Robust Frequency-Aware Instance Segmentation for Serial Tissue Sections

Guodong Sun[1,2], Zejin Wang[1,2], Guoqing Li[1], and Hua Han[1,2,3]

[1] National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China
[2] School of Artificial Intelligence, School of Future Technology, University of Chinese Academy of Sciences, Beijing, China
[3] CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai, China
{sunguodong2019,wangzejin2018,guoqing.li,hua.han}@ia.ac.cn

Abstract. Serial tissue sections are widely used in imaging large tissue volumes. Navigating to each section is indispensable in the automatic imaging process. Nowadays, the locations of sections are labeled manually or semi-manually. Sections are similar and the border is indiscernible if they stick together, which makes it difficult to locate the sections automatically. In this paper, we present frequency-aware instance segmentation framework (FANet), which can extract shape and size information of sections very well. Firstly, FANet uses discrete cosine transform(DCT) . Secondly, each channel extracts an specific frequency component of themselves. Frequency components from all channels is taken as the multi-frequency description of feature map and finally used to model the channel attention. Additionally, we propose a dataset about the serial sections as benchmark, which contains 2708 images in training set and 1193 images in validation set. Experimental results on the benchmark demonstrate our FANet achieves superior performance compared with the current methods. Our code and dataset will be made public.

Keywords: Serial sections · Instance segmentation · Computer vision.

1 Introduction

Serial tissue sections offer superlative opportunities to ascertain biological tissues in three dimensions. A series of methods are proposed to image the serial sections. As a pioneering work, Micheva et al. [23] proposes array tomography(AT), which "array" means arranging serial ultrathin tissue sections in spatial array on a planar solid surface, while "tomography" means imaging the two-dimensional sections to capture of three-dimensional structure using microscope [29]. Therefore, many works combined AT with electron microscope to explore fine details in large tissue volumes [10,18,24,28,30,32], including serial-section transmission electron microscopy (ssTEM) [4,9] and scanning electron microscopy (SEM) [2,17]. Especially, SEM with its imaging method fits well with AT, thus

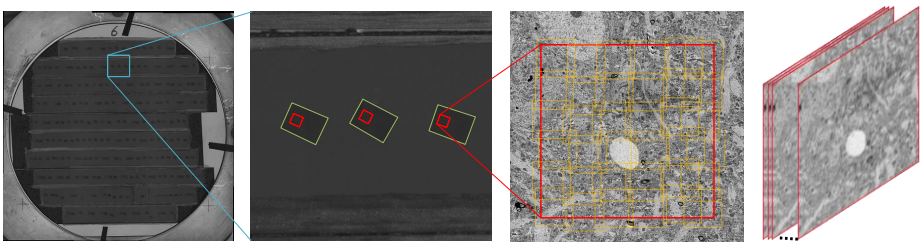


Fig. 1. The imaging process of ATUM-SEM.(a) is the image that silicon wafer which tape strips with sections are adhered to is on SEM stage. (b) is the enlarged view of the blue box of (a). The quadrangle is sections' contour, and the red rectangle is ROI on each section. (c) is SEM imaging ROI on a section. Each yellow square is the field of view (FOV) that SEM image the section once. (d) is sequentially stacking the ROIs on each section to get the 3D structure.

it is most often used with AT. Automatic Tape-collecting UltraMicrotome SEM (ATUM-SEM) is the combination of SEM and AT, using ultramicrotome to cut tissue volumes in ultrathin sections and automatically collecting the sections on tape in sequence. The tape is cut in strips and adhered to silicon wafers which is loaded on SEM stage to be imaged. Then, SEM images hundreds of sections' same region (ROI) on the wafer automatically , as illustrated/shown in Fig 1. In order to accomplish the automation of the imaging process on a wafer, researchers firstly take an image (wafer image) of the wafer, which has hundreds sections on and secondly label the contour of sections which contains ROI and satisfy the condition of imaging, then map wafer image coordinate to SEM stage physical coordinate. Therefore, after setting up the ROI on one section, SEM can navigate to ROIs on every section accurately. Nowadays, the contour of each section is labeled manually or semi-manually [11]. As the volume of tissue that researchers want to explore becoming larger, the number of sections become to tens of thousands [28,32] that requires a lot of people and time to label. Besides, on account of the uncertainty in section preparation processes such as cutting and collecting, the sections on the wafer could be damaged, contaminated with dust or stick together (as shown in Fig 2), bringing difficulty to label sections automatically.

