Color Normalization-based Nuclei Detection in Images of Hematoxylin and Eosin-Stained Multi Organ Tissues

Adam Piórkowski[1][0000−0003−4773−5322]

Department of Biocybernetics and Biomedical Engineering
AGH University of Science and Technology,
A. Mickiewicza 30 Av., 30–059 Cracow, Poland,
pioro@agh.edu.pl

Abstract. This article presents an adaptation of a color transfer method for binarization of cell nuclei in Hematoxylin- and Eosin-stained microscopy images. The aim is to check the ability and accuracy of nuclei detection for multi-organ cases using a public dataset. The results are obtained using the Monte Carlo method and then compared to the ground truth segmentations. Some cases are presented in detail and discussed. This method seems to be promising for the further development of classic and deterministic algorithms for H&E nuclei detection.[1]

Keywords: nuclei detection, nuclei segmentation, cell counting, color transfer, multi-organ tissue, hematoxylin and eosin staining

1 Introduction

Staining with hematoxylin and eosin (H&E) is the most popular method in histopathology. Although this is a very common task, it has still not been fully automated. One of the occurring subproblems is the correct segmentation of cell nuclei.

Classic image processing methods do not always give satisfactory and reproducible results [22]. Many different solutions have been created, although they are limited to the tissues of a given organ [4, 7, 9, 14, 17, 25].

There are no standardized protocols for H&E staining. This results in significant differences in coloration of specimens, which causes problems with the

[1] This is the manuscript of:

Piorkowski, A.: Color Normalization-Based Nuclei Detection in Images of Hematoxylin and Eosin-Stained Multi Organ Tissues. Springer, AISC, vol. 1062, 2020, pp. 57-64.

The final authenticated version is available online at https://doi.org/10.1007/978-3-030-31254-1_8 .

reproducibility of segmentation. Some methods put more emphasis on color normalization in images based on deconvolution of hematoxylin and eosin [5]. Adaptation of different conditions into a unified space enables robust quantitative tissue analysis [6]. Reinhard's [19], Macenko's [13], and Li's [12] color normalization methods are treated as the standard approaches [1,2,11]. In [21] a deconvolution for hematoxylin channel was presented. State-of-the-art color normalization methods can be found in [16].

This article focuses on segmentation and detection of cell nuclei based on the original method and is a continuation of previous works [15,18]. The focus was on the evaluation of a method used for multi-organ tissues, presented in a standardized, public dataset [10]. Further considerations, such as separation of clustered objects [3,8] or the issue of segmentation accuracy [17,20], will be taken into account in further works. Also other color spaces can be considered [23,24].

2 Modified Color Transfer Transformation

The standard color transfer procedure, presented in [19], assumes transformations from RGB space to Lab space using logarithmic operations. The author of the modified method presented in this article decided not to use logarithmic operations because such a simplification increases the sensitivity and precision of the transformations presented below; it is also simpler to implement and much more efficient. In this case, the $LinearLab$ (LL, La, Lb) space can be calculated using combined transformation with matrix multiplication (1).

$$\begin{bmatrix} LL \\ La \\ Lb \end{bmatrix} = \begin{bmatrix} 0.3475 & 0.8265 & 0.5559 \\ 0.2162 & 0.4267 & -0.6411 \\ 0.1304 & -0.1033 & -0.0269 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

Subsequently, the standard color normalization in Lab space is conducted (3) as described in [19], using mean values $(\overline{LL}, \overline{La}, \overline{Lb})$ and standard deviations of channels (σ_s - source and σ_t - target).

