

Semantic Segmentation of Colon Glands with Deep Convolutional Neural Networks and Total Variation Segmentation

Philipp Kainz^{1,2,*}, Michael Pfeiffer², and Martin Urschler^{3,4,5}

¹Institute of Biophysics, Center for Physiological Medicine, Medical University of Graz, Graz, Austria

²Institute of Neuroinformatics, University of Zurich and ETH Zurich, Zurich, Switzerland

³Institute for Computer Graphics and Vision, Graz University of Technology, Graz, Austria

⁴Ludwig Boltzmann Institute for Clinical Forensic Imaging, Graz, Austria

⁵BioTechMed-Graz, Graz, Austria

*philipp.kainz@medunigraz.at

Abstract

Segmentation of histopathology sections is an ubiquitous requirement in digital pathology and due to the large variability of biological tissue, machine learning techniques have shown superior performance over standard image processing methods. As part of the *GlaS@MICCAI2015* colon gland segmentation challenge, we present a learning-based algorithm to segment glands in tissue of benign and malignant colorectal cancer. Images are preprocessed according to the Hematoxylin-Eosin staining protocol and two deep convolutional neural networks (CNN) are trained as pixel classifiers. The CNN predictions are then regularized using a figure-ground segmentation based on weighted total variation to produce the final segmentation result. On two test sets, our approach achieves a tissue classification accuracy of 98% and 94%, making use of the inherent capability of our system to distinguish between benign and malignant tissue.

1 Introduction

The variability of glandular structures in biological tissue poses a challenge to automated analysis of histopathology slides. It has become a key requirement to quantitative morphology assessment and supporting cancer grading. Considering non-pathological cases only, automated segmentation algorithms must already be able to deal with significant variability in shape, size, location, texture and staining of glands. Moreover, in pathological cases gland objects can tremendously differ from non-pathological and benign glands, which further exacerbates finding a general solution to the segmentation problem.

Previous work on gland segmentation in colon tissue has used graphical models [1, 2, 3] or textural features [4]. Others worked on segmentation in prostatic cancer tissue using an integrated low-, high-level and contextual segmentation model [5], probabilistic Markov models [6], k-means clustering and region growing [7], spatial association of nuclei to gland lumen [8, 9]. The reader is referred to the work of Sirinukunwattana *et al.* [3] for a more detailed description of work related to glandular

structure segmentation. Deep learning methods, especially convolutional neural networks (CNNs) [10], have found applications in biomedical image analysis for different tasks: semantic segmentation [11], mitosis detection [12] and classification [13], and blood cell counting [14].

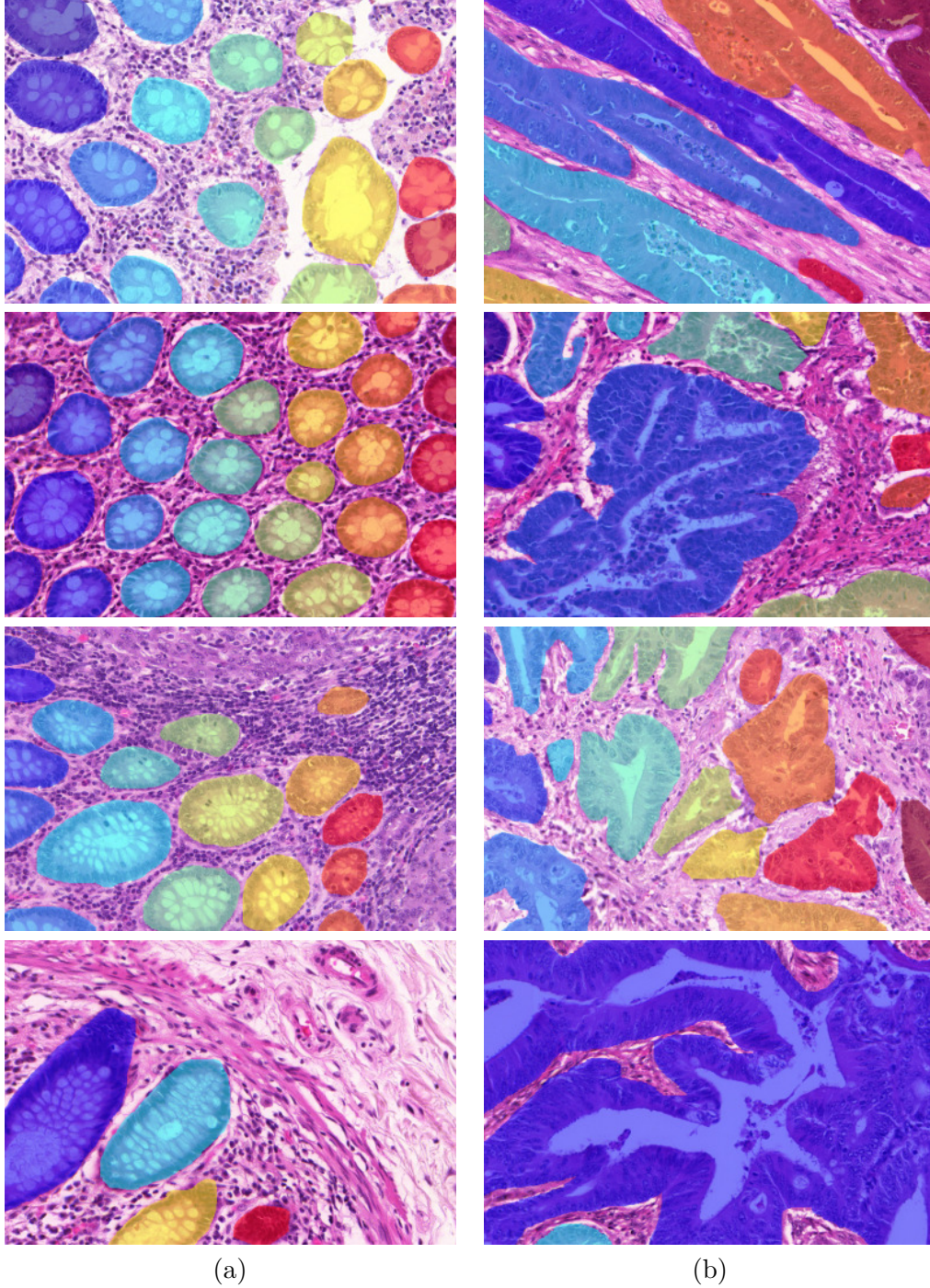


Fig. 1: Samples of (a) benign and (b) malignant colorectal cancer sections in the *Warwick-QU* dataset. Ground truth labels in each image are available for each pixel and overlaid in different colors for individual objects.

