

基于预处理图像惩罚的生成对抗网络水下图像增强

宋巍*, 邢晶晶, 杜艳玲, 贺琪**

上海海洋大学信息学院, 上海 201306

摘要 针对水下图像存在的对比度低、细节模糊、色彩失真问题,提出了一种基于预处理图像惩罚的生成对抗网络(GAN)水下图像增强方法。首先,通过改进的红色通道直方图拉伸算法对水下图像进行预处理,改善图像对比度的同时避免传统直方图拉伸后的局部过增强现象。然后,构建带有预处理图像惩罚的GAN,实现水下图像增强。其中,生成器编码-解码结构中的前两层使用多尺度卷积,以增强网络对细节信息的学习能力。最后,构建多项损失函数,将预处理图像作为伪真值对GAN施加损失惩罚,以提升网络的泛化性能。实验结果表明,相比传统图像增强方法和基于深度学习的图像增强方法,本方法在水下图像的色偏、对比度和细节信息方面的表现更优,且鲁棒性更好。

关键词 图像处理;直方图拉伸;生成对抗网络;惩罚损失;水下图像增强

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP202158.1210024

Underwater Image Enhancement Based on Generative Adversarial Network with Preprocessed Image Penalty

Song Wei*, Xing Jingjing, Du Yanling, He Qi**

Department of Information and Technology, Shanghai Ocean University, Shanghai 201306, China

Abstract Aiming at mitigating the problems of low contrast, blurred details, and color distortion in underwater images, an underwater image enhancement method based on preprocessed image penalty and generative adversarial network (GAN) is proposed in this paper. First, an improved red channel histogram stretching algorithm is used to preprocess the input underwater image to improve the image contrast and avoid over enhancement of local blocks after traditional histogram stretching. Then, GAN with preprocessed image penalty is designed to realize underwater image enhancement. Moreover, multiscale convolution is used for the first two layers of the generator coding-decoding structure to enhance the detailed information learning ability of the network. Finally, a multiloss function is established in which the preprocessed image is used as a false truth value to impose loss penalty on GAN to improve generalization performance of the network. Experimental results show that compared with traditional image enhancement methods and deep learning-based image enhancement methods, the method performs better in terms of color deviation, contrast, and detailed information of underwater images, and has better robustness.

Key words image processing; histogram stretching; generative adversarial network; penalty loss; underwater image enhancement

OCIS codes 100.2000; 100.2980; 100.1830

1 引言

海洋中蕴藏着丰富的物种和资源,水下图像作

为海洋工程和科学研究中的信息载体,是人们的重要研究对象,在海洋军事、海洋环境保护^[1]、海洋资源探索^[2]方面具有重要意义。由于水下环境和光照

收稿日期: 2020-09-07; 修回日期: 2020-10-02; 录用日期: 2020-11-14

基金项目: 国家自然科学基金(61702323, 61972240, 41906179)

* E-mail: wsong@shou.edu.cn; ** E-mail: qihe@shou.edu.cn

条件的复杂性,水下图像增强仍然面临着巨大挑战。为了解决水下图像质量退化的问题,人们提出了多种图像增强方法,如直方图均衡化(HE)^[3]、对比度限制的自适应 HE(CLAHE)^[4]、伽马校正(Gamma correction)通过提高高光图像的全局能见度改善图像质量,灰度世界假设(GWA)、白平衡(WB)和灰度边缘假设(GEA)通过修正图像的颜色和饱和度改善图像质量。Huang 等^[5]根据图像在 RGB (Red, Green, Blue) 和国际照明委员会(CIE)公布的 CIE-Lab(Luminosity, a, b)颜色空间的像素值分布特点,提出了相对直方图拉伸(RGHS)算法。Ancuti 等^[6]提出了一种基于融合(Fusion)的方法,将基于原始图像生成的两幅增强图像进行特征融合。图像复原方法是结合成像原理和先验知识进行单图像恢复,如 He 等^[7]通过大量统计实验提出的暗通道先验(DCP)去雾算法,Drewns 等^[8]基于光在水中传播时不同波长的衰减特性提出的水下 DCP(UDCP)去雾算法,Peng 等^[9]基于模糊先验(BP)估计场景深度,完成图像恢复。Peng 等^[10]进一步改进 BP 算法,基于图像模糊和光吸收(IBLA)对复杂场景下的图像进行恢复。

近年来,随着深度学习^[11]在计算机视觉、自然语言处理等多个领域^[12]的发展,人们开始将深度学习方法应用到水下图像恢复任务中,生成对抗网络(GAN)^[13]的成功打开了水下图像处理领域的新大门。Zhu 等^[14]提出了一种弱监督模型 CycleGAN,用风格转换网络实现了两个不同图像域之间的转换。Fabbri 等^[15]利用 CycleGAN 将无失真的水下图像渲染成质量差、分辨率较低的模糊图像,获得成对的水下图像集,并提出了一种基于像素到像素(Pix2Pix)网络^[16]的水下图像 GAN(UGAN)。Pix2Pix 网络以条件约束 GAN^[17]为基础,允许生成器生成遵循某种模式或属于某个特定类的样本,并构建了一个关注图像块(Patch)的全卷积判别器。Lu 等^[18]结合 DCP^[7]和 CycleGAN^[14]设计了一个自适应水下图像恢复系统,利用 DCP 获取水下图像的深度信息,进一步通过多尺度计算得到增强图像。Chen 等^[19]提出了一种基于滤波和 GAN 的图像恢复方法,在傅里叶域对图像进行恢复。Liu 等^[20]提出了水下极深残差网络(Uresnet),在水下图像增强中引入较深的超分辨率(VDSR)模型^[21],并用异步训练模式改善多项损失函数的模型训练。

本文结合传统图像增强方法和深度学习方法的优点,提出了一种结合预处理图像惩罚的 GAN 水

下图像增强方法。首先,基于红色(R)通道改进的直方图拉伸算法进行图像处理。其次,构建并训练带有预处理图像惩罚的 GAN。网络模型的生成器采用编码-解码结构,该结构的前两层网络使用多尺度卷积,以提取不同的细节信息,同时加入跳跃连接,避免底层信息的丢失;判别器采用 PatchGAN 结构,在减少网络参数的同时增强判别结果的准确性。最后,构建多项损失函数,将预处理图像作为仿真值对 GAN 模型施加惩罚,使网络具有更好的泛化能力。实验结果表明,相比其他图像增强方法,本方法在真实水下图像数据集和合成水下图像数据集上均能有效改善图像的色偏,提升水下图像的质量。

