



智能系统学报

CAAI TRANSACTIONS ON INTELLIGENT SYSTEMS

空洞卷积与注意力融合的对抗式图像阴影去除算法

刘万军, 佟畅, 曲海成

引用本文:

刘万军, 佟畅, 曲海成. 空洞卷积与注意力融合的对抗式图像阴影去除算法[J]. 智能系统学报, 2021, 16(6): 1081–1089.

LIU Wanjun, TONG Chang, QU Haicheng. An antagonistic image shadow removal algorithm based on dilated convolution and attention mechanism[J]. *CAAI Transactions on Intelligent Systems*, 2021, 16(6): 1081–1089.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202011022>

您可能感兴趣的其他文章

基于注意力机制的显著性目标检测方法

Salient object detection method based on the attention mechanism

智能系统学报. 2020, 15(5): 956–963 <https://dx.doi.org/10.11992/tis.201903001>

基于注意力融合的图片描述生成方法

An image caption generation method based on attention fusion

智能系统学报. 2020, 15(4): 740–749 <https://dx.doi.org/10.11992/tis.201910039>

基于生成式对抗网络的道路交通模糊图像增强

Enhancement of blurred road-traffic images based on generative adversarial network

智能系统学报. 2020, 15(3): 491–498 <https://dx.doi.org/10.11992/tis.201903041>

层次化双注意力神经网络模型的情感分析研究

Hierarchical double-attention neural networks for sentiment classification

智能系统学报. 2020, 15(3): 460–467 <https://dx.doi.org/10.11992/tis.201812017>

注意力机制和Faster RCNN相结合的绝缘子识别

Insulator recognition based on attention mechanism and Faster RCNN

智能系统学报. 2020, 15(1): 92–98 <https://dx.doi.org/10.11992/tis.201907023>

基于双向消息链路卷积网络的显著性物体检测

Salient object detection based on bidirectional message link convolution neural network

智能系统学报. 2019, 14(6): 1152–1162 <https://dx.doi.org/10.11992/tis.201812003>

微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.202011022

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20210901.1633.008.html>

空洞卷积与注意力融合的对抗式图像阴影去除算法

刘万军, 佟畅, 曲海成

(辽宁工程技术大学软件学院, 辽宁 葫芦岛 125105)

摘要: 为了解决暗区域、纹理复杂或半影区域的阴影去除效果不明显的问题, 提出了空洞卷积与注意力机制融合的对抗式图像阴影去除算法。该算法基于生成对抗网络的总体思想, 将空洞卷积引入残差网络中, 用自定义的空洞残差块进行特征提取, 扩大了特征提取的感受野。在注意力编码阶段, 加入 4 层相同结构的空洞卷积, 确保最小计算量的情况下为解码阶段提供更抽象、更本质的全局的语义特征。运用多重注意力机制, 引导判别网络对无阴影图像的鉴别, 提高判别网络能力。该算法分别在 ISTD(image shadow triplets dataset) 与 SRD(shadow removal dataset) 公开数据集上进行检验, SSIM(structural similarity) 值达到 97.77%。该算法图像特征信息保存完整, 画面清晰, 暗区域及地物复杂的区域阴影去除效果较好, 对半影区域, 也有具有良好的表现。

关键词: 生成对抗网络; 空洞卷积; 多重注意力; 残差网络; 多尺度; 自编码; 长短记忆法; 阴影去除

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1673-4785(2021)06-1081-09

中文引用格式: 刘万军, 佟畅, 曲海成. 空洞卷积与注意力融合的对抗式图像阴影去除算法 [J]. 智能系统学报, 2021, 16(6): 1081-1089.

英文引用格式: LIU Wanjun, TONG Chang, QU Haicheng. An antagonistic image shadow removal algorithm based on dilated convolution and attention mechanism[J]. CAAI transactions on intelligent systems, 2021, 16(6): 1081-1089.

An antagonistic image shadow removal algorithm based on dilated convolution and attention mechanism

LIU Wanjun, TONG Chang, QU Haicheng

(Software College, Liaoning Technical University, Huludao 125105, China)

Abstract: To solve the problem of the unobvious shadow removal effect in dark areas or complex textured and penumbra areas, an antagonistic image shadow removal algorithm is proposed based on a dilated convolution and attention mechanism. The algorithm is based on the general idea of generative adversarial networks. First, the dilated convolution is introduced into the residual network, and the user-defined hole residual block is used for feature extraction, expanding the receptive field of feature extraction. Second, in the attention coding stage, four layers of dilated convolution with the same structure are added to provide more abstract and essential global semantic features for the decoding phase with a minimum calculation amount. Finally, the multiple attention mechanism is used to guide the discrimination network to identify the unshadowed image; thus, improving the discrimination network's ability. The proposed algorithm is tested on image shadow triplets dataset and shadow removal public datasets and achieves the structural similarity of 97.77%. The image feature information of the algorithm is well preserved, the picture is clear, the shadow removal effect is good in the dark area and complex area, and the algorithm has good performance for the penumbra area.

Keywords: generative adversarial networks; hole convolution; multiple attention; residual network; multi-scale; auto-coder; long short-term memory; shadow removal

收稿日期: 2020-11-20. 网络出版日期: 2021-09-01.