In this work, we propose a learning-based strategy to semantically segment glands in the *Warwick-QU* dataset, presented at the *GlaS@MICCAI2015* challenge¹. It contains 161 annotated images of benign and malignant colorectal adenocarcinoma, stained with Hematoxylin-Eosin (H&E) and scanned at 20× magnification. Fig. 1 shows some example images and their ground truth annotation. In each image, individual objects are annotated with the same label, illustrated by the different colors. To the challenge participants, information on whether an image shows benign or malignant tissue is only available in the training dataset. Three datasets were released during the contest and the total number of non-overlapping images (benign/malignant) in the training set, test set A and test set B is 85(37/48), 60(33/27), and 16(12/4), respectively. These datasets further contained 795, 666, and 91 individual glands.

The contributions of our work are twofold: (i) we present a novel deep learning scheme to generate classifier predictions for malignant and benign object and background pixels accompanied by a dedicated gland-separating refinement classifier that is able to distinguish touching objects, which pose a challenge for later segmentation. (ii) We use these classification results as the input for a simple, yet effective, globally optimal figure-ground segmentation approach based on a convex geodesic active contour formulation that regularizes the classifier predictions according to a minimal contour-length principle. Both technological contributions are described in section 2, while the subsequent sections show and discuss the results of our novel approach applied to the datasets of the *GlaS@MICCAI2015* challenge.

2 Methods

We present a segmentation method for Hematoxylin-Eosin (H&E) stained histopathological sections that proceeds in three steps: The raw RGB images are preprocessed to extract a robust representation of the tissue structure. Subsequently, two classifiers are trained to predict glands (*Object-Net*) and gland-separating structures (*Separator-Net*) from the image. Finally, the outputs of the classifiers are combined and a figure-ground segmentation based on weighted total variation is used to produce the segmentation result.

2.1 Preprocessing H&E Slides

Prior to classification, the RGB images are preprocessed as shown in Fig. 2. A standard color deconvolution [15] is performed for the specific H&E staining used in the provided dataset². It separates tissue components according to their staining, emphasizes the structure and inherently performs data whitening. The first (red) channel of the deconvolved RGB image contains most of the tissue structure information, so the other channels can be omitted. In order to account for different staining contrasts and lighting conditions during image acquisition, contrast limited adaptive histogram equalization (CLAHE) [17] is applied.

¹<http://www2.warwick.ac.uk/fac/sci/dcs/research/combi/research/bic/glascontest/>

²We used the *H&E 2* setting in the implementation of G. Landini, available in Fiji [16].

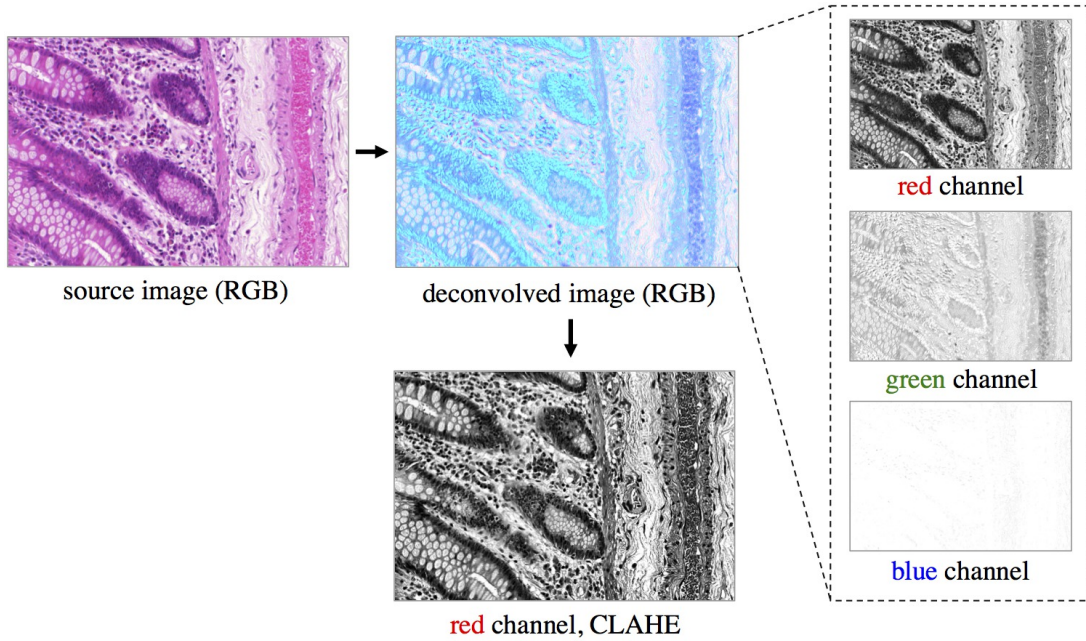


Fig. 2: Preprocessing of the RGB images. Color deconvolution [15] separates the Hematoxylin-Eosin stained tissue components. The red channel of the deconvolved image is processed by CLAHE [17] and taken as input to the pixel classifiers.

2.2 Learning Pixel Classifiers

Given the large variability of both benign and malignant tissue in the *Warwick-QU* dataset, we opted for CNNs due to their recently shown convincing performance in pixelwise classification of histopathology images [12] and to learn a rich set of features directly from images.

The general architecture of both CNNs is motivated by a classical LeNet-5 architecture [18] and consists of $K = 7$ ($k = 1, \dots, K$) layers: four convolutional layers ($\text{Conv}k$) for feature learning and three fully connected ($\text{FC}k$) layers as feature classifier, see Fig. 3. The rectified linear unit (ReLU) nonlinearity ($f(x) = \max(0, x)$) is used as the activation function throughout all layers in the networks. All convolutional layers consist of a set of learnable square 2D filters with pixel stride 1, followed by ReLU activation. Subsampling (max-pooling) layers ($\text{Sub}k, 2 \times 2$), accounting for translation invariance, are used after the first three convolutional layers and are counted as part of the convolutional layer. The final pixelwise classification of an input image is obtained by sliding a window over the image, and classifying the center pixel of that window.

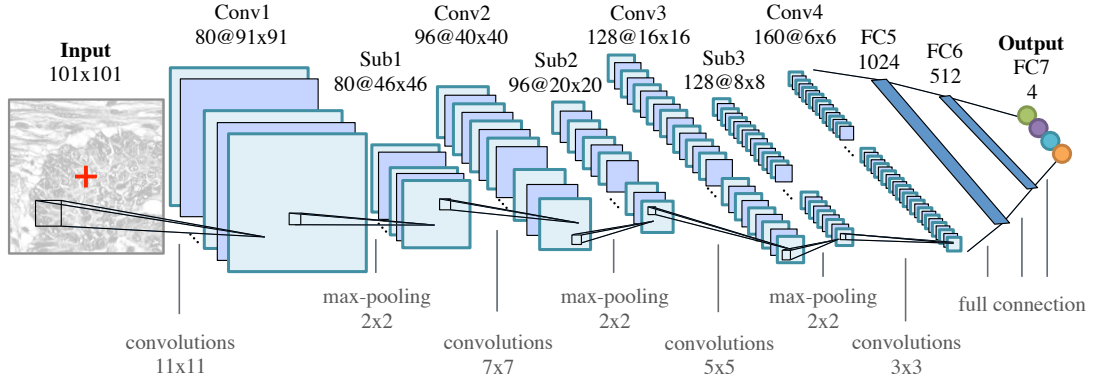
For training minibatch stochastic gradient descent (MBSGD) with momentum, weight decay, and dropout regularization is used to minimize a negative log-likelihood loss function.

