



DOI:10.12404/j.issn.1671-1815.2507317

引用格式:李佩炫,周惠兴,徐崇文,等.基于改进YOLOv11-CUA的建筑门窗视觉测量方法[J].科学技术与工程,2026,26(26):11406-11416.

Li Peixuan, Zhou Huixing, Xu Chongwen, et al. Visual measurement of building windows and doors using improved YOLOv11-CUA framework[J]. Science Technology and Engineering, 2026, 26(26): 11406-11416.

建筑科学

基于改进 YOLOv11-CUA 的建筑门窗视觉测量方法

李佩炫,周惠兴*,徐崇文,李昊禹,吴昊

(北京建筑大学机电与车辆工程学院,北京 100044)

摘 要 针对建筑门窗传统视觉测量方法精度不足和激光点云扫描数据处理复杂的难题,通过将视觉与激光技术相融合,并利用深度学习优化图像处理流程,基于改进YOLOv11(you only look once version11)提出了一种建筑门窗视觉测量方法。利用自主研发的测量设备采集含有深度信息的多视角图像,通过深度信息与位姿信息完成图像校正以拼接并构建包含待测目标墙面的全景图。为提高门窗测量精度的同时减少计算量,采用YOLOv11-CUA(YOLOv11n-C3k2_UIB-ADown)进行目标检测,并结合LSD(line segment detector)直线特征检测算法实现目标区域测量。改进的YOLOv11引入C3K2_UIB模块来增强多尺度特征提取,提升小型门窗检测能力;ADown模块通过轻量化设计减少计算量,替换下采样卷积以增强关键区域的表征能力,减少低对比度场景的细节丢失。在自建门窗数据集上,门窗检测准确率达97.5%, mAP_{50} 为97.2%、参数量仅为2.0 M。相比于YOLOv11n网络,准确率提高了4.7%, mAP_{50} 提高了4.8%,参数量减少了0.6 M。最终尺寸平均测量误差为5.63 mm,最大误差控制在10 mm以内,满足建筑测量规范。为智能建造提供了一种高效的自动化测量方案。

关键词 视觉测量;YOLOv11-CUA;门窗洞目标检测;轻量化模型;门窗洞测量

中图法分类号 TU198;

文献标志码 A

Visual Measurement of Building Windows and Doors Using Improved YOLOv11-CUA Framework

LI Pei-xuan, ZHOU Hui-xing*, XU Chong-wen, LI Hao-yu, WU Hao

(School of Mechanical-Electronic and Vehicle Engineering, Beijing University of Civil Engineering and Architecture, Beijing 100044, China)

[Abstract] To address the challenges of limited accuracy of conventional visual measurement methods for interior doors and windows in buildings, as well as the high computational complexity of laser point cloud data processing, a vision-based measurement method was proposed. Visual and laser technologies were integrated, and deep learning models were employed to optimize and streamline the image processing pipeline. Multi-view images with depth information were captured using a self-developed measuring device. Geometric correction and image stitching were performed using depth information and pose data to construct panoramas of walls containing the target doors and windows. To enhance measurement accuracy while reducing computational load, a YOLOv11-CUA (YOLOv11n-C3k2_UIB-ADown) was utilized for object detection, combined with LSD (line segment detector) algorithm to achieve precise measurement within the target regions. The improved model of YOLOv11 includes the introduction of the C3K2_UIB module to strengthen multi-scale feature extraction and improve detection of small doors and windows. The ADown module employs a lightweight design to reduce computational overhead and replaces conventional downsampling convolutions to strengthen feature representation in key regions, thereby reducing information loss in low-contrast scenarios. Evaluated on a dedicated door-window dataset, the system achieves a detection accuracy of 97.5%, a mAP_{50} of 97.2%, and contains only 2.0 M parameters. Compared with the baseline YOLOv11n, the proposed model shows an improvement of 4.7% in accuracy, 4.8% in mAP_{50} , and a reduction of 0.6 M parameters. The resulting dimensional measurements exhibit an average error of 5.63 mm, with maximum deviations controlled within 10 mm, satisfying relevant building measurement standards. This vision-laser fusion method for measuring interior doors and windows in buildings provides an efficient automated measurement solution for intelligent construction.

收稿日期:2025-10-20; 修订日期:2026-01-19

基金项目:北京建筑大学博士研究生科研能力提升项目(DG2023020)

第一作者:李佩炫(2000—),男,汉族,山西运城人,硕士研究生。研究方向:机器视觉与图像分析技术。E-mail:2108020023036@stu.bucea.edu.cn。

*通信作者:周惠兴(1953—),男,汉族,江苏张家港人,博士,教授。研究方向:建筑机器人与智能装备。E-mail:wm_zhx@126.com。

投稿网址:www.stae.com.cn

[**Keywords**] visual measurement; YOLOv11-CUA; door and window target detection; lightweight model; measurement of door and window openings

在智能建造工程领域,门窗尺寸测量是施工过程中房屋尺寸测量验收的关键环节。常规的手工测量方法高度依赖于操作人员,不仅效率较低,还容易受主观人为因素的干扰而产生较大测量误差^[1]。近年来,市场上主流的自动化测量设备主要采用激光点云扫描仪,这种方法虽精度较高,但存在数据处理复杂且设备昂贵等问题^[2]。而双目视觉测量系统在面对弱纹理或复杂光照场景时,存在测量精度不足、难以满足实时性测量需求等问题^[3]。门窗尺寸测量作为建筑房屋测量的关键部位,精度要求更高,直接影响施工质量与后续安装。因此,实现高精度与实时性协同的门窗测量成为推动房屋尺寸测量发展的重要挑战。

尽管已有研究在提升建筑物门窗洞口尺寸测量的效率与精度方面取得了一定进展,但仍存在诸多亟待解决的难题。Zolanvari 等^[4]提出了一种基于点云切片分析的门窗框提取方法,通过局部点密度突变检测门窗边界,但该方法会产生海量的点云数据,导致测量效率低下。莫玉晓等^[5]基于空洞检测及立面点云来实现门窗检测,尽管在特定场景中提升了检测准确率,但繁重的数据预处理导致耗时较长,限制了实时测量效率。郑晓昱等^[6]将激光测距系统与机器视觉深度融合,构建了一套门窗尺寸自动测量系统,实现了初步的自动化测量,但受限于图像分辨率和复杂光照条件,测量精度难以满足严格的建筑规范,且系统校准与数据处理流程复杂,影响效率。高云龙等^[7]使用模板匹配结合 Joint-Boost 分类器进行窗户提取并优化。但该方法高度依赖门窗几何规整性,在复杂场景下精度明显下降,且模板匹配耗时较长,效率低。

