AUTOMATIC MUCOUS GLANDS SEGMENTATION IN HISTOLOGICAL IMAGES

A. Khvostikov[1]*, A. Krylov[1], I. Mikhailov[2], O. Kharlova[2], N. Oleynikova[2], P. Malkov[2],

[1] Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Moscow, Russia - khvostikov@cs.msu.ru, kryl@cs.msu.ru
[2] University Medical Center, Lomonosov Moscow State University, Moscow, Russia - imihailov@mc.msu.ru, olga.arsenteva@gmail.com, noleynikova@mc.msu.ru, pmalkov@mc.msu.ru

Commission II,WG II/10

KEY WORDS: Image segmentation, Mucous glands, Deep Learning, Convolutional Neural Networks, Histology, Pathology

ABSTRACT:

Mucous glands is an important diagnostic element in digestive pathology. The first step of differential diagnosis of colon polyps in order to assess their malignant potential is gland segmentation. The process of mucous glands segmentation is challenging as the glands not only needed to be separated from a background but also individually identified to obtain reliable morphometric criteria for quantitative diagnostic methods. We propose a new convolutional neural network for mucous gland segmentation that takes into account glands' contours and can be used for gland instance segmentation. Training and evaluation of the network was performed on a standard Warwick-QU dataset as well as on the collected PATH-DT-MSU dataset of histological images obtained from hematoxylin and eosin staining of paraffin sections of colon biopsy material collected by our Pathology department. The collected PATH-DT-MSU dataset will be available at `http://imaging.cs.msu.ru/en/research/histology`.

1. INTRODUCTION

A differential diagnosis criteria of colon polyps are not accurate, there is no quantitative criteria of basal dilation of the crypts and spread of the serration as well as no principles for determining the malignant potential of various benign colon epithelial neoplasms. The same task in vivo, directly during the endoscopic examination is no less difficult. Therefore, the development of auxiliary mathematical models for image recognition that can be used for online detection, endoscopic and morphological characterization of colon epithelial neoplasms is required.

Although a lot of classical semiautomatic (Fernandez-Gonzalez et al., 2004) and fully automatic (WU et al., 2005), (Gunduz-Demir et al., 2010), (Sirinukunwattana et al., 2015) methods were proposed to solve the problem of glands segmentation in histological images, all of them do not provide the required level of segmentation accuracy and tend to give unstable results in some medical cases.

Therefore, applying convolutional neural networks (CNNs) with their good generalization capacity for the problem of histological images segmentation looks more promising.

Almost all CNN-based segmentation methods (Long et al., 2015), (Badrinarayanan et al., 2017), (Ronneberger et al., 2015) use the same idea of convolutional autoencoder (CAE) (Masci et al., 2011). With minor changes these CNN-based segmentation methods can be also applied to histological images. The main problem of the mentioned above approaches of segmentation is the impossibility of the algorithms to separate close or contiguous objects. Various ideas were proposed to solve this problem. In (Chen et al., 2017) a DCAN architecture is proposed using the idea of object detection and separation, but unlike (Kainz et al., 2017) these two steps are performed simultaneously with one FCN-based network that has two outputs. First output predicts probabilities of gland object, while the second predicts the probability map of contours separating glands. The final segmentation masks are calculated using the threshold rule. To strengthen the training process DCAN uses 3 weighted auxiliary classifiers in the 3 deepest layers of the network. The idea of splitting segmented glands got a further development in (Xu and et al., 2017). The authors introduce a CNN with 3 pipelines: a FCN for the foreground segmentation, Faster R-CNN (Ren et al., 2015) for the object detection and HED (Xie and Tu, 2015) for edge detection. All three pipelines a fused into one, and are followed with several convolution layers to predict the final instance segmentation map. This approach leads to the state-of-the-art level of segmentation accuracy.

In this work we propose a CNN-based algorithm for histological images segmentation, that uses multiscale architecture, non-local block and contour-aware loss function. This work is the further improvements of our previous research (Khvostikov et al., 2018).

2. PROPOSED METHOD

We propose a new architecture of a convolutional neural network (CNN) for mucous glands segmentation (Fig. 1) based on U-Net model (Ronneberger et al., 2015) which has proven its good efficiency for segmentation of biomedical images.

