

基于改进YOLOv7-tiny的乒乓球旋转识别与测量算法

张科, 季云峰

(上海理工大学健康科学与工程学院, 上海 200093)

摘要: 旋转是乒乓球运动过程中不可避免的运动状态, 同时也是比赛取得胜利的一大要素, 乒乓球商标识别是检测乒乓球旋转状态和速度的重要指标, 因而对商标的准确识别有着重要意义。为解决球和商标检测难度大、精度低, 容易出现错检漏检的问题, 提出一种基于YOLOv7-tiny的轻量化识别和旋转测量改进算法。在识别算法方面, 首先, 将Biformer注意力机制添加到主干网络中, 通过结合局部注意力和全局注意力, 实现了细节信息和全局信息的双重捕捉; 其次, 替换更加轻量化的CARAFE上采样模块, 与传统的上采样方法相比, CARAFE模块通过内容感知的方式重组特征图, 从而在保留更多细节和上下文信息的同时, 实现高质量的上采样; 最后, 修改损失函数WIOU, 不同的误差分量分配适当的权重, 使得损失计算更加灵活和精确。在旋转测量算法方面, 通过目标识别算法获取乒乓球及其商标信息, 并结合商标补全算法解决因自旋与视角导致的商标不完整问题。利用图像中的商标中心与球心坐标差值、商标面积、深度信息等, 计算球面投影半径、仰角和方位角, 从而修正与补偿商标缺失, 进一步准确计算乒乓球转速。在乒乓球数据集上的实验结果表明, 改进算法准确率和召回率分别提高了2.7%和1.4%, FPS值提升了13帧/s, 旋转速度最大误差为5转/s。该算法在提升其准确率和召回率的同时, 其帧率也相应提升, 极大满足了目标检测的有效性和实时性。

关键词: 乒乓球; YOLOv7-tiny; 小目标识别; Biformer; 旋转测量

DOI: 10.11907/rjdk.251213

扫描二维码阅读全文:



中图分类号: TP391.41

文献标识码: A

文章编号: 1672-7800(2026)005-0196-09

Table Tennis Rotation Recognition and Measurement Method Based on Improved YOLOv7-tiny

ZHANG Ke, JI Yunfeng

(School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: Rotation is an inevitable movement state in the process of table tennis, and it is also a major factor in winning the game. The identification of table tennis trademarks is an important indicator for detecting the rotation state and speed of table tennis. Therefore, accurate identification of trademarks is of great significance. A lightweight improvement algorithm based on YOLOv7-tiny is proposed to solve the problems of difficulty, low accuracy, and easy occurrence of false positives and missed detections in ball and trademark detection. Firstly, the Biformer attention mechanism was added to the backbone network, which achieved dual capture of detail information and global information by combining local attention and global attention. Next, replace the more lightweight CARAFE upsampling module, which is superior to traditional upsampling methods. The CARAFE module recombines feature maps through content awareness, achieving high-quality upsampling while retaining more details and contextual information. Finally, modify the loss function WIOU and assign appropriate weights to different error components to make the loss calculation more flexible and accurate. The experimental results show that the accuracy and recall of the table tennis dataset have been improved by 2.7% and 1.4% respectively, and the FPS value has increased by 13 frames/s, maximum rotational speed error is 5 rev/s. While improving its accuracy and recall, its frame rate also increases accordingly, greatly meeting the effectiveness and real-time requirements of object detection.

Key Words: table tennis; YOLOv7-tiny; small target recognition; Biformer; rotation measurement

收稿日期: 2025-04-15

基金项目: 国家自然科学基金项目(62403319)

作者简介: 张科(2000-), 男, 上海理工大学健康科学与工程学院硕士研究生, 研究方向目标识别与旋转测量; 季云峰(1990-), 男, 博士, 上海理工大学健康科学与工程学院副教授、硕士生导师, 研究方向为乒乓球机器人。本文通讯作者: 张科。

0 引言

乒乓球运动作为一项受欢迎的球类运动,具有简单易学、技术多样性和适应性强的特点。乒乓球作为我国的“国球”,被广泛认可为一项适合各年龄段人群参与的运动项目。我国在乒乓球领域的研究不仅仅体现在体育赛事上,而且对乒乓球机器人的研究也有一定造诣。乒乓球机器人要完成完整的击打任务,首先需要通过视觉系统完成对乒乓球的识别与追踪工作。乒乓球在空中受到外力而改变角动量,从而导致旋球的产生,要想完成对旋转的准确预测,就需要对三维空间中的商标与球体进行准确识别,根据商标相对球体的运动得到乒乓球的旋转状态与速度。由于乒乓球在飞行过程中有着目标小、速度快、变化多的特点,往往会出现目标拖影、遮挡、模糊边界等问题,导致在对其识别捕捉的过程中出现拖影、目标识别丢失等问题。因此,对乒乓球及其商标的准确识别是主要难点之一。

目标检测方法主要分为传统方法和基于深度学习的方法。在传统的算法研究中,Hur等^[1]利用连续运动图像的光学流场与遮挡图前后关键点的对称信息对遮挡区域进行识别预测。Jun等^[2]使用差帧法去除图像上的噪声点,通过轮廓跟踪算法得到商标的运动状态。Lu等^[3]结合混合高斯模型和背景减除法,提高了运动目标检测准确性和鲁棒性,特别是在复杂场景下的表现。Tebbe等^[4]首先通过帧差法提取乒乓球商标的轮廓点,其次将提取到的所有轮廓点的平均值归一化,最后结合圆段拟合法准确地得到商标的中心位置。这些方法虽然可以粗略地得到乒乓球的旋转状态,但由于环境因素、运动遮挡等影响,很难保证在视野范围内得到完整的商标。近年来,随着深度学习的发展,研究人员也逐渐开始将深度学习技术应用到伺服乒乓球机器人击打中。张不凡等^[5]在YOLOv5主干网络中引入跨层特征增强连接,显著增强对小目标的表征能力。董彦强等^[6]通过YOLOv4轻量化模型中的重构特征融合模块,在网络中加入注意力机制模块以提高检测精度。赵英等^[7]通过改进的huffleNetV2网络单元组合重构YOLOv5s主干网络,进一步融合高效通道注意力机制以提高对小目标的检测效果。

