

文章编号 1004-924X(2026)07-1142-14

可靠性自适应引导的红外与可见光图像融合

王琛^{1*}, 马庆禄¹, 周志超¹, 刘明²

(1. 重庆交通大学 交通运输学院, 重庆 400074;

2. 中冶赛迪信息技术(重庆)有限公司, 重庆 401122)

摘要:针对复杂交通场景下光照变化与跨模态干扰导致自动驾驶感知能力退化的问题,提出一种基于可靠性自适应引导的红外与可见光图像融合网络。该方法通过构建像素级可靠性度量机制,联合建模结构一致性与强度异常以动态评估信源可信度,在全局层采用“可信注入”策略校正强度分布,在细节层利用自适应引导滤波实现显著目标与纹理的竞争增强,并结合多约束损失函数协同优化。在 M3FD 与 RoadScene 数据集上的实验结果表明,相较于 DWT, GTF, U2Fusion 及 Umcfuse 等主流算法,本文方法在信息熵、标准差、空间频率、平均梯度、互信息、融合质量、边缘强度及视觉信息保真度等分别平均提高了 1.51%, 16.56%, 42.36%, 52.24%, 38.28%, 80.51%, 21.4% 和 17.6%;在下游目标检测任务中平均精确率达 91.4%, 优于其他融合方法。该方法有效抑制了伪影与噪声,具备优异的场景泛化性与稳定性,能显著提升自动驾驶系统的环境感知精度。

关键词: 图像融合; 红外与可见光; 可靠性自适应引导; 跨模态结构一致性; 可信注入; 自动驾驶感知

中图分类号: TP391.4; U495 **文献标识码:** A

doi: 10.37188/OPE.20263407.1142 **CSTR:** 32169.14.OPE.20263407.1142

Reliability adaptive guided infrared and visible image fusion

WANG Chen^{1*}, MA Qinglu¹, ZHOU Zhichao¹, LIU Ming²

(1. College of Transportation, Chongqing Jiaotong University, Chongqing 400074, China;

2. CISDI Information Technology Co., Ltd., Chongqing 401122, China)

* Corresponding author, E-mail: chenw@mails.cqjtu.edu.cn

Abstract: To mitigate perception degradation in autonomous driving caused by illumination variations and cross-modal interference, an infrared-visible image fusion network with reliability-adaptive guidance is proposed. A pixel-level reliability estimation mechanism is constructed by jointly modeling structural consistency and intensity anomalies, enabling dynamic assessment of source credibility. A “trusted injection” strategy is introduced to correct the global intensity distribution, while adaptive guided filtering enhances the competition between salient objects and texture details in the detail layer; the process is optimized by a multi-constrained loss function. Experiments on the M3FD and RoadScene datasets demonstrate that, compared with DWT, GTF, U2Fusion, and Umcfuse, the proposed method improves EN, SD, SF, AG, MI, Qabf, EI, and VIFF by 1.51%, 16.56%, 42.36%, 52.24%, 38.28%, 80.51%,

收稿日期: 2026-02-04; 修订日期: 2026-03-03.

基金项目: 国家自然科学基金资助项目(No. 52072054); 重庆市交通科技资助项目(No. CQJT-CZKJ2025-07); 重庆市 2025 研究生科研创新项目(No. CYS25535); 重庆市自然科学基金面上项目(No. CSTB2023NSCQ-MSX0551)

21.4%, and 17.6%, respectively. In downstream target detection tasks, an average accuracy of 91.4% is achieved, surpassing existing fusion methods. The proposed approach effectively suppresses artifacts and noise, exhibits strong scene generalization and robustness, and significantly enhances environmental perception accuracy in autonomous driving systems.

Key words: image fusion; infrared and visible; reliability-adaptive guidance; cross-modal structural consistency; trusted injection; autonomous driving perception

1 引言

自动驾驶车辆的安全运行高度依赖于系统对环境的感知能力。但在复杂动态场景下,光学传感器常面临光照不足、局部过曝光及眩光溢出等物理挑战,引发成像信噪比跌落与关键纹理退化,从而制约下游感知算法在复杂工况下的目标识别精度,成为保障自动驾驶行车安全的关键技术瓶颈^[1]。

现有的红外与可见光图像融合算法涵盖传统与深度学习两大体系。其中,传统算法依赖像素级操作、变换域分析及多尺度分解等手段。任鹏百等^[2]构建了一种混合多尺度分解模型并结合改进的脉冲耦合神经网络,有效缓解了边缘细节的受损风险。韩阳等^[3]和孙磊等^[4]分别引入非下采样剪切波变换及自适应模糊集理论,通过显著信息加权与子带增强策略以同步提升图像对比度及清晰度。李英超等^[5]和Lyu等^[6]借助非下采样轮廓波变换优异的方向选择及各向异性特征实现了对图像轮廓边缘细节的精准提取。这些方法在不同程度上解决了融合中的关键问题,但仍存在计算复杂度高、特定场景适应性不足等挑战,难以满足自动驾驶感知对稳定性与效率的双重需求。

随着深度学习在计算机视觉领域的快速发展,研究人员逐渐将数据驱动的建模思想引入红外与可见光图像融合中。针对特征解耦与交互,张永兴等^[7]针对融合过程中红外冗余信息过度引入导致复杂场景信息失衡的问题,提出了一种基于多尺度空间注意力互补的融合网络。Xian等^[8]和Zhao等^[9]均通过嵌入Transformer模块,强化了算法在不同光照下的泛化能力及上下文捕获效率。Peng等^[10]提出了FusionMamba,在保持计算复杂度的同时实现了对多模态全局特征的高效建模。邸敬等^[11]为解决可见光利用率不足难题,

提出了条件扩散与多通道高低频并行模型,显著优化了特征提取的最优分布。针对低照度及复杂光照环境,杨艳春等^[12]设计了基于Nest架构的语义注入式融合网络,通过低光增强分支与梯度流的联合驱动,实现了语义先验对多尺度特征的逐级修正。Shi等^[13]提出了通用的无监督扩散框架VDMUFusion以解决此问题。另外,Tang等^[14]提出的纹理-对比度增强网络试图在弱光增强与信息融合间寻找平衡,实现弱光图像增强与融合的耦合。姚睿等^[15]提出跨模态特征交互融合模块,对红外显著特性和可见光纹理特性进行互补融合,以得到高保真的融合图像。强光环境下,张晓东等^[16]和赵阳等^[17]均通过构建自编码器融合网络,对源图像的特征进行多尺度的提取并进行重构,有效增强了纹理信息和显著性信息。基于深度学习的红外与可见光视觉融合方法在特征表达能力和跨模态信息协同建模方面显著优于传统融合方法,能够更有效地捕获多尺度特征与全局上下文信息,在弱光、强光及复杂环境条件下展现出更强的适应能力。然而,现有方法仍普遍存在对特定场景或数据分布依赖较强、模型结构复杂、泛化能力有待提升等问题。

综上所述,现有方法在处理复杂交通场景中极端光照及红外伪结构与可见光过曝等跨模态退化干扰问题时,存在显著的泛化瓶颈,难以维持全工况下感知的稳定性与时效性。为此,本文构建一种基于可靠性自适应引导的红外与可见光图像融合网络。该方法核心改进在于构建了像素级可靠性度量机制,通过联合建模结构一致性与强度异常分布,动态判别并抑制源图像中的不可靠信息,增强算法对多变光照及跨模态噪声的自适应抑制能力,从物理一致性层面增强了算法对非平稳环境的适应能力;结合全局可信注入与自适应引导增强策略,确保显著目标与纹理特征的深度协同,使模型在未参与训练的未知场景

下依然能保持清晰的轮廓边缘与一致的结构特征,从而显著改善复杂环境下的感知稳定性。

2 可靠性引导的红外与可见光图像融合方法

2.1 网络总体框架

为了提升复杂交通场景下红外与可见光图像的融合质量,克服传统算法及现有深度学习模型在极端光照、噪声干扰下稳定性及泛化性不足的问题,本文提出了一种基于可靠性自适应引导的红外与可见光图像融合框架。

该框架首先通过构建跨模态可靠性映射,联合建模结构一致性得分与强度异常惩罚项,为后续融合过程提供物理可信度的引导约束,该模块生成的像素级可靠性权重图能够动态引导网络在复杂光照和热噪声干扰下识别出高质量的信

