

控制与决策

Control and Decision

基于多尺度轻量化注意力的钢材缺陷检测

周彦, 孟江南, 王冬丽, 谭雅琴

引用本文:

周彦, 孟江南, 王冬丽, 谭雅琴. 基于多尺度轻量化注意力的钢材缺陷检测[J]. *控制与决策*, 2024, 39(3): 901–909.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2022.0997>

您可能感兴趣的其他文章

Articles you may be interested in

[基于自注意力生成对抗网络的图像超分辨率重建](#)

Image super-resolution reconstruction based on self-attention GAN

控制与决策. 2021, 36(6): 1324–1332 <https://doi.org/10.13195/j.kzyjc.2019.1290>

[Anchor-free的尺度自适应行人检测算法](#)

Anchor-free scale adaptive pedestrian detection algorithm

控制与决策. 2021, 36(2): 295–302 <https://doi.org/10.13195/j.kzyjc.2020.0124>

[基于协同聚类和权重注意力稀疏自编码网络的变化检测方法](#)

Change detection approach based on cooperative clustering and weighted-attention sparse autoencoder

控制与决策. 2021, 36(10): 2442–2450 <https://doi.org/10.13195/j.kzyjc.2019.1633>

[一种基于深度学习的时间序列预测方法](#)

A time series prediction method based on deep learning

控制与决策. 2021, 36(3): 645–652 <https://doi.org/10.13195/j.kzyjc.2019.0809>

[一种基于多层语义特征的图像理解方法](#)

An image understanding method based on multi-level semantic features

控制与决策. 2021, 36(12): 2881–2890 <https://doi.org/10.13195/j.kzyjc.2020.0927>

基于多尺度轻量化注意力的钢材缺陷检测

周彦^{1†}, 孟江南¹, 王冬丽¹, 谭雅琴²

(1. 湘潭大学 自动化与电子信息学院, 湖南 湘潭 411105;

2. 湘潭大学 计算机学院·网络空间安全学院, 湖南 湘潭 411105)

摘要: 针对当前 YOLOv5 算法检测钢材表面缺陷精度不高、速度慢等问题, 提出一种基于多尺度轻量化注意力的 YOLO-Steel 钢材表面缺陷检测方法. 首先, 提出一种轻型通道注意力模块, 仅需少量计算成本即可有效关注重要通道; 然后, 利用空洞卷积扩大感受提出一种轻型空间注意力模块, 能够在空间维度上提取有价值信息; 接着, 提出金字塔注意力结构, 利用多级池化放缩特征图在不同分辨率特征图上使用空间注意力模块学习其空间依赖信息, 对多级特征图使用通道注意力模块重构其通道相关信息, 改善检测效果. 实验结果表明, YOLO-Steel 在钢材表面缺陷数据集上平均精度均值 (mAP) 可达 77.2%, 比 YOLOv5s 算法提高 1.8%, 模型时间、空间复杂度与 YOLOv5s 基本持平, 在保证检测速度的基础上能够有效提高精确度.

关键词: YOLOv5; 钢材表面缺陷检测; 注意力机制; 金字塔多级注意力

中图分类号: TP183

文献标志码: A

DOI: 10.13195/j.kzyjc.2022.0997

引用格式: 周彦, 孟江南, 王冬丽, 等. 基于多尺度轻量化注意力的钢材缺陷检测 [J]. 控制与决策, 2024, 39(3): 901-909.

Steel defect detection based on multi-scale lightweight attention

ZHOU Yan^{1†}, MENG Jiang-nan¹, WANG Dong-li¹, TAN Ya-qin²

(1. School of Automation and Electronic Information, Xiangtan University, Xiangtan 411105, China; 2. School of Computer Science · Cyberspace Security, Xiangtan University, Xiangtan 411105, China)

Abstract: Aiming at the problems that the current YOLOv5 algorithm detects steel surface defects with low accuracy and slow speed, a YOLO-Steel steel surface defect detection algorithm is proposed. First, a light-weight channel attention module is proposed, which can effectively focus on important channels with only a small computational cost. Secondly, by using atrous convolution to expand the receptive field, a light-weight spatial attention module is proposed. Finally, a pyramid attention structure is proposed, which uses multi-level pooling to scale feature maps, and uses spatial attention modules on feature maps of different resolutions to learn its spatial dependence information. After splicing in dimensions, the channel attention module is used to reconstruct its channel-related information, which can achieve better detection results for multi-scale detection targets. The experimental results show that the average mean precision (mAP) of YOLO-Steel on the steel surface defect data set can reach 77.2%, which is 1.8 percentage points higher than that of the YOLOv5s algorithm, and the model time and space complexity are basically the same as those of YOLOv5s. On the basis of ensuring the detection speed, the accuracy is effectively improved.

Keywords: YOLOv5; steel surface defect detection; attention mechanism; pyramid multi-level attention

0 引言

钢材在国民生活各个方面都扮演着重要角色, 在制造过程中受限于工艺水平和生产环境等, 表面会产生多种类型的缺陷, 影响钢材抗疲劳性、耐腐蚀性以及抗高压等性能, 对其进行缺陷检测是保证钢材质量的重要步骤. 《中国制造 2025》战略指出: 推进制造过程智能化. 因此, 研究一种智能化的钢材表面缺陷

检测算法具有重要意义.

早期钢材表面缺陷检测方法有频闪法^[1], 工人利用高频高强度的闪光灯发现钢材中存在的缺陷, 检测准确率极大地依赖于工人的工作状态和经验水平. 此外还有红外检测法^[2]、漏磁检测法^[3]、涡流检测法^[4]和激光扫描检测法等. 传统的计算机视觉缺陷检测通过 CCD 相机采集图像信息, 再利用信号处

收稿日期: 2022-06-07; 录用日期: 2022-12-20.

基金项目: 国家自然科学基金项目 (61773330); 湖南省国家应用数学中心项目 (2020YFA0712503); 湖南省教育厅科研项目 (19C1740); 湖南省科技计划项目 (2020GK2036); 上海市科委项目 (19511120900).

†通讯作者. E-mail: yanzhou@xtu.edu.cn.

理程序分析局部异常获得缺陷信息^[5-10]. 这些传统的缺陷检测方法都存在一定的局限性,难以满足高强度现代化高速生产过程中高效率的需求.

