

基于全景图像的社区体检评估技术应用研究

Application Research on the Assessment Technology of Community Physical Examination Based on Panoramic Images

张孝贤 | ZHANG Xiaoxian 周元婧 | ZHOU Yuanjing 李大勇 | LI Dayong

姜冬睿 | JIANG Dongrui 茅明睿 | MAO Mingrui

中图分类号: TU-984.11+4 文献标志码: A 文章编号: 1001-6740(2024)02-0005-08 DOI: 10.12285/jzs.20240402004

摘要: 2023 年住建部全国城市体检工作将城市体检单元细化到了住房、小区(社区)、街区、城区(城市)四个尺度,其中住房、小区(社区)、街区尺度的体检工作需要通过实地调研采集来开展,且体检要求更加精细化。为了提升调研采集效率和数据质量,减低体检评估成本,笔者所在的城市象限团队提出了基于全景图像的社区评估技术方案,通过智能感知背包和电动轮椅的一体化采集装备,完成体检调研过程中全景图像的自动采集;并结合图像算法、物联网数据处理和空间分析技术进行云端数据智能处理,初步实现了包括环境品质、路面平整、无障碍设施、公共空间品质等 50 多项建成环境指标的智能分析计算,探索了基于智能感知的全景图像进行体检指标计算的评估方法。

关键词: 城市体检、智能感知、计算机视觉

Abstract: The 2023 National Urban Physical Examination of the Ministry of Housing and Construction has refined the urban physical examination unit to four scales: housing, district (community), neighbourhood, and urban area (city), of which the physical examination of housing, district (community), and neighbourhood scales needs to be carried out through on-site research and collection, and the physical examination requirements are more refined. In order to improve the research collection efficiency and data quality, and reduce the cost of physical examination and assessment, the author's Urban Quadrant team proposed a community assessment technology programme based on panoramic images, through the integration of intelligent perception backpacks and electric wheelchairs collection equipment, to complete the automatic collection of panoramic images during the physical examination and research process; and combined with image algorithms, Internet of Things (IoT) data processing and spatial analysis technology for cloud-based data intelligent processing, initial The intelligent analysis and calculation of more than 50 built environment indicators, including environmental quality, road surface levelling, barrier-free facilities, public space quality, etc., have been achieved, and the assessment method of calculating medical check-up indicators based on the panoramic image of intelligent perception has been explored.

Keywords: City checkup, Intelligent perception, Computer vision

作者:

张孝贤, 北京城市象限科技有限公司算法工程师;

周元婧, 北京城市象限科技有限公司产品经理;

李大勇, 北京城市象限科技有限公司物联网工程师;

姜冬睿, 北京城市象限科技有限公司售前工程师;

茅明睿(通讯作者), 北京城市象限科技有限公司 CEO, 北京社区研究中心主任。

录用日期: 2024-03

一、研究背景

城市体检是实施城市更新行动、统筹城市规划建设管理、推动人居环境高质量发展的重要抓手。2023 年城市体检工作提出“坚持问题导向、目标导向、结果导向,主客观相结合”的整体

工作要求,将城市体检单元细化到了住房、小区(社区)、街区、城区(城市)四个维度,这一转变意味着社区等微观层面的体检内容占据了主导地位,对微观数据的采集需求和成果要求也随之大幅提升。

然而,由于受到技术发展和数据获取等限制,

传统城市体检评估多采用现场调研、问卷调查等高度依靠人力合作执行的方法。这种方法不仅需要大量的时间和人力成本投入,同时也限制了对微观层面的精细化分析。

随着大数据的推广深化和人工智能产业的蓬勃发展,一种基于“移动感知”的城市建成环境分析方法应运而生^[1]。该方法充分利用了自采集技术的优势,通过移动设备、传感器等工具,可以在较大范围内、高精度地获取城市数据,同时将成本控制在相对低水平。这种数据获取方式不仅能够满足城市体检评估的效率需求,而且还能够适应微观尺度分析的导向。更为重要的是,该方法采用了人为视角,从居民和用户的角度出发,收集建成环境的客观状态数据,从而更好地反映了城市居民的实际感受和需求。

二、研究综述

1. 传统城市体检评估技术

城市体检作为重要的城市治理工作,在评价城市发展健康状况和解决城市问题方面,已经形成了相对完整的技术框架^[2-4]。这一框架涵盖了城市自我体检、第三方体检以及社会满意度调查等多种评估方式,为城市管理提供了重要的参考依据^[5]。地方政府根据自身的实际情况和特点,总结形成了具有地方特色的体检评估工作路径。以上海市为例,该市通过构建完善的指标体系,并采取市、区联动的方式开展数据收集和评估工作,使得城市体检工作更加系统化和科学化。尽管城市体检工作取得了一定的成效,但仍然面临诸多挑战。由于高度依靠人力合作执行,存在工作效率低、调研周期长、成本高等问题,难以在财政资源有限的情况下降低调研成本,支持城市体检工作的常态化开展。此外,多人协作的复杂工作模式容易导致体检成果质量难以控制、工作流程不规范、全程监督留痕困难以及后台核验难等风险。

