中图法分类号:TP391.41　　**文献标识码**:A　　**文章编号**:1006-8961(2022)05-1509-13

论文引用格式:Wang M H, Ke F H, Liang Y, Fan Z and Liao L. 2022. 3D attention and Transformer based single image deraining network. Journal of Image and Graphics,27(05):1509-1521(王美华,柯凡晖,梁云,范衡,廖磊. 2022. 融合3D注意力和Transformer的图像去雨网络. 中国图象图形学报,27(05):1509-1521)[DOI:10.11834/jig.210794]

融合3D注意力和Transformer的图像去雨网络

王美华[1],柯凡晖[1],梁云[1],范衡[2*],廖磊[1]

1. 华南农业大学数学与信息学院,广州　510642;2. 汕头大学工学院,汕头　515063

摘　要:**目的**　因为有雨图像中雨线存在方向、密度和大小等各方面的差异,单幅图像去雨依旧是一个充满挑战的研究问题。现有算法在某些复杂图像上仍存在过度去雨或去雨不足等问题,部分复杂图像的边缘高频信息在去雨过程中被抹除,或图像中残留雨成分。针对上述问题,本文提出三维注意力和Transformer去雨网络(three-dimension attention and Transformer deraining network,TDATDN)。**方法**　将三维注意力机制与残差密集块结构相结合,以解决残差密集块通道高维度特征融合问题;使用Transformer计算特征全局关联性;针对去雨过程中图像高频信息被破坏和结构信息被抹除的问题,将多尺度结构相似性损失与常用图像去雨损失函数结合参与去雨网络训练。**结果**　本文将提出的TDATDN网络在Rain12000雨线数据集上进行实验。其中,峰值信噪比(peak signal to noise ratio,PSNR)达到33.01 dB,结构相似性(structural similarity,SSIM)达到0.927 8。实验结果表明,本文算法对比以往基于深度学习的神经网络去雨算法,显著改善了单幅图像去雨效果。**结论**　本文提出的TDATDN图像去雨网络结合了3D注意力机制、Transformer和编码器—解码器架构的优点,可较好地完成单幅图像去雨工作。

关键词:单幅图像去雨;卷积神经网络(CNN);Transformer;3D注意力;U-Net

3D attention and Transformer based single image deraining network

Wang Meihua[1], Ke Fanhui[1], Liang Yun[1], Fan Zhun[2*], Liao Lei[1]

1. College of Mathematics and Informatics, South China Agricultural University, Guangzhou　510642, China;

2. College of Engineering, Shantou University, Shantou　515063, China

Abstract: **Objective** Vision-based computer systems can be used to process and analyze acquired images and videos in fuzzy weather like rainy, snowy, sleet or foggy. These image quality degradation issues derived from severe weather conditions will significantly distort the image visual quality and reduce the performance of the computer vision system. Hence, it is important to develop computer image deraining automatic processing algorithms. Our research focuses on the issue of single image based removing rain streaks. The traditional image rain removal model is mainly based on the prior information to remove the rain from the image. It regards the rain image as a combination of the rain layer and the background layer, and defines the separation of the rain layer and the background layer by the image deraining task. Due to the existing differences in related to direction, density, and size of rain streaks in rain images, a single image derived de-raining issue is a challenging computer vision task currently. Deep learning has benefited to de-raining images but existing models has challenges like excessive rain removal or insufficient rain removal on complicated images scenario. The high-frequency edge informa-

收稿日期:2021-09-14;**修回日期**:2022-02-21;**预印本日期**:2022-02-28

*通信作者:范衡　zfan@stu.edu.cn

基金项目:国家自然科学基金项目(61976052);广东省基础与应用基础研究基金项目(2019A050510034, 2019B1515210009)

Supported by: National Natural Science Foundation of China (61976052); Guangdong Basic and Applied Basic Research Fund Project (2019A050510034, 2019B1515210009)

tion of some complex images is erased during the rain removal process, or rain components remaining in the rain removal image. We propose this paper proposes the three-dimension attention and Transformer de-raining network (TDATDN) single image rain removal network, which improves the image rain removal network based on the encoder-decoder architecture and integrates 3D attention, Transformer and encoder-decoder take advantages of the structure to enhance the image to the rain effect. Our training dataset consists of 12 000 pairs of training images (including three types of rain images with different rain densities), and 1 200 test set images are used to test the rain removal effect. The input image size is scaled to 256×256 for training and testing. Adam optimizer is used for training and learning. The initial learning rate is set to 1×10^{-4}, and its network epoch number is 100. The learning rate is multiplied by 0.5 when reach 15 times. **Method** Our method melts the three-dimension attention mechanism into the residual dense block structure to resolve the challenge of high-dimensional feature fusion via the residual dense block channel. Then, our proposed three-dimension attention residual dense block as the backbone network to build an encoder-decoder-based architecture image de-raining network, and uses Transformer mechanism to calculate the global contextual relevance of the deep semantic information of the network. The Transformer obtained self-attention feature encoding by is up-sampling operation based on the decoder structure image restoration path. To obtain a rain removal result with richer high-frequency details the up-sampling operation obtains the feature map of the image is spliced in the channel direction with the corresponding encoder-based feature map. For the image high-frequency information loss and the structure information is erased in the rain removal process, our problem solving combines the multi-scale structure similarity loss with the commonly used image de-raining loss function to improve the training of the de-raining network. **Result** Our TDATDN network is demonstrated on the Rain12000 rain streaks dataset. Among them, the peak signal to noise ratio (PSNR) reached 33.01 dB, and the structural similarity (SSIM) reached 0.927 8. A comparative experiment was carried out to verify the fusion algorithm results. The result of the comparative experiment illustrated that our algorithm has its priority to improve the effect of a single image oriented rain removing. **Conclusion** Our image de-raining network takes the advantages of 3D attention mechanism, Transformer and encoder-decoder architecture into account.

