

改进YOLOv5网络的钢材表面缺陷检测算法

张 倩

北京航空航天大学软件学院 北京 100191

摘要: 当前工业钢材表面缺陷检测面临小目标、多尺寸、准确率低等难题, 本文在YOLOv5s算法基础上改进, 添加注意力机制并在输入端改用K-Means++算法预测先验框, 提高了算法对小目标的检测性能。实验表明, 改进算法在数据集上mAP达83.3%, 比原算法提升6.2%, 检测速度达96.5fps/s。改进YOLOv5s算法在钢材表面缺陷检测上达到了高精度实时性的效果。

关键词: 缺陷检测; K-Means++; 注意力机制

Improving the steel surface defect detection algorithm of YOLOv5 network

Qian Zhang

School of Software, Beijing University of Aeronautics and Astronautics, Beijing 100191

Abstract: The current surface defect detection of industrial steel faces challenges such as small targets, multiple sizes, and low accuracy. This paper proposes improvements based on the YOLOv5s algorithm, incorporating an attention mechanism and utilizing the K-Means++ algorithm at the input end to predict anchor boxes, thereby enhancing the algorithm's detection performance for small targets. Experimental results demonstrate that the improved algorithm achieves an mAP of 83.3% on the dataset, which is a 6.2% improvement compared to the original algorithm, while maintaining a detection speed of 96.5fps/s. The improved YOLOv5s algorithm achieves high precision and real-time performance for surface defect detection in steel materials.

Keywords: defect detection; K-Means ++; attention mechanism

引言:

过去常使用人工目测法检验钢材表面缺陷, 随着机器视觉发展, 传统基于图像处理的算法逐渐成为主流, 通过提取缺陷特征进行检测, 但速度慢、泛化适应能力弱。

随着深度学习开始应用, 为克服传统图像处理方法提取特征的局限性, 基于深度学习的钢材表面缺陷检测涌现出很多模型, 为企业提供了更多可能性。杨莉等^[1]学者提出改进的基于Faster R-CNN^[2]的目标检测算法, 为解决使用RPN算法特征融合时存在的不足, 选用GA-RPN算法。Kaiming He等引入实例分割, 提出了Mask R-CNN^[3]网络, 较大提升了检测精度。双阶段目标检测

算法检测精度高但检测速度慢不能满足实时性的要求^[4]。单阶段法提出后提升了目标检测的效率, 李维刚等^[5]基于YOLOv3算法利用加权K-means聚类定位目标, 提高了检测精度且速度保持在50fps。尽管单阶段检测法检测速度有很大的提高, 但是多尺度的小目标上的精度上仍然较低。为了满足企业实际需求, 本文基于YOLOv5网络进行钢材表面缺陷检测算法的研究, 改进锚框聚类添加注意力机制, 提升小目标检测精度, 达到更高精度和更快速度。

一、YOLOv5算法结构

YOLOv5包含4个不同版本的模型, 他们结构相同只是在网络深度和宽度上会有调整。如图1所示, YOLOv5s由输入端、骨干模型、颈部网络和预测头组成。图像输入后采用了Mosaic数据增强进行数据预处理, 增加样本数量。然后采用自适应算法计算先验框中心^[6]。

