

DOI:10.13203/j.whugis20200158



文章编号:1671-8860(2020)11-1660-09

# 利用具有注意力的Mask R-CNN检测震害建筑物立面损毁

眭海刚<sup>1</sup> 黄立洪<sup>1</sup> 刘超贤<sup>1</sup>

1 武汉大学测绘遥感信息工程国家重点实验室,湖北 武汉,430079

**摘要:**震后建筑物损毁信息是灾情快速评估和应急救援的重要决策依据之一。针对传统的建筑物损毁遥感检测技术只关注于顶面信息,导致众多顶面结构完好而中间层、底层倒塌或崩裂的损毁建筑物处于检测盲点的问题,提出了一种融合引入注意力机制的深度学习实例分割模型和图像多尺度分割算法进行震后建筑物立面损毁检测的方法。首先,利用卷积块状注意力模块(convolutional block attention module, CBAM)改进掩模区域卷积神经网络(mask region-based convolutional networks, Mask R-CNN)模型实现了复杂建筑物立面背景中的损毁信息提取;然后,基于建筑物立面影像多尺度分割结果,利用多数投票规则实现了损毁检测结果的后处理优化。实验结果表明,相比传统损毁检测方法,所提方法能够更有效地实现震后建筑物立面损毁信息的精准定位,总体准确率可达到89.15%。

**关键词:**地震;建筑物立面损毁;地面影像;Mask R-CNN;注意力机制;多尺度分割

中图分类号:P237;P208 文献标志码:A

近年来,地震灾害频发,破坏力巨大的强震给人类社会带来了无法弥补的物质和精神损失。建筑物损毁信息作为灾区的典型震害特征,能够为开展震后灾情评估、应急响应以及恢复重建等工作提供重要的决策依据<sup>[1]</sup>。飞速发展的遥感技术凭借全时、全域、高动态、精细化的对地观测能力,在减灾救灾方面的应用逐渐得到重视<sup>[2-3]</sup>。

地震发生后,受制于卫星、航空遥感影像传感器飞行高度和拍摄视角的影响,灾区常存在顶面完好而中间层、底层倒塌或崩裂的损毁建筑物被漏检的情况,如图1所示。而随着传感器技术的发展,利用地面图像传感器对震后建筑物进行近距离影像采集以及精细化损毁检测逐渐成为可能。与卫星、航空遥感影像数据相比,地面影像能够有效克服前者存在的地物遮挡和图像变形等问题,从而为实现建筑物全方位精细化损毁检测提供必要的地面支持。且随着近几年众源地理数据<sup>[4]</sup>的飞速发展,地面影像数据的获取成本低、信息海量和实时性强等特点使其在震后建筑物立面损毁检测中具有广阔的应用前景。

从现有的建筑物立面损毁检测研究来看,相

关学者主要从单时相的角度提出了两种类型的检测方法:(1)基于像素分类的建筑物立面损毁检测。该方法又分为基于纹理特征和统计特征的损毁检测,基于纹理特征的方法主要利用灰度共生矩阵(gray level co-occurrence matrices, GL-CM)及其统计属性进行损毁检测,但受限于建筑物立面背景的纹理复杂度和粗糙度,不适用于检测单时相建筑物立面损毁信息<sup>[5]</sup>;基于统计特征的方法通过利用经济学中的基尼系数判别门窗等部件元素是否整齐排布来判定建筑物立面是否损毁,但其应用场景较为局限,泛化能力较弱<sup>[6]</sup>;(2)面向对象(object-based image analysis, OBIA)分类的建筑物立面损毁检测。该方法首先将建筑物立面分割成多个超像素块,然后通过视觉词袋模型(bag of visual words, BOVW)等手段对其提取损毁特征并进行分类,但由于建筑物立面材质和纹理结构等丰富各异,因此难以制定一套通用的图像分割参数和特征描述标准<sup>[7-9]</sup>。总体上,现有的单时相建筑物立面损毁检测方法主要基于小样本数据开展相关研究,模型的检测能力和泛化能力均存在较大提升空间,且现有研

收稿日期:2020-04-15

项目资助:国家自然科学基金(41771457);国家重点研发计划(2018YFB10046)。

第一作者:眭海刚,博士,教授,主要从事遥感、GIS以及灾害应急相关研究。haigang\_sui@263.net

通讯作者:刘超贤,博士。cx\_leo@whu.edu.cn

究对建筑物立面损毁检测结果的后处理优化方面仍缺乏足够讨论。

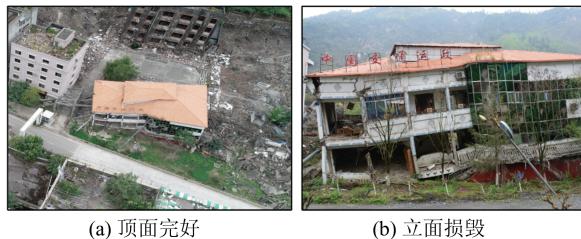


图 1 震后顶面完好而立面破损的建筑物  
Fig.1 Post-Earthquake Building with an Intact Roof Surface and Damaged Façade

