

面向三维点云的域自适应学习

范文辉¹, 林茜¹, 罗欢¹, 郭文忠¹, 汪汉云², 戴晨光²

1. 福州大学 计算机与大数据学院, 福州 350108;

2. 信息工程大学 地理空间信息学院, 郑州 450001

摘要: 三维点云数据在自动驾驶、机器人和高精地图等领域得到了广泛应用。目前, 基于深度学习的三维点云数据处理主要基于有监督学习, 其算法性能依赖于大规模高质量的标注数据集。此外, 仅在单一设备与场景中训练的三维点云数据处理模型难以应用于不同设备与环境, 泛化性能有限。因此, 如何减少三维点云标注数据集的需求以及提高三维点云处理模型的适应性是当前三维点云数据处理面临的重要难题。作为迁移学习的一个重要分支, 域自适应学习旨在不同域间特征分布存在差异的情况下提高模型的适应性, 可为解决上述难题提供重要思路。为便于对点云域自适应学习领域进行更深入有效的探索, 本文主要从对抗学习、跨模态学习、伪标签学习、数据对齐及其他方法5个方面对近年来的三维点云域自适应学习方法进行了系统梳理与分类归纳, 并分析总结每类点云域自适应学习方法所具备的优势及面临的问题。最后, 对三维点云域自适应学习研究领域的未来发展进行了展望。

关键词: 遥感, 三维点云, 域自适应学习, 对抗学习, 跨模态学习, 伪标签学习, 数据对齐

中图分类号: TP391/P2

引用格式: 范文辉, 林茜, 罗欢, 郭文忠, 汪汉云, 戴晨光. 2024. 面向三维点云的域自适应学习. 遥感学报, 28(4): 825-842

Fan W H, Lin X, Luo H, Guo W Z, Wang H Y and Dai C G. 2024. Domain adaptation learning for 3D point clouds: A survey. National Remote Sensing Bulletin, 28(4): 825-842 [DOI: 10.11834/jrs.20233140]

1 引言

随着三维激光扫描技术的发展, 三维点云数据已经具备了采集成本低、获取速度快与数据精度高等优势, 并在自动驾驶、机器人和高精地图等领域得到了广泛应用 (Shi 等, 2022)。近年来, 深度学习在目标检测、形状分类、语义分割等三维点云数据处理重要任务上都取得了较大进展, 出现了多个重要的三维点云数据处理神经网络, 如 PointNet++ (Qi 等, 2017)、RandLA-Net (Hu 等, 2020)、KPCConv (Thomas 等, 2019)。另外, 为了进一步促进点云深度学习技术的发展, 学术界和工业界还公开发布了多个用于不同任务的三维点云数据集。例如, 点云目标检测任务数据集 Waymo (Sun 等, 2020)、NuScenes (Caesar 等, 2020) 和 KITTI (Geiger 等, 2012); 点云形状分类

任务数据集 ModelNet (Wu 等, 2015)、ShapeNet (Chang 等, 2015) 和 ScanNet (Dai 等, 2017); 语义分割任务数据集 Semantic3D (Hackel 等, 2017)、SemanticKITTI (Behley 等, 2019) 和 A2D2 (Geyer 等, 2020); 点云配准任务数据集 WHU-TLS (Dong 等, 2020) 和 Apollo-SouthBay (Lu 等, 2019); 虚拟合成点云数据集 MDLS (Yoon 等, 2019) 和 GTA-LiDAR (Wu 等, 2019)。

基于深度学习的三维点云数据处理方法的性能依赖于大规模高质量的标注数据集。例如, 目标检测任务中, 需要精确标注出场景中感兴趣目标的位置及尺寸; 语义分割任务中需要精确标注出每个三维点的类别信息; 配准任务中需要精确估计出两帧点云数据之间的几何变换参数。这些标注数据集的获取耗时耗力、代价昂贵。同时, 三维点云数据处理模型也难以应用于不同数据集,

收稿日期: 2023-05-04; 预印本: 2023-10-07

基金项目: 国家自然科学基金(编号:42271457); 福建省自然科学基金(编号:2023J01430); 嵩山实验室项目(纳入河南省重大科技专项管理体系)(编号:221100211000-02)

第一作者简介: 范文辉, 研究方向为三维点云处理。E-mail: fanwenhui99@qq.com

通信作者简介: 汪汉云, 研究方向为三维点云智能处理与理解。E-mail: why.scholar@gmail.com

泛化性能有限。因此,如何减少三维点云数据集标注的需求以及提高三维点云数据处理模型的适应性是当前三维点云数据处理面临的重要难题。域自适应DA(Domain Adaptation)学习作为迁移学习的一个重要分支,旨在不同域间特征分布存在差异的情况下提高模型的适应性,为解决此难题提供了一种重要思路(Zhuang等,2021)。具体地,当目标域难以获取足够的高质量标注数据进行模型的有效训练时,可将源域中已训练好的模型通过域自适应方法有效迁移至目标域中,进而完成目标域的相关任务。然而,不同域之间数据存在点密度不一致、点云目标不完整、点云目标

类内几何形状差异大等偏移情况。这些不同域间数据的差异性以及所采用域自适应方法的局限性,使得迁移后的目标域模型可能出现负迁移,难以适应目标域任务(Pan和Yang,2010)。

图1根据研究时间给出了点云域自适应学习相关研究的主要进展。另外,为了便于介绍与梳清点云域自适应学习领域的相关研究,本文将所有方法分为对抗学习、跨模态学习、伪标签学习、数据对齐以及其他方法5个类别。由图1可见,各个类别的方法在近几年都有相关论文发表,都是目前研究的热点方向。表1展示了各种点云域自适应学习方法在任务类别与源域数据类型的区别。

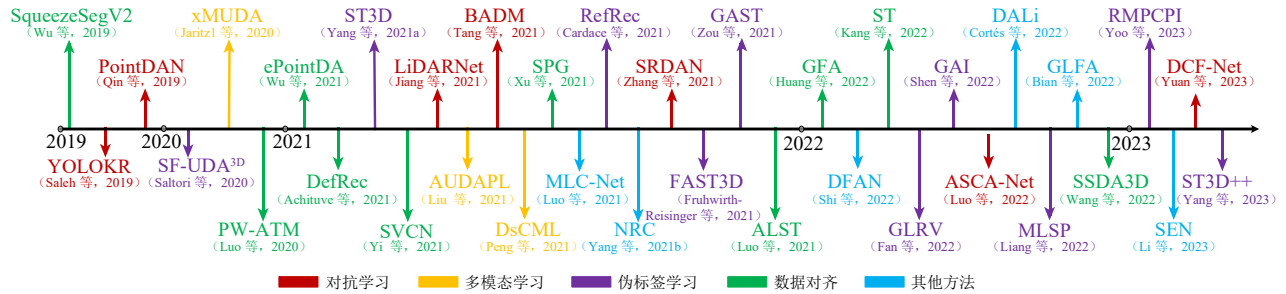


图1 点云域自适应学习研究主要进展(不同的颜色代表不同类别的点云域自适应方法)

Fig.1 Main advances in point cloud domain adaptation learning research adaptation methods(Different colors represent different categories of point cloud domain)

表1 不同点云域自适应方法的区别

Table 1 Differences between different point cloud domain adaptation methods

类别	方法	点云任务			源域类型	
		形状分类	物体检测	语义分割	合成数据	真实数据
对抗学习	PointDAN(Qin等,2019)	√			√	√
	YOLO _{kr} (Saleh等,2019)		√		√	
	LiDARNet(Jiang和Saripalli,2021)			√		√
	SRDAN(Zhang等,2021)		√			√
	BADM(Tang等,2021)	√			√	√
	ASCA-Net(Luo等,2022)	√				√
	DCF-Net(Yuan等,2023)			√		√
跨模态学习	xMUDA(Jaritz等,2020)			√		√
	AUDA(Liu等,2021)			√		√
	DsCML(Peng等,2021)			√		√
伪标签学习	SF-UDA ^{3D} (Saltori等,2020)		√			√
	ST3D(Yang等,2021a)		√			√
	RefRec(Cardace等,2021)	√			√	√
	FAST3D(Fruhvirth-Reisinger等,2021)		√			√
	GAST(Zou等,2021)	√			√	√

续表

类别	方法	点云任务			源域类型	
		形状分类	物体检测	语义分割	合成数据	真实数据
	GLRV(Fan等,2022)	√			√	√
	GAI(Shen等,2022)		√		√	√
	MLSP(Liang等,2022)	√		√	√	√
	ST3D++(Yang等,2023)		√			√
	RMPCPI(Yoo和Jun,2023)	√			√	√
数据对齐	SqueezeSegV2(Wu等,2019)			√	√	
	ePointDA(Zhao等,2021)			√	√	
	PW-ATM(Luo等,2020)			√		√
	DefRec(Achituve等,2021)	√				√
	SVCN(Yi等,2021)			√		√
	SPG(Xu等,2021)		√			√
	ALST(Luo等,2021a)	√		√	√	√
	GFA(Huang等,2022)	√			√	√
	ST(Kang等,2022)	√			√	√
	SSDA3D(Wang等,2022)		√			√
其他方法	NRC(Yang等,2021b)	√			√	√
	MLC-Net(Luo等,2021b)		√			√
	SEN(Li等,2023)	√		√		√
	DFAN(Shi等,2022)	√		√		√
	DALi(Cortés等,2022)		√			√
	GLFA(Bian等,2022)			√		√

2 问题定义

本文首先给出域、任务等相关概念的定义，随后给出点云域自适应学习的定义。

定义1（域）一个域 D 由特征空间 X 和边缘概率分布 $P(X)$ 组成，即 $D = \{X, P(X)\}$ 。其中， X 是一个样本集，表示为 $X = \{\mathbf{x}_i \in X, i = 1, 2, \dots, n\}$ ， \mathbf{x}_i 为特征空间 X 的一个向量。

定义2（任务）一个任务 T 由一个标签空间 Y 和一个类别预测函数 f 组成，即 $T = \{Y, f\}$ 。其中，类别预测函数 f 从样本数据中学习得到。

定义3（点云域自适应学习）给定源域 $D_s = \{X_s, P(X_s)\}$ 和源域任务 $T_s = \{Y_s, f_s\}$ ，目标域 $D_t = \{X_t, P(X_t)\}$ 和目标任务 $T_t = \{f_t\}$ 。其中， $\mathbf{x}_i^s \in X_s$ 与 $\mathbf{x}_i^t \in X_t$ 分别表示源域样本与目标域样本，通常是样本的三维坐标点集合。源域样本 \mathbf{x}_i^s 的标签为

