

DOI:10.19344/j. cnki. issn1671-5276.2024.05.049

港口场景集装箱锁孔目标智能检测研究

杨晓翔, 林云帆, 刘键涛

(福州大学 机械工程及自动化学院, 福建 福州 350116)

摘要:为解决集装箱图像受到外界光线环境的影响造成信息对比度低、暗部细节信息不明显和图像信息难以辨认清等问题,提出一种新的图像增强算法,其重点是利用非线性变换,提高图像的暗部细节,将原本的 RGB 模型转化到 HSV 模型进行均衡化。利用实际工程中港口所提供的测试数据集,使用已训练好的模型进行广泛的实验。实验还探究了图像增强算法对卷积神经网络性能的影响,利用对比度受限的自适应直方图均衡化、伽马校正、拉普拉斯变换以及原始图像与该算法进行对比,使用 10 折交叉验证了该算法的精确率、召回率均大于其他的算法。采用配对 T 检验,比较分析采用各算法间的各项指标差异,结果表明:该算法相比其他算法的效果要好。

关键词:图像增强算法;卷积神经网络;目标检测;直方图均衡化

中图分类号:TP391.4 **文献标志码:**A **文章编号:**1671-5276(2024)05-0234-05

Research on Target Intelligent Detection of Port Container Keyhole

YANG Xiaoxiang, LIN Yunfan, LIU Jiantao

(College of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350116, China)

Abstract: In order to solve the problems of low contrast of information, obscure details and difficulty in identifying image information caused by external light environment in container images, proposes a new image enhancement algorithm, which focuses on improving the dark part details of the image by using nonlinear transformation and converting the original RGB model to the HSV model for equalization. Using the test data set provided by the port in the actual project, a wide range of experiments are carried out with the trained model. The experiment explores the effect of image enhancement algorithm on the performance of convolutional neural network. The contrast limited adaptive histogram equalization, gamma correction, Laplace transform and original image are compared with the proposed algorithm, and the 10-fold cross is used to verify the accuracy rate and recall rate of the proposed algorithm, which are greater than other algorithms. The paired T test is used to compare and analyze the differences of each index among the algorithms. And the results show that the proposed algorithm has better effect than other algorithms.

Keywords: image enhancement algorithm; convolutional neural network; target detection; histogram equalization

0 引言

近些年,智能化集装箱港口逐渐成为全球港口的重点建设项目。港口大型工程机械设备也开始朝着自动化和智能化方向发展。集装箱锁孔识别是在集装箱进闸时的识别,可以实现不停车通闸,节省时间。锁孔识别主要是为了判断单双箱,因为判断单双箱是校验货车集装箱进场逻辑的一个基础,也是集装箱后续的箱号识别、残损识别、危标识别算法的一个基础。只要能识别锁孔,就能判断单双箱,再使用箱号、残损、危标识别算法

来实现这个集装箱进场的识别。判断单双箱不用箱体识别,是因为集装箱数据集是采用拼接图,而使用箱体识别来判断单双箱不一定稳定,因为会受集装箱有残损和拼接效果等的影响。相比而言,锁孔识别会稳定很多。但是有些货箱由于长久使用和被雨水腐蚀的锁孔部分生锈腐蚀,不是很好辨认,并且由于光线的影响所拍摄出来的识别图片在某些地方比较暗,如图 1 所示的集装箱的左下角部分由于光线暗,可能会造成识别的性能不佳。

基金项目:国家自然科学基金资助项目(11972005)

第一作者简介:杨晓翔(1963—),男,内蒙古赤峰人,教授,工学博士,研究方向为力学在工程中的应用和计算力学,937855852@qq.com。



图1 实验数据集中图片的截取

一般在深度学习的数据集中,图像数据集经常会由于拍摄设备、光线照射环境因素或者拍摄时干扰等影响,造成图片质量低和图像的对比度或者图片特征不足。而图像数据集的质量对模型性能有很大的影响。研究表明^[1-2],在训练 CNN 模型时,图像处理是必不可少的,它可以有效地提高 CNN 模型的识别性能。图像增强是图像处理中非常重要的一部分。因此研究图像增强与 CNN 模型之间的关系具有重要意义。图像增强算法是可以将图像的整体或局部特征加强、放大有用特征或者抑制无用特征和加强图像中物体的一些不明显特征^[3]。但是在模型的训练之后使用图像增强算法的很少。提高图像质量,丰富信息,增强图像解释是为了更适合人类视觉系统和体验,但是机器视觉终究是与人类视觉有很大区别的。

