

## 多尺度循环注意力网络运动模糊图像复原方法

王向军<sup>1,2</sup>, 欧阳文森<sup>1,2</sup>

- (1. 天津大学精密测试技术及仪器国家重点实验室, 天津 300072;
2. 天津大学微光机电系统技术教育部重点实验室, 天津 300072)

**摘要:** 在图像采集过程中, 由于拍摄对象运动或相机自身运动造成的图像模糊对于后续的高级视觉任务会产生很不利的影响。针对当前深度学习图像去模糊方法不能兼顾去模糊效果和效率的问题, 提出了一种多尺度循环注意力网络, 使用深度可分离卷积降低参数量, 改进注意力模块合理分配计算资源, 对卷积层进行密集型连接提高参数利用效率, 引入边缘损失提升生成图像边缘细节信息。经过实验验证, 所提方法具有良好的泛化性能和鲁棒性, 在 Lai 数据集和 Köhler 数据集上的 SSIM 和 PSNR 较近年典型方法的最佳效果分别提升了约 1.15%、0.86% 和 0.91%、1.04%, 在 GoPro 数据集上的平均单帧运行速度较同类方法提升约 2.5 倍。

**关键词:** 多尺度循环网络; 注意力机制; 密集型残差网络; 边缘损失

**中图分类号:** TP391.4      **文献标志码:** A      **DOI:** 10.3788/IRLA20210605

## Multi-scale recurrent attention network for image motion deblurring

Wang Xiangjun<sup>1,2</sup>, Ouyang Wensen<sup>1,2</sup>

- (1. State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China;
2. MOEMS Education Ministry Key Laboratory, Tianjin University, Tianjin 300072, China)

**Abstract:** In image acquisition process, the image blur caused by the moving subject or the camera itself will have a negative impact on the subsequent high-level vision tasks. Aiming at the problem that the current deep learning image deblurring method cannot balance the deblurring effect and efficiency, a multi-scale recurrent attention network was proposed, which used separable convolution to reduce the amount of parameters, and improved the attention module to allocate computing resources reasonably. Layers were used for dense connection to improve parameter utilization efficiency, and edge loss was introduced to improve the edge detail information in the generated image. Experiments prove that the proposed method has good generalization performance and robustness. Compared with the typical methods in recent years, the SSIM and PSNR have increased by about 1.15%, 0.86% and 0.91%, 1.04% on the Lai dataset and Köhler dataset, respectively. The average single frame running speed on the GoPro dataset is nearly 2.5 times faster than similar methods.

**Key words:** multi-scale recurrent network; attention mechanism; residual dense block; edge loss

收稿日期: 2021-08-10; 修订日期: 2021-10-15

作者简介: 王向军, 男, 教授, 博士, 主要从事光电传感与测试、计算机视觉与图像分析、微小型光机电系统及 MEMS 方面的研究。

## 0 引言

计算机视觉技术的发展极大地推动了相机在障碍探测<sup>[1]</sup>、目标识别<sup>[2]</sup>、目标跟踪<sup>[3]</sup>、遥感分析等领域的广泛应用。在图像的拍摄过程中不可避免地会由于物体运动或相机抖动而产生成像模糊问题。运动模糊使图像丢失了边缘和细节信息,对于后续的高级视觉任务产生不利影响。因此,对于图像运动模糊去除的研究具有重要的意义。图像模糊的过程可以用公式(1)表示:

$$I_B = I_S \otimes K + N \quad (1)$$

式中:  $I_B$  为模糊图像;  $I_S$  为清晰图像;  $K$  为模糊核;  $N$  为加性噪声。在图像生成的深度学习方法兴起之前,用于图像去模糊的传统方法<sup>[4-6]</sup>主要通过图像中的先验或统计信息,估计出模糊核  $K$ ,通过反卷积的方法恢复出清晰图像  $I_S$ 。参考文献<sup>[6-7]</sup>提出了基于概率统计的方法对模糊核进行估计。模糊核通常具有非一致性,一帧图像中不同像素一般会对应不同的模糊核,所以,求出每一个像素对应的模糊核是一个严重病态的问题。这类方法只对特定的图像有一定的去模糊效果,鲁棒性不强,数学模型较为复杂、计算效率较低。

