

An Approach to Cell Nuclei Counting in Histological Image Analysis

Maryna Lukashevich and Valery Starovoitov^(✉)

Faculty of Economics and Informatics in Vilnius,
University of Bialystok, Vilnius, Lithuania
Lukashevich@bsuir.by, valerystar@mail.ru

Abstract. The paper describes a technique for automated cell nuclei counting. In this study, the primary goal is to provide simple and effective automated scheme of cell nuclei counting. The experiments on public data set of histology images have demonstrated acceptable level of calculation results.

Keywords: Histological analysis · Machine analysis · Cell nuclei counting

1 Introduction and Motivation

The most important and rapidly developing areas in the field of visualization and control of microscopic objects is development of automated computer vision systems. Computer analysis of microscopy cell images has many real-life applications in a wide range of areas including diagnosis of a disease, morphological cell analysis and statistics [1]. It plays an important role in biomedical research and bioinformatics. Microscopic analysis approach is particularly important in solving the following problems: malignant transformation of normal cells and cancer cell detection, morphological changes in cells, dynamic changes in the cells during therapeutic procedures [2].

For that reason, automatic counting of cell nuclei is a key block in systems for microscopic analysis of cell images. Since 1970s automated methods of histological analysis have been developed [1,2]. Detection of cellular structures and cell counting are common tasks for many investigations.

Some researchers have applied thresholding for cell counting in the investigated images. It is a simple way for separating objects of interest from the background. Another approach is to segment images using different edge detection (LoG filter, Laplacian filtering, etc.). However, detection and segmentation of cell nuclei is a challenging task, since the cells have a complex and a nonuniform structure. Another feature of histological image is non-uniform illumination, which also applies limitations on the use of standard approaches.

In recent years many methods have been proposed for cell nuclei segmentation, separation and classification in histological analysis. More sophisticated

approaches consist of several image processing stages. These methods are based on traditional image processing algorithms like adaptive contour model, watershed, morphological operations, k-means, Support vector machine, etc. [3–12].

Recent works suggest combining different approaches for increasing performance. Several excellent reviews about methods for nuclei detection, segmentation and classification can be found in [1, 2].

The main methods for cell nuclei segmentation, separation and classification are presented in Table 1.

Table 1. Short summary of state-of-the-art cell nuclei segmentation, separation and classification

Solvable task	Methods
Segmentation	Adaptive contour model, adaptive thresholding, morphological operations, watershed, k-means, H-maxima transform, region growing, Gaussian mixture model
Separation	Distance transform, watershed transform, concavity detection, edge path selection
Classification	Texture, morphology, texture and morphology, texture/morphology/intensity with Support vector machine, texture/morphology/topology with Bayesian, intensity with k-means clustering, intensity & texture with Adaboost

In our research we study the problem of automatically counting of cell nuclei in histological images of their size and shape.

Depending on the quality of prepared medication and used optical equipment, microscopy image can be both grayscale and color. The large number of different objects in histological images is a serious disadvantage for analysis of this class of images. Unfortunately, cell nuclei analysis is challenging task because of the complexity of the natural objects (noisy images, background heterogeneity, variations in object characteristics). Most of the cell nuclei are part of histological structures with complex and irregular visual features [13]. Samples of microscopy images and their intensity histograms are presented in Fig. 1. Shape of histograms shows that the problem of automatic nuclei detection and segmentation in histological images is difficult by thresholding.