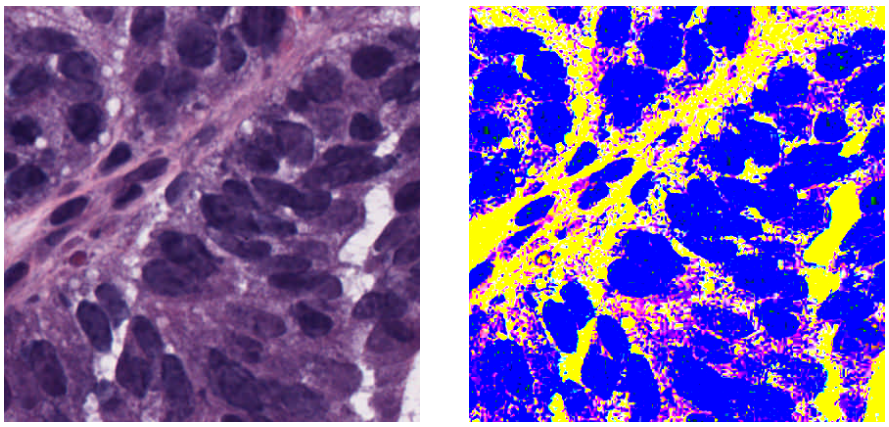
$$\begin{aligned} LL\prime &= \frac{\sigma_t^{LL}}{\sigma_s^{LL}}(LL - \overline{LL}) \\ La\prime &= \frac{\sigma_t^{La}}{\sigma_s^{La}}(La - \overline{La}) \\ Lb\prime &= \frac{\sigma_t^{Lb}}{\sigma_s^{Lb}}(Lb - \overline{Lb}) \end{aligned} \tag{2}$$

Finally, the space transformation back to the new RGB is performed using the combined matrix for multiplication (3).

$$\begin{bmatrix} R\prime \\ G\prime \\ B\prime \end{bmatrix} = \begin{bmatrix} 0.5773 & 0.2621 & 5.6959 \\ 0.5773 & 0.6071 & -2.5452 \\ 0.5833 & -1.0628 & 0.2076 \end{bmatrix} \begin{bmatrix} LL\prime \\ La\prime \\ Lb\prime \end{bmatrix} \tag{3}$$

To carry out a valid $LLab$ to RGB transformation, the RGB values should be trimmed to the 0-255 range (4). This indirect binarization causes backgrounds or structures to disappear; an example is shown in Fig. 1, where the nuclei area is unified in blue, and the background or other structures are shown in yellow or pink.

$$R = \begin{cases} 0, & R < 0 \\ R, & R \in < 0, 255 > \\ 255, & R > 255 \end{cases}, G = \begin{cases} 0, & G < 0 \\ G, & G \in < 0, 255 > \\ 255, & G > 255 \end{cases}, B = \begin{cases} 0, & B < 0 \\ B, & B \in < 0, 255 > \\ 255, & B > 255 \end{cases} \tag{4}$$



(a) TCGA-NH-A8F7 (b) color transfered

Fig. 1. Indirect binarization - removal of selected structures.

3 Nuclei Detection

To detect cell nuclei, the binarization is performed on the converted color image using the formula (5) [18]. The color space parameters are chosen based on the determination of a set of conversion parameters which for a binarized image give the most accurate mask coverage, as described in the formula (6), taking into account pixel-by-pixel comparison. The design of such a formula, based on the standard concept of $ACCURACY$, is dependent of the density of the detected cell nuclei within the entire image, therefore another approach (Jaccard formula, 7) can be also tested.

$$V_3(x, y) = \begin{cases} 0, R(x, y) > 0 \\ 1, R(x, y) = 0 \end{cases} \tag{5}$$

