

刘忠洋¹ 周杰¹ 陆加新² 缪则林¹ 邵根富³ 江凯强¹ 高伟¹

基于注意力机制的多尺度特征融合图像去雨方法

摘要

雨纹分布和形状具有多样性,现有去雨算法在去雨的同时会产生图像背景模糊、泛化性能差等问题.因此,本文提出一种基于注意力机制的多尺度特征融合图像去雨方法.特征提取阶段由多个包含两个多尺度注意力残差块的残差组构成,多尺度注意力残差块利用多尺度特征提取模块提取及聚合不同尺度的特征信息,并通过坐标注意力进一步提高网络的特征提取能力.在组内进行局部特征融合,组间利用全局特征融合注意力模块更好地融合不同层次的特征,通过像素注意力使网络重点关注于雨纹区域.在仿真和真实雨像数据集上与其他现有图像去雨算法相比,本文方法的定量指标有着明显提高,去雨后的图像视觉效果较好且具有良好的泛化性.

关键词

图像去雨;多尺度;特征融合;残差网络;坐标注意力

中图分类号 TP391.4

文献标志码 A

收稿日期 2022-07-18

资助项目 国家自然科学基金(61971167,62101275,62101274);江苏省信息与通信工程优势学科建设项目

作者简介

刘忠洋,男,硕士生,研究方向为图像处理、计算机视觉.20201249541@nuist.edu.cn

周杰(通信作者),男,博士,教授,研究方向为无线通信理论、移动通信理论.zhoujie45@hotmail.com

1 南京信息工程大学 人工智能学院(未来技术学院),南京,210044

2 南京信息工程大学 电子与信息工程学院,南京,210044

3 杭州电子科技大学 自动化学院,杭州,310018

0 引言

雨纹会严重影响图像拍摄质量,含有雨纹的图像往往存在图像模糊、图像对比度和能见度降低、目标遮挡等问题,严重阻碍诸如目标检测、目标跟踪和图像识别等计算机视觉任务的执行.因此,图像雨纹的去除研究近年来在计算机视觉领域引起了极大的关注.现有的图像去雨算法包括视频去雨算法和单幅图像去雨算法.视频去雨算法可通过连续图像帧进行雨纹的检测与去除,而单幅图像去雨算法缺乏时序信息,所以单幅图像去雨更具有挑战性.

现有的基于单幅图像的去雨算法可以分为基于图像先验的算法和基于深度学习的算法.Kang 等^[1]利用双边滤波将图像分解为高频部分和低频部分并在高频部分中应用字典学习去除雨纹.Li 等^[2]使用高斯混合模型作为先验,将雨纹层从背景图像中分离出来.由于雨纹分布密度不均、形状大小不一,所以基于图像先验的算法无法有效捕捉并去除雨纹.近年来,由于卷积神经网络在计算机视觉领域取得成功,研究人员提出许多基于 CNN 的图像去雨方法.Fu 等^[3]提出一种 DerainNet 网络,利用滤波器将图像分解成高频层与低频层并将高频层送入网络进行训练,随后对网络输出和低频层进行图像增强并将其融合得到最终的去雨图像.Ren 等^[4]提出一种渐进式去雨网络 PReNet,每阶段由残差网络与循环层构成,阶段的结果与原始雨图的拼接作为下一阶段的输入.Fu 等^[5]设计出一种基于轻量级金字塔去雨网络 LPNet,该网络引入拉普拉斯金字塔结合递归残差结构实现图像去雨.Wang 等^[6]提出一种基于深度跨尺度融合结构的单幅图像去雨网络 DCSFN,通过跨尺度方式融合不同的子网并使用内尺度连接方式实现雨纹去除.Yi 等^[7]提出一种使用残差通道先验引导的去雨网络 SPDNet,并通过迭代制导的方式直接生成清晰的无雨图像.

虽然基于深度学习方法的去雨性能方面有了显著改进,但部分网络通过增加深度或者利用多分支结构来增强网络性能,使得结构过于复杂并导致网络难以训练.现有方法并未充分利用多尺度信息,对像素以及通道信息也没有有效利用,去雨后的图像存在局部细节模糊现象且不能很好地保留图像的纹理结构.有鉴于此,本文利用全局特征融合结构可以有效融合不同层次特征的特点,提出一种基于注意力机制的多尺度特征融合图像去雨网络(Multi-scale Feature Fusion Attention Network, MFFAN).本文的主要工作如下:

1) 提出一种端到端的结合注意力机制的多尺度特征融合去雨网络,通过双层特征融合结构加强网络各层间的信息流动与特征复用,在有效避免浅层特征丢失的同时为去雨图像提供更多的细节信息.全局特征融合注意力模块进一步提高了去雨图像质量,且整体网络易于训练.

2) 为了充分利用多尺度信息,设计出多尺度注意力残差模块(Multi-scale Attention Residual Block, MARB).首先利用不同扩张卷积提取不同尺度的特征信息,再通过通道分离—卷积—拼接的方式增强网络的特征提取能力,最后坐标注意力(Coordinate Attention, CA)^[8]通过位置信息对通道关系和长期依赖性进行编码,提高网络对于重点特征的关注.

3) 使用均方误差以及结构相似度作为联合损失函数约束网络.在仿真和真实雨像数据集上进行评估,所提算法有着较好的去雨纹效果且保留较多的图像细节信息.