上述方法虽然对小目标检测有一定提升,但在对乒乓球及其商标检测上依然存在如下问题:①乒乓球在运动中存在速度快和随机性强的特点,在对其识别过程中会出现遮挡和拖影问题;②乒乓球及其商标在图像中像素占比过小,对其特征提取过少导致检测性能差;③由于背景、光照等问题,容易出现目标漏检情况。基于上述问题,本文提出了一种基于YOLOv7模型的改进模型。首先,通过数据增强方法对数据集进行遮挡处理;其次,将Biformer注意力机制添加到主干网络中,以提高模型图像特征的提取能

力;再次,修改上采样CARAFE模块以提高上采样质量;最后,替换损失函数WIOU LOSS以更精确地评估预测框与真实框的重叠情况。由于自旋和相机视角等因素,会出现捕捉不完整的商标图像,采用商标补全算法对商标进行补全操作,使其得到更准确的商标中心点。对于坐标混乱、误差较大的情况,该实验采用IQR算法对数据进行处理,以进一步得到其旋转速度。

1 相关算法

1.1 YOLOv7算法

与传统目标算法相比较,YOLOv7将目标检测任务视为一个回归问题,直接在图像级别上预测边界框及其对应的类别概率。为了适应不同的应用场景,YOLOv7根据网络深度和宽度不同,将模型分为YOLOv7、YOLOv7-tiny、YOLOv7-x等^[8]。本文考虑到模型轻量化,选择参数量和计算量相对较少的YOLOv7-tiny作为基础模型进行研究。

YOLOv7-tiny整体网络结构如图1所示,其网络结构包括输出端口(Input)、主干网络(Backbone)、颈部网络(Neck)和检测头(Head)4部分。

1.2 三位重建算法

为了实现乒乓球的识别、商标提取和空间定位,本文采用了汇聚式双目视觉模型。如图2所示,汇聚式模型将两个相机的视野交汇点作为参考点,通过计算两个相机的视角差异以获取深度信息。在乒乓球机器人系统中, P 点表示乒乓球球心位置, P_l 和 P_r 分别表示左、右摄像机获取图像中乒乓球球心的二维图像坐标。基于双目视觉成像原理,在已知 P_l 和 P_r 的情况下,通过标定的摄像机参数,求取 P 点三维坐标的过程被称为三维重建。首先,采用张正友标定法对相机进行标定,获取相机的内外参数等信息,以确保三维坐标计算的精确性;其次,使用相机捕捉的双目图像,通过立体匹配算法计算乒乓球在不同视角下的位置,这些位置信息可以通过视差计算得到相对深度信息,并结合相机的已知参数进行绝对深度计算;最后,将从两个相机中获得的二维像素坐标转换为三维世界坐标系下的位置,进一步分析处理图像以提取乒乓球商标的信息,二维图像坐标转三维直接坐标公式如式(1)所示。

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = s \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = M_1 M_2 \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (1)$$

其中, (x,y) 为图像坐标系下的坐标, (u_0,v_0) 为图像坐

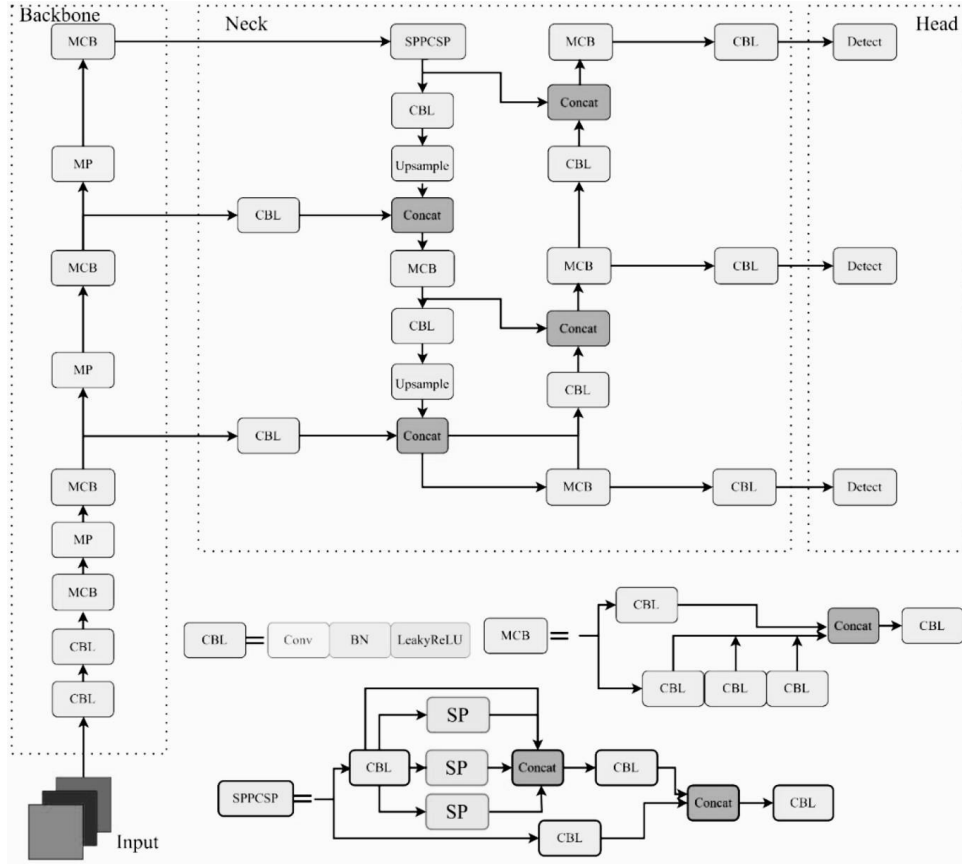


Fig. 1 YOLOv7-tiny network structure

图1 YOLOv7-tiny 网络结构

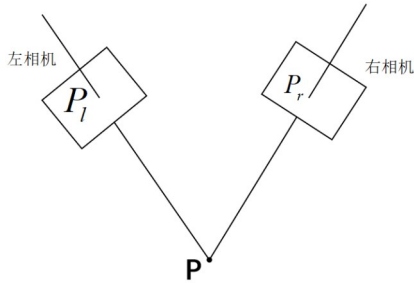


Fig. 2 Schematic diagram of binocular imaging

图2 双目成像示意图

标原点的横纵坐标, M_1, M_2 分别为相机的内参矩阵和外参矩阵, f 为相机的焦距, s 为比例因子, (x_c, y_c, z_c) 为相机坐标系下的坐标, (x_w, y_w, z_w) 为世界坐标系的坐标, d_x 和 d_y 是图像坐标系中 x 方向和 y 方向每个像素的单位长度。

