

基于深度学习的土木工程计算机视觉健康监测

方成¹, 于盛鑫¹, 李永刚², 贾王龙², 杨鹏博², 杨欣悦¹

(1. 同济大学土木工程学院, 上海 200092; 2. 中国二十二冶集团有限公司, 河北唐山 064000)

摘要: 土木工程领域的健康监测对保证工程长期、稳定服务有着重要的意义。相较于传统的监测方法, 基于深度学习的计算机视觉技术具有高效、准确等优势。本文对基于深度学习的计算机视觉技术在土木工程全生命周期健康监测领域中的应用进行系统综述。首先, 借助文献可视化软件对该领域文献进行科学计量分析; 其次, 简要阐述了计算机视觉技术的发展历程, 总结了在构建深度学习数据集过程中数据获取、数据处理和数据标注三个重要环节的方法与内容; 最后, 重点回顾了施工现场安全管理、在役结构局部损伤检测和结构灾后整体损伤评估等应用场景中基于深度学习的计算机视觉技术的发展历程与工程实际应用价值, 并展望了可拓展的应用方向。

关键词: 深度学习; 计算机视觉; 土木工程; 全生命周期; 健康监测

中图分类号: TU317; TP391.4; TU714

文献标志码: A

Deep Learning-Based Computer Vision for Health Monitoring in Civil Engineering

Cheng Fang¹, Shengxin Yu¹, Yonggang Li², Wanglong Jia², Pengbo Yang², Xinyue Yang¹

(1. College of Civil Engineering, Tongji University, Shanghai 200092, China; 2. China MCC22 Group Corporation Ltd., Tangshan 064000, China)

Abstract: Health monitoring in the field of civil engineering is of great significance to ensure the long-term and stable service of infrastructure. Compared with traditional monitoring methods, computer vision technology based on deep learning has the advantages of high efficiency and accuracy. This paper provides a systematic review on the application of deep learning-based computer vision technology in the field of civil engineering life cycle health monitoring. Firstly, a scientific econometric analysis of the literature in this field is carried out with the help of literature visualization software. Secondly, the development process of computer vision technology is briefly described, and the methods of data acquisition, data processing and data annotation in the process of constructing deep learning data sets are summarized. Finally, the development history and practical engineering application value of computer vision technology based on deep learning in construction site safety management, local damage detection of in-service structures and overall damage assessment of structures after disaster are reviewed. In addition, the future application directions are prospected.

Key words: deep learning; computer vision; civil engineering; life cycle; health monitoring

随着基础项目建设的高速发展, 工业民用建筑进入了饱和期和维护期。一方面, 建筑结构在长期荷载、环境腐蚀、材料变形和疲劳破坏等因素的耦合作用下出现大量的损伤破坏。另一方面, 结构在强台风、海啸、破坏性地震等偶发环境外力作用下将发生不同程度的破坏及倒塌。因此, 在工业智能化时代, 如何利用人工智能技术来高效地辅助土木工程建设全生命周期中涉及到的施工安全管理、结构损伤检测及结构重大灾后整体评估等任务一直是学者们不断探索的课题。本文旨在对现有基于深度学习算法的计算机视觉技术在土木工程建设全生命周期中的研究和应用进行针对性总结与归纳, 通过进一步分析各检测体系中所涉及算法的共性问题, 从探索深度学习算法可以完成的计算机视觉任务开始, 到挖掘出更丰富的可以应用的工程领域, 从而实现从算法微观层面到构建整体检测体系的宏观层面的飞跃, 助力人工智能新成果更好地辅助土木工程领域任务。

收稿日期: 2023-06-06

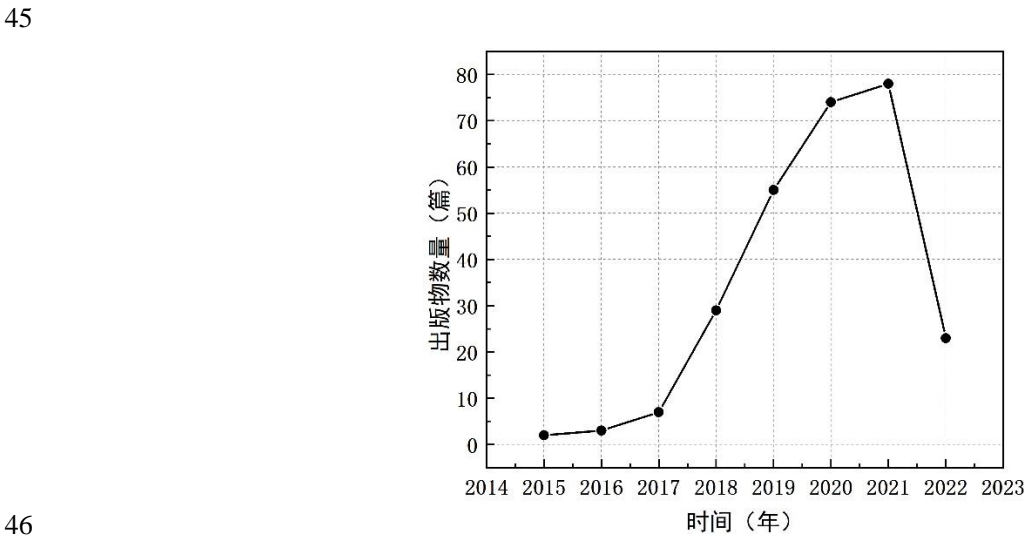
基金项目: 国家自然科学基金项目(52078359, 51820105013)

第一作者: 方成 (1985—), 男, 研究员, 博士, 主要研究方向为韧性结构及智能材料与结构研究。E-mail: chengfang@tongji.edu.cn

通信作者: 于盛鑫 (2000—), 男, 硕士生, 主要研究方向为韧性结构及人工智能与土木工程学科交叉研究。E-mail: ysxnpng@tongji.edu.cn

39 1 文献检索与分析

40 为了系统地识别和分析近年来基于深度学习的计算机视觉技术在土木工程建设全生命周期中的应用，
41 本研究选择 Web of Science 作为检索文章的数据源，通过设置包括“土木工程”且“深度学习”且“计算
42 机视觉”或“结构健康监测”或“结构表面损伤”或“施工安全”或“震后损伤评估”或“灾后影像识
43 别”的关键词或主题，设置跨度为 2015 年 1 月至 2022 年 5 月的发文时间，在剔除不相关文献后，选择使
44 用 271 篇文献进行分析，文献出版时间分布如图 1 所示。



