

周强. 低光图像目标检测的优化研究[J]. 智能计算机与应用, 2026, 16(4): 76-84. DOI: 10.20169/j.issn.2095-2163.25120401

# 低光图像目标检测的优化研究

周强

(学习强国学习平台有限责任公司, 北京 100010)

**摘要:** 针对现有目标检测算法处理低光照场景数据集存在检测精度低、鲁棒性不足的问题, 提出一种基于 YOLOv11n 的低光照目标检测算法模型 YOLOv11n\_CPM。首先, 通过对 SPPF 模块基于跨阶段局部网络 (CSPNet) 结构进行改进, 并引入金字塔压缩注意力, 创新设计出多尺度特征融合模块 CSPP\_SF, 可有效建立更长距离的通道依赖关系, 抑制背景噪声, 实现参数共享、提升计算效率和检测性能、扩大感受野; 其次, 通过引入坐标注意力并替换 CSPP\_SF 模块的 SPPF, 提出 CSPP\_CA 模块, 可有效捕获方向位置和跨通道信息, 减少计算量和参数量, 实现模型轻量化; 最后, 引入 MPDIoU 损失函数, 提升小目标检测效果。实验结果表明, 本文提出的 YOLOv11n\_CPM 算法模型在低光照图像场景实现了更好的检测精度和效果, 在 ExDark 数据集上的 mAP@0.5 和 mAP@0.5:0.95 分别达到 77.1% 和 49.8%, 较原始 YOLOv11n 算法分别提高了 2.1 和 1.2 个百分点, 模型大小降低了约 40.6%, 模型参数量降低了约 41.7%。为进一步验证本模型算法有效性及泛化能力, 在 LoLI-Street 数据集 11 000 张低光 (Dense) 图像上 mAP@0.5 和 mAP@0.5:0.95 分别达到 46.7% 和 33.4%, 较原始 YOLOv11n 算法分别提高了 4.9 和 1.8 个百分点, 模型大小降低了约 40.5%, 模型参数量降低了约 41.5%。

**关键词:** 目标检测; YOLOv11; 金字塔压缩注意力; 坐标注意力; 损失函数

中图分类号: TP389.1

文献标志码: A

文章编号: 2095-2163(2026)04-0076-09

## Research on optimization of object detection in low-light images

ZHOU Qiang

(Xuexi Co., Ltd., Beijing 100010, China)

**Abstract:** In order to solve the problems of low detection accuracy and insufficient robustness of existing object detection algorithms in processing low-light scene datasets, a low-light object detection algorithm model YOLOv11n\_CPM based on YOLOv11n is proposed. Firstly, the multi-scale feature fusion module CSPP\_SF is designed innovatively by improving the structure of the SPPF module based on the Cross Stage Partial Network (CSPNet) and introducing the pyramid squeeze attention. It can effectively establish a longer-distance channel dependence relationship, suppress background noise, realize parameter sharing, improve calculation efficiency and detection performance, and expand the receptive field. Secondly, by introducing coordinate attention and replacing the SPPF of the CSPP\_SF module, The CSPP\_CA module is proposed, which can effectively capture the direction position and cross-channel information, reduce the amount of calculation and parameters, and realize the model lightweight. Finally, the MPDIoU loss function is introduced to improve the small object detection effect. Experimental results show that the proposed YOLOv11n\_CPM algorithm model achieves better detection accuracy and effect in low-light scene images, and the mAP@0.5 and mAP@0.5:0.95 on the ExDark dataset reach 77.1% and 49.8%, respectively. Compared with the original YOLOv11n algorithm, it improves 2.1 and 1.2 percentage points, reduces the model size by about 40.6%, and reduces the number of model parameters by about 41.7%. In order to further verify that effectiveness and generalization ability of the proposed algorithm, the mAP@0.5 and mAP@0.5:0.95 values are 46.7% and 33.4% respectively on 11 000 low-light (dense) images from the LoLI-Street dataset, which are 4.9 and 1.8 percentage points higher than the original YOLOv11n algorithm, respectively. The model size was reduced by about 40.5% and the number of model parameters was reduced by about 41.5%.

**Key words:** object detection; YOLOv11; pyramid squeeze attention; coordinate attention; loss function

## 0 引言

目标检测作为计算机视觉领域的核心研究方

向, 已在车辆检测、遥感监测、医学图像分析、农业、安防监控等多个领域取得成熟应用。然而, 主流目标检测算法大多基于正常光照条件进行相关研究,

作者简介: 周强(1983—), 男, 硕士, 高级工程师, 主要研究方向: 人工智能, 软件工程。Email: zhouqianghit@126.com。

收稿日期: 2025-12-04

哈尔滨工业大学主办 ◆ 系统开发与应用

当应用于低光场景时,该类算法的性能会显著下降。相较于正常光照场景,低光照场景下的图像通常伴随着高噪声、低对比度和细节丢失等问题,导致低光照场景和正常光照场景之间存在差异性,针对正常光照场景的算法往往不能有效应用于低光照场景<sup>[1]</sup>,为此,研究如何提升低光场景下图像目标检测的性能,对于解决实际问题具有重要的理论与实际应用价值<sup>[2]</sup>。

目前,提升低光照场景目标检测性能的方法主要分为3类,第一类研究将可见光与红外图像融合为高质量图像后进行检测,如文献[3]提出一种残差网络将弱光条件的可见光与红外图像融合,然后对融合图像进行检测,但该方法在目标数量较多的场景下,容易出现目标效果融合不均衡的问题。文献[4]提出了一种基于可见光-红外图像的双模态特征融合的目标检测算法,提升了检测精度,但该模型检测的实时性和适用性方面仍需提升;第二类研究通过图像增强技术先对图像进行预处理,提升图像质量后再进行检测,如文献[5-6],其中文献[6]仍需提升密集目标检测场景的检测效果,因此无法确保低光增强算法效果;第三类研究通过设计新的网络结构或检测算法,直接进行目标检测,如文献[7],通过对YOLOv8网络模型改进为CDD-YOLO,提升了检测精度。

