Nuclei Segmentation on Histopathology Images of Breast Carcinoma

V. Y. Ramírez-Guatemala-Sánchez, H. Peregrina-Barreto, G. López-Armas

Instituto Nacional de Astrofísica, Óptica y Electrónica

Abstract

According to WHO data, breast cancer is the most common type of cancer in women, representing 16 % of cases. The diagnosis is made by the expert pathologist through a tissue analysis of histological sections under the microscope. The histological evaluation allows determining the degree of cellular differentiation of the tissue associated with the molecular classification of cancer to identify the most appropriate treatment for a patient. Computer-assisted diagnostic systems have gained relevance in pathology for tasks such as nuclei segmentation and analysis of nuclear pleomorphism, associated with the histopathological classification and grading of breast cancer. In this work, a methodology is proposed for the automatic detection and segmentation of cell nuclei with clump separation based on digital image processing and analysis techniques.The obtained results reached values precision of 0.75 ± 0.09, sensitivity of 0.76 ± 0.12, and a Sørensen-Dice similarity measure of 0.75 ± 0.08.

Introduction

A breast cancer diagnosis is made by the expert pathologist, who performs a visual evaluation of the tissue structures to classify and grade cellular differentiation [**1**] [**2**]. This evaluation is substantial to know the type and degree of malignancy of breast carcinoma [**3**]. The pathologist performs a fundamental task for visual tissue analysis by evaluating the size and shape of the cellular nucleus. The recognition and evaluation of the tissue are done subjectively [**4**]; hence, it is prone to possible diagnosis errors and discrepancies between experts; this problem has been recognized by the WHO [**5**][**6**]. The image analysis techniques and computer vision have been widely recognized in medical research since they provide an advantage in quantitative analysis and aid the final diagnosis of the experts [**1**]. Automatic detection of the cell nuclei remains critical in performing an automated analysis of breast tissue histological sections.

Objective

The objective of the proposed methodology is to perform the identification and automatic segmentation of cell nuclei with clump separation in images of histological sections of mammary carcinoma, which in turn allows a quantitative analysis that assists the histological evaluation of nuclear pleomorphism carried out by a expert pathologist.

Methodology

A. Separation by color

H&E staining system is characterized by highlighting the tissue components in purple tones, being darker in the nucleus area, and pink in the cytoplasm and related tissue surrounding the nucleus. Therefore, the separation of the red channel of the color image helps to highlight the purple hue to extract the components of interest (Fig. 1a-b).

