

文章编号: 1001-4632 (2020) 05-0115-12

基于2阶级联轻量级卷积神经网络的高铁接触网 悬挂紧固件缺陷识别方法

周 雯¹, 史天运², 李 平³, 马小宁³, 吴艳华³

(1. 中国铁道科学研究院 研究生部, 北京 100081;

2. 中国铁道科学研究院集团有限公司, 北京 100081;

3. 中国铁道科学研究院集团有限公司 电子计算技术研究所, 北京 100081)

摘 要: 基于高铁接触网悬挂运行状态监测图像, 分析监测图像与普通图像的区别; 将紧固件缺陷检测问题转换为紧固件检测和运行状态精细识别2个过程, 提出基于2阶级联卷积神经网络的紧固件缺陷识别方法。首先, 设计紧固件检测网络, 由轻量级特征提取网络、全局注意力模块及相互增强的分类器和检测器组成, 实现目标紧固件实例的高效检测; 然后, 搭建1个包含4个卷积层、2个池化层、1个全连接层和1个输出层的多标签分类网络, 完成紧固件运行状态的精细分类, 实现缺陷识别。运用紧固件缺陷识别方法, 对某高铁线路接触网状态监测图像数据进行试验分析, 结果表明: 2阶级联卷积神经网络的方法可以快速准确地检测紧固件的缺陷, 紧固件定位平均检出率达98.2%, 紧固件缺陷平均识别精度达95.8%, 较单一检测网络提高约21.5%。

关键词: 高速铁路; 2阶级联; 卷积神经网络; 接触网; 缺陷识别; 紧固件; 目标检测

中图分类号: U226.8

文献标识码: A

doi: 10.3969/j.issn.1001-4632.2020.05.14

接触网作为高铁牵引供电系统的重要组成部分, 是高铁动车组动力的主要来源, 其服役可靠性是高铁正常运营的重要基础。随着动车组提速和开行数量的增加, 接触网运行环境变得更加复杂, 因关键零部件缺陷而导致的接触网故障时有发生^[1-2]。接触网悬挂状态检测监测装置安装在高速铁路巡检车车顶, 从多个角度拍摄沿线接触网悬挂状态图像, 通过对采集到的图像进行人工查阅, 以检测潜在的缺陷、故障。

为提高图像数据的分析效率、降低人工劳动强度, 现有针对高铁接触网零部件故障诊断的方法是先人工设计零部件特征, 再设计分类器实现单个零部件的定位, 最后采用灰度特征或边缘信息等分析潜在故障。文献[3]针对接触网支撑结构中旋转双耳耳片断裂故障, 利用旋转双耳耳片正负样本的HOG特征训练SVM分类器, 判别检测窗口内是否包含旋转耳片; 利用二维Gabor小波变换能量值对图像中的边缘信息进行筛选, 进而识别耳片断裂故障。文献[4]针对接触网承力索座辅助承力索缺

失的问题, 首先提取图像样本的DHOG特征训练AdaBoost分类器, 给出承力索座的精确定位, 然后通过可接受圆弧检测、圆弧段聚类拟合对图像中的圆形进行检测和统计, 最后根据统计数目给出故障判断。文献[5]针对接触网支撑装置旋转双耳部件销钉的松脱与脱落问题, 利用SIFT算法和改进的RANSAC算法实现双耳部位的定位, 采用Hough变换实现目标图像中双耳套筒倾角的提取及分割, 累加竖直方向像素灰度值, 从而判断销钉的工作状态。文献[6]针对接触网支撑装置中开口销缺失、松脱和张角不足的问题, 在定位阶段, 利用Hough直线检测实现斜腕处开口销的初定位, 利用PBoW模型对初定位图片进行分类, 采用SIFT算法准确定位非斜腕处开口销; 在检测阶段, 使用Hough累加矩阵提取圆形特征, 根据圆环内非连通区域分布规律给出开口销缺失和张角不足检测判据。上述方法虽然能够实现对接触网零部件不良状态的检测, 但是针对不同的零部件需要设计不同的特征提取方法, 存在精度提高难度大、运算效

收稿日期: 2019-11-08; 修订日期: 2020-04-03

基金项目: 国家重点研发计划项目(2018YFB1201403); 中国国家铁路集团有限公司科技研究开发计划重大课题(P2019G003)

第一作者: 周 雯(1986—), 女, 山西河曲人, 博士研究生。E-mail: zwjsj2008@163.com

率低的问题。

近年来,深度学习技术从图像数据出发,通过1个端到端的模型,联合优化特征表示和分类器性能,可最大程度发挥二者联合协助性能,最终得到分类或者检测结果。文献[7]根据高铁巡检车采集的轨腰图像中铁道塞钉图像特点,设计基于深度学习的铁道塞钉自动检测算法。文献[8]针对钢轨伤损检测的问题,将检测数据视为由16个通道二进制矩阵叠加而成的图像,设计基于深度学习的钢轨伤损智能识别模型。文献[9]针对钢轨扣件区域图像的尺寸较小、图像内容相对单一的特点设计基于深度卷积神经网络的扣件缺损识别算法。文献[10]针对无砟轨道扣件缺陷样本相对稀缺的问题,提出基于半监督深度学习的扣件缺陷图像识别方法。文献[11]为及时发现铁路桥梁高强螺栓偶发的延迟断裂,进行基于卷积神经网络的桥梁高强螺栓缺失图像识别方法研究。

本文通过目标检测和分类的2阶级联卷积神经网络,实现接触网悬挂紧固件缺陷的识别。首先,设计端到端自我强化的紧固件检测网络,用于对接触网悬挂支持装置和定位装置紧固件的检测;然后,通过搭建轻量级运行状态分类网络,实现接触网悬挂紧固件的缺陷识别。并进行试验验证。

1 高铁接触网悬挂状态监测图像

高铁接触网悬挂状态监测图像是由巡检车车顶的成像装置从全局和局部多个角度拍摄获得,5个相机分别负责拍摄腕臂连接件和承力索座区域、腕臂区域、水平绝缘子区域、斜绝缘子区域、支持和定位装置全局区域,巡检车及其成像装置如图1所示。

支持和定位装置用以支持接触悬挂和固定接触线的位置,保证接触线在受电弓滑板运行轨迹范围内,其结构如图2所示。支持装置的腕臂连接件包括套管双耳、上套管单耳、下套管单耳,定位装置连接件包括垂直定位环和吊钩定位环。



