

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

跨尺度注意力融合的单幅图像去雨

叶宇超, 陈莹

引用本文:

叶宇超, 陈莹. 跨尺度注意力融合的单幅图像去雨[J]. 光电工程, 2023, 50(10): 230191.

Ye Y C, Chen Y. Single image rain removal based on cross scale attention fusion[J]. *Opto-Electron Eng*, 2023, 50(10): 230191.

<https://doi.org/10.12086/oe.2023.230191>

收稿日期: 2023-07-28; 修改日期: 2023-10-26; 录用日期: 2023-10-26

相关论文

融合暗通道先验损失的生成对抗网络用于单幅图像去雾

程德强, 尤杨杨, 寇旗旗, 徐进洋

光电工程 2022, 49(7): 210448 doi: 10.12086/oe.2022.210448

基于残差和注意力网络的声呐图像去噪方法

赵冬冬, 叶逸飞, 陈朋, 梁荣华, 蔡天诚, 郭新新

光电工程 2023, 50(6): 230017 doi: 10.12086/oe.2023.230017

多尺度注意力与领域自适应的小样本图像识别

陈龙, 张建林, 彭昊, 李美惠, 徐智勇, 魏宇星

光电工程 2023, 50(4): 220232 doi: 10.12086/oe.2023.220232

更多相关论文见光电期刊集群网站 



光电工程
Opto-Electronic Engineering

<http://cn.oeejournal.org/oe>



 OE_Journal



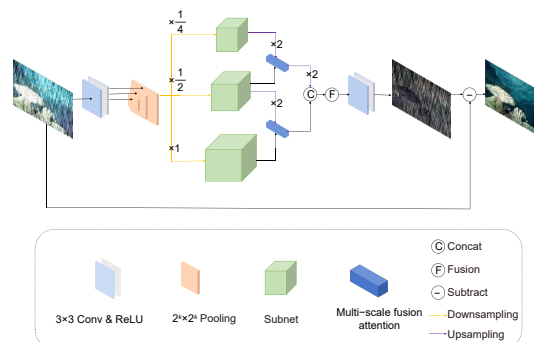
Website

DOI: 10.12086/oe.2023.230191

跨尺度注意力融合的单幅图像去雨

叶宇超, 陈莹*

江南大学, 轻工过程先进控制教育部重点实验室,
江苏 无锡 214222



摘要: 单幅图像去雨算法旨在将有雨图像中的雨纹去除生成高质量无雨图。目前基于深度学习的多尺度去雨算法较难捕获不同层次的细节, 忽视尺度之间的信息互补, 易导致生成图像失真, 雨纹去除不彻底等问题。为此, 本文提出了基于跨尺度注意力融合的单幅图像去雨网络, 在去除密集雨纹的同时尽量保留原本图片的细节, 改善去雨图像的视觉质量。去雨网络由三个子网构成, 每个子网用于获取不同尺度上的雨纹信息。各子网由跨尺度特征提取模块通过稠密连接的方式构成, 该模块以跨尺度融合注意力为核心, 构造不同尺度之间的关联实现信息互补, 使图像兼顾细节与整体信息。实验结果表明, 本文模型在合成数据集 Rain200H 和 Rain200L 上取得显著的去雨效果, 去雨处理后的图片峰值信噪比达到了 29.91/39.23 dB, 结构相似度为 0.92/0.99, 优于一般的主流方法, 并取得了良好的视觉效果, 在保证去雨效果自然的同时保持了图像的细节。

关键词: 机器学习; 图像去雨; 多尺度; 注意力机制; 特征融合

中图分类号: TP391.4

文献标志码: A

叶宇超, 陈莹. 跨尺度注意力融合的单幅图像去雨 [J]. 光电工程, 2023, 50(10): 230191

Ye Y C, Chen Y. Single image rain removal based on cross scale attention fusion[J]. *Opto-Electron Eng*, 2023, 50(10): 230191

Single image rain removal based on cross scale attention fusion

Ye Yuchao, Chen Ying*

Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi, Jiangsu 214122, China

Abstract: Single-image rain removal is a crucial task in computer vision, aiming to eliminate rain streaks from rainy images and generate high-quality rain-free images. Current deep learning-based multi-scale rain removal algorithms face challenges in capturing details at different scales and neglecting information complementarity among scales, which can lead to image distortion and incomplete rain streak removal. To address these issues, this paper proposes an image rain removal network based on cross-scale attention fusion, aiming to remove dense rain streaks while preserving original image details to improve the visual quality of the rain removal image. The rain removal network consists of three sub-networks, each dedicated to obtaining rain pattern information at different

收稿日期: 2023-07-28; 修回日期: 2023-10-26; 录用日期: 2023-10-26

基金项目: 国家自然科学基金资助项目 (62173160)

*通信作者: 陈莹, chenying@jiangnan.edu.cn。

版权所有©2023 中国科学院光电技术研究所

scales. Each sub-network is composed of densely connected cross-scale feature fusion modules. The designed module takes the cross-scale attention fusion as the core, which establishes inter-scale relationships to achieve information complementarity and enables the network to consider both details and global information. Experimental results demonstrate the effectiveness of the proposed model on synthetic datasets Rain200H and Rain200L. The peak signal-to-noise ratio (PSNR) of the derained images reaches 29.91/39.23 dB, and the structural similarity index (SSIM) is 0.92/0.99, outperforming general mainstream methods and achieving favorable visual effects while preserving image details and ensuring natural rain removal.

Keywords: machine learning; image deraining; multi-scale; attention mechanism; feature fusion

1 引言

降雨天气获取到的图片存在着细节丢失、纹理模糊、对比度下降等问题,不利于图像特征提取和辨识,影响了后续户外计算机视觉系统的可靠性和准确性。因此,图像去雨算法具有重要的研究意义和广泛的应用前景。现有的图像去雨算法主要分为基于视频的去雨算法和单幅图像去雨算法。单幅图像去雨算法由于缺少时序信息,相比基于视频的算法去雨难度更大。

一些早期的工作大多是基于图像先验的方法实现单幅图像去雨。Kang^[1]等人通过字典学习和稀疏编码的方法将高频部分中雨纹和非雨成分分离进行去雨。Kim等人^[2]使用自适应的非均值滤波器来获取无雨图像。Li^[3]等人运用高斯混合模型构建先验信息,将背景层和雨纹层通过预先设定的约束进行分离。以上的传统方法受到雨纹复杂性和背景图层多样性的影响,无法彻底去除雨纹并且生成的图片质量较差。

近年来,随着深度学习在计算机视觉领域的发展和运用,基于卷积神经网络的单幅图像去雨算法取得了较大的进展。Eigen等人^[4]的工作首次将卷积神经网络运用到单幅图像去雨中去。Fu等人^[5]提出了一种深度细节网络,图片的高频细节被引导滤波分离,然后利用深度残差网络将其映射到雨纹层,在高层的细节中去除雨纹。Zhang等人^[6]通过设计雨量密度信息获取网络,将雨量密度标签作为指导更加有效地去除图像雨纹。Ren等人^[7]基于残差网络设计了由简单的网络结构组成的渐进循环网络。Qian等人^[8]设计了生成对抗网络,并且在生成器中结合注意力机制,将注意力图和输入图像一起应用生成无雨图像。Wei等人^[9]将CycleGAN的循环一致性运用到图像去雨领域,通过互补的转换过程增强了模型的一致性和稳定性,并首次实现了无监督去雨任务。Huang等人^[10]提出了一种有效的半监督图像去雨方法,将有监督与无监督数据进行记忆存储,并将其用于迁移学习的过程中,

取得了较好的去雨性能。

以上这些方法都将去雨任务放在单尺度的网络中实现。然而,有雨图像的雨纹复杂性和背景多样性导致单尺度网络很难兼顾清雨的彻底性和图片的高质量,因此很多研究者设计了多尺度的网络结构用于图像去雨。Yang等人^[11]运用空洞卷积获取不同尺度的感受野,构造了多通道单幅图像去雨网络,通过闭环反馈的方式多阶段去除雨纹。但该方法使用了固定的权重和简单的融合策略,未考虑不同尺度特征下在不同场景的重要性和贡献度,影响了去雨效果的平衡性和稳定性。Fu等人^[12]利用轻量级拉普拉斯金字塔网络获取不同尺度中的特征信息,构建递归的残差网络快速去除雨纹。该模型虽然在参数量和计算复杂度上具有优势,但是去雨性能一般,无法彻底地去除图像中的雨线。Jiang等人^[13]提出一种多尺度渐进融合网络,通过高斯核进行下采样构建高斯金字塔结构,然后进行特征提取和融合,实现去雨任务。然而该方法并未充分考虑不同尺度之间的信息关联性,导致了特征的重复和冗余,影响了去雨效果的提升。Wang等人^[14]通过将U-Net中的卷积层用Transformer代替,提出了Uformer结构,提升了模型捕获全局依赖关系的能力,首次将Transformer引入图像去雨领域,展现了其在该任务上的潜力。但由于Transformer通常基于查询-密钥(query-key)获取标记信息(tokens)来对注意力关系来聚合特征,标记信息不匹配将会干扰后续的图像恢复清晰度^[15]。尽管多数多尺度去雨算法都取得了优秀的去雨效果,但是仍存在许多问题:1)直接进行跨尺度之间的信息融合,忽略不同尺度信息之间的联系,会导致背景细节丢失;2)低尺度信息在上采样过程易导致混叠效应^[16],图像会发生失真;3)每个尺度特征的重要性不尽相同,如何将其进行融合,保证信息的利用率的同时防止多余信息冗余,也是研究者们需要关注的问题。

