

基于低尺度细节恢复的单幅图像阴影去除方法

吴 文¹, 万 毅²

(1. 新疆理工学院信息工程系, 新疆阿克苏 843100; 2. 温州大学电气与电子工程学院, 浙江温州 325035)

摘 要: 为了在光照复杂、纹理丰富的图像上获得更好的去阴影效果, 基于生成对抗网络提出了一种新颖的阴影去除方法. 首先, 所提网络中的阴影检测子网为阴影图像生成阴影掩膜, 基于该检测结果提出一种光照敏感的多尺度图像分解方法, 在几乎不损失光照信息的同时提取图像纹理信息; 然后, 蒙版生成子网为分解后的低尺度图像生成相应的蒙版用于去除其中阴影; 其次, 边界复原子网修复阴影边界实现友好的过渡; 最后, 使用自适应衰减因子引导图像进行细节恢复以得到纹理丰富的结果. 实验结果表明所提方法可以有效地提高阴影去除效果.

关键词: 图像处理; 阴影去除; 生成对抗网络; 光照敏感; 多尺度分解

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 0372-2112 (2020)07-1293-10

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2020.07.007

Single Image Shadow Removal Using Low-Scale Detail Recovering

WU Wen¹, WAN Yi²

(1. Department of Information Engineering, Xinjiang Institute of Technology, Aksu, Xinjiang 843100, China;

2. School of Electrical and Electronic Engineering, Wenzhou University, Wenzhou, Zhejiang 325035, China)

Abstract: To obtain a better shadow removal result on images with complex illumination and texture, we proposed a novel approach based on generative adversarial networks. Firstly, shadow mask is generated by the shadow detection sub-net from input shadow image. Based on this detection result, we proposed an illumination sensitive multi-scale image decomposition method to extract the texture information with less or no illumination information loss. Secondly, shadow matte is generated by the matte generation sub-net for the low scale shadow image to remove shadows in it. Thirdly, the shadow boundary are naturally recovered by the boundary completion sub-net. Finally, the shadow removal result is obtained using a detail recovering method guided by adaptive attenuation factor. Experimental results show that the proposed method can improve the removal performance effectively.

Key words: image processing; shadow removal; generative adversarial networks; illumination sensitive; multi-scale decomposition

1 引言

阴影虽能为图像深度和物体几何形状预估等计算机视觉研究工作提供重要依据, 但其存在也会加大物体检测、目标跟踪等任务的难度. 因此, 在诸多计算机视觉应用中需要去除图像中的阴影.

已有的阴影去除方法常包含阴影定位和去除两个步骤. 在阴影定位过程中, 文献[1~3]使用阴影检测的方法, 其中基于统计学习的方法^[1]过度依赖自定义的阴影特征, 基于卷积神经网络(Convolutional Neural Networks, CNNs)^[4~6]特征学习的方法因受到数据集规模的限制, 网络层次较浅. 此外, 该方法以区域块的形式处理图像, 因此后续还需要进行全局优化以保持图像的连

贯性. 在阴影去除过程中, 传统的阴影去除方法有梯度域去除法^[7~10]以及颜色域去除法^[1,2,4,11,12]. 其文献[8, 9]基于梯度域仅仅只改变半阴影区域的梯度变量, 对于全阴影区域的光照变量并不适用. 文献[2]基于颜色域去除法使用贝叶斯公式提取蒙版图像以去除图像中的阴影, 而文献[12]使用强度表面恢复法去除阴影.

端到端的阴影去除方法^[13~17]以其运算效率迅速成为主流, 这些方法可分为本征分解法^[13]和深度学习方法^[14~18]. 其中文献[13]基于本征分解绕开了阴影检测环节, 但该方法会改变非阴影区域的颜色, 违背了阴影去除的初衷. 基于深度学习的诸多阴影去除方法^[14~18]中, 文献[14]基于CNNs, 结合图像中阴影的外表与语义信息为阴影图像生成蒙版以得到阴影去除结果. 文献[15]使用

两个条件生成对抗网络(conditional Generative Adversarial Network, cGANs)分别进行阴影检测和阴影去除. 文献[16]基于循环神经网络(Recurrent Neural Network, RNN)利用注意力机制去除图像中的阴影. 文献[17]基于cGANs使用非成对数据集训练模型,大大提高了实验样本场景的多样性. 文献[18]基于cGANs,利用生成器强大的图像生成能力,为阴影图像生成阴影蒙版以去除阴影.

以上基于深度学习的阴影去除方法^[14-18]普遍存在以下问题:

第一,该类方法使用阴影图像和标签无阴影图像组成的图像对用于监督学习训练. 其采集过程是使用拍摄装置先拍摄阴影图像,然后撤去不同形状的遮挡物以拍摄相应的无阴影图像. 而在拍摄图像对的过程中,由于摄像机曝光和姿势以及环境光照已经发生改变,导致图像对在颜色信息、光照信息或空间位置上发生改变. 最终导致模型无法精确的学习阴影和非阴影之间的内在联系.

第二,深度学习方法常需要对输入图像进行下采样以简化特征提取过程并减少神经元个数. 这将导致阴影去除结果细节丢失严重.

为解决以上问题,基于生成对抗网络提出了一种新颖的多任务网络结构和多尺度图像分解方法,能在得到边界过渡自然的去除结果的同时极大的保留其内部纹理信息.

2 多任务生成器对抗网络

生成对抗网络(Generative Adversarial Networks, GANs)^[19]由一个生成器 G 和一个判别器 D 组成. 生成器接受一个随机噪声并生成一幅逼真的图像. 判别器从训练集中学习并判断生成器生成的图像是否为真. 生成器生成图像旨在让判别器难以辨别真假,最终从训练集中学习数据的分布. cGANs^[20]进一步拓展了

GANs^[19],其允许在生成器和判别器中引入额外的条件变量. 文献[21]基于cGANs^[20]使用图像作为条件变量,训练生成器从而生成另一幅图像. 在阴影去除领域,文献[18]将阴影图像作为条件变量,较为简单地将cGANs应用到阴影去除领域,该方法旨在让生成器生成蒙版,然后使用式(1)^[12,14,18]计算得到无阴影图像.

$$\mathbf{x} = \mathbf{y} \cdot \mathbf{s} \quad (1)$$

其中,阴影图像 $\mathbf{x} \in [0,1]^{p \times q \times 3}$,无阴影图像 $\mathbf{y} \in [0,1]^{p \times q \times 3}$,则存在蒙版 $\mathbf{s} \in [0,1]^{p \times q \times 3}$ 满足式(1), \mathbf{s} 表示阴影尺度因子. 如图1所示,通过大量的实验,对比图1(b)和图1(c),可以发现文献[18]并不能为阴影图像生成一幅合理的蒙版图像,这导致了图1(d)中非阴影之外的区域也发生了大量的改变,这违背了阴影去除的初衷. 究其原因,训练样本的误差和深度学习^[14-18]方法对输入图像的下采样,导致最终结果存在明显的阴影边界且伴有全局层面的纹理信息缺失.

