

# 基于深度学习的遥感旋转图像检测

周生翠

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2024年4月27日; 录用日期: 2024年5月23日; 发布日期: 2024年5月31日

## 摘要

旋转目标检测是目前遥感图像目标检测的重要任务之一。遥感图像拥有众多小目标, 检测目标密集, 目标方向任意等特点。针对上述问题, 提出了一种轻量级无锚框旋转目标检测算法YOLOv8-LR。首先, 在主干网络和特征融合之间设计了轻量级通道空间特征增强模块, 增加特征图的通道信息, 提升了对遥感图像的小目标预测能力; 其次, 轻量化特征融合网络, 在不降低检测准确率的情况下, 减少模型计算复杂度; 最后, 增加网络的输出维度, 将角度信息引入损失函数, 使模型具备旋转框检测条件。实验结果表明, 在DOTA数据集上, 该方法检测精度高达77.2%。该模型具有良好的目标检测性能。

## 关键词

遥感图像, 旋转目标检测, 轻量级, 注意力机制, 卷积神经网络

# Remote Sensing Rotation Image Detection Based on Deep Learning

Shengcui Zhou

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 27<sup>th</sup>, 2024; accepted: May. 23<sup>rd</sup>, 2024; published: May. 31<sup>st</sup>, 2024

## Abstract

Rotated object detection is currently one of the important tasks for remote sensing image object detection. Remote sensing images have characteristics such as inconsistent scales, dense detection targets, and arbitrary target directions. In response to the above issues, a lightweight anchor-free rotated object detection algorithm named YOLOv8-LR is proposed. Firstly, a lightweight channel-spatial feature enhancement module is designed between the backbone network and feature fusion to increase the channel information of the feature map and improve the prediction ability

of small targets in remote sensing images. Secondly, the lightweight feature fusion network is used to reduce the computational complexity of the model without reducing the detection accuracy. Finally, the output dimension of the network is increased, and angle information is introduced into the loss function, so that the model has the conditions for rotated bounding box detection. Experimental results show that the detection accuracy of this method reaches 77.2% on the DOTA datasets, respectively. The model has good object detection performance.

## Keywords

Remote Sensing Image, Rotated Object Detection, Lightweight, Attention Mechanism, Convolutional Neural Network

Copyright © 2024 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

遥感图像中任意方向目标的检测对于遥感图像的理解具有重要意义，已受到越来越多的关注。受益于水平目标检测的快速发展，越来越多的旋转目标检测器逐渐涌现，如 RoI Transformer [1]、CFA [2]、S<sup>2</sup>ANet [3]、ReDet [4]、SASM [5]、DAFNe [6]、Oriented RepPoints [7]、FCOSR [8]、DAL [9]、Oriented R-CNN [10]、R<sup>3</sup>Det [11]，其主要设计思路是由相应的水平目标检测器衍生而来，如 FCOS [12]、Faster R-CNN [13]、Focal Loss [14]、RepPoints [15]。YOLOv8 [16]是最新提出的 YOLO 系列算法模型，其首次采用无锚框检测，使得模型可以减小使用锚点带来的计算负担，从而进一步推动检测向实时、高精度方向发展。YOLOv8 在遥感图像水平目标检测中可以达到很高的准确度。因此决定采用 YOLOv8 为设计基线，提出选择目标检测算法。

由于遥感图形中物体的尺度和方向存在巨大变化，旋转物体检测仍然具有挑战性。首先由于遥感图像大都具有高分辨率，且同时存在大、中、小不同尺度的目标，导致许多算法难以有效同时识别不同尺度目标，尤其是小目标，会出现漏检、错检情况。其次，旋转框的表示方法仍然存在一些理论问题。在旋转框目标检测中，目标框表示方法大致可以分为三种：OpenCV 表示法、长边表示法和八参数表示法，如图 1 所示。目前，多数旋转框目标检测采用 OpenCV 表示法，使用这种表示方法时会遇到边界不连续问题[17] [18] [19]。边界不连续问题主要是由角度的周期性和边的可交换性引起的，而后者与特定的旋转框定义(如长边定义)有关。

同时实验中发现，虽然 YOLOv8 设计了轻量化网络，但由于旋转目标检测将角度信息引入，模型的计算复杂度仍然较高，内存占用量较大。这也使得模型对训练资源要求更高。

针对以上问题，一种轻量级无锚框旋转目标检测算法 YOLOv8-LR 本文主要贡献如下：首先在主干网络和特征融合之间设计了轻量级通道空间特征增强模块，增加特征图的通道信息，提升了对遥感图像的多尺度预测能力；其次，轻量化特征融合网络，在不降低检测准确率的情况下，减少模型计算复杂度；然后，增加网络的输出维度，采用无锚框 OpenCV 回归方法对图像进行预测和检测，以适用于旋转目标检测；最后改进 DFL 损失[20]，学习锚框到边框的一般分布，以获得更准确的边框尺度预测。并引入了类似于[9]的 ProbIoU 损失[21]作为回归损失，以避免在预测角度时出现边界不连续问题。

