

文章编号:1671-6833(2023)05-0040-06

# 优化聚类 and 引入 CBAM 的 YOLOv5 管制刀具检测

张震<sup>1</sup>, 陈可鑫<sup>2</sup>, 陈云飞<sup>1</sup>

(1. 郑州大学 电气工程学院, 河南 郑州 450001; 2. 郑州大学 计算机与人工智能学院, 河南 郑州 450001)

**摘要:** 针对传统管制刀具检测方式过度依赖特定设备与环境, 目标检测存在适用范围小、抗干扰能力弱、精度低等问题, 提出一种优化聚类 and 引入卷积注意力模块 (CBAM) 的 YOLOv5 管制刀具检测方法。首先, 通过计算每个点的局部密度和该点与其他具有较高密度点之间的最小距离, 绘制决策图并获取聚类中心; 其次, 计算局部类中所有点与其类中心间的平均距离来提取核心点, 并使用全局搜索分配策略将分类点归类; 最后, 采用统计学习策略分配剩余点, 未被处理的点当作噪声点, 归入到其最近的类中。采用改进密度峰值聚类算法对管制刀具的边界框聚类进行分析, 优化先验框尺寸, 提高先验框与目标物体尺寸的匹配度, 解决 YOLOv5 模型中 K-means 聚类算法聚类效果不稳定以及对大规模数据收敛较慢的问题。此外, 将 Backbone 中的 C3 模块与 CBAM 注意力机制相结合, 改进为 CBAMC3 模块, 提升模型对目标特征的提取能力, 解决 YOLOv5 算法对小目标检测效果不佳的问题, 提高模型精度。实验结果表明: 改进后模型 YOLOv5-Plus 在自定义数据集上的  $P$ 、 $R$ 、 $mAP@0.5$ 、 $mAP@0.5:0.95$  等参数的数值分为 98.14%、95.80%、97.56%、76.68%, 相比于改进前的 YOLOv5 模型分别提高了 1.64、1.59、1.51、3.26 个百分点, 有效提升了公共区域的管制刀具检测能力。

**关键词:** 管制刀具检测; 公共安全; 目标检测; 聚类算法; 注意力机制

**中图分类号:** TP391

**文献标志码:** A

**doi:**10.13705/j.issn.1671-6833.2022.05.015

近年来, 公共安全领域危险事件频发, 引发人们广泛关注。由于公共区域具有人流量大、危险因素多、环境复杂等特点, 公共区域安全难以得到全面保障。当前主流的检测方法主要分为毫米波检测<sup>[1-3]</sup>和目标检测。毫米波检测对静态物品的检测精度高、速度快, 但只能基于特定设备及环境。目标检测不依赖特定设备和环境, 但对小目标的检测效果较差, 增加模型深度又会影响检测时间, 不能满足公共安全领域的实时性要求。人工进行管制刀具检测效率低、成本高、危险度高, 利用监控系统对管制刀具检测可有效克服以上缺点, 因此具有重要的研究意义。

当前主流的目标检测算法主要分为双阶段算法和单阶段算法, 双阶段算法的典型代表有 Faster R-CNN<sup>[4]</sup>、Mask R-CNN<sup>[5]</sup> 和 SPPNet<sup>[6]</sup> 等。Fernandez-carroles 等<sup>[7]</sup> 提出基于 Faster R-CNN 算法的枪支刀具检测系统, 使用 GoogleNet 架构对算法进行改进, 获得刀具检测的最佳结果,  $AP_{50}$  达到 46.68%。虽

然双阶段算法较单阶段算法检测精度高, 但是由于其网络结构是双阶段的, 需要事先提取多个候选区域对应的图像, 检测速度不能满足公共安全领域实时性要求。单阶段算法仅需处理一次图像就能得到目标的位置和分类结果, 避免了双阶段算法中将检测任务分为 2 个阶段进行的复杂操作, 明显提升了检测速度。单阶段算法的典型代表有 YOLOv4<sup>[8]</sup>、YOLOv5<sup>[9]</sup>、SSD<sup>[10-11]</sup> 等。YOLOv5 算法借鉴了 CutMix<sup>[12]</sup> 方法, 在输入部分采用 Mosaic 数据增强的方式对数据进行优化, 并结合自适应锚框计算、自适应图片缩放等技术, 在快速处理图片的同时, 又保证了识别的精度。但是, YOLOv5 算法采用的 K-means 聚类算法鲁棒性较差, 性能依赖于初始类中心的选择, 并需指定类中心个数<sup>[13]</sup>。此外, 管制刀具目标相对较小, 而 YOLOv5 无法有效地识别出小目标。

