

基于RDSRCNN的单幅图像去雨模型研究

傅继彬, 李春辉

(河南财经政法大学 计算机与信息工程学院, 河南 郑州 450046)

摘要: 下雨天气在图像上造成雨痕, 不仅严重影响图像观感, 更会干扰后续图像分析与处理。图像去雨始终是图像复原研究的热点, 于是提出RDSRCNN单幅图像去雨模型。为提升特征提取能力, 以增强型特征提取方法ESIFEM为特征提取手段, 利用其远距离像素关联能力及低局部特征提取代价实现高效特征提取, 同时利用 l_1 与MSS-SIM构造的复合损失函数优化训练效率并保证输出图像对视觉友好, 将以上方法与增强型DSRCNN去雨网络相结合形成单幅图像去雨模型。在Rain100H数据集上的实验结果表明, 该方法在视觉上能将浓密雨分布情况下的图像恢复至细节丰富的干净场景图, 并且相较于大部分对比方法, 虚影与物体边缘变形的情况减少90%以上, 背景的雨痕清除率高于95%。该方法在量化评估中, 峰值信噪比和结构相似性参数在较大部分较对比方法更优, 且在空间复杂度等方面优于Restormer方法。

关键词: 图像复原; 单幅图像去雨; 深度学习; 混合式损失函数

DOI: 10.11907/rjdk.222120

开放科学(资源服务)标识码(OSID):



中图分类号: TP391.41

文献标识码: A

文章编号: 1672-7800(2023)010-0198-07

Research on De-rain Model for Single Image Based on RDSRCNN

FU Jibin, LI Chunhui

(School of Computer & Information Engineering, Henan University of Economics and Law, Zhengzhou 450046, China)

Abstract: Rainy weather will cause rain marks on the image, which will not only seriously affect the look and feel of the image, but also interfere with subsequent image analysis and processing. Image rain removal has always been the focus of image restoration research. In view of this, a single image rain removal model of RDSRCNN is proposed. In order to improve the feature extraction capability, ESIFEM, an enhanced feature extraction method, was used as the feature extraction means to achieve efficient feature extraction by utilizing its remote pixel correlation capability and low local feature extraction cost. Meanwhile, the loss function constructed with the combination of the l_1 and MSS-SIM was used to optimize the training efficiency and ensure the visual friendliness of the output image. The above method is combined with the enhanced DSRCNN rain removal network to form a single image rain removal model. The experimental results on the Rain100H dataset show that this method can visually restore the image with dense rain distribution to a clean scene map with rich details, and compared with most comparison methods, the cases of ghost and object edge deformation are reduced by more than 90%, and the clearance rate of background rain marks is higher than 95%. In quantitative evaluation, the proposed method is superior to the comparison method in most aspects of peak SNR and structural similarity parameters and is superior to Restormer method in space complexity.

Key Words: image inpainting; single image de-rain; deep learning; mixed loss function

0 引言

雨是自然界中的一种常见气候, 但是雨滴在图片上造成的雨痕不仅会劣化人对图像的视觉感受, 而且由于大量

计算机图像处理算法都假定输入图像清晰且可见度高, 因此图片上的雨痕会对诸如目标检测、目标跟踪、图像分割等图像处理工作造成不良影响。综上, 图像去雨模型研究具有重要现实意义。

对于带有雨痕的图像, 即雨图像, 使用最为广泛的模

收稿日期: 2022-09-23

作者简介: 傅继彬(1975-), 男, 博士, CCF高级会员, 河南财经政法大学计算机与信息工程学院副教授、硕士生导师, 研究方向为人工智能、计算机视觉; 李春辉(1993-), 男, 河南财经政法大学计算机与信息工程学院硕士研究生, 研究方向为人工智能。本文通讯作者: 李春辉。

型^[1-2]为:

$$O = B + R \quad (1)$$

其中, O 为雨图像, B 为所期望获得的无雨背景图像, R 为在背景图像上层的雨痕。图像去雨的目标即为从 O 中分离 B 和 R ,以获得无雨背景图像。由于需要求解的未知项是已知项的2倍,故求解本身并不容易。因此,实现彻底完善的图像去雨是计算机图像处理中一项具有挑战的工作。

从现有文献看,雨痕层 R 主要分为3类,分别是线性形雨痕、离散雨滴^[3]和聚积雨雾^[4]。在方法分类上,之前提出的图像去雨方法大致可划分为视频图像去雨和单幅图像去雨2类,其中视频去雨所需的连续帧图像与图像中物体相对稳定的要求在现实拍摄条件下并不容易满足。而单幅图像去雨不仅应用灵活,且可用于增强视频去雨的单帧图像处理性能。因此,自出现起单幅图像去雨就是视频去雨领域的重要研究方向。随着以视觉Transformer(Vision Transformer, ViT)^[5]为代表以Transformer为基础的图像处理方法的出现,以及摆动Transformer(Swin Transformer, ST)^[6]的提出,以ST为基础构建而成的图像处理方法Swin图像恢复(Swin Image Restoration, SwinIR)^[7]能以更优的效能提取图像内的特征,将其与单幅图像去雨方法结合可以有效提升其效能。