息源。在全局层面,实施可信注入策略以消除模态间的灰度分布差异,确保融合图像的亮度尺度自然稳定。在细节层面,引入自适应引导滤波技术,利用可靠性调制的显著性竞争权重,使细节增强始终紧随可信的结构特征,有效避免了伪影的生成。最后,本文将融合过程参数化为可学习映射,通过设计包含强度保持、梯度一致及结构相似性的多约束加权损失函数,实现在复杂动态环境下多准则信息的自适应平衡与协同优化。该方法的泛化性边界主要限于城市及郊区等具有跨模态互补特征的道路工况。其在面对双模态信号同步跌落至成像极限的极端恶劣工况时,依靠空间信息的可靠性度量将趋于饱和失效。通过对此泛化边界的定量约束与定性分析,本文为自动驾驶感知系统在复杂长尾场景下的失效判定与策略切换提供了科学依据,网络框架如图 1 所示。

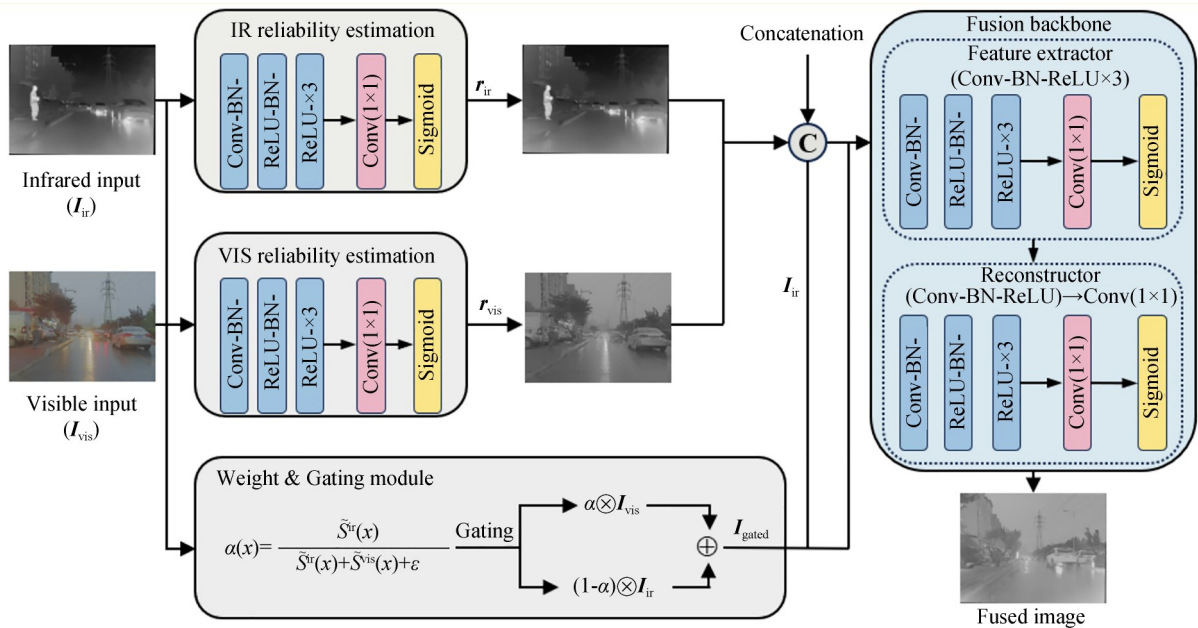


图 1 可靠性自适应引导的红外与可见光融合框架

Fig. 1 Reliability adaptive guided infrared and visible fusion framework

2.2 可靠性约束下的全局强度与结构融合策略

为避免复杂环境下红外噪声、可见光过曝及欠曝对融合结果造成系统性的偏移,通过构建跨模态可靠性映射,并以此约束全局强度与结构信息的融合。首先,考虑输入 I_{ir}, I_{vis} ,为消除不同模态灰度分布差异,先进行全局标准化,得到:

$$\begin{cases} \hat{I}^{(\cdot)}(x) = \frac{I^{(\cdot)}(x) - \mu^{(\cdot)}}{\sigma^{(\cdot)} + \epsilon} \\ \mu^{(\cdot)} = \frac{1}{|\Omega|} \sum_{x \in \Omega} I^{(\cdot)}(x) \\ \sigma^{(\cdot)} = \sqrt{\frac{1}{|\Omega|} \sum_{x \in \Omega} (I^{(\cdot)}(x) - \mu^{(\cdot)})^2} \end{cases}, \quad (1)$$

其中: $\hat{I}^{(\cdot)}$ 为全局归一化结果, ε 为最小常数,防止除零。为了在融合输出中维持可控的整体亮度尺度,定义全局强度基准项,如下:

$$I_g(x) = \eta \hat{I}_{ir}(x) + (1 - \eta) \hat{I}_{vis}(x), \quad (2)$$

其中: $I_g(x)$ 表示在像素点 x 处的全局强度基准,决定融合图像的亮度基座, $\eta \in [0, 1]$ 为控制红外与可见光在全局层面的强度占比。全局对齐后,进一步判断局部结构是否稳定一致。定义梯度幅值与局部结构能量为:

$$\begin{cases} G^{(\cdot)}(x) = \|\nabla \hat{I}^{(\cdot)}(x)\|_2 \\ E^{(\cdot)}(x) = \frac{1}{|\Omega(x)|} \sum_{y \in \Omega(x)} (G^{(\cdot)}(y))^2 \end{cases}, \quad (3)$$

其中: $G^{(\cdot)}$ 描述细节响应, $E^{(\cdot)}$ 描述局部结构稳定性,能量越大,结构越强。由此,构建跨模态结构偏差如下:

$$\begin{cases} \Delta G(x) = |G^{ir}(x) - G^{vis}(x)| \\ \Delta E(x) = |E^{ir}(x) - E^{vis}(x)| \end{cases}, \quad (4)$$

其中: ΔG 表示瞬时边缘差异, ΔE 表示局部结构能量差异, $G^{ir}(x), G^{vis}(x)$ 分别表示红外与可见光图像的细节响应, $E^{ir}(x), E^{vis}(x)$ 分别为红外与可见光图像的局部结构能量。将两类偏差合成为一致性得分:

$$C(x) = \exp(-\alpha \Delta G(x) - \xi \Delta E(x)), \quad (5)$$

其中: $C(x) \in (0, 1]$ 越大表示模态结构越一致, α, ξ 为权重,用于调节两类偏差对最终一致性得分 $C(x)$ 的贡献度。复杂场景下,噪声与过曝引发结构和强度异常。结构一致性与强度异常共同决定像素可靠性。为弥补单一一致性在过曝饱和区的缺陷,引入强度异常惩罚项如下:

$$A^{(\cdot)}(x) = \frac{|\hat{I}^{(\cdot)}(x)|}{\kappa + |\hat{I}^{(\cdot)}(x)|}, \quad (6)$$

其中: $A^{(\cdot)} \in [0, 1)$ 为强度异常惩罚项,其值越接近1表示该像素点在物理一致性层面的异常程度越高,可信度越低; $\hat{I}^{(\cdot)}(x)$ 表示像素点 x 处的灰度均值与该模态全图均值的残差,其值越大表示越偏离全局均值; κ 为控制函数饱和度的常数因子,决定了惩罚项对强度异常的敏感阈值,取值0.01。由于可见光过曝阴影与红外热噪声等

退化,导致结构不稳定和强度异常。通过构建像素级信息可靠性度量,评估跨模态信息可信贡献,并经归一化作为权重引导融合。归一化得到:

$$\begin{cases} R_{ir}(x) = C(x) \cdot \exp(-\rho A^{ir}(x)) \\ R_{vis}(x) = C(x) \cdot \exp(-\rho A^{vis}(x)) \\ \bar{R}_{ir}(x) = \frac{R_{ir}(x)}{R_{ir}(x) + R_{vis}(x) + \varepsilon} \\ \bar{R}_{vis}(x) = 1 - \bar{R}_{ir}(x) \end{cases}, \quad (7)$$

其中: R_{ir}, R_{vis} 分别为两种模态的可靠性,同时受结构一致性与强度异常抑制, $\bar{R}_{ir}(x), \bar{R}_{vis}(x)$ 分别为归一化后的可靠性权重, ρ 为抑制系数。得到全局强度与结构融合的可信形式,如下:

$$F_g(x) = I_g(x) + \lambda_s (\bar{R}_{ir}(x) G_{ir}(x) + \bar{R}_{vis}(x) G_{vis}(x)), \quad (8)$$

其中: I_g 负责全局亮度, λ_s 控制结构增强强度。在复杂交通场景中,融合不仅需要全局结构,还需要突出可行驶区域中的显著目标细节。因此,在可靠性约束下可以进一步引入自适应引导机制,实现细节与显著目标的选择性增强。

