Class-Aware Feature Alignment for Domain Adaptative Mitochondria Segmentation

Dan Yin[1,2], Wei Huang[1,2], Zhiwei Xiong[1,2], and Xuejin Chen[1,2(✉)]

[1] National Engineering Laboratory for Brain-inspired Intelligence Technology and Application, University of Science and Technology of China, Hefei 230027, China
{ydaugust,weih527}@email.ustc.edu.cn, {zwxiong,xjchen99}@ustc.edu.cn
[2] Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 230088, China

Abstract. Unsupervised domain adaptation (UDA) has gained great popularity in mitochondria segmentation, aiming to improve the adaptability of models from the labeled source domain to the unlabeled target domain via domain alignment. However, existing UDA methods only focus on aligning domains on the prediction level, while ignoring the feature space containing more adequate information than the predictions. In this paper, we propose a class-aware domain adaptation method for mitochondria segmentation on the feature level, which relies on the prototype representation to achieve more fine-grained alignment. In particular, we first extract the feature centroids of classes from the source domain as prototypes. Leveraging the extracted prototypes as a bridge, we constrain that features belonging to the same class but from different domains are pulled closer to each other, achieving the class-aware alignment. Meanwhile, we derive a segmentation prediction directly from feature space based on the distance between target features and source prototypes. By incorporating a pseudo label to supervise the learning of this prediction, the feature distribution gap across domains is further reduced. Furthermore, to take full advantage of the potential of target domain, we propose an intra-domain consistency constraint to maintain consistent predictions of samples perturbed differently from the target image. Extensive experiments on different datasets demonstrate the superiority of our proposed method over existing UDA methods. Code is available at https://github.com/Danyin813/CAFA.

Keywords: Unsupervised domain adaptation · Class-aware alignment · Mitochondria segmentation

1 Introduction

Mitochondria segmentation from electron microscopy (EM) images is pivotal to mitochondria morphological analysis [15]. With pixel-wise annotations, existing supervised mitochondria segmentation methods [3,9,10,14] have achieved

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43901-8_23.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14223, pp. 238–248, 2023.
https://doi.org/10.1007/978-3-031-43901-8_23

extraordinary advances in the test data, when the training data and test data come from the same distribution. However, the data distribution of EM images varies in real scenarios due to the diversity in imaging devices, collected organisms and tissues. The different distributions between training and test data, *i.e.*, *domain shift* [13], lead to drastic performance drops on the test data. Manual annotations and model finetuning can ameliorate this problem but at a huge cost. Instead, unsupervised domain adaptive (UDA) mitochondria segmentation methods [2,4,7,11,16,17,21], aiming to transfer the knowledge learned in the labeled dataset (source domain) to unlabeled data (target domain) without any annotations, have gained great popularity in the community.

The mainstream of previous works focus on aligning the distributions of the source and target domains with supervision directly *on the output segmentation maps.* One line of these works use the model pretrained on source domain to obtain **pseudo labels** for target domain [1,21], the performance of which highly relies on the quality of the pseudo labels. Other methods are mainly based on GAN [7,16], where an additional **domain adversarial learning** task, designed to reduce the domain gap between source and target domains, is jointly optimized with segmentation task. The feature space, having higher dimension than the predictions, can express more adequate class-aware knowledge. However, these methods perform domain alignment on the output space, which has insufficient information compared with the feature space, hindering effective alignment.

In this work, to take full advantage of the sufficient class related information in the feature space, we propose a **class-aware domain alignment in the feature space** for mitochondria segmentation, which relies on the *prototype representation* [8,22] to achieve fine-grained feature alignment. Specifically, 1) we first extract the source feature centroid of each class as prototype. To make the prototypes represent source class knowledge better, we minimize the distance between prototype and its within-class source features, as well as push different prototypes away from each other. Also, we select *partial* target features close enough to source prototypes and minimize their distance to align domains at class level. 2) To further supervise *all* target features, we derive the closest prototype as predicted result for each target feature vector based on its distance to prototypes, resulting in a segmentation result directly from the feature space without the segmentation head. A pseudo label is utilized to supervise these predictions for further cross-domain alignment with class knowledge. 3) Though cross-domain alignment can introduce knowledge from source domain to the target domain, there still exists *additional* potential information useful to segmentation in the target domain [18]. Taking this into consideration, we further propose an intra-domain consistency constraint for target samples, where two input images perturbed differently from the same image are enforced to generate the same features and predictions.

