



基于多色域特征与物理模型的水下图像增强

张瑞航, 林森

引用本文:

张瑞航, 林森. 基于多色域特征与物理模型的水下图像增强[J]. *智能系统学报*, 2025, 20(2): 475-485.

ZHANG Ruihang, LIN Sen. Underwater image enhancement based on multicolor space features and physical models[J]. *CAAI Transactions on Intelligent Systems*, 2025, 20(2): 475-485.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202312004>

您可能感兴趣的其他文章

基于生成对抗网络的人脸口罩图像合成

Masked face image synthesis based on a generative adversarial network

智能系统学报. 2021, 16(6): 1073-1080 <https://dx.doi.org/10.11992/tis.202012010>

利用残差密集网络的运动模糊复原方法

Image restoration with residual dense network

智能系统学报. 2021, 16(3): 442-448 <https://dx.doi.org/10.11992/tis.201912002>

多感知兴趣区域特征融合的图像识别方法

Image recognition method based on multi-perceptual interest region feature fusion

智能系统学报. 2021, 16(2): 263-270 <https://dx.doi.org/10.11992/tis.201906032>

基于孪生变分自编码器的小样本图像分类方法

A small-sample image classification method based on a Siamese variational auto-encoder

智能系统学报. 2021, 16(2): 254-262 <https://dx.doi.org/10.11992/tis.201906022>

视觉同时定位与地图创建综述

A survey of VSLAM

智能系统学报. 2018, 13(1): 97-106 <https://dx.doi.org/10.11992/tis.201703006>

基于自编码器的特征迁移算法

Feature transfer algorithm based on an auto-encoder

智能系统学报. 2017, 12(6): 894-898 <https://dx.doi.org/10.11992/tis.201706037>

DOI: 10.11992/tis.202312004

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20250108.0911.002>

基于多色域特征与物理模型的水下图像增强

张瑞航, 林森

(沈阳理工大学自动化与电气工程学院, 辽宁 沈阳 110159)

摘要: 水下智能机器人在探测海洋信息时易受悬浮颗粒和光衰减现象的干扰, 导致视觉图像退化, 造成色彩扭曲、细节模糊等现象。针对上述问题, 提出基于多色域特征与物理模型的水下图像增强。首先, 设计多色域特征聚合网络, 旨在利用不同色域空间提供的信息帮助图像颜色恢复。其次, 为获取到更真实的视觉效果, 对白平衡算法进行推广, 并将深度学习算法与水下光学成像模型结合, 以数据驱动的方式求解清晰图像。最后, 提出多色域轮换模式对网络进行训练, 在不同色域空间中搜索最优解。实验证明, 该方法在色彩平衡、细节恢复方面效果显著, 相比经典算法与前沿算法更具优势, 在特征点匹配与显著性检验任务中满足水下智能机器人视觉系统对图像清晰度的要求。

关键词: 水下图像增强; 成像模型; 深度学习; 多色域空间; 特征聚合; 轮换训练; 算法推广; 卷积神经网络
中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1673-4785(2025)02-0475-11

中文引用格式: 张瑞航, 林森. 基于多色域特征与物理模型的水下图像增强 [J]. 智能系统学报, 2025, 20(2): 475-485.

英文引用格式: ZHANG Ruihang, LIN Sen. Underwater image enhancement based on multicolor space features and physical models[J]. CAAI transactions on intelligent systems, 2025, 20(2): 475-485.

Underwater image enhancement based on multicolor space features and physical models

ZHANG Ruihang, LIN Sen

(School of Automation and Electrical Engineering, Shenyang Ligong University, Shenyang 110159, China)

Abstract: Underwater intelligent robots are susceptible to interference from suspended particles and light attenuation phenomena when detecting oceanic information, which leads to the degradation of visual images, causing color distortion and blurring of details. An underwater image enhancement method based on multicolor domain features and physical models is proposed to address these issues. First, a multicolor space feature aggregation network is designed to leverage information from different color spaces to aid in color recovery. Then, a generalized white balance algorithm is applied to achieve a more realistic visual performance, and deep learning algorithms are combined with underwater optical imaging models to produce clear images in a data-driven manner. Finally, a multicolor space alternation model is introduced to train the network and optimize parameters across different color spaces. Experiment results demonstrate that this method effectively improves color balance and detail recovery, outperforming classical and novel algorithms. The proposed method meets the image clarity requirements of underwater intelligent robot vision systems in tasks such as feature point matching and saliency detection.

Keywords: underwater image enhancement; image formation model; deep learning; multicolor space; feature aggregation; alternate training; algorithm generalization; convolutional neural network

收稿日期: 2023-12-04. 网络出版日期: 2025-01-08.

基金项目: 国家重点研发计划项目 (2018YFB1403303); 辽宁省教育厅高等学校基本科研项目 (LJKMZ20220615); 沈阳理工大学引进高层次人才科研支持计划项目 (1010147000915).

通信作者: 林森. E-mail: lin_sen@126.com.

随着智能探测器领域的快速发展, 水下智能机器人 (underwater intelligent vehicles, UIV) 广泛应用于海洋传输管道监测^[1]、水下涡流分析^[2]和人机协同作业^[3]。由于水下成像环境特殊且照明条

件较差, UIV 的视觉系统很难捕获到高质量清晰图像, 对后续的高级视觉任务造成了极大限制。水下图像退化的原因主要有 2 个方面: 其一, 不同波长的色光在水体传播时的衰减率不同, 在物体反射回 UIV 视觉系统的光线中, 红、绿、蓝色光存在不同程度的衰减, 使捕获的图像颜色扭曲; 其二, 水中悬浮的大量浮游生物、藻类和杂质颗粒阻碍光线的前向散射, 造成图像细节模糊。

早期的研究使用传统方法处理图像, 如使用直方图拉伸^[4]、白平衡^[5]和融合^[6]等方式直接对像素值进行调整, 达到更好的视觉效果。但这些方法忽略了水下图像的成像过程, 使处理后的图像出现过度增强现象。

一些研究者使用光学成像模型^[7]复原图像, 该模型将反射回 UIV 视觉系统的光线分解为前向散射光与后向散射光, Galdran 等^[8]深入分析水下成像的特殊性, 提出红通道先验算法求解成像模型, 通过恢复与短波长相关的颜色进而恢复对比度, 实现图像复原。Drews 等^[9]提出水下暗通道算法, 考虑水体对红通道的快速衰减, 将成像模型与统计先验结合恢复水下退化图像。Peng 等^[10]在此基础上, 对该算法进行推广, 分情况讨论模型参数与形式, 以应对多种场景, 获得更高的鲁棒性。这些方法通常基于先验理论估算模型中的参数, 反推出清晰图像, 因此, 对图像的复原效果有限。

近年来, 人工智能逐渐应用到 UIV 视觉系统中, 基于深度学习的水下图像增强算法取得显著效果。Li 等^[11]构建两阶段水下生成对抗模型, 在模型第一阶段, 利用生成对抗网络^[12]的风格迁移特性^[13]生成合成数据集来训练模型的第二阶段, 并对单目水下图像进行无监督颜色校正。Guo 等^[14]设计多尺度密集连接模块, 并将其添加到生成器中, 提高生成图像的质量与网络性能。Naik 等^[15]提出轻量级连续残差注意力网络提取关键特征, 以较低的参数量获得更清晰的水下图像。Lyu 等^[16]在设计深度学习网络模型的基础上, 将最终预测图像转换到 YUV 颜色空间进行简单的亮度增强。Li 等^[17]将传统方法融入编码器与解码器网络, 旨在提取多个色域空间的特征, 并通过反转深度图进一步补偿高维特征。Islam 等^[18]基于马尔可夫判别器 (patch generative adversarial network, PatchGAN) 设计对抗网络, 有效提高了生成器效率。尽管这些方法在颜色校正方面有良好的效果, 但对色域空间特征的利用率较低, 预测的色彩不自然。

