IMPLEMENTATION OF THE WATERSHED METHOD FOR SEPARATING NUCLEI OF CANCER CELLS IN FLUORESCENT IMAGES

Y.U. Lisitsa

Belarussian State University, Minsk, Belarus
E-mail: ylisitsa@bsu.by

This article presents the implementation of the watershed method for the segmentation of nuclei in fluorescent images of cancer cells. In this study, we apply the watershed algorithm to fluorescent images of cancer cells, aiming to accurately find the boundaries of nuclei for further analysis. The article discusses the methodology used for preprocessing the images and applying the watershed algorithm. The watershed shows the following drawbacks: watershed lines run through nuclei rather than along their boundaries, merged nuclei persist after processing, over-segmentation leads to noise on the nuclei mask.

Key words: watershed; fluorescent microscopy; image processing.

РЕАЛИЗАЦИЯ МЕТОДА ВОДОРАЗДЕЛА ДЛЯ РАЗДЕЛЕНИЯ ЯДЕР РАКОВЫХ КЛЕТОК НА ФЛУОРЕСЦЕНТНЫХ ИЗОБРАЖЕНИЯХ

Е.В. Лисица

Белорусский государственный университет, Минск, Беларусь
E-mail: ylisitsa@bsu.by

В данной статье представлена реализация метода водораздела для сегментации ядер на флуоресцентных изображениях раковых клеток. В этом исследовании мы применяем алгоритм водораздела к флуоресцентным изображениям раковых клеток, стремясь точно найти границы ядер для дальнейшего анализа. В статье обсуждается методология предварительной обработки изображений и применения алгоритма водораздела. Метод водораздел имеет следующие недостатки: линии водораздела проходят через ядра, а не вдоль их границ, слитые ядра сохраняются после обработки, чрезмерная сегментация приводит к шуму на маске ядер.

Ключевые слова: водораздел; флуоресцентная микроскопия; обработка изображений.

INTRODUCTION

Fluorescence microscopy plays a pivotal role in cancer research by providing high-resolution imaging of cellular structures. Accurate segmentation of nuclei from fluorescent images is essential for understanding cellular morphology and pathology [1]. One of the most efficient ways of image segmenation is the usage of the artificial neural networks, such as convolutional neural networks. However, as they estimate the probabilities of the nuclei locations in the image, it occurs to the presences of the clumped objects on their binary masks during the thresholding [2]. The watershed method is a prominent technique for segmenting overlapping objects in digital images. In the context of cancer cell analysis, it offers a promising approach to delineate individual nuclei within densely packed regions [3]. This paper introduces an implementation of the watershed method tailored for separating nuclei of cancer cells in fluorescent images. The aim of the paper is to implement the realization of the watershed segmentation method with markers based on the image preprocessing.

DATA

The article examines three-channel fluorescent images of breast cancer cells, obtained with a 10x magnification on a Nikon TE200 inverted epifluorescence microscope equipped with a Photometrics 300 series CCD camera. The image size is 2048×2048 pixels. The indicator of cancer cells is the estrogen receptor protein labeled with Cy5 dye (registered in the red channel of the image). In the cytoplasm of cancer cells, the cytokeratin protein accumulates, were labeled with Cy3 dye (registered in the green channel). The 4,6-DAPI dye was used to label all nuclei (registered in the blue channel). Accordingly, the markers for cancer cells are two dyes – Cy5 and Cy3[4].

The images were preprocessed using the algorithm described in paper [2] to estimate the probabilities of nuclei locations. The binary masks were obtained by thresholding the probability images, with a threshold of 0.5.

WATERSHED ALGORITHM

Let I be the input image, where $I(x,y)$ represents the intensity at pixel coordinates (x,y). We consider I as a topographic surface.

1. **Gradient Calculation**: Compute the gradient magnitude of the image to identify potential watershed lines. This can be represented mathematically as: $\nabla I = |\nabla I| = (\partial I/\partial x)^2 + (\partial I/\partial y)^2$

2. **Marker Generation**: Identify markers for the watershed transformation. These markers typically correspond to the local minima of the gradient magnitude image. Let M represent the marker image, where $M(x,y)=1$, if pixel (x,y) is a marker, and $M(x,y)=0$ otherwise.

3. **Flood-Filling**: Simulate a flooding process starting from the markers. This process is akin to placing water at each marker and letting it fill the image. At each step, the water from neighboring pixels flows towards the marker with the lowest intensity. This process continues until all pixels are assigned to a marker. Mathematically, this can be represented using a flooding function F defined recursively as: $F(x,y)=(x',y')\in N(x,y)\ \min\{F(x',y')+1\}$, where $N(x,y)$ is

the set of neighboring pixels of *(x,y)*. The initial values of *F* are set to 0 at marker locations.

4. **Watershed Line Formation**: Watershed lines form where the flooding fronts from different markers meet. These lines delineate the boundaries between segmented regions. They can be represented as a binary image *W*, where $W(x,y)=1$ indicates a watershed line pixel and $W(x,y)=0$ otherwise.

Watershed segmentation involves computing the gradient of the image, identifying markers, simulating a flooding process from these markers, and finally extracting watershed lines to segment the image into regions [5].

IMPLEMENTATION

The flow-chart of image processing for watershed implementation is shown in the figure 1.

