

Joint Segmentation and Pairing of Nuclei and Golgi in 3D Microscopy Images

Hemaxi Narotamo^{1,2,*}, Marie Ouarné², Cláudio Areias Franco^{2,+}, Margarida Silveira^{1,+}

Abstract—Blood vessels provide oxygen and nutrients to all tissues in the human body, and their incorrect organisation or dysfunction contributes to several diseases. Correct organisation of blood vessels is achieved through vascular patterning, a process that relies on endothelial cell polarization and migration against the blood flow direction. Unravelling the mechanisms governing endothelial cell polarity is essential to study the process of vascular patterning. Cell polarity is defined by a vector that goes from the nucleus centroid to the corresponding Golgi complex centroid, here defined as axial polarity. Currently, axial polarity is calculated manually, which is time-consuming and subjective. In this work, we used a deep learning approach to segment nuclei and Golgi in 3D fluorescence microscopy images of mouse retinas, and to assign nucleus-Golgi pairs. This approach predicts nuclei and Golgi segmentation masks but also a third mask corresponding to joint nuclei and Golgi segmentations. The joint segmentation mask is used to perform nucleus-Golgi pairing. We demonstrate that our deep learning approach using three masks successfully identifies nucleus-Golgi pairs, outperforming a pairing method based on a cost matrix. Our results pave the way for automated computation of axial polarity in 3D tissues and *in vivo*.

I. INTRODUCTION

Blood vessels nourish tissues with oxygen and nutrients, remove carbon dioxide and waste products away from the tissues, and supply gateways for immune surveillance. Angiogenesis is the formation of new blood vessels from pre-existing ones. It requires two distinct and successive phenomena: sprouting angiogenesis and vascular patterning. On the one hand, during sprouting angiogenesis the vascular network is expanded through migration and proliferation of endothelial cells (ECs), which are the cells lining the interior of blood vessels. The molecular mechanisms involved in sprouting angiogenesis are well established [1], [2]. On the other hand, vascular patterning allows the transformation of the immature vascular network into a functionally hierarchical and perfused network of blood vessels, and it is regulated by blood flow. Cell migration plays an essential role during both processes. During sprouting, ECs polarize and migrate

towards sources of vascular endothelial growth factor A (VEGFA) and are collectively coordinated through tension at adherens junctions [3]. In addition, during patterning ECs polarize and migrate against the flow direction [4]. Thus, unravelling the mechanisms governing EC front-rear polarity is of great importance. Front-rear (axial) polarity of each cell can be computed as a vector that goes from the centroid of the nucleus to the centroid of the corresponding Golgi complex. These vectors are calculated manually, which is time consuming and subjected to human error. Recently, a genetic mouse model (GNrep) was generated to visualize EC axial polarity, and it was proposed to be compatible with automated segmentation and assignment of nuclei and Golgi complexes in mouse tissues [5]. In the present work we present an approach, based on deep learning, that performs joint segmentation and pairing of nuclei and Golgi in microscopy images to later estimate EC polarity.

This paper is organized as follows: in section II a review of prior work on cell organelles segmentation in 3D microscopy images using deep learning and the current methods for nucleus-Golgi pairing are described. In section III the dataset, the approaches for cell nuclei and Golgi segmentation and pairing are described. In section IV the experimental results and discussion are shown. Finally, in section V conclusions and topics for future research are presented.

II. RELATED WORK

Most of the existing deep learning based approaches for cell organelles segmentation in 3D microscopy images use convolutional neural networks (CNNs) or U-Net ([6], [7], [8], [9], [10]). In 2017, Ho et al. [6] proposed a 3D CNN to segment nuclei in fluorescence microscopy images. The CNN was trained with synthetic data and its performance was evaluated on real microscopy data achieving an accuracy of 92.93%. Fu et al. (2017) [7] performed the segmentation of nuclei in 3D microscopy images using a 2D CNN applied to planes extracted along the x, y and z directions of the original image. As a post-processing step, a 3D watershed algorithm was applied to split overlapping nuclei. Although this approach achieves good results (accuracy of 94.23%), a 2D CNN does not fully incorporate the 3D information, thus it can have discontinuities between planes. Later, in 2018, Fu et al. [8] used a 3D U-Net to segment nuclei in fluorescence microscopy images. The U-Net was trained using synthetic data and its performance evaluated on real microscopy images, showing that it can accurately segment the nuclei. Ho et al. [9] (2018) proposed a methodology for 3D nuclei instance segmentation in fluorescence microscopy images.