综上,本文提出了一种基于跨尺度注意力融合的

单幅图像去雨算法 (multi-scale attention fusion network, MSAFNet), 针对图像中在空中的动态雨线以及附着在物体表面上的雨滴及雨纹有着显著的去雨效果。为了获取不同尺度信息, 整个去雨网络由多个子网构成。子去雨网络由跨尺度特征提取模块通过稠密连接方式构成, 将提取到的浅层特征传递至深层, 实现特征的高效利用, 防止浅层特征在网络加深的过程中被遗忘。跨尺度特征提取模块利用跨尺度融合注意力, 加强了相邻尺度信息之间的联系, 并对不同尺度的特征进行权重融合。

2 本文方法

本节主要介绍去雨网络及特征提取子网的结构。跨尺度融合注意力作为不同尺度之间的信息联结模块, 其工作原理和结构也会被着重介绍。

2.1 网络结构

本文提出的基于跨尺度注意力融合的单幅图像去雨算法采用并行式的特征提取、融合结构, 能够最大程度提取不同尺度上的雨纹信息, 保留图片细节的同时对不同程度的雨纹进行去除。该网络基于跨尺度特征提取模块, 可以有效地提取三个尺度上的特征信息。针对已有的多尺度去雨网络的不足, 进行了如下设计: 为了解决忽略尺度相关性而导致的图像退化问题, 该模块中用于提取不同分辨率信息的卷积以跨尺度方式

连接, 增强了在不同分辨率下信息的关联。跨尺度连接中添加的注意力模块用于增强相邻尺度之间的特征传播, 通过权重融合的方式实现了跨不同分辨率级别的信息互补, 减少冗余信息对去雨结果的影响。为了解决上采样操作所产生的图像退化问题, 跨尺度融合注意力以高分辨率信息为主导生成融合特征, 防止混叠效应的产生影响图片质量。

所提出的整体网络结构如图1所示。以池化层将输入图片进行下采样能够迅速扩大感受野, 相较于其他下采样方式网络参数更少。不同分辨率的池化特征送入并行式特征提取子网络, 该子网获取不同分辨率下的雨纹以及背景信息。每个相邻子网输出的特征将进入跨尺度融合注意力模块, 低分辨率的特征将对高分辨率特征进行细节补偿, 将权重系数赋予高分辨率特征, 避免了特征之间直接跨尺度融合导致的混叠效应。同时, 跨尺度融合注意力模块降低了冗余信息对去雨效果的影响, 提高了网络信息流通效率。相邻子网之间的输出将由跨尺度融合注意力模块 (multi-scale fusion attention, MFA) 进行细节补偿, 由于跨尺度融合注意力模块输出的信息尺度不一, 需要通过上采样将特征统一至统一分辨率, 上采样操作采用最近邻插值模式, 相比其他上采样模式速度更快, 在输入图像中找到最近的像素值并将其作为输出的像素值。跨尺度融合注意力输出的多通道特征将通过矩阵拼接、

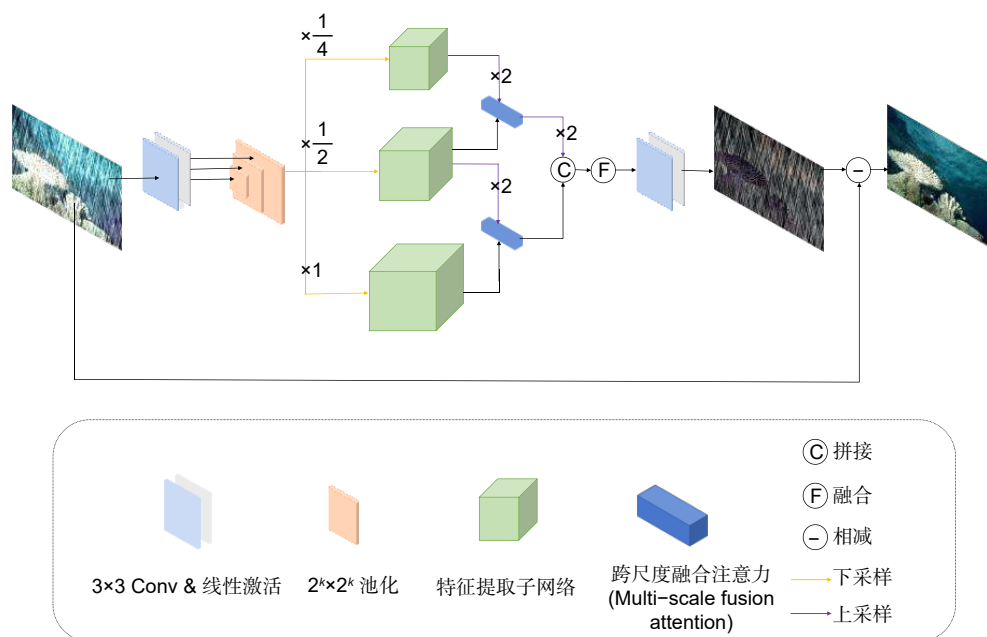


图1 网络结构

Fig. 1 Network structure

融合的方式供后续经过卷积层降维成三通道的特征图, 获取最终的雨纹图像 D , 如式 (1) 所示:

$$D = \mathcal{F}(C[\tilde{D}_1, \mathcal{U}_2(\tilde{D}_{\frac{1}{2}})]), \quad (1)$$

其中: C 表示将不同尺度的特征图在通道维度上拼接, $\mathcal{F}$ 表示将拼接后的特征基于通道维度融合并降维, $\mathcal{U}_2$ 代表系数为 2 的上采样, $\tilde{D}_{\frac{1}{2}}$ 代表在 $\frac{1}{2}$ 尺度大小上 MFA 的输出。

最后的清晰无雨图 O 由输入的带雨图像 X 与特征图 D 通过元素相减的方式获取。操作减去相同分辨率上两个图像对应像素之间的值, O 图中原本雨线位置依然存在图像像素值, 避免了图像引入额外空洞。

$$O = X - D. \quad (2)$$

2.2 跨尺度融合注意力 (MFA)

通过池化操作扩大感受野可以显著减轻网络负担, 降低输入图片的特征尺寸, 但是会导致图片信息丢失, 提取的特征表征能力变差^[17]。为了提升模型对图像信息的理解和利用率以及所提取特征的质量, 许多图像处理任务引入了注意力机制^[18-19]。在本文所提出的去

雨网络中, 由相邻的特征提取子网 (subnet) 输出的特征将通过跨尺度融合注意力模块相连。具有更丰富信息的特征将会被给予更大的权重系数 α , 可以有效地对不同尺度下的去雨结果进行信息补偿。MFA 的结构如图 2 所示。 g 由 $1/(2 \times k)$ 尺度特征提取子网络输出的特征 $D_{\frac{1}{2 \times k}}$ 经过上采样得到, x 代表在 $1/k$ 尺度上特征提取子网络输出的特征 $D_{\frac{1}{k}}$, k 为子网尺度缩放系数, K 为网络中子网的数量, $k = 1, 2, 4, \dots, 2^{K-1}$, 以本文的去雨网络为例, 三个子网提取的分别为 1, $1/2$ 和 $1/4$ 尺度上的信息。 g 与 x 通过元素相加再线性激活的方式, 可以强化特征图对雨纹的检索能力, 去除低分辨率通道中的冗余信息, 如公式 (3) 所示:

$$q_k^{\text{Sub}} = W_{\theta}^T (\sigma_1(W_x^T x + W_g^T g)) + b_{\theta}, \quad (3)$$

