

# 基于视觉提示学习的天气退化图像恢复

文渊博<sup>1)</sup> 高 涛<sup>2)</sup> 安毅生<sup>1)</sup> 李子琦<sup>1)</sup> 陈 婷<sup>1)</sup>

<sup>1)</sup>(长安大学信息工程学院 西安 710064)

<sup>2)</sup>(长安大学数据科学与人工智能研究院 西安 710064)

**摘要** 尽管现有的天气退化图像恢复方法在单一天气去除任务上已经取得良好表现,但其无法适应真实场景下多变的天气类型。为此,本文提出一种基于视觉提示学习的天气退化图像恢复算法,其是预训练语言图像模型与天气退化图像恢复任务结合的新范式。该算法首先设计一个查询提示约束网络(Query Prompt Constrained Network, QPC-Net),其利用对比语言图像预训练模型中的文本编码器和图像编码器来根据给定的退化图像直接编码其对应真实背景的潜在描述特征。同时,该算法还包括一个示例提示引导网络(Example Prompt Guided Network, EPG-Net),其利用给定的示例图像来引导预训练扩散模型去除查询图像上对应的天气退化。相比类似设定的现有算法,本文算法在 8 个天气退化数据集上平均改善峰值信噪比 2.11 dB,平均改善结构相似性 4.74%。

**关键词** 计算机视觉;视觉提示学习;情境学习;图像恢复;扩散模型

**中图法分类号** TP391.4      **DOI 号** 10.11897/SP.J.1016.2024.02401

## Weather-Degraded Image Restoration Based on Visual Prompt Learning

WEN Yuan-Bo<sup>1)</sup> GAO Tao<sup>2)</sup> AN Yi-Sheng<sup>1)</sup> LI Zi-Qi<sup>1)</sup> CHEN Ting<sup>1)</sup>

<sup>1)</sup>(School of Information Engineering, Chang'an University, Xi'an 710064)

<sup>2)</sup>(School of Data Science and Artificial Intelligence, Chang'an University, Xi'an 710064)

**Abstract** Images captured in real-world scenarios often suffer from weather degradations like random occurrences of rain, haze and snow, which may cause detail occlusion and content deterioration, thereby impacting the effectiveness of subsequent advanced computer vision algorithms. Existing methods for weather-degraded image restoration can be categorized into task-specific, task-aligned and all-in-one types. However, the first two types require specific training for different weather degradations and struggle to adapt to the diverse weather conditions encountered in real-world scenes. Although all-in-one methods achieve the competitive performance across adverse weather degradation removal tasks, they also fail to adapt to the unseen weather degradations, resulting in poor generalization performance. To this end, a weather-degraded image restoration algorithm based on visual prompt learning is proposed in this work, which is a novel paradigm that integrates the pre-trained language-image model with the weather degraded-image restoration. Specifically, even text inputs with similar meanings may yield significantly different latent features when processed through the text encoder of contrastive language-image pre-training (CLIP) model. The general expectation of image restoration is to provide a degraded image and have the model generate its corresponding restored image, rather than multiple different reconstructed images. Therefore, directly using text to guide image reconstruction may lead to unstable solution spaces,

收稿日期:2023-11-23;在线发布日期:2024-06-26. 本课题得到国家重点研发计划(2023YFB2504703)、陕西省国际科技合作计划项目(2024GH-YBXM-24)、国家自然科学基金(52172379)、长安大学中央高校基本科研业务费专项资金(300102242901)资助。文渊博,博士研究生,主要研究方向为计算机视觉、智能交通系统。E-mail: wyb@chd.edu.cn。高 涛(通信作者),博士,教授,主要研究领域为人工智能、图像处理。E-mail: gaotao@chd.edu.cn。安毅生,博士,教授,主要研究领域为离散事件系统建模与应用、智能车路协同系统。李子琦,博士研究生,主要研究方向为计算机视觉、交通目标检测。陈 婷,博士,教授,主要研究领域为模式识别、智能交通系统。

often failing to meet the general expectation of image restoration. In response, a query prompt constrained network (QPC-Net) is introduced to utilize the text encoder and image encoder from CLIP to directly encode the latent descriptive features of corresponding ground truth based on the given degraded images. These latent features are further embedded into a pre-trained stable diffusion model using the cross-attention mechanism, thereby constraining the reverse sampling process and facilitating the content reconstruction. QPC-Net consists of two image encoders, with one set of parameters frozen and the other set trainable. Moreover, many existing weather-degraded image algorithms primarily train strict pixel-level mappings between the degraded and clean images, lacking the exploration of knowledge for different image restoration tasks. This limitation makes it difficult for these algorithms to learn the corresponding context for the weather-degraded image restoration tasks not covered in the training dataset, thereby struggling to adapt to different restoration tasks. To address this issue, an example prompt guided network (EPG-Net) is developed to utilize the given example images to guide pre-trained stable diffusion model in learning the context knowledge of corresponding restoration tasks, thereby removing the degradations from query images. Additionally, acquiring suitable example images for complex mixed weather-degraded image restoration tasks are challenging; however EPG-Net can learn the context knowledge from multiple sets of example images. In experimental evaluations conducted on eight seen weather degradation datasets and seven unseen datasets, the proposed algorithm demonstrates significant improvements. Specifically, on the seen weather-degraded datasets, it achieves an average improvement of 2.11 dB in peak signal-to-noise ratio (*PSNR*), 4.74% in structural similarity (*SSIM*), 41.08% in perceptual image block similarity (*LPIPS*) and 24.25% in natural image quality evaluator (*NIQE*) compared to existing algorithm with similar setting. Additionally, on the unseen weather-degraded datasets, it achieves an average improvement of 1.88 dB in *PSNR*, 5.61% in *SSIM*, 21.40% in *LPIPS* and 29.29% in *NIQE*.

**Keywords** computer vision; visual prompt learning; in-context learning; image restoration; diffusion model

## 1 引言

在雨、雾和雪等恶劣天气下获取的图像会因为天气退化的存在造成颜色失真、背景遮挡和细节丢失等问题,进而影响到后续计算机视觉算法的表现<sup>[1-2]</sup>. 在现有工作中,一些方法根据不同天气退化的物理特性建立专门的物理模型<sup>[3-4]</sup>,一些方法为不同的天气退化建立结构统一但需要分别训练的网络模型<sup>[5-6]</sup>. 尽管这些方法在单一的天气退化图像恢复任务中取得良好表现,但却无法有效地适应真实场景中多变的天气退化类型. 因此,探索同一的天气退化图像恢复方案已逐渐成为该领域研究的主流.

在基于视觉提示学习的图像恢复研究中,Bar 等人<sup>[7]</sup>利用网格输入来实现空白图像的修复,Wang 等人<sup>[8]</sup>将示例图像和查询图像一起输入网络来引导干净图像的生成. 随着去噪扩散概率模型(Denoising

Diffusion Probabilistic Model,DDPM)<sup>[9]</sup>的发展,稳定扩散模型(Stable Diffusion Models, SDM)<sup>[10]</sup>逐渐成为图像生成领域表现良好的方案. 近来,Zhang 等人<sup>[11]</sup>利用条件图像来引导预训练扩散模型生成目标图像. 然而,这类仅利用条件图像来引导目标图像生成的方法在天气退化图像恢复任务中无法直接应用. 此外,尽管一些以退化图像作为条件的扩散模型<sup>[12]</sup>能够进行天气退化的去除,但其严重依赖训练数据集覆盖的天气类型,进而无法应对未见的天气退化.

如图 1 所示,为解决以上问题,本文提出一种基于视觉提示学习的天气退化图像恢复算法. 在图像恢复任务中,普遍期望是给定一张退化图像,模型会输出一张对应的干净图像. 但是,稳定扩散模型 SDM 利用对比语言视觉预训练模型(Contrastive Language-Image Pre-training model,CLIP)<sup>[13]</sup>中的文本编码器(Text Encoder, TE)来编码给定的文本提示,进而

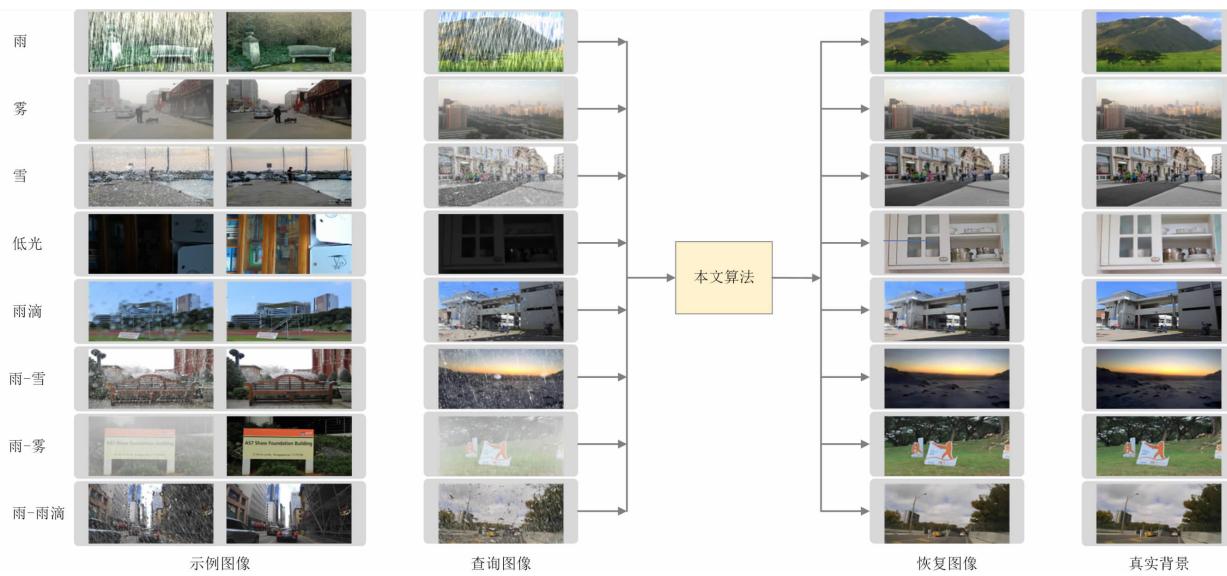


图 1 基于视觉提示学习的天气退化图像恢复示意图

约束反向采样过程并生成基于文本提示的图像。然而,含义接近的不同文本信息却有可能在潜在空间表现出较大差距。因此,这种方法不满足天气退化图像恢复的基本期望。同时,如果直接利用对比视觉预训练模型 CLIP 中图像编码器 (Image Encoder, IE) 来生成退化图像的潜在特征并将其送入预训练扩散模型的反向采样过程中,模型会由于退化特征的引入而无法生成预期的恢复图像。为此,本文提出查询提示约束网络 (Query Prompt Constrained Network, QPC-Net),其利用文本编码器 TE 和图像编码器 IE 来约束退化图像的潜在特征编码和真实背景的潜在描述特征之间的一致性,从而直接根据给定的退化图像生成与真实背景描述一致的潜在特征,进而更好地约束预训练稳定扩散模型 SDM 的反向采样过程。同时,为有效地利用预训练稳定扩散模型 SDM 实现情境学习,本文提出示例提示引导网络 (Example Prompt Guided Network, EPG-Net),其根据示例图像来引导模型去除查询图像上对应的天气退化类型,进而得到高质量的恢复图像。本文算法是预训练语言视觉模型与天气退化图像恢复任务结合的新范式。

## 2 相关工作

### 2.1 天气退化图像恢复

天气退化图像恢复领域的研究主要可以分为三类,即专一型、统一型和同一型。其中,前两者只能去除单独的天气退化,无法有效地处理不同的天气退化图像。在同一型算法中,Li 等人<sup>[14]</sup>利用神经搜索

