applied sciences

MDPI

Article

LCVR-Net: Dual-Attention Visibility Restoration for Traffic Surveillance Under Dust and Fog Degradation

Nuruddin Md [1] and Hosang Lee [2,*]

[1] Department of Global IT Engineering, Kyungsung University, Busan 48434, Republic of Korea; nuruddin.ksu@gmail.com
[2] Department of Computer Engineering, Kyungnam University, Changwon-si 51767, Republic of Korea
* Correspondence: hslee7092@kyungnam.ac.kr

Abstract

Adverse atmospheric conditions, particularly dust and fog, substantially degrade the visibility of traffic surveillance imagery, limiting the reliability of intelligent transportation systems and vision-based traffic monitoring applications. To address these limitations, this paper proposes the Lightweight CCTV Visibility Restoration Network (LCVR-Net), an efficient image restoration framework specifically designed for CCTV surveillance under adverse weather conditions. The proposed architecture adopts a lightweight encoder–decoder backbone based on Residual Depthwise Separable Blocks (RDSBs) to achieve effective feature extraction with low computational complexity. Furthermore, a Dual Attention Refinement Module (DARM) is introduced to enhance degradation-aware feature representation for fog restoration, while a lightweight Color Correction Head (CCH) is incorporated to compensate for atmospheric color distortion and improve perceptual image fidelity. To enable task-specific optimization, two model variants are developed for dust and fog restoration, respectively. Qualitative evaluations on real-world surveillance images further provide supporting evidence of the practical applicability of the proposed framework under naturally occurring atmospheric degradation. These results demonstrate that LCVR-Net provides an effective balance between restoration accuracy and computational efficiency, making it well suited for practical deployment in intelligent transportation and traffic surveillance systems.

Keywords: convolutional neural networks; image restoration; intelligent transportation systems; lightweight neural networks; traffic surveillance; visibility enhancement

Academic Editor: Thomas Lindner

Received: 24 August 2026
Revised: 14 September 2026
Accepted: 15 September 2026
Published: 18 September 2026

1. Introduction

Intelligent Transportation Systems (ITS) have become a fundamental component of modern smart cities, enabling traffic monitoring, incident detection, congestion management, and intelligent transportation services through continuous visual perception [1–3]. Among various sensing modalities, closed-circuit television (CCTV) cameras remain the most widely deployed because they provide extensive spatial coverage, low deployment cost, and continuous monitoring of complex traffic environments. As intelligent surveillance systems increasingly incorporate computer vision and deep learning algorithms, the quality of captured images has become a critical factor influencing the reliability of both human operators and automated traffic analysis. However, outdoor surveillance cameras frequently operate under adverse atmospheric conditions that severely degrade image quality. Dust, fog, and haze scatter and absorb light before it reaches the imaging sensor,

resulting in reduced contrast, color distortion, blurred scene structures, and loss of fine details [4–6]. Such degradations significantly impair downstream vision tasks, including vehicle detection, traffic flow estimation, license plate recognition, object tracking, and incident analysis, thereby limiting the effectiveness of ITS in real-world environments. Among these degradations, dust and fog are particularly challenging because they obscure distant objects and critical road infrastructure while exhibiting complex and spatially varying scattering characteristics. Representative examples of dust- and fog-degraded CCTV images are presented in Figure 1.

Figure 1. Examples of dust- and fog-degraded CCTV surveillance images.

Figure 1 illustrates representative CCTV scenes affected by fog and dust, highlighting their distinct degradation characteristics. Fog causes severe visibility and contrast reduction, whereas dust introduces strong atmospheric attenuation and color distortion. These conditions can significantly impair visual perception and subsequent traffic surveillance tasks.

Early visibility restoration techniques relied on handcrafted image priors and enhancement strategies, including Histogram Equalization (HE), Contrast Limited Adaptive Histogram Equalization (CLAHE), Retinex-based enhancement, Gamma Correction, and the Dark Channel Prior (DCP) [7–10]. Although computationally efficient, these approaches generally depend on fixed assumptions regarding image statistics or atmospheric models, making them susceptible to over-enhancement, color inconsistencies, noise amplification, and poor generalization across diverse surveillance environments. The rapid advancement of deep learning has substantially improved single-image visibility restoration by enabling data-driven feature representation. Representative convolutional models, including DehazeNet [11], AOD-Net [12], FFA-Net [13], AECR-Net [14], and more recent lightweight restoration networks, have achieved remarkable performance on standard dehazing benchmarks. Meanwhile, transformer- and frequency-based restoration frameworks have further advanced restoration quality by modeling long-range dependencies and degradation-aware frequency representations [15,16]. Despite these advances, three important limitations remain. First, most existing methods are developed and evaluated on generic dehazing datasets that differ significantly from real CCTV surveillance scenes in viewpoint, illumination, camera characteristics, and degradation distribution [17]. Second, state-of-the-art restoration networks often achieve higher reconstruction accuracy by substantially increasing model complexity, computational cost, and memory consumption, limiting their applicability to resource-constrained edge surveillance systems. Third, relatively few studies have investigated lightweight architectures specifically designed to restore both dust- and fog-degraded CCTV imagery while maintaining inference capability.

To overcome these limitations, this paper proposes the Lightweight CCTV Visibility Restoration Network (LCVR-Net), a unified and computationally efficient framework tailored for traffic surveillance applications. The proposed architecture adopts lightweight Residual Depthwise Separable Blocks (RDSBs) to achieve efficient hierarchical feature extraction with minimal computational overhead. Furthermore, a Dual Attention Refinement Module (DARM) is introduced to strengthen degradation-aware feature refinement under dense fog, while a lightweight Color Correction Head (CCH) compensates for atmospheric

color distortion and enhances perceptual fidelity. Based on the same backbone architecture, two specialized variants are developed: LCVR-Net-D for dust restoration and LCVR-Net-F for fog restoration. To facilitate practical evaluation, a CCTV-oriented benchmark is further constructed using real traffic surveillance images with synthetically generated dust and fog degradations, complemented by qualitative validation on real-world surveillance scenes. Although LCVR-Net is primarily developed for traffic surveillance, its applicability is not limited to CCTV imagery. The fog-restoration model is also evaluated on general outdoor scenes, to examine its applicability beyond the surveillance domain.

The remainder of this paper is organized as follows. Section 2 reviews related work on image visibility restoration. Section 3 presents the proposed LCVR-Net architecture. Section 4 describes the datasets, implementation details, and experimental settings. Section 5 reports quantitative and qualitative comparisons together with ablation and computational efficiency analyses. Finally, Section 6 concludes the paper.

2. Related Works

2.1. Model-Based Visibility Restoration

Early visibility restoration methods primarily relied on image enhancement techniques and atmospheric scattering models to improve degraded images. Conventional enhancement approaches, including Histogram Equalization (HE), Contrast Limited Adaptive Histogram Equalization (CLAHE), Gamma Correction, and Retinex-based enhancement, improve visual quality by manipulating image intensity distributions or local illumination [7–9]. Although computationally efficient, these methods do not explicitly model atmospheric degradation and therefore often produce over-enhanced images, amplified noise, and inconsistent color reproduction when applied to complex outdoor scenes.

Physics-based restoration methods estimate the degradation process using atmospheric scattering theory. Among these, the Dark Channel Prior (DCP) proposed by He et al. [6] remains one of the most influential single-image dehazing methods. DCP estimates scene transmission from statistical priors and reconstructs the latent clean image using the atmospheric scattering model. Guided filtering was subsequently introduced to refine transmission estimation while preserving edge structures [10]. Other prior-based approaches, including non-local dehazing [18], further improved restoration accuracy by exploiting structural similarities across image regions. Despite their interpretability, model-based methods rely heavily on handcrafted assumptions that frequently fail under dense dust, non-uniform illumination, and complex surveillance environments, leading to halo artifacts, color distortion, and incomplete visibility recovery. Beyond DCP, several prior-based approaches have been developed to improve transmission estimation and scene radiance recovery. Fattal introduced a color-lines prior that exploits the distribution of local image patches in RGB space to estimate scene transmission [19]. Meng et al. incorporated a boundary constraint with contextual regularization to obtain more reliable transmission maps while preserving image details [20]. Zhu et al. proposed the Color Attenuation Prior (CAP), which estimates scene depth from the relationship between brightness and saturation and enables efficient single-image haze removal [21].

2.2. Deep Learning-Based Visibility Restoration

Deep learning has significantly advanced image visibility restoration by learning degradation-aware feature representations directly from large-scale datasets. Unlike traditional methods, convolutional neural networks (CNNs) learn an end-to-end mapping between degraded and clean images without explicitly estimating physical model parameters. DehazeNet [11] pioneered CNN-based single-image dehazing by learning transmission estimation from data, while All –in-One Dehazing Network (AOD-Net) [12] reformulated

the atmospheric scattering model into a lightweight end-to-end framework with improved computational efficiency. Subsequent studies focused on enhancing feature representation through multi-scale learning and attention mechanisms. Feature Fusion Attention Network for single image dehazing network (FFA-Net) [13] introduced feature fusion attention to capture both channel-wise and pixel-level contextual information, whereas Auto-Encoder and Contrastive Regularization Network (AECR-Net) [14] incorporated contrastive learning to improve structural consistency and perceptual restoration. More recently, transformer-based architectures have become increasingly popular because of their ability to model long-range dependencies. Representative methods, including Restormer [15], demonstrate excellent restoration performance by combining self-attention with efficient feature aggregation. In parallel, frequency-domain learning has emerged as an effective strategy for separating degradation components from structural image information, enabling more accurate restoration of complex atmospheric degradations. Comprehensive surveys further highlight the growing importance of transformer- and frequency-based restoration frameworks for adverse-weather image enhancement [22]. Although these methods achieve impressive reconstruction quality, their superior performance is generally accompanied by increased computational complexity, memory consumption, and inference latency, limiting their applicability to surveillance and resource-constrained edge devices.

