RESEARCH   **Open Access**

Clustered nuclei splitting based on recurrent distance transform in digital pathology images

Lukasz Roszkowiak[1*] (ID), Anna Korzynska[1], Dorota Pijanowska[1], Ramon Bosch[2], Marylene Lejeune[3] and Carlos Lopez[3]

*Correspondence:
lroszkowiak@ibib.waw.pl
[1]Nalecz Institute of Biocybernetics
and Biomedical Engineering Polish
Academy of Sciences, Ks. Trojdena 4
Str., 02-109 Warsaw, Poland
Full list of author information is
available at the end of the article

Abstract

The accuracy of the applied technique for automated nuclei segmentation is critical in obtaining high-quality and efficient diagnostic results. Unfortunately, multiple objects in histopathological images are connected (clustered) and frequently counted as one. In this study, we present a new method for cluster splitting based on distance transform binarized with the recurrently increased threshold value and modified watershed algorithm. The proposed method treats clusters separately, splitting them into smaller sub-clusters and conclusively into separate objects, based solely on the shape feature, making it independent of the pixel intensity. The efficiency of these algorithms is validated based on the labeled set of images from two datasets: BBBC004v1 and breast cancer tissue microarrays. Results of initial nuclei detection were significantly improved by applying the proposed algorithms. Our approach outperformed the state-of-the-art techniques based on recall, precision, F1-score, and Jaccard index. The proposed method achieves very low amount of under-segmented, as well as over-segmented objects. In summary, we provide novel and efficient method for dividing the clustered nuclei in digital images of histopathological slides.

Keywords: Biomedical engineering, Digital pathology, Nuclei segmentation, Cluster separation

1 Introduction

In recent years, the development of computational pathology has strongly influenced the progress in quantitative digital pathology [1]. The availability of whole slide image technology has encouraged the development of many systems that perform computer-aided diagnosis. However, certain obstacles and limitations still exist in achieving a reliable result. One of the primary problems is the overlapping of structures, also called clusters of objects (commonly cell nuclei), that are segmented for quantification. The problem arises when multiple objects are counted as one due to clustering. Clustering is observed independently of the applied tissue stains, regardless of the use of hematoxylin and eosin (H&E) or different types of immunohistochemical (IHC) stains. Because brightfield

microscopy provides no information about the depth of the sample, it is only possible to observe the cell nuclei as overlapping flat objects.

In this study, we present a new approach for cluster splitting, which is based on the distance transform and modified watershed algorithm. There is not only no need for pre-processing the images but also the method proposed in this study improves by cluster classification and recurrent approach. We compare our method to the other known methods and validated the efficiency of our method on two datasets: (1) publically available benchmark dataset BBBC0004 [2] and (2) manually labeled set of images of breast cancer cells with IHC staining against FOXP3 with 3,3'-diaminobenzidine and hematoxylin (DAB&H) [3]. Figure 1 shows two examples from the latter dataset.

IHC staining is often employed to augment the diagnosis and allow disease discrimination. The typical examination of the antigen expression is based on the cell's distribution and morphology, as well as on the tissue architecture. Quantitative analysis of such sections can be used to support the diagnosis and evaluation of the disease progression.

The rest of this article is organized as follows. The remaining part of this section describes the related works and reference methods. Section 2 outlines the proposed method along with the dataset description. Section 3 describes the results and Section 4 presents the discussion. Finally, Section 5 presents the conclusion.

1.1   Related works

Cluster splitting is still a troublesome and an unsolved problem in digital pathology. There are many state-of-the-art automated methods for nuclei/cell detection, segmentation, or classification that are reported in the literature, but unfortunately, they are mostly used for typical H&E staining and are difficult to adapt to other staining techniques, such as IHC.

In order to further enhance the capability and accuracy of diagnostic decision, many image analysis techniques have been developed to minimize the problem of over-segmentation and under-segmentation. Some of the researchers tackle the problem of clustering directly, whereas others try to manage it inside their framework. In 2018, Salvi
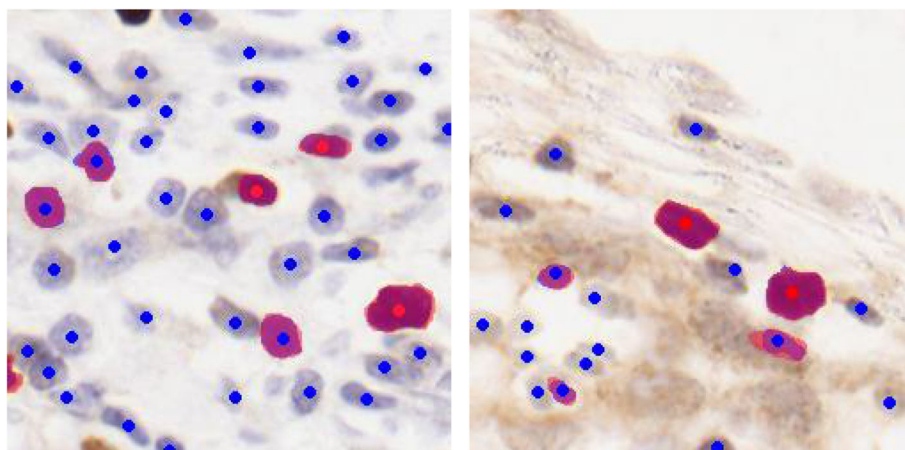
Fig. 1 Two examples of images from dataset with annotations by experts. Positive nuclei are marked with red dots while blue dots represent negative nuclei. The selected nuclei with boundary annotation are shown with blue and red overlay for each expert's marking which results with pink intersection

and Molnari [4] proposed MANA method for nuclei segmentation. Their approach is quite flexible, which means it processes images of different types of tissues and scales but is limited to images with H&E staining. Swiderska et al. [5] presented a method for hot-spot selection where separate nuclei segmentation also played a crucial role. Cheng and Rajapakse [6] developed a solution to process images with fluorescence microscopy in which adaptive H-minima transformation cooperates with watershed-like algorithm. Ali et al. [7] proposed a multilevel set method to solve the problem of clustering. Wienert et al. [8] proposed contour-based minimum model with minimal a priori information. Furthermore, Yan et al. [9–11] proposed various solutions for image and video processing.

Moreover, there are new approaches with deep learning (DL) and convolutional neural networks that not only classify but also localize the nuclei in images [12, 13]. In 2015, Xie et al. [14] proposed an algorithm called deep voting and achieved state-of-the-art results. Recently, Cui et al. [15] reported an automatic end-to-end deep neural network for the segmentation of individual nuclei in high-resolution histopathological images. Unfortunately, it is not available (online) for testing. U-net is a very commonly used baseline model; therefore, in this study, we compare our method with that of a pretrained modified solution, as well as with that of a model trained from scratch.

Apart from the aforementioned methods, literature describes some of the state-of-the-art frameworks for the quantification of nuclei in histopathology. In this study, we compare our proposed method with three such solutions: Tmarker, IHC toolbox, and Qupath. Tmarker [16] was introduced in 2013; it uses color deconvolution, as well as a superpixel-based approach. The ImageJ plugin named IHC toolbox [17] uses oval-fitted nuclei segmentation and quantification functions with automatic processing. Qupath is one of the most popular freeware softwares used in the analysis of whole slide images [18].

