

DOI:10.19651/j.cnki.emt.2519854

针对无人机作战复杂前端目标辨识定位方法^{*}付羿焱^{1,2} 孙传猛^{1,2,3} 孔祥年^{1,2} 李勇³ 靳鸿^{1,2}

(1.中北大学极限环境光电动态测试技术与仪器全国重点实验室 太原 030051;2.中北大学电气与控制工程学院 太原 030051;3.重庆大学煤矿灾害动力学与控制全国重点实验室 重庆 400044)

摘要: 针对复杂战场环境和有限机载资源,导致无人机作战时前端目标辨识定位难以兼顾准确性与实时性的问题,构建了一种针对无人机作战复杂战场环境下前端目标辨识定位方法:以“backbone-neck-head”为基本网络架构,引入非局部注意力扩展模块、全局多尺度解耦网络以及轻量化瓶颈模块,并以 Focal Loss 和 DIoU Loss 为综合损失函数,实现特征建模和多尺度检测增强,以提升对特征的捕捉能力,从而提升准确性;基于依赖图结构化剪枝与通道智能蒸馏,提出一种协同的轻量化策略,从而有效降低了模型复杂度并提升嵌入式可部署性。相关实验表明,本文方法在 mAP@0.5、mAP@0.75 和 mAP@0.5:0.95 分别提升了 6.0%、7.2% 和 5.9%,模型参数量与 GFLOPs 分别降至 17.1% 与 12.0%,精度损失控制在 4.1% 以内。最后在嵌入式硬件平台上的部署验证显示,推理帧率达到了 34 fps,能够较好地满足无人机作战时前端目标辨识与定位的准确度与实时性的需求。

关键词: 无人机;前端目标辨识与定位;非局部注意力;多尺度解耦;轻量化

中图分类号: TP193;TN98 **文献标识码:** A **国家标准学科分类代码:** 520.20

Front-end UAV method for target recognition and localization in complex combat environments

Fu Yiyan^{1,2} Sun Chuanmeng^{1,2,3} Kong Xiangnian^{1,2} Li Yong³ Jin Hong^{1,2}

(1.State Key Laboratory of Extreme Environment Optoelectronic Dynamic Measurement Technology and Instrument, North University of China,Taiyuan 030051, China;2.College of Electrical and Control Engineering, North University of China, Taiyuan 030051, China;3.State Key Laboratory of Coal Mine Disaster Dynamics and Control, Chongqing University,Chongqing 400044, China)

Abstract: To address the challenge of balancing accuracy and real-time performance in front-end target recognition and localization for drones in complex battlefield environments with limited onboard resources, a front-end target recognition and localization method for drone operations in complex battlefield environments was developed: Using a “backbone-neck-head” as the basic network architecture, a non-local attention expansion module, a global multi-scale decoupled network, and a lightweight bottleneck module were introduced. Focal Loss and DIoU Loss were employed as the combined loss functions to achieve feature modeling and multi-scale detection enhancement, thereby improving the ability to capture features and enhancing accuracy; based on dependency graph-structured pruning and channel-wise knowledge distillation, a collaborative lightweight strategy was proposed, effectively reducing model complexity and improving embedded deployability. Experiments show that this method improved mAP@0.5, mAP@0.75, and mAP@0.5:0.95 by 6.0%, 7.2%, and 5.9% respectively, while reducing model parameters and GFLOPs to 17.1% and 12.0%, with precision loss controlled within 4.1%. Finally, deployment validation on embedded hardware demonstrated a frame rate of 34 fps, effectively meeting the accuracy and real-time requirements for front-end target recognition and localization during drone operations.

Keywords: UAV; front-end target recognition and localization; non-local attention; multi-scale decoupling; lightweight

0 引言

在俄乌冲突和纳卡冲突等局部战争中,无人机^[1-3]的广

泛运用正在深刻重塑现代战争的战术体系与交战模式,作为典型的新型作战平台,无人机具有成本低、体积小且杀伤力强的突出优势,能够以有限投入实现高频次出动和规

收稿日期:2025-09-14

*基金项目:煤与煤层气共采全国重点实验室开放基金(2024KF30)、重庆大学煤矿灾害动力学与控制全国重点实验室开放基金(2011DA105287-FW202408)、复杂恶劣环境下履带式车辆主动轮实车扭矩信息获取方法开放基金(202403021211076)、国家自然科学基金青年科学基金(62401521)、仪器科学与动态测试教育部重点实验室开放基金(JYBSYSKFJJ319002)项目资助

模化使用,从而显著提升作战效费比,其小尺寸与低特征赋予良好的隐蔽性和机动性,便于在复杂地形下实施灵活突防与分散部署。在作战效能方面,无人机不仅可承担战场侦察、边境巡防等任务,还能依托边缘智能^[4]在机载端完成目标识别与态势解析,缩短“观察-取向-决策-行动”环路,有效提升实时反应与生存能力。由此,无人机逐渐成为兼具高性价比、灵活突防与智能打击特性的关键力量,并在战术侦察、火力引导及战损评估等任务中展现出不可替代的作战价值。

然而,现代战场背景干扰强、伪装遮蔽多和目标体积小等问题均会不同程度地降低目标辨识定位的准确度,对战场态势感知造成严重干扰,同时,受限于无人机平台搭载资源,其前端感知模块必须具备高度轻量化与嵌入式可部署性。因此,发展一种针对无人机作战环境的融合高精度实时检测能力与可部署的嵌入式战场目标辨识与定位方法,已成为当前军事科学研究领域的研究前沿,对后续战场目标进行实时跟踪、军事预警等战场态势感知技术具有重要理论意义和应用价值。

江波等^[5]提出了一种轻量化的无人机视觉定位系统,通过开发融合状态空间模块的图像特征提取框架,提高检测精度。杨婷等^[6]提出一种基于改进 YOLOv5 的航拍目标检测算法,引入加权双向特征金字塔网络结构以及自适应空间特征融合模块降低漏检率。王宗阳等^[7]提出了一种基于 YOLOv8 改进的无人机高空检测算法,引入自适应空间通道金字塔注意力机制提取多尺度空间信息。上述方法虽具有较高精度,但在战场环境中面对强干扰、遮蔽与小目标等情况时,候选区域生成易受影响,对目标尺度变化适应性较弱,可能导致漏检或检测不准。