$y_i^s \in Y_s$ 。点云域自适应学习是在 $T_s = T_t$ 的情况下，利用源域 D_s 的信息提高类别预测函数 f_t 在目标域 D_t 中的性能。

3 点云域自适应学习方法

3.1 基于对抗学习的方法

对抗学习是目前点云域自适应学习中广泛应用的方法，主要采用的是生成对抗网络GAN(Generative Adversarial Networks)架构(Goodfellow等,2020)。如图2所示，基于生成对抗网络的域自适应学习方法通常包含生成器 G 、分类器 F 和判别器 D 。其中，判别器用于判别不同域的数据；生成器用于生成领域无关的特征以欺骗判别器；分类器用于目标分类任务。通过对抗训练的方式，生成器可实现不同域数据的一致特征表达，分类器可实现不同域数据的正确分类。最终，实现域自适应学习任务。

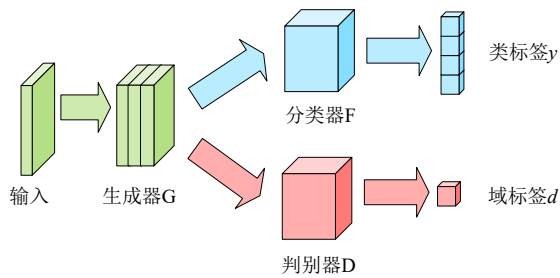


图2 典型的基于生成对抗网络的域自适应学习框架
Fig. 2 Typical GAN based framework of domain adaptation learning method

目前, 已有研究在点云域自适应学习中提出了不同的方法。Qin 等 (2019) 基于对抗学习方法提出一种同时对齐不同域的局部特征与全局特征的域自适应学习网络 PointDAN。为了实现全局特征对齐, PointDAN 设计了最大分类器差异模块 (Saito 等, 2018)。该模块由一个生成器 G 以及两个分类器 F_1 、 F_2 构成。生成器 G 的目标是让两个分类器结果一致, 而分类器 F_1 、 F_2 的目标是找出分类结果不一致的特征。生成器与分类器互相对抗训练实现了全局特征对齐。对于局部特征对齐, PointDAN 使用了自适应 SA (Self-adaptive) 节点模块, 迭代使用偏移量更新节点, 最终得到用于对齐局部特征的局部节点。其中, 偏移量由节点与其邻近点的差异聚合得到。该网络对齐了不同域的局部与全局特征, 表现出了良好的性能。

Jiang 和 Saripalli (2021) 提出了一种边界感知点云域自适应网络 LiDARNet。该网络包含两个分支: 域共享分支和域私有分支。域私有分支负责提取可用于区分不同域的特征, 域共享分支能够提取不同域的共同特征。其中, 域共享分支使用了一种双流卷积神经网络结构 Gated-SCNN (Takikawa 等, 2019) 用于学习边界相关信息, 进而学习语义分割信息。针对目标数据未标记问题, 域共享分支还使用了生成对抗网络架构进一步学习语义信息。

针对三维点云的目标检测问题, Zhang 等 (2021) 提出了基于尺度感知和范围感知的域自适应学习网络 SRDAN。该网络是在三维物体检测模型 SECOND (Yan 等, 2018) 的基础上进行改进的, 其通过引入对抗学习框架, 进而微调特征分布, 以实现两个域之间局部特征和全局特征的联合对齐。针对点云目标大小不一致的问题, SRDAN 设计了基于三维体素的特征金字塔网络结

构, 实现了点云目标在不同尺度下的特征表达。在每个尺度的特征提取过程中, SRDAN 中设计了尺度感知和距离感知的域对齐模块, 对齐了不同域中具有相似三维几何结构 (如尺寸和距离) 的特征。

部分研究在点云域自适应网络结构中设计了包含两个对抗学习模块的方法。Saleh 等 (2019) 提出了一种域自适应学习网络, 用于鸟瞰点云所生成图像的车辆检测。该网络使用了两个对抗模块: 生成器 G_{StoR} 和判别器 D_{R} , 以及生成器 G_{RtoS} 和判别器 D_{S} 。一方面, 生成器 G_{StoR} 以源域点云为输入生成目标域点云, 而判别器 D_{R} 区分目标域的点云与 G_{StoR} 生成的点云; 另一方面, 生成器 G_{RtoS} 与生成器 G_{StoR} 的过程相反, 其输入为生成器 G_{StoR} 所生成的点云, 输出为生成的源域点云。判别器 D_{S} 区分源域的合成点云和从 G_{RtoS} 生成的点云。通过不同域点云数据的模拟生成过程, 可实现点云域自适应学习。

Tang 等 (2021) 提出了一种双对抗差异最小化网络 BADM。该网络设计了对抗学习模块评估不同域的间隔差异, 进而计算差异度量。另外, BADM 利用了熵最小化的方式来防止梯度爆炸或消失, 进一步优化决策边界。

由于类与类之间存在可迁移性差异, 如道路、植被等较容易迁移, 自行车、交通标志等较难迁移。Yuan 等 (2023) 提出了一个类级别的对抗框架, 用于点云语义分割任务的无监督域自适应。该框架利用基于点的表示和基于体素的表示之间的预测一致性来衡量每个点在语义上的对齐程度, 并将两种预测的差异作为权重加入到对抗损失中, 使网络更加关注不易迁移的类别, 提高模型在目标域的性能。此外, 该框架使用了多尺度域条件模块, 进一步提取领域特定的知识, 从而减少由不同 LiDAR 采样模式引起的域差异。该模块一定程度上解决了不同域间点密度不一致导致的域偏移问题。

点云域自适应学习中, 引入对抗学习可以在某种程度上通过生成器与判别器对抗训练的方式缩小源域与目标域特征空间的距离, 实现源域和目标域间特征的隐式对齐。然而, 此种方式忽略了特征对齐后不同类别间特征差异性的保持, 进而可能影响下游分类与识别任务的性能。

3.2 基于跨模态学习的方法

跨模态学习方法是利用多个不同模态的信息来提高模型性能的方法。配备光学相机的激光雷达扫描系统可同时获取同一场景的三维点云与二维图像数据。跨模态学习已经应用于二维图像与三维点云两个模态的特征融合学习中。例如，在

语义分割任务中，将单视图或多视图图像的特征融合到三维点云中，实现了二维图像与三维点云的联合处理（Jaritz等，2019，2020；Chiang等，2019）。在点云域自适应学习中，跨模态学习方法主要用于语义分割任务，其典型网络结构如图3所示。

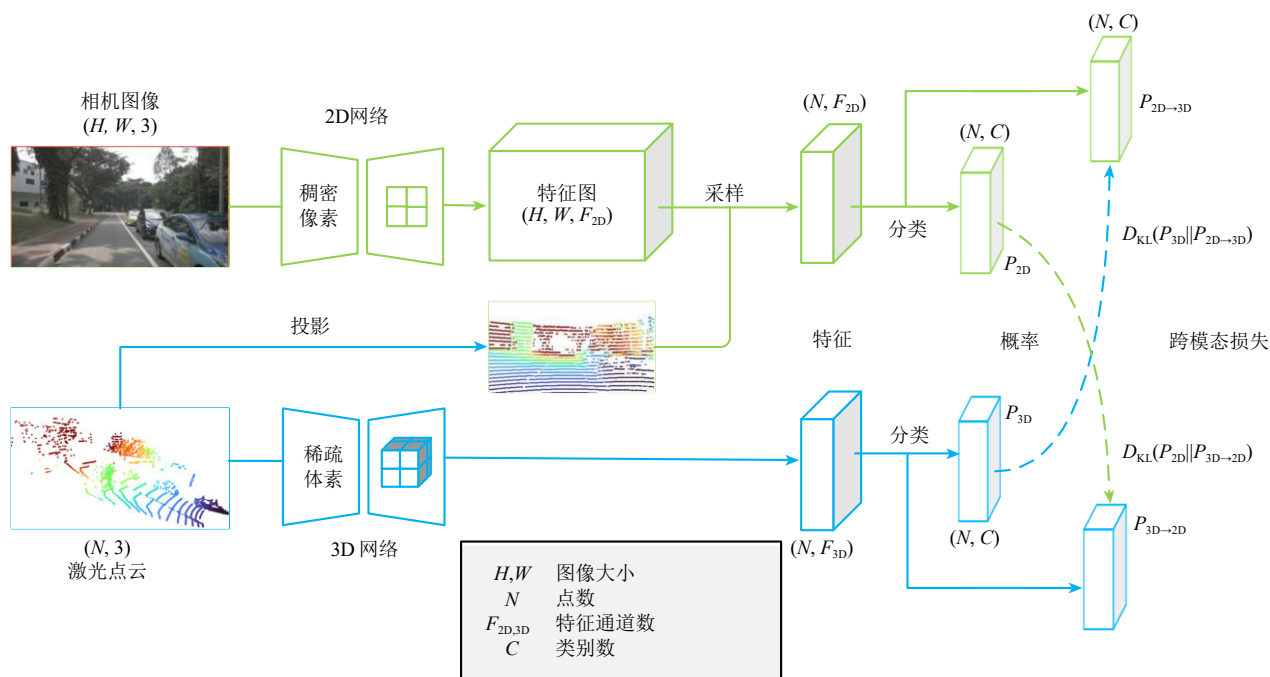


图3 跨模态语义分割网络框架

Fig. 3 Framework of Cross-modal semantic segmentation network

Jaritz等（2020）提出了跨模态无监督域自适应学习框架xMUDA，利用了二维图像与三维点云的类别一致性约束来提升点云域自适应学习的性能。该框架通过KL散度来度量不同模态间类别分布一致性，进而实现不同模态间的相互适配过程。此外，该框架通过分离不同模态的私有和共享信息，可有效避免分类正确率较高的模态采用分类正确率较低的模态的错误预测信息，提高了域自适应方法的性能。

xMUDA框架主要通过利用KL损失来实现二维像素和三维点间跨模态信息的适配。然而，三维点通常比较稀疏，难以用于描述整个二维图像在源域的全局信息。针对此问题，Liu等（2021）在源域中使用完整的二维图像，并提出了一种结合对抗学习和跨模态学习的三维点云语义分割域自适应框架AUDA。另外，为解决数据分布不平衡问题，该模型为不同的类别设置不同的权重，并在

目标域使用权重校准类别的概率，从而提高罕见类别的召回率。

上述方法中使用的跨模态学习方法有效地提升了域自适应学习的性能，但没有充分考虑各模态之间的信息交互。为了提高域自适应学习中跨模态信息交互的能力，Peng等（2021）利用二维图像数据，通过域内和域间的跨模态学习联合实现三维点云的域自适应学习。对于域内跨模态学习，由于二维图像语义分割中相邻像素大多被分类为同一类别，该研究提出了一种跨模态学习方法，在二维图像中动态地为每个三维点捕获像素特征，从而补充三维点云的空间信息。对于域间跨模态学习，该研究提出了基于二维图像和三维点云数据的交叉模态对抗学习方法，针对不同内容、不同模态和不同领域的特征进行学习，通过对抗的方式缓解源域和目标域以及二维模态和三维模态之间的特征分布差异。该研究分别在域内

