

# 基于多尺度前馈结构的单幅图像去雨算法

江智呈<sup>1</sup>, 李志伟<sup>1,2\*</sup>, 陈晨<sup>1</sup>, 周进祥<sup>1</sup>, 周武能<sup>2</sup>

<sup>1</sup>上海工程技术大学电子电气工程学院, 上海 201620;

<sup>2</sup>东华大学信息科学与技术学院, 上海 200051

**摘要** 图像中雨纹的存在将增大目标检测和识别的难度, 而雨纹通常是图像的高频部分, 其中也包含着大量图像细节, 如何在去雨的同时保留有用的细节是需要解决的问题。为避免预处理带来去雨次优效果, 模拟真实场景下物体成像过程, 改进图像复原的通用模型, 丰富网络的感受野, 更精确去雨的同时增强对比度, 提出一种基于多尺度、多层次和多注意力机制的卷积神经网络, 通过多卷积的特征跳跃连接补偿卷积过程中细节信息的丢失以及融合不同层级特征信息, 在网络分支中提取多尺度特征图, 并结合注意力机制形成多个多尺度残差注意力子模块对全局信息在通道维度上进行权重重标定, 去冗余的同时增强有用信息, 将初级特征与高级特征相融合来学习雨图和无雨图之间的映射关系。考虑到实际情况, 真实的雨图无对应的无雨图, 所以本文采用合成的数据集来训练, 并用合成数据集和真实场景图进行验证。实验结果表明, 所提出的网络, 无论雨纹的大小和密度, 都能取得较好的去雨效果, 同时保留了细节信息。

**关键词** 图像去雨; 注意力机制; 多尺度卷积神经网络; 前馈结构

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP202259.1215001

## Multiscale Feedforward Structure-Based Single Image Rain Removal Algorithm

Jiang Zhicheng<sup>1</sup>, Li Zhiwei<sup>1,2\*</sup>, Chen Chen<sup>1</sup>, Zhou Jinxiang<sup>1</sup>, Zhou Wuneng<sup>2</sup>

<sup>1</sup>College of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China;

<sup>2</sup>College of Information Science and Technology, Donghua University, Shanghai 200051, China

**Abstract** The presence of rain patterns in an image increases the difficulty of target detection and recognition. Rain patterns are the high-frequency parts of the image, which contain several image details. However, a major challenge is in removing the rain pattern while retaining useful details. Thus, we proposed a method based on a multiscale convolutional neural network developed using multilevel and multiattention mechanisms. To avoid the suboptimal effect of rain removal caused by preprocessing, we simulated the imaging process of objects in real scenes, improved the general image restoration model, enriched the network's receptive field range, and accurately removed rain while enhancing contrast. Furthermore, we extract multiscale feature maps from the network branch using a multiconvolution feature jump connection to compensate for the loss of detailed information in the convolution process and fusion of different levels of feature information. Additionally, we combined attention to form multiple multiscale residual attention submodules to recalibrate the global information

收稿日期: 2021-04-13; 修回日期: 2021-05-08; 录用日期: 2021-06-02

基金项目: 国家自然科学基金 (61705127)、教育部产学合作协同育人项目 (201701014003, 201802046057)、上海工程技术大学检测技术与自动化装置学科学位点建设项目 (19XXK003)

通信作者: \*zhiwei.li@sues.edu.cn

in the channel dimension, removing redundancy while enhancing useful information and the primary and advanced features were fused to learn the mapping relationship between the rain and no-rain maps. Considering that the real rain map has no corresponding rain-free map, we used a synthetic dataset for training, and used the synthetic dataset and real scene graphs for verification. The experimental results show that our proposed network achieved a good rain removal effect while retaining detailed information irrespective of the size and density of the rain pattern.

**Key words** rain streak removal; attention mechanism; multiscale convolutional neural network; feedforward structure

## 1 引言

户外图像的可见度很大程度上取决于天气的状况,雨纹的存在会覆盖高频信息,造成图像特征信息模糊、质量低下,从而加大后续高级计算机视觉任务的难度,而机器视觉的应用场景无法避免恶劣的天气。因此,图像去雨已成为机器视觉领域中一个重要的研究分支。近年来,传统算法不断得到改进、创新,各种神经网络已取得较好的去雨效果,推动了图像去雨算法研究领域的快速发展。

目前,单幅图去雨算法可以大致分为两类:基于传统模型算法和基于深度学习模型算法。对于传统的去雨算法,文献[1]提出将图像进行雨纹位置定位,通过高频滤波将定位到的雨纹去掉,但该算法基于物理假定模型,应用受到很大的限制,不适合广泛推广。文献[2]通过低通滤波将雨纹图像分解为低频部分和高频部分,从高频部分捕捉无雨图像细节,将无雨图像细节叠加到低频部分,利用暗通道先验算法进一步缓解降雨造成的模糊,但是最终图像的雨痕过于明显。文献[3]提出了一种基于高斯混合模型的先验算法,该算法在大多数不同方向和密度的雨纹上都能取得不错的效果,但难度在于无法准确构建出各种情况下的高斯混合模型。传统的单幅图去雨算法多是基于数学模型优化去雨效果<sup>[1-5]</sup>,运行速度慢且效果有待提升。对于深度学习模型算法,文献[6]通过引导滤波将图像分为基础层和细节层,提出神经网络(DerainNet)对细节层学习去雨负映射,将映射得到的图像叠加到基础层,取得了优异的效果,尽管该卷积神经网络(CNN)增加了图像增强层,但在某些场景下还是会造成颜色失真并且雨纹残留明显。文献[7]提出了一种基于深度卷积的递归神经网络,将不同方向的雨纹看作是多个雨纹层的叠加,通过整合挤压-激励模块(RESCAN),根据强度和透明度为不同雨纹层设置阈值,然后分解成多个阶段逐步达到去雨的效果。文献[8]提

出一种新的基于密度感知的多流密集连接卷积神经网络算法(DID-MDN),利用不同尺度的特征提取不同尺度和形状的雨纹以自动确定雨密度信息,并根据估计的雨密度标签有效去除相应的雨纹。文献[9]提供了一个二元地图,包括雨纹层和背景图,学习二进制雨纹图、雨纹的外观和干净的背景,利用不同尺寸的核形成不同尺度的感受野,每次递推中,一个新的上下文信息的扩展网络被开发(JORDER),从而利用区域上下文信息反复地检测和去除雨纹。目前,单幅图像去雨算法常用滤波得到图像的无雨层和雨痕层,然后对雨痕层求解负残差映射叠加到雨图上,学习像素到像素的映射关系,同时也增加了额外步骤<sup>[7-14]</sup>。

