



硕士学位论文

基于虚实结合的钢轨表面缺陷检测研究

Rail Surface Defect Detection Based on the Fusion of
Virtual and Real Data

作者：崔文凯

导师：李清勇

北京交通大学

2022年6月

学位论文版权使用授权书

本学位论文作者完全了解北京交通大学有关保留、使用学位论文的规定。特授权北京交通大学可以将学位论文的全部或部分内容编入有关数据库进行检索，提供阅览服务，并采用影印、缩印或扫描等复制手段保存、汇编以供查阅和借阅。同意学校向国家有关部门或机构递交论文的复印件和磁盘。学校可以为存在馆际合作关系的兄弟高校用户提供文献传递服务和交换服务。

(保密的学位论文在解密后适用本授权说明)

学位论文作者签名: 崔文凯

导师签名: 孙桂菊

签字日期: 2022 年 6 月 1 日

签字日期: 2022 年 6 月 1 日

北京交通大学

硕士学位论文

基于虚实结合的钢轨表面缺陷检测研究

Rail Surface Defect Detection Based on the Fusion of
Virtual and Real Data

作者姓名：崔文凯

学 号：19120343

导师姓名：李清勇

职 称：教授

学位类别：工学

学位级别：硕士

学科专业：计算机科学与技术

研究方向：计算机视觉

北京交通大学

2022 年 6 月

致谢

三年的时光匆匆而过，转眼之间已经到了毕业的路口。在这段日子里，经历了阳光灿烂的夏天般的得意，也体验过潮湿阴郁的冬季般的压抑。但是，逝者如斯，时间带走了我当时的心境，也洗涤了那些庸俗的杂质。现在回味起来，这旅途之中认识的人，明白的事，收获的道理是如此之多；这趟充满了激情与挑战的研究生之旅带给我的收获和甜蜜是如此之盛。而我深知，这一切的一切离不开我周围的那些亲爱的人和物。因此，我在这里对我的导师、家人、同学、朋友以及母校给予我的帮助、支持，陪伴和鼓励表达我的谢意。

首先，我要感谢我的导师李清勇老师。感谢他带我走进科研的大门，在读研之后，作为一个懵懵懂懂的门外汉突然闯入浩如烟海的科研世界，我变得迷茫且焦躁。而李老师每每总是在关键时候为我的学术研究指引了方向，为我驱散科研路上的迷雾。感谢李老师在个人成长上对我的鞭策和鼓励。李老师是一个温柔严格的人。在我们放松懈怠时，他会用合适的方式来提醒我们黑发不知勤学早，白首方悔读书迟；在我们急于求成的时候，他会告诉我们不积跬步无以致千里，不积小流无以成江海；在我们与人交往时，他会教育我们赠人玫瑰，手有余香；在李老师的引导下，我不仅仅学会了如何做好学术，更是深刻的认知到了自己的性格缺点，也懂得了人生在世要做一个知恩图报的君子。不仅如此，李老师还会经常组织实验室一起出游踏青、聚餐等实验室团建活动来丰富我们的业余生活，让读研的生活更加的多姿多彩。

其次，我要感谢实验室的每一位同学对我的帮助和陪伴。感谢王建柱师兄对我学术研究和论文撰写中的指导与帮助。作为我们课题组的大师兄，他认真负责，耐心细致，在潜移默化中影响着我，促使我成为一个更好的自己。感谢俞浩敏师姐、梁凤娇师姐、张靖师兄和李小宝师兄对我在科研探究当中提供的帮助。感谢祝叶舟、周心源、苏少博和李东等实验室同学的帮助与关心，陪伴我度过了美好的实验室时光。感谢与我一起共同奋战虚实结合项目的彭文娟师姐和黄祺隆师弟，有了你们的参与和帮助，我们的项目才能推进的更快更远。感谢我的室友陈鑫和邓祎程同学，你们幽默风趣，乐观开朗的性格大大影响了我。感谢我的挚友张广阳同学，每当我遇到挫折，你总愿意听我分享和提供帮助，陪我度过人生的低谷。感谢铁科院的王乐博士和王胜春博士在数据和方法研究上的支持与帮助。感谢我的父母。感谢你们为我提供了世界上最温馨的港湾，教会了我知足常乐的人生道理，你们是我最爱的人。

最后，特别感谢百忙之中参加论文评审和毕业答辩的评审老师们，感谢你们给予本文的宝贵建议。

摘要

随着钢轨承载压力的逐年增加，钢轨表面缺陷的发生频率也越来越高。因此对钢轨表面缺陷进行快速，准确的识别和定位也变得愈加重要。基于深度神经网络的钢轨表面缺陷检测方法由于其强大的特征提取能力与学习能力，具有进一步提高钢轨表面缺陷检测能力的巨大潜力。然而，铁路单位对病害钢轨提前预知并替换的工作使得很难收集到足量的病害钢轨图像，从而无法有效完成对神经网络模型的训练。基于此，本文提出了基于虚实结合的钢轨表面缺陷检测方法。

首先设计实现了基于 3D 模型和点云处理方法的虚拟数据生成系统，其可以自动化地生成大量的病害钢轨 3D 模型。在此过程中，分析了影响 3D 模型生成效果的“锯齿现象”，并设计了曲率和反曲率操作来解决这个问题使得生成的钢轨 3D 模型更加的真实。基于 MaxScript 实现了批量表面图像采集软件来自动化对虚拟场景中的钢轨 3D 模型进行渲染获取其表面图像，这些图像可作为神经网络训练的数据集。基于点云处理和模型生成中的特点，设计了标签信息映射方法来生成可用于神经网络训练的标签信息。

其次，基于虚拟数据和真实数据存在较大分布差异的弊端，提出了基于虚实结合的钢轨表面缺陷检测方法，其包括图像对齐和特征对齐两个模块。图像对齐通过风格迁移和纹理叠加在图像级别上使虚拟数据和真实数据在视觉特征上更加相似，初步减轻虚拟数据和真实数据的分布差异。特征对齐则是优化特征提取器，使其提取到的虚拟和真实数据的特征，在分布上趋于一致。最后，通过将图像对齐模块和特征对齐模块加入到 YOLOv3 基底网络当中构建了本文的 DA-YOLO（Domain Adaptation YOLO）检测网络。

在实验结果方面，本文基于仿真图像训练了一批经典和最新的目标检测网络，并将其在真实图像上的检测结果与一些传统图像检测方法进行对比。相比于传统方法，基于生成的仿真图像所训练的神经网络模型有着更高的准确率和 F1-Score 值，综合性能最高提升了 43.4 个百分点，这证明了虚拟数据生成系统的有效性。其次，本文也实验验证了图像对齐模块和特征对齐模块的性能表现。对比基底网络，添加了两个模块的 DA-YOLO 提升了 40.3 个百分点，同时也超出了所有对比方法的检测效果，充分证明了两个模块的优越性。本文也做了相应的消融实验证实了两个模块都是有效的并且效果相辅相成。

关键词：钢轨缺陷检测；3D 模型；点云处理；目标检测；神经网络；域适应

ABSTRACT

As the bearing pressure of the rail increases year by year, the frequency of surface defects of the rail is also increasing. Therefore, it has become more and more important to quickly and accurately identify and locate the surface defects of the rail. The rail surface defect detection method based on deep neural network has great potential to further improve the rail surface defect detection ability due to its powerful feature extraction ability and learning ability. However, the work of railway workers to predict and replace the diseased rails in advance makes it difficult for us to collect enough images of the diseased rails, so that the training of the neural network model cannot be effectively completed. Based on this, we propose a rail surface defect detection method based on virtual data and real data.

Firstly, a virtual data generation system based on 3D model and point cloud processing method is designed and implemented, which can automatically generate a large number of 3D models of diseased rails. During this process, we analyzed the "jaggies problem" that affects the effect of 3D model generation, and designed a curvature operation to solve this problem and make the generated 3D model of the rail more realistic. The batch surface image acquisition software is implemented based on MaxScript. It can automatically render 3D models of rails in virtual scenes to obtain surface images, which can be used as datasets for neural network training. Based on the characteristics of point cloud processing and model generation, a label information mapping method is designed. These labels can be used for neural network training.

In addition, a rail surface defect detection method based on the combination of virtual and real data is proposed, which includes two modules, image alignment and feature alignment. Through style transfer and texture overlay methods, image alignment makes virtual data and real data more similar in visual features. This can initially alleviate the distribution difference between the virtual data and the real data. Feature alignment is about optimizing the feature extractor. It makes the features of the extracted virtual and real data tend to be consistent in the data distribution. Finally, our DA-YOLO detection network is constructed by adding our image alignment module and feature alignment module to the YOLOv3 base network.

In terms of experimental results, this paper trains some classic and state-of-the-art object detection networks on simulated images. Then its detection results on real images are compared with some traditional image detection methods. Compared with traditional

methods, the comprehensive performance of the model trained based on our generated simulated images is up to 43.4 percentage points higher, which proves the effectiveness of our virtual data generation system. This paper also experimentally verifies the performance of the image alignment module and the feature alignment module. Compared with the base network, DA-YOLO with the addition of two modules improves by 40.3 percentage points. At the same time, it also exceeds the detection effect of all contrasting methods, fully proving the superiority of the two modules. We also performed corresponding ablation experiments to confirm that both modules are effective and the effects complement each other.

KEYWORDS: Rail defect detection; Digital model; Object detection; Neural network; Domain adaptation

目录

摘要	iii
ABSTRACT	iv
1 绪论	1
1.1 研究背景与意义	1
1.2 论文的主要工作	3
1.3 各个章节介绍	5
2 相关研究工作与理论基础	7
2.1 钢轨表面缺陷检测研究现状	7
2.1.1 非视觉钢轨表面缺陷检测方法概述	7
2.1.2 视觉钢轨表面缺陷检测方法概述	10
2.1.3 国内外商用钢轨缺陷检测系统	12
2.2 相关理论基础	14
2.2.1 卷积神经网络	14
2.2.2 目标检测	17
2.2.3 域适应	20
3 虚拟数据生成系统设计与实现	21
3.1 引言	21
3.2 虚拟数据生成系统	22
3.2.1 虚拟数据生成系统概述	22
3.2.2 病害钢轨 3D 模型库构建方法	23
3.2.3 基于点云处理的数据增广方法	24
3.2.4 曲率和反曲率操作	26
3.2.5 钢轨 3D 模型表面图像采集	28
3.2.6 标签信息自动提取	29
3.3 实验	30
3.3.1 实验设置	30
3.3.2 实验结果与分析	32
3.4 小结	34
4 基于虚实结合的钢轨表面缺陷检测方法	35
4.1 引言	35
4.2 虚实结合研究方法	36

4.2.1 伪目标域	36
4.2.2 图像对齐模块	37
4.2.3 特征对齐模块	40
4.2.4 适应性模型 DA-YOLO	41
4.3 实验	42
4.3.1 实验设置	42
4.3.2 实验结果与分析	44
4.4 小结	46
5 总结与展望	47
参考文献	49
作者简历及攻读硕士学位期间取得的研究成果	55
独创性声明	56
学位论文数据集	57

1 绪论

1.1 研究背景与意义

铁路是我国举足轻重的一种交通工具，承载着我国众多大宗物资和长途旅客的运输重责。铁路在我国地位的尊崇一方面取决于我国内陆深广的地理位置特点，另一方面也取决于其在远距离运输交通工具中无与伦比的优势。铁路运输成本低、运载大、安全度高、污染小，是内陆交通的最佳方式之一。

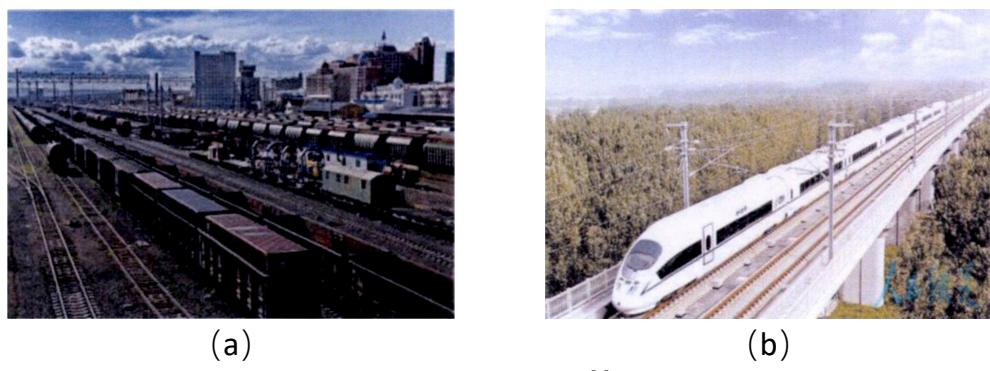


图 1-1 铁路运输^[1]
Figure 1-1 Rail transportation

截止到 2021 年底，中国正在运行的铁路已经达到了 15 万公里左右，其中包括 4 万公里左右的高速铁路，并计划在 2022 年再投产新线 3300 公里。而随着中国铁路运营线路的不断增长，一方面为人民带来了更加便利的出行条件，另一方面对铁路的安全性保障也有了更高的要求。1997 年 4 月 29 日，从昆明开往郑州的 324 次旅客列车在行驶到京广线附近时，发生了 126 人死亡，230 人受伤的重大交通事故。无独有偶，2008 年 4 月 18 日，在行进到胶济铁路和王村期间，北京去往青岛的 T195 次列车发生脱轨，与烟台往徐州的 5034 次列车发生了激烈相撞，引发了 70 人死亡，416 人受伤的惨剧。可以看到，一旦发生铁路交通事故，其造成的人员伤亡和经济损失都是巨大的，因此，对于铁路安全的要求再重视都不为过，对铁路交通事故的预防不容缺失。

钢轨作为整个铁路运输中一个核心组成部分，其不仅负责承载上千吨的重量，还要引导铁路运输的路线和方向。因此，维护钢轨设施正常运行是保障铁路运行安全的关键。在我国的铁道行业标准中对表面质量的检验方式比如落锤实验、超声探伤等大多只能证明当下钢轨状态正常。而随着钢轨在役时间的增加，其受到的挤压和磨损极大概率会使得钢轨出现各式各样的病害。钢轨表面缺陷的产生有两个原因，第一是在锻造过程中由于工艺水平不足导致产生了裂纹、皱褶等肉眼难以分辨

的微型缺陷，投入使用后经过运行设备的不断碾压造成缺陷程度加剧，从而得以暴露。第二则是符合质量要求的钢轨，在长时间和高负荷的运行状态中自然老化，钢轨发生疲劳磨损。常见的钢轨表面缺陷类型众多，包括擦伤、剥落、金属碎裂、拉丝结疤等。在列车高速运动的过程中，这些表面缺陷会使设计良好的力学规则失效，更多的应力集中在局部，导致断轨，脱轨等严重后果，直接危害到列车内人员的生命安全。

因此，为了最大限度的杜绝这种影响铁路安全的隐患，我们有必要持久性地对钢轨表面状态进行监测，并及时的对出现问题的钢轨进行替换，消除安全隐患。同时反馈到钢轨生产线来进行进一步的完善和调整，以增强后续生产钢轨的抗磨损性、抗撞击性、刚度从而降低钢轨表面缺陷发生的频率。

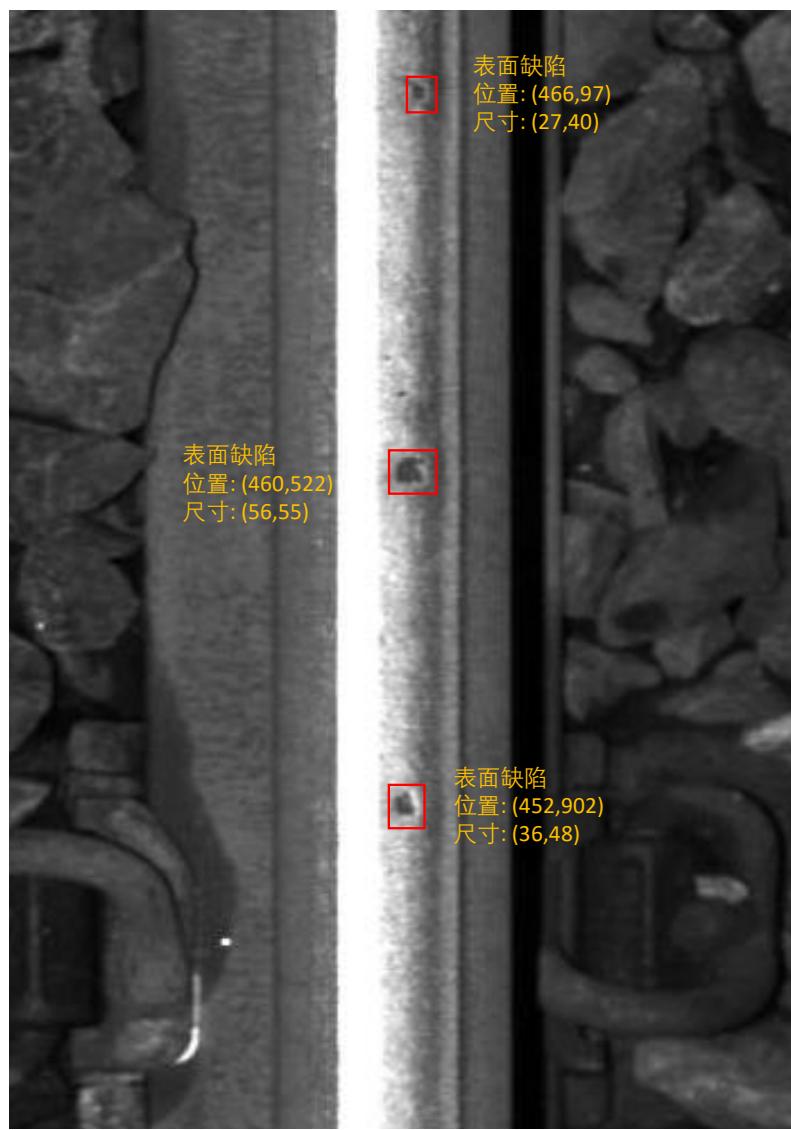


图 1-2 钢轨表面缺陷检测任务示例

Figure 1-2 A sketch map of rail surface's defects detection

钢轨表面缺陷检测指的是对钢轨上存在的缺陷进行识别和定位，图 1-2 展示了钢轨表面缺陷检测任务的一个示例。在过去，主要通过工人巡检来搜寻定位钢轨缺陷。这种方法不仅低效而且对工人的安全性存在着很大的威胁。随着铁路运营长度的大幅度增长，利用人工来对铁路运营状态进行监控也已经变得不切实际。目前，人工检测主要作为辅助检测手段来对自动检测出的钢轨缺陷做进一步的修正与判定。而对于钢轨缺陷的智能化检测，也衍生了超声检测、涡流检测、视觉检测等一系列方法。而视觉检测由于其相比于其他方法的非侵入性和便利性更加受人们的青睐。然而传统的基于视觉的方法通常是手动化设计检测特征，这使得由其衍生的方法大多具有误检率高和鲁棒性低的弊端。

近些年来，随着深度学习理论的推进，大数据的积累以及 GPU 硬件计算的高速发展，基于神经网络的自动化特征提取方法越来越成熟，带来了视觉检测新一波的蓬勃发展。然而美中不足的是成熟的深度学习方法大多需要海量的数据来进行监督训练，这对于钢轨缺陷检测的场景来说，是很难满足的。这是因为良好的维护使得我们所需的病害钢轨图像数量极度稀少和单一，想用这些图像训练一个可用、高效同时又具有较高鲁棒性的模型显得捉襟见肘。因此，设计出一套既不需要大量的病害钢轨数据，又可以充分利用深度学习成果的钢轨表面缺陷检测系统具有着重大的意义。

1.2 论文的主要工作

基于神经网络的钢轨表面缺陷检测问题按照其问题性质可以归结为目标检测问题，即识别一副图像中的目标物体并标记出目标物体的位置和类别。近些年来，随着大数据浪潮的推进、硬件资源的井喷式发展，目标检测的性能也随着深度学习的演进不断进步和突破。目标检测方法按照是否需要区域推荐过程被划分为一阶段目标检测和两阶段目标检测。两阶段目标检测方法首先通过区域推荐过程获取一些粗略的目标框，然后进一步的对目标框的类别进行判定以及对坐标进行进一步的修正^{[2][3][4][5]}。一阶段的目标检测方法则直接对目标框位置和类别进行预测^{[6][7]}。一阶段目标检测方法密集的候选框选取策略使得其面临着严重的正负样本失衡问题，因此提出了 Focal Loss 来缓解这一情况^[8]。通常来说、两阶段的目标检测算法具有更高的准确率而一阶段目标检测算法具有更快的检测速度。近年来，也相继出现了许多新的工作。比如 Anchor Free 的检测算法 CenterNet^[9] 和将 Transformer 引入目标检测任务的 DETR^[10]。尽管这些目标检测算法在公开数据集比如 Pascal VOC^[11]、COCO^[12] 上取得了较为优秀的结果，也被拓展到了安全监控和自动驾驶相关领域进行实际应用。但是如前文所述，缺陷钢轨表面图像的匮乏严重制约了目标检

测方法在铁轨缺陷检测方面一展手脚。一些工作尝试利用目标检测模型进行钢轨表面缺陷检测的工作，尽管具有很好的探究作用，但是由于数据集数量的匮乏，他们的实验结果受到了很大的限制^{[13][14]}。

