文章编号：1003－1251(2024)06－0028－06

基于注意力机制及多尺度特征融合的图像去雨

宋建辉，胡强强，刘晓阳，赵亚威

（沈阳理工大学 自动化与电气工程学院，沈阳 110159）

摘　要：针对去除不同雨纹的同时恢复图像背景细节的问题，提出一种基于注意力机制及多尺度特征融合的图像去雨方法。该网络采用双分支结构，分别用于雨纹去除和背景恢复。雨纹提取模块采用跨空间学习的多尺度注意力机制，通过多尺度上下文信息捕捉、均值计算、权重计算和整体信息综合，帮助改善雨纹去除任务中的图像质量，提高去雨效果。背景恢复模块包括多尺度特征提取部分和特征融合部分，采用多个扩张卷积层，每个卷积层具有不同的扩张因子，以扩大感受野，提取多尺度的图像背景特征；使用大核卷积对提取的多尺度特征信息进行融合调整，从而更准确地进行背景恢复。在多个公开数据集上的实验结果表明：所提方法能够有效去除真实雨图像场景中的雨纹，同时可以更好地恢复图像背景的细节信息。

关　键　词：双分支去雨；多尺度特征融合；注意力机制；扩张卷积

中图分类号：TP391.4　　**文献标志码：**A　**DOI：**10.3969/j.issn.1003－1251.2024.06.005

Image Deraining Based on Attention Mechanism and Multi-scale Feature Fusion

SONG Jianhui，HU Qiangqiang，LIU Xiaoyang，ZHAO Yawei

（Shenyang Ligong University，Shenyang 110159，China）

Abstract：To restore background details while removing different rain patterns，an image rain removal method based on attention mechanism and multi-scale feature fusion is proposed. The network adopts a two-branch structure，which is used for rain stripe removal and background recovery respectively. The rain-stripe extraction module uses a multi-scale attention mechanism of cross-space learning to help improve the image quality in the rain-stripe removal task and improve the rain removal effect through multi-scale context information capture，mean calculation，weight calculation and overall information synthesis. The background recovery module includes multi-scale feature extraction part and feature fusion part，and adopts multiple extended convolution layers，each with different expansion factors，to enlarge the receptive field and extract multi-scale image background features. Large nuclear convolution is used to fuse and adjust the extracted multi-scale feature information，so as to recover the background more accurately. The experimental results on several public data sets show that the proposed method can effectively remove the rain lines in the real rain image scene，and can better recover the details of the image background.

Key words：double branch rain removal；multi-scale feature fusion；attention mechanism；extended convolution

收稿日期：2023－12－11

基金项目：辽宁省教育厅高等学校基本科研项目(LJKZ0275)；沈阳市中青年科技创新人才支持计划项目(RC210247)

作者简介：宋建辉(1981—)，女，教授，博士。

图像去雨是计算机视觉领域的重要任务,主要分为视频图像去雨和单幅图像去雨两大类。视频图像去雨算法通过利用连续图像帧之间的时序信息,能够更有效地分析雨图与背景之间的差异,进而去除雨滴纹理[1-3]。相比之下,单幅图像去雨需要将雨线从背景中分离出来,因其不具备时序信息的帮助,所以相较于视频图像去雨难度更大。

在单幅图像去雨领域,算法主要分为基于传统模型的方法和基于深度学习的方法两大类。传统的图像去雨方法通常利用字典学习[4]、稀疏编码[5]、低秩表示[6]等基于图像先验知识的方法,为有效地进行图像去雨,用户必须多次调整参数以获得满意的结果,降低了实际场景中去除中大雨的效率。针对深度学习方法,Luo 等[7]提出一种循环注意力蒸馏网络,通过综合考虑局部和全局特征提高处理各种复杂背景下雨纹的能力。Wang 等[8]提出一种循环尺度引导网络,通过不同尺度的特征信息相互引导提高去雨性能。Deng 等[9]提出一种上下文聚合细节恢复网络,利用扩张卷积聚合块恢复背景细节。Li 等[10]提出一种多尺度上下文信息和注意力网络,通过跳跃连接传递不同层次的特征信息,重构干净的背景图像。