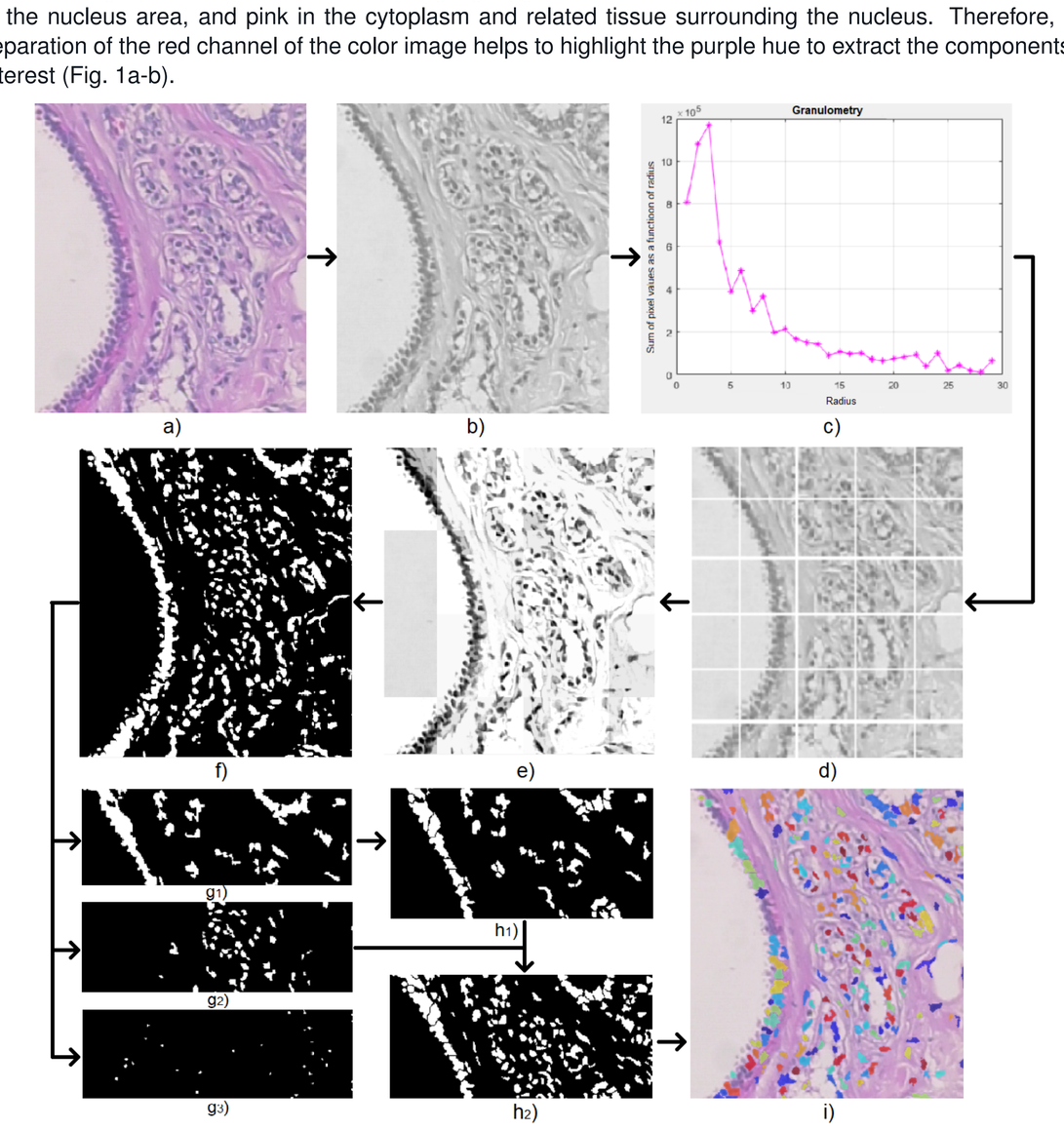

Fig. 1: Scheme of the process for nuclei segmentation. **a)** RGB breast cancer histology image **b)** R channel image, **c)** granulometric analysis, **d)** sub-section separation using granulometry measure as reference, **e)** local contrast adjustment result, **f)** image binarization, $\mathbf{g}_1$**)** identification of big particles and clumps for separation with Watershed, $\mathbf{g}_2$**)** average size particles are saved, $\mathbf{g}_3$**)** small particles, noise and possible artefacts, $\mathbf{h}_1$**)** segmented clumps, $\mathbf{h}_2$**)** resulting particles of previous processes are fused, **i)** resulting segmentation colored.

B. Local contrast adjustment

Once the red channel has been obtained, it seeks to improve the contrast in the regions that correspond to the cell nucleus then, a contrast adjustment is applied which uses a gamma correction to highlight objects of interest in an image. Due to the large variation in gray levels, a better contrast adjustment is made locally by dividing the image into subregions. In this way, it is possible to deal with the gray level differences within a smaller area and enhance the contrast among its components, i.e. cell nucleus and background. To establish an appropriate subregion size, a granulometric analysis is performed to estimate the size distribution of the particles in the image and ensure that the subregion be larger than any cell nucleus on the image (Fig. 1-c). Thus, the subdivision of the image is independent of the magnification. Once the image has been divided into subregions, the contrast adjustment is performed locally (Fig. 1d-e). This result facilitates the binary segmentation of the image, eliminating the background and other tissue parts that are not of interest (Fig. 1-f).

C. Adaptive segmentation

Once that only the objects of interest are present in the binary image, they are analyzed to delete some artifacts such as tissue portions, and clumped or overlapping particles. From the granulometric analysis, the most representative size of the particles r is known, which allows to establish a threshold that eliminates the smaller particles considered non-relevant (Fig. 1-g_3). The particles with a size close to r are of interest (Fig. 1-g_2), therefore, they are reserved and analyzed to obtain an area and circularity threshold. Finally, the particles with a larger size could be associated with clusters, overlapping particles or very close to each other (Fig. 1-g_1). If these particles are removed, important information could be lost. Then, it is important to identify the inner cell nuclei that these components may contain and to divide them. To address this problem, a watershed segmentation was performed (Fig. 1-h_1); all the segmented particles were fused in a single segmented image (Fig. 1-h_2). To obtain the final result, circularity and area threshold were recalculated over the segmented clumps with the aim of preserve only the rounder and with the most significant area particles (Fig. 1-i).

Results

The proposed methodology was tested on a subset of four different databases containing H&E stained microscopic images of histological sections of breast cancer [**7**, **8**, **9**, **10**]. With the aim of test the proposed method under different tissue condition, there were analyzed images of papillary and invasive carcinoma (Fig. 2) as well as benign tumors .

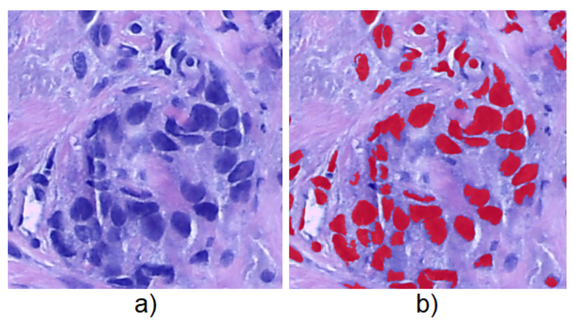

Fig. 2: **a)** Invasive carcinoma histological image [**8**], **b)** the nuclei segmentation results by the proposed method.

