Singapore Computer Society
International Journal of Information Technology
Volume 31, Issue 2, 2025

Sustainable Edge AI for Weather-Adaptive Image Super-Resolution in Safety-Critical Perception

Zhou Zhitong[a,b], Gan Lu[b] and Liu Hongqing[c]

[a]Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798
[b]Brunel University London, Uxbridge, United Kingdom, UB8 3PH
[c]Chongqing University of Posts and Telecommunications, No.2 Chongwen Road, Chongqing, China 400065

zhitong002@e.ntu.edu.sg, lu.gan@brunel.ac.uk, hongqingliu@cqupt.edu.cn

Abstract

Background: Real-world intelligent systems, including autonomous driving and smart mobility infrastructures, require reliable visual perception under adverse environmental conditions to improve public safety. Fog, rain, snow, and low-light scenes can degrade image quality and increase safety risks. At the same time, the high computational cost of deep image restoration models limits their deployment on resource-constrained edge devices, creating a need for more sustainable and deployable AI solutions.

Methods: This study presents a lightweight weather-adaptive image super-resolution framework for safety-critical edge perception. Built upon SwinIR, the proposed model combines model compression and knowledge distillation to reduce computational burden while preserving reconstruction quality. A weather-adaptive loss function is further introduced to guide condition-specific restoration under diverse environmental degradations.

Results: The framework is evaluated on a small-scale adverse-weather subset derived from the ACDC dataset. Experimental results suggest that the proposed method improves reconstruction quality under fog, rain, snow, and night conditions in this evaluation setting. Compared with classical super-resolution baselines that are not specifically adapted to the constructed adverse-weather setting, it achieves higher PSNR/SSIM scores while maintaining a lightweight model design.

Conclusions: Beyond image restoration performance, the proposed framework supports Sustainable AI by reducing model size for safety-critical perception. More reliable and efficient visual perception can contribute to safer intelligent mobility systems while reducing deployment burden on resource-constrained edge devices.

Keywords: Sustainable AI, edge AI, image super-resolution, autonomous driving, adverse weather, model compression, knowledge distillation, safety-critical perception.

I. Introduction

Image super-resolution (SR) overcomes the limitations of traditional imaging systems, opening up new possibilities in fields such as autonomous driving, medical imaging, and remote sensing [1]. In intelligent transportation, SR is closely related to safety-critical perception because visual degradation can affect downstream recognition, segmentation, and decision-making. Autonomous vehicles and smart mobility systems promise economic and social benefits, but their reliability depends on robust perception in real-world environments. Recent surveys further indicate that adverse weather remains a long-standing obstacle for higher-level autonomous driving, as fog, rain, snow, and low-light conditions can degrade sensing quality and reduce perception reliability [2].

At the same time, the growth of deep learning models creates sustainability and deployment challenges. Large models can require substantial computational resources, with associated financial and environmental costs [3]. The Green AI perspective argues that efficiency should be evaluated alongside accuracy, so that AI systems can become more environmentally responsible and more inclusive to deploy and reproduce [4]. This concern is especially relevant to edge intelligence, where AI services are increasingly pushed from central clouds to resource-constrained devices near the data source [5].

In the field of image super-resolution, early interpolation-based algorithms are simple and fast, but suffer from poor quality. Mainstream learning-based algorithms, such as Super-resolution Convolutional Neural Network (SRCNN), Fast Super-resolution Convolutional Neural Network (FSRCNN), and Enhanced Deep Super-resolution Network (EDSR), are suitable for image quality restoration in complex scenes. Nevertheless, they are still limited by excessive computational demands and insufficient detail recovery [6]. Transformer-based algorithms, especially Image Restoration using Swin Transformer (SwinIR), have improved overall structure recovery and noise removal through stronger feature extraction and reconstruction techniques [7]. However, self-attention also increases computation, making even lightweight variants difficult to deploy in some resource-limited edge perception settings.

In addition, the robustness of SR algorithms is challenged by adverse weather. Dehazing and deraining techniques primarily focus on image restoration, but adaptive super-resolution learning algorithms for weather variations remain relatively scarce, often suffering from limitations such as insufficient real-time processing and weak compatibility with multiple weather conditions [8]. There is therefore a need for lightweight weather-adaptive SR methods that balance reconstruction quality, robustness, and deployability for sustainable safety-critical perception.

The aim of this study is to develop and evaluate a lightweight weather-adaptive image super-resolution framework for sustainable and safety-critical edge perception. The proposed model, LiteSwinIR, uses model compression and knowledge distillation to reduce computational burden while maintaining restoration quality for road images under multiple adverse weather conditions.

The main objectives of this study are as follows.

1. To review image super-resolution, model compression, and adverse-weather perception methods relevant to safety-critical edge AI.
2. To train the proposed LiteSwinIR model using knowledge distillation so that computational cost is reduced for resource-constrained deployment.
3. To design a weather-adaptive loss function that adjusts the optimization direction according to different environmental conditions.
4. To preprocess the ACDC dataset according to the model requirements, incorporating weather labels and weather-related degradation features.
5. To compare the proposed model with classical baseline models under a constructed adverse-weather evaluation setting using quantitative and qualitative evaluation.
6. To discuss the proposed model as a Sustainable AI case study by analysing the trade-off between restoration performance and model efficiency.

The implementation includes downsampling the ACDC dataset, adding weather-related degradation features, and extracting sub-images for training. The knowledge distillation framework is combined with model-size reduction techniques to construct a lightweight student model. Weather-adaptive loss terms are designed to improve SR performance under night, foggy, rainy, and snowy conditions. The results are compared with baseline models to evaluate both the restoration performance and the edge-oriented efficiency of LiteSwinIR.

The contribution of this study is threefold. First, it implements a lightweight super-resolution framework by adapting SwinIR through model compression and knowledge distillation, aiming to preserve restoration quality while reducing model size. Second, it introduces a weather-adaptive loss design that adjusts the optimization objective according to different weather conditions, improving robustness in severe visual environments. Third, it evaluates the proposed model on a constructed adverse-weather setting based on the Adverse Conditions Dataset with Correspondences (ACDC) and reports both reconstruction performance and efficiency-related indicators, connecting technical image restoration with the broader goals of Sustainable AI and safety-critical perception.

II. Related Work

A. Image Super-Resolution

Image super-resolution (SR) aims to reconstruct a high-resolution image from a low-resolution input. Early SR methods mainly relied on interpolation-based techniques, such as bicubic interpolation, which are computationally efficient but have limited ability to recover high-frequency details and complex textures [9]. The introduction of deep learning significantly improved reconstruction quality. SRCNN first demonstrated that a convolutional neural network could learn an end-to-end mapping from low-resolution to high-resolution images [10]. Later, deeper residual architectures such as EDSR further improved reconstruction accuracy by increasing network capacity and removing unnecessary batch normalization layers [11].

In addition to pixel-level reconstruction accuracy, perceptual quality has also become an important objective in SR. ESRGAN introduced adversarial learning and perceptual loss to generate visually sharper and more realistic textures [12]. More recently, attention-based and Transformer-based models have achieved strong performance in image restoration. SwinIR combines convolutional feature extraction with Swin Transformer blocks, allowing the model to capture both local structures and long-range dependencies [7]. However, these high-performing models usually require more parameters and computation, which limits their direct deployment in resource-constrained scenarios such as edge devices and autonomous driving platforms.