1.2 Cluster split methods based on the boundary shape

Within the cluster split methods, we can distinguish the group of methods that are based on the values of the pixel intensity and those that are solely based on the shape of the cluster boundary. The advantage of cluster split methods based independently of the pixel intensity values might be crucial in the case of histopathological images as this type of biological data vary in contrast, brightness, and color representation.

1.2.1 *"Mouelhi"*

This algorithm was presented by Mouelhi et al. [19]. It is based on the construction of a concave vertex graph followed by the selection of the shortest path. They treat the separating edges of the watershed algorithm as propositions from which they select the best fit. To do so, they investigated the endpoints of the edges and looked for the points near the concave points of the outer boundary of the cluster based on the calculated Euclidean distance. Then, by applying the Djikstra algorithm, they computed the shortest, optimal path that can separate the clustered nuclei. This method was developed specifically for clustered breast cancer cells and produces very good results without losing any geometrical cell features.

1.2.2 *"Kong"*

This algorithm for cluster splitting was introduced by Kong et al. [20] which uses an iterative process. It starts by finding the **m**ost **l**ikely **c**oncave **p**oints (mlcp), which are

midpoints of the detected concave regions. Cluster splitting is related to the two essential points of the cluster: radial symmetry center and geometrical center. The line between these two points is the basis for selecting the mlcp as the mlcp should be on either side of the said line. Then, the cluster is cut with the line between the selected mlcp. This procedure is iteratively applied until the size of the resulting object is satisfactorily small. This algorithm was a part of the framework that was tested on pathological images of follicular lymphoma and produced incredible results [21].

1.3   Cluster split methods based on intensity

To prove the efficiency of our approach, we compared our results with those of the state-of-the-art cluster splitting algorithms that are based on the intensity of the pixel.

1.3.1   *Seeded watershed algorithm*

This method was applied as presented in MATLAB Marker-Controlled Watershed Segmentation tutorial [22]. It follows a simple procedure. First, we need to use the gradient magnitude as the segmentation function. Then, we need to compute the foreground markers by performing a series of morphological operations (opening, erosion, reconstruction, closing, dilation, reconstruction, and complementing) and finding the local maxima. Next, we need to estimate the background markers with distance transform algorithm. With the foreground and background makers, the watershed transformation needs to be computed on the gradient magnitude with the imposed minima.

1.3.2   *"Huang"*

The algorithm was presented by Huang et al. [23] and is quite similar to the seeded watershed algorithm. However, the difference is in the calculation of the foreground markers. Instead of using the morphological operations, the foreground markers are estimated by finding the local maxima with the use of distance transform algorithm (calculated on the complementary binary image).

1.3.3   *H-minima*

This method [24] is based on applying the H-minima transform algorithm to the distance transform algorithm of binary image. We applied a threshold of suppressed minima relative to the 0.8 of the value in the distance transform algorithm. The H-minima transformation was followed by the watershed algorithm.

1.4   DL approach to cluster splitting

DL algorithm is a novel diagnostic tool in the automated detection analysis of histological images; therefore, we compared our method with two methods applying DL algorithms: a pretrained model proposed by Chen et al. [25] and the U-net model that is trained from scratch on ScienceBowl2018 dataset [26]. Both these solutions were run in Python (training and inference performed with Nvidia GeForce GTX 850M) with implementations that are available online.

1.4.1   *"Chen"*

Chen et al. [25] proposed a modification in U-net model by applying contour enhancement in loss function. The random clipping and rotation was applied for data augmentation. This is a multi-scale model with variations in the size of the input image during

training. Since overlapped patch-based strategy was used, the resulting region in the output mask was combined from the inference of multiple patches. We used implementation available online [27].

1.4.2 U-net

In this study, we used the model of U-net as described in the original paper [28] and also followed the procedure made available online [29]. This is a very challenging dataset consisting of nuclei segmented from a large number of images acquired under a variety of conditions. The nuclei vary in cell type, magnification, and imaging modality (brightfield and fluorescence).

U-net model has been created in Keras. Adam optimizer, with a mini-batch size of 4, was used to train the network for 500 epochs. The binary cross-entropy loss function and dice coefficient metric was evaluated during training. Learning rate was set to 1e–5, and the rest of the hyperparameters were left at default values. Since the prediction maps are outputted as grayscale images, we put a threshold of typical value equal to half of the maximum intensity value, namely 127, to get the binary output maps used for evaluation.

2 Materials and methods

All classical image processing approach algorithms were implemented in MATLAB R2015b and are available at http://ibib.waw.pl/en/scientific-activity/projects/167-umo-2013-11-n-st7-02797 and on Medical Image Analysis Platform (MIAP) [30] and MATLAB File Exchange.

2.1 Proposed cluster split methods

As cluster split methods cannot be applied to raw images, we first use segmentation to separate objects of interest from the background. In our previous work [31], we tested various adaptive threshold methods and concluded that Bradley's [32] method yields the most reliable results. Therefore, we used his method in this investigation. Local threshold value based on the intensity of the pixel and its neighborhood is calculated at every point of an image with sliding window image processing. Moreover, we use simple post-processing to remove small artifacts from the output image. The chosen segmentation algorithm yields reasonably smooth boundaries, so no further preprocessing is performed. We excluded objects consisting of less than 50 pixels and more than 10,000 pixels.

The method presented in this article tackles the difficult problem of distinguishing separate objects based only on the shape of the clustered structure. It is mainly based on the distance transform algorithm, cluster classification, and modified watershed algorithm. Assigning a proper class to the cluster with a particular composition results in increased robustness of the procedure. We further developed the foreground marker estimation procedure and used recursive invoking of the algorithm. Figure 2 presents the workflow of the proposed method.