D. Evaluation

To evaluate the quality of the obtained segmentation, groundtruth images provided by the database in [**9**], were used. To evaluate the obtained results three values were calculated: precision, sensitivity and similarity Sørensen–Dice coefficient (SDC), a measure of similarity between two sets (obtained segmentation and groundtruth) with values in the range $[0, 1]$ [**11**]. Figure 4 shows an example of the original image and the comparison between the obtained segmentation and its corresponding groundtruth. It is observed a high similarity between results although they may differ in how some components are presented. The main difference is the clump separation achieved by the proposed method and the removal of elongated particles. These differences are reflected in the value of the SDC. However, these are positive characteristics in terms of the contribution of the method, ensuring that only well-identified cell nuclei remain in the final segmentation. Table 1 shows the statistics about segmented cell nuclei of the whole database.

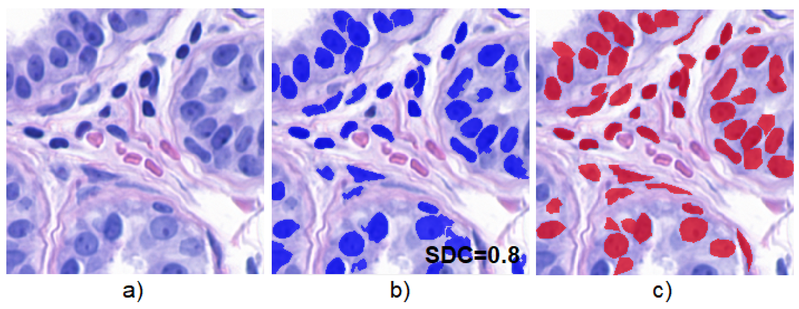

Fig. 3: **a)** Original histological image [**9**], **b)** the segmentation results by the proposed method , and **c)** their corresponding groundtruth.

Table 1: Comparison of core segmentation with and without clump division

Metric	Without Clump division	Clump division
Sørensen-Dice coefficient	0.78 ± 0.07	0.75 ± 0.08
Precision	0.73 ± 0.09	0.75 ± 0.09
Sensitivity	0.83 ± 0.1	0.76 ± 0.12

mean±standard deviation

Conclusions

The high variations in intensity and characteristic forms of the mammary cell tissue represents a challenge for the automatic characterization and segmentation of the cell nuclei. Nonetheless, the general quality of the segmentation depends on the robustness of the method to handle this type of variations. The segmentation results achieved with the proposed method is consistent under different image conditions such as density of cells, magnification and contrast, making it useful in the pathological analysis.

References

[1]- Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M., Yener, B. Histopathological Image Analysis: A Review. IEEE Reviews in Biomedical Engineering, 2, 147–171. doi:10.1109/rbme.2009.2034865. (2009).

[2]- Veta, M. et al. “Breast Cancer Histopathology Image Analysis: A Review.” IEEE Transactions on Biomedical Engineering 61 (2014): 1400-1411.

[3]- B. Ulmer, “Vita ii-active collision avoidance in real traffic,” in Intelligent Vehicles’ 94 Symposium, Proceedings of the. IEEE, 1994, pp. 1–6.

[4]- Slaoui, Mohamed Fiette, Laurence (2011). Histopathology Procedures: From Tissue Sampling to Histopathological Evaluation. Methods in molecular biology (Clifton, N.J.). 691. 69-82. 10.1007/978-1-60761-849-24. (2011)

[5]- Singh, H., Schiff, G. D., Graber, M. L., Onakpoya, I., Thompson, M. J. The global burden of diagnostic errors in primary care. BMJ Quality Safety, 26(6), 484–494. doi:10.1136/bmjqs-2016-005401. (2016).

[6]- Makary, M. A., Daniel, M. Medical error—the third leading cause of death in the US. BMJ, i2139. doi:10.1136/bmj.i2139. (2016).

[7]- Spanhol, F., Oliveira, L. S., Petitjean, C., Heutte, L., A Dataset for Breast Cancer Histopathological Image Classification, IEEE Transactions on Biomedical Engineering (TBME), 63(7):1455-1462, (2016).

[8]-Aresta, Guilherme et al. “BACH: Grand Challenge on Breast Cancer Histology Images.” Medical image analysis 56 (2019): 122-139 .

[9]- Breast Cancer Cell Segmentation. Elisa Drelie Gelasca and Jiyun Byun and Boguslaw Obara and B.S. Manjunath. Avaliable at: http://bioimage.ucsb.edu/research/bio-segmentation

[10] Robert J. Marinelli, Kelli Montgomery, Chih Long Liu, Nigam H. Shah, Wijan Prapong, Michael Nitzberg, Zachariah K. Zachariah, Gavin J. Sherlock, Yasodha Natkunam, Robert B. West, Matt van de Rijn, Patrick O. Brown, and Catherine A. Ball. The Stanford Tissue Microarray Database. Nucleic Acids Res 2008 36(Database issue): D871-7. Epub (2007) Nov 7 doi:10.1093/nar/gkm861

[11]- J. Panduro. "Estudio comparativo de algoritmos disponibles en el ITK para la segmentación de imágenes medicas. Capitulo 5." Avaliable at: http://bibing.us.es/proyectos/abreproy/11863/fichero/PFC% 252FCapitulo+V.pdf