李兆鹏, 董自健, 周逸. 改进 YOLOv5 的长文本箱号轻量化检测算法研究[J]. 智能计算机与应用,2026,16(6):128-135.
DOI:10.20169/j.issn.2095-2163.24090905

改进 YOLOv5 的长文本箱号轻量化检测算法研究

李兆鹏, 董自健, 周 逸
(江苏海洋大学 电子工程学院, 江苏 连云港 222005)

摘 要: 集装箱的进出库都需要通过其箱号进行登记,针对现有长文本箱号检测方法精确度较低、检测效果差、难以部署移动端的问题,本文提出一种长文本箱号检测的轻量化模型 YOLOv5-SLT。首先,基于高效扩展卷积空间金字塔网络 ESPNetv2 设计一种 Cross Stage Partial(CSP)结构的特征提取网络 CSP-ESPNet 作为 YOLOv5 的 Backbone,在减少参数量的同时,使用高效的扩展卷积在更大的感受野中学习表征;针对长宽比例较为极端的箱号文本,在特征提取网络中引入坐标注意力机制 Coordinate Attention(CA),并将 CA 在垂直和水平方向的一维特征编码由行或列平均池化改进为行或列平均池化采样再最大池化,得到优化后的坐标注意力机制(Max Pool-CA,MPCA),MPCA 可以解决由于箱号字符间隔较大使得网络无法捕捉其特征的问题;在网络的检测端加入角度因子,使得只能检测水平框的 YOLOv5 通过旋转一定角度得到更好的箱号检测效果。为验证 YOLOv5-SLT 是否有效,对收集的集装箱数据集进行训练测试。结果表明 YOLOv5-SLT 箱号检测精确率 P 为 96.9%,召回率 R 为 98.3%,综合评价指标 $F1$ - Score 为 97.5%,mAP@50 为 99.2%,相比 YOLOv5 在整体参数量上降低了 44%,准确率提高了 2%。
关键词: 集装箱箱号检测; ESPNetv2; YOLOv5; 坐标注意力; 角度因子
中图分类号: TP391.1 **文献标志码**: A **文章编号**: 2095-2163(2026)06-0128-08

Research on lightweight detection algorithm of YOLOv5 long text box number

LI Zhaopeng, DONG Zijian, ZHOU Yi
(School of Electronic Engineering, Jiangsu Ocean University, Lianyungang 222005, Jiangsu, China)

Abstract: The entry and exit of containers need to be registered by their case number. Aiming at the problems of low accuracy, poor detection effect and difficult deployment of mobile terminal for the existing long text box number detection methods, a lightweight model YOLOv5-SLT for non-horizontal long text box number detection is proposed. Firstly, a feature extraction network CSP-ESPNet with Cross Stage Partial (CSP) structure is designed based on the efficient extended convolution space pyramid network ESPNetv2. CSP-ESPNet, as the Backbone of YOLOv5, uses efficient extended convolution to learn representations in a larger receptor field while reducing the number of parameters. Coordinate Attention (CA) is introduced into the feature extraction network for box number texts with extreme length to width ratio, and the one-dimensional feature coding of CA in vertical and horizontal directions is improved from row or column average pooling to row or column average pooling sampling and then maximum pooling. The optimized coordinate attention mechanism (Max Pool-CA, MPCA) is obtained. MPCA can solve the problem that the network can not capture the characteristics of the container number due to the large character interval. The angle factor is introduced into the detection end of the network, so that the YOLOv5 that can only detect the horizontal frame can get better detection effect of the box number by rotating at a certain angle. In order to verify the effectiveness of YOLOv5-SLT, training tests are carried out on the collected container data set. The experimental results show that the detection accuracy rate P of YOLOv5-SLT box number is 96.9%, the recall rate of R is 98.3%, the comprehensive evaluation index $F1$ - Score is 97.5%, and mAP@50 is 99.2%. Compared with YOLOv5, the overall parameter number is reduced by 44%, the accuracy is increased by 2%.
Key words: container number detection; ESPNetv2; YOLOv5; coordinate attention; angle factor

0 引 言

集装箱作为承载货物的载体之一,具有极高的安全性和规范性,在全球都是通用的;每个集装箱都有唯一的身份标识,由共 11 位英文和数字组合而成;运输部门都是通过该标识对集装箱进行管理,先

作者简介: 李兆鹏(2001—),男,硕士研究生。主要研究方向:计算机视觉应用; 周 逸(2000—),男,硕士研究生。主要研究方向:物联网通信技术。
通信作者: 董自健(1973—),男,博士,教授。主要研究方向:检测与控制,工业自动化。Email:dzjian@126.com。
收稿日期: 2024-09-09

由人工记录到后来借助信标(RFID)进行跟踪和管理[1]；基于深度学习识别集装箱箱号分为 2 个步骤，分别为定位箱号的位置和识别箱号，有效的检测箱号对后续的识别有着很大的影响。除此以外，移动端由于自身较低的硬件配置导致部署的模型要更加轻量化。

