

Chapter 2

Automatic segmentation of lung cancer cells with the new parameters by using methods of image processing and analysis.

Przemysław Jędrusik¹, Robert Koprowski¹, Ilona Bednarek², and Zygmunt Wróbel¹

Abstract. Modern diagnostic methods allow to get multiple information regarding research material. This work focused on the development of an algorithm for automatically determining the correct number of cells. The developed tool allows the detection of cells as individual objects, searching for the objects significantly larger than the sought and checking if they were a combination of objects. The algorithm was based on additional parameters designated in its subsequent steps as well as their respective correcting claimed searched result. Analyzed a large number of images, it was found that there is a close relationship between the surface area of the cells, the degree of extension and the location and correct detection of objects that are neither a cluster of cells, and nothing significant image artifacts. The developed algorithm was written using Matlab software.

Keywords: image processing, algorithms, automatic, lung cancer, cell culture, cell migration

2.1 Introduction

The huge number of medical information obtained in the process of computer - assisted acquisition, allows to reach the growing number of different cases. Existing barriers to obtaining large amounts of data, perform complex and time-consuming

Department of Computer Biomedical Systems, University of Silesia,
Institute of Computer Science,
ul. Będzińska 39, 41-200 Sosnowiec, Poland,
e-mail: pjedrusic@o2.pl, robert.koprowski@us.edu.pl · Department of Biotechnology and Genetic Engineering, Medical University of Silesia,
School of Pharmacy with the Division of Laboratory Medicine in Sosnowiec,
ul. Jedności 8, 41-200 Sosnowiec, Poland,
e-mail: ibednarek@sum.edu.pl

calculations have been mostly eliminated by including in the process a large computing power of computer equipment. This paper refers to the problems connected with behavior and relationships between lung cancer cells. Nowadays the problem of cancer is very timely and important especially. Lung cancers are still at the forefront of medical conditions directly threatening the lives of people every year, in 2012 were over 410.000 new lung cancer cases diagnosed [5]. Understanding the mechanisms of cell behavior significantly affects the effective methods of antitumoral immunotherapies. It was not possible to real analyze the different factors that determine how migration and cell growth without the correct segmentation [8]. Analyzes of changes in saved medical images, are mainly based on the correct separation of visible cells from the background image. Separation of the course has a lot of problems regarding the designation of individual cells as separate objects. This article proposes the use of adaptive method of binarization including the results obtained after the binarization. The results allow properly direct algorithm by eliminating objects, which cannot be the ones we are looking for, because of their morphology.

2.2 Materials in experiment

The study of this work is based on microscopic analysis of medical images A549 lung cancer cells ATCC CCL-185TM. Lung cancer cell line originated from the group ATCC, grown in accordance with the manufacturer's instructions. Cell cultures were routinely propagated in a humidified incubator Hera-Cell (Heraeus), at 37°C, in 5% CO₂ atmosphere. Cells were grown in RPMI-1640 medium supplemented with 10% FBS and gentamicin (20 µg/ml). Cells were seeded in 12-well plates to reach confluency. Cells were allowed to growth and migrate for 24, 48 and 72 hours. Cells were visualized under inverted microscope (OLYMPUS BX-60 microscope with DP50 digital camera) at indicated time points; magnification 100x. The acquisition of microscopic images was performed in the Department of Biotechnology and Genetic Engineering Medical University of Silesia in Katowice. Were registered 120 different images. All obtained images is provided in each of the same area of the cell culture. The images in digital form are saved as a bitmap with a resolution of 2080x1540 pixels. Pixel of acquired images was dimension of approximately 10 nanometers in real. All images were analyzed using a proprietary algorithm, written in Matlab version R2013b software with Image Processing Toolbox version 8.3. The software was running on a computer with an Intel i5 processor clocked at 2.6GHz, along with 8GB of DDR3 RAM, serial ATA hard disk drive and operating system Microsoft Windows 10 Pro (version 1511, compilation 10586.63).