然而,由于相机镜头在制造过程中存在装配和加工误差,使镜头在成像时产生失真,造成相机的非线性畸变,这将导致理想成像平面上的投影点与实际成像点之间存在偏移。非线性畸变主要分为径向畸变和切向畸变。

对于径向畸变,成像装置光学中心处的畸变为0,越靠近透镜,边缘畸变越大。径向畸变大小通常用成像中心位置周围的泰勒级数展开式的前三项加以近似表示,如式(2)所示。切向畸变的产生是由于整个相机装配过程中图像平面所在位置与透镜不平行所导致,如式(3)所示。

$$x_{\text{correcte}} = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \quad (2)$$

$$y_{\text{correcte}} = y(1 + k_1 r^2 + k_3 r^6) \quad (3)$$

其中, x 和 y 是成像点的原始位置; x_{correcte} 和 y_{correcte} 是成像点校正后的新位置。 k_1, k_2 和 k_3 是径向畸变参数, r 为图像像素点到图像中心点的距离。 p_1 和 p_2 是和切向畸变参。至此,从二维图像坐标到乒乓球三维世界坐标的求解方法即完成。

2 算法改进

2.1 YOLOv7-tiny 上采样模块改进

上采样模块主要用于将低分辨率的特征图恢复到高分辨率,从而能够与高分辨率的特征图进行融合。通过这种方式,模型可以结合不同层级的特征,提高目标检测精度和鲁棒性。上采样模块在 YOLOv7-tiny 中用于特征金字塔网络 Feature Pyramid Network (FPN) 或路径聚合网络 Path Aggregation Network (PAN) 的构建,通过最邻近插值方法实现上采样操作,该算法虽然有着简单和速度快的特点,但是其采样后灰度值的不连续性很明显,导致图像质量受损,影响最终识别效果。

为了解决上述问题,本文采用 CARAFE 上采样轻量级模块取代原模块中的 Upsample 结构,该模块分为内容感知重组模块和内核预测模块,它通过使用内容感知的特征重

新装配机制实现上采样,具有更大范围的感受野,从而更好地保留高频细节和语义信息^[9]。其结构如图3所示。

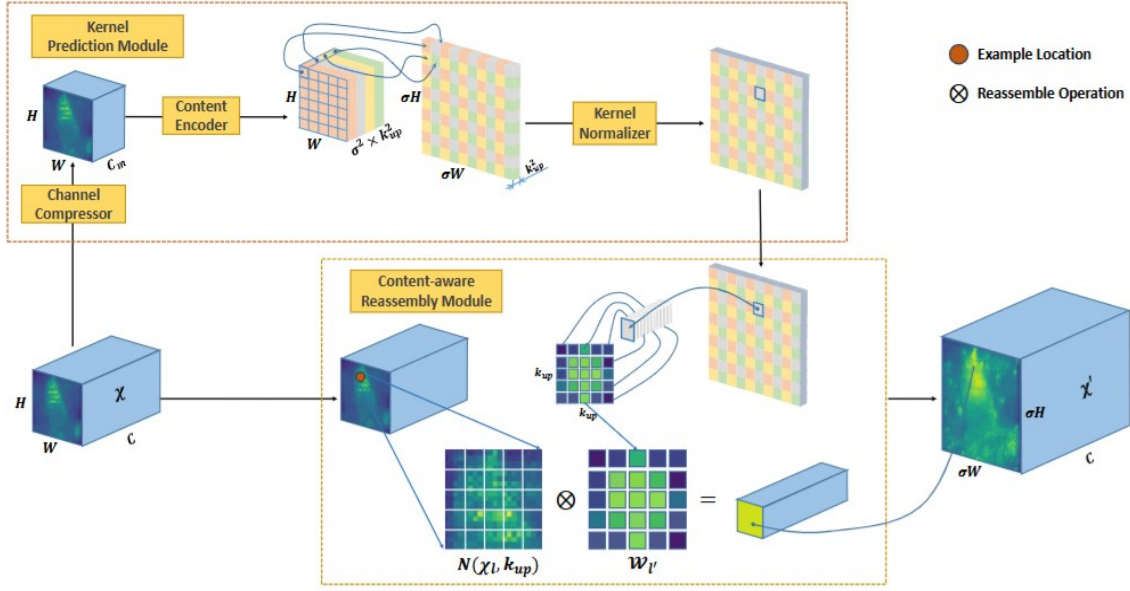


Fig. 3 CARAFE structure

图3 CARAFE结构

2.2 注意力机制模块添加

注意力机制(Attention Mechanism)最早在自然语言处理领域中被提出。随着深度学习的发展,注意力机制也逐渐应用于图像处理,它使得模型可以在图像的不同区域间分配不同的注意力,从而更有效地提取和利用图像中的关键特征。具体而言,注意力机制可以帮助模型在复杂背景中突出目标,或者在多目标场景中区分各目标的重要性。本文将Biformer注意力机制添加到主干网络中,该注意力机制通过结合局部注意力和全局注意力以提高模型在计算机视觉任务中的表现^[10]。Biformer结构如图4所示。