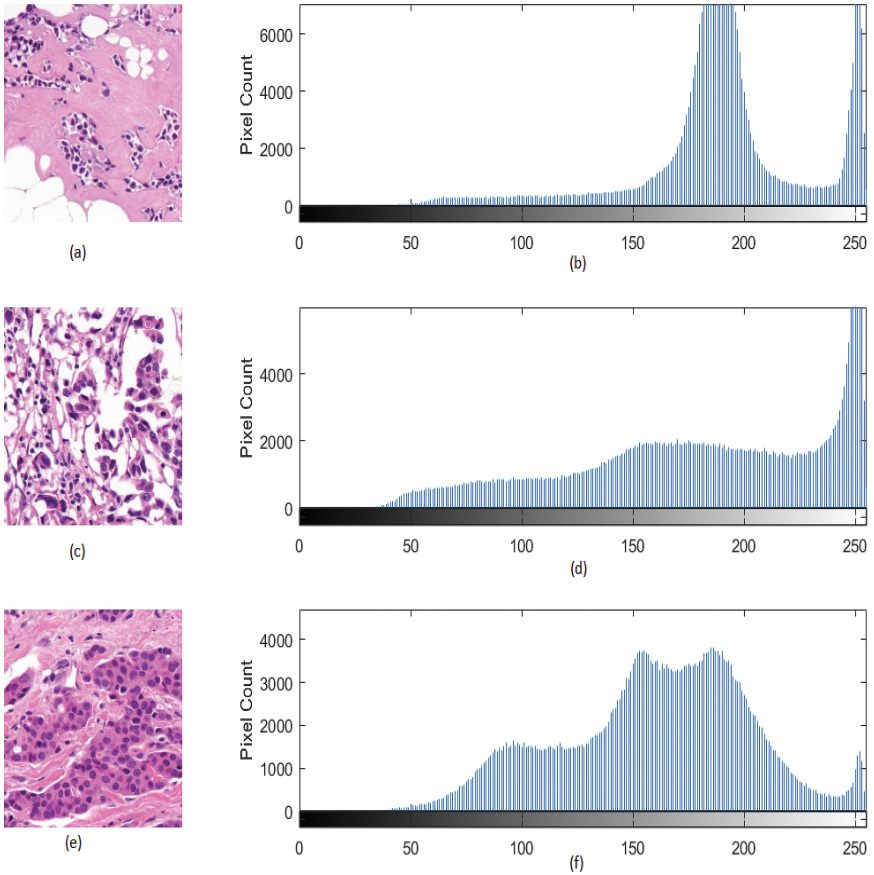


Fig. 1. H&E image samples (a), (c), (e) and their intensity histogram (b), (d), (f) respectively.

2 Methods and Materials

The investigated images are digital microscopy images stained with H&E (Hematoxylin & Eosin) from a publicly available data set [14, 15]. Authors [14, 15] used whole slide and tissue microarrays derived from tissue samples available through routine diagnostic. The histological slides were using the Zeiss Mirax Scan slide scanner. The slide scanner was equipped with Zeiss Plan-Apochromat 20x (numerical aperture = 0.8) and an AVT Marlin F-146C Firewire 1/2" CCD camera with $4.65 \mu\text{m} \times 4.65 \mu\text{m}$ pixel size. Combine with the 20x objective and 1x C-mount adapter the resulting image resolution is $0.23 \mu\text{m} \times 0.23 \mu\text{m}$. All slides were scanned at 20x and performed at full resolution. The resulting images were converted to the Virtual Slide Format with actual image data encapsulated and saved as JPEG image files with 85% JPEG quality. The image sizes vary from

33280×29184 pixels to 70080×159000 pixels, but the analyzed field of the images had a size of 600×600 pixels.

The final dataset contains 7931 cells from 36 color tissue images. The images of different organs (breast, kidney, gastric mucosa, connective tissue, small intestine, etc.) are found in data set. The nuclei center coordinates are used as ground truth data for proposed approach validation. All images were labeled by three pathologists, Fig. 2.

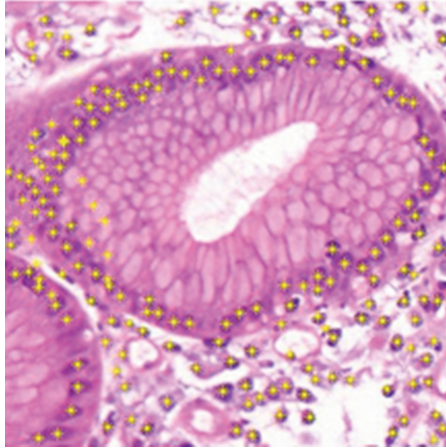


Fig. 2. An example of H&E image with labeled nuclei centers.

2.1 The Main Idea

The proposed approach for cell nuclei counting in histological image analysis consists of the following stages: 1. RGB color space regularization; 2. image intensity adjusting; 3. guided filtering; 4. thresholding; 5. morphological processing.

The preprocessing stage includes several steps. Original images contain a limited number of primary colors such as pink, purple and shades of blue. It is proposed to reduce the number of colors in the original image palette using Minimum Variance Quantization. Minimum variance quantization cuts the RGB color cube into smaller boxes (not necessarily cubical shape) of different sizes, depending on how the colors are distributed in the image. If the input image actually uses fewer colors than the number specified, the output colormap is also smaller, Fig. 3.