The proposed architecture is designed with an attempt to separate glands that stuck together thus performing an instance segmentation. To do it we consider the information about the contours of the glands, but unlike the contour-aware network (Chen et al., 2016), the proposed network uses a combined loss function instead of using 2 outputs to predict glands and their contours. A contour probability map is calculated by applying Sobel filter to the output of the network and the loss function is calculated as the weighted sum of Dice losses of predicted gland map and obtained

*Corresponding author

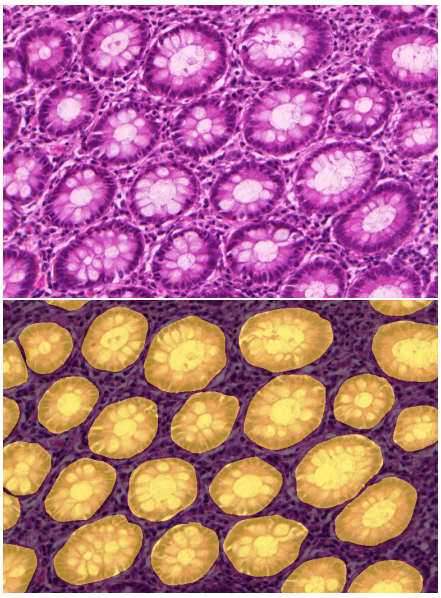


Figure 1. A histological image of colon mucous glands and glands annotation

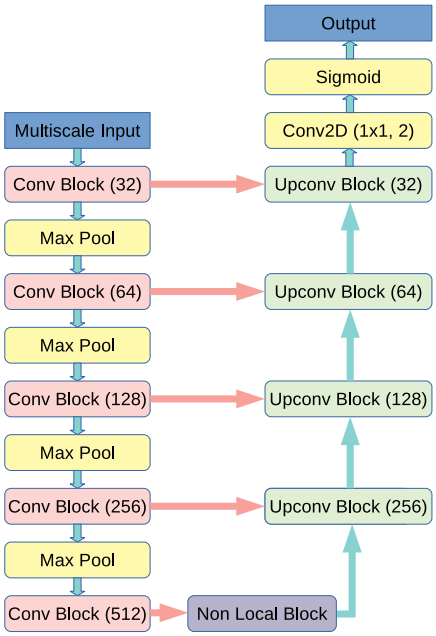


Figure 2. Proposed architecture for mucous glands segmentation

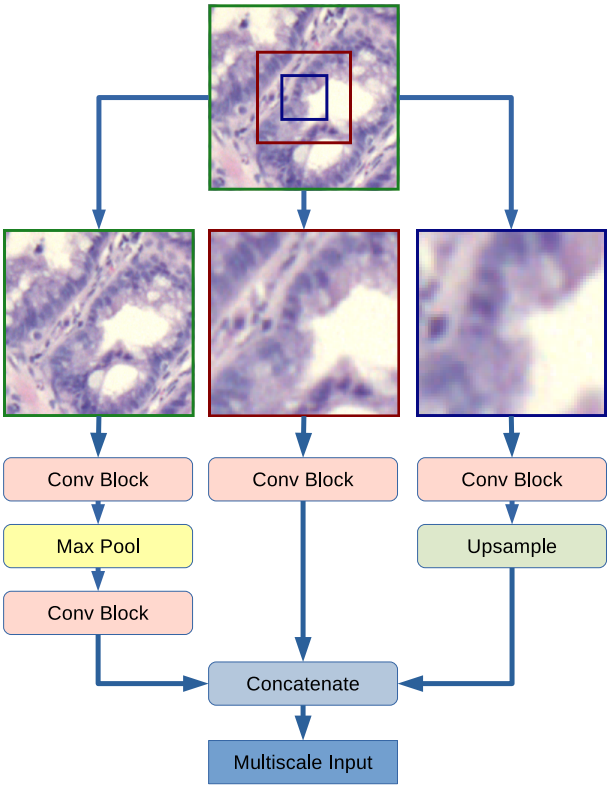


Figure 3. Multiscale input block of the proposed CNN

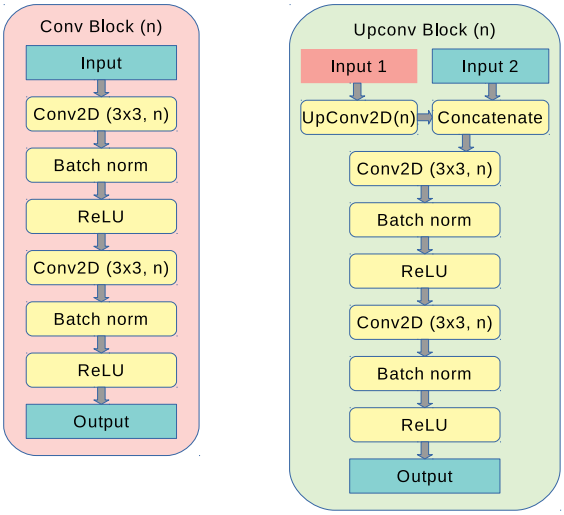


Figure 4. Conv and Upconv blocks of the proposed CNN

contour map.

$$L = \alpha L_c + (1 - \alpha) L_g,$$

where L_c is a Dice loss function applied to the contour map and L_g is a Dice loss function applied to the gland map. Herewith to obtain a more robust training process the weight α changes during training starting from 0 at the first epoch and smoothly increasing up to 0.5 during several epochs.

Although the original U-Net (Ronneberger et al., 2015) as like most CNN architectures used for semantic segmentation do not depend on the input size of the image and can work with image of arbitrary size, we use a fixed-size 256×2556 patch input. It is more convenient in terms of resources allocated for CNN (as full-size histological images can contain more than 4 millions pixels each) and also allows to use batch size values common to the deep learning architectures.

The other distinctive feature of the proposed CNN is the usage of multiscale architecture. Patches of scale 0.5x and 2x (128×128 and 512×512 respectively) are passed to the network's input alongside with the source patch (Fig.3). It allows to better seg-

ment tissue structures at different scales.

The last improvement is the use of a non-local block (Wang et al., 2018) at the deepest layer of the network. Its working principle is similar to classical non-local methods and allows it to capture long-range dependencies in the image structures inside the patch. It also has a positive effect on the time needed for model convergence. The only one non-local block is used due to the limited memory resources of the GPU.