Key words: single image deraining; convolutional neural network (CNN); Transformer; 3D attention; U-Net

0 引 言

雨天会对摄影设备成像质量造成极大影响。单幅图像去雨算法的目标是:输入一幅有雨图像,得到对应图像的去雨结果,图像去雨任务能够有效提高图像质量并增加计算机视觉算法的适应性(Yang 等,2019)。

早期图像去雨算法大多基于图像先验信息完成去雨任务,其将图像分为背景和雨层,把图像去雨任务转化为图层分解任务。使用字典学习、稀疏编码(Li 等,2012;Luo 等,2015)或高斯混合网络(Li 等,2016)学习分解图像,以解决图像去雨任务这一复杂优化问题。然而,手工设计特征不足以获取足够的图像统计学信息,常常导致去雨效果不理想。

深度学习在一系列图像分类(Rawat 和 Wang,2017)、目标检测(Liu 等,2020)和图像分割(Lateef 和 Ruichek,2019)等计算机视觉任务中大放异彩。许多基于深度学习的图像去雨工作取得了不错的效果。例如,基于先验信息引导的图像去雨网络(Fu 等,2017a;Zhang 等,2018;Yasarla 和 Patel,2019)、粗糙到精细的去雨网络(Ren 等,2019)、多任务学习去雨网络(Yang 等,2016)和基于注意力引导的去雨网络(Wang 等,2020,2021)。

基于深度学习的图像去雨网络虽取得了不错的去雨效果,但仍在某些图像上存在着去雨不充分或过度去雨的问题。现有算法在某些去雨后的图像中依旧残留着雨纹;或在去雨过程中抹除了物体边缘信息。现有基于深度学习的去雨网络多是使用卷积操作进行图像去雨。然而,卷积操作不适合建模全局特征依赖,当图像全局信息对去雨过程十分重要时,网络将无法很好地完成去雨任务。图像中低频区域可能出现雨纹残留;或无法区分雨纹与高频信息,导致物体边缘存在涂抹痕迹或雨纹去除不成功。

Transformer 注意力机制首次出现于机器翻译任务(Vaswani 等,2017),其将一个序列作为输入,使用神经元计算序列中每一个元素与其余元素的特征

依赖,这种设计赋予了它强大的特征长期依赖建模能力。基于卷积神经网络(convolutional neural networks,CNN)的去雨方法并不擅长建模特征长距离依赖,因此本文将 Transformer 与编码器—解码器结构应用于去雨任务中,以更好地减少去雨涂抹痕迹和恢复图像高频细节。

本文针对单幅图像去雨任务提出 TDATDN(three-dimension attention and Transformer deraining network)网络,在一定网络空间内寻求更好的网络表达方式。将三维注意力和 Transformer 机制引入图像去雨领域,并将 U 型编码器—解码器网络与 Transformer 结合,以解决引入 Transformer 机制导致的特征分辨率损失问题。TDATDN 具有以下特点:1)使用编码器提取图像特征;2)使用 Transformer 机制在高层语义空间内获取全局内容编码;3)使用解码器结构对 Transformer 获取的自注意力特征编码完成上采样操作;4)使用 Skip-Connection 融合编码器与解码器对应路径特征。

1 相关工作

1.1 单幅图像去雨

1.1.1 传统单幅图像去雨

传统的单幅图像去雨算法将有雨图像视为背景层与雨层的叠加图像。基于此假设,图像去雨工作可认为是背景层和雨层的分离任务。Kang 等人(2012)使用滤波器提取图像高频信息,并使用字典学习和稀疏编码将图像高频成分中的雨分量和非雨分量分离,从而获得去雨图像。Luo 等人(2015)通过学习强互斥性学习字典,使用高可辨性稀疏编码准确分离背景层和雨层。Li 等人(2016)使用高斯混合网络分离背景层和雨层,以适应多方向和不同规模的雨纹。马龙等人(2018)使用最大后验估计建立能量模型,将其与高斯混合模型结合进行去雨。由于背景层和雨层存在复杂的结合关系,传统单幅图像去雨算法难以取得令人满意的结果。

1.1.2 基于深度学习的单幅图像去雨

Fu 等人(2017a)将深度学习与图像去雨结合,在图像高频域内定义有雨图像到去雨图像的转换网络,相比传统去雨算法,显著提升了去雨性能。Fu 等人(2017b)将卷积神经网络和传统图像处理技术结合去雨,以改进视觉效果。但这些粗糙的雨分离算法易导致部分图像高频成分被抹除。Yang 等人(2017)提出去雨多任务学习网络,同时学习获取雨掩膜图、雨线图和去雨图像。针对雨线密度不均问题,Zhang 等人(2018)提出预测雨密度标签引导图像去雨,但其需要额外训练标记数据,限制训练通用性。一些学者将有雨图像中雨线层定义为多个雨层的堆叠,提出对图像递归式去雨,或以多阶段方式图像去雨。Li 等人(2018)将深度卷积和循环神经网络结合,使用通道注意力对雨线特征层分配不同的权重值。Ren 等人(2019)提出一种递归式 ResNet 网络,引入循环层对有雨图像递归式去雨。王美华等人(2020)提出结合选择卷积网络和雨线修正系数的自适应卷积残差修正网络进行去雨。传统去雨算法和部分基于深度学习的去雨算法对图像的去雨操作仅使用浅层特征进行处理,而没有理解图像。

基于编码器—解码器的网络架构将图像内容逐级抽象至高层语义空间,可以更好地理解图像内容,而不仅使用图像浅层特征。一些学者将编码器—解码器架构用于图像去雨中。Guo 等人(2019)提出深度自注意力金字塔网络,将自注意力和多尺度池化模块结合编码器—解码器学习不同雨线信息。Yasarla 和 Patel(2019)考虑雨位置信息,提出不确定性引导的多尺度残差学习网络,但其训练复杂,处理时间长。部分研究者将注意力机制引入图像去雨。Wang 等人(2020)提出注意力机制引导的三阶段编码器—解码器去雨网络,并使用预测损失图辅助训练。Wang 等人(2021)在编码器—解码器引入注意力精细化模块用于引导图像精细化去雨。但这些基于编码器—解码器的去雨网络大多设计复杂,显著增加了处理时间。

1.2 注意力机制

注意力机制在人类视觉感知系统中扮演了一个重要角色。人类视觉系统的一个重要特点是:不会一次性处理一幅图像的整个场景,而是简单快速浏览图像并选择性地关注某些显著区域以获取更好的视觉结构信息(Itti 等,1998)。

1.2.1 传统注意力机制

Wang 等人(2017)提出残差注意力网络,其提出的注意力网络由多层注意力模块堆叠而成,每个注意力模块包含主干分支和掩膜分支,两个分支特征使用点乘操作特征融合。Hu 等人(2020)

提出通道注意力(squeeze-and-excitation network, SE)模块,从通道角度计算特征的关联性,以自适应计算不同特征通道响应性。Woo 等人(2018)提出卷积块注意力模块(convolutional block attention module, CBAM),CBAM 模块将通道注意力和空间注意力级联,分别从通道维度和空间维度生成注意力图。

1.2.2 Transformer 注意力机制

Transformer 注意力机制首次提出于机器翻译领域(Vaswani 等,2017),得益于强大的全局信息建模能力,其在许多自然语言处理任务中取得了最优结果。在计算机视觉领域,研究人员对 Transformer 变体进行研究。Parmar 等人(2018)将其应用于图像生成领域,通过限制 Transformer 注意力机制关注图像局部领域,显著地增加了网络可处理的图像尺度。Child 等人(2019)提出 Sparse Transformer,其在注意力矩阵中使用稀疏分解技术,大大增强了 Transformer 注意力机制对深层网络和长序列数据的全局建模能力。Dosovitskiy 等人(2021)使用纯 Transformer 网络对分块后的输入图像处理,极大加强网络对全局信息的提取能力。本文将 Transformer 注意力机制与编码器—解码器结构结合,以增强网络全局信息建模能力。