2 算法原理

2.1 算法流程

本算法以 GAN 为基础架构,学习失真图像到无失真图像之间的映射函数,算法框架主要包括三部分,如图 1 所示。第一部分为图像预处理,对原始图像进行自适应直方图拉伸,得到对比度高、色彩平衡的预处理图像。第二部分为 GAN 的训练,包括生成器(Generator)和判别器(Discriminator),生成器用多尺度卷积获取不同的感受野,以更好地学习原始图像和参考图像之间的映射函数;结合 L_1 损失和带有梯度惩罚的 Wasserstein GAN(WGAN-GP)损失,能保证生成器生成图像的结构与原始图像一致。第三部分为惩罚损失的设计,用预处理图像作为仿真值对网络进行惩罚,用预训练的极深卷积神经网络 VGG(Visual geometry group)^[22]提取的特征作为惩罚损失。整个网络中的损失函数引导生成器不断更新参数,以学习两个图像域之间的映射函数,最终得到增强图像。

2.2 预处理模型

水下图像的对比度和能见度较低,像素值分布相对集中,不符合人类视觉感知对图像质量的要求。因此,采用改进的直方图拉伸算法对原始低质量图像进行预处理,以自适应弥补全局直方图拉伸增强后图像出现的过饱和现象。直方图拉伸的过程可表示为

$$I_{pre} = (I_r - I_{min}) \left(\frac{O_{max} - O_{min}}{I_{max} - I_{min}} \right), \quad (1)$$

式中, I_r 为输入图像像素的强度, I_{pre} 为经直方图拉伸预处理后输出图像像素的强度, I_{max} 和 I_{min} 、 O_{max} 和 O_{min} 分别为输入图像和理想输出图像像素的最大和最小强度。在全局直方图拉伸中, O_{max} 和 O_{min}

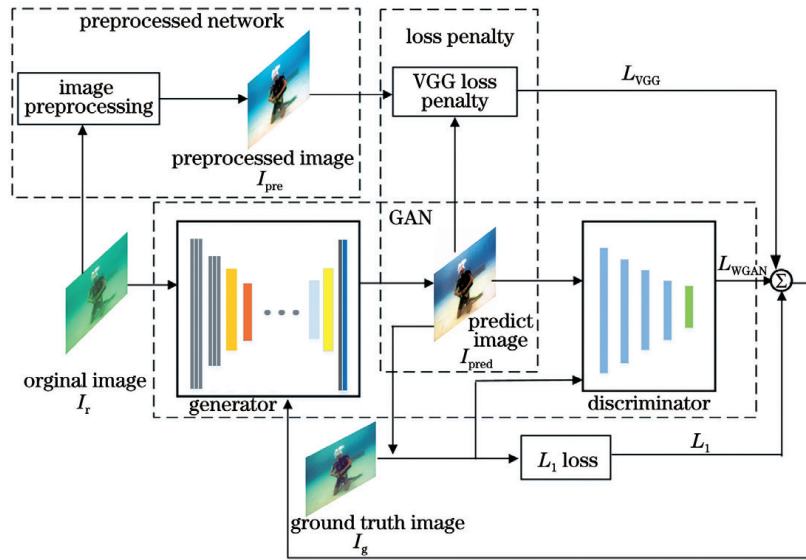


图 1 本算法的框架

Fig. 1 Framework of our algorithm

为常量,分别为 255 和 0。

为避免噪声带来的影响,将每个通道的像素值按升序排列,并对像素值进行截断处理,去除前、后 0.5% (较小和较大) 的像素值,然后确定 (I_{\min} , I_{\max})。借鉴文献[6]中的白平衡处理方法,根据三个通道(R,G,B)像素值的比例关系,动态适应每个通道的最大值和最小值取值范围。对于通道 $c \in (R, G, B)$,其最小值和最大值对应的像素点位置为 (P_{\min}^c, P_{\max}^c),可表示为

$$R^c = \frac{\max(\mu_R, \mu_G, \mu_B)}{\mu_c}, \quad (2)$$

$$P_{\min}^c = 0.005R^c, P_{\max}^c = 1 - P_{\min}^c, \quad (3)$$

式中, μ_c 为通道 c 的像素均值, R^c 为截断比。在 RGB 三通道中,总会有一个通道的像素值均值最大。由(2)式可知,截断比最大可能取值为 1,因此,(3)式可将前、后截断的像素占比限制在总像素数量的 0.5% 以内。对图像像素值进行截断处理可以避免异常值对图像恢复的影响,避免丢失更多信息。

对原始图像进行预处理时,充分考虑了水下图像 RGB 三通道的像素值分布。水下图像以蓝绿色调为主,且 B、G 通道的像素值分布相对均匀,可直接根据确定的参数将其拉伸到全像素空间[0, 255]中。而红光在水下的衰减远大于绿色和蓝色光线,R 通道像素值通常集中分布在直方图左侧偏小的区域,将其拉伸到全像素空间易造成过增强现象。为了抑制过度拉伸,根据 R 通道的均值 μ_R 和方差 σ_R 设置阈值,以限制 R 通道均衡化后的 O_{\max}^R ,设置的 R 通道阈值范围可表示为

$$O_{\max}^R = 3\sigma_R + \mu_R, s.t. \sigma_R < 25. \quad (4)$$

可以发现,像素值的均值和方差分别反映了像素值的集中分布位置和分布范围,统计拉伸后过增强水下图像的像素值发现,R 通道像素值的分布非常集中,需要限制其拉伸的最大范围 O_{\max}^R 。当 $\sigma_R < 25$ 时,(4)式在不同场景和不同损失程度的图像中都能很好地避免图像过增强。方差较大时,表明该通道的像素分布范围较大,其拉伸范围的确定与 G 和 B 通道一致。

图 2 为用传统直方图拉伸和改进 R 通道的直方图拉伸后的图像,可以发现,相比原始图像,经传统直方图拉伸后的图像对比度有明显提升,但也存在严重的过增强现象;且放大了图像的噪声,丢失了部分细节信息。而自适应直方图拉伸算法能简单地提升原始图像的对比度,同时增强了图像的边缘细节。

2.3 生成器模型

生成器采用基于 U-Net^[23] 的编码-解码结构恢复水下图像的细节、对比度和色彩,网络结构如图 3(a)所示。输入 RGB 图像的尺寸为 256 pixel × 256 pixel,在编码阶段,生成器的前两层用卷积块对输入图像进行多尺度^[24-25]特征提取,卷积块由大小为 7 × 7、5 × 5 和 3 × 3 的卷积核构成,具体结构如图 3(b)所示;之后的每层网络采用卷积(Conv)、带泄露修正线性单元(LReLU)激活和批归一化(BN)提取特征。解码阶段不添加 BN 层,每层网络直接用卷积和 LReLU,最后一层网络用 tanh 函数进行非线性激活,同时加入跳跃连接减少特征提取过程

