

## Comparison of HOG, MSER, SIFT, FAST, LBP and CANNY features for cell detection in histopathological images

<sup>1</sup>Şaban Öztürk, <sup>2</sup>Bayram Akdemir

<sup>1,2</sup>Electrical and Electronics Engineering Department, Selçuk University, Konya, Turkey

Email: [sabanozturk@selcuk.edu.tr](mailto:sabanozturk@selcuk.edu.tr), [bayakdemir@selcuk.edu.tr](mailto:bayakdemir@selcuk.edu.tr)

Received: 10<sup>th</sup> Feb 2018, Accepted: 18<sup>th</sup> March 2018, Published: 30<sup>th</sup> April 2018

### Abstract

Cell segmentation and counting has a very important role in diagnosing diseases and in the treatment process. But the complexity of the histopathological images and the differences in cell groups make this process very difficult, even for an expert. In order to facilitate this process, analysis of histopathological images is performed by using computer vision methods. This paper presents the use of different feature extraction methods for cell detection in histopathological images and the comparison of the results of these algorithms. For this reason, HOG, MSER, SIFT, FAST, LBP and CANNY feature extraction algorithms are used. The aim of the study is to determine cells using different feature extraction methods and to determine which of these feature extraction algorithms will be more successful. Firstly, image pre-processing has been applied to clear the noises in the histopathological images. Then, feature extraction algorithms are applied to image, respectively. Finally, the successes of different feature extraction algorithms have been compared.

**Keywords:** HOG, MSER, SIFT, FAST, LBP, canny, histopathological image, cell counting.

### Introduction

Computer aided diagnosis studies in biomedical engineering field are very useful for human health and diagnosis of diseases. In particular, the state of cellular structures and cell forms present useful information about many diseases and cancer types [1]. For this reason, cell forms and behaviors in histopathological images have become one of the main research areas. However, because of the negative and compelling properties of these images, such as unwanted light effects, low contrast, differences in cell shapes etc. analysis takes a long time. Again for this reason, even if these images are evaluated by an expert, histopathological image analysis is quite challenging and time-consuming. Also, it requires skill and experience [2]. In order to overcome all these difficulties, a computer-assisted automatic cell detection system is needed. Image processing algorithms which are used heavily in these systems and directly affect the success of the system have an important effect in this context [3, 4]. Some studies in the literature for cell detection from histopathological images are; compact Hough and radial map [5], logistic regression classifier [6], watershed

segmentation [7], mean-shift clustering based method [8].

In the image segmentation or classification process using image processing techniques, features obtained from images are affecting success. For this reason different features are used for different problems. The characteristics obtained by different feature extraction methods applied on the same problem can lead to different success rates. For this reason, many feature extraction methods have been proposed in the literature. For example, feature extraction methods such as SIFT and SURF are used for interest region detection, and LBP and LPQ are preferred for texture classification operations [9].

In this study, the success of well-known feature extraction methods for cell detection in histopathological images is investigated. For this purpose, HOG, MSER, SIFT, FAST, LBP feature extraction methods and canny edge detection algorithm are used. In order to detect the cells through the features obtained from the histopathological image using the mentioned algorithms, the feature sets need only cover the cells. For this reason, when applied to raw images, they fail because of disturbing factors such as noise. In this study, firstly RGB color space to HSV color space conversion is done. Then, noise in the V component of the HSV color space is cleaned using wiener filter [10]. Then, image sharpening is applied to the image. With the help of pre-processing, feature extraction algorithms have increased the chances of catching only the cells. In the next step, HOG [11], MSER [12], SIFT [13], FAST [14], LBP [15] and canny [16] algorithms are applied respectively. Finally, cells are identified with the help of extracted features and the success of each algorithm is compared with other algorithms.

### System Overview:

This study involves pre-processing step and well-known feature extraction algorithms. Well-known feature extraction algorithms are applied to histopathological images and the success rates of these algorithms are determined. In the proposed system, the image is first prepared for feature extraction. For this, the histopathological image is transformed into HSV color space and inserted into various preprocessing stages. Then feature extraction algorithms are applied to the image in sequence. The mentioned processes are shown in Fig 1.