2 网络结构

提出基于编码器—解码器架构的 TDATDN 去雨网络,将 3 维注意力与残差密集块结合,构建编码器—解码器架构去雨网络,其将输入图像编码至高级语义特征空间,在高层语义特征空间内利用 Transformer 注意力机制计算特征全局关联性,然后解码器逐级恢复至原图分辨率,得到去雨图像。

2.1 整体网络架构

本文整体网络架构见图 1,整体网络为基于 U-Net 的编码器—解码器架构(Ronneberger 等,2015)。在编码器—解码器部分,使用卷积进行下采样操作,采用反卷积作为上采样操作。使用 Skip-Connection 融合编码器和解码器路径特征,对应层级特征图矩阵相加,并使用 3×3 卷积操作融合特征。激活函数为 Leaky ReLU 激活函数(其中,水平负半轴的斜率设置为 0.1),输出层直接使用 3×3 卷积获取最终去雨结果,不额外使用激活函数。其中,编码器路径第 1、2、3、4 个全局残差密集块输入特征通道数分别为 32、64、64、128;解码器路径第 1、2、3 个全局残差密集块和最后 3×3 卷积输入特征通道数分别为 128、64、64、32。

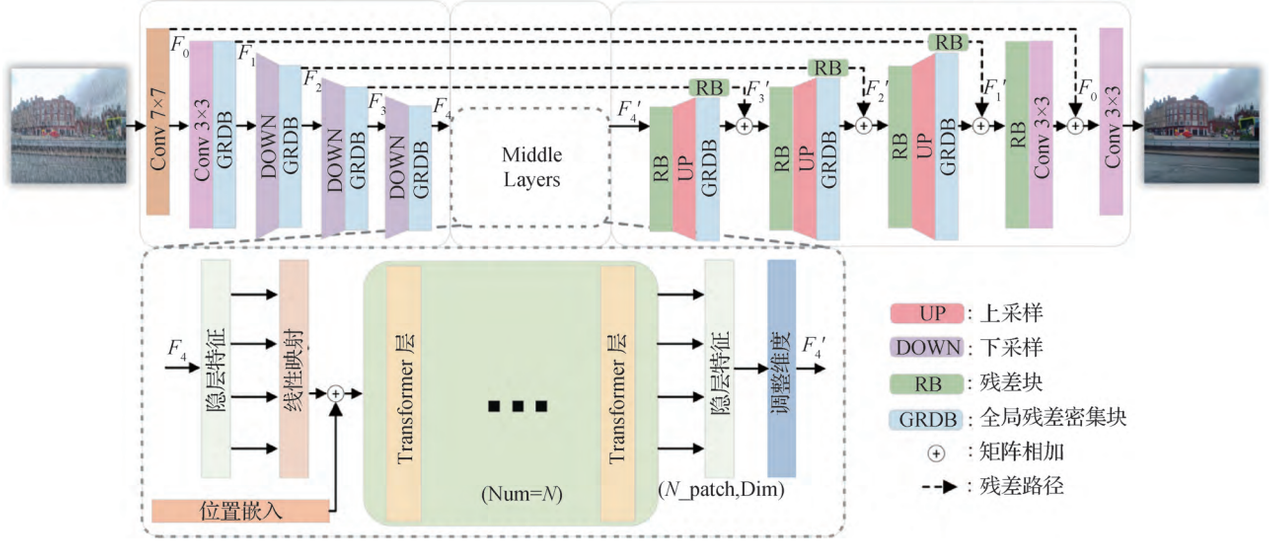


图 1 3 维注意力和 Transformer 去雨网络架构

Fig. 1 The structure of three-dimension attention and Transformer deraining network(TDATDN)

2.2 残差密集块

为了增强网络对于不同雨线位置、方向和差异的去雨能力,受到 RDB(residual dense block)(Zhang 等,2018)工作的启发,提出 TARDB(three-dimension attention residual dense block)作为本文的网络基本单元。Residual(He 等,2016)的加入可有效缓解梯

度消失问题;Dense 机制(Huang 等,2017)对网络中不同层级特征融合;3 维注意力机制(three-dimension attention module, 3DAM)(Yang 等,2021)可通过优化能量函数获取每个神经元重要性,且计算过程中无需额外参数存储。

2.2.1 三维注意力局部残差密集块

传统视觉注意力模块通过输入特征图产生对应注意力权重。传统注意力模块一般可以分为两类:基于 1 维通道方向的通道注意力(Hu 等,2020);基于 2 维空间的空间注意力(Woo 等,2018)。相比 1 维通道注意力或 2 维空间注意力机制,3 维注意力可获取精细的注意力权重图,更好地引导网络去雨任务。本文将 3 维注意力机制(3DAM)与残差密集块结合,提出 3 维注意力局部残差密集块,其包括密集连接层和 3 维注意力特征融合层。3 维注意力局部残差密集块结构见图 2。

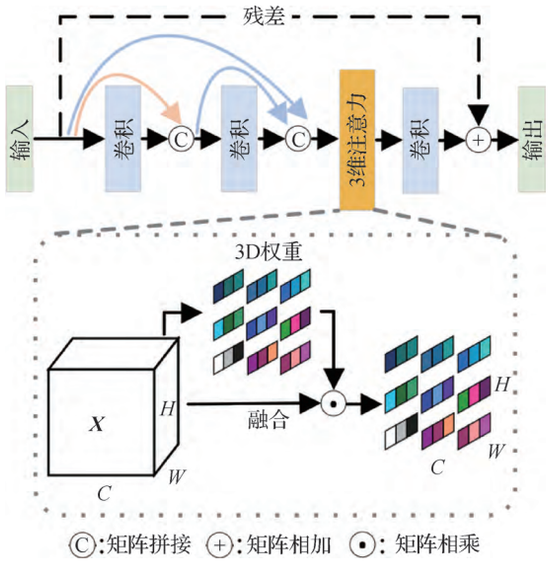


图 2 3 维注意力局部残差块图

Fig. 2 3D attention local residual dense block map

1)密集连接层。为了增加信息在特征图之间的流动性,采用类似 DenseNet 网络(Huang 等,2017)的策略,每一个卷积层将会与先前每一层进行通道方向拼接,然后使用卷积操作特征融合,后接激活函数。定义 $\boldsymbol{X}_i$ 为第 i 个卷积层产生的特征图,$\boldsymbol{X}_0$ 为当前局部残差密集块输入。则第 i 个卷积层产生的特征图输出为

$$\boldsymbol{X}_i = \sigma(\boldsymbol{W}_i[\boldsymbol{X}_0,\boldsymbol{X}_1,\cdots,\boldsymbol{X}_{i-1}]) \tag{1}$$