1 相关工作

1.1 残差网络

卷积神经网络在计算机视觉领域已经取得巨大成功,学者们提出许多经典的卷积神经网络如 LeNet、AlexNet、VGGNet 等.通过增加网络的宽度和深度可以提高网络的性能,但是一味地增加深度会导致梯度弥散/爆炸以及网络退化问题.因此,He 等^[9]提出一种残差网络,并在 ImageNet 比赛中获得分类和识别任务的冠军.残差网络由多个残差块堆叠而成,在残差块中使用一种跳跃连接,有效缓解了深度神经网络中增加深度所带来的问题.

$$H(x) = h(x) + F(x), \quad (1)$$

其中, x 为残差块输入, $F(x)$ 为恒等映射输出, $F(x)$ 为残差映射.前向传播时,输入信号可以从任意低层直接传播到高层,有效解决网络退化问题;反向传播时,信号可以不经过中间权重层直接到达低层,有效缓解梯度弥散的问题.

1.2 注意力机制

人类可以高效地在复杂场景中找到重要区域.受此启发,注意力机制被引入到计算机视觉中,目的是模仿人类视觉系统.注意力机制在图像分类、目标检测、语义分割等视觉任务中进一步提升了网络的性能.康健等^[10]设计出 PFA-LinkNet,利用通道注意力机制对编码阶段的特征图各通道间的相互关系进行建模来增强网络提取水体特征的能力.胡序洋

等^[11]提出一种车道线检测方法 LaneSegNet,使用一种混合注意力网络,将空间与通道注意力相结合来获取丰富的车道线特征,增强网络的性能.张凯等^[12]提出一种用于射频指纹识别的注意力残差卷积神经网络,通过注意力机制,提高网络对于指纹图像中有效信息的关注.

2 本文方法

2.1 网络结构

本文算法 MFFAN 的网络结构如图 1 所示,网络的整体架构包括 5 个部分:浅层特征提取模块(Shallow Feature Extraction Block, SFEB)、多个串联的多尺度注意力残差组模块(Multi-scale Attention Residual Group Block, MARGB)、局部特征融合模块(Local Feature Fusion Block, LFFB)、全局特征融合注意力模块(Global Feature Fusion Attention Block, GFFAB)和图像重建模块(Image Reconstruction Block, IRB).其中每个多尺度注意力残差组由 2 个多尺度注意力残差模块和局部特征融合模块组成,定义输入图像为 x_{input} .

浅层特征提取模块(SFEB)由一层卷积核大小为 3 的卷积层构成,对输入图像进行浅层特征提取.

$$S = \text{Conv}_{3 \times 3}(x_{input}), \quad (2)$$

其中, $\text{Conv}_{i \times i}$ 表示卷积核大小为 i 的卷积, S 代表浅层提取模块的输出.

多个多尺度注意力残差组模块(MARGB)可表示为

$$M = f_m^i(M), i = 1, \dots, n, \quad (3)$$

其中, f_m^i 表示第 i 个 MARGB, M 表示第 i 个 MARGB 的输入和第 $i - 1$ 个 MARGB 的输出.

全局特征融合模块(GFFAB)可表示为

$$G = f_g([M_1, M_2, M_3, \dots, M_n]), \quad (4)$$

其中, f_g 表示 GFFAB, $[]$ 表示通道拼接操作, G 表示特征融合注意力模块的输出.

局部特征融合模块(LFFB)仅使用通道拼接以及 1×1 卷积进行融合.最后使用卷积大小为 3 的卷积核进行图像重建,网络的最终输出可表示为

$$Y = X - \text{Conv}_{3 \times 3}(G), \quad (5)$$

其中, Y 表示网络的最终输出图像.