本文提出一种端到端的去雨 CNN 模型,称为多尺度前馈结构单幅图像去雨 CNN,并不对图像进行滤波等预处理(这些预处理可能会使图形变得更加模糊和扭曲,导致次优的恢复性能)。本算法直接改进现有的图像增强模型,并根据残差网络的优势,在网络的不同层次提取出多细节进行跳跃式拼接,补偿卷积过程中图像细节的丢失,增强像素点与特征间的关联性,以多个多尺度残差注意力子模块(MSRAS)为分支结构,提取卷积过程中丢失的不同层次特征信息的同时,依靠注意力机制给不同尺度的特征通道赋予权重,增加感受野范围,减少特征信息的冗余,让不同像素点之间建立起基于长距离的映射关系,继而实现端到端的单幅图像去雨算法。

## 2 本文算法

改进现有的图像增强模型,学习特定参数,网络结构主体采用跳跃式拼接,分支上采用多个 MSRAS,形成前馈式深层 CNN,将初级特征信息、中层信息与主体框架的末端高级信息相融合,加强像素与距离之间的依赖关系,去除雨纹的同时保留高频信息。

### 2.1 场景建模

目前机器视觉领域中图像复原的通用模型都是利用退化现象的某种先验知识,逆向处理得到恢复图<sup>[15-17]</sup>,在退化矩阵为线性的条件下产生图像的过程为

$$g(x, y) = h(x, y) * f(x, y) + \eta(x, y), \quad (1)$$

式中: $g(x, y)$ 表示观察到的雨纹图像; $h(x, y)$ 表示退化矩阵; $f(x, y)$ 表示无雨图像; $\eta(x, y)$ 表示噪声矩阵;“\*”表示卷积操作。通过傅里叶变换,在频域中,式(1)为

$$G(u, v) = H(u, v)F(u, v) + N(u, v). \quad (2)$$

然而实际得到的图像,不仅有明显的雨点和雨纹,且光线在雨水的作用下会发生折射,从而影响周围的像素值,这个模型过于局限,根据文献[8]所研究的内容,像素点层次研究表明,真实场景下,雨纹所在区域造成高频信息在亮度、形状、位置发生一定的变化即模糊化、变形和衰减,具体来说,雨纹使得高频位置处的像素点数值发生了变化。这种区域像素点连续性变化,在视觉上就是目标的位置发生了偏移,即所谓的“凸镜成像”,偏移量与像素点之间的距离有关。根据文献[18-19]中去雨雾的算法,并结合式(1)和(2),将雨下成像的过程描述成区域像素值加权叠加的结果,这样高频信息保留的同时,有利于将雨痕分离,得到的无雨图像经过对比度增强算法处理,即可得

$$S(i, j) = L(i, j)R(i, j) + [E - L(i, j)]Z(i, j), \quad (3)$$

式中: $i$ 和 $j$ 表示图像像素点的位置; $S(\cdot)$ 表示观测到的雨纹图; $R(\cdot)$ 表示无雨图像; $E$ 表示单位矩阵; $L(\cdot)$ 表示区域内的权重系数矩阵,其元素数值处于0到1之间; $Z(\cdot)$ 表示自适应矩阵。 $Z(\cdot)$ 项将式(2)中的 $N(\cdot)$ 噪声项和图像自适应增强操作结合,系数项矩阵 $(E - L)$ 确保不超过最大像素值,将式(3)变形得到

$$R(i, j) = \frac{S(i, j) - [E - L(i, j)]Z(i, j)}{L(i, j)}. \quad (4)$$

式(4)出现两个重要参数矩阵 $L$ 和 $Z$ ,两次求参过程会使误差变大,为避免中间参数的形成,由式(4)得到

$$R(i, j) = B(i, j)[S(i, j) - E] + bE, \quad (5)$$

$$\text{式中, } B(i, j) = \frac{S(i, j) - Z(i, j) + L(i, j)[Z(i, j) - b]}{L(i, j)[S(i, j) - E]},$$

这样,参数矩阵 $L$ 和 $Z$ 被整合到新变量 $B(\cdot)$ 中,偏置矩阵 $b$ 可为任意值,本算法默认数值为1, $B$ 项使式(4)中的 $L$ 项和自适应矩阵 $Z$ 在实验中联合学习,根据文献[20]中实验数据对比得到,联合学习让2个参数相互补偿减少误差,共同决定最后网络的去雨效果。

### 2.2 通道注意力机制

Hu等<sup>[21]</sup>提出 squeeze-and-excitation networks (SE-Net),该网络专注于特征通道之间互相依赖的关系,通过网络学习的方式来自动获取到每个特征通道的重要程度,最后再为每个通道赋予不同的权重系数,从而来强化重要的特征抑制非重要的特征,如图1所示。通过特征重标定的方式来自适应调整通道之间的特征响应,让计算资源分配给信息中最有用的部分,又因为其额外所占用的计算成本小,很容易嵌入其他深层网络,Basak等<sup>[22]</sup>提出一种基于通道注意力的残差网络实现超分辨率算法,SE block也大大提升了高级计算机视觉任务如目标检测、识别等深层网络的效果<sup>[23-28]</sup>。

图1中, $H$ 、 $W$ 、 $C$ 表示特征的高度、宽度和通道数,GP表示全局池化(global pooling)操作,FC表示全连接(fully connected)操作, $r$ 表示特征压缩率,以此来减少计算的参数量。为不同的特征通道求得对应权重是关键点,首先是 squeeze 操作,将一个通道中整个空间特征编码为一个全局特征,通过 GP 操作将特征图在 $H \times W$ 的空间维度上收缩,定义为

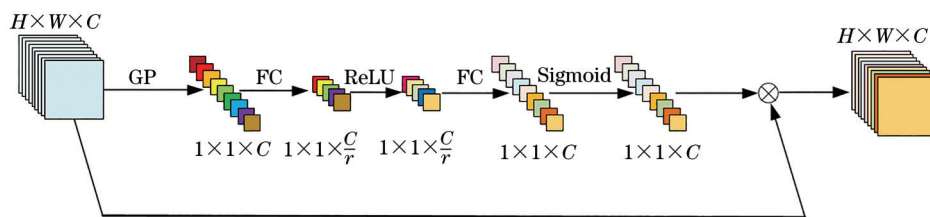


图1 SE block 结构图

Fig. 1 SE block structure diagram

$$Z_c = F_{sq}(x_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j), \quad (6)$$