2.3 Methods

Designed algorithm was divided into 2 stages. In a first step, images treated with the methods of image analysis (pre-processing). A median filter was used to eliminate artifacts in the input images. [1][2]. Filter size was matched to the size of objects that were artifacts of a fixed local window size of 19x19 pixels [7]. The adaptive binarization method was used in the next step with local window, which size was adapted to the size of the areas, collections of pixels of brightness changes [3]. Typically, methods of binarization based on a global binarization threshold determined for all pixels in the image. Modifications of these methods have been used in the article, which analyzes an image with the division into regions, were used a local adaptive thresholding techniques.

An example of input image can be found in Figure 1.

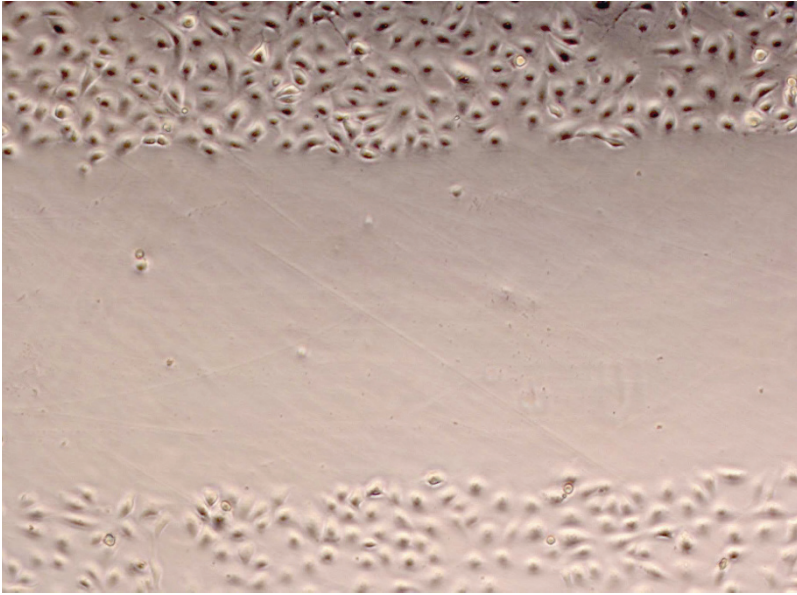


Fig. 2.1: Example source image of lung cancer cells.

$$bw(x,y) = \begin{cases} 0, & \text{if } image(x,y) < th(x,y) \\ 1, & \text{otherwise,} \end{cases} \quad (2.1)$$

where $th(x,y)$ was a threshold for each pixel at location (x,y) .

In this paper was used the Sauvola's method of binarization, based on the standard deviation and average value of the pixel at a specified point [9].

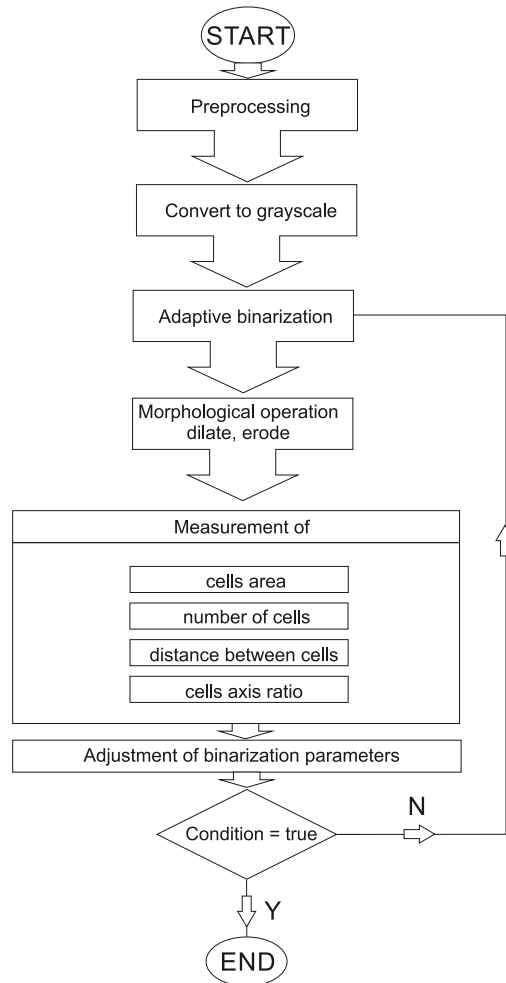


Fig. 2.2: Schematic of the proposed algorithm for automatic segmentation of lung cancer cells.

