Universidade do Minho

Mestrado Integrado em Engenharia Biomédica

Trabalho Prático de Imagiologia
ALGORITMOS DE PROCESSAMENTO DE IMAGENS - OPENCV

Features for the classification of GI tract endoscopic images

Marta Moreira [54459]

Abstract

Computer-aided diagnosis represents a milestone in modern medicine practice, enclosing invaluable short and long-term potential in what regards providing assistance to medical experts and thus improving the quality and accuracy of medical diagnosis. This emergent field strongly relies on the extraction of image descriptive features through computer vision and image processing techniques. In this paper, relevant texture, shape and color features for the classification of gastrointestinal tract endoscopic images are proposed, based on the computation of Gray Level Co-Occurrence Matrices (GLCM), spatial moments and Hue-Saturation color histograms from an experimental data set composed of 78 images. The extracted features are fed to a Support Vector Machine (SVM) classifier, resulting in a combined rate of correctly classified instances of 95% for binary classification and 71% for multi-class classification. The performance of the classifier can potentially be improved by increasing the number of available images, subjecting those images to pre-processing routines and correctly tuning the parameters of the classifier.

1. Introduction

Since medical endoscopy stands as a minimally invasive and relatively painless procedure that allows inspecting the inner cavities of the human body, endoscopes play a key role in the modern practice of medicine spanning the inspection of areas such as the respiratory tract, the female reproductive system and, most commonly, the gastrointestinal tract, among others. Based on the endoscopy of the gastrointestinal tract, physicians are able to detect severe conditions in early development stages and thus contribute to lower the mortality rate for a number of different diseases. Modern endoscopes are regularly used to take digital pictures and record video sequences, abilities that created the field of decision support systems in medical endoscopy [1]. Such systems aim at predicting pathologies and thus assisting medical experts in improving the accuracy of medical diagnosis [2], and strongly rely on the extraction of image descriptive features through computer vision and image processing techniques. In the particular case of gastrointestinal tract images, the most relevant image properties that may lead to an effective discrimination between distinctive gastrointestinal regions essentially include texture, shape and color. While color and shape are quite tangible approaches, the concept of texture is more abstract and subjectively defined and interpreted; however, it embodies valuable information to identify or describe an image [3], describing the structural arrangement of surfaces and their relationship to the surrounding environment. This is the case especially for gastroenterology, where the internal membranes of the digestive tract exhibit strong texture features and distinctive patterns, as shown in Fig. 1 [4].

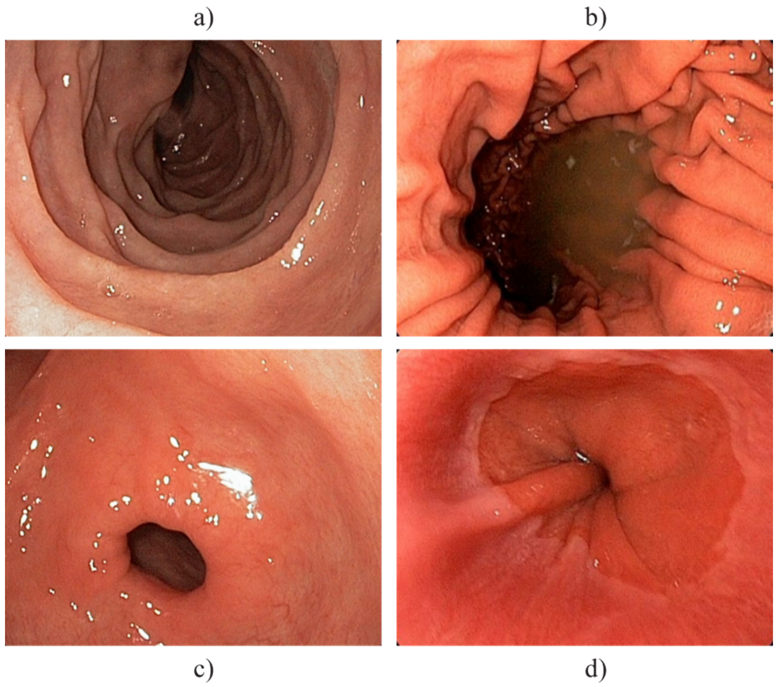

Fig. 1. Digestive tract distinct textures and patterns: a) duodenum, b) body, c) pylorus and d) cardia.

