PubMed

Journal of Pathology Informatics

Publication of the Association for Pathology Informatics

www.jpathinformatics.org

Symposium - Original Research

TMARKER: A free software toolkit for histopathological cell counting and staining estimation

Peter J. Schüffler[1,2], Thomas J. Fuchs[3], Cheng Soon Ong[4], Peter J. Wild[5], Niels J. Rupp[5], Joachim M. Buhmann[1,2]

[1]Institute for Computational Science, ETH Zurich, Switzerland, [3]California Institute of Technology, Pasadena, CA 91125, USA, [4]NICTA, The University of Melbourne, Parkville VIC 3010, Australia, [5]Institute of Pathology, University Hospital Zurich, Switzerland, [2]Competence Center for Systems Physiology and Metabolic Diseases, ETH Zurich

E-mail: *Peter J. Schüffler - peter.schueffler@inf.ethz.ch
*Corresponding author

Abstract

Background: Histological tissue analysis often involves manual cell counting and staining estimation of cancerous cells. These assessments are extremely time consuming, highly subjective and prone to error, since immunohistochemically stained cancer tissues usually show high variability in cell sizes, morphological structures and staining quality. To facilitate reproducible analysis in clinical practice as well as for cancer research, objective computer assisted staining estimation is highly desirable. **Methods:** We employ machine learning algorithms as randomized decision trees and support vector machines for nucleus detection and classification. Superpixels as segmentation over the tissue image are classified into foreground and background and thereafter into malignant and benign, learning from the user's feedback. As a fast alternative without nucleus classification, the existing color deconvolution method is incorporated. **Results:** Our program TMARKER connects already available workflows for computational pathology and immunohistochemical tissue rating with modern active learning algorithms from machine learning and computer vision. On a test dataset of human renal clear cell carcinoma and prostate carcinoma, the performance of the used algorithms is equivalent to two independent pathologists for nucleus detection and classification. **Conclusion:** We present a novel, free and operating system independent software package for computational cell counting and staining estimation, supporting IHC stained tissue analysis in clinic and for research. Proprietary toolboxes for similar tasks are expensive, bound to specific commercial hardware (e.g. a microscope) and mostly not quantitatively validated in terms of performance and reproducibility. We are confident that the presented software package will proof valuable for the scientific community and we anticipate a broader application domain due to the possibility to interactively learn models for new image types.

Key words: Color deconvolution, pathology, nuclei detection, superpixel classification, segmentation staining estimation
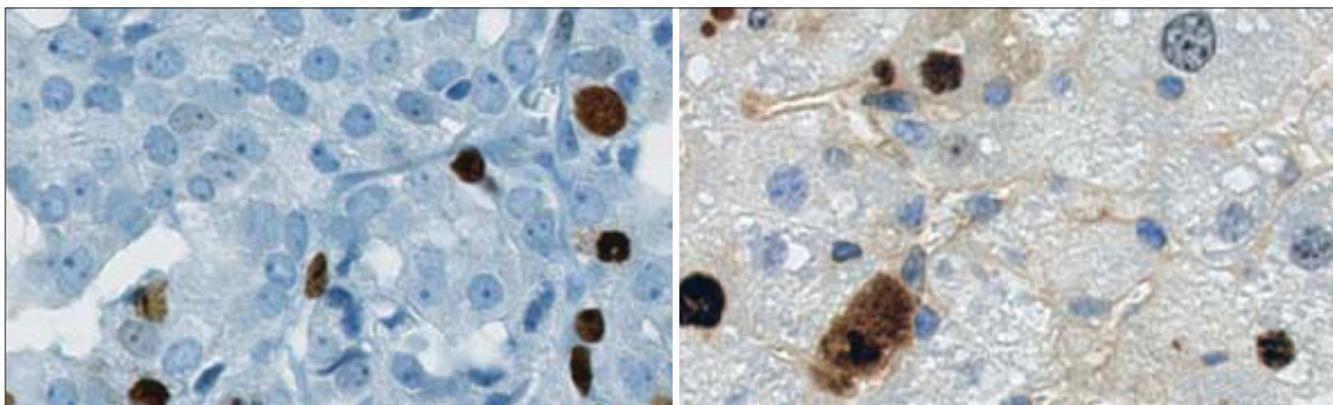
Figure 1: Two parts of typical IHC stained tissue images on which cell nuclei are to be counted and staining percentage is to be estimated. Blue spots are unstained nuclei, brown spots are stained nuclei. For pathology, only cancer cells are relevant for counting