Beyond these representative CNN-based approaches, several architectures have explored joint estimation, multi-scale representation, and feature fusion for improved dehazing. Densely Connected Pyramid Dehazing Network (DCPDN) jointly estimates the transmission map, atmospheric light, and restored image within a densely connected end-to-end framework [23]. Gated Fusion Network (GFN) learns pixel-wise confidence maps to adaptively fuse multiple enhanced representations of the hazy input [24]. GridDehazeNet employs an attention-based multi-scale grid architecture to facilitate information exchange across different feature resolutions without explicitly relying on the atmospheric scattering model [25]. Multi-Scale Boosted Dehazing Network (MSBDN) further introduces multi-scale boosted decoding and dense feature fusion to progressively recover spatial details and improve haze removal [26]. These developments demonstrate the effectiveness of hierarchical and multi-scale feature learning for visibility restoration, although increased architectural complexity can impose additional computational requirements. Gao et al. [27] enhanced degraded sandstorm images using a low-visibility enhancement network based on multiscale channel attention features. To address degraded hazy images, Jia et al. [28] proposed a semi-supervised neural network using disentangled meta-knowledge, yielding high-quality results. Furthermore, Qadir et al. [29] restored distorted sand-dust images using an attention-driven multiscale feature fusion method.

2.3. Lightweight Visibility Restoration and Research Gap

With the rapid deployment of intelligent transportation infrastructure, lightweight restoration networks have received increasing attention. Recent approaches, such as Dehaze-UNet [16], employ simplified encoder-decoder architectures to balance restoration performance and computational efficiency, while frequency-selection-based restoration networks [17] improve feature representation by adaptively emphasizing informative frequency components. These studies demonstrate that lightweight architectures can achieve competitive restoration quality while substantially reducing computational cost.

Nevertheless, several important challenges remain. First, most existing restoration methods are developed and evaluated using generic dehazing benchmarks such as RESIDE [30], OTS [30], and SOTS [30], which differ considerably from practical CCTV surveillance environments in camera viewpoint, scene geometry, traffic density, and degradation characteristics. Second, many lightweight networks primarily target general haze removal

rather than surveillance-oriented restoration under multiple atmospheric degradations. Third, existing approaches generally optimize restoration quality or computational efficiency independently, making it difficult to simultaneously achieve high restoration accuracy, low model complexity, and deployment capability for intelligent transportation systems [30,31]. Motivated by these limitations, this work proposes the Lightweight CCTV Visibility Restoration Network (LCVR-Net), a surveillance-oriented restoration framework specifically designed for dust- and fog-degraded CCTV imagery. The proposed network integrates lightweight Residual Depthwise Separable Blocks (RDSBs) for efficient feature extraction, a Dual Attention Refinement Module (DARM) for degradation-aware feature refinement, and a lightweight Color Correction Head (CCH) for atmospheric color compensation. In addition, a CCTV-oriented benchmark is constructed to facilitate systematic evaluation under realistic traffic surveillance conditions. By jointly considering restoration quality, computational efficiency, and practical deployment requirements, LCVR-Net addresses an important gap between existing image restoration research and real-world intelligent transportation applications.

More recently, DehazeFormer adapts vision transformers specifically to single-image dehazing through modified normalization, activation, and spatial feature aggregation mechanisms, achieving strong restoration performance with improved computational efficiency [32]. Multi-Branch (MB)-TaylorFormer further approximates self-attention using Taylor expansion and incorporates multi-scale refinement to capture long-range dependencies with reduced computational complexity [33]. MixDehazeNet combines multi-scale parallel large-kernel convolutions with enhanced parallel attention to enlarge the effective receptive field while avoiding the computational burden of conventional transformer architectures [34].

The major contributions of this work are summarized as follows.

1. A lightweight surveillance-oriented restoration framework, LCVR-Net, is proposed for efficient visibility enhancement of dust- and fog-degraded CCTV images while maintaining computational performance.
2. Degradation-specific refinement modules are introduced within the lightweight framework. The Dual Attention Refinement Module (DARM) enhances degradation-aware feature representation under dense fog, while the lightweight Color Correction Head (CCH) compensates for atmospheric color distortion with minimal computational overhead.
3. A CCTV-oriented benchmark containing synthetic dust and fog degradations is constructed to enable systematic evaluation under realistic intelligent transportation scenarios.
4. Extensive quantitative, qualitative, cross-dataset, ablation, and computational analyses demonstrate competitive restoration performance with fewer than 0.4 M trainable parameters, supporting the practical applicability of LCVR-Net for traffic surveillance.

3. Proposed Method

The proposed LCVR-Net is motivated by the need for accurate yet computationally efficient visibility restoration in CCTV surveillance. Existing restoration networks often prioritize reconstruction quality at the expense of computational complexity, while lightweight models may struggle with degradation-specific features and atmospheric color distortion. To address these limitations, LCVR-Net combines efficient feature extraction, degradation-aware attention, and lightweight color correction within a unified framework. Its task-specific configurations further accommodate the distinct characteristics of dust and fog while maintaining low computational overhead.

3.1. Overall Architecture of LCVR-NET

The proposed Lightweight CCTV Visibility Restoration Network (LCVR-Net) is designed to restore dust- and fog-degraded traffic surveillance images while maintaining high computational efficiency for deployment. As illustrated in Figure 2, LCVR-Net adopts a lightweight encoder–bottleneck–decoder architecture [35] with skip connections, global residual learning [36], and a lightweight Color Correction Head (CCH).

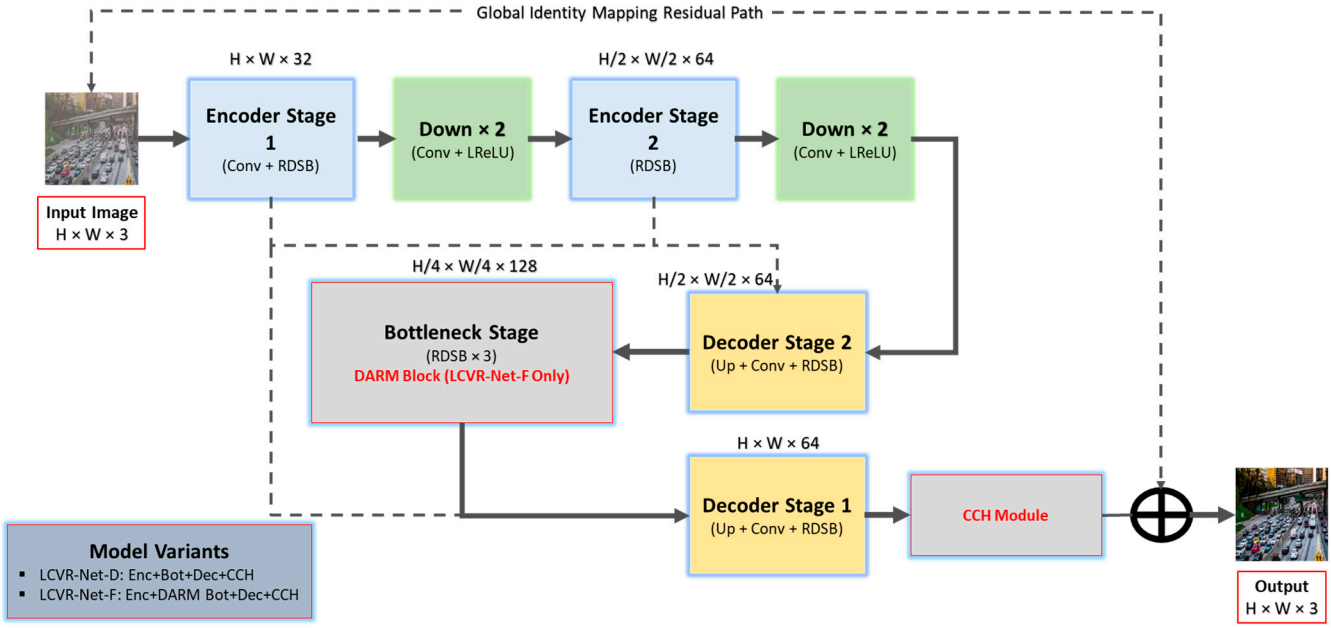

Figure 2. Overall architecture of the proposed LCVR-Net framework.

Given a degraded input image ($I_d \in \mathbb{R}^{H\times W\times 3}$), the network learns a nonlinear mapping to reconstruct the restored image (I_r). The encoder progressively extracts hierarchical multi-scale features using Residual Depthwise Separable Blocks (RDSBs) and downsampling operations. These features are refined in the bottleneck before being reconstructed by the decoder through bilinear upsampling and skip connections, which preserve fine structural information while facilitating efficient feature propagation. To accommodate different atmospheric degradations, two network variants are developed. LCVR-Net-D employs the lightweight backbone for dust restoration, whereas LCVR-Net-F incorporates a Dual Attention Refinement Module (DARM) within the bottleneck to enhance feature representation under severe fog degradation. The restored decoder features are subsequently processed by the proposed Color Correction Head (CCH) to compensate for atmospheric color distortion and improve perceptual consistency. Detailed descriptions of RDSB, DARM, and CCH are provided in the following subsections. Rather than directly predicting the restored image, LCVR-Net adopts global residual learning [36], allowing the network to estimate only the degradation residual. The final output is expressed as:

$$I_r = I_d + R(F_d) + C(F_d), \tag{1}$$

where I_d and I_r denote the degraded and restored images, respectively, $R(\cdot)$ represents the residual reconstruction branch, $C(\cdot)$ denotes the Color Correction Head, and F_d represents the decoder feature representation.

3.2. Residual Depthwise Separable Block (RDSB)

To achieve an effective balance between restoration accuracy and computational efficiency, LCVR-Net employs the Residual Depthwise Separable Block (RDSB) as its

fundamental feature extraction unit. Conventional convolution simultaneously performs spatial filtering and channel mixing, resulting in considerable computational overhead that limits deployment on resource-constrained surveillance systems. To address this limitation, RDSB adopts depthwise separable convolutions, which decompose feature extraction into independent spatial filtering and channel fusion operations, thereby substantially reducing computational complexity while preserving representative feature learning [37,38]. Given an input feature map ($F_{\{in\}}$), the depthwise convolution extracts channel-wise spatial features as:

$$F_d = DW\left(F_{\{in\}}\right), \tag{2}$$

where $F_{\{in\}}$ denotes the input feature tensor, $DW(\cdot)$ represents the depthwise convolution operation, and F_d denotes the resulting depthwise-filtered feature tensor. The resulting features are subsequently fused through a pointwise (1×1) convolution to model inter-channel correlations:

$$F_p = PW(F_d), \tag{3}$$

where $PW(\cdot)$ denotes the pointwise (1×1) convolution operation, and F_p represents the resulting pointwise-fused feature tensor after inter-channel information aggregation.