式中:  $x_c(\cdot)$  表示特征图的第  $C$  个通道;  $F_{sq}(\cdot)$  表示全局池化。得到全局描述特征后, 进行 excitation 操作来抓取特征通道之间的关系, 定义为

$$s = F_{ex}(z, w) = \sigma(\delta(F_{sq}(x_c))), \quad (7)$$

式中:  $\sigma(\cdot)$  代表 Sigmoid 函数, 确保  $s$  值位于  $0 \sim 1$ ;  $\delta(\cdot)$  代表激活函数。最后是 scale 操作, 即将 excitation 操作得到的权重看作是信息重要程度的依据, 定义为

$$\hat{x} = F_{scale}(s, x_c) = s \cdot x_c, \quad (8)$$

式中,  $s$  作为比例因子与原通道相乘, 利用 1 维的稀疏卷积操作来优化 SE 模块中涉及到的全连接层操作来大幅降低参数量并保持相当的性能。为了压缩参数量和计算效率, SE-Net 采用的是“先降维再升维”的策略, 利用两个多层感知机来学习不同通道之间的相关性, 即当前的每一个特征图都与其他特征图进行交互, 是一种密集型的连接<sup>[29]</sup>。

### 2.3 多尺度残差注意力子模块

SE block 能很好地实现特征选择, 但在实际应

用中, 无法用 1 个模型准确描述出雨纹密集度、方向的多变性, 这就要求网络有丰富的感受野适应不同尺度和方向的特征提取, 本文在分支上采用 4 个 MSRAS, 分别为 MSRAS1、MSRAS2、MSRAS3 和 MSRAS4, 具体结构如图 2 所示。

4 个 MSRAS 的大体结构相同, 为了减少网络宽度, 只对上一层输出特征图提取 2 种或者 3 种通道信息通过批量归一化 (BN) 和 ReLU 后相拼接, 在空间维度上压缩, 进行自适应权值重标定。为减少网络计算量, 加快网络收敛速度, 但同时尽可能地保留特征信息, 对初级特征采取大尺寸卷积核, 同时减少卷积核个数; 又为增大感受野的范围, 对初级特征提取 3 种不同尺度的信息, 即  $7 \times 7$ 、 $5 \times 5$  和  $3 \times 3$ , 卷积核个数都为 16。随着网络的深入、特征细化, 需要的感受野也进一步变化。为满足需要, MSRAS 模块的卷积核尺寸相对应逐渐变小, 个数逐渐增加, 能提取出更多细化特征, 这样的设计可以使得网络根据实际需要自适应调节感受野的范围, 不会因为雨纹的大小和密度变化造成网络去雨效果退化, MSRAS 的参数设计如表 1 所示。

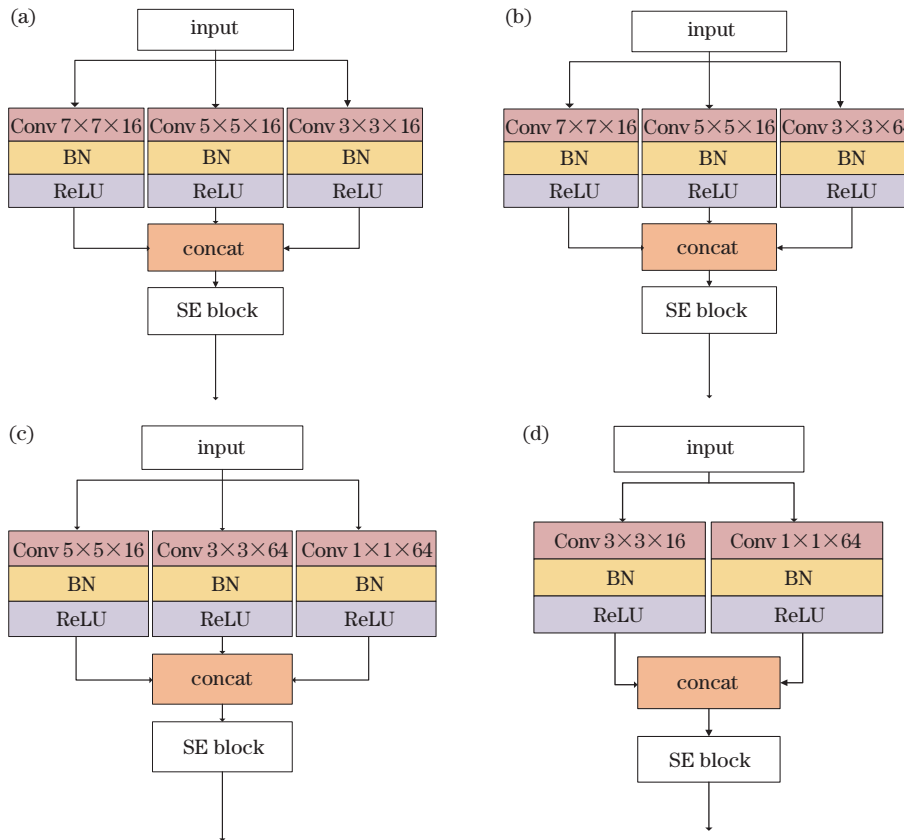


图 2 MSRAS 结构图。(a) MSRAS1; (b) MSRAS2; (c) MSRAS3; (d) MSRAS4

Fig. 2 MSRAS structure diagrams. (a) MSRAS1; (b) MSRAS2; (c) MSRAS3; (d) MSRAS4



表 1 MSRAS 的参数设定表  
Table 1 Parameter setting table of MSRAS

MSRAS	Filter size	Number of filters	Number of output channels
MSRAS1	7 × 7	16	48
	5 × 5	16	
	3 × 3	16	
MSRAS2	7 × 7	16	96
	5 × 5	16	
	3 × 3	64	
MSRAS3	5 × 5	16	96
	3 × 3	16	
	1 × 1	64	
MSRAS4	3 × 3	16	80
	1 × 1	64	

