Low-Resource Cross-Modality Nucleus Detection via One-Sided Domain Adaptation

Fuyong Xing[1] and Siyang Ren[1]

[1] University of Colorado Anschutz Medical Campus
13001 E 17th Pl, Aurora, CO 80045, USA
fuyong.xing@cuanschutz.edu

Abstract. Unsupervised domain adaptation (UDA) based on generative adversarial networks (GANs) or diffusion models has recently demonstrated impressive performance in various medical imaging tasks, including nucleus or cell detection in cross-modality microscopy images. However, these UDA methods often do not consider low-resource settings in the target domain, i.e., only limited target training data is available. Therefore, they suffer from significant performance degradation in low-data regimes. In this paper, we propose a novel one-sided UDA method for low-resource nucleus detection in cross-modality microscopy images. Specifically, we design and incorporate a new masked contrastive learning module into a one-sided GAN for unpaired image translation, which explicitly encourages consistent feature encoding between corresponding source and translated patches. This module is further coupled with a nucleus detector to jointly augment the discriminator and foster semantic content preservation during image translation for domain adaptation. We evaluate the method on multiple public cross-modality microscopy image datasets. In low-data regimes, the method provides significantly superior nucleus detection performance to the reference baseline and recent state-of-the-art approaches, including those diffusion-based models.

Keywords: Domain adaptation · Low resource · Nucleus detection.

1 Introduction

Deep neural networks have been applied to nucleus/cell detection in microscopy imaging data, but they often do not generalize well on cross-modality images due to domain shifts. Domain adaptation methods [13, 15, 3, 28], especially those unsupervised domain adaptation (UDA) approaches based on generative adversarial networks (GANs) or diffusion models, have produced impressive performance in various medical imaging tasks [7, 18, 41, 19, 22, 23, 10, 36, 37, 2, 30], including cross-modality or -domain microscopy image analysis [5, 26, 11, 34].

Current UDA methods often assume that the target domain has sufficient unannotated training data available, but this assumption does not always hold in real-world applications, particularly for underrepresented groups of patients

or rare diseases like neuroendocrine tumors (NETs). Those GAN- or diffusion-based UDA approaches suffer from significant performance degradation in low-resource settings, i.e., target training data is scarce, because training a GAN or diffusion model typically requires a large amount of data [40, 20, 38, 16]. A recent study has improved a dual GAN for domain adaptation with limited target data [34]. However, this method requires two generator-discriminator pairs for image-to-image translation, which exhibit a high computational cost. In addition, its cross-domain prediction consistency [8] relies on high-quality target-to-source translated images to generate good pseudo-labels for target images, but this is challenging due to limited training data. Therefore, the generated pseudo-labels may not serve as reliable supervision signals for model learning.

In this paper, we propose a novel low-resource UDA method for generalizable nucleus detection in cross-modality microscopy images. Instead of relying on a dual GAN that performs bidirectional image translation, we design and incorporate a new masked contrastive learning (MCL) module into a one-sided GAN, which enables fast, single-directional image translation in low-data regimes. We combine this MCL module with a nucleus detector to jointly augment the discriminator and enhance generator learning for semantic content preservation during image translation. In addition, we apply differentiable data augmentation to the data-limited target domain for explicit alleviation of discriminator overfitting. We evaluate the proposed method with multiple public and in-house cross-modality microscopy image datasets. Our method significantly outperforms the reference baseline and multiple state-of-the-art domain adaptation approaches in low-resource settings, including recent diffusion-based models.

2 Methodology

Fig. 1 presents an overview of the proposed low-resource UDA framework. The generator G first translates source images into target-style images, which are then used to learn a target-domain nucleus detector R. A novel MCL module is incorporated into the framework to explicitly promote consistent feature encoding between corresponding patches of real and translated source images. This MCL module cooperates with the nucleus detector R to supplement the discriminator D and provides the generator G with additional supervision signals for enhanced model training with limited data. The framework applies differentiable data augmentation to real target images and translated, target-style images to reduce discriminator overfitting and further improve generator learning.