现有的采用深度学习的目标检测方法为门窗测量提供了一种全新的测量思路。沈乐等^[8]运用 YOLOv2 模型并结合图像与点云映射实现门窗角点定位,但其精度受限于早期模型,且推理速度慢,无法满足实时高效测量的需求。Cheng 等^[9]基于改进 YOLOv11 提出了一种水下目标检测识别系统,通过模块优化显著提升了小目标检测能力。Chen 等^[10]通过可变形 DETR 与多级特征融合提升复杂场景检测精度,为门窗边界识别提供借鉴。文献[11-13]分别基于 YOLOv8 和 YOLOv11-Seg 改进算法,优化了弱纹理环境下的检测与分割效率。这些研究虽非直接针对门窗测量,但其高精度检测与轻量化设计的思路对门窗尺寸测量具有参考价值。

为此,现基于 YOLOv11-CUA 设计一种门窗洞

口视觉测量系统。利用自主研发的建筑测量机器人采集含有深度信息的多视角图像,并依据深度信息与位姿信息完成图像校正与拼接,构建墙面全景图并测量房间尺寸。针对门窗尺寸测量的关键环节,使用 YOLOv11-CUA 进行目标检测,并利用 LSD 直线特征检测算法实现门窗尺寸测量。在自建门窗洞口数据集上训练并验证 YOLOv11-CUA 的性能,并对门窗洞进行实验测量,以验证提出方法的综合性能,为实现门窗尺寸自动化测量提供可行路径,具有良好的工程应用价值。

1 测量原理与图像拼接

本节将详细阐述所提出的基于改进 YOLOv11 的建筑视觉测量新方法的测量原理。

图像校正与拼接是高精度测量的关键步骤^[14]。传统单张图像采集往往受限于相机视场角和透视畸变,导致门窗目标在图像中变形或不完整,无法准确映射到物理尺度。通过图像校正,可以消除透视畸变和几何失真,确保图像符合正交投影;通过拼接多视角图像,则能构建完整的高分辨率全景图,覆盖整个墙面目标区域,避免弱纹理或复杂光照场景下的特征匹配失败问题,更精确地拟合墙面平面方程,从而为后续 YOLOv11-CUA 目标检测和 LSD 直线特征提取提供可靠的输入数据,提高整体测量效率和精度。

1.1 测量原理

门窗尺寸测量的核心在于利用已校正的图像,将检测到的门窗目标边界框坐标转换为物理尺寸。采用的室内实测实量机器人通过视觉激光测距模块与高精度转台,获取门窗所在墙面的多视角深度图像,并结合转台位姿信息完成图像校正与拼接,生成具有真实世界尺度的高分辨率门窗图像。其中,激光测距模块通过提供设备到墙面的精确距离来定义参考平面;通过透视变换进行图像校正与拼接获取门窗全景图后,使用 YOLOv11-CUA 模型检测出门窗目标的边界框,最后运用 LSD 直线特征检测算法对待测门窗进行边框检测提取,完成门窗尺寸测量。

1.2 基于深度参数的图像校正

提出了基于深度参数的校正与拼接方法,融合激光测距值、转台旋转角度及设备设计尺寸,直接推导几何变换,有效规避传统特征点匹配在弱纹理场景下性能显著下降的问题。

1.2.1 基于激光测距的墙面平面建模

由于实际墙面非理想平面,故提出多激光点拟

合平面方法:利用设备对称分布的4个激光测距模块,通过最小二乘拟合最优平面方程。

$$ax + by + cz + d = 0 \quad (1)$$

式(1)中:\$(x, y, z)\$为相机坐标系下激光点的坐标;\$(a, b, c)\$为平面的法向量;\$d\$为相机坐标系原点到该墙面的垂直距离。

1.2.2 单应性变换

基于转台旋转角度计算旋转矩阵\$\mathbf{R}\$,结合经由张正友标定法所获得的相机内参矩阵\$\mathbf{K}\$,计算单应性矩阵\$\mathbf{H}\$。

$$\mathbf{H} = \mathbf{K} \mathbf{R} \mathbf{K}^{-1} \quad (2)$$

1.2.3 几何校正与图像反投影

通过单应性逆变换将旋转图像映射到正对墙面视角。

$$I_{\text{corrected}} = W(I_{\text{rot}}, H^{-1}) \quad (3)$$

式(3)中:\$I_{\text{corrected}}\$为校正后的标准正视图像;\$W\$为双线性插值重采样;\$I_{\text{rot}}\$为旋转拍摄的原始图像;\$H^{-1}\$为单应性矩阵的逆变换。

1.3 基于深度参数的图像拼接

通过建立全局世界坐标系,结合相机内参与外参推导图像间变换关系,有效规避传统特征点匹配方法。图像拼接示意图如图1所示。

坐标系变换公式为

$$p_w = \mathbf{R}_i \mathbf{K}^{-1} p_i + t_i \quad (4)$$

式(4)中:\$p_w\$为点在世界坐标系下的坐标;\$\mathbf{R}_i\$为第\$i\$幅图像的旋转矩阵;\$t_i\$为第\$i\$幅图像的外参;\$p_i\$为第\$i\$幅图像的像素坐标。

将所有图像变换到同一坐标系并对齐后,各图像存在重叠区域,直接叠加可能会产生亮度跳变、纹理断层等现象。为此,采用加权融合策略,权重函数定义为高斯衰减模型。

$$\omega_i(x, y) = \exp\left[-\frac{\|(x, y) - (x_c, y_c)\|^2}{2\sigma^2}\right] \quad (5)$$

式(5)中:\$\omega_i(x, y)\$为基于高斯衰减模型定义的权重函数;\$(x, y)\$为像素坐标位置;\$(x_c, y_c)\$为图像中心坐

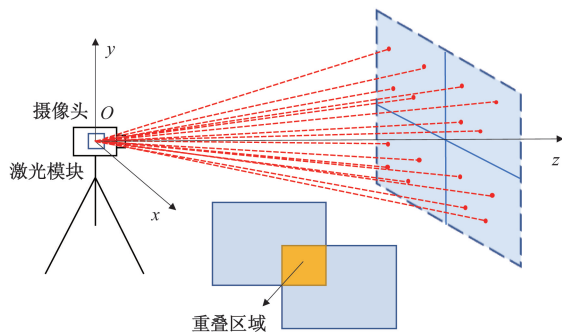


图1 图像拼接示意图

Fig. 1 Schematic diagram of image stitching

标;\$\sigma\$为控制权重衰减速度。

应用拉普拉斯金字塔多频带融合,低频段平滑过渡,高频段保留门窗边缘细节,确保全景图无缝拼接。采用本文提出的方法获得的拼接结果如图2所示。



图2 基于深度参数的图像拼接结果

Fig. 2 Image stitching results based on depth parameters

2 室内测量设备

为验证上述原理的可实施性,设计了一种基于视觉与激光融合的室内实测量机器人,旨在实现房屋尺寸的高精度测量的同时,重点解决门窗尺寸测量的关键需求。如图3(a)所示,设备基于树莓派4B主控制器,视觉模块搭载1300万像素摄像头(分辨率\$3840 \times 2880\$),激光测距模块精度为\$\pm 1\$ mm。系统选择一块3600 mAh的锂电池作为电源,电压为直流24 V。电源接通后,通过降压模块将DC24V电压转换为5 V,用于为控制系统、相机与激光模块供电。当系统启动时,Copley驱动器使用CAN通信控制精密电机配合运动,树莓派4B调用激光与相机来获取室内门窗目标的图像,最后将拍摄照片发送到YOLOv11-CUA网络中进行推理,并结合LSD(line segment detector)直线特征检测结果实现门窗尺寸测量。

