Enhancing Driving Visibility via Semantic-Guided Knowledge Distillation Framework for Adverse Weather Removal

Hanvitha Saraswathi Mukkamala*
International Institute of Information Technology
Hyderabad, Telangana, India
Mahindra University
Hyderabad, India
hanvitha.mukkamala@gmail.com

Shankar Gangisetty
International Institute of Information Technology
Hyderabad, India
shankar.gangisetty@ihub-data.iiit.ac.in

Ananya Kulkarni
International Institute of Information Technology
Hyderabad, India
ananyakulkarni2103@gmail.com

Veera Ganesh Yalla
International Institute of Information Technology
Hyderabad, India
cto@ihub-data.iiit.ac.in

C V Jawahar
International Institute of Information Technology
Hyderabad, India
jawahar@iiit.ac.in

Abstract

Adverse weather such as rain, haze, and low light, severely degrades visual perception in Advanced Driver Assistance Systems (ADAS) and autonomous driving, leading to degraded scene understanding and increased safety risks. We propose a unified, semantic-guided knowledge distillation restoration framework that addresses multi-weather removal while preserving semantics. Our method employs a semantic-guided dual-decoder architecture trained via two-stage multi-teacher knowledge distillation, transferring expertise from multiple high-capacity models into a lightweight student model. Segmentation-aware contrastive learning further aligns low-level restoration with high-level semantic structure, enabling robust detection of roads, vehicles, and pedestrians under challenging conditions. Trained on a mix of synthetic and real-world data with segmentation-guided feature refinement, our framework generalises effectively to real-world unseen environments. Extensive experiments on multiple benchmarks show competitive or superior performance to state-of-the-art methods, with real-time inference suitable for edge deployment. This makes our approach well-suited for safety-critical perception in autonomous and semi-autonomous systems operating in adverse outdoor environments.

CCS Concepts

• **Computing methodologies** → *Machine learning*; **Computer vision**; • **Human-centered computing** → *Human computer interaction (HCI)*; *Accessibility*; Ubiquitous and mobile computing; • **Applied computing** → *Military*.

*Work done while interning at IIIT-Hyderabad.

ICVGIP 2025, Mandi, India

ACM ISBN 979-8-4007-1930-1/25/12
https://doi.org/10.1145/3774521.3774599

Keywords

Adverse Weather Removal, Deraining, Dehazing, Knowledge Distillation, Semantic Guided, Lightweight CNNs, Autonomous Driving

ACM Reference Format:
Hanvitha Saraswathi Mukkamala, Shankar Gangisetty, Ananya Kulkarni, Veera Ganesh Yalla, and C V Jawahar. 2025. Enhancing Driving Visibility via Semantic-Guided Knowledge Distillation Framework for Adverse Weather Removal. In *Indian Conference on Computer Vision, Graphics, and Image Processing (ICVGIP 2025), December 17–20, 2025, Mandi, India.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3774521.3774599

1 Introduction

Adverse weather conditions such as heavy rain, dense haze, snow, low-light environments, and raindrop occlusions pose a significant challenge to both human drivers and autonomous driving systems, often leading to severely degraded visibility and increased risk of accidents. In such scenarios, onboard cameras and sensors in ADAS and autonomous vehicles frequently capture weather-distorted imagery, characterized by rain streaks, water droplets, scattering effects, and poor illumination. These degradations critically hinder key vision tasks such as object detection, semantic segmentation, and anomaly detection, which are essential for safe navigation and scene understanding.

Prior-based weather removal approaches have focused on single-weather restoration tasks [18, 39] such as deraining [16], desnowing [37], and dehazing [18], largely using deep neural networks that address specific types of degradation. Though effective in isolated scenarios, such approaches are insufficient for real-world applications such as autonomous driving, surveillance, and edge-based vision systems, where diverse and unpredictable weather conditions occur. Deploying multiple weather-specific models in such environments incurs substantial computational, storage, and maintenance overheads, making it impractical for large-scale or resource-constrained systems.

To overcome these limitations, recent research has explored all-in-one weather restoration models [12, 37, 46], which leverage a single set of network parameters to handle multiple degradation

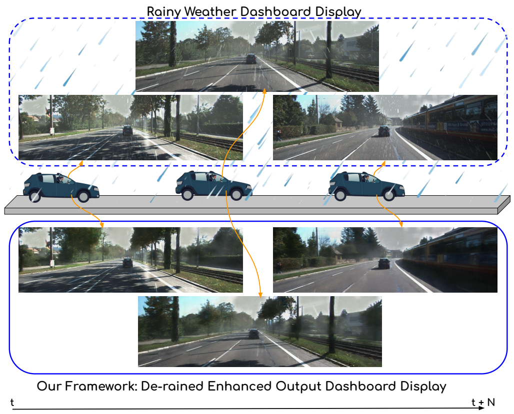

Figure 1: Illustration of our framework on ADAS vehicle in a rainy driving scenario for Adverse Weather Removal. The sequence is from KITTI [14] dataset.

types. However, these unified approaches face several challenges. First, some methods [30, 31, 51] fail to model the *distinct physical and visual characteristics* of each weather condition, resulting to suboptimal restoration in certain scenarios. Second, state-of-the-art unified networks [12, 23, 37, 38, 44, 46] often require *large parameter counts and high computational budgets*, limiting their real-time feasibility. Third, their heavy reliance on synthetic datasets during training causes significant *generalisation gaps* when deployed in real-world driving conditions, where degradations are more complex and less predictable. Fourth, most existing solutions overlook *semantic scene awareness* such as identifying roads, vehicles, and pedestrians which is essential for safety-critical perception tasks in autonomous driving. Finally, CNN-based methods [16, 18, 31, 51], despite their restoration quality, struggle with *inference latency and memory inefficiency*, making them unsuitable for real-time edge deployment.

To address these shortcomings, we propose a unified semantic-guided knowledge distillation framework built around five design factors (refer Sec. 2 of Supplementary for more details on Knowledge Distillation): *(i) Distinct physical and visual characteristics*—we fuse physics-based degradation cues (e.g., rain-streak density, haze scattering) with high-level semantics to preserve structure across conditions; *(ii) Large parameter counts and high computational budgets*—we distill multiple teachers into a compact, low-compute student suitable for real-time embedded deployment; *(iii) Generalisation gaps*—our label-free training encourages robustness to diverse, unseen degradations; *(iv) Semantic scene awareness*—joint restoration and segmentation preserve scene layout and object boundaries so safety-critical entities remain well delineated; *(v) Inference latency and memory inefficiency*—the resulting student sustains low latency and small memory footprint for edge inference.

We propose a semantic-guided knowledge distillation framework that extends the principles of MultiTS [4] through three key advancements. First, unlike MultiTS, which primarily emphasizes multi-teacher supervision for task adaptation, our framework introduces a semantic-guided dual-decoder architecture that jointly leverages image restoration and scene segmentation, enabling stronger semantic fidelity and sharper delineation of safety-critical regions such as roads, vehicles, and pedestrians. Second, a segmentation-aware contrastive regularization mechanism is incorporated to align the restoration process with semantic boundaries, ensuring perceptually accurate and structurally consistent outputs. Third, to achieve cross-weather generalization, we designed a label-efficient multi-teacher distillation strategy that aggregates complementary knowledge from multiple weather-specialized experts into a shared latent space, thereby eliminating the need for explicit weather labels. Through efficient optimization of memory and computation, the resulting lightweight student model attains near-real-time performance suitable for edge and embedded deployment. As illustrated in Fig. 1, under a heavy rain scenario, our model effectively restores visually degraded scenes while preserving semantic structure, enabling robust perception even under severe weather conditions. This unified framework not only delivers high-quality visual restoration but also maintains semantic consistency, ensuring improved robustness and generalization across diverse adverse weather domains.

