



基于条件扩散模型的双阶段恶劣天气图像复原

潘莉文, 王骏, 施俊, 李俊诚

引用本文:

潘莉文, 王骏, 施俊, 等. 基于条件扩散模型的双阶段恶劣天气图像复原[J]. 智能系统学报, 2026, 21(4): 943-951.
PAN Liwen, WANG Jun, SHI Jun, et al. Two-stage adverse weather image restoration based on a conditional diffusion model[J]. *CAAI Transactions on Intelligent Systems*, 2026, 21(4): 943-951.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202509039>

您可能感兴趣的其他文章

利用残差密集网络的运动模糊复原方法

Image restoration with residual dense network

智能系统学报. 2021, 16(3): 442-448 <https://dx.doi.org/10.11992/tis.201912002>

多感知兴趣区域特征融合的图像识别方法

Image recognition method based on multi-perceptual interest region feature fusion

智能系统学报. 2021, 16(2): 263-270 <https://dx.doi.org/10.11992/tis.201906032>

基于孪生变分自编码器的小样本图像分类方法

A small-sample image classification method based on a Siamese variational auto-encoder

智能系统学报. 2021, 16(2): 254-262 <https://dx.doi.org/10.11992/tis.201906022>

一种基于雨线主方向自适应的全局稀疏去雨模型

A global sparse rain removal model based on rain streaks main direction adaptation

智能系统学报. 2020, 15(2): 271-280 <https://dx.doi.org/10.11992/tis.201809042>

基于改进的稀疏表示和PCNN的图像融合算法研究

Image fusion based on the improved sparse representation and PCNN

智能系统学报. 2019, 14(5): 922-928 <https://dx.doi.org/10.11992/tis.201805045>

基于自编码器的特征迁移算法

Feature transfer algorithm based on an auto-encoder

智能系统学报. 2017, 12(6): 894-898 <https://dx.doi.org/10.11992/tis.201706037>

DOI: 10.11992/tis.202509039

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20260311.1635.008>

基于条件扩散模型的双阶段恶劣天气图像复原

潘莉文¹, 王骏¹, 施俊¹, 李俊诚²

(1. 上海大学 通信与信息工程学院, 上海 200444; 2. 华东师范大学 计算机科学与技术学院, 上海 200300)

摘要: 为了解决恶劣天气下图像复原任务中普遍存在的细节模糊、纹理丢失和显存占用高等问题, 本研究提出了一种基于条件扩散模型的双阶段图像复原方法。该方法首先构建了基于双分支特征融合模块的编码器-解码器结构, 充分利用卷积操作的局部感知能力与 Transformer 架构的全局依赖建模优势, 实现对退化图像的初步重建。该阶段同时引入小波变换, 有效提取图像的多尺度频率特征, 在显著降低计算复杂度的前提下保持结构信息的完整性。将条件扩散模型嵌入编码-解码网络之间, 以前一阶段输出作为结构条件, 通过扩散过程的迭代去噪机制实现对纹理与细节的增强。该模块通过隐空间中的概率映射, 引导退化图像稳定逼近真实图像分布, 最终实现视觉感知质量显著提升的复原结果。实验结果表明, 与若干主流图像复原方法相比, 本文方法在去雨、去雪任务上均表现出优异的性能, 在各数据集上的峰值信噪比 (peak signal-to-noise ratio, PSNR) 分别提升了 0.06~1.13 dB 和 0.23~5.72 dB; 同时, 所提方法在模型参数数量与显存占用之间取得了良好的平衡。综上所述, 本研究为恶劣天气下的图像复原任务提供了一种全新、有效且可靠的解决方案。

关键词: 深度学习; 恶劣天气图像恢复; 图像复原; 扩散模型; 小波变换; 特征融合; 多尺度特征; 编码-解码结构中图分类号: TP391 文献标志码: A 文章编号: 1673-4785(2026)04-0943-09

中文引用格式: 潘莉文, 王骏, 施俊, 等. 基于条件扩散模型的双阶段恶劣天气图像复原 [J]. 智能系统学报, 2026, 21(4): 943-951.

英文引用格式: PAN Liwen, WANG Jun, SHI Jun, et al. Two-stage adverse weather image restoration based on a conditional diffusion model[J]. CAAI transactions on intelligent systems, 2026, 21(4): 943-951.

Two-stage adverse weather image restoration based on a conditional diffusion model

PAN Liwen¹, WANG Jun¹, SHI Jun¹, LI Juncheng²

(1. School of Communications and Information Engineering, Shanghai University, Shanghai 200444, China; 2. School of Computer Science and Technology, East China Normal University, Shanghai 200300, China)

Abstract: To address common challenges in image restoration under inclement weather, such as blurred details, texture loss, and high video memory usage, this study proposes a two-stage image restoration method based on a conditional diffusion model. The first stage constructs an encoder-decoder architecture utilizing a dual-branch feature fusion module. This architecture leverages the local perception capabilities of convolutional operations and the global dependency modeling advantages of the transformer architecture to achieve preliminary reconstruction of degraded images. This stage also incorporates a wavelet transform to effectively extract multiscale frequency features, preserving structural information integrity while considerably reducing computational complexity. The second stage embeds a conditional diffusion model between the encoder and decoder networks, using the output of the previous stage as a structural condition. Texture and detail are enhanced through an iterative denoising mechanism via a diffusion process. This module guides the degraded image to steadily approximate the true image distribution through probabilistic mapping in the latent space, ultimately achieving restoration results with considerably improved visual quality. Experimental results demonstrate that our method outperforms several mainstream image restoration methods in rain and snow removal tasks, improving the peak signal-to-noise ratio by 0.06~1.13 and 0.23~5.72 dB, respectively, across various datasets. The proposed method also achieves a good balance between model parameter count and graphics memory usage. In short, this research provides a novel, effective, and reliable solution for image restoration in inclement weather.

Keywords: deep learning; adverse weather image restoration; image restoration; diffusion model; wavelet transform; feature fusion; multi-scale features; encoder-decoder structure

收稿日期: 2025-09-30. 网络出版日期: 2026-03-12.

基金项目: 国家自然科学基金项目 (62301306); 上海市科技创新行动计划项目 (23ZR1422200, 23YF1412800).