如图3(b)和图3(c)所示,自主研发的建筑测量机器人整体采用T字形结构设计,由两个正交装配的高精度伺服电机驱动水平与俯仰转台运动,激光模组A、B整体呈左右对称排布,并与俯仰转台同步旋转。摄像头与激光模组C平行布置,并与水平转台同步旋转,实现墙面图像的拍摄与四点同步测距,能够实时实现图像像素与物理尺寸的精确转换。其中:设备支架臂长尺寸\$L\$为设备核心支撑结构的固定长度;\$(X_c, Y_c, Z_c)\$为相机坐标系下的三维空间坐标轴,水平转台旋转角度\$\theta\$和俯仰转台旋转角度\$\gamma\$分别表示设备绕垂直轴(\$Y\$轴)的旋转角度

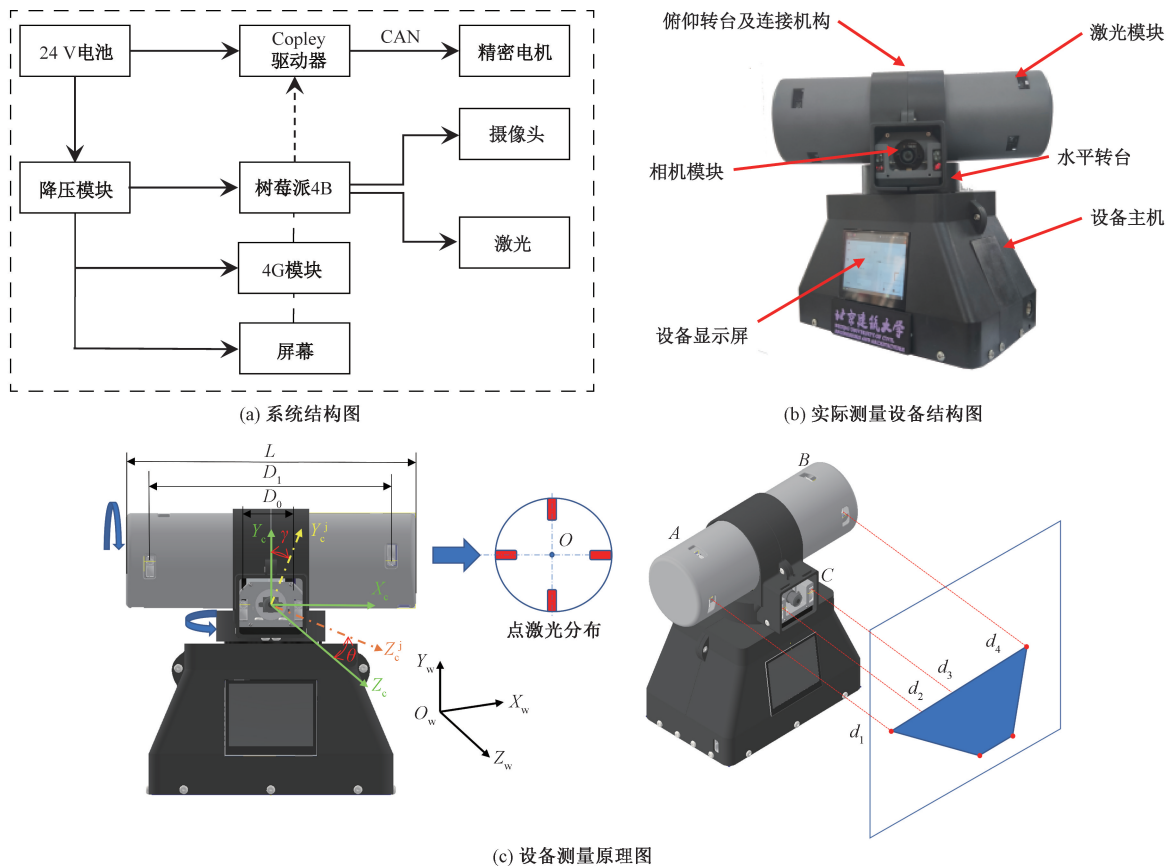


图3 建筑测量设备

Fig. 3 Building surveying equipment

和绕水平轴(X 轴)的旋转角度, Y_e^i 为俯仰转台绕水平轴(X 轴)旋转 γ 角度的坐标轴, Z_e^j 为水平转台绕 Y 轴旋转角度 θ 的坐标轴;激光测距值 d_1 、 d_2 、 d_3 、 d_4 表示多个激光测距模块到目标墙面的测量值;激光安装距离 D_0 、 D_1 为各测距模块的固定相对位置, (X_w, Y_w, Z_w, O_w) 为世界坐标系。

3 改进的 YOLOv11 模型

3.1 YOLOv11 简介

为提升门窗目标检测精度,进而完成门窗尺寸测量,对 YOLOv11n 模型进行改进。YOLOv11 模型是最新一代实时目标检测模型^[15]。相较于前几代 YOLO,它引入了 C3K2 模块代替了 YOLOv8 的 C2f 结构,旨在提高特征表示的有效性。还增加了 C2PSA 模块,这是一种新的结构,包含了注意力机制,旨在提高特征表示的有效性。通过 C2PSA 模块集成注意力机制,使模型将注意力集中在图像的关键部位,从而提高检测的准确性和效率。在检测头内部替换了两个 DWConvs,这一修改在保持检测精度的同时减少了参数数量和计算负荷。通过前面提到的结构优化和注意力机制的加入,模型的处理速度和精度都得到了提升,尤其是 GPU 上的优化,

使得训练和推理更加高效。

3.2 YOLOv11-CUA

针对建筑门窗识别算法存在的精度低、计算量大的问题,以轻量级的 YOLOv11n 模型为基线架构进行了优化。首先,将传统的 C3k2 模块替换为 C3k2_UIB 模块,通过统一倒置瓶颈(UIB)增强通道扩展与压缩,优化梯度流动,提升目标检测能力。随后,ADown 替换传统 Conv,使用并行最大注意力机制,动态调整特征权重,保留关键细节,适合密集或遮挡场景,提高了资源效率,使模型更轻量化,更适合室内测量设备。将此修改版本在本文中命名为 YOLOv11-CUA (YOLOv11n-C3k2_UIB-ADown),其完整组织架构如图 4 所示。