46 图 1 文献出版时间分布

48 Fig.1 Distribution of publication time

49 本研究基于 Citespace^[1]进行了关键词分析，绘制了本研究领域的高频、高中介中心性关键词排序表格，
51 见表 1。可以看出，“deep learning”和“computer vision”位居频率排名的前两名，说明所检索文章较好
52 地反映了基于深度学习的土木工程计算机视觉技术研究；结合高中心性，可以看出“convolutional neural
53 network”与多个研究对象关联紧密，说明其在本研究领域具有重要意义。

54 表 1 关键词排名

55 Tab.1 Keyword ranking

项目	按频率排序		按中介中心性排序	
	关键词	频率	关键词	中介中心性
1	deep learning	112	convolutional neural network	0.26
2	computer vision	62	crack detection	0.19
3	convolutional neural network	38	object detection	0.19
4	neural network	37	recognition	0.17
5	damage detection	36	classification	0.17
6	identification	33	computer vision	0.16

56 2 计算机视觉技术简介

58 计算机视觉技术旨在助力计算机使用算法来理解数字图像和视频并提取有用的信息，重点在于图像
59 数据的特征设计，包括提取特征点、边缘和轮廓等数据特征。最初的特征设计是基于人工实现的，例如，

60 尺度不变特征变换算法实现了显著稳定的图像特征点检测，可有效用于图片匹配；方向梯度直方图算法
61 对图像的几何和光学形变具有良好的不变性，通过与支持向量机分类器结合广泛应用于图像识别中，可
62 实现实时性检测。随着互联网技术的发展以及便携式电子设备的出现，大量图片数据资源被获取，人工
63 特征设计技术衍生成了成熟的数据集图像处理技术，包括降噪、傅利叶变换等，主要负责图像压缩、增
64 强复原和匹配识别三大任务。此外，机器学习在快速发展后被广泛应用于图像识别、结构优化^[2]、自然语
65 言处理等多个领域。由此，基于机器学习的计算机视觉技术在健康监测等领域得到迅速应用与发展。

66 为解决一般机器学习算法效率低、精度不足等问题，有学者提出基于深度学习卷积神经网络（CNN）
67 ^[3]的计算机视觉技术，其具有卓越的学习能力、广泛的覆盖范围和良好的适应性等优势。表 2 列举了与计
68 算机视觉相关的常用深度神经网络。通过与深度神经网络相结合，计算机视觉技术的识别精度得到了大
69 幅提升，应用场景也快速扩展，在图像分类、目标检测、语义分割和实例分割任务中发挥重要作用。

70 表 2 与计算机视觉相关的深度神经网络分类

71 Tab.2 Classification of deep neural networks related to computer vision

领域	算法	细节描述
Object detection	R-CNN、Fast R-CNN、Faster R-CNN	双阶段有锚框深度学习目标检测框架
	YOLO、YOLOX	单阶段有锚框深度学习目标检测框架
	RetinaNet、CornerNet、CenterNet	无锚框深度学习目标检测框架
Pixel level segmentation	FCN、DeepLab、Segnet	语义分割深度学习模型
	Mask R-CNN、RefineNet、U-Net	实例分割深度学习模型
	PointNet、PointCNN、RealPoint3D、VoxelNet	常应用于激光雷达点云数据的处理
Image data processing	PointFusion、PointRCNN、RoarNet	可实现激光雷达信息和图像数据融合
Reconstruction	SRCNN、GAN、f-GAN、EBGAN、InfoGAN	应用于图像、视频等信息处理
	PoseNet、SurfaceNet、MVSNet、CNN-SLAM	常应用于三维重建领域
Model pruning	MobileNet、ShuffleNet、SqueezeNet、EffNet	有效解决模型压缩精简问题

72

73 3 深度学习数据集的构建

74 创建高质量数据集对提高算法训练效率和学习泛化能力具有重要意义。本节将介绍常用的数据获取
75 手段、数据处理技术以及数据标注方式。

76 3.1 数据获取

77 通常将深度学习样本分成训练集、验证集和测试集三部分。训练集用于训练神经网络学习特定特征；
78 验证集用于对网络训练结果进行验证，确定网络参数和最终模型；测试集则用于评估算法的运行状况及
79 检测性能。目前常用的数据获取方法有人工收集、虚拟合成、网络爬取、对抗生成^[4]等，它们可在一定程
80 度上解决深度学习模型训练效果不佳以及学习泛化能力不足等问题。

81 3.2 数据处理

82 除了通过增加新样本来扩充训练集样本数量外，图像处理技术也可增强图像包含的数据信息。例如，
83 通过噪声消除等方法可进一步优化图像对环境变化的鲁棒性，此外，图像处理算法还能减少分辨率差异，
84 优化模型对特征的提取能力和检测性能^[5]。近年来深度学习模型的应用也从二维扩展到三维，实现了扩充

85 数据的目的。以上方法对提升数据驱动型深度学习算法的训练效果和检测性能至关重要。

86 3.3 数据标注

87 为了将数据集中目标对象的位置和像素等信息转化为计算机所能理解的语言，数据标注的过程是必
88 不可少的。目前常用的标注工具大多源自 Github，其功能几乎覆盖了训练深度学习模型所需要的全部数
89 据标注方法。除此之外，各大科技公司也相继开发出了支持人机协作自动化数据标注及人工矫正的云平
90 台。表 3 介绍了常用的标注工具，列举了典型的标注类型和文件导出格式，以便于实际应用。

91 表 3 常用的标注工具及其特点

92 Tab.3 Common annotation tools and their features

应用领域	工具	标注类型	文件导出格式
图像分类、目标检测	LabelImg	2D 框	PASCAL VOC, COCO, CreateML
	Vatic	视频标注	JSON
	Pixel Annotation Tool	多边形分割	JSON
目标检测、语义分割	Labelme	点标注、线标注、2D 框、多边形分割	JSON, VOC, COCO
	VOTT	点标注、线标注、2D框、多边形分割、	PASCAL VOC, CSV, VoTT
	VIA-VGG Image Annotator	语义分割、视频标注	CSV, JSON
	CVAT	点标注、线标注、2D 框、多边形分割、	PASCAL VOC, COCO, CAVT
语义分割、实例分割	Colabeler	语义分割、视频标注、3D 点云	PASCAL VOC
	ModelArts	人机交互协作标注	PASCAL VOC