通信作者: 李俊诚. E-mail: cjunchengli@gmail.com.

©《智能系统学报》编辑部版权所有

智能交通系统在恶劣天气条件下的稳定运行仍面临严峻挑战。在此类环境下采集的交通图像常出现显著的视觉退化, 严重影响语义分割^[1-2]、

目标检测^[3-4]和自动驾驶^[5-6]等下游任务的性能和可靠性。因此,从雨、雪等退化图像中恢复出高质量的清晰图像对计算机视觉任务至关重要。然而,由于光照不均、对比度下降、前景遮挡及感知频率信息丢失等问题,实现高质量图像恢复仍存在较大挑战。

大多数传统方法依赖于先验知识或特定的退化函数来实现图像恢复,但这类方法恢复出的图像质量通常较差且鲁棒性较低。随着深度学习的兴起,通过利用 CNN(convolutional neural network)和 Transformer 等深度网络架构^[7-11]来学习图像之间的映射关系,可以更为高效地实现高质量的图像复原。例如,CNN 通过卷积操作有效提取图像的局部特征,具备平移不变性,能够捕捉图像中的空间结构信息。同时,其共享权重的机制也大大减少了参数数量,降低了模型计算复杂度。但这类模型在捕获全局上下文信息时常常能力不足。Transformer 凭借注意力机制能够实现长距离依赖建模,但因高显存占用导致其在高分辨率图像中扩展性较差。近年来,扩散模型的出现为图像复原任务提供了新的思路和方法,在图像复原领域展现出强大的生成能力。该方法通过模拟数据的逐步扩散和逆扩散过程建模复杂的图像分布,生成结果细节丰富、逼真度高。然而,扩散模型的生成过程往往依赖多次迭代,计算效率低下。

受上述架构启发,同时为解决上述方法所存在的问题,本研究综合考虑以上方案的优缺点,提出一种基于条件扩散模型的双阶段恶劣天气图像复原方法。具体创新点和贡献如下:

1) 设计了小波域下的双分支编码-解码初步恢复机制。通过构建融合 CNN 与 Transformer 优势的双分支特征融合模块(dual-branch feature fusion block, DFFB),在小波域内实现退化图像的初步重建。该结构兼具局部特征保持与全局依赖建模能力,为后续精细化处理提供高质量的结构和语义先验。

2) 提出了条件扩散引导的图像精细化重建方法。在初步恢复基础上,引入条件扩散模型进行端到端联合优化,通过稳定逼近清晰图像的数据分布,逐步增强纹理细节与视觉真实性,在显著提升复原性能的同时保持模型的计算效率。

3) 所提出的方法在雨、雪等复杂退化场景下,多项客观指标均显著优于现有主流方法,同时实现了模型参数量、显存占用与复原质量之间的有效平衡,为实际智能交通系统中的视觉感知任务提供了可靠解决方案。

1 相关工作

1.1 图像恢复

近年来,恶劣天气条件下的图像恢复研究取得了长足进展,研究范式已由早期的模型驱动方法逐渐转向更加复杂的数据驱动方法。传统的模型驱动技术^[12-15]主要依赖于特定的天气先验知识来解决图像退化问题。随着深度学习技术的兴起,图像恢复在重建精度和视觉效果方面均实现了显著提升。

起初,基于 CNN 的图像复原算法实现了令人印象深刻的性能。例如,Dong 等^[16]首次将深度学习引入图像复原领域,提出了 SRCNN(super resolution convolutional neural network)模型,但该模型的重建速度较慢且重建出的图像视觉效果较差。之后,Zhang 等^[17]提出了一种改进的深度卷积神经网络用于图像复原,通过引入残差学习来提升网络性能和收敛速度。尽管如此,由于 CNN 不具备长距离建模能力,且多任务泛化差,算法存在一定的局限性。

随后,基于 Transformer 的网络架构在计算机视觉和图像恢复领域逐渐展现出优于 CNN 的建模潜力。Liang 等^[9]提出了一种基于 Swin Transformer 的图像复原网络,通过引入移位窗口注意力机制,在保持局部特征提取能力的同时,实现跨窗口交互。然而,窗口划分的限制使其在应对恶劣天气下的复杂退化场景时泛化性有限。Valanarasu 等^[18]提出的 TransWeather 网络采用全局自注意力来捕捉天气退化的整体分布模式(如大范围雨滴、雪花分布),保证全局信息捕获。但随着输入图像大小的二次扩展,该网络需要消耗大量计算资源。

1.2 基于小波变换图像恢复算法

小波变换因能同时获得图像空间与频率信息,在图像复原任务中得到广泛应用。研究人员探索了小波变换与 CNN 及 Transformer 架构的深度融合,并取得了令人鼓舞的研究成果。Liu 等^[19]提出一种用于图像复原的 MWCNN(multi-level wavelet-CNN)网络,有效利用了小波变换来进行增强处理。类似地,王维等^[20]提出了一种基于 Transformer 的多尺度小波变换去雨算法。该算法在加快模型计算速度的同时降低了 GPU 内存消耗。值得一提的是,变换前后保持了高质量的复原结果。

1.3 基于扩散模型的图像恢复算法

扩散模型因其强大的生成能力在图像复原领

域获得大量关注。Özdenizci 等^[21]提出了一种全新的基于 patch 的扩散模型图像复原方法, 在 patch 层面上进行扩散建模, 虽然缓解了大分辨率图像计算开销过大的问题, 但改善了恶劣天气条件下的视觉表现。然而, 该方法受到反向采样速度慢的限制, 需要耗费长时间进行采样和推理来生成高质量图像, 这极大限制了其在实际场景中的应用。相对以上算法, Huang 等^[22]提出了一种小波域下基于扩散模型的图像复原算法。该模型通过在小波域中学习干净图像的噪声分布, 大大减少了在空间域中常见的计算开销; 同时提出一种高效的条件采样策略, 其复原效率与传统算法相当, 并且比现有的基于扩散的模型算法快 100 倍以上。即便如此, 该模型参数量高达 124×10^6 , 在模型的轻量化设计上还需进一步优化。