本文提供一种基于DSRCNN的单幅图像去雨模型,利用增强型SwinIR图像特征提取模块优化去雨方法效果,加深DSRCNN网络结构,使用更为合理的损失函数,该方法创新点如下:①利用增强型SwinIR图像特征提取模块(Enhanced SwinIR Image Feature Extraction Module, ESIFEM)取代深度去雨卷积神经网络(Deep Rain Streaks Removal Convolutional Neural Network, DSRCNN)^[8]中的残差块,提取图像中的浅层与深层特征,建立Swin DSRCNN的每级网络,通过更为高效的特征提取优化去雨方法效果,进而通过若干级网络的顺序结合构成残差Swin DSRCNN(Residual Swin DSRCNN, RDSRCNN)网络;②加深DSRCNN网络结构并在除最后一级外的每级网络的输出上应用残差连接,串联本级网络与上一级网络的输出(或原图像),以此增强DSRCNN的去雨能力以及保证输出图像的无雨背景强度,并降低提取特征所需的模型复杂度,以此构成单幅图像DSRCNN去雨模块(Single Image DSRCNN Derain Model, SIDDMM);③在损失函数上, l_1 范数^[9]与MS-SSIM损失函数组成复合损失函数而非单一的 l_2 范数损失函数,这种复合损失函数对图像亮度、颜色和对比度的保留能力优于 l_2 范数损失函数,可增强网络本身的学习效率。

1 相关研究

图像去雨可分为视频去雨和单幅图像去雨两种,其中视频去雨问题最早在文献[10]中被提出。历经十几年发

展,视频去雨领域既有传统的无网络方法^[11-12],也有基于卷积神经网络(Convolutional Neural Network, CNN)的端到端深度学习去雨方法^[12-13]。近期, Jayaraman等^[8]提出一种基于高效视频编码器(High-Efficiency Video Coder, HEVC)后处理优化算法DSRCNN,在此方法中,通过合理利用激活函数等方法避免了CNN在去雨任务下的过拟合问题并加快了网络收敛速度。

单幅图像去雨工作由于其较高的复杂度,研究起步晚于视频去雨。在传统方法中, Kang等^[14]提出通过将去雨问题转换成形态分量分析(Morphological Component Analysis, MCA)问题实现单幅图像去雨痕。该方法利用基于MCA的字典学习和稀疏编码将图像分为有雨和无雨两部分,进而完成去雨。Chen等^[15]提出并建立一种在二维图像上将矩阵变换为张量的低秩雨貌模型以捕获时空相关的雨痕,而后将这种模型泛化到高阶的图像结构(例如图像)上。

2017年后,单幅图像去雨方法中引入了深度学习方法。DerainNet^[16]通过学习无雨图像和雨图像之间的非线性映射关系进行去雨,深度细节网络(Deep Detail Network, DDN)^[17]通过学习负残差细节实现去雨。Yang等^[18]提出联合雨水探测和清除(Joint Rain Detection and Removal, JORDER)的方法,首先利用膨胀网络学习主要特征,然后逐步去雨。随着循环挤压一激励上下文聚合网(Recurrent Squeeze-and-Excitation Context Aggregation Net, RESCAN)^[19]、渐进循环网络(Progressive Recurrent Network, PReNet)^[20]、密度感知的多流密集连接网络(Density-aware Multi-stream Densely Connected Network, DIDMDN)^[21]与空间关注网络(Spatial Attentive Network, SPANet)^[22]等模型提出,深度神经网络已经成为去雨模型的主流算法。

Vaswani等^[23]首次提出Transformer的概念,其在自然语言处理^[24-25]上取得最先进的技术(State of the Art, SOTA)表现。在图像处理方面,Transformer的自注意力机制没有CNN卷积核处理不同的图像区域效果不佳这一问题,而是能捕获全局图像内容间的互动。由此其最初被应用于图像处理^[26]和高级视觉问题处理^[27-28]中。虽然全局自注意力机制能高效捕捉长距离上的像素互动,但其计算复杂度会随着图像分辨率的上升而呈平方级提高。举例而言,一幅大小为 $h \times w$,其图像的全局自注意力处理的计算复杂度 $\Omega(MSA)$ 为(C 为常数):

$$\Omega(MSA) = 4hwC^2 + 2(hw)^2C \quad (2)$$

因此,全局自注意力机制不适用于高分辨率图片。针对这一问题,文献[6-7, 29]给出了大致2种解决方案:文献[6-7]指出只对一个像素四周区域,即一个窗口的区域进行窗口自注意力处理;文献[29]指出将输入图像分解为多个互不重叠的块,而后分别对各块内进行自注意力处理。以上文献指向限制自注意力处理的扩展范围,以求在

