

冷库中基于YOLOv5的识别同伴机器人方法

杜学智¹, 马西良^{1,2}, 马彬¹, 苗文军¹, 朱睿哲¹

¹盐城工学院机械工程学院优集学院, 江苏 盐城

²徐州工程学院机电工程学院, 江苏 徐州

收稿日期: 2023年9月20日; 录用日期: 2023年10月27日; 发布日期: 2023年11月2日

摘要

针对现有目标识别算法因冷库低光照、货物遮挡等因素造成识别丢失的问题, 提出一种基于改进的YOLOv5目标模型。首先, 使用AutomatedMSRCR算法进行对采集到的光照不足图片图像预处理, 提高图像质量。其次, 采用kmeans++算法结合双线性插值法, 重新设计锚框尺寸, 为被遮挡区域填充合理的像素值。最后, 为解决同伴机器人运动时尺度变化导致的识别差, 融合CoordConv卷积提高网络对空间信息和尺寸的敏感度, 提升模型的识别准确率。经实验表明, 相比于现有YOLOv5模型, 本文改进YOLOv5模型的mAP0.5、F1、精确率、召回率、分别提升4.5%、3.1%、3.5%、2.7%, 满足实际应用场景中协作机器人对同伴准确识别的要求。

关键词

目标识别, YOLOv5, 图像增强

A Method for Identifying Companion Robots in Cold Storage Based on YOLOv5

Xuezhi Du¹, Xiliang Ma^{1,2}, Bin Ma¹, Wenjun Miao¹, Ruizhe Zhu¹

¹School of Mechanical Engineering UGS College, Yancheng Institute of Technology, Yancheng Jiangsu

²School of Mechanical & Electrical Engineering, Xuzhou University of Technology, Xuzhou Jiangsu

Received: Sep. 20th, 2023; accepted: Oct. 27th, 2023; published: Nov. 2nd, 2023

Abstract

A modified YOLOv5 object model is proposed to address the issue of recognition loss caused by factors such as low lighting in cold storage and obstruction of goods in existing object recognition algorithms. Firstly, use the AutomatedMSRCR algorithm to preprocess the collected images with insufficient lighting to improve image quality. Secondly, using the kmeans++ algorithm combined

with bilinear interpolation, the anchor box size is redesigned to fill the occluded area with reasonable pixel values. Finally, to address the recognition errors caused by scale changes during the movement of companion robots, CoordConv convolution is integrated to improve the network's sensitivity to spatial information and size, and to improve the recognition accuracy of the model. Through experiments, it has been shown that compared to the existing YOLOv5 model, this improved YOLOv5 model improves mAP0.5, F1, accuracy, and recall by 4.5%, 3.1%, 3.5%, and 2.7%, respectively, meeting the requirements of collaborative robots for accurate peer recognition in practical application scenarios.

Keywords

Object Detection, YOLOv5, Image Enhancement

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

高效的除霜方式是减少冷库能源浪费的有效措施之一，随着冷库产业的不断发展，因冷库结霜造成的能源浪费也日趋增多，霜层的及时处理迫在眉睫。为解决人工除霜时造成的库温波动，二次结霜等问题，使用机器人进行除霜，以提高效率和减少成本。冷库除霜机器人采用双机协作方式进行除霜，实现管上除霜与协同收霜工作，地面协作机器人通过视觉对同伴进行识别、定位、跟随，同伴机器人清除的霜层通过连接管道传递给协作机器人，协作机器人完成对霜的收集与处理。但在冷库中光线不足导致收集的图片质量差，货物堆积等问题导致识别的模型精准率不高，无法及时识别同伴机器人，造成运动不同步，使得两者间的落霜传递管道脱落，造成地面落霜结冰，对工作人员生命安全构成威胁，霜层堆积引发细菌滋生，影响货物的品质，所以协作机器人对于同伴的高识别率至关重要。