在本文中,提出了一种新的图像增强算法,目的是为了提髙目标检测中卷积神经网络的性能,改善原始图片中由于环境原因造成过暗的状况,并能增强有用部分的特征。该方法增强了数据集图像,方便了目标的检测。与其他方法相比,本文的方法具有更好的效果。

1 港口集装箱光照不均匀图像增强方法

用 I 来表示 $M \times N$ 尺寸大小的灰度图像且像素强度为 $\{0, \dots, L-1\}$ 。首先需要对图像进行非线性校正,非线性校正需要将原始图像的 RGB 颜色模型转化为 HSV 颜色模型,而且只对 HSV 中的 V 通道进行校正,为的就是防止破坏图像原有的颜色空间。表示为

$$I'_v = \left(\frac{1}{\alpha} \times I_v \right)^{\frac{1}{\beta}} \quad v=0,1,2,\dots,L-1 \quad (1)$$

式中: I_v, I'_v 分别为图像的输入和输出的灰度值; α, β 为校正常数。

将非线性化的图像划分为 m 个 16×16 固定子块,且每个子块之间是互不影响相互连续的,将每个子块的像素个数平均分到每个灰度级的平均值 N_{mean} :

$$N_{\text{mean}} = \frac{N_x \cdot N_y}{N_n} \quad (2)$$

式中: N_n 为子区域的灰度级的数量个数; N_x 为子区域 x 轴方向的像素个数; N_y 为子区域 y 轴方向的像素个数。

为了限制噪声放大和局部对比度过大,就需要对各个子区域直方图的高度进行限制,再形成输入图像的一维直方图,用所需的强度值替换单个像素的强度。然后根据相应的概率分布计算出所需的强度值。每个子区域内的一维直方图均衡化技术使用如下定义的离散变化来生成输出像素强度。

$$S_t = T(r_t) = \frac{L}{16} \sum_{k=16(m-1)}^{t+16(m-1)} p(r_k) \quad t=0,1,2,\dots,15 \quad (3)$$

式中: S_t 表示输出图像中的均衡化灰度级; r_t 表示输入图像中像素的灰度级; $T(\cdot)$ 表示转换运算符; $p(r_k)$ 表示子区域内灰度级 r_k 出现的概率。

$$p(r_k) = \frac{1}{N_x \cdot N_y} \sum_{k=16(m-1)}^{t+16(m-1)} N(r_k) \quad t=0,1,2,\dots,15 \quad (4)$$

式中 $N(r_k)$ 表示 r_k 在图像中出现的次数。

最后把每个子区域的中点作为参考来估算其灰度值,对图像中的像素使用双线性插值的方法来进行灰度插值。

2 实验方案的设计

2.1 实验数据

实验数据来源于港口货车目标检测图像数据。考虑到摄像头抓拍的货车左右侧视角会有区别,故实验的数据集中有 2 000 个样本,包括 1 000 张左视角和 1 000 张右视角样本。

所选的图像格式都是 JPG。DZIUGAITE 等^[4]的研究表明,神经网络图像分类器易受到对抗性图像的影响, JPG 图像进行压缩操作通常会在很大程度上降低 CNN 分类器的精度。为了避免图像增强过程中对原始图像的二次压缩操作,必须在 JPG 图像进行图像增强之前转换格式,故需要