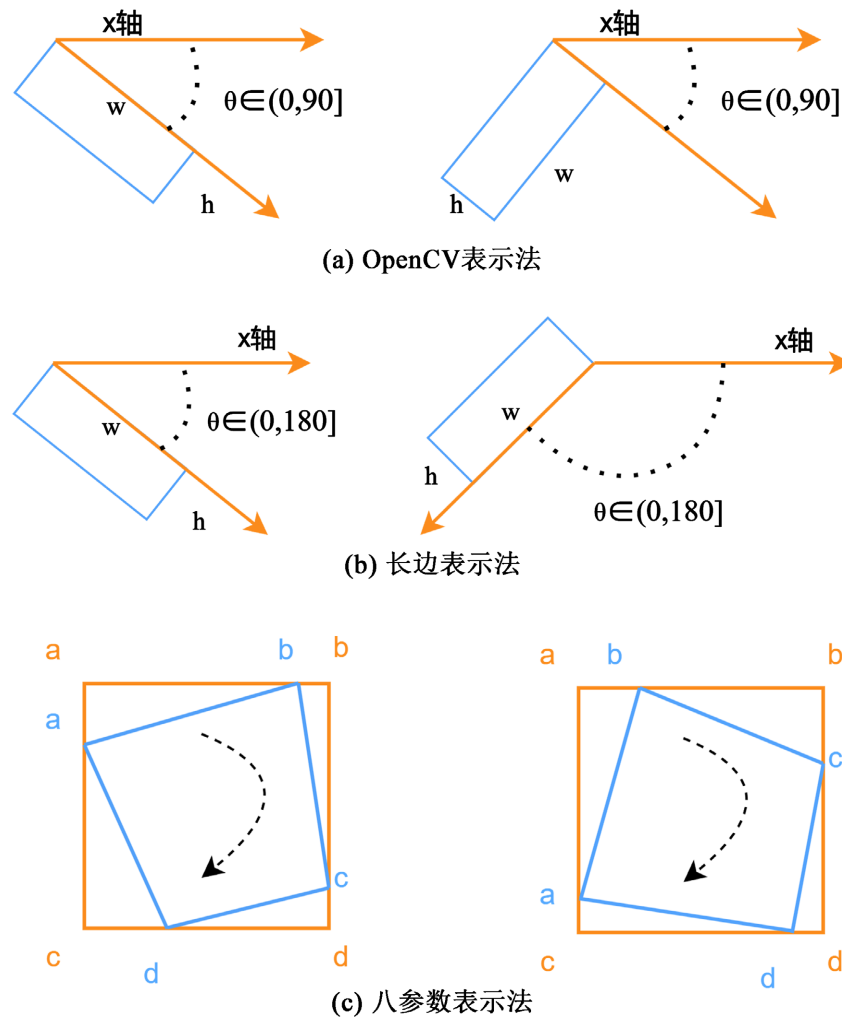


Figure 1. Rotate box definition method  
图 1. 旋转框定义法

## 2. 仿真模型

### 2.1. YOLOv8-LR 整体结构

YOLOv8-LR 算法以 YOLOv8 目标检测算法作为基线结构进行设计，如图 2 所示。相较于 YOLOv8 算法，YOLOv8-LR 算法在通道空间特征增强模块、特征融合网络和损失函数三个方面进行了改进，改进部分在图中用红框标出。

### 2.2. 通道空间特征增强模块

YOLOv8 的特征融合网络结合了特征金字塔(Feature pyramid network, FPN) [22]和路径聚合网络(Path aggregation network, PAN) [23]在一定程度上缓解了信息融合问题。然而，YOLOv8 的骨干网络最后一层快速空间金字塔池化(Spatial Pyramid Pooling-Fast, SPPF) [24]，对不同尺度的特征融合时，采用  $1 \times 1$  卷积来恢复通道数，使其保证输入输出 SPPF 的特征图尺度大小和深度不变。虽然使用 SPPF 提高了模型在处理不同尺寸输入时的速度，但通道信息减少了，导致 FPN 和 PANet 性能不佳。因此，在骨干网络和特征融合网络之间，设计了轻量级通道空间特征增强模块，来改善和共享特征图的通道信息，从而减少了 FPN