To detect the contour of every section automatically, we accomplish this task using instance segmentation of sections on wafer image, which can segment each section in pixel level and classify the section abnormal or not. In recent years, instance segmentation task has been rapidly improved benefiting from convolutional neural networks (CNNs). Among them, , Mask R-CNN [12] is the dominant framework in this task as a two-stage method. Mask R-CNN incorporates the advantages of Faster R-CNN [26] and FCN [22], in consideration of the bounding-boxes, classification and masks for each objects simultaneously. Each bounding-box of an object is used to crop the feature maps using ROIAlign to get the ROI. Then FCN perform semantic segmentation to these ROIs. Many

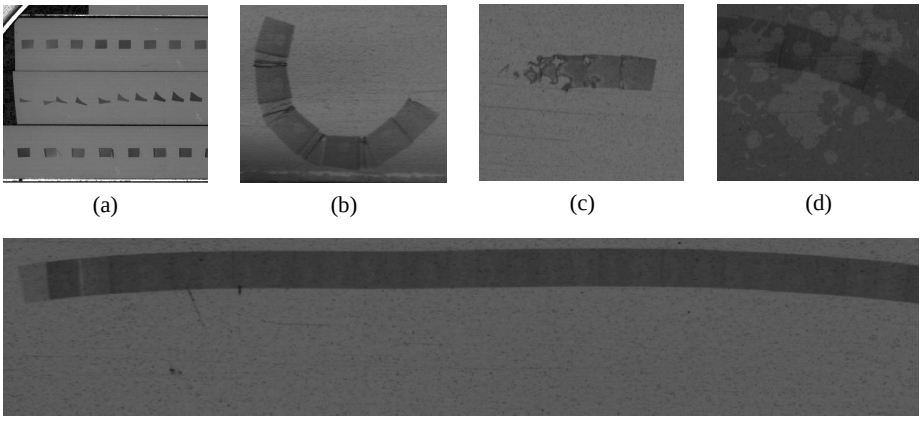


Fig. 2. Some exceptional situations of sections on wafer.(a) incomplete sections resulting from the replacement of diamond knife (b) fold sections stick to good sections. (c) damaged sections. (d) contaminated with dust. (e) sections stick together and the three sections on the far left are incomplete.

works [6,16,21] which have top performance on COCO [20] benchmark dataset are based on Mask R-CNN. However, different from the objects in natural image, the sections in the same wafer image are significantly similar to each other (e.g. in Fig 2, between the good sections or the incomplete and good sections). In addition, if the sections stick together, the boundary between sections is difficult to identify. Moreover, there is little difference between the sections of each category, especially incomplete sections sometimes (Fig 2(e)). Therefore, the performance of these works on wafer image could be sub-optimal, such as incorrect category or boundary. To better distinguish the different categories of sections and identify the boundary between the stick sections more accurately, inspired by [25], we propose a robust frequency-aware network (FANet) which uses multiple frequency components of feature map to model channel attention. FANet aggregates multiple frequency components efficiently and pools frequency information of the feature map. We apply discrete cosine transform (DCT) to feature map and convert it to frequency domain. Then we choose several specific frequency components as a multi-spectral description for feature map. Finally, the frequency components are used as feature map channel importance weight to let network take more attention on important information. Finally, we propose a dataset of wafer image and use it as the benchmark. Experimental results demonstrate that FANet achieves state-of-the-art performance on the benchmark.

In brief, the contributions of this paper are summarized as follows:

- We propose a robust method, FANet, which uses multiple frequency components to take channel attention more efficiently for take instance segmentation on wafer image.

- We propose a dataset of wafer image as benchmark which contains different conditions of sections as more as we can find, comprised of 2708 images as training set and 1193 images as validation set.
- Experiments results on the benchmark demonstrate FANet achieves state-of-the-art performance.

