A General Stitching Solution for Whole-Brain 3D Nuclei Instance Segmentation from Microscopy Images

Ziquan Wei[1,2], Tingting Dan[1], Jiaqi Ding[1,2], Mustafa Dere[1], and Guorong Wu[1,2(✉)]

[1] Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
guorong_wu@med.unc.edu

[2] Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

Abstract. High-throughput 3D nuclei instance segmentation (NIS) is critical to understanding the complex structure and function of individual cells and their interactions within the larger tissue environment in the brain. Despite the significant progress in achieving accurate NIS within small image stacks using cutting-edge machine learning techniques, there has been a lack of effort to extend this approach towards whole-brain NIS from light-sheet microscopy. This critical area of research has been largely overlooked, despite its importance in the neuroscience field. To address this challenge, we propose an efficient deep stitching neural network built upon a knowledge graph model characterizing 3D contextual relationships between nuclei. Our deep stitching model is designed to be agnostic, enabling existing limited methods (optimized for image stack only) to overcome the challenges of whole-brain NIS, particularly in addressing the issue of inter- and intra-slice gaps. We have evaluated the NIS accuracy on top of three state-of-the-art deep models with $128 \times 128 \times 64$ image stacks, and visualized results in both inter- and intra-slice gaps of whole brain. With resolved gap issues, our deep stitching model enables the whole-brain NIS (gigapixel-level) on entry-level GPU servers within 27 h.

Keywords: Image stitching · 3D microscopy image · Whole-brain nucleus instance segmentation · Graph neural network

1 Introduction

Light-sheet microscopy is a powerful imaging modality that allows for fast and high-resolution imaging of large samples, such as the whole brain of the

Supported by NIH R01NS110791, NIH R01MH121433, NIH P50HD103573, and Foundation of Hope.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43901-8_5.

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14223, pp. 46–55, 2023.
https://doi.org/10.1007/978-3-031-43901-8_5

mouse [3,14]. Tissue-clearing techniques enable the removal of light-scattering molecules, thus improving the penetration of light through biological samples and allowing for better visualization of internal structures, including nuclei [2,16]. Together, light-sheet microscopy and tissue-clearing techniques have revolutionized the field of biomedical imaging and they have been widely used for studying the structure and function of tissues and organs at the cellular level.

Accurate 3D nuclei instance segmentation plays a crucial role in identifying and delineating individual nuclei within three-dimensional space, which is essential for understanding the complex structure and function of biological tissues in the brain. Previous [10] and [5] have applied graph-based approaches that model links of voxel and neuron region, respectively, for 3D neuron segmentation from electron microscopy image stacks. However, accurate segmentation of nuclei from light-sheet microscopy images of cleared tissue can be a challenging task due to the presence of complex tissue structures, cell shapes, and variations in nuclei size and shape [1]. Due to the high cost of 3D manual nuclei annotations and the complexity of learning, current end-to-end NIS models are typically limited to training and testing on small image stacks (e.g., $128\times128\times64$). Considering these limitations, one approach for achieving whole-brain NIS is dividing the whole stack into smaller stacks, so the existing NIS methods can handle each piece individually. In such a scenario, constructing the whole-brain nuclei instance segmentation in 3D from these smaller image stacks arises a new challenge. The gaps between these smaller stacks (intra-slice) and the slices (inter-slice) require a robust stitching method for accurate NIS. We show these gaps in Fig. 1. Note, the intra-slice gap, commonly referred to as the boundary gap, arises due to the existence of boundaries in the segmentation outcome of image stacks and poses a challenge in achieving smooth segmentation between neighboring image stacks. Current approaches may, however, undermine the overall quality of the whole-brain NIS results which leads to inaccurate reports of nuclei counts. Figure 1 (left) illustrates the typical examples of the boundary gap issues, where the red circle highlights the incidence of over-counting (both partial nuclei instances are recognized as a complete nucleus in the corresponding image stack), while the dashed blue box denotes the issue of under-counting (none of the partial nuclei instances has been detected in each image stack).