式中,σ 为 Leaky ReLU 激活函数,$\boldsymbol{W}_i$ 为第 i 个卷积层权重,$[\cdot]$ 为将特征图从通道方向拼接特征。

2)3 维注意力特征融合层。将密集连接层获取的特征按照通道方向拼接,然后使用 3DAM 获取注意力图,并将注意力图与之相乘,后接 1×1 卷积完成特征融合和通道压缩。同时,增加残差连接进一步增强信息流动。3 维注意力特征融合层最终输出为

$$\boldsymbol{X}_{\text{out}} = \boldsymbol{W}_{1\times1}(\phi([\boldsymbol{X}_0,\boldsymbol{X}_1,\cdots,\boldsymbol{X}_{i-1},\boldsymbol{X}_i])) + \boldsymbol{X}_0 \tag{2}$$

式中,$\boldsymbol{X}_{\text{out}}$ 为最终输出的特征图,$\boldsymbol{W}_{1\times1}$ 为特征融合的 1×1 卷积,$\phi(\cdot)$ 为 3DAM 注意力机制计算,其计算为

$$\phi(\boldsymbol{X}) = \psi\left(\frac{(\boldsymbol{X}-\boldsymbol{\mu})^2}{4\times\left(\frac{\nu((\boldsymbol{X}-\boldsymbol{\mu})^2)}{w\times h-1}+\lambda\right)+0.5}\right)\odot\boldsymbol{X} \tag{3}$$

式中,$[\cdot]$ 为特征拼接操作,$\boldsymbol{X}$ 为输入特征图,$\phi(\cdot)$ 为 3 维注意力机制计算公式,$\psi(\cdot)$ 为 Sigmoid 激活函数,$\nu(\cdot)$ 为按照各通道内计算特征图和,$\boldsymbol{\mu}$ 为输入特征图 $\boldsymbol{X}$ 按照各通道内计算的特征图平均值,w 和 h 是输入特征图 $\boldsymbol{X}$ 的宽和高,λ 为超参数,$\odot$ 为矩阵点乘操作。

2.2.2 全局残差密集块

全局残差密集块由 3 维注意力局部残差密集块作为基础卷积层,并使用密集连接和残差学习机制提升网络性能。其结构如图 3。计算式为

$$\boldsymbol{Y}_{\text{out}} = \boldsymbol{W}_{1\times1}([\boldsymbol{Y}_0,\cdots,\boldsymbol{Y}_i]) + \boldsymbol{Y}_0 \tag{4}$$

式中,$\boldsymbol{Y}_{\text{out}}$ 为最终输出特征图,$\boldsymbol{W}_{1\times1}$ 为特征融合的 1×1 卷积,$\boldsymbol{Y}_i$ 为第 i 个 3 维注意力局部残差密集块产生的特征图,$[\cdot]$ 为特征拼接操作。

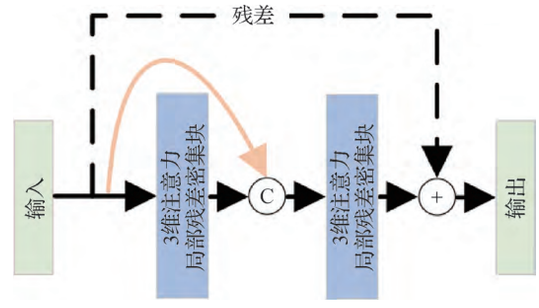


图 3 全局残差块图

Fig. 3 Global residual dense block map

2.3 Middle Layer 层

参考 Dosovitskiy 等人(2021),将输入到 Transformer 层的特征矩阵进行令牌化,重新调整输入矩阵维度变为 2D 特征块 $\{\boldsymbol{X}^i \in \mathbf{R}^{P^2\times C} \mid i=1,\cdots,N\}$,

每一个块的大小为 $P \times P$，C 为特征图通道数。变换之后特征总块数为 $N = HW/P^2$，其中，H 和 W 分别为特征图的高和宽。

2.3.1 Patch Embedding 操作

使用神经网络将向量化的 2D 特征块映射进线性空间。为了将 2D 特征块进行空间位置信息编码，将额外学习位置嵌入信息矩阵，其将与 2D 特征块相加，计算为

$$\boldsymbol{z}_0 = [\boldsymbol{x}_p^1\boldsymbol{E}, \boldsymbol{x}_p^2\boldsymbol{E}, \cdots, \boldsymbol{x}_p^N\boldsymbol{E}] + \boldsymbol{E}_{\text{pos}} \tag{5}$$

式中，$\boldsymbol{z}_0$ 是加入空间位置信息编码的 2D 特征块，$\boldsymbol{x}_p^N$ 为第 N 个大小为 $P \times P$ 的 2D 特征块，$\boldsymbol{E} \in \mathbf{R}^{(P^2 \times C) \times D}$ 为块嵌入映射神经网络权重，D 为线性空间维度，$\boldsymbol{E}_{\text{pos}} \in \mathbf{R}^{N \times D}$ 为空间位置信息编码。

2.3.2 Transformer 层

Transformer 中间层由多层 Transformer 级联而成。单个 Transformer 层包括多头自注意力机制（Vaswani 等，2017）和多层感知器前向神经网络。其结构见图 4。第 i 层 Transformer 层输出为

$$\begin{aligned} z_i' &= \tau(\varphi(z_{i-1})) + z_{i-1} \\ z_i &= \psi(\varphi(z_i')) + z_i' \end{aligned} \tag{6}$$

式中，z_{i-1} 为 Transformer 层输入特征，$\tau(\cdot)$ 为多头自注意力（multi-head self-attention，MSA）机制计算，具体见式（7），$\psi(\cdot)$ 为多层感知器（multi-layer perceptron，MLP）神经网络，具体见式（8），$\varphi(\cdot)$ 为层规范化操作。

$$\begin{aligned} \tau(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) &= \boldsymbol{W}^O[\boldsymbol{h}_1,\cdots,\boldsymbol{h}_i] \\ \boldsymbol{h}_i &= \text{softmax}\left(\frac{\boldsymbol{Q}_i\boldsymbol{K}_i^{\mathrm{T}}}{\sqrt{d}}\right)\boldsymbol{V}_i \end{aligned} \tag{7}$$

式中，$\boldsymbol{Q}$、$\boldsymbol{K}$ 和 $\boldsymbol{V}$ 分别为自注意力的 query 矩阵、key 矩阵和 value 矩阵，$\boldsymbol{Q}_i$、$\boldsymbol{K}_i$ 和 $\boldsymbol{V}_i$ 分别为多头自注意力第 i 头自注意力的 query 矩阵、key 矩阵和 value 矩阵，$\boldsymbol{W}^O$ 为神经网络权重，$\boldsymbol{h}_i$ 为第 i 头特征图，d 为多头注意力中每头特征维度。