在本文中，我们分析了基于深度学习的钢轨表面缺陷检测方法中一些特定的问题，然后设计实现了一套新型的钢轨表面缺陷检测系统并取得了良好的检测效果。如前文所述，良好的维护使得病害钢轨数量稀缺，从而导致无法采集足够数量的病害钢轨图片来进行神经网络的训练。考虑到这个问题，本文从数据和模型两个层面来设计一个钢轨表面缺陷检测系统。

之前，课题组在小样本钢轨缺陷检测领域已经研究颇多，但总归仍需要少量的病害样本参与。因此，为了开拓课题组的研究边界以及获得更无限制的研究方法，我把目光放在了零病害样本钢轨缺陷检测之上。我设计了基于虚实结合的钢轨表面缺陷检测方法，可以充分利用数字生成的虚拟病害钢轨样本以及现实场景中采集的真实无病害钢轨样本来训练一个可用、高效、鲁棒的钢轨表面缺陷检测模型。在基于虚实结合的研究策略下，完全不需要任何病害样本参与到神经网络的训练当中，这种方法显然更加自由以及充满前景。

对于虚拟数据，本文设计实现了基于数字孪生相关技术的生成方案，在数字世界中构建带缺陷的钢轨 3D 模型，并在数字空间中进行钢轨表面图像的采集。通过这样的方法，即可以生成所需的病害钢轨图像。在生成过程中，本文采用了点云处理的方法使得 3D 缺陷钢轨模型的生成能够自动化、快速的进行。其次，提出了曲率反曲率的对偶操作解决了生成过程中遇到的“锯齿问题”。同时，利用标签映射方法让缺陷信息在生成过程中能够同步的生成标注文件。最终就可以自动化地、快速地获取大量的有标签的病害钢轨仿真图像，利用这些仿真数据集进行模型的训练。

然而，仿真图像和真实图像在视觉效果上存在着一定差异，说明两者数据分布存在不一致性。这使得在仿真图像上训练的模型在真实图像下的表现效果受到了影响。为了解决这个问题，本文提出了简单有效的模型优化思路，其可以引入无病害的正常钢轨样本参与到模型的训练当中，来提升神经网络模型的性能。在轨路交通领域，无病害的正常钢轨样本极易采集，这为研究提供了极大的便利和可行性。在研究过程中，本文提出了伪目标域的概念，定义分析了在钢轨缺陷检测这个特定场景下的迁移问题。并基于以上思路、提出了图像对齐和特征对齐两大模块，其可叠加到现有的神经网络模型之上有效提高检测效果。最终，生成了大量的虚拟病害钢轨图像作为训练集，大量的真实无病害钢轨图像作为辅助训练集，以及标记了足量的钢轨表面图像作为测试集来进行实验验证。

基于以上阐述，现将本文的贡献与创新点概括如下：

(1) 设计实现了一套全新的开创性的无需病害钢轨数据参与,却又可充分利用深度学习成果的钢轨表面缺陷检测系统。摆脱了在基于深度学习的钢轨缺陷检测方法中,病害钢轨图片数量不足所带来的掣肘。

(2) 提出了基于 3D 模型和点云处理方法的虚拟数据生成方案,可以自动化地生成大批的高质量病害钢轨表面图像。同时,在生成过程中,可以自动化输出训练神经网络所需的标签信息,大大降低了人工参与的成本。

(3) 提出了图像对齐和特征对齐两个模块,其可以嵌入到现有的大部分网络模型之中。通过这两个模块,可以将极易采集的真实无病害钢轨图像与虚拟数据一起引入到神经网络的训练当中,大幅度提高钢轨缺陷检测的识别效果。

1.3 各个章节介绍

本文共由以下五个章节组成:

第一章 绪论。本章节从我国内陆深广的地理特点出发,引出铁路运输在我国的重要地位。从国内铁轨长度、行驶里程、铁路事故入手表达了对交通安全问题的关切,同时引入交通安全预防问题。之后,自然过渡到对铁路运输中最重要部件钢轨的介绍,同时指出对钢轨表面缺陷进行检测是保障铁路运输安全关键的一点。其次,开始介绍钢轨表面缺陷检测的含义、任务以及目前面临的环境复杂、样本数量少、数据采集困难等多重挑战。最后,基于前文所面临的种种困难,介绍针对这些问题我们所做的工作。

第二章 相关研究工作和理论基础。本章节首先细述钢轨表面缺陷检测领域的传统检测方法,逐步引入到新兴的基于神经网络的钢轨表面缺陷检测思路,然后介绍了一些国内外商用的钢轨缺陷检测系统。其次,总结介绍了本文中需要用到的一些深度学习知识。

第三章 虚拟数据生成系统设计与实现。本章节首先对我们基于 3D 模型和点云处理方法的虚拟数据生成方法进行详细的叙述,同时与其它虚拟数据生成方法进行对比。然后,对虚拟数据生成过程中出现的“锯齿问题”加以分析,并提出曲率和反曲率的对偶操作来解决这个问题。最后,基于生成的虚拟数据来进行网络模型的训练,将实验结果与传统钢轨表面缺陷检测方法加以比较证明方法的有效性。

第四章 基于虚实结合的钢轨表面缺陷检测研究。本章节首先从虚拟病害钢轨图像和真实病害钢轨图像的视觉差异入手,分析其对模型性能的影响。其次借用迁移学习中的相关理论对问题进行形式化叙述,并从数据分布的角度来证明引入真实无病害钢轨图像参与神经网络训练的可能性。之后,对我们的图像对齐模块和特征对齐模块进行详尽的介绍,并对实验的内容加以展示和分析。

第五章 总结。本章节对本文的工作进行系统性、概括性的总结与展望。分析自己研究工作的亮点和不足之处，提出下一步可以继续进步的方向。

2 相关研究工作与理论基础

2.1 钢轨表面缺陷检测研究现状

2.1.1 非视觉钢轨表面缺陷检测方法概述

非视觉的钢轨表面缺陷检测思路，类别各异，方法众多，其包括超声检测、漏磁检测、涡流检测和红外线检测等。

超声检测系统是应用最广泛的无损检测方法之一。它利用超声波换能器将电信号转变为超声波信号，然后让超声波信号在钢轨中传播，通过观察超声波的变换

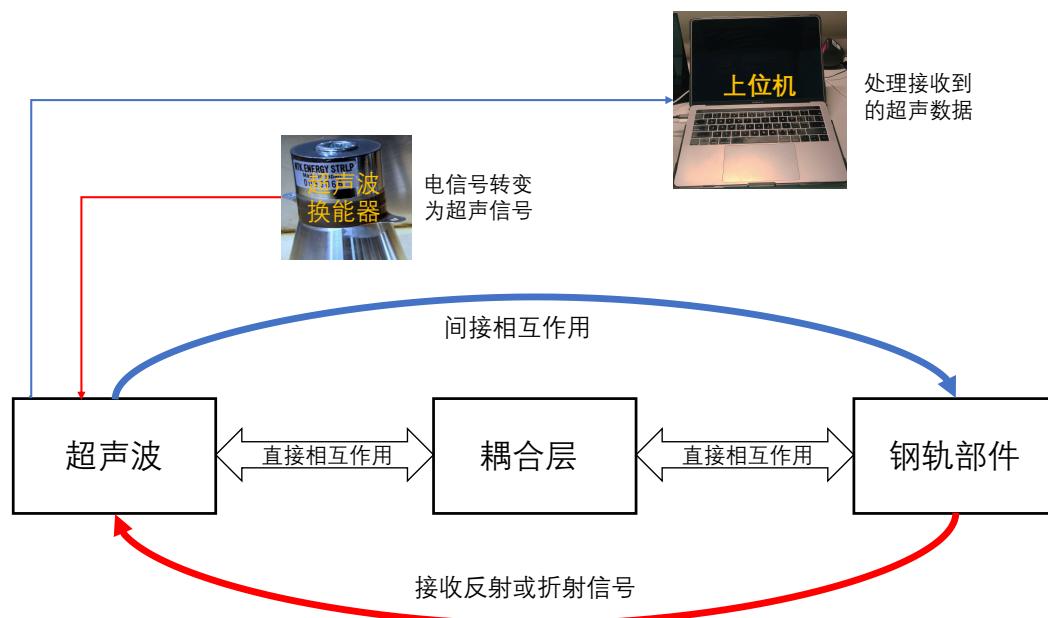


图 2-1 超声检测原理图

Figure 2-1 Schematic diagram of ultrasonic detection

状态来探测钢轨部件是否存在损伤。经过超声波换能器所转换的超声波在钢轨传导的过程中，如果遇到了钢轨缺陷，其就会因为缺陷处介质变化增大，受到的阻力骤变从而发生衰减、折射和反射现象。按照探伤原理，超声波检测方法可以被分成共振法、脉冲透射法和脉冲反射法^[15]。共振法主要被用来计算工件的厚度而非钢轨探伤，其原理是当工件厚度等于超声波的半波长整数倍时会发生共振效应。脉冲透射法则是依据脉冲波或连续波穿透工件之后的能量衰减变化，来分析缺陷信息，其需要将接收探头置于被检工件的两侧。而脉冲反射法是最常用的一种超声探伤方法，它可以只使用一个超声探头来兼顾超声的发射和接收。其通过分析探头接收不

同反射信号的时间差，可以判断出钢轨是否存在缺陷。同时，根据回波信号的其他属性，比如位置，高度我们就可以粗略判断出钢轨缺陷的大小和位置信息。实际的超声检测系统在使用时，为了能够获得较大的检测范围，通常会对发射超声波的角度进行相应的规定。钢轨的轨底、轨腰和轨头分别需要从不同的角度入射。同时，为了使得声能达到更好的传输状态，在钢轨和超声探头之间需要使用耦合液体。

得益于超声检测检出率高、裂纹定位精准和实时性强的优点，一直以来都有许多研究者在超声钢轨检测领域进行研究。2005 年，一些工作基于电磁超声波，使用无监督的方式对钢轨裂纹进行了检测^[16]。2014 年又出现了基于超声换能器所设计的可以准确判别钢轨裂纹的钢轨光声成像系统^[17]。2017 年，华南理工大学的李骏等人基于时频空来对超声波的检测信号进行分析^[18]。在时空中，利用自相关方法对多组检测信号进行特征提取。在频空内则利用小波变换对信号能量分布进行建模并用其作为特征。最后，将综合获取到的超声特征输入到支持向量机之中来进行缺陷检测。2018 年，南昌航空大学的时亚等人则是基于电磁超声检测的有关理论，建立合适的有限元模型来对超声波换能器表面的激励表面波进行优化设计，从而实现对钢轨踏面的 RCF 裂纹的有效检测^[19]。同时，目前也已经有许多成熟的基于超声波的钢轨缺陷检测产品问世，在后面我们会统一介绍。

漏磁检测的机理如图 2-2 所示。钢轨试件在磁铁的影响之下发生了磁化。在钢轨状态良好的时候，其内部的磁力线保持稳定。当钢轨试件出现了缺陷病害，因为磁导率发生了改变，围绕在缺陷周围的磁力线就会从钢轨表面溢出形成漏磁。对泄漏的磁力线的变化情况进行研究和分析，就可以获取到缺陷的位置、大小、形状等各类信息。漏磁检测的原理较为简单，但却具有较强的通用性，并且实现起来也较为容易，因此在除了钢轨缺陷之外的许多其它工业无损检测领域，它也被广泛应用。不过，漏磁检测由于其自身的特征，难以识别出较窄的缺陷，即使钢轨缺陷的长度和深度都已经完全达到了病害的标准，检测的信号也会由于太过微弱而难以探查。

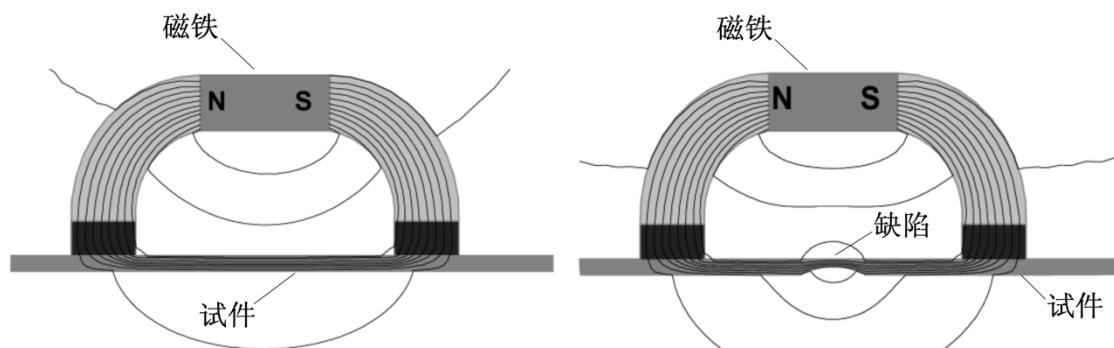


图 2-2 漏磁检测原理图^[26]

Figure 2-2 Schematic diagram of magnetic flux leakage detection

尽管漏磁检测近年来在钢轨缺陷识别领域存在一些限制，但是对于其的研究和改进仍受到许多研究人员的青睐。2006 年，一些研究人员对钢轨裂纹的三维漏磁可视化成像方案进行了研究^[20]，同年，一些工作又研究了漏磁检测中的磁力线在高速运动下的分布情况^[21]。2011 年，研究人员基于有限元分析理论的，成功检测出了斜裂纹^[22]，同年有人通过完善漏磁检测原理，设计出了在当时具有较高检测精度的钢轨漏磁检测方案^[23]。次年，研究人员通过有限元分析的相关软件，对漏磁场进行了数字建模，从而得到了较为准确的磁通密度曲线^[24]。对漏磁信号的特征值进行分析比较，即可了解钢轨是否存在裂纹和裂纹的其它参数信息。2021 年，又有人提出了增强的漏磁信号检测方法，此方法可以检测钢轨上产生的细小裂纹^[25]。

涡流检测包括脉冲涡流检测、常规涡流检测、远场涡流检测以及多频涡流检测。涡流检测的原理如图 2-3 所示。检测探头采用脉冲电流进行激励，而输入的电流通过变化会在空间中产生变化的磁场。变化的磁场即会在金属试件上产生涡流。通过霍尔元件，我们就可以通过检测输入电流和涡流产生的叠加磁场来判断此位置是否存在缺陷。

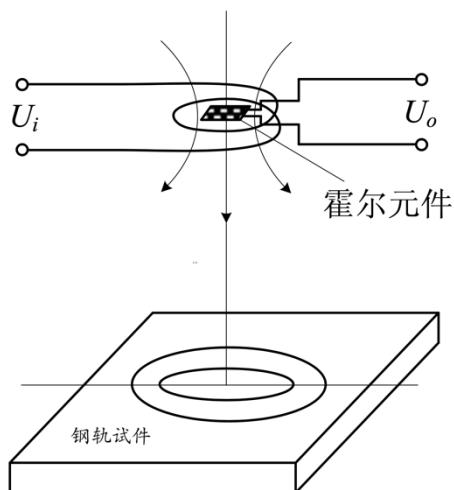


图 2-3 涡流检测原理图^[26]

Figure 2-3 Schematic diagram of eddy current detection

由集肤效应，我们可以知道激励信号的频率越大，钢轨材质的电导率和磁导率越大，产生的涡流就越趋近于钢轨表面，这就使得涡流检测对于检测钢轨表面缺陷具有更高的灵敏度。如果想要检测钢轨深处的缺陷，就需要降低激励电流的频率，但这会对涡流检测的灵敏度产生非常大的影响。综合来看，涡流检测方法在现有的技术条件下仅能对钢轨表面的裂纹缺陷进行检测，并且误报率极高。因此，在大多数情况下，这种方法都会作为钢轨表面缺陷检测的补充方法。

基于涡流检测技术的钢轨缺陷探测方案也有着比较长的历史。2007 年，一些研究人员设计了基于涡流检测技术的钢轨裂纹缺陷探测设备，但是在速度和准确

度上都有所欠缺^[27]。2011 年, 李国厚等人采用减聚类算法改进了径向基函数, 并使用了基于巨磁阻传感器的检测探头, 大幅度提高了涡流检测对钢轨表面微小缺陷和深层缺陷的探测水平^[28]。2013 年, 有人将震动加速度的理论知识引入到涡流检测当中, 使得其设备可以检测出裂纹类型以及裂纹的深度^[29]。2018 年, 浙江大学的彭谢丹^[30]设计出了一套轻型的简易的基于脉冲涡流检测的钢轨裂纹探测系统系统, 可单纯地判断钢轨是否存在裂纹。

红外线检测也是一种无损的钢轨表面缺陷检测方法。其主要分为持续加热法、超声热像法以及脉冲热像法。持续加热法是一种通过利用昼夜温差导致的钢轨表面温度变化差异来识别出缺陷的方法, 它极易受外界环境影响而且检测时间长。超声热像法则是利用大功率的超声波作为外部热源, 超声波的能量在经过缺陷处时, 会发生衰减并在钢轨缺陷附近转化为热能, 因此, 我们就可以通过红外相机探测钢轨上的异常温度来识别出缺陷的位置和大小。然而这种方法由于在实用化上还不完善, 难以达到现场检测所需要的水平。脉冲热像法的检测原理可以由图 2-4 解释, 当激励线圈靠近待检测钢轨的表面时, 钢轨表面会感应出相同频率的涡流。如

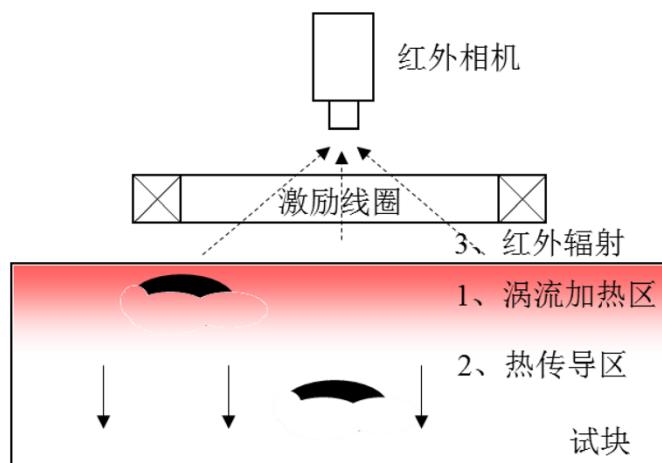


图 2-4 脉冲热像法示意图^[65]

Figure 2-4 Schematic diagram of pulse thermography

果钢轨表面存在裂纹, 钢轨表面连续性就会被破坏, 使得涡流的分布受到扰动, 从而使得由涡流导致的焦耳热呈现一定规律的分布。那么, 我们通过红外相机将钢轨表面温度进行记录下来之后, 就可通过比较正常焦耳热与异常焦耳热分布之间的差异来检测缺陷。同时, 我们还可以对缺陷周围的焦耳热分布进行量化分析来获得诸如大小、形状等更为详细的钢轨缺陷数据。脉冲热像法近些年也得到了充分的发展^{[31][32][33]}并且在很多领域都已经得到了成功的应用^[34]。

2.1.2 视觉钢轨表面缺陷检测方法概述

视觉钢轨病害检测技术相比于上述的非数据检测方案，更加的安全、快速。同时布置简易、价格低廉，在硬件设备上我们仅仅需要一个图像采集设备。在我们对需要探测的钢轨位置拍照之后，即可对所得图片进行算法分析，从而判断拍照位置的钢轨状态是否正常，如果存在病害，则识别出病害的位置。视觉钢轨病害检测技术目前可以大致分为传统图像检测方法和基于神经网络的深度学习检测方法。

(1) 传统图像检测方法

早在 2004 年，就有人通过扫描相机采集图像，并采用光谱图像差分来探测钢轨上是否存在裂纹缺陷^[35]。2009 年，官鑫等人对图像处理技术在钢轨病害检测中的应用进行了分类和理论分析，同时设计了裂纹自动判别分类算法^[36]。2011 年，王海涛等人设计了一套基于机器视觉的软硬件平台，其可探测出钢轨裂纹的位置

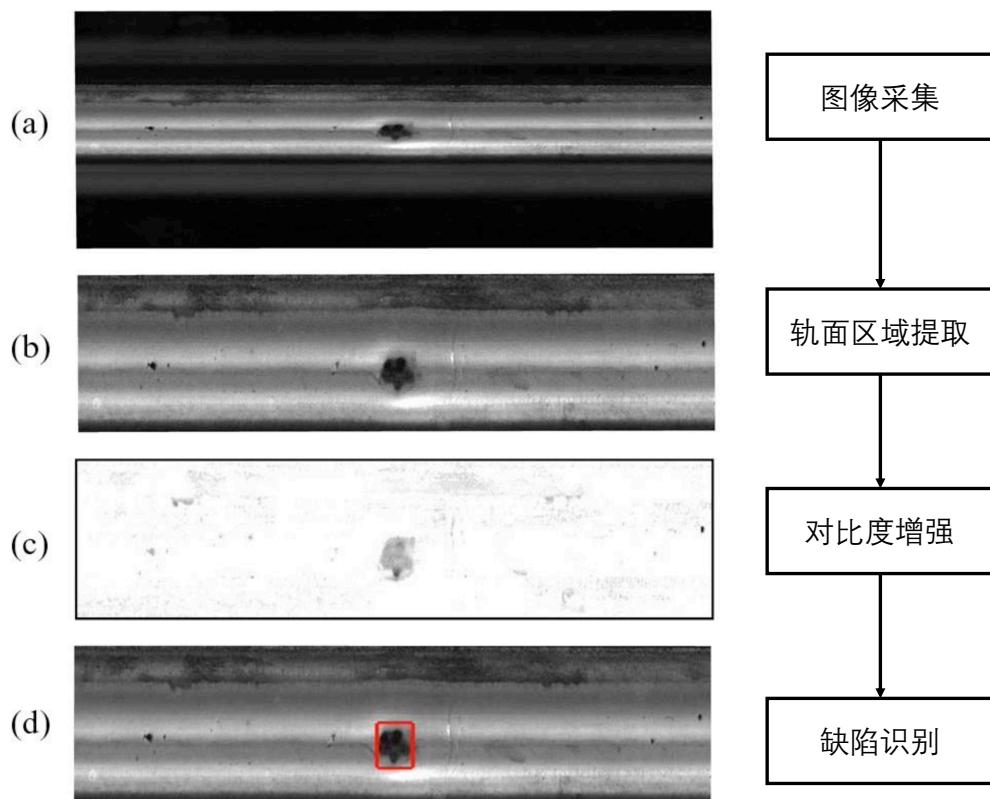


图 2-5 基于局部归一化和投影轮廓的缺陷检测流程图^[38]