Table 1: Representative Image Super-Resolution Methods

Method	**Year**	**Architecture**	**Main Contribution**
Bicubic Interpolation [9]	1981	Traditional	Efficient interpolation-based upsampling with limited detail recovery.
SRCNN [10]	2014	CNN	First end-to-end CNN-based SR framework.
EDSR [11]	2017	Residual CNN	Improved reconstruction accuracy using a deep residual architecture.
ESRGAN [12]	2018	GAN	Improved perceptual quality and high-frequency texture recovery.
IMDN [13]	2019	Lightweight CNN	Balanced reconstruction quality and efficiency through information multi-distillation.
SwinIR [7]	2021	Transformer	Applied Swin Transformer blocks to image restoration and SR.
SwinIRmini [14]	2023	Lightweight Transformer	Reduced SwinIR model size using pruning and knowledge distillation.

B. Lightweight Super-Resolution and Knowledge Distillation

As SR models become deeper and more complex, reducing computational cost has become increasingly important. Lightweight SR methods aim to preserve reconstruction quality while reducing the number of parameters, memory usage, and inference cost. IMDN is a representative lightweight SR model that uses information multi-distillation blocks to extract useful features efficiently [13]. For Transformer-based SR, SwinIRmini reduces the size of SwinIR through pruning and knowledge distillation, showing that compact Transformer-based models can still retain useful restoration performance [14].

Knowledge distillation is widely used for model compression. In a typical teacher-student framework, a compact student model learns not only from the ground-truth image but also from the output or intermediate features of a larger teacher model [15]. This allows the student model to inherit part of the teacher model's representation ability while using fewer parameters. Such a strategy is particularly suitable for edge-oriented SR, where inference efficiency and model size are important. This direction is also consistent with Green AI and edge intelligence, which emphasize that computational cost, latency, and deployment feasibility should be considered together with accuracy [3–5]. However, most existing lightweight SR and distillation methods focus on clean image reconstruction and do not explicitly address weather-specific degradation.

C. Adverse-Weather Image Restoration

Adverse weather conditions such as fog, rain, snow, and night scenes can degrade image quality by reducing contrast, distorting color, introducing noise, and obscuring structural details. These degradations are particularly important in autonomous driving and safety-critical perception systems, where poor image quality may affect downstream tasks such as object detection and semantic segmentation. Existing adverse-weather restoration methods usually focus on image enhancement, dehazing, deraining, or desnowing. Many deep learning-based methods directly learn the mapping from degraded images to clean images using convolutional networks, while more recent methods introduce attention mechanisms or multi-branch structures to focus on weather-affected regions [8, 16, 17].

Loss function design is also important in adverse-weather restoration. Standard reconstruction losses, perceptual losses, and adversarial losses are commonly used in image restoration and SR tasks [6, 18]. In weather-related restoration, some studies further design weather-specific constraints. For example, rain and snow degradation often damages high-frequency texture, so detail-preserving losses can help recover structural information [19]. Foggy scenes usually suffer from contrast reduction and distant-object degradation, where perceptual or feature-level constraints can improve visual restoration quality [20]. Nevertheless, most existing de-weathering methods focus on visibility enhancement or weather removal, while conventional SR methods are usually trained on clean datasets. Therefore, there remains a gap between adverse-weather restoration and lightweight image super-resolution.

D. Datasets and Research Gap

Dataset selection is important for weather-aware SR. DIV2K is a widely used benchmark dataset for single-image super-resolution and provides high-quality images with standard scaling factors [21]. However, it mainly contains clean natural images and does not cover adverse-weather driving scenes. Autonomous-driving datasets such as Foggy Cityscapes and BDD100K include road scenes under different weather conditions and provide annotations for perception tasks [22, 23]. However, they are mainly designed for detection or segmentation and do not provide paired clear high-resolution references suitable for controlled SR evaluation.

The ACDC dataset contains road scenes captured under adverse conditions such as fog, night, rain, and snow, together with corresponding clear-condition references [24]. This makes it more suitable for constructing paired data for weather-aware SR experiments. Based on the above studies, this work focuses on a lightweight weather-adaptive SR framework for adverse driving conditions. Different from standard SR methods trained on clean datasets, the proposed approach combines SwinIR-based model compression, knowledge distillation, and weather-adaptive loss design to improve reconstruction quality while reducing deployment burden.

III. Proposed Method

A. Overall Design

This section presents the methodology and design principles of this study. Overall, the design is divided into three phases. In the Data Preprocessing Phase, the ACDC dataset is reclassified into training, validation and testing sets, and the necessary features and labels are added by downsampling. Then the sub-images are extracted so that the dataset can be used as the correct input of the proposed model.

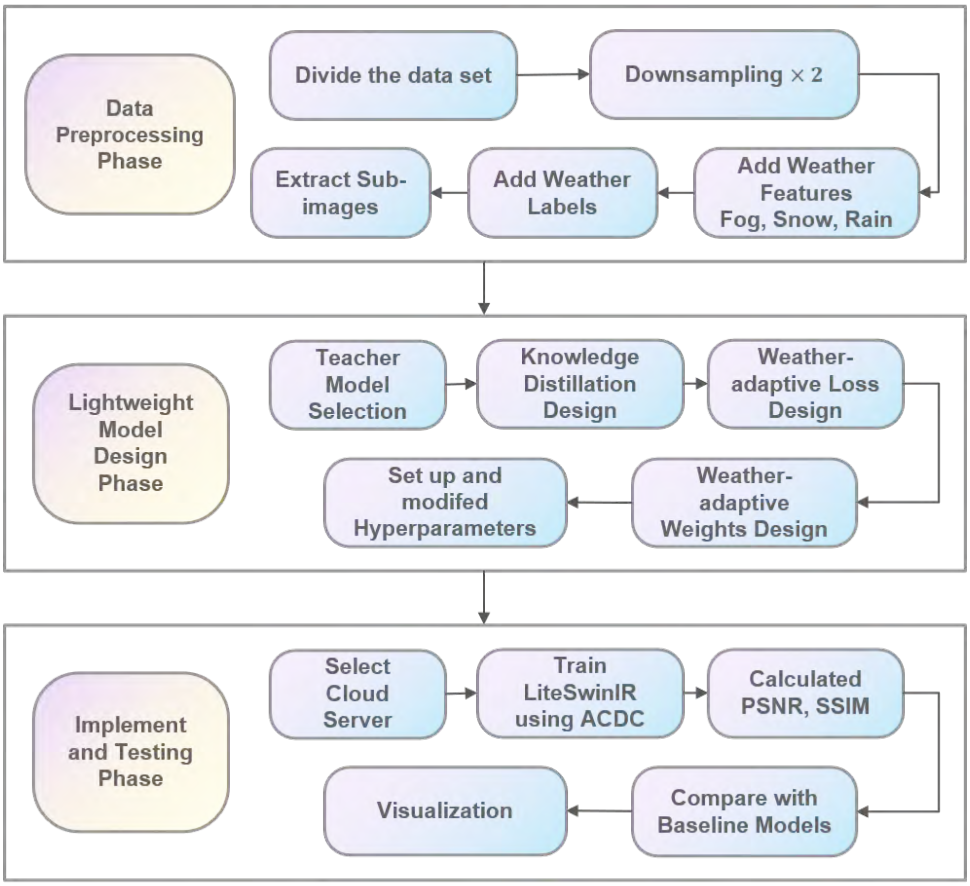

Fig 1: Overall Architecture of LiteSwinIR

In the Lightweight Model Design Phase, the main architecture of model compression and knowledge distillation are designed and the weather adaptive loss function is introduced.The SwinIR lightweight model is selected as the teacher model.

