

# 控制与决策

Control and Decision

## 基于特征共享双头Cascade R-CNN的混凝土细观损伤特征检测

赵亮, 高升伦, 陈俊英, 李佳佳

引用本文:

赵亮,高升伦,陈俊英,李佳佳. 基于特征共享双头Cascade R-CNN的混凝土细观损伤特征检测[J]. *控制与决策*, 2022, 37(7): 1745–1751.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2021.0124>

---

### 您可能感兴趣的其他文章

Articles you may be interested in

#### 复杂背景下全景视频运动小目标检测算法

Panoramic video motion small target detection algorithm in complex background

控制与决策. 2021, 36(1): 249–256 <https://doi.org/10.13195/j.kzyjc.2019.0686>

#### 周围神经MicroCT图像中神经束轮廓获取算法的改进

An improved approach to obtain contours of fascicular groups from MicroCT images of peripheral nerve

控制与决策. 2021, 36(7): 1601–1610 <https://doi.org/10.13195/j.kzyjc.2019.1664>

#### 基于双分支特征融合的场景文本检测方法

A scene text detection based on dual-path feature fusion

控制与决策. 2021, 36(9): 2179–2186 <https://doi.org/10.13195/j.kzyjc.2020.0002>

#### 多目标小尺度车辆目标检测方法

Multi-target and small-scale vehicle target detection method

控制与决策. 2021, 36(11): 2707–2712 <https://doi.org/10.13195/j.kzyjc.2020.0635>

#### 基于卷积神经网络的云雾遮挡舰船目标识别

Obscured ship target recognition based on convolutional neural network

控制与决策. 2021, 36(3): 661–668 <https://doi.org/10.13195/j.kzyjc.2019.0781>

# 基于特征共享双头 Cascade R-CNN 的 混凝土细观损伤特征检测

赵亮<sup>†</sup>, 高升伦, 陈俊英, 李佳佳

(西安建筑科技大学 信息与控制工程学院, 西安 710055)

**摘要:** 混凝土内部损伤破坏形态具有明显的离散性和随机性, 内部损伤特征检测是混凝土细观研究的重要内容. 针对已有混凝土结构内部损伤特征检测模型精度低的问题, 提出一种特征共享双头 Cascade R-CNN 模型对混凝土 CT 图像的损伤特征进行检测. 首先, 为了有效识别损伤特征的空间信息, 构建具有空间敏感性的 fc-head (fully connected head) 与空间相关性的 conv-head (convolution head) 相结合的 Cascade R-CNN 网络模型; 其次, 通过特征共享的方法将检测网络各层级分类信息进行融合, 提升低 IOU (intersection over union) 阈值 (0.5 ~ 0.7) ROI (regions of interest) 检测任务的精度. 实验结果表明, 所提方法在检测混凝土 CT 图像的损伤特征中平均精度达到 91.31%, 比原始的 Cascade R-CNN 提高 3.04%, 低 IOU 阈值 (0.5 ~ 0.7) ROI 平均精度提高 1.49%, 该模型可以较好地检测混凝土 CT 图像中检测出细观损伤部分, 具有精度高、运算简单、易于工程实现等特点.

**关键词:** 深度学习; 目标检测; 混凝土 CT 图像; Cascade R-CNN; double-head; ROI

中图分类号: TP183

文献标志码: A

DOI: 10.13195/j.kzyjc.2021.0124

开放科学(资源服务)标识码(OSID):



**引用格式:** 赵亮, 高升伦, 陈俊英, 等. 基于特征共享双头 Cascade R-CNN 的混凝土细观损伤特征检测[J]. 控制与决策, 2022, 37(7): 1745-1751.

## Feature detection of concrete mesoscopic damage based on feature sharing double-head Cascade R-CNN

ZHAO Liang<sup>†</sup>, GAO Sheng-lun, CHEN Jun-ying, LI Jia-jia

(College of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China)

**Abstract:** The internal damage of concrete has obvious characteristics of discreteness and randomness, and the detection of internal damage characteristics is an important content of concrete mesoscopic research. To solve the problem of low precision of existing models, this paper proposes a double-head Cascade R-CNN model with feature sharing to detect the damage features of concrete CT images. Firstly, a Cascade R-CNN network model, which combines spatially sensitive fc-head (fully connected head) and spatially correlation conv-head (convolution head), is constructed to effectively identify the spatial information of the damage feature. Then, the classification information of each level of the detection network is merged through the feature sharing method, which improves the precision of low IOU (intersection over union) threshold (0.5 ~ 0.7) ROI (regions of interest) detection tasks. The results show that the average precision of the proposed method is 91.31%, which is 3.04% higher than that of the original. The average accuracy of low IOU threshold (0.5 ~ 0.7) ROI is improved by 1.49%. The proposed model can better detect the mesoscopic damage part from the concrete CT image with the features of high precision, simple computation and easy engineering realization.

