Automated evaluation of Her-2/neu status in breast tissue from fluorescent in situ hybridization images

F. Raimondo, M. A. Gavrielides, G. Karayannopoulou, K. Lyroudia,

I. Pitas, *Senior Member* IEEE, I. Kostopoulos

***Abstract*—The evaluation of fluorescent in situ hybridization (FISH) images is one of the most widely used methods to determine Her-2/neu status of breast samples, a valuable prognostic indicator. Conventional evaluation is a difficult task since it involves manual counting of dots in multiple images. In this paper we present a multistage algorithm for the automated classification of FISH images from breast carcinomas. The algorithm focuses not only on the detection of FISH dots per image but also on combining results from multiple images taken from a slice for overall case classification. The algorithm includes mainly two stages for nuclei and dot detection respectively. The dot segmentation consists of a top-hat filtering stage followed by template matching to separate real signals from noise. Nuclei segmentation includes a non-linearity correction step, global thresholding to identify candidate regions and a geometric rule to distinguish between holes within a nucleus and holes between nuclei. Finally, the marked watershed transform is used to segment cell nuclei with markers detected as regional maxima of the distance transform. Combining the two stages allows the measurement of FISH signals ratio per cell nucleus and the collective classification of cases as positive or negative. The system was evaluated with receiver operating characteristic (ROC) analysis and the results were encouraging for the further development of this method.**

***Index Terms*— fluorescent in situ hybridization, breast carcinomas, spot detection, nuclei segmentation, case classification.**

Manuscript received December 1, 2004. This work was supported by the EU project Biopattern: Computational Intelligence for biopattern analysis in Support of eHealthcare, Network of Excellence Project No. 508803.

F. Raimondo is with the Department of Informatics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece.

M. A. Gavrielides was with Department of Informatics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece.

G. Karayannopoulou is with the Department of Pathology, Medical School, Aristotle University of Thessaloniki, 54124, Thessaloniki, Greece.

K. Lyroudia is with the Department of Endodontology, Dental School, Aristotle University of Thessaloniki, 54124, Thessaloniki, Greece.

I. Pitas is with the Department of Informatics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece. (Tel. +30 2310 996304, Fax +30 2310 996304 ; e-mail: pitas@aiia.csd.auth.gr).

I. Kostopoulos is with the Department of Pathology, Medical School, Aristotle University of Thessaloniki, 54124, Thessaloniki, Greece.

I. Introduction

The HER-2/neu (c-erbB2) oncogene is a tyrosine kinase receptor that is overexpressed in approximately 20-30% of high-grade invasive breast carcinomas and has been shown to be a valuable prognostic indicator [1]. HER-2 positive tumours can be more aggressive and their status can predict response to targeting therapy with trastuzumab (Herceptin) monoclonal antibodies and adjuvant chemotherapy. Knowing that a cancer is HER-2/neu positive helps a medical team select the appropriate treatment. Overexpression of the protein product of HER-2/neu gene is usually a consequence of gene amplification, in which multiple copies of the gene appear through the genome. It is thus possible to determine HER-2/neu status by analyzing the numbers of gene copies centrally or the amount of protein peripherically. Currently, the two most widely used technologies to determine HER-2/neu status are immunohistochemistry (IHC) and fluorescence *in situ* hybridization (FISH). IHC uses specific antibodies to stain proteins (products) *in situ,* which allows the identification of many cell types that could be visualized by classical microscopy. FISH imaging allows selective staining of various DNA sequences and thereby the detection, analysis and quantification of specific numerical and structural abnormalities within nuclei. The IHC test measures the protein coded by the HER2 gene, whereas FISH measures the number of copies of the HER-2 /neu gene present in the tumor cells. There are trade-offs in choosing one of these techniques. The IHC test reveals the protein on the cell by staining the product with specific antibody. The advantages of IHC include its wide availability, relatively low cost, easy preservation of stained slides, and use of familiar routine microscopes [2]. On the other hand, the antigenicity of cells can be affected by tissue formalin fixation so the stained slide can cause inaccurate interpretation of IHC results. In addition, IHC testing is subjective: the reader must judge the degree of colour change in the nucleus against a non-standardized chart. A recent study by Bartlett et. al. [3] considered the accuracy, reproducibility and availability of different techniques for the evaluation of Her-2/neu status and recommended screening by immunohistochemistry followed by FISH testing of cases with

intermediate staining intensity (cases scored 2+according to Hercep test). They suggested that the use of automated analysis may increase testing precision and predicted a wider future use of FISH analysis as a more cost-effective technique. A drawback of FISH is that it requires more specialized equipment than immunohistochemistry does and is, therefore, not as widely available yet. However, this technique is likely to have significant impact on diagnosis in the medical practice [4].

The process of evaluating HER-2/neu status from FISH images involves the manual counting of signals in interphase nuclei which become visible as colored dots. The FDA approved PathVision Her2 FISH kit (Vysis, Downers Grove, USA) uses DNA probes, which are small segments of actual DNA material. When applied to a tumor tissue sample, these DNA probes target the HER-2/neu gene and attach themselves to their target sequence. This process is called hybridization. The probes carry special fluorescent markers that emit light, when the probes bind to the HER-2 genes. The Her-2 probes are visible as orange stained spots under a fluorescent microscope. Similarly, probes for centromere 17 (CEP-17), the chromosome on which the gene HER-2/neu is located, are visible as green spots. The sections are counterstained with DAPI, providing a blue background for nucleus body. The conventional analysis involves the scoring of the ratio of HER-2/neu over CEP 17 dots within each cell nucleus and then averaging the scores for a number of ~60 cells. Several images usually need to be read to reach the desired number of dot-including nuclei. A ratio of >= 2.0 of HER-2/neu to CEP 17 copy number denotes amplification[1].

The reading of FISH images is a difficult task since manual dot scoring over a large number of nuclei and over different tissue samples is a time consuming and fatiguing technique. Moreover, it is user-dependent in the clinical setting lacking specific training for this imaging technique. Current analysis of FISH signals in practice is performed in a semi-automated way with the aid of image processing software, which can display the different color channels of a FISH image and apply thresholds for nuclei segmentation. A study by Klijanienko et al. [5] has shown strong correlation of detection results using visual-only and semi-automated methods for evaluating the status of Her-2/neu in breast carcinomas samples. However, the counting of dots in a semi-automatic manner still remains an impractical procedure for a pathologist, since it requires user intervention for excluding poorly segmented, overlapping, clustered or nonepithelial cells [5].

