DOI:10.19951/j.cnki.1672-9331.20231120004 文章编号:1672-9331(2024)05-0136-10
引用格式:赵平,靳丽艳,刘钰.基于双重注意力U-Net的砌体结构震害裂缝检测[J].长沙理工大学学报(自然科学版),2024,21(5):136-145,155.
Citation: ZHAO Ping, JIN Liyan, LIU Yu. Detection of earthquake-induced cracks in masonry structures based on dual attention U-Net[J]. J Changsha Univ Sci Tech (Nat Sci), 2024, 21(5): 136-145,155.

# 基于双重注意力U-Net的砌体结构震害裂缝检测

赵平[1],靳丽艳[1],刘钰[2]
(1.西安建筑科技大学 土木工程学院,陕西 西安 710055;2.西安建筑科技大学 资源工程学院,陕西 西安 710055)

**摘　要:【目的】**砌体结构震害裂缝严重影响结构安全。在震后应急响应阶段,对结构各部位进行裂缝检测是结构安全鉴定的重要依据。为提高检测效率,提出一种改进的多目标语义分割算法,即双重注意力U-Net(dual attention U-Net,DA U-Net)。**【方法】**首先,为支持模型的训练,创建一个手工标记的数据集,并采用数据增强策略保证网络学习到更多的裂缝细节特征;其次,使用VGG16(visual geometry group 16-layer)网络替换U-Net主干特征提取网络,并且在跳层连接处嵌入卷积块状注意力模块和非局部注意力机制,有利于改善裂缝边缘分割不完整和全局特征信息利用不充分的问题,以及有效利用浅层特征信息来加强特征表示;最后,用亚像素卷积取代原上采样操作,补充低分辨率像素缺失的语义信息。**【结果】**改进DA U-Net网络在分类准确率和分割精度上均达到最佳表现。精度评价指标$F_1$分数($F_1$ score)指标$E_{F_1}$、平均像素精度(mean pixel accuracy,MPA)指标$E_{\mathrm{MPA}}$和平均交并比(mean intersection over union,MIOU)指标$E_{\mathrm{MIOU}}$达到84.66%、91.57%和81.50%,相比U-Net网络的分割精度均提高5.00%以上。选择合适的主干网络能更好地捕捉和表征数据的复杂特征,同时对全局空间位置信息关注度也更高。**【结论】**改进算法可显著提升对裂缝识别和定位的准确性,以及对裂缝的整体形态、走向和交融情况的分割完整性,增强多组合裂缝分割精度,为震后砌体结构房屋安全鉴定提供一种有效的检测方法。

**关键词:**砌体结构;地震;裂缝检测;改进U-Net;多目标分割;注意力机制
**中图分类号:**TU362;TU17　　**文献标志码:**A

# Detection of earthquake-induced cracks in masonry structures based on dual attention U-Net

ZHAO Ping[1], JIN Liyan[1], LIU Yu[2]
(1. School of Civil Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China;
2. School of Resources Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China)

**Abstract:** [**Purposes**] Earthquake-induced cracks seriously affect the safety of masonry structures. During the post-earthquake emergency response, crack detection in various components is crucial for the safety assessment of the structures. To enhance detection efficiency, an improved dual attention U-Net (DA U-Net) based on a multi-target semantic segmentation algorithm was proposed. [**Methods**] For effective model training, a manually labeled dataset was created, and data augmentation strategies were utilized to make the network learn about more crack details. Furthermore, VGG16 was employed to replace the backbone feature extraction network of U-Net.

**收稿日期:**2023-11-20;**修回日期:**2024-01-18;**接受日期:**2024-01-18
**基金项目:**国家自然科学基金项目(52208204)
**通信作者:**赵平(1967—)(ORCID:0009-0009-5568-7894),女,教授,主要从事土木工程建造与管理方面的研究。E-mail:13909208708@163.com

Additionally, convolutional block attention modules and non-local attention (NA) mechanisms were embedded into the skip connections. As a result, the problems of incomplete crack edge segmentation and inadequate utilization of global feature information were ameliorated. Shallow feature information was effectively utilized, and feature representation was strengthened. Finally, sub-pixel convolution replaced the original up-sampling operation, supplementing semantic information loss in low-resolution pixels. [**Findings**] The improved DA U-Net achieves optimal performance in both classification accuracy and segmentation precision, with the precision evaluation metrics of $F_1$ score ($E_{F_1}$), mean pixel accuracy ($E_{\mathrm{MPA}}$), and mean intersection over union ($E_{\mathrm{MIOU}}$) reaching 84.66%, 91.57%, and 81.50%, respectively. This represents an improvement of over 5.00% in segmentation accuracy compared to the original U-Net. Selecting an appropriate backbone network enables better capture and representation of complex features in the data, with a heightened focus on global spatial location information. [**Conclusions**] The improved algorithm significantly enhances the accuracy of crack recognition and localization, as well as the completeness of segmentation for the overall morphology, direction, and intersection of cracks. It strengthens the precision of segmenting multi-combined cracks, providing an effective detection method for post-earthquake safety assessment of masonry structures.

**Key words**: masonry structure; earthquake; crack detection; improved U-Net; multi-target segmentation; attention mechanism

**Foundation item**: Project (52208204) supported by the National Natural Science Foundation of China

**Corresponding author**: ZHAO Ping (1967— )(ORCID: 0009-0009-5568-7894), female, professor, research interest: civil engineering construction and management. E-mail: 13909208708@163.com

## 0　引言

我国是大陆地震最多的国家，且地震分布相当广泛，其中绝大多数发生在农村、乡镇等经济不发达地区。这些地区房屋的主要结构为砌体结构[1]，其对应的震害主要以墙体的不同破坏方式为突出特征。作为主要承重构件，墙体在地震时承受拉压、剪切和弯曲等复杂地震作用，产生各种形式裂缝，并且在反复地震作用下，这些裂缝不断发展、增多、加宽，最后导致墙体开裂崩塌甚至房屋塌陷[2]，从而严重威胁人民生命和财产安全。因此，快速准确获取砌体结构建筑实时的灾害信息对应急响应至关重要[3]，不仅有助于确定哪些建筑物需要紧急维修或拆除，以免房屋二次倒塌，还可以评估灾后建筑物的结构破坏程度，提升救灾效率[4]。

