High Resolution Segmentation of Neuronal Tissues from Low Depth-Resolution EM Imagery

Daniel Glasner[1,2,*], Tao Hu[1], Juan Nunez-Iglesias[1], Lou Scheffer[1], Shan Xu[1], Harald Hess[1], Richard Fetter[1], Dmitri Chklovskii[1], and Ronen Basri[1,2,*]

[1] Janelia Farm Research Campus, Howard Hughes Medical Institute
[2] Department of Computer Science and Applied Mathematics, Weizmann Institute of Science

Abstract. The challenge of recovering the topology of massive neuronal circuits can potentially be met by high throughput Electron Microscopy (EM) imagery. Segmenting a 3-dimensional stack of EM images into the individual neurons is difficult, due to the low depth-resolution in existing high-throughput EM technology, such as serial section Transmission EM (ssTEM). In this paper we propose methods for detecting the high resolution locations of membranes from low depth-resolution images. We approach this problem using both a method that learns a discriminative, over-complete dictionary and a kernel SVM. We test this approach on tomographic sections produced in simulations from high resolution Focused Ion Beam (FIB) images and on low depth-resolution images acquired with ssTEM and evaluate our results by comparing it to manual labeling of this data.

Keywords: Segmentation of neuronal tissues, Task-driven dictionary learning, Sparse over-complete representation, Connectomics.

1 Introduction

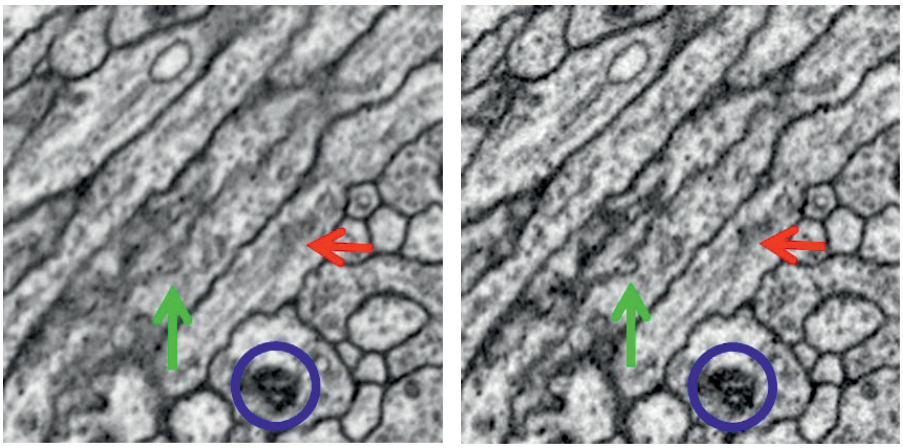
Recent years have seen several large scale efforts to recover the structure of neuronal networks of various animals' brains [1,2]. Detecting every single neuron and its synaptic connections to other neurons in dense neuronal tissues requires both high-resolution and high-throughput imaging techniques. Currently, the only technology that can potentially meet this challenge is high-throughput Electron Microscopy (EM) followed by automated image analysis, and finally manual proofreading [3]. Effective image analysis techniques can greatly speed up this process by reducing the need for manual labour.

High-throughput Electron Microscopy imagery of neuronal tissues can be obtained using serial section Transmission EM (ssTEM) technology. In ssTEM, a fixed and embedded neuronal tissue is sliced into sections of about 50nm in thickness. Each section is then observed using an Electron Microscope producing

* D.G. and R.B. acknowledge the support and the hospitality of the Janelia visitors program under which this work was carried out.

Y. Boykov et al. (Eds.): EMMCVPR 2011, LNCS 6819, pp. 261–272, 2011.

a 2D projection image of the section with a pixel size of about $10 \times 10\text{nm}^2$. Although the images obtained with this method are of high quality, due to the thick sectioning, membranes crossing the section in oblique directions appear blurry (see examples in Figures 1 and 2). Moreover, the same membrane can appear displaced in consecutive sections, making it difficult to link the regions belonging to the same neuron from one section to the next. High depth-resolution can be obtained by Focused Ion Beam (FIB) [4] and serial section electron tomography [5]. However, because of low throughput these techniques are currently limited to small tissue volumes and therefore cannot be used to reconstruct complete neuronal networks.



