

结合无人机遥感数据和 U-Net + + 模型的建筑物足迹提取方法

曾飞

(广州市交通设计研究院有限公司, 广东 广州 511430)

摘要: 针对复杂城市场景下的建筑物足迹提取问题, 使用了更适合于捕捉高分辨率遥感图像细节特征的语义分割 U-Net + + 模型。为了让模型表现出良好的性能和降低模型过拟合化风险, 在训练前期通过 warmup 逐渐提高学习率, 当学习率达到特定的值后选用 LambdaLR 策略。融合了加权交叉熵损失函数, 以此来解决场景中建筑物目标样本不平衡、建筑物和道路目标的相似度较高的问题。在无人机遥感数据集 UDDv6.0 上完成了模型训练, 并在 UDDv6.0 测试集图像上完成了量化指标的评测实验, 实验结果表明, 结合 warmup 和 LambdaLR 策略及整合带权重的损失函数后, U-Net + + 模型在建筑物足迹提取结果上获得了更好的性能。

关键词: 无人机; 遥感; 建筑物足迹; 卷积神经网络

中图分类号: P237

文献标识码: A

A Building Footprint Extraction Method Combining UAV Remote Sensing Data and the U-Net + + Model

ZENG Fei

(Guangzhou Transportation Design & Research Institute Co., Ltd., Guangzhou 511430, Guangdong, China)

Abstract: To address the challenge of building footprint extraction in complex urban environments, a semantic segmentation U-Net + + model was used, as it is particularly effective at capturing fine-grained details in high-resolution remote sensing images. To enhance model performance and reduce the risk of overfitting, the learning rate was gradually increased during the early stages of training through a warm-up process. Once the learning rate reached a predefined value, the LambdaLR strategy was implemented. Additionally, a weighted cross-entropy loss function was employed to tackle the issues of class imbalance—due to the underrepresentation of building samples—and the high similarity between buildings and roads. The model was trained on the UAV remote sensing dataset UDDv6.0, and quantitative evaluation experiments were carried out on the UDDv6.0 test set images. The experimental results demonstrate that, by combining the warm-up and LambdaLR strategies with the weighted loss function, the U-Net + + model achieved significantly improved performance in building footprint extraction.

Key words: unmanned aerial vehicle; remote sensing; building footprint; convolutional neural network

在道路交通路线设计中, 选线工作是道路规划阶段的重要环节, 影响着整个道路的布局、功能和安全性。采用遥感技术提取出城市及乡村区域的建筑物足迹, 可以精确获取建筑物的空间分布及其周围环境信息^[1]。例如: ①可通过分析建

筑物足迹, 有效地避免穿越密集的建筑区, 降低拆迁成本和工程难度。②建筑物的分布信息能帮助设计者选择最优路线和降低对现有建筑的影响。③设计者在规划道路时, 可避开高密度的居住区和商业区, 减少噪音、空气污染以及振动对

收稿日期: 2024-09-30

作者简介: 曾飞, 工程师, 本科, 研究方向为工程测量、工程测绘

基金项目: 广州市交通设计研究院有限公司项目 (编号: BLX-SJY-2023-001); 自然资源部大湾区地理环境监测重点实验室开放基金 (编号: No. 2019002)

周围居民和商业的影响^[2]。近年来,随着无人机及卫星遥感技术的快速发展,其产生的数据越来越丰富,采用无人机及卫星遥感技术提取建筑物足迹成为一种新兴的技术手段^[3]。石茜等针对东亚区域建筑物矢量数据,使用了 0.5 m 分辨率的谷歌影像,完成了 2.8 亿个建筑矢量的提取^[2]。陈韵莹提出了一种基于高分辨率遥感影像的由低到高的分割引导增强网络,建立了南昌市的建筑物足迹数据集^[4]。张洪艳团队提出了一种基于弱监督深度学习的跨分辨率大规模土地覆盖制图网络 (Low-to-High Network, L2HNet), 实现了高分辨率高精度土地覆盖制图,该数据集中包括了建筑物足迹数据^[5]。汪国平发布了城市无人机数据集 (Urban Drone Dataset, UDD), 包含 v5.0 和 v6.0 两个版本^[6], 两个版本的数据集均包括了建筑物足迹数据。

