Automatic Segmentation Framework for Fluorescence in Situ Hybridization Cancer Diagnosis

Marcin Stachowiak, Łukasz Jeleń

Automatic Segmentation Framework for Fluorescence in Situ Hybridization Cancer Diagnosis

Marcin Stachowiak and Łukasz Jeleń

Department of Compuer Engineering,
Wrocław University of Science and Technology,
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland.
e-mail: lukasz.jelen@pwr.edu.pl,

Abstract. In this paper we address a problem of HER2 and CEN-17 reactions detection in fluorescence in situ hybridization images. These images are very often used in situation where typical biopsy examination is not able to provide enough information to decide on the type of treatment the patient should undergo. Here the main focus is placed on the automatization of the procedure. Using an unsupervised neural network and principal component analysis, we present a segmentation framework that is able to keep the high segmentation accuracy. For comparison purposes we test the neural network approach against an automatic threshold method.

Keywords: FISH, pattern recognition, image processing, computer aided diagnosis, breast cancer, nuclei segmentation, HER2, dot counting, SOM, PCA

1 Introduction

According to the data provided by the National Cancer Registry, breast cancer is one of the most often diagnosed cancers among middle–age women [1]. Just in Poland, before 2016, there were 17144 diagnosed cases of breast cancer. This number is increasing year after year. For instance, between 2009 to 2012 there was an increase of 1280 diagnosed cases. The same records show that out of 17144 cancer cases there were 5651 deaths in 2012 which is 341 more than in 2009. Most of these cases could have been fully recovered if the diagnosis would be made in the early stage of the disease, because cancers in their early stages are vulnerable to treatment.

To reduce the number of deaths it is crucial to perform a reliable and fast diagnosis that will allow for the determination of an appropriate treatment. For this purpose when a suspicious growth is found during the screening mammography tests a fine needle aspiration biopsy (FNA) or a core biopsy (CB) is taken. During these examinations a small sample from the questionable breast tissue is extracted and a prognostic factor is evaluated according to the so called Bloom–Richardson scheme [3]. This procedure, called a malignancy grading, allows the

pathologist to describe the type of cancer in detail and estimate its behavior with or without undertaking treatment. Sometimes, when a difficult case is under diagnosis, the above techniques might require additional tests. This is why, for a more accurate diagnosis, a set of different examinations are performed. They will test for the presence of a HER2 gene and HER2 receptors that stimulate the growth of cancer cells. HER2 expression plays an important role in breast cancer diagnosis and the appropriate treatment is chosen accordingly to its status [11]. To determine the status of breast cancer biomarker such as human epidermal growth factor receptor 2 (HER2) a routine ImmunoHistoChemistry (IHC) or Fluorescence in situ hybridization (FISH) tests are performed.

- **ImmunoHistoChemistry** – is a staining process that shows if HER2 receptors and hormone receptors are present on the surface of the cancer cells (see Fig. 1a). This test helps in identification of the antigens in cells. This is possible due to binding of antibodies to the proteins. The final diagnosis is based on the estimation of different markers that may appear within and around the tumor cells [31].
- **Fluorescence in situ hybridization** – is a test that allows for a visualization of genes, in this case HER2 gene [15] (see Fig. 1b). In breast cancer diagnosis it is used to determine if the cancer cells have additional copies of that gene. The rule here is that the more genes one can distinguish, the more HER2 receptors the cells have.
  According to the American Society of Clinical Oncologists (ASCO) and College of American Pathologists (CAP) [13, 25], also known as ASCO/CAP, the complete FISH examination also requires estimation of chromosome 17 centromere enumeration probe (CEP17 or CEN-17) [27]. The final diagnosis is based on the HER2 to CEN-17 ratio [8].