93
94 4 基于深度学习的计算机视觉技术在土木工程中的应用

95 基于深度学习的计算机视觉技术在土木工程结构建设全生命周期中均有广泛而丰富的应用，从对建
96 造施工现场的危险把控和识别，到对结构投入使用过程中的结构连接安全性检测以及结构表面典型缺陷
97 检测，到对老化结构的健康状况监测以及灾后建筑物倒塌检测评估等。

98 4.1 施工现场安全管理

99 基于深度学习的计算机视觉技术在施工现场安全管理中的应用包括施工人员安全、设备安全以及进
100 度管理。

101 （1）施工人员安全监测

102 正确合理佩戴个人防护装备是施工人员预防危险的重要举措，利用深度学习目标检测技术可以检测
103 工人的状况及潜在安全隐患。以安全帽检测为例，已有研究成功实现工人安全帽佩戴检测^[6-7]，通过构建
104 高质量数据集，可显著提高防护设备检测效率^[8]，实现施工场景中人与物的识别^[9-10]。此外，计算机视觉
105 技术在工人活动识别领域也取得较大的进展，研究相继解决了运动传感器布置和混合不安全行为检测^[11]
106 等问题，实现了工人活动模式定义^[12]和行为解释，使施工人员安全管理更为先进高效。表 4 列举了施工人
107 员安全监测任务中使用的多种技术方法，包括工人防护设备检测、潜在环境安全因素检测及工人不安全
108 行为识别和解释等。

109 表 4 施工人员安全监测方面各方法评价

110 Tab.4 Evaluation of construction personnel safety monitoring methods

施工人员安全监测内容	作者及方法	优势	局限/展望
安全帽检测	Shen 等 ^[7] 基于 CNN 并	较 Fang 等 ^[6] 取得了更高的精度	无法精确检测背对监控摄像头工

	引入边界框回归技术	和召回率	人头盔的佩戴情况
包括四种颜色的头盔、工人和安全背心检测	Wang 等 ^[8] 基于 YOLO	可同时检测多个类别防护设备，YOLO v5 模型具有良好的性能	很难检测到小的、阻挡的和奇怪的实例
工人和高空作业安全带检测	Fang 等 ^[9] 基于 Faster R-CNN	Faster R-CNN 检测工人的准确率和召回率均较高	样本大小和安全带的颜色对识别结果造成影响
工人混合不安全行为识别	Ding 等 ^[11] 结合 CNN 和 LSTM	可准确检测现场工人行为	可能会存在运动捕捉耗时多、时空背景不清晰等问题
识别工地图像中的各种建筑活动	Luo 等 ^[12] 基于 Faster R-CNN 和关联网络	成功定义施工人员的各种活动模式	很难区分长期活动和过渡状态；

(2) 施工设备安全监测

施工现场常有大型工程设备，一旦发生不规范设备操作便极易引发安全事故，因此针对工程设备的监测尤为重要。表 5 列举了该监测任务中多种常见的基于深度学习的计算机视觉技术方法，首先解决了工程设备实时检测问题^[13]，并陆续提高了检测精度^[14]，但仍存在遮挡、无法验证距离测量精度等问题^[15]。为此，Zeng 等^[16]基于改进的 YOLO v3 算法实现了对远场视频工程设备的实时多尺度检测与定位，图 2 展示了该研究中各种检测方法关键性能指标的比较，结果表明改进的 YOLO v3 较 Faster R-CNN 和 R-FCN 等在检测速度、精度和鲁棒性方面都得到了显著的提升。值得一提的是，基于深度学习的目标检测算法在设备和工人之间交互行为的安全性检测中也具备突出优势，例如，可实时监控工人与设备之间潜在的碰撞、判断设备与工人在空间中的接近度以及量化发生交互行为时工人的安全水平等。

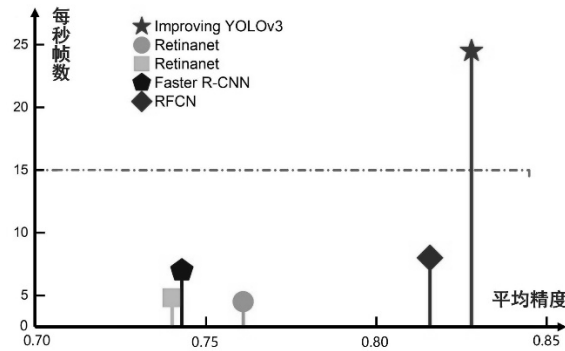


图 2 各类工程设备检测方法的性能比较^[16]

Fig.2 Performance comparison of various testing methods for engineering equipment

表 5 施工设备安全监测方面各方法评价

Tab.5 Evaluation of construction equipment safety monitoring methods

施工设备安全监测内容	作者及方法	优势	局限/展望
挖掘机实时检测	Fang 等 ^[13] 基于改进的 Faster R-CNN	可同时检测工人，检测速度较快，成本较低	存在遮挡问题
自卸车、挖掘机、装载机检测	Kim 等 ^[14] 基于 R-FCN	可同时实现五类施工设备的检测，平均检测精度较高	缺乏时效性
现场设备及工人的姿势状态实时检测	Luo 等 ^[15] 基于 YOLO v2	实现了现场设备及工人的姿势状态（即移动或静止）的及时反馈	假定了恒定的设备运行速度
远场视频中多个工程设	Zeng 等 ^[16] 基于改进的	实现了复杂环境下多个工程设备	进一步可研究更远距离施工监控

（3）施工进度管理

施工进度管理是维持施工秩序、规范建筑活动及控制风险的重要方法，其关键是设备活动识别。表 6 列举了施工进度管理任务中常见的计算机视觉技术方法。将 CNN 和长短期记忆网络(LSTM)结合可实现基本的设备活动识别，在此基础上增加检测、跟踪和分类网络，可以实现长视频中工程设备活动及生产力的评估与分析^[17]，确保更安全的施工操作^[18]、更有效的人力资源配置^[19]以及更高效的施工监管^[20]。此外，自动化施工场景分析也是基于深度学习的计算机视觉技术重要应用领域，将场景分析技术用于安全管理中，可以更清晰地反映当前施工场景的含义^[21]，以便施工进度管理。