基金项目: 国家自然科学基金项目 (41701479); 辽宁省教育厅科学研究经费项目 (LJ2019JL010); 辽宁省自然科学基金面上项目 (20180550529).

通信作者: 佟畅. E-mail: 1879031567@qq.com.

光线穿过一切不透明物体都会产生阴影, 图像中阴影的存在十分常见。图像阴影会对目标检测与识别、图像分割等问题产生一定影响。因此, 去除图像中的阴影是一个十分重要的研究内容。

去除阴影的算法按实现方法可分为基于物理模型的方法、图像自身特征的方法、机器学习和深度学习等方法。从物理模型^[1-3]出发,进行阴影去除,利用光源与障碍物的关系进行建模,具有一定的可行性,但建模过程的计算量大,所需的计算参数精确度欠缺。多尺度与形态学方法^[4]、区域匹配算法^[5]、设置纹理置信区间^[6]等算法在计算过程中存在着固有的误差。归一化植被指数(normalized difference vegetation index, NDVI)阴影影响去除线性模型^[7],需对图片特定区域进行处理。递归式阴影注意力模型(recursive shadow attention model, RSAM)、新的生成器结构及多尺度图像分解法^[8-9]具有局限性。机器学习方法提高了阴影检测与去除的时效性,但监测效果有待提高^[10-11]。深度学习框架下图像阴影去除存在需要大量训练样本^[12]、复杂照明变化任务去除不明显^[13]、黑色物体误检或漏检问题^[14]。当前的主流阴影去除算法,普遍存在模型设计过程中具有局限性以及阴影去除精确度不高的问题。阴影去除算法存在问题可具体总结为两方面:1)对复杂的纹理或与阴影区域相似的暗区域的情况,图像的阴影去除效果不明显;2)半影区域的图像阴影去除效果不明显,去除阴影后的图像留有较小的包围带。为了解决这些问题,提出了空洞卷积与注意力机制融合的对抗式图像阴影去除算法(CDD attention and dilated convolutions of automatic encoding and multiple attention of generative adversarial networks, ADAGAN)。该算法基于生成对抗网络的思想,将空洞卷积与残差网络、注意力编码相结合。用自定义的空洞卷积残差块进行特征提取,提取的特征信息精确性更高,使编码阶段的输入特征信息更加精准。在自动编码阶段,空洞卷积层的加入,代替了池化层,减少了不可逆的信息损失,使编码最后三层的输出包含较大范围的信息。作为临近解码阶段的空洞卷积层,是图像编码阶段全部特征的高度代表,为解码阶段提供更加精确、抽象的特征,即为解码阶段提供更精准的全局的语义特征。在判别网络中,通过多个注意力网络,进行编码的指导,减少了判别网络的误差。该算法提升了去除阴影的准确率,减小了生成的无阴影图像与真实无阴影图像的误差。

1 注意力生成对抗网络

1.1 生成对抗网络

生成对抗网络是蒙特利尔大学的 Goodfellow Ian 于 2014 年提出^[15],由生成器和判别器两部分

构成,生成器的作用是生成可以骗过判别器的图片,判别器的目的是不被生成器生成的图片骗过,生成器与判别器相互制约,共同成长,迭代至生成器生成的图片可以骗过判别器。生成器与判别器的直观感受如图 1 所示。由图 1(a)~(d)显示了生成对抗网络(generative adversarial networks, GAN)训练模型的过程。

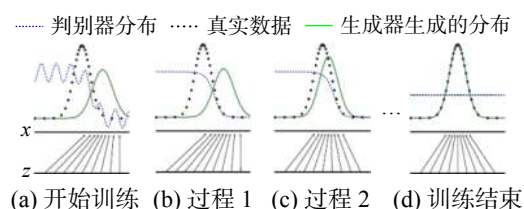


图 1 生成器与判别器对抗图

Fig. 1 Antagonism graph between the generator and discriminator

1.2 注意力机制

注意力机制是在视觉图像领域首先提出的,实际就是对目标数据进行加权变换,扩展了上下文语义信息。把注意力集中放在重要的点上,而忽略其他不重要的因素,而重要程度的判断取决于应用场景。

根据应用场景的不同,注意力机制分为空间注意力机制和时间注意力机制。空间注意力机制一般用于图像处理方面。根据注意力函数的输入序列的位置,分为软注意力机制与硬注意力机制,软注意力机制使用输入序列所有隐藏状态的加权平均值来构建内容向量。软注意力机制、空间注意力机制,形成的二维注意力图的对应权值在 0~1,越重要分配的权值越大。注意力机制可由空洞残差网络与长短期记忆网络(long short term memory, LSTM)组成,可表示为

$$c_t = \varphi(\{c_i\}_{i=1}^M, \{\alpha'_i\}_{i=1}^M) = \sum_{i=1}^M \alpha_i c_i$$

$$\alpha'_i = \frac{\exp(e'_i)}{\sum_{j=1}^M \exp(e'_j)} \quad (1)$$

式中: c_i 表示整体的某一部分; e'_i 表示对应 c_i 在 t 时刻的注意力得分。

长短期记忆网络是 Hochreiter & Schmidhuber 于 1997 年提出的一种特殊的循环神经网络(recurrent neural network, RNN),由遗忘门、输入门和输出门共同组成。

注意力机制的组成及应用过程如图 2 所示。图 2 为注意力机制在生成对抗网络中与生成网络部分结合的构成展示,不仅包含重点关注的内容,也包括周围的环境信息,注意力机制生成注

