自动化学报
ACTA AUTOMATICA SINICA

基于改进CycleGAN的水下图像颜色校正与增强

李庆忠 白文秀 牛炯

Underwater Image Color Correction and Enhancement Based on Improved Cycle-consistent Generative Adversarial Networks

LI Qing-Zhong, BAI Wen-Xiu, NIU Jiong

在线阅读 View online: https://doi.org/10.16383/j.aas.c200510

基于改进 CycleGAN 的水下图像颜色校正与增强

李庆忠 [1] 白文秀 [1] 牛 炯 [1]

摘 要 针对水下观测图像的颜色失真和散射模糊问题, 提出一种基于改进循环一致性生成对抗网络 (Cycle-consistent generative adversarial networks, CycleGAN) 的水下图像颜色校正与增强算法. 为了利用 CycleGAN 学习水下降质图像到空气中图像的映射关系, 对传统 CycleGAN 的损失函数进行了改进, 提出了基于图像强边缘结构相似度 (Strong edge and structure similarity, SESS) 损失函数的 SESS-CycleGAN, SESS-CycleGAN 可以在保留原水下图像的边缘结构信息的前提下实现水下降质图像的颜色校正和对比度增强. 为了确保增强后图像和真实脱水图像颜色的一致性, 建立了 SESS-CycleGAN 和正向生成网络 G 相结合的网络结构; 并提出了两阶段学习策略, 即先利用非成对训练集以弱监督方式进行 SESS-CycleGAN 学习, 然后再利用少量成对训练集以强监督方式进行正向生成网络 G 的监督式学习. 实验结果表明: 本文算法在校正水下图像颜色失真的同时还增强了图像对比度, 且较好地实现了增强后图像和真实脱水图像视觉颜色的一致性.

关键词 水下图像, 深度学习, 循环一致性生成对抗网络, 颜色校正, 图像增强

引用格式 李庆忠, 白文秀, 牛炯. 基于改进 CycleGAN 的水下图像颜色校正与增强. 自动化学报, 2023, **49**(4): 820−829

DOI 10.16383/j.aas.c200510

Underwater Image Color Correction and Enhancement Based on Improved Cycle-consistent Generative Adversarial Networks

LI Qing-Zhong[1] BAI Wen-Xiu[1] NIU Jiong[1]

Abstract Aiming at the problem of color distortion and scattering blurry effect of underwater observation images, an color correction and enhancement algorithm of underwater images is proposed based on improved cycle-consistent generative adversarial networks (CycleGAN). Firstly, in order to learn the mapping relationship between underwater degraded images and in-air images using CycleGAN, the CycleGAN is improved by introducing a new loss function based on strong edge and structure similarity (SESS). The new SESS-CycleGAN can realize color correction and contrast enhancement of underwater degraded images while retaining their edge and structure information. Secondly, for ensuring the color consistency between the enhanced images and the their corresponding in-air images, a dual-level network structure combining SESS-CycleGAN with a forward generation network G is constructed, and a two-stage learning strategy is proposed. In the two-stage learning process, the SESS-CycleGAN is first trained in a weakly supervised manner with a unpaired training set composed of underwater images and in-air images, and then the forward generation network G of SESS-CycleGAN is trained again in a strong supervised manner with a small number of paired training set. Experimental results show that the proposed algorithm can enhance image contrast while correcting color distortion of underwater images, and has better visual color consistency with real images in air.

Key words Underwater image, deep learning, cycle-consistent generative adversarial networks (CycleGAN), color correction, image enhancement

Citation Li Qing-Zhong, Bai Wen-Xiu, Niu Jiong. Underwater image color correction and enhancement based on improved cycle-consistent generative adversarial networks. *Acta Automatica Sinica*, 2023, **49**(4): 820−829

收稿日期 2020-07-09 录用日期 2020-11-04
Manuscript received July 9, 2020; accepted November 4, 2020
国家重点研发计划 (2017YFC1405202), 海洋公益性行业科研专项 (201605002) 资助
Supported by National Key Research and Development Program of China (2017YFC1405202) and National Marine Technology Program for Public Welfare of China (201605002)
本文责任编委 桑农
Recommended by Associate Editor SANG Nong
1. 中国海洋大学工程学院 青岛 266100
1. College of Engineering, Ocean University of China, Qingdao 266100

水下光学图像探测技术是海洋开发、探测常用的关键技术之一, 是水下机器人 (Autonomous underwater vehicle/Remote operated vehicle, AUV/ROV) 进行深海目标监测与生物资源考察的重要方法[1−3], 在海洋工程、海洋军事以及海洋环境保护等方面具有重要应用前景[4].

图像的颜色信息作为反映物体特性的重要信息, 对于水下目标的识别和分类起到重要的作用.

但是水下光线的衰减和散射效应会导致水下图像呈现对比度低、散射模糊和颜色失真现象, 为后续的基于机器视觉的自动识别和理解带来了严重困难. 为此, 本文的目的就是研究基于改进循环生成对抗神经网络的水下图像颜色校正与增强新方法.

根据是否基于水下成像物理模型, 目前水下图像的颜色校正方法可分为基于成像模型的图像复原算法和基于非成像模型的图像增强算法两类.

1) 基于成像物理模型的图像复原方法[5−9] 是对水下图像的退化过程建立数学模型, 并估算出模型的基本参数, 然后通过该模型反演图像的退化过程, 并获得理想状态下的未退化图像. 如 He 等[5] 提出了暗通道先验 (Dark channel prior, DCP) 算法, 该算法通过寻找雾化图像的局部最暗点并结合图像物理模型进行图像复原. 该类算法的共同点是需要建立水下成像模型, 并估计散射光成分和衰减系数. 由于水下环境复杂多变, 水下成像模型很难建立, 所以参数估计精度和模型稳健性都受到一定限制.

2) 图像增强算法不需要对成像过程进行建模和参数估计, 而是直接对降质图像进行增强处理. 主要的代表性方法有以下两种:

a) 传统的图像处理方法. 通过图像处理技术, 直接调整或改变图像的像素值分布以达到改善图像质量的目的[10−13]. 如基于灰度世界 (Gray world, GW) 假设的算法[11], 利用颜色恒常性原理, 将自然场景颜色的平均值设为定值, 并应用于水下图像增强. UCM (Unsupervised color correction method) 算法[12] 先将水下图像在 RGB 空间进行色彩均衡, 然后再在 HSV 空间进行对比度校正. 这类算法虽然简单, 但是易存在噪声过增强、以及引入伪影和颜色失真等问题.