Several methods have been proposed for the automated evaluation of FISH signals, even though they were not applied directly for measuring Her-2/neu gene amplification of breast samples. Most methods focused on automatic spot counting whereas only very few focused on case-based classification of FISH images.

Netten et al. [6] focused on automatic counting of dots per cell nucleus in slides of lymphocytes from cultured blood. The method that they developed consisted of the following steps: a) selecting regions of interest (ROIs) containing at least one nucleus. This was done by global thresholding to separate a nucleus from the background. The ISODATA thresholding algorithm was used to distinguish between nucleus and background within each ROI, followed by morphological operators. Finally, a number of features describing the size, shape and intensity of the segmented nuclei were used to classify the segmented objects. For each segmented nucleus, the hybridized spots were detected using the top hat transform [7] and a nonlinear Laplacian filter. At the end, features describing the area, fluorescent intensity, average and relative intensity of the dots were measured and used to eliminate false positive signals. The method was evaluated on 6 slides using dot distribution, i.e. the percentage of nucleus containing 0,1,2,3, or >3 dots and percentage of correctly counted dots. In a different study [8] the same authors discussed their dot detection algorithm in more detail and evaluated different image processing algorithms.

Solorzano et al. [9] developed a method to study leukocytes in blood samples. For the imaging part, they segmented nuclei using the ISODATA thresholding algorithm. Then, the watershed algorithm incorporating the distance transform was used to isolate nuclei and FISH dots were detected using the top hat transform. A statistical error correction step was then used to improve the detection performance. The method was tested on 9 blood samples including 500 nuclei using the Kolmogorov-Smirnov maximum deviation test.

Kozubek et al. [10] developed a system that acquired 2-D and 3-D FISH images and performed image analysis on both. For the 2-D analysis, the system first segmented the nuclei using bimodal histogram thresholding. Morphological features, including nucleus size, presence of holes, nucleus roundness, and smoothness of boundaries, were used for further binary image processing and removal of false detections. Then, the system detected hybridized dots within each segmented nucleus; a watershed-based algorithm was used for dot segmentation, employing gradual thresholding. Dot features, including maximum intensity, height/size at relative position to the nuclear center were also extracted. 3-D analysis was performed by analyzing the pre-extracted nuclei and dot features for sequential 2-D slices. The authors presented applications of their system for detection of BCR/ABL translocation in interphase nuclei and for measuring distances between centromeres in HL-60 nucleus.

Lerner et al. [11-14] proposed a FISH image classification system based on the properties of in- and out-of-focus images captured at different focal planes. The signals were classified as real or artifacts and the images that contained no artifacts were considered to be the in-focus image. This methodology is in contrast with the methods described above that rely on auto-focusing mechanisms. The authors initially proposed the use of a neural network-based algorithm [11, 13]. This algorithm analyzed multispectral

[1] PathVision HER-2 DNA Probe Kit insert

FISH images in the RGB and HIS color spaces. Color segmentation using global thresholding was applied to each of the RGB channels, and a set of features was extracted for each resulting object. Features included shape (eccentricity), size (area), and spectral features such as maximum and average hue. A neural network (NN) was used to distinguish between real and artifact signals. In a later study [12], the authors employed a Bayesian classifier instead of an NN, to avoid dependency on a large number of parameters and NN architecture settings. A number of methodologies were examined for density estimation.

Recently, Chawla et al [15] developed an automated system for analyzing FISH signals from brain hippocampal and cortical sections. Their objective was to examine temporal gene transcription activity for which manual counting was time-consuming considering that a stack of images had to be examined. The system included a 3D watershed algorithm featuring a gradient-weighted distance transform, followed by a model-based region merging of nuclei. Unwanted glia and non-neuronal nucleus signals were removed with a clustering algorithm using intensity, texture and homogeneity features. The remaining FISH signals were classified as being present in the nucleus or in the cytoplasm, depending on location and morphometric parameters. Finally, FISH signals within the nuclei were measured.

Based on the above, there seems to be a potential for further development of systems for the automated case-based reading of FISH images, particularly for the application of HER-2/neu evaluation in breast carcinomas samples. Such a system should take into account multiple images of a specific case and quantify the HER-2/neu status in a collective manner. In this paper we present a multistage algorithm for the automated classification of FISH images from breast carcinomas samples. The algorithm focuses not only on the detection of FISH dots but on overall case classification. Moreover, the system is evaluated with receiver operating characteristic (ROC) analysis [16] in respect to three main tasks: nuclei segmentation, spot detection, and case classification.

As can be seen in Fig.1, the algorithm consists mainly of two stages for nuclei and dot detection respectively, in a similar fashion to the approach taken by Lerner et. al. [11, 13]. The dot segmentation is performed in the RGB color space and consists of a top-hat filter preprocessing stage followed by grey level template matching to separate real signals from noise. Nuclei segmentation is performed on the blue channel (DAPI) image. After a non-linearity correction step, global thresholding is used to identify candidate regions. A geometric rule is applied to distinguish between holes within a nucleus and holes between different nuclei. Finally, the marked watershed transform, defined in section III-B is used to segment cell nuclei. Combining the two stages allows the measurement of a FISH signal ratio per cell nucleus and consequently, the collective classification of cases, in a manner similar to the clinician's evaluation.