The proposed RDSB consists of two consecutive depthwise separable convolution layers, each followed by a LeakyReLU activation [39]. A residual connection is introduced to facilitate gradient propagation and preserve low-level image information during feature transformation. The block output is defined as:

$$F_r = F_{\{in\}} + H\left(F_{\{in\}}\right), \tag{4}$$

where $H(\cdot)$ denotes the stacked depthwise separable convolution operations, and F_r represents the residual output feature tensor obtained by combining the transformed features with the input $F_{\{in\}}$. By combining lightweight convolution with residual learning, RDSB effectively captures degradation-related features while maintaining stable optimization. The reduced computational cost enables multiple RDSBs to be employed throughout the encoder, bottleneck, and decoder without significantly increasing model complexity, making LCVR-Net suitable for traffic surveillance applications.

3.3. Dual Attention Refinement Module (DARM)

Although the lightweight backbone effectively extracts hierarchical features, restoring dense fog remains challenging because atmospheric scattering suppresses discriminative image information over large spatial regions. Consequently, feature responses associated with distant objects and fine scene structures become progressively weaker. To enhance degradation-aware representation under such conditions, a Dual Attention Refinement Module (DARM) is introduced in the bottleneck stage of LCVR-Net-F, as illustrated in Figure 3. DARM sequentially integrates channel attention and spatial attention [40,41] to refine bottleneck features from complementary perspectives. Given the bottleneck feature map ($F_{\{in\}}$), the Channel Attention Unit (CAU) first estimates the relative importance of each feature channel by aggregating global contextual information through Global Average Pooling (GAP), followed by two (1×1) convolution layers and a sigmoid activation [42]. Channel refinement is formulated as:

$$F_c = F_b \otimes M_c, \tag{5}$$

where M_c denotes the channel attention map and $\otimes$ represents element-wise multiplication. The channel-refined features are subsequently processed by the Spatial Attention Unit (SAU) to identify informative spatial regions. Average-pooling and max-pooling are

first performed along the channel dimension, after which the resulting feature maps are concatenated and passed through a 7 × 7convolution layer followed by a sigmoid activation [42]. The spatial refinement process is expressed as:

$$F_s = F_c \otimes M_s \tag{6}$$

where M_s denotes the spatial attention map. To further enhance feature representation, the attention-refined features are processed by an Attention Refinement Unit (ARU) comprising two consecutive RDSBs. This operation is formulated as:

$$F_r = ARU(F_s), \tag{7}$$

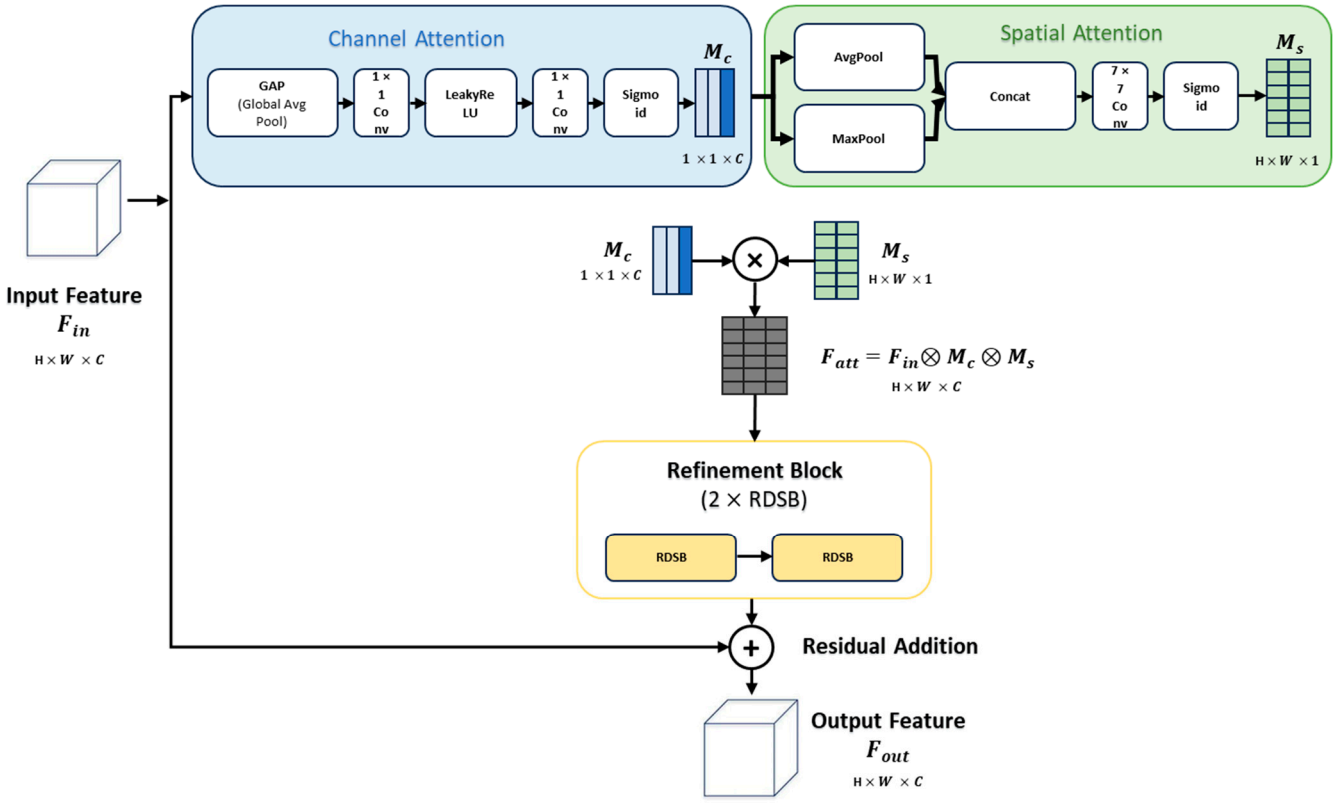

Figure 3. Structure of the proposed Dual Attention Refinement Module (DARM).

Finally, a residual connection is employed to preserve the original bottleneck representation and improve optimization stability. The output of DARM is obtained as:

$$F_{\{DARM\}} = F_b + F_r, \tag{8}$$

where $F_{\{DARM\}}$ denotes the refined bottleneck feature representation.

3.4. Color Correction Head (CCH)

Atmospheric degradation affects not only scene visibility but also color fidelity. Dust commonly introduces yellow or brown color casts, whereas fog reduces saturation and produces illumination inconsistencies through wavelength-dependent scattering. Consequently, recovering structural information alone may yield restored images with residual chromatic distortion. To address this problem, LCVR-Net incorporates a lightweight Color Correction Head (CCH) that jointly models global color bias and spatially varying appearance information.

As illustrated in Figure 4, CCH comprises two complementary branches: a Global Color Branch (GCB) and a Local Refinement Branch (LRB). Given the decoder feature representation FD, where FD denotes the decoder output feature tensor provided as input to the CCH, the GCB aggregates scene-level information using Global Average Pooling (GAP). The pooled representation is subsequently processed by two 1×1 convolution layers separated by a LeakyReLU activation [39], followed by a hyperbolic tangent function that constrains the estimated global correction. The global color representation is computed as:

$$F_g = GCB(F_d), \tag{9}$$

where F_g denotes the global color correction feature map and $GCB(\cdot)$ represents the global color estimation operation.

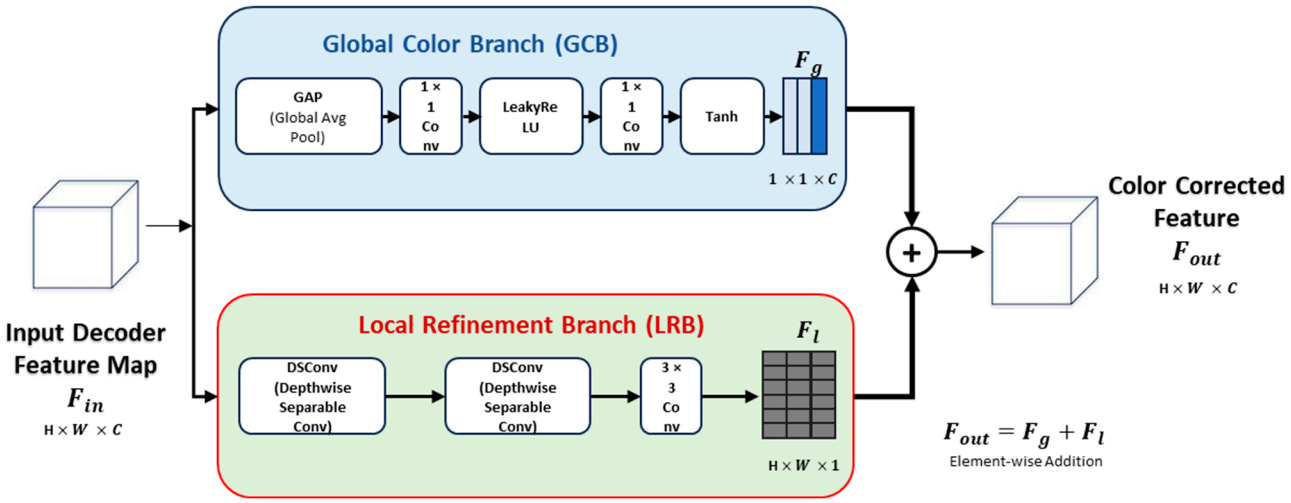

Figure 4. Architecture of the proposed Color Correction Head (CCH).

Although the GCB captures the dominant scene-level color shift, global statistics alone cannot adequately model non-uniform degradation across the image. The LRB therefore performs spatially adaptive correction using two depthwise separable convolution layers followed by a 3×3 convolution. This design preserves local structures and appearance variations while introducing minimal computational overhead. The locally refined representation is expressed as:

$$F_l = LRB(F_d), \tag{10}$$

where F_l denotes the locally refined feature representation. The global and local representations are fused through element-wise addition:

$$F_c = F_g + F_l, \tag{11}$$

where F_c represents the fused color correction feature representation. The resulting feature is integrated with the residual reconstruction branch to obtain the restored image:

$$I_r = I_d + R(F_d) + F_c, \tag{12}$$

where I_d and I_r denote the degraded and restored images, respectively, and $R(\cdot)$ represents the residual reconstruction branch. By combining scene-level chromatic estimation with spatially adaptive refinement, CCH corrects global color shifts without sacrificing local details. Its lightweight implementation introduces only a marginal increase in model complexity, making it suitable for both LCVR-Net variants and surveillance deployment.