注意力图指导编码阶段的无阴影图像的生成。同时, 注意力图受阴影模板 (带阴影的图片与不带阴影的真实图片作差) 的影响。具体图片的注意力图如图 3 所示。

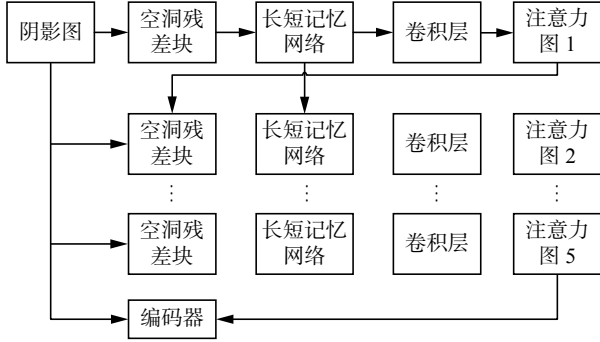


图 2 生成器注意力网络
Fig. 2 Generator attention network

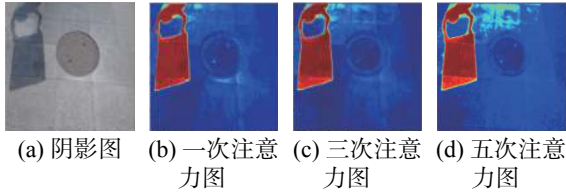


图 3 生成注意力图示例
Fig. 3 Example of generating attention map

1.3 空洞卷积

空洞卷积 (atrous convolutions) 也称扩张卷积 (dilated convolutions), 与普通卷积层相比, 引入了一个称为“扩张率 (dilation rate)”的新参数, 该参数定义了卷积核处理数据时各值的间距^[16]。空洞卷积具有数据结构保留完整和不使用下采样 (down-sampling) 的特性, 优点明显。但多层的空洞卷积也有破坏数据连续性的缺点。

空洞卷积的卷积核是通过在普通的卷积核上填 0 实现的, 增大了网络的扩张系数^[17], 空洞卷积示意如图 4 所示。扩张系数与空洞卷积核的尺寸关系为

$$kd_{size} = (\gamma - 1)(k_{size} - 1) + k_{size} \quad (2)$$

式中: γ 表示空洞卷积的扩张系数; k_{size} 表示普通卷积核尺寸; kd_{size} 表示空洞卷积核尺寸, 当 $\gamma=1$ 时即为普通卷积核。

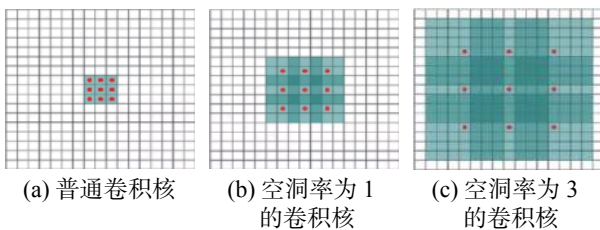


图 4 空洞卷积示意
Fig. 4 Example of dilated convolutions

1.4 损失函数

在生成对抗网络中, 为了达到纳什平衡^[18], 需优化, 见式 (3):

$$\min \max V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_n(z)} [\log (1 - D(G(z)))] \quad (3)$$

式中: G 为生成网络; D 为判别网络; $p_{data}(x)$ 代表真实分布; $p_n(z)$ 代表噪声分布。

注意力机制的损失函数为

$$L_a(\{A\}, M) = \sum_{i=1}^N \theta^{N-i} L_{MSE}(A_i, M) \quad (4)$$

损失的计算是通过比较每次生成的注意力图 (A_i) 与对应图片阴影掩模 (M) 之间的均方误差 (MSE) 进行的。由参考文献 [19-23] 得知, N 取 5, θ 取 0.8 为宜。

编码阶段的损失 (L_e) 由两部分构成, 即图片的真实损失 (L_r) 和模型损失 (L_m)。图片的真实损失为编码生成的图片与真实图片的均方误差, β 取值参考文献 [19]。模型损失为通过 VGG 网络编码生成的图片与真实图片的损失均方误差, 可表示为

$$L_e = \sum_{i=1}^N L_r(O_i, F) + \sum_{i=1}^N L_m(O_i, F)$$

$$L_r(O_i, F) = \beta_i V_{MSE}(O_i, F)$$

$$L_m(O_i, F) = V_{MSE}(VGG(O_i), VGG(F))$$

2 ADAGAN 算法

为了解决图片阴影去除过程中存在的阴影少量遗漏, 半影去除不完全的问题, 提出了 ADAGAN 算法。该算法中 GAN (generative adversarial networks), 即生成对抗网络, 是该算法的主体思想。第 1 个 A (CDD of attention) 指的是生成器中注意力机制特征提取运用 CDD (convolutions and dilated convolutions and dilated convolutions) 空洞残差块; D (dilated convolutions of automatic encoding) 指的是在生成器自编码阶段加入的 4 层相同结构的空洞卷积层; 第 2 个 A (multiple attention) 指的是判别器中的多重注意力网络。

