Hybrid Reinforcement Learning for One-step Degraded Infrared and Visible Image Fusion

Hongwei Yu[1,2,*] **Daoqing Zha**[1,*] **Jiawei Li**[1,†] **Yancheng Bai**[2,†,‡]
Lei Sun[2] **Xiangxiang Chu**[2]

[1]University of Science and Technology Beijing
[2]DreamX Team, Alibaba Group
[*]Equal contribution [†]Corresponding authors [‡]Project lead

https://yuhongwei22.github.io/onehrl-site/

Abstract

Infrared and visible image fusion (IVIF) is essential for robust perception, yet existing methods remain limited under complex degradations. While diffusion models (DMs) offer powerful modeling capabilities to combat such degradations, their iterative process incurs high computational cost. One-step distillation effectively accelerates inference but typically leads to over-smoothed results. To resolve these limitations, we propose a one-step fusion framework with hybrid reinforcement learning, termed OneHRL. We treat degraded inputs as intermediate noisy states along a generative trajectory, enabling a distilled one-step generator to rectify them directly to the clean fusion results. Concurrently, to mitigate the over-smoothed results in distillation, we incorporate the Flow-GRPO paradigm to introduce stochasticity into the deterministic one-step generation. This restores the stochastic exploration necessary for sampling-based reinforcement learning. Furthermore, we design a coarse-to-fine reward that integrates a global semantic reward with an object-level reward. This hybrid reward ensures that the generator captures both naturalness and fine-grained texture. Extensive experiments demonstrate that our method achieves superior fusion quality and efficiency.

1 Introduction

Infrared and visible Image Fusion (IVIF) is a fundamental technique in multi-sensor perception, aiming to integrate the textural details of visible sensors with the thermal information of infrared ones [1; 2; 3]. While existing methods have demonstrated outstanding capabilities in ideal scenarios [4; 5; 6], real-world applications frequently suffer from various degradations. Most previous works are designed for clear inputs and fail to generate robust outputs when facing complex degradations. While some degradation-aware methods have emerged [7; 8; 9], they are often restricted to single-type impairments and remain ineffective against composite degradations with complex distributions.

Diffusion models (DMs) demonstrate powerful probabilistic modeling capabilities [10; 11]. Recently, these capabilities of DMs have been used in IVIF to address degradation challenges. Models such as DRMF [12] and MMAIF [13] have demonstrated remarkable robustness in degraded scenarios. However, as shown in Fig. 1, the iterative sampling process of DMs introduces high computational costs, rendering them impractical for real applications. To achieve acceleration, one-step distillation [14] has been explored in tasks such as super-resolution (SR). Nevertheless, its direct adaptation to the IVIF domain remains non-trivial. IVIF necessitates the simultaneous alignment and fusion of disparate modalities, which increases task complexity compared with SR.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

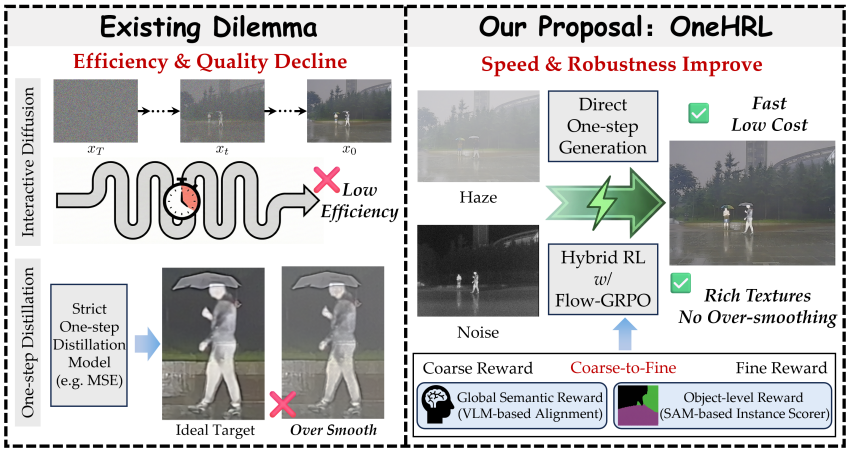

Figure 1: Schematic illustration of existing dilemma and our proposal. It is obvious that our proposed OneHRL can achieve both high inference speed and strong robustness.

Meanwhile, strict one-step distillation (e.g., MSE) typically induces over-smoothed results [15; 16]. To overcome this limitation, Sampling-based reinforcement learning (RL) [17; 18] is introduced to enhance textural richness, leveraging high-level perceptual or physical rewards.

Nevertheless, the deterministic nature of one-step generation poses a significant obstacle. Specifically, since one-step distillation replaces random noise with a fixed degraded representation, the model performs a deterministic mapping. This mapping lacks stochasticity, thereby precluding the exploration required for effective sampling-based RL optimization.

To address these challenges, we propose a novel one-step hybrid reinforcement learning fusion framework, termed OneHRL. The one-step and RL stages are respectively shown in Fig. 2 and 3. Leveraging the concepts of generalized diffusion [19], we treat the degraded VIS/IR fusion process as a map between the corrupted and clean data manifolds. Specifically, we process the degraded VIS/IR pairs through a feature fusion module pre-trained on clean data. We then conceptualize the resulting representation as an intermediate noisy state and perform one-step distillation from a multi-step diffusion model. It enables the generator to rectify this state directly toward the clean fusion manifold in a single step. This mechanism leverages the feature fusion module for cross-modal information aggregation while exploiting the powerful distribution modeling capabilities of diffusion models to restore fidelity from complex degradations. To mitigate over-smoothing, we incorporate the Flow-GRPO [20] paradigm, which reformulates the probability flow ordinary differential equation (ODE) into a stochastic differential equation (SDE) during the training phase. This transformation injects controlled stochasticity into the generative process, restoring the exploration space necessary for sampling-based RL. Finally, we design a coarse-to-fine reward to provide multi-level supervision. For the coarse reward, we leverage the large-scale vision-language model of Qwen2.5-VL [21] for the global semantic reward. A differentiable instance scorer is designed to derive object-level rewards by leveraging the Segment Anything Model (SAM). This hybrid reward enforces both global semantic coherence and object-level structural fidelity, ensuring robust fusion performance even under composite degradations. Our contributions are summarized as follows:

- We propose an efficient one-step IVIF framework, which conceptualizes degraded inputs as intermediate noisy states. It enables the model to achieve joint restoration and modality fusion in a single step.
- We incorporate the Flow-GRPO paradigm into the one-step pipeline. By reformulating the probability flow ODE into an SDE, we successfully bridge the gap between deterministic process and sampling-based RL. It restores the stochastic exploration necessary.
- We design a hybrid coarse-to-fine reward that integrates the semantic intelligence of VLMs with a differentiable instance scorer based on SAM. It ensures the generator captures both global naturalness and object-level prominence, even under complex composite degradations.
- Extensive experiments conducted across multiple datasets and degradation scenarios validate that our OneHRL can achieve robust fusion results and maintain high accuracy in downstream tasks.

2 Related Work and Preliminaries

2.1 Related Works

Infrared-Visible Image Fusion for Degradation Most existing IVIF methods [22; 23; 24; 25; 26; 27; 28; 29] are primarily optimized for clean inputs and struggle in real-world degradations. For the IVIF task under degraded conditions, a robust framework must simultaneously extract salient cross-modal features and mitigate the impacts of environmental artifacts on the fusion performance. The scarcity of degraded scenarios in current benchmarks [30; 31; 32] necessitate the use of simulation techniques to model degradations. [33] constructed an imaging model based on Retinex theory and the atmospheric scattering principle, designed a fine-tuning restoration and fusion network to deal with degraded source images. [34] introduced a language-driven framework for flexible and controllable degradation-robust image fusion. Moreover, [35] utilized intrinsic content extraction, intrinsic content fusion, and appearance representation learning and assignment to conduct a unified robust fusion framework.

Reinforcement Learning Reinforcement Learning has proven effective in aligning Large Language Models (LLMs) with human intent [17; 36]. While its initial success was prominent in aligning large language models with human preferences, recent advancements have extended RL to refine diffusion models for various vision tasks [37; 38; 39]. The remarkable performance of RL in NLP and CV demonstrates its effectiveness in optimizing model performance.

2.2 Preliminaries

Recent advanced image-generation models [21] adopt Rectified Flow framework [40], which connects two distributions via an Ordinary Differential Equation (ODE). Let $z_0 \sim \pi_0$ denote the prior distribution (e.g., standard Gaussian $\mathcal{N}(0, \mathbf{I})$) and $z_1 \sim \pi_1$ denote the data distribution. The dynamics are governed by the flow ODE:

$$dz_t = v(z_t, t)dt, \tag{1}$$

where z_t is the state at time $t \in [0, 1]$, and v is the velocity field. Ideally, the straight-line trajectory between a noise sample $z_0 \sim \pi_0$ and a data sample $z_1 \sim \pi_1$ is defined as:

$$z_t = tz_1 + (1-t)z_0, \quad \text{with } v(z_t, t) = z_1 - z_0. \tag{2}$$

In practice, we train a neural network v_θ to approximate this velocity field. The optimal parameters are obtained by minimizing the flow matching objective:

$$\theta^* = \arg\min_\theta \mathbb{E}\left[\|(z_1 - z_0) - v_\theta(tz_1 + (1-t)z_0, t)\|^2\right]. \tag{3}$$

3 Proposed Method

3.1 Motivation

Existing advanced fusion methods demonstrate outstanding capabilities leveraging the DMs. Despite their remarkable performance, the inherent nature of iterative sampling remains computationally expensive. In SR tasks, recent works apply one-step distillation to compress the generation process. Due to the distinct multimodal inputs of IVIF, directly applying such methods is challenging. A suitable pipeline to achieve one-step fusion is lacking. Moreover, strict distillation (e.g., MSE) often yields over-smoothed results. Sampling-based RL (e.g., GRPO) has demonstrated the ability to improve perceptual quality and structural sharpness. However, applying RL to one-step fusion models is non-trivial. Since the inputs of one-step model are fixed, the outputs lack stochasticity, thereby precluding the exploration required by sampling-based RL methods. To address these issues, we propose a one-step fusion framework with hybrid RL (OneHRL). It conceptualizes degraded inputs as intermediate noisy states for direct restoration. Furthermore, to enhance perceptual performance via RL, we propose the Hybrid RL strategy with Coarse-to-Fine Rewards. For coarse semantic rewards, we introduce Flow-GRPO, a method that changes the ODE to an SDE while strictly guaranteeing the equivalence of the marginal distributions. This transformation enables us to inject controlled stochasticity for sampling, which is then evaluated by Qwen2.5-VL to ensure overall naturalness. For fine-grained rewards, we leverage the SAM to extract object instances and design a Instance Scorer to enforce pixel-level details. This enables us to simultaneously guarantee efficiency and robustness.

