

中图法分类号: TP18; TP391 文献标识码: A 文章编号: 1006-8961(2026)02-0465-14

论文引用格式: Gong J X, Xu J D and Sun H Q. 2026. Dual-channel blind image separation based on a wavelet suppression interactive diffusion model. Journal of Image and Graphics, 31(2):0465-0478(龚嘉鑫, 徐金东, 孙浩钦. 2026. 面向双通道盲图像分离的小波抑制交互扩散模型. 中国图象图形学报, 31(2):0465-0478)[DOI:10.11834/jig.250230]

面向双通道盲图像分离的小波抑制交互扩散模型

龚嘉鑫, 徐金东*, 孙浩钦

烟台大学计算机与控制工程学院, 烟台 264005

摘要: 目的 盲图像分离(blind image separation, BIS)指在未知混合方式及源图像先验的条件下,从单幅观测图像中同时估计并恢复多个独立源图像的逆问题,典型应用包括恶劣天气(雨、雪)去除和反射/阴影层分离等。然而,现有的BIS方法普遍依赖人工先验或基于卷积神经网络(convolutional neural network, CNN)或对抗生成网络(generative adversarial network, GAN)的变体,难以刻画真实场景中源图像复杂且多变的特征分布,导致在强噪声、非线性混合及纹理细节高度耦合的情况下出现源图像分离估计偏差、纹理失真和伪影残留等缺陷。对此,本文创新性地将扩散模型引入双通道盲图像分离,提出一种高效的双通道扩散分离模型(dual channel diffusion separation model, DCDSM)。方法 DCDSM利用扩散模型学习源图像的特征分布,实现对源图像特征结构的有效重建,并根据两个源图像互为耦合噪声的特性,在双分支扩散的反向去噪过程设计了小波抑制模块(wavelet suppression module, WSM),形成交互式的双分支分离网络,提高混合图像细节信息的分离效果。结果 在含雨/雪及复杂混合场景的合成数据集上进行了充分实验,结果表明:1)DCDSM在去雨、去雪的图像复原任务中分别取得PSNR/SSIM指标值为35.002 3 dB/0.954 9和29.810 8 dB/0.924 3,较Histoformer和LDRCNet分别提升1.257 0 dB/0.927 2 dB(PSNR)和0.026 2/0.028 9(SSIM)。2)针对复杂混合图像的双盲分离任务,所恢复的双源图像在纹理保真度与细节完整性上均显著优于其他方法,其PSNR与SSIM指标在两个通道上的平均值分别达到25.004 9 dB和0.799 7,较对比方法提升4.124 9 dB和0.092 6。结论 本文模型在主观视觉和客观指标上相较于对比方法均取得最优的结果,解决了存在雨纹和雪纹残留、混合图像中细节模糊等问题。

关键词: 盲图像分离(BIS); 扩散模型(DM); 小波变换; 傅里叶变换; 图像复原

Dual-channel blind image separation based on a wavelet suppression interactive diffusion model

Gong Jiabin, Xu Jindong*, Sun Haoqin

School of Computer and Control Engineering, Yantai University, Yantai 264005, China

Abstract: Objective Blind image separation (BIS) is the application of blind source separation in the domain of image processing. It refers to the inverse problem of simultaneously estimating and restoring multiple independent source images from a single observation image under conditions of unknown mixing mode and without prior knowledge of the source

收稿日期: 2025-05-28; 修回日期: 2025-09-02; 预印本日期: 2025-09-09

* 通信作者: 徐金东 xujindong@ytu.edu.cn

基金项目: 国家自然科学基金项目(62072391); 烟台市科技创新发展计划基础研究重点项目(2024JCYJ037); 山东省研究生教育教改项目(SDYJSJGC2023075)

Supported by: National Natural Science Foundation of China(62072391); Basic Research Key Project of Yantai Science and Technology Innovation Development Plan(2024JCYJ037); Postgraduate Education Reform Project of Shandong Province(SDYJSJGC2023075)

images. Typical applications include bad weather (rain, snow) removal and reflection/shadow layer separation. Traditional methods, represented by independent component analysis and its improved algorithms, rely on strong assumptions such as “statistical independence of source signals” and “non-Gaussianity,” which often fail for real images owing to strong feature correlation and nonlinear mixing. Although sparse coding and low-rank decomposition, which emerged later, introduced prior constraints, they remained difficult for characterizing complex textures and structures. In recent years, deep learning frameworks have gradually dominated: convolutional neural networks (CNNs) achieved efficient feature extraction with local receptive fields. Generative adversarial networks (GANs) circumvented explicit priors through adversarial training, and variants such as CycleGAN, Attention-GAN, and Transformer-GAN have been proposed successively, achieving significant progress on synthetic datasets. However, in complex real scenes, existing methods based on CNNs and GANs still exhibit insufficient performance in processing such mixed images. The reasons are as follows: 1) uncertainty in source feature distribution (features within the same source category vary significantly in shape, transparency, and scale, resulting in modeling difficulties); 2) complex image mixing (nonlinearity, crosstalk between channels, making the “reverse mapping” nonunique); and 3) irregular noise interference (sensor noise, compression artifacts, motion blur coupled with the source signal, further blurring the separable boundary). These factors make it difficult for models to characterize the complex and variable feature distribution of source images in real scenes. Consequently, under strong noise, nonlinear mixing, and highly coupled texture details, problems, including source image separation estimation bias, texture distortion, and artifact residue, arise. These problems negatively impact the effectiveness of image restoration. To address these challenges, this study proposes a novel dual-channel diffusion separation model (DCDSM). It leverages the powerful generative ability of diffusion models to handle complex mixed images effectively. **Method** DCDSM consists of a dual-branch structure based on a conditional diffusion model. This model exploits the diffusion process to learn the feature distribution of the source images, enabling the reconstruction of the feature structure for initial separation. During the reverse denoising process of the dual-branch diffusion, the design of the wavelet suppression module (WSM) is grounded in the characteristic of mutual coupling noise between the two source images. The structure of the interactive dual-branch separation network enhances the separation of detailed information within mixed images. WSM is composed of two independent wavelet frequency-domain feature extraction networks (WFENs). A WFEN employs a two-dimensional discrete wavelet transform to process high- and low-frequency sub-bands in the time-frequency domain to obtain suppression information. In the time domain, an encoder-decoder structure is utilized to capture global contextual features and reconstruct local details, thereby enhancing the texture and edge information in the low-frequency sub-bands. In the frequency domain, a window-based frequency channel attention mechanism is introduced to process the high-frequency sub-bands. Finally, a two-dimensional wavelet inverse transform is employed to integrate the outputs from both branches to obtain the suppression information. Simultaneously, the noise output from the intermediate process of the other branch is subtracted pixel-wise from the suppression information to decouple the two source images, thereby further improving the model’s performance in image separation tasks. **Result** DCDSM is validated through the construction of synthetic datasets from diverse application scenarios, encompassing rain removal, snow removal, and the simulation of complex mixtures. The experimental results are as follows: 1) in the tasks of rain and snow image restoration, DCDSM achieves quantitative metrics (PSNR/SSIM) of 35.002 3 dB/0.954 9 and 29.810 8 dB/0.924 3, respectively, demonstrating an average improvement of 1.257 0 dB/0.927 2 dB (PSNR) and 0.026 2/0.028 9 (SSIM) over the current state-of-the-art methods. 2) For the dual-blind separation of complex mixed images, the restored dual-source images exhibit significantly better texture fidelity and detail integrity compared with those from other methods, with PSNR and SSIM metrics reaching 25.004 9 dB and 0.799 7, respectively, surpassing those of the comparative methods by 4.124 9 dB and 0.092 6, respectively. Ablation experiments validate the effectiveness of the proposed modules and the interpretability and rationality of the selected hyperparameters. **Conclusion** The experiments of DCDSM on rain, snow, and complex mixed datasets verify the effectiveness of the method in the task of dual-channel BIS. Experimental results show that the proposed method achieves the best subjective and objective indicators compared with other methods and solves the problems of residual rain/snow lines and blurred texture edge details in real complex separation scenes.

