

# Adaptive Method of Glands Segmentation on Histological Images <sup>\*</sup>

Alexander Kosov <sup>[0000-0003-0219-8105]</sup>, Alexander Khvostikov <sup>[0000-0002-4217-7141]</sup>, and  
Andrey Krylov <sup>[0000-0001-9910-4501]</sup>

Faculty of Computational Mathematics and Cybernetics,  
Lomonosov Moscow State University, Moscow, Russia  
mos.cap14@yandex.ru, khvostikov@cs.msu.ru,  
kryl@cs.msu.ru

**Abstract.** Glands segmentation is a very important and yet challenging problem of histological images analysis. Accurate segmentation of mucous glands is a crucial step to obtain reliable morphological statistics and is necessary for the development of high-quality diagnostic algorithms, which are an integral part of timely medical care. In this paper we propose a two-stage segmentation method, which predicts the probability maps of glands boundaries in histological images based on a priori knowledge about the geometric shape of the mucous glands and uses a convolutional neural network (CNN) model to get the final segmentation result based on the predicted probability maps. The proposed method demonstrates good results in separating adjacent glands, which is one of the most challenging aspects in automatic segmentation of histological glands and one of the most complicated for algorithms based on applying convolutional neural networks. The evaluation of the proposed algorithm was performed with *Warwick-QU* dataset, which contains real histological images of colon tissue.

**Keywords:** Image Segmentation, Mucous Glands, Deep Learning, Convolutional Neural Networks, Histology, Pathology

## 1 Introduction

Over the past decade a notable reduction in colorectal cancer mortality was achieved. A significant role in this was played by the appearance of automated systems for segmentation and analysis of histological images which allow to obtain accurate results for large amount of data in a short period of time.

In this paper we propose an adaptive method of glands segmentation on histological images that considers the form of histological glandular structures and evaluate the implementation of the proposed method on a real histological dataset.

---

Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

\* The work was funded by RFBR, CNPq and MOST according to the research project 19-57-80014 (BRICS2019-394)

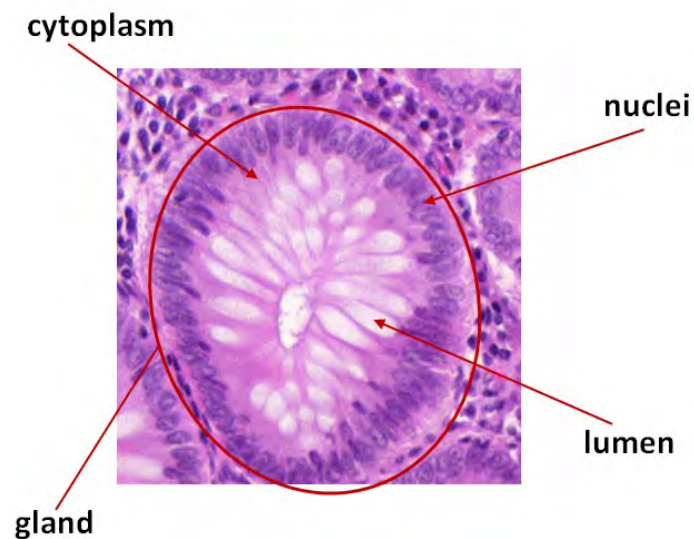
All existing techniques of histological gland segmentation can be divided into three main groups:

1. methods based only on mathematical principles of image processing,
2. methods that are built upon classical image processing methods with embedding of machine learning approaches at one or several stages,
3. methods that are entirely based on neural network model usage.

More detailed overview of existing histological image segmentation methods is given below.

*Classical methods that do not use machine learning* are based mainly on texture analysis of data. For example, in [1] intensity dispersion and Gaussian filters were used to extract structural features from the glands structures. After that K-means clustering [2] was applied. The main disadvantage of such methods is the lack of prior knowledge about the relationships between the glands structural components, therefore these methods can lead to incorrect segmentation.

*Methods that use classical methods and machine learning in conjunction* analyze structural features of glands using a prior knowledge about their relationships [3,4]. This allows to avoid problems that arise in texture based analysis methods.