其中: σ_1 是 ReLU 函数, W 代表卷积和归一化操作, 其卷积核大小由输入 MFA 的特征维度决定。 b_{θ} 为对应卷积 W_{θ} 的偏置项, 以增强其特征表达能力。 W_g 和 W_x 的作用是将输入的特征图 g, x 进行卷积操作并通过批归一化层进行归一化处理, 以增加模型的稳定性和收敛速度。由于经过归一化处理后的特征不需要额外的偏置项, 且为保证模型轻量化, W_g 和 W_x 在此处不

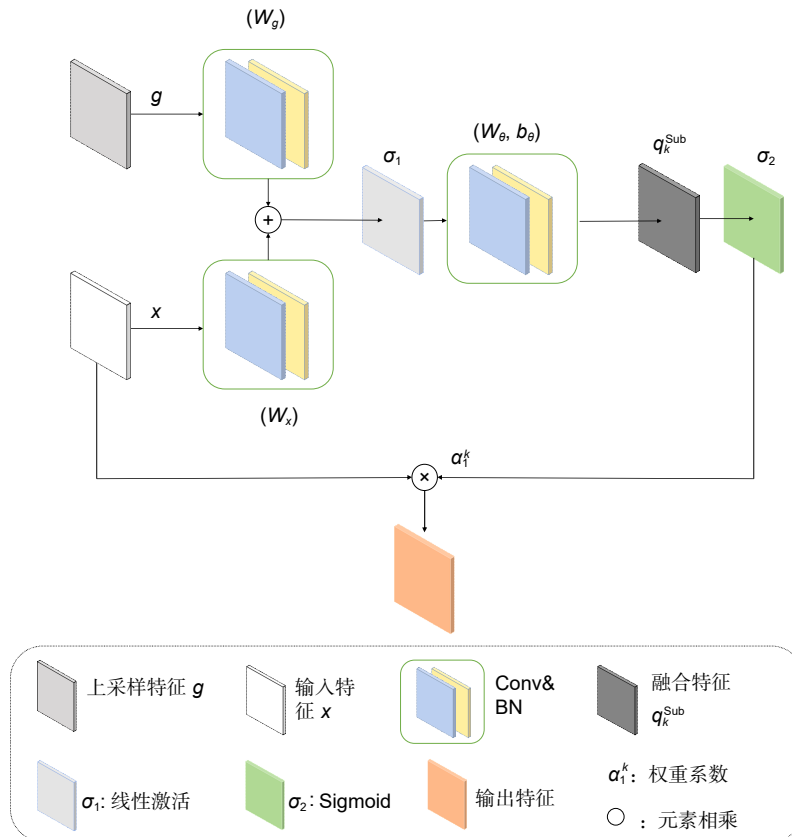


图 2 MFA 结构

Fig. 2 Structure of MFA

引用偏置项。 W_θ 层主要将相加后的特征进行通道降维，将通道数减少为 1 之后获取与输入特征图相同尺寸的权重矩阵，并经过 Sigmoid 函数将权重值的范围限制在 0 到 1 之间，这样可以得到一个像素级别的权重矩阵，用于控制不同位置的特征在加权融合中的贡献程度。

经过注意力模块融合了 $1/k$ 与 $1/(2 \times k)$ 尺度信息的特征用 q_k^{Sub} 表示， q_k^{Sub} 通过 Sigmoid 激活获取权重系数 α_1^k ， D_1 通过与 α_1^k 相乘可以得到来自低分辨率特征的信息补偿，优化高分辨率通道上的特征，用公式表示如下：

$$\alpha_1^k = \sigma_2(q_k^{\text{Sub}}), \quad (4)$$

$$\tilde{x} = \alpha_1^k \times x, \quad (5)$$

式中： σ_2 为 Sigmoid 函数， x 为 $1/k$ 尺度上特征提取子网络的输出， $\tilde{x}$ 代表在 $1/k$ 尺度上经过了跨尺度融合注意力之后输出的特征，MFA 的伪代码如下。

Algorithm 1: 跨尺度融合注意力伪代码

Input: 上采样特征 g ，当前尺度特征 x ，

卷积归一化层 W_g ， W_x ， W_θ ，激活函数 ReLU, Sigmoid

Output: 跨尺度融合特征 $Fusion$

```

1  begin
2     $Feature_g = W_g(g)$ ;
3     $Feature_x = W_x(x)$ ;

```

```

4     $\psi = \text{ReLU}(Feature_g + Feature_x)$ ;
5     $\psi = W_\theta(\psi)$ ;
6     $\alpha = \text{Sigmoid}(\psi)$ ;
7     $Fusion = \alpha * x$ ;
8    return  $Fusion$ ;
9  end

```

为了体现多尺度特征融合的必要性，将跨尺度融合注意力中的输入与输出特征进行可视化分析。从图 3 中可以看出，输入至 MFA 中的高分辨率特征 x 更关注图片中的雨线部分，注重图片的全局信息；经过上采样后的低分辨率特征 g 关注图片中物体的轮廓信息，注意图片中的细节表达。不同尺度的信息帮助 MFA 强化对图片的上下文理解，更好地构建图像中不同物体与结构之间的关系，提取图片中的语义信息，从而解耦雨纹层与背景层。MFA 中所获取的权重系数 α_1^k 作为尺度融合的依据，优化高分辨率特征 x 进一步提取图片中的雨纹信息，输出最终融合特征 $\tilde{x}$ 。

2.3 特征提取子网络

并行结构子网用于提取不同分辨率下的雨纹信息，子网所提取特征的质量决定了生成图片的去雨效果。本节主要介绍图 1 中的特征提取子网络的结构以及其内部组成模块。

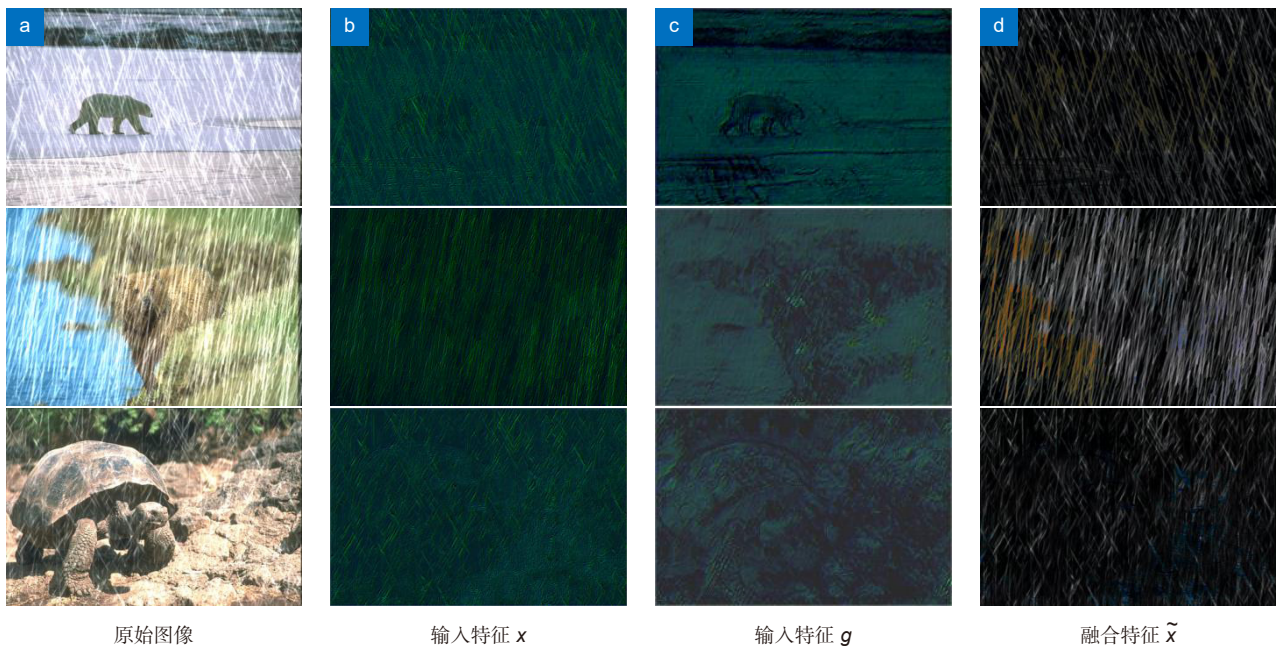


图 3 特征可视化

Fig. 3 Visualization of features

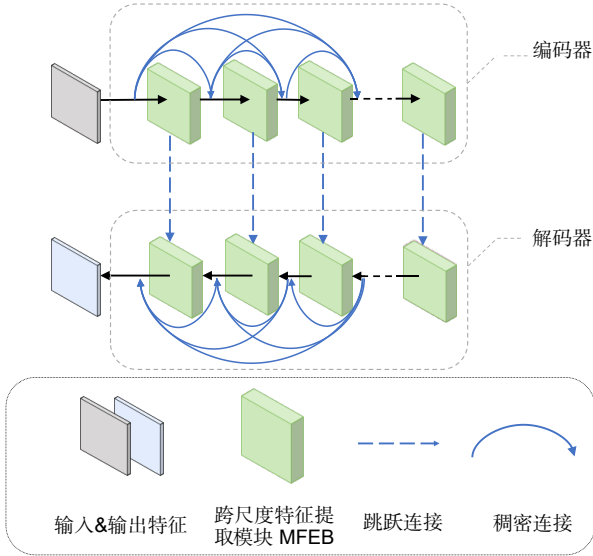


图 4 特征提取子网网络结构

Fig. 4 Feature extraction subnet network structure

2.3.1 网络结构

特征提取子网络以跨尺度特征提取模块 (multi-scale feature extraction block, MFEB) 为核心, 以稠密连接方式^[20]构成了对称式的编码器-解码器结构, 编码器和解码器之间通过跳跃连接完成子网搭建。网络的结构图如图 4 所示。