This paper is organized as follows. Sec. 2 briefly introduces the related work about instance segmentation. Sec. 3 presents the proposed method. Sec. 4 shows the experimental results on the benchmark. Finally, the conclusion is in Sec. 5.

2 Related Work

2.1 Instance Segmentation

Instance segmentation aims at predicting the mask and classification score for each object. Nowadays, the two-stage method Mask R-CNN is still the dominant framework which applies Faster R-CNN to get the bounding-boxes of instances and uses FCN to predict the mask of each instance on the feature map cropped according to the bounding-boxes. [6] combines Cascade R-CNN [5] with Mask R-CNN and interweaves detection and segmentation joint. [21] uses FPN features [19] to enriching the ROI feature. And [16] adds a MaskIOU head to learn the segmentation score. Besides, one-stage methods for instance segmentation are faster than two-stage methods conceptually. [8] takes this task into fully convolutional by predicts position-sensitive score maps. YOLACT [3] uses mask coefficients to distinguish mask for each object on the original mask. Due to the weak accuracy of one-stage method, our method is based on two-stage.

2.2 Channel Attention Mechanism

To date, channel attention mechanism has been widely used in CNNs and achieved satisfactory performance. SENet [15] uses GAP and fully connected layers to get the channel importance weight. And ECANet [31] uses local one-dimensional convolution instead of fully connected to reduce the redundancy. FcaNet [25] states that using GAP to get channel importance weight loses much useful information on feature channels. So it introduces DCT to get more information from different frequency components. However, FcaNet down-samples feature map before calculating frequency components, which much information does not be transformed into frequency domain. And FANet we proposed directly uses DCT to transform information into frequency domain, which can get all the information of specific frequency components. Therefore, FANet can take more information to model channel attention.

3 Method

In this section, we first introduce the DCT frequency on feature map. Then, we elaborate our proposed method FANet and explain why it aggregates multiple frequency components more efficiently and pools more frequency information.

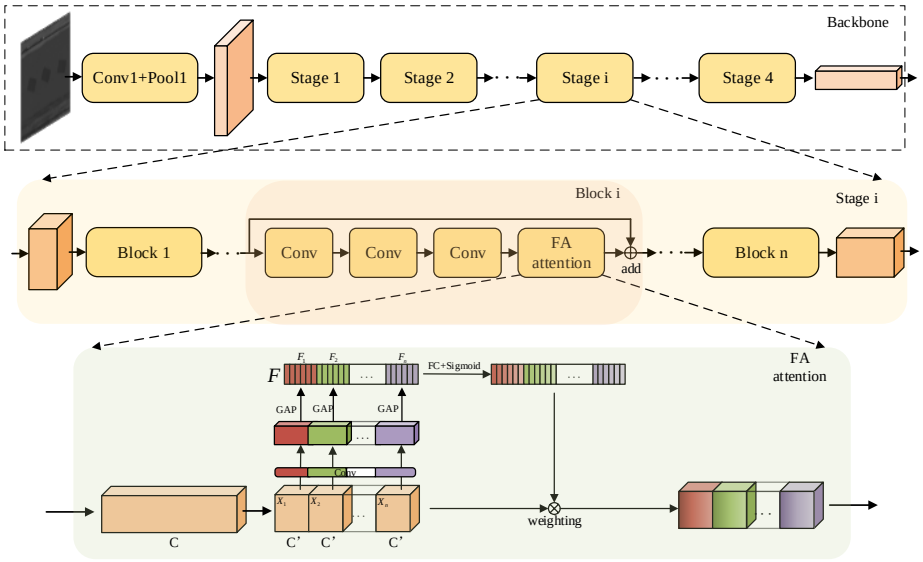


Fig. 3. The architecture of FANet. Frequency aware channel attention is placed in every block. First the input feature map is divided into n parts averagely. Then every channel of feature map in the same part calculates one specific frequency component. The frequency components of all channels are concatenated to a vector F. After FC and Sigmoid, F is used to weighting the Feature map as modeling channel attention. Best viewed in color.