震后结构安全评估的传统方法是分析建筑破坏等级[5]，但这种方法在地震灾害发生后需要投入大量人力、物力，且耗时、易出错。随着人工智能（artificial intelligence，AI）和深度学习（deep learning，DL）的发展，建筑行业正在经历智能化变革[6-9]。深度学习已被广泛应用于混凝土和沥青表观裂缝检测，但在砌体结构的缺陷检测方面的应用较少。砌砖接缝和复杂的背景纹理容易导致砌体表面匀质性变差和噪声变大，使砌体表观检测更加具有挑战性[10]。卷积神经网络（convolutional neural network，CNN）通常用于图像分类和分割，在辅助结构检查和检测应用领域具有广阔前景。HALLEE等[11]在实验室收集了一些砌体结构图片，将砌体裂缝分为裂纹、未裂纹、模糊、局部和无砖5类，并在定制CNN模型中的训练中发现，CNN模型虽比机器学习模型具有更佳的性能，但模型性能在测试真实图片时急剧下降。BRACKENBURY等[12]从砖石桥上拍摄大量图片，用GoogleNet-Inception V3和迁移学习[13]对图像中的裂缝、剥落、砂浆损失、植被等不同缺陷类型进行分类和分割，提出先分割砂浆缝再使用缺陷分割的方式，以便对砌体缺陷区和无缺陷区进行更可靠、更准确的检测。ALI[14]在CNN、支撑向量机

(support vector machines, SVM)和随机森林(random forest,RF)的基础上提出了一套自动视觉裂缝检测系统以简化砌体结构检测,但该系统会混淆砌体结构裂缝和灌浆线。DAIS等[15]在ImageNet数据集上使用预训练的Mobile Net网络获得最高裂缝分类准确率,对补丁级砌体裂缝的分类准确率达到95.30%。因此,研究者提出迁移学习的重要性。全卷积网络(fully convolutional networks,FCN)[16]将CNN结构的全接连层替换为卷积层,被广泛用于语义分割。U-Net[17]是一种深度FCN网络,在边缘检测上表现出良好的性能,成为不同领域图像分割的基准。U-Net使用对称编码器-解码器结构进行像素级语义分割,在较小的数据集中得到广泛的应用[18-20]。IBRAHIM等[21]提出使用分水岭变换对每个砖单元进行分割,再使用U-Net对不同黏结模式砌体结构的砂浆进行分割。DANG等[22]对U-Net、DeepLab V3+和特征金字塔网络(feature pyramid networks,FPN)等模型进行评估,提出通过检测砖块来提取砌体结构裂缝实际长度的方法。其结果表明,基于深度学习的分割效果优于以往方法;但当输入图像的倾斜角度过大时,砖块单元提取精度降低,同时裂缝实际长度检测精度变差。LOVERDOS等[23]基于深度学习,提出对砌体结构墙体进行砖块分割和裂缝检测的自动一体化模型。该模型对砖块检测准确率为96.86%,裂缝检测准确率为79.60%,但无法识别极小尺寸的裂缝。

综上所述,现有砌体结构裂缝检测使用的图像大多来自实验室和特定建筑,且具有相似的分布特征。模型在严格控制的环境下具有较高的性能,但在其他数据集上的泛化能力尚未得到充分评估。关于检测震后砌体裂缝参数(长度、宽度、面积等)的研究较少,但这些信息是震后影响砌体结构安全评估的关键因素。

为提高训练数据集与真实世界的接近程度,本研究收集服从多源数据分布的砌体结构震后真实图片,用迁移学习方法和增强型U-Net架构进行震后砌体结构实时裂缝检测。本研究通过对比分析精度指标来验证改进模型的效果,以期获得专门针对震后砌体结构多裂缝检测的网络模型,并对检测到的裂缝进一步计算实际信息,做到从数字图像中自动记录震后墙体裂缝信息,并评估砌体结构的损伤。

## 1 检测方法

### 1.1 模型整体架构

U-Net多用于二分类语义分割。为提升U-Net模型在处理砌体结构震害图像多类别分割任务中的性能,本研究通过替换主干特征提取网络、增加注意力机制对网络结构进行优化。具体来说,本研究首先将visual geometry group 16-layer(VGG16)网络去掉最后一个池化层和之后的全连接层,然后替换U-Net的特征提取主干网络,接着在U-Net跳层连接处融合通道、空间注意力和非局部注意力(non-local attention,NA),最后形成双重注意力网络。这样使网络在处理数据时重点关注裂缝边缘信息,提高多任务细节分割精度,并改善错分、漏分现象。网络模型如图1所示。

本研究将图片像素调整为512×512作为网络输入。首先,经过二次3×3的卷积激活操作,可得到第一层网络特征图并对之进行2×2最大值池化下采样操作,此时特征图尺寸因通道数加倍变为256×256,再进行卷积操作提取下一层;其次,重复卷积操作和下采样操作,得到5个从大而扁平变得小而立体的特征图,并对其先通过卷积块注意力模块(convolutional block attention module,CBAM)进行特征加强,聚焦裂缝信息且减弱对非裂缝信息的关注;再次,执行非局部注意力操作,使网络在有效利用初始位置信息的基础上注意全局信息;然后,进入解码器阶段后,对特征层执行亚像素卷积上采样操作,使特征图大小加倍通道数减

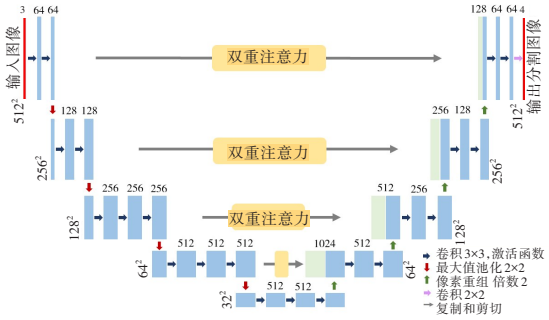



**图1** DA U-Net-VGG16网络总体架构

**Fig. 1** Overall architecture of DA U-Net-VGG16 network

半；接着，将优化后的特征图和上采样特征图进行维度拼接，实现浅层特征复用，增强特征深度延展性，并在重复4次上采样后，继续进行3×3卷积，使特征图恢复到输入时的大小，保证输入、输出尺寸不变；最后，采用1×1卷积将通道数调整为4，对“背景”“微裂缝”“裂缝”“剥落”进行分类输出，以获得最终的震害图像分割结果。

### 1.2　CBAM

WOO等[24]提出CBAM轻量型注意力模块。该模型通过动态调整每个特征通道的权重，使模型更加关注输入图片中的裂缝边缘信息、整体走向，增强特征表示。CBAM包括两块模块，即增强特征通道之间相关性的通道注意力模块（channel attention module，CAM）、对特征图的每个位置进行自适应加权的空间注意力模块（spatial attention module，SAM），如图2所示。