**Keywords:** deep learning; object detection; concrete CT image; Cascade R-CNN; double-head; ROI

## 0 引言

混凝土结构是由水泥砂浆、粗骨料、细骨料、掺杂相、孔洞及裂纹等组成的多相、非均质复合材料,其

细观结构损伤和宏观性能都与内部孔洞、裂纹的扩展有关. 近年来,随着 X 射线计算机断层扫描<sup>[1-3]</sup>(X-CT)技术的发展,无损检测混凝土内部损伤特征成为

收稿日期: 2021-01-21; 录用日期: 2021-04-07.

基金项目: 国家自然科学基金项目(51209167; 12002251); 陕西省自然科学基金项目(2019JM-474); 西安市科技计划项目(2020KJRC0055); 陕西省岩土与地下空间工程重点实验室开放基金项目(YT202004).

责任编辑: 巩敦卫.

<sup>†</sup>通讯作者. E-mail: zhaoliang@xauat.edu.cn.

了混凝土细观研究热点. 文献[4]在CT图像信息上使用迭代局部阈值处理技术分割提取裂纹特征; 文献[5]使用灰度阈值从混凝土CT图像中分割提取了孔洞和裂纹, 并定量分析了损伤过程; 文献[6]开展了针对循环混凝土材料CT图像的特征结构边缘提取研究, 并根据特征结构进行了细观数值模拟研究; 文献[7-8]采用平均灰度值方法分析了混凝土断裂过程. 传统的针对混凝土细观结构检测主要通过对图像像素的处理辨识出混凝土内部的破损结构. 由于混凝土内部结构的特殊性, CT图像中产生的噪声会影响上述方法识别混凝土中孔洞及裂纹的精度, 从而在一定程度上影响混凝土细观破损研究. 文献[9]将混凝土CT图像中的孔隙和裂缝作为检测目标, 采用深度学习的方法构建检测器, 实现了对细观损伤特征的识别, 取得了较好的检测精度, 但在区分前景目标和背景时, 使用单一固定的阈值0.5, 导致训练时产生大量假阳性目标框, 进而影响了目标检测精度.

目标检测是在复杂的图像背景中识别出感兴趣的目标, 标定出一个精确的目标边框(bounding box), 并判断该边框中目标所属的类别<sup>[10-13]</sup>. Krizhevsky<sup>[14]</sup>在2012年提出了卷积神经网络(convolutional neural network, CNN) AlexNet. 卷积神经网络因为权值共享及卷积计算的特点, 能够兼顾噪声抑制和图像细节保护. 目前比较经典的目标检测算法R-CNN<sup>[15]</sup>、fast R-CNN<sup>[16]</sup>、faster R-CNN<sup>[17]</sup>、mask R-CNN<sup>[18]</sup>皆采用two-stage的优化方式, 即首先根据预定义的anchor boxes生成候选框, 然后对候选框进行逐步优化处理. 级联R-CNN(Cascade R-CNN)<sup>[19]</sup>在遵循这一思想的基础上, 采用级联多个不同IOU阈值的检测网络以实现不断优化预测结果的目的.

Cascade R-CNN在COCO、PASCAL VOC目标检测数据集上取得了很好的效果, 但在混凝土CT图像目标检测过程中发现, Cascade R-CNN网络在stage 3训练过程中通常使用较高的IOU(0.7)阈值区分图像的前景和背景, 因此将大量IOU位于0.5~0.7区间ROI标记为背景, 使得这些ROI得不到有效回归. 这种现象导致一些低IOU(0.5~0.7)的ROI可能在前一stage训练中学习得很好, 而在stage 3被非极大值抑制算法(NMS)<sup>[20]</sup>中止学习, 进而影响低IOU(0.5~0.7)的ROI检测精度.

针对上述问题, 本文把混凝土CT实验图像中孔洞及裂纹作为检测目标, 提出一种基于特征共享双头Cascade R-CNN模型(feature sharing double-head Cascade R-CNN, FSDH Cascade R-CNN), 其double-

head<sup>[21]</sup>由fc-head和conv-head组成: fc-head更具有空间敏感性, 有利于分类任务; 而conv-head具有空间相关性, 更加有利于对目标框进行回归. 结合fc-head与conv-head可以更好地执行分类和回归任务. 同时, 本文在分类网络中引入特征共享机制, 每一阶段的预测可以参考其他阶段的预测信息, 旨在保持高IOU的ROI检测精度的基础上提升低IOU的ROI检测精度. 在回归网络中, 通过级联的方法逐步细化ROI, 并通过反向传播不断优化使目标框更为精准, 从而提升整体性能. 实验表明, 所提方法在检测混凝土CT图像的损伤特征中精度达到91.31%, 可有效提升检测任务的精度.