作者简介: 张倩(1996.11—), 女, 硕士生, 研究方向为目标检测, 通讯邮箱: huoshengdewo@163.com

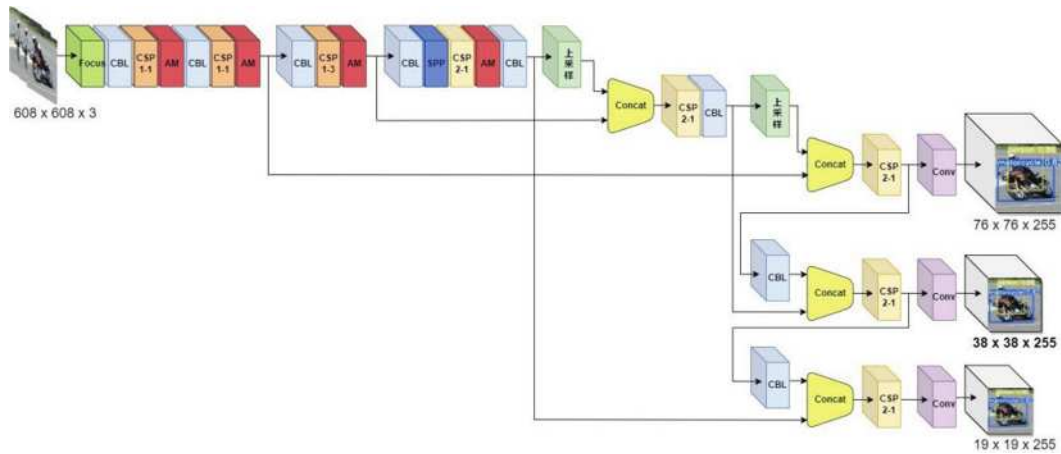


图1 改进后YOLOv5s网络结构

在骨干（Backbone）结构中，采用特征金字塔结构，降采样提取图像特征。在这其中也包含了Focus结构、空间金字塔池化层等进行卷积、池化操作。颈部网络则在Backbone特征提取后进行特征融合操作，除了FPN还加入了自底向上的PANet结构^[7]传递位置信息。预测头部分进行全卷积并concat输出检测结果，损失函数上使用GIoU计算边界框的损失^[8]。

二、改进的缺陷检测算法

1. 预测框重新聚类

本文对YOLOv5s算法中输入端的预测框计算方式进行改进，原YOLOv5自适应聚类方法受初始聚类中心影响，需要多次回归计算才能收敛，增加了计算时长而且检测精度较低，本文改用K-Means++聚类先验框中心。K-Means++聚类算法采用加权距离，更好地选取初始中心，适应多尺度的目标。聚类过程如下图2，首先K设为9，同原算法一致，回归所得预测框中心为（25.1, 37.8）、（35.6, 69.2）、（90.4, 53.7）、（44.1, 128.8）、（31.8, 189.4）、（170.1, 58.1）、（102.2, 100.6）、（91.8, 182.3）、（176.2, 182.5）。以不同中心的纵横坐标作为长和宽可以绘制先验框。

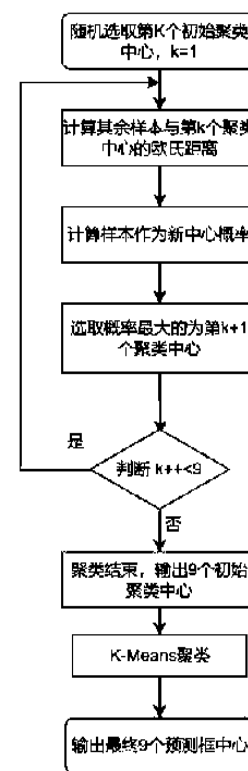


图2 聚类流程图

2. SENet注意力机制

注意力机制可以帮助网络聚焦关键的小目标特征信息，本文采用SENet模块。SENet建立了特征在通道内的联系，根据不同通道的权重进行迭代更新，筛选出通道

上的关键信息，其网络结构如图3。输入首先经过压缩处理，全局平均池化后成C个通道的向量，然后进入两个全连接层网络，采用激活函数归一化后得到各通道的权重系数，最后执行Scale将权重加权到原通道上，得到

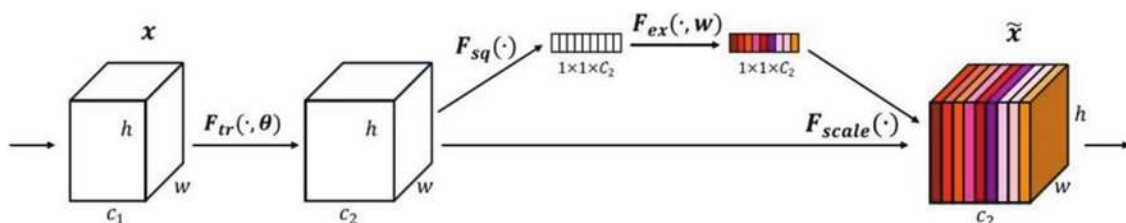


图3 SENet模块结构图

不同特征图的重要度。由于多次全连接，SENet模块参数数量较多，CBMA相对更轻量化一些。本文在YOLOv5s网络Backbone和Neck部分CSPi-X模块的卷积操作中引入SENet注意力机制。

三、实验及结果分析

1. 数据集及实验环境

为了验证本文基于YOLOv5改进的钢材表面缺陷检测算法的有效性，实验采用东北大学发布的带钢表面缺陷数据集NEU-DET，共有6个类别，并混合在企业采集的数据集进行扩充。首先对自采数据集进行LabelImg标注，再以NEU-DET的命名规则命名，后缀名为xml保持图片名称一致。本文按照8:2随机把混合数据集进行划分，如表1所示。