网络和对应不同天气退化类型的编码器来处理天气退化图像。随后,Valanarasu 等人<sup>[15]</sup>引入空间稀疏自注意力机制和任务查询向量来进行多天气退化去除。类似地,Gao 等人<sup>[16]</sup>利用频率信息引导天气退化图像的特征编码并在通道、空间和小波域渐进地优化解码特征的频率表征。近来,Ozdenizci 等人<sup>[12]</sup>利用退化图像作为去噪扩散概率模型 DDPM 的条件来引导其生成去除天气退化的高质量背景图像。同时,Qin 等人<sup>[17]</sup>利用自注意力机制和门控注意力机制来改进扩散模型的噪声估计网络,并在多个天气退化图像恢复任务上取得良好表现。但是,这些方法未针对不同的天气退化提供查询约束或示例引导,从而导致其缺乏对不同天气退化的感知能力。

### 2.2 视觉提示学习

GPT-3<sup>[18]</sup>首次定义一种名为情境学习的学习范式,即一系列自然语言处理任务可以被视为在给定提示的情况下执行许多视觉语言任务,如视觉问答。随着情境学习的发展,视觉提示学习逐渐在下游计算机视觉任务中取得良好表现。例如,Zhang 等人<sup>[11]</sup>通过文本提示来约束预训练稳定扩散模型 SDM 的反向采样过程,进而利用姿态、轮廓和深度等条件图像来引导其生成不同的目标图像。在基于视觉提示学习的图像恢复任务中,Bar 等

人<sup>[7]</sup>基于图像修复的思想将示例图像和查询图像以网格的形式输入生成对抗网络,进而在推理过程中填补空缺以得到预期的恢复图像。然而,该方法严重限制图像的分辨率结构,使其无法实现任意分辨率图像恢复。同时,网格化的图像输入模式导致用户无法端到端地获取恢复图像,从而导致该方法的实用性较差。近来,Wang 等人<sup>[8]</sup>将示例图像和查询图像同时输入网络,并利用权值共享和特征聚合的策略来促使网络生成目标图像。但是,由于表征能力的局限性和缺乏对天气退化潜在特征的探索,该方法所得的恢复图像存在背景模糊和退化残留等问题。

### 2.3 传统扩散模型与稳定扩散模型

近年来,去噪扩散概率模型 DDPM<sup>[9]</sup>在图像生成和多个下游计算机视觉任务上取得良好表现,其主要原理是通过迭代去噪的模式将标准高斯噪声的采样变量恢复到数据分布。具体而言,去噪扩散概率模型 DDPM 可以分为前向扩散过程和反向采样过程。前者通过不断地向训练数据中添加标准高斯噪声来破坏数据,最终使破坏后的数据趋近于标准高斯分布。后者则是利用网络来近似噪声分布,进而通过迭代得到未被噪声破坏的结果。给定输入  $\mathbf{x}_0$ ,其满足初始先验分布  $q(\mathbf{x}_0)$ ,那么前向扩散过程可以定义为

$$q(\mathbf{x}_t | \mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1-\beta_t} \cdot \mathbf{x}_{t-1}, \beta_t \cdot \mathbf{I}) \quad (1)$$

式中,  $\mathbf{x}_t$  为输入  $\mathbf{x}_0$  在  $t$  时刻的噪化状态,  $\beta_t \cdot \mathbf{I}$  是高斯分布的方差,其中  $\beta_t$  为  $t$  时刻的权重系数。根据前向扩散过程的马尔科夫性,  $\mathbf{x}_t$  可以直接由输入  $\mathbf{x}_0$  得到,即

$$q(\mathbf{x}_{1:T} | \mathbf{x}_0) = \prod_{t=1}^T q(\mathbf{x}_t | \mathbf{x}_{t-1}) \quad (2)$$

式中,  $T$  为前向扩散的次数。反向采样过程即学习一个噪声估计网络  $f_\theta(\mathbf{x}_t, t)$  来估计  $t$  时刻的噪声分布  $\epsilon_t$ , 其中  $\theta$  是噪声估计网络的可学习参数。根据前向扩散过程可知,最终的破坏图像近似为标准高斯分布,即  $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ , 那么初始后验分布满足:

$$p(\mathbf{x}_T) = \mathcal{N}(\mathbf{x}_T; \mathbf{0}, \mathbf{I}) \quad (3)$$

此时,估计噪声后验分布满足:

$$p_\theta(\mathbf{x}_{t-1} | \mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_\theta(\mathbf{x}_t, t), \Sigma_\theta(\mathbf{x}_t, t)) \quad (4)$$

式中,  $\mu_\theta(\mathbf{x}_t, t)$  和  $\Sigma_\theta(\mathbf{x}_t, t)$  为噪声估计网络得到的噪声分布参数。同理,根据马尔可夫性,  $\mathbf{x}_0$  的后验分布可以由  $\mathbf{x}_T$  的后验分布表达为

$$p_\theta(\mathbf{x}_{0:T}) = p(\mathbf{x}_T) \prod_{t=1}^T p_\theta(\mathbf{x}_{t-1} | \mathbf{x}_t) \quad (5)$$

然而,去噪扩散概率模型 DDPM 在图像空间添加和去除噪声的操作导致其生成一张高分辨率的图

像非常耗时。为克服以上不足,稳定扩散模型 SDM<sup>[10]</sup>利用编码器来将高分辨率的输入映射到潜在空间,进而在潜在空间进行扩散模型的加噪与去噪过程,最后将结果再通过解码器输出。

## 3 本文算法

### 3.1 查询提示约束网络

稳定扩散模型 SDM 缺乏与天气退化相关的约束和引导来控制反向采样过程,从而导致其无法简单地应用到天气退化图像恢复任务中。在天气退化图像恢复任务中,普遍期望是给定一幅退化图像,模型会去除天气退化并生成对应的恢复图像。根据文献[20]的研究可知,具有相同含义的文本提示在经过文本编码器 TE 后也可能得到不同的潜在特征。如图 2 所示,对于一幅雨天退化图像,我们将其通过图像编码器 IE 得到图像潜在特征,同时将“rain-degraded image”和“rainy image”两种含义近似的正文本提示分别通过文本编码器 TE 得到文本潜在特征。负文本提示均为“rain-free image”。通过对比文本与图像潜在特征的相似度,实验发现两种文本提示的潜在特征存在较大差异。因此,利用文本编码器 TE 来约束模型的反向采样过程会由于文本提示的不唯一性导致解空间的不稳定。然而,对于一幅给定的退化图像,其经过图像编码器 IE 后所得的潜在特征是唯一的,这说明采用图像编码器来约束模型的反向采样过程能够避免解空间的不稳定问题。但是,直接利用预训练的图像编码器 IE 会导致天气退化的潜在特征也被作为约束特征嵌入模型,从而干扰其对天气退化的去除。

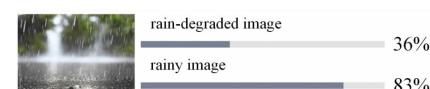


图 2 含义近似的不同文本与同一雨图的潜在空间相似度

为解决以上问题,本文提出查询提示约束网络 QPC-Net。如图 3 所示,该网络利用两组图像编码器 IE 对天气退化图像进行潜在特征的联合编码。其中,一组图像编码器 IE 的参数是冻结的,另一组图像编码器 IE 的参数是可学习的。参数冻结指的是在网络学习过程中预训练的参数不会被改变。可学习图像编码器(Learnable Image Encoder, LIE)的每一层特征均通过一个零初始化卷积与冻结图像编码器(Fixed Image Encoder, FIE)的每一层特征相加并作为 FIE 下一层网络的输入特征。本文使用余弦相似

度损失来训练 LIE 的参数。在训练过程中,本文首先利用自引导语言图像预训练模型(Bootstrapping Language-Image Pre-training model, BLIP)<sup>[21]</sup>生成真实背景的图像描述。该图像描述通过文本编码器 TE 得到当前退化图像对应的真实背景的潜在描述特征。训练的目的是令查询提示约束网络 QPC-Net 根据给定的退化图像直接生成与真实背景的潜在描述特征一致的无退化特征编码。随后,查询提示约束网络 QPC-Net 的网络参数被冻结,进而取代稳定扩散模型 SDM 中的文本编码器 TE,以确保模型满足天气退化图像恢复的普遍期望。随后,查询提示约束网络 QPC-Net 输出的潜在特征通过交叉自注意力机制<sup>[22]</sup>来生成注意力图,进而对稳定扩散模型 SDM 反向采样过程的中间特征进行约束。本文提出的查询提示约束网络 QPC-Net 一方面避免退化因素干扰模型的反向采样过程,另一方面能够促进模型对图像内容的重建。

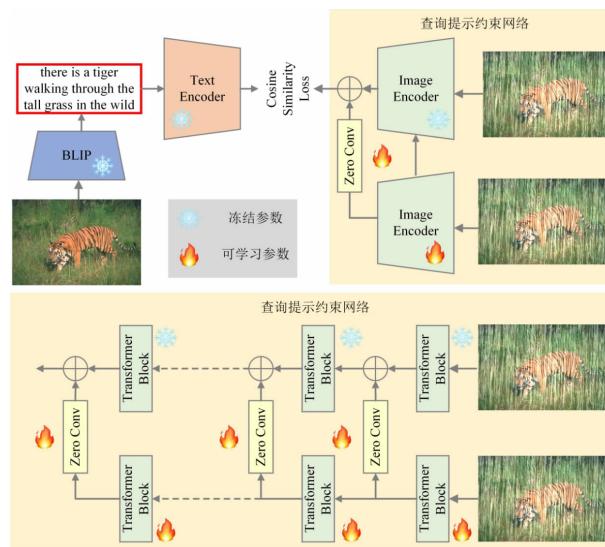


图 3 查询提示约束网络示意图

### 3.2 示例提示引导网络

近来,在基于预训练稳定扩散模型 SDM 的可控图像生成任务中,Zhang 等人<sup>[11]</sup>利用轮廓和姿势等条件图像来引导预训练稳定扩散模型 SDM 的反向采样过程,进而生成受控制的图像。但是,该方法所得的生成图像往往存在较大的解空间,从而导致多次生成的结果存在很大差异,这与天气退化图像恢复任务的普遍期望不符。为此,本文提出示例提示引导网络 EPG-Net,其利用成对的示例图像来引导模型进行情境学习,进而去除查询图像中对应的天气退化。图 4 所示是本文算法的一次反向采样过程。在该过程中,示例提示引导网络 EPG-Net 采用与

稳定扩散模型一样的编码块(Encoder Block, EB)和中间块(Middle Block, MB)。不同的是,在 EPG-Net 中这些网络结构是可学习的,而稳定扩散模型 SDM 中这些网络结构的参数是冻结的。在示例提示引导网络 EPG-Net 中,编码块 EB 和中间块 MB 的输出通过跳跃连接分别输入到零初始化卷积层中。随后,本文通过特征逐层相加的方式将示例提示引导网络的 EPG-Net 的中间特征与稳定扩散模型 SDM 的对应层特征进行融合,从而实现受 EPG-Net 引导的反向采样过程。图 4 中每一层的零初始化卷积的数量实际上与稳定扩散模型中每一层的编码块 EB 的数量是分别对应的,这里为简化表示每一层只画出一个编码块 EB 和一个零初始化卷积。同时,在反向采样过程中,本文算法使用冻结的查询提示约束网络 QPC-Net 来将退化图像的无退化潜在特征嵌入到模型中。稳定扩散模型 SDM 能够利用查询提示约束网络 QPC-Net 将高分辨率的输入映射到潜在空间。但是,示例图像和查询图像均在图像空间,从而无法直接将其嵌入到稳定扩散模型 SDM 的噪声估计网络中。因此,示例提示学习网络 EPG-Net 利用与稳定扩散模型一致的图像编码器得到与潜在空间的分辨率一致的特征。随后查询图像和示例图像的潜在特征通过零初始化卷积相加后输入到编码块 EB 中。