Features for the classification of the gastrointestinal tract images can be roughly grouped into low-level features and high-level features. Low-level features can be further subdivided into spatial and frequency domain features, depending on whether its extraction is direct or involves some sort of transformation of the data into the frequency domain, respectively. Concerning spatial domain features, the simplest approaches use information such as pixel color [5] or a combination of pixel colors with the respective pixel position [6]. Other rather simple techniques use histogram color channels or images, either for direct similarity-based classification or for the indirect extraction of statistical features like those proposed by Haralick et al. in [3], which are then used for classification purposes. Local texture properties are also popular in the field of gastrointestinal endoscopy and are commonly captured by operators such as Local Binary Patterns (LBP) [7], Texture Spectrum (TS) [8] and Scale Invariant Feature Transform (SIFT) [9]. On the other hand, frequency domain features targeted at this field are heavily based on the Discrete Wavelet Transform (DWT) [10], including some of its variants such as the Dual-tree Complex Wavelet Transform (DTCWT) [11], the Curvelet Transform [12] or the Stationary Wavelet Transform (SWT) [13]. Despite the undisputed dominance of wavelet features, other frequency transforms such as the Fast Fourier Transform (FFT) and the Discrete Cosine Transform (DCT) have also been used to extract frequency domain features [1]. High level features usually describe geometrical properties of shapes extracted from images, describing them in a more abstract way and thus using edge detection algorithms like the Canny edge detector [14] or the SUSAN edge detector [15]. Based on such edge image features, different features describing shapes may be used for the classification of endoscopic images [1].

2. Feature Extraction

2.1. Texture Features

Texture analysis can be grouped into four categories: model-based, transform-based, structural-based and statistical-based methods. Model-based approaches seek to predict pixel values based on a mathematical model, while transform methods generally modify the image to some extent in order to obtain a new "response" image that is then analysed as a representative proxy for the original image. On the other hand, structural approaches aim for the understanding of the hierarchal structure of the image and statistical methods describe the image using pure numerical analysis of pixel intensity values [16]. In this paper, the texture of gastrointestinal tract images was explored relying on a specific statistical approach proposed in [3], using co-occurrence matrices, which has previously been proven to feature a potential for effective texture discrimination in biomedical images [17,18].

2.1.1. Gray Level Co-Occurrence Matrix

The Gray Level Co-Occurrence Matrix (GLCM) contains information about the intensities of pixels and their neighbours at fixed distance and orientation, hence exploring the gray level spatial dependencies of texture through the construction of a co-occurrence matrix based on the distance and orientation between image pixels and the subsequent extraction of meaningful statistics from the matrix as texture representation. The concept lies upon scanning the image while iteratively keeping track of the gray levels of each of the two pixels separated by a fixed distance d and direction θ, as shown in Fig. 2. Since one distance and one direction are generally not enough to describe texture features, common practice is to use five distances and four directions. As suggested in [3], distances of 1 through 5 pixels and directions of 0º, 45º, 90º and 135º were used, meaning that each image was represented by 20 matrices. In order to ensure computational efficiency without compromising the outcome [19], the images were first converted to 32 gray levels — leading to 32x32 matrices — without any further preprocessing. The following step involved the extraction of statistical meaningful values from the information enclosed in the matrices, namely the twenty-two features proposed in [3].

a)

0	0	1	2
0	2	2	3
2	3	3	2
1	1	2	3

b)

(i,j)	0	1	2	3
0	#(0,0)	#(0,1)	#(0,2)	#(0,3)
1	#(1,0)	#(1,1)	#(1,2)	#(1,3)
2	#(2,0)	#(2,1)	#(2,2)	#(2,3)
3	#(3,0)	#(3,1)	#(3,2)	#(3,3)

2	1	1	0
1	2	2	0
1	2	2	4
0	0	4	2

c)

2	0	0	0
0	0	1	2
0	1	6	1
0	2	1	2

d)

2	0	2	0
0	0	2	1
2	2	0	6
0	1	6	0

e)

0	0	2	1
0	0	1	1
2	1	2	2
1	1	2	2

f)

Fig. 2. a) Image example, b) construction of co-occurrence matrix and matrix frameworks for c) 0º d) 45º, e) 90º and f) 135º.