2.2 多尺度注意力残差模块

在图像去雨任务中,空间上下文信息对捕捉和去除雨纹十分重要.雨纹的形状和密度都是多样的,使用普通的残差块无法很好地提取到整个图像的雨

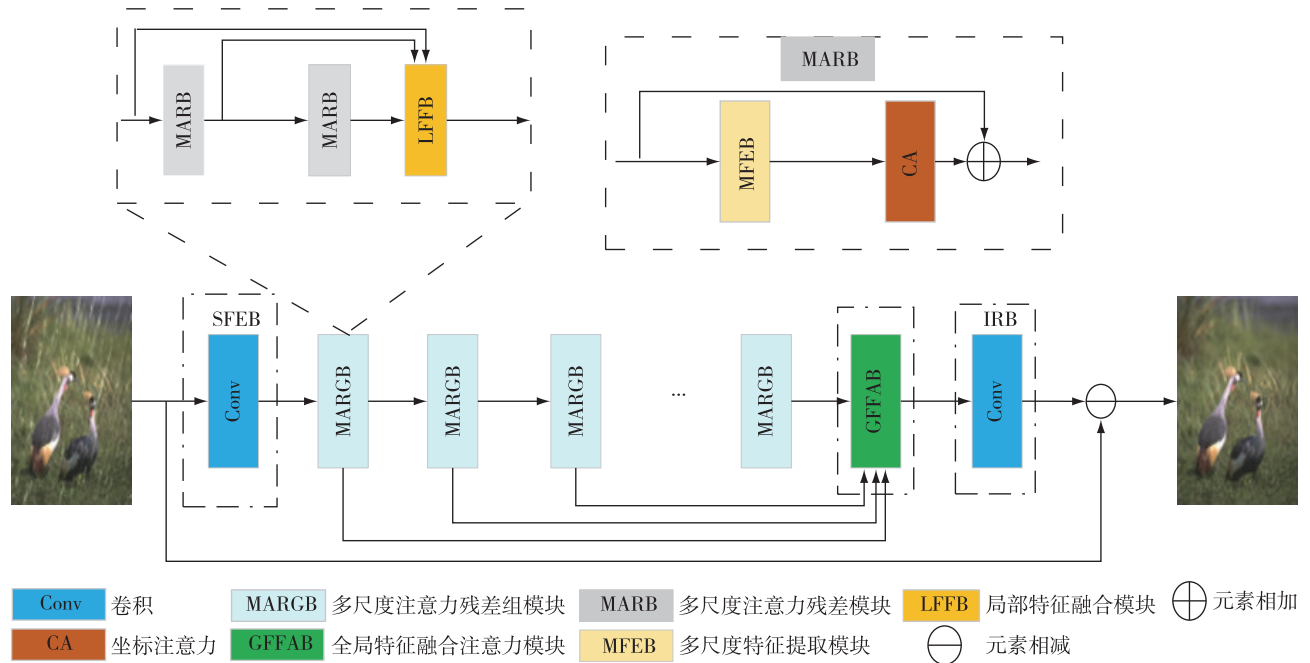


图1 MFFAN 网络框架

Fig. 1 Network framework of the proposed MFFAN

纹特征.因此,本文使用多尺度注意力残差模块提取雨纹特征,利用多尺度特征提取模块(Multi-scale Feature Extraction Block, MFEB)代替残差块中串联的2个 3×3 卷积,并使用坐标注意力进一步细化提取特征,具体结构如图2所示.

对于模块输入 n 首先使用2个大小为3,扩张因子分别为1、2的扩张卷积来提取感受野尺度为 3×3 和 5×5 的特征信息,并使用LeakyReLU激活函数 ϕ .随后将不同尺度的特征进行拼接,使用 1×1 的卷积进行通道降维.

$$k = \text{Conv}_{1 \times 1}[\phi(\text{Conv}_{3 \times 3, 1}(n)), \phi(\text{Conv}_{3 \times 3, 2}(n))], \quad (6)$$

其中, $\text{Conv}_{ixi,j}$ 表示卷积核大小为 i ,扩张系数为 j 的扩张卷积.

为了实现在细粒度层面提取雨纹的多尺度特征,在通道维度使用多级分割卷积拼接方式对雨纹特征 k 进行提取.首先将 k 分割成4个具有同等的通道数的特征 k_1, k_2, k_3, k_4 .第1组特征图 k_1 经过卷积核大小为3的卷积,输出特征为 o_1 .将 o_1 在通道维度对半分割成 $o_{1,1}$ 和 $o_{1,2}$,并将 $o_{1,2}$ 与下一组特征 k_2 进行通道拼接后送入卷积核大小为3的卷积中.这样操作重复3次,使得拼接后的特征得到充分复用.在前面分组中的特征所经历卷积次数较少,其感受野也较小,这能够帮助网络有效关注细节信息.相反,后面

分组中含有的特征经历的卷积次数较多,其感受野也就较大,能够帮助网络关注全局信息.

$$o_i = \begin{cases} \text{Conv}_{3 \times 3}(k_i), i = 1, \\ \text{Conv}_{3 \times 3}([k_i, o_{i-1,2}]), 1 < i \leq 4, \end{cases} \quad (7)$$

最终,将输出特征图 $o_{1,1}, o_{2,1}, \dots, o_4$ 沿通道维度进行拼接,再使用 1×1 卷积降低通道维度:

$$y = \text{Conv}_{1 \times 1}([o_{1,1}, o_{2,1}, \dots, o_4]). \quad (8)$$

接着使用坐标注意力进一步加强网络性能.坐标注意力首先使用尺寸为 $(H, 1)$ 和 $(1, W)$ 的池化核沿着输入特征图 y 的水平坐标及垂直坐标对每个通道使用全局平均池化.因此,高度为 h 的第 c 个通道的输出以及宽度为 w 的第 c 个通道的输出可以表示为

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} y_c(h, i), \quad (9)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} y_c(j, w), \quad (10)$$

其中: $y_c(h, i), y_c(j, w)$ 分别表示输入特征图 y 中坐标为 $(h, i), (j, w)$ 、通道维度为 c 的分量; $z_c^h(h)$ 表示高度为 h 的第 c 个通道的输出分量; $z_c^w(w)$ 表示宽度为 w 的第 c 个通道的输出分量.

之后将 z^h 与 z^w 在通道维度上进行拼接,并使用一个共享的 1×1 卷积进行通道降维,紧接着对其使用H-swish非线性激活函数 δ ,得到中间特征 t :