2016年后,深度学习方法开始更多地应用于图像处理中。Nah<sup>[8]</sup>等人提出使用多尺度卷积神经网络去模糊,创建了目前应用最广泛的 GoPro 数据集,采用“端到端”的训练方式,模型效果和运算速度较传统的非盲去模糊方法提升较大。2018年, Kupyn<sup>[9]</sup>等提出 DeblurGAN,采用 pix2pix 的基本结构,用条件生成对抗网络训练去模糊生成器,获得了比 Nah 更好的去模糊效果和效率。同年, Tao<sup>[10]</sup>等人提出 SRN deblur

net,用尺度循环神经网络实现了“渐进优化”的去模糊过程,去模糊效果更好,达到近年最佳,但该方法单帧运行时间过长。2019年, Kupyn<sup>[11]</sup>等提出了 DeblurGAN-v2,可以选择性使用 Inception、MobileNet 等主干网络,分别达到最佳的去模糊效果或最快的效率,但二者不可兼得。2020年, Zhang<sup>[12]</sup>等连接了两个生成对抗网络 (Generative adversarial network, GAN),提出了利用 GAN 网络生成“清晰——模糊”图像对的方法,但网络过于复杂参数较多,两个 GAN 的协调训练难度较大。2020年, Qi<sup>[13]</sup>等人将注意力机制引入 GAN 网络中,同时在损失函数中引入边缘损失分量,能生成具有更加清晰边缘的清晰图像。

为了保证在一定运动模糊去除效果的前提下尽可能提高网络的运行速度,观察到参考文献<sup>[10]</sup>的方法具有最佳的定量去模糊效果,文中在 SRN deblur net 的基础上对于循环神经网络去除图像运动模糊的方法进行了研究。首先,文中构建了端到端学习的多尺度循环神经网络,避免了模糊核估计对于模糊去除的不良影响。其次,提出了一种密集型残差注意力网络作为总体网络的基础结构,为图像的不同通道和空间分配更加合理地计算资源。最后,设计了基于 Sobel 算子的边缘损失分量作为感知损失和基于欧氏距离 (Euclidean distance) 的  $L_2$  损失作为内容损失,总体损失函数由感知损失和内容损失构成。

## 1 多尺度循环注意力网络

### 1.1 多尺度循环网络架构

文中网络由从粗到精三个尺度的编解码网络构成,每个尺度共享同一套网络参数,如图 1 所示。输入的模糊图像缩放为  $64 \times 64$ 、 $128 \times 128$  和  $256 \times 256$  三

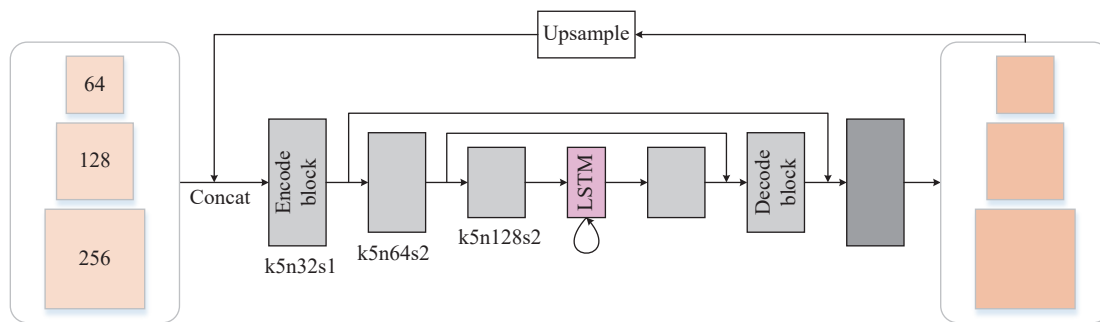


图 1 多尺度循环神经网络架构

Fig.1 Multi-scale recurrent Neural Network Architecture

种尺寸,分别输入到对应尺度的网络中。每个尺度的输出图像经过上采样后,直接与下一个尺度的模糊图像拼接作为输入。每个尺度下使用三个编码网络块(Encode Block)提取特征和三个解码网络块(Decode Block)恢复图像,其中,建立了两个跳跃连接,避免训练过程中的梯度消失,便于梯度的反向传播,加快训练过程,每个尺度的去模糊过程可以用公式(2)表示:

$$I_{ge}^i, h^i = G(I_b^i, I_{ge}^{i+1}, h^{i+1}) \quad (2)$$

式中: $i$ 为尺度编号,大值对应小尺度; $I_{ge}^i$ 和 $h^i$ 分别为第 $i$ 尺度下的生成图像和循环网络隐藏层输出; $I_b^i$ 为当前尺度输入的模糊图像; $G(\cdot)$ 为单尺度网络的函数表达式。在循环网络的结构中,三个尺度被视为三个时序,通过颈部的长短期记忆(Long-short Term Memory, LSTM)模块连接。大尺度网络中的特征选择性地包含了小尺度网络的特征信息,形成了一种具有记忆能力的网络结构。

### 1.2 编、解码网络块

每个尺度的编码模块包含了一个卷积层、一个线性整流函数(Rectified Linear Unit, ReLU)和一个密集型残差注意力网络块(Residual Dense Attention Block, RDA Block),如图2所示。解码模块与编码模块结构类似,最末端的解码模块中取消了 ReLU 层。每个编、解码模块的尺寸参数如图1所示。非特殊标注,网络中所有卷积层使用的卷积核均为 $5 \times 5$ ,使其在不

