

混凝土结构外观质量缺陷识别算法性能测评与分析

郭俊熙¹, 潘鹏^{1,2}, 林佳瑞^{1,2,*}

(1, 清华大学土木工程系, 北京 100084; 2, 住建部数字建造与孪生重点实验室, 北京 100084)

【摘要】混凝土结构作为建筑工程中重要的组成部分, 其安全性和耐久性直接影响建筑的整体质量和使用寿命。然而, 传统的混凝土缺陷识别方法依赖人工, 耗时费力, 亟需探索缺陷自动识别方法。本研究旨在对比不同目标检测算法的性能特点与适用场景, 为混凝土结构外观质量缺陷检测建立更加高效、准确的识别方法。因此, 本研究首先构建了一个包含 1085 张图片及 1760 个缺陷实例的混凝土外观质量缺陷图像数据集; 然后, 选择了 YOLOv5、YOLOv9、SSD 和 EfficientNet 四种目标检测模型进行训练和测试; 最后, 通过对比各模型训练的速度与效率、精确率、召回率与 mAP 等指标对模型进行性能对比与分析。结果表明, 针对混凝土结构外观缺陷识别问题 YOLOv9 在四种模型具有更好的鲁棒性与准确率, mAP 可达 0.71。有关研究为开发施工质量缺陷识别方法及装备提供了有益指导, 并服务施工巡检效率与建造质量提升。

【关键词】混凝土结构; 外观质量缺陷; 目标检测; YOLOv9; 数据集

1 引言

混凝土结构在现代建筑工程中发挥着关键作用, 其安全性和耐久性直接关系到建筑的整体质量和使用寿命。随着城市化和基础设施建设的快速推进, 对混凝土结构安全性的检测变得愈发重要^[1]。

然而, 传统的人工检测方法耗时费力, 且准确性有限。随着计算机视觉和数字孪生技术的发展, 自动化缺陷识别方法逐渐受到关注^[2]。目标检测算法作为计算机视觉的核心技术, 被广泛应用于混凝土外观质量缺陷识别中。因此, 本研究旨在通过训练和测试不同目标检测算法, 评估其在混凝土外观质量缺陷识别中的性能与适用性, 为工程实践提供高效、准确的模型, 提升施工巡检效率和建造质量。

2 研究现状

2019 年以后, 涉及深度学习的混凝土结构图像缺陷自动识别算法成为一个流行话题^[3], 如 Fast R-CNN、Faster R-CNN、SSD、YOLO 等一系列能够快速实现物体识别与定位的目

【基金项目】国家重点研发计划项目(2023YFC3805802)

【作者简介】林佳瑞(1987-), 男, 副研究员。主要研究方向智能审图、智能巡检。E-mail: lin611@tsinghua.edu.cn

标检测网络出现 Cha 等^[4]开发了一个包含 2366 张 500×375 像素的图像的数据库, 然后用 Faster R-CNN 在数据库中进行训练与测试; Wang 等^[5]基于 ResNet101 框架的 Faster R-CNN 模型识别历史砌体结构的两类损伤——风化和剥落; Chen 等^[6]基于金属表面缺陷的特点采用特征金字塔网络对 Faster R-CNN 框架进行优化, 提高了其精度; Ghosh^[7]等将 Faster R-CNN 与 Inception v2、ResNet-50 等特征提取网络结合使用进行震后建筑表面裂纹、等四种缺陷的识别工作。Maeda 等^[8]用安装在汽车上的手机拍摄 9053 张道路缺陷图像, 15435 个缺陷实例中包含横向裂纹, 龟裂等八种类型的缺陷, 他们使用 SSD 准确率均较高。Cao 等^[9]将道路表面的裂缝缺陷细分为五个等级, 然后以 9493 张图像训练和验证 SSD 与 Faster R-CNN 的 8 个识别模型。Yin 等^[10]从闭路电视中提取管道内部的缺陷图像数据, 共有 4056 个样本的数据集对于 YOLOv3 进行训练。Jiang 等^[11]制作了包含 5000 张图片的建筑表面缺陷数据集, 包括裂纹、斑点等类型, 他们用改进的 YOLO 网络以及 MobileNetV3-SSD 在数据集上进行训练, 并且比较了不同 EfficientNet and MobileNet 基准网络的效果。Yu 等^[12]用无人机收集了 376 张桥梁裂缝缺陷的图像, 并且他们利用焦损失对损失函数进行 YOLOv4 的网络优化然后在数据集上进行训练, 改进后的 YOLOv4-fpm 相比 YOLOv4 准确率有所提高。

