OASIS：大规模城市人行道自动评估

Yuxiang Zhang、Suresh Devalapalli、Sachin Mehta 和 Anat Caspi

arXiv:2303.02287v2 [cs.CV] 4 May 2023

摘要—**为监测和维护建成环境，确保社区的步行适宜性、轮椅通行性、安全、主动交通和可持续性，有必要检查公共通行权（PROW）范围内的无障碍障碍物。然而，由调查员或公众进行公共通行权检查，既费力、结果不一致，又成本高昂，且难以大规模开展。智慧城市发展的核心在于运用信息技术评估和管理市政资产。与机动车道路相比，人行道尚未定期纳入信息系统，以优化或指导市政服务。我们开发了开放式自动人行道巡检系统（OASIS），这是一套免费、开源的自动化制图系统，可利用移动物理设备提取人行道网络数据。OASIS 运用神经网络、图像感知、基于位置的方法和紧凑型硬件方面的进展，进行人行道分割与制图，并识别障碍，从而生成地理信息系统行人交通图层，可用于路径规划以及分析和运营报告。我们介绍了一套原型系统，该系统使用真实场景中采集的图像进行训练和测试，同时由当地公共交通路径审查团队的人工调查员参与评估。试点结果显示，路径制图的精确率和召回率表现良好（分别为 0.94 和 0.98）。此外，调查员团队的现场工作效率也有所提升。OASIS 在设计上充分考虑了推广应用的需求，确保系统易于整合到政府路径审查团队的工作流程中，且生成的数据能够与公共数据共享平台互操作。**

索引词—**行人制图、首末公里、图像处理、目标检测、语义分割**

SIDEWALKS 是公共建成环境的核心，提供了一种出行方式，连接着几乎所有其他出行选择和社区活动。然而，并非所有人都能安全、无障碍地使用人行道。许多出行者，尤其是残障人士，会遇到人行道障碍，例如陡坡、损坏的混凝土路面或路缘障碍物，这些问题限制了他们享受城市生活益处的机会。关于路径、路径连通性、可避开的障碍物以及沿途其他便利设施的详细网络化信息，能够帮助有不同行动需求的行人找到适合自己的路线、规划行程，并提前了解他们可能会在实地遇到的情况 [1]——[3]。此外，在美国，根据《康复法》（1973 年）第 504 条以及《美国残疾人法》（ADA）（1990 年）第二编 [4], , 任何受 ADA 联邦法规约束的公共实体都必须对其公共通行权（PROW）进行自我评估，以识别路径中的障碍并制定整改计划。尽管驾驶路线规划、资产管理和交通规划推动了对机动车道路进行广泛的制图、测绘和监测（例如 [5]–[8]），街道旁的人行环境却没有得到定期测绘或监测 [9]。美国聚焦于公共通行权（PROW）的研究表明，政府机构往往不了解自身的公共通行权及其中的障碍；公共通行权包括供公众使用的市政环境和路径（包括人行道和人行步道）[10], [11]。值得注意的是，公共和私营利益相关方都需要描述人行道网络的可路由图，并能从中受益。公共利益相关方需要收集和监测人行道环境数据，涉及急救人员前往紧急地点的路径规划、道路施工协调员、安全上学路线计划、零愿景计划，以及负责管理门到门交通服务并确定乘客资格的辅助公共交通团队。私营用途则包括运输与货运应用（用于识别从路缘到门口的路线），以及街道旁的自主导航设备，例如配送机器人和自动驾驶轮椅。有证据表明，残障行人在人行道上导航所需的制图信息与现有地图信息之间存在的信息鸿沟，会严重降低交通网络的运行效能，并影响出行者的安全与保障 [12]–[14]。我们的目标是推出一套系统，用于自动、持续地监测行人路径网络，并提供一种分析软件解决方案，利用街景影像获取有用、开放、共享的行人网络数据；这些数据具备地理准确性、一致性，并可供公共和私营利益相关方使用。预期成果是开放、共享、可解读的数据，以增强现有工具、可视化内容和指标，从而大规模管理交通网络、无障碍性和可通行性 [15]。

我们推出了一套开源系统，并通过与金县都会交通无障碍服务合作进行测试，完成了概念验证实施；该合作项目也是一项应用案例研究。我们的试点在三个不同的市政环境中进行了测试，以展示该系统在真实环境中以及由实际机构使用时的可行性。

该系统的贡献在于开放式自动人行道检测系统（OASIS）。这是一种分析系统，可通过由面向消费者的现成电子设备组成的人工引导设备，推断并输出人行道网络。该系统包含多个步骤：OASIS 捕获街景图像，对场景中的基础设施和街道设施进行分割，并实时生成用于人行道评估的制图信息；根据用户偏好，还可能避免存储图像数据。OASIS 构建于低功耗、轻量级的边缘设备之上，可轻松集成到其他电动

预印本。审稿中。通讯作者：Yuxiang Zhang（邮箱：yz325@uw.edu）

移动设备中。通常由人工调查员完成的人行道评估流程现已实现自动化，能够以高效且可扩展的方式协助进行人行道制图和审查。我们通过 OASIS 展示了为多个市政当局生成路径评估数据的能力，这些数据采用标准化格式，可供其他利益相关方重复使用，或由下游应用使用。

我们将在以下各节介绍这项工作和案例研究。第I节介绍试点研究的现实背景和合作机构。第II 节讨论户外交通环境制图方面的相关工作和既有研究。第III 节详细介绍构建本系统所采用的实施方案和方法。第IV节介绍概念验证初步实施所用的硬件。试点研究的定量和定性结果见第V节，随后在第VI节进行讨论并总结。

I. 案例研究利益相关方

联邦政府要求交通机构提供符合 ADA 标准的辅助性无障碍交通服务。这类交通服务面向残障人士，用于补充他们无法使用的固定线路服务。金县都会交通局（华盛顿州）通过 Metro Access 项目提供此类辅助交通服务。Metro Access 直接关注公共通行权范围内的行人通行能力评估：（1）与其他区域公共交通服务提供方一样，该机构有法律义务为符合资格、且无法通过公共通行权范围内的道路合理到达其他交通方式的个人提供点对点辅助性无障碍交通服务；（2）该机构必须针对每位个人，有时还需针对每项交通申请，通过当地公共通行权调查确定其服务资格；（3）在可行的情况下，优先引导客户使用固定线路公共交通，沿着对骑行者而言可行且实用的路线出行，但这些路线未必是起点与固定线路公交车站之间最直接的路径。自 2006 年以来，Metro Access 的工作人员组成了一支路径审查员团队，逐步建立起详细的路径及障碍数据库，其中包括不平坦地形、高于八度的高差、无标记交叉路口以及不合规或缺失的路缘坡道。随着时间推移，Metro Access 扩大了这项工作，组建专职团队，针对骑行者频繁出行的特定路径，实地收集数据。人工收集的数据涵盖金县 1019 千米的人行道和过街设施。团队会实地走访行程的起点和目的地，并审查行程两段通往公交车站的路径。审查内容包括现场测量和拍照。资格审查对交通机构和骑行者都有益处。引导骑行者改用固定线路交通系统，可腾出资金用于其他服务（包括投入辅助性无障碍交通系统），也能为提出乘车申请的残障人士提供更好的体验：辅助性无障碍交通系统必须提前数日预约，接送时间范围也较宽；而固定线路公共交通通常可由乘客自行使用，且班次间隔可能更便利。

