An automatic system for cell nuclei pleomorphism segmentation in histopathological images of breast cancer

Pegah Faridi, Habibollah Danyali, Mohammad Sadegh Helfroush
Department of Electrical and Electronics Engineering
Shiraz University of Technology
Shiraz, Iran
p.faridi@sutech.ac.ir

Mojgan Akbarzadeh Jahromi
Department of Pathology, School of Medicine
Shiraz University of Medical Sciences
Shiraz, Iran
akbarzadeh@sums.sc.ir

***Abstract*— Nuclear pleomorphism is one of the criteria for diagnosing and grading breast cancer. The grading that is made by pathologist is subjective and prone to inter, intra observer variations. Furthermore, pathologists may need a huge time for evaluating all cases per day. Therefore, there is a necessity to provide an automatic system for a better diagnosis and detection. This paper proposes an automatic system for detecting and segmenting cancerous nuclei, which is partly different from healthy nuclei segmentation systems. In contrast, our system detects critical nuclei with any shape, border and chromatin density even in higher scores. This system avoids segmenting healthy cell nuclei. It only detects and segments a high percentage of deformed cell nuclei, which are necessary for nuclear pleomorphism scoring even cells with vesicular nuclei that are not detected in any other algorithms.**

Keywords—breast cancer; nuclear pleomorphism, histopathology; level set method; segmentation

I. INTRODUCTION

Breast cancer is one of the most pervasive causes of death among women all around the world. Breast cancer survival rates vary greatly worldwide, ranging from 80% or over in North America to around 60% in middle-income countries and below 40% in low-income countries [1]. The low survival rates can be explained mainly by the lack of early detection programs, as well as by the lack of adequate diagnosis and treatment facilities. To decrease high rate of mortality, precise diagnosis and prognosis is required.

Histologic analysis investigates the presence of cancer and its progression rate. Nowadays, Histopathological images are available in high resolution and high magnification in digital format which can be further processed to extract useful structural information. However, the manual analysis of such huge sets of data can be time consuming [2]. To establish a precise diagnosis, a biopsy examination is required. The biopsy sample is processed and its sections are placed onto glass slides to observe them under microscope for analysis. The current procedure for breast cancer grading is manually performed by pathologists. A Pathologist examine the tissue slides under a microscope and observe it at various magnification levels such as 10X, 20X, 40X pursuant to the structures of interest, which is a subjective and time consuming process. As a result, a computer aided diagnosis system is required to provide a standard and quantitative measurement for breast cancer evaluation.

Each type of a cancer utilizes specific grading scheme. Nottingham grading system, is the benchmark for breast cancer analysis which focuses on three criteria including mitotic count, nuclear pleomorphism and tubule formation.

A large variety of methods are developed for mitosis detection [3]-[5] and segmentation of tubular structures [6]-[11]. However, few methods are presented for nuclear pleomorphism criterion whilst nuclear pleomorphism scoring offers identical contribution in Nottingham grading system as the other two criteria. Cell nuclei detection is a prerequisite for nuclear pleomorphism. Hence, several methods were proposed to tackle this problem [12]–[15]. Most of the available algorithms detect healthy cell nuclei, where an alteration is essential for detecting deformed cell nuclei to carry out nuclear pleomorphism scoring. Few experiments exist that can distinguish healthy cell nuclei from the cancerous one [16], [17].

In this work we propose a system for detecting cancerous cell nuclei and segmenting exact boundaries that is a necessity for nuclear pleomorphism scoring. In [17] only critical cell nuclei that affect the score are segmented using morphological operations and distance transform. Despite creating promising results, it is not able to detect vesicular nuclei. Consequently, accuracy decreases in higher scores. In contrast, our system detects critical nuclei with any shape, border and chromatin density even in higher scores. The remainder of this paper is as follows. In section II, the proposed segmentation method is described. Section III presents evaluation results of the system and finally, the concluding remarks are presented in section IV.

II. THE PROPOSED METHOD

The main aim of the proposed method is to segment deformed cell nuclei boundary accurately for a better scoring of nuclear pleomorphism criterion in Nottingham grading system.

An overview of the segmentation system is shown in Fig.1. The proposed method consists of two internal level:

A. Detection of centroid of cell nuclei
B. Cell nuclei boundary segmentation

A. Detection of center of cell nuceli

Histology images have complex structures such as cells with linked boundaries, overlapping tissues and noisy cell region which absolutely reduces the accuracy of detection and segmentation. Hence, preprocessing is required in order to remove redundant structures and noises so that the accuracy of proposed system increases.