国内外针对低光照收集的质量不佳图片的目标识别进行了广泛研究。例如李雷孝等提出一种利用直方图均衡化高斯滤波等方法对图像增强，通过 YOLOv3 和 FasterRCNN 对夜间安全带进行检测[1]，冯娟等使用 MSRCR (multi-scale Retinex with color restoration)算法对图像增强，使用 YOLOv4 算法对光线不足的水下单环刺螈洞口识别[2]。司永胜等采用 MSRCP (Multi-scaleretinexwithchromaticitypreservation)算法对图像增强后与 YOLOv4 算法结合对夜间奶牛进行检测[3]。但目前并未有冷库环境的相关物体识别研究。

故本文结合现有深度学习研究，针对协作机器人在冷库内因光照不足、货物遮挡和快速运动时空间信息丢失造成的对同伴识别率低问题，选用权重文件小，训练快的 YOLOv5 目标检测方法[4]。在图像预处理时采用具有色彩增益加权的 AutomatedMSRCR 算法[5]。对低光照下的低质量图像进行图像增强；使用 kmeans++算法对数据集重新聚类，提供更好的识别效果[6]；在 YOLOv5 算法中使用双线性插值法 (Bi-Linear interpolation)提供更精细的图像细节[7]。融入 CoordConv 卷积增加模型的位置感知信息更好的捕捉物体的几何形状特征，提升模型的识别精度[8]。

2. 图像获取与预处理

2.1. 图像获取

本文从徐州雨润冷库、七里沟等数十家冷库产品供应商企业与搭建的冷库排管模拟实验室中收集了

冷库除霜机器人照片，获取同伴机器人的在低温冷库内的不同运动姿态，在顶管与侧管等不同位置下与不同空间环境下的机器人图片，采集到的部分数据集图片如图 1 所示，共计收集 1800 张图片数据集，其中 1620 张作为训练集，180 张作为验证集。使用 YOLOv5 算法训练前对图像使用 labeling 对图像进行人工标注，生成含有类别与位置信息的 txt 文件，便于网络的训练与测试。



Figure 1. Image acquisition

图 1. 图像获取

2.2. 图像预处理

冷库中的图像由于光照强度不足，拍摄的图像对比度低、清晰度低、质量退化，需要检测的物体轮廓和细节信息难以清晰表达对物体的识别增加难度。通过使用 AutomatedMSRCR 算法对图像进行色彩增益加权，该方法通过在 R, G, B 三个维度的上下限进行限制，可以去除较大较小的像素值，防止出现两极化使图像的亮度与对比度更加均衡，图像细节更加清晰，上述量化公式如式(1)所示。

AutomatedMSRCR 算法的流程为首先将输入图像分解为不同尺度的图像，估计出每个尺度图像的亮度图像和反射图像，然后对图像进行局部对比度增强，更好的恢复图像的局部细节信息，最后合并所有尺度空间的图像重建出增强后的图像。通过算法可加强同伴机器的物体轮廓与细信息等关键特征点，找回更多的潜在信息。经 AutomatedMSRCR 算法得到增强后的效果图如图 2 所示，可看出通过算法增强后同伴机器人的特征更加明显，腿部和躯体的轮廓更加清晰细节信息更加充分，图像整体质量得到提升，增强后的图像更加自然。

(x, y)-二维图像对应的位置；

max, min-限定范围，

$$R_{\text{MSRCR}_i}(X, Y) = 255 \frac{R_{\text{MSRCR}_i}(X, Y) - \min_i(\min_i R_{\text{MSRCR}_i}(X, Y))}{\max_i(\max_i R_{\text{MSRCR}_i}(X, Y)) - \min_i(\min_i R_{\text{MSRCR}_i}(X, Y))} \quad (1)$$