2. 智能城市体检评估技术

随着新技术的不断涌现,海量的城市数据被利用起来作为城市主客观评价的重要基石,自2018年城市体检评估工作首次在北京市实践以来,除政府部门业务数据和相关统计数据外,城市体检评估中已开始应用网络开源数据、用户生成内容(UGC)等大数据,如百度和高德开放平台数据、手机信令和交通IC卡数据、社交网络和点评数据、12345市民热线数据等,通过地理空间分析和数理统计分析等典型评估方法对城市发展状态及问题展开研究^[6]。

此外,包含了丰富城市建成环境信息的街景图片也在计算机技术的发展下被进一步挖掘,不仅打破了传统城市调研方法的空间和时间局限,还能够为后续城市更新工作提供标准化的调研数据资产,并丰富调研要素集,为城市问题研判和更新项目落位提供工作基础。

3. 计算机视觉技术在建成环境评估中的应用

在建成环境研究领域,常用的计算机视觉技术可以分为目标检测、图像分割与图像分类三大类别。其中,目标检测侧重于识别图像中的特定物体及其位置和大小,如Chen等人^[7]通过街景图片和LDCF算法评估行人检测的表现;Lumnitz等人^[8]利用Mask R-CNN算法在街道级图像中定位单独的树木实例;Garrido-Valenzuela等人^[9]则通过ODM模型识别人员和城市对象来研究城市密度。图像语义分割进一步深化了对图像内容的理解,将图像细分为多个具有特定类别的区域,例如Gong等人^[10]使用PSPNet微调百度地图街景图像以提高分类精度;Tian等人^[11]的Mask R-CNN算法实现了对多个类别和重叠物体的实例分割。图像分类任务则旨在根据图像特征将其分为预定义的类别,如Li等人^[12]使用SVM对不同土地利用类型的建筑物进行分类,Barleta等人^[13]通过ResNet34的迁移学习对图像进行单一分类, Yang等人^[14]则采用

VGG-16图像分类方法进行深度学习分析。

尽管,计算机视觉技术为建成环境评估带来了前所未有的便捷与效率,但在社区体检评估场景中仍没有一套系统性技术框架作为体检指标计算的基础。并且,如何具体地针对社区街景的数据特点对算法模型进行优化,也是目前计算机视觉技术在建成环境评估中的关键挑战。除了算法模型之外,已有研究尚未对体检指标体系中指标与算法处理结果之间的关联关系进行讨论,大多围绕孤立的模型对单一类别问题自动化处理而展开,并未形成完整的数据采集、图像处理、指标计算全流程体检评估技术方案。

4. 移动感知系统在建成环境数据集中的应用

面向城市环境快速变化而出现的數據高频更新问题,研究人员引入实时数据处理技术(如Apache Storm)来提供更高频次的数据更新,以及更精细的空间分辨率,以确保数据的时效性和全面性。在环境检测领域,传感技术的迭代也使得收集到的数据类型变得更加丰富多样。例如声学传感器用于监测噪声水平,水质监测设备用于评估水体质量以及街景采集车辆则负责收集城市街景信息都为城市更新过程中的数据收集提供了新的视角和思路^[15-18]。

然而,移动感知系统在实际应用中尚未充分融入到城市规划和研究人员的日常工作流程中^[19],早期工作提出的移动感知系统框架应支持多种设备和平台,以实现更全面的环境监测,但仍未被实践应用在实际的体检评估工作当中。并且,使用车辆采集的街景图像无法深入社区内部,难以获取到更小尺度的建成环境信息,这样导致移动感知系统在实际应用中的潜力尚未得到充分发挥。

三、技术方法框架

在制定技术框架之前,团队进行了一项前期的可行性评估,以确定体检指标是

否可以通过计算机视觉算法进行识别。这一评估是基于《住房和城乡建设部关于全面开展城市体检工作的指导意见》《完整居住社区建设标准》《北京市无障碍城市设计导则》以及《北京市老旧小区综合整治标准与技术导则》等官方文件进行的。通过对相关文档的详细梳理，团队确定了116项与社区体检相关的指标^{[20][21]}，这些指标涵盖了住宅、小区整体、公共活动空间、公共服务设施和道路等五大类别。在这116项指标中，有88项（占总指标的75.86%）具备通过计算机视觉算法进行识别的可能。