针对以上问题,本文提出了一种融合深度学习实例分割模型和图像多尺度分割算法的震后建筑物立面损毁检测方法。其中:(1)样本集构建方面,通过前往地震遗址实地采集震害建筑物地面影像数据,并进行面向小样本的图像增广来保障模型的训练样本量;(2)模型检测方面,采用深度学习实例分割模型对建筑物立面影像进行损毁特征的深度挖掘和检测识别,能够从复杂的建筑物立面纹理背景中有效提取损毁信息;(3)后处理优化方面,使用建筑物立面影像的多尺度分割结果对其损毁检测结果进行融合处理,能够有效解决因卷积层数叠加所导致的损毁边界模糊问题。实验结果表明,该方法比现有方法精度更高,且能够从复杂建筑物立面背景中有效提取出损毁信息。

## 1 立面损毁检测算法流程

本文提出的震后建筑物立面损毁检测方法包括3个核心步骤:(1)基于实地采集的震后建筑物地面影像数据进行图像增广和样本集制备;(2)将样本集代入引入注意力机制的深度学习实例分割模型中进行训练,然后将待检影像代入训练后的模型中得到检测结果;(3)采用图像多尺度分割算法对待检影像进行同质像元聚类分割,然后基于多数投票规则将多尺度分割结果和模型检测结果进行融合处理,得到最终的建筑物立面损毁检测结果。立面损毁检测算法流程如图2所示。

### 1.1 面向小样本的图像增广

由于目前基于地面影像的震后建筑物立面损毁检测研究开展得较少,且尚无渠道获取对应的开源影像数据集,本文通过前往北川老县城地震遗址实地采集震害建筑物的地面影像数据来

进行研究。为了防止模型在训练过程中过拟合,并尽可能使数据集多样化,采用图像增广的方式对实地采集的小样本数据集进行样本集扩充。图像增广是指通过对训练图像作一系列随机改变来产生相似但又不同的训练样本,从而扩大训练样本集的规模,并提高模型的场景泛化能力。结合震后实地拍摄过程中可能存在因人为失误、设备老化和天气条件等影响所导致的图像模糊、图像噪声和图像过曝、欠曝等问题,本文分别对地面影像数据集添加高斯模糊、椒盐(高斯)噪声和高斯过曝、欠曝噪声等处理来进行基于实景模拟的图像增广(见图3)

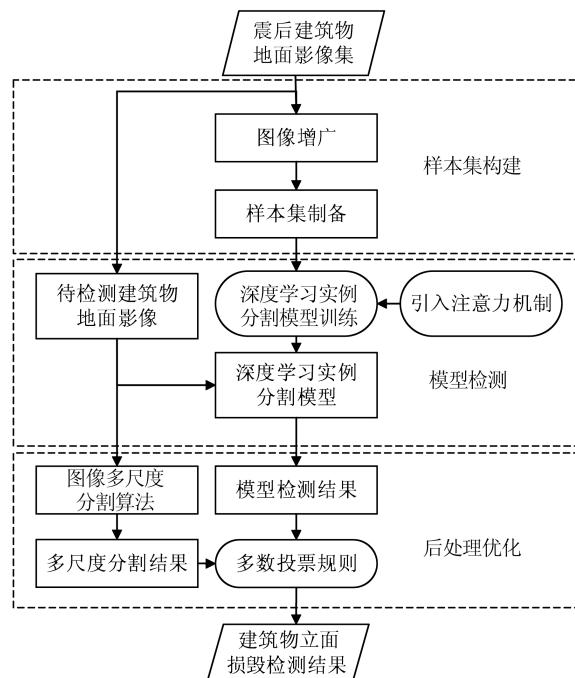


图 2 震后建筑物立面损毁检测流程图  
Fig.2 Flowchart of Post-Earthquake Building Facade Damage Detection

### 1.2 引入注意力机制的 Mask R-CNN

由于建筑物立面损毁信息处于一个相对复杂的背景环境中,使用传统的 GLCM 或 BOVW 等方法难以从建筑物立面背景中有效提取出损毁信息,因此本文选用文献[10]提出的 Mask R-CNN 模型作为震后建筑物立面损毁检测的核心算法。在结构上,Mask R-CNN 模型由更快速区域卷积神经网络(faster region-based convolutional networks, Faster R-CNN)模型和全卷积网络(fully convolutional networks, FCN)模型共同组成。在实现上,首先利用 Faster R-CNN 模型<sup>[11]</sup>从输入图像中检测出感兴趣目标并对其使用标注框定位;然后进一步利用 FCN 模型<sup>[12]</sup>对标注框

内的目标进行像素级的标注分类。由于 Mask R-CNN 模型本身具有较深的卷积层数,其对于建筑物立面损毁信息也具有较为深入的特征挖掘和学习识别能力。因此,相较传统方法,在使用