(1)局部注意力计算。设输入特征图为 $X \in \mathbb{R}^{H \times W \times C}$,其中 H 和 W 分别为特征图的高度和宽度, C 为通道数。 $Q_i = X_i W_Q$, $K_i = X_i W_K$, W_Q 和 W_K 是可学习的权重矩阵, d_k 是缩放因子。局部注意力计算公式如式(4)所示。

$$A_i = \text{softmax} \left(\frac{Q_i K_i^T}{\sqrt{d_k}} \right) \quad (4)$$

(2)局部注意力加权。将注意力权重 A_i 应用于窗口内的特征 X_i ,得到加权特征 Y_i ,如式(5)所示。

$$Y_i = A_i V_i \quad (5)$$

其中, W_V 是可学习的权重矩阵。

(3)全局注意力机制。全局注意力机制在整个特征图范围内计算注意力权重,将输入特征 X 投影到查询、键和值空间,如式(6)一式(8)所示。

$$Q = X W_Q \quad (6)$$

$$K = X W_K \quad (7)$$

$$A = \text{softmax} \left(\frac{Q K^T}{\sqrt{d_k}} \right) \quad (8)$$

(4)全局注意力加权。将注意力权重 A 应用于全局特征 V ,得到加权特征 Y_{global} 。其公式如式(9)所示。

$$Y_{\text{global}} = AV \quad (9)$$

(5)双向融合。将局部特征和全局特征进行融合,生成最终的特征输出,如式(10)所示。

$$Y = \alpha Y_{\text{local}} + \beta Y_{\text{global}} \quad (10)$$

其中, α 和 β 是融合权重,可以是固定值或通过学习得到的权重, Y_{local} 是局部特征, Y_{global} 是全局特征, Y 是输出特征。

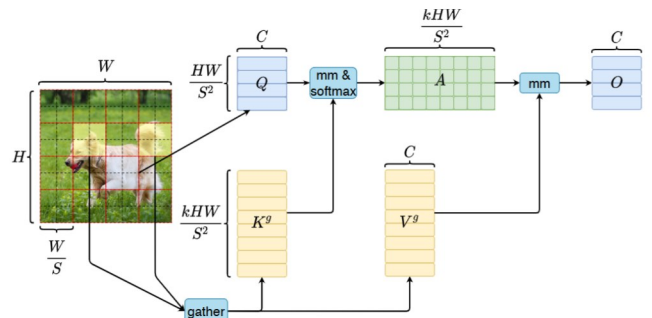


Fig. 4 Biformer structure

图4 Biformer结构

2.3 损失函数改进

在深度学习中,损失函数是用来衡量模型预测结果与实际标签之间差异的函数,其设计目标是让模型通过优化损失值,尽可能准确地学习目标任务。在YOLOv7的目标检测模型中,损失函数的作用是衡量模型预测的边界框(Bounding Box)信息和类别预测与真实标签之间的差异。

目前,常见的损失函数有GIoU^[11]、DIoU^[12]、CIoU^[13]、EIoU^[14]、WIoU^[15]损失函数等。

(1) GIoU Loss (Generalized Intersection over Union Loss)。GIoU Loss 通过引入预测框和真实框的最小外接矩形,以获取预测框、真实框在闭包区域中的比重,从而解决了两个目标没有交集时梯度为零的问题,其计算公式如式(11)所示。

$$L_{\text{GIoU}} = 1 - \text{IoU} + \frac{|C - (A \cup B)|}{|C|} \quad (11)$$

其中, C 为 A 与 B 的最小外界矩阵。

(2) DIoU Loss (Distance Intersection over Union Loss)。DIoU Loss 旨在解决传统的 IoU 和 GIoU 收敛速度慢、预测框回归不稳定(尤其是在包含关系或水平/垂直对齐时)的问题,其计算公式如式(12)所示。

$$L_{\text{DIoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b^{\text{gt}})}{c^2} \quad (12)$$

其中, b, b^{gt} 分别表示预测框和真实框的中心点。 ρ 表示欧几里得距离, c 能够同时包含预测框和真实框的最小外接矩阵的对角线长度。

(3) CIoU Loss (Complete Intersection over Union Loss)。CIoU Loss 是用于优化边界框回归的一种损失函数,它引入了边界框的中心点距离和长宽比的惩罚项,从而能够在多种情况下提供更好的回归性能。其计算公式如式(13)所示。

$$L_{\text{CIoU}} = 1 - \text{IoU} + \frac{d^2}{c^2} + \alpha\rho \quad (13)$$

其中, α 是一个权重平衡项,用于平衡中心点距离和长宽比一致性部分的影响, d 表示预测框和真实框中心点之间的欧氏距离, c 表示外接框的对角线长度, ρ 预测框和真实框的长宽比差异。

(4) EIoU Loss (Efficient Intersection over Union Loss)。EIoU Loss 旨在在进一步提升边界框回归的精度,该损失函数考虑了边界框的交集、距离以及尺度之间的关系。其计算公式如式(14)所示。

$$L_{\text{EIoU}} = 1 - \text{IoU} + \frac{d^2}{c^2} + \frac{w_{\text{diff}}^2}{w_g^2} + \frac{h_{\text{diff}}^2}{h_g^2} \quad (14)$$

其中, d 为预测框和真实框中心点之间的欧氏距离, w_{diff} 为预测框和真实框的宽度差异, h_{diff} 为预测框和真实框的高度差异, w_g 和 h_g 分别为真实框的高度和宽度。

(5) WIoU Loss。本文使用 WIoU (Weighted Intersection over Union Loss) 以改进 IoU 损失函数,旨在更精确地优化边界框预测。WIoU 通过加权考虑边界框的各方面,平衡 IoU、中心点距离、宽高比等因素,从而在不同情况下都能表现良好。其计算公式如式(15)所示。

$$L_{\text{WIoU}} = 1 - \text{IoU} + \alpha \frac{d_{\text{center}}^2}{c^2} + \beta \frac{w_{\text{diff}}^2}{w_g^2} + \gamma \frac{h_{\text{diff}}^2}{h_g^2} \quad (15)$$

其中, α, β 和 γ 是加权系数,用于平衡中心点距离、宽度差异和高度差异的影响。 c 是外接框的对角线长度,用于规范化中心点距离。 d_{center} 为预测框和真实框中心点之间的距离。