According to Hicks and Kulkarni [10], both of the above tests are equivalent for the evaluation of the breast cancer HER2 status. In this study we are focusing

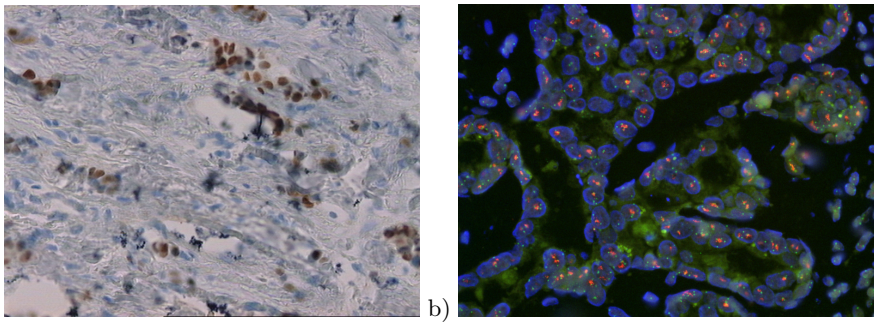


Fig. 1. HER2 slides. a) ImmunoHistoChemistry staining image, b) FISH test image.

on the FISH examination. As reported in literature [2, 23, 26], the problem of

distinguishing the HER2 and CEN-17 reaction within the FISH slides is not an easy task. The main problem here is a segmentation of the reactions, where the HER2 reactions are visible in the image as red dots while the CEN-17 are detectable as green dots. Due to a small size of the regions of interest there is a need to find the best possible segmentation algorithm that will be able to localize both kinds of dots within the slide. Here, we took an opportunity to test two techniques that will be appropriate for this challenging task.
Automatic cancer nuclei detection and segmetnation from medical images has been widely studied in the literature [4, 21]. One can find various reports on applying different imaging techniques [6, 22, 28], segmentation methods [17] or classification approaches [7] to solve this problem. In this study we are concerned on the segmentation of the FISH image to help doctors in localization of the HER2 and CEN-17 reactions. For this purpose, a thresholding based segmentation was investigated and its results were compared with segmentation based on the unsupervised neural network. Application of neural networks for segmentation was widely studied in literature [5, 20]. Substantial portion of the reports deals with segmentations based on self–organized maps (SOMs). Yao *et al.* [30] have successfully applied SOMs to segment sonar images where each pixel of the input image is classified with the proposed neural network. Gorjizadeh *et al.* [9] used a similar idea for segmentation of noisy medical images.
In 2001, Lerner *et al.* [2] described a neural network approach for detection of fluorescence in situ hybridization images. Their method is based on classification between a pair of in– and out–of–focus images. They use the in–focus images for further estimation of the FISH reactions. Results presented by the authors show an accuracy of 83–87%. Another approach, reported by Kiszler *et al.* [15], describes a semi–automated procedure applied to fluorescence in situ hybridization images. This procedure is based on the adaptive thresholding and the final counting is based on the selected areas of the image.

It can be easily noticed that Machine Learning is a popular tool for developing support software for supporting a diagnosis process of specialists. This is why as a main contribution of this paper we propose a fully automatic procedure for segmentation of HER2 and CEN-17 reactions that can be further counted to estimate the final FISH diagnosis. The final decision is based on the ASCO/CAP recommendations [13] that give a full route for HER2 testing in breast cancer.

2 Dataset

For the purpose of this study we have collected a database of 80 fluorescence in situ hybridization images with a size of $1376x1032$ pixels recorded with a resolution of 200 pixels per inch. Example of the images in the database is shown in Fig. 2. Based on these examples it is easy to notice how difficult the automated segmentation of such images is.
Images were recorder with an Olympus BX61 fluorescence microscope with X-Cite series 120Q EXFO fluorescent system. The microscope was equipped with a CCD Olympus XC10 camera working with a Cell-F visualization software. To

capture the image it was required to use fluorescent filter such as 30-151332G-Ov2C146747 filter made by Abbott with a magnification of 60x and 100x.
The database is a courtesy of dr. Anna Lis–Nawara and prof. Michał Jeleń, the head of the Department of Pathology and Oncological Cytology at the Wroclaw Medical University, Wrocław, Poland.

