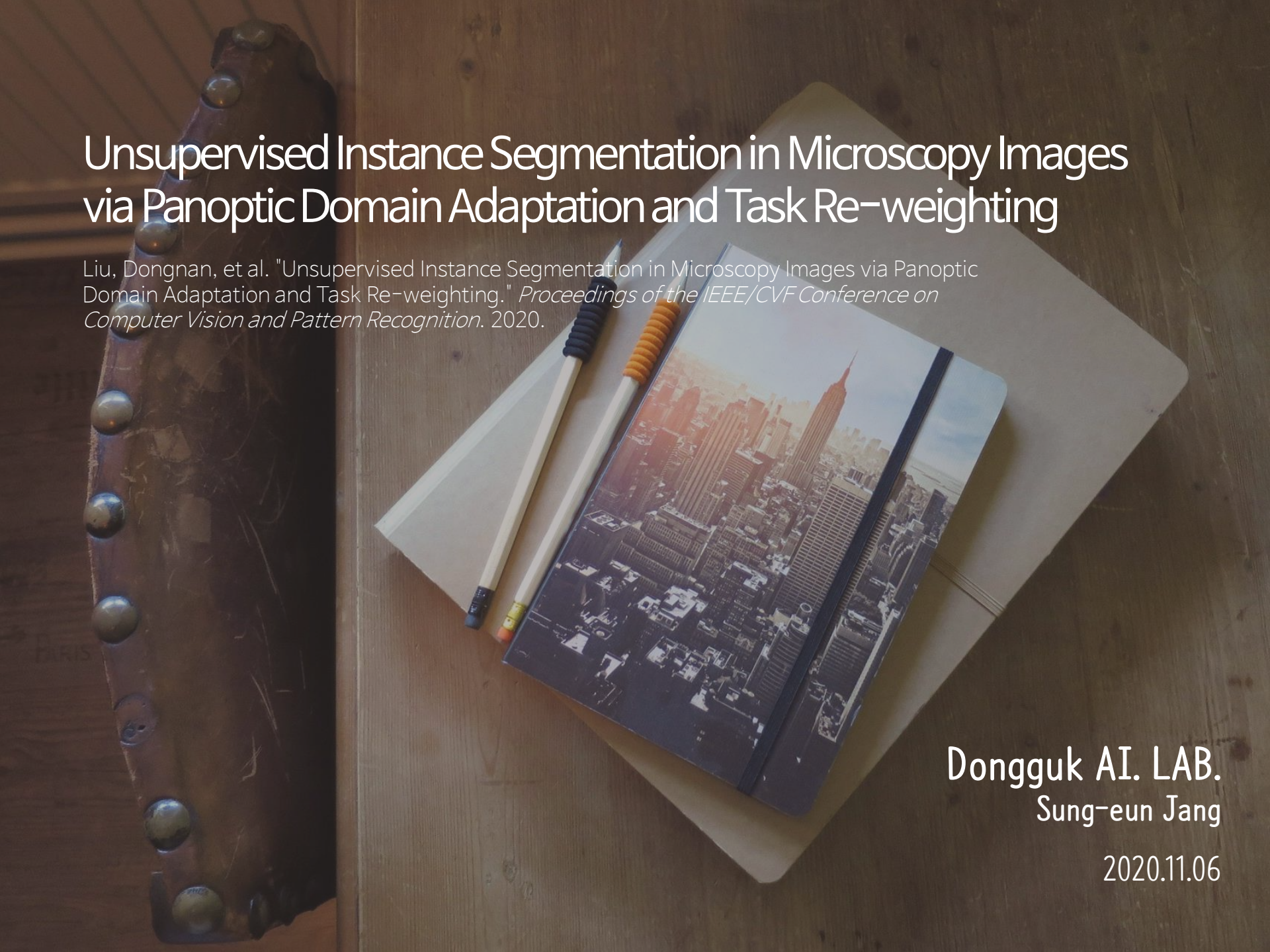
Unsupervised Instance Segmentation in Microscopy Images via Panoptic Domain Adaptation and Task Re-weighting

Liu, Dongnan, et al. "Unsupervised Instance Segmentation in Microscopy Images via Panoptic Domain Adaptation and Task Re-weighting." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020.



