

针对钢包编号分类识别的数据增强方法的研究

吕宜光, 郝明, 孙凯明, 周丽丽

(黑龙江省科学院智能制造研究所, 黑龙江 哈尔滨 150090)

摘要: 针对钢包文字识别中手涂编号识别不佳问题, 提出基于深度学习的分类识别方法对编号进行识别。通过图像变换、添加高斯噪声、Cutout与Gridmask等图像遮挡数据增强方法, 提高了训练模型精度与泛化能力。实验表明方法在真实场景的钢包编号识别中具有较高的准确率。

关键词: 深度学习; 分类识别; 数据增强; 钢包编号

中图分类号: TP391.41 文献标识码: A 文章编号: 1003-7241(2025)01-0185-04

Research on Data Enhancement Methods for Classification and Recognition of Ladle Numbers

LV Yi-guang, HAO Ming, SUN Kai-ming, ZHOU Li-li

(Institute of Intelligent Manufacturing, Heilongjiang Academy of Sciences, Harbin 150090 China)

Abstract: Addressing the issue of poor recognition of manually painted serial numbers in steel ladle text recognition. We propose a classification and recognition method based on deep learning to identify the serial numbers. By utilizing data augmentation techniques such as image transformations, adding Gaussian noise, Cutout, Gridmask, and other image occlusion methods, we have improved the accuracy and generalization capabilities of the trained model. The experiments show that the method achieves high accuracy in identifying steel ladle serial numbers in real scene scenarios.

Keywords: Deep learning; Classification recognition; Data augmentation; Ladle number

0 引言

在数字化浪潮席卷全球的今天, 智能制造已成为我国未来工业发展的重要趋势。炼钢企业作为传统工业的重要支柱, 积极响应时代号召, 不断探索创新路径。目前钢铁冶炼过程纷纷引入机器视觉技术, 实现生产过程中对生产环境、生产过程、生产信息的精准识别与检测, 利用工业互联网技术, 构建高效的信息化平台, 实现生产数据的实时采集与分析。

但由于炼钢厂车间环境复杂, 冶金设备号的识别难度大, 制约了炼钢工艺中的智能化水平。在钢包转运的场景中, 经常会遇到手动涂写钢包编号的情况, 由于这种字体与印刷字体存在显著差异, 因此给字符识别带来了一定的挑战。并且由于钢包设备外形特征以及冶炼场所环境影响, 编号字体存在变形、烟尘干扰、其他设备遮挡等情况如图1所示, 这也对识别造成一定影响。针对此问题, 本文提出了采用深度学习中的分类识别方法来实现钢包编号的自动识别。张勇等提出基于PP-YOLO结合数据增强mixup、颜色扭曲方法进行改进, 在中小目标

检测中有明显提升, 可有效提升病虫害数据样本指标^[1]; 曹义亲等针对金属表面缺陷检测问题提出了一个新的E-YOLOX检测算法, 并采用Cutout数据增强方法, 实现了图像自适应的数据增强^[2]; 陈启鹏等提出了一种基于YOLOv4的工件表面质量在线监测方法, 并且在预处理阶段提出采用Cutmix的数据增强方法, 增加训练样本的多样性, 提高模型的鲁棒性和泛化能力, 以避免训练模型产生过拟合^[3]; 王九鑫等提出了一种基于改进YOLOv5网络的金属表面缺陷分割算法, 采用了直方图均衡、Mosaic数据增强方法增强识别准确性^[4]; 陈千懿等利用Sarsa强化学习算法自动分析巡检作业缺陷, 实现无人机配电网巡检过程中对树障、设备及安全隐患等缺陷的自动识别^[5]。



图1 钢包编号

虽然以上作者的研究内容是针对植物病虫害、金属件缺陷等此类目标特征不规则、外观随机等物体的识别

*基金项目: 黑龙江省省属科研院所科研业务费项目 (CZKYF2022ZN02)

收稿日期: 2024-05-27

方法,与本文所需识别的手涂文字在外观特征上还是有较大差异,但是所利用的数据增强方法提高识别准确率的技术路线为本文研究提供了参考。为此本文提出通过数据增强的方法,进一步提升钢包手涂编号识别的准确性。

1 训练数据合成

在基于深度学习的图像分类中数据集非常重要,由于现场钢包数量有限,采集的编号样本仅有824张,本文采用人工合成的方式生成数据集。本文现场编号样本共23类均为两位号码,人工书写字符模板采用随机图像合成虚拟样本。