显著增加网络参数的前提下尽可能扩大感受野。

编、解码网络块中使用深度可分离卷积<sup>[14]</sup>替代常规卷积,先对每个通道分别执行卷积,再通过逐点卷积将其混合输出。若用 $a$ 表示卷积核边长,单层卷积操作的参数量将降低为常规卷积的约 $1/a^2$ 。深度可分离卷积层更加侧重于提取空间特征。在运动模糊去除任务中,图像空间特征之间是高度相关的,而对于通道特征而言则相对独立,因此,使用深度可分离卷积将减少参数数量,降低计算量,同时避免了特征的大量丢失。

在图像去模糊任务中,一张图像的不同区域可能存在不同程度的模糊,在网络中嵌入注意力机制能有效地将计算资源更多地倾向于模糊区域,减少在其他区域的消耗。文中在每个编、解码模块中嵌入了改进的卷积注意力模块(Convolutional Block Attention Module, CBAM)<sup>[15]</sup>。CBAM由通道注意力模块(Channel Attention Module)和空间注意力模块(Spatial Attention Module)两个子模块拼接而成。通道注意力旨在建模出不同通道之间的相关性,通过网络学习自动获取每个通道的重要程度并分别赋予各自的权重,从而强化重要的特征。空间注意力旨在提升关键区域的特征表达,将原始图片中的空间信息通过空间转换模块,变换到另一个空间中并保留关键信息,为每个位置生成权重掩膜并加权输出,从而增强感兴趣的特定目标区域,同时弱化不相关的背景区域。通道注意力将一个通道内的信息直接进行全局处理,容易忽略空间上的信息交互。空间注意力模块将每个通道内的特征做同等处理,忽略了通道间的信息交互。

参考文献[15]中提出的注意力子模块如图3(a)、(b)所示。通道注意力子模块将输入特征沿 $W$ 、 $H$ 方向上分别通过最大池化(Max Pooling)和平均池化(Average Pooling)进行尺度压缩,共享一个多层感知机(Multi-layer perceptron, MLP)后相加,经过 sigmoid 激活函数后输出一个 $1 \times 1 \times C$ 的向量,即代表了每个通道对应的权值。空间注意力子模块沿通道方向对输入特征分别执行最大池化和平均池化,并在通道维度拼接特征,通过一个 $7 \times 7$ 卷积操作和 sigmoid 激活函数输出一个 $W \times H \times 1$ 的二维空间权值矩阵。

参考文献[15]中提出了3种拼接方式:两子模块先后串联和并联,如图4所示。在实验中发现其对于

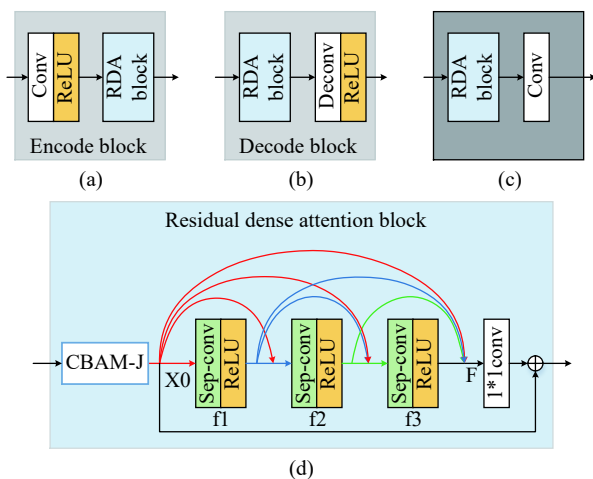


图2 编、解码网络块。(a) 编码块;(b) 解码块;(c) 末端解码块;(d) 密集型残差注意力模块

Fig.2 Encode, decode block. (a) Encode block; (b) Decode block; (c) End decode block; (d) Residual dense attention block