3.1 DCT Frequency on Feature Map

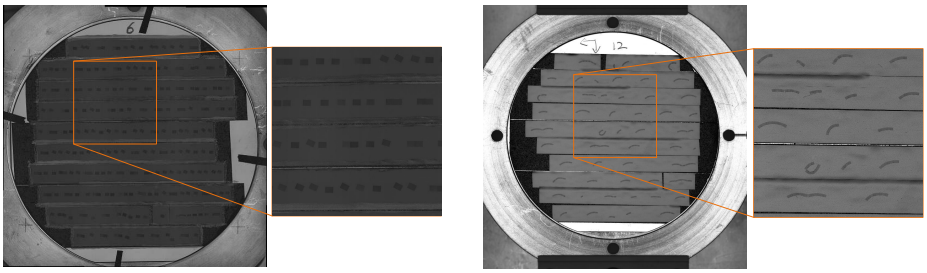
Generalized from one-dimensional (1D) DCT fefinition [1], two-dimensional (2D) DCT can be written as:

$$f_{h,w} = \sum_{u=0}^{H-1} \sum_{v=0}^{W-1} x_{u,v} \cos(\frac{(2u+1)h\pi}{2H}) \cdot \cos(\frac{(2v+1)w\pi}{2W}), \\ s.t. \quad h \in \{0, 1, ..., H-1\}, w \in \{0, 1, ..., W-1\}, \tag{1}$$

Where $f_{h,w}$ is the frequency component (h, w) on the 2D DCT frequency spectrum $f \in \mathbb{R}^{H \times W}$, $x_{u,v}$ is the pixel value on (u, v) of input $x \in \mathbb{R}^{H \times W}$,H and W are the height and width of x respectively. The inverserse 2D DCT can be written as:

$$x_{u,v} = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} f_{h,w} \cos(\frac{(2u+1)h\pi}{2H}) \cdot \cos(\frac{(2v+1)w\pi}{2W}), \\ s.t. \quad u \in \{0, 1, ..., H-1\}, v \in \{0, 1, ..., W-1\}, \tag{2}$$

Eqs. 1 and 2 remove some constant normalization factors, and this has no effect on the results in this work. As we can see in Eq. 2, every pixel value of



(a) (b)

Fig. 4. Wafer images in dataset. (a) The SCN region of the rat brain, (b) Zebrafish whole brain.

the input contains information from all the frequency components. If only use the lowest frequency component of input, h and w are 0 in Eq. 1, we have:

$$\begin{aligned} f_{0,0} &= \sum_{u=0}^{H-1}\sum_{v=0}^{W-1} x_{u,v}\cos(\frac{(2u+1)\cdot 0}{2H})\cdot\cos(\frac{(2v+1)\cdot 0}{2W}) \\ &= \sum_{u=0}^{H-1}\sum_{v=0}^{W-1} x_{u,v} \\ &= GAP(x)\cdot HW \end{aligned} \tag{3}$$

Thus, GAP, which conventional channel attention approaches [15,31] applies to calculate channel importance weight, only differ one constant scaling factor from the lowest frequency component of input. The information from other frequency components is discard. However, deep networks are redundant [14,33]. It is possible that different channels get the same information using GAP. Therefore, as the different frequency components contain different information, utilizing more frequency components may extract more information to model channel attention.

3.2 FANet

From the analysis illustrated in Sec 3.1, conventional channel attention only utilizes the lowest frequency component information of input, which is inadequate. Therefore, as shown in Fig 3, we propose the Frequency-Aware network (FANet), which generalizes the channel attention to more frequency components and models more information in channel attention.

Given a feature map as input $X \in \mathbb{R}^{C\times H\times W}$, we first divide it into n parts averagely along the channel dimension and every part is denoted as $X_i \in \mathbb{R}^{C'\times H\times W}, i \in \{0,1,...,n-1\}, C' = \frac{C}{n}$. In X_i, each channel of feature map are

Table 1. Specific frequency components.