虽然基于深度学习的方法相对于传统的方法提高了单幅图像去雨的整体性能,但雨痕在图像中具有不同的尺度和方向,多尺度雨痕特征提取不足,会导致背景层上存在少量的雨纹。为此,本文提出一种基于注意力机制及多尺度特征融合的图像去雨网络。为了有效感知不同尺度和方向的雨纹,提出一种改进的雨纹提取模块,采用跨空间学习的多尺度注意力机制[11],不仅有助于捕捉更丰富的上下文信息,且结合了均值计算和权重计算,能够较好地融合不同位置的信息,减轻雨纹的影响。为了在去除雨纹的同时有效保留图像细节,本文引入多尺度特征融合模块,其结构包括全局特征提取和多个扩张卷积层。最后,使用大核卷积[12]操作进行特征融合和调整特征幅度,以找回丢失的背景细节,提升去雨效果。

1　本文算法

本文提出一种基于注意力机制及多尺度特征融合的图像去雨网络,其网络结构如图 1 所示。该网络引入雨纹提取和背景恢复两个并行子网络,每个子网络都经历三个主要阶段:对输入图像进行浅层特征提取(CP);特征经过一系列子模块的处理(雨纹提取模块、背景恢复模块)以进一步突出特征中的重要信息;将处理的结果映射回图像空间以生成最终的输出图像(CBC)。

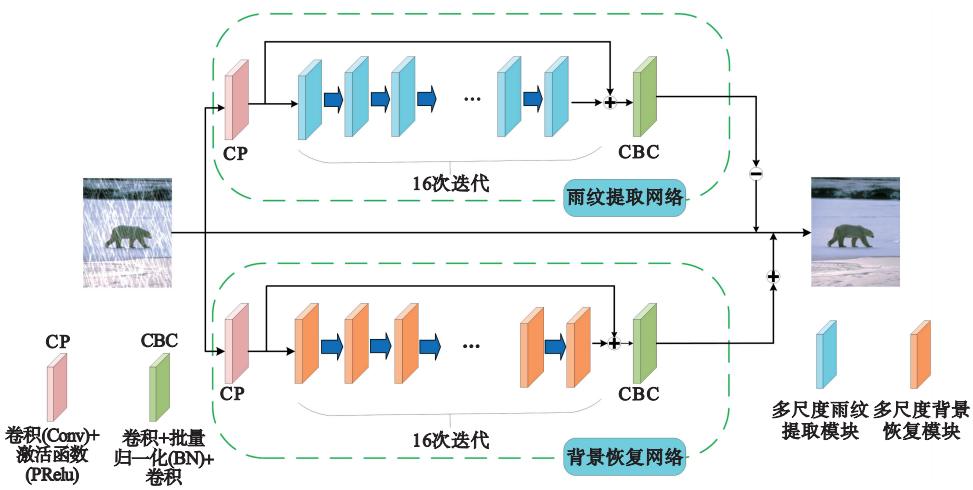


图 1　网络结构

Fig. 1　Network structure

雨纹提取网络通过训练一个雨纹提取子函数从雨天图像中提取雨纹信息;背景恢复网络则专注于细节恢复,通过训练一个背景恢复子函数找回去雨过程中丢失的图像细节,两者协同去除雨纹并恢复去雨造成的丢失细节。

1.1　多尺度雨纹提取模块

文献[9]提出基于通道注意力的图像去雨网络,给模型带来比较显著的性能提升,但是通道注意力通常会忽略对注意力图的位置信息。文献[13]提出基于通道和空间注意力的图像去雨网

络,降低通道数并采用大尺寸卷积以更好地利用位置信息。然而,卷积操作通常仅限于捕获局部相关性,对于图像去雨任务中可能涉及的长程依赖存在一定局限性。文献[14]在图像去雨中引入了坐标注意力(CA),将精确的位置信息嵌入到通道中,同时在空间上捕捉了长距离的相互作用,由此显著提高了图像去雨性能。尽管 CA 在捕获局部和长程信息方面表现出色,但会忽略整个图像空间位置之间的相互作用。