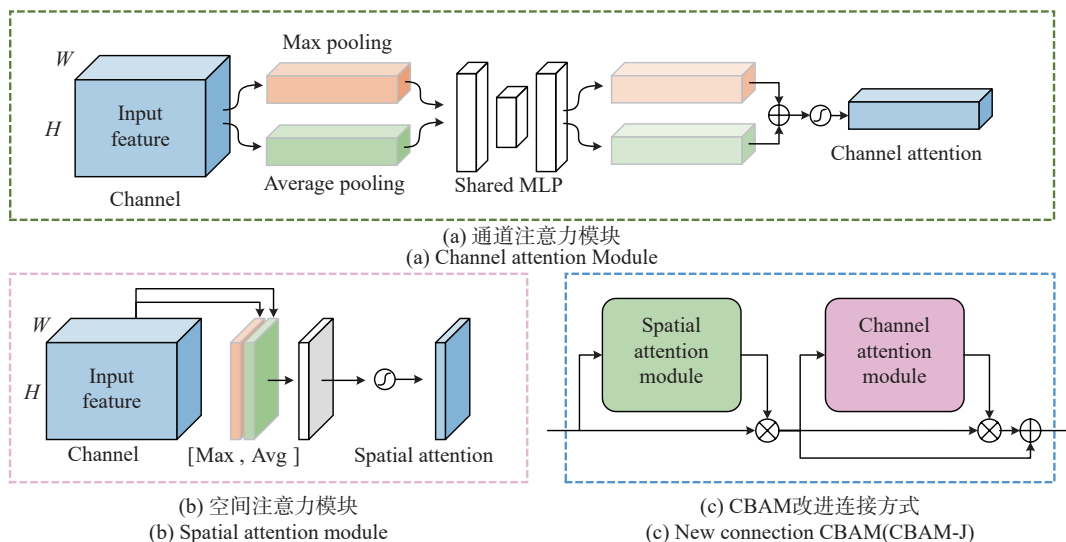


图 3 (a)、(b) CBAM 子模块<sup>[15]</sup>; (c) 文中改进的 CBAM (CBAM-J)

Fig.3 (a), (b) Convolutional block attention submodule<sup>[15]</sup>; (c) Improved CBAM (CBAM-J)

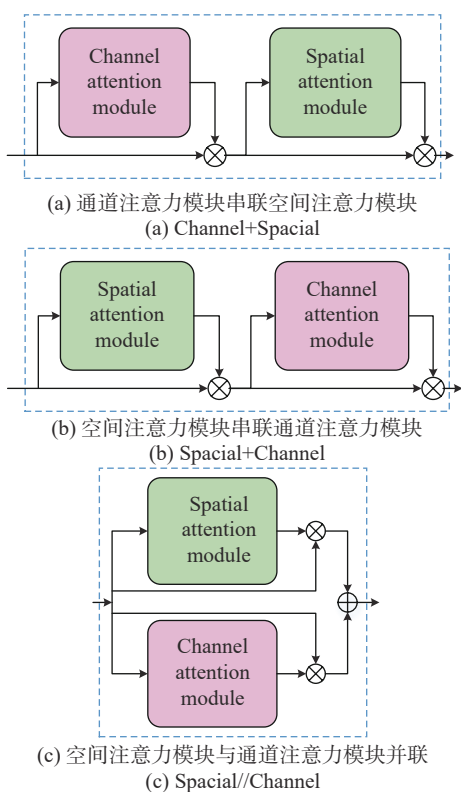


图 4 参考文献 [15] 中 CBAM 的连接方式

Fig.4 CBAM connection modes in literature<sup>[15]</sup>

去模糊模型性能的提升非常有限,甚至可能抑制模型自身性能。根据图像降质公式 (1), 图像运动模糊的产生是空间上的像素点与卷积核进行卷积后产生的结果,而在不同的 RGB 通道之间并未产生相互影响,所以,原始图像的空间特征相比于通道特征会更加重

要。因此,在每个注意力模块中将原始特征优先输入空间注意力子模块中,并将空间注意力子模块的输出跳跃连接至通道注意力模块后以增强其表达。改进的 CBAM 模块 (用 CBAM-J 表示) 如图 3(c) 所示,在空间子模块和通道子模块串联的基础上将空间子模块输出直接跳跃连接至注意力模块整体的末端。经过实验, CBAM-J 注意力模块能在图像去模糊任务中取得最大的提升效果。

表 1 所示为参考文献 [15] 中 3 种子模块连接方式和文中提出的 CBAM-J。 $x$  为输入特征,  $y$  为整体输出,  $S(\cdot)$  和  $C(\cdot)$  分别为空间注意力子模块和通道注意力子模块的函数表达式。可以发现: CBAM-J 是在“Spatial+Channel”的基础上增加了一个跳跃连接  $x \cdot S(x)$ , 反映了加强空间注意力表达的意图。考虑到在特征隐藏层中通道和空间的关系更加复杂, 不适合将空间和通道注意力完全分离, 所以与“Spatial//Channel”相比, CBAM-J 中通道注意力子模块的输入为  $x \cdot S(x)$ , 未直接使用原始特征  $x$ 。

表 1 CBAM 连接方式

Tab.1 Connection modes of CBAM

CBAM connections	Output (y)
<b>Proposed(CBAM-J)</b>	$x \cdot S(x) \cdot (1 + C(x \cdot S(x)))$
Channel+Spacial	$x \cdot C(x) \cdot S(x \cdot C(x))$
Spacial+Channel	$x \cdot S(x) \cdot C(x \cdot S(x))$
Spacial // Channel	$x \cdot S(x) + x \cdot C(x)$