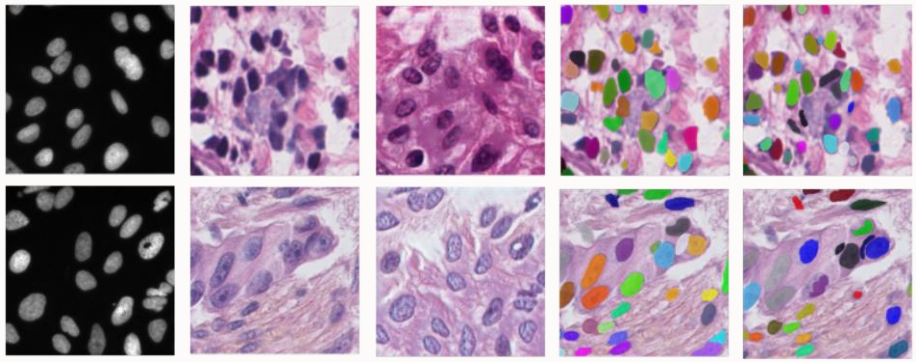
Dongguk AI. LAB.
Sung-eun Jang

2020.11.06

Table of contents

Introduction

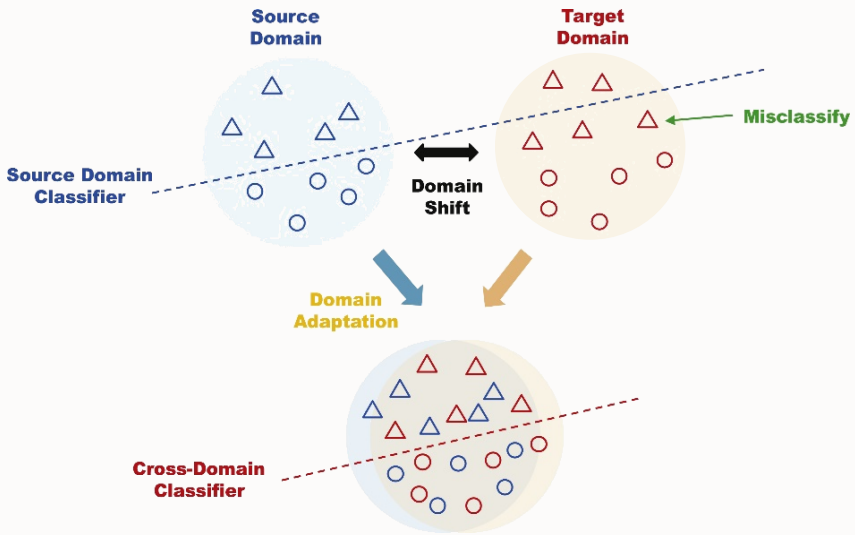
Domain



- Nuclei instance segmentation in histopathology images
- Important step in the digital pathology workflow
- Athologists are able to diagnose and prognose cancers according to…
  - mitosis counts
  - the morphological structure of each nucleus
  - spatial distribution of a group of nuclei

Introduction

Unsupervised Domain Adaptation



- Supervised learning relies on large-scale training data, so requires expertise for annotation
- Unsupervised domain adaptation(UDA) tackle the issue by conducting supervised learning reduces distances between the distribution of feature maps of the source and target domains

Introduction

Limitation of UDA

- Currently, there is a lack of UDA methods specifically designed for instance segmentation and still suffers from challenges

  - Challenge 1: UDA ignore the domain shift at the semantic level, such as …

    the relationship between the foreground and background

    spatial distribution of the objects

  - Challenge 2: UDA object detection methods are multi-task learning

    if the feature extractors fail to generate domain-invariant, then back-propagating the weights according to the task loss functions causes the model bias towards the source domain

Introduction

Proposed method

- Cycle-Consistent Panoptic Domain Adaptive Mask R-CNN (CyC-PDAM) model
    1. A simple nuclei inpainting mechanism to remove false-positive objects in the synthesized images
    2. Domain-invariant features at the panoptic level, by integrating the instance-level adaptation with a semantic-level adaptation module
    3. A task re-weighting mechanism is proposed to alleviate the domain bias towards the source domain