相比于其他图像增强任务, 水下图像增强任

务面临的颜色扭曲问题更具挑战性。另外, 由于缺少高质量标签图像作为参考, 增强后的图像色彩效果不真实。针对当前水下图像增强与现有算法存在的问题, 本文选择深度学习与成像模型相结合的方法对水下退化图像进行增强, 以获取接近真实视觉效果的增强图像, 该方法利用残差思想改写成成像模型, 可降低网络的拟合难度。同时, 基于编码器与解码器^[19]架构设计了对 RGB 与 HSV 色域特征进行提取与融合的特征聚合网络, 该网络以权值共享的方式处理多色域图像, 将图像 HSV 色域特征纳入网络学习, 旨在压缩参数量并消除颜色扭曲。最后, 使用卷积操作将白平衡算法推广, 减少在残差连接中原始退化图像对增强图像的干扰, 取得符合人眼感知的视觉效果。该方法在实验的定性与定量分析中超越了其他先进算法, 展现了模型的优越性。所提出方法主要贡献如下:

1) 将物理模型嵌入多色域特征学习过程, 并基于卷积对传统算法进行推广, 解决传统方法面对复杂水下环境泛化能力差的问题, 同时又利用传统方法赋予模型更强的可解释性, 减少深度学习的参数负担。

2) 设计密集像素注意力模块并融入到编码器与解码器架构, 利用密集连接机制提升对不同尺度特征的处理能力。

3) 提出色域间轮换训练, 旨在从不同色域中拟合网络参数, 扩大解空间范围, 解决增强效果不符合人眼感知问题。

1 基于多色域特征与物理模型的网络结构

图 1 给出了提出网络的整体框架。水下光学成像可以表示为

$$J_c(x, y) = I_c(x, y)T_c + A_c(1 - T_c) \quad (1)$$

式中: $J_c(x, y)$ 表示被 UIV 捕获的水下退化图像, $I_c(x, y)$ 为清晰图像, T_c 代表透射率图, A_c 代表全局背景光, c 代表通道数, (x, y) 代表像素点坐标。为将成像模型更好地嵌入深度学习网络, 对式 (1) 进行改写:

$$\begin{aligned} I_c(x, y) &= \frac{J_c(x, y)}{T_c} - \frac{A_c(1 - T_c)}{T_c} = \\ &\left(\frac{J_c(x, y)}{T_c} - \frac{A_c(1 - T_c)}{T_c} - J_c(x, y) \right) + J_c(x, y) = \\ &\left(\frac{1}{T_c} - 1 \right) (J_c(x, y) - A_c) + J_c(x, y) = \\ &K_c(x, y) + J_c(x, y) \end{aligned} \quad (2)$$

此时, 清晰图像 $I_c(x, y)$ 被分解为非线性项 $K_c(x, y)$ 与 $J_c(x, y)$, $J_c(x, y)$ 代表水下退化图像, 即网络的输入, 现设计残差连接的方式将 $J_c(x, y)$ 作为残差项

与特征聚合网络的输出进行叠加, 仅需通过特征聚合网络拟合非线性项 $K_c(x, y)$ 即可得到 $I_c(x, y)$, 相比于式 (1) 减小了拟合难度。

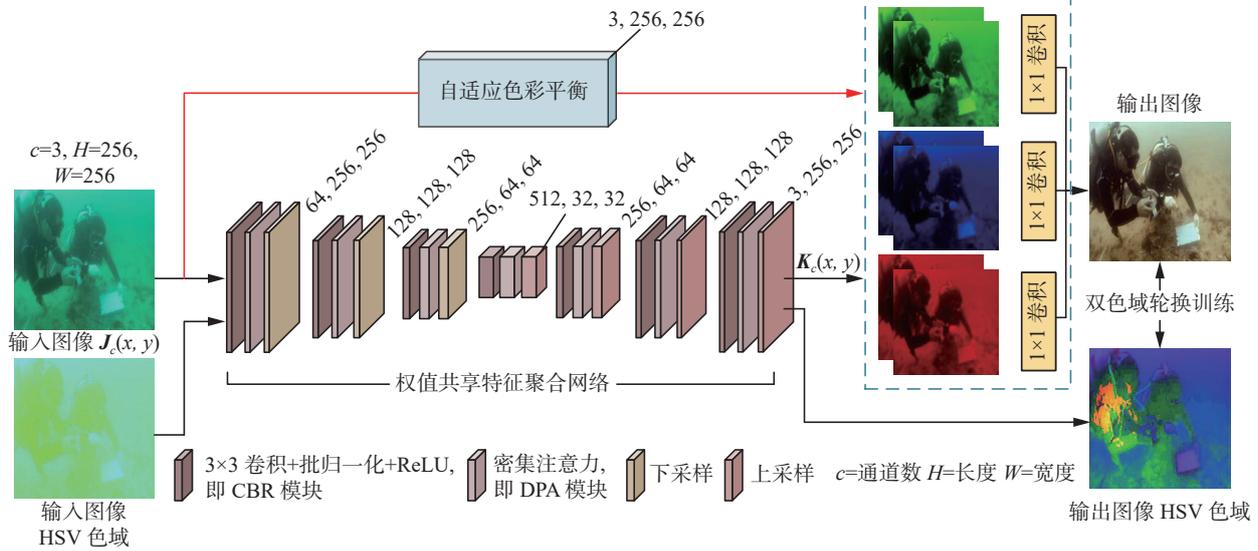


图 1 网络的整体框架

Fig. 1 Overall framework of the proposed network

如图 1 所示, 将输入图像 $J_c(x, y)$ 输入到多色域特征聚合网络对 $K_c(x, y)$ 进行拟合, 在网络的拟合过程中, 多色域特征聚合网络会对输入图像的 HSV 色域空间特征进行提取与聚合, 该操作可以保证增强效果符合人眼真实的感知效果。图中红色连接部分代表残差项 $J_c(x, y)$ 的长跳跃残差连接, 不同于常规的直接叠加, 本文在残差连接中设计了自适应色彩平衡模块, 该模块基于卷积操作对白平衡算法进行推广, 减少原始图像对增强图像的干扰, 因此式 (2) 被改进为

$$I_c(x, y) = K_c(x, y) + \text{WB}(J_c(x, y)) \quad (3)$$

式中: $\text{WB}(\cdot)$ 表示自适应色彩平衡模块对图像的处理, $I_c(x, y)$ 为网络输出的清晰图像。

1.1 权值共享特征聚合网络

在图像增强任务中, 大多数方法通过减小预测图像与标签图像之间的差异来获取网络参数最优解, 但可能出现退化图像与标签图像在 RGB 色域空间中是接近的, 而视觉效果完全相反的情况, 该情况会误导网络训练过程, 并训练出次优的网络参数权重。为解决此问题对预测图像的影响, 提出特征聚合网络同时学习 RGB 与 HSV 色域空间的特征, 不同于常规色彩体系, HSV 空间 3 个分量分别是色调 (hue)、饱和度 (saturation) 和明暗度 (value), 更接近人眼真实感知色彩。

