

# 基于注意力反向知识蒸馏的车轮踏面异常检测

秦荣荣, 高晓蓉\*, 罗林, 李金龙

西南交通大学物理科学与技术学院, 四川 成都 610000

**摘要** 车轮是铁路列车走行部的重要部件, 车轮踏面上产生的缺陷严重危害着铁路列车的安全运行。由于实际中车轮踏面缺陷样本有限, 有监督检测模型对缺陷的检测不具有鲁棒性。针对此问题, 提出使用无监督的知识蒸馏异常检测模型实现对车轮踏面的异常检测任务。首先, 使用 UNet 对踏面区域进行分割, 减少非踏面区域对异常检测模型的影响; 然后, 在多尺度特征聚合之后添加一个注意力机制, 提升反向知识蒸馏结构中孪生网络对正常特征的重建能力, 增强孪生网络对正常特征重建的效果。实验结果表明: 在铁路车轮踏面数据集上, 改进后的模型能够达到 93.8% 的受试者工作特性曲线下的面积、82.3% 的精准率、95.4% 的召回率、87.0% 的准确率。与原模型相比, 改进后的模型检测性能得到提升。

**关键词** 车轮踏面; 无监督; 知识蒸馏; 异常检测; UNet

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP230787

## Wheel Tread Anomaly Detection Based on Attentional Reverse Knowledge Distillation

Qin Rongrong, Gao Xiaorong\*, Luo Lin, Li Jinlong

School of Physical Science and Technology, Southwest Jiaotong University, Chengdu 610000, Sichuan, China

**Abstract** Wheels are an essential part of railway trains; thus, defects on the wheel tread present serious risk regarding the safety of railway trains. Due to the limited samples of wheel tread defects in practice, the corresponding supervised detection model is insufficient. To solve this problem, an unsupervised knowledge distillation anomaly detection model is proposed to detect wheel tread anomalies. Accordingly, UNet is employed to segment the tread region and reduce the influence of non-tread regions on the anomaly detection model. An attention mechanism is then added after the multiscale feature fusion to improve the ability of the student network to reconstruct normal features in the reverse knowledge distillation structure, as well as enhance the reconstruction of normal features. From the experimental results, the improved model achieves the performance indexes of 93.8% area under receiver operating characteristic curve, 82.3% precision, 95.4% recall, and 87.0% accuracy considering the railway wheel tread dataset. Compared with the original model, the detection performance of the model is improved.

**Key words** wheel tread; non-supervision; knowledge distillation; anomaly detection; UNet

## 1 引言

铁路列车车轮在行驶过程中受到惯性作用、轨边应力和热损伤等外界因素的影响, 车轮踏面会产生诸如擦伤、剥离、硌伤等缺陷, 这些缺陷严重危害列车的行车安全, 所以对车轮踏面状况的检测工作至关重要。激光<sup>[1]</sup>、超声和涡流<sup>[2]</sup>检测等依赖人工的方法, 易受人和环境因素的影响, 且检测效率低下、成本较高。随

着计算机领域新技术的发展, 文献[3-10]利用图像处理和视觉检测技术对铁路列车车轮进行检测, 检测精度和速度有所提高, 但这些传统的图像视觉检测技术均是通过提取低层次、人工选定的特征点或边缘进行预测, 所以在检测形态特征多变的检测目标时普适性不强。卷积神经网络在图像处理领域得到了广泛的应用<sup>[11-13]</sup>, 文献[14-17]将深度卷积神经网络应用于铁路列车车轮踏面缺陷检测, 极大地提升了检测精度, 缩

收稿日期: 2023-03-07; 修回日期: 2023-03-28; 录用日期: 2023-05-06; 网络首发日期: 2023-05-16

基金项目: 自然科学基金重点国际(地区)合作与交流项目(61960206010)

通信作者: \*gxrr@vip.163.com

短了检测时间。但是目前大多基于深度学习的图像检测技术是数据驱动型的有监督算法,其需要大量的具有代表性的样本来训练,这些样本需要人工进行标注,耗时费力,而最终模型也只能检测训练中存在过的情况。为了解有监督算法的弊端,克服实际列车行驶过程无法提供大量缺陷样例的难题以及免去人工标注步骤,提出将无监督异常检测算法应用于车轮踏面的检测中。

无监督异常检测算法一直是计算机视觉领域的主流研究方向之一,在产品质量控制<sup>[18]</sup>、医疗诊断<sup>[19]</sup>等领域具有广泛的应用,是视觉检测里的一项重要任务。基于重建的异常检测方法如GANomaly<sup>[20]</sup>,使用生成器将所有图像重建成正常图像,利用对应图像的差异进行判断。然而,此类模型具有强大的泛化能力,导致异常图像也能够被重建,从而降低了检测能力。特征空间可以很好地应用于异常检测,比基于重建的深度学习方法要好<sup>[21]</sup>。基于特征空间的异常检测方法, Multiresolution knowledge distillation (MKD)<sup>[22]</sup>考虑了多尺度特征对齐,以及使用均方误差损失与方向损失的问题。文献<sup>[23]</sup>考虑了传统结构的知识蒸馏存在因相同架构及相同数据流引发异常激活失败的问题,由此提出了反向蒸馏新范式,克服了异常激活差异消失的弊端,也改善了知识蒸馏模型在未知样本表示的

多样性。铁路车轮踏面上的缺陷形态不一、没有明显的分布规律,且复杂的外界检测环境以及不同机位的相机给成像带来很大的影响,若直接使用无监督的异常检测网络进行检测,达不到理想的检测效果。因此,本文提出先分割再检测的思路,为实际工业异常检测避免不相关背景的干扰提供了一种可能。然后以反向知识蒸馏模型为基础,在多尺度特征聚合模块后添加一个注意力机制,降低多尺度融合特征因卷积压缩而模糊纹理特征信息的效果,增强纹理特征的表现。实验结果表明,使用无监督的异常检测改进方法检测效果更优。

