

基于改进 Deeplabv3 + 网络的线缆表面缺陷检测研究^①

陈 亮^② 杨贤昭^③ 刘惠康

(武汉科技大学冶金自动化与检测技术教育部工程研究中心 武汉 430081)

摘 要 为了提高线缆表面缺陷检测正确率,本文提出一种改进 Deeplabv3 + 网络的图像分割方法并将其应用于线缆表面缺陷检测。该方法基于 Deeplabv3 + 网络骨架不变,将空间金字塔结构由 4 个空洞卷积改为 8 个空洞卷积并在其后增加 1×1 的卷积环节;同时在解码融合后用一个并联结构来减少整个网络传输过程的信息丢失。利用改进的算法对线缆表面缺陷图片数据集训练和测试,结果表明改进算法在准确度和平均交并比(MIOU)较原始的 Deeplabv3 + 分析效果更好;相较于边缘分割和阈值分割等算法,改进算法提高了线缆表面缺陷检测的准确率。

关键词 线缆表面缺陷; Deeplabv3 + ; 空洞卷积; 并联结构; 准确度

0 引 言

近年来,我国智能机器人、通讯、智能家居等行业迅速发展,电缆作为一种传输电(磁)能、信息和实现电磁能转换的重要线材产品,是这些行业发展的必需品。线缆的发展成为这些行业的铺垫,其使用频率逐渐增加,这需要严格把关线缆质量。由于线缆外表主要起绝缘、防水、防潮等作用,所以线缆外表检测尤为重要,同时线缆外表也是线缆质量的一个重要指标。它的缺陷主要由于生产设备、生产方式、安装方式、人为因素等原因产生,这些缺陷会造成产品质量得不到保障、施工人员人身安全受到影响、商业经济受到损失等危害。线缆表面缺陷相比墙面和沥青地面产生的裂纹缝隙更小,采用同样方式采集的图像信息更少,得到的分割结果相对较差,这将需要获取更多的图像信息。线缆表面后期检测一般情况下处于高空中,相比于布匹、轨道缺陷的检测,线缆表面的图像信息不易获取。

目前线缆表面缺陷的主要检测方式为人工检

测,但这样的方式误检率太高,不能满足一些特殊行业的要求。而且如果针对处于高空中的线缆,采用人工检测将会存在安全隐患,因此,为了降低检测目标的误检率和保证安全性,利用机器视觉对图像进行判断是十分重要的。

国内外已有部分学者基于机器视觉对线缆缺陷进行研究,其中乔湘洋等人^[1]利用改进的灰度均值和改进的双边滤波与原缺陷图像进行差分,实现图像分割,使其在复杂环境也能被很好地处理。王海芳等人^[2]通过改进高斯滤波算法来建立自适应模板,比较原图和模板的 Pearson 相关系数来判断缺陷。张俊^[3]运用线缆图像的灰度均值垂直投影曲线与其二次拟合曲线方差均值来对图像进行判断。Gao 等人^[4]利用数学几何和线性插值法将电缆转换为平面,再根据图中针孔灰度特征用灰度阈值和梯度幅度来进行判断。同时也有学者使用深度学习的方式来处理裂纹,其中 Song 等人^[5]用 YOLO 算法来检测具体的缺陷位置,后期对具体位置做出量化,但这是针对混凝土裂纹来进行处理。以上处理方式一部分是基于非学习型图像算法,其检测速度上较好,

① 国家重点研发计划(2017YFC0805100)和国家自然科学基金(61703314)资助项目。

② 男,1994年生,硕士生;研究方向:图像识别;E-mail: 1183299896@qq.com

③ 通信作者,E-mail: yangxianzhao@wust.edu.cn

(收稿日期:2020-10-09)