捕捉尽可能长距离像素之间关系的前提下实现高效图像处理。

2 RDSRCNN 方法

本文端到端单幅图像去雨方法 RDSRCNN 的去雨网络为 N 级网络,如图 1 所示。整体网络的输入为由单幅雨图像与无雨真值图像所构成的图像对,将图像对中的雨图像输入特征提取模块进行特征提取并去雨,真值图像用于每一级网络。每级网络的输入 $I_n^{input} \in \mathbb{R}^{c \times h \times w}$ 在经过归一化及特征提取后,取得第 n 个特征张量 $I_n^{feature} \in \mathbb{R}^{c_E \times h_E \times w_E}$ 。特

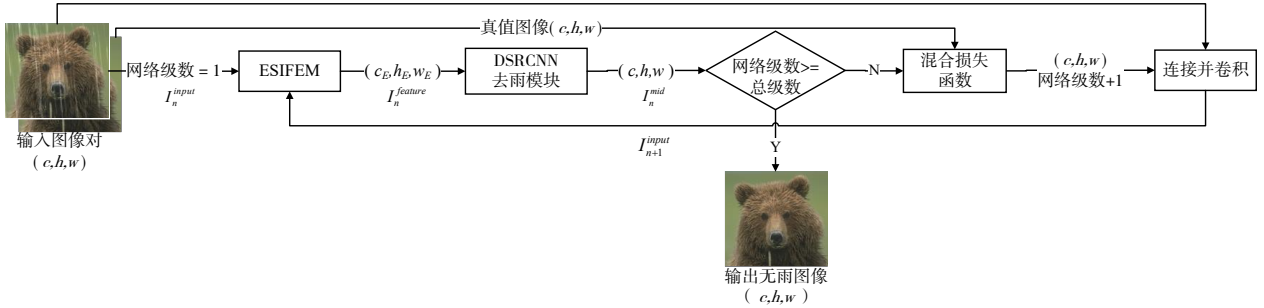


Fig. 1 Structure of the network

图1 网络结构

2.1 ESIFEM

由于文献[8]中 DSRCNN 的原网络主要处理对象是视频图像,如上文所述,视频雨图像的帧间图像具有时空上的连续性,故视频去雨方法可以利用这种连续性以较低的计算代价捕捉雨痕的特征,或者利用背景相对雨痕静止以提取背景特征。对于单幅图像去雨,由于其数据集中的图像之间并不具有时空上的连续性,故若以 DSRCNN 这类视频去雨网络直接处理单幅图像结果会十分不理想,最终通常无法完成去雨任务。为了避免这一问题在 RDSRCNN 中出现,在特征提取方面,本文利用基于 SwinIR 设计的 SwinIR 特征提取模块以提取雨图像中的特征。

增强型 SwinIR 特征提取模块基于 SwinIR 方法设计,其大体结构如图 2(a) 所示,包含浅层特征提取与图 2(b) 中的增强型深层特征提取模块(Enhanced Deep Feature Extraction Module, EDFEM)。浅层特征提取模块与文献[7]中方法相同,使用 3×3 卷积层 $H_{SF}(\cdot)$ 从输入图像 $I \in \mathbb{R}^{H \times W \times C_{in}}$ 中提取出浅层特征 $F_0 \in \mathbb{R}^{H \times W \times C}$,即:

$$F_0 = H_{SF}(I) \quad (3)$$

H 、 W 、 C_{in} 、 C 分别是输入图像的长、宽、通道数和特征张量的通道数。EDFEM 以增强型 RSTB(Enhanced Residual Swin Transformer Block, ERSTB)层(图 2(c))为基础构建,每个 ERSTB 都由受到文献[6]的启发而构建的增强型 STL(Enhanced Swin Transformer Layer, ESTL)(图 2(d))和全局 Transformer 层(Global Transformer Layer, GTL)构成(图 2(e))。ESTL 内部包含 2 组自注意力处理,其中一组以窗口多头自注意力(Window Multi-head Self Attention, W-

征张量的大小 $c_E \times h_E \times w_E$ 根据方法中人为设置的投影维度(本文为 $90 \times 16 \times 16$)确定。而后将特征输入 DSRCNN 去雨模块进行处理,处理结果为第 n 个中间结果图像 $I_n^{mid} \in \mathbb{R}^{c \times h \times w}$,然后将中间结果图像与图像对中的真值图像输入混合损失函数 L^{mix} 并进行反向传播。如此,整个网络中的一级处理结束,若网络级数低于设置数量(本文中为 4),则下一级网络的输入为 $I_{n+1}^{input} = conv(I_n^{input} \oplus I_n^{mid})$,即对本级网络输入与上一级输出连接 1×1 卷积,使用卷积是为了统一输入尺寸。对于没有上一级的第一级网络, I_{n-1}^{input} 为图像对中的单幅雨图像。输出结果则为网络的最终输出结果,即无雨图像 $I^{derain} \in \mathbb{R}^{c \times h \times w}$ 。

