**Michał KRUK[1], Stanisław OSOWSKI[2,3],Wojciech KOZŁOWSKI[4], Robert KOKTYSZ[4]**
**Tomasz MARKIEWICZ[2,4], Janina SŁODKOWSKA[4]**

Warsaw University of Life Sciences (1), Warsaw University of Technology (2)
Military University of Technology, Warsaw (3), Military Institute of Medicine, Warsaw (4)

# Computer-assisted Fuhrman grading system for the analysis of clear-cell renal carcinoma: a pilot study

***Abstract***. *The paper presents an automatic computer system for evaluation of the Fuhrman degree in renal carcinoma, of the accuracy comparable to the human expert. The solution uses the combined methods of mathematical morphology, Hough transform and neural networks for the estimation of Fuhrman degree of the carcinoma clarocellular cells, based on the microscopic kidney image. The results of numerical experiments have shown that the average discrepancy rate between the score of our system and the human expert results estimated on the basis of almost 300 cells is below 10% and this accuracy is acceptable in the medical practice.*

.

***Streszczenie*** *Praca przedstawia podejście komputerowe do automatycznej oceny stopnia skali Fuhrmana w przypadku raka nerki. Ocena dotyczy mikroskopowego obrazu nerek. Proponowane rozwiązanie stosuje zespół metod obejmujących morfologię matematyczną, transformację Hougha, sieci neuronowe oraz grupowanie danych wielowymiarowych. Proponowane rozwiązanie zostało sprawdzone na zbiorze prawie 300 obrazów nerek z różnym stopniem zaawansowania choroby nowotworowej. (**Komputerowy system oceny stopnia zaawansowania raka nerek według skali Fuhrmana**)*

**Keywords:** Fuhrman grading, renal carcinoma cells, mathematical morphology, data clustering and classification, Hough transform.
**Słowa kluczowe:** skala Fuhrmana, rak nerki, morfologia matematyczna, grupowanie i klasyfikacja danych.

## Introduction

The grading schema of clear-cell renal carcinoma (CCRCC) is based on the microscopic image of a neoplasm cells at application of hematoxylin and eosin (H&E) staining. The most popular and widely used system for grading RCC is a nuclear grading system described in 1982 by Fuhrman [1,2]. This grading affects mainly the prognosis of the development of the illness. It is defined on a scale of 1-4, where grade 1 carries the best prognosis and grade 4 the worst. The grade is strictly correlated with stage in that larger tumors tend to be higher grade.

Nuclear grade means that the system is based on just the appearance of the nuclei of the cancer cells, rather than the appearance or structure of the cells as a whole. Fig. 1 presents typical examples of cells corresponding to different grading [9].

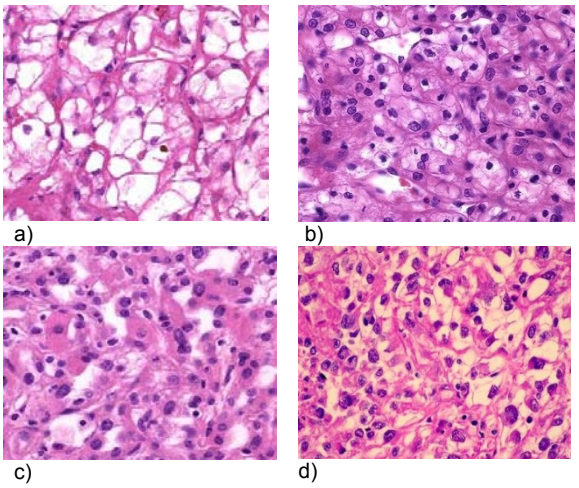


Fig. 1. The typical examples of different grade clear-cell renal carcinoma: a) grade 1, b) grade 2, c) grade 3, d) grade 4

Grade 1 tumors have round, uniform nuclei with inconspicuous nucleoli. They are very little different from normal kidney cells. Nuclear contours at grade 2 are more irregular than grade 1; the nuclei are about 15 microns in diameter. They may be visible at high magnification. At grade 3 the nuclear contours are even more irregular in size and shape. Nuclear diameters can approach 20 microns and the nucleoli are readily seen. The cells at grade 4 look quite different from normal kidney cells. Their size exceeds 20 micrometers, we observe the pleomorphic and hyperchromatic nuclei, and prominent nucleoli in a few cells. The difficulty in accurate nuclear grading arises due, in part, to the heterogeneity of nuclear features in the same tumor.

