

基于RGB-D边缘信息融合的接地设备图像语义分割

董景, 薛黎, 许斌, 孙志强, 王琳

常州常供电力设计院有限公司, 江苏 常州

收稿日期: 2023年11月23日; 录用日期: 2023年12月20日; 发布日期: 2023年12月29日

摘要

传统的接地设备安装质量检测方法主要由工作人员采用手工测量和经验判断的方式来检查, 工作量大, 检测效率低且存在经验误差, 无法准确地测量出安装螺栓的参数信息。针对这一问题, 本文提出了一种基于RGB-D边缘信息融合的接地设备图像语义分割方法。首先将获取到的接地设备相连螺栓的可见光和深度图像进行预处理; 然后引入注意力辅助模块帮助提取可见光图像特征和深度图像特征, 并进行特征融合; 通过金字塔池化模块防止过拟合, 最后利用边缘信息辅助监督, 输出最优的预测图像语义分割结果。实验结果表明, 本文方法能够实现对接地设备相连螺栓图像的精确分割, 并将其应用于接地设备的安装质量检测之中。

关键词

语义分割, 可见光与深度图像, 接地设备, 边缘信息, 金字塔池化模块

Semantic Segmentation of Grounding Equipment Image Based on RGB-D Edge Information Fusion

Jing Dong, Li Xue, Bin Xu, Zhiqiang Sun, Lin Wang

Changzhou Changgong Electric Designing Institute Co., Ltd., Changzhou Jiangsu

Received: Nov. 23rd, 2023; accepted: Dec. 20th, 2023; published: Dec. 29th, 2023

Abstract

The traditional method of inspecting the installation quality of grounding equipment is mainly

文章引用: 董景, 薛黎, 许斌, 孙志强, 王琳. 基于 RGB-D 边缘信息融合的接地设备图像语义分割[J]. 电气工程, 2023, 11(4): 189-196. DOI: 10.12677/jee.2023.114021

performed by manual measurement and empirical judgment. The workload is large, the inspection efficiency is low and there are empirical errors. It is difficult to get accurate parameters of the installation bolts. To address this problem, this paper proposes a semantic segmentation method for grounding equipment images based on RGB-D edge information fusion. Firstly, the visible and depth images of the connected bolts of the grounding equipment are preprocessed. Then, an attention assistance module is introduced to help extract visible image features and depth image features, and perform feature fusion. The pyramid pooling module is used to prevent overfitting, and the edge information is used to assist supervision, and the optimal predictive image semantic segmentation results are output. Experimental results show that this method can achieve accurate segmentation of grounding equipment and apply it to the inspection of the quality of the grounding equipment installation.

Keywords

Semantic Segmentation, Visible and Depth Images, Grounding Equipment, Edge Information, Pyramid Pooling Module

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

传统的螺栓安装质量检测方法通常使用经验观察和尺子测量等方式。这些检测方法的效率低下，且易使工人产生视觉疲劳，进而导致误检、漏检，存在安全隐患[1]。相较于传统的验收方法，使用图像语义分割技术进行螺栓安装验收具有自动化、精度高、花费小的优点，能够使得对接地设备的安装验收更加智能化。

传统的图像语义分割方法使用包括基于阈值、候选区域、聚类等机器学习的方法来构建像素分类器[2]，从而完成图像的语义分割，但由于硬件设备的计算能力有限，传统方法在分割图像的颜色、形状、纹理等低级语义特征时通常效率不高、分割时间较长，并且面对颜色相近且背景复杂的螺栓图像分割任务时，图像的语义分割精确度较低。

随着深度学习技术的兴起，越来越多的学者选择利用深度信息去辅助可见光图像进行图像分割，结合基于深度学习的端到端深度卷积神经网络(Deep Convolutional Neural Networks, DCNN) [3]进行图像语义分割，有效的提高了分割的效率和精度。文献[4]在 DeepLabv3 网络的空洞金字塔池化层中引入了一个全局池化层，更好地提取了图像级特征，能够从螺栓关键区域分割出螺母和螺杆。文献[5]使用了 SSD512 定位算法和 U-net8 模型的实现了对紧顶螺栓图像中螺栓和螺母的分割。文献[6]将阈值分割和 Rammer 算法结合对螺栓轮廓图像进行分割，并利用质点云法确定螺栓轮廓。文献[7]提出了一种基于 Faster-RCNN 卷积神经网络的螺栓松动检测方法，能够更加快速地检测出图像中的螺栓。文献[8]构建了一个协作反卷积神经网络来联合建模，将深度图像作为可见光图像的附加通道输入网络进行语义分割。然而上述方法，要么不能充分地融合深度图像和可见光图像的特征，要么就无法适用于纹理相似且背景复杂的任务场景中，而且还容易受到噪声干扰。