MSRAS 解决了下面几个问题:1)在分支上,提供了一个不同尺度卷积核所带来的特征通道学习机制,使神经元可以自适应调整感受野的大小,适应不同大小和方向的雨纹,增强网络的表达能力,提升去雨效果;2)改进简单的残差网络,加入 SE 模块和改变通道终点,对感受野提取的特征信息进行权值重标定,补偿卷积过程中遗失的重要信息;3)采取残差学习思想,防止梯度消失的同时,使原图的背景信息得以保留;4)传统的残差网络因其结构导致计算成本高,而用分支上的 MSRAS 代替残差模块,可以改善信息流动,加速收敛。

$$\begin{cases} I_1^1 = \delta(N_B(\omega_1^1 * I_{in1} + b_1^1)) \\ I_2^1 = \delta(N_B(\omega_2^1 * I_{in1} + b_2^1)) \\ I_3^1 = \delta(N_B(\omega_3^1 * I_{in1} + b_3^1)) \\ I_4^1 = I_1^1 \oplus I_2^1 \oplus I_3^1 \\ I_1'' = s(I_1^1) \\ I_{out1} = F_{scale}(I_1'', I_1^1) \end{cases}, \quad (9)$$

$$\begin{cases} I_1^2 = \delta(N_B(\omega_1^2 * I_{in2} + b_1^2)) \\ I_2^2 = \delta(N_B(\omega_2^2 * I_{in2} + b_2^2)) \\ I_3^2 = \delta(N_B(\omega_3^2 * I_{in2} + b_3^2)) \\ I_4^2 = I_1^2 \oplus I_2^2 \oplus I_3^2 \\ I_2'' = s(I_2^2) \\ I_{out2} = F_{scale}(I_2'', I_2^2) \end{cases}, \quad (10)$$

$$\begin{cases} I_1^3 = \delta(N_B(\omega_1^3 * I_{in3} + b_1^3)) \\ I_2^3 = \delta(N_B(\omega_2^3 * I_{in3} + b_2^3)) \\ I_3^3 = \delta(N_B(\omega_3^3 * I_{in3} + b_3^3)) \\ I_4^3 = I_1^3 \oplus I_2^3 \oplus I_3^3 \\ I_3'' = s(I_3^3) \\ I_{out3} = F_{scale}(I_3'', I_3^3) \end{cases}, \quad (11)$$

$$\begin{cases} I_1^4 = \delta(N_B(\omega_1^4 * I_{in3} + b_1^4)) \\ I_2^4 = \delta(N_B(\omega_2^4 * I_{in3} + b_2^4)) \\ I_4^4 = I_1^4 \oplus I_2^4 \\ I_4'' = s(I_4^4) \\ I_{out4} = F_{scale}(I_4'', I_4^4) \end{cases}, \quad (12)$$

式中:  $I_{in}^n$  和  $I_{out}^n$  ( $n = 1, 2, 3, 4$ ) 分别表示第  $n$  个 MSRAS 的输入输出;  $\omega$  表示权重参数;  $b$  表示偏置数值; “\*” 表示卷积操作;  $N_B(\cdot)$  表示批量归一化函数;  $\delta(\cdot)$  表示激活函数; “ $\oplus$ ” 表示特征信息拼接,  $s(\cdot)$  表示式(7)函数,  $F_{scale}$  表示式(8)中的函数。

### 3 网络总体框架

采用神经网络中传统的卷积形式而不使用池化层,将输入雨图先经过一个  $3 \times 3$  的卷积层,输出通道为 64,提取浅层特征,然后将  $3 \times 3$  卷积操作+批量归一化处理+ReLU 作为网络隐藏层的基本结构,将每层的输出特征作为下一层的输入,相比于标准的 CNN,本网络多了跳跃拼接结构和 4 个前馈结构的 MSRAS,在第 1 个卷积后连接第一个 MSRAS (MSRAS1),将第 2 层和第 3 层的输出在第 4 层拼接,同时由 MSRAS2 提取不同尺度特征信息进行权值重标定,将第 4 层和第 5 层的输出在第 6 层拼接,连接 MSRAS3, MSRAS3 的卷积核尺寸进一步变小,通道数增加,提取出更多细节,第 10 层将前 8 层经过卷积的信息在此拼接并连接 MSRAS4, MSRAS4 的卷积核尺度变小,通道数继续增大,最后将网络不同层次处经过权值重标定的特征图在第 13 层拼接,经过 2 次卷积和计算层输出无雨效果图,具体结构如图 3 所示,网络总体参数设计如表 2 所示。

