 DOI: 10.3217/978-3-85125-663-5-47

# A Two-Stage Classifier for Collagen in Electron Tomography Images using a Convolutional Neural Network and TV Segmentation

Verena Horak[1,2] and Kristian Bredies[1,2]

***Abstract*— We present an easily realizable practical strategy for the segmentation of tissue types in microscopy images of biological tissue. The strategy bases on a convolutional neural network (CNN) classifier that requires a low amount of manually-labeled data. Spatial regularity of the segmented images is enforced by a total variation (TV) regularization approach. The proposed strategy is applied to and tested on collagen segmentation in electron tomography image stacks.**

## I. INTRODUCTION

In cell biology, one often seeks segmentation of microscopy images according to some biological tissue types or structures. Since manual labeling is very tedious, time consuming, and error-prone, automated classification algorithms are desired. In this respect, many segmentation methods have been proposed and studied, usually focusing on cell segmentation [10] where cell boundaries are clearly defined. In contrast to that, more complex structures such as collagen, for instance, are often only vaguely visible and subject to noise, in particular, if the images were reconstructed from imperfect data, which is often the case in electron tomography. Machine-learning approaches have the potential to overcome this problem [6], [7] but require a sufficient amount of manually-labeled training data which is often not available. One approach is to generate more training data is to consider all possible patches of a specified size within the training images and to train a patch-based classifier. Such a classifier is, however, unaware of the spatial structure of the patches and usually produces binary images with irregular labeled regions. A shape regularization is thus necessary.

In this work, we report on the application of a binary two-stage classifier to automated collagen segmentation that, similar to [8], employs, on the one hand, a patch-based machine-learning approach and, on the other hand, provides regular label regions via total variation (TV) regularization. The method requires only a single manually labeled image stack as training data as well as minimal human interaction.

## II. THE TWO-STAGE CLASSIFIER

The implemented classifier uses the same ideas as [8] for a different application. We assume that a training set of images to label as well as a manually-generated labeling is available, where is latter might the affected by human error. We further assume that only local information is necessary in order to determine whether an image pixel has to be labeled or not.

[1]Institute of Mathematics and Scientific Computing, University of Graz, Austria, Email: {verena.horak, kristian.bredies}@uni-graz.at

[2]BioTechMed-Graz, Austria, https://biotechmedgraz.at/en/

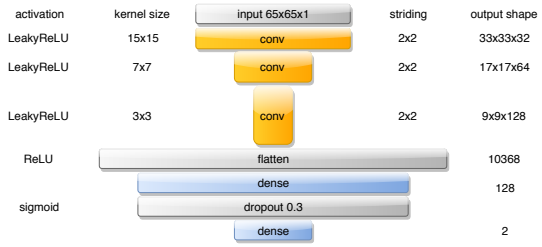


Fig. 1. Example of a weak classifier for patches of size $65 \times 65$.

For this reason, we pursue a patch-based approach which leads to all possible patches and single 0/1 values constituting the training set pairs. This might, however, lead to artifacts and irregular contours. Hence, we propose to learn the basic distribution of collagen based on a patch surrounding the pixel under investigation, and then using a regularization to obtain artifact-free, regular regions.

### A. Stage 1: A weak patch-based machine-learning classifier

Our approach is to predict each pixel in the label image from a so-called patch, that is a sub-image of a specified window size with the corresponding pixel to label as its center. With $r$, $c$, $w$ indicating the number of rows, columns of the image, and the window size, respectively, we obtain from every image $(r-w+1)\cdot(c-w+1)$ of such patches of size $w \times w$, provided that $r \geq w$ and $c \geq w$. Using a certain subset of these patches as training data, a convolutional neural network (CNN) is trained to weakly predict the label corresponding to the center pixel of a given patch, i.e., the outcome is not a binary 0/1-result but a value in $[0,1]$ that can be interpreted as a probability. An example for such a predictor can be found in Fig. 1.

### B. Stage 2: A binary TV-regularized classifier

Due to the patch-based prediction, spatial information between neighboring pixels is not taken into account. This typically yields irregular contours and artifacts. To get rid of these disruptive factors and to obtain the final binary result, a classifier based on a total variation (TV) regularizer is used [2]. More specifically, if $l_0$ is image of predicted labels from Stage 1 with values in $[0,1]$, we solve the problem

$$\min_{l} \int \lambda (b - l_0) \cdot l \, \mathrm{d}x + \mathrm{TV}(l) \quad \text{subject to} \quad 0 \leq l \leq 1, \quad (1)$$

where $b \in [0,1]$ is a bias parameter corresponding to a threshold and $\lambda > 0$ a regularization parameter controlling the regularity of the contours. Afterwards, the minimizer $l^*$ is thresholded (for instance, at 0.5), in order to obtain a

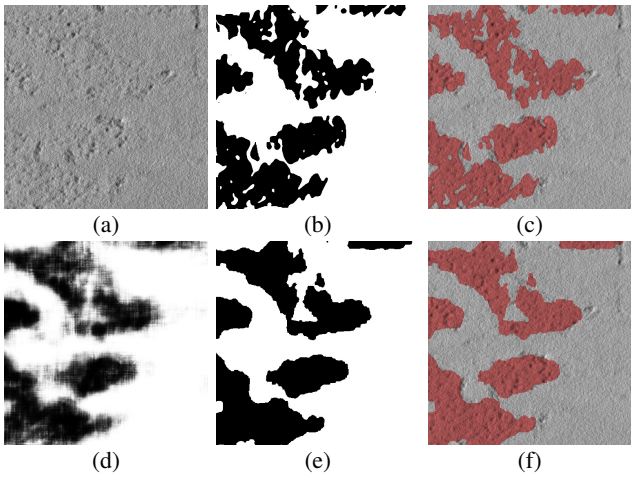

Fig. 2. Visualization of the two-stage classifier for a subregion of a test image. (a) the microscopy image to label, (b) the manually-generated label image, (c) the manual labeling laid over the microscopy image, (d) the weak prediction of Stage 1, (e) the regularized binary prediction of Stage 2, (f) the predicted label image (e) laid over the input image.