Key words: blind image separation(BIS); diffusion model(DM); wavelet transform; Fourier transform; image restoration

0 引言

盲源分离(blind source separation, BSS)旨在通过少量的源和混合过程先验信息,将源信号从混合信号中分离出来,已广泛应用于语音增强、复原与识别、信号降噪等众多领域(徐金东等,2020)。其在图像处理领域中的应用称为盲图像分离(blind image separation, BIS)。在计算机视觉中,获取图像受光线、环境噪声等因素影响,易导致图像混叠、噪声污染等问题,所以如何在缺乏源图像特征分布和混合参数的条件下,从复杂的混叠图像中分离出源图像极具挑战性,对后续的图像分析与识别也具有重要意义。

以独立成分分析(independent component analysis, ICA)(Hyvärinen 和 Oja, 2000)为代表的统计约束 BSS 算法,假设源信号相互独立且呈非高斯分布,而实际采样的源图像之间往往存在较高程度的特征相似性及分布相关性。因此,基于统计约束的盲图像分类算法难以满足复杂图像分离的实际需求。近年来,深度学习(deep learning, DL)在计算机视觉领域大放异彩,众多学者将其引入到图像分离任务。Hoshen(2019)提出一种无监督单通道盲源分离方法,将循环生成对抗网络(cycle-consistent generative adversarial network, CycleGAN)引入盲图像分离,旨在为缺乏足够的先验信息条件下实现分离提供思路。但该方法仅考虑混叠图像与源图像之间的映射,忽略了特征的空间位置关系。对此, Sun 等人(2022)将自注意力机制引入生成对抗网络(generative adversarial network, GAN),利用 GAN 避免对源图像的约束,降低对样本的需求,利用自注意力机制捕获全局依赖以实现高质量的图像分离。为了提高分离精度并保留图像的细节表示, Su 等人(2023)设计一种基于 Transformer 的 GAN。该网络利用 Transformer 在细节处理方面的能力与 GAN 在全局上下文特征提取方面的优势,有效防止了信息丢失,实现局部与整体之间的良好平衡。Liu 等人(2022)设计了一个并行双生对抗网络,建立混合信号与多个源信号之间的关系,在单通道源分离数据集上取得了较好的效果。Jia 等人(2021)针对数据集缺乏的问题,提出 UGAN(U-Net generative adversarial network)与 PAGAN(pixel attention U-Net generative adversarial

network)级联框架,利用 UGAN 生成新样本、PAGAN 分离图像,实现了良好的分离效果。综上所述,尽管 ICA、卷积神经网络(convolutional neural network, CNN)与 GAN 及其诸多变体已为盲图像分离提供了可行思路,但现有方法在真实复杂场景中仍难以有效地分离出严重混叠的区域,从而导致源图像失真。

最近,扩散模型(diffusion model, DM)(Dhariwal 和 Nichol, 2021)在图像处理领域崭露头角,成功应用于图像生成(Lu 等, 2024; Qiao 等, 2025; Zhang 等, 2024a)、图像去噪(Liu 等, 2024a; EI Malki 等, 2025; Yin 等, 2025)、图像超分(Zhang 等, 2024b; Dong 等, 2025)以及图像编辑(Huang 等, 2025; Ma 等, 2025)等多样化任务。相较于以 GAN 的对抗训练范式和 Transformer 的全局注意力机制为基础架构的方法,扩散模型具有强大的渐进式去噪机制与隐式先验知识建模能力,通过反向扩散过程对混合图像进行层级解耦,能够有效处理复杂混合图像中的非平稳噪声分布,同时保持对原始数据分布的强拟合能力。Ho 等人(2020)受非平衡热力学启发提出的去噪扩散概率模型(denoising diffusion probabilistic model, DDPM),首次将离散时间马尔可夫链与变分推断相结合,实现了高质量数据生成。Saharia 等人(2023)利用条件扩散模型生成超分图像并使用 SR3(image super-resolution via iterative refinement)算法迭代细化图像细节。Liu 等人(2024b)提出 Diffusion-based Plugin 框架,使得单个预训练的扩散模型能够在各种低层次的任务中生成高保真度的结果。王义杰等人(2025)提出一种模糊逻辑与扩散模型相结合的方法,增强扩散模型处理文档图像中随机因素的能力,提高文档图像去透射任务的准确性。He 等人(2025)提出一种基于贝叶斯不确定性的扩散概率模型 BUFF(Bayesian uncertainty guided diffusion probabilistic model),通过引入高分辨率不确定性掩码,减少了复杂纹理和细节区域的伪影和模糊。Chen 等人(2025)设计了一种可控采样策略,通过预测初始样本来限制扩散模型的潜在输出范围,从而得到更加准确的去模糊结果。

基于扩散模型优秀的生成范式以及 BIS 的复杂需求,本文首次提出一种基于扩散模型的双通道盲图像分离方法,用来解决两个随机源图像混合双分离。考虑到源图像二者互为耦合噪声,设计了小波抑制模块(wavelet suppression module, WSM),渐进式

地去除了分离过程所产生的伪噪声,提升分离效果。

1 双通道扩散分离模型(DCDSM)

如图1所示,为了更好地分离混合图像结构特征和解耦源图像,本文设计了一个有效的双分支扩散模型框架,从不同的源图像中学习潜在的复杂分布,并利用WSM通过双域协同提取抑制信息辅助解耦,具体实现如下。

1.1 整体网络结构

双通道扩散分离模型(dual channel diffusion sepa-

ration model, DCDSM)是由条件扩散模型组成的双分支结构,整体网络结构如图1所示。混叠图像经过统一的加噪过程生成噪声潜变量,随后并行输入到两个参数独立的去噪分支中协同分离。模型在分支中嵌入WSM,对第 t 个时间步中间结果进行二维离散小波变换,在时域和频域下处理高低频子带获取抑制信息。同时将该信息与另一分支的中间结果进行逐像素差异运算,有效剥离特征中的混合干扰。然后在提取阶段引入时间编码,使提取抑制信息能够自适应不同去噪阶段的噪声分布特性。该模型结合扩散模型损失和约束抑制模块的损失来辅助端到端的训练。

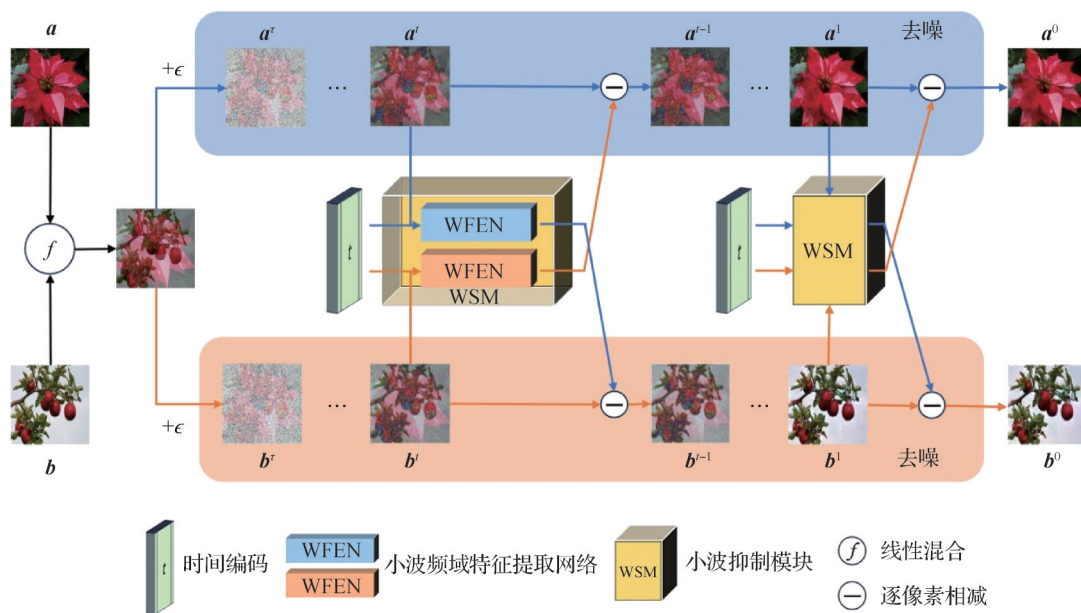


图1 DCDSM的整体网络框架

Fig. 1 Overall network framework of DCDSM

1.2 双通道扩散模型框架

如图1所示,DCDSM可分为前向扩散阶段和反向去噪阶段。首先对混合图像执行统一的前向扩散(加噪过程),以生成噪声干扰图像,使模型学习到图像的潜在特征分布。再通过两个独立的反向去噪分支分别处理加噪后的图像,逐步还原并分离出源图像,实现对两个源的初步分离。

1.2.1 前向扩散阶段

在前向扩散过程中,使用线性调度器 $\{\beta_1, \beta_2, \beta_3, \dots, \beta_T\}$,经过 T 个时间步将输入图像 $\mathbf{x}_0$ 转换为包含噪声的图像数据,这里用马尔可夫链来定义,其表达式为

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

式中, q 表示概率分布, T 为时间步长, $\mathbf{x}_t$ 表示时间步长为 t 时的噪声数据。

在每个时间步,原始图像 $\mathbf{x}_0$ 都会被添加高斯噪声,噪声的量由线性调度器控制,经过 T 个时间步后, $\mathbf{x}_t$ 的数据分布会逐渐转换成接近各向同性的高斯分布。单个时间步加噪过程表达式为

$$q(\mathbf{x}_t | \mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1 - \beta_t} \mathbf{x}_{t-1}, \beta_t \mathbf{I}_c) \quad (2)$$