Slice thickness: 50 nm Slice thickness: 10 nm

Fig. 1. The figure shows part of a tomographic section produced in simulation from high resolution FIB data (left) and the corresponding middle section in the original high resolution image (right). Notice the blurry membranes in the left image, which appear much sharper in the right image. Such membranes are difficult to detect in the low depth-resolution data. An example of a blurry membrane is marked with a green (vertical) arrow, a non-membrane region with similar appearance is marked with a red (horizontal) arrow. A mitochondrion can be seen at the lower part of the image, it is surrounded by a blue circle.

Once the EM imagery is collected, a crucial step in reconstructing the underlying neuronal circuits is to segment each individual neuron in the 3-dimensional images [3,6]. Segmentation of neurons can be difficult since different neurons usually share similar intensity and texture distributions, requiring one to accurately locate their bounding membranes. As neurons are usually very long, each segmentation mistake can lead to significant mistakes in the topology of the recovered network. Moreover, as current, high throughput EM techniques (such as ssTEM) are limited in their depth resolution, relating the different 2D segments across different sections can be challenging.

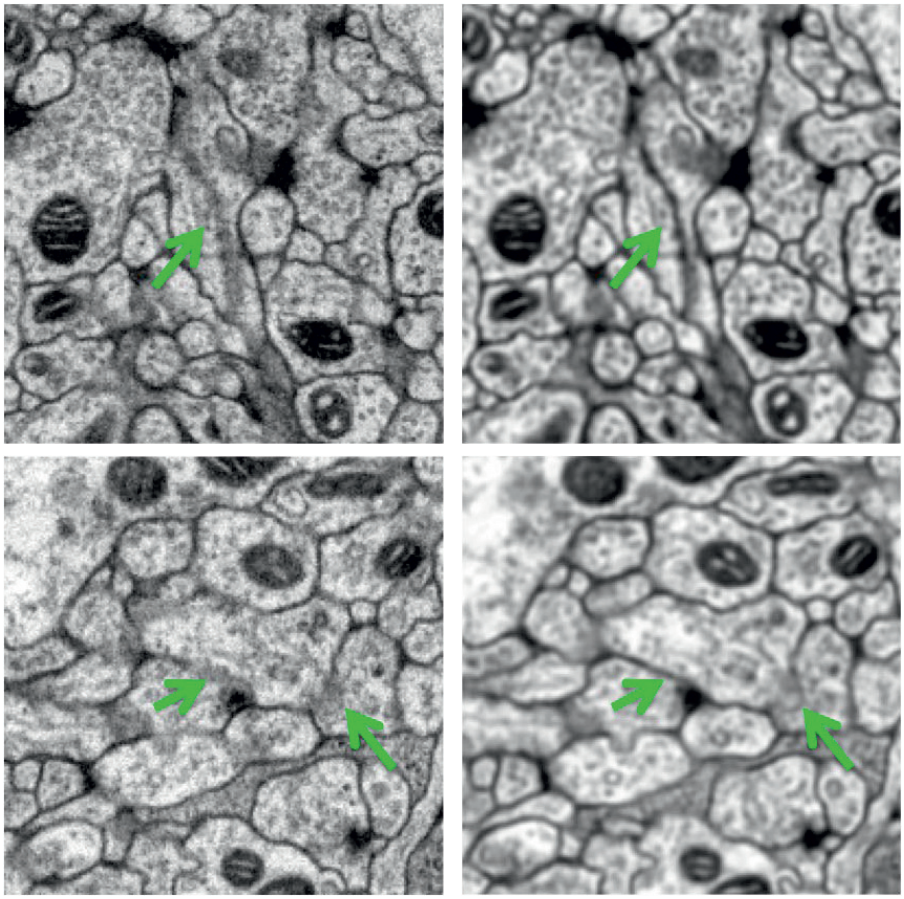


Fig. 2. The figure shows parts of a low depth-resolution serial section Tomographic EM image (left images) and the corresponding middle section in an image reconstructed with super-resolution (right images). Again, notice the blurry membranes in the left images (marked by green arrows), which are somewhat rectified by the super-resolution reconstruction.

A recent approach proposed to improve the depth resolution of ssTEM by using "limited angle tomography" [7]. In this technique, one images each section at only a few angles to maintain the high throughput and uses computational methods to reconstruct the volume structure with high depth resolution. In particular, Veeraraghavan et al. [8] used a sparse representation of the volume using a manually chosen over-complete dictionary [9,10]. [7] used a dictionary learned on high-resolution FIB data. After the volume is reconstructed at higher resolution it can be segmented in 3D.