因此本文借助小波变换, 一是在小波域下利用 CNN 和 Transformer 实现局部特征提取和全局

建模能力; 二是继续探索小波域下扩散模型的可行性, 通过扩散模型强大的纹理细节生成能力, 进一步实现图像的精细化恢复, 得到高质量复原结果。

2 网络架构

恶劣天气下基于频域融合与条件扩散模型图像复原的整体框架如图 1 所示。该方法采用双阶段训练策略: 第一阶段通过一个包含双分支特征融合模块 (DFFB) 的编码-解码结构对退化图像进行初步恢复, 该模块有效融合了局部特征与全局上下文信息, 训练完成后保留编码-解码器权重; 第二阶段加载并冻结第一阶段主干网络的参数, 将条件扩散模型嵌入至编码器与解码器之间进行联合训练, 利用其强大的生成能力进一步细化纹理与结构, 显著提升图像的视觉质量与真实感。

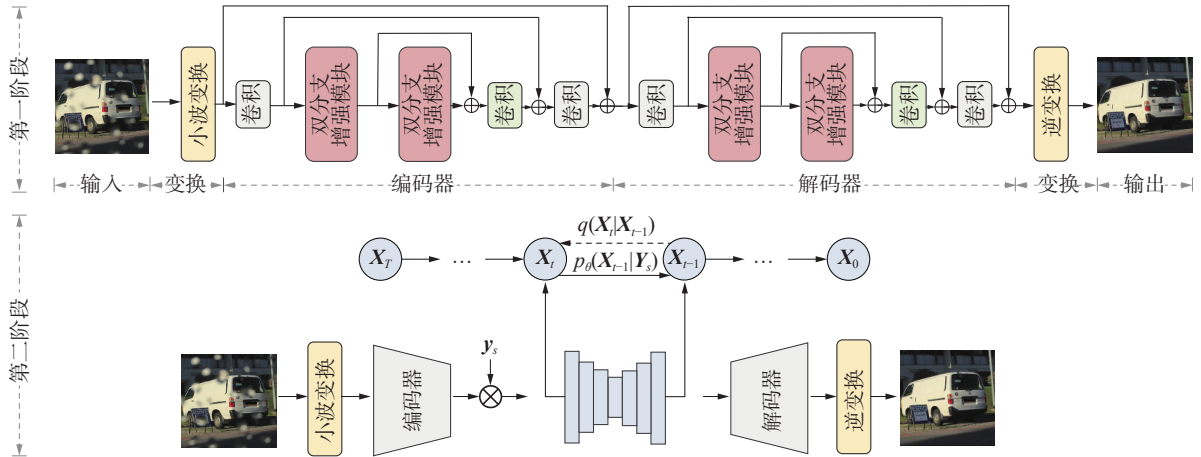


图 1 基于条件扩散模型的双阶段恶劣天气图像复原的整体框架

Fig. 1 Overall framework for two-stage adverse weather image restoration based on a conditional diffusion model

2.1 小波变换原理

小波变换是一种强大的数学工具, 能够将输入图像 $I_{in} \in \mathbf{R}^{H \times W \times C}$ 分解为 4 个子带, 从而同时捕获空间域与频域的信息。其中, 低频子带 I_{LL} 保留图像的整体结构与轮廓信息, 而高频子带 I_{LH} 、 I_{HL} 、 I_{HH} 则包含边缘与纹理等细节特征。离散小波变换 (discrete wavelet transform, DWT) 是小波变换在离散时间域上的实现形式。具体运算为

$$I_{LL}, I_{LH}, I_{HL}, I_{HH} = f_{DWT}(I_{in}) \quad (1)$$

式中: $I_{LL}, I_{LH}, I_{HL}, I_{HH} \in \mathbf{R}^{\frac{H}{2} \times \frac{W}{2} \times C}$ 为 4 个频率子带。因此, 引入 DWT 可以在不丢失原有信息的同时降低输入图像分辨率。

2.2 双分支特征融合模块

本研究提出一种应用于恶劣天气图像复原任务中的双分支特征融合模块。如图 2 所示, 该模

块有两个分支网络, 分别由残差网络和 Transformer 组成。其中残差网络主要由卷积构成, 具备强大的局部特征提取能力。残差网络的基本组成单元是残差块, 其核心特点是通过引入残差连接来缓解深层网络训练中的梯度消失问题, 同时提高网络性能。而 Transformer 网络具备强大的全局建模能力, 因此由其组成的双分支特征融合模块具备优越的局部感知力和全局建模能力。

由图 1 可知, 将小波变换得到的子图通过卷积层得到浅层特征 X_{in} 。 X_{in} 通过双分支特征融合模块得到局部特征 X_{e1} 和全局特征 X_{e2} , 将 X_{e1} 和 X_{e2} 进行相加得到融合特征 X_{out} , 具体运算为

$$\begin{cases} X_{e1} = f_{CNN}(X_{in}) \\ X_{e2} = f_{Trans}(X_{in}) \\ X_{out} = X_{e1} + X_{e2} \end{cases} \quad (2)$$

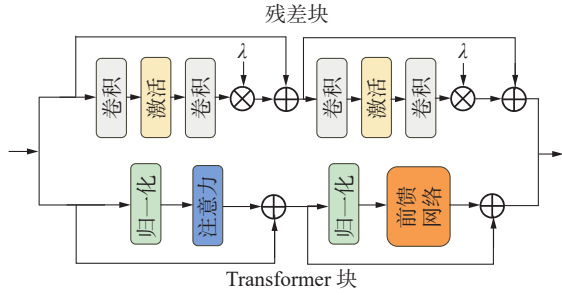


图 2 双分支特征融合模块

Fig. 2 Dual-branch feature fusion block

2.3 扩散模型原理

近年来,扩散模型作为主流生成式模型之一,具备训练稳定、生成质量高等特点,主要包含前向扩散和反向去噪过程。

2.3.1 前向过程

在前向过程中,输入数据 $x_0 \sim q(x_0)$ 通过马尔可夫链在每个时间步 t 上逐步添加噪声 $n \sim N(0,1)$,使其分布逐渐逼近标准高斯分布^[23]。其过程为

$$q(x_t|x_{t-1}) = N(x_t; \sqrt{\alpha_t}x_{t-1}, (1-\alpha_t)I) \quad (3)$$

在图像复原领域中, x_0 表示原始干净图像(通常来源于数据集中提供的无退化图像样本), x_{t-1} 和 x_t 分别表示进行了 $t-1$ 步和 t 步加噪的中间含噪图像。预设参数 $\alpha_t \in (0,1)$ 控制着每步加入噪声的强度。由高斯分布的加性性质可知,多个独立高斯随机变量合并,其结果仍然符合高斯分布。因此,我们可以通过原始干净图像 x_0 和噪声 n 得到任意时刻 t 的含噪图像 x_t , 建模过程为

$$x_t = \sqrt{\alpha_t}x_0 + \sqrt{1-\alpha_t}n \quad (4)$$

式中: $\sqrt{\alpha_t}$ 为式 (3) 从 x_0 到 x_t 递推过程中的累计比例系数, $1-\alpha_t$ 为递推过程中多步噪声合并得到的总方差。