1. 总体技术框架

综合上述社区体检指标的梳理工作，论文尝试使用移动智能感知设备深入采集小区内部的全景图像，并使用计算机视觉算法完成图像数据处理，最后对视觉相关的评估指标进行计算，从而提出一种基于图像算法的社区评估技术方案。论文将指标体系中与视觉相关的指标，按照图像算法的任务目标进行分类，分为目标检测、语义分割以及图像分类（图像评分），部分指标分类结果如表1所示：

本文提出的评估技术主要包括三个阶段：建成环境数据采集端、全景图像处理板块以及体检指标计算。其中，计算机

视觉技术是智能评估基础，集成了多种不同业务功能的图像算法，对传感设备获取到的社区内部全景图像进行处理，作为后续指标的计算基础。总体技术框架如图1所示。

2. 建成环境数据智能采集软硬件技术

尽管百度等LBS服务商提供的街景服务覆盖了主要道路，但受限于车载采集设备，难以获取社区级道路的详细数据，无法精细化完成社区体检评估。此外，由于成本问题，街景数据更新周期长达2~3年，甚至不更新，无法反映最新情况。团

队使用国产微型嵌入式芯片，研发适用于可穿戴、自行车、电动轮椅的便携式社区环境信息采集设备。外出的作业人员可以将设备安装在自行车或电动轮椅上，采用自行车或轮椅出行调研，也可穿戴调研背包，采用步行方式进行调研；实现一次外业，获取目标区域的环境信息、路面颠簸信息和全景照片等多种数据的目的，大幅提升调研的工作效率。

3. 基于计算机视觉的体检分析评估技术

利用计算机视觉算法对采集到的全景图像数据进行处理，是体检评估指标计算

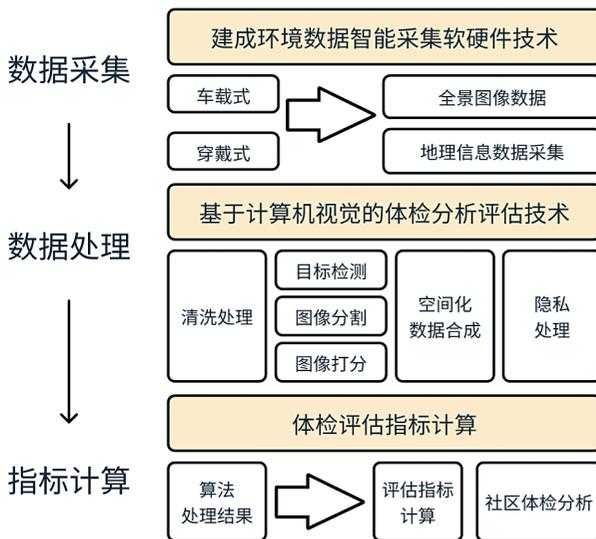


图1：技术方案总览

指标与图像算法任务分类关联关系（部分）

表1

体检指标	算法基础	体检指标	算法基础	体检指标	算法基础
无序占道经营商贩分布分析	目标检测	垃圾收集点分布分析	目标检测	建筑立面视域占比分析	语义分割
人群活力分布	目标检测	未进行垃圾分类的投放点分布分析	目标检测	施工围挡分布分析	语义分割
沿街车辆分布分析	目标检测	休息座椅分布分析	目标检测	路边绿化带分布分析	语义分割
停车位占用分析	目标检测	路面积水点分布分析	目标检测	宽度不达标的人行道路段分布分析	语义分割
停车位缺失分析	目标检测	照明设施分布分析	目标检测	道路侵占路段分布分析	语义分割
机动车充电桩分布分析	目标检测	无障碍停车位分布分析	目标检测	道路铺装不达标分析	语义分割
建筑外立面破损分析	目标检测	消防设施分布分析	目标检测	绿视率分析	语义分割
建筑出入口无障碍配套设施分析	目标检测	路面垃圾堆放点分布分析	目标检测	建筑风貌美观性评价	图像分类
空置及出售商铺分布分析	目标检测	破损围墙分布分析	目标检测	道路空间宽敞程度分析	图像分类
建筑出入口照明设施分析	目标检测	花坛分布分析	目标检测	街道沿街通透性评价	图像分类
建筑出入口安防监测分析	目标检测	安防设施分布分析	目标检测	街道环境卫生性评价	图像分类
人行道入口围堵路段分布	目标检测	人行道缘石坡道缺失分布分析	目标检测	道路开敞程度分析	图像分类
非机动车充电桩分布分析	目标检测	体育器材分布分析	目标检测	……	