和域间使用跨模态学习方法, 进一步提高了不同模态的信息交互, 从而提高域自适应学习的性能。

目前, 基于跨模态学习的点云域自适应学习方法的一般思路描述如下: 首先, 对二维图像进行采样, 得到与较稀疏的三维点云数据相匹配的二维图像数据; 随后, 将得到的二维图像数据用于训练二维模态的域自适应学习网络, 并辅助三维点云领域的自适应学习网络训练。此研究思路虽然获得了一定的成功, 但没有直接考虑不同模态之间信息的互补性。近年来, 面向二维图像的域自适应方法已经得到了快速发展。跨模态学习可考虑直接融合图像与点云的特征信息, 得到更加全面、丰富的特征表示, 同时辅以面向二维图像的域自适应方法, 进而实现三维点云的域自适应学习。但是, 点云与图像的数据和特征分布存在差异, 两者融合后的特征更为复杂, 因此需要重新设计面向点云与图像的融合特征的对齐方法。另外, 某个模态中的错误标注信息, 易对另一个模态的标注结果产生负面影响, 进而降低模型的适应性。因此, 如何有效对齐融合后的特征以及

降低模态间的负面影响是跨模态域自适应学习中待解决的问题。

3.3 基于伪标签学习的方法

伪标签学习是一种半监督学习方法。该方法的主要思想是: 首先, 在标记数据集上训练模型; 然后, 利用训练后的模型预测未标记数据的标签, 即生成伪标签; 最后, 将标注数据和新生成的伪标注数据合并, 作为下次训练的监督信息。在点云处理任务中, 因为源域数据与目标域数据的差异, 直接将源域数据作为监督信息训练网络并不能在目标域上取得良好的效果。伪标签学习将源域数据作为初始监督信息, 在目标域迭代更新伪标签, 因而能更好地适应目标域, 其网络结构如图4所示。在计算机视觉任务中, 伪标签学习在域自适应学习中 (Yan 等, 2017; Zou 等, 2018; Tang 等, 2020) 已经取得不错的效果。因此, 部分学者探索了伪标签学习在点云域自适应学习中的可行性。下文将具体介绍基于伪标签学习的三维点云域自适应学习方法。

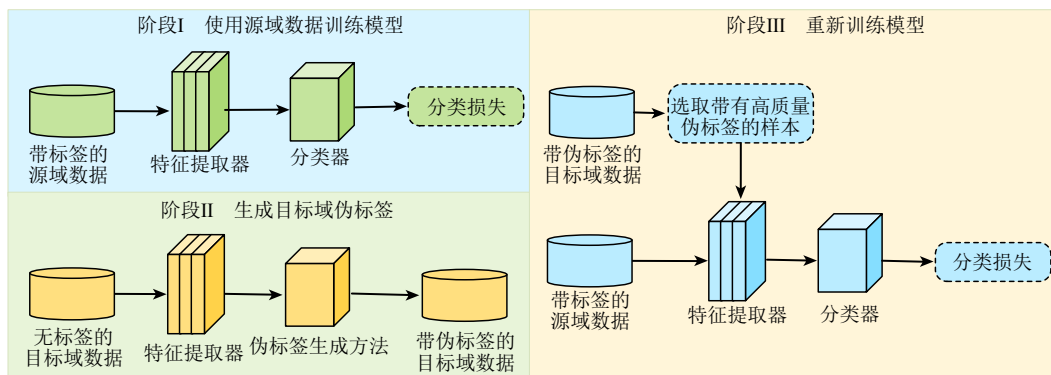


图4 基于伪标签学习的域自适应学习方法框架

Fig.4 The pseudo-label learning based framework of domain adaptation learning methods

由于源域和目标域的数据存在一定的分布偏差 (域偏移), 基于激光雷达的三维检测网络在目标域的泛化能力较差 (Wang 等, 2020)。针对这个问题, Saltori 等 (2020) 提出了基于激光雷达的三维目标检测无监督域自适应学习网络。首先, 该网络使用在源域中预训练的 PointRCNN (Shi 等, 2019) 模型在多个尺度上对未标注的目标数据集进行标注。然后, 通过连续帧的平均体积变化评估其检测的稳定性, 从而评估标注的质量。最后, 对尺度变换进行逆变换, 通过置信度聚合最佳检测标签, 并微调 PointRCNN。该方法在源域数据缺

失的情况下, 使用无标签的目标域数据对源域得到的模型进行微调, 从而达到域自适应的目的。然而, 在源域数据与目标域数据差异较大的情况下, 该方法的伪标签质量较低, 从而导致域自适应的效果不佳。

部分点云域自适应学习工作结合了伪标签与自我训练 (Self-Training) 方法, 并取得了不错的效果。Cardace 等 (2021) 提出了一种结合伪标签和自我训练的无监督域自适应学习方法。该方法通过运行在源域上训练的模型来获得目标域上的预测, 然后将其作为监督信息指导目标域模

型学习目标域的知识。这个过程通常被称为自我训练 (Zou 等, 2018)。然而, 由于域偏移, 源域训练的模型对目标域数据的预测 (即伪标签) 可能是错误的。将错误的伪标签作为监督信息会导致目标域模型的表现不佳。因此, 该方法对伪标签学习进行了两方面的改进: 一方面, 根据置信度将目标域样本分为高置信度样本和低置信度样本, 并通过 K 近邻的方式重新分配低置信度样本的伪标签; 另一方面, 计算每个类在目标域上类似于聚类中心的平均全局描述符 (Pinheiro, 2018), 即原型, 并根据输入形状与其原型的相似性对目标样本的伪标签进行加权。该方法将自我训练与伪标签学习结合并应用于点云域自适应中。同时, 针对错误伪标签影响模型性能的问题, 对伪标签的生成方法进行了改进, 降低了生成错误伪标签的概率。

Yang 等 (2021a) 提出了一种用于三维目标检测的域自适应自我训练过程 ST3D。首先, 该方法使用随机对象缩放技术随机缩放 3D 对象, 并将其用于源域模型的预训练中, 以提高源域模型的泛化能力。然后, 提出构建一个质量感知的三元组记忆库。该记忆库根据交并比 IoU (Intersection over Union) 将所有样本分为了正样本、忽略样本以及负样本 3 类。其中, 正样本作为下次训练的监督信息; 忽略样本不携带类别标签, 可用于识别由于不确定性较高而在模型训练过程中被忽略的区域; 负样本不参与模型训练。最后, 该方法提出了课程数据增强 CDA (Curriculum Data Augmentation) 策略。该策略将模型的训练集从包含具有高置信度伪标签的样本逐步扩充至所有目标域样本。该方法不仅提升了模型的泛化能力, 还从筛选伪标签以及生成伪标签两个方面减低了错误伪标签对模型的影响。

为了进一步减轻了错误伪标签对模型训练的负面影响, Yang 等 (2023) 提出了 ST3D++ 模型。与 ST3D 相比, 该模型有两个方面的改进: 一方面, 该方法提出了混合质量感知三元组存储模块, 通过联合 IoU 和分类置信度两个标准将样本分为正样本、忽略样本以及负样本 3 类, 并以投票的方式利用当前伪标签与历史伪标签更新伪标签, 并将其作为下一次训练的监督信息; 另一方面, 该方法还利用无错误标记的源域数据纠正由错误伪标签引起的不精确梯度, 并通过对源域和目标域数据的联合优

化学习域不变特征, 以保持对不同域中不同类别的区分度。ST3D++ 进一步在筛选伪标签方面做出了改进, 在标准中加入了分类置信度, 并使用源域数据进行纠正, 减少了错误伪标签对模型的影响。

将自我训练方法直接应用于目标域数据中可在一定程度上减少不同域数据的分布差异, 进而减少在环境发生变化时 (例如地理位置、传感器设置、天气条件等) 对高质量标签的需求。为了进一步考虑自动驾驶数据的时间序列特征, Fruhwirth-Reisinger 等 (2021) 基于场景中三维点云的运动信息, 即场景流 (Liu 等, 2019; Wu 等, 2020), 提出了一种具有流感知的自我训练方法, 用于跨域的三维目标检测。该方法包括 4 个步骤: 首先, 利用源域中训练的模型获取目标域数据的初始三维目标检测结果; 其次, 利用场景流的传播得到高置信度的检测结果, 并获得初始检测的运动轨迹。然后, 计算每条轨迹的命中率, 删除命中率 < 0.3 、轨迹长度 < 5 帧或轨迹长度 > 15 帧的轨迹, 并使用场景流恢复部分丢失的轨迹。最后, 提取伪标签进行自我训练, 以改进初始模型。该方法考虑了数据的时间性质, 利用场景流的传播得到了部分高置信度的伪标签。但是, 生成错误的伪标签不可避免, 需要一个良好的方法减少错误伪标签对模型的影响。

由于采集过程不同, 一个物体的点云表示可能会有很大的几何变化, 从而导致域偏移。为了解决这个问题, Zou 等 (2021) 提出了一种具有几何感知的自我训练方法 GAST。GAST 利用自步学习 (Self-Paced Learning) (Zou 等, 2018) 方法, 在每一步迭代中选择下一步学习的样本, 从带有高置信度伪标签的样本逐步扩充到全部样本, 减少错误伪标签的影响。此外, 该方法还引入了旋转角度预测自监督学习与失真位置预测自监督学习两个辅助任务。其中, 旋转角度预测自监督学习任务将源域和目标域的点云数据进行旋转混合, 使模型学习到检测对象的局部信息。失真位置预测自监督学习任务利用曲率预测失真点的位置, 帮助主干网络检测出点云数据中的噪声, 进而提高模型的鲁棒性。该方法使用自步学习的方法降低了错误伪标签对模型的影响, 并通过辅助任务增强模型的泛化能力, 但是仍然需要进一步提高伪标签质量。

Liang等(2022)利用自步学习的方式选择可靠的目标域样本:将最大预测概率对应的类作为每个目标样本的伪标签,并使用熵足够小的目标样本的伪标签监督模型训练。同时,该研究还设计了一个辅助任务来增强模型的泛化能力。该任务随机遮掩源域点云数据的部分区域,要求模型分别通过估计邻域基数(预定义半径内的邻域点的数量)、缺失点的位置和法线来预测区域结构。

