

基于全景域自适应和任务重赋权机制的显微图像无监督实例分割

Dongnan Liu¹, Donghao Zhang¹, Yang Song², Fan Zhang³, Lauren O’ Donnell³, Heng Huang⁴, Mei Chen⁵, and Weidong Cai¹

¹School of Computer Science, University of Sydney, Australia

²School of Computer Science and Engineering, University of New South Wales, Australia

³Brigham and Women’ s Hospital, Harvard Medical School, USA

⁴Department of Electrical and Computer Engineering, University of Pittsburgh, USA

⁵Microsoft Corporation, USA

摘要

在细胞核实例分割中的无监督的域自适应方法(UDA)对于数字病理学有重要作用。这种方法减轻了数据集间劳动密集的标记工作和域变换的负担。在本篇论文中,我们通过学习荧光显微图像,提出了一种在病理学图像中无监督的循环一致性全景域自适应掩模 R-CNN (CyC-PDAM) 的细胞核实例分割方法。具体来说,我们首先突出了一种细胞核修复机制,来移除综合图像中附带生成的对象。其次,我们设计了带有域鉴别器的语义分支,以期达到全景级别的域自适应。最后,为了避免源域带有的偏好特征影响,我们提出了一种任务再赋权机制,对于特定的任务动态的为其损失函数添加权重。在三个数据集上的结果表明,该方法的表现显著超过了当前最先进的 UDA 方法,对于全监督方法来说,该方法也有相似的表现。

1 介绍

在数字病理学 workflows 中,病理图像的细胞核实例分割是重要的一环。病理学家能通过有丝分裂的细胞数,每个细胞核的形态结构,细胞核群的空间

分布等对癌症做出预测和诊断 [7, 25, 5, 1, 34]。当前,基于监督学习的细胞核实例分割方法较为流行,因为这些方法在保持高准确度的情况下效率也高 [24, 35, 3, 9, 33, 50, 29, 28]。但是这些方法的性能高度依赖于大规模的训练数据,而这些数据又要求专业的数据标记。如图 1(b) 所示,复杂的细胞结构和庞大的图像大小导致这个过程耗时耗力。举例来说,标记一个带有 50 张图片,12M 像素的病理图像数据集需要花费一个病理学家 120-230 小时 [16]。而在实际的临床研究中,即使是 40 倍物镜下的全切片图像都包含有 1B 个像素 [10]。因此,研究不依赖于病理学标记的方法是必要的。这能帮助病理学家减少工作量,解决图像缺少病理学标记的问题。

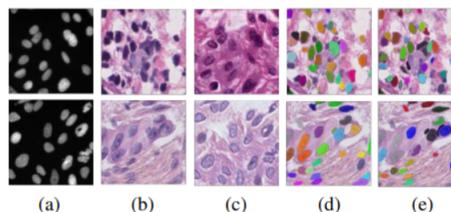


图 1: 我们提出的架构中的实例图像。(a) 荧光显微图像;(b) 真实病理图像;(c) 综合后的病理图像;(d) 我们提出的 UDA 方法生成的细胞核分割结果;(e) 真实标准。

最近提出的无监督域自适应方法通过对源域进

行有监督学习, 来获取针对无标记的目标域的高性能模型 [36, 8, 45]。当前, UDA 方法能减少源域和目标域特征分布图像的距离。此外, 其他方法关注的是从源域图像到目标域图像的像素-像素转换来对齐交叉域图像 [20, 52]。对于这些方法, 由于转换尚存在缺陷, 综合后的图像和真实图像间的分布还存在一些差异 [15, 2, 21]。

一些工作集合了图像变换和 UDA 方法的优点, 来学习目标域和综合后的图像间的域间差异特征 [15, 2, 21]。这些方法在 UDA 分类, 目标检测, 语义分割等任务方面达到了业界先进水平。但是, 当前不存在一种专为实例分割设计的 UDA 方法, 而直接将现有的 UDA 方法运用在细胞核实例分割任务时 [4, 14, 21], 仍然将会面临前述的挑战。首先, 现存的 UDA 目标监测方法关注于在图像层次上 (图像对比度, 亮度等) 和实例层次上 (对象规模, 样式) 减少域间偏差 [21, 4, 14]。这些方法忽略了语义层次上的域变换, 比如前景和背景的关系。其次, 这些 UDA 目标监测方法属于多任务学习范式, 同时优化不同的损失函数。如果在某些训练迭代中, 特征提取器不能生成域间差异特征, 那么基于任务损失函数的权重反向传播将导致源域中出现模型偏差。

为了解决前述问题, 我们提出了无监督的循环一致性全景域自适应掩模 R-CNN (CyC-PDAM) 模型。由于以往的 UDA 模型中没有为实例分割特别设计的方法, 我们基于掩膜 R-CNN [11], 扩展了 CyCADA 模型 [15] 以适用于实例分割。在我们的 CyC-PDAM 模型中, 我们首先提出了一种简单的细胞核修复机制来移除综合后的病理图像中的附带细胞核。其次, 受到全景分割架构的启发 [23, 22], 我们基于前景和背景的关系, 为域差异的特征提出了一种语义级别的适应模组。通过调和语义和实例级别的前景和背景域间差异特征, 我们提出的 CyC-PDAM 模型达到了全景级别的域自适应。此外, 我们还提出了一种任务再赋权机制以重置每个任务损失的权重。在训练时, 如果任务预测的特征不属于域间差异且不是源域差异的, 其损失的重要性将下调; 如果这些差异很难区分, 那么任务损失的重要性将上调。

为验证我们提出的 CyC-PDAM 架构的有效性,

我们在三个公共数据集上开展了实验, 其中, 在两个病理图像数据集上进行了无监督的细胞核实例分割实验, 在一个荧光显微图像数据库上进行了无监督的域自适应实验。与病理图像不同, 由于图像获取技术的不同, 以荧光显微图像为背景的图像中没有与细胞核类似的结构 (如图 1(a) 所示)。与病理图像相比, 荧光显微图像的手动标注更容易获取, 因此它被选为我们的源域。

