

FEW-SHOT ADAPTATION FOR MORPHOLOGY-INDEPENDENT CELL INSTANCE SEGMENTATION

Ram J. Zaveri* Voke Brume* Gianfranco Doretto

Lane Department of Computer Science and Electrical Engineering
West Virginia University, Morgantown, WV 26506 USA

ABSTRACT

Microscopy data collections are becoming larger and more frequent. Accurate and precise quantitative analysis tools like cell instance segmentation are necessary to benefit from them. This is challenging due to the variability in the data, which requires retraining the segmentation model to maintain high accuracy on new collections. This is needed especially for segmenting cells with elongated and non-convex morphology like bacteria. We propose to reduce the amount of annotation and computing power needed for retraining the model by introducing a few-shot domain adaptation approach that requires annotating only one to five cells of the new data to process and that quickly adapts the model to maintain high accuracy. Our results show a significant boost in accuracy after adaptation to very challenging bacteria datasets.

Index Terms— Domain adaptation, cell instance segmentation, few-shot learning, microscopy.

1. INTRODUCTION

The increased frequency and size of microscopy data collections enable rapid scientific discovery but require improved automated tools for quantitative analysis [1]. Image collections present a remarkable variability, depicting cells, membranes, and organelles at various sizes and shapes, collected with different microscope configurations, and from tissues treated differently. In particular, images of bacteria cells present additional challenges due to their smaller size and elongated and non-convex morphologies, including shapes with multiple branches [2].

The typical approach to tasks such as automated cell instance segmentation on these collections is to rely on a method trained on a large *source* dataset [3, 4]. However, because of their variability, new *target* image collections will very likely have different statistical properties than those used for training the segmentation method. Because of this *distribution shift* [5], the segmentation ac-

curacy will deteriorate on new collections. The obvious solution is to retrain the segmentation model on the target data. This is effective for methods like [6, 7, 3] for handling more regular, nearly-convex cell morphologies, whereas for more difficult morphologies (e.g., for bacteria), methods like [8, 4] will be more appropriate. The problem is that this approach requires a labor-intensive data annotation process and a computationally expensive training procedure.

A different solution is to use a *domain adaptation* approach [9], where a portion of the new *target* data is used for “adapting” the model to perform well in spite of the distribution shift. Current approaches are either modality-dependent and not tailored to cell segmentation [10, 11, 12, 13, 14] or computationally demanding because they are unsupervised and need to process large amounts of data [15, 16, 17, 18]. Recently, CellTranspose, a few-shot domain adaptation approach for cell instance segmentation, was proposed [19]. It has the advantage of requiring only a few cells (e.g., one to five) to be annotated by the user. Then, the model is quickly adapted to the target data domain with minimal computational cost. Despite offering an appealing solution, CellTranspose’s ability to handle elongated and non-convex morphologies is limited by the adopted segmentation strategy. This is based on predicting segmentation masks along with gradient flows converging towards the cell centers [3]. This approach has been shown to have limitations with morphologies that are not nearly-convex and for which a global cell center cannot be correctly defined [4].

In this work, we propose a few-shot domain adaptation approach for cell instance segmentation that addresses the limitations expressed above. We do so by starting from a segmentation strategy that handles elongated and non-convex morphologies by predicting segmentation boundaries and a boundary distance field along with its gradient. This requires designing new training losses to support the fast adaptation of a pre-trained model to the target domain with just a few (one to five) annotated cells. The results show that our framework effectively supports adaptation to multiple datasets

*Denotes equal contribution.

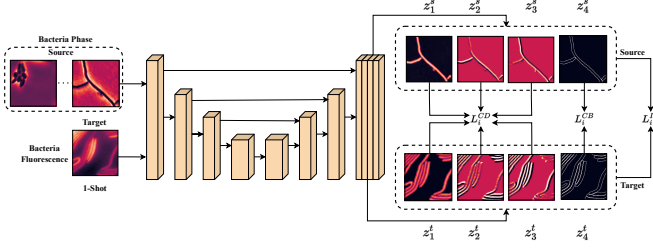


Fig. 1: Architecture approach. Illustration of our contrastive learning-based few-shot cellular instance segmentation approach.

in the case of challenging non-convex morphologies.

