

基于改进 U-Net 卷积神经网络的钢轨表面损伤检测方法

梁波, 卢军*, 曹阳

陕西科技大学机电工程学院, 陕西 西安 710021

摘要 基于卷积神经网络的深度学习方法对钢轨表面损伤的自动化检测起到非常重要的推动作用, 因此提出一种基于卷积神经网络的钢轨表面损伤检测方法。首先, 在经典 U-Net 的收缩路径和扩展路径之间增加一个分支网络, 可以辅助 U-Net 输出理想的分割图。然后, 将 Type-I RSDDs 高速铁路轨道数据集作为检测样本, 使用数据增强的手段扩增检测样本后馈入改进的 U-Net 中进行训练和测试。最后, 采用评价指标对所提方法进行评估。实验结果表明, 所提方法的检测准确率达到 99.76%, 相比于其他方法的最高水平提高 6.74 个百分点, 说明所提方法可以显著提高检测准确率。

关键词 机器视觉; 深度学习; 损伤特征识别; 数据增强; 改进的 U-Net 图形分割网络; 无损检测

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202158.0215009

Rail Surface Damage Detection Method Based on Improved U-Net Convolutional Neural Network

Liang Bo, Lu Jun*, Cao Yang

College of Mechanical & Electrical Engineering, Shaanxi University of Science and Technology, Xi'an, Shaanxi 710021, China

Abstract The deep learning method based on convolutional neural network plays a very important role in promoting the automatic detection of rail surface damage. Therefore, a method based on convolutional neural network for rail surface damage detection is proposed. First, a branch network is added between the contraction path and extension path of the classic U-Net can assist U-Net to output the ideal segmentation graph. Then, the type-I RSDDs high-speed railway track dataset is taken as the test sample, and the test sample is amplified by means of data enhancement and fed into the improved U-Net for training and testing. Finally, the evaluation index is used to evaluate the proposed method. The experimental results show that the detection accuracy of the proposed method reaches 99.76%, which is 6.74 percentage higher than the highest level of other methods, indicating that the proposed method can significantly improve the detection accuracy.

Key words machine vision; deep learning; damage feature recognition; data enhancement; improved U-Net image segmentation network; non-destructive detection

OCIS codes 150.1135; 100.2000; 100.3008

收稿日期: 2020-06-05; 修回日期: 2020-06-26; 录用日期: 2020-08-03

基金项目: 陕西省科技厅自然科学基金(2016GY-049)

* E-mail: lujun@sust.edu.cn

1 引言

随着全球高速铁路的蓬勃发展,其在运输领域中起着举足轻重的作用,因此铁路运输安全直接关系到社会和经济的发展等^[1]。铁路轨道表面损伤是最常见的且最重要的铁路故障之一,又是列车高速运行的潜在危险之一。铁路日常维护中,对铁路轨道表面进行检查是一项必要的工作。通常检查方式主要以机械式检测为主,这不仅造成了严重的人力资源浪费,而且还存在耗时和准确性低等问题,因此轨道表面损伤的自动化检测被提上了日程。

近年来,研究人员针对表面缺陷的自动化检测提出了很多方法,其中基于数据驱动的深度学习方法得到了广泛的关注。文献[2]针对工业铝材缺陷的检测过程中由缺陷样本稀疏带来的训练过拟合和泛化性能差等问题,提出一种基于多任务深度学习的铝材缺陷检测方法,该方法可以很好地解决多任务训练过程中收敛不均衡的问题,并且减少推断时间,提高检测效率。基于机器视觉的缺陷目标检测方案因其具有快速适应新产品的优势,受到了研究人员的广泛关注。文献[3]设计了一种基于机器视觉的聚氯乙烯管材表面缺陷检测算法,并将其应用于工业生产领域。文献[4]针对人工检测编织袋缺陷的正确率低与效率较低的问题,提出一种基于机器视觉的高效在线检测编织袋缺陷的方法。当解决表面质量控制的问题时^[5],基于深度学习的目标检测方法可以获得更为优异的效果。马克杯表面损伤的检测任务中,基于改进 Faster R-CNN (Region-Convolutional Neural Networks) 的深度学习方法可以获得很好的检测效果^[6]。尽管经典的目标检测方法已经获得了相对不错的效果,但目前仍然存在检测图像背景复杂、图像噪声干扰以及检测方法的泛化性和准确率较低等问题。

