



中国科技期刊卓越行动计划项目入选期刊

# 控制与决策

CONTROL AND DECISION



## 面向动态场景去模糊的对偶学习生成对抗网络

纪野, 戴亚平, 廣田薰, 邵帅

引用本文:

纪野, 戴亚平, 田薰, 邵帅. 面向动态场景去模糊的对偶学习生成对抗网络[J]. 控制与决策, 2024, 39(4): 1305–1314.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2022.1454>

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### [基于条件对抗生成孪生网络的目标跟踪](#)

Conditional generative adversarial siamese networks for object tracking

控制与决策. 2021, 36(5): 1110–1118 <https://doi.org/10.13195/j.kzyjc.2019.1215>

#### [基于自注意力生成对抗网络的图像超分辨率重建](#)

Image super-resolution reconstruction based on self-attention GAN

控制与决策. 2021, 36(6): 1324–1332 <https://doi.org/10.13195/j.kzyjc.2019.1290>

#### [基于生成对抗网络学习被遮挡特征的目标检测方法](#)

Object detection via learning occluded features based on generative adversarial networks

控制与决策. 2021, 36(5): 1199–1205 <https://doi.org/10.13195/j.kzyjc.2019.1319>

#### [一种基于多层语义特征的图像理解方法](#)

An image understanding method based on multi-level semantic features

控制与决策. 2021, 36(12): 2881–2890 <https://doi.org/10.13195/j.kzyjc.2020.0927>

#### [自适应直觉模糊相异直方图裁剪的图像增强算法](#)

Adaptive intuitionistic fuzzy dissimilar histogram clipping image enhancement algorithm

控制与决策. 2021, 36(12): 2919–2928 <https://doi.org/10.13195/j.kzyjc.2020.0845>

# 面向动态场景去模糊的对偶学习生成对抗网络

纪野<sup>1,2</sup>, 戴亚平<sup>1,2†</sup>, 廣田薰<sup>1,2</sup>, 邵帅<sup>1,2</sup>

(1. 北京理工大学自动化学院, 北京 100081; 2. 北京理工大学  
复杂系统智能控制与决策国家重点实验室, 北京 100081)

**摘要:** 针对动态场景下的图像去模糊问题, 提出一种对偶学习生成对抗网络 (dual learning generative adversarial network, DLGAN), 该网络可以在对偶学习的训练模式下使用非成对的模糊图像和清晰图像进行图像去模糊计算, 不再要求训练图像集合必须由模糊图像与其对应的清晰图像成对组合而成. DLGAN 利用去模糊任务与重模糊任务之间的对偶性建立反馈信号, 并使用这个信号约束去模糊任务和重模糊任务从两个不同的方向互相学习和更新, 直到收敛. 实验结果表明, 在结构相似度和可视化评估方面, DLGAN 与 9 种使用成对数据集训练的图像去模糊方法相比具有更好的性能.

**关键词:** 动态场景去模糊; 对偶学习; 生成对抗网络; 注意力引导; 特征图损耗函数

中图分类号: TP391 文献标志码: A

DOI: 10.13195/j.kzyjc.2022.1454

引用格式: 纪野, 戴亚平, 廣田薰, 等. 面向动态场景去模糊的对偶学习生成对抗网络 [J]. 控制与决策, 2024, 39(4): 1305-1314.

## Dual learning generative adversarial network for dynamic scene deblurring

Ji Ye<sup>1,2</sup>, Dai Ya-ping<sup>1,2†</sup>, Kaoru HIROTA<sup>1,2</sup>, Shao Shuai<sup>1,2</sup>

(1. School of Automation, Beijing Institute of Technology, Beijing 100081, China; 2. State Key Laboratory of Intelligent Control and Decision of Complex System, Beijing Institute of Technology, Beijing 100081, China)

**Abstract:** For the problem of dynamic scene deblurring, a dual learning generative adversarial network (DLGAN) is proposed in this paper. The network can use unpaired blurry and sharp images to perform image deblurring calculations in the training mode of dual learning, which no longer requires the training image set to be a pair of blurry and their corresponding sharp images. The DLGAN uses the duality between the deblurring task and the reblurring task to establish a feedback signal, and uses this signal to constrain the deblurring task and the reblurring task to learn and update each other from two different directions until convergence. Experimental results show that the DLGAN has a better performance compared to nine image deblurring methods trained with paired datasets in structural similarity and visualization evaluation.

**Keywords:** dynamic scene deblurring; dual learning; generative adversarial network; attention-guided; feature map loss function

## 0 引言

基于深度学习的单幅图像动态场景去模糊方法<sup>[1]</sup>通常是有监督学习方案, 并且明确要求训练数据集必须是成对的 (模糊图像和对应的清晰图像). 然而, 准确获取成对训练数据集是一项艰巨的任务, 即便是用成对的数据集进行了充分的训练, 在针对一幅新的有“动态场景模糊”问题的图片时, 训练好的网

络往往也不具有令人满意的泛化性. 针对一幅“拍摄模糊”的图片, 更实用的解决方案是用一种无监督学习的方法 (仅仅需要一幅模糊图像和任意一幅清晰图像) 进行图像去模糊.

无监督学习只需要非成对的模糊图像与任意清晰图像作为训练数据集, 并且非成对的图像在实际的去模糊学习任务中是容易获得的. Lu 等<sup>[2]</sup>提出一种

收稿日期: 2022-08-12; 录用日期: 2023-01-20.

基金项目: 国铁集团系统性重大项目 (P2021T002); 北京市自然科学基金项目 (L191020).

责任编辑: 胡清华.

†通讯作者. E-mail: daiyaping@bit.edu.cn.

基于分离表征的特定领域无监督图像去模糊方法,该方法对算法合成的模糊类型有效,但无法有效去除真实动态场景模糊;Wen等<sup>[3]</sup>提出一种基于对偶学习的无监督图像去模糊方法,该方法克服了其他方法需要对图像作为训练数据的缺点,并且可以由非成对的另一幅清晰图像自动生成一幅“新”的模糊图像,为后续研究提供更多可用的数据.根据这些学者的研究成果可知,对偶学习是动态场景图像去模糊的合适解决方案之一,值得更加深入的研究.

通常对偶学习的网络结构包含两个学习任务,分别称作初始学习任务和对偶学习任务,其结构的突出特点是将用于初始学习任务的编码器用作对偶学习任务的解码器,而初始学习任务的解码器用作对偶学习任务的编码器,从而实现学习功能上的对偶.对偶学习通过交互学习来产生有效的反馈信号,并使用这个反馈信号约束初始学习任务和对偶学习任务,使得两个学习任务从两个不同的方向互相学习和更新,达到了用非成对数据集完成成对数据集才能完成的图像到图像的转换任务.现有的CycleGAN<sup>[4]</sup>和DRIT++<sup>[5]</sup>为主要代表的对偶学习方法,证明了在图像到图像转换的计算机视觉任务中是有效的.值得注意的是,这些对偶学习方法使用的是非成对训练数据集,并且图像到图像转换任务也是从清晰图像到清晰图像的风格迁移.这些对偶学习方法使用的非成对训练数据集都是领域知识对称的清晰图像,并且清晰图像包含特定领域知识和非特定领域知识两项领域知识.动态场景模糊图像和清晰图像的领域知识是不对称的,假定动态场景模糊图像包含了模糊领域知识和非模糊领域知识两项领域知识,而清晰图像却只有非模糊领域知识一项领域知识.在动态场景模糊图像和清晰图像领域知识不对称的情况下,针对一幅拍摄模糊的图像,如何使用对偶学习来将原模糊图像中的模糊领域知识迁移到另外一幅其他场景下的清晰图像上去,使得去除掉模糊领域知识的原模糊图像转换成一幅清晰图像,这是一个具有挑战性的问题.