2.4 YOLOv7-tiny 改进模型网络结构

本文对 YOLOv7-tiny 的网络结构进行了改进,以提升其目标检测性能。首先,在主干网络的卷积层中引入了 Biformer 注意力机制,通过结合局部和全局注意力,增强模型对图像特征的捕捉能力;其次,将原有的上采样模块替换为 CARAFE 模块,这种基于内容的特征重组方法能够更有效地恢复图像细节,以提升上采样质量;最后,将损失函数修改为 WIoU Loss,这种加权的 IoU 损失函数能够更精确地评估预测框与真实框的重叠情况,提高边界框的定位准确度。通过这些改进,YOLOv7-tiny 在复杂场景下的检测性能得到了显著提升。改进后的 YOLOv7-tiny 结构如图 5 所示。

2.5 旋转测量算法改进

为解决因乒乓球自旋和相机视角引起的商标信息不完整问题,提出了一种基于商标补全的乒乓球转速测量方法,乒乓球在高速运动和自旋过程中,由于球体表面具有复杂的投影变形,且商标会因视角遮挡或自旋呈现不完整,直接通过图像检测算法识别商标的完整信息存在较大挑战。因此,提出了一种基于商标补全与几何修正的球体自旋测量方法。此方法综合目标检测、双目深度估计与球面几何分析,能够有效补偿商标信息缺失,从而实现对乒乓球旋转速度的精确计算。通过目标检测算法提取乒乓球与商标的二维像素坐标后,计算商标的实际面积与标准商标面积的比值,以得到商标的可见完整度。

为了还原三维空间信息,结合双目相机的深度测量功能,利用内参矩阵和匹配的像素点计算球心和商标中心的三维坐标,如式(16)所示。

$$P_b = \left(\frac{(x_b - c_x)Z}{f_x}, \frac{(y_b - c_y)Z}{f_y}, Z \right) \quad (16)$$

其中, Z 为深度信息, f_x 和 f_y 为相机焦距, c_x 和 c_y 为主点光心坐标。

由于商标不完整可能导致中心位置偏移,因此需要引入补全机制。考虑球面上商标的实际投影遵循球面几何投影变换,其中心偏移可通过式(17)一式(19)加以修正。

$$\Delta x = x_m - x_b \quad (17)$$

$$\Delta y = y_m - y_b \quad (18)$$

$$r_p = \sqrt{\Delta x^2 + \Delta y^2} \quad (19)$$

在球面坐标系下,商标投影与球心的夹角 α ,可表示为式(20)。

$$\alpha = \sin^{-1} \left(\frac{r_p}{R} \right) \quad (20)$$

其中, R 为乒乓球半径。

同时,根据商标的不完整面积比例 r_a ,结合球体表面展开模型,估算商标的实际覆盖范围,并进行位置修正。

$$\alpha = \sin^{-1} \left(\frac{r_p}{R} \right) \quad (21)$$

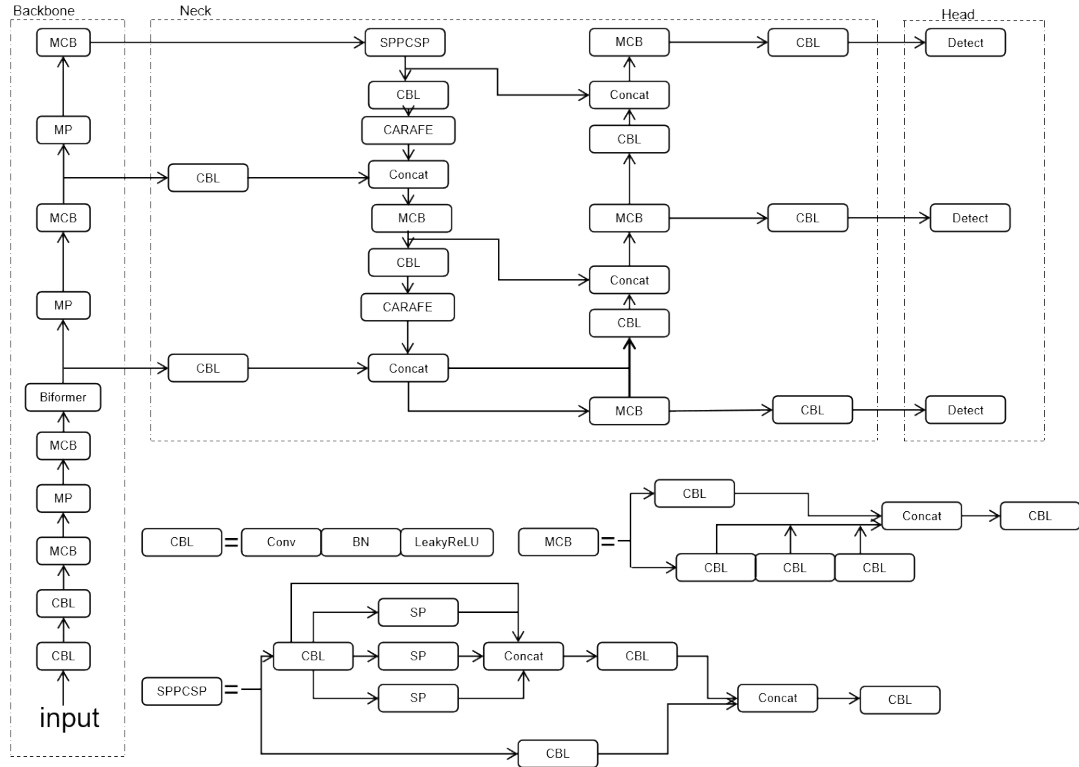


Fig. 5 The improved YOLOv7-tiny structure

图 5 改进后的 YOLOv7-tiny 结构

$$A_{\text{real}} = \frac{A_m}{r_a} \quad (22)$$

$$r_{\text{补全}} = \sqrt{\frac{A_{\text{real}}}{\pi}} \quad (23)$$

通过这一过程,补偿商标中心偏移并获得修正后的商标中心,并将商标中心的投影点映射到球面极坐标系中,计算其仰角和方位角,如式(24)、式(25)所示。

$$\theta = \cos^{-1} \left(\frac{z}{\sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}} \right) \quad (24)$$

$$\theta = \tan^{-2}(\Delta y, \Delta x) \quad (25)$$

商标的运动轨迹通过连续帧图像提取,记录不同时刻 t 的空间位置 (θ_i, φ_i) 。

基于轨迹点计算商标在球面上的弧长位移:

$$\Delta s = R \cdot \cos^{-1}(\sin \beta_1 \sin \beta_2 + \cos \beta_1 \cos \theta_2 \cos(\varphi_2 - \varphi_1)) \quad (26)$$