我们的成果总结如下: (1) 我们提出了一种用于病理图像中, 无监督域自适应细胞核实例分割的 CyC-PDAM 模型。这是就我们所知的第一个 UDA 实例分割方法。(2) 我们提出了一个简单的细胞核修复机制, 用于移除综合后图像的假阳性目标。(3) 我们的 CyC-PDAM 模型通过集成实例水平的域适应和新提出的语义水平适应域模型, 在全景水平上生成域间差异特征。(4) 我们提出了一个任务重赋值机制, 用于减轻源域的域差异。(5) 与当前最先进的 UDA 方法相比, 我们提出的 CyC-PDAM 范式表现更佳。此外, 与当前最先进的全监督细胞核分割方法相比, 我们的模型表现也同样有竞争力。

2 相关工作

2.1 自然图像的域自适应

域自适应的目标在于将从带有标签域中学习的知识转以到其他的不带标注的域中 [36]。最近, 基于在特征级别的内容以及在像素级别的表现, UDA 方法减少了跨域的差异。对于特征层的自适应, 域间差异特征的对抗学习 [8, 45], 最大均值差异最小化 [32], 局部图像校准 [48] 以及跨域协变量对齐等技术 [42] 被广泛引用用于分类任务中。此外, 在其他任务, 如语义分割 [46, 26] 和目标检测 [4, 21, 19, 47], 域自适应也被应用。在语义分割的任务中, 分割结果和中间的特征图必须是存在域间差异的 [26, 46, 44]。此外, ADVENT [46] 下、更进一步为源域和目标域的语义分割预测减少了香农信息熵, 从而减少了域间差异。对于目标检测来说, 一种由图像级和实例级适应组成的域自适应的快速 R-CNN 方法 [40] 通常被运用于全图像和每个对象的域差异特征 [4, 21, 14]。另一

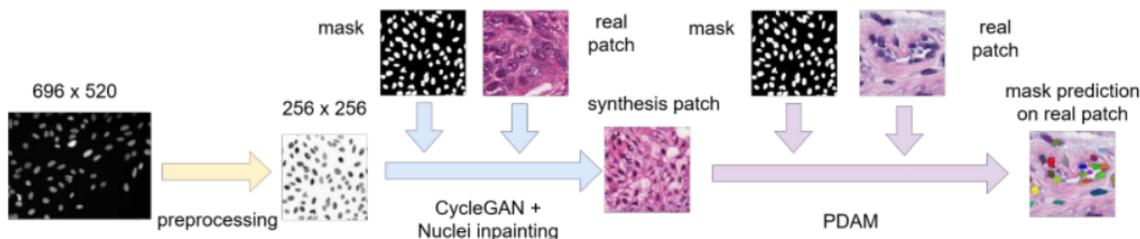


图 2: 我们提出的 CyC-PDAM 的总体架构。真实病理图片块中的标注在实际训练中没有被使用。

方面，图像对图像的转换通过生成类似目标图像以及在其上训练特定任务导向的全监督模型，解决了像素级别的域自适应问题 [30, 17, 20, 52, 33, 37]。然而，由于不完美的转换，域间偏差仍然存在。此外，许多模型被运用于在像素级别上对齐特征层的适应，这是通过学习目标图像和合成图像间的域间差异特征达成的 [15, 21, 2]。

2.2 医学图像的域自适应

医学图像分析中的无监督域自适应很少被探究 [39, 51, 2, 18, 16]。[39] 和 [18] 利用基于 GAN 的架构解决了 UDA 病理图像分类问题。此外，DAM[6] 被用于在 CT 图像终点的 UDA 语义分割，以生成随域变化的中间特征和模型预测。在周期连贯重建的帮助下，TD-GAN[51] 和 SIFA[2] 被用于不同医疗图像的语义分割，包括像素级和特征级的自适应。但是，他们中没有一种是用于 UDA 细胞核实例分割的。即使 Hou 等人 [16] 提出用综合后的病理图像训练基于打磨器的 GAN 和细胞核实例分割模型，他们的范式仅仅只包含像素级别的自适应，并且仍然不能最小化在特征级别域间的间隔。在本次工作中，我们由此提出了一种为 UDA 细胞核实例分割的 CyC-PDAM 范式，这减轻了像素级核特征级的域间偏差。

3 具体方法

我们提出的架构是基于 CyCADA 模型，并将它和实例分割框架掩码 R-CNN 相融合。此外，我们通

过引入细胞核修复机制，全景级别的域自适应，任务重赋权机制等进一步改进了该模型。图 2 阐释了我们方法的整体架构。

3.1 带掩码 R-CNN 的 CyCADA

名称	超参数	输出大小
输入		$256 \times 8 \times 8$
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	$k = (1, 1), s = 1, p = 0$	$2 \times 8 \times 8$

表 1: PDAM 的图像级判别器中每块的参数， k 、 s 、 p 分别表示卷积操作的内核大小、跨度和填充大小。

由于没有针对实例级别分割的 UDA 架构，我们首先设计了一个领域自适应的掩码 R-CNN 网络。我们架构模型的基础是 ResNet[12] 和特征金字塔网络 (FPN)[27]。受到之前用于对象检测的 UDA 方法 [4, 21] 的启发，我们在 FPN 网络之后加入了一个图像层自适应的区分器，在实例分支之后加入了一个实力层自适应的区分器（如图 3 所示）。对于图像层的自适应，FPN 输出的多像素特征图首先被降采样并平均池化到 8×8 的尺度，之后对其进行求和，并用于图像层的区分器。图像层的区分器由四个卷积层（详见表 1）和一个用于对抗学习的逆梯度层 (GRL) 组成。对于实力层的自适应，位于掩码分支的大小为 $14 \times 14 \times 256$ 的特征图被缩小并池化到大小为 $2 \times 2 \times 256$ ，之后被重新定义到大小为 1024×1 ，用于与从边框分支的 1024×1 的特征向量求和。实