Our proposed method treats clusters one-by-one, separately. The input for our method is a cropped binary image containing currently processed cluster. The features of the cluster's shape are calculated and are categorized according to Table 1. We proceed with the calculation of the distance transform on the input image, which is a binary mask of cluster.
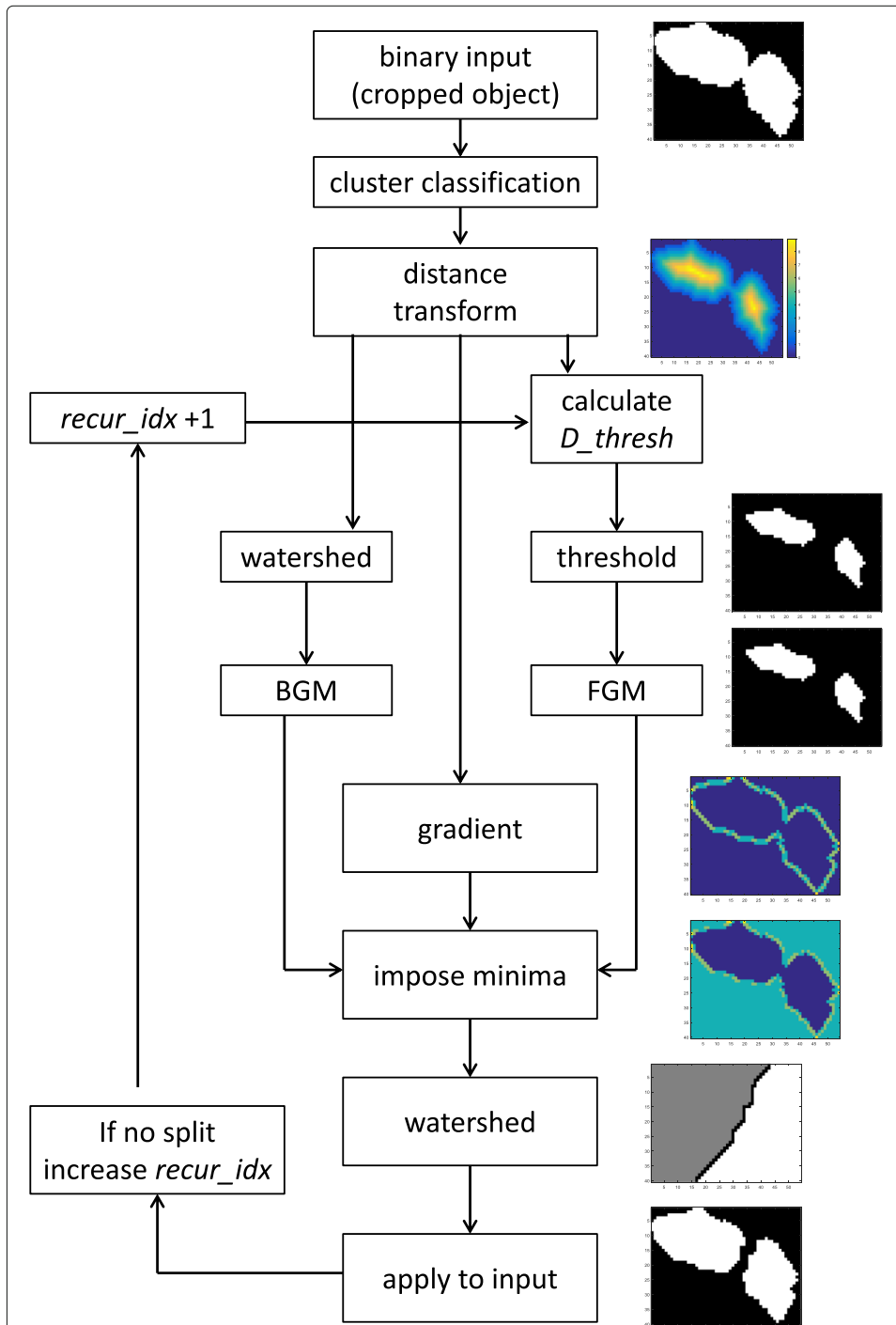
Fig. 2 The scheme of processing using our algorithm with miniature images for better understanding. Cropped binary input contains one cluster. Next, the distance transform (*D*) is calculated. Foreground markers (FGM) are established with distance transform thresholded with *D_tresh* (value based on maximum value present in distance transform map, parameter *T*, and the recurrence index *recur_idx*). Background markers (BGM) are obtained by applying watershed algorithm to distance transform. Then, we impose minima (FGM and BGM) on the gradient magnitude of the input image. If the ridge is present in the result after applying watershed algorithm, then it is used to split the objects. If the resulting objects (tested in terms of area, eccentricity, and roundness) are still classified as clusters, then the algorithm starts again with increased recurrence index

Table 1 The classification of clusters

Cluster class	Criteria	Description
I	Area >average_Area & Eccentricity > 0.8	Generally two overlapping objects, most common
II	Area >average_Area & Eccentricity > 0.6	Three or more objects, with various overlap
III	Area >average_Area & Perimeter > T_Perimeter	Heavily packed objects, with high overlap
IV	Area >3*average_Area	Very big clusters

Where *average_Area* is estimated for dataset based on nonclustered objects of interest; *T_Perimeter* is the perimeter of the circle with *average_Area*. Examples are presented in Fig. 5

Then, we establish a threshold value (D_tresh) based on the following: maximum value present in the distance transform map (D), threshold modifier (T), and the recurrence index (*recur_idx*, initially set to 1). The threshold value calculation is correlated with cluster classification as presented in Algorithm 1. The multiplier in the form of the parameter a allows for the most exhaustive processing of clusters with a vast area. As the threshold increment step is decreased, the separation of clustered objects is more precise. The values of the threshold modifier (T) and parameter a were experimentally set to optimize the number of recurrent iterations. Then, the evaluated threshold value is used on the distance transform map creating foreground markers (FGMs). Concurrently, the distance map is treated with watershed algorithm to obtain background markers (BGMs), which are in most cases nonexistent. Next, we calculate the gradient magnitude of the input image, which is a binary map of cluster. Then, we impose minima on the gradient magnitude so that it has regional minima only at certain desired locations, namely at FGMs. Finally, the watershed algorithm is applied. If the ridge is present in the result of watershed algorithm, then its dilated instance is inserted into the input image to split the objects. The resulting objects are tested in terms of area, eccentricity, and roundness. If

---

Algorithm 1 Calculation of the threshold value depending on the cluster classification

1: **function** CALCULATE $D\_tresh(D, recurr\_idx)$
2:      **if** *cluster_class* $== II$ **then**   ▷ three or more overlapping objects
3:          $T \leftarrow 0.55$
4:          $a \leftarrow 0.2$
5:      **else if** *cluster_class* $== III$ **then**   ▷ heavily packed cluster
6:          $T \leftarrow 0.2$
7:          $a \leftarrow 1$
8:      **else if** *cluster_class* $== IV$ **then**   ▷ very big cluster
9:          $T \leftarrow 0.2$
10:          $a \leftarrow 0.2$
11:      **else**   ▷ most common (two objects overlapping), *cluster_class* $== I$
12:          $T \leftarrow 0.55$
13:          $a \leftarrow 1$
14:      **end if**
15:      $D\_tresh = max(D) * T + (a * (recurr\_idx - 1))$
16:      **return** D_tresh
17: **end function**

---

they are still classified as clusters, then the algorithm starts again for them with the recurrence index set again to 1. Alternatively, if the object was not split, then the algorithm starts again with an incremented recurrence index.