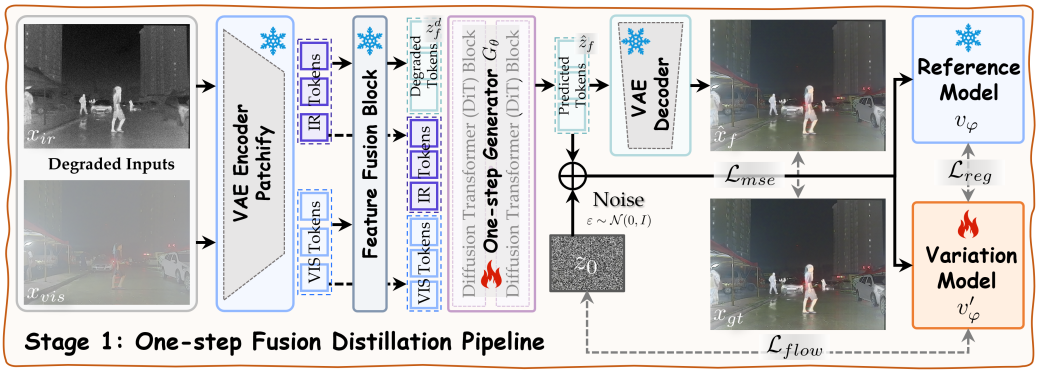

Figure 2: Framework of the proposed OneHRL. **One-step Distillation Stage:** Degraded inputs are mapped to the latent space and integrated into the degraded fused representation z_f^d via the Feature Fusion Block. The generator G_θ then restores z_f^d to a clean state in a single forward pass, optimized by $\mathcal{L}_{mse}$ and $\mathcal{L}_{reg}$. $\mathcal{L}_{flow}$ is utilized to update the variation model.

3.2 One-step Fusion via Distilled Diffusion Priors

To overcome the time cost while retaining the capability of diffusion models, we propose a one-step fusion framework based on rectified flow. Our goal is to learn a deterministic generator G_θ that directly maps degraded inputs to a clean, high-quality fusion image in a single forward pass.

Overview of Distillation Pipeline For notational brevity, we denote the input image space as I and the compressed latent space encoded by the VAE as $z = \mathcal{E}(I)$. We first train a feature fusion block (FFB), denoted as $\mathcal{F}_\phi$, using clean visible (I_{vis}^c) and infrared (I_{ir}^c) pairs. This module serves as a feature aggregator, optimized to maximize the retention of texture and radiation details:

$$z_f^c = \mathcal{F}_\phi(z_{vis}^c, z_{ir}^c), \tag{4}$$

where $z_{vis}^c, z_{ir}^c = \mathcal{E}(I_{vis}^c), \mathcal{E}(I_{ir}^c)$. Simultaneously, we train a diffusion transformer (DiT) conditioned on degraded latent inputs (z_{vis}^d, z_{ir}^d). This model functions as a restoration expert, learning a velocity field v_ψ that governs the probability flow ODE from noise to data:

$$dz_t = v_\psi(z_t, t, z_{vis}^d, z_{ir}^d)dt, \quad z_0 \sim \mathcal{N}(0, \mathbf{I}), \tag{5}$$

where $z_t = tz_1 + (1-t)z_0$ and z_1 represents the encoded clean target $\mathcal{E}(x_{gt})$. The pre-trained parameters of ψ are used to initialize the one-step generator G_θ, the frozen reference model v_ψ, and the learnable variational model $v_{\psi'}$.

In the distillation stage, we freeze the FFB and focus on distilling the restoration capability into a one-step generator. Given degraded inputs (z_{vis}^d, z_{ir}^d), the frozen FFB yields a degraded fused representation, denoted as z_f^d. A key innovation of our method is to conceptualize z_f^d not merely as a semantic condition, but as a specific intermediate noisy state along the generative trajectory. Instead of sampling from pure Gaussian noise, we initialize the reverse process with z_f^d. The one-step generator then projects this state to the clean target manifold, followed by the VAE decoder $\mathcal{D}$ to obtain the final fused image $\hat{x}_f$:

$$\begin{aligned} \hat{x}_f &= \mathcal{D}(\hat{z}_f) = \mathcal{D}(G_\theta(z_f^d, z_{vis}^d, z_{ir}^d)), \\ G_\theta(\cdot) &= z_f^d + \lambda_1 v_\theta(z_f^d, z_{vis}^d, z_{ir}^d), \end{aligned} \tag{6}$$

where v_θ represents the predicted velocity field, and λ_1 is a scaling factor (empirically set to 0.999) controlling the rectification step size. G_θ thus learns a residual mapping to rectify the degraded trajectory in a single forward pass.

The generator is optimized via a composite objective. Specifically, $\hat{x}_f$ is constrained by a pixel-level reconstruction loss $\mathcal{L}_{mse}$ against the ground truth x_{gt}. Simultaneously, $\hat{z}_f$ is perturbed by noise and processed by both v_ψ and $v_{\psi'}$ to derive the $\mathcal{L}_{reg}$. Furthermore, the variational model $v_{\psi'}$ is updated via a flow-matching objective $\mathcal{L}_{diff}$, minimizing the error between its output and the added noise to track the student distribution.

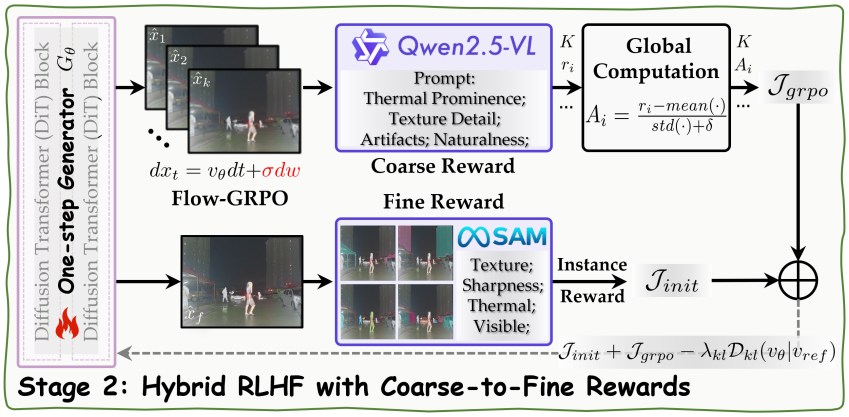

Figure 3: Framework of the proposed OneHRL. **Hybrid RL Stage:** The generator is refined via the Hybrid RL, leveraging a coarse-to-fine reward to enhance textural details.

Loss of One-step Distillation To establish a deterministic one-step generator G_θ, we employ a composite objective function. Our training strategy integrates a pixel-level reconstruction loss $\mathcal{L}_{mse}$ with a regularization term $\mathcal{L}_{reg}$ derived from diffusion priors.

We utilize the Mean Squared Error (MSE) between the generated output and the clean ground truth x_f as the fundamental constraint:

$$\mathcal{L}_{mse} = \|\mathcal{D}(G_\theta(z_f^d, z_{vis}^d, z_{ir}^d)) - x_{gt}\|_2^2. \tag{7}$$

This term ensures the low-frequency structural fidelity of the fusion results, anchoring the content to the ground truth. For regularization term, we employ Variational Score Distillation (VSD) as a regularizer to align the distribution of network outputs with the fusion results $\hat{x}_f$. We instantiate two versions of the pre-trained DiT model: a frozen reference model v_ψ and a learnable variational model $v_{\psi'}$. The frozen model v_ψ preserves the ground-truth velocity field learned from clean data. The variational model $v_{\psi'}$ acts as a dynamic teacher, tracking the student's current flow distribution,

$$\mathcal{L}_{reg} = \mathbb{E}_{t,z_0}\left[||w(t)\left(v_{\psi'}(z_t,t,c) - v_\psi(z_t,t,c)\right)||_2^2\right], \tag{8}$$

where $z_t = t\hat{z}_f + (1-t)z_0$, $\hat{z}_f = G_\theta(z_f^d, z_{vis}^d, z_{ir}^d)$, and $z_0 \sim \mathcal{N}(0, \mathbf{I})$. The total optimization objective for the generator G_θ is defined as:

$$\mathcal{L}_{total} = \mathcal{L}_{mse} + \lambda_{reg}\mathcal{L}_{reg}, \tag{9}$$

where λ_{reg} is a weighting scalar. Furthermore, to ensure $v_{\psi'}$ effectively tracks the student's distribution, it is simultaneously updated via a flow matching objective $\mathcal{L}_{diff}$:

$$\mathcal{L}_{diff} = \mathbb{E}_{t,z_0}\left[\|v_{\psi'}(z_t,t,c) - (\hat{z}_f - z_0)\|_2^2\right]. \tag{10}$$

3.3 Hybrid RLHF with Coarse-to-Fine Rewards

In the RL stage, we propose a Hybrid RLHF framework driven by a Coarse-to-Fine Reward. We integrate a Flow-GRPO optimizer driven by VLMs for semantic alignment, and a differentiable instance scorer based on SAM for object-level precision.

Coarse RL via Flow-GRPO In the one-step generation, the stochastic input z_t is replaced by the fixed input z_f^d. Consequently, the generator outputs a deterministic result $\hat{x}_f$. It precludes the stochastic exploration essential for Sampling-based Reinforcement Learning (RL). To address this, we introduce Flow-GRPO, which reformulates the ODE into an SDE. For brevity, we denote the predicted velocity as $v_\theta(\cdot) = v_\theta(z_f^d, z_{vis}^d, z_{ir}^d)$. The SDE process is reformulated as:

$$dz_t = \left[v_\theta(\cdot) + \frac{\sigma_t^2}{2t}\left(z_f^d + (1-t)v_\theta(\cdot)\right)\right]dt + \sigma_t dw, \tag{11}$$

where dw denotes the Wiener process increments and σ_t controls the level of stochasticity during generation. This transformation injects controlled noise for exploration while theoretically preserving the marginal distributions of the generated samples.