3.5. Task-Specific Network Configurations

Dust and fog exhibit different degradation characteristics and therefore require different levels of feature refinement. To address these differences without constructing entirely separate architectures, two task-specific variants are derived from the same lightweight encoder–decoder backbone. Both variants employ RDSBs and the proposed CCH, whereas DARM is selectively introduced for fog restoration.

3.5.1. LCVR-NET-D for Dust Restoration

Dust degradation is primarily characterized by contrast reduction, partial visibility loss, and pronounced color shifts caused by suspended particles. Because these effects can be effectively addressed through hierarchical feature reconstruction and explicit color correction, LCVR-Net-D employs the shared encoder–decoder backbone together with CCH, while the bottleneck consists only of stacked RDSBs. The dust-restoration configuration is expressed in Figure 3. This configuration avoids additional attention operations, thereby preserving minimal parameter count and computational cost. LCVR-Net-D consequently prioritizes efficient structural recovery and chromatic correction for dust-degraded CCTV imagery.

3.5.2. LCVR-NET-F for Fog Restoration

Fog produces stronger and more spatially extensive visibility attenuation because of atmospheric scattering. Distant objects, road boundaries, and low-contrast structures are therefore more difficult to recover using the lightweight backbone alone. To strengthen bottleneck representation under these conditions, LCVR-Net-F incorporates DARM before image reconstruction. The fog-restoration configuration is formulated in Figure 4. DARM selectively enhances informative channel and spatial responses before the refined features are passed to the decoder. Compared with LCVR-Net-D, this modification introduces only a limited increase in parameters while providing additional representation capacity for severe fog degradation. Both variants retain the same overall design and share most architectural components, simplifying implementation and deployment across different surveillance conditions. LCVR-Net-D emphasizes low-cost restoration and color compensation for dust, whereas LCVR-Net-F provides additional attention-guided refinement for the stronger scattering effects associated with fog.

3.5.3. Loss Function

The proposed LCVR-Net is trained using a hybrid loss function that jointly optimizes pixel-wise reconstruction accuracy and structural similarity. Specifically, the total training loss combines the L1 loss with the Structural Similarity Index Measure (SSIM) [38] loss, enabling the network to preserve both image fidelity and perceptual quality. The pixel-wise reconstruction loss is defined as

$$L_{L1} = \left(\frac{1}{N}\right)\sum_{i=1}^{N}|I_R - I_G|, \tag{13}$$

where L_{L1} denotes the L_1 pixel-wise reconstruction loss, I_R and I_G denote the restored and ground-truth images, respectively, $|\cdot|$ means absolute value operation, and N represents the total number of pixels. To further preserve structural information, the SSIM [43] loss is formulated as:

$$SSIM(I_r, I_g) = \frac{(2\mu_r\mu_g + C_1)(2\sigma_{rg} + C_2)}{(\mu_r^2 + \mu_g^2 + C_1)(\sigma_r^2 + \sigma_g^2 + C_2)}, \tag{14}$$

where SSIM$(\cdot)$ measures the structural similarity between the restored and reference images. The overall optimization objective is defined as:

$$L_{SSIM} = 1 - \mathrm{SSIM}(I_R, I_G), \tag{15}$$

Minimizing L_{SSIM} encourages preservation of structural information and perceptually meaningful image content.

$$L_{total} = L_{L1} + \lambda L_{SSIM}, \tag{16}$$

where λ is the weighting factor that balances the contributions of the two loss terms, L_{total} means total loss. The L_1 loss encourages accurate pixel-level reconstruction while reducing sensitivity to outliers, whereas the SSIM [43] loss preserves structural consistency and perceptual image quality. Their complementary characteristics enable LCVR-Net to recover fine details, suppress atmospheric degradation, and produce visually consistent restoration results without introducing additional computational complexity during inference.

4. Experimental Setup

This section describes the experimental framework used to evaluate the proposed LCVR-Net. The evaluation includes dataset preparation, implementation and training settings, quantitative and qualitative assessment, computational efficiency analysis, and ablation studies. Experiments are conducted for both dust and fog restoration to assess restoration performance, convergence behavior, computational efficiency, and generalization under CCTV surveillance conditions.

4.1. Dataset Preparation

To evaluate the proposed LCVR-Net under realistic traffic surveillance scenarios, a CCTV-oriented benchmark was constructed using publicly available surveillance imagery. Unlike conventional dehazing datasets that mainly contain natural outdoor scenes, the proposed benchmark focuses on traffic surveillance environments where visibility degradation directly affects intelligent transportation applications.

4.1.1. CCTV Traffic Surveillance Dataset

Clean surveillance images were collected from the Traffic Data from Surveillance Cameras dataset [44], which contains diverse traffic scenes captured by fixed CCTV cameras under varying illumination, viewpoints, and traffic conditions. After removing unsuitable samples, 4234 images were retained and resized to 256 $\times$ 256 pixels. The dataset was divided into 3598 training images, 536 validation images, and 100 evaluation images. The evaluation set was completely isolated from training and validation and was used exclusively for performance assessment.

The source dataset does not provide explicit camera identifiers for all images; therefore, a strictly camera-disjoint partition could not be established. Consequently, although the evaluation subset is excluded from training and validation, some similarity in scene or camera viewpoint across the image-level partitions cannot be completely ruled out. To complement this evaluation and reduce reliance on the internal CCTV split, cross-dataset testing was additionally performed on the RESIDE SOTS Outdoor benchmark [30], whose scenes are independent of the CCTV training data. This provides an additional assessment of generalization to previously unseen outdoor scenes.

Publicly available traffic-surveillance datasets with paired clean and atmospherically degraded images remain limited. Therefore, the Surveillance Cameras dataset [44] was used as the primary traffic-surveillance benchmark, as its clean CCTV images enable controlled generation of paired dust- and fog-degraded samples for full-reference evaluation. To complement

this dataset and assess performance beyond the primary benchmark, additional cross-dataset evaluation was conducted on RESIDE SOTS Outdoor [30], while naturally degraded images from the DS Dataset [45] were used for real-world qualitative evaluation.

4.1.2. Synthetic Dust and Fog Datasets

To enable supervised learning, paired degraded images were synthetically generated from the clean Surveillance Cameras dataset [44]. Dust degradation was simulated by introducing atmospheric dust veils, visibility attenuation, contrast reduction, and color shifts, while fog degradation was generated using an atmospheric scattering model to reproduce realistic visibility loss and contrast degradation. The resulting paired datasets were used to train and evaluate LCVR-Net-D and LCVR-Net-F, respectively.

4.1.3. Real World Evaluation Dataset

To assess generalization under practical conditions, additional qualitative evaluations were conducted using naturally degraded images from the DS Dataset [41]. Which contains real-world images degraded by dust and fog under diverse outdoor conditions. Unlike the synthetically generated paired CCTV data used for quantitative evaluation, the DS images represent naturally occurring atmospheric degradation and do not provide corresponding clean ground-truth images. Therefore, selected dust- and fog-degraded samples from this dataset were used exclusively for qualitative evaluation to assess the generalization capability of the proposed LCVR-Net under real-world conditions. Only the fog- and dust-degraded samples were used. Since no corresponding ground-truth images are available, these data were employed exclusively for qualitative comparison and visual analysis.

Since the SOTS Outdoor images originate from a dataset independent of the CCTV training set, this cross-dataset evaluation also provides a scene-independent assessment without overlap between the training and test images. The results therefore complement the internal CCTV evaluation, for which explicit camera-disjoint splitting could not be guaranteed.

4.2. Implementation Details

All experiments were implemented using PyTorch 2.10.0 and conducted on a Kaggle cloud platform equipped with an NVIDIA Tesla T4 GPU. All input images were resized to 256 × 256 pixels. The network was optimized using the AdamW optimizer [46] with an initial learning rate of 2×10^{-4} and a weight decay of 1×10^{-4}. The learning rate was updated using the cosine annealing learning-rate scheduler [47]. Models were trained for 100 epochs with a batch size of 8, using 2 data-loading workers and a fixed random seed of 42. The complete training hyperparameters are summarized in Table 1.

Table 1. Training hyperparameters used in all experiments.

Hyperparameter	Value
Learning Rate	2×10^{-4}
Batch Size	8
Optimizer	AdamW [46]
Weight Decay	1×10^{-4}
Learning Rate Scheduler	cosine annealing learning-rate [47]
Training Epochs	100
Input Resolution	(256 × 256)
Number of Workers	2
Random Seed	42

The proposed models were optimized using the hybrid loss function described in Section 3.5.3 For a consistent comparison, methods retrained in this study were evaluated using the same dataset splits, input resolution, degradation settings, and evaluation metrics as the proposed models. For methods relying on publicly released pretrained models or method-specific training procedures, the official implementations and recommended configurations were used. All methods were evaluated on the same test images using the same PSNR and SSIM calculation protocol. Quantitative performance was assessed using full-reference image quality metrics. The training convergence of both LCVR-Net variants is illustrated in Figures 5 and 6. For LCVR-Net-D and LCVR-Net-F, the training loss progressively decreases, while the validation PSNR and SSIM [43] consistently improve and stabilize toward the later epochs. These trends demonstrate stable optimization and convergence of both models under the adopted training configuration.

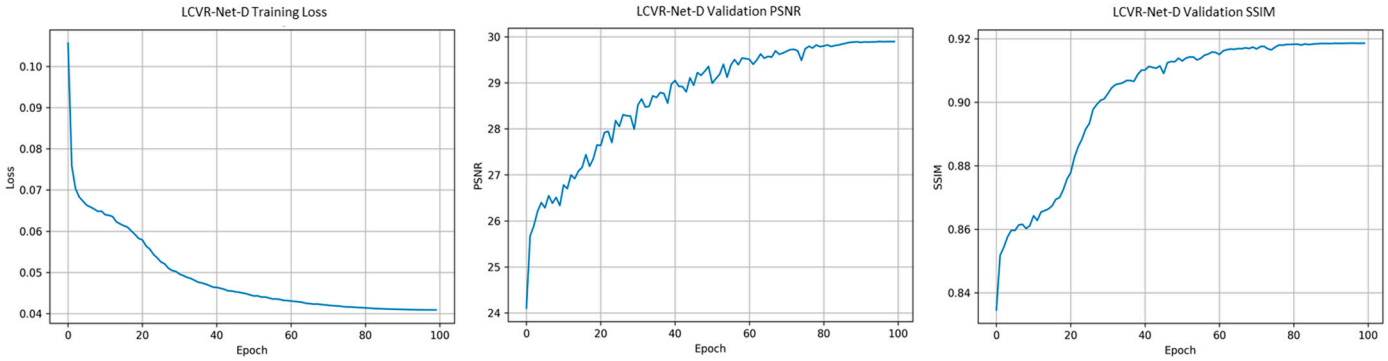

Figure 5. Training convergence of the proposed LCVR-Net-D over 100 epochs showing training loss, validation PSNR, and validation SSIM [43].