## 1 FSDH Cascade R-CNN

在FPN(feature pyramid network)<sup>[22]</sup>的Cascade R-CNN为主要结构的基础上, 引入特征共享的double-head结构以提高检测网络在不同高低IOU阈值下的性能, 同时改善低IOU(0.5~0.7)ROI检测能力不足的问题. 网络结构如图1所示.

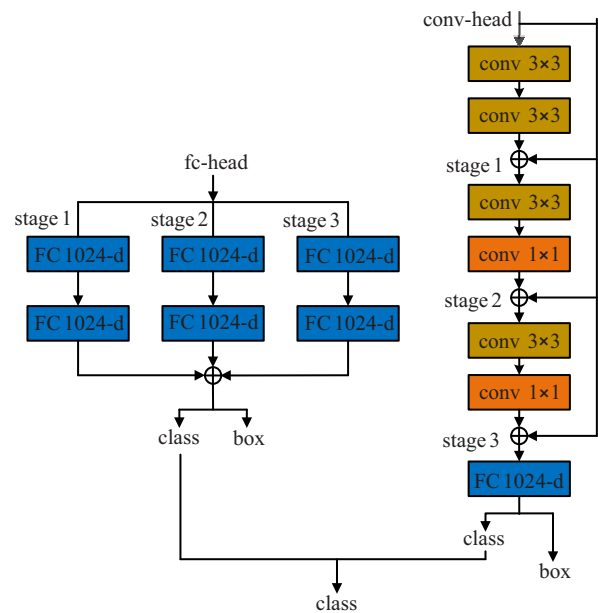


图1 FSDH Cascade R-CNN网络结构

### 1.1 分类网络

目标检测网络需要为每个检测对象分配一个类别标签. 在Cascade R-CNN中, 每个stage在给定IOU阈值下负责预测特定类的置信度分数. 在这个过程中, 当输入一个低IOU(0.5~0.7)目标框时, 它可能在stage 2获得极高的分数, 但在stage 3中却获得较低的分. 对于stage 3, 如果能够提高低IOU目标框的得分概率, 则在训练阶段可以召回更多目标框. 因此stage 3的关键问题是优化低IOU目标框的置信度分数. 基于此问题在分类网络中引入特征共享fc-head

模块. 并行搭建3个不同 IOU (0.5、0.6、0.7) 阈值分类 stage, 对于每个分类 stage, 直接使用 FPN 中的两个 1024-d 全连接层作为共享基, 共享结构如图1左侧所示. 实验中, FPN 池化后的特征 ( $7 \times 7 \times 256$ ) 不仅通过其自身的全连接层, 而且通过其他两个在较低 IOU 阈值上训练的全连接层, 实现过程如下:

$$\begin{aligned} C_1 &= W_1^2(W_1^1(X_i^{\text{cls}})), \\ C_2 &= W_2^2(W_2^1(X_i^{\text{cls}})), \\ &\vdots \\ C_i &= W_i^2(W_i^1(X_i^{\text{cls}})), \\ C &= C_1 + C_2 + \dots + C_i. \end{aligned} \quad (1)$$

其中:  $W_i^1$  和  $W_i^2$  表示第  $i$  阶段的两个全连接层的特征转换,  $X_i^{\text{cls}}$  表示池化后的特征,  $C_i$  表示经两个全连接层转换以后的特征,  $C$  是所有  $C_i$  通过对应元素相加得到的综合特征. 该方法的优点在于分类信息可以跨阶段流动, 并通过反向传播优化所有 IOU 阈值对应的特征转换系数, 从而使训练过程更加稳定.

## 1.2 回归网络

因为回归需要保留有效的空间信息, 所以在回归网络中选用 conv-head 取代 FPN 中常用的全连接层. 虽然引入 conv-head 对目标框的回归有一定的帮助, 但是在低 IOU 阈值下仍然会出现之前提到的问题, 造成 stage 3 和 stage 2 的检测结果差距较大. 因此, 在 conv-head 中进行回归特征共享. 与分类网络中的并行特征共享不同, 回归网络在卷积层间串行共享回归特征, 这种方法的优点在于目标框可以在各 stage 上逐步进行优化, 如图1右侧所示. 在 stage 1 添加了两个  $3 \times 3$  的卷积层, 其他两个 stage 仅包含一个  $3 \times 3$  的卷积层和一个  $1 \times 1$  的卷积层. 为捕捉到回归网络中的分类信息, 在 stage 3 后增加一个 1024-d 全连接层. 实现过程如下:

$$\begin{aligned} B_1 &= X_i^{\text{box}} + F_1^2(F_1^1(X_i^{\text{box}})), \\ B_2 &= X_i^{\text{box}} + G_1(F_2^1(B_1)), \\ &\vdots \\ B_i &= X_i^{\text{box}} + G_{i-1}(F_i^1(B_{i-1})). \end{aligned} \quad (2)$$

其中:  $F_i$  表示  $3 \times 3$  卷积层对应的特征转换,  $G_i$  表示  $1 \times 1$  卷积层对应的特征转换,  $B_i$  用于执行目标框预测.

