

基于双注意力残差循环单幅图像去雨集成网络*

张学锋, 李金晶



(安徽工业大学 计算机科学与技术学院, 安徽 马鞍山 243000)

通讯作者: 李金晶, E-mail: 2424899546@qq.com

摘要: 降雨会严重降低拍摄图像质量和影响户外视觉任务. 由于不同图像中, 雨的形状、方向和密度不同, 导致单幅图像去雨是一项困难的任务. 提出一种新的基于双注意力的残差循环单幅图像去雨集成网络(简称 RDARENet). 在网络中, 因为上下文的信息对于去除雨痕十分重要, 所以首先采用多尺度的扩张卷积网络去获得更大的感受野. 雨痕信息可以认为是多个雨层特征的叠加, 为了更好地提取雨痕的特征和恢复背景图层信息, 运用了通道和空间注意力机制的残差网络. 通道注意力能够反映不同雨层的权重, 而空间注意力则通过相邻空间特征之间的关系增强区域的表征. 随着网络的加深, 防止底层信息的丢失, 采用级联的残差网络和长短时间记忆网络, 将底层特征信息传递到高层中去, 逐阶段地去除雨痕. 在网络的输出部分, 采用集成学习的方式, 将每个阶段的输出结果通过门控网络加权相加, 得到最终的无雨图像. 实验结果表明, 去雨和恢复纹理细节的效果都得到较大提升.

关键词: 单幅图像去雨; 双注意力机制; 残差网络; 门控网络

中图法分类号: TP391

中文引用格式: 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. 软件学报, 2021, 32(10): 3283-3292. <http://www.jos.org.cn/1000-9825/6018.htm>

英文引用格式: Zhang XF, Li JJ. Single image de-raining using a recurrent dual-attention-residual ensemble network. Ruan Jian Xue Bao/Journal of Software, 2021, 32(10): 3283-3292 (in Chinese). <http://www.jos.org.cn/1000-9825/6018.htm>

Single Image De-raining Using a Recurrent Dual-attention-residual Ensemble Network

ZHANG Xue-Feng, LI Jin-Jing

(School of Computer Science and Technology, Anhui University of Technology, Maanshan 243000, China)

Abstract: Rain streaks can severely degrade the quality of captured images and affect outdoor vision. However, due to non-uniform in shape, direction, and density of rain in different images, it is a difficult task to remove rain from a single image. This study proposes a single image de-raining using an ensemble recurrent dual-attention-residual network, called RDARENet. In the network, as contextual information is very important for the process of rain removal, a multi-scale dilated convolution network is firstly adopted to acquire large receptive field. Rain streaks can be regarded as the accumulation of multiple rain streaks layers, the residual of the channel attention and spatial attention mechanisms are used to extract the features of the rain streaks and restore the background layer information. The channel attention can assign different weights to rain streaks layers, and the spatial attention enhances the representation of the area through the relationship between adjacent spatial features. With the deepening of the network, to prevent the loss of low-level information, a cascaded residual network and a long-term memory network are used to transfer low-level feature information to the high-level and remove rain streaks stage by stage. In the output of the network, ensemble learning method is adopted to weight the output of each stage through the gated network, and add to get the clean image. Extensive experiments demonstrate that the effect of removing rain and restoring texture details is greatly improved.

Key words: single image de-raining; dual-attention mechanism; residual network; gated network

* 基金项目: 安徽省教育厅重大课题(KJ2017ZD05); 安徽高校协同创新项目(GXXT-2019-008)

Foundation item: Major Issues of the Anhui Provincial Department of Education (KJ2017ZD05); Anhui University Collaborative Innovation Project (GXXT-2019-008)

收稿时间: 2019-08-11; 修改时间: 2019-11-24, 2020-01-04; 采用时间: 2020-02-07

雨痕会降低图像或者视频的质量,造成一些计算机视觉任务不能正常工作,例如视频监控、自动驾驶、物体检测和图像分类等领域.如何有效地去除图像和视频中的雨痕,一直是计算机视觉方向的研究热点.本文主要关注单幅图像的去雨任务.

在过去的几年中,研究者们提出一些基于图像先验的方法来处理去雨问题,例如高斯混合模型(GMM)^[1]、代表学习(representation learning)^[2]和字典学习(dictionary learning)^[3-5].这些方法认为:雨图像是背景图层和雨痕层相结合的模型,通过将雨痕层从雨图像中分离出来,得到无雨的背景图层.但是由于雨痕和背景图层的复杂组合或者因大雨而产生的薄雾等情况,导致在实际中去雨的效果并不是很好.随着深度学习在底层视觉领域的应用^[6-8],单幅图像去雨再次在深度卷积网络(CNN)的基础上取得了突破性的进展.Fu 等人^[9]首先提出应用三层卷积网络结构直接学习雨图像和无雨图像细节之间的映射关系,通过在高层的细节中移除雨痕,得到无雨的背景图像.随后,在卷积神经网络的基础上,许多模块也被应用在去雨网络中,例如密集连接模块^[10]、残差模块^[11]、扩张卷积^[12]、空间注意力^[13]、多尺度信息^[12,14]和挤压与激励模块^[14,15].通过将模块嵌入去雨网络,使得网络能够更好地去除雨痕和恢复背景的细节信息.在其他方面,一些文献^[12,14-16]通过实验发现:构建多个阶段来逐步地去除雨痕,会提高去雨的效果.

