

# Histological Image Analysis by Invariant Descriptors

Cecilia Di Ruberto, Andrea Loddo, and Lorenzo Putzu<sup>(✉)</sup>

Department of Mathematics and Computer Science,  
University of Cagliari, Via Ospedale 72, Cagliari, Italy  
{dirubert, andrea.loddo, lorenzo.putzu}@unica.it

**Abstract.** In this work we propose a comparative study between different descriptors in analysing histological images. In particular, our study is focused on measuring the accuracy of moments (Hu, Legendre, Zernike), Local Binary Patterns and co-occurrence matrices in classifying histological images. The experimentation has been conducted on well known public datasets: HistologyDS, Pap-smear, Lymphoma, Liver Aging Female, Liver Aging Male, Liver Gender AL and Liver Gender CR. The comparison results show that when combined with co-occurrence matrices and extracted from the RGB images, the orthogonal moments improve the classification performance considerably, imposing themselves as very powerful descriptors for histological image analysis.

**Keywords:** Medical image analysis · Texture descriptors · Moments · Local binary pattern · Co-occurrence matrix · Classification

## 1 Introduction

Histological image analysis is a process that allows to evaluate if microscopic structures at the sub-cellular, cellular, tissue and organs level are affected by diseases, through various computer assisted methods. Tissue image analysis could be used to measure the cancer cells in a biopsy of a cancerous tumour taken from a patient and it can significantly reduce uncertainty in characterizing tumours compared to evaluations done by histologists, or improve the prediction recurrence rate of some cancers. Image analysis involves complex algorithms which identify and characterize cellular colour, shape and quantity of the tissue sample using image pattern recognition technology. In [1] global features are used to automatically discriminate lymphoma, in [2] wavelet features are used for the detection of tumours in endoscopic images and in [3] image texture informations are used to automatically discriminate polyps in colonoscopy images. Over the past few years moment functions have been used in medical image analysis with promising performance. They are statistical measures used to obtain the relevant information of an object. Since the introduction of invariant moments in image analysis [4], moment functions have been widely used in image processing and pattern classification applications, as discriminative descriptors, such as

the geometric moments [5] for texture classification, or the complex moments for texture segmentation [6]. However, both geometric and complex moments contain redundant information and are sensitive to noise, due to the fact that the kernel polynomials are not orthogonal. For these reasons many different moments have been proposed, such as the discrete Tchebichef moments [7], the discrete moments known as Krawtchouk moments [8] or orthogonal moments like Legendre and Zernike moments [9].

Orthogonal moments are shown to be less sensitive to noise and have an efficient capability in feature representation with minimum redundancy. Zernike moments have been widely used in different types of applications, in shape-based image retrieval [10,11] and in pattern recognition [12] tasks. In medical image analysis the orthogonal moments have been used to reconstruct noisy CT, MRI, X-ray medical images [13], to describe the texture of a CT liver image [14] or prostate ultrasound [15], to detect tumours in brain images [16] or in mammography images [17], to recognize parasites [18] and spermatogonium [19].

In this work we propose a comparative study between different descriptors based on texture information for histological image classification. In particular, our study is focused on measuring the accuracy of moments (Hu, Legendre, Zernike), Local Binary Patterns (LPBs) and co-occurrence matrices in classifying histopathological images. The experimental results show that the combination of orthogonal moments with co-occurrence matrices reaches a very high level of accuracy on all the tested datasets, overcoming the most common and used descriptors. The rest of paper is organized as follows. Next section presents the texture descriptors definition used throughout this work. Section 3 presents the experimental evaluation, describing the utilized datasets, showing the experimental measures, the implementation details of each descriptor, and the results achieved for the specific collections. Finally, in Sect. 4 we give the conclusions.

## 2 Texture Descriptors

In this section we describe three important classes of texture descriptors: image geometric and orthogonal moments, cooccurrence matrices and local binary patterns.