鉴于 fc-head 和 conv-head 结构不同, 捕捉到的目标分类信息是不同的. 因此, 将两个分类信息进行如下融合以达到互补的目的, 即

$$p_i^* = p_i^{\text{fc}} + p_i^{\text{conv}}(1 - p_i^{\text{fc}}) = p_i^{\text{conv}} + p_i^{\text{fc}}(1 - p_i^{\text{conv}}). \quad (3)$$

其中:  $p_i^*$  表示 anchor 被预测为目标概率,  $p_i^{\text{fc}}$  和  $p_i^{\text{conv}}$  分别表示 anchor 在 fc-head 和 conv-head 被预测为目标概率.

## 1.3 损失函数设计

分类网络和回归网络都是由区域建议网络 (RPN)<sup>[17]</sup> 端到端联合训练的, 网络的损失函数  $L$  定义如下:

$$L = \omega^{\text{fc}} L^{\text{fc}} + \omega^{\text{conv}} L^{\text{conv}} + L^{\text{rpn}}. \quad (4)$$

其中:  $\omega^{\text{fc}}$  和  $\omega^{\text{conv}}$  分别表示 fc-head 和 conv-head 的权重系数, 本文  $\omega^{\text{fc}} = 2.0$ ,  $\omega^{\text{conv}} = 2.5$ ,  $L^{\text{fc}}$ 、 $L^{\text{conv}}$ 、 $L^{\text{rpn}}$  分别表示 fc-head、conv-head 和 RPN 的损失函数值.

fc-head 的损失包括分类损失和目标框回归损失, 损失函数  $L^{\text{fc}}$  定义如下:

$$L^{\text{fc}} = \lambda^{\text{fc}} L_{\text{cls}}^{\text{fc}} + (1 - \lambda^{\text{fc}}) L_{\text{reg}}^{\text{fc}}. \quad (5)$$

其中:  $L_{\text{cls}}^{\text{fc}}$  和  $L_{\text{reg}}^{\text{fc}}$  分别表示 fc-head 中分类损失和目标框回归损失;  $\lambda^{\text{fc}}$  表示平衡 fc-head 中两个损失之间的权重系数, 本文取  $\lambda^{\text{fc}} = 0.7$ .

同理, conv-head 损失函数定义如下:

$$L^{\text{conv}} = (1 - \lambda^{\text{conv}}) L_{\text{cls}}^{\text{conv}} + \lambda^{\text{conv}} L_{\text{reg}}^{\text{conv}}, \quad (6)$$

其中  $L_{\text{cls}}^{\text{conv}}$  和  $L_{\text{reg}}^{\text{conv}}$  分别表示 conv-head 中分类损失和目标框回归损失. 不同于 fc-head 损失函数, 因为目标框回归是 conv-head 的主要任务, 所以与回归损失  $L_{\text{reg}}^{\text{conv}}$  相乘的平衡权重系数  $\lambda^{\text{conv}}$  取值较大, 本文取  $\lambda^{\text{conv}} = 0.8$ .

fc-head 和 conv-head 中  $L_{\text{cls}}^{\text{fc}}$  和  $L_{\text{cls}}^{\text{conv}}$  均采用交叉熵损失函数, 其定义如下:

$$L_{\text{cls}}(p_i, p_i^*) = -[p_i \log p_i^* + (1 - p_i) \log(1 - p_i^*)]. \quad (7)$$

其中:  $p_i$  表示 anchor 的真实标签, 正 anchor 的真实标签为 1, 负 anchor 的真实标签为 0;  $p_i^*$  表示 anchor 被预测为目标的概率.

fc-head 和 conv-head 中  $L_{\text{reg}}^{\text{fc}}$  和  $L_{\text{reg}}^{\text{conv}}$  均采用 smooth $_{L_1}$  损失函数  $L_{\text{reg}}$ , 其定义如下:

$$L_{\text{reg}}(t_i, t_i^*) = \sum_{i \in x, y, w, h} \text{smooth}_{L_1}(t_i - t_i^*). \quad (8)$$

$$\text{smooth}_{L_1}(x) = \begin{cases} 0.5x^2, & |x| < 1; \\ |x| - 0.5, & \text{otherwise.} \end{cases} \quad (9)$$

其中:  $t_i$  和  $t_i^*$  分别表示预测框的位置和真实目标框的位置.