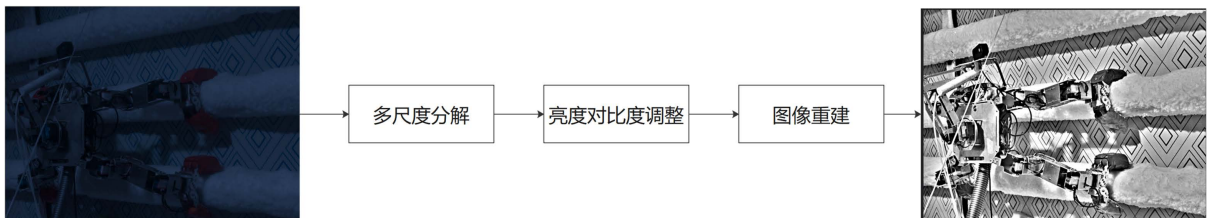


Figure 2. Automated MSRCR image enhancement process

图 2. Automated MSRCR 图像增强过程

3. 改进的同伴机器人识别模型

3.1. 改进的模型结构

YOLOv5 模型的预设锚框是基于 COCO [9]数据集获取，与识别的同伴机器人在尺寸上差距较大，导致识别的精度降低，使用 kmeans++算法重新设计新的锚框尺寸提高模型检测精度；同时为了保留图像的细节和纹理信息，更好地适应物体遮挡的情况，使用双线性插值替换 YOLOv5 模型的最近邻插值法；然后使用含有坐标信息的 CoordConv 卷积替换特征金字塔 FPN (Feature Pyramid Networks)结构的传统卷积层和在 Detect 模块中添加 CoordConv 卷积缓解同伴机器人在运动时造成的空间信息丢失问题得到适用于冷库内的同伴机器人的识别网络。改进后的算法模型结构图如图 3 所示。