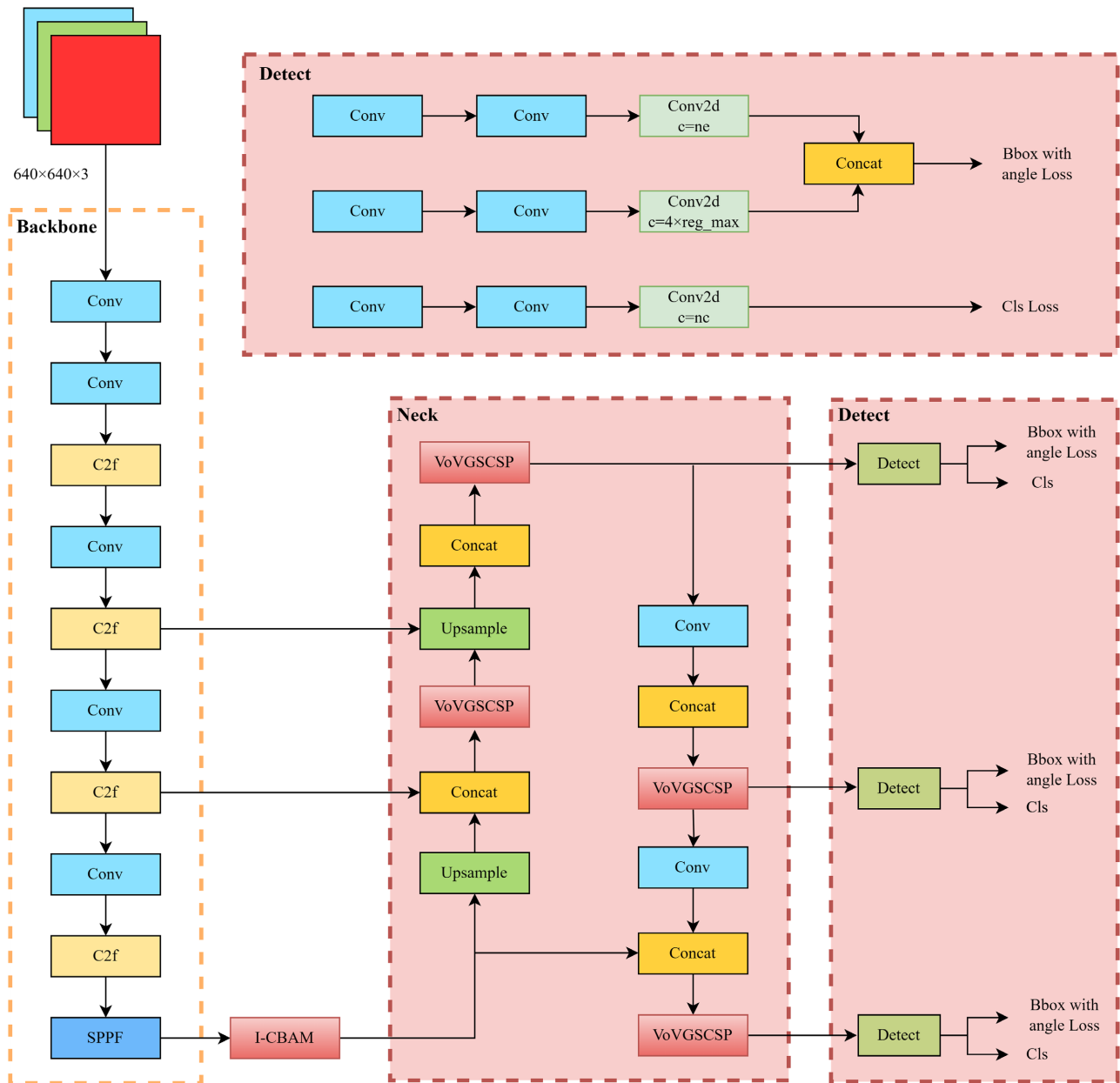


Figure 2. The overall structure of the YOLOv8-LR

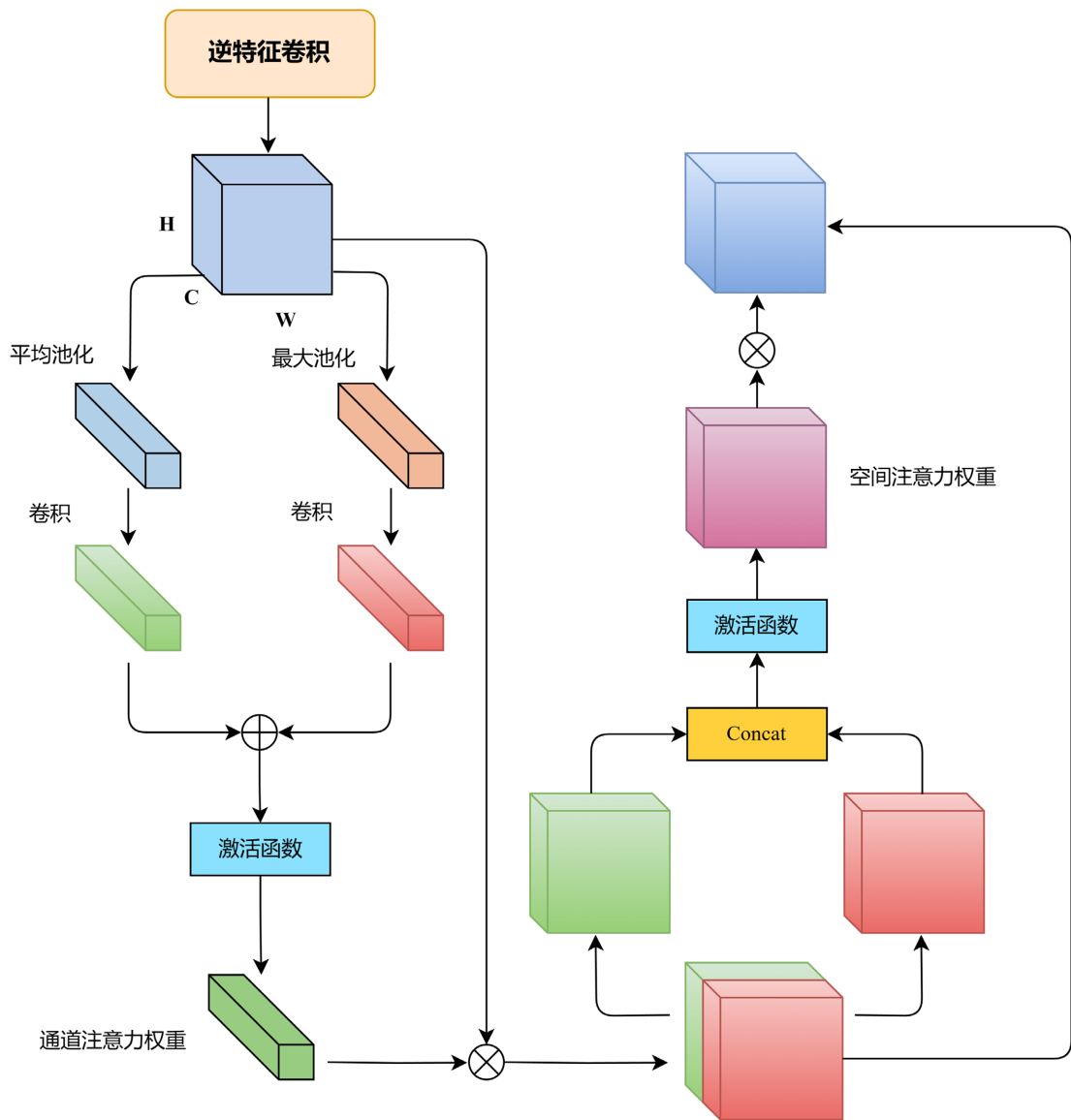
图 2. YOLOv8-LR 的整体结构

初始阶段的信息丢失，同时，还强调了通道和空间域中的重要信息。因此，这种改进有助于增强 FPN 的性能，特别是有利于检测较小尺寸的物体。其网络架构如图 3 所示。

具体来说，首先，将经过  $1 \times 1$  卷积处理后的特征  $X$ ，输入逆特征卷积(Involution) [25]。逆特征卷积在空间域和通道域中对输出特征进行逆变换，恢复原始图像的细节。其定义为：

$$\mathcal{H} \in \mathbb{R}^{H \times W \times K \times K \times G} \quad (1)$$

其中  $H$  和  $W$  为输入特征的高度和宽度； $K$  为卷积核大小； $G$  为使用相同逆特征卷积核的特征组的数量。对于坐标为  $(i, j)$  像素  $X_{i,j} \in \mathbb{R}^C$ ，使用核生成函数  $\Phi$ ，将每个像素位置  $(i, j)$  处的函数映射抽象，由通道



**Figure 3.** Lightweight channel spatial feature enhancement module structure  
**图 3.** 轻量级通道空间特征增强模块结构

排列方式重新排布成以空间排布方式，得到其对应的逆特征卷积核，记作  $\mathcal{H}_{i,j,\dots,g} \in \mathbb{R}^{K \times K}, g = 1, 2, \dots, G$ ，且该核在通道之间共享：

$$\mathcal{H}_{i,j} = \Phi(X_{i,j}) = W_1 \sigma(W_0 X_{i,j}) \tag{2}$$

其中， $W_0 \in \mathbb{R}^{\frac{C}{r} \times \frac{C}{r}}$ ； $W_1 \in \mathbb{R}^{(K \times K \times G) \times \frac{C}{r}}$ ； $\sigma$  为归一化函数。利用逆特征卷积核对输入进行乘加运算其最后的输出表示为：

$$Y_{i,j,k} = \sum_{(u,v) \in \Delta_K} \mathcal{H}_{i,j,k} + \lfloor K/2 \rfloor, v + \lfloor K/2 \rfloor, \lceil kG/C \rceil X_{i+u, j+v, k} \tag{3}$$

其中  $u, v$  为偏移量， $\Delta_K$  为考虑对中心像素进行卷积的邻域偏移量集合。

图 4 以  $G = 1$  为例，输入特征图为  $X$ ，对每个像素位置  $(i, j)$  使用核生成函数  $\Phi$ ，进行通道到空间

的重新排列得到  $\mathcal{H}_{i,j} \in \mathbb{R}^{K \times K \times 1}$ ，然后将  $C$  个通道的张量分别与内卷核相乘，再在  $K \times K$  空间邻域内聚合的求和。