3 数据集构建

本研究构建了一个包含四种常见混凝土表面缺陷的综合数据集, 共包含 1085 张图像和 1760 个缺陷实例。数据集来源于互联网开源数据集、互联网爬取的图像及建筑工地拍摄的图像, 保证了数据的多样性和真实性。数据集中包含 611 个露筋缺陷、598 个剥落缺陷、409 个裂纹缺陷和 142 个蜂窝缺陷, 确保了数据集的平衡性, 有助于提升模型训练效果和实际应用的准确性。此外, 数据集涵盖了路面、桥梁、震损结构和施工阶段等多个场景, 增强了数据集的泛化性, 利于提升模型的鲁棒性。

4 缺陷识别模型对比与选择

本实验选取 YOLOv5、YOLOv9、SSD 和 EfficientNet 四种模型进行训练, 并通过常见的目标检测模型评估指标对各模型进行性能对比, 最终选出最优模型用于实际工程缺陷识别。

4.1 模型对比方法

为了验证各模型的性能, 本研究采用以下步骤:

(1) 数据集划分:

将构建的数据集随机划分为训练集、验证集和测试集, 比例为 7:2:1, 以确保模型训练、验证和测试的公平性。

(2) 模型训练:

在相同的数据集上分别训练 YOLOv5、YOLOv9、SSD 和 EfficientNet 四种模型。训练过程中, 采用适合各模型的训练参数和优化器设置, 以确保结果的可比性。

(3) 性能评估:

为了科学评估各模型的性能,我们选择了精确率 (Precision)、召回率 (Recall) 和平均精度 (Average Precision, AP) 作为评价指标^[13]。

(4) 对比与分析:

对比各模型在精确率、召回率和 mAP 等指标上的表现,分析其优劣。通过对比分析,识别出在不同缺陷类型上的最佳模型,结合实际应用需求,选出综合性能最优的模型用于实际工程缺陷识别。

通过上述步骤,本文可对比各模型在不同缺陷类型上的表现,分析其优劣,最终选择在精确率、召回率和 mAP 等指标上表现最佳的模型用于实际工程应用。

4.2 模型构建与训练

本实验选取 YOLOv5、YOLOv9、SSD 以及 EfficientNet 四种模型进行训练,每种模型均在第三节中构建出的数据集上进行训练,以确保对比结果的公平性。表 1 展示了各模型训练时长、优化器、训练时代数等参数的对比,然后在 4.3 节中会通过各类缺陷的精确率、召回率和 mAP 等指标进一步进行模型性能的对比与分析。

表 1 各模型基本参数与训练策略

模型	输入尺寸	批次	学习率	优化器	训练时间	训练时代	训练硬件
YOLOv5	640x640	16	0.01	Adam	5h	500epoch	NVIDIA GeForce RTX 3090
YOLOv9	640x640	16	0.01	Adam	4h	270epoch	NVIDIA GeForce RTX 3090
SSD	300x300	16	0.001	SGD	8h	220epoch	NVIDIA GeForce RTX 2080Ti
EfficientDet	512x512	16	0.01	Adam	1d	150epoch	NVIDIA GeForce RTX 2080Ti