结合之前学者的工作,本文提出了基于双注意力的残差循环单幅图像去雨集成网络(RDARENet)^[17].网络架构如图 1 所示:去除雨痕的过程分成多个阶段来完成,在每个阶段中,首先使用多尺度的平滑扩张卷积^[18],它能够增大网络感受野,而且极大地减小网格伪影的产生,通过不同的扩张因子卷积聚合特征信息来逐步地提取和细化特征;同时,为了利用前一个阶段有用的去雨信息,使用长短时间记忆网络(LSTM)^[19]来指导下一个阶段的去雨工作;最后,使用通道注意力和空间注意力^[20]的残差网络,它包含 6 个双注意力的残差模块(residual dual attentions).注意力机制能够在不显著增加计算量和参数量的前提下,更好地提升网络模型提取特征的能力,指导网络模型注意雨痕信息,恢复和保留背景的细节特征.在网络的输出层,运用集成学习^[21-23]的思想,将每个阶段的输出结果经过门控网络学习,自动分配各个阶段的权重再进行加权相加,得到最终无雨图像.训练过程只采用单一的负结构相似性(negative SSIM)^[24]作为损失函数训练 RDARENet 网络.

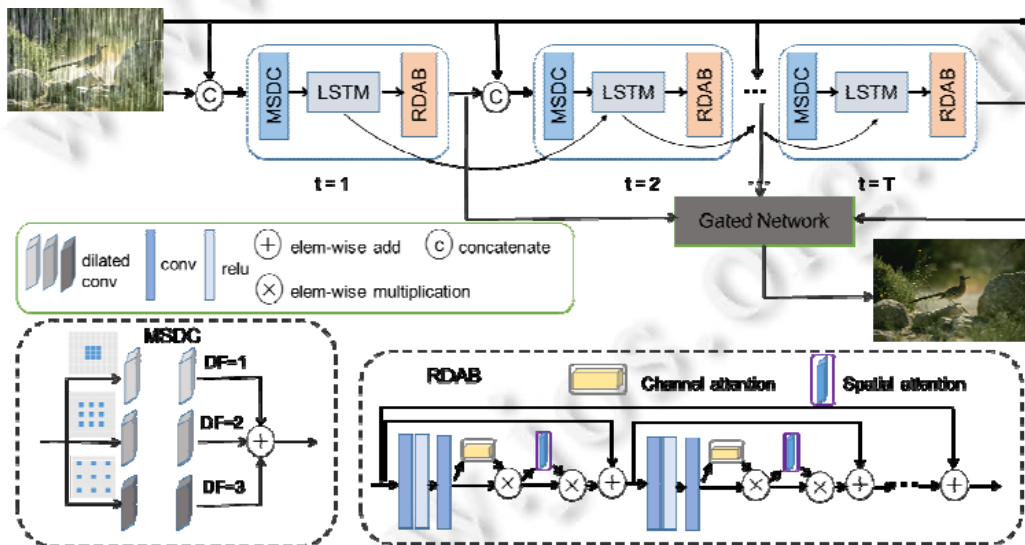


Fig.1 Single image de-raining using a recurrent dual-attention-residual ensemble network structure

图 1 双注意力的残差循环单幅图像去雨集成网络结构图

本文的具体贡献归纳为以下 3 点.

- (1) 提出了一个新的端到端单幅图像去雨网络(RDARENet).通过构建一个由多尺度扩张卷积聚合、长短时间记忆网络和双注意力残差网络组成的基础模块,从浅层块到深层块,逐步地计算雨痕特征的残差

信息;

- (2) 使用集成学习的思想,设计门控网络自动学习各个阶段的权重,将各个阶段的输出结果乘以门控网络对应的权重再线性相加.在网络训练的过程中,只使用单一的负结构相似性作为网络的损失函数;
- (3) 在仿真和真实的雨图像数据集中,本文提出的方法在 SSIM 和 PSNR 评价标准上超过了最新的几种方法,得到了恢复背景图像细节的无雨图像.

本文第 1 节介绍在图像去雨领域方面的相关研究.第 2 节提出单幅图像去雨网络的方法.第 3 节在仿真和真实的雨数据集上,比较本文方法和目前主流的方法.最后总结全文并给出未来的改进目标.

1 相关工作

单幅图像去雨是一项具有挑战性的任务,目前,单幅图像去雨模型主要将雨图像 O 分解为无雨的背景图层 B 和雨痕层 R :

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

通过将雨图像 O 中的雨痕层 R 移除,得到无雨背景图层 B .在之前的方法中,包括字典学习^[3-5]、高斯混合模型^[1]和低秩表现^[25],它们都被用于背景图层或者雨痕层的建模.Kang 等人^[3]首先提出将雨图像分解为高频层和低频层,然后通过字典学习在高频层去除雨痕.Luo 等人^[4]使用判别字典学习,通过很高的判别码强制雨痕层和背景层共享最少的字典原子,分离这两层得到无雨图像.肖等人^[26]提出基于景深和稀疏编码的图像去雨算法.这些模型只是将雨图像认为是一种简单的线性组合,无法处理复杂环境下的雨图像,因此限制了模型去雨的能力.