b) 基于深度学习的水下图像颜色校正增强算法. 该类算法通过深度卷积神经网络的学习, 来建立水下降质图像和真实图像之间的复杂非线性映射关系, 进而达到对水下退化图像颜色校正的目的. 目前主要的代表性方法都是基于生成对抗网络 (Generative adversarial networks, GANs)[14] 和循环一致性生成对抗网络 (Cycle-consistent generative adversarial networks, CycleGANs)[15]. 如 Li 等[16] 首先利用生成对抗神经网络 WaterGAN 以非监督方式生成一个训练集, 该网络的输入是空气中的彩色图像和该图像的深度图, 输出为对应深度的水下生成图像. 然后基于生成的训练集, 利用双级卷积全连接网络进行水下实际图像的颜色校正. 这种方法的特点是必须利用深度图像进行训练学习, 这在实际应用场景中难以实现, 此外深度估计一旦存在误差, 也会影响颜色的校正效果. 此外, Li 等[17] 提出了一种基于循环生成对抗网络 (CycleGAN) 的水下图像颜色校正方法, 该方法利用 CycleGAN 以弱监督方式学习非成对空气图像集和水下图像集之间的颜色映射关系. 为了保留水下图像的内容和结构信息, 在网络训练损失函数中, 新加入了结构相似性损失函数 (Structural similarity index, SSIM). 在缺乏水下成对图像样本集的条件下, 对于校正水下图像的偏色效应取得了较好的实验结果, 并具有普适性强、效率高的优点. 但该方法仍然存在两个问题: 1) 采用的结构相似性损失函数在图像亮度信息、对比度信息、结构信息三方面进行了相似性限制, 这与实际情况不相符合. 因为原始的模糊图像和校正后的清晰图像在亮度、对比度上本应不相似, 若强制相似, 反而起不到对比度增强的目的. 2) 对于校正后图像的颜色没有客观的定量指标进行约束或评价, 故很难保证校正后图像和真实理想图像颜色的一致性.

为了克服文献 [17] 存在的以上两个问题, 本文提出了一种基于改进 CycleGAN 的水下图像颜色校正与增强算法. 该方法的主要贡献体现在两方面: 1) 为了保证在 CycleGAN 学习中, 既能保留原始水下图像的边缘、纹理结构信息, 又能提高增强后图像的对比度, 提出了基于图像强边缘结构相似度 (Strong edge and structure similarity, SESS) 损失函数的 SESS-CycleGAN; 2) 为了保证校正后图像的颜色和真实图像颜色的一致性, 提出了弱监督 SESS-CycleGAN 和强监督生成网络相结合的网络结构并提出了两阶段学习策略. 并通过实验验证了提出方法在颜色校正和对比度增强方面的有效性.

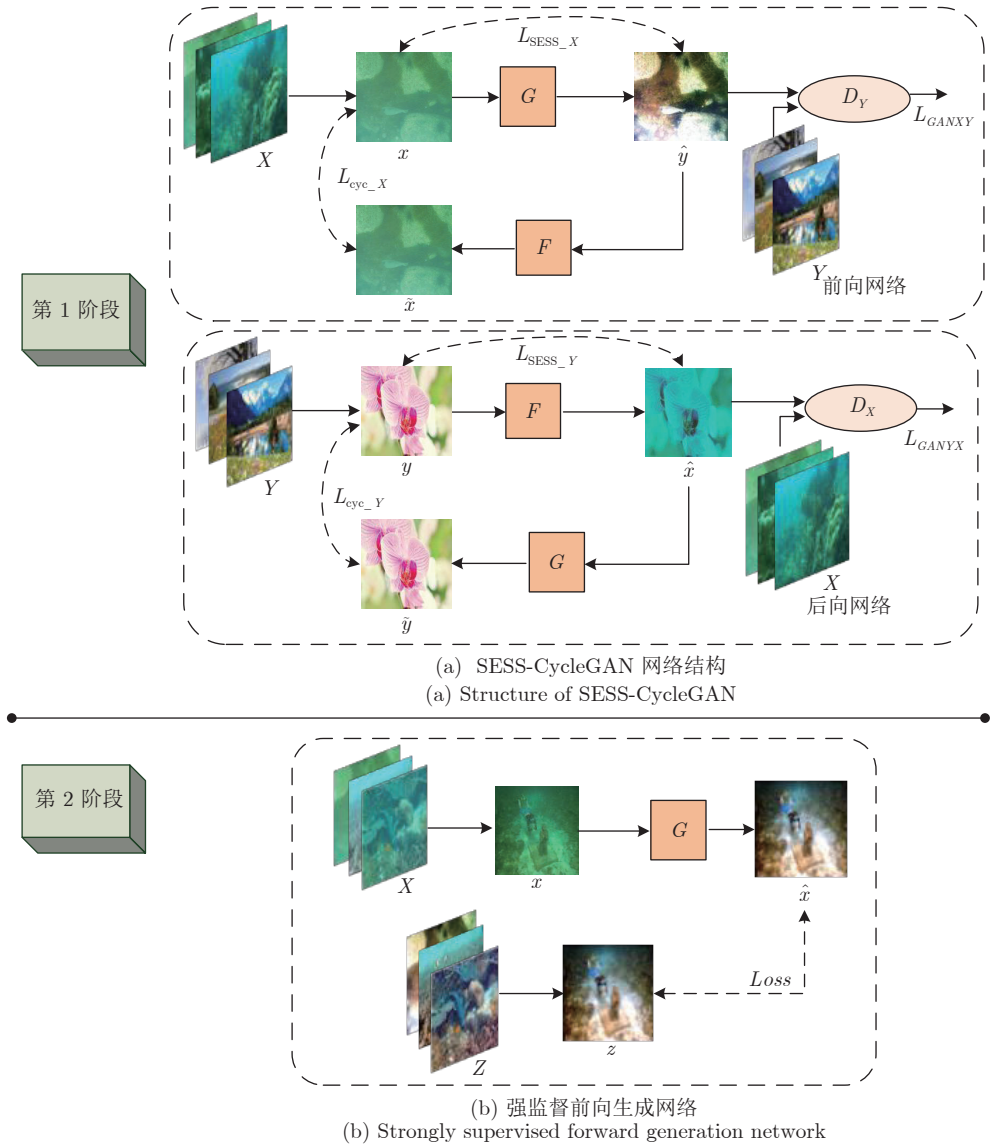
1 提出的方法

本节分别从网络结构组成、网络损失函数确定、和网络学习过程三方面介绍本文提出的基于改进 CycleGAN 的水下图像增强方法.

1.1 网络结构组成

本文提出算法的网络结构组成如图 1 所示. 该网络主要由两部分组成.