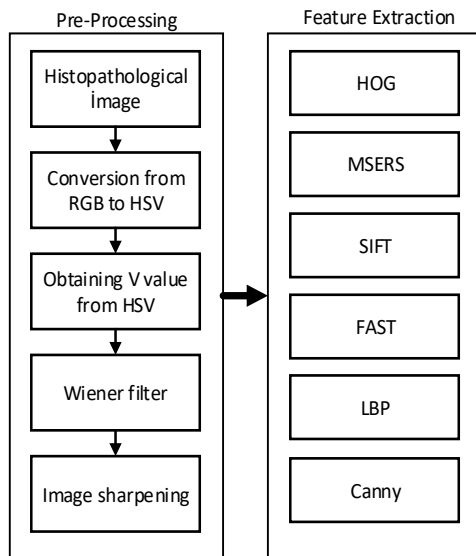


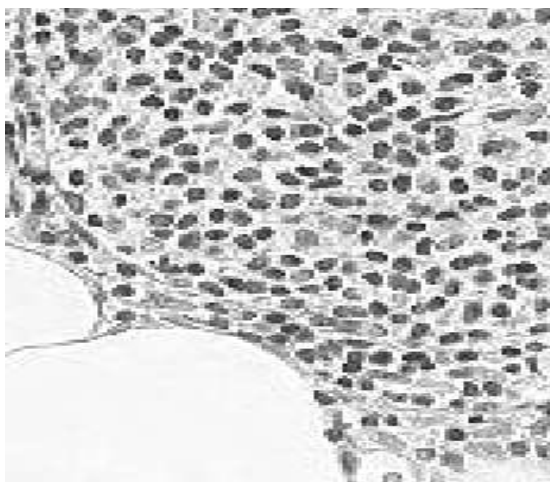
Fig. 1. Proposed system overview

### Pre-Processing of Histopathological images

The main algorithms in image processing applications are mostly sensitive to noise and light variations. For this reason, processing raw images prevents the achievement of desired success. Image pre-processing techniques are used to avoid noise and illumination fluctuations in images and they are used to reduce the effect on image processing algorithms [17]. The noises in the images are usually of different types. For this reason, problem-specific or image-specific image preprocessing techniques are used.

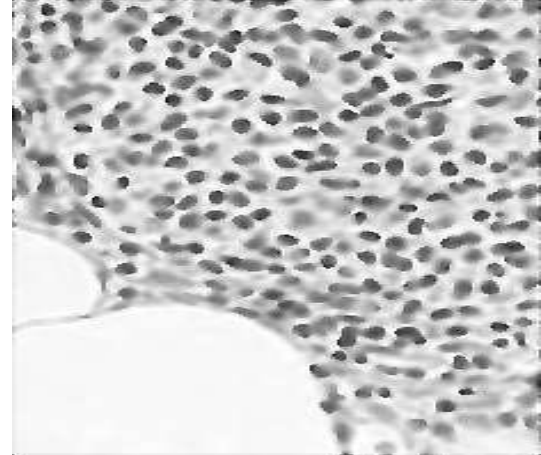
In this study, an image preprocessing solution is presented for low contrast, color problems caused by H & E stained and various noises in histopathological images. First, the image is converted from RGB color space to HSV color space. Then, V component of HSV color space is selected for further processing. V component is shown in Fig 2.

Fig. 2. V component of HSV color space



The Wiener filter [18] is used to remove noise from the V component. The Wiener filter obtains the statistical properties of the image according to the determined neighborhood value. For a great neighborhood value, the smoothing becomes less and the smoothing is big for a small neighborhood value. Finally, the image is sharpened to increase the clarity of the cells. Image after pre-processing is shown in Fig. 3.

