

Automatic Segmentation of the Golgi Apparatus in Volumetric Data with Approximate Labels

Eva Boneš and Matija Marolt

Abstract—The Golgi apparatus (GA) is a cellular organelle involved in the processing and sorting of proteins in eukaryotic cells. Due to its numerous functions, structural complexity, and organizational dynamics, the role of the GA in normal and pathological processes is still under intensive research. In this work, we present an approach to automatic segmentation of the GA in electron microscopy volumetric data, consisting of i) a neural network trained on approximately labelled data, ii) active contours for refining the segmentation, and iii) filtering of the segmented regions. Evaluation on 3D volumes of a urinary bladder epithelial cell shows that the proposed algorithm is able to segment the GA with 89% sensitivity and 99% specificity. Using approximate labels reduced the time needed for manual annotation of the ground truth by a factor of five.

Keywords—Golgi apparatus, Segmentation, Convolutional neural network, Active contours.

I. INTRODUCTION

Eukaryotic cells build all multi-cellular organisms. In order to understand their function, we need a good understanding of the structure and function of the many cellular organelles inside them, including their spatial distribution, shapes, and sizes. The understanding of organelles and their relations will contribute to a better understanding of the functioning of cells and thus to new discoveries in biology and medicine.

In our research, we focus on the umbrella cells that make up the urothelium - the specialized epithelium tissue that covers the bladder. These cells contain fusiform vesicles (FV), which contain, store and transmit urothelial plaques. Plaques are an important factor in establishing the blood-urine barrier of the mammalian bladder. It is still not entirely clear how FVs are formed, but it has been proven that the Golgi apparatus (GA) is involved in the process.

The GA is an organelle of the biosynthetic pathway, where membrane proteins are functionalized and classified. At the same time, it represents the intersection of the exo- and endocytic pathways in the cell. Knowing its spatial distribution, size, and shape within cells can help to better understand the basic cell processes and observe the effects of various pathological conditions on the cell.

The development of new technologies for capturing three-dimensional volumetric microscopic images using the focused ion beam and scanning electron microscopy (FIB-SEM) has enabled direct observation of cellular organelles and expanded the possibilities for understanding their function within cells.

Both authors are with the Faculty of Computer and Information Science, University of Ljubljana, Večna pot 113, 1000 Ljubljana, Slovenia (e-mails: eb1690@student.uni-lj.si, matija.marolt@fri.uni-lj.si).

To establish the positions of individual organelles in three-dimensional volumetric data, segmentation of such volumes is needed. Manual segmentation can be performed by skilled experts that annotate the cellular organelles in individual 2D volume slices, however, due to the large sizes of the volumes (e.g. $1280 \times 1024 \times 1024$ for part of a cell), it is extremely time-consuming. Automatic segmentation approaches are therefore highly desired.

As in all areas of image analysis, most state-of-the-art approaches for biomedical image segmentation are based on deep convolutional neural networks. The first modern segmentation methods, such as the FCN [1] and the U-Net [2], were created to segment 2D data. With the increasing availability of volumetric data, specialized approaches that exploit the information present in 3D volumetric data were soon developed. The 3D U-Net, as an upgrade of the U-Net [3], was developed for segmenting kidney images, the V-Net [4] for prostate segmentation and DeepMedic [5] for segmentation of brain lesions.

Although these architectures are general, they were all developed for specific domains and imaging techniques and may not generalize well to different types of data. As it is important that the characteristics of the target data are considered when developing segmentation methods [6], we chose to use the HighRes3DZMNet [7] architecture in our work. The architecture was recently introduced for segmentation of mitochondria and endolysosomes in volumetric electron microscopy data. As an upgrade of the HighRes3DNet [8], this architecture adds the use of zero-mean convolutional filters at the first levels of the convolutional neural network to equalize the brightness of the input data. Such preprocessing is especially important in microscopic images where uneven lighting artefacts are often present.

In this paper, we present an approach for automatic segmentation of the GA in volumetric electron microscopy data. Although many approaches were introduced for segmentation of different cellular organelles (e.g. mitochondria and endolysosomes [7], endoplasmic reticulum [9], nuclei [10]), no approach, to our knowledge, currently exists for automatic segmentation of the GA in volumetric data. Our approach is based on a deep neural network segmentation model, trained with approximately labelled data, which speeds up the acquisition of training data significantly. We present the approach in Section II and its evaluation in Section III.

