

Detekcja jąder komórkowych w obrazach biomedycznych z uwzględnieniem kodowania barw

Streszczenie. W artykule omówiono proces ekstrakcji parametrów morfometrycznych cyfrowych obrazów histopatologicznych na przykładzie obrazów raka piersi. Wskazano empirycznie wyznaczoną reprezentację kolorów do skutecznego zautomatyzowanego wykrywania jąder komórkowych. Przedstawiono problematykę związaną z komputerowym wspomaganie rozpoznawania obrazów biomedycznych. Przedstawiono, zapisane w pliku csv wyniki dla przetworzonych i rozpoznanych jąder komórkowych (opis liczbowy struktur i obiektów morfologicznych). Wskazano kierunki dalszych.

Abstract. The paper presents process of morphometric parameters extraction of the digital biomedical image of breast cancer. There was present empirical determination of most effective color channel for automated detection of cell nuclei. The problem of computer-aided biomedical image recognition are presented. The results obtained for processed and properly recognized cell nuclei was presented. All features (numerical description of morphological structures and objects) was stored in the csv file. The future work areas are indicated. (*Detection of cell nuclei in biomedical images depending on color coding*).

Słowa kluczowe: analiza kolorów, detekcja obiektów, ekstrakcja cech, parametry morfometryczne.

Keywords: image color analysis, object detection, feature extraction, morphometric parameters.

Wprowadzenie

Kluczową rolę w podejmowaniu decyzji związanych z diagnostyką i rodzajem podejmowanej terapii odgrywa wykrycie zmian patologicznych [1]. W związku z obserwowaną zwiększoną zapadalnością na choroby nowotworowe [2,3] wspomaganie analizy i interpretacji badanego materiału staje się wymaganym kierunkiem rozwoju. Tym bardziej, że kluczową rolę skuteczności terapii, często stanowi czas podjęcia leczenia. Rozwój i wspomaganie interpretowania obrazów biomedycznych staje się bardzo ważnym elementem w diagnostyce szczególnie chorób nowotworowych.

Analiza obrazów cytologicznych nakierunkowana na aspekty ilościowe może być efektywnie wykorzystywana do wspierania decyzji patologów przy rozpoznawaniu i wnioskowaniu o obecności zmian patologicznych, bądź ich braku. Dodatkowo analiza może wesprzeć ocenę postępu choroby oraz skuteczności terapii celowanej. Komputerowo wspomagane systemy diagnozowania bezpośrednio wpływają na wydajność i skuteczność przy klasyfikowaniu zmian patologicznych. Wraz z rozwojem technologii obrazowania cyfrowego, nastąpił postęp w rozwoju rozwiązań sprzętowych - skanerów cyfrowych. Udoskonalenie jakości zapisywania i przeglądania danych biomedycznych (m.in. whole slide imaging) skutkuje zainteresowaniem badaczy problematyką rozpoznawania i klasyfikacji zmian patologicznych (nowotworów) [4,5].

Szczególnie ważnym elementem interpretacji obrazów biomedycznych (histologicznych, cytologicznych) jest detekcja jąder komórkowych. Automatyzacja tego procesu wymaga wydajnego mechanizmu oceny i separacji jąder, które nakładają się na siebie. Jest to złożony problem badawczy, wpływający na skuteczność rozpoznawania jąder na obrazie. Zagadnieniu temu poświęca się wiele uwagi w doniesieniach badawczych [6,7]. Ręczna i automatyczna segmentacja jąder komórkowych jest opisana w literaturze fachowej, jednak pojawiają się nowe metody, które są rozwijane [8,9]. Problem detekcji jąder komórkowych jest aktualny, wymaga dalszego rozpoznania, a metody detekcji wymagają dalszego rozwoju. Warunki klasyfikacji jąder generują informację dotyczącą oceny poziomu choroby, czy też reakcji na proponowane leczenie. W związku z tym, dobranie klasyfikatorów, tak by uchwycić wszelkie odchylenia w budowie jądra komórkowego, to istotny element rozwijanych metod przetwarzania obrazów biomedycznych.