随着深度学习的迅速发展,卷积网络因强大的端到端特征提取能力^[11]以及复杂环境下的强鲁棒性^[12],逐渐被广泛应用于缺陷检测任务^[13-15]. 郭龙源等^[16]提出基于Mask R-CNN的缺陷检测算法,用于解决传统缺陷检测算法难以准确分割其中缺陷的问题. Xing等^[17]改进YOLOv3网络结构,通过CutMix方法扩充小缺陷数量提出CSYOLOv3方法,能够有效检测工件表面的小缺陷. 刘艳菊等^[18]将YOLOv4中特征提取网络CSPDarknet53换为轻量级深层神经网络MobileNetv3以提高检测速度,并重新定义置信度损失以解决正负样本不平衡问题,有效提高了热轧钢条表面的缺陷检测精度. 目前,YOLO系列目标检测算法已经迭代到YOLOv5^[19-23],相比于YOLOv3^[24]精度上有较大提升,相比于YOLOv4^[25]模型体积小速度快,能够很好地满足钢材生产中精度与实时性的要求.

综上所述,为了在检测性能与计算代价之间取得平衡,本文选取YOLOv5作为基网络. 同时,一方面为了减少计算成本,提出了轻型空间、通道注意力模块;另一方面,针对钢材表面缺陷尺度差异性大的问题,提出了金字塔结构注意力模块,利用多级池化放缩特征图,并在多级池化尺度上进行空间注意力运算,通过关注多尺度信息获得更好的检测效果. 最后,结合二者提出一种新的用于钢材表面缺陷的检测网络YOLO-Steel,并在NEU-DET钢材表面缺陷数据集上

验证模型的有效性,与当前主流检测模型相比均取得了有竞争力的结果.

1 算法模型

1.1 网络结构

YOLO-Steel网络框架如图1所示. 其网络结构主要分为3个部分: Backbone、Neck和Head. Backbone主要用来实现特征提取,包括Conv、Focus、C3和SPP-LCBAM模块,其中Conv是基本卷积单元,用来完成卷积运算、归一化和激活运算;Focus首先将特征图slice成4份,再做通道维度上Concat,能够聚焦长宽维度信息到通道维度;C3由Bottleneck模块和CSP结构^[26]组成, Bottleneck先是 1×1 的卷积,然后是 3×3 卷积,若为True,则会有一层残差结构与初始输入相加, CSP结构源自跨阶段局部网络CSPNet,能够优化网络中的重复梯度信息,减少计算量; SPP-LCBAM为本文所提出模块,能够代替基网络YOLOv5s的SPP模块,其中提取空间依赖信息和融合通道信息采用轻型注意力模块实现,这样既能在不增加额外网络模块及复杂性的基础上继承原网络SPP模块的功能,又能充分利用多级池化后不同分辨率的特征图信息,优化其空间局部信息提取能力,从而使整体缺陷检测取得更好的效果. Neck主要用来实现特征融合,包括FPN和PAN,FPN在网络中自上而下传递语义信息,PAN则自下而上传递定位信息,能够有效融合不同深度的特征图信息. Head有3个检测头,分别检测不同深度的特征图,能够获取大、中、小3个尺寸的目标信息.

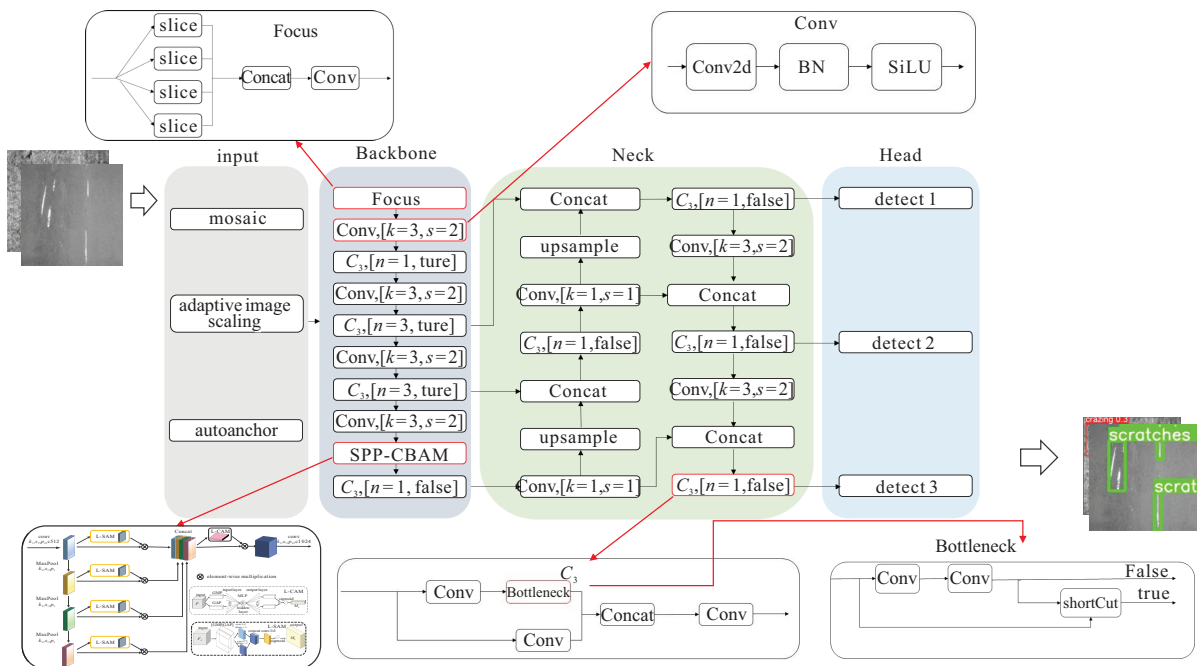


图1 YOLO-Steel网络结构

1.2 轻型注意力模块

本文在CBAM^[23]的基础上进行优化,提出一种轻量级注意力模块,由L-CAM(light channel attention module)和L-SAM(light spatial attention module)两个子模块组成,该模块能有效关注重要通道以及在空间维度上对有价值的信息进行提取.相比于CBAM,能够在几乎不降低原性能的前提下,一定程度上减少参数数量和运算量.