MSA)处理特征张量以提取窗口内特征,另一组以 Swin 窗口自注意力(Swin Window Multi-head Self Attention, SW-MSA)处理第一组的输出以保证窗口自注意力处理不会忽略窗口周围的像素对窗口内图像的影响。在每个残差 Swin Transformer 块(Residual Swin Transformer Block, RSTB)的末尾包含以整个输入张量而非张量窗口为处理对象的全局多头自注意力(Global Multi-head Self Attention, G-MSA)的 GTL 而非 ESTL 作为尾端处理模块以较低代价避免 W-MSA 和 SW-MSA 造成的全局像素关系信息缺失。

若以 $H_{ERSTB_k}(\cdot)$ 表示第 k 个 ERSTB 模块, $H_{CONV_{ERSTB}}(\cdot)$ 为模块尾的卷积层。以 $H_{STL_{k,l}}(\cdot)$ 、 $H_{GTL_k}(\cdot)$ 、 $H_{CONV_{STL-k}}(\cdot)$ 分别表示第 k 个 ERSTB 中的第 l 个 ESTL, k 个 ERSTB 中的 GTL 以及第 k 个 RSTB 中的卷积层。以 X_{W-MSA} 、 X_{SW-MSA} 、 X_{G-MSA} 表示 W-MSA、SW-MSA 和 G-MSA 的输出, $X_{MLP_{ESTL-i}}$ 为 ESTL 中第 i 个 MLP 的输出, $X_{MLP_{GTL}}$ 为 GTL 中的 MLP 输出。则 ESIFEM 的输出 F_{ESIFEM} :

$$F_{ESIFEM} = H_{CONV_{ERSTB}}(F_K), F_k = H_{ERSTB_k}(F_{k-1}), k = 1, 2 \dots K \quad (4)$$

$$F_k = H_{CONV_{ESTL-k}}(F_{CTL_k}) + F_{k-1}, F_{CTL_k} = H_{GTL_k}(F_{k,L}), F_{k,L} = H_{ESTL_{k,l}}(F_{k,L-1}), l = 1, 2 \dots L \quad (5)$$

$$F_{CTL_k} = Conv(MLP(LN(X_{G-MSA})) + X_{G-MSA}), X_{G-MSA} = G-MSA(LN(F_{k,L})) + F_{k,L} \quad (6)$$