3.2.1 C3k2_UIB 模块

在 YOLOv11 架构中,C3k2 模块作为一个特征提取模块,主要作用于 Backbone 和 Neck 结构中,其核心功能是提取不同尺度的关键目标信息。为使其在保持高准确性的同时实现轻量化设计,引入了源自 MobileNetV4 的 UIB (universal inverted bottleneck) 结构。UIB 模块通过扩展-深度可分离卷积-压缩的流程实现高效特征表达^[16]。构建了如图 5 所示的 C3k2_UIB 模块。该模块通过引入两个可选的

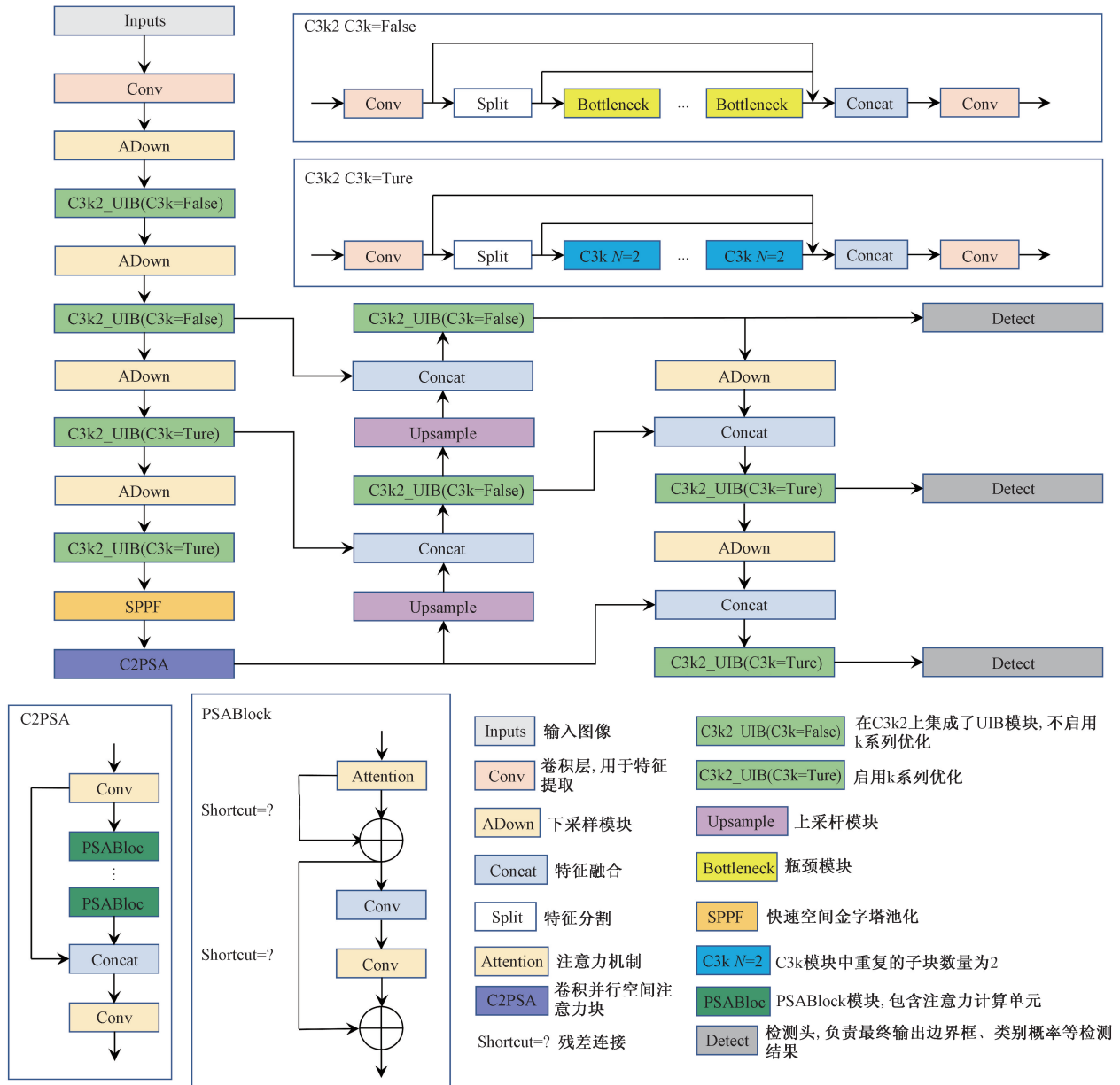


图4 YOLOv11-CPA 组织架构图

Fig. 4 YOLOv11-CPA organizational chart

深度可分离卷积,有效增强了特征提取过程的灵活性与运行效率。为了增强捕获上下文信息的能力,该结构在减轻网络的同时扩展了该层结构的感受野。相较于传统的瓶颈结构,UIB在计算效率和模型轻量化方面具备显著优势,同时可以增强对细节特征的代表能力。具体而言,第一个深度卷积置于扩展层之前,增强对低层次空间细节的捕获能力;第二个深度卷积插入扩展层与投影层之间,增强网络提取高级语义特征的能力。有效地提升了该模型在复杂视觉条件下对门窗目标特征的识别能力,同时兼顾了计算效率与精度,强化了YOLOv11在门窗检测场景下的应用。

3.2.2 ADown 轻量化下采样模块

传统YOLOv11n主干网络采用的步长卷积与最大池化结构在降维过程中容易导致有效信息丢失,进而影响后续特征融合与检测分支的判别性能^[17]。为此,设计了一种高效的、可学习的轻量化自适应下采样模块ADown,用以替代原有卷积下采样路径,以提升模型对关键区域的特征表达能力,减少墙面门窗等低对比度场景下的细节丢失。其结构图如图6所示,处理流程为:首先,对输入通道数为 C 的特征图执行二维平均池化层初步下采样处理后,获得紧凑中间 x ,经过通道切片操作将其沿通道维度均匀分割为通道数为 $C/2$ 的两个特征图 x_1 与 x_2 ;接