的基础。论文按照图像算法的具体任务分类方法，根据指标体系中不同指标所依赖的算法基础，将指标的处理方式分为目标检测、图像分割以及图像评分三个类别。目标检测算法适用于处理无需按照像素大小进行识别的要素，如停车位占用、无障碍设施和安防设施等指标，算法能够检测全景图像中指定目标是否存在或其数量。而对于需要按像素颗粒度进行处理的指标，如建筑立面视域占比、绿化带占比以及道路铺装等，语义分割算法能够更好地应对这类需求。此外，部分评价性指标并不依赖图像中的要素识别，如建筑风貌美观性评价、道路空间宽敞程度分析以及街道沿街通透性评价等，使用图像分类算法完成图像评价价值的计算是被验证可行的方法。最后，团队结合多源异构数据处理技术，整合街景数据和传感器地理位置信息，解决城市体检中数据合成问题。这一技术能高效处理不同数据源信息，实现了社区元素的空间位置映射和建成环境的全面监测。街景数据提供了详细的城市和街区视觉信息，传感器数据提供了实时的环境参数和城市运行状态。通过综合利用这两类数据，能够全面评估城市状况，为城市管理提供有力支持。

4. 体检评估指标计算

论文在前人应用计算机视觉技术在建成环境评估的基础上，梳理了与视觉相关的体检指标，根据图像算法的识别结果结合其他传感器获取到的环境信息对评估指标进行计算，例如结合全景图像中识别到的行人数量以及采集点位采集到的 Wifi 信号强度，综合计算“人群

活力分布”这一指标。通过对指标内涵的拆解，可以智能化指标的计算过程，部分指标体系中的指标所关联的图像算法要素如表 2 所示。

四、核心技术

1. 移动智能采集技术

面向深入社区内部的环境信息采集问题，团队设计了一套小型智能化集设备。采用一种小体积“快装”结构接口，能够方便快捷的通过铆钉和螺丝固定在书包外侧或自行车、电动轮椅的金属框架上，使得传感器以及其他控制设备能够被挂载在书包或交通工具的金属支架上，实现快速安装携带设备的目的。

感知设备的智能化采集和自动化定位技术方案中，团队基于国产的 Risc-V 架构芯片，研发了一套控制终端设备，集成了低功耗蓝牙通讯模块、北斗定位模块和 CAT1 通讯模块。通过传感器设备间的蓝牙通信模块，各模块之间形成小型局域网络，由一个终端统一控制接入网络的全部设备，实现设备的集成。这种将不同通讯端口转换为基于蓝牙通信协议的方式，可以实现各种通讯端口和协议不同的传感器能够被终端统一控制，使得传感器设备开机后，能够自动通过蓝牙连接完成全景图像数据的统一采集。作业人员携带设备到达作业区域打开终端和传感器后，终端控制设备将自动控制各个传感器采集数据，通过北斗定位系统记录作业的位置，然后通过 CAT1 网络，将包含了拍摄位置信息的数据回传到云服务器，完成采集数据的收集。

2. 少样本图像目标检测技术

1) 基于 YOLOV8 算法的半监督目标检测类指标计算

基于全景图像的社区智能化体检评估依赖图像算法的处理结果，诸如垃圾收集点分布、公共活动场所分布分析、无障碍设施建筑配套分析与停车位占用分析等指标，要求图像算法能够检测出特征复杂的相关目标作为指标统计计算的基础。论文采用 YOLO V8 模型作为核心检测算法^[22]，结合 Efficientteacher 半监督框架^[23]对模型进行优化，旨在使用少量标注数据结合大量未标注数据作为训练集得到更高精度的检测模型。

2) 基于 zero-shot 模型的目标检测类指标计算

除了语义较为复杂的检测目标，可以通过标注数据作为监督模型训练的数据集以外，指标体系中还存在花坛分布、休息座椅分布等较为基础的要素识别内容。该类目标的图像特征较为简单，但是类别数目众多，且不同城市体检体系当中的检测类型也有差异，这使得监督模型并不能很好应对这些问题。

面向这类目标的需求，论文采用了基于开放数据集训练的 Grounding Dino 模型^[24]进行目标检测。该模型作为一种 zero-shot 模型，能够在无监督训练的前提下理解简单文本语义，根据文本提示对图像中的指定目标进行检测。引入这种同时考虑文本特征与图像特征的综合性处理方式，使得整体模型具备卓越的可扩展性，能够在无样本训练的情况下拓展更多目标类别。

3. 适应复杂场景的全局特征语义分割技术

目标检测算法能够一定程度上理解图像语义，但在处理不规则目标（例如：人行道、绿化带、建筑立面等）时，这些目标难以用边界框精确表示。图像分割技术对于这些目标能够进行很好的处理，通过