In this paper we follow up on the work presented in [8] and in [7], and propose to segment the neuronal cells directly from low depth-resolution EM images,

while by-passing the reconstruction step. We employ two existing methods for detecting the location of membranes in high resolution. The first method learns a discriminative, over-complete dictionary to relate between the input tomographic projections and the high resolution class labels. We use the algorithm of [11] (for other methods see, e.g., [12,13]). The second method approaches this classification task using a Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel. We test these approaches on two sets of images. The first set includes low depth-resolution images constructed in simulations from high resolution FIB data of fly larva. The second set of images include low-depth resolution images of fly larva obtained with ssTEM technology. We evaluate our results by comparing classification results to manual labeling by proofreaders on both the high resolution FIB images and on super-resolved ssTEM images. We further compare our results to results obtained with other classification methodologies.

2 Approach

2.1 Learning a Discriminative, over-Complete Dictionary

We cast the problem of segmenting the EM imagery as a classification problem. Our objective is to classify the high resolution voxels as either a membrane or non-membrane given the low depth-resolution input images. Inspired by the success of super-resolution methods [8], which demonstrated that neuronal tissues can be effectively reconstructed by using over-complete dictionary, we first approach this problem using sparse representations over a learned dictionary. Our approach is based on the formulation of Mairal *et al.* [11], adapting their method to the resolution available at test time and the desired resolution of the output. We train a dictionary for the low resolution training data, so that our learned dictionary optimizes a classification loss function over the high resolution labels. At test we apply the learned dictionary to the input low resolution data. Below we describe our approach in more detail.

We begin by defining our training objective. The training data includes a collection of labels for 3D high resolution patches. Let $\mathbf{x} \in \mathbb{R}^p$ denote a vector layout of a high resolution patch, and let $\mathbf{y} \in \{-1, 0, 1\}^p$ contain the label for each of the voxels in $\mathbf{x}$, where the label 1 denotes a membrane, -1 a non-membrane, and 0 is unknown. Let $\mathbf{z} = P\mathbf{x} \in \mathbb{R}^q$ be a vector layout of tomographic projections vectors of $\mathbf{x}$, where the (possibly unknown) $q \times p$ matrix P denotes the projection operator. Our task is to learn an association between the low resolution patches and the high resolution class labels.

We train a classifier by learning an over-complete dictionary, below we review the method of [11] as applied to our problem. Let D denote the sought dictionary. D is a matrix of size $q \times k$ where k is the number of dictionary elements (typically $k = 2q$). Given D we decompose our training patch $\mathbf{z}$ over D by optimizing a functional of the following form,

$$\alpha^\star(\mathbf{z}, D) = \underset{\alpha \in \mathbb{R}^k}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{z} - D\alpha\|_2^2 + \lambda_1\|\alpha\|_1 + \frac{\lambda_2}{2}\|\alpha\|_2^2. \tag{1}$$

The first term in this equation seeks a vector of coefficients $\alpha \in \mathbb{R}^k$ that encode the low resolution patch $\mathbf{z}$ in terms of the dictionary D. The remaining terms use the Elastic Net formulation [14] to regularize α. In particular, the second term is the ℓ_1 norm of α, encouraging a sparse encoding over the dictionary. The final term is the squared ℓ_2 norm of α, which is used to provide stability by convexifying the problem. $\lambda_1, \lambda_2 \geq 0$ are constants. We further constrain the encoding coefficients α to be non-negative.

Given an optimal encoding vector $\alpha^\star$ we define our loss function using the logistic loss. Let

$$L(\alpha^\star, \mathbf{y}, W) = \sum_{i=1}^{p} \log\left(1 + e^{-y_i^T \mathbf{w}_i^T \alpha^*}\right), \tag{2}$$

where $\mathbf{y} = (y_1, \ldots, y_p)^T$ is a vector of the provided labels, and the $k \times p$ matrix $W = (\mathbf{w}_1, \cdots, \mathbf{w}_p)$ is a set of learned classification weights[1]. Our training procedure optimizes this loss function:

$$f(D, W) = \mathbb{E}_{\mathbf{y},\mathbf{x}}\left[L(\alpha^\star, \mathbf{y}, W)\right] + \frac{\nu}{2}\|W\|_F^2, \tag{3}$$

where $\mathbb{E}$ denotes expectation taken over the distribution of $(\mathbf{y}, \mathbf{x})$. The rightmost term is a regularization term; $\|.\|_F$ denotes the Forbenious matrix norm, and ν is a predetermined constant. We seek to optimize $f(D, W)$ over all choices of dictionaries $D \in \mathcal{D}$ and weight matrices $W \in \mathbb{R}^{k \times p}$, where

$$\mathcal{D} = \{D \in \mathbb{R}^{q \times k} \mid \|\mathbf{d}_i\|_2 = 1 \ \forall i \in 1, \ldots, k\}. \tag{4}$$

In practice we only learn the labels of a subset $\tilde{p}$ of the labels, $1 \leq \tilde{p} \leq p$, so W is $k \times \tilde{p}$. Our optimization jointly constructs a dictionary D and a weight matrix W that achieve optimal classification over the training data. Further details of the optimization are provided below.