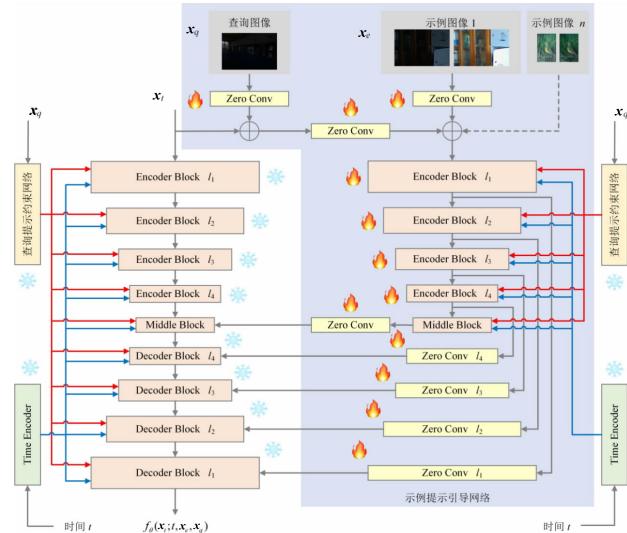


图 4 示例提示引导网络示意图

## 4 实验与分析

### 4.1 实现细节

本文算法的深度学习框架为 Pytorch 1.9,训练环境为 4 张 NVIDIA Tesla A800 GPU,测试环境为

1 张 NVIDIA GeForce RTX 3090 GPU. 训练查询提示约束网络 QPC-Net 的学习率为  $3 \times 10^{-5}$ , 训练次数为 100, 训练耗时为 11 h 37 min. 本文训练示例提示引导网络 EPG-Net 的学习率和迭代次数分别为  $1 \times 10^{-4}$  和 10 000, 训练耗时为 32 h 15 min. 两次训练的参数优化器均选择 AdamW<sup>[23]</sup> 优化器. 两次训练均是全监督的.

## 4.2 数据集

实验分别验证本文算法在单天气和混合天气退化图像恢复两种任务上的表现. 其中, 单天气退化包括雨、雾、雪、低光照和雨滴, 混合天气退化包括雨-雪、雨-雾和雨-雨滴. 实验所采用的数据集的划分细节如表 1 所示. 其中, Snow100K-L 是 Snow100K 数据集中的大雪测试集, Raindrop-A 是 Raindrop 数据集的密集雨滴测试集, RS100K-L 是 RS100K 数据集的密集雨-雪测试集, Test1 是 Outdoor-Rain 数据集的测试集, RainDS-syn 是 RainDS 数据集的合成雨-雨滴测试集. 实验将这 8 种天气退化数据集的训练集整合成一个综合数据集.

表 1 实验数据集的天气类型及数据集划分细节

天气	数据集	训练图像对	测试图像对
雨	Rain200H <sup>[24]</sup>	1800	200
雾	RESIDE-6K <sup>[25]</sup>	6000	1000
雪	Snow100K-L <sup>[26]</sup>	50 000	16 801
低光照	LOL <sup>[27]</sup>	485	15
雨滴	Raindrop-A <sup>[28]</sup>	861	58
雨-雪	RS100K-L <sup>[29]</sup>	50 000	2000
雨-雾	Test1 <sup>[30]</sup>	9000	750
雨-雨滴	RainDS-syn <sup>[31]</sup>	1000	200

## 4.3 评估指标

为验证本文算法在天气退化图像恢复任务中的表现, 实验选择峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR)、结构相似性 (Structural SIMilarity, SSIM)<sup>[32]</sup> 和学习感知图像块相似度 (Learned Perpetual Image Patch Similarity, LPIPS)<sup>[33]</sup> 三个有监督指标来衡量恢复图像与真实背景之间的相似程度. 其中, PSNR 和 SSIM 的值越大代表图像的质量越高, LPIPS 反之. 同时, 实验还选择自然图像质量评估器 (Natural Image Quality Evaluator, NIQE)<sup>[34]</sup> 这一无监督图像质量评价指标来衡量恢复图像的自然度, 其值越小表明恢复图像越接近真实图像.

## 4.4 算法比较

为验证本文算法在各种不同的天气退化图像恢复任务中的表现, 实验分别对比本文算法与现有算法在雨、雾、雪、低光、雨滴、雨-雪、雨-雾和雨-雨滴

8 种天气退化图像上的表现. 在进行每组测试实验时, 本文算法采用测试数据集的第一对图像作为示例图像来引导网络推理. 在以下表格中, 加粗和加下划线的数据分别为排名第一和排名第二的指标. 类型 A、B 和 C 分别代表特定天气退化图像恢复方法、同一型天气退化图像恢复方法和基于视觉提示学习的天气退化图像恢复算法. 这三类算法均在与本文算法一致的综合数据集上进行训练.

### 4.4.1 图像去雨

表 2 所示为不同算法在 Rain200H 数据集上的定量表现. 具体而言, 本文算法在有监督指标 PSNR、SSIM 和 LPIPS 上分别优于 Painter 算法 1.25 dB、5.95% 和 37.64%. 同时, 本文算法在无监督指标上也优于 Painter. 此外, 相比任务特定和任务统一的现有算法, 本文算法在 Rain200H 数据集上也具有明显优势. 为更好地对比本文算法与现有算法在图像去雨任务上的表现, 实验还将不同算法在 Rain200H 数据集上的视觉对比在图 5 中展示. 如图所示, 本文算法能够根据视觉提示去除雨痕, 而 Painter 所得的结果中残留雨痕, 同时图像的细节纹理被破坏.

表 2 不同算法在 Rain200H 数据集上的定量表现

类别	方法	PSNR	SSIM	LPIPS	NIQE
A	Restormer <sup>[35]</sup>	28.11	<u>0.8845</u>	0.1334	5.032
	DehazeFormer <sup>[36]</sup>	27.37	0.8477	0.1758	4.550
	SnowFormer <sup>[37]</sup>	26.42	0.8467	0.1554	4.348
	LLFormer <sup>[38]</sup>	26.62	0.8310	0.2123	4.469
	IDT <sup>[39]</sup>	<u>28.13</u>	0.8593	0.1605	4.633
	SASCFormer <sup>[31]</sup>	26.76	0.8413	0.1774	4.081
	DPENet <sup>[40]</sup>	13.08	0.3734	0.6060	10.700
	UDR-S <sup>2</sup> Former <sup>[41]</sup>	27.67	0.8562	0.1566	4.750
B	All-in-One <sup>[14]</sup>	21.55	0.7440	0.3219	4.310
	TransWeather <sup>[15]</sup>	25.74	0.8302	0.1653	4.881
	AIRFormer <sup>[16]</sup>	24.46	0.7424	0.3154	7.129
	TKL <sup>[42]</sup>	26.14	0.8463	0.2205	5.767
	AWIR-TDM <sup>[17]</sup>	27.87	0.8609	0.1265	<u>3.992</u>
	WeatherDiff <sup>[12]</sup>	26.66	0.8541	<u>0.1071</u>	4.007
C	Painter <sup>[8]</sup>	27.51	0.8434	0.1472	4.630
	本文算法	<b>28.76</b>	<b>0.8936</b>	<b>0.0918</b>	<b>3.856</b>

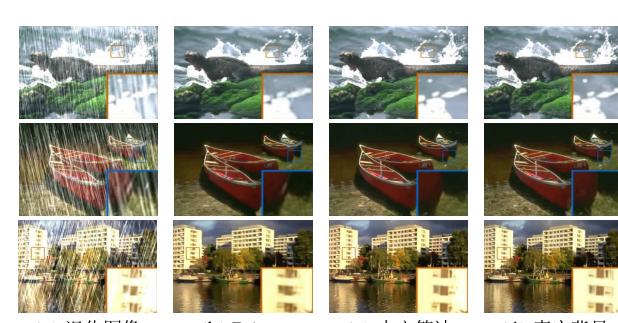


图 5 不同算法在 Rain200H 数据集上的视觉对比

#### 4.4.2 图像去雾

如表 3 所示,在 RESIDE-6K 数据集上,本文算法在除 SSIM 外的其他指标上均取得最佳的分数。同时,相比类似方法 Painter,本文算法所得的指标分数均具有明显优势。图 6 所示为不同算法的视觉对比。其中,Painter 的恢复图像存在明显的雾退化,而本文算法所得结果更接近真实背景。

表 3 不同算法在 RESIDE-6K 数据集上的定量表现

类别	方法	PSNR	SSIM	LPIPS	NIQE
A	Restormer <sup>[35]</sup>	28.96	<b>0.9774</b>	0.0344	4.733
	DehazeFormer <sup>[36]</sup>	28.74	0.9744	0.0333	4.546
	SnowFormer <sup>[37]</sup>	28.45	0.9739	0.0352	4.595
	LLFormer <sup>[38]</sup>	26.94	0.9649	0.0412	4.492
	IDT <sup>[39]</sup>	28.93	0.9735	0.0373	4.450
	SASCFormer <sup>[31]</sup>	22.73	0.9015	0.1273	<b>4.176</b>
	DPENet <sup>[40]</sup>	15.30	0.8246	0.1585	4.473
B	UDR-S <sup>2</sup> Former <sup>[41]</sup>	26.45	0.9652	0.0479	4.407
	All-in-One <sup>[14]</sup>	19.31	0.8761	0.1689	4.559
	TransWeather <sup>[15]</sup>	26.69	0.9653	0.0368	4.406
	AIRFormer <sup>[16]</sup>	25.67	0.9513	0.0639	4.302
	TKL <sup>[42]</sup>	25.52	0.9619	0.0497	4.553
	AWIR-TDM <sup>[17]</sup>	23.52	0.9337	0.0684	4.680
	WeatherDiff <sup>[12]</sup>	24.30	0.9540	0.0514	4.541
C	Painter <sup>[8]</sup>	25.91	0.8446	0.2065	6.150
	本文算法	<b>29.08</b>	0.9601	<b>0.0294</b>	<b>4.089</b>

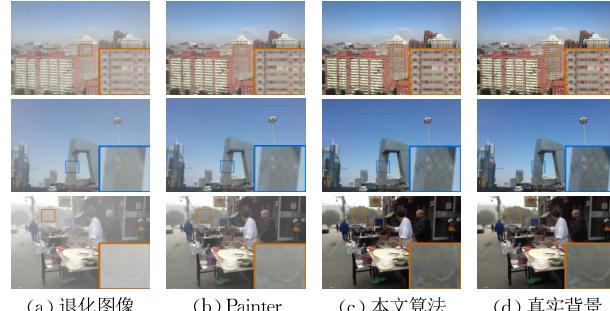


图 6 不同算法在 RESIDE-6K 数据集上的视觉对比

#### 4.4.3 图像去雪

为验证本文算法在图像去雪任务上的有效性,实验对比本文算法与现有算法在 Snow100K-L 数据集上的表现。如表 4 所示,相比 Painter 算法,本文算法的所有指标分数均为最优。具体而言,本文算法在 PSNR 上提升 1.50 dB,在 SSIM 上提升 1.51%,在 LPIPS 上降低 7.93%,在 NIQE 上明显低于 Painter。在 PSNR 和 SSIM 两个指标上,本文算法的结果略差于 A 类算法。同时,实验将不同算法在 Snow100K-L 数据集上的视觉对比在图 7 中进行展示。结果表明,两种方法均能有效去除雪图中的小颗粒雪花。但是,对于大颗粒雪花,本文算法能够有效地去除并恢复背景细节,而 Painter 所得的恢复图像中存在明显的雪退化残留。