将原始数据格式 JPG 转化为 PNG。

2.2 实验方案设计

选取对比度受限自适应直方图均衡化、伽马校正、拉普拉斯变换处理图像以及原始数据集作为对照组。

实验方案如图 2 所示。在数据集中,原始数据集称为 A1, JPG 图像首先转化为 PNG 图像,数据集称为 A2。A2 中所有图像都要从 RGB 通道转换为 HSV 通道,然后再转换回 RGB 通道。该数据集称为 A3, A3 作为实验空白对照组。将 A2 从 RGB 通道转换为 HSV 通道,分别通过对比度受限自适应直方图均衡化方法和伽马校正增强 V 通道,并将增强后 HSV 通道转化为 RGB 通道,此时数据集分别称为 A4 和 A5。通过拉普拉斯操作对 A2 中的图像进行增强,数据集称为 A6。经过本文算法处理的数据集称为 A7。实验将在卷积神经网络(CNN)中比较 A3、A4、A5、A6、A7 的性能。

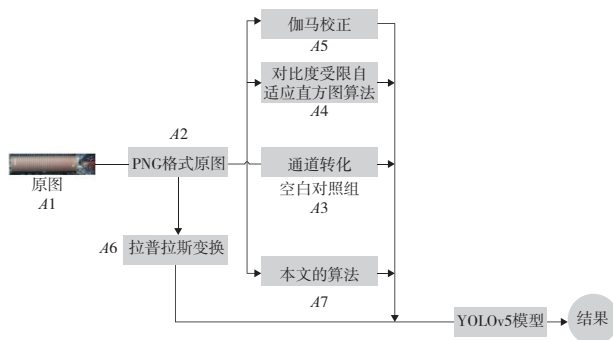


图 2 实验方案

本次实验采用的是 YOLOv5 网络。该模型是由 Ultralytics LLC 公司对 YOLOv4 的改进版本,是目前目标检测优秀的检测网络^[5]。

2.3 模型评价

在本次实验中,如何比较相同 CNN 模型在不同数据集上的性能是关键。但实际上,比较模型性能是非常复杂的,它一般涉及 3 个因素。第一,需要比较 CNN 模型的泛化性能。而一个测试或一个实验只能获得测试集中的性能。测试集的性能不同于泛化性能。第二,测试集的性能会受到测试集的选择、测试集的数量和测试集中样本的影响。第三, CNN 模型并不总是稳定的,即使用相同的参数对同一测试集进行多次测试,结果会有所不同。因此,选择好的评价方法、绩效测量方法和假设检验方法是非常重要的。

1) 评估方法

机器学习中有许多评估方法,比如:留出法、

K 折交叉验证和自举检验等。留出法是最基本的方法,它是将数据集划分为两个互斥集,其中一个集为训练集,另一个集为测试集。由于该方法有较大的偶然性,不同的方法会产生不同的结果,所以该方法是不稳定的。K 折交叉验证通常是与留出法一起使用,将数据集分成 K 个部分,然后进行 K 次检验;每个子集都用作测试集,这确保所有数据都被测试一次。而自举法是将一个数据从原始数据集随机抽取到测试集,重复 m 次得到测试集,将原始数据集作为训练集。当 m 无穷大时,原始数据集中约有 36.8% 的数据不会出现在测试集中^[6]。有论文表明同时使用 K 折交叉验证和留出法在分类算法中具有良好的性能^[7]。所以本文就采用 K 折交叉验证和留出法,其中 K=10。

2) 性能指标

在目标检测领域中,存在着很多精度评价指标。本次实验所用的评价指标是精确率 (Precision)、召回率 (Recall)、F1-Score 和平均精度 (AP)。精确率是整个测试集上,被分类器识别出来的结果中,正确分类个数所占的比率,即衡量的是分类器对数据集错分的情况。召回率是在测试集中,所有正样本被正确识别的概率,即衡量的是分类器对数据集的漏检情况。F1-Score 就是为了解决单一的评价指标精确率或召回率很难宏观地表现分类器的好坏, F1-Score 可以联合评价精确率和召回率。平均精度是 Precision-Recall 曲线下面的面积。

$$e_{\text{Precision}} = \frac{T_p}{T_p + F_p} \quad (5)$$

$$e_{\text{Recall}} = \frac{T_p}{T_p + F_N} \quad (6)$$

$$F_{\text{F1-Score}} = \frac{2}{\frac{1}{e_{\text{Precision}}} + \frac{1}{e_{\text{Recall}}}} = 2 \times \frac{e_{\text{Precision}} \times e_{\text{Recall}}}{e_{\text{Precision}} + e_{\text{Recall}}} \quad (7)$$