2.3.2 反向过程

在反向过程中,借助后验分布实现逐步去噪,过程为

$$p(x_{t-1}|x_t) = N(x_{t-1}; \mu(x_t, t), \tilde{\beta}_t I) \quad (5)$$

式中: $\tilde{\beta}_t = \frac{1-\alpha_{t-1}}{1-\alpha_t}\beta_t$, β_t 为预设参数。在推理阶段的反向过程中,均值 $\mu(x_t, t)$ 是未知的。因此,我们可以通过神经网络的学习进行噪声 n_θ 估计,从而实现数据的生成或重建。其过程为

$$\mu_\theta(x_t, t) = \frac{1}{\sqrt{\alpha_t}}(x_t - \sqrt{1-\alpha_t}n_\theta(x_t, t)) \quad (6)$$

2.3.3 网络设计

基于以上原理,为实现对第 1 阶段重建结果的精细化处理,本研究引入条件扩散模型策略,在加快推理速度的同时提升重建质量。

如图 1 所示,在第 2 阶段的训练过程中,我们

在第 1 阶段训练完成的模型参数上进行微调。由双分支特征融合模块构成的编码器将退化图像输入 I_{in} 映射到小波域中,得到通道数为 12 的特征表示 s :

$$s = f_{\text{Encoder}}(I_{in}) \quad (7)$$

在反向逐步去噪过程中,将特征表示 s 作为条件信息,引导去噪过程。条件建模的反向去噪过程为

$$p_\theta(x_{t-1}|x_t, s) = N(x_{t-1}; \mu(x_t, s, t), \tilde{\beta}_t I) \quad (8)$$

具体来说,在条件引入时,我们采用通道拼接的特征融合方式:将条件信息 s 与当前含噪图像 x_t 在通道维度上进行合并,得到通道数为 24 的特征表示 x'_t , 具体运算为

$$x'_t = \text{Concat}(x_t, s) \quad (9)$$

U-Net 网络参数配置见表 1。将 x'_t 作为网络输入,得到噪声估计 n_θ , 其过程为

$$n_\theta = f_{\text{U-Net}}(x'_t, t) \quad (10)$$

此时,通过噪声估计 n_θ 完成均值 μ_θ 计算过程为

$$\mu_\theta(x_t, s, t) = \frac{1}{\sqrt{\alpha_t}}(x_t - \sqrt{1-\alpha_t}n_\theta(x_t, s, t)) \quad (11)$$

将均值 μ_θ 代入式 (8) 中即可实现反向去噪。

表 1 网络参数设置

Table 1 Network configuration

网络	参数设置	值
噪声估计网络 (U-NET)	输入通道数	24
	信道乘法器数量	{1, 2, 3, 4}
	各尺度残差单元数	2
	时间步 t 向量维度	96
	输出通道数	12

最后,将经过条件扩散模型去噪输出的特征表示 x'_t 输入编码器中,得到重建图像 I_{pred} 。具体运算为

$$I_{\text{pred}} = f_{\text{Decoder}}(x'_t) \quad (12)$$

2.4 损失函数

第 1 阶段训练采用复合损失函数,将像素级约束与感知质量指标相结合,以实现更为有效的监督。

在像素层面,利用平滑 L1 损失 $L_{\text{smooth L1}}$ 来最小化网络输出 O 与真实图像 G 之间的差异,从而保证重建结果在低层次细节上的精确性。具体公式为

$$L_{\text{smooth L1}} = \begin{cases} \frac{1}{2}E^2, & |E| < 1 \\ |E| - \frac{1}{2}, & \text{其他} \end{cases} \quad (13)$$

为了进一步提升重建图像的感知质量, 我们引入基于特征的感知损失 $L_{\text{perceptual}}$ 。该损失通过在高维特征空间中度量预测结果与真实图像的语义差异来实现监督。具体公式为

$$L_{\text{perceptual}} = L_{\text{MSE}}(\text{VGG}(O), \text{VGG}(G)) \quad (14)$$

最终的总损失函数由上述两部分损失加权组合而成, 实现了像素保真度与感知质量的综合优化。具体公式为

$$L_{\text{total}_1} = L_{\text{smooth L1}} + \lambda L_{\text{perceptual}} \quad (15)$$

第 2 阶段训练的重建过程中, 除了 $L_{\text{perceptual}}$ 和 $L_{\text{smooth L1}}$ 的引入, 我们还采用噪声损失 L_{noise} 用于约束模型预测噪声 n_{θ} 与真实噪声 n 之间的一致性, 其定义为

$$L_{\text{noise}} = \|n_{\theta} - n\|_2^2 \quad (16)$$

式中 $\|\cdot\|_2^2$ 表示 L2 范数平方。

3 实验结果与分析

3.1 数据集、评价指标和实现细节

3.1.1 数据集

为了全面评估本算法在雨、雪恶劣天气下的复原性能, 我们在多个数据集上进行了广泛的实验, 包括去雨数据集 RainDrop^[24] (1 100 对图像, 其中 861 对样本用于训练, 239 对样本用于测试) 和去雪数据集 Snow100K^[25] (50 000 对图像, 将其划分为小雪/中雪/大雪测试集, 分别为 16 611/16 588/16 801 张)。这些数据集收集自不同的场景, 如监控的十字路口、人行道和驾驶环境, 可以测试模型在不同环境条件下的泛化能力。