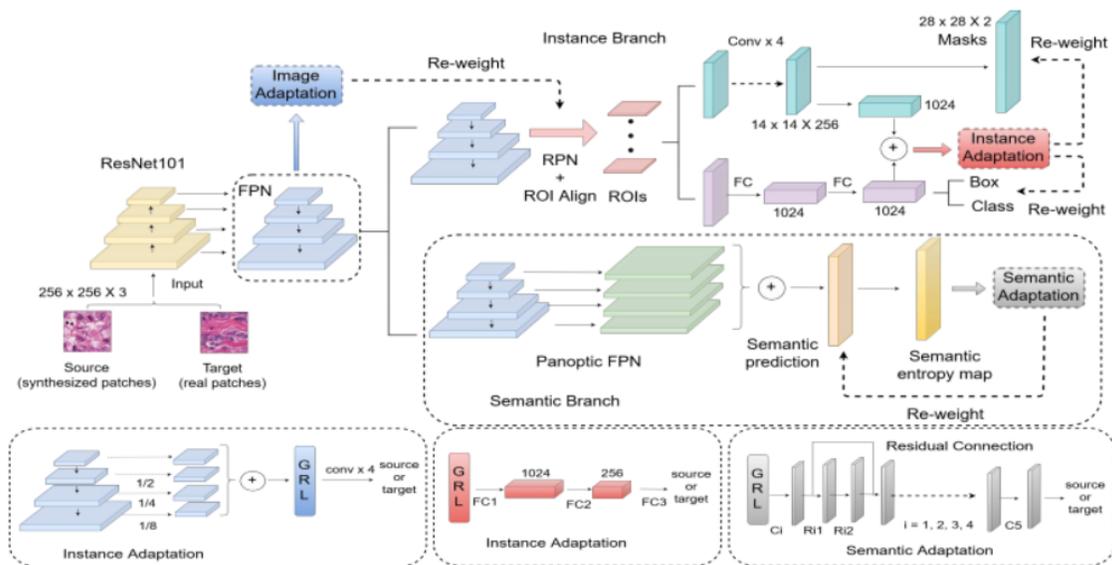


图 3: Panoptic Domain Adaptive Mask R-CNN (PDAM) 的详细说明。\$C_i\$ 和 \$FC\$ 分别代表卷积层和全连接层。\$Ri1\$ 和 \$Ri2\$ 分别指第 \$i\$ 个残差块中的第一和第二卷积层。为简洁起见，省略了每个卷积块后的 \$ReLU\$ 和归一化层。

力层的区分器由 3 个全相连层和一个 GRL 层组成，其输入即之前提到的特征和。

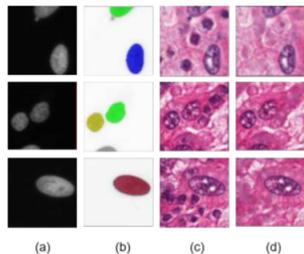


图 4: 细胞核修复机制有效性的可视化结果。(a) 原始的荧光显微图像块;(b) 相应的细胞核标注;(c) 从 CycleGAN 初始合成的图像;(d) 细胞核内修复机制后的最终合成图像。

3.2 细胞核修复机制

虽然 CycleGAN 对于病理图像的综合是有效的，但是由于域间间距过大，且源域与目标域之前的细胞核数目不兼容，生成图像的标签空间有时在从源域迁移时会改变。举例来说，在如图 4 所示的综合图像中出现了冗余的或不期望出现的细胞核。如果这些图像与原始标签直接被用于训练 CNN 网络时，模型将被强制要求把冗余的细胞核看作是背景，即使它们是真的细胞核。

因此，我们提出了一种附属的细胞核修复机制来

移除只出现在综合后的图像中，但又不带有标签的细胞核。我们用 \$S_{raw}\$ 表示原始的由 CycleGAN 网络综合后的病理图像，其对应的掩码用 \$M\$ 表示，我们首先获取所有附属生成的细胞核的掩码预测 \$M_{aux}\$，如下公式所示：

$$M_{aux} = (otsu(S_{raw}) \cup M) - M \quad (1)$$

其中 \$otsu(S_{raw})\$ 表示基于 Ostu 方法对 \$S_{raw}\$ 的二元分割方法。在 \$M_{aux}\$ 中，只有不带标注的附属细胞核有标签。之后，我们得到去除这些细胞核之后的新生成的图像 \$S_{inp}\$，这可以被表示为：

$$S_{inp} = inp(S_{raw}, M_{aux}) \quad (2)$$

其中 \$inp\$ 是一个基于快速行进的修复对象的方法 [43]，这是通过替换在 \$M_{aux}\$ 中被标记的附属细胞核的像素值，用未被标签的背景替换。图 4 表示了我们提出的细胞核修复机制的视觉有效性。但是，有些背景材料在 \$M_{aux}\$ 中被标记为假阳性预测。之间修复这些细胞核会让综合后图像的材质和样貌显得不真实，从而进一步拉大综合图像和真实图像的域间差异。然而，图像级别的域自适应能够通过减缓全

局信息的域间差异来解决该问题，例如曲线、质地和光照。我们细胞核修复机制是省时的，只需要 0.09 秒就能处理一张 256×256 的综合后的病理图像。

3.3 全景级别的域自适应

我们定义一张图像的语义级别特征为其前景和后景的关系。除了图像和特征级别的域差异之外，在语义级别的域漂移也同样存在。由于在综合后和实际的病理图像中的细胞核对象和背景存在差异，在 3.1 节提到的域自适应的掩码 R-CNN 方法会受到语义级别特征的域差异的负面影响，因为掩码 R-CNN 只关注每个对象的局部特征，并且缺少对于整个图像的语义级别的理解。受到之前全景分割方法统一语义和实例分割来对全局和局部特征进行处理的启发，我们提出一种语义级别的自适应方法，来促使模型通过前景和后景的关系学习域间差异特征。通过结合语义和实例级别的适应，我们的全景域适应方法减少了全局和局部之间的跨域差异。

