

基于多图像增强语义分割的传统地块形态识别研究

武文忻¹, 刘华², 李力¹

摘要: 在历史街区肌理修复项目中, 常涉及对于传统建筑形制的考证与恢复。研究提出了一种基于深度学习技术的传统地块形态识别方法, 用于解决由于相关历史资料匮乏导致形态识别困难的问题。通过结合语义分割算法和多模态识别思想, 该方法能有效提高地块特征的识别能力。研究以扬州地籍图为数据集, 进行了模型对比实验, 结果显示引入 Squeeze Excitation (SE) 注意力机制的模型在地块分割任务上表现更好。研究揭示了传统地块形态的拓扑规律具有规则性和普适性特征, 为城市更新和风貌保护提供了有益信息。然而, 研究存在局限性, 尚未在其他城市地籍图上验证模型的泛化性能, 未来需要进一步考虑不同地域的差异, 提高模型的泛化能力。

关键词: 深度学习; 传统地块; 形态研究; 语义分割; 城市更新

Keywords: deep learning; title census; morphological study; semantic segmentation; urban renewal

资助项目情况: 无

引言

地块作为传统城镇用地的基本单元, 其形态控制着传统建筑肌理的特征, 同时其与内部的建筑平面组织也是城市形态学研究的重点内容。而在现代化进程中, 其肌理却遭到了一定的破坏。一方面, 城市规划和管理所依据的规划体系和文物保护体系仍然以物质性的建筑及其形式和风格为主要对象, 忽略了城市肌理的整体性和动态性; 另一方面, 在城市更新和改造中, 决策者对土地单元的合并与修改往往忽略了构成肌理的街坊建筑的地界, 即传统地块形态, 从而导致城市肌理的空间结构和尺度失去原有的秩序和特征^[1]。

面对这样的问题, 东南大学韩冬青教授在《关于城市地块格局的机理认知与设计实践》中^[2], 提出“虚拟地块”的城市设计策略, 以历史地段形态控制为依托, 对缺失肌理的修复进行约束与指导。而传统地块的形态正是这种依托的重要组成部分之一, 这种形态受地界产权等因素的影响, 它们共同构成了城市空间形态和肌理的重要秩序性介质。但是, 由于历史变迁, 很多传统地块已经合并或拆分, 失去了原有的形态, 而大部分古城史料又往往缺少地块的相关记载, 在现实实践中, 需要对原有的形态进行调研与推测, 耗时耗力。如何快速准确判断出传统地块的形态边界, 是当下城市更新改造实践环节中的瓶颈。

针对相关资料匮乏的情况, 李冰、苗力等总结归纳了判定传统地块边界的方法^[3], 其中关于“平面推测法”的研究, 指出设计者可以从卫星图的肌理与 CAD 图纸中暗含的轴线关系、宅前路径等进行对比推测, 并通过相关实例验证了该方法的有效性, 这暗示了从图纸结构信息识别地块

武文忻, 东南大学建筑学院, 硕士研究生

刘华, 东南大学建筑设计研究院有限公司, 高级城乡规划师

李力 (✉), 东南大学建筑学院, 副教授, 101012053@seu.edu.cn

边界的可行性。但这一方法仍然依靠人工的对比与筛选，在实践中有着工作量大、主观性强、准确性低的问题。

而面对上述问题，国内外研究学者运用深度学习的技术手段，结合城市形态与肌理提出了解决方案和创新思路，如：姚佳伟、黄辰宇等总结了基于人工智能的城市肌理识别与评价方法，包括 k-means 聚类、卷积神经网络分类，并运用上述方法对城市的肌理按照密度高低等规则进行分类，提出相关评估^[4]。Lamia El Mendili, Anne Puissant 等则在深度学习框架中，使用 Sentinel-2 数据方法，对城市肌理分类效果进行评估，并提出了一个多时相 FCN 识别网络，该网络在城市肌理的分类上达到了较好的效果^[5]。然而相关研究也存在问题，即对地块形态分类技术不够精细，常被使用的 k-means 等聚类算法，依赖纹理、色彩等低层次特征，忽视了语义和风格。这导致分类结果模糊，受聚类数量和初始中心的影响大，难以精确识别肌理。此外，常见的语义分割技术多依赖单一数据源，当数据源为 CAD 矢量图资料时会缺乏建筑肌理细节，而采用卫星航拍图又无法保证图像的精度，准确性较低。