3.1.2 评价指标

为了定量评估模型的图像复原质量, 本文采用峰值信噪比^[26] (peak signal-to-noise ratio, PSNR)、结构相似性指数^[26] (structural similarity index, SSIM) 和 FID (Fréchet inception distance) 作为主要评价指标。其中, PSNR 衡量重建图像与参考图像在像素强度上的差异, 数值越大表示失真越小; SSIM 则从亮度、对比度和结构信息等方面综合评估两幅图像的相似性, 数值越接近 1 表明感知质量越好; FID 衡量生成图像与真实图像在高维特征空间的统计分布差异, 值越小表示生成图像越接近真实分布。

3.1.3 实现细节

本研究所提出的基于频域融合与条件扩散模型的恶劣天气图像复原方法, 采用基于 PyTorch 的深度学习框架实现, 并在 NVIDIA RTX 3090 显卡上完成所有实验。在训练阶段, 所有训练样本

都统一调整为 256×256 像素大小, 且没有使用任何图像增强手段。

在网络的编码-解码结构中, 其编码器和解码器分别由 2 个双分支特征融合模块构成。前者的融合模块中采用 6 个 Transformer 块, 后者采用 4 个 Transformer 块。第 1 阶段网络迭代训练 400 次, 初始学习率设置为 0.000 3, 当网络训练次数达到 300 次时, 学习率减半, 优化方法采用 Adam 优化器。损失函数中的 λ 设置为 0.01。第 2 阶段训练过程中, 先加载第 1 阶段的训练权重, 解锁最后两层网络参数与扩散模型进行联合训练, 初始学习率设置为 0.000 1, 迭代训练 25 000 次。

3.2 定量结果

本文在去雨、去雪这两类具有代表性且挑战性的恶劣天气图像恢复任务上开展大量对比实验。基于 Raindrop 数据集的实验结果如表 2 所示。可以看出, 与其他方法相比, 本研究方法的 PSNR 平均值和 SSIM 值分别提高了 0.06~1.13 dB 和 0.001~0.032。相比基于非扩散模型的方法有较大性能提升; 而在推理时间和模型轻量化方面, 相比基于扩散模型方法具备显著优势。

表 2 Raindrop 数据集上与现有 SOTA 方法的对比结果
Table 2 Comparison with SOTA methods on Raindrop

方法	PSNR/dB	SSIM	时间/s	显存/10 ⁶
DuRN ^[27]	31.24	0.926	0.09	4 108
CCN ^[28]	31.44	0.947	0.80	<u>3 738</u>
RainAttn ^[29]	31.44	0.926	0.41	8 335
AttnGAN ^[24]	31.59	0.917	0.63	8 797
IDT ^[30]	31.87	0.931	0.39	3 660
TransWeather ^[18]	30.96	0.923	<u>0.16</u>	4 734
MambaIR ^[31]	32.15	0.938	0.38	4 033
PatchDM ^[20]	<u>32.31</u>	0.946	301.35	22 313
WaveDM ^[22]	32.25	<u>0.948</u>	0.30	4 965
本文方法	32.37	0.949	0.65	4 388

注: 加粗数值为最优结果, 下划线数值为次优结果。

去雪任务的实验结果如表 3 所示。其中 Snow100K-S 表示小雪场景下的测试数据集, Snow100K-L 表示大雪场景下的测试数据集。可以看出, 与其他方法相比, 本研究方法在不同退化程度上的 PSNR 平均值分别提高了 0.23~4.92 dB 和 0.68~5.72 dB; SSIM 平均值分别提高了 0.000 5~0.057 2 和 0.006 3~0.105 8。综上所述, 本文方法不仅能够多种恶劣天气条件下有效提升图像细节恢复能力, 同时在保持图像结构一致性方面具有优势, 充分验证了其鲁棒性与泛化能力。

表 3 Snow100K 数据集上与现有 SOTA 方法的对比结果
Table 3 Comparison with SOTA methods on Snow100K

方法	Snow100K-S		Snow100K-L	
	PSNR/dB	SSIM	PSNR/dB	SSIM
JSTASR ^[32]	31.40	0.901 2	25.32	0.807 6
DesnowNet ^[24]	32.33	0.950 0	27.12	0.898 3
DDMSNET ^[33]	34.34	0.944 5	28.85	0.877 2
Restormer ^[10]	36.01	<u>0.957 9</u>	<u>30.36</u>	0.906 8
NAFNet ^[34]	34.79	0.949 7	30.06	0.901 7
All-in-One ^[35]	—	—	28.33	0.882 0
TransWeather ^[18]	34.34	0.934 1	29.31	0.887 9
PatchDM ^[20]	<u>36.09</u>	0.954 5	30.28	0.900 0
Chen et al. ^[36]	34.42	0.946 9	30.22	<u>0.907 1</u>
WGWS-Net ^[37]	34.31	0.946 0	30.16	0.900 7
本文方法	36.32	0.958 4	31.04	0.913 4

注: 加粗数值为最优结果, 下划线数值为次优结果。

3.3 定性结果

除了基于客观评价指标的定量对比外, 本文

还从主观视觉角度进行定性分析, 在测试集中选取图像进行对比分析。为更清晰地展示所提方法在细节特征恢复方面的能力, 本文对重建结果的局部区域进行了放大比较。

图 3 给出了多种方法在 Raindrop 测试集上的效果对比, 其中 Inf-1 是目标图像的 PSNR 和 SSIM 结果。图像中黄框部分为图像局部区域放大图。具体地, TransWeather、WeatherDiff 和 MambaIR 完全不能消除第 1 行图像中由雨滴产生的光斑, 本文方法可以较好地恢复光斑处的细节纹理; 针对第 2 行图像中雨滴的大面积遮挡, 对比方法的重建效果仍比较模糊, 而本文方法的重建效果贴近真实地面图像。

图 4 分别给出了小雪 (第 1 行) 和大雪 (第 2 行) 场景下多种对比方法的重建结果图。在小雪场景下, 各方法恢复较为良好; 但在应对更为复杂的大雪场景中, 其余方法的重建效果仍遗留不同程度的透明雪块, 本文方法重建出较为清晰轮廓。