基于上面的分析,本文提出一种对偶学习生成对抗网络(dual learning generative adversarial network, DLGAN),该网络主要由注意力引导的生成器、多尺度鉴别器和特征图损耗函数组成.为了克服图像去模糊需要成对图像作为训练数据集的问题,本文设计注意力引导的生成器.当注意力引导的生成器作为解码器时,初始学习任务对应去模糊任务;当注意力引导的生成器作为编码器时,对偶学习任务对应重

模糊任务.去模糊任务和重模糊任务之间以对偶学习模式产生有效的反馈信号,这个信号使用感知一致性损耗函数进行优化,约束去模糊任务和重模糊任务从两个不同的方向互相学习和更新.领域知识不对称的图像数据可以通过生成器中的注意力引导模块和特征转换模块,将模糊领域知识和非模糊领域知识编码留存,实现两个学习任务的参数共享.本文设计的多尺度鉴别器能够使用不同分辨率约束下的多重对抗学习形式,避免虚假信息随着分辨率的增加而增加的问题,逐步提高通过注意力引导的生成器获得图像的质量.此外,特征图损耗函数由感知一致性损耗函数和特征感知损耗函数构成,感知一致性损耗函数用来优化两个学习任务中的注意力引导的生成器,能够有效地减少去模糊图像上的伪影特征;特征感知损耗函数用来优化两个学习任务中的多尺度鉴别器,使得网络中不同分辨率的隐藏层受到约束,特征图被迭代优化以产生更高质量的结果.因此,DLGAN的训练过程不需要使用模糊图像与清晰图像成对的训练数据集.本文在GoPro数据集上进行大量实验,实验结果表明,所提DLGAN方法能够使用非成对数据集训练网络,并且与使用了成对数据集训练的图像去模糊方法相比,在定量和定性评估方面精度有明显提升.

## 1 对偶学习生成对抗网络

这一部分将对基于对抗学习体系结构的DLGAN进行具体的理论分析与设计实现.DLGAN体系结构的概述如图1所示.针对动态场景去模糊任务,将生成器、鉴别器和损耗函数的结构进行重构,并在1.1节、1.2节和1.3节中详细介绍.

### 1.1 注意力引导的生成器

注意力引导的生成器 $F$ 和注意力引导的生成器 $G$ 具有相同的网络结构,因此以注意力引导的生成器 $G$ 为例子来阐述.目标是学习给定训练样本 $\{x_i\}_{i=1}^N \in X$ 和 $\{y_j\}_{j=1}^M \in Y$ 的两个域 $X$ 和 $Y$ 之间的映射函数.注意力引导的生成器 $G$ 和注意力引导的生成器 $F$ 分别对应于映射 $G: X \rightarrow \hat{Y}$ 和映射 $F: Y \rightarrow \hat{X}$ .其中:去模糊图像 $\hat{Y}$ 与重模糊图像 $\hat{X}$ 是由注意力引导的生成器生成的图像;模糊图像 $X$ 和清晰图像 $Y$ 是从真实世界获取到的图像.注意力引导的生成器是由特征提取器 $G_E$ 、注意力引导器 $G_A$ 和特征转换器 $G_T$ 三个子网络组成,如图2所示.

下面分别对3个子网络进行详细阐述.

1) 特征提取器 $G_E$ ,目的是从模糊图像中提取底层特征,并将特征参数共享给下游两个子网络.