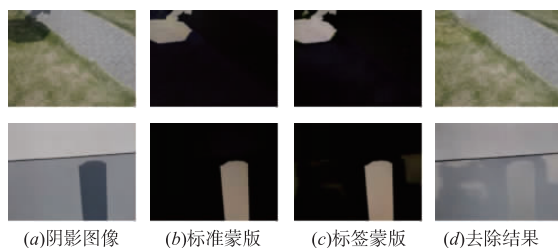


图1 文献[18]的结果

如图2所示,与已有的深度学习阴影去除方法^[14-18]不同,所提方法采用多任务生成器取代了传统的单任务形式. 首先,阴影检测子网检测图像中的阴影得到阴影掩膜;然后,结合阴影检测的结果对原始图像进行光照敏感的多尺度图像分解得到低尺度阴影图像,将其作为蒙版生成子网的输入;其次,蒙版生成子网为低尺度的阴影图像生成相应的蒙版,结合阴影检测

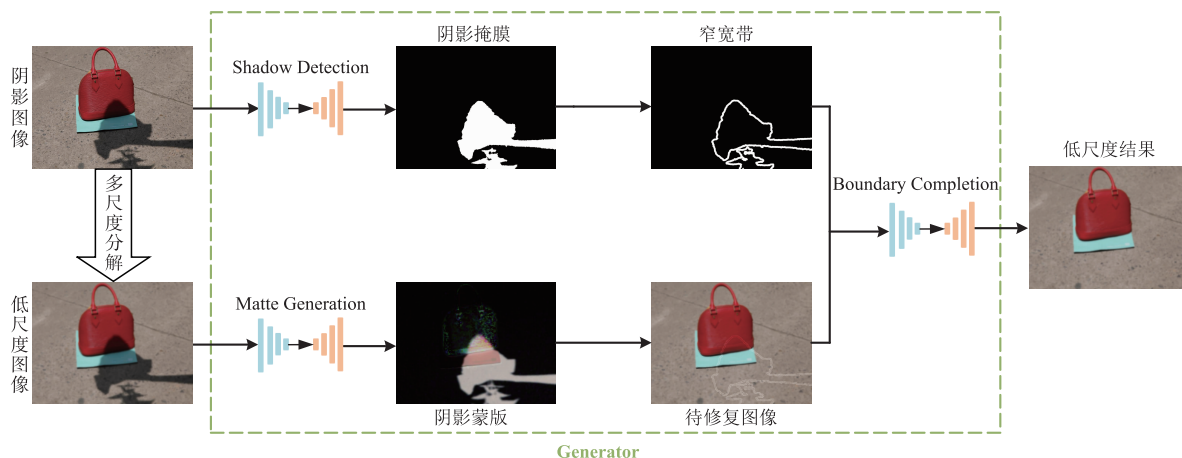


图2 生成器结构

结果对蒙版进行限制,使用式(1)得到待修复的阴影图像;最后,以阴影检测边界为基准构建窄宽带区域并结合待修复图像作为边界复原子网的输入,得到最终的低尺度阴影去除结果。

2.1 阴影检测子网

阴影检测子网可以为阴影图像生成阴影掩膜,其值为 1 时标识阴影区域,其值为 0 时标识非阴影区域。阴影检测子网作为功能较为独立的部分,其预训练的目标函数如式(2)所示。

$$L_M = \frac{1}{N} \sum_{i=1}^N \| \mathbf{M} - \hat{\mathbf{M}} \|^2 \quad (2)$$

其中 N 表示训练样本图像的个数;标签 (Ground Truth, GT) 阴影掩膜 $\mathbf{M} \in \{0, 1\}^{p \times q}$;在预测时,模型通常无法以百分之百的概率判断某个点是否为阴影,因此预测的阴影掩膜 $\hat{\mathbf{M}} \in (0, 1)^{p \times q}$,其值标识着每一个像素点被判定为阴影的概率,其值分布在 0 到 1 之间。子网的内部结构如表 1 所示。其中,卷积层 5(3)表示第五个卷积层重复三次;ReLU 为激活函数;BN 为批标准化^[22];卷积层 1 中第二行的输入项 4/4 表示阴影检测子网的输入为一幅四通道图像,分别由三通道的 RGB 阴影图像

和单通道的阴影掩膜组成;同理,子网中卷积层 12 的第三行输出分别是单通道的阴影检测结果表示为 1/1。如图 3 所示,所提子网可以准确地将阴影检测出来,而更为准确的阴影检测结果将有利于后续的多尺度图像分解以及自适应衰减因子的细节恢复。

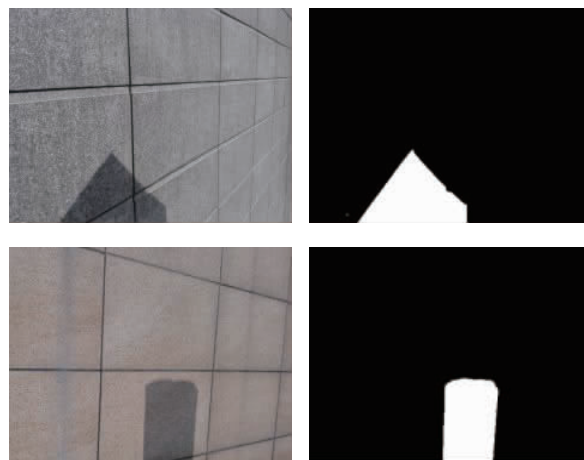


图3 阴影检测结果

表 1 子网结构

层级结构	卷积层 1	卷积层 2	卷积层 3	卷积层 4	卷积层 5(3)	卷积层 6	卷积层 7	卷积层 8(3)	卷积层 9	卷积层 10	卷积层 11	卷积层 12
输入	4/4	64	128	256	512	512	512	1024	1024	512	256	128
输出	64	128	256	512	512	512	512	512	256	128	64	1/1
前邻	-	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU
后接	-	BN	BN	BN	BN	-	BN	BN	BN	BN	BN	Tanh
跳跃	卷积层 12	卷积层 11	卷积层 10	卷积层 9	卷积层 8	-	-	卷积层 5	卷积层 4	卷积层 3	卷积层 2	卷积层 1