In the Implementation and Testing Phase, this study uses cloud servers for model training and compares the proposed model quantitatively and qualitatively with the baseline models. Finally, visualisation results are provided to analyse the training effect.

Fig. 1 provides the specific workflow diagram of this study.

B. Design of LiteSwinIR

B-1. SwinIR Architecture

For safety-critical edge perception, the SR images should exhibit clear edge structures with fine texture details, and the model should maintain a minimal size for efficient edge deployment.

This study initially considered classical methods such as ESRGAN and conducted preliminary tests on a test set with weather features. The results show that while these models have useful restoration capabilities in traditional low-resolution images, they perform poorly in detail texture restoration and lack adaptability to complex and variable environments in low-light and blurry images. On the contrary, the SwinIR model shows advantages in the combination of image de-weatherisation and super-resolution tasks. Due to the ability of the Transformer architecture to capture more complex contextual information, SwinIR can recover structural and detail information of the image when processing weather-affected images, thus providing a suitable teacher model for this study.

Therefore, in this study, a lightweight version of the SwinIR model was finally chosen for further compression, which is based on the Swin Transformer architecture combined with CNN, this makes it one of the relatively state-of-the-art SR model that can focus on improving the image distortion and overall structure[7]. SwinIR contains three main modules, the shallow feature extraction module is a convolutional layer that can extract the basic features of the input low-resolution image. The deep feature extraction module consists of multiple Residual Swin Transformer Blocks (RSTB) and contains many Swin Transformer layers. Global and local features can be captured efficiently through the window self-attention and the shift window self-attention mechanism. The high quality reconstruction module then combines the extracted image features for up-sampling to obtain the model results.

Lightweight SwinIR has 60 channels, 6 blocks, each block has 6 layers. As the lightweight version, it still has 878K parameters, which is overly demanding for some resource-constrained environments. Therefore this work further reduces resource consumption through knowledge distillation.

Compared with the traditional CNN-based model, SwinIR is able to capture contextual information and performs well in recovering high-frequency details. Its generalisation ability also makes it suitable as a teacher model to guide student model in optimizing for different weather conditions.

B-2. LiteSwinIR Architecture

In this study, the LiteSwinIR architecture draws on the proposed channel pruning design in [25] to remove channels with lower importance weight in the teacher SwinIR. To do the pruning, OBProx-SG [26] was used as a sparse optimisation solver, trained for 100 epochs to identify the network layers with lower contribution through sparsity-inducing optimisation.

The density d of nth layer is calculated using the following formula, where p_n is the total number of parameters for layer n.

$$d_l = \frac{\text{\# of nonzero weights in } l}{p_l} \tag{1}$$

The design of network compression is then referred to [14]. Based on the density ratio d_l calculated above, the following equation is used to assess the impact of RSTBs and the number of layers within each block on the model's performance and size.

$$\frac{\hat{N}_b}{N_b} \approx \sqrt[6]{d} \text{ and } \frac{\hat{N}_l + 1}{N_l + 1} \approx \sqrt[6]{d} \tag{2}$$

$$\frac{\hat{N}_c}{N_c} \approx \sqrt[3]{d \frac{N_b (N_l + 1)}{\hat{N}_b \left(\hat{N}_l + 1\right)}} \tag{3}$$

Where d is the density ratio, $\hat{N}_b$ is the number of RSTB blocks, $\hat{N}_l$ is the number of layers in each block, $\hat{N}_c$ is the number of channels.

The final results is $\hat{N}_b$ = 3, $\hat{N}_l$ = 4, $\hat{N}_c$ = 24. The structure of final LiteSwinIR is provided in Fig. 2.

C. Design of Knowledge Distillation

To get ready for the knowledge distillation training, this study first obtains the pre-trained Lightweight SwinIR model as the teacher model and freezes its parameters to ensure they are not updated during distillation training.

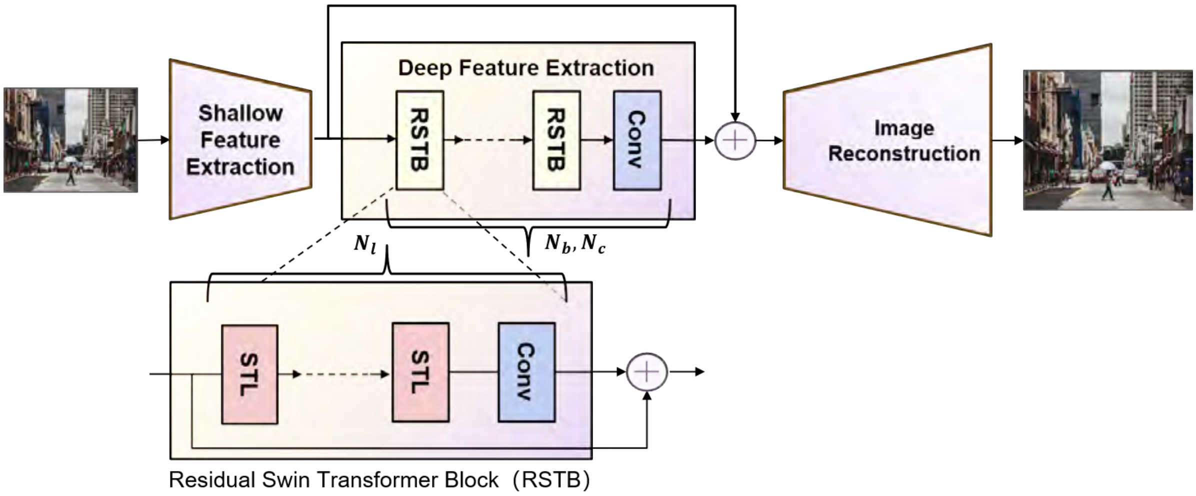

Fig 2: The architecture of LiteSwinIR, with reference from [7]. Here $\hat{N}_c$, $\hat{N}_l$ and $\hat{N}_b$ represent the number of channels, layers and RSTB blocks, STL means the Swin Transformer Layer.

Use LiteSwinIR as the student model to ensured that the student model has similar architecture with the teacher model, avoid causing errors in input images.

During training, the student model learns the features of the teacher model. This enables the final model to maintain reconstruction quality while achieving good edge and texture processing capabilities, also with less parameters.

In terms of dataset setup, the ACDC dataset provides road conditions in bad weather and high-resolution images with weather labels, which is ideal for knowledge distillation in multi-weather conditions. The specific data preprocessing is described in Section IV.A.

The design of the entire knowledge distillation training process is based on the framework in [25], which is divided into forward propagation and back propagation. In the first forward propagation, the low-resolution images will be input to student model and teacher model to generate the super-resolution outputs I_{tea} and I_{stu}. The reconstruction loss between the student output and the ground-truth image, as well as the distillation loss between the student model and the teacher model will be computed. The weather loss will also be calculated according to the different weather labels. During back propagation, the total loss will be calculated based on the different loss term weights and the parameters of the student model will be updated. The teacher model remains unchanged during this process.

