Trainable Segmentation Based on Local-level and Segment-level Feature Extraction

Radim Burget[1], Vaclav Uher[1], and Jan Masek[1]

Brno University of Technology, Czech Republic, European Union

Abstract. This paper deals with the segmentation of neuronal structures in electron microscope (EM) stacks, which is one of the challenges of the ISBI 2012 conference. The data for the challenge consists of a stack of 30 EM slices for training and 30 EM stacks for testing. The training data was labelled by an expert human neuroanatomist. In this paper a segmentation using local-level and segment-level features and machine learning algorithms was used. The results achieved on the ISBI 2012 challenge test set were: the Rand error: 0.139038440, warping error: 0.002641296 and pixel error: 0.102285508. The main criterion for segmentation evaluation was the Rand error.

1 Introduction

The problem with understanding and analysing the human or animal brain[5], [1] lies especially in its complexity. To visualize the brain structure and synapsis connections it is necessary to use special devices such as high-resolution electron microscope (EM). This imaging device produces huge amounts of image data. In order to understand the patterns at both the micro and the macro level, the image data should be segmented according to structural and functional modules. For an expert human neuroanatomist, segmentation of an neuro-images is a trivial task but, unfortunately, it is very time consuming.

This paper deals with the segmentation of neuronal structures in EM stacks, which is one of the challenges of the IEEE International Symposium on Biomedical Imaging (ISBI) 2012. The main objective of this paper is to find new more accurate algorithms for trainable image segmentation of brain structures retrieved from the EM device. The motivation behind the segmentation is that with the use of the 2D segmentation more accurate 3D segmentation could be possible, consequently more accurate model of e.g. 3D brain can be extracted and this can lead to better understanding of the brain. The results achieved are evaluated by pixel error, the Rand error and warping error [2]. More information about the measures is given in section 2 of this paper.

The rest of the paper is structured as follows. The next section describes data used for training and validation. It also describes methods used for the evaluation of segmentation performance. Section 3 describes segmentation training using local-level and also segment-level features. Section 3.6 discusses the results and the last section concludes the paper.

2 Training and testing data set

The data used in this paper were taken from a previous work about micro- and macro-architectural analysis of Drosophila brain [1]. The data consists of two disjunctive sets: training set and testing set, each containing 30 EM slices. The training set was labelled by an expert human neuroanatomist according to the structural and functional modules of the brain. Selected slices and their corresponding labels are depicted in figure 1. The labels for the testing set have not been published but the results can be evaluated automatically by being submitted to a server, which has been provided by the challenge organizers.

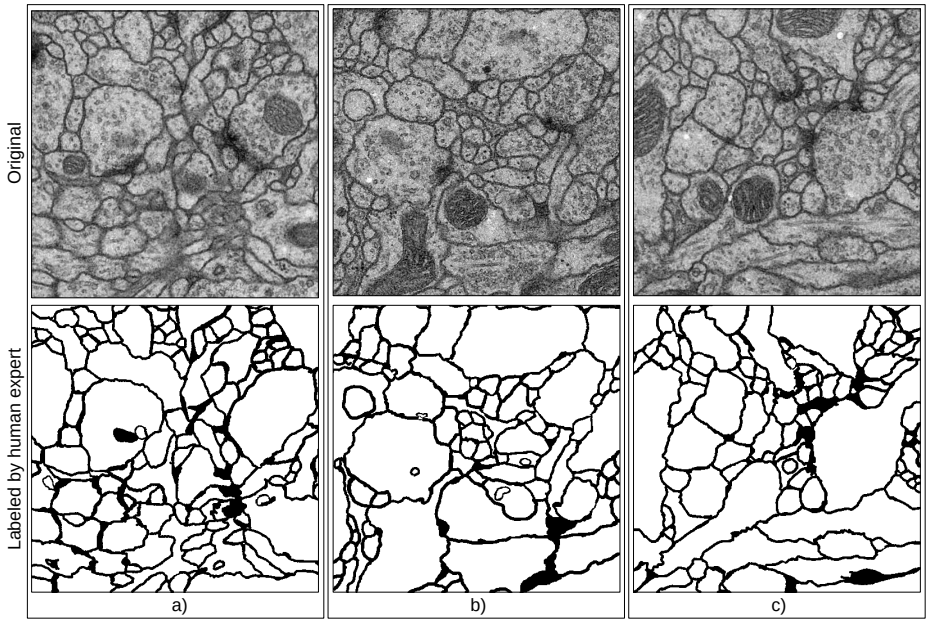

Fig. 1: Training data.