基于视觉的体检指标计算基础（部分）

表 2

体检指标	算法识别要素	其他环境信息
人群活力分布	单元内行人数量	WIFI
停车位占用分析	单元内空停车位、被占用停车位、车辆综合统计	—
道路铺装不达标分析	识别道路为人行道、识别铺装像素不间断	—
路面垃圾堆放分析	单元内路面垃圾、垃圾桶数量	异味指数
……		

对图像中每个区域的像素进行分类，能够计算指定类别在图像中的占比，完成体检评估当中路边绿化带、道路开敞度和建筑立面围合度等指标的计算。

论文采用基于 Transformer 主干网络的 Mask2former 分割模型^[25]，相比于常用于街景语义分割的 Deeplab V3+ 等基于 CNN 构建的模型，该模型在 COCO、ADE20K 等复杂场景数据集上取得不错的成绩，其在多个图像分割任务上的优越性已被验证。该模型能够处理社区全景图像的复杂场景特征，对于人行道与车道、树木与绿化带等特征类似的要素具备较高的处理性能，处理效果如图 2 所示。

4. 基于图像分类算法的评分技术

计算机视觉算法中除了目标检测和语义分割算法被应用在客观指标的计算，图像分类算法也被证实^{[26][27]}能够有效预测街景当中的主观感知指标计算。通过深度卷积神经网络 (CNN) 等模型提取图像的高级特征，并将其映射到预定义的类别标签上，图像算法能够模拟成为街景主观评分的有力工具。

在城市街景的主观评价方面，ConvNeXt 性能优于其他分类模型，而 ConvNeXt V2 模型又在 ConvNeXt 的基础上进行了优化，其在各种视觉识别任务上都取得了不错的性能提升^[28]，因此本文选择 ConvNeXt V2 与随机森林 (RF) 模型作为主要的预测

器，ConvNeXt V2 模型作为图像分类器，能够有效地提取图像特征，而 RF 模型则被应用于回归分析，用于预测图像的主观评分。

团队在 Place Pulse 2.0 数据集上预训练模型，并使用自有数据集对模型进行微调，使得模型能够更好地适应社区体检指标的评分需求。自有数据集将主观评价类别分为：建筑风貌美观性评价、道路空间开敞程度分析、街道沿街通透性评价、街道环境卫生性评价与道路开敞程度分析。评分模型通过将街景的语义特征向量与类概率向量相合并，再使用回归模型实现全景图像的评分任务，其架构如图 3 所示。

5. 物联感知数据空间化和合成技术

由于街景图像的空间坐标是传感设备的地理坐标，而非被拍摄物体的实际坐标，为了将识别到的建筑外立面破损以及建筑出入口无障碍设施等要素与实际建筑进行关联对应，还需要将传感设备的空间坐标转换为拍摄到的建筑物实际空间坐标，以达到更为准确的指标计算结果。

然而，即使基于深度学习的单目景深技术已经较为成熟^[29]，但仅通过单目相机将二维图像中存在的目标映射到三维空间中，获取其真实的绝对物理距离仍存在较大困难。因此，本文考虑结合物联感知的图像定位坐标与先验建筑底图信息计算经纬度之间的距离、方位角等数据，设计四

象限规则完成对于图像中建筑物的空间化与算法处理结果进行合成。

首先，依据前后拍摄点的方向建立局部坐标系，确定建筑所归属的象限，然后计算建筑底图与拍摄点之间的方位角，确定一个距离拍摄点方位角最小的拍摄方向，找到该方向中位置距离最近的建筑物进行关联，规则示意如图 4。

6. 公共空间数据采集隐私保护技术

考虑到公共空间数据可能导致的隐私数据安全问题，团队采用了 YOLO FACE V2 算法对人脸、车牌号码等敏感信息，进行识别并进行模糊处理。该算法在小目标检测任务上，尤其是在面部和车牌号等要素识别，具有显著优势^[30]。在复杂场景（如遮挡、光线变化等）下的识别准确性，而且通过自动对敏感区域进行模糊处理，有效降低了误识别和信息泄露的风险。

五、技术方法实验结果

1. 少样本目标检测模型结果分析

使用 YOLO V8 检测模型结合 Efficient-Net 框架在自有数据集上进行训练，训练集一共包含 28134 张图像数据，其中标注数据为 2366 张，增强过后为 7098 张图像数据，无标注数据占总训练集的 75% 为 21036 张。模型经过 150 轮次 (epochs)