2. METHOD

We are interested in segmenting instances of cells, like bacteria, which can have small dimensions (e.g., a hundred of pixels in area), and can exhibit a wide range of morphologies [20]. Also, new image collections are deeply affected by the different optical features determined by the tissue treatment or microscopy techniques. This means that given new data to segment, a state-of-the-art tool like Omnipose [4] will be affected by the distribution shift [5], if not retrained on such new data, and it will significantly underperform.

To avoid the costly process of annotating large amounts of new data and intense network training for every new segmentation task, we propose a two-step approach. First, we rely on a robust and precise segmentation method that can handle a diversity of morphologies, and that has been pre-trained on a *source* dataset \mathcal{D}^s . Second, we take a *few-shot* learning approach, whereby given a *target* dataset \mathcal{D}^t , distributed differently than \mathcal{D}^s , we require the user to label only a *minimal* amount of target data, let us say K image patches for a K -shot learning, essentially depicting one cell each. The pre-training step needs to be done only once. The *few-shot adaptation* step is fast and low-cost both in manual labor and compute power, and is easily repeatable for new datasets to process.

Note that mere model fine-tuning on the K image patches is not a viable option since this would lead to overfitting [21]. Hence, we introduce a proper few-shot supervised domain adaptation method for cell instance segmentation. This work overcomes the limitations of previous work [19], which could not handle the adaptation to datasets with a large diversity of morphologies. Thus, to the best of our knowledge, this is the first few-shot adaptation for cell instance segmentation that is morphology-independent, meaning it can handle cases where a cell center cannot be correctly defined, which typically occur with shapes with elongated and/or multiple branches.

2.1. Pretrained Model

We rely on Omnipose [4] as the pretrained method because it is the state-of-the-art for handling morphology-independent cell instance segmentation. We summarize it here to introduce notation. Given an image I , a network f produces a dense, pixel-wise feature $\mathbf{Z} = f(I)$, where $\mathbf{Z} = [Z_1, Z_2, Z_3, Z_4] \in \mathbb{R}^{h \times w \times 4}$. For a pixel i , the feature $\mathbf{z} = [z_1, z_2, z_3, z_4] \in \mathbf{Z}$, has the following meaning. $z_1 \doteq \phi$, where ϕ is the *distance field*, which is the distance between pixel i inside a cell region to the closest point on the cell boundary. $\phi = 0$ if pixel i is outside of a cell. Then, $\mathbf{u} \doteq (z_2, z_3)$ is the *gradient flow field*, which is the normalized gradient of the distance field, pointing towards the medial axis (skeleton) of the cell that is defined by the stationary points of the distance field. Finally, $z \doteq z_4$ represents the unnormalized score indicating the probability of pixel i belonging to a cell boundary. With this notation, we rewrite the feature as $\mathbf{z} = [\phi, \mathbf{u}, z]$. The network f is trained in a supervised manner with pixel-wise instance segmentation loss

$$\mathcal{L}_i^{IS} = (\phi - \mathbf{d})^2 + \nu \|\mathbf{u} - (\mathbf{g}_x, \mathbf{g}_y)\|^2 + \mu H(\mathbf{b}, \sigma(z)), \quad (1)$$

where, for pixel i , \mathbf{d} is the ground-truth distance field, $(\mathbf{g}_x, \mathbf{g}_y)$ is the ground-truth gradient flow field with unit ℓ_2 -norm, $\mathbf{b} \in \{0, 1\}$ is the binary mask label indicating absence/presence of a cell border, $\sigma(z) \doteq 1/(1 + \exp(-z))$, H is the binary cross-entropy. ν and μ are hyperparameters set to 0.5 and 1.0. For image I , the contributions $\{\mathcal{L}_i^{IS}\}$ are aggregated into the loss $\mathcal{L}^{IS} = \sum \mathcal{L}_i^{IS} + \mathcal{L}_{IVP}$, where \mathcal{L}_{IVP} is a loss term that takes into account the initial value problem satisfied by the distance field [4].

From \mathbf{Z} , a segmentation head g produces the mask $Y = g(\mathbf{Z})$ based on an Euler integration [4]. For pixel i , the predicted label is $y \in \{0, 1, \dots, N\}$. N is the total number of cell instances segmented. $y = 0$ indicates absence of a cell. See also Figure 1.