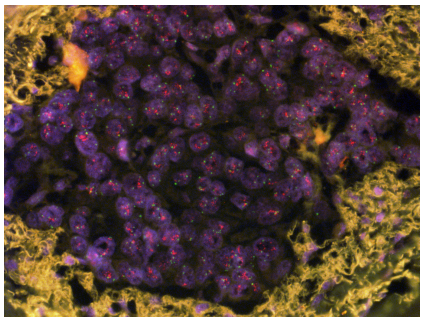
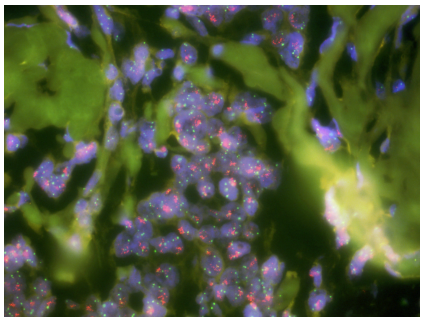
a)  b) 

Fig. 2. Example of images in the database.

3 HER2 and Cen-17 Segmentation

In computer vision, segmentation is a very crucial step that influences all subsequent phases of the classification systems. In medical image processing, the segmentation stage is always a very difficult task and that makes it a very active field of research [12, 18, 19]. In this research, we investigated two segmentation methods that will allow for the proposition of the automated segmentation framework for fluorescence in situ hybridization images. The first method is a simple thresholding were the threshold is automatically selected. The second method is an unsupervised neural network method where we made use of the self–organizing maps for segmentation of the receptors.

3.1 Simple Thresholding

To provide automatization of the segmentation procedure we decided to investigate one of the simplest automatic thresholding techniques proposed by Ridler and Calvard [24]. This method is based on a bimodal histogram of image gray levels. A threshold T is sought on the histogram curve according to equation 1.

$$T = \frac{\mu_1 + \mu_2}{2}, \tag{1}$$

where μ_1 and μ_2 are the means of the components separated by T. These means are calculated iteratively starting form the initial threshold that is typically set to an average gray level of the image.
As mentioned before, HER2 reactions are visible as red dots and CEN-17 reactions are detectable as green dots. If we convert an RGB image to grayscale we would loose that information and therefore we need to apply the above mentioned segmentation to a color image. This means dividing the image into three separate planes (R, G and B). Each plane will represent the main color in the RGB color space. Looking at the nature of the images we can notice that the HER2 reactions are emphasized in a red channel and the CEN-17 reactions are affirmed in the green plane (see Fig. 3). We can take that information into consideration

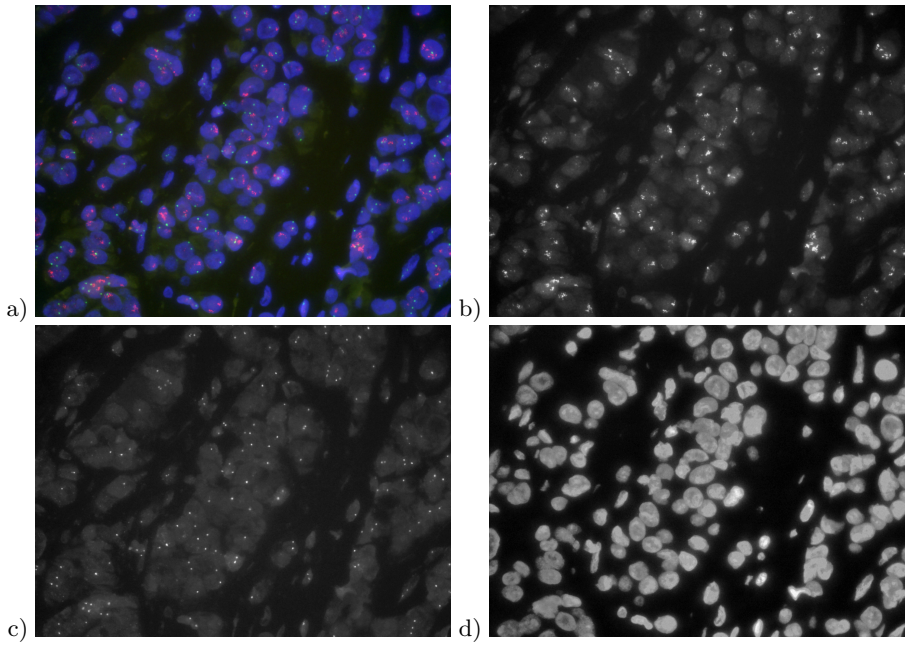


Fig. 3. FISH image RGB channels. a) Original image, b)Red channel, c)Green channel, d)Blue channel.

and apply Ridler and Calvard method to red and green image channels. Results of the application of this simple automatic thresholding method are presented in section 4.1.