为了解决上述存在的问题,科研人员已提出了大量的图像处理技术,其中文献[7]应用了局部频率特征来保证检测局部或微小裂纹的能力。此外,基于频谱特征的离散小波变换(DWT)也被应用在轨道表面损伤检测领域中^[8]。上述目标检测方法依赖于提取的全局或局部特征必须在均匀的纹理轨道表面图像上具有良好的表现,但是在实际工作环境中会产生大量的随机噪声,导致无法检测到具有复杂外观和动态背景的缺陷^[9]。针对上述目标检测方法存在的不足,处理轨道表面图像的另一方法是在图像增强后自动定位缺陷。

文献[10]提出了局部归一化加上基于投影轮廓的缺陷定位(LN+DLBP)方法,该方法可以实时地检测离散缺陷。文献[11]提出了局部类迈克耳孙对比度加上比例强调最大熵(MLC+PEME)的方法,该方法可以检测图像中可能存在的离散缺陷。文献[12]提出了一种新的 Perona-Malik(P-M)逆扩散模型,该模型可以用于图像增强,随后应用自适应阈值二值化可以找出图像中存在的缺陷。文献[13]使用了基于均值漂移算法的粗-细提取器从观察到的序列中沿纵向查找背景,可以找到包含真实缺陷和其他噪声点的异常值,采用集成纵向上下文信息和横向先验信息的粗-细提取器(CFE)方法可以过滤这些噪声点。尽管上述方法可以改善检测性能,但是还会不同程度地受到不规则因素的影响,如噪声点等。

随着深度学习技术的不断发展,文献[14]提出了一种基于改进的 U-Net 的卷积神经网络,该网络最初是用于处理医学图像,其本质是由一个精细的图形分割卷积神经网络构成,架构是完全基于全卷积网络而建立的^[15]。鉴于该网络在医学图像处理领域中表现出来的优异性能,因此本文将改进的 U-Net 图形分割网络用于钢轨表面损伤的检测。经过实验验证,证明所提方法可以很好地克服图像噪声的干扰,同时将所提方法与 LN+DLBP^[10]、MLC+PEME^[11]、DWT^[8]、CFE^[13]和 U-Net^[14]方法进行比较,证明所提方法在钢轨损伤检测的准确率上有明显提高。

2 方法的整体结构

所提方法的总体检测流程如图 1 所示,其中包括图像预处理阶段、图像损伤区域分割提取阶段和图像后处理阶段。

2.1 图像预处理

为了缩短研究周期,选择公开的 Type-I RSDDs 铁路轨道数据集^[13]对所提方法进行有效性验证。Type-I RSDDs 数据集中的样本数量较少,网络训练过程中容易出现过拟合的现象,这会造成分割误差偏大。对于这一问题,采用数据增强的手段对样本进行扩增以降低过拟合发生的概率,从而提升网络输出模型的稳定性。数据增强主要依据图像数据的平移、旋转和缩放不变性原理对原图像数据进行扩充,数据增强后的图像样本如图 2 所示。此外在 U-Net 收缩路径末端的 dropout 层进一步扩充隐式数据,这可以进一步减少过拟合发生的概率。