因而，本研究通过深度学习技术，利用语义分割算法结合多模态识别的思想，以扬州部分地区地籍图为数据集，探讨基于机器学习快速识别传统地块形态规则的可能性，并希望通过机器学习，反向验证图像语义信息下的形态学识别要素。

1 传统地块形态

传统地块形态在中国历史城镇中占有重要地位，通常与一进或多进的传统院落相对应^[6]。在中国传统城市中，地块的概念也与地籍、地界等术语紧密相关，历朝历代都对土地制度和土地管理有着相当的重视^[7]。地界的划分也直接影响到土地的使用、赋税的征收以及城市的布局。地界作为土地所有权的体现，是城市规划和管理的基礎，同时也是城市空间布局和肌理形态的决定性因素之一。因而，地界的布局与特征不仅影响着传统院落式建筑的形成与发展，并且与之共同构建了传统城市布局的内在逻辑与秩序。

而在本文研究中所探讨的传统产权边界的确定实质上是对传统建筑空间单元组织关系边界的界定，如院落关系、建筑组团关系等，而非某一特定期期的确切地界归属。这主要是由于以下几点原因：

1、地籍产权具有历史动态性：受历史经济等因素影响，地块的地籍或产权本身是随着时间不断变化的。然而，由于相关资料的匮乏和缺失，研究者难以找到确切的史料证据，因此需要通过推断的方式进行研究，并更加注重空间的组合关系来辅助实际设计。

2、传统地块具有明显的形态学特征：受中国古代土地私有制度影响，同一地界中房间所围合而成的建筑通常呈现传统院落式布局。因此，留存下的原始院落肌理能在一定程度上反映了原有的传统地块形态。即使院落形状在大小上可能有所不同，但其拓扑关系和基本规则通常不会发生很大变化。

3、现代地块具有推出传统地块的可能性：随着土地改革的成功，土地私有制被废除，一些院落产权转变为区属直管公房^[8]。在房屋重新分配过程中，原有的产权被细分为若干小产权，但原有地块特征仍然保留。因此，通过对现有地块的合并，可以推测传统地块的形态。

综上所述，本文所讨论的“传统地块形态”并非严格意义上受经济、法律、政府规划约束产生的真实历史地籍地块形态，而是偏向于形态学研究，更加注重传统肌理空间的组织关系。

2 识别要素提取

2.1 传统地块形态特征

对于传统地块特征的识别有助于构建数据集，辅助模型学习。尽管不同地区的传统城市结构具有各自独特的风格和特点，但受到普世规则的影响，其地块的形态却展现出相似的特征。其形态特征可归纳为以下几点^[9-11]：

(1) 规整排布：传统地块通常呈现规整的排布形式，如矩形、近似方形或狭长形等。这种规整排布既有助于提高土地利用效率，又方便统一管理和规划。这一特征的来源可追溯至中国古代小农经济的规则映射，反映了历史上土地利用和分配的一致性。

(2) 院落结构性：院落结构也是传统地块的典型特征，即建筑环绕中央院落布局，形成相对封闭的空间。这种布局不仅与古代建筑的典型样式相关，而且反映了人们对私密性与安全性的普遍追求。

(3) 临街门头性：传统地块的另一显著特征在于其与街道之间密切的联系，即临街门头。这些地块往往以建筑沿地块纵深方向布置为主，使得地块呈现出垂直于街道的狭长布局形式。

2.2 传统地块形态平面识别要素提炼

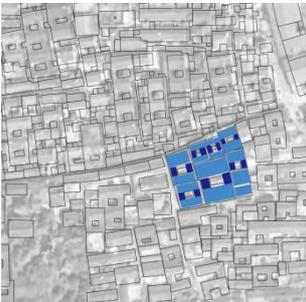
在传统院落的空间组织中，轴线关系、街道与入口布局与建筑肌理的特征，构成了地块形态识别的三大核心要素。这些要素在图像中的体现，对于地块形态的准确识别具有决定性影响。

(1) 轴线关系：院落的空间对称性和秩序性往往围绕中心轴线展开，其中正房通常位居中心轴线之上，而厢房则沿轴线两侧布置。轴线关系在图像中的清晰度，直接影响到对院落布局对称性和秩序性的判读。