The architecture of the proposed CNN is demonstrated on Fig. 2, 3, 4.

The proposed CNN for segmentation is patch-oriented, which means that in order to get the output segmentation for a test image it is split into patches, every patch is passed through the network and then the segmented patches are merged together to get the final result. During the merging we take into account that only the central parts of the segmented patches contain relevant information due to the convolutional padding at multiple input scales. We also perform splitting of patches with the $1/4$ of the patch size overlay and merging with averaging of the output segmentation correspondingly. It makes the predictions more smooth and accurate.

3. EXPERIMENTS AND RESULTS

In this section we describe the data used for training and evaluation of the proposed segmentation algorithm, describe the performed experiments and give the evaluation results along with description of the used evaluation metrics.

3.1 Data selection and preprocessing

In this work we use two different datasets of histological images.

The first one is Warwick-QU dataset (Warwick-QU image dataset description, 2015), which was used for Gland Segmentation Challenge Contest in MICCAI 2015 (Sirinukunwattana and et al., 2017). It contains images acquired by a Zeiss MIRAX MIDI slide scanner from colorectal cancer tissues with a resolution of $0.62\mu m$/pixel. It's worth noting that Warwick-QU contains images of a wide range of histologic grades from benign to malignant subjects, but in this current work we use only a benign subset of Warwick-QU dataset that contains 37 train images and 37 images used for evaluation.

The second dataset is PATH-DT-MSU dataset that was collected and annotated by our Department of Pathology and consists of 20 histological images obtained from hematoxylin and eosin staining of paraffin sections of colon biopsy material. 13 images are hyperplastic polyps (HP); 6 images are sessile serrated adenomas (SSA/P) and one image is normal colon mucous glands. This sample structure is necessary for the further search for quantitative criteria for differential diagnosis between HP and SSA/P. It should also be noticed that the PATH-DT-MSU dataset contains full-size images in contrast to Warwick-QU that contains only central parts of the full-size histological images. It makes the process of automatic segmentation more difficult but also provides us a field for more complicated analysis of histological structures.