但在某些复杂情况下其准确性相对较低;另一部分是基于深度学习算法处理混凝土裂纹,混凝土裂纹相较于本文线缆表面缺陷更容易被识别,有更显著的裂纹缺陷。基于这种情况本文考虑采用改进 Deeplabv3+ 网络对线缆表面缺陷进行识别,目前有很多学者对 Deeplabv3+ 进行研究,Deeplabv3+ 的原型是 Deeplabv1。Deeplabv1 是采用深度卷积神经网络(deep convolutional neural networks, DCNNs)实现对图像的分割。Chen 等人^[6]在 Deeplabv1 的基础上进行多尺度的采样,形成了空间金字塔结构(atrous spatial pyramid pooling, ASPP),产生了 Deeplabv2。薛飞等人^[7]指出,相比于 Deeplabv2, Deeplabv3 改进了串联结构,形成更深结构的空洞卷积,从而进一步提升平均交并比(mean intersection over union, MI-OU)的值。Zhang 等人^[8]在 Deeplabv3 的编码基础上添加了一个解码过程,形成 Deeplabv3+, 使更多的图像信息被利用。其中有一些学者对 Deeplabv3+ 进行了改进, Bhakti 等人^[9]改变空间金字塔结构降

低膨胀率,采用 MoblieNetV2 轻量化架构,使处理图像的速度加快。王亚刚等人^[10]在解码器部分采用 3 层可分离卷积,结合对数据集做预处理,使其结果较好。刘致驿等人^[11]将空间金字塔结构中除空洞率为 1 外的其他 3 个空洞卷积两两连接,取得了较好的效果。

1 基于改进 Deeplabv3+ 网络的线缆缺陷分割算法

1.1 Deeplabv3+ 算法

Deeplabv3+ 是图像分割的一种算法,它能对图像进行像素级分割,这使得它在图像分类上效果较好,它比 Deeplabv3、Deeplabv2、Deeplabv1 增加了解码器结构,形成了编码器-解码器结构,这样能收集更多的像素信息,使分割出来的图像准确率更高。图 1 所示是 Deeplabv3+ 网络对图片识别的过程。

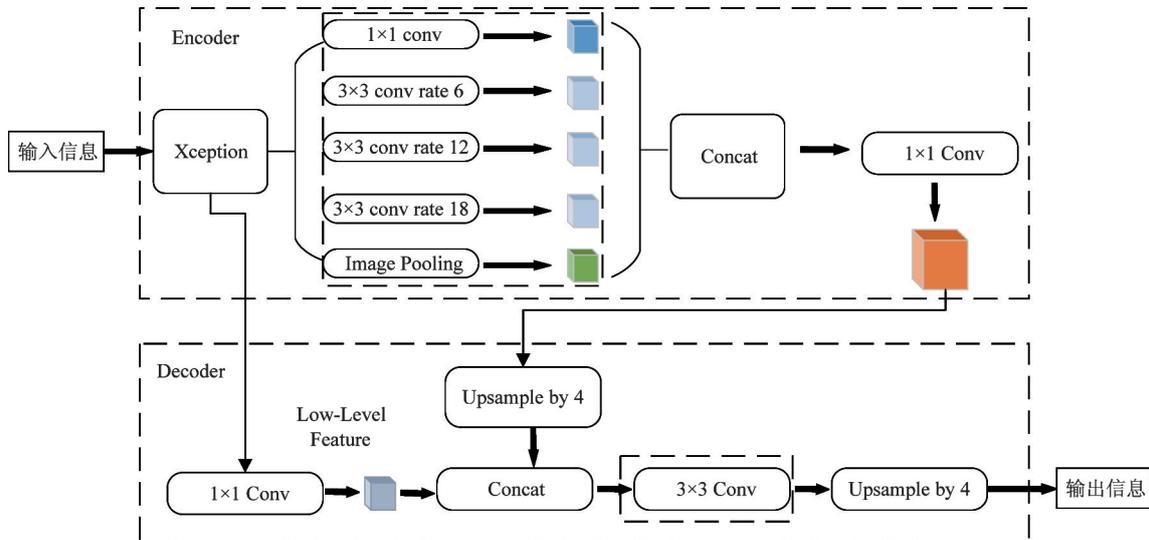


图 1 Deeplabv3+ 网络结构图

图中首先经过编码器部分,这部分先经过主干网络 Xception,其中主干网包括 65 层,引用残差网络结构,将可分离卷积进行串联,形成深度可分离卷积,输出信号进入空间金字塔结构。具体的网络层分布如图 2 所示。