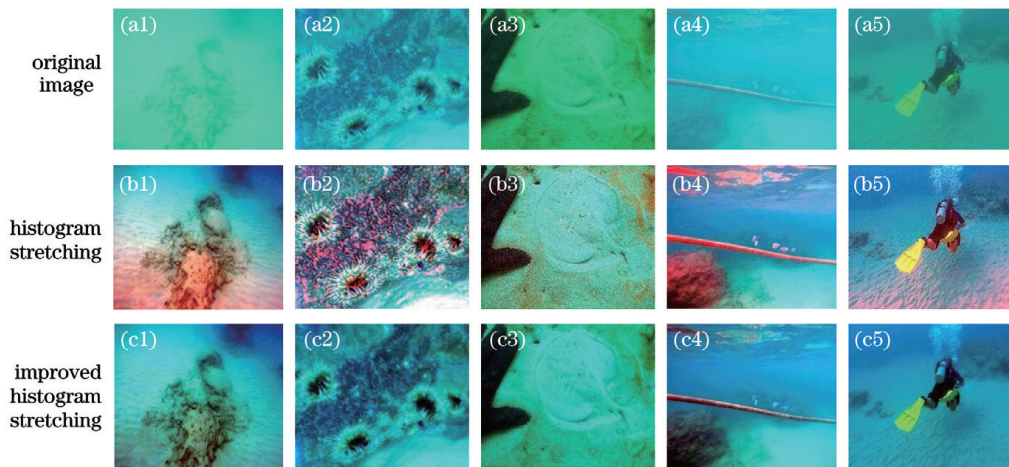


图 2 不同拉伸算法处理后的图像。(a)原始图像;(b)传统直方图拉伸算法处理后的图像;(c)改进直方图拉伸算法处理后的图像

Fig. 2 Images processed by different stretching algorithms. (a) Original image; (b) image processed by traditional histogram stretching algorithm; (c) image processed by improved histogram stretching algorithm improved histogram stretching algorithm

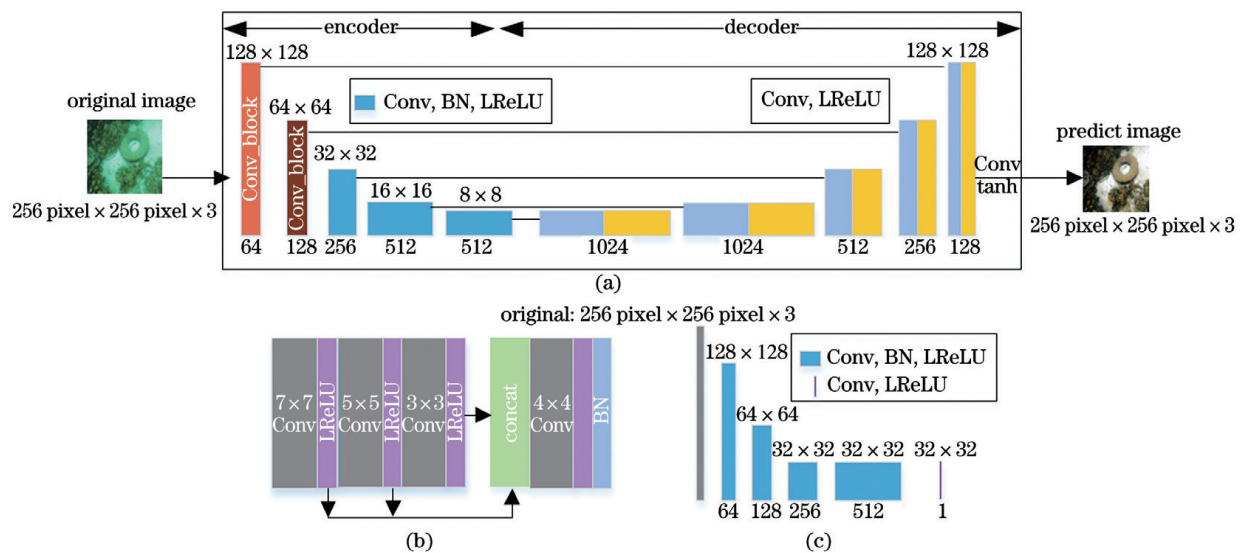


图 3 GAN 的结构。(a)生成器;(b)卷积块;(c)判别器

Fig. 3 Structure of the GAN. (a) Generator; (b) convolution block; (c) discriminator

中的信息丢失问题。除编码阶段的前两层外,生成器网络中其他各层网络均采用大小为 4×4 的卷积核,每层卷积核的数量以 2 的指数形式从 64 逐渐增加到 512;解码过程中使用卷积和步长代替最大池化进行上采样,最终生成器网络输出尺寸为 $256 \text{ pixel} \times 256 \text{ pixel} \times 3$ 的增强图像。

2.4 判别器模型

判别器采用 Markovian PatchGAN^[26] 架构,只根据图像块信息对真实图像和网络生成的图像进行区分。由于图像块计算的参数比全局图像判别要少,因此计算效率更高。如图 3(c)中用 5 个卷积层将尺寸为 $256 \text{ pixel} \times 256 \text{ pixel} \times 3$ 的真实图像和生

成图像转换成尺寸为 $32 \times 32 \times 1$ 的块输出,每层卷积核的尺寸为 4×4 ,步长为 2,采用 LReLU 和 BN 进行非线性激活和批归一化。

2.5 损失函数

通过制定一个目标函数引导生成器 G 学习增强图像,判别器 D 则丢弃生成器 G 生成的与参考图像局部纹理和风格不一致的图像。构建的总损失函数可表示为

$$L_{\text{Loss}} = G \min_{\max} DL_{\text{WGAN-GP}}(G, D) + \lambda_1 L_1 + \lambda_2 L_{\text{VGG}}(G), \quad (5)$$

式中, $L_{\text{WGAN-GP}}(G, D)$ 为带有梯度惩罚的 GAN 损

失, L_1 为生成图像和真实图像的结构损失, $L_{VGG}(G)$ 为预处理图像和生成图像之间的感知损失。可以发现, 损失函数包括三项, 考虑到传统 GAN 训练时的不稳定性, 本方法用带有梯度惩罚的 WGAN-GP^[27] 作为 GAN 的损失函数, 用 L_1 损失最小化生成图像与真实图像之间的误差, 将预处理图像作为伪真值, 以降低感知损失。

1) 对抗损失

在深度学习文献中, GAN 为一类基于博弈论的生成模型, 包括生成器和判别器两部分。生成器 G 的目标是尽可能地生成与真实图像一致的图像以欺骗判别器 D , 而判别器 D 的目标就是尽量区分出生成器 G 生成的图像和真实图像。生成器和判别器的损失可表示为

$$\begin{aligned} \max_D V(G, D) &= E_{I_g \sim p_{\text{train}}(I_g)} [\log D(I_g)] + \\ &E_{I_r \sim p_{\text{gen}}(I_r)} \{\log \{1 - D[G(I_r)]\}\}, \quad (6) \\ \min_G V(G, D) &= E_{I_r \sim p_{\text{gen}}(I_r)} \{\log \{1 - D[G(I_r)]\}\}, \quad (7) \end{aligned}$$