稠密连接网络通过将每一层与之前所有层在特征维度上连接, 使得特征的复用率最大化, 因此不管是在前向传播或者是后向传播, 都可以使信息和梯度在整个网络中保持更好的流动性, 编码器的数学公式可以表示为

$$F_l^E = \mathcal{T}(\mathcal{F}(C[F_0^E, F_1^E, \dots, F_{l-1}^E])), l = 1, 2, \dots, L, \quad (6)$$

其中: $\mathcal{F}$ 和 C 的操作含义和上一节中的保持一致, L 代表特征提取子网中 MFEB 的数量, $\mathcal{T}$ 代表 MFEB 中进行的操作, F_l^E 代表在第 l 层编码器中 MFEB 的输出。解码器的结构和编码器完全对称, 为的是让它们之间进行一次跳跃连接。编码器和解码器之间采用跳跃连接的方式来融合两个过程中对应的特征图, 可以充分利用浅层特征及上下文信息, 并对雨纹特征的偏移量进行修正, 提升模型的去雨效果。解码器的数学公式可以表示为:

$$F_l^D = \mathcal{T}(\mathcal{F}(C[F_0^D, F_1^D, \dots, F_{l-1}^D])) + F_l^E, l = 1, \dots, L. \quad (7)$$

2.3.2 跨尺度特征提取模块 (MFEB)

尽管多尺度特征提取的方法被广泛地运用在一系列图像处理问题中^[21], 但是大多数方法都忽略了不同尺度特征之间的信息差异, 直接进行多尺度的信息融

合可能会导致去雨结果存在一系列问题: 不同尺度的信息具有不同的重要性和贡献度, 忽视差异可能导致部分重要信息被忽略和丢失, 影响最终结果的准确性和完整性; 不同尺度的信息存在着细节差异或者噪声, 直接融合易导致图片质量下降; 低分辨率的信息通过上采样与高分辨率信息融合易导致混叠效应的产生, 致使图片发生模糊和失真; 不平衡的权重分配导致某些尺度的信息被过分强调, 而其他尺度的信息却被忽略, 最终影响去雨结果。为解决上述问题, 本文设计了跨尺度特征提取模块 (图 5), 该模块为上节中特征提取子网的基础组成模块, 可更有效地利用深层雨纹特征信息, 对特征进行逐尺度的细节补偿, 上采样后的特征引导网络获取高分辨率信息, 降低混叠效应对图像质量的影响。

如图 5 所示, 输入的特征图 x 分别经过不同尺寸的池化操作, 完成了特征在对应尺度上的下采样。

$$P_n = \mathcal{P}_n(x), n = 1, 2, \dots, 2^{N-1}, \quad (8)$$

其中: x 是输入特征, N 表示的是跨尺度特征提取模块中尺度的数量。 $\mathcal{P}_n(\cdot)$ 表示大小为 $n \times n$ 的池化层操作, P_n 是指经过了池化层后初始输出的下采样系数为 $1/n$ 的特征。不同于子网之间将输出端通过 MFA 相连的方式, 不同分辨率通道之间通过 MFA 采取递进式连接, 加深不同尺度之间关联。

通过从低到高的递进式连接重视了不同尺度特征之间的相互影响, 能够引导网络在高分辨率通道提取雨纹特征, 对细节信息进行补偿低分辨率的特征会存在噪声, MFA 可以抑制无关区域中的激活, 降低噪声对雨纹特征的干扰, 减少网络传递过程中信息丢失。低分辨率特征将通过 MFA 分配给高分辨率的初始特征权重, 和权重系数相乘后的初始特征将会更加具有表征能力, 提升特征质量, 式 (9) 所示:

$$\begin{aligned} q_n^{\text{MFFB}} &= W_\theta^T (\sigma_1(W_x^T P_n + W_g^T \mathcal{U}_n(Y_{\frac{1}{2n}}))) + b_\theta, \\ \alpha_n^2 &= \sigma_2(q_n^{\text{MFFB}}) \tilde{P}_n = \alpha_n^2 \times P_n, \\ Y_{\frac{1}{n}} &= \text{Conv}_{3 \times 3}^m(\tilde{P}_{\frac{1}{n}}). \end{aligned} \quad (9)$$

式中: q_n^{MFFB} 代表在 MFEB 中融合了 $1/n$ 与 $1/(2 \times n)$ 尺度信息的特征, $W, \sigma_1, \sigma_2, b_\theta$ 的含义与 2.2 节中表示一致, $\text{Conv}_{3 \times 3}^m$ 表示经过 m 层 3×3 卷积操作。 $\tilde{P}_{\frac{1}{n}}$ 为经过 MFA 权重处理后的特征, $Y_{\frac{1}{n}}$ 是指在 $1/n$ 分辨率上, $\tilde{P}_{\frac{1}{n}}$ 经过 m 层 3×3 卷积后进一步提取的特征, 当 $n = 4$ 时, $\tilde{P}_{\frac{1}{n}} = P_{\frac{1}{n}}$ 。 $\mathcal{U}_n(\cdot)$ 指 n 倍系数的上采样。

卷积通道中输出的特征将会与初始特征通过残差连接的方式融合在一起, 如式 (10) 所示, Z 表示经过

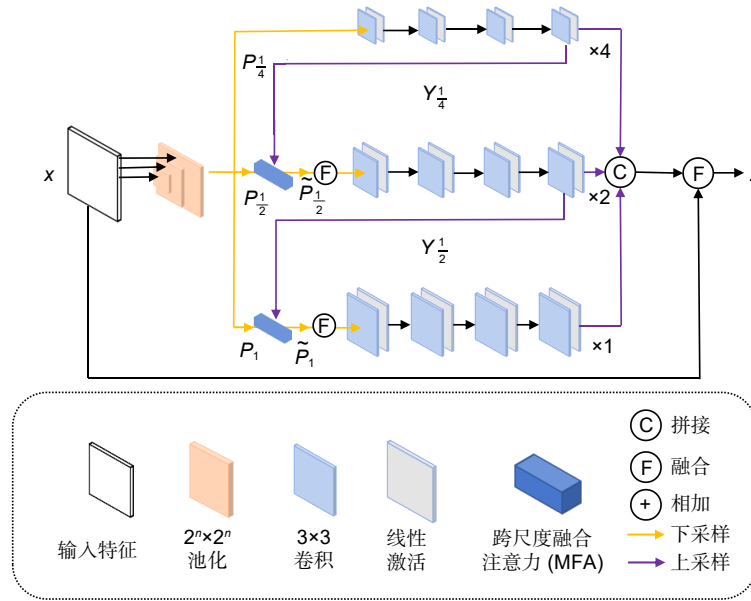


图 5 跨尺度特征提取模块结构

Fig. 5 Multi-scale feature extraction block structure

跨尺度特征提取模块 (MFEB) 后最终输出的特征:

$$Z = \mathcal{F}\left(C\left[Y_1, \dots, \mathcal{U}_{2^{N-1}}\left(Y_{\frac{1}{2^{N-1}}}\right)\right]\right) + x. \quad (10)$$