1) 图 1 中第 1 阶段的改进 CycleGAN 网络, 即基于图像强边缘结构相似度 (SESS) 损失函数的 CycleGAN, 本文称之为 SESS-CycleGAN. 该网络的作用是通过对非成对训练集的弱监督学习, 实现由水下图像到空气中脱水图像颜色风格的转换. 图 1 中, 前向网络的 $L_{\mathrm{SESS\_}X}$ 代表水下图像 x 与经过生成器 G 生成的脱水图像 $\hat{y}$ 之间的差别, $L_{\mathrm{cyc\_}X}$ 代表原水下图像 x 与 $\hat{y}$ 经过生成器 F 生成的水下图像 $\hat{x}$ 之间的差别, L_{GANXY} 代表生成的脱水图像 $\hat{y}$ 与真实的脱水图像域 Y 之间分布的差别; 后向网络同理. 由图 1 可知: SESS-CycleGAN 由两个镜像



(a) SESS-CycleGAN 网络结构
(a) Structure of SESS-CycleGAN

(b) 强监督前向生成网络
(b) Strongly supervised forward generation network

图 1 双级网络结构
Fig. 1 Dual-level network structure

GAN 网络构成一个环形循环网络, 包含两个生成器 (G 和 F)以及两个判别器 (D_X 和 D_Y). 为了实现从水下图像域 X 到空气脱水图像域 Y 的风格转换, 需要训练生成器 G 和 F, 即学习一个映射 G: $X \rightarrow Y$, 使生成的样本 $\hat{y} = G(x)$ 尽可能与真实空气脱水图像域 Y 的分布达到一致; 以及 G 的逆向映射 F: $Y \rightarrow X$, 使 $\tilde{x} = F(G(x))$ 尽可能与真实水下图像集 X 的分布一致, 由此保持循环一致性, 即 $F(G(x)) \approx x$, 以防止 X 的所有图像都被映射为 Y 中的某一幅图像, 与此同时, 还引入了判别器 D_Y 来判别生成图像的类别. 更新 G 和 D_Y, 直到 D_Y 的输出结果接近 0.5, 达到纳什平衡, 学习映射完成. Y 域到 X 域的风格转换与此同理.

生成器的组成如图 2 所示, 主要包含 3 个部分: 编码器、转换器、解码器. 编码器主要用来提取水下图像的不同特征, 将其转化为特征向量; 转换器作用是组合图像的不同相近特征; 解码器和编码器过程完全相反, 通过特征向量重构低级特征, 输出生成的空气脱水图像.

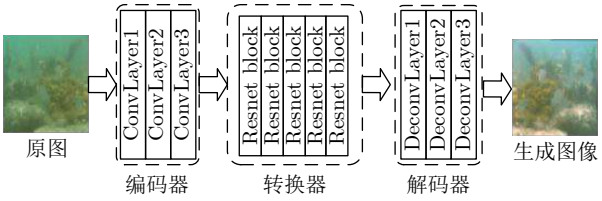


图 2 生成器结构图
Fig. 2 Structure of generation network

而判别器的网络结构如图 3 所示, 由 5 层卷积层构成, 其中前四层用来提取图像特征, 最后一层用来辨别图像的真假. 它将生成的图像作为输入, 对其提取特征并进行图像类型识别, 输出预测图像为真实图像的概率.

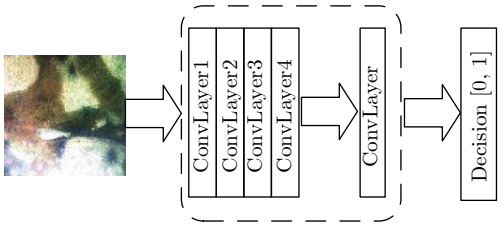


图 3　判别器网络结构

Fig. 3　Structure of discriminator

2) 图 1 中第 2 阶段的强监督生成网络, 其目的是为了生成真实的脱水图像. 利用少量成对的水下图像集 X 和真实脱水图像集 Z, 对该生成网络 G 进行一对一训练, 其网络结构为 Encoder-Transformer-Decoder 框架, 图 4 为该生成器的网络特征流程图.

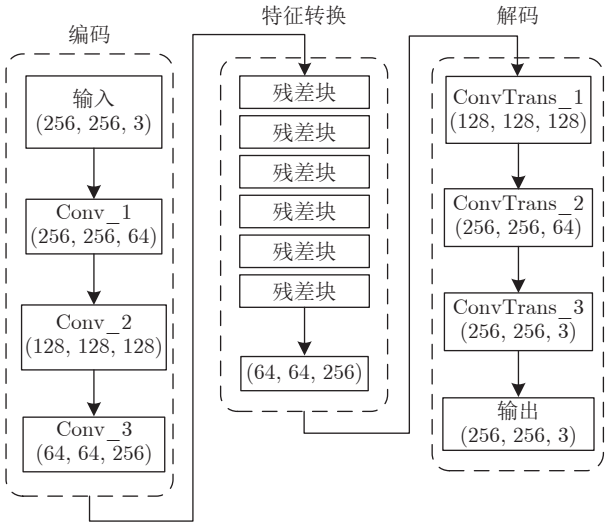


图 4　生成网络特征流程图

Fig. 4　Feature flow chart of generation network

图 4 中编码器由 3 层卷积层构成, 从 $256 \times 256 \times 3$ 的图像中提取 $64 \times 64 \times 256$ 维特征. 转换器由 6 个残差块实现, 将水下图像域 X 图像的 $64 \times 64 \times 256$ 特征向量转换得到对应脱水图像域 Z 图像的 $64 \times 64 \times 256$ 特征向量. 解码器和编码器过程完全相反, 由 3 层反卷积层构成, 将 $64 \times 64 \times 256$ 特征向量输出为 $256 \times 256 \times 3$ 的目标域图像.

1.2　网络损失函数确定

损失函数是卷积神经网络的优化准则, 为网络权重提供更新方向, 整个网络的损失函数包括第 1 阶段 SESS-CycleGAN 的对抗损失、循环一致性损失、边缘损失三部分; 以及第 2 阶段生成网络的一致性损失函数.

1) 对抗损失. 对抗损失主要是为了更准确地提取并融合水下失真图像的特征, 其定义如式 (1) 所示, 代表生成的空气脱水图像 $\hat{y}$ 与真实的空气脱水图像域 Y 之间分布的差别.

$$L_{\mathrm{GAN}}(G, D_Y, X, Y) = \mathrm{E}_{y\sim P_{\mathrm{data}}(y)}[(D_Y(y)-1)^2] + \mathrm{E}_{x\sim P_{\mathrm{data}}(x)}[D_Y(G(x))^2] \tag{1}$$

式中, X, Y 分别代表水下图像域和空气脱水图像域, $x \in X$, $y \in Y$, $P_{\mathrm{data}}(x)$ 代表数据集 X 的分布, $\mathrm{E}_{x\sim P_{\mathrm{data}}(x)}$ 代表在 x 服从 $P_{\mathrm{data}}(x)$ 的情况下求均值. 式 (1) 表示 $X \rightarrow Y$ 的过程, 同理, 式 (2) 表示 $Y \rightarrow X$ 的过程, 为生成的水下图像 $\hat{x}$ 与真实的水下图像域 X 之间的损失.