For each degraded input, we sample a group of K candidate outputs $\{\hat{z}_1, \ldots, \hat{z}_K\}$, creating a local distribution of potential fusion results. We utilize Qwen2.5-VL as a semantic scorer $\mathcal{R}_{sem}$. To ensure robust evaluation, we design a Chain-of-Thought (CoT) prompt that instructs the VLM to analyze **Thermal Prominence**, **Texture Fidelity**, and **Overall Naturalness** before assigning a scalar score r_i. The detailed prompt is shown in Appendix. We then compute the relative advantage A_i for each candidate based on the group statistics:

$$A_i = \frac{r_i - \text{mean}(\{r_1, \ldots, r_K\})}{\text{std}(\{r_1, \ldots, r_K\}) + \delta}, \tag{12}$$

where δ is a constant for numerical stability. The generator is updated to maximize the expected advantage. We employ the clipped surrogate objective to ensure stable policy updates:

$$\mathcal{J}_{GRPO} = \mathbb{E}\left[\frac{1}{k}\sum_{i}^{k}\left(\min\left(\rho_i A_i, \text{clip}(\rho_i, 1-\epsilon, 1+\epsilon)A_i\right)\right)\right], \tag{13}$$

where $\rho_i = \frac{p_\theta(\hat{z}_i | z_f^d, z_{vis}^d, z_{ir}^d)}{p_{\theta_{old}}(\hat{z}_i | z_f^d, z_{vis}^d, z_{ir}^d)}$ denotes the probability ratio between the current and old policies.

Fine RL via Instance Scorer While VLMs offer superior semantic guidance, they may overlook pixel-level distortions. To complement the coarse RL, we introduce an instance scorer that leverages the SAM for fine-grained RL.

For the fused image $\hat{x}_f$, we utilize a pre-trained SAM to extract a set of binary instance masks $M = \{m_1, \ldots, m_N\}$. Detailed extraction method can be found in Appendix. This allows us to transition the evaluation from the global image level to the object level, focusing on regions of interest (e.g., pedestrians, vehicles). For each extracted instance m_j, we construct a composite score based on five differentiable physical dimensions: **Thermal Prominence ($\mathcal{S}_{tp}$)** evaluates saliency via Weber contrast and SNR, **Local Sharpness ($\mathcal{S}_{ls}$)** penalizes blurring using Laplacian gradient energy, **Information Entropy ($\mathcal{S}_{ie}$)** measures information richness via Shannon entropy, **Infrared Visibility ($\mathcal{S}_{iv}$)** assesses ir preservation through intensity correlation, and **Structural Similarity ($\mathcal{S}_{ssim}$)** enforces consistency with source modalities. All operators are computed strictly within the instance masks. The comprehensive formulations of these scores are provided in Appendix. The instance-aware objective is formulated as the mean of these five metrics $\mathcal{K} := \{tp, ls, ie, iv, ssim\}$:

$$\mathcal{J}_{inst} = \frac{1}{N}\sum_{j=1}^{N}\sum_{k\in\mathcal{K}}\mathcal{S}_k(m_j), \tag{14}$$

where N denotes the total number of extracted instances.

Overall RL Optimization Objective The final optimization objective acts as a unified framework that balances semantic alignment, physical fidelity, and training stability. We formulate the total objective to be maximized as:

$$\mathcal{J}_{total} = \mathcal{J}_{GRPO} + \lambda_{inst}\mathcal{J}_{inst} - \lambda_{kl}\mathcal{D}_{KL}, \tag{15}$$

where λ_{inst} and λ_{kl} are weighting hyperparameters. Here, $\mathcal{D}_{KL}$ represents the Kullback-Leibler (KL) divergence penalty, which we approximate via the mean squared error between the velocity predictions of the current policy v_θ and the frozen reference model v_{ref}:

$$\mathcal{D}_{KL} = \mathbb{E}\left[\|v_\theta(z_t, c) - v_{ref}(z_t, c)\|_2^2\right]. \tag{16}$$

This term serves as a crucial regularizer, preventing the generator from deviating excessively from the learned clean data manifold.

4 Experiments

4.1 Experiment Setting

4.1.1 Degradation Scenario Construction

To evaluate model robustness in degradation scenarios, we follow the simulation strategy of [9] to synthesize challenging scenarios by applying diverse degradation conditions to clean source images. All samples are from the M^3FD and MSRS datasets. The comprehensive training details and the degradation setting are in Appendix.

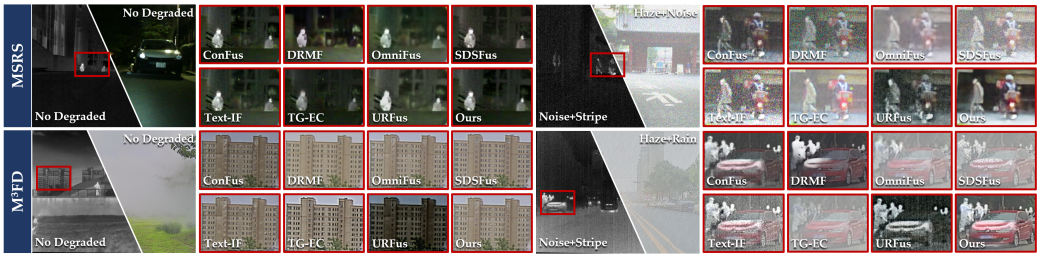

Figure 4: Fusion qualitative comparisons between SOTA methods and our OneHRL. Under different degradation conditions, the proposed OneHRL effectively mitigates adverse artifacts, which exhibit high visual quality.

Table 1: Fusion quantitative comparisons on the M^3FD dataset. Haze and rain degradations are added to VIS, while noise and stripe noise are incorporated into IR.

Method	Metric				
	$CC_\uparrow$	$SCD_\uparrow$	$MI_\uparrow$	$VIF_\uparrow$	$MSE_\downarrow$
ConFus	**0.492**	**1.337**	2.363	**0.487**	0.049
DRMF	0.453	1.304	1.091	0.225	0.048
OmniFus	0.418	0.779	2.335	0.386	0.112
SDSFus	0.311	0.846	0.730	0.216	0.048
Text-IF	0.465	0.833	2.156	0.443	0.118
TG-EC	0.432	1.059	**2.753**	0.454	**0.044**
URFus	0.416	0.810	2.497	0.420	0.117
Ours	**0.519**	**1.544**	**3.088**	**0.498**	**0.039**

Table 2: Fusion quantitative comparisons on the MSRS dataset. Haze and noise degradations are added to VIS, while noise and stripe noise are incorporated into IR.

Method	Metric				
	$CC_\uparrow$	$SCD_\uparrow$	$MI_\uparrow$	$VIF_\uparrow$	$MSE_\downarrow$
ConFus	0.562	**0.920**	1.569	**0.351**	0.126
DRMF	0.558	1.021	1.137	0.320	**0.049**
OmniFus	0.552	0.734	**1.577**	0.342	0.182
SDSFus	0.435	0.752	0.861	0.243	0.149
Text-IF	**0.568**	0.824	1.387	0.333	0.219
TG-EC	0.443	0.740	0.825	0.234	0.097
URFus	0.429	0.648	0.906	0.258	0.222
Ours	**0.611**	**1.527**	**2.384**	**0.465**	**0.033**

4.1.2 Comparison Methods and Metrics

We benchmark our framework against seven SOTA methods for degradation scenario modeling, *i.e.*, ControlFusion [33], TG-ECNet [41], OmniFuse [34], URFusion [35], SDS-Fusion [42], Text-IF [9] and DRMF [12]. Furthermore, the downstream utility of the fused results is evaluated through object detection and semantic segmentation tasks using YOLOv5 [43] and DeepLabV3+ [44], measured by mean average precision (mAP@0.5) and mean intersection over union (mIoU), respectively. For a more equitable comparison, we employ correlation coefficient (CC), the sum of the correlations of differences (SCD), mutual information (MI), visual information fidelity (VIF) and mean squared error (MSE) as metrics for quantitative assessment.

4.2 Comparative Results

4.2.1 Fusion Results Comparison

Fig. 4 illustrates the qualitative comparison of our OneHRL against other methods. In clean scenarios of the MSRS dataset, DRMF and URFusion exhibit over-smoothed and a loss of details. OmniFuse and TG-ECNet fail to preserve background textures. ControlFusion, SDSFusion and Text-IF introduce unnatural color casts. Our OneHRL maintains superior structural fidelity and color consistency. Under composite degradations, SDSFusion, DRMF and Text-IF can alleviate haze but leave the noise influence. Conversely, OmniFuse and TG-ECNet suppress noise but are ineffective against haze interference. ControlFusion and URFusion remain degraded in the IR modality. In contrast, our proposed OneHRL effectively eliminates diverse degradations while maintaining rich-detailed textures and structural fidelity. Similarly, on the M^3FD dataset, OneHRL achieves exceptional fusion performance under clean inputs and maintains robustness across diverse degraded scenarios. Tab. 1 summarizes the quantitative comparisons under the composite degradation of Haze+Rain (VIS) and Noise+Stripe (IR) on the M^3FD dataset. Similarly, Tab. 2 presents the quantitative results for the MSRS dataset under Haze+Noise (VIS) and Noise+Stripe (IR) conditions. It demonstrates that OneHRL consistently achieves top-tier performance across all metrics. These results validate OneHRL's robustness and yields exceptional fused images across diverse degradation scenarios. Additional evaluations, including results under clean conditions, are provided in Appendix.

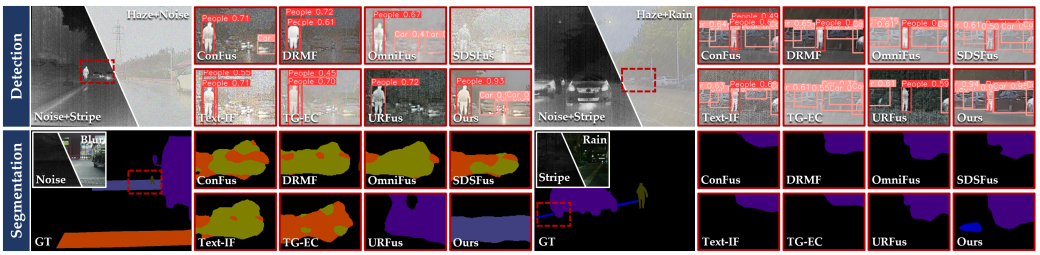

Figure 5: Downstream qualitative comparisons between SOTA methods and our OneHRL. Under different degradation conditions, the proposed OneHRL effectively mitigates adverse artifacts, which can exhibit both high visual quality and reliable downstream performance.

Table 3: Quantitative comparisons of downstream tasks. The subscripts represent the performance drop compared with the corresponding clean results. N-S, H-N, and H-R represent combined degradations involving Noise+Stripe, Haze+Noise, and Haze+Rain, respectively. The best and second-best results are highlighted in red and blue.