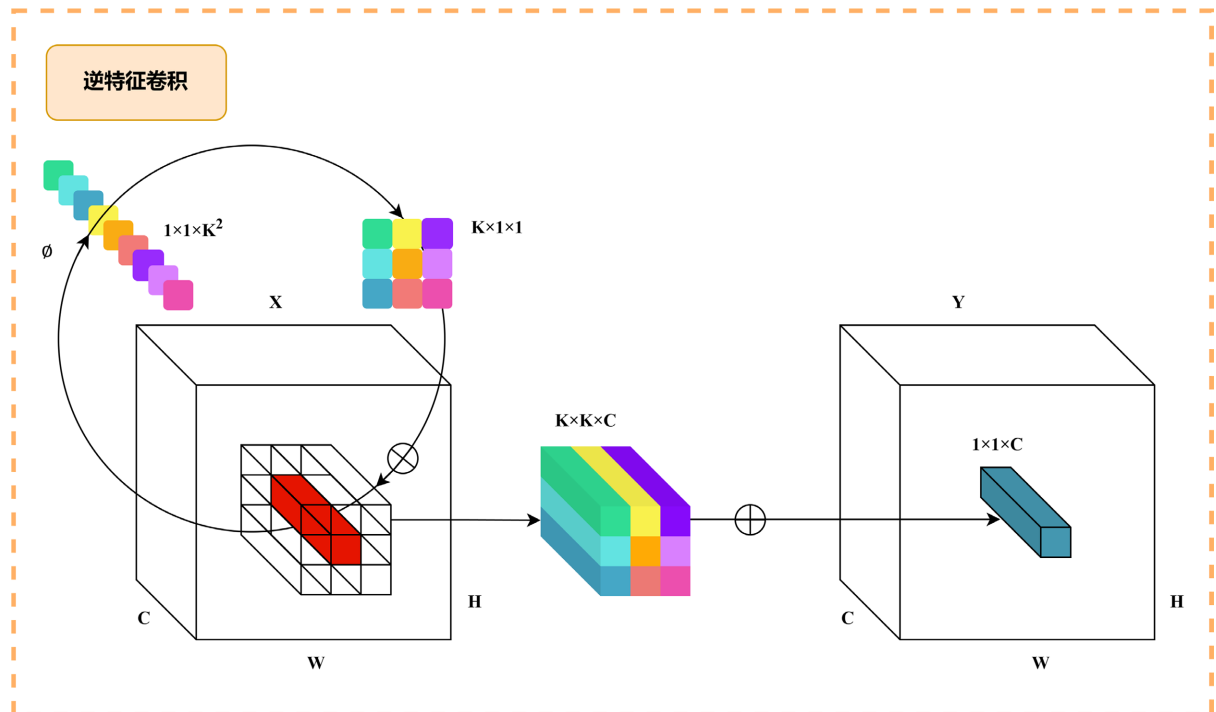


Figure 4. Involution of  $G = 1$   
图 4.  $G = 1$  的逆特征卷积

将补充信息后的特征输入通道与空间混合注意力(channel and spatial mixed attention, CBAM) [26]模块, 利用其双分支进行特征提取和融合, 自适应地将特征进行细化, 选择重要的通道和空间区域。从而突出重要特征, 获取更加丰富的特征信息。

CBAM 由 2 个模块组成: 通道注意力模块和空间注意力模块。这两个模块分别生成通道和空间注意力图, 然后将其与输入特征图相乘, 以便于自适应特征细化。因此, 沿通道和空间轴沿着有意义的特征被强调, 而冗余的特征被抑制。通道注意模块在不同通道上对特征图执行全局最大池化和平均池化, 然后执行逐元素求和, 并且使用 sigmoid 激活。空间注意力模块对不同特征图上相同位置的像素值执行全局最大池化和平均池化, 然后连接两个特征图, 然后进行 Conv2d 操作和 sigmoid 激活。

### 2.3. 轻量化特征融合网络

特征融合网络 FPN 和 PANet 使用的跨阶段部分网络融合模块(Cross Stage Partial Network Fusion, C2f), 需要对多个层级的特征进行融合, 这会增加计算量和内存消耗。特别是在处理大尺寸图像时, 计算复杂度会更高。因此, 引入了一种新的轻量级卷积方法, VoV-GSCSP [27], 设计一种轻量级的特征融合网络。该方法使得卷积计算的输出尽可能接近通道密集卷积的输出, 降低了计算成本, 减轻模型负担同时保持准确性。使模型在准确性和速度之间达到了出色的权衡。其网络架构如图 5 所示。

为了最终加快预测的计算速度, CNN 中的输入图像几乎总是必须在骨干网络中经历类似的转换过程: 空间信息逐步向通道传输。而每次特征图的空间(宽度和高度)压缩和通道扩展都会造成部分语义信息的丢失。通道密集卷积计算最大限度地保留了每个通道之间的隐藏连接, 但通道稀疏卷积完全切断了这些连

接。GSConv 以较低的时间复杂度尽可能多地保留这些连接。通常,卷积计算的时间复杂度由 FLOP 定义。因此, SC(通道密集卷积)、DSC(通道稀疏卷积)和 GSConv 的时间复杂度(无偏差)为:

$$\begin{aligned}
 Time_{SC} &\sim O(W \cdot H \cdot K_1 \cdot K_2 \cdot C_1 \cdot C_2) \\
 Time_{DSC} &\sim O(W \cdot H \cdot K_1 \cdot K_2 \cdot 1 \cdot C_2) \\
 Time_{GSConv} &\sim O\left(W \cdot H \cdot K_1 \cdot K_2 \cdot \frac{C_2}{2}(C_1 + 1)\right)
 \end{aligned}
 \tag{4}$$

其中,  $W$  为输出特征图的宽度;  $H$  为输出特征图的高度;  $K_1 \cdot K_2$  为卷积核的大小;  $C_1$  为每个卷积核的通道数; 也是输入特征图的通道数;  $C_2$  为输出特征图的通道数。