### 2.1 Image Moments

The moments are widely used in many applications for features extraction due to their invariance to scale, rotation and reflection change. The use of moments for image analysis and pattern recognition was inspired by Hu [4]. Hu's, Legendre's and Zernike's are the most common moments.

*Hu moments.* They are derived and calculated from geometric moments. The two-dimensional geometric moments of order  $(p+q)$  of an image of  $M \times N$  pixels with intensity function  $f(x, y)$  are defined as:

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y), \quad (1)$$

where  $p, q = 0, 1, 2, \dots$ . A set of  $n$  moments consists of all  $m_{pq}$  for  $p + q \leq n$ . The corresponding central moments are defined as:

$$\mu_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q f(x, y), \tag{2}$$

where  $\bar{x} = m_{10}/m_{00}$  and  $\bar{y} = m_{01}/m_{00}$  are the coordinates of the centre of mass of the image. The central moments  $\mu_{pq}$  defined in Eq. 2 are invariant under the translation of coordinates. They can be normalized to preserve the invariance by scaling. For  $p + q = 2, 3, \dots$  the normalized central moments of an image are given by:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \quad \text{with} \quad \gamma = \frac{p + q}{2} + 1. \tag{3}$$

Hu defined seven functions that are invariant to scale, translation and rotation changes [4], from the normalized central moments through the order three.

*Legendre Moments.* Legendre moments are orthogonal moments first introduced by Teague [9]. They were used in several patterns recognition [4] tasks. The Legendre moment of order  $(p + q)$  of an image of  $M \times N$  pixels with intensity function  $f(x, y)$  is defined on the square  $[-1, +1] \times [-1, +1]$ , by:

$$L_{pq} = \frac{(2p + 1)(2q + 1)}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P_p(x_i) P_q(y_j) f(x_i, y_j) \tag{4}$$

where  $x_i$  and  $y_j$  denote the normalized pixel coordinates in the range of  $[-1, +1]$ , which are given by:

$$x_i = \frac{2i - (M - 1)}{M - 1}, \quad y_j = \frac{2j - (N - 1)}{N - 1} \tag{5}$$

and

$$P_p(x_i) = \sum_{k=0}^p \frac{(-1)^{\frac{p-k}{2}} x^k (p + k)!}{2^p k! (\frac{p-k}{2})! (\frac{p+k}{2})!} \tag{6}$$

with  $p - k$  even.

*Zernike Moments.* Zernike moments are the mapping of an image onto a set of complex Zernike polynomials. As these Zernike polynomials are orthogonal to each other, Zernike moments can represent the properties of an image with no redundancy or overlapping of information between the moments [9]. Due to these characteristics, Zernike moments have been used as features set in many applications. The computation of Zernike moments from an input image consists of three steps: computation of radial polynomials, computation of Zernike polynomials and computation of Zernike moments by projecting the image onto the Zernike polynomials [20]. The real-valued radial polynomial is defined as:

$$R_{p,q}(r) = \sum_{s=0}^{(p-|q|)/2} \frac{(-1)^s (p-s)! r^{p-2s}}{s! \left(\frac{p+|q|}{2} - s\right)! \left(\frac{p-|q|}{2} - s\right)!} \quad (7)$$

with  $R_{p,q}(r) = R_{p,-q}(r)$ , and  $p, q$  generally called order and repetition, respectively. The order  $p$  is a non-negative integer, and the repetition  $q$  is an integer satisfying  $p - |q|$  even and  $|q| \leq p$ . The discrete form of the Zernike moments of an image of size  $M \times N$  is expressed as follows:

$$Z_{pq} = \frac{p+1}{\lambda} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} R_{pq}(r_{xy}) e^{-jq\theta_{xy}} f(x, y) \quad (8)$$

where  $0 \leq r_{xy} \leq 1$  and  $\lambda$  is a normalization factor. In the discrete implementation of Zernike moments, the normalization factor  $\lambda$  must be the number of pixels located in the unit circle by the mapping transformation and corresponds to the area of a unit circle  $\pi$  in the continuous domain. The transformed  $\theta_{xy}$  phase and the distance  $r_{xy}$  at the pixel coordinates  $(x, y)$  are given by:

$$\theta_{xy} = \tan^{-1} \left( \frac{(2y - (N - 1))/(N - 1)}{(2x - (M - 1))/(M - 1)} \right) \quad (9)$$

$$r_{xy} = \sqrt{\left( \frac{2x - (M - 1)}{M - 1} \right)^2 + \left( \frac{2y - (N - 1)}{N - 1} \right)^2} \quad (10)$$