L-CAM是轻型通道注意力模块,其结构如图2所示.对于输入特征图,该模块能学习到一个通道维度

的权重向量,将该向量与原输入特征图进行逐元素乘法加权运算得到带通道注意力的特征图,目的是对特征图各个通道重要性进行判断,能够有效提升重要通道,抑制非重要通道,其过程表达式为

$$M_c = \sigma\{\text{MLP}[\text{GAP}](F) \oplus \text{MLP}[\text{GMP}(F)]\}. \quad (1)$$

其中MLP为权值共享的多层感知机,其隐层仅含有一个神经元 y ,表达式为

$$y(x) = \sigma\left(\sum_{i=1}^n w_i x_i + b\right). \quad (2)$$

这里: w_i 为可学习权重参数, b 为偏置系数.

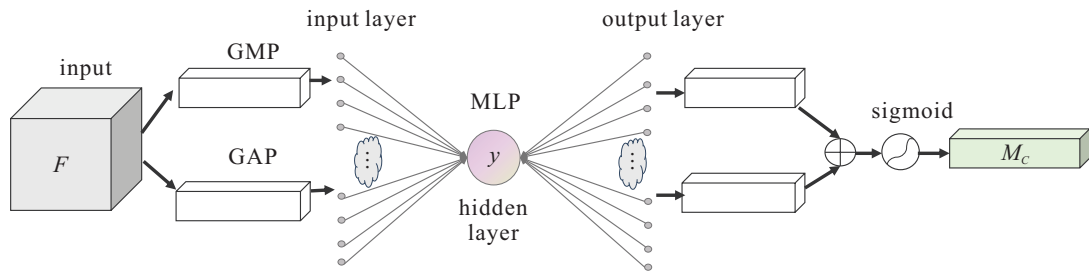


图2 L-CAM示意图

输入特征图首先分别进行全局平均池化和全局最大池化得到两条向量,然后将其作为输入层送入隐层仅为单神经元的多层感知机,多层感知机对于两条向量权值共享,通过隐藏神经元捕捉到输入之间复杂的相互作用.再对其输出向量进行逐元素相加操作并送入激活函数,得到通道注意力掩码.相比于CAM,可以减少多层感知机隐层神经元个数,但是并不改变其基本结构,既保证了其功能性,又突出了轻量化的特点,计算量减少了 $(C_{in} + C_{out}) \times (C_{mid} - 1)$.其中: C_{in} 、 C_{out} 和 C_{mid} 分别为输入层、输出层和隐藏层

神经元个数.

L-SAM是轻型空间注意力模块,其结构如图3所示.相比于SAM利用 7×7 的卷积获取全局空间信息,L-SAM利用一个 3×3 的标准卷积和一个扩张率为 n 的空洞卷积,通过并行运算获得空间信息,再将两个卷积获得的特征图在通道维度上拼接,并通过一个卷积进行特征重要性合并以及还原通道数.其过程表达式为

$$M_s = \text{sig}\{f^{1 \times 1}\{f_{\text{dilate}=n}^{3 \times 3}[\text{GAP}(F_c), \text{GMP}(F_c)] + f_{\text{dilate}=1}^{3 \times 3}[\text{GAP}(F_c), \text{GMP}(F_c)]\}\}. \quad (3)$$

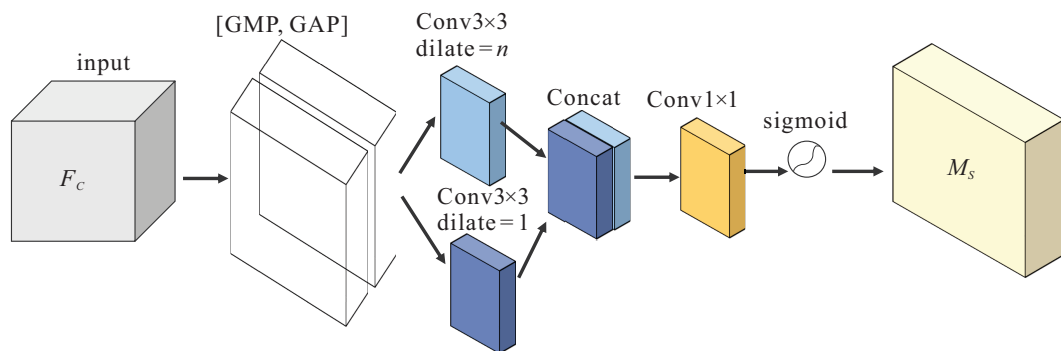


图3 L-SAM示意图

空洞卷积可以在增大感受野的同时不引入额外参数,这样能够在减少运算量的同时获得相同的空间感受野.

假定输入特征图大小为 $D_{in} \times D_{in} \times C_{in}$,输出特征图大小为 $D_{out} \times D_{out} \times C_{out}$,对于卷积核为 7×7 的

标准卷积,其计算量为

$$7 \times 7 \times C_{in} \times C_{out} \times D_{in} \times D_{in}. \quad (4)$$

而采用两个 3×3 卷积并行提取特征,以及后续Concat和 1×1 标准卷积下采样操作,其计算量为

$$\begin{aligned} & 2 \times 3 \times 3 \times C_{in} \times C_{out} \times D_{in} \times D_{in} + \\ & 1 \times 1 \times 2C_{in} \times C_{out} \times D_{in} \times D_{in}. \end{aligned} \quad (5)$$

可以得出计算量比值为

$$\frac{\begin{cases} 2 \times 3 \times 3 \times C_{in} \times C_{out} \times D_{in} \times D_{in} + \\ 1 \times 1 \times 2C_{in} \times C_{out} \times D_{in} \times D_{in} \end{cases}}{7 \times 7 \times C_{in} \times C_{out} \times D_{in} \times D_{in}} = \frac{20}{49} \approx 0.408. \quad (6)$$

采用所提出的结构,能够在获取同样尺度感受野的同时,其计算量仅为原标准7×7卷积的0.408倍.

1.3 金字塔多级注意力结构

不同尺度的特征图具有不同的局部和全局语义信息,大尺度特征图包含更多全局信息,而小尺度特征图保留了更多物体边缘细节信息.因此,本文构建一个金字塔多级注意力结构,通过将局部特征与全局特征相融合,利用不同的分辨率聚合信息丰富特征图表达能力,有利于多尺度多目标的复杂物体检测.尤其是对于钢材表面缺陷检测,缺陷大小尺度丰富、边缘形态复杂,单一尺度容易丢失大量细节信息.