2.2.1 *Object-Net*: Classifying Gland Objects

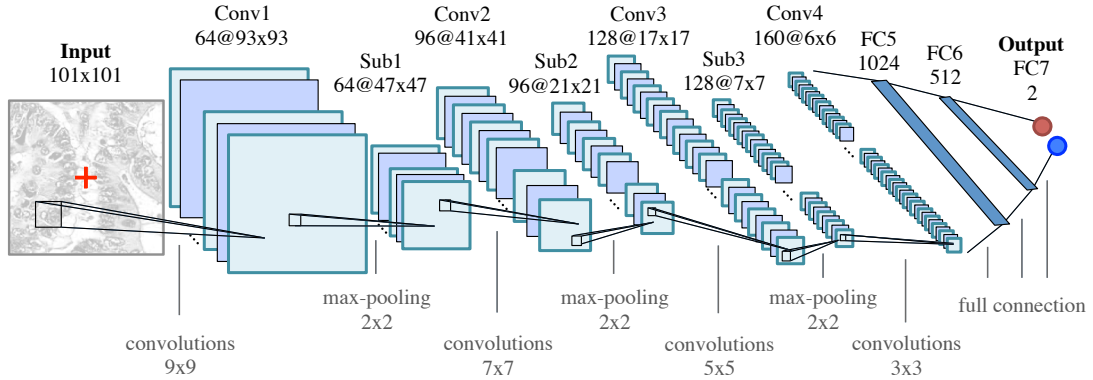
The goal of the *Object-Net* is to predict the probability of a pixel belonging to a gland or background. One could now define a binary classification problem, but malignant and benign tissue express unique features, which are not found in the other tissue type, and which can thus complicate the learning problem. We therefore formulate

an alternative, four-class classification problem, in which we distinguish ($L = 4$, with $l = 0, \dots, L - 1$): background benign (C_0), gland benign (C_1), background malignant (C_2), and gland malignant (C_3). In order to do that it is necessary to transform the provided ground truth labels to reflect benignity and malignancy as well. The annotation images are binarized and a new label is assigned to pixels belonging to each class C_l , see Fig. 4.

The input to the CNN is an image patch $I(\mathbf{x})$ of size 101×101 pixels, centered at an image location $\mathbf{x} = (u, v)^\top$, where $\mathbf{x} \in \Omega$ and Ω denotes the image domain. A given patch $I(\mathbf{x})$ is convolved with 80 filters (11×11) in the first convolutional layer, in the second layer with 96 filters (7×7), in the third layer with 128 filters (5×5), and in the last layer with 160 filters (3×3), see Fig. 3(a). The three subsequent fully connected layers FC5-FC7 of the classifier contain 1024, 512, and four output units, respectively. The output of FC7 is fed into a softmax function, producing the center pixel's probability distribution over the labels. The probability for each class l is stored in a corresponding map $I_{C_l}(\mathbf{x})$.



(a) *Object-Net* architecture



(b) *Separator-Net* architecture

Fig. 3: CNN classifier architectures of (a) the *Object-Net* and (b) the *Separator-Net*. Both architectures have $K = 7$ ($k = 1, \dots, K$) layers, are identical in the number of convolutional ($\text{Conv}k$), max-pooling ($\text{Sub}k$), and fully connected ($\text{FC}k$) layers, but differ in convolution kernel size, size and number of the feature maps, as well as number of output units. The probability distribution over L labels of the center pixel $\mathbf{x} = (u, v)^\top$ (marked as red cross in the input patch) is predicted by the CNNs.

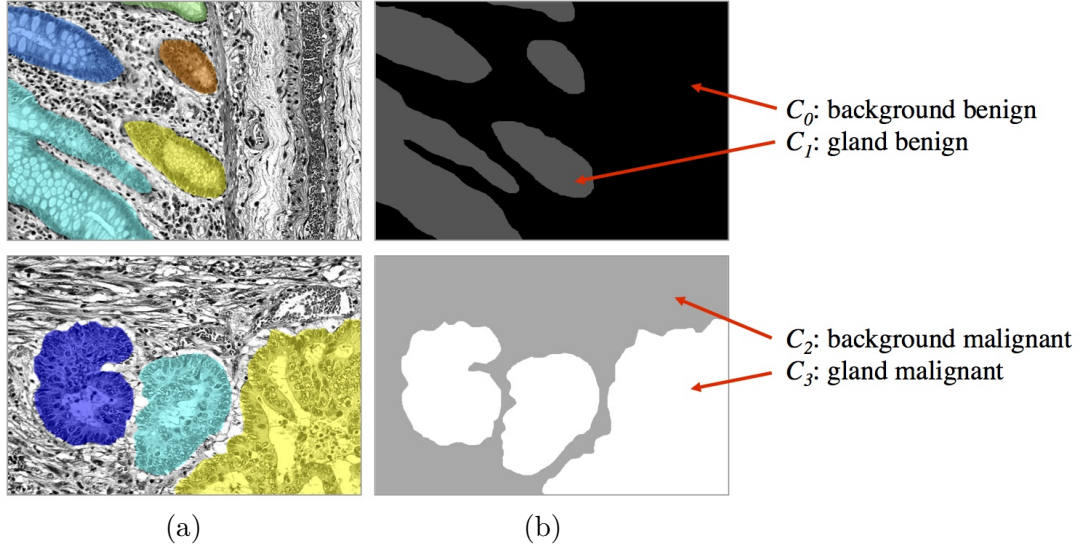


Fig. 4: Ground truth transformation for learning the four-class classification on the preprocessed images with the *Object-Net*. The first row shows a benign case, the second row shows a malignant case. (a) Preprocessed images with overlaid individual ground truth object annotations. (b) Provided annotations were transformed into four labels for benign background (C_0), benign gland (C_1), malignant background (C_2) and malignant gland (C_3).

2.2.2 *Separator-Net*: Classifying Gland-separating Structures

Initial experiments have shown that taking pixelwise predictions only from the *Object-Net* were insufficient in order to separate very close gland objects. Hence, a second CNN, the *Separator-Net*, is trained to predict structures in the image that are separating such objects. This learning problem is formulated as binary classification task.