式中, I_g 为无失真图像(真值), I_r 为失真的原始图像, E 为分布函数的期望值, $p_{\text{train}}(I_g)$ 为真实的样本分布, $p_{\text{gen}}(I_r)$ 为原始图像的分布, $G(I_r)$ 为生成器生成的图像, $V(G, D)$ 为目标函数。模型训练时, 交替更新生成器和判别器的参数, 最终达到动态平衡, 得到增强图像。

实际中 GAN 的训练存在收敛速度慢、梯度消失和梯度爆炸等问题。为了解决上述问题, 文献[28]提出 Wasserstein GAN(WGAN), 使用 Wasserstein 距离改进 GAN 的损失函数, 通过 Lipschitz 限制保证判别器的所有参数有界, 以近似 Wasserstein 距离, 但该方法容易导致参数走向极端。因此, 实验引用 WGAN-GP^[27] 损失训练网络模型, 在处理 Lipschitz 限制条件时设置了一个额外的梯度惩罚损失项。WGAN-GP 损失可表示为

$$\begin{aligned} L_{\text{WGAN-GP}}(G, D) &= E[D(I_g)] - E\{D[G(I_r)]\} + \\ &\lambda_{\text{GP}} E_{x \sim P_x} \left\{ \left[\|\nabla_x D(x)\|_2 - 1 \right]^2 \right\}, \quad (8) \end{aligned}$$

式中, P_x 为真实数据分布和生成数据分布之间的点对点直线样本, λ_{GP} 为权重因子, $E_{x \sim P_x}$ 为样本分布, ∇_x 为梯度。

2) L_1 损失

相比 L_2 均方误差损失, L_1 绝对误差损失受异常值的影响较小。对于由算法生成的水下真实图像, 存在潜在的异常像素值, 因此, 引入 L_1 损失, 使生成图像和真实图像尽可能相似。生成图像和真值

图像之间的 L_1 损失可表示为

$$L_1 = E(I_g, I_r) [\|I_g - G(I_r)\|_1]. \quad (9)$$

3) 感知损失

为了增强网络模型训练的泛化性能, 将预处理后的图像作为伪真值对网络施加惩罚, 计算生成图像 I_{pred} 和预处理图像 I_{pre} 之间的感知损失, 可表示为

$$\begin{aligned} L_{\text{VGG}} &= \frac{1}{CWH} \sum_{c=1}^C \sum_{w=1}^W \sum_{h=1}^H \|\Phi[G(I_{\text{pred}})^{c,w,h}] - \\ &\Phi(I_{\text{pre}})^{c,w,h}\|_2, \quad (10) \end{aligned}$$

式中, C, W, H 分别为特征图的通道、高度和宽度, $\Phi(x)$ 为 ImageNet 数据集^[29] 上预训练的 VGG19 网络^[22] Block5_conv2 层提取的特征映射。

3 实验结果与分析

3.1 数据集

实验使用公开数据集 EUVP^[30] 中成对的水下图像进行网络训练。该数据集来源于 7 个不同的摄像机, 包括多个 GoPros、uEye 摄像机、弱光 USB 摄像机和高清摄像机, 且包含不同位置、不同能见度条件下的水下图像。此外, 该数据集还包括一些公开的 YouTubeTM 视频网站中截取的图像, 以保证数据能体现自然变化的多样性(如场景、水体类型、光照条件)。从主观上检查了图像的色彩、对比度和清晰度等属性, 并考虑了场景是否具有视觉上的可解释性, 最终选取包含 8000 张非成对水下图像的数据集。同时, 严格按照文献[15]中的数据制作流程, 获得包含 12000 张图像质量较好和较差的成对水下图像数据集。利用该数据集进一步训练了一个快速水下图像增强网络 FunieGAN。

为保证实验的一致性, 所有基于深度学习的对比网络均采用 EUVP 成对数据集进行训练。为检测模型的泛化性能, 除了在 EUVP 测试集上进行性能评价外, 还在 Li 等^[31-32] 提出的真实水下图像数据集 UIEBD 和合成水下图像数据集上进行了测试与评价。UIEBD 数据集中真值水下图像的制作过程: 首先, 用 12 种不同图像增强方法对不同场景、不同失真程度的原始水下图像进行增强; 然后, 由具有一定图像处理背景的志愿者对增强后的水下图像进行主观打分; 最后, 从 12 种增强图像中选出评分最高的图像作为原始水下图像的参考真值图像, 将基于 UIEBD 数据集构建的水下图像增强网络称为 Water-Net^[31]。水下合成数据集^[32] 是以大气环境中拍摄的图像为真值图像, 依据水下光学成像模型

及光在海域和近海岸的不同衰减系数进行处理,得到不同失真程度的水下图像数据集。

3.2 实验设置

实验使用的网络代码在 Tensorflow 框架中实现,计算机的 GPU 为 NVIDIA GeForce RTX 2080 Ti,初始学习率为 0.001,用 Adam 算法对网络参数

进行优化,批大小为 4,迭代周期为 30。模型训练时,生成器每更新 1 次参数判别器更新 5 次。当 L_1 损失在图像归一化范围内计算时,为保证多项损失计算后的结果在同等量级,取损失权重 λ_1 、 λ_2 分别为 100、1。

