An effective fully deep convolutional neural network for mitochondria segmentation based on ATUM-SEM

Chi Xiao, Weifu Li, Xi Chen, Hua Han, Qiwei Xie
An Effective Fully Deep Convolutional Neural Networks for Mitochondria Segmentation Based on ATUM-SEM

Chi Xiao\textsuperscript{a,b}, Weifu Li\textsuperscript{a,c}, Xi Chen\textsuperscript{a}, Hua Han\textsuperscript{a,b,c}, and Qiwei Xie\textsuperscript{a}

\textsuperscript{a}Chinese Academy of Sciences, Institute of Automation, Beijing, China
\textsuperscript{b}University of Chinese Academy of Sciences, School of Future Technology, Beijing, China
\textsuperscript{c}Hubei University, Faculty of Mathematics and Statistic, Wuhan, China
\textsuperscript{d}The Center for Excellence in Brain Science and Intelligence Technology, CAS, China

ABSTRACT

Recent studies have empowered that the relation between mitochondrial function and degenerative disorders is related to aging diseases. Due to the rapid development of electron microscope (EM), stacks delivered by EM can be used to investigate the link between mitochondrial function and physical structure. Whereas, one of the main challenges in mitochondria research is developing suitable segmentation algorithms to obtain the shapes of mitochondria. Nowadays, Deep Neural Network (DNN) has been widely applied in solving the neuron membrane segmentation problems in virtue of its exceptional performance. For this reason, its application to mitochondria segmentation holds great promise. In this paper, we propose an effective deep learning approach to realize mitochondria segmentation in Automated Tape-Collecting Ultra Microtome Scanning Electron Microscopy (ATUM-SEM) stacks. The proposed algorithm contains three parts: (1) utilizing histogram equalization algorithm as image preprocessing to keep the consistency of dataset; (2) putting forward a fusion fully convolution network (FCN), which is motivated by the principle the deeper, the better, to build a much deeper network architecture for more accurate mitochondria segmentation; and (3) employing fully connected conditional random field (CRF) to optimize segmentation results. Evaluation was performed on a dataset of a stack of 31 slices from ATUM-SEM, with 20 images used for training and 11 images for testing. For comparison, U-Net approach was evaluated through the same dataset. Jaccard index between the automated segmentation and expert manual segmentations indicates that our method (90.1%) outperforms U-Net (87.9%) and has a preferable performance on mitochondria segmentation with different shapes and sizes.

Keywords: Electron Microscope, Deep learning, Mitochondria segmentation

1. INTRODUCTION

Known as the powerhouse of the cell, mitochondria have proven to carry out all types of important cellular functions by producing the overwhelming majority of cellular adenosine triphosphate (ATP). Meanwhile, they also take substantial responsibility in the regulation of cellular life and death, including disease states. For example, mitochondrial dysfunction has been directly linked to the aging process, which is the largest single risk factor for Alzheimer Disease (AD).\textsuperscript{1} Since morphological alterations usually lead to disturbances in the mitochondrial functions and distribution,\textsuperscript{2} many meaningful research studies have focused their investigations on the relationship between the mitochondrial distribution and shapes, and the corresponding functions. Increasing evidence suggests that the mitochondrial distribution inside a cell can be strikingly heterogeneous.\textsuperscript{3} For example, they are often enriched at the cellular sites where the demands for energy are greater, or where their metabolic functions are required, such as at the level of the synaptic button. Equally, recent studies have shown that the regulation of mitochondrial shapes is crucial for cellular physiology, since changes in mitochondrial shapes have been linked to neurodegeneration, calcium signaling, lifespan, and cell death, which further expounds the crucial role that morphological changes of mitochondria play in the immune system.\textsuperscript{4} Furthermore, it has been established that the function of mitochondria is closely related to cancer.\textsuperscript{5} Some specific examples have

Further author information: (Send correspondence to Hua Han and Qiwei Xie)
Hua Han: E-mail: hua.han@ia.ac.cn Qiwei Xie: E-mail: qiwei.xie@ia.ac.cn
demonstrated that the mitochondria in cancer cells can alter the function of resisting apoptosis,\textsuperscript{6,7} which has naturally led the research studies regarding cancer therapy to focus on mitochondria by stimulating mitochondrial membrane permeability, or by changing the mitochondrial metabolism.\textsuperscript{8} All of these show that the statistics and analysis of the mitochondria is an essential endeavor of neurobiology research.