对此，本文提出了一种基于 RGB-D 边缘信息融合的接地设备图像语义分割方法，充分地融合可见光图像特征以及深度图像信息，并利用边缘信息，大幅提高了螺栓图像分割的准确性。首先对所获取的接

地设备相连螺栓的可见光图像和深度图像进行预处理；然后通过编码模块提取深度图像特征和可见光图像特征，并进行多特征融合；最后利用边缘信息辅助监督，输出最优的预测图像语义分割结果。实验结果证明，本文提出的算法在对颜色纹理相近且边缘轮廓模糊的复杂图像情况下分割准确度较高。对精确分割螺栓图像，保障接地设备的安装质量检测方面有着极大的应用价值。

2. 基于 RGB-D 边缘信息融合的接地设备图像语义分割

本文提出的基于 RGB-D 边缘信息融合的接地设备图像语义分割网络架构如图 1 所示。首先获取包含接地设备相连螺栓的可见光图像和深度图像，对进行预处理，其次将经过预处理的可见光图像和深度图像输入进双路编码 - 解码模块[9]中，其中编码模块有两个平行排列的残差网络 ReSNet [10]组成，采用卷积和池化操作分别得到相应的可见光图像特征与深度图像特征；添加空间金字塔池化层并引用注意力辅助模块获取更加丰富的特征图；将经过特征融合后得到的全局特征图输入至解码模块，此模块由语义分支和边缘分支两个分支构成，语义分支主要负责从可见光图像和深度图像中提取语义信息，并利用这些信息生成语义预测结果。边缘分支主要通过分析可见光图像和深度图像的边缘，确定图像中不同区域之间的边缘，并输出边缘预测结果；最终结合边缘信息进行监督训练，得到最优的预测图像语义分割结果。

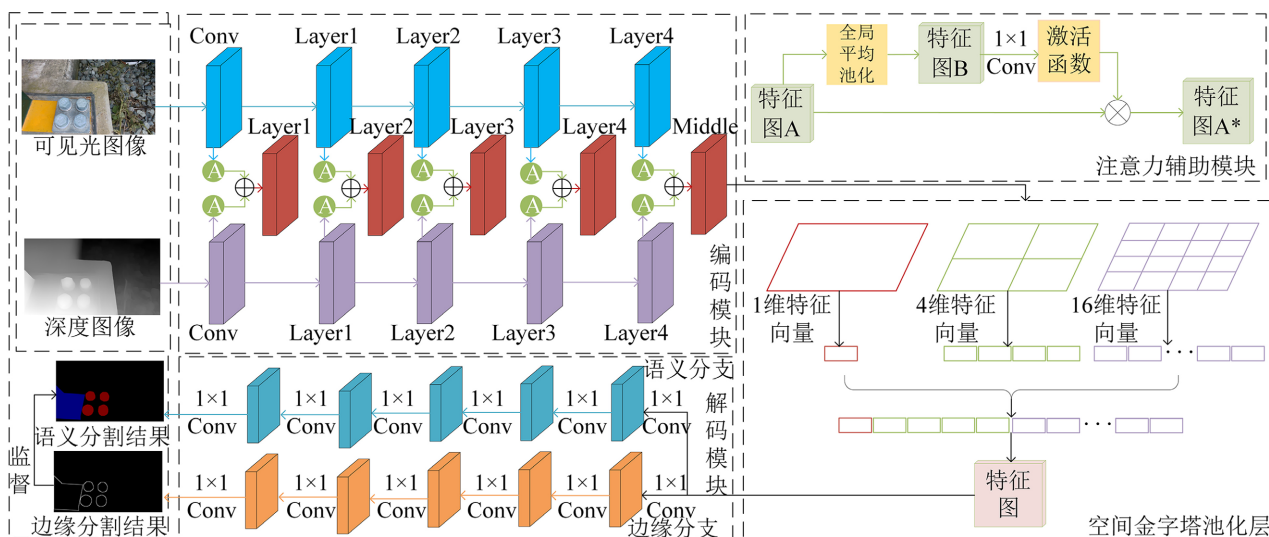


Figure 1. Network architecture of RGB-D image semantic segmentation based on edge information guidance