As depicted in Fig. 3(b), the CNN structure is similar to the *Object-Net*: a given input image patch $I(\mathbf{x})$ of size 101×101 pixels is convolved with 64 filters (9×9) in the first convolutional layer, in the second layer with 96 filters (7×7), in the third layer with 128 filters (5×5), and in the last layer with 160 filters (3×3). The three subsequent fully connected layers FC5-FC7 of the classifier contain 1024, 512, and two output units, respectively. The output of the last layer (FC7) is fed into a softmax function to produce the probability distribution over the labels for the center pixel. The probability for a pixel \mathbf{x} belonging to a gland-separating structure is stored in the corresponding probability map $S(\mathbf{x})$.

2.2.3 Refining CNN Outputs

Once all probability maps have been obtained, the *Object-Net* predictions $I_{C_l}(\mathbf{x})$ are refined with the *Separator-Net* predictions $S(\mathbf{x})$ to emphasize the gland borders and prevent merging of close objects. The subsequent figure-ground segmentation algorithm requires a single foreground and background map to produce the final segmentation result, so outputs are combined as follows.

The foreground probability map p_{fg} is constructed by

$$p_{fg}(\mathbf{x}) = \max \left\{ \left(\sum_{l \in \{1,3\}} I_{C_l}(\mathbf{x}) \right) - \rho S(\mathbf{x}), 0 \right\}, \quad (1)$$

where $\rho \in [0, 1]$ controls the influence of the refinements done by the separator predictions. Similarly, evaluating Eq. (2) produces the background probability map:

$$p_{bg}(\mathbf{x}) = \min \left\{ \left(\sum_{l \in \{0,2\}} I_{C_l}(\mathbf{x}) \right) + \rho S(\mathbf{x}), 1 \right\}. \quad (2)$$

2.3 Total Variation Segmentation

To generate a final segmentation, the following continuous non-smooth energy functional $E_{seg}(u)$ [19, 20] is minimized:

$$\begin{aligned} \min_u E_{seg}(u) &= \min_u \int_{\Omega} g(\mathbf{x}) |\nabla u(\mathbf{x})| d\mathbf{x} + \lambda \int_{\Omega} u(\mathbf{x}) \cdot w(\mathbf{x}) d\mathbf{x} \\ \text{s.t.} \quad u &\in C_{box} = \{u : u(\mathbf{x}) \in [0, 1], \forall \mathbf{x} \in \Omega\} \end{aligned} \quad (3)$$

where Ω denotes the image domain and $u \in C^1 : \Omega \mapsto \mathbb{R}$ is smooth. The first term denotes the g -weighted total variation (TV) semi-norm which is a reformulation of the geodesic active contour energy [21]. The edge function $g(\mathbf{x})$ is defined as

$$g(\mathbf{x}) = e^{-\alpha \|\nabla I(\mathbf{x})\|^\beta}, \alpha, \beta > 0, \quad (4)$$

where $\nabla I(\mathbf{x})$ is the gradient of the input image, thus attracting the segmentation towards large gradients. The second term in Eq. (3) is the data term with w describing a weighting map. The values in w have to be chosen negative if u should be foreground and positive if u should be background. If values in w are set to zero, the pure weighted TV energy is minimized seeking for a minimal contour length segmentation. We use the refined outputs from the previous classification step (Eqs. (1) and (2)) and introduce a threshold τ to ensure a minimum class confidence in a map p :

$$p(\mathbf{x}) = \begin{cases} 0 & \text{if } p(\mathbf{x}) < \tau \\ w(\mathbf{x}) & \text{otherwise} \end{cases}. \quad (5)$$

The weighting map w is derived by applying the logit transformation:

$$w(\mathbf{x}) = \begin{cases} -(\log(p_{fg}(\mathbf{x})) - \log(1 - p_{fg}(\mathbf{x}))) & \text{if } p_{fg}(\mathbf{x}) > p_{bg}(\mathbf{x}) \\ \log(p_{bg}(\mathbf{x})) - \log(1 - p_{bg}(\mathbf{x})) & \text{if } p_{fg}(\mathbf{x}) \leq p_{bg}(\mathbf{x}) \end{cases}. \quad (6)$$

The regularization parameter λ defines the trade-off between our data term and the weighted TV semi-norm. The stated convex problem in Eq. (3) can be solved for its global optimum efficiently using the primal-dual algorithm [22], which can be implemented very efficiently using NVidia CUDA, thus making use of the parallel computing power of recent GPUs. As the segmentation u is continuous, the final segmentation is achieved by thresholding u with a value of 0.5. We optimize the free parameters α , β and λ by performing a grid search in a suitable range of these values ($\alpha \in [0.5, 15]$, $\beta \in [0.35, 0.95]$ and $\lambda \in [0.01, 10]$), where all 85 annotated training images are used to tune these parameters based on the Dice coefficient.

2.4 Implementation Details

2.4.1 Training Dataset Sampling

For the sake of execution speed when using a sliding window approach, the images are rescaled to half resolution prior to classification and upsampled with bilinear interpolation to their original size afterwards. The size of the input patch $I(\mathbf{x})$ is chosen to be 101×101 pixels, such that sufficient contextual information is available to classify the center pixel.

The majority of training images (79) have a size of 775×522 pixels, and resizing reduces them to 387×261 pixels. If we just considered the valid part without border extension for sampling the patches for the training dataset, we would actually lose approximately 46% of the labeled pixels when using a patch size of 101×101 pixels. On the other hand, we would introduce a significant number of boundary artifacts by artificially extending the border to make use of all labeled pixels. Fortunately, most images are tiles of a bigger image and can thus be stitched seamlessly to obtain a total of 19 images³ (Fig. 5), where we can sample enough patches without heavily relying on artificial border extension.

In principle, we pursued the same sampling strategy for the *Separator-Net*, but were required to create the ground truth labels manually. We annotated all pixels that belong to a structure very close to two or more gland borders. The green lines in Fig. 5 illustrate the additional manual annotation of the separating structures. Due to the low number of foreground samples when compared to the *Object-Net*, the number of foreground samples for the *Separator-Net* was artificially increased by exploiting the problem’s requirement for rotation-invariance and adding nine additional rotated versions of the patch, i.e. every 36° .

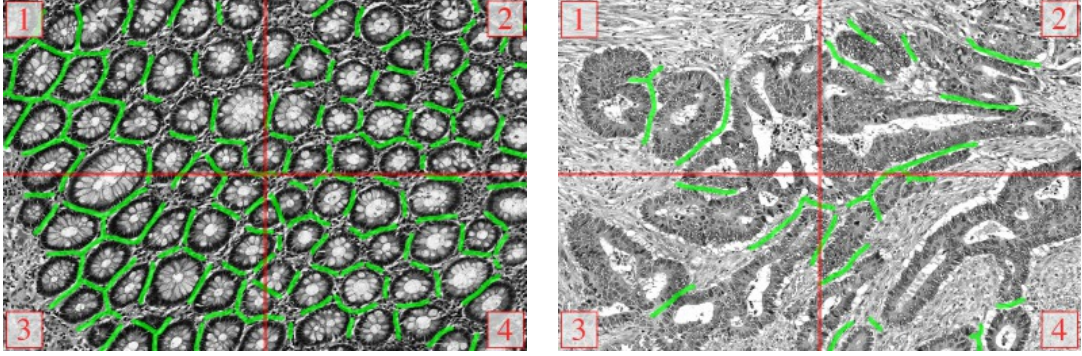


Fig. 5: Manual ground truth annotations for gland-separating structures. Stitched images from four tiles (numbers in red boxes), red lines denote the tile borders. Manual annotations of pixels belonging to gland-separating structures are shown as green lines, the thickness of lines is increased for better illustration.