截至2019年，路径审查团队收集的路径报告并不适合用于创建长期需求评估所需的行人网络：报告遗漏了绝大多数相关的公共行人通道；潜在障碍物和阻碍物以独立于路径的点位形式存储（因此，若不进行大量转换和猜测，就无法通过软件自动评估）。此外，数据的“陈旧程度”难以评估；而且数据收集时有意保持不完整，并带有主观性：审查路径主要是为了满足申请辅助交通服务的行人的需求，因此，凡是被认为与该资格申请无关的信息都不会收集。路径审查团队的许多工作都是重复的，因为他们收集并存储了其他 GIS 服务（如 OpenStreetMap、Google 等）上已 readily 可获取的兴趣点（POI）信息。最后，数据可能会被用来推断个人身份信息（PII），这不仅带来了隐私风险，也增加了共享和利用数据的难度。

Metro Access 是又一个用户案例，说明行人环境中需要全面绘制的地图信息，也呼应了其他研究提出的关注点 [9], [16], [17]。正如 [18] 所述，Metro Access 团队面临的挑战在其他利益相关方中也普遍存在。总体而言，调查和制图任务十分耗时；如果定期依靠人工采集，则难以实施。首先，由于需要大量数据，数据采集与分析的工作量相当可观。其次，即使完成了数据采集，也无法涵盖环境未来的变化，更难以据此轻松维护数据（例如人行道改道或重建）。目前，城市没有设备能够以一致且可扩展的方式采集这类数据，也尚未采用开放数据标准，以便在城市之间以及各组织之间共享数据。

近年来，Metro Access 与合作伙伴一直在使用名为 OpenSidewalks [19] 的开放数据规范，采集地理信息系统（GIS）图层，以存储这些本地调查的结果供日后使用：如果已有已知路径或障碍物，便可通过远程审核确认资格；调查也可聚焦于自上次审核以来基础设施发生的具体变化，而非重复观察 [20]。不过，人工审核的普遍挑战依然存在。作为试点合作伙伴，我们不仅希望解决自动化采集 OpenSidewalks 交通图层及相关测量数据中繁重环节的技术难题，还希望通过应对组织层面对个人身份信息保护的顾虑，以及避免大规模影像存储和归档所涉及的成本与安全问题，提升技术的采用可能性。

II. 相关工作

近年来，部分得益于硬件的改进以及计算上切实可行的神经网络的出现，利用图像数据进行制图已成为一个热门研究领域。这种方法主要应用于汽车环境制图。该领域也提出并实施了一些行人环境制图的尝试。

a) 语义分割与汽车环境制图：汽车环境制图已

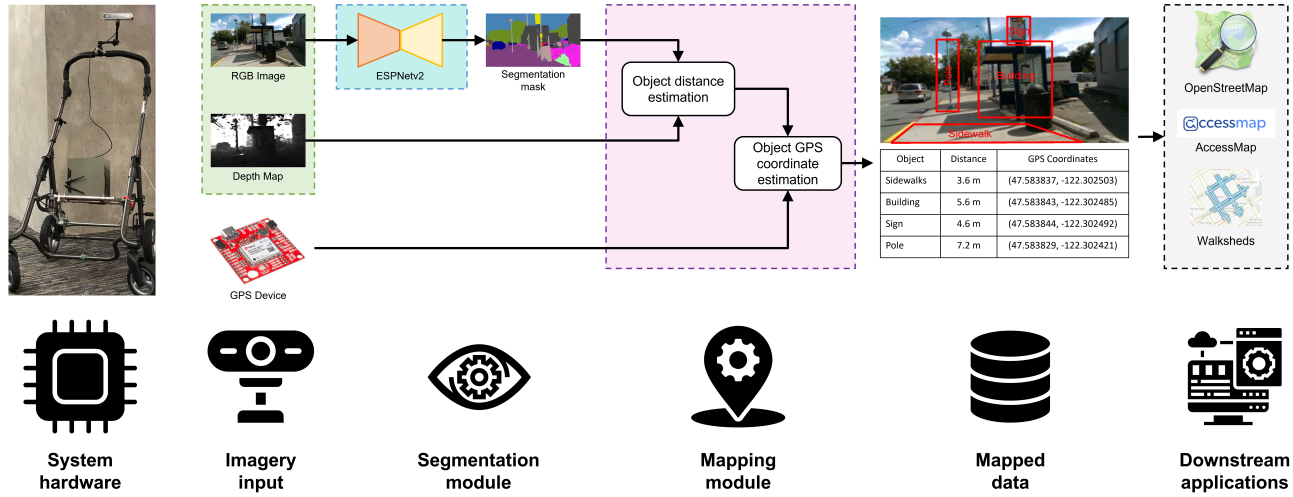

图 1：OASIS 概览：场景中物体的地理位置以及行人路径网络的连通性，均通过立体相机和 GPS设备采集的数据在设备端推断得出。

得到广泛研究，而语义分割在汽车环境制图中发挥着重要作用 [21]–[24]。已有多种语义分割架构被提出，包括编码器-解码器 [25], [26], 、基于区域的架构 [27], [28],以及级联网络 [29]。值得注意的是，用于汽车制图任务的大多数语义分割网络计算开销很大，不适合在边缘设备上进行制图。在本研究中，我们扩展了高效分割架构 ESPNetv 2 [30], ，用于在边缘设备上绘制人行道。作为基础模型，ESPNetv2 在提升能效的同时，达到了与其他先进网络相近的性能，其中包括 MobileNetv2 [31]和 ShuffleNetv 2 [32]。