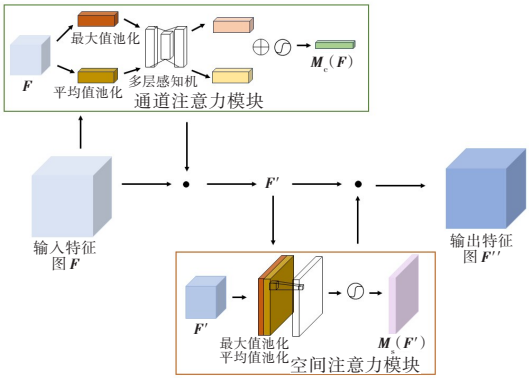



**图2**　CBAM模块示意图

**Fig. 2**　CBAM module

本研究首先对输入特征图$\boldsymbol{F}$进行最大值池化（max pooling，MaxP）和平均值池化（average pooling，AveP）操作，得到两个特征图；其次，将这两个特征图送入多层感知机（multilayer perceptron，MLP）得到输出特征图像素；再次，将输出征图像素加和，再经sigmoid激活函数得到通道注意力特征图$\boldsymbol{M}_c(\boldsymbol{F})$，并与输入特征图$\boldsymbol{F}$点乘得到中间过程特征图$\boldsymbol{F}'$；然后，对$\boldsymbol{F}'$每个通道进行MaxP和AveP操作，并进行通道拼接；接着，送入一个标准卷积层，再降低其通道数为1，并经sigmoid激活得到空间注意力特征图$\boldsymbol{M}_s(\boldsymbol{F}')$；最后，$\boldsymbol{M}_s(\boldsymbol{F}')$与$\boldsymbol{F}'$点乘得到最终输出特征图$\boldsymbol{F}''$。相关公式为

$$\boldsymbol{M}_c(\boldsymbol{F}) = S\left\{\mathrm{MLP}\left\{\mathrm{AveP}(\boldsymbol{F}) + \mathrm{MLP}\left[\mathrm{MaxP}(\boldsymbol{F})\right]\right\}\right\} \tag{1}$$

$$\boldsymbol{F}' = \boldsymbol{M}_c(\boldsymbol{F}) \cdot \boldsymbol{F} \tag{2}$$

$$\boldsymbol{M}_s(\boldsymbol{F}') = S\left\{\sigma\left\{\left[\mathrm{AveP}(\boldsymbol{F}')\right];\ \left[\mathrm{MaxP}(\boldsymbol{F}')\right]\right\}\right\} \tag{3}$$

$$\boldsymbol{F}'' = \boldsymbol{M}_s(\boldsymbol{F}') \cdot \boldsymbol{F}' \tag{4}$$

式中：$\boldsymbol{M}_c(\boldsymbol{F})$为通道注意力特征图；$S$为sigmoid激活函数；MLP为一个共享的两层神经网络；$\sigma$为标准卷积运算；$\boldsymbol{F}$为输入特征图；$\boldsymbol{F}'$为中间过程特征图；$\boldsymbol{F}''$为输出特征图。

### 1.3　非局部注意力（NA）

卷积只对图像中局部区域进行操作获取局部特征信息，忽略了全局图像结构和上下文语境。本研究引入非局部注意力模块[25]，减少信息收缩扩大全像素互动，在重视低级特征边缘信息的同时，有效利用多目标语义分割的上下文信息。非局部注意力模块如图3所示。

在图3中，首先对输入的特征图$\boldsymbol{X}$进行线性映射，生成$\boldsymbol{\theta}$、$\boldsymbol{\varphi}$、$\boldsymbol{g}$三组特征；其次，将这三组特征通道进行合并，对$\boldsymbol{\theta}$和$\boldsymbol{\varphi}$进行矩阵点乘操作；再次，计算每个像素与其他像素之间的自相关性，以衡量它们之间的关联程度；然后，对自相关性特征进行softmax归一化，得到每个像素对其他像素的注意力权重$\boldsymbol{f}$，用于表示像素的重要性；接着，将得到的$\boldsymbol{f}$与$\boldsymbol{g}$相乘，并通过一个1×1卷积扩展回原始通道数；最后，将结果与$\boldsymbol{X}$进行残差连接，得到非局部块的输出$\boldsymbol{X}'$。

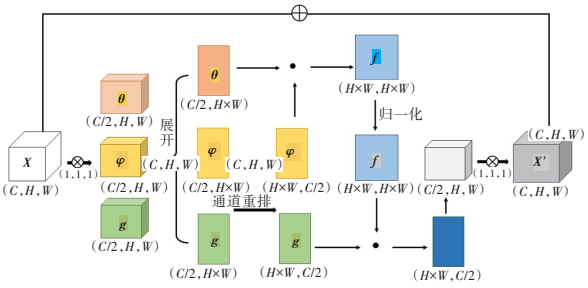



**图3**　非局部注意力模块（$H$:高度，$W$:宽度，$C$:通道数）

**Fig. 3**　NA module ($H$: height; $W$: width; $C$: channel)

### 1.4　亚像素卷积（subpixel convolution）

在解码器结构中，使用亚像素卷积[26]代替原上采样步骤，上采样的倍数为2。亚像素卷积为了将低分辨率的图像映射到高分辨率的图像，首先将每个通道的特征图划分成多个子像素，然后将子像素沿着某一维度堆叠，最后转换为高分辨率特征图输出，从而帮助还原低分辨率图像信息。其中的上采样操作，有利于增加裂缝图像的细节

和清晰度,有助于补充缺失的图像边缘信息。

## 2 创建数据集

### 2.1 数据采集

鉴于现有砌体结构震损数据集较少,本研究创建一个数据集。该数据集包含水平裂缝、竖向裂缝、斜裂缝和交叉裂缝等多种裂缝形式,其中80%来自建筑物震害图集[27-30]和网络,其余来自建筑抗震鉴定现场拍摄。本研究对原始图片进行预处理,保证数据集中所有图片具有相同分辨率。

为增强网络鲁棒性,避免网络训练过拟合,对数据集随机执行上下左右翻转、随机剪裁、旋转、扭曲等数据增强操作,将数据集中的图片扩充至5 000张,并按1∶9划分验证集和训练集。本研究使用标注工具labelme对数据集进行标注,在图像上标注裂缝轮廓并贴上类别标签。根据裂缝的特征,按照裂缝像素宽度将震害类型分为三类:微裂缝,其缝宽极为狭窄,在掩码图显示为一条黄线;裂缝,具有明显宽度,在掩码图上呈现为红色不规则条状;剥落,其特征是整块墙皮掉落,在掩码图以绿色块状形式展现。裂缝标注示例如图4所示。