## 2.2 Co-occurrence Matrices

One of the earliest method for texture descriptors extraction was proposed by Haralick et al. [21]. His method is based on the creation of the grey level co-occurrence matrices, GLCMs, from which features representing some image aspects can be calculated. A GLCM represents the probability of finding two pixels  $i, j$  with distance  $d$  and orientation  $\theta$ . Obviously, the  $d$  and  $\theta$  values can assume different values, but the most used are  $d = 1$  and  $\theta = [0^\circ, 45^\circ, 90^\circ, 135^\circ]$ . A GLCM for an image of size  $M \times N$  with  $N_g$  grey levels is a 2D array of size  $N_g \times N_g$ . Haralick proposed thirteen descriptors that can be extracted from these matrices. Interesting methods have already been presented in order to extend the original implementation of GLCM. In [22] different values for the distance parameter influencing the matrices computation are evaluated, in [23] the GLCM descriptors are extracted by calculating the weighted sum of GLCM elements, in [24] the GLCM features are calculated by using the local gradient of the matrix. Furthermore, the GLCM has been extracted using the colour information from single channels [25] or by combining them in pairs [26, 27]. Considering that invariant descriptors have our main focus in this work, we compute the GLCM only using the grey level intensities, and convert the rotation-dependent descriptors in rotationally invariant by the following approach. We start considering all the possible circular shifts of a feature vector  $f_k = [f_1, \dots, f_m]$ . Then, we construct a matrix of size  $m \times m$  in which all the circular shifts of the vector  $f_k$  are present and disposed regularly, generating a symmetric matrix as follows:

$$\begin{pmatrix} f_1 & f_2 & \cdots & f_{m-1} & f_m \\ f_2 & \cdots & \cdots & f_m & f_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ f_{m-1} & f_m & \cdots & \cdots & \cdots \\ f_m & f_1 & \cdots & \cdots & f_{m-1} \end{pmatrix}$$

Hence, the eigenvalues of such matrix are the new invariant descriptors, GLCMri, as they preserve dimension and direction of the original feature vector.

### 2.3 LBP Descriptors

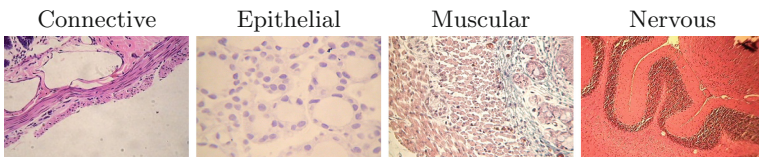
The LBPs, instead, are a quite recent tool for texture analysis, originally proposed in [28] and widely used for grey level texture classification, due to its simplicity and robustness. This operator transforms the image by thresholding the neighbourhood of each pixel and by coding the result as a binary number. The resulting image histogram can be used as a feature vector for texture classification. Moreover, radius and number of neighbourhood pixels are two main parameters needed for the LBP operator. Although the LBP have been extended in many different ways, the most useful version, proposed by the same authors [29], realizes a rotation invariant descriptor, called LBPri. The LBPri is easily obtained through an iterative rotation of the binary digits, until the smallest value has been reached.