$$t = \delta(\text{Conv}_{1 \times 1}(z^h, z^w)). \quad (11)$$

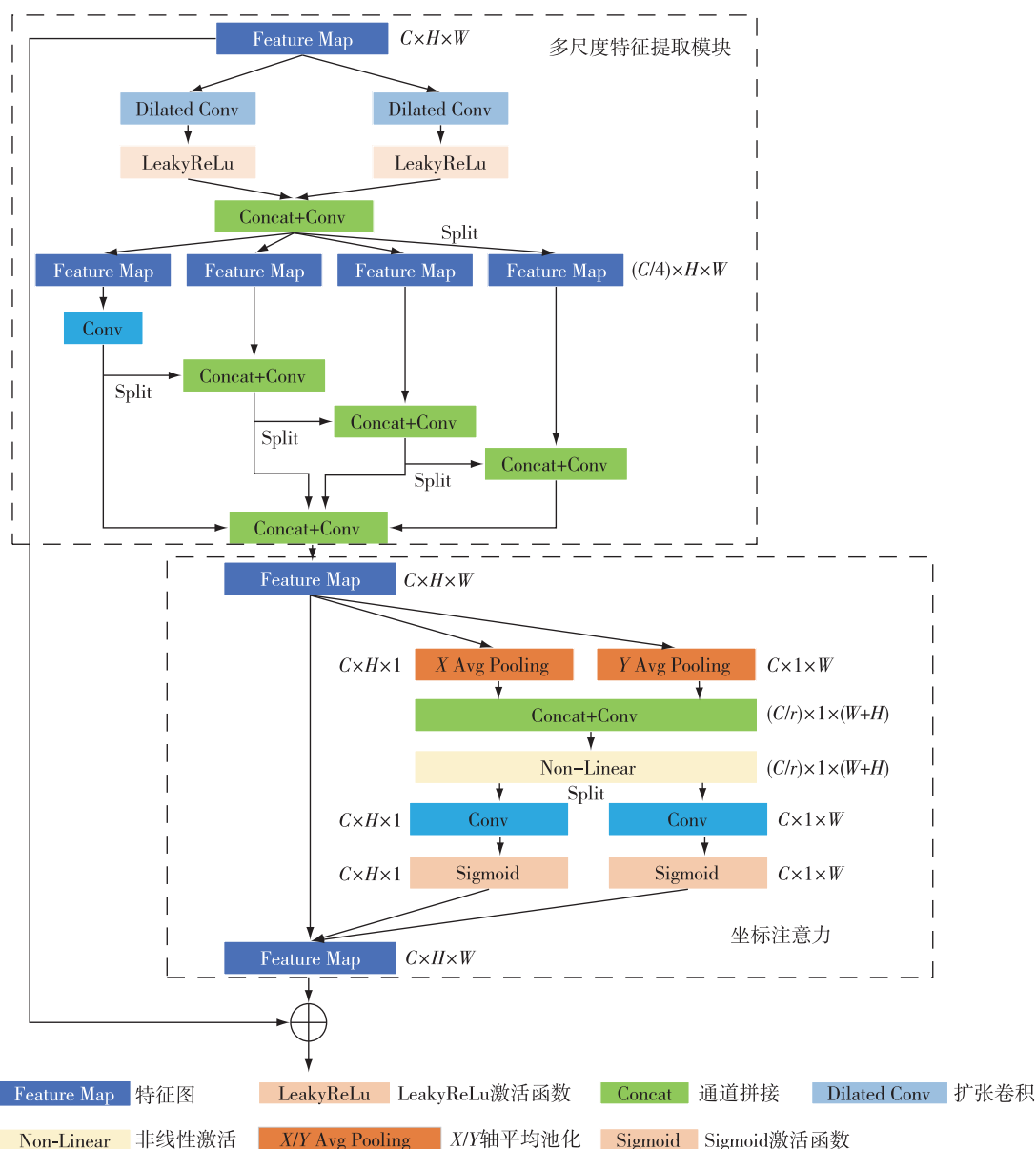


图2 多尺度注意力残差模块

Fig. 2 Multi-scale attention residual block

将 t 沿着空间分割成 2 个独立分量 t^h 和 t^w , 并使用 2 个 1×1 的卷积将 2 个独立分量转换到与输入特征 y 具有相同的通道数的分量, 再使用 Sigmoid 激活函数, 可表示为

$$m^h = \sigma(\text{Conv}_{1 \times 1}(t^h)), \quad (12)$$

$$m^w = \sigma(\text{Conv}_{1 \times 1}(t^w)), \quad (13)$$

m^h 和 m^w 分别表示 2 个坐标轴上坐标的注意力权重, σ 表示 Sigmoid 激活函数. 坐标注意力的输出可以表示为

$$u_c(i, j) = y_c(i, j) \times m_c^h(i) \times m_c^w(j), \quad (14)$$

$u_c(i, j)$ 表示输出特征图中坐标为 (i, j) 、通道维度为

c 的分量.

最终, 整体模块的输出 M 可表示为

$$M = n + u. \quad (15)$$

2.3 全局特征融合注意力模块

简单相加或拼接操作无法很好地融合不同层次的特征. 为了加强特征融合的效果, 本文使用一种像素注意力为图像中的每个像素赋予不同的权重, 使得网络对于雨纹区域的关注度更高. 像素注意力^[13]应用于图像去雾领域用以解决雾特征在不同像素区域分布不均的问题. 首先对不同层次的特征进行通道拼接, 并使用一个 1×1 的卷积进行通道降维, 最后

送入到像素注意力中,模块整体结构如图3所示.

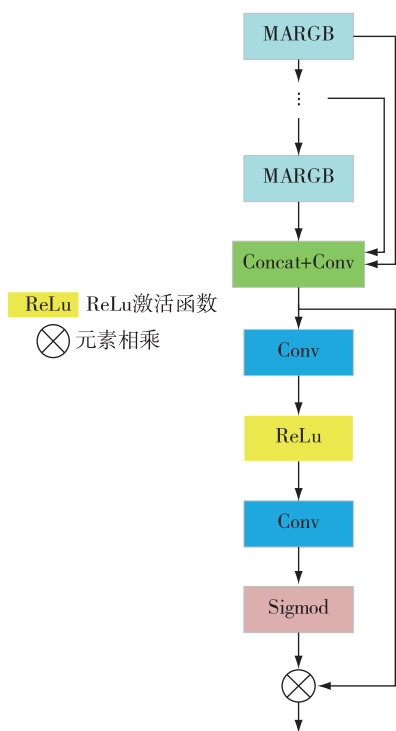


图3 全局特征融合注意力模块

Fig. 3 Global feature fusion attention block

全局特征融合注意力模块(GFFAB)可表示为

$$g = \text{Conv}_{1 \times 1} [M_1, M_2, \dots, M_i], \quad (16)$$

$$P_A = \sigma(\text{Conv}_{1 \times 1}(\varphi(\text{Conv}_{1 \times 1}(g))))), \quad (17)$$

$$G = P_A \otimes g, \quad (18)$$

其中, g 表示聚合的全局特征, φ 表示 ReLU 激活函数, P_A 表示像素注意力模块输出, G 表示整体模块输出.