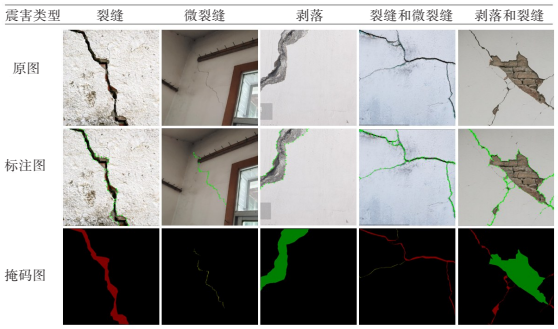



**图4** 标注裂缝图像

**Fig. 4** Labelled crack images

### 2.2 模型训练环境

模型的训练和测试在远程服务器上进行,基于CUDA11.8、Python3.8、PyCharm和Pytorch2.0.0搭建深度学习框架,使用NVIDIA GeForce RTX 4090 GPU对网络进行训练。

### 2.3 网络超参数的选择

训练过程分为两个阶段,总共进行了300次迭代训练,图片尺寸均为512×512像素。第一阶段是冻结训练,将预训练模型主干网络部分参数冻结,冻结迭代次数设置为100,批量大小为8,初始学习率为$1\times10^{-4}$;第二阶段是解冻训练,预训练模型的主干部分的参数被解除固定,并可以被调整或优化。本研究使用Adam优化器,动量设置为0.9,迭代次数为200(其余参数不变),此时模型性能得到进一步提高。整个训练过程采用余弦退火策略调整学习率,可以平稳地调整模型的参数,避免陷入局部最优解。

### 2.4 评价指标

本研究采用平均交并比(mean intersection over union, MIOU)的指标$E_{\mathrm{MIOU}}$、平均像素精度(mean pixel accuracy, MPA)的指标$E_{\mathrm{MPA}}$、损失值$E_{\mathrm{LOSS}}$和$F_1$分数($F_1$ score)的指标$E_{F_1}$验证研究模型的准确性。

为综合考虑模型在多个类别上的分割性能,本研究对每个类别的交并比(intersection over union, IOU)进行平均得到$E_{\mathrm{MIOU}}$。通常情况下,$E_{\mathrm{MIOU}}$越大,表示模型在多类别分割任务中的性能越好;$E_{\mathrm{MPA}}$衡量了模型中所有正确分类的像素数量与总像素数量之间的比例,在多类别图像分割任务中评估模型的整体性能;$E_{\mathrm{LOSS}}$度量了模型预测与实际目标之间的差异,若损失值收敛且在验证集上表现良好,表示模型训练到了合适的程度;$E_{F_1}$的大小反映了模型训练效果的优劣程度。

$$E_{\mathrm{MIOU}} = \frac{1}{K+1}\sum_{i=0}^{K}\frac{P_{ii}}{\sum_{j=0}^{K}P_{ij}+\sum_{j=0}^{K}P_{ji}-P_{ii}} \tag{5}$$

$$E_{\mathrm{MPA}} = \frac{1}{K+1}\sum_{i=0}^{K}\frac{P_{ii}}{\sum_{j=0}^{K}P_{ij}} \tag{6}$$

$$P = \sum_{i=0}^{K}\frac{P_{ii}}{\sum_{j=0}^{K}P_{ij}+P_{ii}} \tag{7}$$

$$R = \sum_{i=0}^{K}\frac{P_{ii}}{\sum_{j=0}^{K}P_{ji}+P_{ii}} \tag{8}$$

$$E_{F_1} = 2\times\frac{P\times R}{P+R} \tag{9}$$

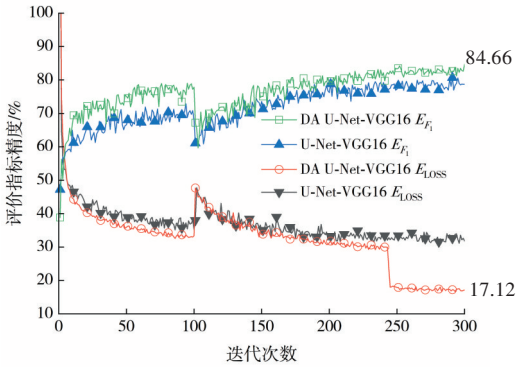
式中:$K$为类别总数;$K$+1为包含背景的类别总数;$P$为查准率;$P_{ii}$为属于$i$类、被预测为$i$类的像素数量,即真正;$P_{ij}$为属于$i$类、被预测为$j$类的像素数量,即假正;$P_{ji}$为属于$j$类、被预测为$i$类的像素数量,即假负;$R$为召回率。

## 3　模型训练与分析

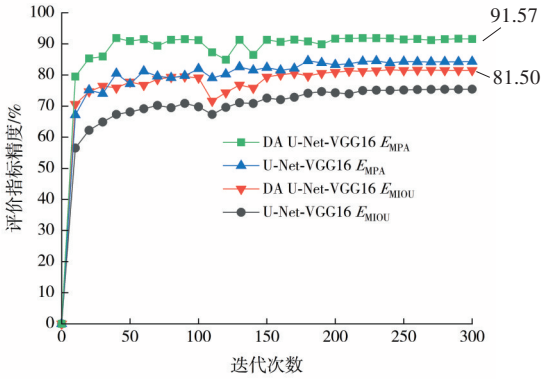
### 3.1　模型训练过程

网络在训练时采用迁移学习技术，将预训练VGG16网络用于模型主干特征提取网络，调用VGG16的预训练权重函数并传入预训练的参数，保持VGG16的强大特征提取能力。在较小数据集上使用迁移学习可以减少网络参数计算量，加速模型对新任务的学习过程。DA U-Net-VGG16和U-Net-VGG16训练结果如图5所示。

从图5可知，随迭代次数增加，损失值$E_{\mathrm{LOSS}}$整体呈现下降趋势，并在第二阶段训练时，突现小幅上升，$E_{F_1}$、$E_{\mathrm{MIOU}}$和$E_{\mathrm{MPA}}$随之下降，表明主干网络在解冻后，自动调整参数以优化模型性能；在迭代250次后，DA U-Net-VGG16的$E_{\mathrm{LOSS}}$再次下跌且最终稳定在17.00%左右，表明增加注意力机制的模型更好地适应训练数据，能更精确执行分割任务；在后200次迭代过程中，$E_{F_1}$、$E_{\mathrm{MIOU}}$和$E_{\mathrm{MPA}}$亦逐渐上升并趋于稳定，最终DA U-Net-VGG16的$E_{\mathrm{MPA}}$稳定在91.00%左右，$E_{\mathrm{MIOU}}$稳定在81.00%左右，$E_{F_1}$




（a）训练过程中的$E_{\mathrm{LOSS}}$、$E_{F_1}$变化




（b）训练过程中的$E_{\mathrm{MIOU}}$和$E_{\mathrm{MPA}}$变化

**图5**　DA U-Net-VGG16和U-Net-VGG16训练过程中的$E_{\mathrm{LOSS}}$、$E_{F_1}$、$E_{\mathrm{MIOU}}$和$E_{\mathrm{MPA}}$变化

**Fig. 5**　Changes of $E_{\mathrm{LOSS}}$, $E_{F_1}$, $E_{\mathrm{MIOU}}$, and $E_{\mathrm{MPA}}$ during training processes of DA U-Net-VGG16 and U-Net-VGG16