(h_i, w_i)			
(0,0)	(0,1)	(6,0)	(0,5)
(0,2)	(1,0)	(1,2)	(4,0)
(5,0)	(1,6)	(3,0)	(0,4)
(0,6)	(0,3)	(3,5)	(2,2)

calculated a specific frequency component (h_i, w_i) , which can be written as :

$$f^j_{h_i,w_i} = \sum_{u=0}^{H-1}\sum_{v=0}^{W-1} X^j_i \cos(\frac{(2u+1)h_i\pi}{2H})\cos(\frac{(2v+1)w_i\pi}{2W}), \tag{4}$$

in which $X^j_i \in \mathbb{R}^{H\times W}, j \in \{0, 1, ..., C'-1\}$, is one of the channel of X_i. All the $f^j_{h_i,w_i} \in \mathbb{R}, j \in \{0, 1, ..., C'-1\}$ of X_i form a C'-dimensional vector $F_i \in \mathbb{R}^{C'}$. For the whole input X, All F_i are concatenated as F:

$$F = cat([F_0, F_1, ..., F_{n-1}]), \tag{5}$$

In which $F \in \mathbb{R}^C$ is the vector containing specific frequency components information of X. Finally, the whole channel importance weight for modeling channel attention of X is obtained as:

$$att = sigmoid(fc(F)), \tag{6}$$

In which $att \in \mathbb{R}^C$. Therefore, from Eqs. 5 and 6, utilizing att to weight X in the channel dimension introduces multiple frequency components information to model the channel attention of X.

As shown in Eq. 2, every pixel of the input contains all the frequency components information. FANet applies DCT to all the pixel of input, which can extract all information of one specific frequency component and model channel attention more efficiently and accurately. Besides, in view of the redundant information of feature map, the channels between feature map are likely to contain the same information. It is believable that each part contains much information of whole feature map. Therefore, obtaining a specific frequency component information from one part is sufficient, which reduces the complexity and size of the method.

4 Experiments

In this section, we first introduce the proposed dataset. Then elaborate on the details of our experiments on the dataset.

4.1 Dataset

To train and evaluate the models, we propose a dataset of wafer image as the benchmark. We choose two biological tissue sections which are suprachiasmatic

Table 2. Instance segmentation results of different methods on wafer image dataset.

Method	Backbone	Detector	AP	AP_{50}	AP_{75}	AP_s	AP_m	AP_L
ResNet [13]	ResNet-34 [13]	Mask R-CNN[12]	34.5	45.1	39.2	27.1	37.5	**28.3**
SENet [15]			38.0	**52.5**	41.1	29.7	42.6	24.2
FcaNet [25]			36.8	48.9	44.2	29.9	40.5	26.5
FANet(ours)			**38.1**	50.9	**45.0**	**36.8**	**43.2**	25.7
ResNet [13]	ResNet-50 [13]		37.7	49.4	**46.5**	26.1	41.6	**32.7**
SENet [15]			38.3	51.4	44.3	27.9	43.5	25.9
FcaNet [25]			38.6	50.1	44.8	**33.3**	44.5	25.3
FANet(ours)			**40.1**	**51.5**	45.7	27.4	**45.8**	24.6
ResNet [13]	ResNet-34 [13]	Mask Scoring R-CNN[16]	37.4	48.4	43.7	**35.0**	39.6	**31.7**
SENet [15]			38.2	50.4	44.9	34.4	43.5	26.1
FcaNet [25]			38.4	50.4	45.3	33.8	44.1	26.2
FANet(ours)			**39.7**	**52.8**	**46.2**	33.6	**45.0**	26.5
ResNet [13]	ResNet-50 [13]		38.9	50.4	47.4	31.2	44.0	24.2
SENet [15]			43.7	56.8	49.3	34.3	49.1	28.8
FcaNet [25]			44.9	58.2	**53.4**	**34.7**	51.6	28.6
FANet(ours)			**46.6**	**60.0**	52.9	33.7	**53.4**	**29.2**

nucleus (SCN) region of the rat brain and Zebrafish whole brain to improve the generalization and robustness of the model. As shown in Fig. 4, these two biological tissue sections have a big difference. After consulting a lot of experts, we identify three most common anomalistic situation of sections, which are incomplete, fold and damaged (Fig. 2). Therefore, there are four classes of sections in the dataset, which are good section and three kinds of bad section. Because of the original wafer images are much large, which are usually more than 10K×10K pixels, we crop the wafer images into sub-images (1024×1024) with overlap of 10%. Finally, the sub-images are randomly divided into training set and validation set. The training set contains 2708 sub-images and validation set is 1193, and both of them contain the two biological tissue sections. All experiments are conducted on the dataset as benchmark.