图 2: Mask2former 分割模型在社区全景图像中的处理效果

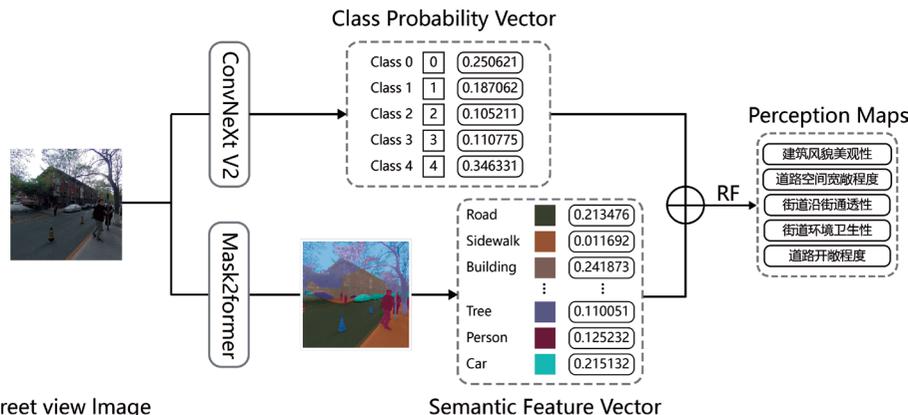


图 3: 图像打分模型结构示意图

训练后，在 720 张验证集的性能测试如表 3 所示。

在少样本数据集检测任务中，Efficient-teacher 半监督框架能够有效提升模型的检测结果，模型检测结果如图 5 所示。

Grounding Dino 模型能够在无样本进行训练的情况下，完成 zero-shot 目标检测任务，本文利用模型对语义简单的目标要素进行识别，例如路椅、路灯以及摊位等目标，结果显示无监督的算法模型对于统计图像中指定要素任务完成度较差，容易忽视图像边缘的小目标导致漏检的情况，但是对于判别这些要素是否存在图像中效果较好。这样使得指标体系当中，无需统计图像中要素数量的体检指标，能够以较低的成本完成计算。Dino 模型在街景中的 zero-shot 检测效果如图 6 所示。

论文使用感知装备采集到的某小区 230 张全景图像，人工对指定要素进行判别，对比零样本检测模型的处理结果，以测试模型是否能够对图像中存在的简单语义信息的目标进行正确判别，测试效果如表 4 所示。

结果表明，zero-shot 模型虽然难以替代监督模型完成目标检测任务，但可以作为判别模型对图像中是否存在指定要素进行较为准确的判别。

2. 复杂场景的图像分割与评分模型结果分析

论文结合基于 transformer 的语义分割模型以及 ConvNeXt 结构，能够更好地处理复杂场景的图像特征，并使用随机森林模型进行回归完成评分任务。在自行打分的 500 张测试数据集上进行推理，对比仅使用 ConvNeXt 模型进行特征提取结合随机森林进行回归的计算结果，以及融合 Mask2Former 计算的语义特征向量的随机森林模型回归结果，结果如表 5 所示。

通过上表可以看出，融合了语义特征向量的评分模型精确度得到提升，尤其是“街道环境卫生性评价”这一指标，在模型理解了图像中的要素含义后，能够更好地

模型训练结果

表 3

模型	precision (B)	recall (B)	mAP50 (B)
YOLO V8	0.71	0.61	46.64
Efficientteacher + YOLO V8	0.81	0.68	55.96