$$F_{k,L} = MLP(LN(X_{SW-MSA})) + X_{SW-MSA}, X_{SW-MSA} = SW-MSA(LN(X_{MLP_{ESTL-i}})) + X_{MLP_{ESTL-i}} \quad (7)$$

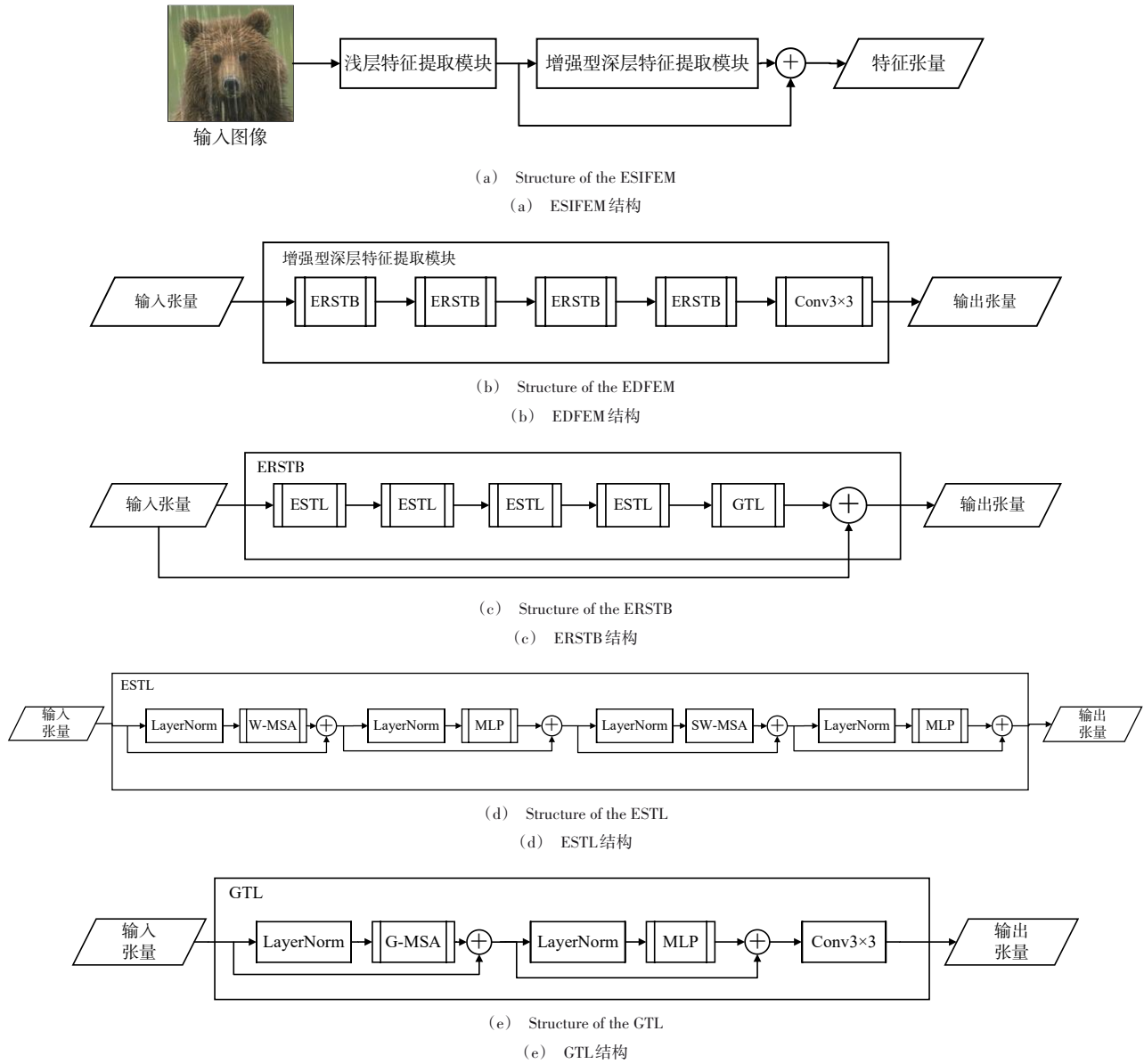


Fig. 2 Structure of ESIFEM and its internal models

图 2 ESIFEM 及其各内部模组结构

$$X_{MLP_{ESTL-l}} = MLP(LN(X_{W-MSA})) + X_{W-MSA}, X_{W-MSA} = W - MSA(LN(F_{k,l-l})) + F_{k,l-l} \quad (8)$$

式(4)~式(8)表述的是 EDFEM、ERSTB、GTL 和 ESTL 处理其输入特征张量的过程。因此得知,利用 SwinIR 在特征提取阶段对于去雨任务的泛化能力,不仅保留了 ESIFEM 对图像特征的高效提取能力,而且通过改进 SwinIR 中的模块而设计的 ESTL 与 GTL 使 ESIFEM 在维持 ST 类图像特征提取方法的低复杂度前提下,还获得了利用窗口外像素辅助进行特征以及一定的全局像素间关系特征的提取能力。 F_{ESIFEM} 的上述特征,给予了 RDSRCNN 处理单幅雨图像的数据基础。

2.2 SIDDM

为使视频去雨方法 DSRCNN 能够有效处理单幅雨图像, SIDDM 对原始 DSRCNN^[8](以下称为原始去雨网络)做

了 3 项改进,即增强网络学习能力、输出迭代和残差连接处理。

在增强网络学习能力方面,以卷积—激活—池化的方式构建复合卷积层代替原始去雨网络中单纯的所有卷积层处理。其目的首先是增强初始化模块对于不同方向、密度和粗细雨痕的探测能力,其次是增加原始去雨网络中负责增强学习能力的批归一化层的深度,利用复合卷积层与池化层使其变为一个 3 层并联深度网络模块,以多种卷积核大小(这里使用的是 $([h_{prev}/2] + 1, [w_{prev}/2 + 1])$) 处理多尺度雨痕问题。

对于原始去雨网络而言,由于输入的帧间图像为多个区分度较低的背景雨图像,若对结果进行迭代处理可能会由于对某些特征的过度学习,加重原始去雨网络的过拟合倾向,故不对去雨结果进行迭代处理。由于 RD-

SRCNN 的输入为除雨痕特征外互不相关的图片,对于输出的迭代不会引起与原始去雨网络一样的严重过拟合情况。相反,由于每个 SIDDMM 输出的都是相对其输入图像的雨痕弱化图像,通过残差连接对雨痕进行增强后输入下一个 SIDDMM 中进行处理即可在维持本 SIDDMM 的雨痕提取效果的前提下通过增强不易被捕捉雨痕的方法提升 SIDDMM 最终雨痕捕捉比例与效果,也即通过逐级增强 SIDDMM 的去雨效果最终达成 RDSRCNN 相对于原始去雨网络在单幅图

像去雨上的性能优势。

最后,通过池化层统一输入特征张量,初始化模块处理张量以及改进归一化模块输出张量的大小,并将整合特征张量输入 3 个串联线性层。由于无论是 ESIFEM 或是 SIDDMM 都更倾向于提取图像中的高频特征,即雨痕,故线性层输出的是雨痕图像 $I_n^{RT} \in \mathbb{R}^{c \times h \times w}$,用输入的雨图像 I_n^{input} 与雨痕图像作差即可输出本 SIDDMM 的雨弱化图像(或最终去雨图像),SIDDMM 的结构如图 3 所示。