The main contributions of our work are:

- Unified semantic-guided setup with a shared encoder and dual decoders for restoration and segmentation, distilled from multiple weather-specialized teachers to preserve structural consistency and stay robust across degradations.
- Label-efficient cross-weather generalisation by distilling complementary knowledge from multiple weather-specialized teachers, eliminating the need for explicit weather labels.
- Broad evaluation on synthetic and real datasets, with reference and non-reference metrics, qualitative visuals, and ablation experiments.
- Model efficiency and real-time deployment with a student-only inference pipeline that achieves low latency, low memory, and stable FPS on edge devices.

2 Related Work

2.1 CNN-based Adverse Weather Removal

Single-Weather Removal. Convolutional neural networks (CNNs) have been widely adopted for adverse weather removal due to their strong local feature extraction and relatively low computational cost. Early single-weather approaches targeted specific degradations such as rain or haze, with methods like EfficientDeRain [22] using dense U-Net style architectures for progressive refinement in deraining, and DehazeNet [8] and AOD-Net [9] exploiting atmospheric priors such as transmission maps and global light estimation for dehazing. Subsequent models, including FFA-Net [10], ACERNet [42], MITNet [34], DEA-Net [5], and the high-capacity Dehaze-XL [3], enhanced restoration quality through attention mechanisms, multi-scale feature fusion, or adaptive enhancement strategies, while strong deraining baselines such as MSPFN [16], MPR-Net [47], CCN [32], and CSUD [7] leveraged multi-stage refinement, progressive residual learning, and context aggregation. Although these models achieve strong performance, their high parameter counts hinder real-time deployment on embedded platforms. To address efficiency, lightweight CNNs like ESRNet [45] with factorized spatial–channel attention, and FastDehazeNet [2] or simplified AOD-Net variants with shallow 1×1 convolutions, have been introduced, offering real-time performance on mobile devices but often sacrificing robustness and contextual understanding, particularly under mixed-weather or low-light conditions.

Multi-Weather Restoration. While single-weather models perform well in controlled conditions, real-world driving scenarios often involve compound or changing weather, where training separate models for each degradation is inefficient. This challenge led to the development of multi-weather restoration frameworks, which aim to generalise across rain, haze, and snow within a single network. AllWeatherNet [31] introduced a shared encoder with weather-specific decoders for adaptive restoration, while WGWS-Net [51] used gated weather-specific feature fusion to dynamically combine information from multiple weather branches. Multi-TS [4] further advanced this idea by distilling knowledge from specialized single-weather teachers into a unified compact student network, improving generalisation while maintaining efficiency.

Semantic-Guided Restoration. Despite these advances, CNN-based methods still face notable limitations such as heavy architectures hindering real-time use, lightweight models often sacrifice robustness in mixed weather, and most frameworks lack high-level semantic understanding of scene elements like roads, vehicles, pedestrians or traffic signs, information that is critical for preserving structural consistency.

To overcome these limitations, recent works have explored combining CNN efficiency with semantic guidance. Incorporating semantic and contextual cues has proven effective in related perception tasks; for example, detection frameworks such as D-YOLO [6] leverage weather-robust features to maintain reliable object detection under challenging conditions, while segmentation-oriented approaches extend these ideas to provide dense contextual priors for restoration. Integrating semantic maps helps restoration focus on perceptually important regions (e.g., roads, vehicles, traffic signs), thereby enhancing structural consistency in adverse conditions [40, 43]. Inspired by this, we extend semantic-guided restoration to multi-weather scenarios using a dual-task CNN trained via multi-teacher distillation. This unified framework combines knowledge from rain, haze, and low-light experts into a single lightweight model, ensuring both high restoration fidelity and real-time feasibility for autonomous driving applications.

2.2 Transformer-based Adverse Weather Removal

Transformer methods have recently advanced weather restoration by leveraging self-attention to capture long-range dependencies, which is particularly beneficial for degradations like haze and snow that require global context. Models such as Restormer [46] uses multi-stage transformer blocks with cross-modality attention for progressive rain removal while preserving fine details. For haze, SwinIR [24] introduces hierarchical shifted-window attention to balance local and global features, improving performance in dense haze. TransWeather [37] and MWFormer-Weather [50] extend this capability to multi-weather conditions by incorporating specialized attention mechanisms for unified restoration. Lightweight variants such as MobileViT-Weather [26] attempt to reduce complexity using mobile-friendly attention mechanisms.

Despite their impressive performance, full transformer models often require large memory (15–30 GB VRAM) and suffer from high inference latency (hundreds of milliseconds per frame), limiting their use in real-time systems. Even optimized versions, though more efficient, may still underperform in preserving structural

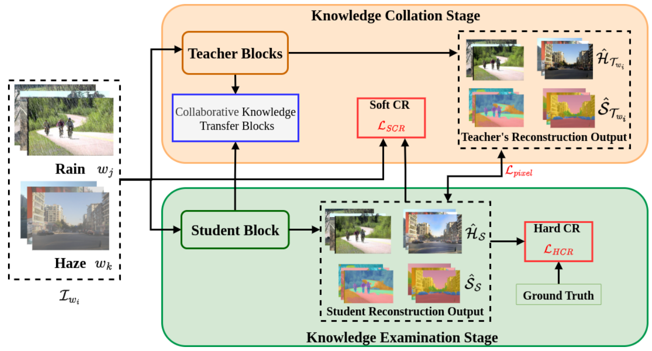

Figure 2: Overview of our semantic-guided knowledge distillation framework. Collaborative distillation from multiple weather-specific teachers (rain, haze) to a unified student via two stages: (1) Knowledge Collation, where soft alignment with teacher reconstructions and segmentation priors, and (2) Knowledge Examination, where enforcing ground-truth consistency with hard constraints. Segmentation maps provide region-aware supervision, emphasizing critical structures like roads and vehicles.

fidelity under mixed or severe degradations. To address this, we utilize transformers not for direct deployment but as teacher models during training. By distilling their rich contextual knowledge into a compact CNN-based student, we retain high-quality restoration and robust generalisation across weather types without incurring the computational and memory overhead of transformer inference.

2.3 Diffusion-based Adverse Weather Removal

Diffusion-based generative models have recently achieved state-of-the-art perceptual quality in image restoration by iteratively denoising toward clean outputs. Their ability to model complex data distributions makes them particularly effective for compound degradations (e.g., rain–fog or rain–snow), where CNNs and transformers often struggle. Multi-weather model such as MDGTNet [23] use dual textual prompts and CLIP-guided semantic cues to preserve scene structure under diverse weather conditions. DiffuWeather [29] addresses the high latency of diffusion by reducing denoising steps (e.g., from 100 to 20), making diffusion more practical for real-world use. DB-STNet [35] extends this trend by adopting a diffusion-based backbone to restore rain, haze, and snow simultaneously, benefiting from the perceptual quality of diffusion while suffering from high inference latency and memory consumption due to iterative denoising. Other approaches, such as DomainTrans [30], apply cross-domain translation to handle multiple weather types without needing explicit labels. However, these models remain computationally intensive. Full diffusion pipelines demand high VRAM and multi-step inference, making them unsuitable for real-time or edge deployment. Additionally, generative translation methods may introduce structural artifacts and often require large paired datasets for training.