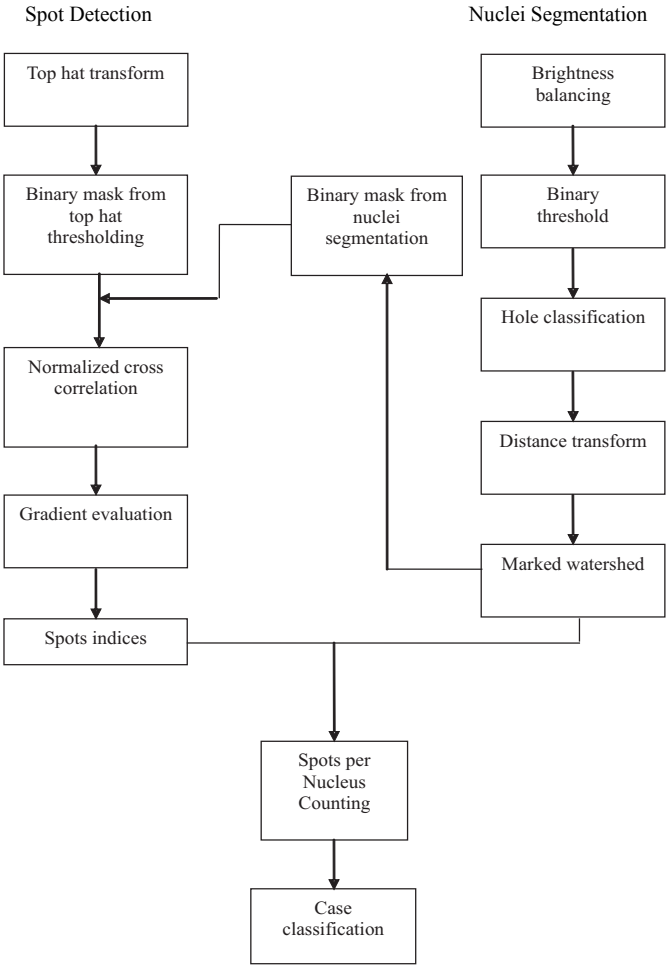


Fig.1: FISH case classification algorithm.

This paper is organized as follows: Section II describes the database used for the development and evaluation of the method. The method is detailed in section III whereas Section IV presents the evaluation of results and a related discussion. Finally, conclusions are stated in Section V.

II. MATERIALS:

Twelve cases, six of which were classified as positive and other six as negative, were available for the algorithm evaluation. A testing set of 40 FISH images was used for the spot detection evaluation. The true location of 887 red spots and 751 green spots was labeled by an expert. The same expert identified the location of 385 true red spots and 334 true green spots in 18 different FISH images that were employed as a training set. FISH images typically contained about 25 cells. The breast tissue slides were prepared using the following procedure.

Paraffin sections of 4μm thickness were incubated overnight at 60^{o}C. Deparaffinization, pretreatment, enzyme digestion and fixation of slides were performed using the Vysis Paraffin Pretreatment kit according to the manufacturer's recommended protocol. Slides were deparaffinized in xylene, dehydrated in 100% ethanol and

immersed in pretreatment solution. Proteolysis of neoplastic cells was performed by immersing the sections in protease solution at 37˚ C for 12 minutes. Tissue sections were denaturated at 85ºC for 2 minutes, then the PathVysion HER-2 DNA Probe (LSI HER-2/CEP17 probe, Abbott GmbH and Company, KG, Wiesbaden-Delkenheim, Germany) was added and hybridization took place at 37˚ C in a moist chamber for 14-18h (overnight incubation). The following day the slides were washed with post-hybridization buffer (2X SSC and 0,3% NP-40) at 72ºC for 2 minutes, followed by counterstaining of the nuclei with 4, 6-diamino-2phenylindole dihydrochloride (DAPI).

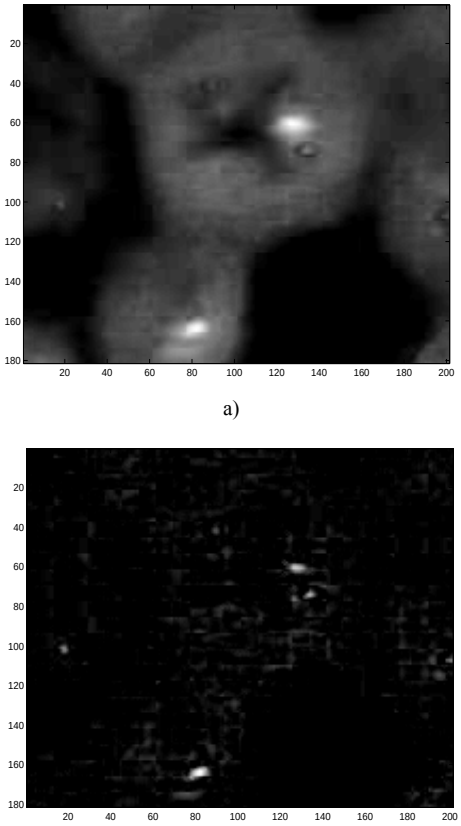
For each case, at least 60 non overlapping nuclei were scored for both Her-2/neu (red spot) and chromosome 17 (green spot) signals by image analysis. Hybridization signals were enumerated utilizing a Zeiss, Axioskop 2 microscope equipped with a 100 Watt mercury lamp (HBO 100) and an automatic filter wheel system with the following filters: BP360/51 DAPI filter, BP485/17 FITC filter-spectrum green, BP560/18 Rhodamine filter-spectrum orange. Plan-Neofluar lens with magnification of x100, NA=1.3 and a pixel size of 0.24 μm were used when reading the images, along with manual focusing. Images were grabbed using a CV-M300 2/3" CCD camera (JAI, Copenchagen, Denmark). The camera had a high S/N ratio of >58dB and an effective pixel resolution of 752(horizontal) x 582 (vertical). During clinical reading, the images were processed using the Meta Systems software (Altlussheim, Germany) in order to adjust contrast in the different color channels. This software contains a shading correction algorithm to account for non-uniform illumination. Her-2/neu gene amplification was determined by a ratio of Her-2/neu gene copies to chromosome 17 centromeres. According to the manufacturer's recommendations the cases with a ratio ≥2 were determined as amplified, while those having a ratio <2 as not amplified.

III. METHOD:

The algorithm for the classification of FISH images was based on the accurate measurement of red/green spot ratio (corresponding to the ratio of HER2/CEP 17) per cell nucleus. For that reason, two combined stages for spot detection and cell nuclei segmentation respectively were developed as described below.

A. FISH spot detection

Despite the fact that the main content of FISH image red and green channels is constituted of spots, many FISH images frequently contain noisy areas consisting of large stains. For this reason, a preprocessing step for noise removal is needed. An effective solution is provided with top-hat filtering, as was proposed in [7]. A disk of 4-pixel radius was chosen as the structuring element of the top-hat transform. The result of the top-hat filtering step is demonstrated using the example image of Fig. 2.



a)

b)

Fig.2: a) Original image red channel; b) Top-hat filtering output.