Because the algorithm is applied recurrently with every cycle, the threshold value is incremented. This results in more restrictive FGMs at every step thereby allowing to discriminate between closely and loosely clumped objects. The increment is set to 1 as it permits relatively dynamic calculations with an acceptable time of processing. Figure 3 presents an example of such recurrent execution.
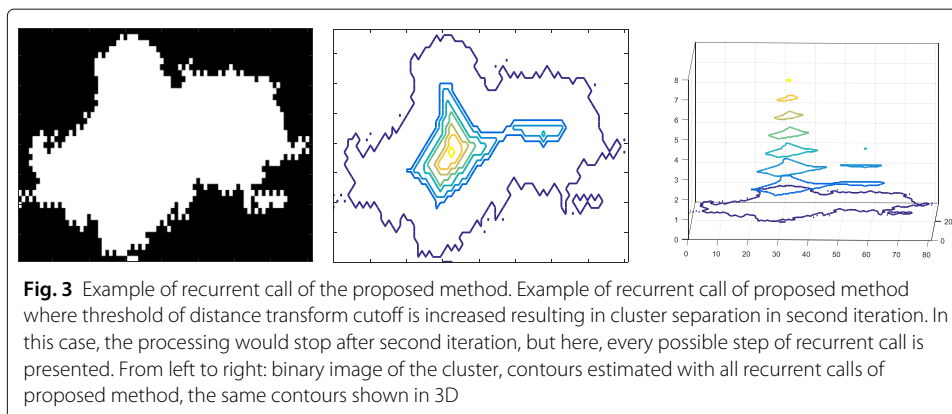
2.2 Datasets

To validate our proposed method, we used two image datasets. Both datasets consisted of images with objects (nuclei) with sparse and compact arrangement including many clustered objects. The dataset (BBBC0004) consisting of the synthetic images provides a more controlled environment, whereas the other dataset (IISPV) allows to test how the implementation is dealing with the natural assignment.

2.2.1 Dataset BBBC0004

We used image set BBBC004v1 [33] from the Broad Bioimage Benchmark Collection [34]. To help assess the performance of the algorithms with regard to cluster splitting, the synthetic image set with known probability of overlap was used. This image set consists of five subsets with increasing degree of clustering, with the following overlap probability: 0, 0.15, 0.30, 0.45, and 0.60. From each subset, we randomly selected five images for evaluation purposes. Each grayscale image contains 300 objects, but the objects overlap and cluster with different probabilities in the five subsets. The images were generated with the SIMCEP simulating platform for fluorescent cell population images [35].

2.2.2 Dataset IISPV

In this study, the images of breast cancer tissue used for the validation of the experiments were obtained from the Molecular Biology and Research Section, Hospital de Tortosa Verge de la Cinta, Institut d'Investigacio Sanitaria Pere Virgili (IISPV), URV, Spain, and from the Pathology Department of the same hospital. The dataset was formerly acquired during the project [36], grant number PI11/0488, conducted at the Instituto de Salud Carlos III, Spain, approved by the Ethics Committee of the Hospital Joan XXIII de Tarragona

Fig. 3 Example of recurrent call of the proposed method. Example of recurrent call of proposed method where threshold of distance transform cutoff is increased resulting in cluster separation in second iteration. In this case, the processing would stop after second iteration, but here, every possible step of recurrent call is presented. From left to right: binary image of the cluster, contours estimated with all recurrent calls of proposed method, the same contours shown in 3D

(reference 22p/2011). The subset used in this study was obtained based on the international cooperation within grant number 2013/11/N/ST7/02797 (Polish National Science Center).

The histological tissue microarrays used for image acquisition were prepared from formalin-fixed, paraffin-embedded tissue blocks of the breast, and auxiliary node biopsies coming from the patients with four major molecular subtypes [37]: Luminal A, luminal B, triple negative, and HER2-enriched. Tissue sample preparation is not crucial for this investigation and was extensively described before [38].

The digital images were captured under $\times$ 40 magnification on an automated whole slide scanning system (Aperio ScanScope XT). Then, they were split into images of separate punches using two designated programs [39, 40]. Although the images are originally 3-channel Red-Green-Blue (RGB) images, we converted them to Hue-Saturaton-Value color-space and used only value layer for processing, similar to the previous research [41].

All methods presented in this study were applied to the dataset that consisted of a total of 7557 nuclei within 13 randomly selected regions of interest (ROI) extracted from the dataset. Images differ in the degree of complexity, architecture compactness, global contrast, and brightness, which is typical for this type of biological data. To crop ROI out of the original tissue punches, we created two virtual circles, centered with tissue punch, with radius of 500 and 2000 pixels. We randomly draw from the coordinates of the points of circles' circumference. This point becomes the upper-left corner of the ROI with size 1000×1000 pixels.

For evaluation purposes, based on the manual annotations performed by the two experts, ground truth templates were generated, see Fig. 1. Each expert marked every location of the positive and negative nuclei with an indicator for object-wise evaluation. Moreover, for pixel-wise evaluation, in each ROI, every immunopositive and 31 randomly selected immunonegative objects' boundaries were marked thereby creating binary mask. The mean of both experts' annotation was assumed to be the ground truth, creating a set of 808 manually marked nuclei (405 immunopositive and 403 immunonegative).

2.3 Evaluation

In this study, we conducted three types of evaluation. Our results are compared with the ground truth and the results of reference methods, both object-wise and pixel-wise. As there is no database with well-known ground truths (benchmark) that are publicly available, the method evaluation is performed on the available experimental material with the ground truth annotated by the experts. Moreover, we compared performance times of the proposed method with that of the state of the art.

In total, there were 7557 manually marked nuclei in 13 images. For randomly selected 808 nuclei, the boundary was marked as well. True positive (TP) objects are those which have one matching object in manually labeled ground truth. A segmented nuclei region is counted as matched if the manual mark is localized within its area.

For the purposes of evaluation on the level of objects, we used the most common criteria including precision or positive predictive value (PPV), recall (TPR), and F1 score. PPV is the ratio of TPs to the number of detected objects (PPV = TP/(TP+FP)). Sensitivity or true positive rate (TPR) shows what proportion of the objects of interest is found (TPR =

TP/(TP+FN)). In case of both metrics, values closer to 1 imply better outcome. F1 score is the harmonic average of precision and recall which ranges between 0 and 1, where higher values represent better method.

Unfortunately, the metrics described above are not always sufficient to discriminate between the two methods. We calculated the following additional pixel-wise criterions introduced by Cui et al. [15]: missing detection rate (MDR), false detection rate (FDR), under-segmentation rate (USR), and over-segmentation rate (OSR). This approach divides the cause for false positives into two types of errors: false detections and over-segmentation. Furthermore, false negatives could be split into two categories: missing detections and under-segmentation. Figure 4 presents examples of all cases.

