Automatic Evaluation System of FISH Images in Breast Cancer

Tomasz Les¹, Tomasz Markiewicz^{1,2}, Stanisław Osowski^{1,3}, Marzena Cichowicz²,
and Wojciech Kozłowski²

¹Warsaw University of Technology

²Military Institute of Medicine

³Military University of Technology, Warsaw, Poland

lestomasz@gmail.com, {markiewt,sto}@iem.pw.edu.pl,
{mcichowicz,wkozlowski}@wim.mil.pl

Abstract. The paper presents the algorithm of an automatic evaluation of the fluorescent in situ hybridization (FISH) images in order to determine HER2 status of the breast cancer samples. The algorithm is based on the accurate measurement of the red/green spot ratio (the ratio of HER2/CEN17) per cell nucleus. The main points of this algorithm is an accurate detection of the nuclei of cells and application of the color map to detect the red and green spots localized in these cells. The results of the numerical experiments concerning these two problems, as well as the assessment of the positivity or negativity of the considered cases are presented and discussed in the paper. They confirm the high efficiency of the proposed solution.

Keywords: image segmentation, object recognition, FISH images.

1 Introduction

The fluorescence in situ hybridization (FISH) test performed on the breast cancer tissue removed during the biopsy allows to visualize the specific genes to discover if extra copies of the HER2 gene have been created. More copies of the HER2 gene mean that more HER2 receptors the cells have, which stimulate the growth of the breast cancer cells.

The probes carry the fluorescent markers emitting the light after binding to the HER2 genes. These HER2 probes are visible as the red/orange stained spots under a fluorescent microscope. On the other hand the probes for the chromosome centromere (CEN17) on which HER2 is located are visible as the green spots. The ratio between the number of the red spots over the number of green spots estimated for the investigated case, decides whether the cancer is positive or negative for HER2 and decides of the way of the medical treatment. Inaccurate results of HER2 test may cause application of the not accurate diagnosis and as a result the not optimal treatment.

The manual analysis of the FISH images is a time consuming and prone to errors, since it involves the manual counting of hundreds of the FISH spots over the tissue slides. Recently a lot of automatic approaches to this problems have been addressed

and published in the scientific papers [1],[2],[3],[4],[5]. The significant problem arises in the assessment of positive cases. The reported accuracy is changing from 74.1% to 95.2% for positive cases [2]. However, all of these results have been obtained for the limited number of analysed images. There is still the need to investigate more images and improve the accuracy of detection of the FISH spots. Such an improvement will result in more accurate final decision of the positivity or the negativity of the analysed cases. This paper will present the new approach to the identification of the red and green spots in the FISH images, which allows to reduce the number of their erroneous localization. Its main novelty is the introduction of the map of colors allowing to get better precision in estimation of the colors of the groups of pixels. The results of the statistical evaluation of the presented approach have confirmed good performance of the developed automatic system.

2 Problem Statement

The typical FISH images of the breast cancer tissues are presented in Fig 1. We can see the groups of blue nuclei cells with the green and red spots, representing the objects under recognition. As we can see the images differ significantly by the distribution of cells, their colors and also the view of the background.

Fig. 1. Two typical examples of the FISH images of the breast cancer representing: a) the negative case, b) the positive case

The aim of the presented analysis is to identify the nuclei regions, and then localize the green and red spots located within their area. These spots represent the HER2 and CEP17 probes. By calculating the average ratio of the number of red (HER2 gene copies) to green (CEP17) spots we can decide whether the case is positive (the ratio higher than 2.2) or negative (the ratio below 1.8).

3 Proposed Solution

The breast tissue FISH slides were prepared using the standard procedure in a similar way as described in [1]. This task was performed in Military Institute of Medicine in

Warsaw. Our developed computerized system reads the multiple images of a specific case and performs their analysis. The solution of the image analysis was split into two tasks. In the first phase the segmentation of nuclei is done, and in the second phase the extracted nuclei are searched for detection of the red and green spots. In the final step we perform the counting of the red and green FISH spots discovered in the previous phase. On the basis of such analysis the overall case classification is made, declaring it as the positive or negative. The results of such computerized analysis are compared to the corresponding results of the human expert estimation.

3.1 Segmentation of Cell Nuclei

Segmentation of the cell nuclei is performed on the blue component of the FISH image. To correct the inhomogeneous intensity of the blue channel the nonlinear correction of color is made first. The segmentation of the nuclei is done in two phases. In the first the Otsu thresholding [6] separates the cells or groups of cells differing significantly from the background. Since the cell nuclei are packed very densely and are overlapping, the next step is needed to separate the touching cells. We have solved this problem by applying the watershed algorithm [7], separating the touching cells.