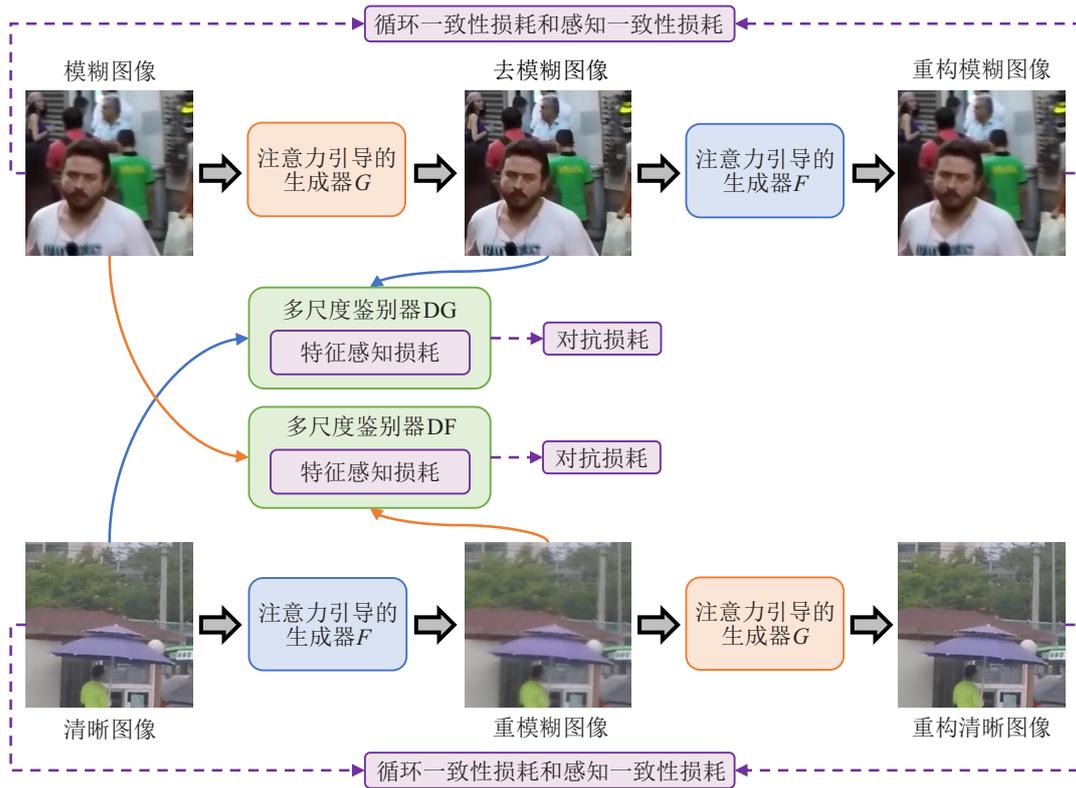


图1 面向动态场景去模糊的DLGAN结构和数据流图

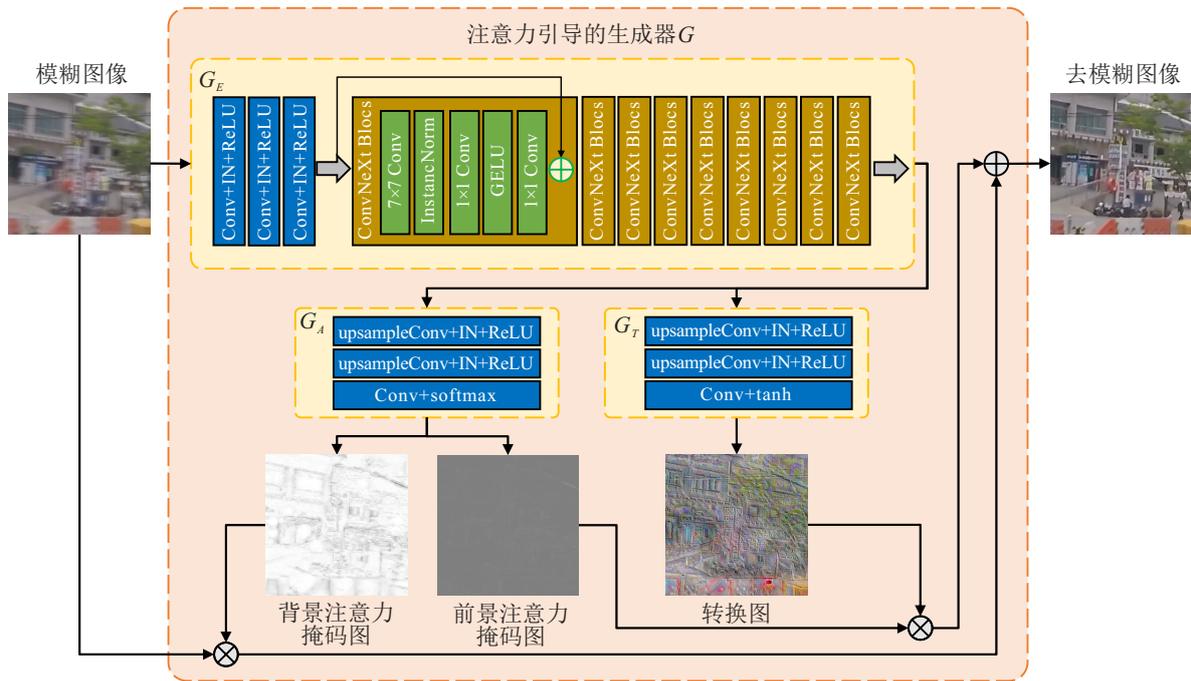


图2 注意力引导的生成器G网络结构设计

特征提取器  $G_E$  包含3个 Conv-InstanceNorm-ReLU 块和9个 ConvNeXt 块, 其中每个 ConvNeXt 块包含1个  $7 \times 7$  Conv 层、1个实例规范化层 (instance normalization, InstanceNorm)、1个高斯误差线性单元 (Gaussian error linear unit, GELU) 非线性变换层和2个  $1 \times 1$  Conv 层。

2) 注意力引导器  $G_A$ , 根据特征提取器  $G_E$  提取

的特征生成前景和背景注意力掩码图, 这使得模糊去除区域集中在动态场景模糊分布附近. 注意力引导器  $G_A$  包含2个 Conv-InstanceNorm-ReLU 块和1个 Conv-Softmax 块, 可以生成注意力掩码图。

3) 特征转换器  $G_T$ , 利用特征提取器  $G_E$  提取的特征生成转换后的中间图像. 特征转换器  $G_T$  包含2个 Conv-InstanceNorm-ReLU 块和1个 Conv-Tanh 块。

经过注意力引导器  $G_A$  获得的前景注意力掩码图与转换后的中间图像做乘积运算, 经过注意力引导器  $G_A$  获得的背景注意力掩码图与输入的模糊图像做乘积运算. 然后将两个乘积运算获得的结果做加法运算, 得到最终的去模糊图像  $\hat{Y}$ . 注意力引导的生成器  $G$  前向传播可用以下公式表示:

$$\hat{Y} = G_T(G_E(x)) \otimes G_A(G_E^F(x)) \oplus x \otimes G_A(G_E^B(x)). \quad (1)$$

其中:  $\otimes$  和  $\oplus$  分别是元素之间的乘法运算符和加法运算符,  $G_E^F(x)$  和  $G_E^B(x)$  分别是前景注意力掩码图和背景注意力掩码图.

为了解释注意力引导器学习到动态场景模糊特征的有效性, 本文将注意力引导器的输出结果进行可视化表示. 一共设置 300 个训练周期, 每隔 50 个周期取出输出的前景注意力掩码图进行可视化表示, 如图 3 和图 4 所示. 由此可见, 注意力引导器的输入可以是