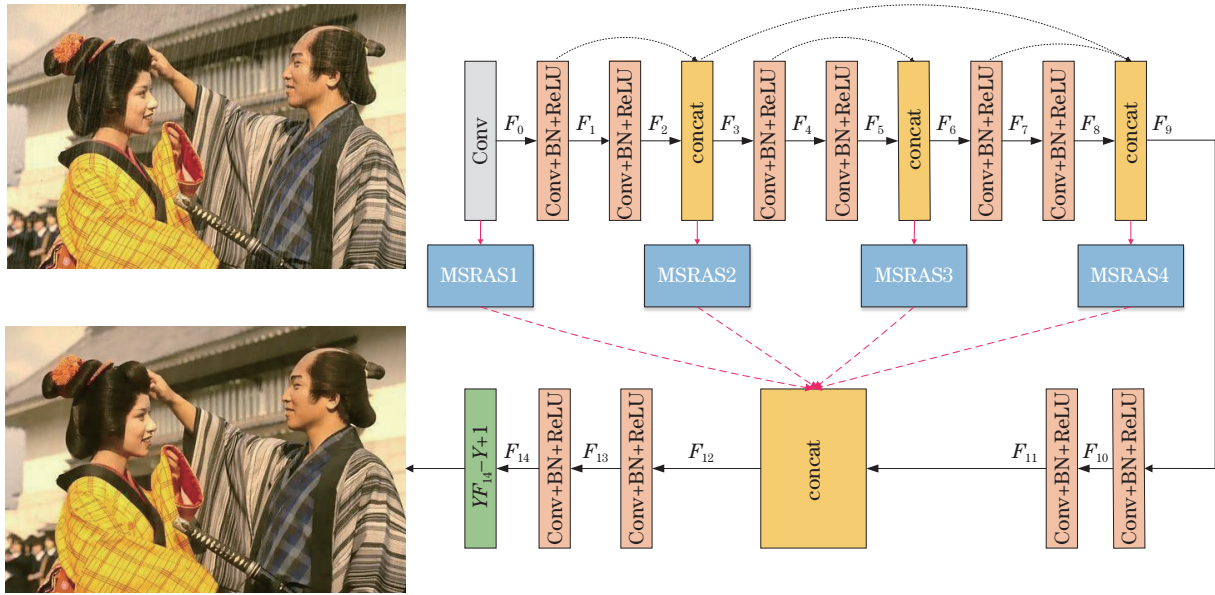


图 3 网络总体结构图

Fig. 3 Overall network structure diagram

表 2 网络参数表

Table 2 Network parameter table

Input size	Filter size	Number of filters	Activation function	Output	Output size
[128, 64, 64, 3]	3 × 3	64	ReLU	$F_0$	[128, 64, 64, 64]
[128, 64, 64, 64]	3 × 3	64	ReLU	$F_1$	[128, 64, 64, 64]
[128, 64, 64, 64]	3 × 3	64	ReLU	$F_2$	[128, 64, 64, 64]
[128, 64, 64, 64]	-	-	-	$F_3$	[128, 64, 64, 128]
[128, 64, 64, 128]	3 × 3	64	ReLU	$F_4$	[128, 64, 64, 64]
[128, 64, 64, 64]	3 × 3	64	ReLU	$F_5$	[128, 64, 64, 64]
[128, 64, 64, 64]	-	-	-	$F_6$	[128, 64, 64, 128]
[128, 64, 64, 128]	3 × 3	64	ReLU	$F_7$	[128, 64, 64, 64]
[128, 64, 64, 64]	3 × 3	64	ReLU	$F_8$	[128, 64, 64, 64]
[128, 64, 64, 64]	-	-	-	$F_9$	[128, 64, 64, 384]
[128, 64, 64, 384]	3 × 3	64	ReLU	$F_{10}$	[128, 64, 64, 64]
[128, 64, 64, 64]	3 × 3	64	ReLU	$F_{11}$	[128, 64, 64, 64]
[128, 64, 64, 64]	-	-	-	$F_{12}$	[128, 64, 64, 384]
[128, 64, 64, 384]	3 × 3	16	ReLU	$F_{13}$	[128, 64, 64, 16]
[128, 64, 64, 16]	3 × 3	3	ReLU	$F_{14}$	[128, 64, 64, 3]

残差网络 (ResNet) 被嵌入到“端到端”的映射中, 虽然它们被共同训练, 但彼此间并不直接依赖, MSRAS 处于网络的不同层次, 保留着残差块的特性, 并有着共同的终点, 加强了该子模块的自适应能力, 相比残差模块彼此间的联系更紧密。传统的前馈结构通过附加映射, 将需要保留的信息在两个隐藏层之间传递, 而 MSRAS 的前馈连接方式确保信息传递的完整性, 提升信息传递的效率, 不使用池化层以避免维度变化带来误差。计算过程为

$$\begin{cases} F_0 = \text{Conv}(Y') \\ F_i = \delta(N_B(\text{Conv}(F_{i-1}))), & (i = 1, 2, 4, 5, 7, 8, 10, 12, 13, 14) \\ F_j = \text{concat}(F_{j-1}, F_{j-2}), & (j = 3, 6) \\ F_9 = \text{concat}(F_k), & (k = 1, 2, 4, 5, 7, 8) \\ F_{11} = \text{concat}(F_{10}, I_{outm}), & (m = 1, 2, 3, 4) \end{cases} \quad (13)$$

式中:  $Y'$  代表输入图像,  $F_i$  和  $F_j$  代表第  $i(j)$  个隐藏层的输出, 本网络为简化计算过程, 加速网络收敛, 每

个隐藏层卷积核个数不超过 64, 不改变特征图维度;  $\text{Conv}(\cdot)$  代表卷积操作;  $\text{concat}$  代表拼接操作;  $I_{\text{out}m}$  代表第  $m$  个 MSRAS 的输出, 计算过程为式(9)~(12), 采用跳跃式拼接结构, 连接低层特征信息, 补偿卷积过程中的信息丢失。网络最后是一个计算层, 通过式(5)得到去雨后的干净图像,  $b$  为默认值 1, 即

$$Y = Y'F_{14} - Y' + 1. \quad (14)$$

## 4 实验与结果分析

### 4.1 评价指标与数据集的选择

采用结构相似性(SSIM)和峰值信噪比(PSNR)作为指标来比较几种方法的去雨效果。考虑到实际中并没有大量的雨图及其对应的标签图, 无法通过神经网络训练来取得非线性的映射关系, 故采用合成数据集来训练网络。本文采用数据集 Rain100L、Rain100H、Rain800、Rain12000 等 4 个合成数据集和 1 个真实雨图数据集来训练和测试, 其中, Rain12000 有 2000 张无雨标签图, 每张标签图有对应 6 张不同方向与大小的雨图, 用 Rain12000 中的 10800 张图片训练网络, 其余 1200 张图片为测试集。

### 4.2 损失函数的选择

采用合成数据集最小化去雨效果图与标签图之间的差异优化网络参数, 即可得

$$L = \frac{1}{N} \sum_{n=1}^N \|R_n - Y_n\|_F, \quad (15)$$

式中:  $N$  为批量样本大小;  $R_n$  为截取的第  $n$  个区域的

对应标签图;  $Y_n$  为网络生成的第  $n$  个效果区域;  $\|\cdot\|_F$  表示求取矩阵的  $F$  范数。

### 4.3 实验配置与参数设置

实验系统为 Windows10, 显卡为英伟达 GeForce GTX 1080Ti, 在开发工具 Pycharm 平台上通过深度学习框架 Tensorflow1. 13. 1 实施, 训练过程中, 为了让批量处理的数据具有差异性, 随机读取 32 幅图像, 并在每幅图像中随机选择 4 个  $64 \text{ pixel} \times 64 \text{ pixel}$  的区域形成一批数据供给网络训练使用。为方便训练, 获取批量样本时应将其统一归一化到  $[0, 1]$ , 训练中采取随机梯度下降(SGD)<sup>[30]</sup>来最小化式(15), 权重衰减设置为  $10^{-10}$ , 动量设置为 0. 9, 学习率初始值为 0. 2, 前 6 万次迭代过程中, 学习率保持不变, 之后每百次迭代后学习率乘以 0. 995, 共训练 30 万次。