## 2 基本原理

### 2.1 基于注意力反向知识蒸馏模型结构

所提的基于注意力反向知识蒸馏模型如图 1 所示,反向知识蒸馏模型又称为反向蒸馏(RD)模型,所提模型是对 RD 模型的改进。该模型使用教师网络来对图像进行特征提取,然后与有对应镜像结构的学生网络进行知识的传递,即特征信息间的交流。在推理时,根据教师网络与学生网络之间获取的特征差异,来对图像进行异常检测。图 1 中:T 表示教师蒸馏层、S 表示学生蒸馏层、L 表示二者之间的损失差。

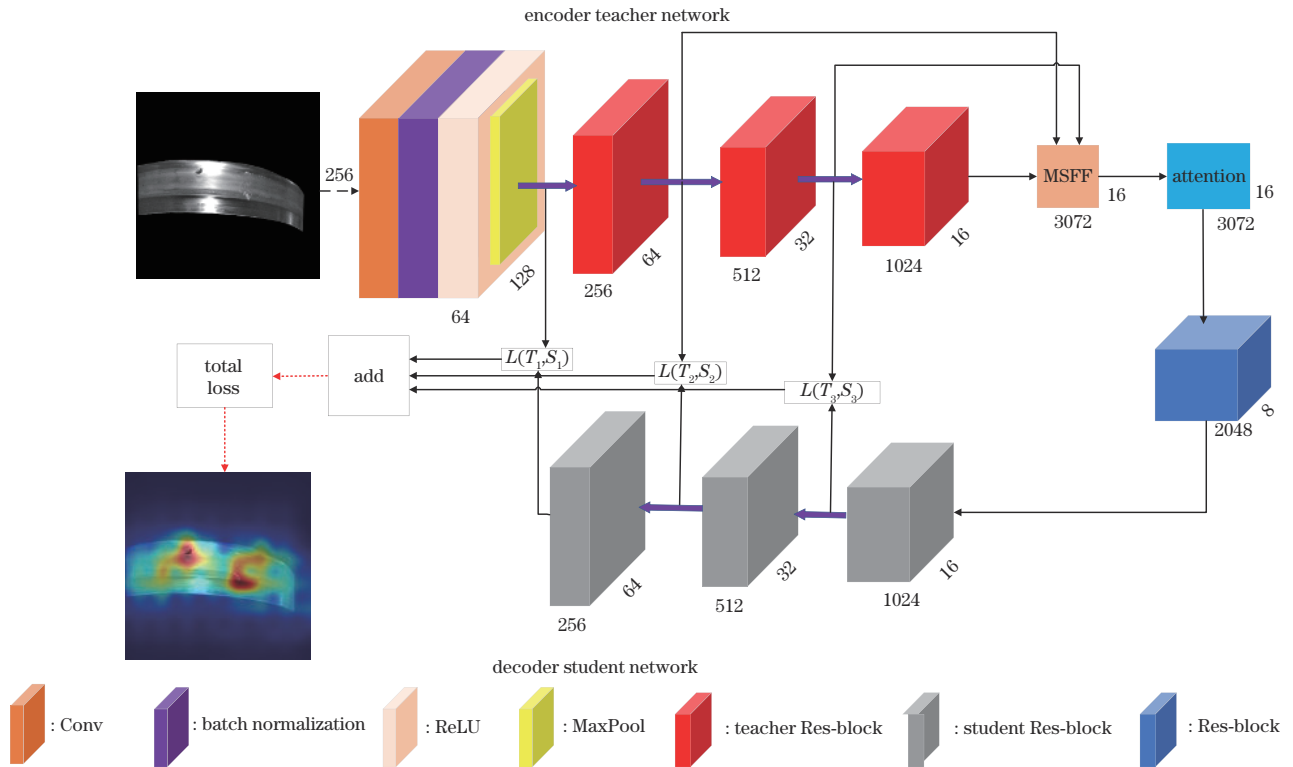


图 1 基于注意力反向知识蒸馏网络结构

Fig. 1 Reverse knowledge distillation network structure based on attention

基于注意力反向知识蒸馏由教师网络、一类瓶颈嵌入模块和学生网络 3 部分组成。教师网络主要由 3 个编码模块(en\_block)堆叠而成,第 4 个 en\_block 作为可训练的一类瓶颈嵌入模块。一类瓶颈嵌入模块使

用 en\_block 将多尺度特征融合模块融合出的特征映射到低维空间,异常特征表述为正常模式上的扰动,有助于阻止异常特征信息传递到学生网络。在一类瓶颈嵌入块中的多尺度特征融合模块后添加了注意力机制,

减轻了多尺度特征融合模块中卷积压缩带来的模糊效果和增强学生网络重建初始特征的纹理信息。为了能够进行特征知识的传递,学生网络是与教师网络对称的镜像解码器结构,该结构由 3 个解码模块(de\_block)堆叠而成。en\_block 由多个瓶颈结构组成,瓶颈结构如图 2 所示。瓶颈结构包含有两个 1×1 卷积和一个 3×3 卷积,每个卷积之后跟着批量归一化(BN)和 ReLU 激活函数,c 代表输入的特征通道数,其在 en\_block 1 至 en\_block 3 中分别为 128、256、512,虚线代表与上一个瓶颈结构进行跳跃连接。在教师网络中,en\_block 1 包含 3 个瓶颈结构,en\_block 2、en\_block 3 分别包含 4、6 个瓶颈结构,3 个 en\_block 提取到的特征通道数分别为 256、512、1024。与教师网络提取特征是下采样的过程相反,学生网络对特征的重建是上采样的过程,学生网络的 de\_block 的瓶颈结构与教师网络的 en\_block 的类似,如图 3 所示,每个 de\_block 的第一个瓶颈结构中的 3×3 卷积变成了 2×2 的反卷积。在学生网络中,de\_block 1 至 de\_block 3 分别包含有 6、4、3 个瓶颈结构块,3 个 de\_block 各自输入的通道数 c 分别为 2048、1024、512,重建后的特征通道数分别为 1024、512、256,其与教师网络提取出的特征是镜像对称的。