Fan等(2022)提出了一种投票方法,为目标域样本分配可靠的伪标签,并进行自我训练:首先,通过共享特征空间中最邻近的源域样本投票得到目标域样本的伪标签。然后,根据源域邻近样本的标签一致性选取部分目标样本作为可靠的训练数据。在网络的训练过程中,自适应地拓展可靠的训练数据,从而实现域自适应任务。此外,该方法还设计了一个辅助任务,辅助主干网络捕获点云的全局结构。该任务通过放大或缩小一个维度的点坐标的方式来学习点云表示,然后根据其他两个不变的维度来预测尺度。

Yoo和Jun(2023)发现GAST(Zou等,2021)的伪标签生成方法存在一个不足之处,即在确定学习方向时可能发生错误。原因是伪标签是由具有最高置信度分数的单个预测结果确定的,忽略了其他预测结果。为了弥补这一缺陷,提出了一种基于回收最大池化RMP(Recycling Max Pooling)模块(Chen等,2022)的自步学习方法。该方法从单个输入中提取多个候选标签生成伪标签。其中,具有高置信度的伪标签将用于监督模型训练。此外,还设计了切割平面识别的方法,将点云在 $x-z$ 或 $y-z$ 方向上分成两部分,然后预测两个平面之间的切割方向,从而学习点云的几何特征。

伪标签学习在点云域自适应学习中通常与自我训练相结合,迭代更新伪标签与特征提取器,缩小源域与目标域之间的距离。在伪标签学习中,最重要的是伪标签的生成以及判断。若将错误的伪标签作为下一轮迭代的监督数据,将影响到神经网络的决策准确率。

3.4 基于数据对齐的方法

数据对齐是一种传统的域自适应学习方法。其思想是通过更改源域数据的方式,使源域和目标域的数据分布尽可能接近,网络框架如图5所示。数据对齐可以分为两种类型,一种是合成数

据与真实数据的对齐,另一种是两个不同场景真实数据的对齐。前者源域与目标域的数据比较相似,可以使用生成点云强度等方法进行特征对齐;后者源域与目标域差异较大,需要使用一些方法增强二者的相似性来对齐特征。下面将介绍近年来各种基于数据对齐的三维点云域自适应学习方法。

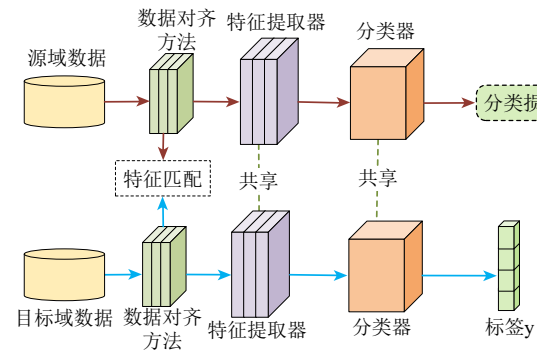


图5 基于数据对齐的域自适应学习方法框架
Fig.5 The data alignment-based framework of domain adaptation learning methods

由于合成数据与真实数据的分布差异,基于合成数据训练的模型通常难以推广到真实世界。为了解决这个问题,Wu等(2019)在SqueezeSeg(Wu等,2018)的基础上,提出了模拟—真实域自适应SRDA(Simulation-to-real Domain Adaptation)框架SqueezeSegV2,用于三维点云的语义分割。其中,SqueezeSeg是一种以球形投影的方式将稀疏、不规则分布的三维点云转换为密集的二维网格表示的语义分割网络。SqueezeSegV2以点云的空间坐标为输入,预测三维点的强度值,并将其加入到点云数据中。在深度神经网络中,输入数据的分布差异可能会引起每一层输出的分布偏移。此类分布偏移将在网络训练过程中逐渐累积,最终导致性能严重下降。为了解决这个问题,SqueezeSegV2采用了渐进域校准PDC(Progressive Domain Calibration)方法,逐层校准,以防止分布偏移在网络中的传播。此外,针对激光雷达点云数据的缺失点(即丢失噪声)严重影响网络性能的问题,提出了上下文聚合模块CAM(Context Aggregation Module)来降低算法对丢失噪声的敏感性。SqueezeSegV2通过学习得到强度值,使合成数据更加接近真实数据,并提出了上下文聚合模块,降低了真实数据中点云缺失对模型的影响。

尽管SqueezeSegV2在模拟—真实域自适应语

义分割任务中达到了当时最先进的性能,但仍存在一些局限性。首先, SqueezeSegV2采用了多阶段处理流程,无法进行端到端训练。其次,它未考虑不同域之间的像素级丢失噪声差异。最后,由于真实世界统计数据难以估计,并且随着数据集的增加和扩展,渐进域校准效率低下且鲁棒性差。基于此, Zhao等(2021)同样使用球形投影将三维点云转换为二维网格表示,并设计ePointDA框架来解决上述问题。该框架提出了一个生成丢失噪声的模块,通过学习使源域的合成数据缺失部分点云,以模拟目标域的真实数据分布。另外,针对渐进域校准效率低下且鲁棒性差的问题,该框架提出了统计不变和空间自适应的特征对齐方法:(1)通过实例规范化提取统计不变特征(Statistics-invariant Features)(Ulyanov等, 2017);(2)通过高阶矩匹配(Higher-order Moment Matching)在高维空间对齐特征映射(Chen等, 2020);(3)通过改进的空间自适应卷积模块SAC(Spatially-adaptive Convolution)生成域不变的空间注意力图(Xu等, 2020)。该方法不仅有助于弥合空间特征的差距,而且不需要事先获取足够的真实数据来获取统计信息,使其能够更好地处理数据集的增加和扩展,从而使其更具鲁棒性和实用性。该方法对SqueezeSegV2存在的一些问题做出了改进,进一步使合成数据的分布接近真实数据,并使用统计不变和空间自适应的特征对齐方法增强了模型的鲁棒性。

以上的两种方法都使用生成点云强度、降低目标域点云数据缺失影响等方法对齐合成数据与真实数据。然而,在对齐两个真实场景时,数据之间可能存在较大的差异,例如郊区场景和城市场景中的房屋高度差异。针对这种情况, Luo等(2020)受PointNet(Charles等, 2017)中变换网络(Tnet)的启发,提出了一种逐点关注转换模块(PW-ATM)。该模块通过学习目标域和源域的缩放系数,自适应地实现了目标域和源域的数据对齐。同时,该方法还引入了基于最大分类器差异(MCD)的对抗学习框架,进一步对齐特征并确保分类准确性,实现了城市车载激光扫描点云语义分割的无监督域自适应学习。该方法所提出的逐点关注转换模块主要解决不同数据集中同类目标高度差异的适配问题,忽略了不同数据集中同类目标具有不同几何形状、点密度差异、点分布差

异等情况(Xu等, 2021)。

此外, Kang等(2022)提出了一个用于点云形状分类的网络,通过搜索点云上的最佳操作来学习点云从源域到目标域的变换,缩小了域间同类目标的几何形状差异,进而实现了点云域自适应。该方法自适应地组合使用压缩、按比例缩放、随机裁剪、随机上采样、随机下采样、点云插值等操作,混淆域鉴别器从而实现域自适应。同时,为了避免转化过程中丢失信息,将变换的输入和变换的输出之间的推土距离EMD(Earth Mover's Distance)(Rubner等, 2000)的平均值作为相似性度量以约束网络训练。

在域自适应方法中,除了对齐两个域的数据外,还可以对源域数据进行处理,以增强模型的泛化能力。Achituve等(2021)通过基于体积的变形、基于特征的变形和基于采样的变形等方法,构造形状缺失的区域,提高源域模型的泛化能力。同时,受到MixUp方法(Zhang等, 2018)的启发,该研究提出了一种应用于源域点云数据的点云混合方法PCM(Point Cloud Mixup),对源域数据变形,增强源域模型的泛化能力。然而,源域和目标域的数据差异较大时,仅仅通过点云变形和点云混合增强模型的泛化能力不足以有效提升域自适应的效果。

基于对三维表面上采样得到的稀疏点云的观察, Yi等(2021)提出了一种不需要手工标注监督数据的用于点云域自适应学习的稀疏体素补全网络。该网络利用对抗学习机制捕获局部曲面的先验信息,实现了不完整三维表面的补全。补全后三维表面有效地减少了源域和目标域数据的几何差异,进而提升域自适应方法的性能。该网络主要解决了不同域点云目标获取不完整情况的适配问题,没有直接考虑源域与目标域同类目标差异较大的情况(Luo等, 2021b)。

由于遮挡、低反射或天气干扰等现象会造成点云数据的缺失,三维目标检测器的性能在跨域时可能会急剧下降。针对这个问题, Xu等(2021)提出了一种语义点生成方法,用于三维点云的目标检测。该方法在预测的前景区域生成语义点,并恢复由遮挡、低反射率或天气干扰等现象造成的前景目标缺失部分。语义点与原始点合并得到增强的点云数据,一定程度上解决了点云目标不完整导致域偏移的问题,提高了网络在目标域上

的性能。

Huang等(2022)使用生成对抗网络(GAN)从源域生成点云数据,从而使输出更接近目标域。该方法充分利用了对抗学习的优势,在生成合成数据的同时保持源域的标签。该方法提出的模型由4个部分组成:生成器、潜在重构模块、判别器和特征编码器/解码器。其中,生成器以点云的编码特征和潜在条件的高斯编码特征(即潜码)作为输入,可以生成多个可能的输出,以提高生成器输出对象的质量。而潜在重构模块在训练过程中恢复潜码,计算潜在重构损失。此外,还有一个额外的鉴别器来限制生成器的输出。该方法使用对抗学习的方法隐式地将源域数据转换为目标域数据,得到带标签的目标域数据,将目标域的无监督学习变为了半监督学习,提升了网络的性能。

Luo等(2021a)提出了一种基于深度神经网络的可学习变换网络,用于可使一个目标缺失部分点云数据并进行恢复的点云破坏-重建任务中。该网络使用对抗学习的方式学习一个连续的非线性变换,生成丰富和复杂的点云,同时保留语义信息,以达到点云破坏-重建的目的。通过点云的破坏-重建,编码器可以提取迁移性较强的局部特征,提升域自适应网络的性能(Alliegro等,2021)。该网络被集成到一种基于自监督的点云无监督域自适应学习体系中(Sun等,2019;Xu等,2019;Achituv等,2021),并取得了不错的效果。该方法的点云破坏-重建任务是一种特殊的数据增强方式,提升了源域模型的泛化能力。

在点云域自适应学习中所采用的数据对齐方法往往只针对原始点云数据进行修改(例如补全、重建等),进而提高源域与目标域中点云数据几何结构的相似性。目前,此类方法的研究在数据对齐过程中,未考虑源域与目标域在特征空间上的对齐,缺乏在数据本身和特征提取这2个不同层面上同时对齐的相关研究。