通过两帧之间的时间间隔 Δt ,估算乒乓球的瞬时自旋速度。

$$\omega = \frac{\Delta s}{\Delta t} \quad (27)$$

商标补全算法显著提高了商标中心的定位精度,减少了旋转速度的测量误差。在应用补全算法前,商标中心的定位误差为 5~10 像素,而在补全后误差降至 2~3 像素。如图 6 所示,红点分别为球中心坐标和商标补全前的坐标,黄点是补全后的商标中心点坐标。此外,与传统基于模板匹配和边缘检测的方法相比,本文算法在商标遮挡严重的情况下仍能保持较高的精度,特别是在商标局部缺失的情

况下,其补偿效果尤为突出。通过这一系列处理,实现了商标信息的精确恢复,并有效提高了乒乓球旋转速度测量的可靠性和稳定性。

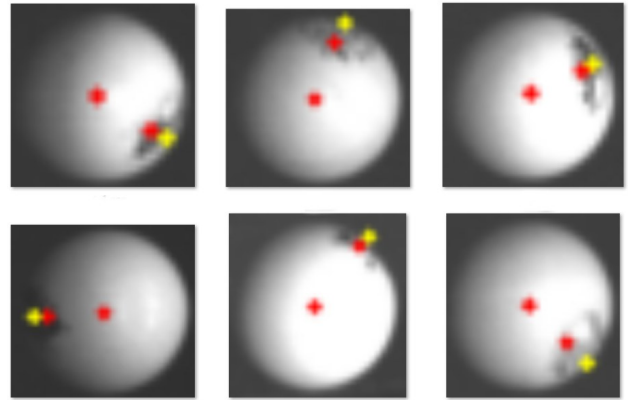


Fig. 6 Trademark supplementary picture

图 6 商标补全图

在通过上述算法获得商标中心点的相对坐标后,由于检测误差或其他噪声的影响,结果可能表现出层次不齐的特性。为进一步提升数据的平滑性和可靠性,可以采用高斯滤波对数据进行平滑处理,并结合四分位距(IQR)算法剔除异常值,得到最终的高质量结果。

IQR(Interquartile Range)方法利用数据的四分位数以识别异常值。四分位数将数据集分为 4 个部分:Q1:第一四分位数(25%),数据集中有 25% 的数据低于这个值;Q2:第二四分位数(50%),即中位数;Q3:第三四分位数(75%),数据集中有 75% 的数据低于这个值。IQR 是 Q3 和

Q_1 之间的差值,用于描述数据的散布程度。

$$IQR = Q_3 - Q_1 \quad (28)$$

通过设定上下限,可以筛选出异常值。通常的设定如式(29)、式(30)所示。

$$\text{下限: } Q_1 - 15 \cdot IQR \quad (29)$$

$$\text{上限: } Q_1 + 15 \cdot IQR \quad (30)$$

它基于数据的四分位数衡量数据分布的扩散程度,并帮助识别远离集中趋势的异常值。它不受极端值影响,因此对非正态分布数据比较鲁棒。

3 实验与结果分析

3.1 数据增强

由于缺乏乒乓球及其商标的开源数据集,为了进一步对乒乓球旋转速度测量进行研究,本文通过两台 200 万像素,最高帧率可达 3 000 fps 的高速相机对旋球拍摄图片,并将其作为数据集。但由于数据的单一性会影响到模型泛化能力,因而本文实验采用旋转、镜像、亮度调整、分块遮挡、高斯噪声等方法对图片进行数据增强处理,最终使用 Labelimg 对 3 614 张图片中的球及其商标进行标注,生成的 xml 文件记录了每个检测框的中心位置及其类别信息。

3.2 实验环境与评估标准

3.2.1 视觉算法验证环境

实验环境是基于 Linux Ubuntu 20.04.6 操作系统搭建,训练所使用的硬件环境是 NVIDIA Corporation GA102GL RTX A6000,编程语言为 Python3.9,深度学习框架为 Pytorch2.0.0+cu117,输入图像的尺寸为 1 920×1 080,,采用随机梯度下降(SGD)进行训练,学习率为 0.001,共训练 300 epoch。

为了评估实验的整体性能,主要使用精确度 P(Precision)、召回率 R(Recall)、平均精准度(mAP)和 FPS 作为模型评估指标。其计算公式如式(31)一式(34)所示。

$$P = \frac{TP}{TP + FP} \quad (31)$$

$$R = \frac{TP}{TP + FN} \quad (32)$$

$$AP = \int_0^1 P(R) dR \quad (33)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (34)$$

其中,TP为真正例,表示正确地将正样本划分为正样本个数;FP为假正例,表示错误地将负样本划分为正样本个数;FN为假负例,表示错误地将正样本划分为负样本个数;TN为真负例,表示正确地将负样本划分为负样本个数。mAP是计算所有类的平均 AP,表示模型在数据集上任务的平均水平。

3.2.2 旋转测量验证装置

该实验利用电机控制乒乓球精确旋转以实现转速的有效验证。该装置通过可调转速电机的转轴与乒乓球连接,且转轴经过乒乓球的球心,将电机设置的转速值作为实际值,通过转速估计值与实际值进行对比以验证旋转测量精度。

3.3 实验结果

为了验证改进算法在乒乓球及其商标识别上的有效性,本文在自建数据集上进行了消融实验。在保证各实验参数相同的前提下,本文将消融实验分为 3 组:第一组验证各改进模块的有效性,第二组通过与当前主流目标检测模型比较以验证改进模型性能,第三组验证损失函数的有效性。

3.3.1 各改进模型比较

为验证本文提出的针对 YOLOv7-tiny 的 3 种改进方法,在自建乒乓球数据增强后的数据集上进行了消融实验。每个数据集有 4 组消融实验,为保证实验准确性,实验环境与参数设置都保持一致。各改进模型比较如表 1 所示。