It is a common practice to use overlapped image stacks to stitch the intensity image (continuous values) by weighted averaging from multiple estimations [12,15]. However, when nuclei are in close proximity and represent the same entity, it becomes crucial to accurately match the indexes of nuclei instances, which refer to the segmentation labels. We call this the nuclei stitching issue. This issue presents a significant challenge in the pursuit of achieving whole-brain NIS. To address this non-trivial challenge, we formulate this problem as a knowledge graph (KP) task that is built to characterize the nuclei-to-nuclei relationships based on the feature presentation of partial image appearance. By doing so, the primary objective of this learning problem is to determine whether to merge two partial nuclei instances that exist across different slices or stacks. Drawing inspiration from recent research on object tracking using graph models,

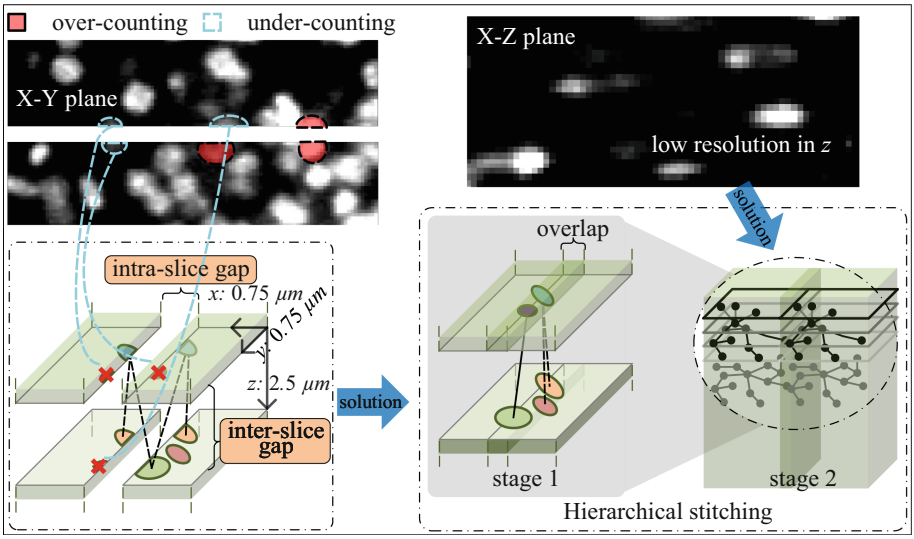

Fig. 1. Top left: Nuclei over-counting (in red) and under-counting (in dashed blue) issues due to the boundary gaps across image stacks. Bottom left: Undetected nuclei instances make stitching difficult due to the absence of partial nuclei instances (indicated by red crosses). Top right: The limited inter-slice resolution presents a challenge for NIS in each image stack, particularly for partial nuclei instances near the stack boundary. Bottom right: Our hierarchal stitching framework leverages overlapping strategy in X-Y plane to reduce the chance of absent nuclei instances in intra-slice stitching. (Color figure online)

we construct a graph contextual model to assemble 3D nuclei in a graph neural network (GNN). In the overlapping area, the complete 2D nucleus instance is represented as a graph node, and the links between these nodes correspond to the nuclei-to-nuclei relationship.

Conventional knowledge graph learning typically emphasizes the learning of the relationship function. In this context, it appears that the process of whole-brain NIS largely depends on the relationship function to link partial nuclei instances along the gap between slices. Nonetheless, a new challenge arises from the NIS backbone due to the anisotropic image resolution where inter-slice Z resolution (e.g., 2.5 μm) is often several times lower than the in-plane X-Y resolution ($0.75 \times 0.75\,\mu m^2$). That is, the (partial) nuclei instances located near the boundary across X-Y planes (along the inter-slice direction with poor image resolution) in the 3D image stack have a large chance of being misidentified, leading to the failure of NIS stitching due to the absence of nuclei instances that are represented as nodes in the contextual graph model. To alleviate this issue, we present a two-stage hierarchical whole-brain NIS framework that involves stitching 2D NIS results in the X-Y plane (stage 1) and then assembling these 2D instances into 3D nuclei using a graph contextual model (stage 2). The conjecture is that the high resolution in the X-Y plane minimizes the risk of missing

2D NIS instances, allowing the knowledge graph learning to be free of absent nodes in establishing correspondences between partial nuclei instances.

Stitching 2D nuclei instances in each image slice is considerably easier than stitching across image slices due to the finer image resolution in the X-Y plane. However, as shown in Fig. 1 (right), the poor resolution in the Z-axis makes it challenging to identify partial nuclei instances along this axis. Thus, pre-stitching in the X-Y plane of the first stage can reduce the probability of having 2D nuclei instances missing along the Z axis at the second stage. Among this, we train the graph contextual model to predict the nuclei-to-nuclei correspondence across image slices, where each node is the 2D nuclei instance without the intra-slice gap issue. Since stage 2 is on top of the existing NIS methods in stage 1, our stitching framework is agnostic and can support any state-of-the-art NIS methods to expand from small image stacks to the entire brain.