2.1 算法总体架构

ADAGAN 算法运用细节信息提取更精细的 VGG-E 作为该网络的预训练模型。基于生成对抗网络架构, 主体可分为图片阴影的特征提取、注意力编码和判别网络 3 部分。其中, 特征提取与注意力编码属于生成器部分, 判别网络属于判别器部分。

在生成器部分, 对图片阴影信息进行特征提取, 特征提取使用 CDD 残差块, 提取特征作为生成器注意力机制的输入。在生成器的自编码阶

段,加入空洞卷积层。判别器部分由多层卷积注意力 (multiple attention) 网络组成。空洞卷积阴影去除算法整体框架如图5所示。其中,注意力编码部分由注意力生成网络和自动编码网络组成。

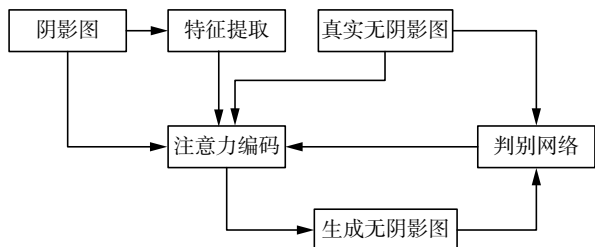


图5 空洞卷积阴影去除算法整体框架图

Fig. 5 Block diagram of the dilated convolution shadow removal algorithm

生成器与判别器部分相互作用,互相影响,形成表现良好的图片去除阴影网络。

2.2 CDD 残差块

为了增加对图片中阴影信息提取的精确度,在网络中引入 CDD 残差块。该方法运用普通卷积与空洞卷积组合的方式,结合残差网络的结构,使生成器注意力机制部分的输入特征更加充分。

空洞卷积能在不增加算法复杂度的情况下,具有更大的感受野,越大的感受野包含越多的上下文关系,但它也会损失信息的连续性,适合大区域的阴影。为了使空洞卷积对小区域的阴影也有很好的特征感知强度,借助残差网络的结构,发现二者的结合,有助于提高特征提取阶段的精确性。残差块对比如图6所示。为了特征提取的表达更加精确细致,改变激活函数 relu 为 lrelu,使值为负数时,也有一定区分,不全为 0。

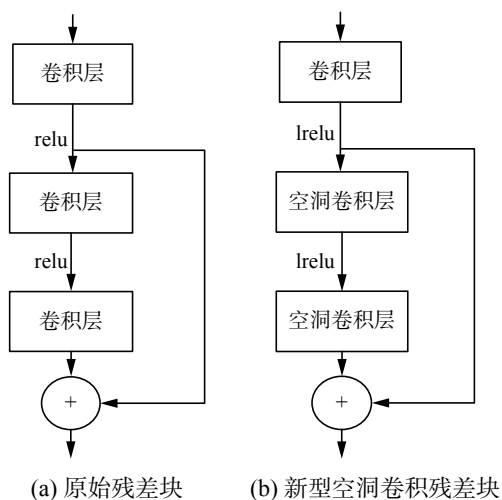


图6 残差块对比

Fig. 6 Residual blocks comparison

2.3 扩展的编码阶段

在编码阶段加入空洞卷积层,使特征表达更

加充分,增大特征的感受范围,减少计算量。该实验在7层卷积后,加入4层同样结构的空洞卷积层,通过多层堆叠的方式增加感受野,为后续的解码过程提取更抽象、更本质的全局的语义特征。

编码阶段可分为编码与解码两个部分,编码由卷积层与空洞卷积完成,解码由反卷积与卷积配合完成。

该实验编码阶段共由11个卷积层,4个空洞卷积层,2个反卷积层以及3个跳跃连接层组成,各层之间的连接关系见表1。其中编码阶段的输入为带阴影的图像与注意力图,每层之间的激活函数为 LRelu, 输出结果为不带阴影的图片。Conv 代表卷积层, Dia_conv 代表空洞卷积层, Deconv 代表反卷积层, Skip 为跳跃连接层。

表1 去除阴影算法编码表

Table 1 Shadow removal algorithm code

名称	输入	输出	卷积核
Conv1	图片+注意力图	Lrelu1	3×3
Conv1_1	Lrelu1	Lrelu1_1	3×3
Conv2	Lrelu1_1	Lrelu2	3×3
Conv3	Lrelu2	Lrelu3	3×3
Conv4	Lrelu3	Lrelu4	3×3
Conv5	Lrelu4	Lrelu5	3×3
Conv6	Lrelu5	Lrelu6	3×3
Dia_conv1	Lrelu6	Lrelu7	3×3
Dia_conv2	Lrelu7	Lrelu8	3×3
Dia_conv3	Lrelu8	Lrelu9	3×3
Dia_conv4	Lrelu9	Lrelu10	3×3
Conv7	Lrelu10	Lrelu11	3×3
Conv8	Lrelu11	Lrelu12	3×3
Deconv_1	Lrelu12	Lrelu13	4×4
Conv_9	Lrelu13+Lrelu3	Lrelu14	3×3
Deconv_2	Lrelu14	Lrelu15	4×4
Conv_10	Lrelu15+Lrelu1	Lrelu16	3×3
Skip_1	Lrelu12+Lrelu10	Skip_1	3×3
Skip_2	Lrelu14	Skip_2	3×3
Skip_3	Lrelu16	Skip_3	3×3