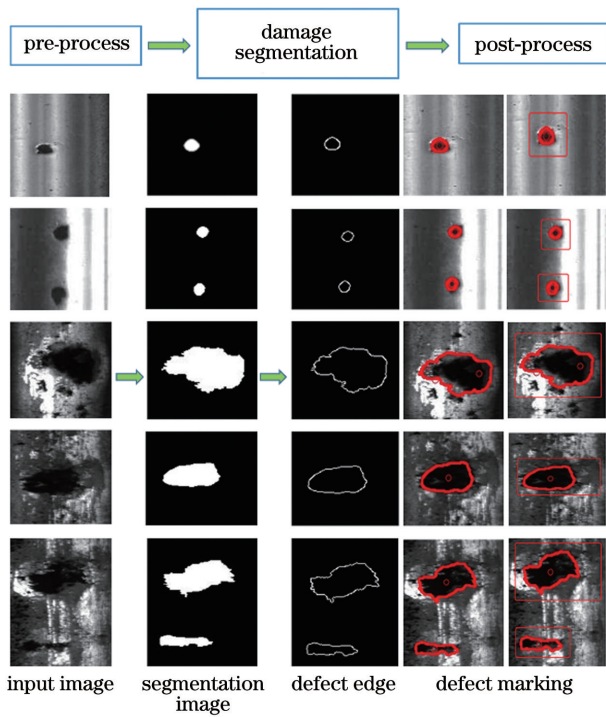


图 1 轨道损伤图像的检测流程图

Fig. 1 Detection flow chart of track damage image

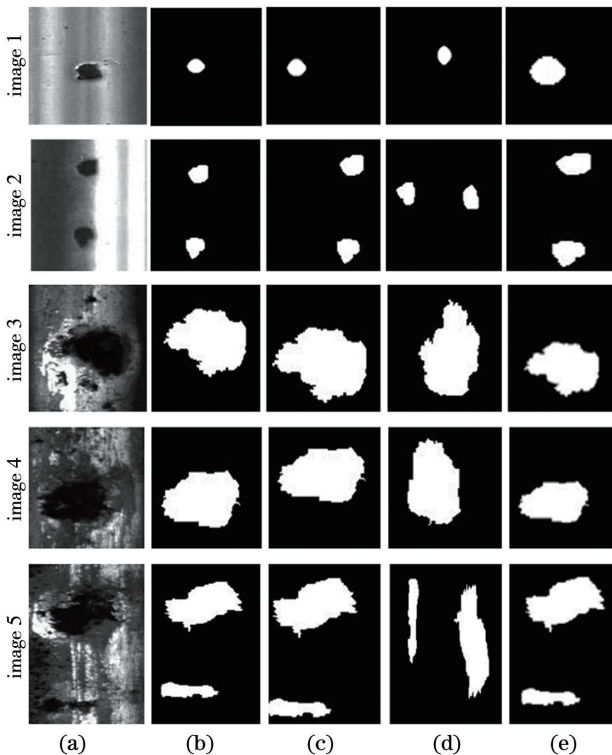


图 2 不同操作扩展后的样本图像。(a)原始图像;(b)注释图像;(c)平移变换;(d)旋转变换;(e)缩放变换

Fig. 2 Sample images expanded by different operations.

(a) Original images; (b) ground truth; (c) translation transformation; (d) rotation transformation; (e) scaling transformation

2.2 损伤区域的分割机制

将改进的 U-Net 卷积神经网络方法作为表面损伤检测的目标分割方法,其运算机制是分割网络在接受到外部馈送的待检测样本后,由于网络中的扩展路径(上采样部分)含有大量的功能通道,这些功能通道允许网络将上下文信息传播到更高分辨率的层。实验在扩展路径之前添加一个分支单元,命名为 GRASP 网络,该分支网络将收缩路径(下采样部分)的输出结果作为输入,将分支网络处理后的预测结果作为输出并将其与扩展路径最终的输出结果结合,旨在优化目标分割的准确率。此外,网络的主干结构中没有全连接层,而且仅使用每个卷积层中的有效部分,所以可以从输入图像中获得完整的上下文信息。

2.3 损伤区域的提取及后处理

后处理阶段是基于改进的 U-Net 计算的分割图在原始的输入图像上,使用经典的图像处理技术对区域进行匹配标记,这可以使工程人员更好地辨识钢轨表面损伤的检测结果,便于工程人员快速且适当地评估检测对象。

