Improvement of YOLOv5 Recognition Model for Small Difference Features Collected by Sensors*

JIA Xiaoqi[1*], ZHU Qing[2]
(1.Department of Computer and Information Engineering, Shanxi Institute of Energy, Jinzhong Shanxi 030604, China;
2.Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China)

Abstract: Difference sensors demonstrate small difference in sensitivity to target features, especially in the complex background, the target has strong scale invariance, and the feature difference is small, resulting in the difficulty of target recognition in the later stage of multi-sensor acquisition. A multi-sensor keypoint pair feature object recognition algorithm based on improved YOLOv5 is proposed. Key points of target images are aligned from different perspectives collected by multiple sensors to enhance sensitivity to target features. In the Backbone of YOLOv5, multi-sensor feature data are fused through slicing and concat operations, and normalized to unify the numerical range. On the basis of feature pyramid networks(FPN), C3STR structure is introduced, combined with cross stage partial network(CSP) and Transformer multi head attention mechanism to enhance sensitivity to small difference features and achieve recognition optimization. The experimental results show that the improved YOLOv5 model has an information entropy value of up to 64 bits after image fusion, significantly improving recognition accuracy and achieving good results.

Key words: improved YOLOv5 algorithm; multi sensor; target recognition; feature fusion; key point extraction
EEACC: 7230 **doi**: 10.3969/j.issn.1004-1699.2025.11.017

传感器采集小差别特征的 YOLOv5 识别模型的改进*

贾晓琪[1*], 朱 青[2]
(1.山西能源学院计算机与信息工程系, 山西 晋中 030604; 2.北京工业大学信息学部, 北京 100124)

摘 要: 由于不同传感器对目标特征的敏感度差异较小, 特别是在复杂背景中, 目标具有强大的尺度不变性, 特征差异微小, 导致多传感器采集的目标后期识别难度大。提出基于改进 YOLOv5 的多传感器关键点对特征目标识别算法。利用多传感器采集的不同视角目标图像关键点进行对齐, 增强对目标特征的敏感度。在 YOLOv5 的 Backbone 中, 通过切片和 Concat 操作融合多传感器特征数据, 并进行归一化处理以统一数值范围。在特征金字塔网络(Feature Pyramid Networks, FPN)基础上引入 C3STR 结构, 结合跨阶段局部网络(Cross Stage Partial Network, CSP)和 Transformer 多头注意力机制, 提升对小差异特征的敏感度, 完成识别优化。实验结果显示, 改进后的 YOLOv5 模型在图像融合后信息熵值高达 64 bit, 识别精度显著提升, 效果良好。

关键词: 改进 YOLOv5 算法; 多传感器; 目标识别; 特征融合; 关键点提取
中图分类号: TP319 **文献标识码**: A **文章编号**: 1004-1699(2025)11-2055-07

在复杂多变的实际应用场景中, 单一传感器往往难以全面、准确地捕捉目标对象的所有特征, 尤其是那些细微且关键的差别特征, 这在高精度要求的领域(如自动驾驶、智能监控、工业检测等)影响巨大。YOLOv5 作为一种流行的实时目标检测模型, 以其出色的检测速度和较高的准确率受到广泛关注[1-2], 但在处理多传感器数据时, 其性能受到限制。

目前, 已有不少研究致力于解决这一问题。例如, 文献[3]提出基于集成卷积神经网络的模型, 通过数据增强和重采样技术扩充训练样本, 采用 Adadelta 算法优化网络。Adadelta 算法在优化网络时, 依赖于对梯度的计算和更新。在多传感器采集特征存在细微差异的情况下, 目标函数的曲面变得更为复杂。从原理上看, 这种细微差异会使目标函数的二阶导数产生难以预测的变化, 进而影响 Adadelta 算法中学习率的自适应调整。文献[4]则利用局部二值模式和 HOG 特征提取, 结合广义判别分析进行特征降维, 实现了多尺度特征融合。在该方法中, 局部二值模式主要关注图像局部的纹理信息, 而方向梯度直方图侧重于图像的边缘和梯度信息。在多传感器采集特征存在细微差异时, 由于不同传感器获取的图像在纹理和边缘信息上的差异并非简单叠加, 局部二值模式和 HOG 特征在融合过程中会相互干扰。文献[5]基于 YOLOv7-tiny 算法, 引