Fig. 2 presents two exemplary images of the segmented cell nuclei after application of both steps of segmentation. The original images have been presented in Fig. 1. As we can see most of the touching cells (from 85% to 90%) have been separated from each other.

Fig. 2. The results of segmentation of the nuclei of two FISH images of the breast cancer representing: a) the negative case, b) the positive case

3.2 Detection of Spots

The detection of spots is done on the red and green channels of the image. To reduce the effect of stains existing in the image we apply its filtering by darkening its central regions using the linear projection of pixel intensities (balancing the intensity of image). This projection has the form of the cone in the centre of the image. In this way we compensate the non-uniformity of lighting of the image, resulting from its

acquisition in the microscope. The next important preprocessing step is the transformation of color channels, leading to the enhancement of the red and green pixels. Two images are then created. In the first one the red and green channels are changed to red, while the blue one to the green. In the second image the red and green channels are changed to the green, while the blue to the red. Thanks to such changes we enhance the visible contrast between the green and the rest and between the red and the rest respectively.

Our main idea of the spot detection is based on the map of colors. This map is created in the form of the parallel vertical and horizontal lines representing equal intensity of one color (for example red) and gradually increasing the intensity (from 0 to 255) of the second one (for example green). Each pixel of the image is associated with the proper place of the map. The pixels are grouped into the clusters representing the potential spots using the k-NN (k-nearest neighbour) technique. The cluster is valid as the spot if all pixels are inside of the circle of the assumed radius (in experiments this radius was chosen as 20 pixels). The same procedure was applied to both transformed images (one image enhancing the recognition of the green, and the second of the red).

Fig. 3. The diagram of the FISH spot detection algorithm

Figure 3 presents in a graphical form the general diagram of the spot detection applied in the solution. It was implemented in an user friendly form using Matlab [8]. In a parallel way we calculate the ratio of the intensity of the red to the green colors (in discovering the red spots) and the green to the red (in discovering the green spots) for the succeeding groups of neighbouring pixels inside of the nuclei. If the average ratio is above some threshold value the cluster is treated as the valid spot of the appropriate color. In the last step we integrate the results of both parallel processing (the union of them) into the final result of the spot detection.

Fig. 4. The examples of the detected red and green spots in the FISH images representing: the negative case, b) the positive case

The exemplary results of the spot detection for the images of Fig. 2 are presented in Fig. 4. The nuclei containing the red and green spots are visible. It is evident that the ratio of the number of the red and green spots is different from the nuclei to nuclei.

In the image representing the negative case this ratio did not exceed the value of 3, while in the positive case we can observe the ratio equal 7. In general there is a significant statistical difference between the number of red and green spots in the FISH images representing the positive and negative cases. Much larger ratio is observed for the positive case.

3.3 Classification of FISH Cases

The most important information obtained from the analysis of the FISH images is determination whether the analysed case is a positive or negative. This is usually done on the basis of the average ratio between the number of the red and green spots among all valid nuclei in the series of images representing the particular case. The case is classified as the positive, if the average ratio is above 2.2, and the negative when this ratio is below 1.8. The gap (1.8 – 2.2) means not the fully determined results and needs further study using the additional methods of diagnosis.

The typical histograms of the red/green ratio, representing the positive and negative cases are shown in Fig. 5. They correspond to 6 analysed images (patients) represented by 20 nuclei in the positive and 20 nuclei in the negative case and depict the number of the nuclei among these 20 analysed nuclei of the particular red/green ratios. As we can see for the negative cases the multiplicity was changing from 0.25 up to 3. In the case of positive cases this ratio was varying from 1 up to 8. It is evident, that in all positive cases the highest stripes correspond to the ratio above 2.2, while in the negative case below 1.8.

Fig. 5. The histograms of the red/green ratio formed for 20 nuclei in the case of the negative (1,2,3) and positive (4,5,6) cases of the FISH images

The estimated average ratios of the red/green in the three negative cases were: 0.82, 1.04, 1.06, respectively, while in the positive cases these ratios were equal: 2.39, 2.58, 2.52 (all above 2.2).

4 Results of Statistical Evaluation of Many Cases

In the numerical experiments we have analyzed 5 positive and 5 negative cases (patients), selected by the experts. In each case we processed up to 3 images in order to obtain 40 nuclei, selected automatically by the program on the basis of the analyzed images. The number of nuclei taken into account in the analysis is in accordance to the Dako HER2 FISH pharmDx™ interpretation guide for the breast cancer. In assessing the individual cases we have applied our automatic computerized system, denoted further by AS. The assessments of the red/green ratio were also done for the same set of images representing each patient by the medical expert. The results presenting the total number of HER2 spots and CEN17 spots as well their HER2/CEN17 ratio for the analysed 10 cases are presented in the Table 1.