此外,近年来卷积神经网络 (Convolutional Neural Network, CNN) 成为遥感图像分割的核心工具,也有较多学术界和工业界研究人员将卷积神经网络 CNN 应用到了无人机及卫星遥感图像语义分割,如 U-Net^[7]、HRNet^[8]、DeepLab V3^[9]、DeepLab V3+^[10]、MANet^[11]等,在语义分割任务中表现出了较好的性能,通过对遥感图像语义分割结果的处理,可以直接获取建筑物足迹数据。目前利用无人机遥感数据提取建筑物足迹的相关研究还较少,同时,卷积神经网络模型不断进步,有必要将最新的卷积神经网络模型与无人机遥感数据结合起来进行建筑物足迹数据的提取。本文结合适用于遥感图像数据的 U-Net++ 模型,以无人机遥感图像数据集 UDD 为训练样本库,在模型中融入加权交叉熵损失函数,采用合理的动态学习率策略和正则化方法,完成了 UDD 测试图像数据中的建筑物足迹提取。

1 U-Net++ 模型的引入及分析

U-Net++ 模型是在 U-Net 模型基础上发展而来的一种语义分割模型,采用的仍然是编码-解码结构。U-Net 在编码上采用卷积层,解码上采用去卷积层。在编码阶段,U-Net 通过卷积和最大池化对数据进行特征提取,提取过程共包括 4 次卷积,每一层都是对数据进行更细一步的特征提取,层次越深,对数据特征提取更细节。在解码阶段,U-Net 进行上采样,对编码过程中提取的数据特征进行融合,每一层只能对上一层的数据特征进行融合,到达最后一层融合完所有特

征作为最后的输出。

在 U-Net 模型基础上,U-Net++ 模型主要改进点在于:结合了类 DenseNet 结构,密集的跳跃连接有助于提升梯度传播的流畅性,增加 U-Net 中的空洞结构和构建编码器与解码器特征图之间的语义连接^[12]。同时,U-Net++ 模型引入了深度监督机制并具备了剪枝功能。高分辨率遥感图像包含许多细节和复杂的地物结构,U-Net++ 网络能够有效地捕捉图像的细节特征,更适合处理高分辨率遥感图像的语义分割任务,能够产生比 U-Net 性能更佳的检测结果^[13]。U-Net++ 模型引入了多尺度的跳跃连接来取代 U-Net 单尺度的跳跃连接。U-Net++ 模型采用长短连接取代了 U-Net 原来的长连接,能够实现不同层次之间进行特征值的信息捕捉,将高分辨率特征与解码器上采样后的输出特征连接起来,密集跳跃连接方式可以减缓编码器与解码器之间的语义差异。

同时,U-Net++ 还使用了深度监督,支持两种模式:①所有分支输出取均值的精准模式;②从所有分支中选择一个分支进行直接输出的快速模式。在采用深度监督方法下的模型训练能够将子网络的输出结果直接输出,从而剪掉网络中的多余部分。U-Net++ 在处理遥感图像时,能够有效区分细节和复杂结构,对于地物分类中的细节识别具有显著优势,综合以上因素,本文选用 U-Net++ 模型完成无人机遥感图像数据中的建筑物足迹提取。U-Net++ 模型结构如图 1 所示。

2 U-Net++ 模型训练策略及无人机数据集 UDD

2.1 模型训练的动态学习率调整

一般而言,如果模型在训练过程中一直使用一个固定学习率,当训练集损失下降到一定程度后,可能会出现停滞的状态,不再继续下降,而是在某范围内震荡。当损失函数收敛到局部极小值附近时,因为学习率可能相对较大,其更新步幅可能会过大。因此在模型训练过程中,动态学习率调整是一种常见的优化策略,通过根据训练情况动态选择适宜的学习率,实现对模型的优化。目前广泛应用的动态学习率调整策略众多,StepLR、Multi-StepLR、ExponentialLR、Cosine AnnealingLR、LambdaLR、OneCycleLR、Poly 都是比较常用的学习率调整策略^[14,15]。实验设置了初始学习率为 0.001, BatchSize 大小为 16。