此文对参考文献 [16] 提出的残差块进行了改进: 包含了 1 个注意力模块、3 个密集型连接的卷积层和 1 个  $1 \times 1$  的卷积层, 如图 2(d) 所示。不同于通过简单的加深网络层数或拓展单层网络的宽度, 密集型残差注意力网络通过特征重用来提高参数的利用率。每个密集连接的卷积层的输入特征都直接由前几层卷积的输出直接拼接而成。如果用  $x$ 、 $y$  表示输入输出, 密集型残差注意力模块可以表示为:

$$x_0 = Att(x) \quad (3)$$

$$y = x_0 + conv1\_1(F) \quad (4)$$

式中:  $Att(\cdot)$  表示 CBAM-J;  $conv1\_1(\cdot)$  为最后一个  $1 \times 1$  卷积层, 用于密集型特征的降维;  $F$  是密集型残差块的输出, 若使用  $\oplus$  表示特征张量的拼接,  $F$  可以表示为:

$$F = x_0 \oplus f_1(x_0) \oplus f_2(x_0 \oplus f_1(x_0)) \oplus f_3(x_0 \oplus f_1(x_0) \oplus f_2(x_0 \oplus f_1(x_0))) \quad (5)$$

在一定的范围内, 密集型连接卷积层的数量越多, 特征提取的效果往往越好, 但过多的密集型连接会导致残差块中特征通道数和网络参数数量快速上升, 网络训练将消耗更大的算力, 同时整个网络的运行速度将大幅降低。为了平衡网络能力和运行速度, 在密集型残差块中使用 3 个密集型连接的卷积层。

在每一个编、解码模块中使用密集型残差块可以最大程度地利用同一维度下每一个卷积层的特征信息, 缓解了梯度消失, 增强了特征的传播, 相对于 3 个卷积层的直接连接, 在不显著增加参数数量的情况下, 提高了特征的利用效率, 提升了网络的性能。

### 1.3 损失函数

文中方法使用的损失函数表达式如下所示:

$$L = L_c + \alpha L_p \quad (6)$$

式中:  $L_c$  为内容损失;  $L_p$  为感知损失;  $\alpha$  为常数系数。经过实验测试,  $\alpha$  取值为 0.05 时, 整体损失函数对于训练的优化效果最佳。

内容损失  $L_c$  对网络每一个尺度的清晰图像和生成图像计算像素平均欧氏距离并进行多尺度间的累加, 在像素层面的相似程度上对于网络进行优化, 表达式如下:

$$L_c = \sum_{i=1}^N \sum_{x=1}^{W_i} \sum_{y=1}^{H_i} \frac{1}{W_i H_i} \left( (I_{gt}^i)_{x,y} - (I_{ge}^i)_{x,y} \right)^2 \quad (7)$$

式中:  $N$  为尺度总数;  $W_i$ 、 $H_i$  为第  $i$  尺度的维度;  $I_{gt}^i$  和

$I_{ge}^i$  分别为第  $i$  尺度下的真实清晰图像和生成图像。文中尝试将图像对输入 ImageNet 数据集预训练好的 VGG-19 网络, 取第七层的输出特征图计算欧式距离作为内容损失的补充项, 实验发现对于网络的训练基本无提升, 因此, 仅使用公式 (7) 用于计算内容损失。

图像去模糊任务的一个定性指标为生成的图像具有清晰的边缘。公式 (7) 表示的内容损失不会对图像边缘锐化进行特别处理。借鉴 Qi 等人 [13] 的思想, 文中在尺度  $256 \times 256$  下, 将生成图像和真实清晰图像转化为灰度图, 再利用 Sobel 算子分别计算两张图像水平和垂直方向上的导数, 构成的损失函数表达式如下:

$$L_p = \frac{1}{WH} \sum_{x=1}^W \sum_{y=1}^H (|\nabla_x(I_{gt})_{x,y} - \nabla_x(I_{ge})_{x,y}| + |\nabla_y(I_{gt})_{x,y} - \nabla_y(I_{ge})_{x,y}|) \quad (8)$$

式中:  $W$  和  $H$  为尺度  $256 \times 256$  的图像维度;  $\nabla_x$ 、 $\nabla_y$  为水平和垂直方向上的图像梯度算子。使用内容损失和边缘感知损失的线性组合作为文中网络的损失函数可以在优化图像像素相似度的基础上, 进一步生成清晰的边缘, 提升图像模糊去除的效果。