3.2 Segmentation with Self Organizing Maps

As mentioned in section 1 self–organized maps (SOMs) can be successfully applied for a segmentation task. They reduce the input space into representative features according to a self–organizing process and are trained in an unsupervised manner [16]. These networks consist of only one layer with a linear transfer function for its neurons. It uses a comprehensive learning algorithm for weights estimation that updates the weight of only one, winning neuron for each input pattern. According to Kohonen [16] the introduction of an additional weight change of the neighboring neurons with smaller step size results in better correspondence to the features of the input data.
To train the SOM network we start with initialization of weights (w) with small random values and for each input data a winning neuron ($\overrightarrow{i}(x)$) is found according to equation 2. Neighboring weights ($w_j(n+1)$) are then calculated according to equation 3.

$$\overrightarrow{i}(x) = \arg\min_j \|\overrightarrow{x}(n) - w_j\|, \tag{2}$$

where $\overrightarrow{x}(n)$ is an input vector.

$$w_j(n+1) = w_j(n) + \eta(n)[x(n) - w_j(n)] \tag{3}$$

where $\eta(n)$ is a neighboring function.
According to Kohonen, the neighborhood taken into consideration should be Gaussian and he suggests the neighborhood description according to equation 4.

$$\Lambda_{j,j^o}(n) = \exp\left(-\frac{|r_j - r_{j^o}|^2}{2\sigma^2(n)}\right) \tag{4}$$

where j^o is the winning neuron and $|r_j - r_{j^o}|$ is a distance between the winning node and the $j-th$ node, and σ^2 is a Gaussian variance.
From this we can notice that this is an adaptive procedure, because the neighborhood and learning rate depend on the current iteration. Due to this fact our neighborhood should be as large as the output space at the start and should be decreasing during the iterations according to the equation 5. We can take the same reasoning for the step size, which should be big at the beginning and progressively decrease according to equation 6 until it reaches zero.

$$\sigma(n) = \frac{1}{c_\sigma + d_\sigma n} \tag{5}$$

$$\eta(n) = \frac{1}{a_\eta + b_\eta n} \tag{6}$$

where a_η, b_η, c_σ and d_σ are constants.
Here we have adopted the SOM methodology to solve a segmentation problem for detection of HER2 and CEN-17 reactions in fluorescence in situ hybridization images. This task is divided into several stages. At first, we extract the red and green components from the RGB image, because these two channels provided

the best localization of the reactions. Then we have applied a morphological reconstruction on both channels to extract the candidates of the red and green dots that represent HER2 and CEN-17 reactions respectively. These are actually the image maxima. Determination of these maxima can provide many misleading results such as multiple one–pixel dots. Some of these would also include background noise and for that reason there is a need to flatten them with reconstruction.
This method is actually a morphological erosion and dilation followed by the reconstruction procedure according to the description of Luc Vincent [29].
According to Vincent, for an image I, called a mask, a reconstructed image (I_r) for a given marker M is calculated as a union of the connected components if the image which contain at least one mask pixel (see eq. 7).

$$I_r = \bigcup_{M \cap I_k \neq 0} I_k \tag{7}$$

To localize candidates for HER2 and CEN-17 indicators we first apply erosion and dilation to appropriate image channels and then we apply a reconstruction algorithm described above. A dilated image was used as a marker and eroded image was treated as a mask (see Fig. 5). As a result a flatten maxima area is obtained.
Knowing the exact location of each of the maxima we have to check if the area it represents is actually red. For this purpose, an color image is divided into areas equivalent to a bounding boxes of the image containing the maxima. All of these areas serve as a feature vector presented as an input to the neural network. If we take an area of 5x5px the obtained feature vector will contain 75 input weights. Such a large number of features suggests a high correlation between the features and therefore a dimension reduction is justified.
To remove redundancy and to reduce a feature vector length, all data was analyzed with principal component analysis (PCA) [14] that used the singular value decomposition to determine the coefficients. Application of the PCA allowed for the reduction of the feature vector down to 3 independent components that characterize the variance changes in 97%.
Such a feature vector is then presented as an input to the neural network which finds the areas concentrated around these tree groups. Analysis of obtained clusters determines the appropriate dots. Results of SOM segmentation are presented in section 4.1.