从图 1 中可以看出,这里是 4 个空洞卷积,膨胀率分别是 1、6、12、18,与池化层一起组合成空间金

字塔结构(ASPP)。假设 $F_k^r(x)$ 表示卷积操作,其中下标 k 表示卷积核大小,上标 r 表示空洞率的大小,用 $G(x)$ 表示池化操作,那么 ASPP 的输出信息如式(1)所示。

$$Y = F_1^1(x) + F_3^6(x) + F_3^{12}(x) + F_3^{18}(x) + G(x) \quad (1)$$

ASPP 输出信息 Y 经过 1×1 的卷积核将通道

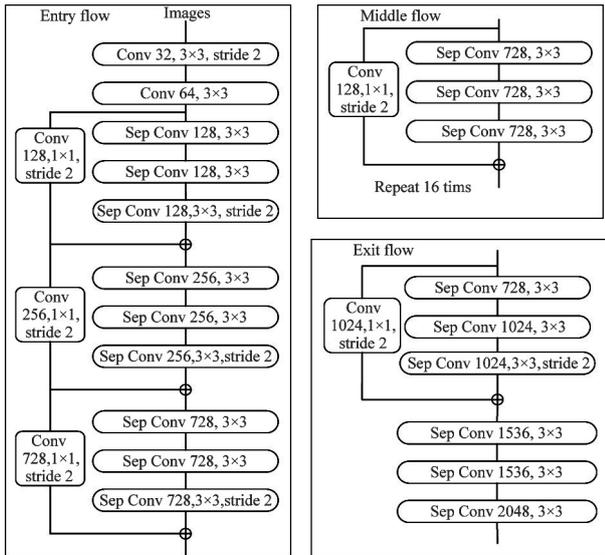


图2 Xception网络结构图

变成一个,这是 Deeplabv3+ 的编码部分。

解码器部分从主干网络 Xception 里面提取低信息的图片,这样处理使分割出来的图片边缘效果更

加清晰。在编码器中出来的图像信息进行4倍双线性插值上采样来提高图片的分辨率,然后将 Xception 里面提取的信息和上采样的图像信息进行组合,输出信号经过 3×3 的卷积核,最后将信息通过4倍双线性插值上采样得出分割的图像。

由于本文的数据集是线缆表面缺陷,有些线缆表面缺陷是小裂纹,原始 Deeplabv3+ 的 ASPP 的空洞率间隔大会导致图像分割断层,且双线性插值上采样会造成信息量少,使图像信息获取不完整。

1.2 改进 Deeplabv3+ 算法

针对上一节提出的问题,本文在 Deeplabv3+ 的基础上进行两方面的改进。

首先,原始空洞卷积的卷积率间隔较大,使输出的图像变得很稀疏,导致一些局部信息丢失,为了解决这个问题本算法提高了 ASPP 里面空洞卷积的卷积率密度,具体改进如图3所示。