## 2 实验验证

### 2.1 数据准备

在图像运动模糊去除领域最常用的 3 个数据集是: GoPro 数据集 [8]、Lai 数据集 [17] 和 Köhler 数据集 [18]。GoPro 数据集包含了 3214 对分辨率为  $1280 \times 720$  的“模糊——清晰”图像。它对采集帧率为 240FPS 的一系列连续图像每 7~13 帧进行叠加平均, 并与中间帧对应形成图像对。GoPro 数据集包含一个训练集和一个测试集, 分别有 2103 对和 1111 对图像, 测试集和训练集的场景不重复、不重叠。Lai 数据集有 100 张图片记录了真实的运动模糊, 可以用于衡量去模糊算法在真实图像上的作用效果。Lai 等人用 4 种模糊核对于 25 张真实清晰图片进行处理, 生成 100 张模糊图像。Köhler 数据集在六自由度下记录相机的轨迹构建 12 个模糊核, 分别施加在 4 张清晰图像上, 是评价图像盲去模糊的基准数据集。

由于 GoPro 数据集包含丰富的场景, 且较为真实地模拟出模糊的产生过程, 其被广泛应用于去模糊网络的“端到端”训练和测试中。文中采用 GoPro 训练集进行训练, 使用 GoPro 测试集、Lai 数据集和 Köhler

数据集进行测试和指标评价。

### 2.2 参数设置

文中网络是在运行环境为 Ubuntu 18.04 操作系统和 TensorFlow 框架上搭建的。实验所用台式机使用了 1 块 AMD(R)Ryzen5 3600x 6-core 3.80 GHz CPU 和 1 块 NVIDIA GeForce RTX 3070 8G GPU。网络将图像对随机裁剪为 256×256 大小输入网络, 训练中使用 Adam 优化器, 其超参数设定值  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ , 初始学习率为  $1 \times 10^{-4}$ , 线性下降至 0, Batch size 取为 4, 共训练 2000 个 epoch 大约 192 小时至收敛。

### 2.3 评价指标

对于网络生成的去模糊图像, 需要将其与对应的清晰图像进行对比, 定量评价模糊去除方法的效果。结构相似性 (Structural Similarity, SSIM) 是基于两张图像的亮度、对比度和结构衡量。两张图像相似程度越高, SSIM 值越接近 1。峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR) 是一个与两张图像像素的均方误差有关的评价指标, 其值越大表示在像素层面上两张图片越相似。若给定  $W \times H$  的清晰图像  $I_{gt}$  和去模糊图

像  $I_{ge}$ ,  $MAX$  为可能的最大像素值 (8 bit 图像为 255), PSNR 可以表示为:

$$MSE = \frac{1}{WH} \sum_{i=1}^W \sum_{j=1}^H (I_{gt}(i, j) - I_{ge}(i, j))^2 \quad (9)$$

$$PSNR = 10 \cdot \log_{10} \left( \frac{MAX^2}{MSE} \right) \quad (10)$$

文中将使用 SSIM 和 PSNR 衡量测试集生成的去模糊图像与对应清晰图像的相似程度, 进而定量评价网络的去模糊效果。

### 2.4 实验与分析

为了验证文中方法的有效性和鲁棒性, 文中以网络在 GoPro 测试集、Lai 数据集和 Köhler 数据集的平均 SSIM 和平均 PSNR 作为定量评价指标, 与模糊去除的典型方法 DeblurGAN<sup>[9]</sup>、DeblurGAN-v2<sup>[11]</sup> 和 SRN deblur net<sup>[10]</sup> 进行定量比较。其中 DeblurGAN-v2 使用了其官方提供的两种主干网络: Inception 和 MobileNet, 分别代表其最佳的去模糊效果和最快的网络运行时间, 实验结果如表 2 所示。

表 2 三个数据集的模糊去除评价结果

Tab.2 Deblurring evaluation results on three datasets

Method	GoPro		Lai		Köhler	
	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB
<b>Proposed method</b>	0.9185	29.0284	<b>0.6674</b>	<b>16.5491</b>	<b>0.7625</b>	<b>19.9943</b>
DeblurGAN	0.8474	25.0200	0.6425	15.8905	0.7447	19.7570
DeblurGAN-v2(Inception)	0.9141	28.2701	0.6514	16.1121	0.7469	19.4994
DeblurGAN-v2(MobileNet)	0.8731	25.9644	0.6598	16.4073	0.7556	19.7882
SRN deblur net	<b>0.9331</b>	<b>30.1513</b>	0.6494	16.1000	0.7505	19.5238

从表中数据可知, 文中方法在 Lai 数据集和 Köhler 数据集上取得了最佳的 SSIM 和 PSNR, 较原网络 SRN deblur net 分别提升约 0.0180、0.4491 dB 和 0.012、0.4705 dB, 在 GoPro 测试集上接近 SRN deblur net, 但仍然有 0.0146 和 1.1229 dB 的差距。网络训练使用 GoPro 训练集, 而文中方法在其他数据集上取得更佳结果, 证明文中方法具有良好的鲁棒性和泛化性能。图 4 中, (a) 为输入的模糊图像, (b)~(f) 为各方法对 GoPro 测试集中的图像进行模糊去除的结果。基于 GAN 的 DeblurGAN 对于模糊图像的复原效果十分有限, 其改进的 DeblurGAN-v2 在使用 Mobile-