2.2. Shape Features

Shape representation generally looks for effective and perceptually important shape features based on either shape boundary information or boundary plus interior content. Various features have been designed, including shape signature, signature histogram, shape invariants, moments and spectral features, among others [20].

2.2.1. Spatial Moments

In particular, spatial moments constitute well-known pattern recognition tools, employed especially in shape analysis, due to their proven sensitivity to spatial details. Depending on the order and type of the chosen moments, different kind of information may be extracted from the input, while the exact effect of these choices on the computed features remains to be investigated [21]. In this paper, spatial moments of order 0 through 4 were retrieved, as well as some parameters which directly derive from them, like elongation and orientation.

2.3. Color Features

Despite the color constancy effect of the human color perception system — whereby the color perception of objects remains relatively constant under varying environmental and visual conditions [22] —, serious color casts exist because of the intervention of an endoscope or camera between the gastrointestinal tract tissue and the physician's eye. For these reasons, color and texture information are usually combined in the interpretation of endoscopic images [4]. Many approaches to color representation have been proposed in the past few years, such as color histogram, color moments and color sets. In this paper, the color histogram approach in the HSV color space was chosen as the color feature, as the histogram is relatively easy to extract and the HSV color space has de-correlated and uniform coordinates which better match the human perception of color, separating both the chromatic and achromatic components. Furthermore, since the value (V) component is easily affected by the lightning conditions, only hue (H) and saturation (S) are used to build a two-dimensional histogram [23].

2.3.1. Hue-Saturation Color Histogram

In most image editing software, a three dimensional representation of the HSV color space is a hexacone, where the central vertical axis represents the intensity component. Hue, which stands for color, is specified as a circular angle in the range $[0, 2\pi]$ relative to the red axis with red at angle 0, green at $2\pi/3$, blue at $4\pi/3$ and red again at 2π; while saturation, which stands for the depth or purity of the color, is measured as a radial distance from the central axis with value between 0 at the centre to 1 at the outer surface [24]. However, the HSV format representation in OpenCV differs from this standard, since the RGB-to-HSV conversion only stores the hue component as an 8-bit integer in the range of 0 to 179. This means that using the cvCvtColor() function to obtain a HSV image compromises the color resolution, as it basically stores hue as a 7-bit number instead of an 8-bit number. Taking this limitation into account, the Hue-Saturation color histogram was constructed (Fig. 3) using the 30 hue levels and 32 saturation levels suggested in the OpenCV documentation and, subsequently, the pixel count for each level was extracted and used to represent the color features.

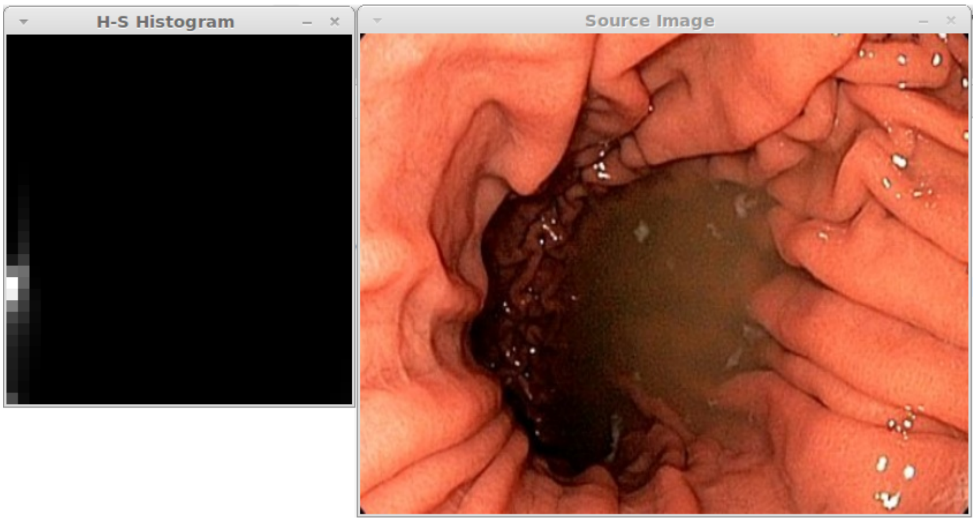

Fig. 3. Hue-Saturation histogram representation as computed in OpenCV.