In the experiments, we have comprehensively evaluated the segmentation and stitching accuracy (correspondence matching between 2D instances) and whole-brain NIS results (both visual inspection and quantitative counting results). Compared to no stitching, Our deep stitching model has shown a significant improvement in the whole-brain NIS results with different state-of-the-art models, indicating its potential for practical applications in the field of neuroscience.

2 Methods

Current state-of-the-art NIS methods, such as Mask-RCNN [6,7], 3D Unet [4,8] and Cellpose [9,11] are designed to segment nuclei instances in a pre-defined small image stack only. To scale up to whole-brain NIS, we propose a graph-based contextual model to establish nuclei-to-nuclei correspondences across image stacks. On top of this backbone, we present a hierarchical whole-brain NIS stitching framework that is agnostic to existing NIS methods.

2.1 Graph Contextual Model

Problem Formulation. Our graph contextual model takes a set of partial nuclei instances, sliced by the inter-slice gap, as input. These nuclei instances can be obtained using a 2D instance segmentation method, such as Mask-RCNN. The output of our model is a collection of nuclei-to-nuclei correspondences, which enable us to stitch together the NIS results from different image slices (by running NIS separately). We formulate this correspondence matching problem as a knowledge graph learning task where the links between nodes in the graph contextual model represent the probability of them belonging to the same nuclei instance. In this regard, the key component of NIS stitching becomes seeking for a relationship function that estimates the likelihood of correspondence based on the node features, i.e., image appearance of to-be-stitched 2D nuclei instances.

Machine Learning Components in Graph Contextual Model. *First*, we construct an initial contextual graph $G = \{\mathbf{V}, \mathbf{E}\}$ for each 2D nucleus instance x (i.e., image appearance vector). The set of nodes $\mathbf{V} = \{x_i | \mathcal{D}(x, x_i) > \delta\}$ includes

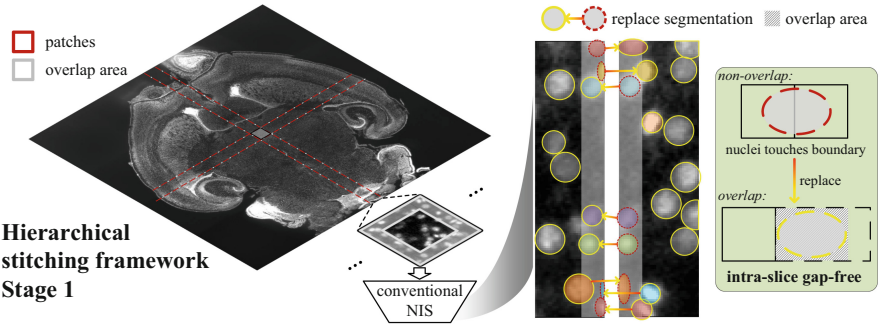

Fig. 2. Stage one of the proposed hierarchical stitching framework for whole-brain NIS. Resolve the intra-slice gap in the X-Y plane by overlap. (Color figure online)

all neighboring 2D nuclei instances, where the distance between the centers of two instances is denoted by $\mathcal{D}$, and δ is a predefined threshold. The matrix $\mathbf{E} \in \mathbb{R}^{N \times N}$ represents the edges between nodes, where N is the number of neighboring instances. Specifically, we compute the Intersection over Union (IoU) between the two instances and set the edge weight as $e_{ij} = IoU(x_i, x_j)$.

Second, we train the model on a set of contextual graphs G to recursively (1) find the mapping function γ to describe the local image appearance on each graph node and (2) learn the triplet similarity function ψ.

- *Graph feature representation learning.* For the k^{th} iteration, we enable two connected nodes to exchange their feature representations constrained by the current relationship topology e_{ij}^{k} by the k^{th} layer of the deep stitching model. In this context, we define the message-passing function as:

$$x_i^{(k+1)} = \gamma^{(k+1)} \left(x_i^{(k)}, \Sigma_{j \in \mathcal{N}(i)} \phi^{(k)}(x_i^{(k)}, x_j^{(k)}, e_{j,i}^{(k)}) \right). \tag{1}$$