本文构建的具有多尺度感知能力金字塔多级注意力结构如图4所示.输入特征图 $F \in R^{C \times H \times W}$,首

先对其使用卷积核为1×1的卷积进行下采样,此时对于生成的特征图 F ,一方面引入空间注意力机制进行空间上重构,使其学习一个空间注意力的掩码,另一方面对其多次进行核为5×5的最大池化操作,每次池化操作后的特征图都分别通过空间注意力调整权重,这样做的好处是每次池化后能获得不同关注尺度的特征图,进而获得不同下采样倍率下带空间注意力的特征图.实际上,串行两个核为5×5的最大池化层其视野域等效为一个核为9×9的最大池化层,串行3个核为5×5的最大池化层其视野域等效为一个核为13×13的最大池化层.相比于直接用大尺度池化核并行进行池化操作,在上一层小尺度池化核得到的特征图上继续用小尺度池化核进行池化操作,其计算效率更高,因为多级池化会更充分地利用到上级池化的结果.不同池化尺度能关注到不同尺度空间物体的边缘细节信息,再对其分别进行空间注意力权重调整,有利于缓解目标尺度变化大以及目标重叠问题.将不同池化尺度获得的带空间注意力的特征图在其通道维度上拼接起来,通过通道注意力机制学习对其进行权重调整,调节各个通道关注度,最后通过一个1×1卷积得到输出特征图 $F' \in R^{C \times H \times W}$.

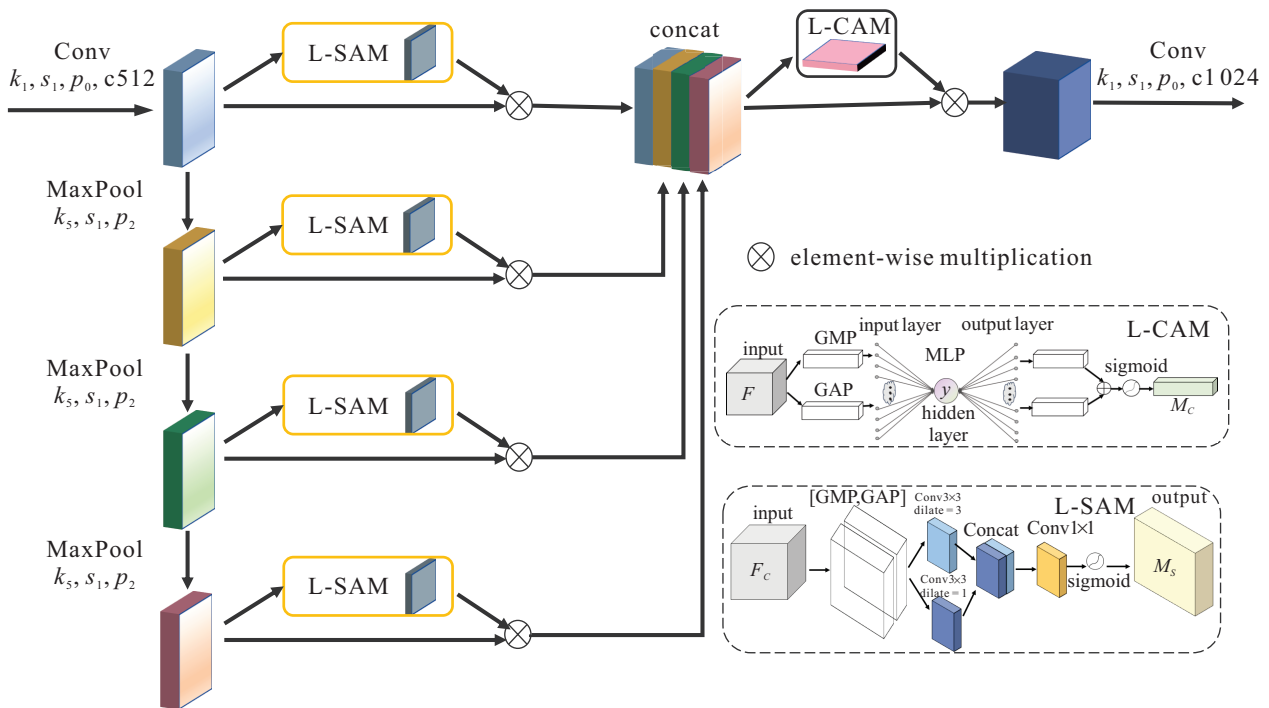


图4 金字塔多级注意力结构

2 实验结果与分析

2.1 数据集

实验使用NEU-DET数据集,如图5所示.该数据集有6类钢材表面缺陷,包括氧化铁(rolled-in scale)、

斑块(patches)、裂纹(crazing)、麻点(pitted surface)、夹杂(inclusion)和划痕(scratches),共1800张,以7:3比例随机抽取划分为训练集与测试集.

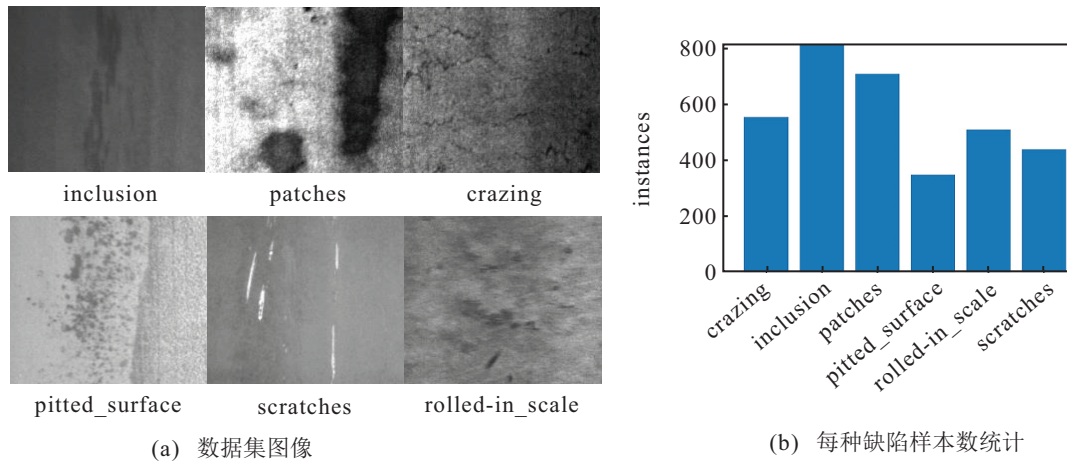


图5 数据集及其各类别样本数统计

2.2 实验参数设置

实验环境如下: Intel i5-9300H 处理器, Nvidia GeForce GTX 1660 Ti 显卡, 16 GB 内存, 操作系统为 Windows 10, 深度学习框架为 PyTorch, 优化器为 SGD, 初始学习率为 0.01, 动量为 0.937, 权重衰减因子为 0.005, Batchsize 为 16. 训练过程如图 6 所示, 经过 275 轮迭代, 模型回归框定位损失 (box_loss)、分类损失 (cls_loss) 和置信度损失 (obj_loss) 均基本收敛。