深度学习实例分割模型进行建筑物立面损毁检测时,能够更好地从复杂建筑物立面背景中有效提取出损毁信息,并对其进行准确定位和像素级标注。



图3 图像增广处理

Fig.3 Image Augmentation Processing

此外,由于建筑物立面不可避免地存在诸如窗、门、阳台等部件元素的干扰,为了使模型更好地聚焦于立面损毁信息,通过引入注意力机制的方式对 Mask R-CNN 模型进行改进。注意力机制本质上是一种聚焦于局部信息的机制,它能够定位图像中的感兴趣信息,并抑制无用信息,结果通常以概率图或概率特征向量的形式展示。本文采用文献[13]提出的卷积块状注意力模块(convolutional block attention module, CBAM),其本质上是一种空间和通道混合型注意力机制。在实现上,首先对输入特征图进行通道维注意力建模,通过对各个特征通道应用相应的注意力权重,学习到不同通道的重要性;然后对特征图进行空间维注意力建模,通过对特征图的不同空间位置应用相应的注意力权重,从而对感兴趣目标进行空间位置上的聚焦。通过引入 CBAM 注意力机制,能够使 Mask R-CNN 模型在训练和预测过程中更好地聚焦于图像中最重要的特征通道和空间位置。

具体地,选取 ResNet-101 网络<sup>[14]</sup>作为 Mask R-CNN 模型的特征提取网络,如图 4 所示。在特征提取阶段,首先对输入影像进行逐层卷积;然后由浅入深地提取其各个阶段所得到的特征图,并记作  $C_1, C_2, C_3, C_4, C_5$ 。在特征融合阶段,首先,对  $C_5$  进行  $1 \times 1$  卷积,输出  $P_5$ ;其次,对  $P_5$  进行  $2 \times 2$  上采样,并与经过  $1 \times 1$  卷积后的  $C_4$  进行对应元素求和操作,得到  $P_4$ ,以此类推,分别得到  $P_3$  和  $P_2$ ;然后对  $P_2, P_3, P_4, P_5$  分别进行  $3 \times 3$  卷积,以消除上采样操作中所带来的混叠效应;最后对  $3 \times 3$

卷积后的  $P_5$  进行池化,得到  $P_6$ ,并将最终输出的  $\{P_2, P_3, P_4, P_5, P_6\}$  作为区域候选网络(region proposal network, RPN)网络的输入特征用于后续训练。

上述过程中,  $P_5, P_4, P_3, P_2$  的计算公式(不包含  $3 \times 3$  卷积)为:

$$P_5 = \text{Conv}_{1 \times 1}(C_5) \quad (1)$$

$$P_i = \text{Conv}_{1 \times 1}(C_i) \oplus f_u(P_{i+1}), i \in \{4, 3, 2\} \quad (2)$$

式中,  $C_i$  表示 ResNet-101 网络第  $i$  阶段所提取的特征图;  $P_i$  表示  $C_i$  经过特征融合后输出的特征图;  $\text{Conv}_{1 \times 1}$  表示  $1 \times 1$  卷积层;  $\oplus$  表示对应元素相加操作;  $f_u$  表示  $2 \times 2$  上采样。

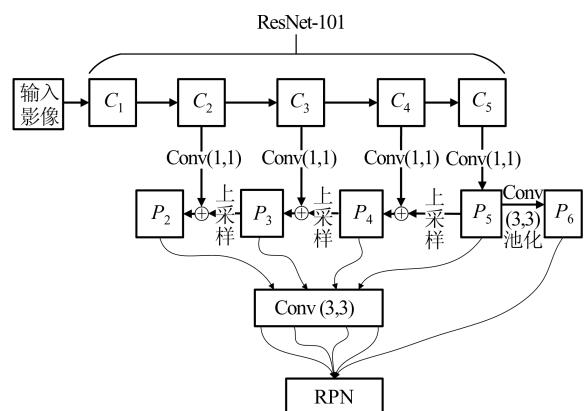


图4 Mask R-CNN 模型特征提取网络结构

Fig.4 Feature Extraction Network of Mask R-CNN

采用 CBAM 注意力机制对 Mask R-CNN 模型特征提取网络的特征融合阶段进行改造,通过在  $C_2, C_3, C_4, C_5$  的  $1 \times 1$  卷积层后分别添加

CBAM 注意力模块<sup>[15]</sup>,以提高 Mask R-CNN 模型对建筑物立面损毁信息的通道和空间聚焦能力,如图 5 所示。由于 CBAM 注意力模块直接嵌入到 Mask R-CNN 模型中,并参与整个模型的网络参数迭代更新,因此,随着模型的不断训练,CBAM 注意力权重矩阵也会不断优化并逐渐收敛得到最优权重矩阵。

改进后的  $P_5$ 、 $P_4$ 、 $P_3$ 、 $P_2$  计算公式为:

$$P_5 = \text{Conv}_{1 \times 1}(C_5) \otimes A_5(\text{Conv}_{1 \times 1}(C_5)) \quad (3)$$

$$P_i = \text{Conv}_{1 \times 1}(C_i) \otimes A_i(\text{Conv}_{1 \times 1}(C_i)) \oplus f_u(P_{i+1}), \quad i \in \{4, 3, 2\} \quad (4)$$