Table 1 Validation of the effectiveness of each improved module

表 1 各改进模块有效性验证

YOLOv7-tiny	CARAFE	Biformer	WIOU	P/%	R/%	mAP/%	FPS/帧/s
√				89.7	84.7	88.7	106
√	√			89.8	81.2	85.5	142
√		√		91.3	80.7	86.8	132
√			√	91.2	84.7	88.9	156
√	√	√	√	92.4	86.1	91.5	119

可以看出,实验中对 YOLOv7-tiny 模型进行了多项改进,包括引入 CARAFE 上采样模块、Biformer 注意力机制和 WIOU 损失函数。实验结果显示,基础模型的平均精度均值(mAP)为 88.7%,帧率为 106 帧/s。引入 CARAFE 模块后,尽管召回率略有下降,但帧率显著提高至 142 帧/s。使用 Biformer 注意力机制提升了精度,帧率也增加到 132 帧/s。采用 WIOU 损失函数后,精度和召回率均有所提升,帧率达到 156 帧/s。最终,综合引入所有改进后,模型平均精度均值提升至 91.5%,帧率保持在 119 帧/s。综上所述,通过合理组合不同的改进方法,YOLOv7-tiny 模型在目标检测任务中的性能显著提升,提供了更优的检测精度和速度。

3.3.2 目标检测模型比较

将改进后的 YOLOv7-tiny 模型与其他主流目标检测模型进行比较,能够评估和验证模型改进效果。因此,本文将改进后的模型与 SSD、YOLOv4-tiny、YOLOv5、YOLOv8 进行比较,结果如表 2 所示。

根据实验结果可知,改进后的 YOLOv7-tiny 模型在各项指标上均优于主流目标检测模型。本文模型在精度(P)、召回率(R)和平均精度均值(mAP)方面分别达到了

Table 2 Comparison of results of different object detection models
表 2 不同目标检测模型结果比较

模型	P/%	R/%	mAP/%	FPS/帧/s
SSD	34.7	16.49	13.3	34
YOLOv4-tiny	56.6	67.4	36.1	40
YOLOv5	67.7	62.8	61.1	23
YOLOv7-tiny	89.7	84.7	88.7	106
YOLOv8	76.4	81.3	62.3	15
Ours	92.4	86.1	91.5	119

92.4%、86.1% 和 91.5%，均显著高于 SSD、YOLOv4-tiny、YOLOv5、YOLOv7-tiny 和 YOLOv8。同时，推理速度(FPS)也达到了 119 帧/s，表现出色。综上所述，改进后的 YOLOv7-tiny 模型在精度、召回率和速度上均取得了显著提升，展示了卓越的综合性能。

3.3.3 损失函数对比

损失函数定义了模型训练过程中需要优化的目标，通过最小化损失函数，模型的预测结果与真实标签之间的差距得以逐步减小。YOLOv7-tiny 模型通过反向传播和梯度下降等优化算法，不断调整网络参数，以减少损失函数值，进而提高检测精度。本文改进多种损失函数以提高模型性能，其不同损失函数改进后对比结果如表 3 所示。

Table 3 Comparison of improvement results for different loss functions
表 3 不同损失函数改进结果对比

损失函数	P/%	R/%	mAP/%	FPS/帧/s
CIoU	89.7	84.7	88.7	106
GIoU	83.7	82.8	83.9	147
DIoU	91.6	83.4	88.7	156
EIoU	87.6	79.8	85.0	158
WIoU	91.2	84.7	88.9	149

基于 YOLOv7-tiny 的不同损失函数对比结果显示，使用 WIoU 损失函数时模型在各项指标上表现更佳，精度(P%)、召回率(R%)和平均精度均值(mAP%)分别达到了 91.2%、84.7% 和 88.9%，同时保持 149 帧/s 的推理速度；而 DIoU 损失函数也表现出色，精度和 mAP 分别达到了 91.6% 和 88.7%，推理速度为 156 帧/s。相比之下，CIoU 和 EIoU 在精度和速度上表现平衡，而 GIoU 的整体性能较低。因此，WIoU 和 DIoU 损失函数在提升检测精度的同时，兼顾了较高的推理速度。

3.4 旋转验证结果分析

实验结果表明，基于商标补全算法的乒乓球旋转速度测量方法具有较高的精度与鲁棒性。如图 7 所示，第一组为整体实验图，高速双目相机分别位于乒乓球台的左上方和右上方，使其尽可能地捕获到乒乓球及其商标的运动状态，第二组和第三组为算法对缺失商标的有效补偿图，对左相机和右相机传输的球及其商标进行捕捉。但是由于乒乓球会产生自旋，并不能保证每一次捕捉商标都是完整的，因此需要对不完整的商标进行补偿，得到更为精准的商标中心点。图 7 中红色点为 YOLO 算法所得的的商标中心点，黄色点为算法补偿点，可以看出该算法显著提高了

商标中心点的定位精度，减少了旋转测量中所产生的误差。第四组为仿真的三维轨迹图，根据左相机和右相机捕获到的画面及其商标补偿算法得到的中心点，通过 MATLAB 仿真，生成了商标中心点相对于球心的三维运动轨迹，轨迹在球面上分布均匀且连续，验证了算法对商标缺失的有效补偿。综合来看，球心拟合误差较小，平均误差仅为 3 个像素点，球半径拟合值与实际值偏差在可接受范围内。补全算法在商标缺失面积超过 40% 的情况下，仍能将补全误差控制在 6 个像素点以内，显著改善了轨迹平滑性。