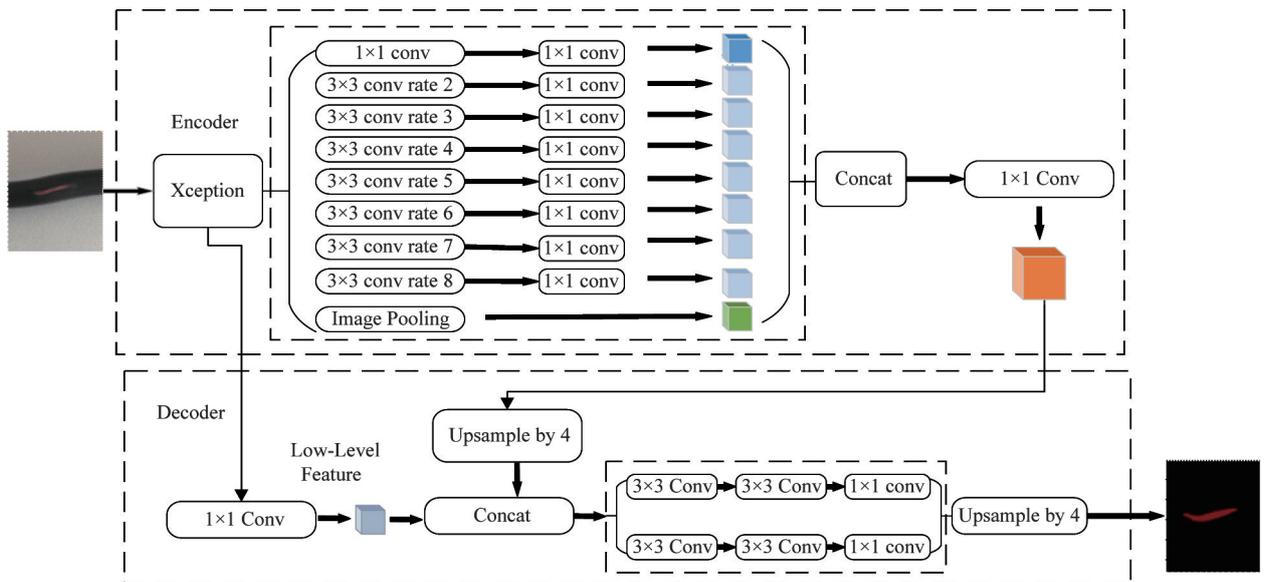


图3 本文算法结构图

从图3中可以看出,将ASPP中的空洞卷积增加到8个来获取更多图像信息,这样能尽量减少信息的丢失,确保识别图片的相对准确性。相比于 Deeplabv3+, ASPP 输出信息如式(2)所示。

$$Y_1 = F_1^1(x) + \sum_{i=2}^8 F_3^i(x) + G(x) \quad (2)$$

比较改进前和改进后 ASPP 输出的信号 Y 和 Y_1 , 不难看出 Y_1 能获得更多的图像信息。一维情况

下空洞卷积的具体操作如式(3)所示。

$$y[i] = \sum_{p=1}^P x[i + r \cdot p] w[p] \quad (3)$$

其中, $y[i]$ 是输出信号, $x[i]$ 是输入信号, $w[p]$ 是卷积核, r 是空洞率, p 是位置, P 表示卷积核的大小。空洞率为 r 时, 则像素为第一个位置、第 $r + 1$ 个位置以及依次加 r 的位置与对应的卷积核相乘后累加, 相当于在卷积核中插入 $r - 1$ 个零元素, 当卷

积核为 k 时,其感受野大小如式(4)所示。

$$d = (r - 1) \times (k - 1) + k \quad (4)$$

当空洞率为 2 且卷积核为 3×3 时,则其感受野为 5。此处的感受野扩张和 Deeplabv3 + 是相同的运算。

在 ASPP 中本实验采用平均池化,取周围像素点的均值,这样可以表达所有像素点的信息。同时在卷积后面加 1×1 的卷积是为了增加整个系统处理非线性对象的能力,使整个系统在面临较复杂的环境时仍然有较高的准确率。具体计算卷积输出的形状大小如式(5)和式(6)所示。

$$OH = \frac{H + 2P - FH}{S} + 1 \quad (5)$$

$$OW = \frac{H + 2P - FW}{S} + 1 \quad (6)$$

其中 H, W 是输入信号的大小, FH, FW 是滤波器的大小, OH, OW 为输出信号的大小, P 是填充, S 是步幅,依据以上公式,当有输入信号的大小、滤波器大小、填充和步幅时就能知道输出信号大小。

其次,在 Deeplabv3 + 网络中采用两次双线性插值上采样,双线性插值的具体矩阵如式(7)所示,以 $(0,0), (1,0), (0,1), (1,1)$ 为例,在这个正方形里面进行 x 轴和 y 轴方向上的插值。

$$f(x, y) \approx [1 - x \ x] \begin{bmatrix} f(0, 0) & f(0, 1) \\ f(1, 0) & f(1, 1) \end{bmatrix} \begin{bmatrix} 1 - y \\ y \end{bmatrix} \quad (7)$$

假设 x, y 分别是目标点的 x 轴坐标和 y 轴坐标,当在 $(0,0), (1,0), (0,1), (1,1)$ 中间进行插值为 (x, y) 时,这个点与 $(0,0)$ 之间的关系为 $f(0,0) \times (1 - x) \times (1 - y)$, 与 $(1,0)$ 的关系为 $f(1,0) \times x \times (1 - y)$, 与 $(0,1)$ 的关系为 $f(0,1) \times (1 - x) \times y$, 与 $(1, 1)$ 的关系为 $f(1,1) \times x \times y$ 。