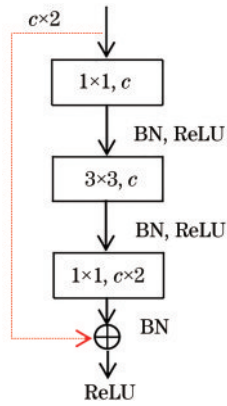


图 2 编码器瓶颈结构

Fig. 2 Structure of encoder bottleneck

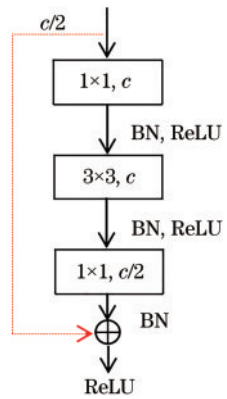


图 3 解码器瓶颈结构

Fig. 3 Structure of decoder bottleneck

计算教师网络与学生网络对应特征之间的余弦相似度,然后再减掉余弦相似性,即可得到两个特征之间的二维异常特征图。最后将 3 组特征异常图进行叠加,得到整张图像在二维平面上的异常特征图。利用在映射特征上积累得到的平均异常值来判断二者提取出的特征差异性,这对只具有小异常区域的样本而言是不公平的,所以将二者在映射特征上积累得到的像素级异常的最大值作为评判样本级异常的标准。

### 2.2 多尺度特征融合模块

多尺度特征融合模块如图 4 所示。对由教师编码器最后一个残差块编码得出的具有语义和结构信息的高级特征进行特征重构,学生解码器很难重构出低级特征。所以模型将教师编码器提取出分别具有低级纹理和高级语义的 3 个特征组,并使用多尺度特征融合模块将其融合。3 个输入特征组分别为  $I_1 \in R^{C_1 \times H_1 \times W_1}$ ,  $I_2 \in R^{C_2 \times H_2 \times W_2}$ ,  $I_3 \in R^{C_3 \times H_3 \times W_3}$ ,为了能够使其进行融合,需将  $I_1$  进行两次下采样, $I_2$  进行一次下采样,才能将尺度缩小到与  $I_3$  同等大小。每一次下采样需通过一个步长为 2 的 3×3 卷积层,然后进行 BN,最后使用 ReLU 函数进行激活。将下采样后的所有特征  $I'_1$  和  $I'_2$  与不进行下采样的特征  $I_3$  先在通道维数上进行拼接,然后再使用步长为 1 的 1×1 卷积层和 ReLU 激活函数,激活 BN 来获得丰富而紧凑的特征。输出特征  $f_{output}$  可表示为

$$I'_1 = \text{ReLU} \left\{ \text{BN} \left\{ \text{Conv}_{3 \times 3} \left\{ \text{ReLU} \left\{ \text{BN} \left[ \text{Conv}_{3 \times 3} (I_1) \right] \right\} \right\} \right\} \right\}, \quad (1)$$

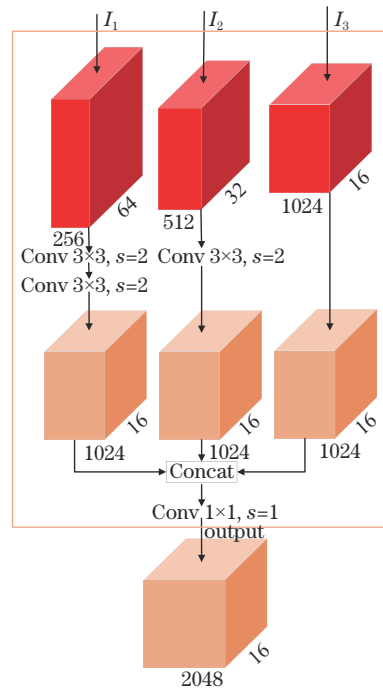


图 4 多尺度特征聚合

Fig. 4 Multi-scale feature fusion

$$I'_2 = \text{ReLU}\left\{\text{BN}\left[\text{Conv}_{3\times 3}(I_2)\right]\right\}, \quad (2)$$

$$f_{\text{output}} = \text{ReLU}\left\{\text{BN}\left\{\text{Conv}_{1\times 1}\left[\text{Concat}(I'_1, I'_2, I_3)\right]\right\}\right\}, \quad (3)$$

式中： $\text{Conv}_{3\times 3}$  为卷积核为 3，步长为 2 的卷积层； $\text{Conv}_{1\times 1}$  为卷积核为 1，步长为 1 的卷积层。 $\text{Concat}$  为特征在通道维数上的拼接， $\text{BN}$  为批量归一化， $\text{ReLU}$  为激活函数。

### 2.3 引入的注意力机制

教师网络与学生网络之间的一类瓶颈嵌入模块是编码-解码结构式的知识蒸馏的关键所在，在推理阶段，该模块通过抑制异常特征的信息传递，使学生网络只会根据在训练阶段学习到的正常特征知识来对异常特征进行重建。在遇到异常样本时，教师网络提取到的特征和学生网络重建出的特征具有差异性，使模型更容易对异常图像进行判断。在一类瓶颈嵌入模块中，多尺度特征聚合模块将教师网络提取出的 3 个不同大小的特征在通道维度上融合成与最后一个特征大小相同的融合特征。然而，在此过程中，具有纹理信息的低层特征经过卷积压缩，会模糊掉一些特征信息，而模糊的正常特征使学生网络不能很好地对其进行重建，这与教师网络的正常特征之间就会有较大差别，一

旦差别过大就会造成模型的误判。所以，为了保证学生网络在重建过程有一个较好的初始特征，以及凸显出异常特征，使一类瓶颈嵌入模块能够发现并对其进行抑制，阻断异常信息的流动，在一类瓶颈嵌入模块中的多尺度特征聚合模块和编码模块之间加入注意力机制，用于增强特征的纹理信息，从而提升模型对异常的鉴别能力。

注意力机制如图 5 所示。该模块由通道注意力 Squeeze-and-excitation networks (SE) 和空间注意力 Split groups spatial enhance (SGSE) 结合组成，其主要作用是在原始特征上叠加筛选而出的细致的纹理特征信息。首先，输入的融合特征先经过 SE，与原始输入特征进行加权叠加形成整个通道注意特征图；然后，通道注意特征图分成两路，一路经过 SGSE 后与另一路进行对应元素乘积，形成空间注意特征图；最后，通道注意特征图与空间注意特征图进行加权叠加形成最终的注意力特征，图 5 中的系数值都为 0.5。原始特征图在通道特征图和空间特征图的混合叠加下，进一步强化了图像上物体的特征，避免了因卷积压缩导致的纹理细节的模糊化效果，提升了模型对物体的检测能力。