At test time given a low resolution patch, $\mathbf{z} = P\mathbf{x}$ for some unknown high resolution patch $\mathbf{x}$, our objective is to recover the labels $\mathbf{y}$ that correspond to the high resolution patch $\mathbf{x}$. To this end we use (1) to find an optimal encoding of $\mathbf{z}$ over the trained dictionary D, and then use the obtained coefficients to recover the sought label values by setting $\mathbf{y} = \text{sign}(W^T\alpha)$. In general, we set $\tilde{p}$ so as to obtain overlapping predictions of $\mathbf{y}$ from neighboring patches. We average those predictions for each high resolution voxel to obtain our final classification.

Optimization: The construction of the dictionary D and training weights W is done using stochastic gradient descent. The gradient of our functional $f(D, W)$ (3) can be written as in [11]

$$\nabla_W f(D, W) = \mathbb{E}_{\mathbf{y},\mathbf{x}}\left[\nabla_W L(\alpha^\star, \mathbf{y}, W)\right] + \nu W \tag{5}$$

$$\nabla_D f(D, W) = \mathbb{E}_{\mathbf{y},\mathbf{x}}\left[-D\beta^\star \alpha^{\star T} + (\mathbf{z} - D\alpha^\star)\beta^{\star T}\right], \tag{6}$$

[1] In our implementation we extend $\alpha^\star$ by appending the entry 1 to allow an affine shift in $W^T\alpha^*$.

where $\beta^\star \in \mathbb{R}^k$ is a vector defined as follows. Let Λ denote the set of non-zero coefficients in $\alpha^\star$ then the Λ entries of β are set to

$$\beta^\star_\Lambda = (D^T_\Lambda D_\Lambda + \lambda_2 I)^{-1} \nabla_{\alpha_\Lambda} L(\alpha^\star, \mathbf{y}, W), \tag{7}$$

and the rest of the entries are set to zero.

Stochastic gradient descent proceeds iteratively as follows:

1. Select an i.i.d. patch sample $\mathbf{z} \in \mathbb{R}^q$ from the training set, along with its corresponding labeling $\mathbf{y} \in \mathbb{R}^{\tilde{p}}$.
2. Compute its sparse coding by solving (1) (e.g. using a modified LARS [15]).
3. Compute the active set Λ.
4. Compute $\beta^\star$ (7).
5. Update W and D by subtracting their respective gradients, scaled by a learning rate ρ_t.

We initialize our dictionary by training an unsupervised representation dictionary D_{init}.

2.2 SVM Classifier

As an alternative we train a Support Vector Machine (SVM) classifier with both a linear and a Radial Basis Function (RBF) kernel. For the SVM classifier, let $\mathbf{z}_1, ..., \mathbf{z}_N$ denote the training patches, and let $y_1, ..., y_N$ denotes the k'th label $(1 \leq k \leq p)$ of each $\mathbf{x}_i$ $(1 \leq i \leq N)$. We train a classifier that optimizes the standard hinge loss, written in dual form as

$$\min_{\mathbf{a}} \frac{1}{2} \sum_{i,j} a_i a_j K(\mathbf{z}_i, \mathbf{z}_j) + c \sum_i \max(0, 1 - y_i \sum_j a_j K(\mathbf{z}_i, \mathbf{z}_j)), \tag{8}$$

where $\mathbf{a} \in \mathbb{R}^N$ are the support weights, c is a constant, and $K(.)$ is a kernel function. For the linear SVM $K(\mathbf{z}_i, \mathbf{z}_j) = \mathbf{z}_i^T \mathbf{z}_j$. The RBF kernel is $K(\mathbf{z}_i, \mathbf{z}_j) = e^{-\gamma \|\mathbf{z}_i - \mathbf{z}_j\|}$, where γ is a scale factor. At test time given a low resolution patch $\mathbf{z}$ we assign the labels $y^k = \sum_i a_i^k K(\mathbf{z}_i, \mathbf{z})$.