4.2 Implementation Details

We evaluate the performance of the proposed FANet with ResNet-34 and ResNet-50 [13] as backbone models. Experiments are implemented in mmdetection [7]. The data augmentation except image scale and hyper-parameter are followed [7]. The input image scale is 1024. The models are trained using stochastic gradient descent (SGD) with momentum of 0.9, a weight decay of 0.0001 and batch size of 8 per GPU. The training epoch is 36 and learning rate is 0.01 with warmup iters of 2000. The learning rate is reduced by a factor of 10 at epoch 24 and 34.

All models are trained with one Tesla P40 GPU.

4.3 The specific frequency components

Because of the limitations of experimental conditions, we follow the experiment results in [25]. The two-dimensional DCT frequency space is divided into 7×7 parts. Using one frequency component each time, it still has a competitive performance on ImageNet [27]. Finally, we choose 16 frequency components (1) with top-16 highest performance in the total 49 experiments results.

4.4 Instance Segmentation on the benchmark

To evaluate our method on the benchmark, we use Mask R-CNN[12] and Mask Scoring R-CNN [16]. FcaNet [25] and FANet are used the same 16 frequency components. As shown in table 2, our method has a superior performance. Especially compared with FcaNet, FANet could outperform it by 1.3% - 1.7% in terms of mAP, which indicates our method aggregates multiple frequency components more efficiently.

5 Conclusions

In this paper, we propose a robust channel attention mechanism using multi frequency components, named FANet, which is utilized to detect the tissue sections on wafer images. Besides, we propose a dataset of wafer images as benchmark and FANet has a superior performance. With satisfying performance, FANet can reduce much people and time to label the sections. In the future work, we will further collect wafer images of different biological tissue, evaluate FANet and improve it.

Acknowledgement

This research was funded by CAS Key Technology Talent Program (No. 292019000126 to X.C.)

References

1. Ahmed, N., Natarajan, T., Rao, K.R.: Discrete cosine transform. IEEE transactions on Computers **100**(1), 90–93 (1974) 5
2. Bogner, A., Jouneau, P.H., Thollet, G., Basset, D., Gauthier, C.: A history of scanning electron microscopy developments: Towards "wet-stem" imaging. Micron **38**(4), 390–401 (2007) 1
3. Bolya, D., Zhou, C., Xiao, F., Lee, Y.J.: Yolact: Real-time instance segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 9157–9166 (2019) 4
4. Briggman, K.L., Bock, D.D.: Volume electron microscopy for neuronal circuit reconstruction. Current opinion in neurobiology **22**(1), 154–161 (2012) 1