近年来,由于强大的特征表达和端到端的数据推断能力,深度卷积网络被广泛应用于单幅图像去雨任务,并实现显著的性能改进.通常将去雨问题建模为基于像素的图像回归过程,该过程通过卷积、池化和非线性等一系列操作,将雨天图像映射为无雨图像或者以端到端的方式映射为负的残差图像.Yang 等人^[12]采用多任务的卷积神经网络框架去联合检测和移除雨痕,基于上下文的扩张卷积网络和循环结构处理多尺度的雨痕.在文献[11]中,结合残差网络^[27],作者提出了深度细节网络(DDN)来学习雨水图像高频部分的残差.Zhang 等人^[10]提出了密度感知的多流密集连接卷积网络,用于联合估算雨密度和消除雨痕.Li 等人^[15]提出使用循环的扩张卷积和挤压与激发块(squeeze-and-excitation)的网络(RESCAN)消除雨痕.Ren 等人^[16]提出渐近式的去雨网络(PReNet),通过多阶段使用残差网络,拼接每个阶段输出和原始雨图像的通道作为残差网络的输入,最后输出负的残差图像.一些轻量级的网络,如拉普拉斯算子的金字塔结构^[28],Fu 等人提高了网络计算的效率但却降低了生成图像的质量.在损失函数方面,Johnson 等人^[29]将感知损失作为新的衡量方法,感知损失是基于像素与像素的操作,被用来计算输出图像与真实图像的特征图差异(例如,一般用预训练的 VGG16 网络提取特征),将它们的差异作为惩罚项,抑制输出中发生伪影图像.上述去雨方法在去除雨痕和图像恢复的效果上都有明显提升,由于网络本身的限制,一些网络输出的图像有伪影或者局部背景模糊,不能很好地保留和恢复图像细节.基于该问题,本文提出的 RDARENet 网络不仅能去除雨痕,而且可以保留和恢复图像的细节.本文也尝试多种损失函数来联合约束网络,实验结果表明,单一的负结构相似度损失函数可使模型实现较好的去雨效果.

2 RDARENet 框架

2.1 网络结构

RDARENet 的整体框架如图 1 所示,MSDC 表示多尺度扩张卷积的聚合模块,RDAB 表示是由多个基于通道注意力和空间注意力残差块组成的残差模块.在单阶段雨痕去除的方法中,生成的图像仍然含有雨信息.为了更好地去除雨痕,本文将去雨的过程分解为多阶段,如公式(2)所示:

$$\begin{cases} Y^t = F_{MLRB}(f_{concat}(X^{t-1}, X^0)), 1 \leq t \leq T \\ X^t = Y^t + X^0 \end{cases} \quad (1)$$

其中, X^0 表示原始有雨图像; F_{MLRB} 表示由多尺度扩张卷积聚合、长时间记忆和基于注意力的残差网络组成的

基础模块; f_{concat} 表示拼接上一阶段输出与原始输入图像的通道; T 表示最大的循环次数; X' 表示原始输入图像与基础模块输出信息的负残差,并作为下一个阶段的输入,逐步地消除雨痕信息.

多阶段去雨的网络架构使得网络不断加深,导致后面阶段无法学习到之前提取的特征信息.本文采用长时间记忆网络(LSTM)使信息得到有效的传递.LSTM网络中引入一个新的内部状态单元,该状态单元既可以线性地循环信息传递,又可以非线性地输出信息给隐藏层的外部状态;同时引入门控机制来控制特征信息积累的速度,它会有选择地加入新的信息并遗忘之前积累的信息,让网络学习到更久远的特征信息.LSTM网络包含输入门 i_t 、遗忘门 f_t 、状态单元 C_t 和输出门 O_t ,具体过程定义如下:

$$\left. \begin{aligned} i_t &= \sigma(W_{xi} \times X_t + W_{hi} \times H_{t-1} + W_{ci} \odot C_{t-1} + b_i) \\ f_t &= \sigma(W_{xf} \times X_t + W_{hf} \times H_{t-1} + W_{cf} \odot C_{t-1} + b_f) \\ C_t &= f_t \odot C_{t-1} + i_t \odot \tanh(W_{xc} \times X_t + W_{hc} \times H_{t-1} + b_c) \\ O_t &= \sigma(W_{xo} \times X_t + W_{ho} \times H_{t-1} + W_{co} \odot C_t + b_o) \\ H_t &= O_t \odot \tanh(C_t) \end{aligned} \right\} \quad (3)$$

其中, X_t 是由MSDC模块提取出来的特征; C_t 编码的状态单元被传送到下一个LSTM网络; H_t 表示当前LSTM的输出,它将会被送入到双注意力的残差块.

2.2 多尺度平滑扩张卷积

上下文的信息对于检测和移除雨图像中的雨痕十分重要,扩张卷积可在不减少分辨率的情况下,指数化地扩大感受野区域.然而,研究者们在工作中发现:原始的扩张卷积会产生网格伪影问题,在扩张率大于1时,输出特征中的相邻单元是由输入元素中完全独立的单元组计算得到,进而导致局部信息的不一致.在文献[18]中,作者针对扩张卷积会产生网格伪影现象,提出了可分离和共享的卷积平滑扩张卷积:“可分离”表示将对应的输出通道和输入通道连接,“共享”表示所有输入和输出通道共享权重.作者通过在输入单元和扩张卷积之间或者扩张卷积和输出单元之间添加一个额外的核为 $(2r-1)$ 的卷积层来克服网格伪影.由于雨痕的多样性和背景的复杂性,为了避免生成图像出现伪影问题,本文使用了这种扩张卷积.如图1所示的聚合模块(MSDC),通过多尺度的平滑扩张卷积聚合来学习雨的特征.在网络中,首先使用3个不同扩张因子的平滑扩张卷积(扩张因子分别为1~3)来聚合不同尺度雨痕的细节,然后通过循环的结构,在下一个阶段逐渐地扩大感受野.本文所用平滑扩张卷积的卷积核大小为 3×3 ,在每个扩张卷积后面,使用一个修正线性单元(ReLU)激活函数.