Problem Definition Let $(\boldsymbol{X}_S, \boldsymbol{Y}_S) = \{(\boldsymbol{x}_S^i, \boldsymbol{y}_S^i)\}_{i=1}^{N_S}$ denote the source training data with N_S pairs of images and associated gold-standard labels. For each source image $\boldsymbol{x}_S^i \in \mathbb{R}^{H\times W\times C}$, where H, W and C represent the height, width and channels of the image respectively, its label $\boldsymbol{y}_S^i \in \mathbb{R}^{H\times W}$ is a 2D heatmap with each position (u, v) calculating its value using the formula [32, 33]: $\boldsymbol{y}_S^i(u, v) = (e^{\alpha(1-d^i(u,v)/\epsilon)}-1)/(e^{\alpha}-1)$ if $d^i(u, v) \leq \epsilon$; otherwise, $\boldsymbol{y}_S^i(u, v) = 0$. The $d^i(u, v)$ is the Euclidean distance between position (u, v) and its nearest annotated nucleus

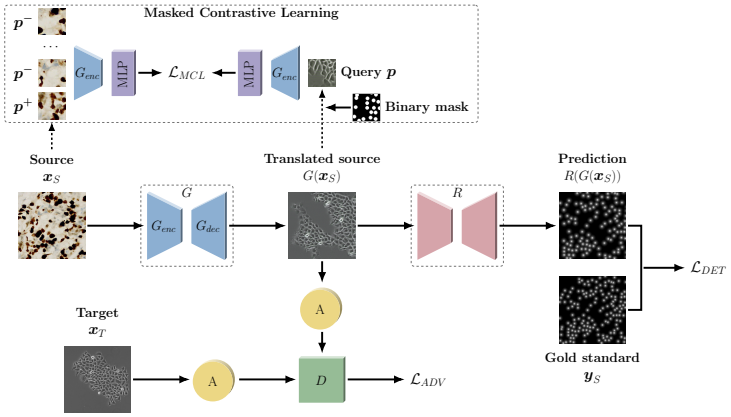


Fig. 1. Overview of the proposed framework. G (composed of G_{enc} and G_{dec}), D and R denote the generator, discriminator and nucleus detector, respectively. MLP is the multilayer perceptron and A represents the differentiable data augmentation module. $\mathcal{L}_{ADV}$, $\mathcal{L}_{DET}$ and $\mathcal{L}_{MCL}$ are the losses for adversarial learning, nucleus detection and contrastive learning, respectively. $\boldsymbol{p}/\boldsymbol{p}^+/\boldsymbol{p}^-$: query/positive/negative patches.

center, ϵ=16 is a distance threshold to define the spatial support, and α=3 controls the exponential decay rate. Let $\boldsymbol{X}_T = \{\boldsymbol{x}_T^i\}_{i=1}^{N_T}$ represent N_T unlabeled target training images, where $N_T \ll N_S$. With $(\boldsymbol{X}_S, \boldsymbol{Y}_S)$ and $\boldsymbol{X}_T$, the proposed framework aims to learn a nucleus detector for unseen target images.

2.1 Masked Contrastive Learning for Low-Resource UDA

Nucleus or cell detection in cross-modality microscopy images often does not have paired data available. Current pixel-space UDA methods usually adopt a dual-model framework for unpaired image translation, mainly based on cycle-consistent GANs [7, 18, 26, 33, 34] or diffusion models [29, 35, 9, 43]. However, these methods are not suitable for low data regimes due to discriminator overfitting in GANs [40, 20] or failure in denoising score matching in diffusion models [38, 16]. Inspired by previous work [4, 25, 6], we address this issue from a different viewpoint: incorporating an auxiliary task of instance discrimination into a one-sided GAN to enable model training with limited data. To this end, we introduce a novel MCL module, in conjunction with the GAN's discriminator, to allow single-directional, source-to-target image translation for fast domain adaptation. Specifically, we tailor a recent patchwise contrastive learning technique [25] for cross-modality nucleus detection with the following significant improvements: 1) Apply a binary mask to patch sampling such that query and positive patches are cropped from only nucleus regions, while negative patches are extracted from other areas. Unlike natural images where foreground objects typically account for a large proportion of each image, nucleus images often have much smaller nucleus regions than the background. Thus, random patch sampling without

masking [25] may prioritize feature learning for image background. 2) Extend the contrastive learning from image translation to object detection by combining it with the task-specific supervised learning of nucleus detection and jointly augmenting the discriminator for enhanced generator learning (see Section 2.2).