2.2. Adaptation of the Pretrained Model

A cell target pixel i of $I \in \mathcal{D}^t$, with label $(\mathbf{d}^t, \mathbf{g}_x^t, \mathbf{g}_y^t, \mathbf{b}^t)$, should have its feature \mathbf{z}^t close to the pixel features from source images in \mathcal{D}^s with the same label. However, this is generally not the case because of the *shift* between source and target domains. Therefore, we design new training losses to reverse the effects of domain shift, which are tailored specifically to the network f introduced in §2.1, and that can adapt f to the target to prevent performance deterioration of the segmentation process.

Contrastive Distance Loss. We jointly align ϕ^t and \mathbf{u}^t with the distance and its gradient features of source pixels by setting up a contrastive prediction task [22]. This differs from [19] where the alignment was needed only for the gradient features.

To that end, we identify a *positive* source pixel inside the boundaries defined by \mathbf{b}_+^s , with distance ϕ_+^s and gradient features \mathbf{u}_+^s that best match the label $(\mathbf{d}^t, \mathbf{g}_x^t, \mathbf{g}_y^t)$, according to a similarity measure. We use $s((\phi, \mathbf{u}), (\psi, \mathbf{v})) \doteq \exp[\frac{-1}{2\sigma}(\phi - \psi)^2] \mathbf{u}^\top \mathbf{v} / \|\mathbf{u}\| \|\mathbf{v}\|$, which combines a radial basis kernel with parameter σ to compare the distance features, and a cosine similarity kernel to compare the gradients, where $\|\cdot\|$ denotes ℓ_2 -norm. We also compose the set of *negative* features \mathcal{N}_i , taken from inside the boundaries \mathbf{b}_-^s , and such that $\mathcal{N}_i = \{(\phi_-^s, \mathbf{u}_-^s) \mid s((\phi_+^s, \mathbf{u}_+^s), (\mathbf{u}_-^s, \phi_-^s)) < \delta\}$, where δ is a threshold that we choose. This allows to define the following *contrastive distance loss*

$$\mathcal{L}_i^{CD} = -\log \frac{e^{\frac{s((\phi^t, \mathbf{u}^t), (\phi_+^s, \mathbf{u}_+^s))}{\tau}}}{e^{\frac{s((\phi^t, \mathbf{u}^t), (\phi_+^s, \mathbf{u}_+^s))}{\tau}} + \sum_{(\phi_-^s, \mathbf{u}_-^s) \in \mathcal{N}_i} e^{\frac{s((\phi^t, \mathbf{u}^t), (\phi_-^s, \mathbf{u}_-^s))}{\tau}}} \quad (2)$$

where τ is a temperature parameter. \mathcal{L}_i^{CD} pulls the positive pair $((\phi^t, \mathbf{u}^t), (\phi_+^s, \mathbf{u}_+^s))$ closer, while pushing apart every negative pair $((\phi^t, \mathbf{u}^t), (\phi_-^s, \mathbf{u}_-^s))$. We also extend the strategy of [19] for mining hard negatives for \mathcal{N}_i .

For a target image and a source image pair we aggregate the loss contributions from the target pixels inside the boundaries \mathbf{b}^t . If \mathcal{B} indicates this set of pixels, then the aggregate loss is $\mathcal{L}^{CD} = \frac{1}{|\mathcal{B}|} \sum_{\mathcal{B}} \mathcal{L}_i^{CD}$.

Contrastive Boundary Loss. We align the unnormalized binary classification score z^t with the scores of the source pixels with same label. We derive a contrastive loss for boundary detection by treating this as a binary classification adaptation problem [9]. This differs from [19] because there they need to focus on cell masks rather than boundaries.