2.2 光照敏感的多尺度分解

所提的光照敏感多尺度分解方法旨在预先保留图像的细节信息,让其在以低尺度的方式去除图像中的阴影。如图 4 所示,多尺度分解旨在将图像分解成一个平滑的低尺度层和多个细节层。低尺度层包含图像主要的颜色信息,而细节层包含逐渐减少且模糊的形状信息。多尺度图像分解过程可以表示为式(3)。

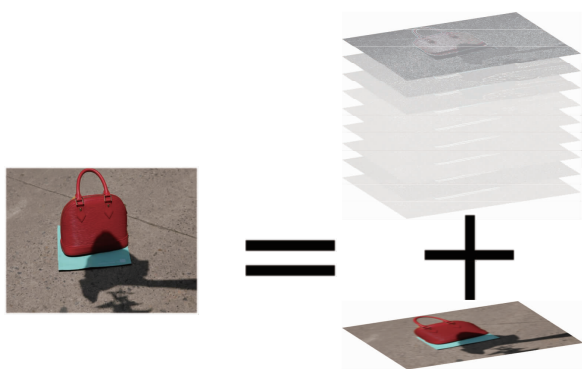


图4 光照敏感的多尺度分解

$$\mathbf{I} = \mathbf{x} + \sum_{i=1}^M \mathbf{L}_i \quad (3)$$

其中输入图像 $\mathbf{I} \in [0, 1]^{p \times q \times 3}$, M 表示分解次数,基础层图像 \mathbf{x} 和第 i 次平滑操作之后的细节层信息 \mathbf{L}_i 均为三通道的 RGB 图像。

为了合理地保留图像中的细节纹理信息,与常见的高斯多尺度分解不一样的是,所提方法将考虑到图像中的光照变化,结合阴影检测结果提出一种光照敏感的多尺度分解方法。因处于不同光照条件下的同一像素,其颜色、亮度会发生改变,但其材质不会发生改变^[1],基于像素之间的颜色相关性、亮度相关性以及色调相关性(式(4))分别为阴影区域和非阴影区域的像素找到相关的像素点集。颜色相关性表示两个具有相同亮度和反射比的像素其颜色是相似的,因此使用颜色相关性判断两个像素点是否具有相似的亮度和材质;亮度相关性表示当光源受到阻挡时,阴影区域的亮度常常低于非阴影区域。阴影和非阴影区域在色调上同样具有较大的差异,因此色调相关性是区分阴影和非阴影区域的一种重要特征。

$$D = D_c \times D_l \times D_h \quad (4)$$

其中颜色相关性 $D_c = \exp\left(-\frac{\|I_i - I_j\|^2}{2\sigma_c^2}\right)$ 、亮度相关性 $D_l = \exp\left(-\frac{(L_i - L_j)^2}{2\sigma_l^2}\right)$ 、色调相关性 $D_h = \exp\left(-\frac{(H_i - H_j)^2}{2\sigma_h^2}\right)$ 。 I_i 和 I_j 分别为点 i, j 在 RGB 空间的像素值; L_i 和 L_j 分别为点 i, j 在 Lab 空间中 L 通道的值; H_i 和 H_j 分别为点 i, j 在 HSI 颜色空间上 H 通道的值。在参数的设置在, 本文先经验性的固定 σ_c, σ_l 和 σ_h 的值, 其值的选取服从以下规律: 过大的 σ_c, σ_l 和 σ_h 将导致颜色、亮度以及色调相差较大的两个像素仍然均有较高的相关性。然后设置不同的相关性阈值 D , 使用式(5)计算像素在经过多尺度分解前后的光照信息损失。

$$Diff = \frac{1}{m} \frac{1}{n} \sum_{j=1}^m \sum_{i=1}^n (L - L')^2 \quad (5)$$

其中 m, n 分别为数据样本的个数和每张样本图像中的像素个数, L 和 L' 分别代表原始图像 I 和基础层图像 x 中各个像素经过光照敏感的多尺度分解前后在 Lab 空间中 L 通道的值。

如表 2 所示, 在可用数据集^[14,15]上进行参数网格化分析。先固定 $\sigma_c = 45, \sigma_l = 45, \sigma_h = 45$, 通过调整不同的阈值分析对应的光照信息损失 $Diff$, 图像中孤立点比例以及最终阴影去除结果与 GT 无阴影图像之间的像素均方根误差 (Root Mean Squared Error, RMSE)。其中, 孤立点比例为某组参数下的样本平均孤立点比例。通过实验量化分析, 发现较低的阈值伴随着较大的光照信息损失, 而过高的阈值则会产生过多的孤立点。在孤立点增多的同时, 光照信息损失虽然在减少, 但因后续

深度学习模型的下采样过程存在着大量的信息丢失, 而孤立点并不会参与光照信息提取过程, 导致存在着大量的细节纹理信息的丢失, 反而会影响最终的阴影去除效果, 其表现在 RMSE 的值会回弹。经过诸多实验分析, 设置 $D = 0.85$ 可以在大多数的阴影场景下获得较好的结果。因此, 当点 i 和 j 的相关系数大于 0.85 时视为相关点, 所有相关点将构成点集 R_i 。

表 2 参数网格化分析

指标	$D=0.7$	$D=0.75$	$D=0.8$	$D=0.85$	$D=0.9$	$D=0.95$
$Diff$	4.87	3.42	1.72	0.48	0.44	0.39
孤立点比例	5.23%	5.45%	5.64%	6.2%	10.2%	17.43%
RMSE	8.47	7.62	6.56	4.45	6.88	7.79

其次, 如式(6)所示, 使用所有相关点 R_i 的平均像素值作为新的像素值, 孤立点的像素值保持不变。

$$I'_i = \frac{\sum_{j \in R_i} \omega_{ij} I_j}{\sum_{j \in R_i} \omega_{ij}} \quad (6)$$

其中 ω_{ij} 是点 i 和 j 之间的相似度, 由两点的像素差异决定权重大小, $\omega_{ij} = \exp\left(-\frac{(I_i - I_j)^2}{2\sigma^2}\right)$ 。计算过程中, 以阴影检测子网的结果作为基准, 先后对阴影区域和非阴影区域进行光照敏感滤波。如图 5 所示, 展示了高斯多尺度分解以及光照敏感的多尺度分解在低尺度层和最终阴影去除结果上的差异。比较图 5(b) 和图 5(d), 可以观察到分区域的进行光照敏感的滤波方法可以更好保留图像中大致颜色、材质、形状信息, 阴影边界保留更为完整, 相较于直接进行全局的高斯滤波图像平滑, 最终的阴影去除效果也更好。