We place the MCL module, termed M, on top of the generator's encoder G_{enc} to encourage consistent feature encoding between corresponding patches in original image $\boldsymbol{x}_S^i$ and its translated output $G(\boldsymbol{x}_S^i)$. This is based on the intuition that a patch in $G(\boldsymbol{x}_S^i)$ should look more similar to its corresponding patch at the same spatial position of the input image $\boldsymbol{x}_S^i$, compared with other random patches. This module takes as input feature maps from multiple layers of the generator's encoder G_{enc}, and then feeds them into a multilayer perceptron (MLP) for instance discrimination. Specifically, for each query patch $\boldsymbol{p}$ in $G(\boldsymbol{x}_S^i)$, its associated positive patch $\boldsymbol{p}^+$ is extracted at the same location of $\boldsymbol{x}_S^i$ and k negative patches $\{\boldsymbol{p}_j^-\}_{j=1}^k$ are cropped at other positions. Note that a binary mask $\boldsymbol{z}_S^i \in \mathbb{R}^{H\times W}$ is used to constrain the patch sampling, i.e., query and positive patches are sampled from nucleus regions, while negative ones are from the background. Here the mask $\boldsymbol{z}_S^i(u,v) = 1$ if $\boldsymbol{y}_S^i(u,v) > 0$, and 0 otherwise. We learn the MCL by promoting similarity between each query and its corresponding positive patch, while driving separation from negative patches. We achieve this by optimizing an InfoNCE loss [24] to solve a $(k+1)$-way classification problem

$$\mathcal{L}_{MCL} = \mathbb{E}_{(\boldsymbol{x}_S,\boldsymbol{y}_S)\sim(\boldsymbol{X}_S,\boldsymbol{Y}_S)}\left[-\log\left(\frac{e^{\boldsymbol{q}\cdot\boldsymbol{q}^+/\tau}}{e^{\boldsymbol{q}\cdot\boldsymbol{q}^+/\tau}+\sum_{j=1}^k e^{\boldsymbol{q}\cdot\boldsymbol{q}_j^-/\tau}}\right)\right], \tag{1}$$

where $\mathbb{E}$ is the expectation operator, $\cdot$ represents the dot product and τ denotes a temperature parameter to control the softness of the distribution. $\boldsymbol{q} = M(G_{enc}(\boldsymbol{p}))$ is the feature presentation of patch $\boldsymbol{p}$. Note that for each query patch, its positive and negative patches are cropped from the same input image $\boldsymbol{x}_S^i$. This internal patch sampling strategy does not require a memory bank or a large dictionary that is adopted in other commonly used contrastive learning methods [14, 31], thus leading to improved memory consumption. Compared with previous work [25], the MCL pays more attention to nucleus regions. In addition, the MCL module enables unpaired image translation with only one generator-discriminator pair, therefore significantly reducing the number of learnable parameters in comparison with those CycleGAN-style approaches [7, 18, 26, 33, 34].

2.2 Joint Discriminator Augmentation

Eqn. (1) uses an auxiliary task of instance discrimination to improve generator learning with limited data for domain adaptation. This does not explicitly enforce the generator G to preserve semantic content with respect to prediction of nucleus labels. To address this issue, we combine the instance discrimination with the primary task, nucleus detection, to jointly enhance generator learning by directly using gold-standard nucleus labels as another supervision signal. In

this way, the generator G receives gradient information from the MCL module M and the nucleus detector R, in addition to the GAN's discriminator D. This joint discriminator augmentation can enhance the gradient flow into the generator G and improve model training in low-resource settings.

For an input source image $\boldsymbol{x}_S^i$, the generator G and the discriminator D compete with each other for model learning by optimizing an adversarial loss [12], $\mathcal{L}_{ADV}$. Meahwhile, the nucleus detector R takes as input the translated image $G(\boldsymbol{x}_S^i)$ to minimize a weighted mean squared error (MSE) [32], $\mathcal{L}_{DET}$, with $\boldsymbol{y}_S^i$ as the supervision label. More importantly, the detector R cooperates with the MCL module M within a single GAN framework for generator learning: the module M encourages consistent feature encoding between corresponding patches, while the detector R enforces the generator G to preserve nuclei during image translation. They augment the discriminator D from different aspects and simultaneously backpropagate gradients to G with additional supervision information. To explicitly alleviate discriminator overfitting, the framework applies a differentiable data augmentation module A [20] to both $G(\boldsymbol{x}_S^i)$ and $\boldsymbol{x}_T^i$ before sending them to D. Formally, the proposed framework is formulated as