b) 行人环境地图绘制：尽管分割模型在汽车环境中的实用性已得到证实，但在行人环境中的应用尝试仍然有限。目前没有广泛使用的基准数据集包含专门针对行人环境采集并标注的街景图像。大多数有关行人路径分割的研究都侧重于使用卫星航拍图像 [33], [34],，或结合卫星图像与街景图像，但仅应用于特定的单一目标类别。例如， [35] 利用卫星图像和 Google街景图像来检测和定位街道过街点。这些应用通常是离线的，主要原因是它们不包含任何关于行人环境的实时信息。其他针对行人环境的在线研究则聚焦于基于SLAM的障碍物检测与避让 [36], ，但这类方法不会生成或存储可供下游应用使用的制图信息，无论是用于行人导航（AccessMap [37],[38],、SidewalkScore [15], 还是 Pathnav [39]）。据我们所知，目前仅有另一种针对步行路径的硬件解决方案，但它只涉及路面完整性指标的标准化（即测量裂缝和不平整的路面扰动），生成专有数据，且不提供行人交通图，以支持下游路径规划应用 [40]。其他大规模绘制行人路径网络的研究完全基于现有街道（道路）数据。例如，Li等人 [41] 提出了一种半自动化方法，使用道路中心线数据。该方法会生成理论上应当存在的人行道，但仍需人工编辑。

III. OASIS：开放式自动人行道巡检系统

本节介绍开放式自动人行道巡检系统（OASIS），该系统用于在通常称为边缘设备的便携式设备上绘制人行道地图。如图 1 所示，OASIS 由两个主要模块组成：（1）分割模块（第 III-A 节）和（2）地图绘制模块（第 III-B 节）。分割模块将输入图像中的每个像素分类为感兴趣的语义类别。地图绘制模块以语义分割掩码、深度图和摄像机的 GPS 坐标为输入，并输出环境中每个物体的 GPS 坐标。这不仅有助于我们了解环境，还能将系统识别到的物体登记到 OpenStreetMap（OSM）等地图数据库中，以便将这些信息用于下游应用。图1 展示了系统的整体结构，以下各小节将介绍 OASIS 的各个组成部分。

A. 分割模块

语义分割将一个场景划分为若干有意义的部分，将其分解为像素，并为每个像素标注其所属的物体类别。OASIS围绕户外行人环境映射场景中具有实际意义的物体类别构建。这些类别包括路径（道路、人行道）以及不可通行障碍物

，后者可能是静态的（建筑物、墙、围栏、杆、交通信号灯、交通标志、植被），也可能是移动的（行人、骑行者、汽车、卡车、公交车、火车、摩托车、自行车）。此外，我们还对天空和地形进行分割，以避免在这些区域将街道两侧的物体误检为目标。每个像素都会被分配一个整数 ID，用于表示其是否属于相应的目标类别。分割过程以一幅图像 $I \in Z^{3\times w\times h}$作为输入，分割模块为图像 I中的每个像素分配这些整数 ID 之一后，会生成一幅名为 $S \in Z^{w\times h}$图像分割掩膜的新图像。需要注意的是，尽管 OASIS 能够对上述所有类别中的物体进行分割，但在该原型系统中，我们选择用于进一步分析的类别仅限于行人环境中的静态对象：杆、交通信号灯、交通标志和建筑物。这些对象通常出现在街道环境中，并且通常难以或过于耗时地手动绘制每个实例。如第II节所述，当前最先进的分割模型是卷积神经网络（CNN）。我们选择 ESPNetv 2 [30] 作为原型系统中的基础分割模型。ESPNetv2 在保持与其他最先进模型相当结果的同时，大幅降低了所需的计算能力，这对于将模型部署到边缘设备上十分重要。双目相机 I 捕获的每幅 RGB图像都作为 ESPNetv2 的输入，并生成相应的分割掩码 S。分割掩码中的每个像素都具有一个整数值，表示模型预测的所有类别中概率最高的类别。

尽管卷积神经网络模型在户外环境中具有较高的已报告分割准确率，但用于户外制图的模型通常是在从车辆仪表盘采集的图像上训练的，例如 Cityscapes [5] 和 KITTI [42]。由这种数据集偏差引入的误差，在我们的步行环境地图绘制领域中不可忽略。我们采用两种方法来解决这一问题。首先，我们的应用旨在摄像机沿人行道行进时执行语义分割。这意味着，捕获并输入 OASIS 的图像在时间上是连续的。我们不再分别使用图像流中的每一帧进行推断，而是提出利用连续帧中的时间信息来帮助改进推断结果。我们计算卷积神经网络在若干时间上相邻帧中生成的掩码的并集。这种直接的方法已被证明能够减少过度分割错误，并提高整体预测准确率。其次，我们观察到，在训练样本与测试样本之间视角发生变化时，人行道是受影响最大的类别，因此我们在更加多样化的数据集 COCO-Stuff 上重新训练模型 [43]，并在包含我们在多种人行道环境中采集的手工标注图像的数据集上进行微调。由于使用从汽车仪表盘采集的图像训练的模型固有地偏向于认为人行道更有可能位于输入图像的两侧，因此在汽车数据集上训练的模型通常会将大部分人行道错误地预测为机动车道（示例见图 4 ）。对于需要了解人行道位置及其连通性的应用而言，这种错误是不可接受的。在我们的以行人为中心的数据集上重新训练后，模型的预测准确率在真实人行道环境中得到了显著提升。我们将在第V-B1节讨论分割模块的评估。

B. 地图绘制模块

本节详细介绍地图绘制模块的设计。利用分割模块的输出，地图绘制模块会跟踪并登记场景中的每个静态对象（第III-B1节），并通过推断每个人行道片段的位置和宽度，推断人行道的步行适宜性和连通性（第III-B2节）。

1）跟踪并登记静态对象：生成每张图像的分割掩码 S I 后，将其与对应的深度图结合，以计算物体与系统之间的距离。深度图是一种图像（数组），其中每个像素都包含该像素处三维场景中各点到摄像机的深度信息，其视点与拍摄对应RGB图像时的视点相同。RGB图像及其对应深度图的示例见图1 。

利用分割掩码，我们找出场景中代表每个已分割物体的像素。计算深度图中这些像素值的算术均值，作为该已分割物体的深度。获取每个已分割物体的距离信息后，将该距离与系统的GPS信息结合，以估算当前场景中每个物体的GPS位置。系统提取拍摄图像时的纬度和经度，然后利用物体的像素坐标和相机朝向推断其相对方向。根据起点（系统位置）的GPS位置、相机朝向以及推断出的距离，使用Haversine公式即可估算摄像机所见每个物体的GPS位置。