表 4 不同算法在 Snow100K-L 数据集上的定量表现

类别	方法	PSNR	SSIM	LPIPS	NIQE
A	Restormer <sup>[35]</sup>	31.51	<b>0.9270</b>	0.0858	4.393
	DehazeFormer <sup>[36]</sup>	30.86	0.9107	0.1060	4.107
	SnowFormer <sup>[37]</sup>	29.86	0.9122	0.1020	4.031
	LLFormer <sup>[38]</sup>	30.74	0.9099	0.1053	3.949
	IDT <sup>[39]</sup>	31.68	0.9187	0.0907	4.017
	SASCFormer <sup>[31]</sup>	30.34	0.9081	0.1067	3.759
	DPENet <sup>[40]</sup>	20.11	0.7015	0.3434	3.857
B	UDR-S <sup>2</sup> Former <sup>[41]</sup>	<b>31.73</b>	0.9215	0.0923	4.204
	All-in-One <sup>[14]</sup>	26.69	0.8829	0.1663	3.946
	TransWeather <sup>[15]</sup>	30.10	0.9039	0.0907	3.854
	AIRFormer <sup>[16]</sup>	28.56	0.8684	0.1678	3.853
	TKL <sup>[42]</sup>	30.55	0.9176	0.0989	4.278
	AWIR-TDM <sup>[17]</sup>	30.41	0.9077	0.1100	3.982
	WeatherDiff <sup>[12]</sup>	28.99	0.8941	0.0880	3.733
C	Painter <sup>[8]</sup>	28.97	0.8965	0.0744	4.918
	本文算法	29.47	0.9100	<b>0.0685</b>	<b>3.463</b>

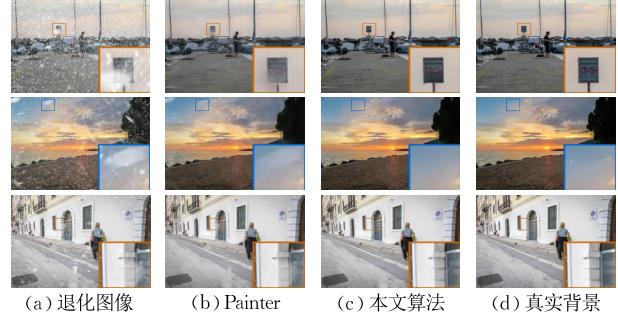


图 7 不同算法在 Snow100K-L 数据集上的视觉对比

#### 4.4.4 图像增强

不同算法在图像增强任务上的定量表现如表 5 所示。由表可知,本文算法在该数据集上的表现优于现有算法 Painter。具体而言,本文算法在有监督指标 PSNR、SSIM 和 LPIPS 上分别改善 1.79 dB、2.31% 和 17.11%,同时这些指标也是所有对比方法中分数最高的。在 NIQE 指标上,本文算法所得指标

表 5 不同算法在 LOL 数据集上的定量表现

类别	方法	PSNR	SSIM	LPIPS	NIQE
A	Restormer <sup>[35]</sup>	21.07	0.8940	0.1700	4.659
	DehazeFormer <sup>[36]</sup>	24.14	0.8907	0.1955	4.976
	SnowFormer <sup>[37]</sup>	21.62	0.8697	0.1915	4.673
	LLFormer <sup>[38]</sup>	24.00	0.8794	0.2036	4.206
	IDT <sup>[39]</sup>	21.18	0.8778	0.1919	<b>3.745</b>
	SASCFormer <sup>[31]</sup>	22.54	0.8744	0.2047	3.880
	DPENet <sup>[40]</sup>	15.00	0.7136	0.3960	4.702
B	UDR-S <sup>2</sup> Former <sup>[41]</sup>	20.39	0.8647	0.2287	4.316
	All-in-One <sup>[14]</sup>	18.95	0.8213	0.4023	4.333
	TransWeather <sup>[15]</sup>	23.46	0.8896	0.1677	4.645
	AIRFormer <sup>[16]</sup>	22.23	0.8312	0.3195	4.146
	TKL <sup>[42]</sup>	20.58	0.8734	0.2235	4.630
	AWIR-TDM <sup>[17]</sup>	17.49	0.8187	0.3094	4.802
	WeatherDiff <sup>[12]</sup>	18.13	0.8426	0.2262	4.962
C	Painter <sup>[8]</sup>	25.84	0.8831	0.1677	5.399
	本文算法	<b>27.63</b>	<b>0.9035</b>	<b>0.1390</b>	4.016

分数排名第三,但优于现有方法 Painter。图 8 所示为不同算法在 LOL 数据集上的增强结果视觉对比。其中,本文算法处理第一幅图所得结果的亮度改善优于 Painter 所得增强图像。第二幅和第三幅图表明,尽管 Painter 能够增强亮度,但所得结果存在细节丢失的问题。

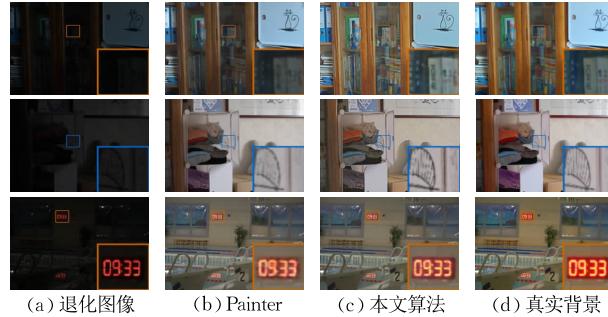


图 8 不同算法在 LOL 数据集上的视觉对比

#### 4.4.5 图像去雨滴

实验将本文算法与对比方法在图像去雨滴任务上的定量对比在表 6 中展示,其中本文算法在 PSNR、LPIPS 和 NIQE 三个评价指标上均优于所有对比方法,在 SSIM 指标上排名第二。具体而言,本文算法在 PSNR 上相比 Painter 取得 1.60 dB 的优势。图 9 所示为不同算法在 Raindrop-A 数据集上的恢复结果视觉对比。可以发现,现有算法 Painter 恢复的结果图像中出现大量的白孔。这表明,尽管 Painter 能够准确定位雨滴的位置,但却无法有效地根据退化图像的上下文信息来重建图像的纹理细节。不同的是,本文算法不仅能够很好地定位雨滴的位置,同时所得图像的纹理细节更加丰富准确。

表 6 不同算法在 Raindrop-A 数据集上的定量表现

类别	方法	PSNR	SSIM	LPIPS	NIQE
A	Restormer <sup>[35]</sup>	<b>30.84</b>	<b>0.9340</b>	0.0874	4.086
	DehazeFormer <sup>[36]</sup>	30.09	0.9187	0.1108	3.898
	SnowFormer <sup>[37]</sup>	30.07	0.9226	0.1040	4.193
	LLFormer <sup>[38]</sup>	30.12	0.9149	0.1257	4.032
	IDT <sup>[39]</sup>	29.72	0.9216	0.1014	3.836
	SASCFormer <sup>[31]</sup>	28.99	0.9081	0.1274	3.640
	DPENet <sup>[40]</sup>	24.53	0.8705	0.1883	3.829
	UDR-S <sup>2</sup> Former <sup>[41]</sup>	29.61	0.9209	0.1145	4.197
B	All-in-One <sup>[14]</sup>	23.41	0.8722	0.2099	4.099
	TransWeather <sup>[15]</sup>	28.83	0.9118	0.0891	4.078
	AIRFormer <sup>[16]</sup>	27.21	0.8777	0.1808	<u>3.619</u>
	TKL <sup>[42]</sup>	27.39	0.9170	0.1085	3.773
	AWIR-TDM <sup>[17]</sup>	28.01	0.9012	0.1457	3.900
C	WeatherDiff <sup>[12]</sup>	25.08	0.8934	0.1096	3.818
	Painter <sup>[8]</sup>	27.88	0.8520	0.1947	4.702
	本文算法	29.48	0.9131	<b>0.0813</b>	<b>3.539</b>

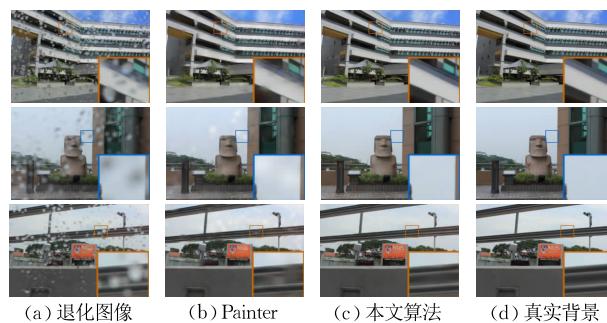


图 9 不同算法在 Raindrop-A 数据集上的视觉对比

#### 4.4.6 图像去雨-雪

为验证本文算法在图像去雨-雪任务上的有效性,实验在表 7 中列出本文算法与对比算法在 RS100K-L 数据集上的定量表现。如表所示,本文算法在 PSNR 上相比 Painter 提升 3.06 dB,在 SSIM 和 LPIPS 上均取得排名最高的分数,同时在 NIQE 上也取得最佳的自然度分数。为更好地分析本文算法的优势,实验还将本文算法与对比算法的视觉对比在图 10 中进行展示。如图所示,两种算法均能有效地去除退化图像中的雨痕,但 Painter 无法去

表 7 不同算法在 RS100K-L 数据集上的定量表现

类别	方法	PSNR	SSIM	LPIPS	NIQE
A	Restormer <sup>[35]</sup>	30.98	<b>0.9205</b>	0.0942	4.465
	DehazeFormer <sup>[36]</sup>	30.35	0.9022	0.1172	4.224
	SnowFormer <sup>[37]</sup>	29.23	0.9032	0.1110	4.123
	LLFormer <sup>[38]</sup>	30.13	0.8998	0.1187	4.060
	IDT <sup>[39]</sup>	<u>31.13</u>	0.9103	0.1017	4.121
	SASCFormer <sup>[31]</sup>	29.95	0.8998	0.1183	3.885
	DPENet <sup>[40]</sup>	18.96	0.6557	0.3855	<u>3.984</u>
	UDR-S <sup>2</sup> Former <sup>[41]</sup>	<b>31.21</b>	0.9131	0.1049	4.317
B	All-in-One <sup>[14]</sup>	26.35	0.8711	0.1825	4.096
	TransWeather <sup>[15]</sup>	29.37	0.8925	0.1026	3.988
	AIRFormer <sup>[16]</sup>	28.01	0.8540	0.1898	4.127
	TKL <sup>[42]</sup>	30.00	0.9081	0.1136	4.396
	AWIR-TDM <sup>[17]</sup>	29.34	0.8990	0.1269	4.274
C	WeatherDiff <sup>[12]</sup>	28.86	0.8891	0.1014	3.993
	Painter <sup>[8]</sup>	25.48	0.8774	0.1025	4.833
	本文算法	28.54	0.9046	<b>0.0844</b>	<b>3.928</b>



图 10 不同算法在 RS100K-L 数据集上的视觉对比

除较大颗粒的雪花。不同的是,本文算法在彻底去除所有退化的同时还能够有效地保护背景细节,进而得到更接近真实背景的恢复结果。

#### 4.4.7 图像去雨-雾

表 8 所示为不同算法在 Test1 数据集上的指标对比。可以发现,本文算法取得优于所有对比方法的表现。具体而言,相比现有算法 Painter,本文算法在有监督指标上分别改善 2.21 dB、4.81% 和 18.21%。同时,在无监督指标 NIQE 的分数表明本文算法所得恢复图像的自然度更好。另外,如图 11 所示,两种方法均能有效地去除退化图像中的雾。然而,Painter 处理所得的恢复图像中存在残留雨痕或丢失纹理细节。如,Painter 处理第一幅退化图像后,图中的栅栏处发生明显模糊,而第二幅和第三幅退化图像的处理结果中均残留雨痕。然而,本文算法不仅能够有效地去除退化图像中的雨痕和雾,同时可以很好地保护图像的细节,从而恢复出更加接近真实背景图像的结果。