综上所述,尽管已有的针对无人机的目标识别方法在一般场景下已经取得了一定的进展,但是在面对干扰强、遮蔽多、多尺度^[8]的复杂战场环境时,仍然面临重大挑战。鉴于此,本文提出一种可以部署于低功耗边缘端嵌入式^[9]设备中的高识别精度、高响应速度的针对无人机复杂作战环境的目标辨识与定位系统,该系统包括目标检测算法改进、轻量化策略设计和边缘端硬件平台部署 3 个关键环节,围绕战场目标辨识和定位的准确性、实时性与边缘端部署可用性两大核心目标,基于 YOLO 算法基本架构,结合多层次算法结构优化与系统级融合部署策略,实现从图像采集、目标检测、三维定位到结果输出的端到端闭环流程。

1 模型设计

1.1 无人机作战环境分析

典型作战环境中,观测条件相对稳定、目标尺度适中,成像清晰度与时间连续性相对容易保障,大型作战平台算力与通信链路相对充裕,可依赖后端处理与多轮复核以提

升辨识与定位精度,感知与决策过程具有一定冗余与容错空间。相比之下,无人机作战环境具有更强的复杂性和约束性,其核心挑战集中在准确性与实时性两方面。

无人机作战环境的复杂性体现在物理扰动频繁,易造成目标特征模糊与掩盖。如图 1(a)所示,爆炸闪光、烟雾粉尘等干扰因素广泛存在,导致图像亮度分布不均与结构畸变,显著削弱关键区域的清晰度与一致性,导致检测算法的准确性降低。同时,伪装手段的普遍应用进一步削弱了目标显性特征,如图 1(b)所示,迷彩涂装与遮蔽物等使目标与背景在颜色、结构及纹理^[10]上差异减弱,表现为边界模糊、轮廓不清和纹理弱化,增加了特征提取^[11]与判别的难度,在 NOMAD 的分档评测中,遮挡从 100% 可见降至 30% 时,主流检测器的 mAP@0.5:0.95 下降到个位数。此外,战场目标类别多样,尺度与结构差异显著,如图 1(c)所示,不同类型目标在空间尺度、形态结构与纹理复杂度上存在较大差异,尤其在远距离观测条件下,目标常以小尺寸形式呈现,像素占比极低,在 VisDrone 数据集上的实验表明,小目标的平均检测精度仅为 26.7%,易被背景淹没,从而导致漏检。



图 1 战场复杂环境图像特征

Fig. 1 Image characteristics in complex battlefield environments

无人机作战环境具有更强的约束性。首先,目标通常具备较强的机动性和环境适应性,无人机须在极短时间内完成信息采集、辨识与定位,并迅速反馈结果,以保障实时锁定和持续跟踪。其次,通信链路极不稳定,常伴随干扰与遮蔽,大型作战平台可依赖 GEO 卫星,但该链路往返延迟通常大于 500 ms,无法实时响应,LEO 系统延迟小于 100 ms 但在电磁干扰或网络拥堵场景中容易失稳,依赖云端或远程服务器进行推理的方式无法满足无人机作战的需求。最后,无人机自身算力、存储及并行处理能力有限,同时还需兼顾飞控、导航等多项任务,使得无人机前端视觉感知系统必须要控制模型复杂度。

1.2 算法改进策略

无人机作战环境中干扰强、遮蔽多且目标尺度小,且前端算力受限,因此检测算法必须同时具备高精度与高实时性。相比其他单阶段检测网络或纯 Transformer^[12] 架构,“Backbone-Neck-Head(BNH)”框架在此类场景下更具优势,Backbone 网络能够逐层提取稳健的语义特征^[13]以抵御爆炸闪光、烟雾粉尘等干扰造成的图像退化。Neck 网络通过跨尺度与上下文融合,有效缓解目标尺寸小、尺度

差异大及部分遮蔽带来的检测困难。Head采用解耦式分类与回归,使类别判别与边界框回归在复杂背景中更加准确。除此之外,其模块化与层次化设计不仅有利于后续插入优化机制,也便于轻量化与压缩,为无人机边缘端的实时部署提供了结构优势。

综上所述,本文选择YOLO系列中精度较高且与TensorRT量化部署链路契合的YOLOv8^[14]作为基线模型,并针对战场目标检测中特征提取弱、噪声干扰强、多尺度建模不足等问题,对网络各层进行针对性改进,改进后的模型(lightweight multi-scale attention you only look once, LMA-YOLO)其结构图如图2所示。

在Backbone网络,引入了基于注意力混合的轻量化瓶颈模块(lightweight attention fuse CSP bottleneck module, LAF Module)和非局部注意力扩展模块(non-local attention extension block, NAE Block);在Neck网络引入多尺度解耦融合模块(multi-scale decoupling fusion block, MDF Block);在Head网络,采用无锚点框架且引入Focal Loss与DIoU Loss综合损失函数,具体改进为:

1) 针对现有模型在战场目标检测任务中存在目标特

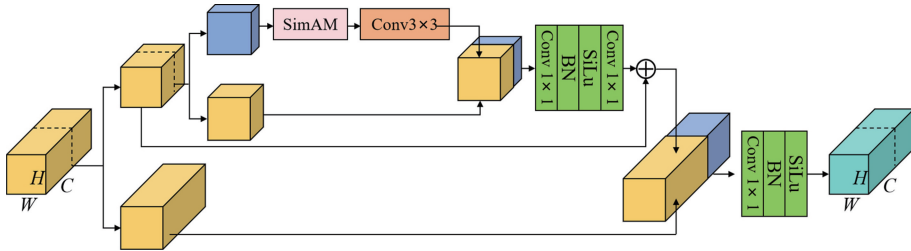


图3 轻量化瓶颈模块结构示意图

Fig. 3 Schematic diagram of the lightweight bottleneck module structure

为衡量特征图中每个神经元的重要性,在SimAM模块构造了基于能量最小化的注意力机制,其核心思想是:若某个神经元与其所在通道的其他神经元差异越大,则其更重要,最小化能量函数如下:

$$e_{\beta}^{\wedge} = \frac{4(\hat{\eta}^2 + \lambda)}{(\beta - \varphi)^2 + 2\hat{\eta}^2 + 2\lambda} \quad (1)$$