(a) 套管单耳



(b) 套管双耳



(c) 定位环

图3 高铁接触网紧固件物理结构和成像结构



(a) 巡检车



(b) 车顶成像装置

图1 巡检车及其成像装置

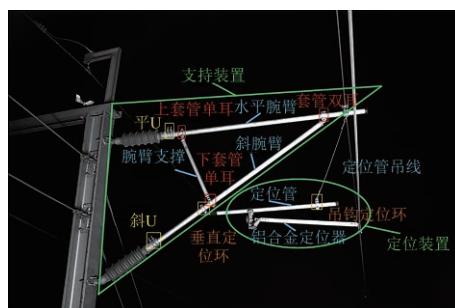


图2 高铁接触网支持和定位装置结构

套管单耳、套管双耳、定位环包含大量的紧固件,其物理结构和成像结构如图3所示。根据中国铁道标准,这些紧固件分为3类,分别是螺栓M12、螺母M12和螺栓销M12。紧固件在接触网装置中起固定作用,而受电弓与接触网装置之间存在复杂的力学作用,因长时间运行过程中的振动疲劳或施工缺陷,使紧固件产生脱落或松动等不良状态,需对其进行实时检测及安全隐患排除。

高铁接触网悬挂状态监测图像与普通图像相比具有如下特点。

(1) 高铁接触网悬挂状态监测图像尺寸为 $(6\,600 \times 4\,400)$ 像素,而公共数据集,如COCO数据集中的自然图像尺寸不超过 $(1\,024 \times 1\,024)$ 像素。相比而言,接触网悬挂状态监测图像为大尺寸图像。

(2) 相比于普通图像而言,高铁接触网悬挂状态监测图像存在大量的无语义信息区域,无语义信息的区域约占整幅图像的 $3/4$ 。

(3) 高铁接触网悬挂状态监测图像中支持和定位装置连接件的紧固件都较小,仅占整幅图像的 0.055% ,属于微小目标。而普通图像中目标占整幅图像比例为 10% 以上。

(4) 实际收集的数据集中,带有缺陷的紧固件样本数量很少,且正常和松脱2个状态的类间差异不大,很难直接训练1个高精度的紧固件缺陷检测

网络。

从上述特点来看,高铁接触网悬挂状态监测图像的独特性使得普通图像的检测方法无法直接被使用,且现有基于深度学习的目标检测方法对小目标的检测精度普遍较低。

2 高铁接触网悬挂紧固件缺陷识别方法

针对监测图像的特点,高铁接触网悬挂紧固件缺陷识别方法由紧固件检测和缺陷识别2个部分组成。首先从基于深度学习的目标检测算法主要组成部分(即特征提取网络、损失函数、检测器)出发,设计紧固件目标检测网络,对高铁接触网悬挂支持装置和定位装置中紧固件进行检测;然后,通过轻量级的运行状态分类网络,完成紧固件缺陷的识别及判定。流程如图4所示。

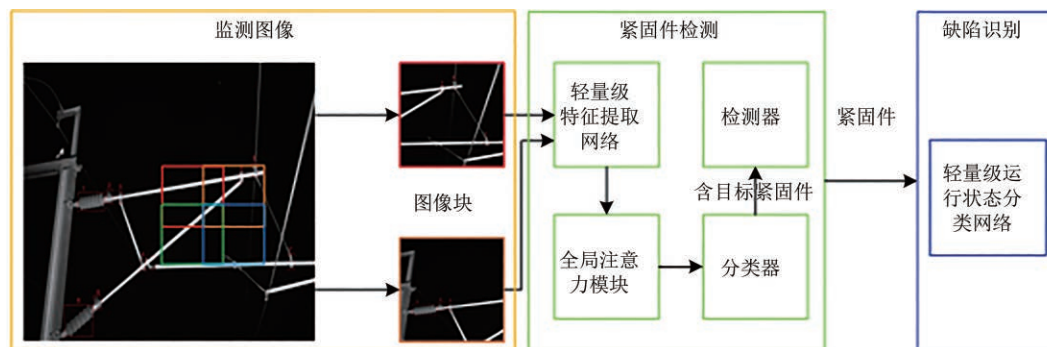


图4 高铁接触网悬挂紧固件缺陷识别方法流程

2.1 紧固件检测

紧固件检测由1个端到端自我增强的紧固件检测网络实现,该网络由轻量级特征提取网络、全局注意力模块及分类器和检测器组成,网络结构如图5所示。大尺寸的高铁接触网悬挂状态监测图像直接作为卷积神经网络的输入会增加计算时间和内存消耗,且图像中存在大量的无语义信息区域,会产生大量无意义的计算,因此,通过小批量处理图像块,可以减少整个网络的计算量,从而提高紧固件检测效率。首先,将 $(6\,600 \times 4\,400)$ 像素的接触网悬挂状态监测图像自动裁剪为 (512×512) 像素的图像块,且各图像块之间的重叠率为 35% ,以保证连接件的完整性;然后,将图像块输入到轻量级特征提取网络中,提取图像块的特征表示;最后,基于这些高辨识度的特征,分类器预测当前图像块中是否存在目标紧固件,若存在则通过检测器

定位目标紧固件。

2.1.1 轻量级特征提取网络

在处理工程领域实际问题时,基于卷积神经网络的目标检测方法通常将ResNet^[12]作为特征提取的主干网络。为了克服数据量少而产生过拟合的问题,主干网络首先在ImageNet数据集上训练得到网络的预训练权重,然后将去掉全连接层和softmax层的卷积网络作为目标检测框架中的特征提取网络。然而,这样的特征提取网络存在以下缺点:第一,参数量大导致检测效率低;第二,用于图像分类的特征分辨率可能不足以进行目标检测^[13]。另外,接触网悬挂监测图像和普通图像之间的差异,也使得预训练网络在实际应用中不够理想^[14],比如,微小的紧固件经过多层卷积后,容易造成特征消失。因此,轻量级特征提取网络调整ResNet-18各阶段特征图的数量为16,32,48和72个,使得