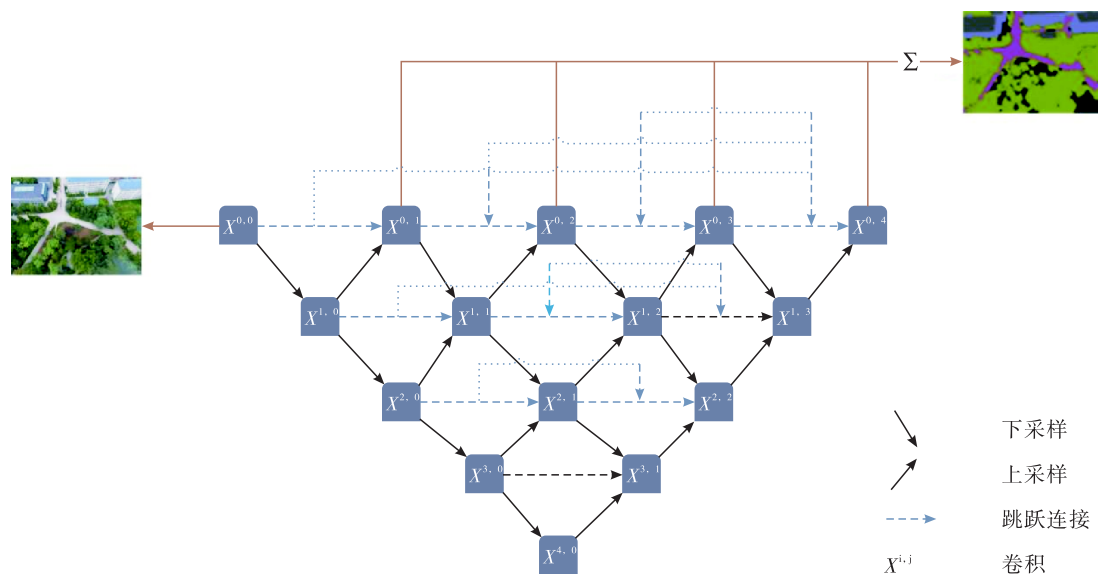


图1 U-Net++模型结构

本文实验采用了结合 warmup 和 LambdaLR 的动态学习率调整策略。在这一策略中,初始学习率较小,模型训练前期通过 warmup 逐渐提高学习率,当学习率达到特定的值后选用 LambdaLR 策略,从而逐渐降低学习率。模型学习率的变化趋势为升高—稳定—降低。相对应地,模型的学习阶段也呈现稳定—学习—稳定的趋势。采用这样的训练策略能够使模型表现出良好的性能和降低模型过拟合化的趋势。

2.2 正则化

一方面,采用正则化能有效地降低模型过拟合化的概率。L1 正则化和 L2 正则化是目前比较常见的策略^[16]。本文实验使用的是 L2 正则化策略,L2 正则化又被称作权重衰减。顾名思义,该策略的基本原理就是设置一个权重损失,在模型训练过程中逐渐降低训练损失率,这样的做法可以有效地降低模型过拟合的可能性,保证该训练模型的泛化能力。另一方面,正则化可以进一步优化梯度的传递,防止在深层网络中出现梯度消失或梯度爆炸的情况,进而改善训练的稳定性。通过适当的正则化,网络可以更有效地学习多尺度信息,并在多个尺度上进行特征融合。此外,正则化还能增强模型的鲁棒性,使其在面对不同的图像分布、噪声或变化时依然能够表现稳定。本文实验在 U-Net++ 中引入 L2 正则化,可以在一定程度上避免网络对训练数据中特定噪声的过拟合,从而提升模型对真实世界数据的适应性。

2.3 加权交叉熵损失函数

在数据集中,由于在不同的场景下,建筑物

目标的分布不均衡,有点的遥感样本图像中建筑物目标在数据集中的比例较小,通过加权交叉熵损失函数,增加小物体类别的权重,模型会在训练过程中更加重视这些小物体,从而提高其在小物体上的分割效果^[17-18]。此外,建筑物目标与道路目标区域存在较高的相似度,加权交叉熵损失函数能够有效地强化对这些难分区域的学习效果。通过对建筑物类别赋予较高的权重,损失函数鼓励模型在这些区域进行更精细的学习,提高这些区域的预测准确性。加权交叉熵损失函数通过优化模型对不同类别的敏感度,使得 U-Net++ 模型在不同条件下的稳定性和鲁棒性更高。例如,在遥感图像中某些物体在空间上被遮挡。加权交叉熵通过为困难类别赋予较高的权重,促使模型在这些困难的分割区域上进行更精确的学习。以本文所选取的数据集 UDD 为例,有的样本中含有的道路比重较大,同时道路和建筑物区域有部分混合和重叠,建筑物部分区域会被植被遮挡,因此要将道路类别的权重设置为小于 1。而建筑物目标类别的权重设置为大于 1。具体公式参见式 (1)。 M 是 BatchSize ($=16$), H 、 W 分别是图像的行列数, C 是类别数。 w_k 是 k 类的权重, y_{mij} 判断像素点是否为 k 类,相同取 1,否则为 0。 p_{mij} 是像素点是否为 k 类的概率。