式中, β_t 是预定方差, $\mathcal{N}$ 表示高斯分布, $\mathbf{I}_c$ 为目标图像。

当 $\beta_t = 1 - \alpha_t$, $\bar{\alpha}_t = \prod_{i=1}^t \alpha_i$ 且 ϵ 从 $\mathcal{N}(0, \mathbf{I})$ 中取样时,根据式(2)递归求得 $\mathbf{x}_t$ 与 $\mathbf{x}_0$ 之间的关系,其表达式为

$$q(\mathbf{x}_i | \mathbf{x}_0) = \mathcal{N}(\mathbf{x}_i; \sqrt{\bar{\alpha}_i} \mathbf{x}_0, (1 - \bar{\alpha}_i) \mathbf{I}_c) \quad (3)$$

1.2.2 反向去噪阶段

在训练过程中的反向去噪阶段, 已知噪声 $\mathbf{x}_i$ 和结果 $\mathbf{x}_0$, 可以精确计算如何从 $\mathbf{x}_i$ 回退到 $\mathbf{x}_0$, 用于在训练时提供监督信号, 其表达式为

$$q(\mathbf{x}_{i-1} | \mathbf{x}_i, \mathbf{x}_0) = q(\mathbf{x}_i | \mathbf{x}_{i-1}, \mathbf{x}_0) \frac{q(\mathbf{x}_{i-1} | \mathbf{x}_0)}{q(\mathbf{x}_i | \mathbf{x}_0)} = \mathcal{N}(\mathbf{x}_{i-1} | \tilde{\mu}_i(\mathbf{x}_i, \mathbf{x}_0), \tilde{\beta}_i \mathbf{I}_c) \quad (4)$$

式中, $q(\mathbf{x}_{i-1} | \mathbf{x}_i, \mathbf{x}_0)$ 为加噪逆操作, $q(\mathbf{x}_i | \mathbf{x}_{i-1}, \mathbf{x}_0)$ 表示加噪分布。由于前向加噪是一个马尔可夫链, 所以 $q(\mathbf{x}_i | \mathbf{x}_{i-1}, \mathbf{x}_0)$ 等价于式(2), $q(\mathbf{x}_{i-1} | \mathbf{x}_0)$ 与 $q(\mathbf{x}_i | \mathbf{x}_0)$ 可由式(3)得到。而在推理过程中, 反向去噪则是从输入的高斯噪声 $\mathbf{x}_i$ 开始, 通过去噪器随时间步逐步去除噪声, 从而拟合与条件图像 $\mathbf{I}_c$ 相关的残差分布, 其表达式为

$$q(\mathbf{x}_{i-1} | \mathbf{x}_i) = \mathcal{N}(\mathbf{x}_{i-1} | \mu_\theta(\mathbf{x}_i, t), \Sigma_\theta(\mathbf{x}_i, t)) \quad (5)$$

式中, $\mu_\theta(\mathbf{x}_i, t)$ 和 $\Sigma_\theta(\mathbf{x}_i, t)$ 分别代表均值和方差。DCDSM 使用网络 $p_\theta(\mathbf{x}_{0:T} | \mathbf{I}_c)$ 拟合整个逆过程。去噪过程中的均值 $\mu_\theta(\mathbf{x}_i, t)$ 可以通过预测噪声 $\epsilon_{\theta(\mathbf{x}_i, t)}$ 间接计算, 此处直接忽略方差足以实现高质量的生成。其表达式为

$$\mu_\theta(\mathbf{x}_i, t) = \tilde{\mu}_i \left(\mathbf{x}_i, \frac{1}{\sqrt{\bar{\alpha}_i}} \left(\mathbf{x}_i - \sqrt{1 - \bar{\alpha}_i} \epsilon_{\theta(\mathbf{x}_i, t)} \right) \right) = \frac{1}{\sqrt{\bar{\alpha}_i}} \left(\mathbf{x}_i - \frac{\beta_i}{\sqrt{1 - \bar{\alpha}_i}} \epsilon_{\theta(\mathbf{x}_i, t)} \right) \quad (6)$$

式中, t 和 $\mathbf{x}_i$ 已知。在去噪过程中加入去噪器 $\epsilon_{\theta(\mathbf{x}_i, t)}$, 得到单步去噪。具体为

$$p_\theta(\mathbf{x}_{i-1} | \mathbf{x}_i, \mathbf{I}_c) = q(\mathbf{x}_{i-1} | \mathbf{x}_i, \epsilon_{\theta}(\mathbf{x}_i, t, \mathbf{I}_c)) \quad (7)$$

然后, 从噪声样本分布 $p(\mathbf{x}_T)$ 开始逐步去噪, 整个逆过程为

$$p_\theta(\mathbf{x}_{0:T} | \mathbf{I}_c) = p(\mathbf{x}_T) \prod_{t=1}^T p_\theta(\mathbf{x}_{t-1} | \mathbf{x}_t, \mathbf{I}_c) \quad (8)$$

1.3 小波抑制模块

在处理复杂且未知的混合源图像时, 单独使用条件扩散模型进行分离仍存在噪声、伪影和分离不完全等问题, 实验如图2所示。考虑到源图像二者互为耦合噪声, 本文提出一种小波抑制模块, 通过双分支扩散模型的交互机制, 使得从两个分支所提取

到的抑制信息能够协同优化分离过程。该模块的核心设计基于小波变换的特性: 其能同时捕捉图像的整体轮廓(低频信息)和局部细节(高频信息, 如边缘、纹理)。具体而言, 模块从一条去噪分支的中间输出中提取抑制信息, 该抑制信息包含难以分离的轮廓和边缘纹理特征, 如图3中的可视化输出所示。并将其与另一分支的中间输出进行像素相减, 有效地剥离其中因分离不充分产生的特征混叠、伪影等噪声。该模块在反向去噪过程中动态插入, 其时间结点为 $t = \alpha$ 和 $t = 0$, α 为设定的超参数, 旨在有效提取中间结果的抑制信息, 以消除另一分支中的噪声。模块由两个独立的小波频域特征提取网络(wavelet frequency domain feature extraction network, WFEN)构成, 分别用于提取两个分支在去噪过程中的抑制信息, 以解耦两个源图像, 进一步增强模型在图像分离任务中的性能。

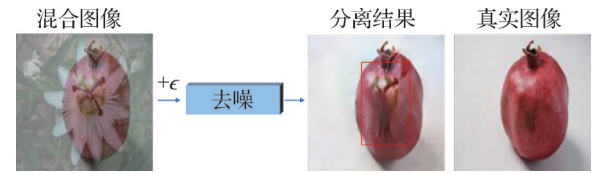


图2 单独使用条件扩散模型分离混合图像结果

Fig. 2 Separate mixed image results using conditional diffusion model alone

小波频域特征提取网络(WFEN)具体结构如图3所示。该网络在反向去噪过程中将上一时间步的含噪结果作为输入, 利用小波变换将图像划分为4个子带: $[LL, LH, HL, HH]$ 。其中, 3个高通子带 $[LH, HL, HH]$ 包含图像的高频边缘和纹理等重要信息。这类信息在另一分支的去噪过程中通常较难处理, 且容易留下伪影。将高通子带按通道拼接, 然后通过 1×1 的卷积操作来调整其尺寸和通道, 使其与原始输入相同, 随后进行逐像素相加, 将处理后的结果输入到U-Net网络中, 借助编码器—解码器的结构捕获全局上下文, 重建局部细节, 以进一步增强边缘和纹理信息。最后, 其输出再次通过小波变换, 得到增强后的高通子带 $[LH', HL', HH']$ 。具体为

$$LH', HL', HH' = DWT_{2D} \left(U \left(\text{Conv} (LH + HL + HH) \right) \right) \quad (9)$$

