAI dataset, $3,840$ patches were randomly selected for the training step and 960 patches were used for the testing set. With respect to the Augsburg database, the training set was composed of $1,216$ random patches, and the test set was composed of the 304 remaining patches.

Concerning the patient patch-based approach, the patient information was used for the patch selection protocol. The available number of patches for this approach was the same available for the previous one (patch-based approach), but the difference here was related to the protocol applied to the definition of training and testing sets. In this experiment, we used the well-known "leave-one-patient-out cross-validation" (LOPO-CV), i.e., $n-1$ patients are used for training and the remaining one is used to evaluate the model, where n stands for the number of patients. This procedure is repeated until all patients have been evaluated.

Finally, the last experiment, i.e., image-based approach, uses the same 20-fold cross-validation protocol applied to the first approach (i.e., patch-based) with 80% for training and 20% for testing purposes. However, the descriptors were now extracted from the full images. Notice that the same protocol was applied to the Augsburg database. Figure 4 illustrates the approaches mentioned above.

IV. Experiments

In this section, we present the experiments used to evaluate the three proposed approaches. The discrimination between positive and negative samples to adenocarcinoma was performed using OPF, SVM-RBF and Bayesian classifier (hereinafter called "Bayes"). The results are presented and discussed for each approach and database. All experiments were conducted on an 8Gb-memory computer equipped with an Intel® Core i5 - 2.30 GHz processor. Additionally, we employed our implementation of the SCM approach in C++ language.

In this work, we adopted the following sensitivity (S), specificity (P), accuracy (A), and $F1$ Score ($F1$) measures:

$$S = \frac{TP}{TP + FN} * 100, \tag{2}$$

$$P = \frac{TN}{TN + FP} * 100, \tag{3}$$

$$A = \frac{TP + TN}{TP + TN + FP + FN} * 100, \tag{4}$$

and

$$F1 = \frac{2 \cdot S \cdot P}{S + P}, \tag{5}$$

where TP and TN stand for the true positives and true negatives, respectively, and FN and FP denote the false negatives and false positives, respectively.

Table I presents the average results regarding sensitivity, specificity, accuracy, and $F1$ score concerning the patch-based approach. Since we considered different values for the radius used in SCM technique, column "Radius" contains the values that lead to the best results. With respect to the Augsburg database, the best results were obtained using the OPF classifier in the blue channel, with sensitivity, specificity, accuracy and $F1$ values of 66.6%, 75.6%, 70.2%, and 70.8%, respectively. Concerning the SVM-RBF classifier, the best results were obtained on the red channel, with values of 58.6% of sensitivity, 87.5% of specificity, 81.4% of accuracy, and 70.2% of $F1$ score. The Bayesian classifier provided the best results in the red channel features as well, with values of 58.1% for sensitivity, 83.5% of specificity, 75.3% of accuracy, and 68.5% of $F1$.

Since the $F1$ score presents a trade-off response between the sensitivity and specificity values, the values in bold stand for the best approaches with respect to such measure. The accuracy can be very dependent on the sensitivity/specificity