➔ Outperforms state-of-the-art UDA & supervised methods

Method

Overall architecture, CyC-PDAM

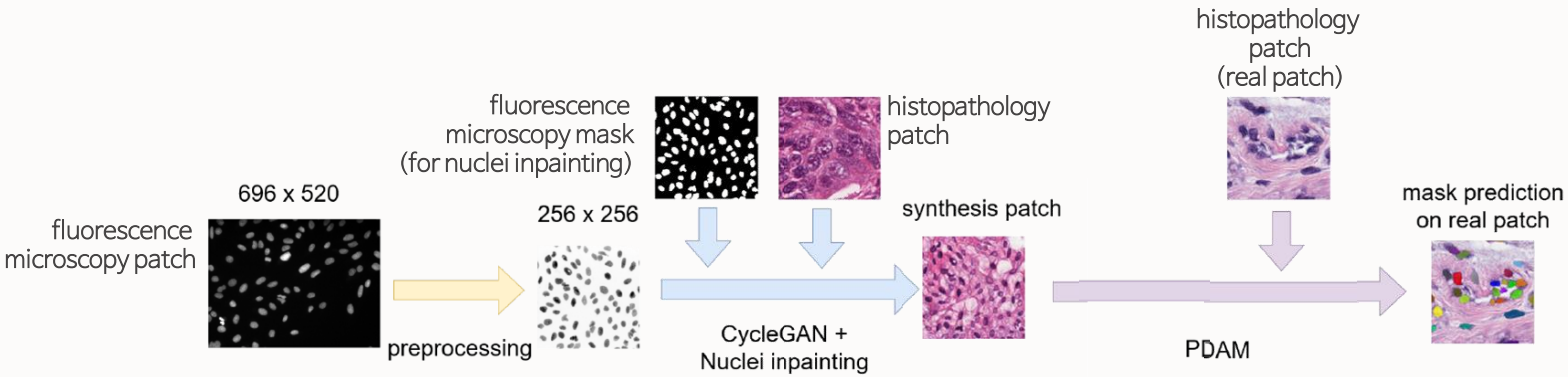
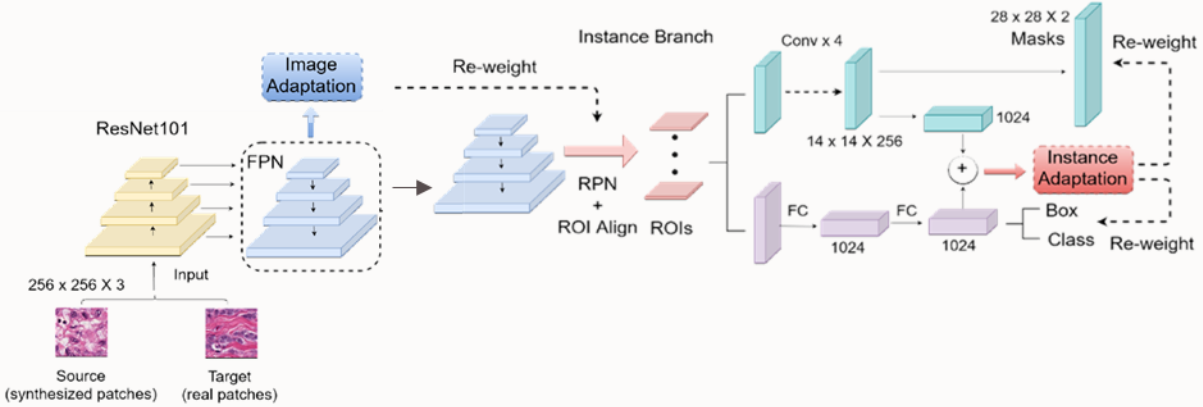


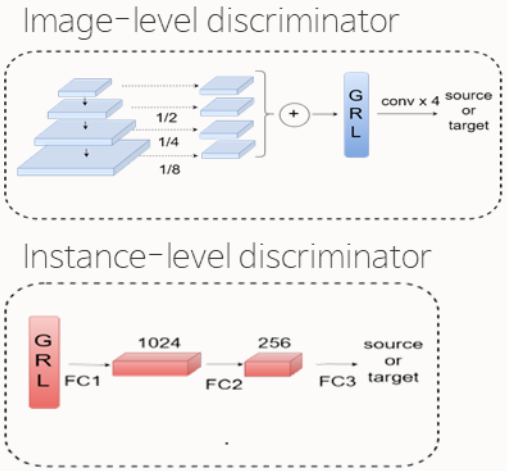
Figure 2. Overall architecture for our proposed CyC-PDAM architecture. The annotations of the real histopathology patches are not used during training.