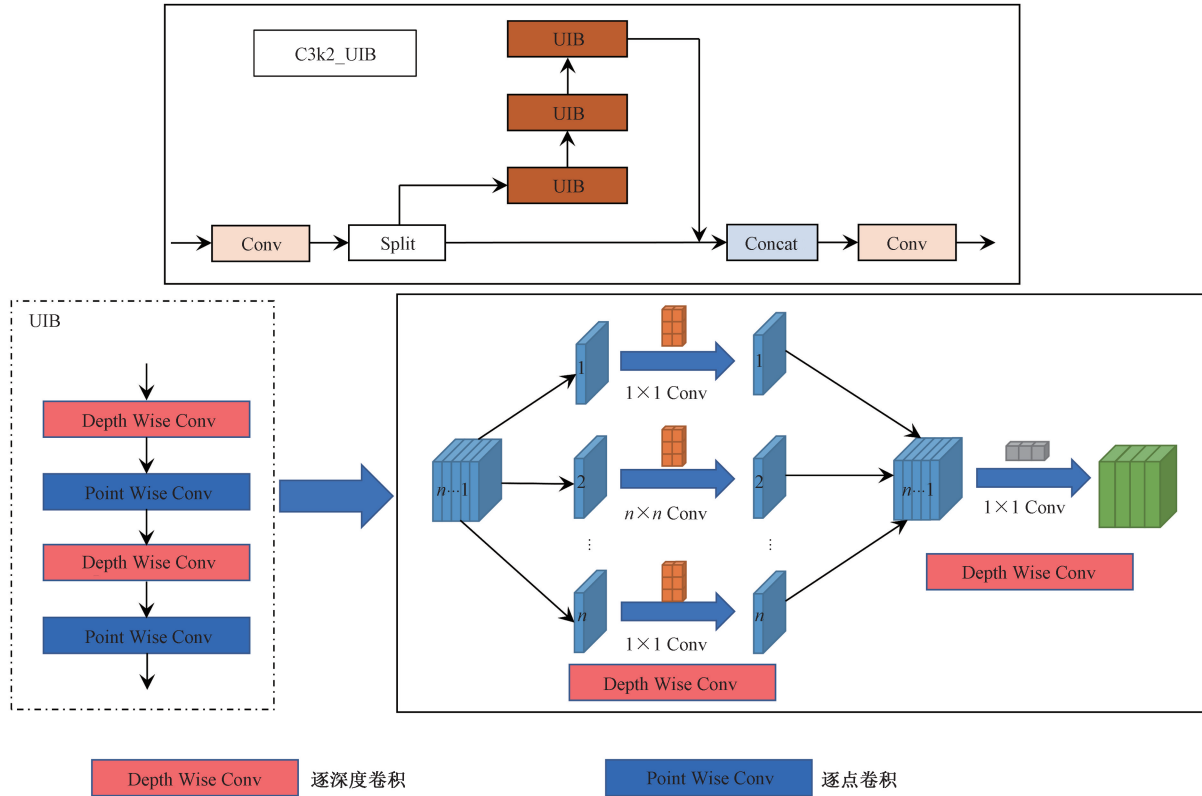


图5 C3k2_UIB 结构图

Fig. 5 Structure diagram of C3k2_UIB

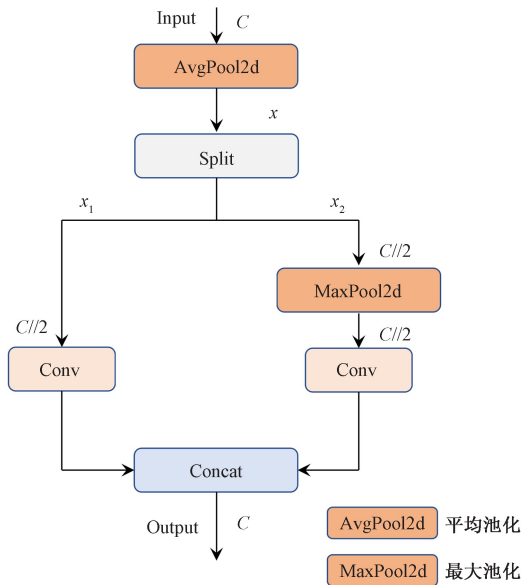


图6 ADown 模块结构

Fig. 6 Structure of the ADown module

着,将其中一个特征图通过二维最大池化层处理来约束其通道数为 $C/2$,并将两个特征图分别经过 Conv 模块处理,最后拼接输出一个与原通道数相同的特征图。通过平均池化与最大池化的有效结合,有助于在降低环境干扰和噪声的同时,保留门窗等

目标的连续性纹理特征,并通过局部显著性增强,突出门窗框等关键几何结构的边缘特征。此外,不同特征的拼接进一步增强了模块对多尺度特征的感知与表征能力,从而提升模型在多样化场景下的检测性能。

主分支通过带有步长的卷积操作完成降采样与局部上下文建模,其输出为

$$F_1 = \text{Conv}_{3 \times 3}^{s=2, p=1}(X_1) \quad (6)$$

式(6)中:Conv 表示核大小为 $k \times k$ 、步长为 s 、填充为 p 的卷积操作。

辅助分支首先通过最大池化操作提取边缘与强响应区域特征,即

$$F_2 = \text{Conv}_{1 \times 1}^{s=1, p=0}[\text{MaxPool}_{2 \times 2}(X_2)] \quad (7)$$

$$Y = \text{Concat}(F_1, F_2) \quad (8)$$

4 门窗尺寸测量

基于上述的门窗采集或者拼接结果,在 YOLOv11-CUA 目标检测后引入直线特征检测步骤,基于精炼后的线段进行物理尺寸计算。

4.1 门窗识别与直线特征检测

使用 YOLOv11-CUA 模型对获取的门窗图进行目标检测,输出门窗的边界框,该步骤提供粗略的门窗定位,为后续直线特征检测提供感兴趣的区域

(region of interest, ROI)。相比于传统的 Hough 变换, LSD 无须参数优化且具有线性时间复杂度, 能够实时检测有限长度的线段, 并在噪声环境下表现出色, 更适合建筑门窗的矩形框结构检测, 因此, 在检测到的边界框内通过 LSD 直线检测算法进行门窗洞口的直线目标检测, 以确认门窗的直线框边界。

LSD 直线特征检测流程如下。

(1) 门窗图像处理: 对设备采集到的门窗图像进行降采样处理; 采用高斯滤波对图像进行平滑处理来抑制锯齿效应与噪声, 为后续的直线特征准确提取奠定基础。

(2) 图像梯度求解与水平线场构建: 对 ROI 区域的灰度图像 $I(x, y)$ 计算每个像素的梯度幅度 $g(x, y)$ 和方向 $\theta(x, y)$, 如式(9)~式(11)所示。

$$\begin{cases} g_x = \frac{\partial I}{\partial x} \\ g_y = \frac{\partial I}{\partial y} \end{cases} \quad (9)$$

$$g(x, y) = \sqrt{g_x^2 + g_y^2} \quad (10)$$

$$\theta(x, y) = \arctan \frac{g_y}{g_x} \quad (11)$$

然后构建水平线场, 即每个像素的水平线角度 $\varphi(x, y) = \theta(x, y) + \pi/2$, 来表示潜在线段方向。

(3) 区域生长与线支持区域分割: 首先依据梯度幅值排序选取种子像素, 然后根据水平线方向相似性准则在八邻域内进行区域生长, 聚合形成空间连通且方向一致的像素集合, 即线支持区域(line support region, LSR)。

(4) 矩形估计与线段精确提取: 该步骤首先对已检测到的直线支撑区域进行最小外接矩形拟合, 进而计算该矩形的核心几何参数(包括中心、方向、长宽), 最终利用这些参数实现对直线的精确参数化描述。