随着深度学习在计算机视觉领域的快速发展，出现许多优秀的目标检测算法，大致分为双阶段和单阶段两种检测算法。相比双阶段目标检测算法，单阶段目标检测算法检测速度很快，网络架构简单，成为各领域的研究对象，最具代表性的算法是 YOLO 目标检测算法[2]。洪炎等学者[3]通过改进 YOLOv8 算法，引入注意力机制和修改检测头实现不同尺度煤矿异物的检测。郑文轩等学者[4]通过引入 MobileNetv3 和 NWD 小目标检测机制来优化 YOLOv7 网络，实现香梨的自动检测。叶涵宇等学者[5]通过引入轻量化模块单元和 CBAM 注意力机制来改进 YOLOv5 算法实现对生活垃圾进行分类。

在集装箱箱号检测任务中，张添添等学者[6]提出一种基于 YOLOv3 的箱号快速定位检测方法，引入 EfficientNetv2 轻型网络和通道剪枝方法，旨在减少网络模型参数量，并增加 SPPF，捕获多尺度上下文信息。沈嘉康[7]通过在 DBNet 网络中引入注意力机制，并与引入空间变换网络的 CRNN 进行串联推理，提高集装箱箱号的检测效率。游索等学者[8]设计了一种基于热图感知的像素级检测架构，该架构结合了 CBAM 与 MobileNetv3 的轻量级初定位网络，以获取目标箱号的预选区域，并使用基于 Transformer 的自适应网络，从中精确提取感兴趣的箱号区域。张炎等学者[9]提出一种基于 YOLOv4 的复杂环境的集装箱箱号检测方法。涂铮等学者[10]将箱号检测和识别结合，改进 DBNet 和二维注意力文本识别方法 SAR，对整个集装箱重量和容积信息来进行检测识别。崔循等学者[11]通过对 Faster R-CNN 进行改进来实现集装箱箱号的定位。孙艳春等学者[12]利用蚁群优化算法进行箱号区域边缘检测，从而实现对箱号区域的有效分割。Zhang 等学者[13]结合基于线分割掩码(LSM)的检测和基于扫描线的识别设计了一个增强 OCR 系统，旨在解决由于文本倾斜和文本粘连的问题。

这些方法在箱号检测方面都有着不错的表现，但针对箱号扭曲和倾斜的场景，在检测精度和网络轻量化两者之间都难以在一个较高水平上实现平衡。因此，本文以检测速度和精度都较高的 YOLOv5 为基础网络架构，引入利用卷积分解思想来降低网络参数量的高效扩展卷积空间金字塔网络 ESPNetv2，并对其进行改进，设计具有 Cross Stage Partial (CSP)结构[14]的高效特征提取骨干网络 CSP-ESPNet，让模型学习更丰富的特征信息；进一步融合 Coordinate Attention(CA)注意力机制[15]，在 CA 注意力机制的输入端添加最大池化(Max Pool)得到 MPCA 模块，以提高特征提取网络对长文本且字符间隔较大的箱号特征信息的提取能力；针对倾斜和扭曲的文本检测，引入角度因子，提高网络预测的检测框效果。

1　轻量化 YOLOv5 目标检测网络设计

在众多的 YOLO 目标检测算法版本中，YOLOv5 有着较高的稳定性和可靠性。在一些轻量级任务中，YOLOv5 的推理速度表现良好，而且 YOLOv5 提供多种大小的模型，部署灵活，因此本研究对 YOLOv5 进行改进。

将改进后的箱号检测网络架构命名为 YOLOv5-SLT，如图 1 所示。集装箱箱号图像经过由 ESPNetv2 和 MPCA 融合的 CSP-ESPNet 特征提取网络，得到长文本箱号的特征信息；接着，进入特征融合网络进行高级语义和低级细节纹理信息的特征融合；然后，通过 Detect 检测头输出带有角度预测的检测框信息，并进行损失计算。

1.1　CSP-ESPNet 特征提取网络设计

ESPNetv2 是 Mehta 等学者[16]针对边缘设备研发的一种十分轻量的网络架构。图 2(a)为组成整个 ESPNetv2 网络的 EESP 单元，整体基于减少-分裂-转换-合并策略，EESP 单元首先使用 Group Point-wise 卷积将高维输入特征映射投影到低维空间。然后，使用不同扩展率的深度扩展可分离卷积并行学习表征，每个分支的不同扩张率使 EESP 单元能够从一个大的有效感受野中学习表征。接着，使用分层特征融合方法解决扩展卷积带来的网格栅影问题，以层次化的方式对上一层扩展卷积学习的特征进行加性融合。最后，再与 EESP 单元输入时的特征直接相加，从而获得丰富的特征信息；为了在多尺度上有效地学习特征，将原本的深度扩展卷积替换成跨行扩展卷积，在输入特征与 EESP 单元的输出特征相加的路径上添加平均池化操作，将原有的加法操作改为串联操作，这样可以有效地扩展特征映射维度、即 Stride EESP(S-ESP2)单元，如图 2(b)所示。

