

Dual Degradation Representation for Joint Deraining and Low-Light Enhancement in the Dark

Xin Lin*, Jingtong Yue*, Sixian Ding, Chao Ren†, Lu Qi and Ming-Hsuan Yang, *Fellow, IEEE*

Abstract—Rain in the dark poses significant challenges to deploying real-world applications such as autonomous driving, surveillance systems, and night photography. Existing low-light enhancement or deraining methods struggle to brighten low-light conditions and remove rain simultaneously. Cascade approaches, like “deraining followed by low-light enhancement” and vice versa, often result in problematic rain patterns or overly blurred and overexposed images. To address these challenges, we introduce a novel two-stage model called L^2 RIRNet, which innovatively integrates low-light enhancement and deraining into a single framework in real-world settings. Our model comprises two key components: a Dual Degradation Representation Network (DDR-Net) and a Restoration Network. The DDR-Net independently learns degradation representations for luminance effects in dark areas and rain patterns in light areas, which are constrained by dual degradation loss and have not been discussed in the previous methods. The Restoration Network restores the degraded image using a Fourier Detail Guidance (FDG) module, which focuses on texture details in frequency and spatial domains to inform the restoration process and leverages near-rainless detailed images. Furthermore, we contribute a dataset containing both synthetic and real-world low-light-rainy images. Extensive experiments demonstrate that our L^2 RIRNet performs favorably against existing methods in synthetic and complex real-world scenarios. All the code and dataset can be found in https://github.com/linxin0/Low_light_rainy.

Index Terms—Low-light-rainy image restoration, Dual degradation loss and representation learning, Detailed image guidance.

I. INTRODUCTION

Rainy nights are common weather conditions that affect many real-world activities, such as autonomous driving, surveillance systems, and photography. While existing single-image rain removal [1], [3]–[9], low-light enhancement [2], [10]–[17] and image restoration [18]–[24] methods have made significant progress in addressing specific degradations, their effectiveness in solving the joint-degradation low-light-rainy task remains limited due to the fact that rain patterns are hidden in the darkness and hard to capture.

*: equal contribution, †: Corresponding author

This work was supported by the National Natural Science Foundation of China under Grant 62171304 and partly by the National Science Foundation of Sichuan Province under Grant 2024NSFSC1423, and Cooperation Science and Technology Project of Sichuan University and Dazhou City under Grant 2022CDDZ-09.

Xin Lin, Jingtong Yue, Sixian Ding, and Chao Ren are with the College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China (e-mail: linxin020826@gmail.com; yuejingtong@stu.scu.edu.cn; dingsixian@stu.scu.edu.cn; chaoren@scu.edu.cn).

Lu Qi and Ming-Hsuan Yang are with the University of California at Merced, Merced, CA 95343 USA (e-mail: qqlu1992@gmail.com; mhyang@ucmerced.edu).

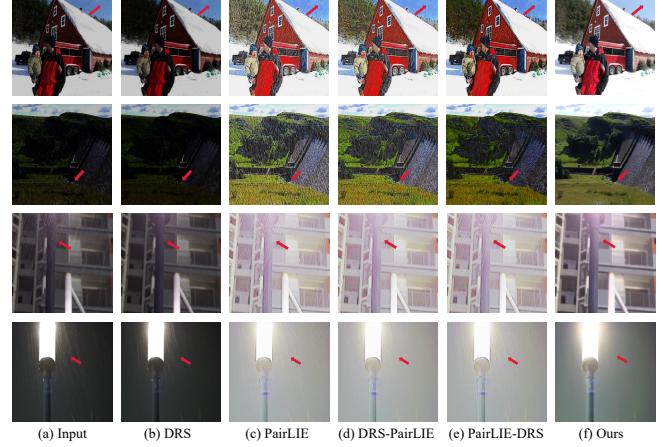


Fig. 1. **Motivation.** Visual results of existing low-light enhancement and deraining methods on synthetic (1st and 2nd rows) and real-world (3rd and 4th rows) low-light-rainy images. (a) Input; (b) DRS [1]; (c) PairLIE [2]; (d) DRS [1] + PairLIE [2]; (e) PairLIE [2] + DRS [1]; (f) L^2 RIRNet (Ours). The single-task models can only enhance low light or remove rain, while cascade methods may result in issues such as overexposure, underexposure, residual rain patterns, and blurring. The results highlight the challenges of existing approaches in effectively addressing low light and rain conditions. Compared with other methods, our L^2 RIRNet can achieve better results.

Considering the large amount of work required for this joint restoration task, we conduct preliminary experiments using a cascade pipeline of two single-task models, but the results were still unsatisfactory. As illustrated in Fig. 1(b), the deraining method DRS [1], trained on the Rain12600 [25] dataset, faces challenges in dark areas where background information is unavailable. In Fig. 1(c), the low-light enhancement approach PairLIE [2] makes rain patterns more visible but fails to remove them. As shown in Figs. 1(d) and 1(e), combining DRS [1] and PairLIE [2] still presents challenges such as over- or under-exposure in specific areas. Specifically, by utilizing the cascade approach, i.e., deraining followed by low light enhancement, the rain streaks become more apparent compared to the input image. This is because the deraining network alone struggles to remove the hidden rain streaks under low-light conditions, leaving them visible after enhancement. Conversely, although the cascade method of low-light enhancement and then deraining can reduce some rain streaks, it leads to other issues. The low-light enhancement model makes the rain streaks more apparent, exposing them to the deraining model. This results in the brightness of the rain streaks appearing abnormally on the image compared to normal rainy images, making it difficult for the deraining network to handle. Meanwhile, it also causes some problems,

such as overexposure or underexposure, further affecting rain-streaking removal. Based on the above discussions, these cascade methods using existing networks do not work well in real-world scenarios, where the distribution of rain streaks is more complex and challenging to capture. Thus, simultaneous low-light enhancement and rain removal is a non-trivial task that needs to be elaborately designed.

To address the challenges mentioned above, we analyze the physical properties of real-world low-light-rainy images and observe distinct degrading representations at different locations. The dark regions, characterized by low illumination, visually obscure rain information, while bright regions exhibit visible rain patterns. That is, different regions suffer from different types of degradation. If we handle them uniformly, it will confuse the model. Therefore, we propose an effective Low-Light-Rainy Image Restoration Network named $L^2RIRNet$ that separately learns representations of rain streaks and low-light conditions. The framework is designed to handle low-light-rainy scenarios by learning hierarchical degradation representations. It mainly contains two key components: Dual Degradation Representation Network (DDR-Net) and Restoration-Net. Moreover, we construct an LLR dataset, comprising two components: synthesized image pairs and real-world images, to train the model.

The novel two-branch representation learning DDR-Net is divided into a rainy branch, which learns rain pattern representations in bright regions, and a light branch, which learns low-light representations in dark regions. The DDRNet facilitates representation learning constrained by presented dual degradation loss (DDLoss).

Rain streaks are hidden in the darkness. Although they are invisible to the naked eye, we can detect their objective existence, and they have a significant impact on image restoration tasks after enhancing brightness with low-light enhancement algorithms. Therefore, removing rain streaks is one of the key points to the success of the low-light-rainy image restoration task. Given the high-frequency characteristics of rain patterns and their impact on the low-frequency content contour structure, we employ different channels for low-frequency and high-frequency information separately. Channel mutual reduction helps attenuate the impact of high-frequency rain patterns. Therefore, we also design a Fourier Detailed Guidance module that utilizes the frequency and spatial domain features of the near-rainless detail image and to guide the restoration process.

In the training process, we have two stages with decoupled targets: one stage to construct the dual degradation representation of low-light-rainy images, and the other stage to eliminate degradation artifacts, generating a high-quality image under the supervision of the degradation representation. Our contributions can be summarized as follows:

- We propose a novel two-stage approach $L^2RIRNet$ to achieve low-light-rainy image restoration in synthetic and real scenarios. Experimental results have proven the effectiveness of our $L^2RIRNet$.
- We develop a DDR-Net containing rainy branch and light branch encoders to learn different degradation representations separately. This design provides better guidance than a single one for the restoration process.

- We present an FDG module that combines spatial and frequency domain information of our designed near-rainless detailed images to provide effective prior for the restoration part.