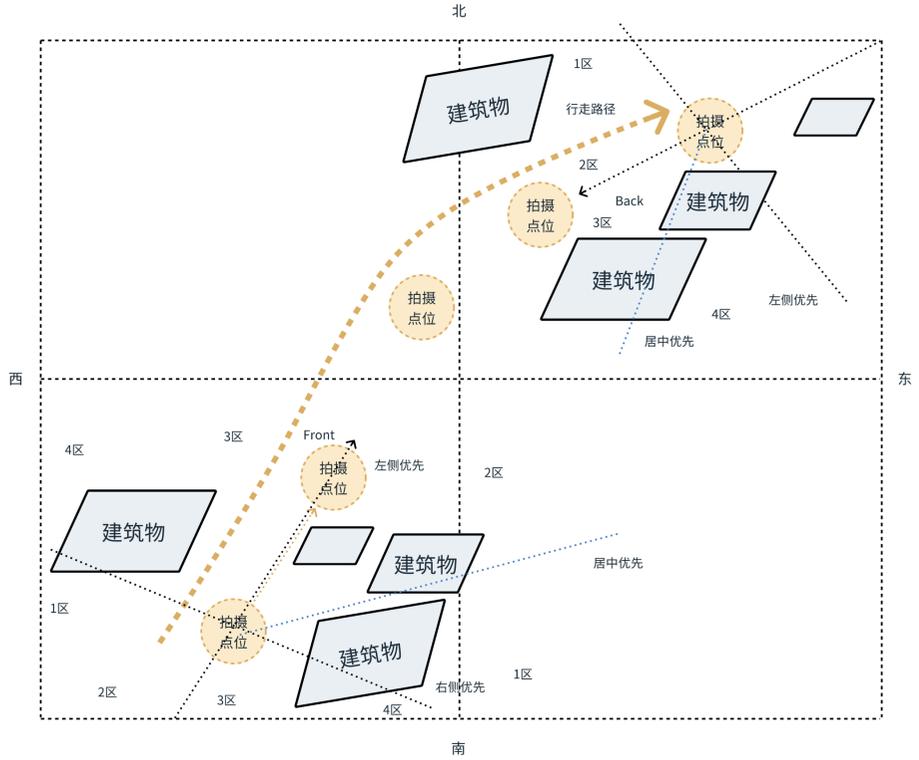


图 4：图像要素关联建筑物底图规则示意图



图 5：半监督 YOLO V8 模型检测结果示意图



图 6. zero-shot 目标检测模型检测效果

grounding dino 模型测试结果

表 4

	precision (B)	recall (B)
花坛	0.75	0.60
长椅	0.73	0.56
摊车	0.69	0.58
安防设施	0.72	0.58
路灯	0.77	0.62
消防设施	0.78	0.56

图像评分模型测试结果

表 5

评分类别	ConvNeXt 模型精确度	结合语义分割的 ConvNeXt 模型的精确度
建筑风貌美观性评价	0.79	0.88
道路空间宽敞程度分析	0.60	0.81
街道沿街通透性评价	0.57	0.68
街道环境卫生性评价	0.48	0.72
道路开敞程度分析	0.59	0.77

地对该指标进行评分。

综上，不同模型在不同的图像处理任务中均能够在少样本的情况下达到不错的效果，随着更多样本的训练集投入到模型训练当中，可以更好地提升模型性能。

六、总结展望

1. 技术贡献

基于住建部城市体检工作对于细化颗粒度的新要求，聚焦社区和住房层面的评估，探讨从移动智能设备进行数据采集，

再到图像算法对图像数据进行处理，最终完成指标计算的智能化评估技术。对比人工调研工作结果，证实了该技术方法的有效性和可靠性，并为后续大规模工程集成应用提供理论基础。具体而言，本文的主要贡献为：

- (1) 本文提出一种集成多传感器的环境信息感知设备方案，并设计为便携式背包或轮椅模式。这使得设备不仅能灵活深入社区内部进行数据精细化采集，而且利用北斗定位系统确保了数据点的精确定位。
- (2) 通过全景图像数据，集成目标检测、语义分割和图像评分三类算法模型，

实现社区体检指标的计算。通过拆解指标内涵，对社区评估场景中各图像算法的选取进行适用性分析，提供建成环境评估的图像模型优化思路。

(3) 讨论了图像算法处理结果与指标体系之间的关联方式。相比以往仅依赖图像处理结果对建成环境进行评估，本文通过整合多源数据并应用图像算法进行综合分析，实现了对社区体检评估指标的精确计算，突破了单一街景数据应用的局限。

2. 研究展望

基于论文的研究技术，在算法模型精度进一步提升之后可以构建城市体检云平台，以云端服务器的形式处理感知设备采集到的数据，并以多元化的形式进行评估结果反馈，从而降低城市体检调查成本，实现体检全过程留痕，提升规范化程度，优化工作流程，并逐步实现城市体检过程的自动化采集和分析。

参考文献

- [1] 王新宇, 李彦, 李伟健, 等. 城市更新视角下的公共空间品质评估方法——基于移动感知技术的探索[J/OL]. 国际城市规划, 1-15.
- [2] 王辉. 谁的城市更新[J]. 建筑师, 2023 (1) : 29-36.
- [3] 张斌. 从城市研究到驿站实践: 关于日常性的思考[J]. 建筑师, 2023 (1) : 45-58.
- [4] 米凯拉·巴罗西奥, 吴红叶. 城市更新: 形态/衍进中的后工业城市[J]. 建筑师, 2023 (1) : 13-19.
- [5] 董亦楠, 沈咏. 为人而更新的城市[J]. 建筑师, 2023 (1) : 4-5.
- [6] 崔博庶, 高硕, 张云金, 等. 基于大数据技术的北京城市体检街区治理评价[J]. 北京规划建设, 2020 (S1) : 130-135.
- [7] CHEN L, LU Y, SHENG Q, etc. Estimating pedestrian volume using Street View images: A large-scale validation test[J/OL]. Computers, Environment and Urban Systems, 2020, 81: 101481.
- [8] LUMNITZ S, DEVISSCHER T, MAYAUD J R, etc. Mapping trees along urban street networks with deep learning and street-level imagery[J/OL]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 175: 144-157.
- [9] GARRIDO-VALENZUELA F, CATS O, VAN