$$L_{\mathrm{GAN}}(F, D_X, Y, X) = \mathrm{E}_{x\sim P_{\mathrm{data}}(x)}[(D_X(x)-1)^2] + \mathrm{E}_{y\sim P_{\mathrm{data}}(y)}[D_X(G(y))^2] \tag{2}$$

2) 循环一致性损失. 式 (3) 为循环一致性损失的定义, 其作用是为了保持转换后的图像与原始图像内容的一致性. 为 $F(G(x))$ 与 x, $G(F(y))$ 与 y 之间的损失, 起到保留图像内容信息的作用.

$$L_{\mathrm{cyc}}(G, F) = \mathrm{E}_{x\sim P_{\mathrm{data}}(x)}[||F(G(x))-x||_1] + \mathrm{E}_{y\sim P_{\mathrm{data}}(y)}[||G(F(y))-y||_1] \tag{3}$$

3) 图像强边缘结构相似度 (SESS) 损失. 对于水下图像的颜色校正与增强来说, 该损失函数的作用至关重要. 因为一般的 CycleGAN 网络, 只能实现图像不同风格的转换, 例如输入是马, 输出的是斑马, 在这种不同风格图像的转换中, 原始图像的结构纹理等都发生改变. 而本文的目的是利用 CycleGAN 网络实现由水下图像到其对应脱水图像的转换, 这里的风格转变主要体现在图像颜色和对比度的变化, 即图像颜色校正和对比度的增强.

为了达到以上目的, 文献 [17] 提出了基于结构相似性 SSIM 的结构性损失函数. SSIM 是通过对比两幅图像在对比度 $l(x, y)$、结构 $s(x, y)$、亮度 $c(x, y)$ 三方面的差异的组合来衡量两幅图像相似性的指标, 将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. 对于两幅图像 x, y, 其 SSIM 指标定义为

$$SSIM(x, y) = l^{\alpha}(x, y)c^{\beta}(x, y)s^{\lambda}(x, y) \tag{4}$$

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \tag{5}$$

$$c(x,\,y)=\frac{2\sigma_x\sigma_y+C_2}{\sigma_x^2+\sigma_y^2+C_2} \tag{6}$$

$$s(x,\,y)=\frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3} \tag{7}$$

式中, $l(x,\,y)$, $s(x,\,y)$, $c(x,\,y)$ 分别代表 x 与 y 的亮度、结构和对比度因子; μ_x 为 x 的均值, μ_y 为 y 的均值, σ_{xy} 为 x, y 的协方差, σ_x^2 为 x 的方差, σ_y^2 为 y 的方差, C_1, C_2, C_3 为常数, 为了防止出现分母为零的情况, $C_1=0.01$, $C_2=0.03$, $C_3=C_2/2$, 参数 α, β, λ 均大于 0, 用来调节亮度, 结构和对比度因子三者之间的重要性, 当 $\alpha=\beta=\lambda=1$ 时, SSIM 指标函数变为

$$SSIM(x,\,y)=\frac{(2\mu_x\mu_y+c_1)(2\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)} \tag{8}$$

文献 [17] 为了保证水下图像转换到空气脱水图像的过程中保留图像原有的内容和结构信息, 只实现颜色转换, 而不改变结构, 通过加入结构性损失函数, 式 (9) 对生成器网络进行约束.

$$L_{\rm SSIM}(x,\,G(x))=1-\frac{1}{N}\sum_{p=1}^{N-1}(SSIM(p)) \tag{9}$$

由 SSIM 的定义式 (4) 可知, 该损失函数可使 CycleGAN 网络生成的脱水图像和输入的水下降质图像, 在亮度、对比度上和纹理结构保持一定的一致性. 从水下图像颜色校正和对比度增强的角度看, 保持图像的纹理结构是必须的, 但是若同时限制图像亮度和对比度的一致性, 显然不合理. 因为输入的水下降质图像一般比较暗且呈现对比度较低的模糊现象, 我们希望生成的图像不仅颜色得到校正, 而且对比度也要得到增强, 即清晰化且颜色不失真. 为克服文献 [17] 中 SSIM 损失函数的缺点, 本文提出了新的基于图像强边缘结构相似度 (SESS) 的损失函数, 该损失函数只是限制输入水下图像和生成脱水图像在边缘纹理结构上保持一致性.

图像强边缘结构相似度 (SESS) 的计算步骤如下:

步骤 1. 对原始图像 X 和生成图像 $fake\_\,Y$ 分别用 Sobel 算子求梯度图像 G_1、G_2, Sobel 算子为

$$G_x=\begin{bmatrix}-1&0&1\\-2&0&2\\-1&0&1\end{bmatrix},\quad G_y=\begin{bmatrix}1&2&1\\0&0&0\\-1&-2&-1\end{bmatrix} \tag{10}$$

步骤 2. 采用 Otsu 算法[18] 自适应确定梯度图像 G_1 的阈值 T_1, 将梯度图像 G_1 分为背景和强边缘前景两部分. 设前景像素点数占整幅图像的比例记为 r_1, 前景平均梯度为 u_1; 背景像素点数占整幅图像的比例记为 r_2, 背景平均梯度为 u_2, 则图像总体均值为

$$u=r_1u_1+r_2u_2 \tag{11}$$

关于阈值 t 的最大类间方差表达式为

$$f(t)=r_1(u_1-u)^2+r_2(u_2-u)^2 \tag{12}$$

当 $f(t)$ 取最大值时, 对应的 t 即为最佳阈值 T_1.

步骤 3. 按照最佳阈值 T_1 对梯度图像 G_1 进行二值分割得到原始强边缘图像 $image_1$; 并计算强边缘像素个数所占整幅图像的比率 $rate$.

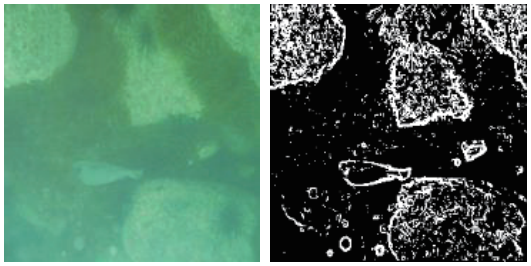
步骤 4. 因为要求原始强边缘比率 $rate$ 和生成图像的强边缘比率相同, 所以由 $rate$ 可确定出生成梯度图像 G_2 的最佳分割阈值 T_2, 并得到生成图像的强边缘图像 $image_2$.