II. RELATED WORK

A. Single Image Rain Removal.

Image deraining poses a significant challenge due to its inherently ill-posed nature. Traditional methods reconstruct the background from rainy images by extracting the rain-free high-frequency component (HFP). Various filtering strategies, such as guided filters [26], multiple guided filtering [27], and non-localized means filtering [28], have been employed for removing the raindrop. For deep-learning-based methods, the CNN-based framework [29] can extract discriminative rain features from the HFP of a single rainy image. A physics-based model [30] can resemble real rain scenes as an initial constraint for deraining. A rain-density classifier guides subsequent refinement. In addition, PFDN [31] is a novel image derivation framework employing rain-related and rain-independent features to eliminate rain streaks and recover contextual details from multiple perspectives. However, the paired datasets are difficult to obtain in the real world, so a GAN-based unsupervised deraining framework [7] is proposed to solve the practical problem. After that, a large rainy database [8] is created to enhance the deraining performance for the real-world scene. Moreover, there are some other deraining methods focusing on model designing. For instance, a multi-scale architecture and attention strategy [6] are presented by Cai et al. Other approaches, such as SGINet [32], construct a new research direction that explores the use of high-level semantic information to enhance rain removal. Additionally, researchers address more complex task settings [33], such as synchronous rain streaks and raindrop removal. In recent advancements, MIRNet [20] uses a novel feature extraction network effective in image restoration and enhancement. Wang et al. [9] present an advanced collaborative network that incorporates multi-scale compact constraints and a bidirectional scale-content mining module to eliminate rain streaks. On the other hand, RCDNet [3] introduces an unfolding method, employing multi-stage training with M-net and B-net, achieving superior performance in image deraining. Recently, Restormer [19] utilizes a transformer-based framework for deraining, capitalizing on transformer benefits while minimizing computational complexity. NAFNet [18] employs a sequence of uncomplicated yet impactful improvements, optimizing the network baseline and fully unlocking its performance advantages. In DRSFormer [1], a multi-experts-based framework provides more accurate detail and texture recovery. Meanwhile, some methods consider the degradation prompt to image restoration and propose a PromptRestorer approach [24]. Most recently, RLP [4] points out a nighttime deraining method emphasizing the significance of incorporating location information regarding rain streaks, especially in night scenes. However, all these methods are tailored for restoring rain images captured in daytime conditions and may struggle to remove rain patterns effectively in low-light conditions.

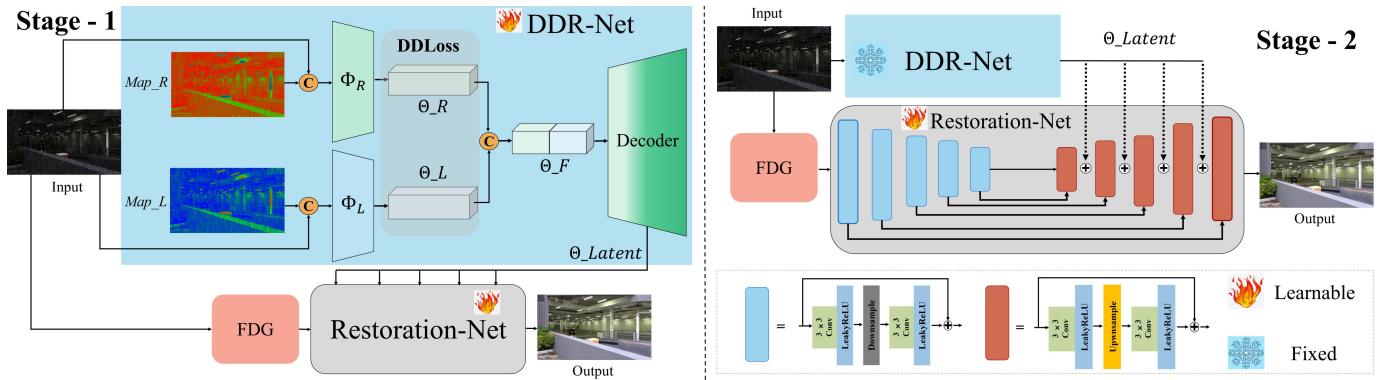


Fig. 2. Overview of our L²RIRNet, consisting of two components: DDR-Net and Restoration-Net. DDR-Net extracts the image's degradation representation of bright and dark regions by rainy and light maps, respectively. The dual degradation loss (DDLoss) is presented to constrain the training process. The decoder generates the multi-channel latent features to guide the restoration process effectively. The Fourier Detailed Guidance (FDG) module uses prior near-rainless information for the Restoration-Net. The training process consists of two phases: 1. Train both DDR-Net and Restoration-Net to get effective representations. 2. Fix the DDR-Net and only train the Restoration-Net to enhance restoration performance further.

B. Low-light Image Enhancement.

Traditional methods employ handcrafted optimization techniques and norm minimization methods to enhance low-light images. These approaches rely on various assumptions, such as dark/bright channel priors [34]–[36], and mathematical models, including Retinex theory [37] and Multi-scale Retinex theory [38], [39], to derive sub-optimal solutions. Learning-based methods fall into four categories: supervised [35], self-supervised [40], unsupervised [12], and zero-shot [10]. KinD [41] considers the impact of noise in low-light images and utilizes a noise removal module to enhance the performance of low-light image enhancement. Zero-shot approaches suggest various approximation strategies to reduce label dependency. Zero-DCE [10], and Zero-DCE++ [11] estimate multiple tone curves from input images to obtain enhanced results. A physical model [15] uses weight coefficients based on the light-scattering rate for low-illumination image enhancement. A strong generalization network, MLLEN-IC [14] is proposed to handle various low-light conditions. After that, a unified Retinex-based zero-reference deep framework [13] is proposed with an effective generative strategy. Moreover, a novel self-supervised framework SCI [40] is introduced with a cascaded illumination learning approach to minimize redundancy and enhance overall efficiency. UHDformer [17] transforms the high-resolution feature to the low-resolution space to achieve a general UHD image restoration. Wang et al. propose a UHD-LOL dataset and an innovative transformer-based approach, LLFormer. PyDiff [42] is the first to consider the diffusion model in low-light enhancement, achieving superior performance and speed compared to existing methods. Meanwhile, LLDiffusion [43] inserts the degradation representation into the diffusion model. Recently, Fu et al. propose a paired-based low-light enhancement technology PairLIE [2], which is composed of the designed priors and a simplified network. In UHDNet [44], it draws inspiration from distinctive features identified in the Fourier domain. However, for low-light-rainy images, these methods only enhance brightness and make rain patterns more visible in dark areas. They can not effectively restore them to a high-quality version.

III. PROPOSED METHOD

A. Overview

Our L²RIRNet consists of two key components: a Dual Degradation Representation Network (DDR-Net) and a Restoration Network. As depicted in Fig. 2, we utilize rainy and light attention maps to learn distinct degradation representations separately. The representation learning process is optimized through dual degradation contrastive loss (DDLoss). Our objective is to capture information that significantly influences the image quality of low-light-rainy images. The DDR-Net encodes the degradation operator into a low-dimensional feature to guide the restoration process, employing a U-Net-like network structure. Overall, the network is trained in two stages. We train the DDR-Net and Restoration-Net in the first stage for more accurate representations. In the second stage, we lock the DDR-Net and only train the Restoration-Net to further improve the restoration performance.

In the case of the DDR-Net, we initiate the process by acquiring a fusion degradation representation through Att_R and Att_L . Subsequently, we derive the latent degradation representation Θ_{Latent} using the Decoder:

$$\begin{cases} Att_R = \text{Concat}[Input, Map_R], \\ Att_L = \text{Concat}[Input, Map_L], \\ \Theta_R = \Phi_R(Att_R), \\ \Theta_L = \Phi_L(Att_L), \\ \Theta_F = \text{Concat}[\Theta_R, \Theta_L], \\ \Theta_{Latent} = \text{Decoder}(\Theta_F), \end{cases} \quad (1)$$

where $\Phi_R(\bullet)$ and $\Phi_L(\bullet)$ denote the encoder in rainy and light branch, respectively. Θ is the degradation representation. Subsequently, the Restoration-Net utilizes the Θ_{Latent} acquired from the DDR-Net to guide the restoration process, resulting in the final output:

$$Output = \text{Restoration}(Input, \Theta_{Latent}). \quad (2)$$

Finally, the restored image is obtained. The key components are elaborated below.

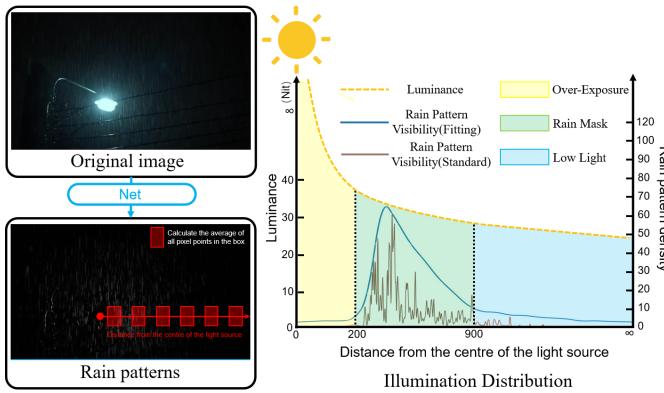


Fig. 3. The illustration of the relationship between rain pattern density (m) and the Euclidean distance (r) from the pixel to the light source. We utilize the pre-trained RCDNet [3] to obtain the rain pattern mask as shown on the left. Furthermore, we select multiple 20×20 patch blocks along a line (the red central line and red blocks shown in the figure) and calculate the rain pattern density for each block. Then, the distance (x-axis) is defined as the Euclidean distance from the center of the light source to the center of each block, and the rain pattern density values of the patch blocks form an illumination distribution curve at the right part. The x-axis represents the distance from the patch to the light source. The yellow curve represents the approximated light intensity distribution based on the Formula 3. The brown curve represents the actual rain density distribution, and the blue curve is the smoothing result of the brown curve.