**Fig. 1.** Typical parts of mucous glands on histological image.

In [5] at the initial stage the Bayesian classifier is used to segment lumen areas (the bright region in the center of the gland), cytoplasm and nuclei (Fig. 1). Then, the

levelset method [6] is applied to the previously classified areas to find desired gland contour. Considering the information about grandular structure allows to achieve better results in comparison to texture based methods, however, a serious limitation of this method lies in the inaccuracy of levelset method. The reason for these inaccuracy is insufficiently smooth and uniform lumen boundary that breaks the levelset method.

The authors of [7] proposed a more effective approach, where boundaries of glands are determined by using the stochastic polygons method. At the initial stage, nuclei are segmented. They are further considered as possible vertices of the desired boundary polygonal approximation. After that a Bayesian-based dropout is performed. Decision of the dropout is based on information about the spatial connectivity and curvature of the contours that can be formed by considered vertexes.

*Methods fully based on the neural network model usage* are represented with works [8–10]. The results in these methods were obtained after passing the initial image through a pre-trained neural network. The architectures used in such methods are quite diverse. Neural networks have great generalizing ability that allows to achieve segmentation with accuracy in some cases close to 1.

Hereinafter, by the accuracy we mean the ratio of the number of correctly segmented image pixels to the total number of pixels.

In [8] multi-level contextual features with auxiliary supervision detection for generating likelihood maps of glands were used. They were found with end-to-end trained fully convolutional network (FCN). Next, deep-contouraware network drawn from FCN for effective gland segmentation by fusing the complementary information of objects and contours was used. This allowed this method to take the first place in *2015 MICCAI Gland Segmentation Challenge* [11, 7].

In [10] two convolutional networks were used. A segmentation network takes an input image and produces segmentation probability maps for this image. An evaluation network takes the segmentation probability maps and the corresponding input image and determines a score indicating the quality of the segmentation.

The main feature of [10] is the use of unannotated data. The authors achieved *Dice score* equal to 0.903 on the *Warwick-QU* dataset:

$$Dice(A, B) = \frac{2 \cdot |A \cap B|}{|A| + |B|}, \quad (1)$$

In [9] the segmentation process was divided into several base stages:

- foreground segmentation stage in which the separation of glands from the background was done,
- edge detection stage in which the glands borders were highlighted,
- object detection stage in which glands were separated from each other,
- fusion stage in which the final segmentation is obtained by combining the data obtained in the previous stages.

All stages were performed with convolutional neural networks.

The authors achieved *Dice score* equal to 0.908 on the *Warwick-QU* dataset.

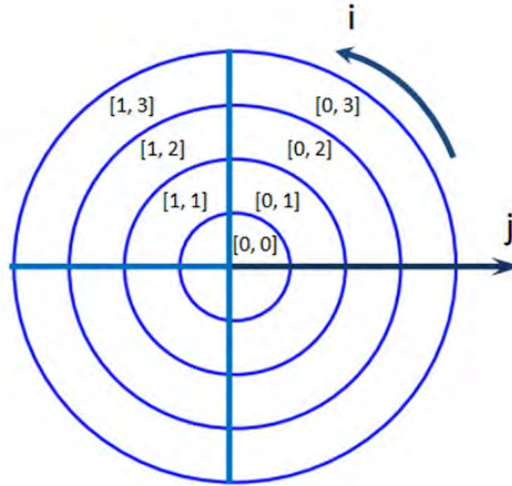
In this paper we propose a method based on a priory information knowledge about structural features of glands. The proposed method constructs a probability map of

glands position in the image and performs post-processing on predicted results receiving initial coarse segmentation. After that the received areas, that are candidates to be glands, are classified with a convolutional neural network.

## 2 Proposed method

The glandular structures are predominantly elliptical in shape. Their borders are expressed in dark color (based on *Warwick-QU*), that is the main criterion for their detection within the proposed method.

Our method consists of two main steps: calculating glands contour probability map that is described in *Section 2.1* and classification and postprocessing of gland candidates that is described in *Section 2.2*.