2.3 自适应引导下的细节与显著目标融合策略

为实现复杂场景下的细粒度增强,通过在细节层引入自适应引导融合策略。在保证可靠性的前提下,让细节增强跟随可信结构,并促使显著目标跨模态自适应竞争。通过对图像基础层、细节层分解,构建可靠性调制的显著性图,并利用自适应引导滤波生成细节层融合结果。先将两模态分解为基础层与细节层,得到:

$$\begin{cases} B^{(\cdot)}(x) = B(\hat{I}^{(\cdot)}(x)) \\ D^{(\cdot)}(x) = \hat{I}^{(\cdot)}(x) - B^{(\cdot)}(x) \end{cases}, \quad (9)$$

其中: $B^{(\cdot)}$ 为低频基础层, $D^{(\cdot)}$ 为高频细节层, $B(\cdot)$ 为边缘保持平滑算子。为了让细节增强不引入噪声,采用局部线性引导模型:

$$F_D(x) = a_k \hat{I}_{vis}(x) + b_k, \quad \forall x \in w_k, \quad (10)$$

其中: a_k, b_k 为待求参数;在窗口 w_k 内,细节融合输出 F_D 表示对引导图的线性变换; $\hat{I}_{vis}(x)$ 为经过预处理的可见光引导信息。因为细节融合需要对值得保留的细节进行判别,先构建显著性响

应,如下:

$$\begin{cases} S^{(\cdot)}(x) = E^{(\cdot)}(x) + \tau |\hat{I}^{(\cdot)}(x) - \mu_{\Omega(x)}^{(\cdot)}| \\ \mu_{\Omega(x)}^{(\cdot)} = \frac{1}{|\Omega(x)|} \sum_{y \in \Omega(x)} \hat{I}^{(\cdot)}(y) \end{cases}, \quad (11)$$

其中: $S^{(\cdot)}(x)$ 为显著性响应, $\mu_{\Omega(x)}^{(\cdot)}$ 表示窗口 $\Omega(x)$ 内像素的平均值,强调局部对比, τ 为调节结构项与对比项的权重。自适应引导权重 $\alpha(x)$ 受显著性竞争与可靠性调制共同支配,核心控制细节注入。它确保增强围绕可信结构,在平衡稳定性的同时自适应引入显著目标,规避噪声伪影。进一步构建显著目标竞争权重并归一化,得到:

$$\begin{cases} \tilde{S}^{(\cdot)}(x) = (\bar{R}^{(\cdot)}(x))^{\beta} S^{(\cdot)}(x) \\ \alpha(x) = \frac{\tilde{S}^{ir}(x)}{\tilde{S}^{ir}(x) + \tilde{S}^{vis}(x) + \epsilon} \end{cases}, \quad (12)$$

其中: $\tilde{S}^{(\cdot)}(x)$ 为经过可靠性调节的显著性响应; β 为非线性指数,可靠性越低,显著性越被压制; $\alpha(x) \in [0, 1]$,红外显著性越强则 α 越大; ϵ 为极小常数。为了让细节增强沿着可信边缘扩散,构造自适应引导图,并在窗口 w_k 内对引导系数 a_k, b_k 进行带正则的最小二乘估计,令目标响应由两模态细节层按显著权重融合,即:

$$\begin{cases} I_g^{ad}(x) = \bar{R}_{vis}(x) \hat{I}_{vis}(x) + \bar{R}_{ir}(x) \hat{I}_{ir}(x) \\ \min_{a_k, b_k} \sum_{x \in w_k} (a_k I_g^{ad}(x) + b_k - T(x))^2 + \epsilon_a a_k^2 \\ T(x) = \alpha(x) D^{ir}(x) + (1 - \alpha(x)) D^{vis}(x) \end{cases}, \quad (13)$$

其中: $I_g^{ad}(x)$ 为构造的自适应引导图,在可见光可靠处更接近可见光,在红外可靠处更接近红外; $T(x)$ 为由两模态细节层按显著权重 $\alpha(x)$ 融合得到的目标响应; ϵ_a 为最小二乘估计中的正则化参数,防止系数 a_k 过大导致伪影。显著目标在红外强则偏向红外细节,反之偏向可见光纹理。即:

$$\begin{cases} a_k = \frac{\text{Cov}_{w_k}(I_g^{ad}, T)}{\text{Var}_{w_k}(I_g^{ad}) + \epsilon_a} \\ b_k = \mu_{w_k}(T) - a_k \mu_{w_k}(I_g^{ad}) \end{cases}, \quad (14)$$

其中: $\mu_{w_k}(\cdot)$ 为窗口均值,Cov,Var为窗口方差与协方差。最终细节层融合输出为:

$$F_d(x) = a(x) I_g^{ad}(x) + b(x), \quad (15)$$

其中: $a(x), b(x)$ 为由窗口重叠平均得到的像素

级系数。虽然在可靠性约束下的细节增强,但若仅依靠显著性竞争,仍可能在极端噪声或强光场景产生局部过增强。为此,将融合过程写为可优化的目标函数,通过强度、梯度、结构相似与可靠性一致性的联合损失来稳定训练与收敛。

2.4 损失函数设计

在复杂场景下,多约束之间存在非线性耦合与冲突,仅依靠解析形式难以全局协调。因此,将融合过程参数化为可学习映射,以实现多约束的自适应平衡。基于可靠性的全局强度与结构融合结果 F_g ,以及自适应引导下的细节与显著目标融合结果 F_d 。为实现二者的协同优化,将最终融合结果 F 表述为可学习映射,如下:

$$F(x) = \Phi(F_g, F_d; \theta), \quad (16)$$

其中: $\Phi(\cdot)$ 为参数化融合映射, θ 为网络参数。为了在可靠区域内最大限度继承各模态的有效信息,同时在不可靠区域主动抑制噪声与伪结构。将整体优化问题建模为一个可靠性加权的多约束最小化问题,得到:

$$\theta^* = \arg \min_{\theta} L(F; I_{ir}; I_{vis}; R_{ir}; R_{vis}). \quad (17)$$

为在可信区域内保持对源模态强度分布的忠实性。对全图积分得到强度保持损失,即:

$$\begin{cases} \epsilon_{int}(x) = \bar{R}_{ir}(x) (F(x) - \hat{I}_{ir}(x))^2 + \\ \bar{R}_{vis}(x) (F(x) - \hat{I}_{vis}(x))^2 \\ L_{int} = \sum_{x \in \Omega} \epsilon_{int}(x) \end{cases}. \quad (18)$$

另外,仅保持强度一致并不能保证结构清晰,进一步引入梯度约束,如下:

$$\begin{cases} G_{ref}(x) = \bar{R}_{ir}(x) G^{ir}(x) + \bar{R}_{vis}(x) G^{vis}(x) \\ G_F(x) = \|\nabla F(x)\|_2 \\ \epsilon_{grad}(x) = C(x) |G_F(x) - G_{ref}(x)| \\ L_{grad} = \sum_{x \in \Omega} \epsilon_{grad}(x) \end{cases}. \quad (19)$$

梯度约束强调局部变化,但对整体结构感知仍不充分,因此引入结构相似性约束,即:

$$\begin{cases} \epsilon_{ssim}(x) = \bar{R}_{ir}(x) (1 - \text{SSMI}(F, \hat{I}_{ir})) + \\ \bar{R}_{vis}(x) (1 - \text{SSMI}(F, \hat{I}_{vis})) \\ L_{ssim} = \sum_{x \in \Omega} \epsilon_{ssim}(x) \end{cases}. \quad (20)$$

由于显著性权重直接影响细节注入强度,若其空间变化过剧,易产生局部振铃或伪增强。因此引入可靠性一致性正则,即:

$$L_{rel} = \sum_{x \in \Omega} |\alpha(x) - \bar{R}_{ir}(x)| G_F(x), \quad (21)$$

其中: L_{rel} 为可靠性一致性正则项,以约束显著性选择权重 $\alpha(x)$ 与物理可靠性 $\bar{R}_{ir}(x)$ 在结构区域保持一致。该项约束显著性选择与可靠性判断在结构区域保持一致,避免显著性偏离物理可信性。综合上述各项,最终融合优化目标为:

$$L = \lambda_1 L_{int} + \lambda_2 L_{grad} + \lambda_3 L_{ssim} + \lambda_4 L_{rel}, \quad (22)$$

其中: $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ 分别为强度、梯度、结构相似性及可靠性一致性损失的加权因子,使模型在复杂环境下具备良好的稳定性与一致性。

