Segmentation Based Classification for Mitotic Cells Detection on Breast Histopathological Images

Tan Xiao Jian[1], Nazahah Mustafa[1], Mohd Yusoff Mashor[1]and Khairul Shakir bin Ab Rahman[2]
[1]*School of Mechatronic Engineering, University Malaysia Perlis, 02600 Arau, Perlis, Malaysia.*
[2]*Hospital Tuanku Fauziah, 01000 Kangar, Perlis, Malaysia.*
xj_0506@hotmail.com

***Abstract*—Breast cancer grading is the standard clinical practice for the prognosis and diagnosis of breast cancer development. The Nottingham Histological Grading (NHG) system is widely used in the breast cancer grading. One of the main criteria in assessing breast cancer is mitotic counts as it reflects the speed of cell division in cancer cells. Detection of mitotic candidates could be performed by implementing image processing techniques. The accuracy of mitotic cells detection is dependent on the number of mitotic candidates. Thus, minimizing the number of mitotic candidates is a crucial step in optimizing accuracy. This study proposed a segmentation based classification method to minimize number of false positive in mitotic candidates. The results show that the proposed segmentation based classification method could provide a promising result by achieving an average effectiveness of 91.85% in minimizing the mitotic candidate number.**

***Index Terms*—Breast Cancer; Classification; Mitotic Cell Candidate; Segmentation.**

I. INTRODUCTION

Breast cancer is a main cause of mortality that commonly occurred among women. It is a complex, heterogeneous and fatal disease [1]. In 2015, 6.42 deaths per 100,000 people were recorded in Malaysia [2].

The grading of breast cancer plays a vital role in prognosis and treatment planning of the cancer. Nottingham Histological Grading (NHG) system is commonly used as grading system for breast cancer worldwide [3]. Patient who is diagnosed with breast cancer is subjected to surgical biopsy. The result from biopsy (ie., breast tissues) is then examined under a light microscope. In routine histological practice, a pathologist performs breast cancer grading manually using a light microscope. This resulted in a tedious and massive workload which is time consuming [4]. Besides, manual grading is found to be not reproducible [5, 6]. The output grading may suffer from inter- and intra- observes variability among pathologist [4]. These issues are then initiated a digitalization era in pathology labs.

One of the main features that is considered during the grading process is mitosis activity. Mitosis is a basic part of cell division which involved a division of single cell into two daughter cells. In breast cancer, the number of mitotic counts in 10 high power field (HPF) is vital as it reflects the speed of cell division in cancer cells. In other words, mitotic count describes how fast a cancer spread. In the recent decades, extensive studies have been conducted on mitotic cells detection. There are also studies that worked on the methods to reduce the number of false positive in mitosis candidates [7 - 9].

Nateghi et al. implemented Teaching-Learning-Based optimization (TLBO) to reduce the mitotic candidates [8]. TLBO represent the number of false positive with a set of cost function. Therefore, by minimizing this function, the majority of the non-mitotic cells can be eliminated. Experimental results demonstrate that TLBO method had a performance of 77.34% F-measure. A study by [9] proposed a Genetic Optimization algorithm to reduce the number of mitotic candidates. The number of non-mitosis candidates are defined as a cast function. By minimizing the function with Genetic Optimization algorithm, most of the non-mitosis candidates could be omitted. This method had achieved 78.47% F-measure. Khan et al. [7] employed the Gamma-Gaussian Mixture Model (GGMM) to reduce the pool of mitotic candidates. The GGMM is a parametric technique to estimate the probability density function of the input images. The Gamma function represented the mitotic cells whereas the Gaussian function represented the non-mitotic region. Besides, a post processing method, Context Aware Post processing (CAPP) was applied to the results to reduce the false positive detection.

The main objective of this study is to minimize the number of mitotic candidates. A minimum number of mitotic candidates could enhance the accuracy in identification of true mitotic cells. It is important to note that the true mitotic cells (ie., ground truth) should not be eliminated at the end of the algorithm. This study takes hyperchromatic nucleus as input to be implemented with the proposed method to obtain a minimum number of mitotic candidates as output.