(5) 线段的有效性验证: 使用 NFA (number of false alarms) 来对线段的有效性进行验证。NFA 计算公式为

$$NFA(l) = N^4 (1-p)^k \binom{M}{k} \quad (12)$$

式(12)中: N 为图像尺寸; p 为随机对齐概率; k 为对齐像素数; M 为 LSR 像素总数; 若 $NFA(l) < \varepsilon$ ($\varepsilon = 1$), 则保留线段, 否则丢弃。

4.2 门窗测量公式

基于 LSD 检测的直线基元, 结合激光测距所获得的到门窗所在墙面距离 d 和相机内参矩阵 $\mathbf{K}$ 中的焦距 f (像素单位, 通过张正友标定法获得), 应用比例转换公式将像素尺寸转换为物理尺寸。

$$d_{\text{physical}} = d_{\text{pixel}} \frac{d}{f} \quad (13)$$

式(13)中: d 为激光测距模块到墙面的距离; d_{pixel} 为像素宽度或高度; d_{physical} 为对应的物理尺寸。

5 实验分析

为验证上述 YOLOv11-CUA 模型的有效性, 通过自建数据集进行门窗检测与实际测量尺寸实验分析。

5.1 数据集创建

为训练 YOLOv11 模型以实现高效的门窗检测, 构建了一个专为门窗目标检测设计的数据集, 旨在提升模型在建筑场景中对门窗目标的识别能力。数据集包含 1 200 张分辨率均为 $1\,920 \times 1\,080$ 的原始图像, 这些图像皆由高分辨率相机在真实场景中采集所得。采集场景涵盖多样化的建筑环境, 主要包括住宅、商业建筑等, 覆盖多种光照条件和天气条件。为保证数据质量, 所有图像经过严格筛选, 确保门窗目标清晰可见且背景具有代表性, 剔除了模糊或严重遮挡的图像。为了模型的训练、优化与评估能够有可靠的数据基础, 本数据集将图像分配至训练集、验证集和测试集 3 个子集, 其数量对应比例为 3:1:1, 分别为 720、240、240 张。数据标注通过采用 Roboflow 标注工具, 生成符合标准化的 YOLO 格式要求的标注文件, 所有标注均经过二次校核, 确保边界框的精确性和一致性, 特别针对小型门窗目标加强标注, 充分考虑门窗检测任务对边缘细节和语义信息的保留需求, 以提升模型的检测能力。

5.2 实验配置及评估体系

设计了相应的实验进行模型训练与评估, 具体的实验环境配置以及训练阶段的关键参数如表 1 所示。

选取了精确率 P 、召回率 R 、平均精度均值 mAP 以及每秒帧数 FPS 作为全面评估模型性能的核心

表 1 实验配置

Table 1 Experimental configuration	
名称	系统参数
中央处理器 (CPU)	Core (TM) i9-10900K
图像处理器	GTX3080Ti
编程软件	Python 3.9.7
深度学习框架	Pytorch1.12.1
优化器	SGD
CUDA 环境	CUDA11.3
训练轮次	150
批量大小	16
输入图像尺寸	640 × 640

指标,用以衡量该模型的推理检测精度与效率。各指标的计算公式为

$$P = \frac{TP}{TP + FP} \times 100\%$$

(14)

$$R = \frac{TP}{TP + FN} \times 100\%$$

(15)

$$AP = \int_0^1 P(R) dR$$

(16)

$$mAP = \sum_{i=1}^N \frac{AP(i)}{N}$$

(17)

$$FPS = \frac{1\ 000}{Pre + Infer + Post}$$

(18)

式中:TP 为真阳性,代表正确预测的门窗数量;FP 为假阳性,代表错误预测的非门窗数量;FN 为假阴性,代表错误预测的门窗数量;AP 为目标检测的准确性;mAP 为所有类别检测的平均精度;N 为类别总数;FPS 用于量化检测速度,即算法每秒能够检测的图像帧数(包括前处理 Pre、模型推理 Infer、后处理 Post 的总时间)。

5.3 对比实验

为了验证本文改进算法的优越性和有效性,在所作数据集上将 YOLO11-CUA 网络模型与其他 YOLO 系列主流算法进行比较(表 2)。从表 2 可知,改进后的算法在各个评估指标上均获得了明显提升。相较于 YOLOv7-tiny、YOLOv8n、YOLOv9n、YOLOv10n 和 YOLOv11n 5 种算法,YOLO11-CUA 的精确率 P 分别提高了 10.8%、7.9%、10.2%、7.3%、4.7%;召回率 R 分别提升了 12.1%、8.6%、7.7%、5.7%、6.6%;相较于基线模型 YOLOv11n,mAP50 提高了 4.8%,参数量减少了 0.6 M,能够满足门窗洞实时检测的要求。综上,经过改进后的 YOLOv11 模型推理速度相对较快,检测精度高,且漏检率下降,相较其余 5 种算法存在明显优势。

表 2 不同检测算法对比

Table 2 Comparison of different detection algorithms

模型	精确率 $P/\%$	召回率 $R/\%$	$mAP_{50}/\%$	参数量/ M	帧率/ FPS
YOLOv7-tiny	86.7	84.7	86.2	6.2	130.0
YOLOv8n	89.6	88.2	87.1	3.4	142.8
YOLOv9	87.3	89.1	88.3	5.7	142.0
YOLOv10n	90.2	91.7	90.1	3.8	145.1
YOLOv11n	92.8	90.2	92.4	2.6	149.3
YOLOv11-CUA	97.5	96.8	97.2	2.0	151.5

注:mAP₅₀为 IoU(intersection over union)阈值为 0.5 时的结果。

5.4 消融实验

为了验证增加的模块对模型整体所产生的影响,对基线模型进行了两个改进模块的消融实验。

如表 3 所示,从基线模型开始,首先使用 C3k2_UIB 模块改进 C3k2 模块,模型的准确率 P 、召回率、mAP₅₀皆有所上升,FPS 产生轻微下降,结果表明仅靠加入这一个模块,模型改进效果一般;其次将下采样部分替换成 ADown 模块,模型的准确率 P 、召回率、mAP₅₀几乎保持不变,但 FPS 明显上升,实现了模型的轻量化处理。将 C3k2_UIB 和 ADown 纳入基线模型后,模型的识别准确率和 mAP₅₀分别达到了 97.5% 和 97.2%,相较于原始的 YOLOv11 模型基线,准确率和 mAP₅₀分别提高了 4.7% 和 4.8%,参数数量减少了约 0.6 M,但造成 FPS 产生轻微下降。综上所述,YOLO11-CUA 模型虽 FPS 产生一定下降,但仍旧保持较高水平,且其余各项指标均呈现显著提升,说明整体上加入的每个模块对最终结果都产生了正向影响,模型改进有效。

