

中国人民大学

《Web 信息处理与应用》课程大作业



题目名称：_____ 基于 3D 点云的零件缺失检测算法 _____

学生姓名：_____ 范肇心 _____

学生学号：_____ 2019000169 _____

专业年级：_____ 信息学院 2019 级博士班 _____

分 数：_____ _____

目录

一、研究背景.....	1
二、项目目标.....	2
三、相关文献综述.....	3
3.1 基于图像的故障检测方法.....	3
3.2 点云语义分割算法.....	4
3.3 点云配准算法.....	6
四、模型介绍.....	7
4.1 模型主要思路.....	7
4.2 编码器介绍.....	8
4.4 解码器介绍.....	9
4.5 损失函数.....	9
五、实验内容.....	9
5.1 数据集.....	9
5.2 模型训练和结果.....	10
六、总结.....	11
参考文献.....	11

一、研究背景

随着动车数量的增长和速度的不断提升,动车故障的检测任务越来越重,安全隐患问题也越来越明显,因此开发一套智能的动车故障检测系统十分必要。如

今人们已经可以借助高速摄像机获取高清稳定的动车表面可视部位图像,但是需要大量工作人员从这些图像中找到动车故障。目标检测是计算机视觉领域的一个重要研究方向,它的主要任务是从图像或视频中识别并定位潜在目标对象,从而在动车可视部位图像中检测出故障并分类定位。但传统的目标检测算法存在大量漏检和误检,因此该技术实际应用效果并不佳。近几年,深度学习得到空前发展,基于深度学习的目标检测算法开始逐渐兴起。对比传统的目标检测算法,深度学习方法不需要手工提取特征,而是从大量样本数据中学习提取特征,得到的特征更加丰富,也更加具有鲁棒性,其在检测精度和速度两方面都远远超越了传统算法。因此深入研究基于深度学习的目标检测算法,然后应用到动车故障检测中是一项十分有意义的工作。

以往的基于深度学习的动车故障诊断方法往往如前文中所描述的一样使用图片作为基本的数据体,然后利用目标检测等算法来诊断故障位置。然而在实际情况中,使用图片来检测故障往往会存在误检等情况。且由于图片只能反映 2D 信息,在进行故障检测时候,会由于待检测零件在空间上的分布特点等原因,产生漏检等情况。因此,使用 3D 数据体来进行该类型的故障诊断对于提高故障诊断的精度有十分重要的意义。

在 3D 众多 3D 数据体中,点云由于其丰富的 3D 空间表达能力,广泛地受到工业界和学术界的关注。云是在同一空间参考系下表达目标空间分布和目标表面特性的海量点集合,在获取物体表面每个采样点的空间坐标后,得到的是点的集合,称之为“点云”(Point Cloud)。相比于网格和体素等数据体,3D 点云占据的内容空间更少,对 3D 场景的感知能力更强。随着激光雷达、深度相机等硬件设施的出现,点云变得更加容易获取且高精度。

基于上述分析,本文旨在探究如何利用激光点云数据体来进行动车故障检测。

二、项目目标

本文主要研究故障问题中的零件缺失问题。如图 1 所示,动车等机械设备在运转过程,其零件由于磨损、机械碰撞等原因,会出现掉落等情况,该掉落会造成极大的安全风险。例如,图 1 中所示位置处为动车刹车片,左图为正常情况,

右图三块刹车片中的左侧刹车片缺失、露出了内测的一块，这会造成动车在刹车时的极大的安全隐患，因此在检修时，这种故障一定要及时被检测出。在检测该类问题时，传统方法往往使用基于图片的 2D 目标检测方法来检测，然而，该类方案并不适用于图中所示情况，因为当前块刹车片缺失时，后面的刹车片会露出并被相机拍到。算法由于检测到后侧的刹车片，因此不会报告零件缺失异常。因此，解决该类问题必须使用基于 3D 的方法来完成。传统的 3D 方法通常使用点云配准等方法，将采集到的异常数据体和标准情况下的数据体进行配准，然后利用基于距离的规则来检测缺失。但是基于传统点云配准的方法存在以下弊端：1. 传统的点云配准方法迭代次数多、运行时间慢、效率低、且受到初始化影响较大。2. 基于点云配准的方法容易受到视角等因素的影响。3. 传统算法没有通用性，针对不同的故障区域和故障零件，可能要设计不同的规则。因此，本文采用基于深度学习的方法来完成零件检测任务，将缺失检测问题转化为语义分割问题来解决。方法细节将在后续章节介绍。