We want to pull together z^t and the unnormalized score z_+^s , of pixels in source images with label $\mathbf{b}^s = \mathbf{b}^t$. We also need to minimize the similarity between z^t and the unnormalized score z_-^s , of pixels in source images with label $\mathbf{b}^s \neq \mathbf{b}^t$. This is achieved by minimizing this *contrastive boundary loss*

$$\mathcal{L}^{CB} = \frac{1}{|\mathcal{P}|} \sum_{\mathcal{P}} d(z^t, z_+^s) + \lambda \frac{1}{|\mathcal{N}|} \sum_{\mathcal{N}} k(z^t, z_-^s). \quad (3)$$

In (3), for a target image and a source image pairs, \mathcal{P} is the set of positive pairs $\{(z^t, z_+^s)\}$, and \mathcal{N} is the set of negative pairs $\{(z^t, z_-^s)\}$. In addition, $d(z^t, z_+^s) = \frac{1}{2}(z^t - z_+^s)^2$, $k(z^t, z_-^s) = \frac{1}{2} \max(0, m - |z^t - z_-^s|)^2$, m is a margin, and λ is a hyperparameter.

Few-shot Adaptation. Let \mathcal{D}_K^t indicate the K labeled samples of the target dataset \mathcal{D}^t . The *K-shot adaptation* training minimizes

$$\mathcal{L}^{ISA} = \sum_{\mathcal{D}_K^t} \left(\mathcal{L}^{IS} + \frac{\gamma_1}{|\mathcal{D}^s|} \sum_{\mathcal{D}^s} \mathcal{L}^{CB} + \frac{\gamma_2}{|\mathcal{D}^s|} \sum_{\mathcal{D}^s} \mathcal{L}^{CD} \right). \quad (4)$$

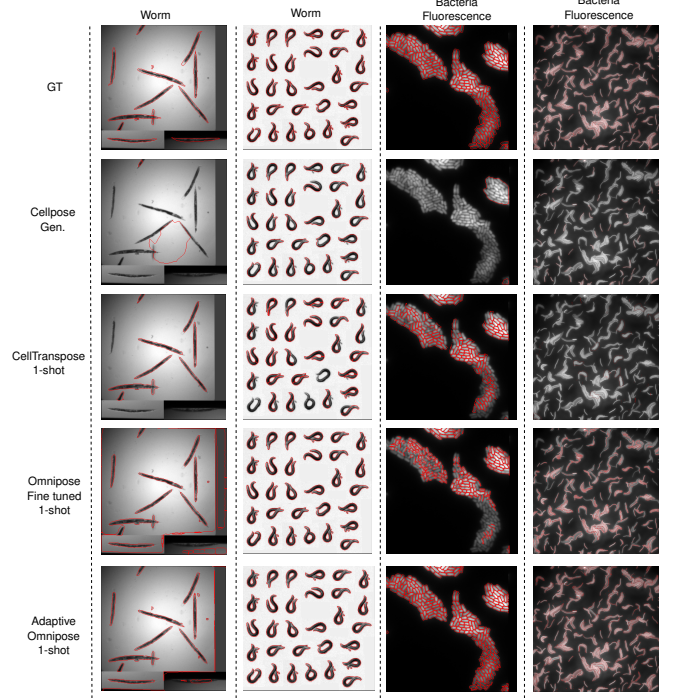


Fig. 2: Qualitative results. Samples from the Worm (left two columns) and Bacteria Fluorescence (right two columns) datasets highlighting the variability of cell morphologies, and the corresponding segmentations by different approaches. GT indicates ground-truth.

The training assumes that the model has already been pretrained. We used Omnipose pretrained for 3800 epochs, and we adapt the model for 5 epochs according to (4). Source images are continuously randomly paired with one of the K target samples without replacement to ensure that even a 1-shot adaptation performs a significant pull of the model towards the target distribution. After adaptation the model is fine-tuned on \mathcal{D}_K^t with \mathcal{L}^{IS} for 5 more epochs with a very low learning rate of $1e-7$. For the definition of one “shot” we followed [19], where the nominal cell size is computed according to [4].

3. EXPERIMENTS

Implementation. We evaluate our approach, named *Adaptive Omnipose*, with different number of shots, $K = 1, 2, 3, 5, 10$. We calculate the average precision (AP) at an intersection over union (IoU) of 0.5. As source datasets \mathcal{D}^s we use separately Bacteria Phase and Bacteria Fluorescence, which are the largest and with the most diverse morphologies among those released by [4].

