image processing, immunohistochemistry,
hyperparathyroidism, light microscopy

Bartłomiej PŁACZEK[1], Tomasz ORCZYK[1], Rafał BUŁDAK[2,3], Marek MICHALSKI[4], Oliwia SEGIET[4], Renata POLANIAK[3]

IMAGE ANALYSIS OF IMMUNOHISTOCHEMICAL STAINS FOR DETECTION OF PARATHYROID DISEASE

In this paper a method is introduced which enables automatic detection of parathyroid hyperplasia and parathyroid adenoma on the basis of immunohistochemical angiogenesis markers expression in micrographs. The proposed method uses digital image processing techniques and classification algorithms to detect diseased tissue. The disease detection is performed by classification of normalized color intensity histograms. Accuracy of this method was evaluated by using micrographs of parathyroid tissue sections obtained from patients that have undertaken surgery due to primary hyperparathyroidism. Use of different color models, various classifiers, and immunohistochemical markers was considered during the experiments. The experimental results show that the introduced method enables accurate detection of parathyroid disease. The most promising results were obtained for k-nearest neighbor and neural network classifiers.

1. INTRODUCTION

Immunohistochemistry (IHC) is a powerful technique based on antigen–antibody interaction, which uses selected antibodies as markers for localization of target antigens in tissues. IHC staining enables visual analysis of the localization and distribution of specific cellular components within cells in proper tissue context. When using the IHC method, the antigen–antibody complexes are visualized in a color image by means of light microscopy. Morphology visualization of the tissue around the specific antigen is obtained by counterstaining. Results of stained IHC markers have important implications for disease diagnosis, drug development and biological research [13]. IHS is widely used in the diagnosis of abnormal cells such as those found in tumors [6].

The use of IHC in pathomorphological diagnosis brought a substantial methodological problem, related with an evaluation of the amount and strength of specific staining reaction [5]. The common practice is visual scoring of the immunostains by expert observers. However,

[1]Institute of Computer Science, University of Silesia, Będzińska 39, 41-200 Sosnowiec, Poland,
e-mail: placzek.bartlomiej@gmail.com, tomasz.orczyk@us.edu.pl
[2]Department of Physiology, School of Medicine with the Division of Dentistry, Medical University of Silesia,
e-mail: rbuldak@sum.edu.pl
[3]Department of Human Nutrition, School of Public Health, Medical University of Silesia,
e-mail: rpolaniak@sum.edu.pl
[4]Department of Histology and Embryology, School of Medicine with the Division of Dentistry, Medical University of Silesia, e-mail: mmichalski@sum.edu.pl

visual scoring is time consuming and subjective. The interpretation of visual scoring causes inconsistencies upon the evaluation process. Computer-assisted IHC staining shortens analysis time and reduces inter-observer variation when the staining intensity levels are evaluated [10], [4].

Digital image processing techniques [7], [8] can be used to overcome the issues of visual scoring. Currently, several image processing approaches have been proposed to automatically extract stained regions and evaluate the IHC stain intensity. Some early computer-assisted methods have involved transforming color images to grey scale and thresholding segmentation [5]. Other methods perform color analysis for IHC quantification by using components of the RGB color space [10], their normalization (nRGB) [2], HSL model [5], and YUV color space [3].

Advantages of CMYK color space for reproducible and unbiased evaluation of IHC stain intensity were discussed in [9]. It was demonstrated that the image analysis method based on yellow color channel (Y) is independent of observer biases for threshold and positive color selection, applicable to different markers, tolerant of counterstaining, and sensitive to small changes in the IHC stain intensity.

Recently, a considerable attention has been paid to color deconvolution approach [12], [17]. According to this approach, the RGB matrix of the input image is multiplied by the inverse of a color spread matrix (color deconvolution matrix) to separate the stain in IHC micrographs. However color deconvolution requires special and precise calibration and determination of pure stain color spectra, which can be impractical [16].

