

中图法分类号: P237 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-13

论文引用格式: Yu Ying, Bai Dong, Zhang Ke, Li Kai, Li Lei, Li Li, Feng Xinyue. Fog density-guided remote sensing image detection of military vehicles in real-world scenarios[J/OI]. Journal of Image and Graphics, XXXX: 1-13. DOI: 10.11834/jig.260239. (于英, 柏栋, 张珂, 李凯, 李磊, 李力, 冯新月. 雾密度引导的真实场景军用车辆遥感图像检测[J/OI]. 中国图象图形学报, XXXX: 1-13. DOI: 10.11834/jig.260239.) [DOI: 10.11834/jig.260239]

雾密度引导的真实场景军用车辆遥感图像检测

于英¹, 柏栋^{1,2}, 张珂¹, 李凯³, 李磊¹, 李力¹, 冯新月¹

1. 信息工程大学 地理空间信息学院, 郑州 450001; 2. 解放军96143部队, 怀化 418000; 3. 军事科学院系统工程研究院, 北京 100171

摘要: 目的 真实场景下, 由于目标自身特性、环境交互及成像条件共同作用, 导致军用车辆目标在可见光遥感成像条件下的可分辨性、可分离性显著降低。尽管基于深度学习的遥感图像目标检测技术近年来取得了显著进展, 但现有研究对遥感图像特殊性的充分适应、效率与精度的平衡、以及真实场景下的泛化能力等方面, 仍有待进一步突破。针对真实场景目标检测的难点, 基于YOLOv11(you only look once)算法提出一种雾密度引导的真实场景军用车辆遥感图像检测算法(YOLO for military vehicle detection under foggy conditions, MVF-YOLO)。方法 在网络输入端增加一个基于雾密度权重的边缘特征恢复模块(fog density-driven edge restoration, FDER), 自适应恢复被雾衰减的边缘信息; 在网络颈部设计了空间注意力引导的频域信息选择模块(spatial attention-based frequency selection, SFS), 增强高频判别特征并抑制低频噪声, 并且利用遥感车辆刚性目标特性和目标长宽比先验知识, 使用基于目标长宽比先验的分类损失函数(aspect ratio binary cross-entropy Loss, ar_BCELoss)优化分类精度。结果 在MVRSD-V2.0数据集上, 所提方法在IoU阈值为0.5时的平均准确率达到87.6%, 较基线模型提升4.4%。结论 试验结果表明, 该方法在真实场景下具有强鲁棒性和军用车辆检测能力。

关键词: 遥感图像; 军用车辆; 目标检测; 真实场景; 雾密度

Fog density-guided remote sensing image detection of military vehicles in real-world scenarios

Yu Ying¹, Bai Dong^{1,2}, Zhang Ke¹, Li Kai³, Li Lei¹, Li Li¹, Feng Xinyue¹

1. Information Engineering University, Institute of Geospatial Information, Zhengzhou 450001, China; 2. Unit 96143 of the Chinese People's Liberation Army, Huaihua 418000, China; 3. Institute of Systems Engineering, Academy of Military Sciences, Beijing 100171, China

Abstract: Military vehicle detection in high-resolution optical remote sensing imagery is a critical link in battlefield situational awareness and intelligence acquisition, and it imposes stringent requirements on both timeliness and accuracy. In remote sensing images, military vehicle targets generally appear at small scale, in dense distributions and with arbitrary orientations, while realistic reconnaissance scenarios are frequently accompanied by complex background clutter and occlusion, artificial camouflage interference, and adverse weather such as haze and fog. These factors are strongly coupled: the physical characteristics of the target itself, its interaction with the surrounding environment, and the imaging conditions act jointly to degrade the observed imagery, so that the distinguishability of a target from its background and the separability

收稿日期: 2026-04-27; 修回日期: 2026-08-28

* 通信作者: 柏栋 baidonglsq001@163.com

基金项目: 国家自然科学基金(42071340); 嵩山实验室项目(纳入河南省重大科技专项管理体系)(221100211000-01)

Supported by: National Natural Science Foundation of China (42071340); Program of Song Shan Laboratory (Included in the

©中国图象图形学报版权所有

among neighboring targets are both markedly reduced. Under foggy conditions in particular, atmospheric scattering attenuates the radiance reflected from the scene and superimposes an additive veil of atmospheric light, and this attenuation grows rapidly with scene depth. As a result, image contrast and edge energy decay severely, directly eroding the contour cues on which the detection of small vehicles most heavily relies. Optical remote sensing detection of military vehicles therefore remains an extremely challenging task. Existing studies have addressed this problem mainly along two lines. For generic detection under complex backgrounds, attention mechanisms and context modeling have been widely adopted to suppress background interference and to compensate for the feature loss caused by occlusion, while saliency modeling and generative adversarial network based approaches have shown potential for enhancing camouflaged and low-resolution targets. For detection under foggy weather, cascaded schemes that first dehaze and then detect can improve accuracy, but they suffer from high computational complexity and error accumulation, because the dehazing stage is optimized for perceptual quality rather than for downstream detection. Parallel dual-branch frameworks that jointly optimize restoration and detection achieve end-to-end learning, yet still encounter bottlenecks such as strong dependence on paired hazy-clear training data, complicated network architectures, and the difficulty of balancing the loss weights between the dehazing and detection sub-tasks. Meanwhile, generic remote sensing detectors are largely built upon natural-image baselines, and their adaptation to the distinctive properties of remote sensing imagery remains limited; the attention modules they employ often incur considerable computational overhead that is hard to reconcile with the real-time demands of military reconnaissance. More fundamentally, the intrinsic prior knowledge carried by vehicle targets in remote sensing images—in particular their rigid geometry and stable aspect-ratio distribution—has not yet been effectively exploited. Overall, current methods still exhibit evident deficiencies in adapting to remote sensing characteristics, in balancing efficiency against accuracy, and in generalizing to real-world scenarios. Motivated by these observations, and taking YOLOv11n (you only look once), which offers a favorable trade-off between detection accuracy and inference speed, as the baseline, this paper proposes a fog-density-guided detection algorithm for military vehicles in real-scenario remote sensing imagery, termed MVF-YOLO (YOLO for military vehicle detection under foggy conditions). The method improves the baseline at the input, the neck and the detection head, and is supported by a purpose-built dataset. At the network input, a fog-density-driven edge restoration module (FDER) is introduced. It fuses the dark channel prior with multi-scale gradient features to estimate a pixel-wise fog density map that characterizes the spatially non-uniform degradation of the scene, and then uses this map to adaptively re-weight and restore the edge responses attenuated by scattering, so that heavily fogged regions receive stronger edge compensation while clear regions remain essentially unchanged and are not over-sharpened. Because FDER carries out edge recovery at the feature level rather than performing image-level dehazing, it requires neither paired hazy-clear supervision nor an additional restoration branch, and it improves feature quality under degraded imaging conditions at very low cost. In the neck, a spatial attention-based frequency selection module (SFS) is designed. It first employs spatial attention to guide the network toward the key regions likely to contain targets, and then performs selective processing in the frequency domain, suppressing low-frequency components that mainly carry smooth background, haze veil and residual noise, and enhancing high-frequency components that encode contours, textures and structural discontinuities—the cues that are most discriminative for rigid vehicle targets. This joint spatial-frequency selection strengthens the recognition of occluded and camouflaged targets whose responses would otherwise be submerged in background energy. In the detection head, a classification loss based on the aspect-ratio prior of the target, denoted $ar_BCELoss$, is proposed. Exploiting the rigid geometric nature of military vehicles, whose aspect ratios are concentrated within a stable interval, the loss modulates the standard binary cross-entropy according to the consistency between the geometry of a candidate and the prior distribution, thereby emphasizing samples that conform to the vehicle shape prior and down-weighting background responses with implausible geometry; the classification boundary is refined and classification accuracy improved with almost no additional computational overhead. Finally, MVRSD, the only publicly available remote sensing dataset for military vehicle detection in China, is improved and extended: style transfer techniques are used to synthesize foggy samples with realistic and diverse fog densities, yielding the MVRSD-V2.0 dataset, which better represents real-world reconnaissance conditions. Experiments on MVRSD-V2.0 show that the proposed method attains an average precision of 87.6% at an IoU threshold of 0.5, an improvement of 4.4 percentage points over the baseline. Ablation studies confirm that each of the three components contrib-

