Received: 13 July 2020 | Accepted: 11 November 2020

DOI: 10.1002/pamm.202000188

Biopolymer segmentation from CLSM microscopy images using a convolutional neural network

Pouyan Asgharzadeh[1,2,*], **Annette I. Birkhold**[1], **Bugra Özdemir**[3], **Ralf Reski**[3,4], and **Oliver Röhrle**[1,2]

[1] Institute for Modelling and Simulation of Biomechanical Systems, University of Stuttgart, Stuttgart, Germany
[2] Stuttgart Center for Simulation Science (SC SimTech), University of Stuttgart, Stuttgart, Germany
[3] Plant Biotechnology, Faculty of Biology, University of Freiburg, Freiburg, Germany
[4] Signalling Research Centres BIOSS and CIBSS, Freiburg, Germany

Confocal microscopy allows visualization of biopolymer networks at the nano scale. Analyzing the structure and assembly of protein networks from images requires a segmentation process. This has proven to be challenging due to multiple possible sources of noise in images as well as exhibition of out-of-focus planes. Here, we present a deep learning-based segmentation procedure for confocal laser scanning microscopy images of biopolymer networks. Utilizing an encoder-decoder network architecture, our deep neural network achieved a dice score of 0.88 in segmenting images of filamentous temperature sensitive Z proteins from chloroplasts of *Physcomitrella patens*, a moss.

1 Introduction

State-of-the-art fluorescence imaging enables investigating details of complex subcellular biopolymer structures such as cytoskeletal protein networks. The high-resolution 3D images from advanced microscopic imaging techniques require segmentation tools for extracting the complex network structure. This segmentation permits quantitative assessment of the protein network structures [1, 2]. Moreover, segmentation is a fundamental step in creating computer models investigating geometrical and mechanical aspects of protein networks [3]. Segmentation of live CLSM images has proven to be challenging due to exhibiting fluorescent overexpression, out-of-focus imaging planes and dynamics of the protein structure captured during live-cell microscopy [4]. Although advanced methods, such as active contour methods, achieved higher accuracy than the classical thresholding algorithms [5], there still exists a substantial gap between outcome of such algorithms and manual expert segmentation. Recently, convolutional neural networks (CNN) have shown success in automatic analysis of high-resolution image data [6]. Specifically, encoder-decoder networks, e.g. UNet, achieve high accuracy in segmentation of biomedical images [7]. This motivates the use of such an approach for the complex task of CLSM image segmentation. Here, we present an automated segmentation process utilizing an encoder-decoder CNN applied to CLSM images of filamentous temperature sensitive Z (FtsZ) protein networks.

2 Materials and Methods

2.1 Confocal microscopy imaging

FtsZ coding sequence, tagged with the coding sequence of Enhanced Green Fluorescent Protein (EGFP), was expressed in moss protoplasts. A total of 40 3D CLSM images of FtsZ protein were acquired (Leica TCS SP8 microscope, Leica Microsystems, Wetzlar, Germany). The images were $0.021\,\mu m$ in x, y- and $0.240\,\mu m$ in z-dimension. Deconvolution was performed using Huygens Professional version 17.04 (Scientific Volume Imaging, Hilversum, Netherlands).

2.2 Preprocessing and segmentation with UNet

The ground truth segmented images are created in a semi-automatic manner. First, images are segmented using an adaptive local thresholding algorithm [1]. Afterwards, the segmented images are modified by two experts in an iterative process until both experts agree upon the segmented labels. Due to the differences of resolution in x-, y- and z-directions, 3D images are transformed into 2D slices parallel to the x-y plane. This increases the size of the training dataset and reduces the computational costs for the training. Then, from the dataset of 2D slices of $n = 40$ 3D image, 85% were randomly selected to form the training data. The remaining form the test set. To homogenize the image sizes, all images were resized to 256×256 pixels. To avoid overfitting, the training dataset size is synthetically increased utilizing standard translational augmentation operations resulted in a training dataset of 15015 2D images. A CNN based on a 2D UNet encoder-decoder network is trained to map the raw image data to the segmented labels (CNN architecture shown in Fig. 1a). Training is carried out for 10 epochs on the training set utilizing a 5-fold cross-validation scheme. At the end of each training epoch, the network is applied on all images of the test set. Networks were trained with the Adam optimization algorithm for minimizing the loss function

* Corresponding author: e-mail asgharzadeh@imsb.uni-stuttgart.de, phone +49 711 685-69254

of binary cross entropy. The accuracy of segmentation was measured by comparing the true segmented label with model prediction using an Intersection-over-Union (IoU) coefficient factor. IoU is the area of overlap between the predicted label and the ground truth divided by the area of union between them. IoU value lies between 0 and 1, with 1 representing the ideal segmentation process. The network was implemented in Keras using Google Tensorflow 1.7 and was trained on a computer with a single Nvidia GTX 1070 GPU. After applying the trained network on the 2D slices belonging to a 3D CLSM image, the slices are put back together to create the segmented 3D image.