In our work, diffusion and generative models serve as teacher networks during training. By distilling their semantic richness and perceptual quality into a compact CNN-based student, we achieve high-fidelity restoration with the efficiency required for real-time, resource-constrained deployment in autonomous driving scenarios.

3 Semantic-Guided Knowledge Distillation Framework

3.1 Overview

Our objective is to build a unified, real-time framework that addresses multiple adverse weather conditions such as rain and haze, while simultaneously enabling high-fidelity semantic segmentation. This dual-task setup is essential for autonomous navigation systems, where both visual clarity and scene understanding are critical for safe operation. In contrast to traditional approaches, which treat restoration and segmentation as separate tasks.

Problem Formulation. Given an input image $\mathcal{I}w_i$ captured under a specific weather condition $w_i \in \mathcal{W} = w_1, w_2, \ldots, w_K$, our goal is to restore a clean image $\mathcal{H}w_i$ using a lightweight student model $\mathcal{S}$ that operates in real-time. To achieve this, we propose a Two-Stage Knowledge Distillation framework inspired from [4], comprising Knowledge Collation (KC) and Knowledge Examination (KE) stages, as illustrated in Fig. 2. In the KC stage, we employ a set of expert teacher models $\mathcal{T}_{w_i},\ i = 1\ to\ K$, where each $\mathcal{T}_{w_i}$ is specialized for a specific weather type w_i. This design allows each teacher to capture degradation-specific priors that are otherwise difficult to encode in a unified teacher trained on heterogeneous data. The student model $\mathcal{S}$ aggregates this diverse knowledge through a Collaborative Knowledge Transfer (CKT) process, aligning its internal representations with the teachers using two key modules: the Progressive Feature Projector (PFP) [4], which ensures fine-grained spatial alignment at early decoder stages, and the Bi-directional Feature Matching (BFM) [4], which enforces semantic and perceptual consistency between the restoration and segmentation outputs. In the KE stage, we further refine the student via task-specific objectives that enhance semantic fidelity and visual quality. Notably, only the student model $\mathcal{S}$ is deployed at inference, enabling real-time processing of continuous image streams with high restoration accuracy. This formulation allows the model to generalise across diverse weather conditions while maintaining efficiency, making it suitable for autonomous driving.

3.2 Knowledge Collation (KC) stage

Dual-Decoder Architecture in Teachers and Student. To enable structured and semantically-aware restoration, we enhance each teacher model $\mathcal{T}_{w_i}$ with a dual-decoder design comprising: (i) an *image restoration decoder* that predicts the clean image $\hat{\mathcal{H}}_{\mathcal{T}_{w_i}}$, and (ii) a *segmentation decoder* that outputs the corresponding semantic map $\hat{\mathcal{S}}_{\mathcal{T}_{w_i}}$. These two outputs are used jointly to supervise the student model through both pixel-level and task-level distillation. The dual-branch *Collaborative Knowledge Transfer (CKT)* block aligns the teacher-student representations in both tasks, encouraging coherent structure across restoration and semantics.

Motivated by this multi-task supervision (image and segmentation tasks), the student model $\mathcal{S}$ adopts a mirrored dual-decoder design, comprising a Restoration Decoder that predicts the clean image $\hat{\mathcal{H}}_{\mathcal{S}}$ and a Segmentation Decoder that outputs the corresponding semantic map $\hat{\mathcal{S}}_{\mathcal{S}}$, denoted as *RD* and *SD*, respectively.

Collaborative Knowledge Transfer (CKT). During the CKT stage, the student model $\mathcal{S}$ is guided by all teachers concurrently as shown in Fig. 3. Let $F^{(q)}_{\mathcal{T}_{w_i}}$ and $F^{(q)}_{\mathcal{S}}$ denote the q-th layer features from teacher $\mathcal{T}_{w_i}$ and student $\mathcal{S}$, respectively. These features are projected into a shared latent space via a learnable Progressive Feature Projector $\phi(\cdot)$:

$$\tilde{F}^{(q)}_{\mathcal{T}_{w_i}} = \phi(F^{(q)}_{\mathcal{T}_{w_i}}), \quad \tilde{F}^{(q)}_{\mathcal{S}} = \phi(F^{(q)}_{\mathcal{S}}) \tag{1}$$

We then compute the Projected Feature Error (LPFE) loss using L_1 distance:

$$\mathcal{L}_{\text{PFE}} = \sum_{q=1}^{Q} \left\| \tilde{F}^{(q)}_{\mathcal{T}_{w_i}} - \tilde{F}^{(q)}_{\mathcal{S}} \right\|_1 \tag{2}$$

To ensure the validity of projected features, we use an inverse projector $\rho(\cdot)$ to reconstruct original features and apply a Projected Feature Verification (LPFV) loss:

$$\mathcal{L}_{\text{PFV}} = \sum_{q=1}^{Q} \left\| \rho(\tilde{F}^{(q)}_{\mathcal{T}_{w_i}}) - F^{(q)}_{\mathcal{T}_{w_i}} \right\|_1 \tag{3}$$

The total collaborative distillation loss is:

$$\mathcal{L}_{\text{CKT}} = \mathcal{L}_{\text{PFE}} + \mathcal{L}_{\text{PFV}} \tag{4}$$

Soft Contrastive Regularization (SCR). Since the student model $\mathcal{S}$ is still in an early learning stage, we adopt a soft contrastive formulation. The teacher output $\hat{\mathcal{H}}_{\mathcal{T}_{w_i}}$, $\hat{\mathcal{S}}_{\mathcal{T}_{w_i}}$ and student output $\hat{\mathcal{H}}_{\mathcal{S}}$, $\hat{\mathcal{S}}_{\mathcal{S}}$ are treated as the positive sample, while degraded images from other weather types $\{\mathcal{I}_{w_j}\}_{j \neq i}$ form the negative set. The SCR loss is defined as:

$$\mathcal{L}_{\text{SCR}} = Q\left(\hat{\mathcal{H}}_{\mathcal{T}_{w_i}},\ \hat{\mathcal{S}}_{\mathcal{T}_{w_i}}, \hat{\mathcal{H}}_{\mathcal{S}},\ \hat{\mathcal{S}}_{\mathcal{S}},\ \{\mathcal{I}_{w_j}\}_{j \neq i}\right) \tag{5}$$

where $Q(\cdot)$ is the contrastive loss with temperature-scaled dot product similarity.

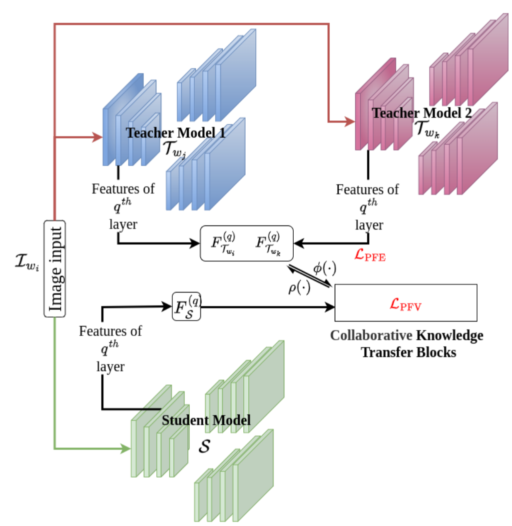

Figure 3: Collaborative Knowledge Transfer: The student learns from multiple weather-specific teachers via feature projection and verification, using $\mathcal{L}_{PFE}$ and $\mathcal{L}_{PFV}$ to align representations for unified restoration and segmentation.