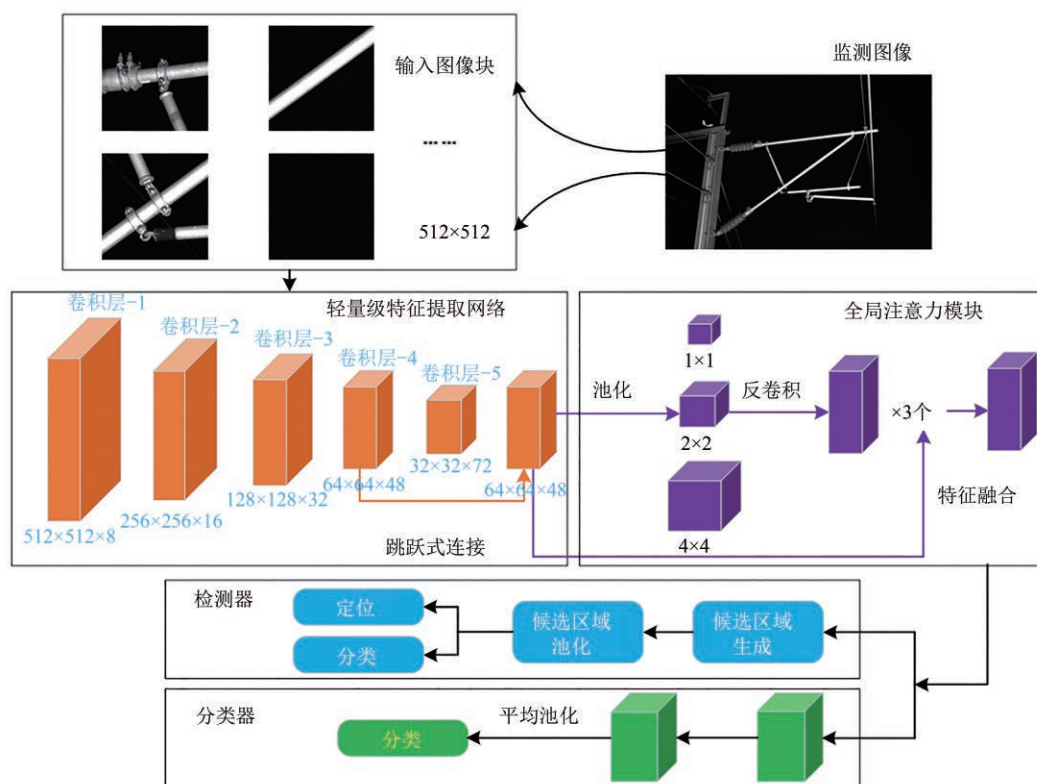


图5 紧固件检测网络结构（单位：像素）

网络参数远小于 ResNet-18，且可以从头开始训练，仅通过1个周期的训练可以得到很好的收敛。各阶段特征图的数量基于计算初始收敛网络中各阶段特征图协方差矩阵的特征谱^[15]，并在训练过程中微调得到。

首先，任意1个特征图的特征谱 Σ 为它的协方差矩阵经过空间分辨率归一化后得到，即

$$\Sigma = \frac{1}{n} \sum_{i=1}^n \frac{1}{W_i H_i} \sum_{x=1}^{W_i} \sum_{y=1}^{H_i} F_{i,x,y} F_{i,x,y}^T \quad (1)$$

式中： n 为图像块的数量； H_i 和 W_i 分别为特征图空间的宽和高； $F_{i,x,y}$ 为第 i 个图像块的特征图中空间坐标为 (x,y) 像素的特征向量。

然后，计算各阶段特征谱 Σ 的特征值，将大于预定阈值的特征值个数作为该阶段特征图的数量。

最后，在训练过程中微调得到最终的特征图数量。

另外，为了能提取到紧固件更精细的特征，减小轻量级特征提取网络总的下采样步长，防止紧固件特征在检测层中消失。卷积层-4先经过残差块得到卷积层-5，再经过上采样得到和卷积层-4尺度相同的特征图，通过跳跃式连接和卷积层-4的特征进行融合，即最终用于检测过程中锚框分布的特征图

的下采样步长为8像素。

2.1.2 全局注意力模块

由于卷积核的感受野有限，会存在一定的误报。如图6(a)和(b)所示，当人眼看到2个外观相似的疑似螺栓脱落缺陷时，如果没有周围图像的信息，无法确定它们是否为真实的脱落缺陷。正如文献[16]中所讨论，卷积神经网络有2种类型的感受野：理论感受野和实际感受野。理论感受野表示理论上影响输出值的输入区域。然而，并非理论感受野中的每个像素对输出值都有相同的影响，只有1个像素子集对输出值有有效的影响，该子集称为有效感受野，有效感受野小于理论感受野，如图6(c)所示。有限的有效感受野导致最终的特征图含有少量的上下文信息，从而导致误报。

根据上述理论，在轻量级特征提取网络后面连接1个特征池化层作为全局注意力模块。特征提取网络输出的特征图首先池化为 (1×1) 像素， (2×2) 像素， (4×4) 像素的金字塔层级特征图，然后使用双线性插值的方法恢复特征图大小，最后进行特征图融合，使得最终的特征图包含更多的上下文信息。全局注意力模块融合了不同金字塔层级的特征，使得检测器可以得到更多包含上下文信息的特征，使假阳性训练样本的置信度明显降低，从

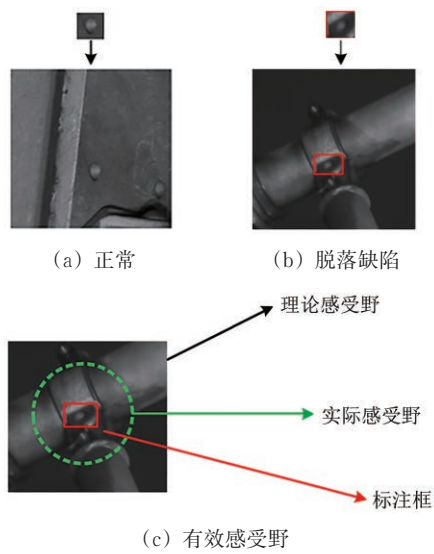


图6 卷积层感受野

而降低误报率。

2.1.3 相互增强的分类器和检测器

分类器和检测器共享轻量级特征提取网络的特征，并在训练过程中通过多任务损失函数实现相互增强。一方面，由于分类器将大量不含目标紧固件的图像块过滤掉，检测器的输入中减少了大量的假阳性样本，从而降低误报率。另一方面，由于检测器可以获取紧固件更精细的特征表示，当前图像块中即使存在一个目标紧固件，分类器也能很好地识别出来，即使当前图像块被分类器识别错误，检测器随后也可以更正这个错误的结果。

整个紧固件检测网络有3个输出：1个是分类器的输出，代表当前图像块中是否存在目标紧固件的概率 m ，另外2个是检测器的输出，分别代表图像块中每个候选区域在 $K+1$ 个类别上的概率分布 $p=(p_0, \dots, p_k)$ 和 K 个类别的边框回归偏移量 $t^k=(t_x^k, t_y^k, t_w^k, t_h^k)$ ，其中， k 为类别索引号， t^k 为相对于标注框中心点 (x, y) 的尺度不变转换及宽高 (w, h) 的空间偏移。参与训练的每个图像块依据是否含有目标紧固件标记为二值标签 b ，检测器中每个候选区域标记为与其匹配的标注框类别标签 u ，边框回归的目标为 v 。多任务损失函数 L 定义如下。

$$L(m, b, p, u, t^u, v) = L_{\text{cls}}(m, b) + \mu[b = 1][L_{\text{cls}}(p, u) + \lambda[u \geq 1]L_{\text{loc}}(t^u, v)] \quad (2)$$