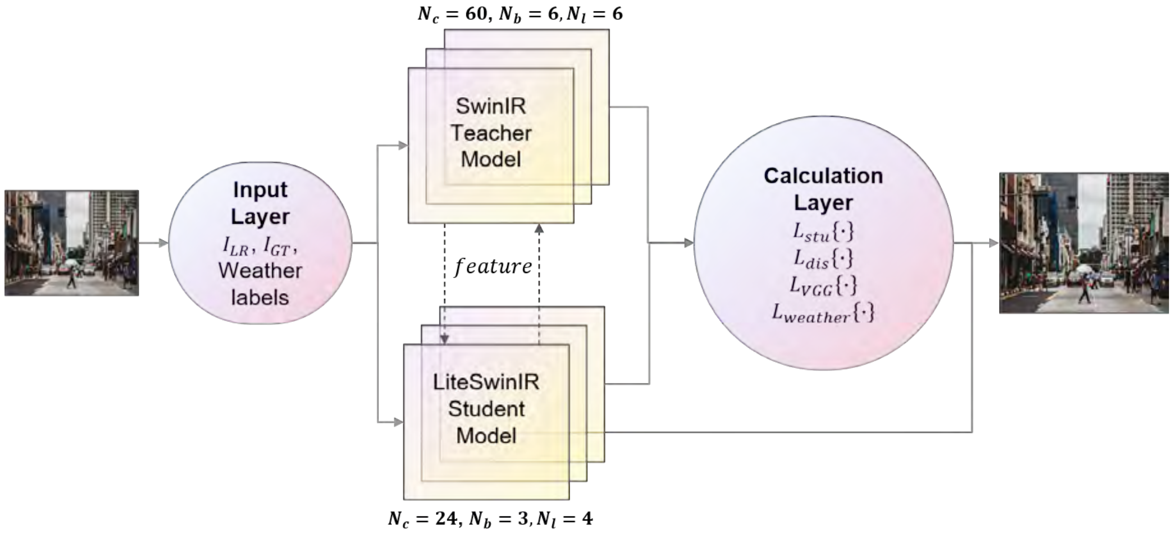

Fig 3: The knowledge distillation structure, where L_stu is the student loss, L_dis is the distillation loss.

D. Design of Weather-adaptive Loss Functions

D-1. Overall Function Expression

In this study, a weather adaptive knowledge distillation loss function is designed to use the student task loss and distillation loss as the basis, combining the perceptual loss and weather loss terms, different loss term weights are also set to balance the total loss. In different weather conditions, specific loss term weights will be used to achieve optimisation for different weather conditions. The overall loss function equation is given by Eq. 4

$$\mathcal{L}_{\text{total}} = \alpha(\omega)\mathcal{L}_{\text{stu}} + \beta(\omega)\mathcal{L}_{\text{feature}} + \theta(\omega)\mathcal{L}_{\text{output}} + \delta(\omega)\mathcal{L}_{\text{VGG}} + \mathcal{L}_{\text{weather}}(\omega) \tag{4}$$

where ω denotes the weather type and $\alpha(\omega)$, $\beta(\omega)$, $\theta(\omega)$, $\delta(\omega)$ represent the corresponding adaptive weighting coefficients.

D-2. Definition of Each Loss Term

1. L_{stu} is the basic **reconstruction loss**, which is the L1 loss defined in Eq. 5.

$$\mathcal{L}_{\text{stu}} = \frac{1}{CHW} \sum_{c=1}^{C} \sum_{h=1}^{H} \sum_{w=1}^{W} |S(x)_{c,h,w} - GT(x)_{c,h,w}| \tag{5}$$

2. L_{feature} and L_{output} is the **distillation loss**, detail expressions are:

$$\mathcal{L}_{\text{feature}} = \frac{1}{CHW} \sum_{c=1}^{C} \sum_{h=1}^{H} \sum_{w=1}^{W} \left(F_s(x)_{c,h,w} - F_t(x)_{c,h,w}\right)^2 \tag{6}$$

$$\mathcal{L}_{\text{output}} = \frac{1}{CHW} \sum_{c=1}^{C} \sum_{h=1}^{H} \sum_{w=1}^{W} \left(S(x)_{c,h,w} - T(x)_{c,h,w}\right)^2 \tag{7}$$

   Here, $F_s(x)$ and $F_t(x)$ represent the intermediate features of the student and teacher models, respectively, while $T(x)$ denotes the output of the teacher model.

3. L_{VGG} is defined in Eq. 8, where $\phi^{(n)}(\cdot)$ is from the pretrained VGG-19 network.

$$\mathcal{L}_{\text{VGG}} = \frac{1}{C_n H_n W_n} \sum_{c=1}^{C_n} \sum_{h=1}^{H_n} \sum_{w=1}^{W_n} \left|\phi^{(n)}(S(x))_{c,h,w} - \phi^{(n)}(GT(x))_{c,h,w}\right| \tag{8}$$

4. L_{weather} is the **specific weather loss term** only calculated in fog, snow and rain conditions.

D-3. Fog

Foggy conditions reduce the edge clarity of an image and weaken its structural features[27]. The Sobel operator can extracts edge information by computing the horizontal and vertical gradients of the image.

The Sobel operator detects image edges using two 3×3 convolution kernels G_x and G_y applied in the horizontal and vertical directions.

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \qquad G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \tag{9}$$

Convolve the convolution kernel with the image to get the gradient of the image in the horizontal and vertical directions, $I_x)$ and I_y.

$$\begin{aligned} I_x &= \text{ Image } * G_x \\ I_y &= \text{ Image } * G_y \end{aligned} \tag{10}$$

The gradient magnitude G is given by:

$$G = \sqrt{I_x^2 + I_y^2} \tag{11}$$

By convolving the gradient magnitude G with the image, we can obtain the edge feature map.

In the fog loss term, the Sobel operator is applied to the student model output and the ground truth image to obtain their edge feature maps Sobel$(S(x))c, h, w$ and Sobel$(GT(x))c, h, w.$ Then the absolute value sum of the pixel differences between the two feature maps are calculated:

$$\mathcal{L}_{\text{fog}} = \frac{1}{CHW} \sum_{c=1}^{C} \sum_{h=1}^{H} \sum_{w=1}^{W} |\text{Sobel}(S(x))c, h, w - \text{Sobel}(GT(x))c, h, w| \tag{12}$$

Setting the fog loss in this way allows the model to recover edge features similar to those of the clear target image, improving the visual clarity of fog-degraded inputs.

D-4. Snow and Rain

The Structure Similarity Index Measure (SSIM) loss can consider the structural similarity of the global image, which is given by:

$$\text{SSIM}(x, y) = \frac{(2\mu_x \mu_y + C_1)(2\sigma_{xy} + C_2)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)} \tag{13}$$

where μ is the mean and σ^2 is the variance, σ_{xy} is the covariance, C_1, C_2 are constants.

Rainy and snowy weather loss is designed based on SSIM loss.

$$\mathcal{L}_{\text{weather}} = 1 - \text{SSIM}(S(x), GT(x)) \tag{14}$$

The noise caused by rain and snow can be better suppressed by setting structural similarity loss [8].

D-5. Weather-adaptive Weights

The weighting factors for different weather conditions are as follows:

Table 2: Weighting Factors For Different Weather Conditions

weather conditions	α	β	θ	δ
fog	1.2	0.15	0.25	0.3
snow and rain	1.0	0.2	0.15	0.2
night	0.9	0.1	0.3	0.15

The weather weights are designed with inspiration from [19, 20]. The L1 loss and perceptual loss are enhanced in foggy weather to improve the quality of distant objects recovery. The distillation coefficient is enhanced in rainy and snowy weather to better preserve the detailed structure of the image. At night, the perceptual loss is reduced and the output loss is enhanced in order to better handle low-light conditions. During training, the loss function weights are dynamically adjusted according to the weather type of the input images.