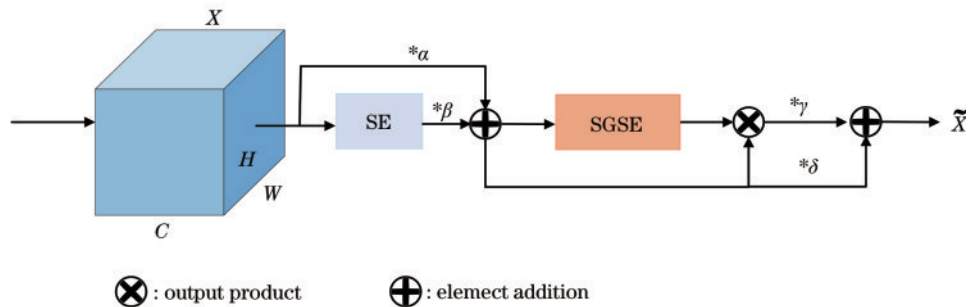


图 5 注意力机制  
Fig. 5 Attention mechanism

SE 模块的结构如图 6 所示。输入的特征经过全局池化，特征图变成有  $c$  个通道的一维向量，接着使用两个通道数相同的全连接层来给每个通道生成对应的权重，然后使用 Sigmoid 对生成的权重进行归一化操作，

最后将归一化后的通道权重系数与输入的原始特征的通道进行相乘，得到最终的特征输出。经过 SE 模块后的特征通道，其重要性程度变得不同，卷积神经网络根据不同的重要性程度来判断该通道是否值得关注。

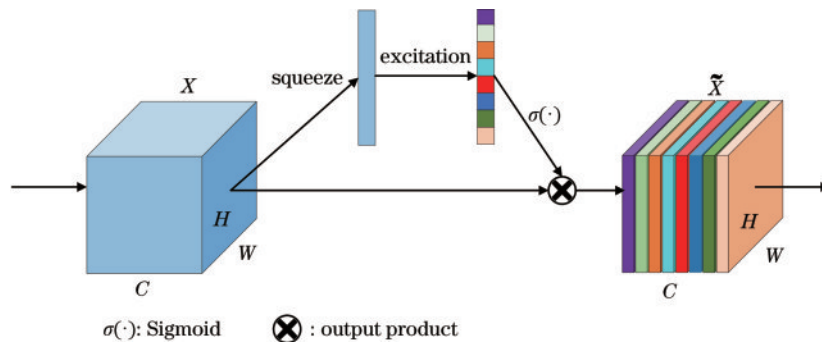


图 6 SE 模块结构  
Fig. 6 SE module structure

SGSE 模块的结构如图 7 所示。输入的特征在通道维数上被划分为多个组,每组都会携带不同语义的子特征。不同于在特征相应空间位置上生成语义特征,针对每一个特定语义的组,使用全局和局部特征的

相关性来生成一个注意力掩码,其在特征的空间维度上产生作用,能够减少特征中可能存在的噪声和干扰,提高区域特征的语义正确性,增强卷积神经网络在特征特定区域细节上学习的可表示性。

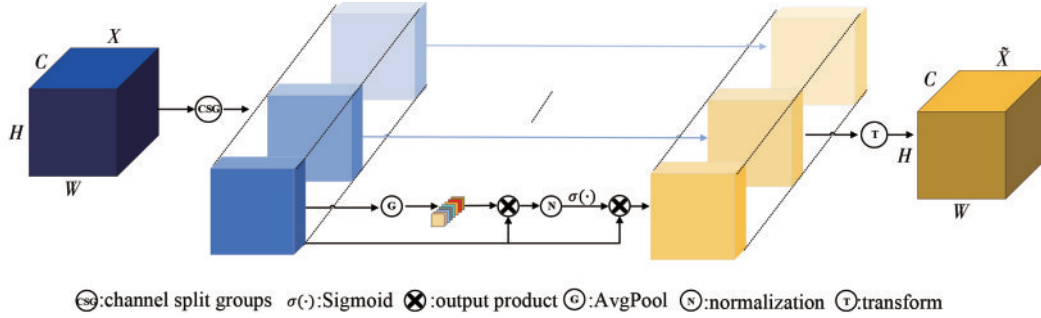


图 7 SGSE 模块结构

Fig. 7 SGSE module structure

### 3 实验设计以及实验结果分析

#### 3.1 实验环境以及评价指标

所有实验均在配备有 Windows 操作系统、RTX3070 显卡、PyCharm 软件、PyTorch 1.8、CUDA 11.0、cuDNN 11.0 的计算环境中进行训练和测试。在分割实验中采用平均交并比 (Mean intersection over union, mIoU) 来评价分割的效果。而在异常检测实验中,采用受试者工作特性曲线下的面积 (Area under receiver operating characteristic curve, AUC)、召回率  $R$ 、精确率  $P$ 、准确率 (Acc) 来共同评价异常检测模型的检测性能。

分割实验采用了经典的 UNet 模型,所有输入的图片大小都被调整为  $512 \times 512$ ,在训练时采用 Adam 优化器来优化分割模型的收敛过程,其参数  $\beta = (0.9, 0.999)$ ,初始学习率为 0.0001,批量大小为 2,总共训练 100 轮。异常检测实验的教师网络使用

WideResnet50-2 在 ImageNet 上预训练后的权重进行初始化,图片输入大小为  $256 \times 256$ ,在训练时也是采用 Adam 优化器来优化,其参数  $\beta = (0.5, 0.999)$ ,初始学习率为 0.0001,批量大小为 4,总共训练 400 轮。

#### 3.2 铁路车轮踏面数据集

铁路车轮踏面是指车轮与钢轨顶面的接触部分,数据集是由 CCD 相机拍摄得到的。所有收集到的铁路车轮踏面数据被分成分割数据集和异常检测数据集。分割数据集共有 3000 张车轮踏面图像,其中训练集有 2400 张,占总数的 80%;测试集有 600 张,占总数的 20%。异常检测数据集共有 7173 张车轮踏面图像,训练集有 4486 张正常类型的,占总数的 62.5%;测试集有 2687 张,占总数的 37.5%,其中正常类型的有 1301 张,异常有缺陷类型的有 1386 张。分割数据集和异常检测数据集互不相同。异常检测数据集中异常类型包含剥离、擦伤、硌伤 3 类常见的缺陷,其数量分别为 202、715、476。车轮踏面图像如图 8 所示。