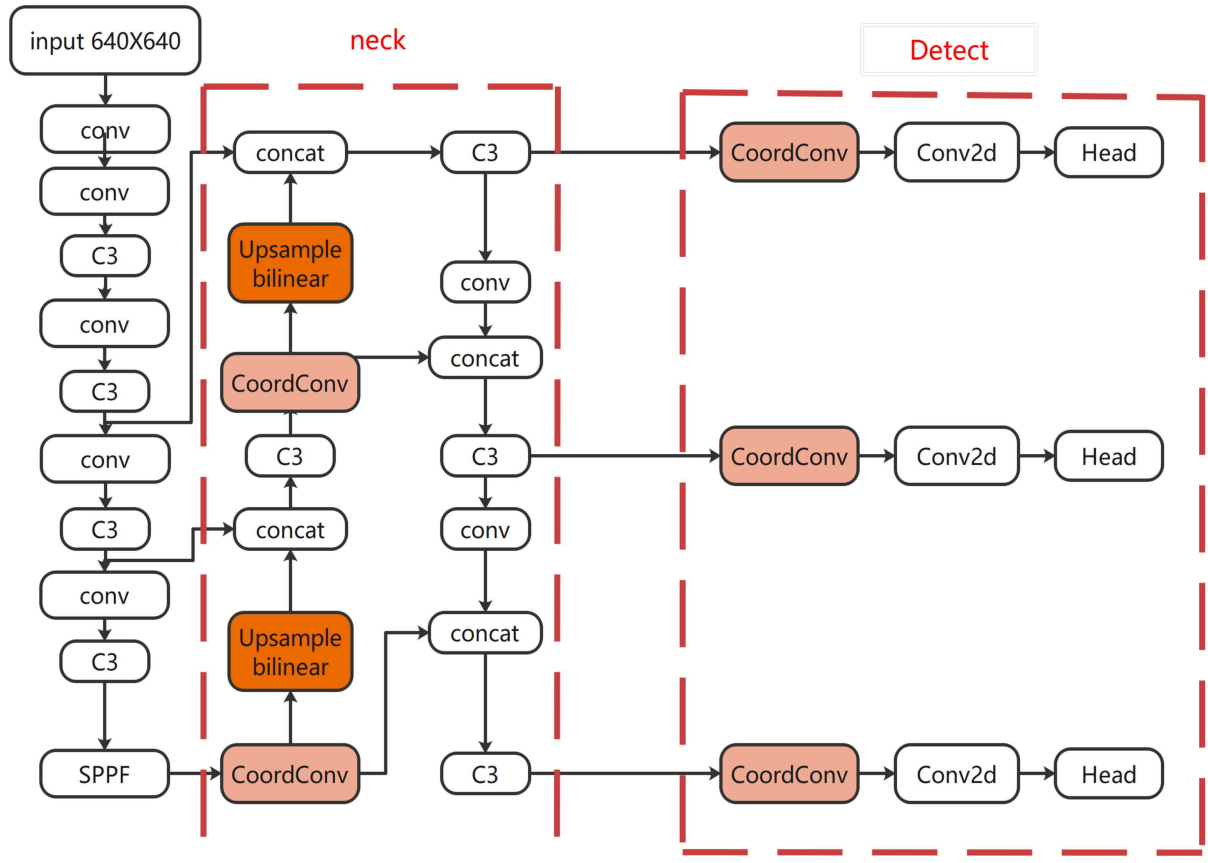


Figure 3. Improved YOLOV5 model
图 3. 改进的 YOLOV5 模型

3.2. 锚框聚类分析

Yolov5 所使用的初始锚框尺寸在检测时并不适用于所有的数据集[10]，为实现冷库机器人的最佳检测效果，采用典型的 kmeans 聚类算法对锚框尺寸进行重新设计，但 kmeans 算法随机性较大对聚类效果会造成影响，选取效果更好随机性较小的 kmeans++算法重新聚类锚框，该算法计算锚框过程中会判断锚框尺寸大小是否还会发生变化，当锚框大小稳定不变时输出锚框。Kmeans++的流程如图 4 所示。通过 kmeans++算法重新聚类得到适合本文同伴机器人的 9 组锚框为(29, 44)、(46, 105)、(73, 229)、(109, 135)、(128, 65)、(143, 273)、(220, 147)、(251, 202)、(302, 328)。

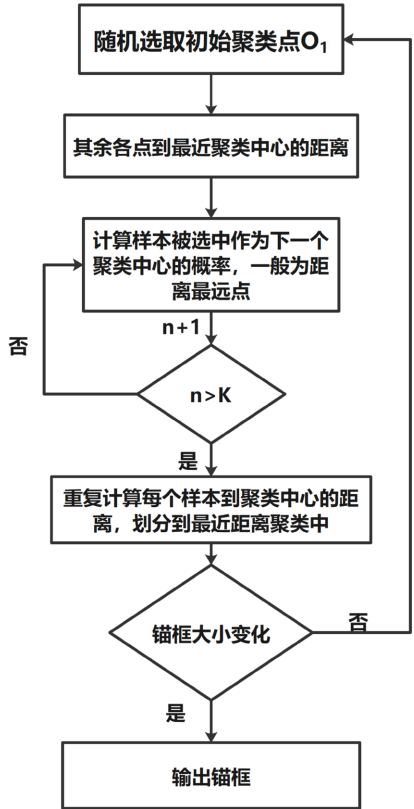


Figure 4. Kmeans++ algorithm flow
图 4. Kmeans++ 算法流程

3.3. 双线性插值

YOLOv5 中上采样方式默认使用的是最近邻插值[11]。最近邻插值法直接复制领边的像素信息容易造成目标的边缘信息丢失，在遇到同伴机器人被货物遮挡的情况时，不能够有效提取特征。而双线性插值通过对图像进行适当的平滑处理和空间扩展可以将遮挡区域填充为合理的像素值，使得物体轮廓更加连续和自然，提供更平滑的图像边缘和更精细的图像细节，故本文模型使用双线性插值替换最近邻插值，提高遮挡时识别精准度。