  Following the popular learning scheme in knowledge graphs [13], we employ Multilayer Perceptron (MLP) to act functions γ, ϕ.
- *Learning the link-wise similarity function to predict nuclei-to-nuclei correspondence.* Given the updated node feature representations $\{x_i^{(k+1)}\}$, we train another MLP to learn the similarity function ψ in a layer-by-layer manner. In the k^{th} layer, we update each 2D-to-3D contextual correspondence $e_{j,i}^{(k+1)}$ for the next layer by

$$e_{j,i}^{(k+1)} = \psi^{(k+1)} \left(x_i^{(k+1)}, x_j^{(k+1)}, e_{j,i}^{(k)} \right). \tag{2}$$

2.2 Hierarchical Stitching Framework for Whole-Brain NIS

Our graph contextual model is able to stitch the NIS results across the intra-slice gap areas in X/Y-Z plane. As demonstrated in Fig. 1, the accuracy of no

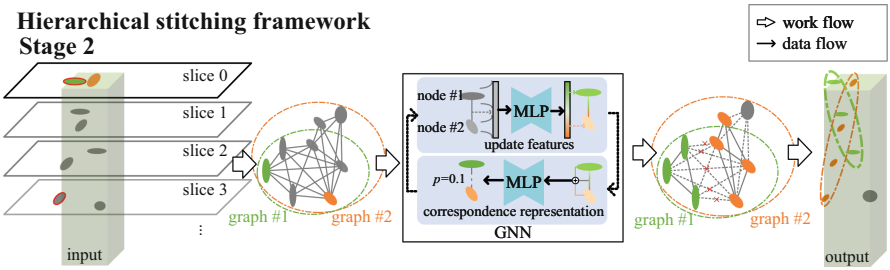

Fig. 3. Stage two of the proposed hierarchical stitching framework for whole-brain NIS. Graph contextual model for inter-slice gap.

stitching is limited by the absence of 2D partial nuclei instances due to the poor inter-slice resolution. With the graph model, as shown in Fig. 2 and 3, we propose a hierarchical stitching framework for whole-brain NIS in two stages.

1. *Resolve intra-slice gap in X-Y plane.* Suppose that each within-stack NIS result overlaps with its neighboring image stack in the X-Y plane. Then, we can resolve the intra-slice gap problem in X-Y plane in three steps: (i) identify the duplicated 2D nuclei instances from multiple overlapped image stacks, (ii) find the representative NIS result from the "gap-free" image stack, and (iii) unify multiple NIS estimations by using the "gap-free" NIS estimation as the appearance of the underlying 2D nuclei. The effect of this solution is shown in Fig. 2 right, where the gray areas are spatially overlapped. We use the arrows to indicate that the 2D nuclei instances (red dash circles) have been merged to the counterpart from the "gap-free" image stack (yellow circles).
2. *Inter-slice stitching using graph contextual model.* At each gap area along Z-axis, we deploy the graph contextual model to stitch the sliced nuclei instances. Specifically, we follow the partition of the whole-brain microscopy image in stage 1, that is a set of overlapped 3D image stacks, all the way from the top to the bottom as shown in the left of Fig. 3. It is worth noting that each 2D nuclei instance in the X-Y plane is complete, as indicated by the red-highlighted portion extending beyond the image stack. Next, we assign a stack-specific local index to each 2D nuclei instance. After that, we apply the (trained) graph contextual model to each 2D nuclei instance. By tracking the correspondences among local indexes, we remove the duplicated inter-slice correspondence and assign the global index to the 3D nuclei instance.

2.3 Implementation Details

We empirically use 18.75 μm for the overlap size between two neighboring image stacks in X-Y plane. The conventional NIS method is trained using 128×128 patches. For the graph contextual model, the MLPs consist of 12 fully-connected layers. Annotated imaging data has been split into training, validation, and testing sets in a ratio of 6:1:1. Adam is used with $lr = 5e - 4$ as the optimizer, $Dropout = 0.5$, and focal loss as the loss function.

3 Experiments

3.1 Experimental Settings

Stitching Methods Under Comparison. We mainly compare our stitching method with the conventional analytic approach by IoU (Intersection over Union) matching, which is widely used in object detection with no stitching. We perform the experiments using three popular NIS deep models, that are Mask-RCNN-R50 (with ResNet50 backbone), Mark-RCNN-R101 (with ResNet101 backbone) and CellPose, with two stitching methods, i.e., our hierarchical stitching framework and IoU-based matching scheme.