$$Loss = -\frac{1}{MHW} \sum_{m=1}^M \sum_{i=1}^H \sum_{j=1}^W \sum_{k=1}^C w_k y_{mij} \log(p_{mij}) \quad (1)$$

2.4 无人机数据集 UDD

无人机数据集 UDD 是由北京大学信息学院汪国平教授团队发布的无人机遥感图像数据集,

该数据集采用了大疆精灵 4 无人机采集，飞行高度为 60 ~ 100 m，场景的采集地点为北京大学、葫芦岛市、河南大学、沧州市。其对应的采集区域覆盖了城市区域及乡村区域，对应的地物类型比较丰富。图像分辨率有两种：4096 × 2160 和 4000 × 3000。地物类型为建筑物、植被、道路、车辆、其他类型。同时，该数据集官方网站发布了几种经典卷积神经网络模型对应的精度指标。本文实验采用的是 UDD v6.0 数据集，该数据集包含训练样本 106 张，验证集包含训练样本 35 张。其对应的样例图像见图 2。



(a) 无人机 UDD 样例图像 1



(b) 无人机 UDD 样例图像 2

图 2 UDD v6.0 数据集训练样本

3 实验环境搭建及结果分析

3.1 实验硬件和软件环境

实验所采用的 GPU 硬件环境：显卡 GeForce RTX3060 (CUDA Core 核心数量 3584，显存容量 12 GB，显存位宽 192 位、加速频率 1.78 GHz)。软件环境：操作系统为 Ubuntu 20.04.3、Pytorch 版本为 1.13.0、CUDA 版本为 11.1。

3.2 实验量化指标及结果

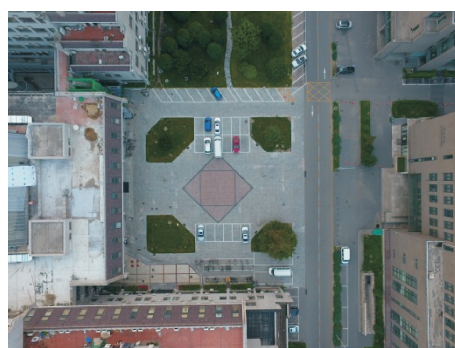
为了和 UDD 的评价指标保持一致，本文采用两个评价指标，分别是像素精度 (pixel accuracy, PA)、平均交并比 (mIoU) 来评估损失函数的性能。

PA 的定义见式 (2)， p_{ij} 表示的是像素点真实类别为 i ，预测类别为 j ，表示分类正确的像素点数和所有的像素点数的比值；mIoU 的定义见式 (3)， k 是总的类别数量，代表各类别交并比 (IoU) 的平均值，是语义分割模型总体性能的重要评测指标。

$$PA = \frac{\sum_{i=0}^k p_{ii}}{\sum_{i=0}^k \sum_{j=0}^k p_{ij}} \quad (2)$$

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ij} - p_{ii}} \quad (3)$$

根据 UDD 数据集发布的评测结果，是将验证集中的无人机遥感图像样本 DJI_0627 挑选出来作为测试样本，见图 3 (a) 测试集样本原始图像。测试样本图像对应的分类参考图见图 3 (b)。模型的预测结果见图 3 (c)。在 UDD v6.0 数据集中，将建筑物分为立面和楼顶两种子类型，因此在最终结果中需要将两者合并为一种类型，结果见图 3 (d)，为了显示更加直观，该图中只包含了建筑和植被两种地物目标。最后，与 UDD 官方发布的评测结果进行量化对比，对比结果见表 1。



(a)



(b)

图 3 UDD 数据集测试集样本



(c)



(d)

图3 UDD 数据集中的测试集样本 (续)