该特征聚合网络结合了权值共享的设计理念。从图像代表的物理意义角度考虑, RGB 色域与 HSV 色域都属于图像信息, 揭示了图像的每个

像素点的红、绿、蓝颜色分量信息与饱和度、色调、亮度特性信息, 常规的深度学习方法仅利用卷积核参数学习 RGB 色域特征, 而该特征聚合网络利用轮换训练的方式使网络的卷积核参数在学习 RGB 色域特征的基础上兼顾了 HSV 色域, 通过同一网络学习 2 种色域的信息, 实现权值共享。

权值共享特征聚合网络分为特征编码阶段与特征解码阶段。输入图像通过特征编码阶段进行特征提取和通道维度扩张, 然后在解码阶段实现特征解码。在编码阶段, 对给定的 RGB 输入图像 X_{RGB} , 利用由卷积、批归一化、ReLU 激活函数构成的组合模块 CBR(convolution batch-normalization ReLU) 对 X_{RGB} 进行维度扩张与特征初步提取:

$$F_{\text{RGB1}} = \text{CBR}(X_{\text{RGB}}) = \text{ReLU}(\text{BN}(\text{Conv}_3(X_{\text{RGB}})))$$

式中: $\text{ReLU}(\cdot)$ 表示激活函数; $\text{BN}(\cdot)$ 表示对特征批归一化处理; $\text{Conv}_3(\cdot)$ 表示核尺寸为 3 的卷积操作; F_{RGB1} 表示得到的特征图, 下标数字 1 表示经过该系列操作的次数, 该系列操作可简称为 $\text{CBR}(\cdot)$ 。

得到初级特征 F_{RGB1} 后, 设计密集像素注意力 (dense pixel attention, DPA) 模块对 F_{RGB1} 的关键信息进行进一步提取, 同时过滤噪声特征。

DPA 模块的设计受密集连接^[20]的启发, 利用密集连接鼓励特征复用的特性来区分不同重要性的信息, DPA 模块的细节如图 2 所示。

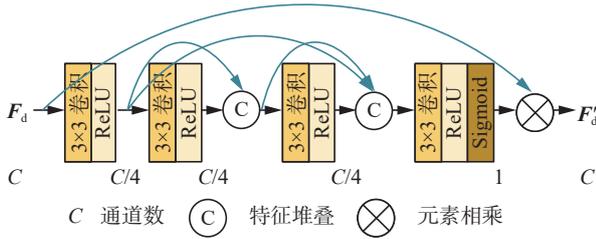


图 2 密集像素注意力模块

Fig. 2 Dense pixel attention module

假设输入特征为 F_d ，通道数为 c ，首先经过密集连接对特征进行差异化处理得到单层特征，然后由 Sigmoid 激活函数将单层特征映射为权重层。最后，权重层与输入相乘，实现对不同重要性特征的再次编码，数学表达为

$$F'_d = \text{DPA}(F_d)$$

式中 $\text{DPA}(\cdot)$ 表示密集像素注意力模块对特征的处理。将 F_{RGB1} 作为输入可得到

$$F'_{\text{RGB1}} = \text{DPA}(F_{\text{RGB1}})$$

经过 DPA 模块计算后，完成对当前分辨率特征的处理，然后对 F'_{RGB1} 进行下采样：

$$F'_{\text{RGB11}} = \text{Max2P}(F'_{\text{RGB1}})$$

式中 $\text{Max2P}(\cdot)$ 代表滑动窗口大小为 2 的全局最大池化，特征图分辨率缩减为原来的一半。编码阶段会对不同分辨率特征进行此操作，旨在多尺度学习图像的高维特征。图 1 中注明了详细的特征图分辨率尺寸。在解码阶段，采用与编码器相对称的设计，复原初始图像的空间分辨率。上采样使用线性插值方法。

为充分提高对多色域信息的利用率，同时压缩整体模型参数， X_{RGB} 在完成前向传播后，其 HSV 空间图像 X_{HSV} 以权值共享的方式再次在多色域特征聚合网络进行前向传播，得到 $X_{\text{HSV}}^{\text{pred}}$ 。最后，整体数学模型可表示为

$$\begin{cases} X_s^{\text{pred}} = \text{Sam}_p(\text{DPA}_m(\text{CBR}_n(X_s))) \\ 0 < m, n, p < 8 \\ s \in (\text{RGB}, \text{HSV}) \\ \text{Sam}_p(\cdot) = \begin{cases} \text{Max2P}(\cdot), 0 < p < 4 \\ \text{chat}(\cdot), 4 \leq p < 8 \end{cases} \end{cases}$$

式中： $\text{chat}(\cdot)$ 表示双线性插值操作， m 、 p 、 n 分别表示对应函数的进行次数。 X_{RGB} 与 X_{HSV} 的维度保持一致且经过相同的数学模型计算，为权值共享提供了运算基础。

1.2 自适应色彩平衡模块

考虑到残差项的直接叠加会影响预测图片的质量，提出自适应色彩平衡模块对残差项进行处理。灰度世界理论^[21]认为，色彩丰富的场景其红、绿、蓝三通道灰度值应趋近平衡。作为一种

有效的颜色校正方法，该理论可表达为

$$\begin{bmatrix} H_R \\ H_G \\ H_B \end{bmatrix} = \begin{bmatrix} E_R & & \\ & E_G & \\ & & E_B \end{bmatrix} \begin{bmatrix} J_R \\ J_G \\ J_B \end{bmatrix} \quad (4)$$

式中： J 表示原始图像， H 表示被灰度世界算法平衡后的图像，系数可表示为

$$E_R = \frac{0.5}{M_R}, E_G = \frac{0.5}{M_G}, E_B = \frac{0.5}{M_B} \quad (5)$$

式中：分母为当前通道的平均灰度，分子为平衡系数。但在 UIV 捕获的水下图像中，场景不总是色彩丰富的，并且平衡系数取值应依赖于三通道之间的相互关联性，因此需要对该方法进行改进。如图 3 所示，利用卷积核大小为 1 的逐点卷积学习每个像素点在三通道之间的关系特征，逐点卷积的卷积核参数可被视为每个像素点三通道的平衡系数。

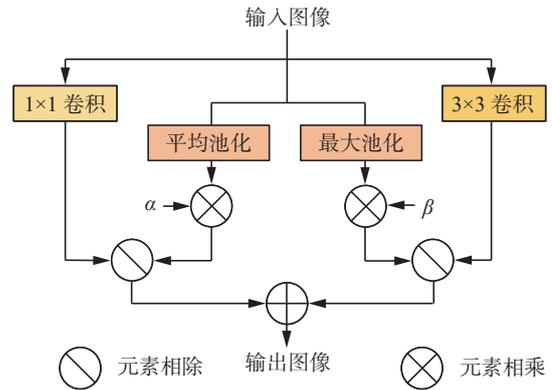


图 3 自适应色彩平衡模块

Fig. 3 Adaptive color balance module

因此，式 (4)、(5) 被初步推广为

$$H_c = \frac{1}{M_c} \text{Conv}_1(J_c) \quad (6)$$

式中 M_c 为红、绿、蓝各通道的平均值，可以利用全局平均池化操作进行求取。为应对不同的水下场景，提高算法泛化性，将最大值白平衡^[22]算法与灰度世界结合。最大值白平衡算法认为图像中最亮的点会对图像的色彩平衡造成影响，将最亮的点的像素值视为白色可对白平衡起到积极效果，可以利用全局最大池化求取图像的最大值点。因此，式 (6) 的分母改写为通道平均灰度与最大灰度的线性组合，式 (4)、(5) 进一步推广为

$$H_c = \frac{\text{Conv}_1(J_c)}{\alpha M_c + \beta P_c}$$

式中： P_c 代表单通道灰度最大值； α 与 β 为可训练参数，调节 2 种方法的比重，两参数之和为 1。最后，利用核尺寸为 3 的卷积扩大感受野，学习相邻像素的相关性，式 (4)、(5) 最终推广为

$$\mathbf{H}_c = \frac{\text{Conv}_1(\mathbf{J}_c) + \text{Conv}_3(\mathbf{J}_c)}{\alpha M_c + \beta P_c} \quad (7)$$