2.1 Performance evaluation

Altogether, three methods were used for the evaluation of segmentation performance : pixel error [2], warping error [2] and the Rand error [6], [8], [2].

The pixel error metric is relatively simple and expresses the square of the number of disagreements between the image and the template (labelled by an expert human radiologist). Unfortunately it suffers from a serious defect. It is overly sensitive to minor displacements in the location of the boundary. However, these disagreements in most of the cases do not lead to qualitative differences in the interpretation of the image.

The warping error metric tolerates disagreements over boundary location, penalizes topological disagreements and can be used directly as a cost function for learning boundary detection. The disadvantage of this approach is definitely its relatively high computational demands. Instead of focusing on the geometric differences (pixel disagreement) between two segments, the warping error focuses on the segments and measures the topological error between them [2].

The Rand error metric is based on the Rand index [6], [8], [2], which is a well-known measure of similarity between two data clusters using the following equation

$$E_{\text{Rand}} = \frac{a+b}{\binom{n}{2}}, \tag{1}$$

, where a denotes the number of pairs of pixels that are in the same object S_1 and in the same object in S_2 (i.e. they have the same label). b denotes the number of pairs of pixels that are in different objects S_1 and S_2 (i.e. they have different labels). n is the number of pixels in the image.

In the challenge, the Rand error was selected as the main criterion in the segmentation evaluation.

3 Trainable segmentation

First, the histograms of all training and testing images were equalized in order to reduce the intensity variability between slices. After the images were equalized, they were segmented using trainable segmentation based on local-level features. The trainable segmentation consists of several steps: 1) first the training points were selected, 2) then different features from the original images were extracted, 3) using the selected points and extracted features the segmentation was trained, 4) in order to achieve more accurate results, the parameters of the transforms were optimized using the genetic algorithm, 5) since the trainable segmentation kept some unwanted objects in the resulting segmented image, the resulting image was segmented again using statistical region merging [4] and different features were extracted from each segment, 6) according to the extracted features, unwanted segments were removed from the resulting image. The following subsections describe in detail each step mentioned above.

3.1 Point selection

The points were selected in order to give higher priority to areas where there was a risk of not recognizing the membrane or where there was a risk of the membrane merging with the nucleus. The points selected are depicted in figure 2, where the red points denote segments and the green points denotes segment boundaries.

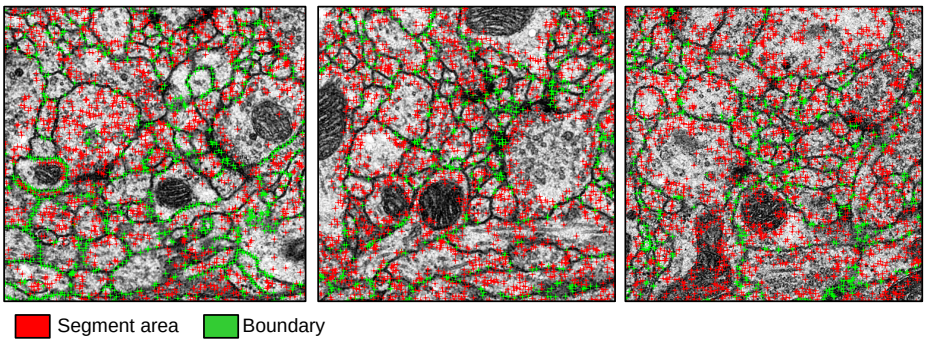

Fig. 2: Training point selection. Red points denotes points which should mark segment boundary, the green points denotes segment bodies.