由于我们的目标是推断人行道环境的地理信息，因此，如果同一静态物体出现在不同帧中，系统就需要能够识别它。为了在连续帧中跟踪每个独特的物体，我们使用质心跟踪方法实现了一个目标跟踪器。利用分割掩码，可以计算出每个分割物体的质心（几何中心）。当摄像机采集到新帧时，目标跟踪器会接收当前帧中由卷积神经网络模型生成的一组新分割物体的质心坐标和多边形。目标跟踪器会尝试将新物体与先前帧中的现有物体进行匹配。具体做法是计算每对新物体质心与现有物体质心之间的欧氏距离。每个新物体都会与欧氏距离最小的现有物体关联。如果现有物体的数量多于新物体，则无法与新物体匹配的现有物体会被标记为暂时丢失；如果它们处于暂时丢失状态的帧数超过设定阈值（在我们的原型系统中为 5 帧），则会被标记为永久丢失。如果现有物体的数量少于新物体，则每个无法与现有物体关联的新物体都会被

分配一个新的对象 ID。由于我们在摄像机移动时跟踪物体，即使物体本身是静态的，场景的移动也会导致物体质心的像素位置发生变化。为纳入这一信息，我们计算相邻帧之间的单应变换，以描述它们在旋转和平移方面的空间关系。然后，我们利用这一变换将每个物体的质心从当前帧映射到下一帧。此步骤使每个现有物体的质心更接近新物体的质心，从而改进对象跟踪过程。第V-B2节评估了静态对象的制图结果。

2）人行道制图：与我们用于静态对象的配准方法不同，我们采用另一种方法来绘制人行道。与其他目标类别不同，对于人行道，我们不仅需要确定其地理位置，人行道的宽度也是衡量行人路径可步行性的重要特征。我们没有尝试在连续帧中跟踪人行道，而是直接使用Haversine公式计算每个人行道片段的 GPS 坐标。随后，通过透视投影，利用相机参数估算每个人行道片段的宽度。

推断出每个人行道片段的地理位置和宽度后，更重要的是，还能推断出行人路径网络的连通性和步行适宜性。为了推断路径连通性，在推断出某个街区的人行道后，如果预测路径网络中任意两个最近的节点之间存在明显间隙，就会将其识别为路径网络断开。推断出的宽度可用于评估人行道的无障碍性和步行适宜性。例如，如果人行道宽度不足以让某些移动辅助设备（如轮椅）通行，就会将其标记为这些使用者无法通行。

IV. 硬件设置

本节介绍我们在试点研究（第 V 节）中使用的原型系统的硬件设置。展示整体硬件设置的示意图见图2。该系统由三个硬件模块组成：（1）成像模块（第 IV-A 节）；（2）GPS 模块（第 IV-A 节）；以及（3）计算模块（第 IV-C 节）。成像模块提供的图像数据和 GPS 模块提供的 GPS 数据（与每张图像同步）在计算模块中进行处理，以生成制图结果。

A. 成像模块

我们使用立体相机和三轴相机夹具作为成像模块。立体相机用于采集 RGB图像及其对应的深度图像（如图 1 所示），这些数据对于推断行人环境中基础设施的地理位置以及人行道连通性至关重要。根据一项比较不同立体相机在实际环境中性能的研究 [44],，为了在多种环境下获得更好的性能，我们使用 ZED 立体相机（ZED）。ZED 立体相机是一种被动立体视觉相机，它通过两侧的两个镜头进行拍摄场景。随后，它会寻找同一场景左视图和右视图之间的匹配点，并通过三角测量方法使用立体视觉算法计算每个感知到的物体的深度。

ZED立体相机与地面平行安装在采集夹具上。RGB图像和深度图的空间分辨率均设置为 1280×720像素。帧率设置为每秒15帧。ZED摄像机的视场角为 $90°\times60°$（水平 × 垂直）。在采集数据之前，按照制造商规格，使用制造商提供的工具对摄像机进行校准。

B. GPS 模块

使用GPS接收器在拍摄图像的同时记录GPS信息。结合图像数据，可利用这些信息推断每个感兴趣对象的地理位置。我们的原型系统所使用的接收器是u-blox ZED-F9P模块（F9P）。在室外行人环境中测试时，其精度为 ±0.75 米。系统的GPS位置每0.5秒使用GPS接收器的数据记录一次。在这一记录频率下，当系统以正常速度在人行道上行进时，将每张捕获的图像与GPS 数据同步后，其位置误差为< 0.7 米，为制图提供了足够的分辨率。

C. 计算模块

该计算设备用于车载语义分割和地理信息推断。由于要充分利用 ZED 摄像机的软件开发工具包 (SDK)，并以最佳性能运行基于CNN的分割模型，需要使用 NVIDIA 的 GPU，因此我们选用功耗低于 10 瓦的 NVIDIA Jetson Nano 作为计算设备。借助该设备，我们能够通过便携式边缘设备绘制人行道地图。

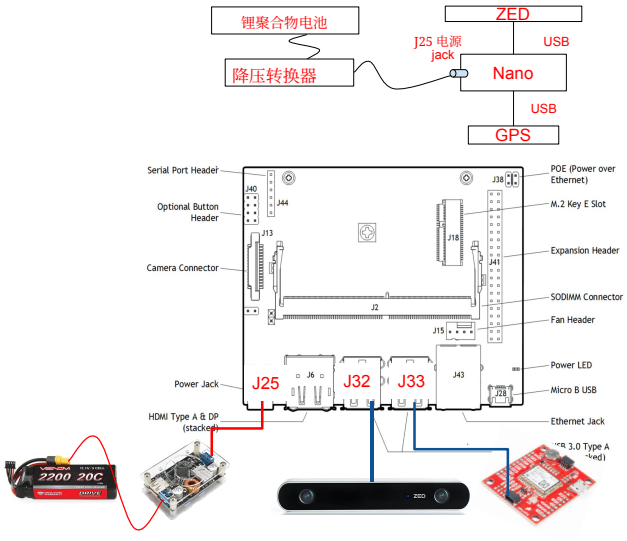

图 2：OASIS 系统示意图

表 I OASIS 制图期间的平均资源使用情况

功耗模式	平均功耗	内存使用量	GPU利用率	CPU利用率
低功耗模式（5w）	3.98 W	3006 MB	82%	11.4%@921MHz
高功耗模式（10w）	5.87 W	3058 MB	84%	10.5%@1479MHz