本文提出一种基于YOLOv11n的优化改进方法,新设计出YOLO11n\_CPM模型算法,对数据直接进行检测。首先,提出多尺度特征融合改进模块CSPP\_SF,通过对SPPF模块进行基于跨阶段局部

网络(CSPNet<sup>[8]</sup>)结构改进,并引入金字塔压缩注意力,可有效建立更长距离的通道依赖关系,抑制背景噪声,实现参数共享提升计算效率和检测性能扩大感受野;其次,通过引入坐标注意力并替换CSPP\_SF模块的SPPF,提出CSPP\_CA模块,并替换YOLOv11n网络结构中的C2PSA模块;最后,引入MPDIoU<sup>[9]</sup>损失函数替换CIoU损失函数。实验结果表明,新提出的YOLO11n\_CPM模型算法在ExDark数据集上的检测性能得到了提升,进一步在LoLI-Street数据集11 000张低光(Dense)图像上验证了本模型算法的有效性及其泛化能力,证明该算法在低光场景下具备更优的检测表现。

## 1 相关工作

### 1.1 卷积神经网络

杨立昆(Yann LeCun)等1998年提出LeNet,形成现代卷积神经网络的雏形,卷积神经网络一般由输入层、交替的卷积层与池化层、全连接层及输出层构成。如图1所示,2012年,多伦多大学的亚历克斯·克里泽夫斯基(Alex Krizhevsky)等采用修正线性单元(Rectified Linear Unit, ReLU)作为激活函数,提出了著名的AlexNet,并取得当年ILSVRC图像分类比赛的冠军,成为深度学习发展史上的重要拐点<sup>[10]</sup>。随后,学术界和工业界通过设计更深的网络层次和更巧妙的创新技术(如残差连接和Inception模块)。深度卷积神经网络实现了快速的发展与结构改进,并相继提出VGGNet、GoogLeNet、ResNet等网络模型<sup>[11]</sup>。

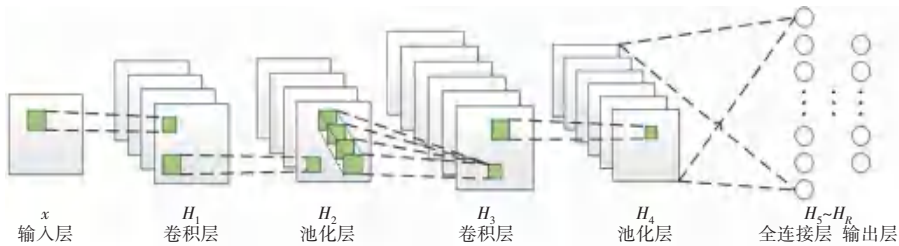


图1 卷积神经网络结构

Fig. 1 Convolutional Neural Network Structure

### 1.2 目标检测

目标检测技术发展大致分为两个阶段,即基于传统算法阶段和基于深度学习算法阶段,在传统算法阶段,检测器完全依靠人设计的先验特征,这类特征难以用一套精准的语言和数学方程定义,因此设计通用且可靠的目标检测器面临较大挑战<sup>[12]</sup>。

2014年,Girshick等<sup>[13]</sup>将基于卷积神经网络特征应用到目标检测场景,提出了R-CNN算法,标志

着目标检测进入基于深度学习算法阶段。此后,目标检测算法的研究,随着深度学习网络的革新实现了飞速发展,相关算法大体分为两种类型:两阶段检测算法和单阶段检测算法。两阶段目标检测算法在检测精度方面表现出色,但计算速度较慢,单阶段检测算法以其更简洁、更具潜力等优势已成为目标检测领域的主流框架<sup>[14]</sup>,其中YOLO系列是应用较广泛、效果较优的目标检测方法。

### 1.3 YOLO 介绍

YOLOv1 (You Only Look Once) 是由 Redmon 等<sup>[15]</sup>于 2016 年采用卷积神经网络实现的端到端目标检测方法,是首个实时目标检测方法,也是最早出现的单阶段目标检测方法。YOLOv1 使用串联的卷积模块组成主干网络,结构简单,检测速度快,泛化能力好,开辟了基于深度学习的目标检测技术路线,

引发了目标检测领域的巨大变革。随后, YOLO 系列模型发布了多个改进版本,包括 YOLOv2 至 YOLOv11,其中, YOLOv11<sup>[16]</sup>是由 Ultralytics 团队于 2024 年 10 月提出,通过引入 C3K2 和 C2PSA 模块,提升了小目标和重叠目标的特征提取能力,但增加了模型的实现复杂度,进而导致训练成本和资源需求显著提升,网络结构如图 2 所示。