3.3 Knowledge Examination (KE) Stage

After sufficient training under teacher supervision, the student model $\mathcal{S}$ enters the KE stage for self-guided refinement. Here, teacher models are detached, and the student is optimized using ground-truth labels. To strengthen robustness and task alignment, we use Hard Contrastive Regularization (HCR).

Let $\mathcal{J}_{\text{GT}}$ denote the ground-truth clean image, and $\mathcal{I}_{w_i}$ be the input degraded image. The positive pair is $(\hat{\mathcal{H}}_{\mathcal{S}}, \mathcal{J}_{\text{GT}})$ and negatives are images degraded by all weather types. The HCR loss is:

$$\mathcal{L}_{\text{HCR}} = Q\left(\hat{\mathcal{H}}_{\mathcal{S}}, \mathcal{J}_{\text{GT}}, \{\mathcal{I}_{w_j}\}_{j=1}^{K}\right) \tag{6}$$

In addition, we compute a pixel-wise reconstruction loss:

$$\mathcal{L}_{\text{Pixel}} = \left\|\hat{\mathcal{H}}_{\mathcal{S}} - \mathcal{J}_{\text{GT}}\right\|_1 \tag{7}$$

The final KE stage loss is a weighted combination:

$$\mathcal{L}_{\text{KE}} = \mathcal{L}_{\text{Pixel}} + \lambda_{\text{HCR}} \cdot \mathcal{L}_{\text{HCR}} \tag{8}$$

The two-stage training thus equips the student model with strong generalisation to mixed and unseen weather conditions, while ensuring real-time inference with a single compact model.

Segmentation Supervision and Loss. To enforce semantic fidelity during training, the segmentation decoder *SD* is supervised using a composite loss function. Let $\hat{\mathcal{S}}_{\mathcal{S}}$ be the predicted semantic map and $\mathcal{J}_{\text{S}}$ the one-hot ground-truth segmentation mask over c classes and spatial dimensions $h \times w$.

We use a weighted sum of pixel-level cross-entropy and global Dice loss:

$$\mathcal{L}_{\text{CE}} = -\sum_{c,h,w} \mathcal{J}_{\text{S}} \log \hat{\mathcal{S}}_{\mathcal{S}} \tag{9}$$

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2\sum_{c,h,w} \mathcal{J}_{\text{S}}\hat{\mathcal{S}}_{\mathcal{S}}}{\sum_{c,h,w} \mathcal{J}_{\text{S}} + \sum_{c,h,w} \hat{\mathcal{S}}_{\mathcal{S}}} \tag{10}$$

$$\mathcal{L}_{\text{Seg}} = \mathcal{L}_{\text{CE}} + \zeta \cdot \mathcal{L}_{\text{Dice}}, \quad \text{with } \zeta = 1.0 \tag{11}$$

This auxiliary segmentation loss $\mathcal{L}_{\text{Seg}}$ plays a critical role in stabilizing training and enhancing spatial consistency. It encourages the student model to maintain semantic integrity, which in turn constrains the restoration decoder *RD* to generate outputs that respect object boundaries and preserve spatial structure.

3.4 Training Objectives

The training process is divided into two stages:

Knowledge Collation (KC): The student model $\mathcal{S}$ is trained using supervision from teacher models. The loss combines:

$$\mathcal{L}_{\text{KC}} = \lambda_{\text{CKT}}\mathcal{L}_{\text{CKT}} + \lambda_{\text{SCR}}\mathcal{L}_{\text{SCR}} + \lambda_{\text{Seg}}\mathcal{L}_{\text{Seg}} \tag{12}$$

where $\mathcal{L}_{\text{CKT}}$ aligns features via collaborative distillation, $\mathcal{L}_{\text{SCR}}$ applies soft contrastive regularization, and $\mathcal{L}_{\text{Seg}}$ supervises the segmentation decoder.

Knowledge Examination (KE): The student is further refined without teacher guidance using:

$$\mathcal{L}_{\text{KE}} = \mathcal{L}_{\text{Pixel}} + \lambda_{\text{HCR}}\mathcal{L}_{\text{HCR}} \tag{13}$$

where $\mathcal{L}_{\text{Pixel}}$ is an L_1 reconstruction loss and $\mathcal{L}_{\text{HCR}}$ enforces hard contrastive regularization using ground truth.

Through this two-stage semantic-guided knowledge distillation process, the lightweight student model acquires the restoration fidelity of high-capacity teacher networks while preserving semantic integrity across adverse weather conditions. The dual-decoder design ensures that both low-level visual clarity and high-level scene understanding are optimized in tandem, enabling the network to prioritize perceptually critical regions such as roads, vehicles, and pedestrians. As a result, the final student model operates in real-time with minimal memory footprint, generalises effectively to unseen and mixed degradations, and is deployment-ready.

4 Experiments and Results

4.1 Experimental Setup

Datasets. We employ a diverse set of datasets to train and evaluate our unified knowledge distillation framework, ensuring robust generalisation across multiple adverse weather conditions. For rain degradations, we use Rain1400 [13], Rain100H [48] and SPA-Data [40], covering both synthetic and real-world rain with varied streak densities. Haze degradations are addressed using the REal-world VIdeo DEhazing dataset (REVIDE) [49] and REalistic Single Image DEhazing (RESIDE) [19] datasets, which includes Synthetic Objective Testing Set (SOTS) along with the Real-world Task-driven Testing Set (RTTS) subset for real-world haze evaluation. For specialized degradations, we incorporate the Raindrop [17] dataset for lens occlusions. To improve robustness to varying severity, we additionally synthesize degraded images by adjusting parameters such as rain streak length and haze scattering coefficients.

Details of the baseline models and implementation settings are provided in the Supplementary Material.

4.2 Quantitative Comparison

Results on Rain/RainDrop Dataset. In Table 1, on synthetic datasets, our method matches WGWS-Net's performance on Rain1400 and surpasses it on Outdoor-Rain by +1.23 dB. Relative to the multi-weather baseline MultiTS, we achieve consistent gains of +1.68 dB on Rain1400 and +1.41 dB on Outdoor-Rain, along with notable SSIM improvements. Overall, the results highlight the effectiveness of semantic guidance and multi-teacher distillation in preserving structure details across diverse rain conditions. On real-world datasets, our model is within 0.12 dB of WGWS-Net on SPA, while outperforming multi-weather baselines such as MultiTS and All-WeatherNet by +2.50 dB to +5.60 dB. On Raindrop, it delivers gains of +1.45 to 3.82 dB over competitive multi-weather approaches, effectively restoring heavily occluded regions. Overall, while specialized single-weather models may have a slight advantage in certain controlled synthetic cases, our framework achieves consistently strong, well-balanced performance across both synthetic and real-world rain, maintaining high visual fidelity and structural consistency under irregular, and spatially varying degradations.