3 损伤区域的分割方法

3.1 改进的 U-Net 图形分割网络结构

主干网络是经典的 U-Net 卷积神经网络,其由一个收缩路径和一个扩展路径构成。收缩路径遵循卷积网络的典型架构,其包括两个 3×3 的卷积,每个卷积层后跟一个 ReLU 函数激活层和一个步长为 2、池化核为 2×2 的最大池化层,该路径可以用于对图像进行下采样操作。收缩路径后添加一个 GRASP 分支网络,该分支网络由两个全连接层和一个输出层构成,将收缩路径的输出结果作为输入,最后在输出层利用边界框机制对损伤区域进行预测并作为输出。GRASP 分支网络利用边界框机制强制更多的下采样层去学习整个网络中更高级别的图像语义信息,其检索的信息包括表面损伤的位置和尺寸,这对辅助经典的 U-Net 输出理想的显著性分割图至关重要。不同于收缩路径,扩展路径中的每个步骤都是对损伤特征图进行上采样操作,经过 2×2 卷积操作后对特征通道的数量进行减半,以便与从收缩路径中裁剪出的相应特征图进行级联,此外再加上两个 3×3 卷积层,且每个卷积层后再跟一个 ReLU 函数激活层。由于卷积的过程中会有边界像素丢失的现象,所以必须对每个卷积单元进行裁剪。在最后一层使用 1×1 卷积将每个 64 通道的

特征向量映射到所需的类上,最终再将主干网络与分支网络的结果结合后输出,故该网络总共有 26 个

卷积层。改进的 U-Net 卷积神经网络的结构如图 3 所示。

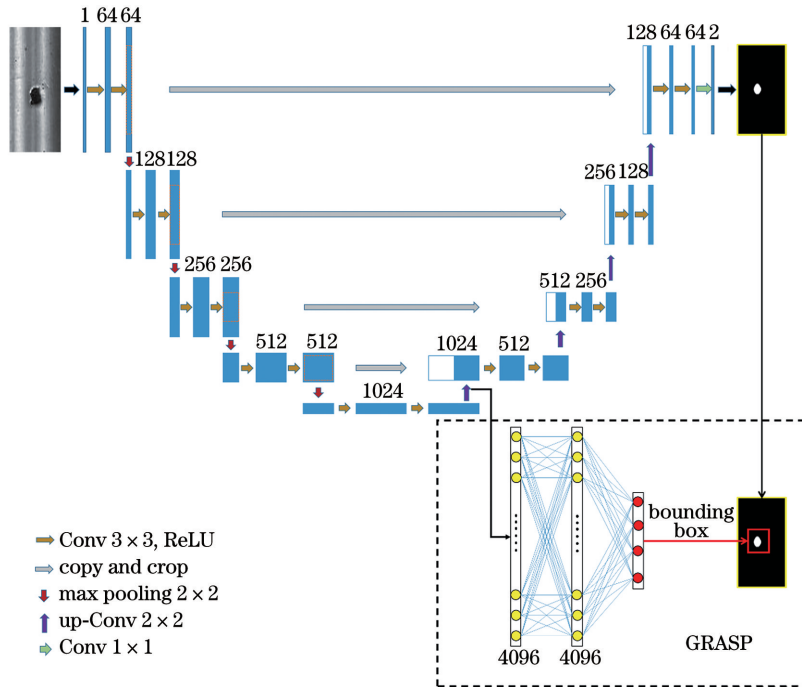


图 3 改进的 U-Net 卷积神经网络的结构

Fig. 3 Structure of improved U-Net convolution neural network

3.2 分割网络模型的算法原理

分割检测模型在训练过程中因采用未填充的卷积进行操作,所以输出图像比输入图像小一个恒定的边框宽度。为了最小化训练开销以及最大化 GPU 内存的利用率,大批量的训练过程中选择高分辨率的输入图像,通过连续卷积的操作来批量减小输入图像的分辨率,此外,为了避免模型在训练过程中陷入局部最优的陷阱,故选择高动量以越过局部最优点。

模型主干网络使用的损失函数由最终特征图上的像素级 Softmax 函数与交叉熵损失函数组合而成。Softmax 函数可表示为

$$p_k(X) = \frac{\exp[a_k(X)]}{\sum_{k=1}^K \exp[a_k(X)]}, \quad (1)$$

式中: X 表示像素点的位置, $X \in \Omega$ 且 $\Omega \subset \mathbf{Z}^2$, 其中 Ω 表示像素点位置所在的区域,即输入的整张待检测图像; $a_k(X)$ 表示第 k 类 X 点的最大激活值; K 表示类的数量; $p_k(X)$ 表示最大似然函数。对于具有最大值的 k , $p_k(X) \approx 1$, 对于其他值的 k , $p_k(X) \approx 0$ 。交叉熵损失函数是 $p_1(X)$ 与 1 的偏差,可表示为

$$E = \sum_{X \in \Omega} \omega \lg[p_1(X)], \quad (2)$$

式中: $l: \Omega \rightarrow \{1, \dots, K\}$ 表示每个像素的真实标签; $\omega: \Omega \rightarrow \mathbf{R}$ 表示引入的权重因子,其可以使一些像素在训练过程中更具有重要性。