3 实验与分析

3.1 实验设置

为验证所提方法在自动驾驶感知场景下的有效性与稳定性,实验选取涵盖显著光照变化及复杂背景的M3FD与RoadScene数据集,严格遵循统一的数据划分(即训练集:验证集为7:3)与预处理标准并进一步将融合结果接入下游检测

任务以量化它对自动驾驶感知系统的实际增益。所有实验均在NVIDIA RTX 4060 GPU上进行。

对比基线涵盖GTF^[18],DWT^[19]等传统算法及U2Fusion^[20],Umcfuse^[21]等深度学习模型,以全方位评估本文方法相对于经典策略及主流框架的性能优势,且所有深度模型均基于公开预训练权重进行统一推理以排除干扰。在评价方式上,本文从信息量、对比度、纹理细节、边缘强度、视觉信息保真度以及结构保持等多个角度对融合结果进行客观评估,选用EN,SD,SF,AG,MI,EI,VIF以及Qabf等评价指标。为避免单一统计量掩盖融合结果在不同场景下的差异性,对各指标在全数据集范围内的分布特性进行分析,从而更全面地反映各方法在复杂环境下的稳定性与适应能力。

3.2 对比分析

从M3FD数据集中选取包括行人、车辆等目标特征的典型应用场景,并覆盖遮挡、强光、复杂背景等共3种具有代表性的红外与可见光图像场景作为源图像进行图像融合实验。通过视觉效果对比和量化指标对比验证本文融合方法的有效性和可行性。图像融合效果如图2所示。

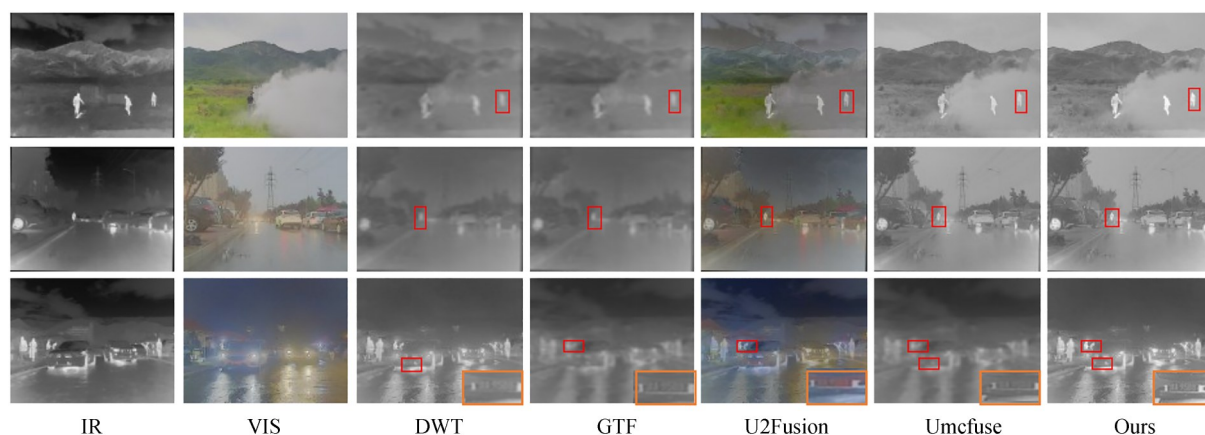


图2 基于M3FD数据集的融合结果对比

Fig. 2 Comparison of fusion results based on M3FD dataset

图2直观揭示了各算法在复杂环境下平衡红外显著目标与可见光纹理时的性能分化。受限于固定频域规则,DWT在场景一、二中未能有效突显红外行人及车辆目标的边缘对比度,且夜间融合结果整体亮度偏低;GTF虽改善了边缘表

现,却因对光照噪声敏感而在场景二强反射区域引发局部亮度失衡。深度学习方法中,U2Fusion虽提升了全局对比度,却倾向于在复杂光照下产生局部过增强导致目标背景层级模糊;Umcfuse虽具备全局建模优势,但在场景三夜间强反射路

面未能充分强化红外目标,致使边界淹没于背景之中。相比之下,本文方法得益于可靠性约束引导,不仅有效抑制了可见光过曝及多源噪声干扰,杜绝了伪影生成,更保证了红外目标的轮廓连续性。场景三的局部放大细节进一步证实了该算法在车牌等细微纹理恢复上的优势,实现了显著目标与背景纹理的深度协同,充分验证了其在各类挑战性场景下规避目标弱化与纹理丢失的稳定性。为客观地对比融合过程中的信息损耗及伪信息引入情况,各融合方法的差值图对比效果如图3所示。

图3进一步揭示了不同策略在信源保持上的差异。其中,DWT因受限于固定频域规则,在红外行人及车辆区域产生显著差异,表明在融合中红外显著目标的能量遭到了大幅衰减。GTF融合结果的结构边缘伴随条带状伪影,其对局部噪声过于敏感导致将反射区域的亮度波动误判为结构信息,诱发了局部的亮度失衡。在主流的深度模型融合中,U2Fusion在高亮反射区的局部集

中高响应表明其在复杂光照环境下存在的过度增强倾向,导致融合结果与源图像产生系统性偏离。Umcfuse的全局差值分布虽相对均匀,但在红外目标处仍存在残留连续响应,反映该方法在跨模态机制中对显著目标的选择性强化能力不足,进而导致关键特征淹没于背景纹理之中。然而,本文方法在各场景融合过程中均能够保持最低的整体响应强度,背景区域差值呈现稀疏不连续分布,验证其跨模态可靠性映射机制的有效性。该方法借助强度异常惩罚项机制精准识别可见光过曝及红外噪声区域,并主动抑制不可靠信息的注入,能够在非目标区域维持与源图像的高度一致。而关键目标区域受控的集中响应特征则表明“可信注入”的策略实现了基于可靠性调制的选择性增强而非无约束放大。通过差值图对比,说明本文方法在降低信息偏移、抑制伪结构及维持融合稳定性方面具有优势,验证了可靠性自适应引导机制在复杂环境下的稳定性。定量结果如图4所示。

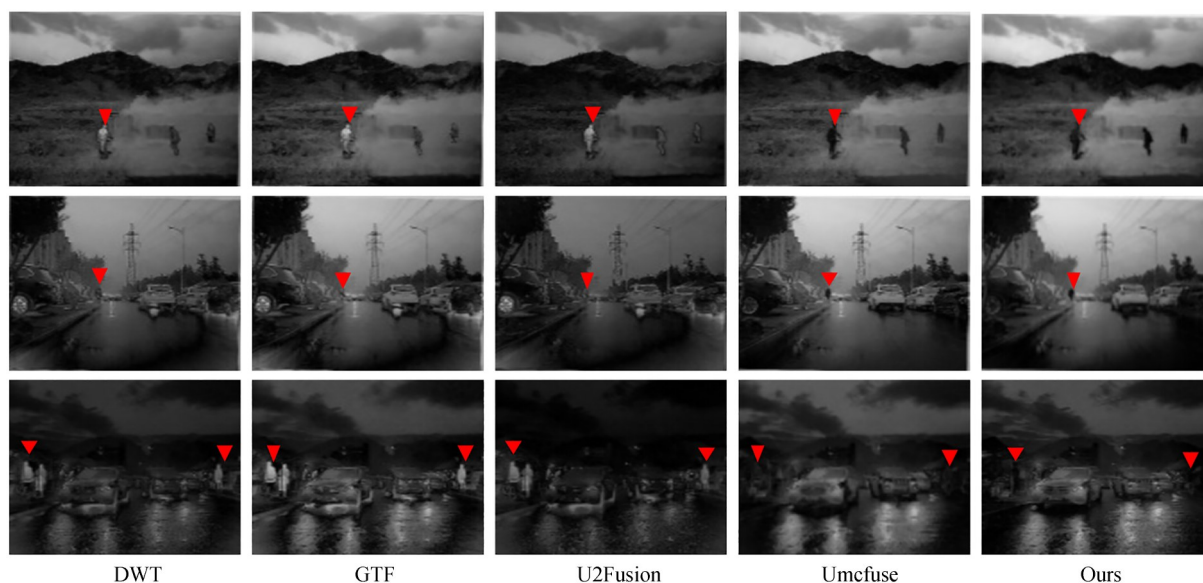


图3 各融合方法差值图对比

Fig. 3 Comparison of difference maps of each fusion method

由图4可知,本文方法在EN与SD指标上始终保持在较高水平且波动较小,表明可靠性映射机制能够从多源场景中稳定地提取有效信息,有效解决了传统算法在极端光照条件下因信息丢失或噪声干扰导致的指标不稳定性。相比之下,U2Fusion在部分样本上的数值起伏较大,表明其