**Fig. 2.** Visualization of a template assembled from circular sectors.

### 2.1 Calculating glands contour probability map

The proposed method, as mentioned earlier, consists of two stages. At the initial stage a uniform grid is placed on the image. After that the center of the described below template is placed in the each of grid nodes. Template represents a circle that is uniformly divided into circular sectors (Fig. 2). Directed weighted graph consisting of vertices lying in the rays is constructed with the help of this template. Each circular sector corresponds to a single vertex lying in the ray corresponding to this sector. Total number of vertices is calculated as:

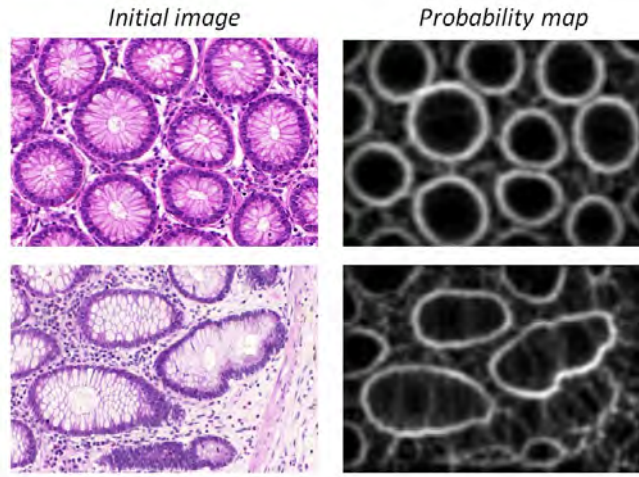
$$N_{vert} = N_{rays} \cdot N_{v.on.ray}, \quad (2)$$

where  $N_{vert}$  is total number of vertices,  $N_{rays}$  is total number of rays and  $N_{v.on.ray}$  is number of vertices on the each ray.

The weight of each vertex is calculated as the average intensity of image pixels lying in corresponding to this vertex circular sector:

$$W_{ij} = \frac{1}{|\Omega_{ij}|} \sum_{u,v \in \Omega_{ij}} I(u,v), \quad (3)$$

where  $\Omega_{ij}$  is circular sector corresponding to the vertex  $(i, j)$  and  $I(u, v)$  is the intensity of the image pixel with coordinates  $(u, v)$ .



**Fig. 3.** Initial images from *Warwick-QU* dataset on the left and the probability maps calculated for these images on the right.

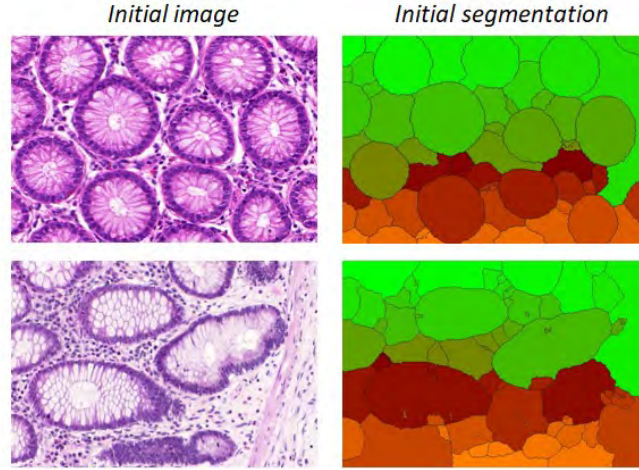
Next, for each graph constructed for grid node a closed cycle with the lowest total weight is searched in the graph. This cycle, transferred to the template, corresponds to the expected gland boundary contour for this node.

Contours obtained for each of the nodes overlap each other thus forming a probability map (Fig. 3) for the glands position in the image. This map is used as a heights map to obtain initial segmentation using *Watershed* algorithm [12].

The initial segmentation (Fig. 4) is built without taking into account whether areas belong to the glands or not, therefore it is necessary to throw out inappropriate areas. Filtering these areas is done on the second stage of the proposed algorithm.