在深度学习的图像分类任务中,训练的数据集扮演着至关重要的角色。鉴于现场钢包数量有限,实际采集的编号样本数量仅为824张,这在一定程度上限制了模型的训练效果。对此,本文采用了人工合成字符的方式以扩充数据集。本文所处理的现场编号样本共涵盖23个类别,均为两位号码。制作多个人工书写的字符模板,如图2,并通过随机图像合成技术生成了虚拟样本,以模拟实际场景中的多样性和复杂性,如图3所示。

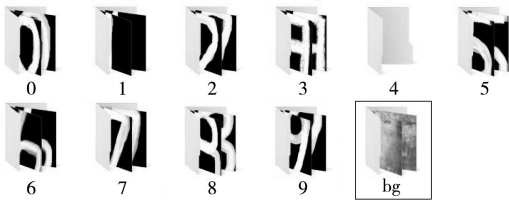


图2 字符样本

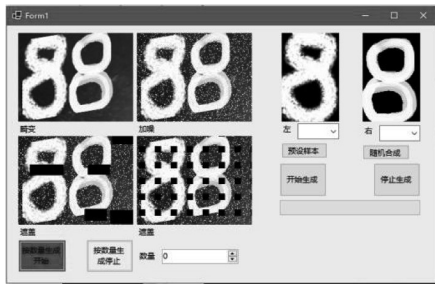


图3 合成软件

2 数据增强

合成的训练集样本和真实样本存在一定差异,为了增强模型的泛化能力和鲁棒性,同时提升其在真实环境下的性能,需要对合成的训练集样本进行数据增强处理,并通过增加数据的多样性来改进模型的整体表现。

2.1 图像变换

钢包本体为锅型结构,编号字符在边缘或底部时在影像上存在倾斜变形。为使得合成样本更贴近实际采用透视变换对其处理,原理是将图片投影到一个新的视平面,也称作投影映射。通用的变换公式为:

$$[x', y', z'] = [u, v, w] * T \tag{1}$$

式中, u, v 为原始图像坐标, T (Transform)为透视变换矩阵其公式为:

$$T = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} T_1 & T_2 \\ T_3 & a_{33} \end{bmatrix} \tag{2}$$

式中, T_1 用于线性变换, T_2 用于透视变换, T_3 用于平移变换。对此本文通过随机性生成透视变换对应的四对像素点坐标,即可获得一个随机透视变换矩阵 T ,图像变换如图4,结果如图5所示。

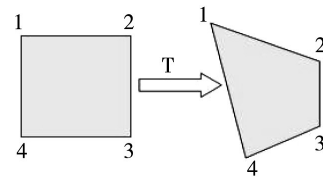


图4 透视变换



图5 变换后样本

2.2 高斯噪声

通过为图像添加高斯噪声的方法模拟现场烟雾环境。高斯噪声也称为白噪声,是一种服从正态分布的随机噪声。并且作为一种数据增强方法在深度学习训练中输入数据加入高斯噪声可以提高模型的鲁棒性和泛化能力。这些添加的噪声可以模拟图像上可能出现的污迹或轻微缺失,使得模型即使在面对与训练数据略有差异的图像时,也能有更高的可能性正确识别它们。高斯分布函数为:

$$P(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \tag{3}$$

式中, z 表示灰度值, μ 表示 z 的平均值或期望值, σ 表示 z 的标准差其平方 σ^2 为 z 的方差,处理后结果如图6所示。



图6 增加噪声后样本

2.3 Cutout方法

真实场景中往往出现钢包编号被吊钩遮挡时的情形,如图7,Cutout能够模拟此种情况被部分遮挡时的分类情形,这有助于提升模型在类似实际场景下的泛化能

力^[6]。此外 Cutout 操作生成了更多的不同遮挡情况的图像,模型需要学习更多的特征来表示这些图像,从而降低了过拟合的风险。



图7 真实场景遮挡情况

Cutout 是直接在合成图像上进行遮挡操作。根据输入图像长 W 高 H 作为阈值,随机生成点坐标 $C(x, y)$ 作为起始点,通过随机生成裁剪区长宽 w 与 h 确定待遮挡区域,将区域内的像素值全部置为零,这相当于在图像中模拟了一个遮挡物,如图8所示。