3.2 Data augmentation

In order to enlarge the amount of data used for training we augment the obtained histological images. The augmentation is performed on the fly. During training process every patch is randomly cropped from the randomly chosen training image. Then random shift, rotation, scale, flip and non-linear operations as well as random change of brightness are applied to the patch. The initial size of the patch is chosen corresponding to all described transformations so that after applying all of them it can be centrally cropped to the size of the proposed network's input. As the input is multiscale the augmentation of each patch is performed at three scales simultaneously. The described augmentation process is controlled by parameter τ describing the number of times the obtained training data is enlarged. In this work we use $\tau = 10$ for both Warwick-QU and PATH-DT-MSU datasets.

3.3 First phase of training

At first we train the proposed network on the train subset of Warwick-QU. Herewith the batch size is chosen as 8, the training is performed using RMSProp optimizer (Tieleman and Hinton, 2012) with initial learning rate $2\mathrm{e}{-3}$ with automatic 10 times decrease when validation loss falls on the plateau and stops decreasing. When the target metric does not change within $1\mathrm{e}{-4}$ range for 10 epochs, the training process is stopped automatically.

3.4 Second phase of training

After the network is trained on Warwick-QU dataset we do a fine-tuning on PATH-DT-MSU dataset to fit the segmentation algorithm to this type of images.

As PATH-DT-MSU dataset contains full-size histological images a problem of processing "open" glands appears. They represent typical mucous glands built from goblet cells and enterocytes with open contour on histological images which are essentially the upper portions of the crypts. Visualized on a histological images they look like glands with the internal lumen merged with the background (Fig. 5). These "open" glands were also annotated in PATH-DT-MSU dataset. The difference in evaluation of the proposed segmentation algorithm trained with and without consideration of the "open" glands is shown in Fig. 7 and Fig. 8. In the future work we plan to treat normal glands and "open" glands as a separate classes.

PATH-DT-MSU dataset differs from Warwick-QU dataset not only by the tissue capture (Warwick-QU contains only central parts of tissue slides, while PATH-DT-MSU consists of full-size images) but also by the image resolution. In order to use the obtained PATH-DT-MSU dataset for transfer learning purposes, we downscaled the images from PATH-DT-MSU dataset by 30% so that the size of histological structures in both datasets are of the same size.

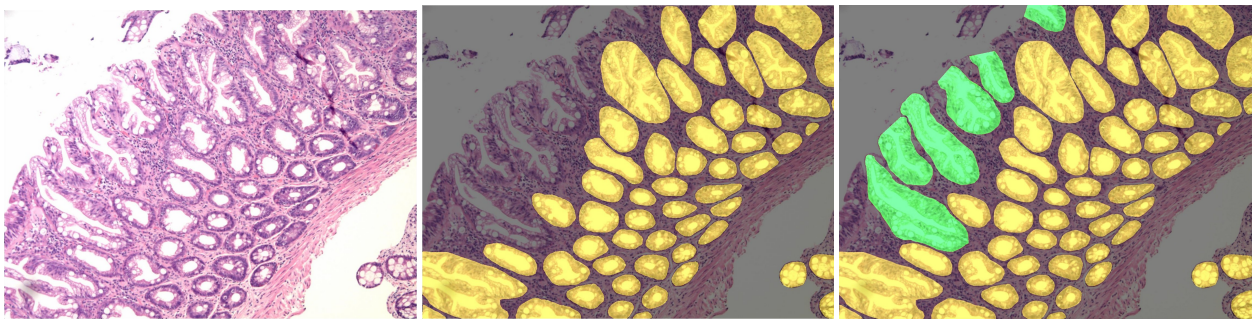
For this fine-tuning we choose the same optimizer as for the first phase but with smaller initial learning rate value of $2\mathrm{e}{-4}$. The rules for automatic learning rate decrease and training termination remain the same.

3.5 Obtained results

The common way of segmentation algorithm evaluation is calculating of Dice score. Given a set of pixels G annotated as a ground truth gland and a set of pixels S predicted as a gland, Dice score can be calculated as

$$D(G,S) = 2\frac{|G \cap G|}{|G| + |S|}.$$

However, this is not suitable for segmentation evaluation on individual objects (instance segmentation). For these reason an



(a) Source image (b) Glands annotations (c) Glands annotations with ”open” glands

Figure 5. Sample image from PATH-DT-MSU dataset.