钢轨表面损伤区域与周围正常组织的像素级是不连续的,为此可以使用形态学运算来计算分割边界。分割权重的表达式为

$$\omega = \omega_c + \omega_0 \cdot \exp\left[-\frac{(d_1 + d_2)^2}{2\sigma^2}\right], \quad (3)$$

式中: $\omega_c: \Omega \rightarrow \mathbf{R}$ 表示权重因子,用于平衡类频率; $d_1: \Omega \rightarrow \mathbf{R}$ 表示损伤区域与最近边界的距离; $d_2: \Omega \rightarrow \mathbf{R}$ 表示损伤区域与次近边界的距离; σ 和 ω_0 表示经验参数。

GRASP 分支网络中,将欧几里得算法^[16]作为损失函数。给定每个缺陷的标记中心坐标 (x_E, y_E) 、宽度 w_E 和高度 h_E ,则损失函数可表示为

$$E_{\text{GRASP}} = (x_E - x_{\text{Pred}})^2 + (y_E - y_{\text{Pred}})^2 + (w_E - w_{\text{Pred}})^2 + (h_E - h_{\text{Pred}})^2, \quad (4)$$

式中: $(x_{\text{Pred}}, y_{\text{Pred}})$ 、 w_{Pred} 和 h_{Pred} 分别表示预测边界框的中心坐标、宽度和高度。

在具有多卷积层和多路径的深度神经网络中,网络权重参数的初始化非常重要,否则网络中的某些部分可能被重复激活,其他部分因处于欠激活的状态而不能发挥作用。理想情况下,应调整初始权

重使网络中的每个特征图都具有近似的单位方差。对于 U-Net 架构,采用标准偏差为 $\sqrt{2/N}$ 的高斯分布实现参数的初始化,其中 N 表示一个神经元的传入节点数。例如,对于一个 3×3 卷积和 64 个特征通道来讲, $N=9 \times 64=576$ 。

4 实验结果及分析

4.1 实验设置

为了评估所提方法的有效性,在 Ubuntu 16.04 操作系统上利用 PyTorch 来搭建网络并进行验证。实验过程中,将数据集中的数据馈送至损伤提取器并在单片 GPU (NVIDIA GeForce GTX1080 Ti-11GiB) 上进行目标提取的训练。选择的数据集包含所有常见的轨道损伤缺陷,其中包含 67 张原图及其注释图像,数据预处理后整个数据集的样本达到 268 张,并将其分为训练集、验证集和测试集,图像数量分别为 187 张、27 张和 54 张。

4.2 实验评估指标

为了验证所提方法,采用 4 种评估指标对其进行评估,指标分别为准确率(P_{Acc})、召回率(P_{Rec})、损失函数曲线和 ROC (Receiver Operating Characteristic) 曲线。准确率是指正确辨识的样本数在所有样本数中的比值,准确率越高,说明分类器越好。召回率是指对整个样本的覆盖程度,该指标可以度量有多少个正样本被正确辨识,又可以衡量分类器对正样本的辨识能力。轨道缺陷的检测过程中,一个正样本通常意味着存在一个危险源,故选择召回率对所提方法进行评价是很有必要的。损失函数曲线可以证明所提方法在轨道损伤缺陷的学习过程中是收敛的。实际的轨道损伤检测过程中,正、负样本的分布变化较大,经常出现类不平衡的现象,即负样本比正样本更多(或者相反),此时可以通过 ROC 曲线对所提方法的可靠性进行评估,ROC 曲线下的 AUC (Area Under Curve) 越大,则说明模型越可靠。除了上述评估指标外,将所提方法与 LN+DLBP^[10]、MLC+PEME^[11]、DWT^[8]、CFE^[13] 和 U-Net^[14] 方法在同一数据集上的检测结果进行比较。其中 P_{Acc} 和 P_{Rec} 的表达式为