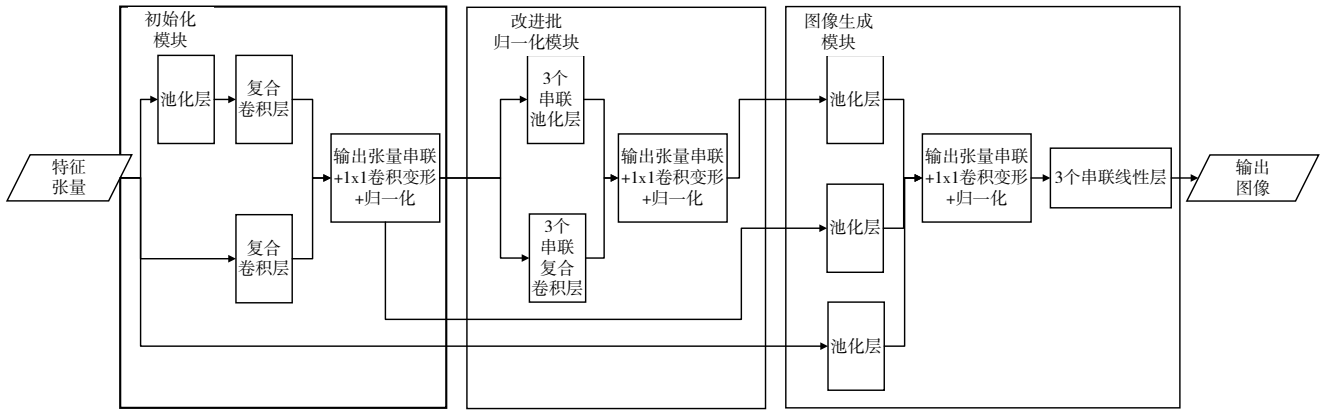


Fig. 3 Structure of SIDDMM

图3 SIDDMM 结构

2.3 损失函数

RDSRCNN 方法采用 Zhao 等^[30]提出的混合损失函数 L^{Mix} 。结构相似性 (Structural Similarity Index Measure, SSIM) 提出自文献[31], SSIM 及其为基础的损失函数是针对视觉感受的感知驱动函数。为了克服 SSIM 作为损失函数对于高斯滤波器标准差的敏感,文献[32]提出 SSIM 的多尺度版本 MS-SSIM (Multi-Scale SSIM, MS-SSIM), 若将尺度设置为 M 级, 某一像素位置的 MS-SSIM 为:

$$MS-SSIM(p) = \left(\frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \right)^\alpha \cdot \prod_{j=1}^M \left(\frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \right)^{\beta_j} \quad (9)$$

式(9)中, μ_x, μ_y 分别为 2 张图片在该位置的均值, σ_x, σ_y 分别为 2 张图片在该位置的方差, σ_{xy} 为联合方差, C_1, C_2 为常数, α 和 β_j 统一设定为 1。由式(9)可得, 以 MS-SSIM 为基础的损失函数在以像素 \tilde{p} 为中心的图像块 P 上的损失为:

$$Loss^{MS-SSIM}(P) = 1 - MS-SSIM(\tilde{p}) \quad (10)$$

不同于 l_2 范数, l_1 范数作为损失函数时不会过度惩罚大误差而容忍小误差。 l_1 范数的损失函数为:

$$Loss^{l_1}(P) = \frac{1}{N} \sum_{p \in P} |x(p) - y(p)| \quad (11)$$

式(11)中, p 为图像块 P 上像素的索引。 $x(p)$ 与 $y(p)$ 是 2 个输入图像块像素 p 的值。

通过联合 $Loss^{l_1}$ 与 $Loss^{MS-SSIM}$ 为 L^{Mix} 使用, 可以同时保留对比度信息、颜色和亮度信息:

$$L^{Mix} = a \cdot Loss^{MS-SSIM} + (1 - a) \cdot Loss^{l_1} \quad (12)$$

根据文献[33]所述, $a = 0.84$ 时效果最佳。相较于 l_2

范数损失函数, 文献[33]中证明混合损失函数 L^{Mix} 在多种图像处理任务上的效能更优。

3 实验与结果分析

3.1 实验环境

为验证本文所提方法的效能, 在 Rain100H 数据集上利用 DerainNet、DID-MDN 与 RDSRCNN 分别进行训练与测试, 给出实验结果并进行效能对比。参与对比的所有方法都统一在 1 800 张图像的训练集上训练 300 轮后, 再根据训练模型进行单幅图像去雨以测试各方法效率。

实验在 Google Colab 上进行。实验软件环境为: Python 3.6, Pytorch 1.12.1, CUDA 11.2; 实验硬件环境为: 内存 52 GB, CPU 主频 2.30 GHz, GPU Tesla P100。