- Based on CyCADA
- Instance segmentation framework Mask R-CNN
- Nuclei inpainting mechanism
- Panoptic-level domain adaptation
- Task re-weighting mechanism

Method

CyCADA with Mask RCNN





GRL: gradient reversal layer
RPN: region proposal network

- Domain adaptive Mask R-CNN with ResNet101 and FPN backbone
- Two discriminators
  - After FPN for the image-level adaptation
  - After the instance branch for instance-level adaptation

Method

Nuclei Inpainting Mechanism

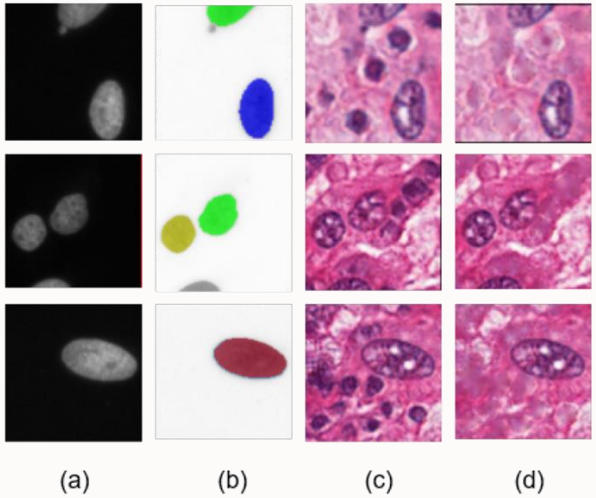


Figure 4. Visual results for the effectiveness of nuclei inpainting mechanism. (a) original fluorescence microscopy patches; (b) corresponding nuclei annotations; (c) initial synthesized images from CycleGAN; (d) final synthesized images after nuclei inpainting mechanism.

- Label space for the generated images sometimes changes after transferring from the source domain

  ➔ can be an obstacle to the training of the model

- Propose an auxiliary nuclei inpainting mechanism to remove the auxiliary objects in the synthesized images, to further avoid false-negative predictions

Method

Nuclei Inpainting Mechanism

$$M_{aux} = (\mathbf{otsu}(S_{\mathbf{raw}}) \cup M) - M$$

$S_{\mathbf{raw}}$: raw synthesized histopathology image by Cycle-GAN

M : corresponding mask

M_{aux} : mask predictions of all the auxiliary generated nuclei

$\mathbf{otsu}(S_{\mathbf{raw}})$: binary segmentation method for S_{raw} based on Otsu threshold

➔ First, obtain the mask predictions M_{aux} of all the auxiliary generated nuclei

➔ In M_{aux} only auxiliary nuclei without annotation is labeled

Method

Nuclei Inpainting Mechanism

$$S_{inp} = inp(S_{raw}, M_{aux})$$

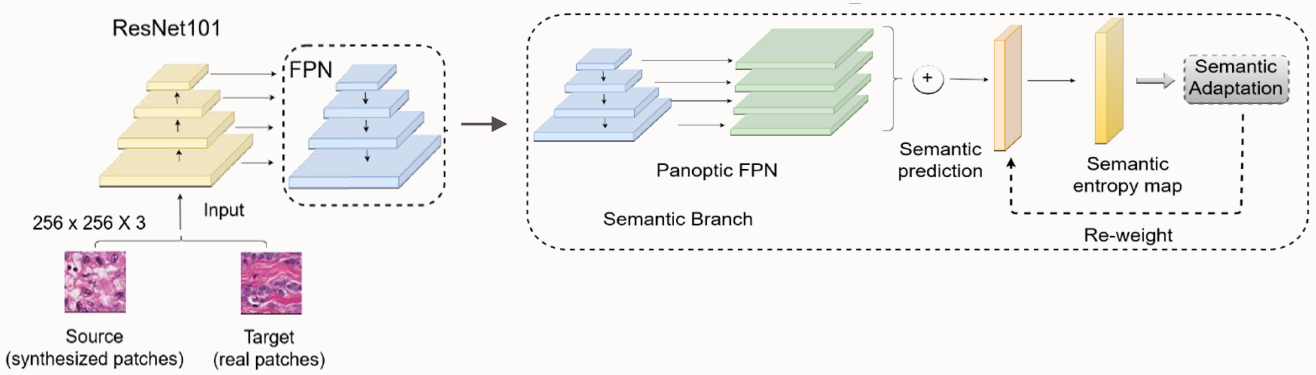
S_{inp} : newly synthesized image

inp : inpainting objects by replacing the pixel values for the auxiliary nuclei labeled in M_{aux} with them for the unlabeled background