Data and Computing Environment. In the following experiments, we first train Mask-RCNN-R50, Mask-RCNN-R101, and CellPose on 16 image stacks ($128 \times 128 \times 64$), which include in total 6,847 manually labeled 3D nuclei. Then we make the methods comparison in two ways. For the stitching comparison based on 16 image stacks, We integrate each NIS model into two stitching methods respectively, which yields six NIS methods and corresponding results. For the stitching comparison based on whole-brain images with the size of $8,729 \times 9,097 \times 1,116 voxel^3$ and the resolution of $0.75\,\mu m \times 0.75\,\mu m \times 2.5\,\mu m$, we firstly partition the one whole-brain image into 89,424 ($69 \times 72 \times 18$) image stacks with the size of $128 \times 128 \times 64$, then employing the best NIS model to segment nuclei instances in each image stack in a parallel manner, finally we deploy our hierarchical stitching framework to stitch the image stacks together. All experiments are run on a Linux server with 4 GPUs (24 GB) and 48 CPUs with 125 GB RAM.

3.2 Evaluation Metrics

We use the common metrics precision, recall, and F1 score to evaluate the 3D NIS between annotated and predicted nuclei instances. Since the major challenge of NIS stitching is due to the large inter-slice gap, we also define the stitching accuracy for each 3D nuclei instance by counting the number of 2D NIS (in the X-Y plane) that both manual annotation and stitched nuclei share the same instance index.

3.3 Evaluating the Accuracy of NIS Stitching Results

Quantitative Evaluation. As shown in Fig. 4, there is a clear sign that NIS models with our hierarchical stitching method outperform IoU-based counterparts on NIS metrics, regardless of the NIS backbone models. In average, our hierarchical stitching method has improved 14.0%, 5.1%, 10.2%, and 3.4% in precision, recall, F1 score, and stitching accuracy, respectively compared with IoU-based results.

Visual Inspection. We also show the visual improvement of whole-brain NIS results before and after stitching in Fig. 5. Through the comparison, it is apparent that (1) the inconsistent NIS results along the intra-slice gap area have been

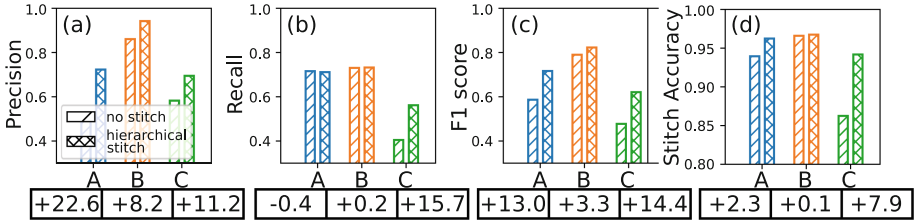

Fig. 4. The NIS precision (a), recall (b), F1 score (c), and stitching accuracy (d) by stitching or not, where the NIS backbones include Mask-RCNN-R50 (A, blue), Mask-RCNN-R101 (B, orange), CellPose (C, green). (Color figure online)

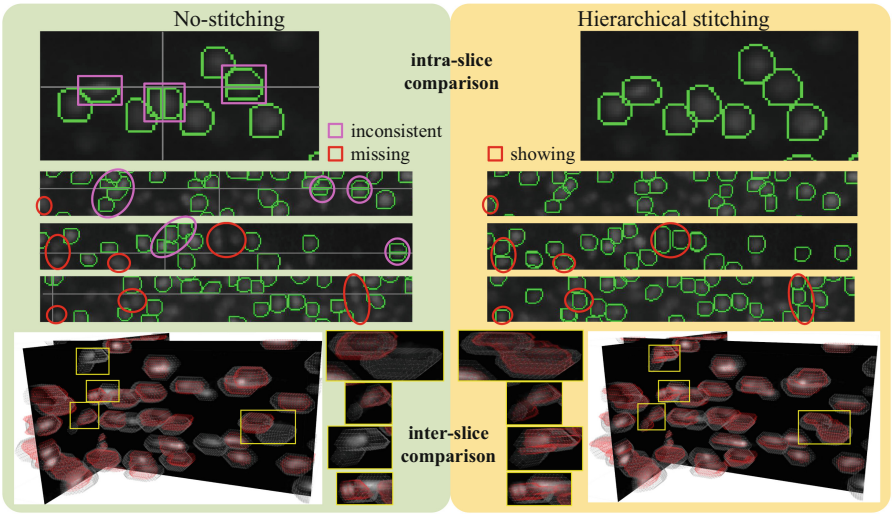

Fig. 5. Visual comparison before (left) and after (right) stitching. Red circles show the missing instances when there is no stitching, and Pink circles indicate inconsistency issues at the intra-slice gap areas. Inter-slice comparison is a zoom-in of 3D visualization (bottom, yellow-highlighted), where red instance denotes NIS results, and gray is GT. (Color figure online)

corrected after stitching (indicated by Pink circles), so there is no apparent boundary between the same nucleus. And (2) the graph contextual model on each nuclei instance alleviates the issue of missing nuclei (indicated by Red circles) by correspondence detection across the inter-slice gap areas.