B. Dual Degradation Representation Network (DDR-Net)

1) *Adaptive Map for Low-light-rainy Images:* As depicted in Fig. 3, we analyze the relationship between light intensity and rain pattern density using a real-world low-light-rainy image containing a light source. The first law of illumination states that “When illuminated by a point source of light, the illumination of a surface is inversely proportional to the square of its distance from the light source (r^2)”. With the development of optics [45], [46], the relationship between light intensity E , the light intensity at the center of the light source E_0 and the distance r can be summarized as:

$$E = \frac{E_0}{r^2}, \quad (3)$$

where r represents the euclidean distance from surrounding pixels (x_1, y_1) to the center (x_0, y_0) of a light source, expressed as

$$r = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2}. \quad (4)$$

To calculate the rain pattern density m , we utilize the pre-trained RCDNet [3] to obtain the rain pattern mask. Then, we generate numerous 20×20 patch blocks intersecting the straight line. Finally, the relationship between m and r is shown in the right part of Fig. 3. Especially in the pixel region between 0 and 200, capturing raindrop information is challenging due to overexposure near the light source; between 200 and 900, raindrop information becomes more visible, and the degraded information is predominantly composed of raindrop details; and in the region between 900 and infinity, capturing raindrop information is also difficult due to image degradation caused by low-light conditions.

Based on the discussion above, we conclude that different locations contain different types of degradation information.

Areas closer to the light source are primarily affected by raindrop degradation, while areas further from the light source are mainly affected by low-light information. Therefore, we hope to explore a method that learns these two types of degradation separately. Inspired by EnlightenGAN [12], the attention map can accentuate the dark areas of the image for low-light image enhancement. So, in our framework, we employ the map as our Map_L , which allows the network to capture information related to low-light degradation effectively. Simultaneously, the complement attention map, Map_R , can be utilized to focus on the bright regions, which correspond to the rain pattern information in those areas. By training the network to independently process the bright and dark areas using the Map_R and Map_L , we can derive the degradation latent and effectively guide the Restoration-Net.

2) *Dual Degradation Contrastive Loss:* Utilizing contrastive learning loss (CL) can not only significantly enhance the performance [47] but also extract and learn degradation representation to guide the downstream process [48]–[50]. In such cases, the mixed degradation representation can be expressed in terms of a multiplicative or additive formula.

Limitation of CL. One of the most classic methods is the contrastive learning scheme of DASR [48], which selects two patches from the same image as the anchor and the positive label and lets patches from the paired clean image be the negative labels. However, our previous analysis in section I and III-B1 reveals significant differences in rainy and dark levels across various positions in a single low-light-rainy image, as depicted in Fig. 4(a). This makes the patch-based approach unsuitable for low-light-rainy image restoration tasks. Thus, in the next section, we introduce different degradation representations of the whole image separately.

Solution. To solve the limitations above, we present a dual degradation loss (DDLoss). As shown in Fig. 4(b), firstly, we abstain from learning the mixed degradation representation and instead concentrate on autonomously acquiring the representations for rainy and light degradation. Following this, we proceed to integrate these representations. Secondly, rather than selecting the anchor and the positive label from the same image, we propose an alternative approach that incorporates a paired clean image Target as the negative label and the augmented version of Input, named Input_{Aug} (rotations of 90, 180, and 270 degrees and horizontal flipping), as the positive label. We put the three images (Input_{Aug}, Input, Target) into the encoder of the DDR-Net(DDR_E) to obtain the dual degradations DD_+ , DD_I and DD_- . Then, we put them into a two-layer MLP projection head to generate three tensors. The process is as follows:

$$DD_+, DD_{Image}, DD_- = DDR_E(\text{Input}_{\text{Aug}}, \text{Input}, \text{Target}), \quad (5)$$

$$T_+, T_{Image}, T_- = \text{MLP}(DD_+, DD_{Image}, DD_-). \quad (6)$$

The learned degradation allows anchor patches to better approach “positive” pairs in a specific metric space and move away from “negative” pairs. The dual degradation loss (DDLoss) function is defined as:

$$L_D = \frac{\|T_{Image} - T_+\|_1}{\|T_{Image} - T_-\|_1}. \quad (7)$$

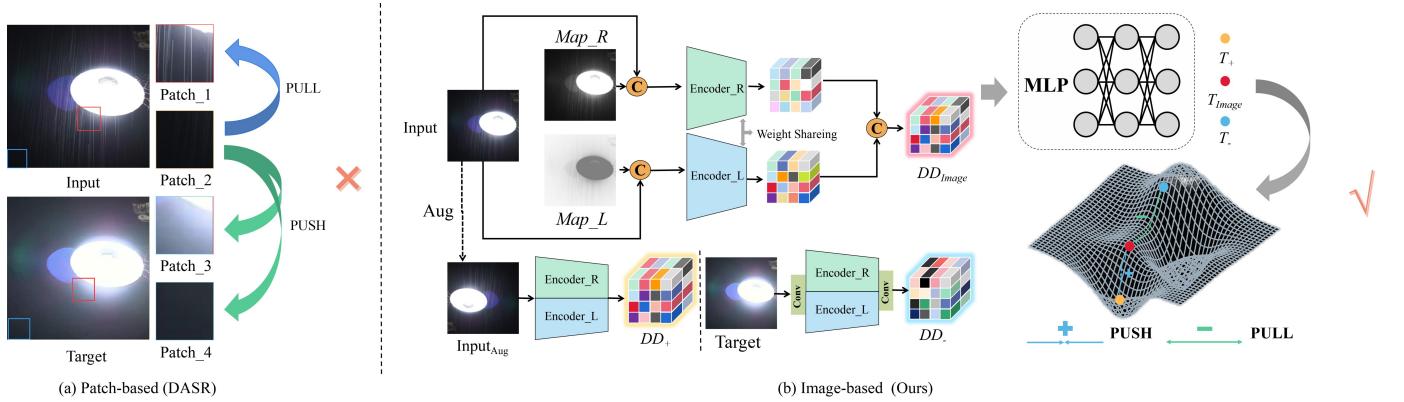


Fig. 4. (a) DASR’s patch-based positive and negative label approach. The selected patches may come from low-light and rain pattern areas, which is challenging to apply as a positive label. (b) Our image-based method learns degradation representation by the dual degradation loss. We present the rainy and light encoder part of the DDR-Net, which takes in three images: a low-light-rainy image (Input), an augmented version of that image ($\text{Input}_{\text{Aug}}$), and a paired clean image (Target). These images are encoded using a parameter-sharing network, with Map_R and Map_L focusing on the degradation representation in bright and dark regions. The DD_+ is the dual degradation representation from $\text{Input}_{\text{Aug}}$, and the DD_- is the dual degradation representation from Target. The DD_{Image} is the dual degradation representation from Input. A multilayer perceptron (MLP) then processes the resulting feature information to a tensor for comparison learning constraint. Our approach aims to make T_{Image} and T_+ similar while keeping T_{Image} and T_- distinct.

3) *DDR-Net Framework*: Based on the discussion above, we propose an effective DDR-Net with a U-Net-like structure for extracting degradation information. It includes components for attention map extraction, an encoder, and a decoder. By utilizing rainy and light maps, we divide the image into bright and dark regions, directing the network’s attention more intensively to each area during training. Specifically, we extract distinct degradation information from the two regions and fuse them.

C. Restoration-Net

1) *Fourier Detail Guidance Module (FDG)*: Low illumination and noise affect not only image content but also the distribution of rain patterns, thereby impacting image visibility. Removing the rain pattern information while recovering the lighting is challenging for the restoration part. We observe that the contents of the rain streaks in three RGB channel maps of the input image are similar. Thus, near-rainless detailed images can be obtained by calculating the difference between these channels, as shown in Fig. 5. Thus, we leverage this characteristic to reduce the interference of rain patterns and deeply explore the image content information. Specifically, we implement a Fourier Detail Guidance (FDG) module. This module utilizes a detailed image of the background without rain patterns to provide well-defined prior information for the restoration process. As shown in Fig. 6, the FDG contains two components: Detailed Image Generation (DIG) and Frequency Spatial Feature Extraction (FSFE).

The input low-light-rainy image is initially split into the RGB (red, green, and blue) channels in the DIG step. Subsequently, the absolute differences between each pair of channels are calculated, resulting in three single-channel maps: $|R - G|$, $|R - B|$ and $|G - B|$. Finally, these maps are concatenated to create a detailed image that captures the unique features of each single-layer map. Fig. 5 illustrates that the detailed image for both synthetic and real-world image obtained by DIG preserves the image’s detailed structure effectively and

removes the rain pattern information. This approach serves as a preliminary rain removal strategy, and the output images can be effectively utilized for the restoration tasks.