步骤 5. 对得到的强边缘图像 $image_1$ 和 $image_2$, 计算边缘结果相似度损失 $L_{\rm SESS}$ 为

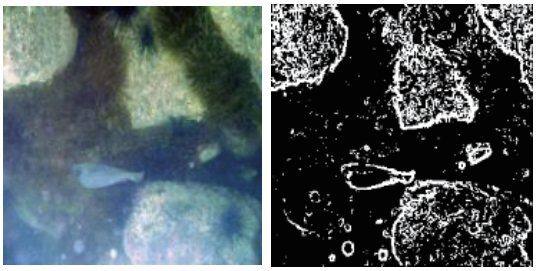
$$\begin{aligned}L_{\rm SESS}=&\,{\rm E}_{x\sim P_{\rm data}(x)}[||G(x)_1-\widehat{x}||_1]+\\&\,{\rm E}_{y\sim P_{\rm data}(y)}[||F(y)_1-\widehat{y}||_1]\end{aligned} \tag{13}$$

其中, 第 1 项为前向网络的边缘损失, $\widehat{x}$ 代表 $image_1$, $G(x)_1$ 代表 $image_2$, 第 2 项为后向网络的边缘损失.

图 5 是经过 SESS-CycleGAN 算法得到的脱水图像与原始水下图像的强边缘图像对比, 其中, 图 5(a) 是原始水下图像和其对应的强边缘图像; 图 5(b) 是生成的脱水图像和其对应的强边缘图像. 可见, 两



(a) 原始图像
(a) Original images



(b) SESS-CycleGAN 图像
(b) SESS-CycleGAN images

图 5 强边缘图像对比
Fig. 5 Comparison of strong edge images

幅图像的强边缘图像基本一致, 即含有类似结构和边缘信息, 但生成脱水图像的对比度和亮度都得到了增强, 颜色也得到了校正, 这说明本文提出的图像强边缘结构相似度 (SESS) 损失函数是有效的.

4) SESS-CycleGAN 的总损失. 总损失函数为对抗损失、循环一致性损失、边缘结构损失三部分的加权组合, 即

$$\begin{aligned} L(G,\,F,\,D_X,\,D_Y) = {} & L_{\rm GAN}(G,\,D_Y,\,X,\,Y) + \\ & L_{\rm GAN}(F,\,D_X,\,Y,\,X) + \\ & \lambda L_{\rm cyc}(G,\,F) + \omega L_{\rm SESS} \end{aligned} \tag{14}$$

式中, 权系数 λ, ω 体现了各个损失函数之间的相对重要性, 为了保证该网络的对抗损失、循环一致性损失和图像强边缘结构相似度损失值在同一数量级, 即对于模型的训练具有同等的重要性, 通过考察计算各损失值的范围, 将权系数设为: $\lambda = 10$, $\omega = 10$. 式中前向网络的生成器 G 的目标是使生成图像 $G(x)$ 尽可能与 Y 域图像更相似, 判别器 D 的目标是将 G 生成的图像 $G(x)$ 尽可能与真实 Y 域图像区别开来. 为了能够达到相应的生成功能和鉴别功能, G 尝试最小化这个损失, 而 D 将尝试最大化该损失, 即: $\min_G \max_{D_Y} L_{\rm GAN}(G,\,D_Y,\,X,\,Y)$, 后向网络同理, 因此, 该网络的训练目标为

$$G^*,\,F^* = \arg\min_{G,\,F}\max_{D_X,\,D_Y} L(G,\,F,\,D_X,\,D_Y) \tag{15}$$

5) 第 2 阶段正向生成网络的一致性损失. 这是强监督学习的损失函数, 利用正向生成器生成的脱水图像 $\hat{z} = G(x)$ 和与 $\hat{z}$ 相对应的真实脱水图像 z 的差来进行正向生成器的二次监督式学习, 因此, 该阶段损失函数为

$$M_{\rm LOSS} = {\rm E}_{z\sim P_{\rm data}(z)}[||G(x) - z||_1] \tag{16}$$

式中, $z \in Z$, Z 为真实脱水图像集, $G(x)$ 为生成的脱水图像, z 为与之对应的真实脱水图像.

1.3　网络学习过程

整个网络的学习过程分为两个阶段: 第 1 阶段为 SESS-CycleGAN 网络非成对样本集的弱监督学习; 第 2 阶段为正向生成网络少量成对样本集的强监督学习. 在训练过程中为了减小振荡, 我们使用缓存的 50 幅历史生成图像而不是最新生成图片来进行判别器的训练. 采用 Adam 优化器, 其利用梯度的一阶矩估计和二阶矩估计动态调整每个参数的学习率, 经过偏置校正之后, 每次迭代的学习率都有一个确定的范围, 能够计算不同参数的自适应学习率, 从而保证网络快速收敛, 稳定高效地完成训练.

图 6 为第 1 阶段 SESS-CycleGAN 网络训练过程中整体损失函数式 (14) 的变化曲线. 可以看出, 该网络在训练开始时的损失函数值较大, 随后迅速收敛至基本稳定的范围, 并且期间波动幅度很小, 由此可以看出该网络具有较强的稳定性.

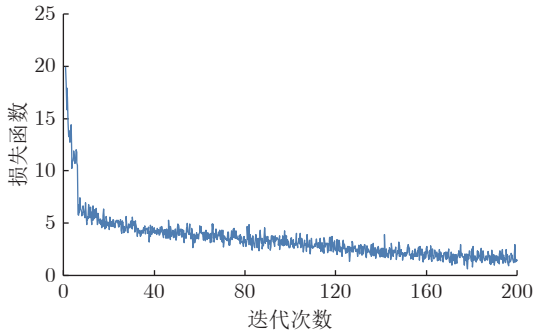


图 6　SESS-CycleGAN 损失函数曲线

Fig. 6　Loss function curve of SESS-CycleGAN

图 7 为第 2 阶段强监督网络损失函数式 (16) 的变化曲线, 该阶段为利用成对数据对第 1 阶段前向生成网络的进一步训练, 计算生成图像与真实脱水图像误差. 由图 7 可以看出, 该阶段网络亦能够快速稳步收敛, 达到训练目标.

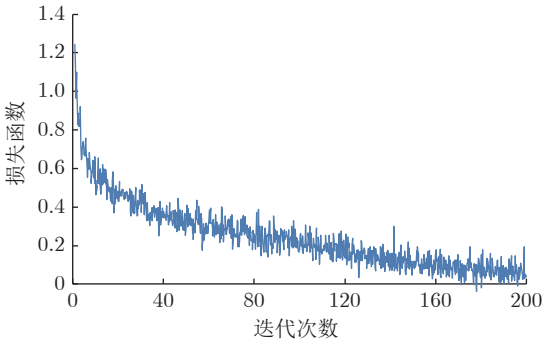


图 7　强监督网络损失函数曲线

Fig. 7　Loss function curve of strongly supervised network

SESS-CycleGAN 网络的生成器和判别器结构与原 CycleGAN 结构相同, 训练参数如下: 批次 (Batch size) 为 1, 训练样本尺寸为 256×256 像素, 动量为 0.5, Epoch 为迭代次数, 将整个数据集每训练一次, Epoch 便增加一次, 一共有 200 个 Epoch. 其中, 前 100 个 Epoch, 学习率为 0.0002; 后 100 个 Epoch, 学习率按照线性递减为 0. 第 1 阶段学习图像集为网络收集到的非成对 3000 幅水下图像和 3000 幅空气图像. 其中水下图像集基本涵盖了所有的水下图像风格, 包括不同水下场景以及水下物种等. 第 2 阶段成对图像集为 800 幅水下图像和对应的

800 幅真实脱水图像, 该图像集来自于文献 [19]. 在以上学习图像集中, 通过其中存在的少量异常数据, 以确保模型训练的鲁棒性; 而模型的泛化性是指网络模型对于未知数据的预测能力, 本文通过数据集的多样性以及 CycleGAN 网络从 X 域到 X 域, 从 Y 域到 Y 域的一致性训练过程来实现网络的泛化性. 测试集为从成对图像中随机抽取的 200 幅水下图像. 本文算法在 Pytorch 下运行, Adam 优化器的其余超参数均为其默认值. 服务器的配置参数为: NVIDIA Quadro P600 GPU、Ubuntu16.4 操作系统.