For the tested data set this procedure is almost visible to the human eye because of the limited number of colors in the difference between original images and same images after regularization is hardly visible (Fig. 3), but it is crucial for further image analysis stages.

After that we try to increase the contrast of the images by a contrast enhancement procedure.

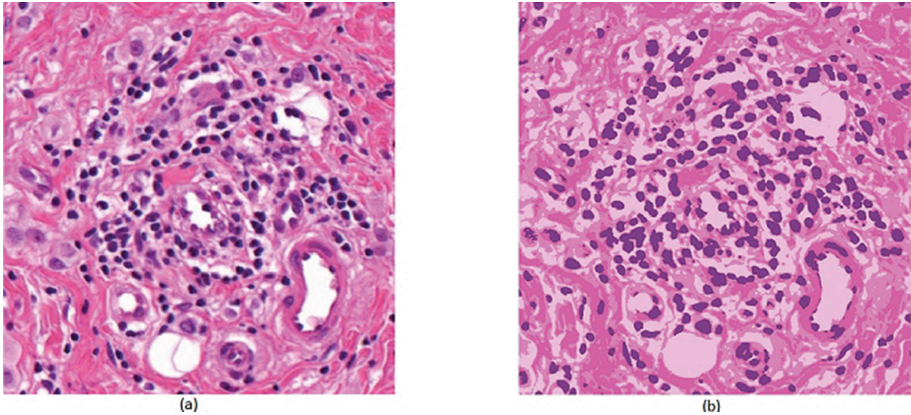


Fig. 3. (a) An input image, (b) the image after requantization. (Color figure online)

The next step is using guided filter. The filter performs edge-preserving smoothing of images, using content of a second image, called a guidance image, to influence the filtering. The guidance image can be the original image itself, a another version of the original image, or a completely different image. Guided image filtering is a neighborhood operation, like other filtering operations, but takes into account pixel statistics in the corresponding spatial neighborhood of the guidance image when calculating the value of the output pixel [16].

Then we apply binarization by Sauvola local image thresholding [17–19] for separation nuclei and background, see Fig. 4.

We clean the obtained binary images by opening and closing with a small structure element. After that we count automatically the number of black blobs corresponding to the cell nuclei.

Opening and closing to the binary images we apply analysis help us to clean the binary images. They are both derived from the fundamental operations of erosion and dilation. Like those operators they are normally applied to binary images, although there are also gray level versions. The basic effect of an opening is somewhat like erosion in that it tends to remove some of the foreground (bright) pixels from the edges of regions of foreground pixels.

2.2 Nuclei Counting and Method Validation

We have tested implementations (in Matlab) of the proposed approach in 36 images from the data set described above.

The typical parameter values used in our experiments are shown in Table 2.

Example of analysis results on H&E images is shown in Fig. 5.

The performance of calculation was calculated using TP (true positive), FN (false negative) and FP (false positive) events. The performance parameters precision and recall calculating using (1)–(2).

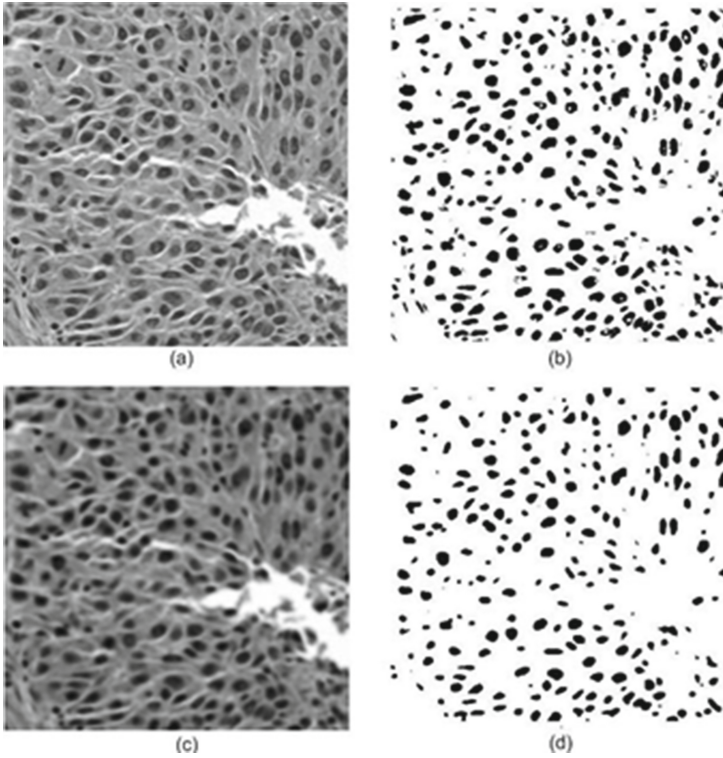


Fig. 4. Improvement in binarization by guided filtering: (a) input grayscale image, (b) thresholding without guided filtering; (c) input grayscale image after guided filtering; (d) thresholding after guided filtering.

