

改进 YOLOv11 的轻量型建筑表面 裂缝检测算法

丁倩^{1,2}, 陶为戈², 孙志刚³

(1. 江苏理工学院 机械工程学院, 江苏 常州 213001;

2. 江苏理工学院 电气信息工程学院, 江苏 常州 213001;

3. 哈尔滨工业大学 电气工程及自动化学院, 哈尔滨 150001)

摘要: 建筑表面裂缝检测是保障建筑工程质量的关键环节; 为应对现有算法在小尺度目标检测准确性不足以及内存占用过高限制了其在边缘设备上应用的问题, 提出了一种基于 YOLOv11 的轻量型裂缝智能检测算法 YOLOv11-LCS; 通过广泛收集建筑表面裂缝图像数据构建 BSCRACK 数据集, 在算法设计上结合轻量化策略与多尺度特征融合, 引入 LDCConv 模块和添加额外卷积操作的 CCFM 模块, 优化了主干特征提取网络和颈部特征融合网络, 显著提升了检测效率和精度; 此外, 采用无参注意力机制 SimAM 构建 C3k2-SimAM 模块, 在不增加额外计算成本的前提下, 进一步增强对小尺度裂缝的检测能力; 实验结果表明, YOLOv11-LCS 的参数数量和计算量分别减少了 42.2% 和 30.2%, 参数量仅为 1.49 M, 同时 $mAP@0.5$ 提升至 97.9%, 有效满足了建筑表面裂缝检测的精度需求, 展现了其在边缘设备上的应用潜力。

关键词: 裂缝检测; YOLOv11; 轻量型; 注意力机制; 多尺度特征融合

Detection Algorithm for Lightweight Building Surface Cracks Based on Improved YOLOv11

DING Qian^{1,2}, TAO Weige², SUN Zhigang³

(1. School of Mechanical Engineering, Jiangsu University of Technology, Changzhou 213001, China;

2. School of Electrical and Information Engineering, Jiangsu University of Technology, Changzhou 213001, China;

3. School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin 150001, China)

Abstract: The detection of surface cracks in buildings is a key link in ensuring the quality of construction projects. To address the issues of insufficient accuracy in detecting small-scale targets and high-memory usage, which limit the application of existing algorithms on edge devices, a lightweight intelligent crack detection algorithm based on YOLOv11, called YOLOv11-LCS, is proposed. By extensively collecting the images of building surface cracks, the Building Surface Crack Recognition and Analysis for Detection (BSCRACK) dataset is created. And in the algorithm design, the lightweight strategy is combined with multi-scale feature fusion. The Linear Deformable Convolution (LDCConv) module and additional convolutional operation in the Cross-Scale Feature Fusion Model (CCFM) module are introduced to optimize both the backbone feature extraction network and the neck feature fusion network, which significantly enhances detection efficiency and accuracy. Furthermore, it employs the parameter-free attention mechanism Similarity-Aware Activation Module (SimAM) to construct the C3k2-SimAM module, which enhances the detection capability for small-scale cracks without increasing additional computational costs. Experimental results show that, The YOLOv11-LCS reduces the number of parameters and computational load

收稿日期: 2025-01-22; 修回日期: 2025-02-21。

基金项目: 国家自然科学基金项目(624B2046); 中央高校基本科研业务费专项资金(HIT. DZJJ. 2024008); 江苏省重点研发计划(BE2019317); 常州市 5G+工业互联网融合应用重点实验室(CM20223015); 2023 年江苏省研究生实践创新训练项目(SJCX23_1625)。

作者简介: 丁倩(1997-), 女, 硕士研究生。

通讯作者: 陶为戈(1973-), 男, 副教授, 硕士生导师。

引用格式: 丁倩, 陶为戈, 孙志刚. 改进 YOLOv11 的轻量型建筑表面裂缝检测算法[J]. 计算机测量与控制, 2025, 33(5): 97-105.

by 42.2% and 30.2%, respectively, with the number of parameters being only 1.49 M, while the $mAP@0.5$ is improved to 97.9%. It effectively meets the precision requirements for building surface crack detection, and demonstrates its application potential on edge devices.

Keywords: crack detection; YOLOv11; lightweight; attention mechanism; multi-scale feature fusion

0 引言

随着经济的迅猛发展,建筑业迎来了前所未有的繁荣,各种城市基础设施不断完善。然而,自然灾害和人为损害等因素导致的建筑坍塌事故频发,严重威胁着公众的生命与财产安全^[1]。在众多导致建筑物结构失效的原因中,建筑表面裂缝是最为常见且关键的问题之一。因此,为了有效控制维护成本、延长建筑物使用寿命并确保其安全性,实施定期的结构健康监测变得尤为重要^[2]。

通过引入先进的人工智能算法,不仅能够提高裂缝识别的准确性和效率,还能为后续的修复工作提供更加详尽的数据支持,从而更好地服务于建筑设施的维护与管理。近年来,基于深度学习的技术在裂缝检测领域得到了广泛应用,并取得了显著成效。这类技术主要分为两大类:两阶段目标检测技术和单阶段目标检测技术。在两阶段方法中,首先生成候选区域,然后对这些区域进行进一步分析以确定是否包含裂缝以及具体类型,代表性的算法包括 R-CNN^[3]、Faster R-CNN^[4]和 Mask R-CNN^[5]等。然而,这种方法的检测速度较慢。相比之下,单阶段方法直接从输入图像中预测出裂缝的位置及其类别,因此检测速度更快,代表性的算法有 SSD^[6]和 YOLO 等系列算法^[7-10]。文献 [11] 提出了一种基于 YOLOv5s 的改进的桥梁裂缝检测方法,该方法通过引入融合视觉注意力机制和自适应空间特征融合模块,使模型能够更好地抑制非关键信息,提高了裂缝目标检测精度。文献 [12] 引入 DGC 新注意力模块,精确定位和检测裂缝,将原 Neck 网络替换为 lowAFPN 进行多尺度融合,使网络能够更全面地捕捉图像中的裂纹信息,提出的裂缝检测算法 DGAP-YOLO 能够满足安全工程和建筑领域裂缝检测的性能要求。以上文献针对裂缝检测提出了不同的算法检测模型,一定程度上提高了检测的精度。然而上述检测算法存在内存资源消耗大、计算效率低的问题,导致其在边缘设备上的移植存在困难。