II. METHODOLOGY

The results of the segmented hyperchromatic nucleus candidates obtained from our previous study were used in this study as input images to detect mitosis candidates. Morphological classification and Otsu thresholding were implemented on the input images to segment the potential mitotic candidates. Morphological classification consists of three stages. Stage 1 was used to characterize the input nucleus candidates based on their morphological features. The nucleus candidates were categorized into five different shapes. Stage 2 is a decision making steps. In this stage, a checking procedure was proposed for each nucleus shape to mark the potential mitotic candidates. In stage 3, the resultant marked mitotic candidates from each group were combined. To further minimize the number of potential mitotic candidates, Otsu thresholding was implemented on the output image of morphological classification. Figure 1 illustrates the flow chart of the proposed method.

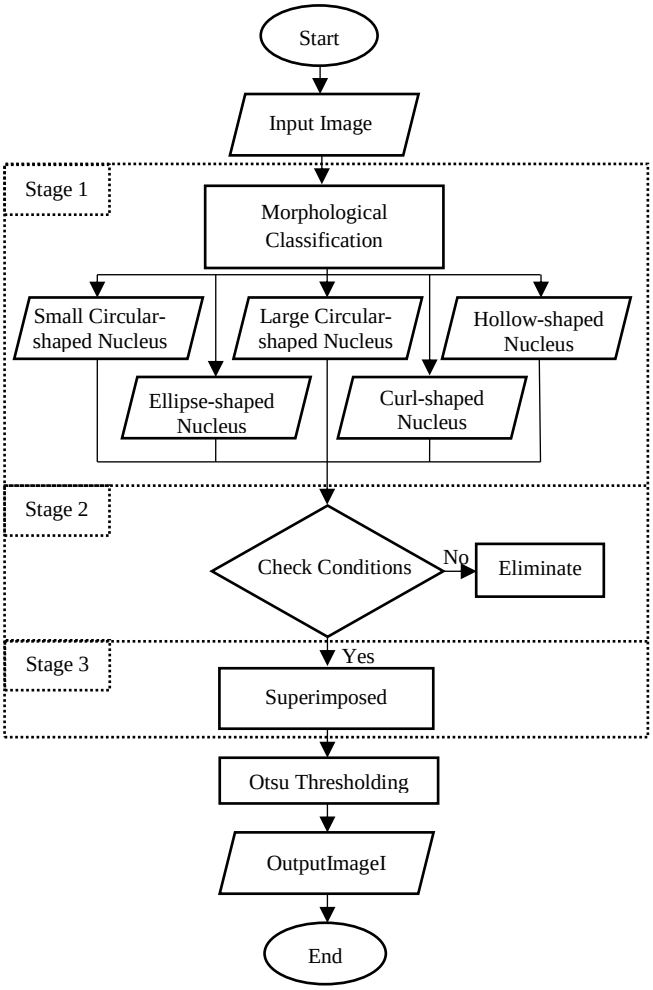

Figure 1: Flow chart of the proposed method.

A. Morphological Classification

Morphological classification consists of three stages and was developed based on the morphological features of the input nucleus candidates. Five parameters: aspect ratio (*AR*), solidity (*S*), ratio of perimeter to area (*PA*), form factor (*FF*) and filled area(*FA*) were first extracted from each of the input nucleus candidates. The parameters are shown in Equation (1) to (5) respectively.

$$AR = \frac{m1}{m2} \tag{1}$$

where: *AR* = Aspect ratio
m1 = Major axis length
m2 = Minor axis length

$$S = \frac{a}{ca} \tag{2}$$

where: *S* = Solidity
a = Area
ca = Convex area

$$PA = \frac{p}{a} \tag{3}$$

where: *PA* = Ratio of perimeter to area
p = Perimeter
a = Area

$$FF = \frac{4 * \pi * a}{p^2} \tag{4}$$

where: *FF* = Form factor
a = Area
p = Perimeter

$$FA = x - y \tag{5}$$

where: *FA* = Filled area
x = Number of pixels enclosed by the object
y = Number of pixels of the object

Stage 1 characterizes the input nucleus candidates into five different shapes based on the proposed parameters. The five different shapes were small circular-shaped nucleus, large circular-shaped nucleus, hollow-shaped nucleus, ellipse-shaped nucleus, and curl-shaped nucleus.

a. Small Circular-shaped Nucleus

In Stage 1, the input nucleus candidate with an*AR* lower than 2 AND a *PA* lower than 0.8 is characterized as small circular-shaped nucleus. In Stage 2, the nucleus in this group were eliminated as these nucleus were assumed as noise.