Stosowane w literaturze reprezentacje barw oraz charakterystyka wskazanego w literaturze materiału badawczego dla obrazów histopatologicznych

Cyfrowe obrazy medyczne, które podlegają interpretacji i ocenie przygotowywane są do przeprowadzenia operacji detekcji i klasyfikacji jąder komórkowych. Przegląd doniesień badawczych [10-21] wskazuje na brak standardów w zakresie akwizycji obrazów cyfrowych, co może wpływać na trudność weryfikacji proponowanych metod przetwarzania obrazów dla danych pozyskanych w skrajnie różniących się warunkach zewnętrznych. Udostępnione bazy danych obrazów cytologicznych są wsparciem do weryfikacji opracowanych metod na zróżnicowanym zbiorze danych (pozyskanych z różnych laboratoriów, innej grupy pacjentów), jednak różnią się one pod względem jakości, wielkości i formatu zapisu danych, czy możliwością pobrania obrazu.

Badacze zajmowali się reprezentacjami przestrzeni barw w zastosowaniach (tele)patologii [10,11]. Segmentacja jąder komórkowych może być wykonana w jednej z wielu przestrzeni, np. HSI, L^*a^*b , RGB. W tabeli 1 - dla wybranych doniesień z badań - znajdują się podstawowe dane dotyczące wykorzystywanego w badaniach materiału cyfrowego. Zestawienie obejmuje prace badawcze, gdzie przedstawiano metody przetwarzania obrazu i detekcji jąder komórkowych.

Problem badawczy

Istnieje cały szereg problemów, z którymi spotykamy się w trakcie komputerowego wspomaganie rozpoznawania obrazów biomedycznych. Przygotowanie materiału badawczego do efektywnego pozyskania danych diagnostycznych, których właściwe przetworzenie może być wsparciem diagnostyki patologów składa się z kilku etapów.

Pierwszy z nich to przygotowanie materiału (pobranie, wybarwienie i utrwalenie materiału badawczego). Na tym etapie występują trudności w postaci braku jednolitych standardów związanych z procesem pobierania materiału i jego przygotowywania, wliczając w to ujednoczone standardy nanoszenia materiału na szkiełka mikroskopowe. Przykłady materiału cyfrowego przedstawiono na rysunku 1. Następnie materiał poddaje się cyfryzacji, w wyniku czego po przetworzeniu uzyskiwany materiał charakteryzuje się dużym rozmiarem a co za tym idzie zwiększonym zapotrzebowaniem na zasoby sprzętowe.

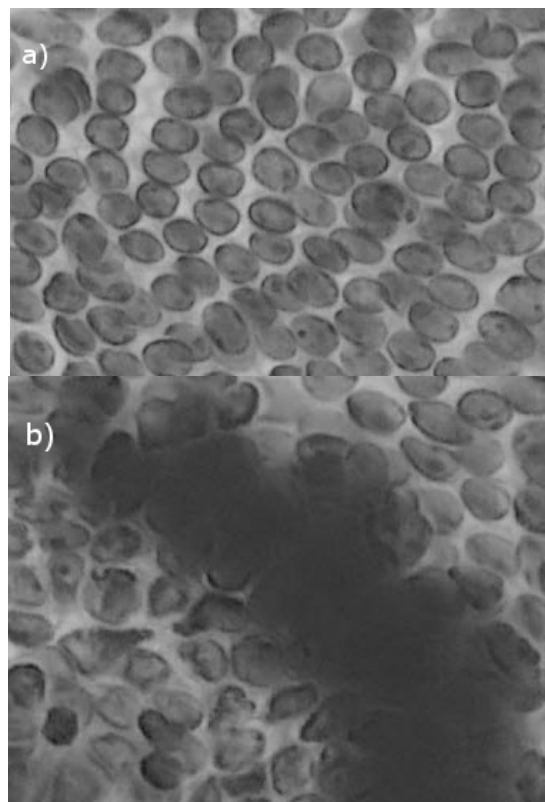
Tabela 1. Wykorzystanie materiału cyfrowego histopatologii w badaniach detekcji jąder komórkowych – opracowanie własne na podstawie [10-21]