图8 Cutout 遮挡后样本

2.4 Gridmask 方法

上述 Cutout 方法存在两个问题,一是过度删除区域可能造成目标主体大部分甚至全部被删除,或者导致上下文信息的丢失,导致增广后的数据成为噪声数据,如图9;其二是保留过多的区域,对目标主体及上下文基本产生不了什么影响,失去增广的意义,如图10。



图9 过度遮盖



图10 无效遮盖

Gridmask 方法的设计初衷在于实现删除信息与保留信息之间的精细平衡, Cutout 方法中可能产生的判别性区域被完全消除或未经修改保留的问题,这些问题潜在地引入不必要的噪声,从而可能对模型的性能产生负面影响。Gridmask 方法通过生成结构化的格子模式,并有选择性地删除格子内的图像信息,以实现这一平衡,如图11所示。

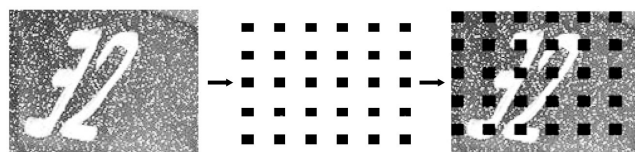


图11 GridMask 方法遮盖

首先通过生成一个与原图具有相同分辨率的掩码(mask),然后将该掩码与原图进行乘法运算,从而得到一幅增强后的图像^[7]。在图12中,灰色区域在掩码中对应的值为1,而黑色区域对应的值为0。这样就实现了对图像中特定区域的遮盖效果,进而达到了图像增强的目的。

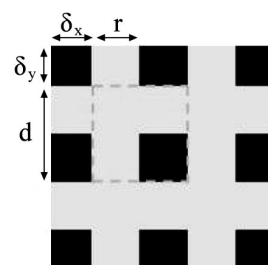


图12 点状方块表示掩码的一个单位

图12中 r 是一个单位中较短的灰色边缘的比例。 d 是一个单位的长度。 δ_x 和 δ_y 是第一个完整单位与图像边界之间的距离。图像遮盖密度及图像信息的保留比例可由如下公式实现:

$$k = \frac{\text{sum}(M)}{H \times W} \quad (4)$$

式中, H 为原图的高, W 为宽, M 为保留下来的像素数, k 为保留比例。 r 可以通过如下公式得到:

$$k = 1 - (1 - r)^2 = 2r - r^2 \quad (5)$$

k 的值越大灰色区域越大遮挡越少,反之遮挡越多, δ_x 和 δ_y 可以通过随机参数 d 的确定:

$$\delta_x(\delta_y) = \text{random}(0, d - 1) \quad (6)$$

3 实验分析

本文使用 Resnet50 作为主干网络进行分类训练, Resnet50 作为一种经典的深度神经网络模型,已经被广泛应用于各种计算机视觉任务中,包括目标检测、图像分类等^[8]。根据公开测试报告 Resnet50 在 ImageNet 验证集上的 Top-1 错误率为 20.73%, 相应的 Top-1 准确率为 79.27%, 接近原始论文报告的 79.8%, 这表明 Resnet50 具有较高的性能和准确性非常适合本文所述方法的实验,其结构如图13所示。

实验环境: CPU Intel Core i9, 操作系统 Windows11 专业版, GPU NVIDIA GeForce RTX4080, Python 3.11.5; 实验数据每组均为: 训练集 16 297 件, 测试集 3 287 件; 测试轮数每组均为: 125 轮。经过测试, 根据 Cutout 训练曲线看到有剧烈波动如图14, 说明有可能存在上文所述