表 8 不同算法在 Test1 数据集上的定量表现

类别	方法	PSNR	SSIM	LPIPS	NIQE
A	Restormer <sup>[35]</sup>	28.01	<b>0.9166</b>	0.1061	4.629
	DehazeFormer <sup>[36]</sup>	29.03	0.8919	0.1302	4.160
	SnowFormer <sup>[37]</sup>	27.81	0.8888	0.1185	4.343
	LLFormer <sup>[38]</sup>	27.72	0.8778	0.1557	4.146
	IDT <sup>[39]</sup>	29.80	0.9016	0.1184	4.328
	SASCFormer <sup>[31]</sup>	26.67	0.8770	0.1394	3.928
	DPENet <sup>[40]</sup>	14.46	0.5906	0.5428	6.579
B	UDR-S <sup>2</sup> Former <sup>[41]</sup>	28.86	0.9001	0.1160	4.158
	All-in-One <sup>[14]</sup>	19.33	0.8012	0.3171	4.436
	TransWeather <sup>[15]</sup>	26.91	0.8682	0.1196	3.941
	AIRFormer <sup>[16]</sup>	24.52	0.7837	0.3025	4.491
	TKL <sup>[42]</sup>	25.76	0.8852	0.1775	4.600
	AWIR-TDM <sup>[17]</sup>	28.65	0.9103	0.0944	3.918
	WeatherDiff <sup>[12]</sup>	25.78	0.8990	0.0736	3.788
C	Painter <sup>[8]</sup>	28.26	0.8801	0.0862	5.079
	本文算法	<b>30.47</b>	<b>0.9124</b>	<b>0.0705</b>	<b>3.693</b>

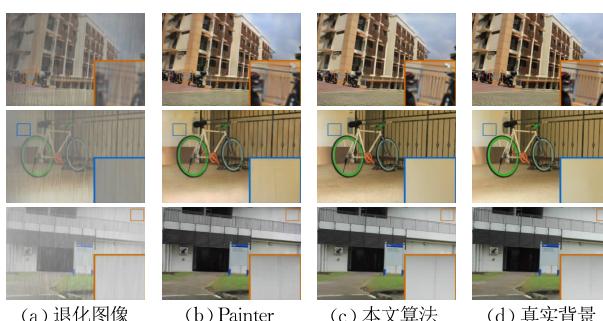


图 11 不同算法在 Test1 数据集上的视觉对比

#### 4.4.8 图像去雨-雨滴

为验证本文算法在雨-雨滴天气退化去除任务上的有效性,实验在表 9 中报告本文算法与对比算

法在 RainDS-syn 数据集上的定量表现。由表可知,本文算法在所有评价指标上均取得目前最好的表现。具体而言,相比现有算法 Painter,本文算法在 PSNR 上提升 2.31 dB,在 SSIM 上提升 1.23%,在 LPIPS 上改善 30.94%。同时,本文算法所得的 NIQE 指标明显低于 Painter。因此,本文算法的定量表现优于现有算法。同理,实验将不同算法在 RainDS-syn 数据集上的视觉对比在图 12 中展示。可以发现,现有算法无法有效地去除雨痕和雨滴,从而导致所得结果存在纹理破坏和细节失真的现象。相比而言,本文算法能够很好地分离退化因素和真实背景,从而恢复出细节丰富且纹理清晰的重建图像。

表 9 不同算法在 RainDS-syn 数据集上的定量表现

类别	方法	PSNR	SSIM	LPIPS	NIQE
A	Restormer <sup>[35]</sup>	29.48	0.9147	0.0983	4.559
	DehazeFormer <sup>[36]</sup>	28.63	0.8693	0.1653	4.678
	SnowFormer <sup>[37]</sup>	27.70	0.8673	0.1683	4.518
	LLFormer <sup>[38]</sup>	27.44	0.8477	0.2016	4.723
	IDT <sup>[39]</sup>	29.30	0.8855	0.1408	4.399
	SASCFormer <sup>[31]</sup>	29.06	0.8779	0.1514	<b>4.074</b>
	DPENet <sup>[40]</sup>	19.26	0.5776	0.4685	6.191
B	UDR-S <sup>2</sup> Former <sup>[41]</sup>	28.63	0.8822	0.1464	4.642
	All-in-One <sup>[14]</sup>	20.81	0.6950	0.3980	5.603
	TransWeather <sup>[15]</sup>	27.03	0.8428	0.1453	5.185
	AIRFormer <sup>[16]</sup>	24.25	0.7174	0.3467	6.511
	TKL <sup>[42]</sup>	26.69	0.8642	0.1820	4.856
	AWIR-TDM <sup>[17]</sup>	29.75	0.9043	0.0984	4.451
	WeatherDiff <sup>[12]</sup>	27.64	0.8981	0.0927	4.257
C	Painter <sup>[8]</sup>	29.19	0.9074	0.1154	4.232
	本文算法	<b>31.50</b>	<b>0.9186</b>	<b>0.0797</b>	<b>3.685</b>

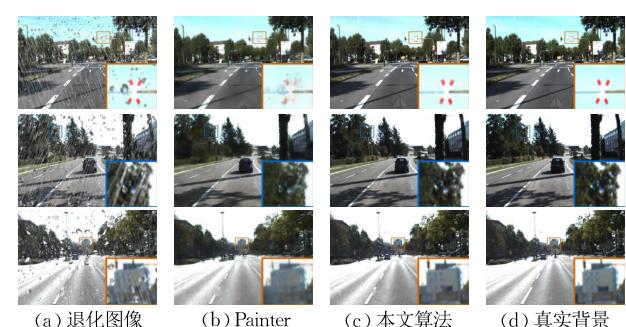


图 12 不同算法在 RainDS-syn 数据集上的视觉对比

#### 4.5 消融实验

##### 4.5.1 网络组件

为验证本文所提网络组件的有效性,实验对不同的组件分别进行消融实验。实验首先随机从 8 个测试数据集中分别挑选 11 对退化图像与对应的真实背景图像组成新的测试数据集,其中第一对图像作为对应天气退化图像恢复任务的示例图像。表 10 所示为本文算法关于网络组件的消融实验结果。其

中, w/o 示例图像表示本文算法使用 QPC-Net 和 EPG-Net, 但不使用示例图像进行提示学习。此时算法的整体表现在 PSNR 上下降 1.39 dB, 因此采用视觉提示学习来引导天气退化图像恢复是有效的。w/o QPC-Net 表示本文算法不使用 QPC-Net, 而是直接使用预训练图像编码器 IE 来编码退化图像。结果表明, 不使用查询提示约束网络 QPC-Net 导致算法的整体表现下降 4.91 dB, 因此该网络确实可以减少退化因素给反向采样过程带来的干扰, 同时有助于模型重建退化图像的内容。w/o EPG-Net 表示本文算法不使用 EPG-Net, 而是直接利用可学习的稳定扩散模型 SDM 来进行天气退化图像恢复。结果表明, 不使用示例提示引导网络导致算法的整体表现下降 3.05 dB。因此, 本文提出的 EPG-Net 可以引导模型进行情境学习, 从而利用示例图像来取得更好地天气退化图像恢复表现。综上所述, 本文算法的网络组件均对最终的图像恢复表现具有正向促进作用。

表 10 本文算法的网络组件的消融实验

模型	PSNR	SSIM	LPIPS	NIQE
w/o 示例图像	29.33	0.9139	0.0941	4.796
w/o QPC-Net	25.81	0.8637	0.2105	6.382
w/o EPG-Net	27.67	0.8956	0.1770	5.698
本文算法	<b>30.72</b>	<b>0.9384</b>	<b>0.0115</b>	<b>4.100</b>

#### 4.5.2 单组示例图像

为判断本文算法的情境学习能力, 实验对比本文算法在同一类天气退化图像恢复时采用不同示例图像的表现。以图像去雨任务为例, 示例图像分别随机选择测试数据集 Rain200H 中的三对图像。表 11 所示为本文算法在不同示例图像引导下的定量表现对比。如表所示, 不同的示例图像会对最终的指标造成微小的变化。具体而言, 在 PSNR 指标上, 三对示例图像引导的定量表现的变化范围在 0.04 dB 和 0.12 dB 范围内。同时, 实验将本文算法使用不同示例图像时的视觉对比在图 13 中展示。由图可知, 尽管示例图像不同, 但最终的恢复结果不论是在雨痕去除还是细节重建上均取得接近的表现。这表明, 本文算法可以根据示例图像进行情境学习, 即学习示例图像的潜在域间映射关系, 而非具体的退化图像与无退化图像的像素间映射关系。

表 11 本文算法使用不同示例图像时的定量表现

结果	PSNR	SSIM	LPIPS	NIQE
1	30.62	0.9136	0.0299	<b>3.757</b>
2	<b>30.70</b>	0.9131	<b>0.0287</b>	3.898
3	30.58	<b>0.9142</b>	0.0305	3.794

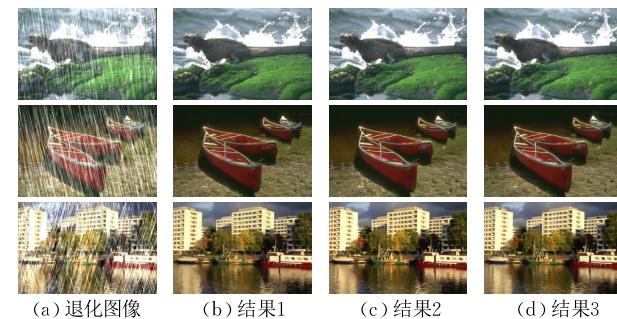


图 13 本文算法使用不同示例图像时的视觉对比

#### 4.5.3 多组示例图像

对于复杂的天气退化图像, 想要获取其对应的示例图像对往往是困难的。以夜间去雾任务为例, 想要得到去除雾气并提高亮度的图像则必须先去雾再增强或者先增强再去雾。然而, 这种串行执行的方式会成倍地增加推理时长。为验证本文算法在同时使用多种类型的示例图像时的表现, 实验选择夜间雾数据集 GTA5<sup>[43]</sup>。去雾示例图像为该数据集中的第一对图像, 增强示例图像为 LOL 数据集中的第一对图像。由于该数据集没有对应的无雾且增强的真实背景图像, 因此实验仅对其视觉对比进行展示。如图 14 所示, 结果 1 为先去雾后增强, 结果 2 为先增强后去雾, 结果 3 为同时去雾和增强。由图可知, 结果 1 和结果 3 均有效地去除图像中的雾气并增强图像的亮度。然而, 结果 2 尽管在图像去雾上有良好的表现, 但最终的结果仍然较暗。实验分析这种现象是由于雾气会导致局部较亮, 在执行增强时模型将其视为亮度良好的区域而不进行增强, 最后执行去雾就会导致所得结果的亮度不高。但是, 可以肯定的是, 同时存在的不同天气类型会互相影响, 使用对应的示例图像来引导模型可以得到更好的表现。然而, 在没有对应示例图像时可以使用多组示例图像来恢复图像。

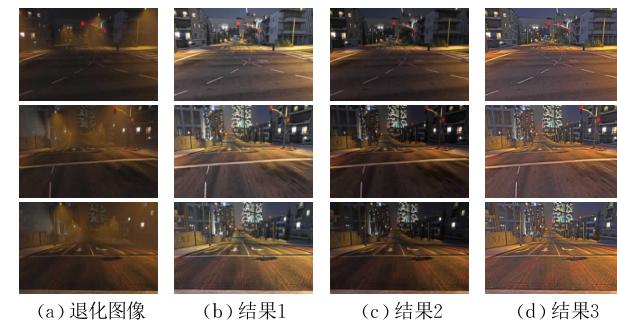


图 14 本文算法使用叠加示例图像时的视觉对比

#### 4.6 泛化实验

为验证本文算法在未知图像恢复任务上的泛化

能力,实验直接将训练好的模型用于七个未参与训练的天气退化类型,分别为交通场景图像去雨-雾、遥感图像去雾、真实图像去雨、水下图像增强、真实图像去雾、真实图像增强和真实图像去雨-雨滴.

