The method for glomerulations detection in histological images of prostate

A A Zavarzin\textsuperscript{1}, A N Pronichev\textsuperscript{1}, O V Rodionova\textsuperscript{1}, E A Komochkina\textsuperscript{1}, E A Prilepskaya\textsuperscript{2}, M V Kovylina\textsuperscript{2}

\textsuperscript{1} National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Kashirskoe shosse 31, 115409, Moscow, Russia
\textsuperscript{2} Moscow State University of Medicine and Dentistry named after A.I. Evdokimov, Delegatskaya street, 20/1, 127473, Moscow, Russia

e-mail: Zavarzin.mephi@gmail.com

Abstract. In the work presented, a method for detecting glomeruli in pictures of histological preparations of the prostate gland is described, the presence of which indicates a malignant neoplasm. Pathological structures at the level of microimages are investigated. The developed method is the result of joint activity of the National Research Nuclear University "MEPhI" and the Moscow State Medical and Stomatological University named after A.I. Evdokimova.

1. Introduction
In the global structure of oncological morbidity, prostate cancer ranks sixth, and among men - the third. Today’s crucial challenges are the development and implementation of automated software systems to diagnose prostate cancer. Solutions to these challenges could allow us to automate the diagnosis of cancer, increase its accuracy and, therefore, accelerate a decision-making process regarding the methods of medical intervention. This pathology is annually diagnosed in more than half a million people, or about one tenth of all oncological diseases in men [1]. At present, there is no counterpart to the developed method for detecting glomeruloid structures.

The task of the research stage is to extract and describe characteristic features of the pathological structures, called glomeruli, and distinguish them from other objects of the histological specimens. Figure 1 presents examples of glomeruli.

\textbf{Figure 1.} Histological images of glomeruli
The formation of glomeruloid bodies is a rare type of cancer differentiation grades. They are made up of solidifying malignant cells, surrounded by crescent-shaped spaces, and then by the wall of an enlarged acinus or duct with a lining of the cancerous epithelium. In accordance with the Gleason index, a tumor of such a differentiation generally corresponds to grade 4. On the contrary, healthy cells produce glandular formations containing spaces of round, ellipsoidal shape. An example of benign glandular structures is shown in Figure 2.

![Figure 2. Histological images of benign glandular structures](image)

The purpose of the work is to develop and automate the procedure for detecting glomeruloid structures in the images of histological specimens of the prostate gland.

2. Materials and methods

Histological specimens of the prostate gland are subjected to detailed analysis. The average size of such images is 1300 x 900 pixels; the color depth is 24 bits.

A binarization method is used to extract ducts of glandular structures (white spaces) contained both in the areas of benign cellular formations and in glomeruli. In the binarization process, the original halftone image having a certain number of brightness levels is converted to a black and white image whose pixels have only two values – 0 and 1 [2]. Threshold processing of the image can be carried out in different ways. To use a binarization method, we convert the original three-channel (RGB) image into a halftone image with a color depth of 8 bits (256 semitones). Further, we apply the binarization method with an upper threshold:

\[
f'(m,n) = \begin{cases} 0, & f(m,n) \geq t \\ 1, & f(m,n) < t \end{cases}
\]

Where \(f(m,n)\) is the pixel half-tone level with coordinates \((m,n)\), \(f'(m,n)\) is the value (0/1) of the pixel as a result of binarization, \(t\) is the threshold value. Due to the study of histological specimens, the threshold value \(t\) can be determined.

![Figure 3. Binarization result](image)
Figure 3 shows the outcome of applying the binarization method. The black channels are the ducts of the prostate. Further we extract the contours of the gaps for subsequent analysis of their shapes. In order to operate with contours, we present them in the form of a chain code of Freeman (Freeman Chain Code) [3]. Figure 4 shows the result of the contour extraction operation.

![Figure 3. Result of the contour extraction](image)

Then we extract crescent-shaped contours. Within the framework of the task set, the criterion for detecting such contours is as follows: if the center of mass of a contour does not belong to its inner area, then the contour is considered to be of crescent shape. With the purpose of determining whether or not the point belongs to the contour- bounded area, the ray tracing method can be used. Assume that it is necessary to define whether point A is related to K contour. To do this, we draw a straight line to point A from some remote point. Along that direction there can occur zero or several intersections of the boundary of the contour: we enter the contour at the first intersection, we exit from it at the second one. If we reach point A with an odd number of intersections of the contour boundary, then point A lies inside the contour, and if an even number of intersections is produced, then the point is outside the contour K. In this case, point A is the contour’s center of mass. As shown in Figure 5, straight lines to point A are drawn along four directions.

![Figure 5. Determination of the point belonging to the contour area Direction of straight lines to the studied point](image)

Figure 6. Result of extraction of crescent- shaped contours
Figure 6 shows the result of extracting crescent-shaped contours. The green rectangle contains a contour satisfying the above criterion, the blue ones - the rest contours.

3. Results

For computer image processing in C++ using the Qt library and the OpenCV computer vision library [4], a software module with a minimal graphical user interface has been developed that allows downloading images and applying automatic analysis, as well as displaying results (Figure 7).

Precision and recall are the metrics used in evaluating the majority of the information extraction algorithms. The system precision within the class is the proportion of documents actually belonging to a given class relative to all documents that has assigned to this class by the system. Recall is the proportion of the classifier's documents belonging to the class relative to all documents of this class in the test set.

Three-channel images of histological specimens of the prostate with a size of 1,300 x 900 pixels with a color depth of 24 bits (8 bits per color component) were used as a test set. These images have been produced as a result of splitting the original image size of 60,000 x 130,000 pixels into tiles. The number of images used for the test set is 40. As a result of testing, the following values for precision and recall have been obtained:

\[
\text{Precision} \approx 90\%, \\
\text{Recall} \approx 70\%.
\]

![Figure 7. Algorithm results](image)

Acknowledgments

This work was supported by Competitiveness Program of National Research Nuclear University MEPhI.

References

[1] Stratsky A. V., Yushko E. I. 2015 Cancer of the prostate (Minsk, BSMU) 16 pages [in Russian]
[2] Shapiro L. 2006 Computer vision (Moscow, Binom) 752 pages [in Russian]
[3] Freeman H. 1961 On the encoding of arbitrary geometric configurations, IRE Transactions on Electronic Computers EC 10 260-268
[4] Bradski G., Kaehler A. 2008 Learning OpenCV: Computer vision with the OpenCV library (O'Reilly Media, Inc.) 556 pages