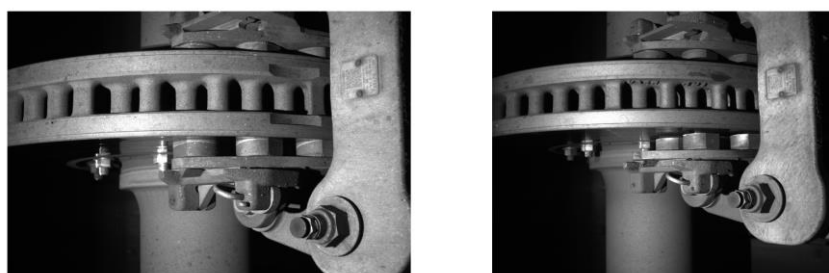


图 1 零件掉块问题示例

三、相关文献综述

3.1 基于图像故障检测方法

传统的基于图像的故障诊断方法往往是利用图片的颜色、边缘、亮度等信息，通过阈值分割等方式来实现故障诊断。例如，张德国提出了一种基于进化计算的模型参数优化算法，综合图像直方图形状、聚类熵和熵等信息，设计了一种多数性混合目标函数，以实现图像的最优分割^[1]。随着深度学习技术的发展，尤其是 2D 目标检测、语义分割等技术的突破，基于的图像的故障检测方法逐渐成为研究主流。本文简单总结了若干种基于图像的机械故障诊断方法。马天兵等人提出了一

种基于机器视觉与激光融合的多方向矿井刚性罐道变形诊断及其定位方法。使用 CCD 相机、激光发射器、荧光屏、LabVIEW 软件和 PC 组成变形诊断及其定位系统实时采集荧光屏上的光斑图像,对采集的光斑图像进行图像增强、阈值分割、图像匹配跟踪和像素质心计算等处理;在 PC 监控界面显示罐道情况及位置^[2]。张江勇将目前主流的深度学习 2D 目标检测算法应用到动车故障检测中,然后根据实际实验结果,对算法进行改进优化,以提高检测的准确率^[3]。谢冠星在已有的故障检测系统基础上设计了一种基于图像处理的列车部件完整性检测方法来进行列车部件的完整性检测^[4]。韩帅利用机器视觉设计了接触网管帽及 U 型抱箍故障检测算法,作何利用图像处理技术对 4C 图像自动识别中的支撑悬挂装置腕臂直线检测问题、管帽精确定位与故障自动识别问题和 U 型抱箍精确定位与故障自动识别问题,进行了深入的算法研究与改进^[5]。倪晨提出使用以 VGG16 为基础网络的 Faster-RCNN 的网络用于输电线路故障图像的识别与检测,并通过对 RPN 网络中 batch_size 大小的不断调整,确定网络中的正负样本比例,从而达到对网络优化的目的,由此形成输电线路故障检测网络模型^[6]。张子健对高铁接触网两个重要部件,吊弦和绝缘子故障检测问题进行了研究,在吊弦故障检测中,使用深度学习目标检测算法 Faster R-CNN 对正常吊弦与明显故障吊弦进行区分,在正常吊弦的定位结果基础上,采用基于 Hough 变换的直线检测算法判断吊弦是否存在轻微松弛故障。在绝缘子故障检测中,使用深度学习目标检测算法 RRPN 对绝缘子进行带有角度信息的定位,在旋转至水平的绝缘子图像基础上,提出一种同一绝缘子逐片比对的故障判别方法,避免了采集图像一致性差的问题^[7]。诸如上述研究的算法非常多,但是都是基于 2D 图像来检测机械故障,容易出现漏检、误检等情况。本文将使用 3D 点云来检测机械故障,并避免解决漏检、误检等问题。