## 2.2 Classification and postprocessing of gland candidates

To perform filtering of glands candidates each of the initial segmentation areas are feed to a pretrained convolutional neural network that estimates the probability of belonging



**Fig. 4.** Initial images from *Warwick-QU* dataset on the left and initial segmentation obtained from a probability maps using the *Watershed* algorithm on the right.

this area to the 'gland' class. Based on the information received, the final segmentation is constructed by eliminating areas that do not belong to 'gland' class. Area is considered as 'not a gland' if the probability predicted by the convolutional neural network is less than 0.5.

The architecture of the used convolutional neural network is sequential and consists of three convolutional blocks and two fully connected layers (Fig. 5). Each convolutional block consists of two convolutional layers as well as a pooling layer and batch normalization layer. First two convolutional blocks are used to highlight features in the image. They use  $3 \times 3$  convolution. First pooling layer use  $4 \times 4$  AVG pooling. Second pooling layer use  $2 \times 2$  MAX pooling. Third convolution block is used to remove insignificant features and uses  $1 \times 1$  convolution and  $2 \times 2$  MAX pooling. Fully connected layers are used to convert the obtained characteristics directly into the probability that the area belongs to the 'glands' class. The first fully connected layer consists of 16 neurons and the second one consists of 1 neuron.

All layers except the last one uses ReLU activation:

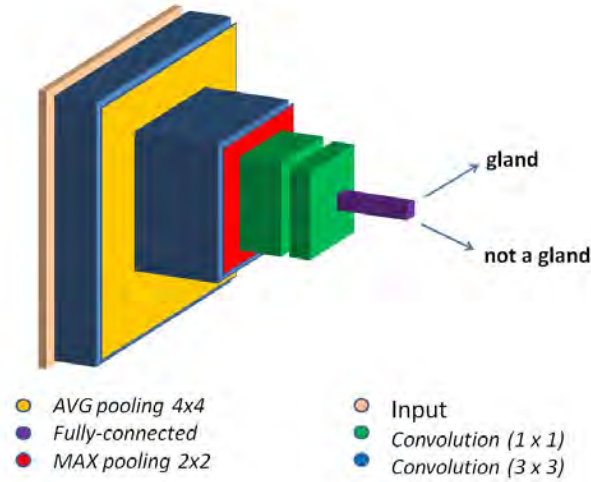
$$ReLU(x) = \begin{cases} x, & x \geq 0, \\ 0, & x < 0. \end{cases} \quad (4)$$

Last layer uses Sigmoid activation for binary classification:

$$sigmoid(x) = \frac{1}{1 + e^{-x}}. \quad (5)$$

Before each pooling layer a batch normalization [13] is used.

This architecture demonstrated the best results among the several other convolutional architectures with two or three convolutional blocks that were tested during the



**Fig. 5.** The architecture of the convolutional neural network that was used to filter glands candidates at the second stage of the proposed method.

development of the algorithm. During the experiments, various sizes and types of pooling layers were used as well as a different number of filters in convolutional layers. All convolutional networks were trained from scratch.

### 3 Experiments and results

#### 3.1 Data selection and preprocessing

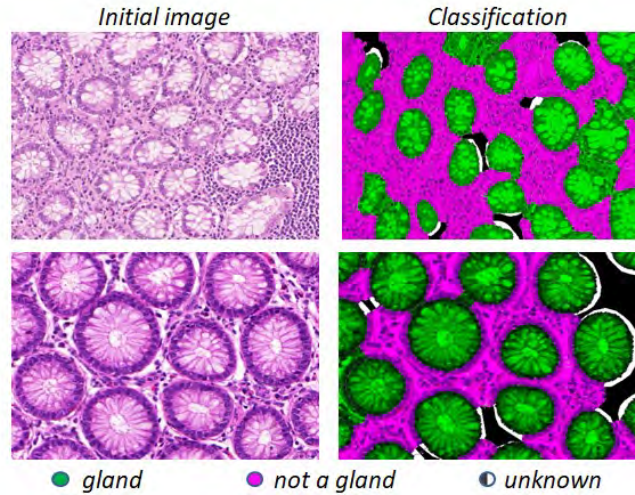
We used *Warwick-QU* [7, 11] dataset as material for the study. It contains 165 colon tissue images obtained with *Zeiss MIRAX MIDI* microscope with a resolution of  $0.62 \mu\text{m}/\text{pixel}$  and a ground truth gland annotation made by medical experts for each of the images.