#### 4.6.1 交通场景图像去雨-雾

如表 12 所示,未在交通场景雨-雾数据集上调优的本文算法在 RainCityscapes 数据集<sup>[44]</sup>上仍然取得良好的泛化表现.具体而言,在四个评价指标上本文算法取得的分数均优于所有对比方法.同时,实验将不同算法的视觉对比在图 15 中进行展示.由图可知,尽管两种方法所得的结果均存在细节丢失和雨痕残留的问题,但 Painter 所得结果中残留大量的雨痕,而本文算法能够得到更加接近真实背景的恢复图像.

表 12 不同算法在 RainCityscapes 数据集上的定量表现

类别	方法	PSNR	SSIM	LPIPS	NIQE
A	Restormer <sup>[35]</sup>	19.85	0.8795	0.1844	4.413
	DehazeFormer <sup>[36]</sup>	20.23	0.8680	0.2180	4.332
	SnowFormer <sup>[37]</sup>	20.00	0.8637	0.1683	4.518
	LLFormer <sup>[38]</sup>	18.25	0.8716	0.1704	4.693
	IDT <sup>[39]</sup>	21.08	0.9017	0.1384	3.936
	SASCFormer <sup>[31]</sup>	16.92	0.8147	0.2281	4.335
	DPENet <sup>[40]</sup>	16.94	0.8274	0.3606	4.345
	UDR-S <sup>2</sup> Former <sup>[41]</sup>	19.72	0.8856	0.1614	4.309
B	All-in-One <sup>[14]</sup>	20.37	0.8640	0.2348	4.695
	TransWeather <sup>[15]</sup>	19.26	0.8701	0.1893	4.358
	AIRFormer <sup>[16]</sup>	20.31	0.8528	0.2224	4.152
	TKL <sup>[42]</sup>	20.58	0.8969	0.1383	4.121
	AWIR-TDM <sup>[17]</sup>	21.07	0.8990	0.1265	4.019
C	WeatherDiff <sup>[12]</sup>	20.91	0.8784	0.1680	4.255
	Painter <sup>[8]</sup>	21.91	0.9080	0.1149	4.095
	本文算法	23.56	0.9277	0.0842	3.813

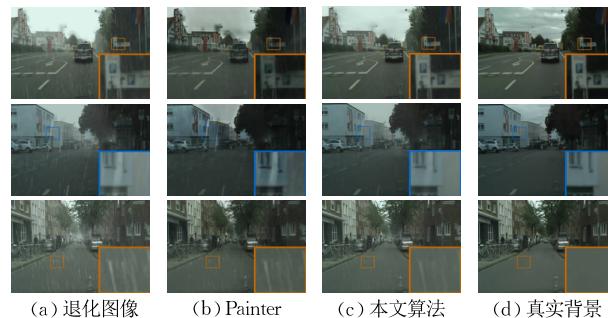


图 15 不同算法在 RainCityscapes 数据集上的视觉对比

#### 4.6.2 遥感图像去雾

实验在遥感图像去雾数据集 DHID<sup>[45]</sup> 上验证本文算法的泛化能力.如表 13 所示,本文算法在有监督指标 PSNR、SSIM 和 LPIPS 上分别相比 Painter 改善 2.09 dB、2.77% 和 18.31%.同时,本文算法在无监督评价指标 NIQE 上也明显低于 Painter,这说

明本文算法所得结果具有更好的自然度.此外,A 和 B 两类算法的指标均明显低于本文算法,因此这两类算法在遥感图像去雾任务上的泛化性能较差.由图 16 的视觉对比可知,Painter 所得结果中仍残留明显的雾,而本文算法在更好地去除雾的同时能够有效地保护恢复图像的颜色保真度.

表 13 不同算法在 DHID 数据集上的定量表现

类别	方法	PSNR	SSIM	LPIPS	NIQE
A	Restormer <sup>[35]</sup>	18.86	0.7974	0.2933	5.570
	DehazeFormer <sup>[36]</sup>	15.56	0.7466	0.3380	5.598
	SnowFormer <sup>[37]</sup>	14.78	0.7186	0.3767	5.790
	LLFormer <sup>[38]</sup>	11.48	0.6461	0.4533	6.092
	IDT <sup>[39]</sup>	15.49	0.7617	0.3339	4.922
	SASCFormer <sup>[31]</sup>	11.40	0.6274	0.4926	6.031
	DPENet <sup>[40]</sup>	11.26	0.6504	0.4363	5.850
	UDR-S <sup>2</sup> Former <sup>[41]</sup>	15.15	0.7570	0.3350	5.134
B	All-in-One <sup>[14]</sup>	11.79	0.6191	0.5278	6.794
	TransWeather <sup>[15]</sup>	15.27	0.7587	0.3080	5.202
	AIRFormer <sup>[16]</sup>	16.08	0.7565	0.3364	4.809
	TKL <sup>[42]</sup>	13.77	0.7194	0.3744	5.673
	AWIR-TDM <sup>[17]</sup>	17.10	0.7591	0.3195	5.088
C	WeatherDiff <sup>[12]</sup>	11.93	0.6685	0.4241	5.816
	Painter <sup>[8]</sup>	19.33	0.8145	0.2600	5.346
	本文算法	21.42	0.8371	0.2124	4.611

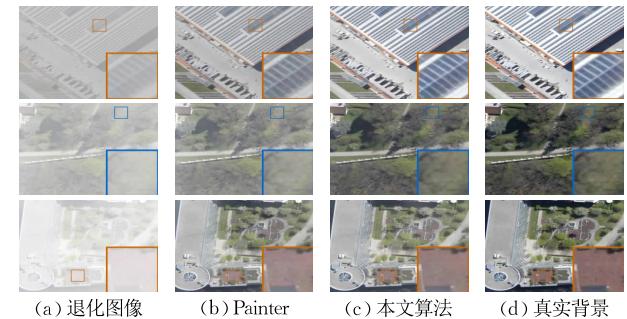


图 16 不同算法在 DHID 数据集上的视觉对比

#### 4.6.3 真实图像去雨

真实雨图中的雨痕分布往往相比合成雨图更加复杂,这导致现有的在合成雨图上表现良好的算法无法有效去除真实雨图中的退化.因此,实验还在真实雨退化数据集 RealRain1K-H<sup>[46]</sup> 上验证本文算法的泛化能力.如表 14 所示,本文算法在 PSNR 和 SSIM 指标上分别优于 Painter 算法 1.73 dB 和 3.65%.同时,本文算法的 LPIPS 和 NIQE 指标也优于现有算法 Painter.此外,本文算法在四个评价指标上均明显优于现有的 A 和 B 两类算法.由图 17 中的视觉对比可知,相比现有算法 Painter,本文算法所得结果中的雨痕明显减少,且图像细节更接近真实背景.

表 14 不同算法在 RealRain1K-H 数据集上的定量表现

类别	方法	PSNR	SSIM	LPIPS	NIQE
A	Restormer <sup>[35]</sup>	25.00	0.7760	0.3881	9.311
	DehazeFormer <sup>[36]</sup>	21.64	0.7561	0.5576	6.807
	SnowFormer <sup>[37]</sup>	20.42	0.7115	0.5817	6.839
	LLFormer <sup>[38]</sup>	20.85	0.7382	0.6033	6.364
	IDT <sup>[39]</sup>	19.30	0.7386	0.5870	6.082
	SASCFormer <sup>[31]</sup>	20.35	0.7259	0.6230	6.297
	DPENet <sup>[40]</sup>	22.44	0.6775	0.5402	10.78
	UDR-S <sup>2</sup> Former <sup>[41]</sup>	24.12	0.7773	0.4171	9.005
B	All-in-One <sup>[14]</sup>	23.07	0.7509	0.4769	9.372
	TransWeather <sup>[15]</sup>	24.11	0.7516	0.4402	9.342
	AIRFormer <sup>[16]</sup>	24.09	0.7527	0.4706	9.420
	TKL <sup>[42]</sup>	22.39	0.7676	0.4065	9.283
C	AWIR-TDM <sup>[17]</sup>	22.09	0.7962	0.4138	7.977
	WeatherDiff <sup>[12]</sup>	22.77	0.7895	0.3788	8.215
	Painter <sup>[8]</sup>	25.78	0.7949	0.3966	9.982
	本文算法	<b>27.51</b>	<b>0.8239</b>	<b>0.3611</b>	<b>5.623</b>

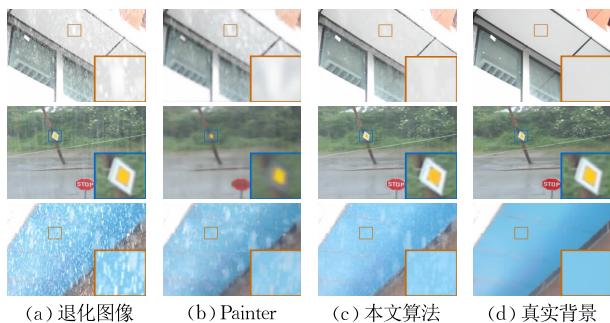


图 17 不同算法在 RealRain1K-H 数据集上的视觉对比

#### 4.6.4 水下图像增强

水下图像增强对于水下计算机视觉任务具有重要意义。因此,实验还验证本文算法在水下图像增强任务上的泛化能力。表 15 所示为不同算法在 LSUI 数据集<sup>[47]</sup>上的定量表现。由表可知,Painter 算法在该数据集上的泛化表现较差,具体在 PSNR 指标上低于本文算法 2.94 dB。同时,实验还将两种方法的视觉对比在图 18 中进行展示。可以发现,Painter 在

表 15 不同算法在 LSUI 数据集上的定量表现

类别	方法	PSNR	SSIM	LPIPS	NIQE
A	Restormer <sup>[35]</sup>	23.76	0.8585	0.3840	4.999
	DehazeFormer <sup>[36]</sup>	23.57	0.8541	0.3907	4.948
	SnowFormer <sup>[37]</sup>	23.61	0.8588	0.3817	4.959
	LLFormer <sup>[38]</sup>	23.60	0.8496	0.3973	5.116
	IDT <sup>[39]</sup>	23.50	0.8499	0.3963	4.959
	SASCFormer <sup>[31]</sup>	23.52	0.8448	0.4007	5.005
	DPENet <sup>[40]</sup>	23.61	0.8547	0.3788	4.906
	UDR-S <sup>2</sup> Former <sup>[41]</sup>	23.45	0.8538	0.3885	4.930
B	All-in-One <sup>[14]</sup>	22.10	0.7850	0.4496	5.548
	TransWeather <sup>[15]</sup>	23.76	0.8520	0.3892	4.989
	AIRFormer <sup>[16]</sup>	23.38	0.8205	0.4331	5.309
	TKL <sup>[42]</sup>	23.46	0.8520	0.3979	4.998
	AWIR-TDM <sup>[17]</sup>	22.42	0.7963	0.4279	4.381
C	WeatherDiff <sup>[12]</sup>	23.33	0.8497	0.3870	4.896
	Painter <sup>[8]</sup>	22.89	0.7984	0.5175	6.417
	本文算法	<b>25.83</b>	<b>0.8877</b>	<b>0.3490</b>	<b>4.536</b>