本实验采用改进密度峰值聚类算法优化先验框尺寸, 增大先验框与实际框的匹配度, 提升目标检测

收稿日期: 2022-03-12; 修订日期: 2022-06-10

基金项目: 国家重点研发计划 (2018YFC0824 \*\*\*)

作者简介: 张震 (1966—), 男, 河南郑州人, 郑州大学教授, 博士, 博士生导师, 主要从事多媒体信息安全、图像处理、模式识别等研究, E-mail: zhangzhen@zzu.edu.cn。

引用本文: 张震, 陈可鑫, 陈云飞. 优化聚类 and 引入 CBAM 的 YOLOv5 管制刀具检测 [J]. 郑州大学学报 (工学版) 2023, 44 (5): 40-45, 61. (ZHANG Z, CHEN K X, CHEN Y F. YOLOv5 with optimized clustering and CBAM for controlled knife detection [J]. Journal of Zhengzhou University (Engineering Science), 2023, 44 (5): 40-45, 61.)

的精度,并将骨干网络 Backbone 中的 C3 模块与 CBAM(convolutional black attention module)结合,改进为 CBAMC3 模块,提升模型捕捉目标物体特征的能力,改善模型对小目标的识别性能。同时,创建自定义管制刀具数据集,并在该数据集上验证了改进后算法的有效性。

## 1 YOLOv5 算法介绍

YOLOv5 网络结构分为输入端 Input、骨干网络 Backbone、颈部 Neck 和预测端 Prediction 等部分。算法包含 4 个模型,模型尺寸由小到大分别为 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x。YOLOv5 算法在 COCO 数据集上的训练结果显示,YOLOv5s 模型能在检测精度较高的前提下保证检测速度,可以满足公共安全领域实时性检测的要求,因此选择基于 YOLOv5s 的模型进行优化。

YOLOv5 算法在输入端 Input 采用 Mosaic 数据增强对数据集进行优化。每次读取数据集中的 4 张原始图片,并对图片进行随机缩放、翻转、改变色域等操作,随后将变化后的图片按照 4 个方向分别拼接。

YOLOv5 算法的骨干网络 Backbone 采用 Focus+CSP 结构。在图片进入 Backbone 之前,需要经过 Focus 结构对图片进行切片,切片操作如图 1 所示,尺寸为  $4 \times 4 \times 3$  的图像切片后变成  $2 \times 2 \times 12$  的特征图。YOLOv5s 模型中原始图尺寸为  $608 \times 608 \times 3$ ,经过切片操作变成  $304 \times 304 \times 12$ ,通过一个 32 个卷积核的卷积后特征图的尺寸变为  $304 \times 304 \times 32$ 。在 CSP 模块中加入 SPP 模块后,多尺度最大池化层大幅度提升了感受野,降低了直接伸缩图片导致信息丢失的可能性,提升了模型精度。YOLOv5 的网络结构借鉴 CSP-Net<sup>[14]</sup>的设计思路,在网络中加入 CSP 结构,骨干网络 Backbone 中使用的是 CSP1\_X 结构,Neck 中使用的是 CSP2\_X 结构。

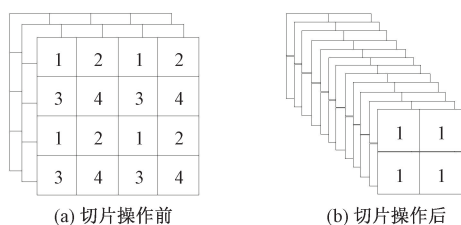


图 1 切片操作

Figure 1 Slicing operation

## 2 算法改进

算法改进主要体现在 3 个方面:采用改进密度

峰值聚类算法优化目标框;引入 CBAM 注意力机制;创建自定义管制刀具数据集。

### 2.1 采用改进密度峰值聚类算法优化目标框

Rodriguez 等<sup>[15]</sup>提出了一种基于快速搜索和发现密度峰值的聚类算法(clustering by fast search and find of density peaks, DPC),可以有效解决 K-means 聚类算法的问题。薛小娜等<sup>[16]</sup>针对 DPC 算法中存在的缺点提出了 IDPCA 聚类算法,当数据集密度差异较大时,DPC 算法中局部密度  $\rho$  和该点与其他具有较高密度点之间的最小距离  $\delta$  不能应对复杂的密度差异,而且 DPC 算法对剩余点的分配策略存在多米诺效应,会导致误差传播。本实验采用 IDPCA 密度峰值聚类算法对目标框聚类。IDPCA 聚类按照以下步骤选取聚类中心。