式中,  $A_i$  代表第  $i$  阶段特征图所对应的 CBAM 注意力权重矩阵;  $\otimes$  代表对应元素相乘操作。

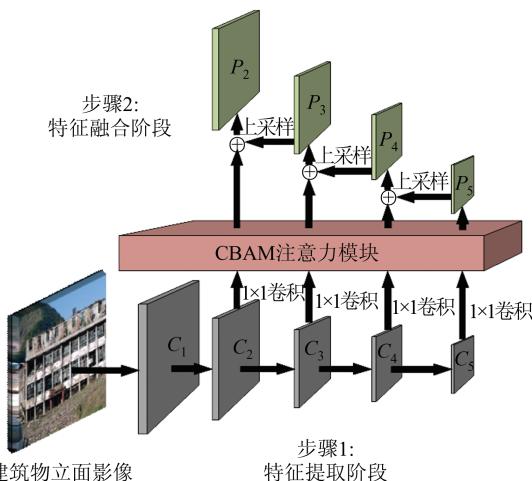


图 5 引入 CBAM 注意力机制的 Mask R-CNN 模型特征提取网络

Fig.5 Feature Extraction Network of Mask R-CNN Model Improved by CBAM Attention Mechanism

### 1.3 融合多尺度分割的检测结果优化

随着卷积神经网络层数的叠加,建筑物立面损毁的边界信息容易丢失。因此,基于建筑物立面影像的多尺度分割结果对其损毁检测结果进行后处理优化,从而借助多尺度分割算法的同质像元聚类属性,尽可能地保留建筑物立面损毁的边界信息。多尺度分割是指在影像信息损失最小的前提下,以任意尺度生成异质性最小、同质性最大的有意义多边形对象。通常分割尺度越大,所生成多边形分割对象的面积就越大而数目越少,反之则面积越小而数目越多。

借助 eCognition 软件中的多尺度分割工具,通过设置不同尺度参数对图像进行多种尺度的分割,并采用多数投票规则作为建筑物立面影像多尺度分割结果和损毁检测结果的融合处理算

法。首先,根据多尺度分割结果将建筑物立面影像划分为不同子区域;然后,计算各个子区域所对应的损毁检测结果中各类别(损毁、非损毁)的像元总数;最后,根据像元数的统计结果,将包含像元总数最多的类别作为该子区域的所属类别,如图 6 所示。计算公式为:

$$L_r = \arg \max_{c=1}^M \sum_i \sum_j \text{sgn}[f(r(i,j))] = c \quad (5)$$

式中,  $L_r$  为子区域  $r$  所属的类别;  $M$  为损毁检测结果的总类别数;  $r(i,j)$  为区域  $r$  内坐标为  $(i,j)$  的像元;  $f(r(i,j))$  为像元  $r(i,j)$  所属的类别;  $\text{sgn}(x)$  函数为数学上的符号函数,如果  $x$  为真,则  $\text{sgn}$  返回 1,否则,返回 0。

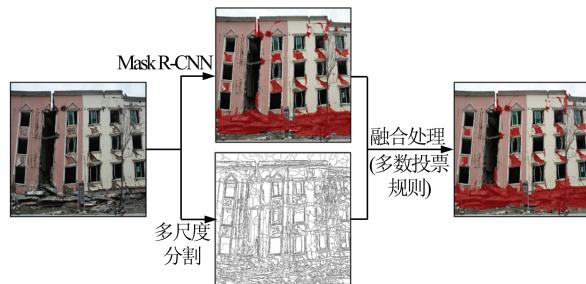


图 6 融合处理流程图  
Fig.6 Flowchart of Fusion Processing

## 2 震后建筑物立面损毁检测

### 2.1 研究区域

2008 年 5·12 汶川 Ms 8.0 地震,共导致 69 227 人遇难,374 643 人受伤,17 923 人失踪,所造成的建筑物损毁以及经济损失不计其数。按照四川省灾后重建总体规划,北川老县城作为地震遗址被保留下来,是目前全世界唯一整体原址原貌保护的规模最宏大、破坏类型最全面、次生灾害最典型的地震灾难遗址区,遗址区位如图 7 所示。研究组于 2019 年 8 月 12 日至 16 日前往该地震遗址开展了震害建筑物地面影像数据采集工作。

### 2.2 建筑物立面损毁检测

首先,对研究组在北川地震遗址采集的震害建筑物地面影像数据,分别添加高斯模糊、椒盐(高斯)噪声和高斯过曝、欠曝噪声等来实现面向小样本的图像增广和样本集制备;然后,将该标注样本集代入引入 CBAM 注意力机制的 Mask R-CNN 模型(简称 CMR 模型)中进行训练;最后,将如图 8(a)、8(b)、8(c)、8(d)所示的待检建筑物地面原始影像输入训练后的 CMR 模型,输出如图 8(e)、8(f)、8(g)、8(h)所示的建筑物立面

损毁检测结果。可以发现,从CMR模型输出的检测结果中,既能获得建筑物立面共有几处损毁及其相应的位置信息,又能获得每一处损毁所覆盖的范围信息;而且在面向不同色彩、材质和纹理结构的建筑物立面背景时,模型都能够从中较为准确地检测出墙面裂缝、墙体剥落、墙面孔洞和倒塌砾石堆等损毁信息。需要说明的是,由于模型对影像中的疑似损毁区域具有较强的敏感性,故部分路面碎石堆也会被检测为倒塌砾石堆的损毁类型而被标注出,如图8(h)中红框所示。

为了验证本文方法的模型可迁移性和鲁棒性,采用研究组在汉旺地震遗址采集的典型震害建筑物地面影像作为测试组,将如图9(a)、9(b)、9(c)、9(d)所示的待检建筑物地面原始影像输入训练后的CMR模型,输出如图9(e)、9(f)、9(g)、9(h)所示的建筑物立面损毁检测结果。可以发现,在CMR模型输出的建筑物立面损毁检测结