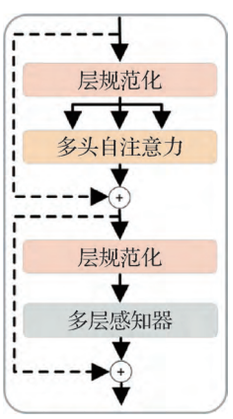


图 4 Transformer 结构图

Fig. 4 Transformer structure map

$$\psi(x) = \omega(0, x\boldsymbol{W}_1 + \boldsymbol{b}_1)\boldsymbol{W}_2 + \boldsymbol{b}_2 \tag{8}$$

式中，$\omega(\cdot)$ 为 ReLU 激活函数，$\boldsymbol{W}_1$ 和 $\boldsymbol{W}_2$ 为感知器权重，$\boldsymbol{b}_1$ 和 $\boldsymbol{b}_2$ 为感知器偏置。

2.4 损失函数

本文损失函数融合了多尺度重构损失、多尺度结构相似度损失和感知损失。网络训练过程更加稳定，可在训练过程中抑制梯度消失和梯度爆炸，加速网络收敛和提升网络去雨效果。

1）多尺度重构损失。为了准确恢复图像细节和捕获不同尺度特征上下文信息，使用 L2 范数建立多尺度重构损失（Wang 等，2021），计算为

$$\boldsymbol{L}_{\text{L}} = \sum_{s=0}^{3} \lambda_s \|\boldsymbol{I}_{\text{d}}^s - \boldsymbol{I}_{\text{c}}^s\|^2 \tag{9}$$

式中，除了 $\boldsymbol{I}_{\text{d}}^0$ 为解码器最后一层卷积层去雨输出，其尺度与输入有雨图像一致，其他 3 个尺度图像 $\boldsymbol{I}_{\text{d}}^1$、$\boldsymbol{I}_{\text{d}}^2$ 和 $\boldsymbol{I}_{\text{d}}^3$ 分别为解码器在第 9 层、第 6 层和第 3 层产生用于深监督训练的中间去雨图像，其图像宽高分别为原始输入有雨图像 1/2、1/4 和 1/8，$\boldsymbol{I}_{\text{c}}^s$ 为相应尺度用于监督训练无雨图像标签。

2）多尺度结构相似度损失。L2 范数收敛速度快，但训练过程不稳定，训练过程中易受到数据离群值影响。即使训练去雨图像与清晰图像计算峰值信噪比可能较高，但是图像部分区域结构仍可能受到破坏。因此，引入结构相似度损失（Ren 等，2019），提出多尺度结构相似度损失函数，以抑制数据离群值影响并加强去雨图像视觉效果。多尺度结构性相似度损失计算为

$$\boldsymbol{L}_{\text{S}} = -\sum_{s=0}^{3} \lambda_s \varpi(\boldsymbol{I}_{\text{d}}^s, \boldsymbol{I}_{\text{c}}^s) \tag{10}$$

式中，$\varpi(\cdot)$ 为结构相似度计算操作。

3）感知损失。参考 Wang 等人（2021）的去雨工作，添加感知损失以获取更好的去雨图像视觉质量。使用的感知损失计算为

$$\boldsymbol{L}_{\text{P}} = \|\rho(\boldsymbol{I}_{\text{d}}^0) - \rho(\boldsymbol{I}_{\text{c}}^0)\|^2 \tag{11}$$

式中，$\rho(\cdot)$ 为 VGG（Visual Geometry Group）16 卷积特征提取器。

网络总损失函数计算为

$$\boldsymbol{L}_{\text{T}} = \boldsymbol{L}_{\text{L}} + \boldsymbol{L}_{\text{S}} + \alpha\boldsymbol{L}_{\text{P}} \tag{12}$$

3 实验与结果分析

3.1 实验细节

本文提出的图像去雨网络使用合成雨线数据进行网络效果训练和测试，与CNN（convolutional neural network）（Fu等，2017a）、RESCAN（recurrent se context aggregation network）（Li等，2018）、PReNet（progressive recurrent network）（Ren等，2019）和AFR-Net（attentive feature refinement network）（Wang等，2021）等网络进行效果对比。

3.1.1 数据集和评估矩阵

图像去雨数据集为合成雨线数据集，首次在Zhang和Patel（2018）的工作中提出。包括低、中和高3种不同级别雨线密度合成有雨图像对，包含12 000对训练图像和1 200对测试图像。采用峰值信噪比（peak signal to noise ratio，PSNR）和结构相似性（structural similarity，SSIM）（Wang等，2004）作为矩阵评估方法。

3.1.2 实验相关设置

合成雨线数据集图像缩放至256×256像素。网络总训练周期数为100次，网络每迭代15次学习率将乘以0.5，不使用额外数据增强策略。网络参数采用Uniform初始化方法，Adam优化器训练网络，Adam参数设置为：$\beta_1=0.9$，$\beta_2=0.999$，$\varepsilon=1\times10^{-8}$，初始学习率为$1\times10^{-4}$。为了控制不同尺度对网络训练影响，参考Wang等人（2021）方法，式（9）和式（10）中λ_0、λ_1、λ_2和λ_3参数分别设置为1.0、0.8、0.6和0.4。式（12）中，感知损失权重α设置为0.1。

运行环境为：Windows 10，Python 3.7，Pytorch 1.6，CPU：Inter（R）Core（TM）I7-7820X，GPU：NVIDIA TITAN Xp，RAM：32 GB。

3.2 消融实验

3.2.1 损失函数与3维注意力

1）损失函数。绝对值（Fu等，2018）、均方差（Fu等，2017a）和结构相似性（Wang等，2004）等损失函数常用于去雨网络训练。结构相似性损失函数考虑了图像亮度、对比度和结构相似程度，更符合人类视觉特性。本文将结构性损失函数与其他常用去雨损失函数结合，并设计实验探究其有效性。首先使用多尺度均方差损失函数和感知损失函数作为baseline，并引入结构性损失函数。实验结果见表1。

表1 损失函数和3维注意力消融实验

Table 1 Results of ablation studies based on loss functions and three-dimension attention

SSIM损失函数		MSE感知损失		3DAM	指标（PSNR（/dB）/SSIM）
单尺度	多尺度	多尺度	单尺度		
×	×	√	√	×	31.57/0.904 9
√	×	√	√	×	32.37/0.923 6
×	√	√	√	×	32.69/0.925 6
×	√	√	√	√	**32.88/0.926 2**

注：加粗字体表示最优结果。

2）3维注意力。3DAM（Yang等，2021）通过优化能量函数计算每个神经元重要性，直接产生3维注意力权重，且无需存储注意力权重。在验证损失函数有效性后，本文使用所提出的损失函数训练加入3DAM的去雨网络，以测试3DAM在去雨任务中的有效性。