值得注意的是, 第 2 阶段监督生成网络的训练有利于生成器学习到水下图像和真实脱水图像更多特征之间的对应关系, 加入真实输出样本对生成器参数的更新方向进行指导, 使生成的脱水图像域图像更加真实、自然、清晰, 亦为生成图像的真实性提供了可靠性的依据.

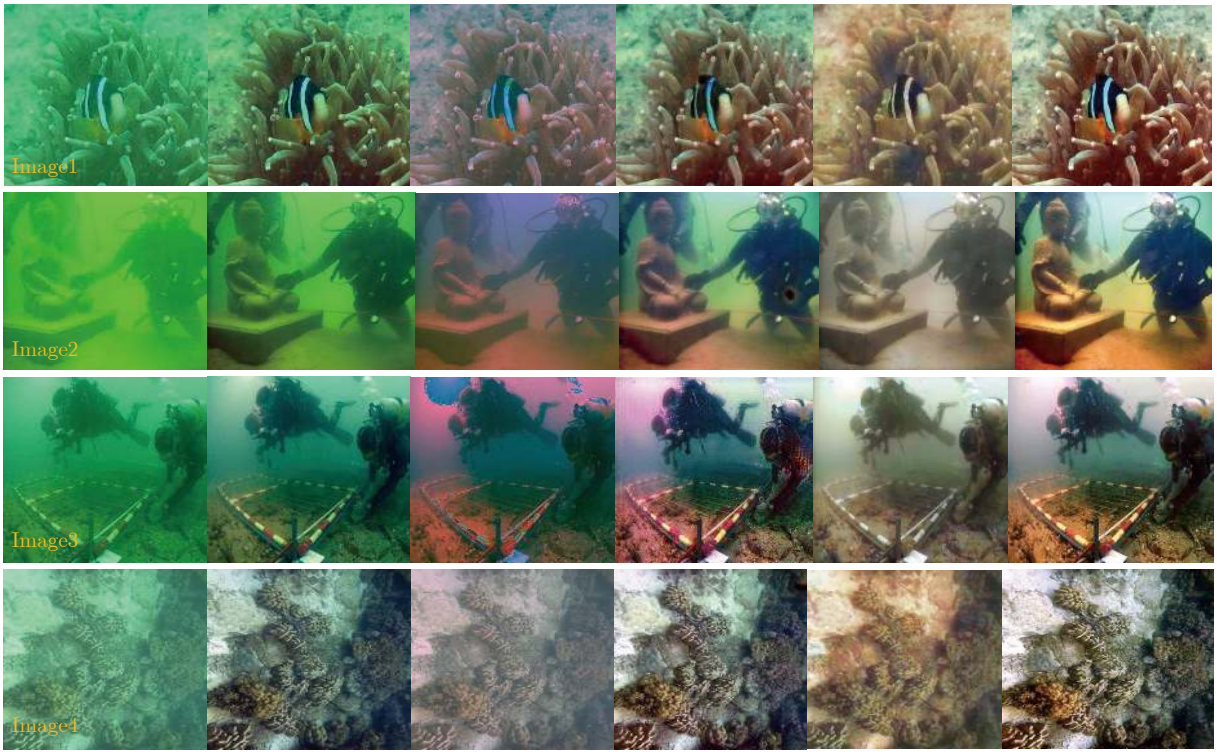
2 实验结果与分析

为了验证本文算法的有效性, 通过主观视觉评价和客观指标评价对校正结果进行分析. 主观上, 将本文算法与暗通道先验算法 (DCP)[5]、GW 算法[11]、循环生成对抗网络 (CycleGAN)[15]、SSIM-CycleGAN[17] 进行视觉效果上的比较; 客观上, 通过信息熵 (Entropy)[20]、水下彩色图像质量评价指标 (Underwater color image quality evaluation, UCIQE)[21] 和峰值信噪比 (Peak signal to noise ratio, PSNR)[13] 作为衡量标准, 对测试集中原始水下图像和经各算法增强后的图像进行质量对比评价.

2.1 主观评价

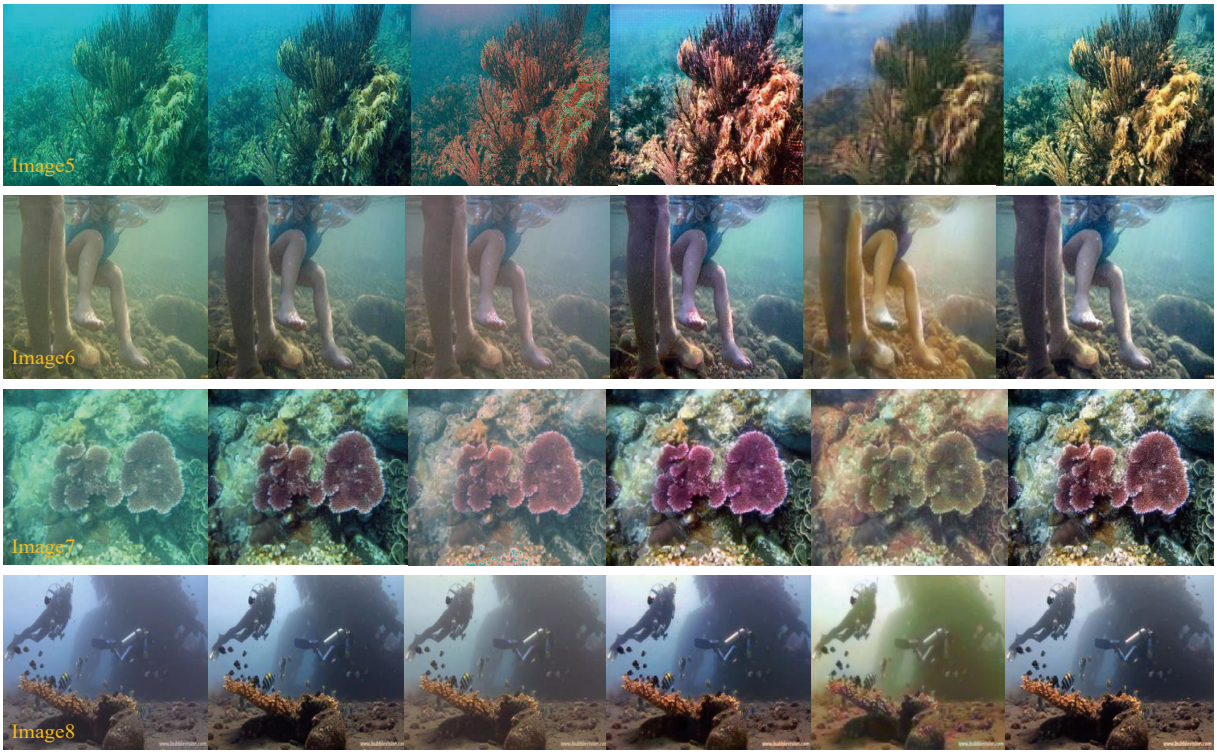
为了显示各种算法的增强效果, 将水下降质图像与 5 种算法的增强结果进行对比, 部分结果图如图 8 所示. 综合两幅图像效果可见, DCP 算法对于水下失真图像颜色校正效果最差, 整体呈现偏蓝色, 但清晰化效果较好, 原因是由于该算法采用了去雾算法, 能够提高图像的对比度, 但对于颜色校正方面没有太大的贡献. GW 算法增强后图像颜色偏红并出现了局部颜色过亮或过暗, 原因是水下环境复杂多变, 但该算法应用颜色恒常性以灰度世界假设为基础, 将物体对于光线的反射均值设为定值, 强制应用于待处理图像, 没有考虑图像局部的颜色的变化, 故对于水下图像的校正效果较差. 由于 CycleGAN 算法旨在实现水下图像和空气图像的风格转换, 因此对于水下图像颜色校正效果一般, 图像清晰度较差.