2.4 损失函数

均方误差(Mean Square Error, MSE)作为损失函数被广泛用于图像去雨任务,但会造成重建图像中的高频纹理过度平滑甚至模糊. 结构相似度 (Structural Similarity, SSIM) 损失作为衡量 2 幅图像之间结构相似性的重要指标,可以使去雨后的图像更符合人类的视觉感官.为了更好地训练网络,将 SSIM 损失和 MSE 损失结合形成一种混合损失函数.

$$L_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^N \|h(x_i) - I_i\|^2, \quad (19)$$

$$L_{\text{SSIM}} = 1 - \text{SSIM}(h(x), I), \quad (20)$$

其中, x 是网络输入图像, $h(x)$ 是网络输出图像, I 是无雨标签图像.

通过组合 L_{MSE} 和 L_{SSIM} , 最终的混合损失函数可以表示为

$$L = L_{\text{MSE}} + rL_{\text{SSIM}}, \quad (21)$$

其中, r 作为平衡 MSE 损失与 SSIM 损失之间的超参数.实验时,将输入图像首先进行归一化处理再送入网络,使得 MSE 损失与 SSIM 损失的范围均在 $[0, 1]$ 之间.根据文献[14],将 r 设置为 0.2,在图像保持像素级相似性的同时保持全局结构,获得与原图更接近的去雨图像.

3 实验与结果

3.1 数据集与训练配置

利用 Rain100L、Rain100H 以及 Rain12 数据集来对 MFFAN 进行训练和测试.Rain100L 是小雨数据集,由一种类型的雨带的合成图像构成,包含 200 对训练图像和 100 对测试图像.Rain100H 是大雨数据集,含有 5 种雨纹方向,包括 1 800 对训练图像以及 100 对测试图像.Rain12 数据集由于数量较少被作为测试集,将 Rain100L 作为训练集进行训练.真实世界的雨图可以有效评估算法的泛化性能.

实验时,将每张图像随机裁成大小 100×100 的图像块进行训练.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 16,卷积核通道数设置为 32.实验在一台装有 Linux 系统,显卡型号为 NVIDIA Tesla T4, Python3.7 版本,pytorch1.2.0 版本的服务器上完成.

3.2 仿真雨像实验与分析

将所提算法在仿真数据集上与 JORDER^[15]、DRN^[16]、PReNet^[4]、ReHEN^[17]、LPNet^[5] 和 BRN^[18] 进行对比.图 4 为本文方法与其他现有方法的去雨图像.表 1 为各方法在仿真数据集上测试的 PSNR (Peak Signal To Noise Ratio, 峰值信噪比)、SSIM 均值,均在 YcrCb 空间取 Y 通道分量进行计算.在 Rain100H、Rain100L 数据集上,本文方法获得了最好的结果,PSNR 较次优值分别提高 0.26 dB 和 0.45 dB,SSIM 分别提高 0.5% 和 0.3%,在 Rain12 数据集上也获得了较高的评价指标值.

Rain100H 由于雨线密集因此更加具有挑战性.从中选取 2 张图像进行效果评价,如图 4 所示: JORDER 去除了大部分雨线,但是仍有雨线残留(如上图水草部分,下图天空部分),同时细节恢复不佳; DRN 有效去除了雨线,但是放大区域内颜色信息没有有效恢复(如上图鱼眼部分,下图建筑部分); PReNet 在放大区域内仍有轻微雨线残留且部分区

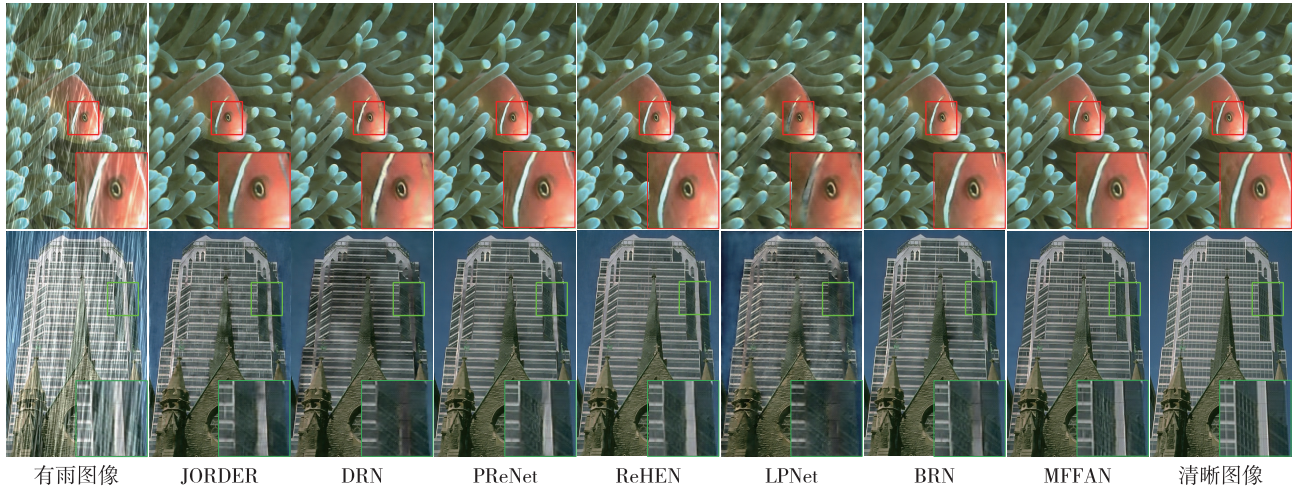


图4 仿真雨天图像多种算法去雨效果对比

Fig. 4 Comparison of rain removal results by various algorithms on simulated rain images