In this paper an image analysis method is proposed which can be applied to micrographs of IHC stained tissue for detection of parathyroid disease. This approach was motivated by the results of previous work by Segiet et al. [14] that suggests usefulness of IHC in diagnosis of primary hyperparathyroidism. According to those results, IHC expression of angiogenesis markers (VEGF, CD31 and CD106) is increased in parathyroid adenomas compared with parathyroid hyperplasias, and higher in hyperplasias compared with healthy parathyroid glands. Therefore, the IHC angiogenesis markers can be useful in distinguishing between parathyroid hyperplasia and neoplasia.

According to the proposed method, the techniques based on nRGB and CMYK color spaces are applied to evaluate expression intensity of the IHC angiogenesis markers in micrographs. Histograms of the IHC stain intensity are then determined and used as input data to a classifier which detects the parathyroid disease.

The paper is organized as follows. Details of the proposed parathyroid disease detection method are presented in Section 2. Section 3 includes results of the experimental evaluation. Finally, conclusions and future research directions are discussed in Section 4.

2. PROPOSED METHOD

In this section the method is presented which enables detection of parathyroid disease by means of image analysis of IHC stained tissue. Block diagram of the proposed image analysis method is shown in Fig. 1. The disease detection is based on classification of image intensity histograms calculated for normalized red (nR) and yellow (Y) color channels. Selection of the analyzed color channels was motivated by the results of previous studies [2], [9]. It should be noted here that high intensities of nR and Y colors correspond with strong expression of the IHC marker in tissue micrographs (Fig. 2).

Preprocessing of the input RGB image includes noise reduction by median filtering and white balance correction to improve the contrast. An example of the image preprocessing results is illustrated in Fig. 2 b. At the next step, the image is converted to nRGB and CMYK
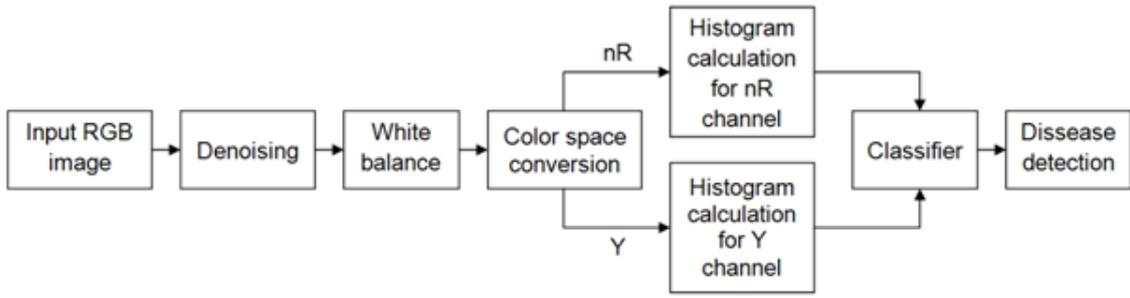
Fig. 1. Block diagram of the proposed image analysis method.
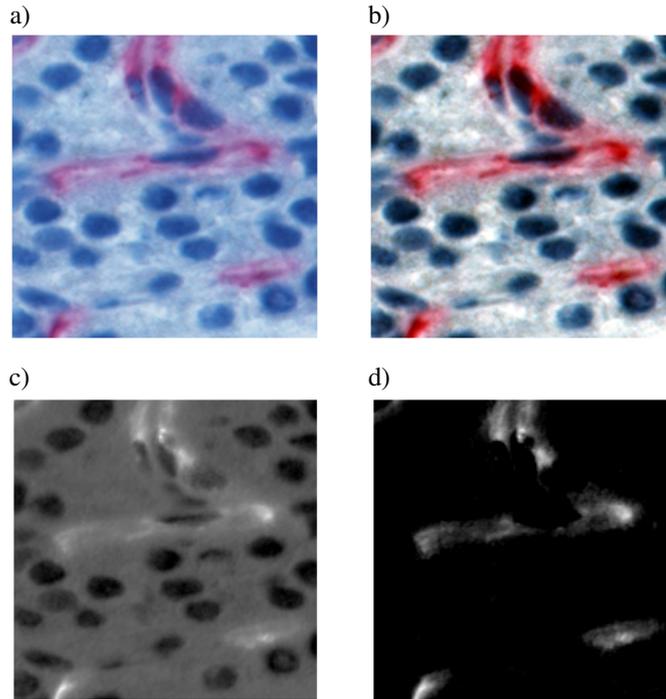
Fig. 2. IHC stained parathyroid tissue: a) input image, b) results of white balance operation, c) intensity of nR channel, d) intensity of Y channel.