It is known that conventional convolution focuses more on local information, and the ability for global representation is limited. Inspired by the Fourier convolution theory [51], some works have observed that learning in the frequency domain achieves broader global capabilities. Therefore, in order to leverage global frequency information to complement local spatial features and enhance the model’s learning ability, we design an FSFE module that effectively facilitates learning global and local information. Specifically, as shown on the right part of Fig. 6, we first send the fusion image F_I (the concatenation of Input and the Detailed Image) into two branches: the frequency domain branch B_F and the spatial domain branch B_S . Then, we further fuse the results obtained from these branches:

$$F_F = B_F(F_I), \quad (8)$$

$$F_S = B_S(F_I), \quad (9)$$

where F_F and F_S denote the frequency and spatial features respectively. Then we fuse them to obtain the feature weights through a Sigmoid function and modulate the features of the two branches separately:

$$S = \text{Sigmoid}(F_F + F_S), \quad (10)$$

$$\widetilde{F}_F = \text{Conv}(F_F \otimes S) + F_F, \quad (11)$$

$$\widetilde{F}_S = \text{Conv}(F_S \otimes S) + F_S, \quad (12)$$

where \widetilde{F}_F and \widetilde{F}_S denote the activated feature of F_F and F_S . Finally, the output \widetilde{F}_I of the FSFE can be obtained by concatenating the features:

$$\widetilde{F}_I = \text{Concat}(\widetilde{F}_F, \widetilde{F}_S). \quad (13)$$

Our experimental results demonstrate that the proposed method effectively removes rain patterns from the image while preserving background content information. As a result, the



Fig. 5. Visual presentation of detailed images. Compared to the input, the detailed image removes most raindrops while maintaining the content information.

FDG module can concentrate on learning the background information, which enhances its ability to remove rain pattern information.

2) *Restoration-Net Framework*: Our Restoration-Net architecture is specifically designed to restore low-light-rainy images to their corresponding clean versions. This is achieved through the FDG module for pre-processing, an encoder, and a decoder.

As detailed in the previous section, the FDG module processes both detail images and low-light-rainy images, concatenating them to form the encoder's input. This pre-processing step can reduce noise and rain patterns, preserve image details,

and enhance contrast. The encoder is responsible for extracting feature information from the input image by increasing its dimension. The encoded feature information is transmitted to the decoder via a mid-level jump connection.

To improve the performance of the restoration process, the Restoration-Net decoder leverages the degradation feature from multiple layers by DDR-Net. We adjust the features from DDR-Net to the same scale as the Restoration-Net through convolution and then add them together. This strategic utilization enables the Restoration-Net architecture to concentrate more on the degraded components of the image, thereby improving the efficiency and performance of the restoration process.

We use the L1 loss and perceptual loss to train our Restoration-Net. Specifically, the loss function is defined as:

$$L_R = \|R(x) - x_{clean}\|_1 + \lambda_p \|\Phi(R(x)) - \Phi(x_{clean})\|_1, \quad (14)$$

where $R(\bullet)$ denotes the Restoration-Net, λ_p is the weight parameters, $\Phi(\bullet)$ denotes the pre-trained VGG19 network. We adopt multi-scale feature maps from layer $\{conv1, \dots, conv4\}$ following the widely-used setting.

The overall loss function is:

$$L = \lambda_D L_D + \lambda_R L_R, \quad (15)$$

where λ_D and λ_R denote weight parameters.

IV. EXPERIMENTS

A. Proposed Dataset

For training the proposed L²RIRNet, we contribute an LLR Dataset containing synthetic and real-world images. It contains 8368 pairs of images for training, 800 synthetic, and 42 real-world images for testing.

Synthetic low-light-rainy images:

(1) Inspired by LEDNet [54], which proposes a pipeline for generating low-light blurred images, we first synthesize 8000 low-light-rainy images by decreasing the light intensity of synthetic rainy images from the DID dataset [52]. Specifically, we use gamma correction to reduce the brightness of the images, where the gamma value is randomly selected between 2 and 3.5 to achieve the different stages of brightness reduction. Then, we add random Gaussian noise to simulate real-world low-light conditions.

(2) Considering the lighting distribution characteristics of real-world environments, we randomly select 4000 images from the first part and enhance the pixels within random patches while reducing the brightness in other areas. The specific steps are illustrated in Fig. 7.

As shown in Fig. 8, we present the results of a real-world low-light-rainy image. Our L²RIRNet is trained on the dataset (1) and (1) + (2), respectively. The addition of (2) improves the resolution of the overexposed areas near the light source, demonstrating the usefulness of this operation.

We also research 368 real-world low-light-rainy images from [53] to construct fake-synthetic low-light-rainy images for training. In Fig. 9, (a) and (b) are a pair of rainy-clean

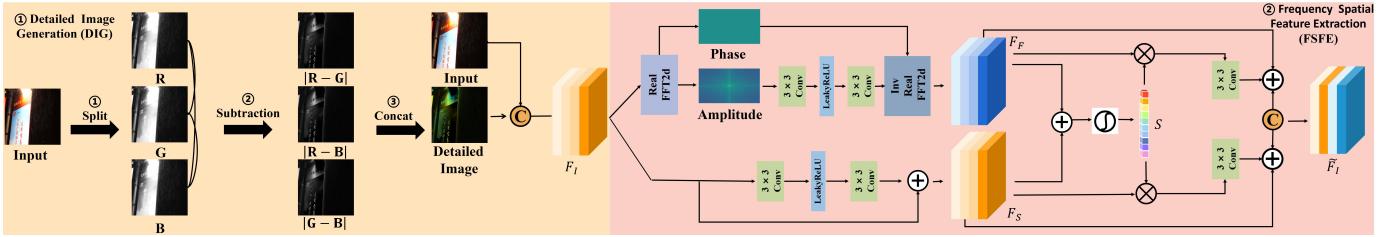


Fig. 6. The architecture of the FDG module is as follows, comprising two main steps: 1. Detailed Image Generation (DIG): split an input low-light-rainy image into RGB channels and subtract the absolute differences between these channels, creating $|R - G|$, $|R - B|$ and $|G - B|$. Then, concatenate them to form a detailed image. 2. Frequency Spatial Feature Extraction (FSFE). FSFE contains two branches: the frequency branch and the spatial branch. The frequency branch focuses on capturing frequency-related information, while the spatial branch is designed to extract spatial features. The fusion of these features is employed to guide the restoration process effectively.

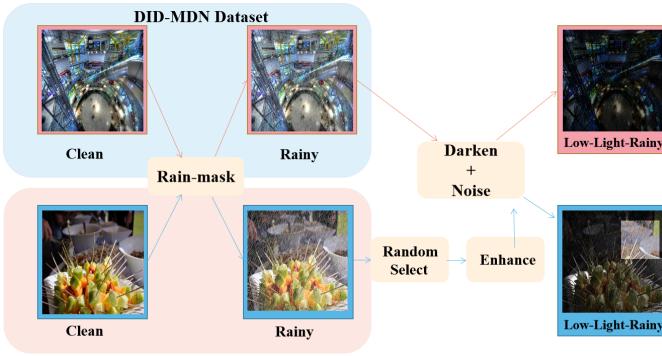


Fig. 7. An overview of synthetic images in the LLR dataset synthesis pipeline. The clean and rainy images belong to the DID-MDN dataset [52]. There are two methods to synthesize low-light-rainy images. (1) Darken and add noise to 4000 images, as shown above. (2) Randomly select patches, enhance their light intensity, and then darken other regions, simulating real-world low-light-rainy images, as shown below.



Fig. 8. A visual comparison on a real-world image of our L²RIRNet training only on (1) and (1) + (2).

images from [53], and we use SCI [40] to enhance the target image to obtain (c) as our target in low-light-enhancement.

Real-world low-light-rainy images:

- (1) We search the real deraining datasets [53], [55] for nighttime scenes.
- (2) We conduct a Google search to find real-world low-light-rainy images.
- (3) We capture some low-light-rainy images in outdoor scenes.

B. Experimental settings

- 1) **Implementation details:** Our DDR-Net and Restoration-Net are both U-Net-like architectures. We give two versions for Restoration-Net: CNN-based (HINet [56]) and Trans-based



Fig. 9. (a) and (b) are a real rainy image and its clean target from [53]. (c) is obtained by enhancing the light of (b) as the target of the low-light-rainy restoration task.

(Restormer [19]). The scale of the DDR-Net will be matched accordingly. To optimize our proposed model, we adopt the Adam optimizer algorithm with $\beta_1=0.9$, $\beta_2=0.999$, and the initial learning rate is set to 2×10^{-4} and gradually decreases to 10^{-6} with the cosine annealing schedule [57]. The used framework is PyTorch and the used GPU is GeForce RTX 3090. For the CNN-based version, we set the batch size to 24, patch size to 256, 200 epochs at the first stage, and 300 epochs at the second stage. For the Trans-based version, batch size is set to 8, and patch size is 128, at the first stage for 90 epochs and the second stage for 200 epochs. We set the weights λ_p , λ_D and λ_R in (14) and (15) to 0.1, 1 and 1, respectively.