(2) 街道与入口：街道作为地块间连接的纽带，其与入口的相对位置和设计，不仅反映了地块与城市空间的互动关系，也为地块形态的界定提供了重要参照。图像中街道与入口的识别，对于地块范围的确定至关重要。

(3) 建筑肌理：建筑的外观特征，包括纹理、结构、材料等，是地块形态识别中不可或缺的微观信息。通过图像分析，可以揭示建筑的平面布局、高度、层数等细节，以及屋顶材质、形式和周围环境的关系，从而为地块形态的全面理解提供支持。

表1 传统地块形态平面识别要素（图片来源：作者自绘）

轴线关系/街道	建筑形态轮廓	建筑肌理
		

综上，在选取数据集时，着重关注了图像对上述特征的反映。同时，采用多数据源进行训练，结合 CAD 图像精准、能更好反映拓扑关系与卫星图像涵盖建筑肌理等优势，构建相关数据集与识别网络。

3 识别网络与数据集构建

本文采用了 U-net 语义分割网络，结合多模态的思想，利用 Squeeze Excitation 注意力机制（下文简称 SE）来提高解码器输出的准确性和判别性。多模态的思想本质上，是从多个模态表达或感知事物，如来自不同传感器拍摄的同一事物的图片、同一事物的图片表达和文字表达等；基于多模态的语义分割则多应用与遥感、医学影像等领域^[12]；研究中，我们将卫星图和 CAD 图视为同一地块的不同表达形式，分别捕捉其拓扑和肌理信息，并将这一方法称为“多图像增强”。研究利用 VGG16 作为编码器来提取 CAD 图的特征，同时 SE 模块用于提炼卫星图的特征，两者在 U-net 框架下融合，增强图像识别分割的效果。SE 机制在解码阶段强化了卫星图像特征的识别，尤其是在区分传统与非传统地块以及产权边界时，可以提识别的准确性。

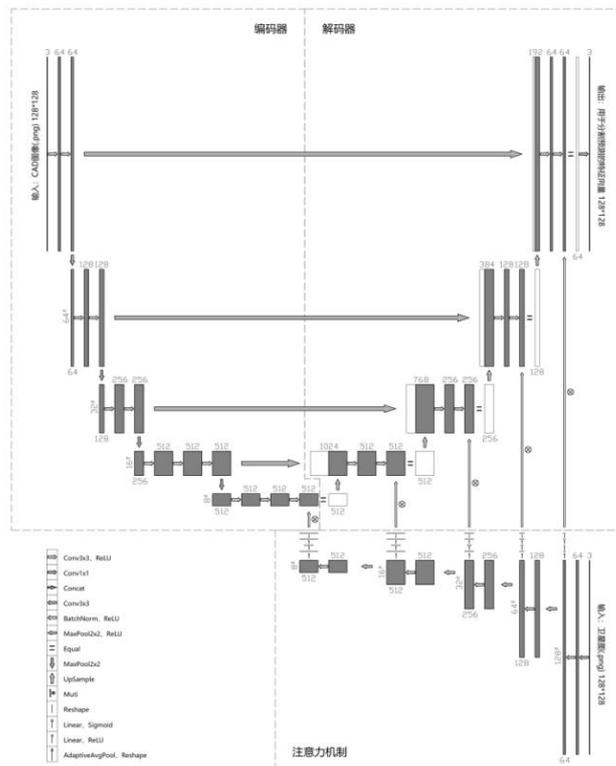


图 1 研究采用的网络结构（图片来源：作者自绘）

而本研究的数据集与标注则来源于设计院提供的根据历史地籍图（产权地块形态）、实地调研风貌所最终推测的传统地块形态而绘制的 CAD。并基于这些数据，进行了清洗工作，主要是剔除地籍图中非传统建筑肌理的部分，如学校、大型商场等，并将一些与卫星图所示位置偏差较大的地方进行重绘，以此使得标注数据更加精准。在此基础上，检查所给产权图与卫星图是否能反映出传统地块形态平面识别要素，对缺失信息，如房屋层数等进行填补。并利用参数化工具对数据进行切割，批量进行标签清理工作。