utes positively and that their combination performs best: FDER mainly recovers the accuracy lost on heavily degraded foggy samples, SFS mainly benefits occluded and camouflaged targets in cluttered backgrounds, and ar_BCELoss mainly reduces false alarms caused by background structures with vehicle-like appearance. Since the added modules introduce only marginal growth in parameters and computation, these gains are obtained while the model retains the lightweight nature and inference speed of the YOLOv11n baseline. Comparisons with mainstream detectors and with cascaded dehazing-then-detection pipelines further indicate that MVF-YOLO achieves a better accuracy-efficiency trade-off without relying on an explicit dehazing stage or on paired hazy-clear training data. The experimental results demonstrate that, by jointly performing fog-density-driven edge restoration, spatially guided frequency selection and shape-prior-constrained classification, the proposed method possesses strong robustness and reliable military vehicle detection capability under the coupled degradations of fog, occlusion and camouflage encountered in real scenarios, providing a practical technical reference for remote sensing reconnaissance applications with real-time requirements.

Key words: remote sensing imagery; military vehicles; object detection; real-world scenarios; fog density

论文引用格式: 于英, 柏栋, 张珂, 李凯, 李磊, 李力, 冯新月. 2025. 雾密度引导的真实场景军用车辆遥感图像检测. 中国图

象图形学报, [DOI:10. 11834/jig. 260239]

0 引言

高分辨率光学遥感图像中的军用车辆检测是战场态势感知与情报获取的关键环节, 对时效性和准确性有着极高要求。然而, 军用车辆目标在遥感影像中普遍呈现尺度小、分布密集、方向任意的特点(Liu 等, 2024), 且真实侦察场景常伴生复杂背景遮挡、人工伪装干扰以及雾霾等恶劣天气影响, 这些因素相互耦合, 导致目标的可分辨性与可分离性显著降低, 使光学遥感军用车辆检测成为极具挑战性的任务。

针对上述问题, 国内外学者已从复杂背景下的通用目标检测和恶劣天气下的专门检测两个维度开展了大量研究。在复杂背景检测方面, 注意力机制与上下文建模被广泛用于抑制背景干扰和补偿遮挡造成的特征缺失(Liu 等, 2021; Li 等, 2024), 基于显著建模和生成对抗网络的方法在伪装目标与低分辨率目标增强上也展现出潜力(Gao 等, 2021)。在雾天检测方面, 级联式先去雾后检测的方案虽能提升精度, 但面临复杂度高、误差累积等问题; 并联双分支联合优化方法实现了端到端学习, 但仍存在对成对有雾-清晰数据依赖强、网络结构复杂以及去雾与检测子任务权重难平衡等瓶颈(Ma 等, 2021; Li 等, 2023; Liu 等, 2023; Song 等, 2023; Li 等, 2024)。与此同时, 通用遥感图像目标检测方法多以自然图

像基线模型为基础, 针对遥感特性的适应性依然有限, 且现有注意力机制常伴随较高计算开销, 难以满足军事侦察的实时性需求。更为关键的是, 遥感图像中车辆目标固有的形状先验等知识尚未被有效利用。总体来看, 现有方法在遥感特性适配、效率与精度平衡及真实场景泛化能力等方面仍存在明显不足(张雨晴 等, 2023)。

本文以检测精度与速度均衡的YOLOv11n为基线, 提出了一种雾密度引导的真实场景军用车辆遥感图像检测算法MVF-YOLO, 解决了真实场景下军用车辆目标检测精度低、鲁棒性不足的问题。所做工作如下: 1) 在网络输入端设计了一种雾密度驱动的边缘特征恢复模块FDER, 通过融合暗通道先验与多尺度梯度特征生成雾密度图, 自适应恢复被雾衰减的边缘信息, 提升图像退化条件下的特征质量; 2) 在颈部网络引入空间注意力引导的频域选择模块SFS, 引导网络聚焦目标关键区域, 并通过滤除低频成分、强化高频判别特征, 增强对遮挡与伪装目标的识别能力; 3) 在检测头中提出基于目标长宽比先验的分类损失函数ar_BCELoss, 利用军用车辆的刚性几何特性优化分类边界, 在几乎不增加计算开销的前提下提升分类精度; 4) 对国内唯一公开的军用车辆检测遥感图像数据集MVRSD进行改进与扩展, 使用风格迁移技术生成雾天数据样本, 构建面向真实场景的MVRSD-V2.0数据集, 实验结果表明, 本文方法在不依赖图像去雾的条件下, 有效提升了模型在真实复杂场景下的鲁棒性与泛化能力。

1 MVF-YOLO算法

本文在YOLOv11n算法网络基础上,通过在网络输入端增加基于雾密度权重的边缘特征恢复模块、在网络颈部设计了空间注意力引导的频域信息

选择模块,分别从图像边缘特征恢复与频域信息选择两个层面提升模型在真实场景下的检测精度,并且利用遥感车辆刚性目标特性和目标长宽比先验知识,使用新的分类损失函数优化分类结果,提出了用于真实场景军用车辆遥感图像检测的MVF-YOLO算法,算法网络结构如图1所示。