3.2 Local-feature extraction

To get more information about each pixel, several transformations were applied to the original input images. Most of the image transforms use for the computation the value of each local pixel and also the values of surrounding pixels. Therefore by using several transforms it is possible to extract more complex information about each area of the image. Several different transformations were used with several different parameters, which were applied to the equalized images. The set of transforms used includes several relatively well-known transforms. In particular, there were used the original equalized image, several blurred images, several edge detection images, skeletonized images, median filter, minimum filter (searching for the minimal values in the given radius) and maximum filter (searching for the maximal value in the given radius) variance filter (highlights edges in the image by replacing each pixel with the neighbourhood variance). The transforms are treated in books about image processing, e.g. [3] and this paper will not discuss their details.

The blurring was performed using a Gaussian filter with σ values: 1.0, 1.5, 2.0, 3.5, 4.0, 5.0, 6.0, 7.0, 8.0, 12.0, 16.0, 24.0, and 32.0. A Sobel filter with the kernel size 3x3 (combination of vertical and horizontal kernels) was also used. Before the application of the Sobel filter the blur using Gaussian kernels with σ values 1.0, 2.0, 3.0, 6.0, 9.0, 12.0 and 24.0 was used. To obtain the skeletonized image, the equalized image was thresholded using the ISOData threshold [7] and the resulting binary skeleton was dilated by 1, 2 and three pixels. Selected results of transforms are depicted in figure 3.

3.3 Trainable segmentation

Using the points obtained as described in section 3.1 and using the transforms as described in section 3.2, a training data set was extracted. Since the training data set was structured in the tabular form, most of the known learning algorithms can be trained using this data.

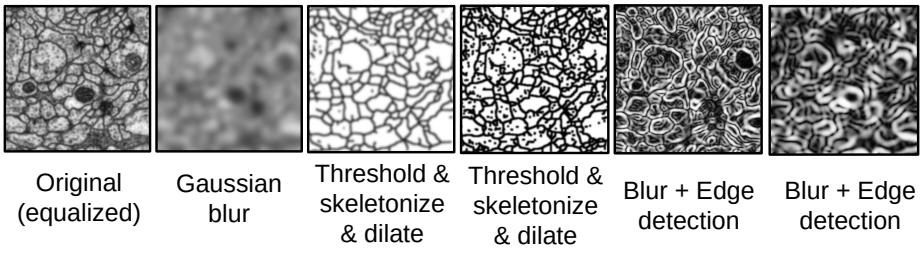

Fig. 3: Selected examples of extracted features.

The principle of transforming images into tabular data is depicted in figure 4. For each point a vector of features is extracted. This vector represents one sample and it forms one row in the resulting table. The values in this vector are integer numbers representing the resulting values at the given point position of a given transform (see figure 4). This sample is extracted for each point. Thus the number of rows in the resulting table equals the number of points.

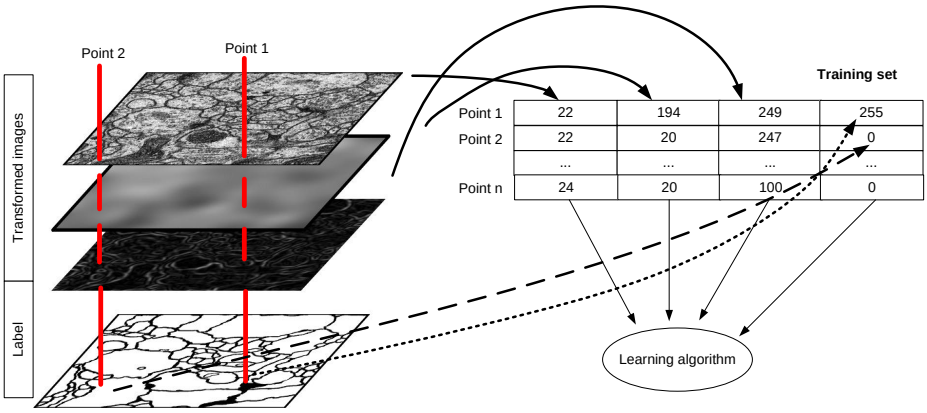

Fig. 4: Principle of segmentation training.