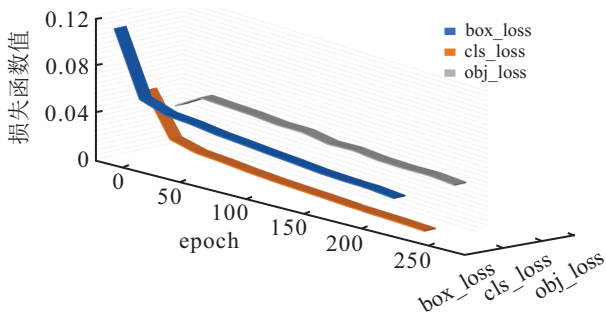


图6 训练损失曲线

2.3 评价指标

实验使用检测速度 FPS (frames per second)、检测精度 mAP (mean average precision)、查准率 Precision、查全率 Recall、计算量 FLOPs (floating-point operations per second) 和参数量 Parameters 等评估模型性能。

查准率是被正确检索的样本数与被检索到样本总数的比值, 查全率是被正确检索的样本数与应当被检索到的样本数的比值, 其表达式分别为

$$P = \frac{TP}{TP + FP}, \quad (7)$$

$$R = \frac{TP}{TP + FN}. \quad (8)$$

其中: 真阳性 (true positive, TP) 表示检测正确的目标数量, 假阳性 (false positive, FP) 表示检测错误的目标数量, 假阴性 (false negative, FN) 表示漏检的目标数量。

FLOPs 表示浮点运算次数, 用来衡量算法时间复杂度, GFLOPs 即为每秒十亿次的浮点运算. 对于卷积层, FLOPs 计算公式为

$$F_{\text{FLOPs}} = H \times W [C_{\text{in}} \times K^2 + (C_{\text{in}} \times K^2 - 1) + 1] \times C_{\text{out}}. \quad (9)$$

Parameters 表示模型含有多少参数, 用来衡量算法空间复杂度. 对于卷积层, 其参数量计算公式为

$$P_{\text{Conv}} = (K_w \times K_h \times C_{\text{in}} + 1) \times C_{\text{out}}. \quad (10)$$

其中: H 和 W 分别为输入特征图的高和宽, C_{in} 和 C_{out} 分别为输入和输出特征图通道数, K_w 和 K_h 分别为卷积核的宽和高, K 为卷积核大小。

2.4 实验评估与对比分析

2.4.1 改进方法对模型性能的影响

在 YOLOv5s 的特征提取部分加入轻型注意力模块, 以学习特征图空间位置信息与通道相互依赖关系. 为了验证轻型注意力模块性能, 讨论轻型注意力模块中膨胀系数对实验结果的影响, 进行不同的参数实验。

NEU-DET 数据集上对 L-CAM 参数 n 的消融实验如表 1 所示. 首先是原始 YOLOv5s, 此时缺少空间注意力模块聚焦位置依赖关系, 在 SPP 模块前加入 CBAM 注意力模块后, mAP 有所上升, 用 L-CBAM 替换 CBAM, 精度略有下降, 但模型参数有一定程度减少. 随着膨胀系数的增加, 模型的空间视野逐步加大, 但 mAP 并没有与膨胀系数的增加呈正相关关系, 在 $n = 3$ 时精度最高, 相比于 CBAM, 取得了性能与计算量之间的一个平衡。

最终选取膨胀系数 $n = 3$ 进行后续实验, 此时模型 mAP 为 76.0%. 由于所提出的金字塔注意力模块在 SPP 模块基础上增加注意力提取特征, 功能性超越了 SPP 模块, 用金字塔注意力模块替换原 YOLOv5 中

的SPP模块,既保留了原有的SPP模块功能,又通过多级注意力机制,使模型能更好地学习到不同尺度的空间相互依赖性信息,对于场景的判断更有利。

表1 NEU-DET数据集上对L-CAM参数n的消融实验

method	Parameters	mAP/%
YOLOv5s	7035811	75.4
YOLOv5s+CBAM	7068677	76.1
YOLOv5s+L-CBAM(n=1)	7036873	75.9
YOLOv5s+L-CBAM(n=2)	7036873	75.9
YOLOv5s+L-CBAM(n=3)	7036873	76.0
YOLOv5s+L-CBAM(n=4)	7036873	75.8

NEU-DET数据集上对金字塔多级注意力模块的消融实验的实验结果如表2所示。算法的时间复杂度和空间复杂度分别由计算量GFLOPs和参数量Parameters进行衡量,其中YOLOv5s+SPP-SAM表示金字塔多级注意力结构中仅加入空间注意力模块SAM, YOLOv5s+SPP-CAM表示金字塔多级注意力结构仅加入通道注意力模块。结果表明,金字塔多级注意力模块仅会增加算法少量时间复杂度和空间复杂度,但能强化原有SPP模块功能并明显提升检测精度,表明了其对于算法改进的有效性。

表2 NEU-DET数据集上对金字塔多级注意力模块的消融实验

method	GFLOPs	Parameters	mAP/%
YOLOv5s	15.9	7035811	75.4
YOLOv5s+SPP-SAM	15.9	7035909	76.7
YOLOv5s+SPP-CAM	16.1	7166883	76.6
YOLOv5s+SPP-CBAM	16.1	7166981	77.5

金字塔多级注意力结构中注意力模块使用所提出的L-CBAM进行消融实验,实验结果如表3所示。可以看到,通道注意力模块会带来模型时间复杂度的提升,而采用所提出轻型注意力模块空间复杂度仅少量提升,且对时间复杂度几乎无影响。相比于直接使用CBAM模块的SPP-CBAM, SPP-LCBAM在仅