2.4 改进判别网络

卷积网络无法有效地捕捉图像的几何结构和形状,而注意力模型可通过不同的权重系数来强调目标的重要性,抑制无关的细节^[20],且不需要监督。注意力模型的最终目的是帮助类似编解码器这样的框架,更好地学到多种内容模态之间的相互关系,从而更好地表示这些信息,克服其无法解释从而很难设计的缺陷。因此可以灵活地感知到全局与局部的联系,提升网络的感知能力,提

高输出的质量。

注意力机制可分为加法注意力机制和乘法注意力机制, 乘法注意力机制是面对加法注意力机制要求编码与解码的隐藏层长度必须相同的条件的改进, 具有更高的灵活性。因此, 在判别器设计中, 运用乘法注意力机制。而在多次网络卷积结构后, 特征信息逐渐精简的同时, 重点可能会被分散。因此, 在每 2 层卷积层中加入乘法注意力机制, 突出重点, 强调特征, 感知全局, 提升质量。判别器设计如图 7 所示。

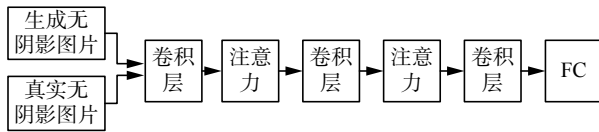


图 7 判别器设计图

Fig. 7 Design of the discriminator

3 实验及结果分析

3.1 实验环境及数据集

实验环境为 ubuntu16.10 系统、GeForce GTX 1080Ti 显卡。数据集选取 ISTD^[21] 和 SRD^[22]。ISTD 数据集共 1870 对阴影与非阴影数据对, 其中训练集 1330 对, 测试集 540 对。SRD 数据集共 3088 对阴影与非阴影数据对, 其中训练集 2680 对, 测试集 408 对。

3.2 实验评价

实验的评价主要从视觉效果和现流行的衡量指标 (结构相似性、峰值信噪比、均方根误差) 进行评价。结构相似性 (structural similarity, SSIM) 基于图像亮度、对比度和结构进行评价, 可表示为

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (5)$$

式中: x 、 y 代表要比较的两张图片; μ_x 为 x 的均值, μ_y 为 y 的均值; σ_x^2 为 x 的方差; σ_y^2 为 y 的方差; σ_{xy} 为 x 和 y 的协方差; $c_1 = (k_1L)^2$ 、 $c_2 = (k_2L)^2$ 为两个常数, 避免除 0; L 为像素范围; 本次实验 k_1 为 0.01; k_2 为 0.03。式 (5) 的值越接近 1, 两图片相似性越强。

峰值信噪比 (peak signal-to-noise ratio, PSNR), 单位是 dB, 数值越大表示效果越好。峰值信噪比基于对应像素点间的误差, 即基于误差敏感的图像质量评价。不考虑人眼的视觉特性, 会出现评价结果与人的主观感觉不一致的情况, 可表示为

$$\text{PSNR} = 10 \lg \left(\frac{\text{MAX}_I^2}{\text{MSE}} \right) \quad (6)$$

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2$$

式中: m 、 n 分别为图像的长与宽; I 为真实图像; K 为生成图像。

均方根误差 (root mean squared error, RMSE) 是一个中间的评价指标, 很多评价指标都是基于均方根误差进行的, 是均方误差 (MSE) 的开根号数, 均方根误差可用来计算阴影图与非阴影图的像素级误差, 可表示为

$$\text{RMSE} = \sqrt{\text{MSE}}$$

3.3 实验结果与分析

实验中针对生成器与判别器的特点, 并参考文献 [23] 来设置学习率参数。其中, 生成器网络学习率为 0.002, 判别器网络学习率为 0.001, 且每进行 10 000 次训练, 学习率缩小 0.1 倍。预训练参数基于 VGG-E 网络。阴影区与非阴影区的像素差的阈值设为 30。

该实验在 ISTD 数据集与 SRD 数据集上进行实验, 经过 100 000 次训练, 实验最佳结果 SSIM 可达到 0.977, PSNR 为 32.2, RMSE 为 6.2。因 ISTD 与 SRD 数据集中各类别图像在数据集中的占比不一致且同一类别中各图像阴影区域的占比也不一致, 因此分别在两个数据集的测试集中选取不同类别。在同一类别中选取一副图片进行测试, 并对它们的值进行展示, 不同数据集效果对比见表 2 (其中, 图片 1~3 为 ISTD 数据集上图片, 图片 4~6 为 SRD 数据集上图片)。

表 2 不同数据集效果对比
Table 2 Comparison of effect of different data sets

评价指标	实验次数					
	1	2	3	4	5	6
SSIM	0.958	0.975	0.914	0.973	0.977	0.967
PSNR	28.88	27.51	30.15	31.32	28.90	32.23
RMSE	9.17	10.73	7.92	6.92	9.15	6.23

该算法分别在 ISTD 和 SRD 数据集上选取 3 个图片进行展示, 可见该算法的阴影去除效果较好, 对相对复杂地物的阴影去除效果也表现良好, 视觉感受上不存在半影情况。不同数据集实验效果如图 8 所示。图 8 共 6 组图片, 前 3 组 (图 (a)~(i)) 为 ISTD 数据集的测试图片, 后 3 组 (图 (j)~(r)) 为 SRD 数据集的测试图片。