Ultimately, this study achieves a balanced scheme by combining the L1 reconstruction loss from the student model and the distillation loss from the teacher model through the knowledge distillation framework. It aims to preserve reconstruction accuracy and indirectly acquire perceptual knowledge through the teacher model. Additionally, it allows for weather-specific optimisation, helping the lightweight model adapt to diverse weather conditions.

IV. Experiments

A. *Experimental Setup*

A-1. Dataset and Preprocessing

This study uses ACDC road-scene images to construct paired data for a controlled super-resolution experiment. The high-resolution target images are selected from the clear-weather reference images in ACDC. The corresponding low-resolution inputs are generated from these high-resolution targets using $\times 2$ downsampling and additional weather-related degradation operations. Therefore, the reported PSNR/SSIM values measure reconstruction quality with respect to the clear high-resolution ground-truth images, rather than with respect to weather-affected targets.

The constructed training set contains 1568 images: 784 clear high-resolution 1920$\times$1080 reference images and 784 low-resolution 960$\times$540 degraded inputs generated via $\times 2$ bicubic downsampling. The low-resolution inputs are divided into three degradation categories: images 1–255 are fog degradations, 256–495 are night degradations, and 496–784 are snow and rain degradations. The specific degradation parameters are given in Algorithm 1, where α is the contrast parameter and β is the brightness parameter.

Algorithm 1 Image Degradation Preprocessing for Different Weather Conditions

```
for each weather ∈ {fog, snow and rain, night}
    Select degradation function based on weather:
if weather = fog
    img_lr ← Bicubic(scale = 2)
    img_lr ← AdjustContrast(α = 0.7, β = 20)
else if weather = snow and rain
    img_lr ← Bicubic(scale = 2)
    img_lr ← AdjustContrast(α = 0.8, β = −20)
    Apply random snow effect with high-intensity pixels
else if weather = night
    img_lr ← Bicubic(scale = 2)
    img_lr ← AdjustContrast(α = 0.6, β = −30)
else
    img_lr ← Bicubic(scale = 2)
end if
    Add img_lr to LR_images collection
end for
```

Examples of the clear high-resolution target images and the generated weather-degraded low-resolution inputs are shown in Fig. 4.

Fig 4: Examples of the downsampling and degradation process.

Weather labels are stored in `.txt` format and verified against the corresponding degradation category. Images with inconsistent labels are excluded. Table 3 shows the label format used for all splits.

Table 3: Weather Labels Format

Filename	Weather	Split	Type
000001.png	fog	train	HR
000002.png	fog	train	HR
000494.png	night	train	HR
000495.png	night	train	HR
000783.png	snowandrain	train	HR
000784.png	snowandrain	train	HR

To satisfy the patch-based training requirement of SwinIR, sub-images are extracted following the strategy in [28]. High-resolution images (1920×1080) are cropped into 480×480 patches with a stride of 240 pixels (overlap ratio 0.5); the corresponding low-resolution images are cropped to 240×240 patches with a stride of 120 pixels. During training, a random crop of size 128×128 (`gt_size`=128) is applied to each patch pair before being fed to the network.

A-2. Implementation Details

The LiteSwinIR student model is trained using the weather-adaptive knowledge distillation framework described in Section III.C. The pretrained SwinIR_LW model is loaded as the frozen teacher, and the student model is initialised with architecture parameters $N_b = 3$, $N_l = 4$, $N_c = 24$ as defined in Section III.B (see Algorithm 2).

Algorithm 2 Initialisation of Teacher and Student Models

Load pretrained SwinIR teacher model:
    upscale = 2, image_size = 64, window_size = 8
    depth = [6,6,6,6], embed_dim = 60, num_heads = [6,6,6,6]
    upsampler = `pixelshuffledirect`
Freeze all teacher parameters; set to evaluation mode.

Initialise student model:
    upscale = 2, image_size = 64, window_size = 8
    depth = [4,4,4], embed_dim = 24, num_heads = [6,6,6]
    upsampler = `pixelshuffledirect`

Training uses the Adam optimiser with an initial learning rate of 1×10^{-4}, $\beta = (0.9, 0.999)$, and no weight decay. The model is trained for 50,000 iterations (10 epochs $\times$ 5,000 iterations) with a batch size of 4 on a single GPU. The learning rate is decayed by a factor of $\gamma = 0.5$ at the 25,000th, 37,500th, and 45,000th iterations via a MultiStepRestartLR scheduler. Initial loss weights are: student loss 1.0, feature distillation loss 0.5, output distillation loss 0.5, and VGG perceptual loss 1.0. The detailed configuration is summarised in Algorithm 3.

Algorithm 3 Training Configuration

Total iterations = 50,000; **batch size** = 4

Optimiser: Adam
    learning rate = 1×10^{-4}, betas = (0.9, 0.999), weight decay = 0

Scheduler: MultiStepRestartLR
    milestones = [25000, 37500, 45000], γ = 0.5

Validation is performed every 1,000 iterations on four weather types (sunny, fog, snow and rain, night) sampled from the ACDC validation split to monitor PSNR and the composite loss value.

A-3. Evaluation Metrics

Two standard image quality metrics are used throughout. **PSNR** measures pixel-level fidelity between the predicted image and the ground-truth reference. **SSIM** [29] captures structural similarity by jointly considering luminance, contrast, and structure. In addition, the total number of model parameters is reported as an indicator of deployment efficiency.

A-4. Baselines and Evaluation Protocol

The proposed LiteSwinIR is compared against the following baseline models, selected to cover GAN-based, ResNet-based, and lightweight Transformer-based paradigms.

- **SwinIR_LW** [7]: The teacher model and direct baseline for LiteSwinIR. Comparing against it quantifies the cost of the compression step.
- **ESRGAN** [12]: A GAN-based model known for high perceptual quality and fine detail reconstruction.
- **EDSR** [11]: A strong ResNet-based model providing high pixel-level accuracy.
- **IMDN** [13]: A lightweight SR model with low parameter count, used to benchmark the efficiency of LiteSwinIR.
- **SwinIRmini** [14]: A compact distillation-based model with a similar training architecture to LiteSwinIR, used as the reference for the ablation study on the weather-adaptive loss function.

Pre-trained weights for all classical baselines use DIV2K [21] as training data, a clean high-definition dataset without weather degradation. Performance comparisons should therefore be interpreted as a domain-shift evaluation rather than a controlled benchmark.

The test set is a small-scale evaluation subset drawn from the ACDC test split, containing 10 images per weather type (fog, snow and rain, night). Weather labels are withheld from all models during inference to simulate unlabeled inputs in practical deployment.

B. Overall Performance

B-1. Training Convergence

The validation PSNR curve during training is shown in Fig. 5, where the model rises rapidly from 19.57 dB to approximately 30 dB in early epochs and reaches a peak of 34.39 dB at iteration 45,000, consistent with the learning rate decay schedule.