本文多尺度雨纹提取模块如图 2 所示。通过引入基于跨空间学习的高效多尺度注意力机制以提取更多的上下文信息。

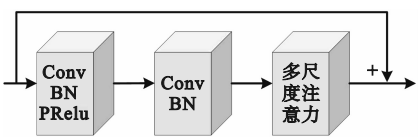


图 2 多尺度雨纹提取模块

Fig. 2 Multi-scale rain pattern extraction module

跨空间学习的高效多尺度注意力机制如图 3 所示。图中:$\boldsymbol{X}$ 表示输入特征张量;$\boldsymbol{X}_{1\times1}$、$\boldsymbol{X}_{3\times3}$ 表示 $\boldsymbol{X}$ 的两个分支;$\boldsymbol{X}_{\mathrm{h}}$、$\boldsymbol{X}_{\mathrm{w}}$ 分别表示在高度、宽度维度上进行平均池化、特征融合和降维处理后的结果;$\boldsymbol{S}_1$、$\boldsymbol{S}_2$ 表示两个不同的空间注意力权重。将经过浅层提取的特征采用分组机制,通过三条平行路径(两条平行路径在 1×1 分支上,第三条路径在 3×3 分支上)以提取不同方向的空间注意力权重(沿一个空间方向捕获远程依赖关系,沿另一空间方向保留精确的位置信息),计算表达式如式(1)~(3)所示。最后,将两个注意力权重融合得到最后的注意力图,突出更多的全局上下文并增强关注对象的表示,表达式如式(4)所示。

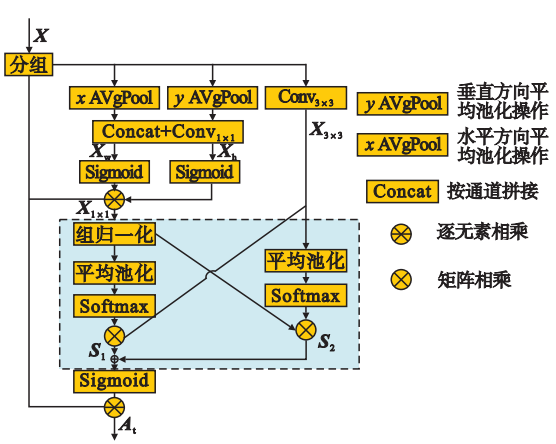


图 3 多尺度跨空间注意力机制

Fig. 3 Multi-scale cross-space attention mechanism

$$\boldsymbol{X}_{1\times1}=\boldsymbol{X}\times\sigma(\boldsymbol{X}_{\mathrm{h}})\times\sigma(\boldsymbol{X}_{\mathrm{w}}^{\mathrm{T}}) \tag{1}$$

$$\boldsymbol{S}_1=s_{\mathrm{f}}\mathrm{Pool}_{\mathrm{avg}}(\boldsymbol{X}_{3\times3})\times(G_{\mathrm{N}}(\boldsymbol{X}_{1\times1}))^{\mathrm{T}} \tag{2}$$

$$\boldsymbol{S}_2=s_{\mathrm{f}}\mathrm{Pool}_{\mathrm{avg}}(G_{\mathrm{N}}(\boldsymbol{X}_{1\times1}))\times(\boldsymbol{X}_{3\times3})^{\mathrm{T}} \tag{3}$$

$$\boldsymbol{A}_{\mathrm{t}}=\sigma((\boldsymbol{X}\times\mathrm{Concat}(\boldsymbol{S}_1,\boldsymbol{S}_2))) \tag{4}$$

式中:G_{N} 表示组归一化;σ 表示 Sigmoid 激活函数;s_{f} 表示 Softmax 激活函数;$\boldsymbol{A}_{\mathrm{t}}$ 表示注意力图;$\mathrm{Pool}_{\mathrm{avg}}$ 表示平均池化;$(\cdot)^{\mathrm{T}}$ 表示维度调整。

1.2 多尺度背景恢复模块

受到文献[9]的启发,本文提出了多尺度背景恢复模块,结构如图 4 所示。先通过多尺度特征提取模块,经过全局特征提取和三种不同扩张率的空洞卷积获得更多的上下文信息,再经过后续的特征融合模块中大核卷积的处理完成背景细节的恢复。

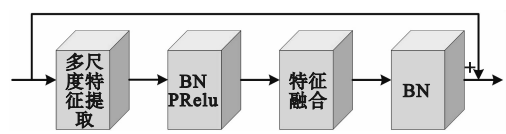


图 4 多尺度背景恢复模块

Fig. 4 Multi-scale background repairing module

多尺度特征提取模块如图 5 所示,图中:r 表示扩张率;$\boldsymbol{I}_{\mathrm{m}}$ 表示提取的全局特征;$\boldsymbol{I}_{\mathrm{s}}$ 表示 $\boldsymbol{I}_{\mathrm{m}}$ 经过注意力机制 SimAm 处理后的结果;C 为特征通道数;H 为图像垂直维度大小;W 为水平维度大小。