为了使网络学习到更多的特征信息，参考

YOLOv5 的 CSP 结构,设计一种 CSP-ESPNet 特征提取网络。CSP-ESPNet 特征提取网络主要由 S-ESP2 模块和 CSP-ESPn 模块构成,CSP-ESPn 模块架构如图 3 所示,该模块参考 YOLOv5 中的 C3 模块结构,输入的特征图通过 1×1 卷积计算分为 2 个相同的特征图,在保留原有特征的同时,对其中的一个特征图进行多次特征提取,最后在输出前将两者进行通道层面上的拼接,这样会在保留原有特征内容的基础上学习原特征图中更高级的特征信息,极大地丰富了特征提取内容。

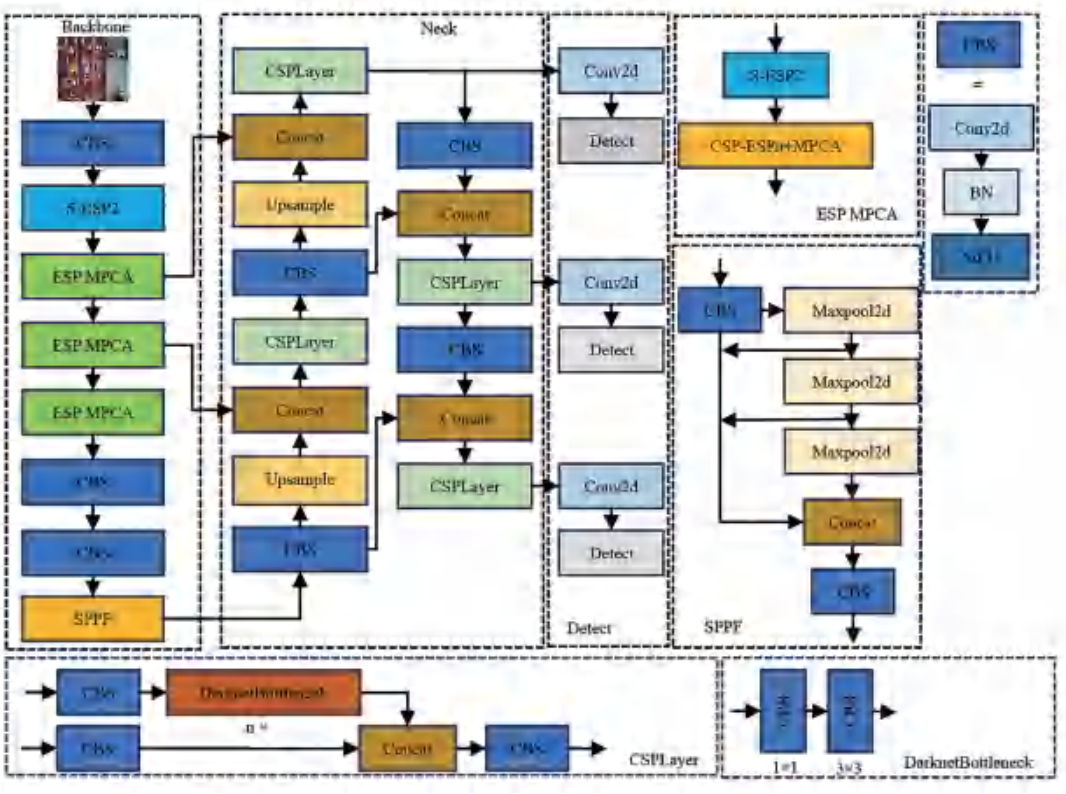
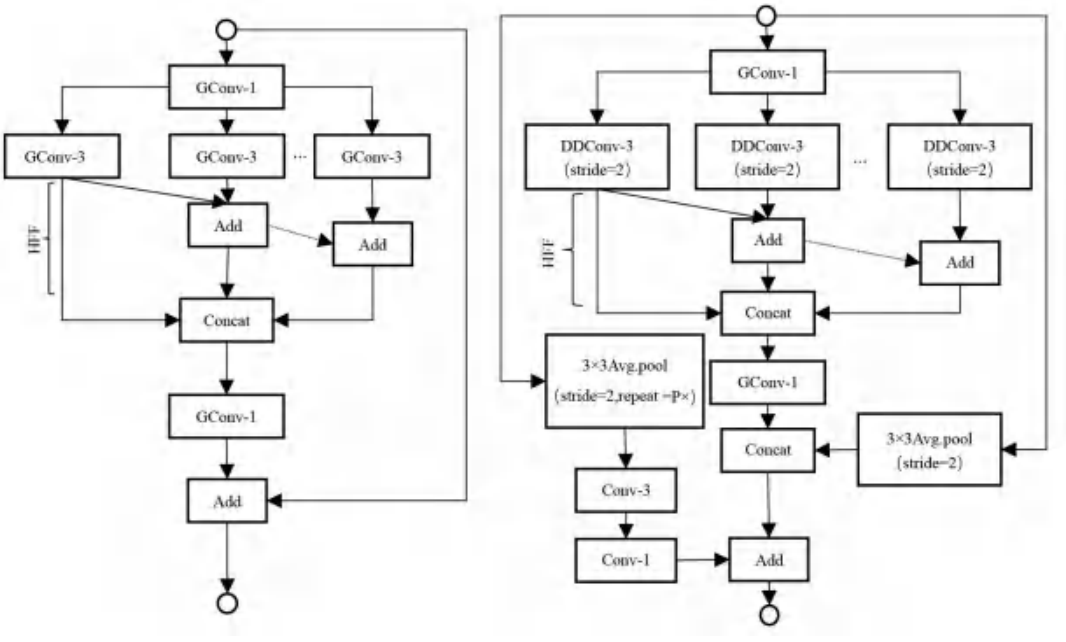