We pretrain our model for 3800 epochs with constant learning rate of 0.03, and the RAdam optimizer with weight decay 10^{-5} and batch size 4. Adaptation is done by following the data splitting guidelines from [19]. We randomly sample K shot samples from the respective

Table 1: AP results at IoU 0.5: Adapting from Bacteria Phase (BP) to Bacteria Fluorescence (BF) and Worm and from Bacteria Fluorescence (BF) to Bacteria Phase (BP) and Worm.

$\mathcal{D}^s \rightarrow \mathcal{D}^t$	BP \rightarrow BF					BP \rightarrow Worm					BF \rightarrow BP					BF \rightarrow Worm				
	1-shot	2-shot	3-shot	5-shot	10-shot	1-shot	2-shot	3-shot	5-shot	10-shot	1-shot	2-shot	3-shot	5-shot	10-shot	1-shot	2-shot	3-shot	5-shot	10-shot
Cellpose-Gen	0.137	0.137	0.137	0.137	0.137	0.173	0.173	0.173	0.173	0.173	0.335	0.335	0.335	0.335	0.335	0.173	0.173	0.173	0.173	0.173
Omnipose-LB	0.009	0.009	0.009	0.009	0.009	0.470	0.470	0.470	0.470	0.470	0.001	0.001	0.001	0.001	0.001	0.002	0.002	0.002	0.002	0.002
Omnipose-FT	0.431	0.511	0.504	0.512	0.578	0.515	0.501	0.511	0.517	0.554	0.238	0.364	0.384	0.375	0.446	0.483	0.501	0.500	0.511	0.547
Omnipose-UB	0.920	0.920	0.920	0.920	0.920	0.847	0.847	0.847	0.847	0.847	0.889	0.889	0.889	0.889	0.889	0.847	0.847	0.847	0.847	0.847
Cellpose-UB	0.927	0.927	0.927	0.927	0.927	0.792	0.792	0.792	0.792	0.792	0.717	0.717	0.717	0.717	0.717	0.792	0.792	0.792	0.792	0.792
Stardist-UB	0.376	0.376	0.376	0.376	0.376	0.326	0.326	0.326	0.326	0.326	0.341	0.341	0.341	0.341	0.341	0.326	0.326	0.326	0.326	0.326
CellTranspose	0.375	0.493	0.516	0.529	0.613	0.531	0.595	0.622	0.645	0.650	0.142	0.181	0.263	0.359	0.403	0.486	0.539	0.520	0.548	0.564
Adaptive Omnipose	0.594	0.601	0.603	0.623	0.642	0.592	0.638	0.642	0.645	0.656	0.409	0.428	0.447	0.461	0.471	0.557	0.567	0.581	0.596	0.585

Table 2: Ablation on 1-shot adaptation. AP at 0.5 IoU.

	BP \rightarrow BF	BP \rightarrow Worm
Omnipose-UB	0.920	0.847
Adaptive Omnipose	0.594	0.592
No \mathcal{L}^{CB}	0.525	0.561
No \mathcal{L}^{CD}	0.569	0.553
No \mathcal{L}^{CB} & No \mathcal{L}^{CD}	0.431	0.515

training set of the target dataset \mathcal{D}^t . Hyperparameters for the adaptation losses include $|\mathcal{N}_i| = 20$, $\tau = 0.1$, $m = 10$, $\gamma_1 = 0.05$, and $\gamma_2 = 0.05$. We used an image patch size of $h = w = 112$ without rescaling since most cellular metrics such as mean diameter and area were consistent among all dataset groups during training. However, we rescale target images using the ratio of mean diameter across the source dataset and mean diameter across the target. Also, we note that the original Omnipose [4] uses a patch size of 224×224 , whereas our implementation of all the algorithms uses 112×112 . The adaptation procedure takes approximately 3 minutes to complete using an NVIDIA RTX 3090 GPU.

Bacteria Phase (BP). BP [4] is a large bacterial phase-contrast dataset consisting of bacterial mutations and treatments causing extreme morphological variations. Also, it includes various wild-type and mutant bacterial growth under cephalixin and aztreonam. It has a total of 27,500 cells in the training set, and 19,500 in the test set. We use the same split as in the original paper to pretrain the base model.