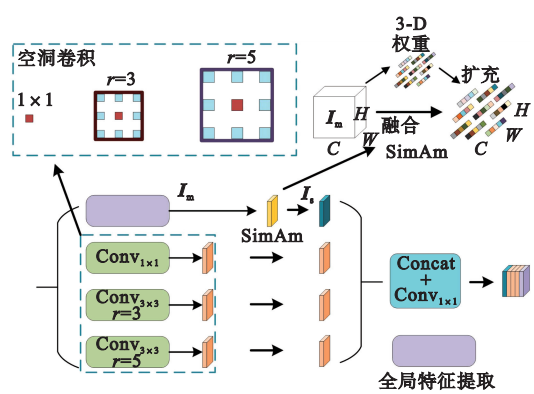


图 5 多尺度特征提取模块

Fig. 5 Multi-scale feature extraction module

首先通过全局平均池化和卷积操作对输入特征图进行全局特征提取,并增强特征信息,最后通过插值将处理后的全局特征融合回原始特征图,以丰富局部特征,表达式如式(5)所示。与此同时,引入无参数的注意力机制 SimAm,使局部特征更具区分性,提高局部细节的感知和表达能力,并使用三种不同扩张率的空洞卷积捕获图像不同尺度的局部信息。最后,将经过增强后的全局特

征和不同尺度的局部特征进行融合,以生成最终的输出特征图,如式(6)所示。

$$\boldsymbol{I}_{\mathrm{m}} = in(\mathrm{Conv}_{1\times1}(\mathrm{Pool}_{\mathrm{avg}}(\boldsymbol{i}))) \tag{5}$$

$$\boldsymbol{MS} = \mathrm{Conv}_{1\times1}(\mathrm{Concat}[\boldsymbol{I}_{\mathrm{s}}, \mathrm{Conv}_{r\times r}]) \quad (r=1,3,5) \tag{6}$$

式中:$\boldsymbol{i}$ 表示输入特征;in 表示上采样操作;$\boldsymbol{MS}$ 表示最后的特征图。

将获取到的多尺度特征信息输入到特征融合模块,如图 6 所示,进行多尺度特征的融合调整并增强特征表示能力。图中:k 表示卷积核的大小;d 表示膨胀率;Norm 表示层归一化;scale 表示缩放调整;$\boldsymbol{x}_0$ 表示作为原始特征的一半输入特征;$\boldsymbol{x}_1$、$\boldsymbol{x}_2$ 表示用作增强处理的另一半输入特征;$\boldsymbol{q}$ 表示 $\boldsymbol{x}_1$、$\boldsymbol{x}_2$ 经过增强处理后拼接的结果。

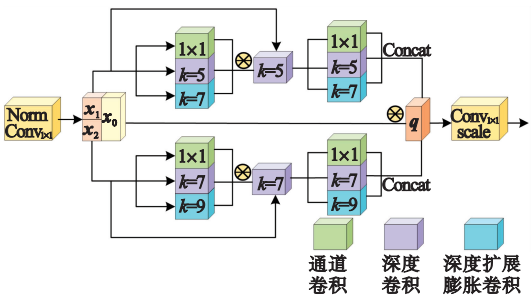


图 6　特征融合处理模块

Fig. 6　Feature fusion processing module

首先将处理过的输入特征标准化并分成两部分:一部分作为原始特征($\boldsymbol{x}_0$),另一部分($\boldsymbol{x}_1$、$\boldsymbol{x}_2$)用作增强处理。增强部分采用多种卷积操作,包括通道卷积、深度卷积(分别为 $k=5$ 和 $k=7$)和深度扩展膨胀卷积(分别为 $k=7$、$d=3$ 和 $k=9$、$d=4$),同时处理以捕捉不同尺度的上下文信息,之后再次通过深度卷积进一步提取特征。最后,将原始特征($\boldsymbol{x}_0$)和增强后的特征($\boldsymbol{q}$)融合,生成最终用于背景恢复的处理的特征。