YOLOv9 的训练时间为 4 小时,是所有模型中训练速度最快者,意味着它能够高效迭代和调整;YOLOv5 训练时间为 5 小时,二者在训练时间指标上相比于 SSD 的 8 小时和 EfficientNet 的一天的训练时间优势非常明显。

4.3 模型对比与测试

在精确率方面,YOLOv9 在综合所有缺陷的表现最佳且均衡,特别是在蜂窝和裂缝识别上,其精确率分别达到 0.955 和 0.904。相较之下,SSD 在精确率方面表现很差,例如其在蜂窝缺陷上的精确率仅为 0.0614,表明其在实际应用中可能会产生大量误报。EfficientNet 虽然在裂缝和露筋识别上的精确率较高,但由于训练时间较长,不适合需要快速响应的实际应用。

在召回率方面,SSD 在综合所有缺陷的表现最佳,YOLOv9 与 YOLOv5 表现相近。SSD 的高召回率意味着其能够识别大多数缺陷,但同时也伴随着较高的误报率。YOLOv9 尽管召回率稍低,但其精确率较高。EfficientDet 召回率方面表现一般,适合需要平衡精确率和召回率的场景。

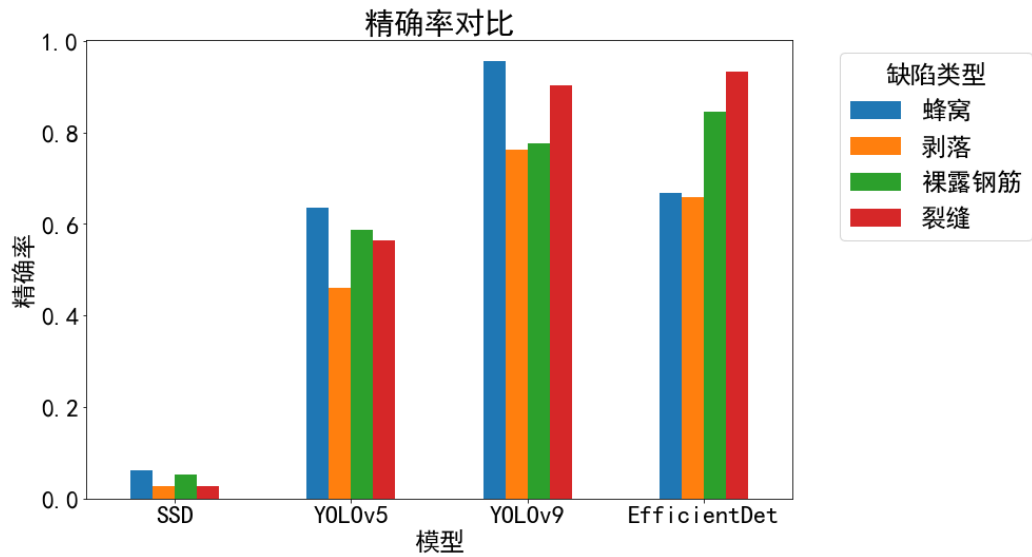


图 1 各模型精确率对比

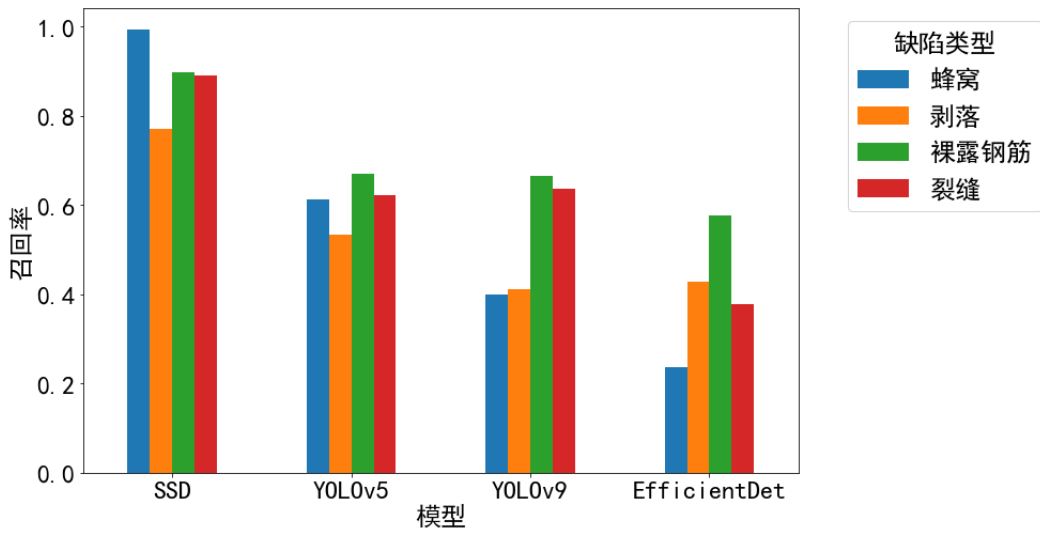


图 2 各模型召回率对比

表 2 各模型训练类别 AP 以及 mAP

模型	各类别表现				mAP0.5
	裂缝	露筋	剥落	蜂窝	
YOLOv5	0.63	0.642	0.477	0.601	0.588
YOLOv9	0.797	0.736	0.566	0.738	0.71
EfficientNet	0.5595	0.7083	0.4626	0.5323	0.5657
SSD	0.883	0.535	0.757	0.773	0.737

SSD 和 YOLOv9 在裂缝、露筋、剥落和蜂窝四个缺陷类别上均表现出较高的 AP 值，二者整体 mAP 达到 0.732 和 0.71，显著高于另外两个模型。