Figure 2-5 Defect detection flow chart based on local normalization and projection profile

和大小^[37]。2012 年，北京交通大学的李清勇等人提出了基于局部标准化来对钢轨进行对比度增强、使用投影轮廓来提取缺陷信息的传统视觉方法，图 2-5 展示了通过此方法提取缺陷的整体步骤和效果^[38]。2015 年，王时丽等人利用线阵相机采集高精度钢轨图片，同时采集钢轨表面的深度信息，设计了三维处理算法来检测裂纹的位置、大小和形状^[39]。2016 年，一些研究人员通过使用 Gabor 滤波器来提取钢轨表面特征，比较钢轨缺陷区域和非缺陷区域的特征差异，从而实现对缺陷区域

的探测^[40]。在 2020 年，又有人提出了一种新型的形态学提取缺陷方法，其可同时保证检测准确率和检测速度^[42]。2021 年，何恬提出了一种边缘算子加权的引导滤波分层图像增强算法，其可以突出钢轨缺陷部分提高后续缺陷检测的准确度^[41]。

(2) 深度学习检测方法

相比于手工设计的特征，深度神经网络提取的特征具有更强的鲁棒性和更高的有效性。因此，近些年来随着计算硬件的高速发展和人工智能理论的不断推进，基于深度神经网络的钢轨缺陷检测方法越来越流行。2017 年，哈尔滨工业大学的张鑫等人提出了一种基于卷积神经网络以及事件概率分析的钢轨状态检测方法，利用 CNN 网络来判断钢轨是出于安全状态还是非安全状态^[43]。2018 年和 2020 年，一些研究人员基于 YOLOv3 目标检测模型，研究设计了定制化的钢轨表面缺陷检测方法^{[13][14]}。其中，后者设计了多目标的铁路安全检测系统，不仅能检测钢轨缺陷，而且可以对其他重要组件比如扣件同时检测。然而，这两个工作由于病害钢轨样本的匮乏，经过训练集和测试集划分之后，测试集中病害样本数量过于稀少。这使得实验结果的权威性受到了些许影响。2020 年，长沙理工大学的宋雅男在学位论文中研究了基于卷积神经网络的多模型融合方法，将 SSD 和 YOLOv3 两种检测算法并连在一起，在不降低检测速度的同时提高了钢轨缺陷识别的精度^[44]。2021 年，研究人员提出了一种改进的轻量级实例分割网络，用于扣件和钢轨的病害识别。同时，其也提出了一种通过分析紧固件掩模的几何特征来检测紧固件缺陷的方法，来克服病害钢轨样本缺乏的问题^[45]。

2.1.3 国内外商用钢轨缺陷检测系统

国外在钢轨缺陷检测商业系统的开发上起步较早。早在 1964 年，日本的东海道铁路公司就开发了第一辆简易自动轨道检测车。该检测车一经研制成功，就克服了以往采用人工拉线测量的不足之处，大大提高了日本在轨道检测方面的效率。1998 年，该公司又在 700 系列的基础上进行了进一步的研制和改进，将检测速度提高到了 270km/h，同时提高了检测精度。这样，从东京到博多接近 1100km 的铁路线路可以在两天内完成全部的检测^[46]。

图 2-6 国外商用轨道检测系统^[47]

Figure 2-6 Foreign commercial track detection system

2001 年，美国的 ENSCO 公司研发出了 VIS 轨道视频综合检测车，其产品样态如图 2-6（a）所示。该检测车在行进过程中会利用相机采集钢轨各部件的图像信息并传输到数据处理中心。在终端服务器上，其会利用识别和定位算法对图片进行扫描来自动判断螺纹扣件是否缺少或者移位等故障。2011 年，位于意大利的 MERMECTEC 公司研制开发了高速多功能检测车，如图 2-6（b）所示，其可以对铁路的轴箱加速度、轨道几何、波浪磨耗和钢轨断面进行非接触式的检测^[47]。其中，对于钢轨尺寸、磨损以及裂缝的检测，其利用的是传统的光学激光原理和超声检测技术。而对于导电弓与接触网的检测，则采用的是图像处理技术^[48]。在 2013 年，法国的 IRIS320 轨道几何检测车，也进行了新一轮的更新升级^[47]。除此之外，德国的 Atlas Elektronik 公司研发了光电式轨道检测系统，其有效的将现代图像处理技术运用到了轨道缺陷检测的任务当中，实现了轨距测量、钢轨变形与磨耗等故障的高效探测^[48]。

国内的轨道检测系统一般是科研单位与国有企业等联合研制，到目前为止也已经基本具备完整的雏形。80 年代初，我国自行设计了 GJ-3 型轨道巡检车，不过该类型轨检车使用的元件仍停留在 100 多年的分离式技术，这使得 GJ-3 检测系统存在着各式各样的问题，比如误差大、集成度低、系统稳定性差等。90 年代之后，GJ-4 型巡检车被研发出来，但是由于并未解决好抖动问题使得在速度高于 120km/h 之后，检测性能大幅度下降。GJ-5 型轨检车是我国在 21 世纪初从美国 ImageMap 公司引进的轨检车型。其不仅应用了惯性基准测量原理，而且采用了当时最先进的激光摄影测量和网络数据库技术。这使得 GJ-5 型轨检车不仅可以测量一些常规的轨道状态项目，还可以测量轨向波长、曲率变化率和轨距变化率等一些额外信息。这些额外项目可以用来评估列车在高速行驶时的舒适度指标和厢内安全性能^[48]。GJ-6 型是我国最新一代的轨检车系列，其更加高效、精确和可靠。解决了高速实时处理、高速图像处理以及高速网络处理的难点，检测速度最高可达到 400km/h。

图 2-7 国内轨道检测系统^[47]

Figure 2-7 Domestic track detection system

同时其无移动部件，悬挂装置经过了精密的应力分析和优化改良，系统经过了长时间的改进，更加的可靠。到目前为止，我国在各个检测部门还在服役的轨检车有 GJ-3 型、GJ-4 型、GJ-4G 型、GJ-5 型以及 GJ-6 型。

2.2 相关理论基础

2.2.1 卷积神经网络

卷积神经网络目前已经被充分的应用到了基于深度学习的视觉检测任务中，几乎所有的视觉任务都会将卷积神经网络视为自动特征提取的一种黄金方法，然而其在被发明的 14 年间，却一直默默无闻。1998 年，LeCun 等人开发了一种新型的支票识别系统，其中使用了一种新型的卷积神经网络架构 LeNet^[49]。然而，受限于当时的硬件资源，神经网络模型难以在大规模的数据集上进行训练以及测试，导致了这种自动化的提取方式在当时并未受到重视。直到 2012 年，Hinton 领导的团队在大规模视觉识别挑战赛 ILSVRC (ImageNet Large Scale Visual Recognition Challenge) 中一举夺冠，并在 ImageNet 数据集 top-5 错误率上超出了亚军 10 个百分点^[50]。而其中最亮眼，作用最大的部件即是多年未被重视的卷积层。自此以后，卷积神经网络开始迎来了高速发展，也带来了人工智能研究的浪潮。

LeNet5 一共有三个卷积层和一个全连接层，是最早的卷积神经网络。其通过参数共享和池化，不仅避免了大量的计算成本，而且增加了网络的泛化性，让利用神经网络检测图像成为了可能。到了 2012 年，让神经网络重新焕发活力的 AlexNet 在 LeNet 的基础上做了进一步的改良^[51]。AlexNet 有 5 个卷积层和三个全连接层，在网络层数上做的更深了，初步说明了深层网络检测的有效性，为后来的研究者提供了研究方向。其次，它使用 Relu 这种非线性激活函数替代了 tanh 和 sigmoid 激活函数，没有饱和区，具有更快的训练速度。同时，Relu 的负半轴为 0，可以使得网络具有稀疏性，有助于缓解过拟合问题。AlexNet 也提出了 Dropout 策略，在每

轮的训练过程中，随机的将一些神经元的值设置为 0，这使得每轮训练的模型都是一个新的模型，最终的结果是多个结果的一个平均，提高了模型过拟合的能力。

2014 年，ILSVRC 的亚军 VGG 为我们带来了神经网络中新的启发。VGG 由牛津大学的 Visual Geometry Group 团队设计，其提出利用多个 3×3 的小卷积核来替代 AlexNet 中 11×11 、 7×7 和 5×5 的大卷积核。通过替换，在感受野不变性能不下降的情况下大大降低参数量和卷积运算的计算资源消耗^[52]。其中，3 层 3×3 卷积操作带来的感受野和一层 7×7 是一样的，因此我们可以尝试计算一下参数量和计算量差异有多大。我们将输入和输出的特征图高、宽、通道数分别用 H 、 W 和 C 表示。为了达到相同的感受野，使用 3×3 卷积核所需参数量是 $3 \times 3 \times C \times C = 27C^2$ ，计算量是 $3 \times H \times W \times 3 \times 3 \times C \times C = 27HWC^2$ 。而使用 7×7 的卷积核参数量则为 $7 \times 7 \times C \times C = 49C^2$ ，计算量为 $H \times W \times 7 \times 7 \times C \times C = 49HWC^2$ 。可以看到，使用小尺寸的卷积核所需的参数量和计算量降低了近一半，并且每一层卷积层后面都会接上一个非线性的激活函数，所以层数更多的小卷积核还提高了模型的非线性表示能力。

硬件资源水平的提高和各种计算复杂度优化方法的提出，让神经网络的层数越来越多，网络越来越深。这暴露出了深度神经网络新的问题，那即是当深度达到一定程度的时候，各项指标提升开始减缓甚至出现指标下降的反常识现象。之后这个问题一直困扰着当时的研究人员，甚至有人觉得神经网络的繁荣即将结束。然而 2015 年，大名鼎鼎的 ResNet 横空出世，并拿下了这一年的 ILSVRC 冠军。其彻底分析了这个问题的根源并提出了残差模块的研究思路解决了网络过深导致的性能衰退问题，随之 ResNet 的深度也一举超过了一百层^[53]。

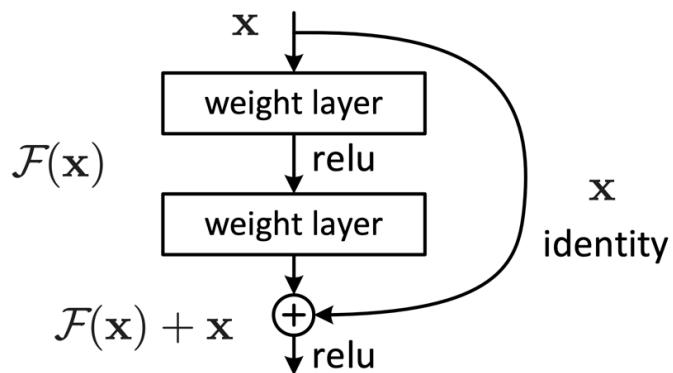


图 2-8 近路连接模块^[53]

Figure 2-8 Residual learning: a building block

ResNet 提出的残差网络最重要的一点即是引入的“近路连接”模块。如图 2-8 所示，“近路连接”模块与以前的模型存在很大的不同。以前的网络大多都是一层的输入经过处理后，再输入到下一层，这使得下一层的输入都是上一层经过处理之后的输出结果。而残差网络的“近路连接”则是在输入到下一层网络的时候，将原本的输入和经过网络层处理后的结果一起输出到下一层，这使得模型中较深网

络层的梯度可以通过旁路 x 来反向传播，解决了网络过深导致的梯度消失问题。从另一个角度看，“近路连接”模块学习到的 $F(x)$ 是潜在映射的残差，而非潜在映射，这使得 ResNet 的性能更加的优秀。

在神经网络高速发展的同时，各式各样的移动设备也开始步入万家。而 VGG 和 ResNet 这种部署在服务端的重型网络架构，由于参数量过大，很难部署在内存空间有限的移动端。因此，大量的研究人员开始设计各式各样的轻量级网络和探索模型压缩技术。在众多轻量化网络中，比较知名和高效的有谷歌公司的 MobileNet 系列和旷世公司（Face++）的 ShuffleNet 系列。其中 MobileNet 系列包括 V1^[54]，V2^[55]和 V3^[56]版本。而 ShuffleNet 系列包括 V1^[57]和 V2^[58]版本。这里我们简单介绍一下 MobileNet V1。MobileNet V1 由谷歌公司提出，其核心是深度可分离卷积。如图 2-9 所示，深度可分离卷积指的是将原本的标准卷积运算分解为深度卷积（逐通道卷积）和逐点卷积。通过这样的操作，可以大幅度降低计算量和参数量。深度卷积与普通卷积的区别在于一个卷积核仅对一个通道起作用，而非作用于所有通

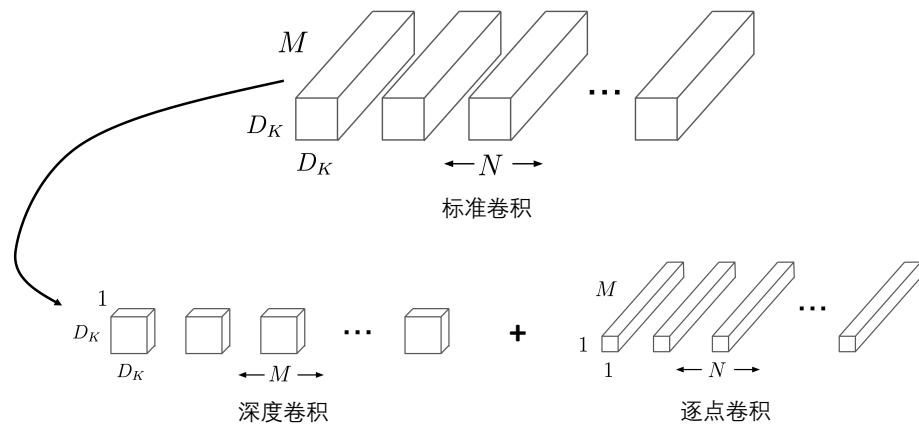


图 2-9 深度可分离卷积^[54]

Figure 2-9 Depthwise Separable Convolution

道。因此，深度卷积在完成卷积操作后，通道数量不会发生改变，并且没有完成对通道之间的融合。因此，在通道卷积的后面又设计了逐点卷积，逐点卷积采用的是尺寸为 $1*1*M$ 的卷积核，这里的 M 指的就是深度卷积之后得到的通道数也即输入深度卷积之前的通道数。通过 N 个这样的 $1*1*M$ 卷积核，就可以实现通道之间的融合和通道数量的改变。下面我们简单比较一下标准卷积分解之后能节省多少的计算量和参数量。我们令 H 和 W 表示输入特征图的高和宽。输入输出的通道数分别用 M 和 N 表示。标准卷积、深度卷积和逐点卷积的参数用图 2-9 所示的字母代号表示。那么可以计算出标准卷积、深度卷积和逐点卷积的参数量与计算量，结果如表 2-1 所示。

表 2-1 各类型卷积操作参数量和计算量统计

Table 2-1 The number of parameters and computation of each type of convolution operation

卷积类型	参数量	计算量
标准卷积	$D_k^2 MN$	$HWD_k^2 MN$
深度卷积	$D_k^2 M$	$HWD_k^2 M$
逐点卷积	MN	$HWMN$

因此，我们可以得到这样的结论，得到深度可分离卷积的参数量和计算量都是标准卷积的 $1/N + 1/D_k^2$ 倍，因此，深度可分离卷积对于构建轻量化网络架构是一个有效的选择。

2.2.2 目标检测

在 1.2 小节中，我们已经对基于深度学习的目标检测的历史和分类进行了介绍。第四章节我们会用到 YOLOv3 作为基底模型，因此本章节我们主要对 YOLO 系列进行叙述。YOLO 系列是一个一阶段的目标检测模型，发展至今已经包括了 V1^[59]、V2^[60]、V3^[6]、V4^[61] 和 V5^[62] 五个版本。其中，前三个版本一般被视为 YOLO 系列的正统，后续的 V4 和 V5 则属于民间优化版本。这也是第四章节中我们选择 YOLOv3 作为我们基底模型验证模块有效性的一个原因。

2016 年，一种不同于 RCNN 系列的新型目标检测算法横空出世，它不再需要先用 RPN（区域推荐网络）初筛可能性目标，再用分类器去判断检测。取而代之的是它将检测问题看成一个回归问题，利用单个神经网络直接对整个图片进行处理，获取目标边界框以及类别概率。因为整个检测流程是单一网络，所以可以很容易的进

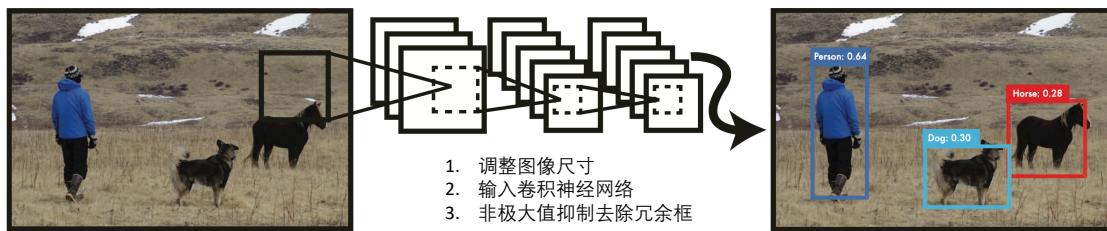
图 2-10 YOLO 检测系统^[59]

Figure 2-10 YOLO Detection System

行端到端的性能优化^[59]。因此，这种算法被作者命名为 YOLO，含义是你只需要看一次（You Only Look Once），实至名归。图 2-10 是 YOLO 的检测系统示意图，可以看出整个检测流程非常简洁。系统的精简带来了检测速度上的飞跃，可以说 YOLO

的出现在当时让实时目标检测正式进入了人们的视野，也让目标检测在工业上的应用大大成为了可能。

YOLO 的 V2 版本称作 YOLO9000，可以检测超过 9000 个类别。作为一个追求实用的网络架构，YOLO 系列的许多新特性都是借鉴的当年深度学习领域比较有效的模块，YOLO9000 也不例外。首先第一个借鉴的是当年比较火热的批标准化操作（Batch Normalization），这个操作主要是对每一批训练数据，统计它们的均值和方差，对数据进行标准化处理。这是一个“白化操作”，其可以使得数据分布保持一定的稳定，从而加快网络的训练和收敛速度、防止梯度爆炸和梯度消失以及有效减弱网络过拟合的风险^[63]。其次，YOLOv2 也借鉴了 Faster R-CNN 基于锚点框（Anchor Box）的预测机制，同时进行了改进。普通的锚点框的大小和比例一般是手动设计的，而 YOLO9000 采用了 K-means 聚类方法来决定锚点框的大小和比例，使得设计的锚点框更加合理。另外，普通的锚点框是基于绝对位置的偏移，使得目标框会偏移到图片的任何地方，YOLO9000 采用了直接预测相对位置的策略，直接回归目标框相对于网格单元左上角的相对坐标，避免了这个问题。除此之外，YOLO9000 也做了很多其它的改进，比如为了 YOLO9000 设计了自己的特征提取网络 Darknet-19，其包括 19 个卷积层和 5 个最大池化层。Darknet-19 只需要 55.8 亿次操作就可以处理一副图像，并且在 ImageNet 上达到了 72.9% 的 Top-1 精度和 91.2% 的 Top-5 精度。

Type	Filters	Size/Stride	Output
Convolutional	32	3×3	224×224
Maxpool		$2 \times 2/2$	112×112
Convolutional	64	3×3	112×112
Maxpool		$2 \times 2/2$	56×56
Convolutional	128	3×3	56×56
Convolutional	64	1×1	56×56
Convolutional	128	3×3	56×56
Maxpool		$2 \times 2/2$	28×28
Convolutional	256	3×3	28×28
Convolutional	128	1×1	28×28
Convolutional	256	3×3	28×28
Maxpool		$2 \times 2/2$	14×14
Convolutional	512	3×3	14×14
Convolutional	256	1×1	14×14
Convolutional	512	3×3	14×14
Convolutional	256	1×1	14×14
Convolutional	512	3×3	14×14
Maxpool		$2 \times 2/2$	7×7
Convolutional	1024	3×3	7×7
Convolutional	512	1×1	7×7
Convolutional	1024	3×3	7×7
Convolutional	512	1×1	7×7
Convolutional	1024	3×3	7×7
Convolutional	1000	1×1	7×7
Avgpool		Global	1000
Softmax			