果中,待检影像中大部分的墙体剥落、墙面孔洞和墙面裂缝等损毁信息均能被检测出来。因此,可以推断出本文方法在应用于震后建筑物立面损毁检测时具有一定的模型可迁移性和场景鲁棒性。

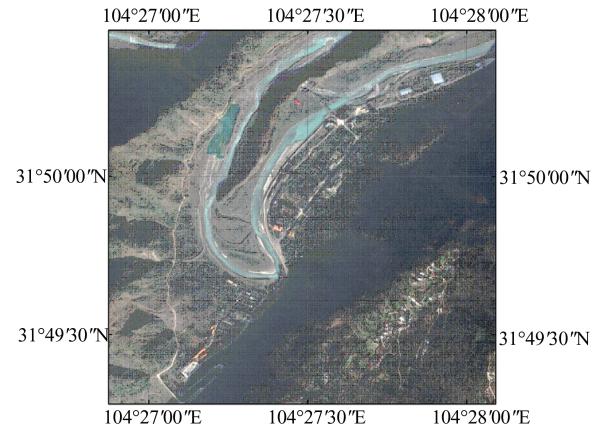


图7 北川老县城地震遗址区位

Fig.7 Location of Beichuan Old County Seismic Site

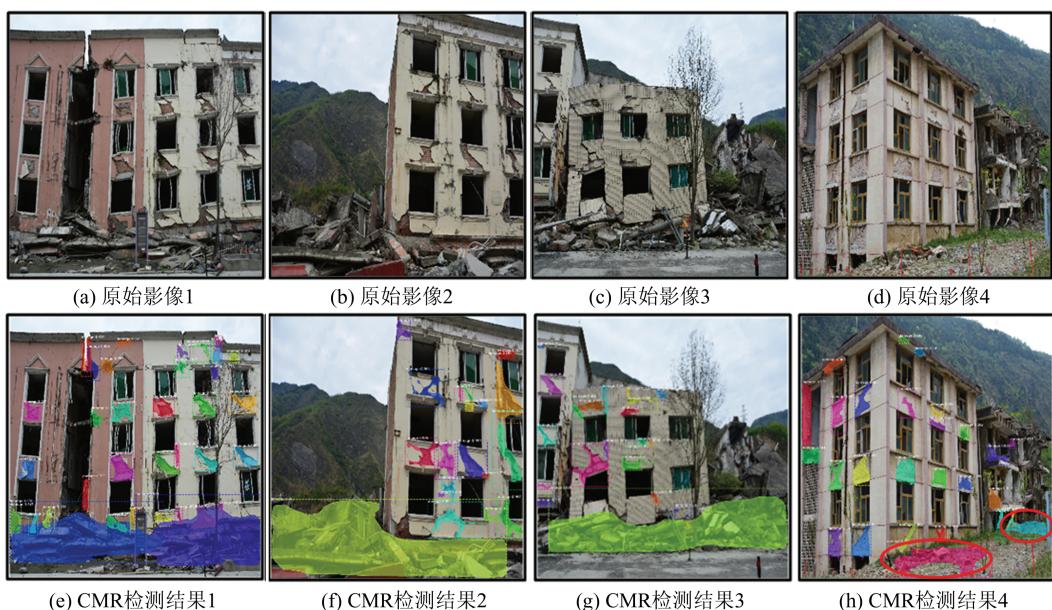


图8 CMR模型输出北川遗址建筑物立面损毁检测结果

Fig.8 Beichuan Building Façade Damage Detection Results Using CMR Model

### 2.3 融合处理

CMR模型虽然能够较为准确地检测出建筑物立面损毁信息,但随着卷积网络层数的叠加,建筑物立面损毁的边界信息会有所丢失。因此,基于多数投票规则,采用建筑物立面影像的多尺度分割结果对其损毁检测结果进行融合处理。为了分析不同分割尺度对损毁检测结果优化的影响,分别采用尺度参数为300、100、50的建筑物立面影像分割结果对CMR模型输出的损毁检测结果进行融合处理,得到的结果如图10(a)、10(b)、10(c)所示。

可以发现,随着分割尺度参数降低(影像过分割程度加深),融合处理后所保留的建筑物立面损毁信息更加完整,且能够对建筑物立面的细节损毁信息(如墙体剥落、墙面裂缝等)进行更为精准的边界提取。因此,在对比实验后,本文最终选定图像多尺度参数为50。

融合处理前,尽管采用CMR模型能够较为准确地定位建筑物立面影像中的损毁信息,但随着卷积网络层数的叠加,立面损毁的边界信息会逐渐模糊或丢失,如图11(b)、11(c)、11(d)中的红圈所示。而利用多尺度分割算法能够实现图

像中的同质像元聚类,并保留各类别之间较为准确的边界信息。因此,在经过多尺度分割融合后处理之后的建筑物立面损毁检测结果中,不仅能够得到建筑物立面损毁信息的精准定位结果,还