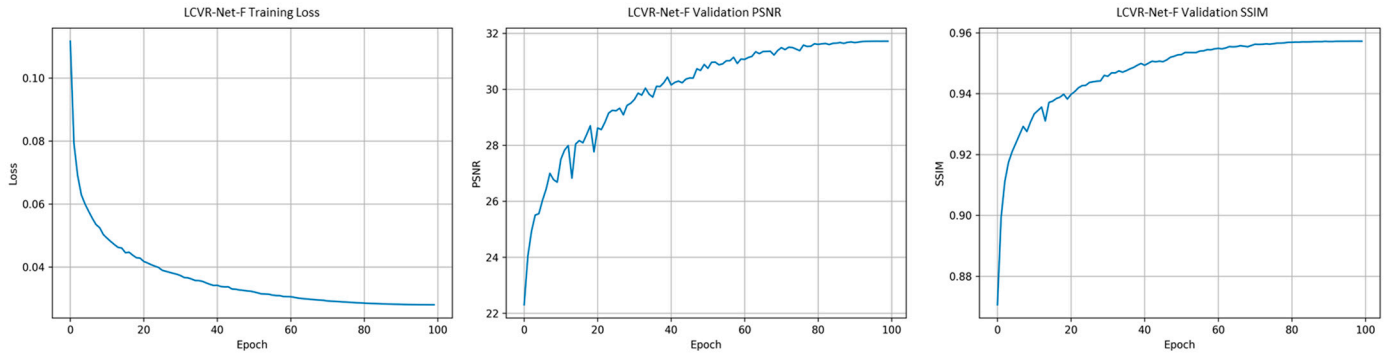

Figure 6. Training convergence of the proposed LCVR-Net-F over 100 epochs showing training loss, validation PSNR, and validation SSIM [43].

5. Experimental Results

This section presents a comprehensive evaluation of the proposed LCVR-Net under dust- and fog-degraded surveillance conditions. The analysis includes quantitative and qualitative comparisons with conventional and deep learning-based restoration methods, followed by computational efficiency and ablation studies. Cross-dataset and real-world evaluations are further conducted to assess the generalization and practical applicability of the proposed framework.

5.1. Quantitative Evaluation on Dust Restoration

The proposed LCVR-Net-D was evaluated on the CCTV dust evaluation set containing 100 previously unseen surveillance images. The proposed method was compared with traditional enhancement techniques, including CLAHE [8], Gamma Correction [7], Retinex [9],

White Balance [48], and DCP [6], as well as representative deep learning-based methods, including AOD-Net [12], DehazeNet [11], FFA-Net [13], AECR-Net [14], Dehaze-UNet [22], and FSNet [15]. Performance was evaluated using PSNR and SSIM. Table 2, summarizes the quantitative results.

Table 2. Quantitative comparison of dust image restoration methods on the CCTV [44] dust evaluation dataset.

Method	PSNR	SSIM [43]
CLAHE [8]	15.44	0.72
Gamma [7]	13.13	0.71
Retinex [8]	18.13	0.77
White Balance [48]	16.01	0.75
DCP [6]	13.81	0.66
AOD-Net [12]	24.81	0.84
DehazeNet [11]	26.62	0.86
FFA-Net [13]	30.85	0.92
AECR [14]	30.79	0.92
DehazeUNet [22]	30.57	0.91
FSNet [15]	31.24	0.92
Proposed Method	30.88	0.93

Traditional enhancement methods achieved limited restoration performance, whereas deep learning-based approaches substantially improved image quality. The proposed LCVR-Net-D achieved the highest SSIM (0.93) and a competitive PSNR (30.88 dB), outperforming all compared methods in structural preservation while maintaining comparable reconstruction accuracy. These results demonstrate the effectiveness of the proposed framework for restoring dust-degraded CCTV surveillance images.

5.2. Quantitative Evaluation on Fog Image Restoration

The performance of LCVR-Net-F was evaluated on both the Surveillance Cameras dataset [44] fog dataset and the RESIDE SOTS Outdoor benchmark [30], is a standard test subset of the RESIDE dataset [30] designed for evaluating single-image dehazing under outdoor conditions. Through Table 3, it appeared.

Table 3. Quantitative comparison of CCTV-trained models on the CCTV [44] fog and sots outdoor datasets [30].

Method	CCTV PSNR	CCTV SSIM [43]	SOTS Outdoor [30] PSNR	SOTS Outdoor [30] SSIM [43]
CLAHE [8]	16.53	0.77	16.43	0.82
Histogram Equalization [7]	16.26	0.66	18.08	0.83
Retinex [9]	15.02	0.76	13.13	0.79
Guided Filter [10]	14.29	0.73	15.91	0.78
DCP [6]	15.68	0.64	16.15	0.84
AOD-Net [12]	23.97	0.89	21.13	0.88
DehazeNet [11]	25.77	0.92	21.74	0.86
FFA-Net [13]	30.99	0.94	22.45	0.88

Table 3. *Cont.*

Method	CCTV PSNR	CCTV SSIM [43]	SOTS Outdoor [30] PSNR	SOTS Outdoor [30] SSIM [43]
AECR [14]	32.27	0.95	21.83	0.89
DehazeUNet [22]	31.48	0.95	21.91	0.89
FSNet [15]	32.34	0.96	22.02	0.89
LCVR-Net-F (Proposed)	32.78	0.96	22.11	0.89

It provides paired hazy and corresponding clean images, enabling objective quantitative evaluation using full-reference metrics such as PSNR and SSIM [43]. In this study, SOTS Outdoor [30] is used to assess the cross-dataset generalization performance of the evaluated restoration models. Comparisons were conducted against traditional enhancement methods, including CLAHE [8], Histogram Equalization (HE) [7], Retinex [9], Guided Filter [10], and DCP [6], as well as representative deep learning-based methods [11–15,22]. As shown in Table 3, LCVR-Net-F achieved the best performance on the CCTV fog benchmark, obtaining 32.78 dB PSNR and 0.96 SSIM. When the same CCTV-trained model was evaluated on the SOTS Outdoor benchmark, it achieved 22.11 dB PSNR and 0.89 SSIM. This result indicates that LCVR-Net-F can also restore fog-degraded images from general outdoor scenes beyond the CCTV surveillance domain, although the performance decrease compared with the CCTV benchmark reflects the effect of domain shift. To further investigate this effect, cross-dataset experiments were performed using models trained on the RESIDE dataset [30].

The results in Table 4, indicate that models trained on CCTV data consistently perform better on surveillance imagery, whereas RESIDE-trained models [30] achieve higher accuracy on the SOTS Outdoor benchmark [30]. These findings highlight the importance of domain-specific training for traffic surveillance image restoration.

Table 4. Cross-dataset evaluation of RESIDE-trained models [30] on the CCTV fog and SOTS outdoor datasets [44].

Method	CCTV PSNR	CCTV SSIM [43]	SOTS Outdoor [30] PSNR	SOTS Outdoor [30] SSIM [43]
AOD-Net [12]	20.45	0.87	23.03	0.89
DehazeNet [11]	19.88	0.85	23.37	0.9
FFA-Net [13]	20.49	0.87	26.81	0.94
AECR [14]	21.15	0.89	28.4	0.96
DehazeUNet [22]	21.02	0.89	27.94	0.95
FSNet [15]	21.22	0.89	28.38	0.96
LCVR-Net-F (Proposed)	21.17	0.89	28.34	0.96

5.3. Qualitative Evaluation on Dust Image Restoration

Representative restoration results are presented in Figure 7. Traditional enhancement methods generally fail to recover severe visibility degradation and often introduce color distortion or contrast imbalance. Although deep learning-based methods significantly improve image quality, residual degradation remains in challenging regions.