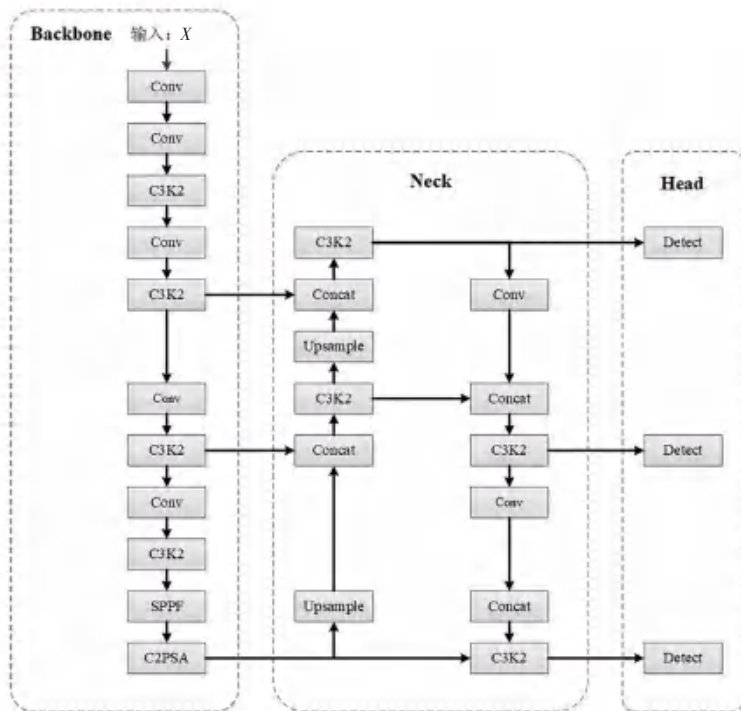


图 2 YOLOv11 网络结构

Fig. 2 YOLOv11 Network Structure

根据网络模型深度与宽度的差异, YOLOv11 分为 5 种不同规格的模型<sup>[17]</sup>, 分别为 YOLOv11n、YOLOv11s、YOLOv11m、YOLOv11l、YOLOv11x。本文选择参数量最少的 YOLOv11n 作为基准模型,通过改进并提出 CSPP\_SF、CSPP\_CA 两个模块,并融合金字塔压缩注意力和坐标注意力机制,进一步引入 MPDIoU 损失函数,旨在探索一种在低光照场景下提升图像检测性能、同时实现更轻量化模型的新方法。

## 2 结构优化

### 2.1 融合注意力

注意力机制起源于对人类视觉机制的研究,最早出现在机器翻译领域,现已成为深度学习的重要理论基础<sup>[18]</sup>。视觉系统会关注图像中有利于判断的信息,并忽略不相关的信息,研究表明,将注意力机制融入卷积神经网络模型可以提升检测性能<sup>[19]</sup>,如文献<sup>[20]</sup>通过引入注意力机制,提升了小目标检

测效果。

#### 2.1.1 坐标注意力

坐标注意力 (Coordinate Attention, CA) 机制<sup>[21]</sup>,是用精确的位置信息对通道关系和远程依赖进行编码,其可以捕获方向感知、位置感知和跨通道的特征信息<sup>[22]</sup>。坐标注意力结构如图 3 所示,可被看作一个计算单元,用于增强网络的特征表达能力,包括坐标信息嵌入和坐标注意力生成两个步骤,假设输入特征张量为  $X = [x_1, x_2, \dots, x_c] \in \mathbb{R}^{C \times H \times W}$  (高度为  $H$ , 宽度为  $W$ , 通道数为  $C$ ), 图中“ $X$  Avg Pool”和“ $Y$  Avg Pool”分别表示一维水平全局池化和垂直全局池化,  $r$  表示缩减因子。

##### (1) 坐标信息嵌入

CA 对输入  $X$  使用池化核的 2 个空间范围 ( $H, 1$ ) 和 ( $1, W$ ), 分别沿水平和垂直方向对每个通道进行编码,池化后的特征图的高度为  $h$  ( $0 \leq h < H$ ), 宽度为  $w$  ( $0 \leq w < W$ ), 因此,第  $c$  ( $0 \leq c < C$ ) 个通道在高度为  $h$  的输出可表示为下式:

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \quad (1)$$

同理, 第  $c$  ( $0 \leq c \leq C$ ) 个通道在宽度为  $w$  的输出可表示为:

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (2)$$

其中,  $x_c(h, i)$  表示输入特征图中通道为  $c$ 、坐标为  $(h, i)$  的分量;  $x_c(j, w)$  表示输入特征图中通道为  $c$ 、坐标为  $(j, w)$  的分量。

上述两个变换分别沿水平和垂直两个空间方向聚合特征, 得到一对方向感知特征图  $z_c^h z_c^w$ 。此外, 这两个转换注意力模块能够捕获沿一个空间方向的长距离依赖关系, 同时保留沿另一个空间方向的精确位置信息, 有助于网络更准确地定位感兴趣目标<sup>[23]</sup>。

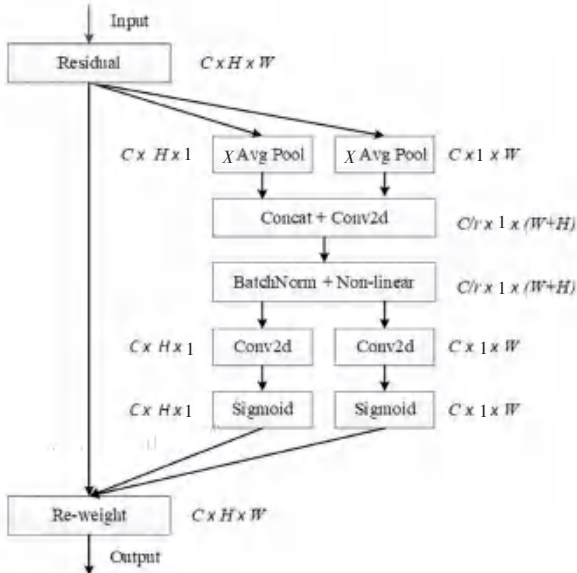


图 3 坐标注意力模块结构

Fig. 3 Structure of the Coordinate Attention Module

(2) 坐标注意力生成

CA 首先对第一步生成的聚合特征图进行级联 ( $z^h$  表示沿  $X$  轴方向平均池化后的输出,  $z^w$  表示沿  $Y$  轴方向平均池化后的输出), 然后将送入共享的  $1 \times 1$  卷积变换函数  $F_1(\cdot)$ , 得到中间特征  $f$  表示如下式:

$$f = \delta(F_1([z^h, z^w])) \quad (3)$$

其中,  $[\cdot, \cdot]$  表示沿空间维度的级联操作;  $\delta(\cdot)$  表示非线性激活函数;  $f \in \mathbb{R}^{C/r \times (H+W)}$  是在水平和垂直方向上编码空间信息的中间特征图;  $r$  为缩减因子。