H.N was supported by the Fundação para a Ciência e Tecnologia (FCT) Doctoral Grant 2020.04511.BD. H.N and M.S were supported by LARSyS - FCT Project UIDB/50009/2020. M.O was supported by EU Horizon 2020 Marie Skłodowska-Curie fellowship (842498). C.A.F was supported by European Research Council starting grant (679368), a grant from the Fondation LeDucq (17CVD03), the FCT funding (grants: PTDC/MED-PAT/31639/2017; PTDC/BIA-CEL/32180/2017; CEECIND/02589/2018).

¹Institute for Systems and Robotics (ISR/IST), LARSyS, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
hemaxi.narotamo@tecnico.ulisboa.pt,
msilveira@isr.tecnico.ulisboa.pt

²Instituto de Medicina Molecular - João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal
{marie.ouarne,cfranco}@medicina.ulisboa.pt

⁺Co-last authors. ^{*}Co-corresponding author.

First, nuclei detection is performed to define potential centers of nuclei using a 3D CNN. Afterwards, for each nuclei seed, a sub-volume of size $16 \times 16 \times 16$ centered at that seed is extracted from the image and passed through a CNN which outputs a binary segmentation volume. The results showed that this method outperformed their previously proposed method [6] achieving an accuracy of 93.82%. In 2019, Ho et al. [10] proposed a new approach for nuclei segmentation in fluorescence microscopy images. It is based on the combination of two networks. The first one is a 3D U-Net which takes as input a patch extracted from a 3D fluorescence microscopy image and outputs the centroid coordinates of the detected nuclei and the binary nuclei segmentation mask. The second CNN takes as input the outputs of the 3D U-Net and returns an individual segmentation mask for each nucleus in the 3D patch. This approach achieves a precision, recall and F1-score of 93.47%, 96.80% and 95.10%, respectively.

Although there are already several approaches for the segmentation of nuclei in microscopy images, to the best of our knowledge deep learning based approaches for the segmentation of Golgi apparatus have not been investigated yet. So far, only traditional automatic methods ([11], [12], [13], [14], [15]) have been used for cell Golgi segmentation. Hence, segmentation of Golgi complex using deep learning is an open area of research. Furthermore, the nucleus-Golgi pair assignment is typically performed manually. This manual assignment is a challenging, time consuming and subjective task. In [5] an automatic method based on a cost matrix was presented to solve the task of nucleus-Golgi assignment. However, this method requires the manual tuning of some parameters (for instance, the maximum distance between a nucleus and a Golgi), and it is based on hand crafted features (the distance between the nucleus and Golgi centroids). Deep learning based approaches don't require the manual tuning of parameters and have the ability to automatically extract meaningful features from the input data [16], [17]. Therefore, in this work we use a deep learning based method to perform nucleus-Golgi pairing. More specifically, we present three approaches to jointly segment and pair nuclei and Golgi in 3D fluorescence microscopy images. The pairing of nuclei and Golgi is achieved by predicting their joint segmentation mask. In one approach, the nuclei, Golgi and joint segmentation masks are predicted at the same time. In the other two approaches first the joint segmentation mask is predicted, it is multiplied by the image, and the resulting image is used to predict the nuclei and Golgi segmentation masks. In these approaches the joint segmentation mask will then be used to perform nucleus-Golgi assignment and compute the polarity vectors.

III. METHODOLOGY

In this section we present the dataset used in this work and the approaches for the segmentation and assignment of nuclei and Golgi.

A. Dataset

We used a dataset of 8 crops extracted from fluorescence microscopy images of mouse retinas. These crops have sizes

varying between $257 \times 505 \times 55$ and $627 \times 818 \times 61$. In these crops nuclei are labeled with green fluorescence protein (GFP) and Golgi are labeled with mCherry. An example of a 3D crop is shown in Fig. 1(a). The ground truth nuclei and Golgi segmentation masks were created manually. The nucleus-Golgi vectors (Fig. 1(b)) were also manually annotated.