表 3 消融实验结果

Table 3 Results of ablation experiment

YOLOv11	C3k2_UIB	ADown	准确率 $P/\%$	召回率 $R/\%$	$mAP_{50}/\%$	参数 量/M	帧率/ FPS
✓			92.8	90.2	92.4	2.6	149.3
✓	✓		93.1	91.3	93.7	2.5	131.5
✓		✓	94.2	93.6	92.4	2.1	169.5
✓	✓	✓	97.5	96.8	97.2	2.0	151.5

注:✓表示采用该项改进措施。

5.5 收敛分析

通过分析图 7 的训练过程,发现该模型边界盒损失、分类损失和分布焦点损失的收敛性都很好,其准确率、召回率和平均精度在大约 100 次 epochs 后趋于稳定,在 150 次 epochs 时停止训练。

6 结果验证

6.1 验证实验

在自建数据集上进行了实验来验证 YOLOV11-CUA 算法的有效性。P-R 曲线如图 8 所示。很明显,YOLOV11-CUA 比 YOLOv11 模型取得了更好的检测性能。为了验证门窗洞目标智能识别系统的性能,进行了系统实验,实验结果如图 9 所示,识别准确率超过 90%,满足了设计要求。

6.2 LSD 直线特征检测

如图 10 所示,针对每条直线特征计算其角度、长度以及原点到直线的距离。

6.3 测量精度验证

在真实建筑场景中选取了 4 个门窗样本进行测量测试来验证设备的实际测量精度。真实尺寸为人工采用专业激光测距仪获得。系统测量基于校正后的全景图像和 YOLOv11-CUA 检测结果计算,如图 11 所示。表 4 展示了测量结果,平均误差与最

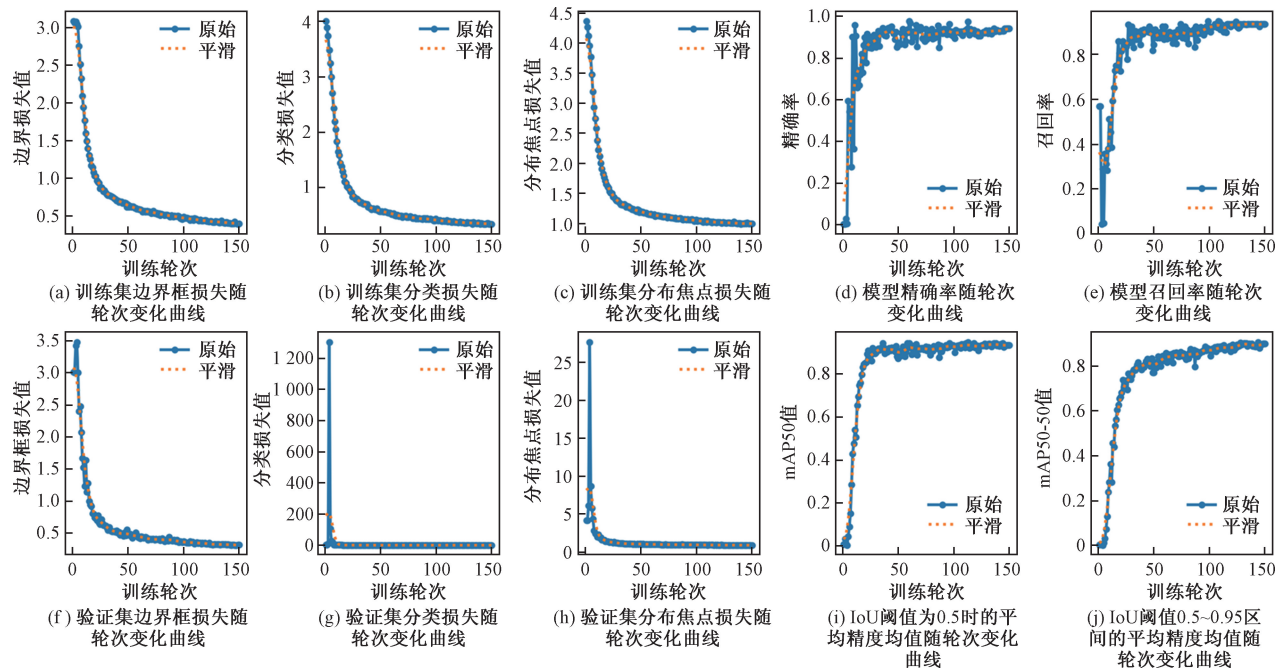


图7 训练过程的收敛曲线
Fig.7 Convergence curve of the training process

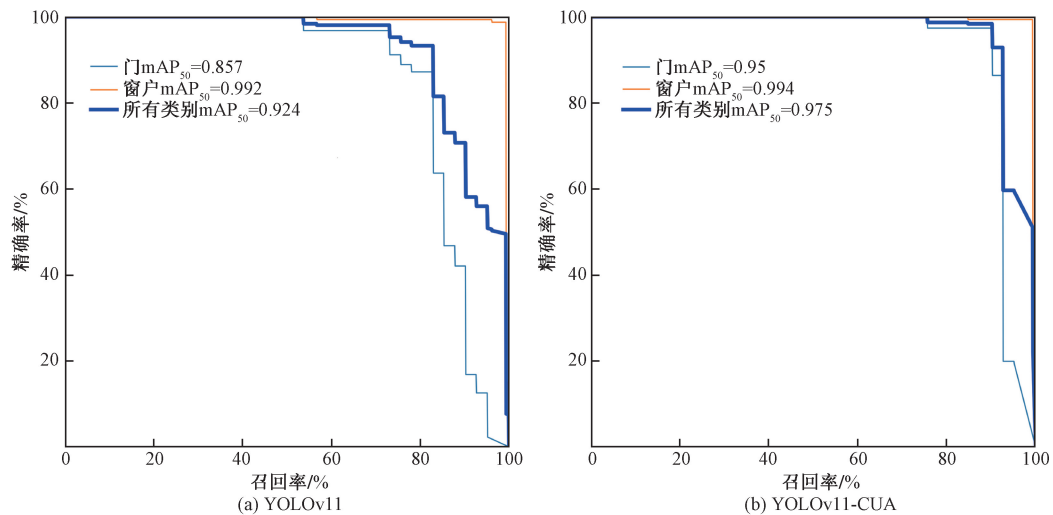


图8 YOLOv11n与YOLOv11n-CUA在数据集上的曲线比较
Fig.8 Comparison of the curves of YOLOv11n and YOLOv11n-CUA on the dataset

表4 窗尺寸测量精度验证

Table 4 Verification of measurement accuracy for window dimensions

样 本	真实 宽度/mm	测量 宽度/mm	宽度 误差/mm	真实 高度/mm	测量 高度/mm	高度 误差/mm
1	1 615	1 623	8	1 507	1 502	-5
2	1 523	1 517	-6	1 409	1 413	4
3	1 515	1 520	5	1 507	1 514	7
4	1 202	1 196	-6	1 007	1 011	4