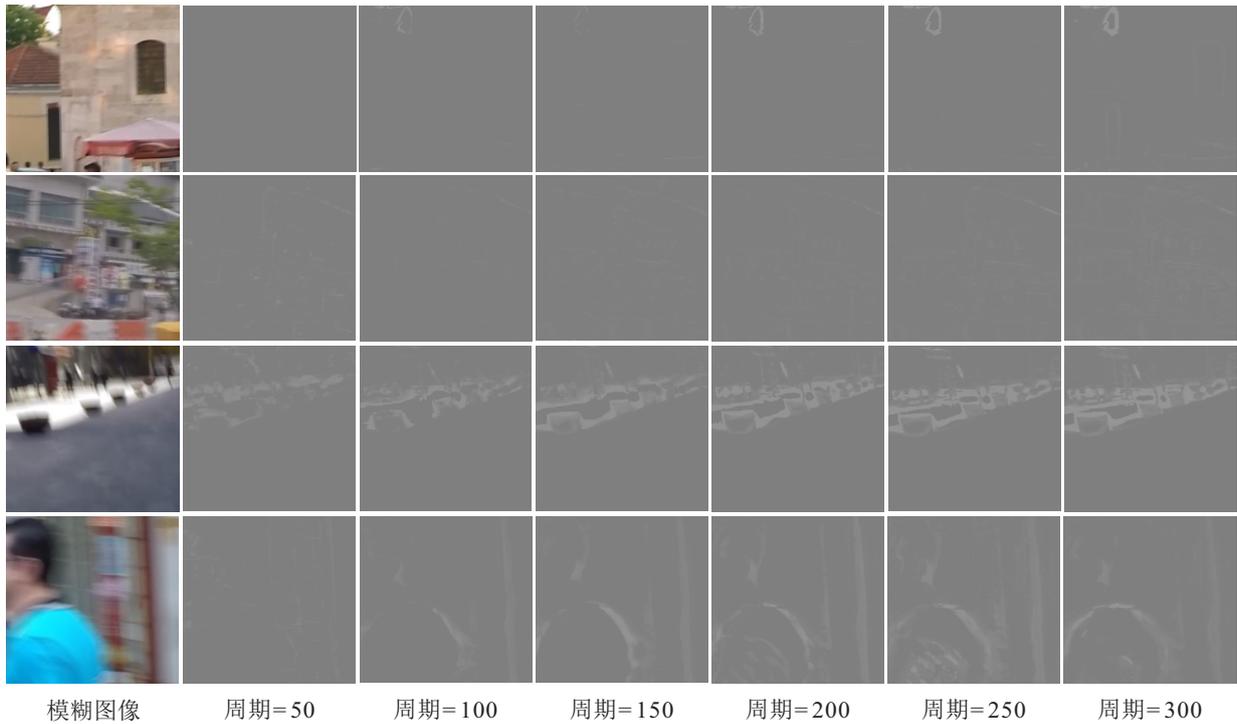


图 3 注意力引导器探测到动态场景模糊图像中的模糊特征掩码

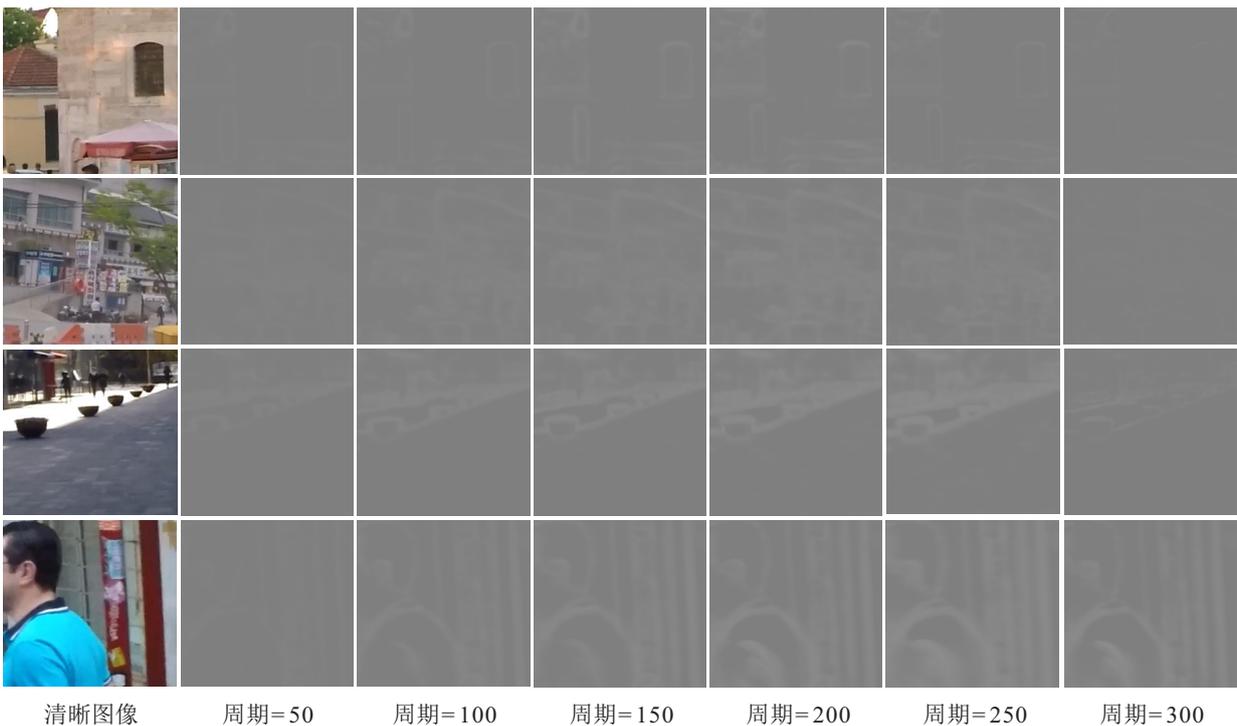


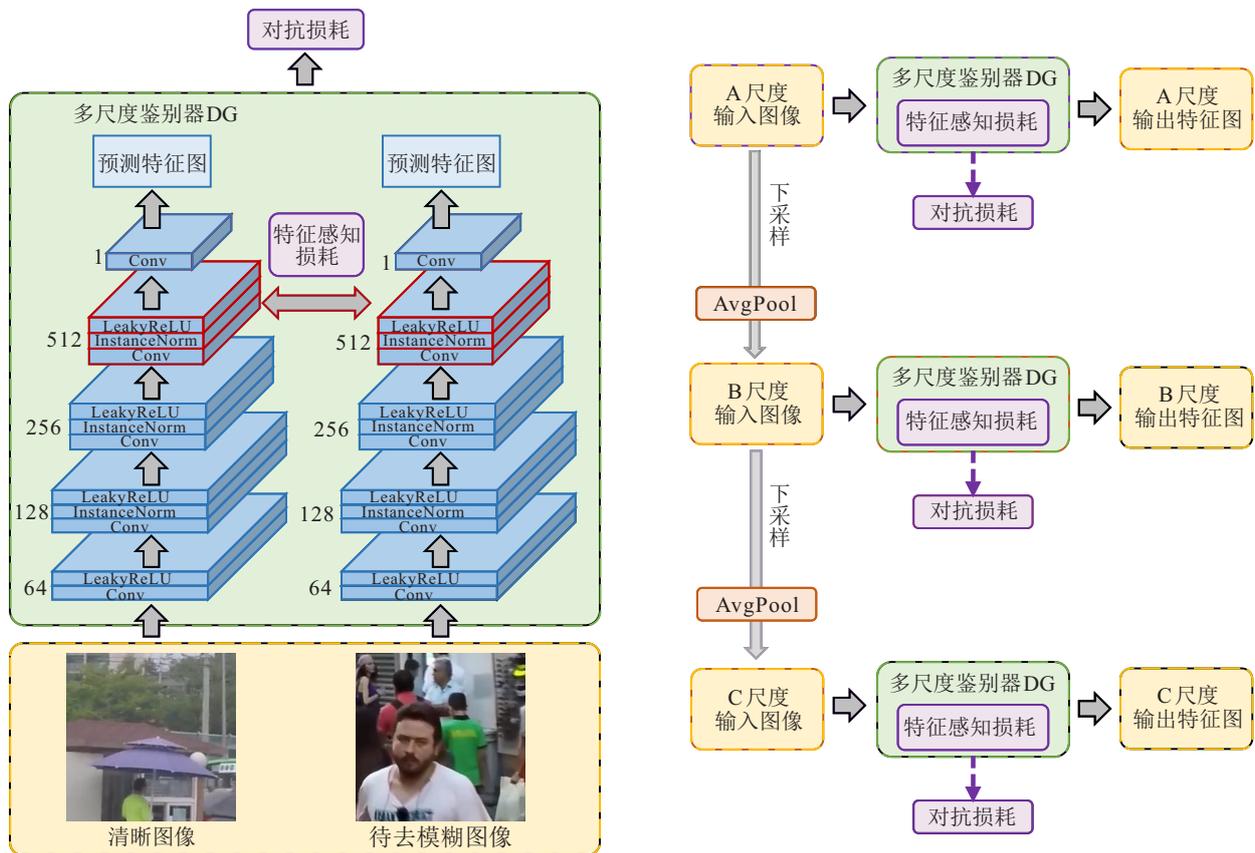
图 4 注意力引导器探测到清晰图像中的非模糊特征掩码

带有或不带有动态场景模糊的任意图像,而输出是根据它对输入图像所探测到的对应特征掩码信息.如图3所示,输入的是动态场景模糊图像,注意力引导器探测到动态场景模糊图像中的模糊特征掩码,从中可以看到随着迭代次数的增加,模糊特征掩码逐渐清晰.如图4所示,输入的是清晰图像,注意力引导器一开始将非模糊特征误认为是动态场景模糊特征,随着学习和引导,逐渐减少对非模糊特征的关注.一般来说,注意力引导器对动态场景模糊特征信息的探测具有鲁棒性,这在已有的模糊特征注意力机制中从未进行过详细的调查与分析.因此,通过所作的大量实验与分析,有效验证了注意力引导器能够从动态场景模糊图像中学习到模糊特征.

### 1.2 多尺度鉴别器

文中设计了两个多尺度鉴别器DG和DF,其中多尺度鉴别器DG是为了鉴别去模糊图像与真实的

清晰图像,多尺度鉴别器DF是为了鉴别重模糊图像与真实的模糊图像.多尺度鉴别器DF的结构与多尺度鉴别器DG的结构类似,因此这部分主要以鉴别器DG为例进行描述.多尺度鉴别器DG的网络结构如图5所示,基本结构设计采用大小为 $70 \times 70$ 的PatchGAN<sup>[6]</sup>结构,因为它能够从局部区域鉴别图像的真实性,比完整图像鉴别器具有更少的参数.但是,通过多次实验发现,仅使用PatchGAN来提高生成图像的质量是不够的.因此,设计了多尺度鉴别器网络结构,包含3个尺度的子鉴别器,3个尺度的输入图像分别是A尺度的原始图像、经过一次下采样得到的B尺度图像和经过两次下采样得到的C尺度图像,并输出对应3个尺度的特征图.其中,下采样是由一个卷积核为3,步长为2的平均池化层构成,每个子鉴别器输出的多尺度特征图再经过平均运算得到最终的输出结果.