## 2 实验过程及结果分析

### 2.1 实验环境

本次实验计算机配置如下:CPU 为 Intel Core i7 10700k@3.80 GHz,显卡为 NVIDIA GeForce GTX 2080Ti,主板为 ROG STRIX Z490-A GAMING,内存为 32 G,操作系统为 Ubuntu18.04, cuda10.0, cudnn 7.6.4. 在 mmdetection 深度学习框架上,采用 Python 编程语言.

### 2.2 实验数据集

本文实验采用规格为 60 mm×120 mm 的一级配混凝土圆柱体试件,砂料是从西安铲河中取得的,骨料是粒径大约为 5~20 mm 的泾河卵石. 采用西安市中心医院影像中心设备 MarconiM8000 螺旋 CT 扫描仪,在不同的静力压缩荷载下( $\delta = 0$  Mpa、24.59 Mpa、34.89 Mpa、32.22 Mpa(卸载)及 12.7 Mpa(宏观破坏)),对混凝土圆柱试件沿横断面连续扫描,扫描厚度大约为 1.0 mm,共扫描 5 次,每次 255 个断面,通过 CT 扫描得到混凝土圆柱体试件的二维断层序列图像后,将其转化成计算机可以接受的数据格式. 混凝土 CT 图像数据标注样本如图 2 所示.

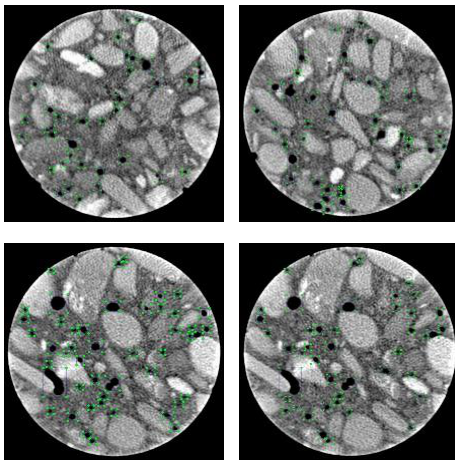


图 2 混凝土 CT 图像数据

对原始输入图像通过随机水平镜像、改变对比度等数据增强(data augmentation)方法扩大训练样本集的个数,平均每幅图像约含 30 个目标,总计约 120 000 个目标. 随机挑选 800 张混凝土 CT 图像作为测试集,同样随机挑选出 800 张混凝土 CT 图像作为验证集,剩下的样本全部用于训练,训练前将所有输入图像大小调整为 640×640.

### 2.3 消融实验

针对本文设计的网络,训练中设置起始的学习率  $lr=0.002$ ,初始的 500 次迭代中设置学习率  $warmup\_ratio=1/3$  进行预热,分别在 4 000 和 6 000 次迭代时

将学习率减小为之前的 0.1 倍,采用随机梯度下降(SGD)算法优化网络参数,动量因子  $momentum=0.9$ ,权重衰减因子  $weight\_decay=0.0001$ . 训练过程中损失函数随迭代次数的变化曲线如图 3 所示.

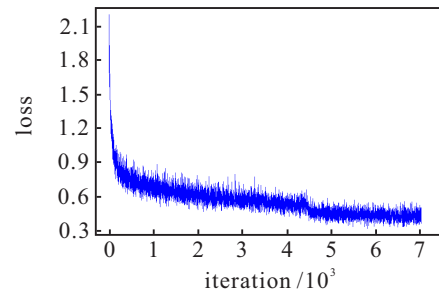


图 3 损失函数随迭代次数的变化

为了验证 fc-head 和 conv-head 对 stage 2、stage 3 及网络整体性能的影响,采用如下方式进行消融实验分析:删除本文网络中的 conv-head 模块,保留 fc-head 模块,将该消融模型记为 fc;删除本文网络中的 fc-head 模块,保留 conv-head 模块,将该消融模型记为 conv. 表 1 为消融实验的检测 AP(average precision)值,图 4~图 6 分别给出了消融实验  $\Delta AP$ (stage 3 与 stage 2 的精度差)值.

表 1 消融实验结果

ablation experiment	fc-head	conv-head	stage 3	stage 2
			AP / %	AP / %
fc	✓	✗	85.37	84.56
conv	✗	✓	86.21	85.58
Cascade R-CNN	✗	✗	85.24	84.69
FSDH Cascade R-CNN	✓	✓	<b>86.73</b>	<b>85.71</b>