2.3 双注意力的残差网络

注意力机制在人的感知领域^[30]非常重要.当人们在看一幅图像的时候,他们会有序地看过整幅图像,并停留在那些比较突出的场景,而不会将关注点放在整幅图像.最近,一些学者将注意力机制引入到卷积神经网络中,明显提高了在大场景下物体分类的精度^[31].在文献[20]中,作者提出注意力的卷积块(convolution block attention module),对于给定的中间特征映射,他们依次沿着通道和空间维度推断出注意力图,然后将注意力图与输入的要图相乘,以自适应地细化特征.综合之前的去雨工作,一些方法在去雨之后,图像细节部分不能保留或者很好地恢复.在设计去雨网络时,为了更好地提取特征、保留和恢复背景的纹理细节,本文将通道和空间注意力引入残差块中.如图1的RDAB模块所示,RDARENet用6个注意力的残差块组成残差模块.在网络的残差块中,输入的要图首先经过两次卷积操作,输出的特征图 $F \in R^{C \times H \times W}$,将它送入到注意力模块,然后依次推断出一维的通道注意力图 $M_c \in R^{C \times 1 \times 1}$ 和二维的空间注意力图 $M_s \in R^{1 \times H \times W}$,具体过程可以总结如下:

$$\left. \begin{aligned} F' &= M_c(F) \otimes F \\ F'' &= M_s(F') \otimes F' \end{aligned} \right\} \quad (4)$$

其中, \otimes 表示元素相乘.通道注意力的值沿着空间的维度传播,同时,空间注意力的值也在每个通道传播, F'' 为注意力输出的特征图,它将与输入要素相加输入到下一个残差块.

通道注意力和空间注意力模块的具体操作如图2所示.在通道注意力模块中,输入的特征图首先经过平均池化(AvgPool)和最大池化(MaxPool)的操作计算出特征,然后将特征送入共享的多层感知机模型(含有一个隐

藏层)产生通道注意力图.空间注意力模块通过聚合平均池化和最大池化的特征图,送入到核为 7×7 的卷积核中卷积,产生二维的空间特征图.通道和空间注意力都在平均池化和最大池化两个方面对特征进行提取聚合,以进一步提高网络的表征能力.

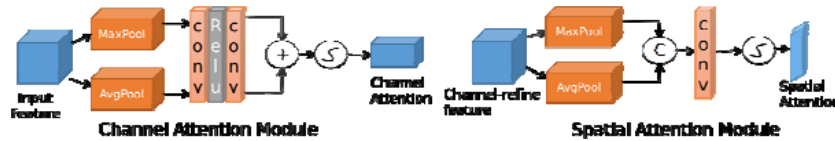


Fig.2 Channel attention module and spatial attention module, σ represents Sigmoid function

图 2 通道注意力和空间注意力模块 σ 表示 Sigmoid 函数

在去雨网络的结构中,通道注意力关注于每一层雨痕的信息,雨痕遮挡的信息需要用周围的信息来逐步地恢复,空间注意力则利用同一通道不同位置特征之间的联系来修复雨痕部位.输入的特征经过双重注意力模块与原始输入特征融合,使图像在雨痕的去除和细节的恢复上明显得到提高,生成的图像更接近原始图像.在残差块中,为了避免随着网络的加深丢失图像低层细节,使用跳跃连接(skip connection)将低层和高层特征连接起来.

2.4 集成学习门控网络

在机器学习领域,集成学习已经运用得很广泛,单一模型的学习能力不强,无法预测出很好的结果.通过将多个学习模型有机地结合在一起,以提高模型的泛化能力.受此启发,在神经网络中,Huang 等人^[22]和 He 等人^[32]将集成学习的思想应用到多种计算机视觉的任务中,明显提升了网络的效果.

RDARENNet 是一个多阶段逐步细化去雨的过程,网络中各个阶段都学习到不同的去雨信息.为了提高网络的泛化能力,文献[33]通过添加额外的门控网络,将粗粒度和细粒度的去雨阶段聚合.将每个阶段输出结果的通道拼接在一起输入到门控网络,门控网络的输出为每一阶段 F_1, F_2, \dots, F_t 对应不同的权重值 $\alpha_1, \alpha_2, \dots, \alpha_t$ (t 表示最大的循环次数),阶段的权重是通过网络自主学习.最后,每个权重乘以对应阶段的输出结果再线性相加:

$$\left. \begin{aligned} (\alpha_1, \alpha_2, \dots, \alpha_t) &= G(F_1, F_2, \dots, F_t) \\ F_{out} &= \alpha_1 \times F_1 + \alpha_2 \times F_2 + \dots + \alpha_t \times F_t \end{aligned} \right\} \quad (5)$$

如公式(5)所示, G 表示卷积核为 1×1 的门控网络, F_{out} 表示集成学习的输出结果.在第 3 节中,我们将对每个阶段和最终融合的输出结果进行比较,实验结果表明,集成学习的效果最好.