4 Experimental investigations

As mentioned in the previous section, to propose an automatic segmentation framework we need to concentrate on counting the visible HER2 and CEN-17 reactions. We propose to convert a color image into three components (red, green, blue) and search for the maxima on the red and green images. By applying the simple thresholding method to these channels we were able to obtain most of the reactions with a threshold value automatically selected for each image.

Unfortunately, this might lead to some problems in situations where some colors can contain higher values of red component that the red color itself. For that reason we split the procedure into two phases. In the first phase the algorithm looks for a red image maxima which serve as red dot indicators. In the next step we classify these indicators with a SOM network to determine which of them represent red dots. We can apply the same procedure for localizing green dots. When all reactions are detected it possible to calculate the a so called FISH coefficient that will be used for further diagnosis. Such a calculation is based on the ASCO/CAP [13] recommendation where a HER2 to CEN-17 ratio is calculated. Using this value we can further decide if the diagnosed case is FISH positive or negative. The main aim of this experimental work was to check the effectiveness of the SOM based segmentation approach applied to the database of the fluorescence in situ hybridization images in order to propose an automatic diagnosis system.

4.1 Results and Discussion

In this section, results of the automatic segmentation techniques described in section 3 are presented. In figure 4 we present an example of the segmentation obtained with simple automatic thresholding. From the example we can see that most of the reactions were localized and could be used for further FISH diagnosis calculations. It is easy to notice that after all, it will lead to some inaccuracies, especially in the red channel. The reason for that is that some of the nuclei assume some shades of red which will be misinterpreted with such a simple thresholding calculation. This will basically mean that the modes of the red channel histogram are not well separated. This scenario is better visualized in Fig. 7a.

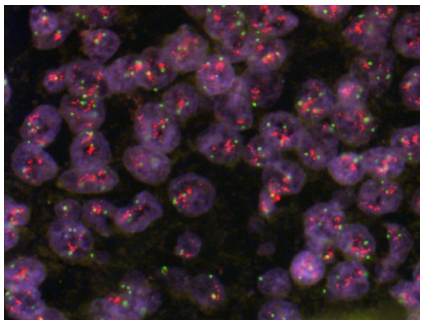
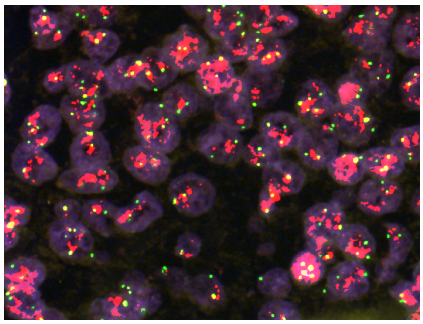
a)  b) 

Fig. 4. Results of simple thresholding.

Taking the above into consideration we noted that application of the global thresholding will lead to misinterpretations due the fact that there are diverse intensity values in different parts of the image. This is why we decided to flatten

the image maxima. This will allow us to use larger maxima areas instead of single peaks. For this purpose we decided to apply erosion and dylation operation followed by reconstruction according to the description in section 3.2. Results of the reconstruction are presented in Fig. 5.
From the results we can notice the detection of multiple one–pixel dots can provide many misleading results. This is because some of them will also include background noise, what could be noticed in case of simple thresholding.
Having the maxima determined and localized we can now check if the areas that are around the maxima contain red or green dot. As already described we proceed with division of the color image into sub–images that contain the maxima and treat it as a feature vector for SOM network. After dimension reduction we classify each of the extracted areas as a background or a red/green dot. A result of this operation is presented in Fig. 6. On the figure we can notice a good discrimination between background and dots. This is represented as a marker of a localization result.