➔ image-level adaptation is able to avoid false-negative predictions by alleviating the domain bias on global visual information, such as curve, texture, and illumination

Method

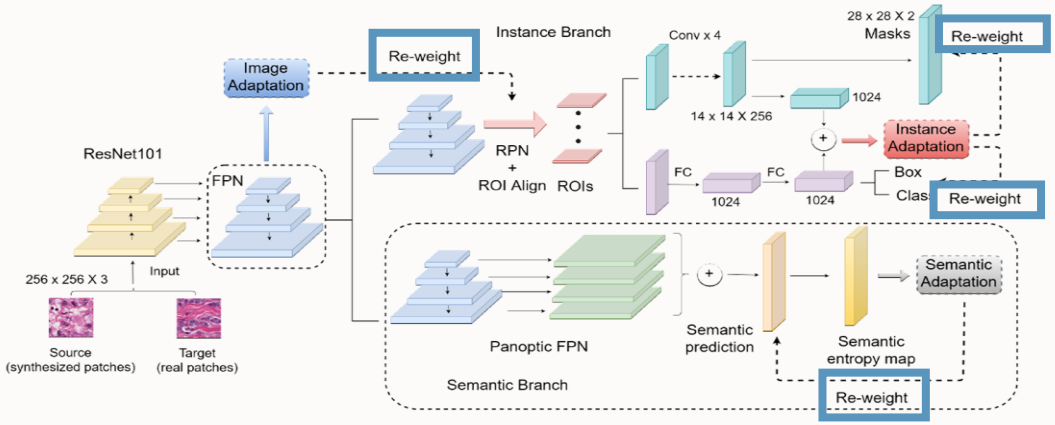
Panoptic Level Domain Adaptation



- Propose a semantic-level adaptation to induce the model to learn domain-invariant features based on the relationship between the foreground and background
- By incorporating semantic- and instance-level adaptation, the model can learn domain-invariant features at the panoptic level

Method

Task Re-weighting Mechanism



$$L_{rw} = \min\left(\frac{1-p_s}{p_s}, \beta\right) L$$

p_s	final task prediction of target domains
L	task-specific loss function
β	threshold value
L_{rw}	re-weighted task-specific loss

- The cross-domain discrepancies of these feature maps are still large in some training iterations
- Propose a task re-weighting mechanism
- According to the prediction of the domain discriminator, add a trade-off weight for each task-specific loss function method

Method

Network Overview and Training Details

- Overall loss function

task re-weighting

$$L_{pdam} = \alpha_{img} L_{rpn} + \alpha_{ins} L_{det} + \alpha_{sem} L_{(sem-seg)} + \alpha_{da} \left(L_{(img-da)} + L_{(sem-da)} + L_{(ins-da)} \right)$$

cross entropy losses
for domain classification at image

L_{rpn} loss function for the RPN

L_{det} loss of class & bounding box & instance mask prediction (Mask R-CNN)

$L_{(sem-seg)}$ cross entropy loss for semantic segmentation

Experiment & Conclusion

Dataset Description and Evaluation Metrics

- Dataset
  - Kumar (histopathology datasets)
  - TNBC (histopathology datasets)
  - BBBC039V1 (fluorescence microscopy dataset)

- Evaluation metrics
  - aggregated Jaccard Index
  - object-level F1 score
  - pixel-level F1 score

Experiment & Conclusion

Experiment Setting

- Two nuclei segmentation tasks
  - adapting from BBBC039V1 to Kumar
  - adapting from BBBC039V1 to TNBC

  As the source domain in two experiments, 100 training images and 50 validation images from BBBC039V1 are used
- Preprocessing
  - normalized into range [0, 255]
  - cropping, rotation, scaling, and flipping
  - patches with fewer than 3 objects are removed
  - inverse the pixel value of foreground nuclei and background for all source fluorescence microscopy patches