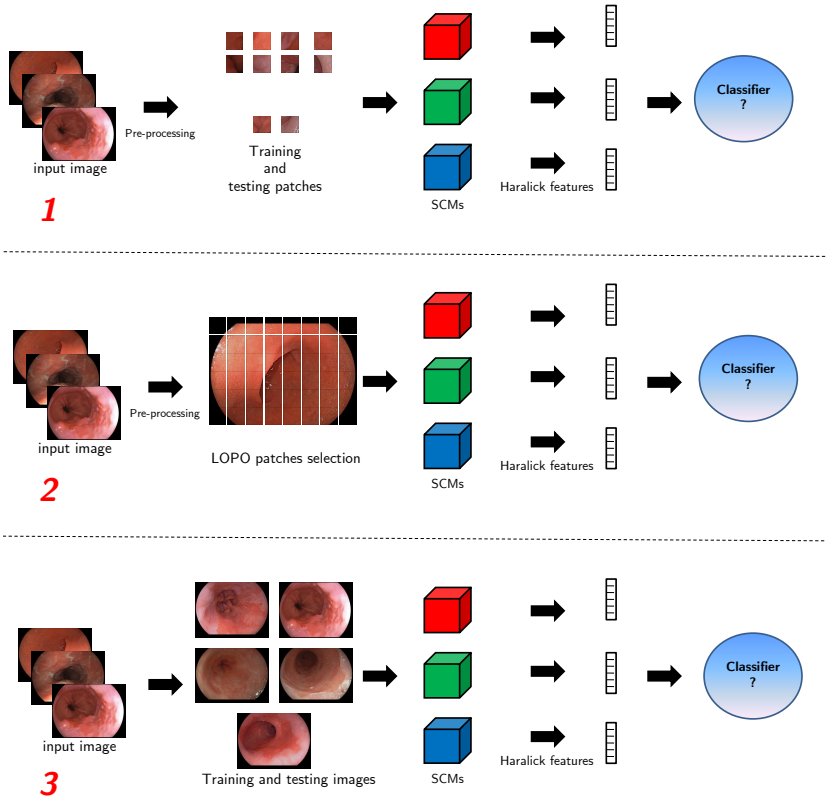

Fig. 4. Approaches used in the experiments: (1) patch-based, (2) patient patch-based, and (3) image-based approach.

values when we have a unbalanced database. We observed that specificity always influenced considerably the accuracy values, once the number of samples positive to cancer was usually lower when compared to the negative examples.

Concerning the MICCAI database, OPF classifier provided the best results in the blue channel once again, with values of sensitivity, specificity, accuracy, and $F1$ equal to 72.6%, 77.2%, 76.6%, and 74.8%, respectively. The SVM-RBF classifier showed the best results in the blue channel as well, with values of 63.9%, 76.7%, 70.1%, and 69.7% for sensitivity, specificity, accuracy, and $F1$ score, respectively. The Bayes classifier obtained the best results in the PCA-RGB features, with sensitivity, specificity, accuracy, and $F1$ values of 58.3%, 80.3%, 75.6%, and 67.6%, respectively.

Table II presents the average results of sensitivity, specificity, and accuracy concerning the patient patch-based approach. Regarding the Augsburg database, the best results considering the $F1$ values were obtained using the OPF classifier in the blue channel, with sensitivity, specificity, accuracy and $F1$ values of 63.0%, 77.4%, 73.5%, and 69.5%, respectively. With respect to the SVM-RBF classifier, the best results were obtained in the green channel, with values of 63.6% for sensitivity, 74.4% for specificity, 70.9% for accuracy, and 68.6% for $F1$. The Bayes classifier provided reasonable results in the green channel as well, with values of 65.8% for sensitivity, 70.5% for specificity, 69.1% for accuracy, and 68.1% for the $F1$ score.

Regarding MICCAI dataset, the OPF classifier provided the best results in the blue channel, with values of sensitivity, specificity, accuracy and $F1$ Score equal to 71.6%, 72.9%, 72.3%, and 72.2%, respectively. The SVM-RBF classifier obtained the best results using the PCA-RGB features, with values of 61.9%, 79.5%, 75.4%, and 69.6% of sensitivity, specificity, accuracy and $F1$, respectively. The Bayes classifier with the PCA-RGB features provided the best results as well, with values of sensitivity, specificity, accuracy and $F1$ of 60.1%, 82.4%, 78.6%, and 69.1%, respectively.

Table III presents the average results of sensitivity, specificity, accuracy and $F1$ score concerning the classification in the image-based protocol. Concerning the Augsburg dataset, the best results were obtained using the OPF classifier in the red channel, with sensitivity, specificity, accuracy and $F1$ values of 63.5%, 66.2%, 64.5%, and 64.82%, respectively. Using the SVM-RBF classifier, the best results were obtained using the PCA-RGB features, with values of 37.7% for sensitivity, 83.4% for specificity, 66.9% for accuracy and 51.93% for the $F1$ score. The Bayes classifier obtained its best results with the PCA-RGB features as well, with values of 55.5% for sensitivity, 77.6% for specificity, 62.7% for accuracy, and 64.72% for $F1$.

With respect to the MICCAI database, OPF classifier obtained the best results in the red channel, with values of sensi-

tivity, specificity, accuracy and $F1$ equal to 60.7%, 70.8%, 68.4%, and 65.36%, respectively. The SVM-RBF classifier achieved the best results using the PCA-RGB features, with values of 49.5%, 78.2%, 71.1%, and 60.62% for sensitivity, specificity, accuracy and $F1$ values, respectively. The Bayes classifier obtained the best results in the red channel, with values of sensitivity, specificity, accuracy and $F1$ equal to 61.7%, 73.5%, 67.0%, and 67.1%, respectively.

V. Discussion and Conclusions

In this paper, we dealt with the problem of computer-assisted Barrett's esophagus and esophageal adenocarcinoma evaluation using endoscopy images. BE stands for an illness that is visually confused with adenocarcinoma, requiring more precise ways for its early detection and prevention.