为了解决这些问题,不少研究者在这方面做出了很多努力。文献 [13] 通过使用轻量级网络替换 YOLOv4 中的特征提取网络,减少骨干层的方法,优化模型网络结构,减少模型计算开销,提高了模型检测速度。文献 [14] 提出了一种基于 YOLOv7-Tiny 的轻量型道路裂缝检测算法 YOLOv7-TPSF,引入部分卷积 PConv,降低模型参数量,同时提出新的特征融合

模块 Bi-FusFPN,减少网络计算量,强化多尺度特征,使得网络的检测准确率和检测速度有了显著提升。文献 [15] 通过在 C2f 模块中融入轻量化多尺度卷积 ALConv,在特征提取网络浅层引入三元注意力机制,运用参数共享理念将原解耦头替换为轻量化解耦头,并将损失函数更换为 MPDIUO 损失函数,构建了一种 YOLOv8-ALTE 轻量化算法,该方法在桥梁裂缝检测中展现出了优异的性能。文献 [16] 基于 YOLOv9 架构提出了一种用于路面裂缝损坏检测的新型轻量级目标检测模型 YOLO9tr,该方法结合了部分注意力机制,提升了模型在复杂场景下的检测能力。YOLO9tr 模型在实时路面状况监测和其他类似应用中具有很大的实际部署潜力。尽管上述文献提出的方法有效促进了模型的轻量化,然而伴随着模型参数的减少以及计算复杂度的下降,模型的检测能力也受到了一定的限制,如何解决建筑表面小尺度裂缝检测准确率低、效率低的问题,还需要做进一步的突破。

在目标检测领域,YOLO 系列算法以其快速迭代和显著的性能提升而著称。目前经过不断更新迭代,已发展至 YOLOv11 版本^[17],相较于之前的版本,YOLOv11 在检测速度上实现了更大的飞跃,并且更加适合在轻量化设备上移植。然而,针对小尺度裂缝的检测与识别,仍有较大的提升空间。本文以 YOLOv11 为基础,提出了一种轻量级算法 YOLOv11-LCS,旨在轻量化网络结构的同时提高对多尺度裂缝,特别是小尺度裂缝的检测和识别能力,从而在检测精度与效率之间实现更好的平衡,极大地拓展了该方法在边缘检测设备上的移植潜力。

YOLOv11-LCS 算法主要对适应建筑裂缝检测需求的 YOLOv11 算法进行了如下改进:

- 1) 引入线性卷积模块 LDConv,替换 C3k2 模块后的传统卷积操作,对特征提取网络进行轻量化处理,改善特征提取网络中的冗余结构,同时强化多尺度特征提取能力。

- 2) 采用添加 Conv 的跨尺度特征融合模块 (CCFM) 优化颈部特征融合网络,将不同尺度的特征进行深度融合,提高对多尺度信息的捕获能力,同时优化颈部网络参数。

- 3) 引入无参注意力机制 SimAM,构建 C3k2-SimAM 模块,进一步提升网络对小尺度裂缝的检测能力,提高网络性能。

1 YOLOv11 算法原理

Ultralytics 公司为满足不同应用场景的个性化需求, 通过精细调整网络的深度和宽度系数, 于 2024 年 9 月推出了 n、s、m、l、x 等系列 YOLOv11 模型。如图 1 所示, YOLOv11 的网络结构主要由负责初步特征提取的主干网络 (Backbone)、用于特征融合与增强的颈部网络 (Neck)、以及执行最终检测任务的头部网络 (Head) 3 个部分组成。YOLOv11 在 Backbone 中采用了对传统 C3 模块深度改进与优化的 C3K2 模块, C3K2 引入了多尺度卷积核 C3K, 其中 K 代表一个可调节的参数, 用于定义卷积核的大小, 这种设计显著拓宽了模型的感受野, 增强了对不同尺度目标的捕捉能力。在结构上, C3K2 模块将输入特征图分为两路进行处理: 一路通过常规卷积操作直接提取基础特征, 另一路则根据条件选择, 当 C3K=True 时, 采用一系列 C3K 模块进行深层次特征挖掘; 而当 C3K=False 时, 则通过 Bottleneck 结构来实现特征的深化处理。随后两路特征拼接合并, 并通过一层 1×1 卷积完成特征融合, 这一机

制不仅保持了模型的轻量化特性, 还极大地促进了深层特征的有效提取。相较于之前版本中的 C2f 块, C3K2 凭借其灵活的可变卷积核策略及高效的通道分离技术, 实现了特征提取能力的飞跃。这种设计不仅提升了计算效率, 还使得 YOLOv11 在图像分析过程中能更迅速地捕获并整合关键信息, 为后续的目标检测任务奠定了坚实基础。

C2PSA 模块是 YOLOv11 中用于增强特征提取的一个高级模块, 该模块不仅有效地结合了 CSP 结构带来的特征多样性和表达能力, 还充分利用了 PSA 注意力机制在增强特征表示方面的优势。具体而言, C2PSA 模块对输入特征进行一次 1×1 卷积后被分成两部分, 一部分直接传递特征, 保留了原始的、未经修改的特征信息, 确保了信息的完整性和直接性。另一部分将特征送入 PSA 注意力模块, 这一模块通过显式的建模特征通道间的相关性, 自适应地重新校准通道特征, 从而突出了对目标检测更为关键的信息。通过这种方式, C2PSA 模块能够在不增加过多计算负担的情况下, 显著提升模型对复杂场景中多尺度物体的检测能力。