式中: T_p 表示被判定为正样本,真实情况也是正样本; T_N 表示被判定为负样本,真实情况也是负样本; F_p 表示被判定为正样本,但真实情况是负样本; F_N 表示被判定为负样本,但真实情况是正样本。

3 实验结果与分析

3.1 视觉评估

图 3 所示为基于每种方法生成的图像增强结果。图 3(a) 所示的原图包括货车的车头和集装箱及一些背景,本文主要关注的是货车的集装箱

部分。如图 3(b)所示,对比度受限自适应直方图均衡化从图像上看只是比原图的颜色更鲜艳,而且原先较暗的区域还是跟原来一样,暗部细节还是不够明显,也就是说这种方法不能改善亮度。如图 3(c)所示,伽马校正改善了直方图的低电平,增强了原图的暗部细节,但改善后的亮度不均匀,总体上给人一种过度明亮且不自然的感觉。如图 3(d)所示,拉普拉斯变换虽增强了局部特征,但是过多的增强了无用部分。如图 3(e)所示,经过本文算法处理的图片,可以看到整体的增强效果要比其他算法好,而且也改善了原始图片中的暗部。



图 3 用不同图像处理方法增强货车集装箱图像对比度的结果

3.2 识别的效果

如图 4—图 8 所示,使用对比度受限的自适应直方图均衡化、伽马校正、拉普拉斯变换算法与原图识别出的锁孔情况一样。这几个算法对集装

箱左下角的锁孔没识别出来,可能因为左下角的暗部细节不够明显,而使用本文算法可以检测出以上算法所遗漏的那个锁孔,这说明本文算法可以提高暗部细节,且能提高 YOLOv5 的性能。



图 4 原始图像



图 5 对比度受限的自适应直方图均衡化



图 6 伽马校正



图 7 拉普拉斯变换



图 8 本文算法

3.3 定量评价

如表 1 所示,使用本文的算法优于其他的算法,而使用其他的图像增强并不能提高锁孔识别的精确率。如表 2 所示,在召回率的指标上,使用本文的算法和对比度受限的自适应直方图均衡化都比原图高,而使用伽马校正和拉普拉斯算法则低于原始图像。

表 1 实验中的各算法 10 折交叉验证的精确率

单位: %

算法	1	2	3	4	5	6	7	8	9	10	平均值
Original	90.67	90.82	92.35	90.18	91.55	90.59	91.25	91.67	92.66	91.30	91.31
CLAHE	91.35	89.89	92.07	90.18	91.14	91.21	90.98	91.57	92.45	90.86	91.17
Gamma	91.09	91.37	91.84	89.13	92.33	90.73	92.08	92.08	91.79	89.54	91.20
Laplace	85.38	89.40	89.02	84.92	85.38	87.66	85.34	87.28	84.65	86.06	86.51
Our	91.04	91.51	92.85	90.49	92.20	91.51	91.48	92.97	92.66	92.33	91.91

表 2 实验中的各算法 10 折交叉验证的召回率

单位: %

算法	1	2	3	4	5	6	7	8	9	10	平均值
Original	90.78	92.93	95.36	89.14	94.44	93.37	94.94	94.90	94.15	92.43	93.24
CLAHE	91.02	91.10	94.73	91.99	93.56	94.01	94.43	95.02	94.04	92.56	93.25
Gamma	90.53	91.71	94.60	88.28	94.19	92.35	94.18	94.65	91.06	90.32	92.19
Laplace	83.62	90.49	91.59	84.61	86.99	87.88	86.96	90.04	83.49	87.34	87.30
Our	91.26	93.29	96.11	91.74	95.45	94.90	95.19	96.48	94.15	94.04	94.26