In addition, we used the Jaccard index to measure how closely the boundary fits the 808 nuclei reference ground truth.
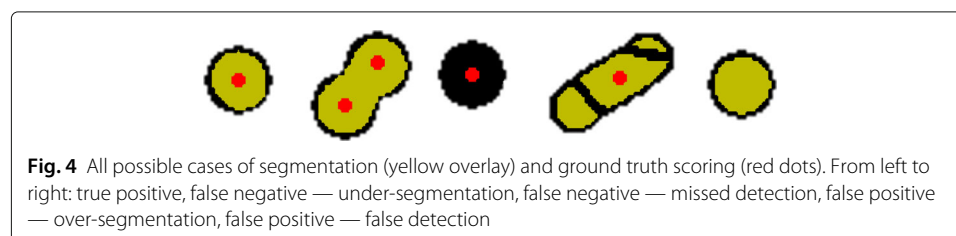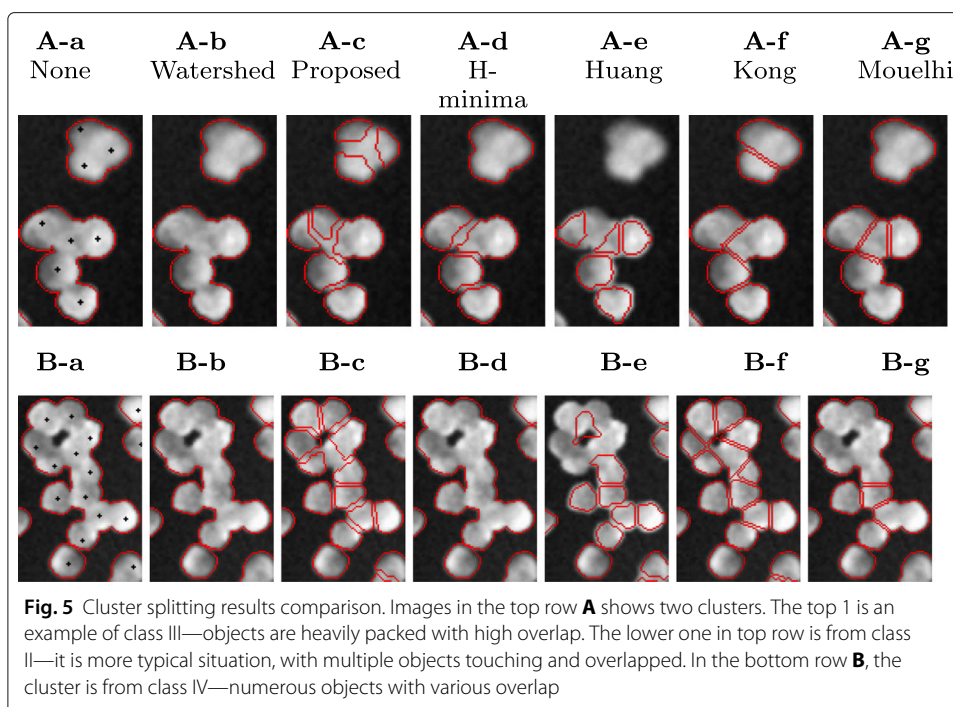
3 Results

In this study, we measured the performance of the algorithm on microscopic images by comparing the number of nuclei and their segmentation results (binary images) with the annotated ground truth. The performance was evaluated using object-wise and pixel-wise metrics.

The process of cluster splitting was first analyzed by the performance comparison on BBBC0004 dataset. It consists of synthetic images, which allows for the most controlled environment for assessing algorithms' performance. The dataset is divided into five subsets with increasing overlap probability, from 0 to 0.6; examples are presented in Fig. 5. Based on our analysis, we were able to distinguish strengths and weaknesses of the cluster splitting methods in relation to increasing clustering. Table 2 presents the comparison.

Furthermore, we evaluated the methods on real images from IISPV dataset. The process of cluster splitting was performed on the results of Bradley thresholding algorithm in 13 ROI. The initial detection was performed by applying threshold to value layer of Hue-Saturation-Value color-space. The mean ratios of possible errors, presented in Fig. 4, were calculated for each method. To prove the superiority of the proposed method, the results were compared with those of segmentation without the application of cluster splitting method and with the results of the baseline seeded watershed algorithm and other methods from the literature. Alongside quality metrics, we also compared the time consumption for the evaluated cluster splitting methods. Table 3 presents the comparison.

In the second part of Table 3, we present the comparison of the proposed method and state-of-the-art frameworks for the quantification in histopathology: *Tmarker* and *Qupath*. We present two versions of results for the latter. First, we used default parameter

Fig. 4 All possible cases of segmentation (yellow overlay) and ground truth scoring (red dots). From left to right: true positive, false negative — under-segmentation, false negative — missed detection, false positive — over-segmentation, false positive — false detection

Fig. 5 Cluster splitting results comparison. Images in the top row **A** shows two clusters. The top 1 is an example of class III—objects are heavily packed with high overlap. The lower one in top row is from class II—it is more typical situation, with multiple objects touching and overlapped. In the bottom row **B**, the cluster is from class IV—numerous objects with various overlap

values (qupath_def). Second, the expert with experience in histopathology and this software adjusted manually the processing parameters to achieve best results (qupath_exSep). We also compared the results with widely available *ImageJ* plugin named *ICHtoolbox* that uses oval-fitted nuclei segmentation and nuclei segmentation and quantification functions are automatically processed. Table 3 also presents the results of DL approaches. The model pretrained by Chen et al. was tested as is. The results of inference were extracted as binary map. The U-net model achieved the following results: loss: 0.0286 and dice coefficient 0.9393, after 500 epochs, on ScienceBowl2018 dataset. The results of inference on IISPV dataset were thresholded with typical in this situation value equal to half of the maximum intensity value, namely 127, to get the binary output maps used for evaluation.

4   Discussion

The performance comparison in controlled environment of the BBBC0004 dataset gave us much insight in strengths and drawbacks of the proposed method. First, we proved that all methods have similar results while there is no overlap (*ov00*) and the objects are only touching (as presented in Table 2 in section *ov00*). With increase in the overlap, the effectiveness diminished in terms of all metrics and all methods. When some of the objects overlap (*ov15* and *ov30*), our method achieved best results with TPR, PPV, and F1-score (over 0.95). For images with more overlapping objects creating complex structures, the Huang method had best recall (TPR) value. However, our proposed method, with much better precision, achieved superior F1-score, as presented in Table 2. As the cluster splitting along with segmentation is usually the preliminary step for object classification, it is critical to lower the missed detections as much as possible. This is because it is possible to discriminate the false positive objects during further processing, whereas false negative (missed) objects will not be further processed.

Table 2 Comparison of performance of cluster splitting methods on BBBC0004 dataset (None—images without applied cluster splitting method)