如图 3 所示，语义分割预测的语义分支被加入到了 FPN 的输出中。我们的语义分支与 [21] 的实现方法相同。由于荧光显微图像和病理图像都可以从组织样本中获取，并且他们都表示出互补和相互关联的信息，综合和实际的病理图像的语义分割标签有着强相似性。除此之外，对齐跨域的熵分布能帮助我们减小目标域的熵预测，这使得模型能够适用于目标图像 [46]。因此，我们使用了 softmax 语义预测的香农熵 [41] 来减少域间差异特性，从而在语义层学习特征。用 P 表示 softmax 的语义预测， $P \in (0, 1)$ ，其香农熵被定义为 $-p \log(p)$ 。

图 3 和表 2 表明了语义自适应层区分器的细节结构。我们用残差相连 CNN 块避免梯度消失 [12, 13]。为了使对抗学习更加稳定，相比于双线性插值，我们采用步伐卷积层来升采样。最后，域标签以一种 16×16 图块的形式被预测。由于图块尺寸较小，基于图块的域标签预测增加了训练样本的数目，以期避免过拟合。

名称	超参数	输出大小
输入		$2 \times 256 \times 256$
C1	$k = (7, 7), s = 2, p = 3$	$64 \times 128 \times 128$
R11&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&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&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&R42	$k = (3, 3), s = 1, p = 1$	$512 \times 16 \times 16$
C5	$k = (1, 1), s = 1, p = 0$	$2 \times 16 \times 16$
输出		$2 \times 16 \times 16$

表 2: PDAM 的语义级判别器中每块的参数， k 、 s 、 p 的含义与表 1 中相同。

3.4 任务重赋权机制

在之前的 UDA 方法中，特定任务的损失函数是基于源域预测的。即使一些对抗域的区分器被用于确保预测特征图是有域差异的，这些特征的跨域差异在一些训练迭代中，当特征与域区分器的决定边界很远时，仍然很大。如果特定任务的损失函数能以更新的方式用这些容易区分的特征优化模型，当用目标数据检测模型时，这些模型的差异将偏向原图像。因此，我们提出一种任务中赋权机制，根据域区分器的预测来为每个特定任务损失函数加入一个取舍权重。用 p_s 和 p_t 表示在最后任务预测属于源和目标域的特征图概率，用 L 表示特定任务的损失函数，这样，重赋权的特定任务损失函数 L_{rw} 为：

$$L_{rw} = \min\left(\frac{p_t}{p_s}, \beta\right) L = \min\left(\frac{1-p_s}{p_s}, \beta\right) L \quad (3)$$

其中 β 是一个阈值，用于避免当 $p_s \rightarrow 0$ 时， $\frac{1-p_s}{p_s}$ 变得太大而导致模型崩溃。根据等式 3，如果决定任务预测的特征图属于源域时 ($p_s \rightarrow 1$)，损失函数将会被减少权重，来减少从模型中学习到的源偏差特征。如图 3 所示，区域提议网络 (RPN) 的损失函数，语义分支和实例分支通过对图像，语义和实例层的域区分器的预测来重赋权。

3.5 网络总览和训练细节

在我们提出的 CyC-PDAM 中, CycleGAN 和最开始的工作 [52] 有着相同的实现。当训练 CycleGAN 时, 最开始的 1/2 的训练中初始学习率被设为 0.0001, 后半部分的 1/2 几乎将为 0。

PDAM 模型通过一个批大小为 1, 每批含有 2 张图像的批训练, 其中一张来自于源域, 而另一张来自于目标域。由于每批的大小较小, 我们将传统的批归一层用组归一层 [49] 代替, 默认组大小为 32, 与 [49] 相同。

PDAM 的总损失函数定义如下:

$$L_{pdam} = \alpha_{img}L_{rpn} + \alpha_{ins}L_{det} + \alpha_{sem}L_{(sem-seg)} + \alpha_{da} (L_{(img-da)} + L_{(sem-da)} + L_{(ins-da)}) \quad (4)$$

其中 L_{rpn} 是 RPN 的损失函数, L_{det} 是族, 方框, 和掩码 R-CNN 实例掩码预测的损失。 $L_{(sem-seg)}$ 是实例分割的交叉熵损失。 $L_{(img-da)}$, $L_{(sem-da)}$ 和 $L_{(ins-da)}$ 是域分类器在图像、语义、实例层的交叉熵损失。 α_{img} , α_{ins} , 和 α_{isem} 根据等式 3 的任务重赋值来计算。我们在实验中将 β 设为 2。 α_{da} 按照如下方式更新:

$$\alpha_2 = \frac{2}{1 + \exp(-10t)} - 1 \quad (5)$$

其中 t 是训练进度, $t \in [0, 1]$ 。因此 α_{da} 逐渐从 0 变换到 1, 来避免早期阶段不稳定的域区分器带来的噪音。

在训练中, PDAM 模型由 SGD 优化, 以 0.9 的动量, 0.001 的权重降低。最开始的学习率为 0.002, 之后 500 次迭代后线性增加。当达到总训练迭代的 3/4 时, 学习率降为 0.0002。在推断过程中, 只有原始掩码 R-CNN 架构采用了适应后权重, 所有用于测试的超参数都在确认集中被优化。我们的所有实验都用 Pytorch 实现, 用到两块 NVIDIA GeForce 1080Ti GPUs。

¹<https://data.broadinstitute.org/bbbc/BBBC039/>

4 实验

4.1 数据集描述和评价标准

我们提出的架构在三个数据集上进行了确认, 分别叫做 Kumar[24], TNBC[35], BBBC039V1[31]。其中, Kumar 和 TNBC 都是病理图像数据集, 而 BBBC039V1 是一个荧光显微图像数据集。Kumar 数据集是从癌症基因图集 (TCGA) 中选取的, 物镜放大倍数为 40 倍, 包含了从 30 张来自不同病人的全切片图像中截取的 30 张带标注的 1000*1000 像素的图片块。所有的这些图片都来自于 18 个不同的医院和 7 种不同的器官 (胸部, 肝脏, 肾脏, 前列腺, 膀胱, 结肠, 胃部)。与在 Kumar 中表现出的疾病多样性不同, TNBC 数据集尤其关注三阴性乳腺癌 (TNBC) [35]。在 TNBC 数据集中, 有从 11 个来自 Curie 研究所不同病人的 50 张 512 × 512 像素的图像块, 物镜放大倍数为 40 倍。BBBC039V1 数据集有关于在高通量化学浆筛渣后的 U2OS 细胞 [31]。它包含 200 个 520 × 696 像素的生物活性化合物图像, 也包含单视野下的 DNA 染色通道数据。