2) **Evaluation Metrics:** We employ the PSNR and SSIM metrics to evaluate the restored results on synthetic images in the LLR Dataset, and for comprehensive comparison, the model parameters and running time are given. For real-world images from the LLR Dataset, since there is no ground truth for calculating PSNR and SSIM, the classical non-reference NIQE [58] and CEIQ [59] are used.

C. Comparisons on the LLR Dataset

We evaluate the effectiveness of the proposed L²RIRNet on the synthetic and real-world images from the LLR dataset quantitatively and qualitatively. As the joint task addressed in this paper is new, no open existing methods can directly compare with ours. To provide a comprehensive evaluation, we carefully select some representative low-light enhancement and deraining methods, resulting in two distinct baseline methods for comparison: cascaded and retrained.

Cascaded Methods. For low-light enhancement methods, we choose the classic Zero-DCE [10], SCI [40], and some

TABLE I

QUANTITATIVE EVALUATION ON THE LLR DATASET. THE BASELINE CONTAINS THE CASCADED METHODS (ENHANCEMENT-DERAINING AND DERAINING-ENHANCEMENT) AND RETRAINED METHODS ON OUR LLR DATASET. '*' INDICATES THE NETWORK IS RETRAINED ON OUR LLR DATASET. THE PARAMETERS AND RUNNING TIME ARE EXPRESSED SEPARATELY IN MILLIONS (M) AND MILLISECONDS (Ms). ALL RUNNING TIME IS EVALUATED ON A 512×512 IMAGE USING A GEFORCE RTX 3090 GPU.

	Methods	PSNR(dB)	SSIM	Param (M)	Running time (Ms)
Cacaded Methods	Zero-DCE \Rightarrow RCDNet	17.58	0.6634	3.5	235.33
	SCI \Rightarrow MIRNet	20.71	0.7186	32.3	198.50
	UHDNet \Rightarrow Restormer	19.91	0.6928	59.8	206.64
	PairLIE \Rightarrow DRSFormer	18.29	0.6495	34.0	555.38
	MIRNet \Rightarrow SCI	19.32	0.6931	32.3	198.50
	RCDNet \Rightarrow Zero-DCE	12.21	0.6175	3.5	235.33
	Restormer \Rightarrow UHDNet	19.57	0.6878	59.8	206.64
Retrained Methods	DRSFormer \Rightarrow PairLIE	19.19	0.6868	34.0	555.38
	Zero-DCE*	21.75	0.7189	0.5	57.81
	RCDNet*	25.04	0.8031	3.0	177.52
	MIRNet*	29.31	0.8872	31.8	193.53
	Restormer*	31.19	0.9174	26.1	190.53
	NAFNet*	29.17	0.8987	17.1	32.67
	PyDiff*	29.06	0.8753	54.5	373.56
	UHDNet*	24.91	0.7914	33.6	16.11
	PairLIE*	19.33	0.6948	0.3	2.42
	LLFormer*	30.96	0.9171	24.52	247.72
	RLP*	29.45	0.8878	5.3	101.72
	DRSformer*	31.58	0.9182	33.7	552.96
	L^2 RIRNet-CNN (Ours)	30.53	0.9030	73.9	58.38
	L^2 RIRNet-Trans (Ours)	32.53	0.9231	37.5	229.01

TABLE II
QUANTITATIVE EVALUATION ON THE REAL-WORLD LLR DATASET.

	Input	RCDNet*	MIRNet*	Restormer*	NAFNet*
NIQE \downarrow	16.93	15.97	14.31	15.55	14.65
CEIQ \uparrow	3.31	3.36	3.43	3.47	3.46
LLFormer	DRSformer*	RLP*	UHDNet*	PairLIE*	L^2 RIRNet (Ours)
14.98	14.82	14.61	16.36	17.34	14.17
3.47	3.45	3.44	3.49	3.36	3.50

state-of-the-art approaches UHDNet [44] and PairLIE [2]. For the deraining part, we choose some recent deraining methods: MIRNet [20], RCDNet [3], Restormer [19], and DRSFormer [1].

Retrained Methods. In addition to the methods mentioned above, we also select other latest approaches for retraining: Pydiff [42], NAFNet [18], MIRNet [20], LLFormer [60] and RLP [4].

1) *Quantitative and Qualitative Evaluations.*: Table I shows the average PSNR and SSIM values of the reported

restoration methods on synthetic images from the LLR Dataset. The results demonstrate that our L^2 RIRNet performs best among all the methods. Importantly, our approach shows a significant advantage compared to the eight cascade approaches. After retraining these state-of-the-art networks on the LLR Dataset, our L^2 RIRNet still performs favorably against them. For instance, compared to DRSFormer [1], there is a similar number of model parameters but increases 0.95 dB/0.0049 in PSNR and SSIM, respectively. As shown in Table II, the proposed L^2 RIRNet achieves the lowest NIQE

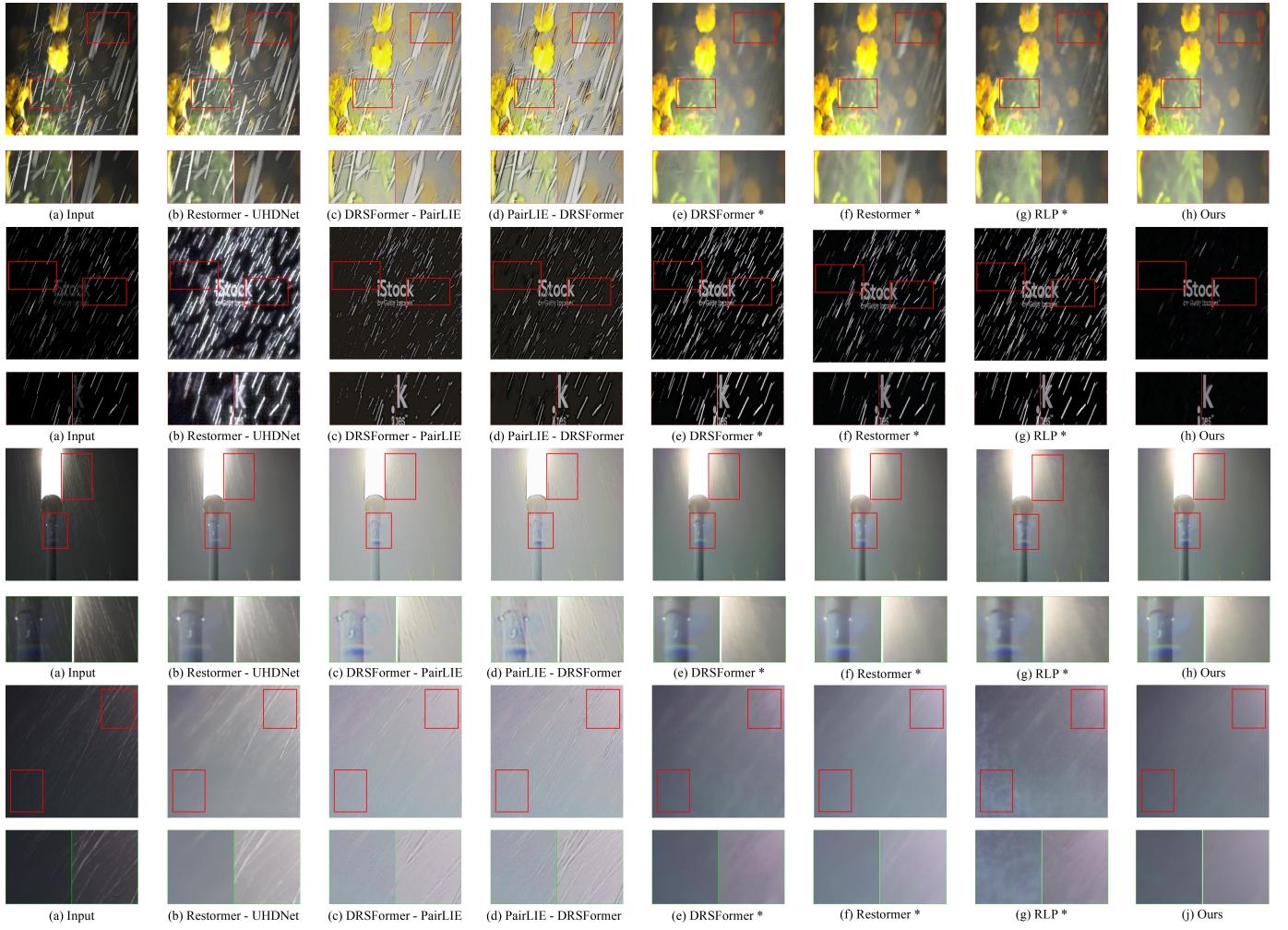


Fig. 10. Visual comparison on a real-world image from the LLR dataset. '*' indicates the network is retrained on our LLR Dataset.

score and the highest CEIQ, indicating that our results are perceptually best.

