胡梦月, 贾东立, 张恒. 融合多尺度膨胀注意力机制的图像去雨算法[J]. 智能计算机与应用,2026,16(6):181-186. DOI:10.20169/j.issn.2095-2163.24090904

融合多尺度膨胀注意力机制的图像去雨算法

胡梦月, 贾东立, 张 恒
(河北工程大学 信息与电气工程学院, 河北 邯郸 056038)

摘 要: 雨天图像由于雨线的影响导致图像背景信息模糊,影响后续计算机任务的执行。针对这一问题,提出了一种融合多尺度膨胀注意力机制的去雨网络(Multi-Scale Dilated Attention Derain Network,MSDA-DNet)。该网络采用多阶段架构,逐步学习有雨图像和无雨图像的映射过程。研究中,提出的多尺度残差块,利用膨胀卷积来扩大感受野,在不增加模型复杂度的情况下,优化雨线检测性能。此外,在网络主干部分引入注意力机制和空间金字塔池化(SPPF),使得网络更有效地学习全局信息和局部信息,减少信息损失。经过实验验证,该算法在主客观评价指标上均优于其他经典及新颖算法。

关键词: 图像去雨; 多阶段架构; 多尺度残差块; 注意力机制; 金字塔池化
中图分类号: TP391.41 **文献标志码**: A **文章编号**: 2095-2163(2026)06-0181-06

Image deraining algorithm with multi-scale expansion attention mechanism

HU Mengyue, JIA Dongli, ZHANG Heng
(School of Information and Electrical Engineering, Hebei University of Engineering, Handan 056038, Hebei, China)

Abstract: The background information of rainy images is blurred due to the influence of rain lines, which affects the execution of subsequent computer tasks. To solve this problem, Multi-Scale Dilated Attention Derain Network (MSDA-DNet) is proposed. The network adopts a multi-stage architecture and gradually learns the mapping process of images with rain and images without rain. During the process, a multi-scale residual block is proposed, which uses expansion convolution to enlarge the receptive field, and optimizes the performance of rainline detection without increasing the complexity of the model. In addition, the introduction of attention mechanism and Spatial Pyramid Pooling Fast(SPPF) in the backbone of the network makes the network learn global and local information more effectively and reduce information loss. The experimental results show that the proposed algorithm outperforms other classical and novel algorithms in both subjective and objective evaluation indexes.

Key words: image deraining; multistage architecture; multi-scale residual block; attention mechanism; Spatial Pyramid Pooling

0 引 言

雨天是一种常见的天气现象。但当雨天出现时,不仅影响人眼视觉判断,还降低图像质量,对计算机后续进行图像分类[1-2]、目标检测[3-4]、视频监控[5-6]和语义分割[7-8]等高级视觉任务造成负面影响。因此,开发高效快速的雨水去除算法是十分必要的。雨图的数学模型可以被建模为:

$$O = B + R \tag{1}$$

其中,O 表示雨天图像;B 表示干净的背景图;R 表示雨层。图像去雨就是要从雨图 O 中把雨层 R 去除,恢复出干净的图像 B。

尽管现有去雨算法已经取得了一定进展,但在处理有较多雨线场景时,仍难以有效保留背景图像的细节。针对这一挑战,提出了一种轻量级的去雨方法,融合了多尺度扩张注意力机制,这种计算方法能够更好地保留图像的纹理细节信息,同时有效去除雨条纹。本次研究主要贡献在于:

(1) 提出了一种新颖的融合多尺度膨胀注意力机制的图像去雨方法,该方法能够充分利用图像中的多种特征信息,通过不同权重特征之间的相互作用来增强去雨效果。

(2) 在网络主干部分引入空间金字塔池化(SPPF),使得网络能够更有效地学习图像中的全局

作者简介: 胡梦月(1999—),女,硕士研究生。主要研究方向:图像处理。
通信作者: 贾东立(1972—),男,博士,教授,硕士生导师。主要研究方向:图像处理技术,智能通信信号处理。Email:jwdsli@163.com。