在模型评价方面, 我们采用了三种被广泛运用的像素级和实物级的评价指标。累积雅卡指数 (AJI) 是用于实物级别评价的雅卡指数的扩展 [24]。实物级别的 F1 分数是每个实物介于精确和回归的调和平均数。对于像素级别的评价, 我们采用了用于二分预测的像素级别的 F1 分数。

4.2 实验设置

我们进行了两个细胞核分割实验, 将在 BBBC039V1 的结果迁移到 Kumar 上, 从 BBBC039V1 的时候结果迁移到 TNBC 上。作为两个实验中的源域, BBBC039V1 中的 100 张图片作为训练集, 50 张图片作为确认集, 并进行了官方的数据分散¹。Kumar 和 TNBC 的标注没有在训练 UDA 的过程中使用, 只被用于评价。

对荧光显微图像的处理分为 3 步。首先, 所有的图像都被归一化到范围 [0,255]。其次, 10K 个 256 × 256 大小的图片块被随机的从 100 个训练图

像中剪切, 随后进行了图像增强, 诸如旋转, 缩放和翻转来避免过拟合。第三步, 移除少于三个目标的图片块。为了更好的综合类似目标的病理图像, 我们最后对所有源荧光显微图片块逆转前景和后景的像素值。对于验证来说, 50 张在 BBBC039V1 验证集的图片通过 CycleGAN 网络的迁移学习来综合病理图像并进行细胞核修复机制。

当目标域是 Kumar 数据集时, 我们采用了和之前工作 [24, 35] 一致的数据分割方法, 并将 16 张图片作为训练图片, 14 张图片作为测试图片。当训练模型时, 10K 个 256×256 大小的图片块被随机的从 16 个训练病理图像中剪切, 随后进行了图像增强, 诸如旋转, 缩放和翻转来避免过拟合。当目标域是 TNBC 数据集时, 10K 个 256×256 大小的图片块被随机的从训练病理图像中提取, 随后进行了基本的诸如翻转和旋转的数据增强操作。

4.3 对比实验

4.3.1 与非监督方法对比

在这一小结中, 我们提出的 CyC-PDAM 模型被用于与其他业界领先的 UDA 方法比较, 包括 CyCADA[15], Chen 等人的方法 [4], SIFA[2], 和 DDMRL[21]。由于最初的 CyCADA 关注于分类和语义分割, 我们结合基于 UDA 的掩码 R-CNN 方法将它扩展, 如 3.1 节所述。Chen 等人的方法 [4] 的初衷是被用于基于快速 R-CNN 的 UDA 目标检测, 通过适应图片集和实例级的特征实现。对于 UDA 实例分割, 我们用与我们架构相同的掩码 R-CNN 以及在 [4] 中提到的原始图片和实例级的自适应替换了原先基于快速 R-CNN 的 VGG16 模型。SIFA 是一种为 CT 和 MR 图像的 UDA 语义分割架构, 能够实现像素和特征级别的自适应。在我们实验中, 我们增加了分水岭算法来分离在 SIFA 语义分割预测中的突出对象, 并以此作为明显的对比。DDMRL[21] 从 UDA 目标检测的各种生成域中学习多重域间的变量特征, 此外, 就像 CyCADA 和 Chen 等人 [4] 的工作一样, 它也被扩展用于实例分割。我们还与 Hou 等人的工作进行了对比, 这项工作特别地被用于病理图像的无监督细胞核分割。它们利用从二分

细胞核掩膜随机生成的综合后的病理图像训练了一个多重任务 (分割, 检测, 打磨) 的 CNN 网络架构。

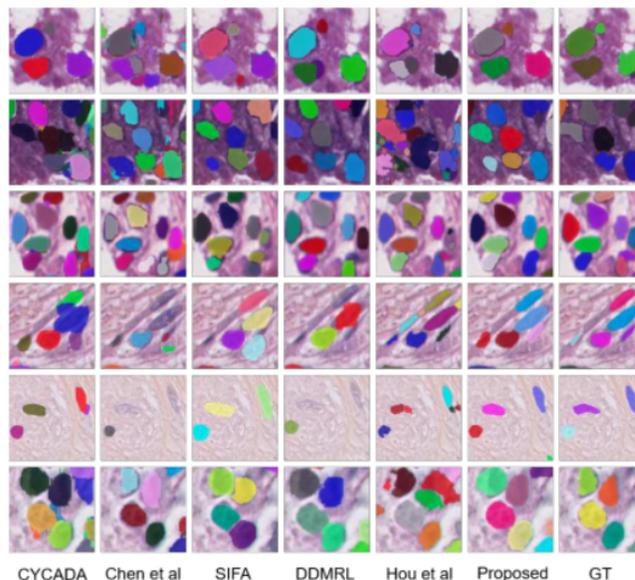


图 5: 对比实验的可视化结果. 前 3 行是来自 Kumar 数据集, 后 3 行来自 TNBC 数据集.

表 3 说明我们提出的方法在不同数据集上的表现大大超过了其他被比较的模型。此外, 我们还运用了单尾检验的 t 检验方法来证明我们模型的改善具有统计意义上的明显性, 其中 p 的值小于。Chen 等人 [4] 人学习了图像级别和实例级别的域间差异特征。然而, 由于荧光显微图像和真实病理图像之间存在的较大的差异, 特征级别的自适应不足以减少域间的差距。对于像素级别的自适应来说, 所有其他方法都表现得更好。与基线 CyCADA[15] 方法相比, 我们得 CyC-PDAM 模型改善了 6%-12%。这是由于我们提出方法的有效性, 包括细胞核修复机制, 全景级别的自适应, 和任务重复权机制。带有 UDA 语义分割结构的 SIFA[2] 关注于图像和语义级别的域间差异特征。由于在病理图像中存在大量的细胞核对象, SIFA 的有效性始终被实例级别的学习和自适应所限制。虽然 DDMRL[21] 只对图像级别的特征做适应, 通过适应多个域的知识, 它的表现和 CyCADA 的水平相同。在所有被比较的方法中, Hou 等人 [16] 的方法实现了第二佳的表现, 但我们的方法仍然在三个衡量指标上超出了它的表现。图 5 是所有被对比方法的可视化例子。