式中： $L_{\text{cls}}(m, n)$ 和 $L_{\text{cls}}(p, u)$ 为softmax损失； L_{loc} 为smooth-L1损失； μ 和 λ 为用于平衡三个任务损失的参数，在试验中设置参数 $\lambda=1$ 和 $\mu=1$ ，训练

过程中，当相应的图像块中检测到目标紧固件时才反向传播检测器的损失。

检测器的结构如图5中蓝色部分所示，它是基于区域的目标检测，主要使用RPN^[16]网络的锚框机制，即使用一组预定义的多尺度边框合理分布于图像平面产生目标候选区域。然而，基于区域的检测器对于微小目标的检测性能显著下降。一方面，随着特征提取网络每个阶段池化层的叠加，下采样步长的增加，微小目标的特征被高度压缩甚至消失，如最小的紧固件在最终的锚框分布特征图中仅为 (1×1) 像素。另一方面，在锚框分布特征图中，如果锚框 i 和标注框 j 的IoU值大于所有锚框和标注框 j 的IoU值，或者该IoU值大于阈值 T_h ，锚框 i 就和标注框 j 匹配。简言之，当特征图中目标紧固件的尺寸和设置的锚框尺寸越接近，将会有越多的锚框与它匹配，目标紧固件就更容易被定位。人脸检测器S³FD^[17]详细分析了锚框数量与目标尺寸之间的匹配关系，如果目标在锚框分布的特征图中的尺寸在平均线(average line)之上，将会匹配到3~4个锚框，足以用于目标的定位；反之，微小目标将匹配1个以下的锚框，从而导致检测的性能大幅下降。

因此，为了在目标紧固件检测过程中，更合理的分配锚框，根据训练集中标注框尺寸用K-means聚类自动选择合适的锚框尺寸。选取聚类个数分别为3, 4, 5的锚框尺寸如下：3个锚框尺寸分别为3.55, 5.34, 7.62像素；4个锚框尺寸分别为2.95, 3.90, 6.13, 8.21像素；5个锚框尺寸分别为1.97, 3.25, 4.9, 7.06, 9.21像素。

分类器的结构如图5中绿色部分所示。分类器仅包含2个卷积层和1个全局平均池化层。采用softmax损失函数，用来引导分类器的训练。

2.2 紧固件缺陷识别

通过紧固件检测网络，实现了紧固件的定位。紧固件缺陷识别通过搭建1个紧固件运行状态分类网络，将紧固件检测网络输出的紧固件图像缩放到 (72×72) 像素作为输入，实现紧固件缺陷的识别。紧固件的运行状态分为正常、松脱和脱落3种类别，如图7所示。

紧固件运行状态识别网络结构如图8所示，包括4个卷积层，2个池化层，1个全连接层和1个输出层，网络的输出是3类紧固件的运行状态，共8个类别。

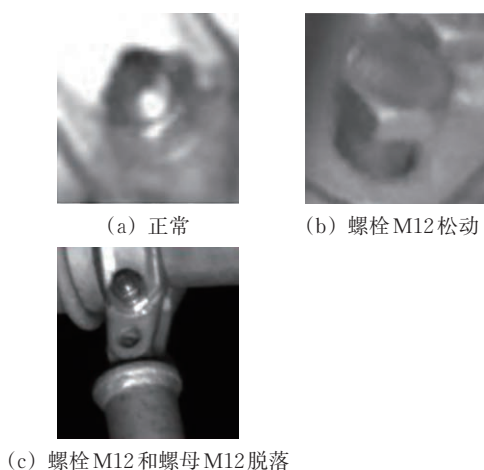


图7 高铁接触网悬挂紧固件运行状态类别

3 试验验证

高铁接触网悬挂紧固件检测网络和运行状态识别网络的训练和测试均采用深度学习框架 TensorFlow 实现, 硬件环境为 GeForce GTX 1080 GPU 服务器, 64 位 Ubuntu16.04 操作系统。

试验中选取某高铁线路接触网监测图像数据,

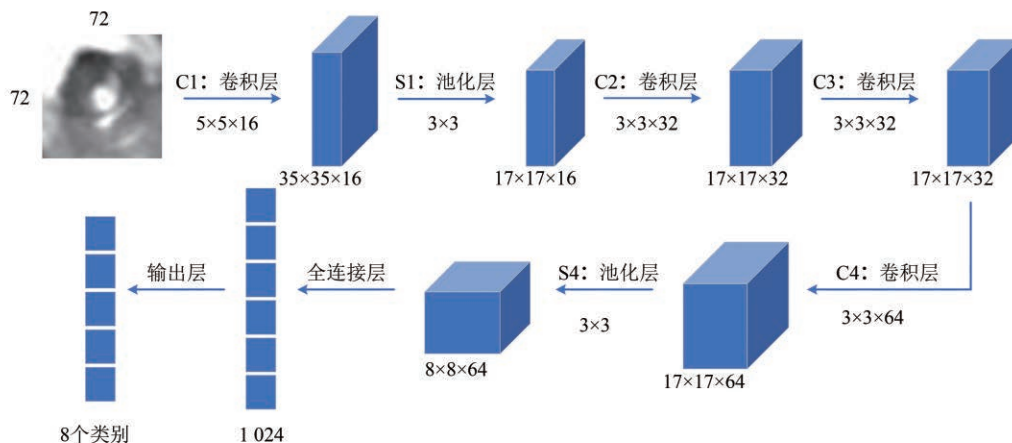


图8 高铁接触网悬挂紧固件运行状态识别网络结构 (单位: 像素)

3.1 网络训练阶段

紧固件检测网络的训练采用随机梯度下降的方法在4块GPU上同时进行, 每个GPU分配1张图像, 每张图像选取512个候选区域用来训练检测器。随机梯度下降的动量为0.9, 权重衰减为0.0005。初始学习率设置为0.001, 每60K个小批量, 学习率减少1/10。由标准差为0.01, 均值为0的高斯分布为所有层随机分配初始权重。因为网络是从头开始训练, 学习率和训练周期迭代2次作为1个循环。且在检测器的RPN网络中, 锚框的大小

包括隧道、岔道、高架桥等不同环境下的图像作为数据集, 以保证模型的泛化能力。数据集中约2100个接触网悬挂支柱图像, 其中约42000个紧固件。所有的图像自动裁剪为 (512×512) 像素的图像块, 图像块间有35%的重叠, 该重叠不仅保证了连接件的完整性, 同时也增强了数据集。将图像块数据集随机选取80%作为训练集, 剩余的20%作为验证集, 对紧固件检测网络进行训练。为了让网络得到更好的收敛, 将包含目标紧固件的图像块和不包含目标紧固件的图像块比例控制在1:3, 从而得到平衡的训练集。包含目标紧固件的数据集标注为3个类别: 螺栓M12、螺母M12和螺栓销M12。