SSIM-CycleGAN 算法对于原始水下失真图像



(a) 增强图像示例 1

(a) Enhanced images 1



(b) 增强图像示例 2
(b) Enhanced images 2

图 8　颜色失真图像增强结果对比

Fig. 8　Comparison of enhanced results for color distortion images

来说效果较好, 但校正后颜色和空气图像颜色仍有偏差, 由于其使用了 SSIM 损失函数, 学习过程中使损失值趋向于零. 对于校正后图像的结构和纹理信息有了保证, 但水下图像和空气图像亮度和对比度的区别会使其结果出现偏差. 而本文算法对此提出 SESS 损失函数, 并提出了弱监督 SESS-CycleGAN 与强监督生成网络相结合的网络结构和两阶段训练方式. 在保证图像内容和结构不变的情况下, 还提高了对比度和亮度, 得到的图像颜色更加真实自然, 清晰度也更高, 图像的视觉质量最高.

图 9 是本文的 SESS-CycleGAN 算法结果与原图像以及真实脱水图像的对比. 由图 9 可见, 本文算法的增强结果对原图像增强效果明显, 与真实脱水图像的视觉效果几乎相同, 由此证明了本文算法的有效性.

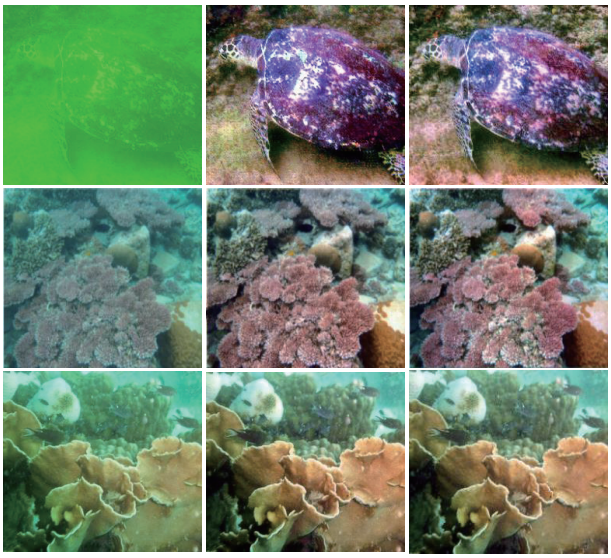


图 9　增强后图像与真实脱水图像对比

Fig. 9　Comparison of enhanced images with real air images

2.2　客观评价

首先, 为了对增强结果的质量进行客观分析比较, 下面将根据信息熵 (Entropy)、UCIQE 和峰值信噪比 (PSNR) 三个方面的指标进行定量分析.

Entropy 体现的是图像信息的丰富度, 其定义为

$$N_{\text{Entropy}} = -\sum_{t=0}^{L} p(t)\log_2 p(t) \tag{17}$$

其中, $p(t)$ 为直方图中灰度值为 t 的概率密度, L 为灰度级, 信息熵的值越大代表图像信息越丰富, 图像越清晰, 质量越好.

UCIQE 是用 CIELab 空间的色度、饱和度、清晰度三者的加权组合来评价图像质量, 其定义为

$$V_{\mathrm{UCIQE}} = c_1 C_V + c_2 S_V + c_3 Q_V \tag{18}$$

式中, c_1, c_2, c_3 为加权系数, 分别为 0.4680, 0.2745, 0.2576; C_V 为色调方差, S_V 为饱和度方差, Q_V 为清晰度方差, UCIQE 的值与图像质量成正比关系.

PSNR 反映了两幅图像之间的相似性, 其定义为

$$L_{\mathrm{MSE}} = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (X[i, j] - Y[i, j])^2 \tag{19}$$

$$K_{\mathrm{PSNR}} = 10\lg\left(\frac{MAX_I^2}{L_{\mathrm{MSE}}}\right) \tag{20}$$

式中, L_{MSE} 为真实脱水图像与增强后图像像素之间的累计平方差, m, n 分别为图像的长和宽, X 为真实脱水图像, Y 为增强后图像, $X[i, j]$, $Y[i, j]$ 为该位置的像素值. MAX_I 代表图像像素点的最大值. L_{MSE} 值越小、K_{PSNR} 值越大表示增强后图像与真实脱水图像越接近, 即真实性越高.

为了显示本文算法增强结果的客观性, 在成对数据集中随机抽取 200 幅水下图像作为测试集. 分别得到 5 种算法的处理结果. 表 1 为 5 种算法对于测试集图像增强结果的 3 个参数的平均测评值. 由表 1 可知, 本文算法的信息熵平均值在 5 种算法中均为最高值, 说明本文算法保留了更多的信息量, 并具有更好的清晰度. 本文的 UCIQE 平均分数在 5 种算法中也为最高, 说明本文算法在均衡水下降质图像的色度、饱和度以及提高清晰度方面有更好的表现, 能够得到整体颜色更自然的增强图像. 不难看出, 与其他算法相比, 本文算法的 PSNR 指标更高. 说明本文算法增强后的图像与真实的脱水图像, 在颜色和对比度等方面更加接近, 即增强效果更佳且真实性更强.