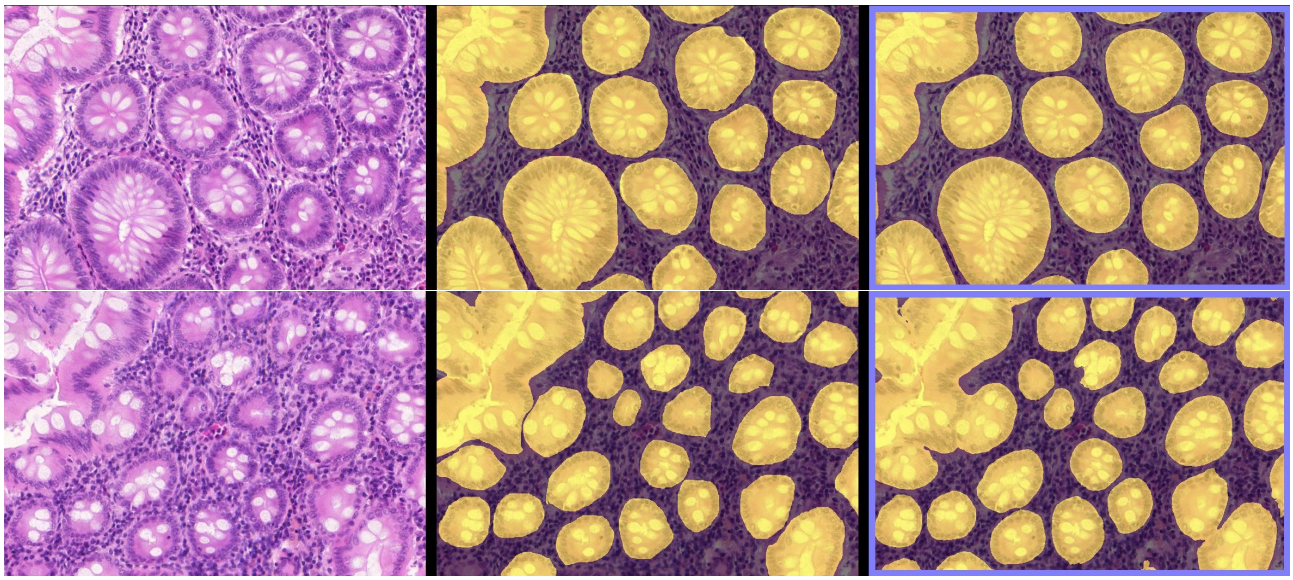
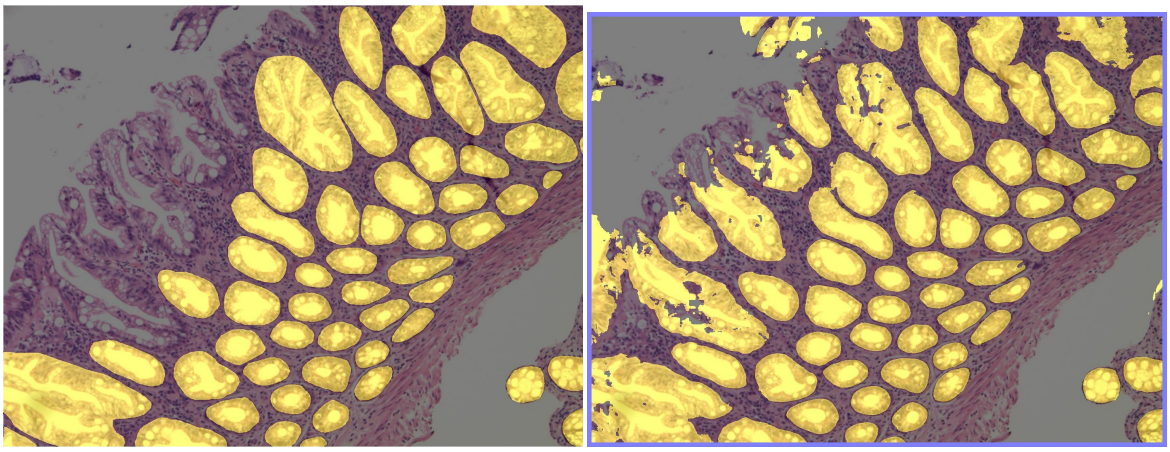


Figure 6. Results of the first phase of training on Warwick-QU dataset. Left to right: source image, ground truth glands, predicted glands.