紧固件运行状态识别网络的训练数据集由紧固件检测网络检测到的紧固件组成, 每类紧固件每种状态分别选取50个样本, 并手工标注其标签。由于缺陷紧固件图像的数量有限, 且目前缺陷数据只收集到一些关于螺栓松脱的数据, 缺少开口销缺失的缺陷数据。训练过程中进行了数据增量, 包括水平翻转、垂直翻转。

在步长为8像素的特征图上使用K-means聚类个数为4的锚框尺寸, 分别为3, 4, 6和8像素, 锚框的长宽比例设置为1。

紧固件运行状态识别网络的损失函数采用多标签 softmax 损失, 同时在全连接层加入50%的 dropout 以减少过拟合。

3.2 网络测试阶段

在铁路目标检测领域, 常用误检率和漏检率两个指标衡量模型的检测效果。通常对二分类问题而言, 可将样本根据真实类别和预测类别划分为真阳

性 (True Positive, TP)、假阳性 (False Positive, FP)、真阴性 (True Negative, TN)、假阴性 (False Negative, FN), 误报率是指假阳性样本在被识别为目标实例中的占比, 即 $1 - P_{\text{recision}}$; 漏检率是指假阴性样本在所有真实值中的比例, 即 $1 - R_{\text{ecall}}$ 。其中 R_{ecall} 和 P_{recision} 分别定义如下。

$$R_{\text{ecall}} = \frac{T_p}{T_p + F_N} \quad (3)$$

$$P_{\text{recision}} = \frac{T_p}{T_p + F_p} \quad (4)$$

式中: T_p , F_N , F_p 分别为 TP, FN 和 FP 的值。

在检测目标时, 每个类别都能够根据 R_{ecall} 和 P_{recision} 绘制 1 条 PR 曲线, 对此曲线做积分, 得到的就是 Average Precision (AP), 即 1 个类别的检测精度, 对多个类别的 AP 求取平均值, 得到检测器的平均检测精度 (Mean AP, mAP)。

测试集包含 100 张接触网悬挂支柱图像, 其中, 每个悬挂支柱包含 4 个螺栓 M12, 12 个螺母 M12, 4 个螺栓销 M12。在紧固件检测网络前向传

播时, 为保证漏检率尽可能小, 也就是召回率尽可能高, 选取置信度为 0.75, 最终得到紧固件检测平均准确率为 97.2%, 召回率为 98.2%, 误报率为 2.8%, 漏检率为 1.8%, 从铁路目标检测领域注重漏检率来看, 表明了紧固件检测网络在实际应用中的潜能。然后, 检测到的紧固件作为输入测试紧固件运行状态识别网络, 得到 95.8% 的平均识别准确率, 识别速率为每秒 430 张图像, 表明了紧固件运行状态识别网络的高效性, 适用于实际工程领域。

高铁接触网悬挂紧固件缺陷识别的检测结果如图 9 所示 (其中绿色框内为螺母, 红色框内为螺栓销, 蓝色框内为螺栓), 识别结果见表 1。在紧固件检测识别阶段, 会自动报告如图 9 (c) 的紧固件脱落故障, 由图 3 物理结构可知, 每个连接件的紧固件数量是固定的, 如果某个紧固件处于脱落状态, 那么将无法检测到该紧固件, 检测时据此可以通过检测紧固件的实际数量与预定数量间的不同, 自动报告紧固件发生脱落故障。

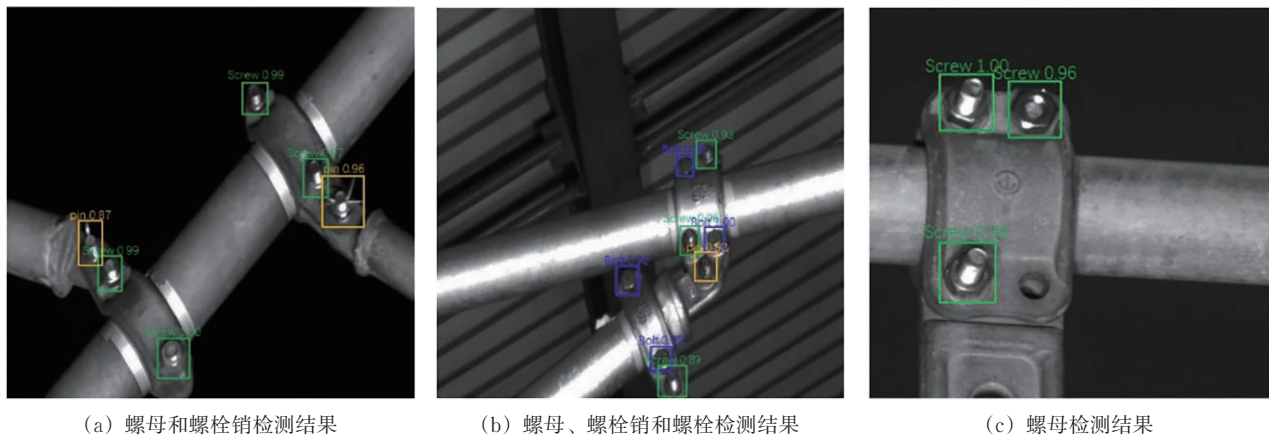


图9 紧固件检测结果

3.2.1 紧固件检测网络性能评估

在测试集上对紧固件检测网络进行综合评估, 分别从紧固件检测的高效性、稳健性、有效性3个方向分析。为了定量评估各组成策略对网络整体性能的影响, 设计表2所示的多组剥离试验, 其中, 分类器、检测器和全局注意力模块分别为图5绿色、蓝色和紫色部分; 端到端联合训练表示分类器和检测器通过损失函数联合训练。Model I 表示训练整个网络, Model II 表示只训练分类器部分, Model III 表示只训练检测器部分, Model IV 表示联合训练除全局注意力模块的分类器和检测器。

为了评估端到端联合训练的分类器在整个紧固

件检测网络中的作用及性能, 将测试集中 100 张接触网悬挂支柱图像自动裁剪为 (512×512) 像素图像块, 其中, 大约有 554 个包含目标紧固件的图像块。分类器评估依据为是否将含有目标紧固件的图像块分类正确, 试验结果见表 3。表中: ResNet-50 是在 ImageNet 预训练后利用紧固件图像数据微调的结果。由表 3 可以看出: 端到端训练的 Model I 分类器性能最好, 其召回率高于 ResNet-50 和 Model II; 只训练分类器部分的 Model II 分类器, 其准确率和召回率较 Model I 分类器均大幅降低, 表明检测器和分类器相互增强机制的重要性, 检测器提取到更精细的特征能抑制分类器的假

表1 紧固件运行状态识别结果

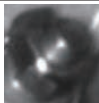

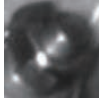
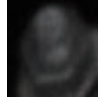
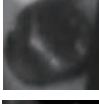
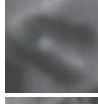