3.5 其他方法

除了上述4种方法以外,还有一些其他类型的点云域自适应方法(Luo等,2021b;Li等,2023;Bian等,2022;Yang等,2021b)。

Luo等(2021b)发现几何不匹配是影响三维目标检测域转移性能的关键因素,并提出了多层次一致性网络。该网络采用均值教师范式

(Tarvainen和Valpola,2018)生成自适应的、可靠的伪标签。教师模型本质上是学生模型的时间集合,教师模型的参数通过对之前迭代的学生模型的参数取均值进行更新。此外,该网络还实现了点、实例和神经统计3个级别的一致性对齐,使学生模型在教师模型的指导下学习域不变表示,从而适应无标签的目标域。在该网络中,学生模型以教师模型生成的伪标签为监督信息进行训练。然而,由于教师模型生成的标签可能存在错误,这可能导致学生模型的性能降低。

Li等(2023)同样使用了均值教师范式,并提出了一种用于三维点云域自适应任务的端到端自集成网络。在源域中,该网络提出了软分类损失,利用教师模型提供可靠的软标签,从而实现学生模型的训练。在目标域中,该网络提出了一致性损失,保证教师模型和学生模型的预测结果一致,从而提高学生网络的鲁棒性。该网络的教师模型还在源域生成软标签,并同真实标签比较,从而校准教师模型的准确度,一定程度地提升了教师模型生成伪标签的质量。然而,源域数据与目标域数据存在较大差异,通过源域数据校准教师模型,并不能有效地保证其生成伪标签的质量。

另一方面,Bian等(2022)提出了一个基于图的框架来实现域之间的局部特征对齐,在域适应过程中保留用于区分类别的语义信息。具体来说,首先通过最远点采样的方式选取质心点,并构建每个质心点的动态局部特征图,获取其局部几何信息。然后,为了丰富源域的特征图,构造了一个特征图记忆库来存储在训练阶段生成的源域特征图。之后,受到点云匹配(Yew和Lee,2020)的启发,采用最优传输代价来度量内存库与目标域之间的图相似度,从而获得可靠的分配矩阵来指导知识从源域转移到目标域。此外,为了进一步提取有区分性的目标域特征,还利用了对比损失来区分不同类别的特征图分布,从而指导网络学习目标域上的判别特征。

由于数据隐私和安全问题,很多公司只提供学到的模型,而不提供客户数据。针对这种情况,Yang等(2021b)发现目标域中无法被分类器正确分类的同类样本在特征空间中仍然存在分布集中现象。因此,他们提出了邻域互惠聚类方法,通过鼓励互惠邻居的标签预测一致来实现无源域自

适应学习。该方法通过定义亲和性值来描述每个数据点与其邻居之间的关联程度，使用最近邻辅助监督来聚合更广泛的局部信息，并提出一种自正则化的方法来减少潜在噪声邻居的负面影响。对于无源域自适应，邻域互惠聚类可有效目标域标签的准确度，从而提升模型的性能。

4 结果与分析

4.1 三维点云形状分类

4.1.1 评价指标

在三维点云目标分类任务中，通常采用准确率（ACC）作为模型性能优劣的评价指标，计算公式如下：

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

式中，TP（True Positive）表示样本预测值与真实值相符且均为正的样本数；FP（False Positive）表示样本预测值为正而真实值为负的样本数；FN（False Negative）表示样本预测值为负而真实值为正的样本数；TN（True Negative）表示样本预测值与真实值相符且均为负的样本数。

4.1.2 数据集

三维点云形状分类的域自适应方法主要在 PointDA-10（Qin 等，2019）数据集上进行测试。该数据集包含 3 个子数据集，分别是 ModelNet-10（M）、ShapeNet-10（S）和 ScanNet-10（S*）。这些子数据集存在 10 个公共类别，可用于域自适应

任务。ModelNet-10 由 3D CAD 模型生成，包含 4183 个训练样本和 856 个测试样本。ShapeNet-10 也是生成的数据集，包含了 17378 个训练样本和 2492 个测试样本。ScanNet-10 是唯一的真实数据集，包含了 6110 个训练样本和 1769 个测试样本。

4.1.3 实验结果分析

表 2 展示了几种典型域自适应方法在 PointDA-10 数据集上的测试结果。Qin 等（2019）发布了 PointDA-10 数据集，并首次在该数据集上进行测试。与 PointDAN 相比，Tang 等（2021）提出的双对抗网络取得了一定的性能提升。可见，相较于仅使用单个对抗模块，适当增加对抗模块可以提高模型性能。在源域数据缺失的情况下，Yang 等（2021b）提出的无源域自适应网络在 PointDA-10 上展现出了良好的性能，证明了邻域互惠聚类方法的可行性。此外，Achituve 等（2021）通过点云变形和点云混合两种方式增强了模型的泛化能力，进一步提高了准确率。随着点云域自适应的发展，出现了一系列结合点云几何信息的域自适应方法，如 GAST（Zou 等，2021）、DFAN（Shi 等，2022）、GLRV（Fan 等，2022）、（Shen 等，2022）和 MLSP（Liang 等，2022）。这些方法在 PointDA-10 上都取得了显著的准确率提升。其中，MLSP 方法通过设计估计邻域基数（预定义半径内的邻域点的数量）、缺失点的位置和法线等辅助任务有效地减少了域间数据分布的差异，在 PointDA-10 上达到了最佳性能。

表 2 不同方法在 PointDA-10 数据集上的准确率 ACC

Table 2 Accuracy of different methods on the PointDA-10 dataset

方法	数据集						平均精度
	M→S	M→S*	S→M	S→S*	S*→M	S*→S	
w/o Adapt	42.5	22.3	39.9	23.5	34.2	46.9	34.9
PointDAN(Qin 等, 2019)	64.2	33.0	47.6	33.9	49.1	64.1	48.7
BADM(Tang 等, 2021)	66.3	33.7	56.3	33.3	53.9	65.8	51.5
NRC(Yang 等, 2021b)	64.8	25.8	59.8	26.9	70.1	69.1	52.6
DefRec(Achituve 等, 2021)	82.7	43.9	79.8	48.0	66.0	67.4	64.6
GAST(Zou 等, 2021)	84.8	59.8	80.8	56.7	81.1	74.9	73.0
DFAN(Shi 等, 2022)	83.7	60.2	84.0	60.3	75.3	76.3	73.3
GLRV(Fan 等, 2022)	85.4	60.4	78.8	57.7	77.8	76.2	72.7
GAI(Shen 等, 2022)	86.2	58.6	81.4	56.9	81.5	74.4	73.2
MLSP(Liang 等, 2022)	86.2	59.1	83.5	57.6	81.2	76.4	74.0

4.2 三维点云目标检测

4.2.1 评价指标

三维点云目标检测任务将平均精度 AP (Average Precision) 作为评价模型性能的指标, 具体公式如下:

$$AP = \int_0^1 P(r) dr \quad (2)$$

式中, $P(r)$ 表示了 PR (Precision Recall) 曲线, 其召回率 Recall 和准确率 Precision 的计算公式如下:

$$Recall = \frac{TP}{TP + FP} \quad (3)$$

$$Precision = \frac{TP}{TP + FN} \quad (4)$$

4.2.2 数据集

三维点云目标检测的域自适应方法主要在 nuScenes 和 KITTI 2 个数据集进行测试。其中, nuScenes 是第一个提供自动驾驶全套传感器数据的大型数据集, 包含了 28130 个训练样本和 6019 个测试样本; KITTI 是一个自动驾驶相关的三维点云目标检测数据集, 由 3712 个训练样本和 3769 个验证样本组成。

4.2.3 实验结果分析

表 3 为在 nuScenes→KITTI 目标检测域自适应任务上不同方法在车辆目标检测上的性能。根据 KITTI 数据集所提供的信息, 我们将车辆检测任务分为简单、中等及困难 3 个级别, 并分别计算在不同级别任务的 AP 值, 进而计算所有任务的 AP 值的平均。由表 3 可知: SF-UDA 的性能略低于 ST3D, 这表明相较于 ST3D 提出的伪标签生成策略, 使用场景流生成伪标签的方法效果稍显不足; MLC-Net 的性能略低于 ST3D, 却高于 SF-UDA。这说明了 MLC-Net 提出的均值教师思路是可行的, 但还有待改进。ST3D++ 在该任务上取得了最佳性能, 验证了所提出的混合质量感知三元组存储 (HQTM) 模块的有效性。

4.3 三维点云语义分割

4.3.1 评价指标

三维点云语义分割任务通常将交并比 IoU (Intersection over Union) 作为评价模型性能的指标。交并比计算了某个类别预测结果与真实标签

之间交集与并集之间的比值, 计算公式如下:

$$IoU = \frac{X \cap Y}{X \cup Y} = \frac{TP}{TP + FP + FN} \quad (5)$$

式中, X 为预测结果, Y 为真实标签, 平均交并比 (mIoU) 为所有类别 IoU 的平均值。

表 3 不同方法在车辆目标检测域自适应任务 nuScenes→KITTI 上的精度 AP

Table 3 AP of different methods on domain adaptation of car detection for nuScenes→KITTI

方法	任务等级			平均精度
	简单	中等	困难	
SF-UDA (Saltori 等, 2020)	68.8	49.8	45.0	54.5
MLC-Net (Luo 等, 2021b)	71.3	55.4	49.0	58.6
ST3D (Yang 等, 2021a)	—	—	—	62.55
ST3D++ (Yang 等, 2023)	—	—	—	65.56

4.3.2 数据集

三维点云语义分割的域自适应方法在 PointSegDA、nuScenes、A2D2、SemanticKITTI 等多个数据集上进行了测试。

PointSegDA 是 Achituve 等 (2021) 提出的人体模型网格数据集, 包含 4 个子数据集: ADOBE (A)、FAUST (F)、MIT (M) 和 SCAPE (S)。这 4 个子数据集都包含八类人体部位 (包括脚、手、头等), 但在点云分布、姿势以及扫描角度等方面存在一定的差异。

SemanticKITTI 是由三维目标检测数据集 KITTI 拓展得到的室外点云语义分割数据集, 将点云目标分为了 19 个不同的类别, 包括但不限于: 道路、建筑物、车辆、行人、自行车等。A2D2 是一个大型的自动驾驶公开数据集, 包含了大量来自汽车上装载传感器的数据。SemanticKITTI 与 A2D2 之间存在 10 个共享的类别, 可用于域自适应研究。

nuScenes 是一个公开的三维点云目标检测数据集, 其标注方式是采用三维外接矩形框。对于每个点云, 如果其位于某个目标的矩形框内, 则会被分配相应的标签, 从而实现三维语义分割的逐点标注; 否则, 该点会被标记为背景。通过这种标注方式, nuScenes 可用于三维点云的语义分割任务。另外, 基于元数据信息可以将 nuScenes 按照白天和夜晚、美国和新加坡两个维度进行拆分, 从而得到 Day→Night 和 USA→Singapore 两个域自