式中, DWT_{2D} 是2D离散小波变换, U 表示3层U-Net网络。

低通子带 LL 包含图像中大部分的结构特征信

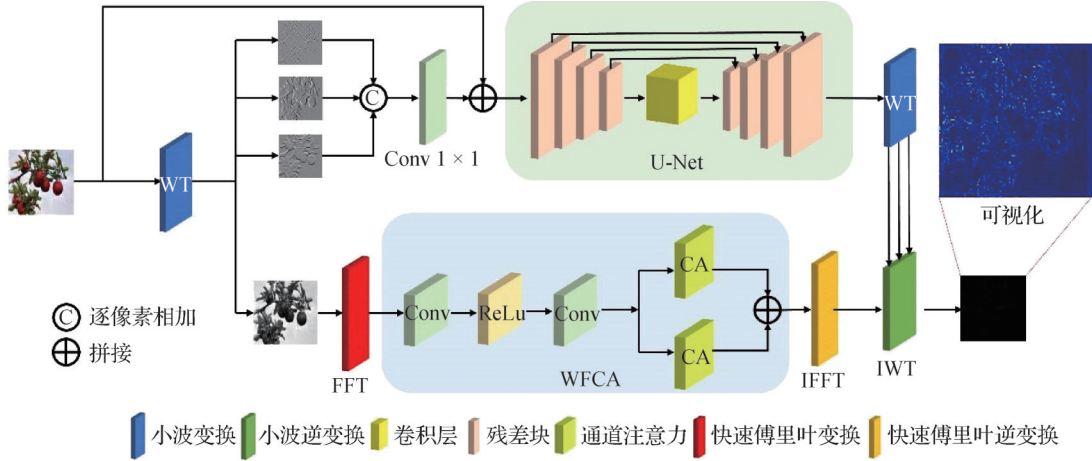


图3 小波频域特征提取模块网络结构

Fig. 3 Network structure of wavelet frequency domain feature extraction module

息,干扰背景的分离,所以网络采用傅里叶变换将该子带转换到频域,通过基于窗口的频率通道注意力机制(window-based frequency channel attention, WFCA)过滤掉不必要的背景信息,WFCA如图3中所示。这一过程旨在保留关键信息,同时防止模块输出在另一分支做逐像素相减时有效信息的流失。 LL' 的具体计算为

$$LL' = IFFT(WFCA(FFT(LL))) \quad (10)$$

式中, FFT 和 $IFFT$ 分别表示快速傅里叶变换和快速傅里叶逆变换,WFCA是基于窗口的频域通道注意力机制。

最后,将式(9)和式(10)所得到的高通子带 $[LH', HL', HH']$ 和低通子带 LL' 放入小波逆变换得到最终输出,具体为

$$\mathbf{x}_{out} = IDWT_{2D}(\{LL', LH', HL', HH'\}) \quad (11)$$

式中, $IDWT_{2D}$ 是离散小波逆变换。

1.4 损失函数

DCDSM训练目标是通过训练去噪器最小化预测噪声 ϵ_θ 与实际噪声之间的距离,损失函数具体表达式为

$$\mathcal{L}_{diff} = E_{\mathbf{x}_0^1, \epsilon_t \sim \mathcal{N}(0, I)} [\|\epsilon_t - \epsilon_\theta(\mathbf{x}_t^1, t, I_c^1)\|^2] + E_{\mathbf{x}_0^2, \epsilon_t \sim \mathcal{N}(0, I)} [\|\epsilon_t - \epsilon_\theta(\mathbf{x}_t^2, t, I_c^2)\|^2] \quad (12)$$

WFEN训练与DCDSM保持同步。为了能更好地约束该网络,确保所提取的抑制信息能够在分支中起到提高分离效果、去除伪影等作用。损失函数采用均方误差计算融合抑制信息和未融合抑制信息的中间输出分别到目标图像的距离之差,并对其指

数化保证非负性,其表达式为

$$\mathcal{L}_{WFEN} = 2 \left(\|\mathbf{x}_t^1 - \mathbf{x}_{out}^1 - I_c^1\|^2 - \|\mathbf{x}_t^1 - I_c^1\|^2 \right) + 2 \left(\|\mathbf{x}_t^2 - \mathbf{x}_{out}^2 - I_c^2\|^2 - \|\mathbf{x}_t^2 - I_c^2\|^2 \right) \quad (13)$$

式中, $\mathbf{x}_t^1$ 和 $\mathbf{x}_t^2$ 分别表示两个分支在时间步 t 的中间输出, $\mathbf{x}_{out}^1$ 和 $\mathbf{x}_{out}^2$ 分别表示 $\mathbf{x}_t^1$ 和 $\mathbf{x}_t^2$ 经过小波频域特征提取模块所得到的抑制信息, I_c^1 和 I_c^2 分别为训练所使用的目标源图像。

在训练中,除了优化扩散模型的目标函数 $\mathcal{L}_{diff}$ 外,还需要兼顾抑制模块的目标函数 $\mathcal{L}_{WFEN}$ 。由于抑制模块的目标函数采用指数的形式,为确保网络在训练阶段不会出现损失收敛失衡的情况,设计的 $\mathcal{L}_{total}$ 表达式为

$$\mathcal{L}_{total} = \gamma \mathcal{L}_{diff} + \mathcal{L}_{WFEN} \quad (14)$$

式中, γ 为超参数。

2 实验

2.1 实验细节和参数设置

实验GPU为NVIDIA GeForce RTX 3090。在训练阶段采用Adam优化器更新模型参数,图像尺寸固定为 256×256 像素,批量大小为8,学习率设置为0.0001,总扩散步长 T 设置为1000,训练迭代次数为100000。

2.2 数据集与评估指标

DCDSM将两个源分别作为真实结果(ground truth, GT)参与训练,在合成数据集上进行验证。从不同的应用场景出发,构建了包括去雨、去雪和模拟复杂混合的合成数据集。

2.2.1 数据集

1)含雨数据集。包含 1 120 幅训练数据和499 幅测试数据。图像选自 CityScape(Cordts 等,2016)数据集,利用雨带掩码图像通过随机混合的方法生成含雨图像。为了验证网络的鲁棒性并更好地模拟真实场景中的雨水效果,采用两种不同的融合策略:第 1 种是直接进行线性混合(Rain),第 2 种是将掩码图像的透明度设置为 0.5 后再进行线性混合(Rain-0.5)。

2)含雪数据集。包含 2 999 幅训练图像和499 幅测试图像。该数据集以 CityScape 图像作为清晰背景,通过与 Snow100K 数据集(Liu 等,2018)中提供的雪掩码进行线性混合,合成了相应的含雪图像数据集(Snow)。未直接采用 Snow100K 中的原始图像作为真实背景的原因在于,该数据集中的图像多数光照较强、天气晴朗,与实际下雪天气常伴随的阴暗光照条件存在显著差异。为避免合成数据与真实雪天成像表现不符,选用 CityScape 数据集作为背景,因其图像整体色调偏暗、多为阴天场景,更符合真实雪天的视觉表现。

3)复杂混合数据集。包含 2 499 幅训练数据和 249 幅测试数据。从花朵和水果两种数据集中随机挑选进行线性混合后分配成训练和测试数据集(complex-mixture)。两个源数据集中的花朵和水果分别含有多个不同品种来模拟实际复杂的未知

盲源。

2.2.2 评估指标

去雨、去雪和混合图像分离实验均采用峰值信噪比(peak signal-to-noise ratio,PSNR)和结构相似性指数(structural similarity,SSIM)评估网络分离能力。PSNR 越高代表混合图像中噪声的分离效果和源的复原效果越好。SSIM 计算网络分离结果图像在亮度、对比度和结构方面与真实图像的相似程度,更大的 SSIM 值即更大的相似性。

2.3 去雨图像实验

表 1 比较了 DCDSM 和现有方法在含雨数据集上的客观度量指标。对比方法包括:DerainNet(Fu 等,2017)、PreNet(Ren 等,2019)、MPRNet(multi-scale pointwise regression network)(Zamir 等,2021)、Restormer(Zamir 等,2022)、BID(blind image decomposition)(Han 等,2022)、DRSformer(deraining network sparse transformer)(Chen 等,2023)、NeRD(neural representations for image deraining)(Chen 等,2024)、LDRCNet(neural representations for image deraining)(He 等,2024)和 Histoformer(Sun 等,2024)。DCDSM 在两个不同程度的雨纹上分别实现了 36.244 3 dB/33.760 2 dB 的 PSNR 和 0.961 5/0.948 2 的 SSIM。结合两组数据的实验结果,DCDSM 实现了最优的去雨结果。由于参与对比方法没有单独提取雨纹掩码的能力,所以在后续单独讨论与其他方

表 1 去雨实验结果对比
Table 1 Comparison of rain removal experiment results

方法	Rain-0.5		Rain		平均值		Param/M	FLOPs/G
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM		
DerainNet	26.396 1	0.814 1	25.482 3	0.794 8	25.939 2	0.804 5	1.25	42.34
PreNet	25.453 8	0.779 7	25.455 2	0.779 6	25.454 5	0.779 7	0.095	10.54
MPRNet	25.642 5	0.764 5	24.983 1	0.734 9	25.183 2	0.749 7	3.64	148.55
Restormer	27.978 8	0.872 7	26.357 4	0.802 2	27.168 1	0.837 5	26.10	140.99
BID	33.475 2	0.904 8	30.458 3	0.851 4	31.966 8	0.878 1	144.9	104.56
DRSformer	26.650 2	0.794 1	22.451 6	0.622 7	24.550 9	0.708 4	33.65	242.91
NeRD	34.949 7	0.945 6	31.533 2	0.907 6	33.241 5	0.922 6	22.89	156.35
LDRCNet	31.526 2	0.894 8	30.760 2	0.887 4	31.148 2	0.891 1	3.64	148.84
Histoformer	36.037 6	0.950 6	31.452 9	0.906 7	33.745 3	0.928 7	16.62	91.59
DCDSM(本文)	36.244 3	0.961 5	33.760 2	0.948 2	35.002 3	0.954 9	16.71	96.61