表3 NEU-DET数据集上对轻型金字塔多级注意力模块的消融实验

method	GFLOPs	Parameters	mAP/%
YOLOv5s	15.9	7035811	75.4
YOLOv5s+SPP-SAM+LCAM	15.9	7037957	76.9
YOLOv5s+SPP-LSAM+CAM	16.1	7166921	77.0
YOLOv5s+SPP-LSAM	15.9	7035849	76.4
YOLOv5s+SPP-LCAM	15.9	7037859	76.8
YOLOv5s+SPP-LCBAM(n=3)	15.9	7037897	77.2

牺牲0.3%精度的基础上,计算量减少了0.2,参数量减少了129084。

NEU-DET数据集上的消融实验如表4所示。可以看出,金字塔注意力结构对于模型mAP提升明显,可以有效获取特征图多尺度空间位置依赖性,相比于基网络,能带来2.1%的精度提升。在金字塔多级注意力结构中使用轻型注意力模块后,算法的时间复杂度下降为与YOLOv5s一致,空间复杂度Parameters减少了129084,有效地降低了计算成本。综合考虑模型性能与计算成本,最终选取SPP-LCBAM即金字塔轻型注意力模块构建钢材表面缺陷检测网络YOLO_Steel。

表4 NEU-DET数据集上的消融实验

method	GFLOPs	Parameters	mAP/%
YOLOv5s	15.9	7035811	75.4
YOLOv5s+SPP-CBAM	16.1	7166981	77.5
YOLOv5s+SPP-LCBAM(n=3)	15.9	7037897	77.2

图7分别验证了YOLOv5s、YOLOv5s+SPP-CBAM、YOLOv5s+SPP-LCBAM(n=3)三种模型在6种类别缺陷上的AP对比结果。可以看出,金字塔注意力模块的加入在各个缺陷类别上检测精度均有提升,且轻量化的金字塔注意力模块在crazing缺陷上检测精度会超过普通金字塔注意力模块,验证了所提出金字塔多级注意力结构的有效性。

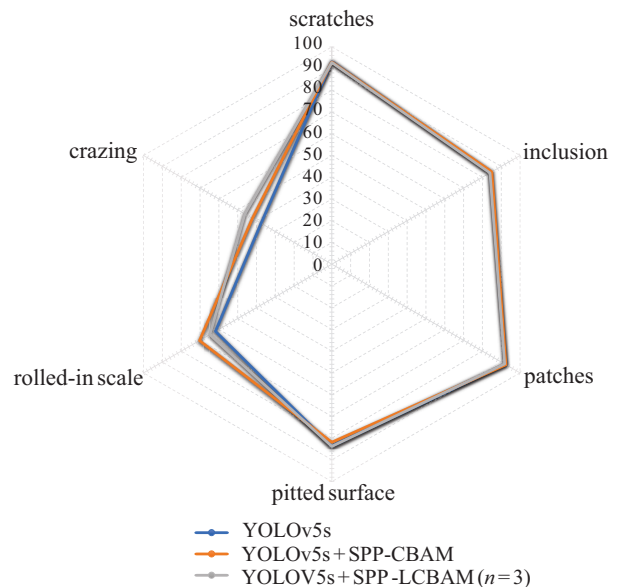


图7 缺陷各类别AP消融实验对比

YOLO_Steel消融实验结果如表5所示。第1组为原YOLOv5s算法,其mAP值为75.4%。第2组在YOLOv5s的SPP模块前加入注意力模块CBAM,可以看到模型mAP提升了0.7%,但是在空间复杂度上

参数量也随之增加了32 866. 第3组在前1组的基础上,将注意力模块CBAM替换为所提出的轻型注意力模块L-CBAM. 实验结果表明,相较于前一组,在mAP几乎不变的情况下,参数量明显减少,相较于原YOLOv5s算法,参数量仅少量增加了1 062,这与直接使用注意力模块其在空间复杂度上增加的存储成本减少了一个数量级,实现了注意力模块的轻量化. 第4组将轻型注意力模块与金字塔结构有效融合构建了SPP-LCBAM模块,用其替换原YOLOv5s网络中的SPP模块,一方面能够继承SPP模块的功能,实现多级池化特征图信息聚合,另一方面利用空间注意力模块分别对多级池化后的多级尺度特征图进行空间依赖信息的有效提取. 实验结果表明,使用金字塔结构的多级注意力提取方式比简单嵌入注意力模块的

模型提取特征能力有明显增强,且对计算参数量仅有少量增加,相较于原YOLOv5s算法,模型检测的查全率Recall提升了1.6%,表明通过金字塔多级注意力模块多级特征提取,模型捕捉检测目标的能力有明显提升. 此外,模型检测能力mAP提升了1.8%,表明了所提出的金字塔多级注意力结构在多尺度空间上提取信息的有效性.

图8为6类缺陷检测效果对比. 第1行为原始图片,第2行为YOLOv5s算法检测效果,第3行为所提出YOLO-Steel算法检测效果. 可以看出,YOLO-Steel相比于原YOLOv5s算法,在检测准确率以及对不同尺度缺陷的检测效果是上均有一定的提升,表明了所提出轻型注意力模块以及金字塔多级注意力结构的有效性.

表5 YOLO_Steel消融实验

method	CBAM	L-CBAM	SPP-LCBAM	Precision/%	Recall/%	Parameters	mAP/%
YOLOv5s	-	-	-	77.1	72.5	7 035 811	75.4
Improved 1	✓	-	-	75.0	72.1	7 068 677	76.1
Improved 2	-	✓	-	79.0	71.3	7 036 873	76.0
Improved 3	-	-	✓	73.9	74.1	7 037 897	77.2