Table 1. The results of counting total number of HER2 and CEN17 signals for 10 cases. AS means the automatic computerized system

HER2		CEN17		HER2/CEN17 ratio		CASE	
AS	Expert	AS	Expert	AS	Expert	AS	Expert
54	57	71	73	0.76	0.78	negative	negative
58	59	39	42	1.49	1.41	negative	negative
71	73	72	74	0.99	0.99	negative	negative
89	93	76	77	1.17	1.21	negative	negative
49	51	46	49	1.07	1.04	negative	negative
127	132	57	59	2.23	2.24	positive	positive
176	184	78	79	2.26	2.33	positive	positive
190	197	83	86	2.29	2.29	positive	positive
162	168	55	56	2.95	3.00	positive	positive
118	125	52	54	2.27	2.31	positive	positive

It should be noted, that our computerized system is fully automatic. Hence the selection of the nuclei which are subject to the estimation of the red/green ratio has been done according to the automatic procedure defined within the program. The user does not interfere in this process. On the other hand the human expert selects the nuclei according to his professional knowledge, blind the selection results of the automatic system. This is the reason why the sets of nuclei subject to further analysis were different in both approaches to the problem.

Different sets of nuclei analyzed by the automatic system and by the expert in each case have caused, that the number of HER2 and CEN17 spots in both approaches are slightly different. The highest discrepancy rate between both results was observed for the HER2 genes and was equal 6.5%. However, in spite of this we can observe the close resemblance of the HER2/CEN17 ratios in both approaches. In most cases

(9 out of 10) the discrepancy rate was below 3%. Only in one (negative) case this rate achieved the value of 5.7%. In spite of this for all analyzed cases we have got the same final assessments of them. In this respect there is a full agreement between the automatic system and the expert results for both positive and negative cases (100% accuracy).

5 Conclusions

The paper has presented the new approach to the automatic evaluation of the HER2 status in the breast samples applying the FISH image analysis. The developed algorithm uses two basic steps of analysis: the segmentation of the cell nuclei and the detection of the red and green spots followed by the calculation of their quantities. As a result the calculation of the average red/green ratio per cell nucleus is performed in the last step. Consequently, the collective classification of cases, in a manner similar to the clinician's evaluation is made.

The results of the numerical experiments performed for 10 patients belonging to negative and positive cases have confirmed good, acceptable accuracy of the estimation of HER2/CEN17 ratio. We have also got full agreement of the final assessments of the considered cases made by our system and by the medical expert. Despite the small number of cases the evaluation results were encouraging for further testing of the developed automatic system.

References

1. Raimondo, F., Gavrielides, M., Karayannopoulou, G., Lyroudia, K., Pitas, I., Kostopoulos, I.: Automated evaluation of HER2/neu status in breast tissue from FISH images. *IEEE Transactions on Image Processing* 14, 1288–1298 (2005)
2. Theodosiou, Z., Kasampalidis, I.J., Ksrayannopoulou, G., Kostopoulos, I., Bobos, M., Bevilacqua, G., Aretini, P., Starita, A., Lyloudia, K., Pitas, I.: Evaluation of FISH image analysis system on assessing HER2 amplification in breast carcinoma cases. *The Breast* 17, 82–86 (2008)
3. Lerner, B., Clocksin, W., Dhanjal, S., Hulten, M., Bishop, C.: Automatic signal classification in FISH images. *Cytometry* 43, 87–93 (2001)
4. Masmoudi, H., Hewitt, S., Patrick, N., Myers, K., Gavrielides, A.: Automated quantitative assessment of HER-2/neu immunohistochemical expression in breast cancer. *IEEE Transaction on Medical Imaging* 28, 916–925 (2009)
5. Lopez, C., Salvado, M., Tomas, B., Llobera, M., Alvaro, T., Bosch, R., Garcia-Rojo, M., Jaen, J., Korzynska, A., Lejeune, M.: Is it necessary to evaluate nuclei in HER2 FISH evaluation? *American Journal of Clinical Pathology* 139, 47–54 (2013)
6. Otsu, N.: A threshold selection method from grey-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics* 9, 62–66 (1979)
7. Soille, P.: *Morphological Image Analysis, Principles and Applications*. Springer, Berlin (2003)
8. Matlab Image Processing Toolbox, user's guide, MathWorks, Natick (2012)