通过对网络结构的针对性设计, 将传统的色彩平衡算法转换为深度学习的可训练模式, 极大增强了算法泛化能力, 传统灰度世界算法与最大值白平衡算法均可被视为式 (7) 的特例。

1.3 线性叠加设计

在得到非线性项 $\mathbf{K}_c(x, y)$ 与残差项 $\mathbf{J}_c(x, y)$ 后, 设计了一种创新性的叠加策略取代直接相加。首先对 $\mathbf{K}_c(x, y)$ 和 $\mathbf{J}_c(x, y)$ 进行 R、G、B 三通道分解, 分别得到 \mathbf{K}_R 、 \mathbf{K}_G 、 \mathbf{K}_B 和 \mathbf{J}_R 、 \mathbf{J}_G 、 \mathbf{J}_B 单通道图像, 然后将相同颜色通道的单层灰度分别堆叠:

$$\mathbf{L}_c = \text{Cat}(\mathbf{K}_c, \mathbf{J}_c)$$

式中 $\text{Cat}(\cdot)$ 表示通道维度特征堆叠。最后经过逐点卷积对 \mathbf{L}_c 进行筛选, 旨在重新获得增强效果最佳的 R、G、B 单层灰度, 在数学上可表示为

$$\mathbf{I}_c = \text{Conv}_1(\mathbf{L}_c)$$

即可得到清晰图像 $\mathbf{I}_c(x, y)$, 逐点卷积对堆叠的特征 \mathbf{L}_c 进行筛选时, 存在以下数学关系:

$$\mathbf{I}_c(x, y) = w_{c=1}(x, y)\mathbf{K}_c(x, y) + w_{c=2}(x, y)\mathbf{J}_c(x, y) \quad (8)$$

式中 $w_{c=1}(x, y)$ 和 $w_{c=2}(x, y)$ 分别表示逐点卷积作用在非线性项和残差项的权重系数。存在以下 3 种特例: 当 $w_{c=1}(x, y) = 1$, $w_{c=2}(x, y) = 0$ 时, 表示在计算清晰图像 $\mathbf{I}_c(x, y)$ 时, 像素点 (x, y) 处 $\mathbf{K}_c(x, y)$ 的值为最优解; 当 $w_{c=1}(x, y) = 0$, $w_{c=2}(x, y) = 1$ 时, $\mathbf{J}_c(x, y)$ 的值为最优解; 当 $w_{c=1}(x, y) = 0.5$, $w_{c=2}(x, y) = 0.5$ 时, 线性叠加公式变为常规的直接相加。因此可将式 (3) 改写为

$$\mathbf{I}_c(x, y) = \mathbf{K}_c(x, y) \oplus \text{WB}(\mathbf{J}_c(x, y))$$

式中 \oplus 代表本小节所设计的线性叠加操作。

1.4 多色域轮换训练

为获得网络参数的全局最优解, 对设计的网络进行多色域轮换训练。输入图像的 RGB 图像在整体网络进行一次前向传播后, 与标签图像的 RGB 图像进行损失计算并更新梯度。然后, 输入图像的 HSV 图像在特征聚合网络进行一次前向传播, 与标签图像的 HSV 图像进行损失计算, 完成对单幅图像的反向传播。损失函数采用 L_1 损失与感知损失的线性组合:

$$L_f = L_1 + L_{\text{VGG}}$$

式中 L_1 代表网络的总损失。利用 L_1 损失函数可以直接学习到预测图像与标签图像的差异性, 表示为

$$L_1 = \sum_{x=1}^H \sum_{y=1}^W |\mathbf{I}(x, y) - \mathbf{I}_{\text{GT}}(x, y)|$$

式中: $\mathbf{I}_{\text{GT}}(x, y)$ 代表标签图像, H 与 W 表示图像的长度和宽度。考虑到部分低质量标签图像对训练

的干扰, 选择感知损失来衡量预测图像与标签图像在特征上的差异, 同时在数值上平衡 L_1 损失。感知损失 L_{VGG} 表示为

$$L_{\text{VGG}} = \sum_{x=1}^H \sum_{y=1}^W |\phi(\mathbf{I}(x, y)) - \phi(\mathbf{I}_{\text{GT}}(x, y))|$$

式中 $\phi(\cdot)$ 表示计算预测图像或标签图像在 VGG16 网络的特征图。为更好理解训练过程, 算法 1 给出了多色域轮换训练的伪代码。

算法 1 轮换训练算法

输入 退化图像 \mathbf{X}_{RGB} , 标签图像 \mathbf{Y}_{RGB} 。

输出 输出图像 \mathbf{I}_{RGB} (RGB 色域), 输出图像 \mathbf{I}_{HSV} (HSV 色域)。

模型实例化 特征聚合网络 Net-1, 剩余网络 Net-2。

- 1) // 将总网络拆分为 Net-1 与 Net-2 以便展示。
- 2) for $i \leq 1$ to n , $i++$, do
- 3) $\mathbf{X}_{\text{HSV}} = \text{RGB_to_HSV}(\mathbf{X}_{\text{RGB}})$ // 图像变换 HSV 色域。
- 4) $\mathbf{I}_{\text{RGB}} = \text{Net-2}(\text{Net-1}(\mathbf{X}_{\text{RGB}}))$
- 5) $L_f = \text{calculate_loss}(\mathbf{I}_{\text{RGB}}, \mathbf{Y}_{\text{RGB}})$
- 6) // 计算总损失 (RGB 色域)
- 7) $L_f.\text{back}()$ // 反向传播。
- 8) $T_{\text{opt}}.\text{step}()$ // 通过优化器 T_{opt} 更新网络参数。
- 9) $\mathbf{I}_{\text{HSV}} = \text{Net-1}(\mathbf{I}_{\text{HSV}})$
- 10) $L_f = \text{calculate_loss}(\mathbf{I}_{\text{HSV}}, \mathbf{Y}_{\text{HSV}})$
- 11) $L_f.\text{back}()$ // 反向传播。
- 12) $T_{\text{opt}}.\text{step}()$ // 更新网络参数。
- 13) $T_{\text{opt}}.\text{zero_grad}()$ // 清空梯度, 为下一次批次做准备。

2 实验分析

本节使用真实数据集对提出方法进行性能评估, 同时与当前先进算法对比: 基于风格内容分离的水下图像增强算法 (underwater image enhancement via content and style separation, UieSS)^[23]、水下图像视觉感知的快速增强算法 (fast underwater image enhancement for improved visual perception, FunieGAN)^[18]、基于暗通道先验推广的图像恢复算法 (generalization of the dark channel prior for single image restoration, GDPC)^[10]、基于模型压缩的水下图像增强网络 (compressed model for underwater image enhancement, ShallowNET)^[15]、水下暗通道先验算法 (underwater depth estimation and image restoration based on single images, UDCP)^[9]、融合 (fusion)^[24]。最后对提出方法进行消融分析, 以验证算法有效性。