Experiment & Conclusion

Comparison Experiments

	$BBBC039 \rightarrow Kumar$			$BBBC039 \rightarrow TNBC$		
Methods	AJI	Pixel-F1	Object-F1	AJI	Pixel-F1	Object-F1
CyCADA [15]	0.4447 ± 0.1069	0.7220 ± 0.0802	0.6567 ± 0.0837	0.4721 ± 0.0906	0.7048 ± 0.0946	0.6866 ± 0.0637
Chen *et al.* [4]	0.3756 ± 0.0977	0.6337 ± 0.0897	0.5737 ± 0.0983	0.4407 ± 0.0623	0.6405 ± 0.0660	0.6289 ± 0.0609
SIFA [2]	0.3924 ± 0.1062	0.6880 ± 0.0882	0.6008 ± 0.1006	0.4662 ± 0.0902	0.6994 ± 0.0942	0.6698 ± 0.0771
DDMRL [21]	0.4860 ± 0.0846	0.7109 ± 0.0744	0.6833 ± 0.0724	0.4642 ± 0.0503	0.7000 ± 0.0431	0.6872 ± 0.0347
Hou *et al.* [16]	0.4980 ± 0.1236	0.7500 ± 0.0849	0.6890 ± 0.0990	0.4775 ± 0.1219	0.7029 ± 0.1262	0.6779 ± 0.0821
Proposed	$\mathbf{0.5610 \pm 0.0718}$	$\mathbf{0.7882 \pm 0.0533}$	$\mathbf{0.7483 \pm 0.0525}$	$\mathbf{0.5672 \pm 0.0646}$	$\mathbf{0.7593 \pm 0.0566}$	$\mathbf{0.7478 \pm 0.0417}$

Table 3. In comparison with other unsupervised methods on both two histopathology datasets.

	AJI			Pixel-F1		
Methods	seen	unseen	all	seen	unseen	all
CNN3 [24]	0.5154 ± 0.0835	0.4989 ± 0.0806	0.5083 ± 0.0695	0.7301 ± 0.0590	0.8051 ± 0.1006	0.7623 ± 0.0946
DIST [35]	0.5594 ± 0.0598	0.5604 ± 0.0663	0.5598 ± 0.0781	0.7756 ± 0.0489	0.8005 ± 0.0538	0.7863 ± 0.0550
Proposed	0.5432 ± 0.0477	0.5848 ± 0.0951	0.5610 ± 0.0982	0.7743 ± 0.0358	0.8068 ± 0.0698	0.7882 ± 0.0533
Upper bound [22]	0.5703 ± 0.0480	0.5778 ± 0.0671	0.5735 ± 0.0855	0.7796 ± 0.0419	0.8007 ± 0.0511	0.7886 ± 0.0531

Table 5. Comparison experiments between our UDA method and fully supervised methods, for BBBC039V1 to Kumar experiment. For CNN3 and DIST, the results of object-level F1 are unknown.

	AJI	Pixel-F1	Object-F1
w/o NI	0.5042 ± 0.1034	0.7336 ± 0.0839	0.6958 ± 0.0832
w/o TR	0.4969 ± 0.0972	0.7654 ± 0.0678	0.6923 ± 0.0778
w/o SEM	0.5046 ± 0.1065	0.7470 ± 0.0754	0.6965 ± 0.0805
proposed	0.5610 ± 0.0718	0.7882 ± 0.0533	0.7483 ± 0.0525

Table 4. Ablation study on BBBC039V1 to Kumar experiment. NI, TR, and SEM represent the nuclei inpainting mechanism, task re-weighting mechanism, and semantic branch, respectively.

Experiment & Conclusion

Comparison Experiments

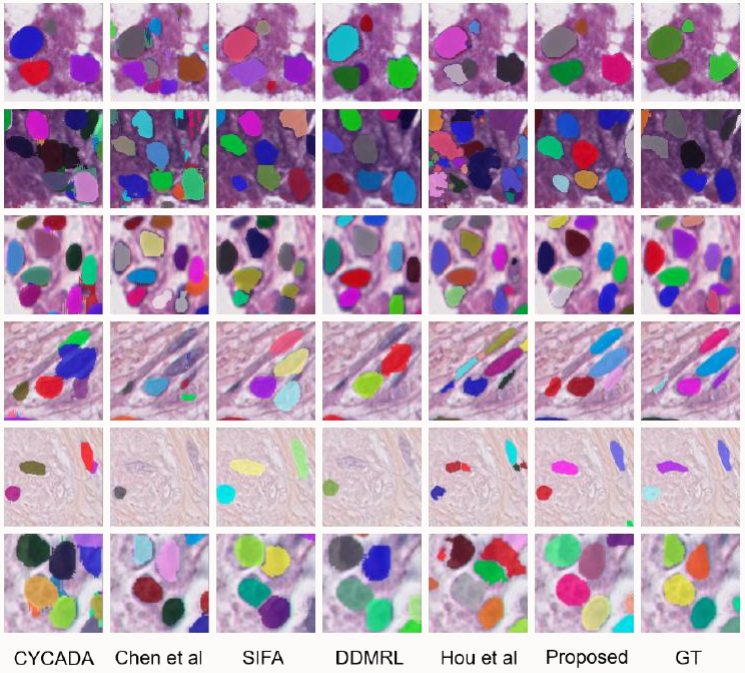


Figure 5. Visualization result for the comparison experiments experiment. The first 3 rows are from Kumar dataset, and the last 3 rows are from TNBC.