图 1 YOLOv5-SLT 集装箱箱号检测框架

Fig. 1 YOLOv5-SLT container number detection framework



(a) EESP 单元 (b) S-ESP2 单元

图 2 EESP 单元和 Stride EESP(S-ESP2)单元

Fig. 2 EESP unit and Stride EESP (S-ESP2) unit

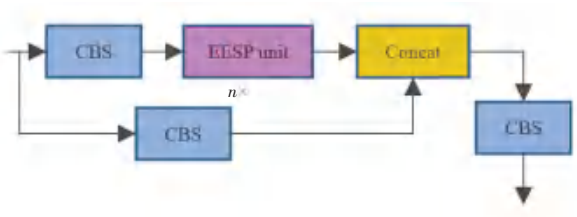


图 3　CSP-ESP*n* 模块

Fig. 3　CSP-ESP*n* module

1.2　MPCA 模块

Coordinate Attention(CA)是一种引入坐标信息来增强网络特征提取的注意力机制，使得模型能够更好地捕捉空间结构和特征分布。但在对集装箱箱号检测时，出现无法将箱号完整框住的情况，如图 4 所示。CA 是分别对垂直和水平两个方向进行一维特征编码(平均池化)，直接对每个通道的特征图的行或列元素进行求和并计算平均值，可能会削弱对较长文本的特征信息捕捉，在检测图 4 中这种单个字符间隔较大的箱号无法捕捉箱号字符之间的联系，为解决这一问题，设计一个 MPCA 模块，如图 5 所示。



图 4　添加 CA 的箱号检测效果

Fig. 4　Detection effect of adding the CA box number

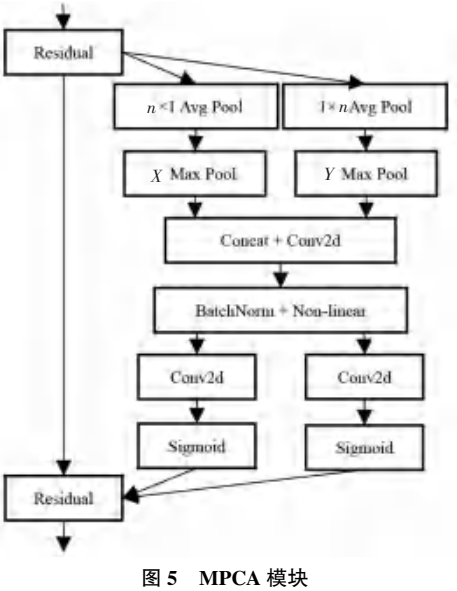


图 5　MPCA 模块

Fig. 5　MPCA module

MPCA 模块是对 CA 注意力模块进行改进；不再直接计算某一行或列的平均值，而是先使用 $n \times 1$ 和 $1 \times n$ 的窗口，通过一定步长，得到多个平均池化结果，这样可以获得每个字符之间的特征联系，再通过最大池化操作，使得网络关注联系较为紧密的字符区域。随着多层特征提取，可以在一定程度上增强网络对字符间隔较大的箱号文本特征关注度。

1.3　Detect 模块

原始集装箱图像经过骨干特征网络和特征融合网络后通过 Detect 检测头输出目标的预测框、置信度、类别信息；由于 YOLOv5 只能预测水平框，对倾斜的集装箱箱号检测效果不好，会含有大量的背景信息，因此引入角度因子使得水平框进行一定角度的旋转，从而更加贴合箱号文本，减少背景信息对后续的箱号识别产生的噪声干扰。

Detect 层加入角度信息后，总损失函数可表示为：

$$L_{\text{loss}} = \alpha_1 L_{\text{bbox}} + \alpha_2 L_{\text{cls}} + \alpha_3 L_{\text{conf}} + \alpha_4 L_{\text{angle}} \tag{1}$$

其中，L_{bbox}，L_{cls}，L_{conf}，L_{angle} 分别表示目标框损失、类别损失、置信度损失和角度损失，α_1，α_2，α_3，α_4 分别表示每个损失函数在总损失的权重系数。