Results on Haze Dataset. In Table 2, on the synthetic RESIDE dataset, our method delivers performance on par with WGWS-Net while clearly surpassing other multi-task baselines, achieving a notable +2.44 dB gain over MultiTS. For O-HAZE dataset, the advantage becomes more pronounced, with improvements of +0.56 dB over WGWS-Net and +1.20 dB over MultiTS, demonstrating the impact of semantic guidance and multi-teacher distillation in retaining fine structural details in haze. In the real-world REVIDE dataset as shown in Table 2, our framework outperforms MultiTS and All-WeatherNet by +1.14 dB to +3.18 dB, and although PSNR is 0.38 dB lower than WGWS-Net, we achieve superior SSIM, reflecting better

Figure 4: Visual comparison of the proposed and existing methods from synthetic datasets (Outdoor-Rain [21], RESIDE [19]) for multi-weather restoration. The image can be zoomed in for improved visualization.

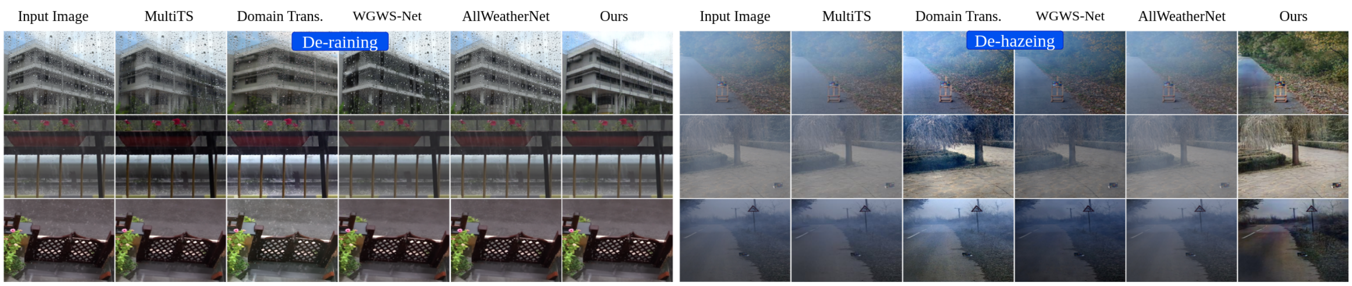

Figure 5: Visual comparison of the proposed and existing methods from real datasets (Raindrop [17], SPA [40] ,O-HAZE [1]) for multi-weather restoration. The image can be zoomed in for improved visualization.

preservation of structural consistency. Overall, these results illustrate that while single-weather dehazing models may have a slight edge in controlled synthetic settings, our approach consistency delivers a more robust trade-off between restoration quality, semantic coherence, and resilience in diverse haze conditions.

Visual Comparison on Synthetic and Real Datasets. In Fig. 4, we compare visual results on synthetic (Outdoor-Rain for de-raining, RESIDE for de-hazing) datasets. Our method removes both dense and fine rain streaks while preserving structure, color fidelity, and semantic clarity (e.g., vehicles, roads), producing sharper textures than MultiTS and Domain Translation, which leave artifacts or blur details. WGWS-Net reduces rain effectively but oversmooths, while AllWeatherNet suffers from color distortion and incomplete streak removal. For de-hazing, our approach restores visibility with balanced contrast, natural tones, and well-defined semantic regions, outperforming MultiTS and Domain Translation, which leave residual haze or muted colors. WGWS-Net over-enhances edges, and AllWeatherNet often produces lower-contrast outputs.

Table 1: Quantitative evaluation of image deraining performance on synthetic and real datasets. Best and second-best results are in bold and underlined, respectively.

Type	Method	Synthetic				Real			
		Rain1400 [13]		Outdoor-Rain [21]		SPA [40]		Raindrop [17]	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Deraining	MSPFN [15]	29.24	0.88	22.13	0.79	28.21	0.83	27.63	0.70
	MPR-Net [47]	31.17	0.91	23.83	0.84	32.55	0.87	28.21	0.72
	CCN [32]	30.12	0.85	21.73	0.82	31.79	0.85	30.51	<u>0.87</u>
	CSUD [7]	**32.48**	**0.93**	**27.58**	**0.92**	34.71	0.93	26.85	0.68
Multi-Task	MultiTS [4]	29.89	0.82	24.94	0.85	36.32	0.94	28.34	0.79
	DomainTranslation [30]	28.14	0.79	23.57	0.81	34.74	0.87	28.12	0.73
	WGWS-Net [51]	31.46	<u>0.91</u>	25.12	0.89	**38.94**	**0.98**	<u>30.71</u>	0.85
	AllWeatherNet [31]	27.31	0.74	22.39	0.82	33.22	0.83	27.67	0.69
	Ours	<u>31.57</u>	0.90	<u>26.35</u>	<u>0.91</u>	<u>38.82</u>	<u>0.96</u>	**32.16**	**0.89**

Table 2: Quantitative evaluation of image dehazing performance on synthetic and real datasets. Best and second-best results are in bold and underlined, respectively.

Type	Method	Synthetic		Real			
		RESIDE [20]		O-HAZE [1]		REVIDE [49]	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Dehazing	FFA-Net [10]	31.17	0.93	20.24	0.62	19.67	0.79
	AECR-Net [42]	30.98	0.91	19.67	0.64	20.65	0.77
	MIT-Net [34]	31.77	0.94	18.21	0.59	21.25	0.76
	DEA-Net [5]	32.74	0.91	19.89	0.67	21.31	0.80
	Dehaze-XL [3]	**33.56**	**0.96**	**21.18**	**0.71**	**23.43**	**0.82**
Multi-Task	MultiTS [4]	27.19	0.85	18.12	0.52	20.10	0.78
	DomainTranslation [30]	28.92	0.91	19.04	0.58	21.21	0.79
	WGWS-Net [51]	29.72	<u>0.94</u>	18.76	0.57	<u>21.62</u>	<u>0.86</u>
	AllWeatherNet [31]	26.21	0.79	17.35	0.49	18.06	0.75
	Ours	<u>29.63</u>	0.92	<u>19.32</u>	<u>0.61</u>	21.24	0.84

In Fig. 5, we present qualitative evaluation on real (Raindrop, SPA for de-raining, O-HAZE for de-hazing) datasets. Our approach completely removes large droplets and fine streaks in rain, and in haze restoration recovers distant scene details with natural colors and preserved semantic boundaries. Competing methods often leave haze, oversmooth textures, or distort colors. These results highlight the benefits of our semantic-guided design, which leverages semantic priors to direct restoration on critical regions, enabling more effective degradation removal while maintaining realistic textures and scene structure in both synthetic and real conditions.