Net 作为主干网络时图像存在较大的失真, 图像边缘不能很好的恢复。在使用 Inception 作为主干网络时, DeblurGAN-v2 能取得较好的去模糊效果, 但仍然存在边缘粘连的情况。从直观角度来看, 文中方法与原网络 SRN deblur net 对模糊车牌的复原效果较接近, 都在一定程度上消除了竖直方向上的模糊, 恢复出了较为清晰的边缘。

为了测试文中方法在真实运动模糊场景下的泛化性能, 文中用 Lai 数据集集中的 100 张真实模糊进行了测试, 并与几种典型方法进行了对比。图 5 中, (a) 为输入的模糊图像, (b)~(f) 代表各典型方法的去模

糊效果。如图 5 所示,在真实模糊数据集中,几种典型方法都有一定的去模糊效果,但存在振铃现象

(图 5(c))、边缘模糊(图 5(e))等问题。如图 6 所示,文中方法生成的图像较其他方法具有较为清晰的边缘,



图 5 Lai 测试集真实模糊结果

Fig.5 Test result on Lai real blur dataset

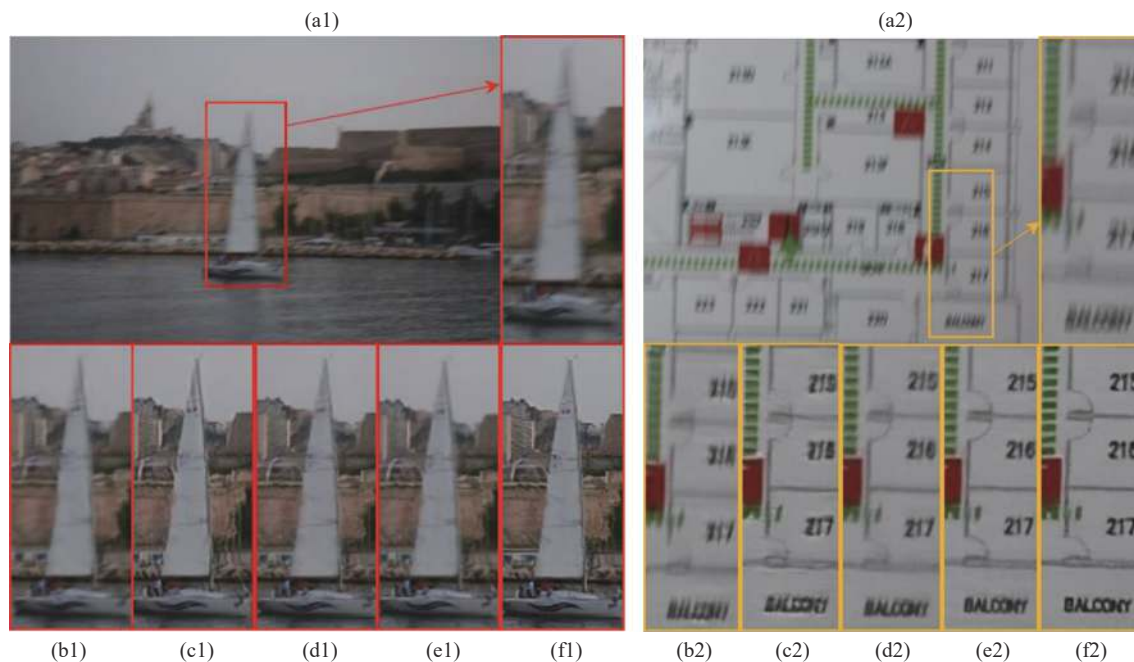


图 6 GoPro 测试集结果

Fig.6 Test result on GoPro testing set

验证了边缘损失函数对于网络性能的优化作用。

图像运动模糊去除方法不仅要求其有良好的复原效果,而且需要算法具有较好的实时性。文中在 GoPro 测试集上对运动模糊去除算法的复杂度、模型文件大小和平均单帧运行时间进行了测算,结果如表 3 所示。为了控制变量,其他典型方法的运行设备均与文中方法一致。

表 3 参数量和单帧运行时间

Tab.3 Number of parameters & execution time per frame

Method	FLOPs/G	Size/MB	Time/s
Proposed method	261.19	<b>12.3</b>	0.206
DeblurGAN	678.29	45.6	0.694
DeblurGAN-v2(Inception)	411.34	244.7	0.212
DeblurGAN-v2(MobileNet)	<b>43.75</b>	13.6	<b>0.068</b>
SRN deblur net	1434.82	78.7	0.501