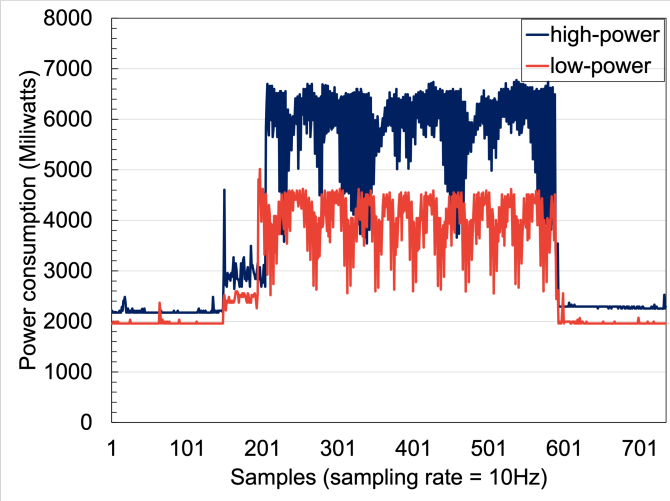

图3：OASIS 在两种不同功耗模式下制图时的功耗

V. 试点研究与实验

本节介绍我们使用 OASIS 开展的试点研究。该研究由两部分组成。第 V-A 节评估了 OASIS 在硬件层面的表现。第V-B节中，我们使用 OASIS 为 3 个街区绘制地图，并将其制图结果与当地辅助交通路径审查团队的人工调查员所绘制的结果进行比较。

A. OASIS 的硬件层面分析

在 OASIS 于行人环境中以两种功耗模式持续制图时，对系统性能进行测试。OASIS 可在两种功耗模式下运行：(1) 低功耗模式（最大功耗：5 W）和 (2) 高功耗模式（最大功耗：10 W）。我们在两种功耗模式下评估了 OASIS，并测量了多项硬件指标，包括功耗、内存使用量、CPU使用率和 GPU使用率，以进行硬件层面分析。图 3展示了 OASIS 在两种不同功耗模式下运行时功耗随时间的变化，同时还比较了 OASIS 在每种模式下空闲时的基础功耗。表 I 展示了 OASIS 在系统持续运行并生成制图结果的 5 分钟内的平均资源使用情况。总体而言，这些指标，尤其是较低的功耗（低功耗模式下约为 4.5 瓦，高功耗模式下约为 6.5 瓦），表明 OASIS 可轻松集成到电动移动设备（例如轮椅或电动代步车）中，为下游应用供电。

B. 使用 *OASIS* 分析行人环境制图

本节介绍我们在真实行人环境中对人行道和基础设施进行制图的研究，并将制图结果与人工调查员的结果进行比较。这些调查员来自当地的辅助交通路径审查团队。研究在三个地理位置各异的地区开展：贝尔维尤、雷德蒙德和西雅图。

1) 分割模块：我们首先通过比较不同的分割模型，评估分割模块在行人环境中的性能。我们以在常用的 Cityscapes数据集上训练的 ESPNetv2 作为基线模型；作为对比，ESPNetv2-PED 指的是我们根据第III-A节所述方法，为行人环境训练的模型。模型的性能如图4所示。中间列展示了 ESPNetv2 的定性分割结果，右侧列展示了 ESPNetv2-PED 的分割结果。由于训练数据是在汽车仪表盘上采集的，ESPNetv2 存在固有偏差，因此会将大多数人行道错误地预测为机动车道，导致无法生成正确的人行道地图信息。另一方面，ESPNetv2-PED 在训练中融入以行人为中心的数据后，能够正确分割人行道，从而正确推断人行道地图信息。在定量评估方面，我们采用三个标准指标来评估分割模块：(1) 逐像素交并比 (IoU)、(2) 精确率和 (3) 感兴趣对象的召回率。如表II所示，常见对象的精确率和召回率较高，表明系统能够成功检测真实行人环境中的对象。我们还观察到，ESPNetv2-PED 在行人环境的所有类别中均优于基线模型。

2）地图绘制模块：当系统沿着行人路径行进时，摄像机拍摄街景图像，系统对其进行处理以生成制图预测。图5a 展示了一些示例，系统会分割出行人感兴趣的基础设施，并推断其距离和地理位置。平均而言，每千米人行道上可识别并定位 62 个可能妨碍行人通行的障碍物。例如，图5a 中的公交车站被预测为建筑物，在导航场景中会被视为障碍物；这类似于建筑物的一部分延伸到人行道上、阻挡行人通行的情况。除了绘制人行道上的基础设施外，更重要的是，我们还对人行道进行制图，以推断行人路径连通性和步行适宜性。图5b展示了绘制出的人行道片段示例，以及推断得到的地理位置和宽度。利用这些片段的推断制图结果，可以生成行人路径网络。

为了定量评估地图绘制模块，我们将每个静态物体的预测地理位置与其在经过人工标注的帧中的真实地理位置进行比较。需要指出的是，我们已在先前的工作中研究了不同环境条件下的距离估算误差 [44]，因此本研究仅测量位置误差。表中展示了我们所绘制的目标物体的位置误差（以两点之间的距离，单位为米，进行测量）III。较低的平均误差表明，OASIS 能够有效绘制人行道上的常见基础设施。更重要的是，为了评估预测的人行道网络图，我们提供了以下定量指标