3 Experiments

To test our approach we have evaluated our method on simulated tomographic projection images constructed from high resolution FIB data. In addition we show results on low depth-resolution ssTEM data.

3.1 Parameter Selection

For the dictionary based method we achieved similar performance using twice and thrice over-complete dictionaries and chose to use twice over-complete in all our experiments. The value of λ_1 was set at $\lambda_1 = 0.03$ for the FIB experiment and $\lambda_1 = 0.05$ for the ssTEM experiment. These values were chosen out of the

set $\{0.15, 0.125, 0.1, 0.075, 0.05, 0.04, 0.03, 0.02\}$ using cross validation. The value of λ_2 was fixed at $\lambda_2 = 0.01$ in all the experiments. In all the experiments we trained using mini batches of size $\eta = 200$ and ran three epochs of $T = 20,000$ iterations each. In each epoch we decreased the learning rate $\rho \in \{0.5, 0.1, 0.01\}$ and ρ_t was then set to $\min(\rho, \rho t_0/t)$, where $t_0 = T/100$. We use the SPAMS toolbox [16,17] to train the unsupervised dictionary and to find an initial W given the unsupervised dictionary. We also use it to solve the lasso during training and testing.

For the SVM with the RBF kernel we used $c = 1$ and $\gamma = 2^{-4}$. These values were chosen using cross validation on a grid of different (c, γ) values. For the linear SVM we used $c = 0.5$ after running cross validation on an extensive set of possible c values.

3.2 Simulations with FIB Data

High resolution 3D images of fly larva were acquired with an Electron Microscope using the Focused Ion Beam (FIB) protocol. The volume included 500^3 voxels of size $10 \times 10 \times 10$ nm^3 each. Proofreaders labeled the thin (1 voxel width) skeletons of the membranes by correcting the results of watershed segmentation. Additional labels were assigned to mitochondria (an example mitochondrion is marked in Figure 1). In the FIB data experiments we ignored the voxels marked as mitochondria. In addition, we ignored voxels of Euclidean distance greater than $\sqrt{2}$ from marked membranes, as those voxels often are membrane voxels, but they are not marked as such by the proofreaders. We used half of the data for training and a disjoint 200^3 part of the volume for testing.

We produced tomographic sections of the volume by averaging each 5 Z-sections in one of 5 directions, parallel the Z-axis and at $\pm 45°$ toward the X and Y directions, see Figure 3. We then selected block patches of size $9 \times 9 \times 15$ (obtaining $p = 1,215$) from the original high resolution volume and used the 5 tomographic projections to produce 2D patches. The parallel tomography sections produced patches of size $9 \times 9 \times 3$. The oblique tomography sections were

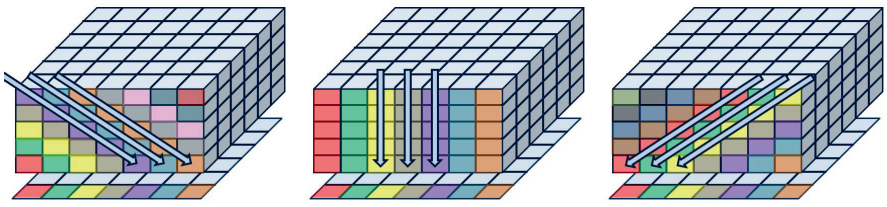


Fig. 3. 2D tomographic sections provide voxel averages in a single direction. This figure illustrates how 2D tomographic sections are produced from a 3D volume in directions $-45°$ (left), 0 (middle), and $+45°$ (right) from vertical in the X-Z plane. In our experiments we used in addition tomographic sections produced in directions $-45°$ and $+45°$ from vertical in the Y-Z plane.

Table 1. Best F-measure achieved by each method on the FIB data

Method	DIC-HR	**DIC-LR**	DIC-SR	**SVM-RBF-LR**	SVM-LIN-LR	LDA-LR
Score	91.28%	**90.07%**	88.68%	**88.26%**	78.16%	76.76%

further intersected with this patch area, producing patches of sizes $5 \times 9 \times 3$ and $9 \times 5 \times 3$. Concatenating these patches we obtained feature vectors of size $q = 783$. With each vector we associate $\tilde{p} = 45$ labels, marking the center $3 \times 3 \times 5$ voxels in the high resolution patches with the proofreaders' labels. Overall the test volume included 662,675 (8.28%) membrane voxels, 5,142,613 (64.28%) non-membrane, and 2,194,712 (27.43%) unknowns.