其中, β 表示当前通道位置的特征值, φ 是同一通道该空间邻域的特征平均值, $\hat{\eta}^2$ 代表该位置对应的方差, λ 为正则化常数。为增强输入特征 \mathbf{X} ,本文通过SimAM使用sigmoid函数归一化能量权重,并对特征图进行加权处理,其公式为:

$$\hat{\mathbf{X}} = \text{sigmoid}(\mathbf{E}_{\beta}) \odot \mathbf{X} \quad (2)$$

其中, \mathbf{E}_{β} 是当前通道所有神经元的重要性权重, $\hat{\mathbf{X}}$ 是加权重标后的特征值。先用 \mathbf{E}_{β} 评估每个位置的重要性,再经sigmoid变成0~1的权重,对原始特征 \mathbf{X} 做逐元素加权得到 $\hat{\mathbf{X}}$,便于网络提取显著性特征。

2) 针对深层卷积网络全局信息特征提取不足的问题,

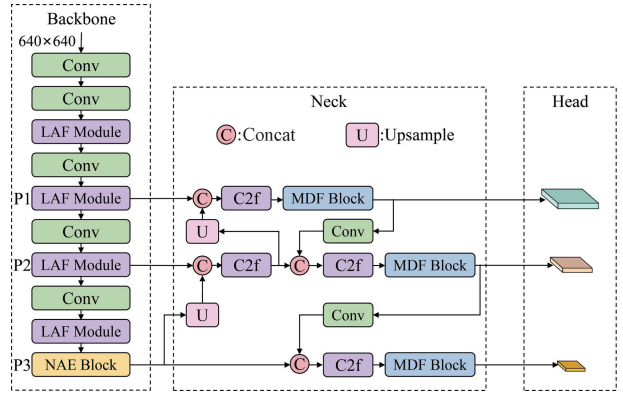


图2 LMA-YOLO结构示意图

Fig. 2 Schematic diagram of the LMA-YOLO structure

征建模效率低等问题,引入基于注意力混合的轻量化瓶颈模块提高模型特征建模效率,结构如图3所示,将输入特征图 $\mathbf{X} \in \mathbf{R}^{H \times W \times C}$ 按照通道进行拆分,并划分成原始通道和注意力通道两部分,其中,注意力通道部分将输入特征进一步划分为 $C/2 \times \vartheta$ 和 $C/2 \times (1 - \vartheta)$ 部分,并应用 $k \times k$ 卷积来提取空间特征。

引入基于非局部注意力的扩展模块,其结构如图4所示。

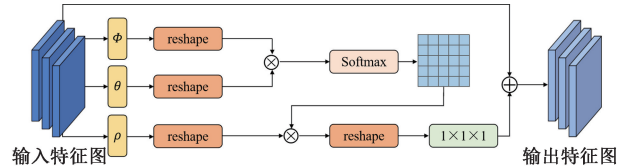


图4 非局部注意力扩展模块结构示意图

Fig. 4 Schematic diagram of the non-local attention expansion module

该模块通过全局表示来捕获深层卷积网络中的长距离依赖关系,并提高模型的全局建模能力,进而提高模型检测不同目标特征的能力和网络的鲁棒性,非局部注意力扩展模块表达式为:

$$\mathbf{Y}_i = \frac{1}{\sum_{v_j} (\mathbf{W}_{\theta} \mathbf{X}_i)^{\top} (\mathbf{W}_{\phi} \mathbf{X}_j)} \sum_{v_j} ((\mathbf{W}_{\theta} \mathbf{X}_i)^{\top} (\mathbf{W}_{\phi} \mathbf{X}_j)) \mathbf{W}_{g_s}(\mathbf{X}_j) \quad (3)$$

其中, i 是输出位置在空间中的索引, j 是所有可能位置的索引, \mathbf{X} 是输入图像信号, \mathbf{X}_j 表示在对应位置的特

征, Y_i 表示模块输出在对应位置的特征。 $(W_\theta X_i)^T (W_\phi X_j)$ 用于计算 i 和 j 之间的标量, 定义为两个点的内积, 表示两位置之间的相似度, W_θ, W_ϕ, W_g 是可学习的权重矩阵。与只考虑局部邻域加权合并的卷积网络不同, 非局部注意力扩展模块考虑了所有位置 $\forall j$, 具备更强的捕获上下文信息的能力和长距离特征关系建立能力。

3) 针对 C2f 模块多尺度信息建模能力不足的问题, 本文设计了一种多尺度解耦融合模块, 旨在提高模型对多尺度特征的提取能力, 并增强网络对小目标的鲁棒性和精度, 其结构如图 5 所示。

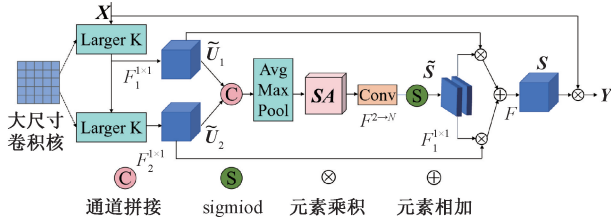


图 5 多尺度解耦融合网络结构示意图

Fig. 5 Schematic diagram of the multi-scale decoupled fusion network

该网络结构可以自适应调整卷积核的大小从而应对不同目标, 并且基于多尺度特征进行多尺度目标的上下文信息重建。为获得不同范围内具有丰富上下文信息的输入图像特征, 本文使用一系列具有不同感受野的深度卷积进行分解, 分解过程为:

$$U_0 = X, U_{i+1} = F_i^{d \cdot w}(U_i) \quad (4)$$

$$\tilde{U}_i = F_i^{1 \times 1}(U_i), i \in [1, N] \quad (5)$$

$$\tilde{U} = [\tilde{U}_1; \dots; \tilde{U}_N] \quad (6)$$

$$\widehat{SA} = F^{2 \rightarrow N}([P_{\text{arg}}(\tilde{U}); P_{\text{max}}(\tilde{U})]) \quad (7)$$

$$Y = F(\sum_{i=1}^N (\widehat{SA}_i \cdot \tilde{U}_i)) \cdot X \quad (8)$$

其中, X 表示输入特征, $F_i^{d \cdot w}(\cdot)$ 是核为 k_i 、膨胀系数为 d_i 的深度卷积。首先将输入特征 X 依次传入多个具有不同膨胀率的深度卷积核 $F_i^{d \cdot w}(\cdot)$, 形成多尺度特征序列。为统一各分支输出的通道数并增强表达能力, 进一步对每个特征 U_i 应用 1×1 卷积 $F^{1 \times 1}(\cdot)$ 进行压缩处理, 得到整合后的特征 \tilde{U}_i 。随后, 将所有分支的特征 \tilde{U}_i 沿通道维度拼接为融合特征 \tilde{U} , 以便进行统一的空间注意力建模。