### 4.4 实验结果分析

从几幅图片的去雨前后视觉效果和数据定量分析(图 4 和表 3)可以看出: 不同结构在不同的图像下去雨效果差别较大, 传统的 ResNet 去雨后雨痕残留明显, 尤其在第一幅图和第四幅图中, 当背景颜色较深时, 留下了大量的雨纹, 高亮度区域的背景模糊, 在去雨纹的同时, 丢失了雨纹下的高频信息; 去掉 MSRAS 后, 相较于 ResNet, 算法在 SSIM 值和 PSNR 值提升了 0. 03~0. 05 和 2~3 dB, 但当背景颜色丰富时, 去雨后图像的 PSNR 数值低于 ResNet, 在视觉上雨痕比 ResNet 少了很多, 但会出现过拟合现

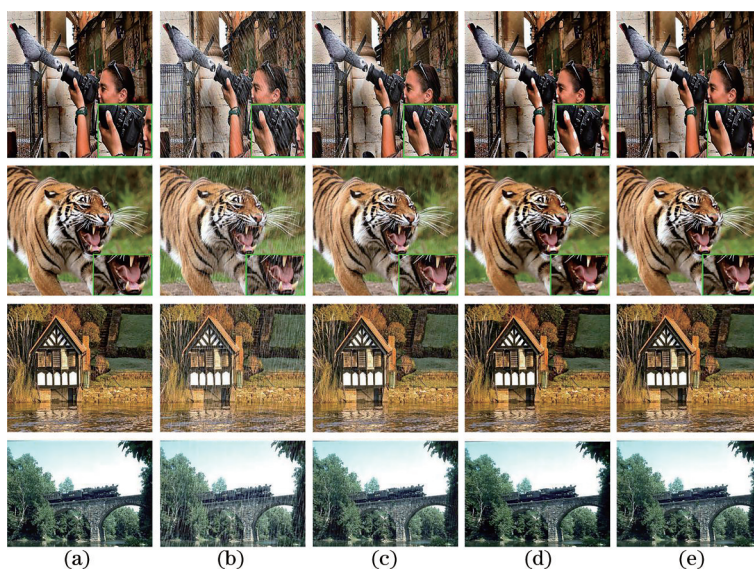


图 4 不同结构去雨效果比较图。(a) 真实图像; (b) 雨图; (c) ResNet; (d) 无 MSRAS; (e) 本文算法

Fig. 4 Comparison of rain removal effects of different structures. (a) Real pictures; (b) rain pictures; (c) ResNet; (d) without MSRAS; (e) ours



表 3 不同结构去雨效果定量值

Table 3 Quantitative values of rain removal effects of different structures

Picture	SSIM			
	Rain picture	ResNet	Without MSRAS	Ours
Photographer	0.7036	0.8915	0.8722	<b>0.9443</b>
Tiger	0.6968	0.8636	0.9057	<b>0.9569</b>
Cottage	0.7823	0.8932	0.9219	<b>0.9788</b>
Bridge	0.7733	0.8953	0.9032	<b>0.9487</b>
200 composite rain pictures	0.7135	0.8722	0.9063	<b>0.9504</b>
Picture	PSNR /dB			
	Rain picture	ResNet	Without MSRAS	Ours
Photographer	19.1911	27.9249	27.4652	<b>33.6541</b>
Tiger	20.0945	26.1495	28.5863	<b>34.0457</b>
Cottage	19.2886	26.3365	30.1487	<b>34.2584</b>
Bridge	18.6967	27.6713	28.9957	<b>32.5586</b>
200 composite rain pictures	18.3256	26.1841	28.9426	<b>33.9953</b>

象,在第一幅图中摄影师的手部放大区域和第二幅图老虎的牙齿放大部分,背景颜色与雨滴接近,在雨密集区域出现了白色斑点;而本文算法较传统的 ResNet来说,PSNR提升了5~7 dB,而SSIM也提升了0.05~0.09,在视觉上去掉雨纹的同时,最大限度地保留了纹理信息,说明MSRAS能增大网络的感受野,同时增强特征的有用维度,抑制无用信息。

分别与几种经典的去雨算法在SSIM、PSNR上作对比,如图5和表4所示。由图5可知,视觉上, DID-MDN仍保留了大量雨纹,无法达到高级机器视觉任务的要求;JORDER未能将暴雨中的雨纹去

除,同时模糊了图像细节,造成了过平滑现象; RESCAN在雨纹颜色和背景相似时,去雨效果急剧下降,在细节上残留了雨痕,负残差映射造成了过拟合,导致图像昏暗,降低了图像的对比度;而本文算法,在视觉上达到了最佳的效果,对不同方向和大小的雨纹都能去除,适应性强。由表4可知, Rain100L雨纹较细、密度较小,几种经典算法都能取得好的效果,但还是略低于本算法;当测试集为雨纹较大、密度大的Rain100H时,几种算法的差距明显变大。数值上,本算法的输出图在SSIM和PSNR上能达到0.85和26 dB以上,而用真实环境

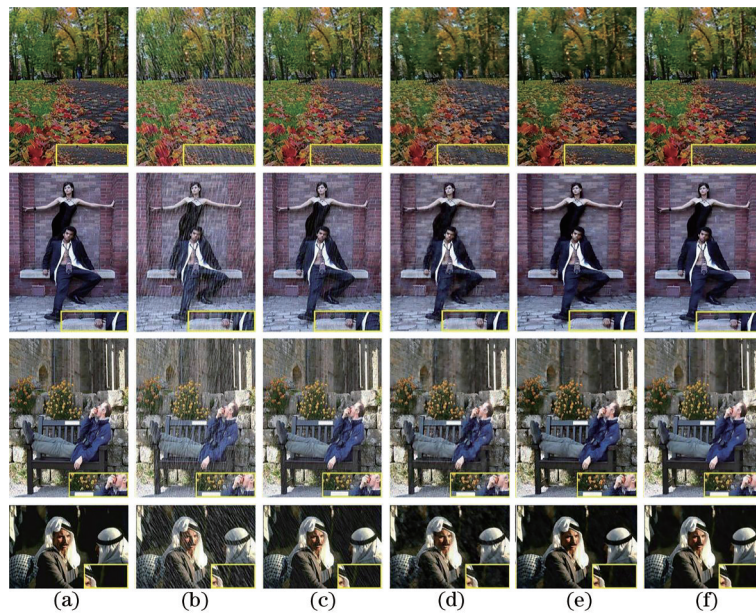


图 5 合成数据集下不同算法的去雨效果图。(a)真实图片;(b)雨图;(c) DID-MDN;(d) JORDER;(e) RESCAN;(f)本文算法  
Fig. 5 Rain removal effect diagram of different algorithms under the composite dataset. (a) Real pictures; (b) rain pictures; (c) DID-MDN; (d) JORDER; (e) RESCAN; (f) ours