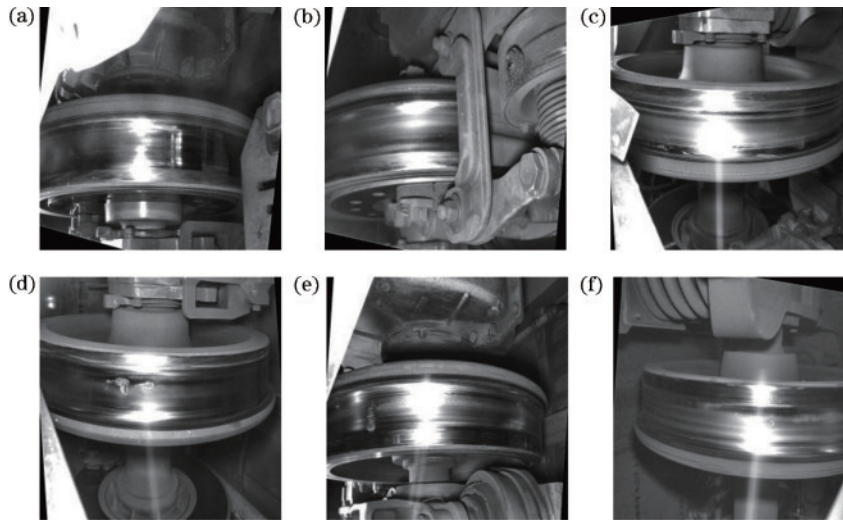


图 8 车轮踏面图像。(a)~(c) 正常车轮踏面图像;(d) 剥离缺陷图像;(e) 擦伤缺陷图像;(f) 硌伤缺陷图像

Fig. 8 Wheel tread image. (a)~(c) Normal wheel tread image; (d) peeling defect image; (e) scratch defect image; (f) injury defect image

### 3.3 分割模型在数据集上的实验结果

在数据集的构建过程中,原始的车轮踏面图像中除了车轮踏面区域,还包含有其他非踏面区域的背景,非踏面区域可能会由于相机拍摄的机位不一致而导致其发生变化。若图像中的背景区域过于复杂,就会带来很多不确定性的变化因素,而只针对车轮踏面上的情况来判断车轮是否存在异常,直接进行异常检测就无法精确地检测踏面区域的状况。所以,为了减少其他非踏面区域带来的不确定性因素的干扰,使用UNet分割模型来对异常检测数据集中的车轮踏面图像进行分割,保留图像中的踏面区域。

分割模型的训练损失曲线如图9所示。训练集共2400张图像,批次为2,轮次为100,则训练即有 $1.2 \times 10^5$ 次。由图9可知,损失曲线刚开始急剧下降,然后逐渐趋近于平缓,即模型达到较好的收敛状态,说明此时分割模型对车轮踏面区域的分割已具备良好的性能。

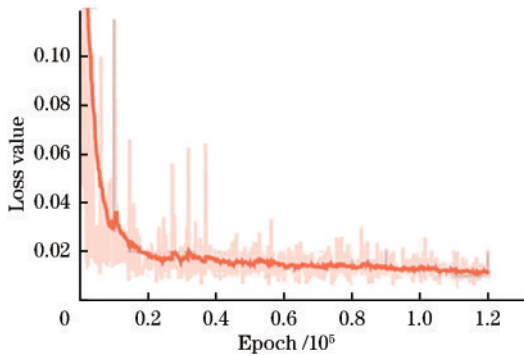


图9 分割模型训练损失曲线

Fig. 9 Segmentation model training loss curve

训练好的分割模型在具有600张车轮踏面图像的测试集上达到97.8%的mIoU。图10为踏面图像踏面区域的分割预测,从ground truth和predict mask来看,UNet分割模型在车轮踏面数据集上进行二分类分割

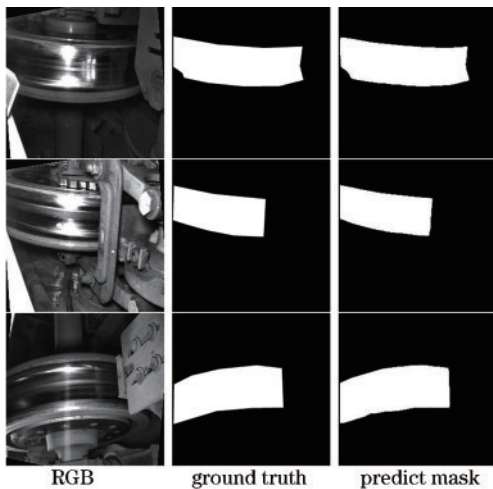


图10 车轮踏面图像分割预测

Fig. 10 Prediction diagram of wheel tread image segmentation

性能表现优异。

踏面区域分割也将作为构建异常数据集最后的预处理操作,分割后的踏面图像即为异常检测实验所用的数据集。

### 3.4 反向知识蒸馏异常检测模型在异常检测数据集上的实验结果

为了验证异常检测车轮踏面图像分割的必要性,使用不加注意力机制的反向知识蒸馏检测模型分别在分割前和分割后的异常检测数据集进行异常检测实验。

如表1所示,反向知识蒸馏模型在分割后的异常检测数据集上的指标除了召回率较低之外,其余指标均有较大提升。其原因是未分割的车轮踏面图像的非踏面区域部分过于复杂,在训练过程中若未学习过的非踏面区域或该角度的图像较少,无监督异常检测模型无法从中统计出一个较好的正常特征分布,所以只要在测试集中有未见过或极少见过的非踏面区域,模型就会认为是异常样本,这导致召回率高和精准率低,也影响了AUC和准确率。车轮踏面图像在进行分割之后,无监督异常检测模型就只会学习踏面上的特征,避免了背景因素的干扰,能有效地提高模型的检测性能。

表1 反向知识蒸馏模型在异常检测数据集分割前后的结果对比

Table 1 Comparison of results of reverse knowledge distillation model before and after anomaly detection dataset segmentation