TABLE I

QUANTITATIVE EVALUATION OF BOTH STAGES OF THE CLASSIFIER. $P_i = v$ REPRESENTS A PREDICTION OF A PIXEL VALUE $v$ IN STAGE $i$, WHEREAS $L = v$ MEANS A PIXEL LABEL OF $v$.

| | $P_1 = 1$ | $P_1 = 0$ | $P_2 = 1$ | $P_2 = 0$ |
|---|---|---|---|---|
| $L = 1$ | 20.63% | 3.42% | 21.10% | 2.94% |
| $L = 0$ | 3.45% | 72.50% | 3.61% | 72.45% |

binary image as final result. This yields a global solution of the corresponding shape optimization problem where the constraints in (1) are replaced by $l \in \{0,1\}$ in each pixel [4]. A primal-dual algorithm [3] is used for the solution of (1).

## III. NUMERICAL EXPERIMENTS

Our numerical studies were carried out on an image stack of 100 images of size $2048 \times 2048$. These images were obtained by tomographic reconstruction from a tilt series of transmission electron microscopy (TEM) images of human aortic tissue. The reconstruction was computed with the IMOD software package [9] and the collagen in this data set was labeled manually, see Figure 2 for a section.

### A. Training the patch-based classifier

For the patch building process, a window size of $65 \times 65$ was chosen, yielding $(2048 - 65 + 1)^2 = 3936256$ patches for each image. For Stage 1, the patches of 80 images were subjected to a random permutation and taken as training data for the CNN described in Fig. 1 with a binary cross entropy loss function. The network was realized by Keras 2.2.4 [5] and Tensorflow 1.12.0 [1], and the computations were performed on a NVIDIA Tesla K40c GPU. The training was stopped after 90 hours of computation.

### B. Classification results

We present quantitative and visual results for both stages of the classifier. Table I, left shows the performance of Stage 1 if a threshold of 0.43 is taken for binary classification. This value was manually chosen to maximize the accuracy and leads to an accuracy of 93.13%, a precision of 85.68%, and a phi coefficient of 81.2%. For Stage 2, the performance is displayed in Table I, right. An accuracy of 93.45%, a precision of 85.41%, and a phi coefficient of 82.25% are obtained by choosing $b = 0.4$ and $\lambda = 4.5$. These values were manually found to be optimal by analyzing the accuracy as well as by visual inspection. The effect of each stage on a region of a test image is shown in Fig. 2. One can see that the TV-regularization step in Stage 2 is indeed beneficial for the visual appearance of the predicted label image. The quantitative results in Table I moreover underline that Stage 2 does not deteriorate the performance of the learned classifier in Stage 1, and even seems to improve it.

## IV. CONCLUSIONS

This study shows the effectiveness of the presented method for labeling problems that, e.g., researchers working with microscopy images commonly face. It can easily be adapted to other types of tissues beyond collagen by providing suitable manually-segmented data. The method could, for instance, enable to significantly increase the throughput for image-based analysis of biological tissue samples.

## ACKNOWLEDGMENTS

The training and test image stack as well as the manual segmentation are courtesy of the Institute of Biomechanics, Graz University of Technology, and the Gottfried Schatz Research Center, Medical University of Graz. The authors thank Anna Pukaluk and Gerd Leitinger for providing the data sets and valuable feedback on the algorithms.

## REFERENCES

[1] M. Abadi *et al.*, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems," https://www.tensorflow.org/, 2015.
[2] V. Caselles, A. Chambolle, and M. Novaga, *Total Variation in Imaging*. Springer New York, 2011, pp. 1016–1057.
[3] A. Chambolle and T. Pock, "A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging," *J. Math. Imaging Vis.*, vol. 40, no. 1, pp. 120–145, 2011.
[4] T. Chan, S. Esedoglu, and M. Nikolova, "Algorithms for Finding Global Minimizers of Image Segmentation and Denoising Models," *SIAM J. Appl. Math.*, vol. 66, no. 5, pp. 1632–1648, 2006.
[5] F. Chollet *et al.*, "Keras," https://keras.io/, 2015.
[6] D. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images," in *Proceedings of NIPS 2012*, pp. 2843–2851.
[7] A. Fakhry, T. Zeng, and S. Ji, "Residual Deconvolutional Networks for Brain Electron Microscopy Image Segmentation," *IEEE Transactions on Medical Imaging*, vol. 36, no. 2, pp. 447–456, 2017.
[8] P. Kainz, M. Pfeiffer, and M. Urschler, "Segmentation and Classification of Colon Glands with Deep Convolutional Neural Networks and Total Variation Regularization," *PeerJ*, vol. 5, p. e3874, 2017. [Online]. Available: https://doi.org/10.7717/peerj.3874
[9] J. R. Kremer, D. N. Mastronarde, and J. McIntosh, "Computer Visualization of Three-Dimensional Image Data Using IMOD," *Journal of Structural Biology*, vol. 116, no. 1, pp. 71 – 76, 1996.
[10] E. Meijering, "Cell Segmentation: 50 Years Down the Road," *IEEE Signal Proc Mag*, vol. 29, 2012.