INTRODUCTION

Accurate immunohistochemical (IHC) staining estimation of human tissue plays a crucial role in various kinds of clinical applications and medical research. One example is the estimation of MIB-1 expression for assessing the proliferation factor of renal clear cell carcinoma or prostate cancer from a biopsy, where pathologists examine morphological and IHC characteristics of a given tissue. These characteristics can be observed by staining thin-tissue slices attached on a glass plate with a protein specific antibody that is linked to a dye. Thus, cell nuclei in the tissue layer that express this protein will exhibit, for example, a dark brown color. Other cells nuclei lose the antibody in the subsequent washing procedure. The whole tissue slice is further unspecifically colored with hematoxylin to reveal morphological structures and cell nuclei.

In medical research, IHC estimation is often applied on tissue microarrays (TMA). Their enormous advantage is the ability to simultaneously stain many samples on one microarray, such that the experimental settings do not change among the samples. The scientific goal is to discriminate groups of samples on the TMA with differing protein expression patterns, which later serve as specific biomarker signatures for the respective patients' groups.

Pathologists generally perform IHC estimation by eye with light microscopy or high resolution scans [Figure 1]. For a typical image, two relevant questions for these types of problems include, (i) are there malignant or benign cells present in the image and, if yes, then where?, and (ii) how many nuclei of the cancer cells express the considered protein?

These questions can be transferred to the field of medical image processing and computer vision, which is an emerging field for various kinds of cancer.[1] Note, that there is a difference between malignant or benign
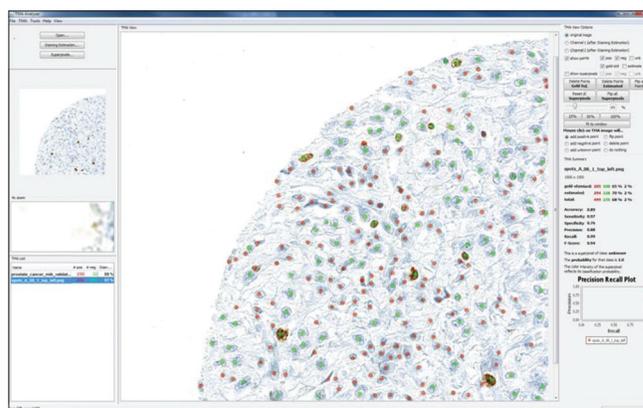

Figure 2: Screenshot of TMARKER with a TMA image of MIB-1 (ki-67) stained renal clear cell carcinoma. Detected cancerous and benign nuclei are marked with red and green points, respectively

cells and stained or unstained cells. The staining of the nuclei only reveals presence of protein. The distinction of malignant and healthy cells is considerably more difficult and relies on parameters like size, shape, and morphology of the cell nuclei but not necessarily on the IHC staining.[2] Stained benign cells (e.g., epithelial cells) should not be considered for staining estimation.

Motivation

High-throughput IHC staining estimation of tissue images poses several challenges in practice. By nature, IHC stained images are much more difficult to analyze than, for example, immunofluorescence (IF) images, where cell nuclei can be easily separated from the comparatively homogeneous background. Furthermore, IF images show almost no morphological structures in the background such as disrupted cells, cell compartments, and extracellular matrix, which might disturb the perception of a nucleus. Manual grading of IHC images is highly subjective, and reproducible cell counts is rare, even by trained human experts.[3] Furthermore, manual assessment of TMAs is very time consuming, prone to errors, and expensive. To this end, several approaches

have been presented to assist or automate the whole estimation process,[2,4] with promising results in nuclei detection and classification. However, while there exist a number of general purpose medical imaging toolkits and commercial solutions, to the best of our knowledge, no software package especially tailored to nuclei counting and nuclear IHC staining estimation of human tissue is available publicly and free of charge for the research community.

Own Contribution

To provide medical researchers and clinical pathologists with a software package that can alleviate the aforementioned tasks, we have presented TMARKER, a user-friendly, freely available and platform-independent toolkit to assist in cell nuclei counting and staining estimation of IHC-stained tissue slices and TMAs [Figure 2]. The software aims to fit the following needs:

- Semiautomatic, reproducible, and fast cell nuclei detection and counting in a given set of IHC stained images
- Automatic classification of cells into malignant and benign (on the basis of the frameworks presented in previous studies[2,4])
- Platform independence, free availability, user-friendly Java-webstart GUI.