4.3 Evaluation on Real World Images

In real-world settings without ground truth as shown in Table 3, we evaluate perceptual quality on the IDD-AW [33] dataset using NIQE [28] and BRISQUE [27], where lower scores indicate natural images. For rain, our method achieves the best results with

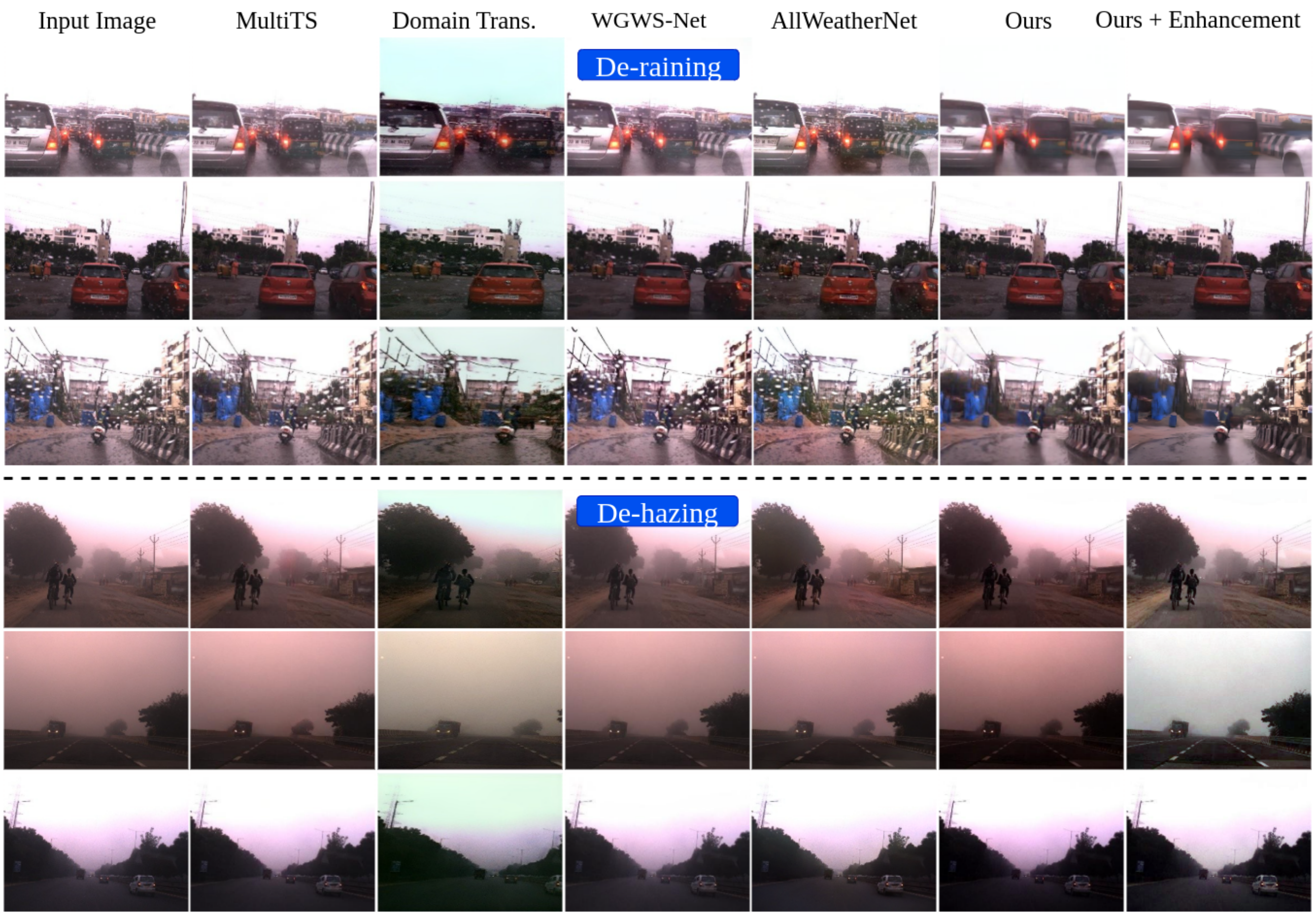

Figure 6: Qualitative comparison of real-world rain and haze images with low-light from IDD-AW [33] dataset.

NIQE 3.11 and BRISQUE 27.4, improving over the next best method WGWS-Net by 0.23 NIQE and 1.5 BRISQUE, and outperforming MultiTS and Domain Translation by more than 1 and 4 points, respectively. For haze, our NIQE score of 10.64 is similar to WGWS-Net, while our BRISQUE score of 27.1 remains competitive, indicating that our method achieves comparable perceptual quality despite more natural restoration with better contrast recovery.

In Fig. 6, qualitative comparisons on IDD-AW [33] dataset show that while baselines often oversmooth or leave residual artifacts, our method effectively recovers fine geometry details such as vehicle edges, traffic signs, and road textures, with segmentation outputs remaining robust to noise for accurate identification of drivable areas and obstacles. For challenging real-world cases where residual rain streaks or droplets persist after restoration, we employ Real-ESRGAN [41] to enhance local texture details and reduce blur. In hazy, low-light scenes, we integrate DarkIR [11], a Retinex-based illumination enhancer, to restore contrast and color fidelity. Both enhancement modules are state-of-the-art in their respective domains and further boost perceptual quality under adverse conditions.

Table 3: Quantitative comparison on real-world derained and dehazed images from IDD-AW [33] using NIQE and BRISQUE (lower is better). Best and second-best results are in bold and underlined, respectively.

Method	**Haze**		**Rain**	
	NIQE ↓	BRISQUE ↓	NIQE ↓	BRISQUE ↓
MultiTS [4]	14.78	33.5	4.51	32.4
DomainTranslation [30]	24.63	32.1	4.40	31.8
WGWS-Net [51]	**10.51**	**19.7**	<u>3.34</u>	28.9
AllWeatherNet [31]	11.73	28.6	3.98	**26.9**
Ours	<u>10.64</u>	<u>27.1</u>	**3.11**	<u>27.4</u>

4.4 Ablation Experiments

Effect of Two-Stage Knowledge Learning. We compare our two-stage distillation strategy with a single-stage baseline to validate its contribution. The two-stage approach achieved an average improvement of +0.8 dB in PSNR and +0.13 in SSIM, confirming that staged learning facilitates better feature transfer from teacher to student networks.

Effect of Segmentation-Aware Guidance. In Table 4, the ablation results on Rain1400 and RESIDE datasets shown that using all three losses, $\mathcal{L}_{\text{SCR}}$, $\mathcal{L}_{\text{HCR}}$, $\mathcal{L}_{\text{SEG}}$, together with both KC and KE

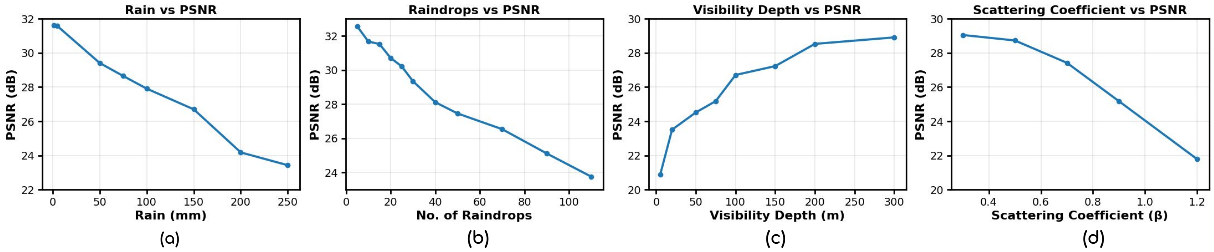

Figure 7: Illustration of performance variation of PSNR across different synthetic rain and haze weather conditions generated using JarvisIR [25] and RainRendering [36] simulations.

achieves the best performance, reaching 31.57/0.90 and 29.63/0.92 in PSNR/SSIM, respectively. Removing $\mathcal{L}_{\text{HCR}}$ reduces PSNR by 0.78 dB on Rain1400 and 0.85 dB on RESIDE, showing that hard contrastive learning strengthens instance discrimination. Omitting $\mathcal{L}_{\text{SEG}}$ results in an even larger drop of 1.68 dB and 1.44 dB, highlighting the importance of segmentation-aware supervision for complex scene restoration. Finally, retaining only $\mathcal{L}_{\text{SCR}}$ with KC produces the lowest scores of -4.25 dB and -3.61 dB compared to the full model, underscoring the complementary role of all components in our unified framework.