收稿日期: 2024-09-09

信息和局部信息,有效减少去雨过程中的信息损失,保留更多背景细节。

(3)提出了一个更加轻量的去雨网络,将递归结构和长短期记忆网络相结合,多阶段地逐步去雨,使用了更少的参数量来有效去除雨痕。并进行了广泛的比较,所提网络在结构相似性指数(SSIM)和峰值信噪比(PSNR)两个评价标准上均超越了最新的几种方法。

1 相关工作

1.1 基于先验的去雨算法

传统的图像去雨算法通过人工标注图像中的雨条纹的特征信息或者使用各种先验知识假设雨痕的大小、方向等进行建模,通过该模型进行雨水的去除,Kang 等学者[9]创新性地将去雨任务转化为形态成分分析的框架内,通过滤波器技术将图像分解为多个频率成分,并利用字典学习方法专门处理高频部分,以消除其中的雨痕。Li 等学者[10]则进一步引入了高斯混合模型,该模型能够自适应地学习雨水条纹的具体方向和尺度特征,从而在去雨效果上实现了显著的改进。

1.2 基于深度学习的去雨算法

近年来,深度学习技术在图像去雨领域展现出了强大的潜力和显著优势。Fu 等学者[11]率先探索了基于卷积神经网络的去雨方法,通过优化网络架构,特别是调整映射策略以缩小学习范围,从而简化了网络训练过程。此外,该深度网络设计尤为注重捕捉图像中的高频细节,这一策略有效提升了雨水去除的效果,为图像去雨任务带来了新的突破。Yang 等学者[12]为了更好地去除各种程度的雨纹,开发了一个多任务学习架构,学习二进制雨痕图来定位雨水区域,利用区域上下文信息循环地检测降雨量,迭代地去除雨条纹。Jiang 等学者[13]考虑到雨痕的模糊程度和分辨率会根据图片的深度而有所不同,提出了一种端到端的多尺度渐进融合网络(MSPFN),采用循环计算不断获取全局纹理,并且构建多尺度金字塔结构,融合空间维度和不同尺度的协同表示,有效进行雨纹去除。Ren 等学者[14]提出了一种更为精简的基线网络模型,通过引入级内递归机制显著降低了网络参数的数量。反复展开浅层残差网络(ResNet),并充分利用不同阶段深层特征之间的互补性,实现了雨水条纹的渐进式去除。Jiang 等学者[15]提出了一种渐进耦合网络(PCNet),使用深度可分离卷积减少模型参数,并且通过多个 CRM 级联,提取多尺度多层次的雨纹特征,逐步分离雨纹和背景内容。

2 本文方法

本节将详细介绍本文所提出的融合多尺度膨胀注意力机制的图像去雨网络(MSDA-DNet)。

2.1 网络总体架构

渐进式多尺度膨胀注意力网络总体框架如图 1 所示。该网络有 6 个阶段,渐进式地去除雨痕,常见的多阶段去雨网络通过堆叠子网络来实现,导致网络的参数增加和过拟合,研究中在阶段间利用递归计算[11,14,16],共享网络参数。在每个阶段的开始有一个卷积层和 ReLU 激活函数,输出一个 32 通道的特征图,随后紧跟一个 SPPF 模块,实现局部特征和全局特征的融合,接着使用 LSTM 有效地传递和表达长时间序列中的信息避免梯度消失/爆炸问题。在每个阶段中都有一个由 5 个多尺度注意力残差块组成的浅层 ResNet,为不同的通道分配不同的加权参数,从而使得有用特征传递到下一阶段。