3.3 主观评价

图4为本方法与几种传统图像增强方法HE^[3]、

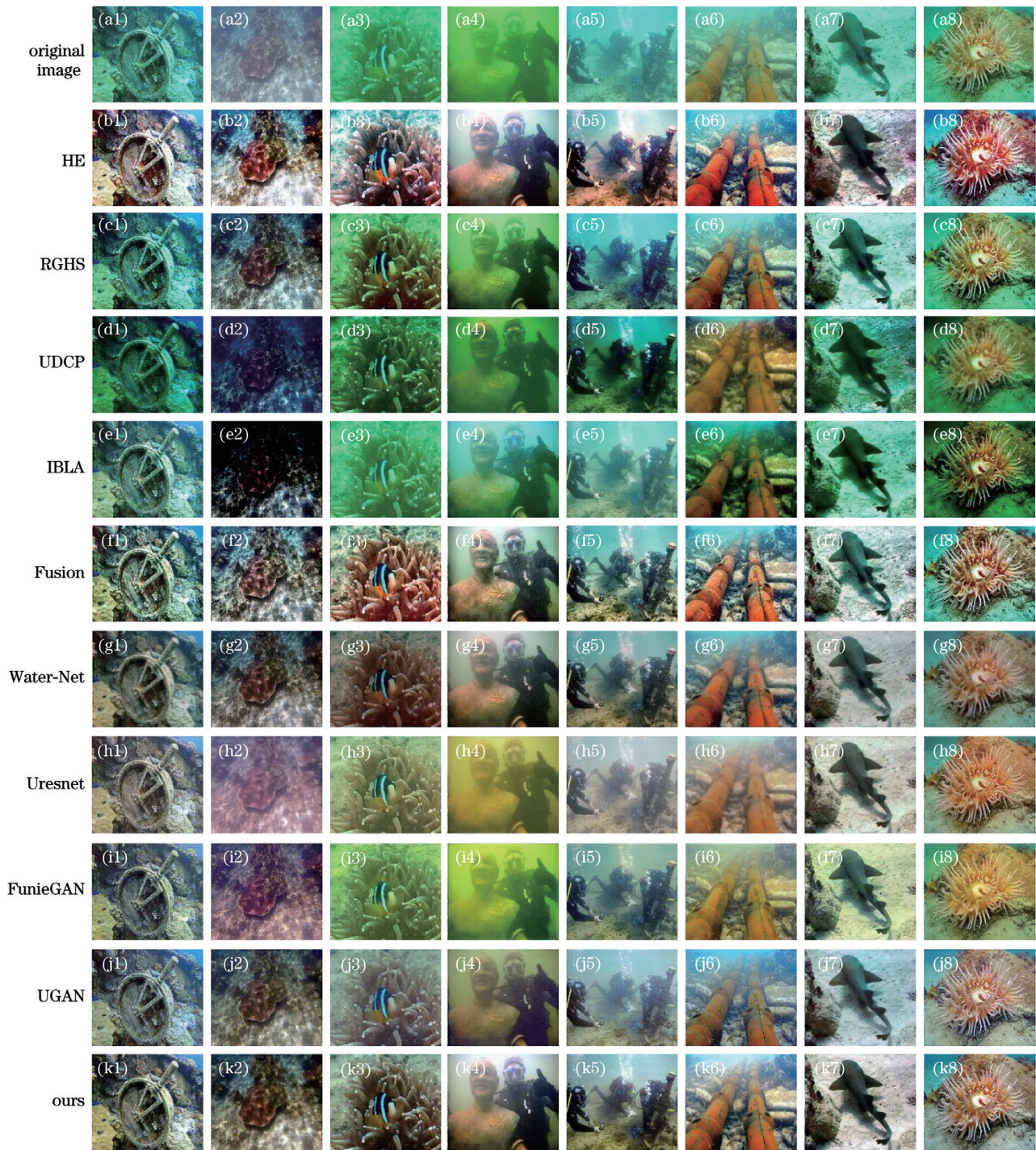


图 4 不同方法对水下图像的增强结果。(a)原始图像;(b) HE;(c) RGHS;(d) UDCP;(e) IBLA;(f) Fusion;(g) Water-Net;(h) Uresnet;(i) FunieGAN;(j) UGAN;(k) ours

Fig. 4 Enhancement results of different methods on underwater images. (a) Original image; (b) HE; (c) RGHS; (d) UDCP; (e) IBLA; (f) Fusion; (g) Water-Net; (h) Uresnet; (i) FunieGAN; (j) UGAN; (k) ours

RGHS^[5]、Fusion^[6]、UDCP^[8]、IBLA^[10]和几种较新的深度学习方法 Water-Net^[31]、Uresnet^[20]、FunieGAN^[30]、UGAN^[15]对水下图像的增强结果。在复现基于深度学习的图像增强方法时,用 EUVP 数据集进行训练和测试。可以发现,HE 方法增强的图像呈压倒性的红色色调,明显出现过饱和现象,且放大了原始图像的噪声,但很好地改善了图像的色彩偏移问题。RGHS 方法一定程度上增强了图像的对比度,但对于色偏严重的图像仍然表现出很大的局限性。IBLA 和 UDCP 方法增强后的图像色调比较相近,整体颜色偏暗,对于对比度较低的图像,增强后的整体亮度降低,很难识别图像中的物体;且引入了新的色调,加重了图像色彩失衡现象,不符合人类的视觉感知。相比上述四种方法,Fusion 方法在对比度、亮度方面有很大提高,但会出现细节丢失问题,且增强后的图像边缘过于突出。Water-Net 方法增强后的图像色调一致性较好,但不能真正去除色偏,细节模糊。Uresnet 方法可去除色偏,对于前景和背景区分明显的图像表现出很好的增强效果,但对于场景复杂的图像,增强后出现模糊,去雾效果较差。FunieGAN 方法处理色偏图像时,表现出较大的限制,会引入新的色偏,同时增

大噪声,在处理光照条件不均匀的水下图像时,容易过度增强,放大物体的边缘信息。UGAN 方法利用风格转移进行图像增强,整体恢复效果较好。而本方法在去除色偏的同时,提高了水下图像的对比度和清晰度,且在不同场景下都表现出了很好的增强效果。

为了验证本方法的泛化性能,利用真实数据集训练的模型对合成数据集^[32]进行评价,结果如图 5 所示。其中,图 5(a1)~图 5(a5)为合成的深海失真图像,图 5(a6)和图 5(a7)为合成的近海岸图像,两类合成图像的失真程度逐渐加剧。对比本方法与 Water-Net、Uresnet、FunieGAN、UGAN 深度学习方法的结果可以发现,Water-Net 方法处理失真程度较小的图像时效果较好,如图 5(b1)所示;对于损失程度较大的图像,增强后出现模糊现象,整体亮度有所降低,如图 5(b5)和图 5(b7)所示。Uresnet 方法处理严重失真的图像时,会加剧图像本身的色调,对比度略有提升,出现局部区域过增强现象,如图 5(c5)和图 5(c7)所示。FunieGAN 方法的整体恢复图像呈橘色调,不符合人类的视觉主观感受。UGAN 方法恢复的图像结果整体较好,而本方法恢复的图像清晰度较高,且保持了原始图像的真实性,在处理不同损失程度图像时的恢复能力均较强。