Table 2. Typical parameters values

Processing stage	Parameters	Different values
Minimum variance quantization	Amount of color space	4
Guided filtering	Mask size balance between data matching and smoothing	33 0.5
Sauvola local image thresholding	Local thresholding with M-by-N neighbourhood threshold	45 × 45 0.35
Morphology	Radius of structuring elements shape of structuring elements	1 disk

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

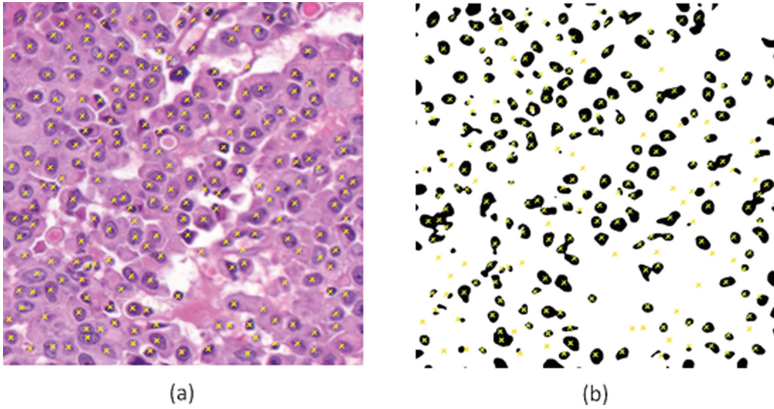


Fig. 5. (a) An example of the original image and (b) its binary variant with the centers of cell nuclei labeled by experts.

A measure that combines precision and recall is the harmonic mean of precision and recall, the traditional F-measure or balanced F-score (3):

$$F = 2 \cdot \frac{\textit{Precision} \cdot \textit{Recall}}{\textit{Precision} + \textit{Recall}} \quad (3)$$

We compared the results with recent advances in this area [7,9,10], Table 3.

Table 3. Detection results of the proposed approach, Wienert's, Al-Kofahi's methods in precision and recall values

	Proposed	Wienert	Al-Kofahi
Precision	0.833 ± 0.01	0.908 ± 0.04	0.707 ± 0.13
Recall	0.920 ± 0.01	0.859 ± 0.04	0.916 ± 0.04
F-measure	0.8743	0.8828	0.7980

3 Discussion

Table 2 shows the best parameters for the tested image database. Developed approach can be adapted to different nuclei counting task. Its main advantage is automatically approach for cell nuclei counting. In future, it is necessary to automatization the process of the initial parametrs settings depending on images. Cell nuclei with small area were lost because we used the guided filter and mathematical morphology. But it allows us to get more accurate results for other cells. This is due to the fact that the improved thresholding results (disappear small areas that are not true cell nuclei). Future work will be focused on improving the segmentation stage and enlargement of the tested image data set. In contrast to the known algorithms, the proposed approach does not segment all of the nuclei in images.

The aim of the experiments was not to improve metrics Precision and Recall. We just calculated them with results of the similar research. In our method the value of F-measure is equal to 0.8743. It is bigger than F-measure of Al-Kofahi, but smaller than F-measure of Wienert results. There are a lot of small nucleus in the images, which have been marked by the experts. These nucleus difficult to correctly identify and count. In this way, we obtain decrease Precision accuracy. More subtle settings allow detecting the missing small nucleus, but this requires coordination with experts about nuclei features.

4 Conclusions

An automated cell nuclei detection and counting method based on the guided filter and morphological operations is proposed. Different to the existing algorithms, the proposed approach does not segment nuclei in images. It only detects the cell nuclei in a histological image and estimate the number of the detected nuclei. It avoids complicated algorithmic computations and provides good accuracy in the cell nuclei counting. The software prototype developed in our study may be considered as an automatic tool for a cell nuclei analysis.