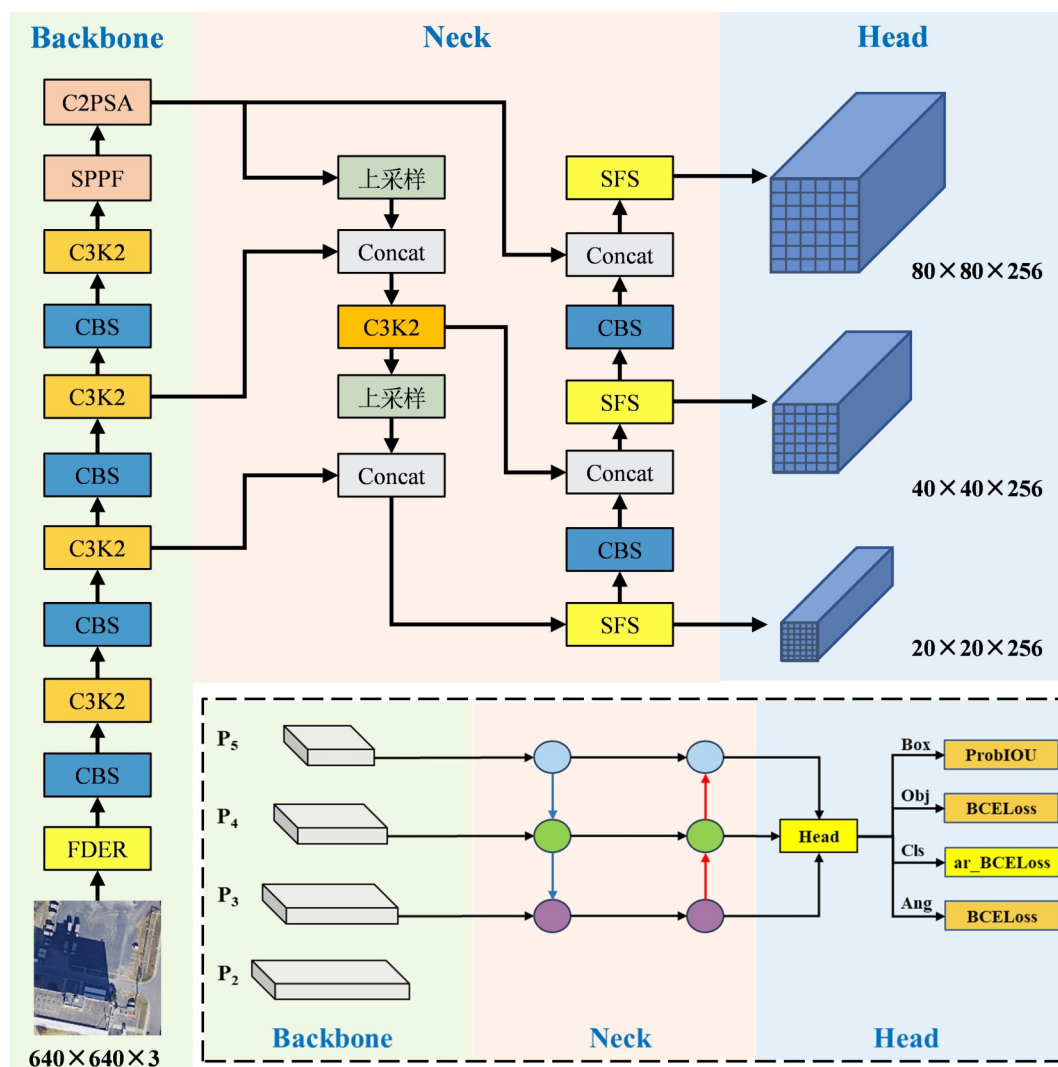


图1 MVF-YOLO网络结构图

Fig. 1 The architecture of the MVF-YOLO

1.2 基于雾密度权重的边缘特征恢复模块

1.2.1 基于雾密度感知的注意力权重生成

遥感图像中的雾分布与自然图像存在显著差异(见图2、表1),现有雾密度计算方法主要针对自然图像设计,难以直接适配(Miclea等,2021;Cao等,2023;Hwang等,2024)。

为此,本文提出一种面向遥感特性的雾密度感知注意力权重生成方法,利用图像统计特征建立雾

密度图,引导网络自适应地关注受雾影响区域。雾密度图生成流程如图3所示,主要包括:

(1)多尺度梯度特征提取。在对图像进行3、9、15、21倍下采样的基础上,分别计算各尺度下采样图的梯度特征 f_{Gi} ;

(2)梯度特征权重矩阵生成。将多尺度梯度特征上采样恢复至输入图像尺寸,融合生成权重矩阵;

(3)暗通道特征分布调整。暗通道特征 f_{dark} 减



(a)有雾遥感图像 (b)有雾自然图像

图2 雾天条件下的遥感图像和自然图像

Fig. 2 Remote sensing images and natural images in foggy weather((a) foggy remote sensing images;(b) foggy natural images)

去上采样后的梯度特征 f'_{G1} ,再与梯度权重矩阵相乘,得到分布调整后的暗通道 f'_{dark} ;

(4)雾密度图生成。分布调整后的暗通道 f'_{dark} 和图像的饱和度-亮度(Saturation and Value,SV)特征 f_{SV} 加权相加,生成雾密度图 f_H ;

整体流程可用下式表示:

$$f_H = \max((f_{dark} - k_1 f_{GI}), 0) \otimes (1 - \frac{1}{4} \sum_{i=1}^4 U(f_{Gi})) + k_2 f_{SV} \quad (1)$$

式中, k_1 和 k_2 是对应特征的权重,本文分别取0.5和0.3, $\otimes$ 代表矩阵相乘, $U(\cdot)$ 代表上采样。

表1 有雾遥感图像和自然图像的对比

Table 1 Comparison of Foggy Remote Sensing Images and Natural Images

类别	遥感图像	自然图像
背景特性	入射角变化范围小	成像角度任意,常包含天空区域
景深分布	成像距离远,景深变化小,局部区域内雾密度变化小	景深变化大,雾密度随景深增大
图像质量	受卫星姿态、传感器自身物理性质等因素影响可能出现模糊	自然图像质量受退化因素影响较小,通常认为是清晰的
雾的空间结构	图像尺寸大,雾多呈非均匀分布	图像尺寸适中,雾在整幅图像上均匀分布,雾浓度仅随景深变化

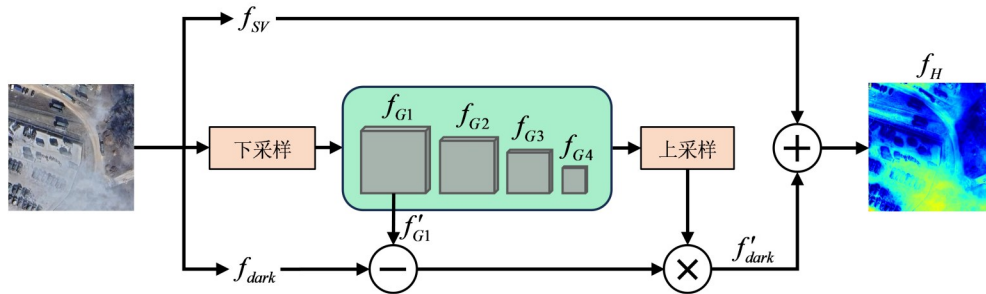


图3 雾密图生成流程

Fig. 3 Fog Density Map Generation Workflow

1.2.2 边缘特征提取

针对传统边缘检测算子在雾天条件下噪声鲁棒性弱、弱边缘响应不足的缺陷,本文设计边缘信息提取模块(Edge Information Extraction Module, EIEM),通过特征图自适应加权实现内容感知的边缘学习,其结构如图4所示。