We observed that only a very few works attempted at coping with the problem of automatic BE identification using image processing and machine learning techniques to date. In this work, we fostered the research towards such area and introduced the use of single channel Color Co-occurrence Matrices in the feature extraction step for automatic BE recognition, as well as we showed how each RGB channel could perform compared to the gray-scale image evaluation. The experimental results were considered over two databases: (i) MICCAI 2015 and (ii) Augsburg. For both scenarios, we evaluated three different approaches and supervised learning techniques for classification purposes.

As one can observe in the previous section, each approach presents a particular and interesting result that deserves attention. Considering the patch-based approach (Table I), the results over Augsburg data highlighted that cancerous patches are harder to be identified than non-cancerous ones, thus explaining low values of sensitivity and higher values of specificity. However, the use of blue-channel SCMs associated with the OPF classifier provided the higher sensitivity and $F1$ values. With respect to MICCAI database, the results presented a similar behavior to those obtained over Augsburg ones. The blue-channel SCMs combined with the OPF classifier achieved the higher $F1$ scores when compared to the other classifiers, thus suggesting it can be a strong learning technique for color-and-texture feature classification. For both databases, the number of non-cancer patches was considerably higher than the cancerous ones, thus explaining the higher specificity values. Although SVM-RBF and Bayes obtained satisfactory results, they were outperformed by OPF.

Concerning the patient patch-based approach (Table II), the best results over Augsburg data were obtained with the OPF classifier and blue channel SCMs. The MICCAI dataset results showed a better performance, with the best values achieved by OPF and blue-channel SCMs as well. Both SVM-RBF and Bayes classifiers were outperformed by OPF. Once again, SVM-RBF provided better sensitivity values in some experiments, but not the overall best results.

Since the image-based approach (Table III) makes use of descriptors obtained from the entire image, the tendency is to achieve the worst results due to the presence of both cancerous and non-cancerous regions in the very same image labeled as cancer. Surprisingly, the results presented satisfactory classification rates for both databases. Concerning the Augsburg and MICCAI databases, the best results were achieved with the OPF classifier and red-channel SCMs. The SVM-RBF showed the worst performance in this approach, with low sensitivity but high specificity values. The image-based approach demonstrates the generalization strength of the SCMs over the entire images. Even with different regions in cancer-labeled images, the obtained feature vector can provide good generalization for the classification step.

Considering all approaches, it is relevant to point out that the results using color-channel features outperformed the gray-scale ones in all experiments, thus corroborating the importance of the color-texture analysis. The blue channel results obtained for the patch and patient-based approaches suggest that, for local evaluation, the blue channel present a more accurate and robust way of description, while the red channel may provide a better global evaluation of the BE and adenocarcinoma problem, according to the global results provided by the image-based approach. The blue results corroborate the ones obtained by Ilgner et al. [31] in which laryngoscopy images presenting or not diseased tissue were classified using colored-texture descriptors, being blue the color channel that provided the best classification rate (81.4%). PCA-RGB features showed a very well performance, achieving the best results between the channels for some classifiers and approaches. With respect to the SCMs, feature vectors calculated with large distances (i.e., higher values of $\mathbf{D}$) showed better results when compared to short distance ones (in this case, $\mathbf{D} = 5$ or $\mathbf{D} = 10$ always presented better results than the ones obtained using $\mathbf{D} = 1$). Such premise is also relevant, reinforcing the importance of the neighboring information during the CMs calculation, suggesting that higher values of $\mathbf{D}$ may provide better generalization abilities for classification purposes. The experiments pointed out that SCMs are suitable to handle BE automatic identification, and there must be a trade-off between the sensitivity and specificity values to compose a cohesive diagnosis result for the BE and adenocarcinoma distinction.

Concerning the previous results obtained for the BE and adenocarcinoma classification in the literature, this one, in particular, proposes a new protocol of image evaluation, in which the images are split into patches, so the labeling problem is changed. It is well-known that once we work with full images, cancerous and non-cancerous regions receive the same label, but when the problem is extended to the patches, the labeling problem becomes less critical. With the patches labeling, even with images that present BE and adenocarcinoma, the misclassification of regions do not exist, once each patch will be labeled according to the previous annotation provided by the experts. Regarding the classification results, handling with patches may improve the results because of this accurate labeling definition that does not happen in the full-image approach. The comparison with previous works can be performed with the third approach (image-based approach)

TABLE I
MEAN VALUES CONCERNING THE PATCH-BASED APPROACH.