Bacteria Fluorescence (BF). The BF [4] dataset involves images with cytosol and membrane fluorescence, which has a different modality compared to BP. It consists of 33,200 cells split in train (18,613 cells) and test (14,587 cells) sets.

Worm. Worm [4] is a non-bacterial dataset consisting of 1,279 worms in the train set and 1,264 worms in the test set. It was acquired from the Open Worm Movement database [23], and the dead *C. elegans* from the BBBC010 dataset [24].

Compared methods. Cellpose-Gen is Cellpose trained on the generalist dataset of [3]. Omnipose-LB is Omnipose trained on \mathcal{D}^s (BP or BF) and represents a

lower bound. Omnipose-FT is Omnipose trained on \mathcal{D}^s and fine-tuned on \mathcal{D}_K^t , i.e., with K -shot samples, data augmentation and slow learning rate. Stardist-UB [6], Cellpose-UB and Omnipose-UB are trained and tested on the respective training and testing splits of the target datasets and represent upper bounds. CellTranspose [19] and Adaptive Omnipose are pretrained on \mathcal{D}^s and then adapted with \mathcal{D}_K^t . See Figure 2 for qualitative results.

BP \rightarrow BF and BF \rightarrow BP. The modality difference between BP and BF adds to the inherent distribution shift: BF images look almost as negative images of BP [4]. This is reflected in Table 1, where Omnipose-LB shows catastrophic performance degradation. Adaptive Omnipose steadily improves, and is very effective even with 1-shot. Also, as expected Adaptive Omnipose consistently surpasses CellTranspose, since CellTranspose builds on Cellpose, which was not designed to handle morphologies where a cell center is not clearly defined.

BP \rightarrow Worm and BF \rightarrow Worm. The target dataset is Worm. The trend is similar to the BP \rightarrow BF and BF \rightarrow BP cases, where Adaptive Omnipose is effective even with 1-shot and consistently improves, and is performing above Omnipose-FT and CellTranspose.

Ablation Study. Since Adaptive Omnipose adds two new losses to the Omnipose framework, we test it by removing one loss at a time on a 1-shot adaptation scheme. Table 2 shows that including even one loss leads to significant improvement over just fine-tuning the model, while adding both leads to the best improvement.

4. CONCLUSIONS

We introduced a few-shot domain adaptation procedure for cell instance segmentation. It improves upon prior work because the model it adapts can handle elongated and non-convex morphologies by design. The procedure requires very limited manual annotation efforts to be effective (even one to five cells is enough), and it requires about 3 minutes with an off-the-shelf GPU. Our experiments based on pretraining on two different sources, and adapting on three different challenging targets fully support our modeling framework, including a favorable comparison with previous work.

5. COMPLIANCE WITH ETHICAL STANDARDS

Ethical approval was not required as confirmed by the license attached with the open access data.

6. ACKNOWLEDGMENTS

Research reported in this publication was supported by the National Institute Of Mental Health of the National Institutes of Health under Award Number R44MH125238. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. This material is also based upon work supported by the National Science Foundation under Grants No. 1920920, 2223793.