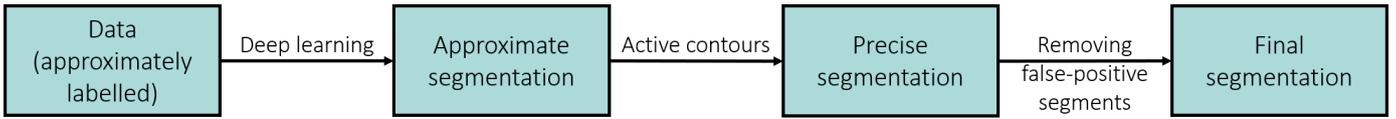


Fig. 1: Our proposed approach to automatic segmentation of the GA.

II. SEGMENTATION

Deep architectures usually need a large amount of annotated data to train properly. Due to the large variety of the GA within the cell and due to their complex shape, their precise manual annotation in volumetric data is extremely time-consuming and very error-prone.

We propose a three-step approach to GA segmentation. First, we train the HighRes3DZMNet deep architecture in a controlled way on approximately labelled data, which yields a model that is able to roughly locate the areas where the GA is located. We then use the active contour method [11] to refine the obtained areas and obtain more precise GA shapes. Finally, we filter the obtained GA segments, to remove false positives. The approach is fully automatic and is outlined in Fig. 1.

We describe the individual steps in the following subsections and visualize their outputs in Fig. 4.

A. Data

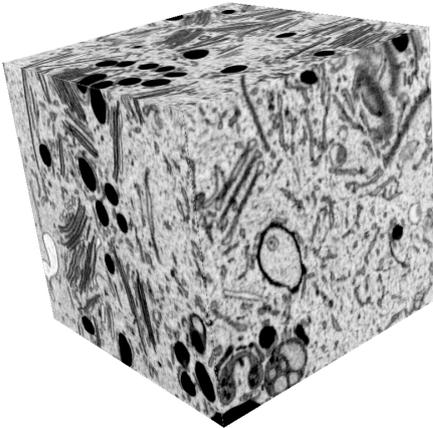


Fig. 2: One of the subvolumes used for testing.

We used the UroCell [7] dataset to train and evaluate our approach. We selected this dataset because it is to our knowledge the only public dataset obtained from urothelial tissue which is the tissue we are interested in. The dataset contains 3D volumes of umbrella cells from the mouse bladder epithelium, obtained with a FIB-SEM electron microscope. The entire volume is divided into subvolumes of size $256 \times 256 \times 256$ voxels, which, given the voxel dimensions of approx. $16 \times 16 \times 15$ nm, yields subvolumes that cover about $4 \times 4 \times 4$ μ m. In this work, subvolumes from different parts of the cell were selected so that the contrast, brightness, content, and artefacts are as diverse as possible. Fig. 2 shows one of the subvolumes.

We manually annotated the GA in 9 subvolumes with the open-source tool Slicer3D [12]. As can be seen in Fig. 3, the shape of the GA is complex, so we only approximately annotated the area they occupy. All the approximate labels were reviewed by an expert in cell biology. With this paper, we are adding GAs to the UroCell dataset publicly available on the following link: <https://github.com/MancaZerovnikMekuc/UroCell>.

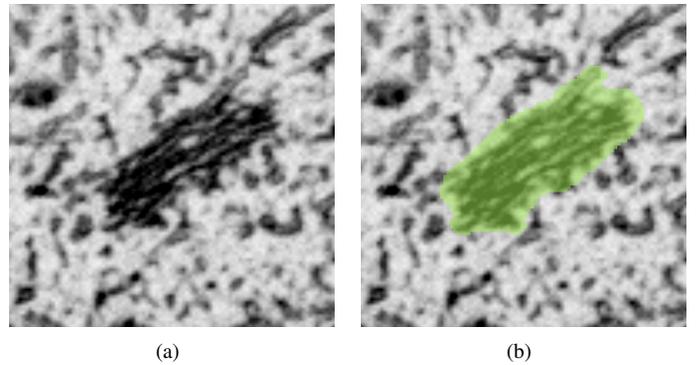


Fig. 3: (a) A volume slice with a GA and (b) the approximately labelled GA area.

B. Deep Learning

To perform the initial approximate segmentation, we used the HighRes3DZMNet deep neural network architecture, which was developed for segmentation of mitochondria and endolysosomes in electron microscopy data (also for the UroCell dataset). The HighRes3DZMNet [7] is a convolutional neural network consisting of 20 layers, which outputs the target class distribution of each individual input voxel.

To train the network, we used 9 subvolumes of the UroCell dataset where we approximately labelled the GA positions. We divided the 9 subvolumes into a training (7) and testing (2) set. Each of the subvolumes contained 0 to 3 GA instances. We used the same parameters and data augmentation techniques for training as in the original HighRes3DZMNet paper, so despite the relatively small dataset, the network generalized well to the unknown data.