式(1)等号右侧从左到右分别为目标框损失、类别损失、置信度损失和角度损失，前面的系数表示每个损失函数在总损失的权重系数。前三者采用 YOLOv5 的框损失函数、类别损失函数、置信度损失函数。对于角度损失函数，研究给出了如下公式：

$$L_{\text{angle}} = 1 - \arccos(\theta_p - \theta_g) \tag{2}$$

其中，θ_p、θ_g 分别表示预测的旋转角度和真实框的旋转角度。

研究中使用式(2)会出现训练过程中损失值突然变大，并变为 NAN 的情况，说明该损失函数设计的不合理，分析其原因可能是角度周期性和边界性问题，假设网络预测的角度为顺时针 150°，但真实角度为逆时针 -30°。实际 2 种检测框的效果是相同的，但损失函数计算是不一样的。

针对角度周期性和边界性问题，采用移相编码(PSC)方法[17]。将角度预测变为余弦值预测，余弦函数的周期为 2π，旋转矩形框的周期为 π，所以两者之间的关系可表示为：

$$\varphi = 2\theta \tag{3}$$

其中，φ 表示余弦函数的角度，θ 表示旋转框角度。

单个余弦值无法求出单个角度，因此 PSC 方法借鉴光学中的移相技术，通过预测多个相位偏移的余弦值来计算出唯一的角度值，具体可表示为：

$$x_{(p,g)n} = \cos\left(\varphi + \frac{2n\pi}{N_{\text{step}}}\right) \tag{4}$$

其中，$n = 1,2,3,\cdots,N_{\text{step}}$，$(p,g)$ 表示(网络预测,真实标注)。

因此网络输出由原本的 $(x,y,w,h,\text{conf},\text{cls},\text{angle})$ 变为 $(x,y,w,h,\text{conf},\text{cls},x_1,x_2,\cdots,x_n)$，前 5 个参数表示目标的检测框信息、置信度和类别;角度损失函数可表示为:

$$L_{\text{angle}} = \frac{1}{n}\sum_{i=1}^{n}(2 \times S(x_{p,n}) - 1 - x_{g,n})^2 \tag{5}$$

$$S(x_{p,n}) = \text{Sigmoid}(x_{p,n}) \tag{6}$$

其中，$x_{p,n}$ 表示第 n 个移相余弦值的预测值，$x_{g,n}$ 则表示第 n 个移相余弦值的真实值，这里 $n = 1,2,3,\cdots,N_{\text{step}}$。

训练好的网络最终输出 n 个移相余弦值，通过反正切函数求得旋转角度，表达式为:

$$\varphi = -\arctan\frac{\sum_{n=1}^{N_{\text{step}}} x_{p,n}\sin\left(\frac{2n\pi}{N_{\text{step}}}\right)}{\sum_{n=1}^{N_{\text{step}}} x_{p,n}\cos\left(\frac{2n\pi}{N_{\text{step}}}\right)} \tag{7}$$

$$\theta = \frac{\varphi}{2} \tag{8}$$

2 实验结果与分析

2.1 实验数据集

实验数据集来自 AI Studio 平台，图片大小为 1 920×1 080;共 2 814 张，其中包括通过添加雾化、雨天、强光的增强数据图片，增强效果如图 6 所示;使用标注工具标注集装箱箱号，并将总数据集按照 7∶2∶1 划分为训练集、验证集、测试集。



(a) 雾化 (b) 强光 (c) 雨天

图 6 数据增强

Fig. 6 Data enhancement

2.2 实验配置

表 1 和表 2 分别为仿真中对比实验、验证实验的各项配置参数，为保证实验的客观性，在相同的数据集和各项配置参数下进行实验验证。

2.3 MPCA 效果实验验证

为了验证优化后的 CA 注意力机制是否能够捕获字符间隔较大的箱号特征，对添加 CA 和 MPCA 的 YOLOv5-SLT 网络进行训练，最终生成的热力图如图 7 所示。

表 1 实验硬件配置

Table 1 Experimental hardware configuration

实验环境	配置
操作系统	Linux
CPU	Intel(R) Xeon(R) Platinum 8255C
GPU	RTX 3080(10 GB)
深度学习框架	Pytorch

表 2 训练参数配置

Table 2 Training parameters configuration

参数	配置
训练迭代次数	300
Batch-Size	16
初始学习率	0.01
权重衰减	0.005
优化器	SGD

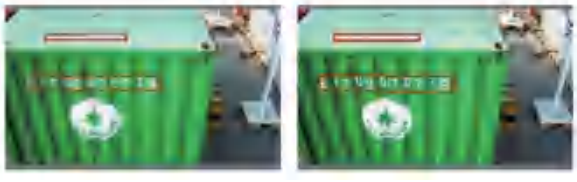


(a) CA (b) MPCA

图 7 CA 注意力机制修改前后热力图对比

Fig. 7 Comparison of heat map before and after modification of CA attention mechanism

从图 7 中可以发现，由于字符“E”距后面字符较大且处于边缘，添加 CA 的检测网络未能捕捉到其特征，从而忽略了字符 E，添加修改后的 MPCA 在一定程度上能够捕捉到字符“E”的特征，并且网络输出检测框完全框住了箱号文本，效果如图 8 所示，充分说明了 MPCA 的有效性。