注:加粗字体表示各列最优结果。

法的分离性能对比。

图 4 展示了 DCDSM 方法与其他对比方法在 Rain-0.5 数据集上的去雨效果。针对暴雨天气雨纹密集场景,如图 4(a)所示,DerainNet 和 Restormer 的去雨效果较差,在图中仍存在明显雨线痕迹。LDRCNet、BID 和 Histoformer 有效去除了雨纹,但图中树叶、花坛等区域纹理细节过度平滑。PreNet 和 NeRD 在去除雨纹后存在明显的伪影残留。本文方法在雨纹的去除和细节的恢复上都要优于其他对比方法。图 4(b)展示了不同方法在小雨天气雨纹稀疏情况下的去雨表现。DerainNet 对稀疏雨纹依然不能有效去除,且相较于 GT 图像,在颜色上存在明

显饱和度偏差。PreNet 和 NeRD 的伪影问题依然存在,在雨痕区域所复原出的场景颜色严重失真。LDRCNet 在车标附近出现雨纹残留,对场景中雨纹相似的特征不能有效区分。DCDSM 在稀疏雨纹下与 GT 图像几乎无差异。

综上,DCDSM 在视觉效果和不同程度雨纹的去除性能上均表现最佳。

2.4 去雪图像实验

各种对比方法去雪效果如图 5 所示,对比方法包括:DesnowNet (Liu 等, 2018)、HDCWNet (hierarchical dual-tree complex wavelet network) (Chen 等, 2021)、TransWeather (Valanarasu 等, 2022)、BID (Han

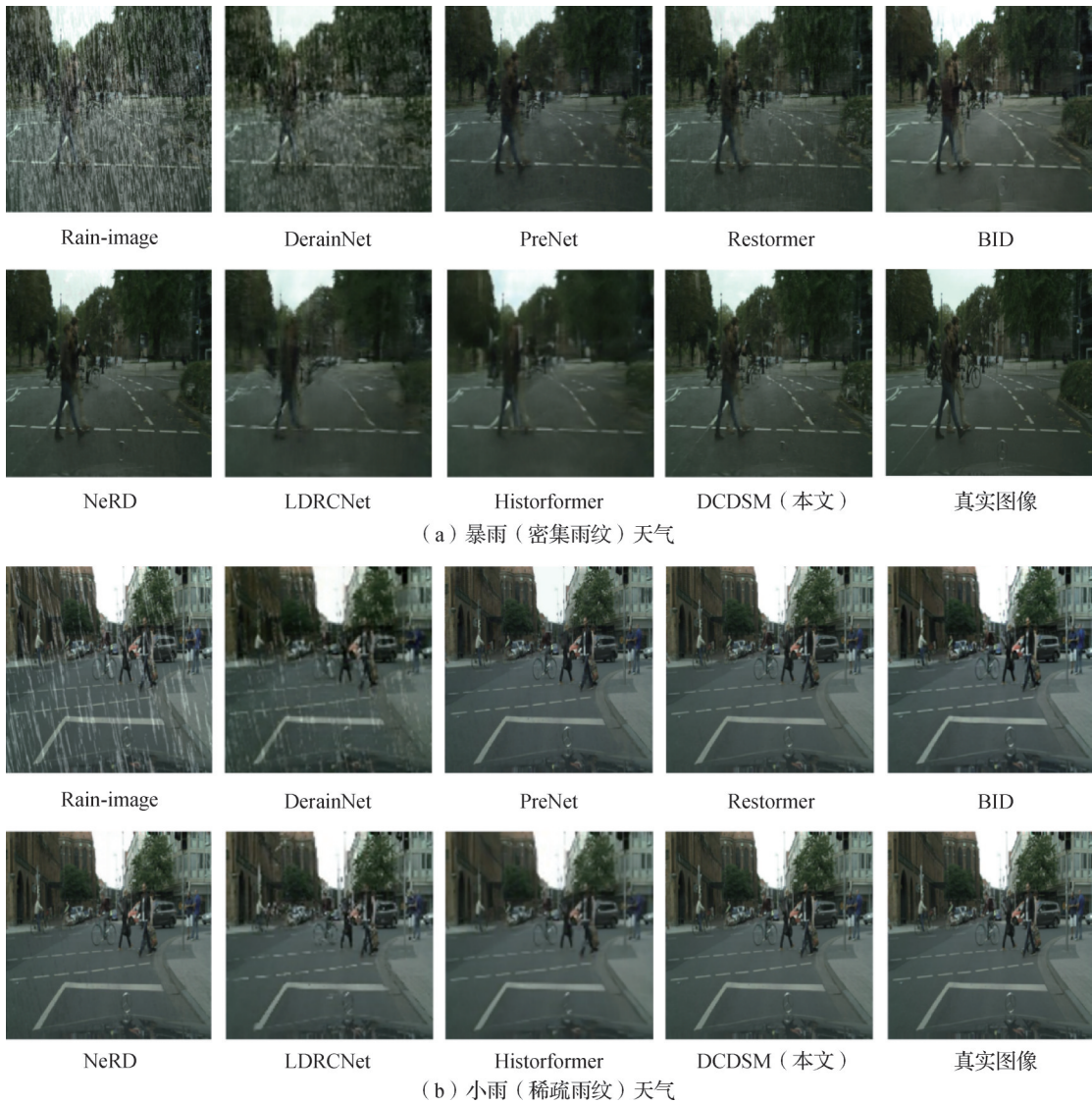


图 4 Rain-0.5 数据集上各方法去雨效果
Fig. 4 Rain removal effect of various methods on Rain-0.5 dataset
((a) rainstorm (dense rain pattern) weather; (b) light rain (sparse rain lines) weather

等, 2022)、SMGARN (snow mask guided adaptive residual network) (Cheng 等, 2023)、LDRCNet (latent degradation representation censtraint network) (He 等, 2024) 和 Histoformer (Sun 等, 2024)。从图中可见: DesnowNet、HDCWNet、TransWeather 和 BID 在去除雪花方面表现不佳, 在图中存在大量雪花痕迹, 且严重失真。SMGARN 仍残留部分伪影和雪花, 尤其是

在天空区域的雪花残留严重。Histoformer 去除雪花后所恢复的场景在颜色上存在差异, 破坏了整体色彩稳定性。LDRCNet 在细节处理上欠佳, 例如将图中车标部分的反光视为雪花进行了淡化处理。本文提出的方法在雪花去除及细节复原方面较对比方法, 显著提升了视觉保真度, 达到了最优的性能指标。上述各种方法去雪的客观度量指标如表 2 所示。

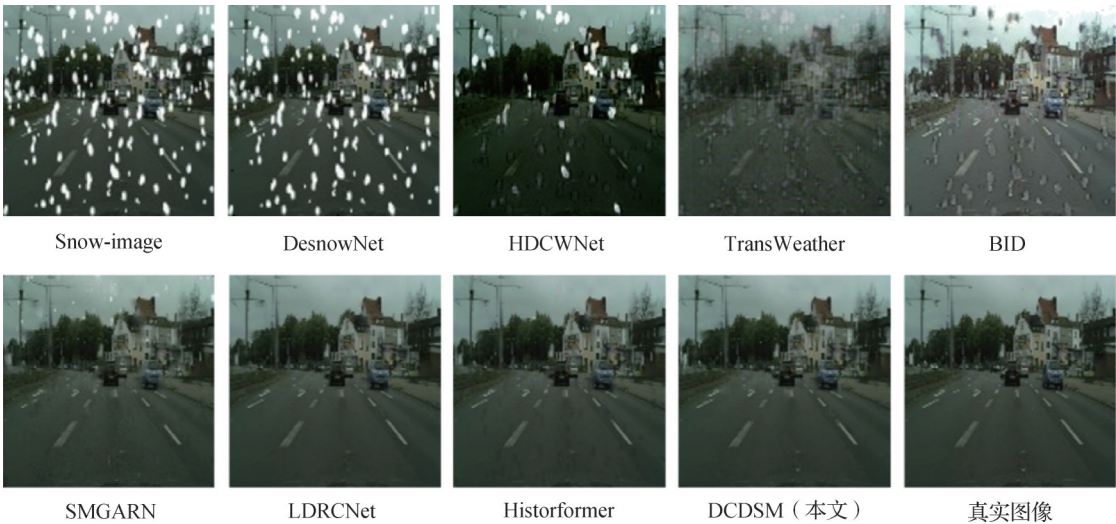


图 5 各种对比方法去雪效果

Fig. 5 Comparison of snow removal effects by various methods

表 2 去雪实验结果对比

Table 2 Comparison of snow removal test results

方法	PSNR/dB	SSIM	Param/M	FLOPs/G
DesnowNet	19.817 9	0.767 9	15.64	1.7K
HDCWNet	17.592 9	0.622 9	6.99	9.78
Snowformer	15.024 0	0.565 5	8.38	19.74
Restormer	20.527 1	0.804 2	26.10	140.99
TransWeather	21.968 4	0.829 9	21.90	5.64
BID	19.858 0	0.757 0	144.9	104.56
SMGARN	27.577 1	0.901 7	6.86	450.30
LDRCNet	28.883 6	0.895 4	3.64	148.84
Histoformer	27.770 0	0.852 4	16.62	91.59
DCDSM(本文)	29.810 8	0.924 3	16.71	96.61