表 4 合成数据集下不同算法的去雨效果定量值

Table 4 Quantitative values of rain removal effects of different algorithms under the composite dataset

Picture	SSIM			
	DID-MDN	JORDER	RESCAN	Ours
Path	0.9069	0.9166	0.9389	<b>0.9392</b>
Model	0.8872	0.9233	0.9258	<b>0.9473</b>
Man	0.9022	0.9341	0.9389	<b>0.9421</b>
Arab	0.8575	0.9178	0.9498	<b>0.9587</b>
Rain100L	0.8821	0.9702	0.9752	<b>0.9821</b>
Rain100H	0.7285	0.7633	0.8726	<b>0.8832</b>
Rain800	0.8858	0.8835	0.8942	<b>0.9289</b>
Rain12000	0.8125	0.8325	0.8432	<b>0.9072</b>

Picture	PSNR/dB			
	DID-MDN	JORDER	RESCAN	Ours
Path	29.8813	30.0288	31.8564	<b>32.0873</b>
Model	28.1575	32.573	33.9847	<b>34.2635</b>
Man	31.2462	34.7834	35.1448	<b>33.2516</b>
Arab	27.1485	33.2175	35.5248	<b>35.6874</b>
Rain100L	25.6564	36.1143	36.2145	<b>36.9326</b>
Rain100H	17.8715	25.7412	26.4121	<b>28.9627</b>
Rain800	26.5586	25.9984	30.2486	<b>31.7585</b>
Rain12000	24.2263	25.8896	26.7423	<b>28.9245</b>

下采集的 Rain800 作测试集时,几种经典算法的效果差距不大,但都要明显低于本算法。为进一步说

明本算法的有效性,本文在真实雨图下做视觉测试。如图 6 所示,从最终效果来看,其他几种算法都

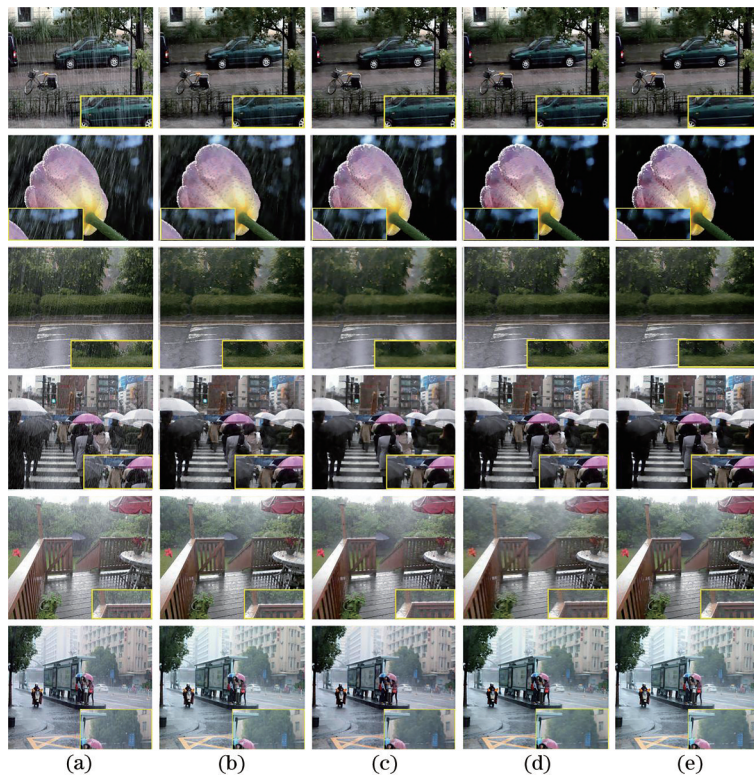


图 6 真实雨图下不同算法的去雨效果图。(a)真实图片;(b) DID-MDN;(c) JORDER;(d) RESCAN;(e)本文算法  
 Fig. 6 Rain removal renderings of different algorithms under real rain maps. (a) Real pictures; (b) DID-MDN; (c) JORDER; (d) RESCAN; (e) ours

存在雨痕残留明显、高频信息丢失、过拟合等问题,尤其是当雨纹较大、方向多变或与背景信息相似时,去雨效果大大降低。而本文算法能有效地解决上述问题,达到更好的去雨效果。本算法的泛化能力较好,能适应不同场景和不同高级机器视觉任务,具有更好的实用性。

#### 4.5 算法运行效率比较

分别在 200 张合成测试图和 6 张真实场景测试图上比较几种算法的运行效率,如表 5 所示,在计算复杂度上,同样的批量处理过程,本算法因为结构优势,比传统残差网络运行速度快 40% 左右,相较于其他几种神经网络算法,本文算法能提速 10%~30%,当处理数据进一步增加时,运行速度差异继续增大。

表 5 不同算法的效率值

Algorithm	ResNet	DID-MDN	JORDER	RESCAN	Ours
6 pictures	140	70	66	57	31
200 pictures	7340	5665	4995	4740	3910

## 5 结 论

本文所提出的单幅图去雨算法,效果上不仅能去掉不同密度、方向的雨痕,也能保住高频信息,不同于传统的负残差映射,本算法不对图像进行高频预处理,在经典的图像增强算法模型上改进,引入自适应项,让网络更新参数的同时,满足输出的需要,跳跃拼接能补偿传统卷积层带来的信息丢失,MSRAS 补充感受野范围的同时也能充当自适应调节器,平衡跳跃拼接所带来的信息冗余,避免在雨纹较小时产生过拟合。本算法的建模是在图像复原的通用模型上变换得到的,未来可以在现有基础上,进一步优化算法的计算成本和输出效果,尤其是真实图像的去雨效果。本算法不仅仅局限于去雨,在进一步研究和优化下,可达到去雨雾、去雨雪和去风暴等效果,在更多恶劣条件下,完成对图像快速又精确的处理。