C. Active Contours

Active contours [11] are often used in the processing of medical images, mainly for segmentation and demarcation of objects. According to the initial estimate of the object boundary, the algorithm iteratively adjusts this boundary, minimizing the energy function and thus approaching the actual boundary.

We use active contours to refine the approximate GA segmentation produced by the HighRes3DZMNet network.

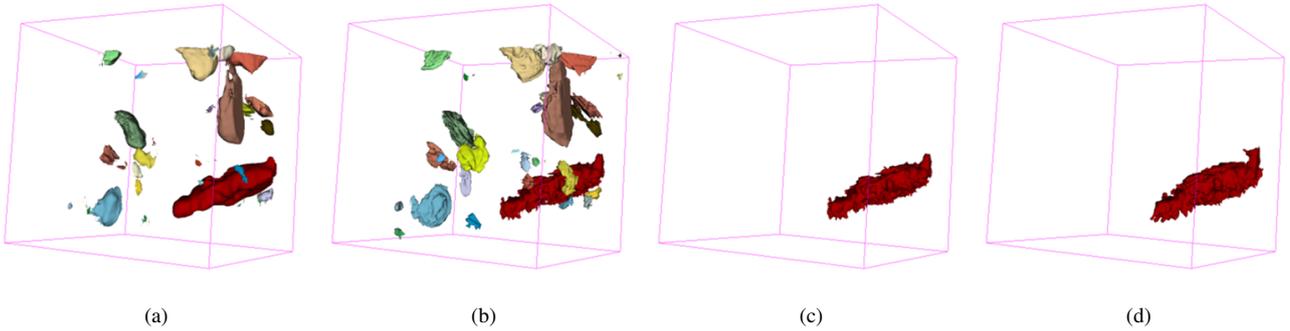


Fig. 4: GA segmentation on a single subvolume: (a) segmentation with the deep model, (b) after active contours, (c) after filtering, (d) the manual ground truth annotation.

Because active contours work on 2D data, we use the method on each of the volume slices separately and across all three dimensions. The HighRes3DZMNet segmentation is used as the initial estimate of the object boundary. We fuse the results of active contours across all the dimensions by assigning a voxel to a GA if it was assigned to the GA in at least one of the dimensions.

The approach has proven to be well suited to our problem. As can be seen in Fig. 5, the approach also works well when part of the initial border is inside the target GA. By minimizing the energy, the border expands outwards at that point.

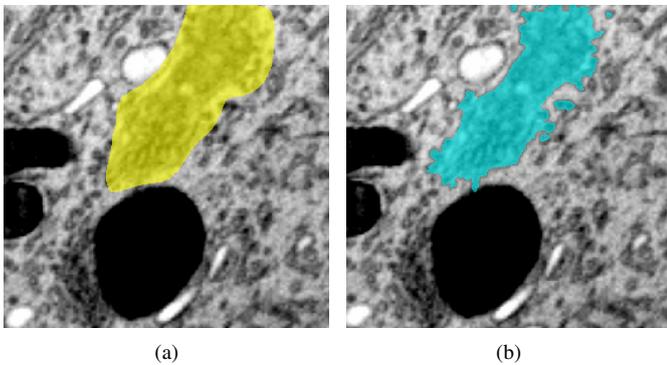


Fig. 5: GA segmentation (a) before and (b) after the use of active contours. On the lower left edge of the yellow annotation, we can see the initial contour inside the GA that the active contours method then expands to the border of the GA seen on the blue annotation.

D. Removing False-positive Segments

In the final step of our approach, we remove excessive (false positive) segments. Most of these occur because stacks of FVs are mistaken for a GA. FVs are specific to the urothelial cells and look very similar in texture to the GA, as can be seen in Fig. 6[a]. Consequently, the deep neural network produces some false-positive segments, where it annotates these stacks as the GA. By analyzing these erroneous segments, we discovered that such errors can be largely eliminated by filtering the segments by size, as the stacks of FVs are usually much smaller than the GA. Thus, we remove all segmented

compartments smaller in volume than $0.19 \mu\text{m}^3$, which is 80% of the size of the smallest GA in our dataset.

III. RESULTS

A. Data

We evaluated our approach on the UroCell dataset, which was divided into the training and test sets as previously described. Although we are aware that the dataset is small and our results may therefore be somewhat optimistic, we are currently not aware of other datasets that we could use to evaluate our approach. Additionally, we must stress that the problem of segmentation in the selected cell type is particularly complex, as the FVs in the cytoplasm of the differentiated urothelial cells are organized into stacks that are very similar in texture to the GA. Therefore, even though we evaluated the approach only on this cell type, we expect that the proposed procedure would be sufficiently general and useful for segmentation of the GA in other cell types as well.