YOLOv11 的检测头引入了 YOLOv10^[18] 的设计思想, 使用深度可分离的方法来减少冗余计算, 提高计算效率。这种方法不仅简化了模型的运算流程, 还大幅减少了模型参数总量与所需的浮点运算次数 (FLOPs), 实现了模型参数量与速度的双重优化。与 YOLOv8 模型相比, YOLOv11 在网络架构上进行了精心调整, 包括网络的层数、深度、宽度以及输出通道的配置比例均有所变化。尽管网络层数有所增加, 但得益于各种轻量化模块的应用, 整体参数量和 FLOPs 反而得到了缩减。这一改进意味着 YOLOv11 在维持甚至提升检测性能的同时, 更加高效地利用了计算资源。YOLOv11 通过这些架构上的改进, 实现了更高的检测精度和更快的处理速度, 使其成为当前实时目标检测任务中的一个强大工具。

但是, YOLOv11 在城市建筑结构体表面裂缝检测中仍存在一下待解决的问题:

1) 在建筑表面裂缝检测领域, 为了促进 YOLOv11 网络模型的广泛应用, 特别是在资源受限的边缘检测设备上, 有必要进一步优化该网络以实现更高的检测精度与更低的参数数量之间的理想平衡, 从而确保检测算法能够在保持高效性能的同时, 更加流畅地运行于各类计算资源受限的平台上。

2) 在实际的建筑表面裂缝检测过程中, 由于拍摄角度和距离变化多样性, 裂缝的视觉特征随即产生较大的差异, 这极大地增加了准确识别的难度。此外, 裂缝本身所具有的复杂拓扑结构, 如曲折蜿蜒、分支繁多、形态细小等特性, 使得在图像分析时容易忽略或遗漏那

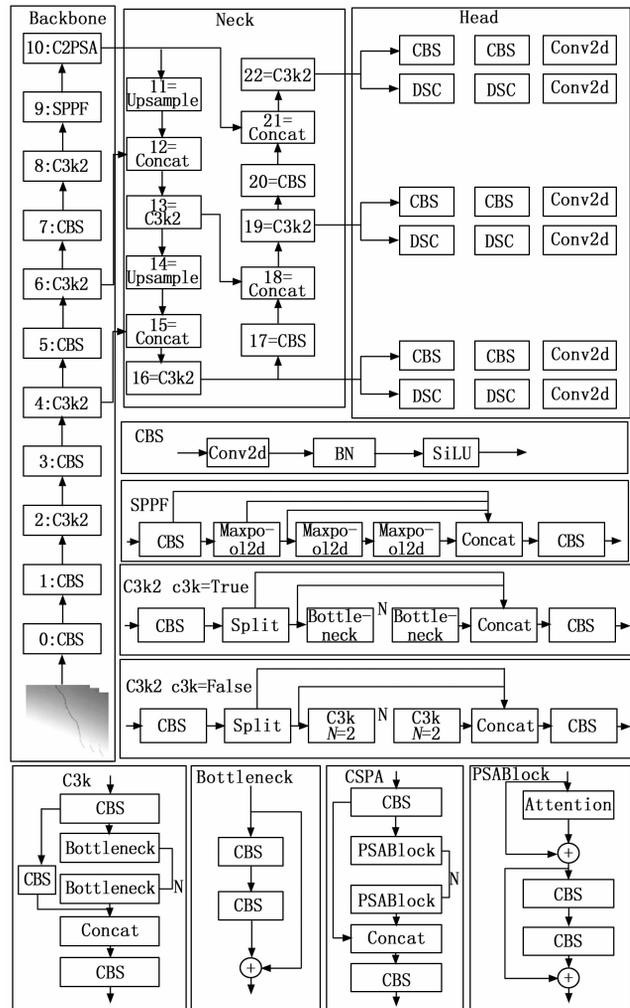


图 1 YOLOv11 算法

些尺度微小但同样关键的裂缝细节。

2 轻量型 YOLOv11-LCS 算法

YOLOv11-LCS 算法基于 YOLOv11n 模型进行了深度优化，主要解决建筑表面裂缝检测任务中存在的准确度不足和模型资源消耗过大等问题。该算法通过降低网络复杂度并强化多尺度特征，尤其是小尺度特征的提取能力，有效弥补原有算法在建筑表面裂缝检测方面存在的缺陷。具体改进措施如下：

首先，采用线性卷积 LDCnv 替代部分传统卷积作为特征提取过程中的下采样方式，降低模型参数的同时提高多尺度特征提取能力。其次，对原有的特征融合网络进行了结构调整与优化设计，进一步大幅度降低计算成本，改善特征融合效率低下问题，提升跨层信息整合效果。最后，在特征融合中，引入了一种无参数注意力机制，构建 C3k2_SmiAM 模块，该模块能够自动学习并聚焦于图像中最具代表性或关键性的部分，从而极大地增强了模型对于目标区域的关注程度。YOLOv11-LCS 算法网络整体结构如图 2 所示。

2.1 主干特征提取网络轻量化

为了实现网络的快速处理并进一步优化网络计算量，本研究着重于网络轻量化策略的实施，同时旨在增强特征提取网络对小尺度目标的捕捉能力。为此，引入了线性可变性卷积 (LDCnv) 技术替换主干网络中 C3K2 模块后的每一个传统下采样操作，如图 2 中

Backbone 部分所示。标准卷积操作通常在固定形状和大小的局部窗口内进行采样，这种方式难以动态适应不同物体的复杂形状与尺寸变化。为了打破这一限制，LDCnv (Linear Deformable Convolution)^[19] 作为一种改进的卷积技术应运而生，它旨在解决传统标准卷积在处理多尺度特征提取时的局限性。LDCnv 通过引入线性可变的参数增长策略，巧妙地平衡了灵活性与计算效率之间的关系。它允许卷积核的参数数量呈线性增长，从而有效克服了可变形卷积参数数量平方增长的问题。这种方法不仅增强了模型对多尺度特征的适应性，还起到了轻量化的作用，使得模型在保持高效特征提取能力的同时，减少了计算资源的消耗。