不同损失函数和加入3DAM训练得到的网络去雨效果见图5。可以看出，使用多尺度均方差和感知损失函数具有一定的去雨效果，但图像中部分区域存在着去雨涂抹痕迹或部分雨线并未被成功去除（图5（c））；加入单尺度结构相似性损失函数时，去雨效果得到加强，图像中大部分涂抹痕迹已被消除（图5（d）），例如图像中上层混凝土围栏和地面立柱顶部三角形区域附近的白色伪影；使用多尺度均方差、多尺度结构相似性损失函数和感知损失相结合方式网络训练时（图5（e）），相对单尺度结构相似性损失，公交车车窗区域获得了画面更纯净的去雨结果，地面立柱顶部三角形区域伪影进一步被消除；加入多尺度结构相似性损失函数和3维注意力训练得到的去雨结果见图5（f），图中更好地还原了上层混凝土围栏边缘细节，地面立柱三角形区域附近颜色过渡更自然。

实验表明，多尺度结构相似性损失在图像获得更好视觉观感的同时减少了局部区域伪影的出现。多尺度均方差、多尺度结构性损失和感知损失相结合的损失函数在网络结构相同情况下取得网络最优结果。本文最终选择此函数作为所提出TDATDN网络训练损失函数。

通过3DAM和密集块结构相结合，输入特征图计算3DAM权重，促使网络加强对去雨不完全区域处理，更好地恢复物体边缘纹理，本文将该机制加入

图 5　不同损失函数和模块消融实验去雨效果对比

Fig. 5　Results of ablation studies on the effects of different loss functions and modules

((a) rainy image; (b) ground truth; (c) baseline; (d) baseline + single-scale SSIM; (e) baseline + multi-scale SSIM; (f) baseline + multi-scale SSIM + 3DAM; (g) baseline + multi-scale SSIM + 3DAM + T1; (h) baseline + multi-scale SSIM + 3DAM + T2; (i) baseline + multi-scale SSIM + 3DAM + T3)

TDATDN 网络中进一步提升网络去雨性能。

3.2.2　Transformer 层数的选择

Transformer 主要应用于序列预测任务，其全局自注意力机制适合建模特征长期依赖，但其不擅长用于处理低层次特征位置信息。本文将 Transformer 模块和 U-Net 网络结合，加入 3DAM 组成 TDATDN 网络，用于单幅图像去雨任务。使用 Transformer 计算卷积编码器得到的高维图像特征块以提取全局语义信息，然后使用解码器逐步恢复图像到原始输入图像分辨率。使用 Transformer 的全局注意力机制计算特征内在关联，弥补了卷积只能对小范围特征依赖建模的缺点。

本文 Transformer 模块为可级联设计，为达到最佳去雨效果，本文提出设计不同数量的级联 Transformer 层以测试网络效果。不同 Transformer 层数训练得到的网络去雨效果见图 5。加入 Transformer 模块使网络具有计算全局特征关联能力，弥补传统卷积操作只能计算小区域特征关联的缺点，减少网络过度去雨和去雨不足现象。当 Transformer 层数为 1 时(即 T1)，图像去雨效果见图 5(g)，相对于没有添加 Transformer 层的网络，公交车车窗玻璃去雨后颜色过渡更加自然，地面立柱顶端三角形区域附近的

去雨伪影面积进一步减少。根据表 2 实验结果，级联多个 Transformer 模块可获得更好的去雨性能。当 Transformer 级联层数为 3 时，PSNR 和 SSIM 评估指标有轻微的下降，在权衡运行速度和去雨效果之后，本文最终选择 Transformer 级联层数为 2。多头注意力头数设置为 16，中间层维度设置为 256。

表 2　不同数量 Transformer 层在 TDATDN 网络的去雨效果

Table 2　Results of the TDATDN by different numbers of Transformer layers

指标	Transformer 层数				
	0	1	2	3	4
PSNR/dB	32.88	32.94	**33.01**	33.00	32.98
SSIM	0.926 2	0.927 5	**0.927 8**	0.927 3	0.927 4

注：加粗字体表示每行最优结果。

3.3　与其他去雨算法对比

3.3.1　合成有雨图像实验结果

所有网络使用相似训练参数重新训练。从实验结果表 3 看出，CNN 算法的 PSNR 值和 SSIM 值最低，其算法去雨性能有待提升；PReNet 去雨网络对比 RESCAN 算法 SSIM 有所提升，但 PSNR 略有降低；本文提出的 TDATDN 去雨算法对比以往单幅图像去雨算法均有明显性能提升，其中 PSNR 和 SSIM 对比 AFR-Net 网络分别提升了 0.67 和 0.010 3，取得了更好的去雨效果。

表 3　不同去雨算法在 Rain12000 数据集的去雨效果

Table 3　Results of different deraining algorithms on Rain12000 dataset

指标	CNN	RESCAN	PReNet	AFR-Net	本文
PSNR/dB	22.55	32.52	32.16	32.34	**33.01**
SSIM	0.812 3	0.911 6	0.919 5	0.917 5	**0.927 8**

注：加粗字体表示每行最优结果。

除了使用 PSNR 和 SSIM 评价指标对比，将现有主流单幅图像去雨算法和本文 TDATDN 算法进行去雨效果图像对比。对比展示见图 6，从结果上看出，CNN 算法去雨效果并不理想，整幅图像残留着大量雨线；RESCAN 算法虽取得不错的去雨效果，但整体图像依然遗留了部分雨雾；PReNet 和 AFR-Net 去雨算法取得较好效果，但原本蓝色的天空部分出现了白色去雨涂抹痕迹；本文算法对比 PReNet 和 RESCAN 去雨算法，极大降低了天空部分白色涂抹痕迹，对于图像高频信息取得较好的恢复效果。

图 6　不同算法的去雨效果对比

Fig. 6　Results of different rain removal methods

((a) rainy image; (b) ground truth; (c) CNN; (d) RESCAN; (e) PReNet; (f) AFR-Net; (g) ours)

从去雨效果细节图(图 7)中看出,RESCAN 和 PReNet 算法在砖墙处虽然去除了雨线,但红砖边缘已经被抹除;AFR-Net 虽保留了边缘信息,但有部分雨纹没有成功去除;本文提出的 TDATDN 对比 RESCAN、PReNet 和 AFR-Net 在地面纹理处和广告牌字母边缘都获得更好的去雨效果,保留了边缘信息,并极大地减少了去雨算法造成的伪影和涂抹痕迹。