The program was implemented in Java v1.6 and can be executed on any client with a Java virtual machine. It is publicly available at http://www.comp-path.inf.ethz.ch.

In Section 2, we have reported algorithmic details as well as implementation details. In Section 3, we have described the results from several real data experiments, which demonstrated a performance comparable to that by trained pathologists. Finally, we have discussed the utility of the software and the achieved results and have given prospects to further developments in Section 3.

MATERIALS AND METHODS

Superpixels and Active Learning-based Nucleus Detection and Classification

Superpixels

In the scope of human tissue images, superpixels can be used to segment cell nuclei as well as other structural compartments. For this purpose, the size of superpixels should roughly cover a typical size of a nucleus. TMARKER finds the correct size and number of superpixels by $n = (w \times h)/(4 \times r^2)$, where w and h are image width and height and r is the nucleus radius. For superpixel formation, we use an adapted Java implementation of the simple linear iterative clustering (SLIC) superpixel algorithm as introduced earlier.[5]

Nucleus Detection

The whole tissue image is first partitioned into superpixels.

The superpixels are then considered as the samples for training a classifier. From each superpixel, a feature vector based on the underlying image properties is calculated. We implemented three feature extractors, which proved valuable for histology in the past: Color histograms (3×16 bins), local binary patterns (LBP)[6] (size 256), and pyramid histograms of oriented gradients (PHOG)[7] (size 338). These features were concatenated resulting in a feature vector of size 642. The default classifier in TMARKER is a random forest[8] (WEKA package[9]), although support for support vector machines (SVM)[10] and Bayesian Networks[11] are implemented. Based on the labels provided by the pathologist, the classifier learns to discriminate between superpixels, which represent a nucleus (foreground) and superpixels belonging to the background. The foreground superpixels are subjected to nucleus classification.

Nucleus Classification

After the detection of the cell nuclei, the goal is to classify them into malignant and benign. To this end, we used the same feature vector as before, but the classifier was now trained only on cell nuclei labeled by the domain expert. The superpixels corresponding to detected nuclei from the step before were hence classified into malignant and benign. These classifications were visualized on the histological image so that the pathologist could correct and retrain the classifier [Figure 3].

Active Learning

Active learning describes a learning method in which the learner is able to choose the (most informative) training samples.[12] TMARKER provides a simple active learning algorithm, which we called "semiautomatic labeling." The user starts labeling the image with positive (i.e., cancerous) and negative (i.e., benign) nuclei. Alternatively, the background of the image can also be labeled in addition. With each label provided by the user, TMARKER retrains the classifier and updates the visualization of the classification results for all superpixels. By iterating this process, the user improves the classification results by continually labeling the image. For active learning, the user labels superpixels with low classification confidence (high uncertainty) and thus improving the discriminative classifier at the decision boundaries. The algorithm can be trained over several images to cover the larger variance among different specimens. Once trained, the classifier can be saved and applied to any new images in a high-throughput manner.