实验是基于 GAN 网络与空洞卷积、注意力结合的思想进行的, 分别在特征提取、自编码、判断网络进行巧妙设计。其中特征提取运用空洞残差块, 自编码加入空洞卷积, 判别网络加入多层乘法注意力网络。逐步改进实验效果见表 3。由表 3

可知, CDD 残差块的引用, 虽具有更大的感受野, 但因破坏了信息的连续性, 造成客观衡量指标的降低, 为了兼顾感受野与信息的连续性, 即主观视觉感受与客观衡量指标, 在运用 CDD 残差块的同时引入注重连续性的注意力机制 (LSTM 为核心模块), 解决了特征提取过程增加感受野与关注信息连续性的问题, 提升阴影区域特征提取的精确性。因此, 综合使用 CDD 残差块与注意力机制时, 客观衡量指标有明显的提升。



图 8 不同数据集实验效果

Fig. 8 Experimental renderings of different data sets

表 3 逐步改进实验效果

Table 3 Gradual improve ment of the experimental effect

编号	空洞卷积	CDD块	注意力	SSIM	PSNR	RMSE
1	√			0.910	22.70	18.67
2	√		√	0.911	23.28	17.48
3	√	√		0.909	22.69	18.70
4	√	√	√	0.916	24.42	15.32

表 3 指标为 SRD 训练集上, 20000 次训练后, 随机选取一幅图像 (MG_4699) 进行测试的结果, 逐步改进实验视觉效果如图 9 所示。

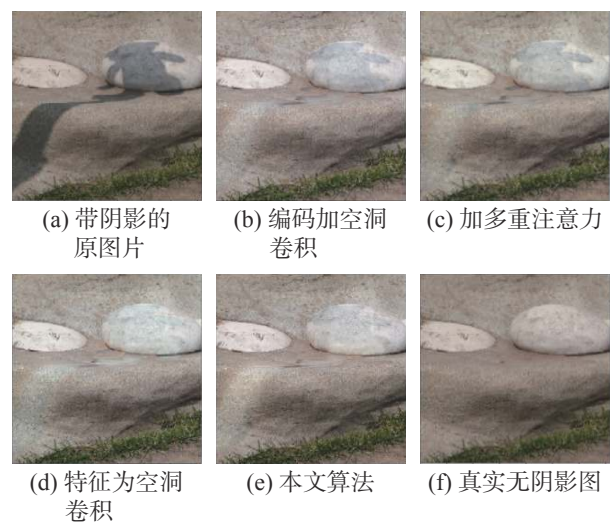


图 9 逐步改进实验视觉效果图

Fig. 9 Gradual improve ment of the visual effect of the experiment

通过图 9 这组实验可以看出: 在判别网络中嵌入多重注意力, 可以提高阴影去除的效果。空洞残差块使特征提取阶段具有更大的感受野, 但同时也破坏了信息的连续性。由表 3 可知, 在加入空洞残差块后, 客观测量指标有所下降。但从图 9 的图像的视觉效果来看, 图像的主观视觉效果比较好。因此, 把两者结合进行实验, 发现实验无论在视觉效果, 还是在图像测试指标上, 都比其他效果要好。

为了证明该算法的有效性, 用该算法与参考文献 [5, 14, 21] 进行对比。文献 [5] 为依靠图像自身特性进行的阴影去除的方法, 按文献 [5] 的思路与参数进行对比实验。文献 [14] 与文献 [21], 均在 ISTD(训练集 1 330 对, 测试集 540 对) 与 SRD(训练集 2 680 对, 测试集 408 对) 数据集上进行训练与测试, 参数配置参考各自原文献。对比实验的方法, 均与本文算法在同实验条件下进行, 视觉效果如图 10、11 所示。图像在 ISTD 数据集