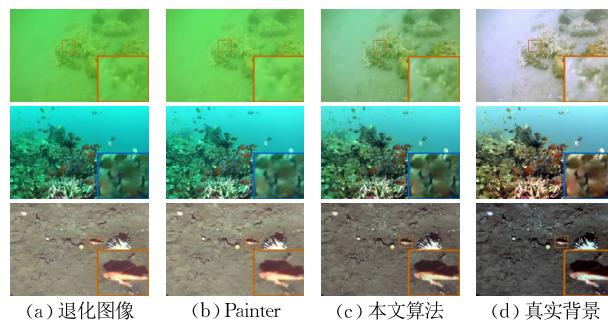


图 18 不同算法在 LSUI 数据集上的视觉对比

水下图像增强任务上几乎没有泛化表现。不同的是,尽管本文算法所得的结果仍然与真实背景存在较大差距,但相比 Painter 所得结果明显更接近真实背景。

#### 4.6.5 真实图像去雾

为验证本文算法在真实世界的雾天退化图像上的泛化表现,实验选择真实的密集雾退化数据集 Dense-Haze<sup>[48]</sup>,其是真实图像去雾领域的一个具有挑战性的数据集。由表 16 可知,尽管本文算法所得的指标分数均优于所有的对比方法,但本文的指标分数仍然较低。同时,A 和 B 两类算法在 Dense-Haze 数据集上的泛化表现较差。如图 19 所示,Painter 和本

表 16 不同算法在 Dense-Haze 数据集上的定量表现

类别	方法	PSNR	SSIM	LPIPS	NIQE
A	Restormer <sup>[35]</sup>	12.45	0.5162	0.7836	7.637
	DehazeFormer <sup>[36]</sup>	13.03	0.5277	0.7703	6.776
	SnowFormer <sup>[37]</sup>	13.15	0.5321	0.7536	6.689
	LLFormer <sup>[38]</sup>	12.63	0.5207	0.7535	7.692
	IDT <sup>[39]</sup>	12.05	0.5007	0.7803	6.831
	SASCFormer <sup>[31]</sup>	11.64	0.4870	0.7675	6.106
	DPENet <sup>[40]</sup>	11.48	0.4971	0.7852	7.277
	UDR-S <sup>2</sup> Former <sup>[41]</sup>	12.33	0.5169	0.7697	6.994
B	All-in-One <sup>[14]</sup>	12.06	0.5082	0.8245	8.131
	TransWeather <sup>[15]</sup>	12.03	0.5135	0.7850	7.452
	AIRFormer <sup>[16]</sup>	12.21	0.4989	0.8155	8.083
	TKL <sup>[42]</sup>	11.39	0.4916	0.7839	7.091
	AWIR-TDM <sup>[17]</sup>	11.94	0.5021	0.7836	6.223
C	WeatherDiff <sup>[12]</sup>	11.86	0.4927	0.7828	6.413
	Painter <sup>[8]</sup>	14.04	0.5107	0.8097	6.985
	本文算法	<b>15.01</b>	<b>0.5504</b>	<b>0.7356</b>	<b>5.698</b>



图 19 不同算法在 Dense-Haze 数据集上的视觉对比

文算法均具有一定的泛化能力,但无法彻底地去除退化图像中的密集雾。对比而言,本文算法所得的视觉表现更好。

#### 4.6.6 真实图像增强

实验选择真实世界的图像增强数据集 LSRW-Huawei<sup>[49]</sup>来验证本文算法在真实图像增强任务上的泛化表现。由表 17 可知,在不针对 LSRW-Huawei 数据集进行训练的前提下,本文算法能够取得明显优于其他所有对比方法的表现。具体而言,相比现有方法 Painter,本文算法在 PSNR 指标上提升 1.56 dB,在 SSIM 指标上提升 1.93%,在 LPIPS 指标上改善 8.86%,同时取得更好的图像自然度分数。此外,A 类算法在该数据集上基本不具备泛化能力,而 B 类算法中的 TransWeather 和 AIRFormer 具有一定的泛化能力,但是仍然弱于本文算法。本文算法与现有算法 Painter 的视觉对比如图 20 所示。如图所示,Painter 所得结果的亮度较低,而本文算法所得结果的亮度较高且细节纹理更接近真实背景。

表 17 不同算法在 LSRW-Huawei 数据集上的定量表现

类别	方法	PSNR	SSIM	LPIPS	NIQE
A	Restormer <sup>[35]</sup>	16.24	0.5575	0.4459	5.428
	DehazeFormer <sup>[36]</sup>	15.11	0.5801	0.4912	5.843
	SnowFormer <sup>[37]</sup>	19.61	0.6267	0.4202	4.312
	LLFormer <sup>[38]</sup>	11.47	0.4432	0.5794	6.710
	IDT <sup>[39]</sup>	15.63	0.5565	0.5479	4.337
	SASCFormer <sup>[31]</sup>	10.17	0.3671	0.7173	6.366
	DPENet <sup>[40]</sup>	13.62	0.5215	0.5357	5.217
B	UDR-S <sup>2</sup> Former <sup>[41]</sup>	16.06	0.5783	0.5123	4.813
	All-in-One <sup>[14]</sup>	10.27	0.3765	0.7101	8.873
	TransWeather <sup>[15]</sup>	21.29	0.6606	0.3671	4.634
	AIRFormer <sup>[16]</sup>	20.66	0.6535	0.4706	4.809
	TKL <sup>[42]</sup>	18.53	0.6290	0.4262	4.539
	AWIR-TDM <sup>[17]</sup>	19.37	0.6198	0.4300	3.764
	WeatherDiff <sup>[12]</sup>	14.18	0.4797	0.5894	3.264
C	Painter <sup>[8]</sup>	23.05	0.6793	0.4280	4.266
	本文算法	<b>24.61</b>	<b>0.6924</b>	<b>0.3901</b>	<b>3.017</b>

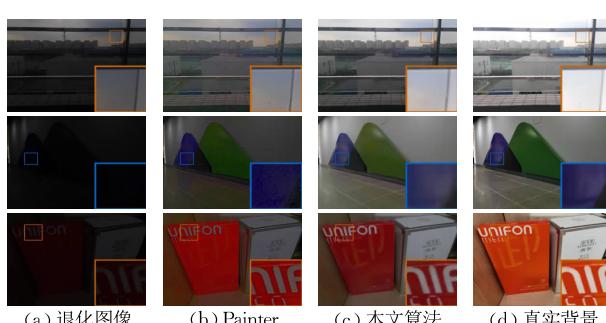


图 20 不同算法在 LSRW-Huawei 数据集上的视觉对比

#### 4.6.7 真实图像去雨-雨滴

实验选择真实世界的存在雨-雨滴混合天气退

化的 RainDS-real<sup>[31]</sup> 数据集来验证本文算法在真实图像去雨-雨滴任务上的表现。如表 18 所示,在不针对 RainDS-real 数据集进行训练的前提下,本文算法所得的指标分数明显优于其他所有对比方法。特别是,相比具有类似设定的现有算法 Painter,本文算法在 PSNR 指标上提升 2.22 dB,在 SSIM 上提升 12.16%,在 LPIPS 指标上改善 46.25%。同时,本文算法所得结果的 NIQE 指标最好。为进一步体现本文算法相比 Painter 算法的优势,实验将视觉对比在图 21 中进行展示。如图所示,两种方法均无法完全去除图像中的雨-雨滴退化,但 Painter 所得结果中存在明显的雨滴和雨痕退化,而本文算法所得结果中残留的退化较少。因此,本文算法在真实图像去雨-雨滴任务上取得更好的泛化表现。

表 18 不同算法在 RainDS-real 数据集上的定量表现

类别	方法	PSNR	SSIM	LPIPS	NIQE
A	Restormer <sup>[35]</sup>	20.77	0.6033	0.3717	3.701
	DehazeFormer <sup>[36]</sup>	21.28	0.6011	0.3771	4.857
	SnowFormer <sup>[37]</sup>	20.16	0.5939	0.3677	4.552
	LLFormer <sup>[38]</sup>	21.72	0.6005	<u>0.3414</u>	4.875
	IDT <sup>[39]</sup>	21.72	<u>0.6071</u>	0.3466	3.802
	SASCFormer <sup>[31]</sup>	20.26	0.5828	0.3656	3.509
	DPENet <sup>[40]</sup>	19.22	0.5682	0.3794	5.318
B	UDR-S <sup>2</sup> Former <sup>[41]</sup>	21.01	0.5998	0.3359	4.702
	All-in-One <sup>[14]</sup>	19.58	0.5712	0.3830	5.426
	TransWeather <sup>[15]</sup>	20.67	0.5786	0.3460	4.449
	AIRFormer <sup>[16]</sup>	20.49	0.5779	0.4090	4.114
	TKL <sup>[42]</sup>	20.14	0.5961	0.3886	3.464
	AWIR-TDM <sup>[17]</sup>	21.45	0.6153	0.3400	3.598
	WeatherDiff <sup>[12]</sup>	20.10	0.5480	0.3753	3.572
C	Painter <sup>[8]</sup>	<u>22.65</u>	0.6013	0.5873	6.047
	本文算法	<b>24.87</b>	<b>0.6744</b>	<b>0.3157</b>	<b>3.211</b>

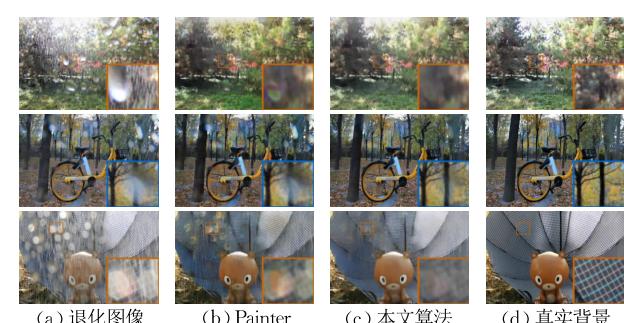


图 21 不同算法在 RainDS-real 数据集上的视觉对比

## 5 局限性

尽管本文算法在多个不同的天气退化图像恢复任务中均取得良好的表现,但其仍然存在两点局限性。如表 19 所示,本文算法在处理分辨率较大的退

化图像时的计算负担较大。同时,本文算法处理分辨率大小为  $512 \times 512$  的退化图像所需时间为 3.04 s,但处理分辨率大小为  $1024 \times 1024$  的退化图像所需时间增加到 14.76 s。因此,本文算法的实时性较差。另外,实验采用的训练数据集仅包括 8 种天气退化,从而导致其无法适应各种真实场景下的复杂天气退化。

**表 19 本文算法处理  $512 \times 512$  大小的图片时的效率分析**

参数数量/M	计算量/G	推理时间/s
9812.52	1161.26	3.04

## 6 结论与展望

本文提出一种有效的视觉提示学习方法,其能够应用于包括图像去雨、去雾、去雪、增强和去雨滴等在内的 15 个天气退化图像恢复任务。具体而言,本文提出查询提示约束网络,其直接根据退化图像得到与真实背景图像对应的潜在特征。同时,本文提出示例提示引导网络,其利用示例图像和查询图像来引导预训练稳定扩散模型去除对应的天气退化。本文算法是预训练语言视觉模型和天气退化图像恢复任务结合的新范式。未来本文将继续探索加速稳定扩散模型反向采样的方法,同时本文将进一步构建覆盖各种复杂场景的统一天气退化图像数据集。