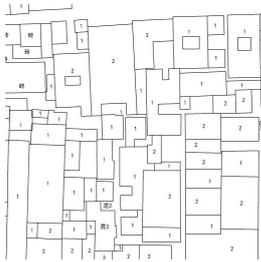
具体来说，以标注地块为中心，绘制多种大小比例的方框，以此作为裁剪框；设置地块边界线像素颜色、传统产权地块像素颜色、非传统产权地块像素颜色，作为处理前的标注标签；利用 grasshopper 中 zoom in、slider 等电池模块，分别对 CAD 图像、卫星图像、标签进行裁剪，获得对应的 CAD 图像、卫星图像、训练预标签。本研究通过数据增强技术，如翻转和旋转，扩展数据集，增加建筑肌理的多角度表示，以丰富特征并提升模型对拓扑关系的学习。这增强了模型

对不同角度拓扑结构的识别能力，提高了数据多样性，减少了过拟合风险，并增强了模型的泛化能力，从而提高了识别准确性和鲁棒性。

4 训练与结果验证

研究设计了两组模型，一组为对照组，另一组实验组则引入了 SE 注意力机制，旨在比较它们在识别与分隔任务上的效果差异。同时，也额外拓展了一组包含房屋形态标注的数据集，作为分割效果的拓展验证。

表 2 数据集-选取单图作为代表（图片来源：作者自绘）

CAD	卫星图	标注（处理前）	总卫星图尺度
			

研究一共采用扬州 2146 张（包含不同比例大小）地籍分割图像作为训练集，238 张图像作为验证集，以及 96 张图像作为测试集，经数据增强处理后，对模型进行了训练。实验结果显示，通过引入 SE 注意力机制来识别城市肌理，在传统地块形态的语义分割任务中取得了显著的效果提升。加入 SE 注意力机制后，模型在识别与分隔任务上的性能得到了显著改善，验证集和测试集上的 MIoU 均有了明显提升。这表明，相较于基于 CAD 的拓扑信息，基于图像维度的肌理特征对于传统地块形态边界的识别至关重要。

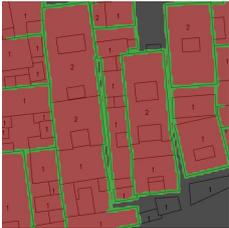
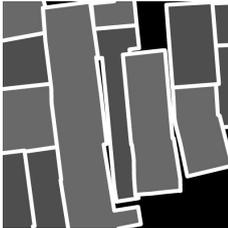
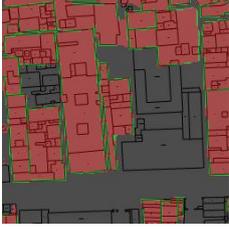
表 3 扬州地籍传统地块形态语义分割 MIoU（平均交并比）、mPrecision（类别平均精确度）、MPA（类别平均像素准确率）

数据集类型	衡量指标	Unet	Unet+SE	Unet+SE+建筑标注
验证集	MIoU	83.29%	86.36%	86.00%
验证集	mPrecision	90.37%	91.95%	92.04%
验证集	MPA	91.35%	92.46%	92.58%
测试集	MIoU	78.54%	80.91%	82.64%
测试集	mPrecision	83.40%	84.25%	89.13%
测试集	MPA	91.93%	94.29%	92.70%

两个模型的结果对比表明了传统地块形态的拓扑关系在城市形态学中具有重要意义，并且这些关系可以被有效地捕捉和表达。然而，结果也显示出，在面对一些非传统的肌理部分时，加

入了 SE 注意力机制的模型表现更为出色，如表 4 所给示例中“含非传统肌理”地块，在加入卫星图肌理后，对非传统的地块部分，如厂房等的识别有所提高。同时，城市形态学具有复杂性，传统的拓扑关系虽然是重要的基础，但对于更加细微和复杂的地块特征的识别，则需要额外的关注和处理；如下表，“复杂肌理”地块示例所示，当仅学习 CAD 特征是，模型较难准确识别边界，尤其是院落和复杂结构，其分割的边界不连续。然而，两种模型对于识别规则地块和建筑轮廓与地块边界基本重合的地块也都具有较高的准确率，但对于入口走廊型地块等则难以准确识别。

表 4 测试集部分类型结果对比（图片来源：作者自绘）

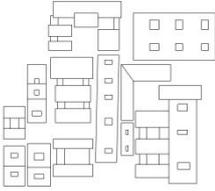
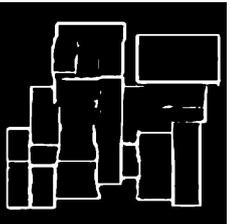
类型	Unet+SE	Unet	卫星图	标注
典型肌理				
含非传统肌理				
复杂肌理				
入口走廊型地块				