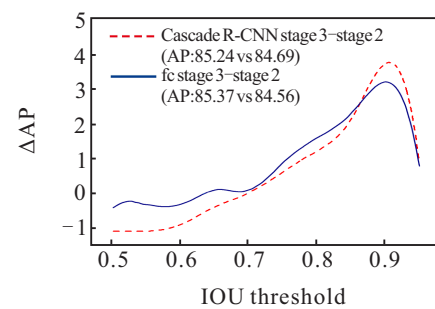


图 4 fc 模块与 Cascade R-CNN 的  $\Delta AP$  对比

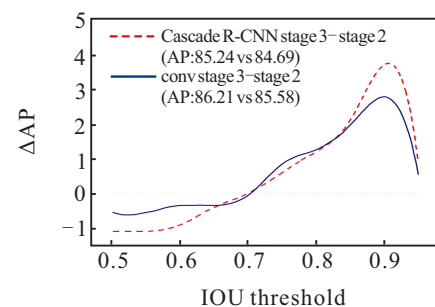


图 5 conv 模块与 Cascade R-CNN 的  $\Delta AP$  对比

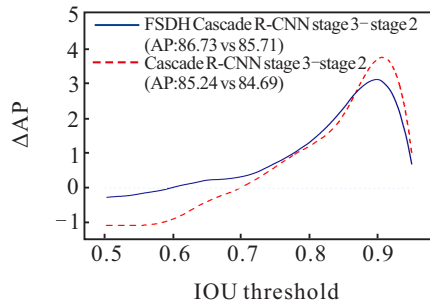


图6 FSDH Cascade R-CNN 与 Cascade R-CNN 的  $\Delta AP$  对比

由表1可以看出,本文网络中使用的fc和conv模块在AP评价指标中都发挥了作用,删除任何一个模块都会导致网络的检测精度下降.其中,当消融实验中缺少conv-head模块时,fc模块stage3的平均精度为85.37%,Cascade R-CNN模型stage3的平均精度为85.24%,相比之下前者的精度优于后者,但相差不大.但是,如图4所示,IOU = 0.5时前者的 $\Delta AP = -0.4$ ,后者的 $\Delta AP = -1.1$ ,由此可见fc模块采用并行特征共享提高了stage3对低IOU目标框的检测精度,使得分更为合理,但对网络的整体性能提升不大,可以在一定程度上与特征共享的回归网络互补.

当消融实验中缺少fc-head模块时,实验结果如图5所示.首先,由于卷积网络充分保留了样本的空间相关性信息,使得网络的整体性能有了很大提高,平均精度提高了0.97%;其次,conv模块采用串行特征共享使stage3对低IOU目标框的检测性能也得到了提升,IOU = 0.5时 $\Delta AP$ 为-0.5,大于-1.1.

图6给出了fc-head和conv-head两个模块同时存在时的结果,由表1可知,stage2和stage3的检测精度有了显著的提高,并且当IOU = 0.5时 $\Delta AP$ 提升到-0.25,改善了stage3对低IOU目标框检测能力不足

的问题,验证了两个模块结合的有效性.

为了进一步验证fc-head模块与conv-head模块结合的有效性,将结合后IOU位于0.5~0.7之间的目标框置信度得分进行可视化.

如图7所示,FSDH Cascade R-CNN在给定的IOU阈值区间内置信度得分更加均衡,可以召回更多低IOU阈值的目标框,提高检测网络的整体性能.

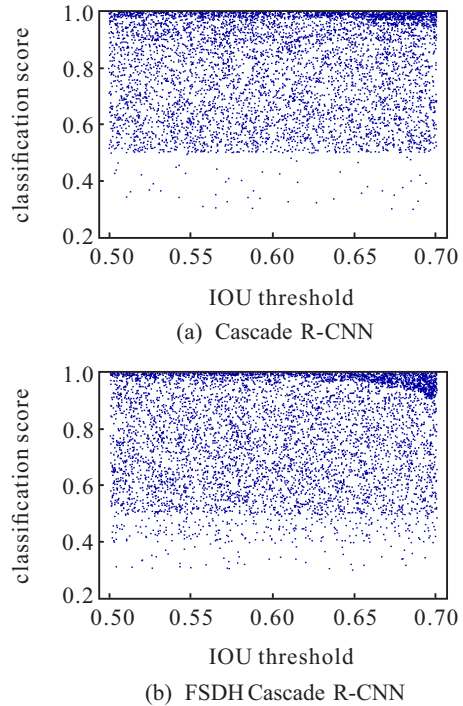


图7 IOU (0.5~0.7) 目标框置信度得分可视化

### 2.4 对比实验结果及分析

本实验使用ResNet-50/101+FPN+FSDH Cascade R-CNN算法和ResNet-50/101+FPN+Cascade R-CNN对混凝土CT图像中孔洞及裂纹数据的训练集迭代7000次,并将训练好的模型在测试集上进行验证,其结果如表2所示.