		OPF					SVM-RBF					Bayes				
Database	Channel	S	P	A	$F1$	Radius	S	P	A	$F1$	Radius	S	P	A	$F1$	Radius
Augsburg	gray	53.7	87.5	80.2	66.6	10	59.9	82.7	77.2	69.5	10	55.5	75.7	70.9	64.4	5
	red	54.4	82.7	77.3	65.7	10	58.6	87.5	81.4	**70.2**	10	58.1	83.5	75.3	**68.5**	10
	green	43.4	93.5	87.8	59.3	5	47.4	90.8	87.2	62.3	10	39.5	89.7	86.8	54.9	10
	blue	66.6	75.6	72.0	**70.8**	5	60.2	77.8	71.0	67.9	10	53.5	80.2	75.3	64.2	10
	PCA-RGB	57.7	83.2	73.3	68.1	5	55.7	82.3	76.7	66.4	10	56.8	81.4	76.8	66.9	10
MICCAI	gray	64.3	85.5	83.4	73.4	10	66.7	79.5	79.1	68.8	5	47.8	78.4	73.4	59.4	10
	red	51.2	90.4	87.5	65.4	10	47.9	91.1	86.9	62.8	10	54.7	85.3	80.0	66.7	10
	green	49.7	90.4	82.7	65.3	5	51.4	90.8	89.1	65.6	10	41.5	87.3	81.5	56.3	5
	blue	72.6	77.2	76.6	**74.8**	10	63.9	76.7	70.1	**69.7**	10	50.0	86.7	81.3	63.4	10
	PCA-RGB	65.1	82.1	77.5	72.6	10	57.5	80.4	73.2	67.0	10	58.3	80.3	75.6	**67.6**	10

TABLE II
MEAN VALUES CONCERNING THE PATIENT PATCH-BASED.

		OPF					SVM-RBF					Bayes				
Database	Channel	S	P	A	$F1$	Radius	S	P	A	$F1$	Radius	S	P	A	$F1$	Radius
Augsburg	gray	52.3	71.0	59.4	60.2	10	54.0	78.2	72.2	63.9	5	50.5	73.4	64.9	59.8	10
	red	52.6	67.4	68.3	59.1	10	56.1	70.9	64.7	62.6	10	54.8	63.9	60.1	59.0	10
	green	70.2	67.2	68.1	68.7	10	63.6	74.4	70.9	**68.6**	10	65.8	70.5	69.1	**68.1**	5
	blue	63.0	77.4	73.5	**69.5**	10	59.7	78.7	72.9	67.9	10	52.0	77.6	70.4	62.3	10
	PCA-RGB	56.2	83.4	76.3	67.2	10	61.4	72.3	66.8	66.4	10	55.2	76.8	68.3	64.2	10
MICCAI	gray	54.3	82.8	73.6	65.6	10	60.9	76.7	67.8	67.9	10	47.0	74.1	63.2	57.5	5
	red	55.3	81.0	64.6	65.7	5	50.2	81.9	62.6	62.2	5	55.1	71.7	60.8	62.3	5
	green	50.9	84.5	71.1	63.5	10	50.0	85.2	73.9	63.0	5	49.9	81.4	69.5	61.9	10
	blue	71.6	72.9	72.3	**72.2**	10	63.8	75.5	69.1	69.1	10	59.6	77.2	69.1	67.3	10
	PCA-RGB	71.1	71.7	71.5	71.4	10	61.9	79.5	75.4	**69.6**	10	60.1	82.4	78.6	**69.1**	5

TABLE III
MEAN RESULT VALUES CONCERNING THE IMAGE-BASED APPROACH.