Our contributions can be summarized as follows: 1) We propose a class-aware feature alignment method for domain adaptive mitochondria segmentation. To our best knowledge, it is the first attempt to align source and target domains on the feature level in UDA for EM mitochondria segmentation. 2) Our class-aware

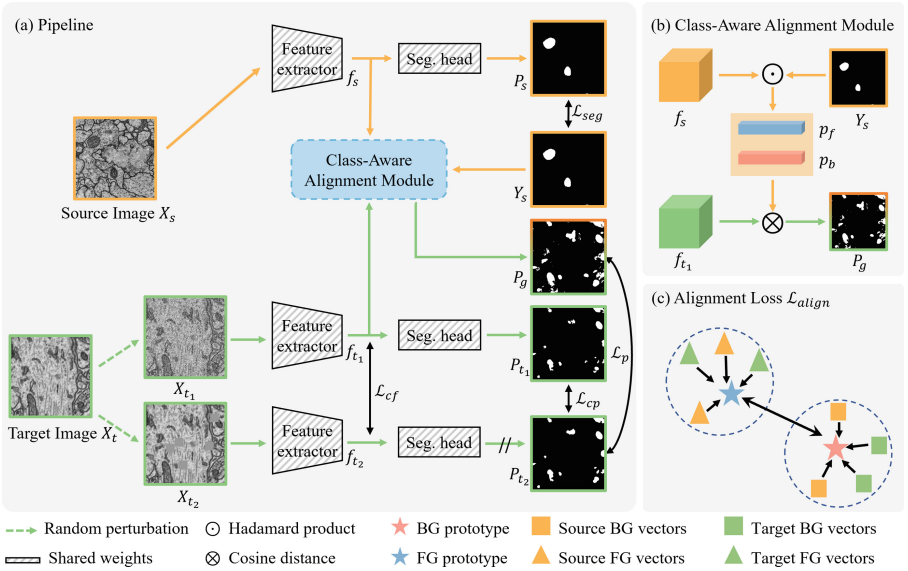

Fig. 1. The overall framework of our proposed method. (a) Images are first fed into the feature extractor to extract image features, which are then used to realize feature alignment by the class-aware alignment module and obtain the segmentation predictions by the segmentation (Seg.) head. We randomly perturb the target image X_t to obtain two augmented counterparts X_{t_1} and X_{t_2} with different augmentations. An intra-domain consistency constraint is incorporated in the feature level ($\mathcal{L}_{cf}$) and the prediction level ($\mathcal{L}_{cp}$). (b) In the Class-Aware Alignment Module, the source feature f_s and its corresponding groundtruth label are used to extract the centroids/prototypes (*i.e.*, p_f and p_b) of each class. Based on the distance between target features f_{t_1} and the prototypes, we can obtain the distance map P_g, representing the segmentation prediction directly from the features, which is further supervised by the pseudo label P_{t_2} ($\mathcal{L}_p$). (c) The illustration of the class-aware alignment loss $\mathcal{L}_{align}$, where the features within the same class are pulled together and the features belonging to different classes are pushed away from each other.

feature alignment relies on the source prototypes, which represent class knowledge from the feature space. With these prototypes, an innovative distance-based alignment and pseudo-labeling are incorporated to achieve class-aware feature alignment. 3) We propose an intra-domain consistency constraint in the target domain to tap into the potential target domain information. 4) We conduct thorough experiments on various EM dataset benchmarks and our proposed method achieves state-of-the-art performance for mitochondria segmentation.

2 Class-Aware Feature Alignment

Problem Formulation. Unsupervised domain adaptation (UDA) aims to transfer the knowledge learned from the labeled source domain to the unlabeled

target domain. In our work, we denote the source domain as $\mathcal{D}_S = \{X^S, Y^S\} = \{(x_i^s, y_i^s)\}_{i=1}^M$ with M samples, where y_i^s is the groundtruth binary segmentation map of the input image x_i^s. The unlabeled target domain is denoted as $\mathcal{D}_T = \{X^t\} = \{x_i^t\}_{i=1}^N$ with N samples. The overall framework of our proposed method is shown in Fig. 1.