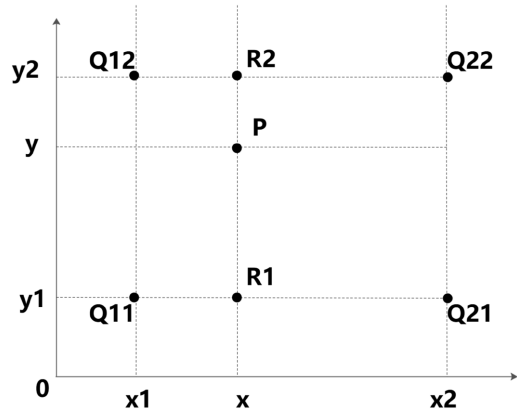


Figure 5. Bilinear interpolation
图 5. 双线性插值

双线性插值法的原理如图 5 通过在像素坐标系中距离目标图像的像素点 $P(x, y)$ 点最近的 $Q11, Q12, Q21, Q22$ 四个点在 x 方向进行插值得到 $R1, R2$ 的像素信息, $R1, R2$ 通过在 y 方向进行线性插值后即可得到目标图像的像素值, 计算公式如(2)~(3)所示。

通过 4 个点在 X 方向的线性插值得到 $R1, R2$ 的像素信息:

$$\begin{aligned} f(R1) &\approx \frac{x2-x}{x2-x1} f(Q11) + \frac{x-x1}{x2-x1} f(Q21) \\ f(R2) &\approx \frac{x2-x}{x2-x1} f(Q12) + \frac{x-x1}{x2-x1} f(Q22) \end{aligned} \quad (2)$$

$R1, R2$ 在 Y 方向的线性插值得到目标像素点 P 的像素信息:

$$f(P) \approx \frac{y2-y}{y2-y1} f(R1) + \frac{y-y1}{y2-y1} f(R2) \quad (3)$$

3.4. CoordConv 卷积

同伴机器人在运动时, 其的空间信息(位置、尺度等)会发生快速变化。YOLOv5 模型通过在 neck 网络使用了 FPN 和 PAN 结构完成对空间变化的感知, 但由于 neck 中特征图经过多次卷积、池化等操作后, 空间分辨率已经较低, 部分细节和空间信息仍可能会丢失。CoordConv 卷积具有引入表达坐标位置信息的特征通道, 可增强模型对坐标信息的敏感度, 提空间位置的抗扰性。

本文在丢失信息的 neck 层中引入 CoordConv 卷积。但若替换 neck 部分全部的卷积, 会使推理的速度变慢, 故本文选择在丢失空间信息较多的 FPN 层中进行替换, 将 FPN 中传统的卷积层替换为 CoordConv 卷积, 在保证推理速度的情况下完成优化。同时为保证获取更多的空间信息在 Detect 中加入 CoordConv 卷积, 有效减少位置偏差问题。在 Detect 层中加入 CoordConv 卷积结构图如图 6 所示。N1, N2, N3 为从 neck 层输出的特征图, 使用 CoordConv 卷积将 N1, N2, N3 层的位置信息提取并逐通道地与输入特征图进行融合, 得到增强版特征图, 增强特征图在传递到后续处理中, 被用来预测目标物体的位置、类别和置信度等信息。能够增强模型对目标特征的感知能力, 有助于提高模型的检测性能。