在此基础上, 本文通过平均池化^[15]与最大池化^[16]提取上下文信息, 并经由卷积映射 $F^{2 \rightarrow N}$ 生成空间注意力图 SA , 用于刻画不同位置的显著性权重。最后, 将每个注意力图与其对应的分支特征相乘加权后进行融合, 并与原始输入特征 X 相乘, 得到最终输出特征 Y 。

4) 为了更好的帮助模型获取多尺度特征, 并提高模型对小目标的检测性能, 本文引入了焦点损失^[17] (Focal Loss) 和 DIoU^[18] 损失。在模型中输出特征图被卷积以生成预测的类别概率和边界框回归值, 对于每个预测框, 计

算它与真实框之间的分类损失, 预测框的类概率为 P_i , 真实标签为 y_i , Focal Loss 的计算公式为:

$$L_{Focal}(p_i) = -\alpha_i(1-p_i)^\gamma \log(p_i) \quad (9)$$

其中, α_i 是样本权重, γ 是调整因子, 当 P_i 接近 1 时, $(1-P_i)y_i$ 接近 0, 从而减少易分类样本对总损失的影响, DIoU 损失的计算公式为:

$$L_{DIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} \quad (10)$$

其中, $\rho(b, b^{gt})$ 表示预测边界框中心点 b 与真实边界框 b^{gt} 之间的欧几里得距离, c 表示两个边界框的最小边界矩形的对角线长度, IoU 是交并比, 用于衡量预测边界框和真实边界框的重合程度。综合损失函数的表达式为:

$$L = \lambda_{Focal} L_{Focal} + \lambda_{DIoU} L_{DIoU} \quad (11)$$

其中, λ_{Focal} 和 λ_{DIoU} 是焦点损失和 DIoU 损失的权重因子, 在实验中取 0.6 和 0.4。综合损失函数的应用使得模型在训练时可以同时关注分类精度和边界框回归精度, 提高了模型的精度并保障了模型的鲁棒性。

1.3 轻量化策略

为满足作战实时性的要求, 本文先对改进后的模型用基于依赖图的结构化模型剪枝和基于通道知识的模型蒸馏方法进行体积压缩, 使其能够部署于边缘端嵌入式设备, 模型剪枝轻量化结构如图 6 所示。

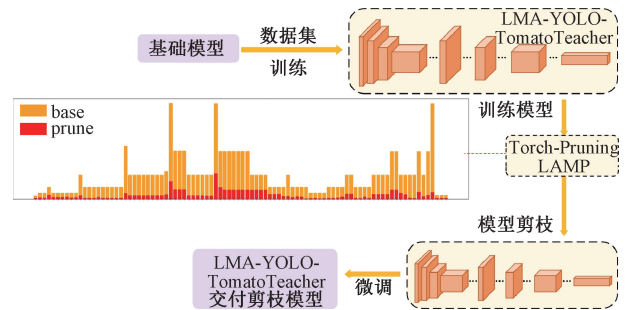


图 6 模型剪枝通道对比图

Fig. 6 Comparison chart of model pruning channels

基于依赖图自适应幅度的剪枝方法在按照升序对局部重要性评分进行排序后, 对其进行归一化以评估相对重要性, 最终全局剪枝基于方法评估的重要性进行, 将整个网络中选择归一化重要性最小的通道进行修建, 直到达到预定的稀疏目标为止。

剪枝后的模型虽然降低了参数量, 但是检测精度也受到影响, 因此本文采用通道智能蒸馏来恢复修剪所造成的性能损失, 将剪枝前的 LMA-YOLO 模型作为教师模型, 剪枝后的 LMA-YOLO-P 模型作为学生模型, 进行通道智能蒸馏, 其结构如图 7 所示。

由于剪枝后的 LMA-YOLO-P 和教师模型之间的通道数量不匹配, 因此使用 1×1 卷积来上采样学生模型的通道数量。每个通道的激活图进行归一化, 得到软概率图, 表示每个位置在通道内的相对重要性或响应强度。将

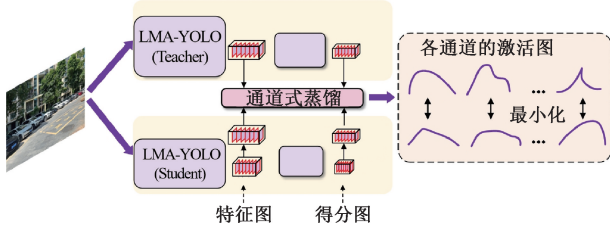


图7 模型蒸馏轻量化结构

Fig. 7 Structure of model distillation for lightweighting

教师和学生网络分别表示为 T 和 S , 相应的激活图表示为 \mathbf{y}^T 和 \mathbf{y}^S , 通道蒸馏损失公式为:

$$\varphi(f(\mathbf{y}^T), f(\mathbf{y}^S)) = \varphi(f(\mathbf{y}_c^T), f(\mathbf{y}_c^S)) \quad (12)$$

其中, $\varphi(\cdot)$ 被应用于将激活值转换为概率分布, 如式

(13) 所示。

$$\varphi(\mathbf{y}_c) = \frac{\exp(\frac{\mathbf{y}_{c,i}}{\tau})}{\sum_{i=1}^{w \cdot h} \exp(\frac{\mathbf{y}_{c,i}}{\tau})} \quad (13)$$

其中, $c = 1, 2, \dots, C$ 表示通道索引; i 表示通道的空间位置。 τ 是温度超参数。然后, 将教师和学生网络对应通道概率分布之间的不对称 Kullback-Leibler (KL) 最小化, KL 散度表示为:

$$\varphi(\mathbf{y}^T, \mathbf{y}^S) = \frac{\tau^2}{C} \sum_{c=1}^C \sum_{i=1}^{w \cdot h} \varphi(\mathbf{y}_{c,i}^T) \cdot \log \left[\frac{\varphi(\mathbf{y}_{c,i}^T)}{\varphi(\mathbf{y}_{c,i}^S)} \right] \quad (14)$$