$$P_{Acc} = \frac{x_{TP} + x_{TN}}{x_{TP} + x_{TN} + x_{FP} + x_{FN}}, \quad (5)$$

$$P_{Rec} = \frac{x_{TP}}{x_{TP} + x_{FN}}, \quad (6)$$

式中: x_{TP} 表示正确划分为正样本的个数; x_{FP} 表示错误划分为正样本的个数; x_{TN} 表示正确划分为负

样本的个数; x_{FN} 表示错误划分为负样本的个数。

4.3 实验执行细节

实验中使用 Adam 优化器进行训练,其中设置一阶矩估计和二阶矩估计的指数衰减率分别为 0.9 和 0.999。损失函数的学习率设为 0.0002。考虑到图像尺寸和 GPU 内存的限制,每次迭代仅使用单幅图像,即批量大小设为 1。此外将(3)式的相关参数设置为 $\omega_0=10$ 和 $\sigma=5$ 。

训练过程中,网络会随机地选择训练样本,为此修改这种选择机制以确保馈送到网络的正、负样本数量一致,其是通过每个偶数次迭代馈送有缺陷的图像以及每个奇数次迭代馈送无缺陷的图像来实现的,这种机制可确保系统以恒定的速率来学习缺陷图像,反之存在更多的非缺陷样本,使学习过程变得不平衡,进而导致学习速率显著降低。网络经过 100 轮的训练后输出目标提取器模型,训练过程中当且仅当一次遍历完整个训练集的所有样本后,才认为一轮结束。

4.4 实验结果分析

采用所提方法对训练集和验证集中的数据进行学习,并在测试集上进行测试,将得到的结果与其他方法进行比较。

4.4.1 方法性能表现

学习过程中输出的训练数据如图 4 所示。从图 4 可以看到,所提方法的 P_{Acc} 值基本在 0.99 左右, P_{Rec} 值在 0.98 左右,整个学习过程中的损失函数曲线趋于稳定收敛,说明所提方法能够实现铁路钢轨损伤的检测,且检测方法准确稳定。