## 3 Datasets

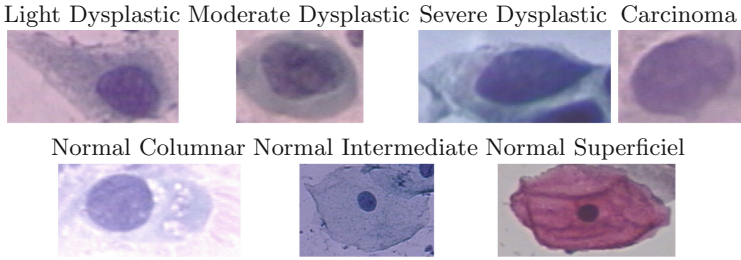
The experimentation has been carried out on seven of the most famous colour histology image databases: HistologyDS, Pap-smear, Lymphoma, Liver Aging Female, Liver Aging Male, Liver Gender AL and Liver Gender CR that represent a set of really different computer vision problems.

**HystologyDS** (HIS) database [30] is a collection of 20,000 histology images for the study of fundamental tissues. It is provided in a subset of 2828 images annotated by four fundamental tissues: connective, epithelial, muscular and nervous. Each tissue is captured in a 24-bit RGB image of size  $720 \times 480$ . Some sample tissue images from HIS database are showed in Fig. 1.

**Pap-smear** (PAP) database [31] is a collection of pap-smear images acquired from healthy and cancerous smears coming from the Herlev University Hospital. It is composed of 917 images containing cells, annotated into seven classes: four



**Fig. 1.** Four different tissues from HistologyDS database.

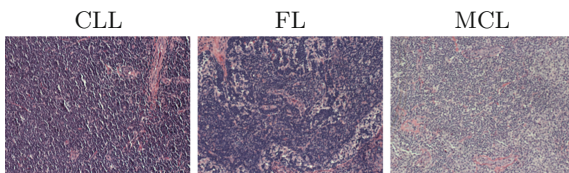


**Fig. 2.** The seven classes of cells belonging to Pap-smear database: first four abnormal and last three normal.

represent abnormal cells and three represent normal cases. Nevertheless, from the medical diagnosis viewpoint the most important requirement corresponds to the general two-class problem of correct separation between normal from abnormal cells. For this reason we have considered only the binary case. Each cell was captured in a 24-bit RGB image without a fixed size, that ranges from about  $50 \times 50$  to about  $300 \times 300$ . Some examples are showed in Fig. 2.

**Lymphoma (LYM)** database [1] is a collection of tissues affected by malignant lymphoma, a cancer affecting lymph nodes. Three types of malignant lymphoma are represented in the set: Chronic Lymphocytic Leukemia (CLL), Follicular Lymphoma (FL) and Mantle Cell Lymphoma (MCL). This dataset presents a collection of samples from biopsies sectioned and stained with (H&E), realized in different laboratories by several pathologists. Only the most expert pathologists specialised in these types of lymphomas are able to consistently and accurately classify these three lymphoma types from H&E-stained biopsies. This slide collection contains significant variation in sectioning and staining and for this reason it is more representative of slides commonly encountered in a clinical setting. This database contains a collection of 374 slides captured in a 24-bit RGB image of size  $1380 \times 1040$ . In Fig. 3 a randomly selected image from each class is showed.

**AGEMAP** Atlas of Gene Expression in Mouse Aging Project [32] is a study by the National Institute on Aging, involving 48 male and female mice, of four ages (1, 6, 16, and 24 months), on ad-libitum or caloric restriction diets. Fifty



**Fig. 3.** Three different kinds of lymphoma belonging to lymphoma database.

colour images from 30 livers were manually acquired using a Carl Zeiss Axiovert 200 microscope and 40x objective, for a total of 1500 images. Each image is of size  $1388 \times 1040$  in TIFF format with a 36-bit RGB colour depth. As the acquisition was done using 12 bits of quantization per colour channel, the histograms have been compressed so as to cover the 8 bits encoding. All the slides were prepared by the same person, thus staining variability in this dataset is very limited. AGEMAP images can be analysed across multiple axis of differentiation: age, gender, diet, or individual mice to construct a variety of classification problems. For these reasons the datasets' authors proposed three different experiments using three different subsets of the original images:

- **Liver Aging Female (LAF)** experiment consists on a 4-way classification problem using the four classes (1, 6, 16 and 24 months) of images of female mice on ad libitum diet. This set is composed of 529 images.
- **Liver Gender AL (LGAL)** experiment consists on a 2-way classifier which classifies the gender of the mouse based on the images of 6-month old male and female mice on ad-libitum diet. This set is composed of 265 images.
- **Liver Gender CR (LGCR)** experiment consists on a 2-way classifier which classifies the gender of the mouse based on the images of 6-month old male and female mice on caloric restriction diet. This set is composed of 303 images.

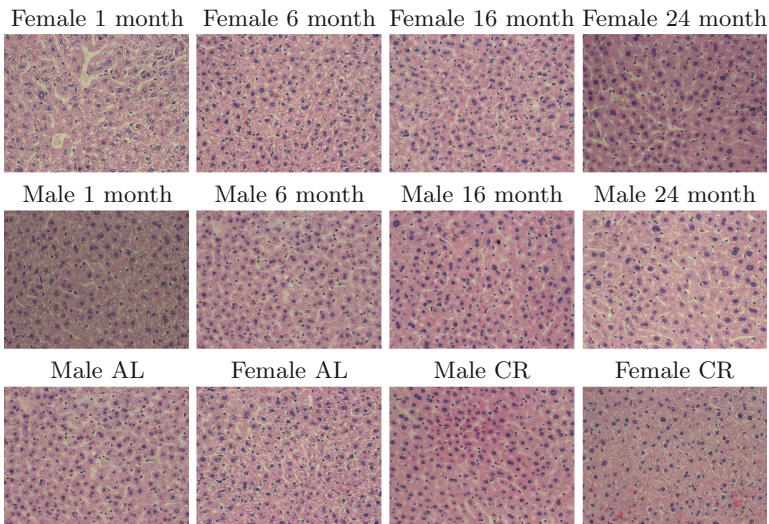
One more experiment has been added, **Liver Aging Male (LAM)**, to the previously mentioned. It consists on a 4-way classification problem, like the first one, even though four classes (1, 6, 16 and 24 months) of images of male mice on an ad libitum diet have been used. This set is composed of 499 images.

## 4 Experimental Evaluation

The performance of the described descriptors has been evaluated following two strategies. The first one has been performed converting each image in grayscale and extracting each descriptor from the converted image. The second one has been performed over the RGB images applying the computation scheme for the three R, G, B channels and linking the descriptors into a single vector in order to take into account the colour information as we proposed in [27]. Classification performances have been evaluated by calculating the accuracy, which offers a good indication of the performance since it considers each class of equal importance. Thus the classification accuracy have been estimated through a k-Nearest Neighbour (k-NN) classifier, with  $k = 1$  and using the euclidean distance. k-NN strategy has been preferred over more complex classifier in order to produce the results more representative of the effectiveness of the descriptors than of the classifiers themselves. Both analysing the grayscale images and the colour images we first tested the Hu, Zernike (up to order 10) and Legendre (up to order 8) moments and GLCMri and LBPri texture descriptors individually to assess the performances of the state-of-the-art methods. Then, we evaluated if the previous descriptors could benefit from a combination of them. In particular, we evaluated