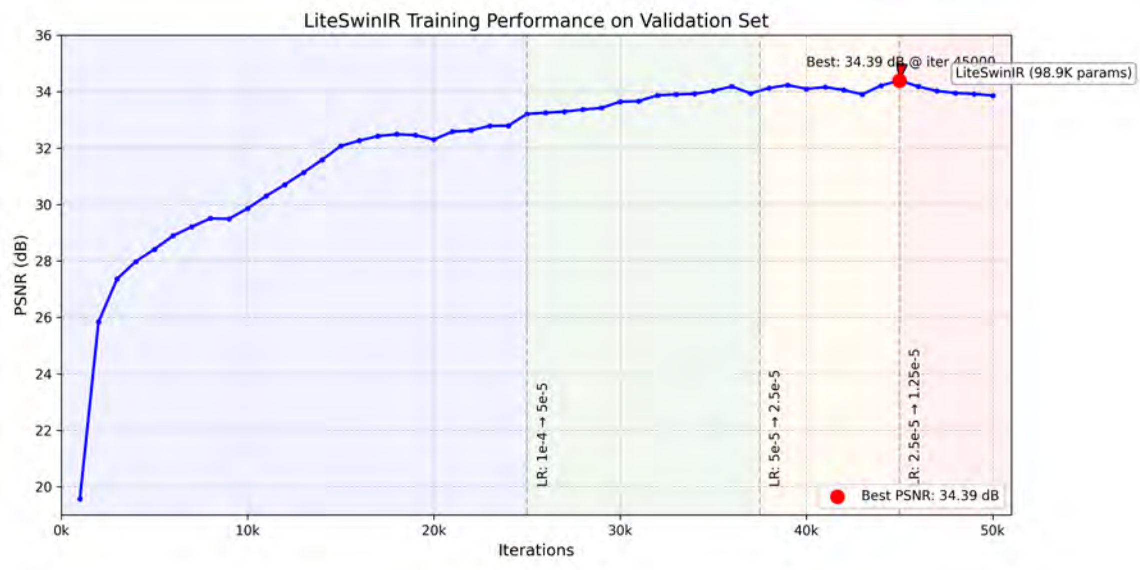

Fig 5: LiteSwinIR validation PSNR curve over 50,000 training iterations.

B-2. Quantitative Comparison

Table 4: Performance Comparison of Super-Resolution Methods under Different Weather Conditions

Methods	Fog		Snow & Rain		Night		Average	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
SwinIR_LW (teacher)	19.44	0.9109	14.18	0.4794	10.77	0.3828	14.80	0.5910
EDSR	**19.48**	**0.9173**	14.15	0.4646	**10.78**	**0.3862**	14.80	0.5893
IMDN	19.46	0.9113	14.17	0.4342	10.77	0.3838	14.80	0.5764
ESRGAN	19.36	0.8600	**15.92**	**0.5502**	10.75	0.3826	**15.34**	**0.5976**
LiteSwinIR (ours)	**29.68**	**0.9359**	**28.50**	**0.8546**	**16.86**	**0.8288**	**25.01**	**0.8731**
Gain vs. 2nd best	+10.20	+0.0186	+12.58	+0.3044	+6.09	+0.4460	+9.67	+0.2755

The quantitative results on the adverse-weather test subset are presented in Table 4. LiteSwinIR outperforms all baselines across all weather conditions on this test subset. Under fog, it achieves

29.68 dB PSNR, a gain of 10.20 dB over the next best model. Under snow and rain, PSNR reaches 28.50 dB (SSIM 0.8546), the largest absolute gain of 12.58 dB. Under night conditions, PSNR reaches 16.86 dB (SSIM 0.8288), the smallest gain, reflecting the greater challenge posed by low-light degradation. Averaged across all weather types, LiteSwinIR achieves 25.01 dB PSNR and 0.8731 SSIM.

The relative weakness on night and snow-rain conditions suggests that the current design leaves room for improvement on low-light and particle-type degradation. Regarding efficiency, the proposed model reduces the parameter count from 878K (SwinIR_LW) to 98.9K, an 88.7% reduction, supporting deployment on resource-constrained platforms.

B-3. Qualitative Comparison

Fig. 6 presents zoomed-in visual comparisons (300×300 crops) across the three weather conditions. Low-resolution inputs are generated synthetically from clear reference images.

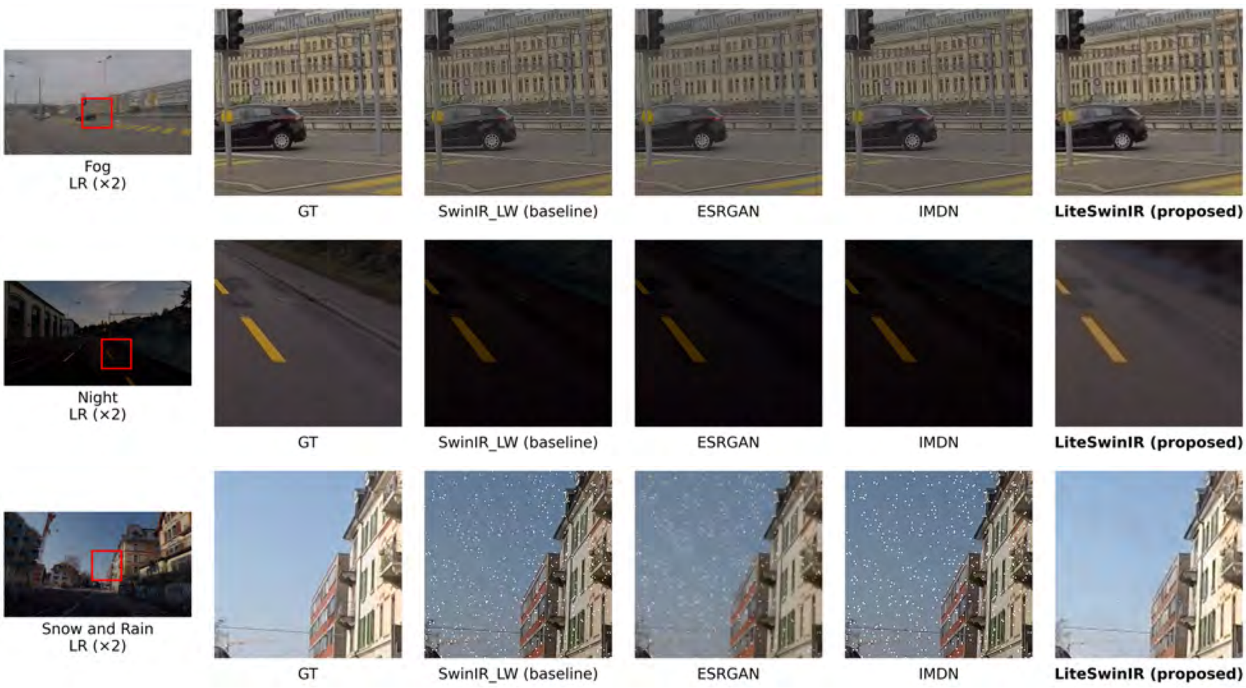

Fig 6: Zoomed-in super-resolution comparisons under fog (top), night (middle), and snow/rain (bottom) conditions.

Under fog, LiteSwinIR better suppresses haze-like degradation, retains clearer image structure, and recovers colors closer to the clear reference. Under night conditions, the model restores image brightness and enhances dark-region details while preserving structural consistency. Under snow and rain, it effectively suppresses synthetic snowflake and rain-streak noise, recovering fundamental structural features of the clear target. Remaining limitations include occasional blurred details and minor artifacts in severely degraded local regions, particularly under rain/snow and night conditions. Given that downstream perception tasks such as detection and segmentation prioritise noise removal and structural fidelity over fine texture [30, 31], these artifacts are unlikely to substantially impair downstream utility.

C. Ablation Study

C-1. Effect of Weather-Adaptive Loss

To isolate the contribution of the weather-adaptive loss function, LiteSwinIR is compared against the fine-tuned SwinIRmini model [14], which shares a similar knowledge-distillation architecture but uses a standard loss combining L1 reconstruction loss and Laplace-based distillation loss without weather-specific terms. SwinIRmini is fine-tuned on the ACDC training set for 10,000 iterations to ensure fair adaptation to adverse-weather inputs.