Filtering is perhaps the most fundamental preprocessing used in Computer aided diagnosis (CAD) systems. In the proposed method, bilateral filtering is used as the first step of preprocessing. A bilateral filter is defined as a weighted average of nearby pixels which takes into account the difference in value with other pixels in the neighborhood so that edges are preserved while smoothing. Bilateral filter denoted as BF[.] is formulated as it is indicated in Eq.1 and Eq.2, in which W_P is the normalization term, I is the original input image, x are the coordinate of the pixels to be filtered inside ω, ω is the window centered in x, f_r is the range kernel for smoothing, and g_s is the spatial kernel for smoothing.

$$BF[I] = \frac{1}{W_P} \sum_{x_i \in \omega} I(x_i) f_r(\|I(x_i) - I(x)\|) g_s(\|x_i - x\|) \quad (1)$$

$$W_P = \sum_{x_i \in \varpi} f_r(\|I(x_i) - I(x)\|) g_a(\|x_i - x\|) \quad (2)$$

Next, gamma correction function is applied on the green channel of the filtered image to extract the cell nuclei. Colors used for staining biopsy samples are generally Hematoxylin-Eosin (H&E), in which H dyes the cell nuclei blue-purple, while E dyes connective tissue and cytoplasm pink. Therefore, green and blue channel have the chance to visualize cell nuclei. On the other hand, deformed cell nucleus refers to variations in shape, size and chromatin density. The more advanced is the cancer, the more nucleus become big in size, malformed in shape and dense in chromatin. Consequently, nucleus with higher score is vascular with some parts of chromatin density which are more colorful and nucleus with lower score contains uniform chromatin which is less colorful. As a result, green channel is a better choice since it indicates all type of cancerous nuclei.

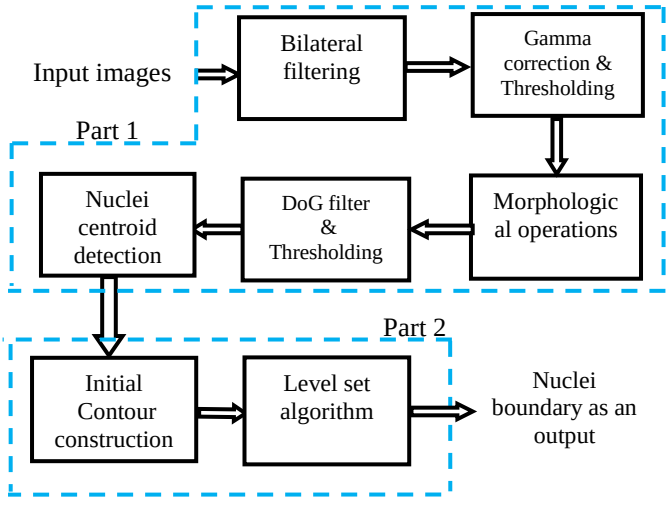


Fig. 1. Block diagram of the proposed segmentation system

The processed image contains closed contour of nuclei in binary image after thresholding. Next, two other morphological operators are applied once on the binary image to dilate and erode it using structuring element with radius equal to one. The following equations are erosion and dilation of image A by structuring element, B, respectively.

$$A \odot B = \{ x \in \varepsilon^2 : \forall s \in B, \exists a \in A \mid x = a - s \} \quad (3)$$

$$A \oplus B = \{ x \in \varepsilon^2 : x = a + s, a \in A, s \in B \} \quad (4)$$

Figure 2, F and G monitor the morphological operations. Dilation, combines nuclei that cluster near each other, whilst isolated nuclei retain isolated. The merged nuclei are disjoint by erosion into individual blobs, while isolated nuclei are eliminated.

Finally, a 2D Difference of Gaussian filter (DoG) with empirically selection of standard deviations equal to 4 and 10, is applied on the blobs detected in the previous step. Figure 2, H shows that the detected blobs are perceived as small edges in DoG filtering, so the objects that are obtained by thresholding the filtered image, are a region inside each nuclei, in which their centers, are nuclei centers as shown in Fig. 3.