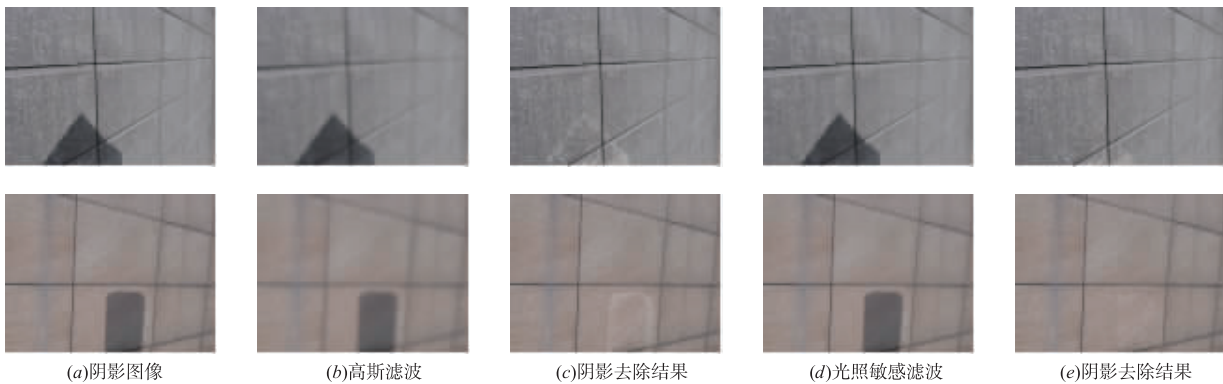


图 5 不同多尺度分解方法的结果

2.3 区域受限的蒙版生成结果

蒙版生成子网可以为输入阴影图像生成相应的蒙版以达到阴影去除的目的。但因普遍存在的数据集图像对之间的光照不一致现象, 导致生成蒙版并不准确, 阴影去除结果在全局均会发生改变。为解决该问题, 基于阴影检测子网的检测结果对蒙版生成结果进行修正(式(7))。

$$\hat{s}_i = \begin{cases} \hat{s}_i, & \text{if } \hat{M}_i = 1 \\ 0, & \text{if } \hat{M}_i = 0 \end{cases} \quad (7)$$

该步骤将仅在阴影区域保留蒙版的值。如图 6(d) 所示, 图 6(c) 展示了区域受限的蒙版结果, 以及使用式(1)计算得到的低尺度阴影去除结果, 见图 6(d)。因为阴影掩膜的特性, 其展示的是阴影检测的硬边界, 也导致低尺

度阴影去除结果中存在少量未经处理的阴影区域.

2.4 边界复原子网

边界复原子网旨在解决阴影边界过渡不均匀的问题^[14-18]. 使用阴影检测边界像素 5 领域的范围构建窄宽带区域 \mathbf{B} 可以很好的覆盖大多数场景下的阴影边界, 将窄宽带区域 \mathbf{B} 作为低尺度阴影去除图像 \mathbf{y}' 的待修复区域. 其网络结构与阴影检测子网一样, 使用式

(8) 对其预训练以降低最终训练的复杂程度.

$$L_M = \frac{1}{N} \sum_{i=1}^N \|\mathbf{B} \cdot (\mathbf{y} - R((1 - \mathbf{B}) \cdot \mathbf{y}'))\|^2 \quad (8)$$

其中窄宽带 $\mathbf{B} \in [0, 1]^{p \times q \times 1}$, \mathbf{y} 为低尺度的 GT 无阴影图像, \mathbf{y}' 为低尺度的阴影去除图像, $R(\cdot)$ 为边界复原子网的输出. 边界复原子网旨在缩小修复区域与 GT 无阴影图像在窄宽带区域 \mathbf{B} 上的像素差异.

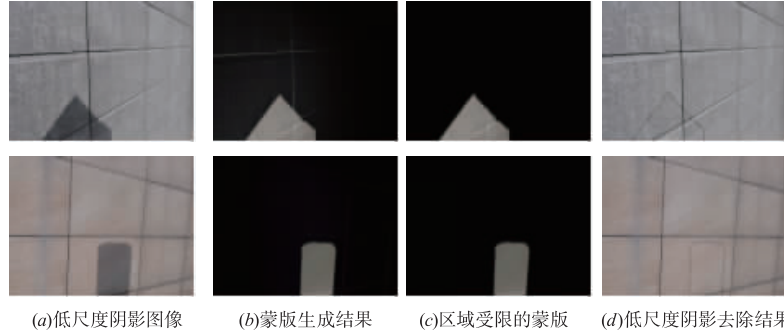


图6 区域受限的蒙版生成结果

受到文献[23]的启发, 若模型仅仅使用式(8)中的像素差异作为优化目标, 这种基于像素的优化方式会忽略了图像的全局相似性, 导致阴影边界修复的半阴影区域视觉效果较差. 本文结合感知损失函数^[23]最小化阴影去除图像与 GT 无阴影图像在图像语义层面的差异, 见式(9). 不同的是, 本文不再使用预先训练好的深度网络提取 GT 无阴影图像和阴影去除图像 \mathbf{y}' 的特征, 而使用后续的判别器提取二者的高级语义特征, 同时大大降低了模型的复杂度与吞吐量.

$$L_p = \frac{1}{C_i H_i W_i} \|D_i(\mathbf{y}') - D_i(\mathbf{y})\|^2 \quad (9)$$

其中 \mathbf{y}' 为低尺度阴影去除图像, D_i 是判别器的第 i 层, C_i 是第 i 层对应的通道数, H_i 和 W_i 是第 i 层特征图的长和宽. 如图 7 所示, 边界复原子网修可以很好地修复窄宽带里面的阴影边界区域.

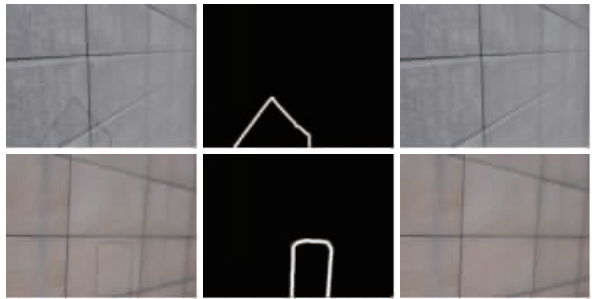


图7 边界复原结果

2.5 判别器

判别器旨在判断生成器为阴影图像生成的低尺度阴影去除图像是否为真, 本质在该阶段为一个图像二

分类问题. 所提方法的判别器与文献[20]类似, 包含多个卷积块, 每个卷积块中, 卷积层都紧跟着批标准化^[22]和激活函数 LeakyReLU. 判别器的最后一层为一个 Sigmoid 函数, 其输出为某对图像为真的概率值, 其网络结构如表 3 所示. 其输入为低尺度阴影图像 \mathbf{x} 和生成的阴影去除图像 \mathbf{y}' 或 GT 无阴影图像 \mathbf{y} , 因为同为三通道的 RGB 图像, 表 3 示为 6/6^[15].