图 1. 基于 RGB-D 边缘信息融合的接地设备图像语义分割网络架构

2.1. 图像预处理模块

传统的语义分割算法在图像处理部分较为粗糙，然而螺栓图像存在对比度低、颜色衰退以及细节模糊等现象，因此需要对螺栓图像进行预处理。图像预处理可以消除图像中无关的信息，恢复有用的真实信息，增强有关信息的可检测性和最大限度地简化数据，从而改进图像分割的可靠性。首先对所获得的可见光图像和深度图像，使用高斯滤波器进行去噪处理；其次利用校正网络(RSN)对去噪之后的图像进行畸变校正，校正网络通过学习训练数据中的畸变模式，预测输入图像的几何畸变，并生成校正后的图像；最后将矫正后的图像进行归一化操作，即将图像的像素值缩放到[0, 1]之间。

2.2. 注意力辅助模块

目前大多数的编码 - 解码模块大都只是简单的从输入图像中提取特征。如文献[11]，使用两个编码

器分别从可见光图像和深度图像中提取特征，并在上采样之前或期间结合两者的特征。另一个类如文献[12]，只是在下采样阶段融合可见光和深度特征。前者不能充分结合图像的特征信息，而后者往往因为融合分支而丢失原始的可见光和深度分支。因此本文在双路编码模块的每一层网络后面引用注意力辅助模块[13] (Attention Assistance Module, AAM)，得到每一层网络新的特征图，从而丰富特征图像的语义特征。

从编码模块得到的某一特征图 $A \in R^{C \times H \times W}$ ，首先对特征图 A 进行全局平均池化，得到一个大小为 $C \times 1 \times 1$ 的特征图 B ，第 m 个通道特征图表示如公式(1)所示：

$$B_m = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W A_m(i, j) \quad (1)$$

其中 $i \in \{1, 2, 3, H\}$ ， $j \in \{1, 2, 3, W\}$ ， H ， W 依次代表特征图 A 的高度和宽度。对于特征图 A ， R 表示特征图维度， C 为通道数。

保持特征图 B 的通道数 C 不变，加入一元卷积层用来确定通道间的权重分布。然后，在卷积结果上使用 sigmoid 激活函数得到约束权重向量的值，并使其保持在 $[0, 1]$ 之间。 σ 是 sigmoid 激活函数，此过程如下所示：

$$S = \sigma(F(B)) \quad (2)$$

最后，将约束值与输入特征图 A 进行叉乘，得到一个包含更多有用信息的特征图 A^* ，其表达式为：

$$A^* = A \times \sigma(F(B)) \quad (3)$$

将可见光特征图像和深度特征图像输出映射按照一定比率输入注意力辅助机制融合语义特征，得到全局特征图。

2.3. 金字塔池化模块

现有的深度卷积神经网络要求尺寸固定的输入图像，这限制了输入图像的纵横比和尺度。当输入图像的尺寸不一致时，大多是通过裁剪或通过扭曲[14]将输入图像拟合到固定尺寸。这种尺度变化造成的不必要的内容丢失和失真，会使识别精度降低。此外，如果要求尺寸发生变化时，预设的尺寸可能又要改变，大大的加深了工作量。因此本文方法将从编码器输出的特征图输入至空间金字塔池化层(Spatial Pyramid Pooling, SPP) [15]，不同大小的目标候选区域的特征图可以被转化为固定大小的特征图，且使用可变尺寸图像进行训练增加了尺度不变性，防止过拟合。

该空间金字塔池化层由 3 个最大池化层组成[16]。对于编码器输入的任意大小的特征图，3 个最大池化层分别以大小为 4×4 、 2×2 和 1×1 的网格将特征图分为 16、4、1 块，然后在每个块上最大池化，提取相应特征。第 1 个池化层提取 16 维特征向量，第 2 个池化层提取 4 维特征向量，第 3 个池化层提取 1 维特征向量，最后将 3 个池化层提取的特征融合，得到 $16 + 4 + 1 = 21$ 维特征向量，从而使任意大小特征图都能转化为 21 维的特征向量，全连接层不再对输入图像的大小有所限制。