LDCnv 允许卷积核具有任意数量的参数和任意的采样形状，这种灵活性使得模型能够更有效地捕捉图像中不同尺度的目标特征，显著提升了多尺度特征提取的能力，其生成过程如下：首先，设计一个任意大小卷积核的生成算法，用于生成卷积核的初始采样坐标 P_0 ；其次，以图像左上角 (0, 0) 作为采样原点，对初始采样坐标偏移 P_n 后，通过卷积运算计算偏移后位置的像素值 P_0, n ，如公式 (1) 和 (2) 所示；最后，将初始采样形状 S 与偏移位置 P_n 相加得到最终的动态采样形状：

$$P_f = (P_0 + P_n) \tag{1}$$

$$P_0, n = \sum_{P_n \in R} WP_n * Pf \tag{2}$$

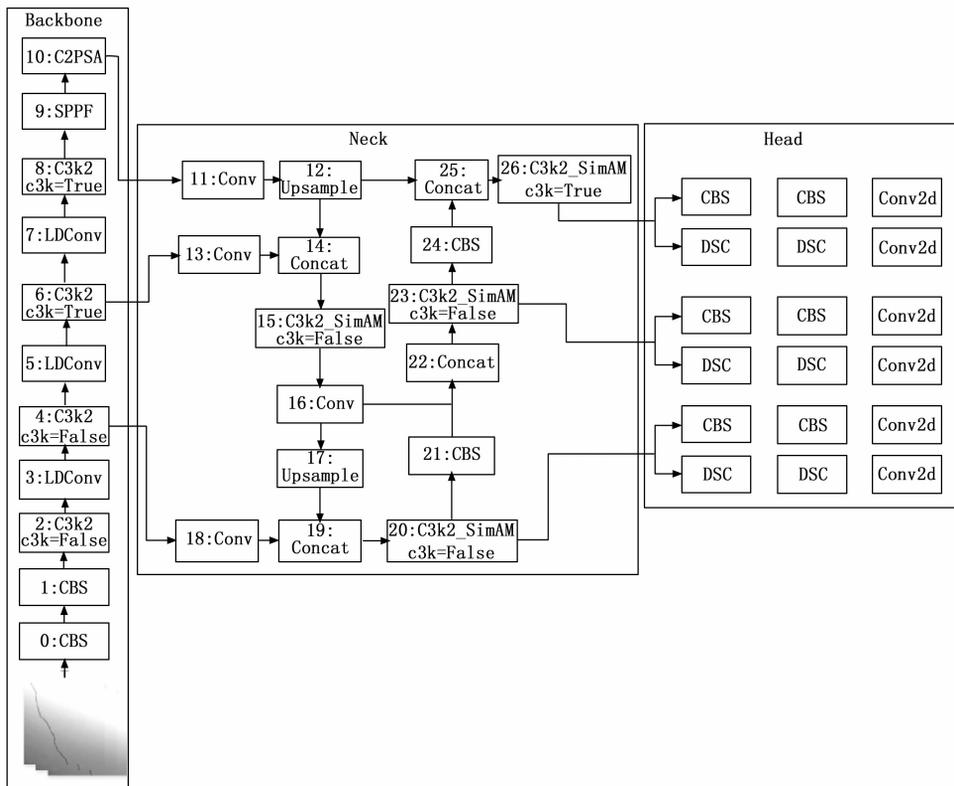


图 2 YOLOv11-LCS 网络结构

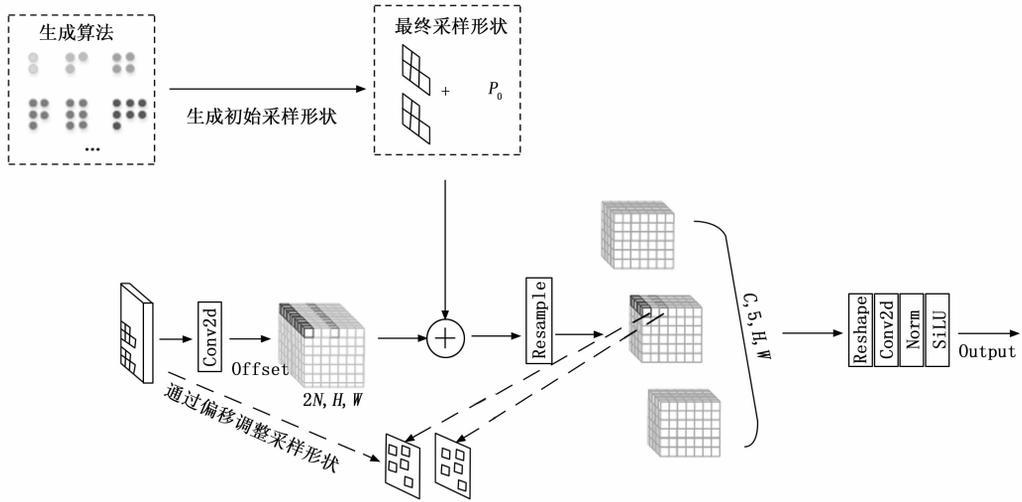


图 3 LDCConv 模块结构图

式中, S 为生成的采样形状, WP_n 为卷积参数, P_f 为初始位置添加偏移量后对应位置像素值。

LDCConv 结构如图 3 所示, 其特征提取过程可以分为 3 个步骤: 首先, 根据卷积核的大小, 使用提出的生成算法生成初始采样形状; 其次, 通过卷积操作获取相应卷积核的偏移量, 通过学习到的偏移量调整初始采样形状; 最后, 调整后的最终采样形状对特征图进行重采样, 重采样后的特征图经过重塑, 再次卷积、归一化处理, 然后通过激活函数 SiLU 输出最终结果。通过上述 3 个步骤, LDCConv 能够完成任意尺寸的卷积操作以提取特征。这种机制使得 LDCConv 在处理图像时更加灵活和有效, 尤其是在需要捕捉细节和局部特征的场景中表现更为出色。