值得注意的是，如表 5 所示，在模型除输出层以外的结构未发生改变的情况下，引入对数据集中建筑形态的标注后，模型能够更好的区分院落与非传统地块，有效地提高了模型训练集和测试集识别分割的准确性。这暗示了建筑形态与地块形态在形态上的关联性。此外，对照组和实验组的测试集结果也间接验证了传统地块的形态具有一定的规则性和普适性特征。在保留拓扑关系的情况下，即使 CAD 细节关系缺失，也能识别出大致的地块形态，这一平面关系的数据补充反映了传统地块形态的识别要素，包括轴线关系、街道与入口以及建筑肌理，为模型更准确地理解和识别传统地块的形态提供了重要参考。

表 5 加入建筑标注后输出结果横向对比（图片来源：作者自绘）

Unet+SE	Unet	Unet+SE+建筑标注	卫星图
			

表 6 部分 CAD 信息缺失下的识别结果（图片来源：作者自绘）

类型	CAD	卫星图	产权边界结果
保留拓扑关系，但 CAD 信息缺失/与训练集画法不一致			

在获得单张图片识别结果后，可利用裁剪与拼贴的思想对更大的图像进行识别与输出，并转化为 SVG 图像辅助后续设计。本研究以扬州老城区地块为例，如图 2 所示，将识别小图合成完整的传统地块形态图，并利用自编 Python 脚本转化为单线骨架后重绘为 SVG 格式。将该识别图像导入 Rhino 或 CAD，可对后续识别为同一传统地块内的建筑，做统一的立面风貌改造。