3. Classification

The modern computer-aided diagnosis systems commonly rely on machine learning techniques to learn hypotheses from a large amount of diagnosed samples and hence extract regularity or some sort of structure from that collection of data. Neural networks (NN) and Bayesian Classifiers are the typical instances to learn such organization from given data observations. A seemingly widespread classifier in the context of supervised classification is the Support Vector Machines (SVM).

3.1.1. Support Vector Machines

The Support Vector Machines (SVM) classifier is based on strong foundations from the broad area of statistical learning theory and has become the classifier of choice of numerous researchers and practitioners for several real-world problems, due to the good degree of accuracy it has demonstrated when faced with the prediction of the unseen or unknown samples, comparing to traditional classifiers as the aforementioned. Some of the advantages of this classifier rely on the facts that it is computationally much less intensive, performs well in higher dimensional spaces and deals satisfactorily with the lack of training data. SVM attempts to evaluate a linear decision boundary or a linear hyperplane between two classes of linearly separable data. Despite being fundamentally developed for binary class supervised classification problems, SVM is extendable for multi-class situations since, theoretically, there exist possibly an infinite number of hyperplanes when the data is linearly separable [25]. In this paper, WEKA‘s implementation of the SVM algorithm (SMO) was used to classify each set of experimental data considering both the binary and the multi-class approaches.

4. Experimental Results

In order to assess whether the GLCM extracted features performed best when isolated or when combined with additional shape and color features, four different experimental data sets were created and stored in an individual Comma Separated Values (CSV) file, considering the isolation of the GLCM values (GLCM) and its combination with either spatial moments (GLCM+SM), Hue-Saturation histogram features (GLCM+HS) or both (GLCM+SM+HS). Subsequently, each experimental data set was individually fed to a SVM classifier employing 11-fold cross validation for evaluation; the remainder parameters adopted the default setting of the SMO package in WEKA. Table 1 encloses the classification rate of correctly classified instances obtained considering a binary classification — where images were bound to the ‘Duodenum’ and ‘Not Duodenum’ classes — and a multi-class classification — involving eight distinct gastrointestinal regions.

Table 1. Accuracy obtained through the binary and multi-class classification of each of the considered experimental sets

Experimental Set	Binary	Multi-class
GLCM	93.5897 %	67.9487 %
GLCM+SM	92.3077 %	62.8205 %
GLCM+HS	93.5897 %	65.3846 %
GLCM+SM+HS	94.8718 %	70.5128 %

Based on the percentage results shown, the combination of GLCM, spatial moments and Hue-Saturation color features amounts to the highest rate of correctly classified instances for both binary and multi-class classifications. This means that out of 78 gastrointestinal tract endoscopy images, 4 and 23 instances are incorrectly classified, respectively. In what concerns GLCM features isolation, it is noticeable that these features perform better on their own than associated with the others, when provided individually.

5. Conclusions and Future Work

In this paper, experimental results describing feature extraction for the classification of gastrointestinal tract endoscopic were presented, based on the computation of Gray Level Co-Occurrence Matrices (GLCMs), spatial moments and Hue-Saturation color histograms for a given set of images. The classification rate of correctly classified instances for both binary classification and multi-class classification suggests that GLCM features could be used as a standalone descriptor of the image texture, although it greatly benefits from the integration of additional features such as shape and color descriptors, especially due to the fact that endoscopic images contain both texture and color information. Regardless of being provided with a limited number of samples, the classifier managed to achieve an acceptable degree of accuracy, which leads to the conclusion that better results can be achieved with a wider set of images. Other details that could be addressed in order to enhance the performance of the classifier are related to the correct tuning of its parameters and also the subjection of images to a pre-processing routine such as histogram equalization or the application of an adequate filter. Ideally, the whole process would be integrated in a single executable program, as to promote the automation of the three feature extraction techniques employed, since the chosen approach proved to be extremely time consuming.

References

[1] Liedlgruber, M. Computer-Aided Classification Endoscopic Images From the Gastrointestinal Tract. PhD Dissertation, University of Salzburg, 2011.

[2] Doi, K. Computer-Aided Diagnosis in Medical Imaging: Historical Review, Current Status and Future Potential. Computerized Medical Imaging and Graphics, vol.4-5, no. 31, pp. 192–211, 2007.