F_a 是输入特征 F 通过平均池化层平滑后得到的低频特征信息,将输入特征与平滑特征相减得到高频信息 F_e , F_e 经卷积层和Sigmoid函数激活后与原始输入相加,得到网络需要的增强边缘特征,EIEM的等效数学表示为:

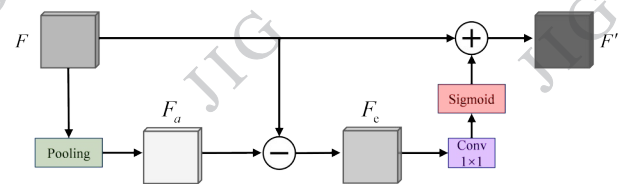


图4 EIEM结构图

Fig. 4 EIEM Structure Diagram

$$F' = F + \sigma(\text{Conv}_1(F_e)) \quad (2)$$

$$F_e = F - F_a \quad (3)$$

式中, Conv_1 代表 1×1 卷积核大小的卷积层, $\sigma(\cdot)$ 代表Sigmoid函数。与固定权重的传统边缘算子相比,

EIEM 能够空间自适应学习,根据图像内容动态调整。

1.2.3 边缘特征恢复

FDER 结合了多尺度边缘特征提取、雾密度注意力权重和卷积操作,如图5所示,其网络结构由卷积分支、边缘特征分支和雾密度注意力分支构成,特征恢复过程为:

(1)卷积分支。通过一个 3×3 的卷积层提取图像局部特征,以保留丰富的空间信息;

(2)边缘特征分支。使用自适应平均池化层 AAP(Adaptive AvgPooling),自动将输入图像转换为指定尺寸的输出特征图,通过提取不同尺度的边缘信息,捕捉图像的多层次特征。这样做主要是考虑

到直接在输入图像上提取边缘信息的局限性:一是原始图像中大量的背景信息会引入干扰噪声,增加网络学习难度;二是该方法会导致参数量过大。此外,多尺度的特征信息能够保持边缘特征和空间信息的一致性;

(3)雾密度注意力分支。通过生成雾密度权重,有选择性地增强高雾密度区域的特征,引导网络在抑制噪声干扰的同时,最大限度地保留检测可能性;

(4)边缘特征恢复与融合。将不同尺度下提取的边缘特征通过插值操作对齐到同一尺度,使用雾密度引导后与局部空间特征融合,输出的特征表示既包含了恢复后的边缘信息,又包含了丰富的空间特征,能够更加全面地表达遥感图像内容。

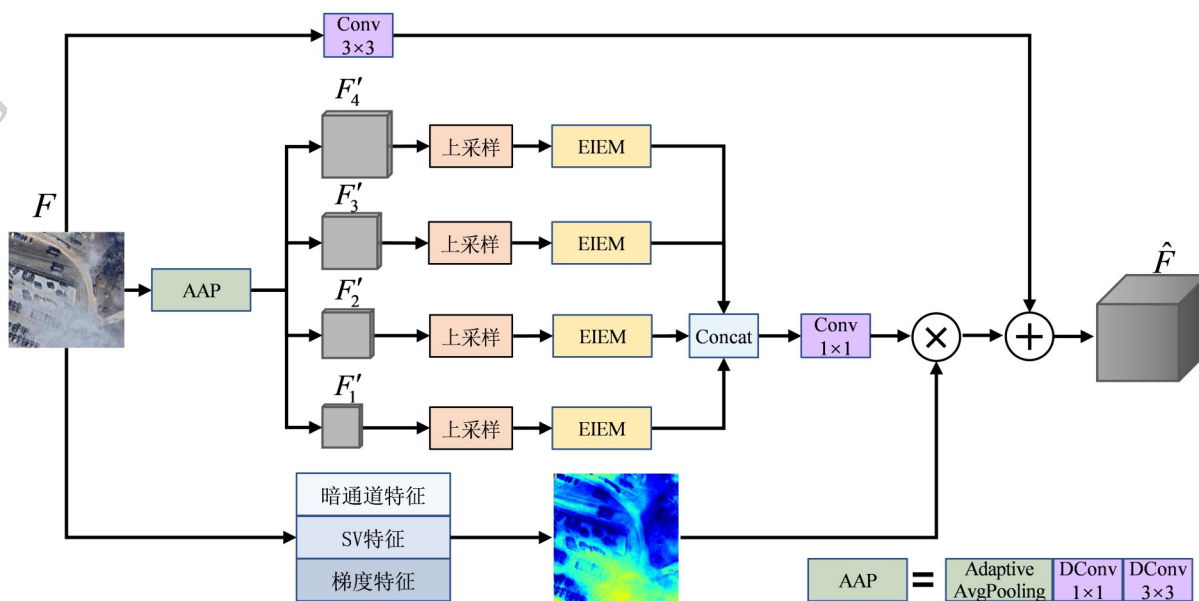


图5 FDER结构图

Fig. 5 FDER Structure Diagram

给定输入 F , FDER 的输出 $\hat{F}$ 可以表示为:

$$\hat{F} = \text{Contact} [\text{EIEM}(\text{U}(F'_i))] \otimes \text{GSD} + \text{Conv}_3(F) \quad (4)$$
 式中, EIEM 代表边缘特征提取, GSD 代表雾密度注意力权重, $\text{U}(\cdot)$ 表示上采样操作, $\otimes$ 表示矩阵相乘, 表示 3×3 卷积核大小的卷积层, F'_i 表示第 i 条分支的自适应池化操作:

$$F'_i = \text{DConv}_3(\text{DConv}_1(\text{AAvgPool}_i(F))) \quad (5)$$

式中, $i \in \{1, 2, 3, 4\}$; $\text{AAvgPool}_i(\cdot)$ 表示自适应平均池化层, 本文选取的平均池化后的特征图尺寸为 $9i \times 9i$, 即分别为 9×9 、 18×18 、 27×27 和 36×36 ; DConv_1 和 DConv_3 分别表示卷积核大小为 1×1 和

3×3 的深度可分离卷积。

1.3 空间注意力引导的频域信息选择模块

在遥感军用车目标检测任务中,背景遮挡和目标伪装目标特征弱化和丢失,通用算法难以准确定位和识别困难目标,影响网络的召回率和精度。为此,本文在网络颈部设计了一个空间注意力引导的频域选择模块 SFS。该模块借助空间注意力引导网络学习特征图上有效信息的空间分布,并通过滤除各通道中的最低频成分,进一步降低无关噪声的干扰,增强模型对目标细节的感知,解决因遮挡、伪装导致的特征缺失和定位偏差,提升复杂背景下的