(a) 雨图　(b) 真实值　(c) RESCAN

(d) PReNet　(e) AFR-Net　(f) 本文算法

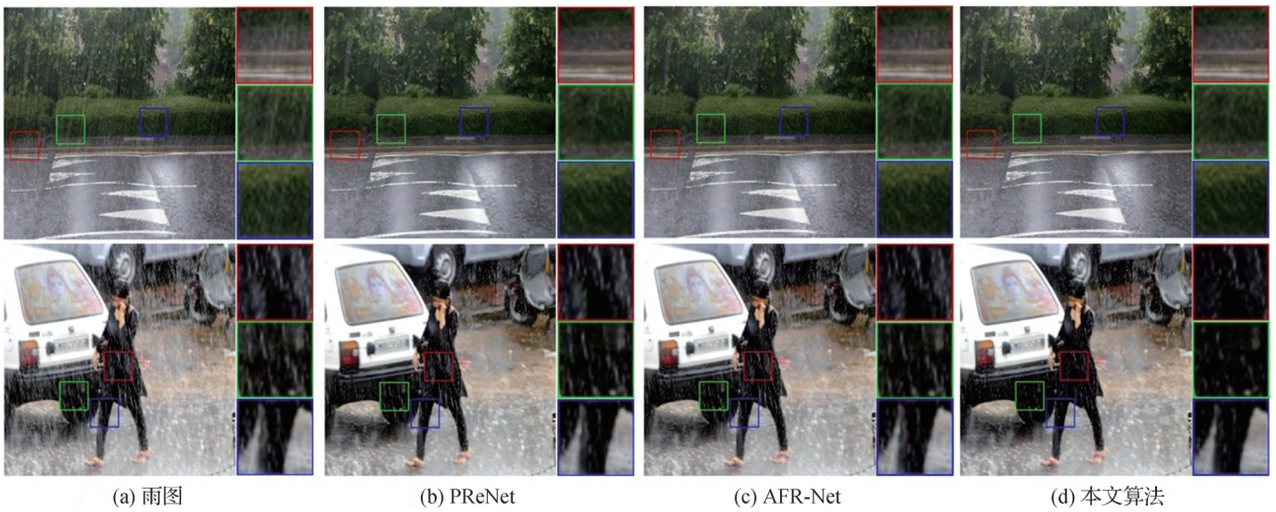
图 7　不同算法的去雨效果细节对比

Fig. 7　Detailed comparisons of rain removal effects of different algorithms

((a) rainy image; (b) ground truth; (c) RESCAN; (d) PReNet; (e) AFR-Net; (f) ours)

3.3.2　真实世界有雨图像实验结果

本文在真实世界有雨图像上进行去雨效果对比,以验证 TDATDN 网络去雨性能。因真实世界有雨图像无对应去雨标签图,本文将去雨图像进行视觉效果对比。对于部分真实世界场景,PReNet、AFR-Net 和本文 TDATDN 算法都能取得不错的去雨效果(如图 8 第 1 行图像);但在某些复杂场景下,如图 8 第 2 行图像,对于汽车轮胎和行人衣服上的



(a) 雨图　(b) PReNet　(c) AFR-Net　(d) 本文算法

图 8　不同算法的真实世界去雨效果细节对比

Fig. 8　Detailed comparisons of rain removal effects of different algorithms on real world

((a) rainy image; (b) PReNet; (c) AFR-Net; (d) ours)

雨线，PReNet 和 AFR-Net 去雨效果并不明显，但本文提出的 TDATDN 去雨算法仍可有效地完成去雨任务，在保留物体纹理信息的同时，较好地去除了复杂位置和物体边缘处雨线。

3.3.3　与其他算法运行时间对比

运行时间是评估去雨算法性能的重要指标，过长运行时间将使算法难以在实际计算机任务中应用。本文将不同去雨算法运行时间进行对比。具体做法为，将每个算法分别在尺寸为256×256像素和512×512像素的有雨图像上完成去雨时间计算，每个用于测试去雨时间的数据集包含200幅有雨图像，求每个数据集单幅图像的平均运行时间作为最终去雨运行时间对比。得到的不同去雨算法运行时间见表4。对比去雨时间相近的 PReNet 网络，本文可得到更好的去雨性能；对比现有去雨性能较优的 AFR-Net 网络，本文算法可在去雨性能提升的同时减少运行时间。TDATDN 算法在去雨性能和运行效率之间取得良好均衡，具有一定实用性。

表4　不同去雨算法运行时间对比

Table 4　Comparisons of running time of different algorithms

/s

分辨率/像素	CNN	RESCAN	PReNet	AFR-Net	本文
256×256	**0.020**	0.031	0.036	0.079	0.041
512×512	**0.060**	0.109	0.140	0.293	0.180

注：加粗字体表示每行最优结果。

4　结　论

针对现有单幅图像去雨网络去雨不完全或过度去雨问题，提出了 TDATDN 单幅图像去雨算法。将3维注意力与密集块结合解决高维度特征融合问题，使用残差密集块构建基于 U-Net 的图像去雨网络，并使用 Transformer 机制计算网络深层语义信息全局上下文关联性，同时将多尺度结构相似性与常用去雨损失函数结合参与网络训练。所提出的 TDATDN 单幅图像去雨网络结合了3维注意力机制、Transformer 和编码器—解码器架构优点，实验结果证明了本文算法的有效性。提出的单幅图像去雨算法并未对 Transformer 内部结构进行精简化，这导致本文算法现在无法很好地应用在实时视频去雨领域。今后研究可以对 Transformer 结构进行精简，在保证去雨效果的同时更好地提升视频去雨效率。

参考文献(References)

Child R, Gray S, Radford A and Sutskever I. 2019. Generating long sequences with sparse transformers [EB/OL]. [2021-08-13]. https://arxiv.org/pdf/1904.10509.pdf

Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X H, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J and Houlsby N. 2021. An image is worth 16×16 words: transformers for image recognition at scale [EB/OL]. [2021-08-13]. https://arxiv.org/pdf/2010.11929.pdf

Fu X Y, Huang J B, Ding X H, Liao Y H and Paisley J. 2017a. Clearing the skies: a deep network architecture for single-image rain removal. IEEE Transactions on Image Processing, 26(6): 2944-2956 [DOI: 10.1109/TIP.2017.2691802]

Fu X Y, Huang J B, Zeng D L, Huang Y, Ding X H and Paisley J. 2017b. Removing rain from single images via a deep detail network//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 1715-1723 [DOI: 10.1109/CVPR.2017.186]

Fu X Y, Liang B R, Huang Y, Ding X H and Paisley J. 2018. Lightweight pyramid networks for image deraining [EB/OL]. [2021-08-13]. https://arxiv.org/pdf/1805.06173.pdf

Guo T A, Dai T, Li J W and Xia S T. 2019. Self-attentive pyramid network for single image de-raining//Proceedings of the 26th International Conference on Neural Information Processing. Sydney, Australia: Springer: 390-401 [DOI: 10.1007/978-3-030-36708-4_32]

He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vega, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]

Hu J, Shen L, Albanie S, Sun G and Wu E H. 2020. Squeeze-and-excitation networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 42(8): 2011-2023 [DOI: 10.1109/TPAMI.2019.2913372]

Huang G, Liu Z, Van Der Maaten L and Weinberger K Q. 2017. Densely connected convolutional networks//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE: 2261-2269 [DOI: 10.1109/CVPR.2017.243]