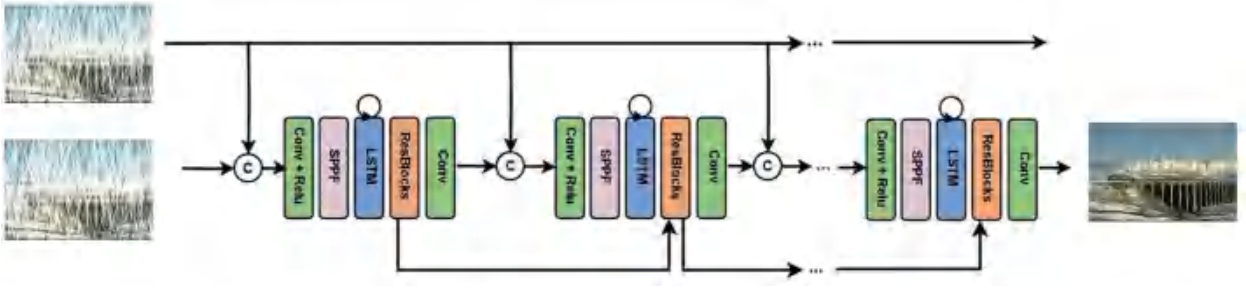


图 1 多尺度膨胀注意力机制单幅图像去雨集成网络结构图

Fig. 1 Single image deraining integrated network structure of multi-scale dilated attention mechanism

2.2 SPPF 模块

空间金字塔池化(SPPF)采用 3 个 5×5 的 MaxPool 串联的方式来达到不同大小的池化操作,其结构如图 2 所示。该算法首先使用 1×1 卷积降

维，然后依次使用 3 个 5×5 池化层进行最大池化操作。在特征融合阶段，采用 Concat 方法将各层经过最大池化处理的特征图无缝拼接，随后利用 1×1 卷积核进行特征精炼，这一过程不仅融合了特征的局部细节与全局视野，还显著提升了模型的泛化能力。此策略有效减轻了雨痕密集带来的图像背景模糊与纹理细节丢失问题，确保了图像恢复后的清晰度。

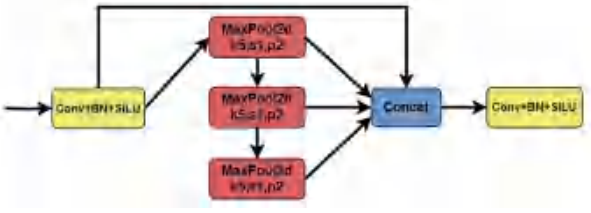


图 2　SPPF 结构图

Fig. 2　Structure diagram of SPPF

2.3　多尺度膨胀注意力残差块

普通卷积构建的去雨网络模型在提取雨条纹特征时易导致图像失真，损失图像的关键信息，扩大卷积核尺寸获取更大范围信息时易导致参数量增加，降低网络效率。由此，本文设计了一个多尺度膨胀注意力残差组模块，该模块由 5 个残差块串联而成，每个残差块都是由 2 个空洞卷积（Dilated Convolution）、激活函数（ReLU）和 SE 注意力组成。SE 注意力机制结构如图 3 所示。通过引入压缩操作和激励操作来自适应地学习每个通道的权重。随后利用全局平均池化将每一个通道的特征图压缩为一个特征向量，用 $Z \in C^R$ 来表示，通道 C 的压缩特征用 Z_c 来表示，公式为：

$$Z_c = F_{sq}(u_c) = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{W} u_c(i,j) \qquad (2)$$

随后，对 Z 进行激励处理，提取 C 个通道的关键权重，从而实现对特征通道信息的选择性增强或抑制。最后，将得到的权重序列与原始输入特征图相乘，对特征通道进行重新加权。

PReNet 网络原始的残差块中只采用单一尺寸的普通卷积提取特征无法很好地提取多个尺度的雨纹特征。因此，本文提出了一个多尺度膨胀注意力残差块，将 PReNet 算法残差块中的普通卷积换成空洞率分别为 1，2，5，1，2 的空洞卷积，使用空洞卷积能有效解决其他通过池化或者大步长的卷积来提取大尺寸的特征图所导致的特征图分辨率降低的问题，帮助模型捕捉到全局和局部之间的关系，提取不同尺度的边缘信息。并且在每个 Resblock 中引入 SE 注意力机制，为不同的通道特征分配不同的权重，增强关键特征，降低非关键特征的冗余，改进后的残差块如图 4 所示。图 4（a）是原始的残差块结构，图 4（b）是改进之后的结构，DConv 表示膨胀卷积。