表1 数据集分类

实验数据集	数量
总数据集	2100
各类数据集	350
训练集	1680
测试集	420

实验环境为Windows10操作系统，处理器为Inter Core i7-9700k，NVIDIA Geforce GTX 1080显卡，学习框架为Pytorch 1.9版本。

2. 评价指标

目标检测的评价指标有很多，本文使用召回率 (Recall)、精确度 (Precision) 和平均精确度均值 (mAP) 三个指标来评估算法的检测精度。各评价指标的计算如下式：

$$mAP = \frac{\sum_{i=1}^{M-1} AP_i}{M} \quad (1)$$

式中：AP_i为每类目标的平均准确度，M代表检测类别数。我们可以使用P-R曲线来同时展示准确率和召回率这两个指标，AP为：

$$AP = \int_0^1 P(r)dr, r \in (0,1) \quad (2)$$

且

$$R = \frac{TP}{TP + FN} \quad (3)$$

$$P = \frac{TP}{TP + FP} \quad (4)$$

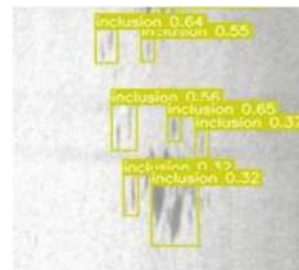
式中：TP是检测正确的样本数，FP为正样本被误检的样本数，FN为正样本被漏检的数量。故召回率R表示模型检测是否完全，精确度P评估预测是否准确。

3. 实验结果分析

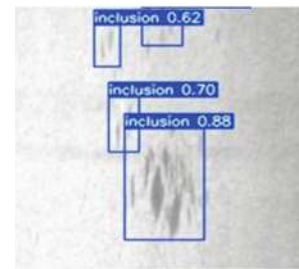
本文使用改进后的YOLOv5算法在混合数据集上进

行消融实验。实验采用在COCO和ImageNet上预训练好的YOLOv5模型，训练时部分重要参数设置为：输入图片大小扩展为640×640，迭代300轮，batch-size为64，优化算法选用SGD，动量因子为0.937，初始学习率为0.01，最终学习率0.0001，采用余弦退火方式调节动态调节学习率。

实验结果如表2。本文进行实验1对比YOLOv5s算法原自适应计算预测框和K-Means++聚类预测框两种方法，验证K-Means++聚类算法在多尺度检测上的有效性。在相同输入条件下，检测效果的对比如图4所示。结果表明使用K-Means++聚类预测框更贴合缺陷目标真实尺寸，更适应多尺度的缺陷目标。



原算法预测框



K-means++ 聚类预测框

图4 不同锚框计算检测效果对比

在网络中引入SENet注意力机制进行实验2，可以看出较原算法mAP提高2.1%，它利用了特征的通道关系，关注目标的位置信息，更满足定位任务，但同时增加了参数量，降低了推理速度。本文将两种策略结合改进YOLOv5s网络进行实验3，mAP提升6.2%，改进后的网络在P和R上都超过了原YOLOv5s网络，P-R曲线比较如图5，可见算法改进后各类缺陷AP都提升了，特别crazing提升到了42.0%和39.7%。

表2 不同改进策略检测精度对比

YOLOv5s	K-Means++	SENet	R _p / %	R _r / %	mAP	速度 fps/s
√			51.9	79.1	77.1	101
√	√		55.3	79.5	79.2	125
√	√	√	57.9	85.2	83.3	96.5