3.2 实验结果评估

对比实验评估分为两部分: ①量化评估, 通过峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 与 SSIM 两个数值, 统计所有方法的去雨表现; ②主观视觉评估, 通过肉眼观察处理图像是否有明显的模糊、伪影、斑纹等缺陷。通过在数值上与视觉上对各方法进行分析以获得较为全面的评价。

3.2.1 量化评估

在量化评估方面, 各对比方法的 PSNR 和 SSIM 值如表 1 所示。

表 1 中的前 3 行展示了 Derain、DID-MDN 和 Restormer^[33] 在 Rain100H 上训练 300 轮后结果的 PSNR 及 SSIM 值。可以看到, 除 Restormer 外的方法在规定训练轮数下的量

Table 1 PSNRs and SSIMs of comparison methods

表 1 各对比方法的PSNR与SSIM值

方法	PSNR	SSIM
DerainNet	15.01	0.61
DID-MDN	14.75	0.59
Restormer	29.02	0.89
RDSRCNN(本文方法 2 × ERSTB)	20.03	0.73
RDSRCNN(本文方法 2 × 总级数)	21.90	0.77
RDSRCNN(本文方法 MSE)	23.06	0.75
RDSRCNN(本文方法 完整网络)	23.07	0.81

化结果,即 PSNR 与 SSIM 均不如最后一行的 RDSRCNN, Restormer 方法虽然在量化结果上具有完全优势,但是在训练期间其内存占用远超 RDSRCNN(内存占用平均比例: Restormer70%, RDSRCNN 50%)且 RDSRCNN 的结构较为简单,更为轻量级,更易于实现。

为了评估本方法中 ERSTB 的网络块数量对于网络去雨性能的影响,在实验中通过分别削减 EDFEM 中 ERSTB 数量为默认值的 50%(每个 EDFEM 中包含 2 个 ERSTB 模块)与削减整个网络中网络级数为默认值的 50%(总级数为 2)进行训练与去雨测试。测试结果见表 1 中 4-5 行,可以看到相对于不进行任何模块与级数削减的完整方法,对 ERSTB 与总网络级数的削减都会导致网络性能下降,也从侧面证明 ERSTB 与适当网络深度对本文网络整体性能的重要性。为评估混合损失函数 L^{Mix} 对于网络性能的影响,实验中在保留完整网络的网络结构的前提下,将 L^{Mix} 换为均方误差损失函数(MSE)进行训练与去雨测试。结果如表 1 第 6 行所示,可以看到在规定条件下,其性能仍不及使用 L^{Mix} 的完整网络,证明本文使用的混合损失函数相较于一般的损失函数 MSE 具有更高的性能。

3.2.2 可视化评估

在可视化评估方面,由图 4 可以看出 Derain、DID-MDN 中的前景仍有未去除的雨痕(0-12 处),除 RDSRCNN 和 Restormer 外没有一个对照方法完成了背景雨痕消除工作。RDSRCNN 的去雨结果显示,图像前景几乎无雨痕(0-1 处),背景雨痕去除率高于 90%,但存在亮度过高的问题。Restormer 虽然去雨效果较为完美,但空间复杂度性能不如 RDSRCNN。

由量化评估与可视化评估结果可以看出,RDSRCNN 作为单张图像去雨方法对系统资源要求较低的同时保证了一定性能,在系统资源有限的场景中有其应用价值。

4 结语

RDSRCNN 方法在网络结构上结合了基于 SwinIR 特征提取模块设计的 ESIFEM 和基于 DSRCNN 设计的 SIDDM,利用 ESIFEM 基于 ST 的特征提取能力弥补了单幅图像相对于视频图像缺乏帧间图像关系的缺陷,利用增强网络学习能力、处理结果迭代和残差连接等方法保证了单幅图像在 SIDDM 上的去雨效果。通过应用更为有效的混合损失

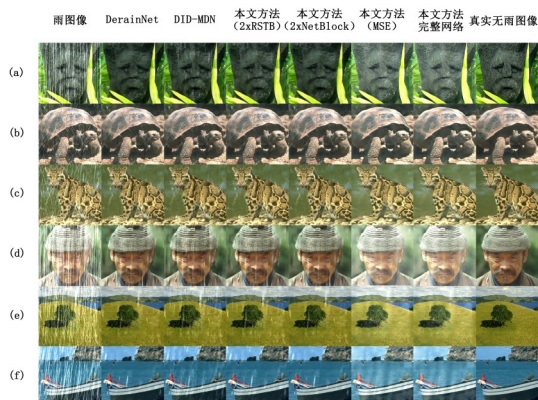


Fig. 4 Visual comparison of processing results by comparison methods

图 4 各对比方法处理结果可视化比较

函数 L^{Mix} ,提升了去雨网络量化评价。并且,通过实验证明了本文提出的增强型方法及其中模块的有效性与当前先进方法相比具有更强的经济性。近年来,随着社会智能化的演进,图像去雨的应用范围也将越来越广泛,将具有重大理论与应用价值,本文方法对于同类研究也颇具借鉴意义。