Reprezentacja koloru	Lit.	Charakterystyka materiału badawczego				L. próbek
		Wybarwienie	Rozdzielczość	Powiększenie	Materiał	
RGB	[12]	H&E	960×512	10x	rak płuc	16
	[13]	H&E	3264×2448	40x	rak płuc	1038
	[14]	H&E	1392 × 1040	40x	rak piersi	234
	[15]	H&E	1024×1024	bd.	rak piersi	6
	[16]	H&E	0.25 µm na piksel	40x	rak piersi	51
	[10]	H&E	0.5 µm na piksel	20x	rak piersi	44
	[11]	DAB	128×128 (1400), 192×192 (1200) 256×256 (800)	40x	rak piersi	3400
	[17]	DAB	6720×4200	20x	rak przelyku, rak okrężnicy	95
	[18]	Giemsa	2592×3872	100x	białaczka limfo-blastyczna	312
HSV	[8]	H&E	704×576	400x	rak piersi	450
	[12]	H&E	960×512	10x	rak płuc	16
	[13]		3264×2448	40x	rak płuc	1083
	[16]	H&E	0.25 µm na piksel	40x	rak piersi	51
	[10]	H&E		20x	rak piersi	44
	[11]	DAB	128×128 (1400), 192×192 (1200) 256×256 (800)	40x	rak piersi	3400
LUV	[18]	Giemsa	2592×3872	100x	białaczka limfoblastyczna	312
	[13]	H&E	3264×2448	40x	rak płuc	1038
LUV	[16]	H&E	0.25 µm na piksel	40x	rak piersi	51
	[19]	H&E	11 whole-slide images, zawierające 1965 jąder komórkowych	40x	rak mózgu/glejak	11
	[13]	H&E	3264×2448	40x	rak płuc	1038
LAB	[16]	H&E	0.25 µm na piksel	40x	rak piersi	51
	[20]	H&E	30×30	40x	rak piersi	35
	[21]	H&E	4000×7000 5000×23000 (zestaw testowy)	20x	rak prostaty	17
	[19]	H&E	11 whole-slide images, zawierające 1965 jąder komórkowych	40x	rak mózgu/glejak	11
	[13]	H&E	3264×2448	40x	rak płuc	1038
CIE-L*a*b	[11]	DAB	128×128 (1400) 192×192 (1200) 256×256 (800)	40x	rak piersi	3400
YCbCr	[17]	DAB	6720×4200	20x	rak przelyku, rak okrężnicy	95
	[11]	DAB	128×128 (1400) 192×192 (1200) 256×256 (800)	40x	rak piersi	3400

Współistniejącym problemem są zamknięte formaty stosowane przy cyfryzacji próbek. W celu zredukowania ilości danych ładowanych do pamięci, wykorzystywane są dedykowane formaty plików, które pozwalają na wczytywanie fragmentów obrazów. Formaty te często bezpośrednio powiązane są z rodzajem oraz producentem sprzętu mikroskopowego. Zamknięte formaty utrudniają bezpośredni dostęp do parametrów kodowania danych obrazu. Ponadto ukryty format kodowania plików oraz może stanowić trudność przy bezstratnym przenoszeniu informacji przy przenoszeniu danych do innych formatów.