Prototype Extraction. Considering there exists more plentiful class-aware information in the feature space than the predictions, we propose the class-aware alignment for better adaptation in the feature space. To achieve class-aware alignment, we first derive the class-aware source prototypes from the source features with the corresponding labels. The prototypes can be calculated as the centroid of each class in the feature space:

$$p_c^s = \frac{\sum_{b=1}^{B_s}\sum_{h=1}^{H_s}\sum_{w=1}^{W_s} f_{b,h,w}^s \mathbb{1}[Y_{b,h,w}^s = c]}{\sum_{b=1}^{B_s}\sum_{h=1}^{H_s}\sum_{w=1}^{W_s} \mathbb{1}[Y_{b,h,w}^s = c]}, \tag{1}$$

where $f_{b,h,w}^s \in \mathbb{R}$ is the source feature vectors, B_s is the batch size, and H_s, W_s is the height and width of the features. c is the index of class number C. The maximum of C is 1. The prototypes can represent the class knowledge in the source domain.

Inter- and Intra-class Constraints. To make the source prototypes represent the class-discriminative source knowledge more accurately, we incorporate inter- and intra-class constraints on the prototypes, which can further help better class-aware alignment across domains. The inter-class loss $\mathcal{L}_{inter}^s$ intends to push the prototypes of different classes far away from each other, which can be implemented by minimizing the average cosine distance of different prototype pairs:

$$\mathcal{L}_{inter}^s = \sum_{i=0}^{C}\sum_{j>i}^{C} \frac{p_i^s p_j^s}{|p_i^s||p_j^s|}. \tag{2}$$

In contrast, the intra-class loss $\mathcal{L}_{intra}^s$ is designed to pull the feature instance point closer to its corresponding prototype, *i.e.*, making the feature distribution of the same class more concentrated/compact. The intra-class loss can be formulated as maximizing the average cosine distance between the prototype and the features belonging to the same class:

$$\mathcal{L}_{intra}^s = 1 - \sum_{c=0}^{C} \frac{f_c^s p_c^s}{|f_c^s||p_c^s|}, \tag{3}$$

It is not straightforward to align target domain to source domain in the class level, considering the lack of groundtruth labels in target domain. To achieve more reliable class-aware alignment for target samples, we only perform alignment on the instances with higher confidence. Specifically, we first calculate the cosine distance between each target feature and all the source prototypes, and

only select instances $\{\tilde{f}^t\}$ the distance of which is closer than a preset threshold τ. The intra-class alignment loss enforces $\tilde{f}_c^t$ to be closer to its corresponding prototype p_c^t:

$$\mathcal{L}_{intra}^t = 1 - \sum_{c=0}^{C} \frac{\tilde{f}_c^t p_c^s}{|\tilde{f}_c^t||p_c^s|}, \tag{4}$$

The class-aware alignment loss $\mathcal{L}_{align}$ is the combination of these three losses, *i.e.*, $\mathcal{L}_{align} = \mathcal{L}_{intra}^s + \mathcal{L}_{inter}^s + \mathcal{L}_{intra}^t$. It is noteworthy that the alignment loss is optimized directly on the feature space instead of the final output predictions, considering there is more abundant information in the feature space.

Pseudo Supervision. The above mentioned alignment loss $\mathcal{L}_{align}$ only affects partial target features with higher confidence. To further force the alignment across domains in the feature space, we incorporate a pseudo supervision on the feature space. Specifically, based on the cosine distance between the feature of each location and the source prototypes, we can attain a distance map P_g, which can be regarded as a segmentation prediction directly from feature space instead of the prediction head. We utilize a pseudo label map P_{t2} as groundtruth to supervise the learning of P_g, leading to alignment directly on feature space. The formulation of P_{t2} will be discussed in the later section. The supervision is the standard cross entropy loss:

$$\mathcal{L}_p = CE(P_g, P_{t_2}). \tag{5}$$

Intra-domain Consistency. The alignment cross domains will borrow the knowledge from source domain to target domain. However, there exists abundant knowledge and information in the target domain itself [18]. To further exploit the sufficient knowledge existed in target domain, we propose an intra-domain consistency constraint in target domain. Specifically, for each target input image I_t, we first augment it with two different random augmentation strategies, resulting in I_{t1} and I_{t2}, which are then fed into the network for segmentation prediction. We incorporate two consistency losses on the feature level $\mathcal{L}_{cf}$ and the final prediction level $\mathcal{L}_{cp}$, respectively:

$$\mathcal{L}_{cf} = MSE(f_{t_1}, f_{t_2}), \quad \mathcal{L}_{cp} = CE(P_{t_1}, P_{t_2}), \tag{6}$$

where MSE denotes the standard mean squared error loss.