For training, several learning algorithms were examined, including Support Vector Machines (SVM), Decision Tree, Random Forests, and k-Nearest Neighbors (k-NN). Since the k-NN is part of the group of so-called lazy learners and the computationally most demanding classification logic is done during the classification and not learning, the k-NN is significantly time demanding during classification. Therefore this algorithm is not very suitable for the problem of image segmentation since for a 512x512 image it means $512^2 = 261144$ classifications of pixels. As mentioned above, not the pixel error of the algorithm was

important but how the algorithm was able to detect boundaries and separate membranes from other unwanted objects. A subjective evaluation method was used and the SVM with dot kernel was selected as the learning algorithm with the best properties.

3.4 Local-feature optimization

In order to optimize the parameters of the transforms used, a genetic algorithm optimization was used. For this purpose the probability of mutation was set to 50 %, the probability of cross-over was set to 25 % using the uniform cross-over operation. As a selection the tournament method was used. The results are depicted in figure 5.

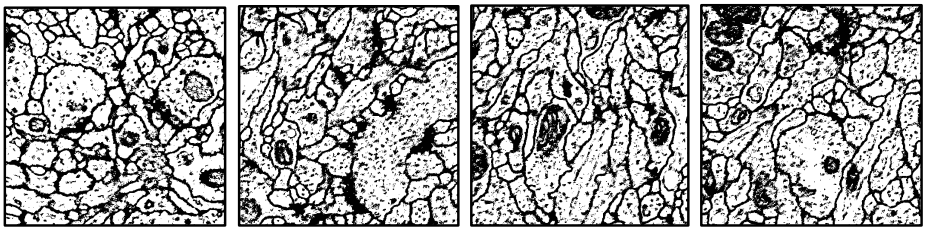

Fig. 5: Results of trainable segmentation on test set. For training the training set was used.

3.5 Segment-level feature extraction

As is clear from figure 5, the resulting images still contain many unwanted objects. In order remove these unwanted objects, the resulting images were segmented again, using the statistical region merging [4] method, and from each segment several segment-level features were extracted, including e.g. size, roundness, minimal gray value, maximal gray value, mean and median values, and several others. Each segment was also labelled as wanted or unwanted. Using the extracted features, a decision tree was trained and the unwanted objects were removed. The decision tree was used because the trained model is relatively easy to understand. The process scheme is depicted in figure 7.

3.6 Results and discussion

As is clear from the results, the segment-level segmentation succeeded in removing small objects but failed to remove some objects. Some bigger objects cannot be easily removed using this method because they are connected to the membrane. In some cases, however, the objects could be removed easily. This is due to the fact that the training set was not sufficient and should be improved.

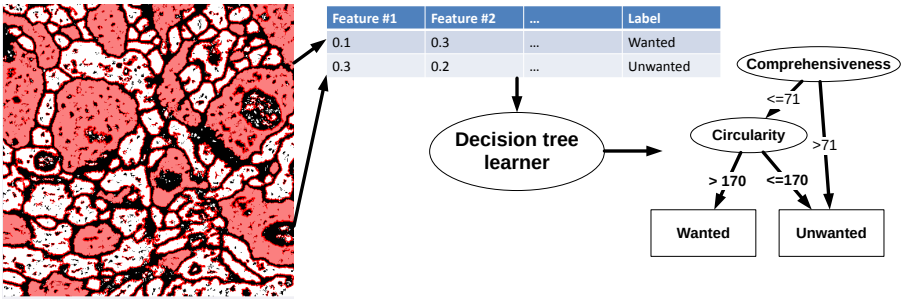

Fig. 6: Scheme of segment-level feature extraction and training.