从公式中可以得出,目标点的灰度值是由周围 4 个像素灰度值取平均得到的,这种方式考虑了周围像素值大小,但是没有考虑相邻点变化率影响,会使放大后部分细节信息丢失。针对这一问题,本文采用图 2 所示的方式,串联两个 3×3 的卷积和一个 1×1 的卷积然后并联一个同样的结构,这样能减少信息丢失。

本实验利用改进的 Deeplabv3 + 和原始 Deep-

labv3 + 对采集的 300 张线缆缺陷图片进行训练、测试,下面对两种算法处理图像的结果进行对比,具体结果如图 4 所示。



图 4 分割结果图

依据分割结果可以看出,Deeplabv3 + 能分辨出大致的缺陷,但相对较浅的地方并没有本文改进的算法好。再将两种方法的精确度和 MIOU 进行对比,具体的实验数据如表 1 所示。

表 1 网络结构数据对比

网络结构	精确值	MIOU
Deeplabv3 +	95.58%	89.56%
改进 Deeplabv3 +	97.65%	93.37%

从上述的结果可以看出,改进的 Deeplabv3 + 相比原始 Deeplabv3 + 网络有更好的准确性和 MIOU,更符合工业的需求,在线缆安全方面也有了更高的保障。

1.3 本文算法流程

图 5 所示是本文算法流程图。

从图中可以看出,先将线缆图片用 Labelme 软件制作成 VOC 的格式标签,里面包含了训练集、测试集和验证集,这样做是为了使 Deeplabv3 + 网络能通过这些标签来识别图片信息。在读取信息后,需要做训练前的准备,调节具体的参数,用训练集训练出网络模型,之后用验证集通过 MIOU 和精确度 P 的数据找到此时相对最优模型。如果验证出 MIOU 和精确度 P 相对原始 Deeplabv3 + 相差较大,则需要返回,继续调节参数,当达到理论标准值后保存此时的最优模型,最后运用最优的模型对测试集进行测试并获得结果。其中具体的 MIOU 和精确度 P 公式如式(8)和式(9)所示。

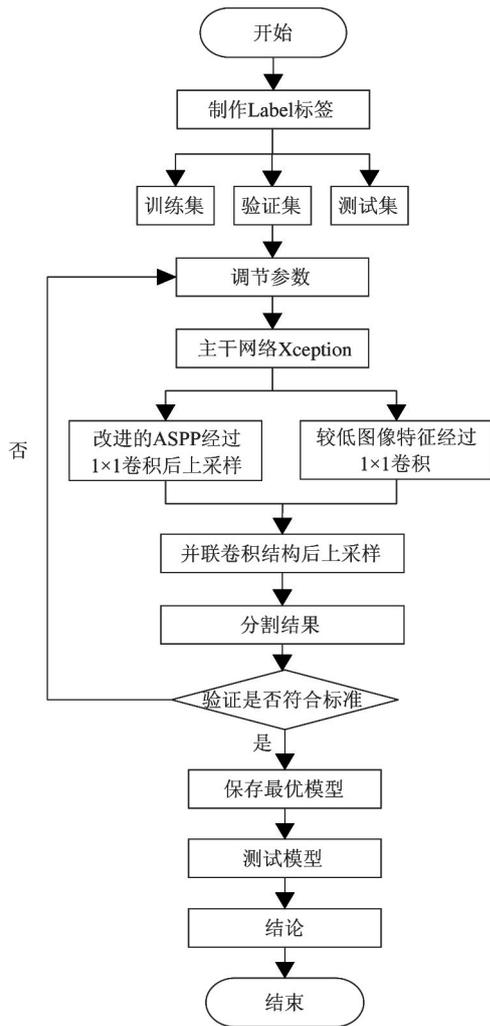


图5 本文算法流程图

$$MIOU = \frac{1}{k} \sum_{i=1}^k \frac{TP}{FN + TP + FP} \quad (8)$$

$$P = \frac{TP}{TP + FP} \quad (9)$$

式中 TP 是被正确分割的线缆缺陷像素数目, FP 是被错误分割为缺陷背景的像素数目, FN 是错误标记为背景的像素数目。MIOU 可以用图 6 表示, 真实值与预测值的交集比上真实值预测值的并集减去交集累加求平均, 精确度 P 就是预测正确的占整个预测样本的值。