Training and Inference. During the training phase, the total training objective $\mathcal{L}_{total}$ is formulated as :

$$\mathcal{L}_{total} = \mathcal{L}_{seg}^s + \lambda_{align}\mathcal{L}_{align} + \lambda_p\mathcal{L}_p + \lambda_{cf}\mathcal{L}_{cf} + \lambda_{cp}\mathcal{L}_{cp}, \tag{7}$$

where $\mathcal{L}_{seg}^s$ denotes the supervised segmentation loss with the cross-entropy loss and $\lambda_{\{align,p,cf,cp\}}$ are the hyperparameters for balancing different terms. Note

Table 1. Quantitative comparisons on the Lucchi and MitoEM datasets. *Oracle* denotes the model is trained on target with groundtruth labels, while *NoAdapt* represents the model pretrained on source is directly applied in target for inference without any adaptation strategy. The results of Oracle, NoAdapt, UALR, DAMT-Net, DA-VSN and DA-ISC are adopted from [7].

Methods	VNC III → Lucchi (Subset1)				VNC III → Lucchi (Subset2)			
	mAP(%)	F1(%)	MCC(%)	IoU(%)	mAP(%)	F1(%)	MCC(%)	IoU(%)
Oracle	-	92.7	-	86.5	-	93.9	-	88.6
NoAdapt	-	57.3	-	40.3	-	61.3	-	44.3
Advent [19]	78.9	74.8	73.3	59.7	90.5	82.8	81.8	70.7
UALR [21]	80.2	72.5	71.2	57.0	87.2	78.8	77.7	65.2
DAMT-Net [16]	-	74.7	-	60.0	-	81.3	-	68.7
DA-VSN [6]	82.8	75.2	73.9	60.3	91.3	83.1	82.2	71.1
DA-ISC [7]	89.5	81.3	80.5	68.7	92.4	85.2	84.5	74.3
Ours	**91.1**	**83.4**	**82.8**	**71.8**	**94.8**	**85.8**	**85.4**	**75.4**
Methods	MitoEM-R → MitoEM-H				MitoEM-H → MitoEM-R			
	mAP(%)	F1(%)	MCC(%)	IoU(%)	mAP(%)	F1(%)	MCC(%)	IoU(%)
Oracle	97.0	91.6	91.2	84.5	98.2	93.2	92.9	87.3
NoAdapt	74.6	56.8	59.2	39.6	88.5	76.5	76.8	61.9
Advent [19]	89.7	82.0	81.3	69.6	93.5	85.4	84.8	74.6
UALR [21]	90.7	83.8	83.2	72.2	92.6	86.3	85.5	75.9
DAMT-Net [16]	92.1	84.4	83.7	73.0	94.8	86.0	85.7	75.4
DA-VSN [6]	91.6	83.3	82.6	71.4	94.5	86.7	86.3	76.5
DA-ISC [7]	92.6	85.6	84.9	74.8	**96.8**	88.5	88.3	79.4
Ours	**92.8**	**86.6**	**86.0**	**76.3**	**96.8**	**89.2**	**88.9**	**80.6**

that the feature extractor and the segmentation head are shared weights in the training phase. Their detailed structures can be found in the supplementary material. During the inference phase, we only adopt the trained feature extractor and segmentation head to predict the target images.

3 Experiments

Datasets. Following the previous work [7], our experiments involve four challenging EM datasets for domain adaptive mitochondria segmentation tasks, *i.e.*, VNC III [5]→ Lucchi (Subset1) [12], VNC III→ Lucchi (Subset2) [12], MitoEM-H [20] → MitoEM-R [20] and MitoEM-R → MitoEM-H. **VNC III** [5] is imaged from the Drosophila ventral nerve cord by ssTEM. The physical resolution of the pixel is $50 \times 5 \times 5$ nm^3. The dataset consists of 20 images, and their resolution is 1024×1024. **Lucchi** [12] is imaged from the hippocampus of mice collected by FIB-SEM. The physical resolution of the pixel is $5 \times 5 \times 5$ nm^3, the training (Subset1) and test (Subset2) sets both have 165 images with 1024×768 resolution. **MitoEM** [20] contains two image volumes imaged by mbSEM, one is from the

Table 2. Quantitative comparisons for the 3D instance segmentation results on the MitoEM dataset.