## 参 考 文 献

- [1] Huang Shu-Ying, Hu Wei, Yang Yong, et al. A low-exposure image enhancement based on progressive dual network model. Chinese Journal of Computers, 2021, 44(2): 384-394 (in Chinese)  
(黄淑英, 胡威, 杨勇等. 基于渐进式双网络模型的低曝光图像增强方法. 计算机学报, 2021, 44(2): 384-394)
- [2] Xiao Jin-Sheng, Wang Wen, Zou Wen-Tao, et al. An image rain removal algorithm via depth of field and sparse coding. Chinese Journal of Computers, 2019, 42(9): 2024-2034 (in Chinese)  
(肖进胜, 王文, 邹文涛等. 基于景深和稀疏编码的图像去雨算法. 计算机学报, 2019, 42(9): 2024-2034)
- [3] Wen Yuan-Bo, Gao Tao, Chen Ting, et al. Frequency-guided dual sparse self-attention algorithm for single image deraining. Acta Electronica Sinica, 2023, 51(10): 2812-2820 (in Chinese)  
(文渊博, 高涛, 陈婷等. 频率引导的双稀疏自注意力单图像去雨算法. 电子学报, 2023, 51(10): 2812-2820)
- [4] Fu X, Huang J, Ding X, et al. Clearing the skies: A deep network architecture for single-image rain removal. IEEE Transactions on Image Processing, 2017, 26(6): 2944-2956
- [5] Zamir S W, Arora A, Khan S, et al. Multi-stage progressive image restoration//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Kuala Lumpur, Malaysia, 2021: 14821-14831
- [6] Chen X, Li H, Li M, et al. Learning a sparse transformer network for effective image deraining//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada, 2023: 5896-5905
- [7] Bar A, Gandelsman Y, Darrell T, et al. Visual prompting via image inpainting. Advances in Neural Information Processing Systems, 2022, 35: 25005-25017
- [8] Wang X, Wang W, Cao Y, et al. Images speak in images: A generalist painter for in-context visual learning//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada, 2023: 6830-6839
- [9] Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. Advances in Neural Information Processing Systems, 2020, 33: 6840-6851
- [10] Rombach R, Blattmann A, Lorenz D, et al. High-resolution image synthesis with latent diffusion models//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA, 2022: 10684-10695
- [11] Zhang L, Rao A, Agrawala M. Adding conditional control to text-to-image diffusion models//Proceedings of the IEEE/CVF International Conference on Computer Vision. Paris, France, 2023: 3836-3847
- [12] Özdenizci O, Legenstein R. Restoring vision in adverse weather conditions with patch-based denoising diffusion models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(8): 10346-10357
- [13] Radford A, Kim J W, Hallacy C, et al. Learning transferable visual models from natural language supervision//Proceedings of the International Conference on Machine Learning. Vienna, Austria, 2021: 8748-8763
- [14] Li R, Tan R T, Cheong L F. All in one bad weather removal using architectural search//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA, 2020: 3175-3185
- [15] Valanarasu J M J, Yasarla R, Patel V M. Transweather: Transformer-based restoration of images degraded by adverse weather conditions//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA, 2022: 2353-2363
- [16] Gao T, Wen Y, Zhang K, et al. Frequency-oriented efficient transformer for all-in-one weather-degraded image restoration. IEEE Transactions on Circuits and Systems for Video Technology, 2024, 34(3): 1886-1899
- [17] Qin Jing, Wen Yuan-Bo, Gao Tao, et al. Transformer-based diffusion model for all-in-one weather-degraded image restoration. <https://doi.org/10.16183/j.cnki.jsjtu.2023.043> (in Chinese)  
(秦菁, 文渊博, 高涛等. 面向多天气退化图像恢复的自注意力扩散模型. <https://doi.org/10.16183/j.cnki.jsjtu.2023.043>)

- [18] Brown T, Mann B, Ryder N, et al. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 2020, 33: 1877-1901
- [19] Alayrac J B, Donahue J, Luc P, et al. Flamingo: A visual language model for few-shot learning. *Advances in Neural Information Processing Systems*, 2022, 35: 23716-23736
- [20] Liang Z, Li C, Zhou S, et al. Iterative prompt learning for unsupervised backlit image enhancement//Proceedings of the IEEE/CVF International Conference on Computer Vision. Paris, France, 2023: 8094-8103
- [21] Li J, Li D, Xiong C, et al. BLIP: Bootstrapping language-image pre-training for unified vision-language understanding and generation//Proceedings of the International Conference on Machine Learning. Maryland, USA, 2022: 12888-12900
- [22] Luo Z, Gustafsson F K, Zhao Z, et al. Controlling vision-language models for universal image restoration//Proceedings of the International Conference on Learning Representations. Vienna, Austria, 2024: 1-21
- [23] Loshchilov I, Hutter F. Decoupled weight decay regularization //Proceedings of the International Conference on Learning Representations. New Orleans, USA, 2019: 1-18
- [24] Yang W, Tan R T, Feng J, et al. Deep joint rain detection and removal from a single image//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Venice, Italy, 2017: 1357-1366
- [25] Liu Y F, Jaw D W, Huang S C, et al. DesnowNet: Context-aware deep network for snow removal. *IEEE Transactions on Image Processing*, 2018, 27(6): 3064-3073
- [26] Li B, Ren W, Fu D, et al. Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing*, 2018, 28(1): 492-505
- [27] Wei C, Wang W, Yang W, et al. Deep retinex decomposition for low-light enhancement//Proceedings of the British Machine Vision Conference. Newcastle, UK, 2018: 1-12
- [28] Qian R, Tan R T, Yang W, et al. Attentive generative adversarial network for raindrop removal from a single image//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salty Lake City, USA, 2018: 2482-2491
- [29] Wen Y, Gao T, Zhang K, et al. Restoring vision in rain-by-snow weather with simple attention-based sampling cross-hierarchy Transformer. *Pattern Recognition*, 2024, 156: 110743
- [30] Li R, Cheong L F, Tan R T. Heavy rain image restoration: Integrating physics model and conditional adversarial learning //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA, 2019: 1633-1642
- [31] Quan R, Yu X, Liang Y, et al. Removing raindrops and rain streaks in one go//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Kuala Lumpur, Malaysia, 2021: 9147-9156
- [32] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612
- [33] Zhang R, Isola P, Efros A A, et al. The unreasonable effectiveness of deep features as a perceptual metric//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salty Lake City, USA, 2018: 586-595
- [34] Mital A, Soundararajan R, Bovik A C. Making a “completely blind” image quality analyzer. *IEEE Signal Processing Letters*, 2012, 20(3): 209-212
- [35] Zamir S W, Arora A, Khan S, et al. Restormer: Efficient transformer for high-resolution image restoration//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA, 2022: 5728-5739
- [36] Song Y, He Z, Qian H, et al. Vision transformers for single image dehazing. *IEEE Transactions on Image Processing*, 2023, 32: 1927-1941
- [37] Chen S, Ye T, Liu Y, et al. Snowformer: Scale-aware transformer via context interaction for single image desnowing. <https://arxiv.org/abs/2208.09703>
- [38] Wang T, Zhang K, Shen T, et al. Ultra-high-definition low-light image enhancement: A benchmark and transformer-based method//Proceedings of the AAAI Conference on Artificial Intelligence. Washington, USA, 2023, 37(3): 2654-2662
- [39] Xiao J, Fu X, Liu A, et al. Image de-raining transformer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 45: 12978-12995
- [40] Wen Y, Gao T, Zhang J, et al. From heavy rain removal to detail restoration: A faster and better network. *Pattern Recognition*, 2024, 148: 110205
- [41] Chen S, Ye T, Bai J, et al. Sparse sampling transformer with uncertainty-driven ranking for unified removal of raindrops and rain streaks//Proceedings of the IEEE/CVF International Conference on Computer Vision. Paris, France, 2023: 13106-13117
- [42] Chen W T, Huang Z K, Tsai C C, et al. Learning multiple adverse weather removal via two-stage knowledge learning and multi-contrastive regularization: Toward a unified model //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA, 2022: 17653-17662
- [43] Yan W, Tan R T, Dai D. Nighttime defogging using high-low frequency decomposition and grayscale-color networks//Proceedings of the European Conference on Computer Vision. Glasgow, UK, 2020: 473-488
- [44] Hu X, Fu C W, Zhu L, et al. Depth-attentional features for single-image rain removal//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA, 2019: 8022-8031
- [45] Zhang L, Wang S. Dense haze removal based on dynamic collaborative inference learning for remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-16

- [46] Li W, Zhang Q, Zhang J, et al. Toward real-world single image deraining: A new benchmark and beyond. arXiv preprint arXiv:2206.05514, 2022
- [47] Peng L, Zhu C, Bian L. U-shape transformer for underwater image enhancement//Proceedings of the European Conference on Computer Vision. Tel Aviv, Israel, 2022: 290-307
- [48] Ancuti C O, Ancuti C, Sbert M, et al. Dense-Haze: A benchmark for image dehazing with dense-haze and haze-free images//Proceedings of the IEEE International Conference on Image Processing. Taipei, China, 2019: 1014-1018
- [49] Hai J, Xuan Z, Yang R, et al. R2RNet: Low-light image enhancement via real-low to real-normal network. Journal of Visual Communication and Image Representation, 2023, 90: 103712



**WEN Yuan-Bo**, Ph.D. candidate.

His current research interests include computer vision and intelligent transportation systems.

**GAO Tao**, Ph.D., professor. His current research interests include artificial intelligence and image processing.

## Background

This work belongs to the weather-degraded image restoration in the field of computer vision. Currently, existing methods can be divided into the task-specific and task-aligned. The former explores special physical models for different weather degradations and designs corresponding weather removal models, while the latter designs aligned networks for different weather degradations and realizes weather removal through separately training. However, these methods can not be applied to the complex and changeable real-world weather removal tasks. For instance, there may be different degradations in the real-world weather-degraded images, such as rain, haze and snow. With the development of prompt learning, several algorithms based on visual prompt learning have gradually emerged, but these methods have weak adaptability and limited representation ability, leading to degradation residue and detail destruction.

To this end, this paper introduces the pre-trained language-image model into weather-degraded image restoration. Specifically, this paper proposes a query prompt constrained network (QPC-Net), which utilizes the image encoders from contrastive language-image pre-training model to directly generate the latent degradation-free features. These features are consistent with the latent caption features generated by the corresponding

**AN Yi-Sheng**, Ph.D., professor. His current research interests include discrete event system modeling and application and intelligent vehicle infrastructure cooperative systems.

**LI Zi-Qi**, Ph.D. candidate. His current research interests include computer vision and traffic target detection.

**CHEN Ting**, Ph.D., professor. Her current research interests include pattern recognition and intelligent transportation systems.

degradation-free image captions. This approach avoids the interference of weather degradations during inference, while facilitating the reconstruction of image contents. Meanwhile, this paper develops an example prompt guided network (EPG-Net), which employs the example images to conduct in-context learning, thereby learning the underlying domain mapping relationships between the weather-degraded images and corresponding ground truth, rather than simple pixel-wise mapping relationships. Our method utilizes the proposed QPC-Net and EPG-Net to guide the pre-trained stable diffusion model to process the weather-degraded images and generate the expectant restored images. Extensive experimental results demonstrate that our method improves the peak signal-to-noise ratio by 2.99 dB on eight publicly available weather-degraded datasets. This work is a novel paradigm that effectively utilizes the pre-trained language-image models to eliminate weather degradations.

This research is partially supported by the National Key R&D Program of China (No. 2023YFB2504703), the Shaanxi International S&T Cooperation Program Project (No. 2024GH-YBXM-24), the National Natural Science Foundation of China (No. 52172379) and the Fundamental Research Funds for the Central Universities (No. 300102242901).