项目来源: 北京市自然科学基金项目(4232017); 2024 年山西省高等教育教改研究项目(J20241557); 山西省科技厅 2024 年第二批面上项目(202403021221353)
收稿日期: 2024-10-11 **修改日期**: 2024-11-11

入 SMSE 注意力机制和 RFB 结构,提升了检测精度。因为目标特征差异微小,SMSE 注意力机制可能无法有效捕捉到多传感器采集特征之间的细微差异,导致无法准确识别并利用这些差异,从而影响检测精度。文献[6]在 YOLOv8 基础上,设计了新的边界框回归损失函数和残差全局注意力机制,并优化了 C2f 模块。在训练过程中,由于不同传感器特征的细微差异,新的边界框回归损失函数会受到影响。从损失函数曲线形状来看,在多传感器数据下,损失函数曲线变得更加不规则。这是因为损失函数将不同传感器特征的细微差异视为重要的预测依据,在测试数据中,当遇到与训练数据中细微差异模式不同的情况时,模型性能会显著下降。

因此,提出基于改进 YOLOv5 的多传感器关键点对特征目标识别算法。其核心是精确提取与对齐多传感器多视角目标图像的关键点以增强识别敏感度。为提升性能,优化了 YOLOv5 的 Backbone 和 Neck 部分来加强特征提取与融合能力。实验表明,该算法使图像融合信息熵显著提升至 64bit,识别精度与效果大幅提高,增强了模型对复杂环境多视角目标的识别能力与系统整体性能。

1 基于多传感器的图像目标关键点对对齐

在复杂背景中,由于不同传感器(如视觉传感器、激光传感器、光纤传感器等)虽然对目标特征的敏感度在理论上存在差异,但在实际应用中,特别是当目标具有强大的尺度不变性且特征差异微小时,这些差异可能会被背景噪声、光照变化、视角变化等因素所掩盖,这使得目标与背景以及不同目标之间的区分变得困难,导致后期识别难度增大。针对这种情况,根据多传感器采集的不同视角目标图像关键点,并将其对齐,可以突出目标在不同传感器图像中的共有特征,同时减少背景噪声和其他非目标特征的干扰。这样做有助于增加小差异特征的敏感度,使得传感器能够更准确地识别目标,而不受环境因素的过多影响[8]。

多传感器采集的信息融合可以充分发挥不同传感器的优势。基于多传感器采集的关键点集合,通过计算关键点对之间的相对角度,构建直方图并找到峰值对应的旋转角度。该旋转角度用于调整不同传感器图像之间的角度,以实现图像的有效对齐和融合,获取完整的目标图像。这一步骤中,旋转角度的确定直接依赖于关键点对之间的相对位置关系,确保了图像融合的一致性和准确性。具体过程如下:

①关键点特征要求:传感器在不同光照条件下捕获的图像会有显著的亮度变化,关键点集合中的特征应对光照变化具有不变性。当传感器从不同角度观察同一目标时,目标在图像中的形状和位置会发生变化。关键点集合中的特征应该对视角变化具有鲁棒性,用(a_i,b_j)表示目标关键点对。

②构建直方图:由于多个传感器噪声会影响单个关键点的准确性,但直方图通过累积多个关键点的角度信息,可以减少噪声的影响,提高特征的鲁棒性。因此,基于构建的关键点集,对关键点实施一一对应,获取关键点对(a_i,b_j);计算关键点对之间的相对角度φ,并将这些角度信息累积到一个直方图$H(\theta)$中。通过分析直方图$H(\theta)$,找到峰值所对应的角度,该角度即为图像之间可能的旋转角度θ的估计值。

③特征子图旋转:传感器在不同位置和角度观察同一目标时,目标在图像中的外观会发生旋转和变形。通过旋转特征子图,可以使特征描述符对视角变化具有不变性。旋转特征子图是确保特征描述符在不同视角下的一致性和匹配精度。通过计算关键点对之间的相对角度并构建直方图,估计出图像间可能的旋转角度。依据此旋转角度对特征子图进行相应旋转,使得特征描述符能够在目标旋转时保持一致。这一过程不仅提高了特征匹配的鲁棒性,还减少了因视角变化导致的误匹配,从而确保了图像融合的一致性和准确性,最终获取到完整且精确的目标图像。