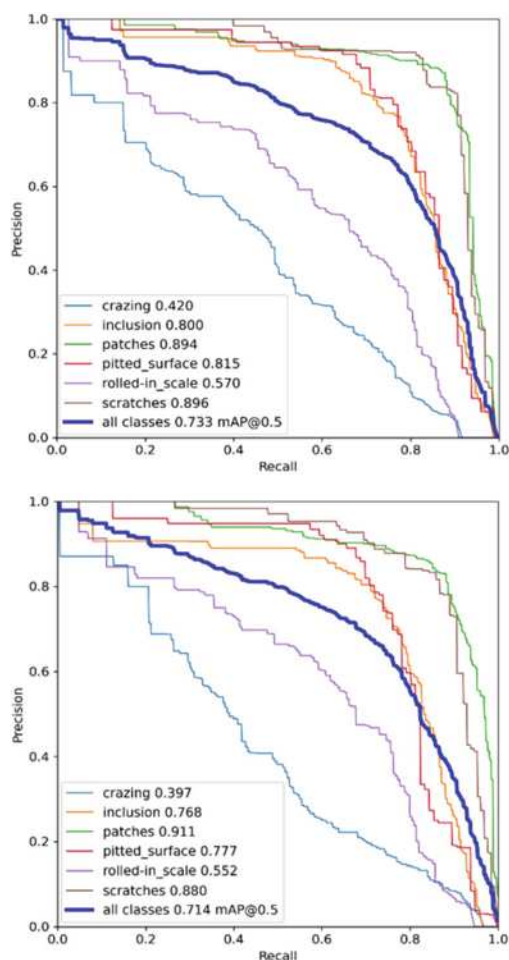


图5 原YOLOv5和改进后的算法P-R曲线性能对比

通过融合实验，我们可以看出融合后的策略检测精度最高达83.3%，因此得到表现最优的改进YOLOv5算法。该算法在检测速度方面大小略有下降，原因考虑注意力机制的引入。它在像素之间建立了通道联系，会增加计算复杂度，需要占用一定时间学习。图6展示了改进后的算法在钢材表面缺陷检测任务中的表现，针对不同的类别，检测到的框与人工标注的框非常接近，表明算法具有较高的准确度，取得更精准实际检测效果。

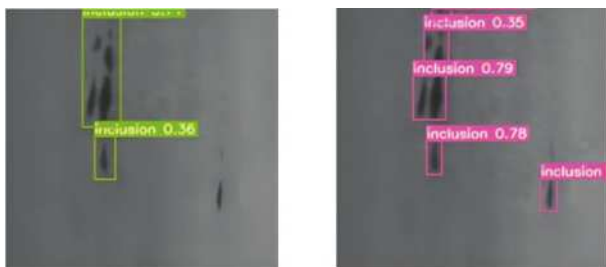


图6 YOLOv5s和改进算法的检测效果对比

四、结束语

当前钢材表面缺陷检测，由于目标小，特征提取的效果差，检测精度偏低，本文引入SENet模块增加小目

标权重，提高检测准确率。同时由于原K-Means算法初始簇心随机指定，容易收敛到局部最优解。为解决该问题，使用K-Means++算法重新聚类先验框中心，使用加权距离寻找相互距离较远的初始聚类中心，优化后所得的先验框能够贴合多种不同尺寸的缺陷，进一步适合钢材表面缺陷，有助于提升检测精度。最后采用自建数据集进行钢材表面缺陷检测，结果表明，改进的YOLOv5s算法在不增加过多参数数量的条件下，融合两种改进策略，检测精度提升了6.2%。但是，增加注意力机制后会使得模型训练时间和存储空间增加。改进后的缺陷检测算法检测速率有所减小，FPS下降了4.5%。虽然在牺牲计算空间的情况下，换取检测精度的提升，但是仍然能够满足工业实际需求。这样的性能提升使得该算法在实际工业生产中具有较高的应用价值。未来可以针对类内差异大的目标实现更精准检测，也可以进一步轻量化模型方便网络部署到实际移动设备终端。

参考文献：

[1]杨莉, 张亚楠, 王婷婷, 等.基于改进FasterR-CNN的钢材表面缺陷检测方法[J].吉林大学学报(信息科学版), 2021, 39(4): 409-415.

[2] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[3] He K, Gkioxari G, Dollár P, et al. Mask R-CNN [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, PP (99): 1-1.

[4]丁瑶.基于融合机制的航拍目标检测与识别[D].厦门: 厦门大学, 2019.

[5]李维刚, 叶欣, 赵云涛, 等.基于改进YOLOv3算法的带钢表面缺陷检测[J].电子学报, 2020, 48(7): 1284-1292

[6]谈世磊, 别雄波, 卢功林, 等.基于YOLOv5网络模型的人员口罩佩戴实时检测[J].激光杂志, 2021, 42(2): 147-150.

[7] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, June 18-23, 2018: 8759-8768.

[8] Wang, H., Jin, Y., Ke, H. et al. DDH-YOLOv5: improved YOLOv5 based on Double IoU-aware Decoupled Head for object detection. J Real-Time Image Proc 19, 1023 - 1033 (2022).