		TPR	PPV	F1	MDR	FDR	USR	OSR
ov00	None	0.968	1.000	0.984	0.00	0.00	0.03	0.00
	Watershed	0.978	1.000	0.989	0.00	0.00	0.02	0.00
	Proposed method	0.989	1.000	0.995	0.00	0.00	0.01	0.00
	H-minima	0.989	1.000	0.995	0.00	0.00	0.01	0.00
	Huang	0.989	0.999	0.994	0.00	0.00	0.01	0.00
	Kong	0.989	0.999	0.994	0.00	0.00	0.01	0.00
	Mouelhi	0.968	1.000	0.984	0.00	0.00	0.03	0.00
ov15	None	0.793	0.997	0.884	0.00	0.00	0.21	0.00
	Watershed	0.811	0.997	0.894	0.00	0.00	0.19	0.00
	Proposed method	0.970†	0.970	0.970†	0.01	0.00	0.02	0.02
	H-minima	0.934	0.991	0.961	0.00	0.00	0.06	0.01
	Huang	0.963	0.955	0.959	0.02	0.00	0.01	0.04
	Kong	0.955	0.972	0.963	0.01	0.00	0.04	0.03
	Mouelhi	0.826	0.995	0.902	0.00	0.00	0.17	0.00
ov30	None	0.723	1.000	0.839	0.00	0.00	0.28	0.00
	Watershed	0.739	1.000	0.850	0.00	0.00	0.26	0.00
	Proposed method	0.959†	0.986	0.972†	0.01	0.00	0.03	0.01
	H-minima	0.890	0.999	0.941	0.00	0.00	0.11	0.00
	Huang	0.950	0.963	0.956	0.04	0.01	0.02	0.03
	Kong	0.926	0.982	0.953	0.00	0.00	0.07	0.02
	Mouelhi	0.774	0.996	0.871	0.00	0.00	0.22	0.00
ov45	None	0.700	1.000	0.824	0.00	0.00	0.30	0.00
	Watershed	0.716	1.000	0.834	0.00	0.00	0.28	0.00
	Proposed method	0.942	0.993	0.967†	0.01	0.00	0.05	0.01
	H-minima	0.873	0.999	0.932	0.00	0.00	0.13	0.00
	Huang	0.946†	0.981	0.963	0.03	0.00	0.02	0.02
	Kong	0.930	0.992	0.960	0.00	0.00	0.07	0.01
	Mouelhi	0.756	0.999	0.861	0.00	0.00	0.24	0.00
ov60	None	0.610	1.000	0.757	0.00	0.00	0.39	0.00
	Watershed	0.626	1.000	0.770	0.00	0.00	0.37	0.00
	Proposed method	0.909	0.983	0.945†	0.01	0.00	0.08	0.02
	H-minima	0.809	0.993	0.891	0.00	0.00	0.19	0.01
	Huang	0.921†	0.973	0.945†	0.05	0.00	0.03	0.03
	Kong	0.893	0.964	0.927	0.01	0.00	0.09	0.04
	Mouelhi	0.696	0.993	0.818	0.00	0.00	0.30	0.01

Best values are marked with †

Moving on from synthetic to real images, we validated the superiority of our proposed method on IISPV dataset. Based on the comparison of the performance on IISPV dataset, it was established that the proposed method outperforms other method, both those based on boundary shape and those incorporating the intensity values. We can assume that nuclei detection was best because of the highest F1 score (0.734).

Table 3 Performance of the proposed method in comparison to the state-of-the-art frameworks

	Methods	TPR	PPV	F1	MDR	FDR	USR	OSR	Jaccard	Time
	None	0.647	0.766	0.686	0.03	0.15	0.33	0.08	0.72	
	Watershed	0.723	0.754	0.727	0.04	0.14	0.25	0.09	0.72	155
	Proposed method	0.839	0.667	0.734	0.06	0.13	0.11	0.18	0.71	105
Classical	H-minima	0.738	0.729	0.724	0.05	0.14	0.23	0.12	0.71	129
	Huang	0.676	0.513	0.576	0.31	0.25	0.02	0.19	0.52	77
	Kong	0.785	0.611	0.672	0.06	0.11	0.17	0.25	0.70	277
	Mouelhi	0.778	0.587	0.652	0.07	0.12	0.16	0.26	0.70	292
	IJ_IHCtoolbox	0.541	0.848	0.643	0.43	0.03	0.05	0.12	0.63	
Frameworks	Qupath_def	0.796	0.510	0.612	0.16	0.25	0.05	0.20	0.66	
	Qupath_exSep	0.869	0.576	0.682	0.08	0.21	0.06	0.18	0.74	
	Tmarker	0.265	0.587	0.334	0.72	0.37	0.05	0.03	0.43	
DL	Chen	0.776	0.700	0.724	0.20	0.25	0.04	0.04	0.55	
	U-net	0.459	0.724	0.555	0.50	0.20	0.09	0.07	0.64	

To differentiate the methods even more precisely, we compared metrics focused on cluster splitting. They rely on object numbers, yet they inform about pixel-level statistics. As presented in Table 3, the use of our proposed method significantly lowers the value of USR. The slight increase in the value of OSR could be remedied by introducing subsequent post-processing step. In general, improvement in terms of under-segmentation is always at the cost of deterioration in terms of over-segmentation. The boundary fitting assessed by Jaccard index stays on the similar level for all methods (except "Huang"). In summary, even though reference methods have better values of OSR and USR, the proposed method is the most balanced solution as proved by superior F1 score results.

According to time evaluation, our proposed method is only outperformed by "Huang" method. Because time evaluation is based on a batch of images with size 1000×1000 pixels, "Kong" and "Mouelhi" methods take longer times, which is discouraging. Apart from these two, the other methods have comparable time of processing.

To prove the reliability of the proposed method in this study, we also compared it to the known state-of-the-art frameworks that are capable of producing quantification and objects' boundary mask for digital histopathology. In this comparison, only Qupath with extensive expert's manual parameter tuning was able to outperform the proposed method according to some metrics, as presented in Table 3. Tmarker's results were drastically low in all metrics. The ImageJ plugin has rather satisfactory results in terms of over- and under-segmentation, but on the downside, there are a lot of missed detections. Overall, our approach has best F1 score and MDR with very good results in terms of false detections.

Nowadays, DL is applied to a variety of computer vision problems, among which is nuclei detection and segmentation. We acknowledge the fact that it is a very powerful tool with vast capabilities, but what we tried to show is that it is not a simple task to adapt DL to a new or even a modified problem. Herein, we performed inference with network pretrained on H&E tissue images (*Chen*). The results are relatively good but do not outperform the proposed method in this study. However, tackling the problem entirely from

the starting point and training the deep network from scratch requires a vast amount of resources and labeled data. With limited access to these resources, the results are quite poor. In summary, establishing a DL model with good performance is not a simple solution while tackling the problem with limited data, such as tissue sections stained with DAB&H from IISPV dataset.

Finally, based on our results, we can say that the proposed method in this study outperforms other methods from literature and state-of-the-art frameworks.

5   Conclusion

The accuracy of the employed automated nuclei segmentation technique is critical in obtaining high-quality and efficient diagnostic performance. Unfortunately, algorithms based on thresholding produce results that almost never perfectly fit to real object's boundary. In addition, the situation where multiple objects are connected (clustered) and counted as one is frequently observed. Therefore, we proposed novel cluster splitting method.

We provide the efficient method for dividing the clustered nuclei in digital images of histopathological slides where results of initial nuclei detection are improved by applying cluster splitting methods. The proposed method in this study achieved better results in terms of average F1 score (0.734) than that of all other referenced methods: seeded watershed, "Huang", H-minima, "Mouelhi", and "Kong." We managed to keep very low both USR and OSR.

Moreover, the established state-of-the-art frameworks also did not outperform our proposed method while automatically processing the data. Only Qupath with manual parameter tuning was able to achieve superior results. Using the DL technique for the instance segmentation of overlapping cell nuclei is not a simple task and requires vast amount of resources and labeled data. Simply using the available solutions results in mediocre performance, especially with even slightly different data (H&E vs. DAB&H staining) as proved in this study.