2.5 损失函数

一些混合损失函数,例如均方误差与边缘损失^[14]、均方误差与结构相似度^[23]和对抗损失^[34],被用来训练去雨网络.边缘损失针对于恢复图像的细节,感知损失是将真实图像和生成图像分别经过卷积得到的特征图相比较,使得在内容与全局结构的高层信息接近.通过在第 3 节中加入均方误差、多阶段损失和感知损失来联合训练 RDARENNet 网络,发现单一的负结构相似度损失函数可以很好地训练网络.损失函数如公式(6)所示, X^{out} 代表集成学习输出图像, X^{gt} 代表相对应的真实无雨图像:

$$Loss = -SSIM(X^{out}, X^{gt}) \quad (6)$$

3 实验

本文在仿真雨数据集和真实雨天数据集测试 RDARENNet 网络.通过结构相似度(SSIM)和峰值信噪比(PSNR)评价不同数据集上的表现.同时,将本文的算法与最新的几种算法进行比较:(a) 混合高斯模型(GMM)(CVPR'16)^[11];(b) 深度细节网络(DNN)(CVPR'17)^[111];(c) 循环的扩张卷积和挤压与激发网络(RESCAN)(ECCV'18)^[115];(d) 渐进式网络(PReNet)(CVPR'19)^[116].

本文基于 Pytorch 库构建的去雨网络,在 NVIDIA GeForce RTX 2080Ti GPU 上训练模型.训练图像的大小为 100×100 像素,批训练大小为 10.采用 ADAM 算法去训练模型,初始化学学习率为 1×10^{-3} ,总共训练迭代 100 次.在

第 30 次、第 50 次、第 80 次迭代时,学习率每次乘以 0.2 降低。

真实环境下很难获得大量成对无雨和有雨的图像,因此使用仿真雨的数据集去训练本文的网络.Yang 等人^[12]收集仿真小雨 Rain100L 和大雨 Rain100H 数据集,每个数据集都有 100 张测试图像.Ren 等人^[16]删除 Rain100H 中与测试集相同的图像,使网络有更好的泛化能力.在本实验中,使用的数据是 Ren 等人提供的数据集、Zhang 等人^[35]的数据集 Rain800(含有 700 张训练图像和 100 张测试图像)以及测试数据集 Rain12^[11].真实的雨图像是从网络下载的,用来测试网络模型在真实情景下的去雨能力。

3.1 融合学习

3.1.1 网络框架

本小节分析 RDARENet 中的两个模块,分别是阶段循环次数、每个阶段的预测结果与集成学习输出结果的比较。

- 循环的次数

网络中使用注意力残差块的个数为 6.所用的训练数据为 Rain100L,见表 1, $T=3,4,5,6,7$ 循环阶段,通过评价指标 PSNR 和 SSIM 值的比较,在循环次数 $T=6$,PSNR 和 SSIM 的值都取得最好且优于其他,所以设置网络阶段的循环次数为 6。

Table 1 Comparison of RDARENet models with different T stages

表 1 比较不同阶段次数的 RDARENet 模型

Model	RDARENet ₃	RDARENet ₄	RDARENet ₅	RDARENet ₆	RDARENet ₇
PSNR	37.887	38.378	38.003	38.547	38.341
SSIM	0.980	0.981	0.981	0.982	0.982

- 阶段和集成输出结果

如图 3 所示,在 Rain100H 训练集上训练网络,分别选取了阶段 1、阶段 3 和阶段 6 的输出图像以及集成输出图像.在 PSNR 和 SSIM 评价指标比较输出图像,集成输出图像的值明显高于其他各阶段.PSNR 和 SSIM 值分别为:图 3(b),19.369/0.705;图 3(c),25.918/0.842;图 3(d),24.621/0.785;图 3(e),**27.424/0.895**。



Fig.3 Comparison of the different stages outputs and ensemble output

图 3 比较不同阶段输出图以及集成输出图

3.1.2 模块的融合

为了验证各个不同模块在网络中的重要性,本文进行了融合学习.如表 2 所示,在 Rain100L 数据集上验证不同模块对网络的效果.首先,使用 6 个连续的残差网络(RB)作为基础的去雨网络.通过在每个残差网络的残差块中添加通道注意力和空间注意力模块,使得网络更好地关注于雨痕和背景图层,细化地提取以及区分雨痕和背景特征.由于雨痕的密度、形状和方向的复杂性,本文在每个残差网络前都添加多尺度扩张卷积(MSDC)来聚合雨痕信息,并且避免因为网络加深造成低层细节的丢失,在多尺度扩张卷积模块和注意力的残差网络之间添加长短时间记忆网络(LSTM).多阶段去雨网络在每个阶段都输出去雨图像,门控网络则可以自动融合各阶段去雨图像之间的差异,使得图像去雨的效果更好。

Table 2 Ablation study using different modules combinations

表 2 融合学习不同组合模块

模块	不同模块的组合				
RB	√	√	√	√	√
Dual attentions	×	√	√	√	√
MSDC	×	×	√	√	√
LSTM	×	×	×	√	√
Gated network	×	×	×	×	√
PSNR	37.382	38.118	38.004	38.452	38.547
SSIM	0.971	0.978	0.981	0.981	0.982

3.1.3 损失函数

在训练网络的过程中,尝试融合不同损失函数实现更好的去雨,其中包括负的结构相似度(negative SSIM)、均方误差(MSE)和感知损失(perceptual losses).如表 3 所示,分别采用负结构相似度(SSIM)、负结构相似度与多阶段均方差损失(SSIM+MSE)和负结构相似度与感知损失(SSIM+PL).在 Rain100L 上的测试结果表明:虽然只用了负结构相似度一种损失函数,但是去雨的效果在 PSNR 和 SSIM 好于其他损失函数的组合.