B. Cell nuclei boundary segmentation

Boundary detection is performed using level set algorithm, that is a set of mathematical methods, which aims at identifying points at which the image brightness changes sharply. Level set method is heavily dependent on an initial guess of contour, which is then moved by image driven forces to the boundaries of the desired objects. The initial contour is defined using Eq.5, which is based on the level set function with the formulation in Eq.6.

$$B = \{(x, y) \mid \phi(t, x, y) = 0\} \quad (5)$$

$$\frac{\partial \phi}{\partial t} + F |\nabla \phi| = 0 \quad (6)$$

In such models, two types of forces are considered - the internal forces, defined within the curve, are designed to keep the model smooth during the deformation process, while the external forces, which are computed from the underlying image data, are defined to move the model toward an object boundary. Each detected center is dilated in order to construct the initial contour for level set algorithm. The structuring element used for the employed morphological operator should be bigger than the size of nuclei. As shown in Fig. 4, the attained binary mask, is an initial boundary and an initial contour for level set segmentation method. As an output of level set, a nice and smooth boundary is segmented for each nucleus and cancerous nuclei are separated from background.

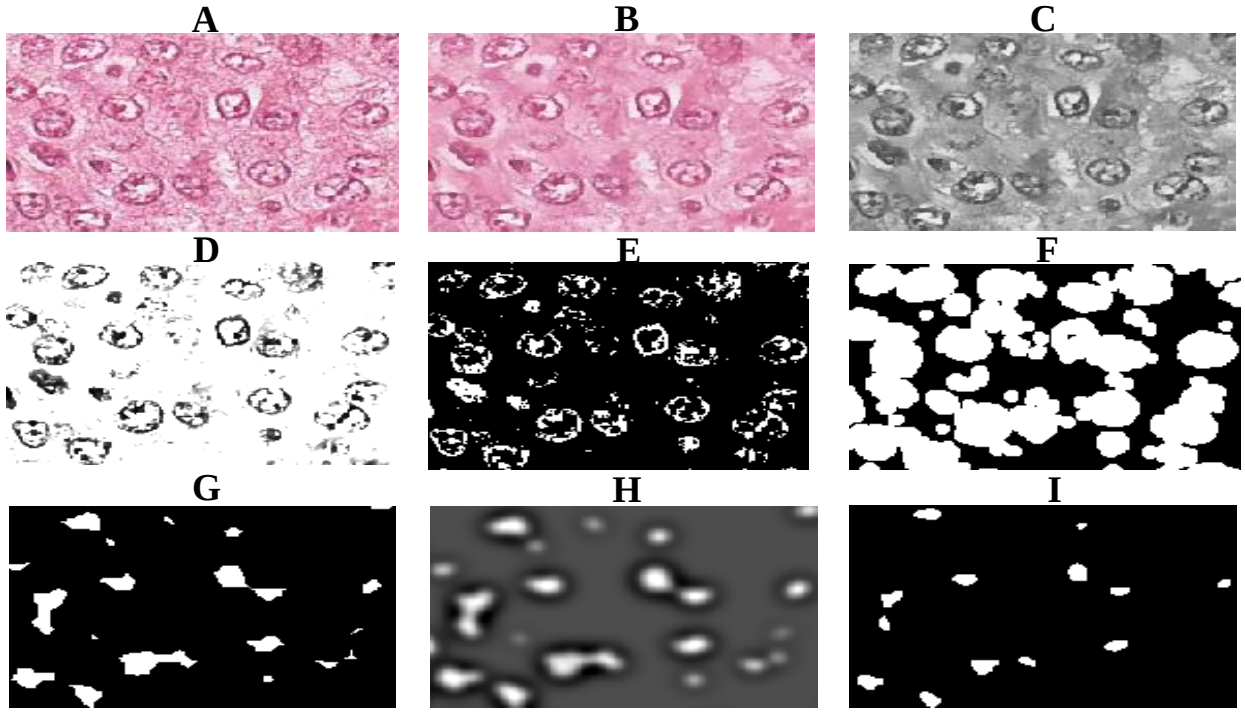


Fig. 2. Detection of center of nuclei, (A) Original image, (B) Filtered image, (C) Green channel of the filtered image, (D) Gamma corrected image, (E) binarized image after applying gamma correction function, (F-G) morphological operations, dilation and erosion, (H) Difference of Gaussian filter applied on the eroded image, (I) threshold mask applied to the image in H

Figure 4 shows the segmentation method is capable of distinguishing the boundary of linked nuclei and considering them as disjointed nuclei.