$$ACC = \frac{TP + TN}{TP + TN + FN + FP} \tag{6}$$

$$JACC = \frac{TP}{TP + FN + FP} \tag{7}$$

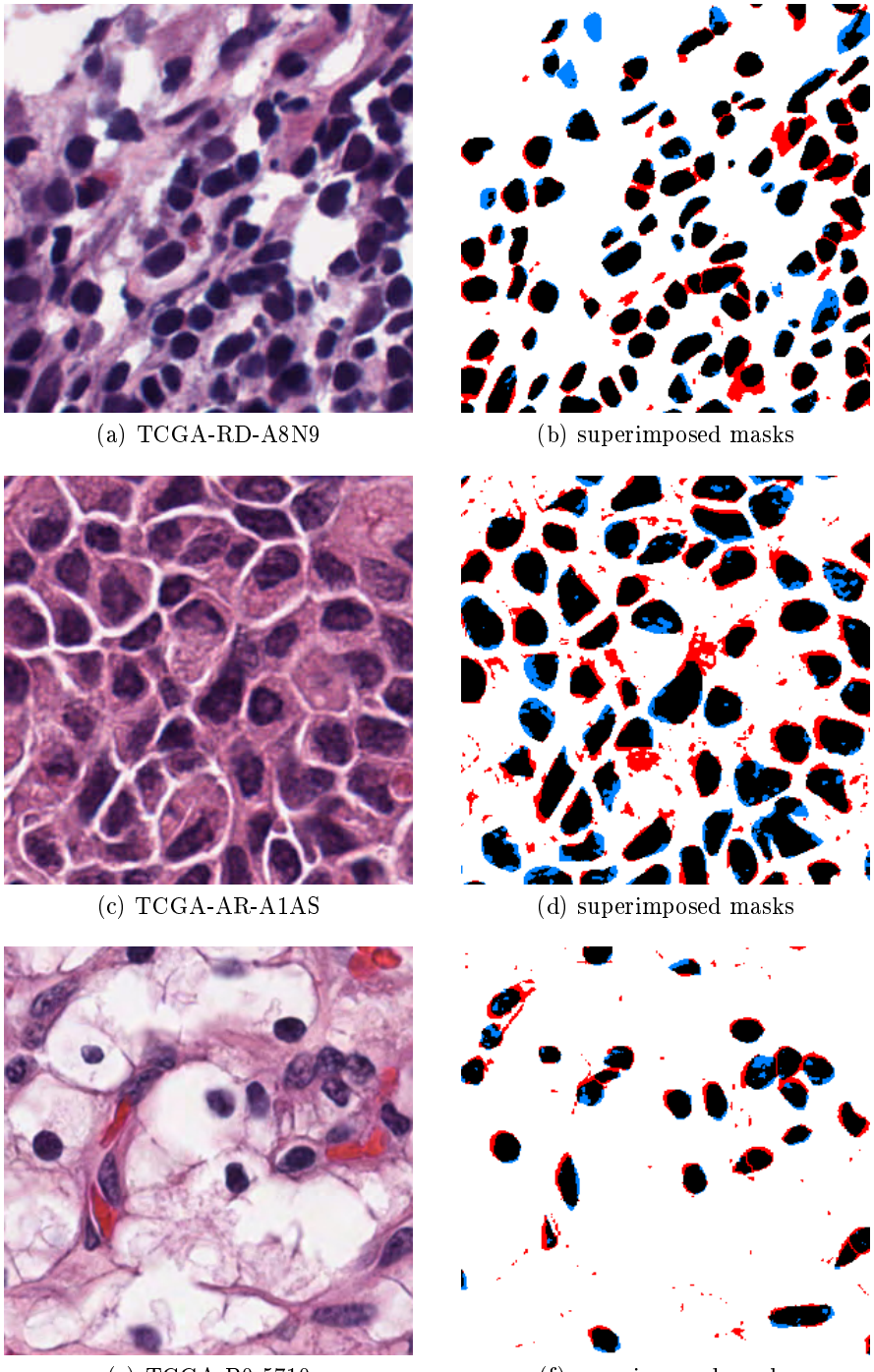
In order to determine the best possible matches, for each image of the dataset 150,000 random target value sets $(\overline{R}, \sigma_t^{LL}, \overline{G}, \sigma_t^{La}, \overline{B}, \sigma_t^{Lb})$ of normalization parameters $(\overline{LL}, \sigma_t^{LL}, \overline{La}, \sigma_t^{La}, \overline{Lb}, \sigma_t^{Lb})$ were generated and tested. The ranges for parameters are presented below.