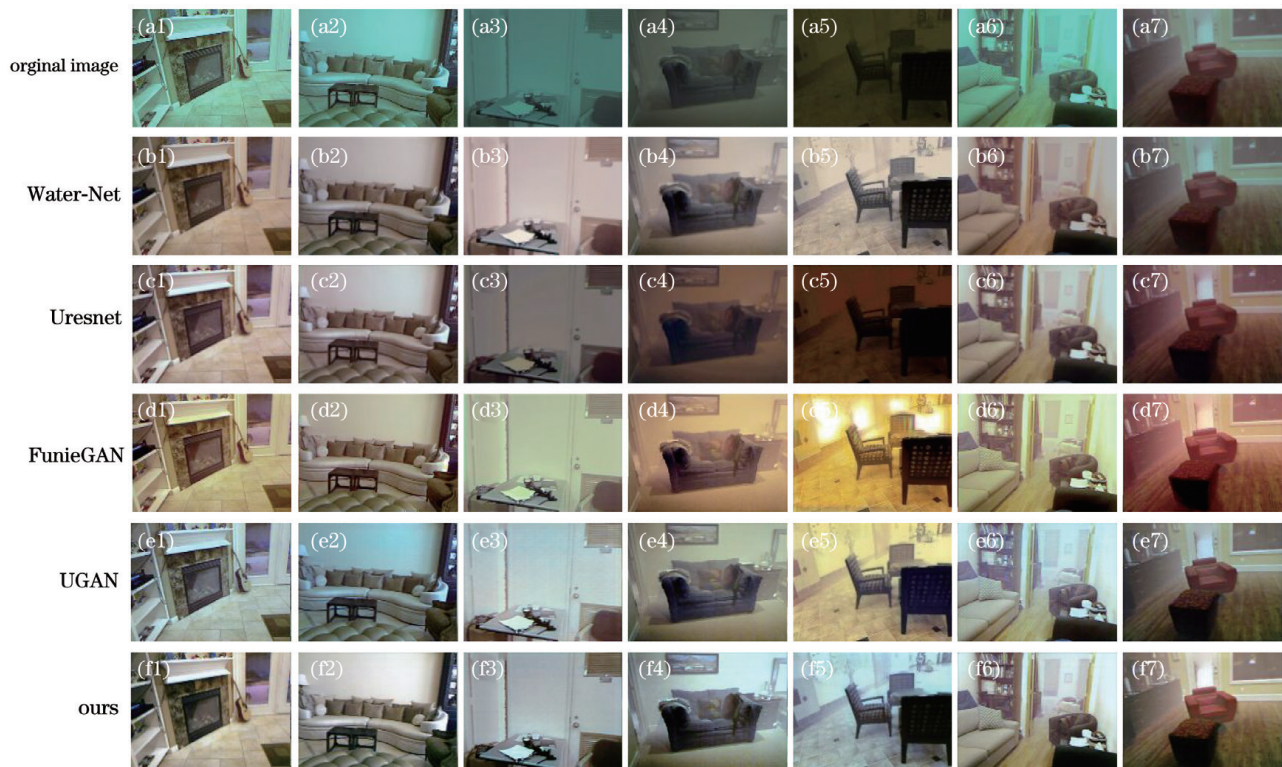


图 5 不同方法在合成数据集上的测试结果。(a)原始图像;(b) Water-Net;(c) Uresnet;(d) FunieGAN;(e) UGAN;(f) ours

Fig. 5 Test results of different methods on the synthetic data set. (a) Original image; (b) Water-Net; (c) Uresnet; (d) FunieGAN; (e) UGAN; (f) ours

为了验证本方法对图像细节的保持情况,将深度学习方法中表现较好且同是基于 GAN 的 FunieGAN 和 UGAN 方法进行对比实验,结果如图 6 所示。图 6(b)、图 6(c)、图 6(d)分别放大了图 6(a)中实线框、点线框和虚线框的图像块。对比发现, FunieGAN 增强后的图像在实线框位置出现明显的条纹,在点线框位置出现失真,部分块的细节直接丢失。UGAN 方法在细节保持上较好,但对图像进行色彩修正时,虚线框位置与原始图像的色彩偏差略大且整体表现相对模糊。相比上述两种方法,本方法在校正图像色偏的同时可以保持图像细节信息,提高图像的清晰度。

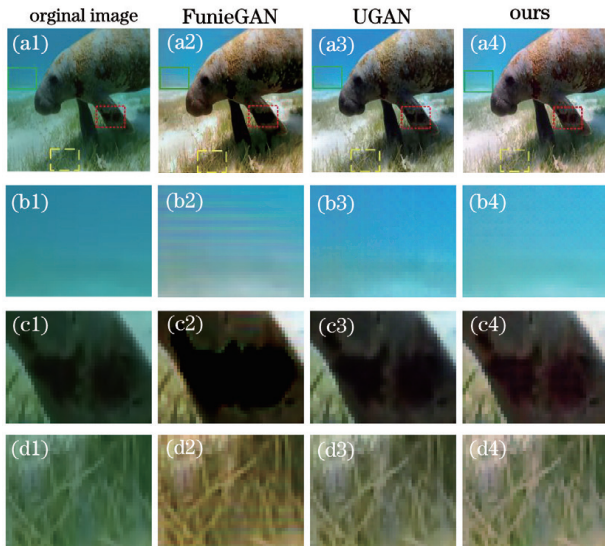


图 6 不同方法的细节保留能力。(a)原始图像;(b)实线框的细节;(c)点线框的细节;(d)虚线框的细节

Fig. 6 Detail retention capabilities of different methods. (a) Original image; (b) details of the solid line frame; (c) details of the dotted line frame; (d) details of the dotted frame

3.4 客观评价

为了客观评价和分析本方法的性能,选取结构相似性(SSIM)和峰值信噪比(PSNR)对图像的增强效果进行定量评价,同时引入水下图像质量评价指标 UIQM^[33]度量生成图像的色彩(UICM)、锐度(UISM)、对比度(UIConM)。由于自然图像是高度结构化的,增强后的图像与真值图像存在很强的结构相关性,因此,用 SSIM 对生成图像的质量进行客观评价。PSNR 是用来衡量信号失真的指标,通常情况下,PSNR 越大,表明被评价图像越接近真实图像。水下图像质量评价指标 UIQM 将色彩、锐度和对比度线性组合在一起,每个属性用于评估水下图像退化的一个方面,可表示为

$$X_{UIQM} = c_1 X_{UICM} + c_2 X_{UISM} + c_3 X_{UIConM}, \quad (11)$$

式中, c_1 、 c_2 、 c_3 为对应参数的权重,参考文献[18]可分别设置为 0.0282、0.2953 和 3.5753。

表 1 为不同方法在 UIEBD 数据集上水下增强图像的 PSNR 和 SSIM,可以发现,基于融合的 Fusion 方法评价指标值最高,与图 4 中的图像恢复效果表现一致。其次,基于深度学习的方法中, Uresnet 和 FunieGAN 方法的得分较低,而 Water-Net、UGAN 方法的得分相对较高,本方法在定量指标上的得分最高。原因是实验使用的测试集 UIEBD 中的真值图像中有 24.72% 是通过 Fusion 方法生成的^[6]。除 Fusion 外,本方法的 PSNR 和 SSIM 指标得分最高,进一步证明了本方法的有效性。

表 1 不同方法在 UIEBD 测试集上的 PSNR 和 SSIM
Table 1 PSNR and SSIM of different methods on the UIEBD test set