大误差分别为 5.63 mm 与 8 mm,完全符合建筑测量规范。

7 结论

(1)利用自主研发的建筑测量机器人,采集含有深度信息的多视角图像,并依据深度信息与位姿信息完成图像校正与拼接,构建包含待测门窗的墙面全景图。

(2)针对门窗测量的这一关键环节,提出了一种基于 YOLOv11-CUA 的门窗检测测量方法。在骨干网络和特征金字塔中引入 C3K2_UIB 模块来增强多尺度特征捕获能力;采用 ADown 模块替换下采样卷积,增强了关键区域的表征能力,减少了墙面等



图9 检测目标可视化
Fig.9 Visualization of the detection target



图10 LSD 直线特征检测
Fig.10 LSD linear feature detection

低对比度场景下的细节丢失。

(3) 在自建门窗洞口数据集上,改进的YOLOv11 算法的门窗检测的准确率为 97.5%, mAP_{50} 达到了 97.2%,参数量仅为 2.0 M。与原始YOLOv11n 网络相比,准确率提高了 4.7%, mAP_{50} 提高了 4.8%,参数数量减少了 0.6 M。

(4)实验表明,窗框测量的最终尺寸平均测量误差仅为 5.63 mm,最大误差小于 10 mm,完全符合



图11 窗框检测可视化结果
Fig.11 Visualization results of window inspection

建筑测量规范要求。

为智能建造领域的自动化测量提供了一种全新的高效自动化测量方法。

参 考 文 献

- [1] 韩鹏圆. 建筑工程测量中测绘新技术的应用[J]. 中华建设, 2025(2): 131-133.
Han Pengyuan. Application of new surveying and mapping technologies in construction engineering surveying[J]. China Construction, 2025(2): 131-133.
- [2] 郝亚东, 赵立峰, 刘士彪, 等. 三维激光扫描技术在房屋建筑“多测合一”中的应用[J]. 北京测绘, 2023, 37(3): 371-376.
Gao Yadong, Zhao Lifeng, Liu Shibiao, et al. Application of 3D laser scanning technology in the integration of multiple measurements in building construction[J]. Beijing Surveying and Mapping, 2023, 37(3): 371-376.
- [3] 赵辉, 韩志伟. 基于深度学习的图像特征匹配方法研究进展[J]. 计算机科学, 2023, 50(1): 144-150.
Zhao Hui, Han Zhiwei. Research progress of image feature matching method based on deep learning[J]. Computer Science, 2023, 50(1): 144-150.
- [4] Zolanvari S M I, Laefer D F, Natanzi A S. Three-dimensional building façade segmentation and opening area detection from point clouds[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018, 143: 134-149.
- [5] 莫玉晓, 徐景中. 基于空洞检测及规则约束的立面点云门窗检测[J]. 激光杂志, 2025, 46(1): 61-67.
Mo Yuxiao, Xu Jingzhong. A method for detecting doors and windows in facade point clouds based on cavity detection and rule constraints[J]. Journal of Lasers, 2025, 46(1): 61-67.
- [6] 郑晓昱, 周惠兴, 王舜, 等. 基于激光点标定和机器视觉的门窗洞尺寸测量[J]. 科学技术与工程, 2022, 22(15): 6204-6209.
Zheng Xiaoyu, Zhou Huixing, Wang Shun, et al. Measurement of door and window opening dimensions based on laser point calibration and machine vision[J]. Science Technology and Engineering, 2022, 22(15): 6204-6209.
- [7] 高云龙, 张帆, 屈孝志, 等. 结合样本自动选择与规则性约束的窗户提取方法[J]. 武汉大学学报(信息科学版), 2018, 43(3): 436-443.
Gao Yunlong, Zhang Fan, Qu Xiaozhi, et al. A window extraction method combining automatic sample selection with regularity constraints[J]. Journal of Wuhan University (Information Science Edition), 2018, 43(3): 436-443.
- [8] 沈乐, 李桂清, 冼楚华, 等. 室内3D点云模型的门窗检测[J]. 计算机辅助设计与图形学学报, 2019, 31(9): 1494-1501.
Shen Le, Li Guiqing, Xian Chuhua, et al. Door and window detection based on indoor 3D point Cloud model[J]. Journal of Computer-Aided Design & Computer Graphics, 2019, 31(9): 1494-1501.
- [9] Cheng S, Han Y, Wang Z, et al. An underwater object recognition system based on improved YOLOv11[J]. Electronics, 2025, 14: 14010201.
- [10] Chen Y, Zhang C, Chen B, et al. Accurate leukocyte detection based on deformable-DETR and multi-level feature fusion for aiding diagnosis of blood diseases[J]. Computers in Biology and Medicine, 2024, 170: 107917.
- [11] 吐尔逊·买买提, 邱建卓, 刘健, 等. 基于改进YOLOv8n的道路裂缝检测算法[J]. 科学技术与工程, 2025, 25(14): 6044-6053.
Tursun Maimaiti, Qiu Jianzhuo, Liu Jian, et al. Road crack detection algorithm based on improved YOLOv8n[J]. Science Technology and Engineering, 2025, 25(14): 6044-6053.
- [12] He L, Zhou Y, Liu L, et al. Research and application of YOLOv11-based object segmentation in intelligent recognition at construction sites[J]. Buildings, 2024, 14: 3777.
- [13] Wang D, Tan J, Wang H, et al. SDS-YOLO: an improved vibratory position detection algorithm based on YOLOv11[J]. Measurement, 2025, 244: 116518.
- [14] 许向阳, 袁杉杉, 王军, 等. 基于全局和局部特征的图像拼接方法[J]. 北京理工大学学报, 2022, 42(5): 502-510.
Xu Xiangyang, Yuan Shanshan, Wang Jun, et al. Image stitching method based on global and local features[J]. Journal of Beijing Institute of Technology, 2022, 42(5): 502-510.
- [15] Khanam R, Hussain M. YOLOv11: an overview of the key architectural enhancements[J]. arXiv Preprint, arXiv: 2410.17725, 2024.
- [16] 胡道婕, 倪翠, 吴春坡, 等. 基于无人机视角的轻量化绝缘子缺陷检测算法[J/OL]. [2025-09-19]. 北京航空航天大学学报. <https://doi.org/10.13700/j.bh.1001-5965.2025.0495>.
Hu Xiaojie, Ni Cui, Wu Chunpo, et al. Light weight insulator fault detection algorithm based on UAV perspective [J/OL]. [2025-09-19]. Journal of Beihang University. <https://doi.org/10.13700/j.bh.1001-5965.2025.0495>.
- [17] 刘英, 范雅慧, 岳辉. YOLO11-DAE: 高分遥感图像露天煤矿复杂场景目标识别方法[J]. 地球信息科学学报, 2025, 27(9): 2230-2249.
Liu Ying, Fan Yahui, Yue Hui. YOLO11-DAE: target recognition method for complex scenes in open-pit coal mines from high-resolution remote sensing images[J]. Journal of Earth Information Science, 2025, 27(9): 2230-2249.