2　实验结果与分析

2.1　数据集与训练配置

本文所有实验均在 Win10 系统下完成,使用两个英伟达显卡 GeForce GTX1080,所提模型在 Pycharm 开发平台上基于 PyTorch 框架实现。使用 Rain200L[15]、Rain200H[15] 和 Rain800[16] 三个合成数据集验证训练网络,其中 Rain800 包含 700 张训练图像和 100 张测试图像;Rain200L 和 Rain200H 都由 1 800 张训练图像和 200 张测试图像组成。实验时,使用 Adam 优化器进行模型参数的优化,初始学习率为 0.001,训练迭代次数设置为 100,在第 30、50、80 次迭代时,学习率乘以 0.2 进行衰减,一个批次大小设置为 8。

评价去雨算法的客观指标主要包括结构相似性(SSIM)和峰值信噪比(PSNR)。其中 PSNR 用来评估经过处理后图像的品质,可以量化去雨算法处理后的图像与原始图像之间的差异,PSNR 值越高,表示图像质量越好。

2.2　对比实验

不同算法在合成数据集 Rain200L、Rain200H、Rain800 上的对比结果以及在 Rain200H 上运行一张 481×321 大小的图片所需的平均时间如表 1 所示(最优值加粗表示,次优值下划线表示)。

实验结果表明,本文所提算法在 Rain200H、Rain200L 和 Rain800 数据集上的各项评价指标均达到最优或次优值,仅个别次实验的 SSIM 值略低于 DRD,处于次优值外,其他指标均优于其他算法。

表 1　不同算法合成数据集上评价指标

Table 1　Evaluation indexes on data sets with different algorithms

算法	Rain200L		Rain200H		Rain800		平均时间/s
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
DRD[9]	37.15	**0.98**	<u>28.16</u>	**0.92**	<u>26.32</u>	**0.90**	0.54
NLEDN[17]	<u>37.53</u>	**0.98**	27.31	0.89	–	–	0.36
JORDER[18]	29.16	0.95	24.37	0.86	21.13	0.81	0.28
DID-MDN[18]	23.23	0.82	17.35	0.52	17.03	0.50	0.53
RESCAN[19]	37.07	**0.98**	26.60	0.89	24.09	<u>0.84</u>	<u>0.25</u>
PReNet[19]	36.69	**0.98**	27.89	0.89	23.31	0.80	**0.15**
本文算法	**37.69**	**0.98**	**28.57**	<u>0.91</u>	**27.06**	<u>0.84</u>	0.78

DRD 采用了两个扩张卷积进行重复循环,强调多次特征增强,旨在捕获更多的特征信息。本文强调全局和局部特征的整合,采用了多尺度注意力机制以及全局和局部特征相结合的方法,并使用像素级注意力机制增强局部特征表达,提高网络对雨天图像中细微变化的感知能力。虽然整体结构相似性稍显减弱,表现为较低的 SSIM,但在像素级别上更好地还原了输入图像,从而呈现出更高的 PSNR。

本文选取了 Rain200H 中 6 张照片的处理结果进行主观对比,如图 7 所示。从图中可以看出:DID-MDN 算法去除雨纹的效果不够理想,图片仍然存在大量雨纹;PReNet 和 RESCAN 算法仍有小部分雨纹残留且存在明显的背景损失,如天空域的黑色阴影以及桥体的细节丢失;NLEDN 算法的雨纹去除效果较好,但会出现小范围的背景模糊,如小女孩发色出现白条、熊的背部条纹模糊;本文算法可以有效去除雨纹的同时保留图像背景信息,具有较好的视觉效果。为进一步验证本文所提方法的鲁棒性,选取 Rain200L 中 5 张照片的处理结果进行主观对比,如图 8 所示。从图中可以看出:DID-MDN、PReNet 和 RESCAN 算法仍有少部分的雨纹残留;NLEDN 出现背景的损失,如老虎后面的树枝颜色发生改变、羊的背部出现白色条纹;本文方法对于雨纹的去除及背景细节保持效果良好。