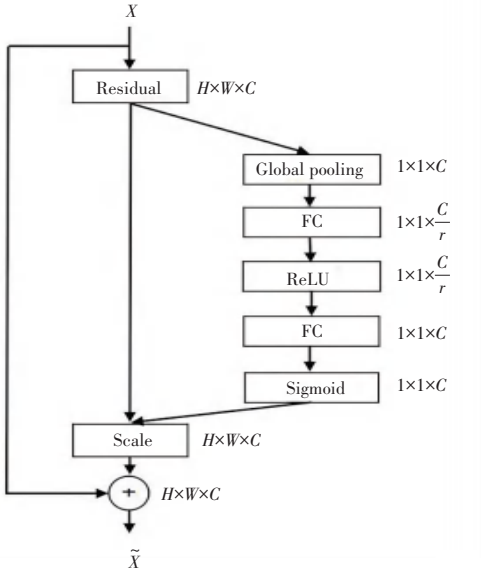
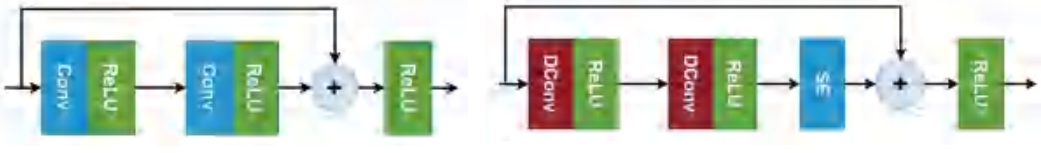


图 3　SE 注意力机制的具体结构

Fig. 3　Specific structure of SE attention mechanism



（a）原始残差块结构　　（b）改进后结构

图 4　原始残差块和改进后的残差块

Fig. 4　Original residual block and improved residual block

2.4　损失函数

图像去雨任务中广泛使用均方误差（Mean Square Error，MSE）来量化模型预测与真实无雨图像之间的差异。然而由于 MSE 中平方项的过度惩罚，通常造成高频纹理过度平滑或者模糊，不利于去雨痕后恢复清晰的图像细节。为了获得更多细节内容，本文使用结构相似度（Structural SIMilarity，SSIM）作为损失函数。SSIM 计算去雨后的输出图像

与无雨的标签图像之间的相似性,并将 SSIM 的负值作为网络的损失函数。对此可以表示为:

$$L = -\sum_{S=1}^{6}[\mathrm{SSIM}(X_S, Y)] \tag{3}$$

其中,S 表示网络中的阶段数;X_S 表示网络经过第 S 个阶段去雨之后的图像;Y 为无雨的标签图像。

3 实验与结果

3.1 数据集与训练配置

本文所提出的去雨网络基于 Pytorch 实现,使用 NVIDIA GeForce RTX 2080Ti GPU 来训练网络模型,训练图像的像素为 100×100,模型的初始学习率为 0.001,训练迭代 100 次,在迭代次数为第 30、50、80 次时,学习率以乘 0.2 的速度减小。Batchsize 设置为 16。采用最常用的评价指标峰值信噪比(PSNR)[17]和结构相似度(SSIM)[18]来衡量图像质量。所有实验都是在 RainTrain100H[12]数据集上进行的,该数据集包含 1 800 张训练图像。本模型在 Rain100H[14],Rain100L[11]和 Rain12[10]合成雨图数据集上进行了性能比较。其中,Rain100L 专注于小雨场景,仅包含单一形态的雨条纹。Rain100H 则针对大雨环境,包含了 5 种不同方向的雨痕。Rain12 仅包含 12 张测试图像被作为测试集。同时将本文算法与最新的几种算法 RESCAN[19]、PreNet[14]、PCNet[15]和 SAPNet[17]进行了比较。同时,该模型也在真实雨景图像上进行了测试,验证了其在实际应用中的效果。