Light FPN PANet

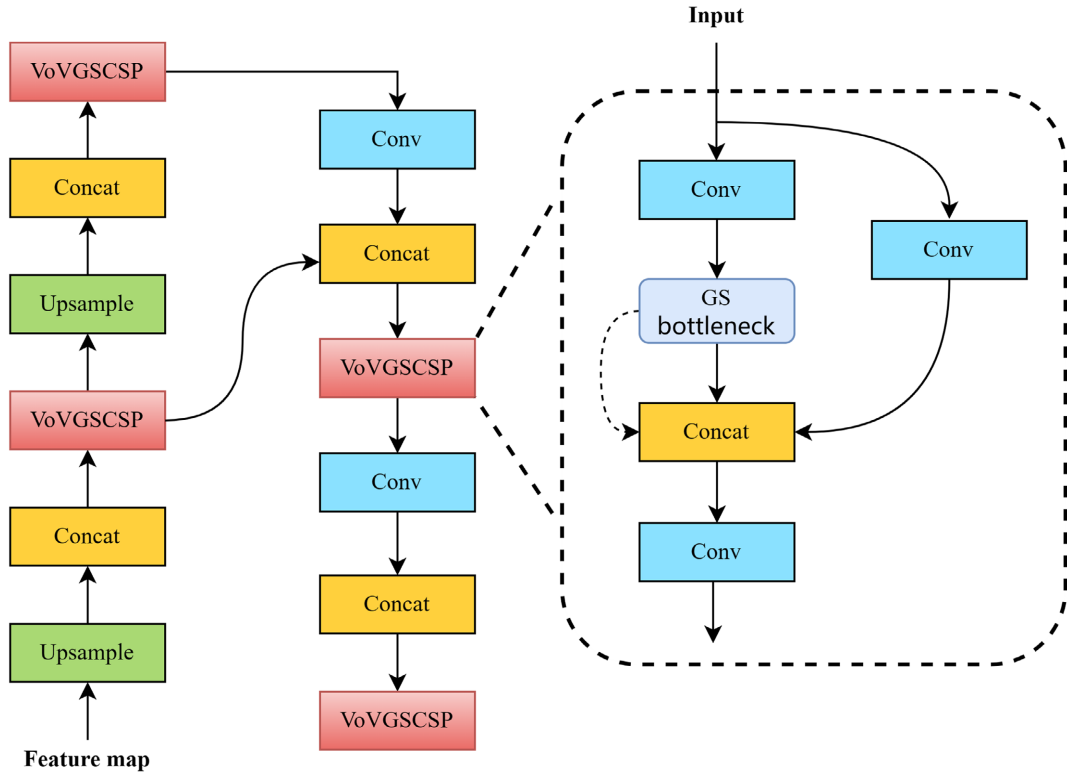


Figure 5. Lightweight feature fusion network structure

图 5. 轻量级的特征融合网络结构

### 2.4. 损失函数设计

YOLOv8 旋转框目标检测的损失函数由三部分组成, 即分类损失(Cls Loss)、定位损失(Bbox Loss)以及旋转目标分布焦点损失(RDFL Loss)三部分具体公式为:

$$L = \lambda_{Cls} l_{Cls} + \lambda_{Bbox} l_{Bbox} + \lambda_{RDFL} l_{RDFL}
 \tag{5}$$

对于定位损失, 由于 CIoU 不考虑目标框旋转角度, 并不适用于旋转框目标检测。旋转交并比(SkewIoU)是评价旋转框目标检测损失的一个非常良好的指标。但是直接计算旋转交并比计算复杂度过高, 且在目标物体位于图像边界时, 检测精度会受到影响。因此目前大多使用基于高斯表示的旋转目标检测损失函数来模拟旋转交并比。例如 Xue 等人提出的 GWD [19]和基于 GWD 的改进算法 KLD 损失函数[28]。但

是上述算法对于超参数的选择过于敏感,但仍存在边界问题。Jeffri 提出的 ProbIoU 考虑了预测边界框的不确定性,从而更准确地评估了检测的性能。

其首先将旋转框转化为二维高斯分布区域:

$$\Sigma = \begin{bmatrix} a & c \\ c & b \end{bmatrix} = R_\theta \begin{bmatrix} a' & 0 \\ 0 & b' \end{bmatrix} R_\theta^T = \begin{bmatrix} a' \cos^2 \theta + b' \sin^2 \theta & \frac{1}{2}(a' - b') \sin 2\theta \\ \frac{1}{2}(a' - b') \sin 2\theta & a' \sin^2 \theta + b' \cos^2 \theta \end{bmatrix} \quad (6)$$

随后计算目标框与预测框之间的巴氏距离(Bhattacharyya Distance):

$$B_C(p, q) = \int_{\mathbb{R}^2} \sqrt{p(x)q(x)} dx \quad (7)$$

其中  $p(x)$  和  $q(x)$  为预测框与目标框的概率密度函数。但是由于巴氏距离并不满足三角不等式,因此进一步将巴氏距离转换为 Hellinger Distance(HD):

$$H_D(p, q) = \sqrt{1 - B_C(p, q)} \quad (8)$$

最终得到 ProbIoU 损失:

$$1 - H_D(p, q) \quad (9)$$

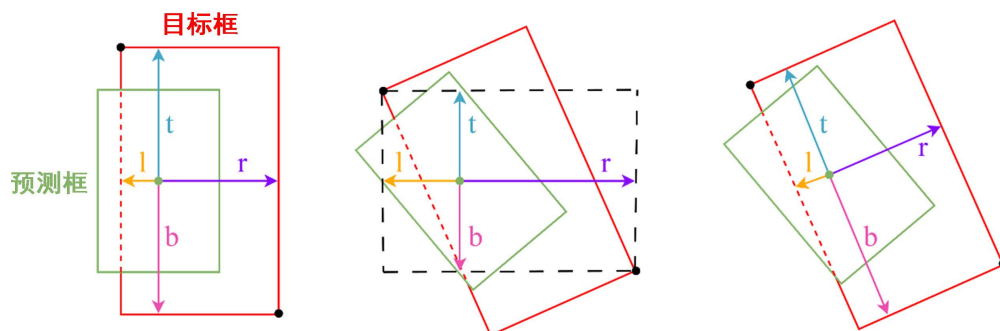
KLD 和 ProbIoU 两种损失函数的性能对比如表 1 所示。实验基于 YOLOv8 模型,分别使用 KLD 和 ProbIoU 两种损失函数训练模型,得到不同模型的 mAP。从表中可以看到无论是在 DOTAv1.0 数据集上还是 DIOR-R 数据集上,使用 ProbIoU 损失的 mAP 都明显高于使用 KLD 损失。这说明 ProbIoU 损失函数能够更好地解决旋转目标的角度边界不连续问题。因此采用 ProbIoU 作为模型旋转框检测的定位损失函数。

**Table 1.** mAP with different loss functions

**表 1.** 不同损失函数的 mAP

损失函数	DOTAv1.0	DIOR-R
KLD	0.704	0.563
ProbIoU	0.732	0.642

DFL 损失部分,将原水平框 DLF 损失直接应用于旋转框,会导致收敛速度较慢。这主要是因为水平框 DLF Loss 缺乏角度信息。如图 6 所示,(a)为水平框计算预测框中心点到真值框四条边框的距离。如果仍旧采用水平框计算距离的方式,旋转框实际匹配的目标框并不是目标框,而是图 6(b)中所示的黑色虚线水平框。



**Figure 6.** DFL and RDFL principles

**图 6.** DFL 及 RDFL 原理



为解决这一问题，计算距离时加入角度信息，计算公式如下：

$$\begin{aligned} l &= (y' - y_1) \sin \theta - (x' - x_1) \cos \theta \\ t &= (y' - y_1) \cos \theta + (x' - x_1) \sin \theta \\ r &= (y_2 - y') \sin \theta - (x_2 - x') \cos \theta \\ b &= (y_2 - y') \cos \theta + (x_2 - x') \sin \theta \end{aligned} \quad (10)$$

其中  $(x', y')$  为预测框中心点坐标， $(x_1, y_1)$  为目标框的左上角点坐标， $(x_2, y_2)$  为目标框的右下角点坐标。这样计算后的距离为预测框中心到目标框四条边框的真实距离，如图 6(c) 所示。

### 3. 数据集及实验设置

#### 3.1. 数据集及实验设置

DOTAv1.0 数据集[29]中有 2806 幅航空图像，15 个类别共包含 188,282 个用旋转框标注的实例目标。类别分别为飞机、棒球场、桥梁、田径场、小型车辆、大型车辆、船舶、网球场、篮球场、油罐、足球场、环岛、港口、游泳池、直升机。每张图片的尺寸都在  $800 \times 800$  像素至  $4000 \times 4000$  像素之间，由于原数据尺寸太大，将 DOTAv1.0 数据集按照 200 像素的重叠度进行切割处理，切割为  $1024 \times 1024$  固定大小的数据集。其中，训练集包含 15,749 幅图像，测试集包含 5297 幅图像。