Table 5: Comparison of Models With and Without Weather-Adaptive Loss

Methods	**Fog**		**Snow & Rain**		**Night**		**Average**	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
SwinIRmini (standard loss)	25.50	0.8545	21.93	0.6951	14.71	0.7303	20.71	0.7600
LiteSwinIR (weather-adaptive loss)	**29.68**	**0.9359**	**28.50**	**0.8546**	**16.86**	**0.8288**	**25.01**	**0.8731**
Gain	+4.18	+0.0814	+6.57	+0.1595	+2.15	+0.0985	+4.30	+0.1131

The weather-adaptive loss yields consistent gains across all conditions. The largest improvement occurs under snow and rain (+6.57 dB PSNR, +0.1595 SSIM), where enhanced feature loss and SSIM-based structural terms encourage texture and structural consistency. Under fog, the Sobel edge loss term contributes a +4.18 dB gain. Night shows the smallest improvement (+2.15 dB), as no dedicated night loss term is used beyond a reduced perceptual loss weight, suggesting that low-light degradation remains an open design challenge.

C-2. Qualitative Comparison

Fig. 7 shows zoomed-in visual comparisons between SwinIRmini and LiteSwinIR under the three weather conditions.

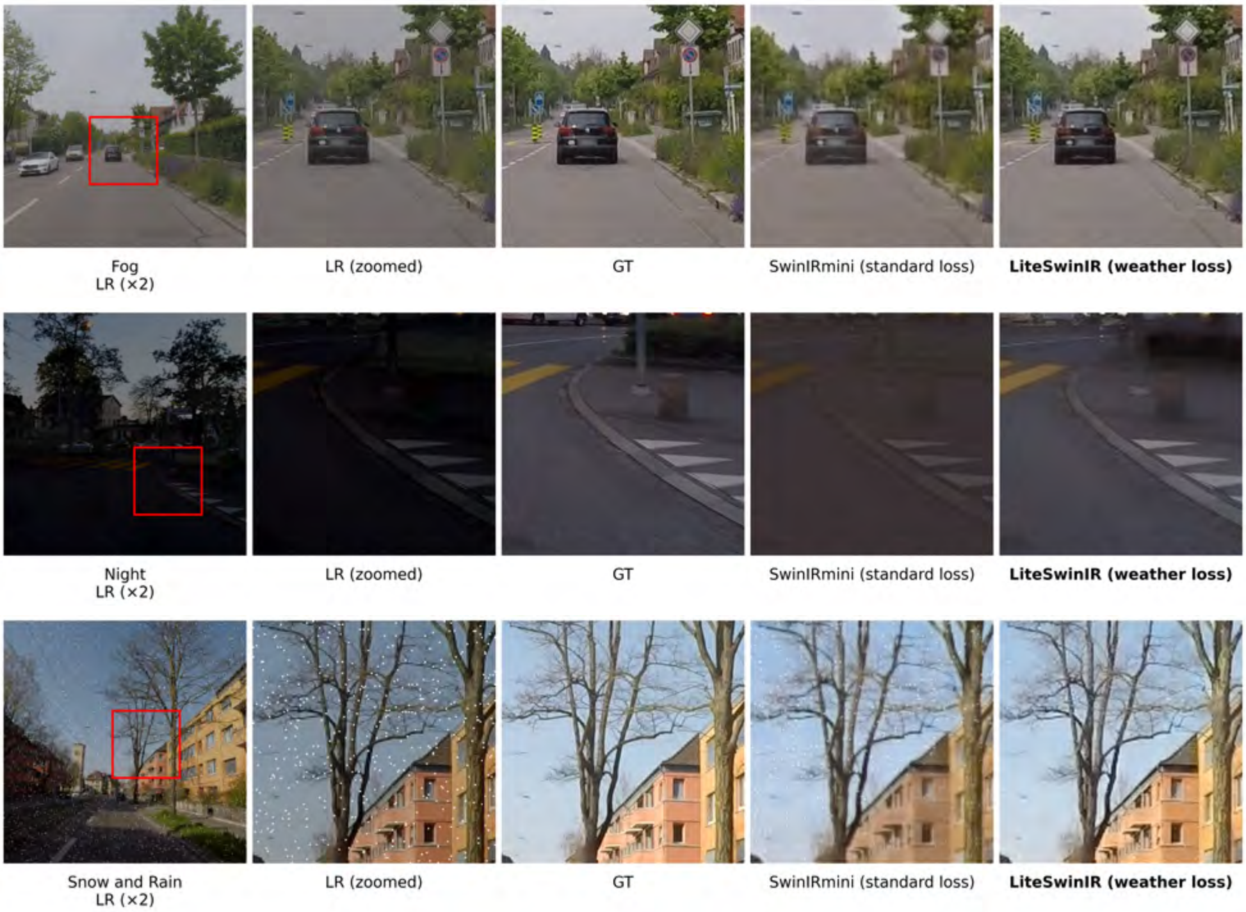

Fig 7: Zoomed-in comparisons between standard-loss SwinIRmini and the proposed weather-adaptive-loss LiteSwinIR under fog (top), night (middle), and snow/rain (bottom) conditions.

Under fog, the standard-loss model leaves residual haze and produces unclear distant objects, whereas LiteSwinIR achieves closer color fidelity and sharper edge definition in vehicle and road regions. Under night conditions, the standard-loss model fails to recover dark-region details, leaving road signage illegible; the weather-adaptive model improves overall exposure and sign clarity. Under snow and rain, the standard-loss model struggles to suppress sparse snow artifacts, while LiteSwinIR removes most snowflakes and rain streaks while preserving structural and textural consistency.

These results provide preliminary evidence that weather-specific loss terms improve reconstruction under targeted degradation conditions. A fully controlled ablation study with larger test sets would be needed to quantify the independent contribution of each loss term.

V. Conclusion and Future Work

A. Conclusions

This study presented LiteSwinIR, a lightweight weather-adaptive image super-resolution framework for safety-critical edge perception. The work addressed two connected challenges: restoring degraded road images under adverse weather and reducing the model size of deep image restoration models for resource-constrained deployment. By combining model compression, knowledge distillation, and weather-adaptive loss design, the proposed framework aims to balance reconstruction quality, robustness, and efficiency.

The experimental evaluation on a small-scale adverse-weather subset derived from the ACDC dataset shows that the proposed model maintains useful reconstruction ability under fog, snow, rain, and night conditions. Quantitative comparisons indicate improved PSNR and SSIM performance over selected classical baselines in this constructed setting, while the architecture design substantially reduces model parameters. These results support the central claim that weather-aware restoration and lightweight model design can be jointly considered for edge-oriented perception tasks, although larger-scale and fully controlled evaluations are still needed.

From the perspective of Sustainable AI, the contribution of this work lies not only in improving image reconstruction quality, but also in treating efficiency as a design objective. This is important for smart mobility systems where perception modules may need to run close to sensors under latency, energy, and hardware constraints. More reliable safety-critical perception can support safer intelligent transportation environments, especially when perception models are expected to operate on resource-constrained edge platforms.

The study also has limitations. First, the evaluation is based on a selected small-scale adverse-weather subset rather than a full benchmark or real-time deployment on embedded hardware. Second, the model still shows weaknesses in some severely degraded local regions, especially under rainy, snowy, and night conditions. Third, the comparisons with classical baselines should be interpreted as adaptation results under the constructed adverse-weather setting rather than as universal superiority over the original baseline models. Future work should therefore include larger test sets, real-world edge deployment tests, broader weather scenarios, and downstream perception tasks such as detection or segmentation to better quantify the safety benefits of the restored images.