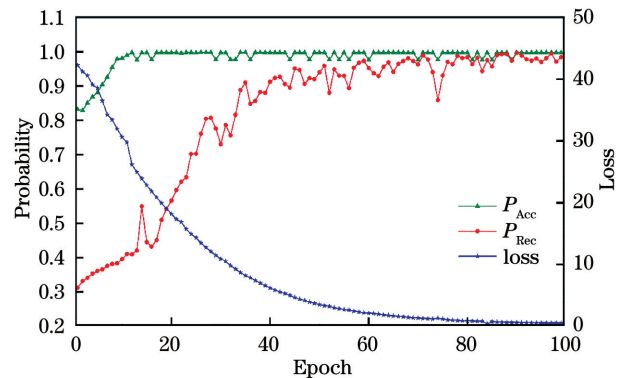


图 4 所提方法在模型训练过程中的性能曲线

Fig. 4 Performance curves of proposed method in model training process

所提方法在训练过程中的 ROC 曲线,如图 5 所示。从图 5 可以看到,所提方法具有良好的可靠性。

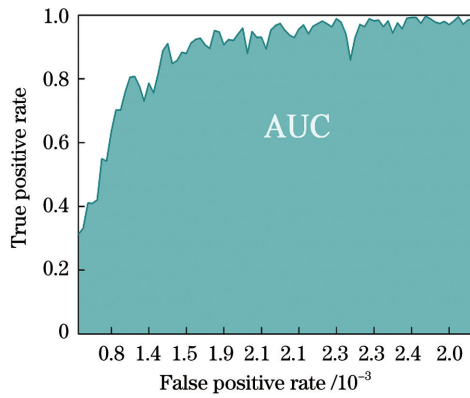


图 5 所提方法的 ROC 曲线

Fig. 5 ROC curve of proposed method

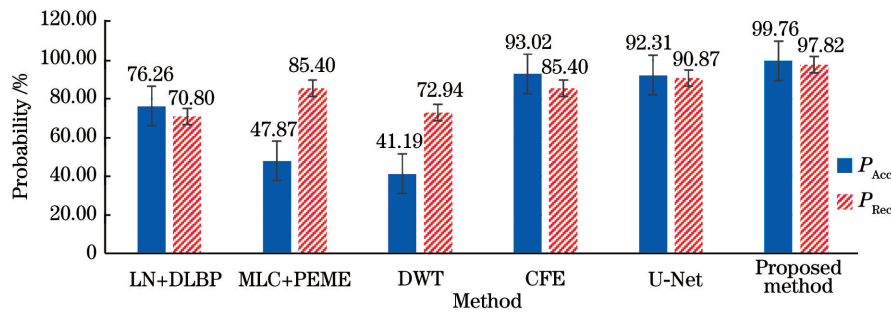


图 6 不同方法的检测结果

Fig. 6 Detection results of different methods

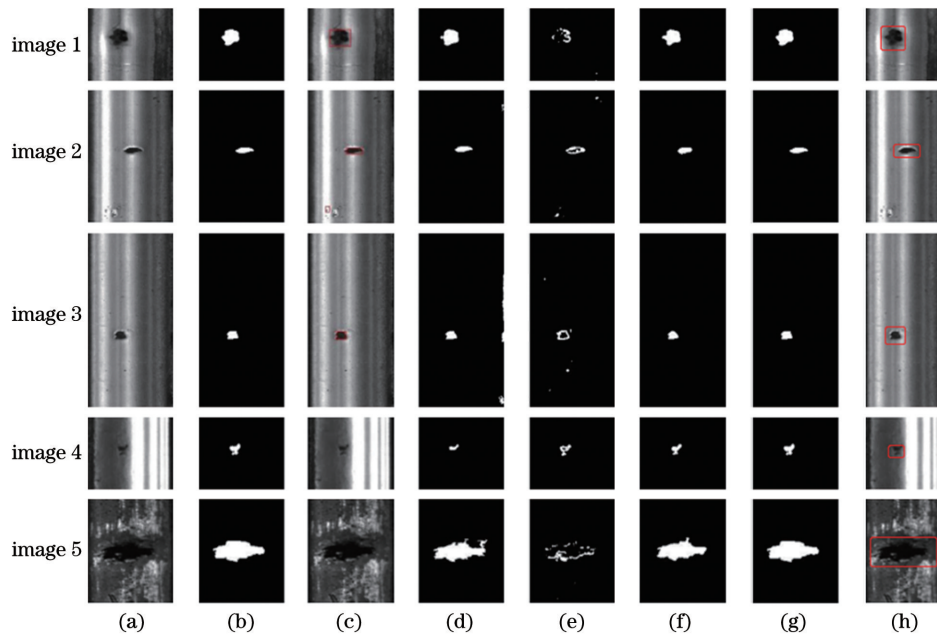


图 7 不同方法的可视化结果。(a) 损伤图像; (b) 注释图像; (c) LN+DLBP; (d) MLC+PEME; (e) DWT; (f) CFE; (g) U-Net; (h) 所提方法

Fig. 7 Visualization results of different methods. (a) Defect images; (b) ground truth; (c) LN+DLBP; (d) MLC+PEME; (e) DWT; (f) CFE; (g) U-Net; (h) proposed method

4.4.2 对比实验

将所提方法在测试集上的测试结果与其他方法进行对比,结果如图 6 所示。从图 6 可以看到,所提方法的 P_{Acc} 值达到 99.76%,召回率 P_{Rec} 值达到 97.82%,较其他方法中的最高水平分别提升 6.74 个百分点和 6.95 个百分点,说明所提方法确实是当前钢轨损伤检测方法体系的更新与扩展。

考虑方法表达的直观性,将可视化检测结果与其他方法进行比较,结果如图 7 所示。从图 7 可以看到,所提方法可以通过最终的检测结果直观地评估检测对象的损伤级别。

5 结 论

本课题组提出基于改进的 U-Net 卷积神经网络

的铁路轨道表面损伤自动化检测方法。将 Type-I RSSDs^[13]数据集作为检测样本,可以使网络模型学习轨道损伤特征并用于实际检测。对准确率等指

标进行评估,证明所提方法的检测准确率可以达到 99.76%,较其他方法的最好成绩提升 6.74 个百分点,这对于实际轨道损伤自动化检测的研究意义重大。因实验存在局限性,在实际的检测工况下未对所提方法进行实验,则这将是下一步工作的重点。