表 3 判别器的内部结构

网络名称	层次结构	卷积层 1	卷积层 2	卷积层 3	卷积层 4	卷积层 5
判别器	输入	6/6	64	128	256	512
	输出	64	128	256	512	1
	前接	-	LReLU	LReLU	LReLU	LReLU
	后邻	-	BN	BN	BN	Sigmoid

2.6 损失函数

与大多数运用 cGANs^[20] 的文章不一样的是, 所提网络的损失函数由最小均方损失主导, 平均绝对误差损失以及感知损失加权组合而成, 如式(10).

$$L = \alpha L_l + \beta L_m + \gamma L_p \quad (10)$$

如表 4 所示, 本文经过诸多实验提供了各种组合类型的损失函数对应的 RMSE 和结构相似性 (Structural Similarity Index, SSIM). 与常规的对损失函数 L_a ^[20] 不一样, 见式(11), 通过观察实验量化数据, 发现使用最小均方损失 L_l ^[24] 为主导可以得到更好的结果, 如式(12)所示.

$$L_a = E_{\mathbf{x}, \mathbf{y} \sim p_{\text{data}}(\mathbf{x}, \mathbf{y})} [\log D(\mathbf{x}, \mathbf{y})] + E_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})} [\log(1 - D(\mathbf{x}, \mathbf{y}'))] \quad (11)$$

$$L_l = E_{x,y \sim p_{\text{data}}(x,y)} [D(x,y)^2] + E_{x \sim p_{\text{data}}(x)} [(D(x,y') - 1)^2] \quad (12)$$

其中 x 表示输入的阴影图像, y 表示 GT 无阴影图像, y' 为低尺度的阴影去除图像, $D(x)$ 是判别器.

表 4 不同种类的损失函数的量化评估

损失函数	L_a	L_l	$L_l + L_m$	$L_l + L_p$	$L_l + L_m + L_p$
RMSE	5.71	5.44	4.83	4.56	4.45
SSIM	0.9717	0.9752	0.9781	0.9815	0.9833

此外,两个平均绝对误差损失分别促进蒙版生成结果和低尺度阴影去除结果与对应的真实训练集更为接近,如式(13).

$$L_m = \frac{1}{N} \sum_{i=1}^N \|s_i - \hat{s}_i\|^2 + \frac{1}{N} \sum_{i=1}^N \|y_i - y'_i\|^2 \quad (13)$$

3 自适应衰减因子的图像细节恢复

在生成器得到边界过渡自然的低尺度阴影去除图像之后,使用多尺度得到的多个细节层对图像进行细节恢复,因为亮度较低的阴影区域常常会弱化细节信息,为了高效的恢复阴影区域的细节,从而引入衰减因子 α 并结合多尺度细节复原,最终的阴影去除结果 $\hat{y} = x + \alpha \sum_{i=1}^N L_i$. 其中衰减因子 α 由阴影区域和非阴影区域的像素值自适应得到, $\alpha = \frac{D_n}{D_s}$. D_n, D_s 分别是细节层中的非阴影区域和阴影区域的平均像素值,细节恢复结果如图 8 所示. 在该实验中,第一行的自适应衰减因子 $\alpha = 1.87$,第二行的衰减因子 $\alpha = 2.41$. 使用该衰减因子作为引导进行多尺度细节复原,能够还原图像的细节纹理信息以得到较好的阴影去除效果.

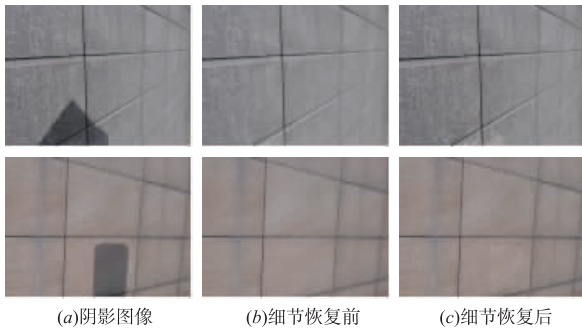


图 8 图像细节恢复

4 实验结果与分析

在数据集部分,挑选 SRD^[14] 和 ISTD^[15] 两个有代表性的数据集中共 4958 组图像进行一系列预处理构建综合筛选数据集. 基于 RMSE 筛选阴影图像和 GT 无阴影图像在非阴影区域的误差较少且场景多样的 1500 组图

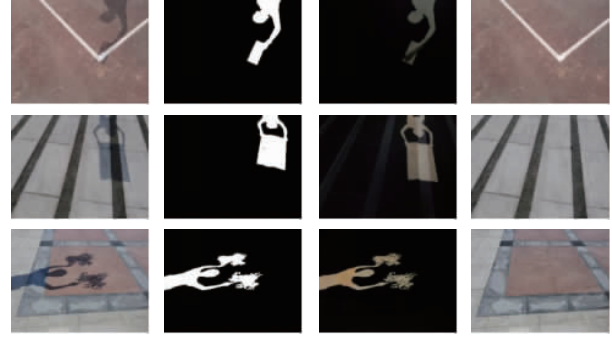


图 9 综合筛选数据集示例

像,如图 9(a) 所示. 其中图 9(a) 和图 9(b) 用于预训练阴影检测子网;以图 9(b) 为基础取 5 领域窄带和图 9(d) 用于预训练边界复原子网. 将整个数据集的 80% 作为训练集,剩余 20% 作为测试集,在测试集上采取 4 折交叉验证的方法以确保超参数的有效性.

4.1 训练细节

基于 Tensorflow^[25] 使用 Python 在 Ubuntu 18.04 下完成编程实验,图像处理器(GPU)采用 NVIDIA GTX 1080Ti. 网络中, LReLU 的斜率设置为 0.25, 使用 Adam 作为目标函数的梯度下降方法. 与传统的 cGANs^[20] 一样,训练时,生成器与判别器交替更新参数,在 600 个循环后停止训练. 按照均值为 0、方差为 0.2 的正态分布初始化生成器和判别器中所有的卷积层和反卷积层的权重,设置偏置值为 0. 数据增大采用图像裁剪的方法,将原始 286×286 的图像裁剪成多个 256×256 的子图像,然后进行水平翻转进一步增大数据集. 通过大量实验,所提网络的参数初始化为 $\alpha = 0.5, \beta = 0.3, \gamma = 0.2$.