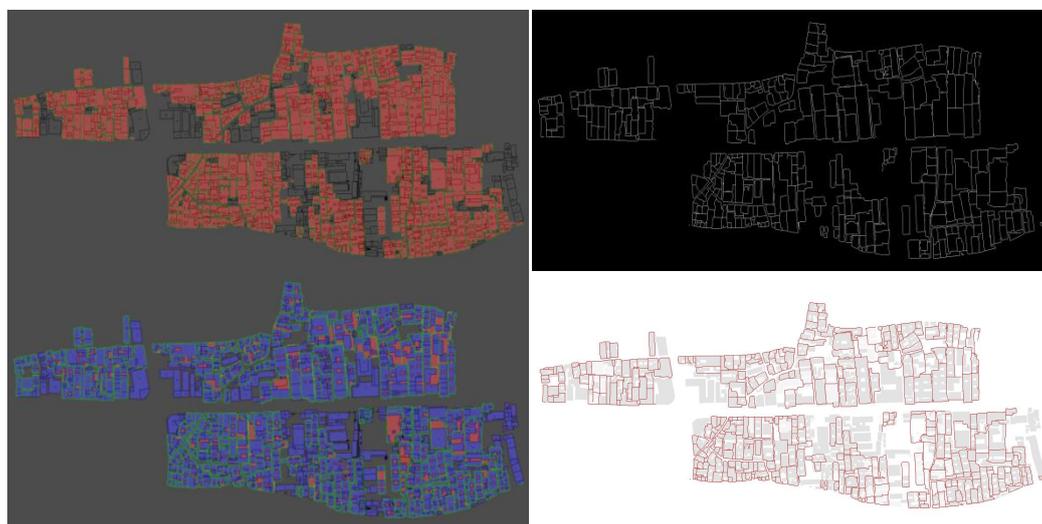


图 2 地块形态分割结果→单线提取→生成 SVG 图→导入建模软件（图片来源：作者自绘）

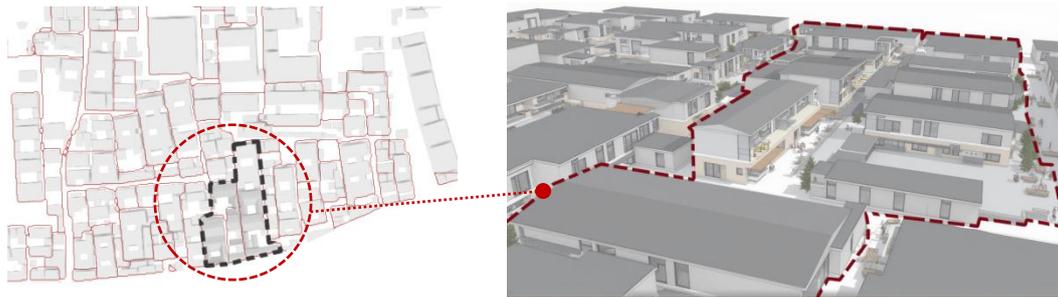


图3 选取同一传统地块进行设计（图片来源：作者自绘）

结语

本研究提出了一种基于深度学习技术的城市传统地块形态识别方法，该方法结合了语义分割算法和多模态识别的思想。通过引入SE注意力机制，新的模型在传统地块形态的语义分割任务中取得了显著的效果提升，验证集和测试集上的MIoU均有了明显提升。这表明了在形态学识别中，利用图像维度的肌理特征相较于基于CAD的拓扑信息具有更大的优势。

机器学习的结果表明，通过平面图和卫星图的结合，算法可以有效辅助设计，帮助建筑师更快速、准确地识别传统地块的形态，从而减少相关工作量。在形态学研究方面，机器学习的反向验证图像语义信息下的形态学识别要素也具有重要意义。它为我们提供了一种全新的视角和方法，使我们能够从图像中挖掘出隐藏的形态学特征，这为形态学研究提供了新的思路和技术手段。通过机器学习的方法，我们可以更全面、更深入地理解形态学的规律和特征，为城市规划和设计提供更科学、更有效的支持。

然而，研究也存在一定的局限性。尽管传统地块形态的拓扑关系具有一定的普适性，但其所包含的肌理信息受地域差异的影响而有所变化。研究的模型在扬州地籍图上表现出了较好的结果，但其泛化性能在其他城市地籍图上尚未进行验证。因此，未来的研究需要进一步考虑不同地域之间的差异，以提高模型的泛化能力，并确保其在不同城市环境下的有效性和可靠性。

参考文献

- [1] 刘鹏, Markus Nepl. 中国历史城市的地块肌理保护研究: 内涵、演变和策略 [J]. 城市规划学刊, 2020, (05): 92-99. DOI:10.16361/j.upf.202005012.
- [2] 韩冬青, 董亦楠, 刘华等. 关于城市地块格局的肌理认知与设计实践 [J]. 时代建筑, 2022, (04): 30-37. DOI:10.13717/j.cnki.ta.2022.04.010
- [3] 李冰, 苗力, 仇一鸣等. 史料及遗存匮乏困境下的传统产权地块判定方法研究——以辽宁省中小型历史城镇为例 [J]. 新建筑, 2022, (01): 102-108.
- [4] 姚佳伟, 黄辰宇, 刘鹏坤等. 基于人工智能的城市肌理识别和评价研究 [J]. 住宅科技, 2019, 39 (11): 9-14. DOI:10.13626/j.cnki.hs.2019.11.002.
- [5] El Mendili L, Puissant A, Chougrad M, Sebari I. Towards a Multi-Temporal Deep Learning Approach for Mapping Urban Fabric Using Sentinel 2 Images. Remote Sensing. 2020; 12(3):423. <https://doi.org/10.3390/rs12030423>
- [6] 郑强. 基于城市形态学的西安明城区更新研究 [D]. 西安建筑科技大学, 2016. DOI:10.27393/d.cnki.gxazu.2016.000390.
- [7] 郭莉. 基于地界的中国传统城市肌理认知与图示研究 [D]. 南京大学, 2020.
- [8] 石场. 院落社会: 北京历史文化街区的生活空间衍化与再生 [M]. 北京: 清华大学出版社, 2022: 37.
- [9] 董亦楠, 韩冬青. 历史地段保护再生实践中的地块分级——以南京小西湖街区为例 [J]. 建筑师, 2022, (02): 55-61.
- [10] 宋丽美, 徐峰, 邓源, 等. 基于科学知识图谱的中国传统村落空间形态研究综述 [J]. 华中建筑, 2020, 38 (02): 94-98. DOI:10.13942/j.cnki.hzjz.2020.02.020.

- [11] 刘铨. 地块肌理——土地划分的形态学 [J]. 建筑师, 2018, (01): 74-80.
- [12] Danfeng H, Bing Z, Hao L, et al. Cross-city matters: A multimodal remote sensing benchmark dataset for cross-city semantic segmentation using high-resolution domain adaptation networks[J]. Remote Sensing of Environment, 2023, 299