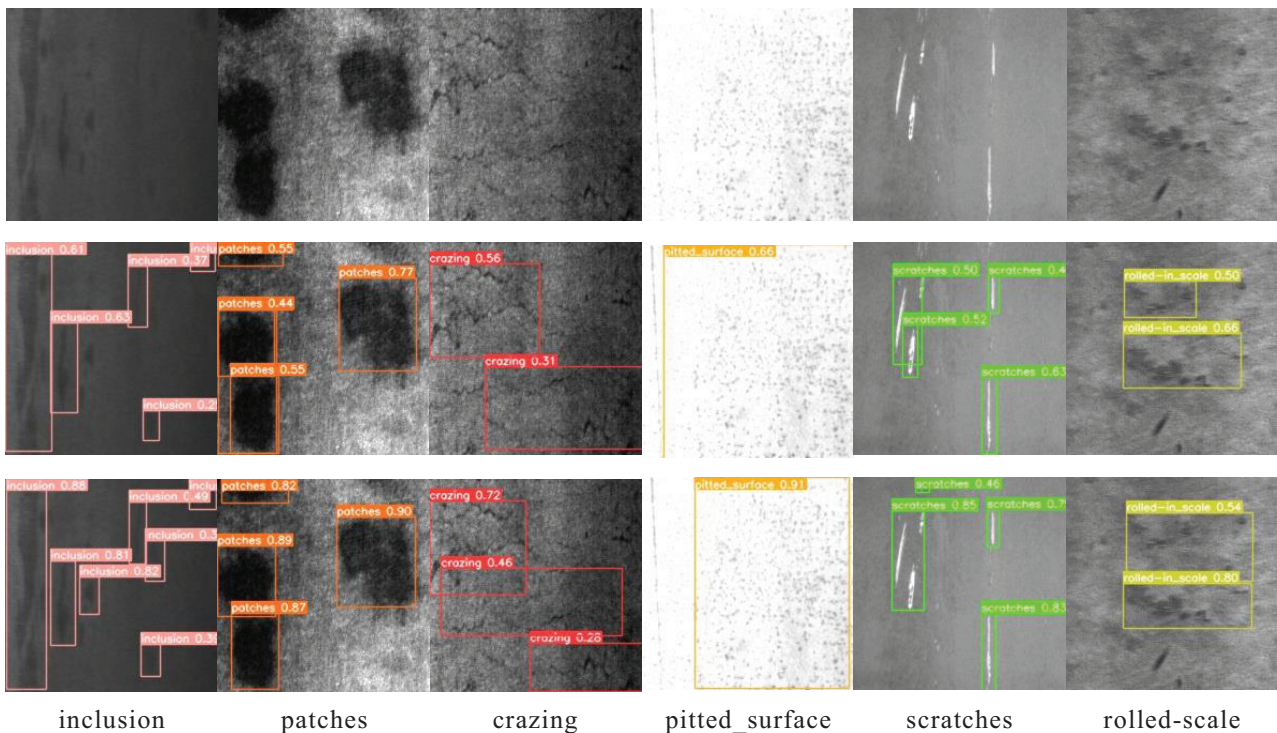


图8 6类缺陷效果对比

2.4.2 与主流目标检测算法性能对比

为了验证所提出YOLO-Steel钢材表面缺陷检测网络的性能,与主流目标检测模型YOLOv3、DDN、Faster R-CNN等进行横向对比,采用mAP和FPS两

项指标进行评估,实验结果如表6所示. 由表6分析可知,YOLO-Steel算法mAP达到了77.2%,比原始YOLOv5s算法提高1.8%,相比其他主流目标检测网络模型,能保持较快的检测速度和较好的检测性能.

表6 与主流模型的对比

network structure	FPS	mAP/%	AP/%					
			Cr	In	Pa	PS	RS	Sc
YOLOv3	12	69.1	44.7	60.8	84.4	74.5	61.1	87.2
DDN (ResNet34) ^[16]	< 20	74.8	48.0	75.9	87.4	78.3	68.4	90.8
DDN (ResNet50) ^[16]	< 10	82.3	62.4	84.7	90.3	89.7	76.3	90.1
Faster RCNN (ResNet34) ^[16]	< 20	70.2	46.7	61.3	82.8	76.5	70.7	83.4
YOLOv5s	119	75.4	37.6	84.2	92.1	83.7	61.8	93.1
YOLO-Steel	115	77.2	46.2	84.0	91.8	83.8	64.8	92.8

3 结论

本文针对钢材表面缺陷检测问题,提出了YOLO-Steel钢材表面缺陷检测网络.为了减少注意力模块的计算成本,提出了一种轻型注意力机制模块,该模块在性能几乎不变的同时有更小的参数量以及更小的计算代价.此外,针对钢材表面缺陷种类尺度变化多的问题,提出了金字塔多级注意力结构,利用多级池化金字塔在不同分辨率的特征图上获取其空间相互依赖性,以达到更好的检测效果.实验结果表明,相比于原始YOLOv5s算法,YOLO-Steel算法可在保持较快检测速度的情况下有效提高钢铁表面缺陷的检测精度,实现对钢铁表面缺陷的智能高效检测.下一步,针对缺陷检测样本量少的问题,尝试将对抗半监督学习与小样本学习结合,实现较少的样本量也能达到较好的检测效果.

参考文献(References)

- [1] 李少波, 杨静, 王铮, 等. 缺陷检测技术的发展与应用研究综述[J]. 自动化学报, 2020, 46(11): 2319-2336. (Li S B, Yang J, Wang Z, et al. Review of development and application of defect detection technology[J]. Acta Automatica Sinica, 2020, 46(11): 2319-2336.)
- [2] 康爽, 陈长征, 罗园庆, 等. 基于微分形态学梯度风力发电机叶片缺陷边缘增强的红外检测研究[J]. 太阳能学报, 2021, 42(6): 432-437. (Kang S, Chen C Z, Luo Y Q, et al. Study on infrared detection edge enhancement of wind turbine blade defects based on differential morphology gradient[J]. Acta Energetica Solaris Sinica, 2021, 42(6): 432-437.)
- [3] 许鹏, 耿明, 方舟, 等. 基于复合电磁的高速轨道缺陷检测方法研究[J]. 机械工程学报, 2021, 57(18): 57-65. (Xu P, Geng M, Fang Z, et al. Study on high-speed rail defect detection method based on combination of EC and MFL testing[J]. Journal of Mechanical Engineering, 2021, 57(18): 57-65.)
- [4] 朱玉龙, 赵迎松, 方阳, 等. 孔边裂纹的旋转涡流检测技术研究[J]. 中国机械工程, <https://kns.cnki.net/kcms/detail/42.1294.TH.20220322.1713.008.html>.
- [5] (Zhu Y L, Zhao Y S, Fang Y, et al. A study on rotating eddy current testing for inspection of cracks at hole edge[J]. China Mechanical Engineering, <https://kns.cnki.net/kcms/detail/42.1294.TH.20220322.1713.008.html>.)
- [5] Liu H W, Lan Y Y, Lee H W, et al. Steel surface in-line inspection using machine vision[J]. First International Workshop on Pattern Recognition, 2016, 10011: 187-191.
- [6] Liang Y, Xu K, Zhou P. Mask gradient response-based threshold segmentation for surface defect detection of milled aluminum ingot[J]. Sensors: Basel, Switzerland, 2020, 20(16): 4519.
- [7] 刘金海, 付明芮, 唐建华. 基于漏磁内检测的缺陷识别方法[J]. 仪器仪表学报, 2016, 37(11): 2572-2581. (Liu J H, Fu M R, Tang J H. MFL inner detection based defect recognition method[J]. Chinese Journal of Scientific Instrument, 2016, 37(11): 2572-2581.)
- [8] Tsai D M, Chen M C, Li W C, et al. A fast regularity measure for surface defect detection[J]. Machine Vision and Applications, 2012, 23(5): 869-886.
- [9] Timm F, Barth E. Non-parametric texture defect detection using Weibull features[C]. Proceedings of SPIE 7877, Image Processing: Machine Vision Applications IV. Piscataway: IEEE, 2011, 7877: 150-161.
- [10] 周恺, 张睿哲, 叶宽, 等. 基于同步压缩小波变换的接地扁钢缺陷电磁超声SH导波检测方法[J]. 清华大学学报: 自然科学版, 2022, 62(12): 2013-2020. (Zhou K, Zhang R Z, Ye K, et al. Electromagnetic ultrasonic SH guided wave detection method for grounded flat steel defects based on synchrosqueezed wavelet transforms[J]. Journal of Tsinghua University: Science and Technology, 2022, 62(12): 2013-2020.)
- [11] Yi L, Li G Y, Jiang M M. An end-to-end steel strip surface defects recognition system based on convolutional neural networks[J]. Steel Research International, 2017, 88(2):