3.2 点云语义分割算法

本文将基于点云的零件缺失检测问题转化为点云语义分割问题,因此在这里简单介绍现有的基于深度学习的语义分割算法。最早直接输入点云数据进行特征学习的网络是在 CVPR2017 上提出的 PointNet^[8]。与其他架构不同,PointNet 不使用卷积提取特征仅由全连接层组成。网络解决了点云无序性、几何旋转的问题,

实验结果也表明其性能相当可观。虽然是突破性的进展，但尚有许多问题，比如对局部空间特征的感知较差。由 2D-CNN 获取灵感，Qi 等人又提出了 PointNet++^[9]，网络由 PointNet 构成的特征提取块组成，并采用了 MSG、MRG 以及特征传播改进网络架构，输入沿着多分辨率层次以逐渐变大的比例捕获特征。虽然在一些数据集上的结果提高不是很多，但是也提高了架构对于稀疏点的鲁棒性。最早对点云进行有序化处理的一个最突出的例子是 OctNet^[10]。该方法将点云数据存储在浅八叉树中并加以索引，卷积运算符是直接在树的结构上定义的，能够处理非定长数据。类似的，通过使用 KD 树结构索引点云，Klokov 等人提出了一种直接在 KD 树上构造的称为 KD-Network 的深度网络^[11]。与 CNN 一样，网络是前馈但可学习与 KD 树中节点的权重相关联的参数。网络还展示了训练和推理的效率以及可扩展性。虽然高效组织点云被证明有效，但需要额外的数据处理步骤，所以在端到端处理上是存在劣势的。Su 等人提出的一种网络架构 SplatNet^[12]很好地解决了这个问题。其灵感来自 permutohedral 网格^[13]，通过 BilateralConvolution Layer (BCL) 卷积层结构堆叠构造起整个网络，把对点云的组织放到了每一次卷积操作中，实现了端到端的处理。而针对分割任务并不受点云顺序影响的观点，Landrieu 和 Simonovsky 提出了一种基于图的方法很好地反驳了这点，网络采用了一种称为超点图 (SuperpointGraphs, SPG^[14]) 的新型数据结构中。该结构包含一组称为超级点 (Superpoint) 的形状单元其灵感来源于超像素点 (Super Pixel)^[15]。文章认为，通过将数据组织为 SPG，并使用边缘条件卷积^[16]有助于分类和分割。网络不仅在小型模型分割上表现良好，在大规模场景分割上也毫不逊色。最近，CVPR2019 的一篇匿名论文^[17]提出了一种采用图注意力卷积 (Graph Attention Convolution, GAC) 的方法。其引入注意力机制组织搜索点云，使网络具有强大的结构化特征学习能力，对室外场景中易遮挡和残缺的对象 (例如汽车、硬景观、低/高植被) 的特征的区分度有较大提升。但是对于点数量较多、较密集并容易混淆的人工和自然地形，权重分配不能很好发挥作用，因此场景总体识别能力提升不大。

3.3 点云配准算法

点云配准算法是过去基于点云的故障检测算法常用的思路,在本研究中也使用了点云配准的思想。因此,在这里也简答回顾若干经典的点云配准算法。点云配准简单说就是将不同视点扫描的点云拼接在同一个扫描坐标系下,也就是找到两个点云之间的对应关系,然后将一个坐标系下的点云转换到另一个坐标系下。配准的主要过程包括寻找同名点对、确定对应关系、解算变换参数矩阵 R 与 T 。人们对利用原始数据的配准算法已研究了几十年之久,其中经典算法之一就是由 Besl P J 和 McKay N D 提出的最近邻点迭代 (Iterative Closest Point, ICP) 算法,或称最近点迭代配准算^[18]。CP 算法本质上是基于最小二乘的最优匹配算法,它重复寻找对应点和最优刚体变换,直到满足某个收敛准则,迭代终止。除 ICP 算法以外,研究学者也提出许多其他配准算法。David Lowe 提出尺度不变特征变换 (Scale Invariant Feature Transform, SIFT) 算法,该算法将极值点引入到图像检测领域,通过极值提取图像的特征,并由相似性确定特征的对应关系,从而求得最终的刚体变换。官云兰、贾凤海^[19]提出一种新的多视点云数据配准方法,该方法首先构建一个统一扫描网坐标系,计算出相邻两个点云数据间的变换参数矩阵和各站点云数据在统一扫描网坐标系下的坐标,然后结合相邻测站间的公共点,以首站点在统一扫描坐标系下的坐标经非线性变形改正后应与首站实际扫描坐标相等、相邻扫描站公共点经非线性变形改正后坐标应相等为条件,借助非线性变形改正模型,计算出各个点云的非线性变形参数,实现对每站扫描点云数据进行改正,降低累积误差。Jayakorn 等人提出一方向 (Inverse) 点云配准算法,即让标准点云去配准目标点云,这种方式大大减少了计算量。近年来,随着深度学习技术的发展,出现了许多基于深度学习的点云配准算法。PointNetLK^[20]提出了一种循环 PointNet[21]算法来完成刚性的点云配准,这种方法相比传统算法,鲁棒性更强。PointGMM[22]提出综合利用深度学习算法和高斯混合模型来进行点云配准。