2.4. 损失函数

在螺栓图像语义分割中，选择一个合理的损失函数去提高准确度尤为关键，同时也可以减少网络的训练时间，帮助网络更好地训练。本文利用 Dice 系数损失函数[17]对语义分割标注真值 Y_1 和预测的图像语义分割结果 Y_2 进行结合分析，得到两者之间的误差 X_1 ，具体公式如下：

$$X_1 = 1 - 2 \times \frac{Y_1 \cap Y_2}{Y_1 \cup Y_2} \quad (4)$$

利用 Boundary LOSS [18] 损失函数对训练图像的边缘信息真值 Y_3 和预测的图像边缘分割结果 Y_4 进行结合分析, 得到两者之间的误差 X_2 , 具体公式如下:

$$X_2 = \max\{0, 1 - Y_3 Y_4\} \quad (5)$$

将由两种不同的损失函数所得到的误差相加得到所需图像语义分割网络的总误差, 具体公式如下:

$$L = A_1 X_1 + A_2 X_2 \quad (6)$$

在公式(6)中 L 表示整个语义分割网络的总误差; X_1 , X_2 分别表示所述预测语义结果与所述先验的语义信息之间的误差以及边缘预测结果与边缘信息之间的误差; A_1 , A_2 为误差 X_1 , X_2 在总误差中所占的比例系数, 两者的大小因语义信息和边缘信息对语义分割网络的重要性而定。利用所得到的误差, 使用随机梯度下降算法去不断的调整参数, 使得预测信息和实际数据的误差最小, 得到一个准确率高的预测结果。

3. 实验结果及分析

针对实际场景下接地设备相连螺栓图像具有不同的类型情况, 本文使用 ZED2 双目相机拍摄了多种场景下的接地设备相连螺栓图像, 并制作数据集。其中, 数据集共包含接地设备相连螺栓 RGB-D 图像 1300 组。部分 RGB-D 图像以及人工标注的真值图像如图 2 所示。训练集与测试集按 7:3 的比例进行划分。其中训练集包含 910 组接地设备相连螺栓的 RGB-D 图像以及真值图像; 测试集包含 390 组接地设备相连螺栓的 RGB-D 图像以及真值图像。

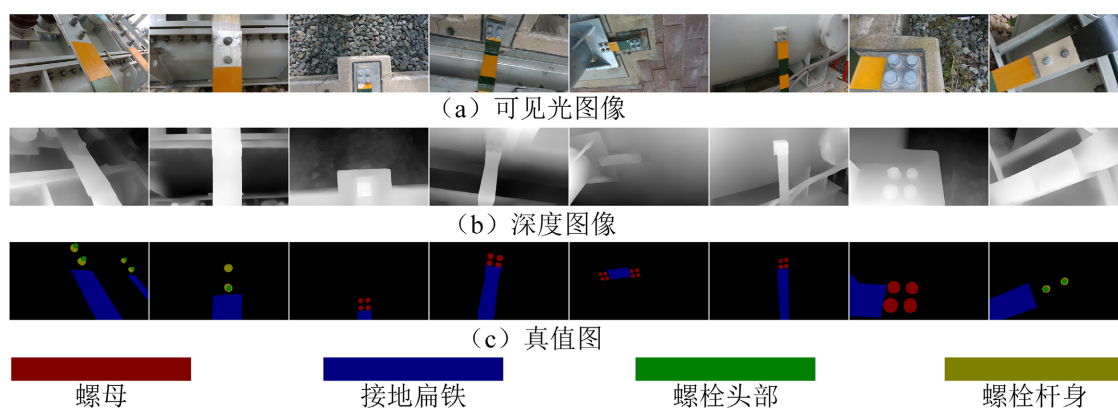


Figure 2. Some RGB-D images and manually marked truth images.