Charakter wykorzystanego materiału badawczego

Materiał badawczy został wybrany z zestawu rzeczywistych obrazów dostarczonych przez patologów. Obrazy zostały pozyskane przy użyciu mikroskopu firmy Olympus VS120 (Virtual Microscopy Slide Scanning System). Pozyskano wirtualne slajdy o wielkości ok. 200000 x 80000 pikseli, z powiększeniem 40-krotnym. Próbki tkanki nowotworowej piersi pobranych metodą biopsji cienkoigłowej. Materiał wybarwiono hematoksyliną i eozyną (H&E). Próbki zostały podzielone przez patologa jako nowotwory złośliwe, łagodne lub gruczolakowłóknaki. Materiał badawczy został wybrany z zestawu rzeczywistych obrazów dostarczonych przez patologów. Obrazy zostały pozyskane przy użyciu mikroskopu firmy Olympus VS120 (Virtual Microscopy Slide Scanning System). Pozyskano wirtualne slajdy o wielkości ok. 200000 x 80000 pikseli, z powiększeniem 40-krotnym. Próbki tkanki nowotworowej piersi pobranych metodą biopsji cienkoigłowej. Materiał wybarwiono hematoksyliną i eozyną (H&E). Próbki zostały podzielone przez patologa jako nowotwory złośliwe, łagodne lub gruczolakowłóknaki.

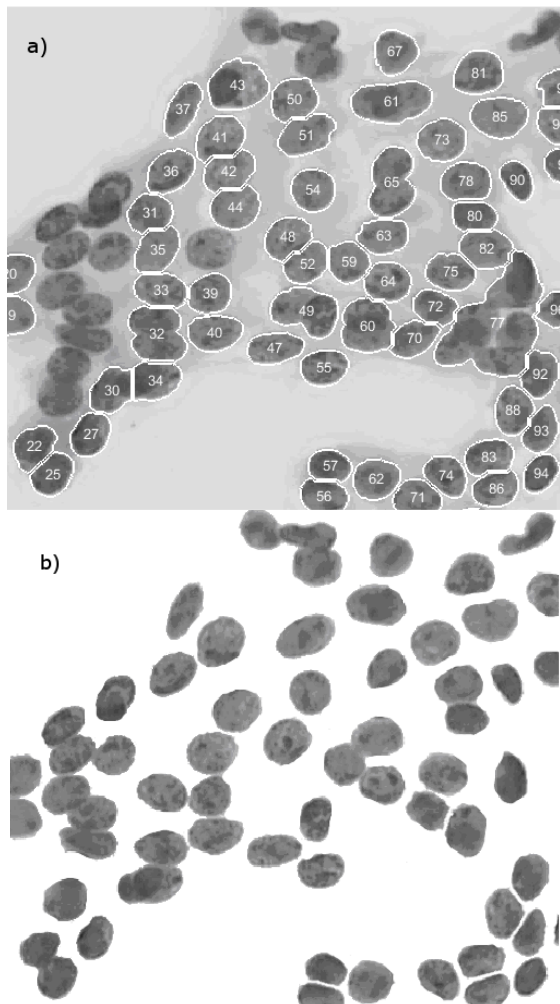


Rys.1. Przykładowy obraz badawczy – fragment czytelny (a) nieczytelny (b)

Obraz źródłowy charakteryzuje się dużymi rozmiarami (rzęd setek MB), ponad to może zawierać tysiące komórek. W związku z ograniczeniami programowymi i sprzętowymi do diagnostyki wybierane są części obrazu, z materiałem,

który dostarcza wymaganych informacji o charakterystyce jąder komórkowych. Obszary takie nazywane są obszarami zainteresowania – ROI (ang. Region of interest). Lekarze podejmują decyzję o klasyfikacji chorób na podstawie analizy ww. obszarów.