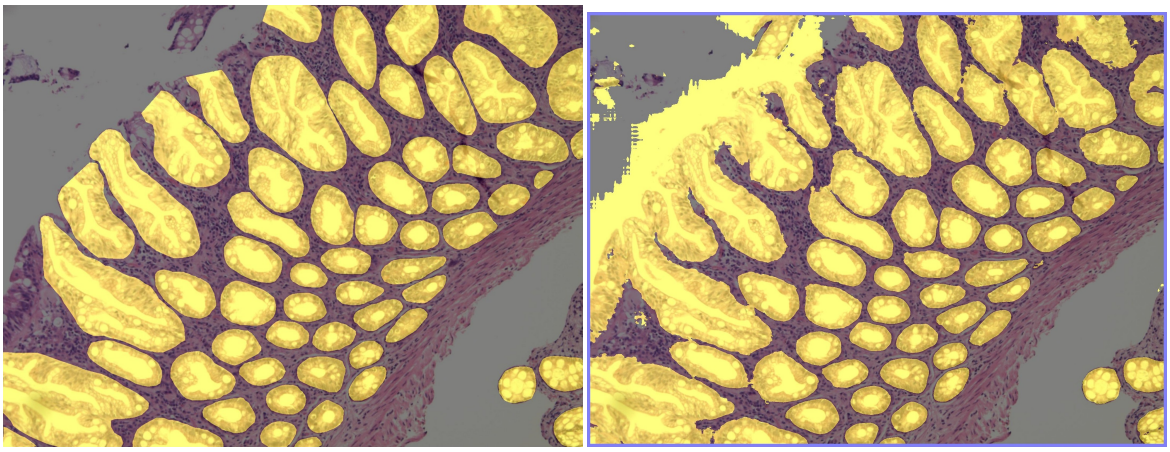


(a) (b)

Figure 7. Results of the proposed CNN for PATH-DT-MSU dataset if trained with normal gland annotations only; (a) is ground truth segmentation, (b) is the predicted result.

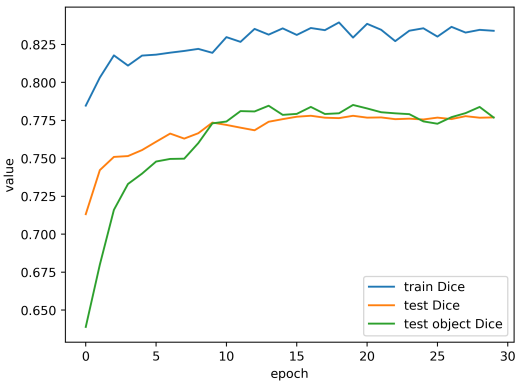
object-level Dice score (or object Dice) is utilized (Sirinukunwattana and et al., 2017) and defined as

$$D_{object}(G,S) = \frac{1}{2}[\sum_{i=1}^{n_S} \omega_i D(G_i, S_i) + \sum_{j=1}^{n_G} \widetilde{\omega_j} D(\widetilde{G_j}, \widetilde{S_j})],$$

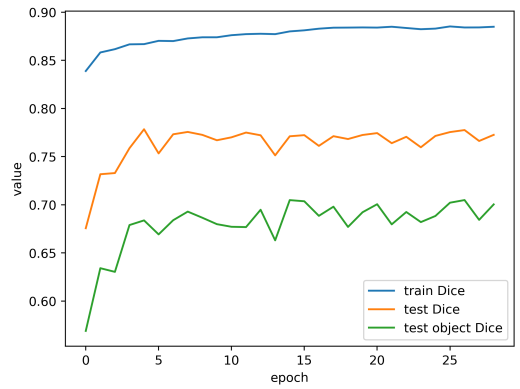


(a) (b)

Figure 8. Results of the proposed CNN for PATH-DT-MSU dataset if trained with both normal and ”open” gland annotations; (a) is ground truth segmentation, (b) is the predicted result.