表1 UDD 数据集量化对比评测结果

模型	PA/%	mIOU/%	模型大小 SIZE/M
GCN	86.54	72.58	302
ENet	85.20	69.44	5.5
DeepLab V3 +	86.90	73.18	721
U-Net + +	87.81	74.75	36.64

从图3及表1的结果可知,相对于DeepLab V3+模型而言,在采用PyTorch框架4字节的量化设置下,U-Net++模型在模型大小方面有较大幅度的降低,模型大小降低为36.64 M;同时,在像素精度PA和平均交并比mIOU两个指标上,均有较大幅度的性能提升(PA上升到87.81%,mIOU上升到74.75%)。从可视化的效果来看,对于建筑物足迹的提取,U-Net++模型在性能上有一定的提升。

4 结论

本研究针对真实城市场景下的建筑物足迹提取问题,采用了U-Net++模型进行语义分割。为提高模型的训练效果,使用了warmup策略逐步提升学习率,使模型在初期阶段平稳收敛,并在学习率达到预定值后切换为LambdaLR策略,实现更灵活的学习率调整。针对建筑物样本在复

杂场景中的不平衡问题,研究引入了加权交叉熵损失函数,有效平衡了不同类别样本的影响,降低了小目标建筑物的漏检率。结果表明,该方法在复杂城市环境中的建筑物足迹提取精度较高,具有良好的鲁棒性与泛化能力,可为交通线路规划等应用提供可靠参考。

参考文献:

- [1] 崔星,杜春兰. 基于GIS+BIM信息协同的景观参数化设计研究——以山地风景环境道路规划设计实验为例[J]. 中国园林, 2023,39(6):39-45.
- [2] SHI Q, ZHU J, LIU Z, et al. The last puzzle of global building footprints—mapping 280 million buildings in east asia based on vhr images [J]. Journal of Remote Sensing, 2024,4:138.
- [3] 李洋,胥海燕,颜青松,等. 单张卫星影像的建筑自动化三维重建[J]. 测绘科学, 2024,49(1):132-42.
- [4] 陈韵莹. 基于LTHNet的高分辨率遥感影像建筑物提取研究[D]. 赣州:江西理工大学, 2023.
- [5] LI Z, HE W, CHENG M, et al. SinoLC-1: the first 1m resolution national-scale land-cover map of China created with a deep learning framework and open-access data [J]. Earth Syst Sci Data, 2023,15(11):4749-80.
- [6] CHEN Y, WANG Y, LU P, et al. Large-scale structure from motion with semantic constraints of aerial images [C]. Lecture Notes in Computer Science, 2018:11256.
- [7] HE M, QIN L, DENG X, et al. Transmission line segmentation solutions for uav aerial photography based on improved UNet[J]. Drones, 2023,7(4):274.
- [8] HUANG S, HAN W, CHEN H, et al. Recognizing zucchinis intercropped with sunflowers in UAV visible images using an improved method based on OCRNet [J]. Remote Sensing, 2021,13(14):2706.
- [9] TALLAM K, NGUYEN N, VENTURA J, et al. Application of Deep learning for classification of intertidal eelgrass from drone-acquired imagery [J]. Remote Sensing, 2023,15(9):2321.
- [10] NJANE S N, TSUDA S, VAN MARREWIJK B M, et al. Effect of varying UAV height on the precise estimation of potato crop growth [J]. Frontiers in plant science, 2023,14:1233349.
- [11] HE P, JIAO L, SHANG R, et al. MANet: Multi-Scale Aware-Relation Network for Semantic Segmentation in Aerial Scenes[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022,60:1-15.
- [12] 袁洲,郭海涛,卢俊,等. 融合UNet++网络和注意力机制的高分辨率遥感影像变化检测算法[J]. 测绘科学技术学报, 2021,38(2):155-9.