2.4 损失函数

混合损失函数常被运用在各类图像去雨任务中, 如: MSE+SSIM^[22-24], L_1 +SSIM^[25-26], L_1 + L_2 ^[27], 然而混合的损失函数提升了网络的参数负担, 因此本文采用单一的损失函数训练网络。结构相似度 (structural similarity, SSIM)^[28] 是用于衡量两张图片相似度的图片衡量指标, 两张图片越相似, 该值则越接近 1。其负值常被作为损失函数被运用在去雨任务中, Ren 等人^[7] 证明了其有效性。损失函数如式 (11), 其中: X^{out} 代表去雨网络最终输出的图像, X^{gt} 代表相对应的真实无雨图像。

$$\text{Loss} = -\text{SSIM}(X^{\text{out}}, X^{\text{gt}}). \quad (11)$$

3 实验与结果分析

将本文提出的方法 MSAFNet 在合成数据集 Rain200H^[11], Rain200L^[11], Rain800^[29], DID-MDN^[6] 以及真实数据集 SPA-DATA^[25] 上进行实验, 并与主流方法进行对比验证其有效性。实验对比的去雨方法如基于多尺度特征提取的去雨方法: JORDER^[11], LPNet^[12], MSPFN^[13]; 递归式去雨网络: BRN^[30], RSGN^[31]; 多阶段渐进式去雨网络: PReNet^[7], RCDNet^[32], MPRNet^[33]; 轻量级高效去雨网络

EfDeRain^[26]。

3.1 实验设置及数据集

本文方法在 PyTorch 深度学习框架下使用编程语言 Python 实现。实验环境为 Ubuntu20.04, CUDA 版本 11.4, 使用双卡 NVIDIA 2080Ti GPU 训练网络。Batchsize 设置为 24, 训练过程中将输入随机裁剪成 64×64 的图像补丁。本文的实验迭代轮数为 800, 初始学习率为 5×10^{-4} , 在总迭代次数的 3/5 和 4/5 时学习率乘以 1/10 进行衰减。

本文采用结构相似性 (structural similarity, SSIM) 和峰值信噪比 (peak signal-to-noise ratio, PSNR) 作为客观评价指标。PSNR 和 SSIM 分别用于计算两张图片之间的像素级误差和内容、纹理相似度, 值越大代表两张图片之间越相似, 客观去雨效果越好。

现实世界中较难获取干净无雨图像和有雨图像的成对样本对, 本文采用 Rain200L, Rain200H, Rain800, DID-MDN 四个合成数据集进行训练, 测试网络对不同类型及大小雨纹的去除效果。前两个数据集都包含了 1800 张训练图像和 200 张测试图像。Rain800 包含了 700 张训练图像和 100 张测试图像。DID-MDN 包含了 12000 张训练图像及 1200 张测试图像。其中 Rain200H 数据集最具挑战, 施加的雨水强度最大, 且包含了五种不同的方向。图片的纹理背景复杂, 难以与雨线彻底分离。图片包含的语义信息多样化, 还原成清晰的原始图像较难。具体图例如图 6 所示, 真

实世界数据集 SPA-DATA 进行的实验用于验证本文方法对真实雨图去雨效果的鲁棒性。

3.2 合成数据集实验结果分析

为验证 MSAFNet 的去雨效果, 本文在 Rain200L, Rain200H 和 Rain800 三个数据集上对 MSAFNet 及其它去雨算法进行实验, 实验均在同一环境下进行以确保公平性, 采用 PSNR 和 SSIM 两个客观评价指标评估去雨图像的质量。与其他方法对比的实验结果如表 1 所示。最好结果用粗体表示, 次优结果用下划线表示。从表 1 中可以看出, 得益于多尺度特征信息的利用和注意力融合模块, MSAFNet 在不同数据集上的所获取的评价指标基本优于其他方法。其中 MSAFNet 在 Rain200L 数据集上的优势更明显, 与 2020 年的 BRN 以及 2021 年的 MPRNet 对比, PSNR 分别提升了 1.52 dB 和 1.88 dB, 与同样采样多尺度特

征提取方法的去雨网络 MSPFN 进行横向对比, 平均每个数据集的 PSNR 提升了 4.1%, SSIM 提升了 2.6%。从参数量对比中也可以发现, MSAFNet 以不显



图 6 Rain200H 典型图例

Fig. 6 Rain200H typical picture examples

表 1 在合成数据集上与其他方法的对比

Table 1 Comparison with other methods on synthetic datasets

	Rain200H		Rain200L		Rain800		Params/M
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	
JORDER ^[11]	23.54	0.805	36.11	0.974	23.47	0.869	3.89
LPNet ^[12]	21.96	0.785	32.12	0.955	22.81	0.820	0.008
MSPFN ^[13]	28.64	0.899	37.67	0.975	25.49	0.861	13.35
PReNet ^[7]	27.81	0.891	36.47	0.982	25.18	0.853	0.16
RCDNet ^[32]	29.27	0.899	<u>38.52</u>	<u>0.985</u>	<u>26.38</u>	0.872	2.99
MPRNe ^[33]	<u>29.32</u>	<u>0.915</u>	37.35	0.980	26.10	<u>0.895</u>	20.6
BRN ^[30]	28.96	0.908	37.71	0.983	25.88	0.857	0.39
RSGN ^[31]	28.31	0.905	37.65	0.981	26.37	0.893	4.16
EfDeRain ^[26]	25.54	0.884	31.51	0.957	24.16	0.839	27.61
MSAFNet	29.91	0.921	39.23	0.991	26.63	0.896	12.8

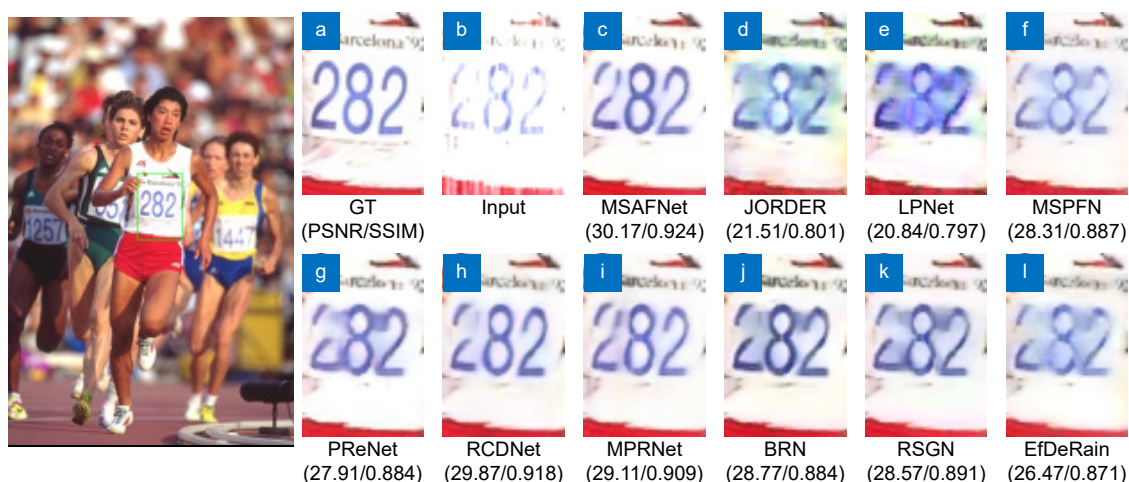


图 7 不同算法在数据集 Rain200H 上的结果

Fig. 7 Results of different algorithms on Rain200H dataset

著增加模型参数量为前提, 极大地提升了模型的性能。

为验证 MSAFNet 的去雨效果与对图像细节的补充, 与其他方法结果通过细节放大方式可视化对比。如图 7 所示, 其中图 7(d) 至图 7(f) 展示不同多尺度去雨网络的结果。MSAFNet 在细节复原和清晰度优于其它多尺度去雨网络, 其中 JORDER 清晰度不够, 与原图存在色差; LPNet 作为轻量级网络存在明显的颜色晕染; MSPFN 号码牌数字较为模糊。图 7(g) 至图 7(i) 展示不同多阶段渐进式去雨网络的结果, 图中可以看出 RCDNet 和 MPRNet 在评价指标中取得了不错的成绩, 但是图片细节的复原程度不如 MSAFNet。图 7(j) 和图 7(k) 展示了不同递归式去雨网络的结果。BRN 的色彩复原程度较差, 对比度偏高; RSGN 存在着晕染问题。图 7(l) 展示的是轻量级的去雨网络 EfDeRain, 其模型注重处理图片的效率, 但图片质量与评价指标偏低。由图 7 可见, 与不同类别的去雨网络进行分类对比, MSAFNet 生成的图片中保留了很多真实图像的细节。图片的纹理以及色彩相较于其它工作更加真实自然, 观感和真实图像接近。