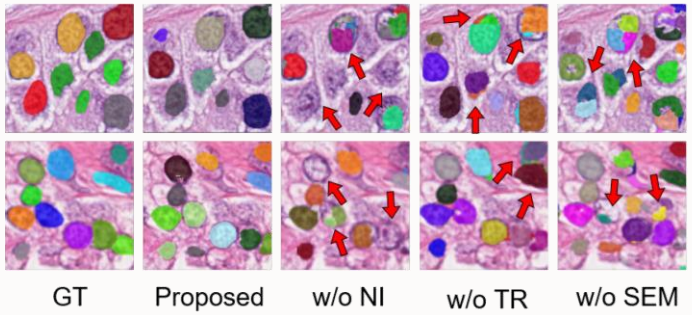


Figure 6. Visualization results for the ablation experiment. NI: nuclei inpainting mechanism; TR: task re-weighting mechanism; SEM: semantic branch.

Experiment & Conclusion

Conclusion

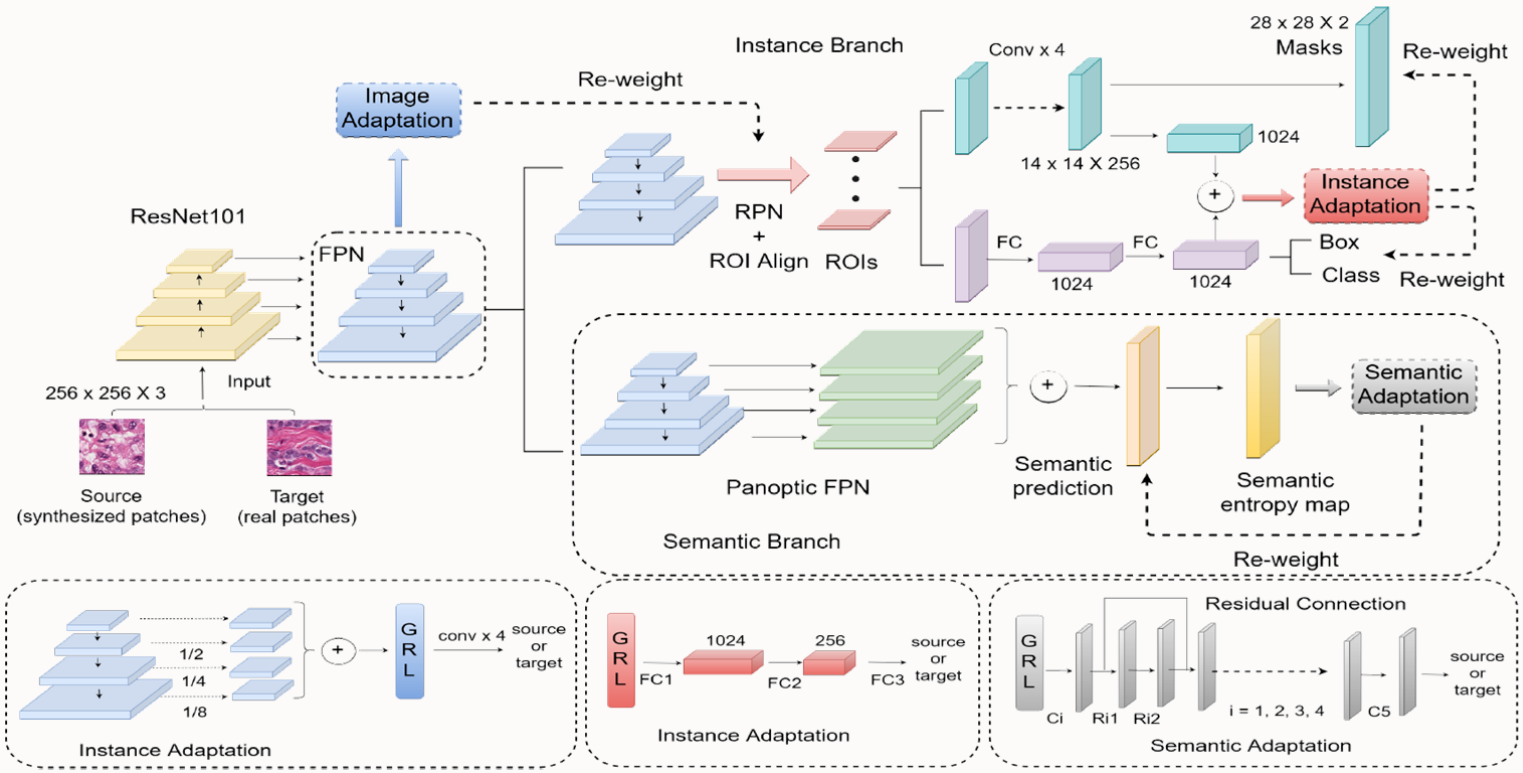
- Propose a CyC-PDAM architecture for UDA nuclei segmentation in histopathology images
  - Design a baseline architecture for UDA instance segmentation, including semantic-, image-, and instance-level adaptation
  - Nuclei inpainting mechanism is designed to remove the auxiliary objects in the synthesized images
  - Task re-weighting mechanism is proposed to reduce the bias