7. REFERENCES

- [1] H. Jeckel and K. Drescher, “Advances and opportunities in image analysis of bacterial cells and communities,” *FEMS Microbiol. Rev.*, vol. 45, no. 4, Aug. 2021.
- [2] A. Ducret, E. M. Quardokus, and Y. V. Brun, “MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis,” *Nat Microbiol*, vol. 1, no. 7, pp. 16077, June 2016.
- [3] C. Stringer, T. Wang, M. Michaelos, and M. Pachitariu, “Cellpose: a generalist algorithm for cellular segmentation,” *Nat. Methods*, vol. 18, no. 1, pp. 100–106, Jan. 2021.
- [4] K. J. Cutler, C. Stringer, T. W. Lo, et al., “Omnipose: a high-precision morphology-independent solution for bacterial cell segmentation,” *Nat. Methods*, pp. 1–11, Oct. 2022.
- [5] H. Shimodaira, “Improving predictive inference under covariate shift by weighting the log-likelihood function,” *J. Stat. Plan. Inference*, vol. 90, no. 2, pp. 227–244, Oct. 2000.
- [6] U. Schmidt, M. Weigert, C. Broaddus, and G. Meyers, “Cell detection with star-convex polygons,” in *MICCAI*, 2018, pp. 265–273.
- [7] N. F. Greenwald, G. Miller, E. Moen, et al., “Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning,” *Nat. Biotechnol.*, Nov. 2021.
- [8] S. Panigrahi, D. Murat, A. Le Gall, et al., “Mistic, a general deep learning-based method for the high-throughput cell segmentation of complex bacterial communities,” *Elife*, vol. 10, Sept. 2021.
- [9] S. Motiian, M. Piccirilli, D. A. Adjeroh, and G. Doretto, “Unified deep supervised domain adaptation and generalization,” in *ICCV*, Oct. 2017.
- [10] H. Guan and M. Liu, “Domain adaptation for medical image analysis: A survey,” *IEEE Trans. Biomed. Eng.*, vol. 69, no. 3, pp. 1173–1185, Mar. 2022.
- [11] S. Xun, D. Li, H. Zhu, et al., “Generative adversarial networks in medical image segmentation: A review,” *Comput. Biol. Med.*, vol. 140, pp. 105063, Nov. 2021.
- [12] M. Goetz, C. Weber, F. Binczyk, et al., “DALSA: Domain adaptation for supervised learning from sparsely annotated MR images,” *IEEE Trans. Med. Imaging*, vol. 35, no. 1, pp. 184–196, Jan. 2016.
- [13] Q. Zhu, B. Du, and P. Yan, “Boundary-Weighted domain adaptive neural network for prostate MR image segmentation,” *IEEE Trans. Med. Imaging*, vol. 39, no. 3, pp. 753–763, Mar. 2020.
- [14] R. Bermúdez-Chacón, P. Márquez-Neila, M. Salzmann, and P. Fua, “A domain-adaptive two-stream U-Net for electron microscopy image segmentation,” in *ISBI*, Apr. 2018, pp. 400–404.
- [15] M. Javanmardi and T. Tasdizen, “Domain adaptation for biomedical image segmentation using adversarial training,” in *ISBI*, Apr. 2018, pp. 554–558.
- [16] T. Zhang, J. Cheng, H. Fu, et al., “Noise adaptation generative adversarial network for medical image analysis,” *IEEE Trans. Med. Imaging*, vol. 39, no. 4, pp. 1149–1159, Apr. 2020.
- [17] S. Yang, J. Zhang, J. Huang, B. C. Lovell, and X. Han, “Minimizing labeling cost for nuclei instance segmentation and classification with cross-domain images and weak labels,” *AAAI*, vol. 35, no. 1, pp. 697–705, May 2021.
- [18] D. Liu, D. Zhang, Y. Song, et al., “PDAM: A Panoptic-Level feature alignment framework for unsupervised domain adaptive instance segmentation in microscopy images,” *IEEE Trans. Med. Imaging*, vol. 40, no. 1, pp. 154–165, Jan. 2021.
- [19] M. R. Keaton, R. J. Zaveri, and G. Doretto, “CellTranspose: Few-shot Domain Adaptation for Cellular Instance Segmentation,” in *IEEE/CVF WACV*, Jan. 2023, pp. 455–466.
- [20] D. T. Kysela, A. M. Randich, P. D. Caccamo, and Y. V. Brun, “Diversity takes shape: Understanding the mechanistic and adaptive basis of bacterial morphology,” *PLoS Biol.*, vol. 14, no. 10, pp. e1002565, Oct. 2016.
- [21] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*, vol. 1, MIT press Cambridge, 2016.
- [22] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, “A simple framework for contrastive learning of visual representations,” in *ICML*, Hal Daumé Iii and Aarti Singh, Eds., 2020, vol. 119, pp. 1597–1607.
- [23] A. Javer, M. Currie, C. W. Lee, J. Hokanson, et al., “An open-source platform for analyzing and sharing worm-behavior data,” *Nature methods*, vol. 15, no. 9, pp. 645–646, 2018.
- [24] V. Ljosa, K. L. Sokolnicki, and A. E. Carpenter, “Annotated high-throughput microscopy image sets for validation,” *Nature methods*, vol. 9, no. 7, pp. 637–637, 2012.