注: 加粗字体表示各列最优结果。

2.5 混合图像分离实验

为了评估 DCDSM 在双通道分离实验中的性能, 在去雨、去雪和复杂混合数据集上分别与 BID (Han 等, 2022) 方法对比验证, 客观度量指标如表 3 所示。

图 7 给出了 DCDSM 与 BID 在 3 个数据集上部分图像的双分离效果。在复杂混合图像的实验中, 如图 6 (a), BID 方法在第 2 列分离结果中出现了明显的黑色斑点及花蕊纹理, 导致了严重的图像失真。而本文方法有效减轻了纹理的干扰, 实现对黑色斑点区域真实特征的准确分离。在颜色准确性和分离效果方面, 本文方法均显著优于 BID 方法, 充分显示了其在提高图像分离质量方面的优势。在去雪和去雨的实验结果中, 见图 6 (b) (c), 虽然 BID 在掩码分离方面表现出色, 但在分离场景的结果中仍然存在雨雪的残留和伪影。相比之下, DCDSM 在这两个任务的分离效果上都展现了良好的视觉质量, 在 3 个数据集下的指标均优于 BID。

2.6 消融实验分析

为了测试所提方法参数的有效性, 在 Rain-0.5 数据集上进行了消融实验。验证了去噪过程融入抑制模块策略中所设置的超参数 α (见图 7) 和损失函数中设置的超参数 γ (见图 8) 的可解释性, 以及所提出的 WSM 的有效性 (见表 4)。

表 3 混合图像双分离实验结果对比

Table 3 Comparison of double separation experimental results of mixed images

方法	复杂混合数据集				Rain				Snow			
	水果		花		清晰图像		雨带掩码		清晰图像		雪花掩码	
	PSNR /dB	SSIM	PSNR /dB	SSIM	PSNR /dB	SSIM	PSNR /dB	SSIM	PSNR /dB	SSIM	PSNR /dB	SSIM
BID	20.780 8	0.725 3	20.979 1	0.688 8	30.458 3	0.851 4	25.529 9	0.739 5	19.858 0	0.757 0	24.386 6	0.922 8
DCDSM	25.509 0	0.813 0	24.500 8	0.786 3	33.760 2	0.948 2	27.313 7	0.829 4	29.810 8	0.924 3	25.636 8	0.939 0

注:加粗字体表示各列最优结果。

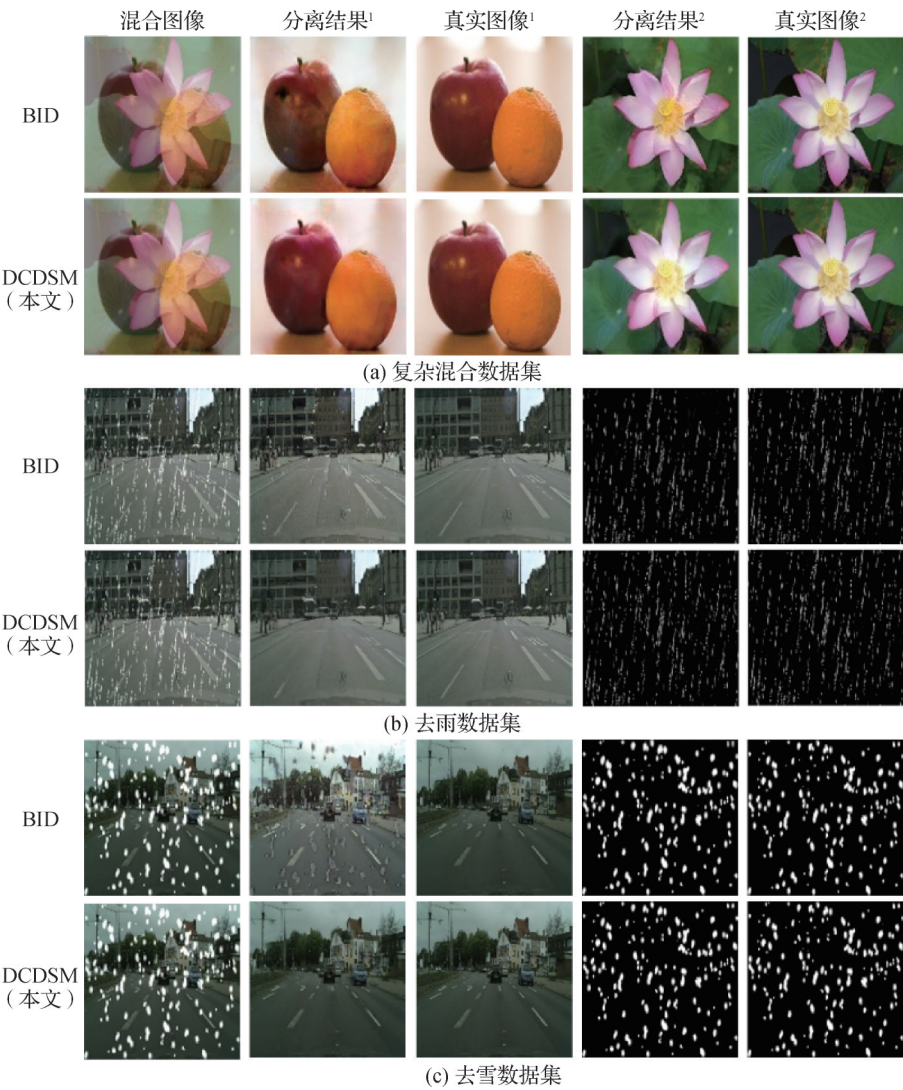


图 6 复杂混合图像、含雨和含雪数据集上双分离实验可视化结果

Fig. 6 Visualization results of double separation experiments on complex mixed images, rain and snow datasets
((a) complex mixed datasets; (b) rain removal datasets; (c) snow removal datasets)

在表 4 中,进行了以下消融实验:(I)基础框架模型,未引入 WSM;(II)在去噪过程的时间步 $t = 5$ 时引入 WSM;(III)在去噪阶段 $t = 0$ 时引入 WSM;(IV)分别在 $t = 5$ 和 $t = 0$ 时引入 WSM。根据表 4 所

示,在时间步 $t = 5$ 和 $t = 0$ 时,引入抑制模块分别使指标 (PSNR/SSIM) 提高 1. 300 4 dB/0. 113 7 和 7. 563 1 dB/0. 138 3;二者同时引入时,使 PSNR 提高 8. 431 8 dB, SSIM 提高 0. 154 8。

图7是抑制模块融合策略中设置超参数 α 的消融实验结果。通过调节 α 的值改变抑制模块在去噪过程中所要插入的时间节点 t 。当 $\alpha = 0$ 时,其等同于表4中的(Ⅲ)。随着 α 的增大,即时间步 t 的增大,中间输出图像中所含的噪声占比越大,这会导致抑制模块无法有效且全面地从中提取出抑制信息,从而对分离效果影响较小。

超参数 γ 的消融实验结果如图8所示。 γ 的作用是

调节扩散模型和抑制模块的损失在训练过程中所占的权重比例, γ 越大代表在训练时模型越偏向于双通道扩散模型的分离性能,反之则表示模型更加注重WSM的作用。从图8中可以看出,随着 γ 增加,模型分离性能主要依赖于基础的条件扩散模型,所以其指标逐渐趋近于表4中的(Ⅰ)的结果。在实验中设置 $\gamma = 2$ 和 $\gamma = 4$ 时,均出现损失收敛失衡导致指标下降的问题,而在 $\gamma = 3$ 时,损失收敛达到最优,获得了最优的效果。