2.4.2 CNN Training

Both CNNs were trained on a balanced training set of 125,000 image patches per class. Patches in the training sets were sampled at random from the available pool of training

³In one case, stitching was not possible, since only 3 tiles were available. These 3 tiles, and the remaining 6 images, that were not part of a bigger image, were treated as individual images.

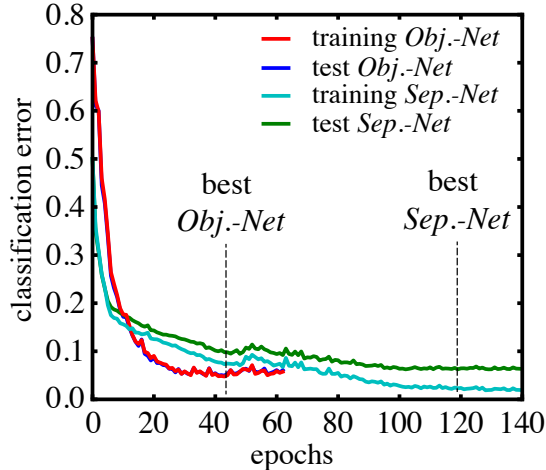


Fig. 6: CNN training progress. Classification error over epochs on a subset of the training data (training error), and on the held-out test set. The *Object-Net* reaches below 4.9% test error after 43 epochs, the *Separator-Net* reaches 6.2% test error after 119 epochs.

images. Training and test sets reflect approximately the same distribution of samples over images. The size of the minibatches in the MBSGD was set to 200 samples and the networks were trained until the stopping criterion was met: no further improvement of the error rate on a held-out test set over 20 epochs. We set the initial learning rate $\eta_0 = 0.0025$, with a linear decay saturating at $0.2\eta_0$ after 100 epochs. For all layers, a weight decay was chosen to be 0.005 and the dropout rate was set to 0.5. We used an adaptive momentum term starting at 0.8 and increasing to 0.99 after 50 epochs, such that with progressing training the updates are influenced by a larger number of samples than at the beginning.

Fig. 6 shows the classification error rate as a function of the training duration in epochs. Each class was represented with 5,000 samples in the test set for the *Object-Net*, and 10,000 for the *Separator-Net*, respectively. The training error is actually estimated on a fixed subset of the training data (20,000 samples), to get an intuition when overfitting starts. The *Object-Net* achieves the best performance after 43 epochs, with a minimum training error of 0.0475 and a minimum test error of 0.0492. Training of the *Separator-Net* continued until the lowest training error of 0.0231 and test error of 0.0624 was reached after 119 epochs. Fig. 7 shows the learned filters of the first convolutional layer in both networks. The CNN models were implemented in Pylearn2 [23], a machine learning library built on top of Theano [24, 25].

3 Results

3.1 Colon Gland Segmentation

The grid search resulted in $\alpha = 10$, $\beta = 0.95$ and $\lambda = 0.1$ as parameters optimizing the TV segmentation based on the Dice score. The confidence threshold for foreground and background was determined empirically and fixed to $\tau = 0.65$. Separator predictions were fully considered for refining the *Object-Net* predictions ($\rho = 1$).

In Table 1, we report performance metrics⁴ for detection (precision, recall, F1-

⁴The evaluation scripts were kindly provided by the contest organizers and are avail-

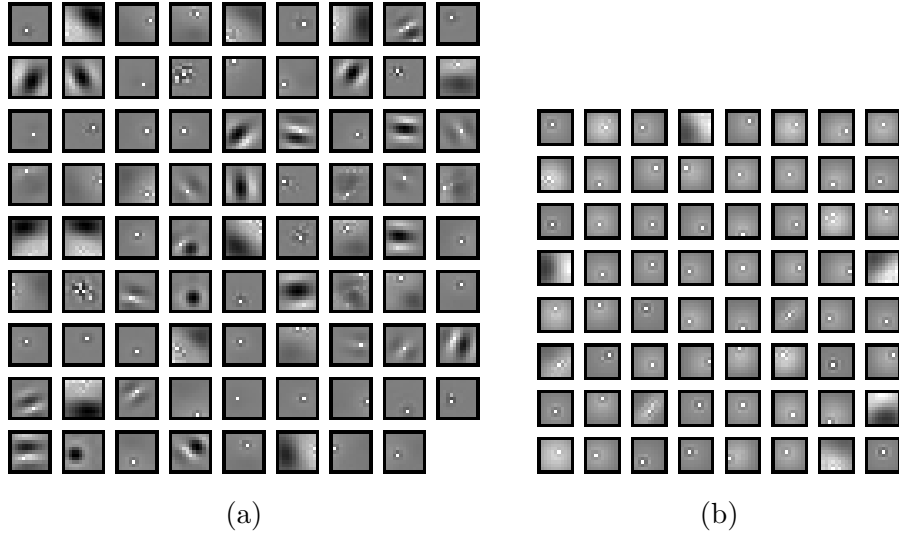


Fig. 7: CNN training results. (a) 80 11×11 filters of the first layer in *Object-Net* and (b) 64 9×9 filters in the *Separator-Net*.

Table 1: Segmentation performance metrics for the *Warwick-QU* dataset used in the *GlaS@MICCAI2015* challenge.

Dataset	Precision	Recall	F1-score	Object-Dice	Hausdorff
<i>without separator refinement</i>					
Training	0.97 (0.09)	0.67(0.21)	0.78(0.17)	0.81(0.16)	116.89(115.18)
Test A	0.83 (0.22)	0.60(0.24)	0.67(0.20)	0.70(0.15)	137.44(78.53)
Test B	0.70 (0.35)	0.48(0.30)	0.50(0.26)	0.58(0.19)	249.37(114.69)
<i>with separator refinement</i>					
Training	0.91(0.15)	0.85 (0.14)	0.87 (0.12)	0.88 (0.09)	61.36 (61.36)
Test A	0.67(0.24)	0.77 (0.22)	0.68 (0.20)	0.75 (0.13)	103.49 (72.38)
Test B	0.51(0.30)	0.70 (0.32)	0.55 (0.28)	0.61 (0.22)	213.58 (119.15)

Metrics are reported as mean and standard deviation, best results are printed in bold. Performance on the training set is reported on all 85 training images. Test set A consists of 60 images, test set B of 16 images. Except for values of the Hausdorff distance, higher values are superior.

score), segmentation (object-level Dice), and shape (Hausdorff distance) on the training set, as well as test set A and B as mean and standard deviation (SD). Blobs with an area less than 500 pixels were removed and all remaining blobs were labeled with unique identifiers before computing the measures.