图 2. 部分 RGB-D 图像以及人工标注的真值图像

3.1. 语义分割结果预测实验

本文算法针对不同类型的接地设备相连螺栓图像进行语义分割实验对比, 并与深度学习算法 SGNet [19]、ACNet [13]、FRNet [20] 与 ESANet [21] 进行比较, 以验证本文所提出的算法的分割效果与精度。对测试集中的 100 组接地设备相连螺栓的 RGB-D 图像进行语义分割结果预测实验, 不同算法下接地设备相连螺栓的语义分割结果如图 3 所示, 其中图 3(a) 为根据接地设备相连螺栓的类型以及拍摄角度分为四种情况; 图 3(b) 为原图中经过人工标注的真实接地设备相连螺栓区域, 作为各个算法识别效果的参考基准。

图 3(c)~(g) 分别为 SGNet、ACNet、FRNet、ESANet 和本文算法对接地设备图像的语义分割结果。SGNet 算法使用的空间信息引导在处理具有复杂背景或者细微边缘的图像时, 会出现分割不准确的现象, 如在处理 B 组图像, SGNet 算法在螺栓边缘分割效果存在“一刀切”的现象, 对螺栓细节分割鲁棒性和表征

能力较差；对于 C 组和 D 组图像，SGNet 算法分割结果边缘线条与轮廓模糊，算法精度低。ACNet 算法更加注重全局特征而非细节特征，导致在处理颜色相近或边缘模糊的区域时会出现错误分类，如 A 组和 B 组图像，ACNet 算法在接地扁铁边缘误识别像素点较多，在颜色较为相近之处出现错误分割的问题；对于 C 组和 D 组图像，ACNet 算法在螺栓与螺栓杆身两区域像素相近时存在部分像素分类的错误，没有精确的分割出螺栓与螺栓杆身，出现了错误的分割。FRNet 和 ESANet 在面对相似纹理和复杂场景下，分割效果较差，存在分割边缘不整齐甚至分割错误的情况，如 B 组和 D 组图像。相比之下，本文算法的预测图像最接近真值图像，对复杂背景中的接地设备相连螺栓的识别精度更高。尤其是在螺栓边缘部分，通过最后边缘分支的监督，使得本文算法在接地设备边缘的分割效果要优于其它算法，从而证明了本文算法具有较高的分割性能。此外本文算法从具有较高的鲁棒性，受背景环境之中的噪声，光照变化影响较小，算法稳定性较高。

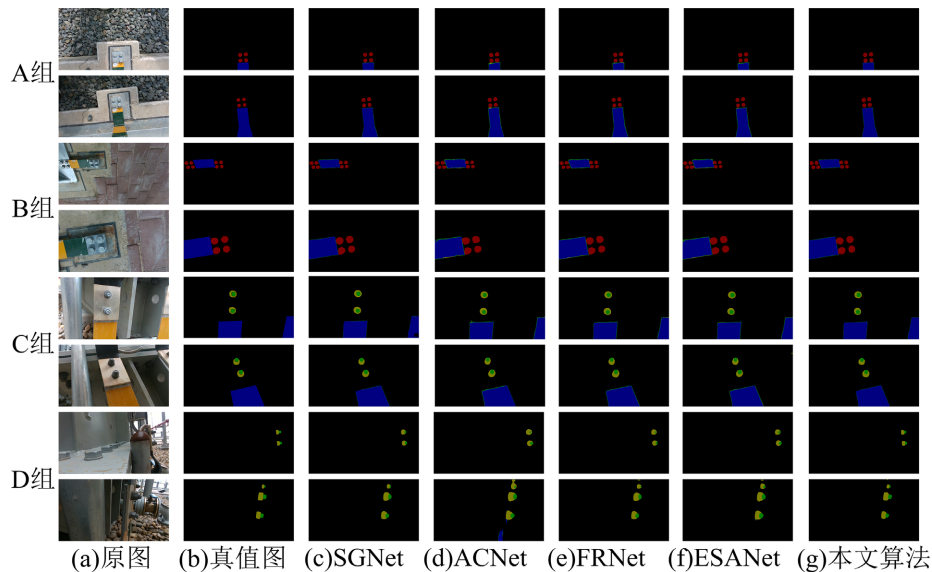


Figure 3. Semantic segmentation results of grounding equipment images with different algorithms
图 3. 不同算法的接地设备图像语义分割结果

3.2. 定量分析

3.2.1. 评价指标

本文选取 Acc [22], IoU [23]作为语义分割算法的性能评价指标。实验中各个算法对 100 组测试图像均做了多次重复实验，结果取平均值。

1) Acc (Accuracy): Acc (准确率)的计算方式如式(7)所示。准确率表示预测正确的样本占所有样本的比例，其中预测正确的可能有正样本也可能有负样本：