适应任务。此外，nuScenes与SemanticKITTI之间有10个共享的类别，可用于域自适应研究。

4.3.3 实验结果分析

表4为不同方法在室外点云数据集nuScenes与SemanticKITTI上语义分割域自适应任务的实现结果。其中，DCF-Net(Yuan等,2023)有效结合了体素和点两种类型的语义分割网络，自适应地赋予难以对齐的点更高的权重以提高对齐速度与精度，进而弥补了对抗学习易忽略局部信息的局限，在两个域自适应任务中均达到最佳性能。

表4 不同方法在数据集nuScenes和SemanticKITTI上的平均交并比mIoU

域自适应任务	方法	mIoU /%
nuScenes→ SemanticKITTI	w/o Adapt	23.5
	SVCN(Yi等,2021)	33.7
	GLFA(Bian等,2022)	—
	DCF-Net(Yuan等,2023)	49.2
SemanticKITTI→nuScenes	w/o Adapt	27.9
	SVCN(Yi等,2021)	31.6
	GLFA(Bian等,2022)	37.3
	DCF-Net(Yuan等,2023)	44.0

表5为不同方法在域自适应任务A2D2→SemanticKITTI上语义分割的比较结果。表6为不

同方法在nuScenes数据集上的语义分割结果的比较。两个表中的方法均属于跨模态学习。其中，3D和2D分别表示对3D和2D网络的评估结果，而2D&3D表示对2D网络和3D网络相融合的评估结果。在这些方法中，AUDA(Peng等,2021)和DsCML(Liu等,2021)两种方法都考虑到了xMUDA(Jaritz等,2020)方法的局限性，并进行了改进。其中，AUDA在A2D2→SemanticKITTI和Day→Night任务中采用对抗学习的方式减小了域间差异，并通过动态增加稀有类别的权重的方式解决了类别不平衡问题，在域自适应任务A2D2→SemanticKITTI和Day→Night上达到了最佳性能。DsCML解决了稠密2D特征和稀疏3D特征之间信息交互不足的问题并通过跨模态对抗学习方法缩小了域差异，在域自适应任务USA→Singapore上达到了最佳性能。

表5 不同方法在域自适应任务A2D2→SemanticKITTI上的平均交并比mIoU

方法	2D网络	3D网络	2D&3D	mIoU /%
w/o Adapt	36.0	36.6	41.8	
xMUDA(Jaritz等,2020)	43.7	48.5	49.1	
AUDA(Liu等,2021)	46.8	51.8	52.4	
DsCML(Peng等,2021)	46.8	48.1	50.6	

注：2D&3D表示对2D和3D网络相融合的评估结果。

表6 不同方法在nuScenes数据集上的平均交并比mIoU

Table 6 The mIoU of different methods on the nuScenes dataset

方法	USA→Singapore			Day→Night			mIoU /%
	2D网络	3D网络	2D&3D	2D网络	3D网络	2D&3D	
w/o Adapt	53.4	46.5	61.3	42.2	41.2	47.8	
xMUDA(Jaritz等,2020)	61.1	54.1	63.2	47.1	46.7	50.8	
AUDA(Liu等,2021)	63.9	56.3	65.1	50.1	48.7	53.0	
DsCML(Peng等,2021)	61.9	54.8	65.6	50.3	49.7	52.6	

注：2D&3D表示对2D和3D网络相融合的评估结果。

表7为几种典型域自适应方法在PointSegDA数据集上的测试结果。Achituve等(2021)发布了PointSegDA，并首次在PointSegDA上实现域自适应任务，是所有方法的基线。Li等(2023)提出的模型SEN的性能比Achituve等(2021)提出的DefRec性能更优越。这表明使用辅助模

型是一种可以有效提高三维点云上的域自适应网络性能的策略。Luo等(2021a)提出的模型ALST通过点云破坏-重建任务增强了模型的泛化能力，在域自适应任务中的性能也优于DefRec。然而，由于个别任务的数据差异较大，其性能提升效果不如SEN显著。DFAN(Shi等，

2022) 采用了多层次的局部特征对齐策略来平衡局部特征的多样性和共性, 并提出了新的损失

函数用于对齐全局特征, 在所有方法中取得了最佳性能。

表7 不同方法在 PointSegDA 数据集上的平均交并比 mIoU
Table 7 The mIoU of different methods on the pointSegDA dataset

方法	数据集							/%
	F→M	F→A	F→S	M→F	M→A	M→S	A→F	
w/o Adapt	60.9	78.5	66.5	33.6	26.6	69.9	38.5	
DefRec(Achituve 等, 2021)	61.8	79.7	67.4	40.1	67.1	72.6	42.5	
SEN(Li 等, 2023)	63.1	75.9	62.8	56.4	78.7	75.7	51.8	
ALST(Luo 等, 2021a)	61.8	80.3	68.5	56.6	60.8	67.8	52.3	
DFAN(Shi 等, 2022)	63.0	74.9	68.6	64.8	73.9	69.7	54.2	
MLSP(Liang 等, 2022)	60.0	80.9	65.5	40.4	67.3	70.8	45.4	

方法	数据集					平均精度
	A→M	A→S	S→F	S→M	S→A	
w/o Adapt	31.2	30.0	64.5	74.1	68.4	53.6
DefRec(Achituve 等, 2021)	28.9	32.3	66.2	66.4	72.2	58.1
SEN(Li 等, 2023)	43.1	43.2	67.1	69.5	64.7	64.7
ALST(Luo 等, 2021a)	38.6	41.0	66.6	67.4	68.0	60.8
DFAN(Shi 等, 2022)	55.4	51.3	71.7	68.7	73.2	65.8
MLSP(Liang 等, 2022)	31.1	38.4	72.5	66.6	74.8	59.5

注: A、F、M、S 分别为 ADOBE、FAVST、MIT 和 SCAPE 4 个子数据集。

5 结 语

本文对点云域自适应学习产生的背景、相关问题的定义以及国内外研究现状进行梳理与总结。总的来说, 现有方法主要从对抗学习、跨模态学习、伪标签学习以及数据对齐 4 个方面开展研究, 一定程度上降低了点云处理算法对大规模高质量标注数据集的依赖, 并在多种点云处理任务上取得了不错的性能。其中, 对抗学习可以隐式地学习域不变特征, 但未考虑特征对齐后不同类别间特征差异性的保持; 跨模态学习利用 2D 图像上的域自适应学习方法辅助训练, 但需要合适的方式进行 3D 模态与 2D 模态的信息交互; 伪标签学习可以在目标域中迭代更新特征提取器, 但需要良好的伪标签生成方法; 数据对齐对原始数据进行处理, 使源域与目标域的数据分布相似, 但目前主要提高的是源域与目标域数据几何结构的相似性, 未考虑对源域与目标域在特征空间上进行对齐。此外, 均值教师范式、图匹配和邻域互惠聚类也是可进一步研究的点云域自适应学习方法。

展望未来, 点云域自适应学习研究可以从以

下两个方面进行探索:

(1) 现有的点云域自适应学习方法大多是借鉴二维图像中的域自适应学习方法。如何针对点云的无序、非结构化等特性设计适用于三维点云的域自适应学习模型是一个需要进一步深入研究的方向;

(2) 点云域自适应学习可以进一步探索无源域的域自适应方法。例如, 在现实场景中, 提供源域模型的公司由于一些原因无法提供源域数据, 如何在仅拥有源域模型的情况下实现域自适应学习是一个值得探索的方向。

参考文献(References)

- Achituve I, Maron H and Chechik G. 2021. Self-supervised learning for domain adaptation on point clouds//2021 IEEE Winter Conference on Applications of Computer Vision (WACV). Waikoloa: IEEE: 123-133 [DOI: 10.1109/WACV48630.2021.00017]
- Alliegro A, Boscaini D and Tommasi T. 2021. Joint supervised and self-supervised learning for 3D real world challenges//2020 25th International Conference on Pattern Recognition (ICPR). Milan: IEEE: 6718-6725 [DOI: 10.1109/ICPR48806.2021.9412483]
- Behley J, Garbade M, Milioto A, Quenzel J, Behnke S, Stachniss C