$$th(x,y) = m(x,y) \left[1 + k \left(\frac{s(x,y)}{R} - 1 \right) \right], \quad (2.2)$$

where threshold $th(x,y)$ is computed using $m(x,y)$ a mean, $s(x,y)$ a standard deviation of the pixel intensities in a $w \times w$ window and where R is the maximum value of standard deviation and k is a positive parameter to modular according the analyzed image. In the next stage used morphological operations: erosion and dilation with selected the structuring element, allowing remove objects that are considered distortion.

In the next step specific cell parameters were determined, that represented their individual properties. According to the proposed algorithm for each of the cells were determined by the following parameters:

- surface area,
- the center of gravity,
- major and minor diagonals.

After considering these parameters were determined by the number of detected cells and the distance between them, specifically between the centers of gravity of the cells. The second stage of the algorithm was repeated until the appropriate values determined parameters. The stop condition for the main loop has been reached a local extreme value which depended on the average size of detected objects and the maximum size of the object. Detected objects (cells) had a surface area less 2.5 times than the average cell size in the picture, and had size not less than half the average size of all detected cells. Characteristic of biological cells is their growth and division [6]. It was assumed that the cell before division for the next can be more than 2 times larger than the average cell size, which at this stage does not exhibit growth or division [4]. Specifically, the object was correctly detected as a single cell, if the size was less than 2.5 times the average cell size. In analogy, the cells which surface area were substantially smaller than the average value for all the cells, were not taken into account. Such objects were treated as artifacts of the image.

$$\forall_{c \in C} (0.5 \cdot C_{meanarea} < area(c) < 2.5 \cdot C_{meanarea}), \quad (2.3)$$

where C was set of all objects (cells), $C_{meanarea}$ was a mean area of all cells in set C .

The research were conducted on a group of 120 pictures cell cultures at different times of proliferation. A wide group of input data allowed to obtain the results shown in the next section of work, and to formulate conclusions.

2.4 Results

Implementation and execution of the designed algorithm for a example input images allowed to obtain the data. Table 1 contains the partial data, collected for the sample input image.

Analyzed table containing the results, it was found that the selected parameter k is associated with an average size of objects and the size of the maximum detected object in the image. The reduction of the average size of the cell area, also decreases the surface area of the maximum field. Determined the appropriate value of k , allowed to obtain an information regarding the change of the surface area of all cells, the average distance between them, the average ratio of the diagonals, and the amount of detected objects. The designated objects do not differ from each other in terms of surface area. The shape of the objects was stretched slightly, the ratio of the diagonals was less than 2. The distance between the cells was significantly higher

Table 2.1: Part of computed data, automated determine k - value

	k	cells	mean area [px]	max cell area [px]	mean distance [px]	major/minor axis
1	0,32	264	662	658	924	1,67
2	0,36	152	685	677	1026	1,88
3	0,22	225	1042	1038	699	1,86
4	0,24	164	1181	1172	887	1,97
5	0,18	282	860	854	1024	1,98
6	0,34	178	832	829	966	1,69
7	0,42	162	797	788	924	1,79
8	0,38	224	660	636	1029	1,66
9	0,26	197	786	781	1104	1,79

than the average cell size, which leads to the conclusion that there were blank areas between cells.