$$\mathcal{L} = \mathcal{L}_{ADV} + \mathcal{L}_{DET} + \lambda \mathcal{L}_{MCL}, \tag{2}$$

$$\mathcal{L}_{ADV} = \mathbb{E}_{\boldsymbol{x}_T \sim \boldsymbol{X}_T}[\log D(A(\boldsymbol{x}_T))] + \mathbb{E}_{\boldsymbol{x}_S \sim \boldsymbol{X}_S}[\log(1 - D(A(G(\boldsymbol{x}_S))))], \tag{3}$$

$$\mathcal{L}_{DET} = \mathbb{E}_{(\boldsymbol{x}_S, \boldsymbol{y}_S) \sim (\boldsymbol{X}_S, \boldsymbol{Y}_S)}[||(\boldsymbol{y}_S + \beta \bar{\boldsymbol{y}}_S)^{1/2} \odot (R(G(\boldsymbol{x}_S)) - \boldsymbol{y}_S)||_F^2], \tag{4}$$

where λ is a weighting parameter for the MCL. The $\bar{\boldsymbol{y}}_S \in \mathbb{R}^{H \times W}$ is a matrix with all elements equal to the mean value of $\boldsymbol{y}_S$. The β controls the importance of nucleus regions, and $\odot$ represents the element-wise multiplication operator. Eqn. (2) not only allows the GAN to perform fast one-sided image translation, but also enforces it to directly adopt gold-standard nucleus labels to guide image translation in low-resource settings. This is different from other augmented GANs [7, 18, 33] that do not consider limited training data or joint discriminator augmentation. Another significant benefit of Eqn. (2) is elimination of the cross-domain prediction consistency constraint [8, 34], which heavily relies on good target pseudo-label generation and is difficult to achieve in low-data regimes. Finally, this modeling applies data augmentation to only the data-scarce target domain but not the data-rich source domain. Unlike the prior work [34] that uses a series of image transformations on both source- and target-domain data, our design avoids additional computational burden during model training.

3 Experiments

Datasets We evaluate the proposed method using 4 cross-modality microscopy image datasets. The immunohistochemistry (IHC)-stained breast cancer dataset

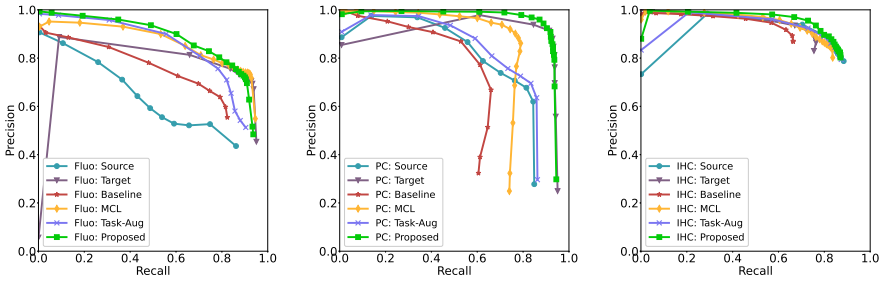


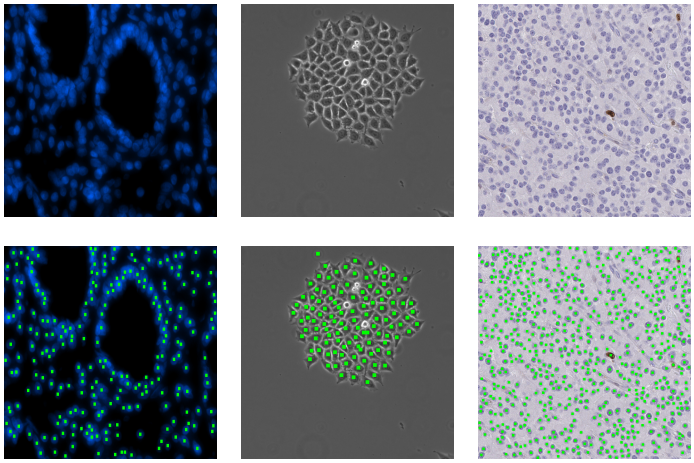
Fig. 2. The precision-recall curves of the ablation study on different target datasets.