#### 3.2. 实验环境与参数

实验在 Windows 系统下，采用 PyTorch 1.13.1 开源深度学习框架搭建网络，在具有 NVIDIA GeForce RTX 3060 GPU 和 6G 显存的计算机上实现。使用 SGD 作为优化器训练，动量因子为 0.9，权重衰减系数为  $10^{-4}$ ，初始学习率为  $10^{-5}$ ，Batch Size 设置为 6。

#### 3.3. 评价指标

为了更好地评价检测方法的表现，实验采用了常用的平均精度 AP (Average Precision) 作为评价指标。检测结果可以分为 3 类，分别是正确预测的正样本 TP (True Positive)，错误预测的正样本 FP (False Positive) 以及错误预测的负样本 FN (False Negative)。用精度指标 P 表示所有检出目标中正确预测的比例，召回指标 R 表示正确预测的目标占应检出目标的比例，则：

$$\left\{ \begin{aligned} P &= \frac{TP}{(TP + FP)} \\ R &= \frac{TP}{(TP + FN)} \end{aligned} \right. \quad (11)$$

AP 则是计算从  $R = 0$  到  $R = 1$  区间内精度 P 的平均值，更高的 AP 值意味着检测效果越好，反之亦然。由于 AP 是针对单个类别，而测试的数据集具有多个类别标签。因此采用多个类别 AP 的平均值用于多类目标检测效果的衡量，这个平均值常用 mAP (Mean Average Precision) 表示。

## 4. 实验结果分析

### 4.1. 对比实验

为了验证 YOLOv8-LR 的有效性，将其与目前一些经典目标检测算法进行对比。表 2 给出了 YOLOv8-LR 与其他目标检测算法在 DOTA 数据集上的检测结果[30]。其他目标检测算法被分为基于回归的目标检测算法、基于感兴趣区域的目标检测算法和基于 DETR 的目标检测算法。其中加粗字体表示所

有检测算法中检测精度最高的结果。

从表 2 中可以看出, 在 DOTA 数据集上, YOLOv8-LR 的 mAP 值达到 77.2%, 相比于其他目标检测算法, YOLOv8-LR 模型拥有更好的目标检测性能, 为所有实验检测算法中最高。与基于回归的目标检测算法 Rotated RetinaNet、R<sup>3</sup>Det、Rotated FCOS 和基于感兴趣区域的目标检测算法 Rotated Faster RCNN、ReDet 相比, YOLOv8-LR 在 DOTA 数据集上 mAP 提高了 3.2%至 8%。除少数类型检测精度略低于上述其他模型外, 如篮球场、网球场, YOLOv8-LR 在其余目标, 尤其是小尺度目标上均有所提升。例如, 相比于 Rotated RetinaNet, YOLOv8-LR 对小型汽车、桥梁、直升机的检测精度分别提升了 27.6%、25.6%、25.8%。

与基于 DETR 的目标检测算法 Rotated D-DETR、AO<sup>2</sup>DETR、ARS-DETR 相比, YOLOv8-LR 的 mAP 提高了 3.4%至 13.8%。除大型车辆、网球场、篮球场等目标, 上述其他模型略高外, 其余目标, YOLOv8-LR 仍然获得了更好的检测结果。其中小目标检测精度同样提升明显, 例如小型车辆, 相较于其他上述模型, YOLOv8-LR 的检测结果可以提高 2.2%至 13.9%。最后与 AO<sup>2</sup>DINO 模型相比, YOLOv8-LR 的 mAP 值提高了 2.2%。且对于大部分目标检测精度, YOLOv8-LR 均高于 AO<sup>2</sup>DINO 模型。

**Table 2.** Comparison of detection accuracy of different types of objects in the DOTAv1.0 dataset  
**表 2.** DOTAv1.0 数据集中不同类别目标的检测精度对比

算法	基于回归的目标检测算法		基于感兴趣区域的目标检测算法			基于 DETR 的目标检测算法				
	Rotated RetinaNet	R <sup>3</sup> Det	Rotated FCOS	Rotated Faster RCNN	ReDet	Rotated D-DETR	AO <sup>2</sup> DETR	ARS-DETR	AO <sup>2</sup> DINO	本文
飞机	0.873	0.892	0.8852	0.891	0.889	0.790	0.880	0.866	0.869	0.912
棒球场	0.789	0.833	0.775	0.783	0.781	0.686	0.795	0.773	0.819	0.883
桥梁	0.465	0.480	0.471	0.489	0.512	0.326	0.457	0.488	0.483	0.721
田径场	0.698	0.725	0.638	0.715	0.728	0.552	0.666	0.668	0.729	0.746
小型车辆	0.677	0.775	0.804	0.740	0.743	0.725	0.789	0.784	0.799	0.953
大型车辆	0.623	0.767	0.805	0.750	0.781	0.578	0.739	0.790	0.632	0.717
船舶	0.736	0.865	0.873	0.859	0.874	0.737	0.733	0.874	0.859	0.886
网球场	0.909	0.909	0.904	0.908	0.908	0.884	0.904	0.906	0.882	0.872
篮球场	0.828	0.823	0.778	0.869	0.808	0.755	0.806	0.828	0.829	0.768
油罐	0.794	0.835	0.841	0.850	0.786	0.793	0.859	0.822	0.869	0.881
足球场	0.596	0.610	0.555	0.580	0.609	0.454	0.552	0.540	0.612	0.706
环岛	0.619	0.631	0.658	0.697	0.642	0.538	0.636	0.626	0.660	0.546
港口	0.650	0.676	0.650	0.681	0.768	0.529	0.518	0.726	0.656	0.724
游泳池	0.678	0.693	0.728	0.713	0.728	0.664	0.702	0.728	0.774	0.601
直升机	0.450	0.495	0.492	0.569	0.549	0.504	0.600	0.650	0.771	0.708
mAP	0.692	0.734	0.724	0.740	0.740	0.634	0.709	0.738	0.750	0.772