		OPF					SVM-RBF					Bayes				
Database	Channel	S	P	A	$F1$	Radius	S	P	A	$F1$	Radius	S	P	A	$F1$	Radius
Augsburg	gray	55.5	60.1	57.5	57.7	10	33.7	84.5	63.8	48.2	5	30.7	64.5	51.3	41.6	5
	red	63.5	66.2	64.5	**64.8**	10	29.7	89.5	67.6	44.6	10	57.1	55.6	56.3	56.3	5
	green	52.8	64.2	59.0	57.9	10	30.3	85.5	64.7	44.7	10	46.8	71.8	59.9	56.7	5
	blue	44.9	69.7	58.8	54.6	10	33.9	85.9	65.6	48.6	10	50.3	53.6	50.1	51.9	5
	PCA-RGB	52.0	64.2	57.8	57.5	10	37.7	83.4	66.9	**51.9**	10	55.5	77.6	62.7	**64.7**	5
MICCAI	gray	59.7	52.7	55.5	56.0	10	50.3	55.4	50.7	52.7	10	47.5	64.8	56.3	54.8	5
	red	60.7	70.8	68.4	**65.4**	10	38.9	86.6	79.2	53.7	10	61.7	73.5	67.0	**67.1**	10
	green	54.9	69.7	64.8	61.4	10	37.8	83.9	73.2	52.1	10	44.4	84.6	72.3	58.2	10
	blue	57.8	63.1	59.6	60.3	10	35.8	86.5	76.4	50.6	5	55.5	60.4	58.4	57.9	10
	PCA-RGB	60.0	66.3	65.2	63.0	10	49.5	78.2	71.1	**60.6**	10	54.9	79.4	72.5	64.9	5

for the MICCAI 2015 database. Souza Jr. et al. conducted two different works with a similar image-based evaluation in such database, but using SURF [9] and SIFT [11] descriptors associated to a large number of classifiers. The accuracy, sensitivity and specificity results obtained in this work could not outperform the ones using SURF and SIFT features for the full-images approach. However, this work provides two very important contributions: (i) the introduction of the single-channel co-occurrence matrices technique for the BE and adenocarcinoma description and (ii) the evaluation based on patch-based approaches in which very promising results could be provided, suggesting that the proposed descriptor, associated with a local representation of the problem, can become a strong resource for the BE and adenocarcinoma context evaluation.

In regard to future works, we aim at considering four major new tasks: (i) the multi-channel implementation of co-occurrence matrices [18] instead of the single-channel approach used in this work; (ii) the reduction of the feature vector dimensionality using feature selection techniques, and (iii) the use of the methodology used in this work as an end-to-end approach to aid physicians during the diagnosis process and; (iv) a scale-evaluation of the BE and adenocarcinoma context using the proposed color co-occurrence matrice descriptor for many levels of image scale, providing descriptors based on a scale-space approach.

ACKNOWLEDGMENT

The authors thank DFG grant PA 1595/3-1, Capes/Alexander von Humboldt Foundation grant number BEX 0581-16-0, CNPq grants 306166/2014-3 and 307066/2017-7, as well as FAPESP grants 2013/07375-0, 2014/12236-1, and 2016/19403-6.

REFERENCES

[1] J. Lagergren and P. Lagergren, "Oesophageal cancer," *BMJ*, vol. 341, 2010.

[2] J. Dent, "Barrett's esophagus: a historical perspective, an update on core practicalities and predictions on future evolutions of management," *Journal of Gastroenterology and Hepatology*, vol. 26, pp. 11–30, 2011.

[3] C. Lepage, B. Rachet, and V. Jooste, "Continuing rapid increase in esophageal adenocarcinoma in england and wales," *The American Journal of Gastroenterology*, vol. 103, pp. 2694–2699, 2008.

[4] P. Sharma, J. J. G. H. M. Bergman, K. Goda, M. Kato, and *et al*, "Development and validation of a classification system to identify high-grade dysplasia and esophageal adenocarcinoma in barretts esophagus using narrow-band imaging," *Gastroenterology*, vol. 150, no. 3, pp. 591 – 598, 2016.

[5] K. N. Phoa, R. E. Pouw, R. Bisschops, O. Pech, K. Ragunath, B. L. A. M. Weusten, and *et al*, "Multimodality endoscopic eradication for neoplastic barrett oesophagus: results of an european multicentre study (euro-ii)," *Gut*, vol. 65, no. 4, pp. 555–562, 2016.

[6] L. A. de Souza Jr., C. Palm, R. Mendel, C. Hook, A. Ebigbo, A. Probst, H. Messmann, S. W., and J. P. Papa, "A survey on barrett's esophagus analysis using machine learning," *Computers in Biology and Medicine*, vol. 96, pp. 203–213, 2018.

[7] F. van der Sommen, S. Zinger, W. L. Curvers, R. Bisschops, O. Pech, and *et al*, "Computer-aided detection of early neoplastic lesions in barretts esophagus," *Endoscopy*, vol. 48, no. 07, pp. 617–624, 2016.

[8] A. R. Hassan and M. A. Haque, "Computer-aided gastrointestinal hemorrhage detection in wireless capsule endoscopy videos," *Computers in Biology and Medicine*, vol. 122, pp. 341–353, 2015.