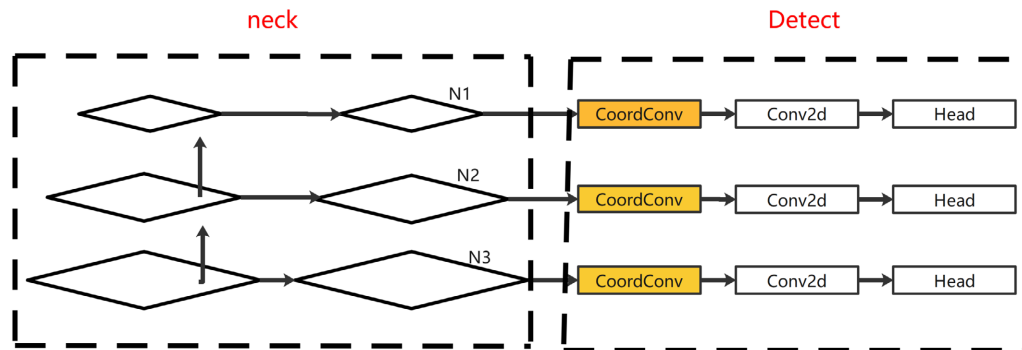


Figure 6. Adding CoordConv to detect
图 6. Detect 中加入 CoordConv

4. 实验结果与分析

4.1. 计算机与训练参数设置

训练所用计算机配置: CPU 使用 IntelCorei5-12500H, 2.50 GHz, 电脑运行内存为 16 G, GPU 采用 NVIDIA GeForce RTX 3050, 应用 Windows11 系统, 使用 pycharm 操作软件, python 版本 3.8, 深度学习

框架为 torch1.9.1, CUDA 11.1。

训练参数的设置: 训练迭代周期为(epochs)100, 批次(batch-size)为 8, 学习率为 0.01, 余弦退火参数为 0.2, 学习率动量为 0.937, 置信度阈值 0.40。本文采用精确率(Precision)、召回率(Recall)、调和平均值(F1)、mAP、几项对模型的性能进行评估。

4.2. 图像增强算法的性能对比

为了验证本文采用 AutomatedMSRCR 算法对低光照图像增强效果的有效性, 本文选取 500 张不同光照强度下的图像分别经过 MSRCR [12]算法、MSRCP [13]算法、AutoMSRCR 算法增强后图像与原始图像在相同软硬件环境下进行目标检测效果对比实验, 实验结果见表 1。

Table 1. Comparison of enhancement algorithms

表 1. 增强算法对比

	mAP0.5%	F1%	P%	R%
原图	83.4	76.7	73.7	79.9
MSRCR	83.9	75.2	81.2	70.1
MSRCP	81.3	75.3	82.5	69.3
AutodMSRCR	83.1	77.8	83.8	72.7

由表 1 可知, 与原始图像相比, MSRCR 算法增强后的图像和精准率值提升 7.3%, mAP0.5 值提高 0.5%, F1 值下降 1.5%, R 值降低 9.8%, 经 MSRCP 增强图像精准率提升 8.8%, mAP0.5 值降低 1.9%, F1 值下降 1.4%, R 值下降 10.6%, 而经过 AutoMSRCR 算法增强的图像精准率提升 10.1%, mAP0.5 值下降 0.3%, F1 值提升 1.1%, R 值降低 7.2%。增强后数据提升明显高于其他算法, 图像质量效果提升最大。综合上述所示, 本文选择 AutomatedMSRCR 算法对于低光照的图像增强效果最好。

4.3. 锚框结果对比

为验证新设锚框的优异性, 采用预设锚框和 kmeans, kmeans++重新设计的锚框分别对数据集进行训练, 结果如表 2, 图 7 所示。

由表 2 和图 7 可知, 不同锚框在迭代的初期都快速上升, 在迭代到 40 次的时候开始趋向稳定, 在 80 次的时候, 新设锚框的优越性开始体现, kmeans++算法的锚框 mAP 曲线明显高于其它 2 组, 模型性能更好。与 YOLOv5 的预设锚框相比, 采用 kmeans 算法重新设计的锚框精准率提升 0.4%, mAP0.5 的值提高 2.4 个百分点, F1 值提升 2.2%, R 值提升 3.3 个百分点。Kmeans++重新聚类的锚框在训练时较预设锚框精准率提升 1.8 个百分点, mAP 值提升 4.3%, F1 值提升 2.9 个百分点, R 值提高 3.6%。通过实验结果表明本文采用的 k-means++对模型提升效果更好。

Table 2. Comparison of results for different anchor frames

表 2. 不同锚框结果对比

	mAP0.5%	F1%	P%	R%
预设	72.4	72.6	81.0	65.9
Kmeans	74.8	74.8	81.4	69.2
Kmeans++	76.7	75.5	82.8	69.5