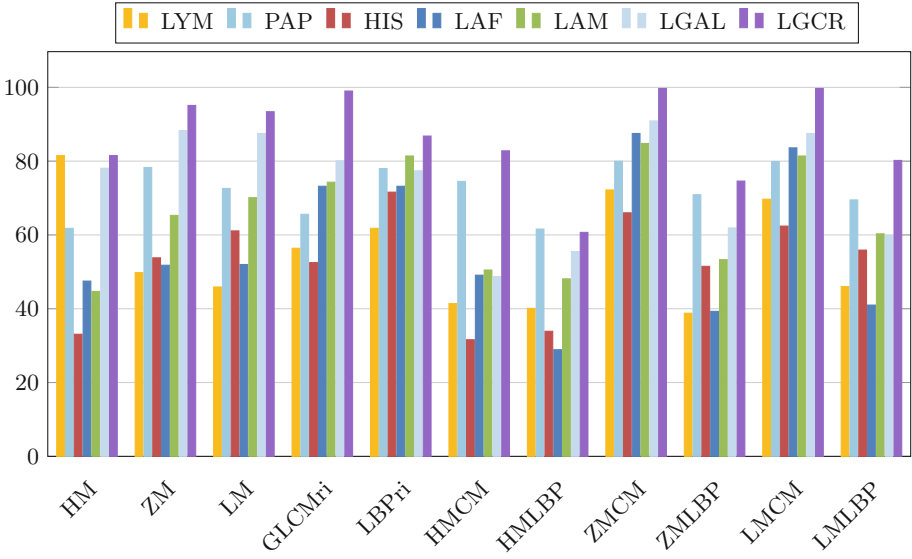


if the invariant moments could be more discriminative if extracted starting from a different representation instead of directly from the original images. Thus, we computed Hu, Zernike and Legendre moments starting from the LBP images and from the GLCM computed with angles  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  and  $135^\circ$ . In order to better understand the behaviour of the single descriptors and their combination we report a plot in Fig. 5 where the average accuracy calculated on every descriptor applied to each of the datasets from the grayscale images is showed. As it can be observed, all the invariant moments, and in particular Zernike and Legendre moments, are more discriminant if extracted from a different representation. However, in order to further improve the classification performances, a second experiment has been conducted. We extracted the features considering the R, G, B, channels colour information, by computing every descriptor for each of the colour channels and then concatenating the results of the three channels in the same feature vector as we proposed in [27]. In that work we demonstrated that the performance of a descriptor depends on the used color model. So, in order to make a fair comparison of our descriptors, in this work we have chosen the RGB color space. A plot that sums up this experiment is presented in Fig. 6. The performance of all the descriptors improves considerably by using colour information.

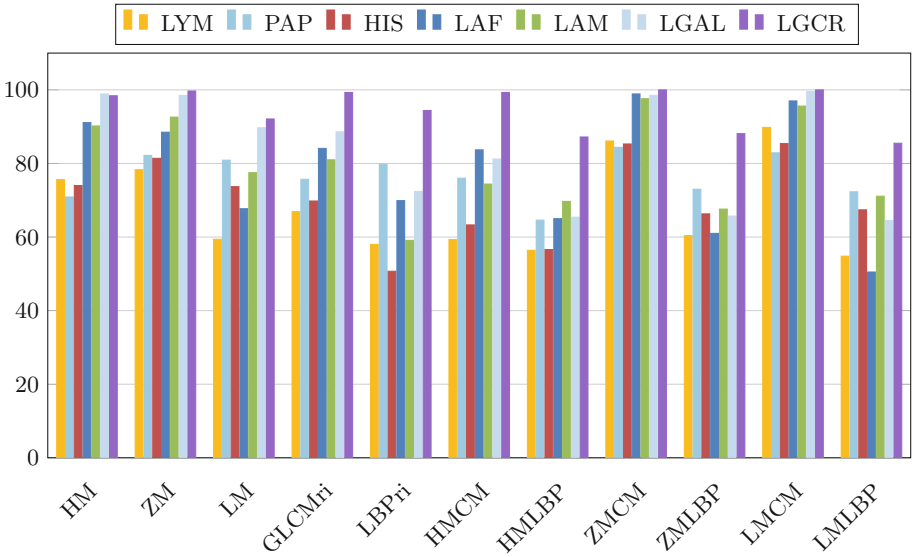


**Fig. 4.** Four liver images representing: female mice of the different ages (top), male mice of the different ages (center), male and female mice on Ad-libitum diet and on caloric restriction diet.