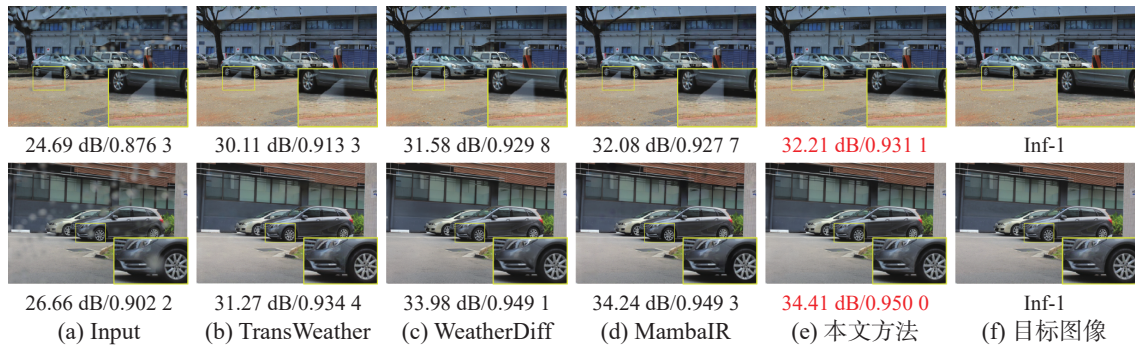


图 3 不同算法在 Raindrop 数据集的重建效果对比

Fig. 3 Contrast of effects of various algorithms on Raindrop dataset

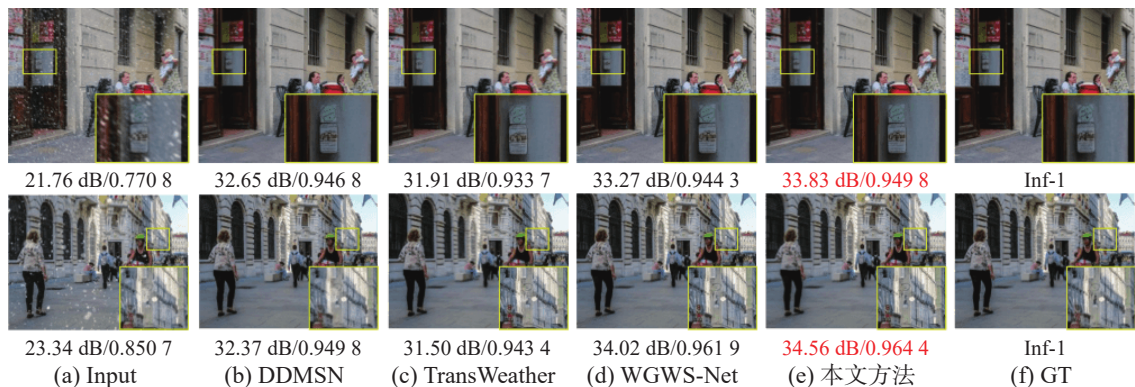


图 4 不同算法在 Snow100K 数据集的重建效果对比

Fig. 4 Contrast of effects of various algorithms on Snow100K dataset

3.4 消融实验

本研究从编码-解码器、小波变换以及条件扩散模型 3 个方面开展消融实验, 以验证其对重建质量的贡献, 如表 4 所示。

在进行消融实验时, 把由编码-解码器组成的网络作为基准模型, 即实验设置 1。将引入小波变换后的编码-解码网络作为实验设置 2。可以看出, 小波变换的引入在保持模型变换前恢复质

量的同时,大大减少了模型推理时间和 GPU 显存资源。实验设置 3 则在前期工作上继续引入条件扩散,进行图像恢复的精细化增强。结果表明,实验设置 3 的网络设计最优,不仅实现了在合理的推理时间和较小的显存资源下恢复了较高的图像质量,其 PSNR 和 SSIM 值分别提高了 0.18~0.24 dB 和 0.004~0.005;其生成图像在分布逼真度

和低频细节上也有显著改善,FID 数值从 23.52 降至 20.21。因此,本文方法一方面验证了扩散模型在频域内的可行性,为恶劣天气条件下的图像复原提供了一种有效方案;另一方面,通过实验结果表明,条件扩散模型能够在保持全局结构的同时进一步实现图像的精细化恢复。

表 4 不同组件的消融实验
Table 4 Ablation studies of different components.

实验设置	编码-解码	小波变换	扩散模型	时间/s	图像分辨率/ 像素×像素	GPU/10 ⁶	PSNR/dB	SSIM	FID
1	√			1.81	64	8 164	<u>32.19</u>	0.933	<u>23.24</u>
2	√	√		0.41	64	3 176	32.13	<u>0.934</u>	23.52
3	√	√	√	<u>0.65</u>	64	<u>4 388</u>	32.37	0.938	20.21

注:加粗数值为最优结果,下划线数值为次优结果。

3.5 模型轻量化分析

为了全面评估所提模型的效率,我们分析了模型参数量与恢复性能之间的平衡。如图 5 所示,该图比较了多个先进方法的参数量和性能指标。结果表明,本研究提出的方法在模型大小与图像重建质量之间实现了最佳平衡,以较少的模型参数量取得了最好的复原结果。

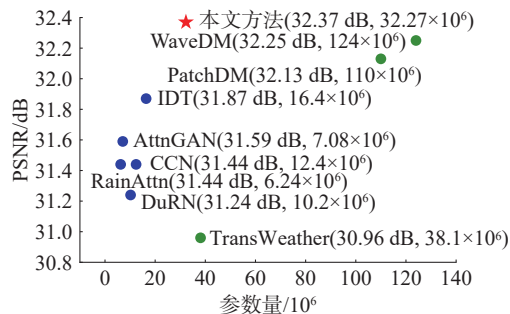


图 5 Raindrop 数据集模型对比

Fig. 5 Raindrop dataset model comparison

3.6 下游任务表现

智能交通系统中的一大关键挑战在于,如何在受恶劣天气退化影响的图像中实现对车辆和行人的准确识别。为验证本文方法的实用性,本文采用先进的视觉模型进行全面评估。其中,利用在 COCO 数据集上预训练的 Detectron2^[38],对去天气处理在目标检测性能上的提升进行了定量分析。如图 6 所示,检测框中的数据为置信度,即模型认为预测结果正确的概率,置信度越高则概率越大。第 1 行为 Raindrop 数据集下的目标检测结果,在输入图和经过 Transformer 方法恢复的图中,检测模型都错误地把雨滴遮挡处识别为 car,且置信度高达 50% 和 71%;而经由本文方法恢复的图像中,检测模型没有发生错误识别。第 2 行