表2 两种检测算法的性能对比

method	backbone	AP / %	train speed / s	test speed / s
Cascade R-CNN	ResNet-50	86.34	0.199	0.107
Cascade R-CNN	ResNet-101	88.27	0.318	0.138
FSDH Cascade R-CNN	ResNet-50	88.86	0.223	0.129
FSDH Cascade R-CNN	ResNet-101	91.31	0.336	0.157

由表2可知:本文提出的改进算法在相同训练情况下平均精度均优于Cascade R-CNN,ResNet-101+FPN+FSDH Cascade R-CNN取得最高的平均精度91.31%;在相同的backbone情况下,引入FSDH后精度均有提高;此外,虽然改进后的Cascade R-CNN算

法采用double-head结构增加了计算量,但添加的 $1 \times 1$ 卷积层有效地抑制了参数增加,从而使检测网络在提升精度的同时没有过多损失训练时间.

为了进一步验证FSDH Cascade R-CNN算法性能,在相同数据集上与主流神经网络模型进行了对

比,表3给出了对比结果。

表3 FSDH Cascade R-CNN与不同检测算法的性能对比

method	backbone	AP/%	train speed/s	test speed/s
SSD <sup>[23]</sup>	VGG16	69.32	0.076	0.065
YOLOv3 <sup>[24]</sup>	DarkNet-53	74.43	0.065	0.022
YOLOv4 <sup>[25]</sup>	CSPDarkNet-53	85.62	0.058	0.018
faster R-CNN <sup>[9]</sup>	ResNet-50	81.39	0.159	0.089
faster R-CNN <sup>[9]</sup>	ResNet-101	87.08	0.203	0.116
FSDH Cascade R-CNN	ResNet-101	91.31	0.336	0.157

从实验结果可以看出:FSDH Cascade R-CNN在检测精度上优于SSD、YOLOv3、YOLOv4、faster R-CNN;而SSD、YOLOv3、YOLOv4由于没有候选区域提取机制,整体网络结构相对轻量化,缩短了训练时间,但也牺牲了一部分检测精度;而YOLOv4采用CSP模块在保证准确率的同时可以减少计算量,提升了检测速度,达到平均每张图像检测时间0.018 s。为了更直观地显示FSDH Cascade R-CNN算法与其他算法的检测效果对比,图8分别展示了FSDH Cascade R-CNN、ResNet-101+FPN+Cascade R-CNN、ResNet-101+FPN+faster R-CNN及YOLOv3在混凝土CT图像中的检测效果。

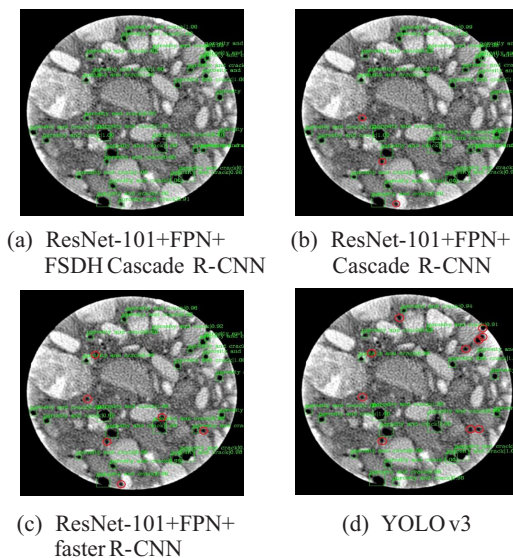


图8 不同算法在混凝土CT图像上的检测效果

图8中将混凝土CT图像测试集孔洞及裂纹命名为porosity and crack。其中:图8(a)中检测出的孔洞及裂纹数量为29,图8(b)中检测出的孔洞及裂纹数量为27,图8(c)中检测出的孔洞及裂纹数量为23,图8(d)中检测出的孔洞及裂纹数量为19。图8中红色标出位置为算法漏检目标。综合上述实验结果可知,FSDH Cascade R-CNN算法在混凝土CT图像损伤特征检测中的检测精度优于其他算法,可检测出更多的目标。

### 3 结论

针对混凝土结构内部损伤特征检测问题,提出了一种基于特征共享双头Cascade R-CNN模型。由具有空间敏感性的fc-head和具有空间相关性的conv-head分别执行分类和回归任务,通过融合fc-head与conv-head的分类信息来提升目标检测网络的整体性能。在分类和回归网络中引入特征共享机制,改善了stage 3低IOU(0.5~0.7)ROI置信度不匹配的问题,将stage 3检测精度提升至86.73%,检测网络整体精度提升至91.31%,并与其他主流的目标检测算法进行了对比实验,结果表明,本文所提模型能够提高混凝土CT图像损伤特征检测精度,为混凝土微观破坏演化机理研究提供有效的技术支撑。FSDH Cascade R-CNN方法对混凝土CT图像损伤特征具有较高的检测精度,但与采用轻量化模型的算法相比,在检测速度上略逊一筹。因此,下一步研究方向是在保证良好检测精度的前提下,进一步优化网络模型,提高网络的检测速度。