[17] is acquired with bright-field microscopy imaging, containing over 1330 images with 803 as the training set. Another IHC dataset (termed IHC) has 60 gastrointestinal NET images. The DAPI-stained fluorescence dataset ($Fluo$) [27] contains 120 colon tissue images, and the cervical cancer dataset (PC) [1] includes 22 phase-contrast microscopy images without chemical staining. We select the large-scale breast cancer dataset as the source domain and the others as target datasets. Following [32, 34], we randomly select 50% of each target dataset for model evaluation, and randomly choose one image from the rest 50% for model training in low-resource settings [34]. We employ precision, recall, F_1 score, and area under precision-recall curve (PRAUC) for model evaluation [32, 34].

Implementation Details We choose the network architectures in CycleGAN [42] to implement our generator and discriminator, use an encoder–decoder network [32] as the nucleus detector, and adopt a two-layer MLP for the MCL module. The detector's encoder has 4 residual learning blocks, each containing two 3×3 convolutional layers, instance normalization and exponential linear units. The decoder mirrors the encoder structure. For differentiable data augmentation in Eqn. (3), we use native PyTorch tensor operations to implement image transformations, including pixel blitting (flipping, right-angle rotation and translation), geometric transformations (scaling and random-degree rotation), and color/intensity adjustments (brightness, contrast, saturation, and hue).

We follow [25, 14] to set $\tau = 0.07$, $k = 255$ in Eqn. (1), use 256 neurons in each layer of the MLP, and extract features from 5 layers of the generator's encoder as the MCL's input [25]. We set $\lambda = 1$ in Eqn. (2) and $\beta = 5$ in Eqn. (4) based on the study [32]. We use the Adam optimizer for model training with the hyperparameter values: learning rate=0.0002, number of epochs=100, and batch size =1. For model testing, we first use a threshold ξ to remove low-valued responses in each prediction map from the nucleus detector and then detect individual nuclei by seeking local maxima in the prediction map.

Ablation Study Fig. 2 shows the precision-recall curves of the ablation study by varying the ξ from 0 to 1. The *Source* and *Target* represent a nucleus detector trained with labeled source and target data, respectively. The *Baseline*



(a) Fluo (b) PC (c) IHC

Fig. 3. Example nucleus detection results using the proposed method. Rows 1 and 2 represent the original images and the detection results (green dots), respectively.

denotes training a one-sided GAN with $\mathcal{L}_{ADV}$ and then a nucleus detector with $\mathcal{L}_{DET}$, MCL means adding the MCL module to the $Baseline$ model, and $Task\text{-}Aug$ represents training a GAN with a joint loss $\mathcal{L}_{ADV} + \lambda\mathcal{L}_{DET}$. We observe that in low-resource UDA, either the MCL or the task augmentation improves the nucleus detection results in comparison with the $Baseline$ model. The MCL also outperforms the random patch sampling-based contrastive unpaired translation (CUT) [25] in Table 1. The proposed method combining both MCL and task augmentation further boosts the performance (PRAUC for $Fluo/PC/IHC$: 84.1/91.8/84.3) and even outperforms the $Target$ model (PRAUC for $Fluo/PC/IHC$: 74.7/86.5/73.6), which uses limited training data (e.g., one image). Fig. 3 shows qualitative results of nucleus detection.

Comparison with State of The Art We compare the proposed method with several recent state-of-the-art UDA approaches including GAN- and diffusion-based models, such as CycleGAN [42], CUT [25], pseudo-labeling-based domain adaptation (PLDA) [33], low-resource dual GAN (LRDGAN) [34], dual diffusion model (CycleDiffusion) [29], energy-guided stochastic differential equation (EGSDE) [39], unpaired neural Schrödinger bridge (UNSB) [21], and contour-guided diffusion model (ContourDiff) [9].

Table 1 summarizes the nucleus detection results with $\xi = 0.5$. The proposed method outperforms the competitors in all the target datasets. In particular, it produces a significantly higher F_1 score than others for almost all the cases with p-value less than 0.05 in two-sample t-tests, except for the LRDGAN on the $Fluo$ dataset. Note that our method uses only one generator-discriminator pair and has a significantly lower computational cost (26.4×10^6 learnable parame-

Table 1. Nucleus detection on different target datasets. Each method is trained with 5 different random seeds, and the mean and standard deviation (std) of each metric over the 5 runs are reported: $^{mean}_{(std)}$. The $*$ means our method has a significantly higher F_1 score than others. The highest F_1 score is in **bold** font. P=Precision, R=Recall.