最高达到84.66%，三者均比U-Net-VGG16分割结果有不同程度提高。虽然模型性能在训练过程存在波动，但最终在更低$E_{\mathrm{LOSS}}$水平上实现了更高精度。这是一个积极迹象，表明模型在整个学习任务的过程中得到不断优化和收敛。

### 3.2　模型对比和讨论

#### 3.2.1　消融试验

本研究为验证增加注意力模块的有效性，设置相同的网络参数对数据集进行训练和测试。以U-Net为基准，并添加CBAM和NA注意力，使用两种主干特征提取网络VGG16和ResNet50（residual network 50-layer）进行消融研究，其结果如表1所示。通过比较所有分割模型的评价指标，可以更直观地分析各个模块对模型改进的有效性。

由表1可看出，主干网络VGG16的整体分割效果比ResNet50的更好，DA U-Net-VGG16的精度评价指标最优，DA U-Net-ResNet50的精度次之。

**表1**　消融研究结果分割结果

**Table 1**　Segmentation results of ablation study　　%

| 模块 | | | | | $E_{\mathrm{LOSS}}$ | $E_{\mathrm{MIOU}}$ | $E_{F_1}$ | $E_{\mathrm{MPA}}$ |
|---|---|---|---|---|---|---|---|---|
| U-Net | VGG16 | ResNet50 | CBAM | NA | | | | |
| √ | √ | | | | 32.00 | 75.36 | 78.68 | 84.19 |
| √ | √ | | √ | | 23.00 | 79.16 | 81.73 | 89.25 |
| √ | √ | | | √ | 18.00 | 76.51 | 84.12 | 85.62 |
| √ | √ | | √ | √ | 17.12 | 81.50 | 84.66 | 91.57 |
| √ | | √ | | | 32.00 | 74.87 | 78.73 | 85.74 |
| √ | | √ | √ | | 30.00 | 78.50 | 82.15 | 88.81 |
| √ | | √ | | √ | 27.00 | 78.14 | 82.94 | 88.23 |
| √ | | √ | √ | √ | 26.00 | 79.77 | 83.08 | 89.34 |

注：表中“√”表示添加了该模块。

这表明,选择合适的主干网络对模型性能优化非常重要。当主干网络使用VGG16时,在添加CBAM后,$E_{\mathrm{MIOU}}$和$E_{\mathrm{MPA}}$大幅提升,NA对$E_{\mathrm{LOSS}}$和$E_{F_1}$的贡献较大;当主干网络使用ResNet50时,精度指标有所改善但效果不显著。因此CBAM和NA模块均能提升模型精度,二者结合后可进一步提升分割效果。DA U-Net-VGG16的$E_{\mathrm{MPA}}$最高可达91.57%,比U-Net-VGG16的提高了7.38个百分点。这表明网络在融入了注意力机制后,能够更准确地捕捉裂缝信息和对像素进行分类,以及改善错分、漏分现象。同时,$E_{F_1}$为84.66%,$E_{\mathrm{MIOU}}$为81.50%,这说明改进后模型性能大幅提高,在多目标分割任务中表现出色。

### 3.2.2 实际分割效果

模型经过像素级标注后,在对图像进行分割时,会识别每个像素并为其分配一个唯一的标签。由于拍摄环境、设备限制或紧急情况下的快速拍摄等原因,震后拍摄的砌体结构震害图片质量不高。同时,剥落通常伴随细小裂缝出现,这使得从图片中准确识别目标变得困难,对模型检测精度要求更高。本研究选择4个精度差距较大的模型,即U-Net-VGG16、U-Net-ResNet50、DA U-Net-VGG16和DA U-Net-ResNet50,并从验证集中抽取6张代表性图片,来验证各模型对砌体结构震害特征的分割性能,如图6所示。