*Warwick-QU* dataset was divided by ourselves into three subsets: test, validation and training:

- training subset consists of 133 images. Images from this group are used to train the CNN,
- validation subset consists of 12 images that are used to evaluate intermediate results during the training process and tune hyper parameters of the network,
- test subset consists of 10 images. They are used for the final assessment of the neural network prediction quality.

The first stage of the proposed algorithm, that was described in Section 2.1, was run for each image in the dataset, though providing the initial segmentation. After that each area on the image was compared to the ground truth annotation and was assigned to one of three classes (Fig. 6):





**Fig. 6.** Initial images from *Warwick-QU* dataset on the left and visualization of the received division into 'gland', 'not a gland' and 'unknown' classes of the initial segmentation areas on the right.

1. gland,
2. not a gland,
3. unknown.

The 'unknown' class was introduced to avoid feeding incorrect patches to the neural network during training process. Patch is incorrect if it's central part consists of pixels corresponding to a class that differs to class of the most pixels in the image. This has significantly improved the prediction. Only those areas, for which it was impossible to establish exact membership to a gland or non-gland class, were assigned to the 'unknown' class.

For each area of initial segmentation a bounding box is constructed. All pixels in this bounding box form a patch. The pixels that are not relevant to considered region in the patch are set to 0 in all RGB channels. All patches received are rectangles of different sizes.

### 3.2 Data augmentation

For augmentation, we used random rotations and reflections, as well as small shifts and zooming. Rotations were carried out from  $-10$  to  $10$  degrees and shifts were carried out no more than 15 percent of the patch size (which is  $256 \times 256$  pixels for *Warwick-QU* dataset).

The number of patches belonging to the 'gland' class is synthetically increased by 1.3 times by creating an extra patch for each third region. Each of these patches contains from 40% to 60% of the original patch area. This is necessary to artificially increase the number of patches which are examples of 'gland' class lying on the image border since

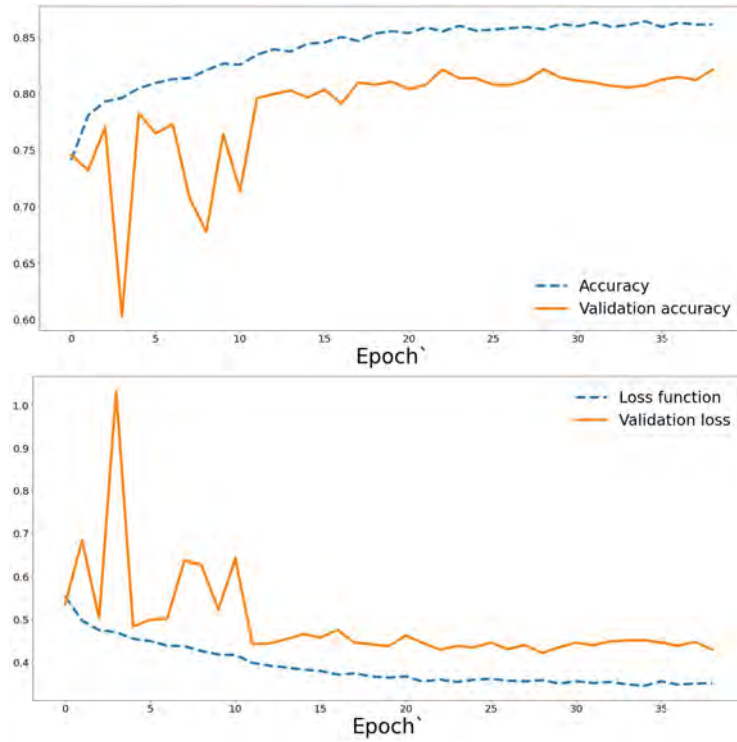


the number if these patches is much smaller than the number of patches that do not lie on the image border.

### 3.3 Training the convolutional neural network

Each patch is rescaled to the  $256 \times 256$  patch and gets a single-number label: 1 if the corresponding area belongs to the class of glands and 0 otherwise.