	Fluo			PC			IHC		
	P	R	F_1	P	R	F_1	P	R	F_1
CycleGAN [42]	45.3	42.4	41.6*	83.2	40.1	53.6*	64.4	43.5	51.3*
	(14.3)	(13.3)	(10.6)	(23.9)	(12.0)	(15.4)	(13.9)	(15.1)	(14.7)
CUT [25]	66.5	44.4	52.1*	63.4	53.1	55.6*	74.7	60.0	65.1*
	(4.5)	(15.3)	(11.8)	(14.5)	(25.1)	(16.4)	(7.6)	(12.2)	(6.9)
PLDA [33]	53.7	78.9	63.5*	65.6	81.2	72.4*	75.6	80.2	77.6*
	(12.3)	(11.3)	(11.5)	(11.7)	(12.3)	(11.5)	(6.1)	(3.1)	(3.3)
LRDGAN [34]	78.7	79.5	79.0	97.9	60.9	74.8*	92.5	72.2	81.0*
	(1.9)	(1.6)	(0.9)	(1.0)	(8.4)	(6.7)	(1.2)	(4.2)	(2.6)
CycleDiffusion [29]	56.9	46.3	50.9*	63.8	60.1	60.4*	77.9	47.4	58.6*
	(8.0)	(3.9)	(4.8)	(4.5)	(21.2)	(13.5)	(4.6)	(7.1)	(4.8)
EGSDE [39]	58.9	49.1	53.4*	67.2	82.4	73.5*	79.1	41.9	54.6*
	(3.0)	(5.0)	(3.3)	(4.3)	(14.6)	(8.2)	(1.6)	(4.9)	(3.6)
UNSB [21]	58.1	48.0	52.1*	67.0	52.7	52.6*	70.8	58.6	64.0*
	(4.1)	(9.0)	(5.6)	(13.0)	(29.0)	(16.4)	(7.7)	(6.6)	(6.6)
ContourDiff [9]	52.1	42.8	47.0*	59.9	74.9	66.1*	75.1	60.0	66.6*
	(7.5)	(7.3)	(7.4)	(5.8)	(11.8)	(5.8)	(6.6)	(6.3)	(6.2)
Proposed	77.0	84.5	**80.5**	95.4	88.2	**91.6**	87.8	83.1	**85.3**
	(2.7)	(3.6)	(1.8)	(1.8)	(6.0)	(3.8)	(1.3)	(2.8)	(0.9)

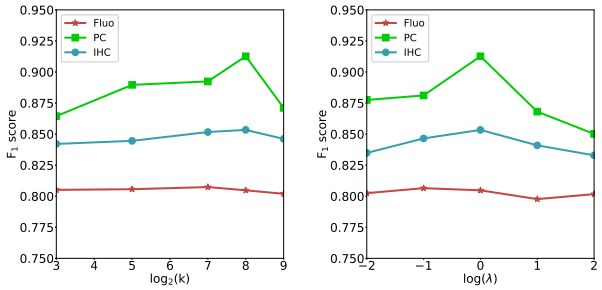


Fig. 4. Nucleus detection with different values of k in Eqn. (1) and λ in Eqn. (2).

ters) than the LRDGAN (51.8×10^6 learnable parameters) [34]. These results demonstrate the superior effectiveness of our method in low-resource settings, compared with other approaches including diffusion-based models.

Effects of Parameters Fig. 4 demonstrates the effects of the number of negative patches k in Eqn. (1) and the weighting parameter λ in Eqn. (2).We note that the nucleus detection performance first improves with respect to k and then decreases when $k > 255$, suggesting that including significantly more negative patches may not be necessary. The results with different λ values indicate the

importance of the MCL module, but an overemphasis with a substantially larger λ value may lead to performance degradation.