Validation

Within the framework, it is important to distinguish between gold-standard nuclei (GS), which were labeled and classified by the user, and estimated nuclei (ES) identified by the system. To evaluate the performance of the presented algorithms, we calculated the match statistics between GS and ES. Two points with the
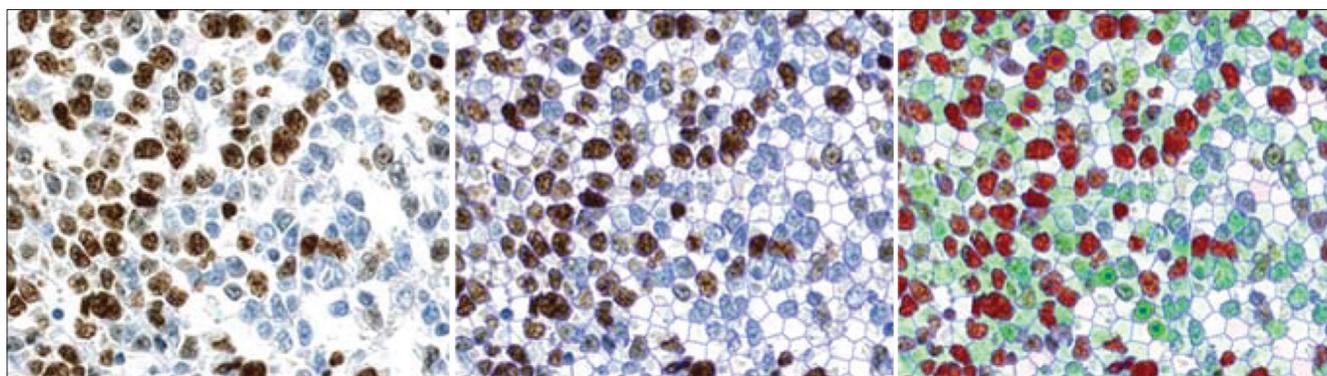
Figure 3: Superpixel algorithm. Left: Part of the original image. Middle: The image is segmented into superpixels. Right: Superpixels are classified into red and green superpixels (positive and negative) as training set serve the labels of the user (red and green circles). The color intensity reflects the classification probability
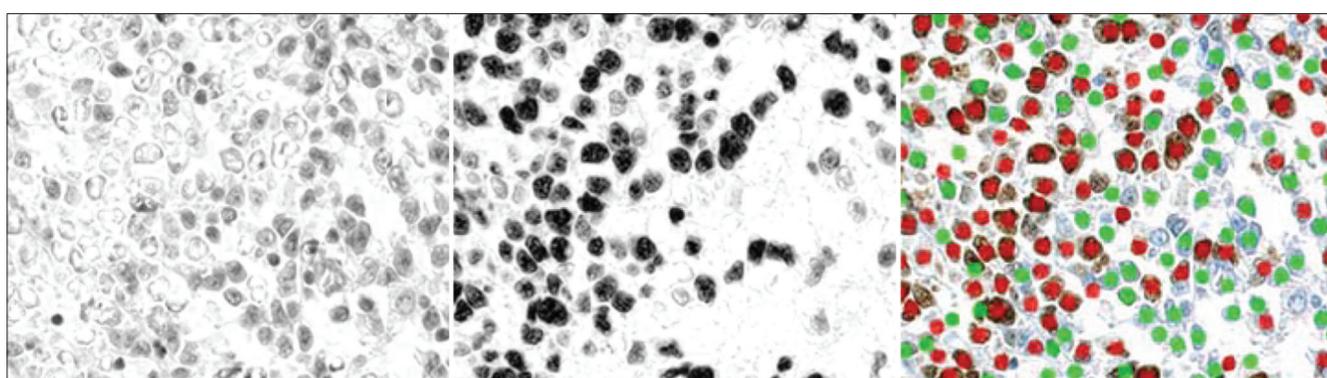

Figure 4: Color deconvolution of the image in Figure 3. Left: The hematoxylin channel image. Middle: The DAB channel image. Right: Found nuclei based on the intensities on the two channels and the nucleus radius *r*

distance d were matched to each other, if $d \leq 2r$, where r is the nucleus radius. Based on this distance, precision, recall, and F−Score were measured. Subsequently, the detected nuclei were classified into malignant and benign and sensitivity, specificity, and overall classification accuracy were calculated.

Color Deconvolution-Based Nucleus Detection

In cases where staining estimation is performed without nucleus classification (i.e., the cells are homogeneous on the image), color deconvolution provides a fast alternative to the superpixel approach. Color deconvolution enables nuclei detection based on the method presented earlier.[13] The image is deconvolved into separate color channels (e.g., hematoxylin channel and DAB channel), which are smoothed with a Gaussian blur filter and subsequently screened for local intensity maxima [Figure 4]. These steps are performed with ImageJ for Java.[14]

Few parameters are needed for local maxima detection: The radius r of cell nuclei describing the size of the local environment and an intensity threshold t per channel, above which a local maximum is accepted. Since these parameters vary between experimental protocols for

the preparation of tissue, TMARKER provides visual assistance to select the parameter values. Interactively, the user gets immediate feedback on changes to the parameters.