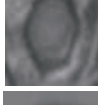

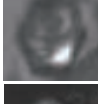
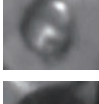



紧固件图像	识别结果	准确率	紧固件图像	识别结果	准确率
	螺母 M12 正常	0.993		螺栓销 M12 正常	0.907
	螺母 M12 正常	0.985		螺母 M12 正常	0.962
	螺栓 M12 松脱	0.926		螺母 M12 脱落	0.958
	螺栓销 M12 正常	0.973		螺栓 M12 正常	0.996
	螺栓 M12 松脱	0.941		螺栓 M12 正常	0.954
	螺母 M12 脱落	0.915		螺母 M12 正常	0.871
	螺栓 M12 松脱	0.896		螺栓销 M12 正常	0.979

表2 紧固件检测网络性能评估剥离试验组

试验组	分类器	检测器	全局注意力模块	端到端联合训练
Model I	✓	✓	✓	✓
Model II	✓	×	✓	×
Model III	×	✓	✓	×
Model IV	✓	✓	×	✓

表3 紧固件检测网络分类器性能对比结果

网络	TP	FP	准确率/ %	召回率/ %	时间消耗/ ms
ResNet-50	487	107	81.99	87.91	27.86
Model II 分类器	383	575	39.94	69.13	8.62
Model I 分类器	538	137	79.62	97.11	8.62

阳性样本；另外，时间消耗表明，轻量级的特征提取网络使得分类器的推理时间为8.62 ms，比复杂的ResNet-50分类器效率高。因此，端到端联合训练的分类器能有效过滤掉大量无目标紧固件的图像块，减少检测器的计算量，同时，检测器提取到更精细的特征能促使分类器更精确。

为了评估紧固件检测网络对微小紧固件检测的有效性，在同样的测试集上测试不同的检测器，其中，Faster R-CNN^[18]的基础网络是ResNet-50，考虑到在下采样步长为8像素的特征图上分布锚

框，ResNet-50需要消耗大量的存储空间，因此，Faster R-CNN在下采样步长为16像素的特征图上分布锚框；同时，也验证了紧固件检测网络在下采样步长为16像素的特征图上分布锚框的检测器。另外，尝试将全局注意力模块连接到卷积层-4 (Conv-4) 特征提取阶段进行对比，试验结果见表4。从表4可以看出：紧固件检测网络检测器的检测精度和召回率均优于其他网络；但当下采样步长为16像素时，检测精度仅为0.836 6，表明下采样步长为16像素时会造成一些目标紧固件的特征消失，从而无法检测到；Conv-4的检测精度相较于紧固件检测网络下降0.020 9，表明卷积层-5 (Conv-5) 阶段可以提取到有利于目标检测的高层语义信息；当选择5个聚类锚框尺寸时，Anchor-5的精度和召回率分别提升0.002 3和0.002 0，但时间消耗增加3.77 ms，该结果表明和数据集紧固件大小匹配精细的锚框尺寸有利于提高检测精度，但锚框数量会影响检测的效率，而Anchor-4能在检测的精度与效率间取得较好的折中。因此，检测网络的最终锚框尺寸为Anchor-4，即表4中紧固件检测网络的性能。

表4 紧固件检测网络检测器性能对比结果

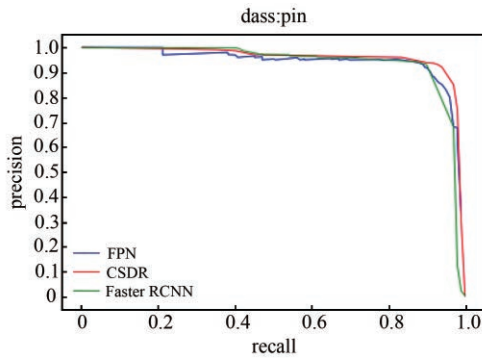
模型	mAP	最大召回率	时间消耗/ms
Faster R-CNN	0.945 6	0.961 4	30.86
步长16(stride-16)	0.836 6	0.853 1	13.57
卷积层-4(Conv-4)	0.951 4	0.962 7	21.45
3个锚框(Anchor-3)	0.948 9	0.956 4	21.97
5个锚框(Anchor-5)	0.974 6	0.983 5	28.62
紧固件检测网络(Anchor-4)	0.972 3	0.981 5	24.85

为了验证紧固件检测网络的有效性，与小目标检测表现较好的FPN^[19]进行对比试验，FPN的基础网络是ResNet-50，试验结果见表5。从表5可以看出：虽然FPN和紧固件检测网络平均精度相近，但紧固件检测网络的召回率比FPN高0.71%，即紧固件检测网络的漏检率低于FPN。

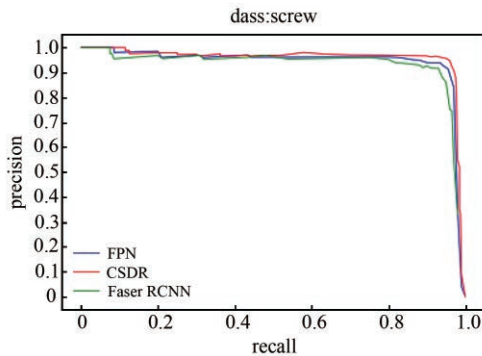
表5 紧固件检测网络与FPN性能比对

网络	不同紧固件的AP			mAP	最大召回率
	螺母 M12	螺栓 M12	螺栓销 M12		
FPN+ResNet-50	0.985 8	0.970 7	0.961 5	0.972 7	0.974 4
紧固件检测网络	0.982 5	0.970 4	0.964 2	0.972 3	0.981 5

紧固件检测网络(CSDR)的PR曲线截图如图10所示。由图10中的红色曲线可知，CSDR的



(a) 螺栓销 M12



(b) 螺母 M12

图 10 检测结果 PR 曲线截图

召回率和检测精度相比 FPN 和 Faster RCNN 都有不错的表现。

为了更好地表明紧固件检测网络 (CSDR) 显著优于对比网络, 进一步对表 4 和表 5 的数据进行了显著性检验, 检验结果见表 6。从表 6 可以看出: 紧固件检测网络显著优于对比的其他网络; 其中, 紧固件检测网络和 FPN 的 mAP 对比不是很明显, 这也和上文中得出的结论相一致, 但是紧固件检测网络的召回率显著优于 FPN 的召回率。