III. SIMULATION AND RESULT

The team of Professor Frédérique Capron, head of the Pathology Department at Pitié-Salpêtrière Hospital in Paris, France, has selected and annotated a set of breast cancer biopsy slides for International Conference on Pattern Recognition (ICPR) contest of detection of mitosis and evaluation of nuclear atypia score in breast cancer histological images. The slides are stained with standard H&E dyes and they have been scanned by Aperio XT scanner with a scale of 0.2456 μm per pixel and magnified at 20X resolution [18]. 20 images are examined in this test, in which they consist of 4 frames scored 1, 8 frames scored 2 and 8 frames scored 3 according to the ground truth. Score for nuclear atypia requires a wide area to be able to evaluate shape and size of a large population of nuclei. For this task, the pathologists have worked at 20X magnification. The score is given by two experienced senior pathologists. When pathologists disagree about the score to be given to a frame, the opinion of a third pathologist has been requested. In that case, the final score for every frame is the score having the majority among the three pathologists [16].

The performance of the experiment is evaluated by measuring the accuracy. The accuracy is estimated by formula in (6):

$$\text{Accuracy (ACC)} = (TP+TN) / (TP+TN+FP+FN) \quad (6)$$

Where the abbreviations are described as follow:

True Positive (TP) = correctly identified

False Positive (FP) = incorrectly identified

True Negative (TN) = correctly rejected

False Negative (FN) = incorrectly rejected

Although scoring ground truth exists in the contest, there is no ground truth available for assurance of the validity of the detected nuclei and segmented boundaries. Therefore, our team provided a subjective ground truth with the association of an expert pathologist who collaborated with us in this project. In the subjective ground truth, all cell nuclei which influenced the scoring were marked by the collaborative pathologist and all detected cell nuclei candidates and segmentation of their boundaries by the proposed method are compared to the mentioned ground truth and manual segmentation done by her. This new generated ground truth contributes to an adequate and accurate comparison between the obtained results by the proposed system and the manually marked and segmented cell nuclei by the collaborating pathologist which ends is achieving reliable accuracy of the proposed system. The average accuracy for the precision of the detected cell nuclei by the proposed system is 86%, in which the maximum accuracy of detected cell nuclei is 92%. The overall accuracy for segmented boundaries of nuclei is 87%. The error rate which is produced in the procedure is mainly occurred due to staining process since staining process leads to the presence of noisy regions in the image. In case of strong unspecified and noisy background Hematoxylin staining, false positive cell nuclei happens, which is mostly the case the data that was analysed in the proposed method. Moreover, all the critical cell nuclei are not detected. However, according to objective scoring by pathologists assessment, the segmentation is authentic for scoring.