3 Results

During training, the segmentation IoU coefficient constantly increased for both training and test sets (Fig. 1b). IoU caps at 0.88 on the test set after 8 epochs and no considerable decrease in loss value after this epoch was gained. Therefore, at this point, the training was stopped to prevent the possibility of overfitting. At the end of the eighth epoch, the loss value for test set wass equal to 0.03 (Fig. 1c).

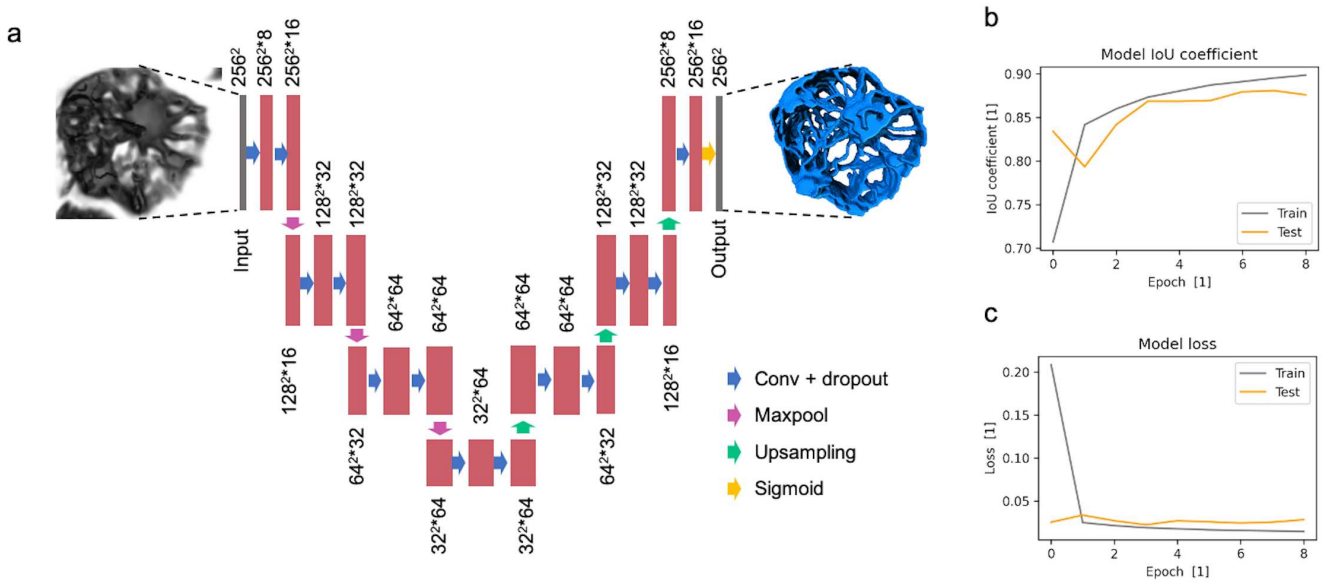

Fig. 1: a) The encoder-decoder CNN to map the raw image data to the segmented labels. b) The IoU coefficient during training vs the training epoch number. c) The cross-entropy loss value vs the training epoch number. The values for applying the model on training and test sets are shown in gray and orange colors, respectively.

4 Discussion

An accuracy comparable to recent active contour segmentation method was reached [5]. For a further improvement in segmentation performance, future studies need to consider an increased training size. Training on 2D slices results in neglecting the geometrical information encoded in z-direction. Therefore, utilizing a 3D CNN directly on 3D images can potentially increase the segmentation performance. However, that leads to a substantial increase in required computational resources.

Acknowledgements This work was funded by the German Research Foundation (DFG) as a part of Transregio/SFB TRR141, project A09.
Open access funding enabled and organized by Projekt DEAL.

References

[1] P. Asgharzadeh, B. Özdemir, R. Reski, O. Röhrle, and A. I. Birkhold, Acta Biomaterialia **69**, 206–217 (2018).
[2] B. Özdemir, P. Asgharzadeh, A. I. Birkhold, S. J. Mueller, O. Röhrle, and R. Reski, Scientific Reports **8**(1), 1–17 (2018).
[3] P. Asgharzadeh, A. I. Birkhold, Z. Trivedi, B. Özdemir, R. Reski, and O. Röhrle, bioRxiv (2020).
[4] V. Westphal and S. W. Hell, Physical Review Letters **94**(14), 143903 (2005).
[5] M. B. Smith, H. Li, T. Shen, X. Huang, E. Yusuf, and D. Vavylonis, Cytoskeleton **67**(11), 693–705 (2010).
[6] P. Asgharzadeh, O. Röhrle, B. M. Willie, and A. I. Birkhold, Acta Biomaterialia **106**, 193–207 (2020).
[7] O. Ronneberger, P. Fischer, and T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, (2015), pp. 234–241.