(1)计算每个点  $x_i$  的局部密度  $\rho_i$  和该点与其他具有较高密度点之间的最小距离  $\delta_i$ ,计算公式<sup>[15]</sup>为

$$\rho_i = \sum_j e^{-\frac{1}{\sigma} \left( \frac{d_{ij}}{\sigma} \right)^2}; \quad (1)$$

$$\delta_i = \min_{j: \rho_j > \rho_i} d_{ij}. \quad (2)$$

式中: $\sigma$  为数据量的 2%; $d_{ij}$  为点  $x_i$  和点  $x_j$  的距离; $r$  为相似性系数,该值越大,距离点  $x_i$  越近的点对该点的  $\rho_i$  权重越大。

(2)通过  $\gamma$  决策图获取  $m$  个聚类中心,其中  $\gamma = \rho\delta$ 。

(3)提取核心点,并采用全局搜索分配策略将分类点归类,提取核心点的方式为计算局部类  $C_m$  中所有点与其类中心  $cen_m$  间的平均距离  $u_m$ 。 $u_m$  的计算公式如式(3)所示,若  $x_i \in C_m$  在  $cen_m$  的  $\varepsilon u_m$  邻域内,则  $x_i$  为核心点,分离阈值  $\varepsilon = N\%$ , $N$  为数据集数量。

$$u_m = \frac{1}{|C_m|} \sum_{i=1}^{|C_m|} d_{i_m}^{cen_m}. \quad (3)$$

(4)采用统计学习策略分配剩余点。

(5)最后未被处理的点可看作噪声点,归入到其最近邻类中。

IDPCA 聚类算法对已标注管制刀具的边界框进行聚类分析,生成多个不同大小的先验框尺寸,使得先验框与实际框的匹配度更高,从而提升目标检测的精度。实验采用的基于 IDPCA 聚类算法的先验框尺寸如表 1 所示,基于 K-means 聚类算法的先验框尺寸如表 2 所示。

### 2.2 引入 CBAM 机制

CBAM<sup>[17]</sup>是由 Woo 等提出的一种分别从通道和空间维度顺序提供注意力图的注意力机制模块,

主要分为通道注意力模块和空间注意力模块,可以使模型提取到的特征更加精炼,有效提升模型的分类效果。通道注意力模块主要关注输入图像中有意义的信息,可以在通道维度不变的情况下压缩空间维度;空间注意力模块关注的是目标的位置信息,可以在空间维度不变的情况下压缩通道维度。CBAM 结构如图 2 所示。

表 1 基于 IDPCA 聚类算法的先验框尺寸

Table 1 Prior box size based on IDPCA clustering algorithm	
目标种类	先验框尺寸
小目标	55, 23, 33, 46, 80, 36
中目标	62, 70, 147, 59, 137, 91
大目标	92, 170, 250, 103, 167, 165

表 2 基于 K-means 聚类算法的先验框尺寸

Table 2 Prior box size based on K-means clustering algorithm	
目标种类	先验框尺寸
小目标	54, 23, 32, 42, 78, 35
中目标	58, 69, 152, 60, 135, 90
大目标	80, 202, 158, 166, 175, 100

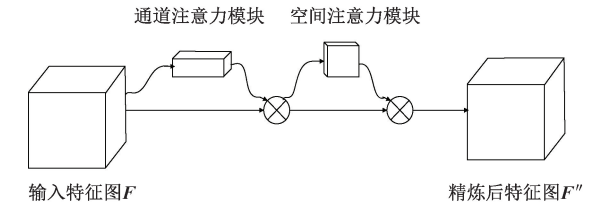


图 2 CBAM 结构  
Figure 2 CBAM structure

通道注意力模块的结构如图 3 所示,在输入端输入特征图  $F(F \in \mathbf{R}^{C \times H \times W})$ ,经过平均池化和最大池化后,将尺寸为  $C \times H \times W$  的特征图变成  $C \times 1 \times 1$  大小,然后再将它们送入神经网络 MLP,第 1 层神经元个数为  $C/r$ , $r$  为下降率,激活函数为 Relu,第 2 层神经元个数为  $C$ ,完成后将结果相加,再经过一个 Sigmoid 函数得到权重系数  $M_c$ 。权重系数的计算方式如式(4)所示,再乘以最初的输入,即可得到缩放后的新特征。

$$M_c(F) = \sigma_1(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c)))。 \quad (4)$$