在图像f_1和f_2中以任意关键点对(a_i,b_j)为中心提取特征子图$T_1(f_1)$、$T_2(f_2)$,并充分考虑旋转角度θ。旋转特征子图可以确保特征描述符在目标旋转时保持一致。对待匹配图像特征子图$T_2(f_2)$实施相应旋转,获取旋转后特征子图$T_2'(f_2)$。从而计算关键点对的对齐度$\zeta(T_1(f_1),T_2'(f_2))$,过程如下式所示:

$$\zeta[T_1(f_1),T_2'(f_2)]=\frac{\eta(u,\varepsilon_{\text{high}})[T_1(f_1)x,y-\iota_1][T_2'(f_2)x,y-\iota_2]}{\sigma^1\sigma^2} \tag{1}$$

式中:特征子图标准差记作σ^1、σ^2形式,图像均值用ι_1、ι_2描述。

④确定匹配关键点对:基于特征子图特性,设定一个相似度阈值T_s,当上述对齐度计算结果高于设定阈值时,认定该对关键点为匹配状态,该关键点对即为待匹配点对。通过对齐的关键点对,完成图像融合,获取完整目标图像。

综上所述,通过不同传感器(如视觉、激光、光纤等)采集到的图像,利用关键点对齐技术,将各传感器视角下的图像进行融合,以获取一个更加完整、信息更丰富的目标图像。这种融合图像不仅包含了来自不同传感器的互补信息,还通过关键点对齐减

少了因视角差异导致的信息错位,为后续的目标识别提供了更可靠的数据基础。

2　基于改进 YOLOv5 的多传感器目标识别方法

2.1　模型的改进

在多传感器中,确保不同传感器对同一目标的识别结果一致性是一个难点。由于传感器的视角、工作原理和环境适应性不同,会导致识别结果存在差异。通过对齐后的图像融合,可以生成一个更加完整、信息更丰富且质量更高的目标图像,作为模型的输入数据。YOLOv5 作为一种在实时目标检测领域表现优异的模型,其设计允许在网络的不同层次上有效提取特征[9-11]。然而,在复杂多变的场景下,YOLOv5 仍可能面临检测精度不足的问题。为了进一步提升模型在复杂背景中的识别能力,需要对 YOLOv5 的 Backbone 网络和 Neck 部分进行优化改进。

在优化后的 YOLOv5 框架中,为了进一步提升模型对目标特征的敏感性,特别是在复杂背景下的识别能力,对模型的多个关键部分进行了深入改进[12]。具体识别模型结构如图 1 所示。

以下是针对 Backbone、Neck 网络及 BottleneckCSP 模块的详细优化描述:

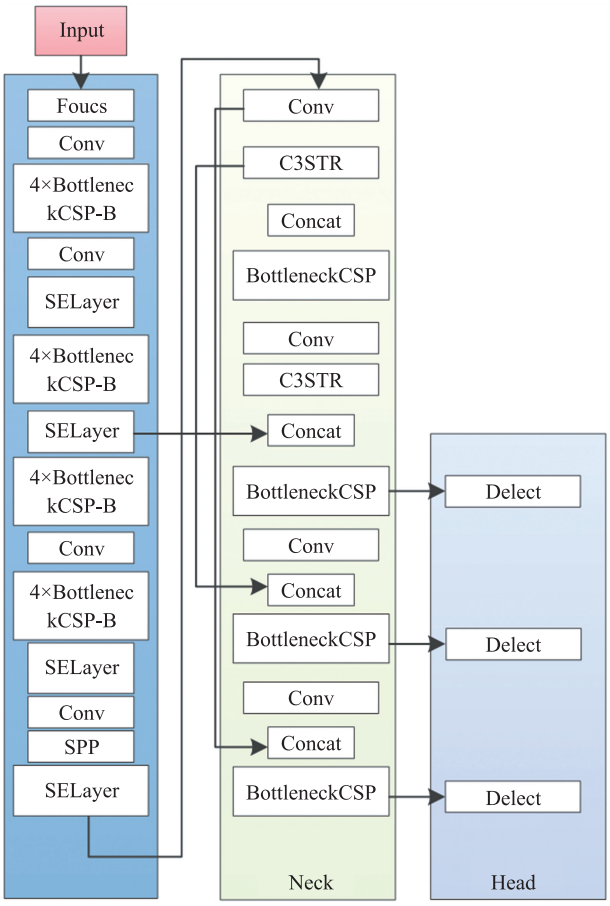


图 1　基于改进 YOLOv5 模型的目标识别架构

①数据预处理与输入

不同传感器获取的数据具有不同的分辨率和格式。其中,视觉摄像头传感器通常提供较高的空间分辨率,可以达到数百万像素,能够捕捉到丰富的颜色和纹理细节,输出的数据格式通常是二维的数字图像;激光雷达能够提供非常高的空间分辨率,尤其是在距离测量上,可以达到毫米级别的精度[13];光纤传感器的分辨率通常在微应变或微米级别,输出的数据通常是光强度、相位、波长或频率的变化。分析图 1 可知,构建的目标识别框架中,Backbone 网络为特征提取网络。此网络架构起始于 Focus 子模块,该模块通过 Concat 技术将以上不同传感器采集图像的特征切片深度融合,有效地结合不同传感器的优势,提高特征的丰富性。通过切片操作处理输入的 3 通道目标图像,并利用 Concat 技术深度融合这些切片,直接生成尺寸优化为 12×320×320 的特征图像。

②Backbone 网络优化

Backbone 网络作为 YOLOv5 模型的核心组成部分,其优化对于提升模型的特征提取能力和敏感度至关重要。在本研究中,对 Backbone 网络进行了以下几方面的改进:

Focus 模块增强:重新设计初始的 Focus 模块,以更好地融合来自多个传感器的数据。引入一种智能的预处理步骤,首先对各传感器的原始数据进行归一化处理,以确保它们处于相近的数值范围,从而消除因动态范围差异导致的潜在问题。随后,Focus 模块利用切片操作将这些归一化后的数据高效地融合在一起。

卷积层优化:Focus 模块的卷积层参数经过调整后,以最大化地利用多传感器数据的互补性。输入通道数被设置为多传感器通道数之和,确保所有传感器的信息都被充分考虑。输出通道数则设定为 64,以平衡特征提取的复杂度和计算效率。卷积核大小选择为 3×3,以捕捉局部特征,同时步长设为 2 以实现特征图的下采样,有助于减少计算量并保留关键信息。

残差块改进:为了进一步增强 Backbone 网络的特征提取能力,在 Backbone 网络中增加残差块数量,每个残差块内嵌入注意力机制。残差块通过多个卷积层和 ReLU 激活函数增强非线性特征提取能力,跳跃连接则有助于缓解梯度消失问题。注意力机制自动学习特征图的空间和通道重要性,使模型更聚焦于关键特征,提高特征的敏感度和准确性。

③Neck 网络优化

Neck 网络作为 Backbone 和 Detection Head 之间的桥梁,其优化对于提升模型的多尺度特征融合能力和检测精度同样重要。在本研究中,对 Neck 部分进行了以下创新性的改进:

C3STR 结构引入:提出 C3STR 结构,该结构在 FPN 的基础上融合了 CSPNet 的跨阶段局部连接思想和 Transformer 的多头注意力机制。C3STR 结构通过局部窗口内的多头注意力计算,有效捕获了不同传感器特征之间的长距离依赖关系,增强了特征的全局表示能力。局部窗口大小设为 7,以平衡计算复杂度和特征感受野。

④BottleneckCSP 模块

采用高效的 BottleneckCSP 模块,作为特征提取的核心。BottleneckCSP 模块的具体参数包括多个卷积层,卷积核大小为 1×1 和 3×3,通过控制输入输出通道数、步长等参数,实现特征的有效提取和压缩。BottleneckCSP 模块结构如图 2 所示。

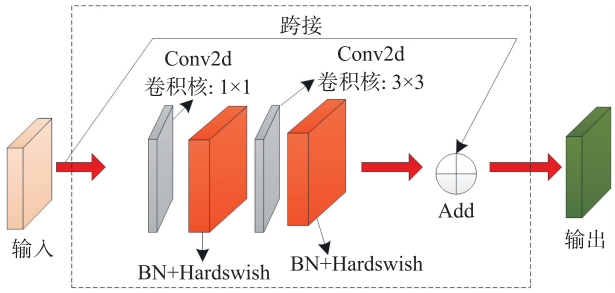


图 2 BottleneckCSP 模块结构

BottleneckCSP 模块通过减少计算量、增强特征多样性以及提高特征提取效率,展现出了在复杂场景和部分遮挡条件下的鲁棒性。此外,结合通道注意力机制,该模块在特征融合过程中能够智能地调整不同通道特征的重要性,进一步提升了特征融合的有效性和准确性。

综上所述,通过对 YOLOv5 模型的 Backbone、Neck 网络及 BottleneckCSP 模块的全面优化,模型在特征提取、融合及敏感度方面均得到了显著提升,为复杂背景下的目标识别任务提供了强有力的支持。