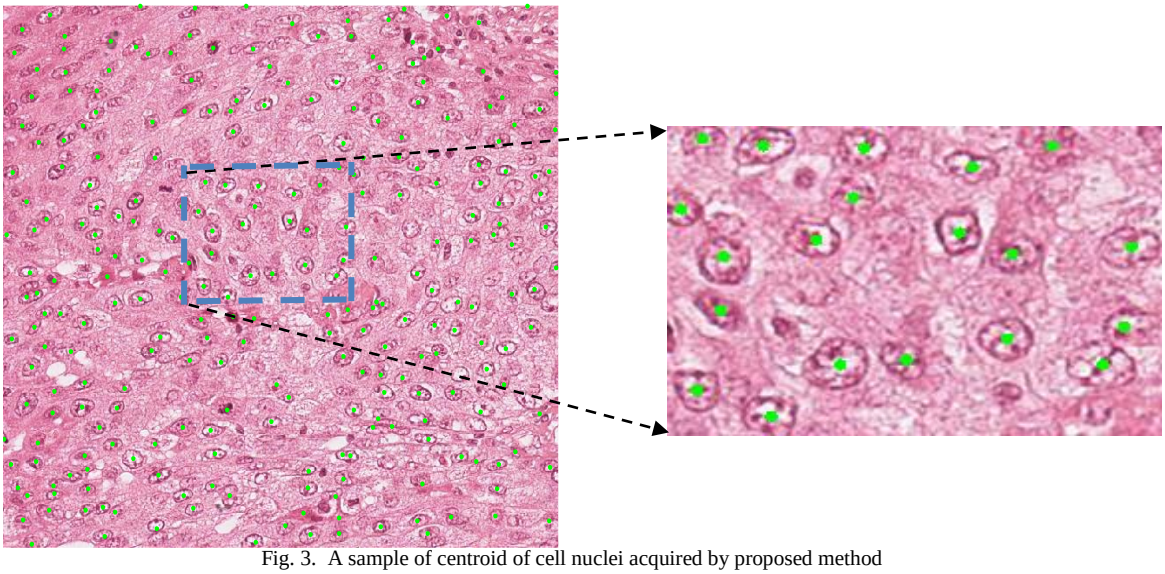
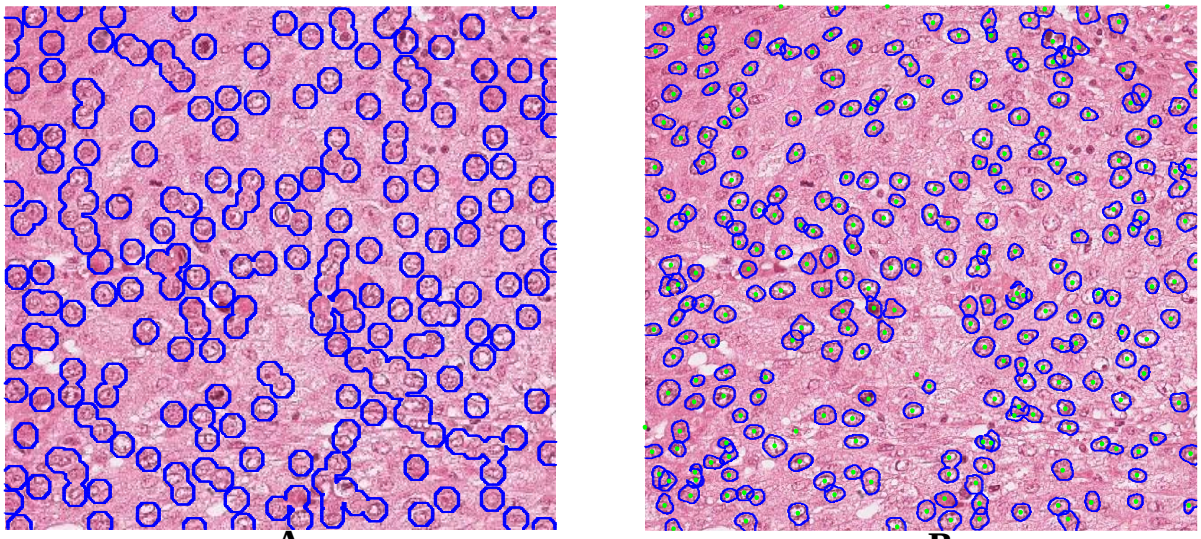


Fig. 3. A sample of centroid of cell nuclei acquired by proposed method



A B

Fig. 4. (A) An initial contour for level set segmentation, (B) A sample of boundary segmentation of nuclei acquired by proposed method

Table 1 shows a Comparison between the result obtained by method proposed in this work and the result achieved by Dalle et al. approach [17], which is the most dominant method in this regard. Since [17] detects critical nuclei with uniform chromatin, it has a higher accuracy error for nuclei scored 2 and 3, but the proposed method is capable of detecting a large variety of vesicular nuclei scored 3, as seen in Table 1.

TABLE I . COMPARISON OF THE PROPOSED METHOD WITH DALLE METHOD

Evaluation results	**Our method**	**Dalle method [17]**
Overall accuracy	87%	84.1 %
Accuracy of frames scored 2	88.5%	84.7%
Accuracy of frames scored 3	85.5%	83.5%

IV. CONCLUSION

Cell nuclei detection is a crucial stage for delineating grade of malignancy based on Nottingham Grading System. Both nuclear pleomorphism and tubule formation require cell nuclei detection as an initial step. In contrast of most existed algorithms, the paper proposed an automatic cell nuclei detection and boundary segmentation method which primary training is not needed and not all the cell nuclei are detected. Only cell nuclei that modify nuclear pleomorphism scoring and are necessary for tubule detection are segmented. Cancerous nuclei are different in several criteria, in which the chromatin density causes the maximum error in detection and segmentation system. In this paper, the mentioned problem is carried out with an average accuracy of 86% for cell nuclei detection and overall accuracy of 87% for boundary segmentation, in a lower error rate compared to other existing methods. We are currently expanding this work to scoring nuclear pleomorphism in a robust and accurate manner, in order to accomplish a single criterion of Nottingham Grading System for breast cancer.

REFERENCES

[1] M. P. Coleman, M. Quaresma, F. Berrino, J.-M. Lutz, R. De Angelis, R. Capocaccia, P. Baili, B. Rachet, G. Gatta, T. Hakulinen, and others, "Cancer survival in five continents: a worldwide population-based study (CONCORD)," *Lancet Oncol.*, vol. 9, no. 8, pp. 730–756, 2008.