(下转第119页)

身紧跟时事政治的特点,其成了开展高职学生国家安全教育的重要载体,通过向学生讲授国内外重要时事政治,特别是结合当前我国面临的国际形势,有助于高职学生更好地认识当前我国面临的国家安全环境,提升学生的国家安全意识与爱国情感。此外,国家安全教育还可依托国防教育进行,以更大力度将国家安全、国际形势等内容融入国防教育中,进一步拓宽国家安全教育渠道。

最后,注重理论教育与实践教育相结合。除课堂教学主渠道外,实践教学也是开展国家安全教育的重要途径。在实际效果上,由于实践教育的实践特性,能够以更形象甚至是“身临其境”的方式让高职学生切身实地地感受到国家安全的概念、范畴、价值等内容,因此相较于理论教育而言,实践教学有助于给学生带来更为震撼、共情的教学体验,因而在提高教学效果上有其独到之处。因此,实践教学可作为课堂教学主渠道的有益补充,实现实践教学与课堂教学的同向同行,推动理论教育与实践教育相结合,助力国家安全教育成效最大化。在具体开展方面,可充分发挥第二课堂主阵地作用。可开展形式丰富的实践课程,包括但不限于与国家安全相关的知识竞赛、“415 全民国家安全教育日”宣传活动、国家安全讲座、校园文化建设等,在这一过程当中,还可与高职学生的学情相结合,结合学生背景、专业等开展针对性实践教学。此外,还可以组织学生开展社会实践,带领学生利用周末或寒暑假时间走访国家安全教育基地、国防教育中心等实践场所,开展国家安全教育现场教学,引导学生培养坚定的国家安全意识与爱国主义情感。

4 结语

当前,国际形势波诡云谲,境外势力主导的

危害我国国家安全的网络负面舆情给我国高职院校国家安全教育带来了众多挑战,深入夯实高职院校国家安全教育具有重大理论与现实价值,对于我国完善以总体国家安全观为代表的国家安全理论体系以及做好高职院校国家安全教育工作意义重大。面对网络负面舆情的冲击,政府部门、网络平台、社会团体、高职院校应正视现状形成治理合力,共同推动高职学生国家安全意识的养成,夯实国家安全教育基石。

参考文献:

- [1] 习近平. 高举中国特色社会主义伟大旗帜 为全面建设社会主义现代化国家而团结奋斗——在中国共产党第二十次全国代表大会上的报告[M]. 北京:人民出版社,2022.
- [2] 中华人民共和国教育部. 2023 年全国教育事业发展基本情况[EB/OL]. (2024-03-01)[2024-06-27]. http://www.moe.gov.cn/fbh/live/2024/55831/sfcl/202403/t20240301_1117517.html.
- [3] 李昌祖,张洪生. 网络舆情的概念解析[J]. 现代传播(中国传媒大学学报),2010(9):139-140.
- [4] 习近平主持召开二十届中央国家安全委员会第一次会议强调加快推进国家安全体系和能力现代化以新安全格局保障新发展格局[J]. 中国军转民,2023(12):6.
- [5] 张荣刚. 以习近平法治思想指引网络安全治理[N]. 中国社会科学报,2024-05-06(A01).
- [6] 赵志云,杜洪涛,张楠. 大数据时代网络社交平台风险治理的国际政策比较与启示[J]. 中国行政管理,2020(2):129-134.
- [7] 中国互联网络信息中心. 第 53 次《中国互联网络发展状况统计报告》[EB/OL]. (2024-03-22)[2024-06-27]. <https://www.cnnic.net.cn/n4/2024/0322/c88-10964.html>.
- [8] 习近平. 思政课是落实立德树人根本任务的关键课程[N]. 人民日报,2020-09-01(01).
- [13] 吕少云,李佳田,阿晓荟,等. Res_ASPP_UNet++:结合分离卷积与空洞金字塔的遥感影像建筑物提取网络[J]. 遥感学报,2023,27(2):502-519.
- [14] 高淑芝,裴志明,张义民. 动态自适应学习率优化的卷积神经网络轴承故障诊断方法[J]. 机械设计与制造,2023(7):211-4,22.
- [15] 张东阳,陆子轩,刘军民,等. 深度模型的持续学习综述:理论、方法和应用[J]. 电子与信息学报,2024,46(10):3849-3878.
- [16] 孙玉春,王思山,童乐言,等. 基于损失函数权重自适应的路径规划优化算法研究[J]. 计算机工程与应用,2024,60(16):85-93.
- [17] 胡峰,杨新瑞,汤成富,等. 基于自适应损失函数的句子级远程监督关系抽取[J]. 智能系统学报,2024,19(3):697-706.
- [18] 徐少平,陈晓军,罗洁,等. 基于相邻特征融合的红外与可见光图像自适应融合网络[J]. 模式识别与人工智能,2022,35(12):1089-1100.

(上接第 56 页)