3.2.2 紧固件运行状态识别网络性能评估

为了评估紧固件运行状态识别网络的性能, 与 AlexNet^[20] 分类网络进行了对比试验。2 个网络的对比试验在相同的数据集上完成, 试验结果见表 7。从表 7 可以看出: 紧固件运行状态识别的轻量级分类网络和 AlexNet 网络的识别准确率十分相似, 但从每秒钟处理的图像数量 (Frames Per Second, FPS) 和训练时间来看, 紧固件运行状态识别网络的推理时间和训练时间显著优于 AlexNet, 更适用于实时的紧固件运行状态识别。

另外, 为了评估紧固件状态识别的准确率, 分析紧固件运行状态识别网络的混淆矩阵, 结果见表 8。表中: 蓝色数字为紧固件运行状态识别网络的精度, 蓝色数字上方为紧固件运行状态识别网络的误报率, 蓝色数字下方为紧固件运行状态识别网络

表 6 试验结果显著性检验

被对比网络 I	对比网络 J	平均值差值	标注误差	显著性	95% 置信区间	
					下限	上限
stride-16	Conv-4	-0.114 80*	0.010 48	0	-0.138 1	-0.091 5
	FasterRCNN	-0.109 00*	0.010 48	0	-0.132 3	-0.085 7
	FPN	-0.136 07*	0.010 48	0	-0.159 4	-0.112 7
	CSDR	-0.138 00*	0.010 48	0	-0.161 3	-0.114 7
Conv-4	strid-16	0.114 80*	0.010 48	0	0.091 5	0.138 1
	FasterRCNN	0.005 80	0.010 48	0.592	-0.017 5	0.029 1
	FPN	-0.021 27	0.010 48	0.070	-0.044 6	0.002 1
	CSDR	-0.023 20	0.010 48	0.051	-0.046 5	0.000 1
Faster R-CNN	strid-16	0.109 00*	0.010 48	0	0.085 7	0.132 3
	Conv-4	-0.005 80	0.010 48	0.592	-0.029 1	0.017 5
	FPN	-0.027 07*	0.010 48	0.027	-0.050 4	-0.003 7
	CSDR	-0.029 00*	0.010 48	0.020	-0.052 3	-0.005 7
FPN	strid-16	0.136 07*	0.010 48	0	0.112 7	0.159 4
	Conv-4	0.021 27	0.010 48	0.070	-0.002 1	0.044 6
	FasterRCNN	0.027 07*	0.010 48	0.027	0.003 7	0.050 4
	CSDR	-0.001 93	0.010 48	0.857	-0.025 3	0.021 4
CSDR	strid-16	0.138 00*	0.01048	0	0.114 7	0.161 3
	Conv-4	0.023 20	0.010 48	0.051	-0.000 1	0.046 5
	FasterRCNN	0.029 00*	0.010 48	0.020	0.005 7	0.052 3
	FPN	0.001 93	0.010 48	0.857	-0.021 4	0.025 3

注: *表示平均值差值的显著水平为 0.05。

的漏检率。从表8可以看出：紧固件运行状态识别网络对于脱落状态的召回率较高，由于松脱和正常状态有可能会很相似，而存在很小的误报率。

表7 紧固件状态识别网络和 AlexNet 网络性能比对

网络	mAP	FPS/ (个·s ⁻¹)	训练时间/ min
AlexNet	0.958 8	85	101
紧固件运行状态识别网络	0.958 1	430	69

表8 3种紧固件分类的混淆矩阵

真实值	预测值								
	螺母 M12			螺栓 M12			螺栓销 M12		
	正常	脱落	松脱	正常	脱落	松脱	正常	脱落	
正常	96.73	0.12	3.15	98.27	0	1.73	94.57	5.43	
脱落	0.74	98.33	0.93	0	100.00	0	2.94	97.06	
松脱	4.26	0.22	95.52	2.39	0	97.61	—	—	

3.2.3 2 阶级联结构性能评估

为了评估2阶级联轻量级卷积神经网络结构的重要性，将紧固件检测和运行状态识别合并到1个阶段，即直接使用紧固件检测网络检测紧固件的缺陷状态，并与2阶级联结构进行对比试验。试验中，选取每类紧固件的每种状态各50个样本，共450个样本组成训练集，并直接标记出训练集中3种紧固件的运行状态及其位置，用于2次训练紧固件检测网络，设置学习率为0.000 5，结果如图11所示。从图11可以看出：由于训练样本有限，用单一的紧固件检测网络直接检测紧固件缺陷的精度

很低且会导致严重的过拟合，特别是松脱缺陷，由于松脱状态和正常状态的差异较小，检测器中的特征不足以将它们正确分类；而2阶级联结构的平均精度较单一检测网络提升21.5%。

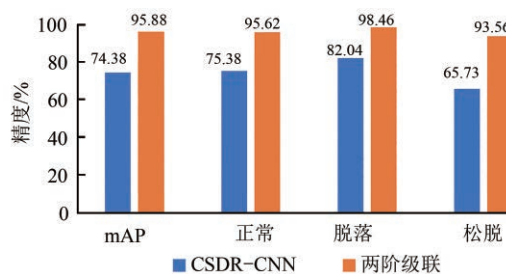


图11 2 阶级联网络与单一检测网络对比

4 结 语

采用基于卷积神经网络的深度学习检测和分类技术，设计2阶级联的卷积神经网络模型实现高铁接触网支持装置紧固件缺陷的识别。紧固件检测网络可以定位接触网悬挂状态监测图像中的目标紧固件，轻量级多标签分类网络能够判断3类紧固件的松动和脱落缺陷。通过对某高铁线路接触网状态监测图像数据进行试验分析，验证提出的2阶级联的高铁接触网支持装置紧固件缺陷识别方法的有效性、准确性和高效性，可以辅助工作人员提高图像数据的分析效率、降低人工劳动强度。

参 考 文 献

- [1] LIU Xiyang, PENG Jinfang, TAN Deqiang, et al. Failure Analysis and Optimization of Integral Droppers Used in High Speed Railway Catenary System [J]. Engineering Failure Analysis, 2018, 91: 496-506.
- [2] 谭德强, 莫继良, 彭金芳, 等. 高速接触网零部件失效问题研究现状及展望[J]. 西南交通大学学报, 2018, 53(3): 610-619.
(TAN Deqiang, MO Jiliang, PENG Jinfang, et al. Research and Prospect on High-Speed Catenary Component Failure [J]. Journal of Southwest Jiaotong University, 2018, 53 (3): 610-619. in Chinese)
- [3] 韩烨, 刘志刚, 耿肖, 等. 基于HOG特征与二维Gabor小波变换的高铁接触网支撑装置耳片断裂故障检测[J]. 铁道学报, 2017, 39(2): 55-60.
(HAN Ye, LIU Zhigang, GENG Xiao, et al. Fracture Detection of Ear Pieces in Catenary Support Devices of High-Speed Railway Based on HOG Features and Two-Dimensional Gabor Transform [J]. Journal of the China Railway Society, 2017, 39 (2): 55-60. in Chinese)
- [4] 张桂南, 刘志刚, 韩烨, 等. 高速铁路接触网承力索座识别与辅助承力索缺失故障检测方法研究[J]. 铁道学报, 2017, 39(5): 40-46.
(ZHANG Guinan, LIU Zhigang, HAN Ye, et al. Loss Fault Detection for Auxiliary Catenary Wire of High-Speed Railway Catenary Wire Holder [J]. Journal of the China Railway Society, 2017, 39 (5): 40-46. in Chinese)