Degradation Type	Metric	Method							
		ConFus	DRMF	OmniFus	SDSFus	Text-IF	TG-EC	URFus	Ours
N-S + H-N	mAP@.5	**0.196**$_{\downarrow 0.459}$	0.165$_{\downarrow 0.413}$	0.160$_{\downarrow 0.440}$	0.081$_{\downarrow 0.405}$	0.156$_{\downarrow 0.503}$	0.134$_{\downarrow 0.535}$	0.172$_{\downarrow 0.474}$	**0.470**$_{\downarrow 0.159}$
N-S + H-R		**0.457**$_{\downarrow 0.198}$	0.367$_{\downarrow 0.211}$	0.249$_{\downarrow 0.351}$	0.175$_{\downarrow 0.311}$	0.285$_{\downarrow 0.374}$	0.231$_{\downarrow 0.438}$	0.153$_{\downarrow 0.493}$	**0.568**$_{\downarrow 0.061}$
Noise + Blur	mIoU	**0.527**$_{\downarrow 0.039}$	0.521$_{\downarrow 0.027}$	0.496$_{\downarrow 0.029}$	0.407$_{\downarrow 0.153}$	0.513$_{\downarrow 0.056}$	0.508$_{\downarrow 0.031}$	0.275$_{\downarrow 0.276}$	**0.601**$_{\downarrow 0.018}$
Stripe + Rain		0.441$_{\downarrow 0.125}$	0.385$_{\downarrow 0.164}$	**0.449**$_{\downarrow 0.076}$	0.374$_{\downarrow 0.185}$	0.390$_{\downarrow 0.180}$	0.352$_{\downarrow 0.187}$	0.401$_{\downarrow 0.150}$	**0.609**$_{\downarrow 0.010}$

4.2.2 Detection Results Comparison

To further evaluate the restoration capability of OneHRL, we display detection comparisons of different degradation in Fig. 5. Under the composite degradation of haze and noise, OneHRL successfully localizes all targets, including one person and three cars. In contrast, DRMF, Text-IF, and TG-ECNet fail to detect the vehicle targets and produce redundant boxes for the person. ControlFusion, URFusion, and OmniFuse correctly identify the pedestrian, but missing vehicle detections. SDSFusion fails to detect any targets. Similarly, in scenarios of haze and rain, ControlFusion, SDSFusion, and TG-ECNet produce false positives. DRMF, OmniFuse, and Text-IF locate error regions. URFusion further suffers from target missing. Tab. 3 summarizes the quantitative evaluation in terms of mAP@.5. Compared with other SOTA methods, OneHRL achieves a competitive mAP of 47.0% under the joint interference of Noise+Stripe (IR) and Haze+Noise (VIS), and 56.8% in the Noise+Stripe (IR) and Haze+Rain (VIS) scenario. These results demonstrate that our OneHRL yields high-quality results which are suitable for the detection task.

4.2.3 Segmentation Results Comparison

As illustrated in Fig. 5, we present two sets of qualitative segmentation comparisons to evaluate the semantic consistency of our fused results. In the left example, it contains a guardrail and a pedestrian. The URFusion misclassifies the guardrail as a vehicle under the degradation interference. Other methods produce segmentations consisting of bump and car stop categories, and exhibit region misalignments relative to the ground truth. The right one is collected from a dark environment. Other methods fail to fit the vehicle contours and completely neglect the road curve. Our OneHRL successfully segments the road curve and achieves more precise alignment with the vehicle boundaries. The quantitative mIoU results for each method are summarized in Tab. 3. Taken together, both qualitative and quantitative analyses confirm that OneHRL maintains high segmentation accuracy despite complex degradations.

4.3 Ablation Analysis

4.3.1 Analysis of Modules

To demonstrate the effectiveness of each proposed component, Fig. 6 presents an ablation study on the M^3FD and MSRS datasets. Here, $\mathcal{D}$ denotes the one-step distillation and $\mathcal{R}$ represents the RL

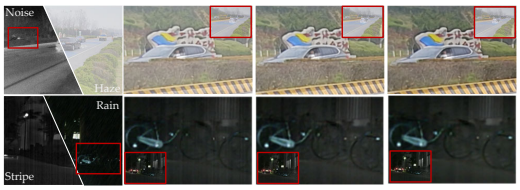

Figure 6: Ablation qualitative analysis of modules. Two scenes i.e., noise with haze and stripe with haze are selected in the M^3FD and MSRS dataset.

Table 4: Ablation quantitative analysis of modules.

Model		Module $\mathcal{D}$	Module $\mathcal{R}$	CC	SCD	MI	VIF	MSE
M^3FD	M1	✗	✗	0.431	1.395	2.876	0.511	0.128
	M2	✓	✗	0.462	1.477	3.009	0.534	0.086
	M3	✓	✓	**0.520**	**1.544**	**3.310**	**0.573**	**0.059**
MSRS	M1	✗	✗	0.573	1.445	2.316	0.555	0.046
	M2	✓	✗	0.591	1.537	2.764	0.632	0.051
	M3	✓	✓	**0.620**	**1.618**	**2.982**	**0.705**	**0.032**

training. We employ a baseline DiT model (M1) with 5-step iterative sampling. Subsequent models progressively incorporate these modules to form the final OneHRL. Under the composite degradation of noise and haze, M2 (M1+D) enhances the robustness against degradations. It also effectively improves the inference speed. By integrating the RL, OneHRL (M3) further refines textural details and visual clarity. The other examples also show the consistent performance gains. As shown in Tab. 4, the results confirm that OneHRL achieves the best performance across all fusion metrics.

4.3.2 Analysis of Rewards



Figure 7: Ablation qualitative analysis of rewards. We show their performance in the detection tasks.

Table 5: Ablation quantitative analysis of modules.

Method	AP@.5 People	Car	Bus	Motor	Truck	Lamp	mAP@.5
w/o F-G	0.516	0.763	0.429	0.117	0.492	0.244	0.426
Ours	**0.637**	**0.825**	**0.630**	**0.338**	**0.601**	**0.335**	**0.561**

In this section, we compare the performance of using only the instance reward (denoted as w/o F-G) against the full configuration on the detection task. We introduce the degradations of Noise (VIS) and Stripe Noise (IR). While employing only the instance reward can improve the contrast for some objects, the semantic style may change. It poses challenges for downstream tasks. As shown in Fig. 7, results of w/o Flow-GRPO miss several detections of target instances. Tab. 5 reports the per class AP@0.5 and mAP@0.5, which observe that the proposed OneHRL stands a leading position.

4.3.3 Analysis of KL Constraints

The KL constraint is a crucial regularization term in reinforcement learning that prevents the model from deviating excessively from the original one. We evaluate its impact through an ablation study on the semantic segmentation task. Fig. 8 illustrates the qualitative results under the interference of Blur (VIS) and Noise (IR). Without the KL constraint, the model exhibits instability, which leads to inaccurate segmentation masks. Specifically, distorted shapes appear in the vehicle regions, and the majority of the road curvature is lost. In contrast, our full method provides precise segmentation results and successfully preserves all semantic elements. In addition, IoU for each class and mIoU is shown in Tab. 6. OneHRL achieves superior performance across all categories.

Table 6: Ablation quantitative analysis of the KL constraint. The evaluation is conducted on the segmentation task.

Method	IoU Bac	Car	Per	Bik	Cur	C S	Gua	C C	Bum	mIoU
w/o KL	0.973	0.653	0.501	0.449	0.156	0.498	0.423	0.379	0.413	0.493
Ours	**0.974**	**0.824**	**0.597**	**0.569**	**0.249**	**0.591**	**0.687**	**0.436**	**0.510**	**0.601**



Figure 8: Ablation qualitative analysis of the KL constraint. We show their performance in the segmentation task.

4.3.4 Analysis of Generation Speed

In this section, we evaluate the inference efficiency of our proposed method compared to the diffusion-based baseline, DRMF. While DRMF requires 25 seconds to process a single image due to its iterative sampling process, OneHRL completes the generation in just 0.4 seconds. This represents a $62.5\times$ speedup, significantly enhancing its practical utility for the fusion task.

5 Conclusion

In this paper, we propose OneHRL, a one-step robust infrared and visible image fusion framework driven by hybrid reinforcement learning. By integrating one-step distillation with a coarse-to-fine reward, OneHRL effectively addresses the speed-robustness dilemma inherent in traditional diffusion-based models. Through the one-step distillation, our method achieves a significant speed improvement compared to existing diffusion-based approaches. Furthermore, our approach utilizes Flow-GRPO to ensure global semantic consistency and an instance scorer to enhance local textural details. Experimental results on the M^3FD and MSRS datasets demonstrate that OneHRL achieves both speed and robustness. Under diverse degradations, our method exhibits superior anti-degradation performance and yields consistent effectiveness across various downstream tasks.

6 Funding Transparency Statement

This work was conducted during Hongwei Yu's internship at Alibaba Group. Alibaba Group provided financial support and computational resources for this work. The authors declare no other competing interests.

References

[1] Jiawei Li, Jiansheng Chen, Jinyuan Liu, Hongwei Yu, Xinlong Ding, and Huimin Ma. Multi-scale spatial channel joint representation for general multi-modality image fusion with self-supervision. *IEEE Transactions on Multimedia*, pages 1–14, 2026.

[2] Jinyuan Liu, Guanyao Wu, Zhu Liu, Di Wang, Zhiying Jiang, Long Ma, Wei Zhong, Xin Fan, and Risheng Liu. Infrared and visible image fusion: From data compatibility to task adaption. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(4):2349–2369, 2024.

[3] Jinyuan Liu, Xingyuan Li, Zirui Wang, Zhiying Jiang, Wei Zhong, Wei Fan, and Bin Xu. Promptfusion: Harmonized semantic prompt learning for infrared and visible image fusion. *IEEE/CAA Journal of Automatica Sinica*, 2024.

[4] Zhiying Jiang, Zengxi Zhang, Jinyuan Liu, Xin Fan, and Risheng Liu. Multispectral image stitching via global-aware quadrature pyramid regression. *IEEE Transactions on Image Processing*, 2024.

[5] Zhiying Jiang, Zengxi Zhang, and Jinyuan Liu. Harmonized domain enabled alternate search for infrared and visible image alignment. *IEEE Transactions on Image Processing*, 2025.

[6] Jiawei Li, Jinyuan Liu, Shihua Zhou, Qiang Zhang, and Nikola K Kasabov. Gesenet: A general semantic-guided network with couple mask ensemble for medical image fusion. *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

[7] Zixiang Zhao, Lilun Deng, Haowen Bai, Yukun Cui, Zhipeng Zhang, Yulun Zhang, Haotong Qin, Dongdong Chen, Jiangshe Zhang, Peng Wang, et al. Image fusion via vision-language model. In *International Conference on Machine Learning*, pages 60749–60765. PMLR, 2024.