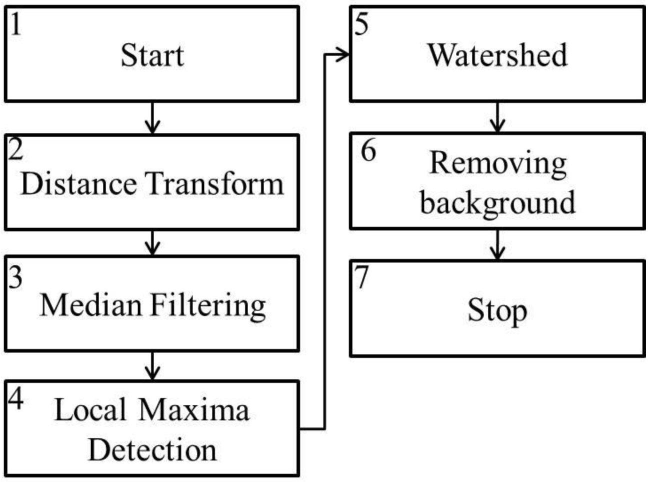

Fig. 1. Watershed implementation flow-chart

1. **Start**. The input consists of a mask of clumped nuclei ($nuclei_{clumped}$), which represents an input image containing clumped nuclei, and the minimum area of nuclei ($area_{min}$).

2. **Distance Transform**. The code computes the distance transform of $nuclei_{clumped}$ using Euclidean distance transform, storing the result in $image_{transform}$.

3. **Median Filtering**. It calculates the radius (r_{min}) based on the $area_{min}$. Then, it applies median filtering to $image_{transform}$ using a square-shaped structuring element with a size determined by r_{min}.

4. **Local Maxima Detection**. It identifies the coordinates of peak local maxima in the negated $image_{transform}$. Than it uses a 3x3 neighborhood and assigns labels based on $nuclei_{clumped}$.

5. **Watershed**. It initializes a mask based on the peak local maxima coordinates and performs watershed segmentation on the negated distance

transform image using these markers. The resulting segmented image is stored in *labels*.

6. **Removing background**. Finally, it removes the background (regions labeled as 0 in *labels*) from $nuclei_{clumped}$, isolating the individual nuclei.

The implementation is made on the Python language and it includes following libraries: numpy, skimage, scipy.

RESULTS

The results of segmentation are shown in the figure 2. The grayscale image (Fig. 2A) is presenting the estimations for the nuclei probability locations. The figure 2B shows only clumped nuclei after thresholding. The distance transform image after filtration is presented in the figure 2C. The separated nuclei by the watershed method are shown in the image 2D.

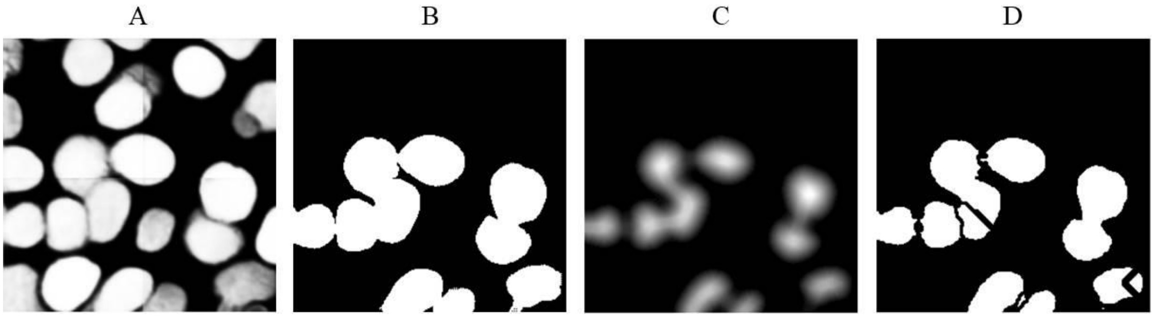

Fig. 2. Segmentation results:
A - Grayscale image; B - Clumped nuclei; C - Distance transform image after filtration; D - Separated nuclei

CONCLUSION

As indicated by the segmentation results, the watershed method allows for separating merged objects. Even though our implementation leverages image processing techniques to enhance segmentation accuracy and efficiency, the disadvantages of the method include:

1. Watershed lines do not run along the boundaries of nuclei but through the nuclei themselves.
2. Merged nuclei remain after processing with the method.
3. Over segmentation, this also results in noise presence on the nuclei mask.

REFERENCES

1. Molecular Biology of the Cell / B. Alberts [et al.] // New York: Garland Science, 2012. P. 1392.
2. *Silun X., Skakun, V.* Fully-automated segmentation of tumor nuclei in cancer tissue images // Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21-24 Sept. 2021, Minsk, Belarus. Minsk : UIIP NASB, 2021. P. 168-172.
3. *R. C. Gonzalez, R. E. Woods, S. L. Eddins.* Gonzalez, R. C. Digital Image Processing Using MATLAB // Gatesmark Publishing, 2020.
4. *Е. В. Лисица, Н. Н. Яцков, В. В. Скакун, П. Д. Кривошеев, В. В. Апанасович.* Отбор информативных геометрических признаков ядер клеток на люминесцентных изображениях раковых клеток // Информатика. 2019. Т. 16, № 2. С.7-17.
5. *L. Najman, M. Schmitt* Watershed of a continuous function // Signal Processing (Special issue on Mathematical Morphology. 1994. Vol. 38, P. 99–112.