- [5] 钟俊平, 刘志刚, 张桂南, 等. 高铁接触网旋转双耳销钉状态检测方法研究[J]. 铁道学报, 2017, 39(6): 65-71.
(ZHONG Junping, LIU Zhigang, ZHANG Guinan, et al. Condition Detection of Swivel Clevis Pins in Overhead Contact System of High-Speed Railway [J]. Journal of the China Railway Society, 2017, 39 (6): 65-71. in Chinese)
- [6] 钟俊平, 刘志刚, 陈隽文, 等. 高速铁路接触网悬挂装置开口销不良状态检测方法研究[J]. 铁道学报, 2018, 40(6): 51-59.
(ZHONG Junping, LIU Zhigang, CHEN Junwen, et al. Defective Condition Detection of Split Pins in Catenary Suspension Device of High-Speed Railway [J]. Journal of the China Railway Society, 2018, 40 (6): 51-59. in Chinese)
- [7] 杜馨瑜, 戴鹏, 李颖, 等. 基于深度学习的铁道塞钉自动检测算法[J]. 中国铁道科学, 2017, 38(3): 89-96.
(DU Xinyu, DAI Peng, LI Ying, et al. Automatic Detection Algorithm for Rail Plug Based on Deep Learning [J]. China Railway Science, 2017, 38 (3): 89-96. in Chinese)
- [8] 孙次锁, 刘军, 秦勇, 等. 基于深度学习的钢轨伤损智能识别方法[J]. 中国铁道科学, 2018, 39(5): 51-57.
(SUN Cisuo, LIU Jun, QIN Yong, et al. Intelligent Detection Method for Rail Flaw Based on Deep Learning [J]. China Railway Science, 2018, 39, (5): 51-57. in Chinese)
- [9] 刘俊博, 黄雅平, 王胜春, 等. 基于机器视觉的多线路钢轨扣件缺损检测方法[J]. 中国铁道科学, 2019, 40(4): 27-35.
(LIU Junbo, HUANG Yaping, WANG Shengchun, et al. Rail Fastener Defect Detection Method for Multi Railways Based on Machine Vision [J]. China Railway Science, 2019, 40 (4): 27-35. in Chinese)
- [10] 戴鹏, 王胜春, 杜馨瑜, 等. 基于半监督深度学习的无砟轨道扣件缺陷图像识别方法[J]. 中国铁道科学, 2018, 39(4): 43-49.
(DAI Peng, WANG Shengchun, DU Xinyu, et al. Image Recognition Method for the Fastener Defect of Ballastless Track Based on Semi-Supervised Deep Learning [J]. China Railway Science, 2018, 39 (4): 43-49. in Chinese)
- [11] 赵欣欣, 钱胜胜, 刘晓光. 基于卷积神经网络的铁路桥梁高强螺栓缺失图像识别方法[J]. 中国铁道科学, 2018, 39(4): 56-62.
(ZHAO Xinxin, QIAN Shengsheng, LIU Xiaoguang. Image Identification Method for High-Strength Bolt Missing on Railway Bridge Based on Convolution Neural Network [J]. China Railway Science, 2018, 39 (4): 56-62. in Chinese)
- [12] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770-778.
- [13] BHARAT Singh, LARRY S. Davis. An Analysis of Scale Invariance in Object Detection-SNIP [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 3578-3587.
- [14] ZAMIR A R, SAX A, SHEN W, et al. Taskonomy: Disentangling Task Transfer Learning [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 3712-3722.
- [15] YOSUKE Shinya, EDGAR Simo-Serra, TAIJI Suzuki. Understanding the Effects of Pre-Training for Object Detectors via Eigenspectrum [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2019.
- [16] LUO W, LI Y, URTASUN R, et al. Understanding the Effective Receptive Field in Deep Convolutional Neural Networks [C]// Advances in Neural Information Processing Systems. Barcelona: IEEE, 2016: 4898-4906.
- [17] ZHANG S, ZHU X, LEI Z, et al. S³fd: Single Shot Scale-invariant Face Detector [C]// Proceedings of the IEEE International Conference on Computer Vision. New York: IEEE, 2017: 192-201.
- [18] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real Time Object Detection with Region Proposal Networks [C]// International Conference on Neural Information Processing Systems. New York: IEEE, 2015: 91-99.
- [19] LIN T, DOLLAR P, GIRSHICK R, et al. Feature Pyramid Networks for Object Detection [C] // IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2017: 2117-2125.
- [20] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet Classification with Deep Convolutional Neural Networks [C]// Advances in Neural Information Processing Systems. Lake Tahoe: IEEE, 2012: 1097-1105.

Defect Recognition Method for Fasteners of Catenary Suspension for High Speed Railway Based on Two-Cascade Lightweight Convolutional Neural Network

ZHOU Wen¹, SHI Tianyun², LI Ping³, MA Xiaoning³, WU Yanhua³

(1. Postgraduate Department, China Academy of Railway Sciences, Beijing 100081, China;

2. China Academy of Railway Sciences Corporation Limited, Beijing 100081, China;

3. Institute of Computing Technology, China Academy of Railway Sciences Corporation Limited, Beijing 100081, China)

Abstract: Based on the operation state monitoring image of high speed railway catenary suspension, the difference between the monitoring image and the ordinary image was analyzed. The problem of fastener defect detection was transferred into two processes of fastener detection and operation state fine recognition, and a fastener defect detection method based on two-cascade convolutional neural network was proposed. Firstly, a fastener detection network was designed, which was composed of a lightweight feature extraction network, a global attention module, a mutually enhanced classifier and a detector, to realize the high efficiency detection of target fastener instances. Then, a multi-label classification network including 4 convolutional layers, 2 pooling layers, 1 fully connected layer and 1 output layer was built to complete the fine classification of fastener operation state and realize defect recognition. The image data of the catenary condition monitoring of a high speed railway was tested and analyzed by using fastener defect detection method. Results show that the method of two-cascade convolutional neural network is fast and accurate in detecting fastener defects. Specifically, the average detection rate of fastener location is 98.2%, and the average recognition accuracy of fastener defects is 95.8%, which is about 21.5% higher than that of single detection network.

Key words: High-speed railway; Two-cascade; Convolutional neural network; Catenary; Defect recognition; Fastener; Object detection

(责任编辑 金 燕)