[8] Haowen Bai, Zixiang Zhao, Jiangshe Zhang, Yichen Wu, Lilun Deng, Yukun Cui, Baisong Jiang, and Shuang Xu. Refusion: Learning image fusion from reconstruction with learnable loss via meta-learning. *International Journal of Computer Vision*, 133(5):2547–2567, 2025.

[9] Xunpeng Yi, Han Xu, Hao Zhang, Linfeng Tang, and Jiayi Ma. Text-if: Leveraging semantic text guidance for degradation-aware and interactive image fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 27026–27035, 2024.

[10] Huaqiu Li, Yong Wang, Tongwen Huang, Hailang Huang, Haoqian Wang, and Xiangxiang Chu. Ld-rps: Zero-shot unified image restoration via latent diffusion recurrent posterior sampling. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 13684–13694. IEEE, 2025.

[11] Rui Lan, Yancheng Bai, Xu Duan, Mingxing Li, Dongyang Jin, Ryan Xu, Dong Nie, Lei Sun, and Xiangxiang Chu. Flux-text: A simple and advanced diffusion transformer baseline for scene text editing. *arXiv preprint arXiv:2505.03329*, 2025.

[12] Linfeng Tang, Yuxin Deng, Xunpeng Yi, Qinglong Yan, Yixuan Yuan, and Jiayi Ma. Drmf: Degradation-robust multi-modal image fusion via composable diffusion prior. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pages 8546–8555, 2024.

[13] Zihan Cao, Yu Zhong, Ziqi Wang, and Liang-Jian Deng. Mmaif: Multi-task and multi-degradation all-in-one for image fusion with language guidance. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 11744–11754, October 2025.

[14] Rongyuan Wu, Lingchen Sun, Zhiyuan Ma, and Lei Zhang. One-step effective diffusion network for real-world image super-resolution. *Advances in Neural Information Processing Systems*, 37:92529–92553, 2024.

[15] Xiangyu He and Jian Cheng. Revisiting l1 loss in super-resolution: a probabilistic view and beyond. *arXiv preprint arXiv:2201.10084*, 2022.

[16] Linwei Dong, Qingnan Fan, Yihong Guo, Zhonghao Wang, Qi Zhang, Jinwei Chen, Yawei Luo, and Changqing Zou. Tsd-sr: One-step diffusion with target score distillation for real-world image super-resolution. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 23174–23184, 2025.

[17] Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, YK Li, Yang Wu, et al. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*, 2024.

[18] Hongling Xu, Qi Zhu, Heyuan Deng, Jinpeng Li, Lu Hou, Yasheng Wang, Lifeng Shang, Ruifeng Xu, and Fei Mi. Kdrl: Post-training reasoning llms via unified knowledge distillation and reinforcement learning. *arXiv preprint arXiv:2506.02208*, 2025.

[19] Arpit Bansal, Eitan Borgnia, Hong-Min Chu, Jie Li, Hamid Kazemi, Furong Huang, Micah Goldblum, Jonas Geiping, and Tom Goldstein. Cold diffusion: Inverting arbitrary image transforms without noise. *Advances in Neural Information Processing Systems*, 36:41259–41282, 2023.

[20] Jie Liu, Gongye Liu, Jiajun Liang, Yangguang Li, Jiaheng Liu, Xintao Wang, Pengfei Wan, Di Zhang, and Wanli Ouyang. Flow-grpo: Training flow matching models via online rl. *arXiv preprint arXiv:2505.05470*, 2025.

[21] Shuai Bai, Keqin Chen, Xuejing Liu, Jialin Wang, Wenbin Ge, Sibo Song, Kai Dang, Peng Wang, Shijie Wang, Jun Tang, et al. Qwen2. 5-vl technical report. *arXiv preprint arXiv:2502.13923*, 2025.

[22] Jiawei Li, Jiansheng Chen, Jinyuan Liu, and Huimin Ma. Learning a graph neural network with cross modality interaction for image fusion. In *Proceedings of the 31st ACM international conference on multimedia*, pages 4471–4479, 2023.

[23] Jiawei Li, Hongwei Yu, Jiansheng Chen, Xinlong Ding, Jinlong Wang, Jinyuan Liu, Bochao Zou, and Huimin Ma. A^2rnet: Adversarial attack resilient network for robust infrared and visible image fusion. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 4770–4778, 2025.

[24] Long Ma, Dian Jin, Nan An, Jinyuan Liu, Xin Fan, Zhongxuan Luo, and Risheng Liu. Bilevel fast scene adaptation for low-light image enhancement. *International Journal of Computer Vision*, 133(11): 8234–8252, 2025.

[25] Xingyuan Li, Zirui Wang, Yang Zou, Zhixin Chen, Jun Ma, Zhiying Jiang, Long Ma, and Jinyuan Liu. Difiisr: A diffusion model with gradient guidance for infrared image super-resolution. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 7534–7544, 2025.

[26] Pengwei Liang, Junjun Jiang, Qing Ma, Chenyang Wang, Xianming Liu, and Jiayi Ma. Fusioninv: A diffusion-based approach for multimodal image fusion. *IEEE Transactions on Image Processing*, 2025.

[27] Wu Wang, Liang-Jian Deng, Qi Cao, and Gemine Vivone. A general image fusion approach exploiting gradient transfer learning and fusion rule unfolding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2026.

[28] Zengyi Yang, Yu Liu, Juan Cheng, Zhiqin Zhu, Yafei Zhang, and Huafeng Li. Customized fusion: A closed-loop dynamic network for adaptive multi-task-aware infrared-visible image fusion. *arXiv preprint arXiv:2604.08924*, 2026.

[29] Xilai Li, Xiaosong Li, Tianshu Tan, Huafeng Li, and Tao Ye. Umcfuse: A unified multiple complex scenes infrared and visible image fusion framework. *IEEE Transactions on Image Processing*, 2025.

[30] Xinyu Jia, Chuang Zhu, Minzhen Li, Wenqi Tang, and Wenli Zhou. Llvip: A visible-infrared paired dataset for low-light vision. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 3496–3504, 2021.

[31] Linfeng Tang, Jiteng Yuan, and Jiayi Ma. Image fusion in the loop of high-level vision tasks: A semantic-aware real-time infrared and visible image fusion network. *Information Fusion*, 82:28–42, 2022.

[32] Jinyuan Liu, Xin Fan, Zhanbo Huang, Guanyao Wu, Risheng Liu, Wei Zhong, and Zhongxuan Luo. Target-aware dual adversarial learning and a multi-scenario multi-modality benchmark to fuse infrared and visible for object detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5802–5811, 2022.

[33] Linfeng Tang, Yeda Wang, Zhanchuan Cai, Junjun Jiang, and Jiayi Ma. Controlfusion: A controllable image fusion network with language-vision degradation prompts. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

[34] Hao Zhang, Lei Cao, Xuhui Zuo, Zhenfeng Shao, and Jiayi Ma. Omnifuse: Composite degradation-robust image fusion with language-driven semantics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025.

[35] Han Xu, Xunpeng Yi, Chen Lu, Guangcan Liu, and Jiayi Ma. Urfusion: Unsupervised unified degradation-robust image fusion network. *IEEE Transactions on Image Processing*, 2025.

[36] Chubin Chen, Sujie Hu, Jiashu Zhu, Meiqi Wu, Jintao Chen, Yanxun Li, Nisha Huang, Chengyu Fang, Jiahong Wu, Xiangxiang Chu, et al. Taming preference mode collapse via directional decoupling alignment in diffusion reinforcement learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12775–12786, 2026.

[37] Kevin Black, Michael Janner, Yilun Du, Ilya Kostrikov, and Sergey Levine. Training diffusion models with reinforcement learning. *arXiv preprint arXiv:2305.13301*, 2023.

[38] Ying Fan, Olivia Watkins, Yuqing Du, Hao Liu, Moonkyung Ryu, Craig Boutilier, Pieter Abbeel, Mohammad Ghavamzadeh, Kangwook Lee, and Kimin Lee. Dpok: Reinforcement learning for fine-tuning text-to-image diffusion models. *Advances in Neural Information Processing Systems*, 36:79858–79885, 2023.

[39] Bryan Sangwoo Kim, Jeongsol Kim, and Jong Chul Ye. Chain-of-zoom: Extreme super-resolution via scale autoregression and preference alignment. *arXiv preprint arXiv:2505.18600*, 2025.

[40] Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. *arXiv preprint arXiv:2209.03003*, 2022.

[41] Yiming Sun, Xin Li, Pengfei Zhu, Qinghua Hu, Dongwei Ren, Huiying Xu, and Xinzhong Zhu. Task-gated multi-expert collaboration network for degraded multi-modal image fusion. In *International Conference on Machine Learning*, pages 57571–57586. PMLR, 2025.

[42] Jun Chen, Liling Yang, Wei Yu, Wenping Gong, Zhanchuan Cai, and Jiayi Ma. Sdsfusion: A semantic-aware infrared and visible image fusion network for degraded scenes. *IEEE Transactions on Image Processing*, 2025.

[43] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.

[44] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)*, pages 801–818, 2018.

[45] Chenfei Wu, Jiahao Li, Jingren Zhou, Junyang Lin, Kaiyuan Gao, Kun Yan, Sheng-ming Yin, Shuai Bai, Xiao Xu, Yilei Chen, et al. Qwen-image technical report. *arXiv preprint arXiv:2508.02324*, 2025.

[46] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.

[47] Mingxing Li, Rui Wang, Lei Sun, Yancheng Bai, and Xiangxiang Chu. Next token is enough: Realistic image quality and aesthetic scoring with multimodal large language model. *arXiv preprint arXiv:2503.06141*, 2025.