#### 参考文献(References)

- [1] Wang P H, Qiao H X, Zhang Y S, et al. Meso-damage evolution analysis of magnesium oxychloride cement concrete based on X-CT and grey-level co-occurrence matrix[J]. Construction and Building Materials, 2020, 255: 119373.
- [2] Wang P H, Qiao H X, Zhang Y S, et al. Three-dimensional characteristics of steel corrosion and corrosion-induced cracks in magnesium oxychloride cement concrete monitored by X-ray computed tomography[J]. Construction and Building Materials, 2020, 246: 118504.
- [3] Xi X, Yang S T. Investigating the spatial development of corrosion of corner-located steel bar in concrete by X-ray computed tomography[J]. Construction and Building Materials, 2019, 221: 177-189.
- [4] Deng H, Fitts J P, Peters C A. Quantifying fracture geometry with X-ray tomography: Technique of iterative local thresholding (TILT) for 3D image segmentation[J]. Computational Geosciences, 2016, 20(1): 231-244.
- [5] Ren W Y, Yang Z J, Sharma R, et al. Three-dimensional in situ XCT characterisation and FE modelling of cracking in concrete[J]. Complexity, 2018, 2018: 1-11.
- [6] Chen A J, Xia X Z, Zhang Q, et al. The meso-level numerical experiment research of the mechanics properties of recycled concrete[J]. Journal of Software, 2012, 7(9): 1932-1940.
- [7] Hong S X, Liu P, Zhang J C, et al. Visual & quantitative identification of cracking in mortar subjected to loads

- using X-ray computed tomography method[J]. *Cement and Concrete Composites*, 2019, 100: 15-24.
- [8] Wan K S, Xue X B. In situ compressive damage of cement paste characterized by lab source X-ray computer tomography[J]. *Materials Characterization*, 2013, 82: 32-40.
- [9] Zhao L, Li C J, Wu X D, et al. Improved damage characteristics identification method of concrete CT images based on region convolutional neural network[J]. *International Journal of Pattern Recognition and Artificial Intelligence*, 2020, 34(6): 2054018.
- [10] Rad M, Lepetit V. BB8: A scalable, accurate, robust to partial occlusion method for predicting the 3D poses of challenging objects without using depth[C]. 2017 IEEE International Conference on Computer Vision (ICCV). Venice, 2017: 3848-3856.
- [11] Kehl W, Manhardt F, Tombari F, et al. SSD-6D: Making RGB-based 3D detection and 6D pose estimation great again[C]. 2017 IEEE International Conference on Computer Vision (ICCV). Venice, 2017: 1530-1538.
- [12] Tekin B, Sinha S N, Fua P. Real-time seamless single shot 6D object pose prediction[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018: 292-301.
- [13] Joshi K A, Thakore D G. A survey on moving object detection and tracking in video surveillance system[J]. *International Journal of Soft Computing & Engineering*, 2012, 2(3): 44-48.
- [14] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [15] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, 2014: 580-587.
- [16] Girshick R. Fast r-cnn[C]. *Proceedings of the IEEE International Conference on Computer Vision*. Washington DC, 2015: 1440-1448.
- [17] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[C]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Piscataway: IEEE, 2015: 1137-1149.
- [18] He K M, Gkioxari G, Dollár P, et al. Mask R-CNN[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 386-397.
- [19] Cai Z W, Vasconcelos N. Cascade R-CNN: Delving into high quality object detection[C]. *The 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, 2018: 6154-6162.
- [20] Neubeck A, Van Gool L. Efficient non-maximum suppression[C]. *18th International Conference on Pattern Recognition (ICPR'06)*. Hong Kong, 2006: 850-855.
- [21] Wu Y, Chen Y P, Yuan L, et al. Rethinking classification and localization for object detection[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, 2020: 10183-10192.
- [22] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, 2017: 936-944.
- [23] 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.
- [24] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, 2016: 779-788.
- [25] Bochkovskiy A, Wang C Y, Liao H Y Mark. YOLOv4: Optimal speed and accuracy of object detection[J/OL]. 2020, arXiv: 2004.10934.

### 作者简介

赵亮(1980—),男,教授,博士,从事智能检测、机器视觉等研究, E-mail: zhaoliang@xauat.edu.cn;

高升伦(1991—),男,硕士生,从事图像处理、人工智能和计算机视觉的研究, E-mail: tg632551422@163.com;

陈俊英(1980—),女,副教授,博士,从事机器学习、智能检测等研究, E-mail: chenjy@xauat.edu.cn;

李佳佳(1989—),男,副教授,博士,从事神经计算与脑认知行为分析等研究, E-mail: lijijia\_dynamics@xauat.edu.cn.

(责任编辑: 闫妍)