对于每个通道, 引导学生模型更关注模仿具有显著激活值的区域, 从而在密集预测任务中实现更准确的定位。

2 实验结果及分析

2.1 实验方案

首先, 数据集增强。本实验所用军事目标数据来源于 Kaggle 科学竞赛平台和中国军事网, 共计图片 4 121 张, 为防止过拟合, 对数据集进行以下增强: 随机裁剪原图像的 0.7~0.9 倍并确保目标核心特征保留 $\geq 85\%$, 围绕中心旋转 $-15^\circ \sim +15^\circ$ 且目标姿态偏移 $\leq 15^\circ$, 叠加均值为 0、方差 0.01~0.03 (信噪比 25~30 dB) 的高斯噪声, 以及密度 0.01~0.02 (污染像素 $\leq 2\%$) 的椒盐噪声, 最终将数据集扩充至 16 806 张, 且将标注后的图像按比例按照分为训练集 (13 446 张) 验证集 (1 680) 和测试集 (1 680)。

然后, 搭建并训练前述模型, 利用测试集验证模型的性能, 同时搭建 YOLOv3、YOLOv5、YOLOv7、YOLOv8 模型训练作为对比实验, 除此之外还设计了消融实验, 分别评估 LAF Module、NAE Block、MDF-Net 以及综合损失函数对模型性能的影响。

最后, 对上述模型进行剪枝和蒸馏, 对比分析 LMA-YOLO 模型、YOLOv8 模型、剪枝未蒸馏的 LMA-YOLO 模型以及剪枝蒸馏后的 LMA-YOLO。本实验在操作系统 Windows 10 下进行, GPU 为 Intel i9-10900K, 主机显存为

24 GB, 主机内存为 64 GB, 编程语言为 Python3.9, 使用 CUDN 11.7 和 CUDNN 8.9.3 对 GPU 进行加速, 基于深度学习框架 PyTorch 2.1.0 进行训练, 训练参数如表 1 所示。

表 1 训练参数设置

Table 1 Training parameter setting

参数	设置
optimize	Adam
image	640
epochs	200
Batch size	10
learning rate	0.008

2.2 评价指标

本实验采用全类别平均精度 (mAP)、参数量 (Params)、浮点运算 GFLOPS 和每秒帧数 (FPS) 来衡量算法的性能。mAP 计算公式如式 (15) 所示。

$$mAP = \frac{\sum_{i=0}^n AP_i}{N} \quad (15)$$

其中, AP 为平均精度, N 为总类别数, mAP@0.5 表示 IOU 阈值为 0.5 的 mAP 值, mAP@0.75 表示 IOU 阈值为 0.75 的 mAP 值, mAP@0.5:0.95 表示 IOU 阈值为 0.5~0.95 且步长为 0.05 的平均 mAP 值。GFLOPs 为每秒十亿次浮点运算, 用来衡量计算量的复杂程度, FPS 表示模型每秒处理图像帧数的速率, 作为检测速度的衡量。

2.3 对比实验

为验证 LMA-YOLO 模型在复杂战场军事目标检测中的性能优越性, 本文设计了对比实验, 部分可视化结果如图 8 和 9 所示, 前两组实验结果中基线模型存在明显的目标漏检现象, 其在复杂战场环境下对多目标、强干扰场景的适应性不足; 第 3 组实验结果中 LMA-YOLO 模型的检测框与目标真实边界贴合度更高, 且目标识别置信度更高, 验证了该模型在目标检测准确性和边界框回归精度上的性能优势, 其检测效果显著优于基线模型。

不同 YOLO 版本对比结果如表 2 所示, 与基线模型相比, 本文所提出的模型在 mAP@0.5 和 mAP@0.75 以及 mAP@0.5:0.95 中分别提升了 6.0%、7.2% 和 5.9%, LAF Module 的通道级噪声抑制、NAE Block 的全局特征建模及 MDF Block 的多尺度解耦融合, 有效弥补了传统 YOLO 模型特征提取不充分、抗干扰弱、多尺度适配差的缺陷, 与 YOLO 系列其他模型相比, LMA-YOLO 模型在战场目标检测任务中取得了最优性能, 且能够高效融合跨尺度特征信息, 弥补 YOLOv8 在目标遮挡与大尺度变化条件下的特征提取不足。

2.4 消融实验

为验证本文所提出的不同改进对模型检测效果的影响, 本文设计了对消融实验, 实验结果如表 3 所示, 引入 LAF

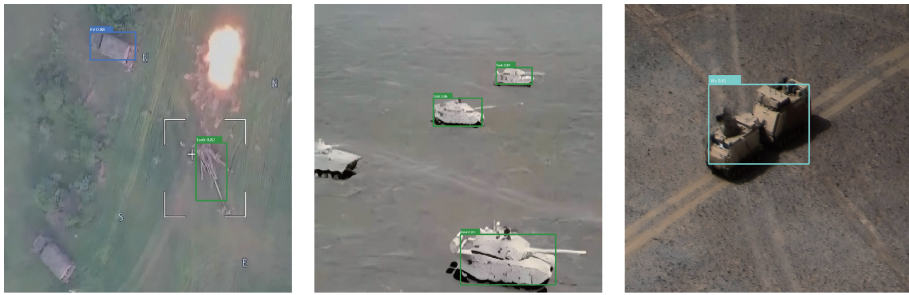


图 8 基于 YOLOv8 的检测结果

Fig. 8 Detection results based on YOLOv8

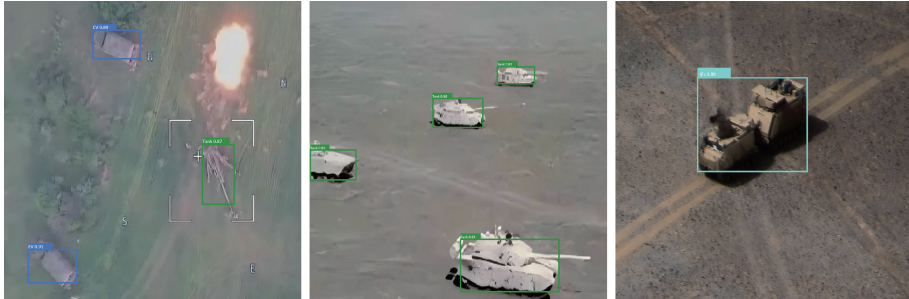


图 9 基于 LMA-YOLO 的检测结果

Fig. 9 Detection results based on LMA-YOLO

表 2 基于战场目标数据集的改进算法性能对比

Table 2 Performance comparison of improved algorithms on the battlefield typical target dataset