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

式中 TP (True Positive)表示本身是正样本，预测也是正样本；预测正确 TN (True Negative)表示本身是负样本，预测也是负样本；预测正确 FP (False Positive)表示本身负样本，但是预测为正样本；预测错误 FN (False Negative)表示本身是正样本，但是预测为负样本，预测错误。

2) IoU (Intersection Over Union): IoU (交占比)用来表示预测框和真实框之间的重叠率，衡量的是预测框与实际框之间的相近程度。记标记框为 A ，检测框为 B ，IOU 的计算公式如下：

$$IoU = \frac{A \cap B}{A \cup B} \quad (8)$$

式中分子部分表示 A 框与 B 框的重叠部分的面积, 分母部分表示 A 框与 B 框的面积总和。由式可得, IoU 的值在 $[0, 1]$ 之间。因此当 IoU 越接近 1 时就说明 A 框与 B 框的重合程度越高, 模型精确度越高, 预测越准确, 反之则越差。

3.2.2. 评价指标实验结果

本文算法与 SNet、ACNet、FRNet、ESANet 进行了比实验, 评价指标如表 1 所示。相较于其他算法, 本文算法在各项评价指标中都表现最优。根据实验结果可知: 本文算法的准确度更高, 对于接地设备相连螺栓图像的分割效果更加优秀。

Table 1. Quantitative analysis of bolt image segmentation results

表 1. 螺栓图像分割结果定量分析

Methods	螺栓头部		螺栓杆身		螺母		接地扁铁		mAcc	mIoU
	Acc	IoU	Acc	IoU	Acc	IoU	Acc	IoU		
SNet	0.9185	0.8544	0.7978	0.5648	0.9301	0.7727	0.9923	0.9016	0.9097	0.7734
ACNet	0.9243	0.8307	0.7768	0.3915	0.9412	0.5867	0.9827	0.8704	0.9063	0.6698
FRNet	0.9123	0.8490	0.7923	0.4228	0.9338	0.6861	0.9852	0.8919	0.9059	0.7125
ESANet	0.9155	0.8462	0.8316	0.4157	0.9272	0.6533	0.9767	0.8852	0.9128	0.7001
本文算法	0.9333	0.8505	0.8345	0.5861	0.9506	0.8552	0.9955	0.8979	0.9285	0.7974

4. 结论

针对传统的接地设备验收方式效率低下, 且接地设备之间粘连以及颜色纹理相近等问题, 提出了一种基于 RGB-D 边缘信息融合的接地设备图像语义分割方法, 模型输入为可见光图像和深度图像, 通过图像预处理增强有用信息, 通过双路残差网络提取图像特征, 引入注意力辅助模块增强二者的特征融合, 采用空间金字塔模块对特征图进行多尺度处理, 最后对输出的特征图进行双分支解码, 得到语义预测结果和边缘预测结果, 并使用边缘预测结果来监督优化语义信息, 输出最优的语义分割结果。实验结果表明, 本文所提出的方法在变电站接地设备场景语义分割任务中具有较好的表现, 能够较好地分割不通的接地设备, 对于接地设备的边缘有着较好的分割结果, 可以应用于接地设备的安装验收任务中。

致 谢

论文作者感谢国网江苏省电力有限公司常州供电分公司对本课题的资助。

基金项目

国网江苏省电力有限公司常州供电分公司科技项目资助。

参考文献

- [1] 王勇, 袁啸阳, 陈铎, 等. 基于多任务卷积神经网络的轨道车辆螺栓异常检测方法[J]. 铁道车辆, 2020, 58(5): 29-32.
- [2] 马文琪, 石颀, 吴宏杰. 深度卷积神经网络语义分割综述[J]. 微电子学与计算机, 2023, 40(9): 55-64.
- [3] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
- [4] 王宝丽, 于龙. 中低速磁浮 F 轨紧固螺栓松动故障检测[J]. 信息技术, 2019, 43(8): 88-92, 97.