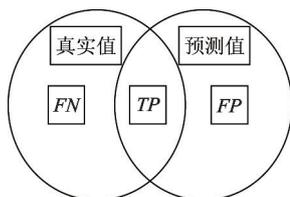


图6 真实值与预测值图示

2 实验分析

本实验共采集实际线缆表面缺陷图 300 张, 通过 Labelme 将数据制作成可识别的标签, 数据的类别分为凹槽、鼓包、裂缝和背景 4 类。本文实验图片由智能机采集, 经过程序处理后图片大小为 513×513 像素, 本文改进 Deeplabv3 + 网络算法程序是在 Windows10 系统、显卡 8 GB 下使用 Pytorch 框架实现的。

实验利用本文提出的改进 Deeplabv3 + 算法模型来识别线缆表面缺陷图, 并将其与 Canny 算子、Sobel 算子、OTSU 阈值割、最大熵阈值分割算法分析结果进行对比, 5 种算法分析结果如图 7 所示。

图 7(b) 表明使用传统的 Canny 算子^[12] 分割能够识别简单的缺陷, 如单个凹槽、鼓包等, 但线缆表面缺陷较浅时很难正确地识别。对比图 7(b) 和 (c) 可以看出, 虽然 Sobel 算子^[13] 效果比 Canny 算子要清晰, 但对于线缆上比较浅的缺陷依然无法清晰识别。图 7(d) 中 OTSU^[14] 阈值分割的结果信息量很少, 只有鼓包这种状态能被清晰识别, 其状态的缺陷都不能被有效地识别。从图 7(e) 中可以看出, 最大熵阈值分割^[15] 效果得到的信息量也很少, 只能识别鼓包。图 7(f) 是本文算法分析结果, 相比其他方法, 凹槽、鼓包、裂缝和复合的环境都能被清晰有效地识别出来。从以上对比实验可以看出, 改进算法面对各种缺陷情况都能准确地识别目标, 相对于其他算法有更好的准确性。

3 结论

本文针对线缆缺陷检测精度问题提出了改进的 Deeplabv3 + 算法, 数据处理部分采用 Labelme 进行数据集的标注。在网络结构部分将 ASPP 中的空洞卷积数量翻倍并在其后增加一个 1×1 的卷积处理环节解决了输出图像可能断层问题和增加系统处理非线性对象的能力。在解码过程中引入并联多层 3×3 的卷积解决了上采样部分信息丢失问题。实验结果表明, 改进算法在精确度、平均交并比方面比原

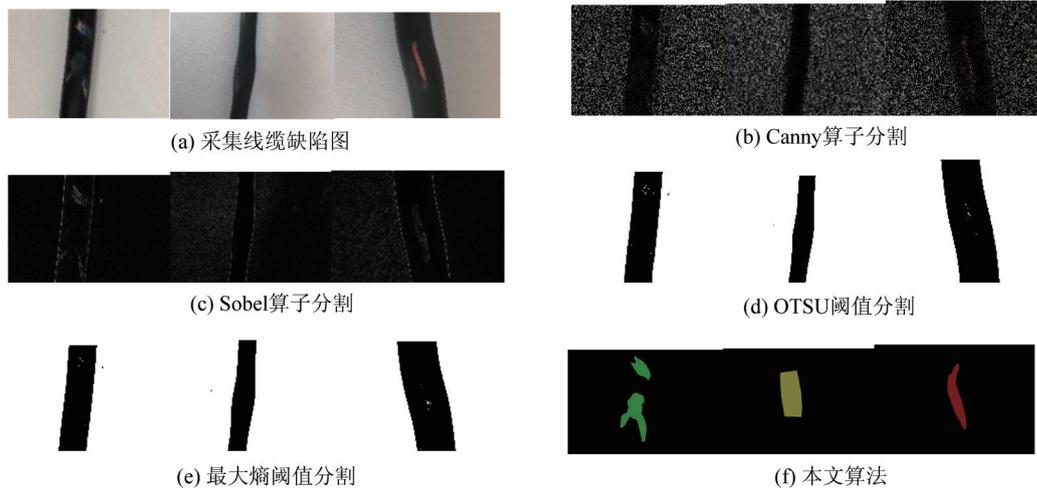


图7 算法对比

始 Deeplabv3 + 效果更好。相较于其他传统算法,改进 Deeplabv3 + 算法虽然在检测精度上表现更好,但是仍然有很多问题需要解决,如本实验相机像素相对较低会影响识别的结果,同时本文算法相对复杂且运算量大,在图片反光或者恶劣环境时分割效果不是很突出。未来的研究将关注如何优化数据以及算法来进一步提高线缆缺陷的检测效果。