然后沿空间维度将  $f$  拆分为两个单独的张量  $f^h \in \mathbb{R}^{C/r \times H}$  和  $f^w \in \mathbb{R}^{C/r \times W}$ 。再通过两个独立的  $1 \times 1$  卷积层  $F_h(\cdot)$  和  $F_w(\cdot)$ , 将  $f^h$  和  $f^w$  分别变换为与输入  $X$  具有相同通道数的张量, 经过 Sigmoid 函数激活之后分别得到特征图在高度和宽度方向的注意力权重  $g^h$  和  $g^w$ :

$$g^h = \sigma(F_h(f^h)) \quad (4)$$

$$g^w = \sigma(F_w(f^w)) \quad (5)$$

接着分别对  $g^h$  和  $g^w$  进行扩展并用作注意力权重, 通过下式计算得到输出特征图:

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (6)$$

由上述过程可得出, CA 注意力具备三方面优势。首先, 实现简单; 其次, 可以利用位置信息, 准确地定位感兴趣的区域; 最后, 可以有效地捕获通道间的关系。

2.1.2 金字塔压缩注意力

金字塔压缩注意力 (Pyramid Squeeze Attention, PSA)<sup>[24]</sup> 是一种高效的多尺度通道注意力, 可有效提取和利用不同尺度特征图的空间信息, 同时可以建立更长距离的通道依赖关系<sup>[25]</sup>。本文假设输入特征图为  $X, S=4, G=4$ , PSA 结构如图 4 所示, 实现过程主要分为 4 个步骤:

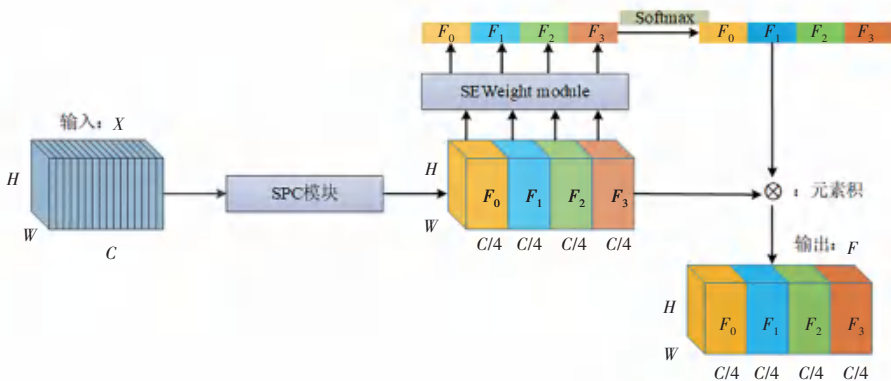


图 4 PSA 结构

Fig. 4 Structure of the PSA

(1)通过,SPC 模块,如图 5 所示,获取输入特征图  $X$  的不同尺度特征图,并进行拼接生成特征图  $F'$ 。

(2)生成通道注意力向量,通过 SEWeight 模块提取特征图  $F'$  的注意力向量。

(3)权重标定(校准),通过 Softmax 重新校准通道注意力向量,获得多尺度通道的重新校准权重  $W'$ 。

(4)将特征图  $F'$  与重新校准权重  $W'$  进行元素积操作。

最后,可以获得具有更丰富信息的多尺度特征图  $F$ 。

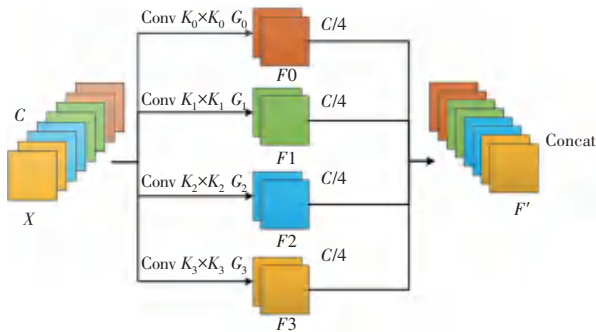


图 5 SPC 模块结构

Fig. 5 Structure of the SPC module

## 2.2 提出新模块

### 2.2.1 CSPP\_SF 模块

快速空间金字塔池化(Spatial Pyramid Pooling-Fast, SPPF)是 YOLOv11 中的模块,结构如图 6 所示,通过池化操作将多个不同尺度的特征图进行融合,生成具有丰富上下文信息的特征图<sup>[26]</sup>。

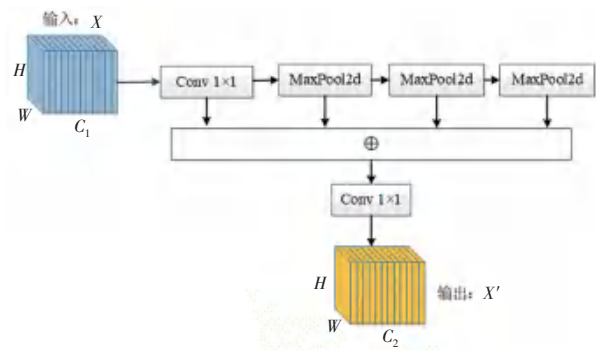


图 6 SPPF 模块结构

Fig. 6 Structure of the SPPF module

本文在 CSPNet 思想基础上,引入金字塔压缩注意力(Pyramid Squeeze Attention, PSA)对 SPPF 进行改进,提出 CSPP\_SF 模块,如图 7 所示,该模块将输入特征图  $X$  分为两部分进行计算,一部分由残差模块和 SPPF 进行运算处理,另一部分做恒等映射,然后将两部分结果拼接实现信息互补,最后通过金字塔压缩注意力提取空间信息,建立更长距离的通道依赖关系,抑制背景噪声,提升检测效果,具有参