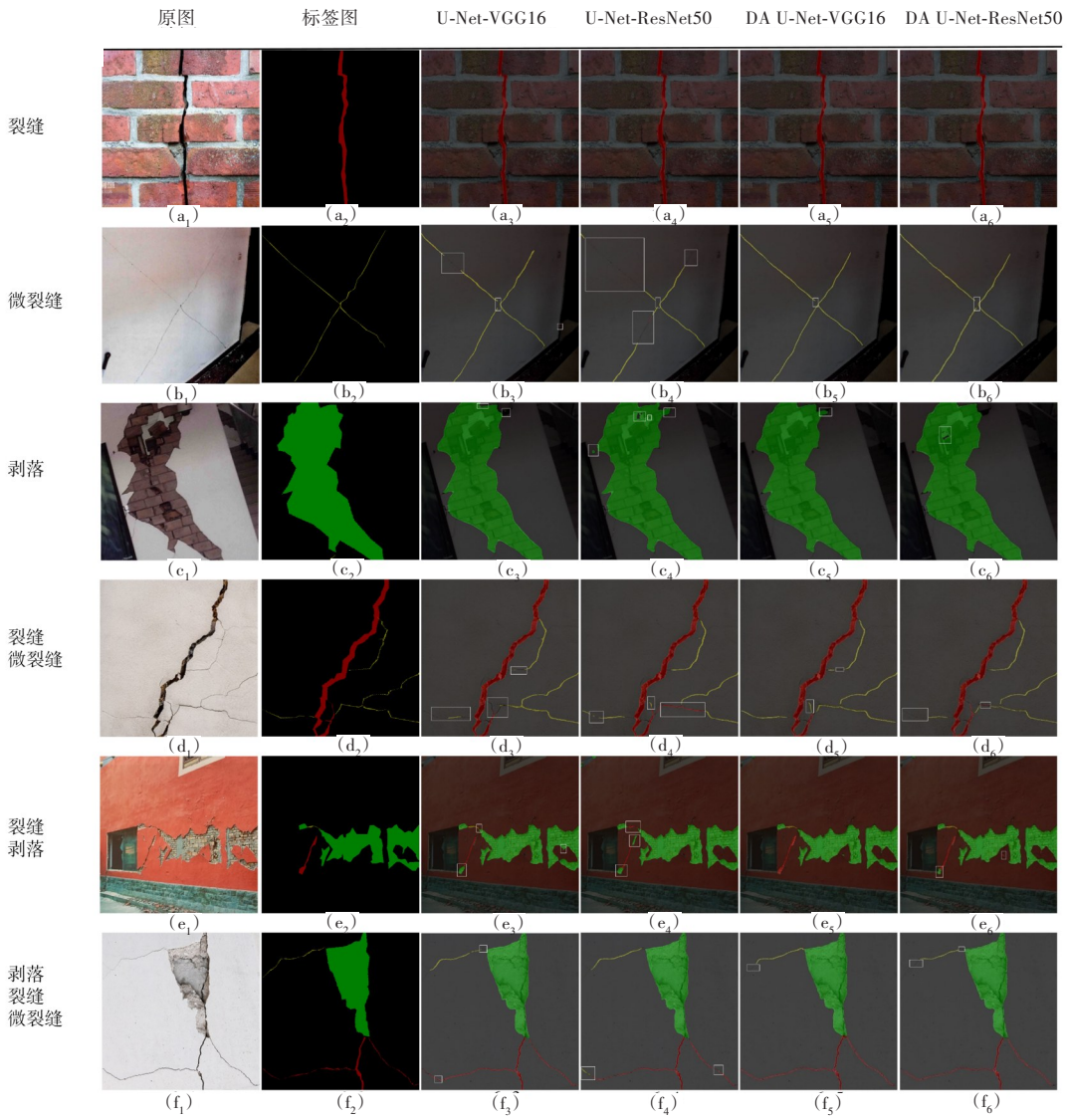



**图6** 砌体结构裂缝分割结果比较

**Fig. 6** Comparison of crack segmentation results of masonry structure

图6($a_1$)中的砖墙仅有竖向裂缝，背景简单，裂缝清晰，4种模型都将裂缝完整分割出来，且保持边界信息完整。图6($b_1$)中的墙面仅有交叉微裂缝，其宽度狭窄甚至肉眼难辨。图6($b_3$)和图6($b_4$)中的墙面出现较多漏分，边界提取不完整。图6($b_5$)和图6($b_6$)虽未能将裂缝交叉点处识别出来，但墙面确认修改以外的其余部分均被正确识别，且微裂缝整体脉络被分割出来。图6($c_1$)中的砖墙剥落面积较大且边缘复杂，其裂缝右上角有监视器，属于噪声干扰。图6($c_3$)的识别结果较准确但未能很好保持边界信息。图6($c_4$)和图6($c_6$)的预测不完整，存在孤岛情况。图6($c_5$)中的墙面边界分割完整，但监视器被误识别为剥落。

4种模型均能将主裂缝准确分割，对图6($d_5$)的分割结果最好，但对支路有不同程度的错分、漏分现象。图6($d_1$)中的墙面斜裂缝周围伴有多支微裂缝，裂缝杂乱且有多处交点，很难能做到完整、准确分割。

图6($e_1$)的外墙墙皮大范围剥落且有裂缝。图6($e_3$)、图6($e_4$)和图6($e_6$)中的裂缝均被误识别为剥落。图6($e_5$)中的分割画面干净且边缘分割完整。图6($f_1$)存在典型的复杂微裂缝。剥落、裂缝、微裂缝这三种类型同时存在的情况在建筑结构震后最为常见。图6($f_3$)、图6($f_4$)和图6($f_6$)均未成功识别微裂缝和剥落连接处。图6($f_5$)将三种裂缝准确识别且对连接点的处理最好，分割最为完整。

分割结果显示，在未添加注意力模块时，对图片裂缝进行分割的U-Net网络能够正确预测大部分像素的类别，但漏分、错分现象较多，容易将裂缝和微裂缝混淆，且边缘检测效果较差。相比之下，在添加CBAM或NA注意力模块之后，模型精度进一步提升，特别是结合两种注意力模块后，分割精度达到了最高水平。分割结果表现出对特征信息的捕捉能力更准确，对边界划分更规整，对细小缺陷的识别度有所提升，分割精度明显提高。分割结果证实了注意力模块在多目标语义分割中的有效性和必要性，为砌体结构震害特征智能检测提供了有力支持。相较于其他模型，本研究提出的DA U-Net-VGG16改进模型对三种震害类型识别最准确，能很好处理裂缝交点问题，并保持裂缝边缘信息，使得分割画面细致完整。

### 3.3　裂缝实际宽度计算

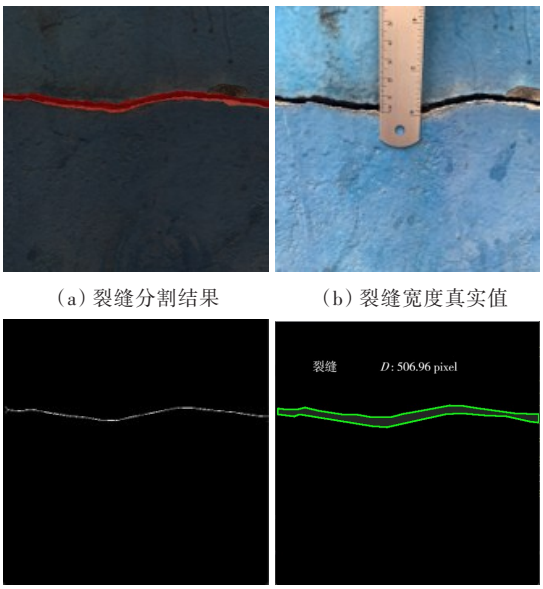
裂缝分割完成后，宽度计算通过提取裂缝骨架进行。首先，对待检测裂缝图像执行骨架化操作，将裂缝宽度压缩至一个像素，以提取裂缝的主干。随后，应用膨胀操作从中心线向外扩展，形成一定宽度线条。但是输出结果以像素为单位，无法用于实际，可通过比例因子将神经网络输出裂缝宽度从像素单位转换为实际物理单位，如式(10)所示：

$$A = k \cdot D \tag{10}$$

式中：$A$为裂缝实际宽度，mm；$k$为比例因子，表示每个像素的实际测量值，mm/pixel；$D$为网络预测裂缝宽度，pixel。

在拍摄照片时，角度、距离、高度及相机参数等都会直接影响裂缝在图像中的大小，例如：同一裂缝在不同拍摄距离下，远距离拍摄时呈现较小的尺寸，而近距离拍摄时则显得更大。为确保准确转换，必须选择适当的$k$为后续的测量和分析提供依据。

实际拍摄图片的测量结果如图7所示。从图7可以看出，裂缝宽度实际测量值$A$为5 mm，模型分割结果为“裂缝”，网络预测裂缝宽度$D$为506.96 pixel，通过计算可知$k$=0.009 86 mm/pixel，该$k$值可用于测量同一拍摄条件下的其他裂缝。