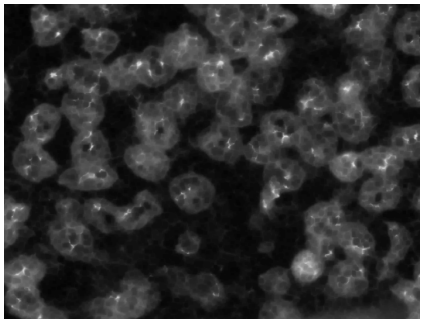
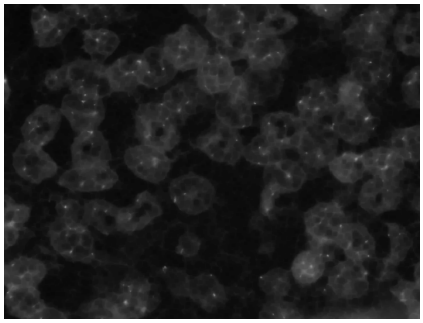
a)  b) 

Fig. 5. Results of morphological reconstruction. a) Red channel, b) Green channel.

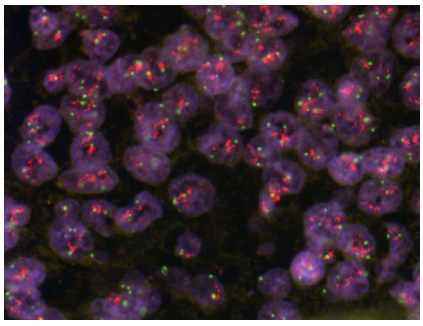
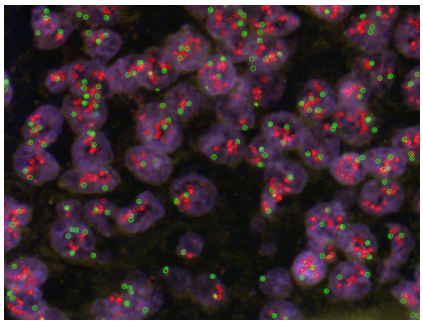
a)  b) 

Fig. 6. Results of SOM segmentation. a) Original image, b) Segmented dots.

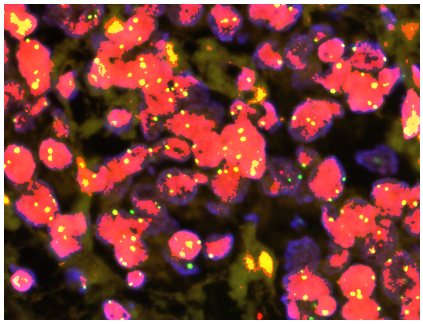
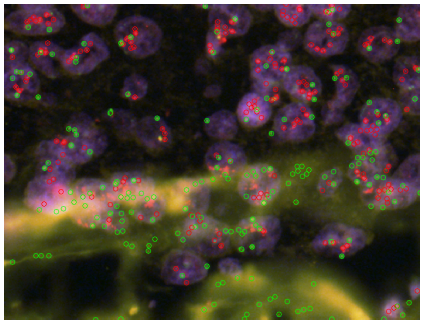
a)  b) 

Fig. 7. Examples of incorrect segmentations. a) Simple thresholding, b) SOM segmentation.

The presented segmentation results show that the SOM neural networks were able to properly distinguish HER2 and CEN-17 receptors. The main problem of this method occurs in a situation shown on Fig. 7b. Here a localization of green dots can be problematic but, in comparison with simple thresholding, we can still notice much better localization, especially in a red component.
To complete the full automatic diagnosis, a number of red and green dots was calculated and HER2 to CEN-17 ratio was calculated. Based on the calculated ratio and according to ACSO regulations a FISH score was evaluated. For the simple thresholding 44 out of 80 images provided similar results as an expert pathologist diagnosis. The other method provided better results were 61 images had similar responds.