Methods	MitoEM-R → MitoEM-H		MitoEM-H → MitoEM-R	
	AP^{50}(%)	AP^{75}(%)	AP^{50}(%)	AP^{75}(%)
Advent [19]	43.6	17.3	56.4	27.0
UALR [21]	56.4	29.1	55.4	33.6
DAMT-Net [16]	55.2	29.5	54.6	25.5
DA-VSN [6]	53.1	24.6	60.2	29.3
DA-ISC [7]	60.0	37.4	63.0	34.7
Ours	**65.6**	**46.3**	**68.5**	**42.4**

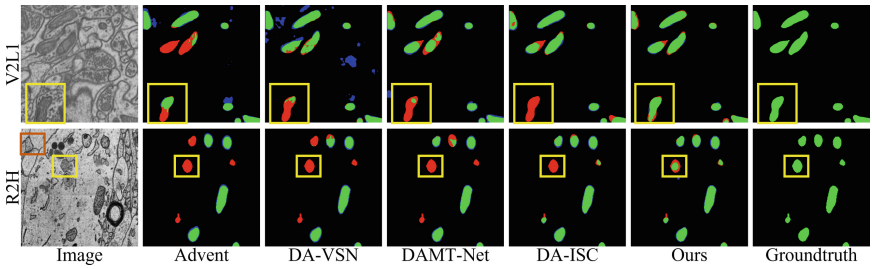

Fig. 2. Visual comparisons of different domain adaptive mitochondria segmentation methods, from VNC III to Lucchi (Subset1), *i.e.*, V2L1, and MitoEM-R to MitoEM-H, *i.e.*, R2H. The pixels in green, red and blue denote the true-positive, false-negative and false-positive segmentation results respectively. More 3D instance segmentation visualizations can be found in the supplementary material. (Color figure online)

rat cortex, named as MitoEM-R, and the other one is from the human cortex, named as MitoEM-H. Each volume contains 1000 images, where the instance-level annotations of the first 500 images are available. The physical resolution of the pixel is $30 \times 8 \times 8\,\mathrm{nm}^3$, and the image size is 4096×4096.

Evaluation Metrics. To thoroughly evaluate the performance of models, we conduct comparisons both on semantic-level and instance-level predictions. 1) Following [7], we compare different methods with mAP, F1, MCC and IoU scores on the 2D binary segmentation. 2) Considering that the quantity, size and morphology of mitochondria are pivotal to related studies, we further evaluate on the 3D instance segmentation task. Following [20], we take AP^{50} and AP^{75} as the metrics to quantitatively compare the performance of different methods.

Implementation Details. Our network architecture is following [7]. We crop each image into 512×512 as patch to input feature extractor. All models are trained using the Adam optimizer with $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The learning

Table 3. Ablation results for the effectiveness of each loss term.

MitoEM-R → MitoEM-H									
Settings	$\mathcal{L}^s_{seg}$	$\mathcal{L}_{align}$	$\mathcal{L}_p$	$\mathcal{L}_{cf}$	$\mathcal{L}_{cp}$	mAP(%)	F1(%)	MCC(%)	IoU(%)
①	✓					87.2	78.7	78.1	64.9
②	✓	✓				88.8	83.3	82.8	71.5
③	✓	✓	✓			91.1	84.2	83.6	72.7
④	✓	✓	✓	✓		91.5	84.2	85.4	75.4
⑤	✓	✓	✓	✓	✓	**92.8**	**86.6**	**86.0**	**76.3**

rate is set at $1e-4$ and is polynomially decayed with a power of 0.9. We train models for 200k iterations in total. The balancing weights λ_{align}, λ_{proto}, λ_{cf}, and λ_{cp} in Eq. 7 are set as 0.1, 0.1, 0.1, and 0.1, respectively. The preset threshold τ is set as 0.7. To obtain 3D instance segmentation results, we adopt the marker-controlled watershed algorithm [20] on the predicted binary predictions.