表 6 施工进度管理方面各方法评价

Tab.6 Evaluation of construction schedule management methods			
施工进度安全管理内容	作者及方法	优势	局限/展望
自动分析多台挖掘机的活动和生产力	Chen 等 ^[17] 基于		
	Faster R-CNN	适用于长视频，且可自动计算	整体活动识别准确率和生产率计算准确
	+Deep SORT +3D ResNet	多台设备生产率	率并不高
土方工程中不规范操作的自动识别	Lin 等 ^[18] 结合 CNN 与 LSTM	实现了土方工程中不规范操作的自动识别	视频分析能力有限，仅可同时识别一台挖掘机和最多两辆自卸卡车
评估工人劳动力消耗以配置人力资源	Li 等 ^[19] 基于 Faster R-CNN	可视化评估工人的劳动力消耗	进一步将应用于大型施工场景
预制墙体施工进度自动监控	Wang 等 ^[20] 基于 Mask R-CNN	实现了预制墙体施工进度监控现场管理的数字化	受摄像头移动和视野范围的影响较为严重

从本节可以看出，目前研究局限于单次解决某个或几个对象特征自动识别的问题，而没有从整体上实现施工现场安全管理中多任务场景内动态关联，从而达到高效的信息获取与分析。未来可以引入视频结构化技术，通过汇聚全量施工现场视频场景记录，在连续帧中检测和关联相同对象来实现动态关联网络，进一步提高检测能力，实现对所获信息的高效利用。

4.2 在役结构损伤检测

基于深度学习的计算机视觉技术在土木工程结构服役时的局部损伤检测中已发挥出积极作用，检测任务主要分为结构连接安全性检测和结构表面典型损伤检测。

（1）结构连接安全性检测

焊缝连接是钢结构、组合结构中常用的连接形式，控制焊缝质量是确保构件连接安全的重点，表 7 列举了多种焊缝缺陷识别计算机视觉技术方法。常用的焊缝缺陷检测多使用普通焊缝图像数据集^[22]，这往往会增加识别成本，影响识别效率。为此学者们提出了利用 X 射线焊缝图像获得理想数据集的思路^[23]，在该基础上结合图像增强技术提高了焊缝缺陷检测效率^[24]。此外，有学者提出更为先进的算法，对缺陷种类进行细分或利用分类器进行精细化缺陷识别，解决了小数据集下网络泛化性能差的问题^[25]并满足了焊缝缺陷实时检测要求^[26]。

153

表 7 焊缝缺陷识别各方法评价

154

Tab.7 Evaluation of weld defect identification methods

焊缝缺陷识别内容	作者及方法	优势	局限/展望
识别正常、重叠、飞溅和带气孔普通焊缝图像	Zhu 等 ^[22] 结合 CNN 和随机森林	采用随机森林算法进行分类，泛化能力强，鲁棒性强	所提方法存在间歇性步骤，影响了识别效率
利用多种图像识别焊缝缺陷	Mery 等 ^[23] 基于 CNN 等多种算法	发现在 X 射线图像上训练的 CNN 模型可以显著提高准确度	可能存在过拟合
识别未焊接、裂纹、气孔等 X 射线焊缝图像	Sizyakin 等 ^[24] 基于 CNN	提高了焊缝缺陷检测效率，减少了误报概率	精度有待提高
识别 11 种焊接缺陷	Yang 和 Jiang 等 ^[25] 基于统一 DNN	小数据集下具有良好的泛化能力，分类能力突出	无法自动获取焊缝缺陷候选区域
识别 8 种不同的焊缝缺陷等级	Yang 等 ^[26] 基于 YOLO v5	检测速度较快，满足钢管生产线上的实时检测要求	不同类型数据集有限

155

156 螺栓连接是另一种常用的结构连接方式，具有连接方便、可靠性强等优点，但螺栓松动、锈蚀和缺
157 失会产生一系列安全隐患。以螺杆伸长为兴趣指标，有学者基于深度学习的计算机视觉技术实现了对螺
158 杆伸长的实时检测，在此基础上，Yuan 等实现了对螺栓松动近实时地识别、分类、分割和计数^[27]。但基
159 于螺杆伸长的方法来检测螺栓初期松动是困难的，为此，有学者结合霍夫线性变换算法实现了任意角度
160 下螺栓微小松动的检测^[28]，进一步实现了螺栓松动角度的量化^[29]。为了提高深度学习模型的检测效率和
161 特征感知能力，Yang、Zhang 等利用一些新的检测算法框架^[30-31]完成螺栓连接微小松动、缺失以及损坏等
162 结构连接安全性检测任务。表 8 列举了螺栓松动自动识别相关技术方法。

163

表 8 螺栓松动识别各方法评价

164

Tab.8 Evaluation of bolt loosening identification methods

螺栓松动识别内容描述	作者及方法	优势	局限/展望
基于螺杆伸长的螺栓松开、拧紧两种检测内容	Yuan 等 ^[27] 基于 Mask R-CNN	实现了螺栓松动近实时识别、分类、像素级分割和计数	暗光下需补光检测；螺栓初期松动检测困难
基于螺母与螺杆相对转动的螺栓松动检测	Huynh 等 ^[28] 基于 R-CNN 和霍夫线性变换	实现了任意角度下的微小松动角度的螺栓检测	对未锈蚀螺栓的适用性有限
量化螺栓松动角度	Qi 等 ^[29] 基于 Faster R-CNN 并采用高斯滤波器	丰富了使用场景，量化了螺栓松动角度	精度有待提高
在合成数据集上训练以识别微小螺栓松动	Yang 等 ^[30] 基于 YOLO v3	大大减少了采集真实图像的时间和成本；实现了微小螺栓松动检测	可进一步优化高水平视角和暗光条件下的错误检测问题
实时检测螺栓损坏	Zhang 等 ^[31] 基于方位感知中心点识别网络	克服了非垂直螺栓定位不准确的问题	可能无法检测螺栓的微小松动

165

166 （2）结构表面典型损伤检测

167 在结构服役期内，裂缝被认为是结构表面损伤中最普遍的破坏类型之一。近年来，已有多种经典的
168 神经网络模型被迁移应用于裂纹检测，如：AlexNet^[32]、GoogLeNet^[33]、ResNet^[34]、Faster R-CNN 和 U-
169 Net^[35]等。此外，有学者开发了新的深度学习网络架构^[36]，显著提高了裂纹检测性能。针对图像数据质量
170 影响检测性能的问题，有研究提出了两阶段联合迁移学习方法^[37]，解决了强环境背景推断条件下的裂纹
171 识别问题，也有研究结合图像生成算法，克服了缺乏裂纹标记图像的缺点。裂纹自动检测技术也可迁移