3.4 Whole-Brain NIS in Neuroscience Applications

One of the important steps in neuroscience studies is the measurement of regional variations in terms of nuclei counts. In light of this, we evaluate the counting accuracy. Since we only have 16 image stacks with manual annotations, we sim-

ulate the stack-to-stack gap in the middle of each image stack and compare counting results between whole-brain NIS with or without hierarchical stitching. Compared to whole-brain NIS without stitching, our hierarchical stitching method has reduced the nuclei counting error from 48.8% down to 10.1%. The whole-brain NIS improvement has vividly appeared in Fig. 5.

In addition, the running time of whole-brain NIS is around 26 h on average for the typical light-sheet microscopy images of the mouse brain. More visual results of the whole-brain NIS can be found in supplementary materials.

4 Conclusion

In this work, we introduce a learning-based stitching approach to achieve 3D instance segmentation of nuclei in whole-brain microscopy images. Our stitching framework is flexible enough to incorporate existing NIS methods, which are typically trained on small image stacks and may not be able to scale up to the whole-brain level. Our method shows great improvement by addressing inter- and intra-slice gap issues. The promising results in simulated whole-brain NIS, particularly in terms of counting accuracy, also demonstrate the potential of our approach for neuroscience research.

References

1. Alahmari, S.S., Goldgof, D., Hall, L.O., Mouton, P.R.: A review of nuclei detection and segmentation on microscopy images using deep learning with applications to unbiased stereology counting. IEEE Trans. Neural Netw. Learn. Syst. (2022)
2. Banerjee, A., Poddar, R.: Enhanced visualization of tissue microstructures using swept-source optical coherence tomography and edible oil as optical clearing agent. Optik **267**, 169693 (2022)
3. Bennett, H.C., Kim, Y.: Advances in studying whole mouse brain vasculature using high-resolution 3D light microscopy imaging. Neurophotonics **9**(2), 021902 (2022)
4. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-net: learning dense volumetric segmentation from sparse annotation. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016, Part II. LNCS, vol. 9901, pp. 424–432. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_49
5. Funke, J., Andres, B., Hamprecht, F.A., Cardona, A., Cook, M.: Efficient automatic 3D-reconstruction of branching neurons from EM data. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1004–1011. IEEE (2012)
6. He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2961–2969 (2017)
7. Iqbal, A., Sheikh, A., Karayannis, T.: DeNerD: high-throughput detection of neurons for brain-wide analysis with deep learning. Sci. Rep. **9**(1), 13828 (2019)
8. Lin, Z., et al.: NucMM dataset: 3D neuronal nuclei instance segmentation at sub-cubic millimeter scale. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12901, pp. 164–174. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87193-2_16

9. Pachitariu, M., Stringer, C.: Cellpose 2.0: how to train your own model. Nat. Methods 1–8 (2022)
10. Pape, C., Beier, T., Li, P., Jain, V., Bock, D.D., Kreshuk, A.: Solving large multicut problems for connectomics via domain decomposition. In: Proceedings of the IEEE International Conference on Computer Vision Workshops, pp. 1–10 (2017)
11. Stringer, C., Wang, T., Michaelos, M., Pachitariu, M.: Cellpose: a generalist algorithm for cellular segmentation. Nat. Methods **18**(1), 100–106 (2021)
12. Vu, Q.D., Rajpoot, K., Raza, S.E.A., Rajpoot, N.: Handcrafted histological transformer (H2T): unsupervised representation of whole slide images. Med. Image Anal. **85**, 102743 (2023)
13. Wang, H., Ren, H., Leskovec, J.: Relational message passing for knowledge graph completion. In: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, pp. 1697–1707 (2021)
14. Yang, B., et al.: DaXi-high-resolution, large imaging volume and multi-view single-objective light-sheet microscopy. Nat. Methods **19**(4), 461–469 (2022)
15. Yang, H., et al.: Deep learning-based six-type classifier for lung cancer and mimics from histopathological whole slide images: a retrospective study. BMC Med. **19**, 1–14 (2021)
16. You, S., et al.: High cell density and high-resolution 3D bioprinting for fabricating vascularized tissues. Sci. Adv. **9**(8), eade7923 (2023)