式中: $\sigma_1$  表示 Sigmoid 函数; $W_0 \in \mathbf{R}^{C \times \frac{C}{r}}$ ;  $W_1 \in \mathbf{R}^{C \times \frac{C}{r}}$ ;  $F_{avg}^c$  表示尺寸为  $1 \times 1 \times C$  的平均池化特征; $F_{max}^c$  表示尺寸为  $1 \times 1 \times C$  的最大池化特征。

空间注意力模块结构如图 4 所示,将上一步得到的结果通过最大值池化和平均池化分成 2 个大小为  $1 \times H \times W$  的通道描述,然后通过连接操作将张量

堆叠在一起,再通过卷积操作和一个 Sigmoid 得到权重系数  $M_s$ 。权重系数的计算方式如式(5)所示,最后用权重系数乘以上一步的输入即可得到缩放后的新特征,完成空间注意力操作。

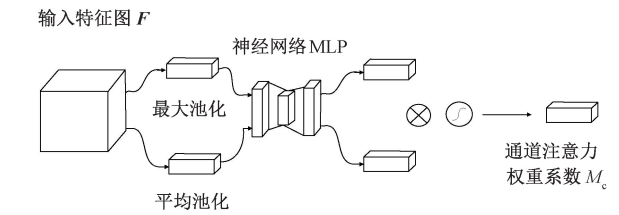


图 3 通道注意力模块结构

Figure 3 Channel attention module structure

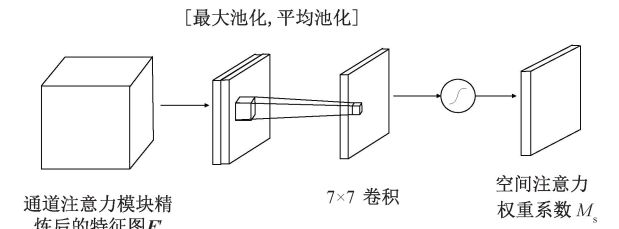


图 4 空间注意力模块结构

Figure 4 Spatial attention module structure

$$M_s(F) = \sigma_1(f^{7 \times 7}([s_{avg}^s; F_{max}^s]))。 \quad (5)$$

式中: $f^{7 \times 7}$  表示一个  $7 \times 7$  的卷积; $\sigma_1$  表示 Sigmoid 函数; $F_{avg}^s$  表示平均池化特征,大小为  $1 \times H \times W$ ;  $F_{max}^s$  表示最大池化特征,大小也为  $1 \times H \times W$ 。

CBAM 注意力机制主要增加在骨干网络 Backbone 中,新版 YOLOv5s 模型中 BottleneckCSP 模块转变为了 C3 模块,因此选择将 Backbone 中的 C3 模块与注意力机制结合,改进为 CBAMC3 模块,改进后的网络结构如图 5 所示。

2.3 创建自定义管制刀具数据集

当前并没有公开的管制刀具数据集,因此选择自定义数据集。素材来源为物流集散中心真实录像、自做管制刀具图片和互联网视频等。数据集中包含水果刀、菜刀、砍刀和斧头等共计 8 060 张各种类型的刀具图片。使用 LabelImg 对筛选后的数据打标签,标签格式为 txt,其中训练集占 60%,验证集占 20%,测试集占 20%。

3 实验

3.1 实验环境与评价指标

实验机器操作系统版本为 Ubuntu18.04, GPU 为 GeForce RTX 3090,显存 24 GB,内存 32 GB,训练框架基于 Pytorch1.8.0,使用 Cuda11.1.1 对 GPU 进行加速。