2.2 颈部特征融合网络结构优化

跨尺度特征融合模块 (CCFM) 是在 RT-DETR^[20] 中提出的, 如图 4 (a) 所示, 该模块通过多个 Fusion 模块的串联和循环实现了特征信息的多次融合和传递, 将不同尺度的特征通过融合操作整合起来, 以增强模型对于尺度变化的适应性和对小尺度的检测能力。这种设计有效地整合了细节特征和上下文信息, 从而提高了模型的整体性能。CCFM 块中的融合块 Fusion 的结构如图 4 (b) 所示, 融合块的作用是将两个相邻的尺度特征融合成一个新的特征。融合块包含两个 1×1 卷积来调整特征通道, 再由 RepConv 组成 N 个 RepBlocks, 其作用是进一步处理输入数据, 可能包括降维、特征提取等操作, 两路径输出由元素加法融合。

如图 5 所示, 为了进一步优化颈部特征融合网络, 减少冗余的同时增强网络多尺度特征融合能力, 在特征融合前引入了 3 个 1×1 卷积层作为预融合模块, 这些卷积层将不同尺度特征的通道维度统一压缩至 256, 在

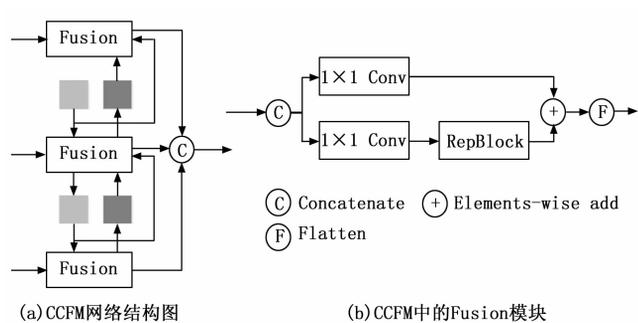


图 4 CCFM 网络结构

保留关键特征信息的同时显著降低了计算复杂度。随后引入了 CCFM 模块, 该模块采用恒定的 256 通道配置, 通过跨层次信息流的深度融合机制, 实现了多分辨率特征图的高效整合。

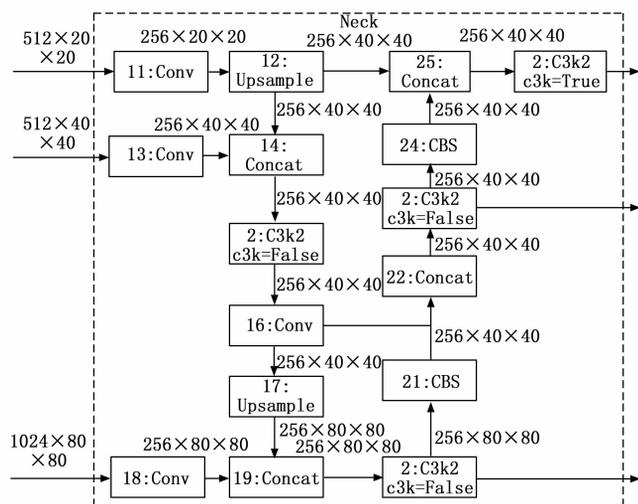


图 5 CCFC 网络结构图

2.3 C3K2_SimAM 特征融合模块

在卷积神经网络 (CNN) 中, 注意力机制的应用日益广泛, 将注意力机制应用于建筑裂缝的检测算法中, 可以引导网络更加专注于裂缝的关键区域, 从而更准确地捕捉裂缝特征, 提升裂缝检测的整体性能。然而, 传统的注意力机制往往伴随着额外参数的引入, 这不仅增加了模型的复杂度, 也提高了计算成本。文献 [21] 提出了一种轻量级、无参的卷积神经网络注意力机制 SimAM, 为解决这一问题提供了新的途径。在图像中, 相邻像素之间通常具有较强的相似性, 而远距离像素之间的相似性则较弱。SimAM 利用这一特性, 通过计算特征图中每个像素与其相邻像素之间的相似性来生成注意力权重。这一过程无需引入任何额外参数, 因此能够在不增加模型复杂度和计算成本的前提下, 有效提升 CNN 的性能。

具体而言, SimAM 首先对输入的特征图进行局部自相似性的计算, 以获取特征之间的相关性信息。然后, 基于这些相关性信息, SimAM 生成相应的注意力权重。最后, 将这些注意力权重应用于特征图上, 从而实现了对关键区域的强化关注。SimAM 注意力机制结构如图 6 所示。

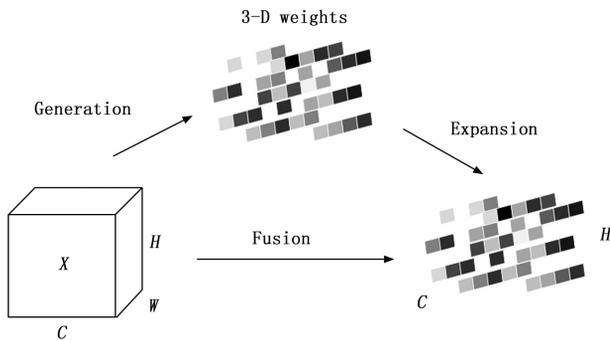


图 6 SimAM 注意力机制

为进一步减少特征融合时造成的特征损失, 提高裂缝检测的准确度, 在网络特征融合中, 对 C3k2 模块进行改进, 添加 SimAM 注意力机制, 构建 C3k2_SimAM 模块, 如图 7 所示。通过该方式, SimAM 不仅能够提高模型对裂缝特征的识别能力, 还能够降低模型的计算负担, 使其更适合在实际工程中的应用。

3 实验结果与分析

3.1 数据集

针对现有建筑表面裂缝数据集中细小裂缝样本较少, 以及实际拍摄中由于距离远或角度倾斜导致裂缝在图像中比例较小、难以识别的问题, 采用从互联网上广泛搜集到的包含建筑表面裂缝的图片作为原始数据源, 随后利用 Labelimg 这一高效的图像标注工具, 对收集到的图片进行精确标注, 另外对初步标注后的数据集进行了一系列的处理和增强操作, 包括但不限于图像的旋转、缩放、裁剪、颜色变换等, 旨在模拟不同拍摄条件下的裂缝形态, 制作的数据集命名为 BSCRACK (Building Surface Crack Recognition and Analysis for Detection)。