问题,造成训练集和验证集相差太大引起震荡。经过Gridmask改进后曲线前期波动较小后期区域平稳如图15,可见此方法对训练有较好的改善。

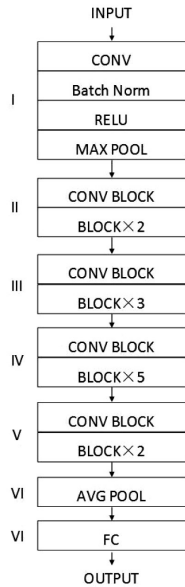


图13 Restnet50网络结构

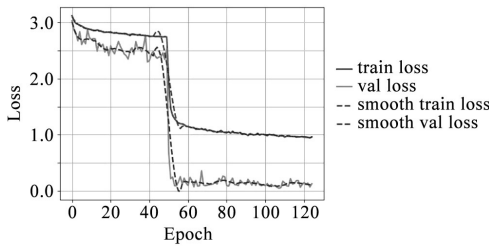


图14 Cutout方法

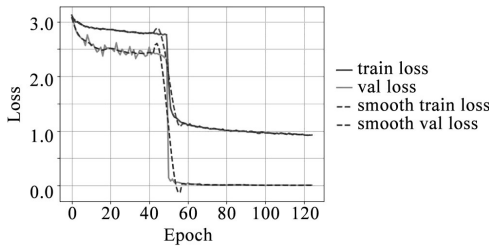


图15 Gridmask方法

数据增强前后对真实钢包编号识别测试对比如图16所示。

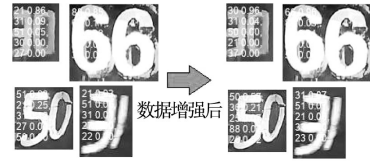


图16 测试结果

4 结束语

本文通过深度学习分类识别的方法对炼钢企业生产过程中的钢包非印刷字体编号进行识别,通过基于深度学习中数据增强的方法不仅扩增了数据集并且增强训练模型的泛化能力,可以使得识别更加贴近真实场景。本文研究内容不仅为炼钢企业提供了更为高效、精确的编号识别手段,同时也为其他类似工业生产领域的识别应用提供了新的思路和方法。

参考文献:

- [1] 张勇. 基于PP-YOLO的农业病虫害识别算法[J], 中国果菜, 2024, 44(5): 80-87.
- [2] 曹义亲. 基于E-YOLOX的实时金属表面缺陷检测算法[J], 图学学报, 2023, 44(4): 677-690.
- [3] 陈启鹏. 基于YOLOV4的工件表面质量在线检测方法研究[J], 包装工程, 2023, 44(3): 148-156.
- [4] 王九鑫. 基于改进YOLOv5的金属表面缺陷分割[J], 组合机床与自动化加工技术, 2024(5): 171-176, 182.
- [5] 陈千懿. 基于强化学习的配电网巡检缺陷自动识别方法[J], 自动化技术与应用, 2023, 42(10): 54-58.
- [6] Terrance Devries. Improved Regularization of Convolutional Neural Networks with Cutout[DB/OL]. [2017], <https://arxiv.org/abs/1708.04552>.
- [7] Pengguang Chen. GridMask Data Augmentation[DB/OL]. [2020], <https://arxiv.org/abs/2001.04086>.
- [8] Kaiming He. Deep Residual Learning for Image Recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016.

作者简介: 吕宜光(1989-), 男, 硕士, 助理研究员, 研究方向: 计算机技术与应用。

通信作者: 周丽丽(1980-), 女, 博士, 研究员, 研究方向: 类脑人工智能。

(上接第149页)

[3] 李宁, 张伟, 王海磊. 高干热现场环境下智能电能表可靠性验证方法[J]. 电测与仪表, 2024, 61(3): 192-198.

[4] 郭利, 周建勇, 何达, 等. CCD图像传感器电极间漏电流自动化测试系统设计[J]. 半导体光电, 2020, 41(2): 287-290.

[5] 李斌, 范玲, 姚斌, 等. 数据自同步条件下的配网馈线电流差动保护[J]. 电力系统及其自动化学报, 2020, 32(10): 1-8.

[6] 张鹏, 王玮, 徐丙垠, 等. 基于自适应功率输出控制的电流互感器取能电源设计方法[J]. 电力系统自动化, 2020, 44(3): 194-200.

[7] 肖遥剑, 冯浩, 仰韩英, 等. 永磁无刷直流电动机换相脉动转

矩及抑制方法[J]. 计量学报, 2019, 40(1): 58-63.

[8] 王思桦, 刘型志, 宣姝, 等. 基于线损等效的视在功率和功率因数计算方法研究[J]. 电测与仪表, 2024, 61(9): 65-70.

[9] 王清亮, 杨博, 高梅, 等. 基于等效电导的矿井电网智能漏电保护方法[J]. 工矿自动化, 2020, 46(6): 59-64.

作者简介: 刘刚(1979-), 男, 本科, 高级工程师, 研究方向: 电力系统分析。