4 Conclusion

We introduce a novel one-sided GAN method for cross-modality nucleus detection in low-resource settings. The MCL module enables efficient single-directional image translation, while the joint discriminator augmentation explicitly enforces semantic content preservation with respect to prediction of nucleus labels for object detection. The proposed method significantly outperforms multiple recent state-of-the-art approaches including diffusion-based models. One potential limitation of the method is its MCL's dependence on source-domain nucleus masks, and noisy annotations may negatively affect the MCL's performance. Our method does not penalize use of more target training data, but its relative advantage may gradually decreases as more training target images are used.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Arteta, C., et al.: Learning to detect cells using non-overlapping extremal regions. In: MICCAI 2012. LNCS, vol. 7510, pp. 348–356. Springer, Cham (2012)
2. Ashraf, T., et al.: D-MASTER: Mask annealed transformer for unsupervised domain adaptation in breast cancer detection from mammograms. In: Linguraru, M.G., et al. (eds.) MICCAI 2024. LNCS, vol. 15011, pp. 189–199. Springer (2024)
3. Basak, H., Yin, Z.: Quest for clone: Test-time domain adaptation for medical image segmentation by searching the closest clone in latent space. In: Linguraru, M.G., et al. (eds.) MICCAI 2024. LNCS, vol. 15008, pp. 555–566. Springer, Cham (2024)
4. Benaim, S., Wolf, L.: One-sided unsupervised domain mapping. In: NeurIPS. p. 752–762 (2017)
5. Bentaieb, A., Hamarneh, G.: Adversarial stain transfer for histopathology image analysis. IEEE TMI **37**(3), 792–802 (2018)
6. Cao, J., Hou, L., Yang, M.H., He, R., Sun, Z.: Remix: Towards image-to-image translation with limited data. In: CVPR. pp. 15013–15022 (2021)
7. Chen, C., Dou, Q., Chen, H., Heng, P.A.: Semantic-aware generative adversarial nets for unsupervised domain adaptation in chest x-ray segmentation. In: Shi, Y., et al. (eds.) MLMI 2018. LNCS, vol. 11046, pp. 143–151. Springer, Cham (2018)
8. Chen, Y.C., Lin, Y.Y., Yang, M.H., Huang, J.B.: CrDoCo: Pixel-level domain transfer with cross-domain consistency. In: CVPR. pp. 1791–1800 (2019)
9. Chen, Y., et al.: Contourdiff: Unpaired medical image translation with structural consistency. MLBI **3**, 711–727 (2025)
10. Du, D., et al.: CAUDA-MI: Cross attention-guided unsupervised domain adaptation with mutual information for cardiac mri segmentation. In: Gee, J.C., et al. (eds.) MICCAI 2025. LNCS, vol. 15965, pp. 46–56. Springer, Cham (2025)
11. Gadermayr, M., et al.: Generative adversarial networks for facilitating stain-independent supervised and unsupervised segmentation: A study on kidney histology. IEEE TMI **38**(10), 2293–2302 (2019)