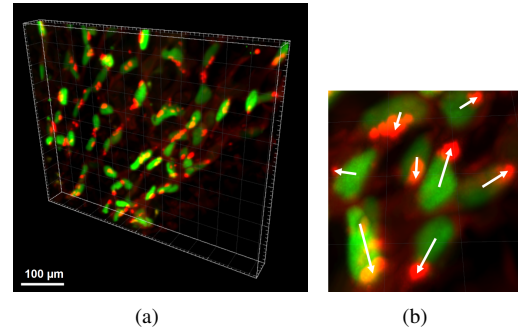


Fig. 1: Dataset. **a)** Example of a 3D crop of a microscopy image of mouse retina (size: $505 \times 618 \times 56$). The green and red channels contain the GFP and mCherry signals, respectively. **b)** Example of nucleus-Golgi vectors annotated manually (shown in white).

B. Segmentation

In this work, we present three approaches based on a 3D U-Net to perform joint segmentation of nuclei and Golgi in fluorescence microscopy images. Additionally, we compare their performance with an approach that performs segmentation based on the U-Net and pairing based on a cost matrix. Thus, the approaches considered in this work are:

- U-Net with 2 classes (U-Net 2C): This U-Net takes as input the 3D image and outputs an image with two channels (red and green) containing the Golgi and nucleus segmentation masks, respectively (Fig. 2(a)).
- U-Net with 3 classes (U-Net 3C): This U-Net is similar to the U-Net 2C, except there is a third output channel (blue) containing the segmentation mask of both nuclei and Golgi (Fig. 2(b)).
- U-Net for Hierarchical Segmentation (U-Net HS): This approach is composed by two U-Nets and based on the idea presented in [18]. The first one, takes as input the 3D image and outputs a binary segmentation mask of nuclei and Golgi. Afterwards, the output of the first U-Net is multiplied by the input image and fed to the second U-Net. The second U-Net outputs the nucleus and Golgi segmentation masks separately (Fig. 2(c)). We present two approaches based on the U-Net HS: U-Net HS A and U-Net HS B, the differences between these approaches are described in subsection IV-A.

C. Assignment

In this work we present a new assignment algorithm (Algorithm 1) based on the joint segmentation mask of nuclei and Golgi (blue channel in Fig. 2(b) and binary mask returned by the first U-Net shown in Fig. 2(c)). First, the connected components in the joint segmentation mask are

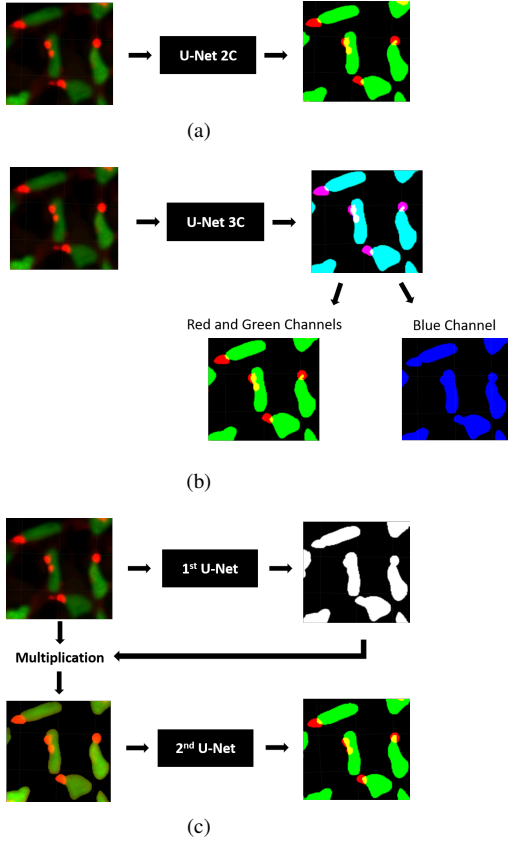


Fig. 2: Graphical representation of the segmentation approaches. **a)** U-Net 2C. **b)** U-Net 3C. **c)** U-Net HS.