b. Large Circular-shaped Nucleus

In Stage 1, the input nucleus candidate with an*AR* lower than 2.0 AND a *PA* lower than 0.8 AND a *FF* greater than 0.3 is characterized in this group. In Stage 2, the checking procedure eliminate nucleus candidates with area smaller than 105 pixels.

c. Hollow-shaped Nucleus

In Stage 1, input nucleus candidate with a *FA* greater than 30 pixels is characterized in this group. In Stage 2, the minor axis length (*m2*) of each nucleus in this group were obtained. A window is developed based on size of *m2*. The window is located at the centroid of respective nucleus candidate. The overall mean intensity of the window is calculated. The checking procedure eliminates nucleus with the mean intensity greater or equal to 0.8.

d. Ellipse-shaped Nucleus

In Stage 1, input nucleus candidate with an*AR* greater than 2.0 is characterized in this group. In Stage 2, the checking procedure eliminates nucleus with an area smaller than 105 pixels.

e. Curl-shaped Nucleus

In Stage 1, the input nucleus candidate with an*AR* lower than 2.0 AND a *PA* lower than 0.8 AND a *FF* lower than 0.3 is characterized in this group. In Stage 2, the minor axis length (*m2*) of each nucleus in the group were obtained. A window is developed based on the size of *m2*. The window is located on the centroid of each nucleus. The overall mean intensity of each window is calculated. The checking procedure eliminates nucleus with an average mean intensity greater or equal to 0.8.

Stage 2 eliminates the non-mitotic candidates based on the proposed conditions. These conditions were developed based on the prior study on the morphological features of mitotic cells. The five extracted parameters were used to build the checking procedure. Stage 3 combines the resultant candidates from each of the groups to form a final pool of mitotic candidates.

B. Otsu Thresholding

An Otsu thresholding [10] was employed to further reduce the pool of mitotic candidates. Based on the prior knowledge, only hyperchromatic cells could probably be the mitotic candidates [11, 12]. Thus, segmentation on low intensity nucleus (ie., hyperchromatic nucleus) could reduce the false positive in the mitotic candidates. The output mitotic candidates obtained after implementation of Morphological classification were used as input to Otsu thresholding. The mean intensities of the mitotic candidates for the whole image was calculated. A constant (k) was selected to obtain the threshold value (T_1) using Equation (6). Equation (7) shows how the Otsu thresholding was implemented.

$$T_1 = m * k \tag{6}$$

where: T_1 = Threshold value
m = Overall mean
k = k value

$$g(x,y) = \begin{cases} 1, & f(x,y) \le T_1 \\ 0, & otherwise \end{cases} \tag{7}$$

where: $g(x,y)$ = Output pixel
$f(x,y)$ = Input pixel
T_1 = Threshold value 1

The effectiveness of the proposed method can be calculated by using Equation (8). The percentage of effectiveness describes how effective is the proposed method to minimize the mitotic candidates.

$$PE = \frac{i - o}{i} x100\% \tag{8}$$

where: PE = Percentage of effectiveness
I = Number of nucleus candidates on the input image
o = Number of mitotic candidates on the output image

III. RESULTS

The breast cancer tissue slides used in this study were provided by the Pathology Department, Hospital TuankuFauziah, Kangar, Perlis. These slides were stained with Hematoxylin and Eosin (H&E). The digital slides were prepared using an Aperio CS2 WSI scanner. Six images were captured from the digital slides under 40x magnification. The six images were named as MC1, MC2, MC3, MC4, MC5 and MC6. The resolution for MC1, MC2 and MC3 are 614 x 1264 pixels whereas the resolution for MC4, MC5 and MC6 are 630 x 1360.

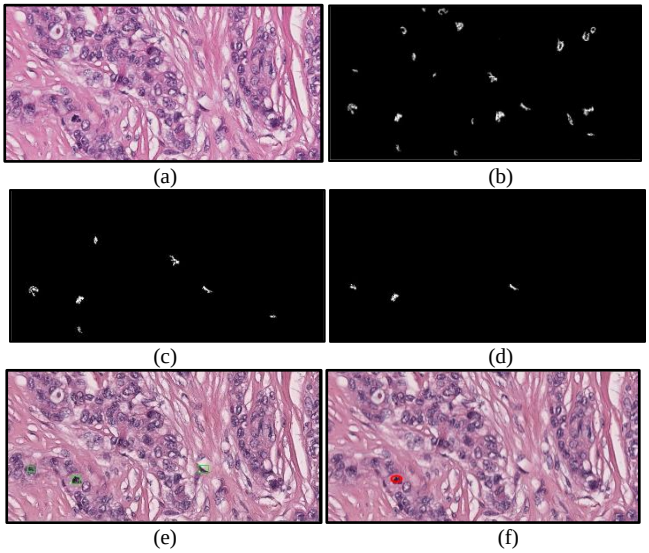

Figure 2: Results of MC1, (a) original histopathological image, (b) input image, (c) output from morphological classification, (d) output from Otsu thresholding, (e) final mitotic candidates superimposed on the original image, (f) ground truth.