参考文献

- [1] 乔湘洋,王海芳,祁超飞,等. 基于机器视觉的线缆表面缺陷检测系统设计及算法研究[J]. 机床与液压, 2020,48(5):49-53
- [2] 王海芳,焦龙,乔湘洋,等. 基于机器视觉的线缆表面缺陷快速检测算法研究[J]. 组合机床与自动化加工技术, 2020(2):119-122
- [3] 张俊. 基于机器视觉的线缆表面缺陷在线检测系统的研究[D]. 成都:电子科技大学光电科学与工程学院, 2017:24-41
- [4] Gao S, Dou X, Bai L, et al. Research on detection and recognition algorithm of optical fiber for surface defects [C]//The 31st Chinese Control and Decision Conference, Nanchang, China, 2019:966-971
- [5] Song E, Seung-Hyun E, Haemin J. Concrete crack detection and quantification using deep learning and structured light[J]. *Construction and Building Materials*, 2020, 252(5):252-256
- [6] Chen L, Papandreou G, Kokkinos I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018,40(4):834-848
- [7] 薛飞,伍岳庆,姚宇,等. 基于 DeepLabv-3 的左心室超声图像分割方法[J]. 计算机应用,2019,39(S2):114-117
- [8] Zhang K H, Liu X H, Chen Y F. Research on semantic segmentation of portraits based on improved Deeplabv3 + [C]//Proceedings of International Conference on AI and Big Data Application, Guangzhou, China, 2019: 425-431
- [9] Bhakti B, Shubham I, Suhas G, et al. Semantic scene segmentation in unstructured environment with modified DeepLabV3 + [J]. *Pattern Recognition Letters*, 2020, 138: 223-229
- [10] 王亚刚,郗怡媛,潘晓英. 改进 Deeplabv3 + 网络的肠道息肉分割方法[J]. 计算机科学与探索, 2020, 14(7):1243-1250
- [11] 刘致驿,孙韶媛,任正云,等. 基于改进 Deeplabv3 + 的无人车夜间红外图像语义分割[J]. 应用光学,2020, 41(2):180-185
- [12] 商景辉. 基于 Canny 算子的边缘检测算法研究[J]. 信息与电脑(理论版),2020, 32(2):39-41
- [13] 张琪,王艺,陈兆飞. 基于 Sobel 算子图像预处理的目标检测算法[J]. 电子技术与软件工程, 2020(5):151-153
- [14] 燕红文,邓雪峰. OTSU 算法在图像分割中的应用研究[J]. 农业开发与装备, 2018(11):103-108
- [15] 王利奉. 基于 OTSU 和最大熵的阈值分割算法的研究[D]. 哈尔滨:哈尔滨理工大学自动化学院, 2018:26-39

Research on cable surface defect detection based on improved Deeplabv3 + network

Chen Liang, Yang Xianzhao, Liu Huikang

(Engineering Research Center for Metallurgical Automation and Measurement Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081)

Abstract

In order to improve the accuracy of cable surface defect detection, an improved image segmentation method of Deeplabv3 + network is proposed and applied to cable surface defect detection. Based on Deeplabv3 + network skeleton unchanged, the spatial pyramid structure is changed from 4 dilated convolutions to 8 dilated convolutions and then 1×1 convolution is added. At the same time, a parallel structure is used to reduce the information loss during the whole network transmission process after decoding and fusion. The improved algorithm is used to train and test the cable surface defect image data set, and the results show that it is better than the original Deeplabv3 + analysis in accuracy and mean intersection over union (MIOU). Compared with edge segmentation, threshold segmentation and other algorithms, the proposed algorithm improves the detection accuracy of cable surface defects.

Key words: surface defects of cable, Deeplabv3 +, dilated convolution, parallel structure, accuracy