**Fig. 5.** Average performances of the descriptors extracted from the grayscale converted images. (Color figure online)



**Fig. 6.** Average performances of the descriptors extracted from the RGB images. (Color figure online)

## 5 Conclusions

In this work we have proposed a comparative study between different descriptors in analysing histological images. We focused the comparison on descriptors invariant to image rotations and, in particular, we measured the accuracy of moments, local binary patterns and co-occurrence matrices in classifying histological images. The experimentation has been conducted on well known public biomedical datasets: HistologyDS, Pap-smear, Lymphoma, Liver Aging Female, Liver Aging Male, Liver Gender AL and Liver Gender CR that represent a set of really different computer vision problems. We observed that, by extracting the invariant moments from the GLCM matrices, the overall accuracy of the invariant moments increases considerably, overcoming the classical LBP and GLCM approaches. In particular, if extracted taking into account colour information, the Zernike and Legendre moments impose themselves as very powerful descriptors for histological image analysis.

## References

1. Shamir, L., Orlov, N., Eckley, D.M., Macura, T., Goldberg, I.G.: A proposed benchmark suite for biological image analysis. *Med. Biol. Eng. Comput.* **46**(9), 943–947 (2008)
2. Karkanis, S.A., Iakovidis, D.K., Maroulis, D.E., Karras, D.A., Tzivras, M.: Computer-aided tumor detection in endoscopic video using colour wavelet features. *IEEE Trans. Inf. Technol. BioMed.* **7**(3), 141–152 (2003)
3. Ameling, S., Wirth, S., Paulus, D., Lacey, G., Vilarino, F.: Texture-based polyp detection in colonoscopy. In: Meinzer, H.P., Deserno, T.M., Handels, H., Tolxdorff, T. (eds.) *Bildverarbeitung für die Medizin 2009*. Informatik aktuell, pp. 346–350. Springer, Heidelberg (2009)
4. Hu, M.K.: Visual pattern recognition by moment invariants. *IRE Trans. Inf. Theory* **8**(2), 179–187 (1962)
5. Tuceryan, M.: Moment based texture segmentation. *Pattern Recogn. Lett.* **15**(7), 659–668 (1994)
6. Bigun, J.: N-folded symmetries by complex moments in Gabor space and their application to unsupervised texture segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **16**(1), 80–87 (1994)
7. Mukundan, R., Ong, S.H., Lee, P.A.: Image analysis by Tchebichef moment. *IEEE Trans. Image Process.* **10**(9), 1357–1364 (2001)
8. Yap, P.T., Raveendran, P., Ong, S.H.: Image analysis by Krawtchouk moments. *IEEE Trans. Image Process.* **12**(11), 1367–1377 (2003)
9. Teague, M.R.: Image analysis via the general theory of moments. *J. Opt. Soc. Am.* **70**(8), 920–930 (1980)
10. Di Ruberto, C., Morgera, A.: A comparison of 2-D moment-based description techniques. In: Roli, F., Vitulano, S. (eds.) *ICIAP 2005*. LNCS, vol. 3617, pp. 212–219. Springer, Heidelberg (2005). doi:[10.1007/11553595\\_26](https://doi.org/10.1007/11553595_26)
11. Di Ruberto, C., Morgera, A.: Moment-based techniques for image retrieval. In: 19th International Workshop on Database and Expert Systems Applications, pp. 155–159 (2008)