5 Conclusions and Future Work

In this study a problem of automatic segmentation of HER2 and CEN-17 reactions from Fluorescence in Situ Hybridization images was addressed. We have proposed a fully automatic segmentation framework that is able to detect dots representing reactions of human epidermal growth factor receptor 2 and chromosome 17 centromere enumeration probe with high precision. Form the results we can draw a conclusion that neural network based method is able to segment these reactions more accurate than the described simple thresholding method. As already mentioned this is caused by multiple one–pixel dots which eliminates a possibility of application of a global thresholding methods. It was shown that utilization of a reconstruction methodology along with morphological erosion and dylation allows for a better localization of the reactions.
Although the results presented here are optimistic there are a few problems that can impair the overall diagnosis. The main issue is a localization of the dots in areas where they are not visible, as shown in Fig. 7. Another problem that can be distinguished here is segmentation of large number of dots as one large single dot. This can lead to incorrect interpretation of the result as several dots are counted as one.

One of the possibilities to improve the accuracy of the diagnosis would be an application of deep and convolutional neural networks which are showing to be very effective in cancer diagnosis. The main problem with these methods is necessity of a large number of training samples. This might be difficult to achieve as not all of the cases are stored digitally. This and other issues that are addressed here will be further researched as a future work.

References

1. National Cancer Registry. http://85.128.14.124/krn/, December 2013. (accessed on 18.04.2016).
2. Lerner B., W.F. Clocksin, S. Dhanjal, M.A. Hultén, and C.M. Bishop. Automatic signal classification in fluorescence in situ hybridization images. *Cytometry*, 43(2):87–93, 2001.
3. H.J.G. Bloom and W.W. Richardson. Histological Grading and Prognosis in Breast Cancer. *British Journal of Cancer*, 11:359–377, 1957.
4. H.D. Cheng, X. Cai, X. Chen, L. Hu, and X. Lou. Computer-aided detection and classification of microcalcifications in mammograms: A survey. *Pattern Recognition*, 36(12):2967–2991, 2003.
5. M. Duraisamy and F.M.M. Jane. Cellular neural network based medical image segmentation using artificial bee colony algorithm. In *Green Computing Communication and Electrical Engineering (ICGCCEE), 2014 International Conference on*, pages 1–6, March 2014.
6. P. Filipczuk, T. Fevens, A. Krzyżak, and R. Monczak. Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies. *IEEE Trans. Med. Imaging*, 32(12):2169–2178, 2013.
7. P. Filipczuk, B. Krawczyk, and M. Woźniak. Classifier ensemble for an effective cytological image analysis. *Pattern Recognition Letters*, 34(14):1748–1757, 2013.
8. T. García-Caballero, D. Grabau, A.R. Green, J. Gregory, A. Schad, E. Kohlwes, I. Ellis, S. Watts, and J. Mollerup. Determination of her2 amplification in primary breast cancer using dual-colour chromogenic in situ hybridization is comparable to fluorescence in situ hybridization: a european multicentre study involving 168 specimens. *Histopathology*, 56(4):472–480, 2010.
9. S. Gorjizadeh, S. Pasban, and S. Alipour. Noisy image segmentation using a self–organized map network. *Advances in Science and Technology Research Journal*, 2015.
10. D.G. Hicks and S. Kulkarni. Her2+ breast cancer: Review of biologic relevance and optimal use of diagnostic tools. *American Journal of Clinical Pathology*, 129:263–273, 2008.
11. D.G. Hicks and L. Schiffhaue. Standardized assessment of the her2 status in breast cancer by immunohistochemistry. *Lab Medicine*, 42(8):459–467, 2011.
12. Ł Jeleń, A. Krzyżak, T. Fevens, and M. Jeleń. Influence of Pattern Recognition Techniques on Breast Cytology Grading. *Scientific Bulletin of Wroclaw School of Applied Informatics*, 2:16–23, 2012.
13. Y. Ji, L. Sheng, X. Du, G. Qiu, B. Chen, and X. Wang. Clinicopathological variables predicting her-2 gene status in immunohistochemistry-equivocal (2+) invasive breast cancer. *Journal of Thoracic Disease*, 6(7):896–904, 2014.
14. I.T Jolliffe. *Principal Component Analysis. 2nd ed.* Springer, 2002.