- 1600068.
- [12] 王红梅, 王晓鹤, 王晓燕. 基于深度学习的复杂背景下目标检测[J]. 控制与决策, 2022, 37(12): 3115-3121.
(Wang H M, Wang X G, Wang X Y. Target detection under complex background based on deep learning[J]. Control and Decision, 2022, 37(12): 3115-3121.)
- [13] 张力, 黄丹平, 廖世鹏, 等. 基于目标检测网络的轮对踏面缺陷检测方法[J]. 激光与光电子学进展, 2021, 58(4): 244-253.
(Zhang L, Huang D P, Liao S P, et al. Wheelset tread defect detection method based on target detection network[J]. Laser & Optoelectronics Progress, 2021, 58(4): 244-253.)
- [14] 张雪荣, 向峰, 李红军, 等. 基于深度学习的钢卷端面缺陷检测系统设计[J]. 计算机集成制造系统, <https://kns.cnki.net/kcms/detail/11.5946.TP.20220306.1105.002.html>.
(Zhang X R, Xiang F, Li H J, et al. Design of rolled steel end defects detection system based on deep learning[J]. Computer Integrated Manufacturing Systems, <https://kns.cnki.net/kcms/detail/11.5946.TP.20220306.1105.002.html>.)
- [15] 刘培勇, 董洁, 谢罗峰, 等. 基于多支路卷积神经网络的磁瓦表面缺陷检测算法[J]. 吉林大学学报: 工学版, <https://doi.org/10.13229/j.cnki.jdxbgxb20210855>.
(Liu P Y, Dong J, Xie L F, et al. Surface defect detection algorithm of magnetic tiles based on multi-branch convolutional neural network[J]. Journal of Jilin University: Engineering and Technology Edition, <https://doi.org/10.13229/j.cnki.jdxbgxb20210855>.)
- [16] 郭龙源, 段厚裕, 周武威, 等. 基于Mask R-CNN的磁瓦表面缺陷检测算法[J]. 计算机集成制造系统, 2022, 28(5): 1393-1400.
(Guo L Y, Duan H Y, Zhou W W, et al. Surface defect detection algorithm of magnetic tile based on Mask R-CNN[J]. Computer Integrated Manufacturing Systems, 2022, 28(5): 1393-1400.)
- [17] Xing J J, Jia M P, Xu F Y, et al. A method for workpiece surface small-defect detection based on CutMix and YOLOv3[J]. Journal of Southeast University: English Edition, 2021(2): 128-136.
- [18] 刘艳菊, 王秋霖, 赵开峰, 等. 基于卷积神经网络的热轧钢条表面实时缺陷检测[J]. 仪器仪表学报, 2021, 42(12): 211-219.
(Liu Y J, Wang Q J, Zhao K F, et al. Real-time defect detection of hot rolling steel bar based on convolution neural network[J]. Chinese Journal of Scientific Instrument, 2021, 42(12): 211-219.)
- [19] Zhu X K, Lyu S C, Wang X, et al. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios[C]. IEEE/CVF International Conference on Computer Vision Workshops. Montreal, 2021: 2778-2788.
- [20] Nepal U, Eslamiat H. Comparing YOLOv3, YOLOv4 and YOLOv5 for autonomous landing spot detection in faulty UAVs[J]. Sensors, 2022, 22(2): 464.
- [21] Fan Y C, Zhang S Y, Feng K, et al. Strawberry maturity recognition algorithm combining dark channel enhancement and YOLOv5[J]. Sensors: Basel, Switzerland, 2022, 22(2): 419.
- [22] Walia I S, Kumar D, Sharma K, et al. An integrated approach for monitoring social distancing and face mask detection using stacked ResNet-50 and YOLOv5[J]. Electronics, 2021, 10(23): 2996.
- [23] Song Q S, Li S B, Bai Q, et al. Object detection method for grasping robot based on improved YOLOv5[J]. Micromachines, 2021, 12(11): 1273.
- [24] Farhadi A, Redmon J. Yolov3: An incremental improvement[C]. Computer Vision and Pattern Recognition. Berlin/Heidelberg: Springer, 2018: 1804-2767.
- [25] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: Optimal speed and accuracy of object detection[J/OL]. 2020, arXiv: 2004.10934.
- [26] Wang C Y, Mark Liao H Y, Wu Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle, 2020: 1571-1580.

作者简介

周彦(1978—), 男, 教授, 博士生导师, 从事视觉导航、多传感器信息融合等研究, E-mail: yanzhou@xtu.edu.cn;

孟江南(1994—), 男, 硕士生, 从事目标检测的研究, E-mail: mjnshizhu@163.com;

王冬丽(1980—), 女, 副教授, 博士, 从事模式识别、机器学习等研究, E-mail: wangdl@xtu.edu.cn;

谭雅琴(1999—), 女, 硕士生, 从事深度学习、生物信息学等研究, E-mail: 1970250317@qq.com.