A typical grey level histogram of the top-hat output presents a strong unimodal trend consisting of a peak followed by a very steep monotonous decrease and a second flat part. A modification of the algorithm proposed in [17] was used to estimate two thresholds for the top-hat red and green channel output respectively. The algorithm assumes that there is one dominant mode in the image histogram. A straight line is drawn from the peak to the high intensity end of the histogram. More precisely, the line starts at the largest frequency bin A and finishes at the first empty bin B of the histogram following the last filled bin. The threshold is selected as the histogram index *Th* that maximises the perpendicular distance between the line and histogram curve. This procedure is illustrated in Fig.3. A different approach [18] would be to plot the derivative of the histogram and select the position of a maximum as a threshold.

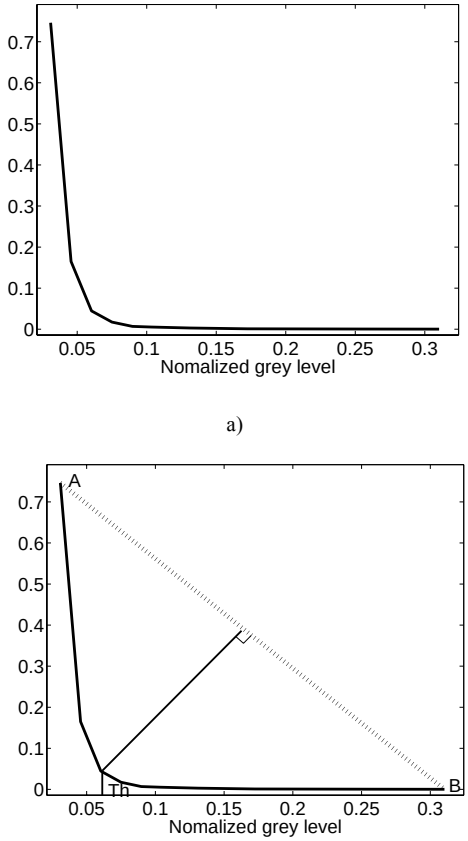


Fig.3: a) Typical grey level histogram of top-hat output; b) threshold selection.

For this application, we modified this algorithm by applying it not on the entire image histogram, but rather on the pixels belonging to the last k bins of the histogram. The value of k is estimated in the following way: if we indicate with N the total number of pixels of the image and with $h(i_n)$ the intensity image histogram value relative to the $n-th$ bin, then we consider the last k bins such that:

$$\sum_{n=N-k}^{N} h(i_n) \geq N \cdot p \qquad (1)$$

The value of p was chosen equal to 0.025. The histogram resulting from the selected pixels still presented a strong unimodal trend. The resulting threshold is proven to be rather insensitive to the value of p .

Even if red and green spots usually have the greatest channel intensity, it is likely that many valid spots have red/green level value smaller than that of false signals, e.g. the ones coming from accidental or non-specific staining. Therefore, even the best threshold choice is not enough to isolate all true spots from false ones using only the red or green channel intensity. For this reason, we employed the characteristic grey level trend of every spot to perform better spot detection.

Let the pixel position be represented by the (x, y) coordinates and the channel intensity be the third coordinate as shown in Fig.4.

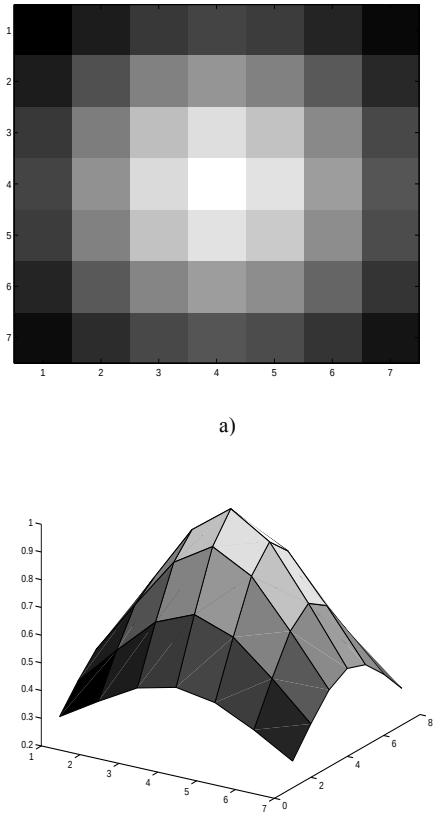


Fig.4: Spot shape: a) channel intensity image; b) perspective intensity plot.

The objective of this step is to compare the grey level "shape" of every candidate spot with the spot shape template obtained from the average shape of a set of valid labeled spots. To measure the similarity between every candidate spot and the spot template we use normalized cross correlation [19]. The estimation of the spot shape template is performed on training FISH images in the following way: considering the red and green channel independently, we estimate the center of every red and green manually labeled spot as the pixel with maximum channel intensity. A 7×7 window positioned on every spot center is saved as a template for spots in the red and green channel respectively. Two spot template windows T_G and T_R are estimated by averaging the respective spot channel intensities:

$$T_R(x, y) = \frac{1}{N_R} \sum_{i=1}^{N_R} f_{R_i}(x, y) \qquad (2)$$

$$T_G(x,y)=\frac{1}{N_G}\sum_{i=1}^{N_G} f_{G_i}(x,y) \qquad (3)$$

where N_R and N_G are the number of used red and green spots, $x=1,...,7$ and $y=1,...,7$ are coordinates in a 7×7 window and f_{R_i} and f_{G_i} are red and green channel intensities of $i-th$ spot image. For each new test spot image, the normalized cross correlation C_R between T_R and the respective channel intensity I_R, is calculated as follows:

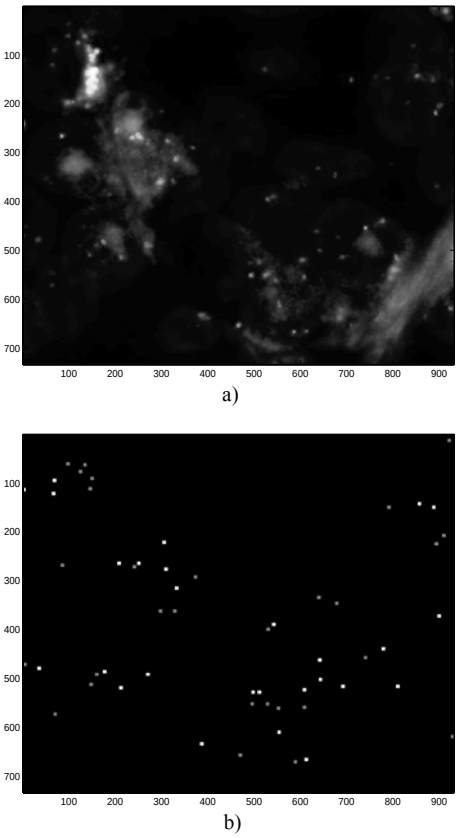
$$C_R(u,v)=\frac{\sum_{x,y}\left[T_R(x,y)-\bar{T}_R\right]\left[I_R(x-u,y-v)-\bar{I}_R(u,v)\right]}{\left\{\sum_{x,y}\left[T_R(x,y)-\bar{T}_R\right]^2\sum_{x,y}\left[I_R(x-u,y-v)-\bar{I}_R(u,v)\right]^2\right\}^{0.5}} \qquad (4)$$

where $\bar{T}_{R_{u,v}}$ is the mean value of T_R, while $\bar{I}_R(u,v)$ is mean value of red channel I_R around the 7-pixel neighborhood of pixel (u,v). The normalized cross correlation $C_G(u,v)$ for the green channel is computed in a similar fashion. As we can see in Fig.1, C_R and C_G were evaluated only for FISH image positions where a) the top-hat output is above threshold and b) a nucleus is present, as found by the nuclei segmentation procedure described below. In order to select red/green spot positions, two positive thresholds Th_R and Th_G are used; spots with a value of C_R and C_G lower then Th_R and Th_G respectively, are discarded while the remaining ones are used as input for the next selection step of the FISH spot detection algorithm.

Finally**,** for every detected spot from the previous step, a channel intensity contrast measure is used. This further processing step is performed to discard spots whose shape is very similar to the template one, but have a low channel intensity contrast with respect to their surrounding pixels, making them appear invisible to the human eye. The contrast measure is performed using the information of the red, green and blue channel. For each spot, two vectors $\boldsymbol{v}_{for}$ and $\boldsymbol{v}_{back}$ are created. Each of the three components of vector $\boldsymbol{v}_{for}$ is estimated considering the average channel intensity of the pixels of a 5×5 window positioned on every spot center, while each of the corresponding three components of vector $\boldsymbol{v}_{back}$ is estimated considering the average channel intensity of the background pixels around binary object perimeter. Then the contrast measure C_M is calculated as:

$$C_M=\frac{\left\|\mathbf{v}_{\mathbf{for}}-v_{back}\right\|}{\left\|v_{back}\right\|}>T_{C_M} \qquad (5)$$

The thresholds Th_R, Th_G and T_{C_M} are empirically chosen in order to minimize spot classification error over the FISH images used for training. An example of original FISH image and the relative output of the spot detection algorithm are shown in Figs.5 a, b.



a)

b)

Fig.5: a) Original FISH image; b) Output of the spot detection algorithm.

B. Cell nuclei segmentation

Cell nuclei segmentation is performed on the FISH image blue channel. For many images, cell nuclei contain an inhomogeneous blue channel intensity. In order to reduce the gray level difference between dark regions and more illuminated ones, a nonlinearity correction step was performed applying the square root function to the blue channel, normalized by its maximum grey level p_{MAX}, as shown in Eq. 6.

$$p_{out}=\sqrt{\frac{p_{in}}{p_{MAX}}} \qquad (6)$$

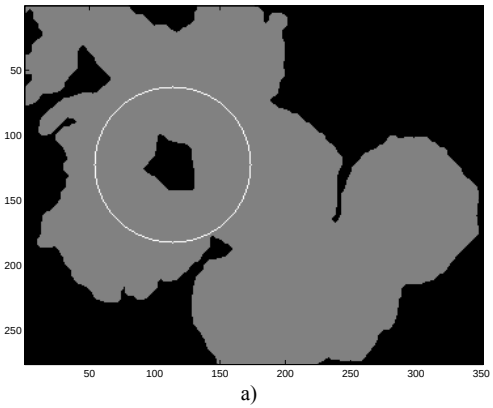
Where p_{out} and p_{in} are the pixel grey levels.

Moreover, gray level peaks due to the presence of spots are made less intense by applying the opening morphological

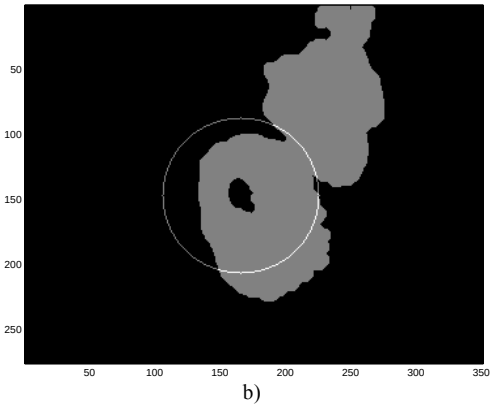
operator to the FISH image blue (DAPI) channel using a disk of 4-pixel radius as structuring element. A further top-hat filtering using as structuring element an 80-pixel radius disk is also performed to reduce the blue channel intensity of regions where non-ideal staining caused color diffusion.

Based on the fact that the histogram of the resulting image has a characteristic bimodal shape, we employ the algorithm by Otsu et al [20] to determine the threshold for initial nuclei segmentation. The binary image resulting from thresholding sometimes contains holes even in a single nucleus body region. This kind of holes has to be filled to enable correct nuclei segmentation. On the contrary, holes present in inter nuclei zones of overlapping nuclei should not be filled. The two types of holes are illustrated in Figs 6a and b respectively. A method to distinguish between these two types of holes is described below.

As the two types of holes can not be separated into two classes using just gray level or morphological features, a geometric approach is employed. Let P be the percentage of the perimeter pixels of a circle of radius R centered on a hole centroid that is contained in the nucleus region. It can be observed from Fig. 6, that the value of P is much higher for the first type of hole (inter nuclei) than the second type of hole (nucleus region). The radius R used is slightly bigger than the average nuclei radius.



a)

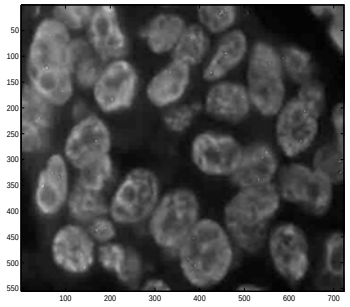


b)

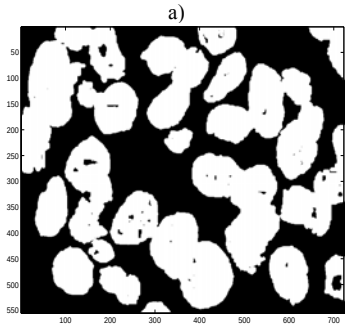
Fig.6: a) Hole due to overlapping nuclei (inter nuclei case); b) Nucleus body hole.