Algorithm 1 Assignment algorithm based on the joint segmentation mask of nuclei and Golgi.

```

1: for object in jointmask do
2:   binmask = zeros(shape(jointmask))
3:   binmask[jointmask==object.label] = 1
   ▷ binary mask which is one for the voxels belonging
   to object, and zero for all the other voxels
4:   binmaskn = binmask*nmask ▷ intersection
   between binmask and nmask (nucleus mask)
5:   binmaskg = binmask*gmask ▷ intersection
   between binmask and gmask (Golgi mask)
6:   NuC = centr(binmaskn) ▷ nucleus centroid
7:   GoC = centr(binmaskg) ▷ Golgi centroid
8:   polvec = GoC - NuC ▷ polarity vector
9:   list.add(polvec)

```

identified. Afterwards, for each object in this mask, the corresponding nucleus and Golgi segmentation masks are extracted from the nuclei and Golgi masks predicted by the U-Net. Thereafter, the centroids of the nucleus and Golgi and the polarity vectors are computed.

The proposed assignment algorithm can be used with the segmentation results obtained with U-Net 3C, U-Net HS A and U-Net HS B. For the U-Net 2C we used the assignment algorithm based on the minimization of a cost matrix as presented in [5]. In this approach, a matrix is built containing the Euclidean distances between the centroid of each nucleus

and the centroid of each Golgi. Thereafter, the cost of this matrix is minimized using the Hungarian method. Finally, a maximum limit to the distance between a nucleus and Golgi centroid is applied.

IV. EXPERIMENTAL RESULTS

In this section we present details regarding the training of the approaches for segmentation, the results and discussion.

A. Training Setup

In this study we compared the performance of four approaches: U-Net 2C (shown in Fig. 2(a)), U-Net 3C (shown in Fig. 2(b)), U-Net HS A and U-Net HS B (approaches based on U-Net HS shown in Fig. 2(c)). The 3D U-Net code is based on a publicly available implementation [19]. All U-Nets were trained from scratch. U-Net 2C, U-Net 3C and the second U-Nets of both U-Net HS A and U-Net HS B were trained using the weighted binary cross-entropy as the loss function. The first net of U-Net HS A was trained using the binary cross-entropy as the loss function. The first net of U-Net B was trained using the following loss function:

$$Loss = \alpha \times \text{bincrossentropy}(GTmask, PREDmask) + \beta \times \text{bincrossentropy}(PROBmap, PREDmask) \quad (1)$$

where bincrossentropy denotes the binary cross entropy loss function, GTmask and PREDmask denote the ground truth and predicted segmentation masks, respectively; PROBmap is a probability map; α and β were set to 0.5. Finally, the second term represents the objectness prior term described in [20]. The probability maps were computed from the distance maps. A distance map is calculated based on the ground truth segmentation mask, where the value of each voxel represents its distance to the closest background pixel. In this study we used the Chebyshev distance.

All the approaches were trained for 200 epochs with learning rate of 1e-3 and another 200 epochs with learning rate of 1e-4. The validation split is equal to 20%. The dataset was divided into a training (6 crops) and a test set (2 crops). The U-Nets were trained with patches of size $128 \times 128 \times 64$ extracted from the training crops, and their performance was tested on the test crops.

B. Comparison between different approaches

The nuclei and Golgi detection results are shown in Table I. To compute the true positives (TP), false positives (FP) and false negatives (FN) a matrix containing the Euclidean distances between the ground truth and predicted object's centroids is built. In this way, an assignment between the ground truth and predicted objects is performed when the distance is below a threshold of 35 voxels: TP correspond to objects that are both in the ground truth and predicted segmentation masks, FP are the objects that only appear in the predicted masks, and FN the ones that only appear in the ground truth masks. We also compute the true positive, false positive and false negative rates TPR, FPR and FNR, respectively. The results in this table show that the addition of a third channel to the U-Net (U-Net 3C) improves the

detection performance compared to the U-Net with two output channels (U-Net 2C). In fact, the number of TP increases and the number of FP and FN decreases. The best approaches are based on the hierarchical segmentation (U-Net HS A and U-Net HS B). U-Net HS A is the approach that presents more TP and U-Net HS B is the one that presents less FP in comparison with the other approaches. These results show that grouping of nuclei and Golgi using deep learning improves the nuclei and Golgi detection.