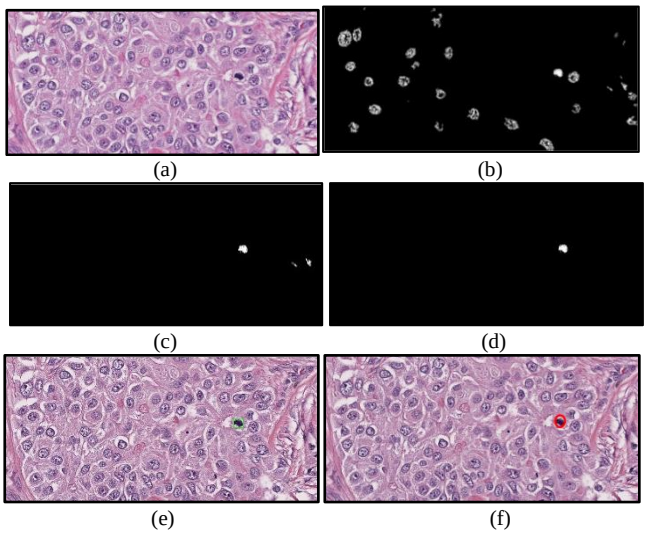

Figure 3: Results of MC4, (a) original histopathological image, (b) input image, (c) output from morphological classification, (d) output from Otsu thresholding, (e) final mitotic candidates superimposed on the original image, (f) ground truth.

Table 1
Comparison Between the Number Nucleus Candidates on the Input Image, Number of Mitotic Candidates after Morphological Classification, Number of Final Mitotic Candidates and the Number of Mitotic Cells in The Ground Truth.

Images	Mitotic Candidates				
	Number of nucleus candidates on input image (*i*)	Number of Mitotic Candidates After Morphological Classification	Number of Final Mitotic Candidates	Number of true mitotic cells (*o*)	Percentage of Effectiveness (*PE*)
MC1	23	8	3	1	86.96%
MC2	16	6	2	1	87.50%
MC3	29	5	2	1	93.10%
MC4	33	3	1	1	96.97%
MC5	53	14	4	1	92.45%
MC6	51	4	3	1	94.12%

Figures 2 and 3 show the results for MC1 and MC4 respectively. For each figure, image (a) shows the original histopathological image, image (b) shows the input image, image (c) shows the result after implementation of morphological classification, image (d) shows the results after the Otsu thresholding, image (e) shows the final mitotic candidates superimposed on the original image and image (f) shows the ground truth. Quantitative results on percentage of classification effectiveness are shown in Table 1. Table 1 also shows the number of nucleus candidates on the input image, number of mitotic candidates after implementation of morphological classification, final mitotic candidates after implementation of Otsu thresholding and the number of mitotic cells in the ground truth for all six images.

IV. Discussion

The main purpose of this study is to minimize the number of mitotic candidates. It is important to note that the true mitotic cells (ie., ground truth) should not be eliminated at the end of the implementation of the proposed method. Based on results in Figures 2 and 3, the proposed method has successfully minimized the number of mitotic candidates while the ground truth is retained. The final mitotic candidates are marked with green square, Figure 2 (e) and Figure 3 (e), whereas the ground truth is marked with red circle, Figure 2 (f) and Figure 3 (f).

Besides, the effectiveness of the proposed method can be calculated based on the number of input nucleus candidates and the final mitotic candidates obtained from the proposed method, Equation (8). The percentage of effectiveness describes how effective is the proposed method to minimize the mitotic candidates while retain the ground truth. The percentage of elimination for MC1, MC2, MC3, MC4, MC5 and MC6 are 86.96%, 92.45%, 87.50%, 94.12%, 93.10% and 96.97%, respectively. The proposed method shows a promising results and proved to be useful for removing non-mitotic candidates. This could significantly enhance the detection of true positive while reducing false positive detection in mitotic cells for breast cancer.