12. Goodfellow, I., et al.: Generative adversarial nets. In: NeurIPS (2014)
13. Guan, H., Liu, M.: Domain adaptation for medical image analysis: A survey. IEEE TBME **69**(3), 1173–1185 (2022)
14. He, K., Fan, H., Wu, Y., Xie, S., Girshick, R.: Momentum contrast for unsupervised visual representation learning. In: CVPR. pp. 9726–9735 (2020)
15. Heo, K.S., et al.: Sparsely labeled fMRI data denoising with meta-learning-based semi-supervised domain adaptation. In: Gee, J.C., et al. (eds.) MICCAI 2025. LNCS, vol. 15966, pp. 583–593. Springer, Cham (2025)
16. Hu, T., et al.: Phasic content fusing diffusion model with directional distribution consistency for few-shot model adaption. In: ICCV. pp. 2406–2415 (2023)
17. Huang, Z., et al.: BCData: A large-scale dataset and benchmark for cell detection and counting. In: MICCAI. LNCS, vol. 12265, pp. 289–298. Springer, Cham (2020)
18. Huo, Y., et al.: Synseg-net: Synthetic segmentation without target modality ground truth. IEEE TMI **38**(4), 1016–1025 (2019)
19. Ji, W., Chung, A.C.S.: Diffusion-based domain adaptation for medical image segmentation using stochastic step alignment. In: Linguraru, M.G., et al. (eds.) MICCAI 2024. LNCS, vol. 15008, pp. 188–198. Springer, Cham (2024)
20. Karras, T., et al.: Training generative adversarial networks with limited data. In: NeurIPS. pp. 12104–12114 (2020)
21. Kim, B., Kwon, G., Kim, K., Ye, J.C.: Unpaired image-to-image translation via neural schrödinger bridge. In: ICLR (2024)
22. Lin, J., et al.: Dsfc: Deformation-aware learning strategy via self-sustaining feedback cycle for medical vision foundation model domain adaptation. In: Gee, J.C., et al. (eds.) MICCAI 2025. LNCS, vol. 15965, pp. 173–182. Springer, Cham (2025)
23. Liu, W., Miao, F., Wang, X.: Solving medical multi-label domain adaptation via wasserstein adversarial learning with class-level alignment. In: Gee, J.C., et al. (eds.) MICCAI 2025. LNCS, vol. 15966, pp. 572–582. Springer, Cham (2025)
24. van den Oord, A., Li, Y., Vinyals, O.: Representation learning with contrastive predictive coding. arXiv:1807.03748 [cs.LG] pp. 1–12 (2018)
25. Park, T., Efros, A.A., Zhang, R., Zhu, J.Y.: Contrastive learning for unpaired image-to-image translation. In: ECCV 2020. pp. 319–345. Springer, Cham (2020)
26. Shaban, M.T., Baur, C., Navab, N., Albarqouni, S.: StainGAN: Stain style transfer for digital histological images. In: ISBI. pp. 953–956 (2019)
27. Tofighi, M., Guo, T., Vanamala, J.K.P., Monga, V.: Prior information guided regularized deep learning for cell nucleus detection. IEEE TMI **38**(9), 2047–2058 (2019)
28. Wang, H., et al.: Advancing UWF-SLO vessel segmentation with source-free active domain adaptation and a novel multi-center dataset. In: MICCAI. pp. 75–85 (2024)
29. Wu, C.H., De La Torre, F.: A latent space of stochastic diffusion models for zero-shot image editing and guidance. In: ICCV. pp. 7344–7353 (2023)
30. Wu, J., et al.: Unsupervised domain adaptation for abdominal organ segmentation using pseudo labels and organ attention cyclegan. In: Ma, J., Wang, B. (eds.) FLARE Challenge 2024. LNCS, vol. 15717, pp. 225–242. Springer, Cham (2024)
31. Wu, Z., Xiong, Y., Yu, S.X., Lin, D.: Unsupervised feature learning via non-parametric instance discrimination. In: CVPR. pp. 3733–3742 (2018)
32. Xie, Y., Xing, F., Shi, X., Kong, X., Su, H., Yang, L.: Efficient and robust cell detection: A structured regression approach. MedIA **44**, 245–254 (2018)
33. Xing, F., Bennett, T., Ghosh, D.: Adversarial domain adaptation and pseudo-labeling for cross-modality microscopy image quantification. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11764, pp. 740–749. Springer, Cham (2019)
34. Xing, F., Cornish, T.C.: Low-resource adversarial domain adaptation for cross-modality nucleus detection. In: MICCAI. LNCS, vol. 13437, pp. 639–649 (2022)

35. Xu, S., et al.: Cyclenet: rethinking cycle consistency in text-guided diffusion for image manipulation. In: NeurIPS. pp. 10359 – 10384 (2023)
36. Zeng, H., et al.: Reliable source approximation: Source-free unsupervised domain adaptation for vestibular schwannoma mri segmentation. In: Linguraru, M.G., et al. (eds.) MICCAI 2024. LNCS, vol. 15010, pp. 622–632. Springer, Cham (2024)
37. Zhang, G., et al.: IPLC: Iterative pseudo label correction guided by SAM for source-free domain adaptation in medical image segmentation. In: Linguraru, M.G., et al. (eds.) MICCAI 2024. pp. 351–360. LNCS, Springer, Cham (2024)
38. Zhang, Z., Hua, Y., Sun, G., Wang, H., Mcloone, S.: Training diffusion-based generative models with limited data. In: ICML. vol. 267, pp. 74950–74965 (2025)
39. Zhao, M., Bao, F., Li, C., Zhu, J.: Egsde: unpaired image-to-image translation via energy-guided stochastic differential equations. In: NeurIPS. pp. 3609–3623 (2022)
40. Zhao, S., Liu, Z., Lin, J., Zhu, J.Y., Han, S.: Differentiable augmentation for data-efficient gan training. In: NeurIPS. pp. 7559–7570 (2020)
41. Zhao, Z., et al.: MT-UDA: Towards unsupervised cross-modality medical image segmentation with limited source labels. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12901, pp. 293–303. Springer, Cham (2021)
42. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networkss. In: ICCV. pp. 2223–2232 (2017)
43. Zou, S., Huang, Y., Yi, R., Zhu, C., Xu, K.: Cyclediff: Cycle diffusion models for unpaired image-to-image translation. IEEE TIP pp. 1–16 (2026)