表 1 增强图像的指标对比

Table 1 Comparison of enhanced images

方法	信息熵	UCIQE	PSNR (dB)
原始图像	6.546	0.437	—
DCP	6.940	0.540	19.537
GW	6.621	0.493	16.491
CycleGAN	7.592	0.613	19.690
SSIM-CycleGAN	7.598	0.541	20.670
本文方法	**7.714**	**0.631**	**24.594**

3 结束语

本文利用深度学习的方法, 提出了一种基于改进循环生成对抗网络的水下图像颜色校正与增强算法. 通过改进循环生成对抗网络的损失函数结构, 提出了基于水下失真图像和生成脱水图像的强边缘结构相似度损失函数的 SESS-CycleGAN, 保证了网络输入图像与输出图像的边缘结构不发生改变. 采用弱监督 SESS-CycleGAN 网络和强监督生成网络相结合的网络结构和双阶段学习模式, 保证了生成图像和真实脱水图像颜色的一致性. 实验结果表明, 本文算法有较高的颜色保真度, 在校正水下图像颜色失真的同时还提高了图像的对比度. 在未来工作中, 我们将进一步研究基于循环生成对抗网络的水下模糊图像清晰化算法研究, 以期达到利用改进的循环生成对抗网络同时实现颜色校正和清晰化增强的目标.

References

1 Pinto A M, Matos A C. A hybrid imaging system for underwater robotic applications. *Information Fusion*, 2020, **55**: 16−29

2 Chadebecq F, Vasconcelos F, Lacher R. Refractive two-view reconstruction for underwater 3D vision. *International Journal of Computer Vision*, 2020, **128**(5): 1101−1117

3 Gallimore E, Terrill E, Pietruszka A, Gee J, Nager A, Hess R. Magnetic survey and autonomous target reacquisition with a scalar magnetometer on a small AUV. *Journal of Field Robotics*, 2020, **37**(7): 1246−1266

4 Ahn J, Yasukawa S, Sonoda T. An optical image transmission system for deep sea creature sampling missions using autonomous underwater vehicle. *IEEE Journal of Oceanic Eengineering*, 2020, **45**(2): 350−361

5 He K M, Sun J, Tang X O. Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**(12): 2341−2353

6 Wen H C, Tian Y H, Huang T J, Gao W. Single underwater image enhancement with a new optical model. In: Proceedings of the IEEE International Symposium on Circuits and Systems (ISC 2013). Beijing, China: IEEE, 2013. 753−756

7 Yang Ai-Ping, Qu Chang, Wang Jian, Zhang Li-Yun. Underwater image visibility restoration based on underwater imaging model. *Journal of Electronics and Information Technology*, 2018, **40**(2): 298−305
(杨爱萍, 曲畅, 王建, 张莉云. 基于水下成像模型的图像清晰化算法. 电子与信息学报, 2018, **40**(2): 298−305)

8 Li C Y, Guo J C. Underwater image enhancement by dehazing and color correction. *Journal of Electronic Imaging*, 2015, **24**(3): Article No. 033023

9 Emberton S, Chittka L, Cavallaro A. Underwater image and video dehazing with pure haze region segmentation. *Computer Vision and Image Understanding*, 2018, **168**: 145−156

10 Liu Y C, Chan W H, Chen Y Q. Automatic white balance for digital still camera. *IEEE Transactions on Consumer Electronics*, 1995, **41**(3): 460−466

11 Ebner M. *Color Constancy*. New York: John Wiley & Sons,

2007.

12 Iqbal K, Odetayo M O, James A E, Salam R A, Talib A Z H. Enhancing the low quality images using unsupervised colour correction method. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics. Istanbul, Turkey: IEEE, 2010. 1703−1709

13 Hitam M S, Awalludin E A, Yussof W N J H W, Bachok Z. Mixture contrast limited adaptive histogram equalization for underwater image enhancement. In: Proceedings of the IEEE International Conference on Computer Applications Technology (ICCAT). Sousse, Tunisia: IEEE, 2013. 1−5

14 Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS). Montreal, Canada: ACM, 2014. 2672−2680

15 Zhu J Y, Park T, Isola P, Efros A A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017. 2242−2251

16 Li, J, Skinner K A, Eustice R M, Johnson-Roberson M. WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images. *IEEE Robotics and Automation Letters*, 2017, **3**(1): 387−394

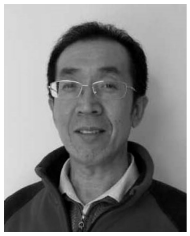
17 Li C Y, Guo J C, Guo C L. Emerging from water: Underwater image color correction based on weakly supervised color transfer. *IEEE Signal Processing Letters*, 2018, **25**(3): 323−327

18 Chen B, Zhang X, Wang R T, Li Z, Deng W. Detect concrete cracks based on OTSU algorithm with differential image. *The Journal of Engineering*, 2019, **23**: 9088−9091

19 Li C Y, Guo C L, Ren W Q, Cong R M, Hou J H, Kwang S, et al. An underwater image enhancement benchmark dataset and beyond. *IEEE Transactions on Image Processing*, 2019, **29**: 4376−4389

20 Ahmad S, Abdul G, Nor A M I. Underwater image quality enhancement through integrated color model with Rayleigh distribution. *Applied Soft Computing*, 2015, **27**: 219−230

21 Yang M, Sowmya A. An underwater color image quality evaluation metric. *IEEE Transactions on Image Processing*, 2015, **24**(12): 6062−6071



李庆忠 中国海洋大学教授. 2000 年获得中国农业大学博士学位. 主要研究方向为信号处理, 图像处理与模式识别.
E-mail: liqingzhong@ouc.edu.cn
(**LI Qing-Zhong** Professor at Ocean University of China. He received his Ph.D. degree from China Agricultural University in 2000. His research interest covers signal processing, image processing, and pattern recognition.)

白文秀 2021 年获得中国海洋大学硕士学位. 主要研究方向为图像处理与模式识别. 本文通信作者.
E-mail: baiwenxiu@ouc.edu.cn
(**BAI Wen-Xiu** Received her master degree in control theory and control engineering from Ocean University of China in 2021. Her research interest covers image processing and pattern recognition. Corresponding author of this paper.)

牛 炯 中国海洋大学高级工程师. 2008 年获得中国海洋大学硕士学位. 主要研究方向为使用高频表面波雷达进行海洋遥感、阵列信号处理和智能仪器仪表研发.
E-mail: niujiong@ouc.edu.cn
(**NIU Jiong** Senior engineer at Ocean University of China. He received his master degree from Ocean University of China in 2008. His research interest covers ocean remote sensing using high-frequency surface wave radar (HFSWR), array signal processing and intelligent instrument and meter.)