Do badań wykorzystano 60 wybranych obszarów (ROI), z czego 41 z nich sklasyfikowane zostały jako złośliwe, 9 jako gruczolakowłókniak, a 10 obszarów jako zmiany łagodne. Ponieważ lekarz podejmuje decyzję na podstawie wizualnej analizy komórek, do celów badawczych wyselekcjonowano jedynie jądra komórkowe, które są wyraźnie widoczne i widoczne są wszystkie granice jądra. Taki zestaw obrazów z ręcznie wyselekcjonowanymi obszarami jąder komórkowych został wykorzystany jako odniesienie do weryfikacji wyników uzyskanych przy użyciu metod komputerowych (bez udziału patologa). Procedura weryfikacji polegała na porównaniu obrazu referencyjnego wytworzonego przez człowieka z obrazem przetworzonym komputerowo. Zliczone zostały prawidłowo wykryte jądra komórkowe. Procent wykrytych jąder został przyjęty jako parametr efektywności i skuteczności komputerowego wspomaganie diagnostyki tego typu zmian.



Rys.2. Wykrywanie jąder komórkowych wspomaganie komputerowo (a) wskazuje przez człowieka (b)

Procedura weryfikacji detekcji jąder komórkowych

W celu oceny skuteczności wykrywania jąder komórkowych wspomaganego komputerowo zastosowano porównanie zautomatyzowanej segmentacji obrazu z wynikami oceny jąder przez patologa.

Niech Nh będzie zbiorem jąder zidentyfikowanych na obrazie przez człowieka (patologa), gdzie $Nh = \{nh_1, \dots, nh_j\}$. Niech nh_i będzie identyfikowane przez (xh_i, yh_i) - parę współrzędnych analizowanego obrazu. Niech Nm będzie zbiorem jąder wykrytych z wykorzystaniem wspomaganie komputerowego, gdzie $Nm = \{nm_1, \dots, nm_j\}$. Niech nm_i będzie identyfikowany przez (xm_i, ym_i) - parę współrzędnych analizowanego obrazu. Zdefiniujemy $AL(x, y)$ jako zbiór przybliżonego położenia punktu (x, y) .

Niech AC będzie współczynnikiem aproksymacji, np. o wartości 10. Para współrzędnych (x_a, y_a) należy do zestawu AL

$$(1) \quad (x_a, y_a) \in AL(x_b, y_b)$$

jeśli spełnione są równania (2) i (3)

$$(2) \quad x_b - AC \leq x_a \leq x_b + AC,$$

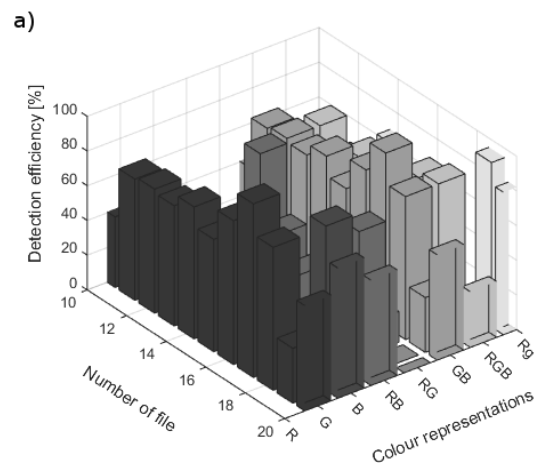
$$(3) \quad y_b - AC \leq y_a \leq y_b + AC.$$

Funkcję wyniku detekcji (DR) można definiować jako:

(4)

$$\sum_j^{Nm} DR_j = \begin{cases} 0 & \text{jeżeli} \quad \bigvee_{h_k} AL(xh_k, yh_k) (xm_j, ym_j) \notin AL(xh_k, yh_k) \\ 1 & \text{jeżeli} \quad \bigvee_{h_k} AL(xh_k, yh_k) (xm_j, ym_j) \in AL(xh_k, yh_k) \end{cases}$$

Funkcja DR służy do określenia, czy jądra komórkowe j wskazane w wyniku komputerowej analizy obrazu są obecne w obrazie, który został poddany manualnej analizie (wykrywanie jąder) przez patologa.