- Extensive experiments on three public datasets indicate proposed method outperforms the state-of-the-art UDA methods by a large margin and reaches the same level as the fully supervised methods

Thank you

Method

Overall architecture, CyC-PDAM



Method

CyCADA with Mask R-CNN

Name	Hyperparamaters	Output size
Input		$256 \times 8 \times 8$
Conv1	$k = (3, 3), s = 1, p = 1$	$256 \times 8 \times 8$
Conv2	$k = (3, 3), s = 1, p = 1$	$512 \times 8 \times 8$
Conv3	$k = (3, 3), s = 1, p = 1$	$512 \times 8 \times 8$
Conv4	$k = (1, 1), s = 1, p = 0$	$2 \times 8 \times 8$

Table 1. The parameters for each block in the image-level discriminator for PDAM. k, s, and p denote the kernel size, stride, and padding of the convolution operation, respectively.

Method

Panoptic Level Domain Adaptation

Name	Hyperparamaters	Output size
Input		$2 \times 256 \times 256$
C1	$k = (7, 7), s = 2, p = 3$	$64 \times 128 \times 128$
R11 and R12	$k = (3, 3), s = 1, p = 1$	$64 \times 128 \times 128$
C2	$k = (5, 5), s = 2, p = 2$	$128 \times 64 \times 64$
R21 and R22	$k = (3, 3), s = 1, p = 1$	$128 \times 64 \times 64$
C3	$k = (5, 5), s = 2, p = 2$	$256 \times 32 \times 32$
R31 and R32	$k = (3, 3), s = 1, p = 1$	$256 \times 32 \times 32$
C4	$k = (5, 5), s = 2, p = 2$	$512 \times 16 \times 16$
R41 and R42	$k = (3, 3), s = 1, p = 1$	$512 \times 16 \times 16$
C5	$k = (1, 1), s = 1, p = 0$	$2 \times 16 \times 16$
Output		$2 \times 16 \times 16$

Table 2. The parameters for each block in the semantic-level discriminator for PDAM. k, s, and p follow the same convention as in Table 1.

Method

Network Overview and Training Details

$$\alpha_{da} = \frac{2}{1 + \exp(-10t)} - 1$$

t is the training progress and $t \in [0, 1]$

gradually changed from 0 to 1, to avoid the noise from the unstable domain discriminators in the early training stage

Experiment & Conclusion

Dataset Description and Evaluation Metrics

- aggregated Jaccard Index

$$\frac{\sum_{i=1}^{N}|G_i \cap P_M^i|}{\sum_{i=1}^{N}|G_i \cup P_M^i| + \sum_{F \in U}|P_F|}$$

P_M^i	generation of mask communication domain
G_i	ith nucleus from the ground truth
N	number of nuclei

- F1 score

$$2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\text{Precision} = \frac{\text{TruePositives}}{\text{TruePositives} + \text{FalsePositives}}$$

$$\text{Recall} = \frac{\text{TruePositives}}{\text{TruePositives} + \text{FalseNegatives}}$$