2.2 基于改进模型的小差异传感目标识别

在针对多传感器目标识别任务中,特别是在需要精确捕捉和识别微小差异的场景下,传统的识别方法识别效果较差。为此,本研究提出基于深度学习和多传感器融合的策略,对 YOLOv5 模型进行了针对性改进,模型改进的核心在于两大方面:一是模型结构的优化,旨在提升特征提取的精度与效率;二是特征的增强与利用,确保模型能够更有效地识别出那些细微差别的目标。

在模型结构优化方面,重新设计了 Backbone 网络,通过增强的 Focus 模块、优化的卷积层参数配置以及引入的残差块与注意力机制[14-16],使得模型能够更充分地融合多传感器数据,提取出更加丰富且有针对性的特征 $M_c(F)$,过程如下式所示:

$$M_c(F)=\varepsilon[\zeta(T_1(f_1),T_2'(f_2))]+W_1(F_{\text{avg}}^c)+W_0(F_{\text{max}}^c) \tag{2}$$

式中:激活函数描述成 ε 形式,池化层连接权重描述成 W_0、W_1 形式。

而在改进模型识别目标的部分[17-18],通过设计的 Head 部分,结合针对微小差异目标优化的损失函数 S_{loss},模型在训练过程中能够更专注于目标本身的细微特征,从而提升在复杂背景下的识别精度。目标识别结果如下式所示:

$$S_{\text{loss}}=1-\left[\text{IOU}-\frac{\gamma^2}{l^2}-\eta\lambda\,\text{Attention}(Q,K,V)\right] \tag{3}$$

式中:真实异常数据与预测值之间重叠面积描述成 IOU 形式,目标实际与预测值之间中心点距离用描述成 γ 形式,数据真实值与预测值之间宽高比相似度标记为 η,影响因子描述成 λ 形式,真实值与预测值最小包围矩形对角线长度描述成 l 形式。

基于改进 YOLOv5 的多传感器目标识别方法,通过优化 Bckbone 和 Neck 部分,结合多传感器数据融合与关键点对齐,引入通道注意力机制和 C3STR 结构,提高了模型在复杂背景下的目标识别精度和效率。

3 实验

为了验证基于改进 YOLOv5 的多传感器关键点对特征目标识别算法的整体有效性,需要对此方法进行测试。分别采用所提方法、文献[3]的集成卷积神经网络、文献[4]的局部二值模式算法展开测试。

3.1 实验设置

①硬件环境:测试过程中,使用 Intel Core i7-9700K 作为处理器,设定内存为 32 GB RAM,储存使用具有 256 GB 容量的 SSD 硬盘,操作系统使用 Ubuntu 18.04 LTS。

②传感器配置:

视觉传感器分辨率为 1 920 pixel×1 080 pixel,其帧率为 30 FPS、视角范围为水平 90°,垂直 60°,覆盖广泛的视野,适应不同场景需求。内置自动曝光和白平衡调整功能,确保在不同光照条件下图像质量稳定。传感器镜头类型选择广角或长焦镜头,以

适应不同距离和视角的目标捕捉需求。

3.2　数据集构建

实验构建了一个包含多种目标(如车辆、行人、动物等)和复杂背景(如城市街道、森林、室内环境等)的多模态数据集。具体步骤如下:①使用上述三种传感器同步采集目标图像和数据,共收集到 6 000 张原始图像,其中 5 400 张用于训练,600 张用于测试。②对每张图像进行标注,标记出目标的位置和类别,采用 PASCAL VOC 格式保存标注信息。③对数据进行预处理,包括图像裁剪、缩放、归一化等,以确保输入模型的数据一致性。

3.3　基于改进 YOLOv5 的参数敏感性分析

①确定关键参数

在基于改进 YOLOv5 的多传感器关键点对特征目标识别算法中,确定以下几个关键参数进行敏感性分析。A. 批量大小:该参数决定了每次迭代中用于训练模型的数据样本数量。B. 学习率:它控制着模型在训练过程中权重更新的步长,对模型收敛速度和最终性能有重要影响。C. 特征融合的权重:在 Neck 部分采用加权和的方式进行特征融合,这些权重直接影响到不同特征在融合后的贡献程度。