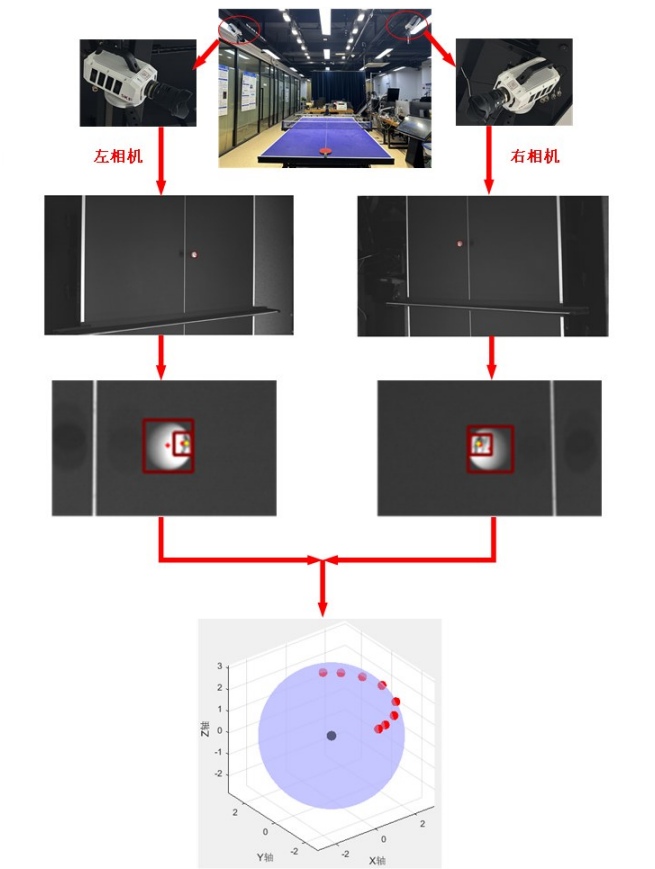


Fig. 7 Three-dimensional motion trajectory of the center point of the trademark after completion

图 7 补全后商标中心点三维运动轨迹

为验证本文乒乓球旋转测量方法在不同转速下的适用性，实验中通过调节电机转速以分析误差。实验装置中，电机的可调转速通过其转轴与乒乓球连接，且转轴确保通过球心，将电机设定的转速作为真实值，进而将其与估计转速进行比较以评估测量精度。表 4 展示了在不同转速条件下乒乓球旋转速度的估计结果及对应误差。从实验结果可知，该方法在旋转速度估计上具有较高的准确性，其最小误差为 1 转/s，最大误差为 5 转/s。

4 结语

本文提出改进型 YOLOv7-tiny 乒乓球识别与转速测

Table 4 Comparison of rotation verification results

表 4 旋转验证结果比较 (转/s)				
转数	1	2	3	4
实际值	67	54	32	10
测量值	62	56	29	11
误差	5	2	3	1

量算法,通过引入Biformer注意力机制、采用CARAFE上采样模块以及优化WIoU损失函数,不仅在乒乓球商标检测上取得了较高的准确率和召回率,而且显著提升了模型实时处理能力。在实验中,该方法在FPS和转速测量精度上均表现出良好的性能,证明了其在高速运动小目标检测与追踪任务中的潜在优势。该算法的实时性和高精度为乒乓球比赛中的智能裁判、训练数据分析和运动员战术改进等实际应用提供了有力支持。同时,算法在小目标识别和运动状态捕捉方面的成功经验,也为其他高速运动场景(如工业自动检测、无人机跟踪等)提供了新的技术思路。然而,在一些复杂背景或存在严重遮挡的场景中,商标检测精度可能会受到影响。尤其在乒乓球被其他物体严重遮挡或背景较为杂乱的情况下,算法准确性可能下降,需进一步优化背景抑制和遮挡处理机制。

参考文献:

- [1] HUR J, ROTH S. MirrorFlow: exploiting symmetries in joint optical flow and occlusion estimation [C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 312–321.
- [2] JUN Z, ZHANG X H. Foreground object detection combining Gaussian mixture model and inter-frame difference in the application of classroom recording apparatus [C]//2018 Proceedings of the 2018 10th International Conference on Computer and Automation Engineering, 2018: 111–115.
- [3] LU X, XU C, WANG L, et al. Improved background subtraction method for detecting moving objects based on GMM [J]. IEEJ Transactions on Electrical and Electronic Engineering, 2018, 13(11): 1540–1550.
- [4] TEBBE J, KLAMT L, GAO Y, et al. Spin detection in robotic table tennis [C]//Paris: Proceedings of the IEEE International Conference on Robotics and Automation, 2020.
- [5] ZHANG B F, YU J H, ZHU X F, et al. Metal-YOLO detection algorithm for metal base defects in coaxial packaging [J]. Progress in Laser and Optoelectronics, 2024, 61(22): 132–140.
- [6] 张不凡, 俞经虎, 朱行飞, 等. 面向同轴封装金属底座缺陷的Metal-YOLO检测算法 [J]. 激光与光电子学进展, 2024, 61(22): 132–140.
- [7] DONG Y Q, CHENG D Q, ZHANG Y H, et al. A lightweight coal mine safety helmet detection method based on attention and reconstruction feature fusion [J]. Computer Engineering and Applications, 2024, 60(15): 297–306.
- [8] 董彦强, 程德强, 张云鹤, 等. 基于注意力和重构特征融合的轻量级煤矿安全帽检测方法 [J]. 计算机工程与应用, 2024, 60(15): 297–306.
- [9] ZHAO Y, WANG Q, SHA J, et al. Lightweight table tennis object detection algorithm based on YOLOv5s [J]. Computer Applications, 2023, 43(S1): 229–234.
- [10] 赵英, 王琦, 沙捷, 等. 基于YOLOv5s的轻量化乒乓球目标检测算法 [J]. 计算机应用, 2023, 43(S1): 229–234.
- [11] JIANG P, ERGU D, LIU F, et al. A review of Yolo algorithm developments [J]. Procedia Computer Science, 2022, 199: 1066–1073.
- [12] WANG J, CHEN K, XU R, et al. Carafe: content-aware reassembly of features [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 3007–3016.
- [13] ZHU L, WANG X, KE Z, et al. Biformer: vision transformer with bi-level routing attention [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 10323–10333.
- [14] REZATOIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: a metric and a loss for bounding box regression [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 658–666.
- [15] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression [C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2020: 12993–13000.
- [16] DU S, ZHANG B, ZHANG P, et al. An improved bounding box regression loss function based on CIOU loss for multi-scale object detection [C]//2021 IEEE 2nd International Conference on Pattern Recognition and Machine Learning (PRML), 2021: 92–98.
- [17] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506: 146–157.
- [18] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism [DB/OL]. <https://arxiv.org/abs/2301.10051>, 2023.

(责任编辑: 孙娟)