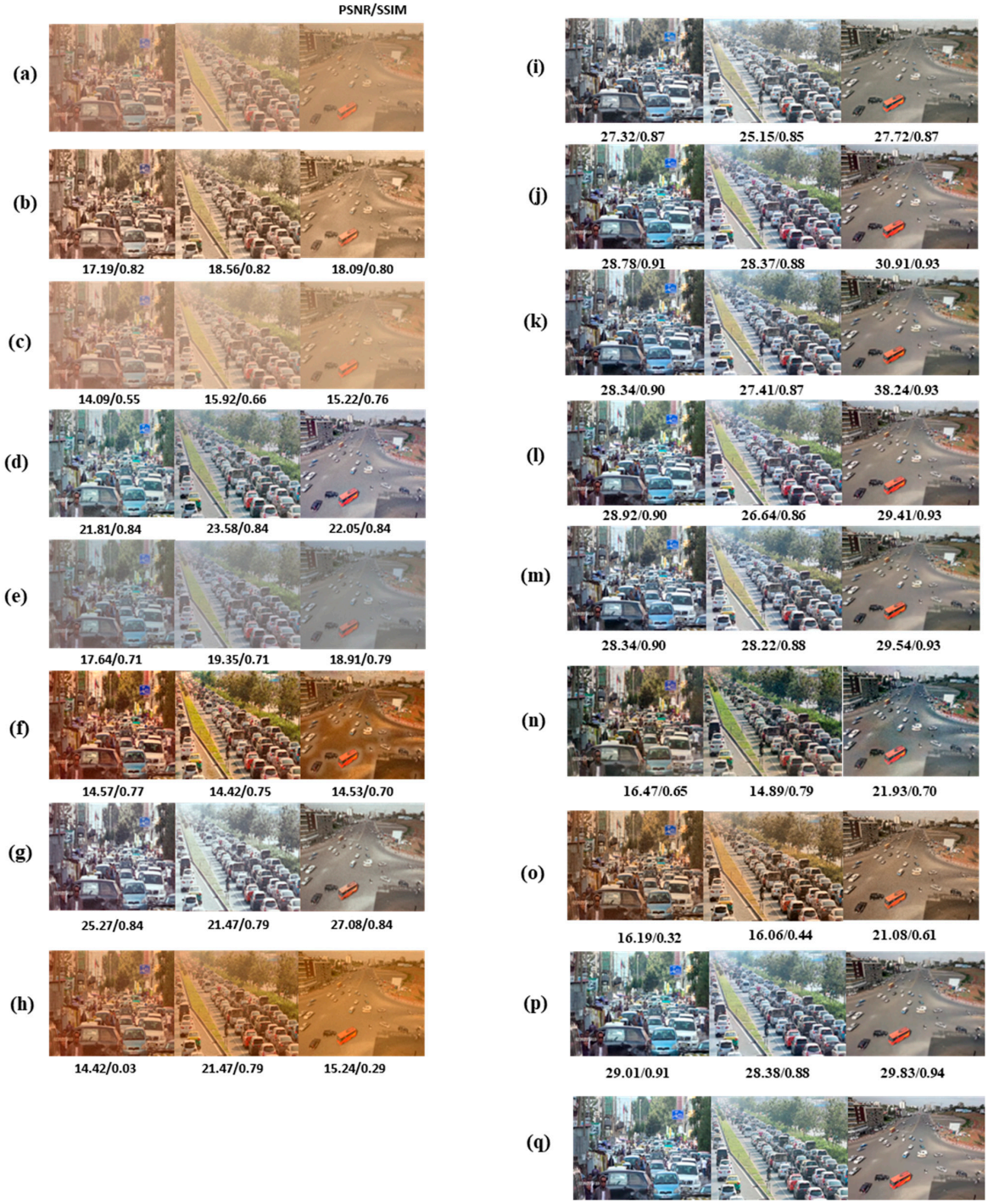

Figure 7. Visual comparison of dust restoration results on CCTV surveillance images: (**a**) Input; (**b**) CLAHE [8]; (**c**) Gamma [7]; (**d**) Retinex [8]; (**e**) White Balance [48]; (**f**) DCP [6]; (**g**) AOD-Net [12]; (**h**) DNDM [28]; (**i**) DehazeNet [11]; (**j**) FFA-Net(13); (**k**) AECR-Net [14]; (**l**) DEhazeUNet [22]; (**m**) FSNet [15]; (**n**) SSDIE [29]; (**o**) TOENet [27]; (**p**) Proposed method; (**q**) Ground truth.

In comparison, LCVR-Net-D produces clearer scene structures, improved color consistency, and enhanced visibility, generating restoration results that are visually closer to the ground-truth images. These observations are consistent with the quantitative results reported in Table 2.

5.4. Qualitative Evaluation on Fog Restoration

Representative visual comparisons are shown in Figure 8. Traditional enhancement techniques exhibit limited capability under dense fog conditions, while existing deep learning-based methods improve visibility but often retain residual haze and reduced contrast. In comparison, LCVR-Net-F restores clearer scene structures, enhances distant objects, and preserves better contrast and color consistency, producing results that are visually closer to the reference images. These qualitative observations are consistent with the quantitative evaluation presented in Table 3.

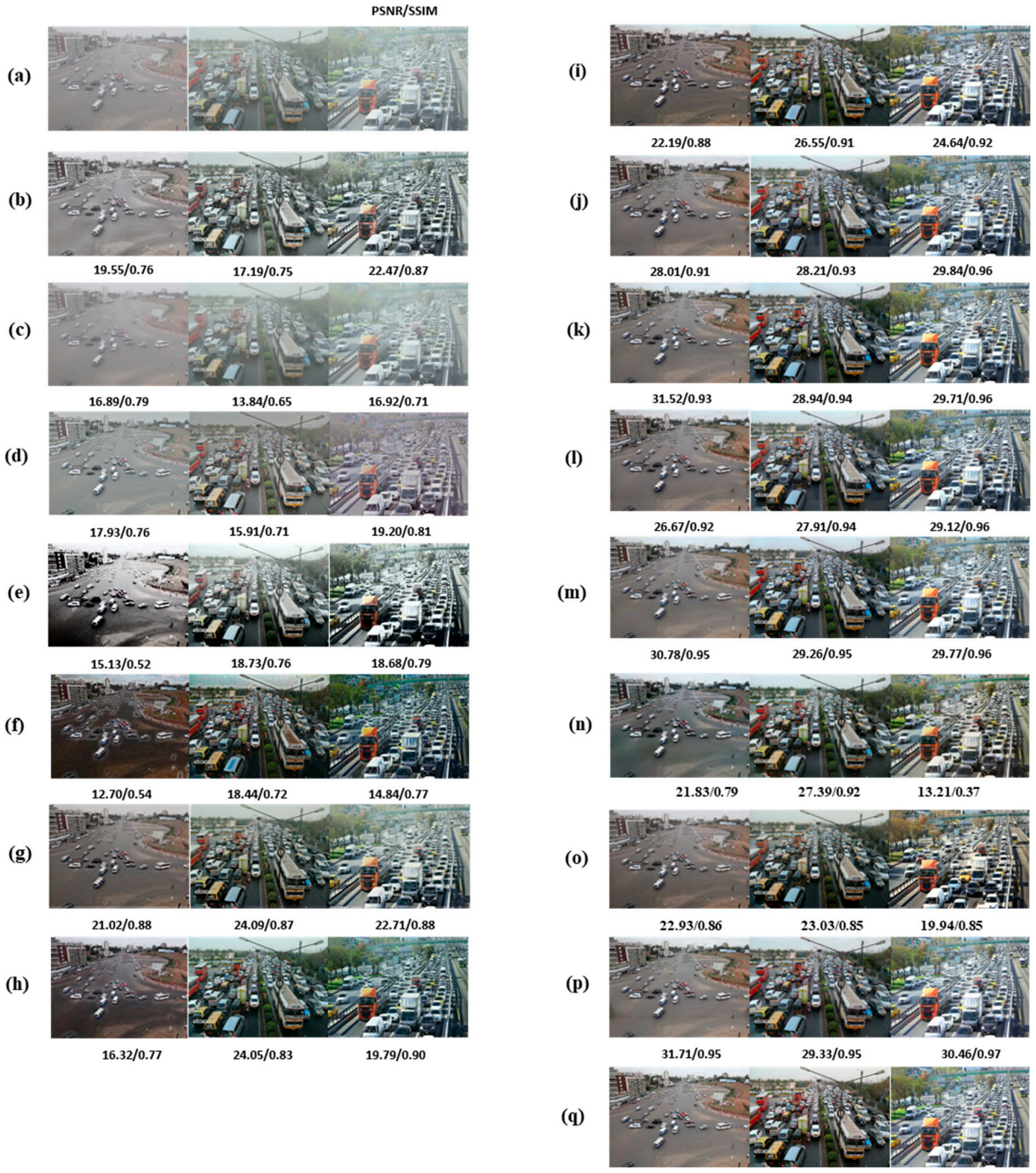

Figure 8. Visual comparison of fog restoration results on CCTV surveillance images: (**a**) Input; (**b**) CLAHE [8]; (**c**) Guided Filter [10]; (**d**) Retinex [9]; (**e**) Histogram Equalization [7]; (**f**) DCP [6]; (**g**) AOD-Net [12]; (**h**) DNDM [28]; (**i**) DehazeNet [11]; (**j**) FFA-Net [13]; (**k**) AECR-Net [14]; (**l**) DehazeUNet [22]; (**m**) FSNet [15]; (**n**) SSDIE [29]; (**o**) TOENet [27]; (**p**) Proposed method; (**q**) Ground truth.

5.5. Real-World Surveillance Image Evaluation

To evaluate practical applicability, qualitative experiments were conducted on naturally degraded surveillance images from the DS Dataset [45]. Representative results are shown in Figure 9. Compared with existing deep learning-based methods, LCVR-Net-F consistently restores clearer scene structures, improves the visibility of traffic-related objects, and enhances overall contrast without introducing noticeable artifacts. Although quantitative full-reference evaluation is not possible due to the absence of corresponding clean ground-truth images, the visual results provide qualitative evidence that the proposed framework can improve visibility under naturally occurring atmospheric degradation. These observations should therefore be interpreted as qualitative support for practical applicability rather than conclusive evidence of robustness or generalization.