B. Evaluation Metrics

For evaluation, we used sensitivity, specificity, and the Dice similarity coefficient, which are the most commonly used measures for evaluating segmentation accuracy in biomedical images.

If we define TP as the number of true positive voxels (correctly labelled as GA), TN as the number of true negative voxels (correctly labelled as background), FP as the number of false-positive voxels (background voxels marked as GA), and FN as the number of false-negative voxels (GA marked as background), we can define the selected performance measures as follows.

The Dice similarity coefficient (DSC):

$$DSC = \frac{2TP}{2TP + FP + FN} \quad (1)$$

The true positive rate - sensitivity (TPR):

$$TPR = \frac{TP}{TP + FN} \quad (2)$$

The true negative rate - specificity (TNR):

$$TNR = \frac{TN}{TN + FP} \quad (3)$$

We also measured the time spent on approximate and precise manual volume annotation.

C. Results

The results that are shown in Table I were obtained on the two subvolumes of the test set. For evaluation purposes, both subvolumes were manually annotated with a precise GA shape.

The results in Table I represent the average score over both test subvolumes. Since we are not aware of other methods for automatic segmentation of the GA in volumetric data, we could not compare our approach with others.

TABLE I:
RESULTS AT DIFFERENT STAGES OF THE PROPOSED METHOD

Method stage	DSC	TPR	TNR
After deep learning	0.515	0.861	0.983
After active contours	0.501	0.876	0.982
Final segmentation	0.926	0.885	0.999

We also estimated the time needed for approximate vs. precise manual annotation of subvolumes. The time needed to precisely annotate an average-sized GA was 2.5 hours, while approximate manual annotation of the same GA took 0.5 hours - a five-fold speed increase.

D. Discussion

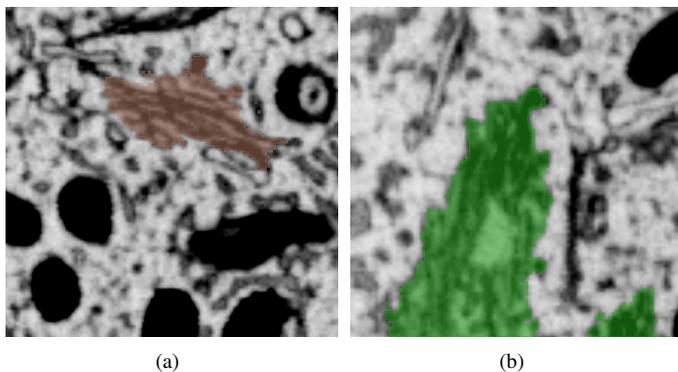


Fig. 6: Errors: (a) FVs annotated as a GA, (b) empty space in the middle annotated as a GA.

The accuracy of our approach after all three steps is high. The DSC is 0.92, the TNR is also very high, only the TPR is slightly lower.

Although we used filtering to remove false-positive segments, several stacks of FVs, labelled as the GA, still remained, as shown in Fig. 6[a].

Empty spaces in the middle of the GA also contribute to errors, as seen in Fig. 6[b]. These were manually annotated as background but were labelled as GA by the deep model. We do not consider this to be a big problem, as the external boundaries of the GA already give us a lot of information, such as the number of the GA bodies in the cell, their spatial distribution, size, and the relationships between them. The

internal structure of the GA is not our focus at the moment, but this is one of the aspects we want to address in the future.

The accuracy of the approach before the filtering step is much lower due to the fact that the stacks of FVs are very similar to the GA and are thus incorrectly labelled. Interestingly, however, we also see a decrease in performance after using the active contour method. This is mainly due to the incorrectly segmented stacks of FVs, where only one part of the stack is annotated as a GA, i.e. the segmentation edge in some parts takes place within the stack. This causes edge propagation with active contours and thus an increase in the number of incorrectly segmented voxels. We can see an example of this in Fig. 7.

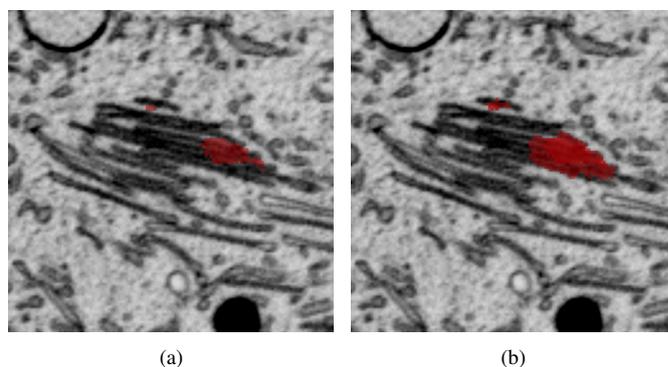


Fig. 7: An incorrectly segmented stack of FVs (a) before and (b) after the use of active contours.