为 Snow100K 数据集下的目标检测结果,未经任何方法恢复的输入图像无法检测出任意船只;同时,经过本文方法恢复的图像中检测的船只数量远多于经 TransWeather 方法恢复的图像。因此,本文方法能够有效去除交通图像中的天气退化效应,同时保留关键目标信息,为后续的识别任务提供重要支持。

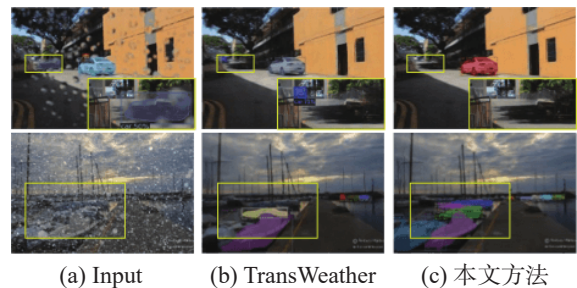


图 6 基于 Detectron2 的检测性能评估

Fig. 6 Detectron2-based detection evaluation

4 结束语

本文提出了一种面向恶劣天气条件的图像复原方法,该方法基于扩散模型与频域融合机制,并在小波域中实现高性能的图像恢复。所提出的网络结构主要包括两个核心组成部分:1)编码-解码器架构,其在小波域下嵌入了双分支特征融合模块(DFFB),融合 CNN 与 Transformer 的优势,兼具局部特征提取与全局依赖建模能力,用于实现图像的初步重建;2)条件扩散模型,用于第 2 阶段的精细化重建。该阶段以初步重建结果作为条件输入,充分利用其中所包含的纹理、结构和光照信息,有针对性地增强退化严重区域。借助该机制,退化图像能够稳定地逼近清晰图像的数据分布,从而显著提升重建质量。为验证所

提方法的有效性,本文选取了8~10种代表性图像复原算法,在两种典型恶劣天气场景下进行对比实验。从多项客观评价指标与视觉对比结果来看,本方法均优于现有对比模型,表明所构建的算法在保持较低复杂度的同时,具有更优异的复原性能。然而,该方法在实际复杂退化场景中仍存在一定局限性,针对退化类型多样且程度不一致的真实环境,其泛化能力仍有待进一步优化与提升。

参考文献:

- [1] ZHAO Hengshuang, SHI Jianping, QI Xiaojuan, et al. Pyramid scene parsing network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 6230–6239.
- [2] 谭睿俊, 赵志诚, 谢新林. 双层残差语义分割网络及交通场景应用[J]. 智能系统学报, 2022, 17(4): 780–787.
TAN Ruijun, ZHAO Zhicheng, XIE Xinlin. Double-residual semantic segmentation network and traffic scenic application[J]. CAAI transactions on intelligent systems, 2022, 17(4): 780–787.
- [3] 陆军, 鲁林超, 翟晓阳, 等. 面向道路交通场景的高效3D目标检测[J]. 智能系统学报, 2025, 20(1): 91–100.
LU Jun, LU Linchao, ZHAI Xiaoyang, et al. High-efficiency 3D object detection for road traffic scenes[J]. CAAI transactions on intelligent systems, 2025, 20(1): 91–100.
- [4] 贾亮, 戚丽瑾, 林铭文, 等. 改进的YOLOv7轻量化目标检测模型[J]. 计算机工程与设计, 2025, 46(8): 2350–2357.
JIA Liang, QI Lijin, LIN Mingwen, et al. Improved YOLOv7 lightweight target detection model[J]. Computer engineering and design, 2025, 46(8): 2350–2357.
- [5] HU Yihan, YANG Jiazhi, CHEN Li, et al. Planning-oriented autonomous driving[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 17853–17862.
- [6] 吕家威, 吕丹璇. 多模态交互技术在自动驾驶舱环境中的应用[J]. 汽车电器, 2025(9): 13–15.
LÜ Jiawei, LÜ Danxuan. Application of multimodal interaction technology in automatic cockpit environment[J]. Auto electric parts, 2025(9): 13–15.
- [7] 王美华, 柯凡晖, 梁云, 等. 融合3D注意力和Transformer的图像去雨网络[J]. 中国图象图形学报, 2022, 27(5): 1509–1521.
WANG Meihua, KE Fanhui, LIANG Yun, et al. 3D attention and Transformer based single image deraining network[J]. Journal of image and graphics, 2022, 27(5): 1509–1521.
- [8] JIANG Kui, WANG Zhongyuan, YI Peng, et al. Multi-scale progressive fusion network for single image deraining[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 8343–8352.
- [9] LIANG Jingyun, CAO Jiezhong, SUN Guolei, et al. SwinIR: image restoration using swin transformer[C]//2021 IEEE/CVF International Conference on Computer Vision Workshops. Montreal: IEEE, 2021: 1833–1844.
- [10] ZAMIR S W, ARORA A, KHAN S, et al. Restormer: efficient transformer for high-resolution image restoration[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 5718–5729.
- [11] WANG Zhendong, CUN Xiaodong, BAO Jianmin, et al. Uformer: a general U-shaped transformer for image restoration[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 17662–17672.
- [12] HE Kaiming, SUN Jian, TANG Xiaoou. Single image haze removal using dark channel prior[J]. IEEE transactions on pattern analysis and machine intelligence, 2011, 33(12): 2341–2353.
- [13] 王平, 张云峰, 包芳勋, 等. 基于雾天图像降质模型的优化去雾方法[J]. 中国图象图形学报, 2018, 23(4): 605–616.
WANG Ping, ZHANG Yunfeng, BAO Fangxun, et al. Multivariate optimized haze removal method based on hazy image degradation model[J]. Journal of image and graphics, 2018, 23(4): 605–616.
- [14] ZHU Lei, FU C W, LISCHINSKI D, et al. Joint bi-layer optimization for single-image rain streak removal[C]//2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 2545–2553.
- [15] YOU Jing, JIA Shaocheng, PEI Xin, et al. DMRVisNet: deep multihead regression network for pixel-wise visibility estimation under foggy weather[J]. IEEE transactions on intelligent transportation systems, 2022, 23(11): 22354–22366.
- [16] DONG Chao, LOY C C, HE Kaiming, et al. Learning a deep convolutional network for image super-resolution[C]//Computer Vision-ECCV 2014. Cham: Springer, 2014: 184–199.
- [17] ZHANG Kai, ZUO Wangmeng, CHEN Yunjin, et al. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising[J]. IEEE transactions on image processing, 2017, 26(7): 3142–3155.
- [18] VALANARASU J M J, YASARLA R, PATEL V M. Transweather: Transformer based restoration of images degraded by adverse weather conditions[C]//2022 IEEE Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 2353–2363.
- [19] LIU Pengju, ZHANG Hongzhi, ZHANG Kai, et al. Multi-level wavelet-CNN for image restoration[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Boston: IEEE, 2018: 886–88609.
- [20] 王维, 王习博, 廖万正, 等. 基于多尺度小波变换的雨天列车外观图像恢复[J/OL]. 北京: 自动化技术与应用, 2025. (2025-08-21)[2025-09-21]. <https://link.cnki.net/urlid/23.1474.tp.20250821.1548.024>.
WANG Wei, WANG Xibo, LIAO Wanzheng, et al. Rainy train appearance image restoration based on mul-