5. Cai, Z., Vasconcelos, N.: Cascade r-cnn: Delving into high quality object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 6154–6162 (2018) 4
6. Chen, K., Pang, J., Wang, J., Xiong, Y., Li, X., Sun, S., Feng, W., Liu, Z., Shi, J., Ouyang, W., et al.: Hybrid task cascade for instance segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4974–4983 (2019) 3, 4
7. Chen, K., Wang, J., Pang, J., Cao, Y., Xiong, Y., Li, X., Sun, S., Feng, W., Liu, Z., Xu, J., Zhang, Z., Cheng, D., Zhu, C., Cheng, T., Zhao, Q., Li, B., Lu, X., Zhu, R., Wu, Y., Dai, J., Wang, J., Shi, J., Ouyang, W., Loy, C.C., Lin, D.: MMDetection: Open mmlab detection toolbox and benchmark. arXiv preprint arXiv:1906.07155 (2019) 8
8. Dai, J., He, K., Li, Y., Ren, S., Sun, J.: Instance-sensitive fully convolutional networks. In: European Conference on Computer Vision. pp. 534–549. Springer (2016) 4
9. Denk, W., Briggman, K.L., Helmstaedter, M.: Structural neurobiology: missing link to a mechanistic understanding of neural computation. Nature Reviews Neuroscience **13**(5), 351–358 (2012) 1
10. Harris, K.M., Perry, E., Bourne, J., Feinberg, M., Ostroff, L., Hurlburt, J.: Uniform serial sectioning for transmission electron microscopy. Journal of Neuroscience **26**(47), 12101–12103 (2006) 1
11. Hayworth, K.J., Morgan, J.L., Schalek, R., Berger, D.R., Hildebrand, D.G., Lichtman, J.W.: Imaging atum ultrathin section libraries with wafermapper: a multi-scale approach to em reconstruction of neural circuits. Frontiers in neural circuits **8**, 68 (2014) 2
12. 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) 2, 8, 9
13. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016) 8
14. He, Y., Zhang, X., Sun, J.: Channel pruning for accelerating very deep neural networks. In: Proceedings of the IEEE international conference on computer vision. pp. 1389–1397 (2017) 6
15. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7132–7141 (2018) 4, 6, 8
16. Huang, Z., Huang, L., Gong, Y., Huang, C., Wang, X.: Mask scoring r-cnn. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 6409–6418 (2019) 3, 4, 8, 9
17. Joy, D.C.: The theory and practice of high-resolution scanning electron microscopy. Ultramicroscopy **37**(1-4), 216–233 (1991) 1
18. Kasthuri, N., Hayworth, K.J., Berger, D.R., Schalek, R.L., Conchello, J.A., Knowles-Barley, S., Lee, D., Vázquez-Reina, A., Kaynig, V., Jones, T.R., et al.: Saturated reconstruction of a volume of neocortex. Cell **162**(3), 648–661 (2015) 1
19. Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S.: Feature pyramid networks for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2117–2125 (2017) 4
20. Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L.: Microsoft coco: Common objects in context. In: European conference on computer vision. pp. 740–755. Springer (2014) 3

21. Liu, S., Qi, L., Qin, H., Shi, J., Jia, J.: Path aggregation network for instance segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 8759–8768 (2018) 3, 4
22. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3431–3440 (2015) 2
23. Micheva, K., Smith, S.: Array tomography: a new tool for imaging the molecular architecture and ultrastructure of neural circuits. Neuron **55**(1), 25–36 (2007) 1
24. Morgan, J.L., Berger, D.R., Wetzel, A.W., Lichtman, J.W.: The fuzzy logic of network connectivity in mouse visual thalamus. Cell **165**(1), 192–206 (2016) 1
25. Qin, Z., Zhang, P., Wu, F., Li, X.: Fcanet: Frequency channel attention networks. arXiv preprint arXiv:2012.11879 (2020) 3, 4, 8, 9
26. Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems **28**, 91–99 (2015) 2
27. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., et al.: Imagenet large scale visual recognition challenge. International journal of computer vision **115**(3), 211–252 (2015) 9
28. Shapson-Coe, A., Januszewski, M., Berger, D.R., Pope, A., Wu, Y., Blakely, T., Schalek, R.L., Li, P., Wang, S., Maitin-Shepard, J., et al.: A connectomic study of a petascale fragment of human cerebral cortex. bioRxiv (2021) 1, 2
29. Smith, S.J.: Q&a: array tomography. BMC biology **16**(1), 1–18 (2018) 1
30. Takemura, S.y., Bharioke, A., Lu, Z., Nern, A., Vitaladevuni, S., Rivlin, P.K., Katz, W.T., Olbris, D.J., Plaza, S.M., Winston, P., et al.: A visual motion detection circuit suggested by drosophila connectomics. Nature **500**(7461), 175–181 (2013) 1
31. Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., Hu, Q.: Eca-net: efficient channel attention for deep convolutional neural networks, 2020 ieee. In: CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE (2020) 4, 6
32. Yin, W., Brittain, D., Borseth, J., Scott, M.E., Williams, D., Perkins, J., Own, C.S., Murfitt, M., Torres, R.M., Kapner, D., et al.: A petascale automated imaging pipeline for mapping neuronal circuits with high-throughput transmission electron microscopy. Nature communications **11**(1), 1–12 (2020) 1, 2
33. Zhuang, Z., Tan, M., Zhuang, B., Liu, J., Guo, Y., Wu, Q., Huang, J., Zhu, J.: Discrimination-aware channel pruning for deep neural networks. arXiv preprint arXiv:1810.11809 (2018) 6