Fuhrman system utilizes few nuclear features for assigning a specific grade: number and size of nucleoli, nuclear size, nuclear shape and chromatin pattern [1].

The main task of this project is to develop the automatic computer system which would evaluate Fuhrman degree with the accuracy comparable to the human expert. The paper presents only the first step in this direction, including the mathematical morphology and Hough transform to extract the cells from the whole image, and neural networks used for the estimation of Fuhrman grade of the clear-cell renal carcinoma, on the basis of the microscopic kidney image.

## The general algorithm for Fuhrman grading

To get the automatic grading system we have to solve the following tasks: segmentation of the individual cells from the image using the mathematical morphology and color-based operations, the identification of the real nuclei cells of carcinoma using Hough transform, generation of features of the extracted cells, selection of the most important features, and the final classification stage relying on the clusterization and classification of data. According to this description the algorithm is composed of the following stages.

- Reading the image,
- Application of morphological operations to extract the cells from the image,
- Application of color-based operations and Hough transform used for nuclei identificatation. After this step only round-shaped objects such as nuclei are left – the other cells without nuclei are eliminated,
- Generation and selection of the geometrical descriptors treated as the numerical features characterizing the cells,
- Fuhrman grade estimation using the neural network,
- Saving the results to the data base.

## Segmentation algorithm

The input data in our experiment is a microscopic digitized kidney image such as that presented in Fig. 1. It is

the image of the biopsy of the tissue at the magnification equal 100. The image are saved in the form of the bitmap file for further processing. The dark objects of the image containing very small oval, dark objects called nuclei are of our interest. The cytoplasm of the cell is clear not visible at H&E staining. The nuclei cells represent the CCRCC and should be extracted from the image of the biopsy.

The first step of processing is the extraction of the dark structures resembling cells through the segmentation of the whole image. This is done by using the morphological operations [4,6,8]. The problem of finding and extracting the real nuclei is not simple, since there are objects very similar to these cells, but not possessing the nucleoli. The nucleoli are of circular shape, hence we recognize them by using circular Hough transform [4]. In further processing only the cells with one or more nucleoli are considered.

After segmentation each nucleus (further called cell) forms the new individual image which should be preprocessed in order to get characterization of it by the proper geometrical descriptors. The examples of the extracted cells with detected nucleoli (the circular dark object inside the nucleus) are depicted in Fig 2.

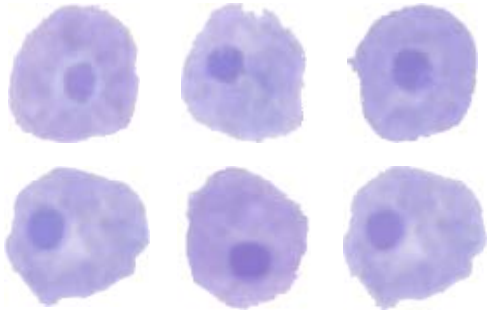

Fig. 2. The examples of extracted nuclei from the images

**The computation of geometrical descriptors**

The extracted cells (nuclei) are subject to the numerical characterization by using the information contained in their geometry. The following geometrical descriptors of nucleus are of interest for us at the estimation of Fuhrman grading:

- Area – real area of (in pixels)
- Radius – short radius and long radius of nucleus
- Factor of circularity (Foc) – ratio of short to long radius
- Real perimeter (in pixels)
- Convex perimeter (in pixels)
- Convex area (in pixels).

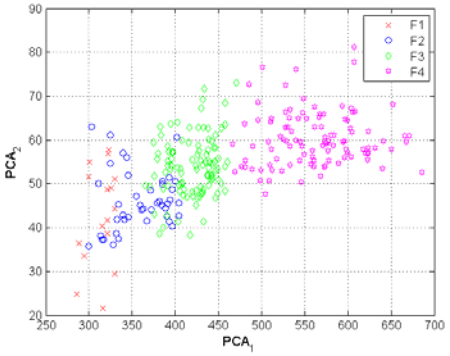


Fig. 3. The distribution of the measured geometrical parameters of the cells mapped into 2D coordinate system using PCA

On the basis of the extracted cells we are able to calculate the mean values and standard deviations of the parameters of all extracted cells, independently for the considered classes. The quantitative analysis was carried out on the collection of 15 CCRCC cases acquired from the archive of the Pathomorphology Dept., the Military Institute of Medicine, Warsaw, Poland. Three analyzed cases represented degree one of Fuhrman grade and the degrees 2, 3 and 4 have been represented by four cases each. 262 nuclei cell images have been extracted from these slides and examined by us.