表1 不同算法在仿真雨天数据集上 PSNR、SSIM 对比

Table 1 Comparison of PSNR and SSIM for different algorithms on simulated rainy day datasets

方法	Rain100H		Rain100L		Rain12	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
JORDER	26.54	0.835	36.61	0.974	33.92	0.953
DRN	26.89	0.860	35.73	0.969	36.12	0.957
PReNet	29.43	0.899	37.39	0.978	36.64	0.960
ReHEN	29.10	0.876	38.41	0.980	36.14	0.946
LPNet	23.73	0.814	34.26	0.956	35.29	0.951
BRN	29.58	0.902	37.82	0.981	36.70	0.959
MFFAN	29.84	0.907	38.86	0.984	36.72	0.957

注:加粗字体表示最优值.

域模糊不清(如下图放大区域墙壁部分);ReHEN在雨线密集部分处理效果较差(如上图鱼嘴部分),并存在部分细节信息丢失现象(如下图建筑玻璃部分);LPNet去雨效果并不是很好,且纹理十分模糊;BRN过分去除雨纹,导致图像中的白色纹理部分被去除(如上图鱼身部分),同时色彩信息恢复较差(如下图放大区域建筑墙壁部分);MFFAN图像中的雨纹去除较为彻底同时能够保留较多的结构纹理,整体图像与相对应的图像结构更相似.

JORDER利用扩张卷积获得多尺度特征,但忽略了对细粒度层面多尺度特征的有效利用;PReNet、DRN和BRN利用相似的递归结构去除雨纹但却没有充分利用通道信息;ReHEN通过SENet为不同的通道赋予不同的权重,提高了去雨的效率但缺乏对于像素信息的利用;LPNet利用多分支结构获得不同尺度的图像,去雨速度较快但是效果并不理想.相

比之下,MFFAN总体网络深度并不深,只有一条主分支避免多分支结构带来的运算成本的提高.MFFAN网络利用多尺度注意力残差模块获得多尺度特征,通过坐标注意力和像素注意力充分利用通道维度和像素维度的信息.简单的双层特征融合结构加强网络各层间流动,且为深度特征提供更多的细节.因此,在3种合成数据集上,MFFAN能够更有效地处理各种形状的雨纹,同时保留图像的细节纹理和色彩信息.

3.3 真实雨像实验与分析

图5为MFFAN与JORDER、DRN、PReNet、ReHEN、LPNet和BRN在真实雨天图像上进行去雨后的结果.JORDER无法彻底去除雨纹在图(第1行第2列)中放大区域内仍有雨纹残留,DRN和BRN有效去除大部分雨纹,但在图(第1行第3列、第2行第3列)左上方区域内产生模糊.PReNet在图(第1行第4列)中放大区域内仍有些许雨纹残留,同时左上方区域细节恢复效果不佳.REHEN无法去除大尺度雨纹,2幅图中放大区域内仍存在雨纹.LPNet去雨后的图像较为平滑,背景纹理细节模糊.而MFFAN展现了较好的泛化性能,去雨后的图像有着很好的视觉效果.

由于真实世界中的雨像缺乏相应的无雨清晰图像,本文仅从定性方面对现有去雨方法及本文方法进行去雨效果的评估.从分析的结果来看,相较于其他现有的方法,本文所提的MFFAN可以更有效地处理真实环境中的雨纹,获得更符合人类视觉感官的去雨图像.