- $\overline{LL}$ - LL calculated for target $(\overline{R}, \overline{G}, \overline{B})$ with range of 0-255,
- σ_t^{LL} - LL standard deviation - 1 - 1000,
- $\overline{La}$ - La calculated for target $(\overline{R}, \overline{G}, \overline{B})$ with range of 0-255,
- σ_t^{La} - La standard deviation - 1 - 1000,
- $\overline{Lb}$ - Lb calculated for target $(\overline{R}, \overline{G}, \overline{B})$ with range of 0-255,
- σ_t^{Lb} - Lb standard deviation - 1 - 1000.

The results are presented in Table 1.

Label	localization	type	$\overline{R}$	σ_t^{LL}	$\overline{G}$	σ_t^{La}	$\overline{B}$	σ_t^{Lb}	ACC	JACC
TCGA-KB-A93J-01A-01-TS1	Stomach	Stomach adenocarcinoma	195	474	165	3	132	2	0.9170	0.7752
TCGA-RD-A8N9-01A-01-TS1	Stomach	Stomach adenocarcinoma	218	457	77	6	219	4	0.9214	0.7612
TCGA-DK-A2I6-01A-01-TS1	Bladder	Bladder Urothelial Carc.	247	530	76	36	28	11	0.9314	0.7460
TCGA-18-5592-01Z-00-DX1	Liver	Lung squamous cell carcinoma	64	886	161	139	118	117	0.7202	0.7310
TCGA-G2-A2EK-01A-02-TSB	Bladder	Bladder Urothelial Carc.	224	270	72	35	55	16	0.9510	0.6452
TCGA-AR-A1AS-01Z-00-DX1	Breast	Breast invasive carcinoma	201	492	29	19	106	2	0.8942	0.6436
TCGA-B0-5710-01Z-00-DX1	Kidney	Kidney renal clear cell carc.	186	160	92	201	97	12	0.9566	0.6318
TCGA-NH-A8F7-01A-01-TS1	Colon	Colon adenocarcinoma	195	456	61	45	194	20	0.8705	0.6192
TCGA-AY-A8YK-01A-01-TS1	Colon	Colon adenocarcinoma	251	584	218	51	92	16	0.8801	0.6149
TCGA-A7-A13F-01Z-00-DX1	Breast	Breast invasive carcinoma	236	404	86	0	113	0	0.9206	0.5924
TCGA-HE-7130-01Z-00-DX1	Kidney	Kidney renal papillary cell carc.	222	255	29	670	57	45	0.8569	0.5831
TCGA-49-4488-01Z-00-DX1	Liver	Lung adenocarcinoma	196	205	5	464	125	48	0.8264	0.5745
TCGA-A7-A13E-01Z-00-DX1	Breast	Breast invasive carcinoma	247	413	83	1	101	0	0.9177	0.5659
TCGA-21-5786-01Z-00-DX1	Liver	Lung squamous cell carcinoma	123	803	69	14	50	23	0.8347	0.5605
TCGA-B0-5711-01Z-00-DX1	Kidney	Kidney renal clear cell carc.	211	199	109	182	100	5	0.9425	0.5571
TCGA-B0-5698-01Z-00-DX1	Kidney	Kidney renal clear cell carc.	187	182	191	122	88	15	0.9424	0.5489
TCGA-G9-6356-01Z-00-DX1	Prostate	Prostate adenocarcinoma	249	284	74	612	0	18	0.8993	0.5435
TCGA-G9-6336-01Z-00-DX1	Prostate	Prostate adenocarcinoma	67	140	149	712	2	34	0.8051	0.5174
TCGA-AR-A1AK-01Z-00-DX1	Breast	Breast invasive carcinoma	223	365	127	72	103	36	0.8679	0.5123
TCGA-50-5931-01Z-00-DX1	Liver	Lung adenocarcinoma	4	670	170	166	163	47	0.7541	0.5050
TCGA-E2-A1B5-01Z-00-DX1	Breast	Breast invasive carcinoma	39	40	0	53	124	1	0.9235	0.4848
TCGA-21-5784-01Z-00-DX1	Liver	Lung squamouscell carcinoma	110	205	217	5	209	21	0.9172	0.4824
TCGA-E2-A14V-01Z-00-DX1	Breast	Breast invasive carcinoma	224	371	55	526	104	11	0.8752	0.4778
TCGA-38-6178-01Z-00-DX1	Liver	Lung adenocarcinoma	238	403	24	388	45	38	0.8571	0.4664
TCGA-G9-6362-01Z-00-DX1	Prostate	Prostate adenocarcinoma	168	436	34	763	187	54	0.8170	0.4627
TCGA-HE-7128-01Z-00-DX1	Kidney	Kidney renal papillary cell carc.	255	370	0	137	68	15	0.9262	0.4516
TCGA-HE-7129-01Z-00-DX1	Kidney	Kidney renal papillary cell carc.	217	383	75	71	47	43	0.8642	0.4374
TCGA-G9-6348-01Z-00-DX1	Prostate	Prostate adenocarcinoma	78	56	1	557	0	14	0.8499	0.4280
TCGA-CH-5767-01Z-00-DX1	Prostate	Prostate adenocarcinoma	205	331	75	856	221	48	0.8528	0.4154
TCGA-G9-6363-01Z-00-DX1	Prostate	Prostate adenocarcinoma	93	246	4	787	84	27	0.8299	0.4027