检测精度。SFS模块可分为空间注意力引导和高频

信息选择两个部分,模块结构如图6所示。

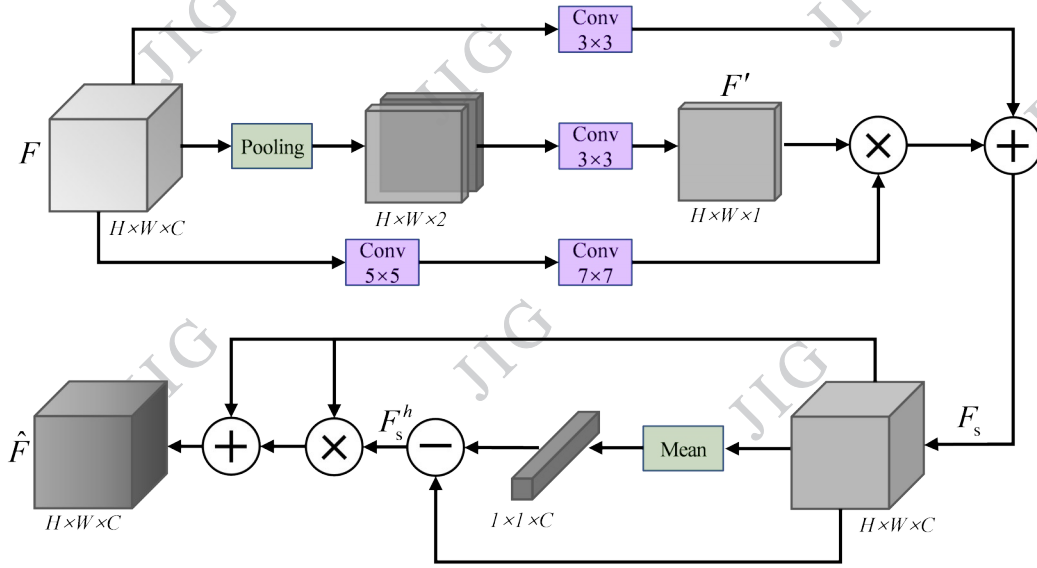


图6 SFS模块结构图

Fig. 6 SFS Structure Diagram

1.3.1 空间注意力引导

对于给定特征 F , 首先沿通道维度通过全局最大池化和全局平均池化进行压缩, 得到两个 $H \times W \times 1$ 的特征图; 对这两个特征图拼接后通过卷积层和 Sigmoid 函数激活, 生成空间注意力权重矩阵 F' , 数学表示如下式:

$$F' = \sigma(\text{Conv}_3(\text{Concat}[\text{AvgPool}(F), \text{MaxPool}(F)])) \quad (6)$$

式中, $\sigma(\cdot)$ 表示 Sigmoid 激活函数, AvgPool、MaxPool 分别表示平均池化和最大池化, Conv_3 是 3×3 卷积核大小的卷积层。

将与空间注意力权重矩阵 F' 相乘后的输入特征, 与通过卷积层生成的局部空间特征相加, 得到关注重点区域的特征 F_s :

$$F_s = \text{Convs}_{5,7}(F) \otimes F' + \text{Conv}_3(F) \quad (7)$$

式中, $\text{Convs}_{5,7}$ 表示 5×5 和 7×7 卷积核大小的级联卷积层, 用于捕获大范围上下文信息; Conv_3 表示 3×3 卷积核大小的卷积层, 用于提取局部细节; $\otimes$ 表示矩阵相乘。

1.3.2 高频信息选择

为引导网络在关注重点区域的同时有针对性增强特征, 通过移除低频信息来获取特征的高频分布, 在减少噪声影响的同时进一步引导网络对目标区高频信息的关注, 同时使用残差连接避免信息的丢失。

具体的, 通过计算并减去 F_s 在空间维度上的均值, 来获得图像的高频特征权重 F_s^h , 过程可描述为:

$$F_s^h = \sigma(F_s - \text{Mean}(F_s)) \quad (8)$$

式中 $\sigma(\cdot)$ 表示 Sigmoid 函数, $\text{Mean}(\cdot)$ 代表对每个通道进行平均池化, 生成 $1 \times 1 \times C$ 的均值低频空间, 并扩展到 $H \times W \times C$ 形状。

将高频特征权重 F_s^h 和原始特征 F_s 相乘, 并通过残差连接生成最终输出 $\hat{F}$:

$$\hat{F} = \alpha(F_s^h \otimes F_s) + \beta F_s \quad (9)$$

式中, α 和 β 是代表高频增强强度和残差保留比例的通道级可学习参数, $\otimes$ 表示矩阵相乘。

1.4 先验知识辅助的目标分类优化

遥感影像具有成像距离远、入射角变化范围小的特点, 车辆作为地面刚性目标, 几何形状受遥感图像拍摄角度和环境变化影响较小, 同类型车辆长宽比分布相对集中, 结合目标先验知识可以优化目标分类损失, 提高检测精度(肖超, 等, 2024)。

本文将目标的长宽比分布信息引入分类损失, 设计了新的分类损失函数 ar_BCELoss, 给模型增加了一个“几何形状”的判断维度。当模型预测的边界框长宽比与某个类别的先验长宽比分布高度匹配时, 增强模型对该类别的置信度(或减少损失); 反之, 当严重不匹配时, 则施加惩罚。

对于给定的目标样本 i , 目标预测框的长宽分别

是 w_i 和 h_i , 则长宽比 ar_i 为:

$$ar_i = \max(w_i, h_i) / \min(w_i, h_i) \quad (10)$$

可根据如下公式计算样本 i 在类别 c 的概率密度 φ_{ic} :

$$\varphi_{ic} = f(ar_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(ar_i - \mu)^2}{2\sigma^2}} \quad (11)$$

式中, μ 和 σ 是目标类别 c 的标注框长宽比均值和标准差。

样本 i 在所有类别的概率密度进行归一化后得到样本 i 在类别 c 的分类预测权重 w_{ic} :

$$w_{ic} = \frac{\varphi_{ic} - \varphi_{\min}}{\varphi_{\max} - \varphi_{\min}} \quad (12)$$

式中, $\varphi_{\max}$ 和 $\varphi_{\min}$ 是样本 i 在所有类别概率密度的最大值和最小值。

对多分类交叉熵损失函数优化后的分类损失 $ar_BCELoss$ 如下式所示:

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^K y_{ic} \log(S(w_{ic} \cdot p_{ic})) \quad (13)$$

式中, N 是总样本数, K 是类别数, p_{ic} 代表第 i 个样本属于类别 c 的概率; $\sum p_{ic} = 1$, $S(\cdot)$ 表示 softmax 函数; $y_{ic} \in \{0, 1\}$, 当第 i 个样本真值属于类别 c 时取 1, 否则取 0。

2 实验结果与分析

2.1 数据集及评价指标

为满足研究对数据质量与任务适配性的要求, 本文在 MVRSD 数据集(柏栋等, 2024)的基础上进行了改进和扩展, 并使用风格迁移技术生成雾天条件下的数据样本, 构建了面向真实场景检测任务的 MVRSD-V2.0 数据集。

MVRSD-V2.0 数据集由真实图像子集和对应的雾合成图像子集构成, 涵盖城市、平原、山地和沙漠等多种地理环境, 包含多类不同伪装程度的军用车辆。真实图像子集采集自亚洲、北美洲和欧洲范围内 80 余个军事场景下的 4500 幅遥感图像, 共有 165 851 个实例, 图像空间分辨率 0.1-0.5m, 尺寸大小为 640×640 像素。数据集将车辆划分为 5 类: 小型军用车辆(Small Military Vehicles, SMV)、大型军用车辆(Large Military Vehicles, LMV)、装甲战车(Armored Fighting Vehicles, AFV)、小型民用车辆(Small Civilian Vehicles, SCV)、大型民用车辆(Large Civilian