However, due to the variety of mitochondrial structures, as well as the presence of noise, artifacts and other sub-cellular structures, mitochondria segmentation in EM has been proven to be a difficult and challenging task. In recent years, various attempts have been made to quantify the important properties of mitochondria from EM data. For some known results, Lucchi \textit{et al.} considered an automated graph partitioning scheme which was incorporated with shape features on focused ion beam scanning electron microscopy (FIB-SEM) data.\textsuperscript{9} In their study, they first over-segmented the image stack into supervoxels to reduce the computational and memory costs by a simple linear iterative clustering (SLIC), followed by embedding the feature vector consisting of Ray descriptors and intensity histograms into the graph. Due to the fact that the Ray features relied on a good binary edge which could not be easily obtained from noisy EM images, they subsequently improved upon their earlier approaches by exploiting the context-aware features instead of the Ray features.\textsuperscript{10} Jorstad \textit{et al.} took advantage of the fact that mitochondria have thick dark membranes, and proposed an active surface-based method for refining the boundary surfaces of mitochondrial segmentation (on FIB-SEM data).\textsuperscript{11} Although these methods\textsuperscript{9–11} have achieved promising performance, they required the image stacks to be isotropic to work properly. As for anisotropic image stacks, Tasel \textit{et al.} first utilized a parabolic arc model to extract membrane structures, and then employed the curve energy based on active contour to obtain roughly outlined candidate mitochondrial regions, and finally obtained the mitochondrial segmentation by way of a validation process (on Transmission Electron Microscope data).\textsuperscript{12} Marquez \textit{et al.} presented a computationally efficient approach that worked with anisotropic voxels, allowing the segmentation of large image stacks (Serial Block-Face Scanning Electron Microscopy data).\textsuperscript{13} Where the conditional random field inference and surface smoothing techniques were adopted to improve the segmentation and visualization. A recent approach, Li \textit{et al.} used ridge detection to acquire the mitochondrial membrane edges in the variational image segmentation model, and further utilized group-similarity in context to optimize the local misleading segmentation (on ATUM-SEM data).\textsuperscript{14} Yet the performances are not satisfying. To this end, we propose an effective fusion deep network to segment mitochondria in ATUM-SEM stacks for its extraordinary performance.

2. METHODS

The proposed mitochondria segmentation approach could be divided into three parts, which are corresponding image preprocessing, mitochondria segmentation with fusion FCN and image post processing through fully connected CRF.

2.1 Image preprocessing

We fist employ the registration method to register adjacent sections by employing modified Moving Least Square (MLS) deformation algorithm and Scale Invariant Feature Transform (SIFT) features.\textsuperscript{15} This algorithm not only reflects the discontinuity around wrinkle areas but also keeps the smoothness in other regions. However, even if the EM images are obtained by ATUM-SEM from the same tissue, there are still different grayscale and contrast distributions among them. To keep the dataset consistent, histogram equalization is adopted to weaken the noise and enhance the contrast of raw images, which provides a stable foundation for the follow-up works.

2.2 Mitochondria segmentation with fusion FCN

2.2.1 Network architecture

Figure 1 is a pictorial illustration of the proposed network architecture. Inspired by previous studies, we combine Resnet,\textsuperscript{16} Pyramid Scene Parsing network (PSPnet)\textsuperscript{17} with Inception-like net\textsuperscript{18} to deepen the network, and then we utilize a simple skip connection to conserve high-resolution feature information as well as propagate gradients back to the early layers.

Given an input image, we employ a pretrained Resnet50 model to extract feature map. As shown in figure 1, the first convolutional layers (kernel size 3×3 with a stride 2) in residual block 1 and 2 are used for downsampling,
Figure 1. The architecture of the proposed fusion fully convolution network. Red blocks represents regular convolutional layers with kernel size and channels \((64, 7, 7)\) means that kernel size is \(7 \times 7\) and channels are 64). Green blocks annotate maxpooling layers over NN patches with stride S, and blue blocks denote deconvolution layers. Each violet block is residual model that consists of three convolutional layers and a residual skip connection. Yellow block represents PSPnet and brown block indicates softmax layer. The red numbers above arrows imply the size of feature maps, while the numbers below arrows imply the channels of feature maps.

while residual block 3 and 4 adopt dilated network strategy, which does not reduce image size. As a result, the output feature map size of Resnet50 model is only \(1/8\) of the input image. In what follows, the feature map is passed to the PSPnet, details are shown in figure 2. PSPnet provides an effective global contextual prior for pixel-level scene parsing, which incorporates the multi-level contextual information with different receptive fields and improves the accuracy of mitochondrial structure segmentation. The inception-like structure is applied to up-sample and refine the per-pixel prediction. In addition, this structure induces topological sparsity and widens network architecture, thereby reducing computational cost and improving the adaptability of the network to different scales.

Figure 2. Overview of the PSPnet. The input feature map are pooled into 4-level pyramid maps, and then the channel number of these pooled maps are reduced to one-quarter through convolutional layer with \(1 \times 1\) kernel size. Subsequently, deconvolution and concatenation layers are employed to form the final feature map, which contains both local and global information.

2.2.2 Data augmentation
Since different ATUM-SEM images share similar orientation-independent textures generally, simple flip and rotation methods are applied to enrich our data and avoid the overfitting case. In addition, we divide one original ATUM-SEM image (size of \(8624 \times 8416\)) into 276 small images (size of \(512 \times 512\)), which greatly augment our dataset. Through data augmentation, the number of training samples is up to 20,000, which is sufficient for deep learning training.

2.2.3 Experimental setup
The proposed network was implemented using Keras deep learning library. Keras is a high-level neural networks API, written in Python and capable of running on top of TensorFlow or Theano, which is developed for speeding
up experimentation. The training and testing tasks are conducted on a server equipped with an Intel i7 CPU of 512 GB main memory and a Tesla K40 GPU.