(a) 裂缝分割结果　(b) 裂缝宽度真实值

(c) 提取裂缝骨干　(d) 网络预测裂缝宽度

**图7**　裂缝检测步骤

**Fig. 7**　Steps for crack detection

《建(构)筑物地震破坏等级划分》(GB/T 24335—2009)规定:砌体结构的裂缝宽度在1.5 mm以下属于微裂缝,对构件承载力无明显影响;宽度在1.5 mm以上,但小于3 mm的裂缝为明显裂缝,裂缝几乎贯穿墙厚;宽度在3 mm以上的裂缝为严重裂缝。业界在进行房屋结构安全评估工作时,需要综合考虑主要承重构件、非承重构件和附属结构的破坏情况。当多数承重墙出现轻微裂缝、部分墙体有明显裂缝、个别墙体有严重裂缝,以及多数附属构件出现明显破坏时,建筑达到Ⅲ级破坏。此时,房屋结构的基本功能受到影响,但经修复可重新投入使用。准确计算裂缝宽度等信息,有助于更精准地评定裂缝破坏程度,进而获取精确的工程震害数据,为应急救援工作、震后房屋快速评定提供一种智能化手段。

## 4 结论

本研究提出一种砌体结构震害裂缝快速检测方法,并基于深度学习语义分割理论,探讨了震后砌体结构破坏特征,量化了裂缝参数,为提高地震后结构安全评估效率提供了技术支持。

本文主要结论如下:

1) U-Net主干网络的选择对模型性能产生显著影响,其中VGG16的表现比ResNet50的更佳。添加CBAM模块可提高$E_{\mathrm{MIOU}}$和$E_{\mathrm{MPA}}$,而添加NA模块则能进一步降低损失。当同时使用CBAM和NA模块时,模型性能得到进一步提升。本研究提出的DA U-Net-VGG16模型在分类准确率和分割精度方面表现最优,其中的$E_{\mathrm{MPA}}$、$E_{F_1}$和$E_{\mathrm{MIOU}}$分别达到了91.57%、84.66%和81.50%。综合利用CBAM和NA模块,可以有效提高分割模型在震害裂缝识别方面的性能。

2) 本研究基于分割结果,对裂缝进行量化分析,采用中轴骨架化方法提取裂缝的骨架结构,凸显裂缝主要特征,并使用比例因子将裂缝像素级参数转化成以毫米为单位的实际参数,进而得到输入图像中每条裂缝的真实长度、宽度等信息。

本研究为裂缝参数量化提供了新的可行性途径,但在精准确定比例因子数值上还存在某些限制。后续研究将考虑不同拍摄条件对裂缝参数的影响,期望得到更精确的测量结果。

### [参考文献]

[1] 陈忠范,陆飞,黄际洸,等.村镇砌体结构建筑抗震技术手册[M].南京:东南大学出版社,2012:1-8.
CHEN Zhongfan, LU Fei, HUANG Jiguang, et al. Handbook of seismic technology for masonry buildings in villages and towns [M]. Nanjing: Southeast University Press, 2012: 1-8.

[2] 郭仕群,吴传文,王亚莉,等.工程结构抗震设计[M].北京:机械工业出版社,2018:175-182.
GUO Shiqun, WU Chuanwen, WANG Yali. Seismic design of engineering structure [M]. Beijing: China Machine Press, 2018: 175-182.

[3] DENG L W, WANG Y. Post-disaster building damage assessment based on improved U-Net [J]. Scientific Reports, 2022, 12(1): 15862. DOI: 10.1038/s41598-022-20114-w.

[4] 闫绪娴,刘杨,王俊丽,等.智慧应急赋能城市减灾:本土建构、现实困境与提升路径[J].灾害学,2024,39(02):201-207. DOI: 10.3969/j.issn.1000-811X.2024.02.029.
YAN Xuxian, LIU Yang, WANG Junli, et al. Smart emergency enabling urban disaster mitigation: local constructs, realistic dilemmas and enhancement paths [J]. Journal of Catastrophology, 2024, 39(02): 201-207. DOI: 10.3969/j. issn. 1000-811X. 2024.02.029.

[5] 中国地震局.建(构)筑物地震破坏等级划分:GB/T 24335—2009[S].北京:中国标准出版社,2009.
China Earthquake Administration. Standard for classification of earthquake damage of buildings: GB/T 24335—2009[S]. Beijing: Standards Press of China, 2009.

[6] 陈赟,李倩.基于复杂网络的隧道施工塌方风险演化研究[J].长沙理工大学学报(自然科学版),2022,19(2):129-137. DOI: 10.19951/j.cnki.1672-9331.2022.02.013.
CHEN Yun, LI Qian. Research on evolution of collapse risk in tunnel construction based on complex network [J]. Journal of Changsha University of Science & Technology (Natural Science), 2022, 19(2): 129-137. DOI: 10.19951/j.cnki.1672-9331.2022.02.013.

[7] 邓露,褚鸿鹄,龙砺芝,等.基于深度学习的土木基础设施裂缝检测综述[J].中国公路学报,2023,36(2):1-21. DOI: 10.19721/j. cnki. 1001-7372.2023.02.001.
DENG Lu, CHU Honghu, LONG Lizhi, et al. Review of deep learning-based crack detection for civil infrastructures [J]. China Journal of Highway and Transport, 2023, 36(2): 1-21. DOI: 10.19721/j.cnki.1001-7372.2023.02.001.

[8] 眭海刚,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):1008-1019. DOI:10.13203/j.whugis20190070.
SUI Haigang, LIU Chaoxian, HUANG Lihong, et al. Application of remote sensing technology in earthquake-induced building damage detection[J]. Geomatics and Information Science of Wuhan University, 2019, 44(7): 1008-1019. DOI: 10.13203/j.whugis20190070.

[9] 李子奇,蒋柱虎,王力,等.基于深度学习的工程结构损伤识别研究进展[J].中国安全生产科学技术,2022,18(12):43-48. DOI:10.11731/j.issn.1673-193x.2022.12.006.
LI Ziqi, JIANG Zhuhu, WANG Li, et al. Research progress in damage identification of engineering structure based on deep learning[J]. Journal of Safety Science and Technology, 2022, 18(12): 43-48. DOI: 10.11731/j.issn.1673-193x.2022.12.006.

[10] ÖZGENEL Ç F, SORGUÇ A G. Performance comparison of pretrained convolutional neural networks on crack detection in buildings[C]//Proceedings of the 35th International Symposium on Automation and Robotics in Construction (ISARC). Berlin, Germany: IAARC, 2018: 693-700. DOI: 10.22260/isarc2018/0094.

[11] HALLEE M J, NAPOLITANO R K, REINHART W F, et al. Crack detection in images of masonry using CNNs[J]. Sensors, 2021, 21(14): 4929. DOI: 10.3390/s21144929.