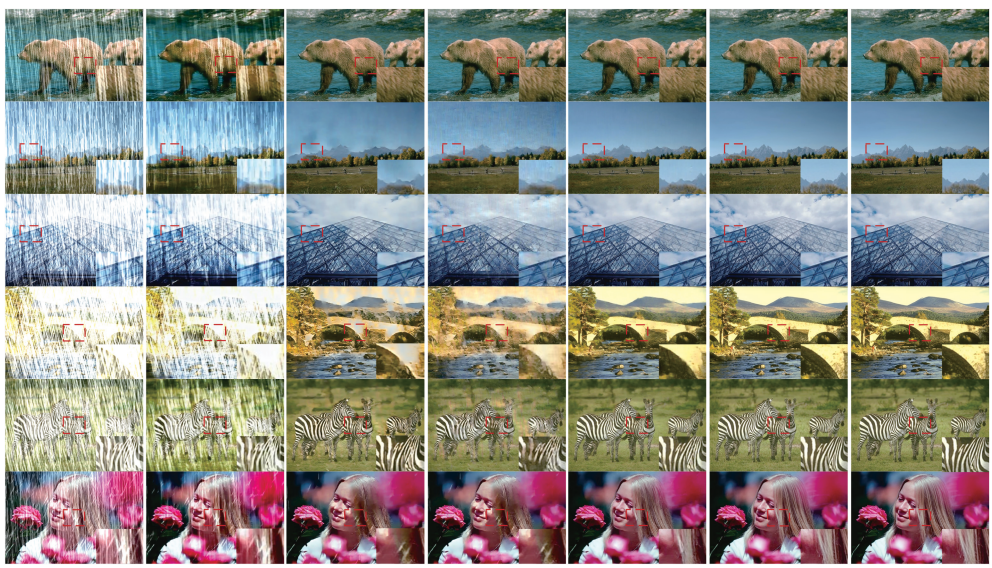


图 7 在合成数据集 Rain200H 上的实验结果

Fig. 7 Experimental results on the dataset Rain200H

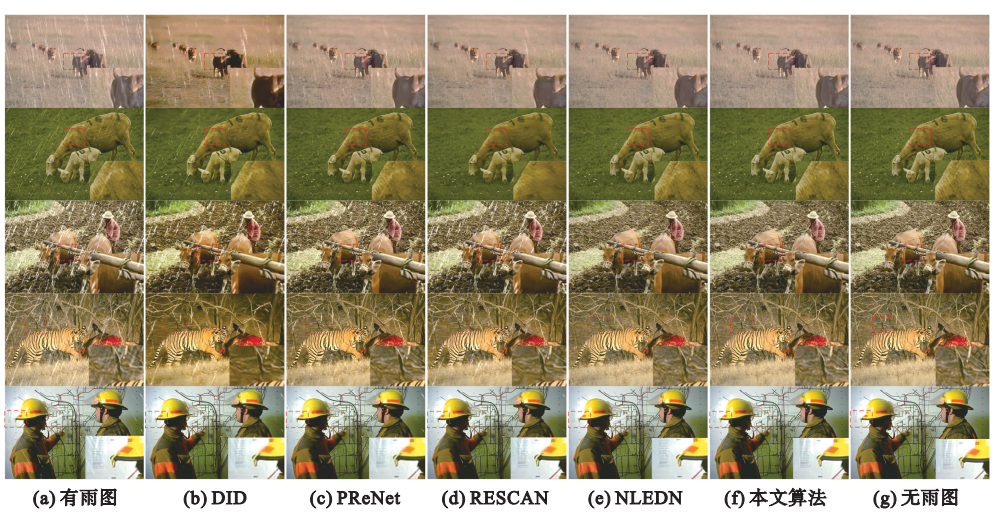


图 8 在合成数据集 Rain200L 上的实验结果

Fig. 8 Experimental results on the dataset Rain200L

为实现更出色的去雨效果,本文采用了多尺度特征提取和整合处理的策略,引入了复杂的子模块结构。尽管算法的运行时间相对较长,但算法在图像质量上取得了显著的提升,表明本设计在性能方面取得了更有价值的成果,为图像去雨任务提供了更为高效的解决方案。

3　结论

本文提出了一种基于注意力机制及多尺度特征融合的图像去雨方法,通过双分支网络结构融合了注意力机制和多尺度特征信息,协同去除雨纹并恢复去雨造成的细节丢失。雨纹提取网络的设计目的是通过对多尺度特征信息的捕获,增强网络对图像中不同区域的关注程度。背景恢复网络的设计旨在平衡全局和局部特征的感知,以全面处理雨天图像。这种平衡性的设计使网络既能捕捉到整体场景的结构,又能对局部细节有针对性地进行处理,从而提高去雨效果并恢复损失的背景细节。

参考文献(References):

[1] SU R, ZHANG L P, ZHANG Y M, et al. Complex scene video frames alignment and multi-frame fusion deraining with deep neural network[J]. Neural Computing and Applications, 2023, 35(7): 5369 – 5380.