Type	Method	PSNR /dB	SSIM
Model-based method	HE	16.28	0.781
	RGHS	19.49	0.832
	Fusion	22.12	0.846
	UDCP	12.64	0.578
	IBLA	12.42	0.463
Deeplearning-based method	Water-Net	19.12	0.813
	Uresnet	17.66	0.740
	FUnieGAN	17.63	0.761
	UGAN	19.60	0.840
	Ours	20.59	0.844

对 EUVP 数据集中非成对的测试集用 UIQM 进行评价,并从增强后图像的色彩、锐度和对比度三个方面对图像进行详细分析,结果如表 2 所示。可以发现,本方法的 UIQM 评分最高。总体来看,基于深度学习的方法优于基于模型的方法,原因是这类方法得到的图像对比度较高,且色彩和锐度的得分比较均衡。基于模型的方法中, Fusion 方法的质量评分最高,生成的图像色彩鲜明,突出了前景和背景,这与图 4 中的表现一致。HE 方法次之,但 HE 方法的分值主要体现在色彩度评分上,且 HE 方法增强后的图像色彩太过艳丽已经失真。基于深度学习的方法中, Water-Net 和 UGAN 方法的评分略低于本方法,原因是 Water-Net 方法对多种预处理图像进行深度融合学习,本方法仅使用了一种预处理

表 2 不同方法在 EUVP 数据集中的 UIQM
Table 2 UIQM of different methods in EUVP data set

Type	Method	UIQM	UICM	UISM	UIConM
Model-based method	HE	2.906	11.127	7.554	0.101
	RGHS	2.169	7.483	5.796	0.069
	Fusion	2.922	8.177	8.004	0.091
	UDCP	1.978	7.711	5.609	0.029
	IBLA	1.718	8.115	4.474	0.046
Deep learning-based method	Water-Net	3.030	5.014	7.545	0.184
	Uresnet	2.884	5.193	6.822	0.202
	FUnieGAN	2.782	6.121	7.021	0.149
	UGAN	3.423	5.988	8.307	0.235
	Ours	3.471	5.483	8.413	0.232

方法,通过像素级融合提升网络的学习性能。相比 UGAN 方法,本方法通过增加 Loss 惩罚项优化了 GAN 的性能,虽然在 UIQM 指标上的提升有限,但结合图 5 可知,本方法的泛化能力较好,能满足水下机器视觉的要求。

3.5 具体分析

除主观和客观质量评价外,高级计算机视觉任务(如水下目标检测)也可用于评价图像增强方法的性能。因此,参考文献[34]中的方法,用 YOLOv3 (You only look once)目标检测器^[35]对原始水下图像和本方法生成的增强图像进行检测,结果如图 7 所示。可以发现,在对比度低、色彩失真的原始图像中,存在大量目标未被识别,而在增强后的图像中可以更准确地检测出海星和海胆。

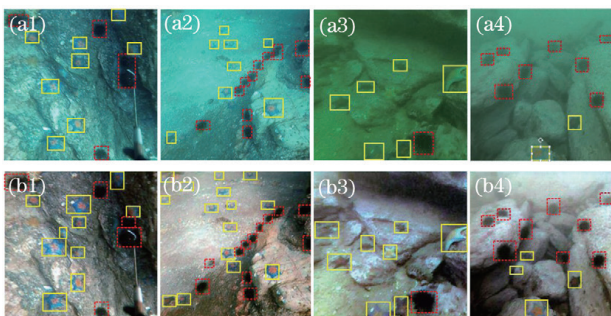


图 7 水下目标的检测结果。(a)原始图像;(b)增强后的图像

Fig. 7 Detection results of underwater targets.

(a) Original image; (b) enhanced image

用预处理图像作为伪真值对网络施加惩罚,能改善网络的图像增强效果,增强网络的泛化能力。为验证预处理图像惩罚的作用,分别在真实水下图

像数据集 EUVP^[30]和合成水下图像数据集^[32]上对比了施加与未施加该惩罚的定量指标,并用 PSNR 和 SSIM 进行评价,结果如表 3 所示。可以发现,对网络施加惩罚的网络模型在真实图像和合成图像上的 PSNR 和 SSIM 均较高,这表明增加图像惩罚能够提升网络的性能。

表 3 惩罚项对图像增强效果的影响

Table 3 Influence of penalty items on the image enhancement effect

Dataset	Method	PSNR /dB	SSIM
EUVP	without penalty	19.71	0.833
	with penalty	20.59	0.844
Synthesis	without penalty	15.25	0.709
	with penalty	15.60	0.716

利用改进的直方图拉伸预处理模块作为惩罚项对网络施加惩罚,对真实水下图像和合成水下图像的增强效果均较好,但并不是所有的预处理算法都能有效提升网络的图像增强效果。因此,用基于 UDCP 的图像恢复方法、基于深度学习网络的 CycleGAN 作为预处理模块进行图像增强对比分析,定量评价结果如表 4 所示。可以发现,基于 UDCP 预处理方法的定量指标最低,甚至会影响网络的增强效果;用 CycleGAN 进行预处理在一定程度上提高了网络的图像增强效果,但与本方法中的自适应直方图拉伸方法相比,网络复杂度较高,且本方法的结构简单、时间成本较低。

表 4 不同预处理方法对图像增强效果的影响
Table 4 Influence of different preprocessing methods on the image enhancement effect

Data set	Method	PSNR /dB	SSIM
EUVP	UDCP	18.65	0.832
	CycleGAN	19.93	0.843
	Ours	20.59	0.844
Synthesis	UDCP	14.19	0.687
	CycleGAN	15.37	0.711
	Ours	15.60	0.716

4 结 论

针对水下图像的色彩失真、对比度低问题,构建了一个结合预处理图像惩罚的 GAN 水下图像增强方法。首先,对原始图像进行改进 R 通道的直方图拉伸,根据图像的均值和方差设置阈值,防止部分像素拉伸后图像局部块出现严重的过增强现象。其次,用多尺度 GAN 进行网络训练,利用不同的卷积核获取不同的感受野,从不同尺度中获得统计信息。最后,将预处理后图像作为整个网络的惩罚项引导生成器的训练,以增强模型的泛化能力。主观和客观评价结果表明,相比传统图像增强方法和其他深度学习网络,本方法在不同场景、不同损失程度的水下图像上的表现均较好;且本方法恢复的图像清晰度更高、图像细节更接近真实图像。本方法充分结合了传统图像增强方法和深度学习方法的优点,能取得比其他深度学习方法更优的水下图像增强效果,但未考虑水下图像的成像机理对图像质量的影响,因此,还需将水下图像成像原理与 GAN 相结合,进一步开展水下图像复原工作。