As a preprocessing step we linearly stretch the values of the input volume after cropping to the range between the 0.001 and 0.999 quantiles of the observed values. We further center each patch by subtracting its mean and scale to have unit ℓ_2 norm. To reduce the dimension of the learning we applied Principal Component Analysis (PCA) to the feature vectors. We chose the number of vectors to account for 95% of the energy in the feature vectors, reducing the dimensionality to 173.

Figure 4 shows a recall-precision plot of our results. These results are also summarized in Table 1, which shows the maximal F-measure (harmonic mean of the recall and precision values) obtained with each method. Our proposed dictionary-based and kernel SVM methods (denoted as DIC-LR and SVM-RBF) achieve F-measures of 90.07% and 88.26% respectively. These values are very close to classification results on the original high resolution data (91.28%, marked by DIC-HR), which can be thought of as a ceiling for our method. We further compare these results to running the dictionary method on super-resolved data (denoted DIC-SR), which achieves an F-measure of 88.68%. This indicates that we can achieve similar or even better classification values if we skip the step of super-resolution reconstruction by classifying the low depth-resolution data directly. Finally, as a base-line we show the results of classifying the membranes using linear SVM (SVM-LIN) and using Linear Discriminant Analysis (LDA).

3.3 ssTEM Data

Low depth-resolution 3D images of fly larva were acquired using the serial section Transmission EM technique. Each section, of width 50nm, was photographed 5 times from roughly the same directions that were simulated with the FIB data (Section 3.2). Each of the obtained 5 volumes included $558 \times 558 \times 16$ voxels of size $10 \times 10 \times 50$ nm^3. To label this data we applied a super resolution

Table 2. Best F-measure achieved by each method on the ssTEM data

Method	SVM-RBF-SR	**DIC-LR**	**SVM-RBF-LR**	SVM-LIN-LR	LDA-LR
Score	88.23%	**87.38%**	**85.75%**	64.28%	64.5%

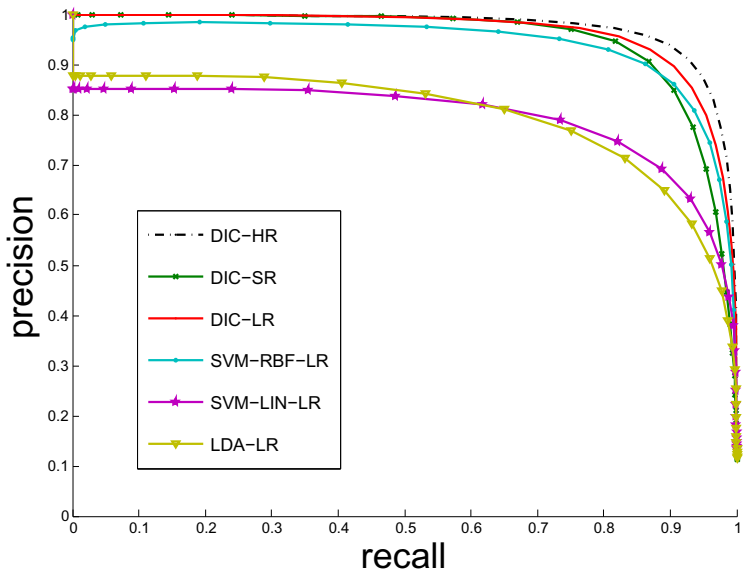


Fig. 4. Results obtained on the FIB data. The figure shows a recall precision plot of our methods, compared to membrane classification on the high resolution and super resolved data.

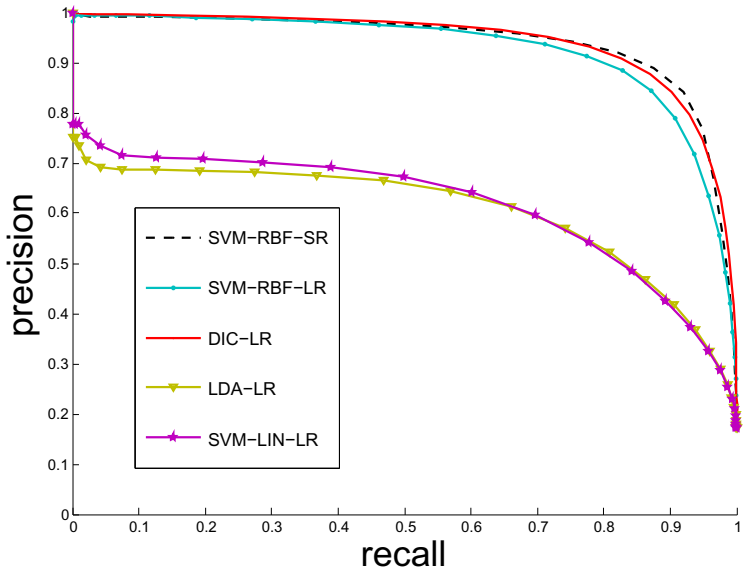


Fig. 5. Results obtained on the ssTEM data. The figure shows a recall precision plot of our methods compared to baseline methods and membrane classification on super resolved data.