Rys.3. Skuteczność wykrywania jąder komórkowych dla wybranych ośmiu reprezentacji kolorystycznych na obrazach reprezentujących zmiany złośliwe (a), łagodne (b)

Tabela 2. Ekstrakcja cech - fragment wygenerowanego pliku csv

Filename	M-11940-x40-11.tif			
Area	63.742286	56.905775	105.95355	142.95014
Centroid1	419.158857	385.72644	437.20036	592.59139
Centroid2	43.581205	13.734020	27.73754	27.492040
MajorAxisLength	70.317614	31.572930	52.10685	56.194247
MinorAxisLength	2.025398	1.266600	1.405229	1.690639
CircularShapeFactor	1.613485	2.298885	1.878567	2.044019
ElongationA	0.619777	0.434994	0.532321	0.489232
ElongationB	0.450630	0.420220	0.514900	0.412479
Roundness	723.000000	18.000000	128.00000	204.00000
Convex Deficiency	0.784778	0.900434	0.846543	0.872154
Eccentricity	211.047000	72.364000	139.24500	147.42400
Perimeter	208.309714	196.92097	205.55193	202.78787
MeanIntensity	134	154	153	151
MinIntensity	231	225	229	233
MaxIntensity	63.742286	56.905775	105.95352	142.95014

Wyniki i propozycje dalszych prac

Na podstawie autorskiego algorytmu zaproponowanego w [22] dokonano wyboru najbardziej skutecznej reprezentacji koloru. Na rysunku 3 przedstawiono wyniki dla wybranych obrazów histopatologicznych. Największa wykrywalność jąder komórkowych dla zmian patologicznych łagodnych i złośliwych uzyskano na kanałach R oraz GB reprezentacji kolorów RGB. Zaproponowano proces ekstrakcji cech jąder komórkowych na cyfrowych obrazach histopatologicznych. Obliczenia przeprowadzono dla jąder komórkowych wyselekcjonowanych na 60 wybranych obrazach. Przykładowy obraz z oznaczonymi jądrami przedstawiono na rysunku 2. Przykładowe dane, które finalnie zapisywane są do pliku w formacie csv, znajdują się w tabeli 2. Wyznaczone parametry morfometryczne jąder komórkowych mogą stanowić bazę do automatycznej klasyfikacji zmian patologicznych, co będzie stanowiło kolejny etap prac. Równolegle prowadzone będą prace z zakresu wspomaganie komputerowego wyodrębniania diagnostycznych fragmentów obrazów histopatologicznych.

Authors: dr inż. Anna Pławiak-Mowna, Uniwersytet Zielonogórski, Wydział Informatyki, Elektrotechniki i Automatyki, Instytut Sterowania i Systemów Informatycznych, ul. Licealna 9, 65-417 Zielona Góra, E-mail: a.mowna@issi.uz.zgora.pl; dr inż. Małgorzata Mazurkiewicz, Uniwersytet Zielonogórski, Wydział Informatyki, Elektrotechniki i Automatyki, Instytut Sterowania i Systemów Informatycznych, ul. Licealna 9, 65-417 Zielona Góra, E-mail: m.mazurkiewicz@issi.uz.zgora.pl.