color spaces. The nR and Y color channels are extracted for further processing. The following formulas are used to determine the color intensities for nR and Y channels:

$$nR = R/(R + G + B), \tag{1}$$

$$Y = \begin{cases} 1 - R/(1 - \max(R, G, B)), & \max(R, G, B) < 1, \\ 1, & \max(R, G, B) = 1, \end{cases} \tag{2}$$

where the color intensities R, G, B, nR, Y are represented by real numbers between 0 and 1 (including the end points). Figs. 2 c and 2 d show an example of nR and Y color channels extracted from the IHC stained micrograph of parathyroid tissue. After color space conversion, normalized intensity histograms are calculated for nR and Y channels. The histogram bin width was set to 0.2. Thus, each histogram is composed of five real numbers $h_i, i = 1, ..., 5$ that belong to interval $[0, 1]$:

$$h_i = n_{[0.2(i-1), 0.2i]}/n_{[0,1]}, \tag{3}$$

where $n_{[a,b]}$ denotes the number of pixels with intensity values belonging to interval $[a, b]$. The large width of histogram bins allows us to reduce construction complexity of classifiers.

Preliminary experiments have shown that the selected bin width is suitable for the considered task of disease detection.

The normalized image histograms are used as inputs to classification. A classifier is applied in the proposed scheme to detect the diseased tissue and recognize type of the primary hyperparathyroidism disease (adenoma, hyperplasia). In this study, three popular classifiers were implemented: decision tree classifier (DT), k-nearest neighbor classifier (kNN), and multilayer perceptron neural network (MLP).

3. EXPERIMENTAL EVALUATION

The considered task is to recognize healthy tissue, hyperplasia, and adenoma based on the IHC stained micrographs. Experiments were conducted in order to evaluate accuracy of the proposed method for different classifiers and various IHC markers. A set of 225 test images was used for the experimental evaluation. Examples of the test images are shown in Fig. 3. The micrographs of parathyroid tissue were captured by using a light microscope with 100x magnification and Nicon DS-Fi1-L2 color camera head. The parameters of test images are as follows: 600 x 600 pixels size, 24 bit depth, RGB color model.

The parathyroid tissue sections were obtained from patients that have undertaken surgery due to primary hyperparathyroidism caused by adenoma and primary hyperplasia. Frozen tissue sections were incubated with purified mouse monoclonal antihuman antibodies: anti-VEGF (clone SP28), anti-CD31 (clone JC/70A) and anti-CD106 (clone 1.4C3). The dilution of the primary antibodies was 1:500 and was verified in a series of pilot experiments. The sections were counterstained with Mayer's haematoxylin. The aim of the IHC staining procedure was to assess the expression of angiogenesis markers, i.e. VEGF, CD31 and CD106.