Figs. 10 and 11 provide a visual comparison of synthetic and real-world images from the LLR Dataset. Whether it is a synthetic or real image, the results of the cascade method are consistently unsatisfactory, often leaving behind rain pattern residue or overexposure. In real-world images, some of the latest methods struggle to remove rain streaks hidden in the darkness. In contrast, our method not only removes the rain streaks (the first and fourth rows of Fig. 10) but also makes the image background clearer (the second row of Fig. 10) and avoids blurring local scenes (the third row of Fig. 10). For synthetic data, ours still achieves the best visual results.

2) *Color correction Evaluations.*: Our method includes capabilities for enhancing low-light conditions. We utilize pixel-level metrics such as PSNR and SSIM, and also compare our model with state-of-the-art methods using color-related metrics to evaluate its color correction performance. Following the previous work [61], we use the Root Mean Square Error (RMSE) in the Lab color space. The results of this comparison are presented in Table III, which demonstrates that our method still achieves the best performance in color correction.

TABLE III
QUANTITATIVE COMPARISON BETWEEN THE PROPOSED METHOD AND SOTA METHODS USING RMSE.

Methods	Input	RLP	Restormer
RMSE↓	7.5127	4.2371	3.7927
Methods	NAFNet	DRSformer	L ² RIRNet (Ours)
RMSE↓	4.3849	3.6592	3.4568

D. Ablation Study

Effectiveness of key architecture. Table. IV shows seven different variants of the model. V1 represents only Retoration-Net. V2 combines the Restoration-Net with the single degradation representation network (SDR-Net). V3 combines the Restoration-Net with the dual degradation representation network (DDR-Net). The V4 and V5 are Restoration + FDG module and Restoration + SDR-Net + FDG module. V6 represents our stage1 of L²RIRNet, and V7 is our complete process with stage1 and stage2.

By carefully comparing these variants, we gain important insights into the performance of our L²RIRNet. Comparing

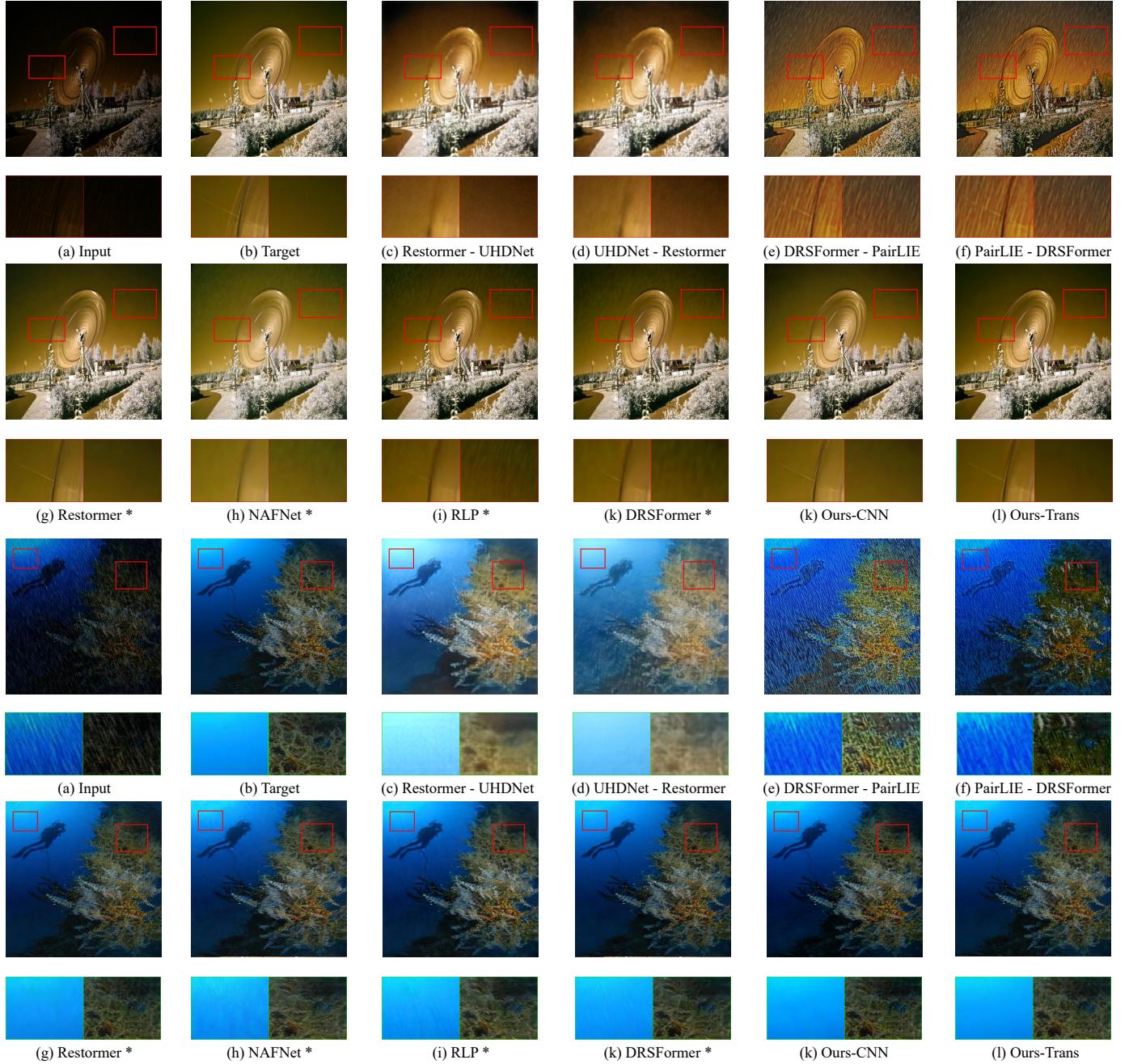


Fig. 11. Visual comparison on synthetic images from the LLR dataset. '*' indicates the network is retrained on our LLR Dataset.

V1 with V2, we find that adding a degradation representation guidance can improve restoration performance. Furthermore, comparing V2 with V3, the PSNR is increased by 0.58 dB. This result proves the vital role of our DDR-Net in extracting degradation information from both bright and dark regions using dual channels, resulting in significant improvement compared to single channels. We also compare V4 and V5 with V1 and V2, and it can be observed that after adding the FDG module, there is a certain improvement across the two basic frameworks. Most importantly, when we compare V3 with V6, 0.84dB is increased in PSNR by adding the FDG module. This result highlights the value of our proposed model, as it demonstrates the importance of adding

prior details and Fourier frequency domain information to the restoration process. V7 increases by 0.57 dB in PSNR over V6. These results demonstrate the effectiveness of each key part of our L^2 RIRNet architecture and the importance of combining them to achieve satisfactory performance.

Effectiveness of the Fourier Detailed Guidance Module. In Table V, we present our L^2 RIRNet without detailed image generation (DIG), frequency, and spatial branch. All four versions are two-stage training methods. The results show that the DIG module enhances the performance from 29.56dB/0.8883 to 30.53dB/0.9030, demonstrating the validity of DIG. The FSFE, without frequency and spatial branches, yield results of 30.12dB/0.8953 and 28.69dB/0.8842, respectively. As we

TABLE IV

EFFECTIVENESS OF THE KEY ARCHITECTURE: **V1**: ONLY THE RESTORATION-NET; **V2**: RESTORATION-NET + SDR-NET (SINGLE DEGRADATION REPRESENTATION NETWORK); **V3**: RESTORATION-NET + DDR-NET (DUAL DEGRADATION REPRESENTATION NETWORK); **V4**: V1 + FDG MODULE; **V5**: V2 + FDG MODULE; **V6**: THE FIRST STAGE OF L²RIRNET; **V7**: FOURIER DETAILED GUIDANCE (FDG) MODULE. **V7 (OURS)**: V4 + SECOND STAGE (L²RIRNET).

Methods	V1	V2	V3	V4	V5	V6	V7
Restoration	✓	✓	✓	✓	✓	✓	✓
SDR-Net	✓			✓			
DDR-Net		✓			✓	✓	
FDG Module			✓	✓	✓	✓	
Stage-2					✓		
PSNR	28.06	28.54	29.12	28.64	29.38	29.96	30.53

TABLE V

EFFECTIVENESS OF THE FOURIER DETAILED GUIDANCE MODULE.

Methods	PSNR (dB)	SSIM
without DIG	29.56	0.8883
without frequency	30.12	0.8953
without spatial	28.69	0.8842
Ours	30.53	0.9030

discussed above, different areas may exhibit different degradation characteristics; areas closer to the light source are mainly affected by rainy degradation, while areas further from the light source are primarily affected by low-light conditions. Therefore, focusing only on the spatial domain features of local regions might overlook these aspects. Paying attention to the global frequency domain can supplement this information. Therefore, adding the frequency domain branch can improve the resolution by almost 0.4 dB.