Fig. 3. V component after pre-processing



### Feature Extraction

For object recognition and classification operations, extracting properties from the image and using these properties is the basic procedure. Basically, feature extraction is done according to orientation, scale and transformations [19]. Many algorithms are used with gray level images, and some algorithms can be used for color images. Extraction of efficient features is very important for image processing problems. However, efficient features may be different for each image problem. For this reason many feature extraction methods are presented. In this study, the success of well-known feature extraction methods is tested on histopathologic images.

#### A. HOG Feature Extraction

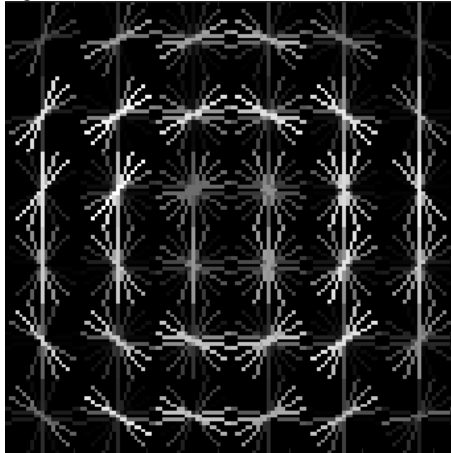
Histogram of gradients (HOG) specifies the magnitude and orientation parameters of feature regions in an image. These parameters are used to determine the objects in the image and image matching. They are independent of the size. Magnitude  $m(x,y)$  and orientation  $\theta(x,y)$  are calculated as in Eq. 1. and Eq. 2.

$$m(x,y) = \sqrt{f_x(x,y)^2 + f_y(x,y)^2} \quad (1)$$

$$\theta(x,y) = \arctan \frac{f_y(x,y)}{f_x(x,y)} \quad (2)$$

A feature image which is trained by cells for cell detection in histopathological images is shown in Figure 4. This feature matrix is applied to the test images to identify the cells.

Fig. 4. HOG feature for cell detection



### B. MSER Feature Extraction

Maximally stable extremal region (MSER) extracts a large number of affine covariant regions from the image. These features are usually elliptical regions. In the MSER algorithm, the threshold values are constantly changed to evaluate the areas whose shape is preserved. Using this algorithm, the capture of elliptical features in the image produces very successful results. For this reason, this algorithm has been chosen to be used for cell detection.

### C. SIFT Feature Extraction

The Scale invariant feature transform (SIFT) algorithm is used to capture local features in the image. The state of the specified properties is not affected by scaling, rotation, translation. First, local maxima and minima are determined by Gaussian function with different values. Equation 3 is used for this operation.

$$D(x, y, \sigma) = G(x, y, t\sigma) - G(x, y, \sigma) \times I(x, y) \quad (3)$$

where  $t$  represents constant factor,  $\sigma$  represents Gaussian difference value. Then, each pixel is compared with the neighboring pixels and the regions with low contrast values are eliminated by examining the key points. The original paper written by Lowe [20] can be examined for detailed information.

### D. FAST Feature Extraction

Features from accelerated segment test (FAST) determines the interest points in the image. Interest points carry local information about the content of interest and can be repeated for different images. Thanks to these features, they can be used for object detection, object tracking and scene matching operations. This algorithm checks whether each pixel is an interest point. The selected pixel intensity value is evaluated according to the intensity value of the neighboring 16 pixels. For a pixel can be an intensity point, the value of this pixel must be different from the neighboring pixels around it after the threshold operation.

### E. LBP Feature Extraction

Local binary pattern (LBP) labels each pixel by thresholding it according to  $n \times n$  neighborhood. A feature histogram representing the visual features is created with the obtained features. This feature histogram is searched in other images and similarities are determined. When similarity is determined, feature histograms are determined in other images and the similarities between two feature histograms are calculated.

### F. Canny Edge Detector

Canny edge detection algorithm is one of the most commonly used edge detection methods. The work of this algorithm is summarized as follows: firstly, the noise in the image is cleared. If the noises are not cleaned, the correct edges cannot be obtained. Then, a two-dimensional derivative of the image intensity in the x and y directions ( $G_x$ ,  $G_y$ ) are calculated. Finally, the gradient of the view is calculated by adding  $G_x$  and  $G_y$  values [21].