Model	Segmentation	AUC / %	P / %	R / %	Acc / %
RD	Before	66.7	55.5	97.9	58.4
	After	93.6	82.1	94.0	86.3

未分割的车轮踏面图像在进行异常检测时,检测模型也会将非踏面区域的某些部位视为异常,这会造成模型的误报,如图11、图12所示。反观分割后的车轮踏面图像,检测区域就限制在车轮踏面区域上,基本

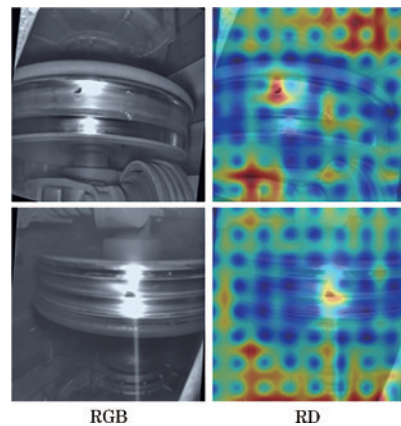


图11 车轮踏面图像分割前在RD模型上的检测视图

Fig. 11 Detection view on RD model before wheel tread image segmentation

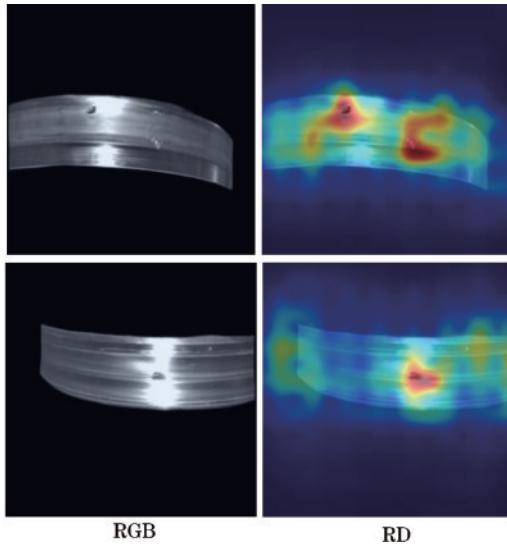


图 12 车轮踏面图像分割后在 RD 模型上的检测视图  
Fig. 12 Detection view on RD model after wheel tread image segmentation

上避免了其余背景的干扰,这对模型的检测是有利的。所以车轮踏面图像在进行异常检测之前,有必要对其踏面区域进行分割。

### 3.5 消融实验

为了验证所设计的注意力机制对检测性能的影响,对模型进行消融实验。消融实验的结果如表 2 所示。

表 2 中,不添加注意力的反向蒸馏模型在异常检

表 2 消融实验结果的对比  
Table 2 Comparison of ablation results

Method	AUC / %	P / %	R / %	Acc / %
RD	93.6	82.1	94.0	86.3
RD+Attention(our)	93.8	82.3	95.4	87.0

测数据集上的 AUC 为 93.6%,精准率为 82.1%,召回率为 94.0%,准确率为 86.3%。在一类瓶颈嵌入模块中添加注意力机制后,AUC 为 93.8%,精准率为 82.3%,召回率为 95.4%,准确率为 87.0%。相比较之下,检测指标性能均提升,这也证明了所提的注意力机制能缓解特征在多尺度聚合后由于卷积压缩造成的模糊效应,一方面为学生网络提供具有更多正常纹理信息的特征,另一方面有利于一类瓶颈嵌入模块淹没掉明显的异常信息,防止异常信息的流动。该结果表明:所提的注意力有助于提升检测模型的性能。

剥离、擦伤和硌伤 3 种缺陷在模型改进前后检测的召回率对比结果,如表 3 所示。由表 3 可知,与反向知识蒸馏模型相比,所提方法有更高的召回率,即能检测出更多的缺陷。这体现出所提的注意力机制能够更明显地增强特征的纹理信息,在推理时,使一类瓶颈嵌入模块更易发现异常信息,并对其进行抑制,导致学生网络难以重建出异常图像的特征信息,加大了教师网络提取的特征和学生网络重建的特征之间的差距,从而能够检测出更多的缺陷。

表 3 不同类型的缺陷在模型改进前后检测的召回率  
Table 3 Recall of different types of defects detected before and after model improvement

Method	Recall of peel / %	Recall of scratch / %	Recall of injury / %
RD	97.5	99.0	85.1
RD+Attention (our)	98.0	99.4	88.2

反向知识蒸馏模型与所提方法在异常检测数据集上检测的可视化结果如图 13 所示。由图 13 可知,正是由于卷积压缩融合特征的操作,模糊了可表示的正常特征的纹理信息,原始方法在重建正常特征上是有所欠缺的。模型在添加注意力后,强化了学生网络对重

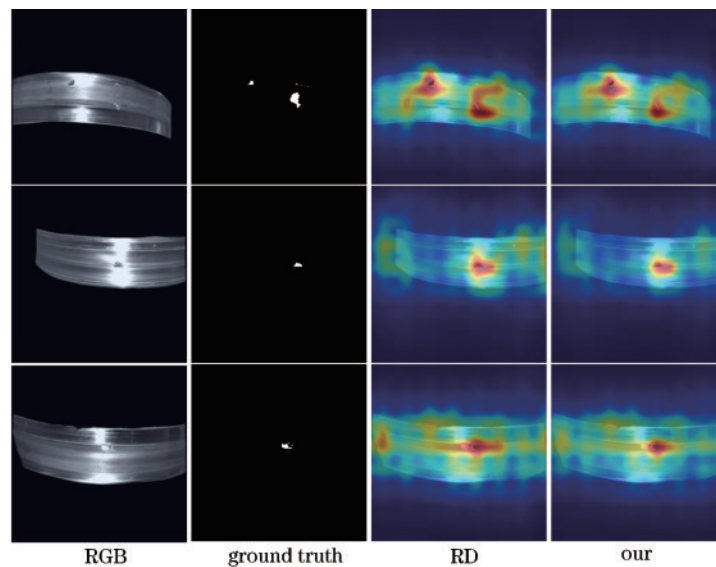


图 13 反向知识蒸馏方法与所提方法在车轮踏面上的检测视图对比  
Fig. 13 Comparison of detection view of wheel tread by reverse knowledge distillation method and the proposed method