Vehicles, LCV)。为进一步拓展研究范围, 采用风格迁移技术构建了配套的雾合成图像子集, 旨在评估检测算法在雾天条件下的性能。数据集部分真实图像和对应雾合成图像见图 7。MVRSD-V2.0 数据集能够支持图像去雾、域自适应迁移及恶劣天气下的鲁棒性检测等前沿研究方向, 可通过 XXX 渠道获取。

本文使用 MVRSD-V2.0 的雾合成图像子集作为实验对象, 按照 7:2:1 的比例随机划分训练集、验证集和测试集; 使用 IoU 阈值为 0.5 时的平均精度(Average Precision, AP)、平均精度均值(mean Average Precision, mAP)来评价模型检测精度; 使用召回率 R 来评估模型的检测能力完备性; 使用模型参数量 Params、浮点运算次数 GFLOPs 和每秒检测图像帧数 FPS(Frames Per Second)评估模型效率。

2.2 实验环境

所有实验过程都是在相同的环境下进行, 网络输入图像尺寸为 640×640, Batch-size 为 32, epoch 为 250。使用 SGD(Stochastic Gradient Descent) 优化器, SGD 优化动量为 0.937, 权重衰减值为 5×10^{-4} 。为提高模型熟练速度和训练效率, 前三个 epoch 进行预热训练, 预热动量设置为 0.8。具体实验环境配置如表 2 所示。

2.3 对比实验

为验证改进算法的精度优势, 本文将所提方法与当前多种代表性的有雾图像目标检测算法进行对比, 包括 IA-YOLO(Liu 等, 2021)、YOLOv5s-Fog(Meng 等, 2023)、MS-DAYOLO(Hnewa 等, 2021)、TogetherNet(Wang 等, 2022), 以及由 AODNet(All-in-One Dehazing Network)与 YOLOv11 级联构成的检测模型。这些方法分别涵盖了图像自适应增强、注意力机制改进、域自适应、联合学习及级联去雾等不同技术路线, 为全面评估性能提供了多维基线。需要特别说明的是, 由于部分对比算法采用水平检测框输出形式, 为避免修改算法结构引起性能变化, 本文对这些算法使用的数据标注文件进行修改, 把所有定向标注框统一取其最小水平外接矩形作为检测标注框, 以避免因检测方式和标注形式不统一而带来的评估偏差。对比结果如表 3 所示。从实验结果可以看出, 在不同类型的有雾图像目标检测算法中, 本文所提出的模型在检测精度上达到了最优性能, 为 87.6%, 同时在参数量、计算复杂度和推理速度之间

取得了良好平衡。



图7 部分真实图像和对应雾合成图像
Fig. 7 Partial Real Images and Corresponding Fog-Synthesized Images

表2 实验硬件环境和实验参数设置

Table 2 Experimental hardware environment and experimental parameter settings		
环境	项目	内容
硬件环境	CPU	Intel(R) Xeon(R) Platinum 8352V
	GPU	NVIDIA GeoForce RTX 4090(24GB)
软件环境	CUDA	Cuda 12.1.1
	Deep Learning Framework	Pytorch 2.4.0
	Operating System	Ubuntu 5.15.0-136-generic
	Programming Languages	Python 3.11

具体而言,IA-YOLO 与 YOLOv5s-Fog 检测精度相对较低,分别为 80.3%和 83.4%。这两种方法主要通过图像自适应增强或增强网络的特征提取与识别能力实现性能提升,虽不依赖清晰图像监督,但对高雾浓度图像的特征提取能力有限,因此在极端天气条件下的泛化能力不及显式图像去雾方法。MS-

DAYOLO 借助无监督域自适应方法避免了有雾图像标注的难题,取得了 83.9%的检测精度。但其模型参数量高达 247.6 M,推理速度仅 5.6 FPS,难以满足实际应用中实时性的要求。此外,该类方法需同时处理源域与目标域数据,训练过程复杂且稳定性较差,最终性能高度依赖于源域数据集规模与质

表 3 在测试集雾合成图像上的实验结果

Table 3 Experimental Results on Fog-Synthesized Test Set Images

网络模型	mAP@0.5(%)	Params(M)	FLOPs(G)	FPS	边界框
IA-YOLO	80.3	62.7	193.89	45.6	水平边界框
YOLOv5s-Fog	83.4	7.2	25.8	53.5	水平边界框
MS-DAYOLO	83.9	247.6	172.4	5.6	水平边界框
TogetherNet	87.4	19.2	79.6	21.8	水平边界框
YOLOv11n	83.2	2.6	6.5	199.5	定向边界框
AODNet+YOLOv11n	84.5	3.5	10.5	42.6	定向边界框
MVF-YOLO(本文模型)	87.6	3.5	17.6	96.3	定向边界框

注:加粗字体为精度最优结果。

量。TogetherNet通过联合学习策略实现图像恢复与目标检测的协同优化,获得了87.4%的高检测精度,但其并联式网络结构引入了较多额外参数与计算量,尽管在推理阶段关闭恢复模块,检测速度仍仅为21.8 FPS,难以在高实时性场景中部署。AODNet与YOLOv11n的级联模型相比YOLOv11n在精度上提升1.3%,说明去雾预处理对检测性能具有积极作用,但由于去雾任务与检测任务之间存在优化目标差异,简单的级联结构难以充分实现两任务间的协同增效。本文提出的检测算法达到87.6%的最优

检测精度,同时保持了较高的推理速度,尽管在输入端引入了图像恢复模块导致FLOPs略有增加,但由于雾密度权重由图像特征直接计算,无需额外的可学习参数,因此模型仍具有较高的计算效率与实用性。

2.4 消融实验

为了系统验证所提方法的有效性和算法中各个组成模块的贡献,采用消融实验,在相同的实验条件下评估各模块对算法的贡献。

表 4 消融实验结果

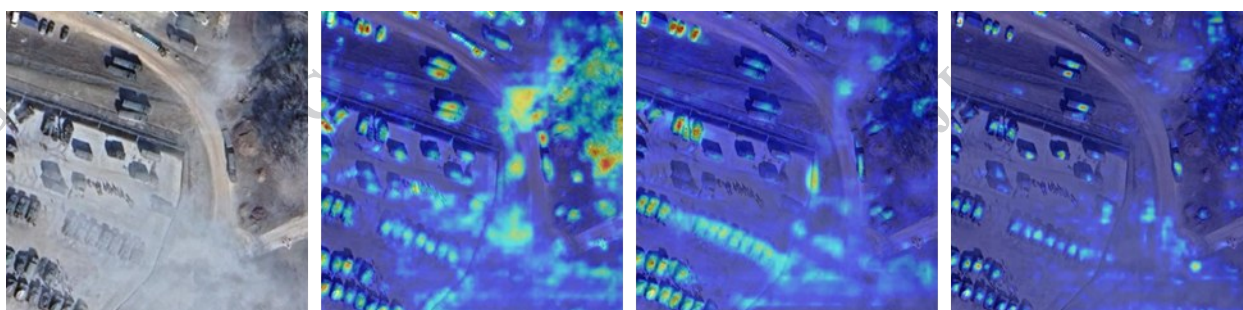
Table 4 Ablation Study Results

网络模型	AP@0.5(%)					mAP@0.5 (%)	Params (M)	FLOPs (G)
	SMV	LMV	AFV	SCV	LCV			
基线模型	81.6	82.4	76.1	92.8	83.1	83.2	2.6	6.5
+FDER	85.6	87.0	80.5	96.0	86.7	87.2	3.6	17.5
+SFS	84.5	85.2	79.9	95.9	86.8	86.5	3.3	7.9
+ar_BCELoss	81.6	83.0	77.1	92.6	83.8	83.6	2.6	6.5
+FDER+SFS	86.0	87.1	80.6	96.4	87.0	87.4	3.5	17.6
+FDER+SFS+ar_BCELoss (MVF-YOLO)	86.0	87.5	80.9	96.4	87.1	87.6	3.5	17.6