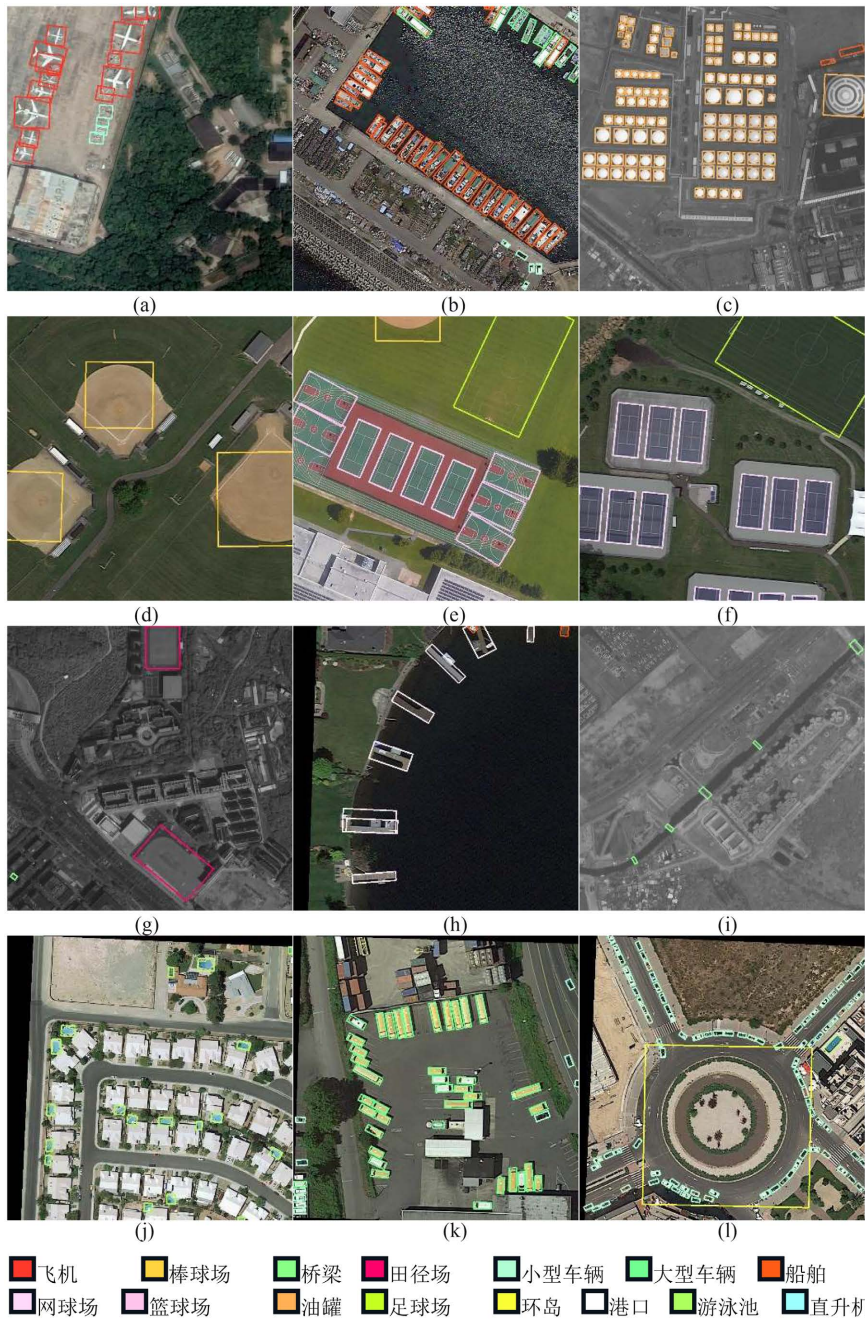
## 4.2. 消融实验

为了验证算法中各改进模块的有效性, 将 YOLOv8 的定位损失函数替换为 Prob 来检测旋转目标, 在此基础上分别添加通道空间特征增强模块(I-CBAM)和轻量化特征融合网络(LFP), 在 DOTA 数据集上进行消融实验。

从表 3 中的实验结果可以看出, 仅添加 I-CBAM 模块后, 在 DOTA 数据集上, 算法的 mAP 从 73.2%

**Table 3.** The impact of different improvements on detection performance  
**表 3.** 不同改进部分对检测性能的影响

I-CBAM	LFP	Prob	mAP	参数量(MB)	计算量(GB)
		✓	0.732	3.022	8.5
✓		✓	0.767	3.027	8.6
	✓	✓	0.738	2.825	7.7
✓	✓	✓	0.772	2.865	7.8



**Figure 7.** Visualization of detection performance of YOLO8-LR on DOTA dataset  
**图 7.** YOLO8-LR 在 DOTA 数据集上检测效果可视化

提高到 76.7%。而参数量和计算量几乎与之前保持一致。这表明 I-CBAM 模块共享通道特征以提取丰富特征信息,有效增强对遥感图像目标检测精度。仅添加 LFP 后,模型大小明显降低,参数量和计算量分别降低了 0.197 MB 和 0.8 GB,为 YOLOv8 的 6.5% 和 9.4%。在数据集上的 mAP 也略有提升。验证了轻量化特征融合网络可以有效降低模型计算成本,且不会对检测精度造成损失。最终算法 YOLOv8-LR 综合所有模块,相较于 YOLOv8,在 DOTA 数据集上 mAP 增加了 4%,参数量和计算量分别降低了 0.157 和 0.7,为 YOLOv8 参数量和计算量的 5.4% 和 9.0%。综上所述,I-CBAM 在提升目标检测精度方面表现出色,而 LFP 则有效的降低了模型的参数量和计算量,使模型更轻量化、更高效。

### 4.3. 可视化检测结果

为了直观显示目标检测效果,将 YOLOv8-LR 在 DOTA 数据集中部分图像检测结果进行可视化,结果如图 7 所示。通过观察可以发现,YOLOv8-LR 可以准确地检测到大多数目标与任意方向。尤其在尺度较小和密集排布的目标检测上效果显著,且漏检的情况较少,如飞机、车辆、船舶、油罐等。

对于不同分辨率图像中的尺度差异较大的同一种类目标(如飞机、油罐等),YOLOv8-LR 能够实现较高精度检测,如图 7(a)和图 7(c)所示。此外也有能力对同一幅图像中尺度差异较大的不同物体准确定位,如图 7(l)中同时存在环岛和车辆的情况。

## 5. 结束语

针对遥感图像目标种类多、密集分布和尺度变化剧烈,导致遥感图像目标多尺度尤其小目标难以检测的问题,本文提出一种轻量级无锚框旋转目标检测算法 YOLOv8-LR。首先,设计了一种轻量级通道空间特征增强模块,将特征图的通道信息共享,有效地增加特征图的细节信息,提升小目标预测能力。其次,构建轻量化特征融合网络,降低计算复杂度而不损失检测精度。最后,增加输出维度并引入角度信息,优化定位损失函数,解决旋转框角度边界问题,并改进动态损失函数以加速模型收敛。实验结果表明,YOLOv8-LR 能有效实现高精度的小目标和密集分布目标的检测,验证了其可行性和有效性。