由于将常规卷积层替换为深度可分离卷积,文中网络的参数数量较原网络大幅下降。与其他方法比较,文中的模型文件最小,由于 3 个尺度共享参数,浮点运算数 (Floating point operations, FLOPs) 和单帧运行时间较运行最快的 DebluGAN-v2 (MobileNet) 略多。文中方法单帧运行时间相比于同类方法 SRN deblur net 缩减了近 60%。结合表 1 中的去模糊评价指标,由于在不同尺度的网络中进行参数共享和强调特征的重用,加之使用注意力模块进行计算资源的合理分配,实验结果证明:文中方法使用最少的参数,取得了几种典型方法中的最佳去模糊效果,同时,运行时间相比同类方法缩减明显。

### 3 结 论

针对图像运动模糊去除效率和效果不能很好统一的问题,文中提出了一种基于多尺度循环注意力网络的图像运动模糊去除方法。文中方法在每个尺度中使用编解码网络结构,LSTM 在颈部进行跨尺度连接,强调多尺度间的信息交流。文中对 CBAM 模块的拼接方式进行了重新设计,在图像上的不同区域更合理地分配计算资源。为了提高参数使用效率,引入卷积层的密集型连接方式。加入边缘损失函数作为

先验指导,有效提高了生成图像的边缘清晰度。实验表明,文中图像去模糊方法的定量指标 SSIM 和 PSNR 较近年典型方法的最佳效果分别提升了 1.15%、0.86% 和 0.91%、1.04%,运行速度比同类方法提高了约 2.5 倍,且文中方法具有更佳的鲁棒性和泛化性能。

### 参考文献:

- [1] Li Anti, Wu Dingjie, Li Chenglong. Survey of autonomous collision avoidance algorithm for unmanned aerial vehicle in low altitude [J]. *Electronics Optics & Control*, 2021, 28(8): 1-8. (in Chinese)
- [2] Lu Feng, Liu Huahai, Huang Changying, et al. Overview of target detection techniques based on deep learning [J]. *Computer Systems Applications*, 2021, 30(3): 1-13. (in Chinese)
- [3] Gao Wen, Zhu Ming, He Baigen, et al. Overview of target tracking technology [J]. *Chinese Optics*, 2013, 7(3): 365-375. (in Chinese)
- [4] Krishnan D, Tay T, Fergus R, et al. Blind deconvolution using a normalized sparsity measure[C]//Proceedings of 2011 IEEE International Conference on Computer Vision and Pattern Recognition, 2011: 233-240.
- [5] Pan J, Sun D, Pfister H, et al. Blind image deblurring using dark channel prior[C]//Proceedings of 2016 IEEE International Conference on Computer Vision and Pattern Recognition, 2016: 1628-1636.
- [6] Whyte O, Sivic J, Zisserman A, et al. Non-uniform deblurring for shaken images [J]. *International Journal of Computer Vision*, 2012, 98(2): 168-186.
- [7] Zhou Luoyu, Zhang Bao, Yang Yang. Estimation of parameter of defocused blurred image using Hough transform [J]. *Infrared and Laser Engineering*, 2012, 41(10): 2833-2837. (in Chinese)
- [8] Nah S, Kim T H, Lee K M. Deep multi-scale convolutional neural network for dynamic scene deblurring [J]. *IEEE Computer Vision and Pattern Recognition*, 2017, 35(1): 257-265.
- [9] Kupyn O, Budzan V, Mykhailych M, et al. DeblurGAN: blind motion deblurring using conditional adversarial networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 8183-8192.
- [10] Tao X, Gao H, Wang Y, et al. Scale-recurrent network for deep image deblurring[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018.
- [11] Kupyn O, Martyniuk T, Wu J, et al. DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better[C]//2019 IEEE/CVF



- International Conference on Computer Vision (ICCV). IEEE, 2019.
- [12] Zhang K, Luo W, Zhong Y, et al. Deblurring by realistic blurring[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020.
- [13] Qi Q, Guo J, Jin W. Attention network for non-uniform deblurring [J]. *IEEE Access*, 2020, 8: 100044-100057.
- [14] Howard A G, Zhu M, Chen B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications[J]. *Computer Vision and Pattern Recognition*, 2017: 10.48550/arXiv.1704.04861.
- [15] Woo S, Park J, Lee J Y, et al. CBAM:Convolutional block attention module[C]//European Conference on Computer Vision, 2018: arXiv:1807.06521.
- [16] Qian P, Wu Y, Zhang X. Dense connected residual generative adversarial network for single image deblurring[C]//2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2021: 461-466.
- [17] Lai W S, Huang J B, Hu Z, et al. A comparative study for single image blind deblurring [C]//Computer Vision and Pattern Recognition, IEEE, 2016: 1701-1709.
- [18] Köhler R, Hirsch M, Mohler B, et al. Recording and playback of camera shake: benchmarking blind deconvolution with a real-world database [C]//Conference on Computer Vision-ECCV, 2012: 27-40.