4.2 客观评价指标

使用 RMSE(式(14))和 SSIM^[26](式(15))作为阴影去除衡量指标. 其中 RMSE 衡量阴影去除图像和 GT 无阴影图像之间的误差,而 SSIM 反映了结构信息,其更符合人类视觉对图像的感知,其从图像的组成的角度将结构信息定义为独立于亮度、对比度,反映场景中物体结构的属性,并将失真建模为亮度、对比度和结构三个不同因素的组合. 其用均值作为亮度的估计,标准差作为对比度的估计,协方差作为结构相似程度的度量.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (I - I')^2} \quad (14)$$

其中 I 和 I' 分别代表某像素经过处理前后的像素值,由 R、G、B 三通道组成, n 为输入图像的像素点个数.

$$\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (15)$$

其中 μ_x 是 x 的平均值, μ_y 是 y 的平均值, σ_x^2 是 x 的方差, σ_y^2 是 y 的方差, σ_{xy} 是 x 和 y 的协方差. $c_1 = (k_1 L)^2$,

$c_2 = (k_2 L)^2$ 是用来稳定的常数,其中 L 是像素值的动态范围,而 $k_1 = 0.01, k_2 = 0.03$. 结构相似性的范围是 0 到 1,当两张图像一模一样时,SSIM 的值等于 1.

4.3 实验结果与量化分析

在实验中,对于基于自定义特征的传统机器学习方法^[1],直接将测试集作为模型输入得到阴影去除结果;对于基于深度学习的阴影去除方法^[15,16],为了保证实验的公平性,使用方式一:保留它们各自的网络结构

表 5 三种深度学习方法在 SRD^[14]数据集上的像素均方根误差与结构相似性

数据集	区域类型	Wang ^[15]		Hu ^[16]		本文方法	
		RMSE	SSIM	RMSE	SSIM	RMSE	SSIM
SRD 数据集 ^[14]	整体	8.77	0.9645	6.67	0.9737	5.64	0.9789
	阴影区域	12.98	0.9417	11.88	0.9489	10.73	0.9518
	非阴影区域	7.95	0.9729	4.94	0.9826	4.51	0.9876

表 6 三种深度学习方法在 ISTD^[15]数据集上的像素均方根误差与结构相似性

数据集	区域类型	Wang ^[15]		Hu ^[16]		本文方法	
		RMSE	SSIM	RMSE	SSIM	RMSE	SSIM
ISTD 数据集 ^[15]	整体	7.47	0.9774	7.21	0.9782	6.71	0.9794
	阴影区域	10.33	0.9423	10.18	0.9454	9.69	0.9513
	非阴影区域	6.93	0.9831	7.07	0.9846	6.17	0.9894

在筛选数据集上将本文与文献[1]、文献[15]以及文献[16]进行比较,其中本文与文献[15]以及文献[16]采用方式一进行比较,其结果如图 10 所示.其中文献[1]基于传统机器学习的方法,使用自定义的阴影特征,先将阴影检测出来,然后基于物理光照模型将阴影去除;文献[15]利用栈式条件生成对抗网络,分别使用两个 cGAN 做阴影检测和阴影去除;文献[16]基于循环神经网络的注意力机制先后进行阴影检测和去除.如图 10(c)中文献[1]因使用缺乏鲁棒性的阴影特征,从而导致阴影检测的不精准,阴影检测不准确将无法得到令人满意的阴影去除结果.图 10(d)和(e)中,文献[15,16]使用深度学习的特征提取能力,相较于文献[1]的方法能够得到更好的结果,但是复原区域存在明显的亮度偏高,图像全局模糊,图像整体过渡不均匀,例如第一行到第四行中依然能够看到遮挡物和人物的轮廓,这说明深度网络对于阴影的检测与识别能力虽然已经超过了基于自定义特征阴影检测能力,但在去除方面仍有提升空间.

其次,采用方式二将以上三种深度学习阴影去除方法在和 ISTD^[15]数据集上进行比较,选择场景多样、纹理复杂(阴影形状复杂、跨物质以及文字与阴影混合等情形)的几幅图像作为展示素材,其结果如图 11 所示.

最后,因光照强度的影响,弱阴影相较于强阴影存在不明显的阴影边界,因此弱阴影去除也一直是衡量阴影去除方法鲁棒性的一个重要参考指标.如图 12 所示,选择几组来自于 SRD^[14]与 ISTD^[15]的弱阴影图像,

和对抗损失函数^[20],而训练数据和测试数据均来自于的所提方法的筛选数据集^[14,15],以体现所提方法在目标函数和处理方法上的可行性.另一方面,使用方式二:将文献[15,16]与所提方法直接在 SRD 和 ISTD 数据集上进行比较,保留其各自的目标函数与训练方式旨在体现筛选数据集步骤的有效性,其量化数据见表 5 和表 6.

所提方法同样展示了其有效性.

将本文方法与文献[1]、文献[15]以及文献[16]进行量化评估,如表 7、表 8 所示.将图像从阴影区域、非阴影区域以及整体区域进行量化分析.表 7、表 8 分别计算测试数据集中 RMSE 与 SSIM 的平均值. RMSE 值越低说明阴影去除图像与 GT 无阴影图像之间的误差越小, SSIM 值越大说明阴影去除图像从结构上与 GT 无阴影图像之间更为相似.相较于文献[16],所提方法在 RMSE 上降低了 19.67%,在结构相似性上提高了 0.38%.