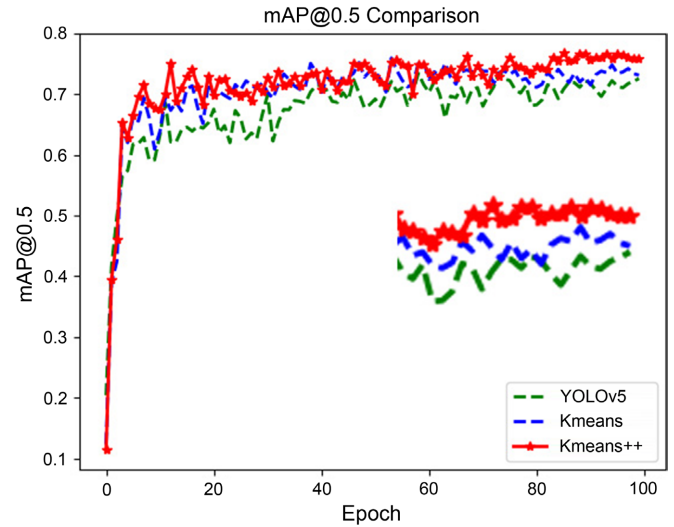


Figure 7. mAP values of different anchor frames
图 7. 不同锚框的 mAP 值

4.4. 不同模型性能指标比较

为进一步验证模型效果，在相同软硬件环境下将改进模型与 YOLOv5，SSD [14]，Faster-RCNN [15]，YOLOV4，YOLOv7-tiny [16]等模型对比实验，对使用 AutomatedMSRCR 算法增强过的数据集进行实验对比，评估结果如表 3。通过表 3 的对比试验结果可知，本文改进模型的检测精度 mAP 相较于 SSD、Faster-RCNN、YOLOv4、YOLOv7-tiny 和 YOLOv5 模型分别提升 3.0%、3.6%、5.0%、4.5%、2.5%，综合上述所示，本文改进后的模型对在低光照冷库环境下运动的同伴机器人识别效果最好，模型的识别精准率更高。

Table 3. Comparison of test results of different models
表 3. 不同模型检测结果对比

	mAP0.5%	F1%	P%	R%
SSD	73.9	73.0	82.1	66.3
Faster-RCNN	73.3	60.0	50.4	75.0
YOLOv4	71.9	65.0	84.2	52.9
YOLOv5	72.4	72.6	81.0	65.9
YOLOv7-tiny	74.4	72.6	73.8	71.5
本文模型	76.9	75.7	84.5	68.6

5. 结论

本文针对机器人在冷库环境中图像质量差，识别物的轮廓与细节信息不够清晰，造成目标丢失等问题，提出使用 AutomatedMSRCR 算法增强图像的质量，优化 YOLOv5 算法加强对识别物的细节信息保留与特征的提取的方法。对比原 YOLOv5 模型在低光照的冷库环境中检测置信度更高，可实现在冷库中对同伴机器人的识别。并以下的结论：

- (1) 提出对于冷库环境内双机协作机器人的同伴识别方法。满足在实际低光照环境中对同伴机器人的

准确识别需求, 实现在冷库检测方面的自动化与智能化, 该方法为在冷库环境内的检测提供思路, 为后续探索冷库内货物检测提供技术支持。

(2) 针对冷库内光照不均匀, 图像质量差的问题, 在图像预处理阶段使用 AutomatedMSRCR 的增强方法, 增强后检测置信度更高, P 值提升 10.1%, F1 值提升 1.1%。

(3) 使用 kmeans++ 算法对同伴机器人锚框尺寸重新聚类, 得到更适合的锚框尺寸, 提升识别的检测精度。使用双线性插值方法代替原有的最近邻插值法, 保留更多的物体细节信息, 使识别物更容易被检测到。在 YOLOv5 的 FPN 层中融合 CoordConv 卷积层, 检测头前增加 CoordConv 层, 增强网络对数据变换的鲁棒性, 加强网络的特征提取能力。改进后 YOLOv5 模型对同伴机器人识别的较原有模型 mAP 提升 4.5%, F1 值提升 3.1%, 精准率提升 3.5%, R 值提升 2.7%, 模型的识别性能得到很大提升。