Table 3 Comparison of RDARENet with different loss functions

表 3 比较不同损失函数的 RDARENet 网络

Loss	SSIM	SSIM+MSE	SSIM+PL
PSNR	38.547	38.430	38.025
SSIM	0.982	0.982	0.981

3.2 仿真雨数据集实验

RDARENet 网络模型分别在 Rain100H, Rain100L 和 Rain800 数据集上训练.如表 4 所示,给出了与 GMM^[11]、DDN^[11]、RESCAN^[15]和 PReNet^[16]这 4 种单幅图像去雨方法进行比较(红色和蓝色分别表示最好和其次)的结果.

Table 4 Average PSNR and SSIM comparison on the synthetic datasets

表 4 在仿真数据集上比较 PSNR 和 SSIM 的平均值

Method	GMM(CVPR'16) ^[11]	DDN(CVPR'17) ^[11]	RESCAN(ECCV'18) ^[15]	PReNet(CVPR'19) ^[16]	Ours
Rain100H	15.21/0.421	25.23/0.764	27.71/0.852	29.41/0.899	29.95/0.903
Rain100L	28.45/0.868	32.34/0.941	38.19/0.980	37.49/0.979	38.55/0.982
Rain12	32.11/0.913	34.93/0.943	36.51/0.953	36.23/0.969	37.02/0.970
Rain800	23.57/0.821	25.71/0.846	28.18/0.884	26.29/0.890	28.03/0.898

相对于其他 4 种去雨方法,在测试集 Rain100H 和 Rain100L, Rain12 和 Rain800 上,本文的方法在 PSNR 上分别提高 0.54dB、0.36dB 和 0.51dB,在 SSIM 上分别提高 0.4%、0.2%、0.1%和 0.8%.图 4 展示的是 Rain100H 测试集去雨之后的图像,RDARENet 网络有效地去除了雨痕,在保留和恢复图像细节方面取得很好的视觉效果.

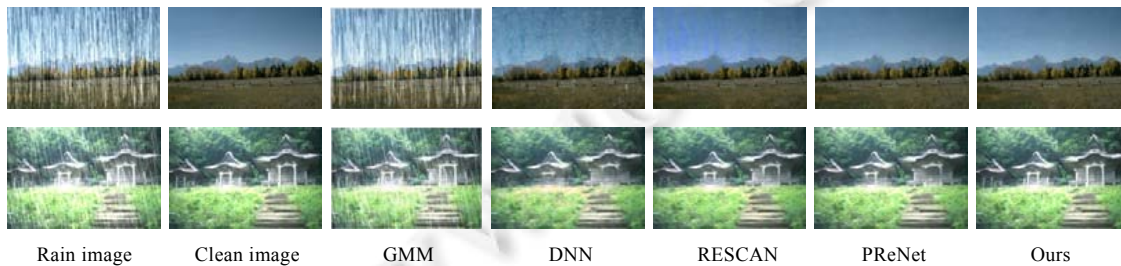


Fig.4 Results of various methods on synthetic images

图 4 仿真数据集多种算法的比较结果

3.3 真实雨数据集实验

如图 5 所示,比较本文算法与 GMM^[11]、DDN^[11]、RESCAN^[15]和 PReNet^[16]方法在真实雨图像的去雨效果.

由于真实场景中无法获得相匹配的有雨图像和无雨图像,所以使用 Rain100H 大雨数据集训练的模型,在真实雨图像上实现去雨.从第 1 幅图像红色矩形框开始对比,DDN、RESCAN 和 PReNet 方法虽然去除了雨痕,但是得到的图像模糊了背景细节;GMM 保留了背景细节,但生成的图像仍然残留不少雨痕.本文的方法不仅很好地去除了雨痕,而且保留了背景的纹理细节.即使在第 2 幅图像真实环境比较复杂的情况下,本文的模型依然可以很好地去除雨痕.总的来说,在真实雨图像中,RDARENet 在雨痕去除和纹理细节恢复上都优于文中的其他方法.