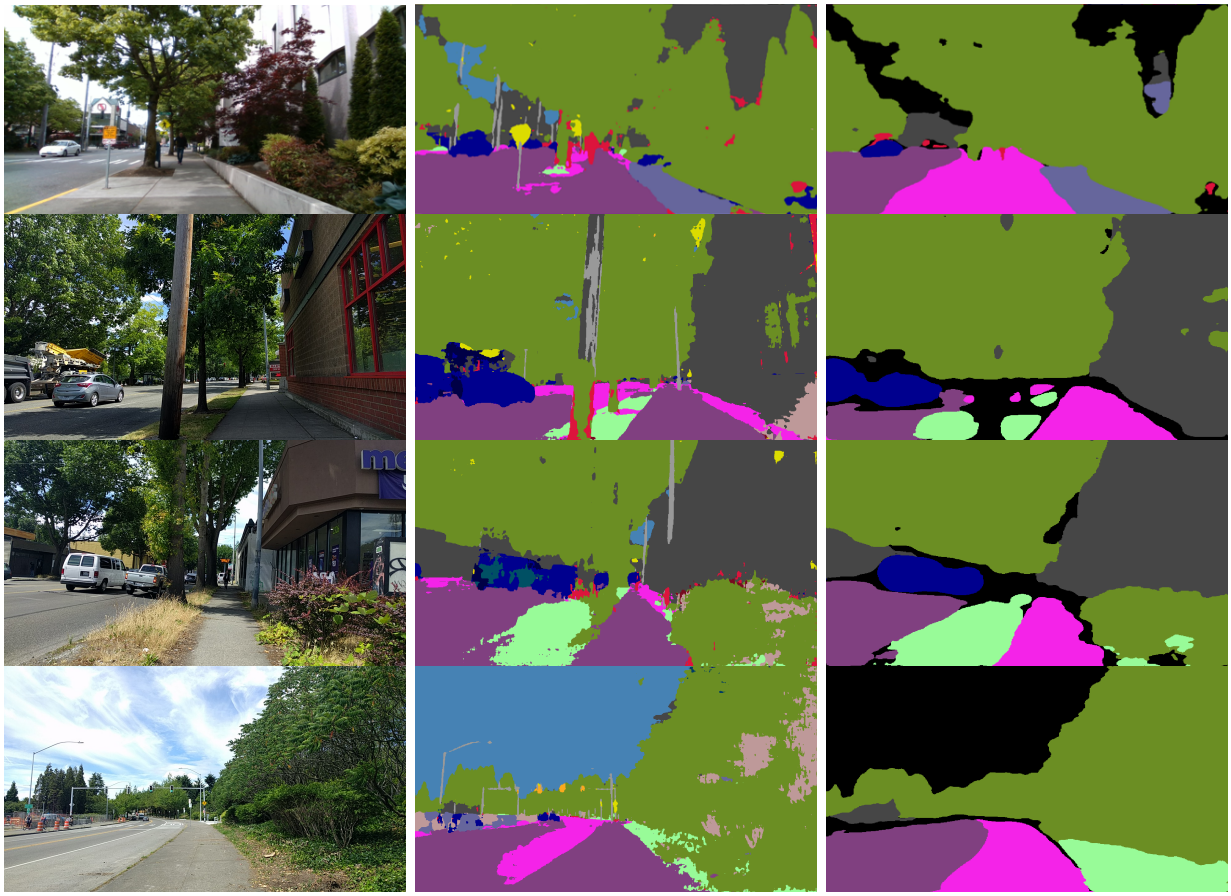

图 4：分割模块的定性表现。左：输入图像。中：ESPNetv2 生成的分割掩码。右：ESPNetv2-PED 生成的分割掩码。在分割掩码中，人行道以洋红色表示，机动车道以深紫色表示，地形以浅绿色表示。

表 II 行人环境中分割模块的性能

	建筑物			Pole			交通信号灯			交通标志			平均值		
	IoU	精确率	召回率	IoU	精确率	召回率	IoU	精确率	召回率	IoU	精确率	召回率	IoU	精确率	召回率
ESPNetv2	0.815	0.654	0.895	0.441	0.841	0.659	0.423	0.577	0.645	0.539	0.781	0.618	0.547	0.713	0.704
ESPNetv2-PED	0.869	0.691	0.932	0.502	0.882	0.721	0.485	0.613	0.713	0.572	0.815	0.686	0.607	0.750	0.763

表中展示了三个城市中每座城市预测的位置和宽度IV。我们将预测结果与城市收集并维护的数据进行比较。较低的均方根误差表明，该系统能够有效预测人行道的位置和宽度。此外，为了在网络图层面评估绘制出的人行道，我们采用了与 [45] 提出的道路拓扑测量方法相似的指标，即根据预测的人行道边（或交叉边）与真实图中对应边的匹配情况计算精确率、召回率和 F1 分数。如表IV最后三列所示，我们的方法在建成环境差异显著的三个区域中均保持了较高的精确率和召回率，证明了该方法的制图效果和准确性。

表 III 映射静态对象位置的误差测量

	建筑物	Pole	交通信号灯	交通标志	平均值
平均误差（米）	0.511	0.620	0.424	0.441	0.496
误差标准差（米）	0.217	0.122	0.102	0.089	0.136
均方根误差（米）	0.554	0.632	0.436	0.449	0.514

表 IV 定量评估：OASIS 绘制的人行道与人工标注的对比

	位置误差（米）			宽度误差（米）			精确率	召回率	F1
	均值	标准差	均方根误差	均值	标准差	均方根误差			
雷德蒙德	0.47	0.12	1.49	0.11	0.03	0.11	0.94	0.98	0.96
贝尔维尤	0.68	0.22	0.71	0.28	0.07	0.29	0.92	0.98	0.95
西雅图	0.53	0.11	0.54	0.14	0.04	0.15	0.96	0.99	0.97
平均值	0.56	0.15	0.91	0.18	0.05	0.18	0.94	0.98	0.96

C. 借助 *OASIS* 实现高效且可扩展的路径审查流程

我们研究中的一个主要考量，是 OASIS 能否切实部署并得到采用，以协助开展路径审查。



（a）制图后的基础设施



（b）制图后的人行道

图5：定性制图结果：（a）人行道上基础设施的预测地理位置。（b）人行道的预测地理位置以及每个人行道片段的推断宽度。

在更大规模的应用中支持团队。在试点研究期间，我们发现 OASIS 有五项优势：(1) 系统可在典型步行速度下引导使用者，而绘制一英里人行道地图的平均时间缩短至原来的六分之一（具体而言，未经辅助的审查员平均需要 120 分钟才能采集一英里路径的数据，而使用 OASIS 只需 20 分钟）。OASIS 通过减少人工调查员的工作量，降低了采集人行道制图数据的时间和成本。(2) 如第 V-B节所述，我们的方法能够提供准确的制图信息，而这类信息难以通过人工一致地采集（例如，采集人行道宽度和场景中的每项基础设施十分耗时），此外还能提供准确的人行道位置和连通性数据。(3) 输出数据可立即按照 Opensidewalk 架构规范，以开放、标准化的格式提供 [46]，从而解决人工调查员采集数据时常见的降噪和后处理问题。这有助于数据录入，并提高数据被其他利益相关方或下游应用重复利用和使用的可能性。(4) OASIS 的输出是行人图层中的地图网络数据；出于公众隐私考虑，原始图像数据可以丢弃。(5) OASIS 基于便携式边缘设备构建，可轻松集成到电动移动设备中，且只需极少的人工干预即可轻松维护数据的时效性。尽管还需要开展更多采用情况研究，但试点研究结果表明，在需要快速、准确评估和重新评估人行道的审核和应用中，OASIS 可以得到采用。此外，OASIS 大规模部署后对熟练调查员的依赖极低，因此适用于更大规模的应用和组织工作。

VI. 讨论与结论

帮助城市管理涵盖各类服务和公共基础设施的数据管道的新技术，将成为未来城市发展的驱动力。行人通道、人行道和步道，是致力于引导城市居民选择主动交通方式而非私人车辆的可持续、无障碍城市所重点关注的对象。本文提出了一种新颖的城市人行道评估方法，利用便携式（边缘）设备上的计算机视觉（机器学习）技术。我们开发并测试了 OASIS，并将其用于人行道环境制图。OASIS 使私人或公共组织能够快速、准确度合理地绘制人行道及其连接关系（以可路由图的形式）。