It was found experimentally that the value for P varied in the range of 90% to 40% for holes of the first type and second type respectively.

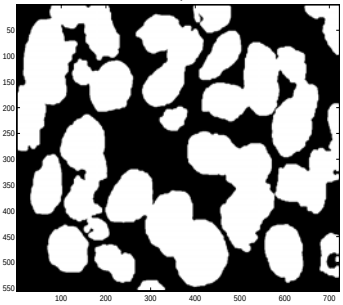
Figs 7 a, b and c show respectively the original blue channel of a FISH image, the binary image after thresholding and the binary image after the hole classification step.



a)



b)



c)

Fig.7: a) blue channel of original image; b) result after thresholding; c) result after hole classification.

The last step of the nuclei segmentation algorithm involves the marked watershed transform [21] which is employed to detect borders in overlapping nuclei clusters. The distance transform [22] is first applied to the binary image obtained from the previous step. For every pixel, this transform produces an intensity value proportional to its Euclidean distance from the closest background pixel. The uneven shapes of binary object borders cause spurious local maxima of the distance transform output that do not correspond to nuclei centers. If all these local maxima were used as markers, an over segmentation would be obtained. In order to reduce this effect we calculate h-dome maxima of the resulting image [23]. H-dome maxima are connected components of pixels in an 8-connected neighborhood with the same intensity value and whose external boundary pixels all have a value less than h. Characteristic values of h are in the range $[0.5 : 2]$. This feature is employed in order to discard insignificant local

maxima present in the distance transformed image. Fig.8 shows examples of the original blue channel image and the final output of the nuclei segmentation step. We note that the nuclei touching the image border are removed because they are not considered in the spots per cell counting.

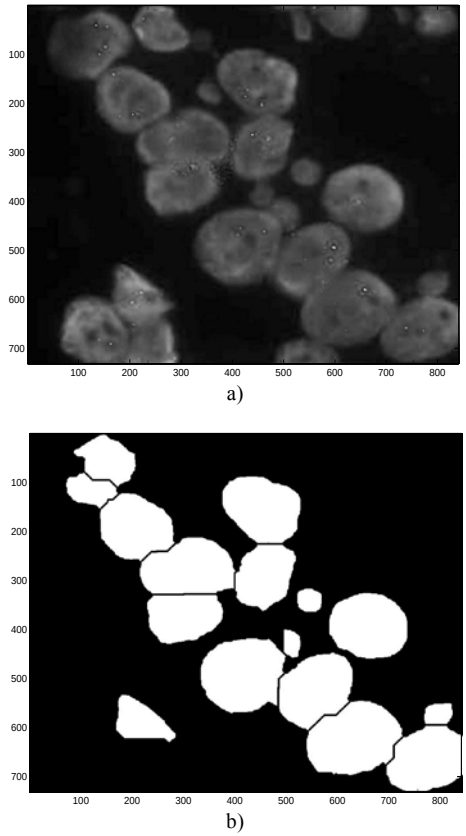


Fig.8: a) Original blue channel image; b) Final output of segmentation algorithm.

IV. RESULTS

The algorithm was evaluated with respect to three different tasks: spot detection, cell nuclei segmentation and case-based classification as will be described below. Receiver Operating Characteristic (ROC) curves were used to describe the performance of each task, as will be shown below.

A. *Spot detection:*

In order to estimate the performance of the algorithm for detecting FISH spots, a testing set of 40 FISH images was used. The true location of 887 red spots and 751 green spots was labeled by an expert. The same expert identified the location of 385 true red spots and 334 true green spots in 18 different FISH images that were employed as training set to estimate spot shape templates. The false positive rate was defined as the ratio between the total number of detected spots not present in the ground truth over the total number of detected spots. The true positive rate was defined as the ratio between the total number of correctly detected spots over the total number of spots present in the ground truth. ROC curves were constructed by collecting pairs of sensitivity (or true positive rate) and false positive rate for different thresholds. Twenty one points of the curve were estimated by varying the threshold applied to top-hat output between 0 and 1 with a step of 0.05 while fixing the thresholds of normalized cross correlation, Th_R and Th_G, and threshold T_{CM} of gradient intensity value to 0.7, 0.65 and 0.3 respectively. Curves were estimated for red and green spots separately using the testing set described above and are displayed in Fig. 9.

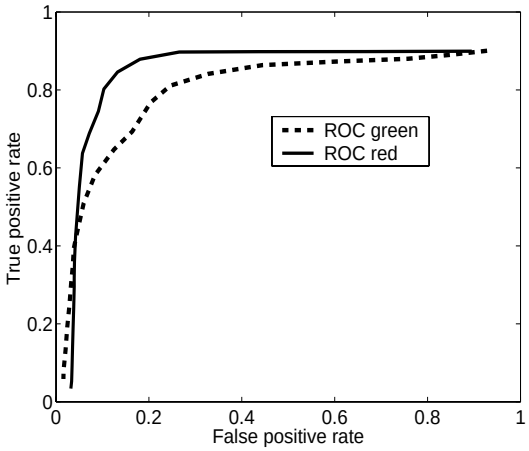


Fig.9: ROC curves relative to red and green spots detection.

It can be seen from Fig.9 that the performance of the algorithm is noticeably better for the detection of red spots. This can be due to the fact that red spots have a more well-defined shape compared to the green ones. As a demonstration point, the algorithm can reach a sensitivity of about 92% and 80% for red and green spots respectively at a false positive rate of about 25%.

B. *Cell nuclei segmentation:*

In order to evaluate the performance of the algorithm for cell nuclei segmentation, the ratio between the area of intersection of segmented nuclei with true nucleus region over the area of the union of the two regions was calculated. The ground truth for the correct nucleus boundaries was determined manually so that 1439 nuclei were labeled. ROC curves were constructed by varying a threshold for the ratio of intersection over union as described above. Fig.10 shows the resulting ROC curves for three values of h {0.5, 1.0, 1.5}.