TABLE I: Nuclei and Golgi detection results (the best results are highlighted in bold).

	Nuclei					
	TP	FP	FN	TPR	FPR	FNR
U-Net 2C	52	30	44	0.54	0.37	0.46
U-Net 3C	57	14	39	0.59	0.20	0.41
U-Net HS A	60	10	36	0.63	0.14	0.38
U-Net HS B	58	4	38	0.60	0.07	0.40
	Golgi					
	TP	FP	FN	TPR	FPR	FNR
U-Net 2C	81	31	15	0.84	0.28	0.16
U-Net 3C	71	16	25	0.74	0.18	0.26
U-Net HS A	81	22	15	0.84	0.21	0.16
U-Net HS B	78	25	18	0.81	0.24	0.19

The nucleus-Golgi assignment results are presented in table II. For the evaluation of the assignment algorithm in addition to calculating the TP, FP, FN, TPR, FPR and FNR, for the TP we also compute the cosine similarity and the distance between the ground truth and predicted vectors. The cosine similarity measures the cosine of the angle between both vectors. The distance between vectors u and v is calculated as $distance = \|u - v\|$. According to this table U-Net HS B is the approach that presents the highest value of TP and the lowest values of FP and FN, showing that the addition of the objectness prior term in the loss function improves the quality of the joint segmentation mask (output of the first U-Net), and consequently the assignment results are improved as well. Moreover, these results enhance again the importance of the deep learning based nuclei and Golgi pairing. The assignment algorithm based on this pairing outperforms the assignment method based on a cost matrix.

TABLE II: Nucleus-Golgi assignment and polarity vectors results (mean \pm std) (the best results are highlighted in bold).

U-Net	Cosine similarity	Distance length (μ m)	TP	FP	FN	TPR	FPR	FNR
2C	0.91 \pm 0.19	2.04 \pm 3.00	49	22	47	0.51	0.31	0.49
3C	0.95 \pm 0.13	1.59 \pm 2.50	48	15	48	0.50	0.24	0.50
HS A	0.95 \pm 0.12	1.76 \pm 2.51	55	12	41	0.57	0.18	0.43
HS B	0.94 \pm 0.16	2.09 \pm 2.92	56	7	40	0.58	0.11	0.42

V. CONCLUSIONS AND FUTURE WORK

In this work we have tackled the important problem of nuclei and Golgi complex segmentation and pairing in 3D fluorescence microscopy images of mouse retina. This step is important to unravel the mechanisms of vascular patterning. We presented three approaches based on deep learning that perform joint segmentation and pairing of nuclei and Golgi in microscopy images. Based on the pairing results we presented a new nucleus-Golgi assignment algorithm. Our results show that the prediction of the joint segmentation mask of nuclei and Golgi improved both the detection and

assignment results. However, the U-Net is not able to split touching objects, thus the FNR is high. In the future, in order to detect each object separately, we will use an instance segmentation architecture (such as the Mask R-CNN) to segment and group nuclei and Golgi. These segmentation results will be used to perform nucleus-Golgi pairing using the algorithms presented in this work.