参考文献

- [1] 李雷孝, 孟闯, 林浩, 等. 基于图像增强与深度学习的安全带目标检测[J]. 计算机工程与设计, 2023, 44(2): 417-24.
- [2] 冯娟, 梁翔宇, 曾立华, 等. 基于改进 YOLO v4 的单环刺螈洞口识别方法[J]. 农业机械学报, 2023, 54(2): 265-74+377.
- [3] 司永胜, 肖坚星, 刘刚, 等. 基于 MSRCR 与改进 YOLO v4 的躺卧奶牛个体识别方法[J]. 农业机械学报, 2023, 54(1): 243-50+62.
- [4] Tang, Z., Zhou, L., Qi, F., et al. (2023) An Improved Lightweight and Real-Time YOLOv5 Network for Detection of Surface Defects on Indocalamus Leaves. *Journal of Real-Time Image Processing*, **20**, Article No. 14. <https://doi.org/10.1007/s11554-023-01281-z>
- [5] 李骏明, 王海瑞, 朱贵富, 等. 基于 RETINEX 的多种水下图像增强算法对比[J]. 中国水运, 2023(2): 96-7.
- [6] 邢晋超, 潘广贞. 改进 YOLOv5s 的手语识别算法研究[J]. 计算机工程与应用, 2022, 58(16): 194-203.
- [7] 于亚龙, 穆远彪. 插值算法的研究[J]. 现代计算机(专业版), 2014(5): 32-5.
- [8] Wei, D., Wei, X., Tang, Q., et al. (2023) RTLseg: A Novel Multi-Component Inspection Network for Railway Track Line Based on Instance Segmentation. *Engineering Applications of Artificial Intelligence*, **119**, 105822. <https://doi.org/10.1016/j.engappai.2023.105822>
- [9] Tong, K., Wu, Y.Q. et al. (2023) Rethinking PASCAL-VOC and MS-COCO Dataset for Small Object Detection. *Journal of Visual Communication and Image Representation*, **93**, 103830. <https://doi.org/10.1016/j.jvcir.2023.103830>
- [10] 庄建军, 叶振兴. 基于改进的 YOLOv5m 电动车骑行者头盔与车牌检测方法[J/OL]. 南京信息工程大学学报(自然科学版): 1-13. <http://kns.cnki.net/kcms/detail/32.1801.N.20230105.0830.001.html>, 2023-10-30.
- [11] 欧阳鼎, 黄华, 李嘉鹏, 等. 改进 YOLOv5s 模型的航拍图像目标检测算法[J]. 福建电脑, 2023, 39(5): 7-15.
- [12] 王利娟, 常霞, 张伯妍, 等. 基于 MSRCR 的自适应低照度图像增强[J]. 现代电子技术, 2022, 45(2): 155-61.
- [13] 阮顺领, 刘丹洋, 白宝军, 等. 基于自适应 MSRCR 算法的煤矿井下图像增强方法[J]. 矿业研究与开发, 2021, 41(11): 186-192.
- [14] Jia, Z., Su, X., Ma, G., et al. (2023) Crack Identification for Marine Engineering Equipment Based on Improved SSD and YOLOv5. *Ocean Engineering*, **268**, 113534. <https://doi.org/10.1016/j.oceaneng.2022.113534>
- [15] Kaihua, Z. and Haikuo, S. (2021) Solder Joint Defect Detection in the Connectors Using Improved Faster-RCNN Algorithm. *Applied Sciences*, **11**, 576. <https://doi.org/10.3390/app11020576>
- [16] 刘浩翰, 樊一鸣, 贺怀清, 等. 改进 YOLOv7-tiny 的目标检测轻量化模型[J]. 计算机工程与应用, 2023, 59(14): 166-175.