The primary disadvantage of the proposed method is the strict categorization of the clusters. It is based on the time-consuming statistical analysis of the processed database. In future, we plan to omit this problem with automatic parameter setting. We also plan to test the algorithm on other datasets consisting of different types of cells.

In summary, the primary achievement of this study is the establishment of the processing step that could be incorporated into image processing framework to improve the results of segmentation .

Authors' contributions
Conceptualization, L.R. and A.K.; methodology, L.R.; software, L.R.; validation, L.R.; formal analysis, L.R. and A.K.; investigation, L.R.; resources, C.L., M.L., and R.B.; data curation, C.L., M.L., and R.B.; writing—original draft preparation, L.R.; writing—review and editing, L.R. and A.K. and D.P.; visualization, L.R.; supervision, D.P.; project administration, L.R. and D.P.; funding acquisition, L.R. and A.K. and D.P. The authors read and approved the final manuscript.

Availability of data and materials
The dataset BBBC0004 [2] analyzed during the current study is available in the Broad Bioimage Benchmark Collection repository, https://data.broadinstitute.org/bbbc/BBBC004/. The dataset IISPV used during the current study is available from the corresponding author on reasonable request. Both datasets and the data that support the findings of this study will be available via MIAP platform [30] upon publication of the manuscript.

Competing interests
The authors declare that they have no competing interests. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Author details
[1]Nalecz Institute of Biocybernetics and Biomedical Engineering Polish Academy of Sciences, Ks. Trojdena 4 Str., 02-109 Warsaw, Poland. [2]Pathology Department, Hospital de Tortosa Verge de la Cinta, Tortosa, Spain. [3]Molecular Biology and Research Section, Hospital de Tortosa Verge de la Cinta, Tortosa, Spain.

References
1.   S. Nam, Y. Chong, C. K. Jung, T.-Y. Kwak, J. Y. Lee, J. Park, M. J. Rho, H. Go, Introduction to digital pathology and computer-aided pathology. J. Pathol. Transl. Med. **54**(2), 125–134 (2020). https://doi.org/10.4132/jptm.2019.12.31
2.   B. Institute, Synthetic cells. https://data.broadinstitute.org/bbbc/BBBC004/. Accession number BBBC004 ·Version 1
3.   M. Takenaka, N. Seki, U. Toh, S. Hattori, A. Kawahara, T. Yamaguchi, K. Koura, R. Takahashi, H. Otsuka, H. Takahashi, N. Iwakuma, S. Nakagawa, T. Fujii, T. Sasada, R. Yamaguchi, H. Yano, K. Shirouzu, M. Kage, FOXP3 expression in tumor cells and tumor-infiltrating lymphocytes is associated with breast cancer prognosis. Mol. Clin. Oncol. **1**(4), 625–632 (2013). https://doi.org/10.3892/mco.2013.107
4.   M. Salvi, F. Molinari, Multi-tissue and multi-scale approach for nuclei segmentation in h&e stained images. BioMed. Eng. OnLine. **17**(1) (2018). https://doi.org/10.1186/s12938-018-0518-0
5.   Z. Swiderska-Chadaj, T. Markiewicz, B. Grala, M. Lorent, Content-based analysis of Ki-67 stained meningioma specimens for automatic hot-spot selection. Diagn. Pathol. **11**(1) (2016). https://doi.org/10.1186/s13000-016-0546-7
6.   J. Cheng, J. C. Rajapakse, Segmentation of clustered nuclei with shape markers and marking function. IEEE Trans. Biomed. Eng. **56**(3), 741–748 (2009). https://doi.org/10.1109/tbme.2008.2008635
7.   S. Ali, R. Veltri, J. I. Epstein, C. Christudass, A. Madabhushi, in *Lecture Notes in Computer Science*, Adaptive energy selective active contour with shape priors for nuclear segmentation and Gleason grading of prostate cancer (Springer, Berlin Heidelberg, 2011), pp. 661–669. https://doi.org/10.1007/978-3-642-23623-5_83
8.   S. Wienert, D. Heim, K. Saeger, A. Stenzinger, M. Beil, P. Hufnagl, M. Dietel, C. Denkert, F. Klauschen, Detection and segmentation of cell nuclei in virtual microscopy images: a minimum-model approach. Sci. Rep. **2**(1) (2012). https://doi.org/10.1038/srep00503
9.   C. Yan, H. Xie, J. Chen, Z. Zha, X. Hao, Y. Zhang, Q. Dai, A fast Uyghur text detector for complex background images. IEEE Trans. Multimed. **20**(12), 3389–3398 (2018). https://doi.org/10.1109/tmm.2018.2838320
10. C. Yan, L. Li, C. Zhang, B. Liu, Y. Zhang, Q. Dai, Cross-modality bridging and knowledge transferring for image understanding. IEEE Trans. Multimed. **21**(10), 2675–2685 (2019). https://doi.org/10.1109/tmm.2019.2903448
11. C. Yan, Y. Tu, X. Wang, Y. Zhang, X. Hao, Y. Zhang, Q. Dai, STAT: spatial-temporal attention mechanism for video captioning. IEEE Trans. Multimed. **22**(1), 229–241 (2020). https://doi.org/10.1109/tmm.2019.2924576
12. H. Irshad, A. Veillard, L. Roux, D. Racoceanu, Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential. IEEE Rev. Biomed. Eng. **7**, 97–114 (2014). https://doi.org/10.1109/rbme.2013.2295804
13. D. C. Cireşan, A. Giusti, L. M. Gambardella, J. Schmidhuber, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*. ed. by K. Mori, I. Sakuma, Y. Sato, C. Barillot, and N. Navab, Mitosis detection in breast cancer histology images with deep neural networks (Springer, Berlin, Heidelberg, 2013), pp. 411–418
14. Y. Xie, X. Kong, F. Xing, F. Liu, H. Su, L. Yang, in *Lecture Notes in Computer Science*, Deep voting: a robust approach toward nucleus localization in microscopy images (Springer, International Publishing, 2015), pp. 374–382. https://doi.org/10.1007/978-3-319-24574-4_45
15. Y. Cui, G. Zhang, Z. Liu, Z. Xiong, J. Hu, A deep learning algorithm for one-step contour aware nuclei segmentation of histopathological images. arXiv (2018). 1803.02786v1
16. P. Wild, N. Rupp, J. Buhmann, P. Schüffler, T. Fuchs, C. Ong, TMARKER: a free software toolkit for histopathological cell counting and staining estimation. J. Pathol. Inform. **4**(2), 2 (2013). https://doi.org/10.4103/2153-3539.109804
17. J. Shu, H. Fu, G. Qiu, P. Kaye, M. Ilyas, in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Segmenting overlapping cell nuclei in digital histopathology images (IEEE, Osaka, 2013). https://doi.org/10.1109/embc.2013.6610781
18. P. Bankhead, M. B. Loughrey, J. A. Fernández, Y. Dombrowski, D. G. McArt, P. D. Dunne, S. McQuaid, R. T. Gray, L. J. Murray, H. G. Coleman, J. A. James, M. Salto-Tellez, P. W. Hamilton, QuPath: open source software for digital pathology image analysis. Sci. Rep. **7**(1) (2017). https://doi.org/10.1038/s41598-017-17204-5
19. A. Mouelhi, M. Sayadi, F. Fnaiech, in *2011 International Conference on Communications, Computing and Control Applications (CCCA)*, Automatic segmentation of clustered breast cancer cells using watershed and concave vertex graph (IEEE, 2011). https://doi.org/10.1109/ccca.2011.6031229