For each IHC marker a separate set of images was prepared and divided into training and testing subsets. 60% of the images were used for classifier training and 40% for testing. In case of DT classifier, the SPRINT algorithm [15] was used to build decision trees. For kNN classifier [1] the number of nearest neighbors used to classify a new instance (k) was set to 3. This setting of k parameter was optimized during preliminary experiments. The MLP neural network classifier with single hidden layer was trained by using the RProp algorithm [11].
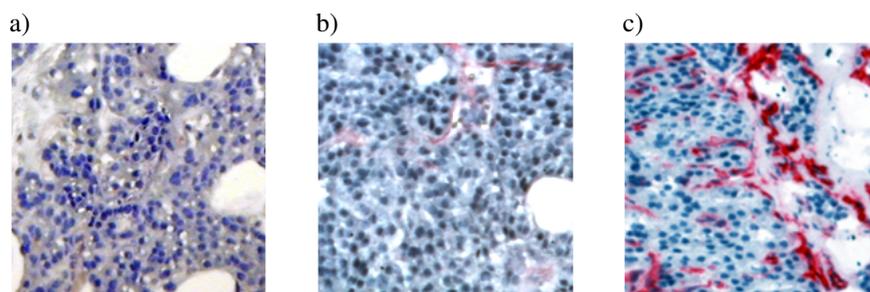
Fig. 3.   Examples of test images (CD31 marker): a) healthy tissue, b) hyperplasia, c) adenoma.

During tests, the classifiers (DT, kNN, and MLP) were applied to categorize the IHC stained images into the 3 classes (healthy, hyperplasia, adenoma). The tests were performed by using 30 images for each IHC marker. Classification was conducted for three sets of the input data: nR histogram, Y histogram, and both histograms (nR + Y). Overall classification accuracy was evaluated (in percents) as the number of correctly classified test images divided by the total number of test images. The results of accuracy assessment are presented in Tab. 1.

The experimental results (Tab. 1) show that it is possible to accurately recognize parathyroid disease by using the proposed method. All test images were correctly categorized (100%

Table 1. Overall accuracy of test image classification.

Classifier input data (histograms)	Classifier type	IHC marker		
		VEGF	CD31	CD106
		Overall accuracy [%]		
nR	DT	97	76	91
nR	kNN	97	88	91
nR	MLP	100	100	97
Y	DT	88	79	79
Y	kNN	100	91	94
Y	MLP	100	94	88
nR+y	DT	97	97	100
nR+y	kNN	100	100	100
nR+y	MLP	100	100	100

accuracy was obtained) for several cases that correspond to different setup of the classifier type, input data set, and IHC marker. In general, the best results were obtained when using MLP and kNN classifiers with input data set containing both histograms. For such settings, each marker enables achieving the maximum classification accuracy. The simplest classifier (DT) provides the highest accuracy only if CD106 marker is applied. Correct classification is also possible when using a single histogram. In case of the input data set which contains only Y histogram, the accuracy of 100% was obtained by using MLP and kNN classifiers for VEGF-stained micrographs. When nR histogram is taken into account then only MLP classifier ensures the correct results. The lowest accuracy was observed when using DT classifier and nR histograms of CD31-stained micrographs.

4. CONCLUSIONS

The proposed method enables detection of parathyroid hyperplasia and parathyroid adenoma by automatic assessment of IHC angiogenesis markers expression in micrographs. The expression of IHC markers is quantified by color intensity analysis. In order to detect the parathyroid diseases, classification of color intensity histograms is performed.

Experimental verification was performed on a set of micrographs of parathyroid tissue sections obtained from patients that have undertaken surgery due to primary hyperparathyroidism. Application of different color models, classifiers, and IHC markers was considered during the experiments. The obtained results reveal a high accuracy of the proposed method. The most promising results were observed for kNN and MLP neural network classifiers.

Further experiments are necessary to verify the proposed method with a larger dataset. Interesting topics for future research include applications of the proposed method to diagnosis of other diseases and its extensions that would enable consideration of generalized color models and recognition of tissue components.

Bibliography

[1] Bernaś M., Płaczek B., Porwik P., Pamuła T. Segmentation of vehicle detector data for improved k-nearest neighbours-based traffic flow prediction. Intelligent Transport Systems, IET, 2015, Vol. 9. pp. 264–274.