12. Haddadnia, J., Ahmadi, M., Faez, K.: An efficient feature extraction method with pseudo-Zernike moment in RBF neural network-based human face recognition system. *J. Appl. Signal Process.* **9**, 890–901 (2003)
13. Hosny, K.M., Papakostas, G.A., Koulouriotis, D.E.: Accurate reconstruction of noisy medical images using orthogonal moments. In: 18th International Conference on Digital Signal Processing (DSP), pp. 1–6, 1–3 July 2013
14. Vijayalakshmi, B., Bharathi, V.S.: Classification of CT liver images using local binary pattern with Legendre moments. *Curr. Sci.* **110**(4), 687 (2016)
15. Wu, K., Garnier, C., Coatrieux, J.L.: A preliminary study of moment-based texture analysis for medical images. In: 32nd Annual International Conference of the IEEE EMBS, pp. 5581–5584 (2010)
16. Iscan, Z., Dokur, Z., Olmez, T.: Tumor detection by using Zernike moments on segmented magnetic resonance brain images. *Expert Syst. Appl.* **37**(3), 2540–2549 (2010)
17. Tahmasbi, A., Saki, F., Shokouhi, S.B.: Classification of benign and malignant masses based on Zernike moments. *Comput. Biol. Med.* **41**, 726–735 (2011)
18. Dogantekin, E., Yilmaz, M., Dogantekin, A., Avci, E., Sengur, A.: A robust technique based on invariant moments - ANFIS for recognition of human parasite eggs in microscopic images. *Expert Syst. Appl.* **35**(3), 728–738 (2008)
19. Liyun, W., Hefei, L., Fuhao, Z., Zhengding, L., Zhendi, W.: Spermatogonium image recognition using Zernike moments. *Comput. Methods Progra. Biomed.* **95**(1), 10–22 (2009)
20. Oujaoura, M., Minaoui, B., Fakir, M.: Image annotation by moments. In: Papakostas, G.A. (ed.) *Moments and Moment Invariants - Theory and Applications*, vol. 1, no. 10, pp. 227–252. Science Gate Publishing (2014)
21. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **3**(6), 610–621 (1973)
22. Gelzinis, A., Verikas, A., Bacauskiene, M.: Increasing the discrimination power of the co-occurrence matrix-based features. *Pattern Recogn.* **40**(9), 2367–2372 (2007)
23. Walker, R., Jackway, P., Longstaff, D.: Genetic algorithm optimization of adaptive multi-scale GLCM features. *Int. J. Pattern Recogn. Artif. Intell.* **17**(1), 17–39 (2003)
24. Chen, S., Chengdong, W., Chen, D., Tan, W.: Scene classification based on gray level-gradient co-occurrence matrix in the neighborhood of interest points. In: *IEEE International Conference on Intelligent Computing and Intelligent Systems*, pp. 482–485 (2009)
25. Benco, M., Hudec, R.: Novel method for colour textures features extraction based on GLCM. *Radioengineering* **4**(16), 64–67 (2007)
26. Di Ruberto, C., Fodde, G., Putzu, L.: Comparison of statistical features for medical colour image classification. In: Nalpantidis, L., Krüger, V., Eklundh, J.-O., Gasteratos, A. (eds.) *ICVS 2015*. LNCS, vol. 9163, pp. 3–13. Springer, Cham (2015). doi:[10.1007/978-3-319-20904-3\\_1](https://doi.org/10.1007/978-3-319-20904-3_1)
27. Di Ruberto, C., Fodde, G., Putzu, L.: On different colour spaces for medical colour image classification. In: Azzopardi, G., Petkov, N. (eds.) *CAIP 2015*. LNCS, vol. 9256, pp. 477–488. Springer, Cham (2015). doi:[10.1007/978-3-319-23192-1\\_40](https://doi.org/10.1007/978-3-319-23192-1_40)
28. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. *Pattern Recogn.* **29**, 51–59 (1996)
29. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)

30. Cruz-Roa, A., Caicedo, J.C., González, F.A.: Visual pattern mining in histology image collections using bag of features. *J. Artif. Intell. Med.* **52**, 91–106 (2011)
31. Jantzen, J., Dounias, G.: Analysis of pap-smear data. In: NISIS 2006, Puerto de la Cruz, Tenerife, Spain (2006)
32. Zahn, J.M., Poosala, S., Owen, A.B., Ingram, D.K., Lustig, A., Carter, A., Becker, K.G., et al.: AGEMAP: a gene expression database for aging in mice. *PLoS Genet.* **3**(11), e201 (2007)