20. H. Kong, M. Gurcan, K. Belkacem-Boussaid, in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, Splitting touching-cell clusters on histopathological images (IEEE, 2011). https://doi.org/10.1109/isbi.2011.5872389
21. H. Kong, M. Gurcan, K. Belkacem-Boussaid, Partitioning histopathological images: an integrated framework for supervised color-texture segmentation and cell splitting. IEEE Trans. Med. Imaging. **30**(9), 1661–1677 (2011). https://doi.org/10.1109/tmi.2011.2141674
22. I. The MathWorks, Marker-controlled watershed segmentation (2019). https://www.mathworks.com/help/images/marker-controlled-watershed-segmentation.html. Online accessed 08 Mar 2020
23. P.-W. Huang, Y.-H. Lai, Effective segmentation and classification for HCC biopsy images. Pattern Recogn. **43**(4), 1550–1563 (2010). https://doi.org/10.1016/j.patcog.2009.10.014
24. C. Jung, C. Kim, Segmenting clustered nuclei using h-minima transform-based marker extraction and contour parameterization. IEEE Trans. Biomed. Eng. **57**(10), 2600–2604 (2010). https://doi.org/10.1109/tbme.2010.2060336
25. K. Chen, N. Zhang, L. Powers, J. Roveda, in *2019 Spring Simulation Conference (SpringSim)*, Cell nuclei detection and segmentation for computational pathology using deep learning (IEEE, 2019). https://doi.org/10.23919/springsim.2019.8732905
26. B. A. Hamilton, Dataset: Science Bowl 2018 (2018). https://www.kaggle.com/c/data-science-bowl-2018/data. Accessed 28 Feb 2019
27. KChen89, Cell-nuclei-detection-and-segmentation. https://github.com/KChen89/Cell-Nuclei-Detection-and-Segmentation. Accessed 28 Feb 2019
28. O. Ronneberger, P. Fischer, T. Brox, in *Lecture Notes in Computer Science*, U-net: convolutional networks for biomedical image segmentation (Springer, International Publishing, 2015), pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
29. kamalkraj, DATA-SCIENCE-BOWL-2018 (2018). https://github.com/kamalkraj/DATA-SCIENCE-BOWL-2018
30. T. Markiewicz, A. Korzynska, A. Kowalski, Z. Swiderska-Chadaj, P. Murawski, B. Grala, M. Lorent, M. Wdowiak, J. Zak, L. Roszkowiak, W. Kozlowski, D. Pijanowska, MIAP – web-based platform for the computer analysis of microscopic images to support the pathological diagnosis. Biocybernetics Biomed. Eng. **36**(4), 597–609 (2016). https://doi.org/10.1016/j.bbe.2016.06.006
31. J. Lukasz Roszkowiak, K. Zak, D. Siemion, A. Pijanowska, *Korzynska in Computer Vision and Graphics: Proc. ICCVG 2020, Nuclei detection with local threshold processing in DAB&H stained breast cancer biopsy images*. (L. J. Chmielewski, R. Kozera, A. Orłowski, eds.) (Springer International Publishing, 2020)
32. D. Bradley, G. Roth, Adaptive thresholding using the integral image. J. Graph. Tools. **12**(2), 13–21 (2007). https://doi.org/10.1080/2151237X.2007.10129236. https://doi.org/10.1080/2151237X.2007.10129236
33. P. Ruusuvuori, A. Lehmussola, J. Selinummi, T. Rajala, H. Huttunen, O. Yli-Harja, in *2008 16th European Signal Processing Conference*, Benchmark set of synthetic images for validating cell image analysis algorithms (IEEE, Lausanne, 2008), pp. 1–5. https://ieeexplore.ieee.org/abstract/document/7080746
34. V. Ljosa, K. L. Sokolnicki, A. E. Carpenter, Annotated high-throughput microscopy image sets for validation. Nat. Methods. **9**, 637 (2012)
35. A. Lehmussola, P. Ruusuvuori, J. Selinummi, T. Rajala, O. Yli-Harja, Synthetic images of high-throughput microscopy for validation of image analysis methods. Proc. IEEE. **96**(8), 1348–1360 (2008). https://doi.org/10.1109/JPROC.2008.925490
36. C. Lopez, C. Callau, R. Bosch, A. Korzynska, J. Jaen, M. Garcia-Rojo, G. Bueno, M. T. Salvado, T. Alvaro, M. Onos, M. d. M. Fernandez-Carrobles, M. Llobera, J. Baucells, G. Orero, M. Lejeune, Development of automated quantification methodologies of immunohistochemical markers to determine patterns of immune response in breast cancer: a retrospective cohort study. BMJ Open. **4**(8), 005643–005643 (2014). https://doi.org/10.1136/bmjopen-2014-005643
37. N. Snoj, P. Dinh, P. Bedard, C. Sotiriou, *Molecular Biology of Breast Cancer*. (W. Coleman, G. Tsongalis, eds.), vol. 9780123744180. (Academic Press, Cambridge, 2010), pp. 341–349. https://doi.org/10.1016/B978-0-12-374418-0.00026-8
38. A. Korzynska, L. Roszkowiak, C. Lopez, R. Bosch, L. Witkowski, M. Lejeune, Validation of various adaptive threshold methods of segmentation applied to follicular lymphoma digital images stained with 3,3'-diaminobenzidine&haematoxylin. Diagn. Pathol. **8**(1), 1–21 (2013). https://doi.org/10.1186/1746-1596-8-48
39. M. Milagro Fernandez-Carrobles, G. Bueno, O. Deniz, J. Salido, M. Garcia-Rojo, L. Gonzalez-Lopez, A cad system for the acquisition and classification of breast TMA in pathology. Stud. Health Technol. Informat. **210**, 756–760 (2015). https://doi.org/10.3233/978-1-61499-512-8-756
40. L. Roszkowiak, C. Lopez, PATMA: parser of archival tissue microarray. PeerJ. **4**, 2741 (2016). https://doi.org/10.7717/peerj.2741
41. L. Roszkowiak, A. Korzynska, D. Pijanowska, in *Proceedings of the 2015 Federated Conference on Computer Science and Information Systems*, Short survey: adaptive threshold methods used to segment immunonegative cells from simulated images of follicular lymphoma stained with 3,3'-diaminobenzidine&haematoxylin (IEEE, Lodz, 2015), pp. 291–296. https://doi.org/10.15439/2015f263

Publisher's Note