- CRANENBURGH S. Where are the people? Counting people in millions of street-level images to explore associations between people's urban density and urban characteristics[J/OL]. *Computers, Environment and Urban Systems*, 2023, 102: 101971.
- [10] GONG Z, MA Q, KAN C, etc. Classifying Street Spaces with Street View Images for a Spatial Indicator of Urban Functions[J]. 2019.
- [11] TIAN H, HAN Z, XU W. Evolution of Historical Urban Landscape with Computer Vision and Machine Learning: A Case Study of Berlin[M/OL]. DE: Wichmann Verlag, 2021[2024-03-02].
- [12] Building block level urban land-use information retrieval based on Google Street View images[J/OL]. *GIScience & remote sensing*, 2017, 54 (6) : 819-835.
- [13] BARLETA L, CARRILLO M, FRANK Z, etc. Ejidos, Urbanization, and the Production of Inequality in Formerly Agricultural Lands, Guadalajara, Mexico, 1975–2020[J/OL]. *Land*, 2020, 9 (12) : 526.
- [14] YANG C, LIU T, ZHANG S. Using Flickr Data to Understand Image of Urban Public Spaces with a Deep Learning Model: A Case Study of the Haihe River in Tianjin[J/OL]. *ISPRS International Journal of Geo-Information*, 2022, 11 (10) : 497.
- [15] AUMOND P, LAVANDIER C, RIBEIRO C, etc. A study of the accuracy of mobile technology for measuring urban noise pollution in large scale participatory sensing campaigns[J/OL]. *Applied Acoustics*, 2017, 117: 219-226.
- [16] GUAN Y, JOHNSON M C, KATZFUSS M, etc. Fine-Scale Spatiotemporal Air Pollution Analysis Using Mobile Monitors on Google Street View Vehicles[J/OL]. *Journal of the American Statistical Association*, 2020, 115 (531) : 1111-1124.
- [17] APTE J S, MESSIER K P, GANI S, etc. High-Resolution Air Pollution Mapping with Google Street View Cars: Exploiting Big Data[J/OL]. *Environmental Science & Technology*, 2017, 51 (12) : 6999-7008.
- [18] MEYERS D, ZHENG Q, DUARTE F, etc. Initial Deployment of a Mobile Sensing System for Water Quality in Urban Canals[J/OL]. *Water*, 2022, 14 (18) : 2834.
- [19] LAPORT-LÓPEZ F, SERRANO E, BAJO J, etc. A review of mobile sensing systems, applications, and opportunities[J/OL]. *Knowledge and Information Systems*, 2020, 62 (1) : 145-174.
- [20] 曾忠忠, 张锦, 张波, 等. 乡村建设评估指标体系研究[J]. *建筑师*, 2019 (5) : 78-84.
- [21] 汪芳. 自然同行. 人文相伴与数字赋能:《城市设计(第4版)》的新时代城市设计范型整合[J]. *建筑师*, 2023 (5) : 126-129.
- [22] TERVEN J, CORDOVA-ESPARZA D. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS[J/OL]. *Machine Learning and Knowledge Extraction*, 2023, 5 (4) : 1680-1716.
- [23] XU B, CHEN M, GUAN W, HU L. Efficient Teacher: Semi-Supervised Object Detection for YOLOv5[J/OL]. arXiv, 2023[2024-03-02].
- [24] LIU S, ZENG Z, REN T, etc. Grounding DINO: Marrying DINO with Grounded Pre-Training for Open-Set Object Detection[M/OL]. arXiv, 2023[2024-03-02].
- [25] CHENG B, MISRA I, SCHWING A G, KIRILLOV A, GIRDHAR R. Masked-attention Mask Transformer for Universal Image Segmentation[J/OL]. arXiv, 2021[2024-03-02].
- [26] LIU L, SILVA E A, WU C, etc. A machine learning-based method for the large-scale evaluation of the qualities of the urban environment[J/OL]. *Computers, Environment and Urban Systems*, 2017, 65: 113-125.
- [27] YAO Y, LIANG Z, YUAN Z, etc. A human-machine adversarial scoring framework for urban perception assessment using street-view images[J/OL]. *International Journal of Geographical Information Science*, 2019, 33 (12) : 2363-2384.
- [28] WOO S, DEBNATH S, HU R, etc. ConvNeXt V2: Co-designing and Scaling ConvNets with Masked Autoencoders[C/OL]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) . 2023: 16133-16142[2024-03-02].
- [29] YANG L, KANG B, HUANG Z, etc. Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data[M/OL]. arXiv, 2024[2024-03-02].
- [30] YU Z, HUANG H, CHEN W, SU Y, LIU Y, WANG X. YOLO-FaceV2: A Scale and Occlusion Aware Face Detector[J/OL]. arXiv, 2022[2024-03-02].

图片来源

图 1~ 图 6: 自绘