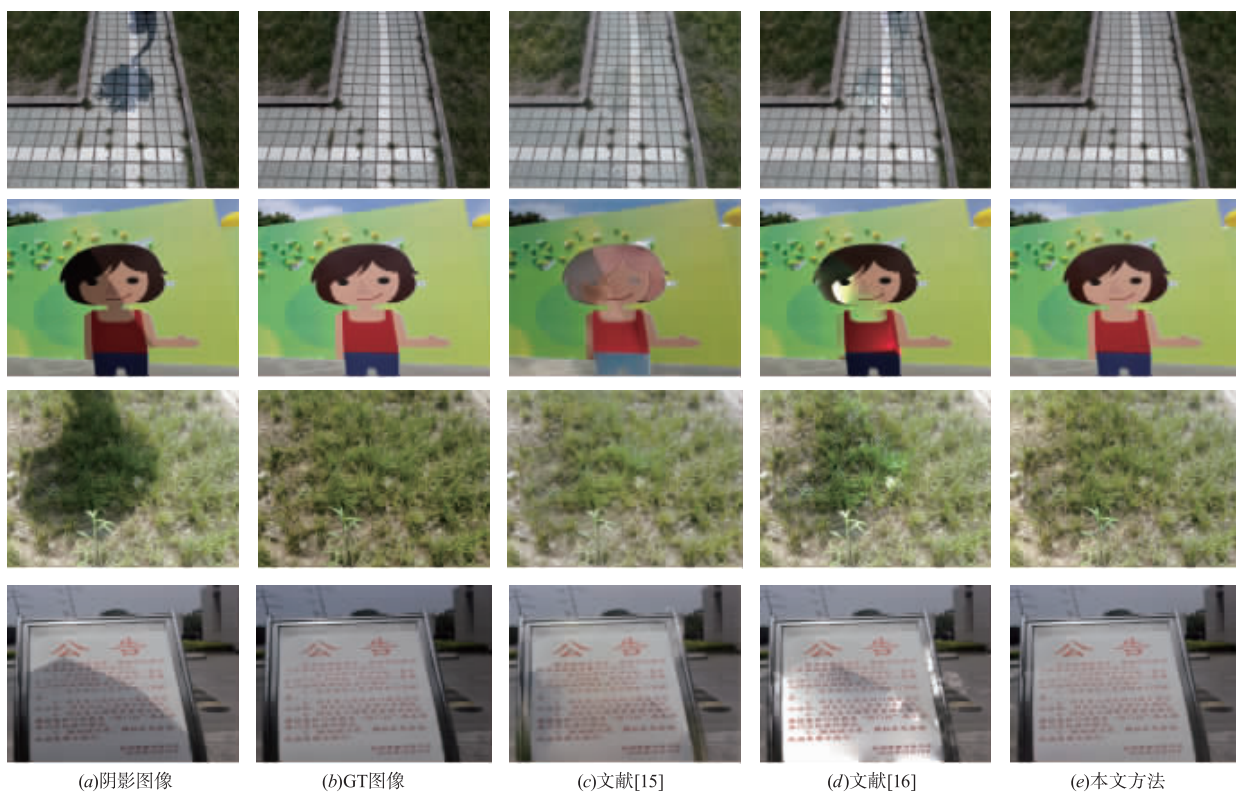
表 7 各个方法在不同区域的像素均方根误差值

数据集	区域类型	Guo ^[1]	Wang ^[15]	Hu ^[16]	本文方法
筛选数据集 ^[14,15]	整体	9.22	6.28	5.54	4.45
	阴影区域	25.65	10.57	9.93	7.24
	非阴影区域	4.74	3.56	3.42	2.67

表 8 各个方法在不同区域的结构相似性

数据集	区域类型	Guo ^[1]	Wang ^[15]	Hu ^[16]	本文方法
筛选数据集 ^[14,15]	整体	0.9672	0.9782	0.9795	0.9833
	阴影区域	0.9348	0.9551	0.9642	0.9774
	非阴影区域	0.9823	0.9842	0.9847	0.9854

采用方式二,在 SRD^[14]数据集和 ISTD^[15]数据集上,分别将所提方法与文献[15]、文献[16]进行量化评估,如表 5、表 6 所示.表 5、表 6 分别在 SRD^[14]和 ISTD^[15]数据集上分析所提方法与文献[15]、文献[16]的 RMSE 与 SSIM 的平均值.根据表 5,在 SRD^[14]数据集上,相较于文献[16],本文在 RMSE 上降低了 15.44%,



在结构相似性上提高了 0.53% ;根据表 6,在 ISTD^[15]数据集上,相较于文献[16],本文在 RMSE 上降低了 6.93%,在结构相似性上提高了 0.12%.

在时间复杂度上,所提方法以深度网络模型为主体,相较于传统的基于机器学习的阴影去除方法^[1]固然有质的提升.而较于文献[15]和文献[16],虽在网络模型的

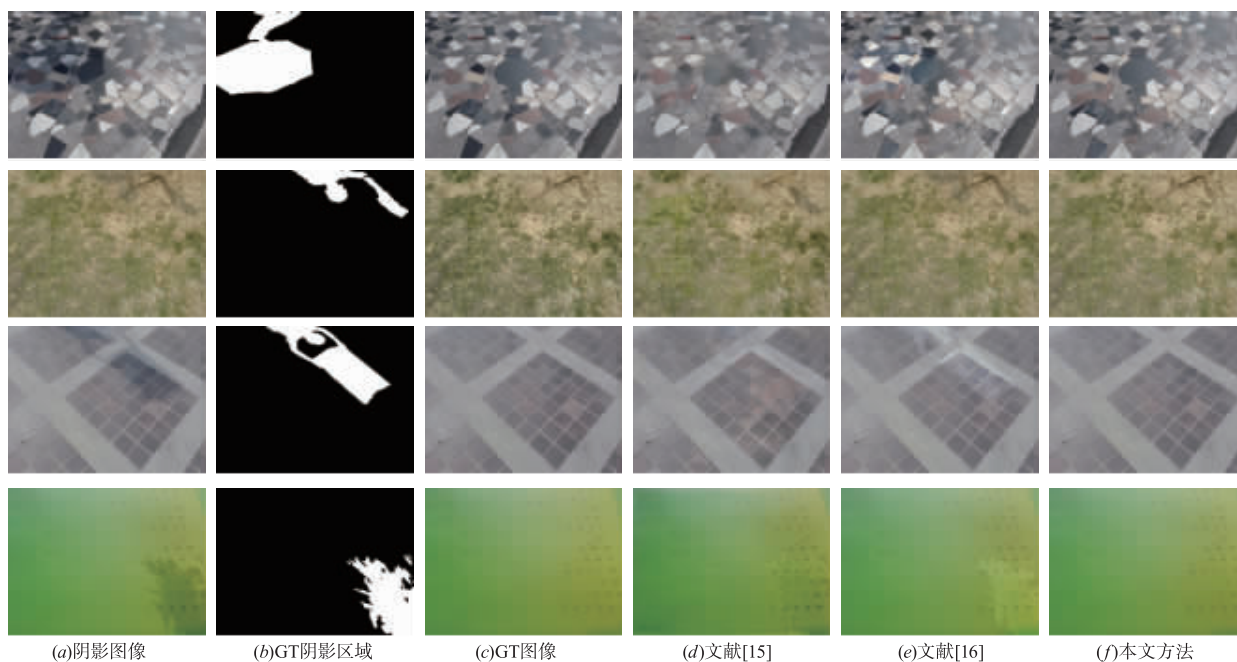


图12 三种深度学习方法的在弱阴影下的阴影去除结果

复杂程度上相差不大,但额外多尺度图像分解以及细节复原分别作为阴影去除的预处理和后处理步骤确实会增加额外的时间复杂度.在本文硬件水平下,对于一幅 840×640 的高清图像,文献[1]的处理时间约 180s;文献[15,16]作为深度学习模型,其主要时间花费于训练过程,处理时间均为 9s;所提方法牺牲了部分的时间复杂度以获取清晰度较高、纹理丰富的结果,处理时间约 17s.