2.1 数据集

实验在英伟达 RTX 3060 显卡与 PyTorch1.7 深度学习框架下进行, 选择不同规模数据集对提出方法训练与测试, 数据集分别为水下大规模配对图像与非配对图像数据集 (fast underwater image enhancement for improved visual perception, EUVP)^[18], 水下图像增强基准数据集 (underwater image enhancement benchmark dataset, UIEB)^[25], 和首个水下超分辨率图像数据集 (UFO)^[26]。其中 UIEB 包含 890 对图像, 作为小规模数据集; UFO 包含 1 500 对图像, 作为中等规模数据集; EUVP 包含 10 000 余张图像, 作为大规模数据集。这些数据集包含各类人机交互与自然场景, 可以很好地测试提出方法的性能。

2.2 衡量指标

在定量效果评估中使用峰值信噪比 (peak signal-to-noise ratio, PSNR)^[27]、结构相似性指数 (structural similarity, SSIM)^[28] 以及水下图像质量评估 (underwater image quality measurement, UIQM)^[29] 作为衡量标准。PSNR 计算预测图像与标签图像在信息内容上的差距, 细节效果越接近, PSNR 分数越高, 计算公式为

$$V_{\text{PSNR}} = 10 \times \lg \left(\frac{V_{\text{max}}^2}{V_{\text{mse}}} \right)$$

式中: V_{max} 表示图像点颜色最大值, V_{mse} 表示 2 幅

图像的均方误差。

SSIM 计算预测图像与标签图像在结构上的差距, 结构相似度越高, SSIM 分数越高, 计算公式为

$$V_{\text{SSIM}}(a, b) = \frac{(2\mu_a\mu_b + c_1)(2\sigma_{ab} + c_2)}{(\mu_a^2 + \mu_b^2 + c_1)(\sigma_a^2 + \sigma_b^2 + c_2)}$$

式中: μ_a 是图像 a 的均值, μ_b 是图像 b 的均值, σ_a^2 是图像 a 的方差, σ_b^2 是图像 b 的方差, σ_{ab} 是 2 幅图的协方差, c_1 和 c_2 是维持稳定的常数。

UIQM 指标包含水下图像色彩测量 (underwater image colorfulness measure, UICM)、水下图像锐度测量 (underwater image sharpness measure, UISM) 和 underwater image contrast measure, UiconM) 3 个分量, 分别衡量颜色均衡性、边缘信息和分块对比度, 各个分量分数越高, 图像增强效果越好, 计算公式为

$$V_{\text{UIQM}} = r_1 V_{\text{UICM}} + r_2 V_{\text{UISM}} + r_3 V_{\text{UiconM}}$$

式中: r_1 、 r_2 、 r_3 为比例系数, 在评估图像质量时, 分别被设置为 0.028 2、0.295 3 和 3.575 3。

2.3 定性定量分析

图 4、5 给出了测试集的部分测试结果, 图中右上角数字分别表示 PSNR/SSIM 或 UIQM 的分数。表 1 为测试集整体的指标分数平均值结果, 可以很好地评价各算法面对多场景时的泛化性能。表 2 为复杂度测试。表中加粗数据代表最优值, “—”表示无数据。

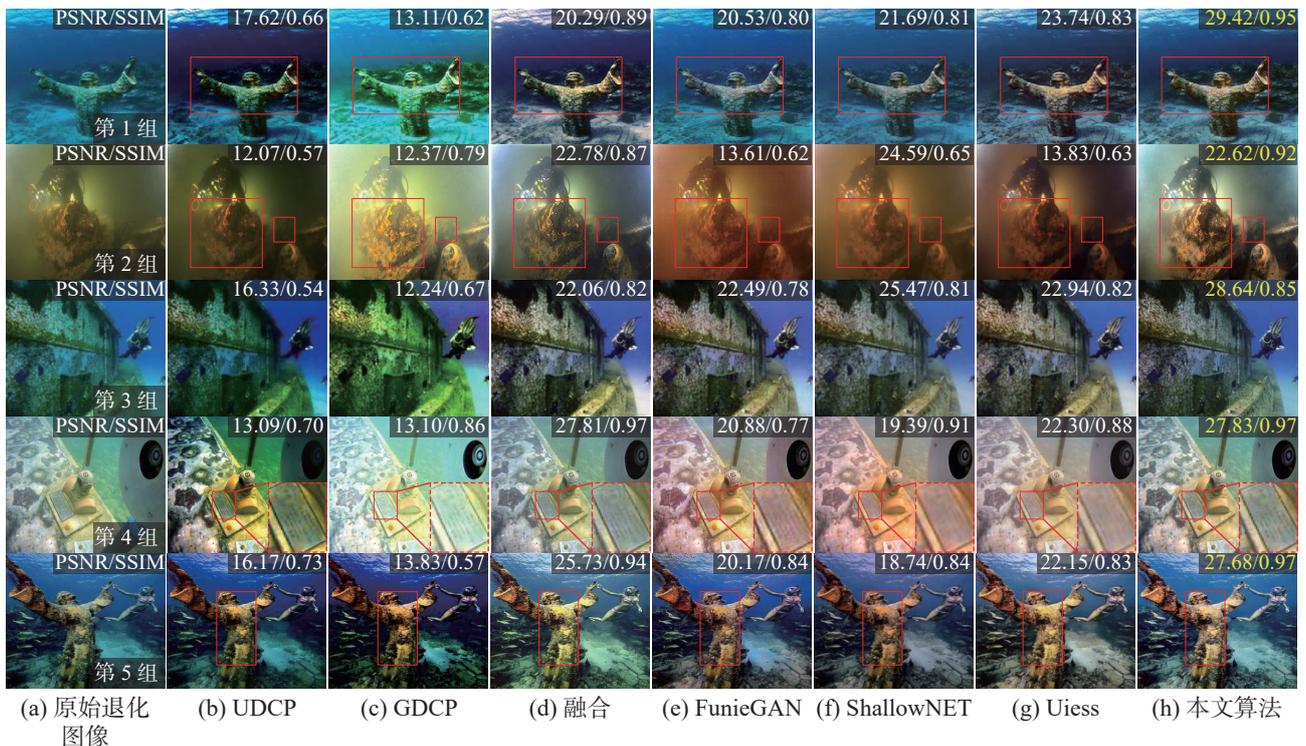


图 4 部分成对测试集的预测结果 (UIEB 与 UFO)

Fig. 4 Some predicted results of paired test sets (UIEB and UFO)

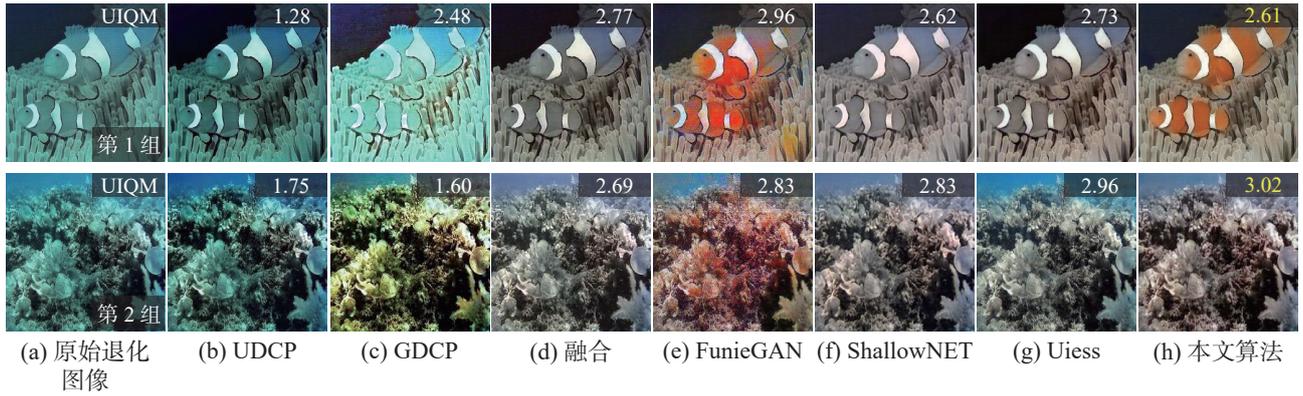


图 5 部分未成对测试集的预测结果 (EUVP)