172 应用到其余表面损伤检测中，同时实现多种缺陷的损伤检测，如：锈蚀、钢筋裸露、混凝土剥落定位检
173 测等^[38]。表 9 列举了多种结构表面典型损伤自动检测技术方法。

174 表 9 结构表面典型损伤检测各方法评价

175 Tab.9 Evaluation of typical damage detection methods for structural surfaces

结构表面典型损伤检测内 容	作者及方法	优势	局限/展望
缺陷的检测和定位	Li 等 ^[32] 提出 MSDRPN	同时实现缺陷的检测和定位	定位子网络对工作条件有要求
裂缝描绘（包括检测和分 割）	Ni 等 ^[33] 基于 CNN 开发 CDN	CDN 不需要手工设计的低级特征，完全 可由自身实现裂缝描绘	对于一些细裂缝，很容易高估其 宽度
裂纹检测、分割和综合定 量评估	Kalfarisi 等 ^[34] 结合 Faster R-CNN 与 SRFED	可对被检测的结构裂纹进行综合可视化 定量评估	进一步将进行多重缺陷的检测和 分割
像素级裂纹检测	Zhang 等 ^[35] 基于 CrackUnet	使用新的损失函数提高模型性能，最高 训练精度优于 FCN 模型	无法从强光背景图像中准确检测 裂缝
道路路面自动裂缝检测	Fan 等 ^[36] 提出 U-HDN	利用具有多个上下文信息的裂纹特征实 现端到端高精度的裂纹检测	神经网络结构复杂，增加了计算 成本
混凝土大坝裂缝分类和弱 监督定位	Li 等 ^[37] 提出两阶段联合 迁移学习	实现以较高的精确率有效识别结构裂缝 且可实现弱监督定位	或不能具体地量化裂缝

176

177 近年来，相关研究学者提出了基于深度学习的三维重建技术，该技术现已在土木工程结构损伤检测
178 领域得到了快速发展与应用。Zhang 等提出了一种基于卷积神经网络的 CrackNet^[39]，可实现三维裂缝像
179 素级别的自动检测，在此基础上有学者提出了 CrackNet II^[40]、CrackDN 等衍生类架构，取得了更好的
180 整体性能。此外，三维重建在生成密集点云方面具有突出优势，其可在建筑立面缺陷、路面裂缝检测、
181 建筑拆除碎片体积评估等方面发挥重要作用^[41-42]。对于桥梁工程结构而言，三维重建技术已发展为可集
182 成小型无人机、移动机器人平台来生成桥梁三维模型的状态管理系统，在此基础上，已有学者结合同步
183 定位和映射方法（SLAM）实现了三维层面上桥梁损伤自动检测和量化^[43]。此外，三维损伤数据还可与
184 桥梁 BIM 模型的缺陷对应^[44]。表 10 列举了一般土木工程结构三维损伤自动检测技术方法。

185 表 10 一般土木工程结构三维损伤检测各方法评价

186 Tab.10 Evaluation of three dimensional damage detection methods for general civil engineering structures

一般土木工程结构三维 损伤检测内容	作者及方法	优势	局限/展望
三维裂缝像素级检测	Zhang 等 ^[40] 提出 CrackNet II	与原始 CrackNet ^[39] 相比，性能提高较多，可 检测到更多的细小或细线裂纹	可能存在像素级不连续问题
基于 3D 点云的混凝土人 行道自动缺陷检测	Jiang 等 ^[41] 基于 U-Net 和 3D 点云	可准确全面地检测垂直位移情况以识别潜在的 路面绊倒危险	进一步可考虑应用于更多复杂道路 场景
混凝土大坝损伤检测和 可视化	Zhao 等 ^[42] 基于 YOLOv5s-HSC	较好应对复杂背景和模糊边界问题；新的投影 方法对损伤识别和定位更加直观准确	可能会出现多映射效应；进一步可 优化拍摄路径
自动检测和量化桥梁混 凝土的分层和剥落情况	MCLAUGHLIN 等 ^[43] 基于 CNN 和 SLAM	结合 CNN 标记图像生成具有彩色和语义标记的 三维结构地图	暂不适用于其他缺陷
3D BIM 桥梁损坏检测	ISAILOVIĆ 等 ^[44] 基于 CNN 并利用	提出结合对 IFC 模型的语义丰富将损伤组件结 合到 BIM 中	基于点云生成的图片可能会造成信 息缺失

本节主要总结了近年来在役结构损伤检测中基于深度学习的计算机视觉技术的应用，可以看出，目前研究缺乏具有可靠文本信息或可视化图像的损伤自动评价体系。未来可考虑集成损伤分析，进一步组织成可供计算机和人理解的可靠文本信息或可视化图形信息，实现“输入图像+输出图文报告”的机制，以提供更多的信息解释和趋势判断能力。

4.3 灾后结构损伤评估

灾后建筑物倒塌损坏识别是实现快速救援及合理重建的关键。随着光学成像、合成孔径雷达等遥感技术的迅速发展及应用，基于深度学习的计算机视觉技术被广泛用于评估灾后建筑物的破坏情况。

(1) 基于卫星影像识别

以卷积神经网络为代表的深度学习算法被广泛应用于灾后备卫星影像物体检测和识别，基于 CNN 框架将卫星遥感图像用于场景分类是可行的，在此基础上，有学者依靠特征提取后处理方法实现了建筑物定量和定性检测^[45]，结合样本平衡方法解决了灾后建筑物识别样本失衡问题^[46]。此外，也有学者结合超参数优化算法、使用深度网络架构以及轻量化网络模型提高了检测精度和速度^[47-48]。近年来，随着超像素分割和特征融合等技术的发展，已有研究同时实现了灾后建筑物多种损伤等级的检测^[49]，克服了特征选择困难、图像分割碎片化等问题。表 11 列举了基于卫星影像的结构灾后整体损伤自动评估技术方法。

表 11 基于卫星影像的结构灾后整体损伤评估各方法评价

Tab.11 Evaluation of global damage assessment methods for post-disaster structures based on satellite images