The Benefit of the Dual Degradation Contrastive Loss

TABLE VI

EFFECTIVENESS OF THE DUAL DEGRADATION CONTRASTIVE LOSS.

Methods	PSNR (dB)	SSIM
Only Restoration-Net	28.06	0.8762
Patch-based (DASR)	27.13	0.8542
Image-based (Ours)	30.53	0.9030

TABLE VII

EFFECTIVENESS OF FORMS OF DEGRADATION REPRESENTATION LEARNING NETWORK.

Methods	PSNR (dB)	SSIM	Params (M)	Runtime (Ms)
DDR-unshared	30.12	0.897	152.62	112.43
DDR-shared (ours)	30.53	0.903	73.88	58.38

(DDLoss). As shown in Table VI, our image-based strategy achieves 30.53dB/0.9030, showing a 2.47dB/0.0268 improvement compared to only using Restoration-Net. In contrast, the patch-based contrastive learning loss from DASR [48] achieves 27.13dB/0.8542, which is lower than 28.06dB/0.8762 obtained by only using Restoration-Net. As depicted in Fig. 4(a), the patch-based approach may result in the selection of blocks with different degradation representations, leading to incorrectly learned representations that can negatively impact the performance of the restoration network.

Effectiveness of Forms of Degradation Representation Network.

To validate the enhancement of our DDR-Net, we compare our dual-branch shared version with the dual-branch unshared version. Table VII demonstrates that a shared DDR-Net can enhance performance, reduce the number of parameters, and decrease runtime.

E. Validate the performance of our model under different brightness levels

In this section, we validate the effectiveness of our methods under different brightness levels. We design five different test set settings. Three of them are aligned with the LLR dataset construction, covering the gamma ranges 2 to 2.5, 2.5 to 3, and 3 to 3.5. Additionally, we compare ranges outside the LLR dataset, which are 1.5 to 2 and 3.5 to 4, to assess the generalization of the proposed model. As illustrated in Fig. 12, we observe that within the gamma range of 2 to 3.5, our method achieves the best performance compared to other methods. Even in the ranges of 1.5 to 2 and 3.5 to 4, our method exhibits considerable advantages, demonstrating its better generalization capability.

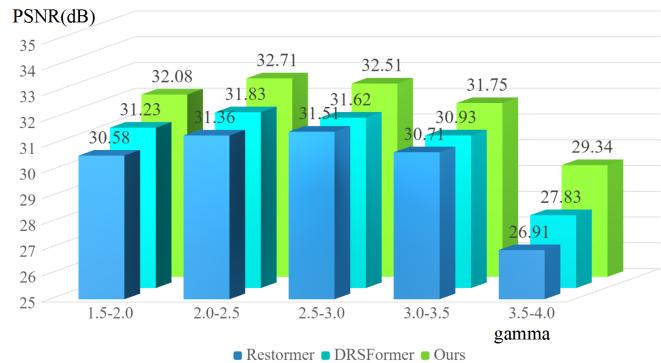


Fig. 12. Comparison of PSNR under different brightness levels of Restormer, DRSFormer, and ours.

V. CONCLUSION

In this paper, we present a two-stage framework for low-light-rainy image restoration that performs favorably against existing solutions for both synthetic and real images. A key component of our approach is the DDR-Net, which can extract distinct degradation representations for the degraded bright and dark regions, thereby accurately capturing the causes of image degradation. Additionally, we introduce an FDG module based on detailed image and frequency spatial feature extraction,

which provides valuable guidance for the restoration process. Our approach offers significant improvements over current methods, highlighting the effectiveness of our framework in addressing the challenges of low-light-rainy image restoration.

VI. LIMITATIONS AND FUTURE WORK

L^2 RIRNet efficiently solves the low-light-rainy image restoration task in synthetic and real-world scenes. Some diffusion-based models may obtain excellent results. Additionally, degradation-robust semantic features may provide better guidance as they focus on the image content information, not incomprehensible degrading factors. Therefore, in the future, a semantic-aware encoder or a large-scale pre-trained model that can provide stable features to guide the restoration process is worth exploring.

REFERENCES

- [1] C. X, L. H, and L. M, "Learning a sparse transformer network for effective image deraining," in *IEEE International Conference on Computer Vision and Pattern Recognition*, 2023, pp. 5896–5905.
- [2] F. Z, Y. Y, and T. X, "Learning a simple low-light image enhancer from paired low-light instances," in *IEEE International Conference on Computer Vision and Pattern Recognition*, 2023, pp. 22252–22261.
- [3] H. Wang, Q. Xie, Q. Zhao, Y. Li, Y. Liang, Y. Zheng, and D. Meng, "Rcdnet: An interpretable rain convolutional dictionary network for single image deraining," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 8668–8682, 2024.
- [4] Z. F, Y. S, and L. Y, "Learning rain location prior for nighttime deraining," in *IEEE International Conference on Computer Vision*, 2023, pp. 13148–13157.
- [5] Q. Wu, L. Wang, K. N. Ngan, H. Li, F. Meng, and L. Xu, "Subjective and objective de-raining quality assessment towards authentic rain image," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 3883–3897, 2020.
- [6] C. L, F. Y, and H. W, "Multiscale attentive image de-raining networks via neural architecture search," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 618–633, 2022.
- [7] Z. H, S. V, and P. V. M, "Image de-raining using a conditional generative adversarial network," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 3943–3956, 2019.
- [8] W. W, L. Y, and L. Z, "Subband differentiated learning network for rain streak removal," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 4675–4688, 2023.
- [9] C. Wang, J. Pan, and X.-M. Wu, "Online-updated high-order collaborative networks for single image deraining," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022, pp. 2406–2413.
- [10] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, and R. Cong, "Zero-reference deep curve estimation for low-light image enhancement," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2020, pp. 1780–1789.
- [11] C. Li, C. Guo, and C. C. Loy, "Learning to enhance low-light image via zero-reference deep curve estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 4225–4238, 2021.
- [12] J. Y, G. X, and L. D, "Enlightengan: Deep light enhancement without paired supervision," *IEEE Transactions on Image Processing*, p. 2340–2349, 2021.
- [13] Z. Z, X. B, and W. L, "Retinexdip: A unified deep framework for low-light image enhancement," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1076–1088, 2021.
- [14] F. G. D, F. B, and G. M, "Multiscale low-light image enhancement network with illumination constraint," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 7403–7417, 2022.
- [15] Y. S. Y and Z. H, "Low-illumination image enhancement algorithm based on a physical lighting model," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 28–37, 2017.
- [16] L. C, W. F, and W. X, "Efnet: Restoration for low-light images via enhancement-fusion iterative network," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 8486–8499, 2022.
- [17] C. Wang, J. Pan, W. W. G. F. S. L. M. W. X.-M. Wu, and J. Liu, "Correlation matching transformation transformers for uhd image restoration," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024, pp. 5336–5344.
- [18] C. L, C. X, and Z. X, "Simple baselines for image restoration," in *European Conference on Computer Vision*, 2022, pp. 17–33.
- [19] Z. S. W, A. A, and K. S, "Restormer: Efficient transformer for high-resolution image restoration," in *IEEE International Conference on Computer Vision and Pattern Recognition*, 2022, pp. 5728–5739.
- [20] Z. S. W, A. A, K. S, H. M., K. F. S., Y. M. H., and L. Shao, "Learning enriched features for fast image restoration and enhancement," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1934–1948, 2022.
- [21] X. Lin, C. Ren, X. Liu, J. Huang, and Y. Lei, "Unsupervised image denoising in real-world scenarios via self-collaboration parallel generative adversarial branches," in *IEEE International Conference on Computer Vision*, 2023, pp. 12642–12652.
- [22] X. Lin, J. Yue, C. Ren, C.-L. Guo, and C. Li, "Unlocking low-light-rainy image restoration by pairwise degradation feature vector guidance," *arXiv:2305.03997*, 2023.
- [23] X. Lin, C. Ren, K. C. Chan, L. Qi, J. Pan, and M.-H. Yang, "Multi-task image restoration guided by robust dino features," *arXiv:2312.01677*, 2023.
- [24] C. Wang, J. Pan, W. Wang, J. Dong, M. Wang, Y. Ju, and J. Chen, "Promptrestorer: A prompting image restoration method with degradation perception," in *Advances in Neural Information Processing Systems*, 2023, pp. 8898–8912.
- [25] X. Fu, J. Huang, D. Zeng, H. Yue, X. Ding, and J. Paisley, "Removing rain from single images via a deep detail network," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, p. 3855–3863.
- [26] X. Jing, Z. Wei, L. Peng, and X. Tang, "Removing rain and snow in a single image using guided filter," in *IEEE International Conference on Computer Science and Automation Engineering*, 2012, pp. 304–307.
- [27] X. Zheng, Y. Liao, X. F. W. Guo, and X. Ding, "Single-image-based rain and snow removal using multi-guided filter," in *International Conference on Neural Information Processing*, 2013, pp. 258–265.
- [28] J. H. Kim, C. Lee, J. Y. Sim, and C. S. Kim, "Single-image deraining using an adaptive nonlocal means filter," in *IEEE International Conference on Image Processing*, 2014, pp. 914–917.
- [29] X. Fu, J. Huang, X. Ding, Y. Liao, and J. Paisley, "Clearing the skies: A deep network architecture for single-image rain removal," *IEEE Transactions on Image Processing*, p. 2944–2956, 2017.
- [30] C. M, G. Z, and R. B, "A two-stage density-aware single image deraining method," *IEEE Transactions on Image Processing*, pp. 6843–6854, 2021.
- [31] W. Q, S. G, and D. J, "Pfdn: Pyramid feature decoupling network for single image deraining," *IEEE Transactions on Image Processing*, pp. 7091–7101, 2022.
- [32] Y. Wei, Z. Zhang, H. Zheng, R. Hong, Y. Yang, and M. Wang, "Sginet: Toward sufficient interaction between single image deraining and semantic segmentation," in *ACM International Conference on Multimedia*, 2022, p. 6202–6210.
- [33] Y. Wei, Z. Zhang, M. Xu, R. Hong, J. Fan, and S. Yan, "Robust attention deraining network for synchronous rain streaks and raindrops removal," in *ACM International Conference on Multimedia*, 2022, p. 6464–6472.
- [34] X. Guo, L. Yu, and H. Ling, "Lime: Low-light image enhancement via illumination map estimation," *IEEE Transactions on Image Processing*, pp. 982–993, 2016.
- [35] X. Fu, D. Zeng, Y. Huang, X. Ding, and X.-P. Zhang, "A variational framework for single low light image enhancement using bright channel prior," in *IEEE global conference on signal and information processing*, 2013, p. 1085–1088.
- [36] W. Zhihong and X. Xiaohong, "Study on histogram equalization," *Intelligence Information Processing and Trusted Computing, International Symposium on IEEE Computer Society*, pp. 177–179, 2011.
- [37] W. C, W. W., Y. W., and L. J., "Deep retinex decomposition for low-light enhancement," *arXiv:2210.05436*, 2018.
- [38] Z.-U. Rahman, D. Jobson, and G. Woodell, "A multiscale retinex for color rendition and dynamic range compression," *Applications of Digital Image Processing*, p. 183–191, 1996.
- [39] D. Jobson, "Retinex processing for automatic image enhancement," *Journal of Electronic Imaging*, pp. 100–110, 2004.
- [40] L. Ma, T. Ma, R. Liu, X. Fan, and Z. Luo, "Toward fast, flexible, and robust low-light image enhancement," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2022, pp. 5637–5646.
- [41] Y. Zhang, J. Zhang, and X. Guo, "Kindling the darkness: A practical low-light image enhancer," in *ACM International Conference on Multimedia*, 2019, pp. 1632–1640.