Technical appendices and supplementary material

A The Details of Our Algorithms

Algorithm 1 One-step Distillation for Image Fusion

Require: Degraded pairs (I^d_{vis}, I^d_{ir}), ground truth x_{gt}, VAE $\{\mathcal{E}, \mathcal{D}\}$, frozen FFB $\mathcal{F}_\phi$.
Require: Initialize $G_\theta, v_\psi, v_{\psi'}$ with pre-trained ψ.

while not converged **do**
  // Step 1: Forward Pass
  $z^d_f = \mathcal{F}_\phi(\mathcal{E}(I^d_{vis}), \mathcal{E}(I^d_{ir}))$; ▷ Encode and fuse degraded inputs
  $\hat{z}_f = G_\theta(z^d_f, \cdot)$; $\hat{x}_f = \mathcal{D}(\hat{z}_f)$; ▷ One-step prediction and reconstruction
  // Step 2: Optimization Objectives
  Sample $t \sim \mathcal{U}(0,1), z_0 \sim \mathcal{N}(0, \mathbf{I})$; Compute $z_t = t\hat{z}_f + (1-t)z_0$;
  $\mathcal{L}_{reg} = ||w(t)(v_{\psi'} - v_\psi)||^2_2$; $\mathcal{L}_{mse} = ||\hat{x}_f - x_{gt}||^2_2$;
  $\mathcal{L}_{diff} = ||v_{\psi'}(z_t, t, c) - (\hat{z}_f - z_0)||^2_2$;
  // Step 3: Parameter Update
  $\theta \leftarrow \nabla_\theta(\mathcal{L}_{mse} + \lambda_{reg}\mathcal{L}_{reg})$; ▷ Update generator
  $\psi' \leftarrow \nabla_{\psi'}\mathcal{L}_{diff}$; ▷ Update variation model for tracking
end while
return Optimized generator G_θ

Algorithm 2 Hybrid RLHF Training for Robust Fusion

Require: Generator G_θ, reference v_{ref}, semantic scorer $\mathcal{R}_{sem}$, and instance extractor (SAM).

while not converged **do**
  // Step 1: Flow-GRPO Group Sampling & Advantage
  for $i = 1$ to K **do**
    Sample $\hat{z}_i$ via SDE; Decode $\hat{x}_i = \mathcal{D}(\hat{z}_i)$;
    $r_i = \mathcal{R}_{sem}(\hat{x}_i)$; ▷ VLM semantic scoring
  end for
  $A_i = \frac{r_i - \text{mean}(\{r_k\})}{\text{std}(\{r_k\}) + \delta}$; ▷ Compute group-wise advantage
  // Step 2: Instance Reward
  $\hat{x}_f = \mathcal{D}(G_\theta(z^d_f, \cdot))$; ▷ Primary fused result
  $\mathcal{J}_{inst} = \text{SAM-Reward}(\hat{x}_f)$; ▷ Object-level fidelity
  // Step 3: Combined Optimization
  $\mathcal{J}_{total} = \mathcal{J}_{GRPO}(A_i) + \lambda_{inst}\mathcal{J}_{inst} - \lambda_{kl}\mathcal{D}_{KL}(v_\theta, v_{ref})$;
  Update θ by maximizing $\mathcal{J}_{total}$;
end while
return Refined generator G_θ

Algorithms 1 and 2 summarize the overall training pipeline of our framework. We first perform one-step distillation in the latent space, where the degraded visible and infrared inputs are encoded by the VAE and fused by the frozen feature fusion block. The generator G_θ then predicts the fused latent in a single forward pass, which is decoded to obtain the reconstructed fusion image. During this stage, the generator is optimized with the reconstruction loss to preserve pixel-level fidelity, while a regularized variation model is updated to track the target diffusion velocity. This distillation stage transfers the generation capability of the diffusion prior into an efficient one-step fusion generator.

After distillation, we further refine the generator through hybrid RLHF training. Specifically, Flow-GRPO samples a group of candidate fused results through the stochastic flow process and computes normalized group-wise advantages according to the semantic rewards produced by the VLM-based scorer. In parallel, an instance-level reward derived from SAM is introduced to encourage object-level structural fidelity. The final optimization objective combines the GRPO reward, the instance reward, and a KL regularization term with respect to the reference model, enabling the generator to improve semantic and instance-level robustness while maintaining stable generation behavior.

B Details of Experiment Setting

B.1 Degradation Scenario Construction

To evaluate model robustness in degradation scenarios, we follow the simulation strategy of [9] to synthesize challenging scenarios by applying diverse degradation conditions to clean source images. Specifically, the degradation types include blur, haze, rain, noise, and stripe. To ensure physical plausibility and prevent excessive information loss, each pristine image is subject to at most two concurrent degradations. This strategy of stochastic composite degradations encourages the model to learn generalized restoration features rather than over-fitting to specific cases. All samples from the M^3FD and MSRS datasets are processed following this protocol.

B.2 Training Details

The generator utilizes a DiT backbone consisting of 8 layers, where each layer features 8 attention heads with a head dimension of 128. The input and output channels are configured as 64 and 16, respectively. We employ a pre-trained VAE encoder from Qwen-Image-Edit [45] for feature extraction. During the training phase, the AdamW optimizer [46] is used to update the parameters of the student network, with the weight decay and learning rate set to $1e^{-2}$ and $1e^{-4}$, respectively. Clean and degraded source images from the M^3FD and MSRS datasets are utilized for training, which adopts a random pairing strategy for enhanced generalization. In the distillation stage, we set the total number of steps to 50k. The hyperparameters λ_{reg} is set to $1e^{-2}$. In the RL stage, we train the model for 1k steps. The hyperparameters λ_{inst} and λ_{kl} are set to $1e^{-2}$ and 0.8. All experiments are performed on a PyTorch framework with 8 NVIDIA H20 Tensor Core GPUs and an Intel(R) Xeon(R) Platinum 8480+ CPU. Notably, the model maintains high efficiency, requiring only 5GB of memory during the inference stage.

C Technical Details of Instance Scorer

C.1 Instance Extraction via SAM

To facilitate object-level evaluation in the reinforcement learning stage, we employ the Segment Anything Model (SAM) with a ViT-H backbone to partition the fused image $\hat{x}_f$ into instances. The extraction process is implemented via the SamAutomaticMaskGenerator, which identifies regions of interest without requiring specific category labels.

We configure the mask generator to prioritize both coverage and mask quality. The key parameters are set as follows:

- Grid Sampling: We utilize a 32×32 point-grid (1024 points in total) sampled across the image to ensure that small yet salient objects (e.g., distant pedestrians) are not overlooked.
- Confidence Filtering: To ensure the instance-level rewards are derived from high-quality masks, we set the predicted IoU threshold to 0.86 and the stability score threshold to 0.92. Masks failing to meet these criteria are discarded to avoid providing noisy feedback to the generator G_θ.
- Area Constraint: We exclude fragmented regions with a total area of fewer than 100 pixels. This prevents the fine-grained RL stage from being distracted by sensor-level noise or trivial artifacts that do not represent physical instances.

For the instance-level evaluation, the Segment Anything Model (SAM) is employed to partition the primary fused output $\hat{x}_f$ into a collection of binary masks $M = \{m_1, \ldots, m_N\}$. These masks identify semantically significant regions of interest, such as pedestrians or vehicles, allowing the reward system to transition from global image assessment to fine-grained object-level precision. Specifically, each extracted mask m_j serves as a spatial operator to isolate instance-specific features for the computation of differentiable physical rewards, including thermal prominence ($\mathcal{S}_{tp}$) and local sharpness ($\mathcal{S}_{ls}$).

C.2 Instance Rewards

To provide fine-grained feedback for the reinforcement learning stage, we define a composite instance-level reward $\mathcal{J}_{inst}$ based on five differentiable physical dimensions. Each metric is computed strictly within the spatial support of the binary mask m_j extracted by SAM. For each extracted mask m_j, we define the localized instance representations as $\hat{x}_f^j = \hat{x}_f \odot m_j$, $x_{ir}^j = x_{ir} \odot m_j$, and $x_{vis}^j = x_{vis} \odot m_j$. The reward $\mathcal{J}_{inst}$ is then computed as a weighted sum of five differentiable metrics, each operating on these masked regions.

1. Thermal Prominence ($\mathcal{S}_{tp}$) This reward ensures infrared targets remain salient relative to their local surroundings. We define an inner target region $\mathcal{M}_{in}$ via morphological erosion and a background region $\mathcal{M}_{bg}$ via dilation of the mask m_j. The score combines three sub-metrics:

- **Weber Contrast**: $C_w = \frac{\mu(\hat{x}_f^j|\mathcal{M}_{in}) - \mu(\hat{x}_f^j|\mathcal{M}_{bg})}{\mu(\hat{x}_f^j|\mathcal{M}_{bg}) + \epsilon}$, measuring the target-to-background distinctness.
- **Intensity Fidelity**: $P_{ir} = \exp(-\frac{|\mu(\hat{x}_f^j|m_j) - \mu(x_{ir}^j|m_j)|}{\mu(x_{ir}^j|m_j) + \epsilon})$, enforcing absolute IR intensity preservation.
- **Signal-to-Noise Ratio**: $SNR = \frac{\mu(\hat{x}_f^j|\mathcal{M}_{in})}{\sigma(\hat{x}_f^j|\mathcal{M}_{bg}) + \epsilon}$, where σ denotes the standard deviation.

The total prominence score is $\mathcal{S}_{tp} = 0.4 \cdot \sigma(2.0 \cdot C_w) + 0.3 \cdot P_{ir} + 0.3 \cdot \sigma(SNR - 3.0)$.

2. Local Sharpness ($\mathcal{S}_{ls}$) We compute the gradient energy to penalize over-smoothed. Applying a Laplacian operator $\mathbf{L}$ to the fused instance $\hat{x}_f^j$, we calculate the masked Tenengrad-variant sharpness:

$$\mathcal{S}_{ls} = \text{Sigmoid}\left(\log\left(\frac{\sum_{p \in m_j} (\mathbf{L} * \hat{x}_f^j)^2}{\sum m_j + \epsilon} + 1\right) - 3\right). \tag{17}$$

3. Information Entropy ($\mathcal{S}_{ie}$) This metric encourages high information density within the instance. Masked pixels are quantized into $L = 256$ bins to compute Shannon entropy:

$$H = -\sum_{i=0}^{L-1} p_i \log_2(p_i + \epsilon), \quad \mathcal{S}_{ie} = \frac{H}{8.0}, \tag{18}$$

where p_i is the normalized intensity frequency within the mask m_j.

4. Infrared Visibility ($\mathcal{S}_{iv}$) To prevent the loss of IR-exclusive features (e.g., heat signatures), we compute the unique signal preservation ratio:

$$\mathcal{R}_{ir} = \frac{\sum_{p \in m_j} \text{ReLU}(\hat{x}_f^j - x_{vis}^j)}{\sum_{p \in m_j} \text{ReLU}(x_{ir}^j - x_{vis}^j) + \epsilon}. \tag{19}$$

This is averaged with the masked Pearson correlation between $\hat{x}_f^j$ and x_{ir}^j and an intensity consistency score to evaluate structural and thermal alignment.

5. Structural Similarity ($\mathcal{S}_{ssim}$) We enforce consistency with both modalities using a masked SSIM index:

$$\mathcal{S}_{ssim} = \frac{1}{2}\left[\text{SSIM}(\hat{x}_f^j, x_{ir}^j|m_j) + \text{SSIM}(\hat{x}_f^j, x_{vis}^j|m_j)\right], \tag{20}$$

where all internal statistics are calculated strictly over the support of m_j.