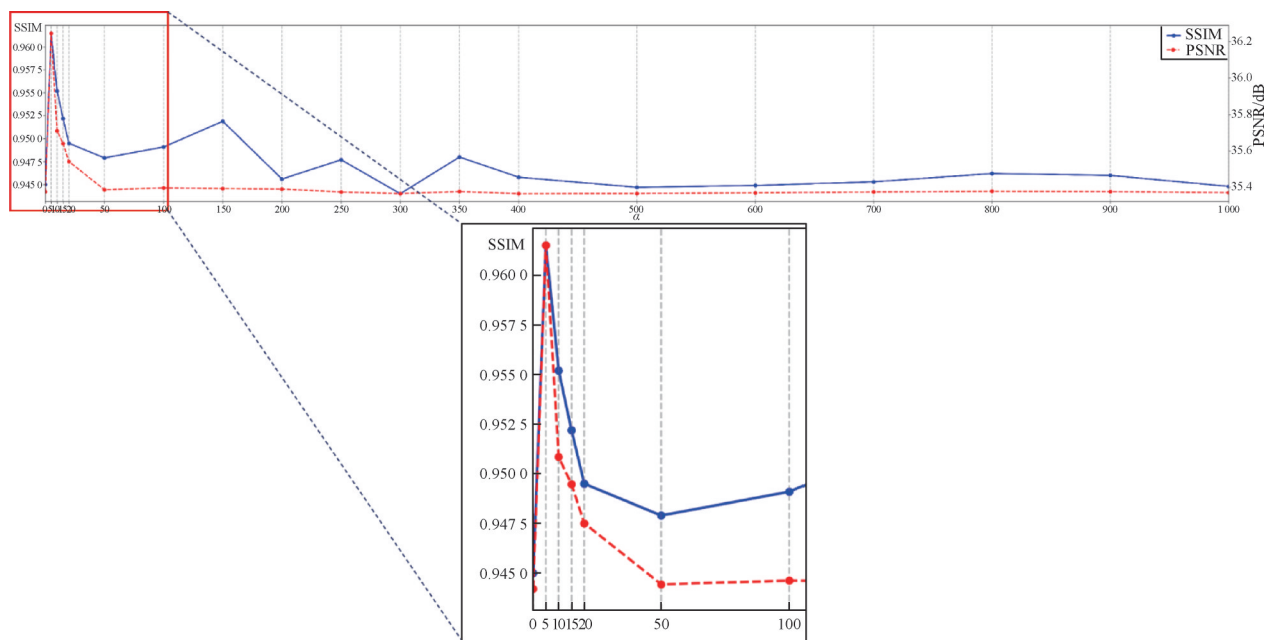


图7 超参数 α 的消融实验

Fig. 7 Ablation experiment of super parameter α

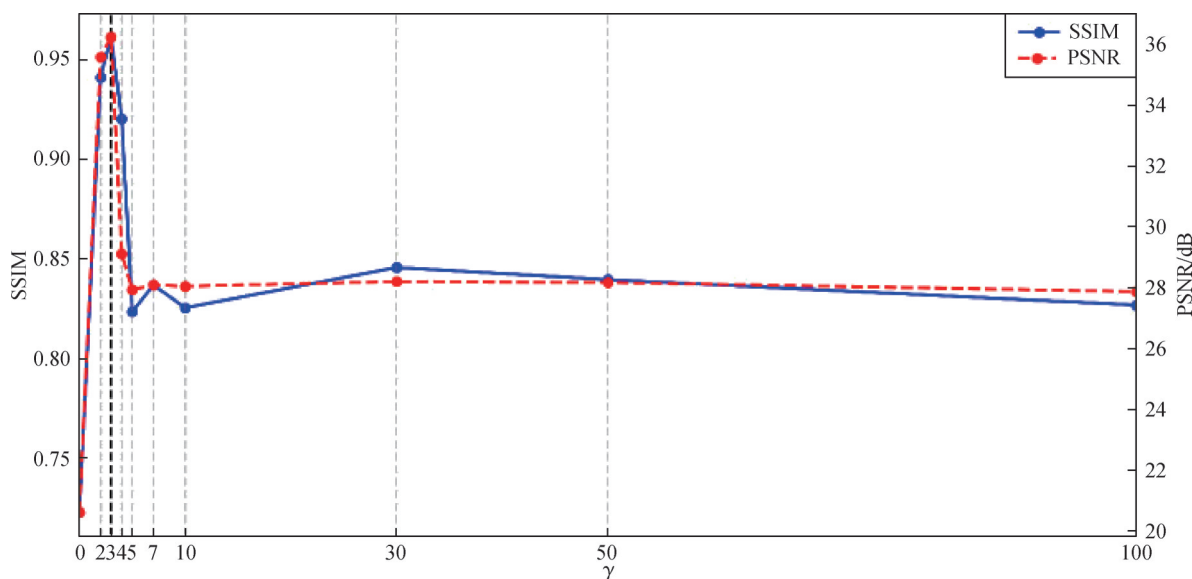


图8 超参数 γ 的消融实验

Fig. 8 Ablation experiment of super parameter γ

表 4 小波抑制模块消融实验
Table 4 Ablation experiment of wavelet suppression module

方法	PSNR/dB	SSIM
(I) DCDSM 基础框架模型, 未引入 WSM	27.812 5	0.806 7
(II) DCDSM 在去噪过程时间步 $t = 5$ 时引入 WSM	29.112 9	0.920 4
(III) DCDSM 在去噪阶段 $t = 0$ 时引入 WSM	35.375 6	0.945 0
(IV) DCDSM 分别在 $t = 5$ 和 $t = 0$ 时引入 WSM	36.244 3	0.961 5

注:加粗字体表示各列最优结果。

3 结 论

双通道盲源分离作为通用性任务,其应用场景广泛,如本文实验中涉及的去雨和去雪等图像处理任务。本文提出一种基于小波抑制扩散模型的双通道盲图像分离网络,旨在解决复杂混合图像双分离问题。提出的网络采用双分支条件扩散模型框架,通过在反向去噪过程中策略性融合小波抑制模块,进一步增强复杂混合图像的双分离效果。设计并引入小波频域特征提取模块,通过小波变换与傅里叶变换的多域特征协同,抑制分离过程中的伪影,提升图像细节还原能力。

为验证 DCDSM 的有效性,在多个合成数据集(去雨/雪、复杂混合)上进行实验,并与当前主流方法进行对比。大量实验结果表明,本文方法在主观视觉和客观指标上均取得较好的结果,解决了存在雨纹和雪纹残留、混合图像中细节模糊等问题。在主观视觉方面,本文方法在去雨、去雪任务中复原的图像视觉效果最佳,有效地消除了雨纹和雪纹的残留,背景细节恢复更为清晰自然。而客观指标方面,在 PSNR 和 SSIM 等关键指标上,本文方法均取得最优结果。在去雨测试集上,PSNR 达到 35.002 3 dB,比次优方法 Histoformer 高出 1.257 dB;SSIM 指标达到 0.954 9,表明本文方法的结果在像素精度和结构相似性上均有显著提升。这些定量结果充分说明了本文方法在分离精度和细节还原方面的优势。

尽管本研究取得了上述成果,但当前方法仍存在一些局限性。首先,由于基于扩散模型,其多步迭

代的去噪过程导致了较高的计算耗时和资源占用。这主要是因为扩散模型本身固有的序列生成特性,难以在推理阶段进行并行化加速。其次,模型在训练数据分布之外的极端混合场景(如暴雨、暴雪)或未知类型的混合模式下,性能可能出现衰减。这源于模型对训练数据中混合先验的依赖,其抑制模块的参数在面对分布外数据时适应性不足。

基于当前研究的成果与不足,未来的工作将重点围绕模型轻量化与加速、增强泛化与鲁棒性和拓展应用边界这 3 点展开。首先,探索潜在扩散模型等先进技术,旨在显著减少推理步数,降低计算成本,推动方法的轻量化部署;其次,改进小波抑制模块的自适应能力,研究引入领域自适应策略,使其能动态适应各种未见的混合退化模式,提升模型的泛化性能;最后,将积极探索本模型在新兴领域的应用潜力。例如,将其应用于透射文档复原任务,通过分离纸张背面的透射文字信息,恢复文档正面的清晰内容,验证其在解决实际分离问题中的价值。

参考文献(References)

Chen W T, Fang H Y, Hsieh C L, Tsai C C, Chen I H, Ding J J, et al. 2021. All snow removed: single image desnowing algorithm using hierarchical dual-tree complex wavelet representation and contradict channel loss//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 4176-4185 [DOI: 10.1109/iccv48922.2021.00416]

Chen X, Li H, Li M Q and Pan J S. 2023. Learning a sparse transformer network for effective image deraining//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 5896-5905 [DOI: 10.1109/CVPR52729.2023.00571]

Chen X, Pan J S and Dong J X. 2024. Bidirectional multi-scale implicit neural representations for image deraining//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE: 25627-25636 [DOI: 10.1109/CVPR52733.2024.02421]

Chen Z Y, Cui G M, Zhao J F and Nie J H. 2025. CDRM: controllable diffusion restoration model for realistic image deblurring. Expert Systems with Applications, 275: #127009 [DOI: 10.1016/j.eswa.2025.127009]

Cheng B D, Li J C, Chen Y and Zeng T Y. 2023. Snow mask guided adaptive residual network for image snow removal. Computer Vision and Image Understanding, 236: #103819 [DOI: 10.1016/j.cviu.2023.103819]