带雨图像形式往往比较复杂, 为了验证模型可以同时处理不同大小与方向上的雨滴, 将其在人工合成数据集 DID-MDN 上进行实验。DID-MDN 数据集可根据图片中的雨线密集程度将其分成大雨、中雨、小雨三组, 每一组图片中存在着多种方向的雨线。

从图 8 中可以看出, 同组图像获得的去雨结果虽然在细节上有着细微的差异: 如图 8(h) 衬衫条纹被抹去, 图 8(j) 中男子头发上残存白色雨线, 但模型最大程度上复原了图片信息与细节, 可见 MSAFNet 对于

处理不同密度, 不同方向的雨线有着较强的稳定性和可靠性。

图 9 显示了 MSAFNet 与另外六个先进的去雨模型在数据集 Rain200L 和 Rain200H 上处理有雨图像时所花费的平均时间与评价指标的对比。为确保实验的公平性, 所有实验数据均来自于同一硬件平台。

从图 9(a) 和 9(b) 中可以看出, MSAFNet 和 RCDNet 在 Rain200L 数据集上在评价指标上的表现优于剩余的模型。EfDerain 处理速度大幅领先剩余模型, 但是去雨效果不佳。Rain200H 数据集中 MSAFNet 与 MPRNet 获得了较高的评价指标, 优于剩余模型。结合两个数据集以及两个评价指标分析, MSAFNet 相比于剩余的去雨模型在评价指标结果上具有更强的稳定性: MPRNet 在 Rain200H 上的表现为次优, 但是在 Rain200L 数据集上的评价指标低于剩余较多的去雨模型; RCDNet 在 Rain200H 数据集中的 SSIM 评价指标低于剩余较多去雨模型。

综合来看, MSAFNet 在不显著增加模型处理有雨图像时间的情况下, 获得了更为稳定且优秀的去雨结果。

3.3 真实数据集上的实验

为了进一步验证 MSAFNet 去雨的鲁棒性以及其自然雨图的泛化性能, 将 Rain200L 数据集上训练的模型在 SPA-DATA 数据集上进行测试, 采用自然图像质量评估方法 (natural image quality evaluator, NIQE)^[34] 和空间-光谱熵质量 (spatial-spectral entropy-based quality, SSEQ)^[35] 两项评价指标对本模型以及其他四个优秀去雨模型的输出进行评估, 其值越小代表



图 8 MSAFNet 在 DID-MDN 数据集上的不同结果

Fig. 8 Different results of MSAFNet on DID-MDN dataset

去雨图像的质量越高。不同算法之间的定量分析结果见表 2。

从表 2 中可以看到, 相比于同样采用多尺度特征提取方法的 MSPFN, 本文方法在 NIQE 以及 SSEQ 的数值上下降了 2.1% 和 4.3%, 存在着较大的提升,

并且结果优于近两年先进的去雨方法 RCDNet 及 MPRNet。

不同算法的自然去雨图可视化分析见图 10。图 10(d) 至图 10(f) 展示了不同基于多尺度特征提取的去雨网络结果, 其中 JORDER 及 LPNet 的去雨效果

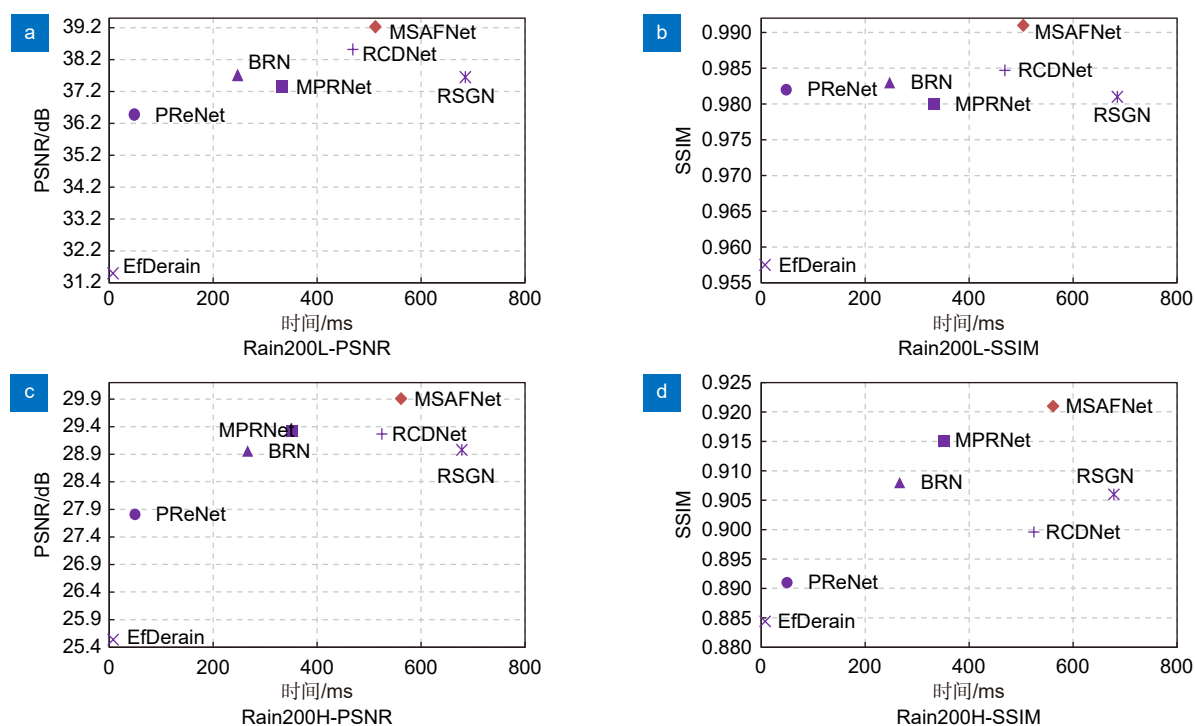


图 9 各模型在 Rain200L 和 Rain200H 数据集上的评价指标与时间效率对比

Fig. 9 Comparison of PSNR/SSIM and time efficiency of each model in Rain200L and Rain200H datasets

表 2 在真实数据集上与其它方法的对比

Table 2 Comparison with other methods on real dataset

方法	MSPFN ^[13]	PReNet ^[7]	RCDNet ^[32]	MPRNet ^[33]	MSAFNet
NIQE↓	4.921	5.013	4.881	4.903	4.816
SSEQ↓	29.26	29.51	28.47	28.67	27.98

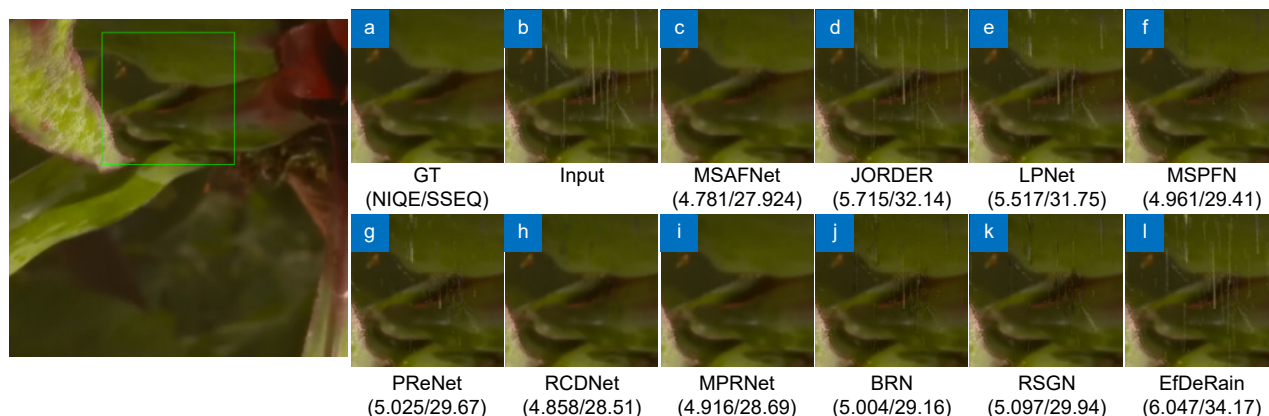


图 10 不同算法在真实数据集 SPA-DATA 上的实验结果

Fig. 10 Experimental results of different algorithms on real-world dataset SPA-DATA

较差, 存在明显雨滴痕迹残留, MSPFN 存在微量残存雨线。图 10(g) 至图 10(i) 展示了不同多阶段式去雨网络的结果, RCDNet 及 MPRNet 获得了良好的视觉效果, 但是评价指标低于 MSAFNet, PReNet 存在明显的雨滴残留。图 10(j) 至图 10(k) 为不同递归式去雨网络的结果, BRN 存在透明绿色空洞, RSGN 去雨不彻底。图 10(l) 为轻量级去雨网络 EfDeRain 的结果图, 能够看出明显的雨线残留。