②参数取值范围设定

A. 批量大小:考虑到硬件资源(32G RAM)和模型的复杂性,设定批量大小的取值范围为[8,32],步长为 4。B. 学习率:基于初始学习率 0.001,设定取值范围为[0.0005,0.002],步长为 0.0005。C. 特征融合的权重:对于参与融合的不同特征,设定权重的取值范围为[0.1,0.9],步长为 0.1。

③进行敏感性分析实验

A. 单一参数变化实验

a. 批量大小:在保持学习率和特征融合权重固定的情况下,对批量大小在[8,32]范围内以步长为 4 进行取值变化实验。实验结果表明,随着批量大小的变化,模型的平均精度均值波动幅度相对较小。这表明模型对批量大小的敏感性较低,在该取值范围内具有一定的鲁棒性。b. 学习率:在固定批量大小和特征融合权重的条件下,将学习率在[0.0005,0.002]范围内以步长为 0.0005 进行取值变化实验。实验结果表明,学习率在取值范围内变化时,模型的平均精度均值虽有波动,但整体仍能保持相对稳定的性能,表现出一定的鲁棒性。不过,相比批量大小的波动幅度稍大一些。c. 特征融合的权重:在设定批量大小和学习率不变的情况下,改变特征融合权重在[0.1,0.9]范围内以步长为 0.1 进行取值变化实验。实验结果表明,当特征融合权重变化时,模型的平均精度均值有一定波动,波动幅度相对前两个参数稍大,但模型仍然能够在不同权重设置下保持一定的性能水平,显示出对特征融合权重的一定鲁棒性。

B. 多参数组合变化实验

选择批量大小、学习率和特征融合权重三个参数进行组合变化实验,批量大小取值为[8,16,32],学习率取值为[0.0005,0.001,0.002],特征融合权重取值为[0.1,0.5,0.9],共形成 27 种参数组合。实验结果表明,从多参数组合变化的实验结果来看,尽管三个关键参数同时发生变化,模型的平均精度均值的波动范围相对单一参数变化时有所增大,但模型仍然能够在不同的参数组合下保持一定的性能水平,进一步证明了模型对不同参数设置具有较好的鲁棒性。

通过以上对模型关键参数的敏感性分析,全面地展示了基于改进 YOLOv5 的多传感器关键点对特征目标识别算法在不同参数设置下的鲁棒性,为模型的实际应用和进一步优化提供了重要依据。

3.4　训练细节

①模型选择:基于 YOLOv5 模型进行改进,主要优化 Backbone 和 Neck 部分,以更好地适应多模态数据的输入和特征融合。

②训练参数:A. 批量大小:16;B. 学习率:初始化为 0.001,采用余弦退火策略进行调整;C. 优化器:AdamW;D. 训练轮次:100;E. 数据增强:随机翻转、旋转、亮度调整等。

③特征融合策略:在 Neck 部分采用加权和的方式进行特征融合。

3.5　实验结果与分析

①多传感器关键点特征提取效果测试

在针对多传感器关键点特征目标识别的实验中,特别关注了所提方法在多传感器提取目标关键点时的完整性和对特征小差异的捕捉能力。通过改进 YOLOv5 模型,在其特征提取网络中嵌入了增强的特征金字塔结构和注意力机制,以更精细地捕捉目标的关键点及细微特征。测试结果如图 3 所示。

分析图 3 可知,所提出的方法在目标关键点提取方面表现卓越。左侧以实线清晰标注了实际目标的轮廓,而右侧则以虚线精确展示了该方法提取的轮廓关键点。所提方法不仅能够有效提取目标的轮廓关键点,而且在关键点完整度上表现出色,与实际目标轮廓高度吻合。在识别具有微小差异的目标时,所提方法展现出了更高的敏感性和准确性。本文方法之所以在目标关键点提取方面表现卓越,主要原因在于

图 3 所提方法目标关键点提取完整度测试结果

根据多传感器采集的不同视角目标图像关键点,并将其对齐,以此提高了不同传感器对目标特征的敏感度,使得模型能够更好地区分并识别这些细微差别。

②多传感器图像融合效果

为了验证多传感器融合在提升识别精度方面的作用,实施了多传感器图像融合实验。通过融合来自不同视角和类型的视觉传感器数据(如普通摄像头、红外摄像头等),获得了更加全面和丰富的目标信息。测试结果如图 4 所示。

分析图 4 可知,所提方法在多传感器融合后,图像信息熵能够稳定维持在较高水平,表明融合过程有效地保留了图像中的关键信息。多传感器的融合不仅增强了目标识别的鲁棒性,还通过不同传感器之间的特征互补,提高了对特征小差异的识别能力。