(a)



(b)

Figure 9. Dice and object Dice scores while training the proposed CNN on PATH-DT-MSU dataset (a) with normal gland annotations only and (b) with both normal and ”open” glands.

where where S_i denotes the ith segmented object, G_i denotes a ground truth object that maximally overlaps S_i, $\widetilde{G_j}$ denotes the jth ground truth object, $\widetilde{S_j}$ denotes a segmented object that maximally overlaps $\widetilde{G_j}$, $\omega_i = |S_i| / \sum_{m=1}^{n_S} |S_m|$, $\widetilde{\omega_j} = |\widetilde{G_j}| / \sum_{n=1}^{n_G} |\widetilde{G_n}|$, n_S and n_G are the total number of segmented objects and ground truth objects, respectively.

For convenience in the current work we give the evaluation results both in Dice score and object Dice score.

For the first phase of training of the proposed CNN on the benign test subset of Warwick-QU dataset we obtained the 0.92 Dice score and 0.88 object Dice score. As it can be seen from Fig.6 although the gland instance segmentation is performed with a relatively good quality it is still not ideal and the main problem of predicted segmentation is that close lying glands can sometimes be merged together.

For the second phase of training the proposed CNN we made two experiments by fine-tuning the CNN on PATH-DT-MSU dataset annotated with and without ”open” glands. We achieved the values of 0.78 Dice and 0.77 object Dice scores for configuration without ”open” glands and 0.77 Dice and 0.7 object Dice scores for the configuration with ”open” glands (Fig. 9). From Figs. 7, 8 it can be seen that the main deviations of the ground truth and predicted annotations are located at near-boundary glands. As it was previously discussed, the standard Warwick-QU dataset represents only the central parts of the histological tissue slides but in case of PATH-DT-MSU dataset it contains full-size images which makes the problem of segmentation of the glands that are adjacent to lumen of the colon much more challenging. In particular, the network trained on the annotations excluding ”open” glands seeks to segment these kind of glands which leads to oversegmentation if compared to the ground truth annotation (Fig.8). At the other side the network trained on the annotations with ”open” glands seeks to segment not only the ”open” glands themselves but also some extra space outside them which leads to merging several ”open” glands into one. Both these cases demonstrate worse results compared to Warwick-QU dataset, but from the medical point of view the segmentation results of second configuration are more preferable. The most obvious way to improve the segmentation of ”open” glands is to perform an analysis of the image at global scale with detection of lumen of colon and muscularis mucosae. So, this is one of the tasks planned for us for the continuation of current research.

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLII-2-W12-103-2019 | © Authors 2019. CC BY 4.0 License.

4. IMPLEMENTATION DETAILS

The proposed segmentation CNN was implemented using open source neural network library Keras (Chollet et al., 2015) with TensorFlow (Abadi et al., 2016) backend. The experiments were performed on two configurations: a personal computer with Intel(R) Core(R) i7-6700HQ CPU and Nvidia GeForce GTX 960M GPU and a FloydHub cloud server with Nvidia Tesla K80 GPU (Soundararaj et al., 2016).

5. CONCLUSIONS

In this work we propose a new convolutional neural network for mucous glands segmentation in histological images. The proposed model is trained and evaluated using Warwick-QU dataset and PATH-DT-MSU dataset obtained by our Pathology department. The multiscale architecture of the proposed CNN makes it less sensitive to the scale of the input image. Due to the specific loss function it is able to detect and separate stuck glands. The used non-linear block enhances the segmentation and has a positive effect on the time needed for model to converge. Altogether this leads to the accurate segmentation of glands on histology images (0.92 Dice and 0.87 object Dice scores for Warwick-QU dataset, 0.78 Dice and 0.77 object Dice scores for the PATH-DT-MSU dataset).

The generalization ability of the proposed algorithm enables it to effectively segment individual glands in histological images. The collected PATH-DT-MSU dataset of histological images of colon biopsy material allows to fine-tune the proposed CNN trained on Warwick-QU dataset and exposes several directions for further development of automatic tools for histological image analysis.

There are several objectives we are going to focus on in the future research. One of them is enlarging the PATH-DT-MSU dataset with new histological images and developing an algorithm for large scale image analysis in order to correctly detect lumen of colon and muscularis mucosae and as a result improve "open" glands segmentation. Another objective is to implement more accurate algorithm to segment individual gland objects, which can be used as post-process segmentation tool. And finally the most ambitious objective is to perform a more complex inner-gland segmentation (detect nuclei, lumen and cytoplasm). The results of this segmentation can be used for the ensuing analysis. In particular, analyzing the histological images of mucous glands helps to detect changes in its lumen shape (serration), in the nuclear-cytoplasmic ratio inside mucus-forming cells, and in the character of the expression of immunohistochemical markers (Oleynikova et al., 2017).