通过与不同类别的去雨网络进行对比, MSAFNet 凭借其深度多尺度特征提取的优势, 对于图像细节的保留和细微雨滴的清除有着较为明显的优势, 图片真实度高, 对自然图像的泛化性好。

3.4 消融实验

为了证明 MFA 对于图像细节的补充效果, 以及多尺度特征提取子网在 MSAFNet 中的有效性, 本文在 Rain200H 数据集上进行了消融实验。跨尺度特征提取模块中尺度的数量以及模型中子网的数量对去雨效果有着较大的影响, 本文进行了多组实验探究子网数量以及 MFEB 中尺度数量对评价指标的影响。

3.4.1 多尺度子网及 MFA 对去雨效果的影响

如表 3 所示, 与单一子网相比, 多个子网构成的去雨网络可以有效提取多尺度特征, 增强网络的去雨能力。由于子网由多个 MFEB 构成, 在 MFEB 中经由 MFA 融合的特征更具有表达能力, 可以有效提升网络的评价指标。子网间的 MFA 由于数目较少, 可以在不显著增加网络负担的前提下, 对低分辨率冗余信息进行剔除, 优化网络性能。

表 3 在 Rain200H 数据集上消融实验结果

Table 3 Results of ablation experiments on the Rain200H dataset

序号	多尺度子网	MFA (Subnet)	MFA (MFEB)	PSNR/dB	SSIM
1	×	×	×	26.76	0.871
2	√	×	×	27.83	0.881
3	×	×	√	27.92	0.884
4	√	√	×	28.51	0.898
5	√	×	√	29.47	0.916
6	√	√	√	29.91	0.921

3.4.2 子网数量及 MFEB 中尺度数量的研究

表 4 中的 K 代表子网数量, S 代表 MFEB 中尺度的数量。为保证公平表中实验都在同一环境使用同一训练参数进行。从表中可以看出, 当 $K=1$ 和 $S=1$ 时, 由于失去 MFA 对尺度信息的补充, PSNR 值略低。

表 4 子网及 MFEB 中尺度数量对 PSNR 的影响

Table 4 The effect of subnet and MFEB scale number on PSNR

	S=1	S=2	S=3	S=4
K=1	26.57	27.22	27.86	27.79
K=2	27.87	28.24	28.56	28.46
K=3	28.23	29.13	29.91	29.64
K=4	28.47	29.22	29.78	29.51

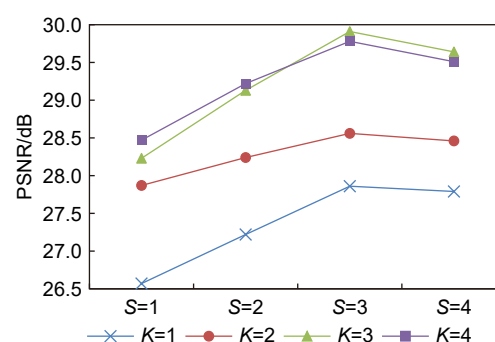


图 11 尺度数量及子网数量折线图

Fig. 11 Curve graph about the number of scales and subnets

当 $K, S=2$ 时, 相较 $K, S=1$ 时, 指标提升量较大。网络在 $K, S=3$ 时到达饱和效果, 再对其进行尺度上的扩充对网络提升效果不大, 再次增加尺度数量甚至出现了负优化。图 11 以折线图的形式直观展现了 K 和 S 对 PSNR 指标的影响。

4 结论

针对多尺度图像去雨网络中存在的混叠效应以及细节丢失等问题, 本文提出了一种基于跨尺度注意力融合的图像去雨算法。网络通过多个子网强化对不同尺度信息的提取能力, 利用跨尺度融合注意力模块解决细节丢失等问题, 根据不同尺度信息之间重要性的不同, 利用低分辨率信息提升高分辨率信息的质量, 显著提升网络性能。本文在多个数据集上进行了实验, 结果表明本文算法能够在去除雨纹的同时保证生成图片的高质量, 保留了图片中的细节信息, 并在客观指标和视觉观感上优于其他去雨算法。

利益冲突: 所有作者声明无利益冲突

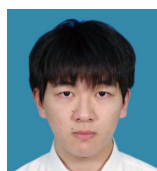
参考文献

- [1] Kang L W, Lin C W, Fu Y H. Automatic single-image-based rain streaks removal via image decomposition[J]. *IEEE Trans Image Process*, 2012, 21(4): 1742–1755.
- [2] Kim J H, Lee C, Sim J Y, et al. Single-image deraining using

- an adaptive nonlocal means filter[C]//*2013 IEEE International Conference on Image Processing*, Melbourne, VIC, Australia, 2013: 914–917. <https://doi.org/10.1109/ICIP.2013.6738189>.
- [3] Li Y, Tan R T, Guo X J, et al. Rain streak removal using layer priors[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 2016: 2736–2744. <https://doi.org/10.1109/CVPR.2016.299>.
 - [4] Eigen D, Krishnan D, Fergus R. Restoring an image taken through a window covered with dirt or rain[C]//*Proceedings of the IEEE International Conference on Computer Vision*, Sydney, NSW, Australia, 2013: 633–640. <https://doi.org/10.1109/ICCV.2013.84>.
 - [5] Fu X Y, Huang J B, Zeng D L, et al. Removing rain from single images via a deep detail network[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 2017: 1715–1723. <https://doi.org/10.1109/CVPR.2017.186>.
 - [6] Zhang H, Patel V M. Density-aware single image de-raining using a multi-stream dense network[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018: 695–704. <https://doi.org/10.1109/CVPR.2018.00079>.
 - [7] Ren D W, Zuo W M, Hu Q H, et al. Progressive image deraining networks: a better and simpler baseline[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, USA, 2019: 3932–3941. <https://doi.org/10.1109/CVPR.2019.00406>.
 - [8] Qian R, Tan R T, Yang W H, et al. Attentive generative adversarial network for raindrop removal from a single image[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018: 2482–2491. <https://doi.org/10.1109/CVPR.2018.00263>.
 - [9] Wei Y Y, Zhang Z, Wang Y, et al. DerainCycleGAN: rain attentive CycleGAN for single image deraining and rainmaking[J]. *IEEE Trans Image Process*, 2021, **30**: 4788–4801.
 - [10] Huang H B, Yu A J, He R. Memory oriented transfer learning for semi-supervised image deraining[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, TN, USA, 2021: 7728–7737. <https://doi.org/10.1109/CVPR46437.2021.00764>.
 - [11] Yang W H, Tan R T, Feng J S, et al. Deep joint rain detection and removal from a single image[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 2017: 1685–1694. <https://doi.org/10.1109/CVPR.2017.183>.
 - [12] Fu X Y, Liang B R, Huang Y, et al. Lightweight pyramid networks for image deraining[J]. *IEEE Trans Neural Netw Learn Syst*, 2020, **31**(6): 1794–1807.
 - [13] Jiang K, Wang Z Y, Yi P, et al. Multi-scale progressive fusion network for single image deraining[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, 2020: 8343–8352. <https://doi.org/10.1109/CVPR42600.2020.00837>.
 - [14] Wang Z D, Cun X D, Bao J M, et al. Uformer: a general U-shaped transformer for image restoration[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, New Orleans, LA, USA, 2022: 17662–17672. <https://doi.org/10.1109/CVPR52688.2022.01716>.
 - [15] Chen X, Li H, Li M Q, et al. Learning a sparse transformer network for effective image deraining[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Vancouver, BC, Canada, 2023: 5896–5905. <https://doi.org/10.1109/CVPR52729.2023.00571>.
 - [16] Liu J J, Hou Q B, Cheng M M, et al. A simple pooling-based design for real-time salient object detection[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, 2019: 3912–3921. <https://doi.org/10.1109/CVPR.2019.00404>.
 - [17] Gao Z T, Wang L M, Wu G S. LIP: local importance-based pooling[C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, Korea (South), 2019: 3354–3363. <https://doi.org/10.1109/ICCV.2019.00345>.
 - [18] Chen L, Zhang J L, Peng H, et al. Few-shot image classification via multi-scale attention and domain adaptation[J]. *Opto-Electron Eng*, 2023, **50**(4): 220232.
陈龙, 张建新, 彭昊, 等. 多尺度注意力与领域自适应的小样本图像识别[J]. *光电工程*, 2023, **50**(4): 220232.
 - [19] Zhao D D, Ye Y F, Chen P, et al. Sonar image denoising method based on residual and attention network[J]. *Opto-Electron Eng*, 2023, **50**(6): 230017.
赵冬冬, 叶逸飞, 陈朋, 等. 基于残差和注意力网络的声呐图像去噪方法[J]. *光电工程*, 2023, **50**(6): 230017.
 - [20] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 2017: 2261–2269. <https://doi.org/10.1109/CVPR.2017.243>.
 - [21] Zhao Y Y, Shi S X. Light-field image super-resolution based on multi-scale feature fusion[J]. *Opto-Electron Eng*, 2020, **47**(12): 200007.
赵圆圆, 施圣贤. 融合多尺度特征的光场图像超分辨率方法[J]. *光电工程*, 2020, **47**(12): 200007.
 - [22] Fan Z W, Wu H F, Fu X Y, et al. Residual-guide feature fusion network for single image deraining[Z]. arXiv: 1804.07493, 2018. <https://doi.org/10.48550/arXiv.1804.07493>.
 - [23] Wu X, Huang T Z, Deng L J, et al. A decoder-free transformer-like architecture for high-efficiency single image deraining[C]//*Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*, Vienna, Austria, 2022: 1474–1480. <https://doi.org/10.24963/ijcai.2022/205>.
 - [24] Chen Z, Bi X J, Zhang Y, et al. LightweightDeRain: learning a lightweight multi-scale high-order feedback network for single image de-raining[J]. *Neural Comput Appl*, 2022, **34**(7): 5431–5448.
 - [25] Wang T Y, Yang X, Xu K, et al. Spatial attentive single-image deraining with a high quality real rain dataset[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, 2019: 12262–12271. <https://doi.org/10.1109/CVPR.2019.01255>.
 - [26] Guo Q, Sun J Y, Xu J F, et al. EfficientDeRain: learning pixel-wise dilation filtering for high-efficiency single-image deraining [C]//*Proceedings of the 35th AAAI Conference on Artificial Intelligence*, 2021: 1487–1495. <https://doi.org/10.1609/aaai.v35i2.16239>.