四、模型介绍

4.1 模型主要思路

假定在应用场景中，我们可以获得某个待检测区域的标准点云数据（不存在任何故障的数据，其维度为 $N \times 3$ ）和待检测故障点云数据（其维度为 $M \times 3$ ）。我们可以将缺失零件检测问题转化为对标准点云的语义分割问题，即对标准点云的每一个点进行二元分类，0 表示该点在待检测数据中缺失，1 表示该点在待检测数据中完好。为了实现这个目的，我们设计了如下神经网络结构。

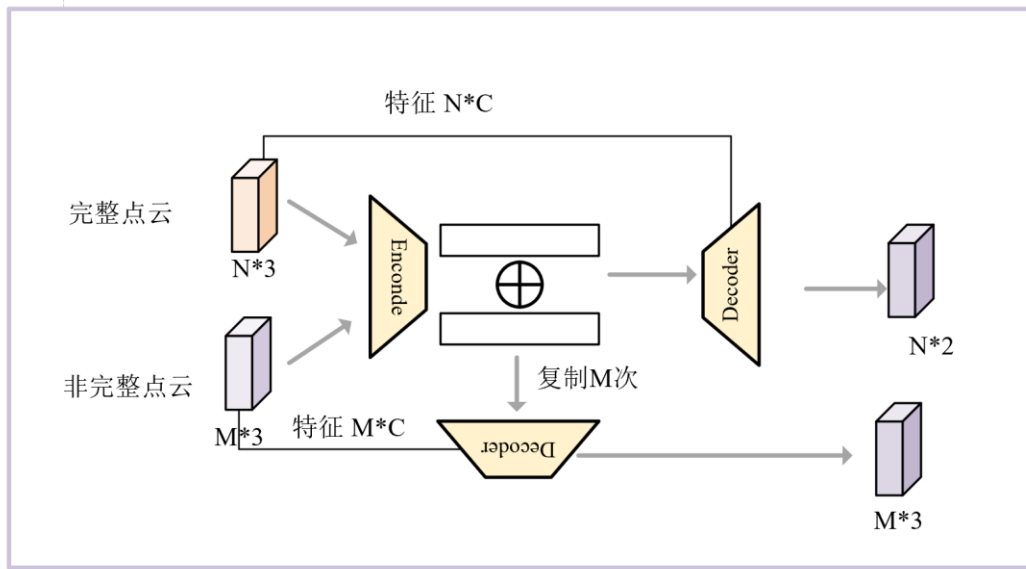


图 2 网络设计基本思想

如图 2 所示，我们使用同一个编码器（encoder）对输入的两个点云进行编码，对于每一个输入的点云，编码器后的输入为一个 $1 \times C$ 的向量（点云编码）和一个 $N \times C$ 或者 $M \times C$ 的矩阵（点特征）。然后，我们用标准点云的编码减去待检测点云的编码，获得另一个 $1 \times C$ 的向量。该向量即可描述两个点云之间的差异，我们把这个向量称为差异向量。在得到改差异向量后，我们将其复制 N 倍后把它同标准点云的点特征矩阵级联起来，获得一个 $N \times 2C$ 的矩阵，该矩阵既反映了标准点云的特征，又反映了两个点云之间的差异。然后基于该矩阵，我们将其输入到一个解码器（decoder）中进行解码，并完成二元分割任务。