灾后建筑物损伤评估内容	作者及方法	优势	局限/展望
基于卫星图像提取建筑物特征	Yang 等 ^[45] 基于 DCNN 及特征提取后处理	定量和定性检测能力优异	性能取决于预测的建筑区域和边界
从震后遥感图像中区分倒塌和未倒塌建筑物	Ji 等 ^[46] 基于 SqueezeNet	在一定程度上解决了样本不平衡问题	对小型建筑物而言，假阳性随着真阳性增加而增加
从震后遥感图像中定位倒塌建筑物	Ma 等 ^[47] 基于改进的 YOLO v3	实现了模型轻量化，提高了检测速度和精度	由于数据量及样本标签准确性的限制会产生误差
超高分辨率遥感图像中震后损坏建筑物检测	Wang 等 ^[48] 基于改进的 FCN 和 OCR 技术等	增强了特征表示，减少了漏检误检	进一步可将该方法应用于其他场景
建筑物多种损坏评估	Shen 等 ^[49] 基于 U-Net 建立了横向特征融合模型	更好地探索灾前和灾后图像之间的相关性	增加了计算成本

(2) 基于航拍影像识别

除了卫星影像外，航拍影像也被广泛应用于灾后救援及重建领域。有学者提出了一种基于航空图像的震区倒塌建筑物自动识别技术，基于此思路，Miura 等利用航拍影像训练 CNN 模型成功实现了不同损伤等级识别^[50]，验证了该技术的应用潜力，值得一提的是，在此基础上引入图像的 3D 特征可以显著提高网络模型的泛化能力和准确率。然而，由于航拍影像下倒塌建筑物形状多样、背景复杂，检测会受到传统网络特征提取能力的限制，为此，有学者引入可变形卷积层提高对任意形状倒塌建筑物的适应性^[51]，或使用 k-means 算法优化锚点的数量和尺寸大小^[52]，这些做法有效缓解了误检和漏检问题。此外，通过结合在线硬样本挖掘技术提高信息利用率^[53]或使用新的网络架构^[54]可获得更好的分类结果。表 12

列举了基于航拍影像的结构灾后整体损伤自动评估技术方法。

表 12 基于航拍影像的结构灾后整体损伤评估各方法评价

Tab.12 Evaluation of global damage assessment methods for post-disaster structures based on aerial images

灾后建筑物损伤评估内容	作者及方法	优势	局限/展望
识别倒塌、未倒塌和蓝色防水布覆盖的建筑物	Miura 等 ^[50] 基于利用航拍影像训练的 CNN	可识别倒塌、未倒塌和蓝色防水布覆盖的建筑物	很难正确分类存在如太阳能电池板等小物体建筑物
基于无人机图像的震后倒塌建筑物检测	Ding 等 ^[51] 基于改进的 NMS 和 Faster R-CNN	有效抑制了冗余边界框，提高了对任意形状倒塌建筑物的适应性	进一步将研究模糊背景下的检测问题
识别航拍图像中背景相似的受损建筑物	Shi 等 ^[52] 基于改进的 YOLO v4	增强了对背景相似倒塌建筑物的检测能力	漏检和误检问题依然存在，模型训练时间较长
震后受损建筑物多级损坏程度检测	Zhan 等 ^[53] 基于改进的 Mask R-CNN 并结合在线硬样本挖掘技术	增强了检测具有相似特征小物体的能力	四种损伤类别的总体分类准确率有待提高
飓风后建筑物损坏评估	Cheng 等 ^[54] 基于堆叠卷积神经网络架构	考虑了损坏级别类别的顺序性	建筑物的定位精度和分类精度不高

（3）结合多种分辨率或多模式灾害影像识别

以上方法大都使用单一分辨率作为输入，而结合多种途径（机载、卫星、航拍图像）或者多种模式（多分辨率、多时相、多感官）获取的高分辨率图像来训练网络可以提高算法的检测精度。较为典型的是 Rudne 等提出的融合多分辨率、多时相和多感官卫星图像训练 CNN 的思想^[55]。表 13 列举了结合多种分辨率或多模式灾害影像的灾后自动识别技术方法。

表 13 结合多种分辨率或多模式灾害影像的灾后识别各方法评价

Tab.13 Evaluation of post-disaster identification methods combined with multi-resolution or multi-modal disaster images

结合多种分辨率或多模式灾害影像的灾后识别内容	作者及方法	优势	局限/展望
生成洪水地图并快速分割被淹建筑物	Rudner 等 ^[55] 基于 Multi3Net	融合多分辨率、多时相和多感官卫星图像数据	将适用于更多灾害识别场景
灾后建筑物损伤检测	Duarte 等 ^[56] 基于 CNN	通过融合分辨率特征，显示出更好的准确性和定位能力	没有关于图像分类任务中每个分辨率级别个体贡献的信息
灾后建筑结构损坏检测	Nex 等 ^[57] 基于改进 CNN	使用涵盖不同位置、空间分辨率的异构和大型数据集，提升了检测精度	可进一步结合使用生成对抗网络缓解由于不同位置缺乏数据而导致的问题
灾后建筑物多级损坏测绘	Adriano 等 ^[58] 基于 Attention U-Net	开发了开源多模态和多时态建筑物损伤数据集	未来将扩展数据集以涵盖多种空间分辨率的遥感数据

从本节可以看出，基于深度学习的计算机视觉技术以高效、低成本、可快速获取感兴趣图像特征等优势正在结构灾后整体损伤评估领域发挥着愈发重要的作用，但不可否认的是，该技术在处理灾后地面现场细致调查时存在精度不足、无法探测结构内部破坏等问题。为此，未来可考虑与工业自动化监测设备集成，在解决灾后细致调查问题的同时使结构健康监测更加智能化和高效便捷。

5 结语

230 本文重点回顾和总结了基于深度学习的计算机视觉技术在土木工程建设全生命周期中的应用。首先
231 利用 CiteSpace 软件对检索到的相关文献进行科学分析与统计；其次简要阐述了计算机视觉技术的发展历
232 程和基本原理，归纳了构建深度学习数据集所涉及的数据获取、数据处理以及数据标注方法；最后回顾
233 了近年来基于深度学习的计算机视觉技术在施工现场安全管理、在役结构损伤检测以及灾后结构损伤评
234 估领域中的应用研究。通过对相关工程应用发展脉络的了解，可以找到研究共性和迁移点，从而启发研
235 究者在未来开发更多类型的实际应用，更好地辅助工程任务。下面列出了可以拓展的应用方向。

236 (1) 进一步完善预制构件的数据集类型，使其更好地服务于装配式建筑施工进程。

237 (2) 引入视频结构化技术，建立施工场景对象动态关联网络。

238 (3) 将可见损伤检测结果与内容分析相结合，进一步组织成可供计算机和人理解的文本信息或可视
239 化图形信息，以提供更多的信息解释和趋势判断能力。

240 (4) 集成自动化监测车以及机器人，使结构健康监测更加智能化和高效便捷，进一步推动基于深度
241 学习的计算机视觉技术在土木工程领域的实际部署及产业化发展。