- [27] Guo X, Fu X Y, Zhou M, et al. Exploring Fourier prior for single image rain removal[C]//*Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*, Vienna, Austria, 2022: 935–941. <https://doi.org/10.24963/ijcai.2022/131>.
- [28] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. *IEEE Trans Image Process*, 2004, 13(4): 600–612.
- [29] Zhang H, Sindagi V, Patel V M. Image de-raining using a conditional generative adversarial network[J]. *IEEE Trans Circuits Syst Video Technol*, 2020, 30(11): 3943–3956.
- [30] Ren D W, Shang W, Zhu P F, et al. Single image deraining using bilateral recurrent network[J]. *IEEE Trans Image Process*, 2020, 29: 6852–6863.
- [31] Wang C, Zhu H H, Fan W S, et al. Single image rain removal using recurrent scale-guide networks[J]. *Neurocomputing*, 2022, 467: 242–255.
- [32] Wang H, Xie Q, Zhao Q, et al. A model-driven deep neural network for single image rain removal[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, 2020: 3100–3109. <https://doi.org/10.1109/CVPR42600.2020.00317>.
- [33] Zamir S W, Arora A, Khan S, et al. Multi-stage progressive image restoration[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, TN, USA, 2021: 14816–14826. <https://doi.org/10.1109/CVPR46437.2021.01458>.
- [34] Mittal A, Soundararajan R, Bovik A C. Making a “completely blind” image quality analyzer[J]. *IEEE Signal Process Lett*, 2013, 20(3): 209–212.
- [35] Liu L X, Liu B, Huang H, et al. No-reference image quality assessment based on spatial and spectral entropies[J]. *Signal Process Image Commun*, 2014, 29(8): 856–863.

作者简介



叶宇超 (1998-), 男, 硕士研究生, 主要研究计算机视觉、图像去雨。

E-mail: 6211924126@stu.jiangnan.edu.cn



【通信作者】陈莹 (1976-), 女, 博士, 教授, 博士生导师。主要研究方向为机器视觉、信息融合、模式识别。

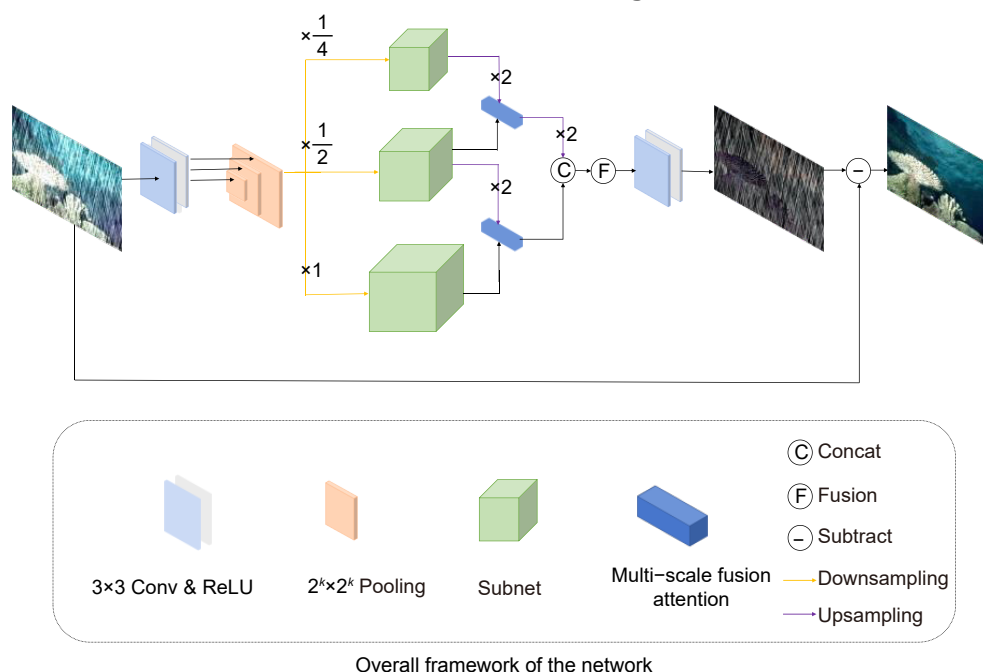
E-mail: chenying@jiangnan.edu.cn



扫描二维码, 获取PDF全文

Single image rain removal based on cross scale attention fusion

Ye Yuchao, Chen Ying*



Overview: Single-image rain removal is an important task in computer vision, aiming to remove rain streaks from rainy images and generate high-quality rain-free images, which has extensive applications in video surveillance analysis and autonomous driving. However, existing rain removal algorithms based on deep learning face challenges in obtaining global information from rainy images, leading to issues such as loss of image details and incomplete rain streak removal. To address these problems, many rain removal algorithms construct multi-scale networks to enhance the detailed information for image deraining. Although these multi-scale deraining algorithms have achieved good results, directly fusing information from different scales without considering the inter-scale relationships may lead to the loss of background details and image distortion during the upsampling process. Therefore, it is important to consider how to establish relationships across different scales to achieve scale feature complementarity, which enables algorithms to balance both details and global information. In response to the above issues, this article proposes an image rain removal network based on cross-scale attention fusion, which aims to remove dense rain streaks while preserving the details of the original image as much as possible, improving the visual quality of the rain removal image. The network is based on a cross-scale feature fusion module, which can effectively extract feature information at three scales. To solve the problem of image degradation caused by neglecting scale correlation, the convolutions used in the module to extract information at different resolutions are connected in a cross-scale manner, enhancing the ability to capture information at different resolutions. The attention module added in cross-scale connections is used to enhance the feature propagation between neighboring scales, achieving information complementarity across different resolution levels. The rain removal network consists of three sub-networks which are composed of densely connected cross-scale feature fusion modules, and each sub-network is used to obtain rain pattern information at different scales. Experimental results demonstrate the effectiveness of the proposed model on synthetic datasets Rain200H and Rain200L. The peak signal-to-noise ratio (PSNR) of derained images reaches 29.91/39.23 dB, and the structural similarity index (SSIM) is 0.92/0.99. These performances outperform the general mainstream methods and achieve better visual effects in terms of preserving image details. In terms of time efficiency, the proposed model also shows advantages compared to some baseline models while ensuring natural deraining effects and maintaining processing speed.

Ye Y C, Chen Y. Single image rain removal based on cross scale attention fusion[J]. *Opto-Electron Eng*, 2023, 50(10): 230191; DOI: 10.12086/oee.2023.230191

Foundation item: Project supported by National Natural Science Foundation of China (62173160)

Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi, Jiangsu 214122, China

* E-mail: chenying@jiangnan.edu.cn