Itti L, Koch C and Niebur E. 1998. A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(11): 1254-1259 [DOI: 10.1109/34.730558]

Kang L W, Lin C W and Fu Y H. 2012. Automatic single-image-based

rain streaks removal via image decomposition. IEEE Transactions on Image Processing, 21(4): 1742-1755 [DOI: 10.1109/TIP.2011.2179057]

Lateef F and Ruichek Y. 2019. Survey on semantic segmentation using deep learning techniques. Neurocomputing, 338: 321-348 [DOI: 10.1016/j.neucom.2019.02.003]

Li X, Wu J L, Lin Z C, Liu H and Zha H B. 2018. Recurrent squeeze-and-excitation context aggregation net for single image deraining//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer: 262-277 [DOI: 10.1007/978-3-030-01234-2_16]

Li Y, Tan R T, Guo X J, Lu J B and Brown M S. 2016. Rain streak removal using layer priors//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE: 2736-2744 [DOI: 10.1109/CVPR.2016.299]

Liu L, Ouyang W L, Wang X G, Fieguth P, Chen J, Liu X W and Pietikäinen M. 2020. Deep learning for generic object detection: a survey. International Journal of Computer Vision, 128(2): 261-318 [DOI: 10.1007/s11263-019-01247-4]

Li W K, Chia W L and Yu H F. 2012. Automatic single-image-based rain streaks removal via image decomposition. IEEE Transactions on Image Processing, 21(4): 1742-1755 [DOI: 10.1109/TIP.2011.2179057]

Luo Y, Xu Y and Ji H. 2015. Removing rain from a single image via discriminative sparse coding//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE: 3397-3405 [DOI: 10.1109/ICCV.2015.388]

Ma L, Liu R S, Jiang Z Y, Wang Y Y, Fan X and Li H J. 2018. Rain streak removal using learnable hybrid MAP network. Journal of Image and Graphics, 23(2): 277-285 (马龙, 刘日升, 姜智颖, 王怡洋, 樊鑫, 李豪杰. 2018. 自然场景图像去雨的可学习混合MAP网络. 中国图象图形学报, 23(2): 277-285) [DOI: 10.11834/jig.170390]

Parmar N, Vaswani A, Uszkoreit J, Kaiser Ł, Shazeer N, Ku A and Tran D. 2018. Image transformer [EB/OL]. [2021-08-13]. https://arxiv.org/pdf/1802.05751.pdf

Rawat W and Wang Z H. 2017. Deep convolutional neural networks for image classification: a comprehensive review. Neural Computation, 29(9): 2352-2449 [DOI: 10.1162/neco_a_00990]

Ren D W, Zuo W M, Hu Q H, Zhu P F and Meng D Y. 2019. Progressive image deraining networks: a better and simpler baseline//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 3932-3941 [DOI: 10.1109/CVPR.2019.00406]

Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]

Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser Ł and Polosukhin I. 2017. Attention is all you need//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: Curran Associates Inc.: 6000-6010

Wang F, Jiang M Q, Qian C, Yang S, Li C, Zhang H G, Wang X G and Tang X O. 2017. Residual attention network for image classification//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE: 6450-6458 [DOI: 10.1109/CVPR.2017.683]

Wang G Q, Sun C M and Sowmya A. 2020. Cascaded attention guidance network for single rainy image restoration. IEEE Transactions on Image Processing, 29: 9190-9203 [DOI: 10.1109/TIP.2020.3023773]

Wang G Q, Sun C M and Sowmya A. 2021. Attentive feature refinement network for single rainy image restoration. IEEE Transactions on Image Processing, 30: 3734-3747 [DOI: 10.1109/TIP.2021.3064229]

Wang M H, He H J and Li C. 2020. Single image rain removal based on selective kernel convolution using a residual refine factor. Journal of Image and Graphics, 25(12): 2484-2493 (王美华, 何海君, 李超. 2020. 自适应卷积的残差修正单幅图像去雨. 中国图象图形学报, 25(12): 2484-2493) [DOI: 10.11834/jig.190682]

Wang Z, Bovik A C, Sheikh H R and Simoncelli E P. 2004. Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4): 600-612 [DOI: 10.1109/TIP.2003.819861]

Woo S, Park J, Lee J Y and Kweon I S. 2018. CBAM: convolutional block attention module//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer: 3-19 [DOI: 10.1007/978-3-030-01234-2_1]

Yang L X, Zhang R Y, Li L D and Xie X H. 2021. SimAM: a simple, parameter-free attention module for convolutional neural networks//Proceedings of the 38th International Conference on Machine Learning. Virtual event: ICML: 11863-11874

Yang W H, Tan R T, Feng J S, Liu J Y, Guo Z M and Yan S C. 2016. Joint rain detection and removal via iterative region dependent multi-task learning [EB/OL]. [2021-08-13]. https://arxiv.org/pdf/1609.07769.pdf

Yang W H, Tan R T, Feng J S, Liu J Y, Guo Z M and Yan S C. 2017. Deep joint rain detection and removal from a single image//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE: 1685-1694 [DOI: 10.1109/CVPR.2017.183]

Yang W H, Tan R T, Wang S Q, Fang Y M and Liu J Y. 2019. Single image deraining: from model-based to data-driven and beyond [EB/OL]. [2021-08-13]. https://arxiv.org/pdf/1912.07150.pdf

Yasarla R and Patel V M. 2019. Uncertainty guided multi-scale residual learning-using a cycle spinning CNN for single image de-raining//

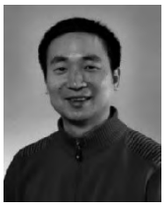
Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE: 8397-8406 [DOI: 10.1109/CVPR.2019.00860]

Zhang H and Patel V M. 2018. Density-aware single image de-raining using a multi-stream dense network//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 695-704 [DOI: 10.1109/CVPR.2018.00079]

Zhang Y L, Tian Y P, Kong Y, Zhong B N and Fu Y. 2018. Residual dense network for image super-resolution//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 2472-2481 [DOI: 10.1109/CVPR.2018.00262]

作者简介

王美华，1970 年生，女，副教授，硕士生导师，主要研究方向为模式识别、机器学习、机器视觉。
E-mail: wangmeihua@scau.edu.cn



范衡，男，通信作者，教授，博士生导师，主要研究方向为人工智能与机器人、设计自动化、进化计算、群体智能、图像处理。
E-mail: zfan@stu.edu.cn

柯凡晖，男，硕士研究生，主要研究方向为机器学习、计算机视觉。E-mail: weian@stu.scau.edu.cn

梁云，女，教授，硕士生导师，主要研究方向为机器视觉、大数据处理、机器学习。E-mail: sdliangyun@163.com

廖磊，男，硕士研究生，主要研究方向为机器学习、计算机视觉。E-mail: ll1423543604@stu.scau.edu.cn