Impact of Weather Severity on Performance. In Fig. 7, PSNR variations are quantified for different synthetic weather severities generated using JarvisIR [25] and RainRendering [36]. For rain streaks (see Fig. 7a), PSNR decreases from 31.63 dB at light rain (1 mm) to 23.43 dB at heavy rain (250 mm), a drop of 8.2 dB. For raindrops (see Fig. 7b), increasing the number of droplets from 5 to 110 reduces PSNR from 32.56 dB to 23.76 dB (−8.8 dB). For haze (see Figs. 7c–d), increasing visibility depth from 5 m to 300 m improves PSNR from 20.87 dB to 28.91 dB, whereas raising the scattering coefficient from 0.3 to 1.2 decreases PSNR from 29.05 dB to 21.78 dB. These selected severities ensure consistent but non-trivial restoration difficulty, and the gradual PSNR decline across all scenarios demonstrates that our framework maintains stable performance and degrades gracefully under extreme weather conditions.

4.5 Model Efficiency and Deployment Feasibility

Model Efficiency. We compare our model's efficiency with recent multi-task weather removal baselines in terms of parameter count and inference time. As shown in Table 5, our model achieves the fastest inference with the lowest parameter count, while maintaining strong restoration performance. In contrast to heavier models like WGWS-Net [51] and AllWeatherNet [31], which rely on complex multi-branch and attention-based designs, our lightweight architecture is better suited for real-time and resource-limited settings.

Deployment Feasibility. To assess the model's feasibility for real-time applications, we conducted latency-focused benchmarks that simulated a video stream by processing video frames sequentially with a batch size of one. Using an NVIDIA RTX 3080 GPU, the fully optimized model, quantized to INT8 precision and compiled with TensorRT, achieved a processing speed of 49.81 FPS, surpassing the 30 FPS threshold for real-time video with a per-frame latency of 20.1 milliseconds. The Jetson Orin Nano demonstrated 6.15 FPS and 292.87 seconds of processing time after INT8 quantization and

Table 4: Ablation study on $\mathcal{L}_{\text{SCR}}$, $\mathcal{L}_{\text{SEG}}$, and $\mathcal{L}_{\text{HCR}}$. Best results are in bold.

Losses			Rain1400 [13]		RESIDE [19]	
$\mathcal{L}_{\text{SCR}}$	$\mathcal{L}_{\text{SEG}}$	$\mathcal{L}_{\text{HCR}}$	PSNR	SSIM	PSNR	SSIM
✓	✗	✗	27.32	0.76	26.02	0.79
✓	✗	✓	29.89	0.82	27.19	0.85
✓	✓	✗	30.79	0.87	28.78	0.88
✓	✓	✓	**31.57**	**0.90**	**29.63**	**0.92**

Table 5: Model efficiency comparison on a 256×256 image.

Metric	MultiTS [4]	WGWS-Net [51]	AllWeatherNet [31]	DomainTrans [30]	Ours
Params (M)	7.0	18.0	19.0	11.0	**7.0**
Inf. Time (ms)	16.6	34.0	38.0	25.5	**15.2**

TensorRT deployment, demonstrating suitable edge deployment. With acceptable performance up to 10–12 FPS achievable through further software optimization, or by upgrading to higher-end hardware (such as Jetson Orin NX, Orin AGX), target real-time FPS can be met, albeit with increased deployment cost. Importantly, this performance maintained visual quality, confirming the model's suitability for low-latency real-world applications. More details in Supplementary Material.

5 Conclusion

In this work, we propose a semantic-guided knowledge distillation framework for removal of adverse weather in safety-critical autonomous driving conditions. Our proposed model employs a semantic-guided dual-decoder architecture trained via a two-stage multi-teacher knowledge distillation strategy, where multiple heavyweight experts guide a lightweight student network. Through extensive experiments on diverse synthetic and real-world datasets, we demonstrate that our approach achieves competitive or superior restoration performance by generalising to real-world unseen driving environments while maintaining low latency and memory usage, suitable for deployment on edge devices. We believe that our framework paves the way for robust perception systems that ensure reliable operation in challenging outdoor environments by significantly improving visibility.

Acknowledgments

The project was supported by iHub-Data and Mobility at IIIT Hyderabad.

References

[1] Codruta O. Ancuti, Cosmin Ancuti, Radu Timofte, and Christophe De Vleeschouwer. 2018. O-HAZE: a dehazing benchmark with real hazy and haze-free outdoor images. In *CVPR NTIRE Workshop*.

[2] Kui Jia Chunmei Qing Bolun Cai, Xiangmin Xu and Dacheng Tao. 2016. DehazeNet: An End-to-End System for Single Image Haze Removal. *IEEE Transactions on Image Processing* (2016).

[3] Jiuchen Chen, Xinyu Yan, Qizhi Xu, and Kaiqi Li. 2025. Tokenize image patches: Global context fusion for effective haze removal in large images. In *CVPR*.

[4] Wei-Ting Chen, Zhi-Kai Huang, Cheng-Che Tsai, Hao-Hsiang Yang, Jian-Jiun Ding, and Sy-Yen Kuo. 2022. Learning Multiple Adverse Weather Removal via Two-stage Knowledge Learning and Multi-contrastive Regularization: Toward a Unified Model. *CVPR*.

[5] Zixuan Chen, Zewei He, and Zhe-Ming Lu. 2024. DEA-Net: Single image dehazing based on detail-enhanced convolution and content-guided attention. *IEEE transactions on image processing* (2024).

[6] Dan Chu. 2024. D-YOLO: Robust Framework for Object Detection in Adverse Weather Conditions. *CoRR* (2024).

[7] Guanglu Dong, Tianheng Zheng, Yuanzhouhan Cao, Linbo Qing, and Chao Ren. 2025. Channel Consistency Prior and Self-Reconstruction Strategy Based Unsupervised Image Deraining. In *CVPR*.

[8] Bohyung Kim et al. 2016. DehazeNet: An End-to-End System for Single Image Haze Removal. In *CVPR*.

[9] Bohyung Kim et al. 2017. AOD-Net: All-in-One Dehazing Network. In *ICCV*.

[10] Qin et al. 2020. FFA-Net: Feature Fusion Attention Network for Single Image Dehazing. In *AAAI*.

[11] Daniel Feijoo, Juan C. Benito, Alvaro Garcia, and Marcos V. Conde. 2025. DarkIR: Robust Low-Light Image Restoration. In *CVPR*.

[12] Alex Frants, Wei Liang, Arun Kumar, Meilin Zhao, and Justin Park. 2025. CMAWR-Net: Multiple Adverse Weather Removal via a Unified Quaternion Neural Architecture. *CoRR* (2025).

[13] Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John Paisley. 2017. Removing rain from single images via a deep detail network. In *CVPR*.

[14] Andreas Geiger, Philip Lenz, and Raquel Urtasun. 2012. Are we ready for autonomous driving The KITTI vision benchmark suite. In *CVPR*.