图 2-11 Darknet-19 特征提取网络^[60]

Figure 2-11 Darknet-19 Feature Extraction Network

到了 2018 年，YOLO 系列的正统关门之作 YOLOv3 正式出炉，带来了更快的速度和更高的性能。YOLOv3 最大的亮点即是重新设计的 Darknet-53 网络架构以及基于 Darknet-53 架构所设计的多尺度目标检测机制。

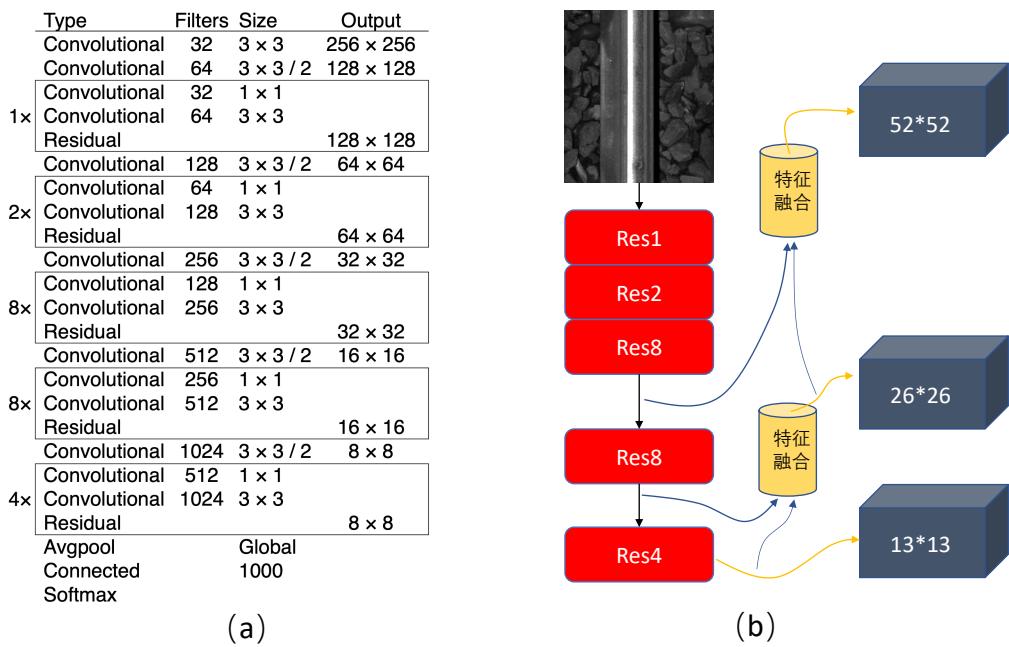


图 2-12 YOLOv3 检测网络

Figure 2-12 YOLOv3

DarkNet-53 的网络架构如图 2-12 (a) 所示，其一共包含 53 个卷积层，其中每个方框框起来的组合操作我们称其为一个残差单元。其综合了 YOLO9000 中用到的 Darknet-19 和残差网络 ResNet 的优点，获得了更为优异的特征提取效果。可以看到，整个 YOLOv3 的网络架构中，完全剔除了常用的池化层和全连接层。因此，特征图大小的变化是采用卷积层调整采样步长实现的。YOLOv3 的多尺度检测机制如图 2-12 (b) 所示，输出了三个不同大小的特征图作为检测结果。其对原始图像有着不同的感受野，这使得模型可以提高对不同尺度大小目标的检测能力。同时，YOLOv3 还借鉴了 FPN 网络的多尺度融合方法^[64]，将 Darknet 的中间层进行上采样，使其与之前某一层输出特征图空间尺寸相同，然后采用 Concat 操作进行两者的拼接。这样，高层特征与底层特征就有了自顶向下的融合，充分利用低层语义信息和高级语义信息获得了更好的检测效果。

YOLOv3 之后，其作者担忧自己的研究成果会被用于军事用途，公开表示之后将不再进行 YOLO 系列模型的更新迭代，自此开始出现一些民间版本。其中，比较知名的有 YOLOv4 和 YOLOv5 网络，两者出现的时间比较接近，并都更偏向于集合已有的各式各样人工智能有效组件组合成一个更快速、准确、鲁棒的通用性目标检

测模型，网络结构也越来越复杂。其中，相比于 YOLOv4, YOLOv5 更加受工业界的青睐。

2.2.3 域适应

在目标检测任务中，我们需要保证训练样本的数据分布和测试集的分布一致，这也是保证机器学习有效的基本假设。因为我们没办法期待一个通过狗的图像来训练的神经网络，能够有效地去检测出图片中的猫。然而，在训练集和测试集分布不一致的情况下训练模型的需求却一直存在，我们希望一个神经网络能够将在某个领域学到的能力和知识能够通过一定的方法来部分或者全部迁移到不同但相关的任务当中。比如，我有一些品种的带训练标签的狗的照片，我希望我训练的网络模型能够检测到狗的种类要多于训练集中所包含的狗的种类。或者我有一些在白天场景下汽车检测的带训练标签的样本，我希望通过这些样本训练出的检测模型也能够有效地检测出深夜场景下的汽车。

在这种需求的催动下，域适应的方法被提了出来。域适应属于迁移学习中比较重要的一个分支，它主要被用来解决域偏移的问题^[66]。所谓域偏移问题，也即是我们的训练集和测试集在数据分布上整体相似但是存在一定的偏差。在域适应领域中，训练集所代表的数据分布域被称为源域，而测试集的数据分布域被称作目标域。那么域适应想要做的即是通过各种处理方法、使源域和目标域数据分布趋于相似，使得在源域上训练的模型可以在目标域同样生效。早期，许多方法尝试着采用一些统计量来表征隐层空间的分布，然后使这些统计量趋于一致^{[67][68][69]}。而随着域适应理论的发展^{[70][71]}，许多方法开始改用域判别器来衡量两个域之间的分布差异，域适应的能力水平也得到了不断地发展^{[72][73][74]}。

3 虚拟数据生成系统设计与实现

3.1 引言

如前文所述，我国在铁路方面发展迅速，到 2021 年铁路总运行长度已经有 15 万公里左右。随着铁路运营线路的不断增长，保障铁路运行安全的挑战也越来越大，对于安全状态的检测也愈加重要。而在整个监测的过程当中，对钢轨的状态维护是重中之重。得益于神经网络提取特征的自动高效，利用深度学习方法来检测钢轨缺陷近些年来成为了新的热点。然而，在实际场景中，我们往往难以获得足量的病害钢轨来训练一个鲁棒有效的神经网络模型。因此，使用虚拟数据来进行神经网络的训练是一个非常值得我们研究的课题。

对于生成仿真数据技术的研究和探索，其它领域已经有许多研究人员做出了尝试，但是在铁路领域的研究还相对较少。在 2014 年，黄茜等人提出了铸件的枝状缩松缺陷仿真图像生成系统，基于直方图匹配思路来生成仿真数据集。但是此方法使用传统的图像处理方法来生成缺陷，局限在二维图像和手工处理之上，获得的仿真数据质量不佳^[75]。在 2019 年，来自浙江大学的伊国栋等人通过使用深度学习中的生成对抗网络模型（GAN）生成了一批仿真数据，并将其用于注塑制品缺陷定位识别的研究当中^[76]。东京大学的研究人员也利用 GAN 来生成仿真的脑部 MR 图像协助进行相关的科学实验^[77]。基于 GAN 的生成方法尽管在真实性上更加具有优势，但它们存在着一些共同的弊端：

- (1) 在训练网络的过程中，我们需要利用许多的真实数据，而这与我们真实数据匮乏的是相冲突的。
- (2) 通过网络生成的数据，其缺陷位置是随机的，我们无法获取精确的定位信息，后期需要人为的标注，增加人力成本。
- (3) 通过网络生成的数据，缺陷的类型也是随机的，可能会出现我们不想要的会干扰网络训练的“负样本”。

除此之外对于我们钢轨表面缺陷的生成，还有一个很致命的缺陷。我们的钢轨缺陷相对于整幅图像来说，往往占比较小。采用网络生成的过程中，其会忽略掉这些微小的缺陷，仅仅生成那些具有较大缺陷的仿真图像，这对我们后期网络训练的影响是巨大的。因此，针对我们的场景来设计一种新型的虚拟数据生成方案不仅具有开创性的意义，而且也在实际场景中能够切实的解决行业痛点。

为验证我们虚拟数据生成系统的可用性，我们选用了两种神经网络模型来基于虚拟数据进行训练，将结果与几种传统的无监督钢轨缺陷检测方法进行比较。结

果表明虚拟数据训练的网络模型可以超过部分传统方法，说明了基于虚拟数据来训练一个钢轨缺陷检测模型具有很大的挖掘潜力，也为我们的第四章奠定了基础。

3.2 虚拟数据生成系统

3.2.1 虚拟数据生成系统概述

基于上述调研的其他仿真数据生成方案，我们了解到想要构建一套能够实际用于神经网络训练的虚拟病害钢轨数据生成系统，我们需要解决以下的挑战：

(1) 这个系统生成的数据需要是可控的，生成的缺陷类型种类需要是我们需要的，否则会成为神经网络训练中无效样本影响训练结果。同时对于固定类型的缺陷还要有一定的随机性，这样才可以满足多样性的需要，训练的模型才会有更高的鲁棒性。

(2) 神经网络的训练结果往往与数据的数量成密切的正相关关系，因此这个系统需要具有在少量人工参与甚至无需人工参与的情况下生成海量数据的能力。

(3) 在有监督学习下，巨量的数据需要消耗难以想象的人力资源来进行标注工作获取每个样本对应的标签。而虚拟数据是我们人为生成的，我们理所当然需要对其有掌控权，因此这个系统需要具有能够自动化提取标签信息的能力。

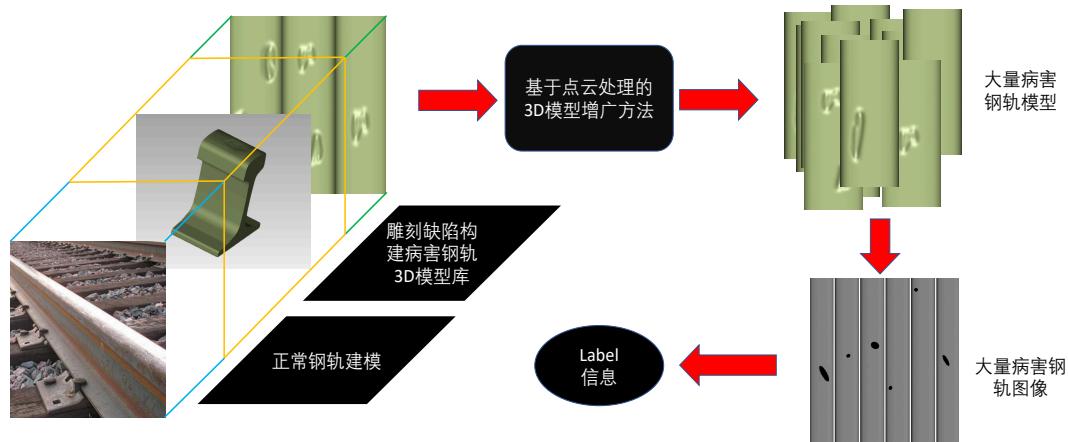


图 3-1 仿真钢轨图像生成系统概览

Figure 3-1 Overview of synthetic rail image generation system

考虑到以上问题，我们提出了基于 3D 模型和点云处理方法的虚拟数据生成系统，这种方式可以优雅的实现以上三个方面。

如图 3-1 所示，我们的生成方案主要分为以下五个步骤：钢轨 3D 模型建模、病害钢轨 3D 模型获取、病害钢轨 3D 模型扩充、病害钢轨 3D 模型表面图像获取、病害钢轨 3D 模型缺陷位置信息自动提取。

为了获取到有各种各样缺陷类型的钢轨 3D 模型，有多种方式。第一种是直接利用 3D 建模工具来构造，这种方式复用性差，每一个带缺陷的钢轨 3D 模型都需要重新构建。另一种方式是首先利用 3D 建模工具根据真实的钢轨特点捏造一段无缺陷的正常钢轨模型，然后再采用模型雕刻工具在正常的钢轨模型上面雕刻我们想要的各种类型的缺陷。这种方式相比第一种方便快捷很多，然而想要生成巨量的带缺陷钢轨 3D 模型还是远远不够。因此，我们提出了基于点云处理方案的带缺陷钢轨 3D 模型扩展方案。在这个方案下，我们首先采用第二种方式来捏造少量的带缺陷钢轨 3D 模型，将这些 3D 模型作为模型仓库。然后通过点云变换对模型仓库中的带缺陷钢轨 3D 模型进行扩增从而达到可以生成无限量所需数据的目的。

其中，钢轨 3D 模型的构建，缺陷的雕刻以及在虚拟空间中对 3D 模型进行表面图像采集偏于交叉学科的工程方面，因此对于这几部分的介绍本文重点会放在详细的工程细节，具体的操作流程以及编写的各个辅助工具的展示。而基于点云处理方法的缺陷钢轨 3D 模型扩充则更倾向计算机学科的理论方面，因此对其的介绍会更侧重于公式化与理论化。通过“因地制宜”的陈述，以期读者能对我们的虚拟数据生成系统有一个全面而深入的了解。

3.2.2 病害钢轨 3D 模型库构建方法

对于带缺陷钢轨 3D 模型库的构建，我们首先使用三维建模软件 Solidworks 来对 60km/h 的标准钢轨进行建模，然后导入到 3D 扫描分析软件 Wrap 以在正常钢轨 3D 模型上面进行雕刻，拟真生成具有各种类型缺陷信息的特殊表面钢轨模型。通过 Wrap，我们可以生成的不同特殊表面钢轨模型包括钢轨表面剥离掉块和钢轨表面擦伤等。表面擦伤和剥落掉块的结果展示如图 3-2 和图 3-3。对于我们钢轨表面缺陷检测场景，我们更重视轨面的状态，因此在实际获取真实钢轨表面图像时，一般会把采集相机置于轨面的正上方进行拍摄，这样就使得钢轨的轨腰部分不会被

采集到。因此，为了加快 3D 模型处理的速度，我们可以考虑将我们 3D 模型的轨腰部分切除，实际上，我们也正是这么做的，切除的钢轨 3D 模型如图 3-4 所示。

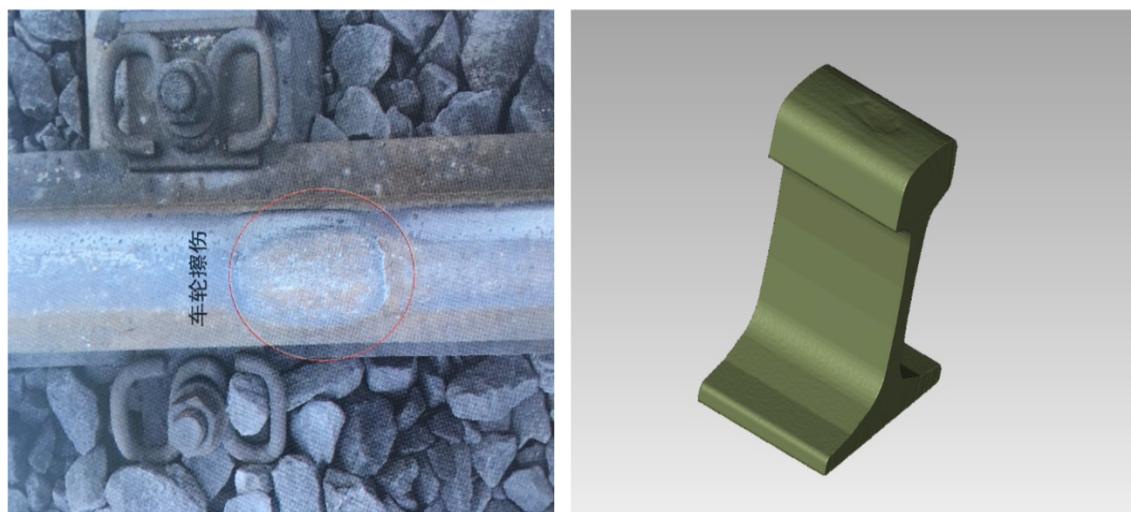


图 3-2 钢轨擦伤真实图像和 3D 模型
Figure 3-2 Real image and 3D model of rail abrasion

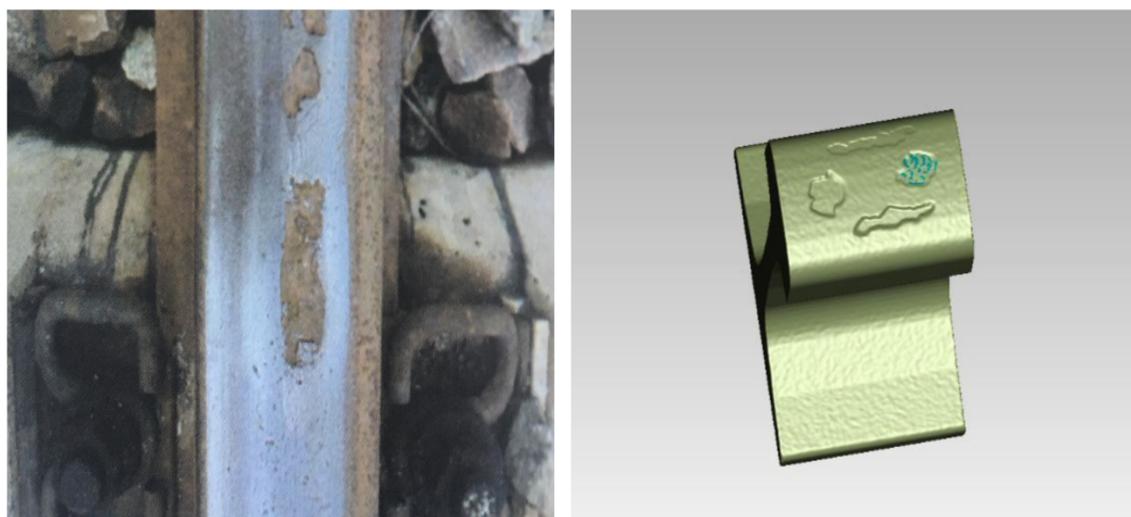


图 3-3 钢轨剥落真实图像和 3D 模型
Figure 3-3 Real image and 3D model of rail surface spalling

3.2.3 基于点云处理的数据增广方法

如本章引言所述，仅仅通过手工的操作方法难以获取巨量的带缺陷钢轨 3D 模型，因此，在本小节中，我们提出了基于点云处理方法的病害钢轨 3D 模型增广方法。

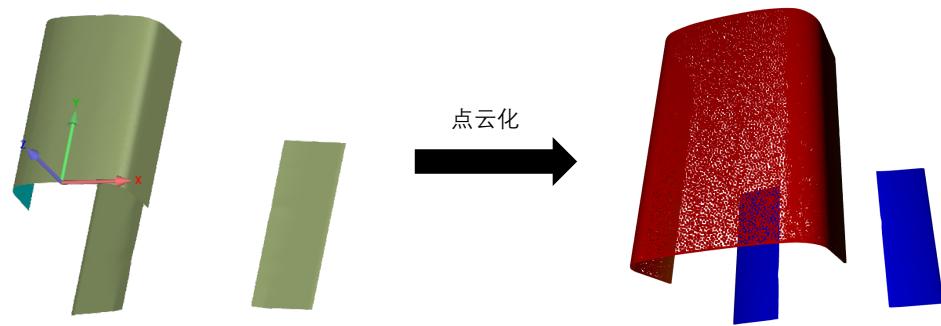


图 3-4 3D 模型点云化示意图

Figure 3-4 3D model point cloud diagram

首先介绍一下什么是点云？所谓点云，简单来说是通过点的堆砌来描绘 3D 物体的一种方法，是将现实世界进行模型化，散列化的一种方式。借助高精度的 3D 点云数据，我们可以更加细粒度地来表现真实世界。通常情况下，点云数据的获取主要有：

(1) 3D 激光扫描法，通常简称为 LiDAR (Light Detection And Ranging)。它发射激光到扫描物体来获取距物体的具体以及检测反射的信息，通过这种方式来采集物体的 3D 信息构建点云数据^[78]。

(2) 图像测量法。这种方式通过从不同的角度，位置拍摄目标物体的交叠图片。然后采用精心设计的算法来评估这些图片的相对位置，最终把这些图片转化为 3D 点云模型，比较常用的摄影测量算法有 SFM 和 MVS^[79]。

(3) 视频测量法。这种方法与图像测量法相似，不同的地方在于它利用视频流来代替了图像。已有研究工作表明，视频中帧与帧之间的连续性对重建 3D 点云更加有益，同时这种方式也大大降低了人工成本^[80]。

(4) RGB-D 相机采集法。RGB-D 相机是在普通的 RGB 相机的基础上添加了物体深度信息的测量，其可以直接采集彩色点云信息，比较知名的是微软的 Kinect。

(5) 双目相机采集法。双目相机是带有两个或多个镜头的图像采集系统，因为两个相机之间的位置和朝向信息是已知的，因此就可以基于采集的 2D 图像来获取 3D 点云数据。比较知名的有 Bumblebee XB3 系统^[81]。

以上几种方式都是在现实世界中的采集方法。在我们的虚拟空间中，由于我们的物体本身即处于数字空间中，所以对 3D 点云的获取更加方便，我们只需对 3D 模型进行采样即可。在本文，我们采用了 Open3D 这个现代的三维数据处理库来实现从 3D 模型到点云和从点云到 3D 模型的转换，图 3-4 是对应的钢轨 3D 模型转为点云文件的形态。