对复杂光照场景下的适应性不足。在AG与SF指标上,本文方法仍处于领先地位,克服了DWT,GTF等算法在降噪过程中引起的细节平滑问题,实现了图像清晰度与稳定性的平衡。另外,MI指标的占优以及Qabf曲线在场景切换时的平滑表现,验证了可靠性加权损失函数在抑制

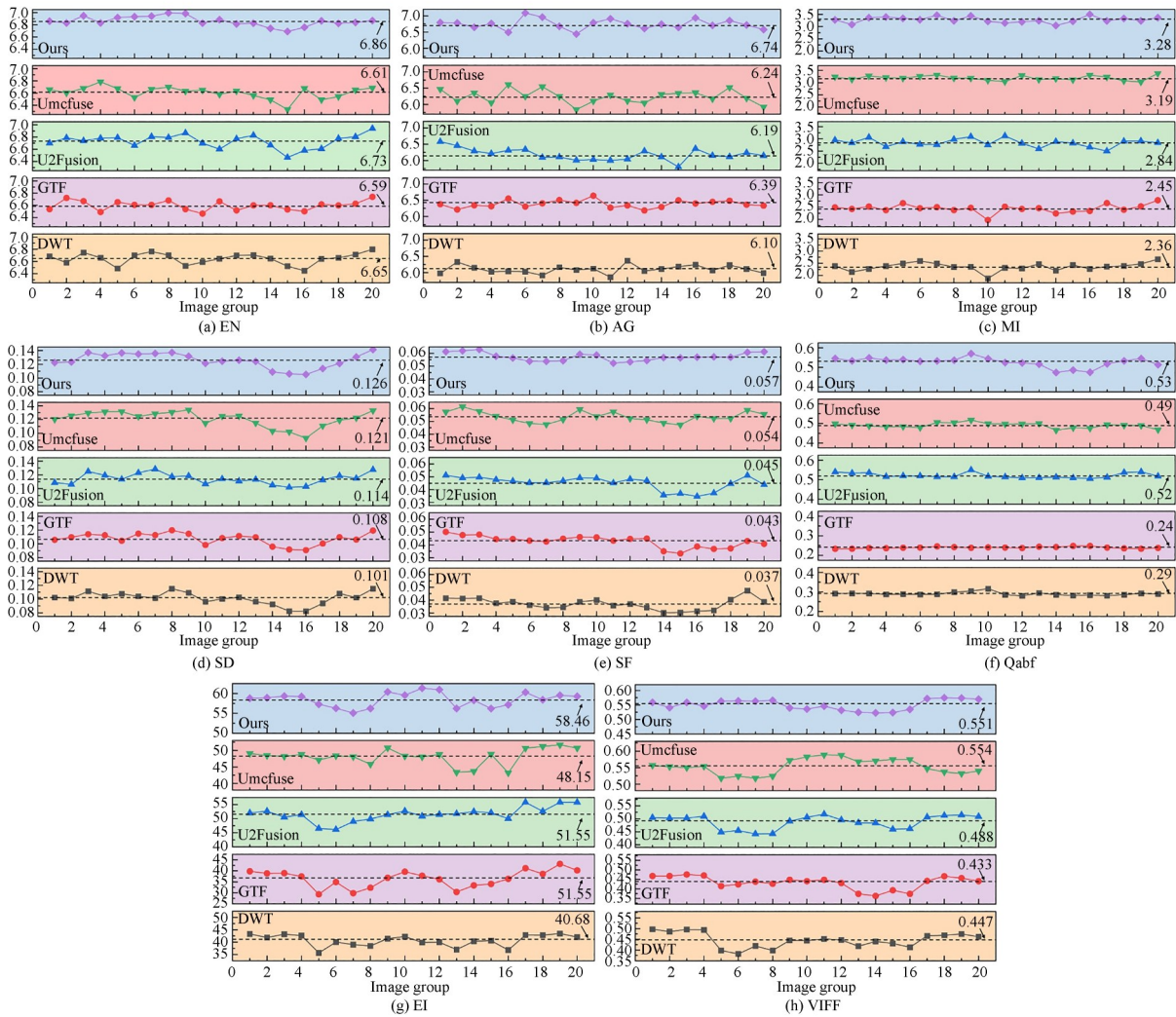


图 4 基于 M3FD 数据集的融合结果对比

Fig. 4 Comparison of fusion results based on M3FD dataset

伪影及保持结构一致性方面的作用。在 EI 指标上,本文方法均保持显著领先,说明自适应引导滤波能够产生更强的边缘响应;在 VIFF 方面,虽然 Umcfuse 的均值最优,但在整体波动幅度却高于本文方法,说明其稳定性不足。综合表明,本文算法具有高均值、低方差的特性,表明其在动态交通环境下具有良好的稳定性。为验证算法的泛化能力,进一步在 RoadScene 数据集上开展泛化实验,实验效果如图 5 所示。

由图 5 可见,DWT 与 GTF 等传统方法因融合规则固定,在处理非平稳噪声时自适应调节能力不足,导致在逆光条件及强干场景下出现亮度失衡。U2Fusion 等模型在跨数据分布时稳定性较差,在强反射区因特定的特征先验易诱发伪

影,致使目标边缘模糊。相比之下,本文方法凭借像素级可靠性感知机制,精准评估信源可信度,在提升对比度的同时能够有效抑制干扰噪声。实验表明,该方法在极端光照及动态干扰下仍能保持清晰的轮廓与纹理,展现出极强的泛化稳定性,定量对比如图 6 所示。

图 6 反映了各算法在 RoadScene 数据集上的性能稳定性差异。在表征信息量与对比度的 EN 和 SD 指标上,本文方法凭借紧凑的四分位区间的分布形态,证实了其在亮度平衡与信息整合上的可靠性,克服了 U2Fusion 等模型指标离散度大的缺陷。针对细节刻画,SF 与 AG 的中位数分布优势明显,验证了本文方法在边缘增强方面的稳定性,有效解决了传统算法因过度平滑导致的