Reward Aggregation Individual metrics are aggregated into an instance score via a Softmax-weighted scheme $\mathbf{w}_{norm}^\top \mathbf{s}_j$. For an image with N instances, the final batch reward is an area-weighted average to prioritize larger, semantically significant objects:

$$\mathcal{J}_{inst} = \sum_{j=1}^{N} \frac{\text{Area}(m_j)}{\sum_{k=1}^{N} \text{Area}(m_k)} \cdot (\mathbf{w}_{norm}^\top \mathbf{s}_j). \tag{21}$$

D VLM Reward Prompt Design

To provide structured semantic feedback for Flow-GRPO, our prompt consists of a quality rubric and a structured assessment instruction. It guides the VLM to assess thermal target prominence, texture preservation, artifacts, and overall naturalness before assigning a numerical score. Related work investigates fine-grained MLLM quality scoring with CoT [47] providing context for our hybrid reward design.

System Prompt for IVIF Evaluation

```
Your role is to evaluate the perceptual quality and information fusion performance of
Infrared and Visible Image Fusion (IVIF) results.
1. Bad: The fusion fails significantly. Severe loss of information from either source
(e.g., thermal targets are invisible or background texture is completely blurred).
Significant artifacts (ghosting, severe noise, or distortion) make the image unusable.
2. Poor: Unbalanced fusion. The image is dominated by one modality, causing a loss
of complementary details. Thermal targets are dim, or visible textures are washed out.
Noticeable artifacts like halo effects or unnatural edges are present.
3. Fair: Acceptable fusion. Key thermal targets and basic background details are
preserved. However, the contrast is low, the image lacks sharpness, or there is slight
visual inconsistency (unnatural look) between the fused layers.
4. Good: Effective fusion. High contrast with clear thermal targets and sharp
background textures. Information from both sources is well-integrated with minimal
artifacts. The image looks natural and informative.
5. Excellent: Optimal synergy. Exceptional clarity and distinctness. Thermal
targets are highly prominent, and background textures are rich and sharp. The fusion
is seamless with zero artifacts, maximizing the complementary information for human or
machine perception.
Please first provide a detailed analysis of the evaluation process within the <Thought>
tag. focus on:
- Thermal Target Prominence: Are the infrared targets (e.g., pedestrians, cars) bright
and distinct?
- Texture Detail Preservation: Are the edges and background details from the visible
image sharp?
- Artifacts: Are there halo effects, ghosting, or noise introduced by the fusion?
- Overall Naturalness: Is the overall brightness and contrast visually pleasing and
natural?
Then, give a final score from 1 to 5 within the <Score> tag.
<Thought>
[Analyze the evaluation process in detail here]
</Thought>
<Score>X</Score>
```

E Extended Experimental Results

Due to space limitations in the main text, we present additional qualitative and quantitative evaluations for both clean and degraded scenarios here. While existing SOTA methods exhibit a pronounced decline in performance under severe environmental perturbations, our method maintains exceptional fusion quality across all conditions. These results further validate the robustness and versatility of our framework in both ideal and challenging real-world environments.

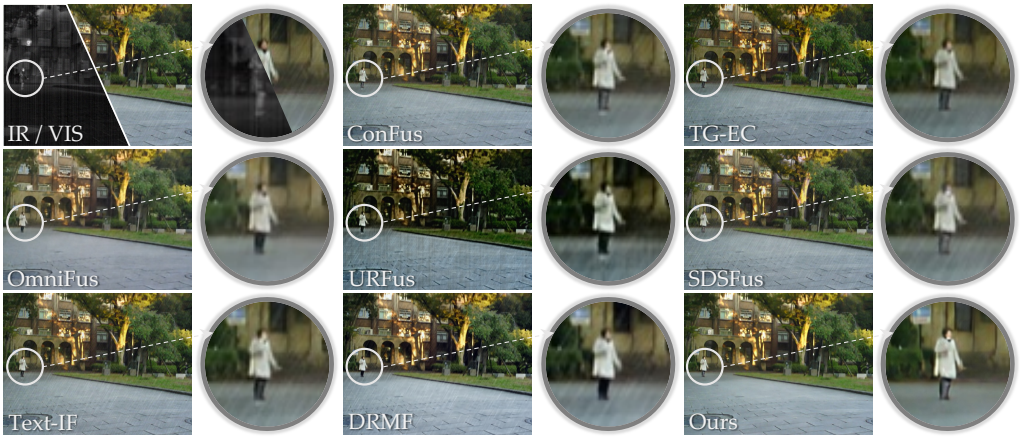

Figure 9: Fusion qualitative comparisons between SOTA methods and our OneHRL. It presents stripe (infrared images) with rain (visible images) degraded scenarios in the MSRS dataset. In all cases, OneHRL achieves well-performance results.

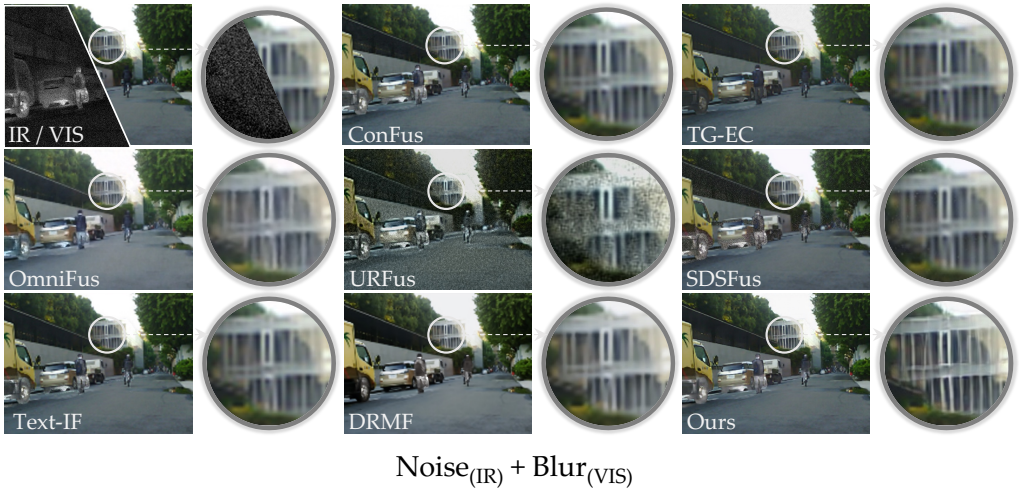

Figure 10: Fusion qualitative comparisons between SOTA methods and our OneHRL. It illustrates noise (infrared images) with blur (visible images) degraded scenarios in the MSRS dataset. In all cases, OneHRL achieves well-performance results.

Table 7: Quantitative comparison of fusion under Noise(IR)-Blur(VIS) scenario.

Method	**M^{3}FD**					**MSRS**				
	CC	SCD	MI	VIF	MSE	CC	SCD	MI	VIF	MSE
ConFus	0.473	1.342	2.748	**0.400**	0.059	0.576	1.421	2.859	0.617	0.043
DRMF	0.414	1.006	**3.360**	0.384	0.067	0.509	1.051	**3.243**	**0.635**	0.068
OmniFus	0.443	1.188	3.011	0.393	0.060	0.529	1.114	2.640	0.605	0.064
SDSFus	0.432	1.165	2.861	0.300	0.059	0.577	1.333	2.265	0.439	0.039
Text-IF	0.387	1.099	1.050	0.200	0.056	**0.586**	**1.444**	2.655	0.555	0.040
TG-EC	0.458	1.233	2.718	0.370	0.051	0.565	1.148	2.413	0.464	**0.034**
URFus	**0.500**	**1.498**	1.278	0.237	**0.050**	0.581	1.361	1.418	0.360	0.050
Ours	**0.511**	**1.506**	**3.116**	**0.452**	**0.049**	**0.614**	**1.549**	**3.713**	**0.673**	**0.032**



$\text{Noise}_{(\text{IR})} + \text{Blur}_{(\text{VIS})}$

Figure 11: Fusion qualitative comparisons between SOTA methods and our OneHRL. It illustrates noise (infrared images) with blur (visible images) degraded scenarios in the M^3FD dataset. In all cases, OneHRL achieves well-performance results.



$\text{Stripe}_{(\text{IR})} + \text{Rain}_{(\text{VIS})}$

Figure 12: Fusion qualitative comparisons between SOTA methods and our OneHRL. It presents stripe (infrared images) with rain (visible images) degraded scenarios in the M^3FD dataset. In allcases, OneHRL achieves well-performance results.

Table 8: Quantitative comparison of fusion under Stripe(IR)-Rain(VIS) scenario.

Method	**M^3FD**					**MSRS**				
	CC	SCD	MI	VIF	MSE	CC	SCD	MI	VIF	MSE
ConFus	0.488	1.381	2.851	0.621	0.056	**0.600**	**1.529**	2.952	**0.757**	0.035
DRMF	0.431	1.081	**3.331**	**0.632**	0.070	0.519	1.103	**3.114**	0.678	0.064
OmniFus	0.472	1.294	2.839	0.555	0.051	0.535	1.105	2.506	0.638	0.062
SDSFus	0.475	1.304	3.280	0.572	0.049	0.597	1.438	2.879	0.672	0.033
Text-IF	0.481	1.394	2.552	0.595	0.056	0.589	1.490	3.049	0.728	0.035
TG-EC	0.502	1.484	2.582	0.603	**0.045**	0.578	1.326	2.480	0.603	**0.029**
URFus	**0.538**	**1.617**	1.998	0.553	0.048	0.587	1.414	1.919	0.626	0.050
Ours	**0.543**	**1.644**	**3.310**	**0.673**	**0.045**	**0.620**	**1.618**	**3.382**	**0.775**	**0.032**

Table 9: Quantitative comparison of detection under no degradation scenario.

Method	AP@.5						mAP@.5
	People	Car	Bus	Motor	Truck	Lamp	
ConFus	0.655	**0.866**	0.733	0.449	0.692	**0.533**	0.655
DRMF	0.590	0.831	0.683	0.342	0.563	0.458	0.578
OmniFus	0.616	0.843	0.690	0.265	0.679	0.506	0.600
SDSFus	0.612	0.769	0.505	0.289	0.449	0.288	0.486
Text-IF	**0.658**	0.858	0.751	0.468	0.701	0.522	0.659
TG-EC	0.614	0.856	**0.796**	**0.507**	**0.715**	0.526	**0.669**
URFus	0.610	0.850	0.779	0.426	0.702	0.513	0.646
Ours	**0.656**	**0.872**	**0.803**	**0.512**	**0.718**	**0.538**	**0.683**

Table 10: Quantitative comparison of segmentation under no degradation scenario.