(a) CA (b) MPCA

图 8 CA 注意力机制修改前后检测框对比

Fig. 8 Comparison of detection boxes before and after modification of CA attention mechanism

2.4 骨干特征提取网络性能验证

降低网络复杂性有多种方法。相比网络模型剪枝和知识蒸馏，轻量化的网络特征提取网络较为简单和便捷，易于部署在深度学习网络中;为验证 ESPNetv2 的高效性，以 YOLOv5 为基础框架，与添

加 GhostNetv2[18]、MobileNetv2[19]、ShuffleNetv2[20]三种轻量化深度学习网络为特征提取网络进行对比，分别从 $F1$ - score、mAP@50、参数量、计算量四个评价指标进行分析。仿真结果见表 3。从表 3 中可发现相比前 3 种，ESPNetv2 特征提取网络在各方面都表现良好。

表 3　轻量化网络性能对比

Table 3　Lightweight network performance comparison

方法	$F1$/ %	mAP@50/%	参数量/10^6 个	计算量/ GFLOPs
GhostNetv2[18]	96.0	99.2	5.07	10.0
MobileNetv2[19]	95.7	99.1	4.63	11.1
ShuffleNetv2[20]	95.9	99.1	4.44	8.5
ESPNetv2	97.5	99.2	3.92	7.0

2.5　消融实验

为了验证 YOLOv5-SLT 网络改进的每一个点是否有效，对 YOLOv5 网络分别添加角度预测(A)、CSP-ESPNet(B)、MPCA(C)进行对比实验，结果见表 4；从表 4 中可以发现，序号 3 参数量和计算量比序号 1 减少了近一半，而且精确率提高了 2.5%，说明 CSP-ESPNet 是有效的，虽然召回率和 mAP@50 有所降低，但相比模型参数量和计算量的显著减少是可忽略的；序号 4 的精确率、召回率相比前三者都比较高，表明添加 MPCA 注意力机制是有效的，能够弥补网络减少参数量导致精准度降低的不足；但在推理时间上，比优秀的 YOLOv5 算法多了 0.03 s，在移动端检测耗时的允许范围内。

表 4　消融实验结果

Table 4　Results of ablation experiments

序号	A	B	C	精确率/%	召回率/%	mAP@50/%	参数量/10^6 个	计算量/ GFLOPs	检测时间/ms
1	×	×	×	94.2	98.0	99.2	7.0	15.8	26.8
2	√	×	×	96.1	98.4	98.8	7.0	14.4	58.4
3	√	√	×	96.7	96.8	98.9	3.9	7.0	56.5
4	√	√	√	96.9	98.3	99.2	3.9	7.0	57.2

2.6　文本检测算法对比

基于深度学习的文本检测算法大致分为基于回归和基于分割两种算法。为了验证本文提出 YOLOv5-SLT 算法在集装箱箱号检测上的性能，将本文算法与基于分割的 DBNet++[21]、PANet[22]和基于回归的 EAST[23]算法进行对比。除此以外，YOLOv5-SLT 算法与加入旋转框预测的 YOLOv8[24]进行对比，结果见表 5。从整体结果分析，基于分割思想的 DBNet++和 PANet 文本检测算法都优于基于回归思想的算法，这是由于基于分割思想的文本检测算法是对图像进行像素级的预测，因此该类算法在准确率上都有着很好的表现，但其网络参数量和权重文件都比较大，不适合部署于移动端，而 YOLOv5-SLT 算法与基于分割思想的文本检测算法在精确度和召回率上相差不是很大，且在参数量和检测时间上，YOLOv5-SLT 网络的权重文件和参数量都很小，在部署于移动端时占有优势；与 EAST 算法相比，YOLOv5-SLT 在准确率和召回率上均优于 EAST；YOLOv8 是 YOLO 算法的一种较新的算法框架，表 5 中可看出 YOLOv8 各项都优于 YOLOv5-SLT，但 YOLOv5 算法在实际应用中，其代码实现和移植部署，相比 YOLOv8 是比较成熟的，因此在硬件资源较差的条件下，一般会优先使用 YOLOv5 算法作为基础架构。