5 结论

为了在场景更为复杂的阴影图像得到细节更为丰富、更加清晰的阴影去除结果.提出了一种新颖的生成器结构,分别由阴影检测子网、蒙版生成子网以及边界复原子网构成;其次,基于阴影检测结果提出一种光照敏感多尺度图像分解方法,该方法可以在几乎不损失光照信息的前提下提取图像中的纹理信息;最后,使用自适应衰减因子的细节恢复方法得到细节丰富的阴影去除图像.未来将研究深度学习方法在面临新数据集时,性能下降的问题,实现阴影去除的域迁移(Domain Shift)方法.

参考文献

[1] GUO Rui-qi, DAI Qie-yun, DEREK H. Paired regions for shadow detection and removal[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(12): 2956 - 2967.

[2] KHAN S H, BENNAMOUN M, SOHEL F, et al. Automatic shadow detection and removal from a single image[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(3): 431 - 446.

[3] VICENTE T F Y, HOAI M, SAMARAS D. Leave-one-out kernel optimization for shadow detection and removal[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(3): 682 - 695.

[4] 舒坚, 张学佩, 等. 基于深度卷积神经网络的多节点间链路预测方法[J]. 电子学报, 2018, 46(12): 2970 - 2977. SHU Jian, ZHANG Xue-pei, et al. Multi-nodes link prediction method based on deep convolution neural networks[J]. Acta Electronica Sinica, 2018, 46(12): 2970 - 2977. (in Chinese)

[5] 李康, 李亚敏, 等. 基于卷积神经网络的鲁棒高精度目标跟踪算法[J]. 电子学报, 2018, 46(9): 2087 - 2093. LI Kang, LI Ya-min, et al. A robust and accurate object tracking algorithm based on convolutional neural network[J]. Acta Electronica Sinica, 2018, 46(9): 2087 - 2093. (in Chinese)

[6] 赵耀霞, 吴桐, 韩焱. 基于卷积神经网络的复杂构件内部零件装配正确性识别[J]. 电子学报, 2018, 46(8): 1983 - 1988. ZHAO Yao-xia, WU Tong, HAN Yan. Identifying the correctness of fit of internal components based on a convolutional neural network[J]. Acta Electronica Sinica, 2018, 46(8): 1983 - 1988. (in Chinese)

[7] LIU Feng, GLEICHER M. Texture-Consistent shadow removal[A]. European Conference on Computer Vision[C]. Berlin, Heidelberg: Springer, 2008. 437 - 450.

[8] FINLAYSON G D, HORDLEY S D, DREW M S. Removing shadows from images[A]. European Conference on Computer Vision[C]. Berlin, Heidelberg: Springer, 2002. 823 - 836.

[9] FINLAYSON G D, HORDLEY S D, et al. On the removal of shadows from images[J]. IEEE Transactions on Pattern

- Analysis and Machine Intelligence, 2006, 28(1):59–68.
- [10] MOHAN A, TUMBIN J, CHOUDHURY P. Editing soft shadows in a digital photograph [J]. IEEE Computer Graphics and Applications, 2007, 27(2):23–31.
- [11] XIAO C, XIAO D, ZHANG L, et al. Efficient shadow removal using subregion matching illumination transfer [J]. Computer Graphics Forum, 2013, 32(7):421–430.
- [12] ARBEL E, HELOR H. Shadow removal using intensity surfaces and texture anchor points [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(6):1202–1216.
- [13] BARRON J T, MALIK J. Shape, illumination, and reflectance from shading [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(8):1670–1687.
- [14] QU Liang-qiong, TIAN Jian-dong, HE Sheng-feng, et al. DshadowNet: A multi-context embedding deep network for shadow removal [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. Honolulu, Hawaii, USA: IEEE, 2017. 4067–4075.
- [15] WANG Ji-feng, LI Xiang, YANG Jian. Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal [A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition [C]. USA: IEEE, 2018. 1788–1797.
- [16] HU Xiao-wei, FU Chi-wing, ZHU Lei, et al. Direction-aware spatial context features for shadow detection and removal [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, arXiv:1712.04142.
- [17] HU Xiao-wei, JIANG Yi-tong, FU Chi-Wing, et al. Mask-ShadowGAN: Learning to remove shadows from unpaired data [J]. arXiv Preprint, 2019, arXiv:1903.10683.
- [18] FAN Hui, HAN Meng, LI Jin-jiang. Image shadow removal using end-to-end deep convolutional neural networks [J]. Applied Sciences, 2019, 9(5):1009.
- [19] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets [A]. Advances in Neural Information Processing Systems [C]. Montreal, Canada, 2014. 2672–2680.
- [20] MIRZA M, OSINDERO S. Conditional generative adversarial nets [J]. arXiv Preprint, 2014, arXiv:1411.1784.
- [21] ISOLA P, ZHU J Y, et al. Image-to-image translation with conditional adversarial networks [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. Honolulu, Hawaii, USA: IEEE, 2017. 1125–1134.
- [22] LOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [A]. International Conference on Machine Learning [C]. Lille, France: IEEE, 2015. 448–456.
- [23] JOHNSON J, ALAHI A, LI Fei-fei. Perceptual losses for real-time style transfer and super-resolution [A]. European Conference on Computer Vision [C]. Amsterdam, The Nether Lands, 2016. 694–711.
- [24] MAO Xu-dong, et al. Least squares generative adversarial networks [A]. Proceedings of the IEEE International Conference On Computer Vision [C]. Venice, Italy: IEEE, 2017. 2794–2802.
- [25] WONGSUPHASAWAT K, SMILKOV D, WEXLER J, et al. Visualizing dataflow graphs of deep learning models in TensorFlow [J]. IEEE Transactions on Visualization and Computer Graphics, 2018, 24(1):1–12.
- [26] 杨学志, 叶铭, 等. 采用混合特征相似性的极化 SAR 图像降噪算法 [J]. 电子学报, 2016, 44(11):2583–2591.
- YANG Xue-zhi, YE Ming, et al. Speckle reduction for Pol-SAR images using hybrid features similarity [J]. Acta Electronica Sinica, 2016, 44(11):2583–2591. (in Chinese)

作者简介



吴文男, 1994年5月出生, 湖北武汉人. 2019年毕业于湖北大学, 获理学硕士学位, 现为新疆理工学院信息工程系教员, 主要研究方向为图像处理、深度学习等.
E-mail: 1119764335@qq.com



万毅(通讯作者)男, 1971年2月出生, 浙江温州人. 系温州大学教授、硕士生导师、IEEE会员. 2012年毕业于西南交通大学工学博士学位. 现为新疆理工学院援疆教授, 主要从事系统可靠性和智能算法方面的研究工作.