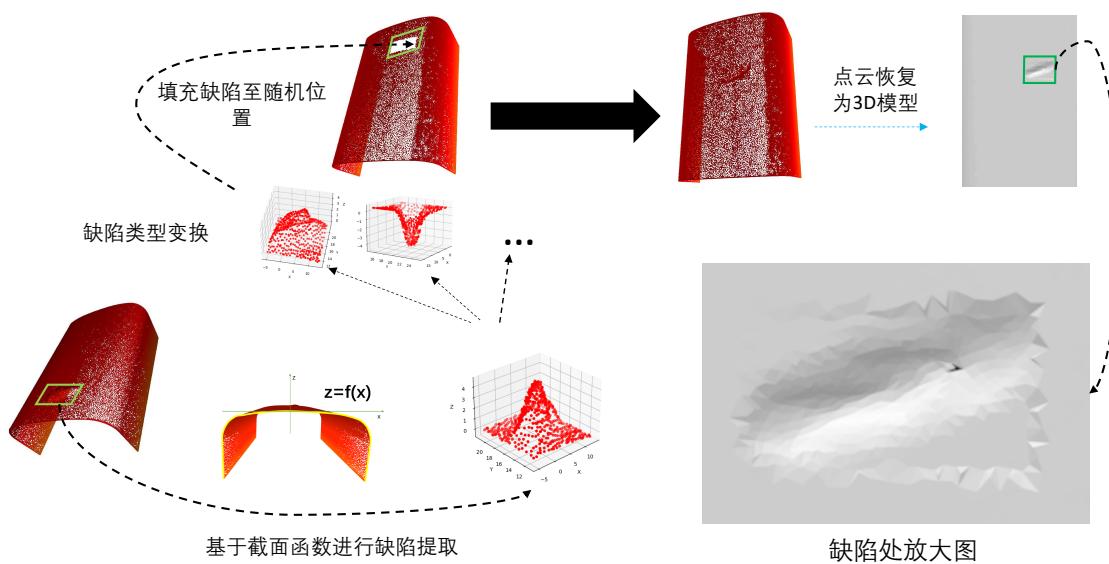


图 3-5 基于点云处理的病害钢轨 3D 模型增广方法

Figure 3-5 Defective Rail 3D Model augmentation method based on point cloud processing

我们基于点云处理的钢轨 3D 模型增广方案的思路如图 3-5 所示，首先我们随机从钢轨 3D 模型库中抽取一个样本。之后，我们将正常钢轨 3D 模型以及抽取的病害钢轨 3D 模型都通过 open3D 这个三维数据处理库转为点云格式。对于缺陷钢轨模型，我们通过缺陷不符合正常钢轨 3D 模型截面函数 $z = f(x)$ 的特点来提取出缺陷点云部分。然后计算包围缺陷点云部分的最小方形包围框，为了鲁棒性，提取最小方形框时在外侧添加 5mm 使框略大于缺陷，之后提取出方形框内的点云信息。

对于提取到的方形框内的点云信息，只改变 x 坐标， y 坐标可实现缺陷的形态变化，改变 z 坐标可使缺陷的深度或者凸起发生变化，而同比例的增大或者缩小三个坐标则可以影响缺陷的大小，因此通过对缺陷点云进行 x, y, z 坐标的变换即可实现缺陷多样性的扩增。同时我们对钢轨可偏移位置进行计算，通过将缺陷点云随机偏移到可选位置即可以实现对于缺陷出现位置的一个增广。

在缺陷点云进行了各式各样的变换之后，我们将其填入到正常钢轨点云之中就得到了一个新的病害钢轨点云文件。之后，我们再将其转化为钢轨 3D 模型来进行后续表面图像的采集。

然而在我们按照上述流程生成一个新的病害钢轨 3D 模型的时候，我们会发现，新生成模型的缺陷与周边存在着严重的过渡不自然问题，如图 3-5 右下角所示。对于这个问题，我们将其定义为“锯齿问题”，其严重影响了生成的病害钢轨的真实性。针对此，我们仔细分析了问题原因并提出了曲率反曲率对偶操作消除了这个弊端，下一小节我们会单独对其进行分析。

3.2.4 曲率和反曲率操作

我们用图 3-6 (a) 来阐述“锯齿问题”出现的原因，图中黑色曲线我们将其示意为正常钢轨点云，红色三角示意为我们的缺陷部分的点云。当我们把经过变换后的缺陷点云填入到钢轨的随机位置的时候，我们可以看到由于钢轨表面存在一定的曲率。这使得缺陷点云与正常钢轨点云之间并不能完全贴合，两者之间存在着一定的间隙。这即是引起“锯齿问题”的原因。

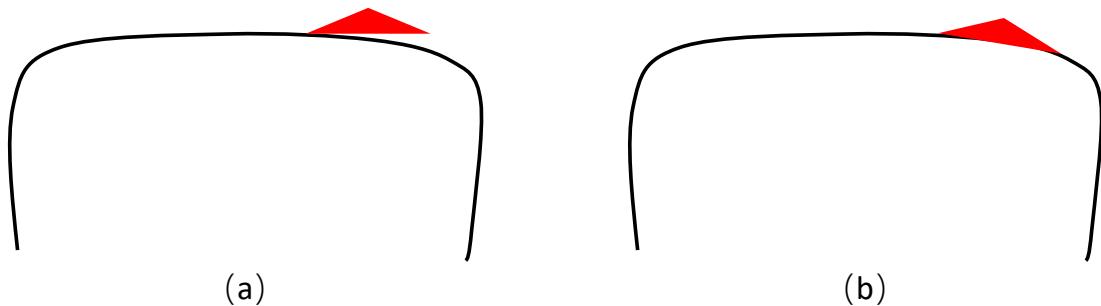


图 3-6 曲率操作示意图

Figure 3-6 Diagram of curvature operation

为了解决这个问题，我们设计了曲率操作。所谓曲率操作，即是给我们提取到的缺陷点云部分添加上所填入位置的曲率信息。如图 3-6 (b) 所示，通过曲率操作来对点进行处理，可以将我们的缺陷点云对齐到钢轨表面，从而来消除所示的“锯齿问题”。经过曲率操作后的钢轨 3D 模型中缺陷处的表现如图 3-8 (b) 所示。

然而，我们可以看到，仅仅通过曲率操作，我们所得的结果并不完美。从图 3-8 (b) 看来，尽管“锯齿问题”得到了一定的改善，但是在缺陷部分的右下角我们可以看到仍旧存在较为严重的锯齿。经过分析，我们认为仅仅执行曲率操作是不充分的，这是因为我们提取的点云不可能是如图 3-6 (a) 中所示的那样，底部是平

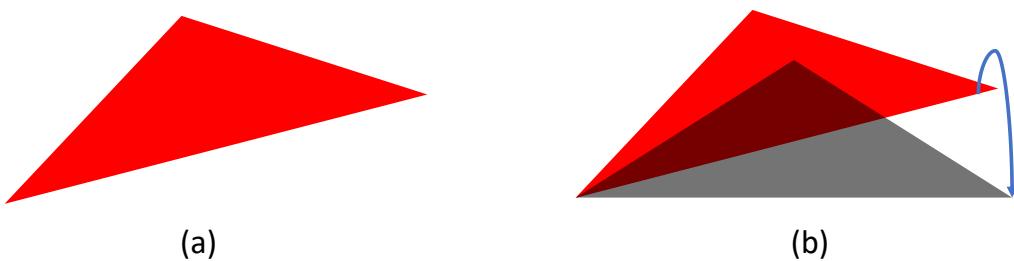


图 3-7 反曲率操作示意图

Figure 3-7 Diagram of anti-curvature operation

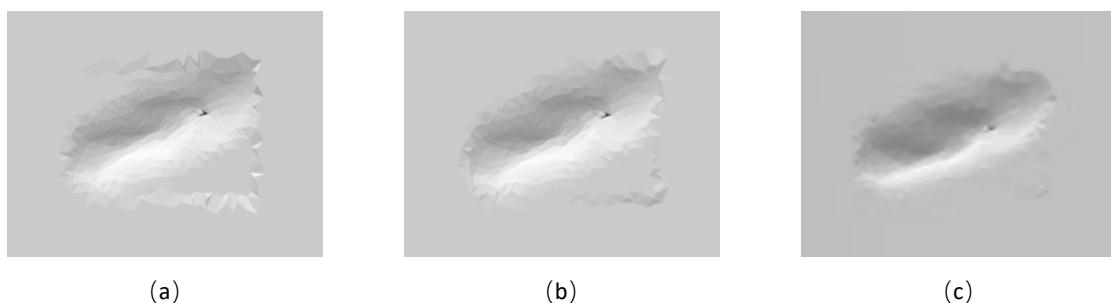


图 3-8 曲率和反曲率操作效果示意图

Figure 3-8 Schematic diagram of curvature and anti-curvature operation effect

滑的，它自己本身也拥有提取位置的曲率信息，实际上它的表现形式应该如图 3-7(a) 所示。因此，我们设计了如图 3-7 所示的反曲率操作来解决这个问题。所谓反曲率操作，即是在将缺陷点云提取出来之后，填入其他位置之前，消除缺陷点云在提取位置的曲率信息。

从最终结果如图 3-8 (c) 所示，经过反曲率操作和曲率操作，钢轨模型在缺陷部分过渡的非常光滑，巧妙的解决了生成钢轨模型时存在的锯齿问题，使得生成的钢轨模型更加逼真。

3.2.5 钢轨 3D 模型表面图像采集

为了对 3D 钢轨模型的表面图像进行采集，我们基于 MaxScript 语言编写了可批量对 3D 钢轨模型进行渲染采图的工具。如图 3-9 所示，我们编写的工具包括以下功能：

- (1) 调整输出图片的大小。我们可以通过设置此参数以对不同长短、大小的钢轨 3D 模型进行表面图像采集，而不至于出现采集不全的情况。
- (2) 可以通过选择或者手动输入确定输出文件的目录，方便我们操作。
- (3) 可以通过选择 3D 模型文件对单个钢轨进行图像采集，也可以选择文件夹来对一整个目录下的钢轨进行表面图像采集。
- (4) 可以对错误添加的钢轨 3D 模型进行移除和批量移除。
- (5) 采集出现错误的钢轨模型会出现在错误列表中，方便我们进行分析。
- (6) 生成图像的文件名和 3D 钢轨模型文件名保持一致，这在后面的标签自动提取步骤至关重要。

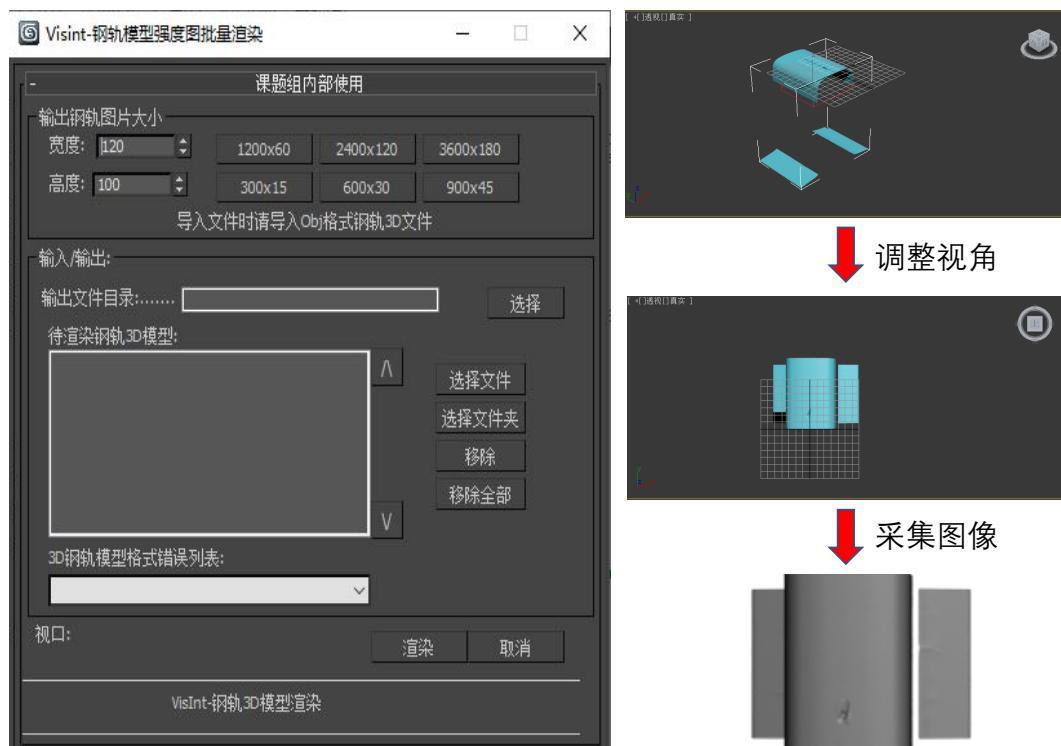


图 3-9 钢轨 3D 模型表面图像采集软件示意图

Figure 3-9 Schematic diagram of rail 3D model surface image acquisition software

3.2.6 标签信息自动提取

当我们同一批模型进行钢轨表面采集时，我们设定的虚拟相机的视角、高度和位置都是一致的。因此，每幅图像中钢轨的位置和大小都是相同的，这为我们进行标签信息自动提取提供了非常有利的条件。我们以示意图 3-10 为例，来说明我们标签自动提取的原理。

首先，我们从这批数据中挑出一副图像，用其来确定我们钢轨轨面区域的左上角距图片左上角的绝对距离 d_1 和 d_2 ，图片的宽高 w_2 和 h_2 ，轨面在图片中的绝对宽高 w_1 和 h_1 。这些信息只要用一幅图像就可以获取到，因为一批数据中的每一幅图像都是这个值。

在基于点云处理的扩增方法中，生成缺陷的信息是由我们指定的。因此，对于缺陷的宽高占钢轨轨面的比例 w 、 h ，缺陷中心距轨面左上角的相对位置坐标 α 和 β ，我们都可以在点云处理阶段进行同步的获取，我们在实际代码中也正是这么做的。

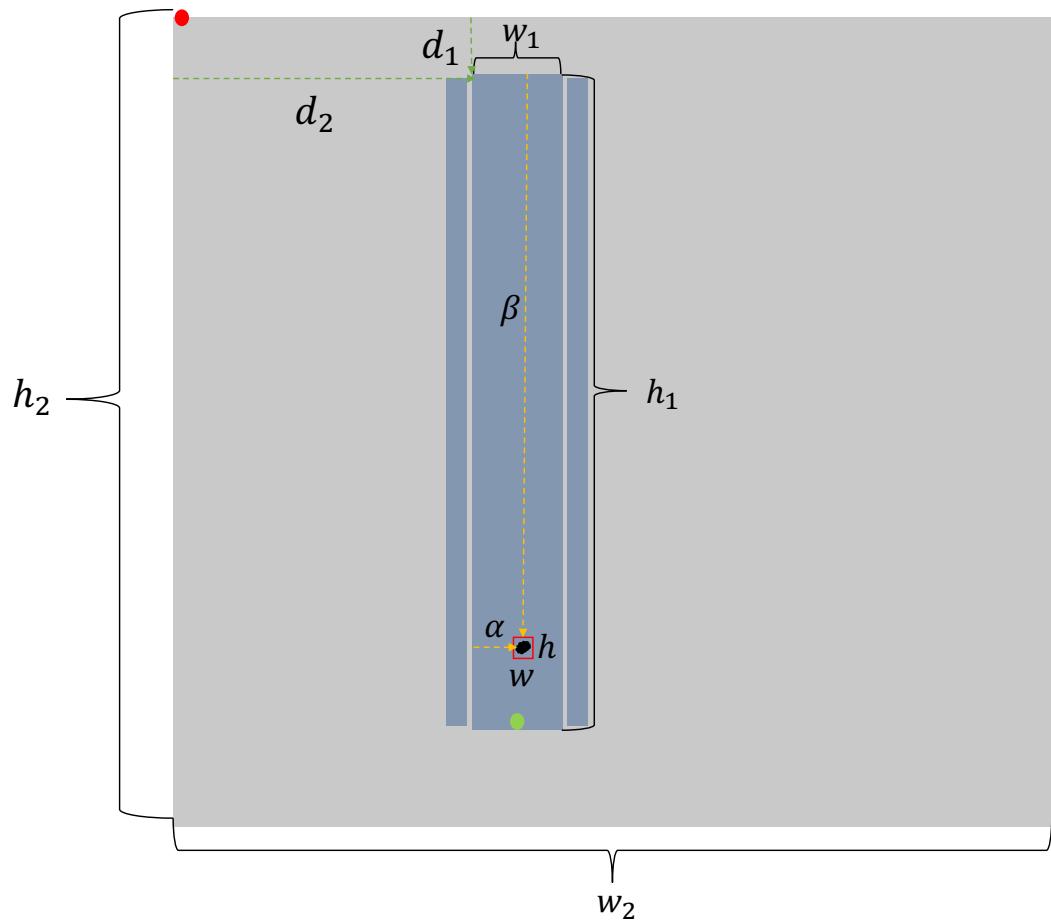


图 3-10 标签自动提取示意图

Figure 3-10 Schematic diagram of automatic label extraction

有了这些已知条件，我们就可以计算出我们的缺陷信息在图片中的相对位置和相对尺寸信息。我们令 X , Y 表示缺陷在图片中的相对位置, W 和 H 表示缺陷的相对图片的宽高比例, 可得: $X = (\alpha w_1 + d_2)/w_2$, $Y = (\beta h_1 + d_1)/h_2$, $W = ww_1/w_2$, $H = hh_1/h_2$ 。

3.3 实验

3.3.1 实验设置

(1) 软硬件环境

本章节中的虚拟数据生成系统的实验均是在一台个人 PC 上完成。其系统为 Windows10 专业版, 硬件设备上配备了双核 2.5GHz 的低压 CPU, 8g 内存以及 HD520 集成显卡。使用的 3D 建模软件版本为 3Dmax 的 V1.4 版本, 模型雕刻使用的软件

为 Wrap 的 2017.0.0.111 版本。在基于点云处理的数据扩增方案中，编程语言为 Python。在图像渲染自动化工具开发中，主要用的编程语言为 MaxScript。

本章节的所有基于深度神经网络模型的训练和测试都是在一台高性能服务器上完成的。其操作系统为 Ubuntu16.04，硬件设备上配备了 18 个 Intel 2.68GHz 的高性能 CPU，256g 内存以及 8 张英伟达的 TITAN Xp 显卡。整个模型的主要编程语言为 Python，使用的深度学习框架是 PytorchV1.2.0。

(2) 数据集

在本章实验中，我们使用了 883 张线阵相机采集的高精度图像，其中包含 143 张病害钢轨图像和 740 张正常的钢轨表面图像。其中 143 张缺陷钢轨图像中共包含离散型缺陷 176 个。同时基于我们的虚拟数据生成系统获得了 757 张仿真图像。

(3) 方法选择与实验对比

在本章的实验过程中，我们选择了 3 个目标检测网络来验证我们虚拟数据的有效性。分别是 YOLOv3^[6]、SSD^[7]、和 YOLOv5^[62]。其中，在 YOLOv3 模型上，我们使用的是 Darknet-53 网络架构的标准 YOLOv3 版本。使用 Adam 优化器来进行网络的训练，其两个关键参数 β_1 和 β_2 分别设置为 0.9 和 0.99。学习率被设置成 0.001，BatchSize（每一批训练的图像数）被设置为 8。在 SSD 检测网络中，使用的特征提取网络为 VGG16 网络。优化器采用的是随机梯度下降（SGD）优化算法，其中初始学习率设置为 0.00001，动量参数（Momentum）设置为 0.9，梯度衰减参数设置为 5×10^{-4} ，BatchSize 大小设置为 24。对于 YOLOv5 检测模型，我们使用的参数和原作者保持一致^[62]。

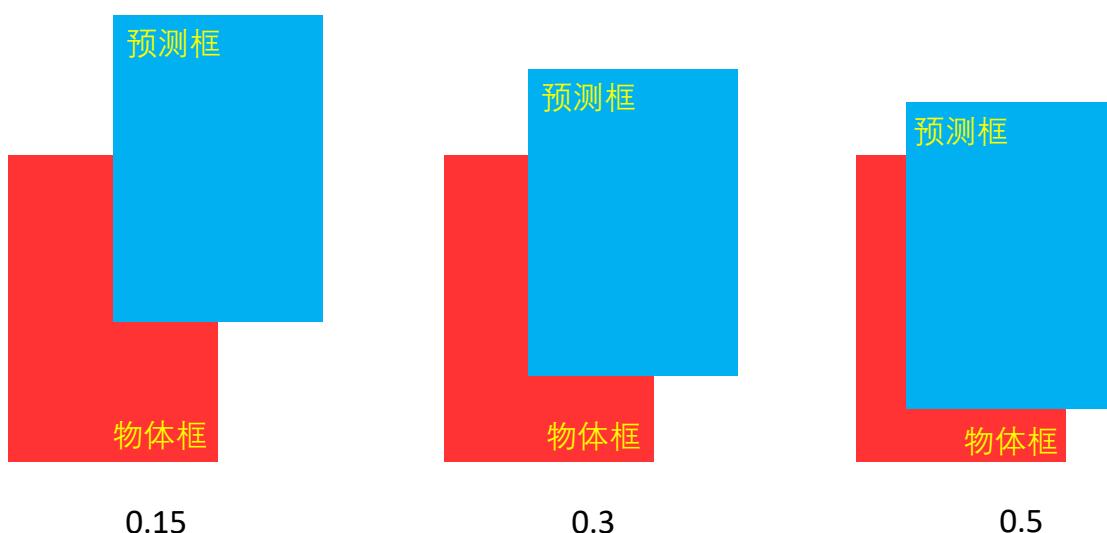


图 3-11 不同大小 IoU 覆盖范围示意图
Figure 3-11 Coverage of different IoU sizes

同时，为了比较基于虚拟数据训练的神经网络的效果，我们也选用了课题组过去设计研制的三种传统图像检测方法来与其进行对比（选用传统方法是因为传统方法跟我们一样，在设计时不需要病害样本的参与）：LN+DLBP^[38]，MLC+PEME^[82]和CTFM^[83]。其中，在LN+DLBP方法中，dp参数被设置为0.3。在MLC+PEME方法中，控制参数 γ 设置为3。在CTFM中，三个参数 w_1, w_2, w_3 分别被设置为0.7, 0.6和1.7。