③实际识别效果测试

在多传感器关键点对特征目标识别中,不同的传感器具有不同的物理特性和工作原理。摄像头捕捉的是可见光图像,对于颜色和纹理丰富的目标识别效果较好;而红外传感器则对温度变化敏感,适合于夜间或热目标的检测。因此,传感器的选择会影响其对特定类别目标的识别能力。选取 600 张图像

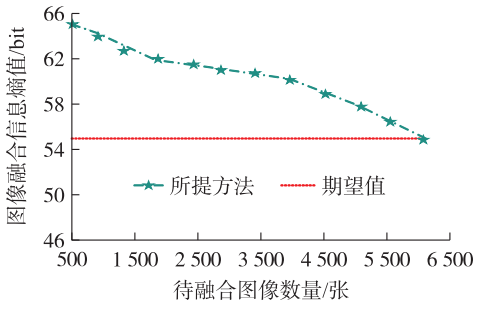


图 4 所提方法多传感器融合效果测试结果

识别样本,以平均交并比(Mean Intersection over Union,mIoU)用于预测识别边界框与真实目标边界框的重叠程度,以检测目标识别的准确性。mIoU 的取值范围在 0 到 1 之间,数值越接近 1 表示识别效果越好。3 种不同方法的结果如表 1 所示。

表 1 不同方法的 mIoU 指标结果对比

图像识别样本数量/张	所提方法	基于集成卷积识别方法	局部二值模式方法
1 000	0.94	0.82	0.81
2 000	0.95	0.80	0.78
3 000	0.90	0.79	0.82
4 000	0.92	0.84	0.75
5 000	0.89	0.83	0.84
6 000	0.96	0.78	0.78

分析表 1 可知,所提出的方法在所有样本数量下均展现出了最高的 mIoU 值,特别是在处理大规模样本集时,其性能优势尤为显著。这一表现充分验证了所提方法在多传感器环境下的有效性,以及对 YOLOv5 模型的成功扩展与优化。这表明该方法在 Bckbone 部分创新性地采用了切片和 Concat 操作,实现了多传感器采集特征数据的高效融合。这一步骤不仅丰富了模型的特征输入,还促进了不同传感器间信息的互补与增强。随后,在特征金字塔网络(FPN)上,引入了 C3STR 结构,这一结构结合了跨阶段局部网络(CSP)的高效特征提取能力和 Transformer 多头注意力机制的强大特征区分能力。这样的设计使得模型在提取丰富特征的同时,能够更加精准地识别并关注细微特征,显著提升了对小差异特征的敏感度。最终,通过特征与关键点的深度融合,该方法成功捕获了目标的轮廓特征,从而在目标识别任务中实现了高精度与优异的效果。

4 结束语

针对复杂环境下目标识别难度增大的挑战,本研究提出一种融合多传感器数据且深度改进 YOLOv5 算法的目标识别框架。该方法整合不同传感器图像关键点,经切片与 Concat 操作在 Backbone 部分高效融合多源特征,结合归一化处理使不同敏感度传感器数据被有效利用,引入 C3STR 结构、CSP 与 Transformer 多头注意力机制增强模型区分小差异特征能力,优化识别效果。

在实际的传感领域,本研究为不同类型传感器(像视觉传感器、红外传感器等)的协同工作树立了典范。在复杂环境下,例如在恶劣天气状况下(无论是白天的强光照射、夜晚的低光照,还是雨、雪、雾

等天气条件),智能安防系统借助本研究成果能够准确识别潜在威胁目标;在自动驾驶领域,即使面临复杂地形(如山区弯道多、丛林地区视野受限等)以及干扰源(如其他车辆的电磁干扰、路边相似物体的干扰等)场景,也能够更好地识别道路上的车辆、行人、交通标志等物体,从而提升行驶的安全性。此外,本研究对于提升传感系统的智能化水平具有重要意义。通过优化算法,不仅提高了目标识别的效率,而且提升了识别的精度,进而为决策系统提供了高质量的数据支持,有力地推动传感领域朝着智能高效的方向发展,为多传感器融合的智能识别系统提供了崭新的思路与方法。

参考文献:

[1] 刘艺,吴路路,邓湘琳,等. 基于改进 YOLOv5s 小目标检测算法[J]. 安徽科技学院学报,2024,38(4):69-77.