(a) 多尺度判别器的网络结构

(b) 多尺度判别器的3个尺度鉴别结构

图5 多尺度鉴别器DG的网络结构设计

### 1.3 损耗函数

本文所提网络的总损耗函数如下:

$$\mathcal{L}_{total} = \mathcal{L}_{adv} + \lambda_{cycle} \mathcal{L}_{cycle} + \lambda_{id} \mathcal{L}_{id} + \mathcal{L}_{fm}, \quad (2)$$

其中 $\lambda_{cycle}$ 和 $\lambda_{id}$ 分别对应于损耗函数 $\mathcal{L}_{cycle}$ 和 $\mathcal{L}_{id}$ 的权重系数.

1) 对抗性损耗函数.以注意力引导的生成器G和多尺度鉴别器DG为例,将一幅运动模糊的图像输入到注意力引导的生成器G中,注意力引导的生成器G将生成去模糊图像,多尺度鉴别器DG区分清晰图像与去模糊图像.对抗性损耗函数的目标是将生成

图像的分布与清晰图像域的分布相匹配. 对抗性损耗函数可以表示如下:

$$\mathcal{L}_{\text{adv}} = \min_G \max_{\text{DG}} \mathbf{E}_y [\log D_y(y)] + \mathbf{E}_x [1 - \log D_y(G(x))]. \quad (3)$$

其中:  $y$  是从真实清晰的图像中采样获得的,  $x$  是从真实动态场景模糊的图像中采样获得的.

2) 循环一致性损耗函数. 由于缺少成对的训练数据集, 仅使用对抗性损耗函数会导致映射空间过大, 因此应用循环一致性损耗函数<sup>[6]</sup>来约束映射空间. 具体而言, 循环一致性损耗函数期望注意力引导的生成器  $G$  生成的  $\hat{Y}$  可以通过注意力引导的生成器  $F$  恢复到输入  $\hat{X}$ , 即  $F(G(x)) \approx x$ , 同理,  $G(F(y)) \approx y$ . 循环一致性损耗函数可表示如下:

$$\mathcal{L}_{\text{cycle}} = \mathbf{E}_x [\|F(G(x)) - x\|_1] + \mathbf{E}_y [\|G(F(y)) - y\|_1]. \quad (4)$$

3) 恒等损耗函数. 为了让注意力引导的生成器更好地学习注意力的映射, 当输入图像为清晰的图像时, 恒等损耗函数<sup>[7]</sup>期望注意力引导的生成器  $G$  不做任何映射, 即  $G(x) \approx x$ . 同理,  $F(y) \approx y$ . 恒等损耗函数可以表示如下:

$$\mathcal{L}_{\text{id}} = \mathbf{E}_x [\|G(x) - x\|_1] + \mathbf{E}_y [\|F(y) - y\|_1]. \quad (5)$$

4) 特征图损耗函数. 为了减少去模糊图像上的伪影特征, 本文使用感知一致性损耗和特征感知损耗构成的特征图损耗函数. 特征图损耗函数如下:

$$\mathcal{L}_{\text{fm}} = \lambda_{\text{pp}} \mathcal{L}_{\text{pp}} + \lambda_{\text{fp}} \mathcal{L}_{\text{fp}}, \quad (6)$$

其中  $\lambda_{\text{pp}}$  和  $\lambda_{\text{fp}}$  分别对应于损耗函数  $\mathcal{L}_{\text{pp}}$  和  $\mathcal{L}_{\text{fp}}$  的权重系数.

① 感知一致性损耗函数. 循环一致性损耗函数通过计算输入图像和重构图像的  $L_1$  损耗来约束映射空间. 然而, 这种计算是在像素空间中进行的, 这使得生成的图像存在部分伪影特征. 因此, 引入感知一致性损耗函数来计算输入图像和重构图像在特征空间的  $L_2$  损耗, 进一步压缩映射空间<sup>[7]</sup>. 感知一致性损耗函数可以表示为

$$\mathcal{L}_{\text{pp}} = \|\phi_l(F(G(x))) - \phi_l(x)\|_2 + \|\phi_l(F(G(y))) - \phi_l(y)\|_2, \quad (7)$$

其中  $\phi_l$  表示 VGG19<sup>[8]</sup> 中的第  $l$  层. 本文的 DLGAN 网络采用 Relu 非线性激活特征, 这种特征来源于预训练 VGG19 网络中的第 4 层和第 17 层.

② 特征感知损耗函数. 由于模糊图像域和清晰

图像域是不对称的, 上述损耗函数不足以恢复非模糊的背景图像. 因此, 使用感知损耗函数<sup>[9]</sup>来度量去模糊图像和清晰图像特征图之间的距离, 以进一步约束特征的映射空间, 因为本文认为生成器生成的去模糊图像应该接近高层特征空间中的清晰图像. 然而, 基于 ImageNet 训练的 VGG 19 网络无法很好地提取复杂的运动模糊特征. 为了解决这一局限性, 用多尺度鉴别器代替 VGG 19 的特征提取网络. 特征感知损耗函数如下:

$$\mathcal{L}_{\text{fp}} = \mathbf{E}_{(x,y)} \sum_{i=l}^t [\|D_y^i(y) - D_y^i(G(x))\|_1], \quad (8)$$

其中  $D_y^i(\cdot)$  的上标  $i$  表示鉴别器中的第  $i$  层, 共提取  $t - l$  个层特征进行计算. 本文使用了多尺度鉴别器 DG 第 8 层到第 11 层的特征图. 同理, 多尺度鉴别器  $D_x^i(\cdot)$  与多尺度鉴别器  $D_y^i(\cdot)$  是类似的.

## 2 实验结果与分析

本节给出了动态场景去模糊实验的数据集准备和图像去模糊实验的实现细节, 并将本文所提 DLGAN 方法的去模糊结果与其他去模糊方法进行定性和定量的比较.

### 2.1 实验数据描述和实验设置

实验是在配备了英伟达 TITAN RTX GPU 的服务器上进行的. 本文提出的 DLGAN 架构是在 PyTorch 深度学习框架的基础上构建的. 本文使用图像运动模糊的通用基准大规模真实数据集 GoPro 进行验证和测试去模糊方法的性能. 训练数据集被随机分配, 以保证网络训练过程中输入的是非成对数据, 而去模糊图像需要与对应的清晰图像进行相似度评估. 训练集中的原始模糊图像在训练期间被随机裁剪成多个  $256 \times 256$  的图像块, 并且训练集经过水平翻转操作, 能够防止由深度神经网络训练引起的过拟合问题. 优化算法采用 ADAM<sup>[10]</sup>, 设置的随机最小批量输入数据为 1. 深度神经网络训练的周期为 300, 超参数设置为  $\lambda_{\text{cycle}} = 10$ ,  $\lambda_{\text{pp}} = 5 \times 10^{-6}$ ,  $\lambda_{\text{id}} = 0.5$  和  $\lambda_{\text{fp}} = 5$ .