参 考 文 献

- [1] Ludvigsen M, Sortland B, Johnsen G, et al. Applications of geo-referenced underwater photo mosaics in marine biology and archaeology [J]. *Oceanography*, 2007, 20(4): 140-149.
- [2] Foresti G L. Visual inspection of sea bottom structures by an autonomous underwater vehicle[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2001, 31(5): 691-705.
- [3] Hummel R. Image enhancement by histogram transformation [J]. *Computer Graphics and Image Processing*, 1977, 6(2): 184-195.
- [4] Zuiderveld K. Contrast limited adaptive histogram equalization[M]//*Graphics Gems*. Amsterdam: Elsevier, 1994: 474-485.
- [5] Huang D M, Wang Y, Song W, et al. Shallow-water image enhancement using relative global histogram stretching based on adaptive parameter acquisition [M]//Schoeffmann K, Chalidabhongse T H, Ngo C W, et al. *MultiMedia modeling*. MMM 2018. Lecture notes in computer science. Cham: Springer, 2018, 10704: 453-465.
- [6] Ancuti C, Ancuti C O, Haber T, et al. Enhancing underwater images and videos by fusion [C]//2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI, USA. New York: IEEE Press, 2012: 81-88.
- [7] He K M, Sun J, Tang X O. Single image haze removal using dark channel prior[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(12): 2341-2353.
- [8] Drews Jr P, do Nascimento E, Moraes F, et al. Transmission estimation in underwater single images [C]//2013 IEEE International Conference on Computer Vision Workshops, December 2-8, 2013, Sydney, NSW, Australia. New York: IEEE Press, 2013: 825-830.
- [9] Peng Y T, Zhao X Y, Cosman P C. Single underwater image enhancement using depth estimation based on blurriness[C]//2015 IEEE International Conference on Image Processing (ICIP), September 27-30, 2015, QC, Canada. New York: IEEE Press, 2015: 4952-4956.
- [10] Peng Y T, Cosman P C. Underwater image restoration based on image blurriness and light absorption[J]. *IEEE Transactions on Image Processing*, 2017, 26(4): 1579-1594.
- [11] Jin W P, Guo J C, Qi Q. Underwater image enhancement based on conditional generative adversarial network [J]. *Laser & Optoelectronics Progress*, 2020, 57(14): 141002. 晋玮佩, 郭继昌, 祁清. 基于条件生成对抗网络的水下图像增强 [J]. *激光与光电子学进展*, 2020, 57(14): 141002.
- [12] Lu H M, Li Y J, Mu S L, et al. Motor anomaly detection for unmanned aerial vehicles using reinforcement learning[J]. *IEEE Internet of Things Journal*, 2018, 5(4): 2315-2322.
- [13] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]//NIPS'14: Proceedings of the 27th International Conference on Neural Information Processing Systems, December 8, 2014,

- Cambridge, MA, US. New York: ACM, 2014, 2: 2672-2680.
- [14] 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 (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2242-2251.
- [15] Fabbri C, Islam M J, Sattar J. Enhancing underwater imagery using generative adversarial networks[C]//2018 IEEE International Conference on Robotics and Automation (ICRA), May 21-25, 2018, Brisbane, QLD, Australia. New York: IEEE Press, 2018: 7159-7165.
- [16] 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 (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 5967-5976.
- [17] Mirza M, Osindero S. Conditional generative adversarial nets[EB/OL]. (2014-11-06) [2020-09-05]. <https://arxiv.org/abs/1411.1784>.
- [18] Lu J Y, Li N, Zhang S Y, et al. Multi-scale adversarial network for underwater image restoration [J]. Optics & Laser Technology, 2019, 110: 105-113.
- [19] Chen X Y, Yu J Z, Kong S H, et al. Towards quality advancement of underwater machine vision with generative adversarial networks[EB/OL]. (2017-12-03) [2020-09-05]. <https://arxiv.org/abs/1712.00736>.
- [20] Liu P, Wang G Y, Qi H, et al. Underwater image enhancement with a deep residual framework [J]. IEEE Access, 2019, 7: 94614-94629.
- [21] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1646-1654.
- [22] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04) [2020-09-05]. <https://arxiv.org/abs/1409.1556>.
- [23] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [24] Chen X Y, Zhang W J, Sun W Z, et al. Super-resolution reconstruction of images based on multi-scale and multi-residual network [J]. Laser & Optoelectronics Progress, 2020, 57(18): 181009. 陈星宇, 张伟劲, 孙伟智, 等. 基于多尺度与多重残差网络的图像超分辨率重建[J]. 激光与光电子学进展, 2020, 57(18): 181009.
- [25] Xu Z G, Yan J J, Zhu H L. Mural image super resolution reconstruction based on multi-scale residual attention network [J]. Laser & Optoelectronics Progress, 2020, 57(16): 161012. 徐志刚, 闫娟娟, 朱红蕾. 基于多尺度残差注意力网络的壁画图像超分辨率重建算法[J]. 激光与光电子学进展, 2020, 57(16): 161012.
- [26] Li C, Wand M. Precomputed real-time texture synthesis with Markovian generative adversarial networks[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9907: 702-716.
- [27] Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of wasserstein GANs[C]//NIPS'17: Proceedings of the 31st International Conference on Neural Information, December 4, 2017, Red Hook, NY, US. New York: ACM, 2017: 5767-5777.
- [28] Arjovsky M, Chintala S, Bottou L. Wasserstein GAN[EB/OL]. (2017-01-26) [2020-09-05]. <https://arxiv.org/abs/1701.07875>.
- [29] Deng J, Dong W, Socher R, et al. ImageNet: a large-scale hierarchical image database[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition, June 20-25, 2009, Miami, FL, USA. New York: IEEE Press, 2009: 248-255.
- [30] Islam M J, Xia Y Y, Sattar J. Fast underwater image enhancement for improved visual perception [J]. IEEE Robotics and Automation Letters, 2020, 5(2): 3227-3234.
- [31] Li C Y, Guo C L, Ren W Q, et al. An underwater image enhancement benchmark dataset and beyond [J]. IEEE Transactions on Image Processing, 2020, 29: 4376-4389.
- [32] Li C Y, Anwar S, Porikli F. Underwater scene prior inspired deep underwater image and video enhancement [J]. Pattern Recognition, 2020, 98: 107038.
- [33] Panetta K, Gao C, Agaian S. Human-visual-system-inspired underwater image quality measures [J]. IEEE Journal of Oceanic Engineering, 2016, 41(3): 541-551.

- [34] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [35] Zhao Q, Li B Q, Li T W. Target detection algorithm based on improved YOLOv3[J]. Laser & Optoelectronics Progress, 2020, 57(12): 121502.
赵琼, 李宝清, 李唐薇. 基于改进 YOLOv3 的目标检测算法[J]. 激光与光电子学进展, 2020, 57(12): 121502.