## Results and Discussion

In this study, the performance of the feature extraction algorithms in the cell detection process is examined. First of all, random image patches are cut from histopathological images. These image patches are cleaned with the preprocessing algorithm and are made feasible for feature extraction process. In the experiments, 12 pieces of histopathological image are examined. The cells of the image patches are marked by an expert. The total number of cells in the marked image patches is determined. Then, feature extraction process is performed according to their own capability and characteristics. The purpose of this process is to determine only the regions containing the cells. The performance of the well-known feature extraction algorithms are shown in Table 1.

**TABLE 1. COMPARISON OF WELL-KNOWN FEATURE EXTRACTION ALGORITHMS**

	Correct Cells	Uncorrected Cells	Undetected Cells	Total Cell number
HOG	2748	76	1176	3924
MSER	3540	22	384	3924
SIFT	2903	428	1021	3924
FAST	3064	294	860	3924
LBP	3489	158	435	3924
CANNY	3322	38	602	3924

In table 1, correctly identified cells are called "correct cell", cells that cannot be detected correctly are called "undetected cells" and regions that are marked as cells when they are not cells are called "uncorrected cells".



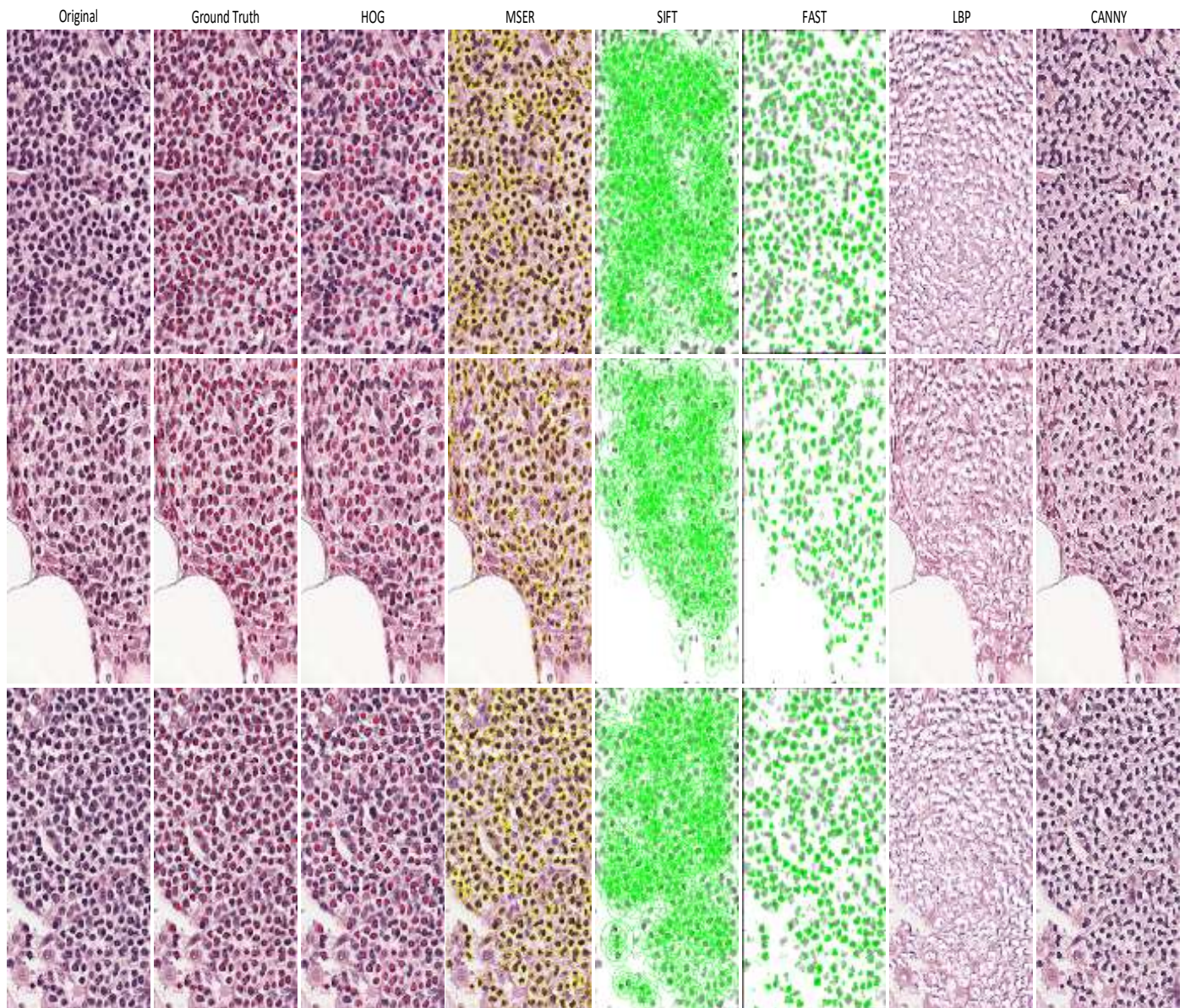


Fig: 5 HOG feature for cell detection

The features determined by well-known feature extraction algorithms are shown in Figure 5. The MSER algorithm produces better results than other algorithms. It could detect many cells and mark around. But the cells that are close to each other are perceived as a single cell by MSER algorithm. The LBP algorithm is another algorithm that can accurately determine the locations of cells. But