[2] 姚珊珊,王静宇,郝斌,等. 基于改进 YOLOv5 的超分辨率和多尺度融合目标检测算法[J]. 光电子·激光,2024,35(8):793-802.

[3] 李汪华,张贞凯.一种基于集成卷积神经网络的 SAR 图像目标识别算法[J]. 电讯技术,2023,63(12):1918-1924.

[4] Sun G,Kuang B,Zhang Y.Fast Target Recognition Method Based on Multi-Scale Fusion and Deep Learning[J]. Traitement du Signal,2022,39(6):2173-2179.

[5] 薛珊,安宏宇,吕琼莹,等. 复杂背景下基于 YOLOv7-tiny 的图像目标检测算法[J]. 红外与激光工程,2024,53(1):269-280.

[6] 张秀再,沈涛,许岱.基于改进 YOLOv8 算法的遥感图像目标检测[J]. 激光与光电子学进展,2024,61(10):298-308.

[7] 宁小娟,李春旭,王嘉豪,等. 基于关键点精确配对的点云曲面匹配方法[J]. 系统仿真学报,2023,35(6):1169-1182.

[8] 张学峰,马行,穆春阳.基于四叉树的 SIFT 与 K-D 树融合的图像匹配研究[J]. 电子测量技术,2021,44(22):121-127.

[9] 李佩,韩芳,杨凯,等. 基于改进 YOLOv5s 算法的桥梁裂缝自动检测及分类[J]. 交通科技,2024(3):53-58.

[10] 王伯涛,周福强,吴国新,等. 基于改进 YOLOv7 的换热器板片故障检测算法研究[J]. 机床与液压,2024,52(11):105-113.

[11] 徐微,郝琦琦,李波波,等. 一种改进 YOLOv5 积木小零件检测算法研究[J]. 电子器件,2024,47(4):1116-1120.

[12] Wei Z,Wang T. Joint attention mechanism with dynamic kernel for YOLOv5 mobile wireless charging coil surface defect identification[J]. Multimedia tools and applications,2024,83(4):12403-12424.

[13] Feng Y N,Zhu J H,Song R X. S2EFT:Spectral-Spatial-Elevation Fusion Transformer for hyperspectral image and LiDAR classification[J]. Knowledge-based systems,2024,283:111190.

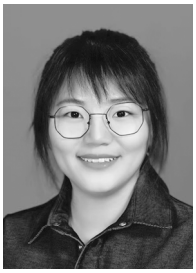
[14] 常开心,侯彦东,陈政权,等. 基于多尺度双注意力网络的植物病虫害识别[J]. 计算机仿真,2024,41(4):175-179.

[15] Chen M,Chen K,Liu C,et al. Theoretical and experimental study on image noise reduction for improving camera-based fire detection performance in thermal environments[J]. Journal of Thermal Analysis and Calorimetry,2023,148(3):1191-1199.

[16] 王娟娟,黄炜,马生菊. 结合卷积注意力机制改进 YOLOv5s 的垃圾检测[J]. 电子器件,2024,47(5):1434-1140.

[17] 陈乔松,周丽,毛彦嵋,等. 基于浅层空间特征融合与自适应通道筛选的目标检测方法[J]. 江苏大学学报(自然科学版),2022,43(1):67-74.

[18] Cheng H,Yang Z,Zhang X,et al. Multimodal Sentiment Analysis Based on Attentional Temporal Convolutional Network and Multi-Layer Feature Fusion[J]. IEEE transactions on affective computing,2023,14(4):3149-3163.



贾晓琪(1992—),女,山西翼城人,本科毕业于太原理工大学,研究生毕业于北京工业大学。现就职于山西能源学院,讲师,主要从事图像处理、数字媒体技术等方面的研究工作。结合图像处理领域和教学改革以第一作者发表论文 6 篇,参与省级自然科学基金项目 1 项,主持及参与省级教改课题 2 项,主持横向课题 1 项,jiaxiaoqi2313@163.com;

朱　青(1965—),北京工业大学教授/博导。主要从事图像处理、计算机视觉、虚拟现实及交互技术等方面的研究工作。发表论文 70 余篇,获授权发明专利 7 项。负责及参与完成了 863 和 973 项目、国家自然科学基金、北京市自然基金等 10 余项纵向科研项目,同时面向企业亟待解决的技术和应用问题开展应用研究,为企业解决实际相关应用项目 20 余项。