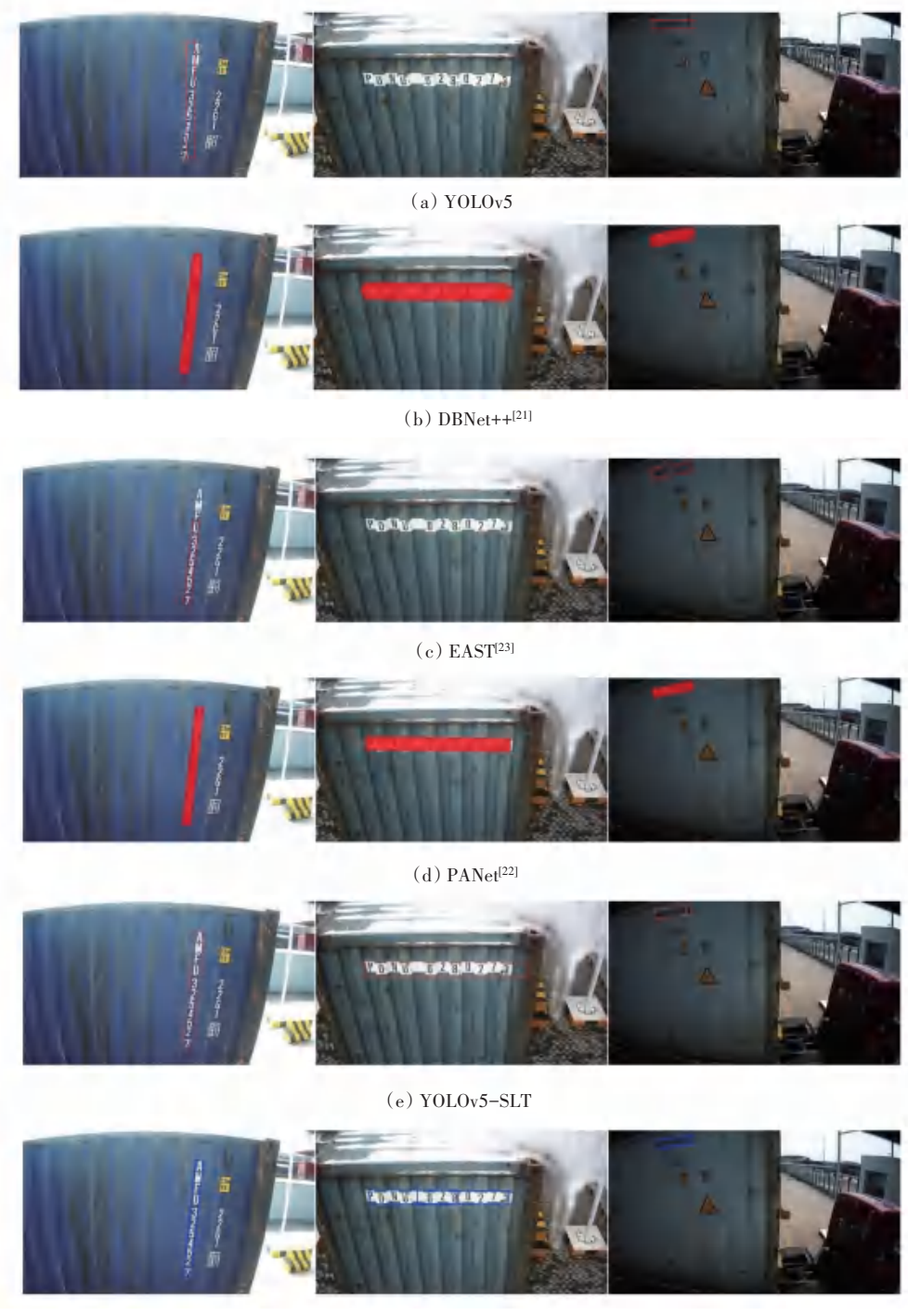
表 5　文本检测算法对比

Table 5　Comparison of text detection algorithms

网络模型	精确率/%	召回率/%	参数量/10^6 个	权重大小/M	检测时间/ms
EAST[23]	95.9	93.4	2.9	18.6	18.9
PANet[22]	97.6	99.4	12.2	151.2	111.0
DBNet++[21]	97.4	99.4	12.3	208.0	119.0
YOLOv8-OBB	97.5	98.8	3.0	6.3	17.0
YOLOv5	94.2	98.0	7.0	13.6	26.8
本文算法	96.9	98.3	3.9	8.3	57.2

图 9 为 YOLOv5-SLT 与基于分割和回归的文本检测算法的箱号检测效果。从图 9 中可看出，YOLOv5-SLT 在较少的参数量和计算量下，与近几年的几种文本检测算法相比，依旧保持着不错的箱

号检测效果；相比 YOLOv5，引入角度预测的 YOLOv5-SLT 的预测框更加贴合箱号，并且在检测字符间距较大的箱号文本时，YOLOv5-SLT 能够很好地框选整个文本，也表明 MPCA 模块是有效的。



(a) YOLOv5

(b) DBNet++[21]

(c) EAST[23]

(d) PANet[22]

(e) YOLOv5-SLT

(f) YOLOv8-OBB

图 9 算法检测效果对比

Fig. 9 Comparison of the detection results of various algorithms

3 结束语

针对集装箱箱号由于拍摄角度导致的文本倾斜和文本长宽比例较大而带来检测效果较差的问题、以及网络模型参数量较大不易部署于移动端的问题，本文提出一种改进 YOLOv5 的轻量化模型 YOLOv5-SLT。使用轻量化网络 ESPNetv2 和 MPCA 注意力机制改进特征提取网络，增强网络对字符间隔较大的长文本箱号特征信息提取能力，引入角度因子，并采用移相编码技术设计角度损失函数，避免

了角度周期性和边界性问题。

YOLOv5-SLT 在集装箱箱号检测任务中能够很好地检测到长文本箱号，为后续的文本识别提供良好的识别箱号文本图像，增加识别的精确度。但所有集装箱箱号印刷不一定都是长文本，也有将箱号分成 2 行的，使用 YOLOv5-SLT 是无法做到完整检测的；因此在后续的研究中，考虑是否能够借鉴图像分割技术来解决这个问题。

参考文献

[1] 张科昌，生绿伟，刘国辉，等. 集装箱码头箱号识别系统研发综述[J]. 港口装卸，2022（2）：39-42.