In total about 17,000 patches were received using the described above procedure. About 14,000 of these patches are from training subset. The number of patches belonging to the class of glands differs by no more than 5% from the number of patches belonging to the class of glands, which is preferable to obtain a robust CNN training.



**Fig. 7.** Values of accuracy and loss function over epoch on training and validation datasets.

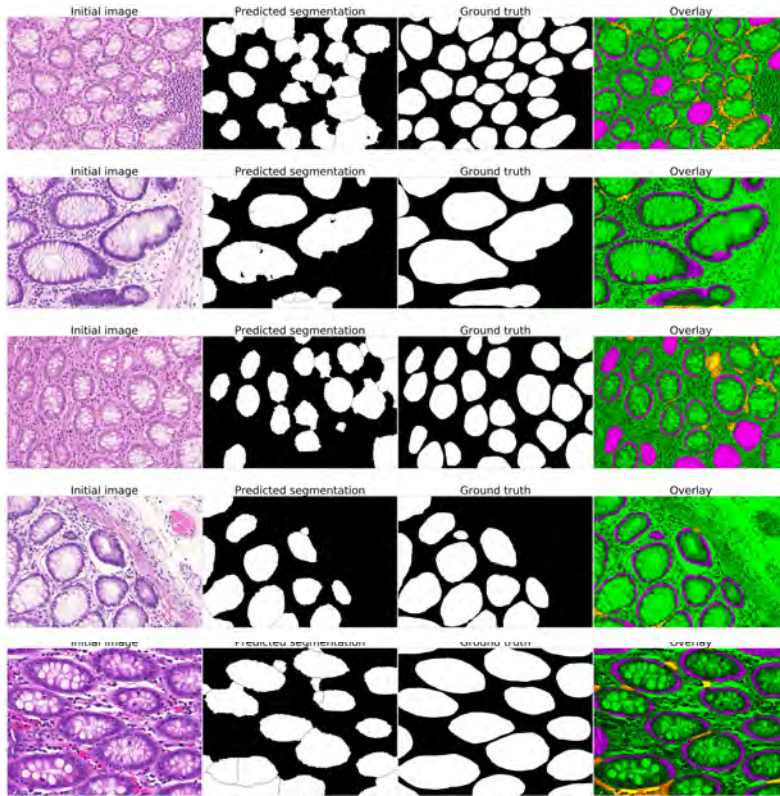
To train the described above CNN we used the Adam optimizer [14], with an initial learning rate equal to 0.001, batch size equal to 16 and binary cross entropy as a loss function:

$$Loss = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i)), \quad (6)$$

where  $y_i$  is the label (1 for 'gland' samples and 0 for 'not a gland' samples),  $p(y_i)$  is the predicted probability of the sample being 'gland' for all  $N$  samples.

At the end of each epoch, prediction accuracy was obtained on a validation set after which a decision was made about the network hyper parameters changing. If the averaged loss function value on the validation dataset did not decrease for 6 epochs in a row, the coefficient of learning rate was decreased by 5 times. If it did not decrease during 8 epochs, the training process was stopped.

### 3.4 Results



**Fig. 8.** The results obtained with the proposed algorithm on Wariwck-QU dataset. Left to right: source image, predicted segmentation, ground truth, overlay. In the overlay green zones are correctly segmented, violet ones are over-segmented, yellow ones are under-segmented.

We achieve accuracy of 0.85 on the test subset. Values of accuracy and loss function over epoch on training and validation datasets showed in Fig. 7.

The average *Dice score* (Equation 1) on the test set was 0.790.

The final results of the proposed algorithm can be observed in the Fig. 8.

## 4 Implementation details

The used segmentational CNN which is used as the send stage of the proposed algorithm was implemented using opensource neural network framework *Keras* with *TensorFlow 2.1.0* [15] backend. *Python 3.7* was used as a programming language. The experiments were performed on a personal computer *AMD Athlon II X4 760K* and *Nvidia GeForce GTX 650*. Neural network training took about 8 hours, processing of one image from *Warwick-QU* took about 120 seconds. The source code of the proposed method can be found at

[https://github.com/moscap/classic\\_gland\\_segmentation](https://github.com/moscap/classic_gland_segmentation).