- [42] D. Zhou, Z. Yang, and Y. Yang, "Pyramid diffusion models for low-light image enhancement," in *International Joint Conference on Artificial Intelligence*, 2023.
- [43] T. Wang, K. Zhang, Z. Shao, W. Luo, B. Stenger, T. Lu, T.-K. Kim, W. Liu, and H. Li, "Lldiffusion: Learning degradation representations in diffusion models for low-light image enhancement," *arXiv:2307.14659*, 2023.
- [44] L. C, G. C. L, and Z. M, "Embedding fourier for ultra-high-definition low-light image enhancement," in *IEEE International Conference on Learning Representations*, 2023.
- [45] S. M, "Simplified design of building lighting," *John Wiley Sons*, 1997.
- [46] G. I, "Elementarnaja fizika," *Nauka, Moskva.*, 1973.
- [47] H. Wu, Y. Qu, S. Lin, J. Zhou, R. Qiao, Z. Zhang, Y. Xie, and L. Ma, "Contrastive learning for compact single image dehazing," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2021, p. 10551–10560.
- [48] L. Wang, Y. Wang, X. Dong, Q. Xu, J. Yang, W. An, and Y. Guo, "Unsupervised degradation representation learning for blind super-resolution," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2021, p. 10581–10590.
- [49] X. Li, C. Chen, and X. Lin, "From face to natural image: Learning real degradation for blind image super-resolution," in *European Conference on Computer Vision*, 2022, p. 376–392.
- [50] K. Zhang, J. Liang, and V. G. L, "Designing a practical degradation model for deep blind image super-resolution," in *IEEE International Conference on Computer Vision*, 2021, pp. 4791–4800.
- [51] A. Golts, D. Freedman, and M. Elad, "Unsupervised single image dehazing using dark channel prior loss," *IEEE Transactions on Image Processing*, p. 2692–2701, 2019.
- [52] H. Zhang and V. M. Petel, "Density-aware single image de-raining using a multi-stream dense network," in *IEEE conference on Computer Vision and Pattern Recognition*, 2018, pp. 695–704.
- [53] L. W, Z. Q, and Z. J, "Toward real-world single image deraining: A new benchmark and beyond," *arxiv*, 2022.
- [54] S. Zhou, C. Li, and C. C. Loy, "Lednet: Joint low-light enhancement and deblurring in the dark," in *European Conference on Computer Vision*, 2022, pp. 573–589.
- [55] B. Y, Z. H, and Y. E, "Not just streaks: Towards ground truth for single image deraining," in *European Conference on Computer Vision*, 2022, pp. 723–740.
- [56] L. Chen, X. Lu, J. Zhang, X. Chu, and C. Chen, "Hinet: Half instance normalization network for image restoration," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2022, pp. 182–192.
- [57] Loshchilov, I., and H. F, "Sgdr: Stochastic gradient descent with warm restarts," in *arXiv:1608.03983*, 2016.
- [58] M. A., S. R., and Bovik, "Making a "completely blind" image quality analyzer," *IEEE Signal processing letters*, pp. 209–212, 2012.
- [59] Y. J, L. J, and F. X, "No-reference quality assessment of contrast-distorted images using contrast enhancement," *arXiv:1904.08879*, 2019.
- [60] T. Wang, K. Zhang, T. Shen, W. Luo, B. Stenger, and T. Lu, "Ultra-high-definition low-light image enhancement: A benchmark and transformer-based method," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2023, pp. 2654–2662.
- [61] Y. Li, K. Xu, G. P. Hancke, and R. W. Lau, "Color shift estimation-and-correction for image enhancement," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2024, pp. 25389–25398.



Xin Lin is currently pursuing a B.Eng. degree in electronics and information engineering from Sichuan University, Chengdu, China, under the supervision of Prof. Chao Ren. Now, he is a Visiting Student at the University of California, Merced, under the supervision of Prof. Ming-Hsuan Yang. He serves as a reviewer of some top-tier conferences and journals such as the CVPR, NeurIPS, ICLR, ECCV, ACM MM, IJCV, and Information Fusion. His research interests include 3D Vision and image restoration.



Jingtong Yue is currently pursuing a B.Eng. degree in Electronics and Information Engineering at Sichuan University in Chengdu, China, under the supervision of Professor Chao Ren. Now she is also involved in research on intelligent robots in space at the Institute of Space and Astronautics, Sichuan University. Her research interests include radar point cloud processing, image restoration, and multi-modal fusion.



Sixian Ding is currently pursuing a B.Eng. degree in Software Engineering at Sichuan University in Chengdu, China, under the guidance of Prof. Yan Wang. Her current research interests include image generation, image denoising, and facial expression recognition.



Chao Ren received the Ph.D degree in communication and information system from Sichuan University, Chengdu, China, in 2017. He is currently an Associate Research Professor at Sichuan University College of Electronics and Information Engineering. His research interests include inverse problems in image and video processing. He has authored more than 60 papers in journals/conferences such as TIP, CVPR, ICCV, NeurIPS etc. He has won the Huawei Spark Award and was selected for the National Post-Doctoral Program for Innovative Talents of China.



Lu Qi works as a postdoc with Prof. Ming-Hsuan Yang at UC Merced and has about 10,000 citations in Google Scholar. He received his Ph.D. degree from The Chinese University of Hong Kong in 2021 and obtained the Hong Kong Ph.D. Fellowship in 2017. His research interests include instance-level detection, image generation, and cross-modal pretraining. He was the senior program chair of AAAI 2023/2024 and area chair of NeurIPS2024 and ICLR2024.



Ming-Hsuan Yang is a professor of Electrical Engineering and Computer Science at the University of California, Merced. Yang serves as a program co-chair of the IEEE International Conference on Computer Vision (ICCV) in 2019, program co-chair of the Asian Conference on Computer Vision (ACCV) in 2014, and general co-chair of ACCV 2016. Yang currently serves as an associate editor-in-chief of IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) and as an associate editor of the International Journal of Computer Vision (IJCV), Image and Vision Computing, and Journal of Artificial Intelligence Research. He received the NSF CAREER award in 2012 and Google Faculty Award in 2009. He is a Fellow of the IEEE and ACM.