3.2 仿真雨像数据集的比较

本文的算法在合成雨图的数据集上得到的 PSNR 和 SSIM 的定量结果见表 1。由表 1 可以看出 MSDA-DNet 在 Rain100H、Rain100L 和 Rain12 数据集上的 PSNR 和 SSIM 都达到了最佳值。为了更直观地验证本文所提方法的性能,本文还随机选取 Rain100H 数据集中的 4 张测试结果对去雨效果进行了评价。测试结果如图 5 所示,可以看出 PCNet 的效果最差,RESNet 和 PRENet 雨线残留明显,而 SAPNet 对于背景的细节恢复较模糊。

表 1 各算法指标值对比表

Table 1 Comparison table of index values of each algorithm

Dataset	PSNR/SSIM				
	RESCAN	PRENet	PCNet	SAPNet	本文模型
Rain100H[14]	27.858 8/0.876 7	28.042 4/0.881 6	19.880 6/0.545 1	28.941 7/0.888 4	**29.017 1/0.889 9**
Rain100L[11]	32.907 0/0.961 1	33.695 9/0.965 5	23.271 4/0.739 8	34.345 2/0.969 6	**35.043 4/0.969 8**
Rain12[10]	34.361 9/0.962 9	34.857 8/0.963 8	24.001 7/0.721 9	35.176 1/0.965 4	**35.859 3/0.966 5**



(a) Input (b) PCNet[15] (c) PreNet[14] (d) RESNet (e) SAPNet[17] (f) The proposed (g) Groundtruth

图 5 Rain100H 数据集上的测试结果

Fig. 5 Test results on the Rain100H dataset

3.3 真实雨像实验与分析

为了进一步验证 MSDA-DNet 的泛化能力,实际雨图数据集测试结果如图 6 所示,本文提出的去雨算法在视觉呈现上显著优于其他前沿方法,展现

了更优的除雨效果。RESCAN[13]、PreNet[14]、PCNet[15]和 SAPNet[17]不能很好地去除雨纹。相比之下，MSDA-DNet 可以去除更多不同形状的雨纹并恢复更干净的背景信息。



(a) Input　(b) PreNet[14]　(c) SAPNet[17]　(d) PCNet[15]　(e) RESNet　(f) The proposed

图 6　真实数据集上的测试结果

Fig. 6　Test results on real-world dataset

3.4　消融实验与分析

3.4.1　残差块数量

分别训练了不同残差块数量下 PSNR 和 SSIM 的值，对比结果见表 2。增加残差块的数量可以提高模型的性能，但当残差块的数量达到一定程度时，模型的性能提升不再明显，所以这里设置 MSDA-DNet 中残差块的数量为 5。

3.4.2　网络组成消融实验

为了验证 MSDA-DNet 算法不同模块去雨的有效性，表 3 列出了添加不同模块后的 PSNR 和 SSIM 值：只使用普通残差块(M0)、加入 SPPF(M1)、多尺度特征提取模块(M2)和注意力机制(本文模型)。可以观察到每个分量都有助于雨水条纹的去除。

表 2　不同数量的残差块的 PSNR/SSIM 对比

Table 2　Comparison of PSNR/SSIM for different number of residual groups

残差组数 N	PSNR/dB	SSIM
2	34.547 8	0.957 7
3	34.896 6	0.962 4
4	35.012 1	0.968 7
5	**35.043 4**	**0.969 8**
6	35.042 2	0.969 6

表 3　融合学习不同组合模块

Table 3　Ablation study using different modules combinations

Dataset	PSNR/SSIM			
	M0	M1	M2	本文模型
Rain100H[14]	28.042 4/0.881 6	28.493 0/0.886 2	28.691 3/0.887 5	**29.017 1/0.889 9**
Rain100L[11]	33.695 9/0.965 5	34.641 7/0.969 2	34.644 4/0.969 5	**35.043 4/0.969 8**
Rain12[10]	34.857 8/0.963 8	35.347 1/0.965 8	35.547 3/0.964 7	**35.859 3/0.966 5**