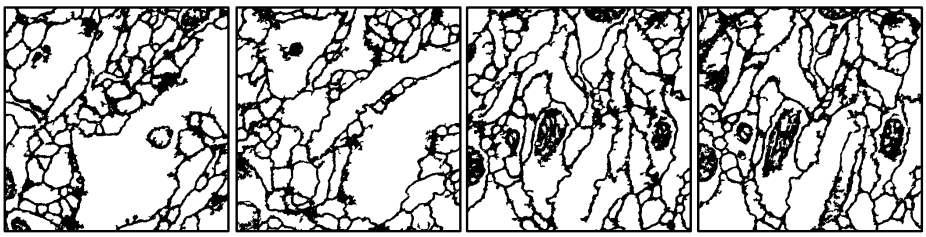

Fig. 7: After unwanted object removal depicted on the test set.

The results achieved on the ISBI 2012 challenge test set were: the Rand error: 0.139038440, warping error: 0.002641296, and pixel error: 0.102285508. The main criterion for the segmentation evaluation was the Rand error.

Further direction to be followed to enhance the results should be to use an extended set of transforms for better feature extraction, since the one used here was relatively simple. Also, preparing a better segment-level training set should give better results especially for the pixel error criterion. Another promising enhancement could be searching for broken lines (i.e. membranes) and trying to reconnect them. Another promising way could be trainable statistical region merging.

4 Conclusion

This paper deals with the segmentation of neuronal structures in EM stacks, which was one of the challenges of the ISBI 2012 conference. The data for the challenge consists of a stack of 30 EM slices for training and 30 EM stacks for testing. The training data was labelled by an expert human neuroanatomist in order to get the desired output of the segmentation. The problem is quite easy for any neuroanatomist, but rather time-consuming. In this paper, segmentation using local-level and segment-level features was used. The results achieved on the ISBI 2012 challenge test set were : the Rand error: 0.139038440, warp-

ing error: 0.002641296 and pixel error: 0.102285508. The main criterion for the segmentation evaluation was the Rand error.

The algorithms and source code including the processes described in this paper is published under open-source licence and can be downloaded[1] for free.

Acknowledgment

This research is part of the project reg. no CZ.1.07/2.3.00/20.0094 "Support for incorporating R&D teams in international cooperation in the area of image and audio signal processing" and is co-financed by the European Social Fund and the state budget of the Czech Republic.

References

1. Cardona, A., Saalfeld, S., Preibisch, S., Schmid, B., Cheng, A., Pulokas, J., Tomancak, P., Hartenstein, V.: An integrated micro- and macroarchitectural analysis of the drosophila brain by computer-assisted serial section electron microscopy. PLoS Biol **8**(10) (10 2010) e1000502
2. Jain, V., Bollmann, B., Richardson, M., Berger, D.R., Helmstaedter, M., Briggman, K.L., Denk, W., Bowden, J.B., Mendenhall, J.M., Abraham, W.C., Harris, K.M., Kasthuri, N., Hayworth, K.J., Schalek, R., Tapia, J.C., Lichtman, J.W., Seung, H.S.: Boundary learning by optimization with topological constraints. In: CVPR, IEEE (2010) 2488–2495
3. Nixon, M., Aquado, A.: Feature Extraction and Image Processing. Newnes (2002)
4. Nock, R., Nielsen, F.: Statistical region merging. IEEE Trans. on Pattern Analysis and Machine Intelligence **26** (2004) 1452–1458
5. Ragan, T., Kadiri, L., Venkataraju, K., Bahlmann, K., Sutin, J., Taranda, J., Arganda-Carreras, I., Kim, Y., Seung, H., Osten, P.: Serial two-photon tomography for automated ex vivo mouse brain imaging. Nat Methods (2012)
6. Rand, W.M.: Objective Criteria for the Evaluation of Clustering Methods. Journal of the American Statistical Association **66**(336) (1971) 846–850
7. Ridler T W Calvard, S.: Picture Thresholding Using an Iterative Selection Method. Systems, Man and Cybernetics, IEEE Transactions on **8**(8) (August 1978) 630–632
8. Unnikrishnan, R., Pantofaru, C., Hebert, M.: Toward objective evaluation of image segmentation algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence **29**(6) (June 2007) 929–944

[1] Download Rapidminer Extension for IMage MIning (IMMI) at http://splab.cz/en