通过分析各指标，发现 SSD 尽管精确率最低，但召回率最高，适用于漏检容忍度低但可接受较多误报的场景。EfficientNet 在精确率和召回率上表现中等，但训练时间最长，适合对训练时间不敏感且只需要中等精确率的应用。YOLOv5 在各项指标上表现适中，适用于需要平衡训练时间和精确度的场景。由于工程实际需要对现场进行全面扫描，会产生海量图片数据，且工程追求经济性，会追求以更小容量的服务器完成任务需求，所以工程实际中混凝土结构表观质量缺陷识别需要识别精确率较高、平均精度较好的模型，防止误报较多进一步生成更多图片数据给服务器带来较大荷载。所以综合考虑识别时间、训练时间、精确率、mAP 等因素，YOLOv9 表现最佳，是综合性能最优的模型。

表 3 各模型综合对比

模型	平均精确率	平均召回率	mAP0.5	训练时间
YOLOv5	0.562	0.609	0.588	5h
YOLOv9	0.849	0.528	0.71	4h
EfficientNet	0.778	0.405	0.5657	1d
SSD	0.0413	0.887	0.737	8h

使用 YOLOv9 模型基于数据集测试集图片进行实验，测试效果如图 3 所示——该图片的标注情况可以作为参照。从两者对比可以看出 YOLOv9 在本次测试中识别露筋、剥落缺陷方面表现良好。例如，在多张含有露筋的图片中，模型均能识别并标记出相应缺陷，且置信度值如 1.0、0.7 等显示出高可靠性。但从本次测试效果观察，模型对于露筋具体范围的检测以及对于剥落和蜂窝的区分能力仍存在不足，需后续继续研究改进。

5 总结

本研究首先构建了一个包含 1085 张图片和 1760 个缺陷实例的混凝土表观质量缺陷图像数据集，并选择了 YOLOv5、YOLOv9、SSD 和 EfficientNet 四种目标检测模型进行训练和测试。通过对比各模型在训练速度、精确率、召回率和 mAP 等指标上的表现，综合评定 YOLOv9 为最优模型。结果表明，YOLOv9 不仅在精确率和鲁棒性方面表现优异，而且在训练时间和实时性上也具有显著优势，适用于实际工程应用。