Table 1. The best JACC values and their color transfer parameters for all pictures of the dataset. The best ACC values are also presented.

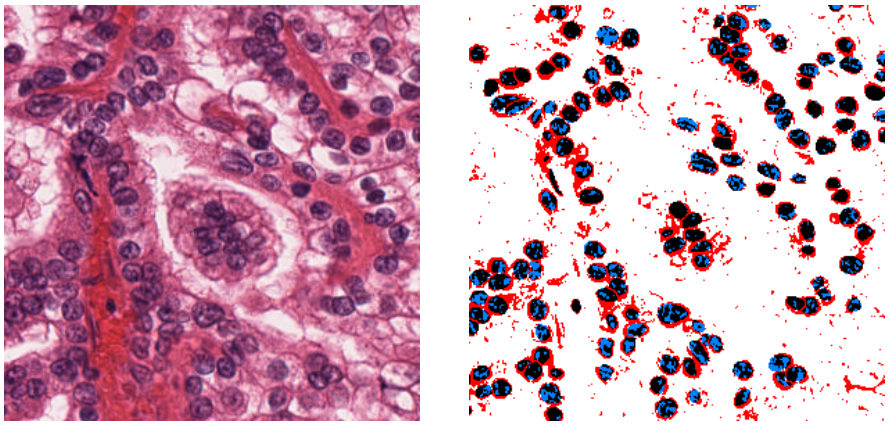


(a) TCGA-RD-A8N9 (b) superimposed masks

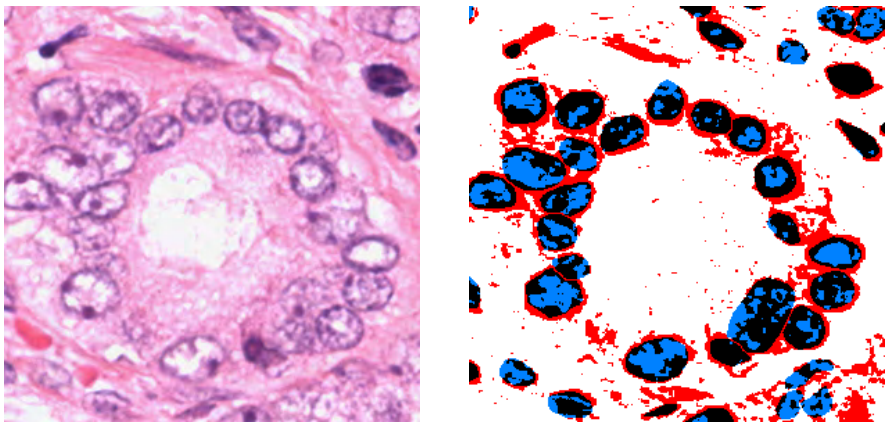
(c) TCGA-AR-A1AS (d) superimposed masks

(e) TCGA-B0-5710 (f) superimposed masks

Fig. 2. The best matches for images with fully filled cell nuclei.



(a) TCGA-HE-7129 (b) superimposed masks



(c) TCGA-G9-6363 (d) superimposed masks

Fig. 3. The worst matches for images with open chromatin nuclei.

4 Conclusions

The presented approach produced rather good results for H&E images in the case of non-damaged cell nuclei that were not involved in the neoplastic process (Fig. 2). The accuracy of the proposed method is very high for a full pixel-by-pixel comparison of objects and backgrounds, as well as for comparison of objects themselves. Unfortunately, for the open chromatin nucleus, this method is unable to generate a full closed cell nucleus mask, which results in poorer outcomes when using the proposed comparison methods (Fig. 3). As a separate observation from the presented research, it can be pointed out that the use of a standard assessment approach based on the accuracy formula (6) fails, and there is a need to reduce the influence of the background area size (7) in order to precisely detect nuclei.

The proposed method will be further developed and also applied in other issues of color image segmentation. The preliminary results of the research confirm its benefits. A sample algorithm demo is available on the site:

http://home.agh.edu.pl/pioro/hecell/