[3] Haralick, R.M. et. al. Textural Features for Image Classification. Journal of Science Communication, vol. SMC-3, no. 6, pp. 610-621, 1973.

[4] Vasileios, C. et al. Enhanced Ulcer Recognition from Capsule Endoscopic Images Using Texture Analysis. New Advances in the Basic and Clinical Gastroenterology, 2012.

[5] Pan, G. et al. Bleeding Detection in Wireless Capsule Endoscopy Based on Probabilistic Neural Network. Journal of Medical Systems, pp. 1–8, 2010.

[6] Alexandre, L. et al. Polyp Detection in Endoscopic Video Using SVMs. In Knowledge Discovery in Databases: PKDD 2007, pp. 358–365. Springer Berlin, 2007.

[7] Ojala, T. et al. A comparative Study of Texture Measures With Classification Based on Feature Distributions. Pattern Recognition, vol.1, no. 29, pp. 51–59, 1996.

[8] Wang, L. and He, D. Texture Classification Using Texture Spectrum. Pattern Recognition, vol. 8, no. 23, pp. 905–910, 1990.

[9] Lowe, D.G. Distinctive Image Features From Scale-Invariant Keypoints. International Journal of Computer Vision, vol. 2, no. 60, pp. 91–110, 2004.

[10] Mallat, S. A Wavelet Tour Of Signal Processing. Academic Press, 2nd Edition, 1999.

[11] Selesnick, I.W. et al. The Dual-Tree Complex Wavelet Transform - a Coherent Framework for Multiscale Signal and Image Processing. IEEE Signal Processing Magazine, vol. 6, no. 22, pp. 123–151, 2005.

[12] Pesquet, J. et al. Time Invariant Orthonormal Wavelet Representations. IEEE Transactions on Signal Processing, vol. 8, no. 44, pp. 1964–1970, 1996.

[13] Candes, E.J. and Donoho, D.L. Curvelets, Multiresolution Representation and Scaling Laws. In Proceedings of the SPIE – Wavelet Applications in Signal and Image Processing VIII, vol. 4119, pp. 277–282, 2000.

[14] Canny, J. A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 6, no. 8, pp. 679–698, 1986.

[15] Smith, S. and Brady, J. SUSAN - a New Approach to Low Level Image Processing. International Journal of Computer Vision, vol. 1, no. 23, pp. 45–78, 1997.

[16] Mohamed, S.S. and Salama, M.M. Computer-Aided Diagnosis for Prostate Cancer Using Support Vector Machines. In Proceedings of the Medical Imaging Conference, California, 2005.

[17] Tjoa, M.P. and Krishnan, S.M. Feature Extraction for the Analysis of Colon Status from the Endoscopic Images. Biomedical Engineering Online, pp. 1-5, 2003.

[18] Kaminishi, M. et al. Endoscopic Classification of Chronic Gastritis Based On A Pilot Study By The Research Society For Gastritis. Digestive Endoscopic, vol.14, pp. 138-151, 2002.

[19] Celebi, E. and Alpkocak, A. Clustering of Texture Features for Content-Based Image Retrieval**.** In Proceedings of the First International Conference on Advances in Information Systems, pp. 216-225. Springer-Verlag London, 2000.

[20] Zhang, D., Lu, G. Review of Shape Representation and Description Techniques. Pattern Recognition, no. 37, pp. 1 – 19, 2004.

[21] Aptoula, E., Lefèvre, S. Spatial morphological covariance applied to texture classification. In Proceedings of the 2006 International Conference on Multimedia Content Representation, Classification and Security, pp. 522-529, Springer-Verlag Berlin, 2006.

[22] Foster, D. Does Color Constancy Exist? TRENDS in Cognitive Sciences, 2003.

[23] Ortega, M. et al. Supporting Ranked Boolean Similarity Queries in MARS. Knowledge and Data Engineering, vol. 10, no. 6, pp. 905 – 925, 1998.

[24] Sural, S. et al. Segmentation and Histogram Generation Using the HSV Color Space for Image Retrieval. In Proceedings of International Conference on Image Processing, pp. 589 – 592, 2002.

[25] Rajpoot, K.M. and Rajpoot, N.M. Wavelets and Support Vector Machines for Texture Classification. *Proceedings of INMIC 2004,* pp.328-333, 2004.