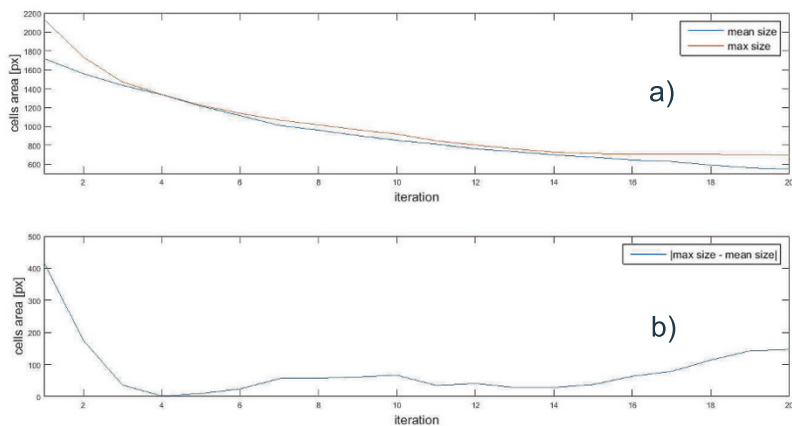


Fig. 2.3: a) mean cells area and area of largest detected cell, b) the absolute value of the difference between mean cells area and largest detected cell

The figure 3 shows example of the dependence of the average cell size of the area and the area of the largest cell in the subsequent steps of the loop algorithm.

The figure 4 shows example of the mean distance between cells, mean major and minor axis ratio and number of cells on iteration steps. For the purposes of that study, loop was stopped at 20-th iteration.

The figure 5 shows the optimum automatic detection of objects - cell lung cancer.

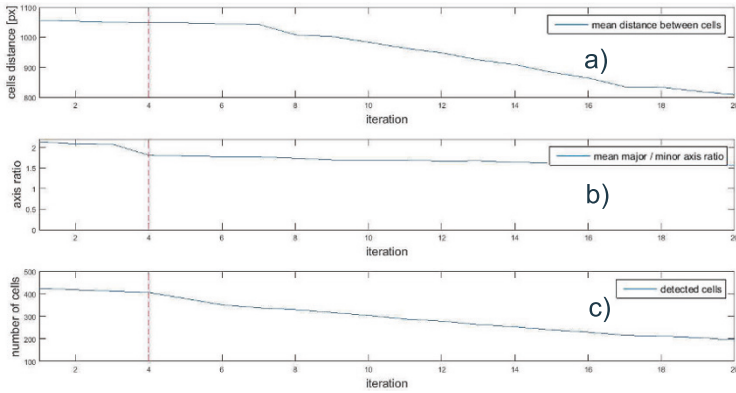


Fig. 2.4: a) mean distance between cells, b) mean major and minor axis ratio, c) number of detected cells

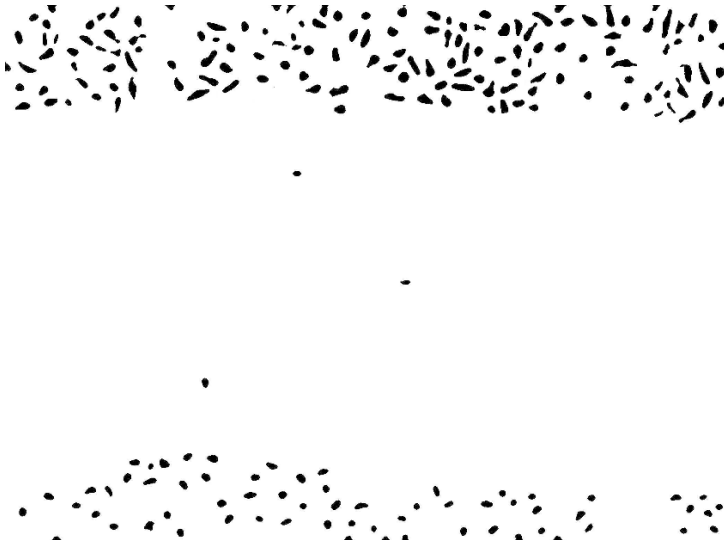


Fig. 2.5: Example of automatic segmentation.