2.3 Image post-processing through fully connected CRF

The prediction of fusion FCN is stunning, yet the score maps are too smooth to produce ambiguities results. In this part, fully connected CRF is applied to refine the coarse prediction as well as recover detailed local structure. According to,\(^{19}\) the energy function of fully connected CRF can be described as

\[
E(x) = \sum_i \theta_i(x_i) + \sum_{ij} \theta_{ij}(x_i, x_j). \tag{1}
\]

The minimum of Eq. (1) could be considered as a good segmentation. Here \(x\) is the label assignment for pixels. The unary potential is defined as \(\theta_i(x_i) = -\log P(x_i)\), where \(P(x_i)\) is the label probability at pixel \(i\) computed by Fusion FCN. The pairwise potentials \(\theta_{ij}(x_i, x_j)\) describe the relationship between pixel \(i\) and pixel \(j\), since the pairwise term for each pair exists no matter the distance of them, the models graph is fully connected. In this model, \(\theta_{ij}(x_i, x_j)\) have the form

\[
\theta_{ij}(x_i, x_j) = \mu(x_i, x_j) \sum_{m=1}^K \omega^{(m)} k^{(m)}(f_i, f_j). \tag{2}
\]

Label compatibility function \(\mu(x_i, x_j)\) evaluates to 0 when \(x_i = x_j\) and 1 otherwise. \(k^{(m)}\) is the Gaussian kernel

\[
k^{(m)}(f_i, f_j) = \exp\left(\frac{-(f_i - f_j)^T A^{(m)} (f_i - f_j) / 2}{\omega^{(m)}}\right), \tag{3}
\]

which depends on feature vectors \(f_i, f_j\) exacted from pixels \(i, j\) and is weighted by \(\omega^{(m)}\). Positive-definite precision matrix \(A^{(m)}\) is adopted to define the shape of kernel.

3. EXPERIMENTS AND RESULTS

The biological specimens used in this paper are taken from the cerebral cortex of mouse. The training dataset consists of a stack of 20 slice from ATUM-SEM dataset, which measures around 17×17×1 microns with a resolution of 2×2×50 nm per voxel. The testing dataset with 11 slices is obtained from a different part of the same specimens. We evaluate the segmentation performance of our approach and compare it with U-Net.\(^{20}\) Performance is measured by Jaccard index, which is used for image segmentation commonly.\(^{21}\) It calculates the pixel-wise overlap between ground truth (\(Y\)) and segmented results (\(X\)) as

\[
Jaccard index (X,Y) = \frac{X \cap Y}{X \cup Y}. \tag{4}
\]

Quantitative results are provided in Table 1. Fusion FCN (Our approach) performs better than the baseline of U-Net. The increased reliability due to the use of histogram equalization and fully connect CRF leads to higher scores for both Fusion FCN and U-Net methods.

<table>
<thead>
<tr>
<th></th>
<th>Raw image</th>
<th>Histogram equalization</th>
<th>Raw image + fully connected CRF</th>
<th>Histogram equalization + fully connected CRF</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-Net(^{20})</td>
<td>85.4%</td>
<td>87.4%</td>
<td>86.0%</td>
<td>87.9%</td>
</tr>
<tr>
<td>Fusion FCN (ours)</td>
<td>88.9%</td>
<td>89.1%</td>
<td>90.1%</td>
<td>90.4%</td>
</tr>
</tbody>
</table>

Figure 3 exhibits the results of mitochondria segmentation. As depicted in this figure, we can easily draw the conclusion that our approach achieves satisfactory results, it yields more accurate results as well as reduces false positives, and most of the mitochondria are segmented correctly.
4. CONCLUSION

In this paper, we propose an effective approach of deep network for mitochondria segmentation. We first adopt histogram equalization method to obtain preprocessing stacks. Subsequently, fusion FCN, an suitable extension of Resnet, PSPnet and Inception-net is proposed to develop a deeper network for accurate end-to-end mitochondria segmentation. At last, fully connected CRF is applied as post processing to obtain more refined results. Experimental results demonstrate the effectiveness of our algorithm in mitochondria segmentation, and confirm that our approach outperforms state-of-the-art methods in standard quality metric. In the future, we plan to take z-continuity of mitochondria into consideration to reduce false positives, which is capable to improve the accuracy of segmentation results.

ACKNOWLEDGMENTS

Authors would like to thank Dr. Yu Kong, Yang Yang and Danqian Liu (Institute of Neuroscience, CAS) for sample preparation and sectioning, and also would like to thank Mr. Lixin Wei and co-leagues (Institute of Automation, CAS) for Zeiss Supra55 Scanning Electron Microscope and technical support. This paper is supported by Scientific Instrument Developing Project of Chinese Academy of Sciences (No.YZ201671), National Science Foundation of China (No. 61201050) and Special Program of Beijing Municipal Science and Technology Commission (No.Z161100000216146). This unnumbered section is used to identify those who have aided the authors in understanding or accomplishing the work presented and to acknowledge sources of funding.

References