注:加粗字体为精度最优结果。

从表4可见,本文提出的检测算法在雾合成图像上取得了最高的平均检测精度87.6%,在所有类别都取得了最佳的检测结果,相比于基线模型,mAP@0.5提高了4.4%,本章所提出的改进模块显著提升了算法在真实场景下的泛化能力与鲁棒性,证实了模块组合的整体有效性。

图8是数据集中的一幅测试集图像在不同网络结构下的热力图可视化结果,从中可以看出,基线模型YOLOv11n受噪声干扰,过多关注非目标信息,易造成漏检误检;网络加入FDER模块后恢复高雾密度区域的边缘特征,但是由于恢复过程由图像的雾密度分布主导,导致部分区域目标特征在信息传递



(a)输入图像 (b)基线模型 (c)+FDER (d)+FDER+SFS

图8 热力图可视化结果

Fig. 8 Heatmap Visualization Results((a) input image;(b) baseline model;(c) +FDER;(d) +FDER+SFS)

过程中丢失;在加入SFS后,网络通过空间注意力和频域选择协同处理,聚焦图像中更重要的区域,自适应的选取更具有任务相关性的特征,不仅缓解了特征信息的丢失,而且能更加准确地描述目标的分布。

2.5 实验结果展示及分析

为了更加直观的展示本文提出的检测效果,选取了部分测试集图像进行识别效果的对比,如图9所示。

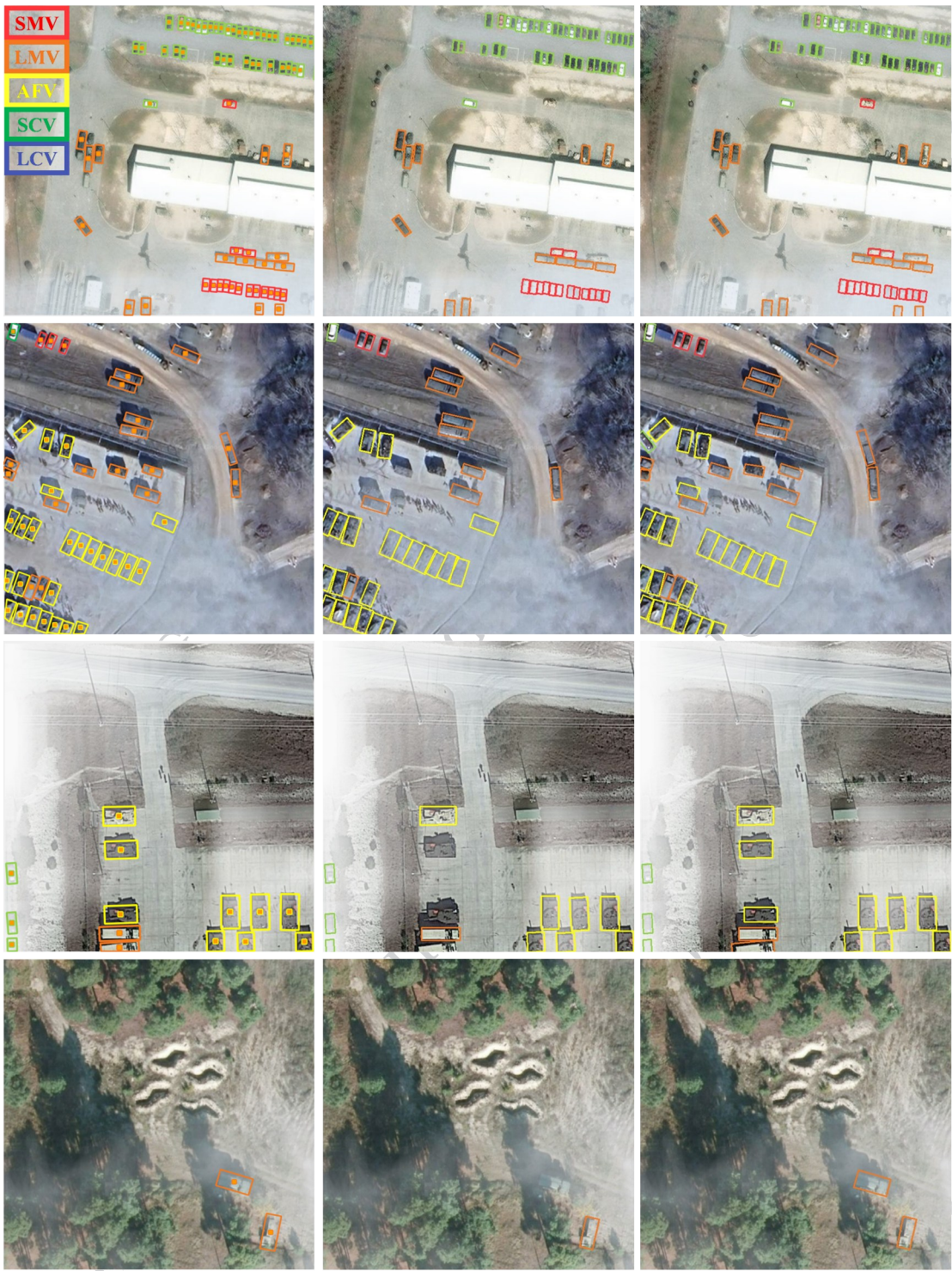
由图9可以看出,真实场景下,遥感图像上目标的边缘与纹理等关键细节被大量噪声掩盖,目标检测算法易出现漏检、误检。本文提出的算法,通过雾密度图引导恢复高雾密度区域的多尺度边缘特征,利用空间注意力和频域选择机制,有选择性、自适应地学习遮挡、伪装目标的关键高频信息,同时避免过度放大无关的高频噪声,有效提升了网络的检测精度和鲁棒性。

从表4可以看出,FDER与SFS组合使用时 $mAP@0.5$ 达87.4%,相比单独使用FDER(87.2%)和SFS(86.5%)均有进一步提升,表明两个模块在功能上具有互补性。FDER侧重于从输入端恢复因雾衰减的边缘信息,为后续特征提取提供更高质量的初始特征;SFS则从频域角度增强高频判别特征并抑制低频噪声,两者分别作用于网络的不同阶段,形成“前端恢复—后端增强”的协同机制。此外, $ar_BCELoss$ 作为分类层的损失函数约束,与前两个特征级模块在作用维度上正交,可叠加使用而无冲突,其引入几乎不增加计算开销,却使大型车辆(LMV、LCV)和特种车辆(AFV)的检测精度明显提升,体现了几何先验知识对数据驱动方法的有效补充。