建的初始特征的纹理信息,在检测时减少正常区域特征对异常值的积累,使得模型对缺陷区域的判断更加精确。

### 3.6 所提方法与其他异常检测网络实验结果的对比

在异常检测数据集上将所提方法与 GANomaly 和 MKD 进行对比,对比结果如表 4 所示。从实验数据来看,基于特征空间的 MKD 和所提方法比基于生成的 GANomaly 方法在检测效果上更好。由图 14 的检测可视化结果可知,GANomaly 深受车轮踏面上反光的影响,该模型也将大部分的反光区域检测成为异常区域,这在很大程度上提高了样本的异常分数,造成高召回率和低精准率的结果。而所提方法一方面继承了

其能够克服异常激活差异消失的弊端和改善未知样本表示的多样性优点,另一方面添加的注意力机制能够增强特征的纹理信息,有利于学生网络对正常特征进行重建和一类瓶颈嵌入模块对异常特征信息的抑制。特殊的知识蒸馏结构以及注意力的协助,使所提方法对车轮踏面上异常检测具有较好的效果。

表 4 不同模型的实验结果的对比

Model	AUC / %	P / %	R / %	Acc / %
GANomaly	57.8	51.9	100.0	52.2
MKD	84.3	73.0	88.9	77.3
Our	93.8	82.3	95.4	87.0

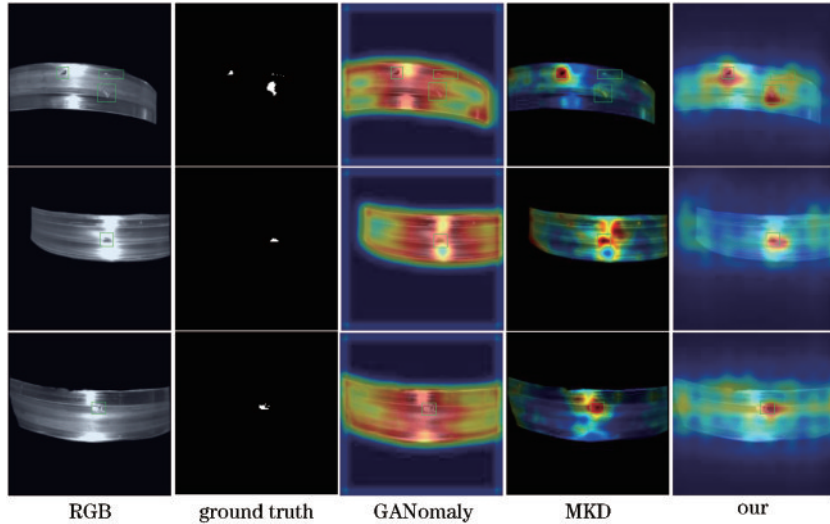


图 14 不同异常检测方法所提方法在车轮踏面上的检测视图对比

Fig. 14 Comparison of detection view of wheel tread by different anomaly detection methods and the proposed method

剥离、擦伤、硌伤 3 种缺陷在不同模型上检测的召回率对比结果,如表 5 所示。GANomaly 方法能够检测出所有的缺陷,但从表 4 对应的精准率和准确率来看,该方法也会将很大一部分的正常样本检测为异常样本,其整体的检测结果差。所提方法在不同类型缺陷的召回率上都比 MKD 方法高,这体现所提方法拥有更好的检测性能。从检测结果来看,有着更为明显特征和更大范围区域的剥离和擦伤两种缺陷被检出的概率是最大的,而硌伤缺陷特征不太明显,所能提取的特征较为微弱,以至于有些异常样本在教师网络和学生网络二者的比较中体现不出差别,漏检的概率较大。

表 5 不同类型的缺陷在不同模型上检测的召回率  
Table 5 Recall of different types of defects detected on different models

Model	Recall of peel / %	Recall of scratch / %	Recall of injury / %
GANomaly	100	100	100
MKD	95.6	91.7	82.1
Our	98.0	99.4	88.2

异常样本检测的可视化效果如图 14 所示。由图 14 可知,GANomaly 的检测效果不比 MKD 和所提方法好,其对正常区域的错误判断,导致检测能力较差。MKD 能够检测出车轮踏面异常区域,但也存在对不明显缺陷漏检和对一些正常区域以及不规则边缘误检的情况。而所提方法降低了对正常区域误检的可能性,同时提高了异常区域在特征上的辨识度,该模型对车轮踏面上异常检测的表现较好。

## 4 结 论

为解决铁路车轮踏面缺陷样本有限的问题,提出了一种基于注意力反向知识蒸馏的车轮踏面异常检测模型。在进行异常检测前,为避免图像中车轮非踏面区域对异常检测模型的影响,使用训练好的 UNet 分割模型对车轮踏面区域进行分割,使异常检测模型对图像的检测能够聚焦在车轮踏面之上。反向知识蒸馏的实验表明,在分割后的车轮踏面图像上进行异常检测,AUC、精准率、召回率和准确率均大幅提升,其检测视图也反映了分割的必要性。在反向知识蒸馏模型



的一类瓶颈嵌入模块中添加了注意力机制,在对铁路车轮踏面异常检测数据集的检测中,达到 93.8% 的 AUC、82.3% 的精准率、95.4% 的召回率、87.0% 的准确率,与原模型相比,AUC 提高了 0.2 个百分点、精准率提高了 0.2 个百分点、召回率提高了 1.4 个百分点、准确率提高了 0.7 个百分点,说明了多尺度特征融合的卷积压缩存在模糊特征的纹理信息,而添加的注意力机制能够对特征的纹理信息进行清晰化处理,有效地弥补信息的不足。在与其他异常检测网络相比时,所提方法在检测中受踏面反光区域的影响小,对正常区域的特征有较强的鉴别能力,同时能够捕捉到不太明显的缺陷特征,其检测能力相对是最强的。综合来说,所提方法的检测指标和检测效果都是较好的。