## 参考文献

- [1] Ding, J., Xue, N., Long, Y., *et al.* (2019) Learning RoI Transformer for Oriented Object Detection in Aerial Images. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, 15-20 June 2019, 2849-2858. <https://doi.org/10.1109/CVPR.2019.00296>
- [2] Guo, Z., Liu, C., Zhang, X., *et al.* (2021) Beyond Bounding-Box: Convex-Hull Feature Adaptation for Oriented and Densely Packed Object Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 19-25 June 2021, 8792-8801. <https://doi.org/10.1109/CVPR46437.2021.00868>
- [3] Han, J., Ding, J., Li, J., *et al.* (2021) Align Deep Features for Oriented Object Detection. *IEEE Transactions on Geoscience and Remote Sensing*, **60**, Article ID: 5602511. <https://doi.org/10.1109/TGRS.2021.3062048>
- [4] Han, J., Ding, J., Xue, N., *et al.* (2021) Redet: A Rotation-Equivariant Detector for Aerial Object Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 19-25 June 2021, 2786-2795. <https://doi.org/10.1109/CVPR46437.2021.00281>
- [5] Hou, L., Lu, K., Xue, J., *et al.* (2022) Shape-Adaptive Selection and Measurement for Oriented Object Detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, **36**, 923-932. <https://doi.org/10.1609/aaai.v36i1.19975>
- [6] Lang, S., Ventola, F. and Kersting, K. (2021) Dafne: A One-Stage Anchor-Free Deep Model for Oriented Object Detection.
- [7] Li, W., Chen, Y., Hu, K., *et al.* (2022) Oriented Reppoints for Aerial Object Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, New Orleans, 18-24 June 2022, 1829-1838. <https://doi.org/10.1109/CVPR52688.2022.00187>
- [8] Li, Z., Hou, B., Wu, Z., *et al.* (2023) FCOSR: A Simple Anchor-Free Rotated Detector for Aerial Object Detection.

- Remote Sensing*, **15**, Article No. 5499. <https://doi.org/10.3390/rs15235499>
- [9] Ming, Q., Zhou, Z., Miao, L., *et al.* (2021) Dynamic Anchor Learning for Arbitrary-Oriented Object Detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, **35**, 2355-2363. <https://doi.org/10.1609/aaai.v35i3.16336>
- [10] Xie, X., Cheng, G., Wang, J., *et al.* (2021) Oriented R-CNN for Object Detection. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Montreal, 11-17 October 2021, 3520-3529. <https://doi.org/10.1109/ICCV48922.2021.00350>
- [11] Yang, X., Yan, J., Feng, Z., *et al.* (2021) R3det: Refined Single-Stage Detector with Feature Refinement for Rotating Object. *Proceedings of the AAAI Conference on Artificial Intelligence*, **35**, 3163-3171. <https://doi.org/10.1609/aaai.v35i4.16426>
- [12] Tian, Z., Shen, C., Chen, H., *et al.* (2019) FCOS: Fully Convolutional One-Stage Object Detection. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, 27 October-2 November 2019, 9627-9636. <https://doi.org/10.1109/ICCV.2019.00972>
- [13] Ren, S., He, K., Girshick, R., *et al.* (2016) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 1137-1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
- [14] Lin, T.Y., Goyal, P., Girshick, R., *et al.* (2017) Focal Loss for Dense Object Detection. *Proceedings of the IEEE International Conference on Computer Vision*, Venice, 22-29 October 2017, 2980-2988. <https://doi.org/10.1109/ICCV.2017.324>
- [15] Yang, Z., Liu, S., Hu, H., *et al.* (2019) Reppoints: Point Set Representation for Object Detection. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, 27 October-2 November 2019, 9657-9666. <https://doi.org/10.1109/ICCV.2019.00975>
- [16] Jocher, G. (2023) Yolov8. <https://github.com/ultralytics/ultralytics>
- [17] Yang, X., Hou, L., Zhou, Y., *et al.* (2021) Dense Label Encoding for Boundary Discontinuity Free Rotation Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 19-25 June 2021, 15819-15829. <https://doi.org/10.1109/CVPR46437.2021.01556>
- [18] Yang, X. and Yan, J. (2020) Arbitrary-Oriented Object Detection with Circular Smooth Label. *Computer Vision-ECCV 2020: 16th European Conference*, Glasgow, 23-28 August 2020, 677-694. [https://doi.org/10.1007/978-3-030-58598-3\\_40](https://doi.org/10.1007/978-3-030-58598-3_40)
- [19] Yang, X., Yan, J., Ming, Q., *et al.* (2021) Rethinking Rotated Object Detection with Gaussian Wasserstein Distance Loss. *International Conference on Machine Learning*. PMLR, 18-24 July 2021, 11830-11841.
- [20] Li, X., Wang, W., Wu, L., *et al.* (2020) Generalized Focal Loss: Learning Qualified and Distributed Bounding Boxes for Dense Object Detection. *Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020*, 6-12 December 2020, 21002-21012.
- [21] Llerena, J.M., Zeni, L.F., Kristen, L.N., *et al.* (2021) Gaussian Bounding Boxes and Probabilistic Intersection-over-Union for Object Detection.
- [22] Lin, T.Y., Dollár, P., Girshick, R., *et al.* (2017) Feature Pyramid Networks for Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 2117-2125. <https://doi.org/10.1109/CVPR.2017.106>
- [23] Liu, S., Qi, L., Qin, H., *et al.* (2018) Path Aggregation Network for Instance Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-22 June 2018, 8759-8768. <https://doi.org/10.1109/CVPR.2018.00913>
- [24] He, K., Zhang, X., Ren, S., *et al.* (2015) Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **37**, 1904-1916. <https://doi.org/10.1109/TPAMI.2015.2389824>
- [25] Li, D., Hu, J., Wang, C., *et al.* (2021) Involution: Inverting the Inherence of Convolution for Visual Recognition. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 19-25 June 2021, 12321-12330. <https://doi.org/10.1109/CVPR46437.2021.01214>
- [26] Chen, B., Zhang, Z., Liu, N., *et al.* (2020) Spatiotemporal Convolutional Neural Network with Convolutional Block Attention Module for Micro-Expression Recognition. *Information*, **11**, Article No. 380. <https://doi.org/10.3390/info11080380>
- [27] Li, H., Li, J., Wei, H., *et al.* (2022) Slim-Neck by GSConv: A Better Design Paradigm of Detector Architectures for Autonomous Vehicles.
- [28] Yang, X., Yang, X., Yang, J., *et al.* (2021) Learning High-Precision Bounding Box for Rotated Object Detection via Kullback-Leibler Divergence. *35th Conference on Neural Information Processing Systems (NeurIPS 2021)*, 6-14 De-

---

cember 2021, 18381-18394.

- [29] Xia, G. S., Bai, X., Ding, J., *et al.* (2018) DOTA: A Large-Scale Dataset for Object Detection in Aerial Images. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 3974-3983. <https://doi.org/10.1109/CVPR.2018.00418>
- [30] 金睿蛟, 王堃, 刘敏豪, 等. 基于 DETR 和改进去噪训练的光学遥感图像多尺度旋转目标检测[J]. 激光与光电子学进展, 2024, 61(2): 356-366.