3.4 统计假设检验

本文采用配对 T 检验进行假设检验。有文献^[8]表示, K 折交叉验证的方差无法无偏估计, 这就要求假设检验应该使用总体方差和总体标准差, 而不是样本方差和样本标准差。在单侧检验中, 配对 T 检验需要比较 T_i 与 $t_{\alpha, k-1}$, 其中 α 是置信区间。当 $T_i > t_{\alpha, k-1}$ 时, 就能拒绝原假设。当 $\alpha = 0.05$ 时, $t_{0.05, 9} = 1.83$ ^[9]。 T_i 的公式为

$$T_i = \left| \frac{\sqrt{K}\mu}{\sigma} \right| \quad (8)$$

式中: μ 为两个差值的平均值; σ 为两个差值的总体标准差。

如表 3 所示, 比较原始图像和对比度受限的自适应直方图均衡化精确率的 $T_i \approx 0.99$ 。由于 0.99 是小于 1.83 的, 所以说原始图像与使用对比度受限的自适应直方图均衡化之后图像集的精确率之间是没有显著差异的。同理对于原始图像和使用伽马校正之后图像集的精确率之间也是没有显著差异的。然而比较原始图像和使用拉普拉斯变换数据集精确率的 $T_i \approx 8.56 > 1.83$, 说明这两个之间的差异是显著的, 再看表 1 中原始图像的平均精确率是明显大于使用拉普拉斯变换的, 可以认为使用拉普拉斯变换的图像集相比于直接原始图像对于卷积神经网络的性能是有下降的。而使用本文算法的 $T_i \approx 4.79 > 1.83$, 说明使用本文算法与直接使用原始图像是有明显差别的。再看表 1 中使用原始图像的精确度比本文的算法小, 故可以得出使用本文算法平均精确率是大于直接使用原始数据集的结论。同理如表 4—表 6 所示, 对比度受限的自适应直方图均衡化与伽马校正和直接使用原始数据集的召回率、F1-Score、精度之间的差异是没有明显差距的。而拉普拉斯变换相比于直接使用原始数据集的召回率、F1-Score、精度是有明显下降的。而本文算法相比于直接使用原始数据集的各项性能指标都有明显差距, 并且平

均值都大于原始数据集。由于本文算法相比于原始图像之间的各项性能指标有显著差异, 故还比较了本文算法与其他算法的各项指标之间差异, 如表 7 所示。

表 3 在实验中比较原始图像与各图像增强算法精确率的 T_i

Precision	CLAHE	Gamma	Laplace	Our
T_i	0.99	0.44	8.56	4.79

表 4 在实验中比较原始图像与各图像增强算法召回率的 T_i

Recall	CLAHE	Gamma	Laplace	Our
T_i	0.01	3.87	8.30	4.20

表 5 在实验中比较原始图像与各图像增强算法 F1-Score 的 T_i

F1-Score	CLAHE	Gamma	Laplace	Our
T_i	0.31	2.36	8.66	5.05

表 6 在实验中比较原始图像与各图像增强算法精度的 T_i

AP	CLAHE	Gamma	Laplace	Our
T_i	0.96	2.66	9.51	5.11

表 7 在实验中比较本文算法与各图像增强算法性能指标的 T_i

T_i	CLAHE	Gamma	Laplace
Precision	3.83	2.46	10.58
Recall	4.22	6.52	10.79
F1-Score	4.23	4.75	10.93
AP	3.92	5.95	11.55

(下转第 272 页)