- [5] 程敦诚, 王倩, 吴福庆, 等. 基于深度学习的接触网顶紧螺栓状态智能检测[J]. 铁道学报, 2021, 43(11): 52-60.
- [6] 周靖, 刘煜, 霍林生. 基于机器视觉的螺栓松动旋转角度测量[J]. 机械设计与研究, 2021, 37(4): 159-162, 175.
- [7] Zhang, Y., Sun, X., Loh, K.J., et al. (2020) Autonomous Bolt Loosening Detection Using Deep Learning. *Structural Health Monitoring*, **19**, 105-122. <https://doi.org/10.1177/1475921719837509>
- [8] Liu, J., Wang, Y., Li, Y., et al. (2018) Collaborative Deconvolutional Neural Networks for Joint Depth Estimation and Semantic Segmentation. *IEEE Transactions on Neural Networks and Learning Systems*, **29**, 5655-5666. <https://doi.org/10.1109/TNNLS.2017.2787781>
- [9] Noh, H., Hong, S. and Han, B. (2015) Learning Deconvolution Network for Semantic Segmentation. *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, 7-13 December 2015, 1520-1528. <https://doi.org/10.1109/ICCV.2015.178>
- [10] He, K., Zhang, X., Ren, S., et al. (2016) Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [11] Cheng, Y., Cai, R., Li, Z., et al. (2017) Locality-Sensitive Deconvolution Networks with Gated Fusion for RGB-D Indoor Semantic Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 3029-3037. <https://doi.org/10.1109/CVPR.2017.161>
- [12] Jiang, J., Zheng, L., Luo, F., et al. (2018) Rednet: Residual Encoder-Decoder Network for Indoor RGB-D Semantic Segmentation. <https://doi.org/10.48550/arXiv.1806.01054>
- [13] Hu, X., Yang, K., Fei, L., et al. (2019) Acnet: Attention Based Network to Exploit Complementary Features for RGBD Semantic Segmentation 2019 *IEEE International Conference on Image Processing (ICIP)*, Taipei, 22-25 September 2019, 1440-1444. <https://doi.org/10.1109/ICIP.2019.8803025>
- [14] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, **60**, 84-90. <https://doi.org/10.1145/3065386>
- [15] He, K., Zhang, X., Ren, S., et al. (2015) Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **37**, 1904-1916. <https://doi.org/10.1109/TPAMI.2015.2389824>
- [16] Emlek, A. and Peker, M. (2023) P3SNet: Parallel Pyramid Pooling Stereo Network. *IEEE Transactions on Intelligent Transportation Systems*, **24**, 10433-10444. <https://doi.org/10.1109/TITS.2023.3276328>
- [17] Chen, L.C., Papandreou, G., Kokkinos, I., et al. (2017) DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **40**, 834-848. <https://doi.org/10.1109/TPAMI.2017.2699184>
- [18] Wu, D., Guo, Z., Li, A., et al. (2023) Conditional Boundary Loss for Semantic Segmentation. *IEEE Transactions on Image Processing*, **32**, 3717-3731. <https://doi.org/10.1109/TIP.2023.3290519>
- [19] Chen, L.Z., Lin, Z., Wang, Z., et al. (2021) Spatial Information Guided Convolution for Real-Time RGBD Semantic Segmentation. *IEEE Transactions on Image Processing*, **30**, 2313-2324. <https://doi.org/10.1109/TIP.2021.3049332>
- [20] Zhou, W., Yang, E., Lei, J., et al. (2022) FRNet: Feature Reconstruction Network for RGB-D Indoor Scene Parsing. *IEEE Journal of Selected Topics in Signal Processing*, **16**, 677-687. <https://doi.org/10.1109/JSTSP.2022.3174338>
- [21] Seichter, D., Köhler, M., Lewandowski, B., et al. (2021) Efficient RGB-D Semantic Segmentation for Indoor Scene Analysis. 2021 *IEEE International Conference on Robotics and Automation (ICRA)*, Xi'an, 30 May-5 June 2021, 13525-13531. <https://doi.org/10.1109/ICRA48506.2021.9561675>
- [22] Valverde-Albacete, F.J. and Peláez-Moreno, C. (2014) 100% Classification Accuracy Considered Harmful: The Normalized Information Transfer Factor Explains the Accuracy Paradox. *PLOS ONE*, **9**, e84217. <https://doi.org/10.1371/journal.pone.0084217>
- [23] Rezatofghi, H., Tsoi, N., Gwak, J.Y., et al. (2019) Generalized Intersection over Union: A Metric and a Loss for Bounding Box Regression. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, 15-20 June 2019, 658-666. <https://doi.org/10.1109/CVPR.2019.00075>