B. Future Work

B-1. Real-time Implementation

This study does not provide a complete deployment analysis on edge devices, but practical embedded inference remains an important direction. Future work can improve inference speed by further optimizing the model structure and compressing the model size. The model can also be modified according to image formats commonly used in edge-enabled devices; for example, it may be adapted to process images without extracting sub-images [28].

B-2. Extending More Weather Types

The model proposed in this study is only optimized for single weather categories such as fog, snow and rain, and night. A future direction is to adapt the model to more complex weather variations and mixed weather types, such as sand, dust, thunderstorms, and compound adverse conditions [32, 33]. In addition, previous work has categorized the same weather condition into different intensity levels, such as clear, moderate fog, and heavy fog conditions, which could support more fine-grained weather-adaptive restoration [34].

B-3. Constructing Real-world Datasets

This study aims to recover weather-degraded road images; however, image super-resolution training requires paired images at different resolutions, so the weather degradation used in this work is synthetically added to clear reference images. In the future, model performance can be further tested by constructing a test set with real weather-affected images and adding both subjective and objective visual evaluation to increase the credibility of the reported performance.

B-4. Introducing Attention Mechanisms

The proposed model is not explicitly optimized for all local details, which could be further addressed by attention mechanisms. By introducing channel attention, such as the DAF-Net design [17], weather weights could be adjusted based on special feature channels to capture pixel dependencies more accurately. This may help the model focus on localized regions and improve the reconstruction of fine details [16].

References

[1] D. C. Lepcha, B. Goyal, A. Dogra, and V. Goyal, "Image super-resolution: A comprehensive review, recent trends, challenges and applications," *Information Fusion*, vol. 91, pp. 230–260, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1566253522001762

[2] Y. Zhang, A. Carballo, H. Yang, and K. Takeda, "Perception and sensing for autonomous vehicles under adverse weather conditions: A survey," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 196, pp. 146–177, 2023.

[3] E. Strubell, A. Ganesh, and A. McCallum, "Energy and policy considerations for deep learning in nlp," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics, 2019, pp. 3645–3650. [Online]. Available: https://aclanthology.org/P19-1355/

[4] R. Schwartz, J. Dodge, N. A. Smith, and O. Etzioni, "Green ai," *Communications of the ACM*, vol. 63, no. 12, pp. 54–63, 2020.

[5] Z. Zhou, X. Chen, E. Li, L. Zeng, K. Luo, and J. Zhang, "Edge intelligence: Paving the last mile of artificial intelligence with edge computing," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1738–1762, 2019.

[6] Z. Wang, J. Chen, and S. C. Hoi, "Deep learning for image super-resolution: A survey," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 10, pp. 3365–3387, 2020.

[7] J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, and R. Timofte, "Swinir: Image restoration using swin transformer," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 1833–1844.

[8] S. Karavarsamis, I. Gkika, V. Gkitsas, K. Konstantoudakis, and D. Zarpalas, "A survey of deep learning-based image restoration methods for enhancing situational awareness at disaster sites: The cases of rain, snow and haze," *Sensors*, vol. 22, no. 13, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/13/4707

[9] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE transactions on acoustics, speech, and signal processing*, vol. 29, no. 6, pp. 1153–1160, 1981.

[10] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 2, pp. 295–307, 2015.

[11] B. Lim, S. Son, H. Kim, S. Nah, and K. Mu Lee, "Enhanced deep residual networks for single image super-resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017.

[12] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, "Esrgan: Enhanced super-resolution generative adversarial networks," in *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, September 2018.

[13] Z. Hui, X. Gao, Y. Yang, and X. Wang, "Lightweight image super-resolution with information multi-distillation network," in *Proceedings of the 27th acm international conference on multimedia*, 2019, pp. 2024–2032.

[14] Y. Jiang, J. Nawała, F. Zhang, and D. Bull, "Compressing deep image super-resolution models," in *2024 Picture Coding Symposium (PCS)*, 2024, pp. 1–5.

[15] J. Gou, B. Yu, S. J. Maybank, and D. Tao, "Knowledge distillation: A survey," *International*

Journal of Computer Vision, vol. 129, no. 6, pp. 1789–1819, 2021.

[16] N. J. Zakaria, M. I. Shapiai, R. A. Ghani, M. N. M. Yassin, M. Z. Ibrahim, and N. Wahid, "Lane detection in autonomous vehicles: A systematic review," *IEEE Access*, vol. 11, pp. 3729–3765, 2023.

[17] X. Hu, C.-W. Fu, L. Zhu, and P.-A. Heng, "Depth-attentional features for single-image rain removal," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[18] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[19] X. Fu, J. Huang, D. Zeng, Y. Huang, X. Ding, and J. Paisley, "Removing rain from single images via a deep detail network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[20] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "Aod-net: All-in-one dehazing network," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

[21] E. Agustsson and R. Timofte, "Ntire 2017 challenge on single image super-resolution: Dataset and study," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017.

[22] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[23] F. Yu, H. Chen, X. Wang, W. Xian, Y. Chen, F. Liu, V. Madhavan, and T. Darrell, "Bdd100k: A diverse driving dataset for heterogeneous multitask learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[24] C. Sakaridis, H. Wang, K. Li, R. Zurbrügg, A. Jadon, W. Abbeloos, D. O. Reino, L. V. Gool, and D. Dai, "Acdc: The adverse conditions dataset with correspondences for robust semantic

driving scene perception," 2024. [Online]. Available: https://arxiv.org/abs/2104.13395

[25] C. Morris, D. Danier, F. Zhang, N. Anantrasirichai, and D. R. Bull, "St-mfnet mini: Knowledge distillation-driven frame interpolation," in *2023 IEEE International Conference on Image Processing (ICIP)*, 2023, pp. 1045–1049.

[26] T. Chen, T. Ding, B. Ji, G. Wang, Y. Shi, J. Tian, S. Yi, X. Tu, and Z. Zhu, "Orthant based proximal stochastic gradient method for l_1-regularized optimization," in *Machine Learning and Knowledge Discovery in Databases*, F. Hutter, K. Kersting, J. Lijffijt, and I. Valera, Eds. Cham: Springer International Publishing, 2021, pp. 57–73.

[27] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "Dehazenet: An end-to-end system for single image haze removal," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5187–5198, 2016.

[28] X. Wang, L. Xie, K. Yu, K. C. Chan, C. C. Loy, and C. Dong, "BasicSR: Open source image and video restoration toolbox," https://github.com/XPixelGroup/BasicSR, 2022.

[29] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[30] C. Hu and H. Wang, "Enhancing rainy weather driving: Deep unfolding network with pgd algorithm for single image deraining," *IEEE Access*, vol. 11, pp. 57 616–57 626, 2023.

[31] M. Hassaballah, M. A. Kenk, K. Muhammad, and S. Minaee, "Vehicle detection and tracking in adverse weather using a deep learning framework," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4230–4242, 2021.

[32] R. Li, R. T. Tan, and L.-F. Cheong, "All in one bad weather removal using architectural search," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[33] J. M. J. Valanarasu, R. Yasarla, and V. M. Patel, "Transweather: Transformer-based restoration of images degraded by adverse weather conditions," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022, pp. 2353–2363.

[34] I. Ogunrinde and S. Bernadin, "A review of the impacts of defogging on deep learning-based object detectors in self-driving cars," *SoutheastCon 2021*, pp. 01–08, 2021.