5 结语

在充分分析和调研现有检修工人攀爬铁塔检修作业的现状和详细流程之后,明确了现有流程存在的主要安全隐患为牵引绳的悬挂。针对牵引绳的自主悬挂问题,结合无人机技术设计了一款输电铁塔高空作业安全防护装置。重点对安全防护装置的自动固定机构和释放机构进行了详细的结构设计,并对整体控制系统做了介绍。在考虑实际工况的情况下,对固定机构支架的可靠性进行了有限元分析,分析结果表明固定机构支架安全可靠,可承受3倍以上负载,并提出固定支架的优化方向。对释放机构的失电型电磁铁进行了选型,其负载大于安全防护装置质量,结构大小适中,选型合理。最后利用ADAMS软件搭建了安全防护装置的虚拟样机,并对其工作过程进行了仿真。结果表明:该安全防护装置设计合理,运行可靠,为提高检修人员的爬塔作业提供了可靠的安全防护,对进一步推动电力行业的发展具有积极意义。

参考文献:

- [1] 杨博,陈义军,姚伟,等. 基于新一代人工智能技术的电力系统稳定评估与决策综述[J]. 电力系统自动化,2022,46(22):200-223.
- [2] 沈玉玲,吕燕,陈瑞峰. 基于大数据技术的电力用户行为分析及应用现状[J]. 电气自动化,2016,38(3):50-52.
- [3] 杨先明,袁欣悦. 西电东送的区域经济增长效应[J]. 财经科学,2022(5):77-89.
- [4] 崔黎. 无人机航拍在铁路事故救援中应用思考[J]. 铁路通信信号工程技术,2016,13(1):48-51.
- [5] 李华锋,赵康康. 多耦合失电型电磁制动器的设计[J]. 电子制作,2022,30(16):67-69,54.
- [6] 葛男男,孙伟,曹界宇. 无人机实时图像传输系统设计[J]. 中国新通信,2021,23(2):61-62.
- [7] 贾文涛,李春涛. 无人机航迹优化与跟踪技术研究[J]. 机械制造与自动化,2020,49(6):156-161.
- [8] 李昀轩,魏雅君,倪博宇,等. 基于ANSYS与ADAMS的六轴机器人整机分析[J]. 机械设计,2020,37(12):13-17.

收稿日期:2022-12-06

(上接第238页)

由表7可以看出本文算法与其他算法的性能指标的 T_i 均大于1.83,且本文算法的性能指标的平均值都大于其他的图像增强算法,说明本文算法与其他算法之间的差异显著。

4 结语

为了提高集装箱锁孔的识别,本文设计了一种增强对比度和提高图像暗部细节的图像增强算法,实验使用适当的度量标准和配对 T 检验来确保结论的有效性。此外,为了减少内部效度的威胁,实验进行了10折交叉验证测试,可以减少实验的偶然性。在实验中,本文的算法明显优于其他几种图像处理算法。

参考文献:

- [1] KOO K M, CHA E Y. Image recognition performance enhancements using image normalization[J]. Human-Centric Computing and Information Sciences, 2017, 7(1):1-11.
- [2] 苏欣,赖复尧,余容平,等. 基于多视角模板匹配的零件图像检索方法[J]. 机械制造与自动化,2023,52(1):222-225,229.
- [3] MAINI R, AGGARWAL H. A comprehensive review of

image enhancement techniques [EB/OL]. (2010-03-22) [2022-09-12]. <https://sites.google.com/site/journalofcomputing>.

- [4] DZIUGAITE G K, GHARAMANI Z, ROY D M. A study of the effect of JPG compression on adversarial images [EB/OL]. (2016-08-02) [2022-09-12]. <https://doi.org/10.48550/arXiv.1608.00853>.
- [5] LIU W T, WANG Z Y, ZHOU B, et al. Real-time signal light detection based on Yolov5 for railway [J]. IOP Conference Series: Earth and Environmental Science, 2021, 769(4):042069.
- [6] 周志华.机器学习[M].北京:清华大学出版社,2016.
- [7] WONG T T. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation[J]. Pattern Recognition, 2015, 48(9):2839-2846.
- [8] GRANDVALET Y, BENGIO Y. Hypothesis testing for cross-validation [J]. Montreal Universite de Montreal, Operationnelle DdIeR, 2006, 1285:1-10.
- [9] MILTON J S, ARNOLD J. Introduction to probability and statistics [M]. New York: McGraw-Hill Education, 2002.

收稿日期:2023-03-06