该数据集专注于建筑表面小尺度裂缝的检测与分析。数据集中涵盖了多种建筑结构类型和不同环境条件下的建筑表面裂缝图像, 包括混凝土结构、涂料涂层结构以及砖砌结构等。为了确保算法在实际应用中的鲁棒性, 数据集中还纳入了多种结构在不同复杂场景下的图像, 例如不同光照条件 (如强光、弱光、阴影等)、覆盖物遮挡 (如污渍、苔藓、广告牌等) 以及其他干扰物 (如窗台、防护网、栅栏等其他建筑设施) 存在的情况下的图像数据, 数据集中建筑表面裂缝图像如图 8 所示。这些多样化的数据能够有效提升算法在复杂环境下的适应性和检测精度。整个数据集共包含 5 005 张高质量的建筑表面裂缝图像, 并按照 7 : 2 : 1 的比例划分为

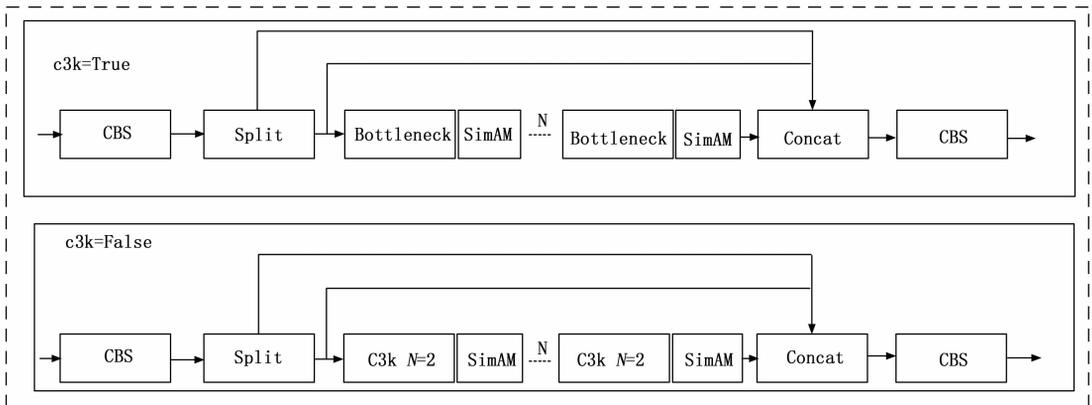


图 7 C3k2_SimAM 模块结构图

训练集、验证集和测试集, 保证了模型训练的有效性, 还为算法的性能评估提供了可靠的数据支持, 从而达到了检测任务所需的良好性能标准。

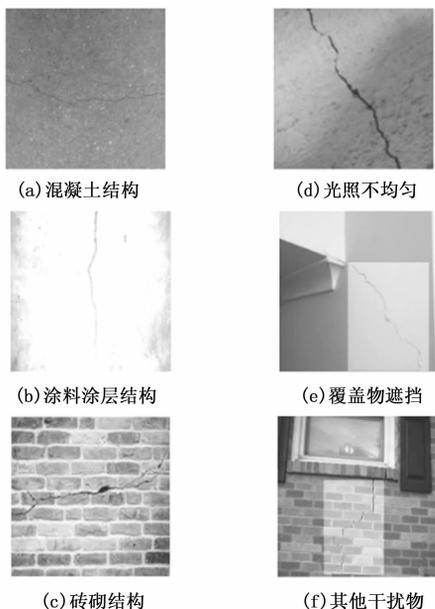


图8 BSCRACK 数据集中涵盖的建筑裂缝图像种类

3.2 实验环境与参数配置

实验条件: 64 位 Windows 11 操作系统、Intel (R) Core (TM) i5-9300H CPU @ 2.40 GHz 处理器、NVIDIA GeForce GTX1650 显卡 (显存 4 GB); 软件环境依托于 PyTorch 1.10.1 深度学习框架, 搭配 CUDA 10.2 和 Python 3.8.20 版本。实验参数设置主要包括: 图像像素分辨率 640×640 , 训练过程中采用批次大小 (batch size) 为 8, 训练轮次 (epoch) 100 个, 优化器选择 Adam, 其余参数均采用默认配置。具体参数设置详见表 1。

表1 参数配置

配置参数名称	配置参数数值
输入尺寸	640×640
训练批次大小	8
训练轮数	100
动量参数	0.937
权重衰减	0.005
初始学习率	0.001
最终学习率	0.001

3.3 评估指标

为了确保建筑裂缝智能检测模型在实际工程中具有高效的应用性能, 重点关注模型的内存消耗、计算速度和计算精度。具体而言, 模型需要具备较少的内存占用, 以便能够轻松移植到资源有限边缘检测设备上; 同

时, 要求模型具有快速的计算能力, 以实现实时检测; 此外, 模型还需具备高计算精度, 以满足工程评估的需求。

因此, 采用了准确率 (Precision, P)、召回率 (Recall, R)、平均精度均值 (mean Average Precision, mAP)、参数数量 (Params)、计算量 (FLOPS) 作为模型的评价指标。这些指标综合考量了模型的性能表现, 为模型的优化和应用提供了有力的支持。 P 、 R 、 AP 和 mAP 的计算公式如 (3) ~ (6) 所示:

$$P = \frac{TP}{TP + FP} \tag{3}$$

$$R = \frac{TP}{TP + FN} \tag{4}$$

$$AP = \int_0^1 P(R) dR \tag{5}$$

$$mAP = \frac{\sum_{i=1}^n AP}{n} \tag{6}$$

其中: n 为图像总类别数, i 为检测次数, TP 为实际是裂缝被预测为裂缝样本的数量, FP 为实际是背景被预测成裂缝样本的数量, FN 为实际是裂缝被预测成背景样本的数量。 AP 表示曲线与准确率 P 为纵轴和召回率 R 为横轴的坐标系下围成的面积。