3 结论

本文设计了一种雾密度引导的真实场景遥感图像检测算法,旨在提升目标检测的精度和鲁棒性;提出一种面向检测任务的遥感图像雾密度计算方法,通过融合暗通道、饱和度-亮度特征与多尺度梯度信息,构建更符合图像实际雾分布情况的雾密度图,并引入雾密度驱动的边缘特征恢复模块,自适应恢复高雾区边缘结构;设计空间注意力引导的频域选择模块,强化任务相关高频特征并抑制噪声,增强网络对遮挡、伪装目标特征的提取能力;利用遥感车辆几何特性优化目标分类损失,在几乎不增加计算量的基础上提升了部分目标类别的检测精度。在MVRSD-V2.0雾合成图像子集上的实验结果表明,本文方法在不依赖图像去雾的前提下,有效提升了真实场景下遥感军用车辆检测的精度与鲁棒性,并在效率与精度之间取得了良好平衡。

但本文的算法也存在局限性,仍有改进空间。本文重点研究了雾天条件下的检测问题,但真实战场环境还可能存在雨、雪、沙尘等多种极端天气,模型在更加极端或未见的场景下的泛化性能仍需进一步验证。未来研究将探索更具普适性的恶劣天气检测算法,并考虑引入SAR、红外等多模态遥感数据,利用其穿透性强、不受光照影响的特点,与光学图像进行互补融合,从根本上提升极端条件下战场目标感知的可靠性与鲁棒性。



(a)标注真实框的图像 (b)基线模型 (c)本文模型

图9 检测效果对比图

Fig. 9 Comparison of Detection Performance((a) image with ground-truth bounding boxes;(b) baseline model;(c) Our model)

参考文献(References)

Bai D, Yu Y, Song L, Cheng B B and Gao H. 2024. Construction and validation of remote sensing image dataset for fine-grained detec-

tion of military vehicles. Journal of Image and Graphics, 29(12): 3564-3577 (柏栋, 于英, 宋亮, 等. 面向军用车辆细粒度检测的遥感图像数据集构建与验证[J]. 中国图象图形学报, 2024, 29(12): 3564-3577) [DOI:10. 11834/jig. 230884]

Dang M, Liu G, Wang Q, Zhang Y Z, Wang D and Pan R. 2025. Review
© 中国图象图形学报版权所有

- of oriented object detection in aerial remote sensing images. *National Remote Sensing Bulletin*, 29(8): 2483-2510 (党敏, 刘刚, 王泉, 等. 航空遥感图像旋转目标检测技术研究综述[J]. 遥感学报, 2025, 29(08): 2483-2510 [DOI: 10. 11834/jrs. 20254504])
- Cao R, Wang X, Li H. Fog density evaluation by combining image gray-scale entropy and directional entropy [J]. *Atmosphere*, 2023, 14(7): 1125.
- Gao P, Tian T, Li L, et al. DE-CycleGAN: An object enhancement network for weak vehicle detection in satellite images [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 3403-3414.
- Hniewa M, Radha H. Multiscale Domain Adaptive Yolo For Cross-Domain Object Detection [C]. 2021 IEEE International Conference on Image Processing (ICIP), 2021: 3323-3327.
- Hwang S H, Kwon K W, Im T H. AEA-RDCP: an optimized real-time algorithm for sea fog intensity and visibility estimation [J]. *Applied Sciences*, 2024, 14(17): 8033.
- Li C, Zhou H, Liu Y, et al. Detection-Friendly Dehazing: Object Detection in Real-World Hazy Scenes [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023: 1-12.
- Li Z, Wang Y, Zhang Y, et al. Context feature integration and balanced sampling strategy for small weak object detection in remote sensing imagery [J/OL]. *IEEE Geosci. Remote Sens. Lett.*, 2024, 21: 1-5. DOI: 10.1109/LGRS.2024.3356507.
- Liu H, Jin F, Zeng H, et al. Image enhancement guided object detection in visually degraded scenes [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2023.
- Liu W, Ren G, Yu R, Guo S, Zhu J, Zhang L. Image-Adaptive YOLO for Object Detection in Adverse Weather Conditions [C]. AAAI Conference on Artificial Intelligence, 2021.
- Liu Y, Li H, Hu C, et al. Learning to aggregate multi-scale context for instance segmentation in remote sensing images [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, 36(1): 595-609.
- Liu Z, et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows [C]// IEEE/CVF International Conference on Computer Vision. 2021: 9992-10002.
- Miclea R C, Ungureanu V I, Sandru F D, et al. Visibility enhancement and fog detection: Solutions presented in recent scientific papers with potential for application to mobile systems [J]. *Sensors*, 2021, 21(10): 3370.
- Meng X, Liu Y, Fan L, et al. YOLOv5s-Fog: An improved model based on YOLOv5s for object detection in foggy weather scenarios [J]. *Sensors*, 2023, 23(11): 5321.
- Ma Y, Cai J, Tao J, et al. Foggy image detection based on dehazenet with improved ssd [C]//Proceedings of the 2021 5th international conference on innovation in artificial intelligence. 2021: 82-86.
- Song Y, He Z, Qian H, et al. Vision transformers for single image dehazing [J]. *IEEE Transactions on Image Processing*, 2023, 32: 1927-1941.
- Wang Y, Yan X, Zhang K, Gong L, Xie H, Wang F L, Wei M. TogetherNet: Bridging Image Restoration and Object Detection Together via Dynamic Enhancement Learning [J]. *Computer Graphics Forum*, 2022, 41:465-476.
- Xiao C, An W, Li Z X, Li B Y, Ying X Y, Lin Z P. 2024. Intelligent detection for moving targets in space-borne optical remote sensing: A review. *National Remote Sensing Bulletin*, 28(7): 1681-1692 (肖超, 安玮, 李朝旭, 等. 天基光学遥感动目标智能检测技术综述[J]. 遥感学报, 2024, 28(7): 1681-1692 [DOI:10. 11834/jrs. 20243277])

作者简介

于英,男,副教授,主要研究方向为遥感智能感知。E-mail: yuying5559104@163.com

柏栋,通信作者,男,硕士研究生,主要研究方向为遥感图像智能解译。E-mail:baidonglsq001@163.com

张珂,女,硕士研究生,主要研究方向为遥感影像智能处理。E-mail:betcha222@163.com

李凯,男,工程师,主要研究方向为军事物联网。E-mail: likai_rs@163.com

李磊,男,讲师,主要研究方向为遥感图像智能处理。E-mail: 3110100798@zju.edu.cn

李力,男,讲师,主要研究方向为遥感图像解译,低空场景三维重建。E-mail:lili315114@163.com

冯新月,女,硕士研究生,主要研究方向为遥感影像智能处理。E-mail:15210779806@163.com