Methods	BBBC039 → Kumar			BBBC039 → TNBC		
	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 <i>et al.</i> [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 <i>et al.</i> [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	0.5610 ± 0.0718	0.7882 ± 0.0533	0.7483 ± 0.0525	0.5672 ± 0.0646	0.7593 ± 0.0566	0.7478 ± 0.0417

表 3: 在两个病理图像数据集上与其他非全监督方法相比的结果。

Methods	AJI			Pixel-F1		
	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

表 4: 对 BBBC039V1 到 Kumar 数据集的子组件研究。NI、TR 和 SEM 分别代表细胞核修复机制、任务重赋权机制和语义分支。

4.3.2 子组件研究

为了测试我们提出的 CyC-PDAM 架构所有组成结构的有效性,我们在 Kumar 数据集上进行了子组件实验。基于我们的 CyC-PDAM 架构,我们移除了细胞修复机制,任务重赋权机制,全景级别的适应语义分支,并用与 4.3.1 节相同的实验设置和数据集训练了子组件模型。表 4 和图 6 表明了子组件实验的细节结果。正如图 6 中所示,没有细胞核修复机制的方法倾向于忽视一些细胞核,这会增加假阴性预测的概率。此外,我们还注意到在 w/o NI 模型中还有一些错误分裂和合并的预测。这是因为增加的假阴性预测对于所有对象的空间分布是有害的,这会进一步影响语义级别自适应的有效性。在所有不包含任务重赋权机制的预测中 (w/o TR),由一些对象表现出不规则的大小。任务重赋权机制阻止模型受到源域本身的特性影响并能够移除这些影响。

因此,这会导致一些基于源域而带有偏差的预测。

与我们的方法相比,没有语义分支的模型 (w/o SEM) 不能学习到语义级域间差异的特征,包括细胞核对象的空间分布和关于背景的细节信息。因此,这些方法不仅保留了一些错误的分割和合并预测,还有一些假阳性和不完美的分割结果。如表 4 所示,在移除了这些模块之后,三个评价标准下的准确程度减少了 4%-6%。此外,单侧检验的 t 测试被用于计算我们提出的方法和子组件方法之间的 p 值。在

加上了这三个组件之后,改善在是统计上显著的 ($P < 0.05$),这进一步证明了我们提出方法的有效性。

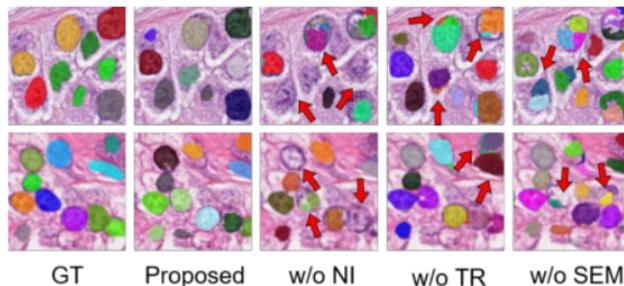


图 6: 子组件研究的可视化结果。NI: 细胞核修复机制; TR: 任务重赋权机制; SEM: 语义分支。

4.3.3 与全监督方法相比较

我们对 Kumar 数据集数据分割的方式与其他业界领先的全监督细胞核分割算法相同,我们对这些方法报告的结果与我们的结果对比。表 5 展示了我们提出的 UDA 架构和其他全监督的比对结果。CNN3[24] 是一个基于边缘的细胞核分割架构,它将细胞核当作除了前景和背景之外的第三类目标。DIST[35] 是一个基于距离图的回归模型。对于全景 FPN[22],我们按照 CNN3 和 DIST 方法直接用 16 个真实的病理图像块训练它,训练结果作为我们非监督方法的上界。在 Kumar 数据库的测试图像被划分为两个子集,一个包含来自于 4 个器官的 8 张图片,称为训练集,一个包含来自 3 个器官的 6 张图

Methods	AJI			Pixel-F1		
	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

表 5: 在 BBBC039V1 到 Kumar 数据集上的实验, 我们的 UDA 方法和完全监督方法的结果对比。对于 CNN3 和 DIST, 对象级 F1 的结果未知。

片, 且不为训练集所知, 被称为不可知集。

如表 5 所示, 我们提出方法的表现优于全监督的 CNN3 和 DIST。这是因为我们提出的方法能够处理在局部级别的每一个 ROI, 但是 CNN3 和 DIST 只能在全局语义级处理图像。通过把语义级别的特征适应到前景和背景, 我们方法的表现与全监督的全景 FPN 在像素级别的 F1 分数达到同一水平。即使我们的 AJI 比全监督的全景 FPN 略微低一些, 我们注意到我们的方法在不可知集上的表现更好。这是由于我们提出的 CyC-PDAM 关注于学习域间差异, 并且避免被域内未知器官的测试照片偏见影响。这些结果表明, 虽然荧光显微图像和病理图像中存在着较大差异, 但是我们提出的 UDA 架构仍能成功的减少它们之间的域间差异, 而且与需要病理图像中细胞核标注的全监督模型相比, 性能更好。

5 结论

我们提出了一种用于 UDA 病理图像中细胞核分割的 CyC-PDAM 架构。我们首先设计了用于 UDA 实例分割的基线架构, 包括对表现级、图像级、实例级的自适应。之后, 我们设计了一种细胞核的修复机制来移除综合后图像中附带的物体, 以此来避免假阴性预测的问题。在特征级别的自适应中, 我们用一个语义分支来在考虑前景和背景的条件下适应特征, 当与语义和实例级的自适应相结合时, 模型能够学习全景级别的域间差异特征。此外, 我们还提出了一种任务重赋权机制来减少域本身的偏见。在三个公共数据集上的广泛实验表明我们提出的方法表现大大超过了当前业界领先的 UDA 算法, 并且与全监督算法处于统一水平。从更高的角度看, UDA 实例分割问题不仅限于病理图像分析。鉴于该工作中接近全监督方法的可观结果, 我们认为我们提出的