RESULTS AND DISCUSSION

Superpixels and Active Learning

We have presented a new software toolkit called TMARKER that is suitable for nucleus classification (malignant/benign and stained/unstained). TMARKER uses a superpixel-based approach for classification. We have shown in Figure 5 (left) that superpixels are suitable for image segmentation and classification of histopathological images. The quantitative detection accuracy of 92% and classification accuracy of 64% touches the range of the inter-pathologists error (97% and 74% on the same dataset, respectively) even at such a difficult problem. This holds true for sensitivity and specificity as well [Figure 5]. Thus, detection and classification of cell nuclei were comparable to those by pathologists. Moreover, we have shown in Figure 5 (right) that an active learning approach profits from user input, especially in borderline cases. The classification accuracy
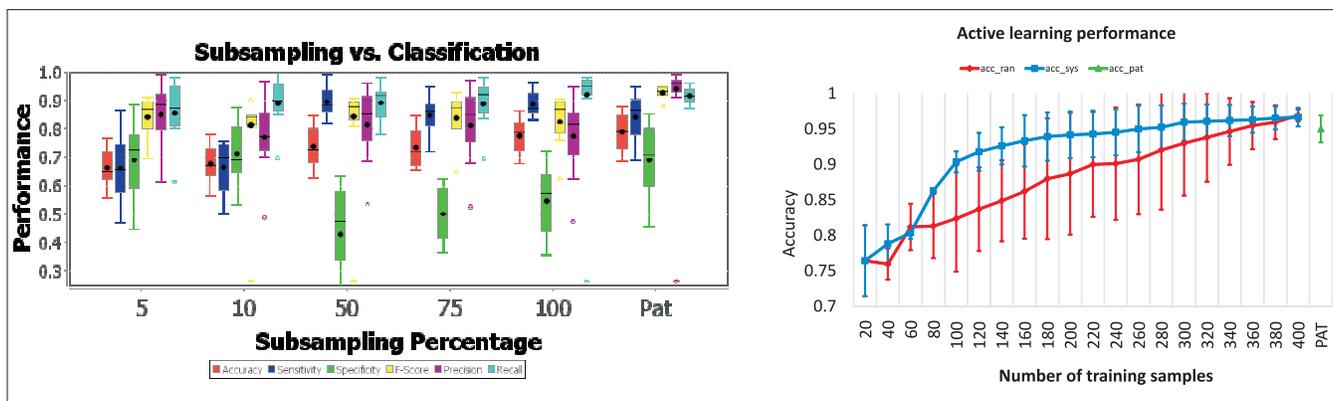
Figure 5: Left: Performance plot for nucleus detection and classification via superpixels. Depicted are precision, recall, and F-score for the nucleus detection as well as sensitivity, specificity, and accuracy for the nucleus classification. Experiments were conducted with training set sizes from 5% to 100% (X-axis) of all nuclei in eight fully labeled TMA spots. Each box represents a leave-one-image-out cross validation run with a SVM (polynomial kernel). The performance stabilizes with 15% of training samples. The inter-expert performances of two pathologists is plotted last ("Pat"). For each of the eight images, pathologist A is taken as reference for pathologist's B guesses. Right: Proof of concept for the active learning approach in TMARKER. For three given TMA images, initially 10 malignant and 10 benign nuclei were selected to train an SVM. The classification result on all nuclei is shown as accuracy on the Y-axis. Consecutively, 20 additional nuclei were added repeatedly to the training (X-axis), thereby improving the classification performance. The additional nuclei are chosen at random ("acc ran") or systematically according to the respective lowest classification score ("acc sys"). The systematic approach saturates much faster. The classification accuracy reaches the level of the two pathologists ("acc pat")
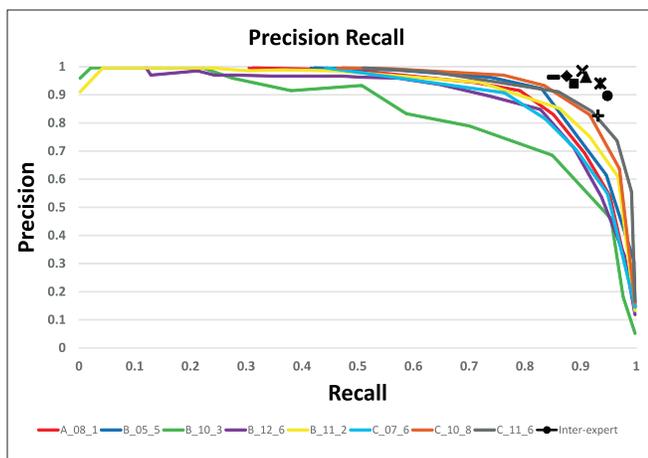

Figure 6: Precision/Recall curve for the nucleus detection via color deconvolution with varying radius r. Each curve represents one out of 8 images of TMA spots. Higher radius r reveals fewer nuclei. The automatically detected nuclei were validated against the labels of a trained pathologist. For all images, the inter-expert values are pointed (top right) as the precision and recall of one pathologist to match the other

saturates already after 160 labels with systematic user input, instead of 360 user interactions with random labeling.