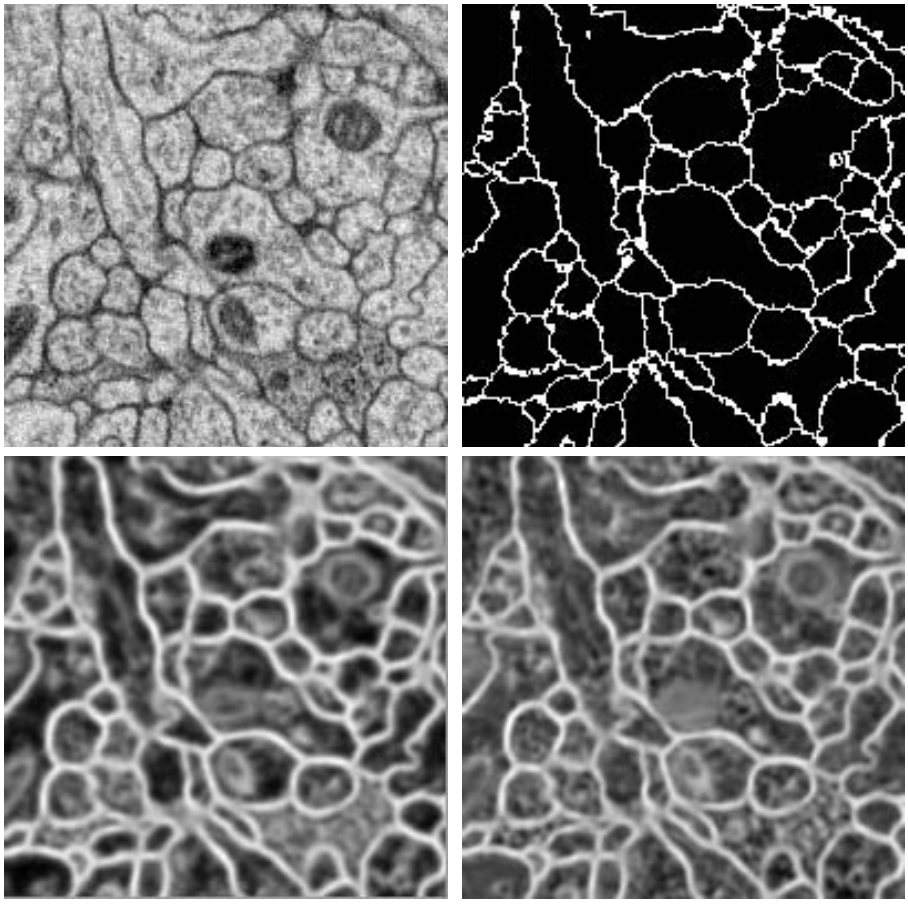


Fig. 6. A classification example. The figure shows part of a section of a ssTEM image (top left), ground truth labeling by a proofreader (top right), and labeling scores obtained with the dictionary-based method (bottom left) and the SVM with RBF kernel (bottom right).

reconstruction using an over-complete dictionary. Proofreaders then labeled the membrane voxels, by marking their skeletons (again, by correcting the results of watershed segmentation). No labeling of mitochondria were available for this data. We used half of the data for training and a disjoint block of size $200 \times 200 \times 65$ for testing.

As before, from the 5 tomographic sections we extracted patches of sizes $9 \times 9 \times 3$ (for the parallel tomographic section) and $5 \times 9 \times 3$ and $9 \times 5 \times 3$ for the other sections, obtaining feature vectors of size $q = 783$. With each vector we associate $\tilde{p} = 45$ labels, marking the center $3 \times 3 \times 5$ voxels in the high resolution patches as either membranes, non-membranes, or unknowns. Overall

the test data included 340,104 (13.08%) membrane voxels, 1,646,992 (63.34%) non-membrane, and 612,974 (23.58%) unknown.

We applied the same preprocessing as for the FIB data (described in Section 3.2). Applying PCA, we reduce the dimension of the feature vectors to 170.