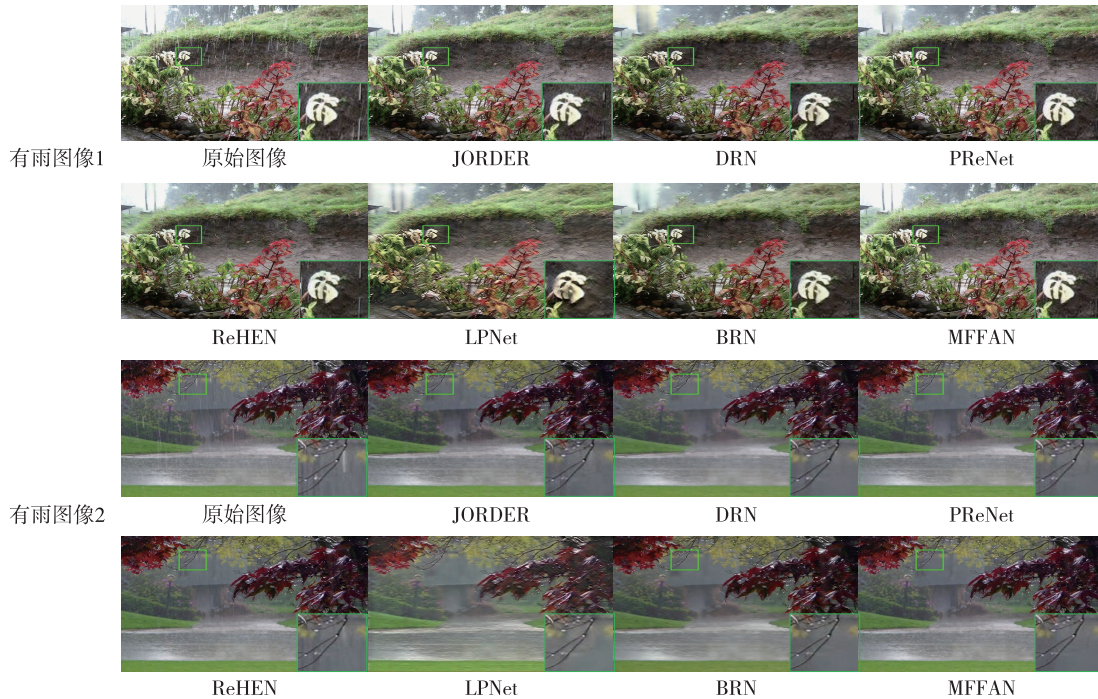


图5 真实雨天图像多种算法去雨效果对比

Fig. 5 Comparison of rain removal results by various algorithms on real rain images

3.4 消融实验与分析

为了验证所提算法的合理性,在同一环境相同配置下进行消融实验.表2展示了不同数量的残差组数 N 在 Rain100L 数据集上的 PSNR、SSIM 值.

表2 不同数量的残差组在 Rain100L 上 PSNR、SSIM 对比

Table 2 Comparison of PSNR and SSIM on Rain100L for different number of residual groups

残差组数 N	PSNR/dB	SSIM
2	38.24	0.981 9
3	38.45	0.982 6
4	38.86	0.984 1
5	38.94	0.984 1

注:加粗字体表示最优值.

由表2可以看出,当多尺度注意力残差组的数量设置为5时获得了最高的PSNR,但是此时的SSIM

值相较于 $N=4$ 时却没有得到增长.这是由于网络深度达到一定的阈值时,其性能并不能得到有效改善.因此,从运算成本以及去雨性能两方面进行综合考虑,将网络中残差组数量 N 设置为4.

为了验证本文所提模块的有效性,对不同模块的组合进行消融实验.如表3所示:组合1使用2个串联的 3×3 卷积代替多尺度特征提取模块并且在模块间进行直连作为基准模型;组合2在多尺度注意力残差模块中使用多尺度特征提取模块;组合3在每两个多尺度注意力残差模块间添加局部特征融合;组合4在图像重建前添加全局特征融合注意力模块;组合5将全局特征融合注意力模块中的像素注意力去除;组合6即本文所提方法.

从表3中可以明显地看出本文所提网络获得了最高的PSNR和SSIM值.基准模型在增加多尺度特

表3 不同组合模块在 Rain100L 上 PSNR、SSIM 对比

Table 3 Comparison of PSNR and SSIM on Rain100L with different combination blocks

组合	3×3 Conv	MFEB	LFFB	GFFAB	GFFB	CA	PSNR/dB	SSIM
1	✓						37.09	0.976 7
2		✓					37.22	0.979 0
3		✓	✓				37.40	0.979 0
4		✓	✓	✓			37.66	0.980 1
5		✓	✓		✓	✓	38.65	0.983 4
6		✓	✓	✓		✓	38.86	0.984 1

注:加粗字体表示最优值.

征提取模块后,有效帮助网络提取丰富的雨纹特征.全局特征与局部特征融合增强了网络内部的信息流动性,评价指标得到提高,在多尺度残差注意力模块中添加坐标注意力及在全局特征融合中增加像素注意力即本文所提方法,增强网络的特征提取能力并且得到最高的指标,表明各模块是缺一不可的.

最后在同环境下验证不同损失函数对去雨效果的影响,结果如表4所示.

表4 不同损失函数在 Rain100L 上 PSNR、SSIM 对比

Table 4 Comparison of PSNR and SSIM on Rain100L with different loss functions

	PSNR/dB	SSIM
MSE	38.46	0.9812
SSIM	38.79	0.9835
本文所用损失函数	38.86	0.9841

注:加粗字体表示最优值.

在表4中,可以看到MSE和SSIM的联合损失函数的PSNR比单一损失函数MSE和SSIM分别高出0.4dB和0.07dB,SSIM则分别高出0.29%和0.06%.两种指标的提高有效表明了混合损失函数可以获得更好的去雨效果.

消融实验从残差组数量 N ,网络中不同模块的组合以及损失函数三方面来对网络进行实验分析,将残差组数量 N 设置为4,使得网络在适当的深度时获得高效的去雨性能,可以避免网络深度过深时,性能却并没有得到提高的问题.本文所提的各模块组合获得了最高的评价指标,说明缺少任意模块都会造成去雨性能的下降,其中多尺度特征提取模块、坐标注意力和像素注意力是本文网络重要的组成模块.最后本文所用的混合损失函数相较于单一的损失函数MSE和SSIM,在Rain100L上取得了最佳的评价指标,表明使用混合损失函数可以获得更好的雨纹处理效果且保留较多的图像细节信息.

4 结论

本文提出一种基于注意力机制的多尺度特征融合图像去雨算法(MFFAN),该算法利用双层特征融合结构将浅层特征进行融合作为后续模块的输入用以充分复用特征,使网络关注局部细节信息的同时关注全局信息.多尺度注意力残差模块充分利用了不同尺度的特征信息,进一步提高网络提取雨纹特征的能力.全局特征融合注意力增强网络对于雨水区域和纹理细节的学习.采用混合损失函数可以获

得更好的图像去雨效果.所提算法在Rain100L、Rain100H、Rain12及真实雨像数据集中获得较高的PSNR和SSIM值,在有效去除雨纹的同时很好地保留和恢复图像的细节纹理与色彩信息,泛化性能较好.如何在更加恶劣的天气情况下有效恢复图像以及如何加强模型在真实雨像下的泛化性能将是以后需要改进和解决的问题.