测试指标对比见表 4。

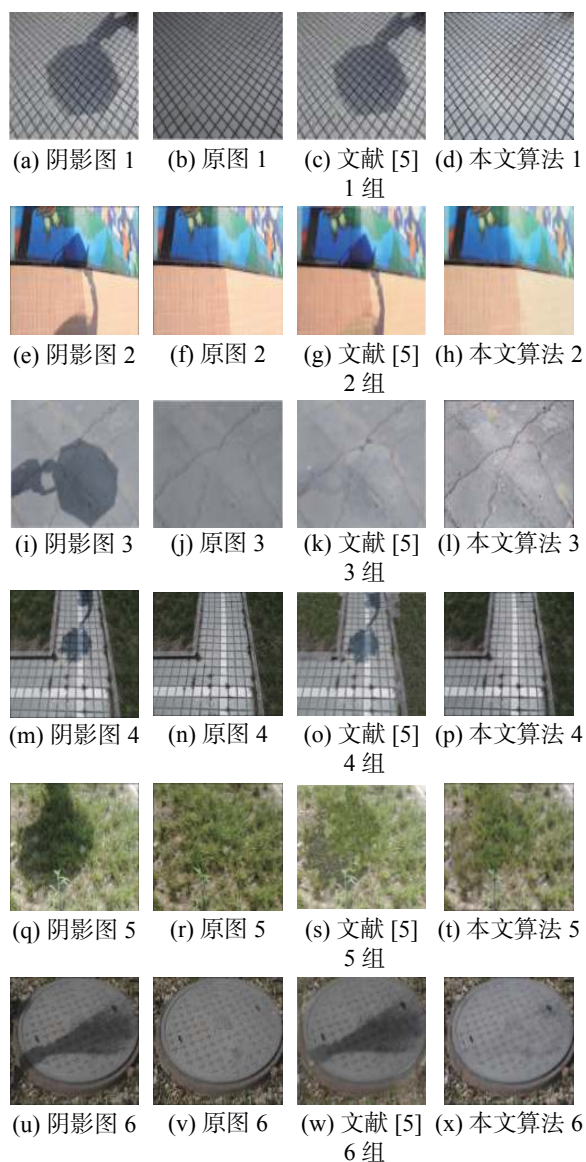


图 10 算法对比视觉效果图 1

Fig. 10 Algorithm comparison visual effect chart 1

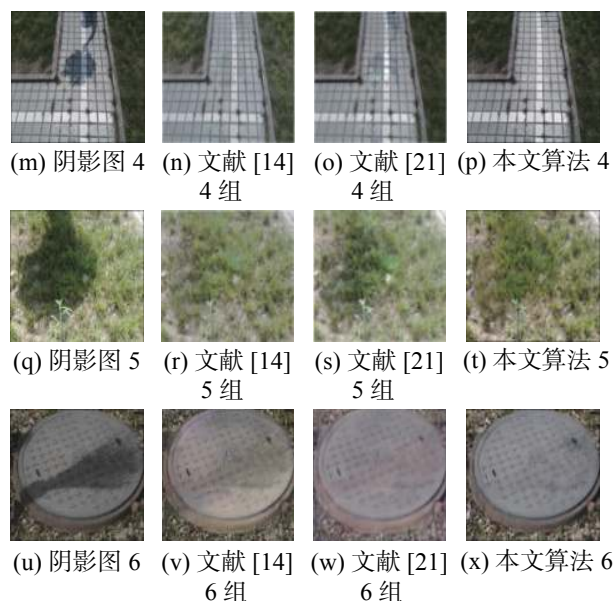
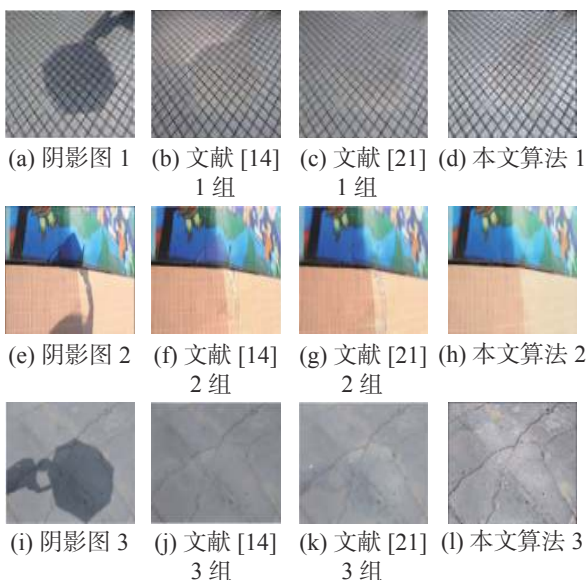


图 11 算法对比视觉效果图 2

Fig. 11 Algorithm comparison visual effect chart 2

表 4 算法测量指标对比

Table 4 Comparison of algorithm measurement indexes

实验	SSIM	PSNR	RMSE
Guo ^[5]	0.934	23.49	14.31
Hu ^[14]	0.961	27.98	8.17
Wang ^[21]	0.949	29.86	7.47
本文算法	0.967	30.12	6.72

4 结束语

本文针对图片阴影去除过程中出现的阴影遗漏现象及半阴影部分去除不全面的问题, 提出空洞卷积与注意力融合的对抗式图像阴影去除算法。该算法用新型的空洞残差块进行特征提取, 加大感受野的同时减少了计算量, 加强了特征感知的强度。编码阶段, 空洞卷积层增加了特征表达的充分性, 具有更为全局的语义信息, 减少计算复杂度。注意力机制的引入加强了网络对全局与部分的把控。本算法的阴影去除效果, 无论从测量指标还是视觉感受, 都达到了比较理想的效果。该算法也可以迁移到其他同类型的监督学习的应用中。

本文提出的算法也存在一些不足之处, 算法对于小规模数据集的效果不够明显, 且生成的图像与原图可能会存在一些细微误差。

参考文献:

- [1] HUANG Xiang, HUA Gang, TUMBLIN J, et al. What

- characterizes a shadow boundary under the sun and sky?[C]//Proceedings of the 2011 International Conference on Computer Vision. Barcelona, Spain, 2011: 898–905.
- [2] FINLAYSON G D, HORDLEY S D, LU Cheng, et al. On the removal of shadows from images[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2006, 28(1): 59–68.
- [3] YU Xiaoming, LI Ge, YING Zhenqiang, et al. A new shadow removal method using color-lines[C]//International Conference on Computer Analysis of Images and Patterns. Ystad, Sweden, 2017: 307–319.
- [4] 林雨准, 张保明, 郭海涛, 等. 结合多尺度分割和形态学运算的高分辨率遥感影像阴影检测 [J]. 中国图象图形学报, 2018, 23(8): 1263–1272.
- LIN Yuzhun, ZHANG Baoming, GUO Haitao, et al. Shadow detection from high resolution remote sensing imagery based on multi-scale segmentation and morphology operation[J]. *Journal of image and graphics*, 2018, 23(8): 1263–1272.
- [5] GUO Ruiqi, DAI Qieyun, HOIEM D. Single-image shadow detection and removal using paired regions[C]//CVPR 2011. Colorado Springs, USA, 2011: 2033–2040.
- [6] FAN Xinyun, WU Wenjun, ZHANG Ling, et al. Shading-aware shadow detection and removal from a single image[J]. *The visual computer*, 2020, 36(10/11/12): 2175–2188.
- [7] 焦俊男, 石静, 田庆久, 等. 多光谱影像的 NDVI 阴影影响去除模型 [J]. 遥感学报, 2020, 24(1): 53–66.
- JIAO Junnan, SHI Jing, TIAN Qingjiu, et al. Research on multispectral-image-based NDVI shadow-effect-eliminating model[J]. *Journal of remote sensing*, 2020, 24(1): 53–66.
- [8] ZHANG Yindan, CHEN Gang, Vukomanovi J, et al. Recurrent Shadow Attention Model (RSAM) for shadow removal in high-resolution urban land-cover mapping[J]. *Remote sensing of environment*, 2020, 247: 111945.
- [9] 吴文, 万毅. 基于低尺度细节恢复的单幅图像阴影去除方法 [J]. 电子学报, 2020, 48(7): 1293–1302.
- WU Wen, WAN Yi. Single image shadow removal using low-scale detail recovering[J]. *Acta electronica sinica*, 2020, 48(7): 1293–1302.
- [10] 张德干, 陈晨, 董悦, 等. 一种基于机器学习的运动目标阴影检测新方法 [J]. 光电子·激光, 2018, 29(12): 1317–1324.
- ZHANG Degan, CHEN Chen, DONG Yue, et al. A new method of moving object shadow detection based on machine learning[J]. *Journal of optoelectronics·laser*, 2018, 29(12): 1317–1324.
- [11] HOSSEINZADEH S, SHAKERI M, ZHANG H. Fast shadow detection from a single image using a patched convolutional neural network[C]//2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Madrid, Spain, 2018: 1–5.
- [12] HU Xiaowei, JIANG Yitong, FU C W, et al. Mask-ShadowGAN: learning to remove shadows from unpaired data[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South), 2019: 2472–2481.
- [13] FAN Hui, HAN Meng, LI Jinjiang. Image shadow removal using end-to-end deep convolutional neural networks [J]. *Applied sciences*, 2019, 9(5): 1009–1026.
- [14] HU Xiaowei, FU C W, ZHU Lei, et al. Direction-aware spatial context features for shadow detection and removal [J]. *IEEE transactions on pattern analysis and machine intelligence*, 2020, 42(11): 2795–2808.
- [15] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada, 2014: 2672–2680.
- [16] 张瑞倩, 邵振峰, PORTNOV A, 等. 多尺度空洞卷积的无人机影像目标检测方法 [J]. 武汉大学学报(信息科学版), 2020, 45(6): 895–903.
- ZHANG Ruiqian, SHAO Zhenfeng, PORTNOV A, et al. Multi-scale dilated convolutional neural network for object detection in UAV images[J]. *Journal of Wuhan University (information science edition)*, 2020, 45(6): 895–903.
- [17] 颜铭靖, 苏喜友. 基于三维空洞卷积残差神经网络的高光谱影像分类方法 [J]. 光学学报, 2020, 40(16): 1628002.
- YAN Mingjing, SU Xiyu. Hyperspectral image classification based on three-dimensional dilated convolutional residual neural network[J]. *Acta optica sinica*, 2020, 40(16): 1628002.

- [18] 贾宇峰, 马力. 条件约束下的自我注意生成对抗网络[J]. 西安电子科技大学学报, 2019, 46(6): 163–170.
- JIA Yufeng, MA Li. Self-attention generative adversarial network with the conditional constraint[J]. Journal of Xidian University, 2019, 46(6): 163–170.
- [19] QIAN Rui, TAN R T, YANG Wenhan, et al. Attentive generative adversarial network for raindrop removal from a single image[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 2482–2491.
- [20] 刘烨, 黄金筱, 马于涛. 基于混合神经网络和注意力机制的软件缺陷自动分派方法[J]. 计算机研究与发展, 2020, 57(3): 459–473.
- LIU Ye, HUANG Jinxiao, MA Yutao. An automatic method using hybrid neural networks and attention mechanism for software bug triaging[J]. Journal of Computer research and development, 2020, 57(3): 459–473.
- [21] WANG Jifeng, LI Xiang, YANG Jian. Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 1788–1797.
- [22] QU Liangqiong, TIAN Jiandong, HE Shengfeng, et al. DeshadowNet: a multi-context embedding deep network for shadow removal[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017: 2308–2316.
- [23] DING Bin, LONG Chengjiang, ZHANG Ling, et al. AR-GAN: attentive recurrent generative adversarial network for shadow detection and removal[C]//Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South), 2019: 10212–10221.

作者简介:



刘万军, 教授, 博士生导师, 主要研究方向为图像与智能信息处理。获国家发明专利 5 项, 出版专著教材 6 部, 发表学术论文 130 余篇。



佟畅, 硕士研究生, 主要研究方向为图像与智能信息处理。



曲海成, 副教授, 博士, 主要研究方向为图像与智能信息处理。