REFERENCES

- [1] M. Potente and T. Mäkinen, "Vascular heterogeneity and specialization in development and disease," *Nature Reviews Molecular Cell Biology*, vol. 18, no. 8, p. 477, 2017.
- [2] S. P. Herbert and D. Y. Stainier, "Molecular control of endothelial cell behaviour during blood vessel morphogenesis," *Nature reviews Molecular cell biology*, vol. 12, no. 9, pp. 551–564, 2011.
- [3] C. G. Fonseca, P. Barbacena, and C. A. Franco, "Endothelial cells on the move: dynamics in vascular morphogenesis and disease," *Vascular Biology*, vol. 2, no. 1, pp. H29–H43, 2020.
- [4] C. A. Franco, M. L. Jones, M. O. Bernabeu, A.-C. Vion, P. Barbacena, J. Fan, T. Mathivet, C. G. Fonseca, A. Ragab, T. P. Yamaguchi, *et al.*, "Non-canonical Wnt signalling modulates the endothelial shear stress flow sensor in vascular remodelling," *Elife*, vol. 5, p. e07727, 2016.
- [5] P. Barbacena, M. Ouarné, J. J. Haigh, F. F. Vasconcelos, A. Pezzarossa, and C. A. Franco, "GNrep mouse: a reporter mouse for front–rear cell polarity," *genesis*, vol. 57, no. 6, p. e23299, 2019.
- [6] D. Joon Ho, C. Fu, P. Salama, K. W. Dunn, and E. J. Delp, "Nuclei segmentation of fluorescence microscopy images using three dimensional convolutional neural networks," in *IEEE CVF*, pp. 82–90, 2017.
- [7] C. Fu, D. J. Ho, S. Han, P. Salama, K. W. Dunn, and E. J. Delp, "Nuclei segmentation of fluorescence microscopy images using convolutional neural networks," in *ISBI 2017*, pp. 704–708, IEEE, 2017.
- [8] C. Fu, S. Lee, D. Joon Ho, S. Han, P. Salama, K. W. Dunn, and E. J. Delp, "Three dimensional fluorescence microscopy image synthesis and segmentation," in *IEEE CVF*, pp. 2221–2229, 2018.
- [9] D. J. Ho, C. Fu, P. Salama, K. W. Dunn, and E. J. Delp, "Nuclei detection and segmentation of fluorescence microscopy images using three dimensional convolutional neural networks," in *ISBI 2018*, pp. 418–422, IEEE, 2018.
- [10] D. J. Ho, S. Han, C. Fu, P. Salama, K. W. Dunn, and E. J. Delp, "Center-extraction-based three dimensional nuclei instance segmentation of fluorescence microscopy images," in *2019 IEEE EMBS BHI*, pp. 1–4, IEEE, 2019.
- [11] G. Srinivasa, M. C. Fickus, Y. Guo, A. D. Linstedt, and J. Kovacevic, "Active mask segmentation of fluorescence microscope images," *IEEE Transactions on Image Processing*, vol. 18, no. 8, pp. 1817–1829, 2009.
- [12] K. K. K. Mai and B.-H. Kang, "Semiautomatic segmentation of plant golgi stacks in electron tomograms using 3DMod," in *Plant Protein Secretion*, pp. 97–104, Springer, 2017.
- [13] G. Srinivasa, M. Fickus, M. N. Gonzalez-Rivero, S. Y. Hsieh, Y. Guo, A. D. Linstedt, and J. Kovacevic, "Active mask segmentation for the cell-volume computation and Golgi-body segmentation of HeLa cell images," in *ISBI*, pp. 348–351, IEEE, 2008.
- [14] T. J. Gniadek and G. Warren, "WatershedCounting3D: a new method for segmenting and counting punctate structures from confocal image data," *Traffic*, vol. 8, no. 4, pp. 339–346, 2007.
- [15] N. Fordyce, S. J. McKenna, C. Hacker, and J. Lucocq, "Finding golgi stacks in electron micrographs," in *MIA*, pp. 59–64, 2014.
- [16] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," *Computational intelligence and neuroscience*, vol. 2018, 2018.
- [17] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and M. K. Khan, "Medical image analysis using convolutional neural networks: a review," *Journal of medical systems*, vol. 42, no. 11, p. 226, 2018.
- [18] S. Eppel, "Hierarchical semantic segmentation using modular convolutional neural networks," *arXiv preprint arXiv:1710.05126*, 2017.
- [19] "3DUnetCNN," <https://github.com/ellisdg/3DUnetCNN>, Visited on March 2020.
- [20] A. Bearman, O. Russakovsky, V. Ferrari, and L. Fei-Fei, "What's the point: Semantic segmentation with point supervision," in *ECCV*, pp. 549–565, Springer, 2016.