V. Conclusion

This paper proposed an effective approach to detect mitotic candidates in breast cancer. The proposed method shows a promising results by achieving an average effectiveness of 91.85% in minimizing the mitotic candidates number. By applying the prior knowledge in mitotic cells, the number of mitotic candidates could be minimized without eliminating the ground truth. The proposed method could facilitate the study of mitotic count in breast cancer in the future.

Acknowledgment

This work is supported by the Fundamental Research Grant Scheme (FRGS) under Ministry of Higher Education Malaysia (9003-00589).The protocol of this study had been approved by Medical Research and Committee of National Medical Research Register (NMRR) Malaysia (NMRR-17-281-34236).

References

[1] Weigelt, B. And Reis-Filho, J. S., "Histological And Molecular Types Of Breast Cancer: Is There A Unifying Taxonomy," Nat Rev ClinOncol, 6, 718-730. 10.1038/Nrclinonc.2009.166, 2009.

[2] GBD Compare. [Online] Http://Vizhub.Healthdata.Org/Gbd-Compare/ (Accessed 12 December 2016).

[3] H. J. G. B. &W.W.Richardson, "Histological Grading And Prognosis Of Breast Cancer," Vol. 22, No. 1, Pp. 36–37, 1957.

[4] O. Sertel, U. V. Catalyurek, H. Shimada, And M. N. Gurcan, "Computer-Aided Prognosis Of Neuroblastoma: Detection Of Mitosis And Karyorrhexis Cells In Digitized Histological Images," Proc. 31st Annu. Int. Conf. Ieee Eng. Med. Biol. Soc. Eng. Futur. Biomed.Embc 2009, Pp. 1433–1436, 2009.

[5] M. Veta, P. J. Van Diest, R. Kornegoor, A. Huisman, M. A. Viergever, And J. P. W. Pluim, "Automatic Nuclei Segmentation In H&E Stained Breast Cancer Histopathology Images," Plos One, Vol. 8, No. 7, Pp. 1–12, 2013.

[6] E. A Rakha, J. S. Reis-Filho, F. Baehner, D. J. Dabbs, T. Decker, V. Eusebi, S. B. Fox, S. Ichihara, J. Jacquemier, S. R. Lakhani, J. Palacios, A. L. Richardson, S. J. Schnitt, F. C. Schmitt, P.-H. Tan, G. M. Tse, S. Badve, And I. O. Ellis, "Breast Cancer Prognostic Classification In The Molecular Era: The Role Of Histological Grade.," Breast Cancer Res., Vol. 12, No. 4, P. 207, 2010.

[7] A. M. Khan, H. Eldaly, And N. M. Rajpoot, "A Gamma-Gaussian Mixture Model For Detection Of Mitotic Cells In Breast Cancer Histopathology Images.," J. Pathol. Inform., Vol. 4, No. Icpr, P. 11, 2013.

[8] R. Nateghi, H. Danyali, M. SadeghHelfroush, And A. Tashk, "Intelligent Cad System For Automatic Detection Of Mitotic Cells From Breast Cancer Histology Slide Images Based On Teaching-Learning-Based Optimization," Comput. Biol. J., Vol. 2014, Pp. 1–9, 2014.

[9] F. P. P. RaminNateghi, HabibollahDanyali, Mohammad SadeghHelfroush, "Automatic Detection Of Mitosis Cell In Breast Cancer Histopathology Images Using Genetic Algorithm," 21st Iran. Conf. Biomed.Eng. (Icbme 2014), Pp. 5–10, 2014.

[10] N. Otsu, "A Threshold Selection Method From Gray Level Histograms.," Ieee Transactions On Systems, Man And Cybernetics, Vol. 9, No. 1, Pp. 62-66, 1979.

[11] V. M., P. J.P.W., V. D. P.J., And V. M.A., Breast Cancer Histopathology Image Analysis: A Review, Vol. 61, No. 5. 2014.

[12] M. Veta, P. J. Van Diest, S. Willems, H. Wang, A. Madabhushi, F. Gonzalez, A. A C. Roa, A. B. L. Larsen, J. S. Vestergaard, B. Dahl, D. C. Cireşan, J. Schmidhuber, A. Giusti, And M. Luca, "Assessment Of Mitosis Detection Algorithms In Breast Cancer Histopathology Images," Med. Image Anal., Vol. 2013, Pp. 1–21, 2013.