### 2.2 对比实验分析

将本文所提出的 DLGAN 方法与 9 种去模糊方法在 GoPro 数据集下进行对比分析. 9 种方法包括: 1 种基于非深度神经网络的方法, 即 Pan 等<sup>[11]</sup>所提出的去模糊方法; 5 种监督学习方法, 即 Nah 等<sup>[12]</sup>所提出的去模糊方法、Gong 等<sup>[13]</sup>所提出的去模糊方法、SRNN<sup>[14]</sup>、DeblurGAN<sup>[15]</sup> 和 DeblurGAN+<sup>[16]</sup>; 3

种非监督学习方法,即UDSDeblur<sup>[2]</sup>、DRIT++<sup>[5]</sup>和CycleGAN<sup>[4]</sup>.

为了验证本文所提DLGAN方法的优势,需要将非成对的模糊图像和清晰图像作为训练数据集,用于训练上述9种方法来作为对比实验. 1种基于非神经网络的方法依赖于对已知模糊核进行线性假设,需要人工构造函数来实现去模糊,所以无法使用非成对训练数据集来自适应学习去模糊任务. 5种监督学习方法和3种非监督学习方法都是基于深度学习的方法,其中5种监督学习方法明确要求训练数据集必须是成对的模糊图像和清晰图像,并且需要使用预训练模型测试去模糊性能,所以监督学习方法使用非成对训练数据集显然是不合适的. 在3种非监督学习方法中,DRIT++<sup>[5]</sup>和CycleGAN<sup>[4]</sup>使用的非成对训练数据集都是领域知识对称的清晰图像,UDSDeblur<sup>[2]</sup>使用的非成对训练数据集是在静态场景下用模糊算法合成的模糊图像. 为了对比实验的有效性,本文需要对3种非监督学习方法在领域知识不对称的非成对训练数据集下自适应学习去模糊任务.

如图6~图9所示,能够对比观察9种去模糊方

法与DLGAN方法的去模糊效果. 基于神经网络的方法得到了比基于非神经网络方法更自然的去模糊图像. 从放大的图像块中可以观察到,由于不能处理空间变化的模糊特征,大多数基于非神经网络的方法无法产生合理的去模糊图像,并且基于非神经网络的方法要么依赖简单的线性假设来有条件地去模糊,要么过高地估计模糊来去模糊. 然而,基于神经网络的方法和本文所提出的DLGAN方法更适合处理动态场景模糊图像的复杂模糊特征,这种复杂的模糊特征是由拍摄相机的快速运动,或者多个拍摄目标的快速运动,甚至是相机与多个拍摄目标快速运动的相互叠加造成的. 上述对比实验可视化结果定性验证了本文所提的DLGAN方法达到了最好的去模糊性能.

如表1所示,基于神经网络的方法比基于非神经网络的方案得到了更高的得分,实现了更好的去模糊性能. 在基于神经网络的方法中,与其他基于CNN、GAN和RNN的方法相比,本文所提DLGAN方法的评估指标得分最高,具有最好的去模糊性能.

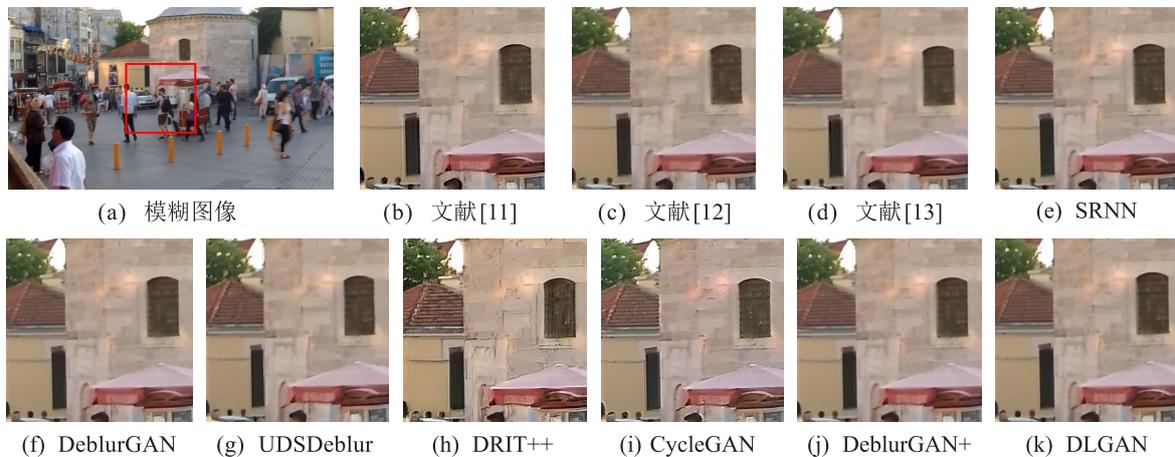


图6 图像去模糊方法对比实验可视化评估(实验1)

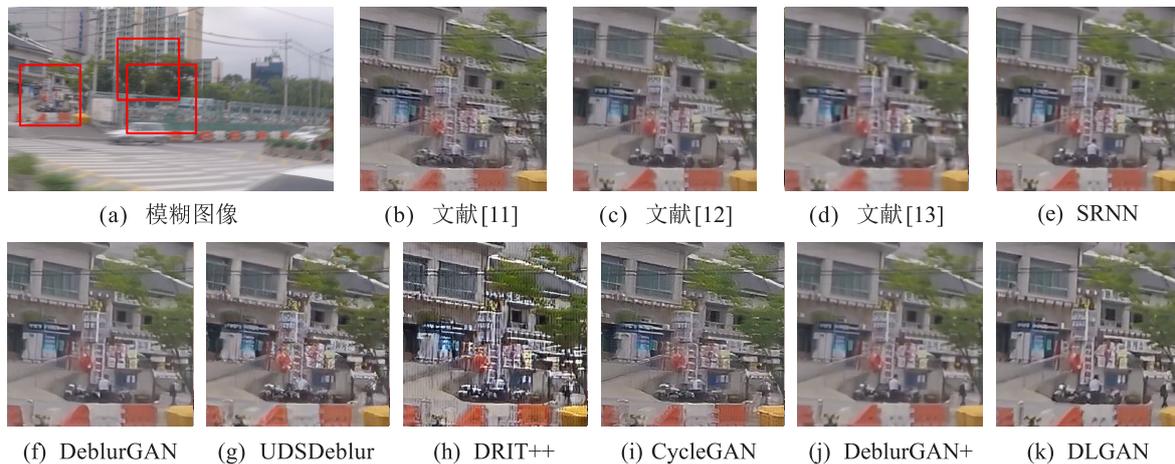


图7 图像去模糊方法对比实验可视化评估(实验2)