方法也可以给一般性图像分析应用带来启发。

参考文献

- [1] Ajay Basavanthally, Michael Feldman, Natalie Shih, Carolyn Mies, John Tomaszewski, Shridhar Ganesan, and Anant Madabhushi. Multi-field-of-view strategy for image-based outcome prediction of multi-parametric estrogen receptor-positive breast cancer histopathology: Comparison to oncotype dx. *Journal of pathology informatics*, 2, 2011.
- [2] Cheng Chen, Qi Dou, Hao Chen, Jing Qin, and Pheng-Ann Heng. Synergistic image and feature adaptation: Towards cross-modality domain adaptation for medical image segmentation. In Association for the Advancement of Artificial Intelligence (AAAI), pages 865–872, 2019.
- [3] Hao Chen, Xiaojuan Qi, Lequan Yu, Qi Dou, Jing Qin, and Pheng-Ann Heng. Dcan: Deep contour-aware networks for object instance segmentation from histology images. *Medical Image Analysis*, 36:135–146, 2017.
- [4] Yuhua Chen, Wen Li, Christos Sakaridis, Dengxin Dai, and Luc Van Gool. Domain adaptive faster R-CNN for object detection in the wild. In Computer Vision and Pattern Recognition (CVPR), pages 3339–3348, 2018.
- [5] Frederic Clayton. Pathologic correlates of survival in 378 lymph node-negative infiltrating ductal breast carcinomas. mitotic count is the best single predictor. *Cancer*, 68(6):1309–1317, 1991.
- [6] Qi Dou, Cheng Ouyang, Cheng Chen, Hao Chen, and PhengAnn Heng. Unsupervised cross-modality domain adaptation of convnets for biomedical image segmentations with adversarial loss. In International Joint Conferences

- on Artificial Intelligence (IJCAI), pages 691–697, 2018.
- [7] Christopher W Elston and Ian O Ellis. Pathological prognostic factors in breast cancer. I. the value of histological grade in breast cancer: experience from a large study with longterm follow-up. *Histopathology*, 19(5):403–410, 1991.
- [8] Yaroslav Ganin and Victor Lempitsky. Unsupervised domain adaptation by backpropagation. In *International Conference on Machine Learning (ICML)*, 2015.
- [9] Simon Graham, Quoc Dang Vu, Shan E Ahmed Raza, Ayesha Azam, Yee Wah Tsang, Jin Tae Kwak, and Nasir Rajpoot. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Medical Image Analysis*, 58:101563, 2019.
- [10] David A Gutman, Jake Cobb, Dhananjaya Somantha, Yuna Park, Fusheng Wang, Tahsin Kurc, Joel H Saltz, Daniel J Brat, Lee AD Cooper, and Jun Kong. Cancer digital slide archive: an informatics resource to support integrated in silico analysis of TCGA pathology data. *Journal of the American Medical Informatics Association*, 20(6):1091–1098, 2013.
- [11] Kaiming He, Georgia Gkioxari, Piotr Dollar, and Ross Girshick. Mask R-CNN. In *International Conference on Computer Vision (ICCV)*, pages 2980–2988, 2017.
- [12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.
- [13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. In *European Conference on Computer Vision (ECCV)*, pages 630–645. Springer, 2016.
- [14] Zhenwei He and Lei Zhang. Multi-adversarial faster-rcnn for unrestricted object detection. In *International Conference on Computer Vision (ICCV)*, pages 6668–6677, 2019.
- [15] Judy Hoffman, Eric Tzeng, Taesung Park, Jun-Yan Zhu, Phillip Isola, Kate Saenko, Alexei A Efros, and Trevor Darrell. Cycada: Cycle-consistent adversarial domain adaptation. In *International Conference on Machine Learning (ICML)*, 2018.
- [16] Le Hou, Ayush Agarwal, Dimitris Samaras, Tahsin M Kurc, Rajarsi R Gupta, and Joel H Saltz. Robust histopathology image analysis: To label or to synthesize? In *Computer Vision and Pattern Recognition (CVPR)*, pages 8533–8542, 2019.
- [17] Xun Huang, Ming-Yu Liu, Serge Belongie, and Jan Kautz. Multimodal unsupervised image-to-image translation. In *European Conference on Computer Vision (ECCV)*, pages 172–189, 2018.
- [18] Yue Huang, Han Zheng, Chi Liu, Xinghao Ding, and Gustavo K Rohde. Epithelium-stroma classification via convolutional neural networks and unsupervised domain adaptation in histopathological images. *IEEE Journal of Biomedical and Health Informatics*, 21(6):1625–1632, 2017.
- [19] Naoto Inoue, Ryosuke Furuta, Toshihiko Yamasaki, and Kiyoharu Aizawa. Cross-domain weakly-supervised object detection through progressive domain adaptation. In *Computer Vision and Pattern Recognition (CVPR)*, pages 5001–5009, 2018.
- [20] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Computer Vision and Pattern Recognition (CVPR)*, pages 5967–5976. IEEE, 2017.
- [21] Taekyung Kim, Minki Jeong, Seunghyeon Kim, Seokeon Choi, and Changick Kim. Diversify and match: A domain adaptive representation learning paradigm for object detection. In *Computer Vision and Pattern Recognition (CVPR)*, pages 12456–12465, 2019.
- [22] Alexander Kirillov, Ross Girshick, Kaiming He, and Piotr Dollar. Panoptic feature pyramid networks. In *Computer Vision and Pattern Recognition (CVPR)*, pages 6399–6408, 2019.
- [23] Alexander Kirillov, Kaiming He, Ross Girshick, Carsten Rother, and Piotr Dollar. Panoptic segmentation. In *Computer Vision and Pattern Recognition (CVPR)*, pages 9404–9413, 2019.