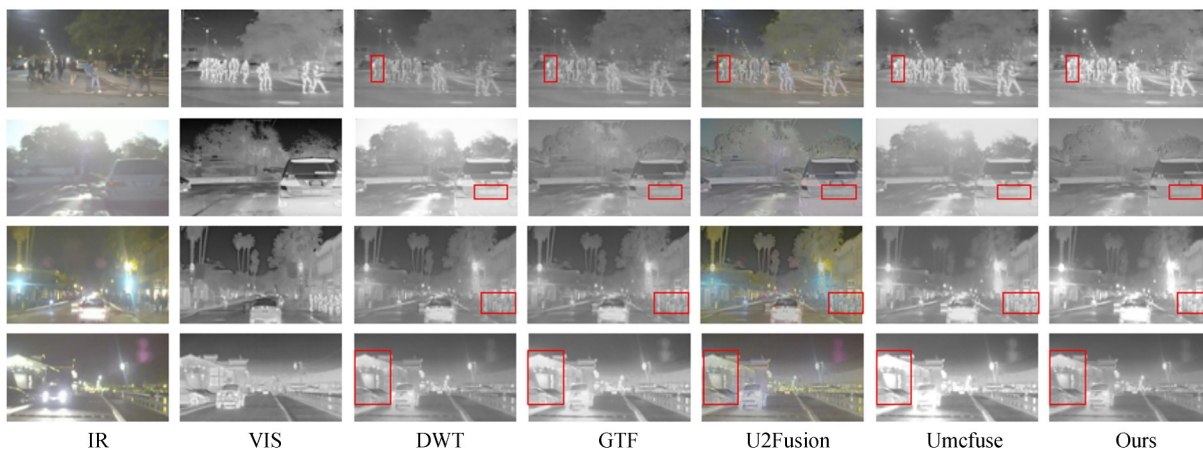


图 5 基于 RoadScene 数据集的融合结果对比

Fig. 5 Comparison of fusion results based on RoadScene dataset

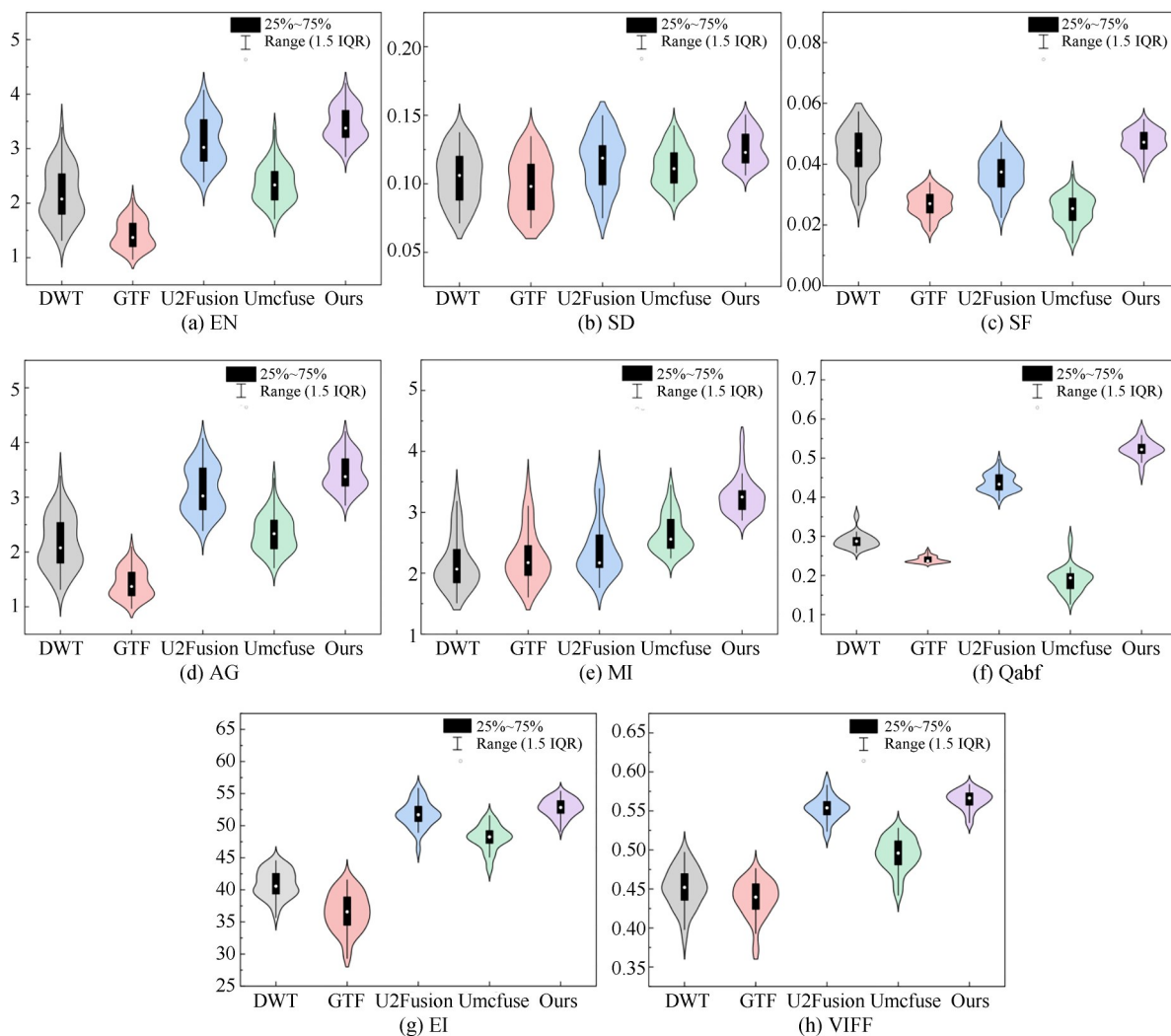


图 6 基于 RoadScene 数据集的评价指标对比

Fig. 6 Comparison of evaluation indexes based on RoadScene dataset

性能不足。此外,MI与Qabf的稳定表现验证了本文方法在信息继承与结构一致性上的优势。进一步观察VIFF与EI指标分布可知,在全数据集范围内,本文方法均呈现出高均值、紧凑分布的形态。特别是在EI指标上,本文方法的四分位区间显著高于对比算法,证实了可靠性自适应引

导机制对边缘特征提取的稳健性。VIFF的高位分布印证了本文策略在保留物理可信信源、提升整体感知保真度方面的优势。为了更全面地量化各算法在跨域场景下的综合性能,各算法在RoadScene数据集上的定量评价结果如表1所示。

表1 RoadScene数据集上的指标均值与效率对比

Tab. 1 Comparison of index mean and efficiency on RoadScene dataset

方法	EN	AG	MI	SD	SF	Qabf	EI	VIFF	FPS	Parameters	FLOPs
DWT	6.64	2.16	1.64	0.105	0.044	0.289	40.9	0.451	32.3	—	—
GTF	6.55	1.43	2.41	0.099	0.027	0.241	36.5	0.437	19.7	—	—
U2Fusion	6.76	3.13	2.87	0.116	0.037	0.437	51.8	0.494	9.2	0.66	42.5
Umcfuse	6.51	2.34	3.13	0.112	0.025	0.191	48.1	0.554	20.3	0.85	26.2
Ours	6.72	3.45	3.27	0.126	0.047	0.523	52.8	0.564	26.8	0.54	28.4

注:表中加粗字体表示此种指标下的最优值。

由表1可知,本文方法在反映图像信息量与边缘特征的SD,SF及EI等多项指标上均取得了最优均值,显著优于DWT与GTF等传统方法。在复杂度方面,由于DWT与GTF为基于解析理论的传统融合方法,不涉及深度学习中的可学习权重更新,故其参数量与FLOPs不参与对比,其时效性优势直接由FPS体现。本文模型的参数量仅为0.54M,最低,虽然引入像素级可靠性权重使得FLOPs略高于Umcfuse,但感知质量获得了极大提升。结合26.8FPS的推理速度,本文方法在保持轻量化的同时,兼顾高性能与实时性,能够满足自动驾驶感知系统对全天候环境实时处理的工程需求。

3.3 消融实验

为量化各功能模块的独立贡献及协同增益,本文构建层级化消融实验:以基础特征提取架构

为基准(A组),依次引入可靠性映射(B组)、可信注入机制(C组)、自适应引导模块(D组)及多约束加权损失(Ours),对比结果见表2。

由表2可知,基准A组引入可靠性映射后,EN由6.43升至6.58,VIFF从0.384提升至0.452,证实了像素级权重对可信信源的筛选作用。随后加入可信注入与自适应引导模块,使反映边缘特性的EI指标从44.6阶跃至53.1,AG从5.98增至6.51,显著强化了图像的结构清晰度与边缘保真。最终,本文方法在EN,AG,MI,SD及VIFF上均达到最高值,分别为6.83,6.78,3.27,0.155及0.564。虽然SF与Qabf较D组略有微调,但VIFF的持续增长验证了多约束损失在抑制伪影、维系视觉感知自然度方面的优化偏向,最终实现了增强力度与感知保真的深度协同。

表2 消融实验结果对比

Tab. 2 Comparison of ablation experimental results

模块	EN	AG	MI	SD	SF	Qabf	EI	VIFF
A	6.43	4.69	1.64	0.048	0.031	0.317	32.5	0.384
B	6.58	5.12	2.41	0.074	0.042	0.366	38.2	0.452
C	6.64	5.98	2.87	0.113	0.047	0.415	44.6	0.491
D	6.72	6.51	3.13	0.141	0.054	0.499	53.1	0.548
Ours	6.83	6.78	3.27	0.155	0.053	0.494	52.8	0.564

注:表中加粗字体表示此种指标下的最优值。

3.4 感知验证

为验证本文方法在复杂环境下的实际效能,将融合结果接入 YOLO11 检测模型。通过多算法对比,分析可靠性引导策略在工程应用中的增益效果与实用价值,实验结果如图 7 所示。

图 7 直观对比了各算法在检测任务中的表现,红色与黄色三角分别标注漏检与误检(彩图见期刊电子版)。在常规场景下,本文方法凭借更强的目标显著性显著优于对比模型。但在图 7(d)远距离雨天场景中,受极低信噪比影响,可靠性度量机制在目标热辐射梯度消失时难以精准判定信源,导致出现对极小目标的漏检。图 7(e)高密度遮挡工况则表明,当边缘结构遭严重

破坏时,即便深度学习模型也难以实现完全识别。这种失效分析界定了算法的泛化性边界,也为后续引入多尺度特征或时序信息优化指明了方向。高质量的融合图像有效滤除了因可见光过曝或红外热噪声引发的伪结构,从而显著降低了检测器对背景干扰的误检率;同时,可靠性自适应引导机制强化了极暗或眩光环境下目标的显著性,有效补全了单模态下难以识别的漏检目标。为进一步验证融合结果在下游感知任务中的实际增益,对不同融合方法的性能进行定量对比分析。选取 $P, R, mAP@0.5$ 及 $mAP@0.5:0.95$ 作为评价指标,系统考察各方法在多场景条件下对目标可检测性与定位准确性的影响,结果如图 8 所示。