Fig. 3 presents the distribution of the measured set of parameters mapped into 2-dimensional coordinate system by using Principal Component Analysis (PCA). F1, F2, F3 and F4 denote the classes of Fuhrman grade.

Four clusters have been created. As it is seen we can observe quite good separation of data of class F4 from the rest. However the other classes interlace each other. It means some difficulties in their automatic recognition.

**Feature selection**

Table 1 presents the mean values and standard deviations of the geometrical parameters corresponding to the set of cells that have been found in these 15 slides of CCRCC tissues. The following columns in the table denote: A – real area, LR, SR – long and short radius of cell, FOC - factor of circularity, P – perimeter, CP – convex perimeter, CA – area of convex perimeter, F1...F4 – Fuhrman grade. All parameter values, except FOC, are given in pixels.

Table 1. The mean values and standard deviations of the geometrical features of the extracted cells

| Class | | A | LR | SR | FOC | P | CP | CA |
|---|---|---|---|---|---|---|---|---|
| F1 | mean | 438.5 | 29.3 | 25.0 | 0.85 | 136.3 | 113.6 | 507.8 |
| | std | 22.1 | 2.6 | 3.2 | 0.09 | 18.8 | 8.1 | 37.03 |
| F2 | mean | 501.9 | 32.8 | 28.6 | 0.87 | 154.9 | 132.7 | 577.8 |
| | std | 47.4 | 3.7 | 3.4678 | 0.04 | 11.1 | 14.2 | 39.3 |
| F3 | mean | 590.7 | 34.9 | 29.3 | 0.84 | 168.0 | 141.7 | 708.6 |
| | std | 36.6 | 2.7 | 3.1 | 0.07 | 17.7 | 17.2 | 64.5 |
| F4 | mean | 793.9 | 40.4 | 32.8 | 0.81 | 178.2 | 154.6 | 896.7 |
| | std | 74.2 | 5.2 | 4.6 | 0.05 | 20.2 | 20.4 | 76.6 |

The next step of processing is to select the most important descriptors on the basis of the measured data [5]. In these investigations we have applied the Fisher criterion [3,6]. The importance of the feature *f* is measured on the basis of the so called discrimination coefficient $S_{AB}(f)$. For two classes A and B the discrimination coefficient of the feature *f* is defined as follows

$$S_{AB}(f) = \frac{|c_A(f) - c_B(f)|}{\sigma_A(f) + \sigma_B(f)} \quad (1)$$

In this definition $c_A$ and $c_B$ are the mean values of the feature *f* in the class *A* and *B*, respectively. The variables $\sigma_A$ and $\sigma_B$ represent the standard deviations determined for both classes. The large value of $S_{AB}(f)$ indicates good potential separation ability of the feature *f* for these two classes. On the other side small value of it means that this particular feature is not good for the recognition between classes A and B.

Table 2 presents the actual values of these Fisher measure coefficients of the features, estimated for all combinations of the classes.

Table 2. The values of Fisher discrimination measure for all features and combinations of classes

| | $S_{12}$ | $S_{13}$ | $S_{14}$ | $S_{23}$ | $S_{24}$ | $S_{34}$ |
|---|---|---|---|---|---|---|
| **A** | 0.91 | 2.59 | 3.68 | 1.05 | 2.40 | 1.83 |
| **LR** | 0.54 | 1.04 | 1.39 | 0.32 | 0.83 | 0.67 |
| **SR** | 0.54 | 0.68 | 0.99 | 0.10 | 0.51 | 0.44 |
| **FOC** | 0.12 | 0.08 | 0.27 | 0.27 | 0.60 | 0.21 |
| **P** | 0.62 | 0.86 | 1.07 | 0.45 | 0.74 | 0.26 |
| **CP** | 0.84 | 1.10 | 1.43 | 0.28 | 0.63 | 0.34 |
| **CA** | 0.91 | 1.97 | 3.42 | 1.25 | 2.75 | 1.33 |

As it is seen the values of $S_{AB}(f)$ are changing from the minimal level equal 0.08 (lack of recognition ability of parameter FOC to recognize classes 1 and 3) to the maximal level of 3.69 (the highest recognition ability of A at

recognition of classes 1 and 4). On the basis of these experiments we have decided to exclude the parameter FOC from further considerations, since its Fisher discrimination measure for all cases is very low.