参考文献

References

- [1] 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, 2012, 21(4): 1742-1755
- [2] Li Y, Tan R T, Guo X J, et al. Rain streak removal using layer priors[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. IEEE, 2016: 2736-2744
- [3] Fu X Y, Huang J B, Ding X H, et al. Clearing the skies: a deep network architecture for single-image rain removal[J]. IEEE Transactions on Image Processing, 2017, 26(6): 2944-2956
- [4] Ren D W, Zuo W M, Hu Q H, et al. Progressive image deraining networks: a better and simpler baseline[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA. IEEE, 2019: 3932-3941
- [5] Fu X Y, Liang B R, Huang Y, et al. Lightweight pyramid networks for image deraining[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(6): 1794-1807
- [6] Wang C, Xing X Y, Wu Y T, et al. DCSFN: deep cross-scale fusion network for single image rain removal[C]//Proceedings of the 28th ACM International Conference on Multimedia. Seattle, WA, USA. New York, NY, USA: ACM, 2020: 1643-1651
- [7] Yi Q S, Li J C, Dai Q Y, et al. Structure-preserving deraining with residue channel prior guidance[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada. IEEE, 2021: 4218-4227
- [8] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA. IEEE, 2021: 13708-13717
- [9] He K M, Zhang X Y, Ren S Q, et al. Identity mappings in deep residual networks[M]//Computer Vision—ECCV 2016. Cham: Springer International Publishing, 2016: 630-645
- [10] 康健, 管海燕, 于永涛, 等. 基于RFA-LinkNet模型的高分遥感影像水体提取[J]. 南京信息工程大学学报(自然科学版), 2023, 15(2): 160-168
KANG Jian, GUAN Haiyan, YU Yongtao, et al. RFA-LinkNet: a novel deep learning network for water-body extraction from high-resolution remote sensing images

- [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2023, 15(2): 160-168
- [11] 胡序洋,高尚兵,汪长春,等. LaneSegNet:一种高效的车道线检测方法[J].南京信息工程大学学报(自然科学版),2022,14(5):551-558
HU Xuyang, GAO Shangbing, WANG Changchun, et al. LaneSegNet: an efficient lane line detection method[J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2022, 14(5): 551-558
- [12] 张凯,郭剑黎,胡军星,等.基于注意力残差网络的Wi-Fi设备的射频指纹识别[J].南京信息工程大学学报(自然科学版),2022,14(3):324-330
ZHANG Kai, GUO Jianli, HU Junxing, et al. Radio frequency fingerprint identification of Wi-Fi device based on attention residual network [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2022, 14(3): 324-330
- [13] Qin X, Wang Z L, Bai Y C, et al. FFA-net: feature fusion attention network for single image dehazing [C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York, USA: AAAI, 2020: 11908-11915
- [14] 马婧婧,黄煜峰,陈翔.多尺度沙漏结构的单幅图像去雨算法研究[J].小型微型计算机系统,2021,42(3):561-565
MA Jingjing, HUANG Yufeng, CHEN Xiang. Multi-scale hourglass network for single image deraining [J]. Journal of Chinese Computer Systems, 2021, 42(3): 561-565
- [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. Honolulu, HI, USA. IEEE, 2017: 1685-1694
- [16] Cai L W, Li S W, Ren D W, et al. Dual recursive network for fast image deraining [C]//Proceedings of the IEEE International Conference on Image Processing. Taipei, China. IEEE, 2019: 2756-2760
- [17] Yang Y Z, Lu H. Single image deraining via recurrent hierarchy enhancement network [C]//Proceedings of the 27th ACM International Conference on Multimedia. Nice, France. New York: ACM, 2019: 1814-1822
- [18] Shang W, Zhu P F, Ren D W, et al. Bilateral recurrent network for single image deraining [C]//2020 IEEE International Conference on Acoustics, Speech and Signal Processing. Barcelona, Spain. IEEE, 2020: 2503-2507

Image rain removal via multi-scale feature fusion based on attention mechanism

LIU Zhongyang¹ ZHOU Jie¹ LU Jiabin² MIAO Zelin¹ SHAO Genfu³ JIANG Kaiqiang¹ GAO Wei¹

¹ School of Artificial Intelligence (School of Future Technology),

Nanjing University of Information Science & Technology, Nanjing 210044, China

² School of Electronics & Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China

³ School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China

Abstract Due to the diversity of the distribution and shape of rain streaks, existing rain removal algorithms produce problems such as blurred image background and poor generalization performance while removing rain. A multi-scale feature fusion image rain removal approach based on attention mechanism is proposed to address these problems. The feature extraction consists of multiple residual groups containing two multi-scale attention residual blocks, which use the multi-scale feature extraction module to extract and aggregate feature information at different scales and further improve the feature extraction capability of the network through coordinate attention. Local feature fusion is performed within groups, and the global feature fusion attention module is used between groups to better fuse features at different levels and to focus the network on rain streak regions through pixel attention. The quantitative metrics of the proposed approach are significantly improved compared with other existing image rain removal algorithms on both simulated and real rain image datasets, and the rain removal images are greatly improved in both visual effects and generalization performance.

Key words image rain removal; multi-scale; feature fusion; residual network; coordinate attention