模型评价指标主要采用精确率  $P$ 、召回率  $R$  和

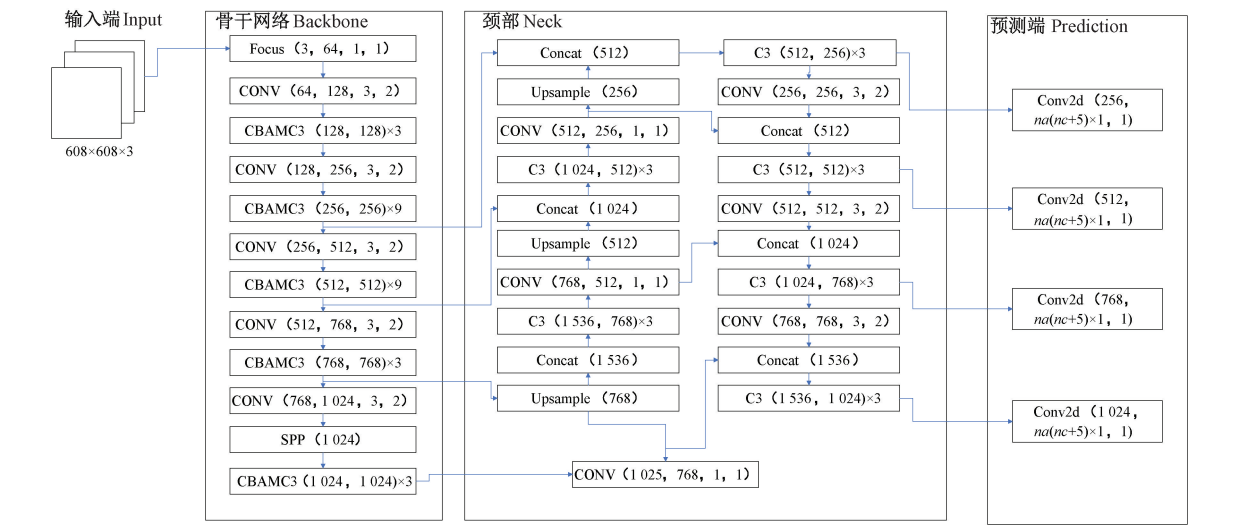


图 5 改进后网络结构

Figure 5 Improved network structure

平均精度均值  $mAP$  等。精确率  $P$  表示在检测出为刀斧的样本中有多少是真正的刀斧,反映了检测结果是否准确;召回率  $R$  表示在全部图片样本中,有多少个刀斧被正确检测出,反映了刀斧数目是否齐全;精度均值  $AP$  大小等于以精确率  $P$  和召回率  $R$  所围成的曲线面积值,平均精度均值  $mAP$  即为所有类别精度均值的平均值。

3.2 实验结果分析

本文实验的训练轮次为 150,批尺寸为 32,采用 Warmup 方法预热学习率。在 Warmup 阶段采用一维线性插值更新每次迭代的学习率,在 Warmup 阶段以后采用余弦退火算法来对学习率进行更新。实验中  $L_0$ 、 $L_1$ 、 $L_2$  分别表示 BN 层、weight 层和 bais 层的学习率,学习率的变化曲线如图 6 所示, $L_0$  和  $L_1$  的曲线相同。

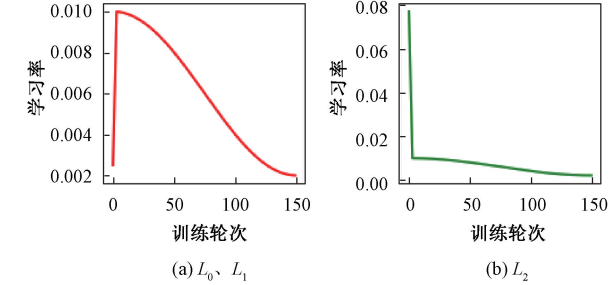


图 6 学习率曲线图

Figure 6 Graph of learning rate

通过消融实验来逐步验证网络结构改变所引起的性能变化,消融实验分为 YOLOv5 + CBAM、YOLOv5 + IDPCA、YOLOv5、YOLOv5-Plus 等测试部分,评价指标  $P$ 、 $R$ 、 $mAP@0.5$ 、 $mAP@0.5:0.95$  指标对比图如图 7~10 所示。

由表 3 的消融实验数据可以看出,在训练自定

义数据集时原始 YOLOv5 模型评价指标  $P$ 、 $R$ 、 $mAP@0.5$ 、 $mAP@0.5:0.95$  等评价指标的数值分别为 96.50%、94.21%、96.05%、73.42%。采用 IDPCA 密度峰值聚类算法优化先验框,使得先验框与实际框的匹配度更高,明显提升检测精度,评价指标的数值分别提升 0.41、1.14、0.60、1.67 个百分点。在模型