Acknowledgement. This work was financed by the AGH – University of Science and Technology, Faculty of EAIIB, KBIB no 16.16.120.773.

References

1. Bautista, P., Hashimoto, N., Yagi, Y.: Color standardization in whole slide imaging using a color calibration slide. Journal of Pathology Informatics **5**(1), 4 (2014). https://doi.org/10.4103/2153-3539.126153
2. Ing, N., Salman, S., Ma, Z., Walts, A., Knudsen, B., Gertych, A.: Machine learning can reliably distinguish histological patterns of micropapillary and solid lung adenocarcinomas. In: Conference of Information Technologies in Biomedicine. pp. 193–206. Springer (2016)
3. Iwaszenko, S., Nurzynska, K.: Rock grains segmentation using curvilinear structures based features. In: Real-Time Image Processing and Deep Learning 2019. vol. 10996 (2019). https://doi.org/10.1117/12.2519580
4. Kłeczek, P., Dyduch, G., Jaworek-Korjakowska, J., Tadeusiewicz, R.: Automated epidermis segmentation in histopathological images of human skin stained with hematoxylin and eosin. In: Medical Imaging 2017: Digital Pathology. vol. 10140, p. 101400M. International Society for Optics and Photonics (2017)
5. Kłeczek, P., Mól, S., Jaworek-Korjakowska, J.: The accuracy of H&E stain unmixing techniques when estimating relative stain concentrations. In: Polish Conference on Biocybernetics and Biomedical Engineering, PCBBE 2017. AISC, vol. 647, pp. 87–97. Springer
6. Korzynska, A., Roszkowiak, L., Pijanowska, D., Kozlowski, W., Markiewicz, T.: The influence of the microscope lamp filament colour temperature on the process of digital images of histological slides acquisition standardization. Diagnostic pathology **9**(1), S13 (2014)