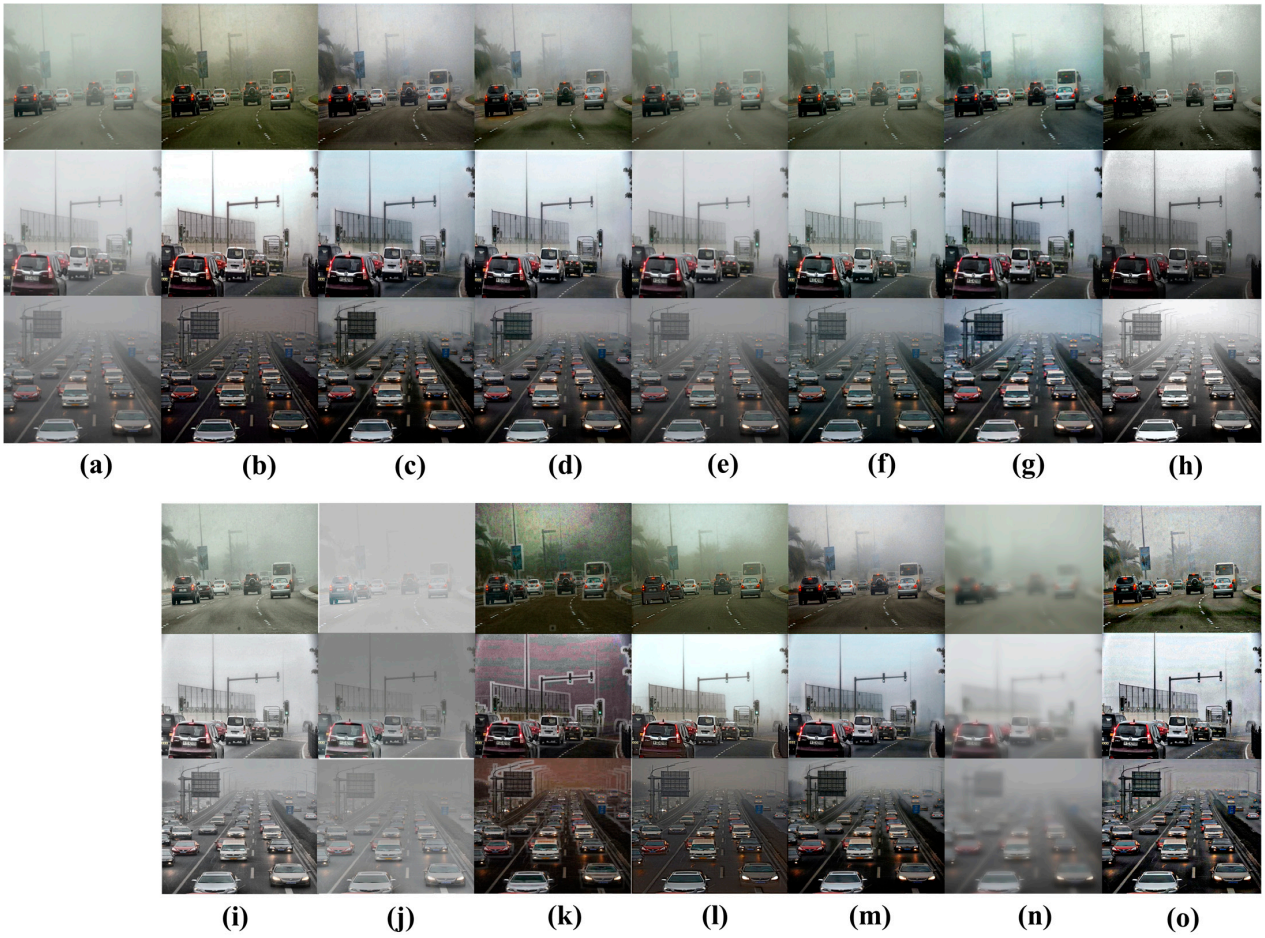

Figure 9. Qualitative comparison of real-world fog-degraded surveillance image restoration results: (**a**) Input; (**b**) DehazeNet [11]; (**c**) FFA-Net [13]; (**d**) AECR-Net [14]; (**e**) DNDM [27]; (**f**) TOENet [27]; (**g**) SSDIE [29]; (**h**) Histogram Equalization [7]; (**i**) CLAHE [8]; (**j**) Retinex [8]; (**k**) DCP [6]; (**l**) AOD-Net [12]; (**m**) DehazeUNet [22]; (**n**) Guided Filter [10]; (**o**) Proposed method.

The representative examples in Figure 9 further illustrate the task-level effect of visibility degradation and restoration. Under degraded conditions, vehicle detections exhibit reduced confidence or are missed entirely, whereas the corresponding restored images recover several of these detections with increased confidence. Together with the quantitative results in Table 5, these observations provide evidence that the proposed restoration framework improves not only image-level quality but also downstream vehicle detection performance in traffic-surveillance scenes. To evaluate the performance of the proposed method against state-of-the-art approaches, this study used the Natural Image

Quality Evaluator (NIQE) [49] and the Underwater Image Quality Measure (UIQM) [50] to assess real-world images. Specifically, UIQM [50] evaluates image sharpness, contrast, and colorfulness, whereas NIQE [49] assesses image naturalness based on contrast. A lower NIQE score indicates better quality, whereas a higher UIQM score reflects superior performance. Furthermore, to refine the results of the proposed method, a guided filter [10] was employed. As shown in Table 5, the proposed method achieves superior enhancement quality for real fog images compared to other methods.

In Table 5, ↓ and ↑ indicate that lower and higher values are better, respectively.

Table 5. Comparison on NIQE [49] and UIQM [50] metrics.

	Input	AECR-Net [14]	AOD-Net [12]	CLAHE [8]	DCP [6]	DehazeNet [11]	DehazeUNet [22]	FFA-Net [13]
NIQE ↓ [49]	24.272	20.087	20.215	19.677	19.418	20.057	19.979	19.979
UIQM ↑ [50]	0.681	1.148	1.022	0.953	1.348	1.229	1.22	1.22
	FSNet [15]	Guided_Filter [10]	Histogram_ Equalization [7]	Retinex [9]	TOENet [27]	SSDIE [29]	DNDM [28]	Proposed Method
NIQE ↓ [49]	20.087	20.323	19.871	20.188	24.12	23.588	24.911	19.195
UIQM ↑ [50]	1.148	0.379	1.104	0.645	1.178	1.22	0.892	1.495

6. Ablation Study

To quantify the contribution of the proposed modules, ablation experiments were conducted for both dust and fog restoration tasks.

6.1. Dust Image Restoration

The ablation results in Table 6 demonstrate that incorporating the Color Correction Head (CCH) improves the PSNR from 29.08 dB to 29.98 dB and the SSIM [43] from 0.9170 to 0.9197, confirming its effectiveness in compensating for dust-induced color distortion and improving restoration quality. In Tables 6 and 7, O denotes the inclusion of the corresponding module, whereas X denotes its absence from the evaluated configuration.

Table 6. Ablation study of the color correction head on dust restoration.

Configuration	CCH	PSNR (dB)	SSIM [43]
Baseline	X	29.08	0.9170
LCVR-Net-D	O	29.98	0.9197

Table 7. Ablation study of DARM and CCH on fog restoration.

Configuration	DARM	CCH	PSNR (dB)	SSIM [43]
Baseline	X	X	26.71	0.9297
Baseline + DARM	O	X	27.01	0.9313
LCVR-Net-F	O	O	31.72	0.9572

6.2. Fog Image Restoration

The results in Table 7 show that introducing the Dual Attention Refinement Module (DARM) improves the baseline performance from 26.71 dB/0.9297 to 27.01 dB/0.9313 (PSNR/SSIM [43]). Incorporating both DARM and CCH further increases the performance to 31.72 dB PSNR and 0.9572 SSIM [43], demonstrating the complementary contributions of the proposed modules to fog restoration.

7. Downstream Vehicle Detection Evaluation

To investigate whether visibility restoration benefits downstream traffic-surveillance analysis, an additional vehicle detection experiment was conducted on the CCTV evaluation set. A fixed pretrained YOLOv8n detector was applied to clean, degraded, and LCVR-Net-restored images without retraining or fine-tuning. The representative examples in Figure 10 further illustrate the task-level effect of visibility degradation and restoration. Under degraded conditions, vehicle detections exhibit reduced confidence or are missed entirely, whereas the corresponding restored images recover several of these detections with increased confidence. Although there are some overlapping regions in the image, there is no scientific issue because it represents the object recognition rate.

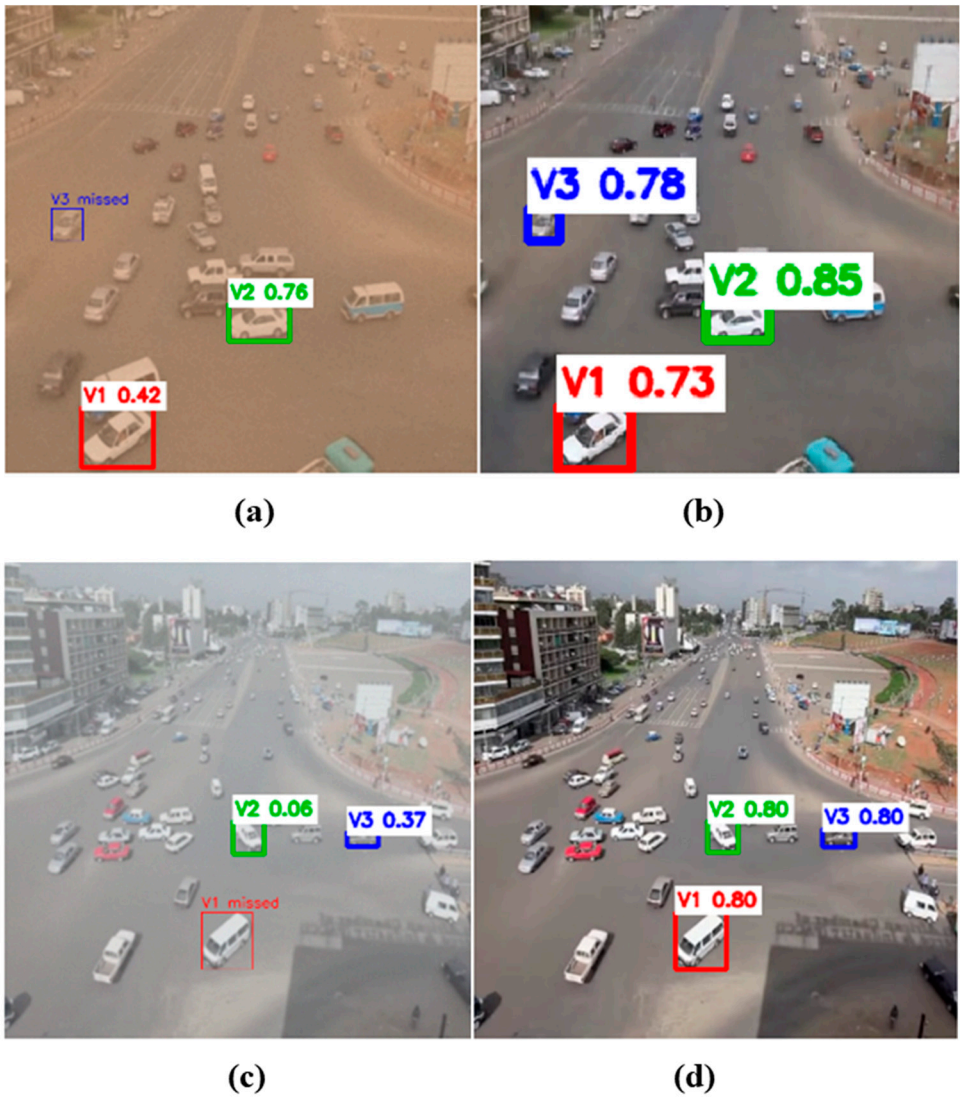

Figure 10. Representative vehicle detection results before and after visibility restoration: (**a**) Dust-degraded; (**b**) LCVR-Net-D; (**c**) Fog-degraded; (**d**) LCVR-Net-F.

8. Conclusions

This paper presented LCVR-Net, a lightweight visibility restoration framework for enhancing dust- and fog-degraded CCTV surveillance images in intelligent transportation systems. The proposed framework employs Residual Depthwise Separable Blocks (RDSBs) to achieve computationally efficient feature extraction, while a Dual Attention Refinement Module (DARM) enhances degradation-aware feature representation for fog restoration and a lightweight Color Correction Head (CCH) compensates for atmospheric color distortion.