[2] ZHANG K H, LI D X, LUO W H, et al. Enhanced spatio-temporal interaction learning for video deraining: faster and better [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(1): 1287 – 1293.

[3] YANG W H, TAN R T, FENG J S, et al. Recurrent multi-frame deraining: combining physics guidance and adversarial learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(11): 8569 – 8586.

[4] WANG Y L, LIU S C, CHEN C, et al. A hierarchical approach for rain or snow removing in a single color image[J]. IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society, 2017, 26(8): 3936 – 3950.

[5] KANG L W, LIN C W, FU Y H. Automatic single-image-based rain streaks removal via image decomposition[J]. IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society, 2012, 21(4): 1742 – 1755.

[6] GUO X L, XIE X Y, LIU G C, et al. Robust low-rank subspace segmentation with finite mixture noise[J]. Pattern Recognition, 2019, 93(C): 55 – 67.

[7] LUO Y, HUANG Q D, LING J, et al. Local and global knowledge distillation with direction-enhanced contrastive learning for single-image deraining[J]. Knowledge-Based Systems, 2023, 268: 110480.

[8] WANG C, ZHU H H, FAN W S, et al. Single image rain removal using recurrent scale-guide networks[J]. Neurocomputing, 2022, 467(C): 242 – 255.

[9] DENG S, WEI M Q, WANG J, et al. Detail-recovery image deraining via context aggregation networks[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA: IEEE, 2020: 14548 – 14557.

[10] LI P, GAI S. Single image deraining using multi-scales context information and attention network[J]. Journal of Visual Communication and Image Representation, 2023, 90: 103695.

[11] OUYANG D L, HE S, ZHANG G Z, et al. Efficient multi-scale attention module with cross-spatial learning[C]//ICASSP 2023 – 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Rhodes Island, Greece: IEEE, 2023: 1 – 5.

[12] GUO M H, LU C Z, LIU Z N, et al. Visual attention network [J]. Computational Visual Media, 2023, 9(4): 733 – 752.

[13] 张学锋,李金晶. 基于双注意力残差循环单幅图像去雨集成网络[J]. 软件学报,2021,32(10):3283 – 3292.
ZHANG X F, LI J J. Single image de-raining using a recurrent dual-attention-residual ensemble network[J]. Journal of Software, 2021, 32(10): 3283 – 3292. (in Chinese)

[14] 刘忠洋,周杰,陆加新,等. 基于注意力机制的多尺度特征融合图像去雨方法[J]. 南京信息工程大学学报(自然科学版),2023,15(5):505 – 513.
LIU Z Y, ZHOU J, LU J X, et al. Image rain removal via multi-scale feature fusion based on attention mechanism[J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2023, 15(5): 505 – 513. (in Chinese)

[15] YANG W H, TAN R T, FENG J S, et al. Deep joint rain detection and removal from a single image[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017: 1685 – 1694.

[16] ZHANG H, SINDAGI V, PATEL V M. Image de-raining using a conditional generative adversarial network[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 30(11): 3943 – 3956.

[17] CUI X, WANG C, REN D W, et al. Semi-supervised image deraining using knowledge distillation[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(12): 8327 – 8341.

[18] TAN Y J, WEN Q, QIN J, et al. Coupled rain streak and background estimation via separable element-wise attention[J]. IEEE Access, 2020, 8: 16627 – 16636.

[19] 杨爱萍,张腾飞,王朝臣,等. 基于方向引导的残差去雨网络[J]. 天津大学学报(自然科学与工程技术版),2023,56(4):391 – 399.
YANG A P, ZHANG T F, WANG C C, et al. Residual deraining network based on the direction guidance[J]. Journal of Tianjin University (Science and Technology), 2023, 56(4): 391 – 399. (in Chinese)

(责任编辑:和晓军)