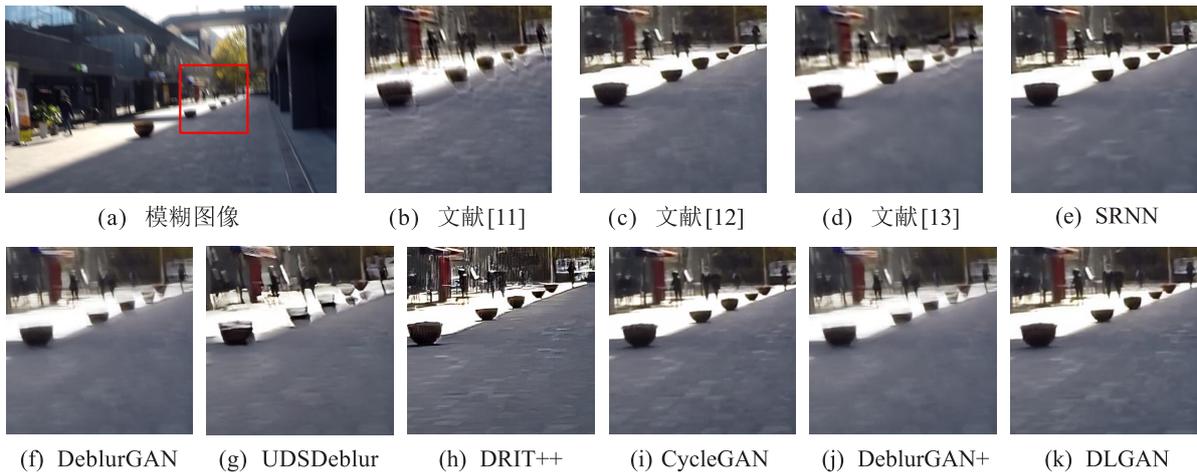


图8 图像去模糊方法对比实验可视化评估(实验3)

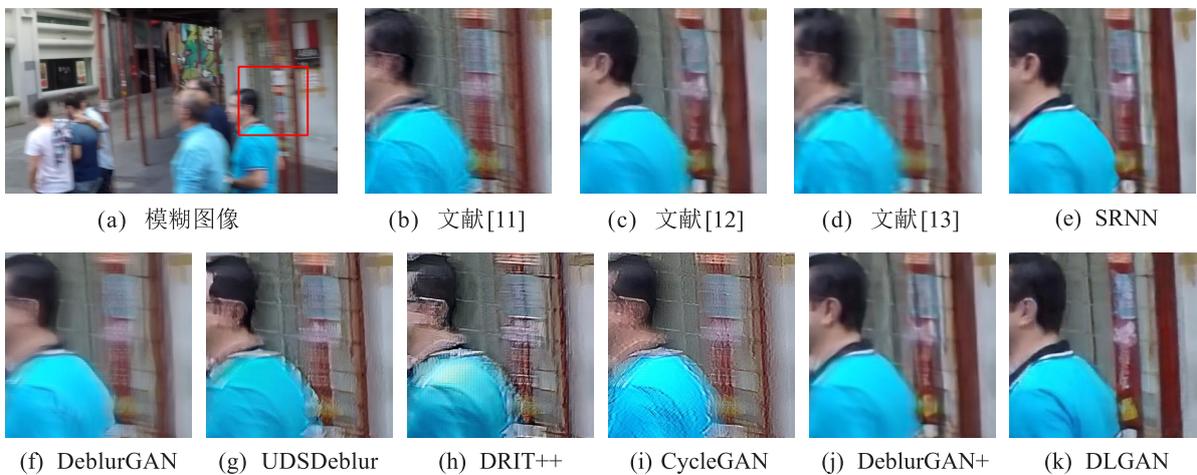


图9 图像去模糊方法对比实验可视化评估(实验4)

表1 图6~图9中对比实验图像相对应的评估指标PSNR、SSIM和FSIM分数比较

方法	图6			图7			图8			图9		
	PSNR	SSIM	FSIM									
文献[11]	25.088	0.787	0.928	24.634	0.762	0.863	24.286	0.716	0.868	24.895	0.816	0.898
文献[12]	28.488	0.829	0.938	28.404	0.823	0.943	28.546	0.852	0.923	28.331	0.841	0.918
文献[13]	28.527	0.861	0.944	26.445	0.843	0.898	26.534	0.798	0.893	26.275	0.848	0.901
SRNN	29.784	0.901	0.943	28.799	0.875	0.901	29.089	0.856	0.928	28.875	0.876	0.902
DeblurGAN	28.057	0.821	0.946	24.817	0.649	0.901	24.686	0.660	0.885	24.832	0.721	0.903
UDSDeblur	27.824	0.863	0.919	25.842	0.843	0.865	26.124	0.787	0.851	24.969	0.834	0.867
DRIT++	18.685	0.471	0.814	19.437	0.479	0.850	15.473	0.580	0.663	19.958	0.575	0.810
CycleGAN	25.478	0.769	0.972	23.900	0.630	0.923	20.919	0.731	0.882	21.772	0.580	0.847
DeblurGAN+	29.019	0.906	0.969	28.612	0.869	0.898	29.451	0.909	0.971	28.918	0.878	0.902
DLGAN	29.698	0.914	0.979	29.044	0.897	0.934	29.526	0.912	0.971	29.310	0.898	0.965

### 2.3 消融实验分析

本节进行一项消融实验研究,以分析本文所提DLGAN方法中感知一致性损耗函数、恒等损耗函数和特征感知损耗函数组件的有效性. 本文将对抗性损耗函数和循环一致性损耗函数作为DLGAN方法的基础损耗函数,而对感知一致性损耗函数、恒等损耗函数和特征感知损耗函数组件的7种组合进行

消融实验分析.

1) DLGAN-v1: 在基础损耗函数上添加恒等损耗函数组件和感知一致性损耗函数组件.

2) DLGAN-v2: 在基础损耗函数上添加感知一致性损耗函数组件和特征感知损耗函数组件.

3) DLGAN-v3: 在基础损耗函数上添加特征感知损耗函数组件和恒等损耗函数组件.

4) DLGAN-v4: 在基础损耗函数上添加感知一致性损耗函数组件.

5) DLGAN-v5: 在基础损耗函数上添加恒等损耗函数组件.

6) DLGAN-v6: 在基础损耗函数上添加特征感知损耗函数组件.

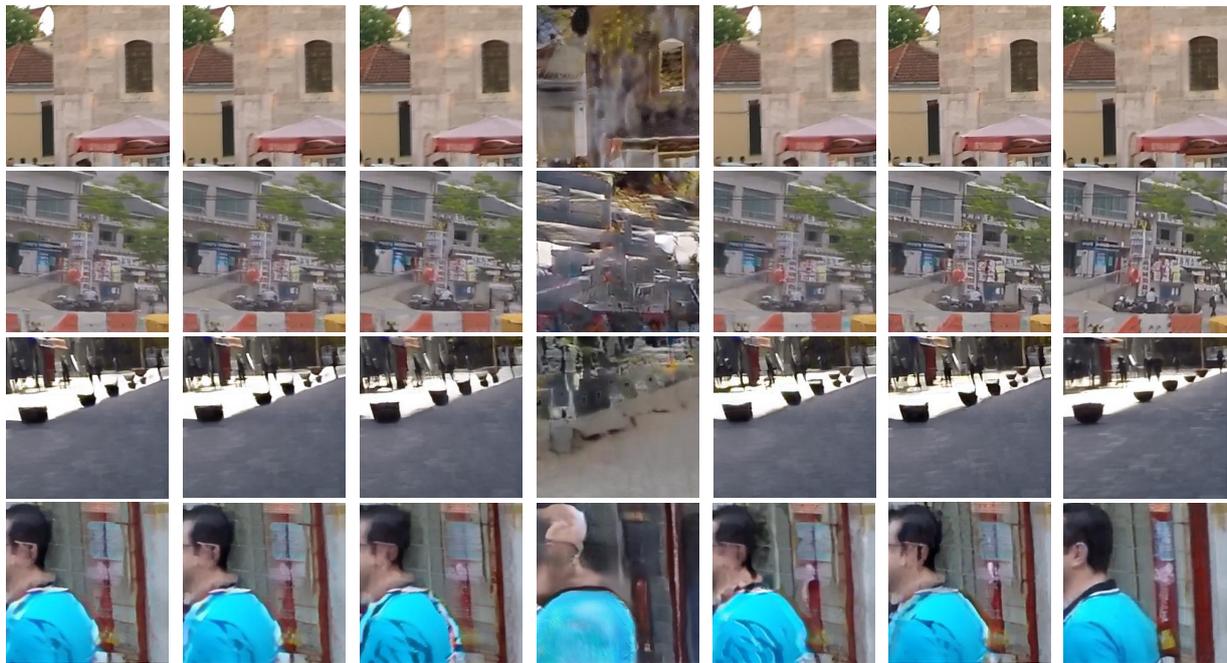
7) DLGAN-v7: 在基础损耗函数上添加感知一致性损耗函数组件、恒等损耗函数组件和特征感知损耗函数组件.