[9] L. A. Souza Jr., C. Hook, J. P. Papa, and . Palm, "Barrett's esophagus analysis using SURF features," in *Bildverarbeitung für die Medizin 2017*. Berlin, Heidelberg: Springer, 2017, pp. 141–146.

[10] R. Mendel, A. Ebigbo, A. Probst, H. Messmann, and C. Palm, "Barrett's esophagus analysis using convolutional neural networks," in *Bildverarbeitung für die Medizin 2017*. Berlin, Heidelberg: Springer, 2017, pp. 80–85.

[11] L. A. Souza Jr, L. C. S. Afonso, C. Palm, and J. P. Papa, "Barrett's esophagus identification using optimum-path forest," in *30th SIBGRAPI Conference on Graphics, Patterns and Images*, 2017, pp. 308–314.

[12] J. P. Papa, A. X. Falcão, and C. T. N. Suzuki, "Supervised pattern classification based on optimum-path forest," *International Journal of Imaging Systems and Technology*, vol. 19, no. 2, pp. 120–131, 2009.

[13] J. P. Papa, A. X. Falcão, V. H. C. Albuquerque, and J. M. R. S. Tavares, "Efficient supervised optimum-path forest classification for large datasets," *Pattern Recognition*, vol. 45, no. 1, pp. 512–520, 2012.

[14] G. Csurka, C. R. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual categorization with bags of keypoints," in *Proceedings of the Workshop on Statistical Learning in Computer Vision*, 2004, pp. 1–22.

[15] X. Peng, L. Wang, X. Wang, and Y. Qiao, "Bag of visual words and fusion methods for action recognition: Comprehensive study and good practice," *Computer Vision and Image Understanding*, vol. 150, pp. 109–125, 2016.

[16] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer Visision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[17] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[18] C. Palm, "Color texture classification by integrative co-occurrence matrices," *Pattern Recognition*, vol. 37, no. 5, pp. 965 – 976, 2004.

[19] L. M. Almond and H. Barr, "Advanced endoscopic imaging in barrett's oesophagus," *International Journal of Surgery*, vol. 10, no. 5, pp. 236 – 241, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1743919112000635

[20] M. Hauta-Kasari, J. Parkkinen, T. Jskelinen, and R. Lenz, "Multi-spectral texture segmentation based on the spectral cooccurrence matrix," vol. 2, pp. 275–284, 11 1999.

[21] K. Song, J. Kittler, and M. Petrou, "Defect detection in random colour textures," *Image and Vision Computing*, vol. 14, no. 9, pp. 667 – 683, 1996.

[22] G. V. de Wouwer, P. Scheunders, S. Livens, and D. V. Dyck, "Wavelet correlation signatures for color texture characterization," *Pattern Recognition*, vol. 32, no. 3, pp. 443 – 451, 1999.

[23] A. Jain and G. Healey, "A multiscale representation including opponent color features for texture recognition," vol. 7, pp. 124 – 128, 02 1998.

[24] C. Palm and T. M. Lehmann, "Classification of color textures by gabor filtering," *Machine GRAPHICS & VISION*, vol. 11, pp. 195–219, 12 2003.

[25] G. Paschos, "Perceptually uniform color spaces for color texture analysis: An empirical evaluation," *IEEE Transactions on Image Processing*, vol. 10, pp. 932 – 937, 07 2001.

[26] P.-H. Suen and G. Healey, "Modeling and classifying color textures using random fields in a random environment," *Pattern Recognition*, vol. 32, no. 6, pp. 1009 – 1017, 1999.

[27] R. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, pp. 610–621, 01 1973.

[28] C. C. Gotlieb and H. E. Kreyszig, "Texture descriptors based on co-occurrence matrices," *Computer Vision, Graphics, and Image Processing*, vol. 51, no. 1, pp. 70 – 86, 1990.

[29] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1–27, 2011, software available at https://www.csie.ntu.edu.tw/ cjlin/libsvm/.

[30] J. P. Papa, C. T. N. Suzuki, and A. X, "LibOPF: A library for the design of optimum-path forest classifiers," software version 2.1 available at http://www.ic.unicamp.br/ afalcao/libopf/index.html.

[31] J. Ilgner, C. Palm, A. G. Schütz, K. Spitzer, M. Westhofen, and T. M. Lehmann, "Colour texture analysis for quantitative laryngoscopy." *Acta oto-laryngologica*, vol. 123 6, pp. 730–734, 2003.