Figure 5 shows a recall-precision plot of our results. The results are also summarized in Table 2, which shows the maximal F-measure obtained with each method. Both our proposed methods (denoted as DIC-LR and SVM-RBF-LR) achieve similar F-measures at 87.38% and 85.75% respectively. These values are slightly lower than the score obtained by running SVM on the super-resolved data (denoted SVM-RBF-SR), which was 88.23%. Note however that the labeling in this experiment is done on the super-resolved data, so it may be biased toward this approach. Finally, as a base line we show the results of classifying the membranes using linear SVM and LDA.

Figure 6 shows an example of the classification scores obtained with the dictionary based and SVM with RBF kernel methods.

4 Conclusion

We presented a system for membrane classification for segmentation of neuronal tissues in low depth-resolution EM imagery. We showed that both a classification method that learns a discriminative, over-complete dictionary as well as SVM with RBF kernel trained over the low depth-resolution EM data with high resolution labeling, can achieve accurate classification of membranes, bypassing the need for an additional step of super-resolution reconstruction. These techniques, therefore, can potentially reduce the amount of manual labor required for reconstructing the topology of the observed cells.

References

1. Varshney, L.R., Chen, B.L., Paniagua, E., Hall, D.H., Chklovskii, D.B.: Structural properties of the Caenorhabditis Elegans neuronal network. PLoS Comput. Biol. 7 (2011)
2. Helmstaedter, M., Briggman, K.L., Denk, W.: 3d structural imaging of the brain with photons and electrons. Curr. Opin. Neurobiol. 18, 633–641 (2008)
3. Chklovskii, D.B., Vitaladevuni, S., Scheffer, L.K.: Semi-automated reconstruction of neural circuits using electron microscopy. Curr. Opin. Neurobiol. 20, 667–675 (2010)
4. Knott, G., Marchman, H., Wall, D., Lich, B.: Serial section scanning electron microscopy of adult brain tissue using focused ion beam milling. J. Neurosci. 28, 2959–2964 (2008)
5. Mcewen, B.F., Marko, M.: The emergence of electron tomography as an important tool for investigating cellular ultrastructure. J. Histochem. Cytochem. 49, 553–563 (2001)
6. Jain, V., Turaga, S.C., Seung, H.S.: Machines that learn to segment images: a crucial technology of connectomics. Curr. Opin. Neurobiol. 20, 653–666 (2010)

7. Hu, T., Nunez-Iglesias, J., Scheffer, L., Xu, S., Hess, H., Fetter, R., Chklovskii, D.B.: Super-resolution reconstruction of brain structure using sparse representation over learned dictionary (2011) (submitted)
8. Veeraraghavan, A., Genkin, A.V., Vitaladevuni, S., Scheffer, L., Xu, S., Hess, H., Fetter, R., Cantoni, M., Knott, G., Chklovskii, D.B.: Increasing depth resolution of electron microscopy of neural circuits using sparse tomographic reconstruction. In: CVPR (2010)
9. Adler, A., Hel-Or, Y., Elad, M.: A shrinkage learning approach for single image super-resolution with overcomplete representations. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010. LNCS, vol. 6312, pp. 622–635. Springer, Heidelberg (2010)
10. Yang, J., Wright, J., Huang, T., Ma, Y.: Image super-resolution via sparse representation. IEEE Trans. on Image Proc. (2011)
11. Mairal, J., Bach, F., Ponce, J.: Task-driven dictionary learning. In: arXiv:1009.5358v1 (2010)
12. Mairal, J., Bach, F., Ponce, J., Sapiro, G., Zisserman, A.: Discriminative learned dictionaries for local image analysis. In: CVPR (2008)
13. Ramirez, I., Sprechmann, P., Sapiro, G.: Classification and clustering via dictionary learning with structured incoherence and shared features. In: CVPR (2010)
14. Zou, H., Hastie, T.: Regularization and variable selection via the elastic net. JRSSB 67, 301–320 (2005)
15. Efron, B., Hastie, T., Johnstone, I., Tibshirani, R.: Least angle regression. Annals of statistics 32, 407–499 (2004)
16. Mairal, J., Bach, F., Ponce, J., Sapiro, G.: Online dictionary learning for sparse coding. In: Proceedings of the 26th Annual International Conference on Machine Learning, pp. 689–696. ACM, New York (2009)
17. Mairal, J., Bach, F., Ponce, J., Sapiro, G.: Online learning for matrix factorization and sparse coding. The Journal of Machine Learning Research 11, 19–60 (2010)