15. G. Kiszler, L. Krecsák, A. Csizmadia, T. Micsik, D. Szabó, V. Jónás, V. Prémusz, T. Krenács, and B. Molnár. Semi-automatic fish quantification on digital slides. *Diagnostic Pathology*, 8(1):1–4, 2013.
16. T. Kohonen. The self–organizing map. In *Proceedings of IEEE*, volume 78, pages 1464–1480, 1990.
17. M. Kowal and P. Filipczuk. Nuclei segmentation for computer-aided diagnosis of breast cancer. *Applied Mathematics and Computer Science*, 24(1):19–31, 2014.
18. M. Kowal, P. Filipczuk, A. Obuchowicz, J. Korbicz, and R. Monczak. Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images. *Computers in Biology and Medicine*, 43(10):1563 – 1572, 2013.
19. O. Lezoray, A. Elmoataz, H. Cardot, G. Gougeon, M. Lecluse, and M. Revenu. Segmentation of cytological images using color and mathematical morphology. *European Conference on Stereology, Amsterdam, Netherlands*, page 52, 1998.
20. J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. *CoRR*, abs/1411.4038, 2014.
21. M. Moghbel and M. Mashohor. A review of computer assisted detection/diagnosis (CAD) in breast thermography for breast cancer detection. *Artif. Intell. Rev.*, 39(4):305–313, 2013.
22. M. B. Nagarajan, M. B. Huber, T. Schlossbauer, G. Leinsinger, A. Krol, and A. Wismüller. Classification of small lesions on dynamic breast mri: Integrating dimension reduction and out-of-sample extension into {CADx} methodology. *Artificial Intelligence in Medicine*, 60(1):65 – 77, 2014.
23. H. Netten, L.J. van Vliet, H. Vrolijk, W.C.R. Sloos, H.J. Tanke, and I.T. Young. Fluorescent dot counting in interphase cell nuclei. *Bioimaging*, 4(2):93–106, 1996.
24. T.W. Ridler and S. Calvard. Picture thresholding using an iterative selection. *IEEE Trans. on Systems, Man and Cybernetics*, 8:630–632, 1978.
25. D.J. Slamon, B. Leyland-Jones, S. Shak, H. Fuchs, V. Paton, A. Bajamonde, T. Fleming, W. Eiermann, J. Wolter, M. Pegram, J. Baselga, and L. Norton. Use of chemotherapy plus a monoclonal antibody against her2 for metastatic breast cancer that overexpresses her2. *New England Journal of Medicine*, 344(11):783–792, 2001.
26. H. J. Tanke, R. J. Florijn, J. Wiegant, A. K. Raap, and J. Vrolijk. Ccd microscopy and image analysis of cells and chromosomes stained by fluorescence in situ hybridization. *The Histochemical Journal*, 27(1):4–14, 1995.
27. L Tibau, A.and López-Vilaró, M. Pérez-Olabarria, T. Vázquez, C. Pons, I. Gich, C. Alonso, B. Ojeda, T. Ramón y Cajal, E. Lerma, A. Barnadas, and D. Escuin. Chromosome 17 centromere duplication and responsiveness to anthracycline-based neoadjuvant chemotherapy in breast cancer. *Neoplasia*, 16(10):861–867, 2014.
28. M. Velikova, P. J. F. Lucas, m. Samulski, and N. Karssemeijer. On the interplay of machine learning and background knowledge in image interpretation by bayesian networks. *Artificial Intelligence in Medicine*, 57(1):73 – 86, 2013.
29. Luc Vincent. Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms. *IEEE Transactions on Image Processing*, 2:176–201, 1993.
30. K.C. Yao, M. Mignotte, C. Collet, P. Galerne, and G. Burel. Unsupervised segmentation using a self-organizing map and a noise model estimation in sonar imagery. *Pattern Recognition*, 33:1575 – 1584, 2000.
31. Dana Carmen Zaha. Significance of immunohistochemistry in breast cancer. *World Journal of Clinical Oncology*, 5(3):382–392, 2014.