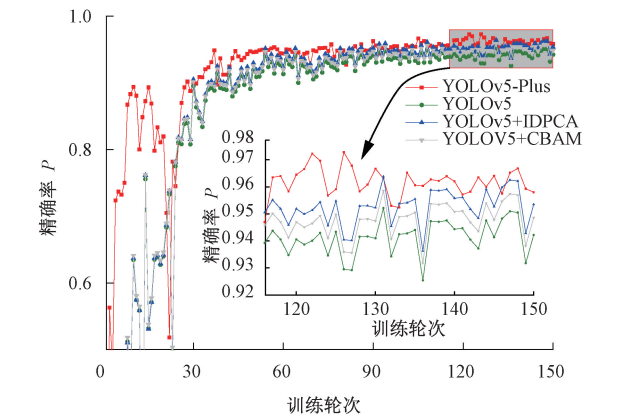


图 7 精确率对比图

Figure 7 Comparison chart of precision

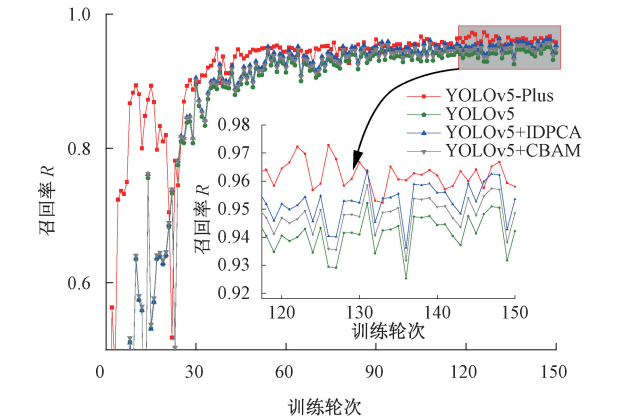


图 8 召回率对比图

Figure 8 Comparison chart of recall



骨干网络加入 CBAM 后,显著提升模型对于对比强调特征等高级特征的提取能力,评价指标的数值分别提升 0.20、0.65、0.45、1.49 百分点。YOLOv5-Plus 算法相比于 YOLOv5 算法,在先验框与实际框的匹配度以及对特征的提取能力均有显著提升,评价指标的数值分别提升 1.64、1.59、1.51、3.26 百分点。

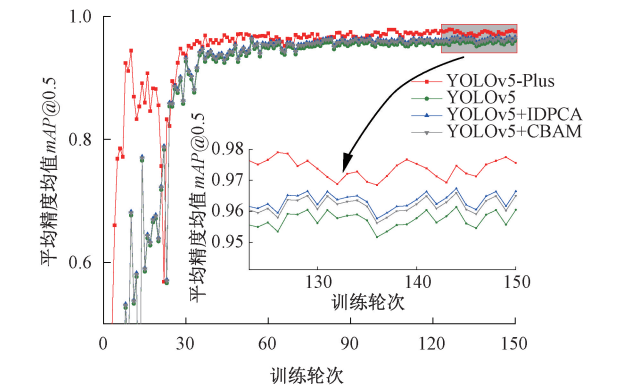


图 9  $mAP@0.5$  对比图

Figure 9 Comparison chart of  $mAP@0.5$

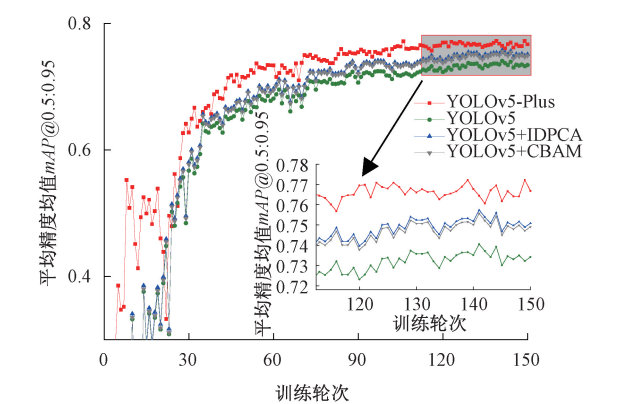


图 10  $mAP@0.5:0.95$  对比图

Figure 10 Comparison chart of  $mAP@0.5:0.95$

表 3 实验结果分析

Table 3 Analysis of experimental results %				
模型	精确率 $P$	召回率 $R$	$mAP@0.5$	$mAP@0.5:0.95$
YOLOv5	96.50	94.21	96.05	73.42
YOLOv5+IDPCA	96.91	95.35	96.65	75.09
YOLOv5+CBAM	96.70	94.86	96.50	74.91
YOLOv5-Plus	98.14	95.80	97.56	76.68