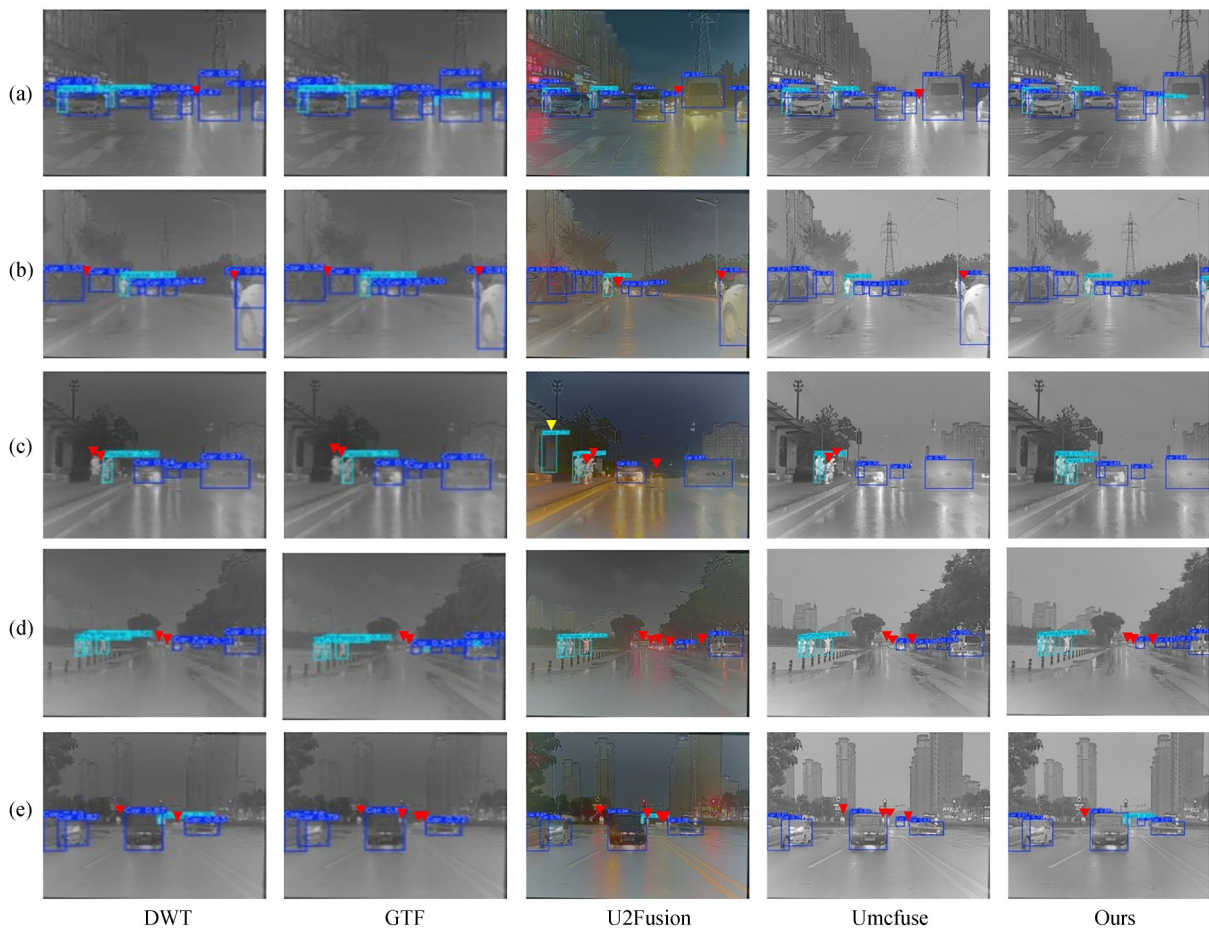


图 7 典型场景与极端工况下的感知验证结果对比

Fig. 7 Comparison of perceptual verification results between typical scenarios and extreme conditions

由图 4,图 6 及图 8 可知,融合图像质量指标的的提升与自动驾驶检测能力的增强呈现出显著

的正向相关性。本文方法在融合指标 $Q_{abf}, AG, SF, VIFF$ 分别平均提高 80.51%, 52.24%,

42.36% 和 17.6% 的基础上,直接带动下游检测任务在复杂工况下的 $P, R, mAP@0.5$ 及 $mAP@0.5:0.95$ 等指标上分别达到了 91.4%, 89.8%, 89.3% 和 82.6%,相较于其他主流融合方法分别

平均提升了 11.6%, 19.7%, 13.9% 和 15.2%。由此表明,该方法能够满足自动驾驶系统在复杂环境下的感知要求,能够为自动驾驶车辆提供更为可靠的检测结果。

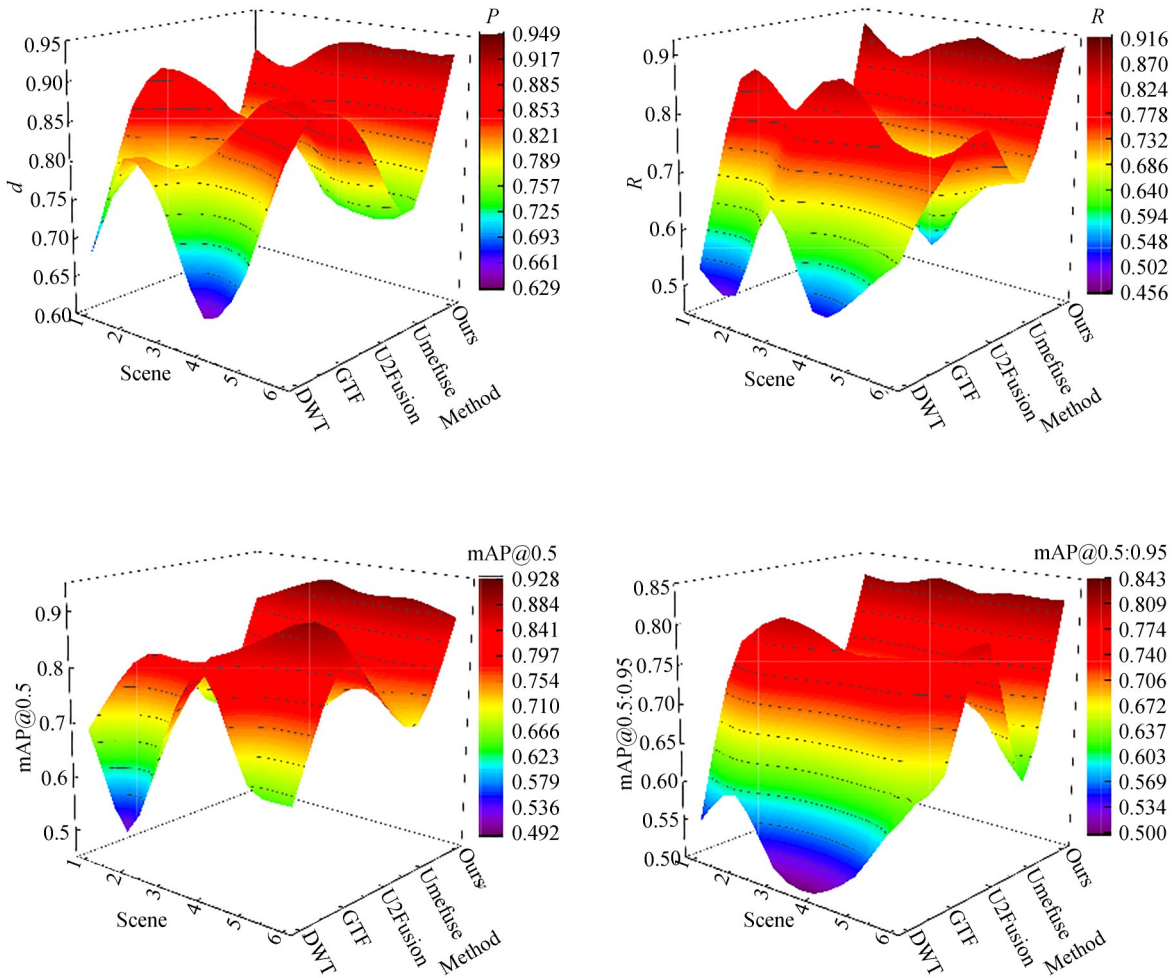


图 8 感知验证评价指标对比

Fig. 8 Comparison of perception verification evaluation indexes

4 结 论

针对复杂交通环境下红外与可见光图像融合中普遍存在的模态冲突及感知能力退化问题,本文提出了一种基于可靠性自适应引导的融合方法。该方法通过构建像素级可靠性度量机制,实现了对源图像信源质量的精准表征,并通过全局“可信注入”与细节层自适应引导策略,解决了