(4) IoU选择

对于我们的目标检测模型而言，其与分类任务不同的是，它需要考虑预测的边界框和对应区域之间的关系，然后再利用分类的评价指标来进行衡量。在定位方面，本文考虑了在IoU（预测框和真实框的交集余并集的比值）为0.3、0.4和0.5的情况下检测精度。如图3-11所示，可以看到，当IoU为0.3的时候就已经满足了我们对钢轨缺陷定位的需要，因此在综合考虑了现实因素之后，我们选择当预测框和缺陷之间的IoU大于0.3的时候，即认为这个预测框是一个准确的预测框，否则，则认为是一个错误的检测结果。在我们的实验方法中，MLC+PEME方法和CTFM方法的实验结果是分割图而非检测框。对于它们，我们会将包围分割结果的最小方形框作为检测框来与真实框进行计算。

(5) 评价指标

使用准确率Precision、召回率Recall以及F1-Score值来评判我们模型的检测效果。准确率、召回率、F1值分别定义如下：

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3.1)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3.2)$$

$$F1 = \frac{2 \times P \times R}{P + R} \quad (3.3)$$

其中TP表示正确检测出的缺陷数量、FP表示误检的缺陷数量、FN表示漏检的缺陷数量。

3.3.2 实验结果与分析

如上文所述，为了验证本章提出的基于虚拟数据训练神经网络的有效性，我们采用了多个神经网络模型基于虚拟数据进行了足量的训练，并选了一些典型实验结果放到图3-12中进行分析。首先图中的(a)和(g)是同一张图分别通过虚拟

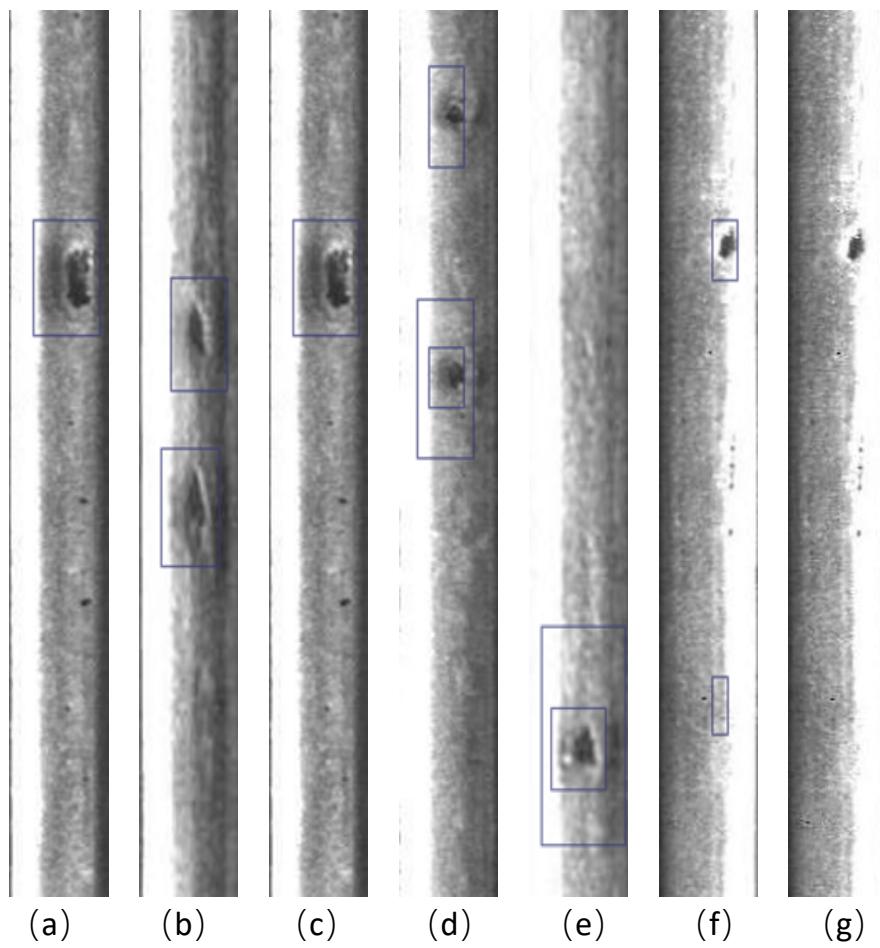


图 3-12 基于虚拟数据训练网络的检测结果

Figure 3-12 Detection results based on virtual data training network

数据训练好的 YOLOv3 网络检测和基础 YOLOv3 网络检测的结果（这里基础 YOLOv3 网络指的是使用 YOLOv3 已经训练好的权重直接检测），从结果可以看出，基于虚拟数据训练的模型可以有效的检测出钢轨缺陷的位置。而采用其他物体训练的

表 3-1 基于虚拟数据训练网络实验结果与传统方法对比图

Table 3-1 Comparison of experimental results between traditional methods and virtual data methods

模型	正确数	误检数	漏检数	准确率	召回率	F1
LN+DLBP	124	410	52	0.232	0.705	0.349
MLC+PEME	156	423	20	0.269	0.886	0.413
CTFM	158	118	18	0.572	0.898	0.699
SSD	129	100	47	0.563	0.733	0.637
YOLOv3	91	159	85	0.364	0.517	0.427
YOLOv5	130	26	46	0.833	0.739	0.783

YOLOv3 基础网络不会检测到任何目标，这证明我们的虚拟数据生成系统是成功的，基于虚拟数据的训练方法也是有效的。但是，它在检测过程中也出现了各式各样其它的问题。比如在 (d) 和 (e) 中出现了检测框重叠的问题，没有很好的过滤掉不完美的重叠的预测框，而在 (f) 中则出现了误检的情况。这些错误的发生都在一定程度上降低了检测的准确率。

为了证明我们的虚拟数据生成系统的的价值，我们也做了与其他传统方法的实验结果对比。实验结果展示如表 3-1 所示。从表中可以看到，当我们把我们的虚拟数据采用 YOLOv5 这个最新的 YOLO 系列模型训练的时候，其准确率达到了惊人的 0.833，大大超出了三种传统方法在这个数据集上的准确率水平。在召回率上，基于虚拟数据训练的 YOLOv5 也仍旧超过了 LN+DLBP。当我们考虑综合指标 F1-Score 值的时候，YOLOv5 对三种传统方法的优势依旧非常的强势，其超出 LN+DLBP 超过百分之 40 的综合性能。而对于传统方法中表现最好的 CTFM，也仍旧有着近百分之 7 的提高效果。这个结果有力地证明了我们虚拟数据生成系统生成的数据完全可以用于神经网络的训练并且可以实现超过传统方法的检测效果。

3.4 小结

本章首先介绍了现有的一些虚拟数据生成系统，概括了它们所具有的一些特征，并严格分析了它们在运用到病害钢轨仿真数据生成时所带来的种种不可接受的弊端。之后，我们提出了一个可用、好用、易用的病害钢轨仿真数据生成系统应该具有的三个特点：可控性和随机性的综合、自动化生成、标签自动提取。并基于这三个特点研发设计实现了我们基于 3D 模型和点云处理方法的虚拟数据生成系统。在点云处理过程中，提出了曲率和反曲率的对偶操作解决了我们面临的“锯齿问题”，让生成的病害钢轨模型更加的逼真。最后，我们把通过虚拟数据生成系统获取的虚拟病害钢轨图像，输入到神经网络中进行训练，在真实数据上进行测试，结果有力的证明了我们方法的有效性。

4 基于虚实结合的钢轨表面缺陷检测方法

4.1 引言

在上一章节中，我们提出了基于虚拟数据的钢轨表面缺陷检测方法，并且取得了不错的结果。但是，科学的进步是无尽头的，因此我们没有停止探索的脚步，而是继续思索了以下问题：在现有的检测方法中我们是否还有什么问题等待解决？我们的结果还能不能提升？我们是否还有其他的方法来进一步提升我们基于虚拟数据检测方法的准确率、召回率和 F1-Score 值？是否还有可以引入到我们方法中的数据？这些问题带着我们一步步完成了这一章节的内容。

如图 4-1，我们比较了虚拟数据和真实数据在视觉上的特征。可以看到，尽管我们已努力让我们生成的虚拟数据在光照、视角差异各个条件上保持与真实数据的一致性。但是虚拟数据和真实数据在视觉表现上仍旧有一定的区分度，这说明，我们的虚拟数据和真实数据存在着一定的域偏移现象。对于域偏移问题的研究，已经有一段时间的历史，在章节 2 中我们也已经进行了简单的介绍。然而，传统域适应要解决的问题和我们有很大的不同。在传统域适应方法中，源域和目标域的样本都很充足，只是目标域没有标签。因此，它们想要解决的问题主要是如何节省大量的标注成本。但是在我们场景却存在诸多限制。首先，我们的目标域就是病害钢轨样本，其数量和多样性上都非常的匮乏，想要用目标域样本来缓解或者说解决虚拟数据和真实数据域偏移的问题显得非常的艰难。其次，对我们来说，我们更想要设

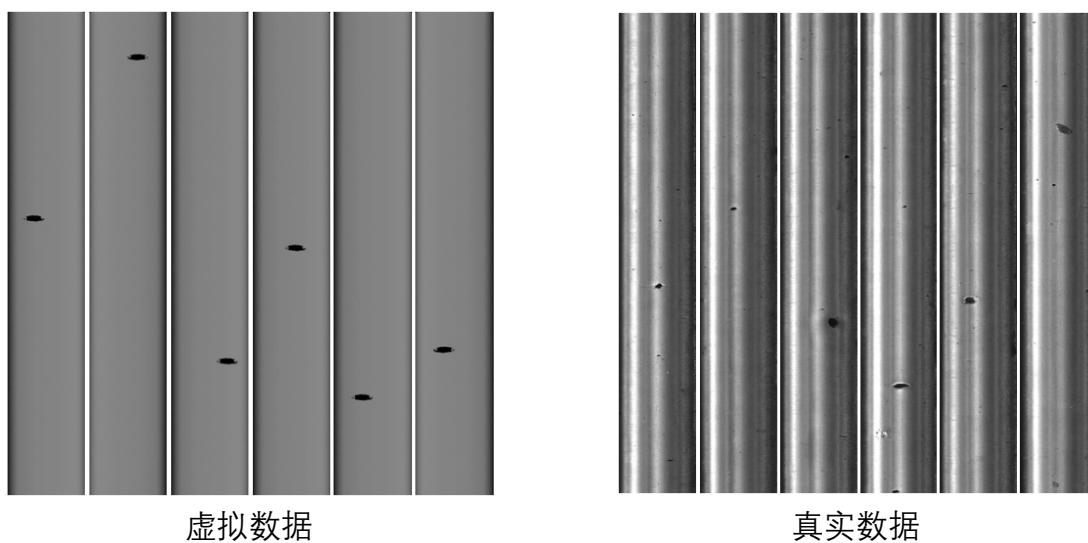


图 4-1 虚拟数据和真实数据视觉对比图

Figure 4-1 Visual comparison of virtual data and real data

计一个完全不需要真实病害钢轨样本参与的钢轨检测模型，如果把病害钢轨样本引入到了训练当中，也违背了我们的初衷。

基于此，我们提出了“伪目标域”的概念，并且通过“伪目标域”的概念，设计了图像对齐和特征对齐两个模块进一步挖掘了我们虚拟数据的潜力，并且在实验部分证实了这两个模块的有效性。

因此，在本章节中，我们首先会利用符号系统和数学语言形式化我们的域偏移问题。然后提出我们的“伪目标域”概念，阐述我们提出“伪目标域”概念的细节和缘由。之后，详述我们提出的图像对齐和特征对齐两个模块，并将其嵌入到YOLOv3 基底模型上进行实验验证。

4.2 虚实结合研究方法

4.2.1 伪目标域

我们定义 $P_s(B, I)$ 为源域中样本的联合分布(joint distribution)，而 $P_T(B, I)$ 为目标域中样本的联合分布，其中 I 代表图像， B 代表图像中的缺陷， S 代表源域， T 代表目标域。在我们的场景中，源域指的是仿真数据构成的训练集，目标域指的是真实的病害钢轨图像构成的测试集。在目标检测任务中、如果我们想要让训练集训练的模型能够在测试集上有效，我们需要保证训练集和测试集上的联合分布是一致的，也即需要保证

$$P_s(B, I) = P_T(B, I) \quad (4.1)$$

然而，由于我们的训练图像来自于我们生成的仿真数据，其与真实图像存在着较大的分布差异如图 4，使得实际情况中

$$P_s(B, I) \neq P_T(B, I) \quad (4.2)$$

利用贝叶斯公式将联合分布进行分解：

$$P(B, I) = P(B|I)P(I) \quad (4.3)$$

像域适应中的任务一样，我们使用协方差偏移假设(covariate shift assumption)。即我们源域和目标域的条件分布 $P(B|I)$ 是相同的，而边缘分布 $P(I)$ 是不同的。因此，通过强制性地消除源域和目标域的边缘分布差异，我们就可以消除源域和目标域的联合分布差异，从而使得在源域上训练的模型在目标域上仍旧有效。

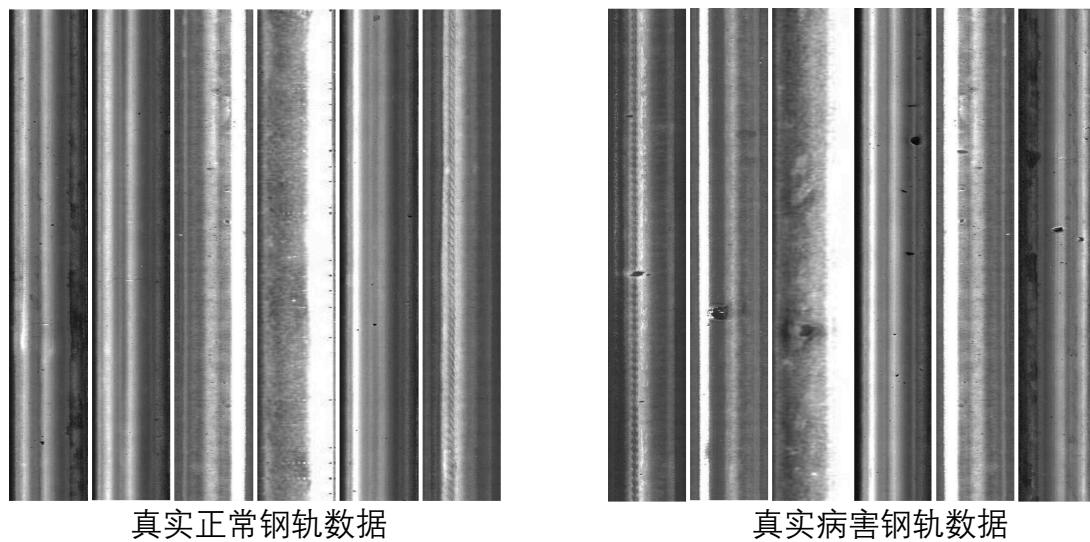


图 4-2 真实正常数据和病害数据视觉对比图

Figure 4-2 Visual comparison of real normal data and disease data

然而，在我们的场景中，我们的目标域代表的是真实的病害钢轨图像，其在训练过程中我们是无法使用的，从而无法利用其来消除源域和目标域之间的分布差异。因此，我们提出了伪目标域的概念。伪目标域我们把它定义为一个与目标域之间分布差异小于源域与目标域分布差异的一个新的数据分布域。伪目标域里用的是真实的正常钢轨图像、其与目标域之间存在着非常小的分布差异如图 4-2，即

$$P_D(I) \approx P_T(I) \quad (4.4)$$

$P_D(I)$ 表示伪目标域中的图像分布，D 代表伪目标域(Dummy target domain)。因为在钢轨采集中，真实的无病害钢轨图像极易采集，所以我们可以得到足量的伪目标域的数据。因此，有了伪目标域，我们就可以通过消除源域和伪目标域的分布差异来间接消除源域和目标域之间的分布差异从而来尝试解决虚拟数据和真实数据的域差问题。（ \rightleftharpoons 符号表示消除两个数据分布之间的分布差异）

$$P_S(I) \rightleftharpoons P_T(I) \Leftarrow \begin{cases} P_S(I) \rightleftharpoons P_D(I) \\ P_D(I) \approx P_T(I) \end{cases} \quad (4.5)$$

4.2.2 图像对齐模块

有了理论的支撑，我们就开始思考怎么解决源域和伪目标域之间的域差问题。我们首先想到的是如何利用真实的无病害钢轨数据来从视觉上让我们的虚拟数据变得和真实数据更加相似。基于这个动机，我们提出了我们的图像对齐模块，即从图像的层面让虚拟数据构成的源域和真实图像构成的目标域达到对齐。

我们的图像对齐模块考虑了两种思路，一种是基于 CycleGAN 的风格迁移方法，另外一种则是传统的图像处理方法，并分析对比了两者在结果上的差异以及选择策略。

(1) 基于 CycleGAN 的图像对齐方法

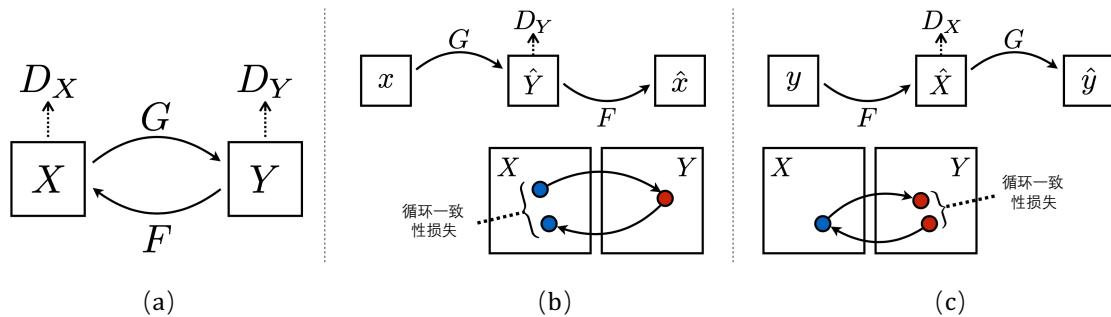


图 4-3 CycleGAN 网络机理^[84]

Figure 4-3 Schematic diagram of CycleGAN's network principle

CycleGAN 是一个无监督的风格迁移模型，其可以学习一个从数据域 X 到数据域 Y 的普适性映射，从而解决一系列视觉领域的问题，比如超分辨率、图像增强以及风格变换等。其原理如图 4-3 所示，(a) 中的 X 和 Y 是两个数据域，而 G 和 F 则是两个生成器，其分别学习 X 到 Y 和 Y 到 X 的映射。 D_X 和 D_Y 是两个判别器，用来判断生成器学习的映射是否准确，指导生成器的生成效果。当训练完成后，我们就得到了一个可以由风格 X 图像生成到风格 Y 图像的生成器，和一个可以由风格 Y 图像生成风格 X 的生成器。然而，只是这样还是不够的，因为我们这样只是做了风格的生成，并没有做风格的迁移。比如我们想训练一个素描画风格转为真实画风格的模型，如果仅仅使用上述方法，我们会发现我们输入了一个素描风格的椅子，它生成的可能是一个真实风格的小狗。因此，为保证我们输入图像除了有风格上的相异性，还应该在高层特征上保证它们的相似性，因此 CycleGAN 设计了图 4-3(b) 和图 4-3(c) 中的循环一致性损失 L_{cyc} 来达到拉近高层特征分布的目的。CycleGAN 的损失函数如公式 4.6 所示，更加清晰的表明了 CycleGAN 的作用机理。

$$L(G, F, D_X, D_Y) = L_{GAN}(G, D_Y, X, Y) + L_{GAN}(F, D_X, Y, X) + \lambda L_{cyc}(G, F) \quad (4.6)$$

我们使用了我们的虚拟病害钢轨图像和真实的正常钢轨图像对 CycleGAN 进行了训练，得到的风格迁移结果如图 4-4 所示。可以看到，基于 CycleGAN 迁移风格后的虚拟钢轨图像数据看起来与真实数据的差异更小。