YOLOv5-plus 与 YOLOv5 检测精度对比如图 11 所示,改进前后检测精度分别为 90%和 95%。实验证明改进后的模型对管制刀具等小目标的检测精度有了较大的提升,在一些环境复杂、目标物体不易发现的场景下提升效果尤为明显,基本满足公共区域对管制刀具检测的速度与精度要求。



图 11 检测精度对比

Figure 11 Detection accuracy comparison

## 4 结论

针对当前公共区域缺乏较为成熟的管制刀具实时检测的算法,传统管制刀具检测方法过度依赖检测设备和环境,且检测精度与检测速度不能满足实时性要求的问题,本文提出一种优化聚类并引入 CBAM 的 YOLOv5 管制刀具检测算法。采用改进密度峰值聚类算法优化先验框尺寸,使得先验框与实际框尺寸的匹配度更高,从而提升目标检测的精度。此外,将骨干网络 Backbone 中的 C3 模块与 CBAM 结合,提升模型对目标特征的捕捉能力,解决 YOLOv5 算法对小目标检测效果不佳的问题。同时,创建自定义管制刀具数据集,并在该数据集上验证了改进后算法的有效性。改进后的算法相比于改进前算法的  $P$ 、 $R$ 、 $mAP@0.5$ 、 $mAP@0.5:0.95$  等评价指标分别提升 1.64、1.59、1.51、3.26 百分点,基本可以满足实用要求。

## 参考文献:

[1] 徐晨玉. 基于 YOLOv3 的危险物品检测模型研究与改进[D]. 兰州: 兰州大学, 2019.  
XU C Y. Research and improvement of a dangerous goods detection model based on YOLOv3 [D]. Lanzhou: Lanzhou University, 2019.

[2] 杨希. 基于深度学习的地铁安检危险物品检测[D]. 武汉: 武汉理工大学, 2020.  
YANG X. Detection of dangerous goods in subway security check based on deep learning[D]. Wuhan: Wuhan University of Technology, 2020.

[3] 王胜. 毫米波图像中危险物品检测[D]. 北京: 清华大学, 2018.

- WANG S. Dangerous objects detection in millimeter wave images[D]. Beijing: Tsinghua University, 2018.
- [4] REN S Q, HE K M, 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.
- [5] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]//2017 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 2980–2988.
- [6] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904–1916.
- [7] FERNANDEZ-CARROBLES M, DENIZ O, MAROTO F. Gun and knife detection based on faster R-CNN for video surveillance[C]//Iberian Conference on Pattern Recognition and Image Analysis. Cham: Springer International Publishing, 2019: 441–452.
- [8] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: optimal speed and accuracy of object detection [EB/OL]. (2020-08-23) [2021-09-13]. <https://doi.org/10.48550/arXiv.2004.10934>.
- [9] Ultralytics. YOLOv5 [EB/OL]. (2020-05-18) [2021-08-12]. <https://github.com/ultralytics/yolov5>.
- [10] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: single shot multibox detector [C] // European Conference on Computer Vision. Berlin: Springer Press, 2016: 21–37.
- [11] 张伟彬, 吴军, 易见兵. 基于 RFB 网络的特征融合管制物品检测算法研究[J]. 广西师范大学学报(自然科学版), 2021, 39(4): 34–46.
- ZHANG W B, WU J, YI J B. Research on feature fusion controlled items detection algorithm based on RFB network[J]. Journal of Guangxi Normal University (Natural Science Edition), 2021, 39(4): 34–46.
- [12] YUN S, HAN D, CHUN S, et al. CutMix: regularization strategy to train strong classifiers with localizable features [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2019: 6022–6031.
- [13] 张震, 李浩方, 李孟洲, 等. 改进 YOLOv3 算法与人体信息数据融合的视频监控检测方法[J]. 郑州大学学报(工学版), 2021, 42(1): 28–34.
- ZHANG Z, LI H F, LI M Z, et al. Video surveillance detection method based on improved YOLOv3 algorithm and human body information data fusion[J]. Journal of Zhengzhou University (Engineering Science), 2021, 42(1): 28–34.
- [14] WANG C Y, MARK LIAO H Y, WU Y H, et al. CSP-Net: a new backbone that can enhance learning capability of CNN [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE, 2020: 1571–1580.
- [15] RODRIGUEZ A, LAIO A. Clustering by fast search and find of density peaks[J]. Science, 2014, 344(6191): 1492–1496.
- [16] 薛小娜, 高淑萍, 彭弘铭, 等. 结合 k 近邻的改进密度峰值聚类算法[J]. 计算机工程与应用, 2018, 54(7): 36–43.
- XUE X N, GAO S P, PENG H M, et al. Improved density peaks clustering algorithm combining k-nearest neighbors[J]. Computer Engineering and Applications, 2018, 54(7): 36–43.
- [17] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module [C]//Proceedings of the European Conference on Computer Vision (ECCV). Berlin: Springer Press, 2018: 3–19.