数共享、提升计算效率、扩大感受野等优点。

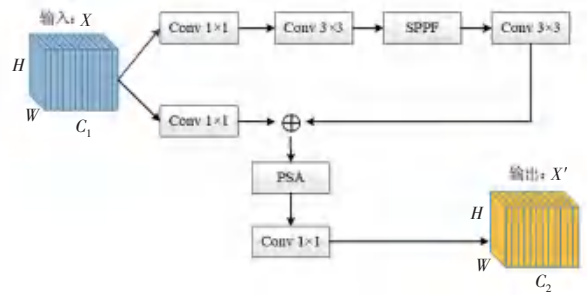


图 7 CSPP\_SF 模块结构图

Fig. 7 Structure of the CSPP\_SF Module

### 2.2.2 CSPP\_CA 模块

YOLOv11 中 C2PSA 模块如图 8 所示,通过引入多头自注意力机制获取长距离依赖,具有并行计算、动态权重等优点,然而多头自注意力也存在计算量大、参数量多的缺点。与 CSPP\_SF 模块同理,提出 CSPP\_CA 模块,如图 9 所示,该模块通过引入坐标注意力并替换多头自注意力,可以获取方向感知、位置感知和跨通道信息,减少计算量,达到更加精准地识别图像中感兴趣的内容,最后通过 PSA 注意力提取空间信息,建立更长距离的通道依赖关系,抑制背景噪声,扩大感受野,从而提升检测效果增强逻辑递进关系。

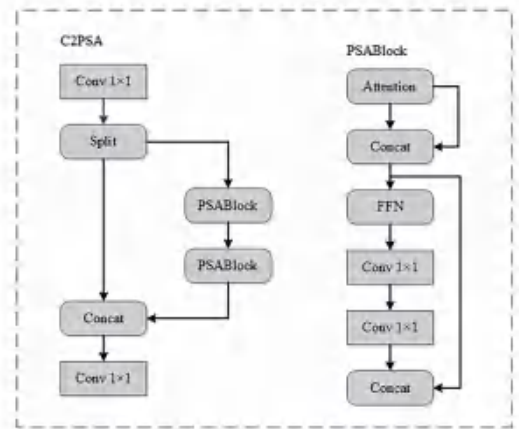


图 8 C2PSA 模块结构

Fig. 8 Structure of the C2PSA Module

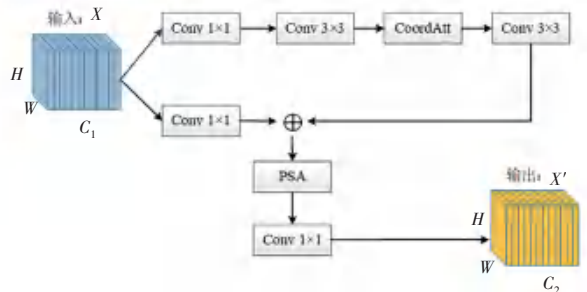


图 9 CSPP\_CA 模块结构

Fig. 9 Structure of CSPP\_CA Module

### 2.2.3 网络结构

通过将 CSPP\_SF 模块和 CSPP\_CA 模块分别替

换 YOLOv11n 网络中的 SPPF 和 C2PSA 后,得到新的网络模型 YOLOv11n\_CPM,网络结构如图 10 所示。

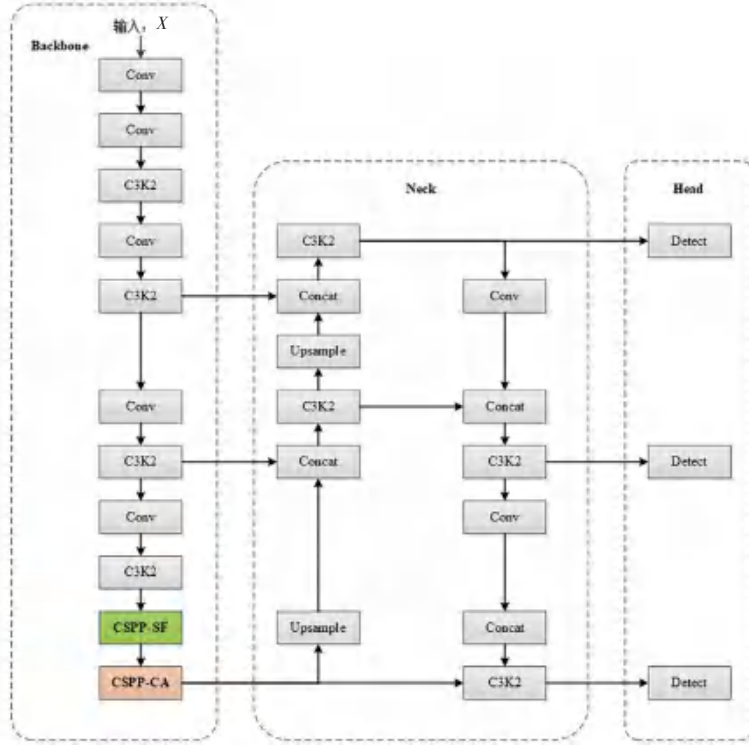


图 10 YOLOv11n\_CPM 模型网络结构

Fig. 10 Network Structure of YOLOv11n\_CPM Model

### 2.3 损失函数

YOLOv11 使用 CIoU<sup>[27]</sup> 作为网络的边界框损失函数,计算公式如下式:

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (7)$$

其中,  $\rho^2(b, b^{gt})$  为预测框  $b$  与真实框  $b^{gt}$  中心点的欧氏距离平方。

CIoU 对于小目标的检测效果较差,MPDIoU 基于几何矩形特性,直接计算预测框与真实框顶点的距离,不仅具有 CIoU 的核心特性,而且解决了小目标定位不足的问题,整体提升了模型的收敛速度和检测精度<sup>[28]</sup>。

MPDIoU 计算公式如下,假设真实框的左上角坐标为  $(x_1^{gt}, y_1^{gt})$ ,右下角坐标为  $(x_2^{gt}, y_2^{gt})$ ,预测框左上角坐标为  $(x_1^{pr}, y_1^{pr})$ ,右下角坐标为  $(x_2^{pr}, y_2^{pr})$ ,  $d_1, d_2$  表示左上角和右下角之间的欧几里德距离,如图 11 所示,变量  $w$  和  $h$  表示图像的宽和高。

$$d_1^2 = (x_1^{pr} - x_1^{gt})^2 + (y_1^{pr} - y_1^{gt})^2 \quad (8)$$

$$d_2^2 = (x_2^{pr} - x_2^{gt})^2 + (y_2^{pr} - y_2^{gt})^2 \quad (9)$$

$$MPDIoU = IoU - \frac{d_1^2}{w^2 + h^2} - \frac{d_2^2}{w^2 + h^2} \quad (10)$$

当使用 MPDIoU 作为损失函数时,计算公式如式:

$$L_{MPDIoU} = 1 - MPDIoU \quad (11)$$

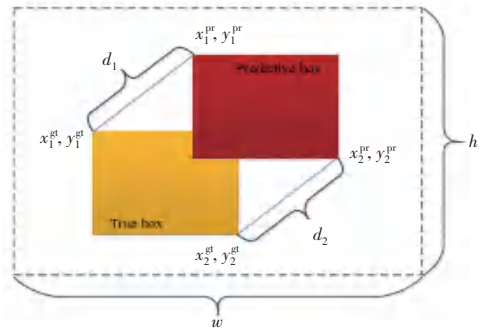


图 11 MPDIoU 损失函数计算示意图

Fig. 11 Schematic calculation illustration on MPDIoU loss function

## 3 实验结果分析

### 3.1 实验环境

本实验在以下硬件与软件环境中完成训练与测试,具体配置见表 1。

表 1 模型训练环境及配置

Table 1 Environment and configuration of model training

参数	实验环境
操作系统	Windows10
CPU	Intel (R) Core(TM) i7-9700
GPU	NVIDIA Quadro P2200
软件	Python3.10
深度学习框架	PyTorch2.0.0

### 3.2 评价指标

为了验证本研究改进算法的性能,采用模型大小、参数量、mAP@0.5%、mAP@0.5:0.95%、GFLOPs作为模型性能评价指标。mAP (mean Average Precision)计算公式如下式:

$$mAP = \frac{\sum AP_c}{N} \quad (12)$$

其中,mAP表示目标检测中所有种类AP的平



图12 ExDark数据集示例图像

Fig. 12 Sample images of the ExDark dataset

LoLI-Street<sup>[30]</sup>数据集包含3种低光强度水平,总计3.3万张图像,为进一步验证本文提出的YOLOv11n\_CPM模型算法效果,引入LoLI-Street数据集中低光场景下的11000张图像,并按9:1比例分为训练集和验证集。

### 3.4 实验分析

#### 3.4.1 消融实验

YOLOv11n\_CPM模型算法在ExDark数据集进行80轮训练实验及消融实验,结果见表2,从结果可以看出,YOLOv11n\_CPM模型算法实现了检测性

均值;N表示种类数量;C表示类别。

### 3.3 数据集

ExDark<sup>[29]</sup>是一个公开数据集,包括12种类别,共计7363张低光照场景图片,该数据集专门用于低光场景下的目标检测和图像增强研究。为验证改进算法的实际效果,本实验将ExDark数据集按9:1比例分为训练集和验证集,其中部分图像如图12所示。

能提升与轻量化,整体mAP@0.5达到77.1%、mAP@0.5:0.95达到49.8%,模型大小降低了约40.6%,模型参数量降低了约41.7%。

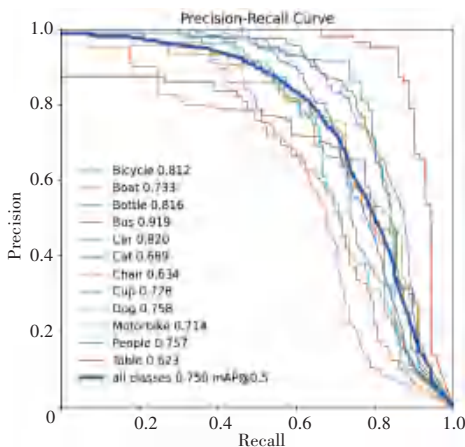
YOLOv11n模型与YOLOv11n\_CPM模型在ExDark数据集训练的PR曲线对比,如图13所示,其中图13(a)为YOLOv11n模型PR曲线、图13(b)为YOLOv11n\_CPM模型PR曲线。

YOLOv11n\_CPM模型算法在ExDark数据集训练检测示例如图14所示。

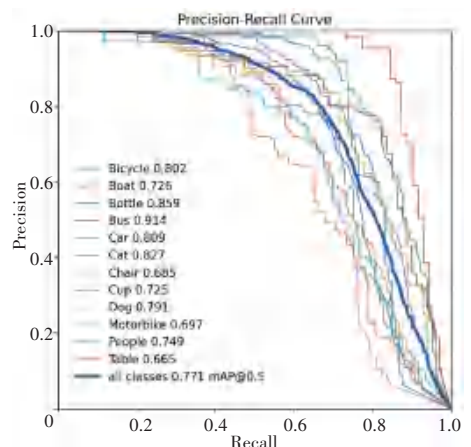
表2 不同模型算法在ExDark数据集结果对比

Table 2 Performance comparison of different model algorithms on the ExDark dataset