[12] BRACKENBURY D, BRILAKIS I, DEJONG M. Automated defect detection for masonry arch bridges[C]//Proceedings of International Conference on Smart Infrastructure and Construction 2019 (ICSIC). Cambridge, UK: ICE Publishing, 2019: 3-9. DOI: 10.1680/icsic.64669.003.

[13] FERGUSON M K, RONAY A, LEE Y T T, et al. Detection and segmentation of manufacturing defects with convolutional neural networks and transfer learning[J]. Smart and Sustainable Manufacturing Systems, 2018, 2: 10.1520/SSMS20180033. DOI: 10.1520/SSMS20180033.

[14] ALI L. Damage detection and localization in masonry structure using faster region convolutional networks[J]. International Journal of GEOMATE, 2019, 17(59): 98-105. DOI: 10.21660/2019.59.8272.

[15] DAIS D, BAL İ E, SMYROU E, et al. Automatic crack classification and segmentation on masonry surfaces using convolutional neural networks and transfer learning[J]. Automation in Construction, 2021, 125: 103606. DOI: 10.1016/j.autcon.2021.103606.

[16] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, US: IEEE, 2015: 3431-3440. DOI: https://doi.org/10.48550/arXiv.1411.4038.

[17] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation [M]. Cham, Germany: Springer International Publishing, 2015: 234-241. DOI: 10.1007/978-3-319-24574-4_28.

[18] 肖成勇,李擎,王莉,等.基于CBAM-Unet的铁矿球团边缘分割试验方法[J].烧结球团,2022,47(2):8-15,23. DOI:10.13403/j.sjqt.2022.02.018.
XIAO Chengyong, LI Qing, WANG Li, et al. Edge segmentation experimental method of iron ore pellets based on CBAM-Unet[J]. Sintering and Pelletizing, 2022, 47(2): 8-15, 23. DOI: 10.13403/j.sjqt.2022.02.018.

[19] 胡帅,李华玲,郝德琛.改进U-Net的多级边缘增强医学图像分割网络[J].计算机工程,2024,50(4):286-293. DOI:10.19678/j.issn.1000-3428.0067779.
HU Shuai, LI Hualing, HAO Dechen. Improved multistage edge-enhanced medical image segmentation network of U-Net[J]. Computer Engineering, 2024, 50(4): 286-293. DOI: 10.19678/j.issn.1000-3428.0067779.

[20] 刘帅,葛浙东,刘晓彤,等.改进UNet++的杉木CT图像年轮分割[J].计算机工程与应用,2024,(5):232-239.
LIU Shuai, GE Zhedong, LIU Xiaotong, et al. Improved UNet++ for tree rings segmentation of cedar wood CT Images[J]. Computer Engineering and Applications, 2024, (5):232-239.

[21] IBRAHIM Y, NAGY B, BENEDEK C. CNN-based watershed marker extraction for brick segmentation in masonry walls [M]. Cham, Germany: Springer International Publishing, 2019: 332-344. DOI: 10.1007/978-3-030-27202-9_30.

[22] DANG L M, WANG H X, LI Y F, et al. Deep learning-based masonry crack segmentation and real-life crack length measurement [J]. Construction and Building Materials, 2022, 359: 129438. DOI: 10.1016/j.conbuildmat.2022.129438.

[23] LOVERDOS D, SARHOSIS V. Automatic image-based brick segmentation and crack detection of masonry walls using machine learning [J]. Automation in Construction, 2022, 140: 104389. DOI: 10.1016/j.autcon.2022.104389.

[24] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [M]. Cham, Germany: Springer International Publishing, 2018: 3-19. DOI: 10.1007/978-3-030-01234-2_1.

(下转第155页)

synergy based on entropy weight-TOPSIS[J]. Journal of Engineering Management, 2022, 36(3): 23-28. DOI: 10.13991/j.cnki.jem.2022.03.005.

[20] 武慧荣, 陈少阳, 崔淑华. 基于熵权TOPSIS模型的大宗货物运输方式综合评价[J]. 重庆理工大学学报(自然科学), 2022, 36(6): 254-260.
WU Huirong, CHEN Shaoyang, CUI Shuhua. Evaluation of bulk goods transportation mode based on entropy TOPSIS method [J]. Journal of Chongqing University of Technology (Natural Science), 2022, 36 (6): 254-260.

[21] 谢庆元, 吴其育, 余亦彧. 不同交通管理策略下异质交通流稳定性分析[J]. 交通科学与工程, 2023, 39 (6): 123-133. DOI: 10.16544/j. cnki. cn43-1494/u. 202207262023121998.
XIE Qingyuan, WU Qiyu, YU Yiyu. Stability analysis of heterogeneous traffic flow under different traffic management strategies[J]. Journal of Transport Science and Engineering, 2023, 39 (6): 123-133. DOI: 10.16544/j.cnki.cn43-1494/u.202207262023121998.

(责任编辑:彭三军;校对:毛娜)

(上接第145页)

[25] WANG X L, GIRSHICK R, GUPTA A, et al. Non-local neural networks[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018: 7794-7803. DOI: 10. 1109/CVPR.2018.00813.

[26] SHI W Z, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]// Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 1874-1883. DOI: 10.1109/CVPR.2016.207.

[27] 中国建筑科学研究院. 2008年汶川地震建筑震害图片集[M]. 北京: 中国建筑工业出版社, 2008: 21-73.
China Academy of Building Research. Photo collection of 2008 Wenchuan earthquake damage to buildings [M]. Beijing: China Architecture & Building Press, 2008: 21-73.

[28] 孙柏涛. 四川省芦山"4·20" 7.0级强烈地震建筑物震害图集[M]. 北京: 地震出版社, 2014: 13-289.
SUN Botao. Atlas of earthquake damage of buildings in Lushan, Sichuan Province "4·20" M 7.0 strong earthquake[M]. Beijing: Seismological Press, 2014: 13-289.

[29] 戴君武. 四川九寨沟7.0级地震之工程震害[M]. 北京: 地震出版社, 2018: 39-288.
DAI Junwu. Engineering damage in Jiuzhaigou M 7.0 earthquake[M]. Beijing: Seismological Press, 2018: 39-288.

[30] 刘广林, 蒋志明, 邓琛, 等. 切缝间距对环向切缝管气隙聚能效应的影响[J]. 交通科学与工程, 2023, 39 (3): 90-97. DOI: 10.16544/j. cnki. cn43-1494/u. 20211105001.
LIU Guanglin, JIANG Zhiming, DENG Chen, et al. Influence of the slotted spacing on air-gap gathering energy effect of the circumferential slit tube[J]. Journal of Transport Science and Engineering, 2023, 39(3): 90-97. DOI: 10.16544/j. cnki. cn43-1494/u. 20211105001.

(责任编辑:彭三军;校对:毛娜)