这项技术及其他类似的基础设施制图技术的一项主要贡献，是能够大规模生成一致、标准化的数据。该领域有机会为构建高效且韧性强的数据生态系统贡献力量，为众多利益相关方提供相关数据，包括交通运输机构、市政当局以及公共和私营的民间组织，使他们能够在城市尺度上分析、量化、确定优先级，并可能管理出行与无障碍性。在边缘设备上完成所有计算，对于那些民众会反对使用云基础设施来获取并可能留存人行道影像的市政当局而言，或许颇具吸引力，因为他们可能认为这侵犯了隐私或危及安全。在便携式边缘设备平台上提供这些数据，也为通过地图为在人行道环境中导航的自动化应用提供信息创造了更多机会，例如自动驾驶轮椅和自动驾驶配送机器人。开放并共享制图结果，是实现这一愿景的关键一步，因为这些结果并不依赖任何专有的街景影像贡献。具体而言，在公共通行权（PROW）范围内的无障碍数据方面，先前研究表明，许多规模较小、资金不足的利益相关方在收集相关且一致的数据时面临着相似的问题。通过协作制图工具贡献数据，并借助 OpenStreetMap (OSM) 等协作数据公地进行分发，可以

改善整体非机动车交通的评估与提升。旅客也能从下游的路径规划和寻路应用中受益（例如 Moovel、TransitApp、AccessMap [19], [37]）。使用协作工具和共享数据，能够让公民参与数据审核和收集，从而增进相关机构与无障碍倡导者之间的信任。总而言之，我们认为，采用自动化人行道评估系统的方法，对于未来城市主动交通网络的开发与管理至关重要。

致谢

如果没有 Metro Access 的 Matthew Weidner 以及 MV Transportation, Inc. 团队的远见与领导，这项工作将无法完成（特别感谢 Eric A. Peterson、Nik Wilson 和 Dwight Sanders）。本研究得到华盛顿州交通部研究资助 GrantT1461-47，以及联邦交通管理局授予 Taskar无障碍技术中心的奖项编号 693JJ32250014 ITS4US 的支持，并获 NVIDIA 捐赠设备。

参考文献

[1] P.-A. Quinones、T. Greene、R. Yang 和 M. Newman，“支持视障导航：一项需求探索研究”，载于*CHI'11*计算机系统中的人为因素扩展摘要，2011年，第1645–1650页。[2] S. Zimmermann-Janschitz、S. Landauer、S. Drexel 和 J. Obermeier，“利用地理信息系统为VIB人士提供独立出行支持”，赋能技术杂志，2021年。[3] S. J. Bosch 和 A. Gharaveis，“独自出行：关于视力或认知衰退老年人寻路的文献综述”，应用人体工程学，第58卷，第327–333页，2017年。[4] 《美国法典》，“《美国残疾人法》”，1990年。[5] M. Cordts、M. Omran、S. Ramos、T. Rehfeld、M. Enzweiler、R. Benenson、U. Franke、S. Roth 和 B. Schiele，“用于城市场景语义理解的Cityscapes数据集”，载于*IEEE*计算机视觉与模式识别会议论文集，2016年，第3213–3223页。[6] G. Varma、A. Subramanian、A. Namboodiri、M. Chandraker 和 C. Jawahar，“IDD：用于探索非受限环境中自动导航问题的数据集 ”，载于*2019*年*IEEE*冬季计算机视觉应用会议（*WACV*）。IEEE，2019年，第1743–1751页。[7] P. Sun、H. Kretzschmar、X. Dotiwalla、A. Chouard、V. Patnaik、P. Tsui、J. Guo、Y. Zhou、Y. Chai、B. Caine 等，“自动驾驶感知的可扩展性：Waymo开放数据集”，载于*IEEE/CVF*计算机视觉与模式识别会议论文集，2020年，第2446–2454页。[8] L. Mi、H. Zhao、C. Nash、X. Jin、J. Gao、C. Sun、C. Schmid、N. Shavit、Y. Chai 和 D. Anguelov，“HDMapGen：高精地图的分层图生成模型”，载于*IEEE/CVF*计算机视觉与模式识别会议论文集，2021年，第4227–4236页。[9] N. Bolten、S. Mukherjee、V. Sipeeva、A. Tanweer 和 A. Caspi，“以行人为中心的数据方法，促进城市基础设施环境的公平可及性”，*IBM*研究与发展杂志，第61卷，第6期，第10–1页，2017年。[10] Y. Eisenberg、A. Heider、R. Gould 和 R. Jones，“美国各社区是否在为残障行人制定规划？地方政府障碍清除计划系统评估的发现”，城市，第102卷，第102720页，2020年。[11] Y. Eisenberg、A. Hofstra、S. Berquist、R. Gould 和 R. Jones，“残障人士的障碍清除计划与行人基础设施公平性”，交通研究*D*辑：交通与环境，第109卷，第103356页，2022年。