Two task-specific variants, LCVR-Net-D and LCVR-Net-F, were developed to effectively address the distinct characteristics of dust and fog degradation within a unified architecture. A CCTV-oriented benchmark with synthetically generated atmospheric degradations was also constructed to facilitate systematic evaluation under realistic traffic surveillance scenarios. Extensive quantitative, qualitative, computational, and ablation experiments demonstrate that the proposed framework achieves competitive restoration performance while maintaining high computational efficiency. LCVR-Net supports inference, making them suitable for deployment on resource-constrained surveillance systems. Furthermore, qualitative evaluation on real-world surveillance images confirms the practical applicability and generalization capability of the proposed framework under naturally degraded atmospheric conditions. The principal contributions of this work are the development of a lightweight surveillance-oriented restoration architecture based on RDSBs, the degradation-aware DARM for enhanced fog restoration, and the lightweight CCH for atmospheric color correction. In addition, the constructed CCTV-oriented dust and fog benchmark enables systematic evaluation under surveillance specific degradation conditions. Together, these contributions provide a compact framework that jointly addresses restoration quality, computational efficiency, and practical deployment requirements. Future work will focus on extending LCVR-Net to handle multiple adverse weather conditions, including rain, snow, and nighttime low-visibility environments, while improving cross-domain generalization through training on larger real-world surveillance datasets. In addition, integrating the proposed restoration framework with downstream traffic analysis tasks, such as vehicle detection, tracking, and traffic flow estimation, represents a promising direction for developing more robust and practical intelligent transportation systems.

Author Contributions: Conceptualization & methodology & validation & formal analysis & visualization & writing—original draft preparation, N.M.; writing—review and editing & supervision, H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data generated or analyzed during this study are included in this published article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Guerrero-Ibáñez, J.A.; Zeadally, S.; Contreras-Castillo, J. Sensor technologies for intelligent transportation systems. *Sensors* **2018**, *18*, 1212. [CrossRef] [PubMed]
2. Yuan, T.; Da Rocha Neto, W.; Rothenberg, C.E.; Obraczka, K.; Barakat, C.; Turletti, T. Machine learning for next-generation intelligent transportation systems: A survey. *Trans. Emerg. Telecommun. Technol.* **2022**, *33*, e4427. [CrossRef]
3. Dilek, E.; Dener, M. Computer vision applications in intelligent transportation systems: A survey. *Sensors* **2023**, *23*, 2938. [CrossRef] [PubMed]
4. Narasimhan, S.G.; Nayar, S.K. Vision and the atmosphere. *Int. J. Comput. Vis.* **2002**, *48*, 233–254. [CrossRef]
5. Fattal, R. Single image dehazing. *ACM Trans. Graph.* **2008**, *27*, 1–9. [CrossRef]
6. He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 2341–2353. [CrossRef] [PubMed]
7. Gonzalez, R.C.; Woods, R.E. *Digital Image Processing*, 4th ed.; Pearson: New York, NY, USA, 2018.
8. Zuiderveld, K. Contrast Limited Adaptive Histogram Equalization. In *Graphics Gems IV*; Heckbert, P.S., Ed.; Academic Press: San Diego, CA, USA, 1994; pp. 474–485. [CrossRef]
9. Land, E.H. The Retinex theory of color vision. *Sci. Amer.* **1977**, *237*, 108–128. [CrossRef] [PubMed]
10. He, K.; Sun, J.; Tang, X. Guided image filtering. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1397–1409. [CrossRef] [PubMed]

11. Cai, B.; Xu, X.; Jia, K.; Qing, C.; Tao, D. DehazeNet: An end-to-end system for single image haze removal. *IEEE Trans. Image Process.* **2016**, *25*, 5187–5198. [CrossRef] [PubMed]
12. Li, B.; Peng, X.; Wang, Z.; Xu, J.; Feng, D. AOD-Net: All-in-one dehazing network. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 4780–4788. [CrossRef]
13. Qin, X.; Wang, Z.; Bai, Y.; Xie, X.; Jia, H. FFA-Net: Feature fusion attention network for single image dehazing. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 11908–11915. [CrossRef]
14. Wu, H.; Qu, Y.; Lin, S.; Zhou, J.; Qiao, R.; Zhang, Z.; Xie, Y.; Ma, L. Contrastive learning for compact single image dehazing. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 10551–10560. [CrossRef]
15. Dudhane, A.; Zamir, S.W.; Khan, S.; Khan, F.S.; Yang, M.H. Restormer: Efficient transformer for high-resolution image restoration. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; pp. 5728–5739. [CrossRef]
16. Zhou, H.; Chen, Z.; Li, Q.; Tao, T. Dehaze-UNet: A lightweight network based on UNet for single-image dehazing. *Electronics* **2024**, *13*, 2082. [CrossRef]
17. Cui, Y.; Ren, W.; Cao, X.; Knoll, A. Image restoration via frequency selection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2024**, *46*, 1093–1108. [CrossRef] [PubMed]
18. Berman, D.; Treibitz, T.; Avidan, S. Non-local image dehazing. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 1674–1682. [CrossRef]
19. Fattal, R. Dehazing using color-lines. *ACM Trans. Graph.* **2014**, *34*, 1–14. [CrossRef]
20. Meng, G.; Wang, Y.; Duan, J.; Xiang, S.; Pan, C. Efficient image dehazing with boundary constraint and contextual regularization. In Proceedings of the Proceedings of the IEEE International Conference on Computer Vision (ICCV), Sydney, Australia, 1–8 December 2013; pp. 617–624. [CrossRef]
21. Zhu, Q.; Mai, J.; Shao, L. A fast single image haze removal algorithm using color attenuation prior. *IEEE Trans. Image Process.* **2015**, *24*, 3522–3533. [CrossRef] [PubMed]
22. Ali, A.M.; Benjdira, B.; Koubaa, A.; El-Shafai, W.; Khan, Z.; Boulila, W. Vision Transformers in Image Restoration: A Survey. *Sensors* **2023**, *23*, 2385. [CrossRef] [PubMed]
23. Zhang, H.; Patel, V.M. Densely connected pyramid dehazing network. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 3194–3203. [CrossRef]
24. Ren, W.; Ma, L.; Zhang, J.; Pan, J.; Cao, X.; Liu, W.; Yang, M.-H. Gated fusion network for single image dehazing. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 3253–3261. [CrossRef]
25. Liu, X.; Ma, Y.; Shi, Z.; Chen, J. GridDehazeNet: Attention-based multi-scale network for image dehazing. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 7314–7323.
26. Dong, H.; Pan, J.; Xiang, L.; Hu, Z.; Zhang, X.; Wang, F.; Yang, M.-H. Multi-scale boosted dehazing network with dense feature fusion. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 2154–2164. [CrossRef]
27. Gao, Y.; Xu, W.; Lu, Y. Let you see in haze and sandstorm: Two-in-one low-visibility enhancement network. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 5023712. [CrossRef]
28. Jia, T.; Li, J.; Zhuo, L.; Yu, T. Semi-supervised single-image dehazing network via disentangled meta-knowledge. *IEEE Trans. Multimed.* **2023**, *26*, 2634–2647. [CrossRef]
29. Qadir, M.S.; Bartani, A.; Mohammed, M.A.; Daneshfar, F. Semi-Supervised sand-Dust image enhancement via attention-Driven multi-Scale feature fusion network. *Digit. Signal Process.* **2026**, *177*, 106093. [CrossRef]
30. Li, B.; Ren, W.; Fu, D.; Tao, D.; Feng, D.; Zeng, W.; Wang, Z. Benchmarking Single-Image Dehazing and Beyond. *IEEE Trans. Image Process.* **2019**, *28*, 492–505. [CrossRef] [PubMed]
31. Asif, M.H.; Bennamoun, M.; Sohel, F.; Togneri, R. Traffic surveillance under adverse weather conditions: A review. *IEEE Access* **2021**, *9*, 94692–94715.
32. Song, Y.; He, Z.; Qian, H.; Du, X. Vision transformers for single image dehazing. *IEEE Trans. Image Process.* **2023**, *32*, 1927–1941. [CrossRef] [PubMed]
33. Qiu, Y.; Zhang, K.; Wang, C.; Luo, W.; Li, H.; Jin, Z. MB-TaylorFormer: Multi-branch efficient transformer expanded by Taylor formula for image dehazing. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France, 1–6 October 2023; pp. 12756–12767. [CrossRef]
34. Lu, L.; Xiong, Q.; Xu, B.; Chu, D. MixDehazeNet: Mix structure block for image dehazing network. In Proceedings of the International Joint Conference on Neural Networks (IJCNN), Yokohama, Japan, 30 June–5 July 2024; pp. 1–10. [CrossRef]

35. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI), Munich, Germany, 5–9 October 2015; pp. 234–241. [CrossRef]
36. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [CrossRef]
37. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 1251–1258. [CrossRef]
38. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. MobileNetV2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520. [CrossRef]
39. Mastromichalakis, S. ALReLU: A different approach on Leaky ReLU activation function to improve Neural Networks Performance. *arXiv* **2020**, arXiv:2012.07564.
40. Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I.S. CBAM: Convolutional block attention module. In Proceedings of the 15th European Conference Computer Vision—ECCV 2018, Munich, Germany, 8–14 September 2018; pp. 3–19. [CrossRef]
41. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141. [CrossRef]
42. Menon, A.; Mehrotra, K.; Mohan, C.K.; Ranka, S. Characterization of a class of sigmoid functions with applications to neural networks. *Neural Netw.* **1996**, *9*, 819–835. [CrossRef] [PubMed]
43. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef] [PubMed]
44. Tanish, J. Traffic Data from Surveillance Cameras. Kaggle. 2024. Available online: https://www.kaggle.com/datasets/tanishjain0604/traffic-data-from-surveillance-cameras (accessed on 15 July 2026).
45. Vijayakumar, V. DS Dataset: Dust and Fog Degraded Image Dataset. Kaggle. 2023. Available online: https://www.kaggle.com/datasets/vijayvkb98/ds-dataset (accessed on 15 July 2026).
46. Loshchilov, I.; Hutter, F. Decoupled weight decay regularization. In Proceedings of the International Conference on Learning Representations (ICLR), New Orleans, LA, USA, 6–9 May 2019.
47. Loshchilov, I.; Hutter, F. SGDR: Stochastic gradient descent with warm restarts. In Proceedings of the International Conference on Learning Representations (ICLR), Toulon, France, 24–26 April 2017.
48. Finlayson, G.D.; Hordley, S.D.; Hubel, P.M. Color by correlation: A simple, unifying framework for color constancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *23*, 1209–1221. [CrossRef]
49. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a "completely blind" image quality analyzer. *IEEE Signal Process. Lett.* **2012**, *20*, 209–212. [CrossRef]
50. Panetta, K.; Gao, C.; Agaian, S. Human-visual-system-inspired underwater image quality measures. *IEEE J. Ocean. Eng.* **2015**, *41*, 541–551. [CrossRef]