随着机器视觉技术的发展,对于铁路钢轨损伤自动化检测的研究主要集中在表面可见损伤的检测,然而对于钢轨内部的不可见损伤因成像方式复杂和信号处理难度大等因素而鲜有研究。将钢轨表面损伤检测的方法应用在真实环境中,还必须考虑检测方法的实时性和稳定性。针对以上钢轨损伤自动化检测存在的问题,未来将还有很多研究工作亟待开展。

参 考 文 献

- [1] Resendiz E, Hart J M, Ahuja N. Automated visual inspection of railroad tracks[J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14(2): 751-760.
- [2] Shen X H, Li Z H, Li M, et al. Aluminum surface-defect detection based on multi-task deep learning[J]. Laser & Optoelectronics Progress, 2020, 57(10): 101501.
沈晓海, 栗泽昊, 李敏, 等. 基于多任务深度学习的铝材表面缺陷检测[J]. 激光与光电子学进展, 2020, 57(10): 101501.
- [3] Li S H, Zhou Y T, Wang D, et al. Surface defect detection of polyvinyl chloride pipes based on machine vision[J]. Laser & Optoelectronics Progress, 2019, 56(13): 131006.
李书华, 周亚同, 王丹, 等. 基于机器视觉的聚氯乙烯管材表面缺陷检测[J]. 激光与光电子学进展, 2019, 56(13): 131006.
- [4] Chi H. On-line defect detection method of woven bag based on machine vision[J]. Laser & Optoelectronics Progress, 2020, 57(20): 201507.
迟欢. 基于机器视觉的编织袋缺陷在线检测方法[J]. 激光与光电子学进展, 2020, 57(20): 201507.
- [5] Chen P H, Ho S S. Is overfeat useful for image-based surface defect classification tasks? [C] // 2016 IEEE International Conference on Image Processing (ICIP), September 25-28, 2016, Phoenix, AZ, USA. New York: IEEE, 2016: 749-753.
- [6] Li D J, Li R H. Mug defect detection method based on improved Faster RCNN[J]. Laser & Optoelectronics Progress, 2020, 57(4): 041515.
李东洁, 李若昊. 基于改进 Faster RCNN 的马克杯缺陷检测方法[J]. 激光与光电子学进展, 2020, 57(4): 041515.
- [7] Li Q Y, Tan Y Q, Zhang H Y, et al. A visual inspection system for rail corrugation based on local frequency features[C] // 2016 IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech), August 8-12, 2016, Auckland, New Zealand. New York: IEEE, 2016: 18-23.
- [8] Caprioli A, Cigada A, Raveglia D. Rail inspection in track maintenance: a benchmark between the wavelet approach and the more conventional Fourier analysis [J]. Mechanical Systems and Signal Processing, 2007, 21(2): 631-652.
- [9] Trinh H, Haas N, Li Y, et al. Enhanced rail component detection and consolidation for rail track inspection [C] // 2012 IEEE Workshop on the Applications of Computer Vision (WACV), January 9-11, 2012, Breckenridge, CO, USA. New York: IEEE, 2012: 289-295.
- [10] Li Q Y, Ren S W. A real-time visual inspection system for discrete surface defects of rail heads[J]. IEEE Transactions on Instrumentation and Measurement, 2012, 61(8): 2189-2199.
- [11] Li Q Y, Ren S W. A visual detection system for rail surface defects[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2012, 42(6): 1531-1542.
- [12] He Z D, Wang Y N, Yin F, et al. Surface defect detection for high-speed rails using an inverse P-M diffusion model[J]. Sensor Review, 2016, 36(1): 86-97.
- [13] Gan J R, Li Q Y, Wang J Z, et al. A hierarchical extractor-based visual rail surface inspection system [J]. IEEE Sensors Journal, 2017, 17(23): 7935-7944.
- [14] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation [M] // Navab N, Hornegger J, Wells W, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [15] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 3431-3440.
- [16] Huang Y, Qiu C, Yuan K. Surface defect saliency of magnetic tile[J]. The Visual Computer, 2020, 36(1): 85-96.