算法	模型大小(MB)	Parameters	mAP@0.5/%	mAP@0.5:0.95/%	GFLOPs
YOLOv11n	14.10	7 276 804	75.0	48.6	10.2
YOLOv11n+MPDIoU	14.10	7 276 804	76.2	49.8	10.2
YOLOv11n+CSPP_SF	15.50	8 014 792	76.1	49.2	10.8
YOLOv11n+CSPP_CA	6.95	3 503 456	74.6	48.3	7.1
本文方法	<b>8.37</b>	<b>4 241 444</b>	<b>77.1</b>	<b>49.8</b>	<b>7.8</b>



(a) YOLOv11n模型PR曲线



(b) YOLOv11n\_CPM模型PR曲线

图13 PR曲线对比图

Fig. 13 Comparison of PR curves



图 14 训练与检测示例图

Fig. 14 Example diagram of training and detection

### 3.4.2 验证实验

为进一步验证 YOLOv11n\_CPM 模型算法的有效性及其泛化能力,在 LoLI-Street 数据集中选取 11 000 张低光(Dense)图像进行 80 轮训练,实验结

果见表 3,其中 mAP@0.5 和 mAP@0.5:0.95 分别达到 46.7%和 33.4%,较原始 YOLOv11n 算法分别提高了 4.9 和 1.8 个百分点,模型大小降低了约 40.5%,模型参数量降低了约 41.5%。

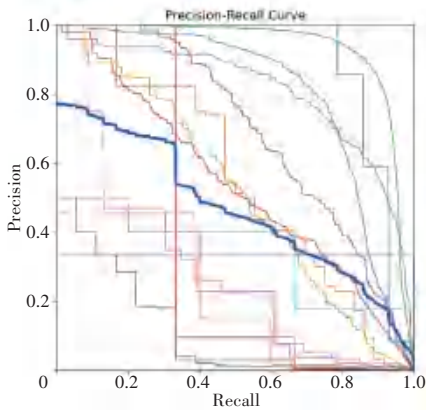
表 3 不同模型算法在 LoLI-Street 子数据集上的性能对比

Table 3 Performance comparison of different model algorithms on the LoLI-Street subset

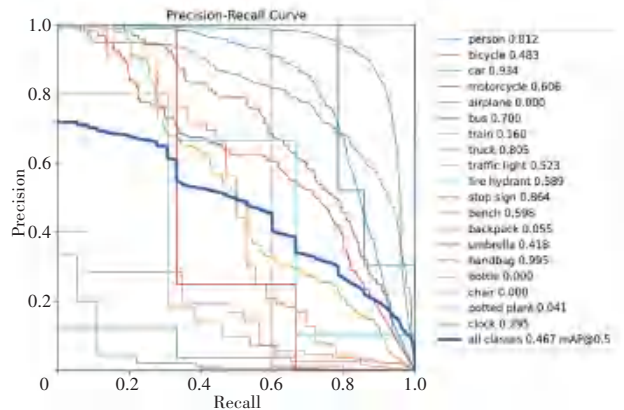
算法	模型大小/MB	Parameters	mAP@0.5/%	mAP@0.5:0.95/%	GFLOPs
YOLOv11n	14.20	7 308 560	41.8	31.6	10.3
本文方法	<b>8.45</b>	<b>4 273 200</b>	<b>46.7</b>	<b>33.4</b>	<b>7.9</b>

YOLOv11n 模型与 YOLOv11n\_CPM 模型在 LoLI-Street 子数据集上训练得到的 PR 曲线对比,

如图 15 所示,其中,图 15(a)为 YOLOv11n 模型 PR 曲线、图 15(b)为 YOLOv11n\_CPM 模型 PR 曲线。



(a) YOLOv11n 模型 PR 曲线



(b) YOLOv11n\_CPM 模型 PR 曲线

图 15 PR 曲线对比图

Fig. 15 Comparison of PR Ccurves

## 4 结束语

为解决低光照场景目标检测精度低、效果差问题,本文提出一种基于 YOLOv11n 的低光照场景目

标检测模型 YOLOv11n\_CPM。首先,设计多尺度特征融合改进模块 CSPP\_SF,通过对 SPPF 模块进行基于跨阶段局部网络(CSPNet)结构进行改进,并引入金字塔压缩注意力,可有效建立更长距离的通道

依赖关系,抑制背景噪声,实现参数共享,提升计算效率和检测性能,扩大感受野;其次,通过引入坐标注意力并结合 CSPP\_SF 模块原理,提出 CSPP\_CA 模块,可有效捕获方向位置和跨通道信息,减少计算量和参数量;最后,引入 MPDIoU 损失函数,基于几何矩形特性,不仅简化计算过程,而且改进了小目标定位不足的问题,整体提升了模型的收敛速度和检测精度。实验结果表明,YOLOv11n\_CPM 模型提升了低光照场景目标检测性能,在 ExDark 数据集上的 mAP@0.5 和 mAP@0.5:0.95 分别达到 77.1% 和 49.8%,较原始 YOLOv11n 算法分别提高了 2.1 和 1.2 个百分点,模型大小降低了约 40.6%,模型参数量降低了约 41.7%。为进一步验证本模型算法有效性及泛化能力,在 LoLI-Street 数据集 11 000 张低光(Dense)图像上 mAP@0.5 和 mAP@0.5:0.95 分别达到 46.7% 和 33.4%,较原始 YOLOv11n 算法分别提高了 4.9 和 1.8 个百分点,模型大小降低了约 40.5%,模型参数量降低了约 41.5%。后续研究可探索在 YOLOv11 其他尺寸模型上的效果验证及改进,以及多模态内容目标检测。

## 参考文献

[1] 赵康伦. 可见光与毫米波深度融合的暗光环境下目标检测算法的研究与实现[D]. 北京:北京邮电大学, 2024.