[12] S. Gallo、D. Chapuis、L. Santos-Carreras、Y. Kim、P. Retornaz、H. Bleuler 和 R. Gassert，“具有多模态触觉反馈的增强型白手杖”，载于2010年第三届IEEE RAS与EMBS生物医学机器人与生物机电学国际会议。IEEE，2010年，第149–155页。[13] R. Pyun、Y. Kim、P. Wespe、R. Gassert 和 S. Schneller，“具有障碍物高度和距离反馈功能的先进增强型白手杖”，载于2013年IEEE第13届康复机器人国际会议（ICORR）。IEEE，2013年，第1–6页。[14] 美国政府问责局，“交通无障碍性：数据缺乏和执法选择有限，制约了联邦监督”，美国政府问责局。[15] N. Bolten 和 A. Caspi，“迈向常态化、城市尺度的无障碍指标：利用个性化行人网络分析，通过图论解释行人通行情况”，PLoS One，第16卷，第3期，文章编号e0248399，2021年。[16] J. Pearlman、E. Sinagra、J. Duvall、R. Cooper、D. Stuckey 和 A. Kortum，“路径测量工具（Pathmet）的开发与特性分析”，技术报告，2014年。[17] J. E. Froehlich、A. M. Brock、A. Caspi、J. Guerreiro、K. Hara、R. Kirkham、J. Sch¨oning 和 B. Tannert，“无障碍地图面临的重大挑战”，Interactions，第26卷，第2期，第78–81页，2019年。[18] N. Bolten 和 A. Caspi，“迈向公共（开放）数据的社区共同生产与管理实践化：行人网络案例研究：社区开放数据实践化中的行人网络案例研究”，载于ACM SIGCAS/SIGCHI计算与可持续社会会议（COMPASS），2022年，第232–247页。[19] N. Bolten 和 A. Caspi，“OpenSidewalks”，2021年。[在线]。网址：https://www.opensidewalks.com/[20] Mass Transit，“金县都会交通局因改善人行道数据荣获全国奖项——CommonPaths为机构提供工具，以便在OpenSidewalks中采集和分发高保真路径数据”，Mass Transit一体化出行最佳实践。[在线]。网址：https://www.masstransitmag.com/safety-security/press-release/21247636/[21] M. Teichmann、M. Weber、M. Zoellner、R. Cipolla 和 R. Urtasun，“MultiNet：用于自动驾驶的实时联合语义推理”，载于2018年IEEE智能车辆研讨会（IV）。IEEE，2018年，第1013–1020页。[22] M. Treml、J. Arjona-Medina、T. Unterthiner、R. Durgesh、F. Friedmann、P. Schuberth、A. Mayr、M. Heusel、M. Hofmarcher、M. Widrich等，“加速自动驾驶中的语义分割”，载于ML-ITS、NIPS研讨会，第2卷，第7期，2016年。[23] D. Wu、M.-W. Liao、W.-T. Zhang、X.-G. Wang、X. Bai、W.-Q. Cheng 和 W.-Y. Liu，“YOLOP：仅需一次观察即可实现全景驾驶感知”，Machine Intelligence Research，第1–13页，2022年。[24] Z. Guo、Y. Huang、X. Hu、H. Wei 和 B. Zhao，“自动驾驶场景理解中基于深度学习的方法综述”，Electronics，第10卷，第4期，文章编号471，2021年。[25] V. Badrinarayanan、A. Kendall 和 R. Cipolla，“SegNet：用于图像分割的深度卷积编码器—解码器架构”，IEEE模式分析与机器智能汇刊，第39卷，第12期，第2481–2495页，2017年。[26] O. Ronneberger、P. Fischer 和 T. Brox，“U-Net：用于生物医学图像分割的卷积网络”，载于医学图像计算与计算机辅助干预国际会议。Springer，2015年，第234–241页。[27] J. Dai、K. He 和 J. Sun，“用于联合目标与背景区域分割的卷积特征屏蔽”，载于IEEE计算机视觉与模式识别会议论文集，2015年，第3992–4000页。[28] H. Caesar、J. Uijlings 和 V. Ferrari，“采用端到端训练的基于区域的语义分割”，载于欧洲计算机视觉会议。Springer，2016年，第381–397页。[29] J. Long、E. Shelhamer 和 T. Darrell，“用于语义分割的全卷积网络”，载于IEEE计算机视觉与模式识别会议论文集，2015年，第3431–3440页。[30] S. Mehta、M. Rastegari、L. Shapiro 和 H. Hajishirzi，“ESPNetv2：轻量、节能且通用的卷积神经网络”，载于IEEE/CVF计算机视觉与模式识别会议论文集，2019年，第9190–9200页。[31] M. Sandler、A. Howard、M. Zhu、A. Zhmoginov 和 L.-C. Chen，“MobileNetV2：倒置残差与线性瓶颈”，载于IEEE计算机视觉与模式识别会议论文集，2018年，第4510–4520页。[32] N. Ma、X. Zhang、H.-T. Zheng 和 J. Sun，“ShuffleNet V2：高效CNN架构设计的实用指南”，载于欧洲计算机视觉会议（ECCV）论文集

，2018，第116–131页。

[33] T. Senlet 和 A. Elgammal，“卫星图像中的遮挡人行道分割”，载于《第21届国际模式识别会议论文集》（ICPR2012）。IEEE，2012年，第805–808页。[34] M. C. Ghilardi、J. C. Jacques 和 I. Manssour，“利用低分辨率卫星图像定位人行横道，以帮助视障人士”，IEEE《计算机图形学与应用》，第38卷，第46期，2016年1月，第30页。[35] D. Ahmetovic、R. Manduchi、J. M. Coughlan 和 S. Mascetti，“留意过街设施：挖掘GIS影像以定位人行横道”，ACM《无障碍计算汇刊》（TACCESS），第9卷，第4期，第1–25页，2017年。[36] Y. Morales、E. Takeuchi、A. Carballo、W. Tokunaga、H. Kuniyoshi、A. Aburadani、A. Hirosawa、Y. Nagasaka、Y. Suzuki 和 T. Tsubouchi，“户外行人道路上的1公里自主机器人导航：‘筑波挑战赛2007’”，载于《2008年IEEE/RSJ智能机器人与系统国际会议论文集》。IEEE，2008年，第219–225页。[37] N. Bolten 和 A. Caspi，“AccessMap”，2021年。 [在线]。网址：accessmap.io [38] N. Bolten 和 A. Caspi，“Accessmap网站演示：大规模个性化无障碍行人出行规划”，载于《第21届国际ACM SIGACCESS计算机与无障碍会议论文集》，2019年，第676–678页。[39] E. Sinagra，“Pathnav的开发：利用智能数据提升无障碍性和步行适宜性的行人导航工具”，技术报告，2019年。[40] D. R. Smith、E. Sinagra 和 J. Duvall，“轮椅使用者可接受的拼块式混凝土路面表面特性。”[41] H. Li、J. Cebe、S. Khoeini、Y. Xu、C. Dyess 和 R. Guensler，“一种半自动生成基于GIS的人行道网络的方法，用于资产管理和行人无障碍评估”，《交通研究记录》，第2672卷，第44期，第1–9页，2018年。[42] A. Geiger、P. Lenz、C. Stiller 和 R. Urtasun，“视觉与机器人技术的结合：KITTI数据集”，《国际机器人研究杂志》，第32卷，第11期，第1231–1237页，2013年。[43] H. Caesar、J. Uijlings 和 V. Ferrari，“COCO-Stuff：情境中的物体与背景类别”，载于《2018年IEEE计算机视觉与模式识别会议（CVPR）论文集》。IEEE，2018年。[44] Y. Zhang 和 A. Caspi，“多样户外环境中的立体图像深度感知：实践考量”，载于《第二届ACM/EIGSCC智慧城市与社区研讨会论文集》，2019年，第1–9页。[45] G. M´attyus、W. Luo 和 R. Urtasun，“DeepRoadMapper：从航空影像中提取道路拓扑”，载于《2017年IEEE国际计算机视觉会议论文集》，第3438–3446页。[46] N. Bolten 和 A. Caspi。（2022）OpenSidewalks模式。 [在线]。网址：https://github.com/OpenSidewalks/OpenSidewalks-Schema