References

1. Irshad, H., Veillard, A., Roux, L., Racoceanu, D.: Methods for nuclei detection, segmentation, and classification in digital histopathology: a review current status and future potential. *IEEE Rev. Biomed. Eng.* **7**, 97–114 (2014)
2. Chen, S., Zhao, M., Wu, G., Yao, C., Zhang, J.: Recent advances in morphological cell image analysis. *Comput. Math. Methods Med.* **2012**, 10 (2012). Hindawi Publishing Corporation
3. Jung, C., Kim, C.: Impact of the accuracy of automatic segmentation of cell nuclei clusters on classification of cell nuclei clusters on classification of thyroid follicular lesions. *Cytometry Part A* **85A**, 709–719 (2014)
4. Saharma, H., Zerbe, N., Heim, D., Wiener, S., Behrens, H., Hellwich, O., Hufnagl, P.: A multi-resolution approach for combining visual information using nuclei segmentation and classification in histopathological images. In: *Proceedings of the 10th International Conference on Computer Vision, Theory and Applications (VIS-APP 2015)*, pp. 37–46 (2015)
5. Alilou, M., Kovalev, V., Taimouri, V.: Segmentation of cell nuclei in heterogeneous microscopy images: a reshapable templates approach. *Comput. Med. Imaging Graph.* **37**, 488–499 (2013)
6. Kowal, M., Filipczuk, P.: Nuclei segmentation for computer-aided diagnosis of breast cancer. *Int. J. Appl. Math. Comput. Sci.* **24**(1), 19–31 (2014)
7. Wienert, S., Helm, D., Saeger, K., Stenziger, A., Beil, M., Hufnagl, P., et al.: Detection and segmentation of cell nuclei in virtual microscopy images: a minimum-model approach. *Nat. Sci. Rep.* **2**, 503 (2012)
8. Zang, C., Xiao, X., Li, X., Chen, Y.-J., Zhen, W., Chang, J., Zheng, C., Liu, Z.: White blood cell segmentation by color-space-based k-means clustering. *Sensors* **14**, 16128–16147 (2014). doi:10.3390/s140916128

9. Song, Y., Cai, W., Huang, H., YueWang, D.D., Feng, M.C.: Region-based progressive localization of cell nuclei in microscopic images with data adaptive modeling. *BMC Bioinf.* **14**, 173 (2013)
10. Coelho, L.P., Shariff, A., Robert, F.: Murphy nuclear segmentation in microscope cell images: a hand-segmented dataset and comparison of algorithms. In: *Proceedings of IEEE International Symposium Biomedical Imaging*, pp. 518–521 (2009)
11. Signolle, N., Revenu, M., Plancoulaine, B., Herlin, P.: Wavelet-based multiscale texture segmentation in application to stromal compartment characterization on virtual slides. *Sig. Process.* **90**(8), 2412–2422 (2010)
12. Lezoray, O., Elmoataz, A., Cardot, H., Gougeon, G., Lecluse, M., Elie, H., Revenu, M.: Segmentation of cytological image using color and mathematical morphology. *Acta Stereologica* **18**, 1–14 (1999)
13. Loukas, C.G., Wilson, G.D., Vojnovic, B., Linney, A.: An image analysis-based approach for automated counting of cancer cell nuclei. *Tissue Sect. Cytometry Part A* **55A**, 30–42 (2003)
14. Al-Kofahi, Y., Lassoued, W., Lee, W., Roysam, B.: Improved automatic detection and segmentation of cell nuclei in histopathology images. *IEEE Trans. Biomed. Eng.* **57**(4), 841–852 (2010)
15. Al-Kofahi, Y., Lassoued, W., Grama, K., Nath, S.K., Zhu, J., Oueslati, R., et al.: Cell-based quantification of molecular biomarkers in histopathology specimens. *Histopathology* **59**(1), 40–54 (2011)
16. Xu, L., Lu, C., Xu, Y., Jia, J.: Image smoothing via L0 gradient minimization. *ACM Trans. Graph.* **30**(6), article 174, December 2011
17. Sauvola, J., Pietikainen, M.: Adaptive document image binarization. *Pattern Recognit.* **33**, 225–236 (2000)
18. Shafait, F., Keysers, D., Breuel, T.M., Efficient implementation of local adaptive thresholding techniques using integral images. In: *Document Recognition and Retrieval XV* (2008)
19. Stathis, P., Kavallieratou, E., Papamarkos, N.: An evaluation technique for binarization algorithms. *J. Univ. Comput. Sci.* **14**(18), 3011–3030 (2008)