- and Gall J. 2019. SemanticKITTI: a dataset for semantic scene understanding of LiDAR sequences//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE: 9296-9306 [DOI: 10.1109/ICCV.2019.00939]
- Bian Y K, Hui L, Qian J J and Xie J. 2022. Unsupervised domain adaptation for point cloud semantic segmentation via graph matching//2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Kyoto: IEEE: 9899-9904 [DOI: 10.1109/IROS47612.2022.9981603]
- Caesar H, Bankiti V, Lang A H, Vora S, Liong V E, Xu Q, Krishnan A, Pan Y, Baldan G and Beijbom O. 2020. NuScenes: a multimodal dataset for autonomous driving//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE: 11618-11628 [DOI: 10.1109/CVPR42600.2020.01164]
- Cardace A, Spezialetti R, Ramirez P Z, Salti S and Stefano L D. 2021. RefRec: pseudo-labels refinement via shape reconstruction for unsupervised 3D domain adaptation//2021 International Conference on 3D Vision (3DV). London: IEEE: 331-341 [DOI: 10.1109/3DV53792.2021.00043]
- Chang A X, Funkhouser T, Guibas L, Hanrahan P, Huang Q X, Li Z M, Savarese S, Savva M, Song S R, Su H, Xiao J X, Yi L and Yu F. 2015. ShapeNet: an information-rich 3D model repository. arXiv: 1512.03012 [DOI: 10.48550/arXiv.1512.03012]
- Charles R Q, Su H, Kaichun M and Guibas L J. 2017. PointNet: deep learning on point sets for 3D classification and segmentation//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE: 77-85 [DOI: 10.1109/CVPR.2017.16]
- Chen C, Fu Z H, Chen Z H, Jin S., Cheng Z W, Jin X Y and Hua X S. 2020. Homm: Higher-order moment matching for unsupervised domain adaptation// Proceedings of the AAAI conference on artificial intelligence. 34(04):3422-3429 [DOI: 10.1609/aaai.v34i04.5745]
- Chen J J, Kakillioglu B, Ren H and Velipasalar S. 2022. Why discard if you can recycle?: A recycling max pooling module for 3d point cloud analysis//2022 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans. IEEE: 559-567.
- Chiang H Y, Lin Y L, Liu Y C and Hsu W H. 2019. A unified point-based framework for 3D segmentation//2019 International Conference on 3D Vision (3DV). Quebec City: IEEE: 155-163 [DOI: 10.1109/3DV.2019.00026]
- Cortés I, Beltrán J, De La Escalera A and García F. 2022. DALi: domain adaptation in LiDAR point clouds for 3D obstacle detection//2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC). Macau, China: IEEE: 3837-3842 [DOI: 10.1109/ITSC55140.2022.9922408]
- Dai A, Chang A X, Savva M, Halber M, Funkhouser T and Niessner M. 2017. ScanNet: richly-annotated 3D reconstructions of indoor scenes//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE: 2432-2443 [DOI: 10.1109/CVPR.2017.261]
- Dong Z, Liang F X, Yang B S, Xu Y S, Zang Y F, Li J P, Wang Y, Dai W X, Fan H C, Hyppä J and Stilla U. 2020. Registration of large-scale terrestrial laser scanner point clouds: a review and benchmark. ISPRS Journal of Photogrammetry and Remote Sensing, 163: 327-342 [DOI: 10.1016/j.isprs.2020.03.013]
- Fan H H, Chang X J, Zhang W Y, Cheng Y, Sun Y and Kankanhalli M. 2022. Self-supervised global-local structure modeling for point cloud domain adaptation with reliable voted pseudo labels//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans: IEEE: 6367-6376 [DOI: 10.1109/CVPR52688.2022.00627]
- Fruhworth-Reisinger C, Opitz M, Possegger H and Bischof H. 2021. FAST3D: flow-aware self-training for 3D object detectors. arXiv: 2110.09355 [DOI: 10.48550/arXiv.2110.09355]
- Geiger A, Lenz P and Urtasun R. 2012. Are we ready for autonomous driving? The KITTI vision benchmark suite//2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence: IEEE: 3354-3361 [DOI: 10.1109/CVPR.2012.6248074]
- Geyer J, Kassahun Y, Mahmudi M, Ricou X, Durgesh R, Chung A S, Hauswald L, Pham V H, Mühlegg M, Dorn S, Fernandez T, Jänicke M, Mirashi S, Savani C, Sturm M, Vorobiov O, Oelker M, Garreis S and Schubert P. 2020. A2D2: audi autonomous driving dataset. arXiv: 2004.06320 [DOI: 10.48550/arXiv.2004.06320]
- Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y. 2020. Generative adversarial networks. Communications of the ACM, 63(11): 139-144 [DOI: 10.1145/3422622]
- Hackel T, Savinov N, Ladicky L, Wegner J D, Schindler K and Pollefeys M. 2017. Semantic3D.net: a new large-scale point cloud classification benchmark. arXiv: 1704.03847 [DOI: 10.48550/arXiv.1704.03847]
- Hu Q Y, Yang B, Xie L H, Rosa S, Guo Y L, Wang Z H, Trigoni N and Markham A. 2020. RandLA-Net: efficient semantic segmentation of large-scale point clouds//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE: 11105-11114 [DOI: 10.1109/CVPR42600.2020.01112]
- Huang J X, Yuan J S and Qiao C M. 2022. Generation for unsupervised domain adaptation: a gan-based approach for object classification with 3D point cloud data//ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Singapore: IEEE: 3753-3757 [DOI: 10.1109/ICASSP43922.2022.9746185]
- Jaritz M, Gu J Y and Su H. 2019. Multi-view PointNet for 3D scene understanding//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Seoul: IEEE: 3995-4003 [DOI: 10.1109/ICCVW.2019.00494]
- Jaritz M, Vu T H, de Charette R, Wirbel E and Pérez P. 2020. XMU-DA: cross-modal unsupervised domain adaptation for 3D semantic segmentation//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE: 12602-12611 [DOI: 10.1109/CVPR42600.2020.01262]
- Jiang P and Saripalli S. 2021. LiDARNet: a boundary-aware domain adaptation model for point cloud semantic segmentation//2021 IEEE International Conference on Robotics and Automation

- (ICRA). Xi'an: IEEE: 2457-2464 [DOI: 10.1109/ICRA48506.2021.9561255]
- Kang D M, Nam Y, Kyung D and Choi J. 2022. Unsupervised domain adaptation for 3D point clouds by searched transformations. *IEEE Access*, 10: 56901-56913 [DOI: 10.1109/ACCESS.2022.3176719]
- Li Q, Peng X J, Yan C, Gao P and Hao Q. 2023. Self-ensembling for 3D point cloud domain adaption. *arXiv: 2112.05301* [DOI: 10.48550/arXiv.2112.05301]
- Liang H X, Fan H H, Fan Z W, Wang Y, Chen T L, Cheng Y and Wang Z Y. 2022. Point cloud domain adaptation via masked local 3D structure prediction//17th European Conference on Computer Vision. *Tel Aviv: Springer: 156-172* [DOI: 10.1007/978-3-031-20062-5_10]
- Liu W, Luo Z M, Cai Y Z, Yu Y, Ke Y, Junior J M, Gonçalves W N and Li J. 2021. Adversarial unsupervised domain adaptation for 3D semantic segmentation with multi-modal learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 176: 211-221 [DOI: 10.1016/j.isprsjprs.2021.04.012]
- Liu X Y, Qi C R and Guibas L J. 2019. FlowNet3D: learning scene flow in 3D point clouds//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). *Long Beach: IEEE: 529-537* [DOI: 10.1109/CVPR.2019.00062]
- Lu W X, Zhou Y, Wan G W, Hou S H and Song S Y. 2019. L3-Net: towards learning based LiDAR localization for autonomous driving//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). *Long Beach: IEEE: 6382-6391* [DOI: 10.1109/CVPR.2019.00655]
- Luo H, Li L K, Fang L N, Wang H Y, Wang C, Guo W Z and Li J. 2022. Domain adaptation for object classification in point clouds via asymmetrical siamese and conditional adversarial network. *IEEE Geoscience and Remote Sensing Letters*, 19: 7004605 [DOI: 10.1109/LGRS.2022.3208589]
- Luo H F, Khoshelham K, Fang L N and Chen C C. 2020. Unsupervised scene adaptation for semantic segmentation of urban mobile laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 169: 253-267 [DOI: 10.1016/j.isprsjprs.2020.10.002]
- Luo X Y, Liu S L, Fu K X, Wang M N and Song Z J. 2021a. A learnable self-supervised task for unsupervised domain adaptation on point clouds. *arXiv: 2104.05164* [DOI: 10.48550/arXiv.2104.05164]
- Luo Z P, Cai Z, Zhou C Q, Zhang G J, Zhao H Y, Yi S, Lu S J, Li H S, Zhang S H and Liu Z W. 2021b. Unsupervised domain adaptive 3D detection with multi-level consistency//2021 IEEE/CVF International Conference on Computer Vision (ICCV). *Montreal: IEEE: 8846-8855* [DOI: 10.1109/ICCV48922.2021.00874]
- Pan S J and Yang Q. 2010. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10): 1345-1359 [DOI: 10.1109/TKDE.2009.191]
- Peng D, Lei Y J, Li W, Zhang P P and Guo Y L. 2021. Sparse-to-dense feature matching: intra and inter domain cross-modal learning in domain adaptation for 3D semantic segmentation//2021 IEEE/CVF International Conference on Computer Vision (ICCV). *Montreal: IEEE: 7088-7097* [DOI: 10.1109/ICCV48922.2021.00702]
- Pinheiro P O. 2018. Unsupervised domain adaptation with similarity learning//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. *Salt Lake City: IEEE: 8004-8013* [DOI: 10.1109/CVPR.2018.00835]
- Qi C R, Yi L, Su H and Guibas L J. 2017. PointNet++: deep hierarchical feature learning on point sets in a metric space. *arXiv: 1706.02413* [DOI: 10.48550/arXiv.1706.02413]
- Qin C, You H X, Wang L C, Kuo C C J and Fu Y. 2019. PointDAN: a multi-scale 3D domain adaption network for point cloud representation. *arXiv: 1911.02744* [DOI: 10.48550/arXiv.1911.02744]
- Rubner Y, Tomasi C and Guibas L J. 2000. The earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2): 99-121 [DOI: 10.1023/A:1026543900054]
- Saito K, Watanabe K, Ushiku Y and Harada T. 2018. Maximum classifier discrepancy for unsupervised domain adaptation//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. *Salt Lake City: IEEE: 3723-3732* [DOI: 10.1109/CVPR.2018.00392]
- Saleh K, Abobakr A, Attia M, Iskander J, Nahavandi D, Hossny M and Nahvandi S. 2019. Domain adaptation for vehicle detection from bird's eye view LiDAR point cloud data//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). *Seoul: IEEE: 3235-3242* [DOI: 10.1109/ICCVW.2019.00404]
- Saltori C, Lathuilière S, Sebe N, Ricci E and Galasso F. 2020. SF-UDA^{3D}: source-free unsupervised domain adaptation for LiDAR-based 3D object detection//2020 International Conference on 3D Vision (3DV). *Fukuoka: IEEE: 771-780* [DOI: 10.1109/3DV50981.2020.00087]
- Shen Y F, Yang Y C, Yan M, Wang H, Zheng Y Y and Guibas L. 2022. Domain adaptation on point clouds via geometry-aware implicit//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). *New Orleans: IEEE: 7213-7222* [DOI: 10.1109/CVPR52688.2022.00708]
- Shi L W, Yuan Z M, Cheng M, Chen Y P and Wang C. 2022. DFAN: dual-branch feature alignment network for domain adaptation on point clouds. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5703412 [DOI: 10.1109/TGRS.2022.3171038]
- Shi S S, Wang X G and Li H S. 2019. PointRCNN: 3D Object Proposal Generation and Detection From Point Cloud//2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). *Long Beach: IEEE: 770-779* [DOI: 10.1109/CVPR.2019.00087]
- Sun P, Kretschmar H, Dotiwalla X, Chouard A, Patnaik V, Tsui P, Guo J, Zhou Y, Chai Y N, Caine B, Vasudevan V, Han W, Ngiam J, Zhao H, Timofeev A, Ettinger S, Krivokon M, Gao A, Joshi A, Zhang Y, Shlens J, Chen Z F and Anguelov D. 2020. Scalability in perception for autonomous driving: waymo open dataset//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). *Seattle: IEEE: 2443-2451* [DOI: 10.1109/CVPR42600.2020.00252]
- Sun Y, Tzeng E, Darrell T and Efros A A. 2019. Unsupervised domain adaptation through self-supervision. *arXiv: 1909.11825* [DOI: 10.48550/arXiv.1909.11825]