Even though we could not compare our results to other approaches, the qualitative analysis shows that the approach works well. Fig. 4 shows the ground truth annotation and the results of the algorithm on one of the test subvolumes. On the test set, the proposed method found all but one GA; the missed GA was at the edge of the subvolume, so only a part of it was included and it was consequently filtered out due to its size. Such errors would not occur when segmenting the entire cell volume.

To test how the method performs on a larger dataset, we ran it on the entire UroCell volume ($1280 \times 1024 \times 1024$ voxels), which is the size of around 80 such subvolumes. We can not quantify these results, as we are lacking the ground truth annotations, however, qualitative inspection showed that the approach works similarly well regardless of the size of the volume.

The advantage of the proposed approach is also that only approximate annotations are needed for training the deep learning model. With an average-sized GA, we spent five times less time with approximate annotations than we would with precise annotations, so we estimate that in total 50 hours of laborious annotating work were saved for producing the entire training and test sets.

IV. CONCLUSION

With the development of technology for capturing increasingly more accurate microscopic data, there is a growing need

for approaches to analyze them. In this paper, we presented an approach for automatic segmentation of the GA in volumetric electron microscopy data, which, to our knowledge, is the first such approach for this type of data. In addition to successful segmentation, our approach reduces the time required for manual annotation of the training and test data by using approximate labels. The proposed approach is general and could also be used for other similar segmentation problems, where objects exhibit complex shapes.

In our future work, we will improve the filtering stage by including additional parameters, such as the shape of the segmented regions, to better distinguish the GA from FV stacks. We will also continue to improve the segmentation itself with larger training datasets and test the approach on other databases when they become available.

With our research, we are enabling scientists to more quickly and accurately analyze 3D cellular data and thus further their research of the Golgi apparatus and other cellular structures, which is crucial for an in-depth understanding of how the cell functions and how outside factors affect its functions. This, in turn, will further a better understanding of the human body and its physiology, which can lead to new medical discoveries.

ACKNOWLEDGMENT

For all the help, support and encouragement during this research, I would like to thank Manca Žerovnik Mekuč from the Laboratory for Computer Graphics and Multimedia at the Faculty of Computer and Information Science and Samo Hudoklin from the Institute of Cell Biology at the Faculty of Medicine in University of Ljubljana.

REFERENCES

- [1] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, pp. 1–1, 2016.
- [2] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *LNCS*, vol. 9351, pp. 234–241, 2015.
- [3] Çiçek, A. Abdulkadir, S. Lienkamp, T. Brox, and O. Ronneberger, "3d u-net: Learning dense volumetric segmentation from sparse annotation," pp. 424–432, 2016.
- [4] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," pp. 565–571, 2016.
- [5] K. Kamnitsas, C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker, "Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation," *Medical Image Analysis*, vol. 36, pp. 61–78, 2017.
- [6] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. [van der Laak], B. [van Ginneken], and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
- [7] M. Žerovnik Mekuč, C. Bohak, S. Hudoklin, B. H. Kim, R. Romih, M. Y. Kim, and M. Marolt, "Automatic segmentation of mitochondria and endolysosomes in volumetric electron microscopy data," *Computers in Biology and Medicine*, vol. 119, p. 103693, 2020.
- [8] W. Li, G. Wang, L. Fidon, S. Ourselin, M. J. Cardoso, and T. Vercauteren, "On the compactness, efficiency, and representation of 3d convolutional networks: Brain parcellation as a pretext task," pp. 348–360, 2017.
- [9] J. Liu, L. Li, Y. Yang, B. Hong, X. Chen, Q. Xie, and H. Han, "Automatic reconstruction of mitochondria and endoplasmic reticulum in electron microscopy volumes by deep learning," *Frontiers in Neuroscience*, vol. 14, p. 599, 2020.
- [10] S. Wienert, D. Heim, K. Saeger, A. Stenzinger, M. Beil, P. Hufnagl, M. Dietel, C. Denkert, and F. Klauschen, "Detection and segmentation of cell nuclei in virtual microscopy images: A minimum-model approach," *Scientific reports*, vol. 2, p. 503, 07 2012.
- [11] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
- [12] R. Kikinis, S. Pieper, and K. Vosburgh, *3D Slicer: A Platform for Subject-Specific Image Analysis, Visualization, and Clinical Support*, 2014, vol. 3.