- [24] Neeraj Kumar, Ruchika Verma, Sanuj Sharma, Surabhi Bhargava, Abhishek Vahadane, and Amit Sethi. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Transactions on Medical Imaging*, 36(7):1550–1560, 2017.
- [25] V Le Doussal, M Tubiana-Hulin, S Friedman, K Hacene, F Spyrtos, and M Brunet. Prognostic value of histologic grade nuclear components of Scarff-Bloom-Richardson (SBR). an improved score modification based on a multivariate analysis of 1262 invasive ductal breast carcinomas. *Cancer*, 64(9):1914–1921, 1989.
- [26] Yunsheng Li, Lu Yuan, and Nuno Vasconcelos. Bidirectional learning for domain adaptation of semantic segmentation. In *Computer Vision and Pattern Recognition (CVPR)*, pages 6936–6945, 2019.
- [27] Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Computer Vision and Pattern Recognition (CVPR)*, pages 2117–2125, 2017.
- [28] Dongnan Liu, Donghao Zhang, Yang Song, Heng Huang, and Weidong Cai. Cell r-cnn v3: A novel panoptic paradigm for instance segmentation in biomedical images. *arXiv preprint arXiv:2002.06345*, 2020.
- [29] Dongnan Liu, Donghao Zhang, Yang Song, Chaoyi Zhang, Fan Zhang, Lauren ODonnell, and Weidong Cai. Nuclei segmentation via a deep panoptic model with semantic feature fusion. In *International Joint Conferences on Artificial Intelligence (IJCAI)*, pages 861 – 868. AAAI Press, 2019.
- [30] Ming-Yu Liu, Thomas Breuel, and Jan Kautz. Unsupervised image-to-image translation networks. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 700 – 708, 2017.
- [31] Vebjorn Ljosa, Katherine L Sokolnicki, and Anne E Carpenter. Annotated high-throughput microscopy image sets for validation. *Nature Methods*, 9(7):637–637, 2012.
- [32] Mingsheng Long, Yue Cao, Jianmin Wang, and Michael I Jordan. Learning transferable features with deep adaptation networks. *International Conference on Machine Learning (ICML)*, 2015.
- [33] Faisal Mahmood, Daniel Borders, Richard Chen, Gregory N McKay, Kevan J Salimian, Alexander Baras, and Nicholas J Durr. Deep adversarial training for multi-organ nuclei segmentation in histopathology images. *IEEE Transactions on Medical Imaging*, 2018.
- [34] Sidra Nawaz and Yinyin Yuan. Computational pathology: Exploring the spatial dimension of tumor ecology. *Cancer letters*, 380(1):296–303, 2016.
- [35] Peter Naylor, Marick Lae, Fabien Reyal, and Thomas Walter. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE Transactions on Medical Imaging*, 2018.
- [36] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359, 2009.
- [37] Taesung Park, Ming-Yu Liu, Ting-Chun Wang, and Jun-Yan Zhu. Semantic image synthesis with spatially-adaptive normalization. In *Computer Vision and Pattern Recognition (CVPR)*, pages 2337–2346, 2019.
- [38] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. *NeurIPS 2017 Autodiff Workshop*, 2017.
- [39] Jian Ren, Ilker Hacihaliloglu, Eric A Singer, David J Foran, and Xin Qi. Adversarial domain adaptation for classification of prostate histopathology whole-slide images. In *International Conference On Medical Image Computing Computer Assisted Intervention (MICCAI)*, pages 201–209. Springer, 2018.
- [40] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 91–99, 2015.

- [41] Claude Elwood Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379-423, 1948.
- [42] Baochen Sun, Jiashi Feng, and Kate Saenko. Return of frustratingly easy domain adaptation. In *Association for the Advancement of Artificial Intelligence (AAAI)*, 2016.
- [43] Alexandru Telea. An image inpainting technique based on the fast marching method. *Journal of Graphics Tools*, 9(1):23-34, 2004.
- [44] Yi-Hsuan Tsai, Wei-Chih Hung, Samuel Schulter, Kihyuk Sohn, Ming-Hsuan Yang, and Manmohan Chandraker. Learning to adapt structured output space for semantic segmentation. In *Computer Vision and Pattern Recognition (CVPR)*, pages 7472-7481, 2018.
- [45] Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell. Adversarial discriminative domain adaptation. In *Computer Vision and Pattern Recognition (CVPR)*, pages 7167 - 7176, 2017.
- [46] Tuan-Hung Vu, Himalaya Jain, Maxime Bucher, Matthieu Cord, and Patrick Perez. ADVENT: Adversarial entropy minimization for domain adaptation in semantic segmentation. In *Computer Vision and Pattern Recognition (CVPR)*, pages 2517-2526, 2019.
- [47] Tao Wang, Xiaopeng Zhang, Li Yuan, and Jiashi Feng. Fewshot adaptive faster r-cnn. In *Computer Vision and Pattern Recognition (CVPR)*, pages 7173-7182, 2019.
- [48] Jun Wen, Risheng Liu, Nenggan Zheng, Qian Zheng, Zhefeng Gong, and Junsong Yuan. Exploiting local feature patterns for unsupervised domain adaptation. In *Association for the Advancement of Artificial Intelligence (AAAI)*, volume 33, pages 5401-5408, 2019.
- [49] Yuxin Wu and Kaiming He. Group normalization. In *European Conference on Computer Vision (ECCV)*, pages 3-19, 2018.
- [50] Donghao Zhang, Yang Song, Dongnan Liu, Haozhe Jia, Siqi Liu, Yong Xia, Heng Huang, and Weidong Cai. Panoptic segmentation with an end-to-end cell r-cnn for pathology image analysis. In *International Conference On Medical Image Computing Computer Assisted Intervention (MICCAI)*, pages 237-244. Springer, 2018.
- [51] Yue Zhang, Shun Miao, Tommaso Mansi, and Rui Liao. Task driven generative modeling for unsupervised domain adaptation: Application to x-ray image segmentation. In *International Conference On Medical Image Computing Computer Assisted Intervention (MICCAI)*, pages 599-607. Springer, 2018.
- [52] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycleconsistent adversarial networks. In *International Conference on Computer Vision (ICCV)*, pages 2223-2232, 2017.