(2) 基于纹理叠加的图像对齐方法

另外一种图像对齐方法，我们使用的是纹理叠加操作。对于我们生成的仿真图像集 S 中的每一个图像 s ，我们都随机的从真实的正常钢轨图像集 D 中选取一副真实

的正常钢轨图像来作为叠加图像 d 。对它们进行像素级的叠加得到对齐后的图像 e ，最后所有的 e 组成的图像集 E 就是我们初步对齐后的数据集。

$$E = \{e | \mu s + (1 - \mu)d, s \in S, d \in D\} \quad (4.7)$$

如图 4-5 所示，通过这样的方法、我们初步缩减了源域和目标域的分布差异。这种方法相比于风格迁移的方法更加简单高效，并且通过实验证明，这种简单的方法在提高模型效果上仍旧有着极大的作用。

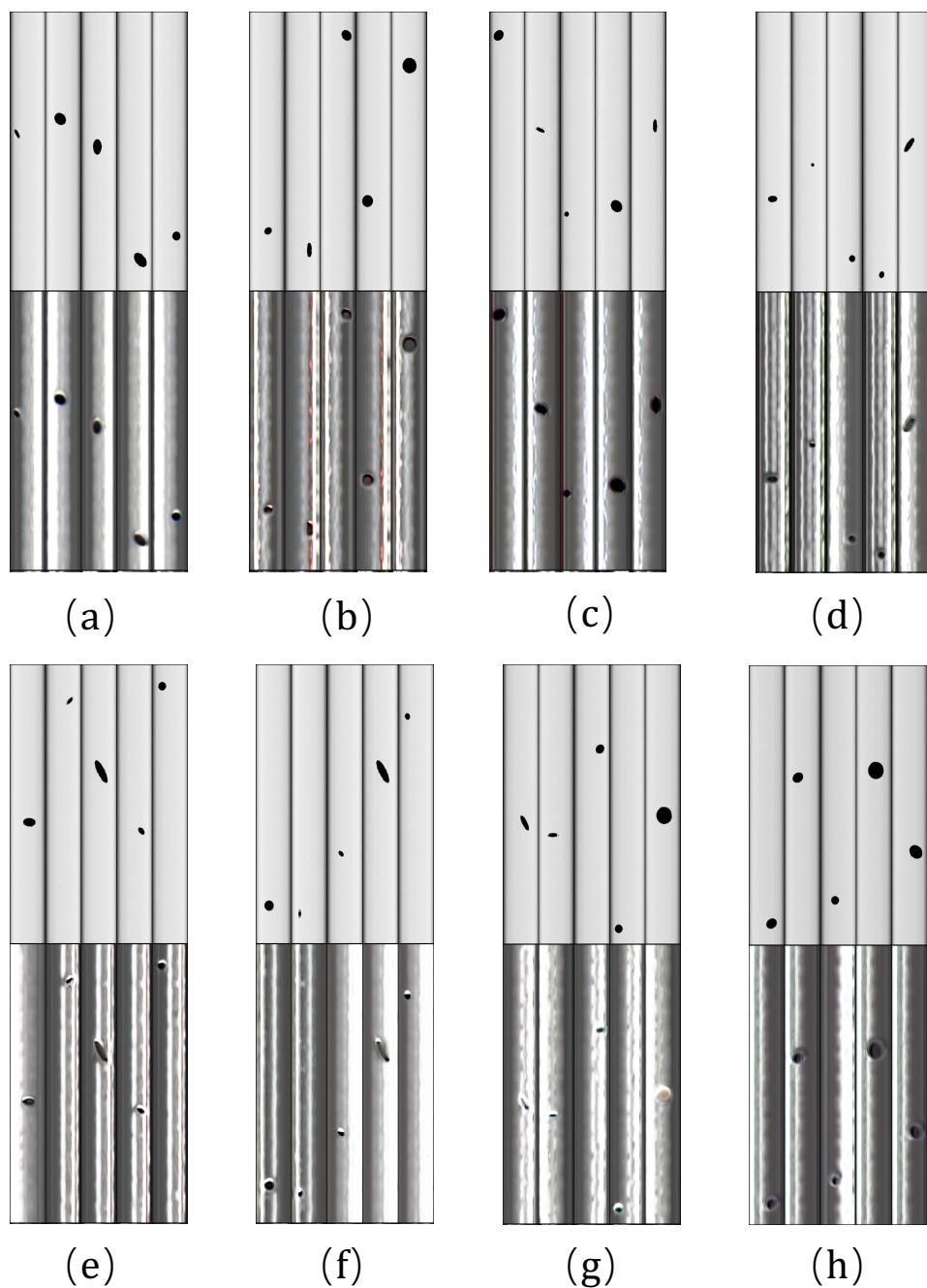


图 4-4 风格迁移效果示意图

Figure 4-4 Sketch of style transfer effect

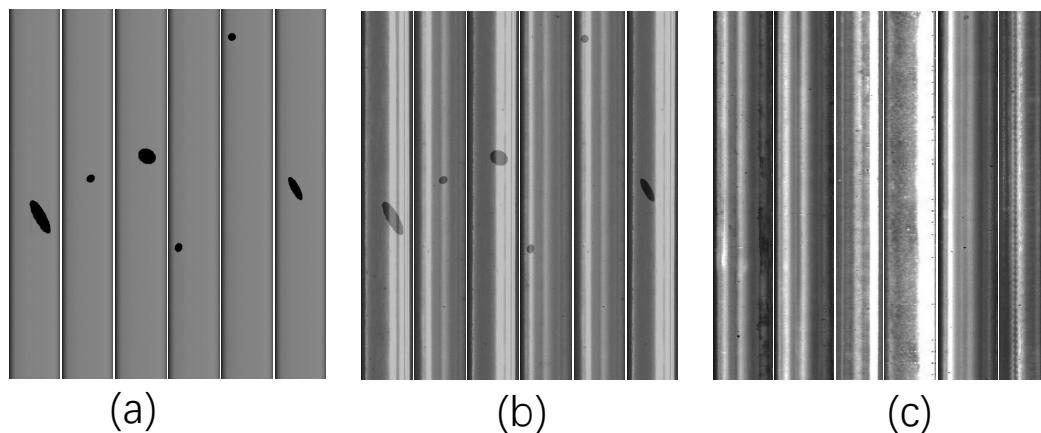


图 4-5 纹理叠加效果示意图
Figure 4-5 Texture overlay effect schematic

4.2.3 特征对齐模块

除了两种图像对齐的模块之外，我们还设计了特征对齐的模块。特征对齐模块通过改造原有神经网络的特征提取层加以实现。通过特征对齐模块，我们想要做的是使得提取到的特征更能体现源域和伪目标域的共同特征。也即是说让源域图像和伪目标域图像在通过改造后的特征提取层之后，两者提取的隐层特征分布尽可能的趋于一致，两个数据分布的距离变得更小。

为了衡量两个特征分布之间的距离，我们借鉴了 Ben-David 在 2010 年提出的 \mathcal{H} 散度^[70]，定义距离公式为

$$dis(S, D) = 2 \left(1 - \min_H (err_S(H(x)) + err_D(H(x))) \right) \quad (4.8)$$

其中， x 表示特征提取网络输出的特征向量。 H 则是一个经过设计的分类器，其用来判断我们的特征向量来自于源域还是伪目标域。 err_S 和 err_D 是分类器 H 在源域和伪目标域上的预测错误率。因此，为了使得源域和伪目标域的特征分布差异减小，我们就需要让我们的特征提取器 F 来降低分类器 H 的错误率 err_S 和 err_D 。

$$\min_F dis(S, D) = \max_F \min_H (err_S(H(x)) + err_D(H(x))) \quad (4.9)$$

因此，由公式 4.9 可以看出，我们的特征对齐模块是由一个分类器和一套对抗网络训练方法所构成的。我们的特征提取网络 F 和域分类器 H 如何才能实现对抗的训练呢？我们使用了被称作 GRL 层（Gradient reverse layer）^[71]的“便携式插件”来实现端对端的对抗训练。

GRL 层被称作梯度反转层，在 2015 年被提出。梯度反转层中不包括任何的参数，被放在我们的特征提取网络和域分类器的中间。在前向传播中，GRL 层表现为一个恒等变换。而在后向传播中，GRL 会从后面的层（域分类器那层）获取梯度，并将其乘上权重因子 $-\lambda$ 传递到前面的层（特征提取层）。在数学上，我们通过公式 4.10 和公式 4.11 两个等式来形式化梯度反转层 $R_\lambda(x)$ 在前向传播和反向传播中的梯度表现，公式中的 I 是一个特征矩阵^[72]。

$$R_\lambda = x \quad (4.10)$$

$$R'_\lambda(x) = -\lambda I \quad (4.11)$$

因此当反向传播经过梯度反转层时，我们的特征提取网络和域分类网络就形成了一个对抗架构。我们的特征提取网络在尽力的提取源域和伪目标域的共有特征，以混淆域分类器。而域分类器也在努力的提高自己的分辨能力以区分越来越相似的源域伪目标域特征。两者相互竞争，相互提高。

4.2.4 适应性模型 DA-YOLO

为了验证我们设计两个模块的有效性，我们基于 YOLOv3 模型设计了融合两个对齐模块的全新网络架构 DA-YOLO。使用 YOLOv3 而非 YOLOv5 作为基底模型的原因是它结构简单清晰、更加利于说明我们每个模块的作用，探究我们每个模块在使用当中需要注意的问题，其中 YOLOv3 的网络架构我们已在 2.2.2 小节进行了详细介绍。

(a) 网络架构

如图 4-6 所示，我们的 DA-YOLO 模型主要包含特征提取器 F ，缺陷检测器 G 以及域判别器 H 。其中 F 是 DarkNet-53 特征提取网络的一个子集，而 F 和 G 两者组成一个完整的 YOLOv3 架构。域判别器 H 是我们设计的一个轻量级自适应的分类网络。

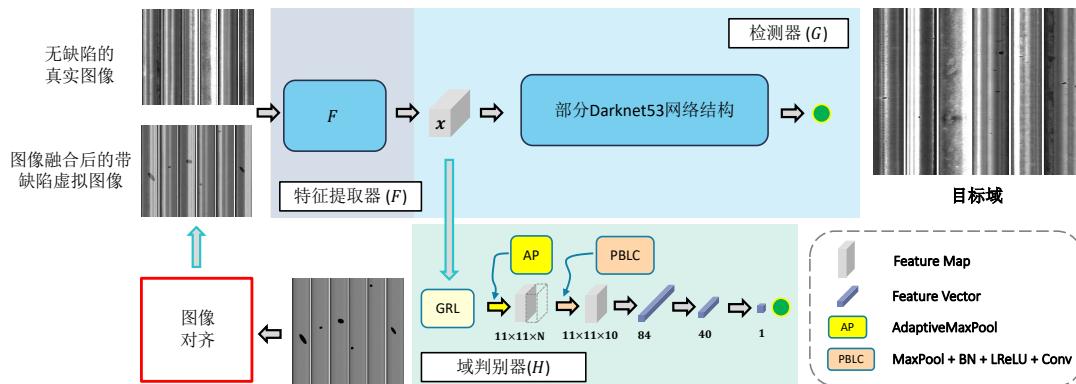


图 4-6 DA-YOLO 网络架构
Figure 4-6 DA-YOLO Network Architecture

其会对经过 F 提取的特征进行进一步的特征提取与处理，然后判断这个特征是来自源域还是伪目标域。我们的域判别器包含两个池化层，一个卷积层和三个全连接层。其中，我们还用到了BN层和Leaky Relu损失函数来使我们的网络更易于训练。

(b) 网络处理流程

在输入图片之后，我们的DA-YOLO首先采用了图像对齐方法对生成的虚拟数据进行调整，使其在数据分布上初步与真实数据保持一致。然后将虚拟病害钢轨图片和真实无病害钢轨图片一起输入到特征提取器 F 中来提取特征。

网络处理流程如图4-6所示。在训练阶段的前向传播中，两者的特征会进行以下处理。对于获取到的虚拟病害钢轨数据特征，会流入到检测器 G 中输出缺陷检测结果，同时其也会经过GRL层之后流进域判别器 H ，输出其判断结果。而对于获取到的真实正常钢轨数据特征，仅仅会流入到 H 中进行域的判断。在反向传播的时候，检测器 G 会根据缺陷检测的结果进行梯度更新，域判别器则会根据域分类的结果来进行梯度更新。而特征提取器则受两个结果的综合的梯度优化，使得自己提取的特征能够兼顾鲁棒性和准确性。在测试阶段，则是直接将真实病害钢轨图片输入到训练好的 F 和 G 中，输出最终的识别结果。

(c) 损失函数

我们的损失函数由两部分组成。第一部分是缺陷检测的损失 L_{det} ，这一部分的公式与YOLOv3中保持一致^[6]。另一部分则是我们的特征对齐损失 L_{align} ，公式为

$$L_{align} = - \sum [C_i \log(p_i) + (1 - C_i) \log(1 - p_i)] \quad (4.12)$$

其中 $C_i \in \{0,1\}$ 表示的是第 i 个图像的域标签、0表示域判别器 H 认为这个特征向量属于源域。 p_i 是 H 的输出。通过GRL层，我们使得 H 来最小化 L_{align} ，而 F 来最大化 L_{align} 。

4.3 实验

4.3.1 实验设置

(a) 软硬件环境

本章节的所有基于深度神经网络模型的训练和测试都是在一台高性能服务器上完成。其操作系统为Ubuntu16.04，硬件设备上配备了18个Intel 2.68GHz的高性能CPU，256g内存以及8张英伟达的TITAN Xp显卡。整个模型的主要编程语言为Python，使用的深度学习框架是PytorchV1.2.0。

(b) 对比实验

为了验证提出的 DA-YOLO 模型在基于虚拟数据上进行缺陷检测的优越性。我们对比了其与原始 YOLOv3, SSD、YOLOv5 以及一些传统方法的实验结果。在 DA-YOLO 中，我们使用 Adam 优化器来进行网络的训练，其两个关键参数 β_1 和 β_2 分别设置为 0.9 和 0.99。学习率被设置成 0.001。BatchSize 设置为 8。而在基于纹理叠加的图像对齐模块中，我们将等式 4.7 中的参数 μ 设置为 0.5。在训练 F , G , H 时我们采用交叉训练的方法。我们首先训练 F 和 G 来提高模型的目标检测能力，等下一个 epoch，我们再继续训练 F 和 H 的域适应能力。以此规则进行交叉训练，保证特征提取层 F 能够提取到既具有域适应能力又有强大缺陷检测能力的特征向量。其余模型和方法的参数设置和章节 3 保持一致。

(c) 消融实验

为了证明我们所提网络中每个模块的有效性，充分挖掘各个模块之间的关系和作用。我们设计了完整的消融实验。其包括：(1) 基准检测网络 (2) 基于纹理叠加的检测网络 (3) 基于风格迁移的检测网络 (4) 基于特征对齐的检测网络 (5) 基于纹理叠加和特征对齐的检测网络 (6) 基于风格迁移和特征对齐的检测网络。

(d) 参数敏感性实验

在我们的特征对齐模块中，有一个比较关键的参数即是我们特征对齐的层数。我们通过设计不同特征对齐层数的网络加以对比来探究分析此参数对结果的影响，对齐的层数指的就是我们在第几层添加域判别器 H 。具体来说，我们分别随机选择对齐 DarkNet-53 架构的前 5 层、前 19 层和前 37 层。

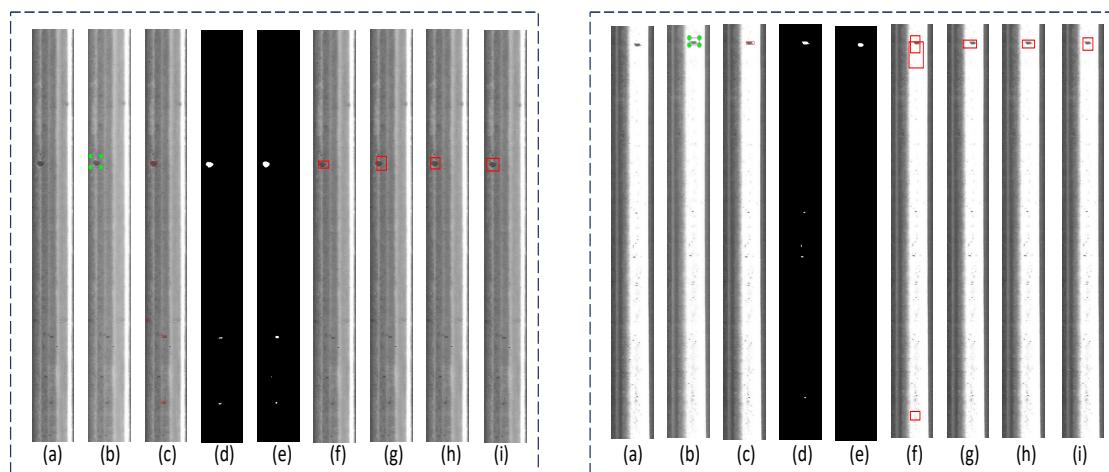


图 4-7 各方法检测图像展示

Figure 4-7 Detection results of different methods. (a) Original image. (b) Ground truth. (c) LN+DLBP. (d) MLC+PEME. (e) CTFM. (f) YOLO-V3. (g) SSD. (h) YOLOv5. (i) Our method.

4.3.2 实验结果与分析

图 4-7 展示了 DA-YOLO 的检测和其他方法检测结果的少部分图像。而具体的准确率、召回率、F1-Score 详细对比结果我们可以参考表 4-1。从表 4-1 中我们可以得到如下的分析结果。

首先，我们的三种传统缺陷检测方法 LN+DLBP、MLC+PEME 和 CTFM 都具有较高的召回率，甚至 CTFM 的召回率在所有的方法中是最高的。但是，在误检数上，它们也是独占鳌头，这导致它们有着较低的准确率，也拉低了他们的综合检测指标 F1-Score。相比于这三个传统方法，我们 DA-YOLO 的优势可以看成是压倒性的。DA-YOLO 在综合评价指标 F1-Score 上领先了 CTFM 13.1 个百分点，超过 LN+DLBP 接近两倍半。

其次，我们的 DA-YOLO 与神经网络模型比较，也有很大的优势。相比于 SSD 网络和 YOLOv3 网络，DA-YOLO 有着至少 19.3 个百分点的综合指标提升，这是因为我们模型融合了图像对齐模块和特征对齐模块，具有更强的域适应能力，训练出的模型也更加的鲁棒。相比于 YOLOv5 的方法，我们的优势则显得没有那么亮眼。这是因为 YOLOv5 作为一个工业界成熟的目标检测模型，其自身包含了许多 Trick 来帮助提高模型的泛化能力和准确度，而我们的 DA-YOLO 是为了说明我们模块的有效性而选择的更简易的 YOLOv3 模型。但即便如此，我们相比于 YOLOv5 模型仍旧有接近 5 个百分点的综合指标提升。因此，综合看来，我们设计的 DA-YOLO 在我们虚拟数据场景下的检测效果更加高效和具有竞争力。

表 4-1 各方法与 DA-YOLO 实验结果对比

Table 4-1 Comparison of experimental results between DA-YOLO and other methods

模型	正确数	误检数	漏检数	准确率	召回率	F1
LN+DLBP	124	410	52	0.232	0.705	0.349
MLC+PEME	156	423	20	0.269	0.886	0.413
CTFM	158	118	18	0.572	0.898	0.699
SSD	129	100	47	0.563	0.733	0.637
YOLOv3	91	159	85	0.364	0.517	0.427
YOLOv5	130	26	46	0.833	0.739	0.783
DA-YOLO	147	31	29	0.826	0.835	0.830

表 4-2 消融实验
Table 4-2 Ablation experiments

所用模块	正确数	误检数	漏检数	准确率	召回率	F1
无	91	159	85	0.364	0.517	0.427
纹理叠加	160	75	16	0.681	0.909	0.779
风格迁移	150	72	26	0.676	0.852	0.754
特征对齐	142	42	34	0.772	0.807	0.789
纹理叠加+特征对齐	147	31	29	0.826	0.835	0.830
风格迁移+特征对齐	143	27	33	0.841	0.813	0.827

消融实验的结果如表 4-2。从表格 4-2 中，我们可以得到如下发现。首先相比于没有添加任何模块的基底模型，我们的纹理叠加和风格迁移两种对齐模块分别可以带来 35.2 和 32.7 个百分点的综合性能提升，而设计的特征对齐模块则可以提高 36.2 个百分点。同时，从实验结果中也可以看到我们的图像对齐模块和特征对齐模块两者具有相互促进的作用。比如纹理叠加+特征对齐相比于仅仅通过纹理叠加提升了至少 5 个点，相比于单独通过特征对齐模块提升了 4 个点。其次，我们也观察到了我们的特征对齐模块在添加到网络之后，为综合性能提升起主要作用的是误检数的下降，而正确数是由一些略微下降的。我们分析这是因为通过特征对齐，网络对噪声的分辨更加灵敏的同时，也对缺陷目标的确认变得更加“谨慎”，从而导致了召回率的下降。

参数敏感性实验的结果如表 4-3。从表格中我们可以看出，对齐浅层网络可有效提升钢轨缺陷检测的各项性能指标。这是因为浅层的网络层一般代表提取的是边缘、颜色等。而一般来说，也正是在这些信息上，我们的源域和伪目标域存在着域差，因此对齐这些底层特征起到了很好的效果，使得综合性能检测指标从 0.427 提升到了 0.830。随着对齐层数的增加，尽管其性能仍旧相比于原始架构有着很大的提升，但可以看到其相比于仅对齐低层特征，在性能上有着明显的下降。这一方面是因为深层网络层提取到的一般是如缺陷位置、大小的高级语义信息，而我们的伪目标域缺乏高层语义代表的目标（钢轨缺陷），这使得其根本就无法与源域在高层语义对齐。另一方面，对齐层数越多，低级特征对齐在整个对齐过程中所占的比重也相应减少，从而影响了实验的效果。