REFERENCES

- Irshad H., Veillard A., Roux L., Racoceanu, D., Methods for Nuclei Detection, Segmentation and Classification in Digital Histopathology: A Review Current Status and Future Potential. *IEEE reviews in biomedical engineering*, 7 (2014), 97-114
- Dyzmann-Sroka A, Malicki J., Cancer incidence and mortality in the Greater Poland Region - Analysis of the year 2010 and future trends. *Reports of Practical Oncology and Radiotherapy*. 19(2014), No.5, 296-300
- Siegel R.L., Miller K.D., Jemal A., Cancer statistics, 2017, *CA Cancer J Clin* 67(2017), No.1, 7-30
- Farris A.B., Cohen C., Rogers T.E., Smith G.H., Whole Slide Imaging for Analytical Anatomic Pathology and Telepathology: Practical Applications Today, Promises, and Perils, *Archives of Pathology & Laboratory Medicine*, 141 (2017), No.4, 542-550
- Elmore J.G., Longton G.M., Pepe M.S., et al., A Randomized Study Comparing Digital Imaging to Traditional Glass Slide Microscopy for Breast Biopsy and Cancer Diagnosis, *Journal of Pathology Informatics*, 8 (2017), 12
- Shu J., Fu H., Qiu G., Kaye P., Ilyas M., Segmenting overlapping cell nuclei in digital histopathology images. *Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 5445-5448, July 2013
- Cui Y., Guiying Z., Liu Z., Xiong Z., Hu.J., A Deep Learning Algorithm for One-step Contour Aware Nuclei Segmentation of Histopathological Images, *CoRR* abs/1803.02786 (2018)
- Kowal M., Filipczuk P., Nuclei segmentation for computer-aided diagnosis of breast cancer, *International Journal of Applied Mathematics and Computer Science*, 24 (2014), No. 1, 19-31
- Liu C., Shang F., Ozolek J.A., Rohde G.K., Detecting and segmenting cell nuclei in two-dimensional microscopy images, *Journal of Pathology Informatics*, 7(2016), 42
- Zarella M.D., Breen D.E., Plagov A, Garcia F.U., An optimized color transformation for the analysis of digital images of hematoxylin & eosin stained slides, *J Pathol Inform*, 2015, 6:33
- Jitaree S. et al., Cell type classifiers for breast cancer microscopic images based on fractal dimension texture analysis of image color layers, *Scanning*, 37 (2015), No. 2, 145-151
- Sammouda M. et al., Cancerous nuclei detection on digitized pathological lung color images, *Journal of Biomedical Informatics*, 35(2002), No. 2, 92-98
- Sajith Kecheril S. et al., Segmentation of lung glandular cells using multiple color spaces, *International Journal of Computer Science, Engineering and Applications*, 2 (2012), No.3, 147-158
- Qi X. et al., Robust segmentation of overlapping cells in histopathology specimens using parallel seed detection and repulsive level set, *IEEE Trans. Biomed. Eng.*, 59 (2012), No.3, 754-765
- Veillard A., Kulikova M.S., Racoceanu D., Cell nuclei extraction from breast cancer histopathology images using colour, texture, scale and shape information. *Diagnostic Pathology*, 8 Suppl 1(2013)
- Dong F, Irshad H, Oh E-Y, Lerwill MF, Brachtel EF, Jones NC, et al., Computational Pathology to Discriminate Benign from Malignant Intraductal Proliferations of the Breast, *PLoS ONE* 9(12) (2014)
- Shu J., Dolman G.E., Duan J., Qiu G., Ilyas M., Statistical colour models: an automated digital image analysis method for quantification of histological biomarkers, *BioMedical Engineering OnLine* (2016);15:46
- Amin M.M et al., Recognition of Acute Lymphoblastic Leukemia Cells in Microscopic Images Using K-Means Clustering and Support Vector Machine Classifier, *Journal of Medical Signals and Sensors*, 5 (2015), No. 1, 49-58
- Kong J. et al., Towards Building Computerized Image Analysis Framework for Nucleus Discrimination in Microscopy Images of Diffuse Glioma, *Conference Proceedings 2011*, 6605-6608
- Khan A.M., Eldaly H., Rajpoot N.M., A gamma-gaussian mixture model for detection of mitotic cells in breast cancer histopathology images, *J Pathol Inform*. 30 (2013), 4-11
- Nguyen K., Jain A.K., Sabata B, Prostate Cancer Detection: Fusion of cytological and textural features, *J Pathol Inform*, 2011, 2:3
- Pławiak-Mowna A., Mazurkiewicz M., On Effectiveness of Human Cell Nuclei Detection Depending on Digital Image Color Representation, *Przegląd Elektrotechniczny* 94(2018), Nr 1, 77-80