能够保留其准确的边界信息,从而显著提升了诸如墙体剥落、墙面裂缝等建筑物立面细节损毁信息的分割精度,如图 11(f)、11(g)、11(h)中的绿圈所示。

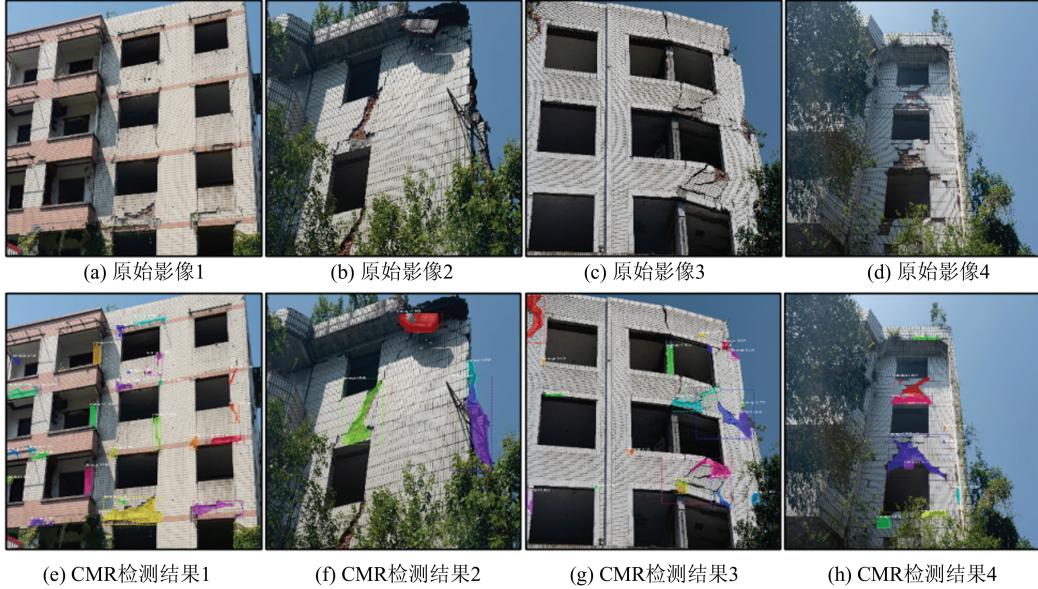


图 9 CMR 模型输出汉旺遗址建筑物立面损毁检测结果

Fig.9 Hanwang Building Façade Damage Detection Results Using CMR Model

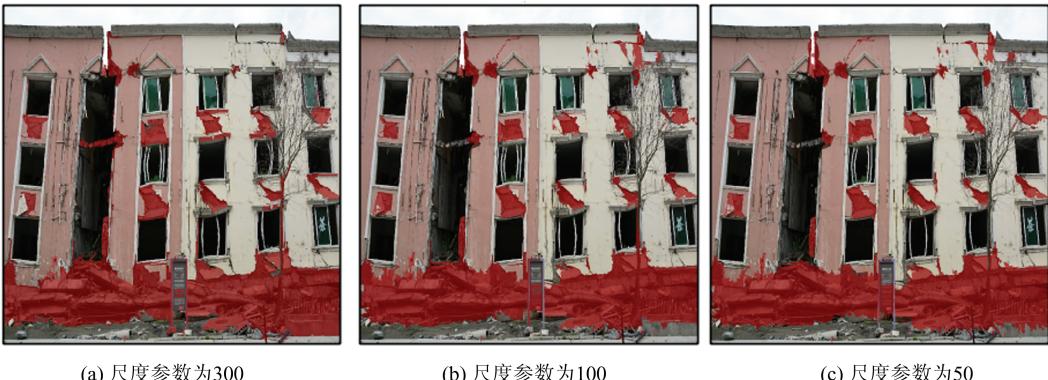


图 10 融合处理结果

Fig.10 Fusion Processing Results

## 2.4 精度评价

本文选取的精度评价指标包括准确率  $A$ 、精确率  $P$ 、召回率  $R$  和  $F_\beta$  分数。

准确率表示预测正确的结果占总样本的百分比,其计算公式为:

$$A = \frac{T_p + T_n}{T_p + F_p + T_n + F_n} \times 100\% \quad (6)$$

式中,  $T_p$  表示实际为正预测为正的像素个数;  $F_p$  表示实际为负预测为正的像素个数;  $T_n$  表示实际为负预测为负的像素个数;  $F_n$  表示实际为正预测为负的像素个数。

准确率虽然能够判断模型的总体正确率,但在正负样本量不均时容易失效,故又引入了精确

率、召回率和  $F_\beta$  分数指标。具体地,精确率表示在所有被预测为正的样本中实际也为正样本的概率,精确率越高代表模型的误检率越低。其计算公式为:

$$P = \frac{T_p}{T_p + F_p} \times 100\% \quad (7)$$

召回率表示在实际为正的样本中被预测也为正样本的概率,召回率越高代表模型的漏检率越低。其计算公式为:

$$R = \frac{T_p}{T_p + F_n} \times 100\% \quad (8)$$

$F_\beta$  分数表示精确率和召回率的综合评价指标。选取  $F_2$  分数作为综合评价指标,计算公式为:

$$F_{\beta} = (1 + \beta^2) \cdot \frac{PR}{(\beta^2 P) + R} \quad (9)$$

从验证集中选取一组建筑物地面影像,分别采用传统方法中具有代表性的BOVW模型、深

度学习实例分割方法中的Mask R-CNN模型和本文方法进行建筑物立面损毁检测实验,并根据各个模型输出的检测结果计算上述精度评价指标,得到如表1所示的计算结果。

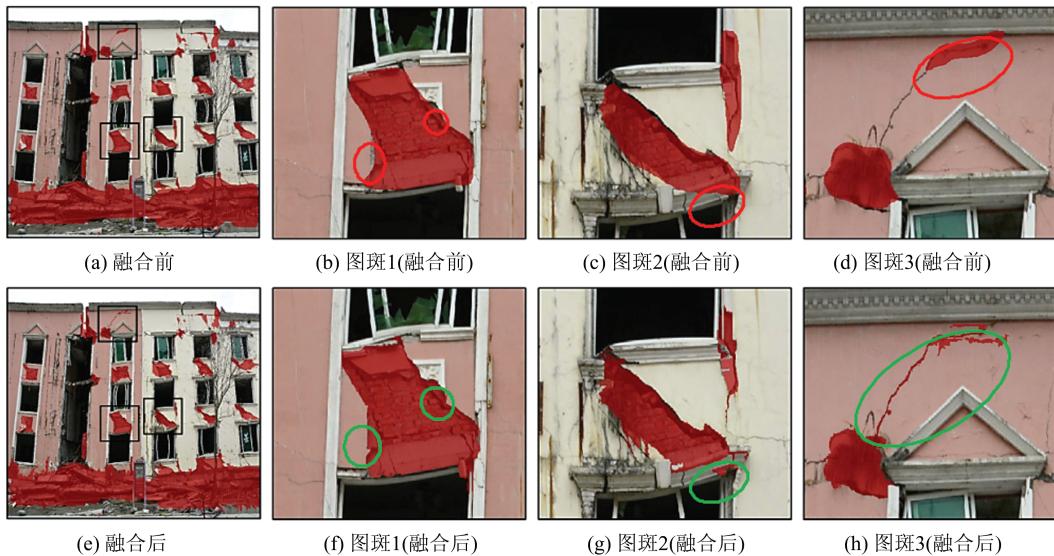


图11 融合处理效果对比

Fig.11 Comparison of Fusion Processing Effects

可以发现,相比于传统方法,深度学习实例分割模型能够大幅提升建筑物立面损毁检测能力;而本文方法又能够通过改进Mask R-CNN模型对建筑物立面损毁信息的通道和空间聚焦能力,有效提高其检测准确率、精确率、召回率和 $F_2$ 分数,从而进一步降低模型的误检率和漏检率。

表1 精度比较/%  
Tab.1 Comparison of the Performance/%

模型	准确率	精确率	召回率	$F_2$ 分数
BOVW	67.78	32.55	40.19	33.74
Mask R-CNN	86.91	65.91	61.72	62.70
本文方法	89.15	72.19	68.01	68.20

### 3 结语

建筑物立面损毁检测对震后建筑物全方位损毁评估有着重要意义。针对目前建筑物损毁检测研究中立面损毁评估缺乏以及智能化、精细化水平不高等缺点,本文提出了一种融合引入注意力机制的深度学习实例分割模型和图像多尺度分割算法进行震后建筑物立面损毁检测的方法。通过实地采集建筑物地面影像并结合灾区情况进行图像增广,有效保障了实例分割模型在训练过程中的必要样本量;引入CBAM注意力模块改进Mask R-CNN模型,显著提高了震后建筑

物立面损毁检测能力,实现了从复杂建筑物立面背景中有效提取损毁信息;此外,采用影像多尺度分割结果对建筑物立面损毁检测结果进行融合处理,解决了因卷积层数累积所导致的损毁边界模糊问题。总体上,本文提出的建筑物立面损毁检测方法能够从复杂的震后建筑物立面背景中有效提取出边界清晰的损毁信息,并有效提升了Mask R-CNN模型对建筑物立面损毁信息的检测能力,为震后建筑物全方位精细化损毁检测提供了重要的地面技术手段。

### 参 考 文 献

- [1] Li Qiang, Zhang Jingfa. Research on Earthquake Damaged Building Extraction by Different Features Fusion [J]. *Journal of Seismological Research*, 2016, 39(3):486-493(李强,张景发.不同特征融合的震后损毁建筑物识别研究[J].地震研究,2016,39(3):486-493)
- [2] Tao Heping, Liu Bintao, Liu Shuzhen, et al. Natural Hazards Monitoring Using Remote Sensing: A Case Study of 5·12 Wenchuan Earthquake [J]. *Journal of Mountain Science*, 2008, 26(3):276-279(陶和平,刘斌涛,刘淑珍,等.遥感在重大自然灾害监测中的应用前景:以5·12汶川地震为例[J].山地学报,2008,26(3):276-279)
- [3] Sui Haigang, Liu Chaoxian, Huang Lihong, et al. Application of Remote Sensing Technology in Earth-