3.4 消融实验

为了进一步验证提出的各种改进方法的有效性, 对相关的改进模块进行一系列的消融实验。旨在通过逐步调整每个改进措施, 观察其对整体性能的影响, 从而明确每个改进点的具体贡献和必要性。通过这种方式, 可以深入理解各种改进策略间的相互作用及其对于提升系统性能的重要性。

表2 消融实验结果

序号	LDCConv	CCFC	C3k2_SimAM	P %	R %	mAP @0.5%	Para ms/M	FLO Ps/G
①	/	/	/	95.8	92.4	97.2	2.58	6.3
②	√	×	×	96.9	90.9	96.6	2.28	5.6
③	×	√	×	96.6	90.3	96.7	1.82	5.3
④	×	×	√	96.1	92.4	97.4	2.45	5.9
⑤	√	√	×	96.4	91.9	97.8	1.49	4.4
⑥	√	√	√	97.4	92.7	97.9	1.49	4.4

从表 2 消融实验结果可知:

1) 从实验②的结果来看, 通过在主干特征提取网络中引入线性卷积 LDCConv 来部分替换传统卷积操作的改进策略, 显著降低了模型参数数量和复杂度, 实现了模型主干特征提取网络的轻量化。具体来说, 相对实验①模型的参数由 2.58 M 下降至 2.28 M, 模型的复杂度由 6.3 G 下降至 5.6 G, 同时精度 P 提高了 1.1%。对比

实验①和③，通过采用额外添加 Conv 的跨尺度特征融合结构 (CCFC) 对模型颈部特征网络进行优化，模型的参数量和计算量分别大幅下降 29.5% 和 15.9%，精度提升了 0.3%。根据实验④对比①的结果可以看出，添加无参注意力机制 SimAM，构建 C3k2_SimAM 模块，能够有效进一步增强模型对于裂缝的关注度，对小尺度裂缝的检测精度有了明显提升。

2) 对比①和⑤实验可知，引入线性卷积和优化颈部特征融合结构后，模型的精度和平均精度 $mAP@0.5$ 分别提升了 0.6% 和 0.2%，参数量和计算量分别下降了 42.2% 和 30.2%，表明两部分的协同效果更为显著。

3) 对比实验②、③和实验⑤的结果可以发现，联合使用 LDConv 模块和 CCFC 结构后，模型的平均精度 $mAP@0.5$ 从原来的 96.6% 和 96.7% 显著提升至 97.8%，相较于单独引入 LDConv 模块或 CCFC 结构，两者的联合使用取得更优的检测效果。这一结果表明，LDConv 模块在主干特征网络中的应用有效提升了多尺度特征的提取能力，而 CCFC 结构在颈部网络中的多尺度特征融合进一步增强了特征表达的丰富性和鲁棒性。此外，多尺度特征融合与轻量化策略的结合使得网络参数量大幅减少至 1.49 M，进一步验证了两者在降低模型复杂度与提升检测精度方面的协同效果。

4) 融合三项改进模块的实验⑥，相较于基础实验①，各项指标均有显著提升，具体来说，模型参数量由原来的 2.58 M 降低至 1.49 M、计算复杂度由原来的 6.3 G 减少至 4.4 G，同时精确率 P 提升至 97.4%，召回率 R 提升至 92.7%，平均精度 $mAP@0.5$ 提升至 97.9%；相较于实验⑤，在参数量和复杂度都保持不变的前提下，模型的 P 、 R 、 $mAP@0.5$ 各项指标均进一步有效提升，表明 C3k2_SimAM 模块在不引入额外计算负担和参数冗余的前提下，有效提升了模型检测的精度，进一步优化了算法的性能表现。综上所述，这些实验表明，采用线性卷积技术和优化特征融合策略能够有效减轻网络负担并提高检测效率；而无参注意力机制的应用则有助于加强细微裂纹等的检出率，整体上促进了算法性能全面提升。

综合分析表 3 的消融实验结果，验证了本文提出的各个改进模块都具有实质效果。

3.5 对比实验

为了验证 YOLOv11-LCS 算法在建筑表面裂缝检测任务中的优越性，以及评估自制数据集 BSCRACK 的泛化能力，在相同的实验环境和条件下，将 YOLOv11-LCS 算法与当前主流的 YOLOv5n 等目标检测算法进行性能对比实验，结果如表 3 所示。

表 3 不同目标检测算法对比实验结果

算法	YOLO	YOLO	YOLO	YOLO	YOLO	Ours
	v5n	v8n	v9t	v10n	v11n	
$P\%$	94.9	95.5	96.7	93.6	95.8	97.4
$R\%$	91.5	92.8	92.2	81.9	92.4	92.7
$mAP@0.5$	96.7	97.5	96.6	92.3	97.2	97.9
Params/M	2.5	3.0	1.9	2.69	2.58	1.49
FLOPs/G	7.1	8.1	7.6	8.2	6.3	4.4

实验结果表明，作为基准算法的 YOLOv11n 满足网络轻量化的要求，并且该算法在各项指标中的综合表现优于 YOLOv5n、YOLOv8n、YOLOv9t、YOLOv10n 算法。具体来说，YOLOv11-LCS 算法检测平均精度分别提高了 1.2%、0.4%、1.3%、5.6%、0.7%。最新的 YOLOv8n、YOLOv9t，虽然在检测精度上有一定的提升，但与 YOLOv11-LCS 算法相比仍然落后很多。从模型模型参数和计算复杂度上看，相比于 YOLOv5n、YOLOv8n、YOLOv9t、YOLOv10n 算法，YOLOv11-LCS 的参数量分别减少了 40.4%、50.3%、21.6%、44.6，复杂度分别降低了 38%、45.7%、42.1%、46.3%。由此可见，与现有的裂缝检测方法相比，该方法在精度、模型尺寸和计算复杂度等方面均具有优势。

为了更直观地体现改进网络的检测效果，对 YOLOv11n 和 YOLOv11-LCS 算法的建筑表面裂缝检测结果进行了可视化对比分析，如图 9 所示。通过 A 组实验图像对比可以看出，改进后的算法在细小裂缝检测中的置信度提升明显；B 组实验结果表明，YOLOv11n 存在部分漏检现象，YOLOv11-LCS 算法有效减少了漏检和错检问题，裂缝检测的识别精度显著提高；C 组实验进一步表明，YOLOv11-LCS 算法在裂缝定位方面更加准确。因此，YOLOv11-LCS 算法更加充分地表达了裂缝目标，对小尺度裂缝的定位精度更高。