ACKNOWLEDGEMENTS (OPTIONAL)

The algorithm development part of this work was supported by Russian Science Foundation grant 17-11-01279.

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M. et al., 2016. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*.

Badrinarayanan, V., Kendall, A. and Cipolla, R., 2017. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence* 39(12), pp. 2481–2495.

Chen, H., Qi, X., Yu, L. and Heng, P.-A., 2016. Dcan: deep contour-aware networks for accurate gland segmentation. In: *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 2487–2496.

Chen, H., Qi, X., Yu, L., Dou, Q., Qin, J. and Heng, P.-A., 2017. Dcan: Deep contour-aware networks for object instance segmentation from histology images. *Medical image analysis* 36, pp. 135–146.

Chollet, F. et al., 2015. Keras. `https://keras.io`.

Fernandez-Gonzalez, R., Deschamps, T., Idica, A., Malladi, R. and de Solorzano, C. O., 2004. Automatic segmentation of histological structures in mammary gland tissue sections. *Journal of biomedical optics* 9(3), pp. 444–454.

Gunduz-Demir, C., Kandemir, M., Tosun, A. B. and Sokmensuer, C., 2010. Automatic segmentation of colon glands using object-graphs. *Medical image analysis* 14(1), pp. 1–12.

Kainz, P., Pfeiffer, M. and Urschler, M., 2017. Segmentation and classification of colon glands with deep convolutional neural networks and total variation regularization. *PeerJ* 5, pp. e3874.

Khvostikov, A., Krylov, A. S., Kharlova, O., Oleynikova, N., Mikhailov, I. and Malkov, P., 2018. CNN-based histological images segmentation of mucous glands. In: *Proceedings of 28th international conference on computer graphics and computer vision. GraphiCon 2018, Tomsk*, pp. 258–261.

Long, J., Shelhamer, E. and Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3431–3440.

Masci, J., Meier, U., Cireşan, D. and Schmidhuber, J., 2011. Stacked convolutional auto-encoders for hierarchical feature extraction. In: *International Conference on Artificial Neural Networks*, Springer, pp. 52–59.

Oleynikova, N., Kharlova, O., Mal'kov, P. and Danilova, N., 2017. Morphological and immunohistochemical classification aspects of serrated formations of the large intestine. *International Journal of Advanced Research* (5), pp. 118–135.

Ren, S., He, K., Girshick, R. and Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. In: *Advances in neural information processing systems*, pp. 91–99.

Ronneberger, O., Fischer, P. and Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical image computing and computer-assisted intervention*, Springer, pp. 234–241.

Sirinukunwattana, K. and et al., 2017. Gland segmentation in colon histology images: The glas challenge contest. *Medical image analysis* 35, pp. 489–502.

Sirinukunwattana, K., Snead, D. R. and Rajpoot, N. M., 2015. A stochastic polygons model for glandular structures in colon histology images. *IEEE transactions on medical imaging* 34(11), pp. 2366–2378.

Soundararaj, S. et al., 2016. FloydHub. fastest way to build, train, and deploy deep learning models. `https://www.floydhub.com`. Accessed: 2019-03-31.

Tieleman, T. and Hinton, G., 2012. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. *COURSERA: Neural networks for machine learning* 4(2), pp. 26–31.

Wang, X., Girshick, R., Gupta, A. and He, K., 2018. Non-local neural networks. In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. 1number 3, p. 4.

Warwick-QU image dataset description, 2015. `https://warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/about/`.

WU, H.-S., Xu, R., Harpaz, N., Burstein, D. and Gil, J., 2005. Segmentation of intestinal gland images with iterative region growing. *Journal of Microscopy* 220(3), pp. 190–204.

Xie, S. and Tu, Z., 2015. Holistically-nested edge detection. In: *Proceedings of the IEEE international conference on computer vision*, pp. 1395–1403.

Xu, Y. and et al., 2017. Gland instance segmentation using deep multichannel neural networks. *IEEE Transactions on Biomedical Engineering* 64(12), pp. 2901–2912.

Revised April 2019