7. Kowal, M., Filipczuk, P., Obuchowicz, A., Korbicz, J., Monczak, R.: Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images. Computers in Biology and Medicine **43**(10), 1563–1572 (2013)
8. Kowal, M., Korbicz, J.: Marked point process for nuclei detection in breast cancer microscopic images. In: Polish Conference on Biocybernetics and Biomedical Engineering, PCBBE 2017. AISC, vol. 647, pp. 230–241. Springer (2018)
9. Kowal, M., Skobel, M., Nowicki, N.: The feature selection problem in computer-assisted cytology. International Journal of Applied Mathematics and Computer Science **28**(4), 759–770 (2018)
10. Kumar, N., Verma, R., Sharma, S., Bhargava, S., Vahadane, A., Sethi, A.: A dataset and a technique for generalized nuclear segmentation for computational pathology. IEEE transactions on medical imaging **36**(7), 1550–1560 (2017)
11. Li, J., Speier, W., Ho, K.C., Sarma, K.V., Gertych, A., Knudsen, B.S., Arnold, C.W.: An em-based semi-supervised deep learning approach for semantic segmentation of histopathological images from radical prostatectomies. Computerized Medical Imaging and Graphics **69**, 125–133 (2018)
12. Li, X., Plataniotis, K.N.: A complete color normalization approach to histopathology images using color cues computed from saturation-weighted statistics. IEEE Transactions on Biomedical Engineering **62**(7), 1862–1873 (2015)
13. Macenko, M., Niethammer, M., Marron, J., Borland, D., Woosley, J.T., Guan, X., Schmitt, C., Thomas, N.E.: A method for normalizing histology slides for quantitative analysis. In: IEEE International Symposium on Biomedical Imaging, ISBI'09. pp. 1107–1110. IEEE (2009)
14. Mazurek, P., Oszutowska-Mazurek, D.: From the Slit-Island Method to the Ising model: analysis of irregular grayscale objects. International Journal of Applied Mathematics and Computer Science **24**(1), 49–63 (2014)
15. Nurzynska, K.: Optimal parameter search for colour normalization aiding cell nuclei segmentation. CCIS, vol. 928, pp. 349–360. Springer (2018)
16. Onder, D., Zengin, S., Sarioglu, S.: A review on color normalization and color deconvolution methods in histopathology. Applied Immunohistochemistry & Molecular Morphology **22**(10), 713–719 (2014)
17. Oszutowska-Mazurek, D., Mazurek, P., Parafiniuk, M., Stachowicz, A.: Method-induced errors in fractal analysis of lung microscopic images segmented with the use of histaenn (histogram-based autoencoder neural network). Applied Sciences **8**(12), 2356 (2018)
18. Piorkowski, A., Gertych, A.: Color normalization approach to adjust nuclei segmentation in images of hematoxylin and eosin stained tissue. In: Information Technology in Biomedicine. ITIB 2018. AISC, vol. 762, pp. 393–406. Springer (2019)
19. Reinhard, E., Adhikhmin, M., Gooch, B., Shirley, P.: Color transfer between images. IEEE Computer Graphics and Applications **21**(5), 34–41 (2001)
20. Roszkowiak, Ł., Korzyńska, A., Siemion, K., Pijanowska, D.: The influence of object refining in digital pathology. In: Image Processing and Communications Challenges 10. IP&C 2018. AISC, vol. 892, pp. 55–62. Springer (2019)
21. Ruifrok, A.C., Johnston, D.A., et al.: Quantification of histochemical staining by color deconvolution. Analytical and quantitative cytology and histology **23**(4), 291–299 (2001)
22. Salvi, M., Molinari, F.: Multi-tissue and multi-scale approach for nuclei segmentation in h&e stained images. Biomedical engineering online **17**(1), 89 (2018)
23. Starosolski, R.: New simple and efficient color space transformations for lossless image compression. Journal of Visual Communication and Image Representation **25**(5), 1056–1063 (2014)

24. Starosolski, R.: Human visual system inspired color space transform in lossy jpeg 2000 and jpeg xr compression. In: International Conference: Beyond Databases, Architectures and Structures. pp. 564–575. Springer (2017)
25. Tosta, T.A.A., Neves, L.A., do Nascimento, M.Z.: Segmentation methods of h&e-stained histological images of lymphoma: A review. Informatics in Medicine Unlocked **9**, 35–43 (2017)