[2] Brey E. M., Lalani Z., Johnston C., Wong M., McIntire L. V., Duke P. J., Patrick C. W. Automated selection of DAB-labeled tissue for immunohistochemical quantification. The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society, May 2003, Vol. 51. pp. 575–84.

[3] Dong J., Li J., Fu A., Lv H. Automatic segmentation for ovarian cancer immunohistochemical image based on yuv color space. Biomedical Engineering and Computer Science (ICBECS), 2010 International Conference on, April 2010. pp. 1–4.

[4] FU R., MA X., BIAN Z., MA J. Digital separation of diaminobenzidine-stained tissues via an automatic color-filtering for immunohistochemical quantification. Biomedical optics express, Feb 2015, Vol. 6. pp. 544–58.

[5] KACZMAREK E., GÓRNA A., MAJEWSKI P. Techniques of image analysis for quantitative immunohistochemistry. Roczniki Akademii Medycznej w Białymstoku (1995), 2004, Vol. 49 Suppl 1. pp. 155–8.

[6] MUNSON P. Immunohistochemistry. Basic Science Techniques in Clinical Practice, 2007. Springer London, pp. 18–30.

[7] PŁACZEK B. Rough sets in identification of cellular automata for medical image processing. Journal of Medical Informatics and Technologies, 2013, Vol. 22. pp. 161–168.

[8] PŁACZEK B., BUŁDAK R. J., POLANIAK R. Automatic immunogold particle detection in transmission electron micrographs of cancer cells. Journal of Medical Imaging and Health Informatics, 2015, Vol. 5. pp. 1350–1357.

[9] PHAM N.-A., MORRISON A., SCHWOCK J., AVIEL-RONEN S., IAKOVLEV V., TSAO M.-S., HO J., HEDLEY D. Quantitative image analysis of immunohistochemical stains using a cmyk color model. Diagnostic Pathology, 2007, Vol. 2. BioMed Central.

[10] PRASAD K., PRABHU G. Image analysis tools for evaluation of microscopic views of immunohistochemically stained specimen in medical research–a review. Journal of Medical Systems, 2012, Vol. 36. Springer US, pp. 2621–2631.

[11] RIEDMILLER M., BRAUN H. A direct adaptive method for faster backpropagation learning: the rprop algorithm. Neural Networks, 1993., IEEE International Conference on, 1993. pp. 586–591 vol.1.

[12] RUIFROK A. C., JOHNSTON D. A. Quantification of histochemical staining by color deconvolution. Analytical and quantitative cytology and histology / the International Academy of Cytology [and] American Society of Cytology, Aug 2001, Vol. 23. pp. 291–9.

[13] SCHACHT V., KERN J. S. Basics of immunohistochemistry. J Invest Dermatol, Mar 2015, Vol. 135. The Society for Investigative Dermatology, Inc, p. e30. Research Techniques Made Simple.

[14] SEGIET O., PIECUCH A., MICHALSKI M., DESKA M., BRZOZOWA M., GAWRYCHOWSKI J., WOJNICZ R. Immunohistochemical assessment of angiogenesis markers in primary hyperparathyroidism. 48 Sympozjum Polskiego Towarzystwa Histochemików i Cytochemików "Od MAKRO do NANO. Nowe horyzonty i nowe możliwości w naukach podstawowych i klinicznych", 2014. p. 50.

[15] SHAFER J., AGRAWAL R., MEHTA M. Sprint: A scalable parallel classifier for data mining. 1996. Morgan Kaufmann, pp. 544–555.

[16] TADROUS P. J. Digital stain separation for histological images. Journal of microscopy, Nov 2010, Vol. 240. pp. 164–72.

[17] VARGHESE F., BUKHARI A. B., MALHOTRA R., DE A. IHC Profiler: an open source plugin for the quantitative evaluation and automated scoring of immunohistochemistry images of human tissue samples. PloS one, 2014, Vol. 9. p. e96801.