## YOLOv5 with Optimized Clustering and CBAM for Controlled Knife Detection

ZHANG Zhen<sup>1</sup>, CHEN Kexin<sup>2</sup>, CHEN Yunfei<sup>1</sup>

(1. School of Electrical Engineering, Zhengzhou University, Zhengzhou 450001, China; 2. School of Computer and Artificial Intelligence, Zhengzhou University, Zhengzhou 450001, China)

**Abstract:** Aiming at the problems that traditional controlled knife detection methods relied too much on specific equipment and environment, and the target detection was applicable to small scope of application, with weak anti-interference ability and low precision, a YOLOv5 model with optimized clustering and CBAM was proposed for controlled knife detection was proposed. Firstly, a decision diagram could be drew and cluster centers could be obtained by calculating the local density of each point and the minimum distance between it and others with higher density. Then the core points could be extracted by calculating all points in the local class the average distance with its class center, the global search assignment strategy was used to classify the test points. Finally, the statistical

(下转第 61 页)

Reinforcement Learning Autonomous Driving Trajectory  
Prediction Based on Directed Graph

CUI Jianming<sup>1</sup>, LIN Fanrong<sup>1</sup>, ZHANG Di<sup>1</sup>, ZHANG Luning<sup>1</sup>, LIU Ming<sup>2</sup>

( 1. School of Information Engineering, Chang'an University, Xi'an 710018, China; 2. National Computer Network Emergency Response Technical Team/Coordination Center of China, Beijing 100029, China )

**Abstract:** As an important part of autonomous driving, trajectory prediction aimed to forecast the vehicle's driving path, so that the vehicle could make path planning according to the driving estimation, so as to make safe and accurate decisions. Firstly, in order to improve the accuracy of vehicle trajectory prediction, the directed graph method was used to construct a high-definition driving scene map, and the directed graph method vectorized the map information to effectively extract the map topology. Secondly, GAIL was used to learn the driving strategy of the dataset through the confrontation game between the generator and the discriminator, so as to adopt the corresponding driving behavior according to the current state. Finally, the multimodal prediction trajectory scheme was obtained by sampling traversal. Simulation was carried out on the nuScenes motion prediction dataset. The quantitative results showed that compared with other methods, when  $K=5$ , the minimum final displacement error  $MinFDE_5$  was increased by 10.8%; when  $K=10$ , the minimum fianl displacement error  $MinFDE_{10}$  increased by 17.53%, the minimum average displacement error  $MinADE_{10}$  increased by 9.52%, and the error rate  $MissRate_{10}$  decreased by 28.26%. The evaluation showed that the generated trajectories were multimodal, could conform to the basic structure of the scene, with improved accuracy.

**Keywords:** autonomous driving; trajectory prediction; directed graph; reinforcement learning; GAIL; attention mechanism; multimodal prediction

( 上接第 45 页 )

learning strategy was employed to allocate the remaining points, these unprocessed points were used as noise points, and they were classified into the class of its nearest neighbor. The improved density peak clustering algorithm was used to analyze the bounding box of the controlled knife, optimize the size of the priori box, and improve the matching degree between the priori box and the size of target object, so as to solve the problem that the clustering effect of the  $K$ -means clustering algorithm in the YOLOv5 model such as its unstability and the low convergence rate of large-scale data. Moreover, The C3 module in the Backbone network was combined with the CBAM attention mechanism, named CBAMC3 module, which could improve the model's ability to extract target features, solve the problem that the YOLOv5 algorithm was not effective for small target detection, and improve the model accuracy. The experimental results showed that the values of  $P$ ,  $R$ ,  $mAP@0.5$ ,  $mAP@0.5:0.95$  of the improved model YOLOv5-Plus on the customed data set were 98.14%, 95.80%, 97.56%, and 76.68%, respectived, which is 1.64%, 1.59%, 1.51%, and 3.26% higher than that of YOLOv5 before improvement, and also verified that the proposed model could effectively improve the detection performance of controlled knife in public areas.

**Keywords:** controlled knife detection; public safety; target detection; clustering algorithm; attention mechanism