Comparisons with Baselines. The binary segmentation result comparisons of our proposed method with previous works on the Lucchi and MitoEM datasets are shown in Table 1. The competitors include UALR [21], DAMT-Net [16], Advent [19], DA-VSN [6], and DA-ISC [7]. Our method achieves the new state-of-the-art results in all cases, which corroborates the superiority of the proposed class-aware alignment in the feature space. Especially, compared with the previous state-of-the-art DA-ISC [7], our method surpasses it by a large margin on the benchmarks of VNC III→ Lucchi (Subset1) (3.1% IoU). The mitochondria in MitoEM-H distribute more densely and are more complex than those in MitoEM-R, leading to the result of MitoEM-R → MitoEM-H is lower than that of MitoEM-H → MitoEM-R. However, our result on MitoEM-R → MitoEM-H has remarkable improvement, owing to that our method not only aligns domain in a fine-grained way but also explore the full potential of target.

As shown in Table 2, we also evaluate the effectiveness of our proposed method on 3D instance segmentation results. We only conduct experiments on the MitoEM dataset due to the lack of groundtruth for 3D instance segmentation in Lucchi. Our method not only deals with the domain adaptation for binary segmentation but also behaves well for the harder 3D instance segmentation, where the latter has rarely been studied in the literature. Furthermore, to further evaluate the effectiveness of our method, we visualize the predicted segmentation of our method and baselines in Fig. 2. Credited to the proposed class-aware feature alignment, our method estimates more fine-grained results on the target domain, substantiating that our method can alleviate the domain shift between source and target domains. In Fig. 2, the yellow and orange boxes represent mitochondria and background, respectively. In R2H, only our method segments the mitochondria in the yellow box. This is because our method is able to extract more representative features of mitochondria and background, and can separate these two types of features more effectively.

Ablation Study for Loss Functions. We conduct thorough ablation experiments to validate the contribution of each loss term in Eq. 7, where the results are shown in Table 3. The experiment ① with only supervised segmentation loss $\mathcal{L}^s_{seg}$ is the *Source-Only* method. All other variants with our proposed losses have superior performance than ①. Specifically, with the proposed class-aware feature alignment ($\mathcal{L}_{align}$ and $\mathcal{L}_p$), ② improves ① by 6.6% IoU and ③ further enlarges the gain to 7.8%. To take advantage of the potential information in target domain with the *intra-domain* losses $\mathcal{L}_{cf}$ and $\mathcal{L}_{cp}$, the final ⑤ improves by 3.6% IoU, leading to the 11.4% IoU improvement in total. To further explore the impact of different components in $\mathcal{L}_{align}$, *i.e.*, $\mathcal{L}^s_{intra}$, $\mathcal{L}^s_{inter}$ and $\mathcal{L}^t_{intra}$, we conduct experiments and the results are shown in the supplementary material. We find that the cross-domain alignment term $\mathcal{L}^t_{intra}$ plays the core role. It is noteworthy that $\mathcal{L}^s_{intra}$ and $\mathcal{L}^s_{inter}$ help to learn better class-aware prototypes.

4 Conclusion

In this paper, for the first time, we propose the class-aware alignment for domain adaptation on mitochondria segmentation in the feature space. Based on the extracted source prototypes representing class knowledge, we design intra-domain and inter-domain alignment constraint for fine-grained alignment cross domains. Furthermore, we incorporate an intra-domain consistency loss to take full advantage of the potential information existed in target domain. Comprehensive experiments demonstrate the effectiveness of our proposed method.

Acknowledgement. This work was supported by the National Natural Science Foundation of China under Grant 62076230.