算法	Params/M	GFLOPs/G	mAP@0.5	mAP@0.75	mAP@0.5:0.95
YOLOv3-s	58.6	155.3	0.672	0.608	0.562
YOLOv3-SPP	59.6	156.1	0.703	0.641	0.584
YOLOv5-s	6.7	16.0	0.725	0.662	0.617
YOLOv5-l	44.0	108.3	0.817	0.751	0.675
YOLOv7	35.4	105.1	0.840	0.766	0.687
YOLOv9	29.8	118.2	0.867	0.781	0.705
YOLOv8-n	2.8	8.2	0.754	0.673	0.608
YOLOv8-s	10.6	28.7	0.809	0.745	0.648
YOLOv8-m(baseline)	24.6	79.1	0.858	0.782	0.686
LMA-YOLO	33.6	98.5	0.918	0.854	0.745

表 3 不同模块的影响

Table 3 The impact of different modules

YOLOv8	LAF Module	NAE Block	MDF Net	Loss	Params/M	GFLOPs/G	mAP@0.5	mAP@0.5:0.95
√	×	×	×	×	24.6	79.1	0.858	0.686
√	√	×	×	×	28.7	88.4	0.876	0.697
√	√	√	×	×	31.4	92.03	0.885	0.709
√	×	√	×	×	28.1	87.9	0.865	0.693
√	√	√	√	×	33.6	98.5	0.904	0.725
√	√	√	×	√	31.4	92.03	0.896	0.722
√	√	√	√	√	33.6	98.5	0.918	0.745

Module 与 NAE Block 后, $mAP@0.5$ 分别提升了 1.8% 和 0.9%, 证明了通道注意力与非局部特征建模的增强效果, 单独引入 NAE Block 时, 模型精度 $mAP@0.5$ 提升了 0.7%, 验证了本文关于 LAF Module 和 NAE Block 的设计, 证明了当 LAF Module 充分提取输入特征图的关键信息时, 模型能够更好的挖掘输入图片的内在特征, 从而提升模型的检测精度。引入 MDF-Net 与综合损失函数, $mAP@0.5$ 分别提升 1.9% 和 1.1%, 表

明所提 Neck 模块与损失函数可有效增强特征表达能力与回归性能。

为更直观的对比模块修改前后模型的性能, 本文给出部分三通道 RGB 热力图, 如图 10 所示, 其中第 1 行为 YOLOv8 模型的热力图, 第 2 行为 LMA-YOLO 模型的热力图, 改进后的模型能够有效的将相关目标的特征进行聚合, 并且热力中心点也基本对准目标中心, 证明了本文改进模块提高了网络的空间信息提取能力。

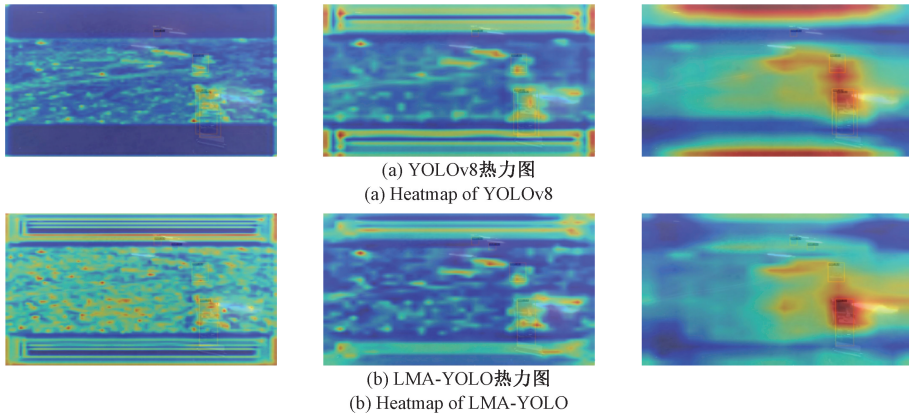


图 10 热力图可视化对比

Fig. 10 Heatmap visualization comparison

为深入探究各改进模块在面对不同战场干扰时的具体贡献, 本文进一步分析了各改进模块在特定干扰下的 $mAP@0.5$ 提升情况, 其结果如表 4 所示。

表 4 不同模块对特定干扰的影响

Table 4 Effects of different modules on specific interference

干扰	YOLOv8	LAF	NAE	MDF	Loss
烟雾	0.801	+0.041	+0.025	+0.032	+0.019
遮挡	0.812	+0.028	+0.055	+0.036	+0.014
小目标	0.739	+0.022	+0.018	+0.068	+0.037

LAF 模块凭借其基于能量最小化的 SimAM 注意力机制, 有效抑制了烟雾、闪光等背景噪声, 显著增强了目标关键特征的响应, 在烟雾遮挡场景下提升最为显著 (+0.041)。NAE 模块则通过建立像素间的长距离依赖, 整合全局上下文信息, 能够关联因伪装或遮蔽而断裂的目

标局部特征, 从而构建出完整结构, 在应对目标伪装时表现最为突出 (+0.055)。而 MDF-Net 通过并行多分支卷积与自适应空间注意力, 协同捕获小目标所需的细节特征与上下文信息, 极大地增强了小尺度目标的表征能力, 使其核心优势明确体现在小目标检测上 (+0.068)。此外, 综合损失函数通过 Focal Loss 调整难易样本权重以聚焦小目标, 并结合 DIoU Loss 优化边界框中心点回归, 二者协同, 对小目标检测带来了显著增益 (+0.037)。

2.5 轻量化分析

由于无人机前端平台算力、存储十分有限, 虽然改进的 LMA-YOLO 在检测精度上表现优异, 但其完整模型的参数量和计算复杂度仍难以直接满足边缘端的实时部署需求。因此, 本文对前述模型采用基于剪枝和蒸馏的协同轻量化策略, 实验结果如表 5 所示, 其中“-P”表示采用模型剪枝后的模型, “-PD”表示剪枝和蒸馏后的模型。

表 5 基于战场典型目标数据集的轻量化算法性能对比

Table 5 Performance comparison of lightweight algorithms on the battlefield target dataset

算法	Params/M	GFLOPs/G	$mAP@0.5$	FPS(GPU)	FPS(CPU)
YOLOv8-n	2.8	8.2	0.754	366	42
YOLOv8-s	10.6	28.7	0.809	267	24
YOLOv8-m(baseline)	24.6	79.1	0.858	222	18
LMA-YOLO	33.6	98.5	0.918	213	11
LMA-YOLO-P	5.76	11.83	0.877	354	36
LMA-YOLO-PD	5.76	11.83	0.894	354	36