too many false positives are marked as positive. The Canny edge detection algorithm determines the edges of the cells very successfully. However, since many regions have an intensity difference, non-cell regions are drawn as a cell by canny algorithm. The HOG algorithm is not very successful at catching all the cells, but the false markup rate is low compared to other algorithms. For this reason, it is more useful than LBP and canny algorithms. The FAST algorithm puts more than one sign on the same cell. The SIFT algorithm cannot detect many cells and marks many non-cell regions.

### Conclusion

In this study, the success of well-known feature extraction methods is investigated for cell detection in histopathological images. It has been understood that feature extraction algorithms cannot achieve the desired success on raw images. For this reason, preprocessing is applied before image properties are extracted. Then feature extraction algorithms are applied to the cleaned image. Each feature extraction algorithm operates on its own characteristic feature. In order for the comparison process to be objective, these characteristics are preserved and applied in the same conditions. At the end of the comparison process, MSER, LBP and canny algorithms detect cells with high success. However, the LBP and canny algorithms incorrectly mark many regions. In this case, the accuracy rate is low. The HOG algorithm is more advantageous than the LBP and canny algorithms because it has a low error rate when it has high

success in determining the cells. In line with this information, the MSER algorithm can be used as the first option and the HOG algorithm can be used as the second option. But, the selection process may vary depending on the application.