242 目前，基于深度学习的计算机视觉技术正在土木工程建设全生命周期中发挥重要作用。未来，通过
243 构建高质量数据集、引入新的图像处理技术以及改进深度学习算法框架等方法，基于深度学习的计算机
244 视觉技术将发挥出更丰富的土木工程应用潜力。

245

246 作者贡献说明：

247 方成：论文想法提出，论文核对

248 于盛鑫：论文框架设计，论文撰写与修改

249 李永刚：论文构思、修改与核对

250 贾王龙：论文构思与核对

251 杨鹏博：论文构思与核对

252 杨欣悦：论文总体规划，论文撰写与核对

253

254 参考文献：

- 255 [1] CHEN C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature [J]. Journal of the
256 American Society for Information Science and Technology, 2006, 57(3): 359.
- 257 [2] FANG C, PING Y, GAO Y, et al. Machine learning-aided multi-objective optimization of structures with hybrid braces-
258 Framework and case study [J]. Engineering Structures, 2022, 269: 114808.
- 259 [3] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition [J]. Neural Computation,
260 1989, 1(4): 541.
- 261 [4] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial
262 networks [J]. arXiv preprint arXiv:1511.06434, 2015.
- 263 [5] YE X W, JIN T, LI Z X, et al. Structural crack detection from benchmark data sets using pruned fully convolutional networks
264 [J]. Journal of Structural Engineering, 2021, 147(11): 04721008.
- 265 [6] FANG Q, LI H, LUO X, et al. Detecting non-hardhat-use by a deep learning method from far-field surveillance videos [J].
266 Automation in Construction, 2018, 85: 1.
- 267 [7] SHEN J, XIONG X, LI Y, et al. Detecting safety helmet wearing on construction sites with bounding-box regression and deep
268 transfer learning [J]. Computer-Aided Civil and Infrastructure Engineering, 2021, 36(2): 180.
- 269 [8] WANG Z, WU Y, YANG L, et al. Fast personal protective equipment detection for real construction sites using deep learning

approaches [J]. *Sensors*, 2021, 21(10): 3478.

[9] FANG W, DING L, LUO H, et al. Falls from heights: A computer vision-based approach for safety harness detection [J]. *Automation in Construction*, 2018, 91: 53.

[10] XIONG R, TANG P. Machine learning using synthetic images for detecting dust emissions on construction sites [J]. *Smart and Sustainable Built Environment*, 2021, 10(3): 487.

[11] DING L, FANG W, LUO H, et al. A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory [J]. *Automation in Construction*, 2018, 86: 118.

[12] LUO X, LI H, CAO D, et al. Recognizing diverse construction activities in site images via relevance networks of construction-related objects detected by convolutional neural networks [J]. *Journal of Computing in Civil Engineering*, 2018, 32(3): 04018012.

[13] FANG W, DING L, ZHONG B, et al. Automated detection of workers and heavy equipment on construction sites: A convolutional neural network approach [J]. *Advanced Engineering Informatics*, 2018, 37: 139.

[14] KIM H, KIM H, HONG Y W, et al. Detecting construction equipment using a region-based fully convolutional network and transfer learning [J]. *Journal of Computing in Civil Engineering*, 2018, 32(2): 04017082.

[15] LUO H, LIU J, FANG W, et al. Real-time smart video surveillance to manage safety: A case study of a transport mega-project [J]. *Advanced Engineering Informatics*, 2020, 45: 101100.

[16] ZENG T, WANG J, CUI B, et al. The equipment detection and localization of large-scale construction jobsite by far-field construction surveillance video based on improving YOLOv3 and grey wolf optimizer improving extreme learning machine [J]. *Construction and Building Materials*, 2021, 291: 123268.

[17] CHEN C, ZHU Z, HAMMAD A. Automated excavators activity recognition and productivity analysis from construction site surveillance videos [J]. *Automation in Construction*, 2020, 110: 103045.

[18] LIN Z H, CHEN A Y, HSIEH S H. Temporal image analytics for abnormal construction activity identification [J]. *Automation in Construction*, 2021, 124: 103572.

[19] LI J, ZHOU G, LI D, et al. Recognizing workers' construction activities on a reinforcement processing area through the position relationship of objects detected by faster R-CNN [J]. *Engineering, Construction and Architectural Management*, 2023, 30(4): 1657-1678.

[20] WANG Z, ZHANG Q, YANG B, et al. Vision-based framework for automatic progress monitoring of precast walls by using surveillance videos during the construction phase [J]. *Journal of Computing in Civil Engineering*, 2021, 35(1): 04020056.

[21] XIAO B, WANG Y, KANG S C. Deep learning image captioning in construction management: a feasibility study [J]. *Journal of Construction Engineering and Management*, 2022, 148(7): 04022049.

[22] ZHU H, GE W, LIU Z. Deep learning-based classification of weld surface defects [J]. *Applied Sciences*, 2019, 9(16): 3312.

[23] MERY D, ARTETA C. Automatic defect recognition in x-ray testing using computer vision [C] // 2017 IEEE winter conference on applications of computer vision (WACV). Santa Rosa: IEEE, 2017: 1026-1035.

[24] SIZYAKIN R, VORONIN V, GAPON N, et al. Automatic detection of welding defects using the convolutional neural network [C] // Automated Visual Inspection and Machine Vision III. SPIE, 2019, 11061: 93-101.

[25] YANG L, JIANG H. Weld defect classification in radiographic images using unified deep neural network with multi-level features [J]. *Journal of Intelligent Manufacturing*, 2021, 32(2): 459.

[26] YANG D, CUI Y, YU Z, et al. Deep learning based steel pipe weld defect detection [J]. *Applied Artificial Intelligence*, 2021, 35(15): 1237.

[27] YUAN C, CHEN W, HAO H, et al. Near real-time bolt-loosening detection using mask and region-based convolutional neural network [J]. *Structural Control and Health Monitoring*, 2021, 28(7): e2741.

[28] HUYNH T C, PARK J H, JUNG H J, et al. Quasi-autonomous bolt-loosening detection method using vision-based deep learning and image processing [J]. *Automation in Construction*, 2019, 105: 102844.

[29] QI Y, LI P, XIONG B, et al. A two-step computer vision-based framework for bolt loosening detection and its implementation on a smartphone application [J]. *Structural Health Monitoring*, 2022: 147592172110499.