## 5 Conclusion

In this work we propose a method of glands segmentation in histological images that constructs a probability map of glands contours based on a priory information about glandular structures and afterwards uses a CNN model to get a final segmentation. The average accuracy on the *Warwick-QU* is 0.825, *Dice score* value is 0.790. The proposed method also showed good separation properties in the case of processing the adjacent glands which is often a problem for methods fully based on convolutional networks.

The continuation of this research will include enhancing the prediction of glands contours maps and evaluating the method on *PATH-DT-MSU* [16, 17] dataset.

## References

1. Farjam, R., Soltanian-Zadeh, H., Jafari-Khouzani, K., Zoroofi, R.A.: An image analysis approach for automatic malignancy determination of prostate pathological images. *Cytometry Part B: Clinical Cytometry: The Journal of the International Society for Analytical Cytology* **72**(4), 227–240 (2007)
2. Jain, A.K.: Data clustering: 50 years beyond K-means. *Pattern recognition letters* **31**(8), 651–666 (2010)
3. Nguyen, K., Jain, A.K., Allen, R.L.: Automated gland segmentation and classification for gleason grading of prostate tissue images. In: 2010 20th International Conference on Pattern Recognition. pp. 1497–1500. IEEE (2010)
4. Fu, H., Qiu, G., Shu, J., Ilyas, M.: A novel polar space random field model for the detection of glandular structures. *IEEE transactions on medical imaging* **33**(3), 764–776 (2014)
5. Naik, S., Doyle, S., Feldman, M., Tomaszewski, J., Madabhushi, A.: Gland segmentation and computerized gleason grading of prostate histology by integrating low-, high-level and domain specific information. In: MIAAB workshop. pp. 1–8. Citeseer (2007)
6. Li, C., Xu, C., Gui, C., Fox, M.D.: Level set evolution without re-initialization: a new variational formulation. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). vol. 1, pp. 430–436. IEEE (2005)
7. Sirinukunwattana, K., Snead, D.R., Rajpoot, N.M.: A stochastic polygons model for glandular structures in colon histology images. *IEEE transactions on medical imaging* **34**(11), 2366–2378 (2015)

8. Chen, H., Qi, X., Yu, L., Heng, P.A.: Dcan: deep contour-aware networks for accurate gland segmentation. In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. pp. 2487–2496 (2016)
9. Xu, Y., Li, Y., Wang, Y., Liu, M., Fan, Y., Lai, M., Eric, I., Chang, C.: Gland instance segmentation using deep multichannel neural networks. *IEEE Transactions on Biomedical Engineering* **64**(12), 2901–2912 (2017)
10. Zhang, Y., Yang, L., Chen, J., Fredericksen, M., Hughes, D.P., Chen, D.Z.: Deep adversarial networks for biomedical image segmentation utilizing unannotated images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 408–416. Springer (2017)
11. Sirinukunwattana, K., Pluim, J.P., Chen, H., Qi, X., Heng, P.A., Guo, Y.B., Wang, L.Y., Matuszewski, B.J., Bruni, E., Sanchez, U., et al.: Gland segmentation in colon histology images: The glas challenge contest. *Medical image analysis* **35**, 489–502 (2017)
12. Saarninen, K.: Color image segmentation by a watershed algorithm and region adjacency graph processing. In: Proceedings of 1st International Conference on Image Processing. vol. 3, pp. 1021–1025. IEEE (1994)
13. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015)
14. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
15. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al.: Tensorflow: A system for large-scale machine learning. In: 12th USENIX symposium on operating systems design and implementation (OSDI '16). pp. 265–283 (2016)
16. Khvostikov, A., Krylov, A., Mikhailov, I., Kharlova, O., Oleynikova, N., Malkov, P.: Automatic mucous glands segmentation in histological images. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* (2019)
17. Oleynikova, N., Khvostikov, A., Krylov, A., Mikhailov, I., Kharlova, O., Danilova, N., Malkov, P., Ageykina, N., Fedorov, E.: Automatic glands segmentation in histological images obtained by endoscopic biopsy from various parts of the colon. *Endoscopy* **51**(04), OP9 (2019)