与其他 YOLOv8 版本相比,本文设计的模型在有效减少了参数量和模型复杂度同时保障了模型的精度。与 LMA-YOLO 模型相比,剪枝后的模型大幅缩减了模型的参数量和复杂度,有效的提高了网络运行帧率,特别是在 CPU 上的实时帧率能达到 36 fps,但是由于模型结构大幅度削减,导致网络的 mAP@0.5 精度降低了 4.1%。因此本文进一步采用了蒸馏模型对网络学习特征进行优化,蒸馏后模型的检测精度得以回升,mAP@0.5 提高了 1.7%。尽管无法使得模型精度与初始状态相同,但与其他模型相比,在 FPS 和 mAP@0.5 精度上取得了出色的平衡。

3 边缘智能部署与实践

3.1 嵌入式硬件平台部署

对算法模型进行轻量化处理后,必须将其部署于边缘端嵌入式设备,以满足无人机作战实时性的需求。为此,本文设计了无人机战场目标辨识与定位系统,具体方案如图 11 所示。

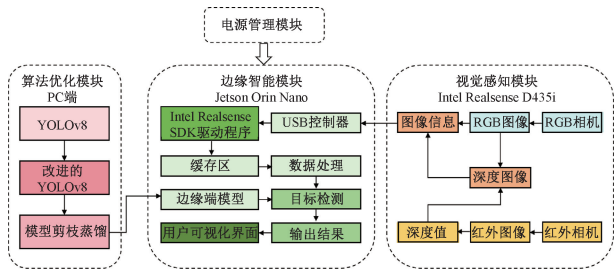


图 11 硬件平台具体方案

Fig. 11 Specific hardware platform solution

从功能上将硬件平台分为 4 个模块:电源管理模块、算法优化模块、边缘智能模块和视觉感知模块。其中算法优化模块前文中已经完成,视觉感知模块用于实现对战场环境的实时感知并采集图像,从深度图像中获取目标的深度值,进而获取目标相对于系统的三维坐标。边缘智能模块作为边缘智能计算终端,可以部署改进后的深度学习模型,代替庞大的工作机服务器,提高系统的便捷性。视觉感知模块和边缘智能模块分别选用 Intel RealSense D435i 深度相机和 Jetson Orin Nano 开发板和实现。

为获得目标三维位置信息,视觉模块硬件选用的是 Intel RealSense D435i 双目深度相机,相机如图 12 所示。



图 12 Intel RealSense D435i 深度相机

Fig. 12 Actual image of the Intel RealSense D435i depth camera

本文基于 D435i 官方的 SDK,结合 OpenCV 与 RealSense API,完成图像采集、像素映射与空间坐标计算

等流程,实现对目标的实时三维定位功能,该方式在无需引入额外空间定位硬件的前提下,为系统提供了可靠的三维空间信息支持,增强了系统在复杂环境下的场景感知能力和定位能力。

为满足实际战场环境所需要的便捷性、实用性,本文使用 TensorRT 在边缘智能终端来进行网络模型推理加速,然后将算法部署在 Jetson Orin Nano 开发板上,如图 13 所示。



图 13 Jetson Orin Nano 4 GB 开发板

Fig. 13 Jetson Orin Nano 4 GB development board

3.2 实践效果分析

为验证系统在实战环境中的目标检测与三维定位能力,本文基于实际搭建的硬件平台开展了实时检测实验,检测对象选用战场目标的高精度仿真模型,以 Jetson Orin Nano 为计算核心,通过 USB 连接 D435i 相机接收图像信息,远程连接 PC 电脑端的交互界面,实现交互控制操作,在右侧信息面板中同步显示检测模式、模型类型、推理条件阈值设置、目标类别、置信度、坐标数据等状态信息。

单一目标检测结果如图 14 所示,系统能够准确检测出目标类别、置信度及三维坐标信息,实时检测延迟为 29.3 ms,帧率约为 34 fps,实时检测性能良好。



图 14 单一目标检测结果

Fig. 14 Single-target detection result

非单一目标检测与定位结果如图 15 所示,从图中可以看出系统依然保持稳定的识别性能,并分别输出各目标的三维坐标,实时检测延迟为 32.4 ms,帧率约为 31 fps,实时检测性能良好。

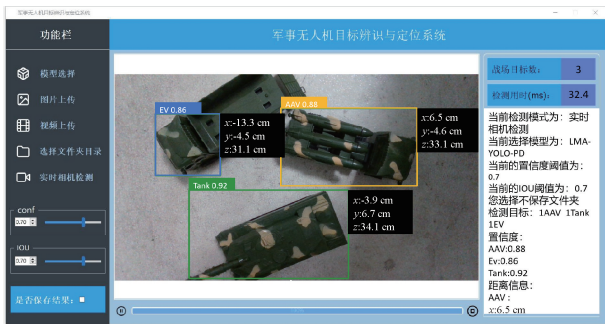


图15 非单一目标检测结果

Fig. 15 Non-single-target detection result

4 结 论

本文围绕复杂战场环境下智能化感知需求,设计实现针对无人机的边缘智能战场目标辨识与目标辨识与定位系统,研究涵盖目标检测算法改进、模型轻量化优化与嵌入式平台部署这3个环节。首先,针对实战中小目标漏检、边缘模糊等问题,提出改进YOLOv8模型,通过引入LAF Module、NAE Block和MDF Net,使mAP等指标显著提升,改善目标识别定位性能。其次,为解决嵌入式平台资源问题,提出依赖图结构化剪枝结合通道智能蒸馏的轻量化策略,大幅减少参数量与浮点数,精度损失控制4.1%以内。最后,基于Jetson Orin Nano与Intel RealSense D435i搭建硬件平台,结合深度信息实现目标三维定位,为后续研究与工程应用奠定基础。

然而,本文所提模型在极端遮挡与强干扰下的鲁棒性仍有提升空间,且目前仅依赖视觉模态,未融合红外、雷达等多源信息。未来工作将聚焦于多模态信息融合、动态场景在线自适应学习,并进一步探索轻量化与精度的平衡,最终通过实战平台集成验证,提升系统在全天候、复杂动态战场中的实际应用能力。

参考文献

- [1] 袁玲玲,陈春梅,朱天鑫,等.基于航拍图像的自适应感知目标检测网络[J].电子测量技术,2025,48(2):57-65.
YUAN L L, CHEN CH M, ZHU T X, et al. Adaptive perception object detection network based on aerial images[J]. Electronic Measurement Technology, 2025, 48(2): 57-65.
- [2] 苏雨蕾,黄丹丹,刘智,等.无人机对地小目标检测方法研究[J].电子测量与仪器学报,2024,38(9):144-154.
SU Y L, HUANG D D, LIU ZH, et al. Research on UAV-Based small target detection method[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(9): 144-154.
- [3] 孙晓永,孙备,郭润泽,等.面向多无人机协同的多模态

目标检测方法[J].仪器仪表学报,2025,46(2):209-220.