Fig. 5 Some predicted results of unpaired test sets(EUVP)

表 1 所有测试集的定量分析

Table 1 Quantitative analysis of all test sets

测试集	指标	UDCP	GDCP	融合	FunieGAN	ShallowNET	Uiess	本文算法
EUVP	PSNR	16.59	14.12	18.55	17.43	17.20	19.12	19.71
	SSIM	0.63	0.65	0.81	0.72	0.73	0.78	0.81
	UIQM	1.50	2.11	2.56	2.69	2.38	2.49	2.83
UIEB	PSNR	12.73	14.38	20.34	16.61	16.44	18.01	21.31
	SSIM	0.61	0.73	0.87	0.72	0.73	0.78	0.90
	UIQM	1.65	1.78	3.13	3.02	2.94	2.88	3.21
UFO	PSNR	16.41	15.18	17.87	21.44	21.79	20.17	23.57
	SSIM	0.64	0.63	0.83	0.74	0.76	0.76	0.84
	UIQM	1.82	2.02	2.89	2.93	2.66	2.80	2.97

表 2 复杂度分析

Table 2 Complexity analysis

指标	UDCP	GDCP	融合	FunieGAN	ShallowNET	Uiess	本文算法
速度/s	0.53	0.66	1.57	0.48	0.11	1.31	0.18
参数量/ 10^6	—	—	—	7.73	1.04	3.83	2.96

图 4 第 1、3 组图分别为 UIV 相机捕获到的海底遗迹和残骸, UDCP 与 GDCP 的预测效果不佳, 主要原因是先验方法错误估算红通道的补偿系数。融合方法减轻了颜色扭曲, 但对阴影处的增强效果有限, 因此 PSNR 分数偏低。FunieGAN、ShallowNET 和 Uiess 有效增强了退化图像, 但图像纹理细节丢失。得益于 HSV 色域提供的信息, 本文算法在纹理对比度、饱和度与亮度上均达到最自然的效果。第 2 组图为弱光场景下的人机协同, 原始图像质量严重退化, UDCP 具有一定的去雾效果, 但对全局颜色校正以及红框处的纹理增强几乎没有作用。GDCP 对红框部分有明显的增强效果但存在对背景的过度补偿, 对比的深度学习方法造成了不同程度的红色色偏, 融合与本文

算法在亮度增强上效果显著。第 4 组图考察不同算法对图像细节的增强效果, 从被放大的区域中可以看出, 本文算法在增强图像的同时保证了图像信息的稳定性。在第 5 组图中, 融合、FunieGAN 和本文算法取得了最清晰的增强结果, 然而, 融合方法对图像全局的无区别增强造成了绿色色偏, FunieGAN 预测结果的色彩不符合人眼的真实感知。

图 5 给出了未成对测试集的部分测试结果。图 5 中, 第 1 组图为海洋鱼类图片, 测试不同算法的色彩复原能力; 第 2 组图为海底珊瑚群图片, 图像全局存在大量纹理细节, 可以很好地考察算法的稳定性。可以看出, 除本文算法与 FunieGAN 外, 其他方法均无法恢复正常色彩, 同时对

第 2 组图的结构造成不同程度的破坏。此外,在 FunieGAN 的预测结果中存在不规则分布的红色伪影。本文算法在稳定性和色彩平衡上取得了最佳效果。

表 1 给出了在不同数据集下各个算法的指标平均值,可以量化算法对测试集整体的预测效果。在传统算法中,基于融合的算法取得良好表现。基于深度学习的算法普遍比传统算法效果更好。本文方法在 PSNR、SSIM 和 UIQM 3 个维度上均优于其他方法。此外,对不同方法的参数量与处理单张图像的平均速度进行测试,除轻量级网络 ShallowNET 外,本文方法在运行时间、参数量上均超过其他方法,满足 UIV 在水下进行实时作业的需求。

2.4 消融实验

为验证提出算法各部分的有效性,在 UIEB 数据集下对模型进行消融实验研究,测试各消融模型在数据集整体上的结果可以得到稳定的数据依据,结果如表 3 所示。表中 w/o 代表去掉该部分,其余网络结构不变。数据表明,HSV 特征信息的加入可以有效提升模型的预测效果,同时 DPA 模块可以帮助网络模型更好地提取多色域特征。

表 3 消融分析
Table 3 Ablation analysis

模块消融	PSNR	SSIM
w/o 自适应色彩平衡	20.32	0.86
w/o DPA 模块	19.53	0.84
w/o HSV 色彩空间	19.71	0.85
w/o 轮换训练	20.65	0.87
w/o 线性叠加	21.14	0.90

图 6 是对网络各部分消融后的可视化效果,可以直观看到,在去掉自适应色彩平衡模块内的算法推广后,常规白平衡方法无法准确平衡图像白点,造成预测结果亮度异常。当网络消融 DPA 模块时,网络整体参数量大幅下降,仅通过常规卷积无法有效处理高维特征,预测图像质量下降。将 HSV 色域学习过程移除后,虽然可以达到较为清晰的效果,但图像全局泛红,视觉效果不自然,此情况存在于大多数的算法中,加入 HSV 色域学习可有效解决该问题。当去掉轮换训练后,没有损失函数的约束网络难以平衡两色域特征,即出现增强失败的情况。在去除线性叠加后,受网络深度的影响,原始图像的部分信息丢失,预测图像的对比度下降,纹理细节模糊。通过实验结果可知,完整网络取得最佳效果。

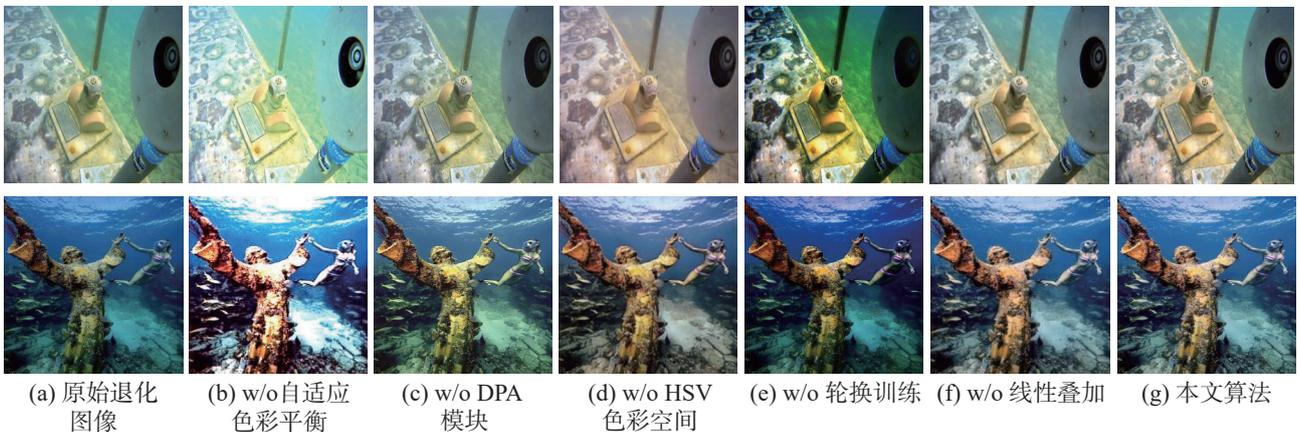


图 6 消融结果
Fig. 6 Ablation result

2.5 高级任务应用实验

水下图像增强的目的是更好地满足 UIV 在高级视觉任务上的需求,如特征点匹配^[30]与显著性检验,实验效果如图 7、8 所示。特征点精确匹配可以确保目标追踪、三维重建以及图像配准等任务的高效完成。图 7 中右上角数字表示可匹配的特征点数量。可以看到原始图像质量退化严重,