[2] A. Veillard, N. Loménie, and D. Racoceanu, "An exploration scheme for large images: application to breast cancer grading," in Pattern Recognition (ICPR), 2010 20th International Conference on, 2010, pp. 3472–3475.

[3] A. Tashk, M. Helfroush, H. Danyali and M. Akbarzadeh-jahromi, "Automatic detection of breast cancer mitotic cells based on the combination of textural, statistical and innovative mathematical features," Applied Mathematical Modelling, vol. 39, no. 20, pp. 6165-6182, 2015.

[4]A. Tashk, M. Sadegh Helfroush, H. Danyali and M. Akbarzadeh, "A Computer-Aided System for Automatic Mitosis Detection from Breast Cancer Histological Slide Images based on Stiffness Matrix and Feature Fusion," CBIO, vol. 10, no. 4, pp. 476-493, 2015.

[5] R. Nateghi, H. Danyali, M. Sadegh Helfroush and A. Tashk, "Intelligent CAD System for Automatic Detection of Mitotic Cells from Breast Cancer Histology Slide Images Based on Teaching-Learning-Based Optimization," Computational Biology Journal, vol. 2014, pp. 1-9, 2014.

[6]K. Nguyen, M. Barnes, C. Srinivas and C. Chefd'hotel, "Automatic glandular and tubule region segmentation in histological grading of breast cancer, " SPIE Medical Imaging, pp. 94200G-94200G, International Society for Optics and Photonics, 2015.

[7] C.-C. Ko, C.-Y. Cheng, and C.-H. Lin, "A Computer-Aided Grading System of Breast Carcinoma: Scoring of Tubule Formation," 2016 30th International Conference on Advanced Information Networking and Applications Workshops (WAINA), 2016.

[8] S. Minaee and Y. Wang, "Screen Content Image Segmentation Using Robust Regression and Sparse Decomposition," IEEE J. Emerg. Sel. Top. Circuits Syst., 2016.

[9] S. Minaee and Y. Wang, "Screen Content Image Segmentation Using Sparse Decomposition and Total Variation Minimization," ArXiv Prepr. ArXiv160202434, 2016.

[10] Maqlin P, Thamburaj R, Mammen JJ, Nagar AK. Automatic Detection of Tubules in Breast Histopathological Images, Springer; 2013, p. 311–21.

[11] Basavanhally A, Yu E, Xu J, Ganesan S, Feldman M, Tomaszewski J, et al. Incorporating domain knowledge for tubule detection in breast histopathology using O'Callaghan neighborhoods. vol. 7963, International Society for Optics and Photonics; 2011, p. 796310.

[12] A. Chekkoury, P. Khurd, J. Ni, C. Bahlmann, A. Kamen, A. Patel, L. Grady, M. Singh, M. Groher, and N. Navab, "Automated malignancy detection in breast histopathological images," in *SPIE Medical Imaging*, pp. 831515–831515, 2012.

[13] M. Veta, P. J. van Diest, R. Kornegoor, A. Huisman, M. A. Viergever, and J. P. Pluim, "Automatic Nuclei Segmentation in H&E Stained Breast Cancer Histopathology Images," *PloS One*, vol. 8, no. 7, p. e70221, 2013.

[14] S. Wienert, D. Heim, K. Saeger, A. Stenzinger, M. Beil, P. Hufnagl, M. Dietel, C. Denkert, and F. Klauschen, "Detection and segmentation of cell nuclei in virtual microscopy images: a minimum-model approach," *Sci. Rep.*, vol. 2, 2012.

[15] Y. Al-Kofahi, W. Lassoued, W. Lee, and B. Roysam, "Improved automatic detection and segmentation of cell nuclei in histopathology images," *Biomed. Eng. IEEE Trans. On*, vol. 57, no. 4, pp. 841–852, 2010.

[16] E. Cosatto, M. Miller, H. P. Graf, and J. S. Meyer, "Grading nuclear pleomorphism on histological micrographs," in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, 2008, pp. 1–4.

[17] J.-R. Dalle, H. Li, C.-H. Huang, W. K. Leow, D. Racoceanu, and T. C. Putti, "Nuclear pleomorphism scoring by selective cell nuclei detection.," in *WACV*, 2009.

[18] Dataset for nuclear atypia scoring in breast cancer histological images: An ICPR 2014 contest. Available from: http://ipal.cnrs.fr/data/mitosis_atypia_2014/icpr2014_Mitosis Atipya_DataDescription.pdf