[2] 王思平. 面向暗光场景的目标增强和检测算法研究[D]. 重庆:重庆理工大学, 2024.

[3] 王昱婷, 刘志明, 万亚平, 等. 基于可见光与红外图像的弱光条件下目标检测[J]. 计算机工程, 2024, 50(8):270-281.

[4] 官敬超, 刘家辉, 赵庆辉, 等. RFC-YOLOv8:可见光和红外图像自适应融合的目标检测[J]. 光电与控制, 2025, 32(9):54-60.

[5] 徐轩, 姚斌, 韩典芝. 基于 Retinex 理论和残差网络的低照度图像增强方法[J]. 智能计算机与应用, 2025, 15(3):164-169. DOI:10.20169/j.issn.2095-2163.250324.

[6] 蒋占军, 吴佰靖, 马龙, 等. 基于金字塔边缘增强的自矫正低光照目标检测[J]. 光学精密工程, 2024, 32(20):3099-3111.

[7] 史丽晨, 杨超, 刘雪超, 等. 基于 CDD-YOLO 的轻量级低光照目标检测算法[J]. 计算机工程与应用, 2025, 61(6):106-117.

[8] WANG Chenyao, LIAO Hongyuan, WU Yuehua, et al. CSPNet: A new backbone that can enhance learning capability of CNN[J]. arXiv, 1911.11929, 2019. DOI:10.48550/arXiv.1911.11929.

[9] 侯阳, 张琼, 赵紫焯, 等. 基于 YOLOv5s 的复杂场景下高效烟火检测算法 YOLOv5s-MRD[J]. 计算机应用, 2025, 45(4):1317-1324.

[10] 李玉鑑. 深度学习:卷积神经网络从入门到精通[M]. 北京:机械工业出版社, 2018:126.

[11] 丛润民, 李锋, 张伟. 人工智能通识[M]. 北京:机械工业出版社, 2025:410.

[12] 涂铭, 金智勇. 深度学习与目标检测:工具、原理与算法[M]. 北京:机械工业出版社, 2021:20.

[13] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2014: 580-587. DOI:10.1109/CVPR.2014.81.

[14] 杨崇尚. 基于多尺度特征融合的目标检测算法研究[D]. 成都:电子科技大学, 2024.

[15] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection [C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2016: 779-788. DOI:10.1109/CVPR.2016.91.

[16] 王宁, 智敏. 深度学习下的单阶段通用目标检测算法研究综述 [J]. 计算机科学与探索, 2025, 19(5):1115-1140.

[17] 张志豪, 厉小润, 陈淑涵. 基于改进 YOLO11 的无人机航拍图像小目标检测算法[J]. 液晶与显示, 2025, 40(6):915-930.

[18] 王冲, 魏子令, 陈曙晖. 基于自注意力机制的无边界应用动作识别方法[J]. 计算机研究与发展, 2022, 59(5):1092-1104.

[19] 张宸嘉, 朱磊, 俞璐. 卷积神经网络中的注意力机制综述[J]. 计算机工程与应用, 2021, 57(20):64-72.

[20] 羊森海, 陈丹. 基于改进 YOLOv7 的抓取图像小目标检测算法 [J]. 智能计算机与应用, 2025, 15(7):99-103. DOI:10.20169/j.issn.2095-2163.250714.

[21] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design [C]//Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2021: 13708-13717. DOI:10.1109/CVPR46437.2021.01350.

[22] 彭晏飞, 孟欣, 李泳欣, 等. 结合坐标注意力与生成式对抗网络的图像超分辨率重建[J]. 计算机工程与科学, 2024, 46(1):122-131.

[23] 李龙岩. 基于视觉的机械臂智能抓取技术研究[D]. 郑州:华北水利水电大学, 2023.

[24] ZHANG H, ZU K, LU J, et al. EPSANet: An efficient pyramid split attention block on convolutional neural network [C]// Proceedings of the Asian Conference on Computer Vision. Cham: Springer, 2023:541-557. DOI:10.1007/978-3-031-26313-2\_33.

[25] 赵东暄, 段喜萍. 基于 EPSANet 的肺部医疗图像识别[J]. 长江信息通信, 2023, 36(10):43-47.

[26] 靳宇风. 基于 YOLO-CLS 算法的经济林病虫害识别及系统设计[D]. 北京:中国地质大学(北京), 2024.

[27] ZHENG Zhaohui, WANG Ping, LIU Wei, et al. Distance-IoU Loss: Aster and better learning for bounding box regression [J]. arXiv, 1911.08287, 2019. DOI:10.48550/arXiv.1911.08287.

[28] 孙小栋, 朱启兵, 徐华伟, 等. 用于超纤革表面瑕疵识别的 MFL-YOLOv8 算法[J]. 光学精密工程, 2025, 33(2):311-323.

[29] LOH Y, CHAN C. Getting to know low-light images with the Exclusively Dark dataset [J]. Computer Vision and Image Understanding, 2018, 178:30-42. DOI:10.1016/j.cviu.2018.10.010.

[30] ISLAM M T, ALAM I, WOO S S, et al. LoLI-Street: Benchmarking low-light image enhancement and beyond [C]// Proceedings of the Asian Conference on Computer Vision. Cham: Springer, 2025:348-365. DOI:10.1007/978-981-96-0917-8\_20.