如表 2 和图 10 所示, 最终版本的 DLGAN (DLGAN-v7) 比前 6 个 DLGAN 版本具有更好的去模糊性能. 根据表 2 中的评估指标得分, DLGAN-v1、DLGAN-v2 和 DLGAN-v3 版本的 3 个评估指标得分相近, 说明感知一致性损耗函数、恒等损耗函数和特征感知损耗函数组件的两两组合对网络学习的约束效果是相似的, 无法给 DLGAN 的图像去模糊性能带来进一步的提升. 类似地, DLGAN-v5 和 DLGAN-v6 版本的 3 个评估指标得分相近, 恒等损耗函数和特征感知损耗函数组件的单独添加对网络学习的约束情况是相似的, 无法给 DLGAN 的图像去模糊性能带

来进一步的提升. 通过进一步观察发现, 损耗函数组件的单独添加获得的评估指标得分低于损耗函数组件的两两组合获得的评估指标得分, 说明损耗函数组件的两两组合能够有效提升 DLGAN 的去模糊性能. 3 个评估指标得分的最低值和最高值分别对应于 DLGAN-v4 和 DLGAN-v7 版本, 与图 10 中的可视化结果对应起来进行验证分析, 这从定性和定量两个方面验证了相较于单独添加和两两组合添加损耗函数组件而言, 3 个损耗函数组件的组合添加获得了最好的去模糊性能. 这进一步说明了对 3 种损耗函数进行组合使用可以优势互补, 有效提升 DLGAN 的去模糊性能.

表 2 7 个 DLGAN 版本复原图像对应的评估指标分数

方法	PSNR	SSIM	FSIM
DLGAN-v1	25.033	0.805	0.889
DLGAN-v2	25.033	0.805	0.899
DLGAN-v3	25.087	0.796	0.888
DLGAN-v4	8.165	0.255	0.680
DLGAN-v5	24.645	0.789	0.884
DLGAN-v6	24.267	0.780	0.879
DLGAN-v7	29.395	0.905	0.962



(a) DLGAN-v1 (b) DLGAN-v2 (c) DLGAN-v3 (d) DLGAN-v4 (e) DLGAN-v5 (f) DLGAN-v6 (g) DLGAN-v7

图 10 消融实验的复原图像可视化结果

### 3 结论

本文针对动态场景去模糊方法需要使用成对训练数据的问题, 提出了一种对偶学习生成对抗网络. 该网络由注意力引导的生成器、多尺度鉴别器、由感知一致性损耗函数和特征感知损耗函数所构成的特征图损耗函数 3 个模块组成, 能够使用非成对训练

数据集以对偶学习的模式实现目标图片的动态场景去模糊, 克服了图像去模糊需要成对图像作为训练数据集的问题. 本文在 GoPro 数据集上进行了大量的实验, 验证了所提 DLGAN 方法能够使用非成对数据集训练网络; 在结构相似度和可视化评估方面, DLGAN 与 9 种使用成对数据集训练的图像去模糊方法相比

具有更好的性能.

### 参考文献(References)

- [1] 余孝源, 谢巍, 陈定权, 等. 一种基于改进的暗通道先验的运动模糊核估计方法[J]. 控制与决策, 2020, 35(7): 1667-1673.  
(Yu X Y, Xie W, Chen D Q, et al. A method of motion blurry kernel estimation based on improved dark channel prior[J]. Control and Decision, 2020, 35(7): 1667-1673.)
- [2] Lu B Y, Chen J C, Chellappa R. Unsupervised domain-specific deblurring via disentangled representations[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, 2020: 10217-10226.
- [3] Wen Y, Chen J, Sheng B, et al. Structure-aware motion deblurring using multi-adversarial optimized CycleGAN[J]. IEEE Transactions on Image Processing, 2021, 30: 6142-6155.
- [4] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]. 2017 IEEE International Conference on Computer Vision. Venice, 2017: 2242-2251.
- [5] Lee H Y, Tseng H Y, Mao Q, et al. DRIT++: Diverse image-to-image translation via disentangled representations[J]. International Journal of Computer Vision, 2020, 128(10/11): 2402-2417.
- [6] Isola P, Zhu J Y, Zhou T H, et al. Image-to-image translation with conditional adversarial networks[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 5967-5976.
- [7] Engin D, Genc A, Ekenel H K. Cycle-dehaze: Enhanced CycleGAN for single image dehazing[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City, 2018: 9380-9388.
- [8] Wei Y Y, Zhang Z, Wang Y, et al. DerainCycleGAN: Rain attentive CycleGAN for single image deraining and rainmaking[J]. IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society, 2021, 30: 4788-4801.
- [9] Johnson J, Alahi A, Li F F. Perceptual losses for real-time style transfer and super-resolution[C]. European Conference on Computer Vision. Cham: Springer, 2016: 694-711.
- [10] Kingma D, Ba J. Adam: A method for stochastic optimization[J/OL]. 2017, arXiv: 1412.6980.
- [11] Pan J S, Sun D Q, Pfister H, et al. Blind image deblurring using dark channel prior[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016: 1628-1636.
- [12] Nah S, Kim T H, Lee K M. Deep multi-scale convolutional neural network for dynamic scene deblurring[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 257-265.
- [13] Gong D, Yang J, Liu L Q, et al. From motion blur to motion flow: A deep learning solution for removing heterogeneous motion blur[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 3806-3815.
- [14] Zhang J W, Pan J S, Ren J, et al. Dynamic scene deblurring using spatially variant recurrent neural networks[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018: 2521-2529.
- [15] 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. Salt Lake City, 2018: 8183-8192.
- [16] Shao W Z, Liu Y Y, Ye L Y, et al. DeblurGAN+: Revisiting blind motion deblurring using conditional adversarial networks[J]. Signal Processing, 2020, 168: 107338.

### 作者简介

纪野(1990—), 男, 博士生, 从事计算机视觉领域的研究, E-mail: jiy521@bit.edu.cn;

戴亚平(1963—), 女, 教授, 博士生导师, 从事图像处理、决策支持系统等研究, E-mail: daiyaping@bit.edu.cn;

廣田薰(1950—), 男, 教授, 博士, 从事计算智能、智能信息处理等研究, E-mail: hirota@bit.edu.cn;

邵帅(1992—), 男, 博士后, 从事模式识别、无线传感器网络等研究, E-mail: shaoshuai@bit.edu.cn.