表 4-3 特征对齐中对齐层数对结果的影响

Table 4-3 The influence of the number of alignment layers in feature alignment on the results

对齐层数	正确数	误检数	漏检数	准确率	召回率	F1
0	91	159	85	0.364	0.517	0.427
5	147	31	29	0.826	0.835	0.830
19	127	25	49	0.836	0.722	0.775
37	133	51	43	0.723	0.756	0.739

4.4 小结

本小节首先从分析数据分布出发，利用公式和符号系统形式化了我们的虚拟病害钢轨数据和真实病害钢轨数据的域偏移问题。为了解决我们目标域样本不足的问题，定义了一个新的概念“伪目标域”，并阐述了其具体定义和在我们场景下的可用性。基于“伪目标域”，设计了图像对齐和特征对齐两个功能模块来提升基于虚拟钢轨数据缺陷检测的性能，其中图像对齐包括了风格迁移和纹理叠加两个方案。然后，我们将两个功能模块嵌入到 YOLOv3 中设计了我们的 DA-YOLO 检测模型，并将 DA-YOLO 检测结果与其他方法对比证实了两个模块的有效性。最后，我们设计了消融实验研究了两个模块之间存在的一些关系，以及设计了参数敏感性实验探究了特征对齐中对齐层数的选择问题。

5 总结与展望

本章主要对前文所述内容进行回顾总结，剖析当前研究工作的不足之处，并展望本课题在未来可以做的一些其他尝试。

本文的研究对象是钢轨，是铁路交通中最重要的部件之一。其表面缺陷会极大影响到铁路交通的运营安全，甚至引发严重事故对旅客的生命和财产安全造成威胁。因此，对其表面缺陷进行有效检测并提前更换病害钢轨，在铁路运营中必不可少。本文聚焦在了钢轨表面缺陷检测中遇到的病害样本严重不足问题，并提出了基于虚实结合的缺陷检测方法。提出的虚实结合方法可以有效利用虚拟的病害钢轨样本和真实的正常钢轨样本，来训练一个性能好、鲁棒性高的目标检测模型。本文的研究方案分为了虚拟数据生成和适应性模型构建两大模块，下面对这两个方法的主要创新点、实验结果进行总结。

(1) 基于 3D 模型与点云处理方法的虚拟数据生成系统

为了生成真实性高、可控性强并可以自动提取标签信息的虚拟病害钢轨图像，我们设计实现了基于 3D 模型和点云处理方法的虚拟数据生成系统。其通过 3D 模型处理软件构建基础的病害钢轨模型库，利用点云操作来对病害钢轨 3D 模型的数量进行扩增，然后对病害钢轨 3D 模型进行表面图像采集获取最终的虚拟样本。在点云处理过程中，我们分析了影响 3D 模型生成效果的“锯齿问题”，并设计了曲率和反曲率操作加以解决。最后，我们利用这些虚拟数据训练目前常用的一些目标检测网络，然后在真实数据集上进行测试，并与传统的图像处理方法进行对比。实验结果表明，在都不需要真实病害钢轨数据参与下，基于虚拟数据训练的 YOLOv5 模型在性能表现上远远超出了对比的三个传统方法（超出 LN+DLBP、MLC+PEME 和 CTFM 分别 43.4、37 和 8.4 个百分点），这证实了在无真实病害钢轨样本监督领域，我们虚拟数据生成系统的优越性。

(2) 基于虚实结合的钢轨表面缺陷检测网络

为了进一步利用我们的虚拟数据，我们又提出了两个网络模块来提高基于虚拟数据训练的检测网络性能。第一个是图像对齐模块，其主要从视觉上拉近我们的虚拟数据和真实数据之间的数据分布差异。我们提出了纹理叠加和风格迁移的两种图像对齐模块，其分别能够提高基底模型 35.2 和 32.7 个百分点的性能。第二个是特征对齐模块，其主要是从神经网络隐层来拉近两者之间的分布差异，使用了 GRL 层来构建对抗架构。我们的特征对齐模块能够提升基底模型 36.2 个百分点。同时，我们设计的将两者都融入在内的 DA-YOLO 模型在综合性能上达到了 0.83，

超出了当前热门目标检测模型 YOLOv5 近 5 个百分点。这些实验结果证明了我们模块的有效性。

本文提出的基于虚实结合的钢轨表面缺陷检测方法很好地解决了病害钢轨样本不足的问题，有着良好的实验结果。但是由于时间、能力的限制，本文提出的研究思路、观点、结论、使用的技术、编程方法都还有着进一步的提升空间。通过对实验结果的分析并与组内成员的讨论，我们认为还有以下几个地方今后可以进一步的探究与加强。

(1) 表面图像采集软件的优化

我们在现实场景中采集钢轨表面图像的时候，光照强度、雾雨天气、相机自身质量以及采集车在行进过程中的颠簸，都会影响到我们采集的图片。这让图片本身更加复杂多变。尽管我们在这方面付出了接近整个研究时间的一半，但是由于研究领域的相异以及时间的关系，我们渲染图像时并没有容纳这么多复杂的因素，这使得我们的虚拟图片相比于真实场景显得更为单一。因此，后期对于采集软件的优化还有着很大的提升空间。

(2) 风格迁移权重选择条件

我们基于风格迁移的图像对齐方法，在衡量虚拟到真实迁移水平的时候，并不是一个自动化的过程。我们需要人工的用肉眼来决定哪一个 epoch 训练的权重迁移效果最好。因此，我们今后可以尝试着探索一些自动化的选择条件。

(3) 多种风格的统一处理

在真实数据集中，可能会含有来自不同轨道线路的图像，其风格各异。而我们的特征对齐和基于风格迁移的图像对齐都是针对一种风格进行的处理。未来可以考虑采用一次训练，多种风格共同处理的思路来进行优化我们的两个模块。

(4) 对抗训练的不稳定性

在特征对齐模块中，本文采用了对抗训练的架构来训练特征提取器 F 和域判别器 H。在训练过程中，会出现我们的 H 能力过强导致对抗训练失效的问题，也即是 H 能够正确识别所有的特征类别使的 loss 为 0 不再变化，无法反向传播优化 F。目前我们的解决方案是人为的添加一些错误数据，将部分虚拟样本的 label 改为真实，这可以部分解决这方面的问题。但此方法并不优雅，所以后期工作中我们还需要寻找解决此问题的最佳方案。

(5) 研究方法的推广

本文提出的基于虚实结合的研究方案不仅仅局限在钢轨缺陷检测领域，其他领域在面临样本不足的情况时，只要满足缺陷在整幅图中所占视觉比例不大，就可以尝试着使用此解决方案来进行研究。因此，今后我们也可以尝试着将我们的研究方法扩展到其他领域，在新的场景中来验证我们方法的有效性。

参考文献

- [1] 石甜. 基于 2D/3D 复合机器视觉的三维钢轨表面缺陷检测技术研究[D]. 湖北: 武汉科技大学, 2017.
- [2] Girshick R, Donahue J, Darrell T, et al. Region-based convolutional networks for accurate object detection and segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 38(1): 142–158.
- [3] Girshick R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. 2015: 1440–1448.
- [4] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. Advances in Neural Information Processing Systems, 2015, 28.
- [5] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 2117–2125.
- [6] Redmon J, Farhadi A. YOLOv3: An incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.
- [7] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector[C]//European Conference on Computer Vision. Springer, Cham, 2016: 21–37.
- [8] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. 2017: 2980–2988.
- [9] Duan K, Bai S, Xie L, et al. Centernet: Keypoint triplets for object detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 6569–6578.
- [10] Carion N, Massa F, Synnaeve G, et al. End-to-end object detection with transformers[C]//European Conference on Computer Vision. Springer, Cham, 2020: 213–229.
- [11] Everingham M, Van Gool L, Williams C K I, et al. The pascal visual object classes (voc) challenge[J]. International Journal of Computer Vision, 2010, 88(2): 303–338.
- [12] Lin T Y, Maire M, Belongie S, et al. Microsoft Coco: Common objects in context[C]//European Conference on Computer Vision. Springer, Cham, 2014: 740–755.
- [13] Wei X, Wei D, Suo D, et al. Multi-target defect identification for railway track line based on image processing and improved YOLOv3 model[J]. IEEE Access, 2020, 8: 61973–61988.

- [14] Yanan S, Hui Z, Li L, et al. Rail surface defect detection method based on YOLOv3 deep learning networks[C]//2018 Chinese Automation Congress (CAC). IEEE, 2018: 1563–1568.
- [15] 李文超, 张丕状. 超声波检测钢轨缺陷及定位的研究[J]. 核电子学与探测技术, 2012, 32(9):4.
- [16] di Scalea F L, Bartoli I, Rizzo P, et al. High-speed defect detection in rails by noncontact guided ultrasonic testing[J]. Transportation Research Record, 2005, 1916(1): 66–77.
- [17] Sun M, Lin X, Wu Z, et al. Non-destructive photoacoustic detecting method for high-speed rail surface defects[C]//2014 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings. IEEE, 2014: 896–900.
- [18] 李骏. 基于时频空特征的钢轨探伤识别研究[D]. 华南理工大学, 2017.
- [19] 时亚. 钢轨踏面裂纹电磁超声表面波换能器优化设计及实验[D]. 南昌航空大学, 2018.
- [20] Wilson J W, Tian G Y. 3D magnetic field sensing for magnetic flux leakage defect characterisation[J]. Insight—Non-Destructive Testing and Condition Monitoring, 2006, 48(6): 357–359.
- [21] Li Y, Tian G Y, Ward S. Numerical simulation on magnetic flux leakage evaluation at high speed[J]. NDT & E International, 2006, 39(5): 367–373.
- [22] Chen Z, Xuan J, Wang P, et al. Simulation on high speed rail magnetic flux leakage inspection[C]//2011 IEEE International Instrumentation and Measurement Technology Conference. IEEE, 2011: 1–5.
- [23] Gao J, Du G, Wei H. The research of defect detection test system based on magnetic flux leakage[C]//Proceedings of 2011 6th International Forum on Strategic Technology. IEEE, 2011, 2: 1225–1229.
- [24] Oliveira H, Correia P L. Automatic road crack detection and characterization[J]. IEEE Transactions on Intelligent Transportation Systems, 2012, 14(1): 155–168.
- [25] Shruthi N, Iype G M, Kavana C P, et al. Rail Track Defect Detection using Enhanced Method of Magnetic Flux Leakage Signal[C]//2021 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C). IEEE, 2021: 277–280.
- [26] 汪开灿. 基于电磁超声的钢轨缺陷检测系统的研究[D]. 哈尔滨: 哈尔滨工业大学, 2010.
- [27] Thomas H M, Heckel T, Hanspach G. Advantage of a combined ultrasonic and eddy current examination for railway inspection trains[J]. Insight—Non-Destructive Testing and Condition Monitoring, 2007, 49(6): 341–344.
- [28] 李国厚, 黄平捷, 陈佩华, 等. 涡流检测在钢轨裂纹定量化评估中的应用[J]. 浙江大学学报: 工学版, 2011, 45(11): 2038–2042.
- [29] Wei Q, Zhang X, Wang Y, et al. Rail defect detection based on vibration acceleration signals[C]//2013 IEEE International Instrumentation and Measurement Technology Conference (I2MTC). IEEE, 2013: 1194–1199.
- [30] 彭谢丹. 面向钢轨踏面斜裂纹损伤的脉冲涡流检测方法研究[D]. 杭州: 浙江大学, 2018.

- [31] Wilson J, Tian G Y, Abidin I Z, et al. Modelling and evaluation of eddy current stimulated thermography[J]. Nondestructive Testing and Evaluation, 2010, 25(3) : 205–218.
- [32] Gao B, Bai L, Woo W L, et al. Thermography pattern analysis and separation[J]. Applied Physics Letters, 2014, 104(25) : 251902.
- [33] Gao B, Bai L, Woo W L, et al. Automatic defect identification of eddy current pulsed thermography using single channel blind source separation[J]. IEEE Transactions on Instrumentation and Measurement, 2013, 63(4) : 913–922.
- [34] 张康. 基于涡流脉冲热成像的闭合疲劳裂纹检测方法研究[D]. 西南交通大学, 2018.
- [35] Deutschl E, Gasser C, Niel A, et al. Defect detection on rail surfaces by a vision based system[C]//IEEE Intelligent Vehicles Symposium, 2004. IEEE, 2004: 507–511.
- [36] 官鑫, 赵智雅, 高晓蓉. 图像处理技术在钢轨表面缺陷检测和分类中的应用[J]. 铁路计算机应用, 2009 (6) : 27–30.
- [37] 王海涛, 甄理, 杨春霞, 等. 基于计算机视觉的铁轨表面缺陷检测系统[D]. 无损检测, 2011.
- [38] Li Q, Ren S. 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.
- [39] 王时丽, 刘桂华. 基于机器视觉的钢轨表面缺陷三维检测方法[J]. 微型机与应用, 2015, 34(19) : 10–13.
- [40] Hajizadeh S, Núñez A, Tax D M J. Semi-supervised rail defect detection from imbalanced image data[J]. IFAC-PapersOnLine, 2016, 49(3) : 78–83.
- [41] 何恬. 基于图像处理的钢轨表面缺陷检测算法研究[D]. 华东交通大学, 2021.
- [42] Nieniewski M. Morphological detection and extraction of rail surface defects[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(9) : 6870–6879.
- [43] Zhang X, Wang K, Wang Y, et al. An improved method of rail health monitoring based on CNN and multiple acoustic emission events[C]//2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC). IEEE, 2017: 1–6.
- [44] 宋雅男. 基于卷积神经网络多模型融合的钢轨表面缺陷检测方法[D]. 长沙理工大学, 2020.
- [45] Tu Z, Wu S, Kang G, et al. Real-Time Defect Detection of Track Components: Considering Class Imbalance and Subtle Difference Between Classes[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1–12.
- [46] 王新国. 便携式轨道检测车——总体设计及机械系统设计[D]. 江苏:南京理工大学, 2008. DOI:10.7666/d.y1367871.
- [47] 金侠挺. 基于深度学习的高铁钢轨表面缺陷检测方法研究[D]. 湖南大学, 2020.
- [48] 贾利红. 基于 DSP 的高速轨道扣件检测算法研究及系统设计[D]. 四川:电子科技大学, 2014. DOI:10.7666/d.D500024.
- [49] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11) : 2278–2324.

- [50]Deng J, Dong W, Socher R, et al. Imagenet: A large-scale hierarchical image database[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. Ieee, 2009: 248–255.
- [51]Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25.
- [52]Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [53]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. 2016: 770–778.
- [54]Howard A G, Zhu M, Chen B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[J]. arXiv preprint arXiv:1704.04861, 2017.
- [55]Sandler M, Howard A, Zhu M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 4510–4520.
- [56]Howard A, Sandler M, Chu G, et al. Searching for mobilenetV3[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 1314–1324.
- [57]Zhang X, Zhou X, Lin M, et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 6848–6856.
- [58]Ma N, Zhang X, Zheng H T, et al. Shufflenet v2: Practical guidelines for efficient cnn architecture design[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 116–131.
- [59]Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 779–788.
- [60]Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 7263–7271.
- [61]Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.
- [62]Jocher G, Stoken A, Borovec J, et al. ultralytics/yolov5: v3. 1-bug fixes and performance improvements[J]. Zenodo, 2020.
- [63]Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]//International Conference on Machine Learning. PMLR, 2015: 448–456.
- [64]Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 2117–2125.
- [65]张康. 基于涡流脉冲热成像的闭合疲劳裂纹检测方法研究[D]. 四川:西南交通大学, 2018.

- [66]Pan S J, Yang Q. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*[J]. 22 (10), 2009, 1345.
- [67]Ghifary M, Kleijn W B, Zhang M. Domain adaptive neural networks for object recognition[C]//Pacific Rim International Conference on Artificial Intelligence. Springer, Cham, 2014: 898–904.
- [68]Tzeng E, Hoffman J, Zhang N, et al. Deep domain confusion: Maximizing for domain invariance[J]. arXiv preprint arXiv:1412.3474, 2014.
- [69]Long M, Cao Y, Wang J, et al. Learning transferable features with deep adaptation networks[C]//International Conference on Machine Learning. PMLR, 2015: 97–105.
- [70]Ben-David S, Blitzer J, Crammer K, et al. A theory of learning from different domains[J]. *Machine Learning*, 2010, 79(1): 151–175.
- [71]Ben-David S, Blitzer J, Crammer K, et al. Analysis of representations for domain adaptation[J]. *Advances in Neural Information Processing Systems*, 2006, 19.
- [72]Ganin Y, Lempitsky V. Unsupervised domain adaptation by backpropagation[C]//International Conference on Machine Learning. PMLR, 2015: 1180–1189.
- [73]Tzeng E, Hoffman J, Saenko K, et al. Adversarial discriminative domain adaptation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 7167–7176.
- [74]Pei Z, Cao Z, Long M, et al. Multi-adversarial domain adaptation[C]//Thirty-second AAAI Conference on Artificial Intelligence. 2018.
- [75]黃茜, 汪玉琳, 李志杰. 铸件的枝状缩松缺陷仿真图像的生成方法及系统[P]. CN104091327A. 2019.
- [76]伊国栋, 李琎. 一种基于迁移学习的注塑制品表面图像缺陷识别方法[P]. CN110111297B.
- [77]Han C, Hayashi H, Rundo L, et al. GAN-based synthetic brain MR image generation[C]//2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE, 2018: 734–738.
- [78]Wang Q, Tan Y, Mei Z. Computational methods of acquisition and processing of 3D point cloud data for construction applications[J]. *Archives of Computational Methods in Engineering*, 2020, 27(2): 479–499.
- [79]Lu Q, Lee S. Image-based technologies for constructing as-is building information models for existing buildings[J]. *Journal of Computing in Civil Engineering*, 2017, 31(4): 04017005.
- [80]Brilakis I, Fathi H, Rashidi A. Progressive 3D reconstruction of infrastructure with videogrammetry[J]. *Automation in Construction*, 2011, 20(7):884–895.
- [81]Son H, Kim C. 3D structural component recognition and modeling method using color and 3D data for construction progress monitoring[J]. *Automation in Construction*, 2010, 19(7): 844–854.

- [82] Li Q, Ren S. 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.
- [83] Yu H, Li Q, Tan Y, et al. A coarse-to-fine model for rail surface defect detection[J]. IEEE Transactions on Instrumentation and Measurement, 2018, 68(3): 656–666.
- [84] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//Proceedings of the IEEE International Conference on Computer Vision. 2017: 2223–2232.

作者简历及攻读硕士学位期间取得的研究成果

一、作者简历

崔文凯，男，河南周口人，汉族

2019. 9–2022. 7 北京交通大学计算机与信息技术学院 计算机科学与技术 硕士

2015. 9–2019. 7 北京交通大学电子与信息工程学院 通信工程 学士

二、发表论文

- [1] Cui W, Wang J, Yu H, et al. From Digital Model to Reality Application: A Domain Adaptation Method for Rail Defect Detection[C]//Chinese Conference on Pattern Recognition and Computer Vision (PRCV). Springer, Cham, 2021: 115–126.

三、参与科研项目与比赛

[1] 北京市自然基金“轨道交通联合”：轨道线路可靠性分析的视觉感知和缺陷识别技术研究，2019–2022.

[2] 北京交通大学：自动 3D 场景生成的研发，2020–2021.

[3] 铁路总公司（原铁道部）：高速铁路基础设施检测监测体系框架及关键技术研究，2018–2020.

[4] 智创工程 AI 挑战赛暨人工智能与轨道安全数据竞赛亚军，2019.

[5] 全国研究生数学建模竞赛二等奖，2020.

四、专利

[1] 李清勇，崔文凯，王建柱等. 一种基于 3D 模型和点云处理的钢轨缺陷图像生成方法[P]. CN113240790A. 2021.

.

.

独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作和取得的研究成果，除了文中特别加以标注和致谢之处外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得北京交通大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名：崔文凯

签字日期：2022年6月1日

学位论文数据集

表 1.1: 数据集页

关键词*	密级*	中图分类号	UDC	论文资助				
钢轨缺陷检测； 3D 模型；点云处理；目标检测； 神经网络；域适应	公开							
学位授予单位名称*		学位授予单位代码*	学位类别*	学位级别*				
北京交通大学		10004	工学	硕士				
论文题名*		并列题名		论文语种*				
基于虚实结合的钢轨表面缺陷检测研究				中文				
作者姓名*	崔文凯		学号*	19120343				
培养单位名称*		培养单位代码*	培养单位地址	邮编				
北京交通大学		10004	北京市海淀区西直门外上园村 3 号	100044				
学科专业*		研究方向*	学制*	学位授予年*				
计算机科学与技术		计算机视觉	3 年	2022 年				
论文提交日期*	2022 年 6 月							
导师姓名*	李清勇		职称*	教授				
评阅人	答辩委员会主席*		答辩委员会成员					
	丁丁		张春 刘海洋					
电子版论文提交格式 文本(√) 图像() 视频() 音频() 多媒体() 其他()								
推荐格式: application/msword; application/pdf								
电子版论文出版(发布)者	电子版论文出版(发布)地			权限声明				
论文总页数*	48							
共 33 项, 其中带*为必填数据, 为 21 项。								