## References

- [1] Mao-Jun, S., Zhao-bin, W., Hong-Juan, Z., & Yi-de, M. (2008, March). A new method for blood cell image segmentation and counting based on PCNN and autowave. In Communications, Control and Signal Processing, 2008. ISCCSP 2008. 3rd International Symposium on (pp. 6-9). IEEE.
- [2] Mölder, A., Sebesta, M., Gustafsson, M., Gisselson, L., Wingren, A. G., & Alm, K. (2008). Non-invasive, label-free cell counting and quantitative analysis of adherent cells using digital holography. *Journal of microscopy*, 232(2), 240-247.
- [3] Bhagavathi, S. L., & Thomas Niba, S. (2016). An Automatic System for Detecting and Counting RBC and WBC using Fuzzy Logic. *ARN Journal of Engineering and Applied Sciences*, 11(11), 6891-94.
- [4] Sommer, C., Straehle, C., Koethe, U., & Hamprecht, F. A. (2011, March). Ilastik: Interactive learning and segmentation toolkit. In *Biomedical Imaging: From Nano to Macro*, 2011 IEEE International Symposium on (pp. 230-233). IEEE.
- [5] Barber, P. R., Locke, R. J., Pierce, G. P., Rothkamm, K., & Vojnovic, B. (2007, January). Gamma-H2AX foci counting: image processing and control software for high-content screening. In *Proceedings of SPIE* (Vol. 6441, pp. 64411M-1).
- [6] Liimatainen, K., Ruusuvaari, P., Latonen, L., & Huttunen, H. (2016, April). Supervised method for cell counting from bright field focus stacks. In *Biomedical Imaging (ISBI)*, 2016 IEEE 13th International Symposium on (pp. 391-394). IEEE.
- [7] Veta, M., van Diest, P. J., Kornegoor, R., Huisman, A., Viergever, M. A., & Pluim, J. P. (2013). Automatic nuclei segmentation in H&E stained breast cancer histopathology images. *PloS one*, 8(7), e70221.
- [8] Qi, X., Xing, F., Foran, D. J., & Yang, L. (2012). Robust segmentation of overlapping cells in histopathology specimens using parallel seed detection and repulsive level set. *IEEE Transactions on Biomedical Engineering*, 59(3), 754-765.
- [9] Kannala, J., & Rahtu, E. (2012, November). Bsif: Binarized statistical image features. In *Pattern Recognition (ICPR)*, 2012 21st International Conference on (pp. 1363-1366). IEEE.
- [10] Benesty, J., Chen, J., Huang, Y. A., & Doclo, S. (2005). Study of the Wiener filter for noise reduction. In *Speech Enhancement* (pp. 9-41). Springer Berlin Heidelberg.
- [11] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* (Vol. 1, pp. 886-893). IEEE.
- [12] Turki, H., Halima, M. B., & Alimi, A. M. (2017, March). A Hybrid Method of Natural Scene Text Detection Using MSERs Masks in HSV Space Color. In *Ninth International Conference on Machine Vision* (pp. 1034111-1034111). International Society for Optics and Photonics.
- [13] Ma, B., Ban, X., Wang, Y., & Liu, X. (2016, August). Recognition of blacklisted vehicle based on SIFT feature. In *Cloud Computing and Intelligence Systems (CCIS)*, 2016 4th International Conference on (pp. 456-460). IEEE.
- [14] Wang, D., Hou, X., Xu, J., Yue, S., & Liu, C. L. (2017). Traffic sign detection using a cascade method with fast feature extraction and saliency test. *IEEE Transactions on Intelligent Transportation Systems*.
- [15] Stella, X. A., Sivagangai, T., & Sujatha, I. D. N. (2016). Performance Analysis of GFE, HOG and LBP Feature Extraction Techniques using kNN Classifier for Oral Cancer Detection. *Journal of Network Communications and Emerging Technologies (JNCET)* www.jncet.org, 6(7).
- [16] Goulart, J. T., Bassani, R. A., & Bassani, J. W. M. (2017). Application based on the Canny edge detection algorithm for recording contractions of isolated cardiac myocytes. *Computers in biology and medicine*, 81, 106-110.
- [17] Heseltine, T., Pears, N., & Austin, J. (2002, July). Evaluation of image pre-processing techniques for eigenface based face recognition. In *Proc. of the Second International Conference on Image and Graphics*, SPIE (Vol. 4875, pp. 677-685).
- [18] Pratt, W. K. "Generalized Wiener filtering computation techniques." *IEEE Transactions on Computers*, 100(7), 1972, 636-641.
- [19] Van De Weijer, J., & Schmid, C. (2006). Coloring local feature extraction. *Computer Vision-ECCV 2006*, 334-348.
- [20] D. Lowe, "Distinctive Image Features from Scaleinvariant VKey Points", *IJCV*, 60(2), 91-110, 2004
- [21] Öztürk, Ş., & Akdemir, B. (2015). Comparison of edge detection algorithms for texture analysis on glass production. *Procedia-Social and Behavioral Sciences*, 195, 2675-2682.