Method	IoU									mIoU
	BG	Car	Person	Bike	Curve	C S	GR	C C	Bump	
ConFus	0.973	0.827	**0.604**	0.616	**0.224**	0.477	0.571	0.327	**0.481**	0.567
DRMF	0.970	0.795	0.564	0.601	0.133	0.489	**0.638**	0.424	0.324	0.549
OmniFus	0.970	0.803	0.521	0.556	0.121	0.483	0.495	0.416	0.365	0.526
SDSFus	0.972	0.813	0.584	0.582	0.161	0.531	0.587	0.371	0.433	0.560
Text-IF	**0.973**	**0.828**	0.587	**0.625**	0.127	**0.535**	0.607	**0.426**	0.425	**0.570**
TG-EC	0.971	0.793	0.594	0.590	0.098	0.493	0.553	0.358	0.404	0.539
URFus	0.971	0.813	0.579	0.597	0.163	0.510	0.543	0.358	0.430	0.552
Ours	**0.976**	**0.842**	**0.623**	**0.626**	**0.271**	**0.590**	**0.699**	**0.445**	**0.502**	**0.619**

Table 11: Quantitative comparison of detection under Noise(IR)-Blur(VIS) scenario.

Method	AP@.5						mAP@.5
	People	Car	Bus	Motor	Truck	Lamp	
ConFus	**0.621**	0.782	0.680	0.310	0.548	0.356	0.550
DRMF	0.514	0.716	0.486	0.269	0.326	0.213	0.421
OmniFus	0.600	**0.814**	0.627	0.226	**0.589**	0.398	0.542
SDSFus	0.457	0.549	0.103	0.065	0.132	0.162	0.245
Text-IF	0.440	0.406	0.053	0.002	0.011	0.039	0.159
TG-EC	0.532	0.773	**0.724**	**0.380**	0.551	**0.412**	**0.562**
URFus	0.501	0.467	0.275	0.045	0.097	0.094	0.247
Ours	**0.629**	**0.798**	**0.756**	**0.394**	**0.652**	**0.497**	**0.621**

Table 12: Quantitative comparison of segmentation under Noise(IR)-Blur(VIS) scenario.

Method	IoU									mIoU
	BG	Car	Person	Bike	Curve	C S	GR	C C	Bump	
ConFus	**0.970**	**0.797**	**0.556**	**0.542**	0.173	0.478	0.449	0.335	**0.444**	**0.527**
DRMF	0.968	0.771	0.533	0.532	**0.186**	**0.482**	0.511	0.374	0.332	0.521
OmniFus	0.967	0.757	0.484	0.499	0.142	0.443	0.348	**0.382**	0.441	0.496
SDSFus	0.959	0.659	0.315	0.279	0.008	0.305	0.423	0.234	0.402	0.407
Text-IF	0.968	0.775	0.541	0.505	0.140	0.461	0.497	0.323	0.412	0.513
TG-EC	0.968	0.778	0.486	0.522	0.105	0.444	**0.541**	0.328	0.403	0.508
URFus	0.950	0.504	0.315	0.251	0.047	0.194	0.005	0.165	0.001	0.275
Ours	**0.974**	**0.824**	**0.597**	**0.569**	**0.249**	**0.589**	**0.687**	**0.436**	**0.510**	**0.601**

Table 13: Quantitative comparison of detection under Stripe(IR)-Rain(VIS) scenario.

Method	AP@.5						mAP@.5
	People	Car	Bus	Motor	Truck	Lamp	
ConFus	**0.653**	**0.851**	0.717	**0.414**	0.635	0.495	0.628
DRMF	0.582	0.786	0.436	0.264	0.506	0.393	0.494
OmniFus	0.598	0.811	0.629	0.238	0.593	0.406	0.546
SDSFus	0.605	0.682	0.261	0.138	0.251	0.215	0.359
Text-IF	**0.653**	0.796	0.641	0.296	0.593	0.463	0.574
TG-EC	0.563	0.801	0.660	0.353	0.575	0.483	0.572
URFus	0.609	0.791	**0.747**	0.391	**0.676**	**0.580**	**0.632**
Ours	0.651	**0.863**	**0.759**	**0.418**	**0.656**	**0.518**	**0.644**

Table 14: Quantitative comparison of segmentation under Stripe(IR)-Rain(VIS) scenario.

Method	IoU									mIoU
	BG	Car	Person	Bike	Curve	C S	GR	C C	Bump	
ConFus	**0.965**	0.711	**0.545**	**0.563**	0.001	**0.392**	0.306	0.326	0.143	0.441
DRMF	0.961	0.681	0.490	0.425	0.003	0.230	0.264	0.290	0.008	0.385
OmniFus	0.965	**0.756**	0.466	0.489	**0.007**	0.317	**0.326**	**0.357**	**0.289**	**0.449**
SDSFus	0.958	0.585	0.460	0.460	0.001	0.289	0.236	0.249	0.115	0.374
Text-IF	0.961	0.661	0.511	0.484	0.001	0.230	0.283	0.264	0.106	0.390
TG-EC	0.956	0.596	0.361	0.439	0.001	0.208	0.243	0.220	0.139	0.352
URFus	0.962	0.685	0.513	0.498	0.003	0.303	0.297	0.173	0.142	0.401
Ours	**0.975**	**0.832**	**0.617**	**0.609**	**0.253**	**0.566**	**0.653**	**0.447**	**0.530**	**0.609**

Table 15: Quantitative comparison of detection under NoiseStripe(IR)-HazeNoise(VIS) scenario.

Method	AP@.5						mAP@.5
	People	Car	Bus	Motor	Truck	Lamp	
ConFus	**0.522**	**0.475**	0.094	0.010	0.041	0.030	**0.196**
DRMF	0.431	0.400	**0.101**	**0.026**	0.034	/	0.165
OmniFus	0.350	0.467	0.022	0.011	**0.069**	0.041	0.160
SDSFus	0.249	0.230	/	/	0.011	/	0.081
Text-IF	0.365	0.411	0.091	/	0.016	**0.054**	0.156
TG-EC	0.316	0.397	0.053	/	0.023	0.012	0.134
URFus	0.477	0.447	0.048	/	0.036	0.022	0.172
Ours	**0.606**	**0.760**	**0.572**	**0.161**	**0.490**	**0.228**	**0.470**

Table 16: Quantitative comparison of segmentation under NoiseStripe(IR)-HazeNoise(VIS) scenario.

Method	IoU									mIoU
	BG	Car	Person	Bike	Curve	C S	GR	C C	Bump	
ConFus	0.942	0.363	0.216	**0.279**	0.001	0.005	/	0.004	/	0.213
DRMF	0.931	0.007	**0.347**	0.007	/	/	/	/	/	0.159
OmniFus	**0.945**	**0.440**	0.272	0.185	/	**0.008**	/	**0.109**	**0.008**	**0.236**
SDSFus	0.927	0.007	/	0.006	/	/	/	/	/	0.119
Text-IF	0.928	0.006	0.003	0.007	/	/	/	/	/	0.123
TG-EC	0.944	0.417	0.232	0.171	/	0.007	**0.001**	0.008	0.007	0.223
URFus	0.933	0.159	0.185	0.212	0.001	**0.001**	/	/	/	0.169
Ours	**0.970**	**0.772**	**0.575**	**0.459**	**0.222**	**0.484**	**0.509**	**0.369**	**0.390**	**0.528**

Table 17: Quantitative comparison of detection under NoiseStripe(IR)-HazeRain(VIS) scenario.

Method	AP@.5						mAP@.5
	People	Car	Bus	Motor	Truck	Lamp	
ConFus	**0.621**	**0.762**	**0.535**	**0.222**	0.305	**0.299**	**0.457**
DRMF	0.535	0.667	0.330	0.170	**0.333**	0.166	0.367
OmniFus	0.393	0.582	0.238	0.026	0.167	0.089	0.249
SDSFus	0.283	0.552	0.038	0.008	0.065	0.101	0.175
Text-IF	0.444	0.621	0.251	0.096	0.142	0.158	0.285
TG-EC	0.202	0.580	0.219	0.041	0.194	0.150	0.231
URFus	0.469	0.400	0.024	/	0.005	0.019	0.153
Ours	**0.636**	**0.820**	**0.623**	**0.330**	**0.599**	**0.400**	**0.568**

Table 18: Quantitative comparison of segmentation under NoiseStripe(IR)-HazeRain(VIS) scenario.

Method	IoU									mIoU
	BG	Car	Person	Bike	Curve	C S	GR	C C	Bump	
ConFus	**0.964**	**0.696**	**0.488**	**0.542**	**0.001**	**0.319**	0.008	**0.287**	**0.182**	**0.397**
DRMF	0.954	0.557	0.438	0.337	/	0.110	0.108	0.174	0.004	0.304
OmniFus	0.949	0.509	0.274	0.247	/	0.134	**0.135**	0.008	0.105	0.271
SDSFus	0.938	0.280	0.129	0.219	/	0.005	0.003	0.007	/	0.193
Text-IF	0.945	0.387	0.262	0.283	/	0.006	0.002	0.106	0.003	0.233
TG-EC	0.940	0.311	0.413	0.243	/	0.005	0.001	0.008	/	0.199
URFus	0.933	0.134	0.212	0.240	/	0.004	/	/	/	0.174
Ours	**0.974**	**0.818**	**0.602**	**0.566**	**0.224**	**0.572**	**0.636**	**0.441**	**0.483**	**0.591**

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction of the paper clearly state the claims made, including the contributions made in the paper and important assumptions and limitations.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [No]

   Justification: The paper do not discusses the limitations of the work.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: All the theorems, formulas, and proofs in the paper have been numbered and cross-referenced, and all assumptions have been clearly stated or referenced in the statement of any theorems.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: The paper has fully disclosed all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and conclusions of the paper.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: We will open access to the code after the paper is accepted.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: We have specified all the training and test details necessary to understand the results.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [No]

   Justification: The paper does not report error bars.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).

- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: The paper provides sufficient information on the computer resource for experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics https://neurips.cc/public/EthicsGuidelines?

   Answer: [Yes]

   Justification: The research conducted in the paper conforms in every respect with the NeurIPS Code of Ethics, ensuring ethical standards are upheld throughout the study.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: The paper discusses both potential positive societal impacts and negative societal impacts of the work performed.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.

- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [Yes]

    Justification: The paper describes safeguards that have been put in place for the responsible release of data or models that have a high risk for misuse, ensuring that appropriate measures are taken to mitigate potential risks.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: The creators or original owners of assets used in the paper are properly credited, and the license and terms of use are explicitly mentioned and properly respected.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.

- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: The new assets introduced in the paper are well documented, and the documentation is provided alongside the assets.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.

- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: The core method development in the paper does not involve LLMs as any important, original, or non-standard components.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.