- Takikawa T, Acuna D, Jampani V and Fidler S. 2019. Gated-SCNN: gated shape CNNs for semantic segmentation//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE: 5228-5237 [DOI: 10.1109/ICCV.2019.00533]
- Tang H, Chen K and Jia K. 2020. Unsupervised domain adaptation via structurally regularized deep clustering//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE: 8722-8732 [DOI: 10.1109/CVPR42600.2020.00875]
- Tang H R, Xu C W and Yang J F. 2021. Bi-adversarial discrepancy minimization for unsupervised domain adaptation on 3D point cloud//2021 International Joint Conference on Neural Networks (IJCNN). Shenzhen: IEEE: 1-8 [DOI: 10.1109/IJCNN52387.2021.9533499]
- Tarvainen A and Valpola H. 2018. Mean teachers are better role models: weight-averaged consistency targets improve semi-supervised deep learning results. arXiv: 1703.01780 [DOI: 10.48550/arXiv.1703.01780]
- Thomas H, Qi C R, Deschard J E, Marcotegui B, Goulette F and Guibas L. 2019. KPConv: flexible and deformable convolution for point clouds//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE: 6410-6419 [DOI: 10.1109/ICCV.2019.00651]
- Ulyanov D, Vedaldi A and Lempitsky V. 2017. Instance normalization: The missing ingredient for fast stylization. arXiv:1607.08022.
- Wang Y, Chen X Y, You Y R, Li L E, Hariharan B, Campbell M, Weinberger K Q and Chao W L. 2020. Train in Germany, test in the USA: making 3D object detectors generalize//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE: 11710-11720 [DOI: 10.1109/CVPR42600.2020.01173]
- Wang Y, Yin J B, Li W, Frossard P, Yang R G and Shen J B. 2022. SS-DA3D: semi-supervised domain adaptation for 3D object detection from point cloud. arXiv: 2212.02845 [DOI: 10.48550/arXiv.2212.02845]
- Wu B C, Wan A, Yue X Y and Keutzer K. 2018. Squeezeseg: Convolutional neural nets with recurrent crf for real-time road-object segmentation from 3d lidar point cloud.//2018 IEEE international conference on robotics and automation (ICRA). Brisbane: IEEE: 1887-1893 [10.1109/ICRA.2018.8462926]
- Wu B C, Zhou X Y, Zhao S C, Yue X Y and Keutzer K. 2019. SqueezeSegV2: improved model structure and unsupervised domain adaptation for road-object segmentation from a LiDAR point cloud//2019 International Conference on Robotics and Automation (ICRA). Montreal: IEEE: 4376-4382 [DOI: 10.1109/ICRA.2019.8793495]
- Wu W X, Wang Z Y, Li Z W, Liu W and Li F X. 2020. PointPWC-Net: cost volume on point clouds for (self-)supervised scene flow estimation//16th European Conference on Computer Vision. Glasgow: Springer: 88-107 [DOI: 10.1007/978-3-030-58558-7_6]
- Wu Z R, Song S R, Khosla A, Yu F, Zhang L G, Tang X O and Xiao J X. 2015. 3D ShapeNets: a deep representation for volumetric shapes//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE: 1912-1920 [DOI: 10.1109/CVPR.2015.7298801]
- Xu C F, Wu B C, Wang Z N, Zhan W, Vajda P, Keutzer K and Tomizuka M. 2020. Squeezesegv3: Spatially-adaptive convolution for efficient point-cloud segmentation//Computer Vision Springer: 1-19. Springer International Publishing [DOI: 10.1007/978-3-030-58604-1_1]
- Xu J L, Xiao L and López A M. 2019. Self-supervised domain adaptation for computer vision tasks. IEEE Access, 7: 156694-156706 [DOI: 10.1109/ACCESS.2019.2949697]
- Xu Q G, Zhou Y, Wang W Y, Qi C R and Anguelov D. 2021. SPG: unsupervised domain adaptation for 3D object detection via semantic point generation//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE: 15426-15436 [DOI: 10.1109/ICCV48922.2021.01516]
- Yan H L, Ding Y K, Li P H, Wang Q L, Xu Y and Zuo W M. 2017. Mind the class weight bias: weighted maximum mean discrepancy for unsupervised domain adaptation//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE: 945-954 [DOI: 10.1109/CVPR.2017.107]
- Yan Y, Mao Y X and Li B. 2018. SECOND: sparsely embedded convolutional detection. Sensors, 18(10): 3337 [DOI: 10.3390/s18103337]
- Yang J H, Shi S S, Wang Z, Li H S and Qi X J. 2021a. ST3D: self-training for unsupervised domain adaptation on 3D object detection//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE: 10363-10373 [DOI: 10.1109/CVPR46437.2021.01023]
- Yang J H, Shi S S, Wang Z, Li H S and Qi X J. 2023. ST3D++: denoised self-training for unsupervised domain adaptation on 3D object detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(5): 6354-6371 [DOI: 10.1109/TPAMI.2022.3216606]
- Yang S Q, Wang Y X, van de Weijer J, Herranz L and Jui S L. 2021b. Exploiting the intrinsic neighborhood structure for source-free domain adaptation. arXiv: 2110.04202 [DOI: 10.48550/arXiv.2110.04202]
- Yew Z J and Lee G H. 2020. RPM-Net: robust point matching using learned features//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE: 11821-11830 [DOI: 10.1109/CVPR42600.2020.01184]
- Yi L, Gong B Q and Funkhouser T. 2021. Complete and label: a domain adaptation approach to semantic segmentation of LiDAR point clouds//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE: 15358-15368 [DOI: 10.1109/CVPR46437.2021.01511]
- Yoo H and Jun K. 2023. Point cloud classification by domain adaptation using recycling max pooling and cutting plane identification. Sensors, 23(3): 1177 [DOI: 10.3390/s23031177]
- Yoon D, Tang T M and Barfoot T. 2019. Mapless online detection of dynamic objects in 3D lidar//2019 16th Conference on Computer and Robot Vision (CRV). Kingston: IEEE: 113-120 [DOI: 10.1109/CRV.2019.00023]
- Yuan Z M, Wen C L, Cheng M, Su Y F, Liu W Q, Yu S S and Wang C.

2023. Category-level adversaries for outdoor LiDAR point clouds cross-domain semantic segmentation. *IEEE Transactions on Intelligent Transportation Systems*, 24(2): 1982-1993 [DOI: 10.1109/TITS.2022.3219853]
- Zhang H Y, Cisse M, Dauphin Y N and Lopez-Paz D. 2018. Mixup: beyond empirical risk minimization. arXiv: 1710.09412 [DOI: 10.48550/arXiv.1710.09412]
- Zhang W C, Li W and Xu D. 2021. SRDAN: scale-aware and range-aware domain adaptation network for cross-dataset 3D object detection//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE: 6765-6775 [DOI: 10.1109/CVPR46437.2021.00670]
- Zhao S C, Wang Y Z, Li B, Wu B C, Gao Y, Xu P F, Darrell T and Keutzer K. 2021. EPointDA: an end-to-end simulation-to-real domain adaptation framework for LiDAR point cloud segmentation. arXiv: 2009.03456 [DOI: 10.48550/arXiv.2009.03456]
- Zhuang F Z, Qi Z Y, Duan K Y, Xi D B, Zhu Y C, Zhu H S, Xiong H and He Q. 2021. A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1): 43-76 [DOI: 10.1109/JPROC.2020.3004555]
- Zou L K, Tang H, Chen K and Jia K. 2021. Geometry-aware self-training for unsupervised domain adaptation on object point clouds//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE: 6383-6392 [DOI: 10.1109/ICCV48922.2021.00634]
- Zou Y, Yu Z D, Vijaya Kumar B V K and Wang J S. 2018. Unsupervised domain adaptation for semantic segmentation via class-balanced self-training//15th European Conference on Computer Vision. Munich: Springer: 297-313 [DOI: 10.1007/978-3-030-01219-9_18]

Domain adaptation learning for 3D point clouds: A survey

FAN Wenhui¹, LIN Xi¹, LUO Huan¹, GUO Wenzhong¹, WANG Hanyun², DAI Chenguang²

1.College of Computer and Data Science, Fuzhou University, Fuzhou 350108, China;

2.School of Surveying and Mapping, Information Engineering University, Zhengzhou 450001, China

Abstract: Three-dimensional (3D) point cloud data have been widely used in many fields, such as autonomous driving, robotics, and high-precision mapping. At present, the state-of-the-art deep learning-based methods for 3D point cloud processing are mainly supervised learning methods. The performance of these methods depends heavily on large-scale, high-quality annotated datasets. However, annotating a large-scale, high-quality, category-diverse, and scenario-rich dataset is time-consuming and labor-intensive. In particular, obtaining sufficiently large numbers of samples for model optimization is also quite difficult in some special cases. In addition, 3D point cloud processing models trained on a single device in a special environment are difficult to generalize to different devices and environments. Their generalizability to various devices and environments is limited. Thus, how to reduce dependencies on high-quality annotated 3D point cloud datasets and how to improve the generalizability of current point cloud processing models are important research topics. In recent years, various kinds of impressive and elaborate technologies, such as meta-learning, few-shot learning, transfer learning, self-supervised learning, semisupervised learning, and weakly supervised learning, have been proposed to solve this problem. As an important research branch of transfer learning, domain adaptive learning aims to eliminate differences in feature distributions across domains and promote the generalization ability of deep learning models, thereby providing a novel solution to address this problem effectively. The academic community has conducted preliminary research on domain adaptive learning for point cloud processing. However, the domain adaptive learning field for point clouds still requires in-depth and effective exploration. Consequently, this study systematically summarizes and classifies recent 3D point cloud domain adaptive learning methods into five categories: adversarial learning, cross-modal learning, pseudo-label learning, data alignment, and other kinds of methods. First, we present the mathematical definition of the domain adaptive learning task and depict the chronological overview of the development of different domain adaptive learning methods to provide readers with a clear understanding. Second, we present the general solution for each category of domain adaptive learning methods and summarize the advantages and disadvantages of the current methods for each category. Third, we compare the performance of current methods on three-point cloud processing tasks, including 3D shape classification, 3D object detection, and 3D semantic segmentation. For each task, we also summarize the commonly used datasets and evaluation metrics for an intuitional comparison. Finally, we conclude the advantages and disadvantages of these five categories of methods and discuss future research directions about the 3D point cloud domain adaptive learning.

Key words: remote sensing, 3D point cloud, domain adaption learning, adversarial learning, cross-modal learning, pseudo-label learning, data alignment

Supported by National Natural Science Foundation of China (No. 42271457); Natural Science Foundation of Fujian Province (No. 2023J01430); Program of Song Shan Laboratory (included in the management of Major Science and Technology Program of Henan Province) (No. 221100211000-02)