#### 参 考 文 献

- [1] 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, September 15-18, 2013, Melbourne, VIC, Australia. New York: IEEE Press, 2013: 914-917.
- [2] Liu J C, Teng S H, Li Z Y. Removing rain from single image based on details preservation and background enhancement[C]//2019 IEEE 2nd International Conference on Information Communication and Signal Processing, September 28-30, 2019, Weihai, China. New York: IEEE Press, 2019: 322-326.
- [3] Li Y, Tan R T, Guo X J, et al. Rain streak removal using layer priors[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 2736-2744.
- [4] Luo Y, Xu Y, Ji H. Removing rain from a single image via discriminative sparse coding[C]//2015 IEEE International Conference on Computer Vision, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 3397-3405.
- [5] 王凯,李志伟,朱成德,等.基于二次引导滤波的局部立体匹配算法[J].激光与光电子学进展,2019,56(8):081004.  
Wang K, Li Z W, Zhu C D, et al. Local stereo matching algorithm base on secondary guided filtering [J]. Laser & Optoelectronics Progress, 2019, 56(8): 081004.
- [6] Fu X Y, Huang J B, Ding X H, et al. Clearing the skies: a deep network architecture for single-image rain removal[J]. IEEE Transactions on Image Processing, 2017, 26(6): 2944-2956.
- [7] Pan J S, Liu S F, Sun D Q, et al. Learning dual convolutional neural networks for low-level vision [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 3070-3079.
- [8] Zhang H, Patel V M. Density-aware single image de-raining using a multi-stream dense network[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 695-704.
- [9] Yang W H, Tan R T, Feng J S, et al. Deep joint rain detection and removal from a single image[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1685-1694.
- [10] 盖杉,王俊生.基于深度学习的非局部注意力增强网络图像去雨算法研究[J].电子学报,2020,48(10):1899-1908.  
Gai S, Wang J S. Image raindrop algorithm research

- using nonlocal attention enhanced network based on deep learning[J]. *Acta Electronica Sinica*, 2020, 48(10): 1899-1908.
- [11] Wang Y L, Chen C, Zhu S Y, et al. A framework of single-image deraining method based on analysis of rain characteristics[C]//2016 IEEE International Conference on Image Processing, September 25-28, 2016, Phoenix, AZ, USA. New York: IEEE Press, 2016: 4087-4091.
- [12] Ding Y, Xue X W, Wang Z Z, et al. Domain knowledge driven deep unrolling for rain removal from single image[C]//2018 7th International Conference on Digital Home (ICDH), November 30-December 1, 2018, Guilin, China. New York: IEEE Press, 2018: 14-19.
- [13] Liu R S, Jiang Z Y, Ma L, et al. Deep layer prior optimization for single image rain streaks removal [C]//2018 IEEE International Conference on Acoustics, Speech and Signal Processing, April 15-20, 2018, Calgary, AB, Canada. New York: IEEE Press, 2018: 1408-1412.
- [14] Li X, Wu J L, Lin Z C, et al. Recurrent squeeze-and-excitation context aggregation net for single image deraining[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11211: 262-277.
- [15] Imtiaz M S, Wahid K A. Image enhancement and space-variant color reproduction method for endoscopic images using adaptive sigmoid function[C]//2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, August 26-30, 2014, Chicago, IL, USA. New York: IEEE Press, 2014: 3905-3908.
- [16] Kim J K, Park J M, Song K S, et al. Adaptive mammographic image enhancement using first derivative and local statistics[J]. *IEEE Transactions on Medical Imaging*, 1997, 16(5): 495-502.
- [17] Ning Y, Li L, Qiu S, et al. Underwater range-gated laser imaging system design with video enhancement processing[C]//2013 2nd International Symposium on Instrumentation and Measurement, Sensor Network and Automation (IMSNA), December 23-24, 2013, Toronto, ON, Canada. New York: IEEE Press, 2013: 760-763.
- [18] 王丙元, 郑芳, 姜建, 等. 一种单幅图像去除雨雾的方法[J]. *激光与光电子学进展*, 2020, 57(14): 141027. Wang B Y, Zheng F, Jiang J, et al. Method for removal of rain and fog in single image[J]. *Laser & Optoelectronics Progress*, 2020, 57(14): 141027.
- [19] 刘增力, 付钰. 基于透射率自适应约束修正的图像去雾算法[J]. *激光与光电子学进展*, 2020, 57(16): 161003. Liu Z L, Fu Y. Image dehazing algorithm based on adaptive constraint correction of transmittance[J]. *Laser & Optoelectronics Progress*, 2020, 57(16): 161003.
- [20] Li B Y, Peng X L, Wang Z Y, et al. AOD-net: all-in-one dehazing network[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 4780-4788.
- [21] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011-2023.
- [22] Basak H, Kundu R, Agarwal A, et al. Single image super-resolution using residual channel attention network[C]//2020 IEEE 15th International Conference on Industrial and Information Systems, November 26-28, 2020, Rupnagar, India. New York: IEEE Press, 2020: 219-224.
- [23] 李聪, 蒋敏, 孔军. 基于多尺度注意力机制的多分支行人重识别算法[J]. *激光与光电子学进展*, 2020, 57(20): 201001. Li C, Jiang M, Kong J. Multi-branch person re-identification based on multi-scale attention[J]. *Laser & Optoelectronics Progress*, 2020, 57(20): 201001.
- [24] Qiu C C, Zhang S Y, Wang C, et al. Improving transfer learning and squeeze-and-excitation networks for small-scale fine-grained fish image classification [J]. *IEEE Access*, 2018, 6: 78503-78512.
- [25] Fu J P, Fan W T, Bouguila N. A novel approach for anomaly event detection in videos based on autoencoders and SE networks[C]//2018 9th International Symposium on Signal, Image, Video and Communications (ISIVC), November 27-30, 2018, Rabat, Morocco. New York: IEEE Press, 2018: 179-184.
- [26] An G Y, Zhou W, Wu Y X, et al. Squeeze-and-excitation on spatial and temporal deep feature space for action recognition[C]//2018 14th IEEE International Conference on Signal Processing, August 12-16, 2018, Beijing, China. New York: IEEE Press, 2018: 648-653.
- [27] Wu J Q, Li Q, Liang S, et al. Convolutional neural



- network with squeeze and excitation modules for image blind deblurring[C]//2020 Information Communication Technologies Conference (ICTC), May 29-31, 2020, Nanjing, China. New York: IEEE Press, 2020: 338-345.
- [28] Wang F, Jiang M Q, Qian C, et al. Residual attention network for image classification[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6450-6458.
- [29] Li K P, Wu Z Y, Peng K C, et al. Tell me where to look: guided attention inference network[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 9215-9223.
- [30] Huang G, Sun Y, Zhang L, et al. Deep networks with stochastic depth[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9908: 646-661.