4　结束语

本文提出了一种融合多尺度膨胀注意力的图像去雨网络(MSDA-DNet)，该网络采用递归计算，运用卷积长短期记忆网络和特征提取模块来提取特征，引入多尺度空洞注意力机制，在不增加模型参数数量的同时，充分提取不同尺度的特征信息，通过注意力机制来映射雨线层的权重信息，有效地提高雨痕的去除能力。实验结果表明，相比其他算法，本算法在高效去除雨痕的同时，保留更多纹理细节，拥有更少的参数量。然而由于数据集的局限性，该算法在应对含有雨滴或雨雾的图像时，其性能会受到影响。为此，未来的研究将着重于收集更为真实的数据集或合成更接近实际情况的数据集。

参考文献

[1] 叶乡凤, 董张玉, 杨学志. 基于扩张图卷积网络的 SAR 图像分类[J]. 智能计算机与应用, 2022, 12(7): 69-73.

[2] 周雨航, 车明亮, 王晓文, 等. 图像分类方法对比评价研究[J]. 智能计算机与应用, 2021, 11(11): 43-48.

[3] 赵景波, 杜保帅. 基于深度学习的小目标检测技术发展[J]. 电光与控制, 2023, 30(2): 63-70.

[4] 车启谣, 严运兵. 基于改进 YOLOv3 的行人检测研究[J]. 智能计算机与应用, 2022, 12(8): 8-13.

[5] SREENU G, DURAI S A M. Intelligent video surveillance: A review through deep learning techniques for crowd analysis[J]. Journal of Big Data, 2019, 6: 48.

[6] ELHOSENY M. Multi-Object Detection and Tracking(MODT) machine learning model for real-time video surveillance systems[J]. Circuits, Systems, and Signal Processing, 2020, 39(2): 611-630.

[7] MO Yujian, WU Yan, YANG Xinneng, et al. Review the state-of-the-art technologies of semantic segmentation based on deep learning[J]. Neurocomputing, 2022, 493: 626-646.

[8] 王瑞绅, 宋公飞, 王明. 引入 ECA 注意力机制的 U-Net 语义分割[J]. 电光与控制, 2023, 30(1): 92-96.

[9] KANG Liwei, LIN Chiawen, FU Yuhsiang. 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.

[10] LI Yu, GUO Xiaojie, LU Jiangbo, et al. Rain streak removal using layer priors[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ:IEEE, 2016: 2736-2744.

[11] FU Xueyang, HUANG Jiabin, ZENG Delu, et al. Removing rain from single images via a deep detail network[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2017: 1715-1723.

[12] YANG Wenhan, TAN R T, FENG Jiashi, et al. Deep joint rain detection and removal from a single image[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2017: 1357-1366.

[13] JIANG Kui, WANG Zhongyuan, YI Peng, et al. Multi-scale progressive fusion network for single image deraining[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2020: 8343-8352.

[14] REN Dongwei, ZUO Wangmeng, HU Qinghua, et al. Progressive image deraining networks: A better and simpler baseline[C]//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2019: 3932-3941.

[15] JIANG Kui, WANG Zhongyuan, YI Peng, et al. Rain-free and residue hand-in-hand: A progressive coupled network for real-time image deraining[J]. IEEE Transactions on Image Processing: A publication of the IEEE Signal Processing Society, 2021, 30: 7404-7418.

[16] FAN Zhiwen, WU Huafeng, FU Xueyang, et al. Residual-guide network for single image deraining[C]//Proceedings of the 26th ACM International Conference on Multimedia. New York: ACM, 2018: 1751-1759.

[17] ZHENG Shen, LU Changjie, WU Yuxiong, et al. SAPNet: Segmentation-aware progressive network for perceptual contrastive deraining[C]//Proceedings of 2022 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW). Piscataway, NJ: IEEE, 2022: 52-62.

[18] WANG Zhuo, CONRAD A B, RAHIM H S, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE transactions on image processing: A publication of the IEEE Signal Processing Society, 2004, 13(4): 600-612.

[19] LI Xia, WU Jianlong, LIN Zhuochen, et al. Recurrent squeeze-and-excitation con-text aggregation net for single image deraining[C]//Proceedings of the European Conference on Computer Vision (ECCV). Cham: Springer, 2018: 254-269.