显著目标提取与背景纹理保留之间的权衡问题。实验结果表明,本文方法在客观量化评价中展现出显著优势。相较于 DWT, GTF, U2Fusion 及 Umefuse 4 种代表性算法,本文方法在保持信息丰度的同时,显著提升了图像的清晰度与对比度,其中标准差与空间频率分别平均提高了 16.56% 和 42.36%;在衡量纹理细节刻画能力与跨模态信息传递效率的平均梯度、互信息及融合

质量指标上,分别提升了 52.24%, 38.28% 和 80.51%;在边缘强度及视觉信息保真度方面分别提升 21.4% 和 17.6%。此外,消融实验进一步证实了可靠性引导机制在多模态特征协同优化中的关键作用。下游感知验证表明,本文方法在 YOLO11 检测任务中达到了 91.4% 的平均精确率,证明了所提算法在提升自动驾驶系统环境感知稳定性方面的工程应用价值。未来的研究重点可针对极端小目标及高比例遮挡导致的感

知时效问题,着重探索引入时域上下文信息或更深层次的语义先验,以进一步增强自动驾驶系统在复杂长尾环境下的稳定性。

作者贡献声明:

王琛:融合方法提出,论文构思与撰写;
马庆禄:实验设计,论文审核与编辑;
周志超:实验数据整理与分析;
刘明:实验数据分析。

参考文献:

- [1] SHA H, SINGH M K, HAOUARI R, *et al.* Network-wide safety impacts of dedicated lanes for connected and autonomous vehicles[J]. *Accident Analysis & Prevention*, 2024, 195: 107424.
- [2] 任鹏百, 雷慧云, 党建武, 等. 基于 HMSD 与改进 PCNN 的红外与可见光图像融合[J]. *光学精密工程*, 2025, 33(9): 1481-1495.
- REN P B, LEI H Y, DANG J W, *et al.* Infrared and visible image fusion based on HMSD and improved PCNN[J]. *Optics and Precision Engineering*, 2025, 33(9): 1481-1495. (in Chinese)
- [3] 韩阳, 杨华. 基于非下采样剪切波变换与显著信息加权的图像融合算法[J]. *科学技术与工程*, 2021, 21(17): 7224-7229.
- HAN Y, YANG H. An image fusion algorithm based on nonsubsampling shearlet transform coupled with significant information weighting [J]. *Science Technology and Engineering*, 2021, 21(17): 7224-7229. (in Chinese)
- [4] 孙磊, 唐倩, 廖一鸣, 等. 浮选泡沫低照度图像颜色深度编解码校正及多尺度增强[J]. *光学精密工程*, 2025, 33(10): 1609-1626.
- SUN L, TANG Q, LIAO Y P, *et al.* Multi-scale enhancement and color depth codec correction of flotation foam low illumination images[J]. *Optics and Precision Engineering*, 2025, 33(10): 1609-1626. (in Chinese)
- [5] 李英超, 赵喆浩, 王祺, 等. 面向地物混杂背景的偏振光谱图像融合方法[J]. *中国光学(中英文)*, 2024, 17(5): 1098-1111.
- LI Y CH, ZHAO ZH H, WANG Q, *et al.* Polarization spectral image fusion method for hybrid backgrounds of ground objects [J]. *Chinese Optics*, 2024, 17(5): 1098-1111. (in Chinese)
- [6] LYU F, LIU X Y, CHEN L, *et al.* Low-amplitude structure recognition method based on non-sampled contourlet transform [J]. *Petroleum Science*, 2024, 21(5): 3062-3078.
- [7] 张永兴, 连博文, 顾乃庭, 等. 基于多尺度空间注意力互补的红外与可见光图像融合[J]. *光学精密工程*, 2025, 33(7): 1152-1168.
- ZHANG Y X, LIAN B W, GU N T, *et al.* Infrared and visible image fusion based on multi-scale spatial attention complementary [J]. *Optics and Precision Engineering*, 2025, 33(7): 1152-1168. (in Chinese)
- [8] XIAN R H, XIONG X, PENG H, *et al.* Feature fusion method based on spiking neural convolutional network for edge detection [J]. *Pattern Recognition*, 2024, 147: 110112.
- [9] ZHAO Y Y, ZHENG Q C, ZHU P H, *et al.* TU-Fusion: a transformer-based universal fusion algorithm for multimodal images [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2024, 34(3): 1712-1725.
- [10] PENG S R, ZHU X Y, DENG H Y, *et al.* FusionMamba: efficient remote sensing image fusion with state space model [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5410216.
- [11] 邱敬, 王鹤然, 梁婵, 等. 条件扩散和多通道高低频并行的红外与可见光图像融合[J]. *光学精密工程*, 2025, 33(1): 148-163.
- DI J, WANG H R, LIANG CH, *et al.* Conditional diffusion and multi-channel high-low frequency parallel fusion of infrared and visible light images [J]. *Optics and Precision Engineering*, 2025, 33(1): 148-163. (in Chinese)
- [12] 杨艳春, 李佳龙, 李毅, 等. 夜间红外与可见光多

- 尺度信息注入式图像融合[J]. 光学精密工程, 2025, 33(2): 282-297.
- YANG Y CH, LI J L, LI Y, *et al.* Multiscale semantic injective fusion of nighttime infrared and visible[J]. *Optics and Precision Engineering*, 2025, 33(2): 282-297. (in Chinese)
- [13] SHI Y, LIU Y, CHENG J, *et al.* VDMUFusion: a versatile diffusion model-based unsupervised framework for image fusion[J]. *IEEE Transactions on Image Processing*, 2025, 34: 441-454.
- [14] TANG L F, XIANG X Y, ZHANG H, *et al.* DIFusion: Darkness-free infrared and visible image fusion [J]. *Information Fusion*, 2023, 91: 477-493.
- [15] 姚睿, 王凯, 郭浩帆, 等. 基于跨模态特征交互和多尺度重建的红外与可见光图像融合[J]. 红外与激光工程, 2025, 54(8): 20250210.
- YAO R, WANG K, GUO H F, *et al.* Infrared and visible image fusion based on cross-modal feature interaction and multi-scale reconstruction [J]. *Infrared and Laser Engineering*, 2025, 54(8): 20250210. (in Chinese)
- [16] 张晓东, 张殿伟, 李媛媛, 等. 掩码引导的二阶段红外与可见光图像融合网络[J]. 光学学报, 2025, 45(15): 1510007.
- ZHANG X D, ZHANG D W, LI Y Y, *et al.* Mask-guided two-stage infrared and visible image fusion network[J]. *Acta Optica Sinica*, 2025, 45(15): 1510007. (in Chinese)
- [17] 赵阳, 杨文贵, 高翠云. 基于全局双组注意力的红外与可见光图像融合[J]. 液晶与显示, 2025, 40(12): 1840-1852.
- ZHAO Y, YANG W G, GAO C Y. Infrared and visible image fusion based on global dual-group attention[J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(12): 1840-1852. (in Chinese)
- [18] 李冰, 叶猛, 顿非凡, 等. 基于自适应光照的输电线路自然暗光图像增强方法[J]. 光学学报, 2025, 45(23): 2310001.
- LI B, YE M, XIE ZH F, *et al.* Adaptive illumination-based natural low-light image enhancement method for transmission lines[J]. *Acta Optica Sinica*, 2025, 45(23): 2310001. (in Chinese)
- [19] MA J Y, CHEN C, LI C, *et al.* Infrared and visible image fusion *via* gradient transfer and total variation minimization[J]. *Information Fusion*, 2016, 31(C): 100-109.
- [20] XU H, MA J Y, JIANG J J, *et al.* U2Fusion: a unified unsupervised image fusion network [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(1): 502-518.
- [21] LI X L, LI X S, TAN T S, *et al.* UMCFuse: a unified multiple complex scenes infrared and visible image fusion framework [J]. *IEEE Transactions on Image Processing*, 2025, 34: 6231-6245.

作者简介:



王琛(2001—),男,河南濮阳人,硕士研究生,2023年于河南工业大学获得学士学位,主要从事机器视觉及智能交通系统与安全方面的研究。E-mail: chenw@mails.cqjtu.edu