在实际的情况中，输入的两个点云由于采集时间、采集设备等因素的不同，往往并不是完全的对齐的，因此我们需要将两个点云进行对齐操作，即点云对齐。

本文采取对待检测点云进行重建的方式来对齐两个点云。具体来说，我们将待检测点云的点云的特征矩阵同复制 M 倍的差异向量级联起来，获得一个 $M \times 2C$ 的特征矩阵。然后，我们使用一个解码器来重建 $M \times 3$ 的一个点云。在训练过程中，我们将使得该点云和标准点云之间的度量距离最小。

通过协同进行语义分割任务和点云配准协同的进行，我们可以让模型学到最能反映该机械区域特点的特征，以及最难反映两个点云之间差异性的特征，从而能够使得模型语义分割的结果中能够检测出标准点云中的缺失点。

4.2 编码器介绍

由于点云存在无序性，因此我们要使用一个能够保证置换不变性的编码器来对点云进行编码，任何一个现有的点云编码器都可以完成该任务，为了简单起见本文采用 PointNet 作为编码器。PointNet 的主要涉及思想如下图所示：

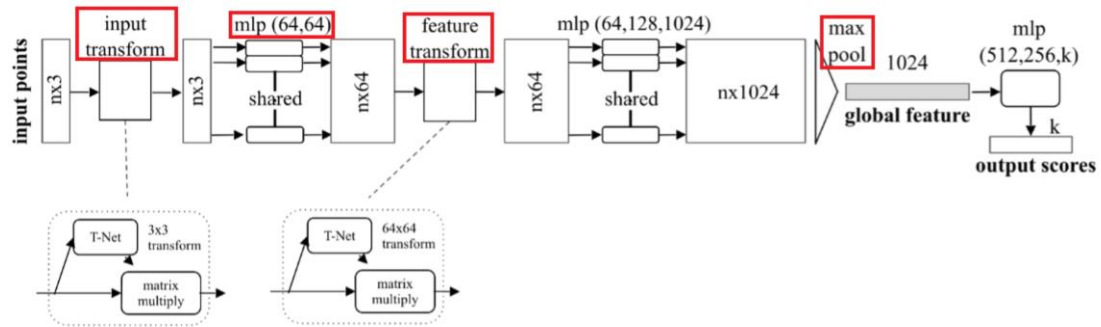


图 3 PointNet 结构

其中，mlp 是通过共享权重的卷积实现的，第一层卷积核大小是 1×3 （因为每个点的维度是 xyz ），之后的每一层卷积核大小都是 1×1 。即特征提取层只是把每个点连接起来而已。经过两个空间变换网络和两个 mlp 之后，对每一个点提取 1024 维特征，经过 maxpool 变成 1×1024 的全局特征。再经过一个 mlp（代码中运用全连接）得到 k 个 score。分类网络最后接的 loss 是 softmax。

4.4 解码器介绍

无论是为了完成点云配准还是点云语义分割，我们都需要将获得的大小为 $N \times 2C (M \times 2C)$ 的特征矩阵进行解码，以进一步提取特征。和编码器一样，解码器同样需要保证点云的置换不变性，在本文中，我们均采用权重共享的多层感知机来对点云进行解码。在语义分割分支中，多层感知机的最终输出为 $N \times 2$ ，每一行表示每个点存在或者缺失的概率。在点云配准分支中，多层感知机的最终输出为 $N \times 3$ ，表示重建后的待检测点云。

4.5 损失函数

为了促使模型能够同时完成语义分割任务和点云配准任务，我们分别为这两个点云设计损失函数。对于语义分割任务，我们使用最为简单的交叉熵损失函数。对于点云配准任务，我们使用如下所示的 chamfer distance 作为重建损失函数。

$$d_{CH}(S, \hat{S}) = \max \left\{ \frac{1}{|S|} \sum_{\mathbf{x} \in S} \min_{\hat{\mathbf{x}} \in \hat{S}} \|\mathbf{x} - \hat{\mathbf{x}}\|_2, \frac{1}{|\hat{S}|} \sum_{\hat{\mathbf{x}} \in \hat{S}} \min_{\mathbf{x} \in S} \|\hat{\mathbf{x}} - \mathbf{x}\|_2 \right\}.$$