- Cordts M, Omran M, Ramos S, Rehfeld T, Enzweiler M, Benenson R, et al. 2016. The cityscapes dataset for semantic urban scene understanding//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 3213-3223 [DOI: 10.1109/cvpr.2016.350]
- Dhariwal P and Nichol A. 2021. Diffusion models beat GANs on image synthesis//Proceedings of the 35th International Conference on Neural Information Processing Systems. Red Hook, USA: Curran Associates Inc.: #672 [DOI: 10.5555/3540261.3540933]
- Dong S Y, Cai Z T, Hangel G, Bogner W, Widhalm G, Huang Y Q, et al. 2025. A flow-based truncated denoising diffusion model for super-resolution magnetic resonance spectroscopic imaging. *Medical Image Analysis*, 99: #103358 [DOI: 10.1016/j.media.2024.103358]
- El Malki I, El Hakoume A, Laghrib A and Hadri A. 2025. Improved image denoising via self-supervised Weickert operator learning and plug-and-play learned Primal Dual. *Neurocomputing*, 620: #129267 [DOI: 10.1016/j.neucom.2024.129267]
- Fu X Y, Huang J B, Zeng D L, Huang Y, Ding X H and Paisley J. 2017. Removing rain from single images via a deep detail network//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 1715-1723 [DOI: 10.1109/cvpr.2017.186]
- Han J L, Li W H, Fang P F, Sun C Y, Hong J, Armin M A, et al. 2022. Blind image decomposition//Proceedings of the 17th European Conference on Computer Vision. Tel Aviv, Israel: Springer Nature: 218-237 [DOI: 10.1007/978-3-031-19797-0_13]
- He Y H, Peng L, Wang L and Cheng J. 2024. Latent degradation representation constraint for single image deraining//Proceedings of 2024 ICASSP IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Seoul, Korea (South): IEEE: 3155-3159 [DOI: 10.1109/icassp48485.2024.10447510]
- He Z H, Zhang S C, Hu R Z, Shen Y H and Zhang Y. 2025. BUFF: Bayesian uncertainty guided diffusion probabilistic model for single image super-resolution//Proceedings of the 39th AAAI Conference on Artificial Intelligence. Philadelphia, USA: AAAI, 39 (4): 3474-3482 [DOI: 10.1609/aaai.v39i4.32360]
- Ho J, Jain A N and Abbeel P. 2020. Denoising diffusion probabilistic models//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc.: #574 [DOI: 10.48550/arxiv.2006.11239]
- Hoshen Y. 2019. Towards unsupervised single-channel blind source separation using adversarial pair unmix-and-remix//Proceedings of 2019 ICASSP IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Brighton, UK: IEEE: 3272-3276 [DOI: 10.1109/icassp.2019.8682375]
- Huang Y, Huang J C, Liu Y F, Yan M F, Lyu J X, Liu J Z, et al. 2025. Diffusion model-based image editing: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47 (6): 4409-4437 [DOI: 10.1109/TPAMI.2025.3541625]
- Hyvärinen A and Oja E. 2000. Independent component analysis: algorithms and applications. *Neural Networks*, 13 (4/5): 411-430 [DOI: 10.1016/s0893-6080(00)00026-5]
- Jia F, Xu J D, Sun X, Ma Y L and Ni M Y. 2021. Blind image separation method based on cascade generative adversarial networks. *Applied Sciences*, 11(20): #9416 [DOI: 10.3390/app11209416]
- Liu H P, Wang Y, Qian B, Wang M and Rui Y. 2024a. Structure matters: tackling the semantic discrepancy in diffusion models for image inpainting//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 8038-8047 [DOI: 10.1109/cvpr52733.2024.00768]
- Liu T, Wang W W, Zhang X F and Guo Y N. 2022. One to multiple mapping dual learning: learning multiple signals from one mixture. *Digital Signal Processing*, 129: #103686 [DOI: 10.1016/j.dsp.2022.103686]
- Liu Y F, Jaw D W, Huang S C and Hwang J N. 2018. DesnowNet: context-aware deep network for snow removal. *IEEE Transactions on Image Processing*, 27(6): 3064-3073 [DOI: 10.1109/tip.2018.2806202]
- Liu Y H, Ke Z H, Liu F, Zhao N X and Lau R W H. 2024b. Diff-plugin: revitalizing details for diffusion-based low-level tasks//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 4197-4208 [DOI: 10.1109/CVPR52733.2024.00402]
- Lu S L, Wang Z L, Li L Y, Liu Y Z and Kong A W K. 2024. MACE: mass concept erasure in diffusion models//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 6430-6440 [DOI: 10.1109/cvpr52733.2024.00615]
- Ma Y S, Zhang D, Li S S, Cambria E and Zhou G D. 2025. Ueco: unified editing chain for efficient appearance transfer with multimodality-guided diffusion. *Expert Systems with Applications*, 270: #126510 [DOI: 10.1016/j.eswa.2025.126510]
- Qiao L H, Wang R X, Shu Y C, Li B B, Li W S, Gao X B, et al. 2025. Fast sampling of diffusion models for accelerated MRI using dual manifold constraints. *IEEE Transactions on Circuits and Systems for Video Technology*, 35 (7): 6549-6560 [DOI: 10.1109/tcsvt.2024.3525015]
- Ren D W, Zuo W M, Hu Q H, Zhu P F and Meng D Y. 2019. Progressive image deraining networks: a better and simpler baseline//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 3932-3941 [DOI: 10.1109/cvpr.2019.00406]
- Saharia C, Ho J, Chan W, Salimans T, Fleet D J and Norouzi M. 2023. Image super-resolution via iterative refinement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4): 4713-4726 [DOI: 10.1109/TPAMI.2022.3204461]
- Su Y Y, Jia D L, Shen Y K and Wang L. 2023. Single-channel blind

- image separation based on transformer-guided GAN. *Sensors*, 23(10): #4638 [DOI: 10.3390/s23104638]
- Sun S Q, Ren W Q, Gao X W, Wang R, and Cao X C. 2024. Restoring images in adverse weather conditions via histogram transformer// *Proceedings of the 18th European Conference on Computer Vision*. Milan, Italy: Springer: 111-129 [DOI: 10.1007/978-3-031-72670-5_7]
- Sun X, Xu J D, Ma Y L, Zhao T Y, Ou S F and Peng L Z. 2022. Blind image separation based on attentional generative adversarial network. *Journal of Ambient Intelligence and Humanized Computing*, 13(3): 1397-1404 [DOI: 10.1007/s12652-020-02637-0]
- Valanarasu J M J, Yasarla R and Patel V M. 2022. TransWeather: transformer-based restoration of images degraded by adverse weather conditions//*Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans, USA: IEEE: 2353-2363 [DOI: 10.1109/cvpr52688.2022.00239]
- Wang Y J, Gong J X, Liang Z B, Chong Q P, Cheng X and Xu J D. 2025. Fuzzy diffusion model for seen-through document image restoration. *Journal of Image and Graphics*, 30(4): 1118-1129 (王义杰, 龚嘉鑫, 梁宗宝, 崇乾鹏, 程翔, 徐金东. 2025. 面向透射文档图像复原的模糊扩散模型. *中国图象图形学报*, 30(4): 1118-1129) [DOI: 10.11834/jig.240350]
- Xu J D, Yan W Q and Ou S F. 2020. *Theory and Application of Blind Image Separation*. Beijing, China: Science Press (徐金东, 阎维青, 欧世峰. 2020. 盲图像分离理论与应用. 北京: 科学出版社)
- Yin S B, Shi Y W, Wang Y B and Yang Y H. 2025. When aware haze density meets diffusion model for synthetic-to-real dehazing. *IEEE Transactions on Circuits and Systems for Video Technology*, 35(5): 4242-4255 [DOI: 10.1109/tcsvt.2024.3520816]
- Zamir S W, Arora A, Khan S, Hayat M, Khan F S and Yang M H. 2022. Restormer: efficient transformer for high-resolution image restoration//*Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans, USA: IEEE: 5718-5729 [DOI: 10.1109/cvpr52688.2022.00564]
- Zamir S W, Arora A, Khan S, Hayat M, Khan F S, Yang M H, et al. 2021. Multi-stage progressive image restoration//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville, USA: IEEE: 14816-14826 [DOI: 10.1109/cvpr46437.2021.01458]
- Zhang D M, Li J Y, Chen Z L and Zou Y T. 2024a. Efficient image generation with contour wavelet diffusion. *Computers and Graphics*, 124: #104087 [DOI: 10.1016/j.cag.2024.104087]
- Zhang Y Z, Zhang J W, Li H, Wang Z X, Hou L W, Zou D Q, et al. 2024b. Diffusion-based blind text image super-resolution//*Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, USA: IEEE: 25827-25836 [DOI: 10.1109/CVPR52733.2024.02440]

作者简介

龚嘉鑫,男,硕士研究生,主要研究方向为计算机视觉和图像复原。E-mail: 202300358064@s.ytu.edu.cn

徐金东,通信作者,男,教授,主要研究方向为图像处理、模式识别和盲源分离。E-mail: xujindong@ytu.edu.cn

孙浩钦,男,硕士研究生,主要研究方向为计算机视觉和图像处理。E-mail: 202300358042@s.ytu.edu.cn