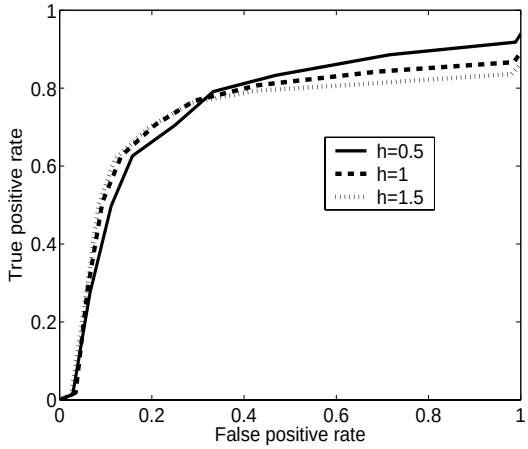


Fig.10: ROC curves relative to the cell nuclei segmentation.

As we can see performances are not much sensitive to the value of *h*, even if a value of h = 1 shows better results in the high sensitivity part of the ROC curve. It has to be noted here that for this application it is more crucial to not discard a true nucleus than to avoid merging overlapping nuclei since dot counting per nucleus is averaged and it is not overly sensitive to two nuclei being counted as one. On the other hand, correct nuclei segmentation definitely reduces the effect of background noise pixels that can be mistaken for FISH dots and affect the dot ratio calculations.

C. *FISH case-based classification:*

Twelve patient cases, six of which were previously classified by an expert as positive and six that were classified as negative, were employed to evaluate the precision of the algorithm in classifying the different cases.

For every segmented nucleus where at least one red spot was present the ratio d was calculated, defined as:

$$d = \frac{N_R}{N_G} \qquad (7)$$

where N_R and N_G are, respectively, the number of red and green spots present in the segmented nucleus. In the instances where the number of green spots was zero, but the number of red spots was nonzero, it was assumed that at least one green spot was present (probably in the same location as a red spot), and consequently, the value of N_G was set to one. In our experiment, considering the total number of segmented nuclei, we had such instances in 16% of all cells. In Figs.11 and 12, the histograms of the calculated red/green spot ratios per nucleus for the six positive and the six negative cases respectively were displayed. As can be seen from the plots, the histograms of the negative cases are more concentrated in the bin corresponding to a ratio red/green equal to one than the corresponding histograms for the positive cases. In order to quantitatively analyze the histogram properties, the probability measure $P(d \geq 2)$ was defined as the probability of having a ratio red/green greater or equal then two. The number of images available and the values of $P(d \geq 2)$ for each of the testing cases are listed in Table 1.

Table 1. Number of available images for each case and values for the estimation of the probability to have a ratio greater or equal to 2 for each of the testing cases.

	# of Images	$P(d \geq 2)$		# of Images	$P(d \geq 2)$
Positive case #1	5	0.5316	Negative case #1	5	0.2398
Positive case #2	5	0.5316	Negative case #2	5	0.2398
Positive case #3	5	0.5316	Negative case #3	5	0.2398
Positive case #4	5	0.5316	Negative case #4	5	0.2398
Positive case #5	5	0.5316	Negative case #5	5	0.2398
Positive case #6	5	0.5316	Negative case #6	5	0.2398

It can be seen from Table 1 that all of our cases can be correctly classified as either positive or negative (100% sensitivity with 0 false positives) by using a proper threshold for the statistic $P(d \geq 2)$. Despite the small number of cases, these preliminary results are encouraging for the further development of a fully automated method to accurately distinguish between normal and abnormal breast tissue samples. A larger database of FISH images of breast tissue is being prepared in order to examine how well these results can generalize in a broader population. Moreover, it will be useful to examine the robustness of our algorithms to changes in the image acquisition step, such as changes in lens magnification and camera resolution and develop automated parameter optimization techniques to account for these changes.

V. CONCLUSIONS

We have developed a method for the automated evaluation of Her-2/status in breast samples by FISH image analysis. Conventional evaluation is a time-consuming task since it involves manual counting of dots in multiple images and, additionally, it is prone to inter-observer variability.

The developed method uses two combined multistage algorithms. The first one, used for the detection of the red and the green spots, includes mainly stages of top-hat filtering, binary thresholding, grey level template matching and contrast evaluation. The second one, used for the cell nuclei segmentation, consists of a non-linear blue channel correction step, a global thresholding by Otsu algorithm, a grey level hole classification by a geometric rule and of the marked watershed transform using local h-dome maxima as markers. The outputs of the two algorithms were merged for estimating the average red/green ratio per cell nucleus, which is the feature used in clinical practice to determine gene amplification.

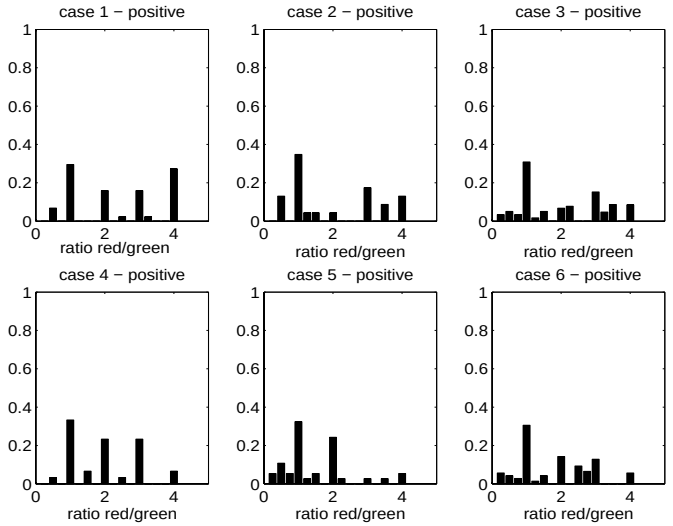


Fig.11: Histograms of the six positive cases.