[2] REDMON J，DIVVALA S，GIRSHICK R，et al. You only lo-ok once：Unified，real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway，NJ：IEEE，2016：779-788.

[3] 洪炎，汪磊，苏静明，等. 基于改进 YOLOv8 的煤矿输送带异物检测[J]. 工矿自动化，2024，50(6)：61-69.

[4] 郑文轩，杨瑛. 基于频域数据增强与轻量化 YOLOv7 模型的成熟期香梨目标检测方法[J]. 农业机械学报，2024，55(5)：244-253.

[5] 叶涵宇，郭来德，李玥娴，等. 改进的 YOLOv5 算法下的智能分类生活垃圾系统设计[J]. 环境工程，2025，43(4)：232-241.

[6] 张添添，周书民，蓝贤桂. 复杂字符干扰场景下铁路集装箱箱号快速定位方法研究[J]. 现代电子技术，2023，46(11)：81-87.

[7] 沈嘉康. 基于改进 DBNet 与改进 CRNN 的集装箱箱号识别系统[J]. 工业控制计算机，2024，37(3)：54-56.

[8] 游索，陈平平，林键辉，等. 基于热图感知的复杂场景港口箱号检测算法[J]. 无线电工程，2023，53(11)：2619-2625.

[9] 张炎，杜培明. 基于 YOLOv4 网络的集装箱箱号目标检测研究[J]. 常州工学院学报，2023，36(2)：33-38.

[10] 涂铮，万志伟. 集装箱重量字符实时视觉识别方法[J]. 港口装卸，2022（6）：39-43.

[11] 崔循，景文博，于洪洋，等. 基于改进 Faster R-CNN 的集装箱箱号定位算法[J]. 长春理工大学学报(自然科学版)，2020，43(1)：45-50.

[12] 孙艳春，李小平. 蚁群优化算法在集装箱箱号定位中应用[J]. 现代计算机(专业版)，2018（13）：54-58.

[13] ZHANG Zhichao，DING Yi，LI Rui，et al. Enhancing OCR with line segmentation mask for container text recognition in container terminal[J]. Engineering Applications of Artificial Intelligence，2024，133：108667.

[14] WANG C Y，LIAO H Y M，WU Y H，et al. CSPNet：A new backbone that can enhance learning capability of CNN[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway，NJ：IEEE，2020：1571-1580.

[15] HOU Qibin，ZHOU Daquan，FENG Jiashi. Coordinate attention for efficient mobile network design[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway，NJ：IEEE，2021：13708-13717.

[16] MEHTA S，RASTEGARI M，SHAPIRO L，et al. ESPNetv2：A light-weight，power efficient，and general purpose convolutional neural network[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway，NJ：IEEE，2019：9182-9192.

[17] YU Yi，DA Feipeng. Phase-shifting coder：predicting accurate orientation in oriented object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway，NJ：IEEE，2023：13354-13363.

[18] TANG Yehui，HAN Kai，GUO Jianyuan，et al. GhostNetV2：Enhance cheap operation with long-range attention[J]. arXiv preprint arXiv，2211.12905，2022.

[19] SANDLER M，HOWARD A，ZHU M，et al. MobileNetV2：Inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway，NJ：IEEE，2018：4510-4520.

[20] MA N，ZHANG X，ZHENG H，et al. ShuffleNetV2：Practical guidelines for efficient CNN architecture design[C]//Proceedings of the European Conference on Computer Vision. Cham：Springer，2018：11218.

[21] LIAO Minghui，ZOU Zhisheng，WAN Zhaoyi，et al. Real-time scene text detection with differentiable binarization and adaptive scale fusion[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，2022，44(11)：7249-7262.

[22] WANG Wenhai，XIE Enze，SONG Xiaoge，et al. Efficient and accurate arbitrary-shaped text detection with pixel aggregation network[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway，NJ：IEEE，2019：8439-8448.

[23] ZHOU Xinyu，YAO Cong，WEN He，et al. EAST：an efficient and accurate scene text detector[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway，NJ：IEEE，2017：2642-2651.

[24] VARGHESE R，SAMBATH M. YOLOv8：A novel object detection algorithm with enhanced performance and robustness[C]//Proceedings of the Advances in Data Engineering and Intelligent Computing Systems. Piscataway，NJ：IEEE，2024：1-6.