[30] YANG X, GAO Y, FANG C, et al. Deep learning-based bolt loosening detection for wind turbine towers [J]. *Structural Control and Health Monitoring*, 2022, 29(6): e2943.

316 [31] ZHANG Y, YUEN K V. Bolt damage identification based on orientation-aware center point estimation network [J]. Structural
317 Health Monitoring, 2022, 21(2): 438.

318 [32] LI R, YUAN Y, ZHANG W, et al. Unified vision-based methodology for simultaneous concrete defect detection and
319 geolocalization [J]. Computer-Aided Civil and Infrastructure Engineering, 2018, 33(7): 527.

320 [33] NI F T, ZHANG J, CHEN Z Q. Pixel-level crack delineation in images with convolutional feature fusion [J]. Structural Control
321 and Health Monitoring, 2019, 26(1): e2286.

322 [34] KALFARISI R, WU Z Y, SOH K. Crack detection and segmentation using deep learning with 3D reality mesh model for
323 quantitative assessment and integrated visualization [J]. Journal of Computing in Civil Engineering, 2020, 34(3): 04020010.

324 [35] ZHANG L, SHEN J, ZHU B. A research on an improved Unet-based concrete crack detection algorithm [J]. Structural Health
325 Monitoring, 2021, 20(4): 1864.

326 [36] FAN Z, LI C, CHEN Y, et al. Automatic crack detection on road pavements using encoder-decoder architecture [J]. Materials,
327 2020, 13(13): 2960.

328 [37] LI Y, BAO T, XU B, et al. A deep residual neural network framework with transfer learning for concrete dams patch-level crack
329 classification and weakly-supervised localization [J]. Measurement, 2022, 188: 110641.

330 [38] JIANG Y, PANG D, LI C. A deep learning approach for fast detection and classification of concrete damage [J]. Automation in
331 Construction, 2021, 128: 103785.

332 [39] ZHANG A, WANG K C P, LI B, et al. Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-
333 learning network [J]. Computer-Aided Civil and Infrastructure Engineering, 2017, 32(10): 805.

334 [40] ZHANG A, WANG K C P, FEI Y, et al. Deep learning-based fully automated pavement crack detection on 3D asphalt surfaces
335 with an improved CrackNet [J]. Journal of Computing in Civil Engineering, 2018, 32(5): 04018041.

336 [41] JIANG Y, HAN S, LI D, et al. Automatic concrete sidewalk deficiency detection and mapping with deep learning [J]. Expert
337 Systems with Applications, 2022, 207: 117980.

338 [42] ZHAO S, KANG F, LI J. Concrete dam damage detection and localisation based on YOLOv5s-HSC and photogrammetric 3D
339 reconstruction [J]. Automation in Construction, 2022, 143: 104555.

340 [43] MCLAUGHLIN E, CHARRON N, NARASIMHAN S. Automated defect quantification in concrete bridges using robotics and
341 deep learning [J]. Journal of Computing in Civil Engineering, 2020, 34(5): 04020029.

342 [44] ISAILOVIĆ D, STOJANOVIC V, TRAPP M, et al. Bridge damage: Detection, IFC-based semantic enrichment and visualization
343 [J]. Automation in Construction, 2020, 112: 103088.

344 [45] YANG H, XU M, CHEN Y, et al. A postprocessing method based on regions and boundaries using convolutional neural networks
345 and a new dataset for building extraction [J]. Remote Sensing, 2022, 14(3): 647.

346 [46] JI M, LIU L, BUCHROITHNER M. Identifying collapsed buildings using post-earthquake satellite imagery and convolutional
347 neural networks: A case study of the 2010 Haiti earthquake [J]. Remote Sensing, 2018, 10(11): 1689.

348 [47] MA H, LIU Y, REN Y, et al. Detection of collapsed buildings in post-earthquake remote sensing images based on the improved
349 YOLOv3 [J]. Remote Sensing, 2019, 12(1): 44.

350 [48] WANG C, QIU X, HUAN H, et al. Earthquake-damaged buildings detection in very high-resolution remote sensing images based
351 on object context and boundary enhanced loss [J]. Remote Sensing, 2021, 13(16): 3119.

352 [49] SHEN Y, ZHU S, YANG T, et al. Cross-directional feature fusion network for building damage assessment from satellite imagery
353 [J]. arXiv preprint arXiv:2010.14014, 2020.

354 [50] MIURA H, ARIDOME T, MATSUOKA M. Deep learning-based identification of collapsed, non-collapsed and blue tarp-
355 covered buildings from post-disaster aerial images [J]. Remote Sensing, 2020, 12(12): 1924.

356 [51] DING J, ZHANG J, ZHAN Z, et al. A precision efficient method for collapsed building detection in post-earthquake UAV images
357 based on the improved NMS algorithm and faster R-CNN [J]. Remote Sensing, 2022, 14(3): 663.

358 [52] SHI L, ZHANG F, XIA J, et al. Identifying damaged buildings in aerial images using the object detection method [J]. Remote
359 Sensing, 2021, 13(21): 4213.

360 [53] ZHAN Y, LIU W, MARUYAMA Y. Damaged building extraction using modified mask R-CNN model using post-event aerial
361 images of the 2016 Kumamoto earthquake [J]. Remote Sensing, 2022, 14(4): 1002.

- 362 [54] CHENG C S, BEHZADAN A H, NOSHADRAVAN A. Deep learning for post-hurricane aerial damage assessment of buildings
363 [J]. Computer-Aided Civil and Infrastructure Engineering, 2021, 36(6): 695.
- 364 [55] RUDNER T G J, RUßWURM M, FIL J, et al. Multi3net: segmenting flooded buildings via fusion of multiresolution, multisensor,
365 and multitemporal satellite imagery [C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2019, 33(01): 702-709.
- 366 [56] DUARTE D, NEX F, KERLE N, et al. Multi-resolution feature fusion for image classification of building damages with
367 convolutional neural networks [J]. Remote Sensing, 2018, 10(10): 1636.
- 368 [57] NEX F, DUARTE D, TONOLO F G, et al. Structural building damage detection with deep learning: Assessment of a state-of-
369 the-art CNN in operational conditions [J]. Remote Sensing, 2019, 11(23): 2765.
- 370 [58] ADRIANO B, YOKOYA N, XIA J, et al. Learning from multimodal and multitemporal earth observation data for building
371 damage mapping [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 175: 132.
- 372