- [1] 眇海刚, 刘超贤, 黄立洪, 等. 遥感技术在震后建筑物损毁检测中的应用[J]. 武汉大学学报·信息科学版, 2019, 44(7): 1 008-1 019.
- [2] Jia Ning, Zhao Pengxiang. Crowdsourcing Geographic Data[J]. *China Science and Technology Information*, 2013(12): 42.
- [3] Jiang Qingxiang, Liu Huiping. Extracting TM Image Information Using Texture Analysis[J]. *Journal of Remote Sensing*, 2004, 8(5): 458-464.
- [4] Tu Jihui, Sui Haigang, Lv Ruipeng, et al. Building Facade Damage Detection Based on the Gini Index from Oblique Aerial Images[J]. *Geomatics and Information Science of Wuhan University*, 2017, 42(12): 1 744-1 748.
- [5] Vettrivel A, Gerke M, Kerle N, et al. Identification of Structurally Damaged Areas in Airborne Oblique Images Using a Visual-Bag-of-Words Approach[J]. *Remote Sensing*, 2016, 8(3): 231-252.
- [6] Kerle N, Fernandez G J, Gerke M. Urban Structural Damage Assessment with Oblique UAV Imagery, Object-Based Image Analysis and Semantic Reasoning[C]. The 35th Asian Conference on Remote Sensing, Nay Pyi Taw, Myanmar, 2014.
- [7] Fernandez G J, Kerle N, Gerke M. UAV-Based Urban Structural Damage Assessment Using Object-Based Image Analysis and Semantic Reasoning[J]. *Natural Hazards & Earth System Sciences*, 2015, 15(6): 1 087-1 101.
- [8] He K, Gkioxari G, Dollár P, et al. Mask R-CNN [C]. The IEEE International Conference on Computer Vision, Venice, Italy, 2017.
- [9] 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): 1 137-1 149.
- [10] Long J, Shelhamer E, Darrell T. Fully Convolutional Networks for Semantic Segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 39(4): 640-651.
- [11] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional Block Attention Module[C]. The European Conference on Computer Vision, Munich, Germany, 2018.
- [12] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, 2016.
- [13] Xie Fei, Mu Yu, Guan Ziyu, et al. Oral Leukoplakia (OLK) Segmentation Based on Mask R-CNN with Spatial Attention Mechanism[J]. *Journal of Northwest University(Natural Science Edition)*, 2020, 50(1): 9-15.
- [14] Sui Haigang<sup>1</sup>, Huang Lihong<sup>1</sup>, Liu Chaoxian<sup>1</sup>.

## Detecting Building Façade Damage Caused by Earthquake Using CBAM-Improved Mask R-CNN

SUI Haigang<sup>1</sup> HUANG Lihong<sup>1</sup> LIU Chaoxian<sup>1</sup>

<sup>1</sup> State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

**Abstract:** **Objectives:** Building damage information can provide an important basis for the decision making of rapid post-earthquake assessment. Traditional building damage detection techniques mainly focus on the roof surface, thus many damaged buildings with an intact roof surface but collapsed middle floors may be neglected. We propose a method of building façade damage detection based on deep learning and multiresolution segmentation algorithm. **Methods:** The method which integrates the instance segmentation and multi-resolution segmentation algorithm is applied to detecting the post-earthquake building façade damage. The

first thing is to collect the ground images of post-earthquake buildings in the field and perform the data augmentation. Secondly, we use the convolutional block attention module (CBAM) to improve Mask R-CNN. Then the dataset is input to the improved model for training, and finally a multiresolution segmentation algorithm is adopted to post-process the building façade damage detection results output by the CBAM-Improved Mask R-CNN. **Results:** The experimental results show: (1) Collecting ground images of buildings in the field and performing image augmentation can effectively guarantee the necessary training sample size of the instance segmentation model. (2) The Mask R-CNN improved by CBAM attention mechanism significantly improves the post-earthquake building facade damage detection capabilities, which realizes the precise extraction of damage information from complex building façade backgrounds. (3) In addition, using the multiresolution segmentation algorithm to post-process the building facade damage detection results can obviously solve the blurred boundary problems caused by the accumulation of convolutional layers. **Conclusions:** The proposed method can significantly improve the capability of post-earthquake building façade damage detection when compared to the traditional methods, which also raises the Mask R-CNN's accuracy, precision, recall and  $F_2$ -score to a certain degree. It can be inferred that the proposed method has the strong potential to be applied to the post-earthquake building façade damage detection and therefore provides an important technical means for detecting the comprehensive and detailed building damage detection caused by earthquake.

**Key words:** earthquake; building façade damage; ground images; Mask R-CNN; convolutional block attention module; multiresolution segmentation

**First author:** SUI Haigang, PhD, professor, specializes in remote sensing, GIS and disaster emergency response. E-mail: haigang\_sui@263.net

**Corresponding author:** LIU Chaoxian, PhD. E-mail: cx\_leo@whu.edu.cn

**Foundation support:** The National Natural Science Foundation of China(41771457); the National Key Research and Development Program of China(2018YFB10046).

**引文格式:** SUI Haigang, HUANG Lihong, LIU Chaoxian. Detecting Building Façade Damage Caused by Earthquake Using CBAM-Improved Mask R-CNN[J]. Geomatics and Information Science of Wuhan University, 2020, 45(11): 1660–1668. DOI: 10.13203/j.whugis20200158(眭海刚, 黄立洪, 刘超贤. 利用具有注意力的Mask R-CNN检测震害建筑物立面损毁[J]. 武汉大学学报·信息科学版, 2020, 45(11):1660–1668.DOI:10.13203/j.whugis20200158)