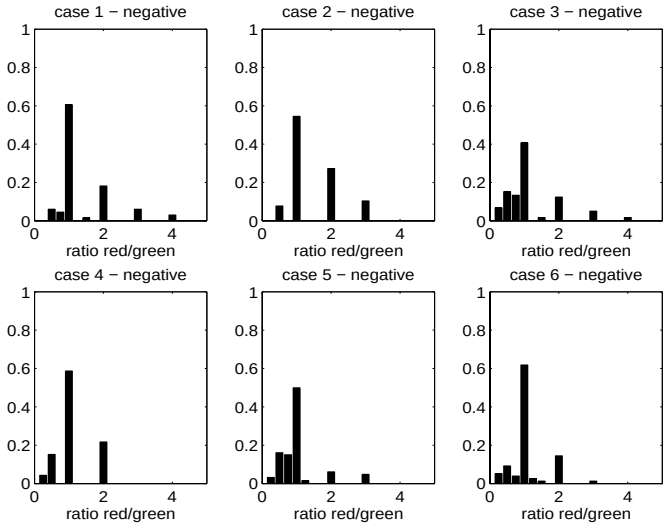


Fig.12: Histograms of the six negative cases.

The performance of the proposed method was evaluated using ROC curves both for the detection of the red and the green spots and for the cell nuclei segmentation on 58 testing FISH images from 12 cases. Moreover, the overall algorithm performance for case-based classification on the 12 testing FISH cases showed the ability of the proposed system to distinguish between all positive and negative cases. Despite the small number of cases, the evaluation results were encouraging for the further testing of the method in clinical trials.

ACKNOWLEDGMENT

This work was supported by the EU project Biopattern: Computational Intelligence for biopattern analysis in Support of eHealthcare, Network of Excellence Project No. 508803.

REFERENCES

[1] P. P. Osin and S. R. Lakhani, “The pathology of familial breast cancer. Immunohistochemistry and Molecular analysis”, *Breast Cancer Research*, vol. 1(1), pp. 36-40, 1999.

[2] J. S. Ross, J. A. Flethcer, G. P. Linette, J. Stec, et. Al, “The Her-2/neu gene and protein in Breast Cancer 2003: Biomarker and target of Therapy”, *The Oncologist*, vol. 8(4), pp. 307-325, 2003.

[3] J. Bartlett, E. Mallon, and T. Cooke, “The clinical evaluation of Her-2 status: which test to use?”, *Journal of Pathology*, vol. 199(4), pp. 411-417, 2003.

[4] J. M. Levsky and R. H. Singer, “Fluorescence in situ hybridization: past, present and future”, *Journal of Nucleus Science*, vol. 116(14), pp. 2833-2838, 2003.

[5] J. Klijanienko, J. Couturier, M. Galut, A.K. El-Naggar, Z. Maciorowski, E. Padoy, V. Mosseri, and P. Vieth, “Detection and Quantitation by Fluorescence in situ hybridization (FISH) and Image Analysis of HER-2/neu gene amplification in Breast Cancer Fine-needle samples”, *Cancer Cytopathology*, vol. 87(5), pp. 312-318, 1999.

[6] H. Netten, I. T. Young, L. J. van Vliet, H. J. Tanke, H. Vroljik, and W. C.R. Sloos, "FISH and Chips: Automation of fluorescent dot counting in interphase cell nuclei". *Cytometry*, vol. 28(1), pp. 1-10, 1997.

[7] F. Meyer, “Iterative Image Transformations for an Automatic Screening of Cervical Cancer”, *Journal of Histochemistry and Cytochemistry*, vol. 27(1), pp. 128-135, 1979.

[8] H. Netten, L. J. van Vliety, H. Vrolijkz, et. al. “Fluorescent dot counting in interphase cell nuclei”, BioImaging, vol. 4, pp. 93-106, 1996.

[9] C. O. de Solorzano, A. Santos, I. Vallcorba, J-M Garcia-Sagredo, and F. del Pozo, "Automated FISH spot counting in interphase nuclei: Statistical validation and data correction", *Cytometry*, vol. 31(2), pp. 93-99, 1998.

[10] M. Kozubek, S. Kozubek, E. Lukasova, A. Mareckova, et. al, "High-resolution cytometry of FISH dots in interphase nucleus nuclei", *Cytometry*, vol. 36(4), pp. 279-293, 1999.

[11] B. Lerner, W. F. Clocksin, S. Dhanjal, M. A. Hulten, C. M. Bishop, "Feature representation and signal classification in fluorescence in-situ hybridization image analysis", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 31(6), pp. 655-665, November 2001.

[12] B. Lerner, "Bayesian fluorescence in situ hybridisation signal classification", *Artificial Intelligence in Medicine*, vol. 30(3): pp. 301-316, 2004.

[13] B. Lerner, W.F. Clocksin, S. Dhanjal, M.A. Hulten, C.M. Bishop, "Automatic signal classification in fluorescence in situ hybridization images". *Cytometry,* vol. 43(2), pp. 87-93, Feb 2001.

[14] W.F. Clocksin and B. Lerner, "Automatic analysis of fluorescence in-situ hybridization images". *Proceedings of the 11th British Machine Vision Conference*, pp. 666-674, Bristol, September, 2000.

[15] M. K. Chawla, G. Lin, K. Olson, A. Vazdarjanova, S. N Burke, et. al., “3D-catFISH: a system for automated quantitative three-dimensional compartmental analysis of temporal gene transcription activity imaged by fluorescence in situ hybridization”, *Journal of Neuroscience Methods*, vol. 139, pp. 13-24, 2004.

[16] J. Swets and R. Pickett, *Evaluation of diagnostic systems: Methods from signal detection theory*. New York, USA: Academic Press, 1982.

[17] P.L. Rosin, "Unimodal thresholding", *Pattern Recognition*, vol. 34(11), pp. 2083-2096, 2001.

[18] GW Zack, JA Spriet, SA Latt, GH Granlund and IT Young, “Automatic detection and localization of sister chromatid exchanges”, J Histochem Cytochem., vol. 24(1), pp. 168-177, 1976.

[19] R.C. Gonzales and R.E. Woods, *Digital Image Processing*, Addison Wesley, 2nd edition, 2003.

[20] N. Otsu, “A thresholding selection method from graylevel histogram”, *IEEE Trans. on Systems, Man and Cybernetics*, vol. 9 (1), pp. 62–66, 1979.

[21] S. Beucher and F. Meyer, “The morphological approach to segmentation: the watershed transformation”, in E.R. Doughterty (Ed.), *Mathematical morphology in image processing*, New-York: Dekker, 1992.

[22] W. K. Pratt, *Digital Image Processing*, New York, John Wiley & Sons, Inc., 1991.

[23] L.Vincent, ”Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms”, *IEEE Trans. Image Processing*, vol. 2(2), pp. 176-201, 1993.