本文构建的多样化数据集，以及对多种目标检测模型的系统性比较，为混凝土结构质量检测领域提供了一定的数据基础，为开发施工质量缺陷识别方法及装备提供了有益指导。未来的研究可以进一步优化 YOLOv9 模型，以提升其对复杂缺陷的识别能力，并将其应用扩展到其他类型的建筑结构质量检测中，进一步推动混凝土表观质量缺陷识别的自动化和智能化发展。

参 考 文 献

- [1] Park, J., Cai, H., & Perissin, D. (2018). Bringing information to the field: automated photo registration and 4D BIM. *Journal of computing in civil engineering*, 32(2), 04017084.
- [2] 周绍杰, 潘鹏, 顾栋炼, 林佳瑞* (2023). 融合 BIM 与三维重建的施工进度数字孪生跟踪方法初探. 第九届全国 BIM 学术会议论文集, 453-457. 中国建筑工业出版社. 中国, 西安.
- [3] Ma, Z., Liu, Y., & Li, J. (2023). Review on automated quality inspection of precast concrete components. *Automation in Construction*, 150, 104828.
- [4] Y.-J. Cha, W. Choi, G. Suh, S. Mahmoudkhani, O. Büyük Öztürk, Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types, *Computer-Aided Civil and Infrastructure Engineering* 33(9) (2018) 731 - 747, <https://doi.org/10.1111/mice.12334>.
- [5] N.Wang, X. Zhao, P. Zhao, Y. Zhang, Z. Zou, J. Ou, Automatic damage detection of historic masonry buildings based on mobile deep learning, *Automation in Construction* 103 (2019) 53 - 66, <https://doi.org/10.1016/j.autcon.2019.03.003>.
- [6] K. Chen, Z. Zeng, J. Yang, A deep region-based pyramid neural network for automatic detection and multi-classification of various surface defects of aluminum alloys, *Journal of Building Engineering* 43 (2021), 102523, <https://doi.org/10.1016/j.jobe.2021.102523>.
- [7] T. Ghosh Mondal, M.R. Jahanshahi, R.T. Wu, Z.Y. Wu, Deep learning-based multiclass damage detection for autonomous post-disaster reconnaissance, *Structural Control and Health Monitoring* 27 (4) (2020), <https://doi.org/10.1002/stc.2507n/a-n/a>.
- [8] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiya, H. Omata, Road damage detection and classification using deep neural networks with smartphone images: road damage detection and classification, *Computer-Aided Civil and Infrastructure Engineering* 33 (12) (2018) 1127 - 1141, <https://doi.org/10.1111/mice.12387>.
- [9] M.-T. Cao, Q.-V. Tran, N.-M. Nguyen, K.-T. Chang, Survey on performance of deep learning models for detecting road damages using multiple dashcam image resources, *Advanced Engineering Informatics* 46 (2020), 101182, <https://doi.org/10.1016/j.aei.2020.101182>.
- [10] X. Yin, Y. Chen, A. Bouferguene, H. Zaman, M. Al-Hussein, L. Kurach, A deep learning-based framework for an automated defect detection system for sewer pipes, *Automation in Construction* 109 (2020), 102967, <https://doi.org/10.1016/j.autcon.2019.102967>.
- [11] Y. Jiang, D. Pang, C. Li, A deep learning approach for fast detection and classification of concrete damage, *Automation in Construction* 128 (2021), 103785, <https://doi.org/10.1016/j.autcon.2021.103785>.
- [12] Z. Yu, Y. Shen, C. Shen, A real-time detection approach for bridge cracks based on YOLOv4-FPM, *Automation in Construction* 122 (2021), 103514, <https://doi.org/10.1016/j.autcon.2020.103514>.
- [13] Everingham, M., Van Gool, L., Williams, C.K.I. et al. The Pascal Visual Object Classes (VOC) Challenge. *Int J Comput Vis* 88, 303-338 (2010). <https://doi.org/10.1007/s11263-009-0275-4>.