Color Deconvolution

The nucleus detection and staining estimation with color deconvolution provides facilitated parameter settings, but no classical machine learning influence. If the pathology goal is not dependent on nucleus type, or the nucleus types are known for the given image set, this method is a fast alternative to the more comprehensive classification. As shown in Figure 6, TMARKER achieves a reproducible

precision and recall in nucleus detection. The performance hereby is still comparable to those of two pathologists, as measured by their inter-precision and inter-recall.

TMARKER is free software with high potential in cell counting and staining estimation of pathological IHC-stained tissue images. A major advantage of TMARKER is the high reproducibility of competitive cell counts. A fast way for staining estimation is provided by the integrated color deconvolution method. When only relevant cells are considered for staining estimation, e.g., with distinction between malignant and benign cells, TMARKER provides modern machine learning methods for nucleus detection and classification. While the potential of TMARKER has been shown, it has to be further validated and improved on larger and different datasets.

REFERENCES

1. Madabhushi A, Dowling J, Huisman H, Barratt D. Prostate cancer imaging. Image analysis and image-guided interventions: International Workshop, Held in Conjunction with MICCAI, Toronto, Canada, Proceedings. Lecture Notes in Computer Science. September 22, Springer; 2011.
2. Schüffler PJ, Fuchs TJ, Ong CS, Roth V, Buhmann JM. Computational TMA analysis and cell nucleus classification of renal cell carcinoma. In: Goesele M, Roth S, Kuijper A, Schiele B, Schindler K, editors. Proceedings of the 32nd DAGM conference on Pattern recognition, Berlin, Heidelberg, Springer-Verlag; 2010. p. 202-11. Available from: http://link.springer.com/chapter/10.1007/978-3-642-15986-2_21 [Last accessed on 21. Jan 2013]

3. Fuchs TJ, Buhmann JM. Computational pathology: Challenges and promises for tissue analysis. Comput Med Imaging Graph 2011;35:515-30.
4. Fuchs TJ, Wild PJ, Moch H, Buhmann JM. Computational pathology analysis of tissue microarrays predicts survival of renal clear cell carcinoma patients. Med Image Comput Comput Assist Interv 2008;11(Pt 2):1-8.
5. Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Suesstrunk S. SLIC Superpixels. Technical report, EPFL; 2010.
6. Ahonen T, Hadid A, Pietikainen M. Face recognition with local binary patterns. In: Pajdla T, Mata J, editors. ECCV 2004. LNCS, Vol. 3021. Heidelberg: Springer; 2004. p. 469-81.
7. Bosch A, Zisserman A, Munoz X. Representing shape with a spatial pyramid kernel. In: Proceedings of the 6th ACM international conference on image and video retrieval. CIVR'07, New York, USA, ACM; 2007. p. 401-8.
8. Breiman L. Random forests. Mach Learn 2001;45:5-32.
9. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The weka data mining software: An update. SIGKDD Explor Newsl 2009;11:10-8.
10. El-Manzalawy Y, Honavar V. WLSVM: Integrating LibSVM into Weka Environment 2005. Software Available from: http://www.cs.iastate.edu/~yasser/wlsvm. [Last accessed on 21 Jan 2013]
11. Friedman N, Geiger D, Goldszmidt M. Bayesian network classifiers. Mach learn 1997;29:131-63.
12. Cohn DA, Ghahramani Z, Jordan MI. Active learning with statistical models. J Artif Intell Res 1996;4:129-45.
13. Ruifrok AC, Johnston DA. Quantification of histochemical staining by color deconvolution. Anal Quantitative Cytol Histol 2001;23:291-9.
14. Burger W, Burge MJ. Digital Image Processing: An Algorithmic Introduction using Java J. 2., berarbeitete Auflage Edn. Berlin, Heidelberg: Springer-Verlag; 2005.