当图像发生偏移时很难配准特征点,本文算法的预测结果可以显著提升特征点匹配数量。在显著性检验中,可以明显地观察到本文算法预测的图像展现了更多的关键性特征,例如在第 1 行遗迹图像中,预测图像显著性检验结果检测出了面积最大的遗迹轮廓,在第 2 行海底管道图像中,预测图像检测出了最完整的管道表面。

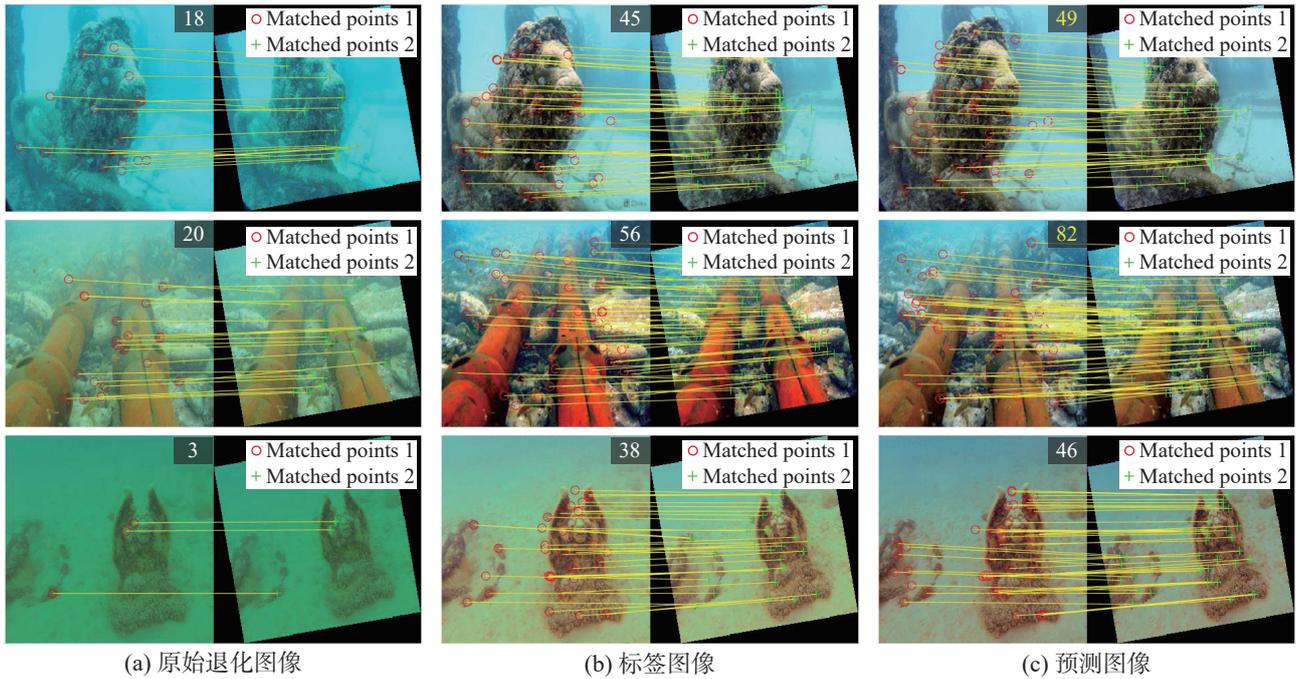


图 7 特征点匹配

Fig. 7 Feature points matching

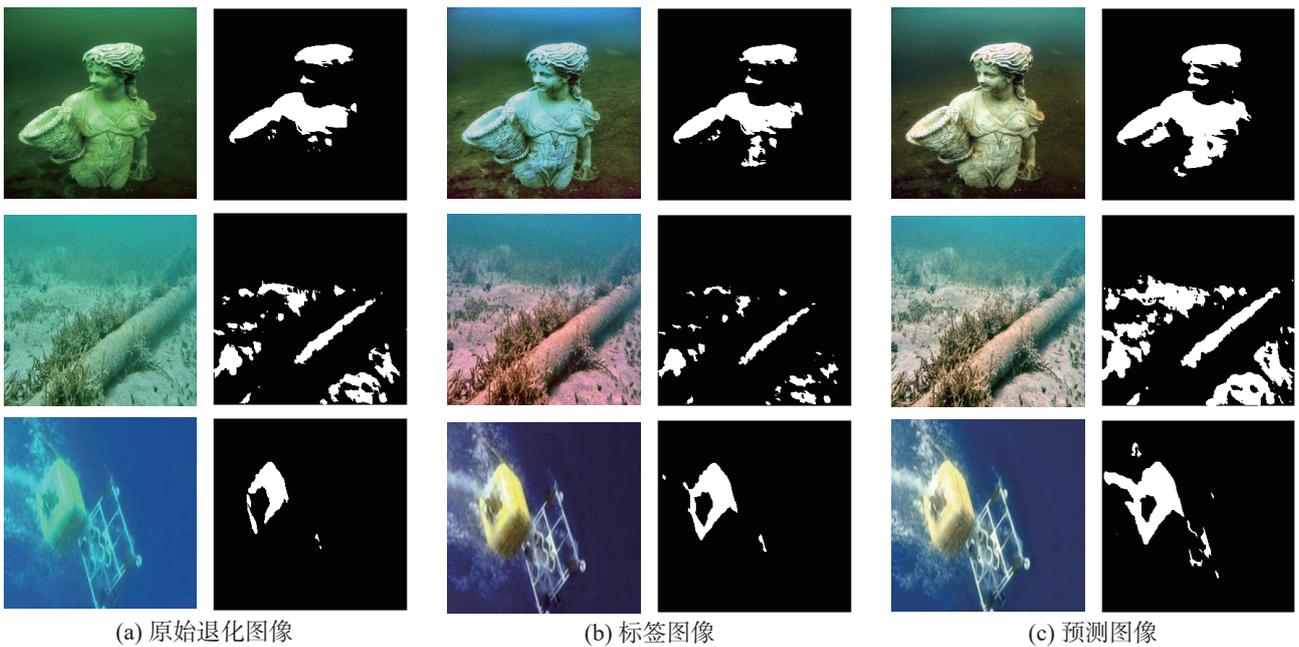


图 8 显著性检验

Fig. 8 Saliency detection

3 结束语

本文针对水下 UIV 视觉系统所面临的图像质量退化问题提出一种图像增强方法, 该方法利用图像额外的色域特征重建图像色彩。同时受到水下成像模型的启发, 分别通过基于权值共享的特征聚合网络和自适应色彩平衡模块拟合模型的非线性项和残差项, 并将传统白平衡算法进行推广, 提出更具泛化能力的自适应色彩平衡模块。

最后在不同色域空间下对网络轮换训练, 得到全局最优解。大量的定性实验与定量实验证明了本文算法对颜色恢复与细节增强的有效性, 可以为后续视觉任务提供良好的图像基础。在未来的工作中, 将对水下非均匀光照问题进行研究, 提升算法的泛化能力。

参考文献:

[1] 马建民, 李晓东, 高天羽, 等. 基于光视觉技术的海底管

- 道自主巡检系统研究[J]. 计算机测量与控制, 2022, 30(6): 32–37,58.
- MA Jianmin, LI Xiaodong, GAO Tianyu, et al. Research on autonomous inspection system of submarine pipeline based on optical vision technology[J]. *Computer measurement & control*, 2022, 30(6): 32–37,58.
- [2] 杜艳玲, 王丽丽, 黄冬梅, 等. 融合密集特征金字塔的改进 R²CNN 海洋涡旋自动检测[J]. 智能系统学报, 2023, 18(2): 341–351.
- DU Yanling, WANG Lili, HUANG Dongmei, et al. Improved R²CNN ocean eddy automatic detection with a dense feature pyramid[J]. *CAAI transactions on intelligent systems*, 2023, 18(2): 341–351.
- [3] 严浙平, 曲思瑜, 邢文. 水下图像增强方法研究综述[J]. 智能系统学报, 2022, 17(5): 860–873.
- YAN Zheping, QU Siyu, XING Wen. An overview of underwater image enhancement methods[J]. *CAAI transactions on intelligent systems*, 2022, 17(5): 860–873.
- [4] 雷小燕, 张卫东, 潘细朋. 基于直方图分布特性的水下图像颜色校正方法[J]. 计算机工程与设计, 2022, 43(8): 2284–2293.
- LEI Xiaoyan, ZHANG Weidong, PAN Xipeng. Underwater image color correction method based on histogram distribution characteristics[J]. *Computer engineering and design*, 2022, 43(8): 2284–2293.
- [5] 张云涛, 刘惠萍, 黄益铭, 等. 基于图像分割和色适应变换白平衡的水下图像增强[J]. 激光与光电子学进展, 2023, 60(14): 37–43.
- ZHANG Yuntao, LIU Huiping, HUANG Yiming, et al. Underwater image enhancement based on image segmentation and color adaptation transformation for white balance[J]. *Laser & optoelectronics progress*, 2023, 60(14): 37–43.
- [6] 林森, 迟凯晨, 唐延东. 基于复原结构与增强纹理融合的水下图像清晰化[J]. 控制与决策, 2022, 37(3): 635–644.
- LIN Sen, CHI Kaichen, TANG Yandong. Underwater image sharpening based on fusion of restored structure and enhanced texture[J]. *Control and decision*, 2022, 37(3): 635–644.
- [7] JAFFE J S. Computer modeling and the design of optimal underwater imaging systems[J]. *IEEE journal of oceanic engineering*, 1990, 15(2): 101–111.
- [8] GALDRAN A, PARDO D, PICÓN A, et al. Automatic red-channel underwater image restoration[J]. *Journal of visual communication and image representation*, 2015, 26: 132–145.
- [9] DREWS P L J, NASCIMENTO E R, BOTELHO S S C, et al. Underwater depth estimation and image restoration based on single images[J]. *IEEE computer graphics and applications*, 2016, 36(2): 24–35.
- [10] PENG Y T, CAO Keming, COSMAN P C. Generalization of the dark channel prior for single image restoration[J]. *IEEE transactions on image processing*, 2018, 27(6): 2856–2868.
- [11] LI Jie, SKINNER K A, EUSTICE R M, et al. WaterGAN: unsupervised generative network to enable real-time color correction of monocular underwater images[J]. *IEEE robotics and automation letters*, 2018, 3(1): 387–394.
- [12] ZHU Junyan, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 2242–2251.
- [13] FABBRIC, ISLAM M J, SATTAR J. Enhancing underwater imagery using generative adversarial networks[C]//2018 IEEE International Conference on Robotics and Automation. Brisbane: IEEE, 2018: 7159–7165.
- [14] GUO Yecai, LI Hanyu, ZHUANG Peixian. Underwater image enhancement using a multiscale dense generative adversarial network[J]. *IEEE journal of oceanic engineering*, 2020, 45(3): 862–870.
- [15] NAIK A, SWARNAKAR A, MITTAL K. Shallow-UWnet: compressed model for underwater image enhancement (student abstract)[C]//Proceedings of the AAAI conference on artificial intelligence. Virtual: AAAI, 2021: 15853–15854.
- [16] LYU Zhangkai, PENG A, WANG Qingwei, et al. An efficient learning-based method for underwater image enhancement[J]. *Displays*, 2022, 74: 102174.
- [17] LI Chongyi, ANWAR S, HOU Junhui, et al. Underwater image enhancement via medium transmission-guided multi-color space embedding[J]. *IEEE transactions on image processing*, 2021, 30: 4985–5000.
- [18] ISLAM M J, XIA Youya, SATTAR J. Fast underwater image enhancement for improved visual perception[J]. *IEEE robotics and automation letters*, 2020, 5(2): 3227–3234.
- [19] SHEN Zhen, XU Haiyong, LUO Ting, et al. UDAformer: underwater image enhancement based on dual attention transformer[J]. *Computers & graphics*, 2023, 111: 77–88.
- [20] PLEISS G, CHEN Danlu, HUANG Gao, et al. Memory-efficient implementation of DenseNets[EB/OL]. (2017–07–21)[2023–12–04]. <https://arxiv.org/abs/1707.06990>.
- [21] MA Xiaoming, CHEN Zhehan, FENG Zhipeng. Underwater image restoration through a combination of improved dark channel prior and gray world algorithms[J]. *Journal of electronic imaging*, 2019, 28: 053033.

- [22] SHI Zhenghao, FENG Yaning, ZHAO Minghua, et al. Normalised gamma transformation-based contrast-limited adaptive histogram equalisation with colour correction for sand-dust image enhancement[J]. *IET image processing*, 2020, 14(4): 747–756.
- [23] CHEN Yuwei, PEI S C. Domain adaptation for underwater image enhancement via content and style separation[J]. *IEEE access*, 2022, 10: 90523–90534.
- [24] ANCUTI C O, ANCUTI C, DE VLEESCHOUWER C, et al. Color balance and fusion for underwater image enhancement[J]. *IEEE transactions on image processing*, 2018, 27(1): 379–393.
- [25] LI Chongyi, GUO Chunle, REN Wenqi, et al. An underwater image enhancement benchmark dataset and beyond[J]. *IEEE transactions on image processing*, 2019, 29: 4376–4389.
- [26] JAHIDUL ISLAM M, LUO Peigen, SATTAR J. Simultaneous enhancement and super-resolution of underwater imagery for improved visual perception[EB/OL]. (2020–02–04)[2023–12–04]. <https://arxiv.org/abs/2002.01155>
- [27] WANG Zhou, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. *IEEE transactions on image processing*, 2004, 13(4): 600–612.
- [28] BAKUROV I, BUZZELLI M, SCHETTINI R, et al. Structural similarity index (SSIM) revisited: a data-driven approach[J]. *Expert systems with applications*, 2022, 189: 116087.
- [29] PANETTA K, GAO Chen, AGAIAN S. Human-visual-system-inspired underwater image quality measures[J]. *IEEE journal of oceanic engineering*, 2016, 41(3): 541–551.
- [30] LIU Xu, LIN Sen, TAO Zhiyong. Learning multiscale pipeline gated fusion for underwater image enhancement[J]. *Multimedia tools and applications*, 2023, 82(21): 32281–32304.

作者简介:



张瑞航, 硕士研究生, 主要研究方向为深度学习、水下图像处理。E-mail: zhangruihang@sylu.edu.cn。



林森, 副教授, 博士, 主要研究方向为深度学习、图像处理与模式识别。主持省部级科研项目 3 项, 获辽宁省自然科学学术成果奖 1 项, 发表学术论文 80 余篇。E-mail: lin_sen6@126.com。