[15] Kui Jiang, Zhongyuan Wang, Peng Yi, Chen Chen, Baojin Huang, Yimin Luo, Jiayi Ma, and Junjun Jiang. 2020. Multi-scale progressive fusion network for single image deraining. In *CVPR*.

[16] Yifan Jiang et al. 2020. MSPFN: Multi-stage progressive fusion network for single image deraining. In *CVPR*.

[17] Yeying Jin, Xin Li, Jiadong Wang, Yan Zhang, and Malu Zhang. 2024. Raindrop Clarity: A Dual-Focused Dataset for Day and Night Raindrop Removal. In *ECCV*.

[18] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. 2017. Aod-net: All-in-one dehazing network. In *ICCV*.

[19] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. 2018. Benchmarking single-image dehazing and beyond. *IEEE transactions on image processing* (2018).

[20] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. 2018. Benchmarking single-image dehazing and beyond. *IEEE transactions on image processing* (2018).

[21] Ruoteng Li, Loong-Fah Cheong, and Robby T Tan. 2019. Heavy rain image restoration: Integrating physics model and conditional adversarial learning. In *CVPR*.

[22] Siyuan Li, Wenqi Ren, Jiawan Zhang, Jinke Yu, and Xiaojie Guo. 2019. Single image rain removal via a deep decomposition–composition network. *Computer Vision and Image Understanding* (2019).

[23] Xiang Li and Jianwu Li. 2025. All-in-one weather removal via Multi-Depth Gated Transformer with gradient modulation. *Pattern Recognit.* (2025).

[24] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. 2021. SwinIR: Image Restoration Using Swin Transformer. In *ICCVW*.

[25] Yunlong Lin, Zixu Lin, Haoyu Chen, Panwang Pan, Chenxin Li, Sixiang Chen, Kairun Wen, Yeying Jin, Wenbo Li, and Xinghao Ding. 2025. JarvisIR: Elevating Autonomous Driving Perception with Intelligent Image Restoration. In *CVPR*.

[26] Sachin Mehta and Mohammad Rastegari. 2022. MobileViT: Light-weight, General-purpose, and Mobile-friendly Vision Transformer. In *ICLR*.

[27] Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik. 2012. No-reference image quality assessment in the spatial domain. *IEEE Transactions on image processing* (2012).

[28] Anish Mittal, Rajiv Soundararajan, and Alan C Bovik. 2012. Making a "completely blind" image quality analyzer. *IEEE Signal processing letters* (2012).

[29] Ozan Özdenizci and Robert Legenstein. 2023. Restoring Vision in Adverse Weather Conditions With Patch-Based Denoising Diffusion Models. *IEEE Trans. Pattern Anal. Mach. Intell.* (2023).

[30] Prashant W. Patil, Sunil Gupta, Santu Rana, Svetha Venkatesh, and Subrahmanyam Murala. 2023. Multi-weather Image Restoration via Domain Translation. In *ICCV*.

[31] Chenghao Qian, Mahdi Rezaei, Saeed Anwar, Wenjing Li, Tanveer Hussain, Mohsen Azarmi, and Wei Wang. 2025. AllWeather-Net: Unified Image Enhancement for Autonomous Driving Under Adverse Weather and Low-Light Conditions. In *International Conference on Pattern Recognition*.

[32] Ruijie Quan, Xin Yu, Yuanzhi Liang, and Yi Yang. 2021. Removing raindrops and rain streaks in one go. In *CVPR*.

[33] Furqan Ahmed Shaik, Abhishek Reddy Malreddy, Nikhil Reddy Billa, Kunal Chaudhary, Sunny Manchanda, and Girish Varma. 2024. IDD-AW: A Benchmark for Safe and Robust Segmentation of Drive Scenes in Unstructured Traffic and Adverse Weather. In *WACV*.

[34] Hao Shen, Zhong-Qiu Zhao, Yulun Zhang, and Zhao Zhang. 2023. Mutual information-driven triple interaction network for efficient image dehazing. In *International conference on multimedia*.

[35] Fuxiang Tan, Yurong Qian, Yuting Kong, Hao Zhang, Daxin Zhou, Yingying Fan, Long Chen, and Zhengqing Xiao. 2022. DBSwin: Transformer based dual branch network for single image deraining. *J. Intell. Fuzzy Syst.* (2022).

[36] Maxime Tremblay, Shirsendu S. Halder, Raoul de Charette, and Jean-François Lalonde. 2020. Rain Rendering for Evaluating and Improving Robustness to Bad Weather. *International Journal of Computer Vision* (2020).

[37] Jeya Maria Jose Valanarasu et al. 2022. TransWeather: Transformer-based restoration of images degraded by multiple weather conditions. In *CVPR*.

[38] Han Wang et al. 2022. WeatherNAS: Neural architecture search for multi-degradation weather restoration. *NeurIPS* (2022).

[39] Hong Wang, Qi Xie, Qian Zhao, and Deyu Meng. 2020. A model-driven deep neural network for single image rain removal. In *CVPR*.

[40] Tianyu Wang, Xin Yang, Ke Xu, Shaozhe Chen, Qiang Zhang, and Rynson W.H. Lau. 2019. Spatial Attentive Single-Image Deraining with a High Quality Real Rain Dataset. In *CVPR*.

[41] Xintao Wang, Liangbin Xie, Chao Dong, and Ying Shan. [n.d.]. Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data. In *International Conference on Computer Vision Workshops (ICCVW)* (2021).

[42] Haiyan Wu, Yanyun Qu, Shaohui Lin, Jian Zhou, Ruizhi Qiao, Zhizhong Zhang, Yuan Xie, and Lizhuang Ma. 2021. Contrastive learning for compact single image dehazing. In *CVPR*.

[43] Weihao Xia, Zhanglin Cheng, Yujiu Yang, and Jing-Hao Xue. 2020. Cooperative Semantic Segmentation and Image Restoration in Adverse Environmental Conditions. *CoRR* (2020).

[44] Yunyang Xiong et al. 2021. Nyströmformer: A Nyström-based algorithm for approximating self-attention. *AAAI* (2021).

[45] Shugo Yamashita and Masaaki Ikehara. 2024. Image Deraining with Frequency-Enhanced State Space Model. *CoRR* (2024).

[46] Syed Waqas Zamir et al. 2022. Restormer: Efficient transformer for high-resolution image restoration. In *CVPR*.

[47] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. 2021. Multi-stage progressive image restoration. In *CVPR*.

[48] H. Zhang and V. Patel. 2024. Rain100H: A Benchmark Dataset for Heavy Rain Scenes. *IEEE Transactions on Image Processing* (2024).

[49] Xinyi Zhang, Hang Dong, Jinshan Pan, Chao Zhu, Ying Tai, Chengjie Wang, Jilin Li, Feiyue Huang, and Fei Wang. 2021. Learning To Restore Hazy Video: A New Real-World Dataset and a New Method. In *CVPR*.

[50] Ruoxi Zhu, Zhengzhong Tu, Jiaming Liu, Alan C. Bovik, and Yibo Fan. 2024. MWFormer: Multi-Weather Image Restoration Using Degradation-Aware Transformers. *IEEE Transactions on Image Processing* (2024).

[51] Yurui Zhu, Tianyu Wang, Xueyang Fu, Xuanyu Yang, Xin Guo, Jifeng Dai, Yu Qiao, and Xiaowei Hu. 2023. Learning Weather-General and Weather-Specific Features for Image Restoration Under Multiple Adverse Weather Conditions. In *CVPR*.