References

1. Bermúdez-Chacón, R., Altingövde, O., Becker, C., Salzmann, M., Fua, P.: Visual correspondences for unsupervised domain adaptation on electron microscopy images. IEEE Trans. Med. Imaging **39**(4), 1256–1267 (2019)
2. Bermúdez-Chacón, R., Márquez-Neila, P., Salzmann, M., Fua, P.: A domain-adaptive two-stream u-net for electron microscopy image segmentation. In: ISBI (2018)
3. Franco-Barranco, D., Muñoz-Barrutia, A., Arganda-Carreras, I.: Stable deep neural network architectures for mitochondria segmentation on electron microscopy volumes. Neuroinformatics **20**(2), 437–450 (2022)

4. Franco-Barranco, D., Pastor-Tronch, J., González-Marfil, A., Muñoz-Barrutia, A., Arganda-Carreras, I.: Deep learning based domain adaptation for mitochondria segmentation on EM volumes. Comput. Methods Programs Biomed. **222**, 106949 (2022)
5. Gerhard, S., Funke, J., Martel, J., Cardona, A., Fetter, R.: Segmented anisotropic ssTEM dataset of neural tissue. Figshare (2013)
6. Guan, D., Huang, J., Xiao, A., Lu, S.: Domain adaptive video segmentation via temporal consistency regularization. In: ICCV (2021)
7. Huang, W., Liu, X., Cheng, Z., Zhang, Y., Xiong, Z.: Domain adaptive mitochondria segmentation via enforcing inter-section consistency. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) MICCAI 2022. LNCS, vol. 13434, pp. 89–98. Springer, Cham (2022)
8. Jiang, Z., et al.: Prototypical contrast adaptation for domain adaptive semantic segmentation. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M., Hassner, T. (eds.) ECCV 2022. LNCS, vol. 13694, pp. 36–54. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-19830-4_3
9. Li, M., Chen, C., Liu, X., Huang, W., Zhang, Y., Xiong, Z.: Advanced deep networks for 3d mitochondria instance segmentation. In: 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI), pp. 1–5. IEEE (2022)
10. Li, Z., Chen, X., Zhao, J., Xiong, Z.: Contrastive learning for mitochondria segmentation. In: EMBC (2021)
11. Liu, D., et al.: PDAM: a panoptic-level feature alignment framework for unsupervised domain adaptive instance segmentation in microscopy images. IEEE Trans. Med. Imaging **40**(1), 154–165 (2020)
12. Lucchi, A., Li, Y., Fua, P.: Learning for structured prediction using approximate subgradient descent with working sets. In: CVPR (2013)
13. Luo, Y., Zheng, L., Guan, T., Yu, J., Yang, Y.: Taking a closer look at domain shift: category-level adversaries for semantics consistent domain adaptation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2507–2516 (2019)
14. Mekuč, M.Ž, Bohak, C., Hudoklin, S., Kim, B.H., Kim, M.Y., Marolt, M., et al.: Automatic segmentation of mitochondria and endolysosomes in volumetric electron microscopy data. Comput. Biol. Med. **119**, 103693 (2020)
15. Mumcuoglu, E., Hassanpour, R., Tasel, S., Perkins, G., Martone, M., Gurcan, M.: Computerized detection and segmentation of mitochondria on electron microscope images. J. Microsc. **246**(3), 248–265 (2012)
16. Peng, J., Yi, J., Yuan, Z.: Unsupervised mitochondria segmentation in EM images via domain adaptive multi-task learning. IEEE J. Sel. Top. Signal Process. **14**(6), 1199–1209 (2020)
17. Roels, J., Hennies, J., Saeys, Y., Philips, W., Kreshuk, A.: Domain adaptive segmentation in volume electron microscopy imaging. In: ISBI (2019)
18. Sohn, K., et al.: Fixmatch: simplifying semi-supervised learning with consistency and confidence. Adv. Neural. Inf. Process. Syst. **33**, 596–608 (2020)
19. Vu, T.H., Jain, H., Bucher, M., Cord, M., Pérez, P.: Advent: adversarial entropy minimization for domain adaptation in semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2517–2526 (2019)
20. Wei, D., et al.: MitoEM dataset: large-scale 3D mitochondria instance segmentation from EM images. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12265, pp. 66–76. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59722-1_7

21. Wu, S., Chen, C., Xiong, Z., Chen, X., Sun, X.: Uncertainty-aware label rectification for domain adaptive mitochondria segmentation. In: de Bruijne, M., Cattin, P.C., Cotin, S., Padoy, N., Speidel, S., Zheng, Y., Essert, C. (eds.) MICCAI 2021. LNCS, vol. 12903, pp. 191–200. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87199-4_18
22. Zhang, P., Zhang, B., Zhang, T., Chen, D., Wang, Y., Wen, F.: Prototypical pseudo label denoising and target structure learning for domain adaptive semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12414–12424 (2021)