五、实验内容

5.1 数据集

由于实际动车故障数据集的保密性，本文使用模拟数据集来验证所设计的算法。我们选择在 ShapeNet 基础上构建模拟数据集。ShapeNet 包含 16,881 个点云物体数据，共 16 类。每个点云包含 2048 个点。每一个点云都可以作为标准点云。然后我们随机选择某点为初始点，其周围 200 个邻居被剔除作为待检测故障数据，并对齐随机旋转和平移来模拟点云没有对齐的情况。最终按照 7:3 划分训练集和测试集。图 5 为数据集的一个示例。左侧为待检测故障数据、右侧为标准点云数据，红色点为缺失点。

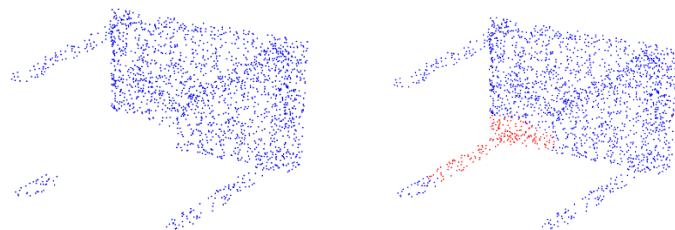


图 5 数据集示例

5.2 模型训练和结果

我们使用 Pytorch 框架来构建模型并训练。使用的优化器为 Adam 优化器。初始学习率为 0.001，每训练 20 个 epoch 我们将学习率衰减一半。我们对模型的每层都使用 Batch norm、并使用了 relu 为激活函数。每轮训练我们都是用点云伸缩、抖动等方式来进行数据扩充。由于我们的目的是检测缺失零件，因此语义分割的准确性是我们的最终检测目标，因此我们选择语义分割的准确率和 mIOU 作为评价指标。模型在 TITAN V GPU 上训练约 10 小时。模型大小为 42.26MB。对测试集的评价结果如下图所示。

```
eval mIoU of data          0.883261
Accuracy is: 0.97698
Class avg accuracy is: 0.92792
Class avg mIOU is: 0.88326
Instance avg mIOU is: 0.88326
```