Compared to using predictions only from the *Object-Net*, the segmentation performance improved with separator refinement. Malignant cases are harder to segment due to their irregular shape and pathological variations in the tissue. Fig. 8 illustrates some qualitative example segmentation results on the training data set, Fig. 9 and Fig. 10 show results on test set A and B, respectively.

able from <http://www2.warwick.ac.uk/fac/sci/dcs/research/combi/research/bic/glascontest/evaluation/>.

The average total runtime for segmenting a 577×522 image is 5 minutes using an NVidia GeForce Titan Black 6GB GPU.

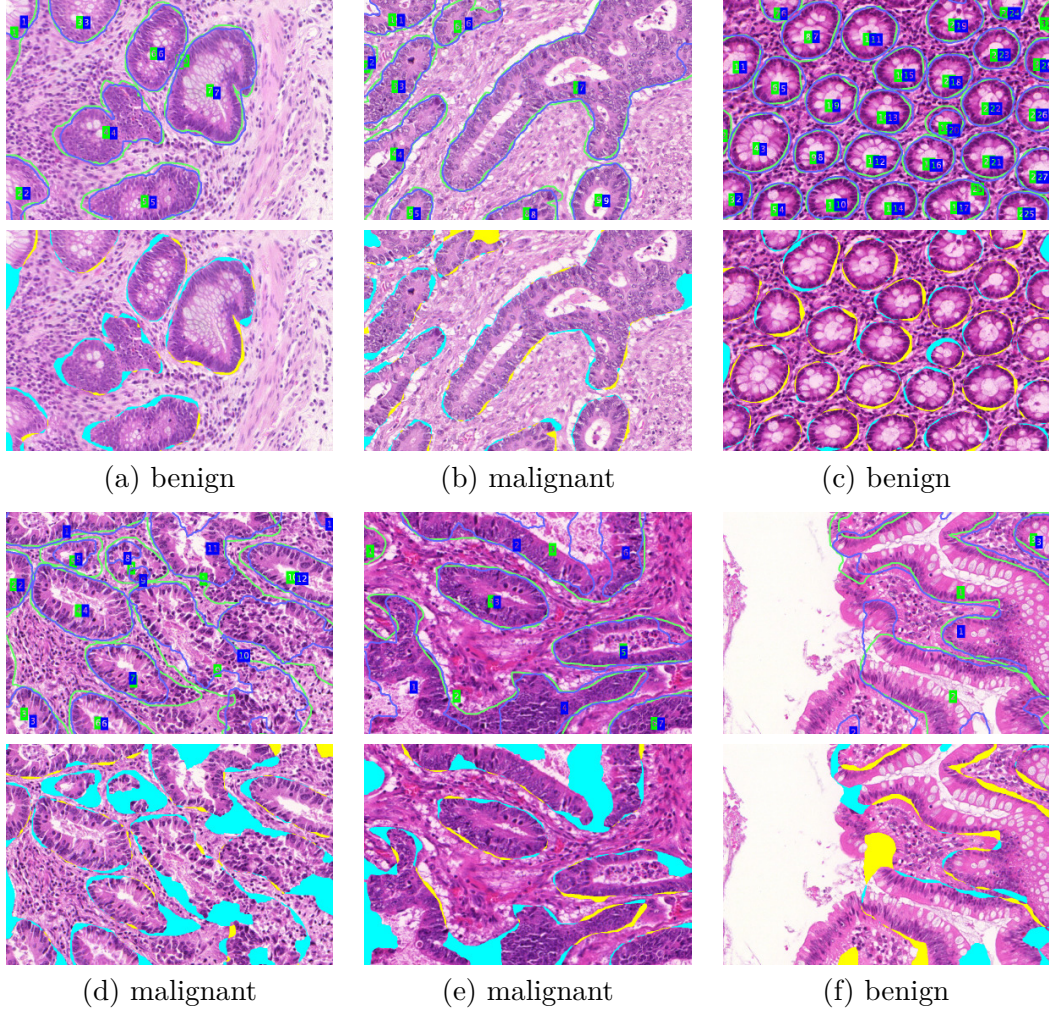


Fig. 8: Qualitative segmentation results on images of the training dataset. Even rows show the outline of ground truth in green and the segmentation result in blue. The numbers refer to the unique objects within the image. Odd rows show the segmentation difference: false negative pixels are colored in cyan, and false positives are colored in yellow. (a-c) show examples, where our segmentation algorithm works well, (d-f) show different types of segmentation errors.

3.2 Benignity and Malignancy Classification

In the proposed approach, the *Object-Net* inherently learns a discrimination of benign ($c = 0$) and malignant ($c = 1$) tissue, since the labels for benign and malignant are available in the training dataset and we defined a four-class classification problem. Instead of combining the probability maps for glands and background as done for segmentation, we combine the maps for benignity and malignancy. Subsequently, the