**Results of classification**

The numerical experiments of classification of cells to the proper class of Fuhrman degree have been performed using the whole database of the microscopic kidney images acquired in the Department of Pathomorphology.

Due to small number of data we have applied the leave-one-out methodology. The Support Vector Machine (SVM) of Gaussian kernel has been applied as a classifier [7]. The statistical results of cell recognition of the testing data not taking part in learning in the form of confusion matrix are presented in Table 3.

Table 3. Confusion matrix of the Fuhrman grade recognition

| | class1 | class 2 | class 3 | class 4 |
|---|---|---|---|---|
| class 1 | 14 | 3 | 0 | 0 |
| class 2 | 4 | 32 | 4 | 0 |
| class 3 | 0 | 3 | 98 | 0 |
| class 4 | 0 | 0 | 0 | 104 |

The results of class recognition have shown that the average discrepancy rate between the score of our system and the human expert results for all examined images is 5.72% and this accuracy is acceptable in the medical practice. To get more reliable results in the future, especially for rare class, we have to increase the number of cells taking part in experiments.

Further work in this subject will be directed to increase the number of images (patients) to represent more uniformly all investigated classes (better balance among different classes), developing more advanced and precise segmentation process of cells, developing some additional features describing nuclei as well as application of the ensemble of classifiers to improve the final step of classification.

**Conclusions**

The paper has presented and discussed the problems concerning the automatic recognition of the Fuhrman grade in clear-cell renal carcinoma. On the basis of the performed experiments we can declare that the proposed computer system is able to recognize the nuclei with the accuracy comparable to the human expert. Actually it is possible to get 5%-10% of discrepancy of both results.

The presented results are only the first step in the research leading to the full automatization of the process of estimation of Fuhrman degree of CCRCC. However, even at this step of research there is a visible correlation between the extracted parameters and the Fuhrman grade.

REFERENCES


[1] Fuhrman S.A., Lasky L.C., Limas C., Prognostic significance of morphologic parameters in renal cell carcinoma. *Am J Surg Pathol*., vol. 6, pp. 655-663

[2] Kontak J.A., Campbell S.C., Prognostic factors in renal cell carcinoma. The Urologic clinics of North America. 2003, vol. 30, pp. 467-480

[3] Duda, R.O., Hart, P.E., Stork, P., Pattern classification and scene analysis, 2003, Wiley, New York

[4] Gonzalez R., Woods R., Digital image processing, 2008, Prentice Hall, New Jersey

[5] Jakubowski J., Ocena możliwości wykorzystania deskryptorów cech lokalnych obrazu twarzy w zadaniu automatycznej identyfikacji osób. *Przegląd Elektrotechniczny*, 2012, vol. 88, pp. 217-221

[6] Kruk M., Osowski S., Koktysz R., Recognition and classification of colon cells applying the ensemble of classifiers. *Computers in Biology and Medicine*, 2009, vol. 39, pp. 156-165

[7] Osowski S., Sieci neuronowe do przetwarzania informacji, 2006, Oficyna Wydawnicza PW

[8] Osowski S., Siroic R., Markiewicz T., et al., Application of Support Vector Machine and Genetic Algorithm for Improved Blood Cell Recognition. *IEEE Trans. on IM*, 2009, vol. 58, pp. 2159-2168

[9] http://www.pathpedia.com/Education/eAtlas/Histopathology/kidney/renal_cell_carcinoma_clear_cell_type.aspx



***Authors:*** *dr. Michał Kruk, Warsaw University of Life Sciences, Warsaw, ul. Nowoursynowska 166, krukm@iem.pw.edu.pl*

*prof. Stanisław Osowski, Warsaw University of Technology, Warsaw, ul. Koszykowa 75, and Military University of Technology, Warsaw, ul. Kaliskiego 2*

*prof. Wojciech Kozłowski, dr Robert Koktysz, prof. Janina Słodkowska, Military Institute of Medicine, Warsaw, ul. Szaserów 128*

*prof. Tomasz Markiewicz, Warsaw University of Technology, Warsaw, ul. Koszykowa 75, and Military Institute of Medicine, Warsaw, ul. Szaserów 128*