图 6 测试集测试结果

可以看出，我们的算法能够以非常高的精度来检测部位的缺失。我们在图 6 中展示了部分可视化结果。

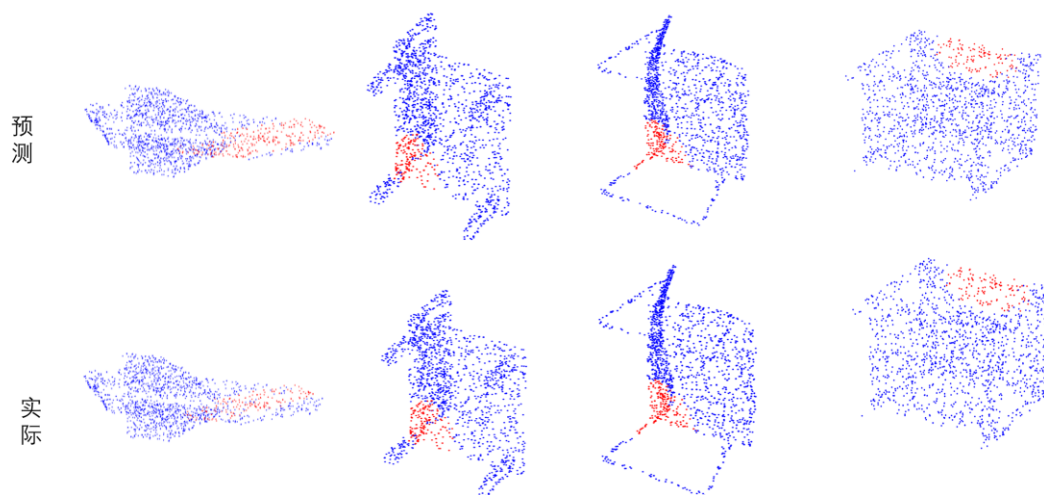


图 6 语义分割结果展示

六、总结

本文设计了一个基于 3D 点云的机械零件缺失检测模型。该模型综合利用了语义分割和点云配准的优势,将两个任务协同到同一个网络中,互相增益,最终完成对缺失零件的检测工作。本文的最核心的思想是,如果使用同一个编码器对两个点云进行编码,那个两个向量之间的差异,可以充分的反映在两个编码后的向量的差上。本文针对所涉及的模型在 ShapeNet 模拟故障数据集上进行了实验,实验表明,我们的方法能够十分有效且高效的检测点云中缺失的零件。

参考文献

- [1]张德国. 基于图像处理的故障检测方法研究[D].西安电子科技大学,2006.
- [2]马天兵,吴强,王鑫泉,王孝东,刘健.基于机器视觉与激光融合的刚性罐道故障定位技术[J].煤矿安全,2020,51(01):134-137.
- [3]张江勇. 基于深度学习的动车关键部位故障图像检测[D].电子科技大学,2019.
- [4] 谢冠星. 基于图像的列车部件完整性检测方法及系统研发[D].长安大学,2019.
- [5] 韩帅. 基于机器视觉的接触网管帽及 U 型抱箍故障检测算法研究[D].石家庄铁道大学,2019.

- [6] 倪晨. 基于卷积神经网络的输电线路故障快速检测与识别[D].内蒙古工业大学,2019.
- [7] 张子健. 面向高铁接触网缺陷检测的智能图像处理关键技术研究[D].浙江大学,2019.
- [8] Charles R Q, Hao S, Mo K, et al.PointNet: Deep learning on point sets for 3D classification and segmentation[C]Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [9] Qi C R, Yi L, Su H, et al.PointNet++: Deep hierarchical feature learning on point sets in a metric space[C]Advances in Neural Information Processing Systems, 2017: 5099-5108.
- [10] Riegler G, Ulusoy A O, Geiger A.OctNet: Learning deep 3D representations at high resolutions[C] Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [11] Klovov R, Lempitsky V.Escape from cells: Deep KD networks for the recognition of 3D point cloud models[C] Proceedings of IEEE International Conference on Computer Vision, 2017.
- [12] Su H, Jampani V, Sun D, et al.SPLATNet: Sparse lattice networks for point cloud processing[C]Proceedings of IEEE Conference on Computer Vision and Pattern Recognition 2018.
- [13] Adams A, Baek J, Davis M A.Fast high- dimensional filtering using the permutohedral lattice[C]//Proceedings of Computer Graphics Forum, 2010.
- [14] Landrieu L, Simonovsky M.Large- scale point cloud semantic segmentation with superpoint graphs[C] Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2018.
- [15] Tighe J, Lazebnik S.SuperParsing: Scalable nonparametric image parsing with superpixels[C]Proceedings of European Conference on Computer Vision, 2010.
- [16] Simonovsky M, Komodakis N.Dynamic edge-conditioned filters in convolutional neural networks on graphs[C] Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [17] Anonymous.Graph attention convolution for point cloud

segmentation[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2019.

[18] 高珊珊.基于三维激光扫描仪的点云配准[D].南京:南京理工大学,2008.

[19] 官云兰,贾凤海.地面三维激光扫描多站点云数据配准新方法[J].中国矿业大学学报,2013(9):880-885.

[20] Vongkulbhisal J , Beñat Irastorza Ugalde, Torre F D L , et al. Inverse Composition Discriminative Optimization for Point Cloud Registration[C]// IEEE/CVF Conference on Computer Vision & Pattern Recognition. IEEE, 2018.

[21] Aoki Y , Goforth H , Srivatsan R A , et al. PointNetLK: Robust & Efficient Point Cloud Registration using PointNet[J]. 2019.

[22] Amir Hertz,Rana Hanocka, Raja Giryes, et. PointGMM: a Neural GMM Network for Point Clouds[C] Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2020.