参考文献:

- [1] HUANG D A, KANG L W, YANG M C, et al. Context-aware single image rain removal [C]// IEEE International Conference on Multimedia and Expo, 2012: 164-169.
- [2] LI Y, TAN R T, GUO X, et al. Rain streak removal using layer priors [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2736-2744.
- [3] YOU S, TAN R T, KAWAKAMI R, et al. Adherent raindrop modeling, detection and removal in video [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 38(9): 1721-1733.
- [4] ZHANG Z, LIU W, MA H, et al. Going clear from misty rain in dark channel guided network [C]// Melbourne: International Joint Conference on Artificial Intelligence-Workshop on AI for Internet of Things, 2017.
- [5] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: transformers for image recognition at scale [C]// Vienna: International Conference on Learning Representations, 2021.
- [6] LIU Z, LIN Y, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 10012-10022.
- [7] LIANG J, CAO J, SUN G, et al. Swinir: image restoration using swin transformer [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 1833-1844.
- [8] JAYARAMAN T, CHINNUSAMY G S. Analysis of deep rain streaks removal convolutional neural network-based post-processing techniques in HEVC encoder [J]. Journal of Circuits, Systems and Computers, 2021, 30(2): 2150020.
- [9] ZHAO H, GALLO O, FROSIO I, et al. Loss functions for image restoration with neural networks [J]. IEEE Transactions on Computational Imaging, 2016, 3(1): 47-57.
- [10] GARG K, NAYAR S K. Detection and removal of rain from videos [C]// Washington: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004.

- [11] ZHANG X, LI H, QI Y, et al. Rain removal in video by combining temporal and chromatic properties [C]// IEEE International Conference on Multimedia and Expo, 2006: 461–464.
- [12] KIM J H, SIM J Y, KIM C S. Video deraining and desnowing using temporal correlation and low-rank matrix completion [J]. IEEE Transactions on Image Processing, 2015, 24(9): 2658–2670.
- [13] LIU J, YANG W, YANG S, et al. D3r-net: dynamic routing residue recurrent network for video rain removal [J]. IEEE Transactions on Image Processing, 2018, 28(2): 699–712.
- [14] 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, 2011, 21(4): 1742–1755.
- [15] CHEN Y L, HSU C T. A generalized low-rank appearance model for spatio-temporally correlated rain streaks [C]// Proceedings of the IEEE International Conference on Computer Vision, 2013: 1968–1975.
- [16] FU X, HUANG J, DING X, 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.
- [17] FU X, HUANG J, ZENG D, et al. Removing rain from single images via a deep detail network [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 3855–3863.
- [18] YANG W, TAN R T, FENG J, et al. Deep joint rain detection and removal from a single image [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1357–1366.
- [19] LI X, WU J, LIN Z, et al. Recurrent squeeze-and-excitation context aggregation net for single image deraining [C]// Proceedings of the European Conference On Computer Vision (ECCV), 2018: 254–269.
- [20] REN D, ZUO W, HU Q, et al. Progressive image deraining networks: a better and simpler baseline [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 3937–3946.
- [21] ZHANG H, PATEL V M. Density-aware single image de-raining using a multi-stream dense network [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 695–704.
- [22] WANG T, 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, 2019: 12270–12279.
- [23] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]// Proceedings of the 31st International Conference on Neural Information Processing System, 2017: 6000–6010.
- [24] BROWN T, MANN B, RYDER N, et al. Language models are few-shot learners [J]. Advances in Neural Information Processing Systems, 2020, 21(33): 1877–1901.
- [25] LIU Y, OTT M, GOYAL N, et al. Roberta: a robustly optimized bert pretraining approach [DB/OL]. <https://arxiv.org/abs/1907.11692>.
- [26] WANG Z, CUN X, BAO J, et al. Uformer: a general u-shaped transformer for image restoration [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 17683–17693.
- [27] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C]// European Conference on Computer Vision, 2020: 213–229.
- [28] TOUVRON H, CORD M, DOUZE M, et al. Training data-efficient image transformers & distillation through attention [C]// International Conference on Machine Learning, 2021: 10347–10357.
- [29] CHEN H, WANG Y, GUO T, et al. Pre-trained image processing transformer [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 12299–12310.
- [30] ZHAO H, GALLO O, FROSIO I, et al. Loss functions for image restoration with neural networks [J]. IEEE Transactions on Computational Imaging, 2016, 3(1): 47–57.
- [31] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600–612.
- [32] WANG Z, SIMONCELLI E P, BOVIK A C. Multiscale structural similarity for image quality assessment [C]// The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003: 1398–1402.
- [33] ZAMIR S W, ARORA A, KHAN S, et al. Restormer: efficient transformer for high-resolution image restoration [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 5728–5739.

(责任编辑:孙娟 周星宇)