- tiscale wavelet transform[J/OL]. Automation technology and application, 2025. (2025-08-21)[2025-09-21]. <https://link.cnki.net/urlid/23.1474.tp.20250821.1548.024>.
- [21] ÖZDENIZCI O, LEGENSTEIN R. Restoring vision in adverse weather conditions with patch-based denoising diffusion models[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2023, 45(8): 10346–10357.
- [22] HUANG Yi, HUANG Jiancheng, LIU Jianzhuang, et al. WaveDM: wavelet-based diffusion models for image restoration[J]. *IEEE transactions on multimedia*, 2024, 26: 7058–7073.
- [23] 管凤旭, 张涵宇, 路斯棋, 等. 扩散模型在计算机视觉领域的研究现状[J]. *智能系统学报*, 2025, 20(2): 265–282. GUAN Fengxu, ZHANG Hanyu, LU Siqi, et al. Research status of diffusion models in computer vision[J]. *CAAI transactions on intelligent systems*, 2025, 20(2): 265–282.
- [24] QIAN Rui, TAN R T, YANG Wenhan, et al. Attentive generative adversarial network for raindrop removal from a single image[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 2482–2491.
- [25] LIU Yunfu, JAW D W, HUANG S C, et al. DesnowNet: context-aware deep network for snow removal[J]. *IEEE transactions on image processing*, 2018, 27(6): 3064–3073.
- [26] HORÉ A, ZIOU D. Image quality metrics: PSNR vs. SSIM[C]//2010 20th International Conference on Pattern Recognition. Istanbul: IEEE, 2010: 2366–2369.
- [27] LIU Xing, SUGANUMA M, SUN Zhun, et al. Dual residual networks leveraging the potential of paired operations for image restoration[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2020: 7000–7009.
- [28] QUAN Ruijie, YU Xin, LIANG Yuanzhi, et al. Removing raindrops and rain streaks in one go[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 9143–9152.
- [29] QUAN Yuhui, DENG Shijie, CHEN Yixin, et al. Deep learning for seeing through window with raindrops [C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2020: 2463–2471.
- [30] XIAO Jie, FU Xueyang, LIU Aiping, et al. Image de-raining transformer[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2023, 45(11): 12978–12995.
- [31] GUO Hang, LI Jinmin, DAI Tao, et al. MambaIR: a simple baseline for Image restoration with State-space model[C]//Computer Vision-ECCV 2024. Cham: Springer, 2025: 222–241.
- [32] CHEN Weiting, FANG Haoyu, DING J J, et al. JSTASR: joint size and transparency-aware snow removal algorithm based on modified partial convolution and veiling effect removal[C]//Computer Vision-ECCV 2020. Cham: Springer, 2020: 754–770.
- [33] ZHANG Kaihao, LI Rongqing, YU Yanjiang, et al. Deep dense multi-scale network for snow removal using semantic and depth priors[J]. *IEEE transactions on image processing*, 2021, 30: 7419–7431.
- [34] CHU Xiaojie, CHEN Liangyu, YU Wenqing. NAFSSR: stereo image super-resolution using NAFNet[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. New Orleans: IEEE, 2022: 1238–1247.
- [35] LI Ruoteng, TAN R T, CHEONG L F. All in one bad weather removal using architectural search[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 3172–3182.
- [36] CHEN Weiting, HUANG Zhikai, TSAI C C, et al. Learning multiple adverse weather removal via two-stage knowledge learning and multi-contrastive regularization: toward a unified model[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 17632–17641.
- [37] ZHU Yurui, WANG Tianyu, FU Xueyang, et al. Learning weather-general and weather-specific features for image restoration under multiple adverse weather conditions[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 21747–21758.
- [38] Wu Y, Kirillov A, Massa F, et al. Detectron2[J/OL]. Facebook AI research, 2019. (2019-10-28)[2025-09-25]. <https://github.com/facebookresearch/detectron2>.

作者简介:



潘莉文, 硕士研究生, 主要研究方向为恶劣天气图像复原。E-mail: 15779365007@163.com。



王骏, 教授, 博士, 主要研究方向为机器学习、人工智能和医学影像智能处理。获江苏省高校科研成果自然科学一等奖、中国产学研创新成果一等奖。发表学术论文 160 余篇, 谷歌学术引用 6000 余次。E-mail: wangjunshu@shu.edu.cn。



李俊诚, 副教授, 博士, 主要研究方向为底层视觉和医学影像智能处理。上海市青年科技英才扬帆计划获得者, 获中国产学研合作创新成果一等奖。发表学术论文 70 余篇, 谷歌学术引用 5000 余次。E-mail: cvjunchengli@gmail.com。

[责任编辑: 孟玮]