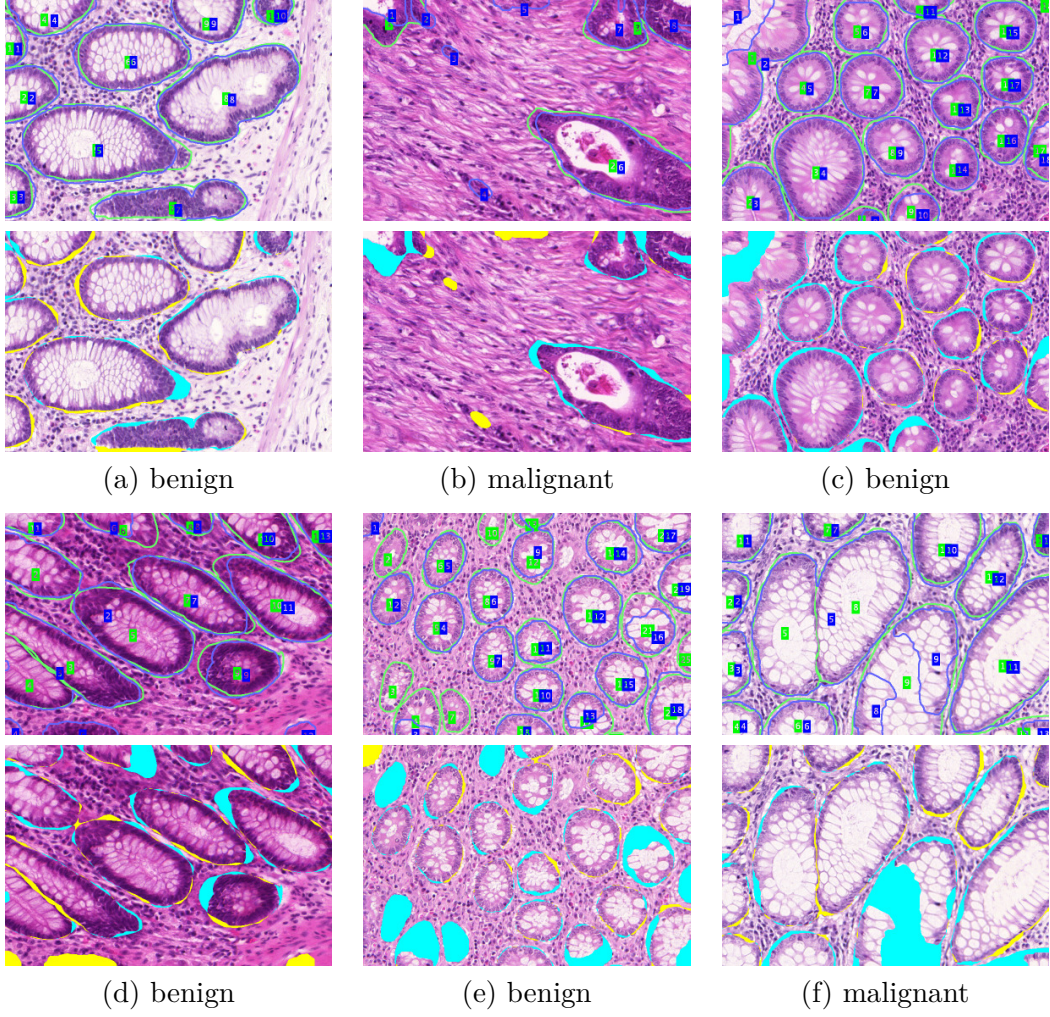


Fig. 9: Qualitative segmentation results on images of test dataset A. Even rows show the segmentation (blue outline) and ground truth (green outline), odd rows show the differences, where false negative pixels are cyan, and false positive pixels are yellow. (a-c) show reasonable segmentation results, in (d-f) different segmentation errors are shown.

average probabilities for a benign case can be computed as

$$\bar{P}(c=0|I_{C_0}, I_{C_1}) = \frac{1}{|\Omega|} \sum_{\mathbf{x} \in \Omega} I_{C_0}(\mathbf{x}) + I_{C_1}(\mathbf{x}), \quad (7)$$

and for a malignant case as

$$\bar{P}(c=1|I_{C_2}, I_{C_3}) = \frac{1}{|\Omega|} \sum_{\mathbf{x} \in \Omega} I_{C_2}(\mathbf{x}) + I_{C_3}(\mathbf{x}), \quad (8)$$

where $|\Omega|$ is the number of pixels in the image domain Ω . The maximum of both values finally indicates the prediction:

$$c^* = \arg \max_c \left\{ \bar{P}(c|\cdot, \cdot) \right\}. \quad (9)$$

We evaluated the classification performance for benign and malignant tissue on the two test sets A and B and achieved an accuracy of 98.33% and 93.75%. The average

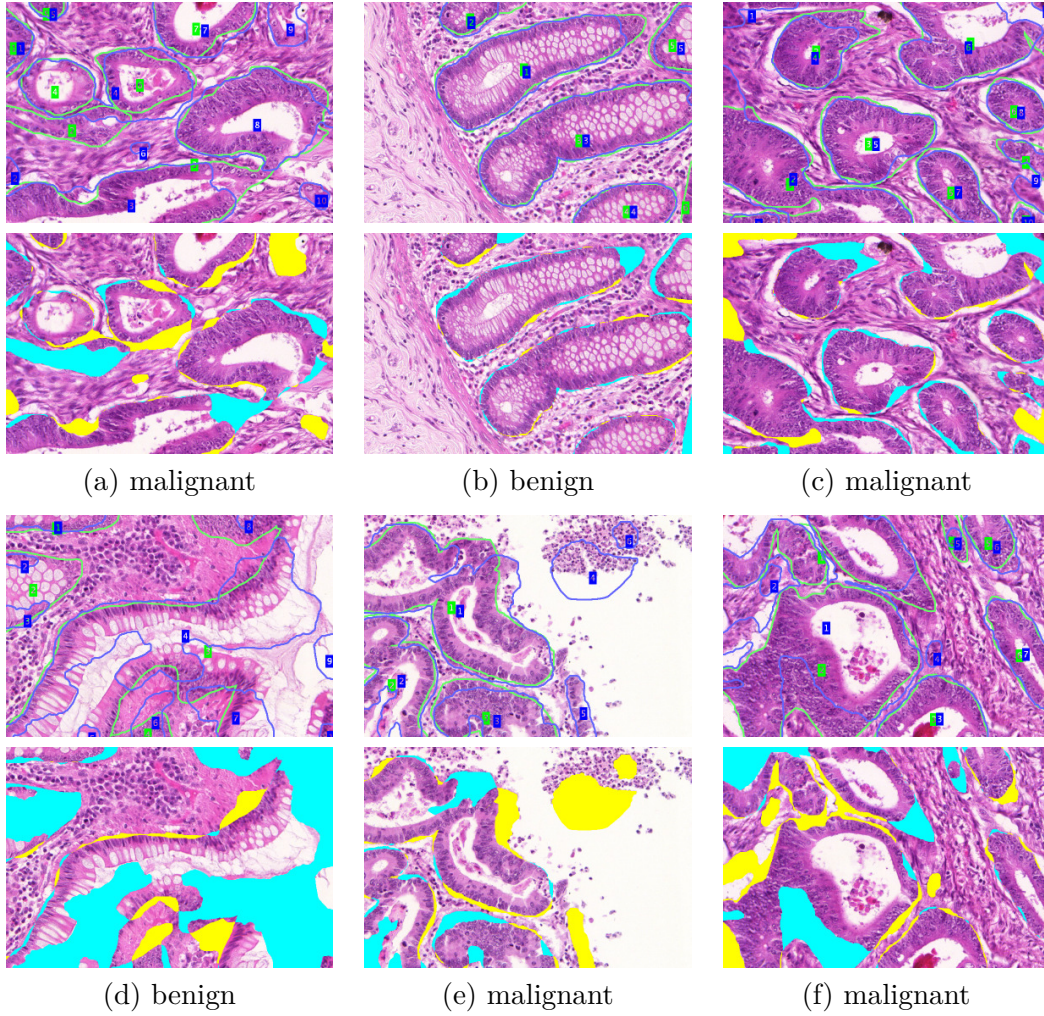


Fig. 10: Qualitative segmentation results on images of test dataset B. Even rows show the segmentation (blue outline) and ground truth (green outline), odd rows show the differences, where false negative pixels are cyan, and false positive pixels are yellow. (a-c) show reasonable segmentation results, in (d-f) different segmentation errors are shown.