SUN X Y, SUN B, GUO R Z, et al. Multi-modal target detection method for multi-UAV cooperation[J]. Chinese Journal of Scientific Instrument, 2025, 46(2): 209-220.

- [4] 涂静正,温晓婧,陈彩莲,等.基于边缘计算的工业视频网络智能感知:挑战与进展[J].自动化学报,2025,51(8):1715-1738.

TU J ZH, WEN X J, CHEN C L, et al. Intelligent perception of industrial video networks based on edge computing: challenges and progress[J]. IEEE/CAA Journal of Automatica Sinica, 2025, 51(8): 1715-1738.

- [5] 江波,屈若崑,李彦冬,等.基于深度学习的无人机航拍目标检测研究综述[J].航空学报,2021,42(4):137-151.

JIANG B, QU R K, LI Y D, et al. Review of UAV aerial target detection based on deep learning[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(4): 137-151.

- [6] 杨婷.基于无人机航拍图像的目标检测算法研究[D].西安:西安工业大学,2023.

YANG T. Research on target detection algorithm based on UAV aerial image [D]. Xi'an: Xi'an University of Technology, 2023.

- [7] 王宗阳,黄莉,江都.基于APW-YOLOv8的无人机高空图像小目标检测[J].计算机系统应用,2025,34(10):195-205.

WANG Z Y, HUANG L, JIANG D. Small target detection of UAV high altitude image based on APW-YOLOv8[J]. Computer System Application, 2025, 34(10): 195-205.

- [8] 孙传猛,魏宇,李欣宇,等.复杂场景下无水尺水位的影像水位反演智能检测方法[J].测绘学报,2024,53(3):558-568.

SUN CH M, WEI Y, LI X Y, et al. Intelligent detection method for image water level inversion without gauge in complex scenes [J]. Journal of Geodesy and Geoinformation Science, 2024, 53(3): 558-568.

- [9] CHEN S Y, TANG X Y, WANG J, et al. An ultra-efficient streaming-based FPGA accelerator for infrared target detection[J]. Journal of Infrared and Millimeter Waves, 2022, 41(5): 914-922.

- [10] XU X H, WANG J, WANG F, et al. Infrared remote sensing image super-resolution network by integration of dense connection and multi-attention mechanism [J]. Journal of Infrared and Millimeter Waves, 2025,

- 44(2): 265-276.
- [11] 陈嘉欣, 孙传猛, 葛耀栋, 等. 基于 CycleGAN 和 Pix2Pix 的轨面缺陷图像智能生成技术[J]. 铁道学报, 2024, 46(2): 122-130.
CHEN J X, SUN CH M, GE Y D, et al. Intelligent generation technology of rail surface defect images based on CycleGAN and Pix2Pix[J]. Journal of the China Railway Society, 2024, 46(2): 122-130.
- [12] 李欣宇, 孙传猛, 魏宇, 等. 融合 Transformer 与残差通道注意力的恶劣场景水位智能检测方法[J]. 电子测量与仪器学报, 2023, 37(1): 59-69.
LI X Y, SUN CH M, WEI Y, et al. Intelligent detection method for water level in severe scenes based on Transformer and residual channel attention [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(1): 59-69.
- [13] ZENG D, WEI J M, ZHANG J J, et al. Progressive spatio-temporal feature fusion network for infrared small-dim target detection[J]. Journal of Infrared and Millimeter Waves, 2024, 43(6): 858-870.
- [14] 熊恩杰, 张荣芬, 刘宇红, 等. 面向交通标志的 Ghost-YOLOv8 检测算法 [J]. 计算机工程与应用, 2023, 59(20): 200-207.
XIONG EN J, ZHANG R F, LIU Y H, et al. Ghost-YOLOv8 detection algorithm for traffic signs [J]. Computer Engineering and Applications, 2023, 59(20): 200-207.
- [15] 马月红, 孔梦瑶. 基于改进快速区域卷积网络的目标检测轻量化算法 [J]. 兵工学报, 2021, 42(12): 2664-2674.
MA Y H, KONG M Y. Lightweight object detection algorithm based on improved fast region convolutional network [J]. Acta Armamentarii, 2021, 42(12): 2664-2674.
- [16] 吕昌, 尹和, 邵叶秦. 基于结构重参数化的目标检测模型[J]. 电子测量技术, 2023, 46(18): 114-121.
LYU CH, YIN H, SHAO Y Q. Object detection model based on structural re-parameterization [J]. Electronic Measurement Technology, 2023, 46(18): 114-121.
- [17] 刘毅, 于畅洋, 李国燕, 等. UAST-RCNN: 遮挡行人的目标检测算法 [J]. 电子测量与仪器学报, 2022, 36(12): 168-175.
LIU Y, YU CH Y, LI G Y, et al. UAST-RCNN: Occluded pedestrian detection algorithm [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(12): 168-175.
- [18] 胡昭华, 王长富. 改进 Faster R-CNN 的遥感图像小目标检测算法 [J]. 计算机工程与科学, 2024, 46(6): 1063-1071.
HU ZH H, WANG CH F. Improved Faster R-CNN algorithm for small target detection in remote sensing images [J]. Computer Engineering & Science, 2024, 46(6): 1063-1071.

作者简介

付羿焱, 硕士, 主要研究方向为深度学习、计算机视觉与图像识别。

E-mail: 18534600404@163.com

孙传猛(通信作者), 博士, 博士生导师, 教授, 主要研究方向为检测技术与自动化装置、模式识别与智能系统、深度学习与机器视觉以及强环境高干扰下的动态测试技术等。

E-mail: suncm@nuc.edu.cn