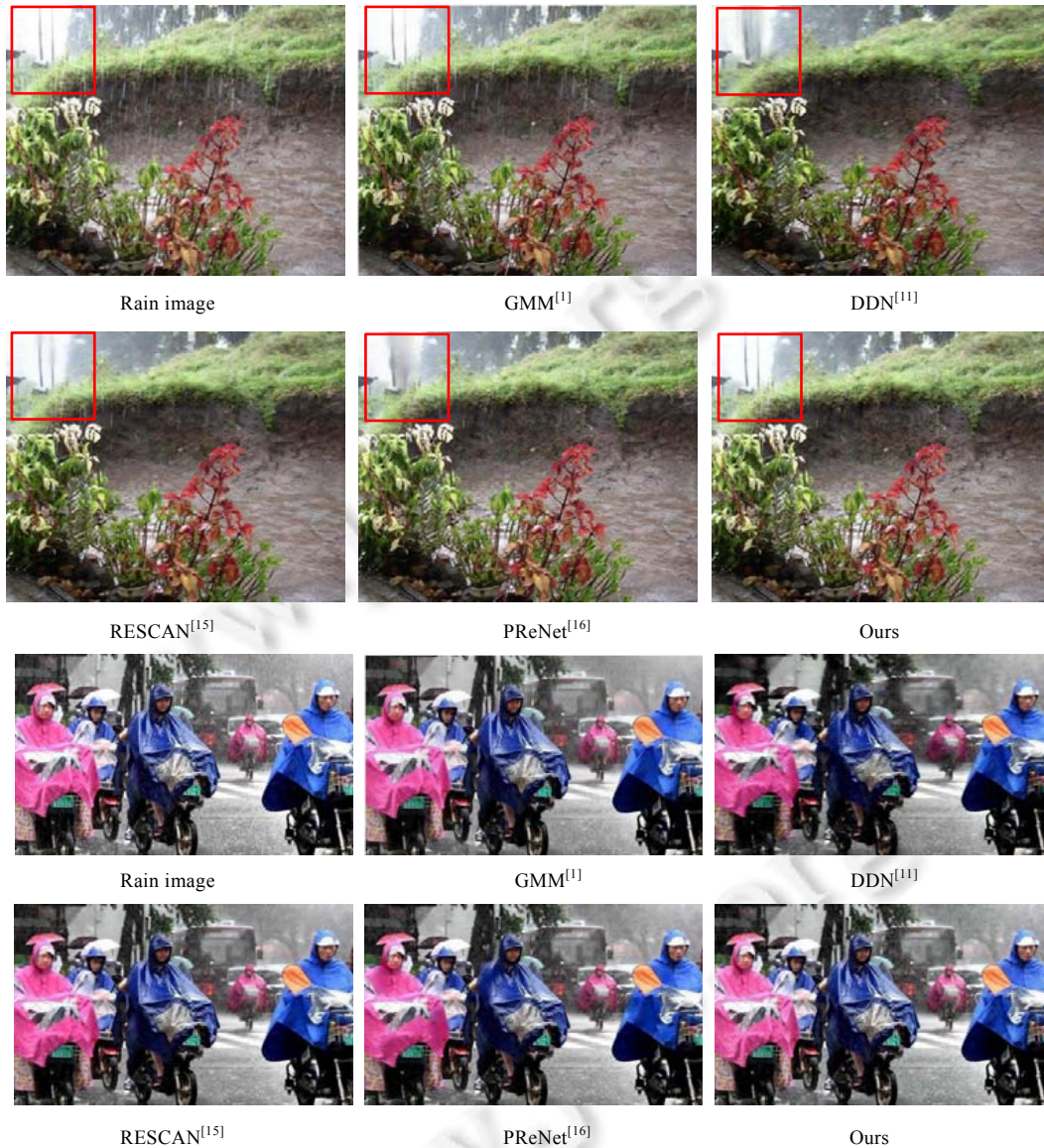


Fig.5 Removal of rain streaks in real rainy images

图 5 真实雨图像的雨痕移除

4 结 论

本文提出了基于双注意力的残差循环单幅图像去雨集成网络,实现了逐阶段去除雨痕的目的.构建了由多尺度扩张卷积聚合、长短时间记忆网络和时空注意力的残差网络组成的基础块.在网络输出层,采用集成学习

的思想提出了门控网络,自主学习和分配各个阶段的权重,将每个阶段的输出结果与门控网络对应的权重相乘相加.通过仿真和真实的雨数据实验,结果表明:本文的方法相对于文中其他图像去雨方法,在雨痕去除和纹理细节保留与恢复上均取得了明显的提升.在未来的工作中,我们计划在非匹配真实的雨天情况下实现单幅图像的去雨,因为仿真雨数据集不能很好地模拟实际复杂雨天的场景,希望通过不匹配的网络学习方式,达到真实情景下的单幅图像去雨任务.

References:

- [1] Li Y, Tan RT, Guo XJ, Lu J, Brown MS. Rain streak removal using layer priors. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2016. 2736–2744.
- [2] Gu SH, Meng DY, Zuo WM, Zhang L. Joint convolutional analysis and synthesis sparse representation for single image layer separation. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2017. 1708–1716.
- [3] Kang LW, Lin CW, Fu YH. Automatic single-image-based rain streaks removal via image decomposition. IEEE Trans. on Image Processing, 2012,21(4):1742–1755.
- [4] Luo Y, Xu Y, Ji H. Removing rain from a single image via discriminative sparse coding. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2015. 3397–3405.
- [5] Wang YL, Liu SC, Chen C, Zeng B. A hierarchical approach for rain or snow removing in a single color image. IEEE Trans. on Image Processing, 2017,26(8):3936–3950.
- [6] Dong C, Loy CC, He KM, Tang XO. Image super-resolution using deep convolutional networks. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2015,38(2):295–307.
- [7] Zhang K, Zuo WM, Chen YJ, Meng DY, Zhang L. Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Trans. on Image Processing, 2017,26(7):3142–3155.
- [8] Tai Y, Yang J, Liu XM. Image super-resolution via deep recursive residual network. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2017. 3147–3155.
- [9] Fu XY, Huang JB, Ding XH, Liao YH, Paisley J. Clearing the skies: A deep network architecture for single-image rain removal. IEEE Trans. on Image Processing, 2017,26(6):2944–2956.
- [10] Zhang H, Patel VM. Density-aware single image deraining using a multi-stream dense network. In: Proc. of the IEEE Conf. on Computer Vision. 2018. 695–704.
- [11] Fu XY, Huang JB, Zeng DL, Huang Y, Ding XH, Paisley J. Removing rain from single images via a deep detail network. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2017. 3855–3863.
- [12] Yang WH, Tan RT, Feng JS, Liu JY, Guo ZM, Yan SC. Deep joint rain detection and removal from a single image. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2017. 1357–1366.
- [13] Wang TY, Yang X, Xu K, Chen SZ, Zhang Q, Lau RW. Spatial attentive single-image deraining with a high quality real rain dataset. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2019. 12270–12279.
- [14] Yang YZ, Lu H. Single image deraining using a recurrent multi-scale aggregation and enhancement network. In: Proc. of the IEEE Int'l Conf. on Multimedia and Expo. 2019. 1378–1383.
- [15] Li X, Wu JL, Lin ZC, Liu H, Zha HB. Recurrent squeeze-and-excitation context aggregation net for single image deraining. In: Proc. of the European Conf. on Computer Vision. 2018. 254–269.
- [16] Ren DW, Zuo WM, Hu QH, Zhu PF, Meng DY. Progressive image deraining networks: A better and simpler baseline. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2019. 3937–3946.
- [17] Li JJ. Research on single image rain removal based on convolutional neural network [MS. Thesis]. Maanshan: Anhui University of Technology, 2020 (in Chinese with English abstract).
- [18] Wang ZY, Ji SW. Smoothed dilated convolutions for improved dense prediction. In: Proc. of the 24th ACM SIGKDD Int'l Conf. on Knowledge Discovery & Data Mining. 2018. 2486–2495.
- [19] Shi XJ, Chen ZR, Wang H, Yeung DY, Wong WK, Woo WC. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In: Advances in Neural Information Processing Systems. 2015. 802–810.

- [20] Woo S, Park JC, Lee JY, Kweon IS. CBAM: Convolutional block attention module. In: Proc. of the European Conf. on Computer Vision. 2018. 3–19.
- [21] Drucker H, Cortes C, Jackel LD, Lecun Y, Vapnik V. Boosting and other ensemble methods. *Neural Computation*, 1994,6(6): 1289–1301.
- [22] Huang G, Liu Z, Weinberger KQ, Maaten L. Densely connected convolutional networks. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2017. 4700–4708.
- [23] Fan ZW, Wu HF, Fu XY, Huang Y, Ding XH. Residual-guide network for single image deraining. In: Proc. of the 26th ACM Int'l Conf. on Multimedia. 2018. 1751–1759.
- [24] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. on Image Processing*, 2004,13(4):600–612.
- [25] Liu GC, Lin ZC, Yan SC, Sun J, Yu Y, Ma Y. Robust recovery of subspace structures by low-rank representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2012,35(1):171–184.
- [26] Xiao JS, Wang W, Zou WT, Tong L, Lei JF. Image deraining algorithm based on depth of field and sparse coding. *Chinese Journal of Computers*, 2019,42(9):2024–2034 (in Chinese with English abstract).
- [27] He KM, Zhang XY, Ren SQ, Sun J. Identity mappings in deep residual networks. In: Proc. of the European Conf. on Computer Vision. Cham: Springer-Verlag, 2016. 630–645.
- [28] Fu XY, Liang BR, Huang Y, Ding XH, Paisley J. Lightweight pyramid networks for image deraining. *The IEEE Trans. on Neural Networks and Learning Systems*, 2019, 1–14.
- [29] Johnson J, Alahi A, Li FF. Perceptual losses for real-time style transfer and super-resolution. In: Proc. of the European Conf. on Computer Vision. Cham: Springer-Verlag, 2016. 694–711.
- [30] Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 2002, 3(3):201–215.
- [31] Wang F, Jiang MQ, Qian C, Yang S, Li C, Zhang HG, Wang XG, Tang XO. Residual attention network for image classification. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2017. 3156–3164.
- [32] He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2016. 770–778.
- [33] Takikawa T, Acuna D, Jampani V, Fidler S. Gated-SCNN: Gated shape CNNs for semantic segmentation. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2019. 5229–5238.
- [34] Qian R, Robby TT, Yang WH, Su JJ, Liu JY. Attentive generative adversarial network for raindrop removal from a single image. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2018. 2482–2491.
- [35] Zhang H, Sindagi V, Patel VM. Image de-raining using a conditional generative adversarial network. *IEEE Trans. on Circuits and Systems for Video Technology*, 2019, 1–10.

附中文参考文献:

- [17] 李金晶. 基于卷积神经网络的单幅图像去雨研究[硕士学位论文]. 马鞍山: 安徽工业大学, 2020.
- [26] 肖进胜, 王文, 邹文涛, 童乐, 雷俊锋. 基于景深和稀疏编码的图像去雨算法. *计算机学报*, 2019,42(9):2024–2034.



张学锋(1978—), 男, 博士, 教授, 主要研究领域为模式识别, 虚拟现实技术, 人工智能.



李金晶(1994—), 男, 硕士, 主要研究领域为计算机视觉, 机器学习, 虚拟现实.