(SD) decision confidence in test set A was 0.84(0.13) for benign and 0.81(0.11) for malignant, and in test set B 0.74(0.11) and 0.86(0.15), respectively.

4 Discussion and Conclusions

This paper presented a method to segment glands in H&E stained histopathological images of colorectal cancer using deep convolutional neural networks and total variation segmentation. As our main contribution, we showed that segmentation results can be greatly improved when the predictions of the *Object-Net* are refined with the learned gland-separating structures of the *Separator-Net*. Adding the separators does not only regulate the trade-off between precision and recall, but generally improves the performance scores for detection (F1-score), segmentation (Dice) and shape (Hausdorff). The final ranking as well as the test set performance results of other

algorithms participating in this challenge are available online at the contest website⁵, which is continuously being updated by algorithms from new participating groups.

Our approach inherently allows to very accurately discriminate benign and malignant cases, because the *Object-Net* was trained on labels for both cases. The average confidence for a decision towards benignity and malignancy is acceptable. Nevertheless, we cannot distinguish more detailed histologic grades among these cases, since there was no information (e.g. high- or low-grade) available in addition to the segmentation ground truth.

Acknowledgements

The authors are grateful to the organizers of the *GlaS@MICCAI2015* challenge for providing (i) the *Warwick-QU* image dataset, and (ii) the MATLAB evaluation scripts for computing performance measures that are comparable among the participating teams. Further thanks goes to Julien Martel for fruitful discussions in early phases of this challenge.

References

- [1] C. Gunduz-Demir, M. Kandemir, A. B. Tosun, and C. Sokmensuer. Automatic segmentation of colon glands using object-graphs. *Medical Image Analysis*, 14(1):1 – 12, 2010.
- [2] A.B. Tosun and C. Gunduz-Demir. Graph run-length matrices for histopathological image segmentation. *IEEE Transactions on Medical Imaging*, 30(3):721–732, March 2011.
- [3] K. Sirinukunwattana, D.R.J. Snead, and N.M. Rajpoot. A stochastic polygons model for glandular structures in colon histology images. *IEEE Transactions on Medical Imaging*, PP(99):1–1, 2015.
- [4] R. Farjam, H. Soltanian-Zadeh, K. Jafari-Khouzani, and R. A. Zoroofi. An image analysis approach for automatic malignancy determination of prostate pathological images. *Cytometry Part B: Clinical Cytometry*, 72B(4):227–240, February 2007.
- [5] S. Naik, S. Doyle, S. Agner, A. Madabhushi, M. Feldman, and J. Tomaszewski. Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology. In *IEEE International Symposium on Biomedical Imaging – ISBI*, pages 284–287, May 2008.
- [6] J. P. Monaco, J. E. Tomaszewski, M. D. Feldman, I. Hagemann, M. Moradi, P. Mousavi, A. Boag, C. Davidson, P. Abolmaesumi, and A. Madabhushi. High-throughput detection of prostate cancer in histological sections using probabilistic pairwise markov models. *Medical Image Analysis*, 14(4):617–629, August 2010.
- [7] Y. Peng, Y. Jiang, L. Eisengart, M. Healy, F. Straus, and X. Yang. Computer-aided identification of prostatic adenocarcinoma: Segmentation of glandular structures. *Journal of Pathology Informatics*, 2(1), 2011.
- [8] K. Nguyen, A. Sarkar, and A. K. Jain. Structure and context in prostatic gland segmentation and classification. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012*, pages 115–123, 2012.
- [9] S. Rashid, L. Fazli, A. Boag, R. Siemens, P. Abolmaesumi, and S. E. Salcudean. Separation of benign and malignant glands in prostatic adenocarcinoma. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, pages 461–468, 2013.

⁵<http://www2.warwick.ac.uk/fac/sci/dcs/research/combi/research/bic/glascontest/results/>

- [10] Y. LeCun, K. Kavukcuoglu, and C. Farabet. Convolutional networks and applications in vision. In *IEEE International Symposium on Circuits and Systems – ISCAS*, pages 253–256, May 2010.
- [11] B. Pang, Y. Zhang, Q. Chen, Z. Gao, Q. Peng, and X. You. Cell nucleus segmentation in color histopathological imagery using convolutional networks. In *Chinese Conference on Pattern Recognition – CCPR*, pages 1–5, October 2010.
- [12] D. C. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber. Mitosis detection in breast cancer histology images with deep neural networks. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, pages 411–418, 2013.
- [13] C. D. Malon and E. Cosatto. Classification of mitotic figures with convolutional neural networks and seeded blob features. *Journal of Pathology Informatics*, 4:9, May 2013.
- [14] M. Habibzadeh, A. Krzyżak, and T. Fevens. White blood cell differential counts using convolutional neural networks for low resolution images. In *Artificial Intelligence and Soft Computing*, pages 263–274. Springer Berlin Heidelberg, 2013.
- [15] A. C. Ruifrok and D. A. Johnston. Quantification of histochemical staining by color deconvolution. *Analytical and Quantitative Cytology and Histology*, 23(4):291–299, August 2001.
- [16] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676–682, July 2012.
- [17] K. Zuiderveld. Contrast limited adaptive histogram equalization. In *Graphics gems IV*, pages 474–485. Academic Press Professional, Inc., 1994.
- [18] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, November 1998.
- [19] C. Reinbacher, T. Pock, C. Bauer, and H. Bischof. Variational Segmentation of Elongated Volumetric Structures. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3177–3184, 2010.
- [20] K. Hammernik, T. Ebner, D. Stern, M. Urschler, and T. Pock. Vertebrae segmentation in 3D CT images based on a variational framework. In *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, pages 227–233. Springer, 2015.
- [21] X. Bresson, S. Esedoglu, P. Vandergheynst, J.-P. Thiran, and S. Osher. Fast global minimization of the active contour/snake model. *Journal of Mathematical Imaging and Vision*, 28(2):151–167, 2007.
- [22] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.
- [23] I. J. Goodfellow, D. Warde-Farley, P. Lamblin, V. Dumoulin, M. Mirza, R. Pascanu, J. Bergstra, F. Bastien, and Y. Bengio. Pylearn2: a machine learning research library. *arXiv preprint arXiv:1308.4214*, 2013.
- [24] J. Bergstra, O. Breuleux, F. Bastien, P. Lamblin, R. Pascanu, G. Desjardins, J. Turian, D. Warde-Farley, and Y. Bengio. Theano: a CPU and GPU math expression compiler. In *Proceedings of the Python for Scientific Computing Conference (SciPy)*, June 2010.
- [25] F. Bastien, P. Lamblin, R. Pascanu, J. Bergstra, I. J. Goodfellow, A. Bergeron, N. Bouchard, and Y. Bengio. Theano: new features and speed improvements. Deep Learning and Unsupervised Feature Learning NIPS 2012 Workshop, 2012.