### 参 考 文 献

- [1] 任锦霞,张凤生,宋腾,等.基于激光位移传感器的车轮踏面磨损检测技术研究[J].机械与电子,2017,35(12):59-62.  
Ren J X, Zhang F S, Song T, et al. Research on wheel tread wear detection technology based on laser displacement sensor[J]. Machinery & Electronics, 2017, 35(12): 59-62.
- [2] 肖乾,姜雄峰,刘海涛,等.铁道车轮踏面损伤实时监测方法研究综述[J].华东交通大学学报,2021,38(4):99-112,126.  
Xiao Q, Jiang X F, Liu H T, et al. Research on real-time monitoring methods for railroad wheel tread defects: a review[J]. Journal of East China Jiaotong University, 2021, 38(4): 99-112, 126.
- [3] Wu K H, Zhang J H, Wu X Q, et al. Dynamic detecting system of the parameters of wheel tread profile based on image processing method[J]. Proceedings of SPIE, 2002, 4925: 604-607.
- [4] 薛琴,陈玮.模式识别在轮对踏面磨耗检测中的应用[J].微计算机信息,2007,23(22):259-261.  
Xue Q, Chen W. The application of pattern recognition in the measurement of wheel tread wear[J]. Microcomputer Information, 2007, 23(22): 259-261.
- [5] 赵勇,方宗德,田丽丽.列车车轮踏面缺陷的图像区域提取[J].光学精密工程,2009,17(4):901-908.  
Zhao Y, Fang Z D, Tian L L. Defect region extraction in images of train wheel tread[J]. Optics and Precision Engineering, 2009, 17(4): 901-908.
- [6] 赵勇.基于视觉的列车车轮踏面擦伤定位方法[J].电子技术应用,2011,37(8):139-141,146.  
Zhao Y. Research on peeling defect location method of wheel tread based on vision[J]. Application of Electronic Technique, 2011, 37(8): 139-141, 146.
- [7] 吴海滨,夏需堂.基于线阵 CCD 的车轮表面缺陷检测系统的图像预处理研究[J].巢湖学院学报,2011,13(3):61-67.  
Wu H B, Xia X T. Based on linear CCD wheel surface defect detection system of image preprocessing[J]. Journal of Chaohu College, 2011, 13(3): 61-67.
- [8] Nan G, Lu S F, Yao J N. Train wheel edge detection and image object region segmentation[J]. Proceedings of SPIE, 2016, 10157: 1015723.
- [9] 赵勇.基于 GA-RBFNN 算法的列车车轮踏面损伤识别[J].计算机工程与应用,2012,48(8):32-34.  
Zhao Y. Recognition of train wheel tread damages based on GA-RBFNN algorithm[J]. Computer Engineering and Applications, 2012, 48(8): 32-34.
- [10] 蒲富鹏,赵军,安喆.基于机器视觉的车轮裂纹识别与提取[J].铁道科学与工程学报,2018,15(8):2113-2122.  
Fu F P, Zhao J, An J. Recognition and extraction of wheel cracks based on machine vision[J]. Journal of Railway Science and Engineering, 2018, 15(8): 2113-2122.
- [11] 孙正,王树雁.深度学习在血管内光学相干层析成像中的应用现状[J].激光与光电子学进展,2022,59(22):2200002.  
Sun Z, Wang S Y. Application status of deep learning in intravascular optical coherence tomography[J]. Laser & Optoelectronics Progress, 2022, 59(22): 2200002.
- [12] 成珂阳,李琦.深度学习用于连续太赫兹同轴数字全息重建[J].中国激光,2023,50(17):1914001.  
Cheng K Y, Li Q. Deep learning for reconstruction of continuous terahertz in-line digital holography[J]. Chinese Journal of Lasers, 2023, 50(17): 1914001.
- [13] 李斌,杨阿坤,孙赵祥,等.基于深度学习的拼接镜共相检测新方法研究[J].中国激光,2023,50(22):2204001.  
Li B, Yang A K, Sun Z X, et al. Research on co-phasing detection new method of segmented mirror based on deep learning[J]. Chinese Journal of Lasers, 2023, 50(22): 2204001.
- [14] He J, Yu H Y, Zhang C F, et al. Damage detection of train wheelset tread using canny-YOLOv3[J]. Journal of Electronic Measurement and Instrumentation, 2019, 31(12): 25-30.
- [15] 杨凯,李锐,罗林,等.基于深度学习的车轮踏面表面缺陷检测研究[J].信息技术,2021,45(7):93-97.  
Yang K, Li R, Luo L, et al. Research on wheel tread surface defect detection based on deep learning[J]. Information Technology, 2021, 45(7): 93-97.
- [16] 郑茹丹,李金龙,张渝,等.基于改进 Faster R-CNN 的车轮踏面缺陷检测[J].中国铁路,2021(3):131-135.  
Zheng R D, Li J L, Zhang Y, et al. Wheel tread defect detection based on improved Faster R-CNN[J]. China Railway, 2021(3): 131-135.
- [17] He J, Hou N, Zhang C F, et al. Diagnosis of train wheelset tread damage based on EPSA-ResNet[J]. China Safety Science Journal, 2022, 32(5): 35.
- [18] Mei S A, Wang Y D, Wen G J. Automatic fabric defect detection with a multi-scale convolutional denoising autoencoder network model[J]. Sensors, 2018, 18(4): 1064.
- [19] Tarassenko L, Hayton P, Cernez N, et al. Novelty detection for the identification of masses in mammograms [C]//4th International Conference on Artificial Neural Networks, June 26-28, 1995, Cambridge, UK. New

- York: IEEE Press, 1995: 442-447.
- [20] Akcay S, Atapour-Abarghouei A, Breckon T P. GANomaly: semi-supervised anomaly detection via adversarial training[M]//Jawahar C V, Li H D, Mori G, et al. Computer vision-ACCV 2018. Lecture notes in computer science. Cham: Springer, 2019, 11363: 622-637.
- [21] Perera P, Patel V M. Deep transfer learning for multiple class novelty detection[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2020: 11536-11544.
- [22] Salehi M, Sadjadi N, Baselizadeh S, et al. Multiresolution knowledge distillation for anomaly detection[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 14897-14907.
- [23] Deng H Q, Li X Y. Anomaly detection via reverse distillation from one-class embedding[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-24, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 9727-9736.