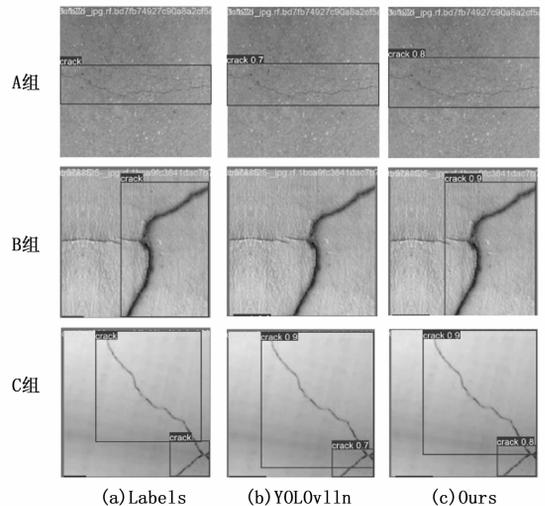


图 9 改进前后检测结果对比

4 结束语

研究提出了一种改进的YOLOv11建筑表面裂缝检测算法。通过引入线性卷积(LDConv)部分替换传统卷积操作,实现了主干特征提取网络的轻量化;同时采用添加额外卷积操作的跨尺度特征融合结构,优化颈部特征融合网络,增强多尺度特征融合能力,降低颈部网络参数;此外还添加无参注意力机制SimAM,构建C3k2_SimAM特征融合模块,提高模型对细微裂纹等小目标物体检测的准确性。实验结果显示,改进后的模型在自制的BSCRACK数据集上表现更出色,相较于YOLOv11n其在精度(Precision)、召回率(Recall)、平均精度(mAP)等关键指标上均有显著提升。同时,模型的参数量和计算复杂度明显降低,有效证明了该方法在建筑表面裂缝图像的目标检测任务中兼具高效性与准确性。

下一步将探索将此算法迁移至资源受限的边缘计算平台上,以便更加灵活高效地部署于各种现场环境之中,同时考虑该算法在建筑表面存在阴影遮挡或其他噪声干扰下目标检测不足的问题,后续针对该类问题做进一步研究。

参考文献:

- [1] KACKER R, SINGH S K, KASAR A A. Understanding and Addressing Multi-faceted Failures in Building Structures [J]. *Journal of Failure Analysis and Prevention*, 2024, 24 (4): 1542 - 1558.
- [2] 余欣, 许益奖, 李薇, 等. 基于图像处理的远距离裂缝检测智能平台开发 [J]. *计算机测量与控制*, 2022, 30 (12): 70 - 77.
- [3] GIRSHICK R. Fast R-CNN [C] // 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015: 1440 - 1448.
- [4] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39 (6): 1137 - 1149.
- [5] HE K, GKIPXARI G, DOLLÁR P, et al. Mask R-CNN [C] // 2017 IEEE International Conference on Computer Vision (ICCV), Venice: Italy, 2017: 2980 - 2988.
- [6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multiBox detector [C] // *Computer Vision-ECCV 2016*, Cham: Springer International Publishing, 2016: 21 - 37.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [J]. <http://arxiv.org/abs/1506.02640>, 2016.
- [8] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [J]. <http://arxiv.org/abs/1612.08242>, 2016.
- [9] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [J]. <http://arxiv.org/abs/2004.10934>, 2020.
- [10] LI C, LI L, JIANG H, et al. YOLOv6: A single-stage object detection framework for industrial applications [J]. 2022.
- [11] 郭佳佳, 董增寿, 常春波. 基于改进YOLOv5的桥梁裂缝模型研究 [J]. *计算机测量与控制*, 2023, 31 (12): 188 - 194.
- [12] SUN Z, LIU J, LI P, et al. DGAP-YOLO: A crack detection method based on UAV images and YOLO [C] // *Advanced Intelligent Computing Technology and Applications*, Singapore, Springer Nature, 2024: 482 - 492.
- [13] ZHANG J. Automated bridge crack detection method based on lightweight vision models [J]. *Intelligent Systems*, 2023.
- [14] 王启涵, 刘超. 改进YOLOv7-Tiny的道路裂缝检测算法 [J/OL]. *计算机工程与应用*, 2024: 1 - 10.
- [15] 杨炜, 方虹苏, 唐湘松, 等. 桥梁裂缝病害检测的轻量化YOLOv8-ALTE算法 [J/OL]. *交通运输工程学报*, 2024: 1 - 17.
- [16] YOUWAI S, CHAIYAPHAT A, CHAIYAPHAT P. YOLO9tr: a lightweight model for pavement damage detection utilizing a generalized efficient layer aggregation network and attention mechanism [J]. *J Real-Time Image Proc*, 2024, 21 (5): 163.
- [17] KHANAM R, HUSSAIN M. YOLOv11: An Overview of the Key Architectural Enhancements [EB/OL]. <http://arxiv.org/abs/2410.17725>, 2024.
- [18] WANG A, CHEN H, LIU L, et al. YOLOv10: Real-Time End-to-End Object Detection [J]. <http://arxiv.org/abs/2405.14458>, 2024.
- [19] ZHANG X, SONG Y, SONG T, et al. LDConv: Linear deformable convolution for improving convolutional neural networks [J]. *Image and Vision Computing*, 2024, 149: 105190.
- [20] ZHAO Y, LV W, XU S, et al. DETRs Beat YOLOs on Real-time Object Detection [J]. <http://arxiv.org/abs/2304.08069>, 2024.
- [21] YANG L, ZHANG R Y, LI L, et al. SimAM: A simple, parameter-free attention module for convolutional neural networks [C] // *Proc. of the 38th International Conference on Machine Learning*, PML, 2021: 11863 - 11874.