2.5 Discussion

Execution of the algorithm loop allowed to automatically determine the values of k occurring in the formula (2). In the next iterations, the average value of the cell area identified as targets has been relevant to the area of the largest object detected. This means that in each successive iterations of the objects that were detected as significantly larger than the other were the appropriate parameters change binarization detected as a greater number of smaller objects. End of the loop gave eliminating

all objects significantly larger than the mean. New objects do not differ in size from mid-sized cells. At the same time the objects significantly smaller than the average size of the cell surface have been removed from the set of designated objects. In addition, determined the objects characterized by reduced the diagonals ratio, where it was less than 2. This means that the objects are not too elongated. It was also observed the relationship between the detection of a valid objects, and the average distance between them. Changing the average distance between the cells changed in a linear with minimal inclination of the trend line to the x-axis.

2.6 Conclusions

The proposed algorithm allows automatic adjustment of imaging parameters of cell cultures. The next steps of the algorithm was based on the information obtained in the previous steps. Parameters determining the results obtained were: the average area of cells, the surface area of the cell with the largest distance between the centers of gravity of the cells, the ratio of the diagonals of the set of cells and the number of assigned cells. The parameters clearly define the appearance of objects obtained, the distribution of the image. The entire process runs automatically, allows the analysis of a large number of images. As a tool to assist assessment of the behavior of tumor cells can be widely applied not only to lung cancer cells, but also to any other cell line. Naturally you must specify the biological properties of growing and migrating cells. Determination of the optimal number of cells further, while maintaining the condition that the cells can overlap or merge. They may not be properly recognized. Creation of this tool also allows to specify automatically confluent cells value, which is a very important determinant performed by the investigator. Information about the confluence was included in the number of cells, but especially in the distance between them. This means that determination of the size of the free space between cells - as one of the parameters of this study. This of course will require the implementation of further research.

References

1. Acharya, T., Ray, A.K.: Image Processing - Principles and Applications. Wiley InterScience; 2006
2. American Cancer Society. Global Cancer Facts & Figures 2nd Edition. Atlanta: American Cancer Society; 2011.
3. Angenent, S., Pichon, E., Tannenbaum, A.: Mathematical methods in medical image processing. Bulletin of the American mathematical society, 43, 365-396; 2006
4. Elizabeth Flate, John R. D. Stalvey: Motility of select ovarian cancer cell lines: Effect of extracellular matrix proteins and the involvement of PAK2. Int J Oncol. 2014 Oct; 45(4): 1401-1411.

5. J. Ferlay, E. Steliarova-Foucher, J. Lortet-Tieulent, S. Rosso, J.W.W. Coebergh, H. Comber, D. Forman, F. Bray; Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012; *European Journal of Cancer* (2013) 49, 1374–1403
6. Jianping Peng, Ganesan Ramesh, Lin Sun, and Zheng Dong: Impaired Wound Healing in Hypoxic Renal Tubular Cells: Roles of Hypoxia-Inducible Factor-1 and Glycogen Synthase Kinase 3 β / β -Catenin Signaling. *J Pharmacol Exp Ther.* 2012 Jan; 340(1): 176–184.
7. Koprowski, R., Korzyńska, A., Zieleźnik, W., Wróbel, Z., Małysek, J., Stępień, B., Wójcik, W.: Influence of the measurement method of features in ultrasound images of the thyroid in the diagnosis of Hashimoto's disease. *BioMedical Engineering OnLine*, 11:91 (2012)
8. Nilendu C Purandare, Venkatesh Rangarajan; Imaging of lung cancer: Implications on staging and management; *Indian J Radiol Imaging.* 2015 Apr-Jun; 25(2): 109–120
9. Sauvola J., Pietikainen M.: Adaptive document binarization; *Pattern Recognition* 33(2), 2000, p. 225-236

Innovations in Biomedical Engineering

Gzik, M.; Tkacz, E.; Paszenda, Z.; Piętka, E. (Eds.)

2018, XXXIII, 410 p. 210 illus., 130 illus. in color.,

Softcover

ISBN: 978-3-319-70062-5